Codes and Ciphers
CODES AND CIPHERS
Contents

Cover

Title Page

Introduction
1 Open communication

2 Secret communication

3 Simple codes

4 Polyalphabetic
ciphers

5 Code wars

6 Codes in cyberspace

Code chronology
Alice is driving to work. She hardly registers the road signs that guide her journey to the train station. Once there, she selects a newspaper and a book of sudoku puzzles, pausing while the bar codes are scanned before she hands over her credit card. She keys in her PIN and
pauses while the scanner securely completes the transaction. On the train she sits next to Bob, whose laptop whirrs into action after he types in his password. She decides to have a go at the crossword in her paper . . .

A WORLD SHAPED BY CODES
We could follow Alice and Bob through their day, continuing, as you will have noticed, to highlight in italics every occasion on which codes enter their lives. Codes are part of our world and of our history, too. Indeed, they have helped to shape it: Greek generals in ancient times used them to give secret orders; Wellington’s defeat of Napoleon in the Peninsular
Wars was partly due to his staff’s ability to read coded messages sent by the French army; and, of course, many people are now aware that the course of World War II was changed by the achievements of Allied cryptographers in breaking German and Japanese ciphers.
CODE OR CIPHER?

• A code uses characters or symbols to represent words or phrases, and requires a code book for encryption and decryption.
• In ciphers, individual letters are moved or changed to conceal the meaning of the message.

Since the emergence of writing, people have at times felt a need to conceal or mystify the meaning of some
of their communications, trying to keep control of who gains access to the knowledge held.

The story of code making is also the story of code breaking. This book begins by looking at ancient writing such as Egyptian hieroglyphics, which, over time, became a code, as the knowledge of how to read
them was washed away with the sands. It also includes the phone texting codes so beloved of many teenagers today.

The book also studies codes devised not to conceal but to allow efficient, fast and often cheaper communication such as Morse code. From there, we take a tour through the world of codes and ciphers
(learning the differences between them on the way) starting with fairly simple methods, such as the Caesar shift used in Roman times.

Where relevant, each code is put into historical context and, although this is not a history book, there is a ‘code chronology’ showing key events in the story of codes and code breaking.
CREATING MESSAGES

This book explains how to create messages using codes, and how they can be broken. People tend to assume that a code they have created will be very hard to break, when, in fact, most basic codes can be cracked in a very short time. *Codes and Ciphers* studies just how this is done,
including the use of cryptanalysts’ tools such as frequency analysis – using knowledge about how common some letters are in comparison to others (see pages). A code-breaking checklist is also provided on pages 171–8.

The vast majority of codes described in this book can be re-created (and broken) using
no more than pencil, paper and perseverance. Undertaking such tasks will take you back in time to the code-breaking ‘black chambers’ (see pages). As explorers and conquerors travelled greater distances in search of power and wealth, their need for secret communication increased, and messengers would walk and ride with encrypted
communications hidden in their clothing, bodies or equipment.

Naturally, this fostered a mini-industry of code breakers working in secret to break the code of any intercepted messages. In an age in which communications are bounced around the globe via satellites, interception has never been easier, and
governments (and others) continue to try to keep sensitive communication secret, and to peep into the inbox of other states and groups, such as terrorists, when the opportunity arises.

LOOKING TO THE FUTURE

The age of pencil and paper
has passed, for since the middle of the 20th century, machines have been employed increasingly to create and to break codes. The demands of code breakers have led directly to the development of the modern computer. The tools of today’s cryptographers read like a code themselves: symmetric encryption, message authentication codes,
public key encryption, one-way hash functions, digital signature schemes and random number generators. Their uses are legion, from allowing you to buy things on the internet knowing that no one will (or should) be able to steal your financial details and your money, to protecting us from a maverick pushing a button and launching a nuclear missile.
Cryptography has become a big business. There is heated debate about the most secure methods of sending sensitive diplomatic, scientific and business information, and a growing controversy about whether and when governments have the right to read private communications. The subject has entered the curriculum of universities and colleges, and data security is
an industry in itself. The subject continues to fascinate, for its history, for the intellectual challenge of creating and breaking codes, and as a leisure activity in puzzles, such as word searches, crosswords and the craze for *su doku* number conundrums.
CONVENTIONS

Throughout this book, as is the convention for code writing, the term ‘plaintext’ refers to the original message, which is altered by encoding or encipherment. Plaintext is written in upper and lower case writing; all codes and ciphers appear in capitals.

Another convention is the use
of the names of Alice and Bob (as seen in this introduction) because it has become a convention in the field of cryptography to use these monikers. This communicative couple only exist in the world of codes and ciphers. There is more on this (and their foe, Eve) on pages 150–1.
1

Open Communication

Codes from the past
Anything that we cannot read is, in effect, a code that we need to break. If it is a language, we can learn it or find someone to translate the message. But what if the language itself has disappeared?

HIEROGLYPHICS
• The word ‘hieroglyphs’ comes from the Greek language and means ‘sacred carvings’.

• They were cut into the stone of significant buildings, such as temples and tombs, or painted onto the interior walls.
The earliest date back at least 3,000 years.

The story that Egyptian hieroglyphics tell is of a highly organized and capable
civilization that collapsed and was forgotten for thousands of years. Much of what we now know about it has been learned from reading its writings on walls and papyrus, and the process of discovering how to read these is similar to the code-breaking methods that helped to shorten World War II. So, in a sense, hieroglyphics became the earliest code,
even though their meaning was not originally disguised. The structures on which hieroglyphics were carved or drawn collapsed, or were buried by the desert sand, while others were defaced by Christians intent on destroying remnants of a pagan past.

Hieroglyphics are pictorial writing: brightly coloured
images that are both simple and complex. Various attempts were made to read the images, but it took many years for people to realize that the pictures stood for sounds (as we might draw a bee to represent the sound ‘b’) of a language that had since died.

The key to unlocking this ancient mystery was the
translation of the Rosetta stone. This is a man-sized black granite rock, inscribed in three different languages, unearthed by French soldiers knocking down a wall in the Egyptian town of Rosetta in 1799. The three-quarter-tonne stone was (reluctantly, after an attempt to sneak it away on a boat) handed over to British occupying forces and taken to the British Museum
in London, where it still stands.
Egyptian hieroglyphs were intricate and colourful and were used to tell stories and demonstrate the power of the pharaohs.

Its value is that the three scripts carry the same message in a trio of languages: Greek, hieroglyphics and demotic (which is a later Egyptian language derived from hieroglyphics). Historians
were able to translate the Greek text and establish that it announced a decree issued by King Ptolemy V in 196 BC. It begins: ‘The new king, having received the kingship from his father . . . ’. For a code breaker, knowing the meaning of the message you wish to unravel is gold dust.

The stone features 1,410 hieroglyphs compared with
486 Greek words, which suggests that individual hieroglyphs do not necessarily represent whole words, and must therefore represent sounds. After some valuable groundbreaking work by Thomas Young, the stone was finally translated by Frenchman Jean François Champollion in 1823. Knowing that royal names were contained in oval shapes
known as cartouches, he made the key discovery that Ptolemy’s name was written bit by bit as p-t-o-l-m-y-s, proving it by finding the same symbols used for the shared sounds ‘p’, ‘t’, ‘o’ and ‘l’ in writings about the famous queen Cleopatra.

This wonderful discovery was the key that opened the door to understanding the language
of hieroglyphics and allowed archaeologists to learn to read other ancient Egyptian writings, shedding fresh light on the world of pharaohs, such as Rameses, and the many gods, such as the sun god Ra.
Among the early forms of written communication is cuneiform writing, in which letters are carved onto tablets (cuneiform means ‘wedge writing’). There is plenty of evidence that meanings were concealed even in such early writings, and the practice continued through the centuries.
AN ANCIENT TRADE SECRET

An early instance of cuneiform writing being enciphered dates from around 1500 BC. A small tablet giving a recipe for a pottery glaze (which must
have been a trade secret at the time) was found on the banks of the River Tigris. The scribe had mixed up different sound symbols to make the text confusing to those outside his profession.

• There are records of secret
writings that were being used for political communication in India in the 4th century BC, and the erotic textbook the *Kama Sutra* lists it as one of the skills women should learn.

- Early antecedents of the Kurds in northern Iraq employed cryptic script in holy books to keep their religion secret from their Muslim neighbours.
• A mixed-up alphabet carved onto a wooden tablet and thought to date from 7th-century Egypt is believed to be the world’s oldest cipher key.

• Medieval monastic scribes entertained themselves by adding messages in simple ciphers to the margins of texts they were copying out.

• An extraordinary 12th-
century nun called Hildegard of Bingen constructed an alternative alphabet and created a new cryptic language, called Lingua Ignota, claiming the inspiration for it came through visions.

- Franciscan Friar Roger Bacon wrote about cryptography in his *Secret Works of Art and the Nullity of Magic* in the 13th
century, listing seven different kinds of secret writing.

- Medieval builders carved masonic symbols into the stone of structures on which they were working, partly as a sort of signature and possibly to help decide how much each should be paid.

- The alchemists of the Middle Ages concealed
their identities and formulae with code marks.

A TRIO OF AMERICAN CODES

Native American tribes used smoke signals to send simple messages over long distances, working to a prearranged code. The practice was common in ancient China,
and is still carried out by boy scouts today.

Many such puffs of smoke were most likely signalling the activities of the cowboys raising cattle on the land. To prove ownership of their beasts, these men also had their own branding alphabet based on three elements: letters or numbers, geometric shapes, and pictorial symbols.
Originating in the 18th century, they were designed to identify cows over long distances and partly to combat cattle rustling (see box, below and illustration overleaf).

Hobos, who travelled around the US in the 19th century, developed a code of marks chalked outside houses to inform fellow tramps of what
kind of reception they could expect. The signs had meanings such as:

- Doctor
- Danger
- Safe camp
- If sick, will care for you
- You can sleep in the hayloft.
Cattle branding has been traced back 6,000 years to the Ancient Egyptians – tomb paintings show it taking place. There is also biblical evidence that Jacob branded his livestock.
Some examples of the cowboy branding alphabet.

Visual systems

Numerous codes have been devised to aid communication, reflecting changes in the world over the last few hundred years: these signal codes enable the passing on of information
over long distances.

**SEMAPHORE**

Semaphore is the common term for a system of signalling using a pair of flags. It was devised in 1817 by Captain Marryat of the British Royal Navy, who adapted it from that organization’s flag code of 1799. Although clearly
developed for use at sea, the Marryat code, or the Universal Code of Signals, as it is also known, is just as effective on land.

The flags are usually divided diagonally into red and yellow, and are moved like independent hands on a clock face. Each position represents a letter, so messages are spelt out. There are also set signals
for recurring content, such as ‘Error’ and ‘End’, while the digits zero to nine are represented by the first eleven letters of the alphabet (J also doubles up as ‘Letters follow’) when preceded by the message ‘Numbers follow’.

It is said that the semaphore flag system was the fastest method of visual
communication at sea – quicker than a flashing light using Morse code (pages 30–3) – and was valuable even in more modern times as it allowed ships to send each other messages while maintaining radio silence.
Semaphore code allows for communication across long distances.
Another flag signalling system used at sea is International Signal Flags, in which each
flag represents a letter or number, and also carries a message for common situations at sea: two-letter signals for emergencies, three letters for general information. This is particularly useful when ships are trying to communicate without a common language.
A/ Diver below (when stationary); I am undergoing a speed trial
B/ I am taking on or discharging explosives
C (affirmative)
D/ Keep clear of me, I am manoeuvring with difficulty
E/ I am altering my course to starboard
F/ I am disabled, communicate with me
G/ I require a pilot
H/ I have a pilot on board
I/ I am altering my course to port
J/ I am going to send a message by semaphore
K/ You should stop your vessel instantly
L/ You should stop, I have something important to communicate
M/ I have a doctor on board
N/ No (negative)
O/ Man overboard
P/ The Blue Peter – all aboard, vessel is about to proceed to sea. (At sea) your lights are out or burning badly

Q/ My vessel is healthy and I request free pratique

R/ The way is off my ship. You may feel your way past me.

S/ My engines are going full speed astern

T/ Do not pass ahead of me

U/ You are standing into danger

V/ I require assistance (not distress)

W/ I require medical assistance

X/ Stop carrying out your intentions and watch for my signals

Y/ I am carrying mails

Z/ To be used to address or call shore stations

**Numeric pennants**

0

1

2

3

4

5

6

7

8

9

**Distress**

**Answering pennant**
These flags enable communication across different languages.
The optical telegraph revolutionized communication.

THE OPTICAL TELEGRAPH

Copper pans, a clock face, a set of pulleys and the telescope all made a contribution to the development of a coded communication system that
had a major impact on 19th-century Europe, though it is now almost forgotten.

Developed by French brothers Claude and René Chappe in the late 18th century, the optical telegraph evolved from the materials listed above into a signalling system of movable black-and-white panels mounted on a beam with angled arms that
mimicked a person holding a pair of flags in different positions. It allowed rapid communication as far as you could see. Soon there were chains of optical telegraphs covering much of western Europe. This was revolutionary at a time when the fastest messages travelled at the speed of a galloping horse. The optical telegraph was a key pan-European
communication system until it was superseded by the electric telegraph, which used Morse code (see pages).

SIGN LANGUAGE

Sign language uses the oldest coding method of all: hand signals. Although used informally since the birth of humankind, the first books describing sign language
appeared in the 17th century, aiming to communicate with the deaf by using hand gestures as an exaggerated ‘mouth’. There are hundreds of sign languages in use around the world, the two most common being American Sign Language (ASL, see illustration, left) and Signed English. In addition, many sports officials use standardized
hand signals for some communication.
American sign language uses an alphabet for spelling out words.

The feeling code: Braille

Braille is a non-secret code that allows the blind to read. Like many codes, its origins lie in the need for the military to communicate without
being detected or understood.

DARK SECRET

The story begins with Artillery Captain Charles Barbier of Napoleon’s French Army in the early 1800s. Frustrated by the difficulty in reading messages safely on the front line (where smoke and the chaos of battle hindered communication, and
lighting a lamp created an easy target for the enemy), he devised a code using 12 raised dots on paper, called night writing. Unfortunately, his fellow soldiers found it too difficult to learn and it was not adopted.

Thinking that it might have a role in helping the blind to read, Barbier started to visit schools for the blind. In 1821
he demonstrated his code to a group of children at the Royal Institution for Blind Youth in Paris. The audience included the 12-year-old Louis Braille, who had been blinded by accident nine years previously. Braille quickly mastered the system but found he could simplify it to use just six dots. Although he published the first Braille book in 1837, it did not catch
on around the world for another 30 years.

Braille has since been adapted to nearly every language on earth and remains the major medium of literacy for blind people everywhere.
The Braille alphabet of raised dots includes numbers and punctuation marks.

- Each Braille character is made up of six dot positions, arranged in a rectangle.
- A dot may be raised
at any of the six positions, and a total of 64 combinations is possible, including the spaces (see opposite).

• There are Braille codes for representing shorthand, mathematics and music.
Braille characters use a system of raised dots.

The world’s most famous code: Morse

Morse code is the most famous code in the world. It allows rapid communication over long distances along
wires or radio waves, or via sound or light in an easy-to-learn system of dots and dashes.

CODE ON TAP

The American Samuel Morse had attempted various inventions and money-making schemes before he chanced across the opportunity to send messages
along wires. He enlisted the help of experts in the field of sending an electric pulse over long distances, and in 1838 designed a code that could be tapped out by hand. Showing admirable understanding of the practicalities of communicating language, he ensured that the most frequently used letters could be entered with the least effort, thus the code for ‘e’ is
a dot, and for ‘t’, it is a dash.

This combination slowly won over a sceptical public and the first telegraph line, following the 40-mile railway track between Baltimore and Washington, launched in 1844 with the message: ‘What hath God wrought’.

Morse code was treated originally as a novelty (public
chess matches were played via it) and struggled to overtake the already established optical telegraph (see pages), but gradually its advantages of cheap and very fast communication were recognized. In England, it enabled the arrest of the murderer John Tawell on 3 January 1845 when his description was sent ahead of the train on which he had
fled, leading to his arrest on his arrival at Paddington Station.
<table>
<thead>
<tr>
<th>A</th>
<th>N</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>O</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>P</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>Q</td>
<td>3</td>
</tr>
<tr>
<td>E</td>
<td>R</td>
<td>4</td>
</tr>
<tr>
<td>F</td>
<td>S</td>
<td>5</td>
</tr>
<tr>
<td>G</td>
<td>T</td>
<td>6</td>
</tr>
<tr>
<td>H</td>
<td>U</td>
<td>7</td>
</tr>
<tr>
<td>I</td>
<td>V</td>
<td>8</td>
</tr>
<tr>
<td>J</td>
<td>W</td>
<td>9</td>
</tr>
<tr>
<td>K</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>Z</td>
<td></td>
</tr>
</tbody>
</table>
Morse code is designed so that the most frequently used letters require the least effort.

BEYOND THE TELEGRAPH

In addition to its use in the telegraph system, Morse code is
also used:

• when armies and navies use a heliograph (which reflects the sun’s rays);

• when messages are sent with the Aldis lamp (a powerful light used at sea);

• in radio messages.
Airline pilots still have to know it, even today.

The telegraph system grew rapidly. Skilled operators soon learned to ignore the paper printouts of messages and instead listened to the clicking of the receiving apparatus to understand the
message. They began using abbreviations for long common words or phrases and many variations of code words were introduced. This was encouraged by a thrifty public keen to keep the length of messages to a minimum as payment was by the word. In the American Civil War (1861–5), the telegraph was used widely for the first time in warfare, and it was
imperative to encrypt important messages because of the ease of interception.

Banks, in particular, were keen to develop secure codes as it allowed them to transfer money electronically. By 1877, nearly $2.5 million was being telegraphed every year in nearly 40,000 separate transactions. Because the messages were legally
binding contracts, even marriage ceremonies were known to be conducted using the equipment, with the bride and groom clicking out ‘I do’ while a congregation of telegraph operators down the line listened in.

Morse code underwent a few refinements as it continued to serve as a code in radio communication. It remained
in use on the seas until it was replaced by a satellite-based communication system in 2000.

**SOS**

The dit-dit-dit-dah-dah-dah-dah-dit-dit-dit SOS distress call does not stand for ‘Save
Our Souls’, as many people believe. It is a code indicating that the operator will no longer be able to send messages. The lack of gaps shows that it does not represent individual letters.
New global codes

Communications technology, such as computers and mobile phones, has spawned a new set of codes aiming to save costs and time by making communication as brief as possible. These are simply modern forms of shorthand, but they also carry great social kudos for the young.
LEET

Around the world, people are creating a new code language at their keyboards, which is the digital equivalent of Pig Latin (see pages) with a few hieroglyphs thrown in for good measure. Widely used in chat rooms, among hackers and computer gamers, leet, or leetspeak (leet is a corruption of ‘elite’), enables rapid
communication to take place by using keyboard characters as short cuts for sounds, words and phrases.

Some basic rules for leet are:

• Numbers can replace the letters they resemble, so ‘leet’ is ‘1337’, ‘B’ (hence ‘be’) is represented by ‘8’, ‘9’ replaces ‘G’, ‘5’ or ‘$’ stands for ‘S’ and ‘4’ is ‘A’. Thus ‘leetspeak’ is
‘13375p34k’.

- Letters substitute for sounds, so words ending in C or K now end in X, and Z supplies ‘S’ and ‘ES’ endings.

However, there are many variations. For example, the first three letters of the alphabet can be shown as:

\[ \text{A: 4,} \Lambda, \text{ @, } /\backslash, \]
One of the rules of leet is that there aren’t really any rules: spelling and grammar conventions are largely ignored and the idiom is constantly evolving. It is particularly popular in communication between computer gamers and hackers.
Here are a few leet words and phrases:

- k3wL = ‘cool’
- m4d sk1llz = ‘mad skills’ – a talent
- n00b = ‘newbie’ – a newcomer
- y0 = ‘yo’, an alternative to ‘hi’.
Txt tlk, or txt lingo, is the language of the mobile phone, so it has similar aims to leet but has been adapted to suit use on a handset. However, it is also used in internet chat rooms along with leet. It developed through the introduction of the Short Messaging Service (SMS), which allows the sending of text messages of up to about 150 characters
and is used for conversations and news and financial information services. The first message was sent in 1992, and just over a decade later texts were being sent at a rate of 500 billion per year.

Some basic rules for txt tlk are:

- Vowels are omitted.
- Whole words can be left
out if the sense is not affected.

- SpacesAreMarkedByCapital

Txt words and phrases:

2 = ‘to’ or ‘too’
4 = ‘for’ as in ‘b4’
C = ‘see’
R = ‘are’
Y = ‘why’
8 = the syllable ‘ate’ as in ‘gr8’ or ‘m8’
bf or gf = ‘boyfriend’ or ‘girlfriend’
thx = ‘thanks’
np = ‘no problem’

Keyboard or handset characters are also used to create pictorial messages such as these ones (try tilting your head if you don’t see how it works):

O:-) = angel
:-! = bored
% ) = confused
Secret Communication

Hidden messages

The best place to hide a message is somewhere
completely innocuous. In the course of history, messages have been concealed inside animals and people, pencils and even coins, and that humble material wax has been a friend to many a spy.

HARES AND HAIRS

The principal motive for sending secret messages was for military or diplomatic
confidentiality. Generals and senior officials wanted to know that even if their message carriers were intercepted, no one would understand what information they had with them.

The early Greek historian Herodotus, writing in about 440 BC, tells a trio of tales involving concealed messages. In the 6th century
BC, Harpagus, a Median soldier, was plotting against Astyages, his own king. He wrote a message to the king’s enemy, Cyrus, promising to change sides, and hid it in the belly of a hare, which was carried by a messenger in the guise of an innocent hunter. Its delivery prompted a Persian revolt against the Medians, and Harpagus did indeed betray his king, who
was replaced by Cyrus. The idea of hiding a message in an animal was still around in 16th-century Italy, when Giavanni Porta records the practice of feeding a message to a dog, which could then be taken on an apparently innocent trip where it would be killed to retrieve the information.

In another complicated plot,
this time by the Greeks against the Persians, Histiaeus earned a place in the pantheon of secret messages. Wishing to prompt a rebellion against Darius, but stranded at court in his role as ambassador, Histiaeus somehow had to encourage his son-in-law Aristagoras to attack the city of Miletus. Clearly prepared to sacrifice speed for secrecy, Histiaeus
shaved a slave’s head, branded his message onto the poor man’s scalp and only sent him on his way when the hair had grown over the writing. It worked: the city was taken by the Greeks and established as a democracy, prompting further rebellions against their Persian rulers.

WAX FACTS
In 480 BC the Persians completed a five-year military build-up and launched an attack on Athens, believing their plan to be a secret. However, a Greek called Demaratus had witnessed their preparations and managed to get a warning to his compatriots. Messages were usually sent on wax tablets, but these would obviously not be kept secret.
So Demaratus instead carved his message into the wooden base of the tablet, which he then covered with wax. The blank tablets arrived without incident, but baffled their recipients until they scraped away the wax and discovered the warning. This allowed them to repel the attack, which was thought to be crucial in preserving their independence.
The Greeks were not alone in their need to communicate secretly, or in their use of wax. The ancient Chinese wrote messages onto silk, which was then scrunched into a little ball and coated with wax. The messenger would hide or swallow the ball to conceal it during his journey. Roman historian Tacitus tells of wounded soldiers concealing writing on
their bandages, of sewing messages into the soles of sandals, and even writing messages on a thin sheet of lead and rolling it into an earring. More recently, Cold War Russian spies hid microfilm in hollowed-out pencils, batteries and coins.
COVERT WRITING

• The term for hidden writing is steganography, from ‘stega’, the Greek word for roof or cover, and ‘graphy’, meaning writing.

• The practice
continues today: in electrical communications it is called transmission security.

What better way to conceal a
message than by rendering it invisible? It saves all the effort of devising and using a code and allows open, apparently innocent communication between people who are being observed. The technical term for this is steganography. Steganography can also be used for secret communication, embedding
data on the screen rather like a microdot hidden in a piece of punctuation. Its advantage is that an interceptor is unlikely to detect the presence of the data, allowing secure communication, albeit of fairly short messages.

HOODWINKING INKS
A number of natural materials have been used for invisible
writing for thousands of years. Both the Greeks and Romans extracted such inks from nuts and plants. For example, in about AD 100 the Roman writer Pliny recorded that he could obtain a liquid from the tithymalus plant (which is part of the euphorbia family). When he wrote with it, the message vanished as the ink dried, but it reappeared when the paper
was gently heated.

Pliny could have chosen one of many other organic liquids such as onion juice, vinegar and apple juice, and any citrus juice. All turn brown when gently heated (a hot iron, hairdryer or light bulb is best – actual fire is too hot). This knowledge has benefited secret agents across the centuries, some of whom
resorted to their own urine when other supplies ran out. Cola drinks also work with this method (use the non-diet kind, as sugar is required).

A large number of chemicals also function as invisible ink (also known as ‘sympathetic ink’). They are activated by another chemical called the ‘reagent’. Examples include iron sulphate solution, which
reacts to solutions of potassium cyanate, or sodium carbonate and copper sulphate, which react to ammonia fumes.

Some of these chemicals can be harmful, so do not try this yourself unless you are a chemist or a trained spy! However, one experiment you can do is to write in milk on thick paper, using a brush
rather than a pen so you don’t leave any indentations on the paper. Watch as your words disappear, then rub dark powder such as ashes or charcoal across the page, and your message returns. In Nazi Germany, ballot forms were secretly numbered in milk to allow checking of how people voted in plebiscites.

Banks and amusement parks
sometimes use invisible inks that shine under ultraviolet radiation known as ‘blacklight’. The inks contain colour-brightening chemicals similar to those found in laundry washing powders.

INVISIBLE DIGITAL WRITING

An intriguing by-product of
the internet age is the re-emergence of secret writing as a valuable tool. If you’ve ever been baffled by how some seemingly unrelated websites appear when you are searching for something on the net, the answer echoes the methods described above. It is possible to hide text on the screen either by typing white text on a white background, concealing it on a non-
printable area of the page, or in graphics or a music file. You can’t see it, but the search engines find it. This embedded text gets their website included in a wider range of search results. Some companies go a step further and hide their competitors’ names on their own websites, meaning that their site pops up on the screen whenever someone keys in the name of
their rival.

**INKY TRICKS**

Sending sheets of ‘blank’ paper is likely to arouse suspicion, so agents writing in invisible ink do so on paper with an innocuous message, or use shopping lists or pictures as a background.
Natural invisible inks are made from organic fluids, which are rich in carbon. The liquid evaporates but some soaks into the paper. When heated, the carbon chars and turns brown.
The use of invisible inks was widespread in the Middle Ages, through the Renaissance and was still an important skill for spies during World War I. German agents were by then disguising possession of such inks by impregnating items of clothing with the liquid and
then activating it by soaking the garment in water. In the later stages of the war, the American Military Intelligence Division MI-8 was testing 2,000 suspicious letters a week for secret ink. Its work led to the capture of German spy Maria de Victorica.

Spoken codes
Speaking directly to someone else is, of course, the most efficient way of communicating, but when there is danger of other ears listening in, a spoken code is required. Most of them are quite simple to break once you know the key.

IGPAY ATINLAY

The simplest spoken code is
Pig Latin, a letter rearrangement code that is particularly popular with children. There are three basic rules:

- Words that start with vowels have ‘ay’ added to the end, so ‘actually’ becomes ‘actuallyay’.

- For words starting with a consonant, that letter is moved to the back, and then
‘ay’ is added at the end, so ‘can’ becomes ‘ancay’.

- If two consonants are at the start, they are moved to the end, adding ‘ay’, so ‘speak’ becomes ‘eakspay’.

Thus, ‘Actually a child can speak Pig Latin well’ is spoken as, ‘Actuallyay aay ildchay ancay eakspay igpay Atinlay ellway’. With practice, Pig Latin can be
spoken and understood at quite a pace. In a variation known as Tut Latin, the sound ‘tut’ is added between each syllable.

Another spoken code language is Opish, in which ‘op’ is added after each consonant. Thus ‘book’ transforms into ‘bopookop’ and ‘code’ is ‘copodop’. Words become very long and
it becomes very hard to decipher meaning. Similar to Opish is Turkey Irish, in which ‘ab’ is added before every vowel sound, so ‘book’ becomes ‘babook’ and ‘code’ is now ‘cabode’. As with Pig Latin, this is a novelty language rather than a code that can be used for meaningful communication.

A more sophisticated spoken
code disguises information. For example, just prior to the German occupation of Norway in 1940, telephone and radio calls from Nazi agents were intercepted. They appeared to be sending sales and tonnage information about fishing, but analysis showed that they were actually communicating details of ships, using the numbers with which vessels
are identified in the shipping bible *Lloyd’s Register*.

Later in World War II, innocent-sounding calls discussing the flower market were also found to be disguising information about which ships were in harbour and the nature of repairs being undertaken.

TAKE A
BUTCHER’S AT THIS

Cockney rhyming slang is a spoken code that has survived for about 200 years. It substitutes (usually) two words for one, with which it rhymes. So ‘butcher’s’ in the heading above is short for ‘butcher’s hook’, meaning ‘look’. Traditional examples are:
• Apples and pears = stairs
• Barnet (fair) = hair
• Brown bread = dead
• Canoes = shoes
• Dickie Dirt = shirt
• Mahatma (Ghandi) = brandy

The slang continues to develop. Here are some more recent introductions:
• Basil (Fawlty) = balti
• Billie (Piper) = sniper or windscreen wiper
• Metal Mickey = sickie

There are various theories about how Cockney rhyming slang started. It was certainly in London’s East End, and definitely devised to prevent other listeners from knowing what was being said. It may
have originated from:

- Villainous builders in the London docks.
- Market vendors talking about customers.
- Prisoners who didn’t want their guards to understand what they were saying.
- Thieves aware that Robert Peel’s newly launched police force might be
NATIVE TONGUES

All of the spoken codes described so far can be broken relatively easily. One spoken code that defied analysis, however, was used by the US forces in the two world wars: a language that none of the opposing forces could understand.
In 1918, orders within D Company, 141st Infantry were openly transmitted by field telephone in complete security because they were spoken by one of eight serving Choctaws, a native American group from Oklahoma.

Other native tongues were also used. The unique advantage was that these
languages had developed, geographically and linguistically, far away from other peoples and conveyed meanings with precise pronunciation and hesitations, which were unintelligible to outsiders.

The practice was repeated in World War II, when as many as 420 Navajo speakers were used by the Marines in the
Pacific combat zone, baffling Japanese intelligence personnel. The story was told in the 2002 film *Windtalkers*, starring Nicolas Cage.

**Words within words**

Hiding words within words has proved to be a popular
method of secret communication as it is very hard to detect, the carrier text being an everyday innocuous message.

READING BETWEEN THE LINES

In 17th- and 18th-century Britain, it was very expensive
to send letters by post, but sending newspapers was cheaper and even, at times, free. Thrifty communicators seized on this and adopted the puncture code that had been described by Greek historian Aeneas the Tactician 2,000 years before.

They would make small pinpricks, or mark tiny dots, over certain letters so that
they could spell out a message, which could then be cheaply dispatched to their correspondent. The practice continued until postage prices were altered in the middle of the 19th century. However, German spies used this exact system during World War I, and again in World War II, using invisible ink.

The disadvantage of this
method is that a lot of the carrier text is redundant: so many words are delivered from which only fairly short messages can be communicated. The next logical step is to write your own ‘carrier’ message, which can be decoded with an agreed formula.

A simple example of this is an acrostic: a sentence in
which the initial letters of words spell out a message. These are called null ciphers and are often employed in puzzles and crosswords. For example, cuddly attack tiger spells CAT. They are also used as memory aids, for example BRASS is an acronym for how to fire a rifle: Breathe, Relax, Aim, Sight, Squeeze, and the sentence Every Good Boy
Deserves Favour sets out the letters on the lines of sheet music when written in the treble clef.

This method is thought to be the inspiration for the Christian sign of the fish. In the early days of Christianity, when it was necessary at times to keep your faith secret, ancient Greek was widely spoken. The phrase
‘Jesus Christ Son of God, Saviour’ rendered in ancient Greek is ‘Iesous Christos Theou Uios Soter’. The first letters spell ICHTHUS, the Greek word for ‘fish’. So followers of Jesus could use a fish sign, or the word ICHTHUS, to show that they were Christian.
In general, devising an innocent-sounding acrostic message that makes sense is very tricky. However, choosing to send a message via, say, every third letter is far easier to devise, and harder to spot. A famous example of this is the English Civil War story of Sir John Trevanion, a royalist locked in Colchester Castle awaiting probable execution by his
Cromwellian captors. He received a letter that began: ‘Worthie Sir John: Hope, that is ye beste comfort of ye afflicted, cannot much, I fear me, help you now. . .’.

On the surface it was just a rather verbose letter, and the guards charged with checking his letters could find nothing suspicious. But take out every third letter after a punctuation
mark and the entire letter creates the more useful message: PANEL AT EAST END OF CHAPEL SLIDES. Sir John promptly asked to be allowed to pray at the chapel, and made his escape. Some historians question details of this tale, but it does illustrate the value of an acrostic code.

A similar strategy was used by a German spy in World
War II, whose message ‘Apparently neutral’ s protest is thoroughly discounted and ignored. Isman hard hit. Blockade issue affects pretext for embargo on by-products, ejecting suets and vegetable oils’, spells out ‘Pershing sails from NY June 1’ if you read only the second letter of each word.

A similar simple encoding
method is to conceal the true message with extra words, and include guidance on how to identify which words to read. For example, the 1/4 at the start of this message instructs the reader to ignore the first and every fourth word: ‘Tom: 1/4 gifts do not arrive often father will be pleased here, Lucy’ so the plaintext reads ‘Tom do not arrive father will be here,
Another method of sending a message concealed on a page of writing is the stencil method, in which the hidden text is read through ‘windows’ cut into card or fabric laid over the letter.
THE CARDANO GRILLE

This code was invented by an Italian doctor and mathematician called Girolamo Cardano in 1550 and is known as the Cardano grille. Small holes were punched in an irregular pattern in a piece of card, which was used as an overlay on top of a piece of writing. This method allows for
reading only single letters at a time, but it can be adapted to use larger holes so that syllables or whole words appear in the window, although this then makes it more difficult to disguise the message.
this method was used in circles for many years
The three stages of creating a Cardano grille
this method was used in circles for many years
Today this is not a method anyone teaches. It was sometimes used by artists in official symbols because squares and circles are better for producing many simple designs without years of practice.
Provided the same card with cut-outs is not used more than a few times, the method is safe, and it was certainly widely used in diplomatic correspondence for hundreds of years after its invention. The system also requires there to be two identical copies of the stencil, or that it be sent to the receiver in some guise. But over-use of
the card would allow a code breaker to identify where the ‘windows’ are, allowing instant reading of the real message. There is also a risk that the message created to ‘camouflage’ the hidden words may be so clumsily worded as to create suspicion.
A refinement of the method is the ‘turning grille’ system, in which the grille is rotated through 180 degrees or flipped every, say, nine
letters. This is remarkably effective, provided the windows never overlap.
3

Simple codes

Making simple codes and ciphers
Simple codes and ciphers can be created with pencil, paper and patience. This chapter describes various examples of simple codes, which, while being quite basic in formulation, are in fact a very effective method of concealing meaning, and have been used by secret agents and spies many times.
Perhaps the simplest way of concealing a written message is a space code, in which the plaintext is broken up into different ‘words’. For example ‘This is an example of a space code’ could be encoded as THI SISA NEXAM PLEO FAS PAC ECO DE. It wouldn’t fool anyone for long, though.
Neither will a backwards code: EDOC SDRAWK CAB A LLIW REHTIEN. This message can be made slightly less recognizable by breaking up the word groups: EDO CSD RAWK CAB ALL IWREH TIEN.
A ‘code’ is a system in which words and/or phrases are changed, and therefore requires a code book, which is like a dictionary. In a ‘cipher’, the substitution is of letters, so no code book is necessary, and deciphering requires...
knowledge of how the letters have been changed. An advantage of ciphers is that the same letter can be changed to many different letters or numbers, making it much harder to detect (see page). Turning a message into code and then enciphering that text is called
A history book written in 14th-century Egypt notes that tax and army officials used the names of perfumes, fruits, birds and flowers to denote certain letters or terms. This kind of code is pretty much unbreakable, but there are difficulties in execution:
• All correspondents need a copy of the code words in use, and this document will be bulky and hard to conceal.

• Anyone able to see this code book can understand your messages. You would also probably need to expand the code book to allow you to use a wide vocabulary, so there would be security issues about this
• If someone studies your conversation or writing, they will be able to make informed guesses about the types of words in the gaps (verbs, nouns, numbers, etc.) and eventually decipher at least parts of it, simply from the context.

• However, it is very easy to create a code allowing you
to talk or write to a friend with no chance of others understanding your communication. All you have to do is agree to change each important word for another word, rendering the plaintext incomprehensible.
TRANSPOSITION CIPHERS

• A cipher in which letters are rearranged, as in the backwards cipher on page 53, is known as a transposition cipher.

• Other examples in
this chapter are the rail fence cipher on pages 64–5 and the Greek scytale (see pages).

• Transposition ciphers are easy to spot if you analyze the letter frequency (see pages).
Codes where words stand for other words or phrases are known as sub rosa codes. A famous story involving their use comes from World War I, when censors who were suspicious of the cable message ‘Father is dead’ amended it to ‘Father is deceased’. This caused confusion with the recipient, who cabled back, ‘Is father dead or deceased?’
The practice was also widespread among spies in World War II. A series of intercepted letters contained detailed concerns apparently about the correspondent’s doll collection such as, ‘A broken doll in a hula grass skirt will have all damages repaired by the first week of February.’ It was eventually established that each doll was code for a different American
One way to avoid having to create and update a code book is to use a dictionary code for all, or some, words. This has proved a popular method for secret agents countless times. When you wish to disguise a
word, you quote the page, column and entry number where it appears in your dictionary: you just have to make sure both parties are using the same edition of the tome. For example, in the tenth edition of *Webster’s Collegiate Dictionary*, the word ‘dictionary’ is on page 322, column 1, and is the third entry, so its code number is 322,1,3. The
comma is required to avoid confusion between the three pieces of information: for example 32213 could mean page 32, column 2, entry 13, which is ‘allow’. The number for ‘code’ is 221,2,17.

Thus the plaintext, ‘This sentence is in dictionary code’ would be encoded as:

‘THIS SENTENCE IS IN
322,1,3 221,2,17’ if you just change two key words, or if you change them all:

1227,1,14   1067,1,13
620,2,11   585,2,6
322,1,3 221,2,17

BREAKING A DICTIONARY CODE

Anyone trying to decode this
message would immediately know how many words it contained by counting the spaces between numbers. The use of three numbers for each word also betrays the ‘page, column, entry’ format, especially as the middle number is always 1 or 2. So you would quickly discern that it is a dictionary code, but would be unable to proceed further without
identifying which dictionary was used.

Refining the Code

Armed with this knowledge, you might decide to conceal the page numbers by making every page number four digits, leaving the column number untouched, but making all the entry numbers two digits, filling in gaps with
zeroes. Now you can run all the numbers together and remove the commas:

12271141067113062021

CODES IN BOOKS
• Jules Verne’s *Voyage to the Centre of the Earth* features a baffling code that turns out to be Latin written in reverse that can be read through the back of the paper.

• William Makepeace Thackeray included a Cardano grille (*see pages*) in *The
History of Henry Esmond.

- H. Rider Haggard used a cipher in *Colonel Quaritch, QV*.

- Agatha Christie used a flower-names code in *The Four Suspects*, solved by the indomitable Miss Marple.
• Mystery writer Dorothy L. Sayers used a message in Playfair cipher (see pages) as a fundamental part of the plot in her Lord Peter Wimsey novel, *Have His Carcass*. He solves it by guessing that the message starts with the name of a city.
and then a year, providing him with a crib.

Another refinement is to use any odd number instead of ‘1’ to indicate the entry is in the left-hand column, and any even number in place of ‘2’ for the right-hand column. This will confuse the decoder.
Yet another trick is to add the same number to every figure. For example, adding five to each figure in the code above would produce this ciphertext:

123261910726180625716059071103276080226722

A determined decoder will still be able to recognize number patterns, although
you can further confound them by changing the procedure slightly and locating the words in the same position but four pages in front of the word in the plaintext. Now your message reads:

**THERAPEUTICS**
**SEMESTER**
**INVESTITURE**
The ‘four pages back in the dictionary’ code is exactly the method unearthed in the scandal surrounding the disputed 1876 American Presidential election in which both Democrats and Republicans were suspected of malpractice. In one
Democrat message, the phrase, MINUTELY PREVIOUSLY READMIT DOLTISH, translates as, ‘Must purchase Republican elector’, via the Household English Dictionary published in 1876.

**BOOK CODE**

Any book can function as a code book in much the same
way as a dictionary, and this again, for centuries, has been a very common encoding method. One problem, though, is in finding the words in the book in the first place. For example, a message using this system from Benedict Arnold, the 18th-century American general who defected to the British, found him culling words from pages as far apart
as 35, 91 and 101. That’s a lot of searching for one word – it would be infinitely easier today using the search facility on a word-processing program.

The disadvantage of using a dictionary or book for creating a code is that many words, such as place names, simply won’t be in the text and will then need to be
laboriously spelled out letter by letter, with the code specifying the page, line, word and letter numbers. This will create a very long message, which occupies hours for both encoder and decoder. However, provided the interceptor never identifies the book used, it is a very secure method, especially if the book is changed frequently, which
explains its popularity. Another benefit is that the shelves of secret agents will not be stacked with incriminating code books or stencils, but merely hold a stock of seemingly innocent books.

THE DOT CIPHER

This is totally different to the practice of dotting or
pinpricking newspapers favoured by thrifty Victorians (see page). The alphabet is written on squared paper with one letter on each line and the receiver needs an exact copy. The ciphertext is created by putting a dot under each letter in your message, working down the page so that each new dot is on a new line. The end result looks like piano roll music for an automatic
piano. So the message, ‘Dots lines and zigzags’ will look like the illustration below. The alphabet can also be written vertically rather than horizontally, in which case the dots will read from left to right.

You can disguise the message by connecting the dots to make lines, a graph or even a crazy picture, or go in
sequence to create a zigzag pattern as you can see in the illustrations below. The dots from which these variations are produced must be positioned precisely to avoid confusion in decoding.
The dot cipher.
A refinement of this system is to set the alphabet in a different, pre-arranged order, such as backwards or with all the vowels first.
The line and zigzag ciphers.
Transposition ciphers

In transposition ciphers, letters are rearranged in a different order, creating an anagram of the message. There are various systems determining how to change the position of the letters to form a cipher.
THE GREEK SCYTALÉ

The scytale is the earliest known piece of cryptographic equipment, dating from 5th-century BC Greece. Probably first used by the Spartans to carry messages around the battlefield, it is a simple transposition machine. A piece of parchment rather like a ribbon was wrapped around
a cylinder, such as a wooden staff. The message was then written unencrypted onto the coiled paper.

Once removed from the rod, the writing was just a jumble of letters that would be meaningless if the enemy captured it. It is possible that messengers wore the fabric as a belt with the writing on the inside. When the message
was delivered, it was wrapped around a cylinder identical in diameter to the first one, and could be read.

A CONE TO DECODE

A scytale message can be decoded by
simply wrapping the material around a cone and sliding it around until the text makes sense.

Part of the scytale’s value was the speed at which communication took place, because no enciphering and deciphering was involved: the
message was written, transported with reasonable security, and read.

THE RAIL FENCE CIPHER

In the 19th century, hundreds of miles of fencing was put up across the US as new territories were taken over. Known as split-rail fences,
they form a zigzag pattern when seen from above, similar to the pattern made by letters in the rail fence cipher. If you write the message, ‘The rail fence cipher makes a zigzag pattern’ in zigzags, it looks like this:
The enciphered message is created by writing each row, choosing, if you wish, to put the letters into groups of four, in which case you will need to fill in the gaps with ‘padding letters’ or ‘nulls’, which are usually X or Z. The first null indicates where the new line starts. In this example, the last two nulls ensure the ciphertext ends with a group of four letters.
<table>
<thead>
<tr>
<th>TEAL</th>
<th>ECCP</th>
<th>EMKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZGAP</td>
<td>TENX</td>
<td>HRIF</td>
</tr>
<tr>
<td>NEIH</td>
<td>RAEA</td>
<td>IZGA</td>
</tr>
<tr>
<td>TRZZ</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To decipher this message, count the letters (40) then divide into two groups. You can now put the letters into order by writing the first letter from each group, then the second, and so on,
ignoring the nulls, and reading the words created.

Refining the rail fence cipher

The cipher can be written in three or more rows, either zigzagging across the page or starting a new column every three letters. This is the code written in zigzag pattern for ‘A three row cipher would
look like this:

This creates the code message:

AEWH OLLT TREO CPEW UDOK IEHS HRIR LOKI
Transposition ciphers require a lot of laborious work to prepare, and there is plenty of scope for mistakes, so they have not been used as widely as some other ciphers in the past.
However, they are effective and the concept lies behind many modern computer-driven cipher algorithms.
When adding nulls to pad out a grid, the number of letters added can be indicated by the final letter’s position in the alphabet. So, a one-letter null would be an A, and if four letters were added, the last one would be D.
DECIPHERING THIS RAIL FENCE CIPHER

The message can be revealed by writing out the eight letters of the top row, the 16 in pairs of the middle row, and the eight of the bottom row, recreating the zigzag pattern of the original. Then the letters are written out in the new order reading down and up.
BREAKING RAIL FENCE CIPHERS

Like other transposition ciphers, the cipher features letter frequencies similar to those of the language in everyday use (see pages), so you can identify them by counting letter frequency. Now you need to unscramble the letters on the page, which is much easier than finding
substitutes for them.

To decipher a transposition cipher you need to identify then ignore the padding nulls – frequency analysis (see pages) will help here – then try reading every second letter, and if that doesn’t work, every third letter, and so on. The longer the message, the bigger the jumps will have to be.
Once the letters are put into a four-column grid, they do not have to be enciphered in the standard left to right, or top to bottom order. For example, your key could be to start at the bottom right corner and spiral clockwise to the centre. This is called a route cipher and makes decryption much
more difficult. The encryption of the same four-row message already used:

```
T I I N O
H S T I W
I W T N S
S R E R A
```

would then begin at the A in the bottom right corner and read as:
Other paths for enciphering include:

• In a spiral from the centre.

• Diagonally (specifying upwards or downwards, left to right, or right to left).

• Up one column, down the next.
Another way to scramble the letters from a grid is to identify the columns with a key word or number. This is called columnar transposition. If you have a four-column grid with the message, ‘This makes it more complicated’ written across, it looks like this:
Two nulls have been added, the last being B to indicate ‘two nulls’. Making the other nulla T rather than, say, an X makes them harder to spot as
imposters.

THE KEY

The encipherer and decipherer agree in advance how many rows the ‘fence’ will have, and, if necessary, the direction of writing
(e.g. forwards, backwards and whether using diagonals or columns). This information is called the key and allows the message to be rapidly deciphered.

Instead of writing the encryption out by following
the column order, you can change it with the four-letter keyword CODE (one letter for each column). In alphabetical order within the word, these letters are 1st, 2nd, 3rd and 4th. This would re-arrange the grid to read:

<table>
<thead>
<tr>
<th>C</th>
<th>0</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>4th</td>
<td>2nd</td>
<td>3rd</td>
</tr>
<tr>
<td>T</td>
<td>S</td>
<td>H</td>
<td>I</td>
</tr>
<tr>
<td>M</td>
<td>E</td>
<td>A</td>
<td>K</td>
</tr>
</tbody>
</table>
In blocks of five (with a three on the end, which could be filled with nulls if you choose), the enciphered message now reads:
Deciphering the Message

The decipherer now works out the column lengths by dividing the key length (four, from the keyword CODE) into the message length (28
letters). This reveals the number of rows as seven, so the content of each column can be identified, then reordered according to the code word, the nulls counted and removed, and the message read.

BOOK CODES
Book, and especially Bible, codes are popular in Hollywood movies.

• In the 1996 film *Mission: Impossible* there are several references to Job 3:14: ‘with kings
and their advisors whose palaces lie in ruins’.

• The 2002 thriller Red Dragon features numerous apparent Bible citations, which turn out to refer to a different book, The Joy of Cooking. This may be less of a surprise when you
consider the film is a sequel to the cannibalistic *The Silence of the Lambs* (1991).

• *National Treasure* (2004) has a plot based around a code hidden in the US Constitution revealing the whereabouts of a treasure buried
during the 1700s.

DOUBLE TRANSPOSITION

A technique for breaking down transposition ciphers is to guess the number of rows, group letters accordingly, then slide the letters around
looking for words or anagrams. Double transposition counteracts this by repeating the scrambling of columns during encryption, usually with a second keyword. Both keywords can be changed at will to protect the cipher from attack.

Double transposition was used by the German Army
during World War I, but it was successfully broken by the French. They were greatly aided by the fact that the Germans, confident of the security of their cipher, used the same key for more than a week at a time – a major sin in the world of cryptography (see pages). Double transposition was widely used in World War II as well, as it was regarded as the most
complex cipher an agent could use as a field cipher.

Substitution ciphers

Transposition ciphers create anagrams of the plaintext by mixing up the words or letters. Substitution ciphers leave the letters in the order
they should be read, but disguise them.

SHIFT CIPHERS

An early example of this comes from 2,000 years ago in a message sent by Roman leader Julius Caesar to Cicero, whose forces were under siege. The Roman letters were substituted with Greek letters, which Caesar
knew Cicero would understand.

MONO TO POLY

- A Substitution cipher that uses one alphabet for encryption so that each plaintext letter
is represented by the same ciphertext letter throughout is described as ‘monoalphabetic’.

• Later ciphers that used more than one alphabet are known as polyalphabetic ciphers (see pages).
Caesar had many reasons to encrypt his messages and did so in many ways. The most famous is the shift cipher, in which each letter is replaced by the letter three places on in the alphabet: ‘a’ becomes D, ‘b’ becomes E, ‘c’ is F and so on. The message, ‘Named after Julius Caesar’ would be written as:
The Caesar shift was widely used for centuries – it was even one of the ciphers being used by Russian forces in 1915. A big advantage of the shift cipher is that it does not require a code book as the method can be easily memorized. It can also be
adapted to shift the letters any number of places from 1 to 25 for a standard 26-letter alphabet through the use of a code number. This is called the St Cyr cipher, after the French national military academy where it was taught in the 1880s.

However, it is also known as the slide rule cipher because it can be created by sliding an
alphabet strip below an identical strip to create the shifted letters. So the code number 7 would indicate a 7-place shift, creating this alphabet:

<table>
<thead>
<tr>
<th>Plain alphabet:</th>
<th>a b c d e f g h i j k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cipher alphabet:</td>
<td>H I J K L M N O P Q R</td>
</tr>
</tbody>
</table>
BREAKING THE SHIFT CIPHER

However, the fact that the cipher is in alphabetical order makes this kind of shift key very easy to break. All you have to do is take one word or set of letters and try out all the possible encryption keys. For example, if the cipher text includes the letters SIFBVE, the word can be
discovered through the process shown in this table:

<table>
<thead>
<tr>
<th>Shift</th>
<th>Produces letters</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>SIFBVE</td>
</tr>
<tr>
<td>1</td>
<td>TJGCWF</td>
</tr>
<tr>
<td>2</td>
<td>UKHDXG</td>
</tr>
<tr>
<td>3</td>
<td>VLIEYH</td>
</tr>
<tr>
<td>4</td>
<td>WMJFZI</td>
</tr>
<tr>
<td>5</td>
<td>XNKGAJ</td>
</tr>
<tr>
<td>6</td>
<td>YOLHBK</td>
</tr>
<tr>
<td>7</td>
<td>ZPMICL</td>
</tr>
<tr>
<td>8</td>
<td>AQNJDM</td>
</tr>
<tr>
<td>9</td>
<td>BROKEN</td>
</tr>
</tbody>
</table>
So shifting the encrypted message nine places along the alphabet solves the cipher.

THE SHADOW

The Shadow magazine published serialized stories for 18 years
from 1931, and remains a cult publication. The tales of the mysterious sleuth, written by newspaperman and magician Walter B. Gibson under the pen name Maxwell Grant, featured various codes in various ways. One story, ‘Chain of Death’, features an
inventive alphabet, and is illustrated below.
<table>
<thead>
<tr>
<th>A</th>
<th>H</th>
<th>O</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>I</td>
<td>P</td>
<td>W</td>
</tr>
<tr>
<td>C</td>
<td>J</td>
<td>Q</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>K</td>
<td>R</td>
<td>Y</td>
</tr>
<tr>
<td>E</td>
<td>L</td>
<td>S</td>
<td>Z</td>
</tr>
<tr>
<td>F</td>
<td>M</td>
<td>T</td>
<td>Space</td>
</tr>
<tr>
<td>G</td>
<td>N</td>
<td>U</td>
<td></td>
</tr>
</tbody>
</table>
So far this is just a graphically pleasing substitution alphabet, but the cipher was later refined with an additional four symbols:

Each symbol indicates
a degree of rotation, adding three levels of transposition to the alphabet. For example, a quarter turn shown by symbol 2 transforms the character for ‘a’ into that of B, and ‘c’ becomes D.

The Shadow was brought to life in a
radio serial in the 1930s, a period when public interest in secret codes was high. Secret decoders were popular toys and promotional gifts from the 1930s onwards, especially from manufacturers of children’s drinks and cereals. The ‘Captain Midnight’ radio serial in the 1930s and 1940s
included secret codes giving clues about the next episode and listeners could send in for a working cipher disc as part of a ‘Spy-Detector Writer’ kit, which could be used to decrypt messages broadcast in the show.
EARLY SUBSTITUTION

A 10th-century Persian substitution alphabet used the names of birds for letters of the alphabet. Another substituted them with names for parts of the night sky.
Another approach is to start the alphabet with a key word, followed by the remaining letters in alphabetical order. This allows regular changing of the cipher by replacing the keyword and enhances...
security of the system. Repeated letters in a code word are omitted (so omitted would be spelt with one ‘t’ as omitted). Here the keyword is ‘scramble’:

Plain alphabet:  a b c d e f g h i j k
Cipher alphabet: S C R A M B L E D F G H I J K
Notice that some letters stay the same in this cipher, a feature best avoided. There are a couple of ways of getting around this. The cipher alphabet can follow the keyword in any agreed order, so one with the keyword 'backwards', followed by the rest of the alphabet in reverse, would look like this:
Alternatively, the keyword need not come at the beginning of the alphabet. So if the keyword is ‘thirteen’ and happens to start on the 13th letter of the alphabet, you would produce this cipher alphabet:
Since one weakness of these substitution ciphers is their alphabetical order, the way to protect them is to put the alphabet into a different, random order. Theoretically this creates 403,291,461,126,605,635,584,
different possible cipher alphabets – more than someone could test in a lifetime, even if they were equipped with a computer.

However, because the cipher alphabet is random, it would have to be memorized by two people (which is unreliable) or written down (which threatens cipher security). It is also, as we will see later,
easily broken by frequency analysis.

**TYPEWRITER CIPHER**

The typewriter cipher substitutes the alphabet for the letters of the keyboard in order from top to bottom (the qwerty keyboard is a random order for which there seems
to be no explanation – it certainly doesn’t reflect frequency of use for its letters. However, have you noticed that the characters of ‘typewriter’ are all on the top row of letters?

Plain alphabet:  a b c d e f g h i j k
Cipher alphabet: Q W E R T Y U I O P A
Alternatives are for each key to represent, say, the one to its left, with the cipher wrapping round to the beginning when the end of the row is reached, or to the left and above, so that ‘typewriter cipher’ would be enciphered as:

5603248534 D80Y34
MORE COMPLEX SHIFT CIPHERS

A cipher with an alphabetical basis can be attacked fairly easily by cryptanalysts because there will be a pattern to at least parts of the cipher for the letters not in the keyword. One option is to re-order the alphabet, for example by writing it backwards:
This produces one of the most famous (and simplest) ciphers in history, Atbash, a cipher in which the first letter becomes the last, the second becomes the second last, reversing the alphabetical order. This was a device used by Hebrew
scribes to encipher parts of the Old Testament. ATBASH is so-called because in the cipher the Hebrew letter A becomes T, B becomes Sh, and so on, hence ATBSh, which is pronounced ATBASH. It is remarkably easy to break because there is only one solution!

Another option is to write all the vowels first, followed by
the consonants:

<table>
<thead>
<tr>
<th>Plain alphabet:</th>
<th>a b c d e f g h i j k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cipher alphabet:</td>
<td>A E I O U B C D F G H</td>
</tr>
</tbody>
</table>

Again, to prevent letters being encrypted as themselves, do this backwards:
If the cipher alphabet is not alphabetical at all, it is harder to break. However, both parties must remember, or keep a record of, the invented alphabet, which can weaken security.
Encrypting a word message into a number cipher is another way to disguise meaning. The most obvious way is to number the alphabet (‘a’ = 1, ‘b’ = 2, etc.), but this will be simple to detect (see frequency analysis, pages 89–95). All the variations in alphabet order described on pages 71–9 could be adapted...
to create a number rather than a letter alphabet.

Another way, which has more possibilities for deception, is a five-by-five grid, giving 25 squares. As there are 26 letters in the alphabet, this requires two to be combined: ‘i’ and ‘j’ will do. The grid looks like this:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
</table>


<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I/J</td>
<td>K</td>
</tr>
<tr>
<td>3</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
</tr>
<tr>
<td>4</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
</tr>
<tr>
<td>5</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
</tr>
</tbody>
</table>

Letters can now have a number created from their grid reference with the horizontal row number appearing first then the vertical column, so ‘b’ would be 12 and ‘s’ would be 43.
Thus the plaintext ‘number cipher’ would be encrypted as:

```
```

The alphabet can be written in any agreed order in the grid, perhaps using a keyword, and decryption is a simple matter of matching
letters to their numbers. One disadvantage of this system is that the ciphertext is twice the length of the plaintext, and writing out the enciphered message is double the work.
Whenever you are encoding, it is crucial to avoid making any mistakes, as these will confuse the recipient of the message. This is especially true when using numbers. Inputting the wrong number or, worse, omitting a digit, will render the message incomprehensible.
THE GREEK SQUARE

The numbering method on such a grid is known as the Greek square, or the Polybius square, after its inventor. Dating back 2,200 years to ancient Greece, it was used to send messages long distances.
by holding up torches, with the number held in each hand indicating the grid reference. There are records of a similar method being used in 16th-century Armenia to add a sense of mystery to religious texts.

The system lends itself to communication with lamps, smoke signals, knots or stitches in string or quilts, and
sounds. Translated into a system of knocks, it is a method for messaging between prison cells. It is said to have been used in this way by captives of the Russian Tsars, and by American prisoners of war during the Vietnam War.

Breaking the Greek Square
Since each letter is matched by one number, this cipher can be broken by frequency analysis (see pages), just like any other monoalphabetic substitution cipher. However, there are possibilities for making it more complex, which are explored on pages 97–100.

Using handwriting
One cipher cunningly conceals its use of the Polybius square in cursive handwriting. The number of characters between each break where the pen is removed creates a numeric code. Apparently this method was also used in letters home from captured German U-boat officers during World War II. Using the ciphertext given on page 80:
The message – in this case, ‘Many secrets can be hidden in seemingly innocuous messages using this coding trick’ – would be written as follows:
A different approach to substitution ciphers is to use an alternative alphabet consisting wholly or partly of
symbols. This has been done many times in history. The earliest record of it is an ancient Greek practice of using dots for vowels (one for alpha, two for epsilon, and so on), with the consonants left unchanged.

An example of this is the King Charlemagne cipher dating from the 8th century and used in his battle reports.
It was a 23-letter symbolic alphabet (there was no ‘j’, ‘v’ or ‘w’ at the time), which recipients of his messages were required to learn.
The King Charlemagne symbol alphabet code.

DIAL C FOR CIPHER

In the days when telephones had dials, most people were familiar with the three
letters that went with the digits two to nine, creating an alphabet with no ‘q’ or ‘z’. This allowed the creation of ciphers where numbers replaced letters. Adding dots or lines indicated which of the possible trio should be read. For example, as the alphabet began with the number two:
A = 2
B = 2.
C = 2

So the word telephone would be encrypted as 83.53.7466.3.
Pigpen cipher
<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>E</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>H</td>
<td>I</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>J</th>
<th>K</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>N</td>
<td>O</td>
</tr>
<tr>
<td>P</td>
<td>Q</td>
<td>R</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
<th>T</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>W</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The pigpen cipher – see page for an example of it in use.

This is an alternative alphabet that was widely used by masons in the 18th century (it is also known as the Freemason cipher), although some put its origins centuries earlier, during the Crusades. The letters of the alphabet are written inside a grid (which looks like an animal pen,
hence the name) (see diagrams above).

Each letter is represented by the graphic symbol from its part of the grid, with dots added to allow each symbol to be used for two different letters. So ‘This is pigpen cipher’ would look like this:
Pigpen cipher was used by
Confederate forces during the American Civil War, possibly because, apparently, many of the generals were Freemasons and so were familiar with the system. It was solved by a former shop worker who recognized the symbols because the same system had been used to mark the prices of goods in the shop in which he had worked before the war!
A variant of the pigpen cipher (also based on a grid), but with the same shape now representing one of three letters, which are identified by placing a dot on the left, centre or right, is known as the Rosicrucian cipher.
<table>
<thead>
<tr>
<th>ABC</th>
<th>DEF</th>
<th>GHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>JKL</td>
<td>MNO</td>
<td>PQR</td>
</tr>
<tr>
<td>STU</td>
<td>VWX</td>
<td>YZ</td>
</tr>
</tbody>
</table>

- A
- B
- C
- D
- E
- F
The Rosicrucian cipher and an example of it in use.

**DEATH BY SUBSTITUTION**

When Mary, Queen of Scots was being held at Chartley Hall in Staffordshire in 1586, she knew that letters in and out would be studied for indications of any plot against her half-sister Queen...
Elizabeth I. Consequently, when she wanted to correspond with a Catholic sympathizer, she took precautions:

- Letters were smuggled in and out hidden in the bung hole of a beer keg.
- She wrote using 23 letter substitutions and 36 code signs for words and phrases. This blend of code
words and cipher is known as a nomenclator.

• She avoided direct references to plots to kill the queen and put her (Mary) on the throne.

Mary was confident that these precautions prevented her being accused of treason. She did not know that:

• Her correspondent, Gilbert
Clifford, was a double agent.

- Even the brewer who supplied the beer kegs was in the pay of English spymaster Francis Walsingham.

- Her letters were routinely opened and copied for analysis.

- Once Mary’s nomenclator was broken, Walsingham
was browsing through her post even before her servants had sneaked incoming letters past the guards at Chartley Hall.

Such was Walsingham’s grasp of the intricacies of Mary’s code system that he even had extra writing forged onto one of her own letters to try to elicit incriminating information. Although she
denied all at her trial, Mary’s letters showed she knew and approved of a plot against Elizabeth. It was enough for her to be executed.

The queen’s spymaster had set up a cipher school in London, where Mary’s code was broken. The method relied on a revolutionary new weapon in western cryptanalysis: frequency
Frequency analysis is the deadly weapon that breaks substitution and transposition ciphers. It was first developed in the 9th century by an Arab religious scholar called Al-
Kindi, who was studying sacred texts of previous civilizations. He realized that in any language some letters are used far more often than others, that some only appear rarely, and that this pattern remains consistent.

Therefore, in a substitution cipher, the most commonly occurring ciphertext letters are likely to represent the
most common letters in the plaintext language. This allows cryptanalysts to make informed guesses about the identity of individual letters following statistical analysis of both the plaintext language and the ciphertext.

Furthermore, once a common letter is known, its position in a word helps identify its partners. For example, once
‘e’ is identified, any three-letter words with which it ends are very likely to be ‘the’, so identifying ‘t’ and ‘h’.

ELEMENTARY, MY DEAR WATSON
Sir Arthur Conan Doyle’s famous fictional detective Sherlock Holmes applied his deductive powers (usually studying letter frequency) to several codes, including:

- A message concealed as every third word in *The ‘Gloria Scott’*. 
• A book code created by arch foe Moriarty using *Whitaker’s Almanac* in *The Valley of Fear*.

• An unusual cipher using stick men symbols in *Adventure of the Dancing Men*. There are 26 of them, so each represents a letter of the alphabet.
(see the illustration below, which spells out ‘Am here Abe Slaney’).
FROM E TO Z

• In the English language, about one in every eight letters is likely to be an ‘e’. So, if about one-eighth of the letters in a ciphertext are Y, it is likely to be an ‘e’ in plaintext.

• The vowels, a, e, i, o
and u, and the part-time vowel ‘y’, make up about 40 per cent of English text.

• The least common letters are k, j, q, x and z, which, between them, account for just over one per cent of English letters used.
COMMON LETTERS

In order of frequency, the English alphabet reads:

• e, t, a, o, n, r, i, s, h, l, d, c, u, f, p, m, w, y, b, g, v, k, x, q, j, z.

• Counting the letters in the
plaintext, if the six most common letters are the first six listed here, and there are very few of the final letters listed, the cipher is likely to be a giant anagram created by transposition.

- By similar logic, if the most frequently occurring letters in the ciphertext are those that only appear rarely in natural text, the cipher is likely to be of the
substitution variety.

MORSE SENSE

The dots and dashes used in Morse code were decided according to frequency analysis,
with the most commonly used letters requiring the least effort to transmit (see pages).

The next step is to look for short words. Only ‘A’ and ‘I’ are one-letter words, but there are many more of two, three and four letters. If you can
identify word lengths, either because they are not disguised or by identifying the spacer null letter, you can guess what they are most likely to be according to their natural frequency of use and the context (some words are unlikely to follow others, or to start or end sentences). It is worth bearing in mind, however, that code messages do not always follow the rules
of sense and grammar. Ciphertext may be shortened to save effort in translating and transmission, and may include code words that are short for phrases or sensitive information.

However, frequency tables are gold dust for a cryptanalyst. For example, it is much easier to make informed guesses about the
ends of words when armed with the information that more than half of all words end with e, s, t or d. Similarly, if two letters in a transposition ciphertext are the same, they are most likely to be (in order): ss, ee, tt, ff, ll, mm, oo. Two letters in a substitution ciphertext may or may not be the same.

The brain is a powerful tool
and is very good at filling in gaps. For example:

-ou a-e –roba-ly ab-e t-rea- t-is te— e-en tho —- a -ot — it i- m—si-g!

This shows how finding some letters, even in a shifting substitution text, can allow you to identify new
DIGRAPHS

A digraph is two letters that together make a single sound. These are common in English, which helps in cryptanalysis because identifying one letter leads us towards the other. In order of frequency they are:

- th, he, an, in, er, on, re, ed,
Trigraphs are parts of words formed by three letters. Their order of frequency is:

- the, and, tha, ent, ion, tio, for, nde, has, nce, tis, oft, men.
The most common two-letter words in English are:

1 of 1

The most common three-letter words in English are:

1 of 1

The most common four-letter words in English are:

1 of 1

the 1 that
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>to</td>
<td>2</td>
<td>and</td>
<td>2 with</td>
</tr>
<tr>
<td>3</td>
<td>in</td>
<td>3</td>
<td>for</td>
<td>3 have</td>
</tr>
<tr>
<td>4</td>
<td>it</td>
<td>4</td>
<td>are</td>
<td>4 this</td>
</tr>
<tr>
<td>5</td>
<td>is</td>
<td>5</td>
<td>but</td>
<td>5 will</td>
</tr>
<tr>
<td>6</td>
<td>be</td>
<td>6</td>
<td>not</td>
<td>6 your</td>
</tr>
<tr>
<td>7</td>
<td>as</td>
<td>7</td>
<td>had</td>
<td>7 from</td>
</tr>
<tr>
<td>8</td>
<td>at</td>
<td>8</td>
<td>her</td>
<td>8 they</td>
</tr>
<tr>
<td>9</td>
<td>so</td>
<td>9</td>
<td>was</td>
<td>9 know</td>
</tr>
<tr>
<td>10</td>
<td>we</td>
<td>10</td>
<td>one</td>
<td>10 want</td>
</tr>
<tr>
<td>11</td>
<td>he</td>
<td>11</td>
<td>our</td>
<td>11 been</td>
</tr>
<tr>
<td>12</td>
<td>by</td>
<td>12</td>
<td>out</td>
<td>12 good</td>
</tr>
<tr>
<td>13</td>
<td>or</td>
<td>13</td>
<td>you</td>
<td>13 much</td>
</tr>
<tr>
<td>14</td>
<td>on</td>
<td>14</td>
<td>all</td>
<td>14 some</td>
</tr>
<tr>
<td>15</td>
<td>do</td>
<td>15</td>
<td>any</td>
<td>15 time</td>
</tr>
<tr>
<td>16</td>
<td>if</td>
<td>16</td>
<td>can</td>
<td>16 very</td>
</tr>
<tr>
<td>17</td>
<td>me</td>
<td>17</td>
<td>day</td>
<td>17</td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>18</td>
<td>my</td>
<td>18</td>
<td>get</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>up</td>
<td>19</td>
<td>has</td>
<td>19</td>
</tr>
<tr>
<td>20</td>
<td>an</td>
<td>20</td>
<td>him</td>
<td>20</td>
</tr>
</tbody>
</table>

**LETTER KNOWLEDGE**

Word breaks are valuable to a cryptanalyst because they can gain clues using the following information:
• Most common first letter in a word, in order:

t, o, a, w, b, c, d, s, f, m, r, h, i, y, e, g, l, n, o, u, j, k.

• Most common third letter in a word, in order:
e, s, a, r, n, i.

• Most common last letter in a word, in order:
e, s, t, d, n, r, y, f, l, o, g, h, a,
k, m, p, u, w.

- Letters most likely to follow the letter ‘e’:
  r, s, n, d.

**WORD FREQUENCY**

Zipf’s law identifies the most common words and in what proportion of text they will appear. Named after Harvard linguist George Kingsley
Zipf, it shows that seven percent of all words are ‘the’, followed by ‘of’ at about half that frequency, then ‘and’. Cryptanalysts can apply this law to make informed guesses about words in context.
Just outside the Central Intelligence Agency’s headquarters in Langley, Virginia, is ‘Kryptos’, a 12-foot-high copper, granite and petrified wood sculpture that has baffled staff and other cryptologists since its
installation in 1990. Sculptor James Sanborn, an ex-CIA worker, inscribed it with some 1,800 letters forming four messages, each in a different cipher.

Three of the ciphers have been broken using frequency analysis and a
Vigènere square (see pages). They reveal a set of coordinates, possibly of a nearby location where Sanborn has buried something. The fourth passage retains its secrets.

After ‘Kryptos’, Sanborn created other coded sculptures,
including ‘The Cyrillic Projector’, a cylindrical installation at the University of North Carolina, which contains text from classified Russian KGB documents in Cyrillic alphabet.
POLYALPHABETIC CIPHERS

Beating frequency
The spread in the use of frequency analysis stimulated cryptographers to create ciphers or codes that defied it. There was enormous pressure to do this as nations continued to fight, trade and negotiate with each other, and there was a clear need for secure communication within governments and armies.
In a null cipher, a message is hidden inside the plaintext (see page). However, another way to use nulls is to adapt the Greek square system outlined on pages 81–2. This allocates each of 25 letters (with two doubling up) a two-digit number, which is actually a grid reference.
Making the grid bigger creates more boxes to put letters in. This enables the encoder to have the same letter represented by more
than one number, which protects it from frequency analysis.

POLYALPHABETIC CIPHERS

Polyalphabetic ciphers use more than one alphabet. ‘Poly’ is
Greek for ‘many’.

For example, increasing the grid from five-by-five to six-by-six creates an extra 11 boxes. These can be randomly filled with the five most common letters used in the English language: e, t, a, o and n (each added twice), with the final gap filled by
another ‘e’ as it is by far the highest frequency letter at 12.7 per cent of all text. Now there are four numbers for ‘e’: 11, 12, 21 and 26, and three numbers each for t, a, o and n. Someone who intercepts the message and counts repeated numbers looking for ‘e’ will be totally misled.
The message ‘Even trees hide secrets’ has seven ‘e’s, which stand out very clearly as the number 15 in this message enciphered from the five-by-five grid:
A decoder would immediately start to decipher the message as:

Four-letter words with E as the first and third letters are:

eden, eked, ekes, epee, erev, even, ever, eves, ewer, ewes, exec, exes, eyed, eyer, eyes

which gives the code breaker some idea of numbers for other letters, and the beginnings of a context for
the whole message.

The same message encrypted from the six-by-six grid reads as follows:

```
```

Here four numbers appear twice: 11, 12, 21 and 53, and
54 appears three times. This is helpful to a code breaker, but it does not give anywhere near as much information as the version encrypted on the five-by-five grid.
The larger the grid, the more possibilities for tricking the enemy there are, such as:

- Enciphering the same letter twice or more.
- Using some numbers for code words or phrases (turning the cipher into a nomenclator).
• Using some numbers as space markers so that the ciphertext does not reveal the correct word lengths.

• Perhaps most useful of all, creating nulls with no meaning whatsoever that the decipherer can safely ignore, but which will force
your opponent into hours of fruitless investigation.

THE GREAT PARIS CIPHER

Numbers were the basis of the great Paris cipher used by Napoleon’s French Army
during the Peninsular Wars (1808–14). The French first used the Army of Portugal code, which grew from 50 to 150 numbers, each giving a short instruction.

Wellington’s self-trained cryptanalyst George Scovell cracked this, but was initially confounded by a new cipher introduced in 1811. This eventually used 1,400
numbers, including more than 130 for ‘e’ alone to combat frequency analysis. Other tricks used were:

• Adding meaningless figures to the ends of words (cryptanalysts study beginnings and ends of words first, looking for patterns).

• Disguising standard phrases by adding meaningless
A VOID

One text that would perplex a frequency analyst is *La Disparition* by Georges Perec, translated from French numbers (see cribs, pages 131–3).
to English as *A Void* by Gilbert Adair. There is not a single letter ‘e’ in the entire 300-page novel in either version.

However, for ease of use and to save space, the French only enciphered key words, leaving the rest as plaintext.
(also known as *en clair*). So a typical message would read (in English): ‘In the letter of 16 March 1207 announced 607.73.432.1181.192.1077.600.530.497.701.711.700 that he considers appropriate.’

This allowed Scovell, a fine linguist, to discern the context of the enciphered words and phrases, so that he could make more informed guesses.
and follow up these hunches. Additionally, as each number represented one letter, Scovell could discount any guesses with the wrong number of letters. It remained a formidable task, but after about a year he had cracked most of the cipher and the information he passed on to Wellington, when combined with other intelligence, was crucial in the final victory.
over the French at Vitoria in 1813 (although Wellington subsequently barely acknowledged it).

FROM EAST TO WEST

The man credited with bringing Arab
knowledge of frequency analysis to the West was Leon Battista Alberti (1404–72), the greatest cryptographer of his day and the first man to suggest superencipherment (see page). Experienced cryptographers equipped with
information on frequency data can solve codes in languages they cannot even read.

ONE AND TWO MAKE TWELVE
When using numbers for a ciphertext that might be merged or re-ordered into false word blocks, it is important to prefix single digit numbers with zero (so one becomes 01) to avoid confusion in decoding – otherwise 1 followed by 2 would read as 12.
Expanding the cipher alphabet to conceal the frequency of letters is particularly effective if the ciphertext is written as numbers. If the alphabet is written out but the six most frequent letters (a, e, n, o, r, t) doubled-up and each character given a number, it
would look like this:

![Image showing a grid with letters and numbers]

You could repeat any letter as many times as you like, creating more ‘e’s, for example, and the numbers can be adjusted so that
consecutive numbers do not describe the same letter – which would be a major giveaway to the cryptanalyst.

An expanded alphabet that accurately reflects the letter frequency pattern, with numbers generated at random, would look like the table below:
<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
<th>j</th>
<th>k</th>
<th>l</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>02</td>
<td>78</td>
<td>89</td>
<td>99</td>
<td>12</td>
<td>68</td>
<td>53</td>
<td>11</td>
<td>77</td>
<td>01</td>
<td>79</td>
<td>18</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>08</td>
<td>91</td>
<td>00</td>
<td>21</td>
<td>80</td>
<td>33</td>
<td>22</td>
<td>85</td>
<td>61</td>
<td>29</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>97</td>
<td>94</td>
<td>09</td>
<td>28</td>
<td>90</td>
<td>26</td>
<td></td>
<td></td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>72</td>
<td>98</td>
<td>52</td>
<td></td>
<td>18</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td>60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>62</td>
<td>87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>51</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>92</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>13</td>
<td>24</td>
<td>35</td>
<td>38</td>
<td>41</td>
<td>44</td>
<td>37</td>
<td>46</td>
<td>27</td>
<td>58</td>
<td>40</td>
<td>36</td>
<td>76</td>
</tr>
<tr>
<td>o</td>
<td>45</td>
<td>06</td>
<td>49</td>
<td></td>
<td>03</td>
<td>57</td>
<td>66</td>
<td>55</td>
<td></td>
<td>05</td>
<td></td>
<td></td>
<td>86</td>
</tr>
<tr>
<td>p</td>
<td>16</td>
<td>23</td>
<td>64</td>
<td>70</td>
<td>42</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q</td>
<td>74</td>
<td>88</td>
<td>39</td>
<td>19</td>
<td>43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>63</td>
<td>65</td>
<td>73</td>
<td>95</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>20</td>
<td>69</td>
<td>04</td>
<td>71</td>
<td>93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td></td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
One hundred numbers have been used in this table, so the number of code numbers given to each letter matches its percentage of usage in everyday language. For example, as more than 12 per cent of all letters are ‘e’s, so it has 12 numbers, while j, k, q, v, x and z all score under 1 per cent so only get one letter. This method completely
nullifies attack by frequency analysis, but there are other ways to attack such ciphers.

A GLIMMER OF HOPE FOR DECODERS

When trying to break this cipher, the decoder knows that each number can only represent one letter. Once identified, that number will always reveal the presence of
the same letter. This can provide clues. For example, if the decipherer was fortunate enough to identify the cipher for ‘q’, this would suggest that the next letter must be ‘u’. Similarly, a three-letter word beginning with ‘a’ is quite likely to be ‘and’. Common two-, three- and four-letter words are listed on page 94. Other spelling-related facts include:
• The high-frequency vowels a, i and o are rarely seen together.

• The most common vowel digraph is ‘ea’.

• The most common repeated letters are:
  ss, ee, tt, ff, ll, mm and oo.

• The pairing ‘th’ is common, ‘ht’ is not.
• Indeed, ‘th’ is the most common letter pair, followed by: he, an, in, er, on and re (see also page 93).

• Words are most likely to start with (top five in order):
  
  t, o, a, w, b.

• Words are most likely to end with (top five in order):
So the decoder may feel that the rules of spelling and grammar are on their side.

The answer to this weakness is not to follow the rules of spelling and grammar. It’s a fair bet that you can read this heading, even though the
vowels and some consonants have been removed, without leaving a gap. The human brain is remarkably good at reading text like this (just look at phone texting codes, see pages), using a mixture of word recognition and context to make sense of it. Indeed, we are able to read words with jumbled letters quite easily, pivoredd the fsrit and lsat ltdtrees are in pclae. So a
plaintext that breaks spelling rules will be comprehensible to the rightful decoder, but will offer fewer clues to the interceptor.

Using more than one cipher alphabet

The breakthrough against
frequency analysis was first suggested by Leon Alberti, an archetypal 15th-century Italian Renaissance man: painter, poet, linguist, philosopher, musician, architect . . . . and cryptographer.

USING TWO CIPHERS
Alberti realized that frequency analysis only works against a monoalphabetic cipher, so his suggestion was to use two ciphers. It would work like this:

**Plain alphabet:**  
abcdefghijklmnopqrstuvwxyz  

**1st cipher alphabet:**  
H I J K L M N O P Q R S T U V W X Y Z A B C D E F G  

**2nd cipher alphabet:**  
Q R S T U V W X Y Z A B C D E F G H I J K L M N O P
Letters of the plaintext message would be enciphered by using the two different ciphers alternately, so, for example, from the table above an alphabet ciphertext would read:

```
H R J T L V N X P Z R
B T D V F X H Z J B L
D N F P
```
He refined this idea with the cipher disc, a pair of copper wheels each inscribed with the alphabet. When one disc is turned, it creates a simple substitution cipher. Alberti’s breakthrough was to suggest that with every few words the wheel would be turned, changing the cipher alphabet. This would be signalled to the decoder with a capital letter in the ciphertext, indicating a
new position for the wheel. Thus the encrypter has a choice of 26 alphabets and can switch between them, making unwanted deciphering very difficult. Alberti’s thinking was centuries ahead of its time and his invention was not widely adopted.
One of the many coding systems used in the American Revolutionary War employed a keyword to create several alphabets that were used concurrently. It was created by James Lovell. The word ‘key’ provides the key here (it could, of course, be a longer word – names were popular, because they are so memorable) and from each letter of the word the alphabet
is continued. The sender then simply works across the columns to encrypt the message. So the plaintext ‘Lovell’ would become the ciphertext 2, 11, 24, 21, 8, 14.
The Vigenère square takes multi-alphabets one step further by setting out the alphabet 26 times, each moved by successive shifts of one place, as shown below. This means that ‘t’, the second most common letter, at nine per cent frequency, is ciphered as W after three
shifts (like the Ceasar alphabet), D on 10 and O on 20. The question for the encoder is how to show which row was used for each letter. One answer is the date shift cipher.

**THE VIGENÈRE SQUARE**
<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
<th>j</th>
<th>k</th>
<th>l</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
</tr>
<tr>
<td>4</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
</tr>
<tr>
<td>6</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
</tr>
<tr>
<td>7</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
</tr>
<tr>
<td>8</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
</tr>
<tr>
<td>9</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
</tr>
<tr>
<td>10</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
</tr>
<tr>
<td>11</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
</tr>
<tr>
<td>13</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
</tr>
<tr>
<td>14</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
</tr>
<tr>
<td>15</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>16</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>17</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>18</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>19</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
</tr>
<tr>
<td>20</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
</tr>
<tr>
<td>21</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
</tr>
<tr>
<td>22</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
</tr>
<tr>
<td>23</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
</tr>
<tr>
<td>24</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
</tr>
<tr>
<td>25</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
</tr>
<tr>
<td>26</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>o</td>
<td>p</td>
<td>q</td>
<td>r</td>
<td>s</td>
<td>t</td>
<td>u</td>
<td>v</td>
<td>w</td>
<td>x</td>
<td>y</td>
<td>z</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>3</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>4</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>5</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>6</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
</tr>
<tr>
<td>7</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
</tr>
<tr>
<td>8</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
</tr>
<tr>
<td>9</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
</tr>
<tr>
<td>10</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
</tr>
<tr>
<td>11</td>
<td>Y</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
</tr>
<tr>
<td>12</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
</tr>
<tr>
<td>13</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>14</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
</tr>
<tr>
<td>15</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
</tr>
<tr>
<td>16</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
</tr>
<tr>
<td>17</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
</tr>
<tr>
<td>18</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
</tr>
<tr>
<td>19</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
</tr>
<tr>
<td>20</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
</tr>
<tr>
<td>21</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
</tr>
<tr>
<td>22</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
</tr>
<tr>
<td>23</td>
<td>K</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
</tr>
<tr>
<td>24</td>
<td>L</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
</tr>
<tr>
<td>25</td>
<td>M</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
</tr>
<tr>
<td>26</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
</tr>
</tbody>
</table>
DATE SHIFT CIPHERS

In the date shift cipher, the encoder might use the date of the message, a birthday or a famous date in history to determine how a message is to be enciphered. For example, take the date of the Gunpowder Plot – 5 November 1605. Numerically, this is
represented as 5/11/1605, producing a code number that is written repeatedly above the plaintext. Each letter is shifted by the number of places indicated by the number above it.

Key number: 511 1605 51 116 055111605 5111
Plaintext: The date of the gunpowder plot
Cipher text: YIF EGTJ TG U1K GZSQPXJEW UMPU
Deciphering polyalphabetic ciphers using word breaks

The ciphertext of this message shows the number of words and their length. In this case the gaps reveal that there are six words: one of two letters, two of three letters, two of four letters, and one of nine. There are a limited
number of words of two or three letters in any language, and sometimes these can be deduced from their context. For example, the three-letter word that begins the message is very likely to be ‘The’, as it so often starts sentences. The most common two-letter word is ‘of’, and here the two-letter word is followed by a three-letter word. Guessing that we have
identified ‘of’, and studying the list of most common three-letter words (see page) in context, removing those that would not make sense, there are only five sensible possibilities: the, her, one, our, all. So by informed guesswork and without even investigating the letters themselves, it already now reads:
HIDING WORD BREAKS

The standard way to conceal word breaks is to write the encrypted message in blocks of five letters. This practice seems to have been introduced by the telegraph operators of the 19th century.
as it made it easier to transmit nonsensical messages (code was used a lot in telegrams to cut costs).

Another way the encrypter can avoid the problem is by merging all the letters so that there are no spaces. However, this could, on occasion, cause confusion in the deciphered message, as there may be more than one interpretation
of where new words begin. Instead, they could use one letter as a space indicator. Using Z in this way, the message reads:

YIFZEGTJZTGZUIKZG

This disguises the word breaks (although the frequent occurrence of the Z gives a clue as to its role). It also
requires the decipherer to make informed guesses as to when $Z$ is working as a cipher letter and when it is a space indicator.

MORE WAYS TO USE THE VIGENÈRE SQUARE

Another way to use the
Vigenère square is to use a keyword such as, well, ‘keyword’. Taking the number of each letter’s place in the alphabet, this would give following values:

<table>
<thead>
<tr>
<th>K</th>
<th>E</th>
<th>Y</th>
<th>W</th>
<th>O</th>
<th>R</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>5</td>
<td>25</td>
<td>23</td>
<td>15</td>
<td>18</td>
<td>4</td>
</tr>
</tbody>
</table>

If this is written repeatedly over the plaintext it means the
ciphering system is now on a cycle of seven different shifts. So the message, ‘This would be very hard to decipher’ is processed like this:

<table>
<thead>
<tr>
<th>Shift</th>
<th>11 5 25 23 15 18 4 11 5 25 23 15 18 4 11 5 25 23 15 18 4 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plaintext</td>
<td>this would be very hard to decipher</td>
</tr>
<tr>
<td>Ciphertext</td>
<td>EMHP LGYWIA BKWVJMZOSLSOJBFEZIC</td>
</tr>
</tbody>
</table>
In this example, the ‘e’ is enciphered as B, W, J and I. For the decipherer equipped with the correct keyword this is no problem: they simply shift each letter back by the indicated number of places.

But for anyone trying to break the cipher without knowledge of the key (which could be a date or keyword), they will not be able to
deduce an alphabetical pattern and frequency analysis will yield no clues. The cipher’s main advantage is that the same letter or numbers can represent more than one letter in the plaintext alphabet – a decoder’s nightmare. The keyword system is the mechanism by which the Vigenère cipher is most used and it remains a popular code among
cryptology enthusiasts.

**HOW DO YOU BREAK VIGENÈRE?**

The simple answer, of course, is to find the keyword.

- If subterfuge does not provide this, sometimes guesswork can (there is a story from World War I of cryptologists trying to find a keyword used in a cipher
created with a mechanical device correctly guessing that the word was ‘machine’).

• However, if that fails, the next option is to look for patterns in sequences that will eventually reveal how many different shifts are evident, and from this, the length of the keyword. For example, common words like ‘the’ or ‘and’ and
common letter strings like ‘ing’ or ‘ted’ will appear many times in a message, so may have been encrypted the same way at some stage in a long communication. Identifying how and when this occurs allows the cryptanalyst to identify how long the keyword is.

• Once you know that the keyword is, say, eight
letters long, you can use frequency analysis for every eighth letter, knowing all are likely to come from the same row of the Vigenère square.

- From this you can begin to identify letters of the keyword itself.

So breaking the Vigenère cipher is time-consuming, but possible, and is easier if
messages are long, providing more ciphertext to study, or keywords are short, reducing the scrambling of letters and making patterns more discernible.

THE UNCRACKABLE CIPHER

The keyword is the weak spot
of a transposition cipher because once its length has been guessed, it can be decrypted using frequency analysis. The answer is to have a key the same length as the plaintext, known as the running key because it runs on rather than repeats. The running key can be any text of sufficient length and all that needs to be agreed between the communicating
parties is the starting point. A rather neat development of this is the use of an ‘autokey’. Here, the start of the message is enciphered with a keyword, after which the revealed plaintext forms the running key to the remainder of the message.

**BREAKING THE RUNNING KEY**
The disadvantage of using a readable text as the key is that it will follow the conventions of grammar, and therefore of word and letter frequency. Since the most common three-letter word is ‘the’, the ciphertext can be attacked by assuming that the key is repeated use of ‘the’, which will eventually yield clues to the message.
THE ONE-TIME PAD

The solution to this weakness is the one-time pad: a set of randomly produced letters that reads like total gibberish. The system was devised by Gilbert S. Vernam and Joseph Mauborgne in 1918 and is also known as the Vernam cipher. Random letters are printed onto a pad, used to encipher the message, then
the sheet is peeled away and destroyed. This creates a genuinely unbreakable cipher, because every letter has been generated at random so there will be no pattern to the key. The drawback is that the decrypter needs a copy of the random key, so, as with codes, anyone on the communication network has a copy of it. Therefore, someone needs to create a set
of ‘one-time pads’ (and it is surprisingly difficult for humans to create random letters), print and distribute them and keep them safe – a logistical problem that prevents this uncrackable Vigenère cipher from being used on more than rare occasions.
Checkerboard ciphers is a collective term for ciphers produced using the principle of the Polybius or Greek square (see pages), in which letters are set out on a grid for encryption.
The Playfair cipher is a highly efficient method for digraphic substitution. It requires a five-by-five square containing a keyword followed by the alphabet with one letter omitted (usually ‘j’ but sometimes ‘q’). So using the keyword ‘playfair cipher’, which with repeated letters left out becomes PLAYFIRCHE, the square would look like this:
<table>
<thead>
<tr>
<th>P</th>
<th>L</th>
<th>A</th>
<th>Y</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>R</td>
<td>C</td>
<td>H</td>
<td>E</td>
</tr>
<tr>
<td>B</td>
<td>D</td>
<td>G</td>
<td>K</td>
<td>M</td>
</tr>
<tr>
<td>N</td>
<td>O</td>
<td>Q</td>
<td>S</td>
<td>T</td>
</tr>
<tr>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Z</td>
</tr>
</tbody>
</table>

Messages are enciphered in pairs of letters following these rules:

- For letters not in the same row or column, the first is replaced by the letter in the
same row and column as the second. The other is substituted by the letter on the same row and column as the first. The four letters thus form a rectangle, which makes the encoding process fast and easy. So in the square shown, ‘ok’ would be SD.

- Letters on the same row are replaced by the letter to their right, with the row
‘wrapping round’ to the start. So ‘pl’ would become LA.

• Letters in the same column are substituted by the letter that lies immediately below. So ‘lr’ would become RD.

• If a pair is formed by the same letter (i.e. ‘ee’) the letters are separated by a dummy ‘null’ letter, such as
X or Z, or, in some variants, the second one is changed to X or Z.

- A single letter at the end of the message is made into a pair using a null such as X or Z.

For example, the surname of the inventor of this cipher, Charles Wheatstone, would become first WH EA TS TO NE, then XC CF NT NQ TI.
Breaking the Playfair Cipher

Messages created with the Playfair cipher can be broken by a form of frequency analysis, since certain pairs of letters (digraphs) occur far more often than others (for example, ‘qu’ – see more on this on pages 93–5), and many others never appear consecutively. However, this
method takes far longer than analyzing single-letter ciphers as there are an impressive 676 possible letter combinations to count.

THE LOVE CODE

In Victorian times,
when courtship was regulated by social etiquette, lovers would communicate through coded messages in newspaper personal columns, especially the *Times*. These were regarded as titillating puzzles to be solved by many readers including
cryptologists Charles Babbage (who is credited with creating the first computer), Sir Charles Wheatstone (who devised the Playfair cipher) and Baron Lyon Playfair (who only promoted it). They even placed mischievous replies of their own, using the
same ciphers, causing enormous confusion.

THE BIFID CIPHER

The Playfair cipher is easy to use and can be adapted into number form, allowing a fiendish twist in the form of the bifid cipher.
As an example, the name of this cipher’s 1901 creator, Felix Delastelle, is enciphered using the same alphabet grid as that on the last page (although it can just
be a mixed alphabet). Each letter is given a grid reference, but these are written vertically, rather than horizontally.

```
 felix delastelle
1 2 1 2 5 3 2 1 1 4 4 2 1 1 2
5 5 2 1 4 2 5 2 3 4 5 5 2 2 5
```

The numbers are then read going across the rows, rather
than down, and put into standard five-digit groups:

12125  32114  42112
55214  25234  55225

This process is called fractionation and in itself creates a tricky cipher, but to make life for the unwelcome decoder even harder, the numbers are then used as grid
references and the cipher returned to text:

12 12 53 21 14 42 11 25 52 14 25

L L W I Y O P E V Y E

So the ciphertext is finally broken up into fixed lengths (called periods), resulting in:

LLWIIY    OPEVY
The bifid cipher works, like all pencil and paper methods, in two dimensions, but it is the father of an unusual ‘three-dimensional’ system called the trifid cipher. Imagine three layers of three-
by-three Polybius squares on top of each other, which form a cube with a mixed alphabet.

<table>
<thead>
<tr>
<th>Layer one</th>
<th>Layer two</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3</td>
<td>1 2 3</td>
</tr>
<tr>
<td>1 M J S</td>
<td>1 B T K</td>
</tr>
<tr>
<td>2 Z Q H</td>
<td></td>
</tr>
<tr>
<td>3 F O X</td>
<td></td>
</tr>
</tbody>
</table>
Now each letter can be given a three-digit number from its layer, row and column. For example:

<table>
<thead>
<tr>
<th>Layer</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>L</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>P</td>
<td>C</td>
</tr>
<tr>
<td>3</td>
<td>R</td>
<td>U</td>
<td>I</td>
</tr>
</tbody>
</table>
The next stage is:

```
trifidisystem
2 3 3 1 3 2 1 2 1 2 3 1
1 3 3 3 3 2 1 2 1 1 2 1
2 1 3 1 3 1 3 2 3 2 1 1
```
This creates:

\[233\ 132\ 121\ 231\ 133\ 332\ 121\ 121\ .\ \text{OZWXUZZKFCB}\]

So the message sent is:

\[.\text{OZWXUZZKFCB}\]

which is decoded by reversing the whole process.
MYSTERIOUS MONUMENT

In the grounds of Shugborough Hall, a stately home in Staffordshire, is the so-called Shepherd’s Monument, which bears some inscriptions that have
baffled people for 250 years. Dating from the 1760s, the marble slab features the letters O, U, O, S, V, A, V and V with a D and an M below, etched under a mirror image of a painting by Nicholas Poussin. Some suggest the painting indicates the letters are supposed to be read in
reverse.

Solutions offered include:

• It is a set of instructions on how to find the Holy Grail.

• The markings are linked to the Masons and African nature worship.

• It is an ancient love
note, and the code is a romantic Latin phrase.

THE STRADDLING CHECKERBOARD

This is a code that was used by Russian spies in the early part of the 20th century,
although it was first employed 400 years previously in papal ciphers.

<table>
<thead>
<tr>
<th>0</th>
<th>9</th>
<th>8</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>T</td>
<td>A</td>
<td>O</td>
<td>N</td>
<td>I</td>
<td>S</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>J</td>
<td>K</td>
<td>L</td>
</tr>
<tr>
<td>2</td>
<td>P</td>
<td>Q</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td></td>
</tr>
</tbody>
</table>

Eight letters (in this case the most common ones, but it could be a keyword or a
random choice) are given single digit values (for example ‘e’ is 0). The rest are represented by two digits as we have seen before (so ‘y’ is 24). The digits 2 and 1 cannot be used individually as this would confuse the decoder as they are the vertical axis numbers as well as being in the horizontal axis. The columns and rows can be numbered differently in any
agreed way in other messages. The ‘/’ symbol indicates a shift from words to numbers or back.

So the (quite true) plaintext, ‘Used in 1937 by Spanish Communists’, would be enciphered as:
In five-digit blocks (plus a final block of two), with each plaintext number digit repeated in case of transmission errors, it now appears as:
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>28401</td>
<td>85621</td>
<td>11993</td>
<td></td>
</tr>
<tr>
<td>37721</td>
<td>10244</td>
<td>20865</td>
<td></td>
</tr>
<tr>
<td>41519</td>
<td>71111</td>
<td>28654</td>
<td>94</td>
</tr>
</tbody>
</table>

The message could be sent as it is, or transposed back into letters using the bifid cipher method.
Field codes and ciphers

The military has long been
the force that pushes for improved cryptography, and phrases such as trench code and field cipher indicate how codes were used in action, requiring simple, fast systems that relied on memory rather than a manual.

CODE OR CIPHER?

One of the first actions the British took in World War I
was to cut the cables laid across the North Sea by the Germans. This forced their enemy to send signals on cables controlled by the Allies, or by radio, both of which allowed them to be intercepted. The move reveals the importance of encrypting and decrypting communication in wartime. The debate about whether to opt for codes or ciphers for
secret communication was particularly heated during World War I. Each had its merits and drawbacks (see overleaf).

CODES

Advantages

- High level of confidentiality if code books are secure.
• Allows use of a wide vocabulary.

• Long phrases used regularly can be communicated concisely with one word or number.

Disadvantages

• Large code books are difficult to transport safely and securely.

• Can be attacked bit by bit
as each code word reveals its secret.

- If your enemy captures a code book without your knowledge, they can read all your messages as fast as you.

- Capture of the code book requires the creation of a new one.

**Ciphers**
**Advantages**

- No bulky code book to transport.
- Can be adjusted regularly, for example by changing keywords, to maintain security.

**Disadvantages**

- Can take longer to decipher, slowing down communication.
• Transmission errors possible, very hard to spot and cause confusion in deciphering.

• Messages sometimes have to be sent twice – gold dust to opposing cryptanalysts.

• Operators of the system get lazy and take short cuts, which offers clues to the enemy.
FIELD CIPHERS

Field ciphers are ciphers that require very little equipment, such as conversion tables or apparatus, and are reliant on an easily remembered and changed keyword. They are ideally suited to communication by an army on a fast-moving battlefield. The ideal field cipher is one that is easily operated by one
person who carries the system mostly in their head.

The most famous field cipher was used by German forces during World War I, and is called the ABFGX cipher, a cunning blend of substitution and transposition that created messages using only five letters. The method required a mixed alphabet grid and a keyword, both of which were
changed daily.

First, a mixed alphabet was created, with ‘i’ and ‘j’ combined:

```
A D F G X
A k r z v l
D f t s x i/j
F g m o h w
G b e q p c
X d a y n u
```
The plaintext, ‘Just five letters’, was fractionated:

```
justfiveletters
dx xx df dd da dx ag gd ax gd dd dd gd ad df
```

Then the resultant monoalphabetic substitution ciphertext was written in columns under the keyword ‘ZEBRA’:
<table>
<thead>
<tr>
<th>Z</th>
<th>E</th>
<th>B</th>
<th>R</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>D</td>
</tr>
<tr>
<td>F</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>A</td>
</tr>
<tr>
<td>D</td>
<td>X</td>
<td>A</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>D</td>
<td>A</td>
<td>X</td>
<td>G</td>
<td>D</td>
</tr>
<tr>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>G</td>
</tr>
<tr>
<td>D</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>F</td>
</tr>
</tbody>
</table>

Finally, the transposed message was read in columns, following the alphabetical order of the letters in the keyword (A = 1, B = 2, E = 3, ...)
R=4, \( Z = 5 \): DAGDGF XDAXDD XDXADA XDGGD DDDDF. It would be transmitted in batches of five letters:

<table>
<thead>
<tr>
<th>DAGDG</th>
<th>FXDAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDXDX</td>
<td>ADAXD</td>
</tr>
<tr>
<td>GGDDD</td>
<td>FDDDD</td>
</tr>
</tbody>
</table>
WHY FIVE?

The letters chosen for the ABFGX cipher were the ones least like each other in Morse code, making mistakes less likely in transmission and decrypting.
This method evolved into the ADFGVX cipher, which created a 36-strong alphabet, allowing the introduction of numbers. A major benefit for its users was that enciphering only required two simple steps, and only six (and earlier, five) letters were used in transmission, so it could be sent rapidly. The Allied forces made huge efforts to solve it and it was broken by
French army lieutenant Georges Painvin. His solution began with finding two messages that started or finished in the same way. This is called a ‘crib’. From this, the Allies were able to decipher first the ciphertext as it would be prior to transposition and then use frequency analysis on the ciphertext. On quiet, low traffic days, they were less
likely to find a pair of messages with identical beginnings or endings.

CRIBS

We are all creatures of habit, and so are many organizations with their agreed practices and procedures. As a result, even encrypted communication can follow patterns and become
predictable. Once this occurs, the code breaker has a crib: a word or a piece of text that they already know is repeated somewhere in an encrypted message. This is extremely valuable as the code breaker can then search for patterns that relate to it, rather than just fishing for clues. Genuine examples of cribs from one or other of the world wars include:
• Stereotyped messages such as ‘Nothing to report’.

• Weather reports that were sent at the same time each day and in the same format.

• Messages beginning, ‘It is my honour to inform your Excellency’.

• Sending well-known proverbs as test messages.

• Using patriotic words such
as ‘Kaiser’ or ‘Deutschland’ as keywords.

GARDENING FOR CRIBS

Such is the value of cribs that code breakers took actions to try to prompt messages with certain words that then became a crib, a practice known as ‘gardening’. For example:

• In World War I, the Allies
bombarded certain trenches knowing it would make the Germans send messages about those parts of the battlefield.

• In World War II, the Americans sent an uncoded message that a ship needed fresh water, knowing the Japanese would intercept and report it, thus revealing their code name for the ship.
MAKING THE MOST OF MISTAKES

Lazy operators not changing their keywords or their style of transmission is another habit that leaves the doors open for code breakers. Even better, operators who make mistakes are invaluable because they are often asked to repeat the message. If they do this with a new cipher or
keyword, the code breaker can create a crib by decrypting the first transmission.

The machine age

After centuries of using pencil and paper methods for encryption and decryption, the impact of cryptography on World War I encouraged
governments and the military to create stronger codes and ciphers. Inventions such as the telegraph, calculator, typewriter and cash register highlighted the new role of technology, and the push was on to invent automated cipher machines.

THE JEFFERSON WHEEL
In fact, the first encrypting machine had been invented, and largely forgotten, in the 1790s, by Thomas Jefferson, an incredibly gifted man who was to become the third President of the United States. The machine is a simple device: a set of between 20 and 40 discs, or wheels, each with the alphabet written on them in random order. The discs are
mounted in a row to form a cylinder. Jefferson was aware that the more wheels there were, the better, and suggested having 36, which gave a huge number of possible settings (nearly 372 followed by 39 zeros!, calculated from $1 \times 2 \times 3 \times \ldots \times 35 \times 36$).

To operate the Jefferson wheel, you would create your
message in plaintext along one row, reading across the discs, then copy out any other row of letters, which would be gibberish. The recipient simply re-creates the line of nonsense, then scans the other rows until he finds a readable message. The cipher could be changed by removing the discs and replacing them in a different agreed order. It was a brilliant invention, years
ahead of its time, yet it was quickly forgotten and was reinvented over a century later.

**RE-INVENTING THE WHEEL**

From 1922 to 1943 the US Army sent ciphered messages using a cylindrical device made up of lettered aluminium wheels, and
known as M-94. It was almost the same as the Jefferson wheel, and had been developed by two officers in 1917. They had been inspired in turn by the Bazières cylinder, an 1891 reinvention of Jefferson’s cipher device.

ENIGMA

The breaking of the Enigma cipher was one of the great
achievements of World War II, and in the history of cryptography. Some historians believe that it shortened the war by a year and saved millions of lives.

THE ENIGMA MACHINE

The Enigma machine was by far the most complicated enciphering system created at the time. It looked like a
typewriter with parts stuck on the sides, which had been put in a wooden box. There were many different versions, commercial and military, and its workings were improved several times during the war.

In essence, Enigma was a set of rotating wheels wired in several ways so that a message whose letters had already been substituted was
repeatedly scrambled to create a monumental number of variations that defied any known analysis because it left no glimmer of a pattern. Depending on the version of the machine used, Enigma made between nine and eleven changes to every letter of the plaintext on its route to the ciphertext.

Its main letter-scrambling
device was a set of three electromagnetic wheels called rotors, each marked with an alphabet, similar to the concept of the Jefferson wheel (see pages 133–4). With single-notched rotors, the ‘period’ of the machine (the possible number of shifts before returning to the original setting) was 17,576 (26 x 26 x 26), helping to protect against overlapping
alphabets. This, in itself, is a relatively simple substitution cipher, vulnerable to frequency analysis. However:

- Each time a key was struck, the first rotor changed position, altering the substituting pattern so that each letter was encrypted with a new encrypting alphabet.

- The rotors were connected
so that when the first rotor had shifted 26 times, the next rotor then advanced one place. With three rotors this meant that the machine only returned to its original setting after 676 letters had been typed in.

- The rotors could be re-ordered (so instead of 123 they could be in 132, 213, 231, 312 or 321 order). This allowed six more
permutations, vastly increasing the number of alphabet combinations \((6 \times 17,576 = 105,456)\).

- Further internal wiring linked the rotors to a plugboard, which exchanged up to six letters, for example swapping around D and F, to make encryption even more complex.
The third rotor was connected to a non-rotating rubber device (confusingly called a reflector), which connected back to the third rotor.

The extra wiring and the reflector also made the machine ‘reciprocal’: this means it was able to both encipher and decipher messages – a huge
convenience. Enigma was also relatively simple to set up and operate, requiring only a key that detailed the rotors to be used, the rotor order and starting positions and the plugboard connections. Every day a new key was taken from a code book that contained 28 new settings, and which was issued every four weeks. A captured Enigma machine was useless
without this code book.

**CHINKS IN THE ENIGMA ARMOUR**

However, Enigma was broken through the dedicated work of hundreds of cryptanalysts based at Bletchley Park in England, the most famous of whom was Alan Turing.
There were a number of characteristics of Enigma and the way it was operated that reduced the number of possibilities requiring investigation of each day’s crop of messages:

- No letter could be encrypted as itself, so a ciphertext A was never a plaintext ‘a’.

- Letters could not be
encrypted as their neighbours in the alphabet, so ‘b’ could never be A or C.

• Enigma was reciprocal, so if ‘f’ was enciphered as K, then ‘k’ would be transformed into F.

• The rotors were not allowed to be in the same place two days running. So if the order one day was
123, it could not be 213 the next day because the third rotor hadn’t moved.

• There was a pattern in how the rotor wheels turned.

While all these were valuable in reducing the number of possibilities requiring investigation, the crucial attack weapon was a crib – some part of the plaintext that was already known or
guessable. For example, weather reports were sent in at the same time, in the same format, each day, so decrypters learned to guess where words such as ‘weather’ were likely to be, giving them a massive head start.

Finally, and some argue most importantly, a powerful weapon for the decrypter was
quite simply operator error:

- In choosing the settings, cipher operators develop habits, such as choosing letters from the same row of the keyboard, or a girlfriend’s initials, when selecting the message key. Code breakers learnt to recognize these and to try popular options first.

- Any mistake by German
communications staff helped the code breakers: the necessity to re-send a message created a valuable crib allowing comparison of two ciphertexts, or a procedural flaw (for example, a sloppy operator sending a second message using the same key settings) weakened the security of the encryption.
Robert Harris’ 1996 novel *Enigma*, about the code-breaking heroics at Bletchley Park in World War II, was made into a film in 2001. There are
other films that feature codes:

• *Cryptonomicon* (1999) by Neal Stephenson features cryptography from World War II, including Enigma.

• Marlene Dietrich sends a message through the notes of the piano music she
performs in the 1931 film *Dishonored*, a re-telling of the Mata Hari spy story from World War I.

• The James Bond book (1957) and film (1963) *From Russia with Love* features a decoding machine that was christened Spektor in the book and then
Lektor in the movie.

- The 1983 film *A Christmas Story* includes a retelling of a myth that the Little Orphan Annie radio show broadcast a secret message that deciphered as ‘Be sure to drink your Ovaltine’, promoting the show’s sponsor.
• The 2001 film *U-571* tells the fictional story of some American submariners obtaining an Enigma machine by hijacking a German submarine.
Other World War II code breaking

PURPLE CODE

American code breakers, known as the SIS (Signals Intelligence Service), broke a Japanese code in the 1930s, dubbing it ‘RED’. Its far
more complicated successor, introduced in 1939, was known in the US as PURPLE.

Since Japanese is too complex to encipher, the Japanese opted to write Japanese sounds phonetically using a Roman alphabet, which they called ‘Romanji’.

The plaintext was fed into a machine that encrypted in a
similar pattern to Enigma, except instead of rotors it used switches like that of a telephone exchange, combined with an extremely complicated wiring system.

BREAKING THE PURPLE CIPHER

The SIS broke PURPLE mainly thanks to a brilliant cryptanalyst, William
Friedman. He got hold of a tabulating machine that used punched cards to monitor accounts information and used it to analyze Japanese communication. PURPLE was broken by:

- Statistical analysis using the tabulating machine.
- Friedman and his staff working out how the Japanese were choosing the
message keys.

- Some of the same messages being sent in both ciphers during the transition from RED to PURPLE, providing a crib.

- The Japanese sending messages in a uniform format with predictable beginnings and endings.

- The amount of communications traffic
between Japan and its embassies increasing hugely in 1940 because of negotiations with the Nazis, providing a large flow of encrypted data to analyze.

**JN-25**

PURPLE code was used for diplomatic communication. Japanese naval messages, however, were sent in a
completely unrelated (genuine) code, dubbed JN-25 because it was the 25th known system used by the Japanese navy.

Only a navy could use such a code, because it comprised more than 30,000 phrases, words and letters, each given a five-digit number – securely storable on a ship, but not by a mobile army. The bulky
encrypting and decrypting code books were changed every few months.

A weakness in large-scale code systems is that since dictionaries are alphabetical, the numerical codes can provide clues to the starting letter. For example, if ‘Attack’ has the number 200, it is likely that lower numbers will be for words appearing
before that word in a dictionary and vice versa.

The JN-25 code counteracted this by altering the numbers with a complicated addition system.

Again, the code was solved partly because of blunders by its operators, who continued to use outdated codes that had been broken by staff at
Bletchley Park, who had discerned a recognizable pattern in the numbers being based on multiples of three. Reading JN-25 code messages gave the US a major advantage in the Battle of Midway on 5 June 1942. The American victory in this battle marked a turning point in the war in the Pacific.
6

Codes in Cyberspace

Computer encryption
The binary number system was first used as a coding method in the 19th century and then became the computer language of the 21st century. Today’s plethora of electronic encryption systems are based on the simple idea that ‘1’ stands for ‘on’ and ‘0’ means ‘off’.
Computer language is in binary code. This is a number system with a base of two, so the only digits used are 0 and 1. In our base-ten decimal number system, ‘1234’ is one thousand, two hundreds, three tens and four units. In base two, the first seven columns have the following values:

64 32 16 8 4 2 1
So ‘2’ is expressed as 10, ‘3’ as 11, ‘4’ as 100, ‘5’ as 101 and ‘6’ as 110. Each digit in binary code is known as a ‘bit’, and each ‘bit’ is either ‘on’ or ‘off’, with a value of ‘true’ or ‘false’.

In 1874, Emile Baudot patented a five-bit code, which allowed representation of characters, numbers and some punctuation using
binary code. This was intended for use in the pulsing ‘on’, ‘off’ telegraphic communication system, and replaced Morse code in the mid 20th century. Baudot’s code had 32 five-digit ‘words’ allowing it to represent every character in the alphabet plus some other symbols, but special signals doubled its capacity by indicating if the subsequent
‘words’ represented letters (11111, also expressed as ++++++ or numbers (11011, also expressed as +++.---++)).

Baudot’s code developed into the seven-bit American Standard Code for Information Interchange (ASCII) code that continues to be used by computers today.
Capital letters are represented by ASCII binary numbers (see chart on page 146). Numbers are expressed as $1 = 0110001$, $2 = 0110010$, $3 = 0110011$ and so on.
British rock band Coldplay used Emile Baudot’s 1870 binary code to create the set of coloured blocks on the artwork of their 2005 *X&Y*. The shapes are loose representations of the letters of the CD’s title achieved by using binary code to decide whether to
have a block or a gap. The colours are irrelevant.

BILATERAL CIPHER

In 1563, Francis Bacon published his
bilateral cipher in which all letters are represented by ‘a’ and ‘b’. It bears an uncanny resemblance to five-bit binary code (see opposite).

The Baconian alphabet is:

A = aaaaa
B = aaaaab
C = aaaba
D = aaabb
E = aabaa
F = aabab
G = aabba
H = aabbb
IJ = abaaa
K = abaab
L = ababa
M = ababb
N = abbba
O = abbab
P = abbbba
Q = abbbbb
R = baaaa
S = baaab
T = baaba
UV = baabb
W = babaa
X = babab
Y = babba
Z = babbb

So ‘Bacon’ is enciphered as

aaaaabaaaaaaaaaaaabaabbbababbbababby

This, of course, creates a
simple code. The message ‘Binary’ is written as:

<table>
<thead>
<tr>
<th>1000010</th>
<th>1001001</th>
</tr>
</thead>
<tbody>
<tr>
<td>1001110</td>
<td>1000001</td>
</tr>
<tr>
<td>1010010</td>
<td>1011001</td>
</tr>
</tbody>
</table>

which can also be written as a single string or broken into blocks of five. The following are two kinds of computer encryption that can still be
done (at an infinitely slower pace than a computer!) by hand, and so could still be used as a simple cipher.

<table>
<thead>
<tr>
<th>ASCII numbers</th>
<th>binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10000001</td>
</tr>
<tr>
<td>B</td>
<td>10000100</td>
</tr>
<tr>
<td>C</td>
<td>10000111</td>
</tr>
<tr>
<td>D</td>
<td>10001000</td>
</tr>
<tr>
<td>E</td>
<td>10001011</td>
</tr>
<tr>
<td>Letter</td>
<td>Binary</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>F</td>
<td>1000110</td>
</tr>
<tr>
<td>G</td>
<td>1000111</td>
</tr>
<tr>
<td>H</td>
<td>1001000</td>
</tr>
<tr>
<td>I</td>
<td>1001001</td>
</tr>
<tr>
<td>J</td>
<td>1001010</td>
</tr>
<tr>
<td>K</td>
<td>1001011</td>
</tr>
<tr>
<td>L</td>
<td>1001100</td>
</tr>
<tr>
<td>M</td>
<td>1001101</td>
</tr>
<tr>
<td>N</td>
<td>1001110</td>
</tr>
<tr>
<td>O</td>
<td>1001111</td>
</tr>
<tr>
<td>P</td>
<td>1010000</td>
</tr>
<tr>
<td>Q</td>
<td>1010001</td>
</tr>
<tr>
<td>R</td>
<td>1010010</td>
</tr>
<tr>
<td>S</td>
<td>1010011</td>
</tr>
<tr>
<td>T</td>
<td>1010100</td>
</tr>
<tr>
<td>Letter</td>
<td>Binary</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>U</td>
<td>1010101</td>
</tr>
<tr>
<td>V</td>
<td>1010110</td>
</tr>
<tr>
<td>W</td>
<td>1010111</td>
</tr>
<tr>
<td>X</td>
<td>1011000</td>
</tr>
<tr>
<td>Y</td>
<td>1011001</td>
</tr>
<tr>
<td>Z</td>
<td>1011010</td>
</tr>
<tr>
<td>1</td>
<td>0110001</td>
</tr>
<tr>
<td>2</td>
<td>0110010</td>
</tr>
<tr>
<td>3</td>
<td>0110011</td>
</tr>
<tr>
<td>4</td>
<td>0110100</td>
</tr>
<tr>
<td>5</td>
<td>0110101</td>
</tr>
<tr>
<td>6</td>
<td>0110110</td>
</tr>
<tr>
<td>7</td>
<td>0110111</td>
</tr>
<tr>
<td>8</td>
<td>0111000</td>
</tr>
<tr>
<td>9</td>
<td>0111001</td>
</tr>
</tbody>
</table>
STREAM CIPHER

In a binary code, changing every bit would be very easy to decipher, but creating a repeating string of bits out of a keyword (known as a key stream sequence) means that ‘0’ can be interpreted as ‘leave’ and ‘1’ as ‘change’. This concept was introduced in 1919 by Gilbert Vernam as
a way of enciphering Baudot messages. It is best expressed in a simple table and is known as the XOR operation:

<table>
<thead>
<tr>
<th>Key</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

| Plaintext | | Ciphertext |
|-----------|---------------|
| 1         | 0             |
| 0         | 1             |
Binary code can be transmitted in ‘on’ or ‘off’ pulses where a timing mechanism checks the line, say, every tenth of a second, to see if a pulse can be sensed or not. Binary code can also be transferred to paper tape by punching, or not punching, holes at regular intervals.

Vernam devised a system in which two punched tapes, one
holding the plaintext, the other a key of random numbers, were fed together into an adapted teletypewriter. If two holes matched up, a hole, or pulse was transmitted. If two holes did not match up, it left a space.

This allowed instant transmission of messages typed in plaintext,
automatically encrypted, and automatically decrypted by a receiver using an identical key tape. This was a huge advance in cryptography, partly because no one had to sit and encrypt or decrypt the message, which was typed in as normal, transmitted in code, yet fed out at the other end as plaintext – you didn’t need a skilled cryptanalyst at either end of the process.
(although you needed someone who could run the machine properly).

Using Vernam’s method, but employing today’s seven-bit ASCII language, the plaintext ‘computers’ can be encrypted using the keyword BAUDOT as follows (the keyword is repeated as often as is needed):
The ciphertext is created by adding the digits of the ASCII message and ASCII keyword, i.e. \(1 + 1 = 0\); \(1 + 0 = 1\); \(0 + 1\)

<table>
<thead>
<tr>
<th>Message</th>
<th>Computers</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASCII message</td>
<td>100011001111001101110100001011010110100100010110100101100111</td>
</tr>
<tr>
<td>Keyword in ASCII</td>
<td>100010100001110101011000100100111111101010010000101010000111010101</td>
</tr>
<tr>
<td>Ciphertext</td>
<td>000001000111000110000010100001110100000000000111001001110000110</td>
</tr>
</tbody>
</table>
1. This is called a stream cipher.

The message is decrypted by reversing this process.

**Block Cipher**

Another encryption method is the block cipher, in which the bits are grouped into threes, which can then be converted into digits, using the ASCII
For example, the same ‘computers’ message can now be encrypted as follows:

<table>
<thead>
<tr>
<th>Block</th>
<th>100 001 110 011 111 001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>4 1 6 3 7 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Block</th>
<th>011 010 100 100 010 111</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>3 2 4 4 2 6</td>
</tr>
</tbody>
</table>
This shorter ciphertext of 416371550253244264523 can still easily be converted back to binary numbers and then into characters.

Key distribution

With the coming of the internet and email electronic communications, messages are bounced off satellites
around the globe. This is fast, but of course such signals can be picked up by anyone, so it is easy for someone with the right knowledge and software to intercept messages. Stream and block ciphers are efficient methods of encryption but suffer the problem of every key-based encryption method: distributing the key.
ALICE AND BOB

By tradition, most explanations about message verification and the use of keys utilize the names Alice and Bob – involving both sexes and offering human-
sounding substitutes for A and B. The evil interceptor is usually referred to as Eve, which is either a biblical reference or short for ‘eavesdropper’.

PUBLIC KEY
Academics and mathematicians struggled to find a way to keep coded communication secure for decades. There is a simple theory of how to achieve message security:

Alice wants to send a secret message to Bob.

• She puts it in a box,
padlocks it and sends it to Bob without the key.

- Bob attaches his own padlock to the box and returns it, again without the key.

- Alice now removes her padlock and returns the box.

- Bob can now open the box with his own key.
This is known as an asymmetric key system, in which a different key (or, if you prefer, combination) is required to decrypt rather than encrypt. In a symmetric key system, Bob would simply use a copy of Alice’s key to open the box – re-introducing the key distribution problem. The asymmetric system is fine if physical keys are used,
because two different keys can (separately) lock one box. However, if the key is a code, the process won’t work because Bob has to put Alice’s box inside another box, which he then locks and dispatches. Now she can’t get inside to her original box.
Prime numbers have no factors apart from 1 and themselves. So 3, 5 and 7 are prime, but 2, 4, 6, 8 and 9 are not as they can be divided by other numbers. The prime numbers used in
public key encryption usually have five or more digits.

MODULAR ARITHMETIC

In modular arithmetic, numbers change after they reach a certain value. The clearest simple example is a clock face. Add on six hours
from 9 o’clock and you reach
15 o’clock, known as 3 o’clock.

In 1976, Martin Hellman hit
upon a way of using modular
arithmetic to allow Alice and
Bob to exchange information
in a similar way to the
example above, using the
results of their calculations as
their key, knowing that no
one listening in could find the
key. The method is called the Diffie–Hellman–Merkle key exchange scheme, named after Hellman and his two colleagues. However, it needed to be made more efficient, with less exchanges of information, to be usable. This was achieved by the RSA and ElGamal systems, which employ the Diffie–Hellman–Merkle method with multi-digit prime
numbers. Detailed explanations of the maths involved in the scheme can be found at these websites: http://en.wikipedia.org/wiki/Diffie-Hellman and www.vectorsite.net/ttcode_10.

PRETTY GOOD PRIVACY

Computer encryption became
a genuinely public tool in the early 1990s with the advent of Pretty Good Privacy (PGP), a computer program that provides cryptographic privacy and authentication, allowing anybody to use the RSA method of encryption without having to deal with its complex mathematics. It was invented by Phil Zimmerman and is mainly used to protect email
communications, which otherwise are completely unencrypted, although it can also be used to encrypt other computer data.

Its introduction was stimulated by the extraordinary rise in internet and email communication, echoing the impact of the invention of the telegram and radio systems in the past.
Communication has never been easier, and nor has interference in it. Digital communication raises a number of issues, stemming from the fact that anyone, anywhere in the world, can send and intercept messages. There are enormous opportunities for fraud, scams and interference in the workings of other computers.
Most purchases made on the internet involve the use of cryptography to protect credit card numbers or other financial information. This is done via the Secure Sockets Layer (SSL) and, increasingly, Transport Layer Security (TLS). Pages protected by SSL have an ‘https’ prefix instead of the conventional ‘http’ one, and use a blend of protocols and
algorithms to enable key exchange, authentication and communication in cipher.

**Passwords: the modern key**

For most of us, thinking up a password or username to key into our computer is the nearest we get to having to
create a secret code, and most people are terrible at it. Any system is only as strong as its weakest link, and passwords are by far the easiest form of encryption to attack.

A GOOD PASSWORD
So a good password should follow these rules:

• Use the maximum possible number of characters.

• Do not use recognized words or names.

• VaRy capiTaliZation.
• Include random numbers or graphic signs as well as letters. This is known as ‘salting’.

There is a basic problem: we are encouraged to create passwords we can remember easily, and discouraged from writing them down. So there
is a strong temptation to use words related to the task (‘password’ is a common password!) or names of family members, and to keep the password short – studies suggest more than one in seven passwords are only three keys long. Forcing people to change their password every month isn’t effective because even if they follow the rule, they tend to
choose passwords they have used already, often simply alternating them with each enforced change.

One of the biggest sins is to choose letter-only passwords, especially those that spell a word. But, of course, this is the easiest way to generate a username that you can be sure of remembering.
Someone who wants to find your password will try the most obvious choices first – family names, maiden name, mother’s maiden name. In the age of the internet, such data is relatively easy to find. The computer can also be used to test out other possible passwords through a
dictionary attack. In this, a computer simply tries every word in the dictionary until it hits the right one. Powerful machines can do this in a matter of seconds. If they fail, they’ll do the same with reversed words, varied capitalization and extra numbers.

A more sophisticated way to find your password is a
timing attack. This notes how long it takes you to key in your password, allowing the computer to calculate the likely number of characters in it. This system can also be used to measure how long it takes for a password to be rejected: the more time, the closer the guess is likely to be.

Another form of attack is
‘password sniffing’. This is when a hacker installs software on your computer that stores the first few keystrokes of every session, which is very likely to include your password.

The trouble is, of course, that a good, complex password is then hard to remember, so you’ll most probably need to write it down. Provided it is
kept in a secure place (not on a note stuck to the computer!) this is a sensible option – your attacker is probably tapping a keyboard thousands of miles away, not snooping around your desk. This method allows you to create trickier passwords without needing to make them memorable. You’ve got the same level of security as the code books that were used for
Encryption is used to protect financial information sent by hole-in-the-wall Automatic Teller Machines (ATMs). The customer places their plastic card with its magnetic strip (and, increasingly, its identifying ‘chip’) in the machine and enters their
Personal Identification Number (PIN). This is communicated to a central computer, which checks the data and ascertains if the customer is permitted to make a withdrawal – so information has to pass both ways.

The four-digit PIN is ‘padded’ with extra digits and all data is sent in encrypted
form using a Data Encryption Standard (DES) cipher. The PIN obviously takes the place of a ‘password’ or ‘key’, and since there are only 10,000 possible combinations of four-digit numbers (a tiny figure compared to the multi-digit encryption possibilities generated by RSA) the ATM only allows three attempts to input the number before retaining the card. This is
starting to seem generous given that a recently devised, extremely complex mathematical attack method can allegedly identify a PIN in about 15 guesses.

QUANTUM CRYPTOGRAPHY

Standard cryptography uses the laws of mathematics.
Quantum cryptography uses the (highly complex) approaches of quantum mechanics and the physics of information. Communication is via photons in optical fibres or electrons in electric current. Since these are measurable and the channel is highly sensitive, any eavesdropping is immediately detected, so communication ceases until it can be kept
safe.

Additionally, quantum computers are theoretically capable of incredibly fast factoring of large numbers, and so may be able to break RSA keys and crack DES and block ciphers far faster than the present generation of conventional computers.

To summarize: quantum
technology may be able to find a way to crack codes faster than ever, but also to create secure, closed communications systems.

QUANTUM MECHANICS

The theory of quantum
mechanics is a completely different way of looking at the world. It replaces Newtonian mechanics and classical electromagnetism at the atomic and subatomic levels and underpins various fields of physics and chemistry, such as condensed matter.
physics, quantum chemistry and particle physics.

CODES IN MUSIC

Composer Edward Elgar loved puzzles
and codes, and managed to create a musical puzzle of his own in his *Enigma Variations*, a series of musical character portraits that is one of his best-loved works. In all, 14 people and a dog are featured in the variations: the people are identified by initials, except for the
13th variation, which may have been about a lover of Elgar’s who had left England. Elgar also revealed that he ‘hid’ a well-known tune in the fabric of the score, and for many years musical detectives tried to find it. At last, in 1991, the musicologist Joseph Cooper solved the
conundrum – the tune is a passage from the slow movement of Mozart’s Prague Symphony.

Elgar’s other cipher

Elgar was the author of the Dorabella cipher, a note sent to his friend Mrs Dora
Powell in 1897. It comprises 87 squiggly characters at various angles in three neat lines. No one has managed to decipher it and it is thought it may have been linked to the mystery surrounding his *Enigma Variations.*

**Do, re, mi, fa,**
Composers have long been able to use letters identified with musical notation to build words into their music. There are two methods:

• The Sol Fa scale creates the syllables do, re, mi, fa, sol, la, and si/ti, represented by the notes C, D, E,
F, G, A and B.

- Western notation uses the letters A to G but in German musical tradition B is also known as H, and E flat is represented by S, providing a somewhat limited but usable alphabet.

Baroque composers
often used these letters to weave the names of friends or places into their music, and Johann Sebastian Bach was particularly fond of spelling out his surname. Robert Schumann’s *ABEGG Variations* records the name of a woman he was in love with: Meta Abegg.
Russian composer Dmitri Shostakovich frequently represented himself with the musical motif DSCH. He also used the sequence EAEDA to represent his student Elmira Nasirova, creating the ‘word’ E, La, Mi, Re, A.
From black chambers to Cheltenham

The new doughnut-shaped spy centre in Cheltenham, the centre for signals intelligence in the UK, is the latest instalment in an international saga of concealing and probing communications that
grows back for centuries.

INTERCEPTION AND DECEPTION

As explorers found new lands and places with which to trade, countries needed diplomats to negotiate with other governments around the world. Inevitably, their messages would be
intercepted, so they started using codes to conceal their content. Thus, the ‘black chambers’ were born. In many countries from the 16th century onwards, missives sent by foreign diplomats were routinely intercepted (often through bribing lowly-paid or greedy officials), opened, copied, re-sealed and sent on their way while clerks began breaking their codes.
The practice became even more widespread when Britain separated from the Catholic Church, as European countries and the papacy discussed the significance of the move and manoeuvred for political advantage from it.

This culture of secrecy and subterfuge stimulated cryptological endeavour. In 16th-century Venice, there
were specialist schools on the subject, such was the demand for cryptography in this commercial and diplomatic centre. In England, the success of Elizabeth I’s spymaster Walsingham in trapping her rival Mary, Queen of Scots (see pages) was down to his efficient team of code breakers. In 1703, William Blencowe became the first Englishman
to get a regular salary for cryptanalysis, receiving £100 a year and taking on the title Decrypter. Actually the job had been going for years, and Blencowe was taking over from his grandfather, John Wallis, who had trained him up in the dark art of unravelling secrets from coded writing.
NEEDLING OUT INFORMATION

One efficient method for seeing letters sealed in envelopes used a long needle rather like an old-fashioned sardine tin opening key. The
needle was slipped into a corner of the envelope and turned, rolling up the paper inside. This could then be removed, copied and returned with the same method, leaving the envelope seal undisturbed. The only evidence of tampering was a small hole in the corner of the envelope.
At this time, the most active and efficient black chamber in the world was the *Geheime Kabinets-Kanzlei* in Vienna, Austria. Here the day’s intercepted missives arrived at 7 o’clock each morning, and were immediately dictated to secretaries to prepare the copy that the
cryptanalysts would set to work decoding while the original message went on its way. Staff received financial incentives to master new languages and successful decryptions earned a substantial bonus, paid in person by the grateful king, Karl V. They even got compensation for lost payment opportunities if one of their spy colleagues
succeeded in stealing solutions direct from the embassies!

The analysts worked one week on, one week off in their Viennese office in recognition of the mental strain of their job. Indeed, there is a long history of rapid weight loss, stress and even nervous breakdown associated with the people
working in this field. Cryptanalysts have often reported difficulty in sleeping, and recurrent dreams in which they are faced with impossibly big searches, like finding the right pebble on a beach.

Throughout the 18th and early 19th centuries, the black chambers were a secret mini industry that recruited the
brightest minds and trained them to puzzle out the secret messages of friends and foes. Such was their expertise and value that they would often be kept on by new administrations and monarchs even when other officials who had served the previous government were disposed of.
NO SUCH AGENCY

The National Security Agency (NSA) is said to be the world’s major employer of top mathematicians, to own the largest group of supercomputers, and to have a bigger
budget than the CIA. This is despite it being surrounded by such secrecy that its initials were said to stand for ‘No Such Agency’.

Do they know what we know?
A recurring issue in the history of code breaking is what to do with the information gathered, because of the risk that your opponent will realize their codes have been broken and so change them. Diplomats would remain tight-lipped as other ambassadors expressed opinions known to be the opposite of the view expressed in secret.
communication with their superiors. At one point, the Spanish government was a laughing stock among diplomats because its codes were so easy to break. They chuckled behind their hands, however, as the information they were receiving was so useful. In
A number of universities in the UK and the US now offer courses about information security in which much of the content is about cryptography.
Graduates tend to go on to work as IT security managers or consultants.

World War II, the British sometimes took no action over decrypted messages for fear of alerting the Germans to their ability to read Enigma messages, and at the time of
the Pearl Harbor attack, the Americans knew the Japanese were going to break off diplomatic relations before their own ambassador could deliver the message.

An amusing variation on this theme of whether to reveal or conceal what you know is the story from Henry II’s siege of Réalmont in 1628. A decoded intercepted message revealed
the defenders had few supplies left. Henry sent in the decoded plaintext of their letter and they surrendered, knowing they had no chance of success.

WHERE ARE THEY NOW?

After World War I, the US founded MI-8, a code-breaking team under the guise
of a New York commercial code production company. It was briefly closed down in 1931 on the orders of Secretary of State Henry Stimson, allegedly with the comment ‘Gentlemen don’t read other gentlemen’s mail’. This deprived its key worker, Herbert Yardley, of an income so he wrote a book about his work, which alerted the Japanese to the fact that
America had broken its codes, which were subsequently completely redeveloped.

Later reformed, MI-8 became the National Security Agency (NSA), combining its work with the Central Security Service (CSS).

Britain’s black chamber moved to Room 40 at the
Admiralty in London during World War I. It later evolved into Bletchley Park and various other sites, and today the Government Communications Headquarters (GCHQ) at Cheltenham is the UK centre for signals intelligence and information protection.

Both the American and British secrets-busting
organizations run websites explaining some of what they do (see pages).
## A code chronology

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>c. 1900 BC</td>
<td>Some Egyptian hieroglyphs are written in non-standard characters, a code apparently intended to</td>
</tr>
<tr>
<td></td>
<td>add a little mystery. Possibly as little as 250 years later a small part of the Leiden, or</td>
</tr>
<tr>
<td></td>
<td>Ipuwer, papyrus (dates much</td>
</tr>
</tbody>
</table>
disputed) was partly in cipher

c. 1500 BC  A Sumerian pottery glaze recipe is written in code

600–500 BC  Hebrew scribes use the ATBASH cipher

487 BC  Greek use of the scytale device is recorded

60–50 BC  Julius Caesar’s shift cipher is used

AD 0–400 (date)  The *Kama Sutra* of Vatsayana lists cryptography as the 44th
unknown) and 45th of 64 arts (yogas) for men and women

805–873 Lifespan of Abu Al-Kindi, the first genuine cryptanalyst

c. 1214–94 Lifespan of Roger Bacon (Dr Mirabilis), who described ciphers in use

1379 Gabrieli di Lavinde publishes the first-known nomenclators

*Treatise on the Astrolabe*, attributed to English poet Geoffrey Chaucer,
contains some enciphered passages

Subh al-a‘sha, a 14-volume encyclopedia written by Shihab al-Din al-Qalqashand, includes material on cryptology

Leon Battista Alberti invents the cipher disc to allow encryption using two alphabets

Religious disputes make secret communication more important,
16th century stimulating the use of nomenclators and ciphers and the growth of the deciphering ‘black chambers’

1516 First printed book on cryptology, *Steganographia* by Johannes Trithemius, is published

1563 Giovanni Battista Porta creates the first-known polygraphic substitution cipher
1586 Blaise de Vigenère publishes his Vigenère square

1587 Mary, Queen of Scots is executed after her codes are broken

17th century Antoine and Bonaventure Rossignol develop the Great Cipher (date of creation unknown)

1623 Francis Bacon produces his bilateral code

1781 Benjamin Franklin invents the homophonic
substitution cipher

Thomas Jefferson invents the Jefferson wheel, then forgets about it.

1790s

1791 Optical telegraph is demonstrated

1811 Major George Scovell cracks the French codes, helping Wellington win the Peninsular War

1838 Morse code is invented

1844 Invention of the electric telegraph stimulates new
interest in code making

1854 Playfair cipher is invented by Charles Wheatstone

1891 Bazières cylinder develops as a reinvention of the Jefferson wheel

1914–18 World War I encourages the development and use of ciphers, field codes and cryptanalysis

1914 Code-breaking Room 40 is set up at Admiralty House, London
1917  US enters the war as a result of the deciphered Zimmerman telegram

1918  Gilbert S. Vernam and Joseph Mauborgne devise the Vernam cipher one-time pad

1924  Enigma machine first shown

1939–45  Codes and ciphers play an important role in concealing and revealing communications during World War II
1974  Story of how Enigma had been solved is told

1976  Diffie–Hellman–Merkle key exchange scheme introduces the idea of public key encryption

1977  RSA algorithm makes public key encryption feasible

1990s The rapid global rise of the internet and email communication highlights the issue of digital cryptography
<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>First research is published on quantum cryptography.</td>
</tr>
<tr>
<td>1991</td>
<td>Phil Zimmerman releases his Pretty Good Privacy program.</td>
</tr>
</tbody>
</table>
A variety of code-breaking methods are mentioned in this book where relevant to certain types of codes or ciphers. However, a summary of approaches can be useful.
Gather all the encoded messages together. The more material you have at this stage, the more leads there are to follow.

IS IT A SIMPLE CODE?
Study the text carefully to see if any of it makes sense. Look for the balance of vowels and
consonants. Could it be an anagram? Check for embedded words or letters by looking at initial letters, and try counting every second, third or fourth (etc.) letter. Look at the start of the message and consider if any of it could contain instructions for the decoder: 1/5, for example, might indicate that the plaintext is contained in every fifth letter.
There may also be clues if a keyword or date has been used (see ‘Identifying the key’ on page 175).

THE METHOD

The scientific process for code breaking is:

- Analysis (letter
counting, etc.)

- **Hypothesis** (guesswork)
- **Prediction** (if you find ‘e’, other letters become clearer)
- **Verification** (you were right!) or
- **Refutation** (start again, checking if what you think you
This method takes you through three stages:

• Identification of the code type

• Breaking the code to see how it works

• Setting, which is the term for decrypting individual messages.
CODE SUPPOSITIONS

If it seems likely that groups of characters each represent a word or phrase, remember they were put together in a dictionary, most probably in alphabetical order and so lower value numbers are likely to refer to words beginning with letters that appear earlier in the alphabet, and vice versa.
WORD BREAKS

It is a major advantage if the ciphertext is still broken down into words, in which case it will be in groups varying in size between one and about eight characters, with the majority of words being three, four or five letters long. If the groups are larger but still irregular, each letter may be represented by
more than one character as in a Baconian or binary code (see page).

If word lengths are clear, look for one-, two- and three-letter words. Check them against the list of the most common ones on page 94. Look for the definite or indefinite article (‘the’, ‘a’ or ‘an’) at the start of sentences.
In continuous ciphertext, the word (or, more often, sentence) breaks may be indicated by a null, the most obvious being an X or some other repeated character or pattern. Identifying the starts of sentences allows you to make more informed guesses about likely words, and therefore initial letters, at those points in the ciphertext.
CRIBS AND CONTEXT

If part of the message is in plaintext (i.e. not encoded or enciphered), read up to the enciphered words and consider the context. Could they be names, places or numbers? If not, it is likely to be more technical language that is being concealed, so consider the overall context.
of the message: is it about battle preparations, finance, romance, etc.?

You may be able to find a crib. Are any sections of two messages the same, in word length or character pattern? If you know something of the context, what words or phrases might you expect to find? Does the message (or, far more usefully, do the
messages) appear to start or finish in a formal way? Sometimes you can create your own crib. Assume that the message starts with the letter ‘a’. Work out the key letter that could have encrypted ‘a’ into the first ciphertext letter. Try that key on the first letters of the other messages. Then repeat, assuming the first letter is ‘b’. Each time, look for patterns
in the lists of letters.

**LETTER FREQUENCY**

Make a frequency chart for each character in the ciphertext. Refer to the letter frequency information on pages 91–5. The most common letter in English is ‘e’, which will appear
noticeably more often than others in a monoalphabetic cipher. The frequency table will give you clues as to other common letters.

Once you can identify a common letter, you can make more informed guesses about its neighbours (for example a three-letter word ending in ‘e’ is very likely to be ‘the’). You can also look for
common letter patterns such as ‘ing’ (which is particularly useful as it ends a word, giving you less options to check for the start of the next word).

If the encryption is polyalphabetic, there will be no marked differences in character frequency, but there may still be discernible patterns. These can help you
identify the key.

IDENTIFYING THE KEY

If you suspect a key was used to encrypt, look for any indications from the sender of what the keyword is – it may not have been agreed in advance.

Encryption using a keyword
can be attacked by guessing the length of the keyword. This can be achieved by exploiting the fact that certain patterns (such as ‘th’ or ‘ed’ are likely to appear very frequently in the plaintext and so may at some stages have been encoded with the same ciphertext letters. Number each character by position (1st, 2nd, 3rd, and so on). Now look for any repeated
patterns (say, repeated pairs or triplets of letters or numbers).

Count how many characters there are between the start of each repeated pattern. If you find that certain pairs of characters appear after 54, then 120, then 96, then 186 ciphertext letters, you can turn to mathematics to help. Try dividing each of these
numbers by three, then four, then five, and so on. Disregard any answers that are not whole. In this case, you will find that each number is divisible by both three and six. This suggests the keyword used to encipher was either three or six letters long. The lower figure seems unlikely, so try six first. Now you can test the assumption that the enciphering pattern is
repeated every six letters by putting the ciphertext into six columns and carrying out frequency analysis on each column, because those letters are likely to have been enciphered with the same letter.

For example, if part of the ciphertext read JAQLZSOUFBLWPNIAFYHX, and you assume from
analysis of the full text that the keyword length is six, arranging the letters of this section into six columns produces:

```
J  A  Q  L  Z  S
O  U  F  B  L  W
P  N  I  A  F  Y
H  J  B  I  W  L
V  C  E  U  F  C
H  B  V  V  X  X
```
Frequency analysis of the full ciphertext might then reveal that the letters in the first three columns were shifted 3, 9 and 16 places respectively on a Vigenère square, producing a plaintext of:

```
g r a ? ? ?
l l y ? ? ?
m e s ? ? ?
 e s t ? ? ?
s t o ? ? ?
```
This would show that the first three letters of the keyword are CIP. A combination of continued frequency analysis and guesswork of the keyword would reveal it as CIPHER, with the full plaintext in columns now known to be:
and the plaintext can be written as, ‘Gradually the message starts to make sense’.

This example shows that
sometimes you will be able to work out part of the keyword used and from this discover the whole keyword. For example, if you are pretty sure the keyword has the letters ENC___T_ON, you can guess the full keyword is ENCRYPTION.
HISTORY OF DECRYPTION

The history of decryption is as long as that of encryption and there is only space here to show the initial basic steps. More information can be found at:
www.bbc.co.uk/dna/h2g2/alabaster/A613135
and
www.vectorsite.net/ttcode_01.html
Glossary

algorithm: A set of mathematical instructions forming a step-by-step procedure to encrypt or decrypt information.

ASCII: American Standard Code for Information Interchange, used to represent text in
computers.

**asymmetric key system:** A cryptological system where a different ‘one-way’ key is needed for encrypting and decrypting.

**authentication:** The process of confirming identity.

**bigram:** Pair of letters, syllables or words, commonly used as the
basis for statistical analysis of text.

**binary code:** Code system using only two characters or numbers, 0 and 1.

**black chamber:** General term for the cryptological offices set up by various governments since the 16th century where intercepted messages
were studied for decryption.

**block cipher:** A cipher in which blocks of text are enciphered in groups, usually each of 64 bits.

**Caesar shift cipher:** Cipher in which a letter is replaced by another, a set number of places along in the alphabet.

**cipher:** A process in which
individual letters are reordered or replaced to conceal the meaning of a text.

ciphertext: Enciphered text.

clear text: Another term for plaintext or *en clair*.

code: A system in which words or phrases are reordered or replaced for concealment. The word ‘code’ comes from the Latin *for ‘book’,*
‘codex’.

code book: The crucial ‘dictionary’ giving words and phrases and the character(s) to be used to represent them. For larger-scale codes, a reverse dictionary is required for decryption.

crib: A section of known plaintext, which can be used to break a code or cipher.
cryptanalysis: The art of breaking codes and ciphers.
cryptography: The art of devising codes and ciphers.
cryptology: The general term for cryptanalysis and cryptography.
cryptosystem: A system for encrypting and decrypting data.
decipher: To turn enciphered text into the original message, or plaintext.

decode: To turn a coded text into the original message, or plaintext.

decryption: The process of turning encoded or enciphered text into plaintext.

DES: Data Encryption Standard, the algorithm widely used for data
encryption, adopted in 1976.

**Diffie–Hellman–Merkle key exchange:** Process for establishing a secret key through public discussion.

**digital signature:** Electronic identification of a person, using a public key algorithm.

**digraph:** Two letters representing one sound,
such as ‘ph’ or ‘th’, forming common pairings that are useful in decryption.

**encode:** To turn plaintext into a coded message.

**encipher:** To turn plaintext into a cipher message.

**Enigma:** Most famous enciphering system in recent history, used by the Germans in World
War II and, crucially, broken by the British. **en clair:** Plaintext, un-encoded message.

**fractionation:** Process in which plaintext symbols are converted into new symbols prior to transposition, creating a more complex cipher.

**frequency analysis:** Decryption strategy in
which ciphertext letters are counted to identify patterns, which relate to how often letters occur in natural text.

**Greek square:** Device for changing characters into numbers, using the grid references of a square. Also known as the Polybius square or the Greek checkerboard.
homophonic substitution: A cipher where a letter can be represented by several different characters, thus combating frequency analysis.

key: The set of characters that determines how a text is to be encrypted.

key length: The number of characters or bits in the
key. The longer it is, the harder it is to decrypt.

**monoalphabetic substitution cipher:** Cipher in which the plaintext is encrypted using one alphabet.

**nomenclator:** Encryption using a mixture of homophonic substitution and codes for certain words and phrases. It
was the main cryptological method for several centuries until World War I.

**null**: Part of the ciphertext, which indicates ends of sentences, or which can be ignored in decryption because it is there to confuse enemy code breakers.

**one-time pad**: The only
known totally secure encryption method – a running key of totally random characters, used only once.

plaintext: A message before it is encrypted and after it is decrypted.

Polybius checkerboard: see Greek square.

polyalphabetic cipher: Cipher in which the
plaintext is encrypted using more than one alphabet.

Pretty Good Privacy (PGP): A method for secure, encrypted email communication, developed by Phil Zimmerman.

private key: The ‘secret’ part of an asymmetric key system, also known as the decryption key.
public key: The ‘open’ part of an asymmetric key system, also known as the encryption key.

quantum cryptography: The use of quantum physics to create random bits on a computer, which can be used to create a one-time pad cipher.

running key: A key as long
as the plaintext, as in a book cipher.

**RSA**: Rivest, Shamir and Adleman’s system enabling public key cryptography, invented in 1977.

**steganography**: Greek for ‘hidden writing’, this is the art of hiding the message itself, rather than concealing its
meaning.

substitution cipher: Encryption system in which letters are replaced but remain in the correct position.

superencipherment: Encrypting a message twice, either with the same or with a different method for the second process. Also known as superencryption.
Transposition cipher: Encryption system in which letters change position, creating a giant anagram (in which the letters are also changed).

Vigenère square: The first polyalphabetic cipher made using a tabula recta combined with a keyword.
Further reading

BOOKS

Codes and Ciphers, Robert Churchhouse (CUP, 2002) gives a good briefing with many examples.

Codes, Ciphers and Secret Writing, Martin Gardner (Dover Publications, 2002) is good on simple codes.

Cryptography: A Very Short Introduction, Fred Piper and Sean Murphy (OUP,
2002) is a good briefing on modern trends


**General histories**

_The Code Book_, Simon Singh (Fourth Estate, 1999) covers the story of codes

_The Codebreakers_, David
Khan (Scribner, 1996) is the classic work on the history of codes.

Specific histories

*The Man Who Broke Napoleon’s Codes*, Mark Urban (Faber and Faber, 2001) is the story of William Scovell

*Navajo Weapon: The Navajo Code Talkers*, Sally McClain (Rio Nuevo,
Navajo Code Talkers, Andrew Santella (Compass Point Books, 2004)

The Victorian Internet, Tom Standage (Weidenfeld and Nicolson, 1998) tells the story of the telegraph system
The history of codes
http://fly.hiwaay.net/~paul/cryptography covers the early history of cryptography
www.axsmith.net/encryption.html
www.bbc.co.uk/history/ancient
www.brooklynmuseum.org/ex
www.freemaninstitute.com/Ga
has examples of simple substitutions
www.jproc.ca/crypto/crypto_hist.html
has a timeline of cryptography
www.murky.org/archives/cryptography/lots on history and development

www.simonsingh.com/ website of the author of the excellent The Code Book

www.thebritishmuseum.ac.uk/id=OBJ67home.ecn.ab.ca

www.vectorsite.net/idsearch.html tells the story of codes with plenty on their American history and on
breaking codes
www.world.std.com/~cme/htn
has a chronology of cryptography

Sites for children and beginners
http://search.looksmart.com/p/
has links to sites for children
www.google.com/intl/xx-piglatin/
translates messages into Pig Latin
www.scouting.org.za/codes/
www.10ticks.co.uk/s_codebreaker

allows you to email
coded messages to your
friends

Sites with specialist
information

crypto.pdf is a
comprehensive site on
all aspects of modern
cryptography
http://members.tripod.com/~m shows the cowboy branding alphabet
http://starbase.trincoll.edu/~cr for rail fence and scytale ciphers
www.aldertons.com/index.htm is a cockney rhyming slang dictionary
www.cia.gov/cia/information/ the CIA website on ‘Kryptos’
www.cockneyrhymingslang.co
is a cockney rhyming slang dictionary

www.codesandciphers.org.uk/includes the 1944 Bletchley Park cryptographic dictionary

www.comsoc.org/livepubs/cil/is an article giving an overview on modern public cryptography

www.cs.dartmouth.edu/~jford/gives background on quantum cryptography
www.elonka.com/kryptos is a site about the ‘Kryptos’ sculpture
www.fas.org/irp/world/uk/gchq is an unofficial website about GCHQ
www.gchq.gov.uk/ website of the UK signals intelligence headquarters
www.geocities.com/Vienna/4056/cipher.html shows the Elgar code
www.history.navy.mil/faqs/faqs/faq2.htm contains
information on Navajo code talkers

www.infosyssec.org/infosyssec has lots of background and current information on security

www.nationalarchives.gov.uk

www.profactor.at/~/wstoec/rsa has a step-by-step guide to public encryption

Miscellaneous other sites

http://elonka.com/UnsolvedCc
http://eprint.iacr.org/ an archive of papers on cryptology
http://mad.home.cern.ch/frode lists many papers on post-war codes
www.faqs.org/faqs/cryptography-faq/part01/index.html answers your questions on cryptography
www.gchq.gov.uk/codebreaking is the site of the British government’s
communications experts, and has puzzles and games

www.nsa.gov/history/index.cfm is the website of the USA National Security Agency

www.nsa.gov/museum/index.cfm is the website of the national cryptologic museum in the US

www.puzz.com/cryptoquotes.html has cryptograms of
sayings by famous people

www.spymuseum.org/index.asp is the website of the international spy museum in the US

www.und.nodak.edu/org/crypto/crypto/.sample-issue.html is the website of the American cryptogram association, which was set up in the 1920s
Computer short cuts and symbols

en.wikipedia.org/wiki/Leet has an alphabet of leet symbols

http://tronweb.supernova.co.jp/characcodehis has a brief history of ASCII character codes

www.acronymfinder.com/ gives definitions for acronyms and abbreviations
txt lingo

is a parents’ guide to computer slang

is a dictionary of text short cuts with symbols

is about text short cuts
The pagination of this electronic edition does not match the edition from which it was created. To locate a specific entry, please use your e-book reader’s search tools.

ABFGX cipher 129–30
acrostic codes 48–9
ADFGVX cipher 131
Alberti, Leon Battista 102, 107–8
Alice and Bob 150, 151–2
alphabet, expanding 103–5
American Standard Code for Information Interchange (ASCII) 144, 146, 148, 149
Army of Portugal code 101
Atbash 78–9
Bacon, Francis 145
Baudot, Emile 144, 147
Bazières cylinder 134
bifid cipher 120–1
bilateral cipher 145
binary code 143–6, 147
black chambers 160–5
Blencowen, William 161
Bletchley Park 137, 142, 165
block cipher 149
book code 60
books, codes in 58
Braille 27–9
Caesar shift 72
Cardano grille 50–2
Central Security Service (CSS) 164
Chappe, Claude and René 25
checkerboard ciphers 118–26
Chinese, ancient 39
cipher definition 54
cipher disc 108
ciphers 128–9
combining 107–8
Cockney rhyming slang 45–6
code chronology 167–70
code definition 54
codes 128–9
Cold War 39
columnar transposition 68–9
common letters 91–3
computer encryption 143–9
cowboy branding 17, 18
cribs 131–3, 174
cuneiform writing 15–16

Data Encryption Standard
(DES) cipher 156
date shift ciphers 112
decryption, history of 178
Delastelle, Felix 121
dictionary code 56–9
Diffie–Hellman–Merkle key exchange 152
digital writing, invisible 41–2
digraphs 93–5
dot cipher 61–2
double transposition 70–1
Doyle, Sir Arthur Conan 90
en clair 102
Enigma 135–8, 164
field ciphers 129–31
field codes and ciphers 127–33
films, codes in 70
frequency analysis 89–95, 105, 174–7
beating 97–106
frequency tables 92, 94
Friedman, William 140–1
Government Communications Headquarters (GCHQ)

165

great Paris cipher 101–2

Greek scytale 63–4

Greek square 81–2, 97–102, 118

Hellman, Martin 151–2

Herodotus 37

hieroglyphs 11–13

Histiaeus 38

hobos, 17
inks, for invisible writing 40–1, 43, 47
international signal flags 22–3
invisible writing 40–3
Jefferson wheel 133–4, 135
Jefferson, Thomas 133
JN-25 141–2
key distribution 149–53
key stream sequence 147
keyword identification 175–7
keywords 108–9
using 76–78
King Charlemagne cipher 82, 83
‘Kryptos’ 96
leet 34–5
line cipher 62
Marryat, Captain 19
Mary, Queen of Scots 88–9, 161
Mauborgne, Joseph 117
MI-8 164
modular arithmetic 151–2
monoalphabetic cipher 72
Morse code 30–3, 92, 131, 144
music, codes in 158
National Security Agency (NSA), 164
nomenclator 88
null cipher 48, 97
nulls 66
number ciphers 79–80
number keys see word keys
one-time pad 117
Opish 44
optical telegraph 24–5

Painvin, Georges 131
password attacks 155–6
password sniffing 155
passwords 153–6

Personal Identification Number (PIN) 156
Pig Latin 43
pigpen cipher 85–6
Playfair cipher 118–19
Pliny 40
polyalphabetic cipher 72, 98, 112–13
Polybius square 81–2, 118, 122
Porta, Giavanni 38
Pretty Good Privacy (PGP) 152–3
public key enciphering 150
PURPLE code 140–1
quantum cryptography 157
quantum mechanics 157
rail fence cipher 64–6
Romanji 140
Room 40 165
Rosicrucian cipher, 86, 87
running key 116–17
Scovell, George 101
Secure Sockets Layer (SSL)
153

Semaphore 19–21

Shepherd’s Monument 124

Shift ciphers 71–3

Sign language 25–7

Signals Intelligence Service (SIS) 140

Slide rule cipher 73

Space code 53–5

Spoken codes 43–7

St Cyril cipher 72–3

Steganography 39, 40
stencil code 50
straddling checkerboard 125–6
stream cipher 147–8
sub rosa codes 56
substitution ciphers 71–89
symbols, using 82–3
Tacitus 39
texting 35–6
The Shadow magazine 74–5
transmission security 39
Transport Layer Security (TLS) 153

transposition ciphers 55, 63–71

transposition with a key 67

transposition, double 70–1

trifid cipher 122–3

Turing, Alan 137

typewriter cipher 77–8

Vernam cipher 117

Vernam, Gilbert 117, 147
Vigenère square 110–16
Wallis, John 161
wax tablets 38–9
Wheatstone, Charles 119
word breaks 112–13
word frequency 95
word keys 68–71
World War I codes 127–33
World War II codes 132, 133–39, 140–2, 164
World War II films and codes
139

XOR operation 147

Zimmerman, Phil 152

Zip’s law 95
If you enjoyed *Codes and Ciphers*, check out these other great Collins Gem titles.
Buy the ebook here
Buy the ebook here
Buy the ebook here
Collins Gem
COMPLEMENTARY THERAPIES

a mine of information
Buy the ebook here
Buy the ebook here
Disaster Survival

Be prepared for any eventuality
Buy the ebook here
HarperCollinsPublishers has made every reasonable effort to ensure that any picture content and written content in this ebook has been included or removed in accordance with the contractual and technological constraints in operation at the time of publication.
About the Publisher