Maths delivers!

A guide for teachers – Years 11 and 12

RSA Encryption

AMSI
AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

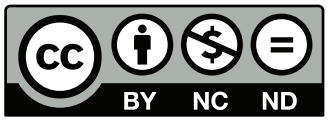
Maths delivers!
RSA Encryption

Dr Michael Evans AMSI

Editor: Dr Jane Pitkethly, La Trobe University

Illustrations and web design: Catherine Tan, Michael Shaw

For teachers of Secondary Mathematics



Australian Mathematical Sciences Institute
Building 161
The University of Melbourne
VIC 3010
Email: enquiries@amsi.org.au
Website: www.amsi.org.au

maths
delivers!

RSA encryption

Introduction

These notes accompany the video *Maths delivers! RSA encryption*. In this video, we see how encryption is used in defence, banking and internet transactions.

Encryption plays a crucial role in the day-to-day functioning of our society. For example, millions of people make purchases on the internet every day. Each time you submit your credit-card details online, there is a risk that this information may be stolen. So how can the information be sent securely?

A shopper's credit-card details need to be encrypted before they are transmitted over the internet, and so the method of encryption needs to be made public. But the method of decryption should be known only to the bank that is processing the payment.

For all of the ciphers in use before RSA, the methods of encryption and decryption were known to both the sender and the receiver of the message. With RSA, the instructions for how to encrypt a message can be made public, without compromising the security of the method of decryption. This was the big breakthrough that came with RSA encryption.

The Caesar cipher

We start with a discussion of a very simple method of encryption, the **Caesar cipher**, which is thought to have been used by Julius Caesar.

For example, if we choose a key of 1, then the letter A is concealed as B, the letter B is concealed as C, and so on. We view the alphabet as a loop. So, with a key of 1, the letter Z is concealed as A.

Number the letters of the alphabet in order from 0 to 25.

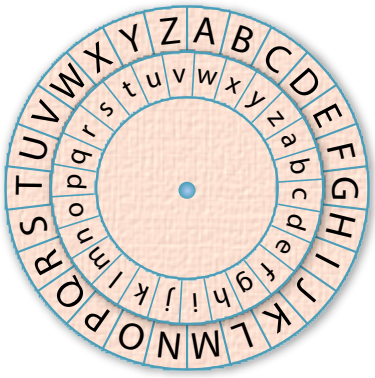
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Then using the key 1 means that we encrypt each letter by adding one.

If we use the Caesar cipher with key 22, then we encrypt each letter by adding 22. For example, since Q has number 16, we add 22 to obtain 38. But we want a number between 0 and 25 inclusive. Since $38 - 26 = 12$, the number 38 identifies the same place in the alphabet as the number 12, which is M. So we encrypt Q as M. The following table shows this for every letter of the alphabet.

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

We can also represent the Caesar cipher with key 22 using two concentric wheels, as shown in the following diagram. The inner wheel can be spun around for other keys.



Example

The following message has been encrypted using the Caesar cipher with key 22:

YHKOA PDA ZKKN.

Decrypt it.

Solution

Each letter has been encrypted by adding 22. So we decrypt by subtracting 22. (We will pretend that we haven't got the table above.)

- Y is numbered 24. Subtract 22 to obtain 2, which is C.
- H is numbered 7. Subtract 22 to obtain -15. Since $-15 + 26 = 11$, the number -15 identifies the same place in the alphabet as the number 11, which is L.

We continue in this way to obtain the original message: CLOSE THE DOOR.

Exercise 1

The Caesar cipher with key 15 has been used to encrypt a word as

PBHX.

What is the word?

Once we have introduced modular arithmetic, we will be able to explain Caesar ciphers in a different context.

How RSA encryption works

In order to be able to make online purchases securely, you need to be able to encrypt your credit-card details before transmitting them over the internet, so that only the authorised recipient can access these details. This means that we require a system where the method of encryption can be made public, while the method of decryption is kept secret.

This would not be possible using an early cryptosystem such as the Caesar cipher, where decryption is simply the opposite of encryption. Such a system is called a **symmetric cipher**, because encryption and decryption are symmetrical.

In an **asymmetric cipher**, there are two distinct keys: the public key and the private key. The public key is used for encryption, and the private key for decryption. The instructions for encrypting a message may be made public, without compromising the secrecy of the method of decryption. This is also called **public-key encryption**.

One of the first and most widely used algorithms for public-key encryption is **RSA**. The algorithm is named after its inventors Ron Rivest, Adi Shamir and Leonard Adleman, who published it in 1977 while working at MIT.

(From left) Ronald Rivest, Adi Shamir and Leonard Adelman. Photo by Steve Maller

The three originators of RSA encryption.

Prime numbers

RSA encryption is based on a special property of the prime numbers.

The prime numbers

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, \ldots$$

are natural numbers greater than 1 which cannot be expressed as a product of smaller natural numbers. That is, a **prime number** is a natural number greater than 1 whose only positive factors are itself and 1.

Each natural number greater than 1 can be factorised as a product of powers of primes. Moreover, if we ignore the order of the prime powers, then there is only one way to do this. For example, we can write

$$60 = 2^2 \times 3 \times 5.$$

There are infinitely many prime numbers. Using a computer, it is relatively easy to find lots of large prime numbers. At present, however, it is very difficult to find the prime factorisation of a very large number. This is what makes RSA encryption so hard to crack.

In the example we look at next, we work with very small prime numbers. This would be easy to crack.

A first look at RSA: How to encrypt and decrypt

We first go through the steps for producing the necessary numbers for encryption and decryption in RSA. We will then go through a simple example to understand how the processes of encryption and decryption are implemented. At this stage, we make no attempt to explain why it works, as we will first need a little bit of number theory.

Setting up

Step 1. Choose two primes p and q.

Example: $p = 3$ and $q = 11$.

Step 2. Let $n = p \times q$.

Example: $n = 3 \times 11 = 33$.

Step 3. Let $A = (p-1)(q-1)$.

Example: $A = 2 \times 10 = 20$.

Step 4. Choose an integer E with $1 < E < A$ such that E and A have no common factors other than 1.

Example: Choose $E = 7$.

Step 5. Find the integer D with $1 < D < A$ such that $D \times E - 1$ is a multiple of A.

Example: As $(3 \times 7) - 1 = 20$, we have $D = 3$.

The numbers n and E are the **public key**; they can be shared with anyone. The number D is the **private key**; it must be kept secret.

How do we use it?

Lily wants to be able to receive a secret message from Frank. She sends him the values of n and E. It doesn't matter if this message is intercepted; the values of n and E can be made public. Only Lily knows the value of D.

When data is sent securely over the internet, it is protected by a system like this. Anyone can encrypt, but only the authorised recipient can decrypt. Let us see how to employ it.

Before a message can be encrypted, it must be converted into a number (or a sequence of numbers). The number must be less than n. Since n is so small in our example, we will assign distinct numbers (each less than n) to the letters of the alphabet, and encrypt our message one letter at a time.

Suppose the message is 'HELP'. For simplicity, we suppose that the letters H, E, L and P were assigned the numbers 2, 3, 4 and 5, respectively.

H E L P

2 3 4 5

The next step is to raise each of these numbers to the power $E = 7$.

128 2187 16 384 78 125

Then find the remainders when each of these numbers is divided by $n = 33$. For example, we have $128 = 3 \times 33 + 29$, and so 128 gives a remainder of 29.

29 9 16 14

The encrypted message to be sent is 29 9 16 14.

The method for decryption is similar, but we use D instead of E. Start with the encrypted message.

29 9 16 14

Raise each number to the power $D = 3$.

24 389 729 4096 2744

Find the remainders when each number is divided by $n = 33$.

2 3 4 5

This is the original message.

Some security issues

There are two very big problems with this example. The first problem is the size of the primes. The number $n = 33$ is public. It is clear that our two primes must be 3 and 11, and so $A = 20$. Since $E = 7$ is also public, is it easy to work out the decryption key D.

This problem is overcome by using much larger primes. The two primes p and q should be chosen to have roughly the same size; in practice, they would each have over 150 decimal digits.

The second problem is that we have encrypted one letter at a time. This means, of course, that each letter will always have the same value once encrypted. This immediately makes the cipher easy to break by using a simple frequency count.

This problem is overcome by grouping letters together, and encrypting each group as a single number. For example, the ASCII codes for the letters H, E, L and P are 72, 69, 76 and 80, respectively. So, if n were large enough, we could convert the word 'HELP' into the single number 72 697 680.

Modular arithmetic

Modulo 24

We will first work with a 24-hour digital clock. If it is 6:00 am now and someone asks you what time it will be in 130 hours, it seems like a bit of work has to be done. First, you would see how many days there are in 130 hours. Since $24 \times 5 = 120$, there are 5 days and 10 hours left over. So the time will 16:00 on a 24-hour clock, which is 4:00 pm.

You can see we are really interested in the remainder when we divide by 24. We write

$$130 = 5 \times 24 + 10.$$

There is a notation for this which we will use throughout the remainder of these notes. We write

$$130 \equiv 10 \pmod{24}.$$

We read this as '130 is congruent to 10 modulo 24'.

There are clearly infinitely many numbers which are congruent to 10 modulo 24. We can form new ones by adding 24 to 10, and continuing to add 24 to each result: $10, 34, 58, \ldots$.

When we divide by 24, there are 24 possible remainders: $0, 1, 2, \ldots, 23$.

Modulo 12

The idea we have just introduced for 24-hour digital clocks can be adapted to 12-hour analogue clocks.

The clock shows 4 o'clock; the hour hand points to this position. Where will the hour hand point in 654 hours?

When we divide 654 by 12, the remainder is 6:

$$654 = 54 \times 12 + 6.$$

Using the same notation as above, we write

$$654 \equiv 6 \pmod{12}.$$

So the time on the clock in 654 hours will be the same as the time in 6 hours, which will be 10 o'clock. You can picture the hour hand as first completing 54 full revolutions, and then moving forward by 6 hours.

Modulo 7

Finally, as another familiar example, we look at days of the week. Number the days of the week from zero to six, starting from Sunday.

If today is Tuesday, then what day of the week will it be in 347 days? Since Tuesday corresponds to 2, we want to find the day of the week corresponding to $2 + 347 = 349$. We have

$$349 = 49 \times 7 + 6,$$

and so we can write

$$349 \equiv 6 \pmod{7}.$$

Thus in 347 days it will be Saturday.

Exercise 2

a It is 3:00 pm. What time will it be in 876 hours?

b It is Wednesday. What day of the week will it be in 1089 days?

Modulo n

In general, we can consider numbers modulo n, where n is any positive integer. This means that we divide a number by n, and just look at the remainder. The quotient is discarded. Each remainder has to be a whole number r such that $0 \leq r < n$.

For example,

$$23 \equiv 1 \pmod{11}.$$

This is read as '23 is congruent to 1 modulo 11'. When we divide by 11, there are 11 possible remainders: $0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$.

Here are some other examples:

$$80 \equiv 8 \pmod 9, \qquad 40 \equiv 2 \pmod{19}, \qquad 20 \equiv 2 \pmod 3.$$

The idea can be extended to the negative integers. For example, we divide -50 by 6 as follows:

$$-50 = -9 \times 6 + 4.$$

Here the remainder has to be a whole number r such that $0 \leq r < 6$. We can write

$$-50 \equiv 4 \pmod 6.$$

Congruence modulo n

The integers 18 and 58 are both congruent to 2 modulo 8, so we say they are congruent to each other:

$$58 \equiv 18 \pmod 8.$$

We could have checked this directly by subtracting 18 from 58: we have $58 - 18 = 40$, which is a multiple of 8.

In general, two integers a and b have the same remainder modulo n if and only if $a - b$ is a multiple of n. This leads us to the following definition.

Definition

Let n be a positive integer. For integers a and b, we write

$$a \equiv b \pmod n$$

if $a - b$ is a multiple of n.

For example:

$$96 \equiv 24 \pmod 8, \qquad 100 \equiv 28 \pmod 8, \qquad 2 \equiv 20 \pmod 3.$$

Theorem

Let n be a positive integer. For all integers a, b and c, we have

a $a \equiv a \pmod n$

b $a \equiv b \pmod n$ *implies* $b \equiv a \pmod n$

c *if* $a \equiv b \pmod n$ *and* $b \equiv c \pmod n$, *then* $a \equiv c \pmod n$.

Proof

We will only prove part c. Assume that $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$. Then, by definition,

$$a \equiv b \pmod{n} \iff a - b \text{ is a multiple of } n$$

$$b \equiv c \pmod{n} \iff b - c \text{ is a multiple of } n.$$

Thus there exist integers k_1 and k_2 such that $a - b = k_1 n$ and $b - c = k_2 n$. Therefore

$$\begin{aligned} a - c &= a - b + b - c \\ &= k_1 n + k_2 n \\ &= (k_1 + k_2) n. \end{aligned}$$

Hence, $a - c$ is a multiple of n. That is, $a \equiv c \pmod{n}$. □

The previous theorem says that congruence modulo n is an **equivalence relation** on the integers; it has the following three properties.

- **Reflexivity:** $a \equiv a \pmod{n}$.
- **Symmetry:** $a \equiv b \pmod{n}$ implies $b \equiv a \pmod{n}$.
- **Transitivity:** if $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$.

Theorem (Equivalence relation)

The relation $a \equiv b \pmod{n}$ is an equivalence relation on the set of integers. It has n distinct equivalence classes $[0], [1], \ldots, [n-1]$.

For $n = 5$, the equivalence classes are

$$[0] = \{\ldots, -10, -5, 0, 5, 10, \ldots\}$$

$$[1] = \{\ldots, -9, -4, 1, 6, 11, \ldots\}$$

$$[2] = \{\ldots, -8, -3, 2, 7, 12, \ldots\}$$

$$[3] = \{\ldots, -7, -2, 3, 8, 13, \ldots\}$$

$$[4] = \{\ldots, -6, -1, 4, 9, 14, \ldots\}.$$

Exercise 3

List the equivalence classes for $n = 7$.

There is further discussion of these equivalence classes in the *Appendix*.

Arithmetic modulo n

The following result is essential for the explanation of why RSA encryption works.

Theorem (Addition and multiplication modulo n)

If $a \equiv b \pmod{n}$ *and* $c \equiv d \pmod{n}$, *then*

a $a + c \equiv b + d \pmod{n}$

b $ac \equiv bc \pmod{n}$.

Proof

a Since $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, we know that n divides both $a - b$ and $c - d$. Since

$$(a + c) - (b + d) = (a - b) + (c - d),$$

it follows that n divides $(a + c) - (b + d)$. Hence, $a + c \equiv b + d \pmod{n}$.

b Again, we know that n divides both $a - b$ and $c - d$. Now

$$\begin{aligned} ac - bd &= ac - bc + bc - bd \\ &= (a - b)c + (c - d)b, \end{aligned}$$

and thus n divides $ac - bd$. Hence, $ac \equiv bc \pmod{n}$. □

Also, by a simple inductive argument, you can see that

$$a \equiv b \pmod{n} \implies a^m \equiv b^m \pmod{n},$$

for any positive integer m. This is very useful for computation, as the following example shows.

Example

Show that

1 $56^{100} \equiv 1 \pmod{5}$

2 $5^{100} \equiv 2 \pmod{7}$.

Solution

1 First note that $56 \equiv 1 \pmod{5}$. Thus $56^{100} \equiv 1^{100} \pmod{5}$, and so $56^{100} \equiv 1 \pmod{5}$.

2 Note that $5^2 \equiv 4 \pmod 7$, and so $5^4 \equiv 4^2 \pmod 7$. Continuing in this way, we have

$$5^4 \equiv 2 \pmod 7 \qquad \text{since } 4^2 = 16 \equiv 2 \pmod 7$$

$$5^{12} \equiv 2^3 \pmod 7$$

$$5^{12} \equiv 1 \pmod 7 \qquad \text{since } 2^3 = 8 \equiv 1 \pmod 7$$

$$5^{96} \equiv 1 \pmod 7 \qquad \text{since } 5^{96} = (5^{12})^8.$$

Hence, by combining the first and last equations above, we have

$$5^{100} = 5^{96} \times 5^4 \equiv 2 \pmod 7.$$

Exercise 4

Find a in each of the following, where a is a non-negative integer and as small as possible.

a $2468 \equiv a \pmod 7$

b $2^{242} \equiv a \pmod 3$

c $3^{245} \equiv a \pmod 5$

d $654^5 \equiv a \pmod{12}$.

Two non-zero integers a and b are **relatively prime** if their highest common factor is 1.

The following result will be very useful. It gives conditions under which we are allowed to cancel when doing modular arithmetic.

Theorem (Cancellation)

If $ab \equiv ac \pmod n$ and a is relatively prime to n, then $b \equiv c \pmod n$.

Proof

If $ab \equiv ac \pmod n$, then n divides $ab - ac = a(b-c)$. Since a and n are relatively prime, it follows that n divides $b-c$. Hence, $b \equiv c \pmod n$. □

The Caesar cipher using modular arithmetic

We will see how modular arithmetic can help us to understand the Caesar cipher and then to define new ciphers. Recall that we assign the numbers 0 to 25 to the letters of the alphabet. For the Caesar cipher with key 22, we encrypt a letter by adding 22 to its assigned number, dividing by 26 and considering the remainder, and then writing down the matching letter. This can be described as follows.

Let M be an element of $\{0,1,2,\ldots,25\}$. Then a new number in the set is assigned by

$$C \equiv M + 22 \pmod{26}.$$

We choose the number C which is in the set $\{0,1,2,\ldots,25\}$.

For example, we encrypt the letters D and P as follows:

- The letter D is assigned the number $M = 3$. Then

$$\begin{aligned} C &\equiv 3 + 22 \pmod{26} \\ &\equiv 25 \pmod{26}. \end{aligned}$$

  So D is encrypted as Z.
- The letter P is assigned the number $M = 15$. Then

$$\begin{aligned} C &\equiv 15 + 22 \pmod{26} \\ &\equiv 11 \pmod{26}. \end{aligned}$$

  So P is encrypted as L.

The decryption is obtained by subtracting 22:

$$M \equiv C - 22 \pmod{26}.$$

We can use this to recover the original letters in the two examples above:

- For $C = 25$, we have

$$\begin{aligned} M &\equiv 25 - 22 \pmod{26} \\ &\equiv 3 \pmod{26}. \end{aligned}$$

- For $C = 11$, we have

$$\begin{aligned} M &\equiv 11 - 22 \pmod{26} \\ &\equiv -11 \pmod{26} \\ &\equiv 15 \pmod{26}. \end{aligned}$$

We note again that the values of M and C are in the set $\{0,1,2,\ldots,25\}$.

Developing other ciphers using modular arithmetic

By making a careful choice, we can create a cipher using multiplication instead of addition. An example would be

$$C \equiv 5M \pmod{26}.$$

The number 5 is relatively prime to 26, with

$$5 \times 21 \equiv 1 \pmod{26}.$$

We can use this to encrypt and decrypt:

$$C \equiv 5M \pmod{26} \quad \text{and} \quad M \equiv 21C \pmod{26}.$$

To encrypt 17, we calculate

$$\begin{aligned} C &\equiv 5 \times 17 \pmod{26} \\ &\equiv 85 \pmod{26} \\ &\equiv 7 \pmod{26}. \end{aligned}$$

So 17 is encrypted as 7. To decrypt, we calculate

$$\begin{aligned} M &\equiv 21 \times 7 \pmod{26} \\ &\equiv 147 \pmod{26} \\ &\equiv 17 \pmod{26}. \end{aligned}$$

The reason this works is that 21 is the **multiplicative inverse** of 5 modulo 26. That is, if we are working modulo 26, then multiplication by 21 reverses (or undoes) multiplication by 5. This follows from the fact that

$$5 \times 21 \equiv 1 \pmod{26}.$$

Not every number has a multiplicative inverse modulo 26. That is why we needed to start from a number that is relatively prime to 26.

We can make the cipher more complicated by undertaking addition as well. For example, we can work with

$$C \equiv 5M + 1 \pmod{26} \quad \text{and} \quad M \equiv 21(C - 1) \pmod{26}.$$

Exercise 5

The number $M = 38$ is encrypted using

$$C \equiv 5M + 3 \pmod{49}.$$

Find the encrypted number C, and describe the method of decryption.

Two theorems

In order to understand why RSA works, we will need to use two basic theorems of number theory.

Fermat's little theorem

Pierre de Fermat was one of the greatest number theorists of all time. He was also one of the founders of probability theory. This famous French mathematician was an 'amateur' in that he practised law for a living.

Pierre de Fermat (1601–1665).

One of the most famous theorems in number theory is **Fermat's last theorem**, which states that, for any integer $n \geq 3$, the equation

$$x^n + y^n = z^n$$

has no solution with x, y and z all positive integers. This theorem was first conjectured by Fermat in 1637, in the margin of a copy of the text *Arithmetica*. No successful proof was published until 1995.

In this section, we deal with **Fermat's little theorem**, which is central to our discussion of RSA encryption. This theorem states that

$$a^p \equiv a \pmod{p},$$

for every prime p and integer a. Fermat stated this theorem in 1640, but did not supply a proof. There are several different ways to prove the theorem. We present one proof here, and a further proof in the *Appendix*.

The proof given in this section involves the binomial theorem. Recall that the binomial coefficients are defined by

$$\binom{n}{r} = \frac{n!}{(n-r)!\,r!},$$

for integers n and r with $0 \le r \le n$. For example,

$$\binom{7}{3} = \frac{7!}{4!3!} = \frac{7\times6\times5\times4\times3\times2\times1}{(4\times3\times2\times1)\times(3\times2\times1)} = \frac{7\times6\times5}{3\times2\times1} = 35.$$

Since the formula for the binomial coefficient must give an integer result, we know that all the numbers in the denominator will have to cancel out against the numbers in the numerator. Note in the formula for $\binom{7}{3}$ above that, since 7 is prime and greater than every number in the denominator, the 7 in the numerator will not be cancelled, and so the number $\binom{7}{3}$ must be divisible by 7. We use this idea to prove the following lemma.

Lemma

For any prime p, we have

$$(x+y)^p \equiv x^p + y^p \pmod{p}.$$

Proof

Let p be a prime, and consider the binomial coefficient

$$\binom{p}{i} = \frac{p!}{(p-i)!\,i!},$$

where $0 < i < p$. Then p occurs in the numerator and is greater than each number in the denominator (as $p > p-i$ and $p > i$). Since p is prime, it follows that $\binom{p}{i}$ is divisible by p. So $\binom{p}{i}$ is congruent to zero modulo p, for $0 < i < p$.

The binomial theorem states that

$$\begin{aligned}(x+y)^p &= \sum_{i=0}^{p}\binom{p}{i}x^{p-i}y^i \\ &= \binom{p}{0}x^p + \binom{p}{1}x^{p-1}y + \binom{p}{2}x^{p-2}y^2 + \cdots + \binom{p}{p-1}xy^{p-1} + \binom{p}{p}y^p.\end{aligned}$$

The middle terms (all except the first and last) are congruent to zero modulo p. Since $\binom{p}{0} = \binom{p}{p} = 1$, it follows that $(x+y)^p \equiv x^p + y^p \pmod{p}$. □

We now prove Fermat's little theorem for positive integers, which is all we will need for explaining RSA.

Theorem (Fermat's little theorem)

Let p be a prime. Then, for every positive integer a, we have

$$a^p \equiv a \pmod{p}.$$

Proof

The proof is by induction on a. If $a = 1$, then $a^p = 1^p = 1 = a$. So the result holds for $a = 1$.

Inductive step. Assume that

$$k^p \equiv k \pmod{p}$$

and consider $(k+1)^p$. By the previous lemma, we have

$$(k+1)^p \equiv k^p + 1 \pmod{p}.$$

Hence, by the induction assumption, we obtain

$$(k+1)^p \equiv k+1 \pmod{p}.$$

By the principle of induction, it follows that $a^p \equiv a \pmod{p}$, for every positive integer a. □

The Chinese remainder theorem

This theorem first appeared in the 3rd century in the book *The Mathematical Classic of Sun Zi*. We give the result for a special case which is sufficient for our purposes.

Theorem (Chinese remainder theorem)

Let m_1 and m_2 be two positive integers that are relatively prime. Given any two integers a and b, there exists an integer x such that

$$x \equiv a \pmod{m_1}$$

$$x \equiv b \pmod{m_2}.$$

Furthermore, any two solutions of these equations are congruent to each other modulo $m_1 m_2$.

Proof

Consider the numbers

$$a, \quad a+m_1, \quad a+2m_1, \quad \ldots, \quad a+(m_2-1)m_1.$$

Each of these numbers is congruent to a modulo m_1.

Suppose that two of these numbers are congruent to each other modulo m_2. We will show that this leads to a contradiction.

Let the two congruent numbers be $a+im_1$ and $a+jm_1$, where $0 \le i < j \le m_2-1$, so that

$$a+jm_1 \equiv a+im_1 \pmod{m_2}.$$

This implies that

$$jm_1 \equiv im_1 \pmod{m_2}.$$

Since m_1 and m_2 are relatively prime, it follows by the cancellation theorem that $j \equiv i \pmod{m_2}$. So m_2 divides $j-i$. But this is impossible, as $0 \le i < j \le m_2-1$.

We have shown that no two numbers from the list

$$a, \quad a+m_1, \quad a+2m_1, \quad \ldots, \quad a+(m_2-1)m_1.$$

are congruent to each other modulo m_2. That is, these numbers give distinct remainders on division by m_2. There are m_2 numbers in the list, and there are m_2 possible remainders. So we must obtain every possible remainder. In particular, there is an integer k with $0 \le k \le m_2-1$ such that

$$a+km_1 \equiv b \pmod{m_2}.$$

Now we can take $x = a+km_1$, and it is easy to see that x satisfies the two required equations.

To prove the final statement of the theorem, assume that both x_1 and x_2 are solutions to the equations. Let $z = x_1 - x_2$. Then

$$z \equiv 0 \pmod{m_1} \qquad \text{and} \qquad z \equiv 0 \pmod{m_2}.$$

Thus z is divisible by both m_1 and m_2. Since m_1 and m_2 are relatively prime, it follows that z is divisible by m_1m_2. Hence, $x_1 \equiv x_2 \pmod{m_1m_2}$. □

Why RSA encryption works

We start with another example to refresh our memories of the process. We again use small numbers to make the example easier to read, but of course this is not secure.

Example

Step 1. Choose two primes p and q. *Example:* $p = 5$ and $q = 13$.

Step 2. Let $n = pq$. *Example:* $n = 5 \times 13 = 65$.

Step 3. Let $A = (p-1)(q-1)$. *Example:* $A = 4 \times 12 = 48$.

Step 4. Choose an integer E with $1 < E < A$ such that E and A are relatively prime.

Example: Choose $E = 11$.

Step 5. Let D be the multiplicative inverse of E modulo A. That is, find the unique integer D with $1 < D < A$ such that $D \times E - 1$ is a multiple of A.

Example: As $(35 \times 11) - 1 = 384 = 8 \times 48$, we have $D = 35$.

To encrypt a number, we first raise it to the power $E = 11$, and then find its remainder when we divide by $n = 65$. For example, to encrypt 3, we find

$$3^{11} = 177\,147.$$

Since $177\,147 = 2725 \times 65 + 22$, the remainder is 22. So 3 encrypts as 22. We can write

$$3^{11} \equiv 22 \pmod{65}.$$

To decrypt, we take 22 to the power $D = 35$, and find its remainder when we divide by $n = 65$. We can use some of the properties of modular arithmetic which we developed earlier to help with this. We will break up 35 as $24 + 10 + 1$.

We have $22^2 \equiv 29 \pmod{65}$ and, conveniently, $29^6 \equiv 1 \pmod{65}$. So $22^{12} \equiv 1 \pmod{65}$ and therefore $22^{24} \equiv 1 \pmod{65}$.

Since $22^2 \equiv 29 \pmod{65}$ and $29^5 \equiv 9 \pmod{65}$, we have $22^{10} \equiv 9 \pmod{65}$.

Finally, we obtain

$$\begin{aligned} 22^{35} &= 22^{24} \times 22^{10} \times 22 \\ &\equiv 1 \times 9 \times 22 \pmod{65} \\ &\equiv 3 \pmod{65}. \end{aligned}$$

Public-key encryption

The public key is the pair (n, E). These two numbers may be published and shared with anyone. A plaintext message is a number M with $0 < M \leq n-1$. We encrypt M as the ciphertext message C by

$$C \equiv M^E \pmod{n},$$

where $0 \leq C \leq n-1$.

Private-key decryption

Your private key is the number D, and it must be kept secret. Decrypt the ciphertext C by

$$M \equiv C^D \pmod{n}.$$

The primes p and q must also be kept secret, since anyone who knows p, q and E can compute D. The secrecy of D relies on the fact that it is very difficult to factorise large numbers. We need to choose the primes p and q large enough so that, even though their product n is public, it is not computationally feasible to factorise n and discover p and q.

The mathematics of RSA

To understand why RSA works, we must understand why

$$M \equiv (M^E)^D \pmod{n},$$

where $0 \leq M \leq n-1$.

Here the numbers E, D and n are chosen according to steps 1 to 5 above, and the number M is being encrypted. Recall that D is chosen so that

$$ED \equiv 1 \pmod{A}.$$

Thus there exists an integer k such that

$$\begin{aligned} ED &= 1 + kA \\ &= 1 + k(p-1)(q-1). \end{aligned}$$

Therefore

$$\begin{aligned} (M^E)^D &= M^{ED} \\ &= M^{1+k(p-1)(q-1)} \\ &= M \times M^{k(p-1)(q-1)}. \end{aligned}$$

By Fermat's little theorem, we have

$$M \times M^{p-1} \equiv M \pmod{p}.$$

Applying this result $k(q-1)$ times gives

$$M^{ED} = M \times M^{k(p-1)(q-1)} \equiv M \pmod{p}.$$

By symmetry, we can also argue that

$$M^{ED} = M \times M^{k(p-1)(q-1)} \equiv M \pmod{q}.$$

Thus both $x = M$ and $x = M^{ED}$ are solutions to the system of congruences

$$x \equiv M \pmod{p}$$

$$x \equiv M \pmod{q}.$$

By the Chinese remainder theorem, the solution is unique modulo $pq = n$. So we can conclude that

$$M \equiv (M^E)^D \pmod{n}.$$

Code-breaking in World War II

Bletchley Park

Bletchley Park is in Buckinghamshire, England. During the Second World War, Bletchley Park was the site of the United Kingdom's main cryptanalysis establishment. Here several important ciphers were cracked, including those generated by the German Enigma and Lorenz machines.

The work carried out at Bletchley Park was of great importance. It is believed by some that this work shortened the war by two to four years, and that the outcome of the war would have been uncertain without it. Alan Turing was the leader of a group at Bletchley Park which cracked the Enigma cipher.

Alan Turing (1912–1954)

Alan Turing is known not only for his contribution to the breaking of the Enigma cipher at Bletchley Park, but also for his contribution to the foundation of computer science. In a paper in 1937, Alan Turing described a theoretical computing machine which would be able to compute according to any rules fed to it. Such a theoretical machine is now called a *Turing machine* and could, if given enough time and space, perform any calculation which could be described as algorithmic.

What was probably the first programmable electronic computer, named Colossus, became operational in 1943 at Bletchley Park.

Turing was found dead on 7 June 1954. A court determined that the cause of death was suicide.

Sydney University, T. G. Room and code-breaking in WW II

The following is the introduction to an article by Peter Donovan (UNSW) and John Mack (University of Sydney); see `www.maths.usyd.edu.au/u/ww2codes/gazette.html`.

Introduction

Thomas Gerald Room (1902–1986), FRS and FAA, was Professor of Mathematics at the University of Sydney from 1935 to 1968, succeeding H. S. Carslaw. He was one of the founders of the Australian Mathematical Society and one of its early Presidents, as well as being the first Editor of its Journal.

At the time of his death in 1986, very little factual information on military intelligence work in WW2 was available, except for some books and articles on Bletchley Park in the UK and its work on the German Enigma codes. In the obituary notice written for the Royal Society in 1987, the following extracts summarise what could be gleaned at that time:

> Arrangements were made early in 1941 for a small group at the University of Sydney to study Japanese codes. They were Room (as leader), his colleague Lyons and two members of the Greek Department, A. D. Trendall and A. P. Treweek. In mid-1941, the Australian government set up a cryptographic analysis unit at Victoria Barracks in Melbourne. Its job was to work on the deciphering of Japanese diplomatic codes. The Sydney group was recruited into this unit by an intelligence officer, Captain T. E. Nave, the University of Sydney having agreed to its secondment to the Defence Department. Room was a senior member of the unit. Later in 1941, he went to the British Far East Combined Bureau in Singapore to gain experience in British code-breaking methods. A letter to the Registrar of Sydney University, Walter Selle, on Christmas Eve indicates that he had been hard at work learning Japanese: 'The two terms I have spent under Miss Lake have proved as useful in my present job as the twenty years' mathematics!'
>
> ... the unit devised codes for the coast-watchers. Room was among those who took part in this work.

> When General MacArthur set up his headquarters in Brisbane in 1942, ... a joint signal intelligence section, called the Central Bureau, [was formed] in conjunction with it. Room was transferred to [it], where he worked until the end of the war on the decoding of Japanese military signals. ... His work was described as 'spectacularly successful'.

Appendix: Congruence classes

In this section, we follow on from the discussion of equivalence classes in the section *Modular arithmetic*.

Definition

Let n be a positive integer, and let a be an integer. The **congruence class** of a modulo n, denoted by $[a]$, is the set of all integers that are congruent to a modulo n. That is,

$$[a] = \{b \in \mathbb{Z} : b \equiv a \pmod{n}\}$$

$$= \{b \in \mathbb{Z} : b = a + kn, \text{ for some } k \in \mathbb{Z}\}.$$

Theorem

$a \equiv b \pmod{n}$ *if and only if* $[a] = [b]$.

Proof

Assume $a \equiv b \pmod{n}$. Let $c \in [a]$. Then $c \equiv a \pmod{n}$. By transitivity, we have $c \equiv b \pmod{n}$. Thus $c \in [b]$, and so we have shown that $[a] \subseteq [b]$. In the same way, it can be shown that $[b] \subseteq [a]$. Hence, $[a] = [b]$.

To prove the converse, assume $[a] = [b]$. By reflexivity, we have $a \equiv a \pmod{n}$, and therefore $a \in [a] = [b]$. Hence, $a \equiv b \pmod{n}$. □

It is easily seen that, when we write the list $[0], [1], [2], \ldots, [n-1]$, we are accounting for all congruence classes modulo n. Every integer belongs to exactly one of these congruence classes. We will denote this set of congruences classes by Z_n.

We recall the following theorem.

Theorem (Addition and multiplication modulo n)

If $a \equiv b \pmod{n}$ *and* $c \equiv d \pmod{n}$, *then*

a $a + c \equiv b + d \pmod{n}$

b $ac \equiv bc \pmod{n}$.

We can introduce the idea of addition and multiplication on the set Z_n of congruence classes to show that we have a **commutative ring**. This will not be done here in detail, but the following indicates the pathway to show this.

Addition is defined by

$$[a] + [b] = [a + b],$$

and multiplication by

$$[a] \times [b] = [ab].$$

It can be shown that both operations are well defined and satisfy the associative and commutative laws, and that the distributive law of multiplication over addition holds.

We can use congruence classes to give an alternative proof of Fermat's little theorem.

Theorem (Fermat's little theorem)

Let p be a prime. Then, for every integer a, we have

$$a^p \equiv a \pmod{p}.$$

Proof

Let Z_p be the set of congruence classes modulo p. We can write these congruence classes as $[0], [1], [2], \ldots, [p-1]$.

First assume that $[a] = [0]$. Then $a \equiv 0 \pmod{p}$, and so it follows easily that $a^p \equiv a \pmod{p}$. We can now assume that $[a] \neq [0]$. This implies that a and p are relatively prime.

Consider all products $[a][0], [a][1], [a][2], \ldots, [a][p-1]$, where multiplication in Z_p is defined as discussed above. Suppose that two of these products are equal. Then $[a][i] = [a][j]$, for some $0 \leq i < j < p$. This gives $ai \equiv aj \pmod{p}$ and so, by the cancellation theorem, we have $i \equiv j \pmod{p}$, which is a contradiction. We have shown that the products $[a][0], [a][1], [a][2], \ldots, [a][p-1]$ in Z_p are distinct.

Since $[a][0] = [0]$, it now follows that the products $[a][1], [a][2], \ldots, [a][p-1]$ must be $[1], [2], \ldots, [p-1]$ in some order. Since multiplication in Z_p is associative and commutative, this implies that

$$[a][1][a][2] \cdots [a][p-1] = [1][2] \cdots [p-1].$$

Using the definition of multiplication in Z_p, this gives

$$a^{p-1}(p-1)! \equiv (p-1)! \pmod{p}.$$

By the cancellation theorem, we can conclude that $a^{p-1} \equiv 1 \pmod{p}$. □

Answers to exercises

Exercise 1

The encrypted word is 'PBHX'. The corresponding numerical values are $15, 1, 7, 23$. We subtract 15 to obtain $0, -14, -8, 8$. These can be adjusted by adding 26 onto the negative numbers to obtain $0, 12, 18, 8$. So the word is 'AMSI'.

Exercise 2

a Using a 24-hour clock, 3:00 pm is 15:00. We have $15 + 876 = 891 = 37 \times 24 + 3$, and so

$$15 + 876 = 891 \equiv 3 \pmod{24}.$$

So 876 hours later it will be 3:00 am.

b Wednesday is assigned the number 3. We have

$$3 + 1089 = 1092 \equiv 0 \pmod{7}.$$

So 1089 days later is Sunday.

Exercise 3

For $n = 7$, the equivalence classes are

$$[0] = \{\ldots, -14, -7, 0, 7, 14, \ldots\}$$

$$[1] = \{\ldots, -13, -6, 1, 8, 15, \ldots\}$$

$$[2] = \{\ldots, -12, -5, 2, 9, 16, \ldots\}$$

$$[3] = \{\ldots, -11, -4, 3, 10, 17, \ldots\}$$

$$[4] = \{\ldots, -10, -3, 4, 11, 18, \ldots\}$$

$$[5] = \{\ldots, -9, -2, 5, 12, 19, \ldots\}$$

$$[6] = \{\ldots, -8, -1, 6, 13, 20, \ldots\}.$$

Exercise 4

a We have $2468 = 352 \times 7 + 4$. Hence,

$$2468 \equiv 4 \pmod{7},$$

and therefore $a = 4$.

b We have

$$\begin{aligned} 2^{242} &= (2^2)^{121} \\ &\equiv 1^{121} \pmod 3 \qquad \text{since } 2^2 \equiv 1 \pmod 3 \\ &\equiv 1 \pmod 3. \end{aligned}$$

Therefore $a = 1$.

c We have

$$\begin{aligned} 3^{245} &= 3^{4\times 61+1} \\ &\equiv 1^{61} \times 3 \pmod 5 \qquad \text{since } 3^4 \equiv 1 \pmod 5 \\ &\equiv 3 \pmod 5. \end{aligned}$$

Therefore $a = 3$.

d We have

$$\begin{aligned} 654^5 &\equiv 6^5 \pmod{12} \qquad \text{since } 654 \equiv 6 \pmod{12} \\ &\equiv 0 \pmod{12}, \end{aligned}$$

and so $a = 0$.

Exercise 5

Since we have

$$5 \times 38 + 3 = 193 \equiv 46 \pmod{49},$$

the number 38 is encrypted as 46.

For the decryption, we first note that

$$5 \times 10 \equiv 1 \pmod{49}.$$

Hence, the decryption is given by

$$M \equiv 10(C - 3) \pmod{49}.$$

Links and references

Links

- A BBC website on Bletchley Park:
  `www.bbc.co.uk/history/places/bletchley_park`
- The Australian Mathematics Trust website has an excellent biography of Alan Turing:
  `www.amt.edu.au/biogturing.html`
- A University of Sydney website devoted to code-breaking in WW II:
  `www.maths.usyd.edu.au/u/ww2codes/`

References

- P. Donovan and J. Mack, 'Sydney University, T. G. Room and codebreaking in WW II, Part I', *The Australian Mathematical Society Gazette* 29 (2002), 76–84.
- P. Donovan and J. Mack, 'Sydney University, T. G. Room and codebreaking in WW II, Part II', *The Australian Mathematical Society Gazette* 29 (2002), 141–148.
- A. Hodges, *Alan Turing: The Enigma*, Centenary edition, Princeton University Press, 2012.
- J. Humphreys and M. Prest, *Numbers, Groups and Codes*, Second edition, Cambridge University Press, 2004.
- N. Koblitz, *A Course in Number Theory and Cryptography*, Second edition, Springer, 1994.


0
1
2
3
4
5
6
7
8
9
10
11
12