Chapter 7

Cryptography and Group Theory

In this chapter we discuss one of the main methods of encrypting data, the RSA encryption system. The algebraic structure that is at the heart of this method is that of a group. Group theory, perhaps the first algebraic structure to be studied abstractly, is one of the most fundamental of structures. As we shall see, rings, fields, and vector spaces are all special examples of groups. What distinguishes groups from these other structures is that a group has only a single operation.

Cryptography is the subject of transmitting private data in a secure manner. If you make a purchase on the internet, you need to send to the merchant a credit card number. If somebody were to intercept the transmission of this information, they would have your number. Because of this, most internet sites encrypt such data. By doing so, anybody intercepting the transmission will see a useless string of digits instead of a valid credit card number. If, however, the interceptor were to know how the merchant replaces credit card numbers with other numbers, they would have a way of recovering the number. Because of this, merchants must use methods of encryption that are very difficult to “break”. We will discuss one such system, the RSA encryption system.

7.1 The RSA encryption system

The RSA encryption system was invented in 1978 by Ron Rivest, Adi Shamir, and Leonard Adleman, and is one of the most common methods of encrypting data used today. To describe the RSA system, one starts with the following data:

- distinct prime numbers $p$ and $q$
- an integer $e$ relatively prime to $(p - 1)(q - 1)$.

From this data we will build an encryption system. Let $n = pq$. We will restrict our attention to encrypting numbers. This is satisfactory, since any text message can be converted to numbers by replacing each letter with an appropriate number. Let $M$ be an integer,
considered to be a message we wish to encrypt. We then calculate \( M^e \mod n \), the remainder after dividing \( n \) into \( M^e \). This remainder is our encrypted message.

For example, let
\[
\begin{align*}
p &= 3486784409, \\
q &= 282429536483, \\
e &= 19.
\end{align*}
\]
Then \( n = pq = 984770904450021093547 \). Also, \((p - 1)(q - 1) = 984770904164104772656 \). To encrypt the message 12345, we calculate
\[
12345^{19} \mod n,
\]
which comes out to be 123355218486796132288. Therefore, if we wish to transmit the number 12345, we would instead transmit 123355218486796132288.

How does somebody receiving 123355218486796132288 know that this number represents 12345? First, by our assumption that \( e \) is relatively prime to \((p - 1)(q - 1)\), we know that \( e \) has an inverse modulo \((p - 1)(q - 1)\); that is, there is an integer \( d \) with \( ed \equiv 1 \mod (p - 1)(q - 1) \). If an encrypted number \( N \) is received, then one calculates \( N^d \mod n \); the result returns the original message. For example, from a Maple computation, we can see that
\[
d = 207320190350337846875.
\]
Thus, to recover the original message 12345, we compute
\[
123355218486796132288^{207320190350337846875} \mod 984770904450021093547 = 12345
\]
While this calculation looks formidable, Maple can do it virtually instantaneously. In fact, on an average personal computer, Maple can calculate \( M^d \mod n \) in a couple of seconds even if \( d \) and \( n \) are 400 digit numbers, so the calculations in the RSA system are easy to do even with very large numbers.

To summarize, the RSA encryption system starts with two prime numbers \( p \) and \( q \) and an integer \( e \) satisfying \( \gcd(e, (p - 1)(q - 1)) = 1 \). One then calculates a positive integer \( d \) satisfying \( ed \equiv 1 \mod (p - 1)(q - 1) \). One then encrypts an integer \( M \) by replacing it by
\[
N = M^e \mod n.
\]
To decrypt \( N \), one sees that
\[
M = N^d \mod n.
\]
Why this method works will be addressed in our study of group theory. The theoretical fact that the inventors of the RSA system, Rivest, Shamir, and Adleman, needed was Euler's theorem. This is a special case of Lagrange’s theorem of group theory. We will introduce groups, see several examples, and prove Lagrange’s theorem in the next section. To prove this theorem we will revisit cosets, a subject we have studied for both vector spaces and for rings.
7.2 Groups

To motivate the definition of a group, we discuss the main example used in the RSA encryption system. Let $R$ be a ring. Recall that a unit of $R$ is an element having a multiplicative inverse. Recall also that if $a, b$ are units of $R$, then so is $ab$, since $ab$ has $b^{-1}a^{-1}$ as its multiplicative inverse. We denote by $R^*$ the set of all units of $R$. In other words,

$$R^* = \{a \in R : \text{there is a } c \in R \text{ with } ac = ca = 1\}.$$ 

By the statement above, if we multiply two elements of $R^*$, the result is another element of $R^*$. Therefore, multiplication induces a binary operation on the set $R^*$. We note three properties of this binary operation: multiplication on $R^*$ is associative; $1 \in R^*$, so $R^*$ has an identity; and each element of $R^*$ has an inverse in $R^*$. It is these properties that make up the definition of a group.

**Definition 7.1.** Let $G$ be a nonempty set together with a binary operation $*$ on $G$. Then the pair $(G, *)$ is said to be a group if

1. $a * (b * c) = (a * b) * c$ for all $a, b, c \in G$;
2. there is an $e \in G$ such that $e * a = a * e = a$ for all $a \in G$;
3. for each $a \in G$ there is an element $b \in G$ with $a * b = b * a = e$.

We give several examples of groups.

**Example 7.2.** Let $R$ be a ring. Then, as noted above, $(R^*, \cdot)$ is a group. The identity of $R^*$ is the multiplicative identity $1$.

**Example 7.3.** Let $R$ be a ring. Then $(R, +)$ is a group. To see this, we remark that the definition of a ring includes the three properties of a group when considering addition. Note that the identity of $(R, +)$ is $0$, the additive identity of $R$.

**Example 7.4.** Let $V$ be a vector space. Ignoring scalar multiplication and only considering addition, we see that $(V, +)$ is a group. Again, looking at the definition of a vector space and only considering addition, we have the properties in the definition of a group.

**Example 7.5.** For some special cases of the examples above, the following are groups under addition: $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, and $(\mathbb{R}, +)$. Furthermore, under multiplication, $(\mathbb{Z}^*, \cdot)$, $(\mathbb{Q}^*, \cdot)$, and $(\mathbb{R}^*, \cdot)$ are all groups. To better understand these examples, we see that $\mathbb{Z}^* = \{1, -1\}$; the only elements of $\mathbb{Z}$ that have multiplicative inverses in $\mathbb{Z}$ are $1$ and $-1$. However, since $\mathbb{Q}$ and $\mathbb{R}$ are fields, we have $\mathbb{Q}^* = \mathbb{Q} - \{0\}$ and $\mathbb{R}^* = \mathbb{R} - \{0\}$. That is, every nonzero element of $\mathbb{Q}$ or of $\mathbb{R}$ has a multiplicative inverse (in $\mathbb{Q}$ or $\mathbb{R}$, respectively). This leads to the next example.
Example 7.6. Let $F$ be a field. Then $(F - \{0\}, \cdot)$ is a group under multiplication. This is because $F$ is a ring and $F^* = F - \{0\}$ by the definition of a field.

Example 7.7. Let $R = M_n(\mathbb{R})$, the ring of all $n \times n$ matrices with real number entries. Then $R^*$ is usually denoted by $\text{GL}_n(\mathbb{R})$, and is called the general linear group of $n \times n$ matrices. By recalling properties of determinants, we may write

$$\text{GL}_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det(A) \neq 0\}.$$ 

This is a group under matrix multiplication. Note that multiplication in this group is not commutative.

Example 7.8. Let $X$ be any set. We denote by $P(X)$ the set of all permutations of $X$. That is,

$$P(X) = \{f : X \to X : f \text{ is a 1-1, onto function}\}.$$ 

We define an operation on $P(X)$ by composition of functions; note that if $f, g$ are functions from $X$ to $X$, then so is $f \circ g$, and that if $f$ and $g$ are both 1-1 and onto, then so is $f \circ g$. Therefore, composition is indeed a binary operation on $P(X)$. We leave it as an exercise to show that $P(X)$ is a group under composition.

We now focus on the example needed for the RSA system. Let $n$ be a positive integer, and consider the ring $\mathbb{Z}_n$. We can then form the group $\mathbb{Z}_n^*$, a group under coset multiplication. By Corollary ??, if $\overline{a} \in \mathbb{Z}_n$, then there is a solution to the equation $\overline{a} \cdot \overline{x} = \overline{1}$ if $\gcd(a, n) = 1$. That is, $\overline{a}$ has a multiplicative inverse in $\mathbb{Z}_n$ if $\gcd(a, n) = 1$. The converse is also true; if $\gcd(a, n) = 1$, then we may write $1 = ax + by$ for some integers $x, y$, by Proposition ??. Then $\overline{1} = \overline{a} \cdot \overline{x}$, so $\overline{a}$ has a multiplicative inverse. We thus have the following description:

$$\mathbb{Z}_n^* = \{\overline{a} \in \mathbb{Z}_n : \gcd(a, n) = 1\}.$$ 

How many elements does this group have? For example, $\mathbb{Z}_4^* = \{1, 3\}$ has two elements, $\mathbb{Z}_6^* = \{1, 5\}$ also has two elements, and $\mathbb{Z}_{11}^* = \mathbb{Z}_{11} - \{0\}$ has ten elements. In general, the number of elements of $\mathbb{Z}_n^*$ is equal to the number of integers $a$ satisfying $1 \leq a \leq n$ and $\gcd(a, n) = 1$. This number is referred to as $\phi(n)$, and the function $\phi$ is called the Euler phi function. While there is a general formula for computing $\phi(n)$ from the prime factorization of $n$, we concern ourselves only with two cases. First, if $p$ is a prime, then $\phi(p) = p - 1$. This is clear since all integers $a$ with $1 \leq a \leq p$ except for $p$ itself are relatively prime to $p$. The next case we state as a lemma.

Lemma 7.9. Let $p$ and $q$ be distinct primes. Then $\phi(pq) = (p - 1)(q - 1)$.

Proof. Set $n = pq$. To count $\phi(n)$, we first count the number of integers between 1 and $n$ and not relatively prime to $n$. If $1 \leq a \leq n$, then $\gcd(a, n) > 1$ only if $p$ divides $a$ or $q$ divides $a$. The multiples of $p$ between 1 and $n$ are then

$$p, 2p, \ldots, (q - 1)p, qp = n,$$
so there are \( q \) multiples of \( p \) between 1 and \( n \). The multiples of \( q \) in this range are

\[ q, 2q, \ldots, (p - 1)q, pq = n, \]

so there are \( p \) multiples of \( q \) in this range. The only number on both lists is \( n \); this follows from unique factorization. Therefore, there are \( p + q - 1 \) integers between 1 and \( n \) that are not relatively prime to \( n \). Since there are \( n = pq \) numbers total in this range, we see that

\[
\phi(n) = pq - (p + q - 1) = pq - p - q + 1 \\
= p(q - 1) - (q - 1) = (p - 1)(q - 1),
\]

as desired.

To summarize, the group \( \mathbb{Z}_n^* \) has \( \phi(n) \) elements, and if \( n = pq \) is the product of two distinct primes, then \( \mathbb{Z}_{pq}^* \) has \( (p - 1)(q - 1) \) elements. The significance of this result and its application to the RSA encryption system will become clear when we prove Lagrange’s theorem. To do this, we first need to discuss subgroups. This concept is the analogue in group theory of subspaces of a vector space.

Before going further, we make a note about notation and terminology. Some of our examples of groups use addition as the operation and other examples use multiplication. Unless we have a specific group in which the operation is addition, we will use multiplicative notation for the operation. In particular, we will write \( ab \) or \( a \cdot b \) for the product of \( a \) and \( b \) and \( a^{-1} \) for the inverse of \( a \). As for terminology, unless the symbol for the operation needs to be specified, we will refer to a group by a single symbol, such as \( G \), rather than by a pair such as \( (G, \ast) \). We need to be careful to remember that a group is not just a set but is a set together with a binary operation.

**Definition 7.10.** Let \( G \) be a group. A nonempty subset \( H \) of \( G \) is said to be a subgroup of \( G \) if the operation on \( G \) restricts to an operation on \( H \), and if \( H \) is a group with respect to this restricted operation.

For example, consider the group \( \mathbb{Z} \) under addition and let \( H \) be the set of even integers. Addition restricts to an operation on \( H \) because the sum of two even integers is again even. Just as there is a theorem helping us determine when a subset of a vector space is a subspace (Lemma ???), there is a result that helps us determine when a subset of a group is a subgroup.

**Lemma 7.11.** Let \( G \) be a group, and let \( H \) be a nonempty subset of \( G \). Then \( H \) is a subgroup of \( G \) provided that the following two conditions hold: (i) if \( a, b \in H \), then \( ab \in H \), and (ii) if \( a \in H \), then \( a^{-1} \in H \).

**Proof.** Suppose a subset \( H \) of a group \( G \) satisfies the two conditions in the statement. The first says that the operation on \( G \) restricts to an operation on \( H \), so we have a binary operation on \( H \). We need to verify for \( H \) the three axioms in the definition of a group. Associativity is clear; if \( a, b, c \in H \), then \( a, b, c \in G \), so \( a(bc) = (ab)c \) since \( G \) is a group.
Next, Condition (ii) ensures that every element of $H$ has an inverse in $H$. So, the only thing remaining is to see that $H$ has an identity. We do this by proving that if $e$ is the identity of $G$, then $e \in H$. To see why this is true, first note that there is an element $a \in H$ because $H$ is nonempty. By Condition (ii), $a^{-1} \in H$. Then, by Condition (i), $a \cdot a^{-1} \in H$. But $a \cdot a^{-1} = e$, so $e \in H$.

There is a particularly nice construction of subgroups. Let $G$ be a group and let $a \in G$. Consider trying to build a subgroup containing $a$. By Condition (i) of the Lemma above, we see that the subgroup has to contain $a \cdot a = a^2$. Using the condition again, the subgroup has to contain $a^2 \cdot a = a^3$. The subgroup must also contain $a^{-1}$. Then by condition (i), it must contain $a^{-1} \cdot a^{-1} = a^{-2}$. Continuing with this idea leads us to the following definition. If $G$ is a group and $a \in G$, then the cyclic subgroup generated by $a$ is the set $\langle a \rangle = \{a^n : n \in \mathbb{Z}\}$.

To formally define $a^n$, we first define $a^0 = e$, the identity of $G$. If $n$ is a positive integer, then we define, inductively, $a^{n+1} = a^n \cdot a$. Therefore, $a^1 = a^0 \cdot a = e \cdot a = a$, and $a^2 = a \cdot a$, and so on. For negative exponents, if $n > 0$, we set $a^{-n} = (a^n)^{-1}$. The following laws of exponents are consequences of the definition of a group (these are left as homework problems): $a^n \cdot a^m = a^{n+m}$, $(a^n)^m = a^{nm}$, and $(a^n)^{-1} = (a^{-1})^n$ for any $n, m \in \mathbb{Z}$. From these it follows that $\langle a \rangle$ is a subgroup of $G$ by the lemma.

**Example 7.12.** Let $G = \mathbb{Z}^*_8$. Then $G = \{1, 3, 5, 7\}$. We calculate the cyclic subgroups of $G$. First, since $1^3 = 1$ for all $n$, we have $\langle 1 \rangle = \{1\}$ contains only the identity $1$. Next, for $3$, we see that

- $3^1 = 3$,
- $3^2 = 9 = 1$,
- $3^3 = 27 = 3$,
- $3^4 = 81 = 1$.

You may recognize a pattern. Since $3^2 = 1$, the identity of $G$, if $n = 2m$ is even, then $3^n = 3^{2m} = (3^2)^m = 1^m = 1$. If $n = 2m + 1$ is odd, then

$$3^n = 3^{2m+1} = 3 \cdot 3^{2m} = 3 \cdot 1 = 3.$$  

Therefore, the only powers of 3 are 1 and 3. Thus, $\langle 3 \rangle = \{1, 3\}$. If we do similar calculations for $5$ and $7$, we will see similar patterns, for $5^2 = 1$ and $7^2 = 1$. Thus, the only powers of 5 will be 1 and 5, and the only powers of 7 are 1 and 7.

**Example 7.13.** Let $G = \mathbb{Z}^*_10$. Then $G = \{1, 3, 7, 9\}$ has four elements. As with the previous example (and any group in fact), the cyclic subgroup generated by 1 is just $\{1\}$.
Next consider $3$. We have
\[
\begin{align*}
3^1 &= 3, \\
3^2 &= 9, \\
3^3 &= 27 = 7, \\
3^4 &= 81 = 1.
\end{align*}
\]
We have therefore produced all four elements of $G$ as powers of $3$; therefore, $\langle 3 \rangle = G$. A similar calculation will show that $\langle 7 \rangle = G$; this can also be seen by the fact that $7 = 3^{-1}$ and $\langle a \rangle = \langle a^{-1} \rangle$ for any group $G$ and any $a \in G$. Finally, for $9$, we have
\[
\begin{align*}
9^1 &= 9, \\
9^2 &= 81 = 1.
\end{align*}
\]
As with the previous example, even powers of $9$ are $1$ and odd powers of $9$ are $9$, so $\langle 9 \rangle = \{1, 9\}$.

We make some numerical observations from the previous two examples. The first thing to notice is that the number of elements in a given cyclic subgroup $\langle a \rangle$ turned out to be the same as the first positive number $n$ for which $a^n$ was the identity. Second, the number of elements in each cyclic subgroup was a divisor of the number of elements of the given group. Both of these facts are not coincidences; they are general facts that we will now prove. The second is in fact Lagrange’s theorem. For a piece of notation, we will write $\vert X \vert$ for the number of elements in a set $X$. For a group $G$, the number $\vert G \vert$ is often called the order of $G$.

**Lemma 7.14.** Let $G$ be a finite group and let $a \in G$. If $n = \min \{m : m > 0, a^m = e\}$, then $n = \vert \langle a \rangle \vert$, the number of elements in the cyclic subgroup generated by $a$.

**Proof.** We will prove the lemma by proving that $\langle a \rangle = \{a^r : 0 \leq r < n\}$ and that these elements are all distinct. First, any element of $\langle a \rangle$ is of the form $a^s$ for some integer $s$. By the division algorithm, we may write $s = qn + r$ with $0 \leq r < n$. Then
\[
a^s = a^{qn+r} = a^m a^r = (a^n)^q (a^r) = a^r
\]
since $a^n = e$, so $(a^n)^q = e$. Therefore, $a^s$ can be written as a power $a^r$ of $a$ with $0 \leq r < n$. This proves the first claim. For the second, suppose that $a^r = a^t$ with $0 \leq r, t < n$. Suppose that $r \leq t$. Then, by the laws of exponents, $e = a^t a^{-r} = a^{t-r}$. Since $n$ is the smallest positive integer satisfying $a^m = e$, and since $0 \leq t - r < n$, we must have $t - r = 0$. Thus, $t = r$. So, the elements $a^0, a^1, \ldots, a^{n-1}$ are all distinct. Since they form $\langle a \rangle$, we have proved that $\vert \langle a \rangle \vert = n$. \qed
We now consider Lagrange’s theorem. To do so we need to look at a concept we have
seen twice before in the case of vector spaces and for rings. If \( H \) is a subgroup of a group
\((G, \ast)\), and if \( a \in G \), then the coset of \( H \) generated by \( a \) is the set
\[
H \ast a = \{ h \ast a : h \in H \}.
\]
Cosets are equivalence classes for the following equivalence relation: for \( a, b \in G \), define
\( a \sim b \) if \( ab^{-1} \in H \). Then \( \sim \) is an equivalence relation (see homework problems) and the
equivalence class of \( a \) is the coset \( Ha \). Therefore, the cosets of \( H \) form a partition for the
group \( G \). If we write the group operation as multiplication, we will usually write
\( Ha \) for the coset of \( a \).

**Theorem 7.15 (Lagrange).** Let \( G \) be a finite group and let \( H \) be a subgroup of \( G \). Then
\( |H| \) divides \( |G| \).

**Proof.** We prove this by showing that each coset has \( |H| \) elements. From this it will follow
that \( |G| \) is equal to \( |H| \) times the number of cosets, and this will prove the theorem. To do
this, let \( a \in G \). We wish to prove that \( |Ha| = |H| \). One way to prove that two sets have
the same size is to produce a 1-1 onto function between them. We do this here by defining
a function \( f : H \rightarrow Ha \) by \( f(h) = ha \). This is 1-1 since if \( f(h) = f(k) \), then \( ha = ka \).
Multiplying both sides on the right by \( a^{-1} \) yields \( h = k \), so \( f \) is 1-1. The function \( f \) is also
onto since, if \( x \in Ha \), then \( x = ha \) for some \( h \in H \), and so \( x = f(h) \). Since \( f \) is then a 1-1
onto function from \( H \) to \( Ha \), we have proven that \( |H| = |Ha| \), as desired.

We can combine Lagrange’s theorem with the previous lemma to get a result key in the
RSA encryption system.

**Corollary 7.16.** Let \( G \) be a finite group with \( n = |G| \). If \( a \in G \), then \( a^n = e \).

**Proof.** Let \( m = |\langle a \rangle| \). By Lagrange’s theorem, \( m \) divides \( n \), so \( n = mt \) for some integer \( t \).
By the lemma, \( a^m = e \). Therefore, \( a^n = a^{mt} = (a^m)^t = e^t = e \), as desired.

As a special case of this corollary, we obtain Euler’s theorem.

**Corollary 7.17 (Euler’s Theorem).** Let \( n \) be a positive integer. If \( a \) is an integer with
gcd\((a, n) = 1\), then \( a^{\phi(n)} \equiv 1 \mod n \).

**Proof.** If gcd\((a, n) = 1\), then \( a \in \mathbb{Z}_n^* \), a group of order \( \phi(n) \). The previous corollary tells us
that \( a^{\phi(n)} = 1 \). By definition of coset multiplication, \( a^{\phi(n)} = a^\phi(n) \). The equation \( a^{\phi(n)} = 1 \) is
equivalent to the relation \( a^{\phi(n)} \equiv 1 \mod n \).

We are now in a position to see how group theory will tell us that the method of decrypting
in the RSA system recovers the original message. Let \( G = \mathbb{Z}_n^* \), where \( n = pq \) is the product
of two distinct prime numbers. As we have seen, \( |G| = \phi(n) = (p-1)(q-1) \). We have
an integer \( e \) satisfying \( \gcd(e, (p-1)(q-1)) = 1 \). Therefore, there is an integer \( d \) satisfying
$ed \equiv 1 \mod \phi(n)$. We may write $1 = ed + s\phi(n)$ for some integer $s$. The claim of the RSA system is that, for any message $M$, we have $(M^e)^d \mod n = 1$. Written another way, it claims that $(\overline{M})^d = \overline{M}$. Assuming that $M$ is not divisible by $p$ or $q$, we have $M \in \mathbb{Z}_n^*$. Therefore

$$\overline{M} = \overline{M}^{ed + s\phi(n)} = \overline{M}^{ed} \overline{M}^{s\phi(n)} = \overline{M}^{ed} (\overline{M}^{\phi(n)})^s = \overline{M}^{ed}$$

since $\overline{M}^{\phi(n)} = 1$ by the corollary to Lagrange’s theorem. Thus, $(\overline{M}^e)^d = \overline{M}$, and so the decryption in RSA recovers the original message.

In the argument above, we assumed that $M$ was not divisible by either $p$ or $q$ in order to conclude that decryption would recover $M$. This is not a necessary assumption, but it makes the argument a little simpler.

### 7.3 Secure Signatures with RSA

One issue of data transmission is the ability to verify a person’s identity. If I send a request to a bank to transfer money out of an account, the bank wants to know if I am the owner of the account. If I make the request over the internet, how can the bank check my identity? The RSA encryption system gives a method for checking identities, which is one of the important features of the system.

Suppose that person $A$ transmits data to person $B$, and that person $B$ wants a method to check the identity of person $A$. To do this, both person $A$ and $B$ get sets of RSA data; person $A$ has a modulus $n_A$ and an encryption exponent $e_A$. These are publicly available. That person also has a decryption exponent $d_A$ that remains private. Person $B$ similarly has data $n_B$, $e_B$, and $d_B$. In addition, person $A$ has a signature, a publicly available number $S$. To convince person $B$ of his identity, person $A$ first calculates $T = S^{d_A} \mod n_A$ and then $R = T^{e_B} \mod n_B$. He then transmits $R$ to person $B$. Person $B$ then decrypts $R$ with her data, recovering $T = R^{d_B} \mod n_B$. Finally, she encrypts $T$ with person $A$’s data, obtaining $T^{e_A} \mod n_A = S$. By seeing that this result is the signature of person $A$, the identity has been validated.

For example, suppose that the data for person $A$ is

- $n_A = 2673157$
- $e_A = 23$
- $d_A = 2437607$
- $S = 837361$
and the data for person $B$ is

\begin{align*}
n_B &= 721864639 \\
e_B &= 19823 \\
d_B &= 700322447
\end{align*}

Person $A$ then calculates

\begin{align*}
837361^{2437607} \mod 2673157 &= 1216606,
\end{align*}

and then

\begin{align*}
1216606^{19823} \mod 721864639 &= 241279367.
\end{align*}

Person $A$ then transmits $241279367$ to person $B$. When person $B$ receives this, she calculates

\begin{align*}
241279367^{700322447} \mod 721864639 &= 1216606,
\end{align*}

and finally recovers $S$ as $S = 1216606^{23} \mod 2673157$.

To explain why this works, we denote by $\text{encrypt}_A(M)$ and $\text{decrypt}_B(M)$ the integers $M^{e_A} \mod n_A$ and $M^{d_A} \mod n_A$, respectively. We similarly have $\text{encrypt}_B(M)$ and $\text{decrypt}_B(M)$.

The validity of the RSA system says that

\begin{align*}
\text{decrypt}_A(\text{encrypt}_A(M)) &= M, \\
\text{encrypt}_A(\text{decrypt}_A(M)) &= M.
\end{align*}

Similar equations hold for $B$. With this notation, person $A$ calculates

\begin{align*}
R &= \text{encrypt}_B(\text{decrypt}_A(S))
\end{align*}

and then person $B$ calculates

\begin{align*}
\text{encrypt}_A(\text{decrypt}_B(R)).
\end{align*}

Therefore, person $B$ will calculate

\begin{align*}
\text{encrypt}_A(\text{decrypt}_B(\text{encrypt}_B(\text{decrypt}_A(S)))) &= \text{encrypt}_A(\text{decrypt}_A(S)) = S
\end{align*}

because of the equations above. Therefore, person $B$ does recover the signature of person $A$.

The reason that this method validates the identity of person $A$ is because only person $A$ can calculate $\text{decrypt}_A(S)$. If another person tries to claim he is person $A$, tries to substitute a number $F$ in place of $\text{decrypt}_A(S)$, he will transmit $\text{encrypt}_B(F)$ to person $B$. Person $B$ will then calculate

\begin{align*}
\text{encrypt}_A(\text{decrypt}_B(\text{encrypt}_B(F))) &= \text{encrypt}_A(F).
\end{align*}

However, in order to have $\text{encrypt}_A(F) = S$, we must have

\begin{align*}
\text{decrypt}_A(S) &= \text{decrypt}_A(\text{encrypt}_A(F)) = F,
\end{align*}

which means that this person has to have the correct decrypted number $\text{decrypt}_A(S)$; he cannot send any other number without person $B$ realizing it is a fake number.