Notes on the RSA Algorithm, and HW assignments #10, #11, and #12

These notes concern the "RSA," or Rivest-Shamir-Adelman, algorithm, which is a method for encoding and decoding messages using some number theory ideas.

The Exercises at the ends of Parts A, B, and C of these Notes constitute HW #10. The Exercises at the ends of Parts D and E of these Notes constitute HW #11, and those at the end of Part F constitute HW #12.

Throughout, by "prime," or "prime number," we will mean a *positive* integer p whose only positive integer factors are 1 and p.

The RSA algorithm is based upon the simple idea that, while multiplying together two large primes is relatively easy, *factoring* such a product is much harder.

More specifically: given a pair of large primes p and q, a decent computer can, in general, calculate $m = pq$ quite easily, even if p and q have hundreds of digits. But given only the product m of two such prime numbers, it's generally *not so easy*, even with *lots* of computing power, to figure out of which two primes m is a product (even with the advance knowledge that m *is*, in fact, a product of *some* pair of primes).

For example, Mathematica 13.3.1.0, running on an M1 iMac, took 0.000012 seconds of CPU time to multiply together the primes

$p =$28,012,569,795,147,037,305,920,963,277,749,628,914,662,527,590,314,892,381,540,899,557,658,
727, 561,627,073,596,629,516,007,733,350,970,901,196,381,503,333,712,077,626,705,499,954,515,
577,260,792,348,632,533,889,368,689,260,551

and

$q =$409,979,012,803,156,684,026,992,824,311,225,162,850,617,662,647, 990,082,269,707,895,322,401,
233,158,338,554,223,937,364,652,604,454,924,195,546,130,462,715,574,033,228,042,504,577,902,
809,413,850,720,086,027,157,221,973,957,611,016,318,502,032,623,823.

(Both p and q *are*, in fact, prime.) It's been working on factoring $m = pq$ since approximately 10 AM on Tuesday, November 12, and is not likely to succeed. You'll be notified if it does. (It won't.)

Here is how RSA works.

Part A: Encoding. We start with a message; we'll assume, for the sake of simplicity, that the message consists only of upper-case English letters A, B, . . . , Z. We encode our message as follows.

1. First, we convert the message to a natural number n. To do so, we'll use this "numerization key:"

A	B	C	D	E	F	G	H	I	J	K	L	M
11	12	13	14	15	16	17	18	19	20	21	22	23
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
24	25	26	27	28	29	30	31	32	33	34	35	36

For example, the message "HI" would become the integer $n = 1819$.

Remark: A more complex message might involve digits, lower-case letters, punctuation, spaces, etc. We could numerize those symbols as well; there are plenty of two-digit numbers left to numerize symbols with. But to keep things simple, we'll assume, again, that we're working only with the symbols A through Z.

Also, the numerization code is easily cracked; it should not be considered a fundamental part of the RSA encryption scheme. Rather, it's just a simple way of putting everything into the form of an integer, so that we may perform the "integer arithmetic" to be described below.

2. Next, we raise the natural number n to another natural number k. For example, if $n = 1819$ as above, and we choose $k = 17$, then

$$n^k = 1819^{17} = 26{,}130{,}991{,}223{,}692{,}189{,}568{,}654{,}731{,}688{,}484{,}952{,}572{,}018{,}097{,}043{,}964{,}584{,}557.$$

3. Next, we choose *another* natural number m, and divide m into n^k, yielding a quotient q and a remainder r, where $0 \le r < m$. That is, we write

$$n^k = m \cdot q + r \qquad (0 \le r < m). \tag{1}$$

By the division algorithm, the numbers q and r here are uniquely determined.

For example, if n and k are as above and we choose $m = 8927$, then one can compute that

$$\begin{aligned} n^k &= 1819^{17} \\ &= 8927 \cdot 2{,}927{,}186{,}201{,}825{,}046{,}457{,}862{,}745{,}792{,}369{,}771{,}767{,}897{,}176{,}772{,}035{,}911 + 7042, \end{aligned} \tag{2}$$

so our remainder r, in this case, is $r = 7042$.

4. The encoded message, then, *is* the remainder r. So again, in the above example, the encoded version of the original message "HI" (or its "numerized" alias 1819) is 7042.

We used a computer to get the equation (2). For smallish numbers, even a pocket calculator will work: see the Exercises at the end of this section.

In real-life implementation of RSA, the numbers n, k, and m will typically be *much* bigger, and therefore the calculations required to find r will be unmanageable on a computer, even a very powerful one, without some added "tricks."

To move towards an understanding of such tricks, note that equation (2) above tells us that

$$1819^{17} - 7042 = 8927 \cdot 2{,}927{,}186{,}201{,}825{,}046{,}457{,}862{,}745{,}792{,}369{,}771{,}767{,}897{,}176{,}772{,}035{,}911,$$

which in turn tells us that

$$8927 | (1819^{17} - 7042). \tag{3}$$

More generally, equation (1) above tells us that

$$n^k - r = m \cdot q,$$

which in turn tells us that

$$m|(n^k - r).$$

The moral of the story is that we should be studying phenomena of the form $m|(a-b)$. We do so in the next section.

Exercises for Part A. You'll need a calculator for these exercises.

1. (a) Numerize the single-letter message "L," using the numerization key above. Call your numerization n: $n =$ _____.

   (b) Compute n^k, with $k = 5$. Just plug n^k into your calculator, and write down the number you get. Answer: $n^k =$ _____.

   (c) Let $m = 577$. Find natural numbers q and r, with $0 \leq r < m$, such that $n^k = m \cdot q + r$. Hint: plug n^k/m into your calculator. Write your answer in decimal form:

   $n^k/m =$ _____.

   Your answer should have some stuff to the left of the decimal, and some stuff to the right: that is, your answer should look like $q.y$, where q and y are natural numbers. Then q *is* your quotient q. To find your remainder r, subtract $m \cdot q$ from n^k.
   Write your answer here:

$$n^k = 577 \cdot \_\_\_\_\_ + \_\_\_\_\_.$$

2. Repeat problem 1 with the message "A," the exponent $k = 7$, and the divisor $m = 223$:

$$n = \_\_\_\_\_; \quad n^k = \_\_\_\_\_; \quad n^k/m = \_\_\_\_\_;$$
$$n^k = 223 \cdot \_\_\_\_\_ + \_\_\_\_\_.$$

3. Repeat problem 1 with the message "ME," the exponent $k = 2$, and the divisor $m = 1137$:

$$n = \_\_\_\_\_; \quad n^k = \_\_\_\_\_; \quad n^k/m = \_\_\_\_\_;$$
$$n^k = 1137 \cdot \_\_\_\_\_ + \_\_\_\_\_.$$

Part B: Congruences. As noted near the end of Part A, it will be useful to consider how to treat situations where $m|(a-b)$, for integers a, b, m. To this end, we begin with:

Definition 1. Let $a, b, m \in \mathbb{Z}$. We say "a is congruent to b mod m," and write

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

if $m|(a-b)$.

For example:

$$\begin{aligned}
31 &\equiv 1 \pmod{10} && \text{(since } 31-1=10\cdot 3\text{, so } 10|(31-1)\text{);}\\
3^5 &\equiv 3 \pmod{24} && \text{(since } 3^5-5=240=24\cdot 10\text{, so } 24|(3^5-5)\text{);}\\
132617 &\equiv 617 \pmod{132} && \text{(since } 132617-617=132000=132\cdot 1{,}000\text{);}\\
-24 &\equiv 48 \pmod{9} && \text{(since } -24-48=72=9\cdot 10\text{);}\\
k &\equiv 1 \pmod{2} && \text{if } k \text{ is odd, since then } k-1 \text{ is divisible by 2;}\\
1819^{17} &\equiv 7042 \pmod{8927} && \text{(by equation (3) above);}\\
732597^{48} &\equiv 1 \pmod{65} && \text{(we'll see why in Part D below);}
\end{aligned}$$

and so on.

In general, a relation of the form $a \equiv b \pmod{m}$ is called a *congruence*. In such a congruence, we call m the *modulus*. And when manipulating congruences, we say that we are doing *modular arithmetic*.

In the next section, we'll need to do a fair amount of modular arithmetic. The following proposition will allow us to do so.

Proposition 1.

(a) Let $a, b, c, m \in \mathbb{Z}$. Then

$$a \equiv b \pmod{m} \quad \text{and} \quad b \equiv c \pmod{m} \Rightarrow a \equiv c \pmod{m}.$$

(b) Let $a, b, c, d, m \in \mathbb{Z}$. Suppose $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$. Then:

(i) $a + c \equiv b + d \pmod{m}$;

(ii) $a - c \equiv b - d \pmod{m}$;

(iii) $ac \equiv bd \pmod{m}$.

Proof.

(a) Let $a, b, c, m \in \mathbb{Z}$, and suppose that $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m}$. Then by definition of congruence, $m|(a-b)$ and $m|(b-c)$. But then m divides the sum $(a-b)+(b-c)$; that is, $m|(a-c)$. So $a \equiv c \pmod{m}$.

(b) (Part (b)(i) only; for the rest, see the Part B Exercises below.) Assume $a, b, c, d, m \in \mathbb{Z}$, $a \equiv b \pmod{m}$, and $c \equiv d \pmod{m}$. Then by definition of congruence, $m|(a-b)$ and $m|(c-d)$. So $m|\big((a-b)+(c-d)\big)$ or, rearranging terms, $m|\big((a+c)-(b+d)\big)$. But then, by definition of congruence, $a + c \equiv b + d \pmod{m}$. □

Remark. In this proof, we've used the results Exercise B(i)-3 in S-POP. You may use these results in the Exercises below. You don't need to cite that S-POP Exercise directly, as long as you use it properly.

Example 1. We can compute that $46 \equiv -6 \pmod{13}$ and $63 \equiv -2 \pmod{13}$. By part (b)(i) of the above proposition, we can add these two congruences together to get

$$109 \equiv -8 \pmod{13},$$

which we can check by noting that $109 - (-8) = 117 = 13 \cdot 9$.

Example 2. Suppose we want to compute the remainder of $11^4 \pmod{57}$, by which we mean the remainder of 11^4 after division by 57. We first note that $11^2 = 121$, and we compute easily that $121 = 57 \cdot 2 + 7$, so $11^2 \equiv 7 \pmod{57}$. But by part (iii) of the above proposition, we can multiply this congruence by itself, to get

$$11^2 \cdot 11^2 \equiv 7 \cdot 7 \pmod{57},$$

or

$$11^4 \equiv 49 \pmod{57}.$$

This last identity tells us that $11^4 = 57 \cdot q + 49$ for some integer q. And since $0 \leq 49 < 57$, we see that 49 must be the remainder of $11^4 \pmod{57}$.

This method is, arguably, easier than actually trying to divide 57 into 11^4 directly.

Example 3. What is the remainder of $11^8 \pmod{57}$? Well, by Example 2 directly above, $11^4 \equiv 49 \pmod{57}$, so by the same kind of argument as was used in that example,

$$11^4 \cdot 11^4 \equiv 49 \cdot 49 \pmod{57},$$

meaning

$$11^8 \equiv 2401 \pmod{57}.$$

Now 2401 is larger than 57, so 2401 can't be a remainder after division by 57. But we can easily divide 57 into 2401: we compute that $2401 = 57 \cdot 42 + 7$, so

$$2401 \equiv 7 \pmod{57}.$$

By part (a) of Proposition 1 above, we can string together the above two congruences $11^8 \equiv 2401 \pmod{57}$ and $2401 \equiv 7 \pmod{57}$ to get

$$11^8 \equiv 7 \pmod{57}.$$

Since 7 *is* less than 57, 7 *is* our remainder (mod 57).

Exercises for Part B.

1. Use your answers to the Exercises for Part A, above, to fill in each of the following blanks:

   (a) $22^5 \equiv$ _____ (mod 577).

   (b) $11^7 \equiv$ _____ (mod 223).

   (c) $2315^2 \equiv$ _____ (mod 1137).

2. Use the methods and results of Examples 2 and 3 in Part B above to compute the remainder of 11^{16} (mod 57).

3. Prove parts (b)(ii,iii) of Proposition 1 above.

   Hint for part (b)(ii): $m|(a-b)$ and $m|(c-d)$, so $m|\big((a-b)-(c-d)\big)$.

   Hint for part (b)(iii): $m|(a-b)$ and $m|(c-d)$, so $m|\big(c(a-b)+b(c-d)\big)$.

Part C: Successive squaring.

In the Exercises for Part A, above, we described a method for obtaining remainders (mod m). This method works fine for numbers n and k that are relatively small. But it fails when these numbers are in, say, the hundreds of digits. This is because, for numbers n and k of such a magnitude, n^k can be astronomical, to the point where even the best of computers can't compute it explicitly. So for such numbers, we need another strategy.

The strategy that we develop here builds on the techniques of Examples 2 and 3 from Part B above. The important idea behind those examples is that of "successive squaring."

Among other things we saw, in those examples, that we could compute 11^8 (mod 57) without ever having to deal with a number as large as 11^8 *explicitly*. We were able to do so by first

computing $11^2 \pmod{57}$, then using this information to compute $11^4 \pmod{57}$, and finally using that information, in turn, to compute $11^8 \pmod{57}$. And the largest number we had to encounter explicitly, in these investigations, was 2401, which is much smaller than $11^8 = 214358881$.

A similar, but somewhat expanded, strategy will allow us to compute n^k efficiently, even for numbers n and k in the hundreds of digits. We will illustrate this strategy, below, with substantially smaller numbers n and k, so that we can do most of the computations "by hand" (or at worst with a pocket calculator). But the same ideas apply to much much larger numbers.

Example 1. Compute $13^{27} \pmod{15}$.

Solution. Our method here will comprise three main steps.

Step 1: We compute the "binary expansion" of the exponent 27. That is, we express 27 as a sum of powers of 2. Such an expansion of a natural number k always exists.

To compute the binary expansion of 27, we first ask: what's the largest power of 2 that "goes into" 27, in the sense of being less than 27? In this case, the answer is $16 = 2^4$. Subtract that power of 2 from 27 to get 11, and ask: what's the largest power of 2 that goes into 11? The answer is $8 = 2^3$. Subtract 8 from 11 to get 3, and ask: a what's the largest power of 2 that goes into 3? The answer is $2 = 2^1$. Subtract 2 from 3 to get $1 = 2^0$, and we're done: we've found that

$$27 = 16 + 8 + 2 + 1.$$

To summarize our thought process: we computed that

$$27 = 16 + 11 = 16 + 8 + 3 = 16 + 8 + 2 + 1.$$

We kept "breaking off" powers of 2 until there were none left to break off.

Step 2. Make a list of the base, 13, raised to successive powers of 2 (starting with $2^0 = 1$), (mod 15). Keep going until you've raised the base to the largest power of 2 appearing in Step 1. Each entry in the list is found by squaring, and reducing (mod 15), the previous entry, as follows.

$$\begin{aligned}
13 &\equiv 13 \pmod{15},\\
13^2 &\equiv 169 \equiv 15 \cdot 11 + 4 \equiv 4 \pmod{15},\\
13^4 &\equiv (13^2)^2 \equiv 4^2 \equiv 16 \equiv 1 \pmod{15},\\
13^8 &\equiv (13^4)^2 \equiv 1^2 \equiv 1 \pmod{15},\\
13^{16} &\equiv (13^8)^2 \equiv 1^2 \equiv 11 \pmod{15}.
\end{aligned}$$

Step 3. Put Steps 1 and 2 together to compute $13^{27} \pmod{15}$, reducing along the way. Like this:

$$\begin{aligned}
13^{27} &\equiv 13^{16+8+2+1} \equiv 13^{16} \cdot 13^8 \cdot 13^2 \cdot 13^1\\
&\equiv 1 \cdot 1 \cdot 4 \cdot 13 \equiv 52 \equiv 15 \cdot 3 + 7 \equiv 7 \pmod{15}.
\end{aligned}$$

To summarize the strategy for finding $n^k \pmod{m}$:

- **Step 1.** Compute the binary expansion of k (write k as a sum of powers of 2, including, if necessary, the power $2^0 = 1$).
- **Step 2**. Make a list of the base n raised to successive powers of 2 (starting with $2^0 = 1$), (mod m). Keep going until you've raised n to the largest power of 2 appearing in Step 1. Each entry in the list is found by squaring, and reducing (mod m), the previous entry.
- **Step 3.** Put Steps 1 and 2 together to compute $n^k \pmod m$, reducing along the way to keep numbers small.

Here's another example.

Example 2. Compute $24^{37} \pmod{57}$.

Solution. Step 1: Compute the "binary expansion" of the exponent 37:

$$37 = 32 + 4 + 1.$$

Step 2. Raise the base 24 to successive powers of 2, (mod 57). Keep going through the largest power of 2 – namely, 32 – appearing in Step 1:

$$\begin{aligned}
24 &\equiv 24 \pmod{57},\\
24^2 &\equiv 576 \equiv 57 \cdot 10 + 6 \equiv 6 \pmod{57},\\
24^4 &\equiv (24^2)^2 \equiv 6^2 \equiv 36 \pmod{57},\\
24^8 &\equiv (24^4)^2 \equiv 36^2 \equiv 1296 \equiv 57 \cdot 22 + 42 \equiv 42 \pmod{57},\\
24^{16} &\equiv (24^8)^2 \equiv 42^2 \equiv 1764 \equiv 57 \cdot 30 + 42 \equiv 54 \pmod{57},\\
24^{32} &\equiv (24^{16})^2 \equiv 54^2 \equiv 2916 \equiv 57 \cdot 51 + 9 \equiv 9 \pmod{57}.
\end{aligned}$$

Step 3. Put Steps 1 and 2 together to compute $24^{37} \pmod{57}$, reducing along the way to keep numbers small. Like this:

$$\begin{aligned}
24^{37} &\equiv 24^{32+4+1} \equiv 24^{32} \cdot 24^4 \cdot 24^1\\
&\equiv 9 \cdot 36 \cdot 24 \equiv 324 \cdot 24 \equiv (57 \cdot 5 + 39) \cdot 24\\
&\equiv 39 \cdot 24 \equiv 936 \equiv 57 \cdot 16 + 24 \equiv 24 \pmod{24}.
\end{aligned}$$

Note that, in Steps 2 and 3 of Example 2 directly above, we had to compute some remainders that weren't immediately obvious. For such remainders, one can use the method of the Exercises from Part A above.

For example, we computed in Step 2 above that $2916 \equiv 57 \cdot 51 + 9 \equiv 9 \pmod{57}$. How did we find this? We divided 2916 by 57 on a calculator: we got $2916/57 = 51.15789474$. This tells us

that the quotient q is 51: that is, $2916 = 57 \cdot 51 + r$, where r is the desired remainder. To find r, we now just subtract: $r = 2916 - 57 \cdot 51 = 9$. So $2916 = 57 \cdot 51 + 9$, so $2916 \equiv 9 \pmod{57}$.

We also note that, in Steps 2 and 3 of both examples above, we used the symbol "$\equiv$" exclusively, even though we could have used "$=$" in some places. For example, we wrote $24^{37} \equiv 24^{32+4+1}$, even though both sides are, in fact, equal. It's safe to use "$\equiv$" always, when computing an answer (mod m), since if two numbers are equal, they're certainly congruent (mod m), for any $m \in \mathbb{Z}$ (since $a = b \Rightarrow a - b = 0 \Rightarrow m|(a-b) \Rightarrow a \equiv b \pmod{m}$, no matter what m is).

It's not always safe to go the other way though: we can certainly have $a \equiv b \pmod{m}$ without having $a = b$.

Exercises for Part C.

Using the method of successive squaring:

1. Compute $3^{42} \pmod{15}$.

2. Compute $27^{84} \pmod{38}$.

3. Numerize the message "HI," using the numerization key on the first page, and encode it using the exponent $k = 17$ and the modulus $m = 8927$. Note: you'll come up with some relatively large numbers here, which you may want to reduce (mod m) in the way described in the Exercises for Part A.

For example, you will have to reduce $1819^2 \pmod{8927}$. Type $1819^2/8927$ into your calculator to get something like $370.646\ldots$. So your quotient is 370. Then enter $1819^2 - 8927 \cdot 370$, to get 5771, so 5771 is your remainder, so $1819^2 \equiv 5771 \pmod{8927}$. And so on.

You might want to check your answer against equation (2) on page 2 of these Notes.

Part D: Prelude to decoding: some number theory.

When we encode a message n by computing the remainder r of $n^k \pmod{m}$, it seems as though we're losing a lot of information. After all, to say that $n^k \equiv r \pmod{m}$, where $0 \leq r < m$, is to say that

$$n^k = m \cdot q + r$$

for integers q and r with $0 \leq r \leq m$. And if only transmit the remainder r, then the receiver will not know the quotient q. Without this, how can they recover n, or even n^k (even knowing m and k)?

Surprisingly, there *is* a way, under appropriate circumstances. To see how, we'll need some relatively basic number theory notions.

We begin with:

Definition 1. Let $a, b \in \mathbb{Z}$; assume that either $a \neq 0$ or $b \neq 0$. We define the *greatest common divisor*, or gcd, of a and b, denoted gcd(a,b), by

$$\gcd(a,b)=\text{the largest natural number that divides both } a \text{ and } b.$$

We also define $\gcd(0,0) = 0$.

The reason for the special definition in the case $a = b = 0$ is this. If $a = b = 0$, there *is* no largest natural number dividing both a and b, since *any* natural number divides 0. So the general definition given above does not make sense in this case. We need a separate definition, and defining $\gcd(0,0) = 0$ makes some calculations turn out nicely, as we'll see below.

For some quick examples, we have:

$$\begin{aligned}
\gcd(15,21) &= 3; \\
\gcd(-15,-21) &= 3; \\
\gcd(19327,-19327) &= 19327; \\
\gcd(143,11) &= 11 && \text{(since } 143 = 11 \cdot 13); \\
\gcd(a,b) &= \gcd(b,a) && \forall a,b \in \mathbb{Z}; \\
\gcd(1,b) &= 1 && \forall b \in \mathbb{Z}; \\
\gcd(a,0) &= |a| && \forall a \in \mathbb{Z};
\end{aligned}$$

and so on. Note that our definition $\gcd(0,0) = 0$ makes the last identity above true even for $a = 0$.

There are various methods for *finding* the gcd of a given pair of integers a and b. One way would be to list all divisors of a, list all divisors of b, and take the largest number that's common to both lists. For example, 600 has divisors

$$1, 2, 3, 4, 5, 6, 8, 10, 12, 15, 20, 24, 25, 30, 40, 50, 60, 75, 100, 120, 150, 200, 300, 600$$

and 14640 has divisors

$$1, 2, 3, 4, 5, 6, 8, 10, 12, 15, 16, 20, 24, 30, 40, 48, 60, 61, 80, 120, 122, 183, 240,244, 305, 366, 488, 610, 732, 915, 976, 12201464, 1830, 2440, 2928, 3660, 4880, 7320, 14640.$$

Scanning both lists, we see that the largest common entry is 120, so $\gcd(600,14640) = 120$.

As this example illustrates, though, writing out all divisors of a and b can be cumbersome. A perhaps better way is this: write both a and b as powers of distinct primes. Then the largest product of (positive) powers of primes that is common to both factorizations *is* the greatest common divisor of a and b. If there is no such product, then $\gcd(a,b) = 1$.

For example: it's clear that $600 = 6 \cdot 100$. Moreover, since $6 = 2 \cdot 3$ and $100 = 10^2 = (2 \cdot 5)^2 = 2^2 \cdot 5^2$, we have $600 = 2 \cdot 3 \cdot 2^2 \cdot 5^2 = 2^3 \cdot 3 \cdot 5^2$. (When factoring integers into primes, it's conventional, though not necessary, to write the prime powers in order of ascending base.)

Factoring 14640 is harder, but we can begin by noting that this integer ends in zero, so we can factor out a 10: $14640 = 1464 \cdot 10 = 1464 \cdot 2 \cdot 5$. Now we see that 1464 is even, so we factor out a 2: $1464 = 732 \cdot 2$, so $14640 = (732 \cdot 2) \cdot 2 \cdot 5 = 732 \cdot 2^2 \cdot 5$. Clearly 732 is even: $732 = 366 \cdot 52$, so $14640 = (366 \cdot 2) \cdot 2^2 \cdot 5 = 366 \cdot 2^3 \cdot 5$. And again: $366 = 183 \cdot 2$, so $14640 = (183 \cdot 2) \cdot 2^3 \cdot 5 = 183 \cdot 2^4 \cdot 5$. Now it is not hard to see that $183 = 3 \cdot 61$, and that 61 is prime. So finally,

$$14640 = (3 \cdot 61) \cdot 2^4 \cdot 5 = 2^4 \cdot 3 \cdot 5 \cdot 61,$$

which is a factorization into primes. Putting it all together gives

$$\gcd(600, 14640) = \gcd(2^3 \cdot 3 \cdot 5^2, 2^4 \cdot 3 \cdot 5 \cdot 61) = 2^3 \cdot 3 \cdot 5 = 120,$$

since the largest product of primes powers that is common to both $2^3 \cdot 3 \cdot 5^2$ and $2^4 \cdot 3 \cdot 5 \cdot 61$ is $2^3 \cdot 3 \cdot 5 = 120$.

In general, if there is *no* product of positive powers of prime factors that is common to both a and b, then the only positive integer dividing both a and b is 1. In this case, we say a and b are *coprime* (or *relatively prime*).

For example, 28 and 45 are coprime, since $28 = 2^2 \cdot 7$ and $45 = 3^2 \cdot 5$, so no product of prime powers is common to 28 and 45. As this example illustrates, neither a nor b needs to be prime for a and b to be coprime. On the other hand, if a and b are prime and *distinct*, then they are coprime, since they have no common prime factors. (If a and b are prime but *not* distinct – say $a = b = p$ for some prime p – then $\gcd(a,b) = p$.)

The notion of coprime integers is central to the following two theorems, which will in turn be central to RSA decoding.

Theorem RSA$_1$. If $a, b \in \mathbb{N}$ are coprime, then $\exists x, y \in \mathbb{N}$ with

$$ax - by = 1.$$

Example 1.

(a) $a = 5$ and $b = 7$ are distinct primes and therefore are coprime, as noted above. Let $x = 10$ and $y = 7$. Then

$$ax - by = 5 \cdot 10 - 7 \cdot 7 = 50 - 49 = 1,$$

so $x = 10$ and $y = 7$ satisfy the conclusion of the theorem.

(b) $a = 28$ and $b = 45$ are coprime, as noted above. Let $x = 37$ and $y = 23$. Then

$$ax - by = 28 \cdot 37 - 45 \cdot 23 = 1036 - 1035 = 1,$$

so $x = 37$ and $y = 23$ satisfy the conclusion of the theorem.

(c) Since $a = 28$ and $b = 45$ are coprime, so, of course, are $a = 45$ and $b = 28$. Let $x = 5$ and $y = 8$. Then

$$ax - by = 45 \cdot 5 - 28 \cdot 8 = 225 - 224 = 1,$$

so in this case $x = 5$ and $y = 8$ satisfy the conclusion of the theorem.

(d) If $a = 1$ and b is any natural number, then a and b are coprime. Let $x = b + 1$ and $y = 1$. Then

$$ax - by = 1 \cdot (b + 1) - b \cdot 1 = b + 1 - b = 1,$$

so $x = b + 1$ and $y = 1$ satisfy the conclusion of the theorem.

In Part F below, we'll see why Theorem RSA_1 is true, and will also produce an algorithm for producing the natural numbers x and y described in that theorem.

Our final result for this section is the following theorem, which we present without proof. A proof may be found, for example, in *A Friendly Introduction to Number Theory* by Joseph H. Silverman.

Theorem RSA_2. Let $m = pq$ be a product of *distinct* primes p and q; define $\varphi(m) = (p-1)(q-1)$. Then for any $a \in \mathbb{Z}$ that is coprime to m, we have

$$a^{\varphi(m)} \equiv 1 \pmod{m}.$$

Example 2.

(a) Let $p = 5$, $q = 11$, and $a = 17$. Then a and $m = pq = 55$ are coprime. Moreover,

$$\phi(m) = (5-1)(11-1) = 4 \cdot 10 = 40,$$

so by Theorem RSA_2,

$$a^{\varphi(m)} = 17^{40} \equiv 1 \pmod{55}.$$

Let's check that this is correct, by successive squaring: we have

$$40 = 32 + 8$$

and

$$\begin{aligned}
17 &\equiv 17 \pmod{55};\\
17^2 &\equiv 289 \equiv 55 \cdot 5 + 14 \equiv 14 \pmod{55};\\
17^4 &\equiv (17^2)^2 \equiv 14^2 \equiv 196 \equiv 55 \cdot 3 + 31 \equiv 31 \pmod{55};\\
17^8 &\equiv (17^4)^2 \equiv 31^2 \equiv 961 \equiv 55 \cdot 17 + 26 \equiv 26 \pmod{55};\\
17^{16} &\equiv (17^8)^2 \equiv 26^2 \equiv 676 \equiv 55 \cdot 12 + 16 \equiv 16 \pmod{55};\\
17^{32} &\equiv (17^{16})^2 \equiv 16^2 \equiv 256 \equiv 55 \cdot 4 + 36 \equiv 36 \pmod{55}.
\end{aligned}$$

So

$$17^{40} \equiv 17^{32}17^8 \equiv 36 \cdot 26 \equiv 936 \equiv 55 \cdot 17 + 1 \equiv 1 \pmod{55},$$

as the theorem implies.

(b) Let $p = 101$, $q = 103$, and $m = pq = 10403$. One checks that p and q are prime, so $\varphi(m) = (p-1)(q-1) = 100 \cdot 102 = 10200$. Let $a = 11011 = 7 \cdot 11^2 \cdot 13$: since $7, 11$, and 13 are prime, and are distinct from p and q, we see that a and m are coprime. So by Theorem RSA_2,

$$a^{\varphi(m)} = 11011^{10200} \equiv 1 \pmod{10403}.$$

One can, of course, check this by successive squaring, though it would take some work.

Exercises for Part D.

For these exercises, you might find the following list of the first 60 primes useful:

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163, 167, 173, 179, 181, 191, 193, 197, 199, 211, 223, 227, 229, 233, 239, 241, 251, 257, 263, 269, 271, 277, 281

1. Find gcd(9060,333000). Hints: $111 = 3 \cdot 37$; $906 = 3 \cdot 302$.

2. Find gcd(777777,4949). Hints: $777777 = 77 \cdot 10101$; $10101 = 91 \cdot 111$; $4949 = 49 \cdot 101$.

3. 18 and 55 are coprime; note that

$$18 \cdot (-3) - 55 \cdot (-1) = 1.$$

   The problem is that -3 and -1 are not natural numbers. Can you find natural numbers x and y such that $18x - 55y = 1$? Hint: add 55 to -3, and add something to -1 to compensate.

4. Suppose $a, b \in \mathbb{N}$ are coprime; let $x = x_0$ and $y = y_0$ be a pair of natural numbers satisfying

$$ax - by = 1.$$

   (Such x and y exist by Theorem RSA_1 above.)

   (a) Show that the integers

$$x = a - y_0 \quad \text{and} \quad y = b - x_0$$

   satisfy the equation

$$bx - ay = 1.$$

   (b) 41 and 27 are coprime; note that

$$41 \cdot 2 - 27 \cdot 3 = 1.$$

   Use part (a) of this exercise, above, to find natural numbers x and y such that

$$27x - 41y = 1.$$

5. Let $a = 10$ and $m = 33$.

   (a) Check that the hypotheses (that is, the conditions) of Theorem RSA_2 are met for this a and m.

   (b) Compute $a^{\varphi(m)}$ by successive squaring, to confirm that the conclusion of Theorem RSA_2 holds in this case.

Part E: RSA decoding.

Suppose a message n is encoded according to the method described above: that is, the encoded message, which we'll call b, is the remainder of n^k after division by m. In other words, $b = n^k \pmod{m}$, where m and k are known publicly (so that anyone can encode a message).

For the purposes of decoding, we will need to assume further that:

1. $n < m$;
2. n and m are coprime;
3. m is a product of two distinct primes: $m = pq$, where p and q are prime and are *not* made public, but are known to the intended receiver/decoder of the coded message b;
4. k and $\varphi(m) = (p-1)(q-1)$ are coprime.

Conditions 1 and 2 can actually be avoided, but for simplicity, let's assume that they hold.

How does the receiver/decoder decode the message? That is, how do they recover n from b?

The following simple two steps provide the answer:

Step 1: Find natural numbers x and y such that

$$kx - \varphi(m)y = 1. \qquad (*)$$

Such natural numbers are guaranteed to exist by Theorem RSA_1. (The theorem does not tell us how to find x and y. There *is* an efficient method for this, called the *Euclidean algorithm*, but we'll leave this issue aside for now.)

Step 2: Compute $b^x \pmod{m}$: the result *is* the original message n.

That's it! And why does this method work? The justification is simple: since $b = n^k$, we have

$$b^x = \left(n^k\right)^x = n^{kx},$$

by properties of exponents. But by equation $(*)$ above, $kx = 1 + \varphi(m)y$. So

$$b^x = n^{1+\varphi(m)y} = n^1\left(n^{\varphi(m)}\right)^y,$$

again by properties of exponents. But $n^{\varphi(m)} \equiv 1 \pmod{m}$ by Theorem RSA_2, so we get

$$b^x \equiv n^1 1^y \equiv n \pmod{m},$$

and we're done!

The RSA method is secure, if p and q are very large (on the order of several hundred digits each), because factoring $m = pq$, when m is known but p and q are not, is generally very hard, even for the most powerful computers.

Example. A two-letter message is numerized using the key from page 1 above, and is RSA-encoded with $k = 5111$ and $m = 71 \cdot 73 = 5183$. (Note that $k = 19 \cdot 269$ and $\varphi(m) = 70 \cdot 72 = 5040 = 2^4 \cdot 3^2 \cdot 5 \cdot 7$, so $\gcd(k,\varphi(m)) = 1$.) The resulting coded message is 4647. Find the original message (in letters).

SOLUTION. We first need to find natural numbers x and y such that

$$kx - \varphi(m)y = 5111x - 5040y = 1.$$

One checks that $x = 71$ and $y = 72$ work. So we need to compute $b^x = 4647^{71} \pmod{5183}$.

We do so by successive squaring (a calculator, at least, is certainly helpful here): we have

$$71 = 64 + 4 + 2 + 1,$$

and

$$\begin{aligned}
4647 &\equiv 4647 \pmod{5183};\\
4647^2 &\equiv 21594609 \equiv 2231 \pmod{5183};\\
4647^4 &\equiv \left(4647^2\right)^2 \equiv 2231^2 \equiv 4977361 \equiv 1681 \pmod{5183};\\
4647^8 &\equiv \left(4647^4\right)^2 \equiv 1681^2 \equiv 2825761 \equiv 1026 \pmod{5183};\\
4647^{16} &\equiv \left(4647^8\right)^2 \equiv 1026^2 \equiv 1052676 \equiv 527 \pmod{5183};\\
4647^{32} &\equiv \left(4647^{16}\right)^2 \equiv 527^2 \equiv 277729 \equiv 3030 \pmod{5183};\\
4647^{64} &\equiv \left(4647^{32}\right)^2 \equiv 3666^2 \equiv 9180900 \equiv 1807 \pmod{5183}.
\end{aligned}$$

So

$$\begin{aligned}
4647^{71} &\equiv 4647^{64} \cdot 4647^4 \cdot 4647^2 \cdot 4647\\
&\equiv 1807 \cdot 1681 \cdot 2231 \cdot 4647\\
&\equiv 3037567 \cdot 10367457 \equiv 329 \cdot 1457 \equiv 479353 \equiv 2517 \pmod{5153}.
\end{aligned}$$

So the original message is $n = 2517$ which, "denumerized," is the message "OG." (Get it? The original message is OG. :))

Exercises for Part E.

1. Decode the message $b = 23$, with $k = 19$ and $m = 51$, using successive squaring. Hint:

$$19 \cdot 27 - 32 \cdot 16 = 1.$$

2. (a) Encode the message "A," with $k = 35$ and $m = 65$, using numerization and successive squaring.

   (b) Suppose your answer to part (a) of this problem is b. Pretending that you don't know where b came from, decode b to get the original message n. Make sure you get what you should. Hint:
   $$35 \cdot 11 - 48 \cdot 8 = 1.$$

3. (a) Encode the message "V," with $k = 23$ and $m = 77$, using numerization and successive squaring.

   (b) Suppose your answer to part (a) of this problem is b. Pretending that you don't know where b came from, decode b to get the original message n. Make sure you get what you should. Hint:
   $$23 \cdot 47 - 60 \cdot 18 = 1.$$

Part F: Greatest common divisors and RSA: the Euclidean algorithm

In Part E above we saw that, to implement the RSA decoding algorithm with exponent k and modulus m, we needed to find natural numbers x and y satisfying the equation

$$kx - \varphi(m)y = 1.$$

The existence of such numbers x and y is guaranteed by Theorem RSA_1, assuming, again, that k and $\varphi(m)$ are coprime. However, we have, so far, neither proven Theorem RSA_1, nor demonstrated how to actually *find* x and y. The following strategy not only demonstrates why the theorem is true, but also gives an algorithm for finding x and y (and yields an analogous result even when a and b are not coprime).

We have, first of all,

Theorem/Algorithm GCD(a). To find the gcd (greatest common divisor) of two natural numbers a and b:

1. Divide the smaller of these two numbers into the larger.
2. Divide the remainder from the previous step into the divisor from the previous step.
3. Repeat step 2 until you obtain a remainder of zero.
4. When this happens, the *previous* remainder **is** gcd(a,b).

Before explaining *why* the above algorithm works, let's do an example to illustrate *how* it works.

Example 1. Find gcd(245,182).

Solution. First write
$$245 = 182 \cdot 1 + 63.$$

Next, divide the above remainder, 63, into the above divisor, 182:

$$182 = 63 \cdot 2 + 56.$$

Now repeat until a remainder of zero occurs:

$$63 = 56 \cdot 1 + 7,$$
$$56 = 7 \cdot 8 + 0.$$

The above theorem claims that the next-to-last remainder, namely 7, *is* gcd(245,672). This may readily be verified by noting that $245 = 5 \cdot 7^2$ and $672 = 2^5 \cdot 3 \cdot 7$.

The fact that this algorithm really does produce the greatest common divisor can be understood by tracing backwards through the above steps. In particular, the fourth – that is, the last – of the above "remainder equations" shows that 7 divides 56. But then 7 divides both 56 and 7, so by the third remainder equation, 7 divides 63. But then 7 divides both 63 and 56, so by the second remainder equation, 7 divides 182. But then, finally, 7 divides both 182 and 63, so by the first remainder equation, 7 divides 245. So we've shown through the algorithm that 7 divides both 245 and 182. Similar reasoning may be used to show that 7 is the *greatest* common divisor of these two numbers.

Here is another example.

Example 2. Find gcd(2222,286).

Solution. We divide successively, as follows:

$$2222 = 286 \cdot 7 + 220,$$
$$286 = 220 \cdot 1 + 66,$$
$$220 = 66 \cdot 3 + 22,$$
$$66 = 22 \cdot 3 + 0.$$

Since the final remainder before a remainder of zero is 22, we conclude that

$$\gcd(2222,286) = 22.$$

Of course, we do have another way of finding gcd(a,b): we can simply factor a and b into prime powers, and take the largest product of prime powers that divides both a and b. But as already noted, factoring can be very slow, whereas the above method can be shown to be *very fast* (even for very large numbers a and b).

Moreover, the above method can also be used to yields results that are very useful for RSA decoding (among other things). The following theorem explains this.

Theorem/Algorithm GCD(b). Given natural numbers a and b, there are integers x and y such that $ax - by = \gcd(a,b)$. To find these integers x and y:

1. Take the next-to-last of the "remainder equations" that you produced in finding gcd(a,b), and solve this equation for its remainder (which, again, *is* gcd(a,b)).
2. Solve the *previous* remainder equation for the remainder there, and plug this result into the formula just derived for gcd(a,b). Then simplify by collecting like terms.
3. Repeat step 2 until you're done.

Example 3. Let's see how this algorithms works in the context of Example 1 above. Solving the next-to-last remainder equation for the remainder 7 (= gcd(182,245)), we get

$$7 = 63 - 56 \cdot 1.$$

Solving the prior remainder equation for its remainder gives

$$56 = 182 - 63 \cdot 2.$$

Plugging this expression for 56 into the above expression for 7, and collecting like terms, we get

$$7 = 63 - (182 - 63 \cdot 2) \cdot 1 = 63 - 182 \cdot 1 + 63 \cdot 2 \cdot 1 = 63 \cdot 3 - 182 \cdot 1.$$

Next, solving the prior remainder equation for its remainder gives

$$63 = 245 - 182 \cdot 1.$$

Plugging this expression for 63 into the above expression for 7, and collecting like terms, we get

$$7 = (245 - 182 \cdot 1) \cdot 3 - 182 \cdot 1 = 245 \cdot 3 - 182 \cdot 3 - 182 \cdot 1 = 245 \cdot 3 - 182 \cdot 4.$$

So we have expressed gcd(182,245) (namely, 7) in terms of 182 and 245. (It's worth checking that the formula $7 = 245 \cdot 3 - 182 \cdot 4$ is correct.)

Example 4. Find integers x and y such that $2222x - 286y = 22$.

Solution. Using the remainder equations from Example 2 above, we have:

$$\begin{aligned}
22 &= 220 - 66 \cdot 3 && \text{(by third remainder equation of Example 2)} \\
&= 220 - (286 - 220 \cdot 1) \cdot 3 && \text{(by second remainder equation of Example 2)} \\
&= 220 - 286 \cdot 3 + 220 \cdot 3 = 220 \cdot 4 - 286 \cdot 3 && \text{(simplify)} \\
&= (2222 - 286 \cdot 7) \cdot 4 - 286 \cdot 3 && \text{(by first remainder equation of Example 2)} \\
&= 2222 \cdot 4 - 286 \cdot 7 \cdot 4 - 286 \cdot 3 = 2222 \cdot 4 - 286 \cdot (7 \cdot 4 + 3) && \text{(simplify)} \\
&= 2222 \cdot 4 - 286 \cdot 31. && \text{(simplify)}
\end{aligned}$$

Theorem/Algorithm GCD(a) and Theorem/Algorithm GCD(b) together constitute what's known as the Euclidean Algorithm.

Exercises for Part F.

1. Use the Euclidean algorithm to:
   (a) Find gcd(5005,210).
   (b) Find integers x and y such that $\gcd(5005,210) = 5005x - 210y$.
2. Use the Euclidean algorithm to:
   (a) Find gcd(234,432).
   (b) Find integers x and y such that $\gcd(234,432) = 234x - 432y$.
3. Let $k = 19$ and $m = 143$.
   (a) Compute $\varphi(m)$.
   (b) Use the Euclidean algorithm to find $\gcd(k, \varphi(m))$.
   (c) Use the Euclidean algorithm to find integers x and y such that
   $$\gcd(k, \varphi(m)) = kx - \varphi(m)y.$$
   (d) Suppose a message is encoded using RSA with $k = 19$ and $m = 143$. Suppose the coded message is the integer 110. Use your answer to part (c) of this problem, along with the usual RSA decoding algorithm, to decode this message. Express your answer both as an integer and as a "denumerized" message consisting of a single letter.
4. Let $m = 143$ as in the previous problem, but this time, let $k = 17$.
   (a) Use the Euclidean algorithm to show that $\gcd(k, \varphi(m)) = 1$.
   (b) Use the Euclidean algorithm to show that
   $$k(-7) - \varphi(m)(-1) = 1.$$
   (c) The problem with the answer in part (b) is that the coefficient -7 that's multiplying k is not a natural number. We need this coefficient to be a natural number if we want to use the RSA decoding algorithm.
   What do you need to add to the numbers -7 and -1 appearing in part (b) of this problem to get an equation of the form
   $$kx - \varphi(m)y = 1,$$
   where x and y are both *natural numbers*? Hint: you might want to look at Exercise 3 from Part D above.
   (d) Suppose a message is encoded using RSA with $k = 17$ and $m = 143$. Suppose the coded message is a natural number b. What natural number x would you raise b to (mod m), to decode the message? (Use your answer to part (c) of this problem.) You don't need to actually do the decoding, just specify (as an actual explicit natural number, like 43 or 107) what exponent you would use to do so.