CS 70 Discrete Mathematics and Probability Theory
Summer 2019 Course Notes **Note 6.5**

1 Chinese Remainder Theorem

It is worth stepping back for a moment and looking at what the EGCD revealed to us. It said that the GCD could be expressed as $ax+by$ for two numbers x,y. To interpret this, we can imagine the number line, starting at zero and stretching out infinitely in both directions. Imagine that we are only allowed to take steps that are either x or y long. So, if $x=5$ and $y=7$, then we can either move to the right or left by 5 units or 7 units. Suppose we start at zero, and want to know everywhere we can reach by taking a sequence of such moves.

Intuitively, if we can reach a number z, we can reach any multiple of z by simply repeating the steps it took to get to z over and over again. The fact that we can execute the steps of the Euclid's GCD algorithm tells us that anything we can reach by taking steps of x and y must share all the common factors of x and y. This means that we can only reach any multiple of the GCD of x and y. The set of points that we can reach with such operations is called a "lattice" and this lattice-width interpretation of the GCD is interesting[1].

When the GCD is 1, it means that we can reach all points on the integer lattice in this manner. Those who have taken linear algebra will notice a very striking intellectual "rhyme" with the ideas of a basis and span. When their GCD is 1, it is as though the numbers x and y span all the integers[2]. The Chinese Remainder Theorem (CRT) can be interpreted as a way to make this interpretation even more striking.

Suppose we wanted to understand all the numbers mod pq where p and q are relatively prime to each other. If we had to arrange these numbers onto a sheet of paper, how would we do so? Going back to elementary school, it is natural to associate a product pq with a rectangle: p long on one side and q long on the other. So now, we know that we can place the pq numbers from 0 to $pq-1$ on this rectangle. But how? In what order? Given a number, how can you find its "x-coordinate" as something from $0,1,\ldots,p-1$ and its "y-coordinate" as something from $0,1,\ldots,q-1$? The natural first guess is to take a number z and just compute z mod p and z mod q to get two "coordinates" for z.

At this point, it is very useful to do a little exercise for yourself. Suppose $p=3$ and $q=5$ and just place all the numbers from 0 to 14 on this grid. You will see the coordinates as $0=(0,0), 1=(1,1), 2=(2,2), 3=(0,3), 4=(1,4), 5=(2,0), 6=(0,1), 7=(1,2), 8=(2,3), 9=(0,4), 10=(1,0), 11=(2,1), 12=(0,2), 13=(1,3), 14=(2,4)$. When writing them out, you will see that all the numbers lie on a diagonal line that wraps around the rectangle until it fills it. Notice that no two numbers from 0 to 14 have the same coordinates. Furthermore, notice that doing component-wise mod $(3,5)$ addition on the coordinates corresponds to doing mod 15 addition on the numbers themselves. Perhaps more interestingly, doing component-wise mod $(3,5)$ multiplication on the coordinates corresponds to doing mod 15 multiplication on the numbers themselves. (E.g., $3*4=12$ and $(0,3)*(1,4)\equiv(0,2)$). This means that operations can be equivalently performed component-wise in the tuple-representation.

[1] This interpretation also makes short work of the classic family of puzzles of the form "you have a 5 oz cup and a 7 oz cup, an infinite reservoir of water, and a unlimited size mixing bowl. Can you manage to pour exactly z oz of water into a jar?" Do you see how such puzzles can be solved using EGCD?

[2] And when the GCD is 2, we can reach all even numbers. The even numbers behave in a way analogous to a subspace in linear algebra.

Furthermore, we notice that there are two special tuples $(1,0) = 10$ and $(0,1) = 6$. The corresponding numbers act like "orthonormal basis elements" do in linear algebra. They provide an easy way to map from coordinates back to numbers. So (a,b) in coordinates represents the same number as $10a + 6b \bmod 15$. For example, $(2,1) \to 20 + 6 = 26 \equiv 11 \pmod{15}$. So, not only can we easily move from numbers to coordinates (by just taking mods), we can also easily move from coordinates to numbers (by using these special basis elements). Before we state the general form of the Chinese Remainder Theorem, it is useful to observe that the basis element 10 corresponding the first coordinate (obtained by modding by 3) is a multiple of the other modulus 5. This has to be true because its representation in coordinates is designed to have a zero in that other coordinate. Similarly, 6 corresponds to the second coordinate (obtained by modding by 5) and is a multiple of 3.

With this example in hand, we are ready to generalize and to state the result more formally.

Theorem 6.6 (Chinese Remainder Theorem)**.** *Let $n_1, n_2, ..., n_k$ be positive integers coprime to each other, and let $N = n_1 \cdot n_2 \cdot ... \cdot n_k$. Then for any sequence $(a_1, a_2, ..., a_k)$ such that $a_i \in \mathbb{Z}_{n_i}$, there exists a unique $x \in \mathbb{Z}_N$ such that*

$$\begin{aligned} x &\equiv a_1 \pmod{n_1} \\ x &\equiv a_2 \pmod{n_2} \\ &\vdots \\ x &\equiv a_k \pmod{n_k} \end{aligned}$$

Proof. We will first show that such an x exists by explicitly constructing it. Suppose that we have a sequence of integers $(b_1, b_2, ..., b_k)$ such that for each i, $b_i \equiv 1 \pmod{n_i}$ but $b_i \equiv 0 \pmod{n_j}$ for any $j \neq i$. I claim that $a_1b_1 + a_2b_2 + ... + a_kb_k$ satisfies all k of equations. Indeed, if we imagine taking this summation modulo n_1, every term except the first will drop out, as $b_j \equiv 0 \pmod{n_1}$ for all $j \neq 1$. This means we are only left with the first term — and since $b_1 \equiv 1 \pmod{n_1}$, this is equivalent to $a_1 \bmod n_1$, as desired. We can repeat this same argument for all k congruences. We cannot immediately take $x = a_1b_1 + ... + a_kb_k$ though, as this may be larger than N. However, adding or subtracting a multiple of n_i from x will not affect the ith equation; since N is a multiple of all the n_is, adding or subtracting a multiple of N will not affect any equation. Thus, we can take $x = a_1b_1 + ... + a_kb_k \pmod{N}$ as our solution.

Of course, this is all based on the assumption that we had a sequence of integers $(b_1, ..., b_k)$ with the desired properties. Thus, in order to complete the existence proof, we have to show how to construct such integers. In order to construct b_i, we note that for any integer c, $c \cdot \prod_{j \neq i} n_j$ will be a multiple of n_j, and hence will be 0 mod n_j for any $j \neq i$. Thus, we just need to choose c to ensure that $c \cdot \prod_{j \neq i} n_j \equiv 1 \pmod{n_i}$. But this is exactly saying that c should be the multiplicative inverse of $\prod_{j \neq i} n_j$ modulo n_i! How do we know such an inverse exists? Well, we know that each n_j shares no prime factors with n_i, so certainly their product doesn't either. Thus, $\prod_{j \neq i} n_j$ is relatively prime to n_i, and so will have an inverse as required. Putting this all together, we construct our sequence of integers by taking $b_i = \left(\prod_{i \neq j} n_j\right) \cdot \left(\left(\prod_{i \neq j} n_j\right)^{-1} \pmod{n_i}\right)$.

We now show that this solution is unique modulo N; we have two arguments that we could use. The simplest argument is by counting. There are $N = \prod_{i=1}^{k} n_i$ possible values for the $(a_1, a_2, \ldots, a_k)$ tuples and the N numbers from 0 to $N - 1$ each land in exactly one of these. If two landed in one bin, then that means that another bin must be empty. But we can construct an x corresponding to that bin and so it cannot be empty. This means that there must be a bijection from the coordinate tuples $(a_1, a_2, \ldots, a_k)$ and the N numbers from 0 to $N - 1$.

Alternatively, suppose that some y also solves these congruences. Consider $z = y - x$. Clearly $z \bmod n_i$ is zero for all the n_i. This means that z is a multiple of n_i for each i and since they are all coprime, z is a multiple of N, their product. But the difference of two numbers ranging from 0 to $N-1$ must have an absolute value of at most $N-1$. This means that the only multiple of N that z can be is 0. This means that $y = x$ and so indeed, the given solution is unique.

□

The Chinese Remainder Theorem (CRT) is a very powerful tool since it lets us move between numbers and their coordinates for the purpose of doing computations. This means that instead of doing one large calculation, we may be able to get away with doing several smaller calculations and combining their results at the end. Depending on the constraints of your system, this may give you much more power than you would otherwise have.

2 Bijections

A function f with domain D and range R, denoted $f : D \to R$, is simply a way of assigning an element of R to each element of D. It is perfectly allowed to have two elements of D assigned to the same element of R, or to have no element in D assigned to some particular element of R[3], so long as each element in D gets assigned *exactly one* element in R. However, in the rest of this note and the next, we often need functions for which one or both of those do not happen. In order to refer to such functions, we give the following definitions.

Definition 6.1. *Let f be a function from D to R. We say*
(1) f is one-to-one (injective) if for all $x, x' \in D$ such that $x \neq x'$, $f(x) \neq f(x')$.
(2) f is onto (surjective) if for all $y \in R$, there exists an $x \in D$ such that $f(x) = y$.
(3) f is bijective if f is both one-to-one and onto.

Intuitively, f is one-to-one if we don't allow two elements of D to be mapped to the same point in R; f is onto if we make sure there is no element of R that doesn't get mapped to by something.

As an example of where this is helpful, we notice that the Chinese Remainder Theorem gives us a bijection between $\mathbb{Z}_{n_1} \times ... \times \mathbb{Z}_{n_k}$ and $\mathbb{Z}_N$.[4] And indeed, with a little bit more thinking, we realize that it also gives us a bijection in the other direction! In particular, the function that maps x to $(x \bmod n_1, ..., x \bmod n_k)$ is a bijection from $\mathbb{Z}_N$ to $\mathbb{Z}_{n_1} \times ... \times \mathbb{Z}_{n_k}$. The existence part of the CRT tells us that this function is onto, while the uniqueness part tells us it is one-to-one.

What this example hints at is another way of characterizing bijections: bijections are precisely those functions that have inverses. In the CRT example above, we see that the map from $\mathbb{Z}_{n_1} \times ... \times \mathbb{Z}_{n_k}$ to $\mathbb{Z}_N$ and that in the other direction are doing precisely the opposite of one another. We formalize this intuition in the following theorem:

Theorem 6.7. *Let f be a function from D to R. Then f is a bijection if and only if it has an inverse function f^{-1}; that is, if and only if there exists a function $f^{-1} : R \to D$ such that $f(f^{-1}(y)) = y$ for all $y \in R$ and $f^{-1}(f(x)) = x$ for all $x \in D$.*

[3] In other contexts, you may have seen the word range defined to be only those elements which the function can actually take on, while R here would be called the *co-domain*. For the purposes of this class, we do not make this distinction.

[4] In fact, this bijection is also an *isomorphism*, meaning that it doesn't matter if we add/multiply and then apply the function or apply the function first, then add/multiply. This property is precisely what allows us to do arithmetic calculations in the smaller moduli before combining them into the final answer in the larger modulus.

Proof. We first prove the "if direction". Suppose that we have an inverse function f^{-1}. For any $y \in R$, consider $f^{-1}(y)$, which is an element of D. By the definition of an inverse function, $f(f^{-1}(y)) = y$, so there is an element of D that maps to y. We can do this with any $y \in R$, so f must be onto. To prove that f is one-to-one, suppose that we had two values $x, x' \in D$ such that $f(x) = f(x')$. If we plug the same value into f^{-1} twice, we'll get out the same result, so $f^{-1}(f(x)) = f^{-1}(f(x'))$. By the definition of the inverse, we know that the left side is x while the right side is x', so $x = x'$. Thus, for any $x \neq x'$, we have that $f(x) \neq f(x')$, and so f is one-to-one. Since f is both one-to-one and onto, we have that f is bijective.

We now proceed with the "only if" direction. Suppose that f is one-to-one and onto. Since f is onto, we know that for every $y \in R$, there is an $x \in D$ such that $f(x) = y$; since f is one-to-one, this x is unique. Thus, we can define a function f^{-1} that maps each y to its corresponding $x \in D$. By definition then, we will indeed have that $f(f^{-1}(y)) = y$ for all $y \in R$ and $f^{-1}(f(x)) = x$ for all $x \in D$, so f^{-1} is indeed f's inverse.

□

If we look closely at the definition of an inverse function, we notice that $(f^{-1})^{-1}$ is just f itself. This tells us that the inverse of any bijection is itself a bijection, as it has an inverse. Hence, if there is a bijection from D to R, there is also a bijection from R to D; this is why we sometimes say there is a bijection *between D* and R rather than *from* one *to* another.

3 Fermat's Little Theorem

Now that we have discussed bijections, we can give a proof of a very famous theorem, known as Fermat's Little Theorem. While its statement may seem esoteric at first, we will see in the next note that it is the key to why the RSA cryptosystem, and by extension much of modern-day e-commerce, works.

Theorem 6.8 (Fermat's Little Theorem)**.** *Let p be prime and $a \not\equiv 0 \pmod{p}$. Then $a^{p-1} \equiv 1 \pmod{p}$.*

Proof. Let S_p denote the set of non-zero numbers modulo p; that is, $S_p = \{1, 2, ..., p-1\}$. We first prove that the function $f(x) = ax \pmod{p}$ is a bijection from S_p to itself. First, we have to show that for any $x \in S_p$, $f(x) \in S_p$ — if this were not the case, f wouldn't be a function from S_p to itself, much less a bijection! Since $x \in S_p$, we know that x is not divisible by p. We can say the same thing for a, as $a \not\equiv 0 \pmod{p}$. Since p is a prime, we know that ax will not be divisible by p so long as neither a nor x is, so indeed $f(x) \in S_p$.

Now consider the set $S'_p = \{a \bmod p, 2a \bmod p, ..., (p-1)a \bmod p\}$. In other words, $S'_p = \{f(x) | x \in S_p\}$. But $f(x)$ is a bijection, so every element in S_p appears exactly once in S'_p. In particular, this means that S'_p and S_p are *the same set* — it's just that we listed their elements in different orders! But order doesn't matter when multiplying, so we must have that the product of all the elements in S_p is the same as that for S'_p:

$$\prod_{x=1}^{p-1} x = \prod_{x=1}^{p-1} \big(ax \bmod p\big) \tag{1}$$

Recall that for modular equivalences, it does not matter if we take the modulus at an intermediate step or at the end. Thus, if we take (1) modulo p and wait to do the modulus until the end, we get

$$\prod_{x=1}^{p-1} x \equiv \prod_{x=1}^{p-1} ax \equiv a^{p-1} \prod_{x=1}^{p-1} x \pmod{p} \tag{2}$$

where in the last step we factored the a out of each term in the product. But now we're effectively there — if we simply multiply both sides of the equation by the inverse of the product, we'll have $a^{p-1} \equiv 1 \pmod{p}$

as desired. But how do we know that the inverse of that product exists? Since p is prime, we know that none of the numbers in S_p share a factor with it, so their product won't either.[5] In other words, we have that the product is coprime to p, and hence has an inverse. Thus, it is indeed valid for us to multiply both sides of (2) by the inverse of the product in order to get our desired result.

□

In addition to the applications to cryptography we'll see in the next note, Fermat's Little Theorem can also help us speed up modular exponentiation when our modulus is prime. For example, if we wanted to calculate $2^{122} \pmod{11}$, we could note that it is equivalent to $(2^{10})^{12} \cdot 2^2 \equiv 4 \pmod{11}$; this is much faster even than our repeated squaring algorithm from the previous note.

[5] This is the key part of the proof where we use the fact that p is prime. It turns out that there is a generalization of Fermat's Little Theorem, known as *Euler's Totient Theorem*, which allows us to work with non-prime moduli.