Fall 2021 – Math 328K – 55385

Homework 11 Solutions

Problem 1. Here is a way to construct a primitive root modulo p. Let the prime decomposition of $\phi(p) = p - 1$ be

$$p - 1 = q_1^{t_1} \cdots q_k^{t_k}$$

where $q_1, \ldots, q_k$ are distinct prime factors of $p - 1$.

a) Let $a, b \in (\mathbb{Z}/p\mathbb{Z})^\times$ such that $\operatorname{ord}(a)$ and $\operatorname{ord}(b)$ are coprime. Show that $\operatorname{ord}(ab) = \operatorname{ord}(a)\operatorname{ord}(b)$.

b) Suppose that for each $i \in \{1, \ldots, k\}$ we have $a_i \in (\mathbb{Z}/p\mathbb{Z})^\times$ with $\operatorname{ord}(a_i) = q_i^{t_i}$. We showed in class that such an integer exists (in fact, we proved there are $\phi(q_i^{t_i})$ of these). Show that then $a = a_1 \cdots a_k$ is a primitive root.

Solution 1.

a) Let $m = \operatorname{ord}(a)$, $n = \operatorname{ord}(b)$ and $k = \operatorname{ord}(ab)$. Then $(ab)^{mn} = (a^m)^n(b^n)^m = [1]$, so $\operatorname{ord}(ab) \leq mn$.

   On the other hand, we have that $(ab)^k = [1]$. Then $a^k = b^{-k}$, so $\operatorname{ord}(a^k) = \operatorname{ord}(b^{-k}) = \operatorname{ord}(b^k)$. We know that $\operatorname{ord}(a^k) = m/(m,k)$ and $\operatorname{ord}(b^k) = n/(n,k)$. Writing $\ell = \operatorname{ord}(a^k) = \operatorname{ord}(b^k)$ this means $\ell \mid m$ and $\ell \mid n$. But m and n are assumed to be coprime, so $\ell = 1$. This means that $a^k = [1]$ and $b^k = [1]$, so $m \mid k$ and $n \mid k$. As m and n are coprime, this implies that $mn \mid k$. We showed above that $k \leq mn$, so $k = mn$.

b) By repeatedly applying part a) we get

$$\operatorname{ord}(a) = \operatorname{ord}(a_1) \cdots \operatorname{ord}(a_k) = q_1^{t_1} \cdots q_k^{t_k} = p - 1 = \phi(p),$$

   so a is a primitive root.

Problem 2. Alice tries a few improvements to Caesar's cipher ($c_i = m_i + k$). Find the keys using the frequencies of letters and decrypt the ciphertexts. The spaces are only there for readability, they don't correspond to spaces in the plain text.

E is by far the most frequent letter in the English language (11.2%), followed by A (8.5%), R (7.5%), I (7.5%), O (7.1%), T (7.0%), N (6.7%), S (5.7%) and L (5.5%).

a) The text is encrypted with a Caesar cipher, and then the result is encrypted again with Caesar cipher, using a different key.

$$\texttt{FTUEU EQCGU HMXQZ FFAME UZSXQ OMQEM DOUBT QD}$$

b) The cipher $c_i = am_i + b$, using two keys $a \in (\mathbb{Z}/26\mathbb{Z})^\times$ and $b \in \mathbb{Z}/26\mathbb{Z}$. *(This one is a bit more difficult. It might be helpful to use a computer.)*

$$\texttt{FQKVF ZQQUL ZLSNI VFZVZ QUVSH BWVIJ MAJZZ LWKVT VXZ}$$

c) The cipher $c_i = am_i + i$, using a key $a \in (\mathbb{Z}/26\mathbb{Z})^\times$ (the indices start with 0).

$$\texttt{PMWJY LTDMF HVSXG NNBHE UILII FPBPE ECHZI TQLJF IUBBP}$$

Solution 2.

a) Applying to Caesar ciphers in a row is equivalent to a single Caesar cipher. Counting the letters in the cipher text gives us:

$$5 \times \texttt{Q}, \texttt{U}, \quad 4 \times \texttt{E}, \texttt{M}, \quad 3 \times \texttt{F}, \quad 2 \times \texttt{D}, \texttt{O}, \texttt{T}, \texttt{X}, \texttt{Z}, \quad 1 \times \texttt{A}, \texttt{B}, \texttt{C}, \texttt{G}, \texttt{H}, \texttt{S}$$

So we would guess that either Q or U translate to E, which would correspond to the keys M or Q. If we try to decrypt the message with the keys M and Q, respectively, we get

$$\texttt{THISISEQUIVALENTTOASINGLECAESARCIPHER}$$

$$\texttt{PDEOEOAMQERWHAJPPKWOEJCHAYWAOWNYELDAN}$$

The first one looks better.

b) This is a bit trickier as we have two unknowns, but can still be broken with a frequency analysis. Here is one way (out of many) to do this. We have

$$7 \times \texttt{Z}, \quad 6 \times \texttt{V}, \quad 4 \times \texttt{Q}, \quad 3 \times \texttt{F,L}, \quad 2 \times \texttt{I,J,K,S,U,W}, \quad 1 \times \texttt{A,B,H,M,N,T,X}.$$

We sometimes write letters for the corresponding elements of $\mathbb{Z}/26\mathbb{Z}$. Instead of the affine transformation $c_i = am_i + b$ which maps a letter of plain text into cipher text, it is convenient to consider the inverse decryption transformation $m_i = cc_i + d$, where $c = a^{-1} \in (\mathbb{Z}/26\mathbb{Z})^\times$ and $d = -a^{-1}b \in \mathbb{Z}/26\mathbb{Z}$.

Let's focus on the two most frequent letters, Z and V. Let z be the plain text letter which gets encrypted to Z, and v the letter which is encrypted to V. This means

$$z = c\texttt{Z} + d, \quad v = c\texttt{V} + d$$

Subtracting these equations, we get

$$z - v = c(\texttt{Z} - \texttt{V}) = [4]c.$$

Let's make a table of the possible values for $z - v$, for all $c \in (\mathbb{Z}/26\mathbb{Z})^\times$.

c	1	3	5	7	9	11	15	17	19	21	23	25
$4c$	4	12	20	2	10	18	8	16	24	6	14	22

We guess that one of Z and V translates to $\mathtt{E} = 4$, and the other to one of

$$\mathtt{A} = 0,\ \mathtt{R} = 17,\ \mathtt{I} = 8,\ \mathtt{O} = 14,\ \mathtt{T} = 19,\ \mathtt{N} = 13,\ \mathtt{S} = 18,\ \mathtt{L} = 11.$$

For each of these combinations (z, v), we compute $z - v$ and use the table above to obtain c. Often there is no suitable c, which means that this guess for (z, v) cannot occur. Then we can also compute d by $d = z - c\mathtt{Z} = z - c \cdot (-1) = z + c$. For each of these possibilites, we also look at the third most frequent letter, Q, and compute which plain text it corresponds to, that is $c\mathtt{Q} + d$.

We get the following table:

(z, v)	(E,A)	(E,R)	(E,I)	(E,O)	(E,N)	(E,S)	(E,L)	(E,T)
$z - v$	4	13	22	16	17	12	19	11
c	1		25	17		3		
d	5		3	21		7		
$c\mathtt{Q} + d$	V		N	H		D		

(z, v)	(A,E)	(R,E)	(I,E)	(O,E)	(N,E)	(S,E)	(L,E)	(T,E)
$z - v$	22	13	4	10	9	14	7	15
c	25		1	9		23		
d	25		9	23		15		
$c\mathtt{Q} + d$	J		Z	L		T		

Let's restrict to the three cases where Q resolves to a frequent letter, N, L, or T. Trying to decrypt the sentence with these keys gives

YNTIYENNJSESLQVIYEIENJILWCHIVURDUEESHTIKIGE $c = 25, d = 3$

QLJEQOLLVSOSDKREQOEOLVEDIGNERABXAOOSNJEMEWO $c = 9, d = 23$

ATLEASTTHISINCREASESTHENUMBEROFPOSSIBLEKEYS $c = 23, d = 15$

So it seems that the correct decryption key is $c = 23, d = 15$.

c) We can just revert the addition of i, i.e. compute $c'_i = c_i - i$, and obtain

PLUGUGNWEWXKGKSYXKPLANPLKGPANBAXBSAKGAXSUFLKX

The most common letter here is $\mathtt{K} = 10$, it appears 6 times. If the corresponding plain text is $\mathtt{E} = 4$, then the decryption key a^{-1} must be one of the two solutions of $10a^{-1} = 4$, that is $a^{-1} \in \{3, 16\}$. These two possibilities would result in the plain texts

THISISNOMORESECURETHANTHESTANDARDCAESARCIPHER

```
GUISISAOMOEESECUEEGUAAGUESGAAQAEQCAESAECICUEE
```

of which the first one is certainly what we want.

Problem 3. Two texts were encrypted with a Vigenère cipher, both using the same key. The ciphertexts are

We found out that the first message begins with HELLOBOB, and ends with ALICE, but we don't know what is in between.

ALLYY TUJRM ZKXAT BXAMP OIGKE LVRXS ZBCMH YNHLC VSIMV TVIX

AHVRK FOKZJ EKTR

Decrypt the second message.

Solution 3. To get the i–th letter c_i of the ciphertext form the i–th letter m_i of the plain text, we compute

$$c_i = m_i + k_{i \bmod l}$$

where $k_1, \dots, k_l$ are the letters of the key. So when we know the plaintext and the ciphertext, we can recover parts of the key with $k_{i \bmod l} = c_i - m_i$.

For example, when we subtract HELLOBOB from ALLYYTUJ, we get THANKSGI. Similarly, subtracting ALICE from VTVIX gives VINGT. We still don't know the key length, but knowledge of the context let's us guess THANKSGIVING as the key (a more random key might have been more secure).

Problem 4. Let $[a]_p \in (\mathbb{Z}/p\mathbb{Z})^\times$ be a primitive root. We want to show that either $[a]_{p^2}$ or $[a+p]_{p^2}$ is a primitive root in $\mathbb{Z}/p^2\mathbb{Z}$.

a) Let $n = \operatorname{ord}[a]_{p^2}$. Show that either $n = p(p-1)$ or $n = p-1$. The same holds for the order of $[a+p]_{p^2}$.

   Hint: Show that $n \mid p(p-1)$ and $p-1 \mid n$.

b) Show that at least one of $[a]_{p^2}$ and $[a+p]_{p^2}$ is not a solution of the equation $x^{p-1} = 1$, and is therefore a primitive root.

Solution 4.

a) Let $x = [a]_{p^2}$. We know that $(x \bmod p)^n = x^n \bmod p = [1]_p$, so n is a multiple of $\operatorname{ord}(x \bmod p) = p-1$. On the other hand, the order of any element of $(\mathbb{Z}/p^2\mathbb{Z})^\times$ divides $\phi(p^2) = p(p-1)$. So $n \mid p(p-1)$ and $p-1 \mid n$, so n is either $p-1$ or $p(p-1)$.

b) We use Hensel's Lemma for the polynomial $f(x) = x^{p-1} - 1$ with derivative $f'(x) = (p-1)x^{p-2}$. $[a]_p$ is a solution of $f(x) = 0$ in $\mathbb{Z}/p\mathbb{Z}$, and $f'([a]_p) = [(p-1)a^{p-2}]_p \neq [0]_p$, so a unique lift of $[a]_p$ to $\mathbb{Z}/p^2\mathbb{Z}$ solves the equation $f(x) = 0$. Both $[a]_{p^2}$ and $[a+p]_{p^2}$ are such lifts, and are different, so only one of them can be a solution. This means the other one can not have order $p-1$, and so has order $p(p-1)$ by part a). That is, it is a primitive root.