CPSC 418 / MATH 318 — Introduction to Cryptography
ASSIGNMENT 3

Name: Harjee Johal
Student ID: 30000668

Problem 1 — Flawed MAC designs, 13 marks

(a) The message M_2 is of the form $M_2 = M_1||X$, where X is an arbitrary n-bit block. From the question description we also know that $M_1 = P_1||P_2||\ldots||P_L$, where $P_1, P_2, \ldots, P_L$ are each n-bit blocks. In this case,

$$\begin{aligned}\text{PHMAC}_K(M_2) &= ITHASH(K||M_1||X) \\ &= ITHASH(K||P_1||P_2||\ldots||P_L||X)\end{aligned}$$

From the algorithmic description of the ITHASH function, we can see that the result of a given iteration ITHASH($K||P_1||P_2||\ldots||P_{i-1}||P_i$), is equal to $f(H, P_i)$ where f is a public compression function, and H is the value of ITHASH computed in the previous iteration: ITHASH($K||P_1||P_2||\ldots||P_{i-1}$), and P_i is the i^{th} n-bit block of the message being hashed.

From this, we can see that:

$$\begin{aligned}\text{PHMAC}_K(M_2) &= ITHASH(K||M2) \\ &= ITHASH(K||P_1||P_2||\ldots||P_L||X) \\ &= f(ITHASH(K||P_1||P_2||\ldots||P_L), X) \\ &= f(ITHASH(K||M_1), X) \\ &= f(\text{PHMAC}_K(M_1), X)\end{aligned}$$

Since both $\text{PHMAC}_K(M_1)$ and X are known, and f is a publicly accessible function, that means that we can compute $\text{PHMAC}_K(M_2)$ as $f(\text{PHMAC}_K(M_1), X)$, which does not require any knowledge of the value of K. Thus, the computational resistance of PHMAC is thwarted.

(b) Since ITHASH is defined as being not weakly collision resistant, that means that given a pair (M, ITHASH(M)), it's possible to find another pair (M', ITHASH(M')) such that

$$ITHASH(M) = ITHASH(M')$$

and

$$M \neq M'$$

Since the compression function used in ITHASH is public, that means that it's possible to compute ITHASH(M_1). Then, given (M_1, ITHASH(M_1)), we can find a value M_2 such that $M_1 \neq M_2$ and ITHASH(M_1) = ITHASH(M_2). By definition,

$$\text{AHMAC}_K(M_2) = ITHASH(M_2||K)$$

In the previous part, we demonstrated that

$$ITHASH(K||P_1||P_2||\ldots||P_{i-1}||P_i) = f(ITHASH(K||P_1||P_2||\ldots||P_{i-1}), P_i)$$

We know that the key K is also n-bits long. That means that $\text{AHMAC}_K(M_2)$ can be re-written as:

$$\begin{aligned}\text{AHMAC}_K(M_2) &= ITHASH(M_2||K) \\ &= f(ITHASH(M_2), K)\end{aligned}$$

By definition, ITHASH(M_2) = ITHASH(M_1). This means that $\text{AHMAC}_K(M_2)$ can once again be re-written as:

$$\begin{aligned}\text{AHMAC}_K(M_2) &= f(ITHASH(M_1), K) \\ &= ITHASH(M_1||K)\end{aligned}$$

By definition, $\text{AHMAC}_K(M_1) = ITHASH(M_1||K)$. This means that $\text{AHMAC}_K(M_1) = \text{AHMAC}_K(M_2)$. Thus, we have demonstrated that given a message/ AHMAC pair (M_1, AHMAC(M_1)), it's possible to find a message/ AHMAC pair (M_2, AHMAC(M_2)) without knowledge of K. Thus, computational resistance has been defeated in this case.

Problem 2 — Fast RSA decryption using Chinese remaindering, 8 marks)

We are told that $d_p \equiv d \mod p-1$. This means that d_p can be written in the form: $d_p = d + j(p-1)$, where j is an integer. We are also told that $d_q \equiv d \mod q-1$. This means that d_q can be written in the form: $d_q = d + k(q-1)$, where k is an integer.

Next, we are told that $M_p \equiv C^{d_p} \mod p$. Since we know that $d_p = d+j(p-1)$, that means that we can re-write this as: $M_p \equiv C^{d+j(p-1)} \mod p$. We can manipulate this using power rules to obtain:

$$\begin{aligned} & C^{d+j(p-1)} \mod p \\ \equiv & C^d C^{j(p-1)} \mod p \\ \equiv & C^d (C^{p-1})^j \mod p \end{aligned}$$

Since p is a prime, that means that $\phi(p) = p-1$. Furthermore, since we're told that $gcd(n, C) = 1$, that means that C and n share no prime factors. Since n's prime factors are p and q, that means that $gcd(p, C)$ and $gcd(q, C)$ must also be 1. Since $\phi(p) = p-1$ and $gcd(p, C) = 1$, we can apply Euler's theorem on C^{p-1}, meaning that $C^{p-1} \equiv 1 \mod p$. Using this, we can further simplify M_p:

$$\begin{aligned} & C^d (C^{p-1})^j \mod p \\ \equiv & C^d (1)^j \mod p \\ \equiv & C^d (1) \mod p \\ \equiv & C^d \mod p \end{aligned}$$

Thus, we can see that $M_p \equiv C^d \mod p$. That means that $M_p = C^d + sp$, where s is an integer.

Furthermore, we are told that $M_q \equiv C^{d_q} \mod q$. Since we know that $d_q = d + k(q-1)$, that means that we can re-write this as: $M_q \equiv C^{d+k(q-1)} \mod q$. We can manipulate this using power rules to obtain:

$$\begin{aligned} & C^{d+k(q-1)} \mod q \\ \equiv & C^d C^{k(q-1)} \mod q \\ \equiv & C^d (C^{q-1})^k \mod q \end{aligned}$$

Since q is a prime, that means that $\phi(q) = q-1$. Earlier, we demonstrated that since $gcd(n, C) = 1$, then $gcd(q, C)$ must also equal 1. Since $\phi(q) = q-1$ and $gcd(q, C) = 1$, we can apply Euler's theorem on C^{q-1}, meaning that $C^{q-1} \equiv 1 \mod q$. Using this, we can further simplify M_q:

$$
\begin{aligned}
&C^d(C^{q-1})^k \mod q \\
\equiv &C^d(1)^k \mod q \\
\equiv &C^d(1) \mod q \\
\equiv &C^d \mod q
\end{aligned}
$$

Thus, we can see that $M_q \equiv C^d \mod q$. That means that $M_q = C^d + tq$, where t is an integer.

Lastly, we are told that $M' \equiv pxM_q + qyM_p \mod n$, where M' is the value derived from this version of RSA decryption. We can substitute M_p with $C^d + sp$ and M_q with $C^d + tq$ to obtain:

$$
\begin{aligned}
M' &\equiv pxM_q + qyM_p \mod n \\
&\equiv px(C^d + tq) + qy(C^d + sp) \mod n \\
&\equiv pxC^d + pqtx + qyC^d + pqsy \mod n \\
&\equiv C^d(px + qy) + pq(tx + sy) \mod n
\end{aligned}
$$

Where $tx + sy$ is an integer. We know that $n = pq$, and from step 3 of the algorithm we know that $px + qy = 1$. Using this information, we obtain:

$$
\begin{aligned}
M' &\equiv C^d(px + qy) + pq(tx + sy) \mod n \\
&\equiv C^d + n(tx + sy) \mod n \\
&\equiv C^d \mod n
\end{aligned}
$$

We can remove $n(tx + sy)$ from the expression since it's a multiple of n, and the modular arithmetic is being done modulus n. Therefore, we obtain $M' \equiv C^d \mod n \equiv M^{ed} \mod n$. In RSA, integers e and d are determined so that $ed \equiv 1 \mod n$. Therefore, that means that $M' \equiv M^{ed} \mod n \equiv M^1 \mod n \equiv M \mod n$. Therefore, the M' that we obtain using this method of decryption is the same as the M determined during "normal" RSA decryption.

Problem 3 — RSA primes too close together, 18 marks)

(a) We are told that $y > 0$. We are also told that $x + y = p$. This means that $y = p - x$. Since $y > 0$, then that means that $p - x > 0$ as well. Since $p - x > 0$, then $p > x$.

We are told that $n = x^2 - y^2$. This means that $y^2 = x^2 - n$. Since $y > 0$, that means that $y^2 > 0$. Since $y^2 = x^2 - n$, and $y^2 > 0$, then that means $x^2 - n > 0$ as well. If $x^2 - n > 0$, then $x^2 > n$, which means that $x > \sqrt{n}$.

From this, we can see that $p > x$ and that $x > \sqrt{n}$. Thus, we can see that $p > x > \sqrt{n}$.

(b) We're told that in the Fermat factorization algorithm there is a *while* loop that computes both $a = a + 1$ and $b = \sqrt{a^2 - n}$. This loop continues while the computed value of b isn't an integer.

We know that $n = x^2 - y^2$. Therefore, we can re-write this equation as: $b = \sqrt{a^2 - (x^2 - y^2)}$. When $a = x$, then we get:

$$
\begin{aligned}
b &= \sqrt{a^2 + y^2 - x^2} \\
b &= \sqrt{(x)^2 + y^2 - x^2} \\
b &= \sqrt{y^2} \\
b &= y
\end{aligned}
$$

Thus, when $a = x$, then $b = y$. Since by definition y is an integer, that means that the *while* loop will terminate when $a = x$.

We're also told that the algorithm outputs $a - b$ once it's completed. Since the *while* loop terminates when $a = x$, and $b = y$ when $a = x$, then that means $a - b = x - y$. By definition, $q = x - y$, meaning that when this algorithm terminates, it outputs q.

We can also show that the *while* loop won't terminate for a value of a less than x using a proof by contradiction. Suppose that the loop does terminate for some integer $a < x$. Then that means that for that value of a, we get $b = \sqrt{a^2 - n}$, where b is an integer. We can re-arrange the equation $b = \sqrt{a^2 - n}$ to obtain:

$$
\begin{aligned}
b &= \sqrt{a^2 - n} \\
b^2 &= a^2 - n \\
n &= a^2 - b^2 \\
n &= (a + b)(a - b)
\end{aligned}
$$

From this, we can see that $n = (a + b)(a - b)$. Since $n = pq$ and $n > p > q > 0$, then that means $p = a + b$ and $q = a - b$. From these two equations, we can see that $a = \frac{p+q}{2}$. By definition, $x = \frac{p+q}{2}$. Therefore, that would mean that $a = x$. However, we defined a such that $a < x$. This is a contradiction. Therefore, this shows that the loop does not terminate for any value of $a < x$.

(c) The value of a is initialized as $a = \lceil\sqrt{n}\rceil$. During the first iteration of the loop, the algorithm first performs $a = a + 1$, and then computes $b = \sqrt{a^2 - n}$. It continues to do this until it finally reaches $a = x$. This means that the first iteration of the loop is performed with $a = \lceil\sqrt{n}\rceil + 1$, and the last iteration of the loop is done with $a = x$. The number of iterations between the first and last values of a is thus $x - \lceil\sqrt{n}\rceil$. The last iteration of the loop is after $a = x$. The condition at the top of the *while* loop is checked one last time. Since b is an integer when $a = x$, that means that the condition at the top of the *while* loop will not be satisfied on the last iteration, meaning that the *while* loop will be skipped. However, this iteration is still counted. Therefore, there are $x - \lceil\sqrt{n}\rceil$ iterations where the loop is entered, and 1 iteration where the loop is not entered. Therefore, there are $x - \lceil\sqrt{n}\rceil + 1$ iterations overall.

(d) We are asked to prove that $x - \lceil\sqrt{n}\rceil < \frac{y^2}{2\sqrt{n}}$. We know that $n = x^2 - y^2$. We can rearrange this to get $y^2 = x^2 - n = (x + \sqrt{n})(x - \sqrt{n})$. We can re-arrange this equation to get: $(x - \sqrt{n}) = \frac{y^2}{x+\sqrt{n}}$.

In part (a), we showed that $x > \sqrt{n}$. From this, we can see that $x + \sqrt{n} > 2\sqrt{n}$. Since $2\sqrt{n} < x + \sqrt{n}$, that means that $\frac{y^2}{2\sqrt{n}} > \frac{y^2}{x+\sqrt{n}}$. Since $(x - \sqrt{n}) = \frac{y^2}{x+\sqrt{n}}$, that means that $(x - \sqrt{n}) < \frac{y^2}{2\sqrt{n}}$.

By definition, $\lceil\sqrt{n}\rceil \geq \sqrt{n}$. Since $\lceil\sqrt{n}\rceil \geq \sqrt{n}$, then that means $x - \sqrt{n} \geq x - \lceil\sqrt{n}\rceil$. We can use this to show that since $(x - \sqrt{n}) < \frac{y^2}{2\sqrt{n}}$, that means $(x - \lceil\sqrt{n}\rceil) < \frac{y^2}{2\sqrt{n}}$. Thus, we have proven that $x - \lceil\sqrt{n}\rceil < \frac{y^2}{2\sqrt{n}}$.

(e) Suppose that $p - q < 2B\sqrt[4]{n}$. We can re-arrange this to obtain:

$$\begin{aligned}
\frac{p-q}{2} &< B\sqrt[4]{n} \\
y &< B\sqrt[4]{n} \qquad (y = \frac{p-q}{2}) \\
y^2 &< B^2\sqrt{n} \\
\frac{y^2}{\sqrt{n}} &< B^2 \\
\frac{y^2}{2\sqrt{n}} &< \frac{B^2}{2}
\end{aligned}$$

Thus, we can see that $\frac{y^2}{2\sqrt{n}} < \frac{B^2}{2}$. In part (d), we showed that $x - \lceil\sqrt{n}\rceil < \frac{y^2}{2\sqrt{n}}$. Using this statement, we can see that $x - \lceil\sqrt{n}\rceil < \frac{B^2}{2}$. Since $x - \lceil\sqrt{n}\rceil < \frac{B^2}{2}$, then that means that $x - \lceil\sqrt{n}\rceil + 1 < \frac{B^2}{2} + 1$. We showed in part (c) that the number of loops iterations performed by Fermat's Factorization algorithm to factor n is also $x - \lceil\sqrt{n}\rceil + 1$. Therefore, we can see that algorithm factors n after at most $\frac{B^2}{2} + 1$ iterations.

Problem 4 – The El Gamal public key cryptosystem is not semantically secure, 12 marks

We are asked to prove that El Gamal is not semantically secure. We do this by proving 6 assertions. The assertions are as follows:

Assertion 1: If $\left(\frac{y}{p}\right) = 1$ and $\left(\frac{C_2}{p}\right) = 1$, then $C = E(M_1)$:

Since $C_2 \equiv My^k \mod p$, this implies that $\left(\frac{C_2}{p}\right) = \left(\frac{My^k}{p}\right) = \left(\frac{M}{p}\right)\left(\frac{y}{p}\right)^k$. Since $\left(\frac{y}{p}\right) = 1$ and $\left(\frac{C_2}{p}\right) = 1$, we can substitute them into the equation to get: $1 = \left(\frac{M}{p}\right)(1)^k = \left(\frac{M}{p}\right)$. Since $\left(\frac{M}{p}\right) = 1$, then that means that M is a quadratic residue modulo p. We are told in the question that M_1 is a quadratic residue modulo p, and M_2 is not a quadratic residue modulo p. Therefore since the M in $C_2 \equiv My^k \mod p$ is a quadratic residue modulo p, it must be M_1, meaning that in this case $C = E(M_1)$. Thus the assertion is true.

Assertion 2: If $\left(\frac{y}{p}\right) = 1$ and $\left(\frac{C_2}{p}\right) = -1$, then $C = E(M_2)$:

Since $C_2 \equiv My^k \mod p$, this implies that $\left(\frac{C_2}{p}\right) = \left(\frac{My^k}{p}\right) = \left(\frac{M}{p}\right)\left(\frac{y}{p}\right)^k$. Since $\left(\frac{y}{p}\right) = 1$ and $\left(\frac{C_2}{p}\right) = -1$, we can substitute them into the equation to get: $-1 = \left(\frac{M}{p}\right)(1)^k = \left(\frac{M}{p}\right)$. Since $\left(\frac{M}{p}\right) = -1$, then that means that M is not a quadratic residue modulo p. We are told in the question that M_1 is a quadratic residue modulo p, and M_2 is not a quadratic residue modulo p. Therefore since the M in $C_2 \equiv My^k \mod p$ is not a quadratic residue modulo p, it must be M_2, meaning that in this case $C = E(M_2)$. Thus the assertion is true.

Assertion 3: If $\left(\frac{y}{p}\right) = -1$ and $\left(\frac{C_1}{p}\right) = 1$ and $\left(\frac{C_2}{p}\right) = 1$, then $C = E(M_1)$:

Since $y \equiv g^x \mod p$, then that means that $\left(\frac{y}{p}\right) = \left(\frac{g}{p}\right)^x$. Since we know that $\left(\frac{y}{p}\right) = -1$, that means that $-1 = \left(\frac{g}{p}\right)^x$. From this, we can see that $\left(\frac{g}{p}\right) = -1$, and x must be an odd integer. Next, we see that $C_1 \equiv g^k \mod p$, meaning that $\left(\frac{C_1}{p}\right) = \left(\frac{g}{p}\right)^k$. We know that $\left(\frac{C_1}{p}\right) = 1$ and $\left(\frac{g}{p}\right) = -1$, meaning that $1 = (-1)^k$. Therefore, k must be an even number.

Next, we can see that since $C_2 \equiv My^k \mod p$, then $\left(\frac{C_2}{p}\right) = \left(\frac{M}{p}\right)\left(\frac{y}{p}\right)^k$. We know that $\left(\frac{C_2}{p}\right) = 1$ and that $\left(\frac{y}{p}\right) = -1$ and that k is an even number. Therefore, we can see that $1 = \left(\frac{M}{p}\right)(-1)^{even} = \left(\frac{M}{p}\right)$. Since $\left(\frac{M}{p}\right) = 1$, then that means that M is a quadratic residue modulo p. We know that M_1 is a quadratic residue modulo p, and M_2 is not a quadratic residue modulo p. Therefore since the M in $C_2 \equiv My^k \mod p$ is a quadratic residue modulo p, it must be M_1, meaning that in this case $C = E(M_1)$. Thus the assertion is true.

Assertion 4: If $\left(\frac{y}{p}\right) = -1$ and $\left(\frac{C_1}{p}\right) = 1$ and $\left(\frac{C_2}{p}\right) = -1$, then $C = E(M_2)$:

Since $y \equiv g^x \mod p$, then that means that $\left(\frac{y}{p}\right) = \left(\frac{g}{p}\right)^x$. Since we know that $\left(\frac{y}{p}\right) = -1$, that means that $-1 = \left(\frac{g}{p}\right)^x$. From this, we can see that $\left(\frac{g}{p}\right) = -1$, and x must be an odd integer. Next, we see that $C_1 \equiv g^k \mod p$, meaning that $\left(\frac{C_1}{p}\right) = \left(\frac{g}{p}\right)^k$. We know that $\left(\frac{C_1}{p}\right) = 1$ and

$\left(\frac{g}{p}\right) = -1$, meaning that $1 = (-1)^k$. Therefore, k must be an even number.

Next, we can see that since $C_2 \equiv My^k \mod p$, then $\left(\frac{C_2}{p}\right) = \left(\frac{M}{p}\right)\left(\frac{y}{p}\right)^k$. We know that $\left(\frac{C_2}{p}\right) = -1$ and that $\left(\frac{y}{p}\right) = -1$ and that k is an even number. Therefore, we can see that $-1 = \left(\frac{M}{p}\right)(-1)^{even} = \left(\frac{M}{p}\right)$. Since $\left(\frac{M}{p}\right) = -1$, then that means that M is not a quadratic residue modulo p. We know that M_1 is a quadratic residue modulo p, and M_2 is not a quadratic residue modulo p. Therefore since the M in $C_2 \equiv My^k \mod p$ is not a quadratic residue modulo p, it must be M_2, meaning that in this case $C = E(M_2)$. Thus the assertion is true.

Assertion 5: If $\left(\frac{y}{p}\right) = -1$ and $\left(\frac{C_1}{p}\right) = -1$ and $\left(\frac{C_2}{p}\right) = 1$, then $C = E(M_2)$:

Since $y \equiv g^x \mod p$, then that means that $\left(\frac{y}{p}\right) = \left(\frac{g}{p}\right)^x$. Since we know that $\left(\frac{y}{p}\right) = -1$, that means that $-1 = \left(\frac{g}{p}\right)^x$. From this, we can see that $\left(\frac{g}{p}\right) = -1$, and x must be an odd integer. Next, we see that $C_1 \equiv g^k \mod p$, meaning that $\left(\frac{C_1}{p}\right) = \left(\frac{g}{p}\right)^k$. We know that $\left(\frac{C_1}{p}\right) = -1$ and $\left(\frac{g}{p}\right) = -1$, meaning that $-1 = (-1)^k$. Therefore, k must be an odd number.

Next, we can see that since $C_2 \equiv My^k \mod p$, then $\left(\frac{C_2}{p}\right) = \left(\frac{M}{p}\right)\left(\frac{y}{p}\right)^k$. We know that $\left(\frac{C_2}{p}\right) = 1$ and that $\left(\frac{y}{p}\right) = -1$ and that k is an odd number. Therefore, we can see that $1 = \left(\frac{M}{p}\right)(-1)^{odd} = -\left(\frac{M}{p}\right)$, meaning that $\left(\frac{M}{p}\right) = -1$. Since $\left(\frac{M}{p}\right) = -1$, then that means that M is not a quadratic residue modulo p. We know that M_1 is a quadratic residue modulo p, and M_2 is not a quadratic residue modulo p. Therefore since the M in $C_2 \equiv My^k \mod p$ is not a quadratic residue modulo p, it must be M_2, meaning that in this case $C = E(M_2)$. Thus the assertion is true.

Assertion 6: If $\left(\frac{y}{p}\right) = -1$ and $\left(\frac{C_1}{p}\right) = -1$ and $\left(\frac{C_2}{p}\right) = -1$, then $C = E(M_1)$:

Since $y \equiv g^x \mod p$, then that means that $\left(\frac{y}{p}\right) = \left(\frac{g}{p}\right)^x$. Since we know that $\left(\frac{y}{p}\right) = -1$, that means that $-1 = \left(\frac{g}{p}\right)^x$. From this, we can see that $\left(\frac{g}{p}\right) = -1$, and x must be an odd integer. Next, we see that $C_1 \equiv g^k \mod p$, meaning that $\left(\frac{C_1}{p}\right) = \left(\frac{g}{p}\right)^k$. We know that $\left(\frac{C_1}{p}\right) = -1$ and $\left(\frac{g}{p}\right) = -1$, meaning that $-1 = (-1)^k$. Therefore, k must be an odd number.

Next, we can see that since $C_2 \equiv My^k \mod p$, then $\left(\frac{C_2}{p}\right) = \left(\frac{M}{p}\right)\left(\frac{y}{p}\right)^k$. We know that $\left(\frac{C_2}{p}\right) = -1$ and that $\left(\frac{y}{p}\right) = -1$ and that k is an odd number. Therefore, we can see that $-1 = \left(\frac{M}{p}\right)(-1)^{odd} = -\left(\frac{M}{p}\right)$, meaning that $\left(\frac{M}{p}\right) = 1$. Since $\left(\frac{M}{p}\right) = 1$, then that means that M is a quadratic residue modulo p. We know that M_1 is a quadratic residue modulo p, and M_2 is not a quadratic residue modulo p. Therefore since the M in $C_2 \equiv My^k \mod p$ is a quadratic residue modulo p, it must be M_1, meaning that in this case $C = E(M_1)$. Thus the assertion is true.

We've proved that all of Mallory's assertions are true. Therefore, El Gamal is not semantically secure.

Problem 5 — An IND-CPA, but not IND-CCA secure version of RSA, 12 marks

We are asked to show that the version of RSA specified in this question is not IND-CCA secure. We start by choosing two different plaintexts, M_1 and M_2, and receive a ciphertext C that is an encryption of one of them. The ciphertext can be represented as:

$$C = (s||t) = (r^e(mod \quad n)||H(r) \oplus M_i)$$

where i = 1 or i = 2. The value r is a random k-bit value such that $r < n$. In this case, n also has k-bits. H is a public random function that maps $\{0,1\}^k$ to $\{0,1\}^m$, where m is the bit-length of the message being encrypted.

We then compute a new ciphertext, C' from C, such that:

$$C' = (s||t')$$

where $t' = t \oplus M_1$. From here, we have two cases:

Case 1: C is an encryption of M_1.

In the question, we are told that decryption is done via $M \equiv H(s^d(mod \quad n)) \oplus t$. The decryption of C' would therefore be: $M \equiv H(s^d(mod \quad n)) \oplus t'$. Since $t' = t \oplus M_1$, and in this case we know that $t = H(r) \oplus M_1$, that means that $t' = H(r) \oplus M_1 \oplus M_1$. A number XOR'd with itself equals zero, meaning that $t' = H(r) \oplus 0 = H(r)$. Therefore, the decryption of C' can be simplified to:

$$\begin{aligned} M &\equiv H(s^d(mod \quad n)) \oplus t' \\ &\equiv H(s^d(mod \quad n)) \oplus H(r) \end{aligned}$$

We know that $s = r^e$, meaning $s^d = r^{ed}$. By the definition of e and d, we know that $ed = 1 + k\phi(n)$, meaning that $r^{ed} = r^{1+k\phi(n)} = r(r^{\phi(n)})$. We can use Euler's theorem to show that $r^{\phi(n)} \equiv 1 \mod n$ (the probability of $gcd(r,n) \neq 1$ is very low). Thus, we can see that $r^{ed} \equiv r \mod n$. Therefore, $H(s^d(mod \quad n)) = H(r)$. Thus, we can once again modify the decryption process to see that:

$$\begin{aligned} M &\equiv H(s^d(mod \quad n)) \oplus H(r) \\ &= H(r) \oplus H(r) \\ &= 0 \end{aligned}$$

Therefore, when C is an encryption of M_1, then the decryption of C' is 0.

Case 2: C is an encryption of M_2.

In the question, we are told that decryption is done via $M \equiv H(s^d (mod \quad n)) \oplus t$. The decryption of C' would therefore be: $M \equiv H(s^d (mod \quad n)) \oplus t'$. Since $t' = t \oplus M_1$, and in this case we know that $t = H(r) \oplus M_2$, that means that $t' = H(r) \oplus M_1 \oplus M_2$. Therefore, the decryption of C' can be re-written as:

$$\begin{aligned} M &\equiv H(s^d (mod \quad n)) \oplus t' \\ &\equiv H(s^d (mod \quad n)) \oplus H(r) \oplus M_1 \oplus M_2 \end{aligned}$$

From the previous case, we showed that $H(s^d (mod \quad n)) = H(r)$. Thus, we can once again modify the decryption process to see that:

$$\begin{aligned} M &\equiv H(s^d (mod \quad n)) \oplus H(r) \oplus M_1 \oplus M_2 \\ &= H(r) \oplus H(r) \oplus M_1 \oplus M_2 \\ &= M_1 \oplus M_2 \end{aligned}$$

Therefore, when C is an encryption of M_2, then the decryption of C' is $M_1 \oplus M_2$. Mallory can easily compute this value, since she selects both M_1 and M_2.

From this, we can see that Mallory has a method to easily identify whether C is the ciphertext of M_1 or M_2 based on the decryption of C'. If the decryption of C' is 0, then C is an encryption of M_1. Otherwise, if the decryption of C' is $M_1 \oplus M_2$, then C is an encryption of M_2.

Problem 6 — An attack on RSA with small decryption exponent, 25 marks

(a) By definition, $e, d > 1$. The minimum value that ed could possibly take occurs when $e = d = 2$, where we get $ed = 4$. This means that it's not possible to pick values for e, d such that $ed = 1$. Therefore, the minimum value that ed can take on such that $ed \equiv 1 \mod \phi(n)$ is $ed = 1 + \phi(n)$. In this case, $k = 1$, since $ed = 1 + (1)\phi(n)$. Thus, we have demonstrated that $k \geq 1$.

We can show that $k < d$ through a proof by contradiction. Suppose that $k \geq d$. We know that since $ed \equiv 1 \mod \phi(n)$, then $ed = 1 + k\phi(n)$. This can be re-arranged to obtain $ed - k\phi(n) = 1$. Since $k \geq d$, then that means that in order for this equation to hold, $e > \phi(n)$ must also be true. This is because if $e \leq \phi(n)$ that would imply that $ed \leq k\phi(n)$, which would make $ed - k\phi(n) = 1$ impossible. However, $e > \phi(n)$ is not possible, since by definition $1 < e < \phi(n)$. This is a contradiction, meaning that it's not possible to have $k \geq d$. Thus, it must be the case that $k < d$.

From the previous two explanations, we can see that $1 \leq k < d$.

Next, we must prove that $gcd(d, k) = 1$. We can prove this by using the properties of modular inversion. By definition, a number a has an inverse modulo m, meaning that there exists some integer x such that $ax \equiv 1 \mod m$, if and only if $gcd(a, m) = 1$. Furthermore, if $ax \equiv 1 \mod m$, then that means $ax - my = 1$, for some integer y.

Now, let's revisit $ed = 1 + k\phi(n)$. We can re-write this as $ed - k\phi(n) = 1$, which is similar to the form $ax - my = 1$. From this, we can draw two potential conclusions. The first conclusion is that $ed \equiv 1 \mod \phi(n)$, meaning that d has an inverse modulo $\phi(n)$, e. Thus, $gcd(d, \phi(n))$.

The second conclusion is that $ed \equiv 1 \mod k$. This holds because we can write

$$ed - km = 1$$

for some integer m (in this case $m = \phi(n)$). Therefore, since $ed \equiv 1 \mod k$ that means that e is d's inverse modulo k. Since d has an inverse modulo k, that means that $gcd(d, k) = 1$ by the definition of a modular inverse.

(b) In this question, we are asked to prove that $2 \leq n - \phi(n) < 3\sqrt{n}$. Since $n = pq$ and $\phi(n) = (p-1)(q-1)$, we can re-write $n - \phi(n)$ as:

$$\begin{aligned} n - \phi(n) &= pq - (p-1)(q-1) \\ &= pq - (pq - p - q + 1) \\ &= pq - pq + p + q - 1 \\ n - \phi(n) &= p + q + 1 \end{aligned}$$

Now that we've established this, we can find the minimum value of $n - \phi(n)$ by finding the minimum values for p and q. We are told that both p and q are odd primes, and that $q < p < 2q$. The two smallest values that satisfy these conditions are $q = 3$ and $p = 5$. We then plug these values into the equation:

$$\begin{aligned} n - \phi(n) &= p + q - 1 \\ &= 5 + 3 - 1 \\ &= 7 \end{aligned}$$

Therefore, the minimum value for $n - \phi(n)$ under these constraints is 7. From inspection, we can see that $7 \geq 2$, meaning that we have demonstrated that $n - \phi(n) \geq 2$.

Next, we must prove that $n - \phi(n) < 3\sqrt{n}$. We've already shown that:

$$n - \phi(n) = p + q - 1$$

We also know that p and q must be odd primes such that $q < p < 2q$. Since we must pick $p < 2q$, that means that our value for $n - \phi(n)$ is also bounded such that:

$$\begin{aligned} n - \phi(n) &= p + q - 1 \\ &< (2q) + q - 1 = 3q - 1 \\ n - \phi(n) &< 3q - 1 \end{aligned}$$

Thus, we see that $n - \phi(n) < 3q - 1$. We know that $n = pq$. Since $p < 2q$, then that means

$$\begin{aligned} n &< pq \\ &< (2q)q \\ &< 2q^2 \end{aligned}$$

Thus, we know that $n < 2q^2$. From this, we can see that $\sqrt{n} < \sqrt{2}q$, and from this we see that $3\sqrt{n} < 3\sqrt{2}q$.

Now, we compare these two pieces of information. We know that the upper bound of $n - \phi(n)$ is $3q - 1$, and we know that the upper bound of $3\sqrt{n}$ is $3\sqrt{2}q$. We also know that the upper bound for both of these values is defined by the constraint $p < 2q$. By inspection, we can see that $3\sqrt{2}q > 3q - 1$ for all values of q, meaning that $3\sqrt{n} > 3q - 1$. Since we know that $n - \phi(n) < 3q - 1$ and $3q - 1 < 3\sqrt{n}$, then that must mean that $n - \phi(n) < 3\sqrt{n}$.

We have proven that $2 \leq n - \phi(n)$ and that $n - \phi(n) < 3\sqrt{n}$. Therefore, we have proven that $2 \leq n - \phi(n) < 3\sqrt{n}$.

(c) In this part we are asked to prove that $0 < kn - ed < 3d\sqrt{n}$. From the question description, we know that $ed = 1 + k\phi(n)$. We can use this equation to obtain:

$$\begin{aligned} kn - ed &= kn - (1 + k\phi(n)) \\ &= kn - 1 - k\phi(n) \\ &= k(n - \phi(n)) - 1 \end{aligned}$$

From part (b), we know that $2 \leq n - \phi(n)$. From this, we can see that if $2 \leq n - \phi(n)$, then $2k - 1 \leq k(n - \phi(n)) - 1$. From part (a), we know that $1 \leq k$. Therefore, the minimum value k can take is $k = 1$. If we plug this into the inequality, we obtain:

$$2k - 1 = 2(1) - 1 = 1 \leq k(n - \phi(n)) - 1$$

Thus, we have shown that $1 \leq k(n - \phi(n)) - 1$. From this, we can conclude that $0 < k(n - \phi(n)) - 1$

Next, we must show that $k(n - \phi(n)) - 1 < 3d\sqrt{n}$. From part (b), we know that $n - \phi(n) < 3\sqrt{n}$. From this, we can see that $k(n - \phi(n)) - 1 < 3k\sqrt{n} - 1$. In part (a), we determined that $k < d$. Therefore, it must be true that $3k\sqrt{n} - 1 < 3d\sqrt{n} - 1$. Since $3d\sqrt{n} - 1 < 3d\sqrt{n}$, that means that $3k\sqrt{n} - 1 < 3d\sqrt{n}$ must also be true. Since we know $k(n - \phi(n)) - 1 < 3k\sqrt{n} - 1$, and $3k\sqrt{n} - 1 < 3d\sqrt{n}$, we can see that $k(n - \phi(n)) - 1 < 3d\sqrt{n}$.

Thus we have also proven both $0 < k(n - \phi(n)) - 1$ and $k(n - \phi(n)) - 1 < 3d\sqrt{n}$, meaning that we have proven:

$$0 < k(n - \phi(n)) - 1 < 3d\sqrt{n}$$

Since $k(n - \phi(n)) - 1 = kn - ed$, that means we've proven:

$$0 < kn - ed < 3d\sqrt{n}$$

(d) We are asked to show that $0 < \frac{k}{d} - \frac{e}{n} < \frac{1}{2d^2}$. From part (c), we know that

$$0 < kn - ed < 3d\sqrt{n}$$

If we divide the entire inequality by dn, which we can do since $0 < d, n$ by definition, we get:

$$0 < \frac{k}{d} - \frac{e}{n} < \frac{3}{\sqrt{n}}$$

By inspection, we can already see that $0 < \frac{k}{d} - \frac{e}{n}$. Thus, we must prove that $\frac{k}{d} - \frac{e}{n} < \frac{1}{2d^2}$. In the preliminary information for the question, we are told that $d < \frac{\sqrt[4]{n}}{\sqrt{6}}$. We can re-arrange this inequality as follows:

$$
\begin{aligned}
d &< \frac{\sqrt[4]{n}}{\sqrt{6}} \\
d^2 &< \frac{\sqrt{n}}{6} \\
6d^2 &< \sqrt{n}
\end{aligned}
$$

From this, we can conclude that $6d^2 < \sqrt{n}$. If $6d^2 < \sqrt{n}$, then that means $\frac{3}{\sqrt{n}} < \frac{3}{6d^2}$. Since $\frac{k}{d} - \frac{e}{n} < \frac{3}{\sqrt{n}}$ and $\frac{3}{\sqrt{n}} < \frac{3}{6d^2}$, then that means $\frac{k}{d} - \frac{e}{n} < \frac{3}{6d^2}$. We can simply $\frac{3}{6d^2}$ into $\frac{1}{2d^2}$, meaning that $\frac{k}{d} - \frac{e}{n} < \frac{1}{2d^2}$.
Thus, we have demonstrated that $0 < \frac{k}{d} - \frac{e}{n} < \frac{1}{2d^2}$.

(e) In this question, we are asked to prove that when we apply the Euclidean theorem to e and n, $k = A_i$ and $d = B_i$ for some $i \in \{1, 2, \ldots, m\}$, where A_i, B_i are the associated sequences defined in the question. First, let us define our positive rational number $r = \frac{a}{b} \in \mathbb{Q}$ with $a, b \in \mathbb{N}$ as $r = \frac{e}{n}$ as specified in the question. We can then use the theorem defined in the question. The theorem states if we let $r = \frac{a}{b} \in \mathbb{Q}$ and let $\frac{A}{B} \in \mathbb{Q}$ be a fraction in the lowest terms such that

$$\mid r - \frac{A}{B} \mid < \frac{1}{2B^2}$$

Then $A = A_i$ and $B = B_i$ for some $i \in \{1, 2, \ldots, m\}$. In our case, $r = \frac{e}{n}$, meaning that we must find values A and B such that

$$\mid \frac{e}{n} - \frac{A}{B} \mid < \frac{1}{2B^2}$$

Suppose we set $A = k$ and $B = d$. Then $\frac{A}{B} = \frac{k}{d}$. We already know that $\frac{k}{d}$ is in lowest terms, since we proved that $gcd(d, k) = 1$ in part (a). Then if we could show that

$$\mid \frac{e}{n} - \frac{k}{d} \mid < \frac{1}{2d^2}$$

we could prove that $k = A_i$ and $d = B_i$ for some $i \in \{1, 2, \ldots, m\}$. By the definition of absolute values, we know that

$$\mid \frac{e}{n} - \frac{k}{d} \mid \quad = \quad \mid \frac{k}{d} - \frac{e}{n} \mid$$

Furthermore, we know from part (d) that $\frac{k}{d} - \frac{e}{n} > 0$. From this, we can conclude that

$$\left| \frac{k}{d} - \frac{e}{n} \right| \quad = \quad \frac{k}{d} - \frac{e}{n}$$

Which ultimately means that

$$\left| \frac{e}{n} - \frac{k}{d} \right| \quad = \quad \frac{k}{d} - \frac{e}{n}$$

Thus, the inequality in the theorem now becomes:

$$\frac{k}{d} - \frac{e}{n} < \frac{1}{2d^2}$$

We already proved that this inequality is true in part (d). Therefore, we have shown that when we let $\frac{A}{B} = \frac{k}{d}$, we get a fraction in lowest terms such that

$$\left| \frac{e}{n} - \frac{k}{d} \right| < \frac{1}{2d^2}$$

Thus according to the theorem, that means that $k = A_i$ and $d = B_i$ for some $i \in \{1, 2, \ldots, m\}$. Therefore, we have proven what was asked of us.

(f)

Problem 7 — Universal forgery attack on the El Gamal signature scheme, 12 marks)

(a)

(b)

(c)

Problem 9 — Columnar transposition cryptanalysis, 10 marks