Exercise 8: A forgery attack on GHASH. GHASH is used in GCM in TLS and GCMP in WiFi, as well as EIA1 in 4G-LTE. In theory, it has been proved it is secure under the assumption that nonce cannot be reused. As you have seen, in the real world, in both 4G-LTE and WiFi, the nonce can be forced to repeat. Hence, an attacker is able to forge the authentication generated by GHASH. In the following, we will assume that a GHASH polynomial is evaluated in finite field GF(2^4), defined by $t(x) = x^4 + x + 1$, a primitive polynomial, and α is a root of t(x) in GF(2^4). We give the following two pairs of plaintext and ciphertext.

plaintext	ciphertext
M = 001100101111	C = 101000111001
M' = 100000110000	C' = 001011100101

where the right most bit is LSB and each ciphertext is generated by a random cipher.

(a) ** Let H = 0101 in GCMP, compute GHASH(C, H) and GHASH(C', H). Find a ciphertext which has a valid hash value.

$$\boldsymbol{C_1 = 1010, C_2 = 0011, C_3 = 1001}$$

$\boldsymbol{Y_1} = C_1 * H = 1010 * 0101 = \alpha^9 * \alpha^8 = \alpha^{17} = \alpha^2 = \mathbf{0100}$
$\boldsymbol{Y_2} = (C_2 + Y_1) * H = (0011 + 0100) * 0101 = (\alpha + 1 + \alpha^2) * \alpha^8 = 0111 * \alpha^8 = \alpha^{10} * \alpha^8$
$= \alpha^{18} = \alpha^3 = \mathbf{1000}$
$\boldsymbol{Y_3} = (C_3 + Y_2) * H = (1001 + 1000) * 0101 = (\alpha^3 + 1 + \alpha^3) * \alpha^8 = 1 * \alpha^8 = \mathbf{0101}$

$\boldsymbol{GHASH(C, H) = Y_3 = 0101}$

$$\boldsymbol{C'_1 = 0010, C'_2 = 1110, C'_3 = 0101}$$

$\boldsymbol{Y'_1} = C'_1 * H = 0010 * 0101 = \alpha * \alpha^8 = \alpha^9 = \mathbf{1010}$
$\boldsymbol{Y'_2} = (C'_2 + Y'_1) * H = (1110 + 1010) * 0101 = (\alpha^3 + \alpha^2 + \alpha + \alpha^3 + \alpha) * \alpha^8 = \alpha^2 * \alpha^8$
$= \alpha^{10} = \mathbf{0111}$
$\boldsymbol{Y'_3} = (C'_3 + Y'_2) * H = (0101 + 0111) * 0101 = (\alpha^2 + 1 + \alpha^2 + \alpha + 1) * \alpha^8 = \alpha * \alpha^8$
$= \alpha^9 = \mathbf{1010}$

$\boldsymbol{GHASH(C', H) = Y'_3 = 1010}$

Given the linearity of GHASH, the following relation holds:

$$\boldsymbol{GHASH(C, H) \oplus GHASH(C', H) = GHASH(C \oplus C', H)}$$

A valid ciphertext with a valid hash value can be generated as follows:

$\boldsymbol{C'' = C \oplus C' = 101000111001 \oplus 001011100101 = 100011011100}$

(b) ** IN EIA1, let P = 1111, Q = 0001 and OTP = 0011 (i.e. without truncating), compute GHASH(M, P), and GHASH(M', P), the GHASH component in EIA1 for message M and M'.

$$\boldsymbol{M_1 = 0011, M_2 = 0010, M_3 = 1111}$$

$\boldsymbol{Y_1} = M_1 * P = 0011 * 1111 = \alpha^4 * \alpha^{12} = \alpha^{16} = \alpha = \mathbf{0010}$
$\boldsymbol{Y_2} = (M_2 + Y_1) * P = (0010 + 0010) * 1111 = \mathbf{0000}$
$\boldsymbol{Y_3} = (M_3 + Y_2) * P = (1111 + 0000) * 1111 = \alpha^{12} * \alpha^{12} = \alpha^{24} = \alpha^9 = \mathbf{1010}$

$$\boldsymbol{GHASH(M, P) = Y_3 \oplus LEN(M) \otimes Q \oplus OTP}$$

$\boldsymbol{LEN(M)} = 12\ bits = \mathbf{1100}$

$\boldsymbol{GHASH(M, P)} = 1010 \oplus 1100 \otimes 0001 \oplus 0011 = \mathbf{0101}$

$$\boldsymbol{M'_1 = 1000, M'_2 = 0011, M'_3 = 0000}$$

$\boldsymbol{Y'_1} = M'_1 * P = 1000 * 1111 = \alpha^3 * \alpha^{12} = \alpha^{15} = \mathbf{0001}$
$\boldsymbol{Y'_2} = (M'_2 + Y'_1) * P = (0011 + 0001) * 1111 = (\alpha + 1 + 1) * \alpha^{12} = \alpha * \alpha^{12} = \alpha^{13}$
$= \mathbf{1101}$
$\boldsymbol{Y'_3} = (M'_3 + Y'_2) * P = (0000 + 1101) * 1111 = \alpha^{13} * \alpha^{12} = \alpha^{25} = \alpha^{10} = \mathbf{0111}$

$$\boldsymbol{GHASH(M', P) = Y'_3 \oplus LEN(M') \otimes Q \oplus OTP}$$

$\boldsymbol{LEN(M')} = 12\ bits = \mathbf{1100}$

$\boldsymbol{GHASH(M', P)} = 0111 \oplus 1100 \otimes 0001 \oplus 0011 = \mathbf{1000}$

(c) IGNORE

(d) ** Show that after attacker intercepts the MAC-I(M) and MAC-I(M'), he can forge a valid MAC-I(M_{new}) where $\boldsymbol{M_{new} = 0010 \cdot (M + M') + M}$. (Hint. Show that $\boldsymbol{MAC - I(M_{new}) = \alpha^5[MAC - I(M) + MAC - I(M')] + MAC - I(M)}$)

Given the linearity of GHASH, the following relation holds:

$$\boldsymbol{GHASH(C, H) \oplus GHASH(C', H) = GHASH(C \oplus C', H)}$$

$$\boldsymbol{M_{new} = 0010 * (M + M') + M}$$

Because the MAC-I function is based on GHASH, MAC-I is also linear. As a result, it can be shown that **MAC-I(M_{new})** is linearly based on MAC-I(M) and MAC-I(M'):

$$
\begin{aligned}
MAC - I(M_{new}) &= MAC - I(0110 * (M + M') + M) \\
&= MAC - I\big(0110 * (M + M')\big) + MAC - I(M) \\
&= 0110 * MAC - I(M + M') + MAC - I(M)
\end{aligned}
$$

$$\boldsymbol{MAC - I(M_{new}) = \alpha^5[MAC - I(M) + MAC - I(M')] + MAC - I(M)}$$

Therefore, if an attacker intercepts both MAC-I(M) and MAC-I(M'), using their linear combination the attacker can forge $MAC - I(M_{new})$.