Solution of exercise sheet n° 5 :

5-1.

Channel capacity : $C = \max_{p(x)} I(X:Y)$. There are three ways to calculate $I(X:Y)$:

1. $I(X:Y) = \sum_{x,y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}$.
2. $I(X:Y) = H(X) - H(X|Y)$.
3. $I(X:Y) = H(Y) - H(Y|X)$.

We remark that for the (memoryless) additive noise channel where the input X and the noise Z are uncorrelated we can use the relation $H(Y|X) = H(X+Z|X) = H(Z)$. Therefore, for the calculation of the capacity we can use in this exercise the equation

$$C = \max_{p(x)}\{H(Y)\} - H(Z), \tag{1}$$

where the second term $H(Z)$ no longer depends on X (and thus, p(x)).

In the following we need to distinguish three cases : I) $a = 0$, II) $a > 1$ and III) $a = 1$.

- I) $a = 0$ **:** No noise is added, thus $Y = X$ and $H(Z) = 0$. The capacity is therefore $C = \max_{p(x)} H(p, 1-p) = 1$ bit.
- II) $a > 1$ **:** The output alphabet reads $\mathcal{Y} = \{0, 1, a, 1+a\}$. For the input variable X we define the general probability distribution $P(X=0) = p, P(X=1) = 1-p$. Then we can compute probability distribution of the output Y, *i.e.* $P(Y=0) = P(X=0)P(Z=0) = p/2$, $P(Y=1) = P(X=1)P(Z=0) = (1-p)/2$, $P(Y=a) = P(X=0)P(Z=a) = p/2$ and $P(Y=a+1) = P(X=1)P(Z=a) = (1-p)/2$. We conclude that each output can be associated to a unique combination of input X and noise Z and thus, we make no error. We confirm that indeed like in I) again $C = 1$ bit by injecting the probability distribution $p(y)$ in equation (1) :

  $$C = \max_p\{-p\log_2(p/2) - (1-p)\log_2(\frac{1-p}{2})\} - 1 \text{ bit.}$$

  This is maximized by $p = 1/2$ for which $C = 1$ bit.
- III) $a = 1$ **:** In this case, $\mathcal{Y} = \{0, 1, 2\}$. We see like above that if one obtains $Y = 0, 2$ one does not do any error when estimating X. However, $Y = 1$ corresponds to either $X = 0, Z = 1$ or $X = 1, Z = 0$. We find the output probability distribution $p(y) = \{p/2, 1/2, (1-p)/2\}$. Injected in the capacity formula (1) we find that $C = 1/2$ bits for $p = 1/2$.

5-2. This exercise can be solved in the same way as in ex. **5-1**, *i.e.* because the input X and noise Z are independent we can use again Eq. (1) (careful : in general (1) is **is not valid !**).

(a) We again parametrize the input probability distribution, but now, as the input takes 4 values we set it to $p(x) = \{a, b, c, d\}$ where $a+b+c+d = 1$ (alternatively one can include the constraint into the parametrization, so write $p(x) = \{a, b, c, 1-a-b-c\}$). We remark, that $-1 \mod 4 = 3$, so the output alphabet reads $\mathcal{Y} = \{0, 1, 2, 3\}$. Now, we have to find the parameters a, b, c, d that maximize the output entropy $H(Y)$. We can express (similar to ex. **5-1**) $p(y)$ as a function of a, b, c, d, which reads $p(y) = \{\frac{a}{4}+\frac{c}{4}+\frac{d}{2}, \frac{b}{4}+\frac{c}{2}+\frac{d}{4}, \frac{a}{4}+\frac{b}{2}+\frac{c}{4}, \frac{a}{2}+\frac{b}{4}+\frac{d}{4}\}$. We try now to find a, b, c, d such that the (optimal) uniform distribution $p(y) = \{1/4, 1/4, 1/4, 1/4\}$ is reached (careful : in general it may **not be possible** to achieve this! Then one needs to apply the method of Lagrange multipliers). We have thus, 4 equations + 1 equation for the constraint $a+b+c+d=1$ to solve (one equation is linear dependent on the others). We find that $a = c$ and $b = d$. Namely, there is an infinite number of solutions and among them $a = b = c = d = 1/4$, *i.e.*, the uniform distribution for X is a solution.

(b) When we inject the solution of (a) in Eq. (1) we find $C = 1/2$ bits.

(c) We need to sum both noises : $Z_{total} = Z_1 + Z_2$ and again follow a calculation like in (a). One obtains the new probability distribution (attention : $-2 = 2$ since we apply "mod 4"); Z_{total} takes the values : -1, 0, 1, 2 with probabilities 1/4, 3/8, 1/4, 1/8. We have thus $H(Z) = H(1/4, 3/8, 1/4, 1/8)$ and we find $C = \log_2 4 - 1.91 = 0.09$ bits.

(d) $C_{total} = C_1 + C_2 = 1$ bit. (The capacity is additive!)

5-3.

(a) If $p(X = 1) = p$ and $p(X = 0) = 1 - p$. One obtains :

$$I(X : Y) = H(1 - p/2, p/2) - p$$

Taking into account that :

$$\frac{\partial H(x, 1-x)}{\partial x} = \log_2 \frac{1-x}{x} \tag{2}$$

The maximum is found for $p^* = 2/5$ and $C = I_{p=2/5}(X, Y) = \log_2 5 - 2$ bits.

(b) The binary symmetric channel has the capacity $C = 1 - H(\alpha)$ bits, where α is the error rate of the channel, because $I(X : Y) = H(Y) - \sum p(x) H(Y|X = x) = H(Y) - H(\alpha) \leq 1 - H(\alpha)$ bits.

(c) We have $C = \max_{p(x)} I(X : Y)$. Où $I(X : Y) = H(Y) - H(Y|X)$. The entropy at the output is given by $H(Y) = H(1 - pq, pq)$ and the conditional entropy reads $H(Y|X) = pH(q, 1-q)$. $I(X : Y)$ is maximal if $\frac{\partial I(X:Y)}{\partial p} = 0$, which implies

$$q \log_2 \frac{1 - pq}{pq} = H(q, 1-q). \tag{3}$$

To simplify notations we write $H(q, 1-q) = H$. The distribution p that maximizes $I(X : Y)$ is

$$p = \frac{1}{q(1 + 2^{H/q})}. \tag{4}$$

We finally obtain the capacity of the channel :

$$C = \log_2(1 + 2^{H/q}) - \frac{H}{q}. \tag{5}$$

To check consistency we can test Eq. (5) for $q = 0.5$. Since $H(q = 0.5) = 1$ we confirm the result of (a), *i.e.* $C(q = 0.5) = \log_2 5 - 2$ bits.

5-4.

(a) It does not attain the capacity because the equiprobable distribution does not maximize the mutual information of the channel of exercises 5-3.

(b) For a probability distribution $p(x)$, the maximal transmission rate R is bounded from above by $I(X:Y)$. For the channel of exercise 5-3 : $R_{p=1/2} < I_{p=1/2}(X:Y) = 0.3113$ bits.

5-5.

(a) $C = \max_{p(x)} I(X:Y)$ and $\tilde{C} = \max_{p(x)} I(X:\tilde{Y})$.
We have $I(X:Y,\tilde{Y}) = H(X:Y) + H(X:\tilde{Y}|Y)$,
and $I(X:Y,\tilde{Y}) = H(X:\tilde{Y}) + H(X:Y|\tilde{Y})$.
Since $H(X:Y|\tilde{Y}) \geq 0$ and $H(X:\tilde{Y}|Y) = 0$ (see exercise 2-3), we deduce that
$I(X:\tilde{Y}) \leq I(X:Y)$. Thus, $\tilde{C} > C$ is impossible.

(b) One requires $H(X:Y|\tilde{Y}) = 0$. This chain thus must satisfy $X \to \tilde{Y} \to Y$. This is only possible if $Y \leftrightarrow \tilde{Y}$, *i.e.* iff $\tilde{Y} = f(Y)$ is a bijective function.