

Correcteur de facteur de puissance

Corrigé UPSTI

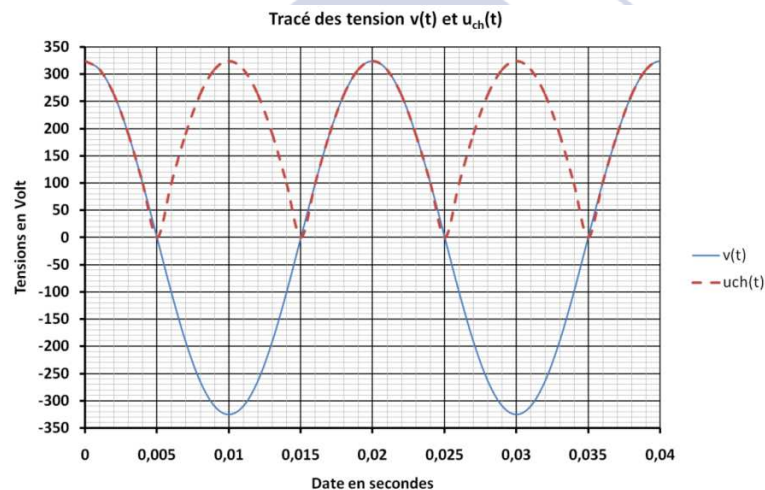
I - Partie I – Redresseur de capacité en tête

I.1 - Présentation du montage

Question 1 On admet que si la charge ne comporte pas de condensateur alors :

- ☞ Les diodes D_1 et D_4 sont passantes et les diodes D_2 et D_3 sont bloquées lorsque $v(t) > 0$
- ☞ Les diodes D_2 et D_3 sont passantes et les diodes D_1 et D_4 sont bloquées lorsque $v(t) < 0$.

Tracer alors l'allure de la tension $u_{ch}(t)$ si la charge n'est constituée que du résistor de résistance R .



Question 2 On admet que la tension aux bornes de la charge $u_{ch}(t)$ est quasiment constante lorsque le temps mis par le condensateur à se décharger dans le résistor est très long devant la période de $u_{ch}(t)$. Déterminer une inégalité entre R , C et T pour qu'il en soit ainsi.

Le temps de charge du condensateur étant $R.C$, il faut que : $T \gg R.C$

I.2 - Spectre du courant de ligne

Question 3 Compléter les lignes de l'extrait de programme donné dans le tableau 1.

- 1- # Fréquence d'échantillonnage
- 2- $F_e = n/t$
- 3- # Nombre d'échantillons
- 4- $N = \text{len}(\text{courant})$
- 5- # Calcul de la transformée de Fourier discrète
- 6- $\text{fft_courant} = \text{np.fft.rfft}(\text{courant})$
- 7- # Module de la transformée de Fourier discrète
- 8- $\text{mod_fft_courant} = \text{np.abs}(\text{fft_courant})$
- 9- # Vecteur fréquences
- 10- $\text{freq} = (2/N) * \text{mod_fft_courant}$

```

11-
12- # Tracé du spectre
13- plt.figure(1,figsize=(8,6))
14- plt.plot(np.linspace(0,(len(freq)-1)*Fe/N,len(freq)),freq,linewidth=2)
15- plt.xlim(0,Fe/2) # Fe/2 en vertu du critère de Shannon
16- plt.xlabel('Fréquences (hz)',fontsize=16)
17- plt.ylim(0,max(freq)+0.5)
18- plt.ylabel('Amplitudes du spectre de Courant (A)',fontsize=16)
19- plt.tick_params(axis='both',labelsize=16)
20- plt.grid()
21- plt.show()

```

Question 4 Donner les valeurs des coefficients $\hat{I}_n$ qui peuvent être déduites graphiquement du spectre en amplitude de $i(t)$ de la figure 3, obtenu à partir du programme. Que peut-on dire des harmoniques paires ?

Par lecture sur le graphique de la figure 3

$$\begin{aligned} \hat{I}_0 &= 0 \text{ A} & \hat{I}_1 &= 3 \text{ A} & \hat{I}_2 &= 0 \text{ A} & \hat{I}_3 &= 1,3 \text{ A} & \hat{I}_4 &= 0 \text{ A} & \hat{I}_5 &= 0,55 \text{ A} & \hat{I}_6 &= 0 \text{ A} & \hat{I}_7 &= 0,45 \text{ A} \\ \hat{I}_8 &= 0 \text{ A} & \hat{I}_9 &= 0,3 \text{ A} & \hat{I}_{10} &= 0 \text{ A} & \hat{I}_{11} &= 0,15 \text{ A} & \hat{I}_{12} &= 0 \text{ A} & \hat{I}_{13} &= 0,15 \text{ A} & \hat{I}_{14} &= 0 \text{ A} \end{aligned}$$

On s'aperçoit que les harmoniques paires sont nulles : $\hat{I}_k = 0$ lorsque k est pair.

Question 5 Retrouver par le calcul la constatation faite à la Q4 en remarquant que le courant d'intensité $i(t)$ vérifie la symétrie de glissement : $i(t) = -i(t - T/2)$.

La D.S.F s'écrit : $i(t) = \sum_{n=0}^{+\infty} (a_n \cos(n\omega_0 t) + b_n \sin(n\omega_0 t)) = \sum_{n=0}^{+\infty} \hat{I}_n \cos(n\omega_0 t + \varphi_n)$ avec : $\hat{I}_n = \sqrt{a_n^2 + b_n^2}$

$$a_0 = \frac{1}{T} \cdot \int_{-T/2}^{T/2} i(t) \cdot dt \quad \text{Et pour } n \geq 1 : \quad a_n = \frac{2}{T} \cdot \int_{-T/2}^{T/2} i(t) \cdot \cos(n\omega_0 t) \cdot dt \quad b_n = \frac{2}{T} \cdot \int_{-T/2}^{T/2} i(t) \cdot \sin(n\omega_0 t) \cdot dt$$

$$\text{On a donc : } a_n = \frac{2}{T} \cdot \left[\int_{-T/2}^0 i(t) \cdot \cos(n\omega_0 t) \cdot dt + \int_0^{T/2} i(t) \cdot \cos(n\omega_0 t) \cdot dt \right]$$

On effectue un changement de variable sur la 1^{ère} intégrale avec : $t = t' - \frac{T}{2} \quad t' = t + \frac{T}{2} \quad dt = dt'$

$$\text{On a alors : } a_n = \frac{2}{T} \cdot \left[\int_0^{T/2} i\left(t' - \frac{T}{2}\right) \cdot \cos\left(n\omega_0 \cdot \left(t' - \frac{T}{2}\right)\right) \cdot dt + \int_0^{T/2} i(t) \cdot \cos(n\omega_0 t) \cdot dt \right]$$

$$\text{Or : } i\left(t' - \frac{T}{2}\right) = -i(t') \quad \text{Et : } \omega_0 = \frac{2\pi}{T} \Rightarrow \cos\left(n\omega_0 \cdot \left(t' - \frac{T}{2}\right)\right) = \cos(n\omega_0 t' - n\pi) = (-1)^n \cdot \cos(n\omega_0 t')$$

$$\text{On en déduit : } a_n = \frac{2}{T} \cdot \left[\int_0^{T/2} -i(t') \cdot (-1)^n \cdot \cos(n\omega_0 t') \cdot dt + \int_0^{T/2} i(t) \cdot \cos(n\omega_0 t) \cdot dt \right]$$

$$a_n = \frac{2}{T} \cdot \left((-1)^{n+1} + 1 \right) \cdot \int_0^{T/2} i(t) \cdot \cos(n\omega_0 t) \cdot dt \quad \text{Or pour pair : } (-1)^{n+1} + 1 = 0$$

Donc pour n pair on a : $a_n = 0$

$$\text{Par un calcul similaire on obtient : } b_n = \frac{2}{T} \cdot \left((-1)^{n+1} + 1 \right) \cdot \int_0^{T/2} i(t) \cdot \sin(n\omega_0 t) \cdot dt$$

Donc pour n pair on a : $b_n = 0$

$$\text{Or : } \hat{I}_n = \sqrt{a_n^2 + b_n^2} \quad \text{Donc pour } n \text{ pair on a : } \hat{I}_n = 0 \quad \text{Les harmoniques paires sont donc nulles}$$

I.3 - Puissance moyenne fournie par le réseau

Question 6 Établir l'expression de P en fonction de $\hat{V}$ et certains paramètres du DSF de $i(t)$. Commenter.

$$P(t) = v(t) \cdot i(t) = \hat{V} \cdot \cos(\omega_0 t) \cdot \sum_{n=0}^{+\infty} \hat{I}_n \cdot \cos(n \cdot \omega_0 t + \varphi_n) = \hat{V} \cdot \sum_{n=0}^{+\infty} \hat{I}_n \cdot \cos(\omega_0 t) \cdot \cos(n \cdot \omega_0 t + \varphi_n)$$

$$P = \langle p(t) \rangle = \langle \hat{V} \cdot \sum_{n=0}^{+\infty} \hat{I}_n \cdot \cos(\omega_0 t) \cdot \cos(n \cdot \omega_0 t + \varphi_n) \rangle = \hat{V} \cdot \sum_{n=0}^{+\infty} \hat{I}_n \cdot \langle \cos(\omega_0 t) \cdot \cos(n \cdot \omega_0 t + \varphi_n) \rangle$$

$$\text{Or si } \omega_1 \neq \omega_2 \text{ alors : } \langle \cos(\omega_1 t) \cdot \cos(n \cdot \omega_2 t + \varphi_n) \rangle = 0$$

$$\text{Donc : } P = \hat{V} \cdot \hat{I}_1 \cdot \langle \cos(\omega_0 t) \cdot \cos(\omega_0 t + \varphi_1) \rangle = \hat{V} \cdot \hat{I}_1 \cdot \langle \frac{1}{2} \cos(2 \cdot \omega_0 t + \varphi_1) + \cos \varphi_1 \rangle$$

$$P = \frac{\hat{V} \cdot \hat{I}_1}{2} (\langle \cos(2 \cdot \omega_0 t + \varphi_1) \rangle + \langle \cos \varphi_1 \rangle) \quad \text{Or : } \langle \cos(2 \cdot \omega_0 t + \varphi_1) \rangle = 0 \quad \text{et : } \langle \cos(\varphi_1) \rangle = \cos \varphi_1$$

$$\text{Donc : } P = \frac{\hat{V} \cdot \hat{I}_1}{2} \cdot \cos \varphi_1$$

Question 7 Calculer numériquement P en supposant que $\varphi_1 = 0$.

$$\text{On a vu à la question 4 que } \hat{I}_1 = 3 \text{ A} \quad P = \frac{325 \times 3}{2} \cdot \cos 0 = 1\,987,5 \text{ W} \approx 2 \text{ kW}$$

I.4 - Modélisation du réseau

Question 8 On souhaite enregistrer le chronogramme de la force électromotrice $e(t)$ à l'aide d'un oscilloscope numérique. Expliquer comment procéder.

On mesure avec l'oscilloscope la tension $v(t)$ et l'intensité $i(t)$. Sachant que l'on a : $e(t) = v(t) + r \cdot i(t) + \lambda \cdot \frac{di(t)}{dt}$ on peut obtenir $e(t)$.

Question 9 Expliquer pourquoi les chronogrammes de la **figure 5** conduisent à considérer $\lambda = 0$ en première approximation. Évaluer ensuite le paramètre r .

On voit sur les chronogrammes de la figure 5 que $e(t) - v(t) = r \cdot i(t) + \lambda \cdot \frac{di(t)}{dt}$ est quasiment proportionnel à $i(t)$. On peut donc avoir l'approximation $\lambda = 0$.

$$\text{On a alors : } r = \frac{e(t) - v(t)}{i(t)}$$

On se place à une date de la période où $i(t) \neq 0$ et où $\frac{di(t)}{dt}$ est minimal.

Cela est vérifié autour des dates $k \cdot 0,02$. On a alors $e(t) - v(t) = 15 \text{ V}$ et $i(t) = 4,2 \text{ A}$

$$\text{On obtient alors : } r \approx \frac{15}{4,2} = 3,6 \, \Omega$$

I.5 - Pertes en ligne

Question 10 Donner l'expression littérale de P_{pertes} . Numériquement on trouve $P_{\text{pertes}} \approx 21\text{W}$. Expliquer comment est obtenue cette valeur numérique à partir des valeurs des coefficients $\hat{I}_n$ obtenues à la Q4 (le calcul numérique complet n'est pas demandé).

$$P_{\text{pertes}} = \frac{1}{T} \int_0^T (e(t) - v(t)) \cdot i(t) \cdot dt \quad \text{Soit pour } \lambda = 0 : \quad P_{\text{pertes}} = \frac{1}{T} \int_0^T r \cdot i^2(t) \cdot dt = r \cdot \frac{1}{T} \int_0^T i^2(t) \cdot dt$$

$$\text{Sachant que : } i(t) = \sum_{n=0}^{+\infty} \hat{I}_n \cdot \cos(n \cdot \omega_0 \cdot t + \varphi_n) \quad P_{\text{pertes}} = r \cdot \left\langle \sum_{n=0}^{+\infty} \hat{I}_n \cdot \cos(n \cdot \omega_0 \cdot t + \varphi_n) \right\rangle^2 >$$

$$\text{Or si } \omega_1 \neq \omega_2 \text{ alors : } \langle \cos(\omega_1 \cdot t) \cdot \cos(n \cdot \omega_2 \cdot t + \varphi_n) \rangle = 0 \quad \text{Donc : } P_{\text{pertes}} = r \cdot \left\langle \sum_{n=0}^{+\infty} \hat{I}_n^2 \cdot \cos^2(n \cdot \omega_0 \cdot t + \varphi_n) \right\rangle$$

$$\text{Or : } \langle \cos^2(\omega \cdot t + \varphi) \rangle = \frac{1}{2} \quad \text{D'où : } P_{\text{pertes}} = \frac{r}{2} \cdot \sum_{n=0}^{+\infty} \hat{I}_n^2 = \frac{3,6}{2} \cdot (3^2 + 1,3^2 + 0,55^2 + 0,45^2 + 2 \times 0,15^2) \approx 21 \text{ W}$$

Question 11 Pour une même puissance moyenne fournie par le réseau électrique, expliquer pourquoi un courant de ligne $i(t)$ sinusoïdal en phase avec la tension du réseau $v(t)$ permet de réduire les pertes en ligne par rapport au cas d'un courant impulsionnel. Pour le redresseur à capacité en tête étudié, quelle réduction de pertes en ligne obtiendrait-on si le courant de ligne $i(t)$ était sinusoïdal en phase avec $v(t)$? On donnera le résultat sous forme d'un pourcentage.

Pour un courant de ligne $i(t)$ sinusoïdal en phase avec la tension du réseau $v(t)$ l'intensité n'a qu'une seule harmonique $\hat{I}_1$. Donc le terme $\sum_{n=0}^{+\infty} \hat{I}_n^2$ est réduit par rapport au courant impulsionnel. En revanche la puissance fournie à la charge est identique car celle-ci ne dépend que de $\hat{I}_1$.

$$\text{Pour le redresseur à capacité en tête les pertes en ligne seraient de : } P_{\text{pertes}} = \frac{r}{2} \cdot \hat{I}_1^2 = \frac{3,6}{2} \cdot 3^2 = 16,2 \text{ W}$$

$$\text{Soit une réduction des pertes en ligne de : } \frac{21 - 16,2}{21} = 0,23 = 23 \%$$

II - Correcteur de facteur de puissance

II.1 - Hacheur survolteur

Question 12 Déterminer $u_{Tr}(t)$, $u_D(t)$, $i_L(t)$, $i_{Tr}(t)$ et $i_D(t)$ pour $t \in [0, \alpha \cdot T_H]$. On justifiera soigneusement l'état de la diode D .

Dans ce cas le transistor est fermé donc : $u_{Tr}(t) = 0$

On a donc : $u_D(t) = -u_{Ch}(t)$ d'où $u_D(t) < 0$, la diode est donc bloquée : $i_D(t) = 0$

$$\text{On a également : } i_L(t) = i_{Tr}(t) \quad \text{avec : } U_0 - L \cdot \frac{d i_L(t)}{dt} = 0 \quad \text{Soit : } i_L(t) = i_{Tr}(t) = I_{\min} + \frac{U_0}{L} \cdot t \quad (a)$$

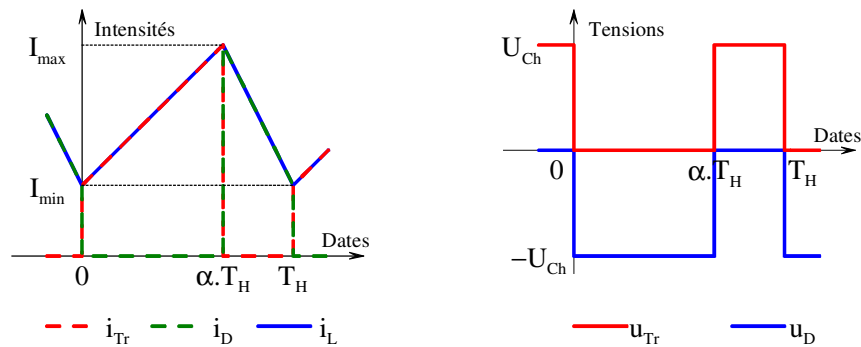
Question 13 Même question pour $t \in [\alpha \cdot T_H, T_H]$.

Le transistor est ouvert donc : $i_{Tr}(t) = 0$ $i_L(t) = i_D(t)$ la diode est donc passante

On a donc d'une part : $u_D(t) = 0$ et : $u_{Tr}(t) = U_{Ch}$

$$\text{Et d'autre part : } U_0 = L \cdot \frac{d i_L(t)}{dt} + U_{Ch} \quad \text{Soit : } i_L(t) = i_D(t) = I_{\max} - \frac{U_{Ch} - U_0}{L} \cdot (t - \alpha \cdot T_H) \quad (b)$$

Question 14 Tracer alors les allures des chronogrammes correspondants.



Question 15 Établir la relation existant entre α , U_{Ch} et U_0 dans ce régime de fonctionnement.

De la relation (a) on a : $i_L(\alpha.T_H) = I_{max} = I_{min} + \frac{U_0}{L}.\alpha.T_H \Rightarrow I_{max} - I_{min} = \frac{U_0}{L}.\alpha.T_H$ (c)

De la relation (b) on a : $i_L(T_H) = I_{min} = I_{max} - \frac{U_{Ch} - U_0}{L}.(T_H - \alpha.T_H) \Rightarrow I_{max} - I_{min} = \frac{U_{Ch} - U_0}{L}.(1 - \alpha).T_H$ (d)

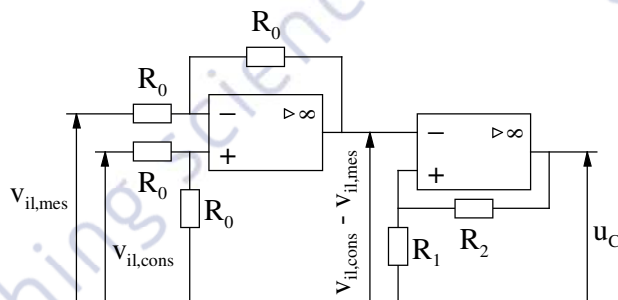
Des relations (c) et (d) on obtient : $\frac{U_0}{L}.\alpha.T_H = \frac{U_{Ch} - U_0}{L}.(1 - \alpha).T_H$ Soit après calcul : $U_{Ch} = \frac{U_0}{1 - \alpha}$

Question 16 On a mesuré $\Delta I = I_{max} - I_{min} = 0,50$ A pour $\alpha = 0,50$, $U_0 = 80$ V et $T_H = 1,0.10^{-3}$ s. Calculer numériquement la valeur de L .

De la relation (c) on obtient : $L = \frac{U_0}{I_{max} - I_{min}}.\alpha.T_H = \frac{80}{0,5} \times 0,5 \times 1,0.10^{-3} = 0,08$ H = 80 mH

II.2 - Principe du correcteur de facteur de puissance et commande du transistor

Question 17 Proposer un circuit de commande du transistor T_r mettant en jeu deux des trois montages de la figure 8. Expliciter alors le paramètre Δ en fonction des résistances utilisées et de V_{sat} .



Avec : $\Delta = \frac{R_1}{R_1 + R_2}.V_{sat}$

II.3 - Mise en œuvre

Question 18 Commenter l'allure des chronogrammes de la figure 9. Comment le paramètre Δ se manifeste-t-il dans ces chronogrammes ? Évaluer la valeur de Δ_1 dans le cas de la figure 9. Comment évolue la valeur du paramètre Δ de la figure 9 à la figure 11 ?

Sur ces chronogrammes, le courant présente une allure sinusoïdale avec du bruit. Lequel bruit est dû à la valeur de Δ . Sur la figure 9, ce bruit peut être estimé à $1,5$ A = $\frac{2.\Delta_1}{k}$. On a donc : $\Delta_1 \approx \frac{1,5 \times 1}{2} = 0,75$ V

Ce bruit est plus faible sur les figures 10 et 11. Donc entre la figure 9 et la figure 11 le paramètre Δ diminue.

Question 19 Décrire l'allure du spectre en amplitude du courant de ligne $i(t)$. Comment l'allure du spectre est-elle modifiée lorsque le paramètre Δ passe de la valeur Δ_1 (figure 9) à la valeur Δ_3 (figure 11). On distinguera clairement les cas des parties "basses fréquences" et "hautes fréquences" du spectre.

Le spectre a un pic principal de 325 V à la fréquence $f = 50$ Hz (basse fréquence) puis un autre pic de valeur qui dépend de Δ à une fréquence (haute fréquence) qui dépend également de Δ .

Plus Δ est petit plus cette haute fréquence est élevée et plus la valeur du pic est faible.

Question 20 Quelle caractéristique des ALI des montages de la figure 8 constitue une limite technologique à la diminution de Δ ?

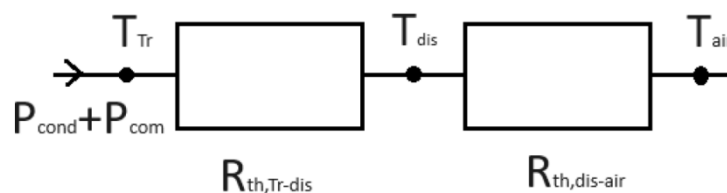
La vitesse de commutation des ALI limite la diminution de Δ

II.4 - Dissipation thermique dans le transistor

Question 21 Le modèle affine choisi est-il en cohérence avec les mesures effectuées ? Justifier.

Le modèle affine est cohérent car en considérant les incertitudes de mesures, l'ensemble des points trouvés expérimentalement se trouvent sur la droite.

Question 22 Proposer, sous la forme d'un schéma électrique équivalent, une modélisation du comportement thermique du transistor en régime permanent. En déduire les expressions de paramètres a et b .



$$(P_{\text{cond}} + P_{\text{com}}) \cdot R_{\text{th,dis} \rightarrow \text{air}} = T_{\text{dis}} - T_{\text{air}}$$

$$\text{or : } P_{\text{com}} = E_{\text{On/Off}} \cdot f_{\text{cm}} \quad \Rightarrow \quad T_{\text{dis}} = T_{\text{air}} + P_{\text{cond}} \cdot R_{\text{th,dis} \rightarrow \text{air}} + E_{\text{On/Off}} \cdot f_{\text{cm}} \cdot R_{\text{th,dis} \rightarrow \text{air}}$$

$$\text{Donc : } a = E_{\text{On/Off}} \cdot R_{\text{th,dis} \rightarrow \text{air}} \quad \text{et : } b = T_{\text{air}} + P_{\text{cond}} \cdot R_{\text{th,dis} \rightarrow \text{air}}$$

Question 23 Déterminer les valeurs numériques de P_{cond} et de $E_{\text{on/off}}$.

$$E_{\text{on/off}} = \frac{a}{R_{\text{th,dis} \rightarrow \text{air}}} = \frac{5,5}{4} = 1,4 \text{ SI} \quad P_{\text{cond}} = \frac{b - T_{\text{air}}}{R_{\text{th,dis} \rightarrow \text{air}}} = \frac{40 - 20}{4} = 5 \text{ SI}$$

III - Contrôle du correcteur de facteur de puissance par Intelligence artificielle

III.1 - Présentation du problème

Question 24 Les différents éléments de cette installation sont-ils câblés en série ou en dérivation ? Justifier ce choix.

Les éléments de cette installation sont câblés en dérivation afin qu'ils soient alimentés avec la tension définie par l'onduleur. S'ils étaient câblés en séries ils ne seraient pas alimentés par la tension pour laquelle ils sont conçus. D'autre part débrancher un élément reviendrait à tous les couper.

Question 25 Quelle est l'allure des courants absorbés par le radiateur électrique et le chauffe-eau lorsqu'ils sont connectés directement au réseau électrique ?

Le radiateur et le chauffe eau sont des résistances pures. Le courant qu'ils absorbent a donc une allure sinusoïdale en phase avec la tension.

Question 26 Donner l'ordre de grandeur de la puissance électrique maximale consommée par un radiateur électrique et un chauffe-eau correspondant aux besoins d'un studio.

L'ordre de grandeur de puissance des ces deux éléments est d'environ 1 KW

Question 27 Combien y a-t-il de combinaisons possibles entre les différents modes de fonctionnement des quatre éléments ?

Il y a $3^4 = 81$ combinaisons possibles

Question 28 Dans le tableau 3, expliquer pourquoi $\Delta_{opt} = -1$ pour les combinaisons telles que $x_1 = 0$ et $x_2 = 0$.

Pour toutes ces combinaisons seuls le radiateur et le chauffe-eau absorbent un courant. Or ceux-ci absorbent un courant sinusoïdal en phase avec la tension. Le but du correcteur de facteur de puissance étant d'avoir un courant sinusoïdal en phase avec la tension, celui-ci est inutile lorsque $x_1 = 0$ et $x_2 = 0$. D'où le choix de $\Delta_{opt} = -1$.

Question 29 Commenter le choix $\Delta_{opt} = -1$ pour la combinaison telles que $(x_1, x_2, x_3, x_4) = (2, 2, 2, 2)$.

Je ne vois pas ce qu'attendent les auteur·rice·s du sujet.

III.2 - Contrôle du paramètre Δ avec un réseau de neurones

Question 30 Proposer un mode de communication pour que les différents appareils électriques de l'installation puissent transmettre leurs modes de fonctionnement au contrôleur. Aucun détail technique n'est attendu pour cette question.

On peut utiliser un réseau Wi-Fi ou alors un réseau par Courant Porteur en Ligne (CPL)

Question 31 Quels sont les avantages et inconvénients du choix d'un réseau de neurones pour modéliser la commande ?

Les réseaux de neurones permettent de traiter un grand nombre de données et de résoudre des problèmes complexes. En revanche, ils ont un traitement des ces données opaque qui ne permet pas de comprendre le fonctionnement de l'apprentissage.

Remarque : Le choix d'un réseau de neurones est ici peu pertinent car on a que très peu de données.

Question 32 L'apprentissage supervisé est-il pertinent ? Justifier votre réponse.

Un apprentissage supervisé est ici adapté car on a des données étiquetées : Couples entrée/sortie connus.

III.3 - Définition du perceptron

Question 33 Calculer la dérivée de la fonction f' notée $f'(x)$ et écrire une fonction Python notée f_prime qui prend en argument x et qui renvoie la dérivée $f'(x)$. L'argument d'entrée x est de type tableau.

Par dérivation de la fonction $f(x) = \frac{1}{1 + e^{-x}}$, on obtient : $f'(x) = \frac{e^{-x}}{(1 + e^{-x})^2}$

D'où le code de la fonction Python :

```
def f_prime(x) :  
    return np.exp(-x)/(1+np.exp(-x))**2
```

III.4 - Phase d'inférence

Question 34 Donner les dimensions des matrices W_2 , W_3 , B_2 et b_3 .

W_2 : Dimensions 4×4

W_3 : Dimensions 1×4

B_2 : Dimensions 4×1

b_3 : Dimensions 1×1

Question 35 Donner l'expression de la matrice W_1 qui vérifie :

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = W_1 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} - & - & - & - \\ - & - & - & - \\ - & - & - & - \\ - & - & - & - \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$$

$$W_1 = \begin{bmatrix} w_{1,1} & w_{1,2} & w_{1,3} & w_{1,4} \\ w_{2,1} & w_{2,2} & w_{2,3} & w_{2,4} \\ w_{3,1} & w_{3,2} & w_{3,3} & w_{3,4} \\ w_{4,1} & w_{4,2} & w_{4,3} & w_{4,4} \end{bmatrix}$$

Question 36 Ecrire en Python la fonction *inference_couche* qui va prendre en argument le vecteur d'entrées noté X , la matrice des poids notée W , le vecteur de biais noté B et qui renvoie le vecteur des sorties de la couche noté A .

```
def inference_couche(X,W,B) :  
    return f(np.dot(W,X) + B)
```

Question 37 Calculer $A = f(Y)$ dans le cas où :

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad W_1 = \begin{bmatrix} 0,3 & 0,2 & 0,25 & -0,04 \\ 0,4 & -0,3 & 0,6 & -0,3 \\ -0,4 & 0,3 & -0,6 & -0,4 \\ -1 & -0,75 & -0,1 & -0,5 \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$Y = \begin{bmatrix} 0,3 & 0,2 & 0,25 & -0,04 \\ 0,4 & -0,3 & 0,6 & -0,3 \\ -0,4 & 0,3 & -0,6 & -0,4 \\ -1 & -0,75 & -0,1 & -0,5 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0,2 \\ -0,3 \\ 0,3 \\ -0,75 \end{bmatrix} \quad A = f(Y) = \begin{bmatrix} \frac{1}{1 + e^{-0,2}} \\ \frac{1}{1 + e^{0,3}} \\ \frac{1}{1 + e^{-0,3}} \\ \frac{1}{1 + e^{0,75}} \end{bmatrix}$$

III.5 - Rétropropagation

Question 38 Montrer que : $\frac{\partial \mathcal{L}}{\partial w_{ij}} = \frac{\partial \mathcal{L}}{\partial a_i} \cdot f'(y_i) \cdot x_j$.

$$\frac{\partial \mathcal{L}}{\partial w_{ij}} = \frac{\partial \mathcal{L}}{\partial a_i} \cdot \frac{\partial a_i}{\partial w_{ij}} = \frac{\partial \mathcal{L}}{\partial a_i} \cdot \frac{\partial f(y_i)}{\partial w_{ij}} = \frac{\partial \mathcal{L}}{\partial a_i} \cdot \frac{\partial f\left(\sum_{k=1}^n (w_{ik} \cdot x_k) + b_i\right)}{\partial w_{ij}} = \frac{\partial \mathcal{L}}{\partial a_i} \cdot \frac{\partial \left(\sum_{k=1}^n (w_{ik} \cdot x_k) + b_i\right)}{\partial w_{ij}} \cdot f'\left(\sum_{k=1}^n (w_{ik} \cdot x_k) + b_i\right)$$

où n est le nombre d'entrées du $i^{\text{ème}}$ neurone de la couche considérée.

Or : $\frac{\partial \left(\sum_{k=1}^n (w_{ik}x_k) + b_i \right)}{\partial w_{ij}} = x_j$ Donc : $\frac{\partial \mathcal{L}}{\partial w_{ij}} = \frac{\partial \mathcal{L}}{\partial a_i} \cdot x_j \cdot f' \left(\sum_{k=1}^n (w_{ik}x_k) + b_i \right)$ avec : $y_i = \sum_{k=1}^n (w_{ik}x_k) + b_i$

On en déduit : $\frac{\partial \mathcal{L}}{\partial w_{ij}} = \frac{\partial \mathcal{L}}{\partial a_i} \cdot x_j \cdot f'(y_i)$

Question 39 En utilisant la relation précédente, donner la relation matricielle entre $\frac{\partial \mathcal{L}}{\partial W}$, $\frac{\partial \mathcal{L}}{\partial A}$, $f'(Y)$ et X , puis entre $\frac{\partial \mathcal{L}}{\partial W}$, $\frac{\partial \mathcal{L}}{\partial A}$, W , X et B .

En notant m le nombre de neurones et n le nombre d'entrées des neurones de la couche considérée, on a :

$$\left(\frac{\partial \mathcal{L}}{\partial A} * Y \right) \cdot X^T = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial a_1} \\ \vdots \\ \frac{\partial \mathcal{L}}{\partial a_i} \\ \vdots \\ \frac{\partial \mathcal{L}}{\partial a_m} \end{bmatrix} * f' \begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} \cdot [x_1 \dots x_j \dots x_n] = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial a_1} \cdot f'(y_1) \\ \vdots \\ \frac{\partial \mathcal{L}}{\partial a_i} \cdot f'(y_i) \\ \vdots \\ \frac{\partial \mathcal{L}}{\partial a_m} \cdot f'(y_m) \end{bmatrix} \cdot [x_1 \dots x_j \dots x_n]$$

$$\left(\frac{\partial \mathcal{L}}{\partial A} * Y \right) \cdot X^T = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial a_1} \cdot f'(y_1) \cdot x_1 & \dots & \frac{\partial \mathcal{L}}{\partial a_1} \cdot f'(y_1) \cdot x_j & \dots & \frac{\partial \mathcal{L}}{\partial a_1} \cdot f'(y_1) \cdot x_n \\ \vdots & & \vdots & & \vdots \\ \frac{\partial \mathcal{L}}{\partial a_i} \cdot f'(y_i) \cdot x_1 & \dots & \frac{\partial \mathcal{L}}{\partial a_i} \cdot f'(y_i) \cdot x_j & \dots & \frac{\partial \mathcal{L}}{\partial a_i} \cdot f'(y_i) \cdot x_n \\ \vdots & & \vdots & & \vdots \\ \frac{\partial \mathcal{L}}{\partial a_m} \cdot f'(y_m) \cdot x_1 & \dots & \frac{\partial \mathcal{L}}{\partial a_m} \cdot f'(y_m) \cdot x_j & \dots & \frac{\partial \mathcal{L}}{\partial a_m} \cdot f'(y_m) \cdot x_n \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial w_{11}} & \dots & \frac{\partial \mathcal{L}}{\partial w_{1j}} & \dots & \frac{\partial \mathcal{L}}{\partial w_{1n}} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial \mathcal{L}}{\partial w_{i1}} & \dots & \frac{\partial \mathcal{L}}{\partial w_{ij}} & \dots & \frac{\partial \mathcal{L}}{\partial w_{in}} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial \mathcal{L}}{\partial w_{m1}} & \dots & \frac{\partial \mathcal{L}}{\partial w_{mj}} & \dots & \frac{\partial \mathcal{L}}{\partial w_{mn}} \end{bmatrix}$$

On en déduit donc que : $\frac{\partial \mathcal{L}}{\partial W} = \left(\frac{\partial \mathcal{L}}{\partial A} * f'(Y) \right) \cdot X^T$ puis : $\frac{\partial \mathcal{L}}{\partial W} = \left(\frac{\partial \mathcal{L}}{\partial A} * f'(W \cdot X + B) \right) \cdot X^T$

On notera que le « * » représente un produit terme à terme alors que le « . » représente le produit matriciel.

Question 40 Ecrire la fonction Python rétropropagation-couche qui prend en argument :

- la valeur d'entrée d'une couche, notée X ,
- la matrice de poids, d'une couche, notée W ,
- le vecteur de biais de la couche, noté B ,
- le vecteur d'erreur en sortie de la couche, noté Ea ,
- la valeur du coefficient d'apprentissage, notée α .

Cette fonction doit calculer :

- $E_w = \frac{\partial \mathcal{L}}{\partial W}$
- $E_B = \frac{\partial \mathcal{L}}{\partial B}$
- $E_X = \frac{\partial \mathcal{L}}{\partial X}$

Cette fonction doit renvoyer :

- la matrice de poids W mise à jour,
- la matrice de biais B mise à jour,
- le vecteur gradient du coût $E_X = \frac{\partial \mathcal{L}}{\partial X}$

```
def inference_couche(X,W,B,Ea,alpha) :
```

```
    Y=np.dot(W,X)+B
```

```
    F_prime_Y=f_prime(Y)
```

```
    Ew=np.dot(F_prime_Y*Ea,X.T)
```

```
    Eb=F_prime_Y*Ea
```

```
    Ex=np.dot(W,F_prime_Y)*Ea
```

```
    W=W-alpha*Ew
```

```
    B=B-alpha*Eb
```

```
    return (W,B,Ex)
```

III.6 - Analyse des résultats

Question 41 Analyser les résultats. Ce réseau de neurones est-il capable de bien prédire la valeur optimale de Δ ?

On voit sur les données test que la valeur prédite de Δ (0,015) est très loin de la valeur optimale (0,22). Cela représente 20 % du jeu de test pour laquelle la valeur prédite est éloignée de la valeur réelle. On ne peut donc pas dire que ce réseau de neurones puisse prédire correctement la valeur optimale de Δ .

Question 42 Pour améliorer le modèle est-il préférable d'augmenter le nombre d'itérations de l'entraînement ou d'augmenter le nombre de données de test ? Justifier votre réponse.

On voit sur la courbe d'évolution de la fonction d'erreur que celle-ci se stabilise très rapidement dès 10 à 15 itérations. Il n'est donc pas nécessaire d'augmenter le nombre d'itération pour améliorer les performances de l'algorithme de prédiction.

En revanche, il était assez évident qu'avec seulement 9 données d'entraînement l'apprentissage ne pouvait pas être efficace. Il est donc préférable d'augmenter le nombre de données d'entraînement. (Et pas de données test comme le suggère l'énoncé).