TP

Effectuer une régression linéaire

I Méthode pour faire une régression linéaire

Nous introduisons la notion de régression linéaire à travers l'exemple du TP 5. On étudie un montage multivibrateur astable, et on s'intéresse en particulier à la période des oscillations du signal produit par le circuit.

Côté modélisation théorique : quelle est la loi que l'on veut vérifier ?

L'étude théorique prédit la loi suivante pour la période T des oscillations :

$$T = \frac{4RR_2C}{R_1}. \tag{1}$$

On peut choisir de vérifier plusieurs choses. Par exemple :

- Vérifier que la période T est bien proportionnelle à C. ($\rightarrow$ c'est ce que vous allez faire en TP)
- Vérifier que la période T est bien proportionnelle à R_2. ($\rightarrow$ c'est l'exemple que l'on développe ici)

▶$_{\text{étape 1}}$ On met la loi à vérifier sous la forme $\boxed{y = a_{\text{théo}}x + b_{\text{théo}}.}$

Ici on pose $y = T$, $x = R_2$ (ce sont les paramètres qui varient).

La loi 1 s'écrit alors $y = \dfrac{4RC}{R_1} \times x$.

On identifie alors $a_{\text{théo}}$ et $b_{\text{théo}}$: ici on a $a_{\text{théo}} = \dfrac{4RC}{R_1}$ et $b_{\text{théo}} = 0$.

▶$_{\text{étape 2}}$ On fait l'application numérique pour $a_{\text{théo}}$ et $b_{\text{théo}}$, avec incertitudes si besoin.

⋆ Ici on choisit $R = R_1 = 10\,\text{k}\Omega$, $C = 100\,\text{nF}$.

Donc $a_{\text{théo}} = 400\,\text{nF}$.

On prend une incertitude de 2% pour R et R_1, et de même sur C, et on utilise la formule de propagation des incertitudes pour obtenir $\Delta a_{\text{théo}} = 0.2\,\text{nF}$.

On a donc finalement $\boxed{a_{\text{théo}} = (4.0 \pm 0.2) \times 10^2\,\text{nF}.}$ ⋆ On a de plus $\boxed{b_{\text{théo}} = 0\,\text{s}}$.

Côté expériences : comment vérifier la loi ?

▶$_{\text{étape 3}}$ On réalise des mesures expérimentales de y (ici la période T, à l'aide de l'oscilloscope) pour plusieurs valeurs de x (donc ici de R_2).

On entre ces valeurs sous le logiciel Regressi (par exemple) dans un tableau à deux entrées.

Remarque : On peut entrer également l'incertitude sur chaque mesure expérimentale (mais c'est facultatif), en cliquant sur "incertitude" en haut. Ici on peut prendre 2% pour R_2 et 2% pour la mesure de T à l'oscilloscope. On trace alors la variable y (donc ici les valeurs de T) en fonction de la valeur de x (donc ici les valeurs de R_2).

▶étape 4 On confirme d'abord *visuellement* que les points forment bien approximativement une droite.

– Si c'est le cas, on peut déjà **conclure que le modèle linéaire ou affine est le bon**, et on passe à l'étape suivante.

– Si ce n'est pas le cas on arrête ici et on dit que le modèle linéaire ou affine n'est pas en accord avec les mesures.

▶étape 5 On demande au logiciel d'effectuer une modélisation, affine dans notre cas ($y = ax + b$).

Le logiciel utilise alors un algorithme qui calcule l'équation de la droite qui minimise la distance entre les points expérimentaux et la droite. (Il trouve a et b tels que l'écart entre la droite et les points soit minimal, cet écart étant défini comme $C(a,b) = \sum_i [y_i - (ax_i + b)]^2$.)

Il retourne les valeurs de a et de b, assortis d'une incertitude (même si on n'a pas entré d'incertitude sur les valeurs expérimentales).

Il s'agit donc des **valeurs expérimentales** de a et b, que l'on va noter $a_{\text{exp}} \pm \Delta a_{\text{exp}}$ $b_{\text{exp}} \pm \Delta b_{\text{exp}}$.

Ici on trouve, avec le jeu de données utilisé : $a_{\text{exp}} = (376 \pm 48)$ S.I. et $b_{\text{exp}} = (28.9 \pm 89) \times 10^{-6}$ S.I.. On sait par ailleurs que l'unité de a est le Farad et celle de b les secondes.

On écrira donc : $\boxed{a_{\text{exp}} = (3.8 \pm 0.5) \times 10^2\,\text{nF}}$ et $\boxed{b_{\text{exp}} = (3 \pm 9) \times 10^{-5}\,\text{s}}$

Comparaison entre théorie et expérience

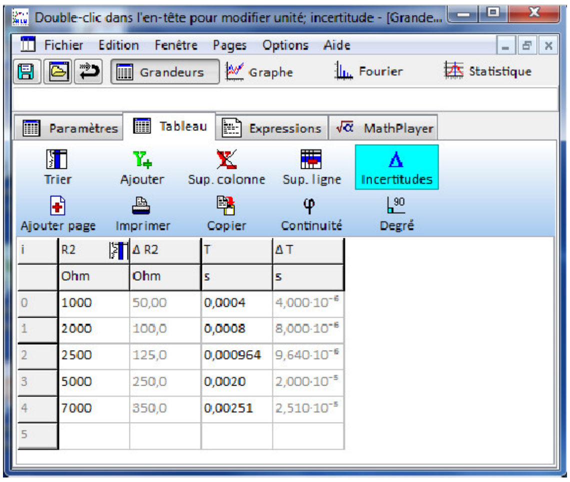
▶étape 6 On conclue en comparant théorie et expérience.

On voit dans le cas présent que les intervalles $a_{\text{exp}} \pm \Delta a_{\text{exp}}$ et $a_{\text{théo}} \pm \Delta a_{\text{théo}}$ ont des valeurs en commun. De même, $b_{\text{théo}}$ est inclus dans l'intervalle $b_{\text{exp}} \pm \Delta b_{\text{exp}}$.

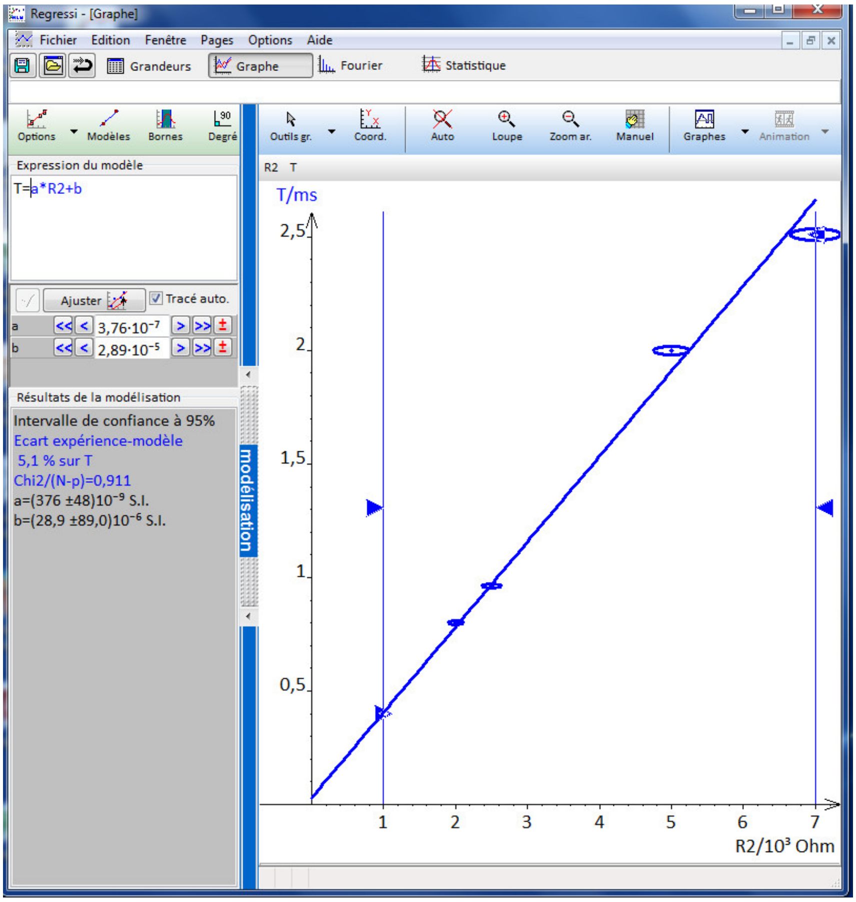
On peut donc conclure que *la modélisation est en accord avec la théorie.*

Remarque : Si il n'y a pas accord entre expérience et théorie *de façon grossière*, il faut dire qu'on a probablement fait une erreur quelque part, ou qu'un composant est défectueux. Si il n'y a pas accord, mais que les valeurs sont tout de même proche, il faut dire que : soit on a fait une erreur, soit le modèle théorique ne prend pas en compte tous les aspects expérimentaux. (Par exemple dans le cas présent, pour des petites périodes on voit à l'oscilloscope que la sortie de l'hystérésis met du temps à basculer entre $\pm V_{\text{sat}}$, alors que dans le modèle on suppose ceci instantané.)

Remarque : Il est aussi possible de faire directement une régression avec une loi linéaire (du type $y = ax$). Il faut alors vérifier visuellement qu'il n'y a pas d'ordonnée à l'origine. Avec les données prises ici le logiciel donne $a_{\text{exp}} = (389 \pm 24) \times 10^{-9}\,\text{nF}$, ce qui convient.



Capture d'écran du logiciel Regressi, montrant les valeurs utilisées.



II Cas où la loi à vérifier n'est pas linéaire

Il est possible que la loi à vérifier ne soit pas linéaire. Par exemple supposons que l'on veuille vérifier que la vitesse du son dans l'air est bien proportionnelle à la racine carrée de la température : $c_s = \alpha\sqrt{T}$, avec α une constante.

- On effectue des mesures de la vitesse du son c_s pour différentes températures T. On les rentre dans Régressi ou dans la calculatrice.
- On essaie de se ramener à une loi linéaire. La raison est que le cerveau humain est doué pour voir si des points sont alignés, mais pas du tout pour estimer s'ils suivent une loi en racine carrée ou autre.

  Ici par exemple on peut décider de tracer c_s en fonction de $\sqrt{T}$. En effet, si on pose $y = c_s$ et $x = \sqrt{T}$, alors la relation à vérifiée devient $y = \alpha x$, et on peut employer la méthode précédente.

  Concrètement, sous Régressi on crée une nouvelle variable dans l'onglet expression : $x = T^{0.5}$. Le logiciel calcule alors x pour toutes les valeurs de T.

  On trace alors c_s en fonction de x, et on fait la régression linéaire sur ce graphique.

Il se peut dans certains cas qu'on ne puisse pas se ramener à une loi affine ou linéaire. Dans ce cas on utilise les autres modèles disponibles dans le logiciel, tout en regardant à l'œil si le modèle est vraisemblable.

III Exemple

On reprend l'exemple précédent de la température. On a le tableau de valeurs expérimentales suivant :

c_s (m/s)	310	325	331	345	355
T (K)	240	260	280	300	320

À l'aide de votre calculatrice, vérifier si ces données expérimentales sont en accord avec la loi théorique $c_s = 20.05\sqrt{T}$ (T en K, c_s en m/s).