Karnaj

Calcul approché d'intégrales

15 juin 2020

Table des matières

L'intégration est un domaine très utile des mathématiques. Il sert notamment en physique (le calcul d'une intégrale permet de calculer des positions, des travaux, etc.). Le calcul d'intégrales est donc important et les mathématiciens ont développé plusieurs techniques pour calculer de façon assez simple la valeur exacte d'une intégrale (calcul de primitives, intégration par parties, etc.).

Ces techniques permettent de calculer plusieurs intégrales, mais pas toutes, et il existe des fonctions dont on est incapable d'exprimer les primitives à l'aide des fonctions usuelles, même si on sait qu'elles en admettent une (parce qu'elles sont continues par exemple). L'un des exemples le plus connu est (cette intégrale apparaît en probabilité) :

$$\int_a^b \mathrm{e}^{-t^2}\mathrm{d}t.$$

En informatique, calculer la valeur exacte d'une intégrale est encore plus compliqué puisqu'on ne peut pas (en tout cas pas facilement) appliquer les techniques dont nous venons de parler.

Une solution pour quand même avoir la valeur de notre intégrale est d'en calculer une valeur approchée. Dans ce tutoriel, nous allons donc voir plusieurs méthodes pour calculer une valeur approchée d'une intégrale.

> Ce tutoriel requiert quelques notions en mathématiques et plus particulièrement en analyse: il faut savoir calculer une intégrale et faire une intégration par partie. D'autres connaissances peuvent être un plus mais ne sont pas nécessaires pour comprendre le tutoriel. Des connaissances à propos de la complexité des algorithmes (voir ce tutoriel) pourront par exemple être utile mais ne sont pas indispensables. Ce tutoriel s'adresse principalement à un public en fin de lycée.

1. Introduction au problème

Dans ce tutoriel, nous allons nous fixer pour but de calculer numériquement n'importe quelle intégrale. Pour illustrer ce tutoriel, nous allons nous baser sur le calcul de l'intégrale de la fonction f qui à x associe e^{-x^2} sur l'intervalle $[-1, 3]$. Nous allons alors trouver des algorithmes permettant d'approcher le mieux possible la valeur de son intégrale.

Bien sûr, puisque nous ne connaissons pas la valeur de cette intégrale, nous ne pourrons pas vérifier que les algorithmes vus donnent bien le bon résultat. Nous allons alors les vérifier sur une autre intégrale, à savoir celle de la fonction cosinus sur l'intervalle $[0, \pi]$.

1.1. Présentation du problème

Nous voulons calculer numériquement l'intégrale de f, le problème étant que nous n'en connaissons pas de primitive explicite. La situation est mauvaise, et pour le moment, nous n'avons aucune idée de comment calculer cette intégrale.

1.1.1. Changeons d'aire

Pourtant, un simple changement de point de vue rend ce problème plus simple. Plaçons-nous d'un point de vue graphique en nous rappelant que **l'intégrale d'une fonction correspond à l'aire sous sa courbe représentative**. Ainsi, calculer l'intégrale d'une fonction revient à trouver la surface sous sa courbe représentative. La question qui vient alors est la suivante.

> ? Comment calculer cette aire?

Nous pourrions tourner en rond en disant qu'il s'agit de l'intégrale de la fonction associée à cette courbe, mais ce ne serait pas très productif…

Il nous faut ruser. Nous avons le graphe de la fonction à notre disposition, nous devrions alors être capable d'en tirer quelque chose et de calculer l'aire voulue. Bien sûr l'histoire n'est pas aussi simple et obtenir une **valeur exacte** de cette aire est compliqué.

Nous devrons nous contenter d'une **valeur approchée** de cette aire. Pour cela, nous allons approcher la courbe de notre fonction par la courbe d'une autre fonction, dont nous connaissons la valeur de l'aire. Dis comme ça, ça à l'air simple, mais encore faut-il savoir par quelle fonction nous allons l'approcher. Nous allons voir qu'il existe plusieurs manières de le faire. Certaines sont plus précises et permettent d'avoir une meilleure approximation en moins de calcul (bien sûr, tout cela dépend également de la forme de la fonction que l'on veut intégrer et de ce qu'on connaît d'elle).

1.2. Approximation de la courbe

Nous avons décidé d'approcher la courbe de notre fonction par la courbe d'une autre fonction dont nous connaissons l'intégrale. Il nous reste à trouver quelle fonction nous allons utiliser pour cela. Regardons la courbe représentative de la fonction, peut-être nous donnera-t-elle une idée.

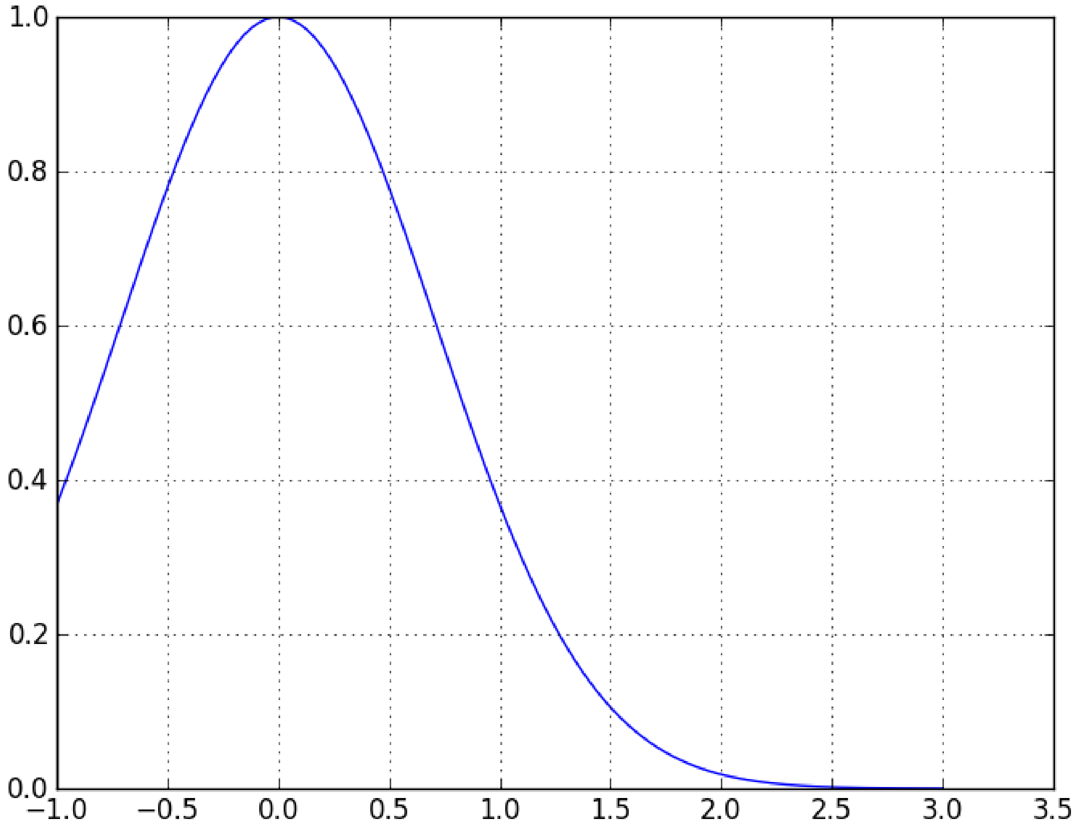

FIGURE 1.1. – La fonction f sur l'intervalle $[-1, 3]$.

Trouver une fonction autre que f dont la courbe approche suffisamment bien cette courbe semble compliqué. Là encore, il va nous falloir ruser.

1.2.1. Découpage de la fonction

Plutôt que de chercher une fonction qui approche f sur tout l'intervalle, nous allons découper l'intervalle en plusieurs morceaux et approcher la fonction sur chacun des intervalles par une fonction différente. Le travail est alors plus facile.

Par exemple, pour intégrer f de -1 à 3, on peut l'approcher par une fonction différente sur chaque intervalle de longueur 0,5. Le plus simple est de l'approcher par une fonction constante de manière à avoir à chaque fois l'aire d'un rectangle à calculer. On obtient alors cette approximation.

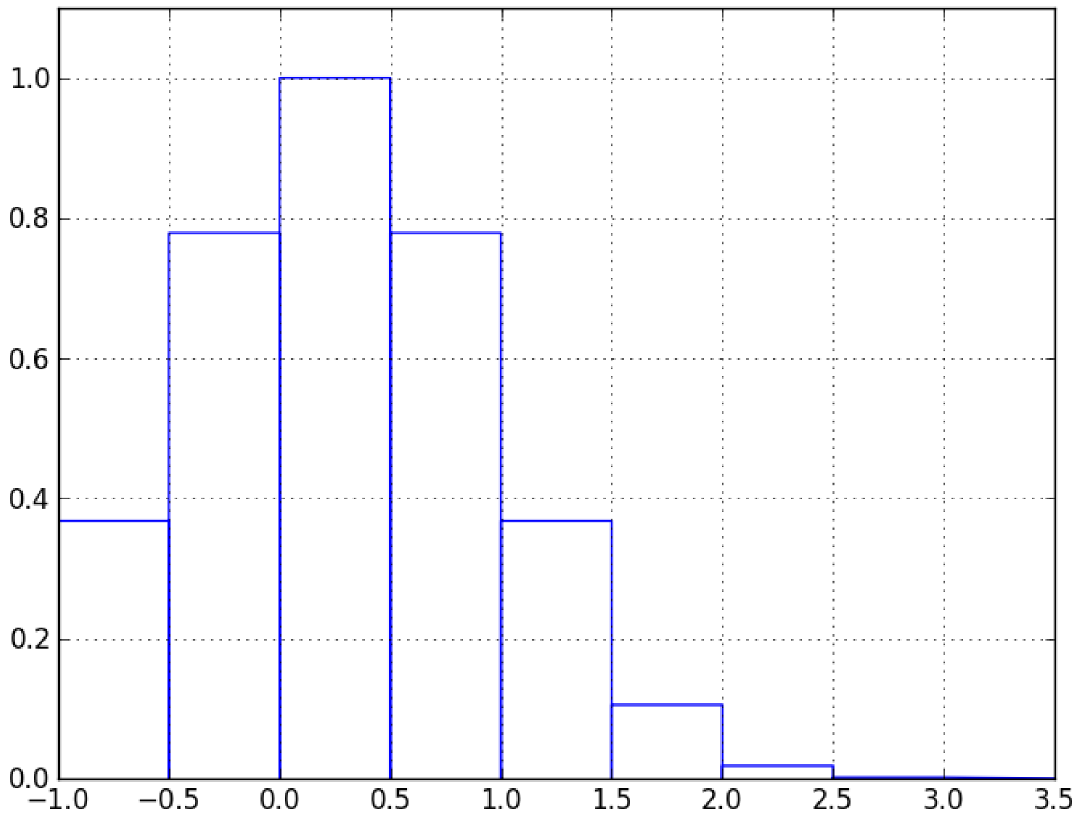

Figure 1.2. – On essaie avec des rectangles?

?

Quoi? Mais ça n'approche pas bien du tout la courbe. Les aires seront beaucoup trop différentes.

C'est vrai que les deux courbes ne sont pas trop proches. Mais en diminuant la longueur des intervalles, les deux courbes sont de plus en plus proches, comme on peut le voir sur le graphe ci-dessous (en bleu la courbe représentative de la fonction et en vert les rectangles).

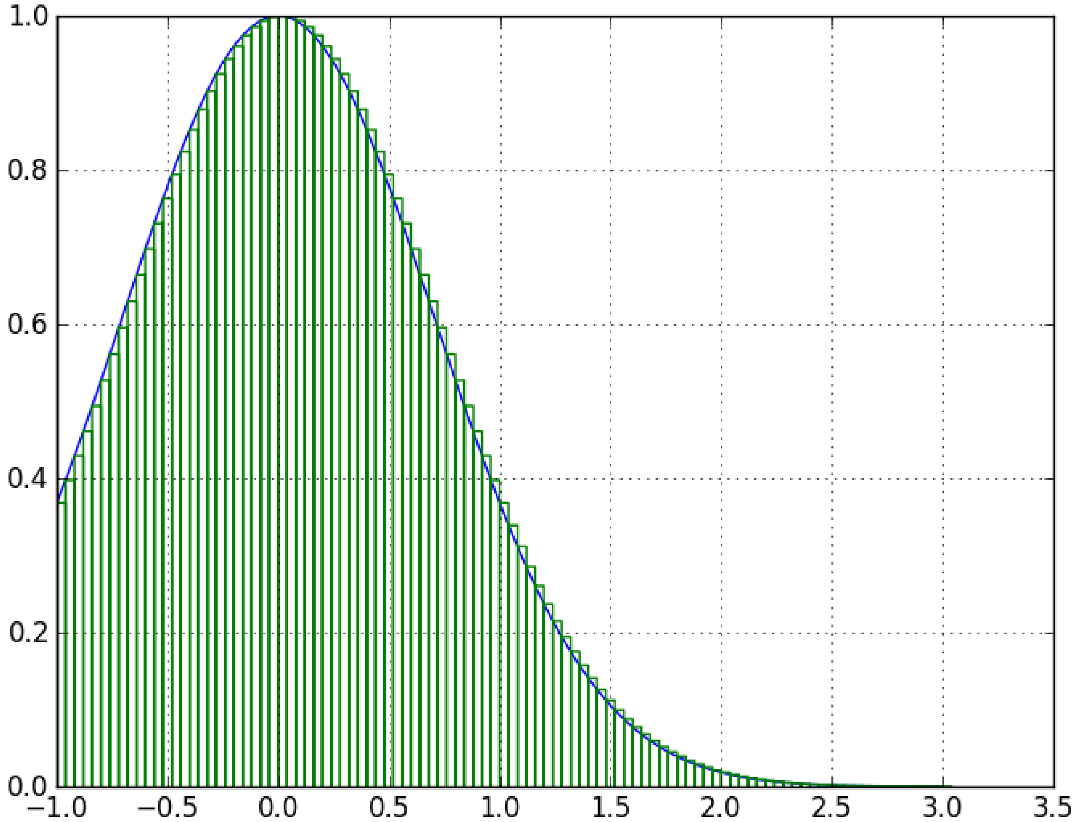

FIGURE 1.3. – Là c'est déjà mieux.

1.2.2. Calcul de l'aire

Il nous faut maintenant calculer l'aire approchée. On va séparer notre intervalle en 8 et calculer l'aire des 8 rectangles obtenus. On a que la largeur de chaque rectangle est $\frac{3-(-1)}{8} = 0{,}5$. Il nous faut maintenant les hauteurs:

— la hauteur du premier rectangle est $f(-1)$;
— la hauteur du deuxième rectangle est $f(-0{,}5)$;
— la hauteur du troisième rectangle est $f(0)$.

En fait, la hauteur du k$^{\text{ème}}$ rectangle est $f(-1+0{,}5k)$. On additionne ces différentes aires.

$$ I \approx \sum*{k = 0}^{8} 0{,}5 \times f(-1 + 0{,}5 k) = 0{,}5 \sum*{k = 0}^{8} f(-1 + 0{,}5 k). $$

> *i*
> Le symbole $\sum$ introduit ici est le symbole de la somme. $\sum_{k=0}^{8} f(-1+0{,}5k)$ se lit comme «la somme pour k allant de 0 à 10 des $f(-1+0{,}5k)$», c'est-à-dire comme
> $$ f(-1) + f(-0{,}5) + f(0) + \ldots + f(2{,}5) + f(3).$$

1.3. Formalisation

Maintenant que nous avons une méthode, il ne nous reste plus qu'à la formaliser. Considérons alors les fonctions continues sur l'intervalle $[a, b]$. Cet intervalle étant un intervalle fermé borné, alors, d'après le **théorème des bornes**, l'image de $[a, b]$ par f est un intervalle fermé borné. Ceci nous assure que f ne «parte pas vers l'infini» (pour ne pas avoir de rectangle d'aire infinie).

!

Certaines fonctions ont des limites infinies et leur intégrale est quand même finie ($\int_0^1 \frac{\sqrt{t}}{t} \mathrm{d}t$ par exemple) mais nous n'allons pas en tenir compte ici.

Une fois que ceci est assuré, nous pouvons appliquer la méthode que nous avons trouvée. Rappelons-nous qu'elle consiste à faire deux choses.

1. Diviser l'intervalle $[a, b]$ en n petits intervalles de même longueur.
2. Remplacer l'aire de f sur ces intervalles par l'aire de petits rectangles.

1.3.1. Division de l'intervalle

Pour diviser $[a, b]$ en intervalles, posons pour tout entier k de $[\![0, n]\!]$,

$$x_k = a + k\frac{b-a}{n}.$$

On a ainsi une **subdivision** de notre intervalle. Les x_k sont les abscisses des points de cette subdivision.

Nommons un peu tout ce que nous venons d'utiliser:

— n est le nombres de tranches de la subdivision;
— $\delta = \frac{b-a}{n}$ est le pas de la subdivision (c'est la distance entre deux points).

Les x_k peuvent également être définis par récurrence:

$$\begin{cases} x_0 = a \\ \forall k \in [\![0, n-1]\!], x_{k+1} = x_k + \delta \end{cases}$$

1.3.2. Calcul de l'intégrale

Maintenant que nous avons subdivisé notre intervalle, nous pouvons utiliser la **relation de Chasles** pour remarquer ceci:

$$I(f) = \int_a^b f(t)\mathrm{d}t = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(t)\mathrm{d}t.$$

Il ne reste plus qu'à remplacer la fonction f sur chacun des intervalles $[x_k, x_{k+1}]$ par une fonction g_k constante.

On obtiendra alors une approximation de $I(f)$.

$$I(f) \approx \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} g_k(t)\mathrm{d}t.$$

Pour correspondre à la méthode que nous avons utilisée, il faut $g_k(x) = f(x_k)$. On a alors que l'aire de chaque rectangle est

$$R_k(f) = \int_{x_k}^{x_{k+1}} g_k(t)\mathrm{d}t = \int_{x_k}^{x_{k+1}} f(x_k)\mathrm{d}t.$$

L'intégrale d'une constante sur un intervalle est cette constante multipliée par la longueur de l'intervalle, donc

$$R_k(f) = (x_{k+1} - x_k)f(x_k) = \delta f(x_k).$$

On peut maintenant passer à l'approximation de l'aire totale sous la courbe:

$$I(f) \approx \sum_{k=0}^{n-1} R_k(x) = \sum_{k=0}^{n-1} \delta f(x_k).$$

On peut sortir la constante δ de la somme et on obtient comme résultat final:

$$\boxed{I(f) \approx \delta \sum_{k=0}^{n-1} f(x_k)}.$$

La méthode que nous venons de trouver porte un nom. Il s'agit de la **méthode des rectangles** qui consiste à remplacer la fonction f sur chaque $[x_k, x_{k+1}]$ par une fonction constante. Notons qu nous aurions aussi pu choisir $g_k(x) = f(k_{i+1})$. On peut alors faire ces deux méthodes des rectangles:

— la **méthode des rectangles à gauche** où $g_k(x) = f(x_k)$;
— la **méthode des rectangles à droite** où $g_k(x) = f(x_{k+1})$.

La méthode des rectangles correspond aux **sommes de Riemann** qui peuvent être utilisées pour définir la notion d'intégration.

1.4. Algorithme et vérification

Maintenant que nous avons trouvé une formule, utilisons-la pour écrire un algorithme en Python. Écrivons la méthode des rectangles à gauche.

```python
def rectangle(f, a, b, n):
    somme = 0
    pas = (b - a) / n
    x = a
    for i in range(0, n): # On calcule la somme des f(x_i)
        somme += f(x)
        x += pas
    return somme * pas    # On retourne cette somme fois le pas
```

Ici, `rectangle` prend en paramètre la fonction `f` à intégrer, les deux bornes d'intégration `a` et `b` et le nombre de tranches `n` à utiliser pour la subdivision.

Pour vérifier si notre algorithme fonctionne, utilisons le sur des fonctions dont nous connaissons l'intégrale. Par exemple,

$$\int_0^\pi \cos(t)\mathrm{d}t = [\sin(xt)]_0^\pi = 0.$$

On peut alors essayer de voir ce que donne `rectangle(cos, 0, pi, 100)`. Diviser l'intervalle en 100 semble suffisant pour obtenir une bonne valeur. Nous obtenons **0,0314**, associé à ce graphe.

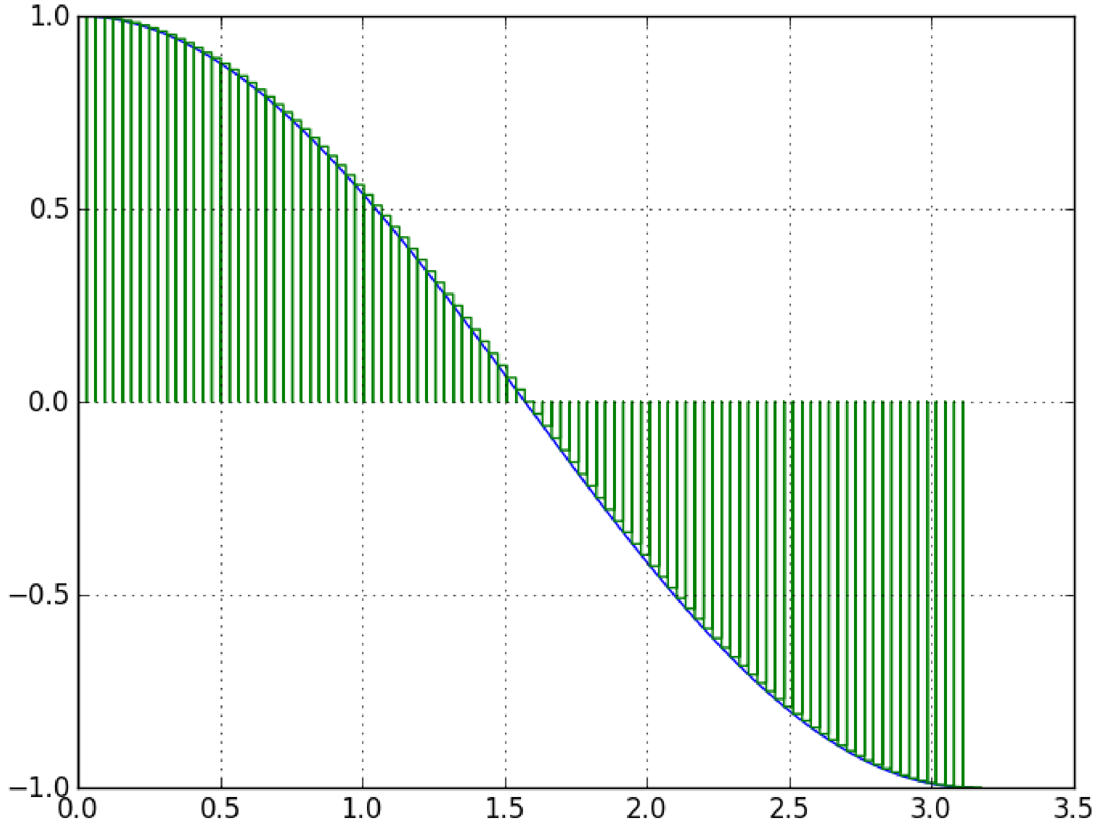

FIGURE 1.4. – La méthode des rectangles sur la fonction cosinus.

La méthode semble alors fonctionner. Mais ce petit test ne peut pas suffire à nous convaincre qu'elle donne des résultats satisfaisants dans tous les cas. Nous pouvons alors la tester sur d'autres fonctions pour voir ce qu'on obtient.

Reste à voir ce qu'elle donne sur notre exemple.

```python
from math import exp

def f(x):
    return exp(-(x*x))

print(rectangle(f, 0, 1, 100))
```

On obtient environ 0,749979.

2. Performance et précision

2.1. Les pentes c'est dangereux

2.1.1. Test de la méthode des rectangles

La méthode des rectangles semble fonctionner assez bien. Nous l'avons testé sur la fonction cosinus et en divisant l'intervalle en 100, nous avions obtenu un résultat vrai jusqu'à la première décimale. Mais qu'en est-il d'autres fonctions? Le résultat est-il toujours aussi bon?

Il paraît évident que si nous testons la méthode des rectangles sur une fonction constante, le résultat sera bon. En fait, dans ce cas, le résultat est exactement la valeur de l'intégrale et pas une valeur approchée. De même, si nous la testons sur une fonction en escalier, nous avons de bonnes chances d'obtenir un bon résultat (pourvu que le pas soit bien choisi).

Cependant, nous pouvons voir que si nous testons notre fonction sur une fonction qui croît rapidement par exemple, le résultat sera moins probant. Et plus la fonction croîtra vite, moins le résultat sera probant.

2.1.2. Augmenter la précision du résultat

Notre but est alors d'augmenter la précision de notre résultat. Pour cela, rappelons-nous du chapitre précédent. En diminuant le pas, nous augmentons la précision de notre résultat. En faisant de même ici, on devrait améliorer notre résultat.

Bien sûr, cette solution fonctionne, mais si l'on augmente encore la pente, on se retrouve encore avec une différence notable. De plus, augmenter le pas augmente également le temps de calcul. Sur certaines courbes, cette solution n'est pas viable. Il nous faut donc trouver une autre méthode pour gérer ce cas des fonctions à forte pente.

> *i* En fait, nos ordinateurs font des erreurs d'approximation lors des calculs, de sorte que si en théorie un pas infiniment petit est possible, en pratique, ça ne l'est pas. Pour plus d'informations à ce sujet, nous pouvons nous renseigner sur les erreurs de calcul des ordinateurs (voir cet article et celui-là).

2.2. La méthode des trapèzes

2.2.1. Principe de la méthode

La méthode des trapèzes consiste à remplacer la fonction f sur chaque intervalle par une fonction qui sera affine. Pour qu'elle approche bien la courbe, les fonctions que nous allons utiliser sont toutes trouvées: pour chaque x_k de la subdivision, nous allons prendre la fonction qui passe par les points $M_k(x_k, f(x_k))$ et $M_{k+1}(x_{k+1}, f(x_{k+1}))$.

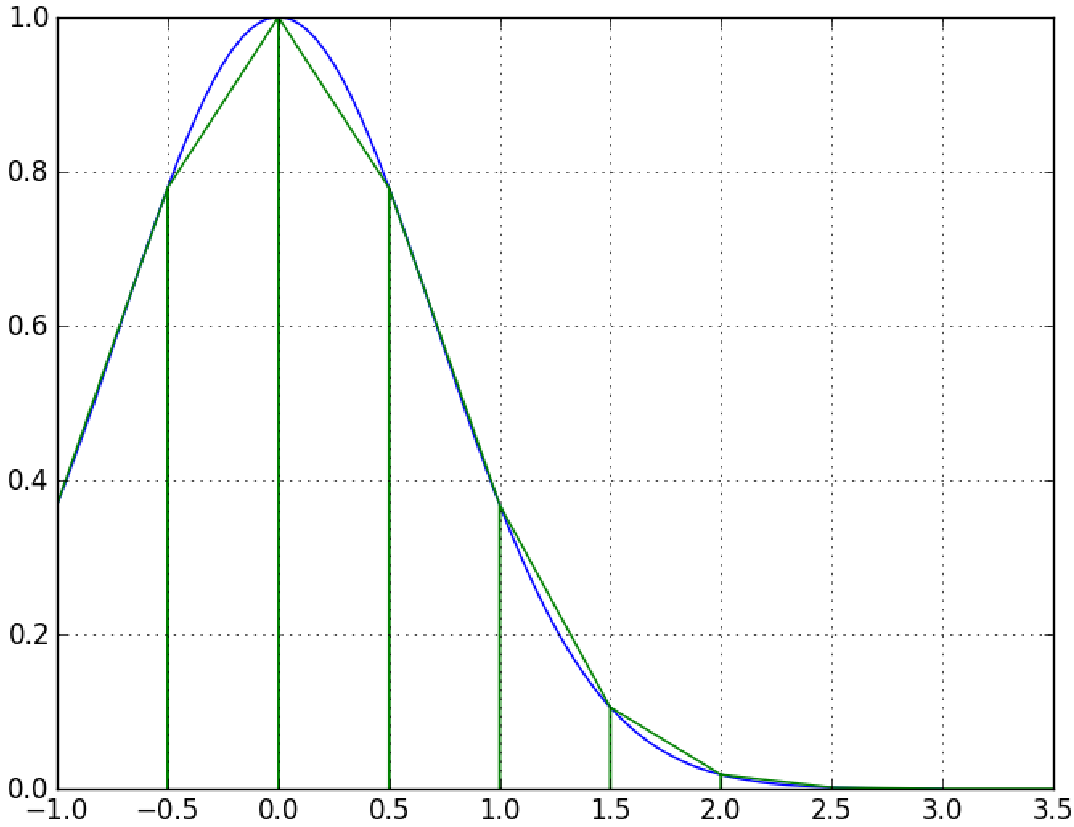

Figure 2.1. – Avec des trapèzes c'est mieux.

Nous voyons qu'en approchant la courbe de notre fonction par des trapèzes, on s'en rapproche plus que lorsque l'on utilisait des rectangles. Sur cette fonction, la méthode des trapèzes donne une meilleure approximation que celle des rectangles.

Maintenant que nous avons vu les choses graphiquement, trouvons les fonctions g_k correspondantes. On veut une fonction affine (donc de la forme $g_k(x) = cx + d$) telle que

$$\begin{cases} g_k(x_{k+1}) = f(x_{k+2}) \\ g_k(x_k) = f(x_k) \end{cases}.$$

Ces deux informations nous permettent de trouver la fonction. On a

$$g_k(x) = \frac{f(x_{k+1}) - f(x_k)}{x_{k+1} - x_k}(x - x_k) + f(x_k).$$

2.2.2. Calcul de l'intégrale

La connaissance des fonctions g_k nous permet de calculer les aires T_k des différents trapèzes:

$$T_k(f) = \int_{x_k}^{x_{k+1}} g_k(t)\mathrm{d}t = \int_{x_k}^{x_{k+1}} \frac{f(x_{k+1}) - f(x_k)}{x_{k+1} - x_k}(t - x_k) + f(x_k)\mathrm{d}t.$$

Nous pouvons nous amuser à calculer cette intégrale si nous le voulons (comment ça, ce n'est pas amusant), mais ce sera inutile. En effet, nous sommes malins. Et donc, nous préférons calculer l'aire du trapèze, plutôt que cette intégrale.

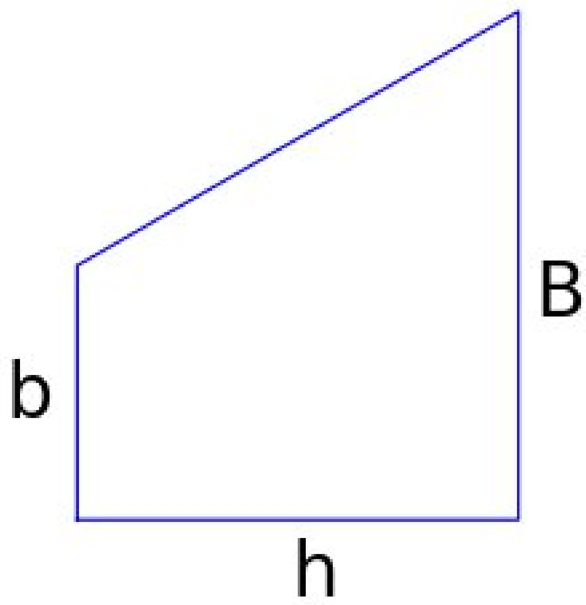

FIGURE 2.2. – L'aire du trapèze.

Pour ceux qui l'auraient oublié, l'aire d'un trapèze est: $h\frac{b+B}{2}$ (pour la retrouver, on peut découper le trapèze en un rectangle et un triangle rectangle).

On a donc dans notre cas:

- $h = x_{k+1} - x_k$;
- $b = f(x_k)$;
- $B = f(x_{k+1})$.

Et donc:

$$T_k(f) = x_{k+1} - x_k \frac{f(x_k) + f(x_{k+1})}{2}.$$

Et comme pour la méthode des rectangles, puisque $x_{k+1} - x_k = \delta$, on a:

$$T_k(f) = \delta \frac{f(x_k) + f(x_{k+1})}{2}.$$

Nous pouvons maintenant approcher $I(f)$:

$$I(f) \approx \sum_{k=0}^{n-1} T_k(f) = \sum_{k=0}^{n-1} \delta \frac{f(x_k) + f(x_{k+1})}{2}.$$

On peut sortir la constante $\frac{\delta}{2}$ de la somme:

$$I(f) \approx \frac{\delta}{2} \sum_{k=0}^{n-1} f(x_k) + f(x_{k+1}).$$

On remarque alors que tous les x_k sont comptés deux fois sauf le premier (x_0) et le dernier (x_n). Donc:

$$I(f) \approx \frac{\delta}{2} \left(f(x_0) + f(x_n) + 2 \sum_{k=1}^{n-1} f(x_k) \right).$$

Et finalement, en multipliant toute la somme par le 2, on obtient

$$\boxed{I(f) \approx \delta \left(\frac{f(x_0) + f(x_n)}{2} + \sum_{k=1}^{n-1} f(x_k) \right)}.$$

2.3. Algorithme de la méthode des trapèzes

Écrivons cet algorithme en Python.

```python
def trapèze(f, a, b, n):
    somme = (f(a) + f(b)) / 2   # On initialise la somme à (f(a) + f(b)) / 2
    pas = (b - a) / n
    x = a + pas                 # La somme commence à x_1
    for i in range(1, n):       # On calcule la somme des f(x_i)
        somme += f(x)
        x += pas
```

```
    return somme * pas          # On retourne cette somme fois le
        pas
```

Vérifions qu'il fonctionne en vérifiant (comme pour la méthode des rectangles) qu'il nous donne un résultat proche de 0 pour l'intégrale de 0 à π de la fonction cosinus. En découpant l'intervalle en 100 (donc avec `trapèze(cos, 0, pi, 100)`) on obtient un résultat de l'ordre de **-10^{-15}** là où avec la méthode des rectangles on obtenait un résultat de l'ordre de **10^{-2}**. La méthode des trapèzes nous donne un résultat beaucoup plus précis.

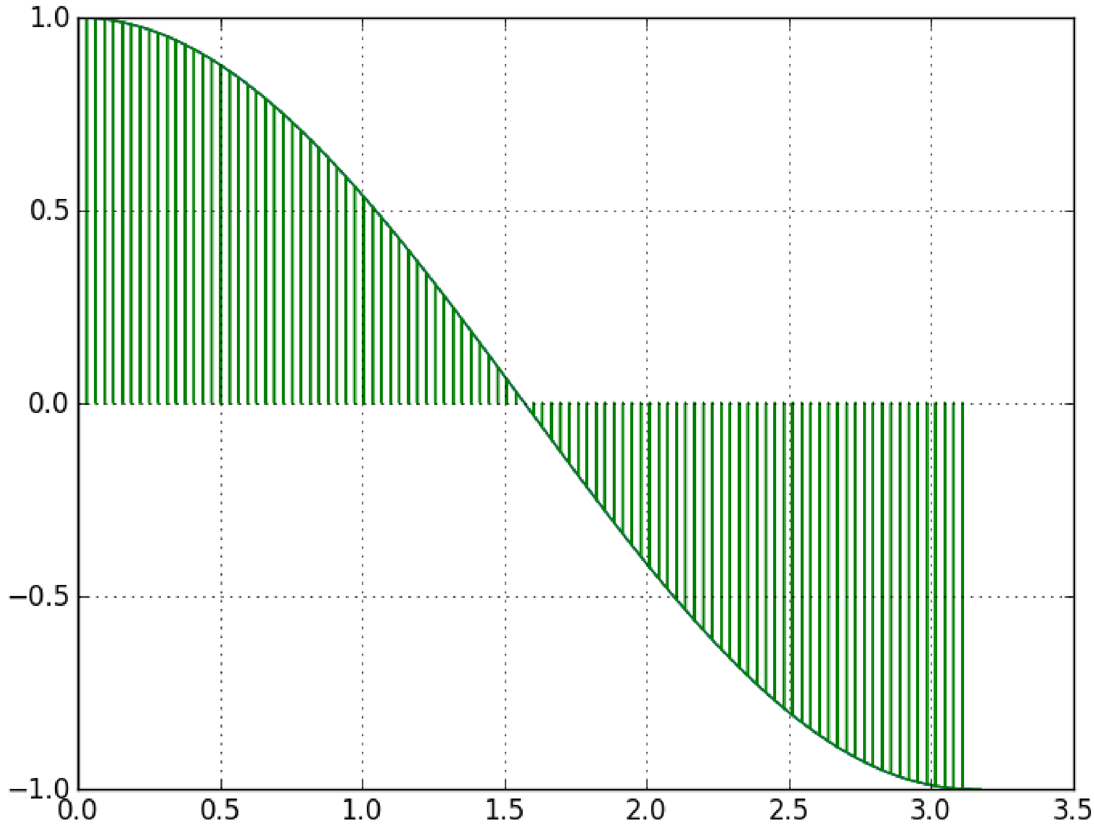

FIGURE 2.3. – La méthode des trapèzes sur la fonction cosinus.

Nous voyons qu'en effet, la courbe de la fonction est plus épousée par les trapèzes que par les rectangles.

Voyons voir ce qu'on obtient sur notre exemple.

```
print(trapèze(f, 0, 1, 100))
```

On obtient environ 0,746818.

3. La méthode de Simpson

Comme nous l'avons vu dans le chapitre précédent, la méthode des trapèzes est plus précise que la méthode des rectangles. Notre but dans ce chapitre est de trouver une manière de se rapprocher encore plus de la courbe.

3.1. Principe de la méthode

La méthode de Simpson consiste à remplacer la fonction f sur chaque intervalle par une fonction qui sera un trinôme. En effet, nous allons l'approcher par un polynôme de degré 2 qui passe pas les points $M_k(x_k, f(x_k))$, $M_{k+1}(x_{k+1}, f(x_{k+1}))$ et par le point au milieu $M_{k,5}(x_{k,5}, f(x_{k,5}))$ avec $x_{k,5} = \frac{x_k+x_{k+1}}{2}$.

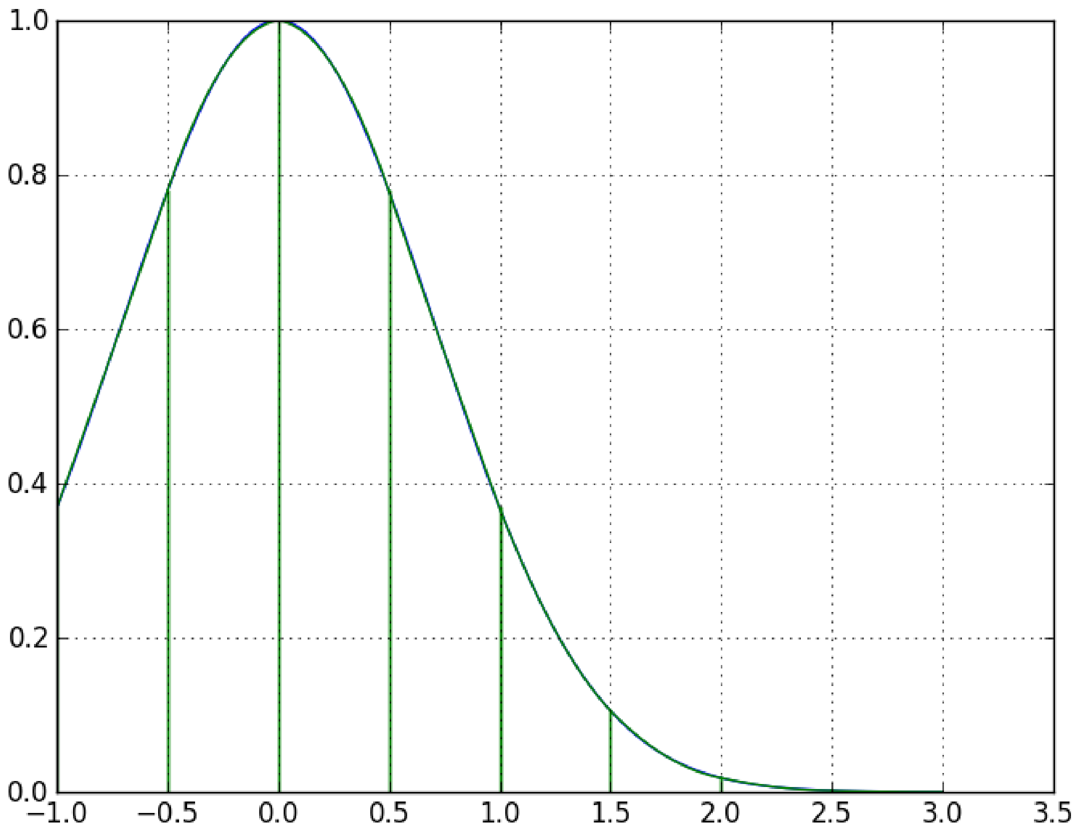

Figure 3.1. – Les trinômes, ça marche.

La courbe de la fonction f est quasiment confondue avec les courbes des fonctions g_k.

Cependant, nous ne savons toujours pas que sont nos polynômes. En fait, il existe un unique polynôme de degré 2 qui vérifie les conditions que nous lui imposons. Pour le construire nous allons utiliser les **polynômes interpolateurs de Lagrange** . Si vous ne les connaissez pas, ça ne fait rien, nous allons construire notre polynôme pas par pas.

Récapitulons ce que nous demandons au polynôme Q_k.

- L'image de x_k par la fonction polynomiale associée doit être $f(x_k)$.
- L'image de $x_{k,5}$ par la fonction polynomiale associée doit être $f(x_{k,5})$.
- L'image de x_{k+1} par la fonction polynomiale associée doit être $f(x_{k+1})$.

Les autres points ne nous importent donc pas.

Pour construire ce polynôme, nous allons construire trois polynômes :

- L_k tel que $L_k(x_k) = 1$, $L_k(x_{k,5}) = 0$ et $L_k(x_{k+1}) = 0$.
- $L_{k,5}$ tel que $L_{k,5}(x_k) = 0$, $L_{k,5}(x_{k,5}) = 1$ et $L_{k,5}(x_{k+1}) = 0$.
- L_{k+1} tel que $L_{k+1}(x_k) = 0$, $L_{k+1}(x_{k,5}) = 0$ et $L_{k+1}(x_{k+1}) = 1$.

On les construit de cette manière :

$$\begin{aligned}
L_k &= \frac{(X - x_{k+1})(X - x_{k,5})}{(x_k - x_{k+1})(x_k - x_{k,5})}.\\
L_{k,5} &= \frac{(X - x_{k+1})(X - x_k)}{(x_{k,5} - x_{k+1})(x_{k,5} - x_k)}.\\
L_{k+1} &= \frac{(X - x_k)(X - x_{k,5})}{(x_{k+1} - x_k)(x_{k+1} - x_{k,5})}.
\end{aligned}$$

Nous pouvons vérifier, ces polynômes font bien ce qu'on leur demande. Par exemple :

$$\begin{aligned}
L_k(x_k) &= \frac{(x_k - x_{k+1})(x_k - x_{k,5})}{(x_k - x_{k+1})(x_k - x_{k,5})} = 1.\\
L_k(x_{k+1}) &= \frac{(x_{k+1} - x_{k+1})(x_{k+1} - x_{k,5})}{(x_k - x_{k+1})(x_k - x_{k,5})} = 0.\\
L_k(x_{k,5}) &= \frac{(x_{k,5} - x_{k+1})(x_{k,5} - x_{k,5})}{(x_k - x_{k+1})(x_k - x_{k,5})} = 0.
\end{aligned}$$

Maintenant, il est très simple de construire Q_k :

$$Q_k = f(x_k)L_k + f(x_{k,5})L_{k,5} + f(x_{k+1})L_{k+1}.$$

On obtient bien un polynôme qui satisfait aux contraintes que nous lui avions imposées.

3.2. Calcul de l'intégrale

Il nous reste maintenant le plus gros du travail à faire. Il nous faut calculer l'intégrale de Q_k. Et pour cela, il nous faut calculer les intégrales de L_k, de $L_{k,5}$ et de L_{k+1}.

> Pour faciliter l'écriture, nous allons poser pour le calcul de ces intégrales $i{,}5 = m$.

On a:

$$\int_{x_k}^{x_{k+1}} L_k(t)\mathrm{d}t = \int_{x_k}^{x_{k+1}} \frac{(t-x_{k+1})(t-x_m)}{(x_k-x_{k+1})(x_k-x_m)}\mathrm{d}t.$$

Nous pouvons sortir la constante:

$$\int_{x_k}^{x_{k+1}} L_k(t)\mathrm{d}t = \frac{1}{(x_k-x_{k+1})(x_k-x_m)}\int_{x_k}^{x_{k+1}} (t-x_{k+1})(t-x_m)\mathrm{d}t.$$

En faisant une intégration par partie:

$$\int_{x_k}^{x_{k+1}} L_k(t)\mathrm{d}t = \frac{1}{(x_k-x_{k+1})(x_k-x_m)}\left(\left[(t-x_m)\frac{(t-x_{k+1})^2}{2}\right]_{x_k}^{x_{k+1}} - \int_{x_k}^{x_{k+1}} \frac{(t-x_{k+1})^2}{2}\mathrm{d}t\right).$$

Et là nous savons faire l'intégration et donc:

$$\int_{x_k}^{x_{k+1}} L_k(t)\mathrm{d}t = \frac{-(x_k-x_{k+1})^2(x_k-x_m)}{2(x_k-x_{k+1})(x_k-x_m)} - \frac{1}{2(x_k-x_{k+1})(x_k-x_m)} \times \left[\frac{(t-x_{k+1})^3}{3}\right]_{x_k}^{x_{k+1}}.$$

Nous poursuivons le calcul pour obtenir:

$$\int_{x_k}^{x_{k+1}} L_k(t)\mathrm{d}t = \frac{(x_{k+1}-x_k)}{2} + \frac{(x_k-x_{k+1})^2}{6(x_k-x_m)}.$$

Et finalement, sachant que $x_{k+1}-x_k=\delta$ et que $x_k-x_m = -\frac{x_{k+1}-x_k}{2} = -\frac{\delta}{2}$, on a:

$$\int_{x_k}^{x_{k+1}} L_k(t)\mathrm{d}t = \left(\frac{1}{2}-\frac{1}{3}\right)\times\delta = \frac{\delta}{6}.$$

Le même calcul mène à

$$\int_{x_k}^{x_{k+1}} L_{k+1}(t)\mathrm{d}t = \frac{\delta}{6}.$$

De la même manière, nous calculons l'intégrale de L_m:

$$\int_{x_k}^{x_{k+1}} L_m(t)\mathrm{d}t = \int_{x_k}^{x_{k+1}} \frac{(t-x_{k+1})(t-x_k)}{(x_m-x_{k+1})(x_m-x_k)}\mathrm{d}t.$$

Toujours en sortant la constante:

$$\int_{x_k}^{x_{k+1}} L_m(t)\mathrm{d}t = \frac{1}{(x_m-x_{k+1})(x_m-x_k)}\int_{x_k}^{x_{k+1}} (t-x_{k+1})(t-x_k)\mathrm{d}t.$$

Et là, on fait la même intégration par partie que précédemment:

$$\int_{x_k}^{x_{k+1}} L_m(t)\mathrm{d}t = \frac{1}{(x_m-x_{k+1})(x_m-x_k)}\left(\left[(t-x_k)\frac{(t-x_{k+1})^2}{2}\right]_{x_k}^{x_{k+1}} - \int_{x_k}^{x_{k+1}} \frac{(t-x_{k+1})^2}{2}\mathrm{d}t\right).$$

D'où:

$$\int_{x_k}^{x_{k+1}} L_m(t)\mathrm{d}t = \frac{1}{2(x_m-x_{k+1})(x_m-x_k)} \times \left[\frac{(t-x_{k+1})^3}{3}\right]_{x_k}^{x_{k+1}}.$$

Nous poursuivons le calcul pour obtenir:

$$\int_{x_k}^{x_{k+1}} L_m(t)\mathrm{d}t = \frac{1}{2(x_m-x_{k+1})(x_m-x_k)} \times \frac{-(x_k-x_{k+1})^3}{3}.$$

Et finalement, sachant que $x_{k+1}-x_k=\delta$, que $x_m-x_k=\frac{x_{k+1}-x_k}{2}=\frac{\delta}{2}$ et que $x_m-x_{k+1}=-\frac{x_{k+1}-x_k}{2}=-\frac{\delta}{2}$, on a:

$$\int_{x_k}^{x_{k+1}} L_m(t)\mathrm{d}t = -\frac{4}{2\delta^2}\times\frac{-\delta^3}{3} = \frac{2\delta}{3}.$$

Grâce à ces deux calculs, nous pouvons calculer l'intégrale de Q_k:

$$\begin{aligned}
\int_{x_k}^{x_{k+1}} Q_k(t)\mathrm{d}t &= \int_{x_k}^{x_{k+1}} f(x_k)L_k(t)\mathrm{d}t + f(x_m)L_m(t)\mathrm{d}t + f(x_{k+1})L_{k+1}(t)\mathrm{d}t \\
&= f(x_k)\int_{x_k}^{x_{k+1}} L_k(t)\mathrm{d}t + f(x_m)\int_{x_k}^{x_{k+1}} L_m(t)\mathrm{d}t + f(x_{k+1})\int_{x_k}^{x_{k+1}} L_{k+1}(t)\mathrm{d}t \\
&= \frac{\delta f(x_k)}{6} + \frac{2\delta f(x_m)}{3} + \frac{\delta f(x_{k+1})}{6} \\
&= \frac{\delta}{3}\left(\frac{f(x_k)+f(x_{k+1})}{2} + 2f(x_m)\right).
\end{aligned}$$

Ouf, c'était un travail de longue haleine, mais le plus dur est maintenant fait.

3.2.1. Calcul de l'intégrale complète

Il ne nous reste plus qu'à calculer la somme des intégrales de Q_k pour enfin avoir l'approximation de notre intégrale (on revient à notre «i,5»):

$$I(f) \approx \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} Q_k(t)\mathrm{d}t = \sum_{k=0}^{n-1} \frac{\delta}{3} \left(\frac{f(x_k) + f(x_{k+1})}{2} + 2f(x_{k,5}) \right).$$

Comme d'habitude on sort la constante de la somme:

$$I(f) \approx \frac{\delta}{3} \sum_{k=0}^{n-1} \left(\frac{f(x_k) + f(x_{k+1})}{2} + 2f(x_{k,5}) \right).$$

On remarque que tous les x_k sont comptés deux fois sauf le premier (x_0) et le dernier (x_n). Donc on peut les sortir de la somme (mais ça nous oblige à sortir le $f(x_{0,5})$ de la somme):

$$\boxed{I(f) \approx \frac{\delta}{3} \left(\frac{f(x_0) + f(x_n)}{2} + 2f(x_{0,5}) + \sum_{k=1}^{n-1} (f(x_k) + 2f(x_{k,5})) \right)}.$$

Et c'est cette écriture que nous allons garder.

3.3. Algorithme et tests

Écrivons cet algorithme en Python (pour obtenir les ,5, il nous suffit d'ajouter $\frac{\delta}{2}$) :

```python
def simpson(f, a, b, n):
    pas = (b - a) / n
    somme = (f(a) + f(b)) / 2 + 2 * f(a + pas / 2)  # On initialise la somme
    x = a + pas              # La somme commence à x_1
    for i in range(1, n): # On calcule la somme
        somme += f(x) + 2 * f(x + pas / 2)
        x += pas
    return somme * pas / 3   # On retourne cette somme fois le pas / 3
```

Comme pour la méthode des rectangles et celle des trapèzes, essayons-la avec la fonction cosinus sur $[0, \pi]$, en découpant l'intervalle en 100. On obtient alors ce graphe.

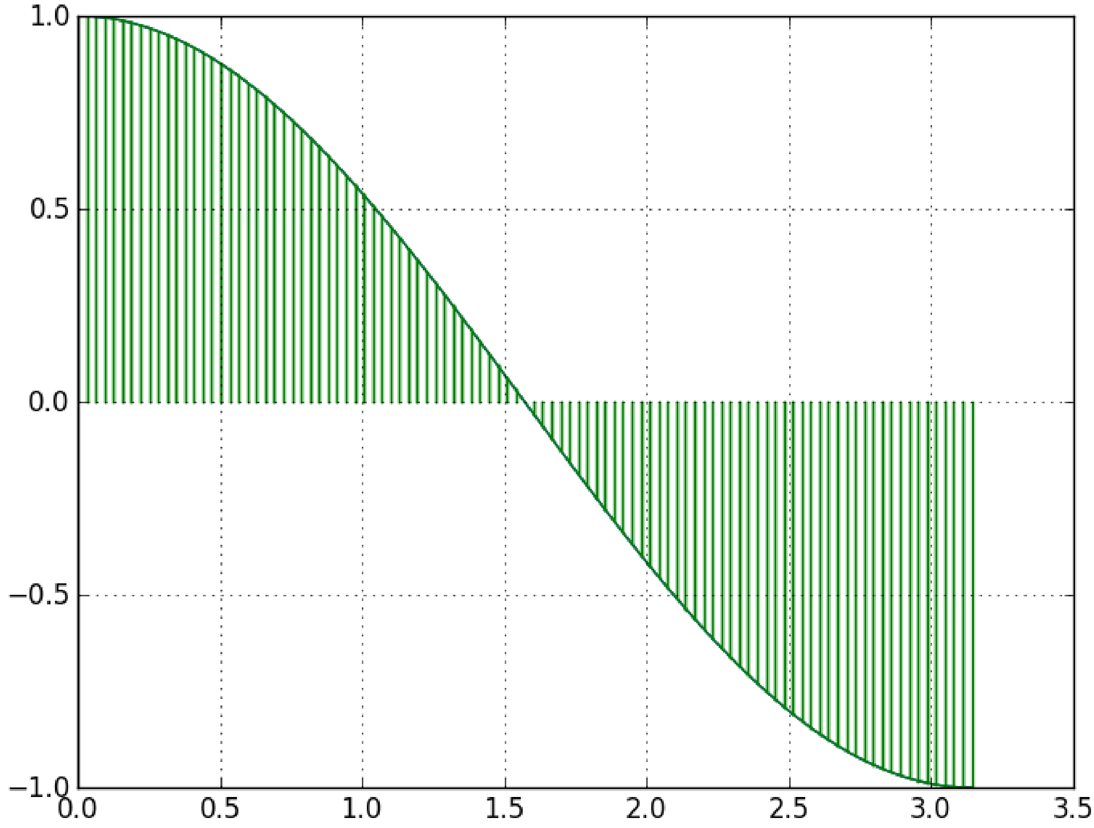

FIGURE 3.2. – La méthode de Simpson sur la fonction cosinus.

Le graphe nous permet déjà de savoir que le résultat obtenu sera proche de la vraie valeur de l'intégrale, et en effet, nous obtenons un résultat de l'ordre de $\mathbf{10^{-15}}$. Ce résultat est bien proche de 0.

En fait, la méthode de Simpson est moins précise que la méthode des trapèzes dans le cas particulier de la fonction cosinus. Dans le bloc secret qui suit, nous verrons pourquoi. Il ne s'agit que d'un cas particulier, mais généralement, sur une fonction suffisamment régulière (voir le chapitre suivant), la méthode de Simpson est plus précise.

Contenu masqué n°1

Il est maintenant temps de voir ce que cette méthode nous donne sur notre exemple.

```
print(simpson(f, 0, 1, 100))
```

Avec elle on obtient environ 0,746824.

Contenu masqué

Contenu masqué n°1

En fait, la méthode des trapèzes est censé donner un résultat nul pour la fonction cosinus sur $[0, \pi]$. En effet, appliquer la méthode des trapèzes avec un pas n dans ce cas consiste à calculer

$$\frac{\pi}{n}\left(\frac{\cos(0)+\cos(\pi)}{2}+\sum_{k=1}^{n-1}\cos\left(\frac{k\pi}{n}\right)\right).$$

On a $\cos(0)+\cos(\pi)=0$, il nous reste à calculer la somme (appelons la S_n). On a en faisant le changement d'indice $l = n - k$,

$$\begin{aligned} S_n &= \sum_{l=1}^{n-1}\cos\left(\frac{(l-n)\pi}{n}\right)\\ &= \sum_{l=1}^{n-1}\cos\left(\frac{l\pi}{n}-\pi\right)\\ &= \sum_{l=1}^{n-1}\cos\left(-\frac{l\pi}{n}\right)\\ &= \sum_{l=1}^{n-1}-\cos\left(\frac{l\pi}{n}\right)\\ &= -S_n. \end{aligned}$$

On en déduit que S_n est nul, et donc on obtient bien que la méthode des trapèzes donne un résultat nul dans ce cas particulier (l'ordinateur ne nous donne pas 0 à cause des erreurs d'approximations).

Un autre fait que nous pouvions remarquer est que (toujours pour la fonction cosinus entre 0 et π), la méthode des rectangles nous donne $\frac{\pi}{n}$, car elle consiste à calculer

$$\begin{aligned} \frac{\pi}{n}\left(\sum_{k=0}^{n-1}\cos\left(\frac{k\pi}{n}\right)\right) &= \frac{\pi}{n}\left(\cos(0)+S_n\right)\\ &= \frac{\pi}{n}. \end{aligned}$$

Retourner au texte.

4. Calcul de l'erreur

Dans ce chapitre, nous allons faire des calculs pour déterminer **l'erreur commise** (en fait, nous allons seulement la majorer). Cela signifie que nous allons majorer la différence entre la vraie valeur et la valeur que la méthode nous donnerait avec un pas n. Notons que nous sommes censés trouver que cette erreur tend vers 0 quand le pas tend vers l'infini.

Pour majorer l'erreur ε_n, nous allons majorer toutes les petites erreurs sur chaque petit intervalle (notons $\varepsilon_{n,k}$ l'erreur sur l'intervalle $[x_k, x_{k+1}]$). Par exemple, pour la méthode des rectangles, $\varepsilon_{n,k}$ correspond à la différence entre l'intégrale de f sur $[x_k, x_{k+1}]$ et l'aire du petit rectangle. En additionnant toutes les petites erreurs, on obtient l'erreur sur l'intervalle $[a, b]$ où l'on intègre (donc sur tout l'intervalle).

$$\varepsilon_n = \sum_{k=0}^{n-1} \varepsilon_{n,k}.$$

4.1. La méthode des rectangles

Avant de faire le moindre calcul, nous allons établir une propriété qui nous sera très utile pour la suite.

4.1.0.1. Propriété

Soit n un entier naturel et f une fonction de classe C^n sur un intervalle $[a, b]$. Alors, il existe une constante M telle que pour tout (x, y) appartenant à $[a, b]$,

$$|f^{(n-1)}(y) - f^{(n-1)}(x)| \leq M|y - x|.$$

En effet, on a que $f^{(n)}$ est continue et donc bornée sur $[a, b]$ d'après le **théorème des bornes**. Donc, il existe une constante M telle que pour tout x appartenant à $[a, b]$, $|f^{(n)}(x)| \leq M$. En intégrant $f^{(n)}$ sur $[a, b]$ (possible car $f^{(n)}$ est continue sur cet intervalle), on obtient alors

$$f^{(n-1)}(y) - f^{(n-1)}(x) = \int_x^y f^{(n)}(t)\mathrm{d}t \leq \left| \int_x^y |f(t)| \,\mathrm{d}t \right|.$$

D'où, puisque $f^{(n)}$ est bornée,

$$|f^{(n-1)}(y) - f^{(n-1)}(x)| \leq \left| \int_x^y M\mathrm{d}t \right| = M|y - x|.$$

Passons maintenant au calcul de l'erreur dans le cas de la méthode des rectangles. On va supposer f de classe C^1 (cela nous permettre d'appliquer notre propriété). Avec la méthode des rectangles, l'aire du rectangle k est $R_k(f) = (x_k - x_{k+1})f(x_k)$ d'où l'erreur sur l'intervalle $[x_k, x_{k+1}]$ est

$$\varepsilon_{n,k} = \left(\int_{x_k}^{x_{k+1}} f(t)\mathrm{d}t\right) - (x_k - x_{k+1})f(x_k).$$

Pour x appartenant à $[a, b]$, on pose

$$E(x) = \left(\int_{x_k}^{x} f(t)\mathrm{d}t\right) - (x_k - x)f(x_k).$$

Pour majorer $\varepsilon_{n,k}$, il nous suffit de majorer E (car $\varepsilon_{n,k} = E(x_{k+1})$).

En dérivant E, on obtient $E'(x) = f(x) - f(x_k)$ d'où, en appliquant notre propriété, il existe une constante M telle que $E'(x) \leq M|x - x_k|$. En intégrant sur $[x_k, x]$ (avec $E(x_k)$ nul), on obtient

$$E(x) \leq \frac{M|x - x_k|^2}{2}.$$

Ceci nous permet d'obtenir en évaluant E en x_{k+1},

$$\varepsilon_{n,k} \leq \frac{M(x_{k+1} - x_k)^2}{2}.$$

Or on sait que $x_{k+1} - x_k = \delta = \frac{b-a}{n}$. Donc :

$$\varepsilon_{n,k} \leq \frac{M(b-a)^2}{2n^2}.$$

Grâce à cela, on peut majorer l'erreur ε_n :

$$|\varepsilon_n| \leqslant \sum_{k=0}^{n-1} |\varepsilon_{n,k}| \leqslant \sum_{k=0}^{n-1} \frac{M(b-a)^2}{2n^2}.$$

Et finalement :

$$|\varepsilon_n| \leqslant \frac{M(b-a)^2}{2n}.$$

La méthode des rectangles converge bien vers $I(f)$ puisque $|\varepsilon_n| \underset{n\to+\infty}{\longrightarrow} 0$ et on sait qu'elle a une convergence en $O\left(\frac{1}{n}\right)$.

> La convergence est trouvée en négligeant certains termes tels que les constantes, cela nous donne juste une idée de la vitesse de convergence de la méthode.

4.2. La méthode des trapèzes

Pour majorer l'erreur de la méthode des trapèzes, nous allons faire comme pour la méthode des rectangles. Supposons f de classe C^2 (donc f'' est bornée sur $[a, b]$ par une constante M. L'aire du trapèze k est

$$T_k(f) = (x_{k+1} - x_k)\frac{f(x_{k+1}) - f(x_k)}{2}$$

de sorte que

$$\varepsilon_{n,k} = \int_{x_k}^{x_{k+1}} f(x)\mathrm{d}t - (x_{k+1} - x_k)\frac{f(x_{k+1}) - f(x_k)}{2}.$$

Cette fois, on considère

$$E(x) = \int_{x_k}^{x} f(t)\mathrm{d}t - (x - x_k)\frac{f(x) - f(x_k)}{2}.$$

Là, encore, majorer E sur $[a, b]$ nous permet de majorer $\varepsilon_{n,k}$. Essayons comme tout-à-l'heure de dériver E.

$$E\text{'}(x) = \frac{1}{2}\left(f(x) - f(x_k) - (x - x_k)f\text{'}(x)\right).$$

Cela ne suffit pas, dérivons une nouvelle fois.

$$E\text{''}(x) = \frac{1}{2}\left(f\text{'}(x) - f\text{'}(x) - (x - x_k)f\text{''}(x)\right) = -\frac{(x - x_k)f\text{''}(x)}{2}.$$

On a alors puisque f'' est bornée

$$|E\text{''}(x)| \leq \frac{M(x - x_k)}{2}.$$

Il ne nous reste plus qu'à intégrer deux fois sur $[x_k, x]$ (sachant que $E\text{'}(x_k)$ et $E(x_k)$ sont nuls) pour obtenir une majoration. On obtient $|E\text{'}(x)| \leq \frac{M(x-x_k)^2}{4}$ et

$$|E(x)| \leq \frac{M(x - x_k)^3}{12}.$$

Et donc en évaluant en x_{k+1}, on obtient

$$|\varepsilon_{n,k}| \leq \frac{M(x_{k+1}-x_k)^3}{12} = \frac{M(b-a)^3}{12n^3}.$$

Ce résultat nous permet de majorer l'erreur ε_n :

$$|\varepsilon_n| \leqslant \sum_{k=0}^{n-1} |\varepsilon_{n,k}| \leqslant \sum_{k=0}^{n-1} \frac{M(b-a)^3}{12n^3}.$$

Et donc :

$$|\varepsilon_n| \leqslant \frac{M(b-a)^3}{12n^2}.$$

$|\varepsilon_n| \underset{n\to+\infty}{\longrightarrow} 0$ donc la méthode des trapèzes converge bien vers $I(f)$ et elle a une convergence en $O\left(\frac{1}{n^2}\right)$. Elle a donc une meilleure vitesse de convergence que la méthode des rectangles.

> Cela ne veut pas dire qu'à pas égal la méthode des trapèzes donne un résultat plus précis que celle des rectangles. On a juste qu'elle converge plus vite lorsqu'on augmente le pas.

4.3. La méthode de Simpson

Pour la méthode de Simpson, le calcul de l'erreur est un peu plus long à faire. Cette fois, nous considérons f de classe C^4. L'intégrale du $k^{\text{ème}}$ polynôme considéré est

$$\int_{x_k}^{x_{k+1}} L_k(t)\mathrm{d}t = \frac{x_{k+1}-x_k}{3}\left(\frac{f(x_k)+f(x_{k+1})}{2} + 2f(c)\right)$$

et donc l'erreur à considérer est

$$\varepsilon_{n,k} = \int_{x_k}^{x_{k+1}} f(t)\mathrm{d}t - \frac{x_{k+1}-x_k}{3}\left(\frac{f(x_k)+f(x_{k+1})}{2} + 2f(x_{k,5})\right).$$

En considérant $c = \frac{x_k+x_{k+1}}{2} = \frac{\delta}{2}$ et le point milieu $x_{k,5}$, nous pouvons transformer l'expression précédente de la manière suivante.

$$\varepsilon_{n,k} = \int_{x_{k,5}-c}^{x_{k,5}+c} f(t)\mathrm{d}t - \frac{c}{3}\Big(f(x_{k,5}-c) + f(x_{k,5}+c) + 4f(x_{k,5})\Big).$$

On va maintenant considérer (toujours sur $[a,b]$)

$$E(x) = \int_{x_{k,5}-x}^{x_{k,5}+x} f(t)\mathrm{d}t - \frac{x}{3}\Big(f(x_{k,5}-x) + f(x_{k,5}+x) + 4f(x_{k,5})\Big).$$

Et on dérive (cette fois, il sera nécessaire de dériver trois fois).

$$\begin{aligned}
E'(x) &= \frac{2}{3}\Big(f(x_{k,5}+x)+f(x_{k,5}-x)\Big) - \frac{x}{3}\Big(f'(x_{k,5}+x)-f'(x_{k,5}-x)\Big) - \frac{4}{3}f(x_{k,5}).\\
E''(x) &= \frac{1}{3}\Big(f'(x_{k,5}+x)-f'(x_{k,5}-x)\Big) - \frac{x}{3}\Big(f''(x_{k,5}+x)+f''(x_{k,5}-x)\Big).\\
E^{(3)}(x) &= \frac{x}{3}\left(f^{(3)}(x_{k,5}-x)-f^{(3)}(x_{k,5}+x)\right).
\end{aligned}$$

On peut maintenant utiliser notre propriété (car on a supposé f de classe C^4). Donc, il existe une constante M telle que $|f^{(3)}(x_{k,5}-x)-f^{(3)}(x_{k,5}+x)| \leq M|(x_{k,5}-x)-(x_{k,5}+x)| = 2M|x|$. On a alors

$$\left|E^{(3)}(x)\right| \leq \frac{2M|x|^2}{3}.$$

Il nous reste à intégrer trois fois pour obtenir la majoration de E. En faisant ces trois intégrations sur $[0,x]$ (on a $E(0)$, $E'(0)$ et $E''(0)$ nuls), on obtient successivement $|E''(x)| \leq \frac{2Mx^3}{9}$, $|E'(x)| \leq \frac{Mx^4}{18}$ et finalement

$$|E(x)| \leq \frac{M|x|^5}{90}.$$

En évaluant E en c, on obtient

$$\varepsilon_{n,k} \leq \frac{M(x_{k+1}-x_k)^5}{2^5 * 80} = \frac{M(b-a)^5}{2880n^5}.$$

On peut maintenant majorer l'erreur,

$$|\varepsilon_n| \leq \sum_{k=0}^{n-1}|\varepsilon_{n,k}| \leq \sum_{k=0}^{n-1}\frac{M(b-a)^5}{2880n^5}.$$

On obtient finalement

$$|\varepsilon_n| \leq \frac{M(b-a)^5}{2880n^4}.$$

On a $|\varepsilon_n| \underset{n\to+\infty}{\longrightarrow} 0$, donc la méthode de Simpson converge bien. De plus, elle a une convergence en $O\left(\frac{1}{n^4}\right)$ (donc elle converge plus vite que les autres lorsqu'on diminue le pas).

Avec les résultats que nous avons obtenu, il semble logique de penser que la méthode de Simpson est plus efficace que les autres. Il nous faut cependant garder un œil critique sur ces résultats et sur les hypothèses que nous avons formulé pour les trouver. On peut alors pointer ces différents points.

— Nous avons supposé que la fonction était d'une certaine régularité.

- La méthode de Simpson demande la connaissance de deux fois plus de points que la méthode des rectangles.
- La méthode de Simpson demande légèrement plus de calculs.

Au vu de ces différents points, nous pouvons juste conclure que la méthode de Simpson converge plus vite lorsqu'on diminue le pas.

Nous avons testé les trois méthodes pour calculer

$$\int_0^1 \mathrm{e}^{-t^2} \mathrm{d}t$$

avec un pas de $\frac{1}{100}$ à chaque fois. Nous avons obtenu dans l'ordre environ 0,749979, environ 0,746818 et environ 0,746824. Un logiciel de calcul comme Wolfram Alpha nous donne comme résultat 0,746824. On voit donc que la méthode de Simpson donne un résultat assez satisfaisant sur cet exemple.

C'est maintenant la fin de ce tutoriel. Nous avons vu trois méthodes pour calculer les intégrales. La méthode de Simpson est celle des trois qui converge le plus vite. Mais il existe beaucoup d'autres méthodes pour calculer les intégrales. Voyez la page Wikipédia à ce sujet, elle n'est pas complète et pourtant elle est déjà très longue. Le chemin à parcourir est encore long.

Dans tout le tutoriel nous avons calculé l'intégrale entre 0 et 1 de $x \mapsto \mathrm{e}^{-x^2}$. L'intégrale de cette fonction est très utilisée, notamment son intégrale entre $+\infty$ et $-\infty$. Pour des informations à propos de ce type d'intégrale, à savoir les **intégrales de Gauss**, nous pouvons regarder la page Wikipédia à ce sujet, et utiliser les différents programmes que nous avons codés pour obtenir quelques résultats.

Nous devons noter qu'un ordinateur est une machine qui fait des calculs avec une précision limitée. Ceci mène également a des erreurs d'approximation comme nous pouvons le voir dans certains contenus.

Je remercie @Holosmos et @Vayel pour leurs conseils d'écriture et leurs corrections. Merci également à @adri1 pour ses conseils et son courage pour valider ce tutoriel.