25 Intégration

Le but de ce chapitre est de définir **rigoureusement** ce qu'est une intégrale.
Bien entendu, vous avez l'impression[1] de déjà savoir ce qu'est une intégrale.
Mais la définition d'intégrale dont nous disposons pour l'instant (à l'aide d'une primitive) a plusieurs lacunes :

- on a admis jusque là l'existence de primitive(s) d'une fonction continue
- elle n'explique en rien le lien entre l'aire et l'intégrale qu'on vous a appris en terminale. Et d'ailleurs, qu'est-ce qu'une aire ?
- Elle nécessite des fonctions continues, et par exemple ne nous autorise pas à calculer $\displaystyle\int_0^2 \lfloor x \rfloor\, dx$, alors que graphiquement, on se fait une bonne idée de ce que doit valoir cette intégrale.

Nous allons donc reconstruire l'intégrale en partant de zéro, et retrouver toutes les propriétés que nous connaissons déjà.
Pour la culture, mentionnons qu'il existe plusieurs théories de l'intégration, et que l'intégrale que nous construisons ici se nomme l'intégrale de Riemann.
C'est probablement la plus facile à construire, mais elle souffre de plusieurs lacunes :

1. elle est délicate à étendre à des intégrales de fonctions définies sur autre chose qu'un segment (par exemple $]0, 1]$ ou $[0, +\infty[$), sujet auquel vous consacrerez un peu de temps en seconde année
2. certains grands théorèmes (au programme de seconde année) sont pénibles à prouver dans le contexte de l'intégrale de Riemann.
3. elle ne nous permet pas de définir par exemple $\displaystyle\int_0^1 \mathbb{1}_{\mathbf{Q}}(t)\, dt$.

Une autre théorie de l'intégration très célèbre se nomme l'intégrale de Lebesgue. Elle est bien plus difficile à construire[2], mais elle possède de nombreux avantages, notamment de répondre aux trois problèmes soulevés ci-dessus (les preuves des grands théorèmes d'intégration sont presque triviales avec le vocabulaire de l'intégrale de Lebesgue), et de fournir un cadre très général dans lequel faire des probabilités.
Un certain nombre d'entre vous auront l'occasion d'en reparler en école, et pas uniquement ceux qui suivront des cursus de maths pures.

Dans toute la suite du chapitre, en l'absence de précisions, la lettre $\mathbf{K}$ désigne indifféremment $\mathbf{R}$ ou $\mathbf{C}$.

[1] Et à juste titre.

[2] Et c'est la raison pour laquelle elle n'est pas enseignée en prépa.

25.1 Uniforme continuité

Définition 25.1 – Soit $f : D \to \mathbf{K}$ une fonction définie sur une partie $D \subset \mathbf{R}$. On dit que f est **uniformément continue sur** D si :

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall (x, y) \in D^2,\ |x - y| < \eta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Notons tout de suite la différence avec la continuité : une fonction f est continue sur D si elle est continue en tout point de D. Soit si et seulement si

$$\forall x \in D,\ \forall \varepsilon > 0,\ \exists \eta > 0,\ \forall y \in D,\ |x - y| < \eta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

La différence réside donc uniquement dans l'ordre des quantificateurs : pour une fonction continue, le η dépend du point x choisi[3], alors que pour une fonction uniformément continue, le même η convient pour tout x.

[3] Et bien entendu du ε.

Proposition 25.2 : *Une fonction uniformément continue est continue.*

Démonstration. C'est une simple permutation de quantificateurs : si il existe un η qui convient pour tout x, alors pour tout x, il existe un η tel que...
Dans le détail : fixons $\varepsilon > 0$. Alors $\exists \eta > 0$ tel que $\forall (x, y) \in D^2, |x-y| < \eta \Rightarrow |f(x)-f(y)| < \varepsilon$.
Soit donc $x \in D$ fixé. Alors, pour tout $y \in D$, si $|x - y| < \eta$, alors $|f(x) - f(y)| < \varepsilon$.
Donc f est continue en x, et ceci étant vrai pour tout $x \in D$, f est continue sur D. □

La réciproque est fausse : une fonction peut être continue sans être uniformément continue, c'est-à-dire que le η dépend réellement de x.
Par exemple considérons la fonction carré $f : x \mapsto x^2$ sur $\mathbf{R}^+$.
Prenons $\varepsilon = 1$. Pour tout $x \in \mathbf{R}^+$ et $h > 0$, on a

$$|f(x + h) - f(x)| = (x + h)^2 - x^2 = 2xh + h^2.$$

Et donc en particulier, pour $\eta > 0$ fixé, prenons $x = \frac{1}{\eta}$ et $h = \frac{\eta}{2}$.
Alors les deux réels x et $y = x+h$ vérifient $|x-y| = \frac{\eta}{2} < \eta$ mais $|f(x)-f(y)| = 2\frac{1}{\eta}\frac{\eta}{2} + \frac{\eta^2}{4} > 1$.
Ainsi, on a prouvé que

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists (x, y) \in \mathbf{R}_+^2, |x - y| < \eta \textbf{ et } |f(x) - f(y)| \geqslant \varepsilon.$$

Ce qui est bien la négation de f uniformément continue.

Proposition 25.3 : *Une fonction lipschitzienne est uniformément continue.*

Démonstration. Soit $f : D \to \mathbf{K}$ une fonction k-lipschitzienne, avec $k > 0$.
Soit $\varepsilon > 0$, et soit $\eta = \dfrac{\varepsilon}{k}$.
Alors pour tout $(x, y) \in D^2$, si $|x - y| < \eta$, alors

$$|f(x) - f(y)| \leqslant k|x - y| \leqslant k\frac{\varepsilon}{k} \leqslant \varepsilon.$$

Et donc f est uniformément continue. □

On a donc les implications f lipschitzienne $\Rightarrow$ f uniformément continue $\Rightarrow$ f continue.
Mais aucune des deux implications réciproques n'est vraie[4].

Le résultat qui suit est vraiment fondamental dans la suite :

Théorème 25.4 (de Heine) : *Soit f une fonction continue sur un segment $[a, b]$. Alors f est uniformément continue.*

Démonstration. Soit $f : [a, b] \to \mathbf{K}$. Supposons par l'absurde que f n'est pas uniformément continue sur $[a, b]$, de sorte que

$$\exists \varepsilon > 0, \forall \eta > 0,\ \exists (x, y) \in [a, b]^2,\ |x - y| < \eta \textbf{ et } |f(x) - f(y)| \geqslant \varepsilon.$$

En particulier, pour tout $n \in \mathbf{N}^*$, en prenant $\eta = \frac{1}{n}$, il existe x_n, y_n tels que $|x_n - y_n| < \frac{1}{n}$ et $|f(x_n) - f(y_n)| \geqslant \varepsilon$.
Puisque (x_n) est bornée[5], par le théorème de Bolzano-Weierstrass, il existe une extractrice φ telle que $(x_{\varphi(n)})_{n\in\mathbf{N}}$ converge vers un réel x.
Notons que x est encore dans $[a, b]$ car $a \leqslant x_{\varphi(n)} \leqslant b$, et que les inégalités sont préservées par passage à la limite.
Puisque $\dfrac{1}{\varphi(n)} \underset{n\to+\infty}{\longrightarrow} 0$, et que $|x_{\varphi(n)} - y_{\varphi(n)}| \leqslant \dfrac{1}{\varphi(n)}$, on a également $y_{\varphi(n)} \underset{n\to+\infty}{\longrightarrow} x$.
Par caractérisation séquentielle de la continuité[6], $f(x_{\varphi(n)}) \underset{n\to+\infty}{\longrightarrow} f(x)$, et de même pour $f(y_{\varphi(n)})$.
Et donc en passant à la limite $|f(x) - f(x)| \geqslant \varepsilon$, ce qui est absurde. □

[4] Nous avons déjà prouvé qu'il existe des fonctions continues non uniformément continues, voir un peu plus loin pour une fonction uniformément continue et non lipschitzienne.

[5] Elle est à valeurs dans $[a, b]$.

[6] Et c'est là qu'on utilise la continuité de f.

Rappelons que nous avions prouvé que la fonction racine carrée, n'est pas lipschitzienne sur $[0,1]$. En revanche, étant continue sur le segment $[0,1]$, elle est[7] uniformément continue. Et nous avons donc là un exemple de fonction uniformément continue non lipschitzienne.

[7] Par le théorème de Heine.

25.2 Intégrale des fonctions en escalier

Définition 25.5 – Si $[a,b]$ est un segment, avec $a \leqslant b$, on appelle **subdivision de** $[a,b]$ toute partie finie $\sigma = \{x_i, 0 \leqslant i \leqslant n\}$ de $[a,b]$ avec $a = x_0 < x_1 < \cdots < x_n = b$.
Le réel $\max\limits_{0\leqslant i\leqslant n-1} x_{i+1} - x_i$ est appelé le pas de σ.

Autrement dit, se donner une subdivision de $[a,b]$, c'est juste se donner un «découpage» de $[a,b]$ en un nombre fini de segments plus petits.

L'exemple le plus classique de subdivision est, pour $n \in \mathbf{N}^*$, la subdivision à n éléments obtenue en posant $\forall k \in [\![0,n]\!]$, $x_k = a + k\dfrac{b-a}{n}$.
On parle de subdivision à **pas régulier**.

25.2.1 Fonctions en escalier

Définition 25.6 – Soit $f : [a,b] \to \mathbf{R}$. On dit que f est une **fonction en escalier** s'il existe une subdivision $(x_i)_{0\leqslant i\leqslant n-1}$ de $[a,b]$ telle que pour tout $i \in [\![0, n-1]\!]$, $f_{|]x_i,x_{i+1}[}$ soit constante.
Une telle subdivision σ est dite **adaptée** à la fonction f.
On note $\mathscr{E}([a,b])$ l'ensemble des fonctions en escalier sur $[a,b]$.

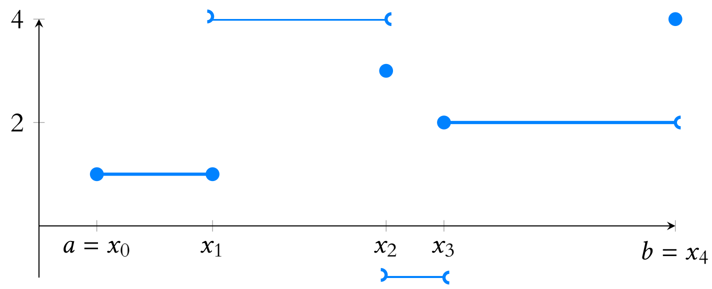


Figure 25.1 – Un exemple de fonction en escalier.

Remarques. ▶ Notons qu'on n'impose rien sur la valeur des $f(x_i)$. La fonction f peut être continue à droite en x_i, elle peut être continue à gauche, ou elle peut n'être ni l'un ni l'autre. Voir par exemple la fonction ci-dessus.
▶ Il n'y a absolument pas unicité d'une subdivision adaptée à f : si $(x_i)_{0\leqslant i\leqslant n}$ est adaptée à f, alors toute subdivision plus fine[8] l'est encore.
En effet, la restriction d'une fonction constante à une partie est encore une fonction constante.

[8] C'est-à-dire contenant tous les x_i.

Proposition 25.7 : *L'ensemble $\mathscr{E}([a,b])$ des fonctions en escalier est à la fois un sous-espace vectoriel et un sous-anneau de $\mathscr{F}([a,b],\mathbf{R})$.*

Démonstration. La preuve n'est pas si dure, mais plutôt désagréable à écrire.
La seule chose à laquelle il faut être soigneux lorsqu'on manipule deux fonctions en escalier f et g est qu'elles ont a priori chacune une subdivision qui leur est adaptée, mais il n'y a aucune raison qu'il s'agisse de la même.
En revanche, l'union de ces deux subdivisions, réarrangée par ordre croissant, et après élimination des éventuels doublons[9] est une subdivision de $[a,b]$ plus fine que les deux

[9] Par exemple a et b sont déjà dans les deux subdivisions de départ.

subdivisions adaptées à f et à g, et donc adaptée aux deux à la fois.

Une fois établie l'existence d'une telle subdivision, il n'y a plus grand chose à faire. Prouvons par exemple que $\lambda f + g$ est encore une fonction en escalier.
Soit donc $a = x_0 < x_1 < \cdots < x_n = b$ une subdivision de $[a, b]$ adaptée à la fois à f et à g.
Pour tout $i \in [\![0, n-1]\!]$, notons alors μ_i (resp. ν_i) l'unique valeur que prend f (resp. g) sur $]x_i, x_{i+1}[$.
Alors pour tout $t \in]x_i, x_{i+1}[$, $(\lambda f + g)(t) = \lambda\mu_i + \nu_i$, et donc $\lambda f + g$ est bien constante sur $]x_i, x_{i+1}[$.
Et donc $\lambda f + g \in \mathscr{E}([a, b])$. Notons que nous avons prouvé au passage qu'une subdivision adaptée à la fois à f et g est adaptée à $\lambda f + g$.
La preuve de stabilité de $\mathscr{E}([a, b])$ par produit est du même acabit. □

25.2.2 Intégrale des fonctions en escalier

Définition 25.8 – Soit f une fonction en escalier sur $[a, b]$, et soit $x_0 < x_1 < \cdots < x_n$ une subdivision adaptée. Alors on appelle **intégrale de f entre a et b**, et on note $\displaystyle\int_{[a,b]} f$, $\displaystyle\int_a^b f$ ou $\displaystyle\int_a^b f(t)\,dt$ le réel défini par

$$\int_a^b f(t)\,dt = \sum_{i=0}^{n-1} (x_{i+1} - x_i)\lambda_i$$

où pour tout $i \in [\![0, n-1]\!]$, λ_i est la valeur que prend f sur $]x_i, x_{i+1}[$.
Cette quantité ne dépend pas de la subdivision adaptée à f.

Remarque
On a
$$\lambda_i = f\left(\frac{x_i + x_{i+1}}{2}\right).$$

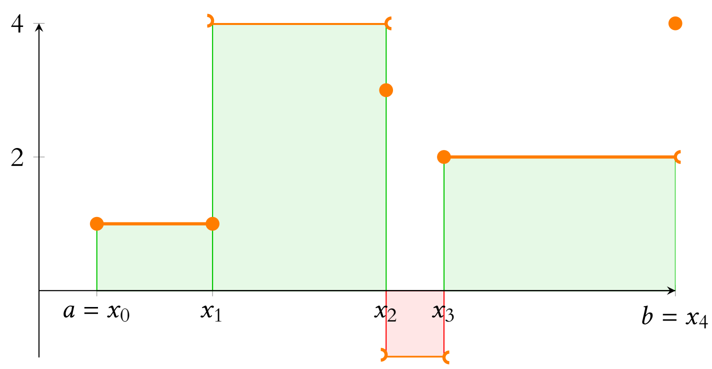


FIGURE 25.2 – L'interprétation en termes d'aire de l'intégrale d'une fonction en escalier est assez évidente. Notons que la valeur de f en les points qui composent la subdivision n'a aucune importance.

Avant de prouver que cette quantité est bien indépendante de la subdivision adaptée choisie, notons que l'interprétation en terme d'aire est assez évidente, du moins si on accepte que l'aire d'un rectangle est bien donnée par `longueur × largeur`.
En effet, $\lambda_i(x_{i+1} - x_i)$ est bien l'aire du rectangle délimité par l'axe des abscisse, la courbe de f et les droites d'équations $x = x_{i+1}$ et $x = x_i$.

Plus précisément
Nous parlons là d'une aire algébrique, qu'on autorise à être négative si $\lambda_i < 0$.

Démonstration. Il s'agit de prouver[10] que cette quantité ne dépend pas de la subdivision adaptée à f.

Étant donnée une subdivision $\sigma = (x_i)_{0 \leqslant i \leqslant n-1}$ adaptée à f, notons $I_\sigma(f) = \displaystyle\sum_{i=0}^{n-1}(x_{i+1} - x_i)\lambda_i$, où $\lambda_i = f\left(\dfrac{x_{i+1} + x_i}{2}\right)$.
Commençons par ajouter que si σ' est une subdivision obtenue à partir de σ par ajout d'un point, alors $I_{\sigma'}(f) = I_\sigma(f)$.
Notons par exemple c le point ajouté, et soit k tel que $x_k < c < x_{k+1}$.

[10] C'est assez facile à comprendre, plus pénible à écrire...

Alors σ' est encore adaptée à f, et on a

$$I_{\sigma'}(f) = \sum_{i=0}^{k-1}(x_{i+1}-x_i)f\left(\frac{x_{i+1}+x_i}{2}\right)+(c-x_k)f\left(\frac{c+x_k}{2}\right)+(x_{k+1}-c)f\left(\frac{x_{k+1}+c}{2}\right)+\sum_{i=k+1}^{n-1}(x_{i+1}-x_i)f\left(\frac{x_{i+1}+x_i}{2}\right).$$

Mais f est constante sur $]x_k, x_{k+1}[$ et donc $f\left(\frac{x_k+c}{2}\right) = f\left(\frac{c+x_{k+1}}{2}\right) = f\left(\frac{x_k+x_{k+1}}{2}\right)$.
Par conséquent,

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

et donc $I_\sigma(f) = I_{\sigma'}(f)$.
Une récurrence triviale prouve que l'ajout d'un nombre fini de points à une subdivision ne change pas la valeur de l'intégrale de f.

Considérons à présent σ_1 et σ_2 deux subdivisions de $[a,b]$ adaptées à f. Alors $\sigma_1 \cup \sigma_2$ est encore une subdivision adaptée à f, obtenue à partir de σ_1 par ajout d'un nombre fini de points[11].
Et donc $I_{\sigma_1}(f) = I_{\sigma_1\cup\sigma_2}(f)$. De même $I_{\sigma_2}(f) = I_{\sigma_1\cup\sigma_2(f)}$.
Il vient donc bien $I_{\sigma_1}(f) = I_{\sigma_2}(f)$. □

[11] Ceux de σ_2 qui n'étaient pas déjà dans σ_1.

Cette définition de l'intégrale des fonctions en escalier jouit déjà de bonnes propriétés qu'on attendrait pour une intégrale[12] :

[12] Et que nous avons prouvé pour les fonctions continues… en admettant le théorème fondamental de l'analyse !

Proposition 25.9 : *Soient $(f,g) \in \mathscr{E}([a,b])$ et soit $\lambda \in \mathbf{R}$. Alors :*

1. $\displaystyle\int_a^b (\lambda f(t)+g(t))\,dt = \lambda\int_a^b f(t)\,dt + \int_a^b g(t)\,dt$ *(linéarité de l'intégrale)*
2. *Si $f \geqslant 0$, alors* $\displaystyle\int_a^b f(t)\,dt \geqslant 0$.
3. *Si $f \leqslant g$, alors* $\displaystyle\int_a^b f(t)\,dt \leqslant \int_a^b g(t)\,dt$.
4. *Pour tout $c \in [a,b]$,* $\displaystyle\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt$ *(relation de Chasles).*

Remarque
Il n'y a pas équivalence : l'intégrale d'une fonction qui n'est pas de signe constant peut être positive.

Démonstration. 1. Si σ_1 est une subdivision adaptée à f et σ_2 une subdivision adaptée à g, alors $\sigma_1\cup\sigma_2$ est adaptée à f et à g à la fois, travaillons donc avec une telle subdivision $x_0 < x_1 < \cdots < x_n$.
Nous avons déjà prouvé que $\lambda f + g$ est encore constante sur chacun des $]x_i, x_{i+1}[$.
Et alors

$$\begin{aligned}\int_a^b (\lambda f(t)+g(t))\,dt &= \sum_{i=0}^{n-1}(x_{i+1}-x_i)\left(\lambda f\left(\frac{x_{i+1}+x_i}{2}\right)+g\left(\frac{x_i+x_{i+1}}{2}\right)\right)\\ &= \lambda\sum_{i=0}^{n-1}(x_{i+1}-x_i)f\left(\frac{x_i+x_{i+1}}{2}\right)+\sum_{i=0}^{n-1}(x_{i+1}-x_i)g\left(\frac{x_i+x_{i+1}}{2}\right)\\ &= \lambda\int_a^b f(t)\,dt + \int_a^b g(t)\,dt.\end{aligned}$$

2. Si f est positive, alors

$$\int_a^b f(t)\,dt = \sum_{i=0}^{n-1}\underbrace{(x_{i+1}-x_i)}_{\geqslant 0}\underbrace{f\left(\frac{x_i+x_{i+1}}{2}\right)}_{\geqslant 0} \geqslant 0.$$

Mieux
Le résultat reste valable si on demande f positive sauf en un nombre fini de points.

3. Il suffit d'appliquer la positivité à $g-f$, puis la linéarité de l'intégrale.
4. Si $c = a$ ou $c = b$, il n'y a pas grand chose à dire, si ce n'est constater que $\displaystyle\int_a^a f(t)\,dt = 0$.

Supposons donc $c \in]a, b[$. Considérons une subdivision $x_0 < x_1 < \cdots < x_n$ de $[a, b]$ adaptée à f, qui contient[13] c, et soit $k \in [\![0, n-1]\!]$ tel que $c = x_k$. Alors

$$\begin{aligned}\int_a^b f(t)\,dt &= \sum_{i=0}^{n-1}(x_{i+1} - x_i) f\left(\frac{x_i + x_{i+1}}{2}\right)\\ &= \sum_{i=0}^{k-1}(x_{i+1} - x_i) f\left(\frac{x_i + x_{i+1}}{2}\right) + \sum_{i=k}^{n-1}(x_{i+1} - x_i) f\left(\frac{x_i + x_{i+1}}{2}\right).\end{aligned}$$

Mais $x_0 < x_1 < \cdots < x_k = c$ est une subdivision de $[a, c]$ adaptée à f, de sorte que la première somme ci-dessus vaut $\int_a^c f(t)\,dt$.

Et de même, $x_k < \cdots < x_n$ est une subdivision de $[c, b]$ adaptée à f et la seconde somme vaut donc $\int_c^b f(t)\,dt$.

□

[13] Il est toujours possible de l'ajouter à une subdivision.

25.3 INTÉGRALE DES FONCTIONS CONTINUES PAR MORCEAUX

25.3.1 Fonctions continues par morceaux sur un segment

Les fonctions que nous allons pouvoir intégrer ne se limitent pas aux fonctions continues, nous allons autoriser également des fonctions qui ont un nombre fini de points de discontinuité. Mais attention pas toutes les fonctions qui ont un nombre fini de points de discontinuité, uniquement celles de la forme suivante :

Définition 25.10 – Une fonction $f : [a, b] \to \mathbf{R}$ est dite **continue par morceaux** s'il existe une subdivision $(x_i)_{0 \leqslant i \leqslant n}$ de $[a, b]$ telle que pour tout $i \in [\![0, n-1]\!]$, $f_{|]x_i, x_{i+1}[}$ est continue sur $]x_i, x_{i+1}[$ et prolongeable par continuité en x_i et en x_{i+1}.
Une telle subdivision est dite adaptée à f.
On note $\mathscr{C}_m([a, b])$ l'ensemble des fonctions continues par morceaux sur $[a, b]$.

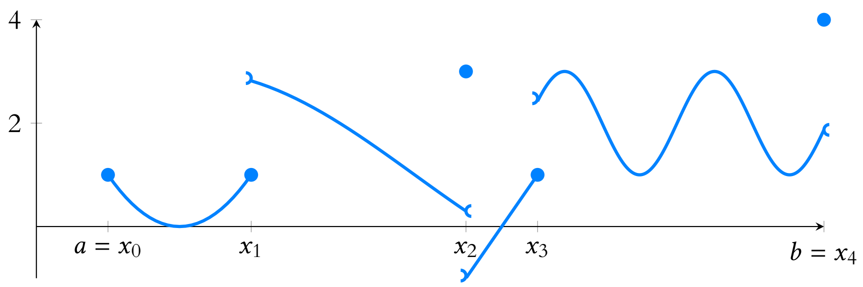


FIGURE 25.3 – Un exemple de fonction continue par morceaux.

Remarques. ▸ Comme pour les fonctions en escalier, il n'y a pas de contrainte sur la valeur de f aux points de la subdivision.
▸ En revanche, le fait que $f_{|]x_i, x_{i+1}[}$ soit prolongeable par continuité en x_i et en x_{i+1} se reformule de la manière suivante : f possède des limites finies à droite et à gauche en tous les x_i.
▸ Et comme pour les fonctions en escalier, il n'y a pas unicité d'une subdivision adaptée.
▸ Une fonction en escalier est continue par morceaux : $\mathscr{E}([a, b]) \subset \mathscr{C}_m([a, b])$.

Exemples 25.11

▸ Toute fonction continue sur le segment $[a, b]$ y est continue par morceaux[14]
▸ La restriction de la fonction partie entière à un segment est continue par morceaux sur ce segment.

[14] Et il suffit de prendre la subdivision à deux points $a < b$.

Une fonction continue par morceaux n'est pas seulement une fonction avec un nombre fini de points de discontinuité, il y a vraiment la condition sur les limites latérales en les x_i.

Par exemple, la fonction définie sur $[-1,1]$ par $x \mapsto \begin{cases} \frac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$ n'est pas continue par morceaux.

Proposition 25.12 : *L'ensemble $\mathscr{C}_m([a,b])$ est un sous-espace vectoriel et un sous-anneau de $\mathscr{F}([a,b],\mathbf{R})$.*

Démonstration. La preuve est à peu près la même que pour les fonctions en escalier.
Si f et g sont deux fonctions continues par morceaux, et si $a = x_0 < x_1 < x_n = b$ est une subdivision adaptée à la fois à f et à g, alors pour tout $\lambda \in \mathbf{R}$, $(\lambda f + g)_{|]x_i,x_{i+1}[}$ et $(fg)_{|]x_i,x_{i+1}[}$ sont continues sur $]x_i, x_{i+1}[$.
De plus, $f_{|]x_i,x_{i+1}[}$ et $g_{|]x_i,x_{i+1}[}$ possèdent des limites à droite en x_i et à gauche en x_{i+1}, donc il en est de même de $(\lambda f + g)_{|]x_i,x_{i+1}[}$ et $(fg)_{|]x_i,x_{i+1}[}$. □

Proposition 25.13 : *Une fonction continue par morceaux sur un segment est bornée.*

Démonstration. Soit f une fonction continue par morceaux sur $[a,b]$, et soit $(x_i)_{0\leqslant i\leqslant n}$ une subdivision adaptée à f.
Alors pour tout $i \in [\![0, n-1]\!]$, le prolongement par continuité de $f_{|]x_i,x_{i+1}[}$ à $[x_i, x_{i+1}]$ est une fonction continue sur le segment $[x_i, x_{i+1}]$ et donc est bornée. Notons M_i un majorant de sa valeur absolue, de sorte que pour tout $x \in]x_i, x_{i+1}[$, $|f(x)| \leqslant M_i$.
Si on pose $M = \max(M_0, M_1, \dots, M_{n-1}, |f(x_0)|, \dots, |f(x_n)|)$, alors pour tout $x \in [a,b]$, $|f(x)| \leqslant M$. □

Définition 25.14 – Plus généralement, une fonction f définie sur un intervalle[15] I est dite continue par morceaux sur I si pour tout segment $[a,b]$ de I, $f_{|[a,b]}$ est continue par morceaux sur $[a,b]$.

[15] Qui peut être ouvert et/ou non borné.

Exemple 25.15

La fonction partie entière est continue par morceaux sur $\mathbf{R}$.

25.3.2 Approximation uniforme par des fonctions en escalier

Définition 25.16 – Si f est une fonction bornée sur $[a,b]$, on pose $\|f\|_\infty = \sup\limits_{[a,b]} |f|$.

Terminologie
On parle de la «norme infinie» de f.

Notons qu'en particulier, $\|f\|_\infty$ est bien définie si f est continue sur $[a,b]$, ou si f est continue par morceaux sur $[a,b]$.
Graphiquement, $\|f\|_\infty$ est le plus petit réel M tel que le graphe de f soit compris entre les deux droites horizontales d'équations $y = M$ et $y = -M$.
Les propriétés qui suivent justifient que l'on appelle $\|\cdot\|_\infty$ une norme, mais vous n'étudierez ces objets qu'en seconde année.
Pour vous faire une intuition, remplacez dans l'énoncé suivant les fonctions par des vecteurs de $\mathbf{R}^2$ ou $\mathbf{R}^3$ et les $\|\cdot\|_\infty$ par les normes usuelles[16] des vecteurs.

[16] C'est-à-dire les longueurs !

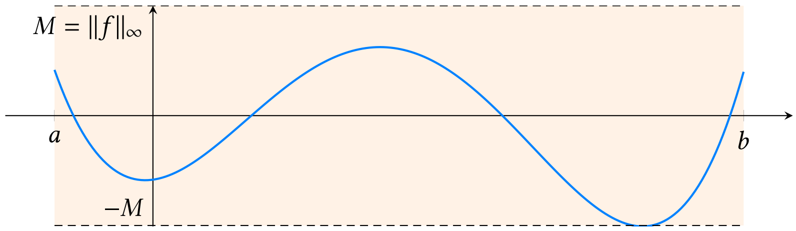


Figure 25.4 – Interprétation de la norme infinie.

Proposition 25.17 : *Soient f, g deux fonctions bornées sur $[a, b]$ et soit $\lambda \in \mathbf{R}$. Alors*

1. $\|\lambda f\|_\infty = |\lambda| \cdot \|f\|_\infty$
2. $\|f + g\|_\infty \leqslant \|f\|_\infty + \|g\|_\infty$ *(inégalité triangulaire)*
3. $\|f\|_\infty = 0 \Leftrightarrow f = 0$

Démonstration.
1. Pour tout $t \in [a, b]$, on a $|\lambda f(t)| = |\lambda| \cdot |f(t)|$.
   Donc en passant au sup, $\|\lambda f\|_\infty = |\lambda| \cdot \|f\|_\infty$.
2. Pour tout $t \in [a, b]$, on a $|f(t) + g(t)| \leqslant |f(t)| + |g(t)| \leqslant \|f\|_\infty + \|g\|_\infty$.
   Et donc, en passant au sup[17], $\|f + g\|_\infty \leqslant \|f\|_\infty + \|g\|_\infty$.
3. Il est évident que si f est la fonction nulle, alors $\|f\|_\infty = 0$.
   Inversement, si $\|f\|_\infty = 0$, alors, pour tout $t \in [a, b]$, $-0 \leqslant f(t) \leqslant 0$, et donc $f(t) = 0$, de sorte que f est la fonction nulle.

□

[17] Rappelons qu'il s'agit du plus petit des majorants.

Définition 25.18 – Si f et g sont deux fonctions bornées sur $[a, b]$, on appelle **distance uniforme entre f et g** le réel positif $\|f - g\|_\infty$.

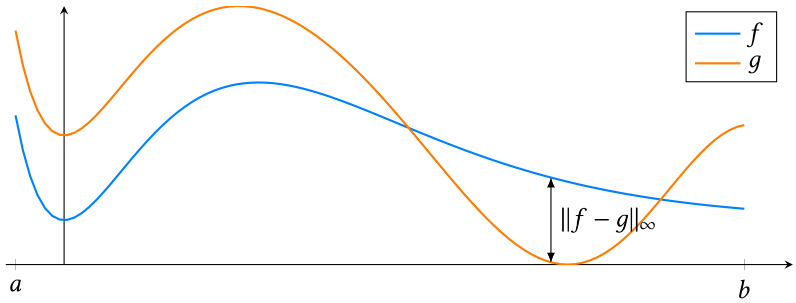


Figure 25.5 – Graphiquement, $\|f - g\|_\infty$ est la plus grande distance verticale entre les graphes de f et de g.

Théorème 25.19 (Approximation uniforme par des fonctions en escalier) : *Soit $f \in \mathscr{C}_m([a, b])$, et soit $\varepsilon > 0$. Alors il existe deux fonctions en escalier φ_ε^- et φ_ε^+ telles que*

1. $\varphi_\varepsilon^- \leqslant f \leqslant \varphi_\varepsilon^+$
2. $\|\varphi_\varepsilon^+ - \varphi_\varepsilon^-\|_\infty \leqslant \varepsilon$.

Pour le dire autrement, il est toujours possible de «coincer» f entre deux fonctions en escalier, qui ne sont jamais à distance supérieure à ε l'une de l'autre.

Notons qu'en particulier, en vertu de l'inégalité triangulaire,

$$\|f - \varphi_\varepsilon^-\|_\infty \leqslant \varepsilon \text{ et } \|f - \varphi_\varepsilon^+\|_\infty \leqslant \varepsilon.$$

Ce qui signifie que les deux fonctions[18] φ_ε^+ et φ_ε^- tiennent entièrement dans la bande de demi-largeur ε centrée autour de f.

[18] Ou plutôt leurs graphes...

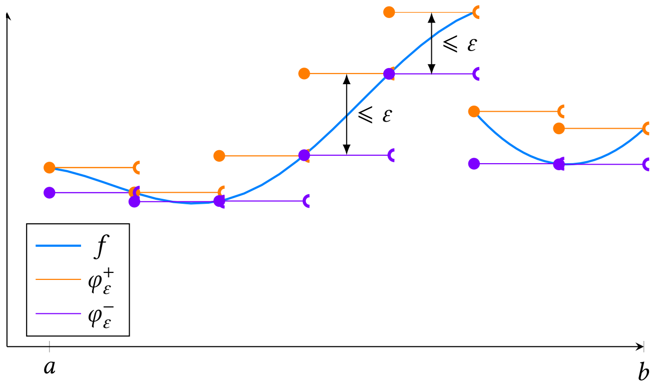


FIGURE 25.6 – Une fonction continue par morceaux «coincée» entre deux fonctions en escalier à distance moins de ε l'une de l'autre.

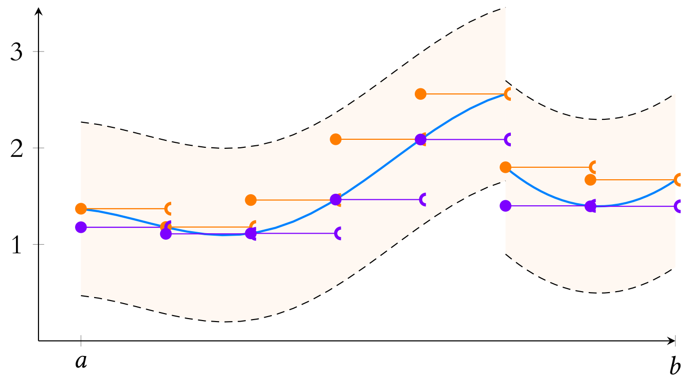


FIGURE 25.7 – La distance entre f et les deux fonctions φ_ε^+ et φ_ε^- est inférieure à ε.

Démonstration. ► Commençons par le cas où f est continue sur $[a, b]$.
Par le théorème de Heine, il existe $\eta > 0$ tel que pour $|x - y| \leqslant \eta$, $|f(x) - f(y)| \leqslant \varepsilon$.
Soit alors $n \in \mathbf{N}^*$ tel que $\dfrac{b-a}{n} < \eta$, et posons alors $(x_k)_{0\leqslant k\leqslant n}$ la subdivision régulière de $[a, b]$ obtenue en posant $x_k = a + k\dfrac{b-a}{n}$.
Pour tout $i \in [\![0, n-1]\!]$, posons $M_i = \max\limits_{t\in[x_i,x_{i+1}]} f(t)$ et $m_i = \min\limits_{t\in[x_i,x_{i+1}]} f(t)$.
Posons alors

$$\varphi_\varepsilon^+ : t \mapsto \begin{cases} M_i & \text{si } t \in [x_i, x_{i+1}[\\ f(b) & \text{si } t = b \end{cases} \quad \text{et} \quad \varphi_\varepsilon^- : t \mapsto \begin{cases} m_i & \text{si } t \in [x_i, x_{i+1}[\\ f(b) & \text{si } t = b \end{cases}$$

Il est alors évident que φ_ε^- et φ_ε^+ sont deux fonctions en escalier, et que $\varphi_\varepsilon^- \leqslant f \leqslant \varphi_\varepsilon^+$.
Reste à vérifier que $\|\varphi_\varepsilon^+ - \varphi_\varepsilon^-\|_\infty \leqslant \varepsilon$.
Par le théorème des bornes atteintes, M_i et m_i sont des valeurs atteintes par f en des points T_i et t_i de $[x_i, x_{i+1}]$. Mais on a alors $|T_i - t_i| \leqslant |x_{i+1} - x_i| \leqslant \dfrac{b-a}{n} \leqslant \eta$.
Et donc, par définition de η, $|M_i - m_i| \leqslant |f(T_i) - f(t_i)| \leqslant \varepsilon$.
Et par conséquent, pour tout $t \in [x_i, x_{i+1}[$,

$$|\varphi_\varepsilon^+(t) - \varphi_\varepsilon^-(t)| = \varphi_\varepsilon^+(t) - \varphi_\varepsilon^-(t) \leqslant M_i - m_i \leqslant \varepsilon.$$

Puisqu'en $t = b$, φ_ε^- et φ_ε^+ coïncident, on a donc bien

$$\forall t \in [a, b],\ |\varphi_\varepsilon^+(t) - \varphi_\varepsilon^-(t)| \leqslant \varepsilon.$$

Et donc $\|\varphi_\varepsilon^+ - \varphi_\varepsilon^-\|_\infty \leqslant \varepsilon$.

► Passons à présent au cas général d'une fonction continue par morceaux.
Soit donc $(x_i)_{0\leqslant i\leqslant n}$ une subdivision adaptée à f.
Pour tout $i \in [\![0, n-1]\!]$, notons $\widetilde{f_i}$ le prolongement par continuité de $f_{|]x_i,x_{i+1}[}$ à $[x_i, x_{i+1}]$.
Par ce qui précède, il existe donc deux fonctions en escalier $\varphi_{\varepsilon,i}^-$ et $\varphi_{\varepsilon,i}^+$ telles que

1. $\forall t \in]x_i, x_{i+1}[,\ \varphi^-_{\varepsilon,i}(t) \leqslant \underbrace{\widetilde{f_i}(t)}_{=f(t)} \leqslant \varphi^+_{\varepsilon,i}(t)$.
2. $\|\varphi^+_{\varepsilon,i} - \varphi^-_{\varepsilon,i}\|_\infty \leqslant \varepsilon$, c'est-à-dire que

$$\forall t \in]x_i, x_{i+1}[,\ \varphi^+_{\varepsilon,i}(t) - \varphi^-_{\varepsilon,i}(t) \leqslant \varepsilon.$$

Si on pose alors

$$\varphi^+_\varepsilon : t \mapsto \begin{cases} \varphi^+_{\varepsilon,i}(t) & \text{si } t \in]x_i, x_{i+1}[\\ f(x_i) & \text{si } \ t = x_i \end{cases} \text{ et } \varphi^-_\varepsilon : t \mapsto \begin{cases} \varphi^-_{\varepsilon,i}(t) & \text{si } t \in]x_i, x_{i+1}[\\ f(x_i) & \text{si } \ t = x_i \end{cases}$$

alors on vérifie sans grande difficultés qu'on a bien les propriétés attendues. □

Vous interpréterez l'an prochain ce résultat en parlant de la *densité* de l'ensemble des fonctions en escalier dans l'ensemble des fonctions continues par morceaux : on peut approcher une fonction continue par morceaux d'aussi près[19] qu'on le souhaite par des fonctions en escalier.

[19] Au sens de la norme infinie.

25.3.3 Intégrale des fonctions continues par morceaux

Définition 25.20 – Soit $f \in \mathscr{C}_m([a,b])$. On note :

1. $\mathscr{E}^+(f) = \{\varphi \in \mathscr{E}([a,b]) \mid f \leqslant \varphi\}$ l'ensemble des fonctions en escalier supérieures à f
2. $\mathscr{E}^-(f) = \{\varphi \in \mathscr{E}([a,b]) \mid f \geqslant \varphi\}$
3. $I^+(f) = \left\{\int_a^b \varphi(t)\,dt,\ \varphi \in \mathscr{E}^+(f)\right\}$ l'ensemble[20] des intégrales des fonctions en escalier supérieures à f.
4. $I^-(f) = \left\{\int_a^b \varphi(t)\,dt,\ \varphi \in \mathscr{E}^-(f)\right\}$

[20] De réels

Théorème 25.21 : *Soit $f \in \mathscr{C}_m([a,b])$. Alors $\inf I^+(f) = \sup I^-(f)$.*

On appelle intégrale de f entre a et b et on note $\int_a^b f(t)\,dt$ la valeur commune de ces deux bornes.
Si $f \in \mathscr{E}([a,b])$, alors cette définition de l'intégrale coïncide avec celle donnée précédemment.

Démonstration. Commençons par remarquer qu'il est à peu près évident que $\mathscr{E}^+(f)$ et $\mathscr{E}^-(f)$ sont non vides, il suffit par exemple de prendre les fonctions constantes égales à $\pm\|f\|_\infty$.
Par conséquent, $I^+(f)$ et $I^-(f)$ sont non vides.
Et puisque $\forall g \in \mathscr{E}^-(f), \forall h \in \mathscr{E}^+(f)$, $g \leqslant h$, et donc par croissance de l'intégrale[21], $\int_a^b g(t)\,dt \leqslant \int_a^b h(t)\,dt$.
Autrement dit tout élément de $I^+(f)$ est supérieur à tout élément de $I^-(f)$.
Mais alors $\int_a^b g(t)\,dt \leqslant \int_a^b h(t)\,dt$, et ceci étant valable pour tout $g \in \mathscr{E}^-(f)$, $\sup I^-(f)$ existe et $\sup I^-(f) \leqslant \int_a^b h(t)\,dt$.
Et ceci valant pour tout $h \in \mathscr{E}^+(f)$, $\sup I^-(f) \leqslant \inf I^+(f)$.
À présent soit $\varepsilon > 0$. Alors il existe deux fonctions en escalier φ^+_ε et φ^-_ε telles que $\varphi^-_\varepsilon \leqslant f \leqslant \varphi^+_\varepsilon$ et $\|\varphi^+_\varepsilon - \varphi^-_\varepsilon\|_\infty \leqslant \varepsilon$.
Alors $\varphi^-_\varepsilon \in \mathscr{E}^-(f)$, $\varphi^+_\varepsilon \in \mathscr{E}^+(f)$ et

$$\begin{aligned}\int_a^b \varphi^+_\varepsilon(t)\,dt - \int_a^b \varphi^-_\varepsilon(t)\,dt &= \int_a^b \left(\varphi^+_\varepsilon(t) - \varphi^-_\varepsilon(t)\right)\,dt \\ &\leqslant \int_a^b \varepsilon\,dt\end{aligned}$$

[21] Pour les fonctions en escalier.

C'est la linéarité de l'intégrale **pour les fonctions en escalier.**

C'est la croissance de l'intégrale **pour les fonctions en escalier.**

$$\leqslant (b-a)\varepsilon.$$

En particulier, $\inf I^+(f) \leqslant \displaystyle\int_a^b \varphi_\varepsilon^+(t)\,dt$, et donc

$$\inf I^+(f) - \int_a^b \varphi_\varepsilon^-(t)\,dt \leqslant (b-a)\varepsilon.$$

Mais cette fois, $\displaystyle\int_a^b \varphi_\varepsilon-(t)\,dt \leqslant \sup I^-(f)$, et donc $\inf I^+(f) - \sup I^-(f) \leqslant (b-a)\varepsilon$.
Ceci étant vrai pour tout $\varepsilon > 0$,

$$\inf I^+(f) - \sup I^-(f) \leqslant 0 \Leftrightarrow \inf I^+(f) \leqslant \sup I^-(f).$$

Comme nous avions déjà prouvé l'inégalité réciproque, on a bien l'égalité annoncée.

Enfin, si f est une fonction en escalier, alors $f \in \mathscr{E}^+(f)$ et $f \in \mathscr{E}^-(f)$.
Mais par croissance de l'intégrale sur les fonctions en escalier, $\displaystyle\int_a^b f(t)\,dt$ (intégrale au sens des fonctions en escalier) est un minorant de $I^+(f)$. C'est donc le plus petit élément de $I^+(f)$, et donc $\inf I^+(f)$. □

On a alors une bonne interprétation de l'intégrale en termes d'aire.
En effet, nous comprenons bien ce que doit être l'aire d'un rectangle. Autrement dit, on accepte volontiers que la formule «*base* × *hauteur*» définisse ce qu'on a envie d'appeler l'aire d'un rectangle. Par exemple, qu'un rectangle de 2m par 1m possède une aire de 2m^2.
Dès lors, l'intégrale d'une fonction en escalier est bien l'aire de la partie du plan sous la courbe, qui est une union de rectangles.

Ce que nous dit la définition ci-dessus, c'est que l'intégrale de f est[22] la borne supérieure des aires des parties du plan situées sous la courbe de f et qui sont union de rectangles. On a légitimement envie d'appeler cette quantité l'aire sous la courbe de f.

[22] Au moins dans le cas de fonctions à valeurs positives.

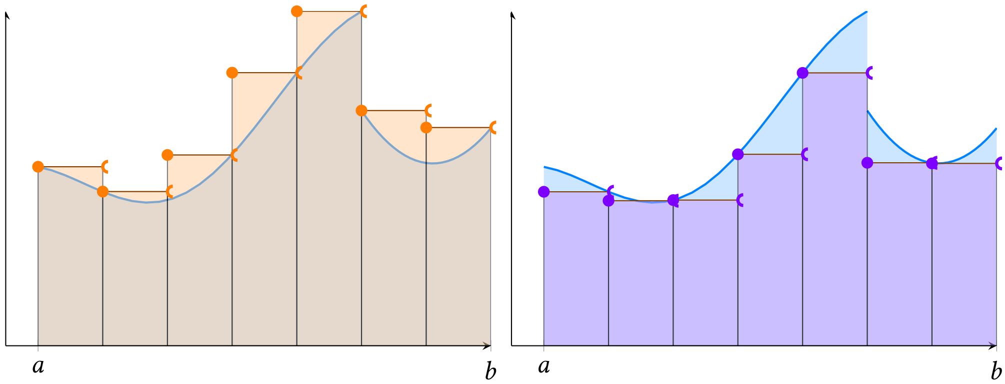


FIGURE 25.8 – L'aire sous la courbe de f est encadrée par les aires sous les courbes des fonctions de $\mathscr{E}^+(f)$ et celles de $\mathscr{E}^-(f)$.

Nous n'irons pas plus loin dans la définition de ce qu'est une aire, et il nous faudrait aller vers l'intégrale/la mesure de Lebesgue pour donner un sens plus précis à cette notion.
Autrement dit, nous ne donnons un sens qu'à l'aire de parties du plan qui sont délimitées par deux droites verticales[23], l'axe des abscisses et la courbe représentative d'une fonction continue par morceaux.
Par exemple, nous ne disons rien de l'aire d'un cercle. Même s'il est possible de ruser un peu : l'aire d'un demi-cercle étant égale à $\displaystyle\int_{-1}^1 \sqrt{1-t^2}\,dt$.

Enfin, prouvons que l'intégrale ainsi définie possède bien les propriétés qu'on lui connaît dans le cas des fonctions continues.
Les preuves les plus désagréables sont celles des résultats qu'on pense les plus évidents : la linéarité et Chasles. Ces preuves sont assez techniques...

Déjà fait ?

Avons-nous déjà prouvé ces propriétés pour les fonctions continues ? Oui et non !
Les preuves données dans le chapitre 8 sont correctes, mais reposent toujours sur le théorème fondamental de l'analyse, qui n'a toujours pas été prouvé (patience, plus que quelques pages à attendre...).

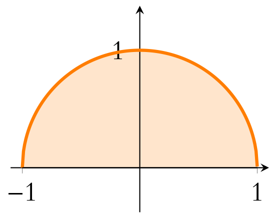


FIGURE 25.9 – La fonction $x \mapsto \sqrt{1-x^2}$.

Proposition 25.22 (Propriétés de l'intégrale) : *Soient $f, g \in \mathscr{C}_m([a,b])$. Alors*

1. $\forall \lambda \in \mathbf{R}$, $\displaystyle\int_a^b (\lambda f(t)+g(t))\,dt = \lambda \int_a^b f(t)\,dt + \int_a^b g(t)\,dt$ *(linéarité de l'intégrale)*
2. *si* $f \geqslant 0$, *alors* $\displaystyle\int_a^b f(t)\,dt \geqslant 0$ *(positivité de l'intégrale)*
3. *si* $f \leqslant g$, *alors* $\displaystyle\int_a^b f(t)\,dt \leqslant \int_a^b g(t)\,dt$ *(croissance de l'intégrale)*
4. $\displaystyle\left|\int_a^b f(t)\,dt\right| \leqslant \int_a^b |f(t)|\,dt$ *(inégalité triangulaire)*
5. *si f et g ne diffèrent qu'en un nombre fini de points, alors* $\displaystyle\int_a^b f(t)\,dt = \int_a^b g(t)\,dt$.
6. *pour tout* $c \in]a,b[$, $\displaystyle\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt$.

Remarque
Si f est continue par morceaux, $|f|$ l'est aussi.

Démonstration. 1. Fixons $\varepsilon > 0$, et soient $\varphi_\varepsilon^+, \varphi_\varepsilon^-, \psi_\varepsilon^+, \psi_\varepsilon^-$ des fonctions en escalier telles que

$$\begin{cases} \varphi_\varepsilon^- \leqslant f \leqslant \varphi_\varepsilon^+ \\ \|\varphi_\varepsilon^+ - \varphi_\varepsilon^-\| \leqslant \varepsilon \end{cases} \quad \text{et} \quad \begin{cases} \psi_\varepsilon^- \leqslant g \leqslant \psi_\varepsilon^+ \\ \|\psi_\varepsilon^+ - \psi_\varepsilon^-\| \leqslant \varepsilon \end{cases}$$

Si $\lambda \geqslant 0$, alors $\lambda\varphi_\varepsilon^- + \psi_\varepsilon^- \leqslant \lambda f + g \leqslant \lambda\varphi_\varepsilon^+ + \psi_\varepsilon^+$, et les fonctions $\lambda\varphi_\varepsilon^- + \psi_\varepsilon^-$ et $\lambda\varphi_\varepsilon^+ + \psi_\varepsilon^+$ sont des fonctions en escalier.
On a donc[24]

$$\int_a^b (\lambda\varphi_\varepsilon^-(t) + \psi_\varepsilon^-(t))\,dt \leqslant \int_a^b (\lambda f(t) + g(t))\,dt \leqslant \int_a^b (\lambda\varphi_\varepsilon^+(t) + \psi_\varepsilon^+(t))\,dt.$$

[24] C'est la définition de l'intégrale, comme borne sup de $I^-(\lambda f + g)$ et comme borne inf de $I^+(\lambda f + g)$.

Mais pour les mêmes raisons,

$$\lambda \int_a^b \varphi_\varepsilon^-(t)\,dt \leqslant \lambda \int_a^b f(t)\,dt \leqslant \int_a^b \varphi_\varepsilon^+(t)\,dt \quad \text{et} \quad \int_a^b \psi_\varepsilon^-(t)\,dt \leqslant \int_a^b g(t)\,dt \leqslant \int_a^b \psi_\varepsilon^+(t)\,dt.$$

Donc par somme d'inégalités, en notant

$$A = \int_a^b (\lambda f(t) + g(t))\,dt - \lambda \int_a^b f(t)\,dt - \int_a^b g(t)\,dt,$$

il vient

$$\int_a^b (\lambda\varphi_\varepsilon^-(t)+\psi_\varepsilon^-(t))\,dt - \lambda \int_a^b \varphi_\varepsilon^+(t)\,dt - \int_a^b \psi_\varepsilon^+(t)\,dt \leqslant A \leqslant \int_a^b (\lambda\varphi_\varepsilon^+(t)+\psi_\varepsilon^+(t))\,dt - \lambda \int_a^b \varphi_\varepsilon^-(t)\,dt - \int_a^b \psi_\varepsilon^-(t)\,dt. \qquad (\star)$$

Intéressons-nous au membre de droite : par linéarité de l'intégrale **pour les fonctions en escalier**, on a

$$\begin{aligned}\int_a^b (\lambda\varphi_\varepsilon^+(t) + \psi_\varepsilon^+(t))\,dt - \lambda \int_a^b \varphi_\varepsilon^-(t)\,dt - \int_a^b \psi_\varepsilon^-(t)\,dt &= \lambda \int_a^b (\varphi_\varepsilon^+(t) - \varphi_\varepsilon^-(t))\,dt + \int_a^b (\psi_\varepsilon^+(t) - \psi_\varepsilon^-(t))\,dt \\ &\leqslant \lambda \int_a^b \varepsilon\,dt + \int_a^b \varepsilon\,dt\end{aligned}$$

$$\leqslant (\lambda+1)(b-a)\varepsilon.$$

Sur le même principe, on prouve que le membre de gauche de $(\star)$ est $-(\lambda+1)\varepsilon(b-a)$, et donc que $|A| \leqslant (1+\lambda)\varepsilon(b-a)$.
Ceci étant vrai pour tout $\varepsilon > 0$, on a donc $A = 0$, de sorte que

$$\int_a^b (\lambda f(t) + g(t))\,dt = \lambda \int_a^b f(t)\,dt + \int_a^b g(t)\,dt.$$

On procède de même dans le cas où $\lambda < 0$, mais en prenant soin de changer le sens des inégalités lorsqu'on multiplie $\varphi_\varepsilon^- \leqslant f \leqslant \varphi_\varepsilon^+$ par λ.
On obtient au final $|A| \leqslant (1-\lambda)(b-a)\varepsilon$, et on conclut de même.

2. Si f est positive sur $[a,b]$, alors la fonction nulle sur $[a,b]$ est dans $\mathscr{E}^-(f)$.
Et donc $0 = \displaystyle\int_a^b 0\,dt \in I^-(f)$.
Et par conséquent, $\displaystyle\int_a^b f(t)\,dt = \sup I^-(f) \geqslant 0$.
3. Il suffit de cumuler les deux points précédents : la fonction $g - f$ est positive, et donc d'intégrale positive.
Mais par linéarité de l'intégrale, $\displaystyle\int_a^b (g(t)-f(t))\,dt = \int_a^b g(t)\,dt - \int_a^b f(t)\,dt$, et donc
$$\int_a^b g(t)\,dt - \int_a^b f(t)\,dt \geqslant 0 \Leftrightarrow \int_a^b g(t)\,dt \geqslant \int_a^b f(t)\,dt.$$
4. On a $-|f| \leqslant f \leqslant |f|$.
Et donc par croissance et linéarité[25] de l'intégrale,
$$-\int_a^b |f(t)|\,dt \leqslant \int_a^b f(t)\,dt \leqslant \int_a^b |f(t)|\,dt.$$
Et donc $\displaystyle\left|\int_a^b f(t)\,dt\right| \leqslant \int_a^b |f(t)|\,dt$.
5. Si f et g ne diffèrent qu'en un nombre fini de points, alors $f-g$ est une fonction en escalier, dont l'intégrale est clairement nulle.
Donc par linéarité de l'intégrale,
$$\int_a^b (g(t)-f(t))\,dt = 0 \Leftrightarrow \int_a^b g(t)\,dt = \int_a^b f(t)\,dt.$$
6. Définissons trois fonctions f_1, f_2 et f_3 en posant
$$f_1(x) = \begin{cases} f(x) & \text{si } a \leqslant x < c \\ 0 & \text{si } c \leqslant x \leqslant b \end{cases},\ f_2(x) = \begin{cases} 0 & \text{si } a \leqslant x \leqslant c \\ f(x) & \text{si } c < x \leqslant b \end{cases},\ f_3(x) = \begin{cases} f(c) & \text{si } x = c \\ 0 & \text{sinon} \end{cases}$$
de sorte que f_1, f_2, f_3 sont continues par morceaux, avec $f = f_1 + f_2 + f_3$.
Donc par linéarité de l'intégrale, $\displaystyle\int_a^b f(t)\,dt = \int_a^b f_1(t)\,dt + \int_a^b f_2(t)\,dt + \int_a^b f_3(t)\,dt$.
Mais f_3 est une fonction en escalier, nulle sauf en c, de sorte que $\displaystyle\int_a^b f_3(t)\,dt = 0$.
Prouvons que $\displaystyle\int_a^b f_1(t)\,dt = \int_a^c f_1(t)\,dt$.
Soit φ une fonction en escalier sur $[a,c]$, supérieure à f sur cet intervalle.
Soit alors $\widetilde{\varphi}$ la fonction définie sur $[a,b]$ par $\widetilde{\varphi}(x) = \begin{cases} \varphi(x) & \text{si } a \leqslant x \leqslant c \\ 0 & \text{si } c < x \leqslant b \end{cases}$.
Alors $\widetilde{\varphi}$ est une fonction en escalier majorant f, et par la relation de Chasles pour les fonctions en escalier,
$$\int_a^b \widetilde{\varphi}(t)\,dt = \int_a^c \widetilde{\varphi}(t)\,dt + \underbrace{\int_c^b \widetilde{\varphi}(t)\,dt}_{=0} = \int_a^c \varphi(t)\,dt.$$

Remarque
Si f est continue par morceaux, $|f|$ l'est aussi.

[25] Pour sortir le signe moins de l'intégrale.

Astuce
Un moyen pratique de définir ces fonctions serait d'écrire
$f_1 = f \times \mathbb{1}_{[a,c[}, f_3 = f\mathbb{1}_{\{a\}}, \dots$

Autrement dit
On prend φ dans $\mathscr{E}^+(f_{|[a,c]})$.

Mais par définition de l'intégrale comme borne inférieure de $I^+(f)$, on a donc $\displaystyle\int_a^b f_1(t)\,dt \leqslant \int_a^b \widetilde{\varphi}(t)\,dt$.

Et donc $\displaystyle\int_a^b f_1(t)\,dt$ est un minorant de $I^+((f_1)_{|[a,c]})$.

Donc $\displaystyle\int_a^b f_1(t)\,dt \leqslant \int_a^c f_1(t)\,dt$.

Sur le même principe, à l'aide de fonctions en escalier inférieures à f_1 sur $[a,c]$, on prouve l'inégalité réciproque, et donc l'égalité.

Sur un principe similaire, on prouve que $\displaystyle\int_a^b f_2(t)\,dt = \int_c^b f_2(t)\,dt$.

Et donc il vient bien

$$\int_a^b f(t)\,dt = \int_a^c f_1(t)\,dt + \int_c^b f_2(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt.$$

□

Enfin, ajoutons une propriété qui est propre aux fonctions continues[26], et qui aurait pu être prouvée plus tôt.

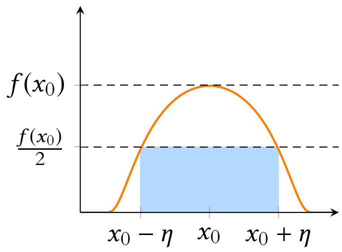
[26] Et pas aux fonctions continues par morceaux.

Proposition 25.23 : *Soit $f \in \mathscr{C}([a,b],\mathbf{R})$, de signe constant.*
Alors $\displaystyle\int_a^b f(t)\,dt = 0$ si et seulement si f est la fonction nulle.

Démonstration. Le sens $\Leftarrow$ est évident.
Pour l'autre sens, supposons, quitte à changer f en son opposé que f est positive.
Prouvons la contraposée en montrant que si f n'est pas la fonction nulle, alors son intégrale est non nulle. Comme on sait déjà[27] que cette intégrale est positive, il s'agit de prouver qu'elle est strictement positive.
Soit donc $x_0 \in [a,b]$ tel que $f(x_0) > 0$.
Puisque f est continue, on peut supposer que $x_0 \in]a,b[$ est un point intérieur à $]a,b[$.
Par définition de la continuité, en prenant $\varepsilon = \frac{f(x_0)}{2}$, il existe $\eta > 0$ tel que pour $x \in]x_0-\eta, x_0+\eta[\cap [a,b]$, $|f(x) - f(x_0)| \leqslant \varepsilon$.
Quitte à diminuer η, on peut supposer que $]x_0-\eta, x_0+\eta[\subset [a,b]$.
Et alors pour $x \in]x_0-\eta, x_0+\eta[$, $f(x) \geqslant \dfrac{f(x_0)}{2}$.
Ainsi, f est minorée par la fonction en escalier $\varphi : x \mapsto \begin{cases} \frac{f(x_0)}{2} & \text{si } x \in]x_0-\eta, x_0+\eta[\\ 0 & \text{sinon} \end{cases}$.
Mais l'intégrale de φ vaut $\eta f(x_0) > 0$, donc

$$\int_a^b f(x)\,dx \geqslant \int_a^b \varphi(x)\,dx > 0.$$

[27] C'est la positivité de l'intégrale.



□

25.3.4 Extension au cas $b \leqslant a$

Les définitions données ci-dessus d'intégrale ne sont valables que pour des intégrales dont les bornes sont «dans le bon sens».
Si $b < a$, par définition, on pose

$$\int_a^b f(t)\,dt = -\int_b^a f(t)\,dt.$$

La linéarité est alors conservée, mais ce n'est pas le cas de toutes les propriétés qui ont trait à la relation d'ordre (donc la positivité, la croissance et l'inégalité triangulaire).
Si on a besoin d'utiliser des inégalités dans des intégrales, on commencera par remettre les bornes dans le bon sens.

En revanche, la relation de Chasles est conservée :

Proposition 25.24 : *Soit f une fonction continue par morceaux sur un intervalle I. Alors* $\forall (a,b,c) \in I^3$, $\displaystyle\int_a^c f(t)\,dt = \int_a^b f(t)\,dt + \int_b^c f(t)\,dt$.

Démonstration. Le résultat a déjà été prouvé si $a < b < c$, et il est trivial si $a = b$ ou $b = c$.
Notons que pour tout $(m, x, y) \in I^3$ avec $m \leqslant x$ et $m \leqslant y$, on a

$$\int_x^y f(t)\,dt = \int_m^y f(t)\,dt - \int_m^x f(t)\,dt.$$

En effet, si $x \leqslant y$, alors par Chasles[28]

$$\int_x^y f(t)\,dt + \int_m^x f(t)\,dt = \int_m^y f(t)\,dt.$$

[28] Version «bornes dans le bon sens».

Et si $y < x$, alors

$$\int_m^y f(t)\,dt + \int_y^x f(t)\,dt = \int_m^x f(t)\,dt.$$

Posons donc $m = \min(a, b, c)$, de sorte que

$$\begin{aligned}\int_a^b f(t)\,dt + \int_b^c f(t)\,dt &= \int_m^b f(t)\,dt - \int_m^a f(t)\,dt + \int_m^c f(t)\,dt - \int_m^b f(t)\,dt \\ &= \int_m^c f(t)\,dt - \int_m^a f(t)\,dt \\ &= \int_a^c f(t)\,dt.\end{aligned}$$

□

25.4 Intégrales et calcul différentiel

25.4.1 Le théorème fondamental de l'analyse

Il est enfin grand temps de prouver un théorème annoncé depuis longtemps :

Théorème 25.25 : *Soit I un intervalle de $\mathbf{R}$, soit $a \in I$ et soit $f \in \mathscr{C}^0(I, \mathbf{K})$. Alors* $F : x \mapsto \displaystyle\int_a^x f(t)\,dt$ *est l'unique primitive de f sur I qui s'annule en a.*

Remarque
Notons que F est bien définie, puisque f est une fonction continue (et donc continue par morceaux) sur le segment $[a, x]$ (ou $[x, a]$ si $x \leqslant a$).

Démonstration. L'unicité d'une telle primitive ayant déjà été prouvée, il s'agit surtout de prouver que F est bien une primitive de f.
Soit donc $x_0 \in I$, et soit $\varepsilon > 0$.
Puisque f est continue en x_0, il existe $\eta > 0$ tel que

$$\forall t \in I \cap]x_0 - \eta, x_0 + \eta[, \ |f(t) - f(x_0)| \leqslant \varepsilon.$$

Donc en particulier, pour $x \in I \cap]x_0 - \eta, x_0 + \eta[, x \neq x_0$, on a

$$\begin{aligned}\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| &= \left|\frac{1}{x - x_0}\int_{x_0}^x f(t)\,dt - f(x_0)\right| \\ &= \left|\frac{1}{x - x_0}\int_{x_0}^x f(t)\,dt - \frac{1}{x - x_0}\int_{x_0}^x f(x_0)\,dt\right| \\ &= \frac{1}{|x - x_0|}\left|\int_{x_0}^x (f(t) - f(x_0))\,dt\right|.\end{aligned}$$

Si $x \geqslant x_0$, alors on peut utiliser l'inégalité triangulaire[29] :

$$\left|\int_{x_0}^x (f(t) - f(x_0))\,dt\right| \leqslant \int_{x_0}^x |f(t) - f(x_0)|\,dt \leqslant \int_{x_0}^x \varepsilon\,dt \leqslant \varepsilon(x - x_0).$$

[29] Celle-ci nécessite vraiment d'avoir les bornes de l'intégrale «dans le bon sens».

Si $x < x_0$, il faut prendre quelques précautions supplémentaires :

$$\begin{aligned}\left|\int_{x_0}^{x}(f(t)-f(x_0))\,dt\right| &= \left|\int_{x}^{x_0}(f(t)-f(x_0))\,dt\right|\\ &\leqslant \int_{x}^{x_0}|(f(t)-f(x_0))|\,dt\\ &\leqslant (x_0-x)\varepsilon \leqslant |x-x_0|\varepsilon.\end{aligned}$$

Dans tous les cas, on obtient que pour $|x-x_0|<\eta$,

$$\left|\frac{F(x)-F(x_0)}{x-x_0}-f(x_0)\right| \leqslant \frac{1}{|x_0-x|}|x_0-x|\varepsilon \leqslant \varepsilon.$$

Et donc c'est bien la définition de $\displaystyle\lim_{x\to x_0}\frac{F(x)-F(x_0)}{x-x_0}=f(x_0)$.
Donc F est dérivable en x_0, avec $F'(x_0)=f(x_0)$.
Ceci étant vrai pour tout $x_0\in I$, F est bien une primitive de f sur I. □

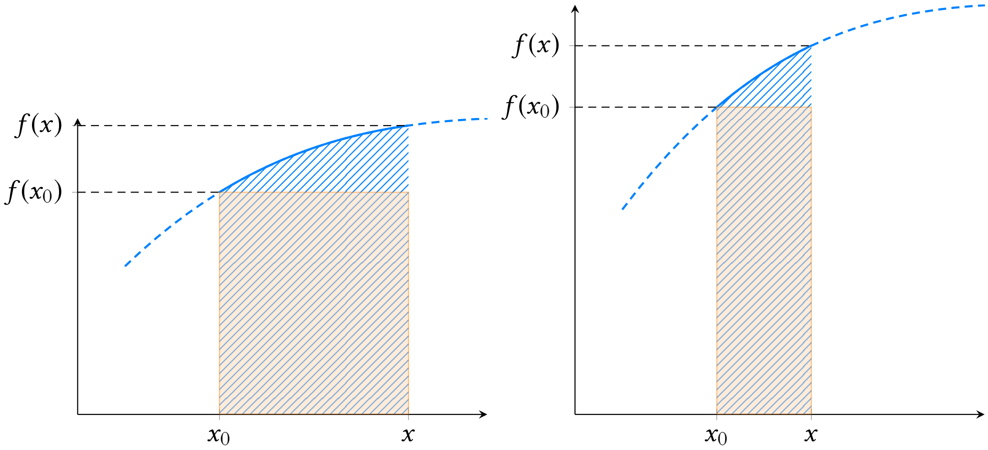
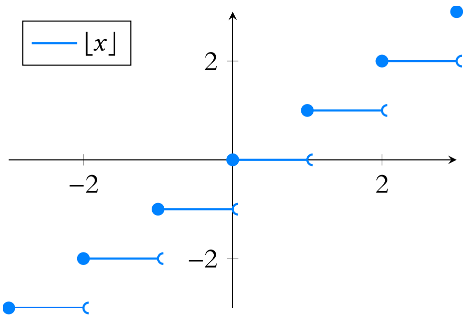
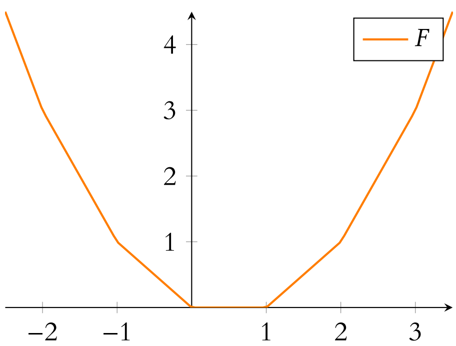


FIGURE 25.10 – Plus x est proche de x_0, et plus $\displaystyle\int_{x_0}^{x}f(t)\,dt$ est proche de l'aire d'un rectangle de largeur $|x-x_0|$ et de hauteur $f(x_0)$.

⚠ Ce théorème n'est plus valable pour des fonctions qui ne sont pas continues mais uniquement continues par morceaux.
Par exemple, $F : x\mapsto \displaystyle\int_0^x \lfloor t\rfloor\,dt$ n'est pas une primitive de la fonction partie entière. En effet, il est assez facile de se convaincre[30] que F est dérivable à gauche et à droite en $k\in\mathbf{Z}$, mais avec $F_g'(k)=k-1$ et $F_d'(k)=k$, donc F n'est pas dérivable en k. Une autre raison, un peu





plus subtile est un résultat[31] connu sous le nom de théorème de Darboux : la dérivée d'une fonction dérivable sur un intervalle doit satisfaire la propriété des valeurs intermédiaires, au sens où l'image d'un intervalle doit être un intervalle. Or ce n'est clairement pas une

[30] Faire le calcul !

[31] Qui figurait dans le TD 20.

propriété vérifiée par la fonction partie entière.

Soyons bien clair : ce théorème n'est pas une autorisation à dériver sans la moindre précaution toutes les fonctions qui seraient définies par une intégrale.

Tel quel, il ne dit pas que $x \mapsto \displaystyle\int_x^{x^2} e^{-t}\,dt$ ou $x \mapsto \displaystyle\int_1^2 \sqrt{t+\cos(x)}\,dt$ sont dérivables, et donne encore moins leurs dérivées.

L'énoncé du théorème fondamental de l'analyse ne vaut que pour des fonctions de la forme $x \mapsto \displaystyle\int_a^x f(t)\,dt$, c'est-à-dire que :

1. la borne «du bas» de l'intégrale ne dépend pas de x
2. celle «du haut» est égale à x
3. l'intégrande[32] ne dépend pas de x

[32] Nom (*masculin*) désignant la fonction intégrée.

Exemples 25.26

▶ $x \mapsto \displaystyle\int_0^x e^{-t^2}\,dt$ est l'unique primitive de $x \mapsto e^{-x^2}$ qui s'annule en 0.

▶ En revanche, moyennant quelques arguments supplémentaires, le théorème fondamental de l'analyse peut aider à dériver d'autres fonctions définies par des intégrales.

Par exemple, considérons la fonction f définie sur $\mathbf{R}_+^*$ par $f(x) = \displaystyle\int_0^1 \sqrt{1+x^2e^{2t}}\,dt$.

Alors, grâce au changement de variable $u = xe^t$, pour lequel $t = \ln(u) - \ln(x)$ et donc $dt = \dfrac{du}{u}$, il vient, pour tout $x > 0$, $f(x) = \displaystyle\int_x^{ex} \frac{\sqrt{1+u^2}}{u}\,du$.

Notons que cette fois, l'intégrande ne dépend plus de x, mais en revanche les bornes de l'intégrale ne permettent pas encore directement d'appliquer le théorème fondamental de l'analyse.

Posons alors $F : x \mapsto \displaystyle\int_1^x \frac{\sqrt{1+u^2}}{u}\,du$. Par le théorème fondamental de l'analyse, F est $\mathscr{C}^1$ sur $\mathbf{R}_+^*$, avec $F'(x) = \dfrac{\sqrt{1+x^2}}{x}$.

On a alors $f(x) = F(ex) - F(x)$. Donc déjà f est $\mathscr{C}^1$ par opérations sur les fonctions $\mathscr{C}^1$, et de plus

$$f'(x) = eF'(ex) - F'(x) = e\frac{\sqrt{1+e^2x^2}}{ex} - \frac{\sqrt{1+x^2}}{x} = \frac{\sqrt{1+e^2x^2} - \sqrt{1+x^2}}{x}.$$

Lorsqu'il n'est pas possible de se ramener au théorème fondamental de l'analyse pour étudier la dérivabilité d'une fonction définie par une intégrale, le dernier recours est bien souvent[33] le retour à la définition de la dérivée comme limite du taux d'accroissement.

[33] En réalité, vous aurez en seconde année des théorèmes assez puissants qui permettent l'étude de nombreuses fonctions définies par des intégrales.

Corollaire 25.27 – *Soit f une fonction continue sur un intervalle I. Alors f admet des primitives sur I.*
De plus, pour tout $(a,b) \in I^2$, et pour toute primitive F de f sur I,

$$\int_a^b f(t)\,dt = F(b) - F(a).$$

Démonstration. Pour l'existence d'une primitive, il suffit de considérer $x \mapsto \displaystyle\int_a^x f(t)\,dt$.

Pour le second point, nous avons déjà prouvé que deux primitives diffèrent d'une constante, et donc que la quantité $F(b) - F(a)$ ne dépend pas de la primitive F de f choisie.

En particulier, si on prend $F : x \mapsto \displaystyle\int_a^x f(t)\,dt$, il vient $F(b) - F(a) = F(b) = \displaystyle\int_a^b f(t)\,dt$. □

25.4.2 Techniques de calcul d'intégrales

Nous ne donnerons pas dans ce chapitre de nouveaux outils pour le calcul d'intégrales, mais notons que maintenant que l'intégrale est correctement définie, tout ce qui a été fait précédemment, notamment sur le changement de variable et l'intégration par parties reste valable.

Attention toutefois aux hypothèses : l'intégration par parties nécessite toujours des fonctions de classe $\mathscr{C}^1$, et le changement de variable un intégrande continu.

Si on souhaite utiliser l'une ou l'autre de ces méthodes pour calculer des intégrales de fonctions continues par morceaux, il faudra commencer par utiliser la relation de Chasles afin de se ramener à des morceaux où l'utilisation de ces théorèmes est légitime.

Ajoutons tout de même les résultats suivants, graphiquement évidents.

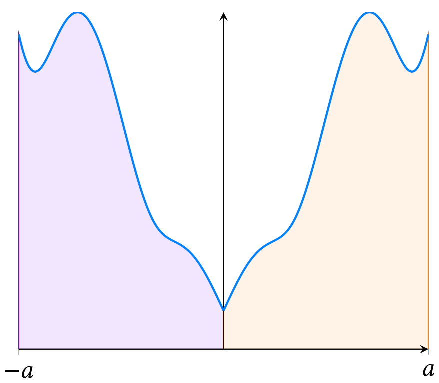
Proposition 25.28 (Intégrale des fonctions paires/impaires/périodiques) :

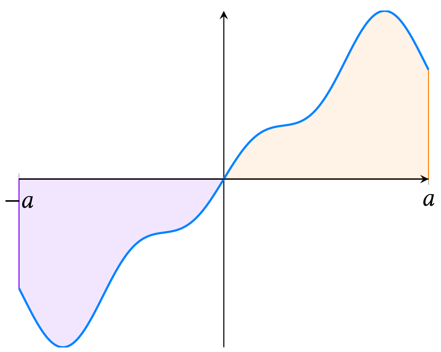
1. *Soit* $f \in \mathscr{C}_m([-a,a],\mathbf{R})$ *une fonction paire.*
   Alors $\displaystyle\int_{-a}^{a} f(t)\,dt = 2\int_0^a f(t)\,dt$.
2. *Soit* $f \in \mathscr{C}_m([-a,a],\mathbf{R})$ *une fonction impaire.*
   Alors $\displaystyle\int_{-a}^{a} f(t)\,dt = 0$.
3. *Soit* $f \in \mathscr{C}_m(\mathbf{R},\mathbf{R})$ *une fonction* T*-périodique. Alors pour tout* $a \in \mathbf{R}$, $\displaystyle\int_a^{a+T} f(t)\,dt = \int_0^T f(t)\,dt$.

Démonstration. Prouvons-le pour une fonction continue, le cas général s'en déduira à l'aide de la relation de Chasles.
Pour les deux premières assertions, procédons au changement de variable $x = -t$. On a alors, dans le cas d'une fonction paire,

$$\int_{-a}^{0} f(t)\,dt = -\int_a^0 f(-x)\,dx = \int_0^a f(x)\,dx.$$

Et dans le cas d'une fonction impaire, $\displaystyle\int_{-a}^{0} f(t)\,dt = -\int_a^0 f(-x)\,dx = -\int_0^a f(x)\,dx$.
Puis la relation de Chasles $\displaystyle\int_{-a}^{a} f(t)\,dt = \int_{-a}^{0} f(t)\,dt + \int_0^a f(t)\,dt$ permet de conclure.





Pour le cas d'une fonction périodique, une preuve concise est la suivante : si $F : x \mapsto \displaystyle\int_x^{x+T} f(t)\,dt = \int_0^{x+T} f(t)\,dt - \int_0^x f(t)\,dt$, alors F est dérivable[34], de dérivée $x \mapsto f(x+T) - f(x) = 0$.

[34] C'est le théorème fondamental de l'analyse.

Donc F est constante.Et en particulier

$$\int_a^{a+T} f(t)\,dt = F(a) = \int_0^T f(t)\,dt.$$

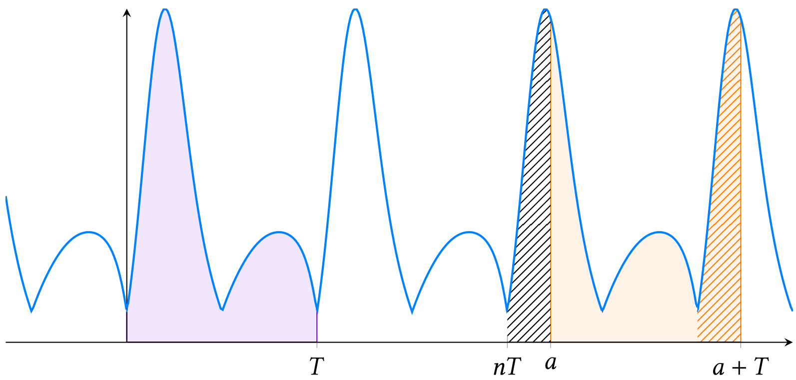
Pour autant, j'aimerais bien en donner une autre preuve, plus laborieuse, mais bien plus claire si l'on réfléchit graphiquement.
Soit donc $a \in \mathbf{R}$, et soit $n = \left\lfloor \frac{a}{T} \right\rfloor$, de sorte que $nT \leqslant a < (n+1)T$.
Alors par le changement de variable $x = t - T$, on a

$$\int_{(n+1)T}^{a+T} f(t)\,dt = \int_{nT}^{a} f(x+T)\,dx = \int_{nT}^{a} f(x)\,dx.$$

Et donc

$$\int_a^{a+T} f(t)\,dt = \int_a^{(n+1)T} f(t)\,dt + \int_{(n+1)T}^{a+T} f(t)\,dt = \int_a^{(n+1)T} f(t)\,dt + \int_{nT}^{a} f(x)\,dx = \int_{nT}^{(n+1)T} f(t)\,dt.$$

Le changement de variable $u = t - nT$ permet alors aisément de conclure.



□

25.4.3 Les formules de Taylor

Théorème 25.29 (Formule de Taylor avec reste intégral) : *Soit f une fonction de classe $\mathscr{C}^{n+1}$ sur un intervalle I et soit $a \in I$. Alors pour tout $x \in I$,*

$$\begin{aligned} f(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\,dt \\ &= f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\,dt. \end{aligned}$$

Démonstration. La preuve se fait par récurrence sur n.
À l'ordre 0, il s'agit de remarquer que pour f de classe $\mathscr{C}^1$, $\displaystyle\int_a^x f'(t)\,dt = f(x) - f(a)$ et donc $f(x) = f(a) + \displaystyle\int_a^x f'(t)\,dt$.
Supposons donc le résultat vrai pour une fonction de classe $\mathscr{C}^{n+1}$, et soit donc f une fonction de classe $\mathscr{C}^{n+2}$. Alors pour tout $x \in I$,

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\,dt.$$

Procédons alors à une intégration par parties sur cette dernière intégrale, en posant $u(t) = f^{(n+1)}(t)$ et $v(t) = -\dfrac{(x-t)^{n+1}}{(n+1)!}$, qui sont deux fonctions de classe $\mathscr{C}^1$ sur I, avec

$u'(t) = f^{(n+2)}(t)$ et $v'(t) = \dfrac{(x-t)^n}{n!}$. Alors

$$\int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\,dt = \left[-f^{(n+1)}(t)\frac{(x-t)^{n+1}}{(n+1)!}\right]_a^x + \int_a^x f^{(n+2)}(t)\frac{(x-t)^{n+1}}{(n+1)!}\,dt$$
$$= f^{(n+1)}(a)\frac{(x-a)^{n+1}}{(n+1)!} + \int_a^x f^{(n+2)}(t)\frac{(x-t)^{n+1}}{(n+1)!}\,dt.$$

Et donc

$$f(x) = \sum_{k=0}^{n+1}\frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x f^{(n+2)}(t)\frac{(x-t)^{n+1}}{(n+1)!}\,dt.$$

Donc la propriété est vraie au rang $n+1$, et donc par le principe de récurrence, est vraie pour tout $n \in \mathbf{N}$. □

Remarques. ► Vous aurez sûrement reconnu que la somme est exactement la partie régulière du $DL_n(a)$ de f.
► La grosse différence[35] entre cette formule et la formule de Taylor-Young est que la formule avec reste intégral est une formule exacte.
En effet, elle nous donne une formule pour la différence entre f et son développement limité, valable sur I tout entier, là où Taylor-Young nous donne juste une information sur le comportement de cette différence *au voisinage de a.*
► La formule de Taylor-Young peut se retrouver (dans le cas d'une fonction de classe $\mathscr{C}^{n+1}$) en prouvant que le reste intégral est négligeable devant $(x-a)^n$.

Le reste intégral[36] décrit l'erreur que l'on comment en approchant f par son $DL_n(a)$.

[35] Outre le fait que l'une nécessite f de classe $\mathscr{C}^n$ quand l'autre nécessite f de classe $\mathscr{C}^{n+1}$.

[36] Nom généralement donné à l'intégrale dans la formule ci-dessus.

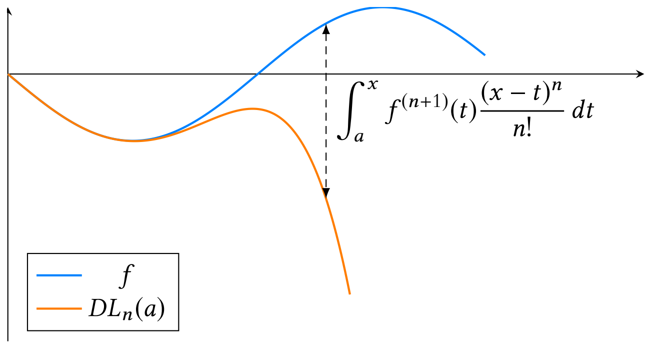


FIGURE 25.11 – Ici l'écart entre la fonction $-\sin$ et son $DL_5(0)$.

Corollaire 25.30 (Inégalité de Taylor-Lagrange) : *Soit f une fonction de classe $\mathscr{C}^{n+1}$ sur un intervalle I, et soient $(a,b) \in I^2$. Notons J le segment d'extrémités a et b. et soit M un majorant de $|f^{(n+1)}|$ sur J. Alors*

$$\left|f(b) - \sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(b-a)^k\right| \leqslant M\frac{|b-a|^{n+1}}{(n+1)!}.$$

Notations

Il serait tentant de dire que $J = [a, b]$, mais dans le cas où $b < a$, $J = [b, a]$.

Démonstration. Par la formule de Taylor avec reste intégral, on a

$$f(b) - \sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(b-a)^k = \int_a^b f^{(n+1)}(t)\frac{(b-t)^n}{n!}\,dt.$$

Il s'agit donc de majorer la valeur absolue de cette intégrale. Si $a \leqslant b$, alors

$$\left|\int_a^b f^{(n+1)}(t)\frac{(b-t)^n}{n!}\,dt\right| \leqslant \int_a^b \left|f^{(n+1)}(t)\frac{(b-t)^n}{n!}\right|\,dt$$
$$\leqslant \int_a^b M\frac{(b-t)^n}{n!}\,dt$$

Inégalité triangulaire.

Pour $t \in [a, b]$, $b - t \geqslant 0$, donc on peut se passer de la valeur absolue.

$$\begin{aligned} &\leqslant M \int_a^b \frac{(b-t)^n}{n!}\,dt \\ &\leqslant M \left[-\frac{(b-t)^{n+1}}{(n+1)!} \right]_a^b \\ &\leqslant M \frac{(b-a)^{n+1}}{(n+1)!}. \end{aligned}$$

Dans le cas où $b < a$, le principe est le même, mais il y a quelques précautions à prendre car l'inégalité triangulaire et la croissance de l'intégrale nécessitent des bornes dans le bon sens.

$$\begin{aligned} \left| \int_a^b f^{(n+1)}(t) \frac{(b-t)^n}{n!}\,dt \right| &= \left| \int_b^a f^{(n+1)}(t) \frac{(b-t)^n}{n!}\,dt \right| \\ &\leqslant \int_b^a \left| f^{(n+1)}(t) \frac{(b-t)^n}{n!} \right| dt \\ &\leqslant M \int_b^a \frac{(t-b)^n}{n!}\,dt \\ &\leqslant M \left[\frac{(t-b)^{n+1}}{(n+1)!} \right]_b^a \\ &\leqslant M \frac{(a-b)^{n+1}}{(n+1)!} \leqslant M \frac{|b-a|^{n+1}}{(n+1)!}. \end{aligned}$$

Danger ! Cette fois pour $t \in [b, a]$, $b - t \leqslant 0$, et donc il y a des précautions à prendre pour enlever la valeur absolue.

□

Remarque. Bien entendu, si $f^{(n+1)}$ est bornée sur I tout entier, alors on peut prendre pour M un majorant de $|f^{(n+1)}|$, mais la formule reste valable y compris si un tel majorant sur I n'existe pas.

Exemple 25.31

Considérons la fonction $f : t \mapsto e^t$. Alors, pour tout $x \in \mathbf{R}$, sur le segment d'extrémités 0 et x, $|f^{(n+1)}| = f^{(n+1)}$ est majorée par $e^{|x|}$.
Et donc par l'inégalité de Taylor-Lagrange, puisque tous les $f^{(k)}(0)$ valent 1,

$$\left| e^x - \sum_{k=0}^n \frac{x^k}{k!} \right| \leqslant e^{|x|} \frac{|x|^{n+1}}{(n+1)!}.$$

Mais par croissances comparées, on a $\dfrac{|x|^n}{n!} \underset{n\to+\infty}{\longrightarrow} 0$.
On en déduit que

$$\lim_{n\to+\infty} \left| e^x - \sum_{k=0}^n \frac{x^k}{k!} \right| = 0 \Leftrightarrow \lim_{n\to+\infty} \sum_{k=0}^n \frac{x^k}{k!} = e^x.$$

25.5 Extension aux fonctions à valeurs complexes

La plupart des résultats qui précèdent s'étendent sans difficulté aux fonctions à valeurs complexes.

Définition 25.32 – Soit $f : [a, b] \to \mathbf{C}$ une fonction à valeurs complexes. On dit que f est continue par morceaux sur $[a, b]$ si les deux fonctions[37] $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ sont continues par morceaux.
On note $\mathscr{C}_m([a, b], \mathbf{C})$ l'ensemble des fonctions à valeurs complexes qui sont continues par morceaux sur $[a, b]$.

[37] À valeurs réelles.

Définition 25.33 – Pour $f \in \mathscr{C}_m([a,b],\mathbf{C})$, on pose

$$\int_a^b f(t)\,dt = \int_a^b \mathrm{Re}(f)(t)\,dt + i\int_a^b \mathrm{Im}(f)(t)\,dt.$$

Notons en particulier que $\mathrm{Re}\left(\int_a^b f(t)\,dt\right) = \int_a^b \mathrm{Re}(f)(t)\,dt$ et $\mathrm{Im}\left(\int_a^b f(t)\,dt\right) = \int_a^b \mathrm{Im}(f)(t)\,dt$.

Notons qu'on perd ici toute interprétation de l'intégrale comme une aire !
En séparant partie réelle et partie imaginaire, on prouve aisément que la linéarité de l'intégrale et la relation de Chasles restent valables pour les fonctions à valeurs complexes.
Bien entendu, la positivité et la croissance n'ont alors plus aucun sens, en l'absence d'une relation d'ordre naturelle sur $\mathbf{C}$.
En revanche, l'inégalité triangulaire reste valable.

Proposition 25.34 : *Pour* $f \in \mathscr{C}_m([a,b],\mathbf{C})$, $\left|\int_a^b f(t)\,dt\right| \leqslant \int_a^b |f(t)|\,dt$.

Démonstration. Notons θ un argument de $\int_a^b f(t)\,dt$, de sorte que

$$\int_a^b f(t)\,dt = e^{i\theta}\left|\int_a^b f(t)\,dt\right|.$$

Posons alors $F(x) = \int_a^x f(t)\,dt$ et $G(x) = \mathrm{Re}\left(e^{-i\theta}F(x)\right)$.
Alors $G(b) = \mathrm{Re}\left(e^{-i\theta}\int_a^b f(t)\,dt\right) = \left|\int_a^b f(t)\,dt\right|$.
Mais par ailleurs, $G(b) = \mathrm{Re}\left(\int_a^b e^{-i\theta}f(t)\,dt\right) = \int_a^b \mathrm{Re}(e^{-i\theta}f(t))\,dt \leqslant \int_a^b |\mathrm{Re}(e^{-i\theta}f(t))|\,dt$.
Or, $|\mathrm{Re}(e^{-i\theta}f(t))| \leqslant |e^{-i\theta}f(t)| \leqslant |f(t)|$.
Et donc

$$G(b) = \left|\int_a^b f(t)\,dt\right| \leqslant \int_a^b |f(t)|\,dt.$$

□

Rappel
Pour tout complexe z,
$$|Re(z)| \leqslant |z|.$$

Notons que le théorème fondamental de l'analyse reste lui aussi valable pour une fonction continue à valeurs complexes.
Enfin, les formules de Taylor (reste intégral et Taylor-Lagrange) restent valables, avec des preuves inchangées[38].

[38] Si ce n'est qu'il faut remplacer les valeurs absolues par des modules, ce qui est totalement indolore.

25.6 Calcul approché d'intégrales

25.6.1 Sommes de Riemann

Définition 25.35 – Soient $a < b$ deux réels. On appelle **subdivision pointée** de l'intervalle $[a,b]$ la donnée d'un couple $s = (\sigma, (t_i)_{1\leqslant i\leqslant n})$ où :

- $\sigma = (x_i)_{0\leqslant i\leqslant n}$ est une subdivision de $[a,b]$: $a = x_0 < x_1 < \cdots < x_n = b$
- $\forall i \in [\![1,n]\!]$, $t_i \in [x_{i-1}, x_i]$.

On appelle alors pas de s et on note $\delta(s)$ le pas[39] de la subdivision σ.

[39] Qui, rappelons-le, est la distance maximale entre deux points consécutifs de la subdivision.

Autrement dit, se donner une subdivision pointée, c'est se donner une subdivision et choisir un point dans chacun des intervalles de cette subdivision.

Définition 25.36 – Si $f \in \mathscr{C}_m([a,b],\mathbf{R})$ et si $s = (\sigma, (t_i)_{1\leqslant i\leqslant n})$ est une subdivision pointée de $[a,b]$, alors on note $R(f,s) = \sum_{i=1}^{n}(x_i - x_{i-1})f(t_i)$.
Le réel $R(f,s)$ est appelé **somme de Riemann** associée à f et s.

Proposition 25.37 : *Soit f une fonction continue sur $[a,b]$, à valeurs réelles. Pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour toute subdivision pointée s de $[a,b]$, si $\delta(s) \leqslant \eta$, alors*

$$\left| \int_a^b f(t)\,dt - R(f,s) \right| \leqslant \varepsilon.$$

Démonstration. Puisque f est continue sur $[a,b]$, par le théorème de Heine, elle y est uniformément continue.
Et donc il existe $\eta > 0$ tel que $\forall (x,y) \in [a,b]^2, |x-y| \leqslant \eta \Rightarrow |f(x) - f(y)| \leqslant \dfrac{\varepsilon}{b-a}$.
Soit alors $s = (\sigma, (t_i)_{1\leqslant i\leqslant n})$ une subdivision pointée de $[a,b]$, avec $\delta(s) \leqslant \eta$ et où la subdivision σ est donnée par $a = x_0 < x_1 < \cdots < x_n = b$. Alors

$$\begin{aligned}
\left| \int_a^b f(t)\,dt - \sum_{i=1}^{n}(x_i - x_{i-1})f(t_i) \right| &= \left| \int_a^b f(t)\,dt - \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(t_i)\,dt \right| \\
&= \left| \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(t)\,dt - \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(t_i)\,dt \right| && \text{Relation de Chasles.} \\
&= \left| \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} (f(t) - f(t_i))\,dt \right| \\
&\leqslant \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |f(t) - f(t_i))|\,dt && \text{Inégalité(s) triangulaire(s).} \\
&\leqslant \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} \frac{\varepsilon}{b-a}\,dt \\
&\leqslant \frac{\varepsilon}{b-a} \int_a^b dt \leqslant \varepsilon.
\end{aligned}$$

□

Remarques. Le résultat reste valable pour des fonctions qui ne sont que continues par morceaux, mais la preuve en est plus désagréable, car les points de discontinuité de f ne sont pas forcément des points de la subdivision σ.

► Ce que dit ce résultat, c'est que quand le pas tend vers 0, alors $R(f,s)$ tend vers $\int_a^b f(t)\,dt$.

La preuve ci-dessus et même la définition de subdivision adaptée ne sont pas du tout importantes.
Ce qui l'est en revanche est ce qui suit, où l'on considère des subdivisions régulières de $[a,b]$, c'est-à-dire avec $x_k = a + k\dfrac{b-a}{n}$.
Dans ce cas, parmi toutes les manières de considérer des subdivisions pointées, trois sont sans doute un peu plus «naturelles» que les autres :

- prendre $t_i = x_{i-1}$, la borne de gauche de $[x_{i-1}, x_i]$. On parle alors de la *méthode des rectangles à gauche.*
- prendre $t_i = x_i$, la borne de droite de $[x_{i-1}, x_i]$. On parle alors de la *méthode des rectangles à droite.*
- prendre $t_i = \dfrac{x_{i-1} + x_i}{2}$ le milieu de $[x_{i-1}, x_i]$.

Graphiquement, $R(f,s)$ est l'approximation de l'aire sous la courbe de f par des rectangles de hauteur $f(t_i)$.

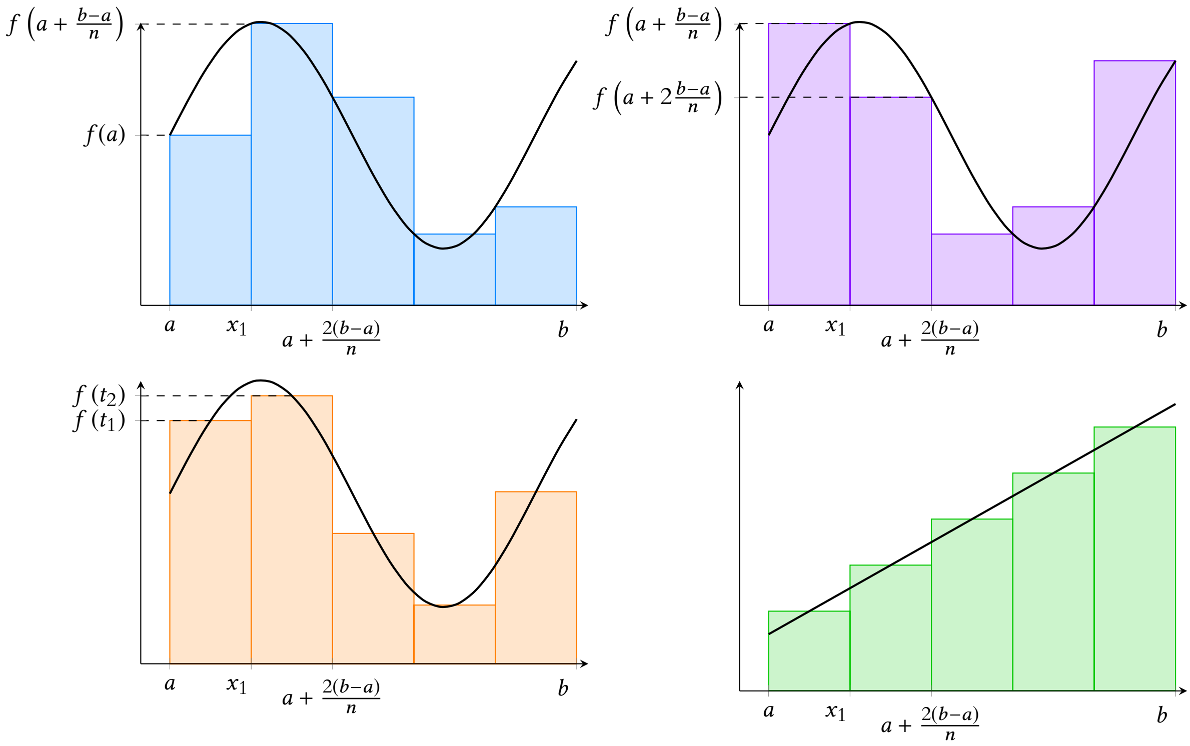


FIGURE 25.12 – Rectangles à gauche, rectangles à droite et point milieu.
Notons que la méthode du point milieu est exacte pour les fonctions affines.

Théorème 25.38 (Convergence des sommes de Riemann) : *Soit f une fonction continue sur $[a, b]$. Alors*

$$\int_a^b f(t)\,dt = \lim_{n\to+\infty} \frac{b-a}{n}\sum_{k=0}^{n-1} f\left(a+k\frac{b-a}{n}\right)$$
$$= \lim_{n\to+\infty} \frac{b-a}{n}\sum_{k=1}^{n} f\left(a+k\frac{b-a}{n}\right).$$

Il est vraiment important de réussir à bien comprendre en quoi les deux sommes ci-dessus correspondent aux deux premiers dessins de la figure 25.12.
En particulier, il est hors de question de les apprendre par cœur, alors qu'il suffit de les comprendre et de savoir les retrouver avec un dessin.
La première limite correspond aux rectangles à gauche, la seconde aux rectangles à droite.
Dans les deux cas, le $\frac{b-a}{n}$ est la longueur de chacun des intervalles de la subdivision, c'est-à-dire la longueur de nos petits rectangles.

Graphiquement, cela signifie que plus on considère de points dans notre subdivision régulière, mieux on approche l'intégrale de f avec les sommes de Riemann.

Démonstration. Soit $\varepsilon > 0$. Alors, il existe $\eta > 0$ tel que pour toute subdivision pointée s de pas inférieur à η, $\left|\int_a^b f(t)\,dt - R(f,s)\right| \leqslant \varepsilon$.
Soit donc $N \in \mathbf{N}^*$ tel que pour $n \geqslant N$, $\dfrac{b-a}{n} \leqslant \eta$.
Alors en particulier, pour $n \geqslant N$ les subdivisions pointées correspondant aux rectangles à gauche et à droite ont un pas inférieur à η, et donc les sommes de Riemann associées sont à distance moins de ε de $\displaystyle\int_a^b f(t)\,dt$.

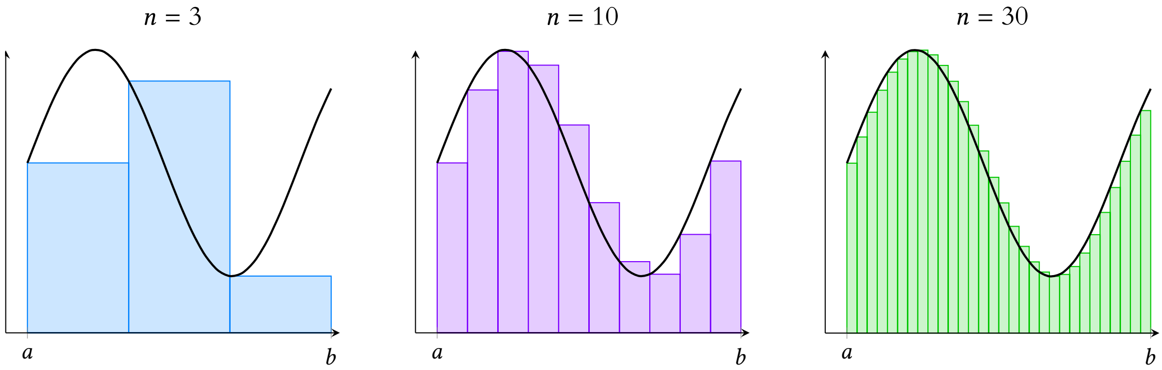


Figure 25.13 – Illustration de la convergence des sommes de Riemann (ici avec des rectangle à gauche) vers l'intégrale.

Or, $\dfrac{b-a}{n}\displaystyle\sum_{k=0}^{n-1} f\left(a+k\dfrac{b-a}{n}\right)$ est la somme de Riemann pour les rectangles à gauche. Donc pour $n \geqslant N$,

$$\left|\int_a^b f(t)\,dt - \frac{b-a}{n}\sum_{k=0}^{n-1} f\left(a+k\frac{b-a}{n}\right)\right| \leqslant \varepsilon.$$

C'est exactement la définition[40] de $\displaystyle\lim_{n\to+\infty} \frac{b-a}{n}\sum_{k=0}^{n-1} f\left(a+k\frac{b-a}{n}\right) = \int_a^b f(t)\,dt$.

On raisonne de même pour les rectangles à droite. □

[40] Pour tout $\varepsilon > 0$ il existe N tel que...

Remarque. Bien entendu, le même résultat vaut pour la méthode du point milieu, et je vous laisse trouver la formule correspondante, mais je ne la mentionne pas car elle sert bien moins souvent que les deux précédentes.

Méthode

Le plus difficile dans les sommes de Riemann sera de les reconnaître. Et même une fois qu'on a pensé à une somme de Riemann, il faut encore trouver le a, le b et la fonction f.
Généralement, le plus simple est de prendre $[a, b] = [0, 1]$, de faire apparaître un $\frac{1}{n}$ devant la somme, et alors il n'y a plus de choix pour f.

Exemple 25.39

Soit $(u_n)_{n\geqslant 1}$ la suite définie par $u_n = \displaystyle\sum_{k=1}^{n} \frac{n}{n^2+k^2}$. Alors

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

Posons alors $a = 0$, $b = 1$ et $f : x \mapsto \dfrac{1}{1+x^2}$, de sorte que

$$u_n = \frac{b-a}{n}\sum_{k=1}^{n} f\left(a+k\frac{b-a}{n}\right) \underset{n\to+\infty}{\longrightarrow} \int_a^b f(t)\,dt = \int_0^1 \frac{1}{1+t^2}\,dt = \frac{\pi}{4}.$$

25.6.2 Majoration de l'erreur dans la méthode des rectangles (IPT)

Proposition 25.40 : *Si f est lipschitzienne sur $[a, b]$, alors*

$$\left|\int_a^b f(t)\,dt - \frac{b-a}{n}\sum_{k=1}^{n} f\left(a+k\frac{b-a}{n}\right)\right| \underset{n\to+\infty}{=} O\left(\frac{1}{n}\right).$$

Démonstration. Soit donc $K > 0$ tel que f soit K-lipschitzienne.
Posons alors $x_k = a + k\dfrac{b-a}{n}$ pour $0 \leqslant k \leqslant n$.

On a alors

$$\begin{aligned}
\left|\int_a^b f(t)\,dt - \frac{b-a}{n}\sum_{k=1}^n f\left(a+k\frac{b-a}{n}\right)\right| &\leqslant \left|\sum_{k=1}^n \int_{x_{k-1}}^{x_k} f(t)\,dt - \sum_{k=1}^n \int_{x_{k-1}}^{x_k} f(x_k)\,dt\right| \\
&\leqslant \left|\sum_{k=1}^n \int_{x_{k-1}}^{x_k} (f(t)-f(x_k))\,dt\right| \\
&\leqslant \sum_{k=1}^n \int_{x_{k-1}}^{x_k} |f(t)-f(x_k)|\,dt \\
&\leqslant \sum_{k=1}^n \int_{x_{k-1}}^{x_k} K|t-x_k|\,dt \\
&\leqslant K\sum_{k=1}^n \int_{x_{k-1}}^{x_k} (x_k-t)\,dt \\
&\leqslant K\sum_{k=1}^n \frac{(x_k-x_{k-1})^2}{2} \\
&\leqslant K\sum_{k=1}^n \frac{(b-a)^2}{2n^2} \leqslant K\frac{(b-a)^2}{2n} \\
&\underset{n\to+\infty}{=} O\left(\frac{1}{n}\right).
\end{aligned}$$

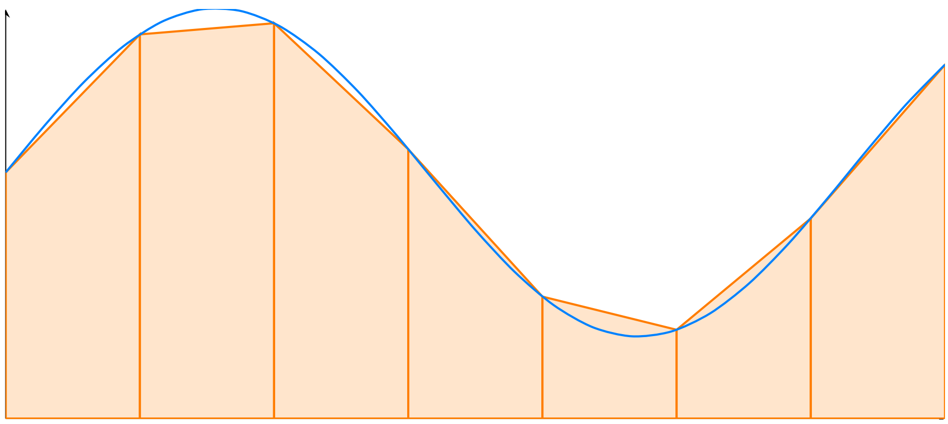
□

Corollaire 25.41 – *Si f est de classe $\mathscr{C}^1$ sur $[a,b]$, alors l'erreur de la méthode des rectangles (à gauche ou à droite) est un $O\left(\frac{1}{n}\right)$.*

Démonstration. Il suffit de remarquer que si f est $\mathscr{C}^1$, sur $[a,b]$, alors par le théorème des bornes atteintes, sa dérivée est bornée, et donc f est lipschitzienne. □

25.6.3 Méthode des trapèzes (IPT)

Les sommes de Riemann (c'est-à-dire la méthode des rectangles) consistent à approximer la fonction f par une fonction constante par morceaux, constante sur chacun des $[x_k, x_{k+1}]$. La méthode des trapèzes consiste en l'approximation de f par des fonctions affines par morceaux. Plus précisément, sur chacun des $[x_k, x_{k+1}]$, on approche f par l'unique fonction affine dont le graphe passe par $(x_k, f(x_k))$ et $(x_{k+1}, f(x_{k+1}))$.



Il n'est alors pas très difficile[41]

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

[41] Pour peu qu'on se souvienne que l'aire d'un trapèze est donnée par $\frac{(\text{pte base} + \text{gde base})\times \text{hauteur}}{2}$

Proposition 25.42 : *Si f est de classe $\mathscr{C}^2$ sur $[a,b]$, alors l'erreur commise lorsqu'on approche $\int_a^b f(t)\,dt$ par la méthode des trapèzes est un $O\left(\frac{1}{n^2}\right)$. Autrement dit*

$$\left|\int_a^b f(t)\,dt - \frac{b-a}{n}\left(\frac{f(a)+f(b)}{2} + \sum_{k=1}^{n-1} f\left(a+k\frac{b-a}{n}\right)\right)\right| \underset{n\to+\infty}{=} O\left(\frac{1}{n^2}\right).$$

Démonstration. Majorons l'erreur commise sur $[x_k, x_{k+1}]$, il suffira ensuite de sommer les erreurs.
La fonction affine qui passe par $(x_k, f(x_k))$ et $(x_{k+1}, f(x_{k+1}))$ est

$$f_k : x \mapsto \frac{f(x_{k+1})-f(x_k)}{x_{k+1}-x_k}(x-x_k)+f(x_k).$$

Alors, comme annoncé précédemment, $\displaystyle\int_{x_k}^{x_{k+1}} f_k(t)\,dt = \frac{f(x_k)+f(x_{k+1})}{2}\frac{b-a}{n}$.
Fixons $t \in]x_k, x_{k+1}[$, et considérons la fonction $g_k : x \mapsto f(x)-f_k(x)-A_{k,t}\dfrac{(x-x_k)(x-x_{k+1})}{2}$, où $A_{k,t}$ est choisi de telle sorte que $g_k(t)=0$. Puisque $g_k(x_k)=g_k(t)=g_k(x_{k+1})=0$, par le théorème de Rolle, g' s'annule deux fois dans l'intervalle $]x_k, x_{k+1}[$, et donc g'' s'annule une fois en un point $c_t \in]x_k, x_{k+1}[$.
Mais $g''(x) = f''(x) - A_{k,t}$, et donc $g''(c_t) = 0 \Leftrightarrow A_{k,t} = f''(c_t)$.
Puisque f est de classe $\mathscr{C}^2$, f'' est bornée sur $[a,b]$. Soit donc M un majorant de $|f''|$ sur $[a,b]$.
On a donc, pour tout $t \in]x_k, x_{k+1}[$,

$$|f(t)-f_k(t)| = |A_{k,t}|\cdot\left|\frac{(t-x_k)(t-x_{k+1})}{2}\right| \leqslant M_2\frac{(t-x_k)(x_{k+1}-t)}{2}.$$

Et donc

$$\begin{aligned}
\int_{x_k}^{x_{k+1}} |f(t)-f_k(t)|\,dt &= \frac{M_2}{2}\int_{x_k}^{x_{k+1}} (t-x_k)(x_{k+1}-t)\,dt \\
&\leqslant \frac{M_2}{2}\underbrace{\left[(t-x_k)\frac{(t-x_{k+1})^2}{2}\right]_{x_k}^{x_{k+1}}}_{=0} + \frac{M_2}{2}\int_{x_k}^{x_{k+1}}\frac{(t-x_k)^2}{2}\,dt \\
&\leqslant \frac{M_2}{2}\left[\frac{(t-x_k)^3}{6}\right]_{x_k}^{x_{k+1}} \\
&\leqslant \frac{M_2}{12}(x_{k+1}-x_k)^3 \leqslant \frac{M_2}{12}\frac{(b-a)^3}{n^3}
\end{aligned}$$

Intégration par parties.

Donc en sommant les erreurs[42] sur chacun des intervalles $]x_k, x_{k+1}[$, on en vient à une erreur dans la méthode de trapèzes majorée par

$$n\frac{M_2}{12}\frac{(b-a)^3}{n^3} \leqslant \frac{M_2}{12}\frac{(b-a)^3}{n^2}.$$

Comme annoncé, on a bien un $O\left(\frac{1}{n^2}\right)$. □

[42] Une inégalité triangulaire s'est cachée par là, saurez-vous la retrouver ?

Remarque. Vous l'aurez sans doute remarqué, mais la somme de la méthode des trapèzes ressemble beaucoup à celles de la méthode des rectangles, et n'en diffère que pour les termes extrémaux.
Il est donc fort probable que le majorant de l'erreur obtenu dans la méthode des rectangles soit loin d'être optimal dans la plupart des cas, et que l'erreur soit alors elle aussi un $O\left(\frac{1}{n^2}\right)$.