Intégration I

Licence de mathématiques, 4$^{\text{e}}$ semestre
Université Aix-Marseille 1

J-Y. Briend

Janvier 2006

1 Introduction

Les deux théorèmes dont l'énoncé suivent sont, peu ou prou, les seuls résultats généraux que vous connaissiez sur l'intégration des fonctions d'une variable réelle.

Théorème 1.1. — *Si f est une fonction dérivable sur l'intervalle compact $[a, b]$,* et si f' est continue, *alors*

$$\int_a^b f'(x)\,\mathrm{d}x = f(b) - f(a).$$

De plus, toute fonction continue g sur $[a, b]$ y admet une primitive G, unique à l'addition d'une constante près, donnée par la formule

$$\mathrm{G}(t) = \int_a^t g(x)\,\mathrm{d}x.$$

Ce premier énoncé, connu sous le nom de *théorème fondamental du calcul différentiel et intégral,* permet le calcul explicite d'intégrales *via* le calcul de primitives. Le second, qui suit, donne une interprétation de l'intégrale en terme d'aire, notion géométrique intuitive :

Théorème 1.2. — *Soit f une fonction continue positive sur l'intervalle compact $[a, b]$. Alors l'aire A de la région plane contenue entre l'axe des abscisses et le graphe de f est donnée par*

$$\mathrm{A} = \int_a^b f(x)\,\mathrm{d}x.$$

Cet énoncé s'étend aux fonctions de signe variable en définissant une aire « algébrique ».

Ces deux résultats, qui constituent par ailleurs les seules définitions sérieuses données au lycée pour l'intégrale, ont un grave défaut : ils ne permettent d'intégrer que des fonctions continues. Or les théorèmes d'appoximation des fonctions continues font tous appel à la convergence uniforme, forme de convergence bien trop forte dans de nombreuses applications pratiques du calcul intégral. Il en est ainsi, par exemple, en traitement du signal : on approxime un signal très bruité par des échantillons obtenus par superposition de signaux sinusoïdaux, mais ces échantillons ne convergent que rarement de manière uniforme vers le signal original. La convergence a lieu en un sens plus faible. On a donc besoin, pour approcher des invariants du signal définis par une intégrale (l'énergie par exemple), d'avoir des théorèmes de permutation de limite et d'intégrale qui fonctionnent dans le cadre de ce type de convergence.

La théorie de l'intégrale qui répond de manière satisfaisante aux deux théorèmes énoncés plus haut est celle initiée par Bernhard Riemann au milieu du dix-neuvième siècle. Elle n'admet cependant pas de bons théorèmes de convergence, rendant son application concrète difficile : les meilleurs théorèmes la concernant font toujours l'hypothèse que la fonction limite est intégrable, ce que l'on ne sait en général pas démontrer. Par ailleurs, l'intégrale de Riemann ne permet pas de traîter de manière satisfaisante l'intégration de fonctions non bornées, ou définies sur des intervalles non compacts.

C'est pour combler ces lacunes qu'Henri Lebesgue inventa, au début du vingtième siècle, l'intégrale qui porte maintenant son nom. Le théorème de convergence dominée de Lebesgue, point d'orgue de cette théorie, fait qu'elle est actuellement l'intégrale universellement utilisée, aussi bien en mathématiques qu'en physique. Elle a cependant quelques défauts. L'un d'eux est que sa construction, même pour les fonctions réelles d'une variable réelle, est très difficile, et cela rend déraisonnable sa présentation dans un cours de Licence. Par ailleurs, l'intégrale de Lebesgue ne possède pas d'intégrales semi-convergentes, comme par exemple

$$\int_0^{+\infty} \frac{\sin x}{x} \,\mathrm{d}x.$$

Du coup, les rapports entre la théorie de l'intégrale et le calcul différentiel (illustrés par le théorème 1.1) sont plutôt ambigus, voire carrément mauvais lorsque l'on essaye d'obtenir un résultat général reliant intégrales et primitives.

Dans ce cours, nous allons présenter une théorie de l'intégrale dûe à Jaroslav Kurzweil et Ralph Henstock, à la fin des années 1950. Elle est basée sur une modification *a priori* anodine de l'intégrale de Riemann.Elle permet, en un coup et avec des outils simples, de palier aux défauts des intégrales de Riemann et de Lebesgue. Elle est basée sur une idée très simple : on calcule des aires à l'aide de sommes de Riemann, sommes basées sur des subdivisions de l'intervalle *dont on adapte le pas à la fonction à étudier.* Cette démarche, qui semble naïve, permet en fait de démontrer une version très générale du théorème fondamental, d'avoir les théorèmes de convergence de l'intégrale de Lebesgue, et enfin de traîter les intégrales semi convergentes et impropres, inaccessibles aussi bien à Riemann qu'à Lebesgue.

Table des matières

2 Des aires aux primitives, et *vice versa*

Nous savons tous calculer l'aire de figures simples, triangles ou rectangles. Pour des polygones plus compliqués, nous prenons des ciseaux et essayons de nous ramener à un nombre fini de petits bouts simples (procédé utile pour les cadastres, par exemples) ; il s'agit finalement d'une méthode d'approximation, et nous allons rafiner cette idée pour les régions plus compliquées traîtées dans le théorème 1.2.

Soit f une fonction, disons, pour simplifier, positive et continue sur l'intervalle $[a, b]$. Le théorème 1.2 prétend calculer l'aire A du domaine R contenu sous le graphe de f.

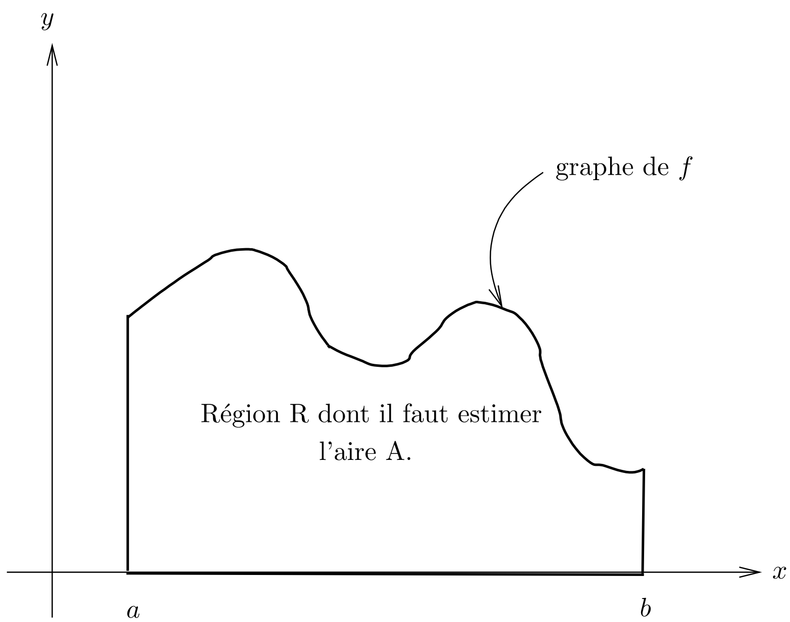

FIG. 1: L'intégrale comme aire d'une région plane.

L'idée la plus simple, pour calculer cette aire de manière approchée, est de découper la région R en rectangles verticaux, choisis, dans un premier temps *au pif*. Soyons simplistes et un peu fainéants : contentons nous de trois rectangles, basés respectivement sur les intervalles $\mathrm{I}_1 = [a = x_0, x_1], \mathrm{I}_2 = [x_1, x_2], \mathrm{I}_3 = [x_2, x_3 = b]$. Choisissons un point dans chacun des ces intervalles : $t_1 \in \mathrm{I}_1, t_2 \in \mathrm{I}_2$ et $t_3 \in \mathrm{I}_3$, et dessinons des rectangles des base I_j, de hauteur $f(t_j)$ (*voir* la figure 2).

Nous obtenons ainsi trois rectangles d'aires respectives

$$\mathrm{A}_1 = (x_1 - x_0)f(t_1),\ \mathrm{A}_2 = (x_2 - x_1)f(t_2) \text{ et } \mathrm{A}_3 = (x_3 - x_2)f(t_2),$$

soit une approximation de l'aire de la région R par la somme

$$\mathrm{A} \approx \mathrm{A}_1 + \mathrm{A}_2 + \mathrm{A}_3 = \sum_{k=1}^{3}(x_k - x_{k-1})f(t_k) = \sum_{k=1}^{3} |\mathrm{I}_k| f(t_k),$$

où l'on a noté $|\mathrm{I}_k|$ la *longueur* de l'intervalle I_k. Nous venons de définir deux notions qui seront centrales dans toute la suite de ce cours, à savoir celle de *subdivision*

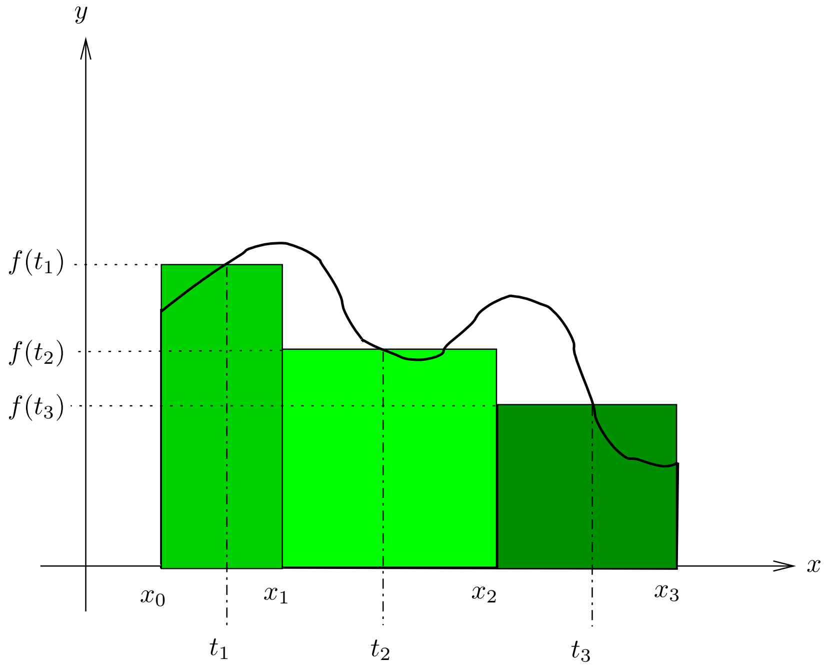

FIG. 2: Approximation de l'aire de A par des rectangles.

pointée (la famille des intervalles $\mathrm{I}_1, \mathrm{I}_2, \mathrm{I}_3$ et le choix d'un point t_j dans chacun des I_j), et celle de *somme de Riemann* (la somme $\mathrm{A}_1 + \mathrm{A}_2 + \mathrm{A}_3$) associée à cette subdivision.

Définition 2.1. — *Soit* $\mathrm{I} = [a, b]$ *un intervalle compact. Une subdivision pointée* $\mathcal{P}$ *de* I *est la donnée d'une subdivision* $\mathrm{I}_1 = [x_0, x_1], \mathrm{I}_2 = [x_1, x_2], \ldots, \mathrm{I}_n = [x_{n-1}, x_n]$ *de* I *(avec* $a = x_0 \leq x_1 \leq \cdots \leq x_n = b$*) et d'un « pointage »de cette partition,* i.e. *des points* $t_1 \in \mathrm{I}_1, \ldots, t_n \in \mathrm{I}_n$. *On la notera* $\mathcal{P} = \{(\mathrm{I}_1, t_1), \ldots, (\mathrm{I}_n, t_n)\}$, *et on appellera les* t_i *des points de marquage de* $\mathcal{P}$.

Notons au passage que les subdivisions que nous considérerons seront toujours supposées ordonnées par l'ordre usuel de $\mathbf{R}$. Les subdivisions pointées de l'intervalle $[a, b]$ permettent naturellement, comme nous venons de le voir, d'approcher l'aire du domaine A sous le graphe d'une fonction f définie sur $[a, b]$.

Définition 2.2. — *Soient* f *une fonction numérique* quelconque *définie sur l'intervalle* $[a, b]$, *et* $\mathcal{P} = \{(\mathrm{I}_1, t_1), \ldots, (\mathrm{I}_n, t_n)\}$ *une subdivision pointée de* $[a, b]$. *On appelle somme de Riemann de* f *associée à* $\mathcal{P}$ *la quantité :*

$$\mathrm{S}(f, \mathcal{P}) = \sum_{k=1}^{n} |\mathrm{I}_k| f(t_k).$$

Nous disposons maintenant d'un outil simple d'approximation des intégrales au sens du théorème 1.2, et il serait agréable que ces appoximations fonctionnent aussi pour les intégrales données par le théorème fondamental 1.1. Voyons de plus près ce qu'il en est.

Considérons donc une fonction f, dérivable sur $[a, b]$ (en a et b, il s'agit de dérivées à droite et à gauche respectivement). Donnons-nous également une subdivision pointée $\mathcal{P} = \{(\mathrm{I}_1, t_1), \dots, (\mathrm{I}_n, t_n)\}$ de $[a, b]$, et formons la somme de Riemann de f' associée à $\mathcal{P}$:

$$\mathrm{S}(f', \mathcal{P}) = \sum_{k=1}^{n} (x_{k-1} - x_k) f'(t_k),$$

où, bien entendu, on a noté $\mathrm{I}_k = [x_{k-1}, x_k]$. La quantité $(x_{k-1} - x_k) f'(t_k)$ n'est pas sans rappeler la formule des accroissements finis (conséquence du théorème de Rolle). Celle-ci stipule que, sous les hypothèses faites ici, il existe un point $c_k \in [x_{k-1}, x_k]$ tel que

$$f(x_k) - f(x_{k-1}) = (x_k - x_{k-1}) f'(c_k).$$

En termes plus géométriques, la tangente au graphe de f au point $(c_k, f(c_k))$ est parallèle à la *corde* entre les points $(a, f(a))$ et $(b, f(b))$ (*voir* la figure 3).

Si, par le plus grand des hasards, et grâce à une chance incroyable, nous avions $t_k = c_k$ pour tous les k, la somme de Riemann de f' associée à $\mathcal{P}$ deviendrait une *somme télescopique* se simplifiant par miracle :

$$\begin{aligned} \mathrm{S}(f', \mathcal{P}) &= (x_0 - x_1) f'(c_1) + (x_1 - x_2) f'(c_2) + \cdots + (x_n - x_{n-1}) f'(c_n) \\ &= (f(x_1) - f(x_0)) + (f(x_2) - f(x_1)) + \cdots + (f(x_n) - f(x_{n-1}), \end{aligned}$$

soit :

$$\mathrm{S}(f', \mathcal{P}) = f(x_n) - f(x_0) = f(b) - f(a).$$

Nous aurions ainsi trouvé *exactement* la formule du théorème 1.1 ! Bien-sûr, comme nous parlons d'approximation, il est insensé de croire que la situation $c_k = t_k$ puisse réellement avoir lieu, sauf dans le cas où f est affine.

Dans tous les autres cas, nous ne pouvons espérer qu'une seule chose : qu'en prenant les intervalles I_k suffisament petits, les points c_k et t_k soient tellement proches que l'on ait alors $f'(c_k) \approx f'(t_k)$. Cela donne une première idée pour définir l'intégrale de f' : ce serait la limite, si elle existe, des sommes de Riemann $\mathrm{S}(f', \mathcal{P})$, quand on prend des subdivisions de pas de plus en plus petit. On pourrait par exemple prendre pour sommets des intervalles les nombres du type $a + j(b-a)2^{-l}$, $j = 0, \dots, 2^l$, c'est à dire découper $[a, b]$ en 2^l intervalles de longeur égale, et ensuite faire tendre l vers l'infini. Mais ce n'est, *a priori,* pas suffisant : cela est illustré par la figure 3. Dans le premier cas, le « bon », prendre t_k différent de c_k donnera une bonne approximation de

$$(x_k - x_{k-1}) f'(c_k) = f(x_k) - f(x_{k-1})$$

par

$$(x_k - x_{k-1}) f'(t_k),$$

car f' (*i.e.* la pente du graphe) varie peu dans l'intervalle. Mais, dans le second, un choix malencontreux de t_k, par exemple en un point où la dérivée de f est grande, donnera une approximation déplorable. Vous verrez en travaux dirigés qu'il existe des fonctions f dérivables, dont la dérivées varie au voisinage de certains points si fortement que l'usage d'intervalles de longueur constante ne permet pas de faire converger les sommes de Riemann vers une limite. Il faut assouplir le processus de

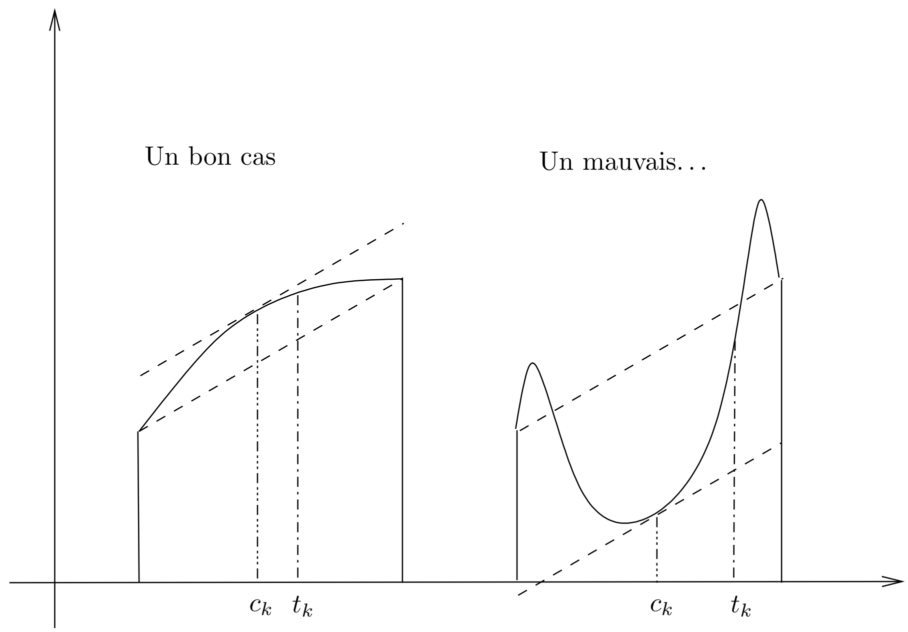

FIG. 3: Deux cas de la formule des accroissements finis.

passage à la limite et s'autoriser à *adapter* la longueur des intervalles des subdivisions en fonction des variations locales de f'. Ces considérations nous mènent tout naturellement à la définition centrale de tout ce cours, celle de *jauge* et de subdivision δ-fine :

Définition 2.3. — *Un pas, ou* jauge, *est une fonction δ définie sur $[a,b]$ et à valeurs dans $]0,+\infty[$. Une subdivision pointée $\mathcal{P} = \{(\mathrm{I}_1,t_1),\dots,(\mathrm{I}_n,t_n)\}$ sera dite adaptée au pas, ou encore δ-fine, si, pour tout $1 \le k \le n$ on a*

$$\mathrm{I}_k \subseteq [t_k - \frac{\delta(t_k)}{2}, t_k + \frac{\delta(t_k)}{2}].$$

On remarque en particulier que $|\mathrm{I}_k| \le \delta(t_k)$.

Le but du jeu est maintenant de trouver une telle jauge δ, qui nous assure que l'on a une bonne approximation de $f(b)-f(a)$ par $\mathrm{S}(f',\mathcal{P})$ dès que la subdivision $\mathcal{P}$ est δ-fine. Commençons par nous donner un $\varepsilon > 0$, éventuellement petit, par exemple $\varepsilon = 1/1000$. C'est l'ordre auquel on veut approximer notre formule. La discussion qui précède nous indique qu'une bonne jauge devrait permettre de dire que l'on est dans le « bon cas »de la formule des accroissements finis sur chaque I_k d'une subdivision δ-fine.

Une jauge qui satisfairait à la condition suivante (notons cette condition $\star$) serait *au poil :* pour tout $x \in [a,b]$, et tout $u \neq v$ satisfaisant

$$x - \delta(x)/2 \le u \le x \le v \le x + \delta(x)/2,$$

on a

$$|f(v) - f(u) - f'(x)(v-u)| < \frac{\varepsilon}{b-a}(v-u).$$

En effet, si $\mathcal{P}$ était une partition δ-fine pour une telle jauge δ, un petit coup de télescope et l'inégalité triangulaire donneraient :

$$\begin{aligned}|(f(b)-f(a)-\mathrm{S}(f',\mathcal{P})| &= \left|\sum_{k=1}^{n}(f(x_k)-f(x_{k-1}))-\sum_{k=1}^{n}(x_k-x_{k-1})f'(t_k)\right| \\ &\leq \sum_{k=1}^{n}|f(x_k)-f(x_{k-1})-f'(t_k)(x_k-x_{k-1})|,\end{aligned}$$

soit, du fait que δ satisfait l'hypothèse $\star$,

$$|(f(b)-f(a)-\mathrm{S}(f',\mathcal{P})| \leq \varepsilon.$$

Nous voyons que nous approchons du but : d'une part définir une intégrale à partir d'approximation d'aires, en utilisant des sommes de Riemann et des subdivisions à pas adapté, et d'autre part démontrer que cette intégrale vérifie, sous des hypothèses très générales, le théorème 1.1.

A-t-on espoir de trouver une jauge δ satisfaisant l'hypothèse $\star$? Si $x \in [a,b]$, f est dérivable en x. Ainsi

$$\lim_{t\to x}\frac{f(t)-f(x)}{t-x} = f'(x).$$

En réécrivant cette identité avec des *epsilon* et des *delta,* on obtient l'existence d'un $\delta(x)>0$ tel que, pour tout t satisfaisant $|t-x|<\delta(x)$, on ait

$$\left|\frac{f(t)-f(x)}{t-x}-f'(x)\right| < \varepsilon' = \frac{\varepsilon}{b-a},$$

soit

$$|f(t)-f(x)-f'(x)(t-x)| < \varepsilon'|t-x|.$$

Soient alors u, v comme dans la condition $\star$. Le théorème $a=a+b-b$ et l'inégalité triangulaire nous permettent d'écrire :

$$\begin{aligned} & |f(v)-f(u)-f'(x)(v-u)| \\ = \ & |[f(v)-f(x)-f'(x)(v-x)]-[f(u)-f(x)-f'(x)(u-x)]| \\ \leq \ & |f(v)-f(x)-f'(x)(v-x)|+|f(u)-f(x)-f'(x)(u-x)| \\ < \ & \varepsilon'(v-x)+\varepsilon'(x-u) = \frac{\varepsilon}{b-a}(v-u).\end{aligned}$$

Voila, nous avons trouvé une jauge δ satisfaisant $\star$, et nous sommes maintenant en mesure de passer à la définition générale de fonction intégrable et d'intégrale, *via* les sommes de Riemann et les subdivisions δ-fines.

3 Fonctions intégrables, intégrale

Nous avons vu, dans la section précédente, les notions de subdivision pointée, de somme de Riemann associée, ainsi que la notion de subdivision δ-fine pour une jauge δ sur $[a,b]$. Nous sommes donc en mesure de définir la notion d'intégrale et d'intégrabilité :

Définition 3.1. — *Soit f une fonction numérique définie sur $[a,b]$. La fonction f est dite intégrable s'il existe un nombre réel* S *tel que, pour tout $\varepsilon > 0$, il existe une jauge δ_ε sur $[a,b]$ telle que, pour toute subdivision δ_ε-fine $\mathcal{P}$, on ait*

$$|\mathrm{S}(f,\mathcal{P}) - \mathrm{S}| < \varepsilon.$$

On notera $\mathcal{I}([a,b])$ l'ensemble des fonctions intégrables sur $[a,b]$. Le nombre S *ci-dessus est appelé intégrale de la fonction f sur $[a,b]$ et est noté*

$$\int_{[a,b]} f(x)\,\mathrm{d}x \text{ ou encore } \int_a^b f(x)\,\mathrm{d}x, \text{ voire } \int_{[a,b]} f \text{ ou } \int_a^b f.$$

La stratégie pour montrer qu'une fonction donnée est intégrable est alors claire : on se fixe un réel $\varepsilon > 0$ et l'on cherche à construire la jauge δ_ε. Nous allons voir dans un instant, sur des exemples, que cela marche très bien pour un grand nombre de fonctions, commes les polynômes par exemple. Mais il nous faut tout d'abord régler un détail. Si jamais il existait une jauge δ n'admettant pas de subdivision δ-fine, alors toute fonction serait intégrable ! Cette chausse-trappe est heureusement évitée grâce au lemme suivant, dit lemme de Cousin :

Lemme 3.1. — *Pour toute jauge $\delta > 0$ sur $[a,b]$, il existe une subdivision δ-fine.*

Démonstration. — L'intervalle $[a,b]$ étant compact, il vérifie la propriété de Borel-Lebesgue. En particulier, commes les intervalles $]t-\delta(t)/2, t+\delta(t)/2[, t \in [a,b]$ recouvrent $[a,b]$ et sont ouverts, on peut trouver un nombre fini d'entres eux, $]t_i - \delta(t_i)/2, t_i+\delta(t_i)/2[, t_1 < t_2 < \cdots < t_n$, qui recouvrent $[a,b]$: nous venons simplement de dire que de tout recouvrement de $[a,b]$ par des ouverts de $\mathbf{R}$, on peut extraire un sous-recouvrement fini.

Nous supposerons de plus que ce sous-recouvrement est minimal, dans le sens où, si on lui enlève un de ses intervalles, ce qui reste ne recouvre plus tout $[a,b]$. Construisons maintenant une subdivision pointée, avec les t_i pour points de marquages. Pour cela, il nous suffit de prendre définir $x_0 = a, x_n = b$, et de choisir, pour tout $1 \leq i < n$, un point $x_i \in [\max\{t_{i+1} - \delta(t_{i+1})/2\}, \min\{t_i+\delta(t_i)/2, t_{i+1}\}]$. En effet, le précédent intervalle est non vide du fait que les $]t_i - \delta(t_i)/2, t_i+\delta(t_i)/2[$ recouvrent, donc on peut trouver un point x_i dedans. Enfin, on a clairement $t_{i+1} \in [x_i, x_{i+1}]$ et, par minimalité du recouvrement, on a $|x_{i+1} - x_i| \leq \delta(t_{i+1})$. La subdivision $\{([x_i, x_{i+1}], t_{i+1}), 0 \leq i \leq n-1\}$ est donc δ-fine.

Ouf ! Nous venons de voir, pour la première fois, le besoin d'un argument fin de topologie de la droite réelle. Ce n'est pas la dernière ! Avant de passer aux premiers exemples de fonctions intégrables, commençons par en donner la première propriété, à savoir l'unicité : il serait en effet embêtant que l'intégrale d'une même fonction puisse avoir plusieurs valeurs possibles.

Proposition 3.1. — *Si f est intégrable sur $[a, b]$, alors son intégrale* S *est unique.*

Démonstration. — C'est une conséquence immédiate de la séparation de **R**, qui dit, entre autre, que si une suite converge vers l et vers l' alors $l = l'$. Supposons donc que f ait deux intégrales, S et S'. Soit $\varepsilon > 0$. Par définition de S, il existe une jauge δ telle-que, pour toute subdivision δ-fine $\mathcal{P}$ on ait

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}| < \varepsilon.$$

Par ailleurs, il existe une autre jauge δ' telle-que, pour toute subdivision δ-fine $\mathcal{P}'$, on ait

$$|\mathrm{S}(f, \mathcal{P}') - \mathrm{S}'| < \varepsilon.$$

Ce qui serait maintenant bien, c'est de trouver une jauge δ'' telle que toute subdivision $\mathcal{P}''$ qui soit δ''-fine soit aussi fine pour δ et δ'. En prenant alors une telle subdivision $\mathcal{P}''$, on aurait, en additionnant les deux inégalités ci-dessus,

$$|\mathrm{S} - \mathrm{S}'| \leq |\mathrm{S} - \mathrm{S}(f, \mathcal{P})| + |\mathrm{S}(f, \mathcal{P}) - \mathrm{S}'| < 2\varepsilon,$$

et, comme cette inégalité est vraie pour tout ε, la proposition serait démontrée : $\mathrm{S} = \mathrm{S}'$. En fait, un jauge telle que δ'' existe toujours, c'est une conséquence de l'utile lemme suivant :

Lemme 3.2. — *Soient $\delta_1, \dots, \delta_l$ un nombre fini de jauges sur $[a, b]$. Alors $\delta = \min\{\delta_1, \dots, \delta_n\}$ est une jauge. De plus, toute subdivision pointée δ-fine est aussi δ_i-fine pour $1 \leq i \leq l$.*

La démonstration sera vue en travaux dirrigés, en guise d'entrainement au maniement des jauges.

Premiers exemples de fonctions intégrables, intégrales de Riemann et de Lebesgue.

1. Commençons par les fonctions constantes :

   Proposition 3.2. — *Soit $c \in \mathbf{R}$, et notons encore c la fonction constante égale à c sur $[a, b]$. Alors c est intégrable et*

   $$\int_a^b c \,\mathrm{d}x = (b - a)c.$$

   C'est une évidence une fois que l'on a remarqué que les sommes de Riemann ne dépendent pas de la subdivision choisie, et valent toutes $(b - a)\mathrm{C}$ (sommes télescopiques).

2. Essayons maintenant d'intégrer un monôme de la forme $f(x) = x^n$, pour un entier $n \geq 0$. Bien-sûr, nous connaissons le résultat : l'intégrale de x^n entre a et b vaut $\mathrm{F}(b) - \mathrm{F}(a)$, où $\mathrm{F}(x) = (n+1)^{-1}x^{n+1}$ est une primitive de f. Nous revenons donc aux idées de la section précédente : si $\mathcal{P} = \{([x_{k-1}, x_k], t_k), 1 \leq k \leq n\}$ est

une subdivision pointée, alors la formule des accroissements finis nous garantit, pour tout k, de l'existence d'un nombre $u_k \in [x_{k-1}, x_k]$ tel que

$$\mathrm{F}(x_k) - \mathrm{F}(x_{k-1}) = f(u_k)(x_k - x_{k-1}),$$

soit, en donnant à f et F leurs vrais significations,

$$\frac{1}{n+1}(x_k^{n+1} - x_{k-1}^{n+1}) = u_k^n(x_k - x_{k-1}).$$

Additionons ces identités, et nous obtenons une somme télescopique :

$$\frac{1}{n+1}(b^{n+1} - a^{n+1}) = \sum_{k=1}^{n} u_k^n(x_k - x_{k-1}),$$

soit, en faisant la différence avec la somme de Riemann associée à $\mathcal{P}$:

$$\frac{1}{n+1}(b^{n+1} - a^{n+1}) - \mathrm{S}(f, \mathcal{P}) = \sum_{k=1}^{n}(u_k^n - t_k^n)(x_k - x_{k-1}).$$

Fixons $\varepsilon > 0$. Il existe un nombre $\delta > 0$ tel que, pour tout $x, y \in [a, b]$ vérifiant $|x - y| < \delta$, alors $|x^n - y^n| < \varepsilon$. Ceci exprimme *l'uniforme continuité* de la fonction continue f sur l'intervalle compact $[a, b]$. Dans le cas de $f(x) = x^n$, nous pouvons expliciter δ en fonction de ε, en prenant, par exemple, $\delta = \max\{n|b|^{n-1}, n|a|^{n-1}\}$. Si l'on note encore δ la jauge constante égale à δ, on peut majorer la différence ci-dessus par

$$\begin{aligned} \left|\frac{1}{n+1}(b^{n+1} - a^{n+1}) - \mathrm{S}(f, \mathcal{P})\right| &\leq \sum_{k=1}^{n}(u_k^n - t_k^n)(x_k - x_{k-1}) \\ &\leq \sum_{k=1}^{n} \varepsilon(x_k - x_{k-1}) = \varepsilon(b - a), \end{aligned}$$

car $u_k, t_k \in \mathrm{I}_k$ et donc $|u_k^n - t_k^n| < \varepsilon$ par construction de δ. Nous pouvons conlure cette discussion par la proposition suivant :

Proposition 3.3. — *Soient $[a, b]$ un intervalle compact de* $\mathbf{R}$*, et* $n \in \mathbf{N}$*. La fonction monomiale $f(x) = x^n$ est intégrable sur $[a, b]$ et*

$$\int_a^b x^n \,\mathrm{d}x = \frac{1}{n+1}(b^{n+1} - a^{n+1}),$$

ce qui est le résultat attendu.

3. Dans l'exemple ci-dessus, nous avons vu qu'il suffisait, pour montrer l'intégrabilité de x^n, de considérer des jauges constantes. Nous venons de retrouver, comme cas particulier des fonctions intégrables, les fontions *intégrables au sens de Riemann,* dites encore *Riemann-intégrables.*

Définition 3.2. — *Une fonction f définie sur $[a, b]$ est dite Riemann-intégrable s'il existe* $\mathrm{S} \in \mathbf{R}$*, vérifiant que, pour tout $\varepsilon > 0$, il existe un* nombre réel $\delta > 0$ *tel que, pour toute partition δ-fine (pour la jauge* constante δ*), on ait*

$$|\mathrm{S} - \mathrm{S}(f, \mathcal{P})| < \varepsilon.$$

Il est bien évident que toute fonction Riemann-intégrable est intégrable, mais nous allons voir dans un instant que la réciproque est fausse : les fonctions intégrables au sens de Riemann doivent présenter une certaine uniformité dans leurs variations locales.
Soit $\chi : [0,1] \longrightarrow \mathbf{R}$ la fonction de Dirichlet, ou fonction indicatrice des nombres rationnels :

$$\chi(x) = 1 \text{ si } x \in \mathbf{Q} \cap [0,1],\ \chi(x) = 0 \text{ sinon.}$$

Montrons que χ est intégrable, d'intégrale nulle. Pour cela, remarquons que $\mathrm{D} = \mathbf{Q} \cap [0,1]$ est *dénombrable,* car $\mathbf{Q}$ lui-même est dénombrable. On peut donc énumérer D en $\mathrm{D} = \{r_1, r_2, \ldots\}$. Fixons $\varepsilon > 0$, et définissons la jauge δ_ε par

$$\begin{cases} \delta_\varepsilon(x) = 1 \text{ si } x \notin \mathrm{D} \\ \delta_\varepsilon(x) = \varepsilon 2^{-j} \text{ si } x = r_j, j = 1, 2, \ldots \end{cases}$$

Si $\mathcal{P} = \{(\mathrm{I}_k, t_k), 1 \leq k \leq n\}$ est une subdivision δ_ε-fine, alors on aura $\chi(t_k)|\mathrm{I}_k| = 0$ si $t_k \notin \mathrm{D}$ et $\chi(t_k)|\mathrm{I}_k| = \varepsilon 2^{-j}$ si $t_k = r_j$. Ainsi,

$$\mathrm{S}(\chi, \mathcal{P}) = \sum_{k=1}^{n} \chi(t_k)|\mathrm{I}_k| \leq \varepsilon \sum_{j=1}^{+\infty} 2^{-j} \leq \varepsilon.$$

La fonction χ est donc intégrable, d'intégrale nulle. Elle n'est cependant pas intégrable au sens de Riemann. En voici, succinctement, la raison. Si $\mathcal{P} = \{(\mathrm{I}_k, t_k)\}$ est une subdivision δ-fine, pour une jauge *constante* δ, alors les subdivisions déduites de $\mathcal{P}$ en gardant les intervalles I_k mais en perturbant les points de marquage t_k sont encore δ-fines. Comme D et $[0,1] - \mathrm{D}$ sont denses dans $[0,1]$, on peut trouver des partitions δ-fines telles que tous les t_k soient hors de D et d'autres telles que tous soient dans D. Ainsi, pour toute constante $\delta > 0$, il existe des partitions δ-fines telles que $\mathrm{S}(f, \mathcal{P}) = 0$ et d'autres telles que $\mathrm{S}(f, \mathcal{P}) = 1 > 0$. En fait, en choisissant bien les intervalles et leurs points de marquage, on peut obtenir à la limite toute valeur entre 0 et 1. La fonction χ ne peut donc être Riemann-intégrable.

4. Pour finir, donnons deux exemples de fonctions dont on peut montrer à la main qu'elles sont intégrables. La première est non-bornée sur son intervalle de définition, la seconde est non bornée et d'intégrale non absolument convergente (*voir* le chapitre « Intégrales impropres »). Ces deux exemples montrent que l'intégrale présentée dans ce cours est une extension stricte à la fois de l'intégrale de Riemann et de l'intégrale de Lebesgue (*voir* plus loin). Les démonstration font l'objet d'exercices de travaux dirigés.
Soit $f : [0,1] \longrightarrow \mathbf{R}$ la fonction définie par

$$f(0) = 0,\ f(x) = \frac{1}{\sqrt{x}} \text{ si } x \in]0,1].$$

Alors f est intégrable sur $[0,1]$ et

$$\int_0^1 \frac{1}{\sqrt{x}}\, \mathrm{d}x = 2.$$

Considérons pour finir la série de terme général $(-1)^n/n$. C'est une série convergente mais non absolument convergente. On sait en particulier qu'en réordonnant les termes de cette série on peut la faire converger vers n'importe quelle limite : elle n'est pas commutativement convergente (en fait, une série de nombres réelles est commutativement convergente si, et seulement si, elle est absolument convergente). Construisons, sur l'intervalle $[0,1]$, une fonction en escalier dont l'intégrale soit la somme de cette série :

$$g(x) = \begin{cases} (-1)^{k+1}2^k/k \text{ si } x \in [1-2^{-(k-1)}, 1-2^{-k}[,\ k \in \mathbf{N}^* \\ 0 \text{ si } x = 0. \end{cases}.$$

Entre 0 et 1/2, g vaut 2/1, entre 1/2 et 3/4 g vaut $-2^2/2$ *etc...* En fait, g est intégrable sur $[0,1]$, et

$$\int_0^1 g(x)\,\mathrm{d}x = \sum_{k=1}^{+\infty} \frac{(-1)^{k+1}}{k},$$

on a donc un exemple d'intégrale dite semi-convergente : f est intégrable, mais pas $|f|$. Celà nous amène à la définition suivante :

Définition 3.3. — *Une fonction f définie sur* $\mathrm{I} = [a,b]$ *est dite intégrable au sens de Lebesgue, ou* Lebesgue-intégrable, *si f* ET *$|f|$ sont intégrables.*

La classe des fonctions Lebesgue-intégrable est très importante en analyse, en particulier parce-qu'elle peut se généraliser à d'autres espaces que $\mathbf{R}$: dimension plus grandes, courbes, surfaces, espaces probabilisés *etc...* Nous y reviendrons plus en détail dans le chapitre sur les théorèmes de convergence.

Finissons cette section par deux procédures utiles dans le maniement des jauges et des subdivisions. La première consiste à couper une subdivision selon ses points de marquage, la seconde consiste à forcer certains points à être de marquage.

Coupage–collage d'une subdivision le long de ses points de marquage. Soient δ une jauge sur $[a,b]$ et $\mathcal{P} = \{([x_{k-1}, x_k], t_k), k = 1, \dots, n\}$ une subdivision δ-fine. Formons la subdivision $\mathcal{P}^\star$ obtenue en coupant $\mathcal{P}$ en ses points de marquages :

$$\mathcal{P}^\star = \{([x_0,t_1],t_1),([t_1,x_1],t_1),([x_1,t_2],t_2),\dots,([x_{n-1},t_n],t_n),([t_n,b],t_n)\}.$$

Il est clair que $\mathcal{P}^\star$ est δ-fine. Mais par ailleurs, pour toute fonction f définie sur $[a,b]$, on a

$$\mathrm{S}(f,\mathcal{P}^\star) = \sum_{k=1}^{n} [f(t_k)(x_k - t_k) + f(t_k)(t_k - x_{k-1})] = \sum_{k=1}^{n} f(t_k)(x_k - x_{k-1}) = \mathrm{S}(f,\mathcal{P}),$$

et donc passer de $\mathcal{P}$ à $\mathcal{P}^\star$, et réciproquement, est sans effet, ni sur la finesse ni sur les sommes de Riemann. Remarquons que l'on aurait pu faire une procédure partielle, en ne demandant qu'à certains points t_k de devenir des extrémités.

Forçage des points de marquage. Considérons pour l'instant un cas simple, celui d'un point $c \in [a,b]$, et définissons la jauge $\delta_c(t)$ par

$$\begin{cases} \delta_c(c) = 1/4 \\ \delta_c(t) = |t-c|/2 \text{ si } t \neq c. \end{cases}$$

Soit alors $\mathcal{P}$ une partition δ-fine et k l'indice tel que $c \in \mathrm{I}_k$. Comme $\mathcal{P}$ est δ-fine, on a

$$0 \leq |t_k - c| \leq \delta(t_k).$$

Si l'on avait $t_k \neq c$, on aurait $|t_k - c| = 2\delta(t_k) > 0$, contredisant l'inégalité ci-dessus. Donc $t_k = c$, et l'on a forcé c à être un point de marquage de $\mathcal{P}$. Ce procédé peut

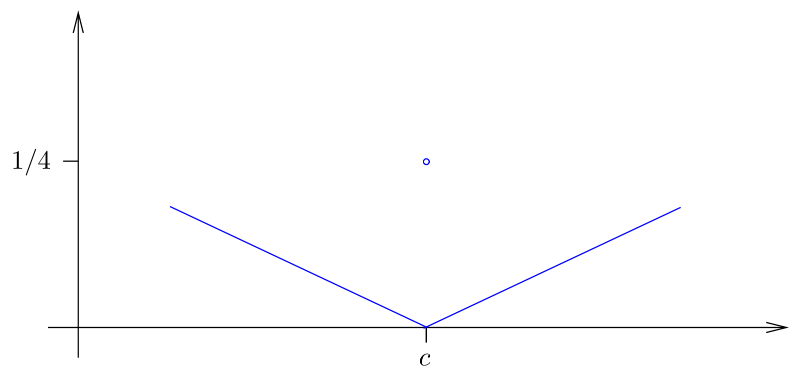

FIG. 4: Jauge forçant c à être point de marquage.

se généraliser grandement. Soit δ une jauge et $c_1 < \cdots < c_l$ des points de $[a, b]$. Définissons alors une jauge δ' par

$$\delta'(t) = \min\{\delta(t), \delta_{c_1}(t), \ldots, \delta_{c_l}(t)\}.$$

Alors toute subdivision δ'-fine est aussi δ-fine, mais en plus les points $c_1, \ldots, c_l$ sont des points de marquage de $\mathcal{P}$.

Application : les fonctions en escalier sont intégrables. Soient $a < c < b$, α, β deux constantes, et f la fonction en escalier valant α sur $[a, c[$ et β sur $[c, b]$.

Le point de discontinuité, en c, est celui qui pose problème par rapport au cas des fonctions constantes. Fixons $\varepsilon > 0$ et soit δ la jauge définie par

$$\begin{cases} \delta(x) = \dfrac{1}{2}|x - c| \text{ si } x \neq c, \\ \delta(c) = \dfrac{\varepsilon}{\max\{1, |\beta - \alpha|\}}. \end{cases}$$

Si $\mathcal{P}$ est une subdivision δ-fine, alors c est un des points de marquage. Quitte à couper $\mathcal{P}$ le long de son point de marquage c, on peut supposer que c est sommet des deux intervalles successifs $[x_{k-1}, c]$ et $[c, x_k]$ de $\mathcal{P}$. Mais alors les sommes de Riemann se « télescopent »sur chacun des intervalles $[a, c[$ et $[c, b]$, et l'on obtient

$$\mathrm{S}(f, \mathcal{P}) = \alpha(x_{k-1} - a) + \beta(b - x_{k-1}) = \alpha(c - a) + \beta(b - c) + \alpha(x_{k-1} - c) + \beta(c - x_{k-1}),$$

soit, par le choix de la jauge δ,

$$|\mathrm{S}(f, \mathcal{P}) - (\alpha(c - a) + \beta(b - c))| \leq \varepsilon.$$

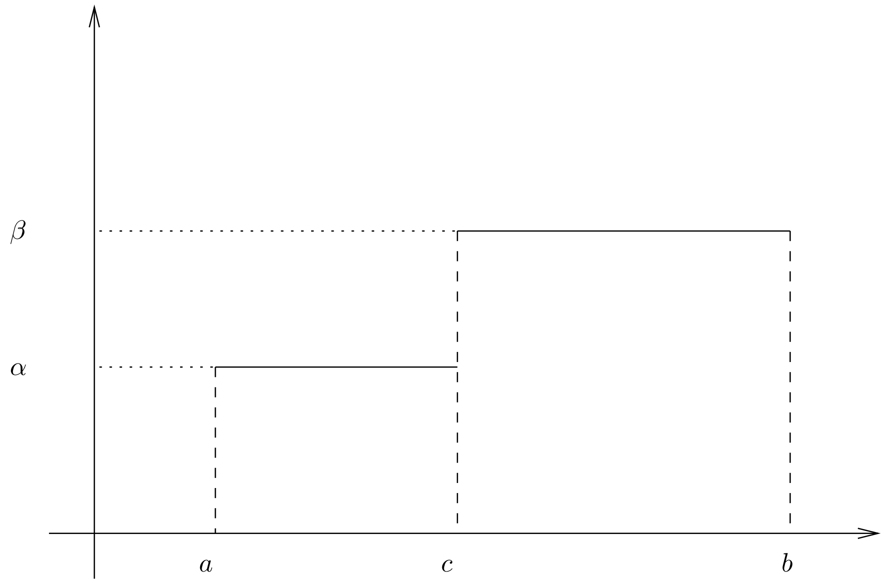

Fig. 5: Une fonction en escalier.

Proposition 3.4. — *Soit f une fonction en escalier sur $[a, b]$. Supposons que f prenne les valeurs $\alpha_1, \ldots, \alpha_n$ sur les intervalles $\mathrm{I}_1 \leq \mathrm{I}_k \leq \cdots \leq \mathrm{I}_n$, qui forment une partition une $[a, b]$. Alors f est intégrable et*

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{k=1}^{n} \alpha_k |\mathrm{I}_k|.$$

4 Critère de Cauchy, applications

L'intégrale est définie par un processus de passage à la limite. Jusqu'à maintenant, nous avons démontré des résultats d'intégrabilité pour des fonctions dont nous connaissions, *a priori,* le caractère intégrable et la valeur de l'intégrale. Pour démontrer des résultats plus généraux, il est indispensable de disposer d'un critère d'intégrabilité qui ne fasse pas appel à la valeur de l'intégrale. Pour la convergence des suites de nombres réels (et plus généralement d'éléments d'un espace complet), un tel critère existe, c'est le critère de Cauchy. Celui-ci permet souvent de démontrer qu'une suite converge sans que l'on connaisse *a priori* sa limite. Nous disposons, dans le cadre de l'intégrale, d'un outil analogue.

Théorème 4.1. — *Une fonction $f : \mathrm{I} = [a, b] \longrightarrow \mathbf{R}$ est intégrable si, et seulement si, pour tout $\varepsilon > 0$, il existe une jauge δ_ε telle que, si $\mathcal{P}$ et $\mathcal{Q}$ sont deux subdivisions δ_ε-fines, alors*

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}(f, \mathcal{Q})| \leq \varepsilon.$$

Démonstration. — Si f est intégrable, fixons $\varepsilon > 0$ et trouvons une jauge δ telle que pour toute subdivisions δ-fines $\mathcal{P}$ et $\mathcal{Q}$, on ait

$$\left|\mathrm{S}(f, \mathcal{P}) - \int_\mathrm{I} f\right| \leq \varepsilon/2 \quad \text{et} \quad \left|\mathrm{S}(f, \mathcal{Q}) - \int_\mathrm{I} f\right| \leq \varepsilon/2.$$

Alors $|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}(f, \mathcal{Q})| \leq \varepsilon$, et le critère de Cauchy est vérifié. Réciproquement, supposons que f vérifie le critère de Cauchy. Soit (δ_n) une suite de jauges telles que si $\mathcal{P}, \mathcal{Q}$ sont δ_n fines, alors

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}(f, \mathcal{Q})| \leq \frac{1}{n}.$$

Quitte à prendre le minimum des $\delta_i, i = 1, \ldots, n$, on peut supposer (δ_n) décroissante. Choisissons, pour chaque n, une subdivision $\mathcal{P}_n$ qui soit δ_n-fine. Pour tout $m > n$, $\mathcal{P}_m$ est aussi δ_n fine. De plus, $|\mathrm{S}(f, \mathcal{P}_m) - \mathrm{S}(f, \mathcal{P}_n)| \leq 1/n$: la suite $(\mathrm{S}(f, \mathcal{P}_n))$ est de Cauchy dans $\mathbf{R}$. Elle converge donc vers une limite, que nous noterons $\mathcal{S}$. Fixons maintenant un $\varepsilon > 0$, et prenons $n > 2/\varepsilon$, ainsi qu'une jauge $\mathcal{Q}$ qui soit δ_n-fine. Alors

$$|\mathrm{S}(f, \mathcal{Q}) - \mathcal{S}| \leq |\mathrm{S}(f, \mathcal{Q}) - \mathrm{S}(f, \mathcal{P}_n)| + |\mathrm{S}(f, \mathcal{P}_n) - \mathcal{S}| \leq 1/n + 1/n \leq \varepsilon,$$

ce qui montre que f est intégrable sur I.

Comme première application du critère de Cauchy, on peut par exemple donner une caractérisation des fonctions intégrables par un procédé d'encadrement. Celui-ci permet ensuite de traîter de vastes classes de fonctions, comme les fonctions réglées, ou encore de montrer facilement le théorème d'interversion de limite et d'intégrale pour la convergence uniforme.

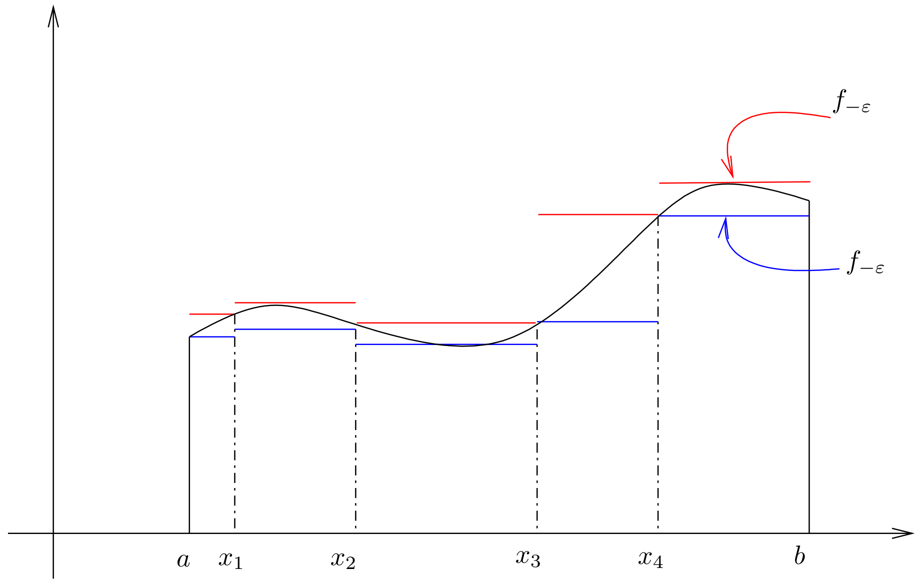

FIG. 6: Encadrement d'une fonction continue par deux fonctions en escalier.

Théorème 4.2. — *Une fonction f est intégrable sur* I *si et seulement si, pour tout $\varepsilon > 0$, il existe deux fonctions intégrables $f_{-\varepsilon}$ et $f_{+\varepsilon}$ telles que*

$$f_{-\varepsilon} \leq f \leq f_{+\varepsilon} \quad \textit{et} \quad \left| \int_{\mathrm{I}} f_{+\varepsilon} - \int_{\mathrm{I}} f_{-\varepsilon} \right| \leq \varepsilon.$$

DÉMONSTRATION. — Si f est intégrable, elle vérifie clairement les conclusions du théorème, en prenant $f_{-\varepsilon} = f = f_{+\varepsilon}$. Supposons maintenant que f vérifie la condition d'encadrement ci-dessus. Si l'on se donne $\varepsilon > 0$, on trouve $f_{-\varepsilon}$ et $f_{+\varepsilon}$, ainsi que des jauges δ^+ et δ^- pour ces deux fonctions, qui sont telles que pour toute subdivision $\delta^{\pm}$-fine $\mathcal{P}$, on ait

$$\left| \mathrm{S}(f_{\pm\varepsilon}, \mathcal{P}) - \int_{\mathrm{I}} f_{\pm\varepsilon} \right| \leq \varepsilon.$$

On pose $\delta = \min\{\delta^+, \delta^-\}$. Pour toute subdivision pointée $\mathcal{P}$, on a

$$\mathrm{S}(f_{-\varepsilon}) \leq \mathrm{S}(f, \mathcal{P}) \leq \mathrm{S}(f_{+\varepsilon}, \mathcal{P}).$$

Si donc $\mathcal{P}, \mathcal{Q}$ sont δ-fines, alors

$$\int_{\mathrm{I}} f_{-\varepsilon} - \varepsilon \leq \mathrm{S}(f, \mathcal{P}) \leq \int_{\mathrm{I}} f_{+\varepsilon} + \varepsilon \quad \text{et} \quad -\int_{\mathrm{I}} f_{+\varepsilon} - \varepsilon \leq -\mathrm{S}(f, \mathcal{Q}) \leq -\int_{\mathrm{I}} f_{-\varepsilon} + \varepsilon,$$

ce qui implique immédiatement, vues les hypothèses sur $f_{\pm\varepsilon}$, que

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}(f, \mathcal{Q})| \leq \int_{\mathrm{I}} (f_{+\varepsilon} - f_{-\varepsilon}) + 2\varepsilon \leq 3\varepsilon.$$

La fonction f vérifie donc le critère de Cauchy, et est intégrable.

Donnons un corollaire important de ce théorème :

Corollaire 4.1. — *Une fonction continue sur* $[a, b]$ *est intégrable (et même Riemann-intégrable).*

Voici l'idée de la démonstration : f est continue sur $[a, b]$, qui est compact, donc elle y est uniformément continu. On dispose donc d'un module de continuité uniforme δ_ε qui permet, en prenant des subdivisions $\delta_{-\varepsilon}$-fine, de construire deux fonctions en escaliers qui encadrent f (*voir* la figure 6) et vérifient les hypothèses du théorème 4.2. On retrouve au passage une méthode d'intégration dérivée de celle de Riemann et dûe à Gaston Darboux.

Corollaire 4.2. — *Soit* (f_n) *une suite de fonctions intégrables sur* $[a, b]$*, et* f *une fonction définie sur* $[a, b]$*. Si* (f_n) *converge uniformément vers* f*, alors* f *est intégrable et*

$$\lim_{n \to +\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Nous verrons dans le chapitre sur les théorèmes de convergence un énoncé, le Théorème de convergence dominée de Lebesgue, qui est bien plus général et beaucoup plus puissant.

5 Propriétés élémentaires de l'intégrale

La propriété la plus simple à démontrer est la linéarité de l'intégrale en la fonction à intégrer.

Théorème 5.1. — *Si f et g sont des fonctions intégrables sur* I, *et si λ est un nombre réel, alors $f+g$ et λf sont intégrables, et l'on a les identités*

$$\int_{\mathrm{I}} (f+g) = \int_{\mathrm{I}} f + \int_{\mathrm{I}} g \ \textit{et} \ \int_{\mathrm{I}} \lambda f = \lambda \int_{\mathrm{I}} f.$$

Démonstration. — Fixons $\varepsilon > 0$. Il existe deux jauges δ_1, δ_2 telles que pour toute subdivision $\mathcal{P}$ qui soit δ_1-fine (*resp.* δ_2-fine), on ait

$$|\mathrm{S}(f, \mathcal{P}) - \int_{\mathrm{I}} f| \leq \varepsilon \quad \textit{resp.} \quad |\mathrm{S}(g, \mathcal{P}) - \int_{\mathrm{I}} g| \leq \varepsilon.$$

On pose alors $\delta = \min\{\delta_1, \delta_2\}$, qui est encore une jauge, et l'on note que si $\mathcal{P}$ est δ-fine, elle est aussi δ_1- et δ_2-fine. On conclut alors aisément.

Il est tout aussi facile de démontrer que l'intégrale est une fonctionnelle positive :

Théorème 5.2. — *Si $f \in \mathcal{I}(\mathrm{I})$ vérifie $f \geq 0$, alors*

$$\int_{\mathrm{I}} f \geq 0.$$

En particulier, si $f, g \in \mathcal{I}(\mathrm{I})$ vérifient $f \geq g$, alors

$$\int_{\mathrm{I}} f \geq \int_{\mathrm{I}} g.$$

Afin de justifier la notation avec les bornes des intervalles pour l'intégrale, il nous faut démontrer la relation de Chasles :

Théorème 5.3. — *Soient $f : [a,b] \longrightarrow \mathbf{R}$ et $c \in]a,b[$. Alors f est intégrable sur $[a,b]$ si, et seulement si, ses restrictions à $[a,c]$ et $[c,b]$ le sont. Dans ce cas, on a la relation*

$$\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f.$$

C'est pourquoi l'on note aussi l'intégrale de f sur l'intervalle $[a,b]$ sous la forme

$$\int_{[a,b]} f = \int_a^b f.$$

Démonstration. — Supposons pour commencer f intégrable sur $\mathrm{I}_1 = [a,c]$ et $\mathrm{I}_2 = [c,b]$. Montrons qu'alors f est intégrable sur $\mathrm{I} = [a,b]$, et que l'on a la relation de Chasles. Pour cela, fixons $\varepsilon > 0$ et donnons-nous deux jauges δ_1 et δ_2 sur I_1 et I_2 respectivement, telles que pour toute subdivision pointée δ_i-fine $\mathcal{P}_i$ de I_i on ait

$$|\mathrm{S}(f, \mathcal{P}_i) - \int_{\mathrm{I}_i} f| \leq \varepsilon.$$

Définissons maintenant une jauge δ sur I en entier, qui forcera toute subdivision δ-fine à être d'une part δ_i-fine, $i = 1, 2$, et d'autre part à avoir c comme point de marquage :

$$\delta(x) = \min\{\delta_i(x), \frac{1}{2}|x - c|\}, x \in \mathrm{I}_i,\ \delta(c) = \min\{\delta_1(c), \delta_2(c)\}.$$

Par le procédé de marginalistaion des points de marquage, on peut supposer que si $\mathcal{P}$ est δ-fine, c est borne d'un des intervalles de $\mathcal{P}$, et donc $\mathcal{P}$ induit sur chacun des intervalles I_i une subdivision pointée $\mathcal{P}_i$ qui est $\delta_{|\mathrm{I}_i}$-fine, donc δ_i-fine. Comme alors $\mathrm{S}(f, \mathcal{P}) = \mathrm{S}(f, \mathcal{P}_1) + \mathrm{S}(f, \mathcal{P}_2)$, il vient immédiatement que

$$|\mathrm{S}(f, \mathcal{P}) - (\int_{\mathrm{I}_1} f + \int_{\mathrm{I}_2} f)| \leq |\mathrm{S}(f, \mathcal{P}_1) - \int_{\mathrm{I}_1} f| + |\mathrm{S}(f, \mathcal{P}_2) - \int_{\mathrm{I}_2} f| \leq 2\varepsilon,$$

ce qui démontre que f est intégrable, et que la relation de Chasles est vérifiée.

Pour démontrer l'implication réciproque, le mieux est d'utiliser le critère de Cauchy. Fixons donc $\varepsilon > 0$ et soit δ une jauge sur I qui vérifie le critère de Cauchy pour ε. Si $\mathcal{P}_1$ et $\mathcal{Q}_1$ sont deux subdivisions $\delta_{|\mathrm{I}_1}$-fines, on peut, comme on l'a fait ci-dessus, leur rajouter des intervalles de sorte que c soit un point de marquage et une borne de l'intervalle de la subdivision auquel il appartient. On prolonge ensuite $\mathcal{P}_i$ et $\mathcal{Q}_i$ en deux subdivisions $\mathcal{P}$ et $\mathcal{Q}$ de I qui sont δ-fines, en utilisant les mêmes points et intervalles sur I_2. Alors

$$\mathrm{S}(f_{|\mathrm{I}_1}, \mathcal{P}_1) - \mathrm{S}(f_{|\mathrm{I}_1}, \mathcal{Q}_1) = \mathrm{S}(f, \mathcal{P}) - \mathrm{S}(f, \mathcal{Q}),$$

et comme $\mathcal{P}$ et $\mathcal{Q}$ sont δ-fine, on en déduit que

$$|\mathrm{S}(f_{|\mathrm{I}_1}, \mathcal{P}_1) - \mathrm{S}(f_{|\mathrm{I}_1}, \mathcal{Q}_1)| \leq \varepsilon,$$

ce qui, par le critère de Cauchy, nous permet de conclure que f est intégrable sur I_1.

Ce théorème a pour corrolaire que la restriction d'une fonction intégrable sur I à un sous-intervalle est encore intégrable. Enfin, on peut définir l'intégrale entre des bornes pas nécessairement mises dans le bon ordre, en prenant les bonnes conventions de signe. Cela sera utile lors de l'utilisation de changements de variables décroissants, et vous est déjà très familier.

Inégalités de Cauchy-Schwarz, Hölder et Minkowski.

Soient f et g deux fonctions Lebesgue-intégrables sur $[a, b]$, et telles que f^2 et g^2 soient aussi intégrables. On dit que f et g sont des fonctions « L^2 », et on note $f, g \in \mathcal{L}^2([a, b])$[1]. Si $x \in [a, b]$, on a bien-sûr

$$0 \leq (f(x) - g(x))^2 = f(x)^2 + g(x)^2 - 2f(x)g(x).$$

On en déduit que

$$|f(x)g(x)| \leq \frac{1}{2}(f(x)^2 + g(x)^2),$$

[1]En toute rigueur, il faudrait d'abord définir la notion de fonction mesurable, mais celà nous mènerait trop loin ; la version présentée ici est suffisante pour beaucoup d'applications.

et donc que fg et $|fg|$ sont intégrables (ici, je triche un petit peu). De plus,

$$\int_a^b |f(x)g(x)|\,\mathrm{d}x \leq \frac{1}{2}\left(\int_a^b f(x)^2\,\mathrm{d}x + \int_a^b g(x)^2\,\mathrm{d}x\right).$$

Soit maintenant $t > 0$. Le même argument que ci-dessus, utilisant l'identité remarquable $(\alpha-\beta)^2 = \alpha^2+\beta^2-2\alpha\beta$, donne, pour tout $x \in [a,b]$:

$$0 \leq (\sqrt{t}f(x) - \sqrt{t^{-1}}g(x))^2 = tf(x)^2 + \frac{1}{t}g(x)^2 - 2f(x)g(x),$$

soit

$$2|f(x)g(x)| \leq tf(x)^2 + \frac{1}{t}g(x)^2.$$

Ainsi,

$$2\int_a^b |f(x)g(x)|\,\mathrm{d}x \leq t\int_a^b f(x)^2\,\mathrm{d}x + \frac{1}{t}\int_a^b g(x)^2\,\mathrm{d}x. \tag{1}$$

Le second membre de cette inégalité prend sa valeur minimale pour

$$t = \left(\int_a^b g(x)^2\,\mathrm{d}x\right)^{1/2} \Big/ \left(\int_a^b f(x)^2\,\mathrm{d}x\right)^{1/2},$$

ce qui, reporté dans l'inégalité 1, donne :

$$\int_a^b |f(x)g(x)|\,\mathrm{d}x \leq \left(\int_a^b f(x)^2\,\mathrm{d}x\right)^{1/2}\left(\int_a^b g(x)^2\,\mathrm{d}x\right)^{1/2}.$$

Cette inégalité est connue sous le nom *d'inégalité de Cauchy–Schwarz.* Elle est un cas particulier des inégalités de Hölder, qui se démontrent de la même manière, et qui ont pour corollaire l'inégalité de Minkowski. Nous ne démontrons pas ici ces deux dernières. Pour ne pas entrer dans des détails délicats, nous énonçons toutes ces inégalités dans le cas des fonctions continues :

Théorème 5.4. — *Soient $[a,b]$ un intervalle compact et f, g des fonctions continues sur $[a,b]$.*

1. **Inégalité de Cauchy-Scwharz :**
$$\int_a^b |f(x)g(x)|\,\mathrm{d}x \leq \left(\int_a^b f(x)^2\,\mathrm{d}x\right)^{1/2}\left(\int_a^b g(x)^2\,\mathrm{d}x\right)^{1/2}.$$
2. **Inégalité de Hölder :** *soit $p > 1$, et q tel que $1/p+1/q = 1$ (*i.e.* $q = p/(p-1)$, on dit que p et q sont conjugués). Alors*
$$\int_a^b |f(x)g(x)|\,\mathrm{d}x \leq \left(\int_a^b f(x)^p\,\mathrm{d}x\right)^{1/p}\left(\int_a^b g(x)^q\,\mathrm{d}x\right)^{1/q}.$$
3. **Inégalité de Minkowski :** *soit $p > 1$, alors*
$$\left(\int_a^b |f(x)+g(x)|^p\,\mathrm{d}x\right)^{1/p} \leq \left(\int_a^b |f(x)|^p\,\mathrm{d}x\right)^{1/p} + \left(\int_a^b |g(x)|^p\,\mathrm{d}x\right)^{1/p}.$$

On peut étendre ces inégalités à de larges classes de fonctions, mais les énoncer réclammerait l'introduction de notions qui dépasseraient notre propos.

6 Intégrales et primitives

Revenons brièvement, dans cette section, sur notre introduction. Nous allons démontrer, comme annoncé, que l'intégrale que nous avons introduite ici répond complètement au problème du calcul inverse des dérivées. Nous ne démontrerons pas en cours l'énoncé le plus général dans cette direction, même si celui-ci est accessible assez facilement. Nous nous contenterons de prouver le théorème suivant, qui est déjà une amélioration sensible de l'énoncé analogue dans les théories de Riemann et de Lebesgue.

Théorème 6.1. — *Soit* $\mathrm{F} : [a, b] \longrightarrow \mathbf{R}$ *dérivable sur* $[a, b]$*, de dérivée* f*. Alors* f *est intégrable et l'on a*

$$\int_a^b f = \mathrm{F}(b) - \mathrm{F}(a).$$

Démonstration. — Précisons ici les hypothèses du théorème : F doit être continue sur $[a, b]$, dérivable sur $]a, b[$, et admettre une dérivée à droite en a (*resp.* à gauche en b).

La dérivabilité de F, définie par des limites ponctuelles, nous fournit une jauge naturelle. Fixons $\varepsilon > 0$. Nous avons vu plus haut que pour tout $x \in [a, b]$, il existe un $\delta(x) > 0$ tel que, pour tout $y, z \in [a, b]$ vérifiant $z - y \leq \delta(x)$ et $y \leq x \leq z$, on ait

$$(\dagger) \qquad |\mathrm{F}(z) - \mathrm{F}(y) - (z - y)f(x)| \leq \varepsilon(z - y).$$

Si $\mathcal{P}$ est une subdivision pointée δ-fine, on a par définition

$$|\mathrm{S}(f, \mathcal{P}) - (\mathrm{F}(b) - \mathrm{F}(a))| = \left|\sum_{k=1}^{n} (\mathrm{F}(x_k) - \mathrm{F}(x_{k-1})) - (x_k - x_{k-1})f(t_k))\right|,$$

et, en sommant les inégalités ($\dagger$) pour chaque terme du second membre, on trouve la majoration

$$|\mathrm{S}(f, \mathcal{P}) - (\mathrm{F}(b) - \mathrm{F}(a))| \leq \varepsilon \sum_{k=1}^{n} (c_k - c_{k-1}) = \varepsilon(b - a),$$

ce qui achève la démonstration du théorème.

Passons maintenant brièvement en revue les énoncés qui vous sont familiers depuis la terminale, et commençons par introduire les intégrales indéfinies. Soit f une fonction intégrable sur $[a, b]$; on sait, comme l'intégrale a la propriété de restriction aux sous-intervalles, que f est intégrable sur $[a, x]$ pour tout $x \in [a, b]$, et l'on peut donc définir

$$\psi_f(x) = \int_a^x f, \ x \in [a, b].$$

La relation de Chasles et le théorème fondamental impliquent immédiatement le théorème suivant, base du calcul intégral au lycée :

Théorème 6.2. — *Toute fonction continue* f *admet une primitive, égale à l'addition d'une constante près à* ψ_f*.*

On peut résumer tout celà de manière condensée (mais peu rigoureuse) :

– Si f est dérivable, on a

$$\int_a^x f'(t)\,\mathrm{d}t = f(x) + \text{constante}.$$

– Si f est continue,

$$\left(\int_a^x f(t)\,\mathrm{d}t\right)' = f(x).$$

Le théorème fondamental a pour corollaires immédiats des formes assez générales de formules d'intégration par partie et de changement de variables :

Théorème 6.3. — *Soient* F *et* G *deux fonctions dérivables sur* $[a, b]$. *Alors* $\mathrm{F}'\mathrm{G}$ *est intégrable si et seulement si* $\mathrm{F}\mathrm{G}'$ *l'est. Dans ce cas, on a la formule*

$$\int_a^b \mathrm{F}'\mathrm{G} = [\mathrm{F}\mathrm{G}]_a^b - \int_a^b \mathrm{F}\mathrm{G}'.$$

Démonstration. — Comme F et G sont dérivables, FG l'est aussi et

$$(\mathrm{F}\mathrm{G})' = \mathrm{F}'\mathrm{G} + \mathrm{F}\mathrm{G}'.$$

Le théorème fondamental dit par ailleurs que $(\mathrm{F}\mathrm{G})'$ est intégrable, et le théorème s'en déduit facilement.

Le théorème de changement de variable a une preuve tout aussi simple.

Théorème 6.4. — *Soient* g *une application dérivable sur l'intervalle compact* $[a, b]$, *et* f *une fonction admettant une primitive sur l'intervalle* $g([a, b])$. *Alors*

$$\int_{g(a)}^{g(b)} f(x)\,\mathrm{d}x = \int_a^b f(g(t))g'(t)\,\mathrm{d}t.$$

Il suffit en effet d'appliquer le théorème fondamental à $f \circ g$ et f. Il faut remarquer ici qu'on ne suppose pas que g est une bijection. Cependant, l'intervalle $g([a, b])$ pourrait être plus grand que $[g(a), g(b)]$, et l'hypothèse que f admet une primitive sur $g([a, b])$ est donc relativement forte.

Finissons ce chapitre par un théorème utile, la *formule de la moyenne* pour les intégrales. Ce résultat ne concerne pas les liens entre intégration et primitives, mais il est proche de la formule des accroissements finis, c'est pourquoi nous l'énonçons ici.

Théorème 6.5. — *Soient* f, g *deux fonctions définies sur* $[a, b]$. *On suppose* f, g *et* fg *intégrables,* g *positive, et* f *bornée. Alors*

$$\inf\{f(x), x \in [a, b]\} \int_a^b g(x)\,\mathrm{d}x \le \int_a^b f(x)g(x)\,\mathrm{d}x \le \sup\{f(x), x \in [a, b]\} \int_a^b g(x)\,\mathrm{d}x.$$

Si de plus f *est continue, alors il existe un point* $c \in [a, b]$ *tel que*

$$\int_a^b f(x)g(x)\,\mathrm{d}x = f(c) \int_a^b g(x)\,\mathrm{d}x.$$

Remarquons, avant de donner la démonstration, que l'hypothèse d'intégrabilité faite sur le produit fg est indispensable : il existe des fonctions f, g intégrables telles que fg ne le soit pas.

DÉMONSTRATION. — Soit m (*resp.* M) la borne inférieure (*resp.* supérieure) de f sur $[a, b]$. On a, pour tout $x \in [a, b]$, l'encadrement

$$mg(x) \leq f(x)g(x) \leq \mathrm{M}g(x),$$

et la première partie du théorème découle de la positivité de l'intégrale. Quant à la deuxième, elle découle du théorème des valeurs intermédiaires.

Corollaire 6.1. — *Si f est continue sur $[a, b]$, alors il existe $c \in]a, b[$ telle que*

$$\int_a^b f(x)\,\mathrm{d}x = (b-a)f(c).$$

7 Quelques méthodes pratiques de calcul d'intégrales et de primitives

Le théorème fondamental et ses corollaires donne la manière la plus efficace pour calculer explicitement des intégrales et des primitives de fonctions construites à partir de fonctions usuelles. Nous allons, dans cette section, rappeler les primitives usuelles, puis donner quelques exemples de méthode de calcul d'intégrales. Les travaux dirigés complèteront ce cours avec une grand nombre de calculs explicites.

7.1 Primitives usuelles

La liste suivante devrait être connue par cœur par tout étudiant sérieux. Pour fixer les notations, nous noterons la primitive F d'une fonction continue f, là où elle est définie, sous la forme
$$\mathrm{F} = \int f(x)\,\mathrm{d}x + \mathrm{C},$$
où C désigne une constante quelconque.

$$\int x^\alpha\,\mathrm{d}x = \frac{1}{\alpha+1}x^{\alpha+1} + \mathrm{C},\ \alpha \neq -1 \qquad \int \frac{1}{x}\,\mathrm{d}x = \mathrm{Log}\,|x| + \mathrm{C}$$
$$\int \mathrm{e}^{(a+ib)x}\,\mathrm{d}x = \frac{1}{a+ib}\mathrm{e}^{(a+ib)x} + \mathrm{C},\ a+ib \neq 0 \qquad \int \sin x\,\mathrm{d}x = -\cos x + \mathrm{C}$$
$$\int \cos x\,\mathrm{d}x = \sin x + \mathrm{C} \qquad \int \tan x\,\mathrm{d}x = -\mathrm{Log}\,|\cos x| + \mathrm{C}$$
$$\int \frac{1}{\sin x}\,\mathrm{d}x = \mathrm{Log}\left|\tan\frac{x}{2}\right| + \mathrm{C} \qquad \int \frac{1}{\cos x}\,\mathrm{d}x = \mathrm{Log}\left|\tan\left(\frac{\pi}{4}+\frac{x}{2}\right)\right| + \mathrm{C}$$
$$\int \mathrm{cotan}\,x\,\mathrm{d}x = \mathrm{Log}\,|sinx|| + \mathrm{C} \qquad \int \frac{1}{\cos^2 x}\,\mathrm{d}x = \tan x + \mathrm{C}$$
$$\int \frac{1}{\sin^2 x} = -\mathrm{cotan}\,x + \mathrm{C} \qquad \int \frac{1}{\sin x\cos x}\,\mathrm{d}x = \mathrm{Log}\,|\tan x| + \mathrm{C}$$
$$\int \mathrm{sh}\,x\,\mathrm{d}x = \mathrm{ch}\,x + \mathrm{C} \qquad \int \mathrm{ch}\,x\,\mathrm{d}x = \mathrm{sh}\,x + \mathrm{C}$$
$$\int \mathrm{th}\,x\,\mathrm{d}x = \mathrm{Log}\,\mathrm{ch}\,x + \mathrm{C} \qquad \int \frac{1}{\mathrm{sh}\,x}\,\mathrm{d}x = \mathrm{Log}\left|\mathrm{th}\,\frac{x}{2}\right| + \mathrm{C}$$
$$\int \frac{1}{\mathrm{ch}\,x}\,\mathrm{d}x = 2\arctan(\mathrm{e}^x) + \mathrm{C} \qquad \int \frac{1}{\mathrm{th}\,x}\,\mathrm{d}x = \mathrm{Log}\,|\mathrm{sh}\,x| + \mathrm{C}$$
$$\int \frac{1}{\mathrm{ch}^2 x}\,\mathrm{d}x = \mathrm{th}\,x + \mathrm{C} \qquad \int \frac{1}{\mathrm{sh}^2 x}\,\mathrm{d}x = -\coth x + \mathrm{C}$$
$$\int \frac{1}{\mathrm{sh}\,x\,\mathrm{ch}\,x}\,\mathrm{d}x = \mathrm{Log}\,|\mathrm{th}\,x| + \mathrm{C} \qquad \int \frac{1}{x^2+a^2}\,\mathrm{d}x = \frac{1}{a}\arctan\frac{x}{a} + \mathrm{C}$$
$$\int \frac{1}{x^2-a^2}\,\mathrm{d}x = \frac{1}{2a}\mathrm{Log}\left|\frac{x-a}{x+a}\right| + \mathrm{C} \qquad \int \frac{1}{\sqrt{a^2-x^2}}\,\mathrm{d}x = \arcsin\frac{x}{|a|} + \mathrm{C}$$
$$\int \frac{1}{\sqrt{x^2+a^2}}\,\mathrm{d}x = \mathrm{Log}(x+\sqrt{x^2+a^2}) + \mathrm{C} \qquad \int \frac{1}{\sqrt{x^2-a^2}}\,\mathrm{d}x = \mathrm{Log}\,|x+\sqrt{x^2-a^2}| + \mathrm{C}$$

$$\int \frac{1}{(x^2+a)^{3/2}}\,\mathrm{d}x = \frac{x}{a\sqrt{x^2+a}} + \mathrm{C} \qquad \int \frac{1}{(a-x^2)^{3/2}}\,\mathrm{d}x = \frac{x}{a\sqrt{a-x^2}} + \mathrm{C},\ a \neq 0$$

Ces formules se déduisent facilement des formules de trigonométries (usuelle ou hyperbolique) ainsi que des formules donnant les dérivées des fonctions usuelles. De manière générale, le calcul de primitives ou d'intégrales, hormis quelques méthodes connues pour des cas particuliers, demande du nez, et donc beaucoup d'entrainement. Il est à noter qu'il n'existe pas d'algorithme général permettant de donner la primitive explicite d'une fonction définie par des fonctions usuelles. Il existe même des fonctions très simples (comme $\exp(-x^2)$) dont la primitive n'admet pas d'expression algébrique en termes de fonction usuelles. C'est pourquoi, quand on rencontre un problème de calcul explicite d'intégrale, on fait souvent appel, suivant le type d'étude entreprise, aux méthodes numériques, ou bien à des méthodes très avancées d'algèbre, de géométrie ou d'analyse.

7.2 Intégration des fractions rationnelles

Vous verrez l'an prochain, dans l'uv intitulée « structures algébriques », la théorie générales des polynômes et des fractions rationnelles. Pour nous, une fraction rationnelle sera simplement une fonction définie comme le quotient P/Q de deux polynômes P et Q à coefficients complexes. Cette fonction est définie sur **C** privé des racines de Q, que l'on appelle les *pôles* de la fraction rationnelle P/Q. Nous admettrons le théorème fondamental suivant, conséquence du fait que **C** est *algébriquement clos :* tout polynôme non constant à coefficients complexes admet au moins un racine dans **C**.

Théorème 7.1. — *Soient* P *et* Q *deux polynômes à coefficients complexes sans racine commune. Soient* $a_1, \ldots, a_r$ *les pôles de* P/Q. *Alors il existe un polynôme* E, *des nombres entiers* $\alpha_i \in \mathbf{N}^*, 1 \leq i \leq r$ *(*α_i *est appelé multiplicité du pôle* a_i*), et des nombres complexes* $\mathrm{A}_{j,k}, 1 \leq k \leq r, 1 \leq j \leq \alpha_k$, *tels que*

$$\frac{\mathrm{P}(x)}{\mathrm{Q}(x)} = \mathrm{E}(x) + \sum_{k=1}^{r}\sum_{j=1}^{\alpha_k} \frac{\mathrm{A}_{j,k}}{(x-a_k)^j}.$$

La décomposition ci-dessus s'appelle *décomposition de* P/Q *en éléments simples.* Le problème du calcul pratique de cette décomposition est délicat, et nous ne traîterons ici que deux exemples simples. D'autres exemples seront faits en travaux dirigés, et vous verrez qu'ils peuvent donner lieu à des calculs pénibles. Avant de passer aux exemples, notons que comme il n'existe pas d'algorithme permettant de calculer les racines d'un polynôme général, le théorème ci-dessus est d'une portée limitée dans le calcul explicite de primitives ou d'intégrales. Par contre, nous verrons aussi que l'on ramène souvent un calcul d'intégrale, *via* des changements de variables ou tout autre méthode, à un calcul d'intégrale de fraction rationnelle.

Premier exemple. Soit à calculer l'intérgrale suivante, pour $1 < a < x$:

$$\mathrm{J} = \int_a^x \frac{1}{t^3(1+t^3)}\,\mathrm{d}t.$$

Il faut commencer par vérifier que la fonction sous le signe intégrale est bien intégrable. Pour celà, remarquons que le dénominateur ne s'annule qu'en 0 et -1, donc la fonction à intégrer est intégrable sur $[a, x] \subset \mathbf{R} - \{0, 1\}$, car continue. Avant de partir tête baissée dans le calcul du développement en éléments simples de la fonction, cherchons à la simplifier. Ainsi, on remarque que

$$\frac{1}{t^3(1+t^3)} = \frac{1}{t^3} - \frac{1}{1+t^3}.$$

Le premier terme du second membre est facile à intégrer, si l'on se rappelle de la règle de dérivation des puissances :

$$\int_a^x \frac{1}{t^3}\,\mathrm{d}t = \left[-\frac{1}{2}\frac{1}{t^2}\right]_a^x,$$

où la notation $[f(t)]_a^b$ signifie $f(b) - f(a)$, comme à l'acoutumée. Il nous reste donc à calculer l'intégrale du second terme. Remarquons que $1 + t^3$ a pour racines, dans $\mathbf{C}$, $-1, \mathrm{e}^{i\pi}$ et $\mathrm{e}^{-i\pi}$. Ainsi

$$1 + t^3 = (t+1)(t - \mathrm{e}^{i\pi/3})(t - \mathrm{e}^{-i\pi/3}).$$

Les pôles de $(1+t^3)^{-1}$ sont de multiplicité 1 (ce qui se voit ici en écrivant $1+t^3 = (t+1)^1(t+\mathrm{e}^{i\pi})^1(t+\mathrm{e}^{-i\pi})^1$, et l'on parle alors de pôles *simples*), et donc la décomposition en éléments simples prend la forme

$$\frac{1}{1+t^3} = \frac{a}{t+1} + \frac{b}{t - \mathrm{e}^{i\pi/3}} + \frac{c}{t - \mathrm{e}^{-i\pi/3}}.$$

On détermine les trois nombres complexes a, b, c en multipliant successivement l'identité ci-dessus par les monômes $t + 1$, $t - \mathrm{e}^{i\pi/3}$ et $t - \mathrm{e}^{-i\pi/3}$, et en faisant tendre la variable vers chacune des racines correspondantes de $1 + t^3$. Commençons par déterminer a. L'identité ci-dessus implique en particulier que

$$\frac{t+1}{1+t^3} = \frac{1}{(t - \mathrm{e}^{i\pi/3})(t - \mathrm{e}^{-i\pi/3})} = a + (t+1)\left(\frac{b}{t - \mathrm{e}^{i\pi/3}} + \frac{c}{t - \mathrm{e}^{-i\pi/3}}\right),$$

et l'on peut faire tendre t vers -1 dans cette dernière égalité, pour obtenir

$$\frac{1}{(-1 - \mathrm{e}^{i\pi/3})(-1 - \mathrm{e}^{-i\pi/3})} = a,$$

soit $a = 1/3$. De même, l'identité plus haut, multipliée par $t - \mathrm{e}^{i\pi/3}$ donne

$$\frac{1}{(t+1)(t - \mathrm{e}^{-i\pi/3})} = b + (t - \mathrm{e}^{i\pi/3})\left(\frac{a}{t+1} + \frac{b}{t - \mathrm{e}^{-i\pi/3})}\right),$$

et, si l'on fait tendre t vers $\mathrm{e}^{i\pi/3}$, on trouve enfin

$$\frac{1}{(1 + \mathrm{e}^{i\pi/3})(\mathrm{e}^{i\pi/3} - \mathrm{e}^{-i\pi/3})} = b,$$

soit

$$b = \frac{1}{i\sqrt{3}(1+\mathrm{e}^{i\pi/3})}.$$

On calcule de même c, et l'on trouve finalement

$$\frac{1}{1+t^3} = \frac{1/3}{1+t} + \frac{1}{i\sqrt{3}(1+\mathrm{e}^{i\pi/3})}\frac{1}{t-\mathrm{e}^{i\pi/3})} - \frac{1}{i\sqrt{3}(1+\mathrm{e}^{-i\pi/3})}\frac{1}{t-\mathrm{e}^{-i\pi/3})}.$$

Le premier terme est facile à intégrer, mais les deux suivants, qui semblent tout aussi facile, comportent des nombres complexes. Remarquons cependant que les deux derniers termes sont *conjugués complexes.* Ceci est un fait général, dû au caractère réel du polynôme $1+t^3$: les racines complexes d'un polynômes réel sont deux-à-deux conjuguées. Il en découle que dans la décomposition en éléments simples, on peut ajouter deux-à-deux les termes correspondants à un couple de pôles conjugués, et obtenir au final une expression réelle ! Additionons-donc les deux derniers termes de l'identité ci-dessus :

$$\frac{1}{i\sqrt{3}(1+\mathrm{e}^{i\pi/3})}\frac{1}{t-\mathrm{e}^{i\pi/3})} - \frac{1}{i\sqrt{3}(1+\mathrm{e}^{-i\pi/3})}\frac{1}{t-\mathrm{e}^{-i\pi/3})} = \frac{1}{3}\frac{2-t}{t^2-t+1}.$$

On peut donc déjà écrire J sous la forme :

$$\mathrm{J} = \int_a^x \frac{1}{t^3(1+t^3)}\,\mathrm{d}t = \int_a^x \frac{1}{t^3}\,\mathrm{d}t - \frac{1}{3}\int_a^x \frac{1}{1+t}\,\mathrm{d}t - \frac{1}{3}\int_a^x \frac{2-t}{t^2-t+1}\,\mathrm{d}t.$$

Les deux premiers termes du second membre s'intègre facilement. Nous allons voir qu'il en est de même du troisième. Pour celà, la première chose à faire est de faire apparaître la dérivée du dénominateur au numérateur, et de compenser habilement le reste pour obtenir la somme d'une dérivée logarithmique et d'un e dérivée de fonction circulaire inverse (ici l'arc-tangente). Voilà comment l'on procède : remarquons tout d'abord que

$$\frac{\mathrm{d}}{\mathrm{d}t}(t^2-t+1) = 2t-1.$$

Or on ne dispose au numérateur que de $-t = -1/2(2t)$. Faisons donc apparaître $2t-1$ au numérateur :

$$\begin{aligned}
\frac{2-t}{t^2-t+1} &= \frac{-1/2(2t)+2}{t^2-t+1} \\
&= \frac{-1/2(2t)-1/2(-1)+2+1/2(-1)}{t^2-t+1} \\
&= -\frac{1}{2}2t-1t^2-t+1+\frac{3/2}{1}t^2-t+1.
\end{aligned}$$

Le premier terme du dernier membre s'intègre aisément, car on y reconnait une dérivée logarithmique. Pour le second terme, de la forme

$$\frac{1}{t^2+\alpha t+\beta},$$

on s'arrange pour le mettre sous la forme

$$\frac{1}{(t+\eta)^2+1}.$$

Remarquons que t^2-t est le début du développement de $(t-1/2)^2$. Ainsi

$$t^2-t+1=(t-1/2)^2+3/4,$$

et ainsi

$$\frac{2-t}{t^2-t+1}=-\frac{1}{2}\frac{2t-1}{t^2-t+1}+\frac{2}{\frac{3}{4}\left(t-\frac{1}{2}\right)^2+1}.$$

On peut alors se reporter à la liste des intégrales usuelles, et obtenir :

$$\mathrm{J}=\left[-\frac{1}{2}\frac{1}{t^2}-\frac{1}{3}\operatorname{Log}|1+t|+\frac{1}{6}\operatorname{Log}|t^2-t+1|-\frac{\sqrt{3}}{3}\arctan\frac{2}{\sqrt{3}}\left(t-\frac{1}{2}\right)\right]_a^x.$$

Ouf ! Comme on peut le voir, les calculs sont fastidieux, mais, dès que l'on connait les pôles de la fraction rationnelle (*i.e.* les racines du dénominateur), ils deviennent automatiques, et l'on pourrait écrire un programme permettant de calculer l'intégrale.

Exercice vache. En utilisant les formules donnant les racines des polynômes de degrés $1,2,3$ et 4, écrire un programme informatique donnant la primitive d'une fraction rationnelle de degré au plus 4.

Deuxième exemple. Essayons maintenant de calculer, avec les mêmes hypothèses sur a et x que plus haut, l'intégrale

$$\mathrm{J}=\int_a^x\frac{3t^3+10t^2-2t}{(t^2-1)^2}\,\mathrm{d}t=\int_a^x\frac{\mathrm{P}(t)}{\mathrm{Q}(t)}\,\mathrm{d}t.$$

Remarquons tout d'abord que Q est de degré plus grand que P. Cela implique que le quotient de P par Q dans la division euclidienne des polynômes est 0. C'est la partie entière de P/Q, notée E dans le théorème de décomposition en éléments simples. Celui-ci nous indique donc qu'il existe des nombres complexes a,b,c,d tels que

$$\frac{\mathrm{P}(t)}{\mathrm{Q}(t)}=\frac{a}{(t-1)^2}+\frac{b}{t-1}+\frac{c}{(t+1)^2}+\frac{d}{t+1},$$

puisque $(t^2-1)^2$ admet -1 et 1 pour racines doubles (*i.e.* -1 et $+1$ sont des pôles d'ordre 2 de P/Q). Remarquons que ni -1 ni 1 ne sont des racines de P. On calcule a en multipliant l'identité ci-dessus par $(t-1)^2$, ce qui donne

$$\frac{3t^3+10t-2t}{(t+1)^2}=a+(t-1)\left[b+\frac{c(t-1)}{(t+1)^2}+\frac{d(t-1)}{t+1}\right].$$

En faisant $t=1<$ on obtient $a=11/4$. De même, on obtient c en multipliant par $(t+1)^2$ et en faisant $t=-1$ dans l'égalité obtenue : $c=9/4$. Pour trouver encuite

b et d, on affecte à t diverses valeurs, comme par exemple $t = 0$, ce qui donne $a - b + c + d = 0$. On peut aussi multiplier l'identité plus haut par t, et faire tendre t vers $+\infty$, ce qui donne $b + d = 3$. Toute autre cuisine est bien-sûre possible, et au final on obtient un système de deux équations linéaires dont b et d sont solutions :

$$b - d = 5 \text{ et } b + d = 3,$$

ce qui donne $b = 4$ et $d = -1$. On peut donc maintenant calculer l'intégrale J :

$$\begin{aligned} \mathrm{J} &= \frac{11}{4}\int_a^x \frac{1}{(t-1)^2}\,\mathrm{d}t + \frac{9}{4}\int_a^x \frac{1}{(t+1)^2}\,\mathrm{d}t + 4\int_a^x \frac{1}{t-1}\,\mathrm{d}t - \int_a^x \frac{1}{t+1}\,\mathrm{d}t \\ &= \left[-\frac{1}{2}\frac{10t+1}{t^2-1} + \mathrm{Log}\,\frac{(t-1)^4}{|t+1|}\right]_a^x. \end{aligned}$$

7.3 Applications de la formule d'intégration par parties

Premier exemple. Soit n un entier naturel, posons

$$\mathrm{I}_n = \int_0^{\pi/2} \sin^n t\,\mathrm{d}t,$$

intégrale portant le nom du mathématicien anglais John Wallis (1616–1703). Pour calculer une telle intégrale, une méthode courante consiste à trouver, par une ou plusieurs intégrations par parties, une relation de récurrence satisfaite par la suite (I_n). Si $n \geq 2$, on a $\sin^n t = \sin^{n-1} t \sin t$, et l'on peut appliquer la formule d'intégration par parties à cette décomposition :

$$\begin{aligned} \mathrm{I}_n &= \int_0^{\pi/2} \sin^n t\,\mathrm{d}t = \int_0^{\pi/2} \sin t \sin^{n-1} t\,\mathrm{d}t \\ &= \left[-\cos t \sin^{n-1} t\right]_0^{\pi/2} - \int_0^{\pi/2} (-\cos t)(n-1)\cos t \sin^{n-2} t\,\mathrm{d}t \\ &= (n-1)\int_0^{\pi/2} (1 - \sin^2 t)\sin^{n-2} t\,\mathrm{d}t \\ &= (n-1)\mathrm{I}_{n-2} - (n-1)\mathrm{I}_n, \end{aligned}$$

ce qui donne, pour tout $n \geq 2$:

$$\mathrm{I}_n = \frac{n-1}{n}\mathrm{I}_{n-2}.$$

Deux cas se présentent alors : si $n = 2p$ est pair, on obtient

$$\begin{aligned} \mathrm{I}_{2p} &= \frac{(2p-1)(2p-2)\dots 1}{(2p)(2p-2)\dots 2}\mathrm{I}_0 \\ &= \frac{\pi}{2}\frac{(2p)!}{(2^p p!)^2}. \end{aligned}$$

De même, si $n = 2p + 1$ est impair, on obtient

$$\mathrm{I}_{2p+1} = \frac{(2^p p!)^2}{(2p+1)!}.$$

Deuxième exemple. Soit $n \geq 1$ un entier naturel, on cherche à calculer une primitive F_n de $f_n(x) = x^n \cos x$. La fonction f est déjà écrite comme produit de deux fonctions, qu'en plus l'on sait intégrer. Intégrons par parties :

$$\begin{aligned}
\int x^n \cos x \, \mathrm{d}x &= [x^n \sin x] - \int n x^{n-1} \sin x \, \mathrm{d}x \\
&= x^n \sin x - n \left([x^{n-1}(-\cos x)] - \int (n-1) x^{n-2} (-\cos x) \, \mathrm{d}x \right) \\
&= x^n \sin x + n x^{n-1} \cos x - n(n-1) \mathrm{F}_{n-2}(x),
\end{aligned}$$

relation de récurrence qui permet de calculer F_n pour tout n.

7.4 Applications de la formule de changement de variable

Commençons par rappeler un moyen mnémotechnique utile dans l'utilisation de la formule de changement de variable. Si l'on pose $x = \varphi(t)$, alors on obtient $\mathrm{d}x$ par la formule $\mathrm{d}x = \varphi'(t) \, \mathrm{d}t$.

Un premier exemple amusant. Essayons de déterminer une primitive de

$$\frac{1}{\sqrt{1+x} + \sqrt[3]{1+x}}.$$

On remarque que l'on a les mêmes quantités sous les radicaux. On va donc essayer de s'en débérasser en exprimant ces deux radicaux comme puissances d'une même quantité. Le ppcm de 2 (pour la racine carrée) et 3 est 6. Ainsi on a

$$(\sqrt[6]{1+x})^2 = \sqrt[3]{1+x} \text{ et } (\sqrt[6]{1+x})^3 = \sqrt{1+x}.$$

On s'empresse donc de faire le changement de variable $y = \sqrt[6]{1+x} = (1+x)^{1/6}$. Comme $y^6 = 1 + x$, il vient $\mathrm{d}x = 6y^5 \, \mathrm{d}y$, et donc

$$\int \frac{1}{\sqrt{1+x} + \sqrt[3]{1+x}} \, \mathrm{d}x = \int \frac{6y^5}{y^3 + y^2} \, \mathrm{d}y,$$

et cette dernière intégrale se calcule facilement. On commence par effectuer une division euclidienne de y^5 par $y^3 + y^2$, ce qui donne

$$y^5 = (y^2 - y + 1)(y^3 + y^2) - y^2,$$

soit

$$\frac{y^5}{y^3 + y^2} = y^2 - y + 1 - \frac{1}{y+1},$$

et il sort alors de tout ça le merveilleux résultat suivant :

$$\int \frac{1}{\sqrt{1+x}+\sqrt[3]{1+x}}\,\mathrm{d}x = 2\sqrt{1+x} - 3\sqrt[3]{1+x} + 6\sqrt[6]{1+x} - 6\,\mathrm{Log}\,|1+\sqrt[6]{1+x}| + \mathrm{C}.$$

Il ne faut pas oublier, à la fin du calcul, de revenir à la variable x en remplaçant partout y par sa valeur.

Deuxième exemple. Cet exemple donne en fait une méthode générale de calcul des primitives de fractions rationnelles en *sinus* et *cosinus*. Soient P et Q deux polynômes de deux variables, à coefficients réels. On cherche à intégrer une fonction de la forme $\mathrm{P}(\cos x, \sin x)/\mathrm{Q}(\cos x, \sin x)$, et pour celà on rappelle les formules de trigonométrie suivantes : si l'on pose $t = \tan(x/2)$, alors

$$\sin x = \frac{2t}{1+t^2} \text{ et } \cos x = \frac{1-t^2}{1+t^2}.$$

On obtient alors facilement le $\mathrm{d}x$, qui est donné par

$$\mathrm{d}x = \frac{2\,\mathrm{d}t}{1+t^2},$$

et l'on en déduit que

$$\int \frac{\mathrm{P}(\cos x, \sin x)}{\mathrm{Q}(\cos x, \sin x)}\,\mathrm{d}x = \int \frac{\mathrm{P}((1-t^2)(1+t^2)^{-1}, (2t)(1+t^2)^{-1})}{\mathrm{Q}((1-t^2)(1+t^2)^{-1}, (2t)(1+t^2)^{-1})}\,\frac{2\,\mathrm{d}t}{1+t^2},$$

qui est une intégrale de fraction rationnelle en t, et que l'on est donc en mesure, en théorie, de calculer.

Exercice. Faire la même chose que ci-dessus pour une fraction rationnelle en ch et sh.

8 Intégrales impropres

Intégrer des fonctions définies sur des intervalles compacts s'avère très vite insuffisant, tant en mathématiques (arithmétique, analyse...) qu'en physique. On y est en effet confronté à des intégrales du type

$$\int_0^1 \frac{\mathrm{d}x}{\sqrt{x}} \text{ ou } \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x, \text{ voir } \int_0^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x.$$

Or la première fonction intégrée n'est pas définie en 0, ni bornée sur $[0,1]$, tandis que les deux autres intégrale font intervenir un intervalle non borné. Nous allons voir, dans cette section, que l'on peut étendre la définition de l'intégrale de manière à prendre en compte ces deux situtations. De plus, nous énoncerons un théorème dû à Heinrich Hake, qui dit que l'intégrale de Kurzweil et Henstock contient ses intégrales impropres (*i.e.* obtenues par passage à la limite dans les bornes).

Nous ne donnerons que peu de preuves ici, et nous bornerons le plus souvent à donner les énoncés des résultats les plus importants. En fait, tout marche comme dans le cas de l'intégration des fonctions définies sur un compact, quasiment mot pour mot. Enfin, nous nous limiterons à un intervalle de la forme $[a,+\infty[$, les autres cas s'en déduisant aisément. Les travaux dirrigés permettront de travailler plus en détail sur des exemples explicites.

8.1 Intégrabilité et intégrale sur $[a,+\infty[$

Soit f une fonction définie sur l'intervalle $[a,+\infty[$. On aimerait pouvoir étendre la définition de l'intégrabilité de f (et de son intégrale, si elle est intégrable) en utilisant des sommes de Riemann et des subdivisions pointées. L'idée la plus simple consiste à *tronquer* l'intervalle $[a,+\infty[$ à droite, de manière à se ramener à des sous-intervalles compacts. Pour cela, nous allons procéder par étapes.

– Commençons par *compactifier* $[a,+\infty[$ en lui rajoutant $+\infty$, c'est-à-dire en prenant son adhérence dans la droite achevée $\overline{\mathbf{R}}$. Ainsi, nous allons travailler sur $\mathrm{I} = [a,+\infty]$.
– Posons $f(+\infty) = 0$.
– Définissons maintenant les subdivisions pointées :

$$\mathcal{P} = \{([a = x_0, x_1], t_1), ([x_1, x_2], t_2), \ldots, ([x_{n-1}, x_n = +\infty], t_n)\}.$$

– Il nous faut maintenant pour voir définir la somme de Riemann de f associée à $\mathcal{P}$. L'intervalle non-borné $[x_{n-1},+\infty]$ pose problème. Pour celà, nous allons étendre f à $[a,+\infty]$ en posant $f(+\infty) = 0$. Ensuite, nous allons étendre la notion de finesse de manière à ce que le dernier point de marquage t_n soit toujours égal à $+\infty$. De cette manière, le dernier terme de la somme de Riemann disparait, et celle-ci sécrit :

$$\mathrm{S}(f,\mathcal{P}) = \sum_{k=1}^{n-1} (x_k - x_{k-1}) f(t_k).$$

- Une jauge sur I est toujours une fonction $\delta : \mathrm{I} \longrightarrow]0, +\infty[$.
- La subdivision $\mathcal{P}$ est dite δ-fine si, pour tout $k = 1, \ldots, n-1$ on a

  $$[x_{k-1}, x_k] \subset [t_k - \delta(t_k)/2, t_k + \delta(t_k)/2],$$

  et, pour le dernier intervalle,

  $$x_{n-1} > 1/\delta(+\infty).$$

  Ainsi, plus l'on prend $\delta(+\infty)$ petit, plus x_{n-1} doit être grand. Remarquons que, comme $[x_{n-1}, +\infty]$ est le seul intervalle de la subdivision à contenir $+\infty$, cette définition de finesse impose forcément que $t_n = +\infty$.

Munis de ces notations et conventions, il semble naturel d'adopter la

Définition 8.1. — *La fonction $f : [a, +\infty[\longrightarrow \mathbf{R}$ est intégrable si, et seulement si, il existe un nombre $\mathrm{S} \in \mathbf{R}$ tel que, pour tout $\varepsilon > 0$, il existe un jauge δ sur $[a, +\infty]$ tel que, pour toute subdivision δ-fine $\mathcal{P}$, on ait*

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}| < \varepsilon.$$

Si f est intégrable, le nombre S ci-dessus est appelé l'intégrale de f sur $[a, +\infty[$ et est noté

$$\int_{[a,+\infty[} f, \text{ ou } \int_a^{+\infty} f(x)\,\mathrm{d}x.$$

On dit enfin que f est Lebesgue-intégrable sur $[a, +\infty[$ si f et $|f|$ sont intégrables sur $[a, +\infty[$.

L'intégrale sur $[a, +\infty[$ jouit de propriétés analogues à celles sur un intervalle compact. Bien-sûr, du fait que l'intervalle est non-borné (ou que la fonction est non bornée sur un intervalle borné), déterminer si la fonction est intégrable peut s'avérer problématique. Voici quelques exemples de fonctions intégrables :

$$\frac{1}{\sqrt{t}} \text{ sur } [0, 1], \quad \frac{1}{x^2} \text{ sur } [1, +\infty], \quad \frac{x \sin x}{1 + x^2} \text{ sur } [0, +\infty].$$

Avant de voir comment les théorèmes vrais pour l'intégrale sur un intervalle compact se généralisent au cas étudié ici, énonçons le théorème de Hake.

Théorème 8.1. — *La fonction $f : [a, +\infty[\longrightarrow \mathbf{R}$ est intégrable si, et seulement si, pour tout $c \in [a, +\infty[$, f est intégrable sur $[a, c]$, et si la limite*

$$\lim_{c \to +\infty} \int_a^c f(x)\,\mathrm{d}x = \mathrm{L}$$

existe. On a alors

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \mathrm{L}.$$

Ce théorème est admis ici, même si sa preuve, un peu délicate, est tout-à-fait accessible : elle est donnée en appendice. Il est très utile dans les estimations ou les calculs explicites d'intégrales. Ainsi, combiné au théorème d'intégration par parties, il permet de montrer que $x^{-1}\sin x$ est intégrable sur $[0,+\infty[$.
Remarque : l'intégrale de Riemmann proprement dite ne permet pas de traîter les intervalles non compacts ni les fonctions non bornées, et l'intégrale de Lebesgue ne possède pas d'intégrale semi convergente comme

$$\int_0^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x.$$

L'intégrale de Kurzweil-Henstock développée dans ce cours est à ce propos supérieure.

8.2 Extension de quelques théorèmes aux intégrales impropres

Nous revenons ici sur quelques résultats démontrés dans le cas d'une fonction définie sur un intervalle compact, et les étendons au cas non compact. Les démonstrations ne sont pas données, mais, soit elles se calquent sur celles déjà faites, soit elles sont une application facile du théorème de Hake.

La première remarque à faire est que, comme dans le cas compact, le nombre S intervenant dans la définition de l'intégrabilité est, s'il existe, unique. Le critère de Cauchy 4.1, ainsi que le théorème d'encadrement 4.2 passent sans changement. De même, l'intégrale est linéaire, monotone et vérifie la relation de Chasle. Celle-ci permet de définir des intégrales entre moins l'infini et plus l'infini, par exemple :

Proposition 8.1. — *La fonction $f:\mathbf{R}\longrightarrow\mathbf{R}$ est intégrable si, et seulement si, la limite*

$$\lim_{a\to-\infty,b\to+\infty}\int_a^b f(x)\,\mathrm{d}x=\mathrm{L}$$

existe. Dans ce cas, $f\in\mathcal{I}(\mathbf{R})$ et

$$\mathrm{L}=\int_{\mathbf{R}} f(x)\,\mathrm{d}x=\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x.$$

Toutes les variations sur ce thème sont bien sûr valables.

De même, on peut baser les intégrales indéfinies en une borne infinie de l'intervalle de définition, et le théorème 6.2 est vrai avec $a=-\infty$. La formule de la moyenne 6.5 est valable avec $b=+\infty$. Enfin, les autres résultats de la section « Intégrales et primitives »se généralisent, avec quelques modifications, au cas de bornes infinies. On peut ainsi énoncer la formule d'intégration par partie pour des fonctions définies et dérivables sur $[a,^{\infty}[$ de la manière suivante : soient F,G deux fonctions dérivables sur $[a,+\infty[$, et telles que $\mathrm{F}(x)\mathrm{G}(x)$ admette une limite finie quand x tend vers $+\infty$. Si FG' est intégrable sur $[a,+\infty[$, alors $\mathrm{F}'\mathrm{G}$ l'est et

$$\int_a^{+\infty}\mathrm{F}'(x)\mathrm{G}(x)\,\mathrm{d}x=\lim_{x\to+\infty}[\mathrm{FG}]_a^x-\int_a^{+\infty}\mathrm{F}(x)\mathrm{G}'(x)\,\mathrm{d}x.$$

Pour finir, notons le fait commode suivant : soit f une fonction définie sur un intervalle (borné ou non) $[a,b]$. On peut prolonger f à $\mathbf{R}$ tout entier en posant $\bar{f}(x)=0$ pour tout $x \notin [a,b]$ et $\bar{f}(x)=f(x)$ pour $x\in[a,b]$. On a alors, du fait de la relation

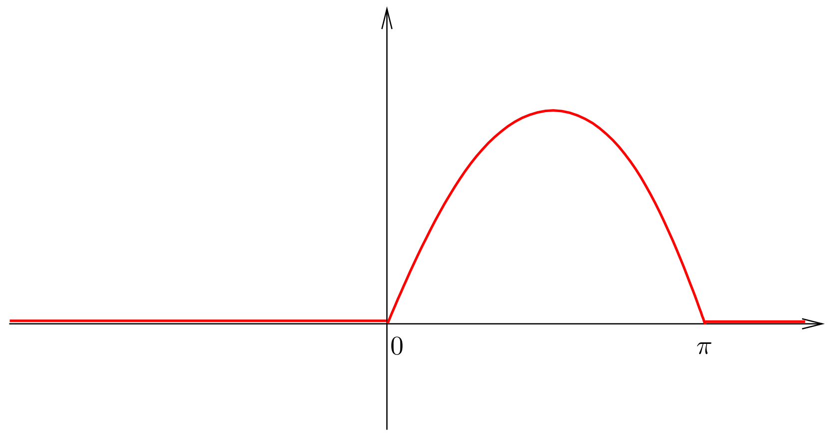

Fig. 7: Prolongement de sin sur $[0,\pi]$ à $\mathbf{R}$ entier.

de Chasles, le fait suivant :
Fait. La fonction f est intégrable sur $[a,b]$ si, et seulement si $\bar{f}$ est intégrable sur $\mathbf{R}$. La même chose est vraie avec l'intégrabilité au sens de Lebesgue.

8.3 Critères de convergence

Commençons par introduire la terminologie historique (et encore employée de nos jours) concernant les intégrales sur des intervalles non compacts.

Soit f une fonction définie sur l'intervalle $[a,+\infty[$. On dit que f est *localement intégrable* sur l'intervalle $[a,+\infty[$ si elle est intégrable sur tout intervalle *compact* contenu dans $[a,+\infty[$. Cela revient à dire, d'après la relation de Chasles, que pour tout $\mathrm{X}\in[a,+\infty[$, f est intégrable sur $[a,\mathrm{X}]$.

Définition 8.2. — *Soit f une fonction localement intégrable sur $[a,+\infty[$. L'intégrale*

$$\mathrm{I}=\int_a^{+\infty} f(x)\,\mathrm{d}x$$

est dite convergente *si f est intégrable sur $[a,+\infty[$. Elle est dite* absolument convergente *si f et $|f|$ sont intégrables sur $[a,+\infty[$,* i.e. *si f est Lebesgue-intégrable sur $[a,+\infty[$. On dit qu'elle est* semi-convergente *si elle est convergente mais pas absolument convergente. Si elle n'est pas convergente, on dit qu'elle est* divergente. *Le cas des autres types d'intervalle donne lieu à une terminologie analogue.*

Maintenant, on aimerait disposer de critères assurant la convergence d'une intégrale impropre. Il s'agit d'appliquer le théorème de comparaison 4.2, valable aussi bien dans le cas d'intervalles compacts que non compacts. Typiquement, le problème se pose en ces termes : « l'intégrale

$$\mathrm{I} = \int_a^{+\infty} f(x)\,\mathrm{d}x$$

est-elle *convergente ?* »On ne sait *a priori* répondre à cette question que pour une toute petite classe de fonctions. L'idée est alors, dans un cas plus général, de *majorer* la valeur absolue de $f(x)$ pour x suffisament grand, par une de ces fonctions dont on sait qu'elle admet une intégrale convergente en $+\infty$.

Proposition 8.2 (Critère de convergence en un point infini). — *Soit f une fonction localement intégrable sur l'intervalle $[a, +\infty[$.*

1. *Si $|f(x)| \leq x^{-\alpha}$ pour un certain $\alpha > 1$ et pour tout x assez grand, alors* I *est absolument convergente.*
2. *Si $|f(x)| \leq \exp(-\eta x)$ pour un certain $\eta > 0$ et pour tout x assez grand, alors* I *est absolument convergente.*
3. *Si $|f(x)| \geq x^{-\alpha}$ pour un certain $\alpha \leq 1$ et pour tout x assez grand, alors* I *est divergente.*

On peut énoncer des critères analogues pour la convergence (non nécessairement absolue), mais il faut pour celà procéder à des encadrements de f, et non plus à une majoration de $|f|$.

Exemple. Montrons que

$$\mathrm{J} = \int_1^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x$$

converge. La proposition ci-dessus ne s'applique pas directement, car la majoration brutale

$$\frac{|\sin x|}{x} \leq x^{-1},$$

ne rentre dans aucune case. Remarquons cependant que l'on veut intégrer un produit de deux fonctions que l'on sait aussi bien dériver qu'intégrer. Si, en particulier, on dérive le x^{-1}, on va tomber sur un x^{-2} qui lui pourrait nous permettre d'appliquer le théorème. Fixons donc $\mathrm{X} > 1$ et intégrons par partie :

$$\int_1^{\mathrm{X}} \frac{\sin x}{x}\,\mathrm{d}x = [\frac{-\cos x}{x}]_1^{\mathrm{X}} - \int_1^{\mathrm{X}} \frac{-\cos x}{-x^2}\,\mathrm{d}x,$$

ce qui permet de conclure : le premier terme du second membre converge vers $-\cos(1)$, tandis que le second terme tend vers

$$-\int_1^{+\infty} \frac{\cos x}{x^2}\,\mathrm{d}x$$

qui existe car $|x^{-2}\cos x| \leq x^{-2}$. On aurait tout aussi bien pu appliquer directement le théorème d'intégration par parties énoncé un peu plus haut, mais il est tout aussi simple, et plus prudent, de l'appliquer entre des bornes finies puis d'invoquer le théorème de Hake. En guise d'exercice, vous pouvez montrer que, si J est convergente, en revanche $|x^{-1}\sin x|$ n'est pas intégrable sur $[1,+\infty[$: l'intégrale J est semi-convergente.

On peut bien-sûr enoncer la proposition ci-dessus pour le problème de la convergence en une borne finie de l'intervalle d'intégration.

Proposition 8.3 (Critères de convergence en un point fini). — *Soit f une fonction localement intégrable sur $]a,b]$.*

1. *Si $|f(x)| \leq 1/(x-a)^\alpha$ pour un certain $\alpha < 1$ et pour tout x assez proche de a, alors* I *est absolument convergente.*
2. *Si $|f(x)| \leq |\eta \operatorname{Log}|x-a||$ pour un certain $\eta > 0$ et pour tout x assez proche de a , alors* I *est absolument convergente.*
3. *Si $|f(x)| \geq 1/(x-a)^\alpha$ pour un certain $\alpha \geq 1$ et pour tout x assez proche de a, alors* I *est divergente.*

Pour finir ce chapitre, le théorème suivant permet de ramener l'étude de certaines intégrales à l'étude de séries, ce qui permet parfois de simplifier les problèmes, mais amène surtout à de jolies formules.

Théorème 8.2 (Comparaison séries-intégrales). — *Soit f une fonction* positive *sur l'intervalle $[a,+\infty[$. Pour que l'intégrale*

$$\mathrm{I} = \int_a^{+\infty} f(x)\,\mathrm{d}x$$

soit (absolument) convergente, il faut et il suffit qu'il existe une suite croissante (x_n), tendant vers $+\infty$, et telle que la série de terme général

$$u_n = \int_{x_n}^{x_{x+1}} f(x)\,\mathrm{d}x$$

soit (absolument) convergente.

9 Ensembles de mesure nulle et notion de « presque partout »

Nous entrons maintenant dans le vif du sujet, à savoir *la théorie de l'intégration.* La première étape consiste à définir une classe d'ensembles, dits de mesure nulle, qui jouent un rôle particulier : ils sont négligables dans les problèmes d'intégration. Ainsi, nous verrons que les ensembles dénombrables sont de mesure nulle, et que l'on peut changer les valeurs d'une fonction sur un ensemble de mesure nulle sans changer son caractère intégrable ni son intégrale. Cela confère à l'intégrale une souplesse inestimable dans les applications.

9.1 Généralités

Définition 9.1. — *Une partie* N *de* $\mathbf{R}$ *est de mesure nulle si pour tout* $\varepsilon > 0$*, il existe une suite d'intervalles ouverts* $(\mathrm{J}_k)_{k\in\mathbf{N}}$*, telle que*

$$\mathrm{N} \subset \bigcup_{k\in\mathbf{N}} \mathrm{J}_k \quad \text{et} \quad \sum_{k=0}^{+\infty} |\mathrm{J}_k| \leq \varepsilon.$$

Une fonction f *définie sur une partie* A *de* $\mathbf{R}$ *est dite nulle presque partout si* $\{x \in \mathrm{A}, f(x) \neq 0\}$ *est une partie de mesure nulle. On dit qu'une propriété est vraie presque partout etc...*

Résumons quelques propriétés fondamentales des ensembles de mesure nulles.

Proposition 9.1. — *Une partie dénombrable est de mesure nulle. Une union dénombrable d'ensembles de mesure nulle est de mesure nulle. Une partie incluse dans un ensemble de mesure nulle est de mesure nulle. Un intervalle d'intérieur non vide n'est pas de mesure nulle.*

La démonstration sera faite en travaux dirrigés.

9.2 Intégrale et ensembles de mesure nulle

Avec un peu plus de travail, nous pouvons montrer le théorème suivant, qui relie la notion d'intégrale à celle de mesure nulle.

Théorème 9.1. — *Soit* f *une fonction nulle presque partout sur* $[a, b]$*. Alors* f *est intégrable, d'intégrale nulle.*

DÉMONSTRATION. — Notons E l'ensemble de mesure nulle où f est non-nulle. L'ensemble E est l'union dénombrable des ensembles de mesure nulle $\mathrm{E}_m = \{x, m-1 \leq |f(x)| < m\}$, et donc, étant donné $\varepsilon >$ et $m \geq 0$, il existe une collection d'intervalles *ouverts* $\{\mathrm{J}_k^m, k \in \mathbf{N}\}$, tels que

$$\mathrm{E}_m \subset \bigcup_k \mathrm{J}_k^m \text{ et } \sum_{k=0}^{+\infty} |\mathrm{J}_k^m| \leq \frac{\varepsilon}{m2^m}.$$

Si $x \in \mathrm{E}$, il existe un unique entier $m(x)$ tel-que $x \in \mathrm{E}_m$, et un entier minimal $k(x)$ tel que $x \in \mathrm{J}_k^m(x)$. On définit alors une jauge δ sur I par

$$\delta(x) = 1 \text{ si } x \notin \mathrm{E}, \quad \delta(x) = \frac{1}{2}\mathrm{dist}(x, \partial \mathrm{J}_{k(x)}^{m(x)}) \text{ si } x \in \mathrm{E},$$

de telle sorte que si $x \in \mathrm{E}$, l'intervalle $[x - \delta(x), x + \delta(x)]$ soit inclus dans J_k^m (la notation $\mathrm{dist}(x, \partial\mathrm{J})$ désigne la distance entre x et le bord de $\mathrm{J} = [c, d]$, *i.e.* $\min\{|x - c|, |x - d|\}$). Soit maintenant $\mathcal{P} = \{(\mathrm{I}_l, t_l)\}$ une subdivision pointée δ-fine. Alors

$$\mathrm{S}(f, \mathcal{P}) = \sum_{l=1}^{n} f(t_l)|\mathrm{I}_l| = \sum_{t_l \in \mathrm{E}} f(t_l)|\mathrm{I}_l| \leq \sum_m \sum_k \sum_{t_l \in \mathrm{Z}_m \cap \mathrm{J}_k^m} f(t_l)|\mathrm{I}_l|.$$

Mais, pour les l tels que $t_l \in \mathrm{Z}_m \cap \mathrm{J}_k^m$, les I_l sont disjoints, de longeur totale plus petite que $|\mathrm{J}_k^m|$, et $|f(t_l)| < m$, donc on en déduit que

$$|\mathrm{S}(f, \mathcal{P})| \leq \sum_m \sum_k m|\mathrm{J}_k^m| \leq \sum_m m \frac{\varepsilon}{m 2^m} \leq \varepsilon,$$

et cela achève la démonstration du théorème.

Il admet une réciproque valable pour les fonctions positives.

Théorème 9.2. — *Une fonction intégrable, positive, et d'intégrale nulle est nulle presque partout.*

Nous admettrons ce théorème ici, car sa démonstration fait appel à un lemme de recouvrement technique, ce qui la rend un peu longue. Notons cependant le corollaire suivant, conséquence de la linéarité de l'intégrale :

Corollaire 9.1. — *Si f est intégrable sur $[a, b]$, et si g est une fonction sur $[a, b]$ égale à f presque partout, alors g est intégrable et*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b g(x)\,\mathrm{d}x.$$

Remarque. Beaucoup d'énoncés de la théorie de l'intégration comportent, dans leur formulation, la mention « pour presque-tout »ou « presque partout ». Cela signifie que la condition que l'on demande ou la conclusion que l'on énonce sont vrais pour tout x, sauf pour un ensemble de mesure nulle (*i.e.* négligéable) d'entres eux. Plusieurs des théorèmes énoncés jusqu'ici peuvent ainsi être « presque partoutisés ».

9.3 L'ensemble triadique de Cantor et l'escalier du diable.

Nous allons ici construire un ensemble de mesure nulle qui n'est pas dénombrable, par un procédé itératif de retrait de sous-intervalles de $[0, 1]$. Cet ensemble admet une multitude de caractérisations. Lui et ses congénères, appelés génériquement *ensembles de Cantor,* sont d'une importance cruciale dans de nombreux domaines. Son inventeur, Georg Cantor, est un des fondateurs des mathématiques modernes.

La construction la plus simple de l'ensemble de Cantor K consiste à commencer par couper l'intervalle $[0,1]$ en trois intervalles de longueur $1/3$:

$$\mathrm{I}_0^1 = [0, 1/3],\ \mathrm{I}_1^1 =]1/3, 2/3[,\ \mathrm{I}_2^1 = [2/3, 1].$$

On retire alors l'intervalle du milieu (d'où le nom anglosaxon de *middle-third Cantor set*), et l'on pose $\mathrm{K}^1 = \mathrm{I}_0^1 \cup \mathrm{I}_2^1$. On répète ce processus de découpage et de retrait aux deux intervalles restants, I_0^1 et I_2^1. Ainsi,

$$\mathrm{I}_0^1 = \mathrm{I}_{00}^1 \cup \mathrm{I}_{01}^2 \cup \mathrm{I}_{02}^2 = [0, 1/3^2] \cup]1/3^2, 2/3^2[\cup [2/3^2, 1/3],$$

et

$$\mathrm{I}_2^1 = \mathrm{I}_{20}^2 \cup \mathrm{I}_{21}^2 \cup \mathrm{I}_{22}^2 = [2/3, 7/3^2] \cup]7/3^2, 8/3^2[\cup [8/3^2, 1].$$

On pose alors

$$\mathrm{K}^2 = \mathrm{I}_{00}^2 \cup \mathrm{I}_{02}^2 \cup \mathrm{I}_{20}^2 \cup \mathrm{I}_{22}^2.$$

L'ensemble K^3 est obtenu en retirant à chacun des 2^2 intervalles constituant K^2 l'intervalle tiers milieu. Ainsi, K^3 est l'union de 2^3 intervalles fermés disjoints, tous de longueur 3^{-3}. L'itération de ce procédé nous donne, à chaque étape, un compact K^n qui est l'union de 2^n intervalles fermés disjoints, tous de longueur 3^{-n} (*voir* la figure 8).

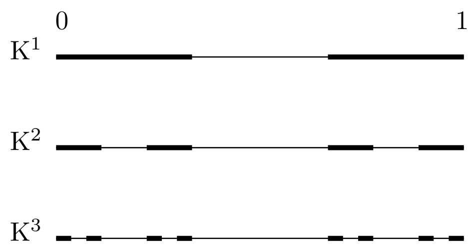

Fig. 8: Les trois premières étapes de construction de K.

La suite des sous-ensembles K^n est décroissante, et l'ensemble de Cantor est leur intersection :

$$\mathrm{K} = \bigcap_{n \geq 1} \mathrm{K}^n.$$

Remarque. On peut donner une autre description de K, en termes de développement d'un nombre réel en base 3. En effet, tout nombre réel entre 0 et 1 peut s'écrire en base 3, de même qu'il admet un développement décimal. Ainsi,

$$0 = 0,000\ldots,\ 1/3 = 0,1000\ldots,\ 2/3 = 0,20000\ldots, 1/2 = 0,1111111\ \textit{etc.}\ldots$$

Tout nombre $x \in [0,1]$ admet ainsi une écriture du type

$$x = \sum_{i=1}^{+\infty} \alpha_i 3^{-i},$$

où les « décimales »triadiques α_i valent 0, 1 ou 2. C'est ce développement que nous avons utilisé dans la notation des interavalles I_0^1 et autres. En effet, I_0^1 est constitué des nombres dont la première décimale triadique est un 0, I^1_1 ceux pour lesquelles celle-ci est un 1, et I^1_2 ceux qui commencent par un 2. Du coup, on code de la même manière les intervalles $\mathrm{I}^n_{\alpha_1,\alpha_2,\ldots,\alpha_n}$: c'est l'ensemble (a au plus deux exceptions près, mais passons sur ce détail) des nombres $x \in [0,1]$ tels que le développement de x en base 3 commence par $0,\alpha_1 \ldots \alpha_n$. Il est alors facile de décrire K, au vu de ces remarques : c'est l'ensemble des nombres réels $x \in [0,1]$ dont le développement en base 3 *ne comporte aucun* 1.

L'ensemble K est un fermé de $[0,1]$, car c'est une intersection de fermés. Il est non vide, car il s'agit d'une intersection décroissante de fermés dans un compact. Mais on a mieux :

Théorème 9.3. — *L'ensemble de Cantor* K *est un compact de mesure nulle et non dénombrable.*

DÉMONSTRATION. — La suite K^n est décroissante, et chaque K^n est l'union de 2^n intervalles I^n_k, fermés et de longueur 3^{-n}. Chaque I^n_k est inclus dans un intervalle ouvert J^n_k de longueur $(2,9)^{-n}$, et ainsi l'on a, pour tout n,

$$\mathrm{K} \subset \mathrm{K}^n \subset \bigcup_{k=1}^{2^n} \mathrm{J}^n_k,$$

avec

$$\sum_{k=1}^{2^n} |\mathrm{J}^n_k| \leq \sum_{k=1}^{2^n} (2,9)^{-n} = \left(\frac{2}{2,9}\right)^n,$$

et donc K est de mesure nulle.

Montrons maintenant que K est non dénombrable. Pour cela, raisonnons par l'absurde et supposons K dénombrable. On peut alors l'énumérer : $\mathrm{K} = \{x_1, x_2, \ldots, x_n, \ldots\}$. Construisons alors un point x qui sera dans K mais pas dans l'énumération $x_1, \ldots$ Pour cela, procédons par induction, et utilisons le fait que dans la construction de K, le nombre de choix possibles dans les intervalles I^n_k croit exponentiellement vite avec n. Tout d'abord, x_1 est soit dans I^1_0, soit dans I^1_2. Posons J_1 égal à celui des deux où x_1 *n'est pas.* Supposons maintenant construit J_{n-1}, et posons J_n égal à l'intervalle de $\mathrm{J}_{n-1} \cap \mathrm{K}^n$ où x_n *n'est pas.* On obtient ainsi une suite décroissante d'intervalles fermés non vides, dont l'intersection est donc non vide et contient un point $x \in \mathrm{K}$. Pour tout n, on a $x_n \notin \mathrm{J}_n$ et $x \in \mathrm{J}_n$, donc $x \neq x_n$ et l'on a achevé la démonstration du théorème.

Une fois construit l'ensemble de Cantor, nous pouvons introduire la fonction dite « escalier du diable », car, comme nous le verrons, même s'il permet de monter d'un étage, il y a toujours, entre deux marches différentes, une infinité de marches. Construisons cette fonction par approximation. Définissons φ_1 comme étant constante égale à 2^{-1} sur I^1_1, $\varphi_1(0) = 0$ et $\varphi_1(1) = 1$, et enfin, sur les 2 intervalles restants, φ_1 est affine, de manière à ce qu'elle soit continue. Il n'y a qu'une seule fonction φ_1 vérifiant ces hypothèses (*voir* la figure 9 pour comprendre l'évidence de cette affirmation). De manière générale, φ_n est l'unique fonction continue, affine par morceaux, telle que $\varphi_n(0) = 0, \varphi_n(1) = 1$ et φ_n prend successivement, sur les $2^n - 1$ intervalles retirés de $[0,1]$ pour construire K^n, les valeurs $1/2^n, 2/2^n, \ldots, (2^n - 1)/2^n$.

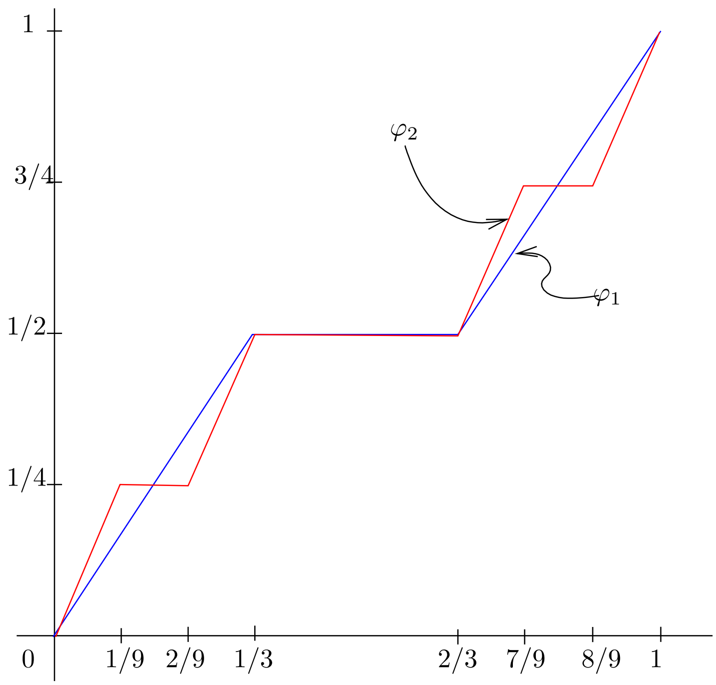

FIG. 9: L'escalier du diable de Cantor–Lebesgue.

La suite K^n est décroissante, donc la suite $[0,1]-K^n$ est croissante, et pour tout $x \in [0,1]-K$, $\varphi_n(x)$ est constante à partir d'un certain rang $n(x)$. En fait, pour tout $n \geq 1$, les graphes de φ_n et φ_{n+1} soient coïncident, soit sont dans la même bande horizontale de hauteur $1/2^n$. On en déduit que

$$\text{pour tout } x \in [0,1], \text{ pour tout } n \geq 1, |\varphi_n(x) - \varphi_{n+1}(x)| \leq 2^{-n}.$$

La suite $\varphi_n(x)$ est de Cauchy, et elle admet donc une limite $\varphi(x)$, qui satisfait de plus :

$$\text{pour tout } x \in [0,1], \text{ pour tout } n \geq 1, |\varphi_n(x) - \varphi(x)| \leq 2^{-(n-1)}.$$

Ainsi, nous avons montré que (φ_n) converge uniformément vers φ sur $[0,1]$. Comme les φ_n sont continues et croissantes, φ l'est aussi, et de plus $\varphi(0) = 0$, $\varphi(1) = 1$.

Soit maintenant x choisi hors de l'ensemble de mesure nulle K. Comme φ_n est égale à une constante indépendante de $n \geq n(x)$ au voisinage de x, φ est elle aussi constante au voisinage de x. En particulier, φ est dérivable en x et $\varphi'(x) = 0$.

De toute cette discussion, nous déduisons le théorème suivant :

Théorème 9.4. — *La fonction φ, dite « escalier du Diable », est continue surjective de $[0,1]$ sur $[0,1]$, croissante. Elle est par ailleurs presque partout dérivable, avec $\varphi'(x) = 0$ presque partout.*

Nous voyons ici apparaître une pathologie dans les liens entre intégration et primitivation. Ainsi, φ est presque partout dérivable, de dérivée nulle. En d'autres termes, φ est une « presque »primitive de 0. Mais si l'on appliquait le théorème fondamental à φ, on obtiendrait

$$\int_0^1 \varphi'(x)\,\mathrm{d}x = \varphi(1) - \varphi(0) = 1,$$

mais le terme de gauche est nul car φ' est nulle presque partout. Cela est impossible. On ne peut donc espérer généraliser le théorème fondamental avec des presque partout dedans. Si l'on veut faire cela, il faut faire une hypothèse supplémentaire sur les variations locales de φ.

10 Les théorèmes de convergence. Applications

10.1 Le lemme de Henstock

La théorie un peu plus avancée de l'intégrale repose presque entièrement sur un lemme facile à démontrer, mais d'une puissance remarquable. Pour des raisons de simplicité, nous ne ferons les démonstrations que dans le cas où l'intervalle d'intégration est compact.

Une *sous-subdivision pointée* de l'intervalle compact I est une collection finie J_k d'intervalles fermés, d'intérieurs disjoints, inclus dans I, et la donnée d'un point de marquage $t_k \in \mathrm{J}_k$ pour tout k : on ne suppose plus que I est l'union des J_k. Les définitions de finesse, de somme de Riemann... passent immédiatement aux sous-subdivisions. Si $\mathcal{S}_0 = \{(\mathrm{J}_k, t_k), k = 1 \dots n\}$ est une sous-subdivision pointée, on pose

$$\mathrm{D}(\mathcal{S}_0) = \bigcup_{k=1}^{n} \mathrm{J}_k.$$

Théorème 10.1 (Lemme de Henstock). — *Soient f une fonction intégrable sur* I *et $\varepsilon > 0$. Soit δ_ε une jauge telle que pour toute subdivision δ-fine, on ait $|\mathrm{S}(f, \mathcal{S}) - \int f| \leq \varepsilon$. Alors, pour toute sous-subdivision δ-fine $\mathcal{S}_0 = \{(\mathrm{J}_k, t_k), k = 1, \dots, n\}$, on a :*

$$\left| \sum_{k=1}^{n} \left(f(t_k)|\mathrm{J}_k| - \int_{\mathrm{J}_k} f \right) \right| = \left| \mathrm{S}(f, \mathcal{S}_0) - \int_{\mathrm{D}(\mathcal{S}_0)} f \right| \leq \varepsilon.$$

En d'autres termes, si l'on a un contrôle sur les subdivisions, on a le meme contrôle sur les sous-subdivisions : c'est une espèce d'uniformité spatiale de l'intégrale. Celà permet, dans la pratique, de *concentrer l'étude sur des parties de l'intervalle,* où se posent des problèmes quant aux sommes de Riemann de f.

Démonstration. — Soient $\varepsilon > 0$ fixé et $\delta = \delta_\varepsilon$ comme dans l'énoncé. L'idée est de compléter la sous-subdivision $\mathcal{S}_0$ en une subdivision de I en entier, en prenant la sous-subdivision supplémentaire suffisament fine pour que ses contributions dans les sommes de Riemann soit très petite. Soient donc $\mathrm{K}_1, \dots, \mathrm{K}_m$ l'adhérence des composantes connexes de $\mathrm{I} - \mathrm{D}(\mathcal{S}_0)$. Soit $\alpha > 0$ arbitrairement petit. Comme f est intégrable sur chacun des intervalles K_i, il existe une jauge δ_i sur K_i telle que, pour toute subdivision δ_i-fine $\mathcal{Q}_i$ de K_i, on ait

$$|\mathrm{S}(f, \mathcal{Q}_i) - \int_{\mathrm{K}_i} f| \leq \frac{\alpha}{m}.$$

Quitte à diminuer les jauges, on peut toujours supposer que $\delta_i \leq \delta$. Soit maintenant $\mathcal{P}^*$ la subdivision pointée de I définie par $\mathcal{P}^* = \mathcal{S}_0 \cup \mathcal{Q}_1 \cup \cdots \cup \mathcal{Q}_m$. Elle est δ-fine, et de plus

$$\mathrm{S}(f, \mathcal{P}^*) = \mathrm{S}(f, \mathcal{S}_0) + \mathrm{S}(f, \mathcal{Q}_1) + \cdots + \mathrm{S}(f, \mathcal{Q}_m) \text{ et } \int_{\mathrm{I}} f = \int_{\mathrm{D}(\mathcal{S}_0)} f + \int_{\mathrm{K}_1} f + \cdots + \int_{\mathrm{K}_m} f.$$

Il s'ensuit que

$$\left| \mathrm{S}(f, \mathcal{S}_0) - \int_{\mathrm{D}(\mathcal{S}_0)} f \right| \leq \left| \mathrm{S}(f, \mathcal{P}^*) - \int_{\mathrm{I}} f \right| + \sum_{i=1}^{m} \left| \mathrm{S}(f, \mathcal{Q}_i) - \int_{\mathrm{K}_i} f \right| \leq \varepsilon + \alpha,$$

et comme α a été choisi arbitrairement petit, l'inégalité du lemme de Henstock en découle.

Corollaire 10.1. — *Sous les mêmes hypothèses, on a en plus les inégalités suivantes :*

1.

$$\sum_{k=1}^{n} \left| f(t_k)|\mathrm{J}_k| - \int_{\mathrm{J}_k} f \right| \leq 2\varepsilon.$$

2.

$$\left| \sum_{k=1}^{n} \left(|f(t_k)||\mathrm{J}_k| - |\int_{\mathrm{J}_k} f| \right) \right| \leq 2\varepsilon.$$

Démonstration. — Pour obtenir la première inégalité, on casse $\mathcal{S}_0$ en deux sous-subdivisions, $\mathcal{S}_0^+$ et $\mathcal{S}_0^-$, définies par le signe de $f(t_i)|\mathrm{J}_i| - \int_{\mathrm{J}_i} f$. On peut appliquer le lemme de Henstock à $\mathcal{S}_0^+$ et $\mathcal{S}_0^-$:

$$\sum_{\mathcal{S}_0^+} \left(f(t_j) - \int_{\mathrm{I}_j} f \right) \leq \varepsilon \quad \text{et} \quad -\sum_{\mathcal{S}_0^-} \left(f(t_j) - \int_{\mathrm{I}_j} f \right) \leq \varepsilon.$$

L'addition de ces deux inégalités donne la première du corollaire. La deuxième s'obtient grâce à la première et à l'inégalité triangulaire.

Un exemple d'application. Le lemme de Henstock est donc très simple à démontrer, mais nous allons voir qu'il a de nombreuses et profondes applications, comme par exemple la continuité des intégrales indéfinies :

Théorème 10.2. — *Si f est une fonction intégrable sur $\mathrm{I} = [a, b]$, alors ses intégrales indéfinies sont continues.*

Démonstration. — Il suffit de considérer l'intégrale indéfinie s'annulant en a :

$$\psi_f(x) = \int_a^x f.$$

Soit $c \in [a, b[$, et montrons que ψ_f est continue à droite en c. La démonstration de la continuité à gauche pour $c \in]a, b]$ sera la même, quitte à faire le changement de variable $x \mapsto -x$. Fixons $\varepsilon > 0$ et soit δ une jauge comme dans les hypothèses du lemme de Henstock. Modifions la jauge afin de forcer c à être un point de marquage : soit δ' définie par

$$\delta'(x) = \min\{\delta(x), \frac{1}{2}|x - c|\}, x \neq c, \ \delta'(c) = \min\{\delta(c), \frac{\varepsilon}{|f(c)| + 1}\}.$$

Soient $\alpha < \delta'(c)$ et $\mathcal{S}_0$ la sous-subdivision pointée constituée de la seule paire $([c, c+\alpha], c)$. Elle est δ-fine, et donc la première inégalité du lemme de Henstock nous dit que

$$\left| \alpha f(c) - \int_c^{c+\alpha} f \right| \leq \varepsilon.$$

Mais $\alpha \leq \varepsilon(|f(c)|+1)^{-1}$, et donc

$$|\psi_f(c+\alpha) - \psi_f(c)| = \left|\int_c^{c+\alpha} f\right| \leq |f(c)|\alpha + \varepsilon \leq 2\varepsilon,$$

ce qui achève la démonstration.

***A parte* historique.** Il est bien beau ce théorème, mais ce que l'on aimerait vraiment savoir maintenant, c'est si ψ_f est dérivable, de dérivée f. En d'autres termes, toute fonction intégrable admet-elle une primitive ? La réponse est donnée par le théorème suivant, dûe initialement à Arnaud Denjoy et Oskar Perron.

Théorème 10.3. — *Si f est intégrable sur $[a, b]$, alors son intégrale indéfinie est continue et admet presque partout une dérivée égale à f.*

Nous ne démontrerons pas ce théorème ici, car la preuve fait appel à un lemme de recouvrement extrêmement utile, mais un peu délicat. On peut résumer les liens entre intégration et dérivation de la manière suivante :

- Si l'on dérive une intégrale indéfinie, alors on retrouve, presque partout, la fonction dont on est parti.
- Si l'on intègre une vraie dérivée, alors on retrouve la fonction qu'on a dérivée.
- Si l'on intègre un dérivée définie seulement presque partout, alors on peut très bien ne pas pouvoir retrouver la fonction de départ. L'exemple de l'escalier du diable est à ce titre éclairant (dérivée nulle presque partout, mais fonction continue surjective de $[0, 1]$ sur lui-même).

Pour avoir un théorème complet de commutation entre dérivation et intégration, il faut se restreindre à une classe de fonctions présentant de faibles variations locales.

10.2 Fonctions Lebesgue-intégrables

Comme vous l'avez vu en travaux dirrigés, il existe des fonctions intégrables dont la valeur absolue n'est pas intégrable. En d'autres termes, comme il existe des séries semi-convergentes, il existe des intégrales semi-convergentes (qui n'existent par dans les théories de Riemann et Lebesgue). Cependant, comme dans la théorie des séries, les fonctions à valeur absolue intégrable jouent un rôle particulier de par leur stabilité dans une large classe de processus de passage à la limite.

Définition 10.1. — *Une fonction f définie sur un intervalle quelconque* I *est dite* Lebesgue -intégrable, ou intégrable au sens de Lebesgue *si f* ET *$|f|$ sont intégrables. On dit aussi que f est absolument intégrable.*

Remarque : l'hypothèse d'intégrabilité sur f ne peut être retirée : l'implication $|f|$ intégrable implique f intégrable est indécidable...

Le premier théorème, très utile, concernant les fonctions Lebesgue-intégrables, est le principe de comparaison suivant :

Proposition 10.1. — *Soient f, g deux fonctions intégrables. Si, pour presque tout $x \in \mathrm{I}$, on a $|f(x)| \leq g(x)$, alors f est absolument intégrable et de plus*

$$\left|\int_{\mathrm{I}} f\right| \leq \int_{\mathrm{I}} |f| \leq \int_{\mathrm{I}} g.$$

Démonstration. — On peut toujours, quitte à modifier g sur un ensemble de mesure nulle, supposer que l'inégalité $|f(x)| \leq g(x)$ est vraie pour tout x. Le théorème est alors une conséquence immédiate de la positivité de l'intégrale, puisque, pour tout $x \in [a, b]$, on a

$$-|f(x)| \leq f(x) \leq |f(x)|.$$

La classe des fonctions absolument intégrables est suffisament importante pour mériter une notation propre :
Notation : Si I est un intervalle quelconque, nous noterons $\mathcal{L}(\mathrm{I})$ l'ensemble des fonctions intégrables au sens de Lebesgue sur I.

Proposition 10.2. — *L'espace $\mathcal{L}(\mathrm{I})$ est un espace vectoriel. De plus, $f \in \mathcal{L}(\mathrm{I})$ si et seulement si $f^+ = \max\{f, 0\}$ et $f_- = \min\{f, 0\}$ sont intégrables. Enfin, si $f, g \in \mathcal{L}(\mathrm{I})$, alors $\min\{f, g\}$ et $\max\{f, g\}$ sont aussi intégrables au sens de Lebesgue.*

Nous laissons la preuve de cette proposition en exercice, en indiquant simplement que le théorème de comparaison vu plus haut peut être utile. On peut ici faire une remarque qui simplifie la vie : si f est Lebesgue-intégrable sur un intervalle I, alors on peut prolonger f à $\mathbf{R}$ tout entier en posant $\bar{f}(x) = 0$ si $x \notin \mathrm{I}$ et $\bar{f}(x) = f(x)$ sinon. On a alors équivalence entre $f \in \mathcal{L}(\mathrm{I})$ et $\bar{f} \in \mathcal{L}(\mathbf{R})$. On pourra ainsi dans la suite se limiter au cas de fonctions intégrables sur $\mathbf{R}$ en entier, et déduire l'énoncé correspondant pour les intervalles quelconques par une simple restriction.

10.3 Théorèmes de convergence

L'intégrale de Riemann classique (celle avec les jauges constantes) est intéressante à cause de son contenu intuitif. Elle souffre cependant d'un grave défaut : elle se comporte mal par approximation des fonctions. En mathématiques, on ne sait faire des calculs explicites que pour une petite classe de fonctions : polynômes et fonctions en escalier le plus souvent. On a donc souvent besoin, pour attaquer des problèmes généraux, de passer par des approximations. En d'autres termes, on se donne une fonction f sur $[a, b]$, et une suite (f_n) qui converge vers f en un sens à préciser. Si cette convergence est très bonne, uniforme par exemple, alors la limite des intégrales de f_n (au sens de Riemann) est égale à l'intégrale de f. Cependant, la convergence uniforme est trop forte pour attraper, par approximation, toutes les fonctions intégrables, en particulier celles, un peu pathologiques, qui sortent à tous bouts de champs en physique, chimie, systèmes dynamiques et autres. Supposons par exemple que nous ne sachions faire des calculs explicites que sur des polynômes. Par passage à la limite uniforme sur un intervalle compact $[a, b]$, nous n'attraperons que des fonctions continues (en fait, toutes les fonctions continues, c'est le théorème de Stone-Weierstraß). Impossible d'approcher ainsi la fonction de Dirichlet ou l'indicatrice d'un Cantor. Enfin, les théorèmes de convergence de l'intégrale de Riemann, mêmes les plus généraux, font toujours l'hypothèse de l'intégrabilité de la limite : impossible ainsi d'étendre la classe des fonctions intégrables par passage à la limite.

L'intégrale de Lebesgue (et plus généralement l'intégrale de Kurzwil et Henstock) permet de s'affranchir de cette limitation rédibitoire de l'intégrale de Riemann, car elle possède des théorème de convergence valables sous de très faibles hypothèses. Ce sont ces théorèmes qui rendent cette théorie si féconde et utile : c'est elle que l'on utilise dans toute l'analyse moderne (équations au dérivées partielles, distributions, théorie des opérateurs *etc...*) ainsi que dans toutes les branches des mathématiques où le signe « $\int$ »fait son apparition. Enfin, c'est l'absence d'hypothèse d'intégrabilité sur la fonction limite qui rend ces énoncés si puissants.

Commençons par démontrer le théorème de convergence monotone, aussi connu sous le nom de théorème de Beppo-Levi.

Théorème 10.4. — *Soient* I *un intervalle quelconque,* (f_n) *une suite de fonctions définies sur* I*, et* f *une fonction définie sur* I*. On suppose que*

1. *les fonctions* f_n *sont Lebesgue-intégrables,*
2. *pour presque tout* $x \in \mathrm{I}$*, la suite* $(f_n(x))$ *est croissante,*
3. *la suite* (f_n) *converge vers* f *presque partout,* i.e. *pour presque-tout* $x \in \mathrm{I}$*, on a* $\lim_{n\to+\infty} f_n(x) = f(x)$*,*
4. *et la suite des intégrales* $\left(\int_{\mathrm{I}} f_n\right)$ *est majorée.*

Alors f *est Lebesgue-intégrable et*

$$\lim_{n\to+\infty} \int_{\mathrm{I}} f_n(x)\,\mathrm{d}x = \int_{\mathrm{I}} \lim_{n\to+\infty} f_n(x)\,\mathrm{d}x = \int_{\mathrm{I}} f(x)\,\mathrm{d}x.$$

Le même énoncé est vrai en remplaçant « croissante »par « décroissante »et « majorée »par « minorée ».

Démonstration. — Nous nous placerons dans le cas où $\mathrm{I} = [a, b]$ est compact, et où la suite $f_n(x)$ converge simplement partout vers $f(x)$. La suite des intégrables

$$\int_{\mathrm{I}} f_n$$

est croissante, et comme elle est majorée, elle converge. On pose

$$\mathrm{A} := \lim_{n\to+\infty} \int_{\mathrm{I}} f_n = \sup\{\int_{\mathrm{I}} f_n, n \in \mathbf{N}\}.$$

Fixons $\varepsilon > 0$ et $r > 0$ tels que $2^r > 4/\varepsilon$ et

$$0 \le \mathrm{A} - \int_{\mathrm{I}} f_r < \varepsilon.$$

Remarquons que par monotonicité, ces conditions sont en fait vérifiées pour tous les entiers plus grand que r. Pour tout $n \in \mathbf{N}$, il existe une jauge δ_n telle que, pour toute subdivision δ_n-fine $\mathcal{P}$ de I, on ait

$$|\mathrm{S}(f_n, \mathcal{P}) - \int_{\mathrm{I}} f_n| \leq 2^{-n}.$$

La suite (f_n) converge simplement et de manière croissante vers f, donc pour tout $x \in \mathrm{I}$, il existe un entier $n(x) \geq r$ tel que

$$0 \leq f(x) - f_{n(x)}(x) \leq \varepsilon,$$

et l'on pose alors $\delta(x) = \delta_{n(x)}(x)$, qui est une jauge sur I, qui va nous permettre de montrer que f est intégrable. Une fois encore, cela illustre l'extraordinaire puissance de cette notion de jauge... Soit $\mathcal{P} = \{(\mathrm{I}_k, t_k), k = 1, \ldots, l\}$ une subdivision δ-fine de I. Comme

$$\mathrm{S}(f, \mathcal{P}) - \mathrm{A} = \sum_{k=1}^{l} \left[\Big(f(t_k)|\mathrm{I}_k| - f_{n(t_k)}(t_k)|\mathrm{I}_k| \Big) + \left(f_{n(t_k)}(t_k)|\mathrm{I}_k| - \int_{\mathrm{I}_k} f_{n(t_k)} \right) + \left(\int_{\mathrm{I}_k} f_{n(t_k)} - \mathrm{A} \right) \right],$$

on en déduit, grâce aux hypothèses faites et à l'inégalité triangulaire, que

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{A}| \leq \sum_{k=1}^{l} \varepsilon |\mathrm{I}_k| + \sum_{k=1}^{l} \left| f_{n(t_k)}(t_k)|\mathrm{I}_k| - \int_{\mathrm{I}_k} f_{n(t_k)} \right| + \left| \sum_{i=1}^{l} \int_{\mathrm{I}_k} f_{n(t_k)} - \mathrm{A} \right|.$$

Le premier terme du membre de droite est majoré par $\varepsilon(b - a)$. Pour majorer le second, nous allons utiliser le lemme de Henstock en cassant la subdivision $\mathcal{P}$ en sous-subdivisions $\mathcal{P}_p$, définies par la condition $n(t_k) = p$, $p = r, \ldots, s = \max\{n(t_1), \ldots, n(t_l)\}$. La sous-subdivision $\mathcal{P}_p$ est clairement δ_p-fine, et la première inégalité du corollaire 10.1 au lemme de Henstock nous dit que

$$\sum_{n(t_k)=p} \left| f_{n(t_k)}|\mathrm{I}_k| - \int_{\mathrm{I}_k} f_{n(t_k)} \right| \leq \frac{1}{2^{p-1}},$$

ce qui, sommé sur les p variant de r à s, donne

$$\sum_{k=1}^{l} \left| f_{n(t_k)}|\mathrm{I}_k| - \int_{\mathrm{I}_k} f_{n(t_k)} \right| \leq \varepsilon.$$

Il nous reste à estimer le troisième terme du second membre. Pour cela, on remarque que $f_r \leq f_{n(t_k)} \leq f_s$, et donc

$$\int_{\mathrm{I}} f_r \leq \sum_{k=1}^{l} \int_{\mathrm{I}_k} f_{n(t_k)} \leq \int_{\mathrm{I}} f_s.$$

Par hypothèse sur r, cela implique

$$\mathrm{A} - \varepsilon \leq \sum_{k=1}^{l} \int_{\mathrm{I}_k} f_{n(t_k)} \leq \mathrm{A},$$

et donc le troisième terme en question est lui aussi majoré par ε. Ainsi

$$|\mathrm{S}(f,\mathcal{P}) - \mathrm{A}| \leq (b-a+2)\varepsilon,$$

et f est intégrable, d'intégrale $\mathrm{A} = \lim \int_{\mathrm{I}} f_n$. Pour montrer que $|f|$ est intégrable, et donc que f est Lebesgue-intégrable, on utilise la deuxième inégalité du corollaire 10.1, et un argument analogue.

Voila, le plus dur est fait ! Nous pourrons, à partir de maintenant, récolter nos semences. Commençons par démontrer deux autres résultats de convergence. Le premier, dit lemme de Fatou, est très utile dans bien des situations.

Théorème 10.5. — *Soient f_n, g des fonctions Lebesgue-intégrables telles que $g \leq f_n$ pour tout n et* $\liminf \int_{\mathrm{I}} f_n$ *soit finie. Alors* $\liminf f_n$ *est Lebesgue-intégrable et*

$$(\dagger) \qquad -\infty < \int_{\mathrm{I}} \liminf f_n \leq \liminf \int_{\mathrm{I}} f_n < +\infty.$$

Démonstration. — Pour tout $k \geq n$, posons $g_{n,k}(x) = \min\{f_n(x), f_{n+1}(x), \ldots, f_k(x)\}$, et $h_n(x) = \inf\{f_i(x), i \geq n\}$. L'hypothèse que $f_n \geq g$ implique que h_n existe. Enfin, on a, par définition,

$$h_n(x) = \lim_{k\to+\infty} g_{n,k}(x).$$

La suite $(g_{n,k})_{k\geq n}$ est décroissante, et de plus

$$\int_{\mathrm{I}} g \leq \int_{\mathrm{I}} g_{n,k} \leq \int_{\mathrm{I}} f_n, \text{ pour tout } k \geq n,$$

donc, d'après le théorème de convergence monotone, la fonction h_n est Lebesgue-intégrable, et, comme

$$\int_{\mathrm{I}} h_n \leq \int_{\mathrm{I}} f_n,$$

on a

$$\lim_{n\to+\infty} \int h_n \leq \liminf_{n\to+\infty} \int_{\mathrm{I}} f_n.$$

Le second membre est fini par hypothèse, donc on peut encore appliquer le théorème de convergence monotone, et l'on obtient :

$$\int_{\mathrm{I}} \liminf_{n\to+\infty} f_n = \int_{\mathrm{I}} \lim_{n\to+\infty} h_n = \lim_{n\to+\infty} \int_{\mathrm{I}} h_n,$$

ce qui permet de conclure.

Le second théorème de convergence est la grande victoire de la théorie de Lebesgue : le théorème de convergence dominée. Une visite approfondie de ce haut lieu de l'analyse mathématique s'impose.

Théorème de convergence dominée de Lebesgue. — *Soit* (f_n) *une suite de fonctions Lebesgue-intégrables sur* I, *convergeant presque partout vers une fonction* f. *On suppose qu'il existe une fonction Lebesgue-intégrable* g *telle que*

$$|f_n(x)| \leq g(x), \quad \textit{pour presque-tout}\, x \in \mathrm{I}, \ \textit{et}\, \forall n \in \mathbf{N}.$$

Alors f *est Lebesgue-intégrable sur* I, *et*

$$\lim_{n\to+\infty} \int_{\mathrm{I}} f_n = \int_{\mathrm{I}} \lim_{n\to+\infty} f_n = \int_{\mathrm{I}} f.$$

Démonstration. — On applique le lemme de Fatou aux suites de fonctions $f_n - g$ et $h - f_n$.

Ces théorèmes nous ouvrent la porte d'un monde merveilleux : celui de la *mesure de Lebesgue de* $\mathbf{R}$. Nous entrouvrirons cette porte pendant les travaux dirigés. Citons deux corollaires immédiats :

Corollaire 10.2. — *Soit* (f_n) *une suite de fonction Lebesgue-intégrables sur l'intervalle compact* $[a, b]$. *Si la suite* (f_n) *est uniformément bornée,* i.e. *il existe une constante* $\mathrm{M} > 0$ *telle que* $|f_n(x)| \leq \mathrm{M}, \forall x \in [a, b], \forall n \in \mathbf{N}$, *et si* (f_n) *converge simplement vers* f, *alors*

$$\lim_{n\to+\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Corollaire 10.3. — *Soit* (f_n) *est une suite de fonctions Lebesgue-intégrables sur* $[a, b]$ *qui converge uniformément vers* f *sur* $[a, b]$. *Alors* f *est intégrable et*

$$\lim_{n\to+\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

10.4 Intégrales dépendant d'un paramètre et dérivation sous le signe somme

Les théorèmes de passage à la limite, ou de différentiation sous le signe somme, font souvent appel, dans le cadre de la théorie de Riemann, à des hypothèses d'uniformité. Nous allons ici obtenir les mêmes résultats sous des hypothèses plus lâches de domination. Nous ne les donnerons pas dans leur plus grande généralité, car il faudrait faire appel à la notion de fonction mesurable.

Nous considérerons une fonction $(x, t) \mapsto f(x, t)$ de deux variables, définie sur un pavé $\mathrm{I} \times \mathrm{T}$, où $\mathrm{T} = [c, d]$ est borné et où I est un intervalle quelconque. La variable t sera le paramètre.

Théorème 10.6. — *En gardant les mêmes notations, soit* t_0 *un point de* T. *On suppose que :*

1. *pour tout* $t \neq t_0$*, la fonction* $x \mapsto f(x,t)$ *est Lebesgue-intégrable en la variable* x *sur* I.
2. *Pour presque-tout* $x \in \mathrm{I}$*, on a* $\lim_{t\to t_0} f(x,t) = f(x,t_0)$.
3. *Il existe une fonction* g*, Lebesgue-intégrable sur* I*, telle que, pour tout* $t \neq t_0$ *et presque-tout* $x \in \mathrm{I}$ *on ait*
$$|f(x,t)| \leq g(x).$$

Alors $x \mapsto f(x,t_0)$ *est Lebesgue-intégrable sur* I *et*

$$\lim_{t\to t_0} \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x = \int_{\mathrm{I}} f(x,t_0)\,\mathrm{d}x.$$

Démonstration. — C'est une application immédiate du théorème de convergence dominée ; si t_n est une suite d'éléments de T tendant vers t_0, on pose $f_n(x) = f(x,t_n)$ et on applique à la suite (f_n) le théorème de Lebesgue.

Théorème 10.7 (continuité sous le signe intégral). — *Soient* $\mathrm{I} = [a,b]$ *et* $\mathrm{T} = [c,d]$*, et* $f(x,t)$ *une fonction définie sur* $[a,b] \times [c,d]$*. On suppose que :*

1. *Pour tout* $t \in \mathrm{T}$*, la fonction* $x \mapsto f(x,t)$ *est Lebesgue-intégrable en la variable* x *sur* I,
2. *pour tout* $x \in \mathrm{I}$*, la fonction* $t \mapsto f(x,t)$ *est continue en la variable* t *sur* T,
3. *il existe une fonction Lebesgue-intégrable* g *sur* I *telle que*

$$|f(x,t)| \leq g(x), \quad \forall (x,t) \in \mathrm{I} \times \mathrm{T}.$$

Alors la fonction

$$t \mapsto \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x$$

est continue sur $\mathrm{T} = [c,d]$.

Démonstration. — On applique le théorème 10.6 de passage à la limite, en tout point t_0 de T.

Voici maintenant l'énoncé que nous attendons tous, à savoir celui qui permet d'intervertir l'intégration et la dérivation partielle, ou « théorème de dérivation sous le signe somme ».

Théorème 10.8 (Dérivabilité sous le signe intégral). — *Avec les même notations que plus haut, on suppose que :*

1. *pour tout* $t_0 \in \mathrm{T}$*, la fonction* $x \mapsto f(x,t_0)$ *est Lebesgue-intégrable sur* I.
2. *Pour tout* $x \in \mathrm{I}$*, la fonction* $t \mapsto f(x,t)$ *est dérivable, de dérivée*

$$f_t(x) = \frac{\partial f}{\partial t}(x,t).$$

3. *Il existe g, Lebesgue-intégrable sur* I*, telle que*

$$|f_t(x)| \leq g(x),\ \forall x \in \mathrm{I}\ et\ \forall t \in \mathrm{T}.$$

Alors pour tout t, les fonctions

$$x \mapsto f(x,t)\ et\ x \mapsto \frac{\partial f}{\partial t}(x,t)$$

sont Lebesgue-intégrables sur $[a,b]$, et la fonction

$$t \mapsto \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x$$

est dérivable sur T*, de dérivée*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathrm{I}} f(x,t)\,\mathrm{d}x = \int_{\mathrm{I}} \frac{\partial f}{\partial t}(x,t)\,\mathrm{d}x.$$

Démonstration. — Fixons $t \in \mathrm{T}$ et soit t_n une suite d'éléments de $\mathrm{T} - \{t\}$ telle que $t_n \longrightarrow t$. Alors, pour tout $x \in \mathrm{I}$, on a

$$f_t(x,t) = \lim_{n\to+\infty} \frac{f(x,t_n) - f(x,t)}{t_n - t}.$$

La formule des accroissements finis nous dit par ailleurs qu'il existe s_n (dépendant de x, t et t_n) entre t_n et t tel que

$$\frac{f(x,t_n) - f(x,t)}{t_n - t} = f_t(x, s_n).$$

Posons alors $\varphi_n(x) = f_t(x,s_n)$. La suite φ_n vérifie les hypothèses du théorème de convergence dominée, et donc $x \mapsto \lim_n \varphi_n(x) = f_t(x,t)$ est intégrable, et

$$\lim_{n\to+\infty} \int_{\mathrm{I}} \varphi_n(x) = \int_{\mathrm{I}} f_t(x,t)\,\mathrm{d}x.$$

Mais

$$\int_{\mathrm{I}} \varphi_n(x)\,\mathrm{d}x = \int_{\mathrm{I}} \frac{f(x,t_n) - f(x,t)}{t_n - t}\,\mathrm{d}x = \frac{1}{t_n - t}\left(\int_{\mathrm{I}} f(x,t_n)\,\mathrm{d}x - \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x\right),$$

et le théorème est démontré.

Nous allons finir cette partie par un théorème utile mais rarement vu dans les cours de licence, la formule de Leibniz de dérivation d'intégrale dont l'intégrande *et* les bornes dépendent d'un paramètre.

Théorème 10.9. — *Soient $f : \mathbf{R} \times \mathrm{T} \longrightarrow \mathbf{R}$ et $u, v : \mathrm{T} \longrightarrow \mathbf{R}$ des fonctions satisfaisant les hypothèses suivantes :*

1. *f est continue sur $\mathbf{R} \times \mathrm{T}$.*

2. *La dérivée partielle* $(x,t) \mapsto f_t(x,t)$ *existe, et est continue sur* $\mathbf{R} \times \mathrm{T}$.
3. u *et* v *sont dérivables sur* T.

Alors la fonction $\mathrm{G} : \mathrm{T} \longrightarrow \mathbf{R}$, *définie par*

$$\mathrm{G}(t) = \int_{u(t)}^{v(t)} f(x,t)\,\mathrm{d}x$$

existe, et est dérivable sur T. *De plus, sa dérivée est donnée par*

$$\mathrm{G}'(t) = \int_{u(t)}^{v(t)} \frac{\partial f}{\partial t}(x,t)\,\mathrm{d}x + f(v(t),t)v'(t) - f(u(t),t)u'(t).$$

11 Appendice

Nous regroupons dans cet appendice deux compléments au cours : la démonstration du théorème de Hake et une courte introduction à la méthode d'Euler pour le calcul approché des primitives.

11.1 Preuve du théorème de Hake

Nous avons, dans la section sur les intégrales impropres, énoncé le théorème de Hake sous la forme suivante :

Théorème 11.1. — *La fonction $f : [a,+\infty[\longrightarrow \mathbf{R}$ est intégrable si, et seulement si, pour tout $c \in [a,+\infty[$, f est intégrable sur $[a,c]$, et si la limite*

$$\lim_{c \to +\infty} \int_a^c f(x)\,\mathrm{d}x = \mathrm{L}$$

existe. On a alors

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \mathrm{L}.$$

Comme nous n'avons pas développé la théorie générale de l'intégrabilité sur les intervalles infinis outre mesure, nous allons démontrer ce théorème en remplaçant la borne $+\infty$ par un réel fini $b > a$. Quitte à poser $f(b) = 0$, nous obtenons l'énoncé suivant :

Proposition 11.1. — *Soit $f : [a,b] \longrightarrow \mathbf{R}$ une fonction numérique. Alors f est intégrable sur $[a,b]$ si, et seulement si, elle est intégrable sur tout sous-intervalle de la forme $[a,c]$, avec $a < c < b$, et si la limite*

$$\lim_{c \to b} \int_a^c f(x)\,\mathrm{d}x = \mathrm{L}$$

existe. Dans ce cas, on a

$$\int_a^b f(x)\,\mathrm{d}x = \mathrm{L}.$$

La preuve dans le cas d'une borne infinie n'est guère différente de celle que nous allons donner maintenant.

Démonstration. — Commençons par montrer l'implication directe, et supposons donc f intégrable sur $[a,b]$. La relation de Chasles dit que f est intégrable sur $[a,c]$ pour tout $c \in [a,b]$, et le théorème de continuité des intégrales indéfinies permet de conclure que

$$\lim_{c \to b} \int_a^c f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x,$$

ce qui est la conclusion recherchée.

Montrons maintenant l'implication réciproque, et reprenons les notations du théorème. Soi (c_k) une suite strictement croissante telle que $a = c_0$ et $b = \lim_{k\to+\infty} c_k$. Soit également $\varepsilon > 0$. Il existe $r \in \mathbf{N}$ tel que, pour tout $k \geq r$, on ait

$$b - c_k \leq \frac{\varepsilon}{|f(b)| + 1} \quad \text{et} \quad \left|\int_a^t f - \mathrm{L}\right| \leq \varepsilon,\ \forall t \in [c_r, b[.$$

Pour tout $k \in \mathbf{N}$, f est intégrable sur $\mathrm{I}_k = [c_{k-1}, c_k]$, et il existe donc une jauge δ_k sur I_k telle que, pour toute subdivision δ_k-fine $\mathcal{P}_k$ de I_k, on ait

$$\left|\mathrm{S}(f, \mathcal{P}_k) - \int_{c_k}^{c_{k+1}} f\right| \leq \frac{\varepsilon}{2^k}.$$

Quitte à diminuer les jauges δ_k, on peut supposer que

1. $\delta_1(c_0) \leq \frac{1}{2}(c_1 - c_0)$,
2. $\delta_{k+1}(c_k) \leq \min\{\delta_k(c_k), \frac{1}{2}(c_k - c_{k-1}), \frac{1}{2}(c_{k+1} - c_k)\}$,
3. $\delta_k(t) \leq \min\{\frac{1}{2}(t - c_{k-1}), \frac{1}{2}(c_k - t)\}, t \in]c_{k-1}, c_k[$.

La famille $\{[c_{k-1}, c_k[, k \geq 1\}$ est une partition de $[a, b[$, on peut donc définir la jauge δ sur $[a, b]$ par

$$\delta(t) = \delta_k(t) \text{ si } t \in [c_{k-1}, c_k[, \quad \text{et} \quad \delta(b) = b - c_r.$$

Soit $\mathcal{P}$ une subdivision δ-fine. Par construction, le point b est point de marquage pour le dernier intervalle $[x_{n-1}, b]$ de la subdivision. Par ailleurs, on a $c_r = b - \delta(b) \leq x_{n-1}$, car $[xn-1, b] \subset [b - \delta(b)/2, b]$. Si donc s est le premier entier pour lequel $c_s \geq x_{n-1}$, alors $s \geq r$. Par hypothèse sur δ, tout intervalle de $\mathcal{P}$ contenant un c_k admet c_k pour point de marquage. Quitte à faire un coupage–collage le long des ponts de marquage, on peut supposer que les points $c_0, \ldots, c_{s-1}$ sont aussi des bornes d'intervalles de $\mathcal{P}$. On peut donc restreindre $\mathcal{P}$ aux intervalles $[c_{k-1}, c_k]$, et l'on pose

$$\mathcal{Q}_1 = \mathcal{P} \cap [c_0, c_1], \ldots, \mathcal{Q}_{s-1} = \mathcal{P} \cap [c_{s-2}, c_{s-1}], \mathcal{Q}_s = \mathcal{P} \cap [c_{s-1}, x_{n-1}], \mathcal{Q}_{s+1} = \{([x_{n-1}, b], b)\}.$$

Chaque subdivision $\mathcal{Q}_k$, $k = 1, \ldots, s-1$, est δ_k-fine, et ainsi

$$\left|\mathrm{S}(f, \mathcal{Q}_k) - \int_{c_{k-1}}^{c_k} f\right| \leq \frac{\varepsilon}{2^k}.$$

Enfin, $\mathcal{Q}_s$ est une *sous-subdivision* δ_s-fine de $[c_{s-1}, c_s]$, et le lemme de Henstock implique donc que

$$\left|\mathrm{S}(f, \mathcal{Q}_s) - \int_{c_{s-1}}^{x_{n-1}} f\right| \leq \frac{\varepsilon}{2^s}.$$

Pour $\mathcal{Q}_{s+1}$, on a $|\mathrm{S}(f, \mathcal{Q}_{s+1})| = |f(b)|(b - x_{n-1})$, qui, par construction de r, est plus petit que ε. Comme on a

$$\mathrm{S}(f, \mathcal{P}) = \mathrm{S}(f, \mathcal{Q}_1) + \cdots + \mathrm{S}(f, \mathcal{Q}_s) + \mathrm{S}(f, \mathcal{Q}_{s+1}),$$

il vient

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{L}| \leq \left|\sum_{k=1}^{s} \mathrm{S}(f, \mathcal{Q}_k) - \int_a^{x_{n-1}} f\right| + |\mathrm{S}(f, \mathcal{Q}_{s+1})| + \left|\int_a^{x_{n-1}} f - \mathrm{A}\right| \leq 3\varepsilon.$$

Celà achève la démonstration du théorème.

11.2 Méthode d'Euler

Nous n'avons pas abordé, dans ce cours, le problème du calcul numérique approché des intégrales. Le cours de troisième année est fait pour celà, et vous y verrez plusieurs méthodes de calcul approché d'intégrales. Nous présentons cependant ici l'algorithme le plus naïf qui soit pour calculer une approximation d'une primitive : la *méthode d'Euler.* Celle-ci est en effet très proche de la théorie des jauges de Kurzweil et Henstock, dont elle est en quelque sorte l'avatar numérique. On pourrait très bien, pour introduire l'intégrale, partir de la méthode d'Euler et montrer comment, en la poussant dans ses derniers retranchements théoriques, elle mène à l'intégrale que nous avons présenté ici. Le problème est le suivant : étant donné une fonction f définie sur $[a, b]$, peut-on déterminer une approximation d'une solution au problème suivant :

$$(\dagger)\begin{cases} \mathrm{F}'(t) = f(t), t \in [a, b] \\ \mathrm{F}(a) = y_0, \end{cases}$$

où y_0 est la condition initiale ou condition au bord du problème $(\dagger)$. Comme il s'agit de calcul numérique, nous ne pouvons calculer une solution approché que sur un nombre fini de points. Ces points $x_0 = a < x_1 < \cdots < x_\mathrm{N} = b$ forment une subdivision de $[a, b]$. On pose $h_i = x_{i+1} - x_i$, $\mathrm{H} = \max\{h_i, i = 0, \ldots, \mathrm{N} - 1\}$ et $h = \min\{h_i\}$. Une bonne méthode devrait assurer que si H tend vers 0, alors toute solution approchée calculée sur les points x_i approxime de mieux en mieux les valeurs d'une vraie solution. Enfin, pour des raisons pratiques de précision et de temps de calcul, nous devons écrire un algorithme où h est minoré par une constante, petite, dépendant de la machine sur laquelle on travaille. L'idée va être de construire pas-à-pas les points x_i, en les prenant plus proche quand l'erreur risque d'augmenter, et plus éloignés quand on est sûr que l'erreur diminue. Mettons tout celà en ordre.

Nous supposerons dorénavant f dérivable, même si elle peut, par endroit, présenter de brusques sautes d'humeur. Soit Ψ une solution exacte de $(\dagger)$, et supposons-nous donnée la subdivision $\{x_i\}$. Un développement de Taylor de Ψ au second ordre donne

$$\Psi(x_{i+1}) = \Psi(x_i + h_i) = \Psi(x_i) + h_i f(x_i) + \frac{1}{2} h_i^2 f'(x_i) + o(h_i^2).$$

La méthode d'Euler consiste à prendre pour solution approchée la suite des valeurs donnée par

$$\begin{cases} \mathrm{F}(0) = y_0 \\ \mathrm{F}(x_{i+1}) = \mathrm{F}(x_i) + h_i f(x_i). \end{cases}$$

Dans la pratique, on ne peut pas prendre exactement $\mathrm{F}(0) = y_0$, mais nous verrons plus loin que ce problème n'en est pas vraiment un.

On appelle *erreur de consistance* de la méthode à l'étape i, le réel

$$\mathrm{e}_i = \Psi(x_{i+1}) - \mathrm{F}(x_{i+1}),$$

si l'on suppose que $\mathrm{F}(x_i) = \Psi(x_i)$. En reportant dans le développement de Taylor ci-dessus, on trouve

$$\mathrm{e}_i = \frac{h_i^2}{2} f'(x_i) + o(h_i)^2.$$

Ainsi, plus f varie violemment au voisinage de x_i (*i.e.* $|f'(x_i)|$ grand) plus il va falloir prendre h_i petit et donc augmenter le nombre de points d'évaluation, et ainsi le temps de calcul. En revenant à la définition de e_i, on a

$$\mathrm{e}_i = \Psi(x_{i+1}) - \Psi(x_i) - h_i f(x_i),$$

et le théorème des accroissements finis nous assure de l'existence d'un $c_i \in [x_i, x_{i+1}]$ tel que

$$\mathrm{e}_i = h_i f(c_i) - h_i f(x_i) = h_i[f(c_i) - f(x_i)].$$

Comme f est continue sur l'intervalle compact $[a, b]$, elle y est uniformément continue, et donc, si $\varepsilon > 0$ est fixé, il existe $\delta > 0$ tel que, si $\mathrm{H} < \delta$, alors

$$\sum_{i=1}^{\mathrm{N}-1} |\mathrm{e}_i| \leq \sum_{i=1}^{\mathrm{N}-1} h_i |f(c_i) - f(x_i)| \leq \varepsilon(b-a).$$

On a montré la proposition suivante :

Proposition 11.2. — *La méthode d'Euler est* CONSISTANTE,*, en d'autres termes, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que, si* $\mathrm{H} < \delta$*, alors*

$$\sum |\mathrm{e}_i| \leq \varepsilon.$$

Dit autrement : si la méthode converge, elle converge vers une vraie solution. Reste maintenant à savoir si, avec des erreurs d'arrondi successives ou une erreur dans la condition initiale influent sur le résultat trouvé : c'est le problème de la STABILITÉ de la méthode.

En effet, dans la pratique, on fait, à chaque étape (et dès la condition initiale) des erreurs d'arrondis. Supposons nous donnée une suite d'erreurs ε_i, et donnons-nous une solution, perturbée par ces erreurs, de l'algorithme d'Euler :

$$\left\{ \tilde{\mathrm{F}}(x_0) = \tilde{y}_0 = y_0 + \varepsilon_0 \tilde{\mathrm{F}}(x_{i+1}) = \tilde{\mathrm{F}}(x_i) + h_i f(x_i) + \varepsilon_i. \right.$$

On aimerait savoir si $\tilde{\mathrm{F}}$ diffère beaucoup de F, et donc évaluer

$$\tilde{\mathrm{E}} = \max_i |\tilde{\mathrm{F}}(x_i) - \mathrm{F}(x_i)|.$$

Mais $|\tilde{\mathrm{F}}(x_{i+1}) - \mathrm{F}(x_{i+1})| \leq |\tilde{\mathrm{F}}(x_i) - \mathrm{F}(x_i)| + |\varepsilon_i|$, et donc

$$\tilde{\mathrm{E}} \leq |\tilde{y}_0 - y_0| + \sum_{i=1}^{\mathrm{N}-1} |\varepsilon_i|,$$

et l'on a ainsi démontré la

Proposition 11.3. — *La méthode d'Euler est* STABLE, *de constante de stabilité* $\mathrm{S} = 1$ *: pour toute solution perturbée* $\tilde{\mathrm{F}}$ *comme ci-dessus, on a*

$$\max_i |\tilde{\mathrm{F}}(x_i) - \mathrm{F}(x_i)| \leq \mathrm{S}\left(|\tilde{y}_0 - y_0| + \sum_{i=1}^{\mathrm{N}-1} |\varepsilon_i|\right).$$

On peut maintenant conclure cette étude générale de la mnéthode d'Euler, en montrant qu'elle est CONVERGENTE :

Théorème 11.2. — *Pour toute solution exacte* Ψ *de* (†) *et toute solution approchée donnée par la méthode d'Euler :*

$$\begin{cases} \mathrm{F}(0) = \tilde{y}_0 \\ \mathrm{F}(x_{i+1}) = \mathrm{F}(x_i) + h_i f(x_i), \end{cases}$$

l'errreur globale E *vérifie*

$$\mathrm{E} = \max_i |\mathrm{F}(x_i) - \Psi(x_i)| \to 0$$

quand $|\tilde{y}_0 - y_0|$ *et* H *tendent vers* 0. *On dit que la méthode est* CONVERGENTE.

Cela découle immédiatement du fait que la méthode est consistante et stable.

Finissons cette section par une possible mise en pratique sur machine. Comme nous l'avons vu lors du calcul de l'erreur de consistance e_i, il faudrait adapter à chaque étape le pas h_i de notre subdivision pour contrer les possibles variations brusques de f. Les conditions de mise en pratique (précision de la machine, temps de calcul...) nous fixent un intervalle de contrôle $[h_{\min}, h_{\max}]$ dans lequel le pas devra se trouver. Fixons-nous par ailleurs $\varepsilon > 0$, qui va être la précision que l'on exige sur la solution approchée. On doit commencer par pouvoir prendre $\tilde{y}_0$ à moins de $\varepsilon/(b-a)$ de y_0. Ensuite, on choisit h_{i+1} de telle sorte que

$$\mathrm{e}_i^* = \frac{1}{2}(f(x_{i+1}) - f(x_i)) \approx \mathrm{e}_i$$

vérifie

$$\frac{\mathrm{e}_i^*}{h_i} \leq \frac{\varepsilon}{b-a}.$$

On obtient ainsi une erreur globale majorée par ε. Remarquons que comme on calcule les $f(x_i)$, le calcul de e_i^* ne perturbe pas beaucoup le temps de moulinage du programme.

Si l'erreur e_i^* est petite, on augmente le pas, si elle est grande, on le diminue, et si elle explose en imposant que h_{i+1} sorte de l'intervalle de contrôle, on arête le programme. On pourrait par exemple faire comme suit :

– si $\varepsilon/(3(b-a)) \leq |\mathrm{e}_i^*| \leq \varepsilon/(b-a)$, on pose $h_{i+1} = h_i$.

– si $|\mathrm{e}_i^*| < \varepsilon/(3(b-a))$, on pose $h_{i+1} = \max\{5h_i/4, h_{\max}\}$.

– si $|\mathrm{e}_i^*| > \varepsilon/(b-a)$, on pose $h_{i+1} = 4h_i/5$,

et si dans la procédure on obtient $h_{i+1} < h_{\min}$, on arrête tout.

Nous Avons regroupé ci-dessous quelques références autour de l'intégrale de Kurzweil et Henstock. On y trouve d'abord les articles originaux de Denjoy, Perron, Kurzweil et Henstock. Enfin, quelques livres traitant ce sujet. Un seul, celui de Mawhin, est en français. Il existe par ailleurs de très nombreux cours d'analyse pour la licence, livres que vous trouverez facilement à la bibliothèque universitaire ou en librairie. On trouve, d'occasion, d'excellents livres destinés aux DEUG, aux premiers cycles universitaires, ou aux classes préparatoires. C'est un moyen peu onéreux d'acquérir des livres de mathématiques qui, malgré leur age, gardent toute leur valeur.

Références

[1] Robert G. Bartle. *A Modern Theory Of Integration*, tome 32 de *Graduate Studies in Mathematics* (American Mathematical Society, Providence, RI), 2001.

[2] Arnaud Denjoy. *Une extension de l'intégrale de m. lebesgue*. Comptes Rendus Acad. Sci. Paris, **154**, p. 859–862, 1912.

[3] Ralph Henstock. *The efficiency of convergence factors for functions of a continuous real variable*. J. London Math. Soc., **30**, p. 273–286, 1955.

[4] Jaroslaw Kurzweil. *Generalized ordinary differential equations and continuous dependance on a parameter*. Czech. Math. J., **7 (82)**, p. 418–446, 1957.

[5] Jean Mawhin. *Analyse. Fondements, Techniques, Évolutions* (De Broeck Université, Bruxelles), deuxième édition, 1997.

[6] Robert M. McLeod. *The Generalized Riemann Integral*, tome 20 de *Carus Monograph* (Mathematical Association Of America, Washington, DC), 1980.

[7] Oskar Perron. *Über den integralbegriff*. Sitzber. Heidelberg Akad. Wiss., Math.-Naturw. Klasse Abt. A, **16**, p. 1–16, 1914.

[8] Lee Peng Yeng et Rudolf Výborný. *The Integral. An Easy Approach after Kurzweil And Henstock* (Cambridge University Press, Cambridge, UK), 2000.