

Théorie élémentaire de l'intégration : l'intégrale de Kurzweil-Henstock

Jean-Pierre Demailly

Université Joseph Fourier Grenoble I

<http://www-fourier.ujf-grenoble.fr/~demailly/books.html>

version du 12 août 2011

Introduction

L'objectif de ce texte est de proposer une « trame » pour l'enseignement de l'intégration depuis le lycée jusqu'aux premières années de l'université. Pour cela, nous développons les bases de la théorie de l'intégration telles qu'elles ont été posées et grandement éclaircies par Jaroslav Kurzweil et Ralph Henstock à la fin des années 1950. Au fil des chapitres, le niveau mathématique de l'exposition glisse progressivement d'un niveau très élémentaire jusqu'à un niveau de second cycle universitaire.

Même si au début une partie des résultats doit être admise ou démontrée de manière heuristique du fait des contraintes de temps ou des prérequis, nous pensons que la démarche mathématique utilisée dans l'enseignement doit respecter chaque fois que cela est possible les principes d'une progression par généralisations successives compatibles avec les exposés qui ont précédé (« progression concentrique »). Il est donc très utile d'introduire dès le début des définitions qui sont susceptibles d'être rendues rigoureuses et formalisées, et d'adopter un ordre de présentation des concepts et une articulation logique tels que la théorie n'aura pas à être substantiellement changée le jour où viendra la formalisation complète.⁽¹⁾

-
- (1) À l'heure actuelle, l'intégration est souvent abordée en plusieurs étapes qui sont les suivantes :
- a) En terminale, une première étape qui consiste à introduire l'intégrale comme « aire sous la courbe », avec un état d'esprit qui se ressent obligatoirement de l'appauvrissement continu de la conceptualisation depuis 2 décennies. Il en résulte qu'il est devenu très difficile de formaliser complètement la théorie, ce qui veut probablement dire qu'on ne peut guère espérer que le cours donne de véritables démonstrations, mais seulement au mieux quelques indications de preuves complétées par des considérations heuristiques.
 - b) Dans les premières années d'université, les enseignants présentent en général une version plus ou moins édulcorée de la théorie de « l'intégrale de Riemann », à l'aide d'encadrements par des fonctions en escalier, et des preuves qui tendent de plus en plus à disparaître du fait du recul des connaissances fondamentales requises (pratique des ε et δ , continuité uniforme, ...)
 - c) En Master 1ère année apparaîtra dans les bons cas une théorie plus solide de l'intégration reposant sur la théorie de la mesure et l'intégrale de Lebesgue. Mais il est de notoriété publique que ce sujet qui fâche a tendance aujourd'hui à être de moins en moins traité, surtout dans les filières dites « Master enseignement ».
- Cette situation présente de nombreux inconvénients. La théorie de l'intégrale de Riemann n'est pas une « bonne théorie », ni du point de vue didactique ni du point de vue mathématique. La présentation n'en est pas très simple : il faut manipuler constamment des encadrements de fonctions, utiliser la continuité uniforme pour démontrer l'intégrabilité des fonctions continues, faire des découpages de ε et δ parfois peu éclairants. Il y a de nombreuses restrictions ou pathologies, des théorèmes essentiels comme ceux de la convergence monotone ne sont pas valables, etc. Plus tard, l'introduction de l'intégrale de Lebesgue viendra balayer ce travail en montrant qu'il s'agissait en fait d'une théorie bancale et incomplète. Si les étudiants échappent à l'intégrale de Lebesgue – comme cela arrive à un nombre de plus en plus grand de CAPESiens – ils n'auront donc jamais eu l'occasion de se voir exposer une théorie « sérieuse » de l'intégration, ce qui est préoccupant.

Le choix de l'intégrale de Kurzweil-Henstock présente l'avantage de fournir des définitions assez simples – peut-être plus simples que celle de Riemann puisque les encadrements de fonctions ne sont plus nécessaires, que l'on n'a plus besoin de la continuité uniforme, que tous les théorèmes de base se démontrent en quelques lignes – et, en même temps, d'être assez puissante pour contenir les parties élémentaires de la théorie de Lebesgue ... Dans ces conditions, il paraît quelque peu anachronique que la théorie n'ait pas encore trouvé sa juste place dans l'enseignement !

Nous suivons ici d'assez près le livre « Introduction to gauge integrals » [Sw] de Charles Swartz, en l'allégeant autant que faire se peut, pour atteindre très rapidement la preuve des théorèmes fondamentaux (rapport entre intégration et primitive, intégration par parties, changement de variable, dérivabilité des intégrales indéfinies de fonctions continues...). Nous développons ensuite les résultats plus spécifiques à la théorie de Kurzweil-Henstock pour atteindre les théorèmes de convergence monotone et dominée. L'existence de la mesure de Lebesgue et ses propriétés fondamentales figurent parmi les conséquences directes, sans qu'il y ait besoin d'introduire au préalable le langage général des tribus et des mesures dénombrablement additives. Les étudiants auront donc en fait un premier exemple motivant sous la main lorsque ces notions plus générales seront introduites. Enfin, les définitions ad hoc des diverses intégrales généralisées, impropres et semi-convergentes deviennent également accessoires, puisque toutes ces notions peuvent être exprimées en une seule définition naturelle contenant les cas utiles : l'intégrale de Kurzweil-Henstock autorise par nature même la « semi-convergence » ...

Les premières étapes nous paraissent éventuellement utilisables au lycée, à condition d'admettre quelques-unes des démonstrations – et en prenant comme perspective que c'est l'intégrale de Kurzweil-Henstock qui sera développée ensuite, de sorte que les énoncés présentant des hypothèses artificielles superflues n'ont pas lieu d'être.

Le cours qui suit a été à l'origine inspiré par des notes synthétiques rédigées par Eric Charpentier [Ch] à l'Université de Bordeaux autour de 2002, et fait des emprunts à de multiples sources (cf. bibliographie). Ces notes ont été ensuite développées sous forme de cours photocopié par Jean-Yves Briend à Marseille [Br] (après que je l'ai informé des suggestions d'Eric Charpentier). Le présent texte a fait l'objet de plusieurs rédactions successives depuis l'automne 2005, et a lui-même inspiré ultérieurement des cours ou manuels mis en chantier par plusieurs collègues. Je voudrais dans ce cadre signaler d'utiles remarques et questions formulées par Xavier Buff, auteur du chapitre sur l'intégration pour le L2 dans la collection de manuels "Licence-Tout-en-Un" dirigée par Jean-Pierre Ramis et André Warusfel [RW], et remercier ces deux derniers pour leur intérêt et leurs encouragements.

Table des matières

Introduction	1
Table des matières	3

Chapitre I

Définitions et résultats fondamentaux. Cas des fonctions d'une variable

1. Sommes de Riemann	5
2. Définition de l'intégrale d'une fonction	9
3. Propriétés élémentaires de l'intégrale	14
4. Le théorème fondamental de l'Analyse	18
5. Méthodes de calcul des primitives et des intégrales	22
6. Calcul d'aires et de volumes	30
7. Encadrement par des fonctions en escalier et intégrabilité	33
8. Convergence uniforme, continuité et dérivabilité en fonction de paramètres	41
9. Intégrales à valeurs vectorielles et formule de Taylor	46
10. Vrai ou faux	51
11. Exercices	54

Chapitre II

Calcul des intégrales en plusieurs variables

1. Intégration sur un pavé fermé borné	63
2. Intégration sur un domaine ouvert et formule du jacobien en plusieurs variables	70
3. Intégration des formes différentielles et formule de Stokes	77

Chapitre III

Théorèmes généraux de convergence, espaces L^p et fonctions mesurables

1. Lemme de Henstock et théorème de Hake	95
2. Fonctions absolument intégrables	104
3. Le théorème de convergence monotone	106
4. Mesure de Lebesgue et ensembles négligeables	108
5. Lemme de Fatou et théorème de convergence dominée	111
6. Lemme de recouvrement de Vitali et différentiabilité des intégrales indéfinies	114
7. Ensembles et fonctions mesurables	118
8. Densité des fonctions continues à support compact	125
9. Variation totale, fonctions à variation bornée	131
10. Fonctions de puissance k-ième intégrable	137
11. Caractérisation de l'intégrabilité au sens de Riemann	141
12. Fonctions approximativement continues et primitives	142

Chapitre IV

Compléments historiques

1. Archimède et la quadrature de la parabole	151
2. La famille ibn Qurra.	152

3. Calcul numérique des intégrales	153
4. Les différentes intégrales.	155
5. La mésaventure de Chasles	156
Références bibliographiques	159

Chapitre I

Définitions et résultats fondamentaux

Cas des fonctions d'une variable

L'objectif de ce chapitre est de proposer une présentation rigoureuse et complète des premiers éléments de la théorie de l'intégration. Compte tenu de cette ambition, l'exposé est nécessairement théorique, et donc beaucoup plus exigeant que d'autres textes ou manuels qui se contentent de donner des techniques de calcul, et qui s'appuient seulement sur une intuition de la notion d'aire en lieu et place de justifications mathématiques complètes. Un prérequis indispensable est d'avoir déjà assimilé l'art de couper les ε en quatre – ou d'être prêt à faire l'effort de creuser la question. Le public visé est celui des élèves de Terminale très motivés – la théorie de base ne fait jamais appel à aucune notion qui dépasse le niveau du lycée. La plupart des notes de bas de page sont destinées à des lecteurs plus avancés et ne peuvent normalement pas être comprises par des personnes qui aborderaient la théorie pour la première fois ; de même, les passages marqués d'un * ou de deux ** peuvent être omis en première lecture ; les sections 7, 8 et 9 sont nettement moins élémentaires et relèvent de l'université.

1. Sommes de Riemann

Dans toute cette section, a et b désignent deux réels tels que $a < b$. Soit f une fonction, définie sur l'intervalle $[a, b]$, à valeurs dans \mathbb{R} . La portion du plan comprise entre le graphe de f et l'axe horizontal est l'ensemble des couples (x, y) tels que

$$0 \leq y \leq f(x) \text{ si } f(x) \geq 0, \quad f(x) \leq y \leq 0 \text{ si } f(x) \leq 0.$$

Pour une fonction f suffisamment régulière, nous souhaitons évaluer l'aire A de cette portion de plan, en comptant positivement les surfaces situées au-dessus de l'axe horizontal, et négativement celles situées au-dessous (Fig. 1). Nous parlerons de « l'aire algébrique »⁽²⁾ située sous le graphe de f .

(2) L'approche des intégrales par les aires nous paraît infiniment préférable à celle qui consiste à introduire a priori l'intégrale par le calcul des primitives, ne serait-ce que parce que cette façon de voir dépouille l'intégrale de son sens géométrique (et qu'en outre elle escamote l'unique façon de démontrer en général l'existence des primitives...)

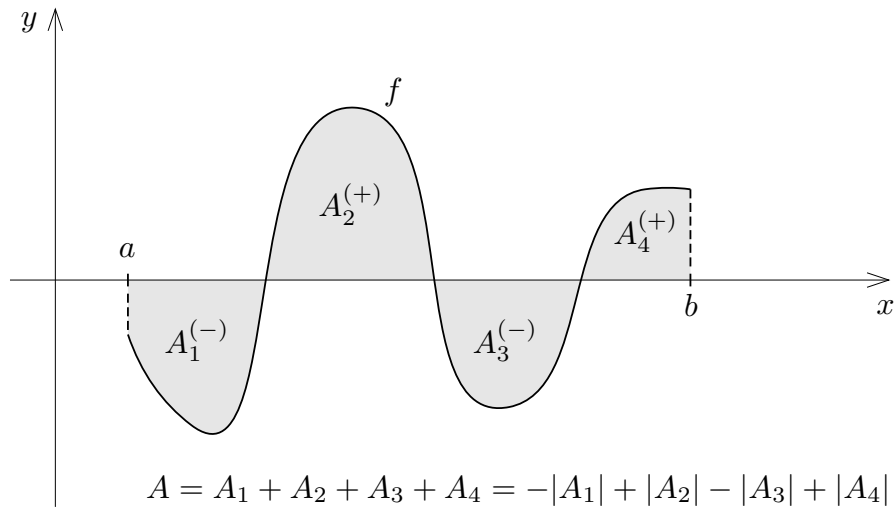


Fig. 1. Aire algébrique située sous le graphe de f .

L'idée est de découper l'intervalle $[a, b]$ au moyen d'une subdivision en sous-intervalles $[a_j, a_{j+1}]$, puis de sommer les aires de rectangles basés sur les intervalles $[a_j, a_{j+1}]$. La figure 2 ci-dessous résume le procédé.

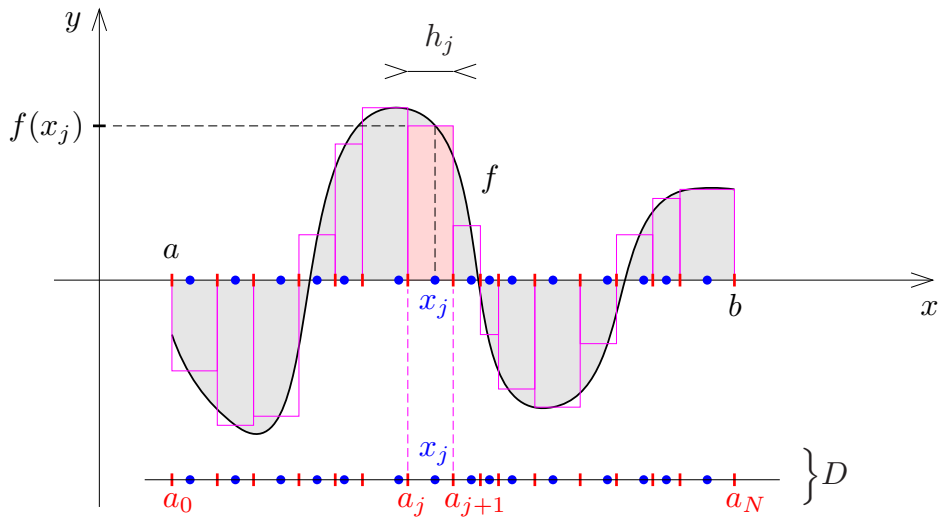


Fig. 2. Somme de Riemann associée à f sur D .

La somme des aires des rectangles figurés ci-dessus est donnée par $\sum_{j=0}^{N-1} (a_{j+1} - a_j) f(x_j)$, et on est ainsi conduit à poser la définition suivante.

(1.1) Définition.

(a) On appelle « subdivision pointée » de l'intervalle $[a, b]$ la donnée de $N + 1$ points

$$a = a_0 < a_1 < \dots < a_{N-1} < a_N = b$$

et de N points x_0, x_1, \dots, x_{N-1} tels que

$$\forall j = 0, 1, \dots, N-1, \quad x_j \in [a_j, a_{j+1}].$$

Elle sera notée

$$D = \{([a_j, a_{j+1}], x_j)\}_{0 \leq j < N}.$$

Les réels $h_j = a_{j+1} - a_j$ (amplitudes des intervalles) sont les « pas » de la subdivision.

- (b) Soit D une subdivision pointée de l'intervalle $[a, b]$ et f une fonction de $[a, b]$ dans \mathbb{R} . On appelle « somme de Riemann » associée à f sur D , le réel

$$S_D(f) = \sum_{j=0}^{N-1} (a_{j+1} - a_j) f(x_j) = \sum_{j=0}^{N-1} h_j f(x_j).$$

La somme de Riemann $S_D(f)$ est l'aire algébrique de la réunion des rectangles de largeur h_j et de hauteur $f(x_j)$ (Fig. 2). Il s'agit bien d'une aire algébrique, puisque $h_j f(x_j)$ est compté positivement si $f(x_j) > 0$ et négativement si $f(x_j) < 0$.

Intuitivement l'aire A cherchée est la limite de $S_D(f)$ quand les pas h_j tendent vers 0. Un choix possible consiste par exemple à prendre une subdivision en sous-intervalles égaux

$$a_j = a + jh = a + j \frac{b-a}{N}, \quad 0 \leq j \leq N \quad \text{où} \quad h_j = h = \frac{b-a}{N},$$

que l'on peut combiner avec un choix quelconque des points x_j .

(1.2) Exemple. Comme premier exemple, considérons la fonction identité $f(x) = x$ sur l'intervalle $[0, 1]$. Pour $N \geq 1$, posons :

$$a_0 = 0, \quad a_1 = \frac{1}{N}, \dots, \quad a_j = \frac{j}{N}, \dots, \quad a_N = 1$$

Les pas de cette subdivision sont tous égaux à $1/N$. Voici trois calculs de sommes de Riemann, selon que l'on place les points x_j au début, au milieu ou à la fin des intervalles $[a_j, a_{j+1}]$ (rappelons que la somme des N premiers entiers vaut $N(N+1)/2$).

$$(1.2 \text{ a}) \quad x_j = a_j : \quad S_D(f) = \sum_{j=0}^{N-1} \frac{1}{N} \frac{j}{N} = \frac{1}{N^2} \sum_{j=0}^{N-1} j = \frac{N-1}{2N},$$

$$(1.2 \text{ b}) \quad x_j = \frac{a_j + a_{j+1}}{2} : \quad S_D(f) = \sum_{j=0}^{N-1} \frac{1}{N} \frac{2j+1}{2N} = \frac{1}{2N^2} \sum_{j=0}^{N-1} 2j+1 = \frac{1}{2},$$

$$(1.2 \text{ c}) \quad x_j = a_{j+1} : \quad S_D(f) = \sum_{j=0}^{N-1} \frac{1}{N} \frac{j+1}{N} = \frac{1}{N^2} \sum_{j=0}^{N-1} j+1 = \frac{N+1}{2N}.$$

La seconde somme est égale à $1/2$ pour tout N , les deux autres tendent vers $1/2$ quand N tend vers l'infini. L'aire du triangle sous le graphe de la fonction est bien $1/2$:

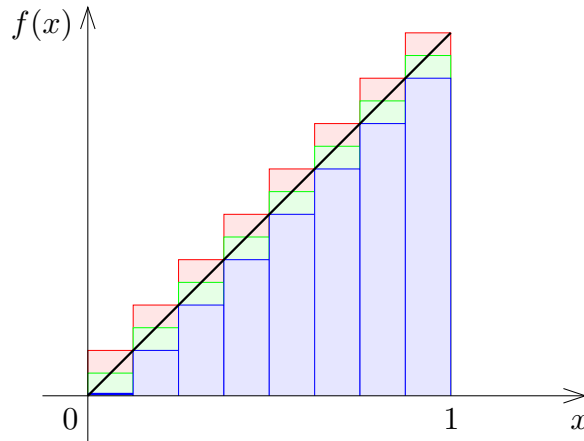


Fig. 3. Sommes de Riemann associées à l'identité sur $[0, 1]$.

(1.3) Exemple*. Nous considérons ici une situation où la fonction f n'est plus bornée. On prend $[a, b] = [0, 1]$ et f telle que $f(x) = 1/\sqrt{x}$ si $x \in]0, 1]$ et $f(0) = 0$. On introduit une subdivision pointée $D = \{([a_j, a_{j+1}], x_j)\}_{0 \leq j < N}$ telle que $a_j = (j/N)^2$ pour $0 \leq j \leq N$ et on pose $x_j = (t_j/N)^2$ avec $t_j \in [j, j+1]$. On a alors $a_{j+1} - a_j = (2j+1)/N^2$ et $f(x_j) = N/t_j$ si $t_j > 0$, d'où

$$S_D(f) = \frac{1}{N} \sum_{0 \leq j \leq N-1, t_j > 0} \frac{2j+1}{t_j}.$$

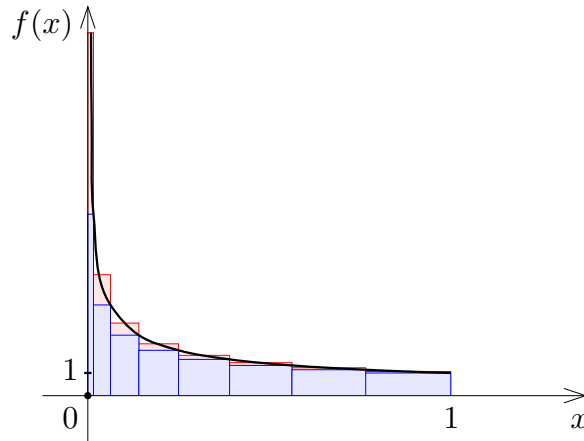


Fig. 4. Sommes de Riemann associées à $f(x) = 1/\sqrt{x}$ sur $[0, 1]$.

Le choix le plus simple est $t_j = j + \frac{1}{2}$, qui donne $\frac{2j+1}{t_j} = 2$ et donc $S_D(f) = 2$. Si on choisit plutôt $t_j = j + 1$, on obtient la valeur minimale possible pour $S_D(f)$, mais comme $\frac{2j+1}{j+1} \rightarrow 2$ quand $j \rightarrow +\infty$, il est facile de voir que l'on a encore $S_D(f) \rightarrow 2$ pour $N \rightarrow +\infty$ (on peut observer par exemple que $\frac{2j+1}{j+1} = 2 - \frac{1}{j+1} \geq 2 - 1/\sqrt{N}$ pour $\sqrt{N} \leq j \leq N$). Comme $a_{j+1} - a_j \leq (2N-1)/N^2 \rightarrow 0$ quand $N \rightarrow +\infty$, ce calcul amène à penser que l'aire du domaine non borné défini par $0 < x \leq 1$ et $0 \leq y \leq 1/\sqrt{x}$ est bien finie et égale à 2 (ce qui reste toutefois à démontrer). \square

Dans la suite, nous aurons besoin pour des raisons à la fois théoriques et pratiques de considérer des sommes de Riemann sur des subdivisions arbitraires. Il est facile de voir à partir de la définition 1.1 (b) que ces sommes vérifient les propriétés suivantes.

(1.4) Propriétés fondamentales.

- (a) Linéarité. Si $f, g : [a, b] \rightarrow \mathbb{R}$ sont des fonctions quelconques et λ, μ des constantes réelles, alors

$$S_D(\lambda f + \mu g) = \lambda S_D(f) + \mu S_D(g).$$

- (b) Monotonie. Si $f, g : [a, b] \rightarrow \mathbb{R}$ sont des fonctions quelconques, alors

$$f \geq g \quad \Rightarrow \quad S_D(f) \geq S_D(g).$$

En particulier, si $f \geq 0$, alors $S_D(f) \geq 0$.

- (c) Formule de Chasles. Soient $a < b < c$ des réels et f une fonction définie sur $[a, c]$. Si D_1 est une subdivision pointée de $[a, b]$ et D_2 une subdivision pointée de $[b, c]$, alors $D_1 \cup D_2$ est une subdivision pointée de $[a, c]$ et

$$S_{D_1 \cup D_2}(f) = S_{D_1}(f) + S_{D_2}(f).$$

2. Définition de l'intégrale d'une fonction

La première idée qui vient à l'esprit est de considérer des subdivisions pointées $D = \{([a_j, a_{j+1}], x_j)\}_{0 \leq j < N}$ dont les pas $h_j = a_{j+1} - a_j$ sont tels que $0 < h_j \leq \delta$ avec δ tendant vers 0, et de regarder si les sommes de Riemann de Riemann $S_D(f)$ convergent bien vers une limite A . Cette limite sera alors interprétée comme étant l'aire cherchée. On aboutit à la définition suivante, qui est la définition historiquement introduite par Cauchy et Riemann.

(2.1) Définition. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction quelconque. On dit que f est intégrable au sens de Riemann (ou Riemann-intégrable) s'il existe un réel A qui représente l'aire algébrique située sous le graphe de f , tel que pour toute marge d'erreur $\varepsilon > 0$ donnée a priori, on peut trouver un réel $\delta > 0$ tel que pour toute subdivision pointée $D = \{([a_j, a_{j+1}], x_j)\}$ de $[a, b]$ on ait

$$h_j = a_{j+1} - a_j \leq \delta \quad \Rightarrow \quad |S_D(f) - A| \leq \varepsilon.$$

Le nombre réel A de la définition précédente est appelé intégrale de f sur $[a, b]$ et noté

$$A = \int_a^b f(x) dx.$$

et on dit que $\int_a^b f(x) dx$ est la limite des sommes de Riemann $S_D(f)$, lorsque le pas de la subdivision tend vers 0.

Dans cette définition, on peut par exemple se contenter de prendre une subdivision en N sous-intervalles de pas constant $h = \frac{b-a}{N}$, et on obtient alors la conséquence immédiate suivante.

(2.2) Convergence des sommes de Riemann. Si $f : [a, b] \rightarrow \mathbb{R}$ est une fonction intégrable au sens de Riemann, on a

$$\begin{aligned} \int_a^b f(x) dx &= \lim_{N \rightarrow +\infty} \frac{b-a}{N} \sum_{j=0}^{N-1} f\left(a + \frac{j}{N}(b-a)\right) && [\text{cas } x_j = a_j] \\ &= \lim_{N \rightarrow +\infty} \frac{b-a}{N} \sum_{j=1}^N f\left(a + \frac{j}{N}(b-a)\right) && [\text{cas } x_j = a_{j+1}] \\ &= \lim_{N \rightarrow +\infty} \frac{b-a}{N} \sum_{j=0}^{N-1} f\left(a + \frac{2j+1}{2N}(b-a)\right) && \left[\text{cas } x_j = \frac{a_j + a_{j+1}}{2}\right]. \end{aligned}$$

Si l'aire $\int_a^b f(x) dx$ située sous le graphe de f est connue d'une manière ou d'une autre, on peut alors en déduire la valeur des limites correspondantes ; il faut savoir pour cela que la fonction f est intégrable au sens de Riemann, on démontrera plus tard (cf. théorèmes 7.6 et 7.10) que c'est le cas si f est continue ou continue par morceaux.

(2.3) Exemple. Considérons la somme

$$\frac{1}{N^2} \left(\sqrt{1(N-1)} + \sqrt{2(N-2)} + \dots + \sqrt{(N-1)1} \right)$$

qui peut aussi s'écrire :

$$\frac{1}{N} \left(\sqrt{\frac{1}{N} \left(1 - \frac{1}{N}\right)} + \sqrt{\frac{2}{N} \left(1 - \frac{2}{N}\right)} + \dots + \sqrt{\left(1 - \frac{1}{N}\right) \frac{1}{N}} \right).$$

C'est une somme de Riemann associée à la fonction $x \rightarrow \sqrt{x(1-x)}$ sur l'intervalle $[0, 1]$. Sa limite, lorsque n tend vers $+\infty$, est égale à : $\int_0^1 \sqrt{x(1-x)} dx$. Or, le graphe $y = \sqrt{x(1-x)}$ de cette fonction est un demi-arc du cercle $x^2 - x + y = 0$, soit encore $(x - \frac{1}{2})^2 + y^2 = \frac{1}{4}$, c'est donc un demi-cercle de centre $\frac{1}{2}$ et de rayon $\frac{1}{2}$. La limite de la sommation est égale à l'aire du demi-disque, qui vaut $\pi/8$. \square

Cependant, on s'aperçoit assez vite que la définition de l'intégrabilité au sens de Riemann impose des restrictions assez gênantes sur la fonction f :

(2.4) Condition nécessaire. Toute fonction f Riemann-intégrable est bornée.

Démonstration. En effet, selon la définition 2.1, prenons $\delta > 0$ donnant une erreur au plus ε . Pour une subdivision D de pas constant $h \leq \delta$, à savoir $h = (b-a)/N$ avec $N \geq (b-a)/\delta$, nous devons avoir la majoration $|S_D(f) - A| \leq \varepsilon$, donc

$$|S_D(f)| = \left| \frac{b-a}{N} \sum_{0 \leq j < N} f(x_j) \right| \leq |A| + \varepsilon \Rightarrow \left| \sum_{0 \leq j < N} f(x_j) \right| \leq \frac{N}{b-a} (|A| + \varepsilon),$$

ceci pour tout choix des points $x_j \in [a_j, a_{j+1}]$. Choisissons pour l'un des points x_j un point $x \in [a, b]$ quelconque, et pour les autres les points a_j de la subdivision. Il vient alors

$$\forall x \in [a, b], \quad |f(x)| \leq \sum_{0 \leq j < N} |f(a_j)| + \frac{N}{b-a}(|A| + \varepsilon),$$

ce qui montre que f doit être bornée. \square

Cette restriction que f soit bornée est très gênante, puisqu'on a vu à l'exemple (1.3) qu'il existait des fonctions non bornées pour lesquelles l'aire située sous le graphe est finie et calculable sans difficulté. D'autre part, d'un point de vue concret de calcul numérique, on peut être amené à faire des calculs d'aires pour des fonctions bornées qui « oscillent plus » à certains endroits qu'à d'autres :

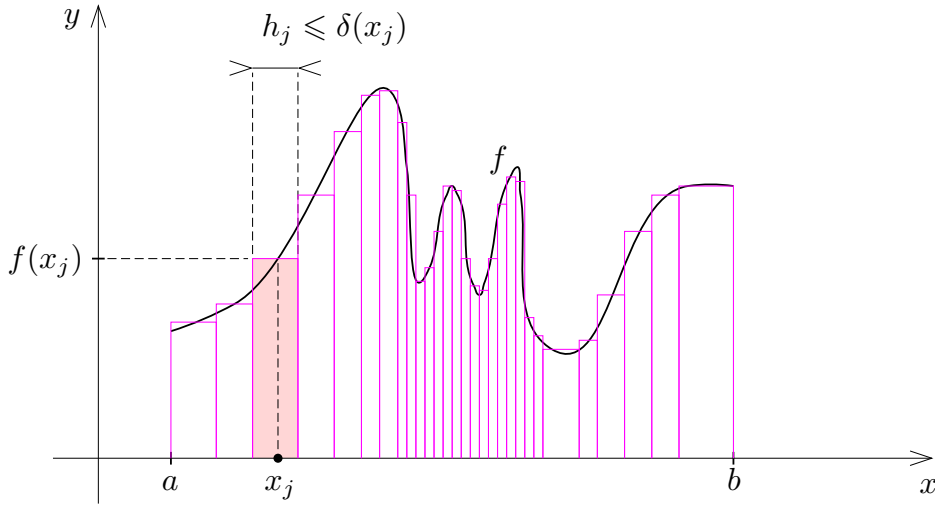


Fig. 5. Somme de Riemann à pas variable.

Dans ce cas, on sent bien intuitivement que l'on a intérêt à resserrer davantage les pas h_j aux endroits où f oscille davantage. Plutôt que de supposer que $h_j \leq \delta$ où δ est un réel positif fixé, il vaudra donc mieux demander que les pas h_j satisfassent une condition $h_j \leq \delta(x_j)$ où $\delta(x_j)$ est une quantité positive assez petite dépendant de l'endroit où l'on prend le rectangle de hauteur $f(x_j)$. On choisira alors des fonctions $\delta : [a, b] \rightarrow \mathbb{R}_+^*$ positives qui serviront à majorer les pas h_j . Une telle fonction sera appelée une *jauge* sur $[a, b]$.

(2.5) Définition. Soit $\delta : [a, b] \rightarrow \mathbb{R}_+^*$ une fonction positive quelconque. Une subdivision pointée $D = \{([a_j, a_{j+1}], x_j)\}_{0 \leq j < N}$ de $[a, b]$ sera dite δ -fine si

$$\forall j = 0, 1, \dots, N-1, \quad h_j = a_{j+1} - a_j \leq \delta(x_j).$$

Si δ_* et δ sont deux jauges telles que $\delta_* \leq \delta$, alors toute subdivision δ_* -fine est aussi δ -fine. Le résultat suivant, appelé *lemme de Cousin*, affirme que la définition précédente n'est jamais vide de contenu.

(2.6) Lemme. Soit $\delta : [a, b] \rightarrow \mathbb{R}_+^*$ une jauge. Alors il existe une subdivision pointée D de l'intervalle $[a, b]$ qui est δ -fine.

*Démonstration.** Elle est basée sur un procédé de dichotomie. Nous allons raisonner par l'absurde, en supposant que $[a, b]$ n'admet pas de subdivision pointée δ -fine. Posons $a_0 = a$ et $b_0 = b$. Divisons l'intervalle $[a_0, b_0]$ en deux, et considérons les deux moitiés $[a_0, \frac{a_0+b_0}{2}]$ et $[\frac{a_0+b_0}{2}, b_0]$: si chacune des deux admettait une subdivision δ -fine, la réunion de ces deux subdivisions serait une subdivision δ -fine de $[a_0, b_0]$. Donc l'une des deux moitiés au moins n'admet pas de subdivision δ -fine : on note celle-ci $[a_1, b_1]$.

On itère ensuite le procédé, de manière à construire des intervalles emboîtés $[a_k, b_k]$, de longueur $(b - a)/2^k$, dont aucun n'admet de subdivision δ -fine. Les suites (a_k) et (b_k) sont adjacentes par construction, donc elles convergent vers la même limite. Soit x_0 cette limite. Puisque $\delta(x_0) > 0$, il existe k_0 tel que

$$[a_{k_0}, b_{k_0}] \subset [x_0 - \frac{1}{2}\delta(x_0), x_0 + \frac{1}{2}\delta(x_0)].$$

Donc la subdivision pointée de $[a_{k_0}, b_{k_0}]$ formée seulement de l'intervalle tout entier et du point x_0 est δ -fine. D'où la contradiction.⁽³⁾ \square

(2.7) Définition. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction quelconque. On dit que f est intégrable au sens de Kurzweil-Henstock (ou KH-intégrable) s'il existe un réel A qui représente l'aire algébrique située sous le graphe de f , tel que pour toute marge d'erreur $\varepsilon > 0$ donnée a priori, on peut trouver une jauge $\delta : [a, b] \rightarrow \mathbb{R}_+^*$ en sorte que pour toute subdivision pointée $D = \{([a_j, a_{j+1}], x_j)\}$ de $[a, b]$ on ait

$$D \text{ } \delta\text{-fine, i.e. } \forall j, a_{j+1} - a_j \leq \delta(x_j) \Rightarrow |S_D(f) - A| \leq \varepsilon.$$

(une telle jauge δ sera dite ε -adaptée à f). Le nombre réel A de la définition précédente est appelé intégrale de f sur $[a, b]$, on écrit

$$\int_a^b f(x) dx = \lim_{\text{KH}, D} S_D(f) = \lim_{\text{KH}, D} \sum_{j=0}^{N-1} (a_{j+1} - a_j) f(x_j)$$

et on dit que $\int_a^b f(x) dx$ est la limite (au sens de Kurzweil-Henstock) des sommes de Riemann, lorsque la subdivision D devient de plus en plus fine⁽⁴⁾.

-
- (3) Le lecteur ayant quelques connaissances de topologie aura reconnu un cas élémentaire du théorème de Borel-Lebesgue relatif à la compacité du segment $[a, b]$. Cependant, notre but est de rester aussi élémentaire que possible – sans éluder les difficultés – donc il nous a paru préférable de faire la démonstration dans le cadre strict des subdivisions pointées. [On notera que celle-ci permet en fait d'énoncer un résultat un peu plus précis : pour toute jauge δ il existe une subdivision pointée δ -fine de $[a, b]$ formée d'intervalles dont les longueurs sont de la forme $(b - a)/2^{k_j}$. En effet, il suffit d'utiliser un raisonnement par l'absurde avec ce type d'intervalles pour aboutir à une contradiction.] Quoi qu'il en soit, la démonstration de 2.6 est probablement assez difficile pour la classe Terminale, c'est pourquoi nous l'avons marquée d'un astérisque *, comme toutes les parties plus délicates qui vont suivre.
- (4) L'intégrale de Kurzweil-Henstock est parfois appelée aussi *intégrale de jauge* ou encore *intégrale de Riemann généralisée*. Elle a été introduite au milieu des années 1950, alors que l'intégrale de Riemann remonte au 19ème siècle. Bien qu'en apparence la définition de Kurzweil-Henstock diffère très peu de celle de l'intégrale de Riemann, il se trouve qu'elle jouit de propriétés beaucoup meilleures, tout en procurant des démonstrations souvent plus simples et en éliminant beaucoup d'hypothèses superflues, par exemple le fait que f soit nécessairement bornée.

$\int_a^b f(x) dx$ se lit « somme de a à b de $f(x) dx$. ».

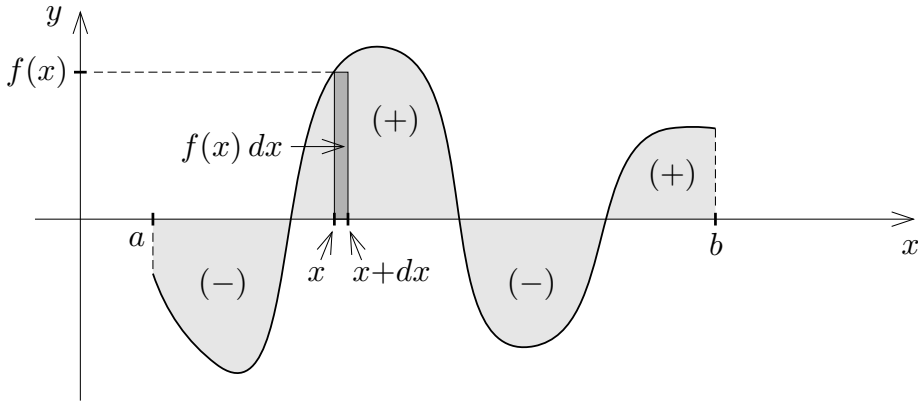


Fig. 6. Intégrale et son élément différentiel $f(x) dx$.

existe des points $(u_i)_{0 \leq i \leq p}$ de $[a, b]$ et des constantes réelles c_i telles que

$$a = u_0 < u_1 < \dots < u_p = b \quad \text{et} \quad f(x) = c_i \quad \text{sur }]u_i, u_{i+1}[,$$

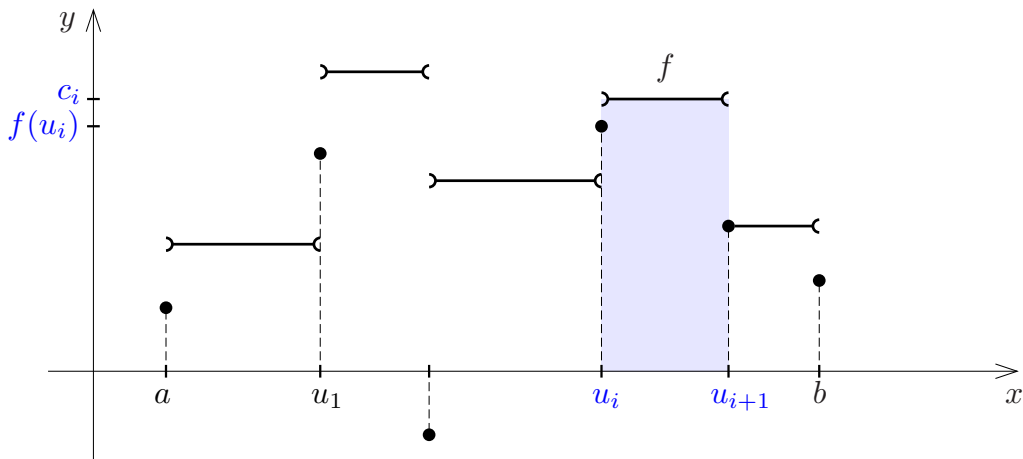
valeurs c_i .

Fig. 7. Une fonction en escalier.

Nous affirmons :

3. Propriétés élémentaires de l'intégrale

Les propriétés énoncées dans cette section sont, dans l'ordre, la linéarité, la monotonie et la relation de Chasles. Elles sont obtenues par passage à la limite à partir des propriétés analogues des sommes de Riemann $S_D(f)$ (propriétés 1.4 (a,b,c)), et valent pour l'intégrabilité au sens de Riemann et de Kurzweil-Henstock indifféremment.

(3.1) Linéarité. Si $f, g : [a, b] \rightarrow \mathbb{R}$ sont des fonctions intégrables sur $[a, b]$ et λ, μ des constantes réelles, alors $\lambda f + \mu g$ est intégrable sur $[a, b]$ et

$$\int_a^b (\lambda f(x) + \mu g(x)) dx = \lambda \int_a^b f(x) dx + \mu \int_a^b g(x) dx.$$

Démonstration. En termes de limites de sommes de Riemann sur des subdivisions pointées D de $[a, b]$, nous pouvons écrire

$$\begin{aligned} \int_a^b (\lambda f(x) + \mu g(x)) dx &= \lim_{\text{KH}, D} S_D(\lambda f + \mu g) = \lim_{\text{KH}, D} \lambda S_D(f) + \mu S_D(g) \\ &= \lambda \lim_{\text{KH}, D} S_D(f) + \mu \lim_{\text{KH}, D} S_D(g) \\ &= \lambda \int_a^b f(x) dx + \mu \int_a^b g(x) dx. \end{aligned} \quad \square$$

Pour analyser plus en détail cet argument, reprenons le calcul en termes de jauges, une marge d'erreur $\varepsilon > 0$ étant fixée a priori. Posons

$$A = \int_a^b f(x) dx, \quad B = \int_a^b g(x) dx.$$

Il existe par hypothèse des jauges δ_1, δ_2 telles que si D est δ_1 -fine alors $|S_D(f) - A| \leq \varepsilon$ et si D est δ_2 -fine alors $|S_D(g) - B| \leq \varepsilon$. Prenons une subdivision D δ -fine avec $\delta = \min(\delta_1, \delta_2)$. Comme $S_D(\lambda f + \mu g) = \lambda S_D(f) + \mu S_D(g)$ on en déduit

$$|S_D(\lambda f + \mu g) - (\lambda A + \mu B)| \leq (|\lambda| + |\mu|)\varepsilon.$$

Ceci montre que $\lambda f + \mu g$ est intégrable et que son intégrale est bien $\lambda A + \mu B$.

(3.2) Remarque. Si $f : [a, b] \rightarrow \mathbb{R}$ est une fonction qui est nulle partout sauf en un nombre fini de points $u_i \in [a, b]$, alors f est Riemann-intégrable d'intégrale nulle⁽⁵⁾. C'est en effet un cas très particulier de fonction en escalier, et on peut appliquer la formule de l'exemple 2.8 avec $c_i = 0$. Il en résulte que si deux fonctions $g, h : [a, b] \rightarrow \mathbb{R}$ diffèrent en un nombre fini de points, alors l'intégrabilité de l'une équivaut à l'intégrabilité de l'autre et on a $\int_a^b g(x) dx = \int_a^b h(x) dx$ (puisque $f = g - h$ est d'intégrale nulle). \square

⁽⁵⁾ Pour l'intégrabilité au sens de Kurzweil-Henstock, ce résultat est vrai plus généralement dans le cas où f est nulle partout en dehors d'un ensemble dénombrable de points $E \subset [a, b]$. Voir l'exercice 9.22.

(3.3) Monotonie. Si $f, g : [a, b] \rightarrow \mathbb{R}$ sont des fonctions intégrables sur $[a, b]$

$$f \geq g \quad \Rightarrow \quad \int_a^b f(x) dx \geq \int_a^b g(x) dx.$$

Démonstration. Cela résulte de l'inégalité sur les sommes de Riemann $S_D(f) \geq S_D(g)$, par passage à la limite. \square

(3.4) Relation de Chasles. Soient $a < b < c$ des réels et $f : [a, c] \rightarrow \mathbb{R}$ une fonction. Si f est intégrable sur $[a, b]$ et intégrable sur $[b, c]$, alors f est intégrable sur $[a, c]$ et

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx.$$

Ceci vaut aussi bien pour l'intégrabilité au sens de Kurzweil-Henstock que pour l'intégrabilité au sens de Riemann.

Démonstration. Fixons un réel $\varepsilon > 0$ et choisissons des jauges $\delta_1 : [a, b] \rightarrow \mathbb{R}_+^*$, $\delta_2 : [b, c] \rightarrow \mathbb{R}_+^*$ ε -adaptées à f sur $[a, b]$ et $[b, c]$ respectivement. Autrement dit, si on écrit

$$A_1 = \int_a^b f(x) dx, \quad A_2 = \int_b^c f(x) dx, \quad A = A_1 + A_2,$$

on aura $|S_{D_1}(f) - A_1| \leq \varepsilon$ et $|S_{D_2}(f) - A_2| \leq \varepsilon$ pour toutes subdivisions pointées δ_1 -fines D_1 de $[a, b]$ et δ_2 -fines D_2 de $[b, c]$. Si nous supposons démontrée l'intégrabilité de f sur $[a, c]$, c'est-à-dire l'existence de la limite $\lim_{\text{KH}, D} S_D(f)$ lorsque D parcourt les subdivisions pointées de $[a, c]$, alors on peut prendre une jauge δ sur $[a, c]$ telle que $\delta \leq \delta_1$ sur $[a, b]$ et $\delta \leq \delta_2$ sur $[b, c]$, et une subdivision $D = D_1 \cup D_2$ δ -fine égale à la réunion d'une subdivision D_1 δ_1 -fine de $[a, b]$ et d'une subdivision D_2 δ_2 -fine de $[b, c]$. On obtient dans ces conditions $S_D(f) = S_{D_1}(f) + S_{D_2}(f)$, donc $|S_D(f) - A| \leq 2\varepsilon$ et la relation désirée

$$\int_a^c f(x) dx = \lim_{\text{KH}, D} S_D(f) = A = A_1 + A_2 = \int_a^b f(x) dx + \int_b^c f(x) dx$$

s'ensuit en prenant ε arbitrairement petit. La seule difficulté qui reste est de démontrer l'existence de la limite $\lim_{\text{KH}, D} S_D(f)$ pour des subdivisions $D = \{([a_j, a_{j+1}], x_j)\}$ de $[a, c]$ ne comprenant pas nécessairement $a_j = b$ comme l'un des points intermédiaires, ce qui est a priori requis pour prouver l'intégrabilité de f sur $[a, c]$. La méthode consiste à redécouper l'intervalle $[a_j, a_{j+1}]$ de D qui contient le point b , et à estimer de combien on modifie ainsi la somme de Riemann $S_D(f)$.

(3.5)* Preuve de l'intégrabilité de f sur $[a, c]$ sous les hypothèses de 3.4. Dans le cas de l'intégrabilité au sens de Kurzweil-Henstock, la preuve est très simple. On définit une jauge δ sur $[a, c]$ en posant

$$\delta(x) = \begin{cases} \min(\delta_1(x), (b-x)/2) & \text{si } x \in [a, b[, \\ \min(\delta_2(x), (x-b)/2) & \text{si } x \in]b, c], \\ \min(\delta_1(b), \delta_2(b)) & \text{si } x = b. \end{cases}$$

de sorte que $\delta \leq \delta_1$ sur $[a, b]$ et $\delta \leq \delta_2$ sur $[b, c]$. Soit $D = \{([a_j, a_{j+1}], x_j)\}$ une subdivision δ -fine. Si $x_j < b$, alors on a $a_{j+1} \leq x_j + \delta(x_j) < x_j + (b - x_j) = b$. De même, si $x_j > b$, alors $a_j \geq x_j - \delta(x_j) > x_j - (x_j - b) = b$. Ceci montre que le seul cas où $[a_j, a_{j+1}]$ peut contenir le point b est le cas $x_j = b$, ce qui permet de découper l'intervalle pointé $([a_j, a_{j+1}], b)$ en les deux intervalles pointés $([a_j, b], b)$ et $([b, a_{j+1}], b)$. On produit ainsi une subdivision pointée D_1 δ_1 -fine de $[a, b]$ et une subdivision pointée D_2 δ_2 -fine de $[b, c]$ telles que $S_D(f) = S_{D_1}(f) + S_{D_2}(f)$. On a donc $|S_D(f) - A| \leq 2\varepsilon$ et la conclusion s'ensuit comme précédemment. \square

Dans le cas de l'intégrabilité au sens de Riemann, la preuve donnée ci-dessus n'est pas valable, car la jauge δ produite n'est pas constante. On sait cependant de plus que f est bornée, disons $|f| \leq M$ (si f est Riemann-intégrable sur $[a, b]$ et $[b, c]$, elle y est nécessairement bornée d'après la condition 2.4). Prenons pour D une subdivision pointée δ -fine avec $\delta \leq \min(\delta_1, \delta_2)$. Si l'un des intervalles $[a_j, a_{j+1}]$ contient b en son intérieur, on remplace son point de marquage x_j par $x'_j = b$, ce qui donne après découpage comme ci-dessus une nouvelle subdivision $D' = D_1 \cup D_2$ δ -fine telle que

$$S_{D'}(f) = S_{D_1}(f) + S_{D_2}(f) \Rightarrow |S_{D'}(f) - A| \leq 2\varepsilon.$$

On a de plus

$$|S_D(f) - S_{D'}(f)| = |h_j(f(x_j) - f(b))| \leq 2M\delta,$$

par conséquent $|S_D(f) - A| \leq 4\varepsilon$ dès que $\delta \leq \min(\delta_1, \delta_2, \varepsilon/M)$. Ceci entraîne bien l'intégrabilité de f au sens de Riemann sur l'intervalle $[a, c]$, ainsi que la formule de Chasles⁽⁶⁾. \square

- (6) On peut donner une démonstration alternative un peu plus subtile, valable à la fois pour l'intégrabilité au sens de Kurzweil-Henstock et pour l'intégrabilité au sens de Riemann. Pour cela, on fixe un réel positif θ assez petit et, avec les notations ci-dessus, on définit une jauge δ sur $[a, c]$ en posant

$$\delta(x) = \begin{cases} \min(\delta_1(x), \theta) & \text{si } x \in [a, b[, \\ \min(\delta_2(x), \theta) & \text{si } x \in]b, c], \\ \theta & \text{si } x = b, \end{cases}$$

de sorte que $\delta \leq \delta_1$ sur $[a, b]$, $\delta \leq \delta_2$ sur $[b, c]$ et $\delta(x) \leq \theta$ sur $[a, c]$ (la valeur précise de θ sera attribuée plus loin.) La difficulté qui se pose est d'étudier ce qui se passe au point de jonction $x = b$. Ceci fait l'objet du lemme suivant.

Lemme. *Supposons θ choisi tel que $\theta \leq \min(\delta_1(b), b - a)$. Si $x \in [a, b]$ est tel que $|x - b| \leq \delta(x)$, alors $\delta(x) |f(x) - f(b)| \leq 2\varepsilon$.*

En effet, choisissons grâce au lemme 2.6 une subdivision pointée δ_1 -fine D_* de l'intervalle $[a, b - \delta(x)]$ (notons que $b - \delta(x) \geq b - \theta \geq a$). On pose $D_1 = D_* \cup \{([b - \delta(x), b], x)\}$ et $D'_1 = D_* \cup \{([b - \delta(x), b], b)\}$. Ce sont bien des subdivisions pointées δ_1 -fines de $[a, b]$ du fait que $\delta(x) \leq \delta_1(x)$ et $\delta(x) \leq \theta \leq \delta_1(b)$. Or $S_{D_1}(f) - S_{D'_1}(f) = \delta(x) (f(x) - f(b))$ par définition des sommes de Riemann. Comme $|S_{D_1}(f) - A_1| \leq \varepsilon$ et $|S_{D'_1}(f) - A_1| \leq \varepsilon$, le lemme s'ensuit par différence. \square

Notons que par un raisonnement symétrique sur $[b, c]$, le lemme est également valable pour $x \in [b, c]$ si on suppose $\theta \leq \min(\delta_2(b), c - b)$. On choisira donc

$$\theta = \min(\delta_1(b), \delta_2(b), b - a, c - b).$$

Considérons maintenant une subdivision pointée δ -fine $D = \{([a_j, a_{j+1}], x_j)\}$ de $[a, c]$. Si D contient $a_j = b$ comme point intermédiaire, D induit une subdivision δ_1 -fine D_1 de $[a, b]$ et une subdivision δ_2 -fine D_2 de $[b, c]$ telles que $S_D(f) = S_{D_1}(f) + S_{D_2}(f)$; puisque $|S_{D_i}(f) - A_i| \leq \varepsilon$, $i = 1, 2$, on voit alors que $|S_D(f) - A| \leq 2\varepsilon$. Sinon, l'un des intervalles $[a_j, a_{j+1}]$ contient le point b en son intérieur, et comme $h_j = a_{j+1} - a_j \leq \delta(x_j) \leq \theta = \delta(b)$, on peut remplacer $([a_j, a_{j+1}], x_j)$ par l'intervalle pointé $([a_j, a_{j+1}], b)$. Ceci donne une nouvelle subdivision δ -fine D' de $[a, c]$ telle que

$$|S_D(f) - S_{D'}(f)| = (a_{j+1} - a_j) |f(x_j) - f(b)| \leq \delta(x_j) |f(x_j) - f(b)| \leq 2\varepsilon$$

Pour des réels a, b qui ne vérifient pas nécessairement $a < b$, on pose

$$(3.6) \quad \int_a^b f(x) dx = 0 \quad \text{si } a = b, \quad \int_a^b f(x) dx = - \int_b^a f(x) dx \quad \text{si } a > b.$$

On peut vérifier que la relation de Chasles reste valable dans tous les cas, quel que soit l'ordre des réels a, b, c , à condition bien sûr que f soit intégrable sur chacun des intervalles mis en jeu.

Nous souhaitons maintenant étudier l'intégrabilité en restriction à un sous-intervalle. Pour cela, nous avons besoin du critère d'intégrabilité de Cauchy, qui donne une condition nécessaire et suffisante d'intégrabilité sans qu'il soit nécessaire de faire intervenir la valeur A de l'intégrale.

(3.7) Critère de Cauchy.** Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction définie sur un intervalle $[a, b]$ fermé borné. Pour que f soit KH-intégrable (resp. Riemann-intégrable), il faut et il suffit que pour tout $\varepsilon > 0$ il existe une jauge $\delta : [a, b] \rightarrow \mathbb{R}_+^*$ (resp. une constante $\delta > 0$), telle que pour toutes subdivisions pointées D et D' δ -fines on ait $|S_{D'}(f) - S_D(f)| \leq \varepsilon$.

Démonstration. Si f est KH-intégrable d'intégrale A , pour chaque jauge δ qui est $\varepsilon/2$ -adaptée à f , les inégalités $|S_D(f) - A| \leq \varepsilon/2$ et $|S_{D'}(f) - A| \leq \varepsilon/2$ pour D, D' δ -fines entraînent $|S_{D'}(f) - S_D(f)| \leq \varepsilon$. La réciproque est une conséquence de la complétude de \mathbb{R} (cf. § II.1, et en particulier II.1.6, pour plus de détails). \square

Nous pouvons maintenant énoncer le théorème sur l'intégrabilité des restrictions.

(3.8) Théorème.** Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction KH-intégrable (resp. Riemann-intégrable). Alors la restriction $f|_{[c, d]}$ de f à tout intervalle $[c, d] \subset [a, b]$ est encore KH-intégrable (resp. Riemann-intégrable).

Démonstration. On utilise le critère de Cauchy. Étant donné $\varepsilon > 0$, il existe une jauge δ sur $[a, b]$ telle que $|S_D(f) - S_{D'}(f)| \leq \varepsilon$ pour toutes subdivisions pointées δ -fines D et D' de $[a, b]$. Soient maintenant Δ et Δ' deux subdivisions pointées de $[c, d]$ qui sont $\delta|_{[c, d]}$ -fines. En considérant grâce au lemme 2.6 des subdivisions de $[a, c]$ et de $[d, b]$ qui sont δ -fines, on peut compléter Δ et Δ' en des subdivisions pointées D et D' de $[a, b]$ qui sont elles-mêmes δ -fines. On obtient alors

$$|S_\Delta(f) - S_{\Delta'}(f)| = |S_D(f) - S_{D'}(f)| \leq \varepsilon.$$

Le théorème est démontré. Pour l'intégrabilité au sens de Riemann la preuve est tout à fait analogue. \square

d'après le lemme. En décomposant $([a_j, a_{j+1}], b)$ en les deux intervalles pointés $([a_j, b], b)$ et $([b, a_{j+1}], b)$, on obtient maintenant une subdivision δ_1 -fine D'_1 de $[a, b]$ et une subdivision δ_2 -fine D'_2 de $[b, c]$ telles que $S_{D'}(f) = S_{D'_1}(f) + S_{D'_2}(f)$, et par conséquent $|S_{D'}(f) - A| \leq 2\varepsilon$. Dans le pire des cas nous obtenons $|S_D(f) - A| \leq 4\varepsilon$, d'où $\lim_{KH, D} S_D(f) = A$. Ceci achève la preuve de l'intégrabilité de f au sens de Kurzweil-Henstock. Dans le cas de l'intégrabilité au sens de Riemann, on voit d'après ce qui précède qu'on peut choisir la jauge constante $\delta = \min(\delta_1, \delta_2, b - a, c - b)$.

4. Le théorème fondamental de l'Analyse

On appelle théorème fondamental de l'analyse, le fait que l'intégration (calcul d'aires) et la dérivation (calcul de tangentes et de différentielles) sont des opérations inverses l'une de l'autre.

(4.1) Théorème. *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction dérivable. Alors f' est intégrable au sens de Kurzweil-Henstock et*

$$\boxed{\int_a^b f'(x) dx = f(b) - f(a).}$$

Démonstration. Commençons par une preuve heuristique (c'est-à-dire simplifiée, mais non rigoureuse). Soit $D = \{[a_j, a_{j+1}], x_j\}$ une subdivision pointée assez fine. On a par définition

$$\int_a^b f'(x) dx \simeq \sum_{j=0}^{N-1} (a_{j+1} - a_j) f'(x_j).$$

Mais $f'(x_j)$ peut être vu comme une approximation de la pente de f sur l'intervalle $[a_j, a_{j+1}]$ et on a donc $(a_{j+1} - a_j) f'(x_j) \simeq f(a_{j+1}) - f(a_j)$, d'où

$$\int_a^b f'(x) dx \simeq \sum_{j=0}^{N-1} (f(a_{j+1}) - f(a_j)) = f(b) - f(a).$$

Il n'est pas difficile de rendre cet argument rigoureux. L'hypothèse de l'existence de $f'(x) = \lim_{y \rightarrow x} \frac{f(y) - f(x)}{y - x}$ signifie par définition que pour tout point $x \in [a, b]$ et tout $\varepsilon > 0$ il existe un réel $\delta(x) > 0$ tel que

$$y \in [a, b], \quad 0 < |y - x| \leq \delta(x) \Rightarrow \left| \frac{f(y) - f(x)}{y - x} - f'(x) \right| \leq \varepsilon, \quad \text{et donc}$$

$$y \in [a, b], \quad y \in [x - \delta(x), x + \delta(x)] \Rightarrow |f(y) - f(x) - (y - x)f'(x)| \leq \varepsilon |y - x|$$

(puisque l'inégalité est vraie aussi de manière évidente pour $y = x$). Prenons une subdivision pointée $D = \{[a_j, a_{j+1}], x_j\}$ δ -fine, c'est-à-dire telle que $h_j = a_{j+1} - a_j \leq \delta(x_j)$. En appliquant l'égalité ci-dessus à $x = x_j$ et $y = a_j$ ou $y = a_{j+1}$, il vient

$$\begin{aligned} |f(a_j) - f(x_j) - (a_j - x_j)f'(x_j)| &\leq \varepsilon |a_j - x_j| = \varepsilon (x_j - a_j), \\ |f(a_{j+1}) - f(x_j) - (a_{j+1} - x_j)f'(x_j)| &\leq \varepsilon |a_{j+1} - x_j| = \varepsilon (a_{j+1} - x_j). \end{aligned}$$

En faisant la différence, l'inégalité $|v - u| \leq |u| + |v|$ donne

$$|f(a_{j+1}) - f(a_j) - (a_{j+1} - a_j)f'(x_j)| \leq \varepsilon (a_{j+1} - a_j).$$

La sommation de ces inégalités pour $j = 0, 1, \dots, N - 1$ implique en définitive

$$|f(b) - f(a) - S_D(f')| \leq \varepsilon (b - a),$$

par conséquent

$$\int_a^b f'(x) dx = \lim_{\text{KH}, D} S_D(f') = f(b) - f(a). \quad \square$$

Notons qu'on retrouve ainsi le résultat important suivant qui, alternativement (et de manière plus classique) est démontré au moyen du théorème des accroissements finis.

(4.2) Corollaire. *Soit f une fonction dérivable sur un intervalle I de \mathbb{R} , telle que $f' = 0$ (resp. $f' \geq 0$, $f' \leq 0$). Alors f est constante (resp. croissante, décroissante).*

Démonstration. En effet, pour tous points $a < b$ dans I , on voit que $f(b) - f(a)$ est égal à 0 (resp. positif ou nul, négatif ou nul). \square

Le théorème fondamental peut aussi se s'énoncer comme une formule de calcul d'une intégrale à partir de la primitive d'une fonction.

(4.3) Calcul des intégrales au moyen de primitives. *Si $f : [a, b] \rightarrow \mathbb{R}$ admet une primitive F (c'est-à-dire si F est dérivable et $F' = f$), alors f est intégrable et*

$$\boxed{\int_a^b f(x) dx = F(b) - F(a) \quad \text{encore noté} \quad [F(x)]_a^b.}$$

Pour exploiter cette formule⁽⁷⁾, il convient de connaître une liste suffisante de primitives de fonctions usuelles – en connaître un certain nombre par cœur devient vite indispensable pour être en mesure de calculer les intégrales de manière efficace. Compte tenu de la formule 4.3, la primitive F d'une fonction continue f , là où elle est définie, sera notée sous la forme

$$(4.4) \quad F(x) = \int f(x) dx + C,$$

où C désigne une constante quelconque. Une écriture de la forme $\int f(x) dx$ est parfois appelée *intégrale indéfinie*. On sous-entend (contrairement à la notation utilisée pour une intégrale définie) que la variable x utilisée pour la fonction primitive $F(x)$ est la même que celle utilisée sous le signe \int .

(4.5) Liste de primitives usuelles. Voici une liste de primitives qui permet déjà de calculer un bon nombre d'intégrales usuelles. Les variables α , a , b représentent ici des

⁽⁷⁾ On notera que pour en arriver là, on a eu besoin de nettement moins de théorie que dans l'approche classique de l'intégrale de Riemann. En effet, dans cette approche, on est amené à supposer que f est continue – ce qui, outre le fait d'amener des hypothèses superflues de continuité, nécessite de démontrer (ou d'admettre) l'intégrabilité des fonctions continues au sens de Riemann. Tous les résultats du paragraphe 5 sont également affectés, comme la formule d'intégration par parties.

nombre réels quelconques avec $a \neq 0$.

$$\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C, \quad \alpha \neq -1$$

$$\int \ln |x| dx = x \ln |x| - x + C$$

$$\int \cos(ax+b) dx = \frac{1}{a} \sin(ax+b) + C$$

$$\int \frac{1}{\cos^2 x} dx = \tan x + C$$

$$\int \frac{1}{x^2+a^2} dx = \frac{1}{a} \arctan \frac{x}{a} + C$$

$$\int \frac{1}{x} dx = \ln |x| + C$$

$$\int e^{ax} dx = \frac{1}{a} e^{ax} + C$$

$$\int \sin(ax+b) dx = -\frac{1}{a} \cos(ax+b) + C$$

$$\int \frac{1}{\sin^2 x} dx = -\cotan x + C$$

$$\int \frac{1}{x^2-a^2} dx = \frac{1}{2a} \ln \left| \frac{x-a}{x+a} \right| + C$$

Voici une autre liste de fonctions un peu moins élémentaires

$$\int \frac{x}{\sqrt{a^2-x^2}} dx = -\sqrt{a^2-x^2} + C$$

$$\int \frac{x}{\sqrt{x^2+a^2}} dx = \sqrt{x^2+a^2} + C$$

$$\int \frac{x}{\sqrt{x^2-a^2}} dx = \sqrt{x^2-a^2} + C$$

$$\int \frac{x}{(ax^2+b)^{3/2}} dx = \frac{-1}{a(ax^2+b)^{1/2}} + C$$

$$\int \sinh(ax+b) dx = \frac{1}{a} \cosh(ax+b) + C$$

$$\int \frac{1}{\cosh^2 x} dx = \tanh x + C$$

$$\int \frac{1}{\sqrt{a^2-x^2}} dx = \arcsin \frac{x}{|a|} + C$$

$$\int \frac{1}{\sqrt{x^2+a^2}} dx = \ln(x + \sqrt{x^2+a^2}) + C$$

$$\int \frac{1}{\sqrt{x^2-a^2}} dx = \ln(x + \sqrt{x^2-a^2}) + C$$

$$\int \frac{1}{(ax^2+b)^{3/2}} dx = \frac{x}{b(ax^2+b)^{1/2}} + C$$

$$\int \cosh(ax+b) dx = \frac{1}{a} \sinh(ax+b) + C$$

$$\int \frac{1}{\sinh^2 x} dx = -\cotanh x + C.$$

(4.6) Compléments**.

(a) Si on souhaite énoncer le théorème 4.1 dans le cadre de la théorie de l'intégrale de Riemann, il convient d'ajouter l'hypothèse que la fonction dérivée f' soit intégrable au sens de Riemann. Dans ce cas, on peut aussi invoquer le théorème des accroissements finis pour démontrer la formule (4.1) : on sait qu'il existe $c_j \in]a_j, a_{j+1}[$ tel que $f(a_{j+1}) - f(a_j) = (a_{j+1} - a_j)f'(c_j)$ et donc par sommation $f(b) - f(a) = S_D(f')$ pour la subdivision pointée particulière $D = \{[a_j, a_{j+1}], c_j\}$; ceci montre bien à la limite que

$$f(b) - f(a) = \int_a^b f'(x) dx.$$

(b) Considérons l'exemple classique de la fonction $f : [0, 1] \rightarrow \mathbb{R}$ telle que

$$f(x) = x^2 \sin(x^{-n}) \quad \text{pour } x \neq 0, \quad \text{et } f(0) = 0.$$

Il est facile de voir que $f'(0) = \lim_{x \rightarrow 0} f(x)/x = 0$ et donc que f est partout dérivable sur \mathbb{R} . Cependant, pour $n \geq 2$, $f'(x) = 2x \sin(x^{-n}) - nx^{1-n} \cos(x^{-n})$ est non bornée au voisinage de 0 et donc non intégrable au sens de Riemann sur $[0, 1]$.⁽⁸⁾

(8) En fait, il est facile de voir que f n'est pas non plus intégrable au sens de Lebesgue sur $[0, 1]$ pour $n \geq 2$, puisque les intégrales $\int_0^1 x^{1-n} |\cos(x^{-n})| dx$ et $\int_0^1 |f'(x)| dx$ divergent. Pour que le théorème (4.1) soit valable en théorie de Lebesgue, il faut là encore faire l'hypothèse supplémentaire que f soit intégrable au sens de Lebesgue. C'est donc un résultat particulièrement impressionnant de la théorie de Kurzweil-Henstock.

(c) On démontrera plus loin (cf. (11.6 b)) qu'il existe une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ partout dérivable telle que f' prend des valeurs positives, négatives et nulles sur tout intervalle de \mathbb{R} , avec de plus $|f'| \leq 1$. Une telle fonction f n'est intégrable au sens de Riemann sur aucun intervalle $[a, b]$ de \mathbb{R} . En effet, si elle l'était, elle serait également intégrable sur tout sous-intervalle $[c, d] \subset [a, b]$, et en prenant des points de subdivisions x_j tels que $f'(x_j) > 0$ (resp. $f'(x_j) < 0$), on conclurait à la limite que $f(d) - f(c) = \int_c^d f'(x) dx$ est ≥ 0 et aussi ≤ 0 , ce qui impliquerait que f est constante sur $[a, b]$, en contradiction avec l'hypothèse sur la variation du signe de f' sur tout intervalle.⁽⁹⁾

5. Méthodes de calcul des primitives et des intégrales

Il convient d'observer qu'il est en général impossible d'expliciter en termes de fonctions usuelles la primitive de beaucoup de fonctions pourtant relativement simples. Ainsi on peut montrer que e^{x^2} ou $\ln(\sin x)$ ont des primitives qui ne peuvent pas s'exprimer en termes des fonctions trigonométriques, puissances, exponentielles, logarithmes et leurs composées. Lorsque le calcul est possible, on s'appuie le plus souvent sur l'une des deux formules importantes qui suivent.

(5.1) Formule d'intégration par parties. Soient $u, v : [a, b] \rightarrow \mathbb{R}$ deux fonctions dérivables. Le produit uv est alors dérivable et $(uv)' = u'v + uv'$. Par conséquent $uv' = (uv)' - u'v$ est intégrable si et seulement si $u'v$ l'est, et dans ce cas

$$[u(x)v(x)]_a^b = \int_a^b (uv)'(x) dx = \int_a^b u'(x)v(x) dx + \int_a^b u(x)v'(x) dx.$$

On a donc la formule

$$\boxed{\int_a^b u(x)v'(x) dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x) dx,}$$

qui peut encore se récrire de manière plus abrégée

$$\boxed{\int_a^b u dv = [uv]_a^b - \int_a^b v du.}$$

En termes de primitives et avec la notation des intégrales indéfinies, on peut aussi écrire

$$\boxed{\int u(x)v'(x) dx = u(x)v(x) - \int u'(x)v(x) dx.}$$

(5.1 a) Exemple. Soit $n \geq 1$ un entier naturel, on cherche à calculer une primitive F_n de $f_n(x) = x^n e^{-x}$. La fonction f_n est déjà écrite comme produit de deux fonctions, que

⁽⁹⁾ On peut démontrer aussi qu'il existe une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ continue et dérivable presque partout, telle que f' ne soit intégrable au sens de Lebesgue sur aucun intervalle de \mathbb{R} , cf. Proposition (11.8). Ces « défauts » de la théorie de Lebesgue avaient amené A. Denjoy [Dj1] à proposer en 1912 une théorie de « l'intégrale totale » qui lui permit de rétablir la validité du Théorème 4.1 (cf. aussi les travaux de O. Perron [Pe]). Cette théorie se révélera a posteriori essentiellement équivalente à la théorie de Kurzweil-Henstock, mais avec un formalisme et des justifications beaucoup plus compliquées.

l'on sait intégrer. Intégrons par parties en prenant $u(x) = x^n$ et $v'(x) = e^{-x}$ (d'où, par exemple, $v(x) = -e^{-x}$):

$$\begin{aligned} F_n(x) &= \int x^n e^{-x} dx = [x^n(-e^{-x})] - \int n x^{n-1}(-e^{-x}) dx \\ &= -x^n e^{-x} - n \int x^{n-1} e^{-x} dx = -x^n e^{-x} + n F_{n-1}(x) \quad (+ C). \end{aligned}$$

Comme $F_0(x) = -e^{-x}$, cette relation de récurrence permet de calculer F_n pour tout entier $n \geq 1$:

$$F_n(x) = -\left(\sum_{p=0}^n n(n-1)\dots(p+1)x^p\right)e^{-x} \quad (+ C).$$

(5.1 b) Exemple. De même, pour $a \in \mathbb{R} \setminus \{-1\}$ et $n \in \mathbb{N}$, en posant $u'(x) = x^a$ et $v(x) = (\ln x)^n$, on trouve

$$\begin{aligned} \int x^a (\ln x)^n dx &= \frac{x^{a+1}}{a+1} (\ln x)^n - \int \frac{x^{a+1}}{a+1} \frac{n}{x} (\ln x)^{n-1} dx \\ &= \frac{x^{a+1}}{a+1} (\ln x)^n - \frac{n}{a+1} \int x^a (\ln x)^{n-1} dx \\ &= \frac{x^{a+1}}{a+1} (\ln x)^n - \frac{n}{a+1} \frac{x^{a+1}}{a+1} (\ln x)^{n-1} + \frac{n(n-1)}{(a+1)^2} \int x^a (\ln x)^{n-2} dx. \end{aligned}$$

Par récurrence, ceci donne

$$\int x^a (\ln x)^n dx = x^{a+1} \sum_{p=0}^n (-1)^p \frac{n(n-1)\dots(n-p+1)}{(a+1)^{p+1}} (\ln x)^{n-p}.$$

(5.2) Formule de changement de variable. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction admettant une primitive F . On a alors

$$\int_a^b f(x) dx = [F(x)]_a^b = F(b) - F(a).$$

Il est souvent commode de changer de variable, en posant $x = \varphi(t)$ pour une certaine fonction $\varphi : [\alpha, \beta] \rightarrow [a, b]$ dérivable telle que $\varphi(\alpha) = a$ et $\varphi(\beta) = b$ (on ne suppose pas nécessairement $a \leq b$ ni $\alpha \leq \beta$). Il vient alors

$$\int_a^b f(x) dx = F(\varphi(\beta)) - F(\varphi(\alpha)) = [F \circ \varphi(t)]_\alpha^\beta = \int_\alpha^\beta (F \circ \varphi)'(t) dt = \int_\alpha^\beta f(\varphi(t)) \varphi'(t) dt$$

On a donc la formule fondamentale

$$\boxed{\int_a^b f(x) dx = \int_\alpha^\beta f(\varphi(t)) \varphi'(t) dt = \int_\alpha^\beta f \circ \varphi d\varphi.}$$

En pratique, on effectue les substitutions

$$\boxed{x = \varphi(t), \quad dx = \varphi'(t) dt, \quad a = \varphi(\alpha), \quad b = \varphi(\beta),}$$

en prenant soin de changer les bornes comme indiqué. En termes d'intégrales indéfinies, si F désigne une primitive de f , on peut écrire

$$\boxed{F(x) = F(\varphi(t)) = \int f(\varphi(t))\varphi'(t) dt.}$$

(5.2 a) Exemple. Le cas particulier $f(x) = \frac{1}{x}$, $F(x) = \ln |x|$ implique

$$\int \frac{\varphi'(t)}{\varphi(t)} dt = \ln |\varphi(t)| + C.$$

De même, le cas $f(x) = \frac{1}{1+x^2}$, $F(x) = \arctan x$ implique

$$\int \frac{\varphi'(t)}{1+\varphi(t)^2} dt = \arctan \varphi(t) + C.$$

Ceci permet par exemple de trouver successivement les formules

$$\begin{aligned} \int \tan x \, dx &= \int \frac{\sin x}{\cos x} \, dx = -\ln |\cos x| + C \\ \int \cotan x \, dx &= \int \frac{\cos x}{\sin x} \, dx = \ln |\sin x| + C \\ \int \frac{1}{\sin x \cos x} \, dx &= \int \left(\frac{\cos x}{\sin x} + \frac{\sin x}{\cos x} \right) dx = \ln |\tan x| + C \\ \int \frac{1}{\sin x} \, dx &= \int \frac{1}{2 \sin \frac{x}{2} \cos \frac{x}{2}} \, dx = \ln \left| \tan \frac{x}{2} \right| + C \\ \int \frac{1}{\cos x} \, dx &= \int \frac{1}{\sin(x + \frac{\pi}{2})} \, dx = \ln \left| \tan \left(\frac{x}{2} + \frac{\pi}{4} \right) \right| + C. \end{aligned}$$

Le cas des fractions rationnelles trigonométriques générales sera traité plus loin.

(5.2 b)* Un exemple plus élaboré. Soit à calculer la primitive

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

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

$$y = \sqrt[6]{1+x} = (1+x)^{1/6}.$$

Ainsi on a $\sqrt{1+x} = y^3$ et $\sqrt[3]{1+x} = y^2$. Comme $y^6 = 1+x$, il vient $6y^5 dy = dx$ et donc

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

Cette dernière intégrale se calcule facilement. On commence par effectuer une division euclidienne de y^3 par $y+1$, ce qui donne

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

et par conséquent

$$\begin{aligned} \frac{y^3}{y+1} &= y^2 - y + 1 - \frac{1}{y+1}, \\ \int \frac{y^3}{y+1} dy &= \frac{1}{3}y^3 - \frac{1}{2}y^2 + y - \ln|1+y| + C. \end{aligned}$$

Il ne faut pas oublier, à la fin du calcul, de revenir à la variable x en remplaçant partout y par sa valeur. Il sort alors de tout ça le résultat attendu :

$$\int \frac{1}{\sqrt{1+x} + \sqrt[3]{1+x}} dx = 2\sqrt{1+x} - 3\sqrt[3]{1+x} + 6\sqrt[6]{1+x} - 6\ln|1 + \sqrt[6]{1+x}| + C.$$

(5.3) Intégration des polynômes trigonométriques. Pour calculer l'intégrale d'un polynôme trigonométrique $\int P(\cos x, \sin x) dx$ on utilise ce qu'on appelle la *méthode de linéarisation*, qui consiste à trouver une combinaison linéaire de la forme

$$P(\cos x, \sin x) = a_0 + \sum_{1 \leq n \leq N} a_n \cos nx + b_n \sin nx.$$

Pour cela, on utilise les formules d'Euler

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}, \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i},$$

afin d'exprimer $P(\cos(x), \sin(x))$ comme combinaison linéaire des fonctions e^{inx} et e^{-inx} . On remplace finalement ces fonctions par $e^{\pm inx} = \cos nx \pm i \sin nx$ (ou bien, si P est réel, on prend la partie réelle du résultat. Par exemple

$$\begin{aligned} \cos^2 x \sin^3 x &= \left(\frac{e^{ix} + e^{-ix}}{2} \right)^2 \left(\frac{e^{ix} - e^{-ix}}{2i} \right)^3 \\ &= \frac{i}{2^5} (e^{2ix} + 2 + e^{-2ix}) (e^{3ix} - 3e^{ix} + 3e^{-ix} - e^{-3ix}) \\ &= \frac{i}{32} (e^{5ix} - e^{3ix} - 2e^{ix} + 2e^{-ix} + e^{-3ix} - e^{-5ix}) \\ &= \frac{1}{16} (-\sin 5x + \sin 3x + 2\sin x). \end{aligned}$$

On obtient alors

$$\int \cos^2 x \sin^3 x \, dx = \frac{1}{16} \left(\frac{1}{5} \cos 5x - \frac{1}{3} \cos 3x - 2 \cos x \right) + C.$$

Plus généralement, toute expression polynomiale $P(e^{a_j x}, \cos b_k x, \sin c_\ell x)$ se ramène à une combinaison linéaire d'exponentielles complexes, et peut donc se traiter d'une manière analogue.

(5.4) Intégration des fractions rationnelles. Une fraction rationnelle est une fonction définie comme le quotient P/Q de deux polynômes P et Q à coefficients réels ou complexes (en algèbre on s'intéresse aussi au cas où les coefficients sont pris dans un corps quelconque; le cas du corps \mathbb{Q} des rationnels est bien sûr très usuel). La fonction P/Q est définie sur \mathbb{C} privé des racines de Q , que l'on appelle les *pôles* de la fraction rationnelle. La division euclidienne de P par Q s'écrit $P = EQ + R$ avec un quotient E et un reste R qui sont des polynômes tels que $\deg R < \deg Q$. On a alors

$$\frac{P}{Q} = E + \frac{R}{Q}.$$

Le quotient E s'appelle aussi *la partie entière* de la fraction rationnelle, et R/Q la *partie polaire*. Elle est telle que $\lim_{x \rightarrow \infty} R(x)/Q(x) = 0$.

(5.4 a) Cas où P/Q possède un seul pôle. C'est le cas où $Q(x) = (x - \alpha)^m$. En écrivant R comme un polynôme en $X = x - \alpha$, soit $R(x) = \sum_{0 \leq j < m} c_j (x - \alpha)^j$, on voit que P/Q admet une écriture de la forme

$$\frac{P(x)}{Q(x)} = E(x) + \sum_{j=1}^m \frac{\lambda_j}{(x - \alpha)^j}.$$

Si α est réel, la primitive se calcule sans peine puisque

$$\int \frac{1}{x - \alpha} \, dx = \ln |x - \alpha| + C, \quad \int \frac{1}{(x - \alpha)^j} \, dx = -\frac{1}{j-1} \frac{1}{(x - \alpha)^{j-1}} + C \quad \text{si } j > 1.$$

(5.4 b) Cas où Q est un trinôme du second degré. Si $Q(x) = ax^2 + bx + c$, le calcul des racines réelles ou complexes donne une factorisation $Q(x) = a(x - \alpha)(x - \beta)$. Si les racines α, β sont réelles et distinctes, on écrit

$$\frac{1}{(x - \alpha)(x - \beta)} = \frac{1}{\alpha - \beta} \left(\frac{1}{x - \alpha} - \frac{1}{x - \beta} \right),$$

ce qui donne aussitôt

$$\int \frac{1}{(x - \alpha)(x - \beta)} \, dx = \frac{1}{\alpha - \beta} \ln \left| \frac{x - \alpha}{x - \beta} \right| + C.$$

Dans le cas d'un trinôme réel $ax^2 + bx + c$ de discriminant $\Delta = b^2 - 4ac < 0$, les racines sont complexes et on procède différemment. On a en effet

$$Q(x) = a \left(\left(x + \frac{b}{2a} \right)^2 - \frac{\Delta}{4a^2} \right) = a((x - \gamma)^2 + \delta^2)$$

avec $\gamma = -b/2a$ et $\delta = \sqrt{|\Delta|}/2|a|$, les racines complexes étant $\alpha = \gamma + i\delta$, $\bar{\alpha} = \gamma - i\delta$. Après division euclidienne de P par Q , on est ramené à intégrer des expressions de la forme

$$\int \frac{\lambda x + \mu}{(x - \gamma)^2 + \delta^2} dx.$$

Dans ce cas on écrit $\lambda x + \mu = \lambda(x - \gamma) + (\lambda\gamma + \mu)$ et on observe que $(x - \gamma)$ est la moitié de la dérivée du dénominateur, ce qui donne

$$\begin{aligned} \int \frac{\lambda x + \mu}{(x - \gamma)^2 + \delta^2} dx &= \int \frac{\lambda(x - \gamma)}{(x - \gamma)^2 + \delta^2} dx + \int \frac{\lambda\gamma + \mu}{(x - \gamma)^2 + \delta^2} dx \\ &= \frac{1}{2}\lambda \ln((x - \gamma)^2 + \delta^2) + (\lambda\gamma + \mu) \int \frac{1}{(x - \gamma)^2 + \delta^2} dx \\ &= \frac{1}{2}\lambda \ln((x - \gamma)^2 + \delta^2) + \frac{\lambda\gamma + \mu}{\delta} \arctan \frac{x - \gamma}{\delta} + C. \end{aligned}$$

(5.4 c) Cas général.** Un théorème fondamental d'algèbre affirme que tout polynôme non constant Q de degré d à coefficients complexes admet exactement d racines complexes comptées avec multiplicité, ce qui permet de le factoriser sous la forme $Q(x) = c \prod_{1 \leq j \leq s} (x - \alpha_j)^{m_j}$ avec $m_1 + \dots + m_s = d$. On va alors montrer par récurrence sur le degré d de Q que la partie polaire R/Q , $\deg R < \deg Q$, s'écrit comme combinaison linéaire de termes $1/(x - \alpha_j)^\ell$ n'ayant comme pôle qu'un seul des α_j à la fois :

$$\frac{R(x)}{Q(x)} = \sum_{j=1}^s \sum_{\ell=1}^{m_j} \frac{\lambda_{j,\ell}}{(x - \alpha_j)^\ell}.$$

Le résultat est vrai (et trivial) pour $d = 1$, puisque dans ce cas $Q(x) = c(x - \alpha)$ et R doit être une constante. D'après le raisonnement (5.4 a), le résultat est encore vrai si Q n'a qu'une seule racine. Supposons donc que Q ait au moins 2 racines distinctes α_1, α_2 , et cherchons à simplifier l'écriture de chaque terme

$$\frac{x^k}{(x - \alpha_1)^{m_1}(x - \alpha_2)^{m_2} \dots (x - \alpha_s)^{m_s}}, \quad k < d,$$

provenant des différents monômes x^k de $R(x)$. Pour cela, on effectue la substitution

$$\begin{aligned} \frac{1}{(x - \alpha_1)(x - \alpha_2)} &= \frac{1}{\alpha_2 - \alpha_1} \left(\frac{1}{x - \alpha_1} - \frac{1}{x - \alpha_2} \right) & \text{si } k = 0, \\ \frac{x}{(x - \alpha_1)(x - \alpha_2)} &= \frac{1}{\alpha_1 - \alpha_2} \left(\frac{\alpha_1}{x - \alpha_1} - \frac{\alpha_2}{x - \alpha_2} \right) & \text{si } k \geq 1, \end{aligned}$$

ce qui fournit deux termes dont dénominateur est de degré $d' = d - 1$ et dont le numérateur est de degré $< d'$. Par hypothèse de récurrence, ces termes se décomposent sous la forme voulue. Le résultat cherché s'ensuit d'après l'hypothèse de récurrence.

La décomposition de P/Q qui en résulte, c'est-à-dire

$$\frac{P(x)}{Q(x)} = E(x) + \sum_{j=1}^s \sum_{\ell=1}^{m_j} \frac{\lambda_{j,\ell}}{(x - \alpha_j)^\ell}$$

s'appelle *décomposition de P/Q en éléments simples*. Dans le cas où la fraction P/Q est réelle, il peut tout de même y avoir des racines complexes, mais si $c/(x - \alpha)^\ell$ est un terme complexe, on peut écrire que $P/Q = \operatorname{Re}(P/Q)$ et remplacer $c/(x - \alpha)^\ell$ par

$$\begin{aligned} \operatorname{Re} \left(\frac{c}{(x - \alpha)^\ell} \right) &= \frac{1}{2} \left(\frac{c}{(x - \alpha)^\ell} + \frac{\bar{c}}{(x - \bar{\alpha})^\ell} \right) = \frac{1}{2} \left(\frac{c(x - \bar{\alpha})^\ell + \bar{c}(x - \alpha)^\ell}{(x - \alpha)^\ell (x - \bar{\alpha})^\ell} \right) \\ &= \frac{A(x)}{((x - \gamma)^2 + \delta^2)^\ell} \end{aligned}$$

où $\alpha = \gamma + i\delta$ et où A est un polynôme réel de degré ℓ . En posant $y = x - \gamma$ et en faisant des divisions euclidiennes de A (exprimé comme polynôme de la variable y) par $y^2 + \delta^2$, on se ramène de manière évidente à une somme de termes de la forme

$$(*) \quad \frac{uy + v}{(y^2 + \delta^2)^j}, \quad j \leq \ell.$$

Ceci donne une *décomposition en éléments simples réels* du type

$$\frac{P(x)}{Q(x)} = E(x) + \sum_{j=1}^p \sum_{\ell=1}^{m_j} \frac{\lambda_{j,\ell}}{(x - \alpha_j)^\ell} + \sum_{j=1}^q \sum_{\ell=1}^{m'_j} \frac{u_{j,\ell}(x - \gamma_j) + v_{j,\ell}}{((x - \gamma_j)^2 + \delta_j^2)^\ell}$$

où les α_j sont les racines réelles de multiplicités m_j , et les $\gamma_j \pm i\delta_j$ les racines complexes (deux à deux conjuguées) de multiplicités m'_j . Pour calculer les primitives des termes $(*)$ de la dernière sommation, on observe qu'on a d'une part

$$\int \frac{y}{(y^2 + \delta^2)^\ell} dy = -\frac{1}{(2\ell - 2)} \frac{1}{(y^2 + \delta^2)^{\ell-1}} + C,$$

et d'autre part

$$\begin{aligned} \int \frac{1}{(y^2 + \delta^2)^\ell} dy &= \frac{1}{\delta^2} \int \frac{(y^2 + \delta^2) - y^2}{(y^2 + \delta^2)^\ell} dy \\ &= \frac{1}{\delta^2} \int \frac{1}{(y^2 + \delta^2)^{\ell-1}} dy - \frac{1}{\delta^2} \int \frac{y \cdot y}{(y^2 + \delta^2)^{\ell-1}} dy. \end{aligned}$$

Une intégration par parties fournit

$$\int \frac{y \cdot y}{(y^2 + \delta^2)^\ell} dy = -\frac{1}{2(\ell - 1)} \frac{y}{(y^2 + \delta^2)^{\ell-1}} + \frac{1}{2(\ell - 1)} \int \frac{1}{(y^2 + \delta^2)^{\ell-1}} dy.$$

En combinant ces égalités, on trouve

$$\int \frac{1}{(y^2 + \delta^2)^\ell} dy = \frac{1}{2(\ell - 1)\delta^2} \frac{y}{(y^2 + \delta^2)^{\ell-1}} + \frac{2\ell - 3}{2\ell - 2} \frac{1}{\delta^2} \int \frac{1}{(y^2 + \delta^2)^{\ell-1}} dy,$$

ce qui permet un calcul de la primitive par récurrence sur ℓ . On substitue *in fine* $y = x - \gamma$ pour revenir à l'expression cherchée en la variable x .

Le calcul pratique de la décomposition en éléments simples peut être pénible, et nous ne traiterons ici qu'un exemple.

(5.4 d)* Exemple. Soit à calculer l'intégrale indéfinie

$$F(x) = \int \frac{1}{x^3(1+x^3)} dx.$$

Le dénominateur ne s'annule qu'en 0 et -1 , donc la fonction à intégrer est continue sur $\mathbb{R} \setminus \{0, -1\}$, de sorte qu'elle y admet des primitives sur les trois intervalles $] -\infty, -1[$, $] -1, 0[$, $] 0, +\infty[$. On peut commencer par observer que

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

Pour le second terme on remarque que $1+x^3$ a pour racines dans \mathbb{C} les nombres complexes -1 , $e^{i\pi/3}$ et $e^{-i\pi/3}$. Ainsi

$$1+x^3 = (x+1)(x-e^{i\pi/3})(x-e^{-i\pi/3}) = (x+1)(x^2-x+1).$$

Les pôles de $1/(1+x^3)$ sont de multiplicité 1 (les racines de $1+x^3$ sont distinctes) – on parle alors de *pôles simples* – et donc la décomposition en éléments simples prend la forme

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

avec des nombres complexes a, b, c à déterminer. Pour trouver leurs valeurs, on observe que $a/(x+1)$ est un équivalent de la partie polaire quand $x \rightarrow -1$. Or, un équivalent de $1+x^3$ au voisinage d'une racine $x = \alpha$ est précisément $3\alpha^2(x-\alpha)$ puisque la dérivée vaut $3x^2$. Ceci donne pour $1/(1+x^3)$ l'équivalent $\frac{1}{3\alpha^2} \frac{1}{x-\alpha}$ et on en déduit

$$a = \frac{1}{3}, \quad b = \left(\frac{1}{3\alpha^2} \right)_{\alpha=e^{i\pi/3}} = -\frac{1}{3}e^{i\pi/3}, \quad c = \left(\frac{1}{3\alpha^2} \right)_{\alpha=e^{-i\pi/3}} = -\frac{1}{3}e^{-i\pi/3}.$$

On obtient par conséquent la décomposition en éléments simples

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

On aurait pu aussi chercher directement la décomposition en éléments simples réels sous la forme

$$\frac{1}{1+x^3} = \frac{a}{x+1} + \frac{bx+c}{x^2-x+1},$$

et procéder par identification à partir de l'égalité $1 = a(x^2-x+1) + (bx+c)(x+1)$ qui donne $a+c=1$, $-a+b+c=0$ et $a+b=0$. On aboutit ainsi à

$$\frac{1}{x^3(1+x^3)} = \frac{1}{x^3} - \frac{1}{3} \frac{1}{1+x} - \frac{1}{3} \frac{2-x}{x^2-x+1}.$$

Pour le dernier terme, on écrit

$$\frac{2-x}{x^2-x+1} = \frac{2-x}{(x-\frac{1}{2})^2 + \frac{3}{4}} = \frac{\frac{3}{2} - (x-\frac{1}{2})}{(x-\frac{1}{2})^2 + \frac{3}{4}}$$

de façon à faire apparaître la dérivée du dénominateur au numérateur. Ceci donne

$$\begin{aligned} \int \frac{2-x}{x^2-x+1} dx &= \frac{3}{2} \frac{1}{\frac{\sqrt{3}}{2}} \arctan \frac{x-\frac{1}{2}}{\frac{\sqrt{3}}{2}} - \frac{1}{2} \ln \left((x-\frac{1}{2})^2 + \frac{3}{4} \right) + C \\ &= \sqrt{3} \arctan \frac{2x-1}{\sqrt{3}} - \frac{1}{2} \ln (x^2-x+1) + C \end{aligned}$$

à l'aide de la méthode de (5.4 b). Au total nous trouvons

$$\int \frac{1}{x^3(1+x^3)} dx = -\frac{1}{2} \frac{1}{x^2} - \frac{1}{3} \ln |1+x| - \frac{1}{\sqrt{3}} \arctan \frac{2x-1}{\sqrt{3}} + \frac{1}{6} \ln(x^2-x+1) + C.$$

Comme on peut le voir, les calculs sont fastidieux, mais, dès que l'on connaît les pôles de la fraction rationnelle (i.e. les racines du dénominateur), ils deviennent automatiques.

(5.4 e)* Fractions rationnelles trigonométriques. Une méthode générale pour le calcul des primitives de fractions rationnelles $P(\cos x, \sin x)/Q(\cos x, \sin x)$ en sinus et cosinus consiste à utiliser les formules de trigonométrie suivantes : si l'on pose $t = \tan(x/2)$, alors

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

On obtient par ailleurs $x = 2 \arctan t$, d'où

$$dx = \frac{2 dt}{1+t^2}.$$

On en déduit que

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

C'est une intégrale de fraction rationnelle en t , et on sait donc en calculer une primitive d'après ce qui précède. Pour une fraction rationnelle en $\cosh x$ et $\sinh x$, on peut poser de même $t = \tanh x/2 \in]-1, 1[$, ce qui donne $dx = 2dt/(1-t^2)$ et

$$\cosh x = \frac{1+t^2}{1-t^2}, \quad \sinh x = \frac{2t}{1-t^2}.$$

6. Calcul d'aires et de volumes

Nous commencerons par une application concrète très importante qui est le calcul des aires et des volumes.

(6.1) Calcul d'aires⁽¹⁰⁾. Par définition même, l'intégrale d'une fonction f calcule l'aire algébrique située sous son graphe. Pour un domaine plan quelconque, la frontière est en général délimitée par les graphes de plusieurs fonctions (deux au moins, davantage si le domaine comporte des « trous » ou des « renforcements »).

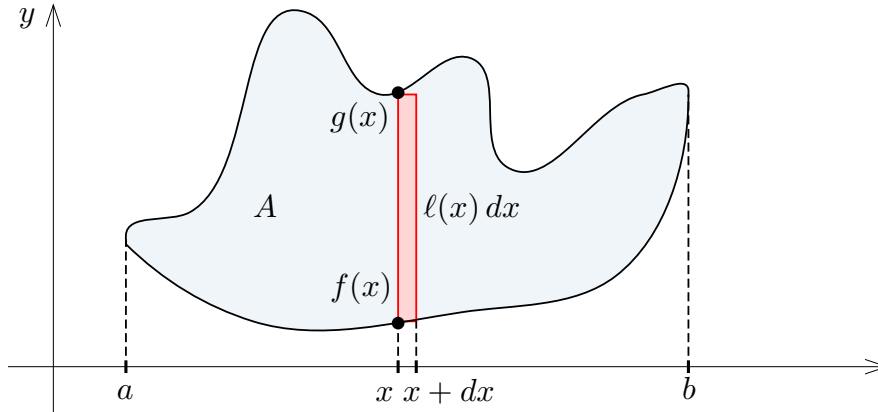


Fig. 9. Aire d'un domaine délimité par des graphes de fonctions.

Comme on le voit par soustraction des aires situées sous les graphes de f et g respectivement (et en remplaçant si nécessaire f , g par $f + C$, $g + C$ pour avoir des fonctions positives), l'aire du domaine plan est égale à la différence

$$A = \int_a^b g(x) dx - \int_a^b f(x) dx.$$

L'aire d'un tel domaine plan est donc donnée par l'intégrale par rapport à x de la longueur $\ell(x) = g(x) - f(x)$ des sections « verticales » du domaine :

$$A = \int_a^b \ell(x) dx = \int_a^b (g(x) - f(x)) dx.$$

À titre d'exemple, cherchons l'aire délimitée par l'ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, de demi-axes a et b . Le domaine délimité par l'ellipse est compris entre les graphes des fonctions $g(x) = b\sqrt{1 - (x/a)^2}$ et $f(x) = -g(x)$ pour $x \in [-a, a]$, ce qui donne

$$A = \int_{-a}^a 2g(x) dx = 2b \int_{-a}^a \sqrt{1 - \frac{x^2}{a^2}} dx.$$

Le changement de variable $x = a \sin t$ avec $t \in [-\pi/2, \pi/2]$ donne $dx = a \cos t dt$, d'où

$$A = 2ab \int_{-\pi/2}^{\pi/2} \cos^2(t) dt = 2ab \int_{-\pi/2}^{\pi/2} \frac{1 + \cos 2t}{2} dt = \pi ab. \quad \square$$

(6.2) Calcul de volumes⁽¹⁰⁾. La procédure est essentiellement la même : on effectue des sections planes (disons par exemple en coupant par un plan horizontal $z = \text{Cte}$), et on suppose connue l'aire $S(z)$ d'une telle section plane (cette aire aura en général été évaluée elle-même par un calcul intégral comme expliqué au paragraphe 6.1).

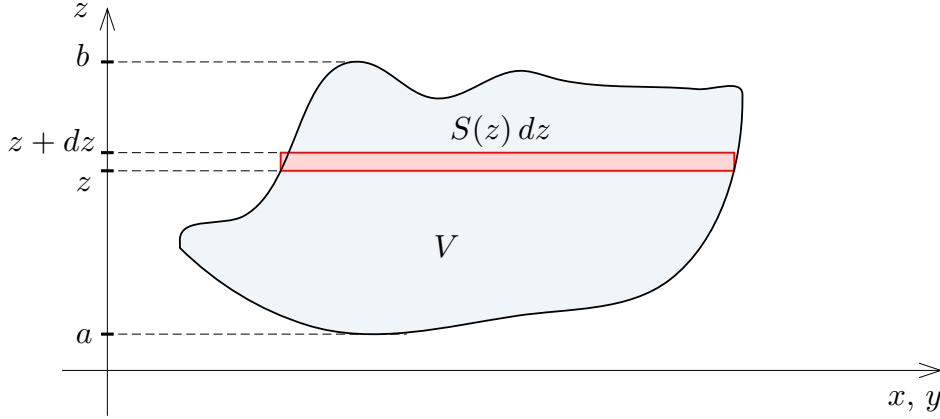


Fig. 10. Calcul d'un volume par tranchage dans la direction z .

Dans ce cas, le volume est donné par

$$V = \int_a^b S(z) dz.$$

Une application immédiate est le calcul du volume de la boule de rayon R d'équation $x^2 + y^2 + z^2 \leq R^2$. L'aire de la section $z = \text{Cte}$ est l'aire $S(z) = \pi(R^2 - z^2)$ du disque $x^2 + y^2 \leq R^2 - z^2$ de rayon $\sqrt{R^2 - z^2}$. On trouve donc

$$V = \int_{-R}^R \pi(R^2 - z^2) dz = 2\pi \int_0^R (R^2 - z^2) dz = 2\pi [R^2 z - z^3/3]_0^R = \frac{4}{3}\pi R^3.$$

Une autre application est la formule donnant le volume d'un cône – éventuellement « oblique » – ayant une base d'aire A connue et une hauteur h . On entend ici par cône la réunion des segments issus d'un point (sommet), ayant pour autre extrémité un point quelconque d'un domaine plan (choisi comme base). Pour faire le calcul, il est commode de choisir le sommet du cône comme origine et de fixer un repère orthonormé tel que la base du cône soit contenue dans le plan $z = h$.

(10) La définition de l'aire (et de même du volume), pour des domaines « mesurables » arbitraires de \mathbb{R}^2 et de \mathbb{R}^3 requiert des concepts sophistiqués de la théorie de l'intégration que nous ne pouvons pas développer ici. La démonstration complète des formules 6.1 et 6.2 s'appuie ainsi sur le célèbre théorème dit de Fubini. Nous renvoyons à la section 9 du chapitre III pour une preuve détaillée dans le cadre de la théorie des ensembles mesurables. Il n'est pas incorrect ici de considérer que les formules données pour la surface et le volume sont en quelque sorte des définitions de ces grandeurs.

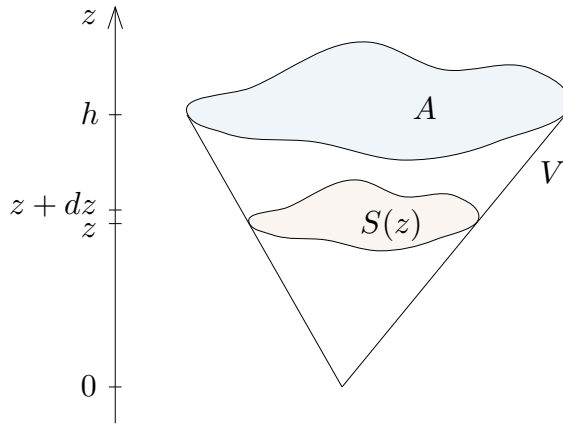


Fig. 11. Calcul du volume d'un cône.

Comme l'aire de l'homothétique de rapport λ d'un domaine plan est multipliée par λ^2 , et comme le domaine d'aire $S(z)$ est homothétique du domaine d'aire A dans le rapport $\lambda = z/h$, on a

$$S(z) = \frac{z^2}{h^2} A.$$

On en déduit que le volume du cône vaut

$$V = \int_0^h S(z) dz = \int_0^h \frac{z^2}{h^2} A dz = \left[\frac{1}{3} z^3 \right]_0^h \frac{1}{h^2} A,$$

ce qui prouve la formule classique

$$V = \frac{1}{3} Ah.$$

On peut appliquer le même raisonnement pour un cône sphérique, c'est-à-dire un cône ayant pour sommet le centre d'une sphère, et pour base un domaine de cette sphère (plutôt qu'un domaine plan). Si R est le rayon de la sphère et A l'aire du domaine sphérique, on trouve alors comme précédemment $V = \frac{1}{3} AR$.

Dans le cas où la base est la sphère toute entière, le cône est la boule elle-même, d'où $V = \frac{4}{3} \pi R^3$. On en déduit ainsi la valeur de l'aire de la sphère

$$A = 4\pi R^2.$$

7. Encadrement par des fonctions en escalier. Intégrabilité des fonctions continues ou monotones par morceaux⁽¹¹⁾

Nous nous proposons de démontrer ici que quelques classes usuelles de fonctions (notamment les fonctions continues ou monotones par morceaux) sont intégrables.

⁽¹¹⁾ À partir de ce point, le niveau théorique de l'exposé augmente sensiblement, nous considérons donc que cela relève plutôt de l'enseignement supérieur – le théorème (7.10) peut être admis d'emblée en Terminale si l'on souhaite ensuite être en mesure d'énoncer et de démontrer le corollaire 7.7.

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction bornée. L'idée est de considérer des encadrements

$$(7.1) \quad \varphi \leq f \leq \psi$$

de f par des fonctions en escalier φ, ψ , comme figuré sur le schéma ci-dessous.

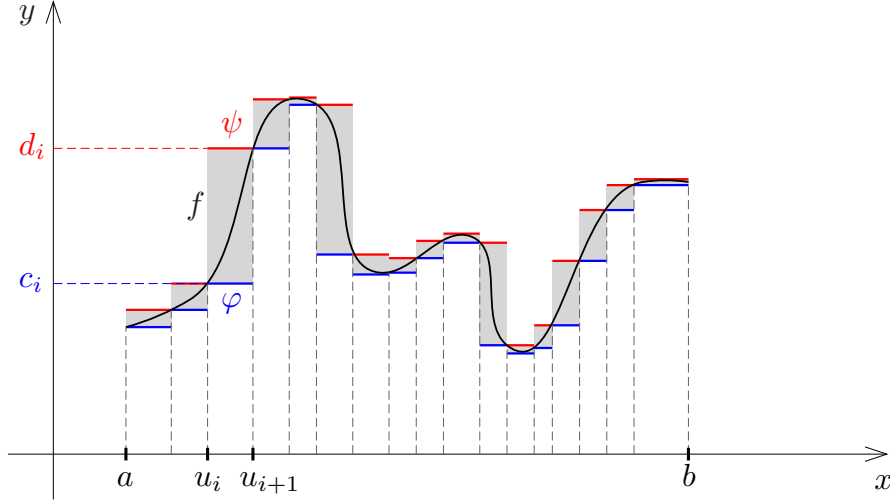


Fig. 12. Encadrement d'une fonction bornée f par des fonctions en escalier.

On peut supposer ici que les fonctions en escalier φ et ψ sont exprimées au moyen de la même subdivision $a = u_0 < u_1 < \dots < u_p = b$, sinon il est toujours possible de redécouper les subdivisions qui les définissent pour arriver à une subdivision commune [de plus, comme les valeurs $\varphi(u_i), \psi(u_i)$ ne jouent pas de rôle dans les intégrales, on peut choisir par exemple $\varphi(u_i) = \psi(u_i) = f(u_i)$]. Dans cette situation, l'erreur due à l'encadrement de l'aire est précisément

$$(7.2) \quad \int_a^b \psi(x) dx - \int_a^b \varphi(x) dx = \sum_{0 \leq i \leq p-1} (u_{i+1} - u_i)(d_i - c_i)$$

(somme des aires des rectangles dessinés en grisé sur la Fig. 12). On obtient ainsi le

(7.3) Critère d'intégrabilité au sens de Riemann. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction. On suppose que pour tout $\varepsilon > 0$ il existe un encadrement $\varphi \leq f \leq \psi$ de f par des fonctions en escalier φ, ψ telles que

$$\int_a^b \psi(x) dx - \int_a^b \varphi(x) dx = \sum_{0 \leq i \leq p-1} (u_{i+1} - u_i)(d_i - c_i) \leq \varepsilon.$$

Alors f est intégrable au sens de Riemann sur $[a, b]$ et on a

$$\int_a^b f(x) dx = \sup \left\{ \int_a^b \varphi(x) dx ; \varphi \leq f \right\} = \inf \left\{ \int_a^b \psi(x) dx ; \psi \geq f \right\}.$$

Démonstration. Considérons les ensembles de valeurs prises par les intégrales de fonctions en escalier φ qui minorent et ψ qui majorent, c'est-à-dire

$$E_1 = \left\{ \int_a^b \varphi(x) dx ; \varphi \leq f \right\}, \quad E_2 = \left\{ \int_a^b \psi(x) dx ; \psi \geq f \right\},$$

et $S = \sup E_1$, $I = \inf E_2$. Il est clair que pour tout $A_1 = \int_a^b \varphi(x) dx \in E_1$ et tout $A_2 = \int_a^b \psi(x) dx \in E_2$ on a $A_1 \leq A_2$. L'hypothèse de la condition signifie que pour un choix adéquat de φ , ψ on a $A_2 - A_1 \leq \varepsilon$. Ceci entraîne bien $S = \sup E_1 = I = \inf E_2$. Comme toute fonction en escalier est intégrable avec des jauges constantes (cf. exemple 2.8), il existe $\delta_1 > 0$, $\delta_2 > 0$ tels que

$$\begin{aligned} \forall D = \{([a_j, a_{j+1}], x_j)\}, \quad h_j \leq \delta_1 \text{ pour tout } j &\Rightarrow |S_D(\varphi) - A_1| \leq \varepsilon, \\ \forall D = \{([a_j, a_{j+1}], x_j)\}, \quad h_j \leq \delta_2 \text{ pour tout } j &\Rightarrow |S_D(\psi) - A_2| \leq \varepsilon. \end{aligned}$$

Or nous avons $A_1 \leq S = I \leq A_2$ et $A_2 - A_1 \leq \varepsilon$. Prenons une subdivision pointée $D = ([a_j, a_{j+1}], x_j)$ vérifiant $h_j \leq \delta = \min(\delta_1, \delta_2)$. Nous obtenons dans ces conditions $S_D(\varphi) \leq S_D(f) \leq S_D(\psi)$, donc

$$S_D(f) \leq S_D(\psi) \leq A_2 + \varepsilon \leq I + 2\varepsilon, \quad S_D(f) \geq S_D(\varphi) \geq A_1 - \varepsilon \geq S - 2\varepsilon.$$

Par suite en posant $A = S = I$ il vient $|S_D(f) - A| \leq 2\varepsilon$. La suffisance du critère d'intégrabilité est démontrée. \square

(7.4)* Complément. Le critère 7.3 est en fait une condition nécessaire et suffisante pour que f soit intégrable au sens de Riemann sur $[a, b]$.

Démonstration. Pour voir la nécessité de la condition, supposons $f : [a, b] \rightarrow \mathbb{R}$ Riemann-intégrable et soient $\varepsilon > 0$ et $\delta > 0$ satisfaisant la définition 2.1. Considérons une subdivision pointée δ -fine $D = \{([a_j, a_{j+1}], x_j)\}$ quelconque. En faisant varier indépendamment chaque x_j dans l'intervalle $[a_j, a_{j+1}]$, on trouve

$$\begin{aligned} \sup_{\{x_j\}} S_D(f) &= \sup_{\{x_j\}} \sum_{j=0}^{N-1} (a_{j+1} - a_j) f(x_j) = \sum_{j=0}^{N-1} (a_{j+1} - a_j) M_j, \\ \inf_{\{x_j\}} S_D(f) &= \inf_{\{x_j\}} \sum_{j=0}^{N-1} (a_{j+1} - a_j) f(x_j) = \sum_{j=0}^{N-1} (a_{j+1} - a_j) m_j \end{aligned}$$

où $M_j = \sup_{x \in [a_j, a_{j+1}]} f(x)$, $m_j = \inf_{x \in [a_j, a_{j+1}]} f(x)$. Puisque $A - \varepsilon \leq S_D(f) \leq A + \varepsilon$, on voit en passant au sup et à l'inf que

$$A - \varepsilon \leq \sum_{j=0}^{N-1} (a_{j+1} - a_j) m_j \leq \sum_{j=0}^{N-1} (a_{j+1} - a_j) M_j \leq A + \varepsilon.$$

Ceci entraîne en particulier qu'aucune des bornes M_j ne peut être égale à $+\infty$ et qu'aucune des bornes m_j ne peut être égale à $-\infty$, et par conséquent que f est bornée. Si on définit les fonctions en escalier φ , ψ par

$$\varphi(x) = m_j, \quad \psi(x) = M_j \quad \text{si } x \in]a_j, a_{j+1}[, \quad \varphi(a_j) = \psi(a_j) = f(a_j),$$

on obtient un encadrement $\varphi \leq f \leq \psi$ tel que

$$\int_a^b \psi(x) dx - \int_a^b \varphi(x) dx \leq 2\varepsilon,$$

et on voit que la condition du critère 7.3 est bien satisfaite. \square

Une application directe de 7.3 est la preuve de l'intégrabilité des fonctions monotones ou continues (dans la suite de ce paragraphe, tous les résultats concerneront donc l'intégrabilité au sens de Riemann.)

(7.5) Théorème. *Toute fonction $f : [a, b] \rightarrow \mathbb{R}$ monotone est intégrable au sens de Riemann sur $[a, b]$.*

Démonstration. Supposons par exemple f croissante et soit $a = u_1 < \dots < u_p = b$ une subdivision δ -fine de $[a, b]$, où $\delta > 0$ est une constante. De manière évidente, on définit un encadrement $\varphi \leq f \leq \psi$ de f par des fonctions en escalier en posant

$$\varphi(x) = f(u_j) \text{ sur }]u_j, u_{j+1}[, \quad \psi(x) = f(u_{j+1}) \text{ sur }]u_j, u_{j+1}[$$

(et $\varphi(u_j) = \psi(u_j) = f(u_j)$ comme déjà convenu). Ceci donne

$$\begin{aligned} \int_a^b \psi(x) dx - \int_a^b \varphi(x) dx &= \sum_{0 \leq j \leq N-1} (u_{j+1} - u_j)(f(u_{j+1}) - f(u_j)) \\ &\leq \delta \sum_{0 \leq j \leq N-1} (f(u_{j+1}) - f(u_j)) = \delta(f(b) - f(a)). \end{aligned}$$

On voit ainsi que la condition 7.3 est vérifiée en prenant δ assez petit. \square

(7.6) Théorème *Toute fonction $f : [a, b] \rightarrow \mathbb{R}$ continue est intégrable au sens de Riemann sur $[a, b]$.*

Démonstration. Soit $\varepsilon > 0$. La continuité de f en tout point $x \in [a, b]$ implique l'existence d'un réel $\delta(x) > 0$ tel que

$$\forall x' \in [a, b], \quad x' \in [x - \delta(x), x + \delta(x)] \Rightarrow |f(x') - f(x)| \leq \varepsilon.$$

Soit $D = ([a_j, a_{j+1}], x_j)_{0 \leq j < N}$ une subdivision pointée δ -fine. En prenant $x = x_j$ et en faisant parcourir à x' l'intervalle $[a_j, a_{j+1}] \subset [x_j - \delta(x_j), x_j + \delta(x_j)]$, on déduit de la majoration $|f(x') - f(x_j)| \leq \varepsilon$ que les bornes inf et sup

$$m_j = \inf_{x' \in [a_j, a_{j+1}]} f(x'), \quad M_j = \sup_{x' \in [a_j, a_{j+1}]} f(x')$$

sont comprises entre $f(x_j) - \varepsilon$ et $f(x_j) + \varepsilon$, par conséquent $M_j - m_j \leq 2\varepsilon$. On obtient ainsi un encadrement $\varphi \leq f \leq \psi$ de f par des fonctions en escalier φ, ψ telles que

$$\varphi(x) = m_j, \quad \psi(x) = M_j \quad \text{si } x \in]a_j, a_{j+1}[, \quad \varphi(a_j) = \psi(a_j) = f(a_j),$$

et de plus

$$\begin{aligned} \int_a^b \psi(x) dx - \int_a^b \varphi(x) dx &= \sum_{0 \leq j \leq N-1} (a_{j+1} - a_j)(M_j - m_j) \\ &\leq 2\varepsilon \sum_{0 \leq j \leq N-1} (a_{j+1} - a_j) = 2\varepsilon(b - a). \end{aligned}$$

Ceci montre que le critère 7.3 est satisfait, donc f est intégrable.⁽¹²⁾ □

(7.7) Théorème. *Toute fonction continue $f : [a, b] \rightarrow \mathbb{R}$ admet une primitive F , donnée par*

$$F(x) = \int_a^x f(t) dt, \quad x \in [a, b].$$

Les autres primitives sont les fonctions de la forme $F_1(x) = F(x) + C$ où C est une constante.

Démonstration. Nous savons que l'intégrale donnant $F(x)$ existe par le Théorème 7.6. La relation de Chasles donne

$$F(x+h) - F(x) = \int_x^{x+h} f(t) dt \quad \Rightarrow \quad \frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f(t) dt.$$

Pour tout $\varepsilon > 0$, l'hypothèse de continuité dit que $f(x) - \varepsilon \leq f(t) \leq f(x) + \varepsilon$ pour $|t - x| \leq \delta(x)$, on a donc

$$f(x) - \varepsilon = \frac{1}{h} \int_x^{x+h} (f(x) - \varepsilon) dt \leq \frac{F(x+h) - F(x)}{h} \leq \frac{1}{h} \int_x^{x+h} (f(x) + \varepsilon) dt = f(x) + \varepsilon$$

pour $|h| \leq \delta(x)$, ce qui signifie que

$$F'(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = f(x).$$

Si on a une autre primitive F_1 , il vient $(F_1 - F)' = f' - f' = 0$, donc $F_1 - F = C$ constante. □

(7.8) Proposition.

- (a) *Si $f : [a, b] \rightarrow \mathbb{R}$ est une fonction continue positive ou nulle, on a $\int_a^b f(x) = 0$ si et seulement si $f = 0$.*
- (b) *Si $f, g : [a, b] \rightarrow \mathbb{R}$ sont continues et $f \leq g$, on a $\int_a^b f(x) dx < \int_a^b g(x) dx$ dès que f et g se sont pas égales.*

⁽¹²⁾ On remarquera que dans cette preuve l'utilisation explicite de la propriété de continuité uniforme de la fonction f n'a pas été nécessaire, ce qui est une simplification appréciable susceptible de rendre la preuve abordable dès la classe terminale [avec des vitamines tout de même!].

Démonstration. (a) Si f n'est pas nulle, il existe un point $x_0 \in [a, b]$ tel que $f(x_0) > 0$, et on pose alors $\varepsilon = f(x_0)/2$. La continuité de f en x_0 implique qu'il existe un intervalle $[x_0, x_0 + \eta]$ (ou $[x_0 - \eta, x_0]$, si $x_0 = b$) sur lequel $|f(x) - f(x_0)| \leq \varepsilon$. On voit donc qu'il existe un sous-intervalle $[c, d]$ de $[a, b]$ de longueur $d - c = \eta > 0$ sur lequel $f(x) \geq f(x_0) - \varepsilon \geq \varepsilon$. Par suite $\int_a^b f(x) dx \geq \int_c^d f(x) dx \geq (d - c)\varepsilon > 0$.

(b) Si $f \leq g$ et $f \neq g$, alors $h = g - f \geq 0$ n'est pas nulle, donc $\int_a^b h(x) dx > 0$ d'après (a). \square

(7.9) Formule de la moyenne. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Alors il existe un point $c \in]a, b[$ tel que la « valeur moyenne de f sur $[a, b]$ » soit égale à $f(c)$:

$$\frac{1}{b-a} \int_a^b f(x) dx = f(c).$$

Démonstration. Soit $m = \min_{[a,b]} f$, $M = \max_{[a,b]} f$. Supposons d'abord f non constante. D'après la proposition 7.8 (b) appliquée aux inégalités $m \leq f \leq M$, nous avons

$$m(b-a) < \int_a^b f(x) dx < M(b-a), \quad \text{soit} \quad m < \frac{1}{b-a} \int_a^b f(x) dx < M.$$

La formule de la moyenne est donc une conséquence du théorème des valeurs intermédiaires, puisque $f(]a, b[)$ est un intervalle qui contient l'intervalle $]m, M[$. Si f est égale à une constante C , le résultat est évident, les deux membres de la formule étant égaux à C quel que soit le choix de $c \in]a, b[$. \square

(7.10) Une généralisation. On dit qu'une fonction $f : [a, b] \rightarrow \mathbb{R}$ est continue (resp. monotone) par morceaux s'il existe une subdivision

$$a = \alpha_0 < \alpha_1 < \alpha_2 < \dots < \alpha_{N-1} < \alpha_N = b$$

telle que f soit continue (resp. monotone) sur chaque intervalle $]\alpha_j, \alpha_{j+1}[$ et possède des limites à droite et à gauche finies en chaque point α_j tel que $j < N$ (resp. $j > 0$). Toute fonction continue ou monotone par morceaux est intégrable au sens de Riemann.

Démonstration. En effet, la restriction $f|_{[\alpha_j, \alpha_{j+1}]}$ diffère d'une fonction continue (resp. monotone) par une fonction en escalier nulle sur $]\alpha_j, \alpha_{j+1}[$ (et prenant des valeurs adéquates en α_j et α_{j+1}). Par conséquent $f|_{[\alpha_j, \alpha_{j+1}]}$ est intégrable au sens de Riemann, et l'intégrabilité de f sur $[a, b]$ résulte de la relation de Chasles. \square

Nous démontrons maintenant une généralisation du théorème fondamental 4.1. Observons d'abord que comme une fonction nulle sauf sur un ensemble fini est d'intégrale nulle (remarque 3.2), on peut envisager d'intégrer des fonctions f qui sont seulement définies sur le complémentaire $[a, b] \setminus L$ d'un ensemble fini Z : on se ramène au cas d'une fonction partout définie en étendant f arbitrairement sur $[a, b]$, par exemple en

posant $\tilde{f}(x) = 0$ sur Z . L'intégrale ainsi calculée est indépendante du prolongement \tilde{f} choisi.

(7.11) Théorème. *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. On suppose qu'il existe un ensemble fini $Z = \{u_i; 1 \leq i \leq p\}$ tel que f soit dérivable sur $[a, b] \setminus Z$. Alors f' (prolongée arbitrairement aux points u_i) est KH-intégrable sur $[a, b]$ et on a*

$$\boxed{\int_a^b f'(x) dx = f(b) - f(a).}$$

*Démonstration.** On reprend la démonstration du théorème 4.1. La dérivabilité de f sur $[a, b] \setminus Z$ implique l'existence d'une fonction $\delta : [a, b] \setminus Z \rightarrow]0, +\infty[$ telle que

$$y \in [a, b], \quad y \in [x - \delta(x), x + \delta(x)] \Rightarrow |f(y) - f(x) - (y - x)f'(x)| \leq \varepsilon|y - x|$$

pour tout $x \in [a, b] \setminus Z$, ce qui donne

$$|f(a_{j+1}) - f(a_j) - (a_{j+1} - a_j)f'(x_j)| \leq \varepsilon|a_{j+1} - a_j|$$

pour toute subdivision pointée $D = \{([a_j, a_{j+1}], x_j)\}$ δ -fine, lorsque $x_j \in [a, b] \setminus Z$. D'autre part, si $x_j = u_i \in Z$, la continuité de f au point u_i entraîne l'existence de $\delta_i > 0$ tel que tout point $x \in [u_i - \delta_i, u_i + \delta_i]$ satisfasse $|f(x) - f(u_i)| \leq \varepsilon 2^{-i}$. On choisit de plus δ_i assez petit pour que $\delta_i|f'(u_i)| \leq \varepsilon 2^{-i}$, et on pose $\delta(u_i) = \delta_i$. Dans ce cas il vient

$$|f(a_{j+1}) - f(a_j)| \leq |(f(a_{j+1}) - f(u_i)) - (f(a_j) - f(u_i))| \leq 2\varepsilon 2^{-i}$$

si $x_j = u_i$ et $a_{j+1} - a_j \leq \delta(x_j) = \delta(u_i) = \delta_i$, ce qui implique

$$|f(a_{j+1}) - f(a_j) - (a_{j+1} - a_j)f'(u_i)| \leq 3\varepsilon 2^{-i}.$$

En sommant toutes ces inégalités, il vient

$$|f(b) - f(a) - S_D(f')| \leq \varepsilon(b - a) + 3\varepsilon,$$

(car $\sum_{1 \leq i \leq p} 3\varepsilon 2^{-i} \leq 3\varepsilon$), et le théorème est démontré. \square

(7.12)* Remarque. D'après la preuve ci-dessus, on voit que le théorème 7.11 est en fait encore vrai avec un ensemble $Z = \{u_i\}_{i \in \mathbb{N}}$ dénombrable, cf. aussi l'exercice 10.22.

(7.13) Corollaire. *Si $f :]a, b[\rightarrow \mathbb{R}$ admet une primitive F sur $]a, b[$, et si cette primitive F admet une limite à droite $F(a + 0)$ en a et une limite à gauche $F(b - 0)$ en b , alors f est KH-intégrable sur $[a, b]$ (si f n'est pas a priori définie en a et b , on peut lui attribuer des valeurs $f(a), f(b) \in \mathbb{R}$ arbitraires), et on a*

$$\boxed{\int_a^b f(x) dx = F(b - 0) - F(a + 0) = \lim_{x \rightarrow b-0} F(x) - \lim_{x \rightarrow a+0} F(x).}$$

Démonstration. Il suffit en effet de prolonger F par continuité sur $[a, b]$ en posant $F(a) = \lim_{x \rightarrow a+0} F(x)$ et $F(b) = \lim_{x \rightarrow b-0} F(x)$, puis d'appliquer le théorème 7.11 à sa dérivée F' qui est définie sur $]a, b[= [a, b] \setminus Z$ avec $Z = \{a, b\}$. \square

Un exemple typique d'application du corollaire 7.13 est celui de la fonction

$$f(x) = \frac{1}{\sqrt{1-x^2}} \quad \text{sur l'intervalle }]-1, 1[.$$

Dans ce cas, on a en effet une primitive $F(x) = \arcsin(x)$ qui se prolonge en une fonction continue sur $[-1, 1]$. On obtient par conséquent

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} dx = \arcsin(1) - \arcsin(-1) = \pi.$$

Plus généralement, le corollaire 7.13 nous amène à la définition des « intégrales impropres ».

(7.14) Intégrales « impropres ». Soit $I = [a, b[\subset \mathbb{R}$ un intervalle semi-ouvert, où $b \in \mathbb{R} \cup \{+\infty\}$, et soit $f : [a, b[\rightarrow \mathbb{R}$ une fonction intégrable (par exemple continue ou monotone par morceaux) sur tout intervalle $[a, \beta] \subset [a, b[$. On dit que l'intégrale $\int_a^b f(x) dx$ est convergente au point b si la limite

$$A = \lim_{\beta \rightarrow b-} \int_a^\beta f(x) dx$$

existe dans \mathbb{R} (en particulier finie), et on pose alors $\int_a^b f(x) dx = A$.

On donne une définition analogue dans le cas d'un intervalle semi-ouvert $]a, b]$, $a \in \mathbb{R} \cup \{-\infty\}$, en considérant la limite $\lim_{\alpha \rightarrow a+} \int_\alpha^b f(x) dx$ avec $[\alpha, b] \subset]a, b]$, et on dit qu'on a convergence sur un intervalle ouvert $]a, b[$ si les intégrales sur $]a, c]$ et $[c, b[$ convergent pour tout point intermédiaire $c \in]a, b[$.

(7.15) Exemples. On vérifiera que les intégrales

$$\int_0^1 \frac{1}{x^a} dx = \frac{1}{1-a}, \quad \int_1^{+\infty} \frac{1}{x^b} dx = \frac{1}{b-1}$$

convergent respectivement pour $a < 1$ et $b > 1$, et que

$$\int_0^{+\infty} e^{-cx} dx = \frac{1}{c}$$

converge pour tout $c > 0$. En utilisant le calcul de la primitive de $x^n e^{-x}$ donné au point (5.1 a), on obtient également le résultat classique

$$\int_0^{+\infty} x^n e^{-x} dx = n!.$$

(7.16) Convergence absolue.** Soit $f : [a, b[\rightarrow \mathbb{R}$ une fonction KH-intégrable sur tout intervalle fermé borné $[a, \beta]$. On dit que l'intégrale $\int_a^b f(x) dx$ est absolument convergente si de plus $|f|$ est KH-intégrable sur chaque intervalle $[a, \beta] \subset [a, b[$ et si la limite

$$\lim_{\beta \rightarrow b-} \int_a^\beta |f(x)| dx$$

existe dans \mathbb{R} (donc en particulier si elle est finie).

On donne une définition analogue dans le cas d'un intervalle semi-ouvert $]a, b]$, $a \in \mathbb{R} \cup \{-\infty\}$, en considérant la limite $\lim_{\alpha \rightarrow a+} \int_\alpha^b f(x) dx$ avec $[\alpha, b] \subset]a, b]$, et on dit qu'on a convergence sur un intervalle ouvert $]a, b[$ si les intégrales sur $]a, c]$ et $[c, b[$ convergent pour tout point intermédiaire $c \in]a, b[$.

Le critère de Cauchy* permet de voir qu'il y a convergence (resp. convergence absolue) de l'intégrale sur $]a, b[$ si et seulement si pour tout $\varepsilon > 0$ il existe $\beta_\varepsilon < b$ tel que pour tous $\beta, \gamma \in [\beta_\varepsilon, b[$, $\beta < \gamma$, on ait

$$\left| \int_\beta^\gamma f(x) dx \right| \leq \varepsilon, \quad \text{resp.} \quad \int_\beta^\gamma |f(x)| dx \leq \varepsilon.$$

Comme $\left| \int_\beta^\gamma f(x) dx \right| \leq \int_\beta^\gamma |f(x)| dx$, il est clair que la convergence absolue implique la convergence.

8. Convergence uniforme, continuité et dérivabilité d'intégrales en fonction de paramètres**

Soit $f_n : [a, b] \rightarrow \mathbb{R}$ une suite de fonctions réelles. On dit que la suite (f_n) converge uniformément vers une limite $f : [a, b] \rightarrow \mathbb{R}$, si la distance uniforme de f_n à f tend vers 0 c'est-à-dire si

$$d(f_n, f) = \sup_{x \in [a, b]} |f_n(x) - f(x)| \rightarrow 0.$$

(8.1) Théorème de convergence uniforme. Soit $f_n : [a, b] \rightarrow \mathbb{R}$ une suite de fonctions réelles convergeant uniformément vers une limite f . Si les fonctions f_n sont Riemann-intégrables (resp. KH-intégrables), alors f est Riemann-intégrable (resp. KH-intégrable) et

$$\lim_{n \rightarrow +\infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx.$$

Démonstration. Posons $A_n = \int_a^b f_n(x) dx$. Soit $\varepsilon > 0$ et n_0 tels que $|f_n - f| \leq \varepsilon$ pour $n \geq n_0$. Nous avons $|f_p - f_q| \leq 2\varepsilon$ pour $p, q \geq n_0$, d'où $|A_p - A_q| \leq 2\varepsilon(b-a)$. La suite (A_n) est donc une suite de Cauchy, par conséquent la limite $A = \lim_{n \rightarrow +\infty} A_n$ existe. De plus, il existe une jauge δ (resp. une jauge constante dans le cas de l'intégrabilité au sens de Riemann) telle que pour toute subdivision pointée D de pas $h_j \leq \delta(x_j)$ on ait $|S_D(f_{n_0}) - A_{n_0}| \leq \varepsilon$. Il vient à la limite $|A - A_{n_0}| \leq 2\varepsilon(b-a)$ tandis que

$|S_D(f) - S_D(f_{n_0})| \leq \sum h_j |f(x_j) - f_{n_0}(x_j)| \leq \varepsilon(b-a)$. En mettant bout à bout ces inégalités il vient

$$|S_D(f) - A| \leq \varepsilon + 3\varepsilon(b-a),$$

donc f est bien intégrable sur $[a, b]$ d'intégrale $A = \lim A_n$. \square

(8.2) Remarque. Pour obtenir l'intégrabilité au sens de Kurzweil-Henstock de la limite $f = \lim f_n$, il suffirait d'avoir une estimation de la forme $|f_n - f| \leq \varepsilon_n g$ avec $g \geq 0$ KH-intégrable (non nécessairement bornée) et $\varepsilon_n \rightarrow 0$. Nous obtiendrons de toutes façons des théorèmes de convergence beaucoup meilleurs encore dans ce cas, cf. chapitre III.

Rappelons qu'une fonction $f : I \rightarrow \mathbb{R}$ définie sur un intervalle I est dite *réglée* si elle admet une limite à droite et à gauche en tout point de l'intérieur I° , ainsi qu'une limite à droite en $a = \inf I$ si $a \in I$ et à gauche en $b = \sup I$ si $b \in I$.

(8.3) Lemme. *Toute fonction réglée sur un intervalle fermé borné $[a, b]$ est limite uniforme de fonctions en escalier.*

(8.4) Corollaire. *Toute fonction réglée sur un intervalle fermé borné $[a, b]$ y est Riemann-intégrable.*

Démonstration du lemme et de son corollaire. Soit $\varepsilon > 0$. Pour tout $x \in [a, b]$, il existe $\delta(x) > 0$ et des constantes $c^+ = \lim_{\xi \rightarrow x+0} f(\xi)$, $c^- = \lim_{\xi \rightarrow x-0} f(\xi)$ telles que $|f(x) - c^-| \leq \varepsilon$ sur $[x - \delta(x), x] \cap [a, b]$ et $|f(x) - c^+| \leq \varepsilon$ sur $[x, x + \delta(x)] \cap [a, b]$. D'après le lemme I.2.6, on peut trouver une subdivision pointée $D = ([a_j, a_{j+1}], x_j)$ δ -fine. Ceci nous permet de définir une fonction en escalier φ en posant

$$\begin{aligned} \varphi(x_j) &= f(x_j), \quad \varphi(a_j) = f(a_j), \\ \varphi|_{]a_j, x_j[} &= c_j^- = \lim_{\xi \rightarrow x_j-0} f(\xi), \quad \varphi|_{]x_j, a_{j+1}[} = c_j^+ = \lim_{\xi \rightarrow x_j+0} f(\xi). \end{aligned}$$

Par construction nous avons $|f - \varphi| \leq \varepsilon$ sur $[a, b]$. Il en résulte que f est Riemann-intégrable comme limite uniforme de fonctions en escalier. \square

On va maintenant tirer du théorème de convergence uniforme les propriétés usuelles de continuité et de dérivabilité sous le signe somme des intégrales dépendant de paramètres. Des résultats beaucoup plus forts sont vrais, mais on aura besoin pour cela de théorèmes de convergence plus subtils que le théorème 8.1 (cf. Chapitre III, section 5).

(8.5) Lemme. *Soit $T \subset \mathbb{R}^d$ une partie quelconque et $f : [a, b] \times T \rightarrow \mathbb{R}$, $(x, t) \mapsto f(x, t)$ une application continue. Alors, si $t_0 \in T$, la famille de fonctions $f_t : x \mapsto f(x, t)$ converge uniformément vers la fonction $f_{t_0} : x \mapsto f(x, t_0)$ quand $T \ni t$ tend vers t_0 . Autrement dit, pour tout $\varepsilon > 0$, il existe un voisinage V de t_0 tel que pour $t \in T \cap V$, on ait $|f(x, t) - f(x, t_0)| \leq \varepsilon$ pour tout $x \in [a, b]$.*

Démonstration. Pour tout $x_0 \in [a, b]$, l'hypothèse de continuité au point (x_0, t_0) implique qu'il existe un voisinage V_{x_0, t_0} de t_0 (dépendant de x_0) et un voisinage

$$U_{(x_0, t_0)} = [x_0 - \delta(x_0), x_0 + \delta(x_0)] \times V_{x_0, t_0} \quad \text{de } (x_0, t_0)$$

tel que $(x, t) \in U_{(x_0, t_0)} \cap ([a, b] \times T)$ implique $|f(x, t) - f(x_0, t_0)| \leq \varepsilon/2$. En appliquant ceci en particulier pour $t = t_0$ et en faisant la différence, on voit que pour tout $t \in T \cap V_{x_0, t_0}$ et tout $x \in [x_0 - \delta(x_0), x_0 + \delta(x_0)]$ on a $|f(x, t) - f(x, t_0)| \leq \varepsilon$. D'après le lemme I.2.6, il existe une subdivision pointée $D = ([a_i, a_{i+1}], x_i)$ de $[a, b]$ qui est δ -fine. Par conséquent $[a_i, a_{i+1}] \subset [x_i - \delta(x_i), x_i + \delta(x_i)]$, et en prenant $t \in V = \bigcap_i V_{x_i, t_0}$ on voit que la conclusion du lemme est vérifiée. \square

De là, on déduit aussitôt le théorème de continuité des intégrales dépendant de paramètres.

(8.6) Continuité sous le signe somme. Soit $T \subset \mathbb{R}^d$ une partie quelconque et $f : [a, b] \times T \rightarrow \mathbb{R}$, $(x, t) \mapsto f(x, t)$ une application continue. Alors l'application

$$F(t) = \int_a^b f(x, t) dx$$

est continue sur T .

Démonstration. Avec les notations du lemme 8.5, on trouve

$$|F(t) - F(t_0)| \leq \int_a^b |f(x, t) - f(x, t_0)| dx \leq (b - a)\varepsilon$$

pour $t \in T \cap V$, ce qui prouve la continuité de F au point t_0 . \square

(8.7) Dérivation sous le signe somme. Soit T un intervalle de la droite réelle et $f : [a, b] \times T \rightarrow \mathbb{R}$, $(x, t) \mapsto f(x, t)$ une fonction telle que

- (a) f est continue sur $[a, b] \times T$;
- (b) f admet en tout point une dérivée partielle $\frac{\partial f}{\partial t}(x, t)$ qui est elle-même continue sur $[a, b] \times T$.

Alors l'application $F(t) = \int_a^b f(x, t) dx$ est de classe C^1 sur T et pour tout $t_0 \in T$ on a

$$F'(t_0) = \int_a^b \frac{\partial f}{\partial t}(x, t_0) dx.$$

Démonstration. Soit $t \in T$. En appliquant le théorème des accroissements finis à $t \mapsto f(x, t)$ sur l'intervalle $[t_0, t]$, on voit que

$$\frac{F(t) - F(t_0)}{t - t_0} = \int_a^b \frac{f(x, t) - f(x, t_0)}{t - t_0} dx = \int_a^b \frac{\partial f}{\partial t}(x, c_{t,x}) dx$$

pour un certain point $c = c_{t,x} \in]t_0, t[$. Soit $\varepsilon > 0$. En appliquant le lemme 8.5 à la fonction continue $g(x, t) = \frac{\partial f}{\partial t}(x, t)$, on voit qu'il existe un voisinage $V =]t_0 - \eta, t_0 + \eta[$ de t_0 tel que $|\frac{\partial f}{\partial t}(x, t) - \frac{\partial f}{\partial t}(x, t_0)| \leq \varepsilon$ pour tout $x \in [a, b]$ et tout $t \in V$. Sous cette

hypothèse on a également $c(x, t) \in]t_0, t[\subset V$, donc $|\frac{\partial f}{\partial t}(x, c_{t,x}) - \frac{\partial f}{\partial t}(x, t_0)| \leq \varepsilon$ et par suite

$$\left| \frac{F(t) - F(t_0)}{t - t_0} - \int_a^b \frac{\partial f}{\partial t}(x, t_0) dx \right| = \left| \int_a^b \left(\frac{\partial f}{\partial t}(x, c_{t,x}) - \frac{\partial f}{\partial t}(x, t_0) \right) dx \right| \leq (b - a)\varepsilon.$$

On a donc bien $F'(t_0) = \int_a^b \frac{\partial f}{\partial t}(x, t_0) dx$, et cette relation montre que F' est continue d'après le théorème 8.6. \square

Pour un paramètre $t = (t_1, \dots, t_d) \in \mathbb{R}^d$, nous avons le résultat analogue suivant.

(8.8) Différentiabilité sous le signe somme. Soit $T \subset \mathbb{R}^d$ une partie ouverte et $f : [a, b] \times T \rightarrow \mathbb{R}$, $(x, t) \mapsto f(x, t)$ une fonction telle que

- (a) f est continue sur $[a, b] \times T$;
- (b) f admet en tout point des dérivées partielles $\frac{\partial f}{\partial t_j}(x, t)$ qui sont elles-mêmes continues sur $[a, b] \times T$.

Alors l'application $F(t) = \int_a^b f(x, t) dx$ est de classe C^1 sur T et on a

$$\boxed{\frac{\partial F}{\partial t_j}(t) = \int_a^b \frac{\partial f}{\partial t_j}(x, t) dx.}$$

Démonstration. Il suffit en effet d'appliquer le théorème de dérivation sous le signe somme séparément pour chaque variable t_j pour voir que les dérivées partielles $\partial F / \partial t_j$ existent, et d'invoquer ensuite le théorème de continuité pour voir que ces dérivées partielles sont continues sur T . \square

Il est bon parfois de connaître également la formule de différentiation sous le signe somme dans le cas où l'intégrale est calculée sur des intervalles dépendant du paramètre.

(8.9) Formule générale de dérivation sous le signe somme. Soient $I \subset \mathbb{R}$ et $T \subset \mathbb{R}$ des intervalles. On considère une intégrale de la forme

$$F(t) = \int_{a(t)}^{b(t)} f(x, t) dx$$

où $a, b : T \rightarrow I$ sont des fonctions de classe C^1 et $f : I \times T \rightarrow \mathbb{R}$, $(x, t) \mapsto f(x, t)$ une fonction continue admettant une dérivée partielle $\frac{\partial f}{\partial t}$ continue sur $I \times T$. Alors l'application F est de classe C^1 sur T et on a

$$\boxed{F'(t) = \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t}(x, t) dx + b'(t)f(b(t), t) - a'(t)f(a(t), t).}$$

Démonstration. On pose

$$\Phi(t, u, v) = \int_u^v f(x, t) dx.$$

Plaçons-nous sur un sous-intervalle fermé borné $T' = [t_1, t_2] \subset T$, et soit

$$A = \min_{t \in T'} a(t) \in I, \quad B = \max_{t \in T'} b(t) \in I.$$

Nous avons $[A, B] \subset I$, donc les hypothèses des théorèmes 8.5 et 8.6 sont satisfaites sur $[A, B] \times T'$. On voit par conséquent que Φ admet des dérivées partielles

$$\frac{\partial \Phi}{\partial t} = \int_u^v \frac{\partial f}{\partial t}(x, t) dx, \quad \frac{\partial \Phi}{\partial v} = f(v, t), \quad \frac{\partial \Phi}{\partial u} = -f(u, t)$$

(pour cette dernière égalité, il suffit d'échanger les bornes u et v). De plus, ces trois dérivées partielles sont continues en les variables (t, u, v) . Pour $\frac{\partial \Phi}{\partial u}$ et $\frac{\partial \Phi}{\partial v}$ c'est clair par hypothèse, tandis que pour $\frac{\partial \Phi}{\partial t}$ cela résulte de l'estimation

$$\left| \frac{\partial \Phi}{\partial t}(t, u, v) - \frac{\partial \Phi}{\partial t}(t_0, u_0, v_0) \right| \leq M(|u - u_0| + |v - v_0|) + \left| \int_{u_0}^{v_0} \left(\frac{\partial f}{\partial t}(x, t) - \frac{\partial f}{\partial t}(x, t_0) \right) dx \right|$$

dans laquelle M est un majorant de $|\frac{\partial f}{\partial t}|$ sur la pavé $[A, B] \times T'$. On déduit de tout cela que $\Phi(t, u, v)$ et $F(t) = \Phi(t, a(t), b(t))$ sont de classe C^1 . De plus

$$F'(t) = \frac{\partial \Phi}{\partial t}(t, a(t), b(t)) + \frac{\partial \Phi}{\partial u}(t, a(t), b(t)) a'(t) + \frac{\partial \Phi}{\partial v}(t, a(t), b(t)) b'(t)$$

ce qui donne la formule (8.9) attendue. □

9. Intégrales à valeurs vectorielles et formule de Taylor**

L'objet de cette section est d'étendre la définition des intégrales au cas des fonctions $f : [a, b] \rightarrow E$ à valeurs dans un espace E de Banach, c'est-à-dire un *espace normé complet* $(E, \|\cdot\|)$ sur \mathbb{R} ou \mathbb{C} . On va voir que la théorie de l'intégrale de Kurzweil-Henstock permet de généraliser, d'unifier et de simplifier considérablement des énoncés classiques comme le théorème des accroissements finis ou les différentes formules de Taylor pour les fonctions à valeurs vectorielles. On définit d'abord les sommes de Riemann associées à une subdivision pointée $D = \{([a_j, a_{j+1}], x_j)\}_{0 \leq j < N}$ par

$$(9.1) \quad S_D(f) = \sum_{j=0}^{N-1} (a_{j+1} - a_j) f(x_j) = \sum_{j=0}^{N-1} h_j f(x_j),$$

la seule différence avec le cas réel étant qu'il s'agit désormais d'une combinaison linéaire à valeurs dans E . On pose alors comme précédemment la définition suivante.

(9.2) Définition. Soit $f : [a, b] \rightarrow E$ une fonction quelconque à valeurs dans un espace de Banach E . On dit que f est *intégrable au sens de Kurzweil-Henstock* (ou *KH-intégrable*) s'il existe un élément $A \in E$ qui représente la valeur de l'intégrale, tel que pour tout $\varepsilon > 0$ donné a priori, on peut trouver une jauge $\delta : [a, b] \rightarrow \mathbb{R}_+^*$ en sorte que pour toute subdivision pointée $D = \{([a_j, a_{j+1}], x_j)\}$ de $[a, b]$ on ait

$$D \text{ } \delta\text{-fine, i.e. } \forall j, a_{j+1} - a_j \leq \delta(x_j) \Rightarrow |S_D(f) - A| \leq \varepsilon.$$

L'élément $A \in E$ de la définition précédente est appelé *intégrale de f sur $[a, b]$* , et on écrit

$$\int_a^b f(x) dx = \lim_{\text{KH}, D} S_D(f) = \lim_{\text{KH}, D} \sum_{j=0}^{N-1} (a_{j+1} - a_j) f(x_j).$$

Comme pour (3.8), on voit que la KH-intégrabilité de f sur $[a, b]$ entraîne la KH-intégrabilité sur tout sous-intervalle $[c, d] \subset [a, b]$; il suffit d'appliquer le critère de Cauchy et d'utiliser la complétude de E . Par ailleurs, l'inégalité triangulaire donne aussitôt $\|S_D(f)\| \leq S_D(\|f\|)$, d'où :

(9.3) Proposition. Soit $f : [a, b] \rightarrow E$ une fonction KH-intégrable à valeurs dans un espace de Banach E . Alors on a

$$\left\| \int_a^b f(x) dx \right\| \leq \int_a^b \|f(x)\| dx$$

dès que $\|f\|$ est KH-intégrable sur $[a, b]$ (on notera que la KH-intégrabilité de f n'entraîne pas nécessairement celle de $\|f\|$).

On considère maintenant des fonctions différentiables à valeurs dans des espaces de Banach. Soient G, E des espaces de Banach, U une partie de G et $f : U \rightarrow E$ une

fonction. On dit que f est différentiable en un point $x \in U$ s'il existe une application linéaire continue $\ell = df(x) \in \mathcal{L}_c(G, E)$ telle qu'on puisse écrire

$$f(y) = f(x) + \ell(y - x) + o(\|y - x\|)$$

pour $y \in U$ voisin de x , c'est-à-dire si pour tout $\varepsilon > 0$, il existe $\delta(x) > 0$ tel que $y \in U$, $\|y - x\| \leq \delta(x)$ implique

$$(9.4) \quad \|f(y) - (f(x) + \ell(y - x))\| \leq \varepsilon \|y - x\|.$$

Dans le cas où U est une partie de la droite réelle $G = \mathbb{R}$, ceci équivaut à l'existence d'une dérivée $f'(x) \in E$, i.e.

$$(9.4') \quad f'(x) = \lim_{y \in U, y \rightarrow x} \frac{f(y) - f(x)}{y - x} \in E,$$

la relation entre $f'(x)$ et $\ell = df(x)$ étant que $f'(x) = \ell(1)$. Le « théorème fondamental de l'Analyse » s'énonce alors :

(9.5) Théorème. *Soit $f : [a, b] \rightarrow E$ une fonction continue en tout point de $[a, b]$ et dérivable en dehors d'un ensemble dénombrable $Z = \{u_i\}_{i \in \mathbb{N}}$. Alors la dérivée $f' : [a, b] \setminus Z \rightarrow E$, prolongée arbitrairement sur Z , est intégrable au sens de Kurzweil-Henstock sur $[a, b]$ et*

$$\boxed{\int_a^b f'(x) dx = f(b) - f(a).}$$

Démonstration : c'est une copie pure et simple de la preuve de (7.11), à l'invocation près de la norme de E à la place de la valeur absolue usuelle. La dérivabilité de f sur $[a, b] \setminus Z$ implique l'existence d'une fonction $\delta : [a, b] \setminus Z \rightarrow]0, +\infty[$ telle que

$$y \in [a, b], \quad y \in [x - \delta(x), x + \delta(x)] \Rightarrow \|f(y) - f(x) - (y - x)f'(x)\| \leq \varepsilon |y - x|$$

pour tout $x \in [a, b] \setminus Z$. En prenant $y = a_j$ puis $y = a_{j+1}$ et en sommant, ceci donne

$$|f(a_{j+1}) - f(a_j) - (a_{j+1} - a_j)f'(x_j)| \leq \varepsilon |a_{j+1} - a_j|$$

pour toute subdivision pointée $D = \{([a_j, a_{j+1}], x_j)\}$ δ -fine, lorsque $x_j \in [a, b] \setminus Z$. D'autre part, si $x_j = u_i \in Z$, la continuité de f au point u_i entraîne l'existence de $\delta_i > 0$ tel que tout point $x \in [u_i - \delta_i, u_i + \delta_i]$ satisfasse $\|f(x) - f(u_i)\| \leq \varepsilon 2^{-i}$. On choisit de plus δ_i assez petit pour que $\delta_i \|f'(u_i)\| \leq \varepsilon 2^{-i}$, et on pose $\delta(u_i) = \delta_i$. Dans ce cas il vient

$$\|f(a_{j+1}) - f(a_j)\| \leq \|(f(a_{j+1}) - f(u_i)) - (f(a_j) - f(u_i))\| \leq 2\varepsilon 2^{-i}$$

si $x_j = u_i$ et $a_{j+1} - a_j \leq \delta(x_j) = \delta(u_i) = \delta_i$, ce qui implique

$$\|f(a_{j+1}) - f(a_j) - (a_{j+1} - a_j)f'(u_i)\| \leq 3\varepsilon 2^{-i}.$$

En sommant toutes ces inégalités, il vient

$$\|f(b) - f(a) - S_D(f')\| \leq \varepsilon(b - a) + 6\varepsilon,$$

(car $\sum_{i \in \mathbb{N}} 3\varepsilon 2^{-i} = 6\varepsilon$), et le théorème est démontré. \square

Supposons maintenant que $f : U \rightarrow E$ soit une fonction différentiable sur une partie U d'un espace de Banach G , et soit $\gamma : [a, b] \rightarrow U$ une fonction continue et disons de classe C^1 par morceaux. Alors $f \circ \gamma$ est continue sur $[a, b]$ et différentiable en dehors des points anguleux de γ (en nombre fini par hypothèse). Comme $(f \circ \gamma)' = (df \circ \gamma)(\gamma')$, il vient

$$(9.6) \quad f(\gamma(b)) - f(\gamma(a)) = \int_a^b (df \circ \gamma)(\gamma'(t)) dt$$

et en passant au sup de la norme pour l'intégrande $\|(df \circ \gamma)(\gamma'(t))\| \leq \|df(\gamma(t))\| \|\gamma'(t)\|$, on en déduit l'inégalité dite des *accroissements finis*

$$(9.6') \quad \|f(\gamma(b)) - f(\gamma(a))\| \leq \sup_{z \in \text{Im}(\gamma)} \|df(z)\| \times \text{longueur}(\gamma)$$

où $\text{longueur}(\gamma) = \int_a^b \|\gamma'(t)\| dt$. En particulier, si $[x, y]$ est un segment contenu dans l'ensemble U , paramétré par $\gamma(t) = x + t(y - x)$, $t \in [0, 1]$, on trouve

$$(9.6'') \quad \|f(y) - f(x)\| \leq \sup_{z \in [x, y]} \|df(z)\| \times \|y - x\|.$$

On notera qu'en fait (9.6), (9.6') et (9.6'') sont vérifiés dès lors que γ est continu sur $[a, b]$ et différentiable en dehors d'un ensemble dénombrable, et qu'il suffit de même que f soit continue sur $\text{Im}(\gamma)$ et différentiable en dehors d'une partie dénombrable de $\text{Im}(\gamma)$ (resp. de $[x, y]$). Le cas particulier de (9.6) avec $\gamma(t) = x + th$, $t \in [0, 1]$, donne également

$$(9.6''') \quad f(x + h) = f(x) + \int_0^1 df(x + th)(h) dt$$

dès que f est différentiable sur le segment $[x, x + h]$. Par récurrence sur p , on définit la dérivée p -ième

$$d^p f : U \rightarrow \mathcal{L}_c(G^p, E) \quad \text{par } d^p f = d(d^{p-1} f),$$

où $\mathcal{L}_c(G^p, E) \simeq \mathcal{L}_c(G, \mathcal{L}_c(G^{p-1}, E))$ désigne l'espace des applications p -multilinéaires continues de $G^p = G \times \dots \times G$ dans E . La dérivée p -ième est donc définie sous l'hypothèse que $d^{p-1} f$ existe et soit elle-même différentiable sur $U \subset G$; on dit alors que f est p fois différentiable sur U . Dans ce cas, on obtient la

(9.7) Formule de Taylor avec reste intégral. Soit $f : U \rightarrow E$ une fonction p fois différentiable sur U . Alors pour tout segment $[x, x + h] \subset U$ on a

$$\boxed{f(x + h) = f(x) + df(x)(h) + \dots + \frac{1}{(p-1)!} d^{p-1} f(x)(h)^{p-1} + \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} d^p f(x + th)(h)^p dt.}$$

On notera que (9.7) se réduit à (9.6''') pour $p = 1$ et que la preuve s'obtient immédiatement par récurrence à l'aide d'une intégration par parties, sachant que la dérivée de $t \mapsto d^p f(x + th)(h)^p$ est $t \mapsto d^{p+1} f(x + th)(h)^{p+1}$:

$$\int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} d^p f(x + th)(h)^p dt = \frac{1}{p!} d^p f(x)(h)^p + \int_0^1 \frac{(1-t)^p}{p!} d^{p+1} f(x + th)(h)^{p+1} dt.$$

Dans le cas où $E = \mathbb{R}$, la version générale de la « formule de la moyenne » démontrée à l'exercice 11.31 (c) implique qu'il existe $\theta \in]0, 1[$ tel que

$$\begin{aligned} \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} d^p f(x + th)(h)^p dt &= \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} dt \times d^p f(x + \theta h)(h)^p \\ &= \frac{1}{p!} d^p f(x + \theta h)(h)^p, \end{aligned}$$

d'où

$$(9.7') \quad f(x + h) = f(x) + df(x)(h) + \dots + \frac{1}{(p-1)!} d^{p-1} f(x)(h)^{p-1} + \frac{1}{p!} d^p f(x + \theta h)(h)^p.$$

La formule (9.7') se réduit à $f(x + h) - f(x) = df(x + \theta h)(h)$ pour $p = 1$. Ces derniers résultats ne sont pas vrais pour une fonction à valeurs vectorielles $f : U \rightarrow E$ avec $\dim_{\mathbb{R}} E \geq 2$. Ainsi pour $p = 1$, $E = \mathbb{C}$ et $f(x) = e^{ix}$ sur l'intervalle $[0, 2\pi]$, on a

$$f(2\pi) - f(0) = 0 \neq f'(e^{2i\pi\theta}) \times 2\pi, \quad \forall \theta \in]0, 1[.$$

Sous une hypothèse plus faible, la formule de Taylor avec reste intégral montre l'existence d'un développement limité :

(9.8) Formule de Taylor-Young. *Supposons que $f : U \rightarrow E$ soit $(p-1)$ fois différentiable sur U et que $d^{p-1} f$ soit différentiable en x . Alors pour $h \rightarrow 0$ on a*

$$f(x + h) = f(x) + df(x)(h) + \dots + \frac{1}{p!} d^p f(x)(h)^p + o(\|h\|^p).$$

Démonstration. En effet, on peut appliquer (9.7) à l'ordre $p-1$ et estimer terme à terme le reste intégral

$$\int_0^1 \frac{(1-t)^{p-2}}{(p-2)!} d^{p-1} f(x + th)(h)^{p-1} dt,$$

au moyen de l'égalité de définition

$$d^{p-1} f(x + th) = d^{p-1} f(x) + d^p f(x + th)(th) + \|th\| \varepsilon(th),$$

qui exprime l'hypothèse de différentiabilité de $d^{p-1}f$ au point x . Étant donné $\eta > 0$, il existe $\delta > 0$ tel que $\|h\| \leq \delta$ implique $\|\varepsilon(h)\| \leq \eta$. Comme $\int_0^1 t \frac{(1-t)^{p-2}}{(p-2)!} dt = \frac{1}{p!}$, on trouve alors

$$\begin{aligned} \int_0^1 \frac{(1-t)^{p-2}}{(p-2)!} d^{p-1}f(x)(h)^{p-1} dt &= \frac{1}{(p-1)!} d^{p-1}f(x)(h)^{p-1}, \\ \int_0^1 \frac{(1-t)^{p-2}}{(p-2)!} d^p f(x)(th)(h)^{p-1} dt &= \frac{1}{p!} d^p f(x)(h)^p, \\ \left\| \int_0^1 \frac{(1-t)^{p-2}}{(p-2)!} \|th\| \varepsilon(th)(h)^{p-1} dt \right\| &\leq \frac{\eta}{p!} \|h\|^p \quad \text{pour } \|h\| \leq \delta, \end{aligned}$$

ce qui conclut la preuve de (9.8). \square

On va maintenant donner une expression commode de la dérivée p -ième en termes de différences finies itérées. Pour $h \in G$, on introduit «l'opérateur aux différences finies»

$$\Delta_h f(x) = f(x+h) - f(x).$$

On suppose ici f définie et différentiable sur U , avec U ouvert, de sorte que $\Delta_h f$ est bien définie au voisinage de tout point $x \in U$ si h est assez petit. Pour $x \in U$ et $k \in G$ assez petit, la formule (9.6''') donne

$$(9.9) \quad \Delta_k f(x) = f(x+k) - f(x) = \int_0^1 df(x+tk)(k) dt.$$

Supposons f deux fois différentiable au point x (c'est-à-dire f différentiable au voisinage de x et df différentiable en x). On peut alors écrire

$$(9.9') \quad \Delta_h \Delta_k f(x) = \Delta_k f(x+h) - \Delta_k f(x) = \int_0^1 (df(x+h+tk) - df(x+tk))(k) dt.$$

L'hypothèse de différentiabilité de df en x permet d'écrire

$$df(x+h) = df(x) + d^2 f(x)(h) + \|h\| \varepsilon(h)$$

et on trouve donc par différence

$$\begin{aligned} df(x+h+tk) - df(x+tk) &= (df(x) + d^2 f(x)(h+tk) + \|h+tk\| \varepsilon(h+tk)) \\ &\quad - (df(x) + d^2 f(x)(tk) + \|tk\| \varepsilon(tk)) \\ &= d^2 f(x)(h) + \|h+tk\| \varepsilon(h+tk) - \|tk\| \varepsilon(tk). \end{aligned}$$

En évaluant cette différence (qui est un élément de $\mathcal{L}_c(G, E)$) sur le vecteur k , on obtient

$$\begin{aligned} &\| (df(x+h+tk) - df(x+tk))(k) - d^2 f(x)(h, k) \| \\ &\leq (\|h\| + t\|k\|) \|k\| \times \|\varepsilon(h+tk)\| + (t\|k\|) \|k\| \times \|\varepsilon(tk)\|. \end{aligned}$$

Pour tout $\eta > 0$ on a $\|\varepsilon(h)\| \leq \eta$ si $\|h\| \leq \delta$ est assez petit. Pour $\|h\| \leq \delta/2$ et $\|k\| \leq \delta/2$, on en déduit alors

$$\|(df(x+h+tk) - df(x+tk))(k) - d^2f(x)(h,k)\| \leq \eta(\|h\| + 2t\|k\|)\|k\|.$$

Par intégration sur $[0, 1]$ et compte tenu de l'égalité (9.9') il vient

$$\|\Delta_h \Delta_k f(x) - d^2f(x)(h,k)\| \leq \int_0^1 \eta(\|h\| + 2t\|k\|)\|k\| dt = \eta(\|h\| + \|k\|)\|k\|.$$

Si nous remplaçons (h, k) par $(\lambda h, \lambda k)$ avec λ tendant vers 0, alors η peut être pris arbitrairement petit, et on en déduit après division par λ^2 que

$$(9.10) \quad d^2f(x)(h, k) = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda^2} \Delta_{\lambda h} \Delta_{\lambda k} f(x).$$

Comme

$$\Delta_h \Delta_k f(x) = f(x+h+k) - f(x+h) - f(x+k) + f(x),$$

on voit que les opérateurs Δ_h et Δ_k commutent. Par suite $d^2f(x)(h, k) = d^2f(x)(k, h)$ et nous pouvons énoncer :

(9.11) Corollaire. *Si f est deux fois différentiable en un point x d'un ouvert U , alors $d^2f(x) \in \mathcal{L}_c(G^2, E)$ est toujours une application bilinéaire symétrique.*

Plus généralement, si f est p fois différentiable en $x \in U$, on a

$$(9.12) \quad \boxed{d^p f(x)(h_1, \dots, h_p) = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda^p} \Delta_{\lambda h_1} \dots \Delta_{\lambda h_p} f(x)}$$

et donc $d^p f(x) \in \mathcal{L}_c(G^p, E)$ est p -multilinéaire symétrique. La démonstration est très analogue au cas $p = 2$, on écrit par récurrence

$$\Delta_{h_2} \dots \Delta_{h_p} f(x) = \int_0^1 \dots \int_0^1 d^{p-1}f(x + t_2 h_2 + \dots + t_p h_p)(h_2, \dots, h_p) dt_2 \dots dt_p$$

à l'aide d'une itération de (9.9), ce qui donne

$$\Delta_{h_1} \dots \Delta_{h_p} f(x) = \int_0^1 \dots \int_0^1 (d^{p-1}f(x + h_1 + t_2 h_2 + \dots + t_p h_p) - d^{p-1}f(x + t_2 h_2 + \dots + t_p h_p))(h_2, \dots, h_p) dt_2 \dots dt_p.$$

Le reste de la preuve se poursuit de manière analogue en exprimant la condition de différentiabilité de $g = d^{p-1}f$ en x , qui fait apparaître par différence $d^p f(x)(h_1, \dots, h_p)$ et des termes d'erreur contrôlés par $\eta(\sum \|h_j\|)\|h_2\| \dots \|h_p\|$ au second membre. \square

10. Vrai ou faux ⁽¹³⁾

10.1. On considère la fonction $f : x \mapsto x$ sur l'intervalle $I = [0, 2]$. Parmi les affirmations suivantes, lesquelles sont vraies, lesquelles sont fausses et pourquoi ?

- ☐ $\sum_{i=1}^{2n} \frac{i}{n}$ est une somme de Riemann associée à f sur I .
- ☒ $\sum_{i=1}^{2n} \frac{i}{n^2}$ est une somme de Riemann associée à f sur I .
- ☐ $\frac{2}{n} \sum_{i=1}^n \frac{2i}{n^2}$ est une somme de Riemann associée à f sur I .
- ☒ $\sum_{i=1}^n \frac{4i}{n^2}$ est une somme de Riemann associée à f sur I .
- ☐ $\sum_{i=1}^{2n} \frac{i}{n^2}$ tend vers $\frac{1}{2}$ quand n tend vers l'infini.
- ☒ $\sum_{i=1}^{2n} \frac{i}{n^2}$ tend vers 2 quand n tend vers l'infini.

10.2. Toutes les fonctions considérées sont supposées intégrables sur l'intervalle considéré. Parmi les affirmations suivantes, lesquelles sont vraies, lesquelles sont fausses et pourquoi ?

- ☒ L'intégrale sur $[0, 1]$ d'une fonction négative ou nulle est négative ou nulle.
- ☐ L'intégrale d'une fonction paire sur $[0, 1]$ est positive ou nulle.
- ☒ L'intégrale d'une fonction impaire sur $[-1, 1]$ est nulle.
- ☐ L'intégrale sur $[0, 1]$ d'une fonction minorée par 1 est inférieure ou égale à 1.
- ☐ L'intégrale sur $[-1, 1]$ d'une fonction majorée par 1 est inférieure ou égale à 1.
- ☒ L'intégrale sur $[-1, 1]$ d'une fonction majorée par 2 est inférieure ou égale à 4.
- ☒ Si une fonction f est telle que pour tout $x \in [-1, 1]$, $f(x) < x^3$, alors son intégrale sur $[-1, 1]$ est strictement négative.
- ☒ Si l'intégrale d'une fonction f continue sur $[0, 1]$ vaut y , alors il existe $x \in [0, 1]$ tel que $f(x) = y$.
- ☐ Si l'intégrale d'une fonction f sur $[-1, 1]$ vaut y , alors il existe $x \in [0, 1]$ tel que $f(x) = 2y$.

10.3. Parmi les affirmations suivantes lesquelles sont vraies, lesquelles sont fausses, et pourquoi ?

⁽¹³⁾ Les questions et les exercices qui suivent proviennent du photocopié de Bernard Ycart, proposé aux étudiants de L1 à l'Université de Grenoble I.

- ☐ Toute fonction intégrable sur $[a, b]$ est continue.
- ☒ Si une fonction est continue sur $[a, b]$, sauf en un point, alors f admet une primitive.
- ☒ Toute fonction continue sur $[a, b]$ admet une primitive qui s'annule en b .
- ☐ Toute primitive d'une fonction continue sur $[a, b]$ s'annule en un point de $[a, b]$.
- ☒ Toute primitive d'une fonction continue sur $[a, b]$ est dérivable sur $]a, b[$.
- ☐ Toute primitive d'une fonction continue sur $]a, b[$ est dérivable à droite en a .

10.4. Parmi les affirmations suivantes lesquelles sont vraies, lesquelles sont fausses, et pourquoi ?

- ☐ Toute primitive d'une fonction positive ou nulle est positive ou nulle.
- ☒ Toute primitive d'une fonction négative ou nulle est décroissante.
- ☐ Toute fonction continue est la primitive d'une fonction continue.
- ☐ Si f est une fonction continue, alors $-\cos(f(x))$ est une primitive de $\sin(f(x))$.
- ☐ Si f est une fonction continûment dérivable et ne s'annulant pas, $\ln(f(x))$ est une primitive de $\frac{f'(x)}{f(x)}$.
- ☒ Si f est une fonction continûment dérivable, $\arctan(f(x))$ est une primitive de $\frac{f'(x)}{1+f^2(x)}$.
- ☐ Il existe des primitives de $(1 - x^2)^{-1/2}$ définies sur l'intervalle $[0, 2]$.
- ☒ Il existe des primitives de $(1 - x^2/2)^{-1/2}$ définies sur l'intervalle $[0, 1]$.
- ☒ Il existe des primitives de $|1 - x^2|^{1/2}$ définies sur l'intervalle $[-2, 2]$.
- ☐ Il existe des primitives de $|1 - x^2|^{-1/2}$ définies sur l'intervalle $[-2, 2]$.

10.5. Parmi les égalités suivantes lesquelles sont vraies, lesquelles sont fausses, et pourquoi ?

- ☒ $\int_0^\pi x \sin(x) dx = - \int_0^\pi \frac{x^2}{2} \cos(x) dx .$
- ☐ $\int_0^\pi x \sin(x) dx = \int_0^\pi \cos(x) dx .$
- ☒ $\int_0^\pi x \sin(x) dx = \pi - \int_0^\pi \cos(x) dx .$
- ☐ $\int_0^\pi x \sin(x) dx = \pi - 2 .$
- ☒ $\int_0^\pi x \sin(x) dx = \pi .$
- ☒ $\int_0^\pi x \cos(x) dx = -2 .$
- ☐ $\int_0^\pi x \sin(2x) dx = \left[\sin(2x) \right]_0^\pi - 2 \int_0^\pi \cos(2x) dx .$

$$\boxtimes \quad \int_0^\pi x \cos(2x) dx = 0 .$$

$$\boxtimes \quad \int_0^\pi x \sin^2(x) dx = \int_0^\pi x \cos^2(x) dx .$$

$$\square \quad \int_0^\pi x \sin^2(x) dx = \frac{\pi^2}{2} .$$

10.6. Parmi les égalités suivantes lesquelles sont vraies, lesquelles sont fausses, et pourquoi ?

$$\square \quad \int_0^{\frac{\pi}{2}} \sin(2x) dx = \int_0^{\frac{\pi}{2}} \sin(u) \frac{du}{2} .$$

$$\square \quad \int_0^{\frac{\pi}{2}} \sin(2x) dx = \int_0^{\frac{\pi}{4}} \sin(u) \frac{du}{2} .$$

$$\boxtimes \quad \int_0^{\frac{\pi}{2}} \sin(2x) dx = \int_0^{\frac{\pi}{2}} \sin(u) du .$$

$$\boxtimes \quad \int_0^{\frac{\pi}{8}} \sin(2x) dx = \int_0^{\frac{\sqrt{2}}{2}} \frac{u}{2\sqrt{1-u^2}} du .$$

$$\square \quad \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \sin(2x) dx = \int_{\frac{1}{\pi}}^{\frac{2}{\pi}} \frac{\sin(\frac{1}{u})}{u^2} du .$$

$$\boxtimes \quad \int_0^{\frac{\pi}{2}} \sin(2x) dx = \int_1^{e^\pi} \frac{\sin(\ln(u))}{2u} du .$$

11. Exercices

11.1. Calculer les intégrales suivantes.

$$(a) \quad \int_0^1 (\sqrt{u} + 7u) du \quad (b) \quad \int_0^2 \frac{u^3}{u^4 + 7} du \quad (c) \quad \int_0^{\sqrt[3]{\pi}} u^2 \sin(u^3 + \pi) du$$

$$(d) \quad \int_0^1 \frac{u}{(u^2 + 1)^3} du \quad (e) \quad \int_0^1 \frac{x}{x^2 + 1} dx \quad (f) \quad \int_0^{e-1} \frac{\ln(x+1)}{x+1} dx$$

11.2. Calculer les intégrales suivantes par des changements de variables judicieux.

$$(a) \quad \int_0^7 \frac{dz}{(z+7)^2 + 3} \quad (b) \quad \int_0^3 \frac{dt}{1 - 2t^2} \quad (c) \quad \int_0^1 \frac{u^5}{u^{12} + 1} du$$

$$(d) \quad \int_1^2 (\sqrt{u} + 1) \ln(\sqrt{u} + 1) \frac{du}{\sqrt{u}} \quad (e) \quad \int_a^b \frac{dt}{\sqrt{(a+t)(b-t)}}$$

$$(f) \quad \int_2^4 \frac{\sqrt{x^2 - x + 1}}{\sqrt{x^2 - x + 3}} dx$$

11.3. Calculer les primitives et intégrales suivantes en utilisant une ou plusieurs intégrations par parties.

- (a) la primitive de $\ln(x)$ nulle en 0
 (b) la primitive de $\arctan(x)$ nulle en $\pi/4$
 (c) $\int_0^e \ln(x^3 + 1) dx$ (d) $\int_0^1 x \ln(x^2 + 1) dx$
 (e) $\int_0^5 w e^{2w+1} dw$ (f) $\int_0^\pi x^2 \cos(x) dx$

11.4. Étudier les fonctions suivantes.

- (a) $f(x) = \int_7^x \arctan(z + 1) dz$ (b) $g(x) = \int_{-x^2}^2 \frac{2v + 3}{v + 1} dv$
 (c) $h(x) = \int_x^{2x} \ln(u) du$

11.5. Pour chacune des fonctions suivantes, donnez son domaine de définition puis ses éventuelles primitives (on utilisera des changements de variables adéquats).

- (a) $x \mapsto \frac{(\ln x)^p}{x}$ ($p \in \mathbb{Z}$) (b) $x \mapsto \frac{x}{\sqrt{1 - x^4}}$ (c) $x \mapsto \frac{\cos \ln x}{x}$
 (d) $x \mapsto \frac{x}{\sqrt{x^4 - 1}}$ (e) $x \mapsto \frac{e^x}{1 + e^{2x}}$

11.6. Pour chacune des fonctions suivantes, donnez son domaine de définition puis ses éventuelles primitives (on utilisera des intégrations par parties adéquates).

- (a) $x \mapsto \arcsin x$ (b) $x \mapsto x^p \ln x$ ($p \in \mathbb{N}$)
 (c) $x \mapsto \frac{x}{\cosh^2 x}$ (d) $x \mapsto x \tan^2 x$

11.7. Calculer les primitives de chacune des fractions rationnelles suivantes.

- (a) $x \mapsto \frac{1}{8 - x^3}$ (b) $x \mapsto \frac{1}{1 + x^4}$ (c) $x \mapsto \frac{1}{8 - x^3}$
 (d) $x \mapsto \frac{x^5 + 2}{x^5 - x}$ (e) $x \mapsto \frac{2}{x(x^2 + 1)^3}$

11.8. Calculer les primitives de chacune des fractions rationnelles trigonométriques suivantes.

- (a) $x \mapsto \frac{1 - \tan x}{1 + \tan x}$ (b) $x \mapsto \frac{1}{1 + 2 \sin^2 x}$ (c) $x \mapsto \frac{1}{\sin^4 x + \cos^4 x}$

11.9. Calculer les « intégrales abéliennes » suivantes.

- (a) $x \mapsto \frac{1}{x + \sqrt{x - 1}}$ (b) $x \mapsto \frac{1}{x} \sqrt{\frac{x^2 - 1}{4 - x^2}}$

11.10. Déterminer les primitives $\int e^{ax} \cos(bx) dx$ et $\int e^{ax} \sin(bx) dx$ où $a, b \in \mathbb{R}$.

11.11. Calculer

$$(a) \int_0^1 \frac{1 + \arctan x}{1 + x^2} dx \quad (b) \int_a^b \frac{dx}{x \ln x} \quad (a, b > 1)$$

$$(c) \int_a^{1/a} \frac{\ln x}{1 + x^2} dx \quad (a > 0) \quad (d) \int_0^a e^{-\sqrt{t}} dt$$

11.12. Démontrer les résultats suivants.

$$(a) \lim_{n \rightarrow +\infty} \sum_{k=0}^{n-1} \frac{1}{n+k} = \int_0^1 \frac{1}{1+x} dx = \ln(2)$$

$$(b) \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n \frac{n^2}{(n+k)^2} = \int_0^1 \frac{1}{(1+x)^2} dx = \frac{1}{2}$$

$$(c) \lim_{n \rightarrow +\infty} n \sum_{k=n}^{2n-1} \frac{1}{k^2} = \int_0^1 \frac{1}{(1+x)^2} dx = \frac{1}{2}$$

$$(d) \lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{1}{\sqrt{n^2 + 2kn}} = \int_0^1 \frac{1}{\sqrt{1+2x}} dx = \sqrt{3} - 1$$

$$(e) \lim_{n \rightarrow +\infty} \sum_{k=0}^{2n} \frac{k}{k^2 + n^2} = \int_0^2 \frac{x}{1+x^2} dx = \frac{1}{2} \ln(5)$$

$$(f) \lim_{n \rightarrow +\infty} \sum_{k=1}^n \sqrt{\frac{n-k}{n^3 + n^2 k}} = \int_0^1 \sqrt{\frac{1-x}{1+x}} dx = \frac{\pi}{2} - 1.$$

11.13. Évaluer les limites de sommes de Riemann suivantes

$$(a) \lim_{n \rightarrow +\infty} \frac{1}{n^3} \sum_{k=1}^n k^2 \sin\left(\frac{k\pi}{n}\right)$$

$$(b) \lim_{n \rightarrow +\infty} \prod_{k=1}^n \left(1 + \frac{k^2}{n^2}\right)^{1/n}$$

$$(c) \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) f'\left(\frac{k+1}{n}\right) \text{ où } f \in C^1([0, 1]).$$

11.14. Soit $f \in C^0([a, b])$ telle que $f \geq 0$. Montrer que

$$\lim_{n \rightarrow +\infty} \left(\int_a^b f(x)^n dx \right)^{1/n} = \sup_{x \in [a, b]} f(x).$$

11.15. Soit $f \in C^0([0, 1])$. Déterminer $\lim_{n \rightarrow +\infty} n \int_0^1 x^n f(x) dx$.

11.16. Calculer $\lim_{x \rightarrow 0} \int_x^{3x} \frac{\cos t}{t} dt$.

11.17. Calculer $\int_0^1 \left(E\left(\frac{2}{x}\right) - 2 E\left(\frac{1}{x}\right) \right) dx$ où $E(\cdot)$ désigne la fonction partie entière.

11.18. On considère la fonction $f : x \mapsto x^{-2}$, sur l'intervalle $[1, 2]$. Soit $n \geq 1$. Soit D la subdivision régulière de $[1, 2]$ en n intervalles égaux.

(a) Démontrer que

$$S_D(f) = n \sum_{j=0}^{n-1} \frac{1}{(n+j)^2}.$$

(b) Démontrer que

$$n \sum_{j=0}^{n-1} \frac{1}{(n+j)(n+j+1)} < S_D(f) < n \sum_{j=0}^{n-1} \frac{1}{(n+j-1)(n+j)}.$$

(c) En déduire que

$$\frac{1}{2} < S_D(f) < \frac{n(n-2)}{n(2n-1)}.$$

Indication : remarquer que

$$\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}.$$

(d) En déduire que

$$\lim_{n \rightarrow \infty} S_D(f) = \frac{1}{2}.$$

11.19. Soient a et b deux réels tels que $a < b$. On considère la fonction exponentielle $f : x \mapsto e^x$.

(a) Soit $n \geq 1$ un entier. On considère la subdivision D_1 de $[a, b]$ en n intervalles égaux, pointée par la borne de gauche de chaque intervalle. Démontrer que

$$S_{D_1}(f) = \frac{(b-a)e^a}{n} \frac{e^{b-a} - 1}{e^{(b-a)/n} - 1}.$$

(b) Soit D_2 la subdivision de $[a, b]$ en n intervalles égaux, pointée par la borne de droite de chaque intervalle. Démontrer que

$$S_{D_2}(f) = e^{(b-a)/n} S_{D_1}(f).$$

(c) Démontrer que

$$\lim_{n \rightarrow \infty} S_{D_1}(f) = \lim_{n \rightarrow \infty} S_{D_2}(f) = e^b - e^a.$$

11.20. Calculer le volume du paraboloïde tronqué $0 \leq \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq \frac{z}{h} \leq 1$ avec $a, b, h > 0$.

11.21. Calculer le volume du tore de grand rayon R et de petit rayon $r < R$ engendré par la révolution autour de l'axe Oz du cercle de centre $(R, 0, 0)$ et de rayon r situé dans le plan xOz .

11.22. Convergence de sommes de Riemann de fonctions non bornées.

Soit $f : [0, 1] \rightarrow \mathbb{R}$ la fonction f telle que $f(x) = \ln x$ si $x \in]0, 1]$ et $f(0) = 0$. On introduit la subdivision pointée $D = \{([a_j, a_{j+1}], x_j)\}_{0 \leq j < N}$ telle que $a_0 = x_0 = 0$ et, pour $j \geq 1$, $a_j = p^{j-N}$ où $p > 1$ est un réel quelconque. On prend enfin $x_j = a_j$, et on définit de même une subdivision D' en prenant $x_j = a_{j+1}$.

(a) Montrer que

$$S_D(f) = -(p-1) \ln p \sum_{i=1}^{N-1} ip^{-i}.$$

(b) Vérifier que

$$\sum_{i=1}^{N-1} ix^i = x \left(\sum_{i=0}^{N-1} x^i \right)' = x \frac{1 + (N-1)x^N - Nx^{N-1}}{(1-x)^2}$$

et en déduire l'expression de $S_D(f)$ en fonction de p et de N . Calculer de même $S_{D'}(f)$.

(c) On choisit $N \geq 2$ et p tel que $p^N = N^2$. Montrer que $\lim_{N \rightarrow +\infty} S_D(f) = -1$ (on pourra observer que $p \rightarrow 1_+$ et que $\lim_{p \rightarrow 1_+} \ln p / (p-1) = (\ln p)'_{p=1} = 1$).

(d) Justifier l'intégrabilité de f au sens de Kurzweil-Henstock et montrer que $\int_0^1 f(x) dx = -1$.

11.23. Ensembles exceptionnels dénombrables.

(a) Montrer qu'une fonction $f : [a, b] \rightarrow \mathbb{R}$ qui est nulle en dehors d'une partie dénombrable $E = \{u_i\}_{i \in \mathbb{N}}$ de $[a, b]$ est intégrable au sens de Kurzweil-Henstock et que son intégrale $\int_a^b f(x) dx$ est nulle.

Indication : choisir une fonction jauge δ telle que $\delta(u_i) \leq 2^{-i}\varepsilon / (1 + |f(u_i)|)$.

(b) Montrer que si deux fonctions $g, h : [a, b] \rightarrow \mathbb{R}$ diffèrent uniquement sur une partie dénombrable $E = \{u_i\}_{i \in \mathbb{N}}$ de $[a, b]$, alors l'intégrabilité de l'une au sens de Kurzweil-Henstock équivaut à l'intégrabilité de l'autre, et que leurs intégrales sont égales.

11.24. Soit f la fonction indicatrice de \mathbb{Q} , telle que

$$f(x) = 1 \quad \text{si } x \in \mathbb{Q}, \quad f(x) = 0 \quad \text{si } x \notin \mathbb{Q}.$$

On rappelle que tout intervalle ouvert non vide de \mathbb{R} contient un rationnel et un irrationnel. Soit n un entier strictement positif. Pour $j = 0, \dots, n$, on pose $a_j = j/n$.

- (a) Montrer à l'aide de l'exercice précédent que f est intégrable au sens de Kurzweil-Henstock sur l'intervalle $[0, 1]$.
- (b) Soit $([a_j, a_{j+1}])$ une subdivision de $[0, 1]$. Montrer que pour tout $j = 0, \dots, n-1$, il existe $x_j, y_j \in [a_j, a_{j+1}]$ tels que $f(x_j) = 1$ et $f(y_j) = 0$.
- (c) On considère les deux subdivisions pointées

$$D_1 = \{([a_j, a_{j+1}], x_j)\}_{0 \leq j < n} \quad \text{et} \quad D_2 = \{([a_j, a_{j+1}], y_j)\}_{0 \leq j < n}.$$

Montrer que $S_{D_1}(f) = 1$ et $S_{D_2}(f) = 0$.

- (d) En déduire que f n'est pas intégrable au sens de Riemann sur $[0, 1]$.

11.25. Soit f une fonction de $[a, b]$ dans \mathbb{R} et n un entier. On note A_n l'intégrale suivante, si elle existe.

$$A_n = \int_a^b f(x) \sin(nx) dx.$$

- (a) Soit f une fonction dérivable sur $[a, b]$, dont la dérivée est continue et bornée sur $]a, b[$. Démontrer que pour tout $n \in \mathbb{N}$, A_n existe. En utilisant une intégration par parties, démontrer que A_n tend vers 0 quand n tend vers l'infini.
- (b) Soit f une fonction en escalier sur $[a, b]$. Démontrer que pour tout $n \in \mathbb{N}$, A_n existe. En utilisant la relation de Chasles, démontrer que A_n tend vers 0 quand n tend vers l'infini.
- (c) Soit f une fonction continue sur $[a, b]$. Démontrer que pour tout $n \in \mathbb{N}$, A_n existe. En utilisant un encadrement par des fonctions en escalier, démontrer que A_n tend vers 0 quand n tend vers l'infini.
- (d) Soit f une fonction continue par morceaux sur $[a, b]$. Démontrer que pour tout $n \in \mathbb{N}$, A_n existe. Démontrer que A_n tend vers 0 quand n tend vers l'infini.

11.26. Soit f une fonction continue de $[a, b]$ dans \mathbb{R} , telle que :

$$\int_a^b f(x) dx = (b-a) \sup_{x \in [a, b]} f(x).$$

Démontrer que f est constante.

11.27. Soit f une fonction continue sur $[0, 1]$ telle que $\int_0^1 f(x) dx = 1/2$. Montrer qu'il existe $a \in [0, 1]$ tel que $f(a) = a$.

11.28. Soit f une fonction continue sur $[0, 1]$.

- (a) Soit g une fonction en escalier de $[0, 1]$ dans \mathbb{R} . Démontrer que la fonction fg est intégrable sur $[0, 1]$.

(b) On suppose que pour toute fonction en escalier g de $[0, 1]$ dans \mathbb{R} , on a

$$\int_0^1 f(x)g(x) dx = 0.$$

Montrer que la fonction f est nulle.

11.29. Soient a, b, c trois réels tels que $a \leq c < b$. Pour $n \in \mathbb{N}^*$, on définit la fonction g_n sur $[a, b]$ par :

$$g_n(x) = \begin{cases} n & \text{si } x \in [c, c + 1/n] \\ 0 & \text{sinon.} \end{cases}$$

(a) Soit f une fonction continue sur $[a, b]$. Démontrer que

$$\lim_{n \rightarrow \infty} \int_a^b f(x)g_n(x) dx = f(c)$$

(b) Soit f une fonction croissante sur $[a, b]$. Démontrer que

$$\lim_{n \rightarrow \infty} \int_a^b f(x)g_n(x) dx = \lim_{x \rightarrow c+} f(x)$$

11.30. Soient f et g deux fonctions continues de $[a, b]$ dans \mathbb{R} .

- (a) Démontrer que les fonctions f^2 , fg , g^2 sont intégrables sur $[a, b]$.
 (b) Soit λ un réel quelconque. Démontrer que la fonction $(f + \lambda g)^2$ est intégrable sur $[a, b]$ et exprimer son intégrale en fonction des intégrales de f^2 , fg et g^2 .
 (c) En observant que l'intégrale de $(f + \lambda g)^2$ est positive ou nulle pour tout $\lambda \in \mathbb{R}$, démontrer l'inégalité de Cauchy-Schwarz :

$$\left(\int_a^b f(x)g(x) dx \right)^2 \leq \left(\int_a^b f^2(x) dx \right) \left(\int_a^b g^2(x) dx \right).$$

(d) Démontrer que l'égalité a lieu si et seulement si f et g sont proportionnelles.

11.31. On considère deux fonctions $p, f : [a, b] \rightarrow \mathbb{R}$ et on suppose $p \geq 0$ (une telle fonction p est appelée « fonction poids »).

(a) (*Formule de la moyenne généralisée*) Si p et f sont continues, montrer qu'il existe un point $c \in]a, b[$ tel que

$$\int_a^b p(x)f(x) dx = f(c) \int_a^b p(x) dx.$$

Indication : Si $p = 0$ il n'y a rien à montrer. Sinon, posons

$$m = \inf_{[a,b]} f, \quad M = \sup_{[a,b]} f \quad \text{et} \quad A = \int_a^b p(x) dx.$$

Observer que $m \leq \frac{1}{A} \int_a^b p(x)f(x) dx \leq M$. Dans le cas où l'une des inégalités est une égalité, montrer que f doit être constante égale à m (resp. M) sur l'ouvert $\{p(x) > 0\}$ et conclure que l'on peut prendre dans tous les cas $c \in]a, b[$.

- (b) Montrer que la dérivée $f(x) = F'(x)$ d'une fonction différentiable satisfait le théorème des valeurs intermédiaires, c'est-à-dire que l'image d'un intervalle est un intervalle.

Indication : Il suffit de montrer que pour tout segment $[\alpha, \beta] \subset [a, b]$, l'image $f([\alpha, \beta])$ contient l'intervalle $[f(\alpha), f(\beta)]$. Pour cela on appliquera le théorème des valeurs intermédiaires aux deux fonctions continues

$$h(x) = \begin{cases} \frac{F(x) - F(\alpha)}{x - \alpha} & \text{si } x \in]\alpha, \beta], \\ F'(\alpha) = f(\alpha) & \text{si } x = \alpha, \end{cases} \quad \tilde{h}(x) = \begin{cases} \frac{F(x) - F(\beta)}{x - \beta} & \text{si } x \in [\alpha, \beta[, \\ F'(\beta) = f(\beta) & \text{si } x = \beta, \end{cases}$$

et on conclura par le théorème des accroissements finis élémentaire .

- (c) (*Variante de la formule de la moyenne*). On suppose ici que p est de classe C^1 et que $f = F'$ est la dérivée d'une fonction différentiable. Montrer à l'aide d'une intégration par parties que la fonction $pf = pF'$ est KH-intégrable, et démontrer qu'il existe un point $c \in]a, b[$ tel que

$$\int_a^b p(x)f(x) dx = f(c) \int_a^b p(x)dx.$$

Indication : Si on a $\int_a^b p(x)(M - f(x)) dx = 0$ et si $[\alpha, \beta]$ est un intervalle sur lequel $p(x) > 0$, montrer par dérivation de $\int_\alpha^x p(t)(M - f(t)) dt$ que l'on doit avoir $f(x) = M$ sur $[\alpha, \beta]$. En déduire que l'on peut prendre dans tous les cas $c \in]a, b[$.

Chapitre II

Calcul des intégrales en plusieurs variables

L'objectif de ce chapitre est d'abord de définir les intégrales à plusieurs variables par la méthode de Kurzweil-Henstock, mais c'est surtout d'établir les principales formules permettant le calcul effectif de telles intégrales :

- Théorème de Fubini
- Formule du changement de variable et jacobien
- Formes différentielles et calcul extérieur
- Formule de Stokes.

Ces formules sont indispensables en Physique, notamment en dimensions 2 et 3. Il est cependant beaucoup plus commode de se placer d'emblée en dimension arbitraire de manière à unifier les notations – et accessoirement, pour éviter la répétition fastidieuse de démonstrations particulières qui occultent la simplicité de la preuve générale. Nous situons ce Chapitre comme étant au niveau du L2 dans l'idéal (et de manière plus réaliste, au niveau du L3 dans les conditions actuelles, ce qui vient malheureusement trop tard pour faire de la Physique...)

1. Intégration sur un pavé fermé borné

La plus grande partie des résultats que nous avons obtenus au Chapitre I fonctionnent tout aussi bien avec des fonctions f de plusieurs variables, $(x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n)$, définies sur des pavés fermés bornés

$$(1.1) \quad P = [a_1, b_1] \times \dots \times [a_n, b_n], \quad a_i < b_i.$$

Le volume et le diamètre du pavé P sont par définition

$$(1.2) \quad \text{vol}(P) = \prod_{1 \leq i \leq n} (b_i - a_i), \quad \text{diam}(P) = \max_{1 \leq i \leq n} \{b_i - a_i\}$$

(il sera commode d'utiliser la norme $\|x\| = \max_{1 \leq i \leq n} \{|x_i|\}$ et la distance associée $d(x, y) = \max_{1 \leq i \leq n} \{|x_i - y_i|\}$ sur \mathbb{R}^n dans tout le reste de ces notes). Dans le cas de la dimension $n = 1$, le volume se réduit naturellement à la notion de longueur d'un intervalle, et dans le cas de la dimension 2 à celui de l'aire d'un rectangle.

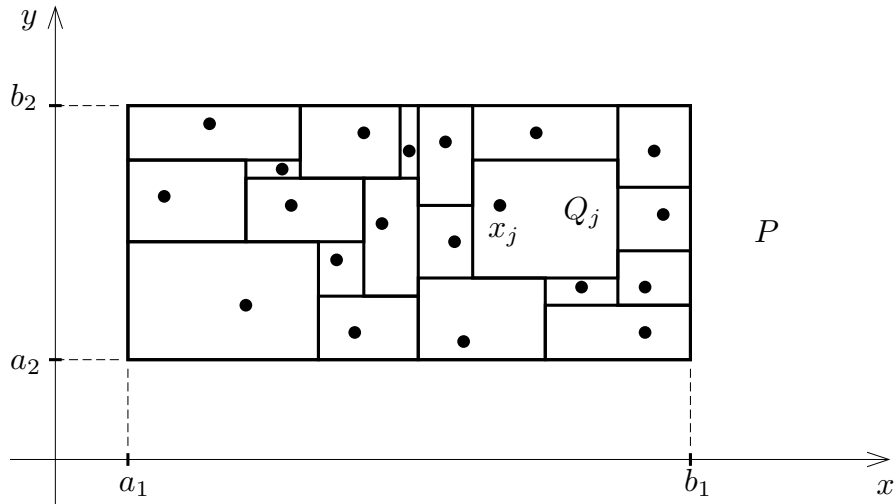


Fig. 13. Subdivision pointée d'un pavé P de \mathbb{R}^n .

On appelle *subdivision* de P toute famille $D = (Q_j)_{0 \leq j < N}$ formée de pavés fermés $Q_j \subset P$ d'intérieurs disjoints (et non vides) tels que $P = \bigcup Q_j$. On dit que l'on a une *subdivision pointée* de P si on s'est donné de plus des points $x_j \in Q_j$ dans chacun des pavés de la subdivision, et on note $D = (Q_j, x_j)_{0 \leq j < N}$ une telle famille.

Notons qu'une subdivision (Q_j) n'est pas nécessairement une subdivision produit (c'est-à-dire une subdivision obtenue en prenant des subdivisions de chacune des arêtes $[a_i, b_i]$ de P et les pavés produits correspondants). Cependant, en considérant les projections de tous les sommets des pavés Q_j sur chacun des axes, on définit une subdivision produit (Q'_k) plus fine que (Q_j) (c'est-à-dire telle que chaque Q_j est la réunion d'un nombre fini des Q'_k). Dans tous les cas on a

$$(1.3) \quad \text{vol}(P) = \sum_{0 \leq j < N} \text{vol}(Q_j) ;$$

en effet, si (Q_j) est une subdivision produit, c'est évident par distributivité de la multiplication par rapport à l'addition, et le cas général s'y ramène en considérant la subdivision produit (Q'_k) de P qui décompose simultanément tous les pavés Q_j .

La *somme de Riemann* associée à f sur $D = (Q_j, x_j)_{0 \leq j < N}$ est définie comme

$$(1.4) \quad S_D(f) = \sum_{0 \leq j < N} \text{vol}(Q_j) f(x_j).$$

Si $\delta : P \rightarrow \mathbb{R}_+^*$ est une *jauge* (c'est-à-dire une fonction positive arbitraire), on dit que la subdivision D est δ -fine si $\text{diam}(Q_j) \leq \delta(x_j)$ pour tout j . L'existence de subdivisions δ -fines se généralise aussitôt à la dimension n , en utilisant une dichotomie comme dans la preuve du lemme I.2.6 (on divise chacune des arêtes en 2 moitiés selon les axes x_1, \dots, x_n , ce qui produit 2^n pavés à chaque étape). La définition I.2.7 peut se calquer mot à mot dans le cas des fonctions de plusieurs variables.

(1.5) Définition. Soit $f : P \rightarrow \mathbb{R}$ une fonction définie sur un pavé fermé borné.

- (a) On dit que f est intégrable au sens de Kurzweil-Henstock (ou KH-intégrable) sur P s'il existe un réel A tel que pour tout $\varepsilon > 0$ donné a priori, on peut trouver une jauge $\delta : P \rightarrow \mathbb{R}_+^*$ (dite ε -adaptée à f), en sorte que pour toute subdivision pointée $D = \{(Q_j, x_j)\}$ de P on ait

$$\text{diam}(Q_j) \leq \delta(x_j) \text{ pour tout } j \Rightarrow |S_D(f) - A| \leq \varepsilon.$$

- (b) De même, on dit que f est intégrable au sens de Riemann (ou Riemann-intégrable) sur P si on peut de plus choisir la jauge δ constante, c'est-à-dire si pour tout $\varepsilon > 0$ il existe une constante $\delta > 0$ telle que

$$\text{diam}(Q_j) \leq \delta \text{ pour tout } j \Rightarrow |S_D(f) - A| \leq \varepsilon.$$

Dans l'un ou l'autre cas, on note

$$\int_P f(x_1, \dots, x_n) dx_1 \dots dx_n,$$

ou même simplement $A = \int_P f(x) dx$, et on appelle A l'intégrale de f sur le pavé P .

L'un des principaux critères d'intégrabilité que nous utiliserons dans la suite est le critère de Cauchy.

(1.6) Critère de Cauchy. Soit $f : P \rightarrow \mathbb{R}$ une fonction définie sur un pavé P fermé borné. Pour que f soit KH-intégrable (resp. Riemann-intégrable), il faut et il suffit que pour tout $\varepsilon > 0$ il existe une jauge $\delta : P \rightarrow \mathbb{R}_+^*$ (resp. une constante $\delta > 0$), telle que pour toutes subdivisions pointées D et D' δ -fines on ait $|S_{D'}(f) - S_D(f)| \leq \varepsilon$.

Démonstration. Si f est KH-intégrable d'intégrale A , pour chaque jauge δ qui est $\varepsilon/2$ -adaptée à f , les inégalités $|S_D(f) - A| \leq \varepsilon/2$ et $|S_{D'}(f) - A| \leq \varepsilon/2$ pour D, D' δ -fines entraînent $|S_{D'}(f) - S_D(f)| \leq \varepsilon$. La réciproque est une conséquence de la complétude de \mathbb{R} , ou, de façon équivalente, de l'existence de bornes supérieures et inférieures pour les parties bornées de \mathbb{R} . En effet, supposons que pour tout entier $n \geq 1$ il existe une jauge δ_n telle que

$$|S_{D'}(f) - S_D(f)| \leq \varepsilon_n = 1/n \quad \text{lorsque } D \text{ et } D' \text{ sont } \delta_n\text{-fines.}$$

Quitte à remplacer δ_n par $\min(\delta_1, \delta_2, \dots, \delta_n)$ on peut supposer la suite δ_n décroissante. L'encadrement précédent montre que les quantités

$$M_\ell = \sup \{S_{D'}(f); D' \text{ } \delta_\ell\text{-fine}\}$$

sont bornées et vérifient $|M_\ell - S_D(f)| \leq 1/n$ pour tout $\ell \geq n$ et toute subdivision D δ_n -fine. De plus la suite (M_ℓ) est décroissante bornée, et si on pose $A = \inf\{M_\ell\}$, on trouve $|A - S_D(f)| \leq 1/n$ pour toute subdivision D δ_n -fine. Ceci implique

$$\lim_{\text{KH}, D} S_D(f) = A = \inf_{\ell > 0} M_\ell, \quad \text{donc } f \text{ KH-intégrable d'intégrale } A. \quad \square$$

On dit qu'une subdivision pointée $\tilde{D} = (\tilde{Q}_k, \tilde{x}_k)$ est *emboîtée* dans une autre subdivision pointée $D = (Q_j, x_j)$ si tout pavé Q_j de D se décompose comme une réunion

$$Q_j = \bigcup_{\tilde{Q}_k \subset Q_j} \tilde{Q}_k$$

de pavés de \tilde{D} (cf. Fig. 14).

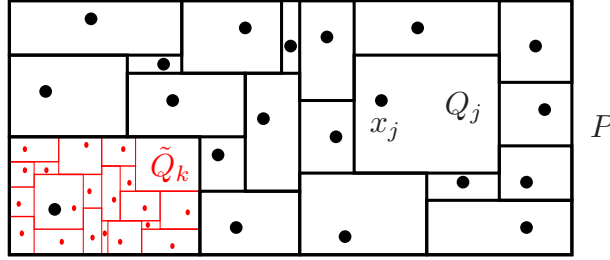


Fig. 14. Subdivisions emboîtées D et \tilde{D} .

Étant donné deux subdivisions pointées $D' = (Q'_j, x'_j)$ et $D'' = (Q''_k, x''_k)$ δ -fines, on peut produire une subdivision pointée $\tilde{D} = (\tilde{Q}_\ell, \tilde{x}_\ell)$ δ -fine emboîtée à la fois dans D' et D'' en considérant tous les pavés intersections $Q'_j \cap Q''_k$ d'intérieur non vide, et en prenant une subdivision pointée δ -fine de chacune de ces intersections. Si le critère de Cauchy est satisfait pour des subdivisions emboîtées, on a alors $|S_{D'}(f) - S_{\tilde{D}}(f)| \leq \varepsilon$ et $|S_{D''}(f) - S_{\tilde{D}}(f)| \leq \varepsilon$, donc $|S_{D''}(f) - S_{D'}(f)| \leq 2\varepsilon$ pour D' et D'' δ -fines quelconques.

(1.7) Conséquence. *Pour appliquer le critère de Cauchy, il suffit de considérer des paires de subdivisions D, D' emboîtées.*

On exploite cette observation à l'aide du lemme schématisé par la figure ci-dessous.

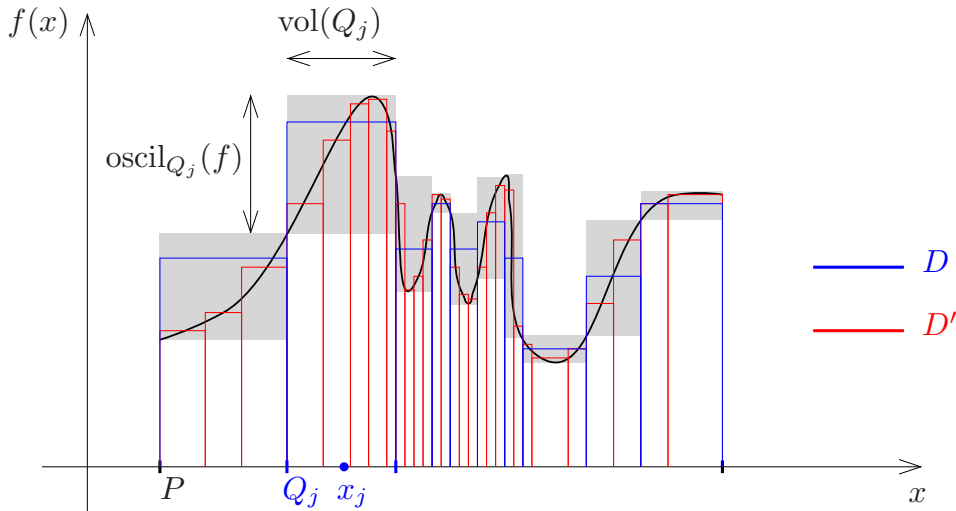


Fig. 15. Écart entre sommes de Riemann emboîtées.

(1.8) Lemme. Si D' est emboîtée dans $D = (Q_j, x_j)$, on a la majoration

$$|S_{D'}(f) - S_D(f)| \leq \sum_j \text{vol}(Q_j) \text{oscil}_{Q_j}(f)$$

où $\text{oscil}(f) = \sup_{Q_j} \sup_{x,y \in Q_j} |f(x) - f(y)|$ désigne l'oscillation de f sur Q_j .

Démonstration. Chaque terme $\text{vol}(Q_j)f(x_j)$ de $S_D(f)$ correspond dans $S_{D'}(f)$ à une sommation $\sum_{Q'_k \subset Q_j} \text{vol}(Q'_k)f(x'_k)$. Comme $\text{vol}(Q_j) = \sum_{Q'_k \subset Q_j} \text{vol}(Q'_k)$, on a

$$\begin{aligned} \left| \text{vol}(Q_j)f(x_j) - \sum_{Q'_k \subset Q_j} \text{vol}(Q'_k)f(x'_k) \right| &= \left| \sum_{Q'_k \subset Q_j} \text{vol}(Q'_k)(f(x_j) - f(x'_k)) \right| \\ &\leq \left(\sum_{Q'_k \subset Q_j} \text{vol}(Q'_k) \right) \text{oscil}_{Q_j}(f) = \text{vol}(Q_j) \text{oscil}_{Q_j}(f). \end{aligned}$$

L'inégalité du lemme s'obtient en sommant sur l'indice j . □

(1.9) Corollaire. Si f est KH-intégrable sur le pavé fermé borné P et si $D = (Q_j, x_j)$ est une subdivision pointée de P , on a

$$\left| \int_P f(x) dx - S_D(f) \right| \leq \sum_j \text{vol}(Q_j) \text{oscil}_{Q_j}(f).$$

Démonstration. Ceci résulte du lemme 1.8 et du fait que $\int_P f(x) dx = \lim_{KH, D'} S_{D'}(f)$ en prenant la limite sur les subdivisions pointées D' emboîtées dans D . □

(1.10) Théorème. Toute fonction continue $f : P \rightarrow \mathbb{R}$ définie sur un pavé fermé borné est intégrable au sens de Riemann (et donc KH-intégrable) sur P .

Démonstration. Pour tout $x \in P$, la continuité de f en x montre qu'il existe $\delta(x) > 0$ tel que pour tout pavé Q contenant x de diamètre $\leq \delta(x)$ on ait $|f(y) - f(x)| \leq \varepsilon$, et donc $\text{oscil}_Q(f) \leq 2\varepsilon$. Si D et D' sont des subdivisions δ -fines emboîtées, le lemme 1.8 implique $|S_{D'}(f) - S_D(f)| \leq 2\varepsilon \sum \text{vol}(Q_j) = 2\varepsilon \text{vol}(P)$, ce qui démontre l'intégrabilité au sens de Kurzweil-Henstock. En réalité, f est uniformément continue et on peut remplacer δ par une jauge constante : si η est le minimum des arêtes des pavés intervenant dans une subdivision δ -fine quelconque D_0 fixée, tout pavé Q de diamètre $\leq \eta$ va rencontrer l'un des pavés Q_j^0 de D_0 et sera contenu dans la réunion de Q_j^0 avec les pavés Q_k^0 qui lui sont adjacents, de sorte que $\text{oscil}_Q(f) \leq 4\varepsilon$. Ceci implique que $|S_{D'}(f) - S_D(f)| \leq 4\varepsilon \text{vol}(P)$ pour tout couple de subdivisions D et D' emboîtées η -fines. L'intégrabilité au sens de Riemann s'ensuit. □

D'autre part, nous avons les résultats basiques suivants qui généralisent la formule de Chasles.

(1.11) Proposition. Soit $f : P \rightarrow \mathbb{R}$ une fonction KH-intégrable (resp. Riemann-intégrable) sur un pavé fermé borné P . Alors la restriction $f|_Q$ de f à tout pavé fermé $Q \subset P$ est encore KH-intégrable (resp. Riemann-intégrable).

Démonstration. On utilise le critère de Cauchy. Étant donné $\varepsilon > 0$, il existe une jauge δ sur P telle que $|S_D(f) - S_{D'}(f)| \leq \varepsilon$ pour toutes subdivisions pointées δ -fines D et D' de P . Soient maintenant Δ et Δ' deux subdivisions pointées de Q qui sont $\delta|_Q$ -fines. Comme $P \setminus Q$ est une réunion finie de pavés d'intérieurs disjoints, on peut compléter Δ et Δ' en des subdivisions pointées D et D' de P qui sont δ -fines (et on peut naturellement prendre le même découpage et les mêmes points x_j sur $P \setminus Q$). On obtient alors

$$|S_\Delta(f) - S_{\Delta'}(f)| = |S_D(f) - S_{D'}(f)| \leq \varepsilon.$$

La proposition est démontrée. Pour l'intégrabilité au sens de Riemann la preuve est tout à fait analogue. \square

(1.12) Proposition (relation de Chasles). *Étant donné un découpage $P = \bigcup P_i$ d'un pavé fermé borné P en pavés fermés P_i d'intérieurs disjoints, une fonction $f : P \rightarrow \mathbb{R}$ est KH-intégrable sur P si et seulement si elle est intégrable sur chaque pavé P_i , et on a alors*

$$\int_P f(x) dx = \sum_i \int_{P_i} f(x) dx.$$

Démonstration. Pour simplifier l'exposé, on va procéder d'une façon un peu plus « brutale » que ce que nous avons fait en dimension 1 (bien qu'une méthode analogue reste possible). Choisissons des fonctions jauge ε -adaptées δ_i sur chaque pavé P_i . On considère une jauge δ sur $P = \bigcup P_i$ telle que $\delta \leq \delta_i$ sur P_i , en imposant la condition supplémentaire suivante: soit ∂P_i le bord de P_i (réunion de chacune des faces pour lesquelles l'une des coordonnées x_k est constante); pour tout $x \in P$, on demande que δ satisfasse

$$\delta(x) \leq \min_{i, \partial P_i \not\ni x} d(x, \partial P_i).$$

De cette manière, si $x \in P$ et si $Q \subset P$ est un pavé contenant x de diamètre $\text{diam}(Q) \leq \delta(x)$, ou bien x est dans l'intérieur P_i° d'un des pavés, auquel cas $\delta(x) \leq d(x, \partial P_i)$ et donc $Q \subset P_i$, ou bien x appartient à certains bords $\partial P_{i_1}, \dots, \partial P_{i_N}$. Dans ce cas x n'appartient à aucun autre pavé P_i , et les inégalités $\text{diam}(Q) \leq \delta(x) \leq d(x, \partial P_i)$ impliquent $Q \subset \mathbb{C}P_i^\circ$. Il en résulte que $Q \subset P_{i_1} \cup \dots \cup P_{i_N}$ et donc dans tous les cas $Q = \bigcup_{i, P_i \ni x} P_i \cap Q$; on peut évidemment limiter cette union aux indices i tels que l'intérieur $(P_i \cap Q)^\circ$ soit non vide. Soit alors $D = (Q_j, x_j)$ une subdivision δ -fine. D'après ce qui précède, la famille $D_i = \{(P_i \cap Q_j, x_j)\}_{j, x_j \in P_i, (P_i \cap Q_j)^\circ \neq \emptyset}$ constitue une subdivision pointée δ_i -fine de P_i . L'intérêt est qu'il n'est pas nécessaire de changer les points de marquage x_j , parce que ceux-ci appartiennent nécessairement aux faces ∂P_i lorsque les pavés Q_j se découpent suivant plusieurs pavés P_i . On a donc $S_D(f) = \sum_i S_{D_i}(f)$ avec $|S_{D_i}(f) - \int_{P_i} f(x) dx| \leq \varepsilon$ et la proposition 1.12 s'ensuit par passage à la limite. \square

(1.13) Remarque. On peut étendre sans difficulté la définition de l'intégrale $\int_P f(x) dx$ au cas où P est réunion finie de pavés fermés bornés. On se ramène aussitôt au cas des pavés grâce à la relation de Chasles. \square

Nous terminons ce paragraphe par la preuve d'une version élémentaire du théorème de Fubini.

(1.14) Thorème de Fubini (version élémentaire). Soit $f : [a_1, b_1] \times [a_2, b_2] \rightarrow \mathbb{R}$ une fonction continue. Alors

$$(1.14 \text{ a}) \quad \boxed{\int_{[a_1, b_1] \times [a_2, b_2]} f(x, y) \, dx \, dy = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x, y) \, dy \right) dx.}$$

Commentaires. Cet énoncé mérite quelques commentaires préalables. L'intégrale de gauche $\int_{[a_1, b_1] \times [a_2, b_2]} f(x, y) \, dx \, dy$ est celle introduite par la définition 1.5. Le calcul de l'intégrale de droite $\int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x, y) \, dy \right) dx$ se décompose comme suit : on commence par évaluer

$$F(x) = \int_{a_2}^{b_2} f(x, y) \, dy$$

pour tout $x \in [a_1, b_1]$, ce qui, d'après le théorème 6.6 produit une fonction continue $F : [a_1, b_1] \rightarrow \mathbb{R}$. On calcule ensuite l'intégrale ordinaire $\int_{a_1}^{b_1} F(x) \, dx$. Le théorème de Fubini stipule que les deux résultats coïncident. Puisqu'on peut échanger le rôle des coordonnées x et y , on aura aussi comme conséquence que

$$(1.14 \text{ b}) \quad \boxed{\int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x, y) \, dy \right) dx = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x, y) \, dx \right) dy.}$$

Démonstration. Fixons $\varepsilon > 0$. Par continuité uniforme, il existe alors $\delta > 0$, c'est-à-dire une jauge constante, telle que $\text{oscil}_Q(f) = \sup_{\xi, \xi' \in Q} |f(\xi) - f(\xi')| \leq \varepsilon$ sur tout pavé $Q \subset P = [a_1, b_1] \times [a_2, b_2]$ de diamètre $\text{diam}(Q) \leq \delta$. Soit D une subdivision produit δ -fine obtenue en découpant $[a_1, b_1]$ et $[a_2, b_2]$ en N_1 (resp. N_2) sous-intervalles $[x_j, x_{j+1}]$ (resp. $[y_k, y_{k+1}]$) de longueur $\leq \delta$. On note $\ell = (j, k)$, $0 \leq j < N_1$, $0 \leq k < N_2$ et

$$Q_\ell = [x_j, x_{j+1}] \times [y_k, y_{k+1}], \quad \xi_\ell = (x_j, y_k) \in Q_\ell.$$

Avec ces notations, nous avons

$$S_D(f) = \sum_{\ell=(j,k)} \text{vol}(Q_\ell) f(\xi_\ell) = \sum_{j,k} (x_{j+1} - x_j)(y_{k+1} - y_k) f(x_j, y_k)$$

et le corollaire 1.9 donne

$$(1.14 \text{ c}) \quad \left| \int_{[a_1, b_1] \times [a_2, b_2]} f(x, y) \, dx \, dy - S_D(f) \right| \leq \varepsilon \, \text{vol}(P).$$

On a d'autre part

$$|F(x') - F(x)| = \left| \int_{a_2}^{b_2} (f(x', y) - f(x, y)) \, dy \right| \leq \varepsilon(b_2 - a_2) = \varepsilon_2$$

pour tous $x, x' \in [x_j, x_{j+1}]$, et le corollaire 1.9 appliqué à $F : [a_1, b_1] \rightarrow \mathbb{R}$ entraîne

$$(1.14 \text{ d}) \quad \left| \int_{a_1}^{b_1} F(x) \, dx - \sum_j (x_{j+1} - x_j) F(x_j) \right| \leq \varepsilon_2(b_1 - a_1) = \varepsilon(b_1 - a_1)(b_2 - a_2).$$

Enfin, le même raisonnement appliqué à la fonction d'une variable $y \mapsto f(x_j, y)$ relativement à la subdivision pointée $([y_k, y_{k+1}], y_k)$ de $[a_2, b_2]$ implique

$$\begin{aligned} \left| F(x_j) - \sum_k (y_{k+1} - y_k) f(x_j, y_k) \right| \\ = \left| \int_{a_2}^{b_2} f(x_j, y) dy - \sum_k (y_{k+1} - y_k) f(x_j, y_k) \right| \leq \varepsilon (b_2 - a_2), \end{aligned}$$

pour tout indice j , donc en multipliant par $(x_{j+1} - x_j)$ et en sommant on trouve

$$(1.14 \text{ e}) \quad \left| \sum_j (x_{j+1} - x_j) F(x_j) - S_D(f) \right| \leq \varepsilon (b_1 - a_1)(b_2 - a_2).$$

En combinant (1.14 d,e) on trouve

$$(1.14 \text{ f}) \quad \left| \int_{a_1}^{b_1} F(x) dx - S_D(f) \right| \leq 2\varepsilon (b_1 - a_1)(b_2 - a_2).$$

D'après (1.14 c,f) il vient par différence

$$\left| \int_{[a_1, b_1] \times [a_2, b_2]} f(x, y) dx dy - \int_{a_1}^{b_1} F(x) dx \right| \leq 3\varepsilon \text{vol}(P),$$

ce qui conclut la démonstration puisque ε est arbitraire. \square

Il est facile de généraliser le théorème de Fubini aux cas des fonctions continues sur un pavé $P = \prod [a_j, b_j] \subset \mathbb{R}^n$ en toute dimension. On démontre ainsi que

$$(1.15) \quad \boxed{\int_P f(x) dx = \int_{[a_1, b_1] \times \dots \times [a_{n-1}, b_{n-1}]} \left(\int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \right) dx_1 \dots dx_{n-1}}$$

et on en déduit par récurrence sur n qu'une intégration dans \mathbb{R}^n consiste en un calcul itéré d'intégrations à une seule variable :

$$(1.16) \quad \boxed{\int_P f(x) dx = \int_{a_1}^{b_1} \left(\dots \int_{a_{n-1}}^{b_{n-1}} \left(\int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \right) dx_{n-1} \dots \right) dx_1.}$$

L'ordre des intégrations est indifférent puisque la définition de l'intégrale que nous avons donnée est invariante par permutation des coordonnées de \mathbb{R}^n .

2. Intégration sur un domaine ouvert et formule du jacobien en plusieurs variables

Soit Ω une partie ouverte de \mathbb{R}^n . On se propose d'étudier l'intégration des fonctions continues $f : \Omega \rightarrow \mathbb{R}$ et, dans ce cadre, d'établir la formule de changement de variable. L'hypothèse de continuité de la fonction f pourra en fait être levée ultérieurement sans trop de difficultés, par des arguments de passage à limite.

Pour calculer une intégrale sur Ω , on considère des domaines K_p qui sont des réunions finies de pavés fermés d'intérieurs disjoints, de sorte que

$$(2.1) \quad K_0 \subset K_1 \dots \subset K_p \subset \dots \subset \Omega, \quad K_p \subset K_{p+1}^\circ \quad \text{et} \quad \Omega = \bigcup K_p.$$

Une telle suite sera appelée un *pavage exhaustif* de Ω . Une manière simple d'obtenir un pavage exhaustif consiste à construire K_p par récurrence en prenant les cubes constituant $K_p \setminus K_{p-1}^\circ$ dans le « quadrillage » associé à un certain réseau $2^{-s_p} \mathbb{Z}^n$ de côté 2^{-s_p} , pour une suite d'entiers s_p strictement croissante adéquate (on ajoute à K_{p-1} les cubes $2^{-s_p}([0, 1]^n + u)$, $u \in \mathbb{Z}^n \cap [-4^{s_p}, 4^{s_p}[^n$, qui sont contenus dans $\Omega \setminus K_{p-1}^\circ$).

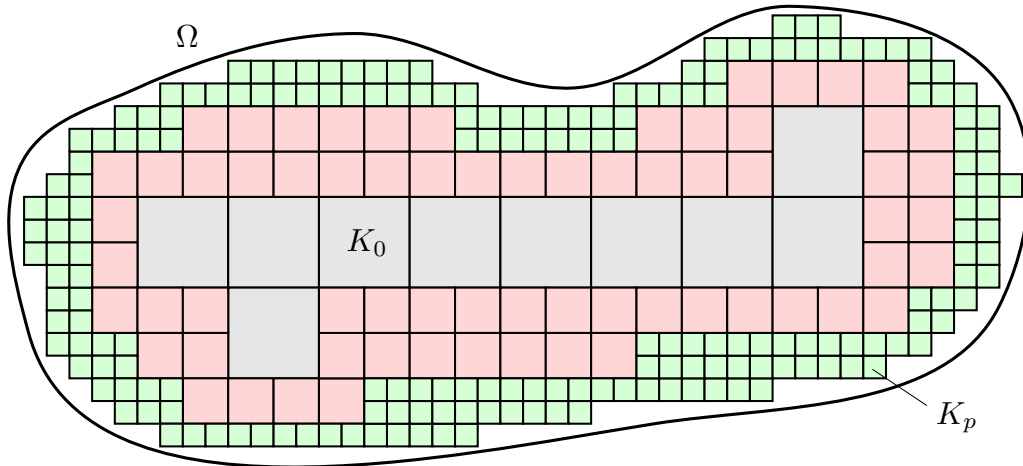


Fig. 16. Pavage exhaustif d'un ouvert Ω de \mathbb{R}^n .

On cherche alors à définir

$$(2.2) \quad \int_{\Omega} f(x) dx = \lim_{p \rightarrow +\infty} \int_{K_p} f(x) dx$$

en supposant qu'il y a *convergence absolue*, c'est-à-dire que la suite croissante d'intégrales $I_p = \int_{K_p} |f(x)| dx$ reste bornée. Comme le domaine K_p est une réunion finie de pavés d'intérieurs disjoints, l'intégrale sur K_p est définie comme la somme des intégrales sur ces pavés (et cette somme est bien indépendante du découpage choisi d'après la proposition 1.12). Dans ce cas le critère de Cauchy montre que pour tout $\varepsilon > 0$ il existe un entier p_0 tel que pour $q \geq p \geq p_0$ on ait $0 \leq I_q - I_p \leq \varepsilon$, donc

$$(2.3) \quad \left| \int_{K_q} f(x) dx - \int_{K_p} f(x) dx \right| = \left| \int_{K_q \setminus K_p^\circ} f(x) dx \right| \leq \int_{K_q \setminus K_p^\circ} |f(x)| dx \leq \varepsilon$$

(le domaine $K_q \setminus K_p^\circ$ étant lui aussi réunion finie de pavés d'intérieurs disjoints). Si K'_p est un autre pavage exhaustif, il existe des indices $q(p)$ et $q'(p)$ tels que $K'_p \subset K_{q(p)}$ et $K_p \subset K'_{q'(p)}$, donc $K_p \setminus (K_p \cap K'_p) \leq K'_{q'(p)} \setminus K'_p$ et $K'_p \setminus (K_p \cap K'_p) \leq K_{q(p)} \setminus K_p$, ce qui montre que la limite

$$\lim_{p \rightarrow +\infty} \int_{K_p} f(x) dx = \lim_{p \rightarrow +\infty} \int_{K_p \cap K'_p} f(x) dx = \lim_{p \rightarrow +\infty} \int_{K'_p} f(x) dx$$

ne dépend pas du pavage exhaustif choisi.

Dans ce qui suit, nous aurons besoin d'une autre forme de « découpage », à savoir la possibilité d'écrire f , sur tout compact K_p , comme une somme finie $f = \sum f_j$ où les f_j sont continues et à *support compact* dans des cubes aussi petits que l'on veut, contenus dans Ω . Pour cela, on écrit par exemple que sur \mathbb{R} on a $1 = \sum_{j \in \mathbb{Z}} \theta(x - j)$ où θ est la fonction sinusoïdale tronquée définie par

$$(2.4 \text{ a}) \quad \theta(x) = \begin{cases} \cos^2(\frac{\pi}{2}x) & \text{si } x \in [-1, 1], \\ 0 & \text{si } x \in \mathbb{R} \setminus [-1, 1]. \end{cases}$$

On a en effet $\theta(x - 1) = \sin^2(\frac{\pi}{2}x)$ si $x \in [0, 2]$, et donc $\theta(x) + \theta(x - 1) = 1$ sur $[0, 1]$.

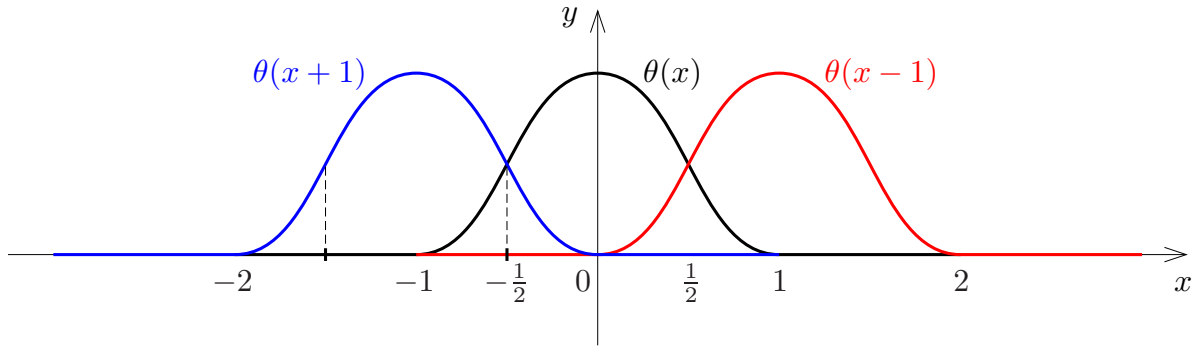


Fig. 17. Découpage par des fonctions continues à support compact.

En n variables on écrit que $\prod_{1 \leq k \leq n} (\sum_{j_k \in \mathbb{Z}} \theta(x_k - j_k)) = 1$. Après une homothétie de rapport assez petit 2^{-t_p} , ceci donne l'analogie en dimension n , à savoir que pour $j = (j_1, \dots, j_n) \in \mathbb{Z}^n$ et $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ on a

$$(2.4 \text{ b}) \quad \sum_{j \in \mathbb{Z}^n} \theta_{j,p}(x) = 1 \quad \text{avec} \quad \theta_{j,p}(x) = \prod_{1 \leq k \leq n} \theta(2^{t_p} x_k - j_k).$$

Il est facile de voir que $\theta_{j,p}$ est à support dans le cube fermé $C_{j,p}$ de centre $2^{-t_p} j$ et de côté 2^{1-t_p} . Alors $f = \sum_{j \in \mathbb{Z}^n} f \theta_{j,p}$ et $f \theta_{j,p}$ est à support dans $C_{j,p}$. Si on restreint la somme à l'ensemble J_p des indices j correspondant aux cubes $C_{j,p}$ contenus dans $\Omega \cap [-2^{t_p}, 2^{t_p}]^n$, alors on obtient une fonction continue à support compact $f \chi_p$, avec $\chi_p = \sum_{j \in J_p} \theta_{j,p}$, $0 \leq \chi_p \leq 1$, χ_p égale à 1 sur K_p pour $t_p \in \mathbb{N}$ assez grand. Dans ces conditions, nous avons

$$(2.5) \quad \int_{\Omega} f(x) dx = \lim_{p \rightarrow +\infty} \int_{\Omega} f(x) \chi_p(x) dx = \lim_{p \rightarrow +\infty} \sum_{j \in J_p} \int_{C_{j,p}} f(x) \theta_{j,p}(x) dx.$$

(2.6) Formule du changement de variable dans \mathbb{R}^n . Soit $\varphi : \Omega \rightarrow \Omega'$ une bijection de classe C^1 entre deux ouverts de \mathbb{R}^n , telle que le « déterminant jacobien »

$$\text{Jac } \varphi(x) = \det \left(\frac{\partial \varphi_i}{\partial x_j} \right)_{1 \leq i, j \leq n}$$

soit partout non nul. Alors pour toute fonction continue $f : \varphi(\Omega) = \Omega' \rightarrow \mathbb{R}$ d'intégrale absolument convergente sur $\varphi(\Omega)$ on a

$$\int_{\varphi(\Omega)} f(y) dy = \int_{\Omega} f(\varphi(x)) |\text{Jac } \varphi(x)| dx$$

et l'intégrale du membre de droite est absolument convergente sur Ω .

Dans la pratique on pose $y = \varphi(x)$, et on écrit que les éléments de volume se transforment par la formule $dy = |\text{Jac } \varphi(x)| dx$.

Démonstration. En dimension 1 et dans le cas d'un intervalle $\Omega =]a, b[\subset \mathbb{R}$, la formule se réduit à celle obtenue en (4.5) avec $|\text{Jac } \varphi(x)| = |\varphi'(x)|$; la valeur absolue provient de ce qu'on ne tient pas compte de l'ordre des bornes pour l'intervalle image $\varphi([a, b])$. Pour obtenir le cas général en dimension quelconque, l'idée est la suivante : tout changement de variable peut s'obtenir comme une composition de changements de variables dans lesquels on change une seule coordonnée à la fois, et dans ce cas, la preuve se ramène au cas élémentaire d'une seule variable à l'aide du théorème de Fubini⁽¹⁴⁾. Pour cela, on procède en plusieurs étapes.

Première étape. On considère le cas particulier d'un changement de variable

$$x = (x_1, \dots, x_n) \xrightarrow{\varphi} y = (u(x_1, \dots, x_n), x_2, \dots, x_n)$$

portant uniquement sur la première coordonnée, avec x dans un pavé $P = \prod [a_i, b_i]$ et $x_1 \mapsto u(x_1, \dots, x_n)$ strictement monotone sur $[a_1, b_1]$. On écrit pour simplifier $x = (x_1, x')$ avec $x' = (x_2, \dots, x_n)$, donc $y = \varphi(x) = (u(x), x')$. On suppose en outre que la fonction f à intégrer est à support compact dans l'image $\varphi(P^\circ)$ du pavé ouvert P° , de sorte que $f \circ \varphi$ est nulle au voisinage du bord ∂P et sur le complémentaire $\mathbb{C}P$. Nous avons

$$\varphi(P) \subset Q \quad \text{avec } Q = [m, M] \times P', \quad m = \min_P u, \quad M = \max_P u, \quad P' = \prod_{i \geq 2} [a_i, b_i].$$

(14) Bien entendu, il pourra éventuellement être judicieux, dans un premier temps, de ne donner la preuve qu'en dimension 2, ce qui permet de simplifier les notations. Curieusement, le fait que l'on puisse ramener la preuve de la formule du jacobien à la dimension 1 par des manipulations purement algébriques semble ne pas être « passé dans les mœurs », de sorte que les livres traitant de l'intégration développent souvent des démonstrations analytiques nettement plus élaborées. Celles-ci consistent en général à faire des approximations linéaires uniformes modulo des versions effectives adéquates du théorème d'inversion locale, en énonçant d'abord la formule du déterminant pour le cas d'un changement linéaire. Ceci apparaîtra en fait plutôt comme une conséquence du résultat général dans la présente approche, qui contient le cas linéaire comme cas très particulier (cf. remarque 2.6).

Il revient au même d'intégrer f sur l'ouvert image $\varphi(P^\circ)$ ou sur le pavé Q , puisque f est nulle par hypothèse sur $Q \setminus \varphi(P^\circ)$. On obtient par conséquent

$$(2.6 \text{ a}) \quad \begin{aligned} \int_{\varphi(P^\circ)} f(y) dy &= \int_Q f(y) dy = \int_{P'} \left(\int_m^M f(y_1, y') dy_1 \right) dy' \\ &= \int_{P'} \left(\int_m^M f(y_1, x') dy_1 \right) dx' \end{aligned}$$

à l'aide du théorème de Fubini et du changement de variable trivial $y' = x'$. Effectuons dans l'intégrale en dy_1 le changement $x_1 \mapsto y_1 = u(x) = u(x_1, x')$. Comme la fonction $y_1 \mapsto f(y_1, x')$ est nulle en dehors de l'intervalle image $[u(a_1, x'), u(b_1, x')]$, il vient par le théorème de changement de variable en dimension 1

$$\begin{aligned} \int_m^M f(y_1, x') dy_1 &= \varepsilon \int_{u(a_1, x')}^{u(b_1, x')} f(y_1, x') dy_1 = \varepsilon \int_{a_1}^{b_1} f(u(x), x') \frac{\partial u}{\partial x_1}(x) dx_1 \\ &= \int_{a_1}^{b_1} f(u(x), x') \left| \frac{\partial u}{\partial x_1}(x) \right| dx_1, \end{aligned}$$

avec $\varepsilon = +1$ si u est croissante en x_1 et $\varepsilon = -1$ si u est décroissante en x_1 . On a donc

$$(2.6 \text{ b}) \quad \begin{aligned} \int_{P'} \left(\int_m^M f(y_1, x') dy_1 \right) dx' &= \int_{P'} \left(\int_{a_1}^{b_1} f(u(x), x') \left| \frac{\partial u}{\partial x_1}(x) \right| dx_1 \right) dx' \\ &= \int_P f(u(x), x') \left| \frac{\partial u}{\partial x_1}(x) \right| dx \end{aligned}$$

en appliquant de nouveau le théorème de Fubini. Compte tenu du fait évident que $\varphi : x \mapsto (u(x), x_2, \dots, x_n)$ a pour jacobien $\text{Jac } \varphi = \partial u / \partial x_1$, la combinaison de (2.6 a) et (2.6 b) donne bien l'égalité attendue

$$\int_{\varphi(P^\circ)} f(y) dy = \int_P f(u(x), x') \left| \frac{\partial u}{\partial x_1}(x) \right| dx = \int_{P^\circ} f(\varphi(x)) |\text{Jac } \varphi(x)| dx.$$

Deuxième étape. On suppose ici que le changement de variable $x \mapsto y = \varphi(x)$ est de la forme

$$(2.6 \text{ c}) \quad x = (x_1, \dots, x_n) \xrightarrow{\varphi} y = (\varphi_1(x), \dots, \varphi_r(x), x_{r+1}, \dots, x_n)$$

pour un certain $r \geq 1$ (c'est-à-dire que $\varphi_i(x) = x_i$ pour $i \geq r+1$; bien entendu, pour $r = n$, il s'agit d'un changement de variable complètement général). On montre dans ce cas par récurrence sur r que φ est, localement au voisinage de tout point $a \in \Omega$, de la forme $\varphi = \sigma \circ \psi_1 \circ \dots \circ \psi_r$ où σ est une permutation des coordonnées (x_1, \dots, x_n) qui permute (x_1, \dots, x_r) et laisse invariants (x_{r+1}, \dots, x_n) , et où ψ_i est un changement de variable local agissant sur la seule variable x_i :

$$(2.6 \text{ d}) \quad x = (x_1, \dots, x_n) \xrightarrow{\psi_i} y = (x_1, \dots, x_{i-1}, u_i(x), x_{i+1}, \dots, x_n).$$

(ceci revient à dire qu'un changement de variable tel que (2.6 c) s'obtient en changeant les coordonnées x_1, \dots, x_r une par une, successivement). Si $r = 1$, le résultat est évident avec $\sigma = \text{Id}$, $u_1 = \varphi_1$, $\varphi = \psi_1 = \sigma \circ \psi_1$. Supposons maintenant $r \geq 2$ et soit $a \in \Omega$. Comme

$$\text{Jac } \varphi(a) = \det \left(\frac{\partial \varphi_i}{\partial x_j}(a) \right)_{1 \leq i, j \leq r} \neq 0,$$

on voit que l'une des dérivées $\partial \varphi_i / \partial x_r(a)$ est non nulle. Quitte à remplacer φ par $\varphi^\theta = \theta \circ \varphi$ avec une permutation θ adéquate des coordonnées (y_1, \dots, y_r) , on peut supposer que $\partial \varphi_r^\theta / \partial x_r(a) \neq 0$. On considère alors le changement de variable ψ_r défini par (2.6 d) avec $u_r(x) = \varphi_r^\theta(x)$. Comme le jacobien de ψ_r en a est égal à $\partial \varphi_r^\theta / \partial x_r(a) \neq 0$, le théorème d'inversion locale montre que ψ_r est bien un changement de variable inversible au voisinage de a . De plus, $\tilde{\varphi} = \varphi^\theta \circ \psi_r^{-1}$ est de la forme (2.6 c) à l'ordre $r - 1$ au voisinage de $b = \psi_r(a)$, puisque si $y = \psi_r(x)$ alors $z = \tilde{\varphi}(y) = \varphi^\theta(\psi_r^{-1}(y)) = \varphi^\theta(x)$ est tel que $z_i = x_i = y_i$ pour $i \geq r + 1$ tandis que $z_r = \varphi_r^\theta(x) = y_r$. On peut donc écrire par hypothèse de récurrence

$$\tilde{\varphi} = \tilde{\sigma} \circ \tilde{\psi}_1 \circ \dots \circ \tilde{\psi}_{r-1}$$

au voisinage de b , ce qui entraîne que

$$\varphi = \theta^{-1} \circ \varphi^\theta = \theta^{-1} \circ \tilde{\varphi} \circ \psi_r = (\theta^{-1} \circ \tilde{\sigma}) \circ \tilde{\psi}_1 \circ \dots \circ \tilde{\psi}_{r-1} \circ \psi_r$$

est de la forme voulue au voisinage de a .

Troisième étape. On montre ici que si la formule est vraie pour les changements de variables φ et ψ , alors elle est encore vraie pour le changement de variable composé

$$\psi \circ \varphi : x \mapsto y = \varphi(x) \mapsto z = \psi(y),$$

chaque fois que $\varphi : \Omega \rightarrow \Omega'$ et $\psi : \Omega' \rightarrow \Omega''$ sont des bijections de classe C^1 ayant des jacobiens non nuls. Sous cette hypothèse $\psi \circ \varphi : \Omega \rightarrow \Omega''$ a les mêmes propriétés, et comme $(\psi \circ \varphi)' = (\psi' \circ \varphi) \cdot \varphi'$, on a $\text{Jac}(\psi \circ \varphi) = ((\text{Jac } \psi) \circ \varphi) \times \text{Jac } \varphi$. En posant $g(y) = f(\psi(y)) |\text{Jac } \psi(y)|$, on trouve successivement

$$\int_{\psi(\varphi(\Omega))} f(z) dz = \int_{\varphi(\Omega)} g(y) dy = \int_{\Omega} g(\varphi(x)) |\text{Jac } \varphi(x)| dx,$$

et on a bien

$$g(\varphi(x)) |\text{Jac } \varphi(x)| = f(\psi(\varphi(x))) |\text{Jac } \psi(\varphi(x))| |\text{Jac } \varphi(x)| = f(\psi \circ \varphi(x)) |\text{Jac } \psi \circ \varphi(x)|.$$

Quatrième étape. On démontre ici la version locale de la formule, à savoir sa validité, pour tout point y_0 de Ω , lorsque f est à support compact dans un voisinage V assez petit de y_0 . Dans ce cas $f \circ \varphi$ est à support compact dans $U = \varphi^{-1}(V)$ qui est un voisinage de $x_0 = \varphi^{-1}(y_0)$. Lorsque U et V sont assez petits, on sait d'après la deuxième étape qu'on peut écrire φ sur U comme la composée $\varphi = \sigma \circ \psi_1 \circ \dots \circ \psi_n$ d'une permutation σ de coordonnées et de changements ψ_i portant sur une seule des variables. Puisque la formule de changement de variable est vraie trivialement pour

σ et pour chaque ψ_i (étape 1), elle est vraie pour φ d'après la troisième étape. Pour justifier cela de façon précise, on s'arrange pour trouver successivement des pavés P_0, P_1, \dots, P_n assez petits pour que ψ_i soit définie sur P_i , avec $\psi_i(P_i) \subset P_{i-1}$ pour tout i , et on prend $U = P_n^\circ$, $V = \varphi(U)$.

Cinquième étape. La formule est vraie en toute généralité, lorsque f est continue sur Ω' et d'intégrale absolument convergente. Pour cela, on utilise un découpage de Ω' en cubes assez petits, et la formule (2.5) qui donne

$$\int_{\Omega'} f(y) dy = \lim_{p \rightarrow +\infty} \sum_{j \in J_p} \int_{C_{j,p}} f_{j,p}(y) dy$$

avec $f_{j,p} = f\theta_{j,p}$ continue à support compact dans $C_{j,p}$. Lorsque le découpage est assez fin, la quatrième étape donne

$$\int_{C_{j,p}} f_{j,p}(y) dy = \int_{\Omega'} f_{j,p}(y) dy = \int_{\Omega} f_{j,p}(\varphi(x)) |\text{Jac } \varphi(x)| dx$$

Avec les notations de (2.5) et après sommation sur j il vient $\sum_{j \in J_p} f_{j,p} = f\chi_p$, d'où

$$\int_{\Omega'} f(y)\chi_p(y) dy = \int_{\Omega} f(\varphi(x))\chi_p(\varphi(x)) |\text{Jac } \varphi(x)| dx.$$

Comme les intégrales sont absolument convergentes et que $f = \lim_{p \rightarrow +\infty} f\chi_p$ avec $0 \leq \chi_p \leq 1$ et $\chi_p = 1$ sur K_p , nous obtenons après passage à la limite la formule de changement de variable cherchée (2.6). \square

(2.7) Remarque. Un cas très particulier de la formule de changement de variable est celui où la transformation $\varphi : x \mapsto Ax$ est linéaire, définie par une matrice inversible A . Dans ce cas, la formule de changement de variable s'écrit $y = Ax$, $dy = |\det A| dx$ et on trouve donc

$$(2.7 \text{ a}) \quad \boxed{\int_{A(\Omega)} f(y) dy = \int_{\Omega} f(Ax) |\det A| dx.}$$

La preuve est bien entendu un cas particulier de ce qui précède, mais elle peut aussi s'obtenir de manière directe comme ci-dessus en observant que A est un produit de matrices élémentaires de la forme

$$D_i(\lambda) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad A_{ij}(\mu) = \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & \mu & \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{pmatrix}.$$

En effet, celles-ci fournissent des changements de variables qui ne modifient qu'une seule des coordonnées à la fois :

$$D_i(\lambda) : x_i \mapsto x'_i = \lambda x_i, \quad A_{ij}(\mu) : x_i \mapsto x'_i = x_i + \mu x_j.$$

En appliquant la formule (2.7 a) à la fonction constante $f(x) = 1$, on obtient la formule usuelle de transformation du volume

$$(2.7 \text{ b}) \quad \boxed{\text{vol}(A(\Omega)) = |\det A| \text{vol}(\Omega).}$$

Dans le cas encore plus particulier où Ω est le cube unité, on voit que le volume du parallélogramme $\Pi_{a_1, \dots, a_n} = \{\sum x_i a_i, 0 \leq x_i \leq 1\}$ engendré par les vecteurs colonnes (a_1, \dots, a_n) de la matrice A est donné par

$$(2.7 \text{ c}) \quad \boxed{\text{vol}(\Pi_{a_1, \dots, a_n}) = |\det(a_1, \dots, a_n)|.}$$

La formule 2.7 (c) peut s'obtenir de manière élémentaire comme suit : d'une part les matrices $D_i(\lambda)$ dilatent les volumes dans le facteur $|\lambda| = |\det(D_i(\lambda))|$, d'autre part les « transvections » $A_{ij}(\mu)$ (qui sont de déterminant 1 et qui envoient les pavés rectangulaires sur des parallélépipèdes) ne modifient pas les volumes. Ceci peut se voir directement à partir de la définition de l'intégrale et de l'invariance du volume par translation, en observant le schéma usuel bien connu depuis l'enseignement primaire :

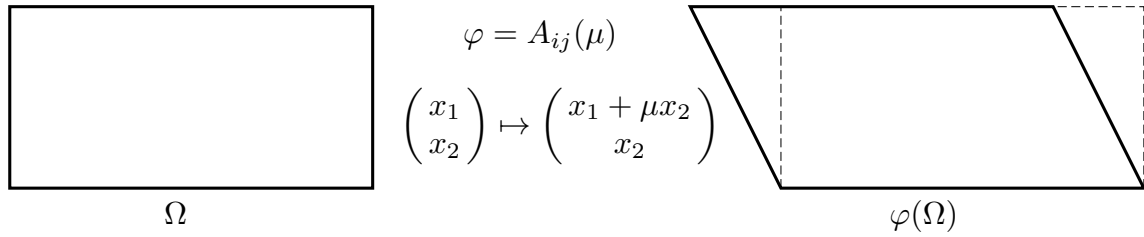


Fig. 18. Invariance de l'aire et du volume par translation.

(Le schéma figuré ici correspond au choix $i = 1, j = 2, \mu = -1/2$). ⁽¹⁵⁾

3. Intégration des formes différentielles et formule de Stokes

L'objet de ce paragraphe est d'introduire la théorie de l'intégration sur les courbes, surfaces ou plus généralement sur les sous-variétés quelconques tracées dans \mathbb{R}^n . L'outil fondamental est la notion de forme différentielle et le calcul différentiel « extérieur » qui lui est associé. Plutôt que de viser à la généralité la plus grande possible, nous essaierons d'introduire les notions fondamentales de manière concise, en nous concentrant sur les méthodes de calcul et les principales formules.

⁽¹⁵⁾ Une fois qu'on en est arrivé à ce point, c'est-à-dire en supposant acquise la formule (2.7 c) par une voie directe, voici l'autre approche possible de la preuve du théorème de changement de variable 2.6 (peut-être plus classique que celle que nous avons suivie) : elle consiste à utiliser une approximation du changement de variable φ par son application linéaire tangente, combinée à des estimations de la distorsion non linéaire relativement à des découpages en pavés suffisamment fins. Au total, la démonstration n'est pas nécessairement plus simple que celle « purement algébrique » que nous avons adoptée, puisque des estimations analytiques sont nécessaires.

Étant donné une fonction f de classe C^k sur un ouvert Ω de \mathbb{R} ($k \geq 1$), on lui associe sa différentielle $df = \sum_{1 \leq j \leq n} \frac{\partial f}{\partial x_j} dx_j$, qui est une application

$$(3.1) \quad df : \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x; h) \mapsto df(x; h) = \sum_{1 \leq j \leq n} \frac{\partial f}{\partial x_j} h_j,$$

de classe C^{k-1} par rapport à la variable $x = (x_1, \dots, x_n) \in \Omega$ et linéaire par rapport à $h = (h_1, \dots, h_n) \in \mathbb{R}^n$. En particulier la fonction coordonnée $f(x) = x_j$ a bien comme différentielle la forme linéaire $dx_j(h) = h_j$ (indépendante du point $x \in \mathbb{R}^n$, et (dx_1, \dots, dx_n) constitue une base de l'espace $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ des formes linéaires sur \mathbb{R}^n . De manière générale, on introduit la notion de forme différentielle de degré 1 comme suit.

(3.2) Définition. On appelle forme différentielle de degré 1 et de classe C^k sur l'ouvert $\Omega \subset \mathbb{R}^n$ toute application

$$\alpha : \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x; h) \mapsto \alpha(x; h)$$

de classe C^k en x et linéaire par rapport à la variable h . Une telle application peut donc s'écrire sous la forme

$$\alpha(x; h) = \sum_{1 \leq j \leq n} \alpha_j(x) h_j \quad \text{ou encore} \quad \alpha = \sum_{1 \leq j \leq n} \alpha_j(x) dx_j,$$

avec des coefficients $\alpha_j(x)$ qui sont des fonctions de classe C^k .

En dimension 1, une forme différentielle s'écrit $\alpha = f(x) dx$, et on lui associe son intégrale sur un intervalle $[a, b] \subset \mathbb{R}$ en posant simplement $\int_{[a, b]} \alpha = \int_a^b f(x) dx$. L'élément différentiel de longueur est donné quant à lui par

$$d\ell(h) = \|h\| = \sqrt{\sum_{1 \leq i \leq n} h_i^2} = \sqrt{\sum_{1 \leq i \leq n} dx_i(h)^2}, \quad \text{soit} \quad \boxed{d\ell = \sqrt{\sum_{1 \leq i \leq n} dx_i^2}}.$$

Il ne s'agit pas d'une forme différentielle puisque $d\ell(\lambda h) = |\lambda| d\ell(h)$ pour tout scalaire $\lambda \in \mathbb{R}$ et qu'on n'a donc pas linéarité en h .

(3.3) Intégrales curvilignes. Si $\alpha = \sum_{1 \leq j \leq n} \alpha_j(x) dx_j$ est une forme différentielle continue sur un ouvert $\Omega \subset \mathbb{R}^n$, et

$$g : [a, b] \rightarrow \Omega, \quad t \mapsto x = g(t) = (g_1(t), \dots, g_n(t))$$

un chemin de classe C^1 (ou C^1 par morceaux) tracé dans Ω , on peut considérer l'expression différentielle $g^* \alpha$ obtenue par la substitution $x_j = g_j(t)$ dans $\alpha(x)$, à savoir

$$g^* \alpha = \sum_{1 \leq j \leq n} \alpha_j(g(t)) g'_j(t) dt.$$

Si $[g]^+$ désigne le chemin g orienté dans le sens des paramètres croissants (avec $a < b$), on définit l'intégrale curviligne de α sur $[g]^+$ par

$$(3.3 \text{ a}) \quad \int_{[g]^+} \alpha = \int_a^b g^* \alpha = \int_a^b \sum_{1 \leq j \leq n} \alpha_j(g(t)) g'_j(t) dt$$

et la longueur du chemin g par

$$(3.3 \text{ b}) \quad \text{longueur}(g) = \int_{[g]^+} d\ell = \int_a^b \sqrt{\sum_{1 \leq j \leq n} g'_j(t)^2} dt.$$

Il résulte de la formule de changement de variable (3.5) que ces intégrales ne dépendent pas de la paramétrisation, c'est-à-dire qu'elles sont inchangées si on remplace $g : [a, b] \rightarrow \Omega$ par $g \circ \varphi : [a', b'] \rightarrow \Omega$ où $\varphi : [a', b'] \rightarrow [a, b]$ est un C^1 -difféomorphisme (on doit supposer en outre φ strictement croissant dans le cas de (3.3 a), si on ne veut pas avoir à modifier l'ordre croissant des bornes). Un cas particulier important est celui où la forme α est la différentielle $\alpha = df$ d'une fonction f de classe C^1 dans Ω . Dans ce cas $\alpha_j = \partial f / \partial x_j$ et

$$g^*(df) = \sum_{1 \leq j \leq n} \frac{\partial f}{\partial x_j}(g(t)) g'_j(t) dt = (f \circ g)'(t) dt,$$

d'où la formule fondamentale

$$(3.3 \text{ c}) \quad \text{si } g : [a, b] \rightarrow \Omega, \quad \int_{[g]^+} df = \int_a^b g^*(df) = f(g(b)) - f(g(a)),$$

et en particulier $\int_{[g]^+} df = 0$ si le chemin est fermé (i.e. $g(a) = g(b)$). □

(3.4) Éléments différentiels de dimension supérieure à 1 (aires, volumes, ...)

Comme le montre la formule (2.7 b), le déterminant, et aussi la valeur absolue du déterminant, jouent un rôle important en dimension quelconque. En particulier, si $u_1, \dots, u_p \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ sont des formes linéaires, on introduit l'application notée $u_1 \wedge \dots \wedge u_p : (\mathbb{R}^n)^p \rightarrow \mathbb{R}$, appelée *produit extérieur* de u_1, \dots, u_p , telle que

$$(3.4 \text{ a}) \quad u_1 \wedge \dots \wedge u_p(h_1, \dots, h_p) = \det(u_i(h_j))_{1 \leq i, j \leq p}.$$

C'est une forme p -linéaire alternée, c'est-à-dire linéaire en chaque h_j , $1 \leq j \leq p$, et telle que $u_1 \wedge \dots \wedge u_p(h_{\sigma(1)}, \dots, h_{\sigma(p)}) = \text{sign}(\sigma) u_1 \wedge \dots \wedge u_p(h_1, \dots, h_p)$, pour toute permutation $\sigma \in \mathfrak{S}_p$. Il est clair que $u_1 \wedge \dots \wedge u_p = 0$ si deux des formes u_i coïncident, ou plus généralement si les formes u_i sont linéairement dépendantes. On a

de plus $u_{\sigma(1)} \wedge \dots \wedge u_{\sigma(p)} = \text{sign}(\sigma) u_1 \wedge \dots \wedge u_p$ pour tout $\sigma \in \mathfrak{S}_p$. En particulier si $I = (i_1, \dots, i_p) \in \{1, \dots, n\}^p$ est un « multi-indice », on considère le produit extérieur

$$(3.4b) \quad \boxed{\begin{array}{l} dx_I = dx_{i_1} \wedge \dots \wedge dx_{i_p}, \quad \text{i.e.} \\ dx_I(h_1, \dots, h_p) = \det(dx_{i_k}(h_\ell))_{k,\ell} = \det(h_{i_k,\ell})_{k,\ell} \end{array}} \quad \text{où } h_\ell = \begin{pmatrix} h_{1,\ell} \\ \vdots \\ h_{n,\ell} \end{pmatrix},$$

pour tous $h_1, \dots, h_p \in \mathbb{R}^n$. On a $dx_I \neq 0$ si et seulement si les indices i_k sont 2 à 2 distincts. Dans le cas d'une paire $I = (i, j)$ on trouve par exemple

$$(3.4c) \quad dx_i \wedge dx_j(h_1, h_2) = h_{i,1}h_{j,2} - h_{i,2}h_{j,1}.$$

D'après la formule (2.7 c), la valeur absolue $|dx_{i_1} \wedge \dots \wedge dx_{i_p}(h_1, \dots, h_p)|$ représente l'aire p -dimensionnelle du parallélogramme de \mathbb{R}^p construit sur les projections des vecteurs $h_1, \dots, h_p \in \mathbb{R}^n$ dans l'espace des coordonnées $(x_{i_1}, \dots, x_{i_p}) \in \mathbb{R}^p$. Si α est une forme p -linéaire alternée et (e_1, \dots, e_n) la base canonique de \mathbb{R}^n , on pose

$$(3.4d) \quad \alpha_I = \alpha_{(i_1, \dots, i_p)} = \alpha(e_{i_1}, \dots, e_{i_p}) \quad \text{pour } I = (i_1, \dots, i_p) \in \{1, \dots, n\}^p,$$

ce qui permet d'écrire

$$\alpha(h_1, \dots, h_p) = \alpha\left(\sum_{i_1=1}^n h_{i_1,1}e_{i_1}, \dots, \sum_{i_p=1}^n h_{i_p,p}e_{i_p}\right) = \sum_{I \in \{1, \dots, n\}^p} \alpha_I h_{i_1,1} \dots h_{i_p,p}.$$

Comme $\alpha_I = 0$ si deux des indices i_k coïncident et $\alpha_{(i_{\sigma(1)}, \dots, i_{\sigma(p)})} = \text{sign}(\sigma)\alpha_{(i_1, \dots, i_p)}$ pour tout $\sigma \in \mathfrak{S}_p$, on peut regrouper la sommation en fonction des multi-indices croissants $i_1 < \dots < i_p$ de $\{1, \dots, n\}^n$, dont on note $\mathcal{C}_p(n)$ l'ensemble. On trouve alors

$$\alpha(h_1, \dots, h_p) = \sum_{I \in \mathcal{C}_p(n)} \alpha_I \sum_{\sigma \in \mathfrak{S}_n} \text{sign}(\sigma) h_{i_{\sigma(1)},1} \dots h_{i_{\sigma(p)},p} = \sum_{I \in \mathcal{C}_p(n)} \alpha_I dx_I(h_1, \dots, h_p).$$

On voit ainsi que $\alpha = \sum_{I \in \mathcal{C}_p(n)} \alpha_I dx_I$ et que la famille $(dx_I)_{I \in \mathcal{C}_p(n)}$ constitue une base des formes p -linéaires alternées sur \mathbb{R}^n .

(3.5) Définition. Soit Ω un ouvert de \mathbb{R}^n . On appelle p -forme différentielle (ou forme différentielle de degré p) et de classe C^k sur l'ouvert Ω toute application

$$\alpha : \Omega \times (\mathbb{R}^n)^p \rightarrow \mathbb{R}, \quad (x; h_1, \dots, h_p) \mapsto \alpha(x; h_1, \dots, h_p)$$

de classe C^k telle que $(h_1, \dots, h_p) \mapsto \alpha(x; h_1, \dots, h_p)$ soit une forme p -linéaire alternée pour tout $x \in \Omega$. On peut donc l'écrire en coordonnées

$$\boxed{\alpha = \sum_{I \in \mathcal{C}_p(n)} \alpha_I(x) dx_I,}$$

avec des coefficients $\alpha_I(x)$ qui sont des applications de classe C^k sur Ω , la somme étant prise sur l'ensemble $\mathcal{C}_p(n)$ des multi-indices croissants $I = (i_1 < \dots < i_p)$ de $\{1, \dots, n\}^n$.

Nous allons avoir besoin d'une notion légèrement différente (et d'une certaine façon plus générale). Cherchons en effet l'expression algébrique de l'aire euclidienne p -dimensionnelle du parallélogramme porté par des vecteurs $h_1, \dots, h_p \in \mathbb{R}^n$ linéairement indépendants. On sait d'après la formule (2.7 c) que cette aire est la valeur absolue du déterminant $|\det_{(a_1, \dots, a_p)}(h_1, \dots, h_p)|$ calculé dans une base orthonormée (a_1, \dots, a_p) du sous-espace vectoriel V engendré par h_1, \dots, h_p , ce déterminant étant bien indépendant au signe près du choix de la base orthonormée. C'est donc la valeur absolue du déterminant $\det(\Lambda)$ où $\Lambda = (\lambda_{ij})$ est la matrice $p \times p$ des coefficients tels que $h_j = \sum_i \lambda_{ij} a_i$. Si H et A sont les matrices $n \times p$ formées respectivement par les vecteurs colonnes (h_1, \dots, h_p) , (a_1, \dots, a_p) , nous avons $H = A\Lambda$ et donc le « déterminant de Gram » des vecteurs h_i vaut

$$\det(h_i \cdot h_j) = \det({}^t H H) = \det({}^t \Lambda {}^t A A \Lambda) = \det({}^t \Lambda \Lambda) = \det(\Lambda)^2$$

puisque ${}^t A A = (a_i \cdot a_j)$ est la matrice unité. L'aire cherchée est donc $(\det(h_i \cdot h_j))^{1/2}$, expression qui a le mérite de ne plus dépendre du choix de la base orthonormée (a_i) . Nous affirmons que

$$(3.6) \quad \det(h_i \cdot h_j) = \sum_{I \in \mathcal{C}_p(n)} (dx_I(h_1, \dots, h_p))^2$$

et plus généralement

$$(3.6') \quad \det(h_i \cdot h'_j) = \sum_{I \in \mathcal{C}_p(n)} dx_I(h_1, \dots, h_p) dx_I(h'_1, \dots, h'_p)$$

pour tous systèmes de p -uplets de vecteurs (h_i) , (h'_j) de \mathbb{R}^n . En effet, comme l'expression $\det(h_i \cdot h'_j)$ est p -linéaire alternée en (h_1, \dots, h_p) et en (h'_1, \dots, h'_p) , nous avons nécessairement une écriture de la forme

$$\det(h_i \cdot h'_j) = \sum_{I, J \in \mathcal{C}_p(n)} \beta_{I, J} dx_I(h_1, \dots, h_p) dx_J(h'_1, \dots, h'_p)$$

suivant la base (dx_I) , pour des coefficients $\beta_{I, J}$ adéquats. En prenant (h_i) et (h'_j) extraits de la base canonique (e_1, \dots, e_n) , on constate immédiatement que $\beta_{I, J} = 1$ si $I = J$ et $\beta_{I, J} = 0$ si $I \neq J$. Il résulte de la formule (3.6) que l'aire euclidienne p -dimensionnelle d'un p -uplet de vecteurs (h_1, \dots, h_p) est donnée par l'élément différentiel

$$(3.7) \quad dS = \sqrt{\sum_{I \in \mathcal{C}_p(n)} dx_I^2}.$$

En particulier, si $p = 1$, on retrouve l'élément de longueur $d\ell = \sqrt{\sum_{1 \leq i \leq n} dx_i^2}$. L'expression (3.7) ne définit plus une p -forme différentielle, mais ce qu'on appelle une p -densité différentielle :

(3.8) Définition. Soit Ω un ouvert de \mathbb{R}^n . On appelle p -densité différentielle de classe C^k sur l'ouvert Ω toute application

$$\alpha : \Omega \times (\mathbb{R}^n)^p \rightarrow \mathbb{R}, \quad (x; h_1, \dots, h_p) \mapsto \alpha(x; h_1, \dots, h_p)$$

définie et de classe C^k sur les p -uplets de vecteurs linéairement indépendants, vérifiant en outre la propriété

$$\alpha(x; \Lambda(h_1, \dots, h_p)) = |\det(\Lambda)| \alpha(x; h_1, \dots, h_p)$$

pour toute transformation par combinaison linéaire $(h_j) \mapsto (\sum_i \lambda_{ij} h_i)$ des vecteurs h_i (ce qui entraîne que $\alpha(x; h_1, \dots, h_p)$ est nulle sur des vecteurs (h_i) non linéairement indépendants).

Si $\alpha_1, \dots, \alpha_s$ sont des p -formes différentielles et f_1, \dots, f_s sont des fonctions, il est clair que

$$\alpha = \sum_{1 \leq j \leq s} f_j |\alpha_j| \quad \text{et} \quad \beta = \sqrt{\sum_{1 \leq j \leq s} \alpha_j^2}$$

sont des p -densités différentielles. Lorsque $p = n$, tout n -uplet (h_1, \dots, h_n) est l'image de la base canonique (e_1, \dots, e_n) par l'application linéaire Λ dont la matrice est fournie par les vecteurs colonnes h_j , et on voit que toute n -densité différentielle α est donnée par

$$\alpha(x; h_1, \dots, h_n) = |\det(\Lambda)| \alpha(x; e_1, \dots, e_n) = f(x) |\det(h_1, \dots, h_n)|$$

en posant $f(x) = \alpha(x; e_1, \dots, e_n)$, soit encore

$$(3.9) \quad \alpha(x) = f(x) |dx_1 \wedge \dots \wedge dx_n|.$$

Nous introduisons maintenant les opérateurs algèbro-différentiels fondamentaux pour les formes différentielles.

(3.10) Produit extérieur. Si $\alpha = \sum \alpha_I dx_I$ et $\beta = \sum \beta_J dx_J$ sont des formes différentielles de degrés respectifs p, q , on définit le produit extérieur $\alpha \wedge \beta$ comme la forme différentielle de degré $p + q$ telle que

$$\boxed{\alpha \wedge \beta = \sum_{I \in \mathcal{C}_p(n), J \in \mathcal{C}_q(n)} \alpha_I \beta_J dx_{IJ}.}$$

Autrement dit, on pose $dx_I \wedge dx_J = dx_{IJ}$, ce terme étant nul si I et J ont des indices en commun ; de plus IJ n'est pas toujours ordonné – même si I et J le sont – il faut donc éventuellement permuter et changer les signes des coefficients pour se retrouver dans la base ordonnée croissante habituelle.

On voit aussitôt que le produit extérieur est associatif et que l'on a

$$(3.11) \quad \beta \wedge \alpha = (-1)^{pq} \alpha \wedge \beta \quad \text{si } p = \deg \alpha, \quad q = \deg \beta,$$

en particulier $\beta \wedge \alpha = -\alpha \wedge \beta$ si α et β sont de degrés impairs.

(3.12) Changement de variable. Soit $\alpha(x) = \sum \alpha_I(x) dx_I$ une p -forme différentielle sur un ouvert $\Omega \subset \mathbb{R}$ et $\varphi : \Omega' \rightarrow \Omega$, $t \mapsto x = \varphi(t)$ un « changement de variable » de classe C^1 sur un ouvert Ω' de \mathbb{R}^m . On appelle image inverse de α par φ la p -forme différentielle sur Ω' notée $\beta = \varphi^* \alpha$, obtenue par substitution de $x = \varphi(t)$ dans l'expression de $\alpha(x)$:

$$(3.12 \text{ a}) \quad \varphi^* \alpha(t) = \sum \alpha_I(\varphi(t)) d\varphi_I(t) = \sum \alpha_I(\varphi(t)) d\varphi_{i_1}(t) \wedge \dots \wedge d\varphi_{i_p}(t).$$

De manière alternative, c'est la forme $\varphi^* \alpha : \Omega' \times (\mathbb{R}^m)^p \rightarrow \mathbb{R}$ telle que

$$(3.12 \text{ b}) \quad \varphi^* \alpha(t; h_1, \dots, h_p) = \alpha(\varphi(t); \varphi'(t)(h_1), \dots, \varphi'(t)(h_p)),$$

obtenue en remplaçant $t \mapsto \varphi(t)$ via φ et $h_i \mapsto \varphi'(t)(h_i)$ via l'application linéaire tangente à φ . Cette deuxième définition (3.12b) est encore valable pour une p -densité différentielle.

Si on fait une deuxième substitution $t = \psi(u)$ avec $\psi : \Omega'' \rightarrow \Omega'$, il est clair que ceci est équivalent à substituer directement $x = \varphi(t) = \varphi(\psi(u)) = \varphi \circ \psi(u)$, donc on a

$$(3.13) \quad \psi^*(\varphi^* \alpha) = (\varphi \circ \psi)^* \alpha.$$

Par ailleurs, l'opération de substitution est compatible avec le produit extérieur, c'est-à-dire que pour deux formes α, β sur Ω quelconques on a

$$(3.14) \quad \varphi^*(\alpha \wedge \beta) = \varphi^* \alpha \wedge \varphi^* \beta.$$

(3.15) Différentiation extérieure. Si $\alpha = \sum \alpha_I(x) dx_I$ est une p -forme différentielle de classe C^k , $k \geq 1$, on définit sa différentiation extérieure comme la $(p+1)$ -forme $d\alpha$ de classe C^{k-1} telle que

$$d\alpha(x) = \sum_{I \in \mathcal{C}_p(n)} d\alpha_I(x) \wedge dx_I = \sum_{I \in \mathcal{C}_p(n)} \sum_{\ell=1}^n \frac{\partial \alpha_I(x)}{\partial x_\ell} dx_\ell \wedge dx_I.$$

La première propriété essentielle est que

$$(3.16) \quad d^2 \alpha = d(d\alpha) = 0 \quad \text{pour toute forme } \alpha \text{ de classe } C^k, k \geq 2.$$

En effet, un calcul direct donne

$$\begin{aligned} d(d\alpha)(x) &= \sum_{I \in \mathcal{C}_p(n)} \sum_{1 \leq k, \ell \leq n} \frac{\partial^2 \alpha_I(x)}{\partial x_k \partial x_\ell} dx_k \wedge dx_\ell \wedge dx_I \\ &= \sum_{I \in \mathcal{C}_p(n)} \sum_{1 \leq k < \ell \leq n} \left(\frac{\partial^2 \alpha_I(x)}{\partial x_k \partial x_\ell} - \frac{\partial^2 \alpha_I(x)}{\partial x_\ell \partial x_k} \right) dx_k \wedge dx_\ell \wedge dx_I \end{aligned}$$

en échangeant les rôles de k et ℓ lorsque $k > \ell$ et en tenant compte du fait que $dx_\ell \wedge dx_k = -dx_k \wedge dx_\ell$, et $dx_k \wedge dx_k = 0$. Or le théorème de Schwarz implique que $\partial^2 \alpha_I(x) / \partial x_k \partial x_\ell = \partial^2 \alpha_I(x) / \partial x_\ell \partial x_k$ pour tous k, ℓ .

On a par ailleurs la *formule de Leibnitz* de différentiation d'un produit extérieur

$$(3.17) \quad d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^{\deg \alpha} \alpha \wedge d\beta$$

pour toutes formes α, β de classe C^1 . Ceci résulte immédiatement du fait que $\alpha \wedge \beta = \sum_{I,J} \alpha_I \beta_J dx_{IJ}$, et de la règle de Leibnitz usuelle $d(\alpha_I \beta_J) = \alpha_I d\beta_J + \beta_J d\alpha_I$ pour les fonctions, appliquée aux coefficients de $\alpha \wedge \beta$. Enfin, la différentiation extérieure *commute avec les changements de variable* :

$$(3.18) \quad d(\varphi^* \alpha) = \varphi^*(d\alpha) \quad \text{pour } \alpha, \varphi \text{ de classe } C^1.$$

Dans le cas d'une forme de degré 0, c'est-à-dire d'une fonction $\alpha = f$, ceci signifie que $d(f \circ \varphi) = \varphi^*(df)$, or

$$d(f \circ \varphi)(t; h) = (f \circ \varphi)'(t)(h) = f'(\varphi(t))(\varphi'(t)(h)) = (df)(\varphi(t); \varphi'(t)(h)) = \varphi^*(df)(t; h)$$

d'après la règle de différentiation des fonctions composées. Le cas général en résulte grâce à (3.14, 3.16, 3.17) : si $\alpha = \sum \alpha_I dx_I$, on a $d(d\varphi_I) = d(d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_p}) = 0$ par (3.16) et (3.17), tandis que (3.17) et (3.14) donnent

$$\begin{aligned} d(\varphi^* \alpha) &= d\left(\sum (\alpha_I \circ \varphi) d\varphi_I\right) = \sum d(\alpha_I \circ \varphi) \wedge d\varphi_I + (\alpha_I \circ \varphi) \wedge d(d\varphi_I) \\ &= \sum d(\alpha_I \circ \varphi) \wedge d\varphi_I = \sum \varphi^*(d\alpha_I) \wedge \varphi^*(dx_I) = \varphi^*(d\alpha). \end{aligned}$$

(3.19) Cas de la dimension 3. Nous allons voir qu'en dimension 3 le formalisme du calcul différentiel extérieur rejoint (en le généralisant et en le simplifiant considérablement), le formalisme usuel de la physique mathématique. Soit Ω un ouvert de \mathbb{R}^3 . Si $f : \Omega \rightarrow \mathbb{R}$ est une fonction de classe C^1 , on lui associe son gradient $\overrightarrow{\text{grad}} f$ qui est par définition le champ de vecteurs sur Ω tel que

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz = \overrightarrow{\text{grad}} f \cdot \overrightarrow{dM} \quad \text{où } \overrightarrow{\text{grad}} f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix}, \quad \overrightarrow{dM} = \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}.$$

Soit maintenant $\overrightarrow{V} : \Omega \rightarrow \mathbb{R}^3$, $M \mapsto \overrightarrow{V}(M)$ un champ de vecteurs de classe C^1 arbitraire sur Ω . On notera $M = (x, y, z)$ les coordonnées et $\overrightarrow{V} = (V_x, V_y, V_z)$ les composantes de \overrightarrow{V} dans la base canonique. On peut de deux manières différentes associer une forme différentielle à ce champ de vecteurs. La première manière est de considérer la 1-forme de « *circulation infinitésimale* » du champ de vecteurs \overrightarrow{V} :

$$\alpha = \overrightarrow{V}(M) \cdot \overrightarrow{dM} = V_x dx + V_y dy + V_z dz.$$

La deuxième est d'introduire *l'élément de surface vectoriel* \vec{dS} défini comme le produit vectoriel euclidien

$$\vec{dS}(h, k) = \vec{h} \wedge \vec{k} = \begin{pmatrix} h_y k_z - h_z k_y \\ h_z k_x - h_x k_z \\ h_x k_y - h_y k_x \end{pmatrix} = \begin{pmatrix} dy \wedge dz \\ dz \wedge dx \\ dx \wedge dy \end{pmatrix} (\vec{h}, \vec{k})$$

c'est-à-dire

$$\vec{dS} = \begin{pmatrix} dy \wedge dz \\ dz \wedge dx \\ dx \wedge dy \end{pmatrix}$$

(on notera que $dS = \|\vec{dS}\|$ est bien la densité d'aire euclidienne d'après (3.7)). On peut alors considérer la 2-forme de « *flux infinitésimal* » du champ de vecteurs \vec{V} :

$$\beta = \vec{V}(M) \cdot \vec{dS} = V_x dy \wedge dz + V_y dz \wedge dx + V_z dx \wedge dy.$$

Un calcul immédiat donne

$$d\alpha = \left(\frac{\partial V_y}{\partial z} - \frac{\partial V_z}{\partial y} \right) dy \wedge dz + \left(\frac{\partial V_z}{\partial x} - \frac{\partial V_x}{\partial z} \right) dz \wedge dx + \left(\frac{\partial V_x}{\partial y} - \frac{\partial V_y}{\partial x} \right) dx \wedge dy,$$

c'est-à-dire

$$d\alpha = \text{rot } \vec{V} \cdot \vec{dS} \quad \text{où} \quad \text{rot } \vec{V} = \begin{pmatrix} \frac{\partial V_y}{\partial z} - \frac{\partial V_z}{\partial y} \\ \frac{\partial V_z}{\partial x} - \frac{\partial V_x}{\partial z} \\ \frac{\partial V_x}{\partial y} - \frac{\partial V_y}{\partial x} \end{pmatrix}.$$

On a enfin la formule

$$d\beta = \left(\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z} \right) dx \wedge dy \wedge dz = \text{div } \vec{V} \, dx \wedge dy \wedge dz$$

où

$$\text{div } \vec{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}.$$

Ceci montre que le calcul de la différentielle extérieure $d\alpha$ pour des formes différentielles α de degré 0, 1, 2 données respectivement par f , $\vec{V} \cdot d\vec{M}$, $\vec{V} \cdot \vec{dS}$ correspond au calcul de $\text{grad } f$, $\text{rot } \vec{V}$, $\text{div } \vec{V}$.

(3.20) Intégration des densités différentielles sur une nappe paramétrée. Soit Ω un ouvert de \mathbb{R}^n . On suppose donnés un ouvert borné $U \subset \mathbb{R}^p$ et une application $g : \overline{U} \rightarrow \Omega$, $t \mapsto x = g(t)$ de classe C^k sur \overline{U} , $k \geq 1$.

$$g(\overline{U}) \subset \Omega \subset \mathbb{R}^n$$

$$U \subset \mathbb{R}^p$$

$$g$$

$$(t_1, t_2) \in [-\pi, \pi] \times [0, \pi] \mapsto \begin{cases} x_1 = (2 + \cos t_1) \cos(t_2) - 0.055/(t_1^2 + t_2^2 + 10^{-9}) \\ x_2 = (2 + \cos t_1) \sin t_2 \\ x_3 = \sin t_1. \end{cases}$$

Fig. 19. Un exemple de nappe paramétrée (non injective).

On dira qu'une telle application définit une *nappe paramétrée* de dimension p et de classe C^k dans l'ouvert Ω . Si $h : \overline{V} \rightarrow \Omega$, $\tau \mapsto h(\tau)$ est une autre nappe paramétrée, on dira que les nappes associées à g et h sont C^k -équivalentes s'il existe un C^k -difféomorphisme $\varphi : \overline{U} \rightarrow \overline{V}$, $t \mapsto \tau = \varphi(t)$ tel que $g = h \circ \varphi$. Il est aisé de vérifier qu'il s'agit bien d'une relation d'équivalence. De plus, si $g \sim h$, les images $g(\overline{U}) \subset \Omega$ et $h(\overline{V}) \subset \Omega$ coïncident. On notera $[g]$ la classe d'équivalence de g .

Maintenant, si $x \mapsto \alpha(x)$ est une p -densité différentielle à coefficients continus sur Ω , on définit

$$(3.20 \text{ a}) \quad \boxed{\int_{[g]} \alpha = \int_U g^* \alpha.}$$

Ici $g^* \alpha$ est le résultat de la substitution $x = g(t)$ dans $\alpha(x)$. On obtient ainsi d'après (3.9) une expression de la forme $g^* \alpha(t) = f(t_1, \dots, t_p) |dt_1 \wedge \dots \wedge dt_p|$, ce qui permet de poser par définition

$$(3.20 \text{ b}) \quad \boxed{\int_U g^* \alpha = \int_U f(t_1, \dots, t_p) |dt_1 \wedge \dots \wedge dt_p| = \int_U f(t_1, \dots, t_p) dt_1 \dots dt_p.}$$

Autrement dit, on confond dans les notations la p -densité $|dt_1 \wedge \dots \wedge dt_p|$ avec l'élément de volume $dt_1 \dots dt_p$ de \mathbb{R}^p , ce qui paraît naturel compte tenu de la formule (2.7 c). Un point tout à fait essentiel est que l'intégrale (3.20 b) est indépendante de la paramétrisation, c'est-à-dire que si $g = h \circ \varphi$ comme ci-dessus, alors on a bien $\int_{[g]} \alpha = \int_{[h]} \alpha$ comme la notation le laisse supposer (autrement dit l'intégrale ne dépend que de la classe d'équivalence $[g]$ de g). Pour le voir, on écrit $h^* \alpha(\tau) = F(\tau) |d\tau_1 \wedge \dots \wedge d\tau_p|$ et on applique le changement de variable $\tau = \varphi(t)$ pour obtenir

$$\begin{aligned} g^* \alpha &= (h \circ \varphi)^* \alpha = \varphi^* (h^* \alpha) = \varphi^* (F(\tau) |d\tau_1 \wedge \dots \wedge d\tau_p|) \\ &= F \circ \varphi(t) |Jac \varphi(t)| |dt_1 \wedge \dots \wedge dt_p|. \end{aligned}$$

Or la formule (2.6) implique

$$\int_U F \circ \varphi(t) |\text{Jac } \varphi(t)| dt_1 \wedge \dots \wedge dt_p = \int_V F(\tau) |d\tau_1 \wedge \dots \wedge d\tau_p|,$$

d'où l'égalité souhaitée $\int_U g^* \alpha = \int_V h^* \alpha$. Dans le cas de la p -densité d'aire euclidienne $dS = \sqrt{\sum dx_I^2}$, on obtient la formule

$$(3.20 \text{ c}) \quad \boxed{\text{aire}([g]) = \int_{[g]} dS = \int_U g^* dS = \int_U \sqrt{\sum_{I \in \mathcal{C}_p(n)} (\text{Jac } g_I(t))^2} dt_1 \dots dt_p}$$

où

$$\boxed{\text{Jac } g_I(t) = \det \left(\frac{\partial g_{i_k}}{\partial t_\ell} \right)_{1 \leq k, \ell \leq p}.$$

Les considérations qui précèdent montrent que l'aire d'une nappe paramétrée g ne dépend bien que de la classe d'équivalence $[g]$. Dans le cas d'une courbe $g : U \rightarrow \Omega$ (avec $p = 1$, $U \subset \mathbb{R}$), on retrouve l'expression de la longueur déjà donnée par (3.3 b)

$$(3.20 \text{ d}) \quad \text{longueur}([g]) = \int_U \sqrt{\sum_{1 \leq i \leq n} g'_i(t)^2} dt.$$

(3.21) Intégration des formes différentielles. Pour toute nappe paramétrée $g : \overline{U} \rightarrow \Omega$ de dimension p et de classe C^1 , et toute p -forme différentielle α continue sur Ω , on peut utiliser essentiellement la même approche et poser

$$\int_U g^* \alpha = \int_U f(t_1, \dots, t_p) dt_1 \wedge \dots \wedge dt_p = \int_U f(t_1, \dots, t_p) dt_1 \dots dt_p$$

après évaluation de la p -forme $g^* \alpha = f(t_1, \dots, t_p) dt_1 \wedge \dots \wedge dt_p$ en fonction des paramètres t_j . Cependant, l'effet d'un changement de variable $t = \varphi(\tau)$ fera intervenir $\text{Jac } \varphi$ (et non pas $|\text{Jac } \varphi|$) dans l'expression de $\varphi^*(g^* \alpha)$, de sorte que l'intégrale va changer de signe si $\text{Jac } \varphi < 0$. L'intégrale n'est donc la même que pour les classes d'équivalence $[g]^+$ de nappes paramétrées *orientées*, avec par définition $g \sim^+ h$ si $g = h \circ \varphi$ pour un certain difféomorphisme $\varphi : \overline{U} \rightarrow \overline{V}$ tel que $\text{Jac } \varphi > 0$ (on définit de même la classe d'orientation opposée $[g]^-$ comme la classe d'équivalence des $h = g \circ \psi$ avec $\text{Jac } \psi < 0$, classe qui bien sûr ne contient pas g elle-même). Il est licite de poser par définition

$$(3.21 \text{ a}) \quad \boxed{\int_{[g]^+} \alpha = \int_U g^* \alpha,}$$

et avec les notations précédentes on a

$$(3.21 \text{ b}) \quad \boxed{\int_{[g]^-} \alpha = - \int_{[g]^+} \alpha.}$$

Dans le cas des courbes (dimension $p = 1$), considérer une classe d'orientation donnée $[g]^+$ correspond à n'autoriser que les changements de variables strictement croissants. Si $g : \overline{U} \rightarrow \Omega$ est une courbe, $[g]^+$ sa classe d'équivalence orientée et $\alpha = \sum_{1 \leq j \leq n} \alpha_j(x) dx_j$ une 1-forme différentielle sur Ω , on retrouve l'*intégrale curviligne*

$$(3.21 \text{ c}) \quad \int_{[g]^+} \sum_{1 \leq j \leq n} \alpha_j(x) dx_j = \int_U g^* \alpha = \int_U \sum_{1 \leq j \leq n} \alpha_j(g(t)) g'_j(t) dt$$

(où le choix de l'orientation $[g]^+$ implique que l'on range les bornes de l'intervalle U dans l'ordre croissant).

Nous en arrivons maintenant au point culminant de ce paragraphe, à savoir la formule de Stokes pour des sous-variétés à bord orientées de dimension quelconque dans \mathbb{R}^n .

(3.22) Notions de sous-variété et de sous-variété à bord. Une partie M de \mathbb{R}^n est appelée sous-variété (sans singularités et sans bord) de dimension p et de classe C^k si tout point $x_0 \in M$ possède, après permutation éventuelle des coordonnées (laquelle permutation dépend a priori du point x_0), un voisinage parallélépipédique $P_{x_0} = \prod]a_j, b_j[$ tel que $M \cap P_{x_0}$ soit un graphe d'équation

$$(3.22 \text{ a}) \quad x_i = w_i(x_1, \dots, x_p), \quad p+1 \leq i \leq n,$$

pour des fonctions $w_i : P'_{x_0} = \prod_{j \leq p}]a_j, b_j[\rightarrow]a_i, b_i[$ de classe C^k , $p+1 \leq i \leq n$. [D'après le théorème des fonctions implicites, il est équivalent de supposer que $M \cap P_{x_0}$ peut être défini par $n-p$ équations $f_j(x) = 0$, $1 \leq j \leq n-p$, ayant des différentielles df_j linéairement indépendantes sur V .] On utilise le terme de « courbe » lorsque $p = 1$, « surface » lorsque $p = 2$ et « hypersurface » lorsque $p = n-1$. Pour obtenir une paramétrisation $g : U \rightarrow M \cap P_{x_0}$ de $M \cap P_{x_0}$ sur un ouvert $U \subset \mathbb{R}^p$, il suffit de choisir un C^k difféomorphisme arbitraire

$$\varphi = (\varphi_1, \dots, \varphi_p) : U \rightarrow P'_{x_0}, \quad t \mapsto (x_1, \dots, x_p) = \varphi(t)$$

et de définir $g : U \rightarrow P$ par

$$(3.22 \text{ b}) \quad g(t) = (\varphi_1(t), \dots, \varphi_p(t), w_{p+1}(\varphi(t)), \dots, w_n(\varphi(t))).$$

Le choix $\varphi = \text{Id}_{P'_{x_0}}$ sur $U = P'_{x_0}$, c'est-à-dire $x_i = \varphi_i(t_1, \dots, t_p) = t_i$, $1 \leq i \leq p$, est naturellement permis. On voit ainsi que toute sous-variété de dimension p peut se définir localement comme une nappe paramétrée g de dimension p telle que l'une des projections p -dimensionnelle $\varphi = (g_{i_1}, \dots, g_{i_p})$ soit un difféomorphisme local sur un ouvert U de \mathbb{R}^p . *Orienter la sous-variété M* consiste à choisir une famille de paramétrisations locales recouvrant la totalité de M et ayant tous « la même orientation », c'est-à-dire tels que les changements de paramétrisation soient tous à jacobien > 0 (ceci peut se faire par exemple en considérant des projections $x \mapsto (x_{i_1}, \dots, x_{i_p})$ et en permutant les coordonnées si nécessaire).

$$\begin{array}{ll} p = 2 & x_3 = w_3(x_1, x_2) \\ n = 3 & \end{array}$$

$$x_0 \in \partial M \qquad M$$

$$x_1 < w_1(x_2)$$

Fig. 20. Un exemple de sous-variété à bord $(M, \partial M)$.

Plus généralement (cf. Fig. 20), on appelle sous-variété à bord de dimension p et de classe C^k une paire de sous-variétés $(M, \partial M)$ de dimensions respectives p et $p - 1$ et de classe C^k , disjointes, avec $\partial M \subset \overline{M}$, de sorte que pour tout point $x_0 \in \partial M$, il existe, après permutations éventuelle des coordonnées, un voisinage parallélépipédique $P_{x_0} = \prod]a_j, b_j[$ tel que $M \cap P_{x_0}$ soit un graphe d'équation

$$(3.22 \text{ c}) \quad \begin{cases} x_1 < w_1(x_2, \dots, x_p), & \text{resp. } x_1 > w_1(x_2, \dots, x_p), \\ x_i = w_i(x_1, \dots, x_p), & p + 1 \leq i \leq n, \end{cases}$$

et $\partial M \cap P_{x_0}$ le graphe d'équations

$$(3.22 \text{ d}) \quad \begin{cases} x_1 = w_1(x_2, \dots, x_p), \\ x_i = w_i(x_1, \dots, x_p), & p + 1 \leq i \leq n, \end{cases}$$

pour des fonctions de classe C^k

$$\begin{aligned} w_1 &: \prod_{2 \leq j \leq p}]a_j, b_j[\rightarrow]a_1, b_1[, \\ w_i &: \prod_{1 \leq j \leq p}]a_j, b_j[\cap \{ \pm (x_1 - w_2(x_2, \dots, x_p)) \geq 0 \} \rightarrow]a_i, b_i[, \quad p + 1 \leq i \leq n. \end{aligned}$$

[On prendra garde au fait que dans cette notation, en général, ∂M n'est pas la frontière de M au sens topologique, puisque M est d'intérieur vide si $p < n$.]

Dans cette situation, une paramétrisation très commode de M consiste à utiliser les variables $t = (t_1, \dots, t_p)$ telles que

$$(3.22 \text{ e}) \quad \begin{cases} t_1 = \varepsilon(x_1 - w_1(x_2, \dots, x_p)) \\ t_2 = x_2 \\ \vdots \\ t_p = x_p \end{cases} \Leftrightarrow \varphi(t) = \begin{cases} x_1 = \varepsilon t_1 + w_1(t_2, \dots, t_p) \\ x_2 = t_2 \\ \vdots \\ x_p = t_p \end{cases}$$

où $\varepsilon = \pm 1$ est choisi de manière que l'orientation attribuée à M soit fournie par la paramétrisation φ [φ est bien un difféomorphisme de $\mathbb{R} \times \prod_{2 \leq j \leq p}]a_j, b_j[\subset \mathbb{R}^p$ sur lui-même, de jacobien constant ε], avec de plus

$$(3.22 \text{ f}) \quad x_j = w_j(x_1, \dots, x_p) = w_j(\varphi(t)), \quad p+1 \leq j \leq n.$$

Si (quitte à prendre η assez petit et à restreindre les intervalles $]a_j, b_j[$, $2 \leq j \leq p$) on pose $U' = \{0\} \times \prod_{2 \leq j \leq p}]a_j, b_j[$ et

$$U =]0, \eta[\times \prod_{2 \leq j \leq p}]a_j, b_j[, \quad \text{resp.} \quad U =]-\eta, 0[\times \prod_{2 \leq j \leq p}]a_j, b_j[,$$

ceci conduit à une paramétrisation de classe C^k

$$(3.22 \text{ g}) \quad \begin{aligned} g : \mathbb{R}^p \supset U \cup U' &\rightarrow (M \cup \partial M) \cap P_{x_0}, \\ t &\mapsto g(t) = (\varphi(t), w_{p+1}(\varphi(t)), \dots, w_n(\varphi(t))) \end{aligned}$$

telle que

$$(3.22 \text{ h}) \quad \begin{aligned} U &\subset]-\infty, 0[\times \mathbb{R}^{p-1} \quad \text{ou} \quad U \subset]0, +\infty[\times \mathbb{R}^{p-1}, \\ U' &\subset \overline{U} \cap \{0\} \times \mathbb{R}^{p-1}, \quad g(U) = M \cap P_{x_0}, \quad g(U') = \partial M \cap P_{x_0}. \end{aligned}$$

En effet, par construction, la partie $M \cap P_{x_0}$ va être définie par la condition $t_1 > 0$ ou $t_1 < 0$, et sa frontière $\partial M \cap P_{x_0}$ sera donnée par la condition $t_1 = 0$.

(3.23) Convention d'orientation du bord. Soit $(M, \partial M)$ une sous-variété à bord de dimension p dans \mathbb{R}^n , munie d'une orientation. Au voisinage d'un point x_0 de ∂M , on peut paramétriser M à l'aide d'une nappe $g : U \cup U' \rightarrow M \cup \partial M$ satisfaisant (3.22 g,h), et telle que M soit positivement orientée par (t_1, \dots, t_p) (on remplace au besoin t_1 par $-t_1$). Alors on convient d'orienter ∂M comme suit : si M est donnée par $\{t_1 < 0\}$, on oriente ∂M à l'aide des paramètres (t_2, \dots, t_p) , tandis que si M est donnée par $\{t_1 > 0\}$ on choisit l'orientation inverse de (t_2, \dots, t_p) [si $p \geq 2$, on peut ainsi choisir par exemple $(-t_2, t_3, \dots, t_p)$; si $p = 1$, le bord ∂M consiste en des points isolés, orienter ces points consiste juste à leur affecter le coefficient $+1$ ou -1 suivant que M est donnée par $t_1 < 0$ ou $t_1 > 0$, respectivement.]

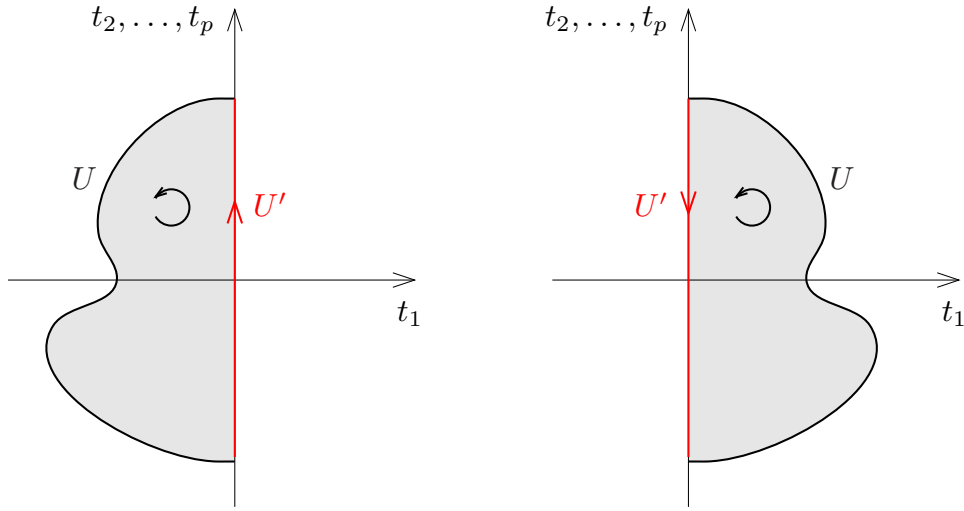


Fig. 21. Convention d'orientation du bord.

Il est facile de voir que l'orientation ainsi obtenue sur ∂M ne dépend que de celle de M et pas de la paramétrisation g choisie, sous réserve bien sûr que g satisfasse les hypothèses précédentes.

(3.24) Intégration sur les sous-variétés orientées. Soit Ω un ouvert de \mathbb{R}^n , α une p -forme différentielle sur Ω à coefficients continus, et $(M, \partial M)$ une sous-variété à bord orientée de dimension p et de classe C^k , $k \geq 1$, contenue dans Ω . On suppose que l'intersection $K = (M \cup \partial M) \cap \text{Supp } \alpha$ du support de α avec $M \cup \partial M$ est compact. Alors K peut être recouvert par un nombre fini de pavés $P_j \subset \Omega$, $1 \leq j \leq N$, dans lesquels on a des paramétrisations orientées $g_j : U_j \cup U'_j \rightarrow (M \cup \partial M) \cap P_j$ comme ci-dessus. Soit (θ_j) une partition de l'unité sur K supposée aux pavés P_j , c'est-à-dire une collection de fonctions θ_j continues à support compact dans P_j telles que $0 \leq \theta_j \leq 1$ et $\sum \theta_j = 1$ sur K (il suffit d'utiliser un quadrillage assez fin, et de considérer les fonctions $\theta_{j,p}$ définies par (2.4), qui sont en fait de classe C^1). Dans ces conditions on peut écrire $\alpha = \sum \alpha_j$ avec $\alpha_j = \theta_j \alpha$ à support dans P_j , et on pose

$$\int_M \alpha = \sum_j \int_{M \cap P_j} \alpha_j = \int_{U_j} g_j^* \alpha_j$$

conformément à la définition (3.21 a). L'invariance des intégrales par changement de variable orienté montre aisément que le résultat est en réalité indépendant du découpage et des paramétrisations orientées choisies (si on a une autre partition de l'unité θ'_k à support dans P'_k , on redécoupe suivant $\alpha_{jk} = \theta_j \theta'_k \alpha$ qui est à support dans $P_j \cap P'_k$ et on utilise la linéarité de l'intégrale...). Nous pouvons maintenant énoncer la

(3.25) Formule de Stokes. Soit Ω un ouvert de \mathbb{R}^n , α une $(p-1)$ -forme différentielle sur Ω de classe C^1 , et $(M, \partial M)$ une sous-variété à bord orientée de dimension p et de classe C^k , $k \geq 1$, contenue dans Ω , telle que $(M \cup \partial M) \cap \text{Supp } \alpha$ soit compact [le cas usuel est celui où la variété à bord $M \cup \partial M$ est elle-même compacte]. Alors

$$\int_M d\alpha = \int_{\partial M} \alpha,$$

à condition que les orientations de M et ∂M soient choisies l'une en fonction de l'autre comme il a été spécifié au paragraphe (3.23).

Démonstration. D'après la définition (3.24), il suffit de montrer que l'on a

$$\int_{M \cap P_j} d\alpha_j = \int_{\partial M \cap P_j} \alpha_j$$

pour tout j et de faire la somme (si les θ_j sont choisies de classe C^1 , ce qui est possible, alors $\alpha_j = \theta_j \alpha$ est aussi de classe C^1). Maintenant, on écrit que par définition

$$\int_{M \cap P_j} d\alpha_j = \int_{U_j} g_j^*(d\alpha_j) = \int_{U_j} d(g_j^* \alpha_j), \quad \int_{\partial M \cap P_j} \alpha_j = \int_{U'_j} g_j^* \alpha_j$$

où

$$g_j : U_j \cup U'_j \rightarrow (M \cup \partial M) \cap P_j, \quad t = (t_1, \dots, t_p) \mapsto g(t)$$

est une paramétrisation choisie comme dans 3.22 (e,f,g,h). En faisant les calculs des images inverses $\beta = g_j^* \alpha_j$ mises en jeu dans les coordonnées (t_1, \dots, t_p) , on est ramené à traiter le cas d'une forme β de degré $(p-1)$ en (t_1, \dots, t_p) , qui s'écrira donc

$$\beta(t) = \sum_{i=1}^p f_i(t) dt_1 \wedge \dots \wedge dt_{i-1} \wedge dt_{i+1} \wedge \dots \wedge dt_p.$$

En outre, quitte à translater l'origine, on peut supposer que β est à support compact dans un pavé $P = \prod_{1 \leq i \leq p}]-c_i, c_i[\subset \mathbb{R}^p$, et qu'on est dans l'un ou l'autre des deux cas suivants :

$$(3.25 \text{ a}) \quad M = P, \quad \partial M = \emptyset,$$

$$(3.25 \text{ b}) \quad M = P_- =]-c_1, 0[\times \prod_{2 \leq i \leq p}]-c_i, c_i[, \quad \partial M = \{0\} \times \prod_{2 \leq i \leq p}]-c_i, c_i[.$$

[Si $M = P_+ =]0, c_1[\times \prod_{2 \leq i \leq p}]-c_i, c_i[$, on se ramène aisément au cas $M = P_-$ en changeant t_1 en $t'_1 = -t_1$, ce qui a juste pour effet de changer simultanément l'orientation de M et de ∂M .]

Or on a

$$d\beta = \sum_{1 \leq i \leq p} (-1)^{i-1} \frac{\partial f_i}{\partial t_i} dt_1 \wedge \dots \wedge dt_p,$$

donc

$$\int_M d\beta = \int_M \sum_{1 \leq i \leq p} (-1)^{i-1} \frac{\partial f_i}{\partial t_i} dt_1 \dots dt_p.$$

Dans le cas (3.25 a), une intégration à une variable fournit pour tout $i = 1, \dots, p$ l'égalité

$$\begin{aligned} & \int_{-c_i}^{c_i} \frac{\partial f_i}{\partial t_i}(t_1, \dots, t_p) dt_i \\ &= f_i(t_1, \dots, t_{p-1}, c_i, t_{p+1}, \dots, t_p) - f_i(t_1, \dots, t_{p-1}, -c_i, t_{p+1}, \dots, t_p) = 0 \end{aligned}$$

du fait que f_i est à support compact dans P et donc nulle sur ∂P . Il vient par conséquent $\int_M \partial f_i / \partial t_i dt_1 \dots dt_p = 0$ grâce au théorème de Fubini, et comme $\partial M = \emptyset$ on a bien

$$\int_M d\beta = 0 = \int_{\partial M} \beta.$$

Dans le cas (3.25 b) on voit qu'on a encore $\int_M \partial f_i / \partial t_i dt_1 \dots dt_p = 0$ pour $i \geq 2$ par le même raisonnement que ci-dessus, tandis que

$$\int_{-c_1}^0 \frac{\partial f_1}{\partial t_1}(t_1, \dots, t_p) dt_1 = f_1(0, t_2, \dots, t_p),$$

puisque la valeur de $f_1(t_1, \dots, t_p)$ en $t_1 = -c_1$ est nulle [cette valeur n'est bien sûr pas nécessairement nulle dans le cas $t_1 = 0$, qui correspond à un point $t \in \partial M$]. D'après le théorème de Fubini, on obtient

$$\int_M d\beta = \int_{P_-} \frac{\partial f_1}{\partial t_1} dt_1 \dots dt_p = \int_{\Pi_{2 \leq i \leq p}[-c_i, c_i]} f_1(0, t_2, \dots, t_p) dt_2 \dots dt_p.$$

Or, le résultat de la substitution $t_1 = 0$ dans β donne la $(p-1)$ -forme

$$\beta|_{\{t_1=0\}} = f_1(0, t_2, \dots, t_p) dt_2 \wedge \dots \wedge dt_p,$$

et on a donc aussi

$$\int_{\partial M} \beta = \int_{\Pi_{2 \leq i \leq p}[-c_i, c_i]} f_1(0, t_2, \dots, t_p) dt_2 \dots dt_p.$$

Ceci conclut la démonstration. \square

(3.26) Remarque. La formule de Stokes est encore vraie si ∂M est de classe C^1 par morceaux, c'est-à-dire si ∂M est localement C^1 -difféomorphe à un domaine polyédral. Pour cela, il suffit d'observer que si S est l'ensemble des arêtes $(p-2)$ -dimensionnelles et S_ε l'ensemble des points situés à distance inférieure ou égale à ε de S , alors $\text{aire}_{p-1}(\partial M \cap S_\varepsilon) \leq C\varepsilon$ et $\text{aire}_p(M \cap S_\varepsilon) \leq C'\varepsilon^2$. On peut par conséquent se ramener au cas d'une forme α nulle au voisinage de S en considérant $\alpha_\varepsilon = (1 - \theta_\varepsilon)\alpha$ où θ_ε est égale à 1 au voisinage de S , à support dans S_ε et telle que $|d\theta_\varepsilon| \leq C''\frac{1}{\varepsilon}$. On voit alors aisément que $\int_{\partial M} \alpha_\varepsilon$ et $\int_M d\alpha_\varepsilon$ convergent vers les limites attendues quand ε tend vers 0.

(3.27) Cas particuliers de la formule de Stokes. Observons tout d'abord que dans le cas d'une courbe orientée ($\dim M = p = 1$), la formule de Stokes se réduit à une formule équivalente à (3.3 c), à savoir

$$(3.27 \text{ a}) \quad \int_{\partial M} df = \sum_{x \in \partial M} \varepsilon_x f(x),$$

avec $\varepsilon_x = +1$ si $x \in \partial M$ est une extrémité d'arc orienté et $\varepsilon_x = -1$ si $x \in \partial M$ est une origine d'arc orienté (M pouvant comprendre a priori plusieurs arcs orientés). Nous examinons maintenant le cas où $p \geq 2$, lorsque la dimension ambiante est petite.

Lorsque $p = n = 2$, le cas typique est celui d'un domaine borné $M \subset \mathbb{R}^2$ à bord de classe C^1 . Avec l'orientation canonique de \mathbb{R}^2 et l'orientation induite sur ∂M , on obtient la formule dite de *Green-Riemann*

$$(3.27 \text{ b}) \quad \int_{\partial M} P(x, y) dx + Q(x, y) dy = \iint_M \left(\frac{\partial Q}{\partial y} - \frac{\partial P}{\partial x} \right) dx dy,$$

valable pour toute forme $\alpha = P(x, y) dx + Q(x, y) dy$ de classe C^1 sur \overline{M} (on a bien $d\alpha = \left(\frac{\partial Q}{\partial y} - \frac{\partial P}{\partial x} \right) dx \wedge dy$).

Le cas suivant est $p = 2, n = 3$. On se donne une 1-forme $\alpha = \overrightarrow{V(x)} \cdot \overrightarrow{dx}$ de classe C^1 sur un ouvert Ω contenant une surface compacte orientée à bord $M \cup \partial M \subset \mathbb{R}^3$. Comme $d\alpha = \overrightarrow{\text{rot}} \overrightarrow{V} \cdot \overrightarrow{dS}$, on obtient la formule attribuée originellement à *Stokes*

$$(3.27 \text{ c}) \quad \int_{\partial M} \overrightarrow{V(x)} \cdot \overrightarrow{dx} = \iint_M \overrightarrow{\text{rot}} \overrightarrow{V} \cdot \overrightarrow{dS}.$$

Enfin, pour $p = n = 3$, il s'agit du cas d'un domaine à bord compact $M \cup \partial M \subset \mathbb{R}^3$. On considère dans ce cas une 2-forme $\alpha = \overrightarrow{V} \cdot \overrightarrow{dS}$ de classe C^1 sur \overline{M} et sa différentielle extérieure $d\alpha = \text{div } \overrightarrow{V} dx_1 dx_2 dx_3$. Ceci donne la formule dite de *Green-Ostrogradski*

$$(3.27 \text{ d}) \quad \iint_{\partial M} \overrightarrow{V(x)} \cdot \overrightarrow{dS} = \iiint_M \text{div } \overrightarrow{V} dx_1 dx_2 dx_3.$$

Chapitre III

Théorèmes généraux de convergence

Espaces L^p et fonctions mesurables

Nous établissons ici un pont direct entre la théorie de Kurzweil-Henstock et la théorie de la mesure. Ceci se fait en observant que l'intégrale de jauge satisfait les théorèmes de convergence fondamentaux que sont le théorème de convergence monotone et le théorème de convergence dominée. Ceci permet d'établir de manière naturelle l'existence de la mesure de Lebesgue dans \mathbb{R}^n . La substance de ce chapitre correspond à un (solide) niveau de L3, mais peut aussi se traiter en Master 1, par exemple comme introduction à la théorie générale de la mesure.

1. Lemme de Henstock et théorème de Hake

Le but du lemme de Henstock est d'obtenir des estimations fines pour les sommes de Riemann calculées sur des familles de sous-pavés d'un pavé P qui ne constituent pas nécessairement un découpage complet de P . Ces estimations ont elles-mêmes de nombreuses conséquences importantes.

(1.1) Lemme de Henstock. *Soit P un pavé fermé borné de \mathbb{R}^n et $f : P \rightarrow \mathbb{R}$ une fonction KH-intégrable sur P . Soient $\varepsilon > 0$ et δ une jauge ε -adaptée à f sur P . Soient enfin $(Q_i)_{1 \leq i \leq N}$ des sous-pavés de P d'intérieurs deux à deux disjoints, et $x_i \in Q_i$, $1 \leq i \leq N$, des points choisis dans ces pavés. Si ceux-ci sont δ -fins, c'est-à-dire si $\text{diam}(Q_i) \leq \delta(x_i)$, alors*

$$(a) \quad \left| \sum_{i=1}^N \text{vol}(Q_i) f(x_i) - \sum_{i=1}^N \int_{Q_i} f(x) dx \right| \leq \varepsilon ;$$

$$(b) \quad \sum_{i=1}^N \left| \text{vol}(Q_i) f(x_i) - \int_{Q_i} f(x) dx \right| \leq 2\varepsilon.$$

Démonstration. Soit $\eta > 0$ arbitrairement petit. Le complémentaire $P \setminus \bigcup Q_i^\circ$ peut se décomposer en la réunion d'un nombre fini de pavés fermés R_j d'intérieurs disjoints (ayant pour sommets des points de \mathbb{R}^n dont les coordonnées sont des projections de sommets de P ou des Q_i). Sur chaque R_j on peut trouver une jauge $\delta_j \leq \delta|_{R_j}$ telle que

$$\left| S_{D_j}(f) - \int_{R_j} f(x) dx \right| \leq \eta$$

pour toute subdivision pointée δ_j -fine D_j de R_j . En prenant la réunion des pavés pointés (Q_i, x_i) et des subdivisions D_j des pavés R_j , on obtient une subdivision pointée δ -fine de $[a, b]$, par conséquent

$$\left| \sum_i \text{vol}(Q_i) f(x_i) + \sum_j S_{D_j}(f) - \int_P f(x) dx \right| \leq \varepsilon.$$

En soustrayant toutes les inégalités précédentes et en tenant compte du fait que $\int_P f(x) dx = \sum_i \int_{Q_i} f(x) dx + \sum_j \int_{R_j} f(x) dx$, il vient

$$\left| \sum_i \text{vol}(Q_i) f(x_i) - \sum_i \int_{Q_i} f(x) dx \right| \leq \varepsilon + r\eta$$

où r est le nombre de pavés R_j mis en jeu. Comme $\eta > 0$ est arbitraire, l'inégalité (a) s'ensuit.

Pour obtenir (b), on applique séparément l'inégalité (a) aux pavés Q_i pour lesquels $\text{vol}(Q_i) f(x_i) - \int_{Q_i} f(x) dx \geq 0$ (resp. ≤ 0), et on fait la somme. \square

Nous commencerons par un corollaire élémentaire en dimension 1, puis nous démontrons son analogue, sensiblement plus subtil, en dimension supérieure.

(1.2) Théorème. *Pour toute fonction f KH-intégrable sur un intervalle $[a, b]$, l'intégrale indéfinie $x \mapsto \int_a^x f(t) dt$ est continue sur $[a, b]$.*

Démonstration. Il s'agit de prouver que $\forall x \in [a, b]$, $\int_x^{x+h} f(t) dt$ tend vers 0 avec h . Soit $\varepsilon > 0$ et δ une jauge ε -adaptée à f sur $[a, b]$. Prenons h tel que $|h| \leq \delta(x)$. En appliquant le lemme de Henstock 4.1 (a) à l'unique intervalle $[a_1, b_1] = [x, x+h]$ avec $x_1 = x \in [a_1, b_1]$, il vient

$$\left| h f(x) - \int_x^{x+h} f(t) dt \right| \leq \varepsilon,$$

donc

$$\left| \int_x^{x+h} f(t) dt \right| \leq \varepsilon + |h f(x)| \leq 2\varepsilon$$

pour $|h| \leq \min(\delta(x), \varepsilon/|f(x)|)$. Le corollaire est démontré. \square

(1.3) Théorème. *Soit $P' = \prod [a'_k, b'_k]$ est un pavé fermé contenu dans $P = \prod [a_k, b_k]$. Nous écrivons que P' tend vers $\tilde{P} = \prod [\tilde{a}_k, \tilde{b}_k] \subset P$ si chaque borne a'_k, b'_k de P' tend vers la borne correspondante \tilde{a}_k, \tilde{b}_k de \tilde{P} . Alors pour toute fonction f KH-intégrable sur P , on a*

$$(a) \quad \lim_{\text{vol}(P') \rightarrow 0} \int_{P'} f(x) dx = 0.$$

$$(b) \quad \lim_{P' \subset P, P' \rightarrow \tilde{P}} \int_{P'} f(x) dx = \int_{\tilde{P}} f(x) dx.$$

Démonstration. (a) Si (a) n'était pas vrai, on pourrait trouver une suite de pavés P_ν tels que $\text{vol}(P_\nu)$ tend vers 0 et $\left| \int_{P_\nu} f(x) dx \right| \geq \varepsilon_0 > 0$. Quitte à extraire une sous-suite

convergente de chacune des suites constituant les bornes de P_ν , on peut supposer que $P_\nu \rightarrow \tilde{P}$ pour un certain pavé \tilde{P} . On obtiendra une contradiction en montrant que $\int_{P_\nu} f(x) dx$ tend vers 0. Comme $\text{vol}(\tilde{P}) = 0$, le pavé \tilde{P} vérifie $\tilde{a}_{i_0} = \tilde{b}_{i_0}$ pour un certain indice i_0 . Il n'est pas restrictif de supposer que $P_\nu = \prod_i [a_{\nu,i}, b_{\nu,i}]$ contient sa limite $\tilde{P} = \prod_i [\tilde{a}_i, \tilde{b}_i]$. En effet, on a sinon $a_{\nu,i} > \tilde{a}_i$ ou $b_{\nu,i} < \tilde{b}_i$ pour un certain i . Si par exemple $a_{\nu,i} > \tilde{a}_i$, on écrit que

$$\int_{P_\nu} f(x) dx = \int_{P'_\nu} f(x) dx - \int_{P''_\nu} f(x) dx$$

ou P'_ν (resp. P''_ν) est la suite de pavés dont le i -ième facteur est $[\tilde{a}_i, b_{\nu,i}]$ (resp. $[\tilde{a}_i, a_{\nu,i}]$) qui convergent tous deux vers des pavés $\tilde{P}' = \tilde{P}$ et \tilde{P}'' de volume nul, avec cette fois les bonnes inégalités $a'_{\nu,i} \leq \tilde{a}'_i = \tilde{a}_i$, $a''_{\nu,i} \leq \tilde{a}''_i = \tilde{a}_i$ (qui sont en fait des égalités). Même raisonnement si $b_{\nu,i} < \tilde{b}_i$. En répétant la procédure pour chacun des axes de coordonnées et en utilisant la formule de Chasles, on se ramène ainsi par différences successives au cas où chaque pavé P_ν de la suite contient sa limite \tilde{P} .

On choisit maintenant δ une jauge ε -adaptée à f vérifiant la condition supplémentaire $\delta(x) \leq \frac{1}{2}d(x, \tilde{P})$ si $x \notin \tilde{P}$ et $\delta(x) \leq \varepsilon e^{-2n(|x|+\varepsilon)}/(1+|f(x)|)$ si $x \in \tilde{P}$. Pour toute subdivision pointée δ -fine $D = \{(Q_j, x_j)\}_{0 \leq j < N}$ de P , ceci force les pavés Q_j pointés par un point $x_j \notin \tilde{P}$ à être contenus dans le complémentaire de \tilde{P} , par conséquent la réunion $\bigcup_{x_j \in \tilde{P}} Q_j$ constitue un voisinage de \tilde{P} dans P , et on a donc $P_\nu \subset \bigcup_{x_j \in \tilde{P}} Q_j$ pour $\nu \geq \nu_\varepsilon$ assez grand. Il en résulte que la famille $\{(P_\nu \cap Q_j, x_j)_{x_j \in \tilde{P}}\}$ est une subdivision pointée δ -fine de P_ν pour $\nu \geq \nu_\varepsilon$. D'après le lemme de Henstock, nous en déduisons

$$\left| \sum_{x_j \in \tilde{P}} \text{vol}(Q_j) f(x_j) - \int_{P_\nu} f(x) dx \right| \leq \varepsilon.$$

Si \tilde{P} est tel que $\tilde{a}_{i_0} = \tilde{b}_{i_0}$, alors il est contenu dans l'hyperplan $\{x_{i_0} = \tilde{a}_{i_0}\}$ et le lemme 1.4 ci-dessous implique que

$$\sum_{x_j \in \tilde{P}} \text{vol}(Q_j) |f(x_j)| \leq 2\varepsilon.$$

On a donc $|\int_{P_\nu} f(x) dx| \leq 3\varepsilon$ pour $\nu \geq \nu_\varepsilon$. Comme $\varepsilon > 0$ est arbitraire, nous avons obtenu la contradiction désirée et la propriété (a) s'ensuit.

(b) se déduit immédiatement de (a) et de la formule de Chasles, puisque si $P' \rightarrow \tilde{P}$, les différences $P' \setminus \tilde{P}$ et $\tilde{P} \setminus P'$ s'expriment comme des réunions finies de pavés dont le volume tend vers 0. □

(1.4) Lemme. Soit $f : P \rightarrow \mathbb{R}$ une fonction définie sur un pavé fermé borné P de \mathbb{R}^n et $E \subset P$ une partie finie contenue dans un hyperplan de coordonnées $H = \{x_{i_0} = c\}$. On suppose que (Q_j, x_j) est une famille δ -fine de pavés pointés d'intérieurs disjoints, avec $Q_j \subset P$, $x_j \in E$, et $\delta(x_j) \leq \varepsilon e^{-2n(|x_j|+\varepsilon)}/(1+|f(x_j)|)$. Alors

$$\sum_{x_j \in E} \text{vol}(Q_j) |f(x_j)| \leq 2\varepsilon.$$

Démonstration. Supposons pour simplifier l'écriture que $i_0 = n$. On pose alors

$$Q_j = Q'_j \times [\alpha_j, \beta_j] \subset \mathbb{R}^{n-1} \times \mathbb{R}, \quad c \in [\alpha_j, \beta_j].$$

Quitte à remplacer $[\alpha_j, \beta_j]$ par $[\alpha_j, c]$ (resp. $[c, \beta_j]$) et à regrouper les pavés ainsi obtenus, on peut supposer que les pavés Q_j sont adjacents à l'hyperplan $\{x_n = c\}$ et situés dans le même demi-espace délimité par celui-ci (ceci donne alors éventuellement deux sommes à considérer, ce qui multiplie la constante finale par 2 au plus). Sous cette hypothèse, les pavés $Q'_j \subset \mathbb{R}^{n-1}$ sont eux-mêmes d'intérieurs disjoints. Comme

$$\beta_j - \alpha_j \leq \text{diam}(Q_j) \leq \delta(x_j) \leq \varepsilon e^{-2n(|x_j|+\varepsilon)} / (1 + |f(x_j)|),$$

nous avons

$$\text{vol}(Q_j) \leq \text{vol}(Q'_j) \varepsilon e^{-2n(|x_j|+\varepsilon)} / (1 + |f(x_j)|)$$

et par conséquent

$$\text{vol}(Q_j) |f(x_j)| \leq \varepsilon \text{vol}(Q'_j) e^{-2n(|x_j|+\varepsilon)}.$$

Or $\text{vol}(Q'_j) e^{-2n(|x_j|+\varepsilon)} \leq \int_{Q'_j} e^{-2n|x'|} dx'$ du fait que $\text{diam}(Q'_j) \leq \text{diam}(Q_j) \leq \varepsilon$, et puisque $n|x'| = n \max_{i \leq n-1} |x_i| \geq \sum_{i \leq n-1} |x_i|$ on voit que pour $M > 0$ assez grand

$$\begin{aligned} \sum_j \text{vol}(Q'_j) e^{-2n(|x_j|+\varepsilon)} &\leq \int_{[-M, M]^{n-1}} e^{-2n|x'|} dx' \leq \int_{[-M, M]^{n-1}} e^{-2 \sum |x_i|} dx' \\ &\leq \prod_{i \leq n-1} \int_{-M}^M e^{-2|\xi_i|} d\xi_i \leq 1 \end{aligned}$$

par le théorème de Fubini. Ceci conclut la preuve. \square

(1.5) Remarque. Le lemme 1.4 permet de voir également que l'intégrale $\int_P f(x) dx$ n'est pas modifiée si on change les valeurs de f sur une tranche hyperplane $P \cap \{x_{i_0} = c\}$ (ou une réunion finie de telles tranches), en particulier que cette intégrale ne dépend pas des valeurs prises par f sur le bord ∂P . Il suffit comme ci-dessus de prendre des jauges δ telles que $\delta(x) \leq \frac{1}{2}|x_{i_0} - c|$ si $x_{i_0} \neq c$ et $\delta(x) \leq \varepsilon e^{-2n(|x|+\varepsilon)} / (1 + |f(x)|)$ sur l'hyperplan $\{x_{i_0} = c\}$. \square

On se propose maintenant de donner la définition de l'intégrale de Kurzweil-Henstock sur un pavé quelconque P de \mathbb{R}^n (= produit d'intervalles quelconques, non nécessairement fermés ni bornés). Lorsque P est non compact, le principe consiste à introduire son compactifié, à savoir son adhérence \overline{P} dans le « compactifié d'Alexandrov » $\hat{\mathbb{R}}^n = \mathbb{R}^n \cup \{\infty\}$. Si $f : P \rightarrow \mathbb{R}$ est une fonction arbitraire, on l'étend à \overline{P} en posant $f(x) = 0$ pour $x \in \overline{P} \setminus P$ (donc en particulier $f(\infty) = 0$, si P est non borné); en fait, d'après le lemme 1.4 et la remarque 1.5, les valeurs prises par f sur ∂P ne joueront aucun rôle. Les fonctions jauge sont également supposées définies sur \overline{P} , valeur à l'infini $\delta(\infty)$ y comprise (là, il est important que δ soit définie sur un ensemble compact pour assurer l'existence de subdivisions pointées δ -fines).

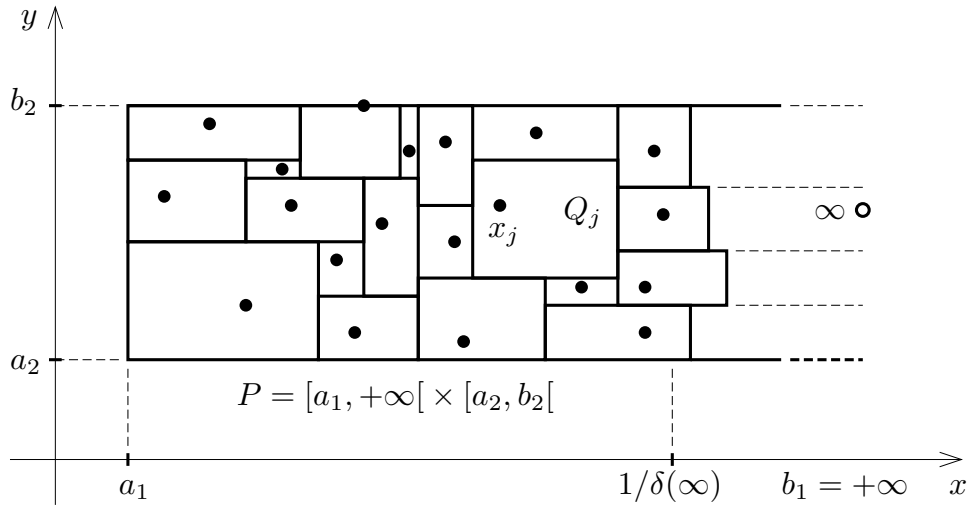


Fig. 22. Subdivision pointée d'un pavé P non borné.

(1.6) Définition. Une subdivision pointée $D = \{(Q_j, x_j)\}_{0 \leq j < N}$ de P est une famille finie de pavés Q_j fermés dans P , d'intérieurs disjoints, tels que $P = \bigcup Q_j$ (ou, de façon équivalente, tels que $\overline{P} = \bigcup \overline{Q_j}$), avec des points marqués $x_j \in \overline{Q_j}$. On dit que D est δ -fine si $\text{diam}(Q_j) \leq \delta(x_j)$ pour $x_j \neq \infty$, et $Q_j \subset V_{\infty, \delta} = \{x; \|x\| \geq 1/\delta(\infty)\}$ pour $x_j = \infty$ [Fig. 22].

Par définition, un pavé Q_j non borné ne peut être marqué que par le point $x_j = \infty$, et alors le terme $\text{vol}(Q_j)f(x_j)$ de la somme de Riemann doit être considéré comme nul (bien que le volume $\text{vol}(Q_j)$ soit infini). Ceci revient à dire que les pavés Q_j non bornés ne sont jamais pris en compte dans les sommes de Riemann considérées. L'existence de subdivisions δ -fines provient de la compacité de \overline{P} (ou de $\overline{P} \setminus V_{\infty, \delta}^\circ$, si on préfère). Avec ces conventions, la définition I.2.7 peut être reprise à l'identique.

(1.7) Définition. Une fonction $f : P \rightarrow \mathbb{R}$ définie sur un pavé quelconque $P \subset \mathbb{R}^n$ est dite intégrable au sens de Kurzweil-Henstock (KH-intégrable) s'il existe un réel A tel que pour toute erreur $\varepsilon > 0$ donnée a priori, on puisse trouver une jauge $\delta : \overline{P} \rightarrow \mathbb{R}_+^*$ telle que pour toute subdivision pointée $D = \{(Q_j, x_j)\}$ δ -fine de P on ait $|S_D(f) - A| \leq \varepsilon$. Dans ce cas, on note

$$A = \int_P f(x) dx,$$

et on appelle A l'intégrale de f sur P .

Il est clair, compte tenu du lemme 1.4 et de la remarque 1.5, que les valeurs prises par f sur ∂P ne jouent aucun rôle. La notion d'intégrabilité est inchangée si on passe d'un pavé P à un pavé $P' \subset P$ ayant les mêmes bornes (c'est-à-dire tel que $\overline{P'} = \overline{P}$). Les résultats suivants s'obtiennent avec des preuves rigoureusement identiques à celles que nous avons déjà données (cf. (II.1.6), (II.1.7)). Nous les énoncerons donc sans commentaires supplémentaires.

(1.8) Critère de Cauchy. Pour qu'une fonction $f : P \rightarrow \mathbb{R}$ soit KH-intégrable sur un pavé quelconque P , il faut et il suffit que pour tout $\varepsilon > 0$ on puisse trouver une

jauge δ sur P telle que pour toutes subdivisions δ -fines D, D' de P on ait

$$|S_{D'}(f) - S_D(f)| \leq \varepsilon$$

(et on peut se restreindre, comme on l'a déjà observé, au cas de subdivisions D et D' telles que D' est emboîtée dans D .)

(1.9) Proposition. Soit $f : P \rightarrow \mathbb{R}$ une fonction KH-intégrable sur un pavé P quelconque. Alors la restriction $f|_Q$ de f à tout pavé Q contenu dans P est encore KH-intégrable.

(1.10) Proposition (relation de Chasles). Étant donné un découpage $P = \bigcup P_i$ d'un pavé P en pavés P_i d'intérieurs disjoints, une fonction $f : P \rightarrow \mathbb{R}$ est KH-intégrable sur P si et seulement si elle est intégrable sur chaque pavé P_i , et on a alors

$$\int_P f(x) dx = \sum_i \int_{P_i} f(x) dx.$$

(1.11) Lemme de Henstock. Le lemme de Henstock (1.1) est valide pour un pavé P quelconque.

Notons maintenant la caractérisation simple suivante, qui montre que calculer des intégrales KH sur des pavés P non compacts est la même chose que de calculer ce que l'on appelle parfois des « intégrales impropres », à savoir des limites d'intégrales prises sur des parties compactes adéquates de plus en plus grandes.⁽¹⁶⁾

On considère dans le pavé P les parties R compactes, pavables et bien remplies au sens suivant : une telle partie R est une réunion finie de pavés compacts $P_i \subset P$ d'intérieurs disjoints non vides, telle que le complémentaire $P \setminus R^\circ$ soit lui-même réunion finie de pavés Q_j fermés dans P et d'intérieurs disjoints, non compacts (c'est-à-dire non bornés ou adhérents à $\partial P \setminus P$). Si $\delta : \overline{P} \rightarrow \mathbb{R}_+^*$ est une jauge, est dit que la partie compacte $R \subset P$ est δ -remplie si on peut trouver des pavés marqués (Q_j, x_j) δ -fins, fermés dans P , non compacts et d'intérieurs disjoints, tels que $P \setminus R^\circ = \bigcup Q_j$.

Par exemple, en dimension 1, si $P = [a, b]$, les parties compactes pavables et bien remplies sont les intervalles $[a, \beta]$, $\beta \in]a, b[$, et cette partie est δ -remplie dès lors que $b - \beta \leq \delta(b)$: il suffit de marquer l'intervalle $Q = [\beta, b]$ par le point b pour le voir. En dimension plus grande, ces régions peuvent avoir une forme plus compliquée :

(16) Le mot « impropre » n'est utilisé que parce que le résultat correspondant au théorème 1.12 ci-dessous n'est vrai ni pour l'intégrale de Riemann ordinaire, ni même pour l'intégrale de Lebesgue. Là encore, l'intégrale de Kurzweil-Henstock se révèle être à la fois plus souple, les intégrales impropres deviennent des intégrales « normales » ! Bien entendu – et surtout si on se limite à la dimension 1 – il est possible d'alléger l'exposé de la théorie en prenant plutôt le théorème 1.12 ci-après comme définition de $\int_P f(x) dx$.

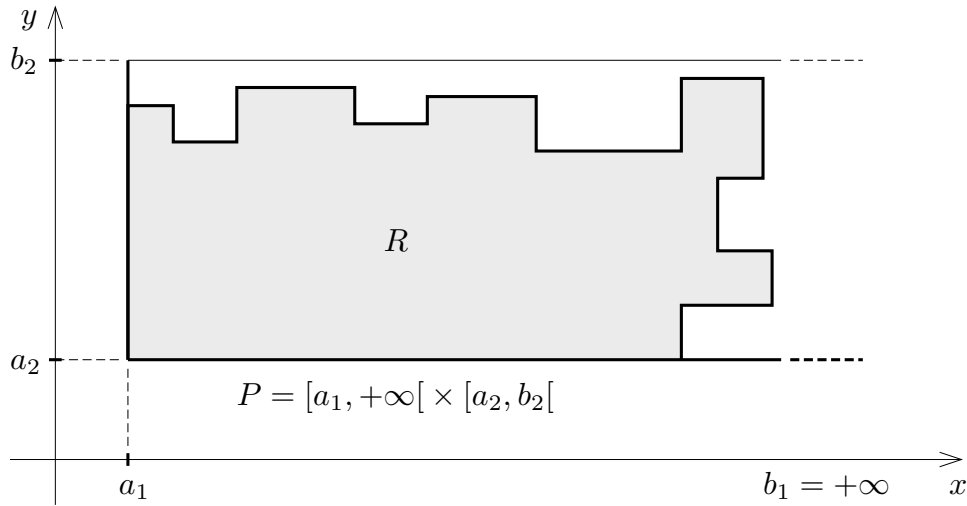


Fig. 23. Partie R compacte, pavable et bien remplie dans P .

Désignons par $\mathcal{R}(P)$ l'ensemble des parties compactes pavables et bien remplies dans P et soit $R \in \mathcal{R}(P)$ une partie δ -remplie. Il est commode de supposer que δ vérifie la propriété supplémentaire $\delta(x) \leq \frac{1}{2}d(x, F)$ pour tout face F de ∂P et tout $x \in \overline{P} \setminus F$ (de sorte qu'on a en particulier $\delta(x) \leq \frac{1}{2}d(x, \partial P)$ pour $x \in P^\circ$). En effet, sous cette hypothèse, tout pavé pointé (Q_j, x_j) tel que $x_j \in P^\circ$ est compact et contenu dans P° , donc pour toute subdivision pointée δ -fine $D = \{(Q_j, x_j)\}$ de P , les seuls pavés Q_j non compacts correspondent à des points que $x_j \in \partial P \cup \{\infty\}$. On voit même que x_j doit être sur une des faces F de $\partial P \setminus P$, car sinon l'hypothèse relative à la distance aux faces entraînerait de nouveau que Q_j serait compact. Si on suppose $\sup \delta \leq \frac{1}{k}$, on voit alors que la réunion R des pavés compacts $Q_j \subset P$ contient le pavé compact

$$P_k = \{x \in P; d(x, \partial P \setminus P) \geq 1/k, \|x\| \leq k\}$$

(c'est le pavé des points x dont les coordonnées x_j vérifient $|x_j| \leq k$ et dont la distance est au moins $1/k$ aux faces $F = \{x_i = c_i\}$ qui ne sont pas contenues dans P). On notera que (P_k) est une suite croissante de pavés compacts telle que $P = \bigcup P_k$.

(1.12) Théorème de Hake. Soit $f : P \rightarrow \mathbb{R}$ une fonction définie sur un pavé quelconque dans \mathbb{R}^n . Il y a équivalence entre

- (a) f est KH-intégrable sur P .
- (b) f est KH-intégrable sur tout pavé compact contenu dans P et la limite

$$\lim_{R \in \mathcal{R}(P), R \rightarrow P} \int_R f(x) dx \quad \text{existe}$$

(la limite étant prise suivant le filtre des parties R compactes pavables δ -remplies).

Dans ce cas, on a

$$\boxed{\int_P f(x) dx = \lim_{R \in \mathcal{R}(P), R \rightarrow P} \int_R f(x) dx}.$$

Démonstration. (a) \Rightarrow (b). Soit $\varepsilon > 0$ et δ une jauge ε -adaptée à f sur P . On supposera en outre que

$$\delta(x) \leq \begin{cases} \frac{\varepsilon e^{-2n|x|}}{1 + |f(x)|} & \text{lorsque } x \in \partial P, \\ \frac{1}{2}d(x, \partial P) & \text{lorsque } x \in P^\circ, \end{cases}$$

Si R est une partie compacte pavable δ -remplie, on peut par définition trouver une subdivision pointée δ -fine $D = \{(Q_j, x_j)\}_{0 \leq j < N}$ telle que $R = \bigcup Q_j$ soit la réunion des pavés Q_j compacts. Pour les Q_j non compacts, on a nécessairement $x_j \in \partial P \cup \{\infty\}$ comme on l'a déjà vu, et le terme associé $\text{vol}(Q_j)f(x_j)$ satisfait par suite l'estimation du lemme 1.5 (à moins que $x_j = \infty$, auquel cas ce terme est nul). Ceci donne

$$\left| S_D(f) - \sum_{Q_j \text{ compact}} \text{vol}(Q_j)f(x_j) \right| \leq 4n\varepsilon$$

(puisqu'il y a au plus $2n$ faces mises en jeu dans $\partial P \setminus P$). Le lemme de Henstock implique par ailleurs

$$\left| \sum_{Q_j \text{ compact}} \text{vol}(Q_j)f(x_j) - \int_R f(x) dx \right| \leq \varepsilon.$$

On obtient donc

$$\left| S_D(f) - \int_R f(x) dx \right| \leq (4n + 1)\varepsilon,$$

ce qui montre que (b) est vrai.

(b) \Rightarrow (a). Posons

$$A = \lim_{R \in \mathcal{R}(P), R \rightarrow P} \int_R f(x) dx.$$

Soit $\varepsilon > 0$ et $\hat{\delta} : \overline{P} \rightarrow \mathbb{R}_+^*$ une jauge telle que

$$\left| \int_R f(x) dx - A \right| \leq \varepsilon$$

pour toute partie compacte pavable $R \in \mathcal{R}(P)$ $\hat{\delta}$ -remplie. Fixons une suite croissante $R_0 \subset R_1 \subset R_2 \subset \dots$ de parties compactes pavables telles que $R_k \supset P_k$, $P = \bigcup R_k$ et

$$\left| \int_{R_k} f(x) dx - A \right| \leq 2^{-k}\varepsilon.$$

Pour tout $k \geq 0$, la différence $T_k = R_k \setminus R_{k-1}^\circ$ (avec $T_0 = R_0$) est pavable par des pavés $Q_{k,\ell}$ en nombre fini, et on peut trouver des jauges $\delta_{k,\ell}$ sur $Q_{k,\ell}$ telles que

$$\left| S_{D_k}(f) - \int_{T_k} f(x) dx \right| \leq 2^{-k}\varepsilon$$

pour toute subdivision pointée δ_k -fine D_k de T_k obtenue en réunissant des subdivisions des $Q_{k,\ell}$. On choisit maintenant une jauge δ sur P en prenant $\delta(\infty) = \hat{\delta}(\infty)$ et

$$\delta(x) = \begin{cases} \min \left(\hat{\delta}(x), \inf_{Q_{k,\ell} \ni x} \delta_{k,\ell}(x), \frac{1}{2} \inf_{F \ni x, F \subset \partial Q_{k,\ell}} d(x, F) \right) & \text{si } x \in P, \\ \hat{\delta}(x) & \text{si } x \in \partial P \setminus P, \end{cases}$$

où les F désignent les faces des pavés $Q_{k,\ell}$. Les inf sont bien strictement positifs du fait que la famille des $Q_{k,\ell}$ est localement finie dans P . Soit $D = \{(Q_j, x_j)\}$ une subdivision δ -fine de P . Les pavés Q_j non-compacts ont pour réunion une partie compacte R $\hat{\delta}$ -remplie, et on a donc $|\int_R f(x) dx - A| \leq \varepsilon$. Or, on a nécessairement $R \subset R_{k_0}$ pour un certain k_0 assez grand. Le choix de δ vis à vis de la distance aux faces des pavés $Q_{k,\ell}$ implique que D induit des subdivisions pointées partielles D_k δ_k -fines des T_k pour un nombre fini de parties T_k , $k \leq k_0$. Par le lemme de Henstock on en déduit

$$\left| S_{D_k}(f) - \int_{R \cap T_k} f(x) dx \right| \leq 2^{-k} \varepsilon,$$

et, en faisant la somme,

$$\left| S_D(f) - \int_R f(x) dx \right| \leq \sum 2^{-k} \varepsilon = 2\varepsilon.$$

Ceci entraîne $|S_D(f) - A| \leq 3\varepsilon$ et montre que f est KH-intégrable sur P , d'intégrale $\int_P f(x) dx = A$. \square

Bien entendu, comme l'intégrabilité de Kurzweil-Henstock sur P et sur P° sont équivalentes, on peut tout aussi bien approcher l'intégrale $\int_P f(x) dx$ par des intégrales $\int_R f(x) dx$ sur des parties compactes δ -remplies $R \subset P^\circ$, si on le souhaite. En dimension 1, on a le cas particulier suivant beaucoup plus simple du théorème de Hake.

(1.13) Théorème de Hake (dimension 1). *Soit $f : [a, b[\rightarrow \mathbb{R}$ une fonction définie sur un intervalle non compact, $b \in \mathbb{R} \cup \{+\infty\}$. Alors f est KH-intégrable sur $[a, b[$ si et seulement si elle est KH-intégrable sur tout intervalle $[a, \beta]$, $\beta \in [a, b[$, et si la limite $\lim_{\beta \rightarrow b-0} \int_a^\beta f(x) dx$ existe. Dans ce cas*

$$\int_a^b f(x) dx = \lim_{\beta \rightarrow b-0} \int_a^\beta f(x) dx. \quad \square$$

En dimension plus grande, une partie de ce résultat est encore valable.

(1.14) Proposition. *Si $f : P \rightarrow \mathbb{R}^n$ est une fonction KH-intégrable sur un pavé P quelconque, on a*

$$\int_P f(x) dx = \lim_{P' \subset P, P' \rightarrow P} \int_{P'} f(x) dx$$

où $P' \subset P$ est un pavé quelconque dont les bornes convergent vers celles de P .

Démonstration. Quitte à prolonger f par 0 sur $\overline{P} \setminus P$, on peut supposer P fermé dans \mathbb{R}^n . Il suffit alors de combiner le résultat du théorème de Hake appliqué à un pavé compact $P_1 \subset \overline{P}$ assez grand avec le lemme de Henstock et les estimations du lemme 1.4 sur le bord ∂P . Les détails sont laissés en exercice pour le lecteur. \square

On notera cependant qu'en dimension > 1 la réciproque de la proposition 1.14 dans laquelle on considérerait les seuls pavés compacts $P' \subset P$ δ -remplis n'est pas vraie. Par exemple, la fonction $f : P \rightarrow \mathbb{R}$ définie sur $P = [0, 1] \times \mathbb{R}$ telle que $f(x, y) = 1$ si $x \in [0, 1/2]$ et $f(x, y) = -1$ si $x \in]1/2, 1]$ est bien KH-intégrable d'intégrale nulle sur tout pavé $P' = [0, 1] \times [0, A]$, mais elle n'est pas KH-intégrable sur P (puisqu'à l'évidence elle ne l'est pas sur $Q = [0, 1/2] \times \mathbb{R}$).

2. Fonctions absolument intégrables

Étant donné une fonction f KH-intégrable sur un pavé P , il peut fort bien se produire que l'intégrale $\int_P |f(x)| dx$ soit divergente, en d'autres termes, que la fonction $|f|$ ne soit pas KH-intégrable. Un exemple classique est celui de l'intégrale $\int_0^{+\infty} \frac{\sin x}{x} dx$, ou encore $\int_0^1 \frac{1}{x} \sin \frac{1}{x} dx$. Ceci justifie la définition suivante.

(2.1) Définition. On dit qu'une fonction $f : P \rightarrow \mathbb{R}$ définie sur un pavé $P \subset \mathbb{R}^n$ est absolument intégrable sur P si à la fois f et $|f|$ sont KH-intégrables sur P , ou, de façon équivalente, si $f_+ = \max(f, 0)$ et $f_- = \max(-f, 0)$ sont KH-intégrables sur P .

L'équivalence des deux conditions résulte en effet des formules immédiates⁽¹⁷⁾

$$f_+ = \frac{1}{2}(f + |f|), \quad f_- = \frac{1}{2}(|f| - f), \quad f = f_+ - f_-, \quad |f| = f_+ + f_-.$$

Si f est absolument intégrable sur P , on a

$$-\int_P f_-(x) dx \leq \int_P f(x) dx = \int_P f_+(x) dx - \int_P f_-(x) dx \leq \int_P f_+(x) dx,$$

et comme les membres de droite et de gauche sont majorés par $\pm \int_P |f(x)| dx$ on en déduit

$$(2.2) \quad \left| \int_P f(x) dx \right| \leq \int_P |f(x)| dx.$$

Un critère commode pour l'intégrabilité de $|f|$ est le suivant.

(2.3) Critère d'intégrabilité absolue. Soit $f : P \rightarrow \mathbb{R}$ une fonction KH-intégrable. Alors on a

$$\int_P |f(x)| dx = \sup_D \sum_{0 \leq j < N} \left| \int_{Q_j} f(x) dx \right|$$

⁽¹⁷⁾ On remarquera qu'il ne suffit pas de supposer $|f|$ KH-intégrable dans cette définition, du moins dans la théorie des ensembles de Zermelo-Fraenkel augmentée de l'axiome du choix (ZFC - la théorie la plus couramment utilisée ...). Dans cette théorie, il existe en effet des parties $E \subset [0, 1]$ qui sont non mesurables, et la fonction $f = \chi_E - \chi_{[0,1] \setminus E}$ est de valeur absolue $|f| = 1$ KH-intégrable sur $[0, 1]$, tandis que f n'est pas KH-intégrable.

où le \sup est pris sur toutes les subdivisions pointées $D = (Q_j, x_j)_{0 \leq j < N}$ de P , la fonction $|f|$ étant KH-intégrable sur P si et seulement si le membre de droite est fini.

Démonstration. Si $|f|$ est KH-intégrable, on a d'après (2.2)

$$\left| \int_{Q_j} f(x) dx \right| \leq \int_{Q_j} |f(x)| dx,$$

donc on voit immédiatement que

$$\sup_D \sum_{0 \leq j < N} \left| \int_{Q_j} f(x) dx \right| \leq \int_P |f(x)| dx.$$

Dans l'autre sens, soit S le supremum figurant dans le membre de gauche, supposé fini. Étant donné $\varepsilon > 0$, soit $D = \{(Q_j, x_j)\}_{0 \leq j < N}$ une subdivision pointée de P qui réalise le supremum à ε près. On choisit une jauge δ adaptée à ε , telle qu'on ait en outre $\delta(x) \leq \min_{\partial Q_j \ni x} d(x, \partial Q_j)$. Toute subdivision $D' = \{(Q'_k, x'_k)\}_{0 \leq k < N'-1}$ δ -fine de P se décompose alors (quitte à redécouper si nécessaire les pavés Q'_k tels que $x'_k \in \partial Q_j$) en des subdivisions de chacun des pavés Q_j par ceux des pavés Q'_k qui sont contenus dans Q_j . Le lemme de Henstock 1.1 (b) combiné à l'inégalité triangulaire $||u| - |v|| \leq |u - v|$ implique

$$\left| \sum_{0 \leq k < N'-1} \left(\text{vol}(Q'_k) |f(x'_k)| - \left| \int_{Q'_k} f(x) dx \right| \right) \right| \leq 2\varepsilon.$$

Par ailleurs, la formule de Chasles donne que chaque intégrale $\int_{Q_j} f(x) dx$ est la somme des d'intégrales $\int_{Q'_k} f(x) dx$ telles que $Q'_k \subset Q_j$, donc

$$S - \varepsilon \leq \sum_{0 \leq j < N-1} \left| \int_{Q_j} f(x) dx \right| \leq \sum_{0 \leq k < N'-1} \left| \int_{Q'_k} f(x) dx \right| \leq S.$$

On en déduit $\left| \sum_k \text{vol}(Q'_k) |f(x'_k)| - S \right| \leq 3\varepsilon$, donc $|f|$ est bien KH-intégrable d'intégrale $\int_P |f(x)| dx = S$. \square

(2.4) Corollaire. Si $f, g : P \rightarrow \mathbb{R}$ sont KH-intégrables et si $|f| \leq g$, alors f est absolument intégrable sur P .

Démonstration. On applique le critère 2.3, en observant que

$$\sup_D \sum_{0 \leq j < N-1} \left| \int_{Q_j} f(x) dx \right| \leq \sup_D \sum_{0 \leq j < N-1} \int_{Q_j} g(x) dx \leq \int_P g(x) dx < +\infty. \quad \square$$

(2.5) Corollaire. L'ensemble $\tilde{L}^1(P)$ des fonctions absolument intégrables est un espace vectoriel, et l'intégrale de la valeur absolue

$$\|f\|_1 = \int_P |f(x)| dx$$

définit une semi-norme sur $\tilde{L}^1(P)$, c'est-à-dire que

$$\|\lambda f\|_1 = |\lambda| \|f\|_1, \quad \|f + g\|_1 \leq \|f\|_1 + \|g\|_1$$

pour tout scalaire $\lambda \in \mathbb{R}$ et tous $f, g \in \tilde{L}^1(P)$.

Démonstration. L'intégrabilité absolue de $f + g$ découle de l'inégalité triangulaire $|f + g| \leq |f| + |g|$ dans laquelle on sait que le membre de droite est KH-intégrable. Ceci implique également l'inégalité triangulaire pour la semi-norme $\|\cdot\|_1$. \square

(2.6) Remarque. $\|\cdot\|_1$ n'est pas une norme sur $\tilde{L}^1(P)$ car il existe des fonctions f non nulles telles que $\|f\|_1 = 0$. Il suffit par exemple qu'il existe un ensemble dénombrable E tel que $f(x)$ soit nul sur $P \setminus E$. On verra au paragraphe 5 comment on peut néanmoins construire un espace normé complet à partir de $\tilde{L}^1(P)$.

(2.7) Remarque. Si $f, g : P \rightarrow \mathbb{R}$ sont KH-intégrables, il n'est pas vrai en général que $\max(f, g)$ et $\min(f, g)$ sont intégrables (ce n'est même pas vrai si $g = 0$!). Cependant, c'est vrai si $f, g \geq 0$ ou si f, g sont absolument intégrables ; pour le voir il suffit en effet de poser

$$\max(f, g) = \frac{1}{2}(f + g) + \frac{1}{2}|f - g|, \quad \min(f, g) = \frac{1}{2}(f + g) - \frac{1}{2}|f - g|.$$

Plus généralement, c'est vrai s'il existe une fonction h KH-intégrable telle que $h \leq f$ et $h \leq g$, car on peut alors écrire

$$\min(f, g) = h + \min(f - h, g - h), \quad \max(f, g) = h + \max(f - h, g - h),$$

et de même s'il existe une fonction h KH-intégrable telle que $f \leq h$ et $g \leq h$, car on peut alors écrire

$$\min(f, g) = h - \max(h - f, h - g), \quad \max(f, g) = h - \min(h - f, h - g).$$

3. Le théorème de convergence monotone

L'un des points centraux de la théorie de l'intégration est de comprendre ce qui se passe pour la limite d'une suite d'intégrales de fonctions. Le cas le plus fondamental est celui d'une suite monotone de fonctions.

(3.1) Théorème. Soit $f_k : P \rightarrow \mathbb{R}$ une suite croissante de fonctions KH-intégrables sur un pavé $P \subset \mathbb{R}^n$, convergeant vers $f : P \rightarrow \mathbb{R}$ en tout point de P . Alors f est KH-intégrable sur P si et seulement si la suite croissante $A_k = \int_P f_k(x) dx$ est majorée, et alors

$$\int_P f(x) dx = \lim_{k \rightarrow +\infty} \int_P f_k(x) dx.$$

Démonstration. (a) Si f est KH-intégrable sur P , alors

$$A_k = \int_P f_k(x) dx \leq \int_P f(x) dx < +\infty,$$

par conséquent la suite croissante (A_k) est bornée et elle admet une limite

$$A = \lim_{k \rightarrow +\infty} \int_P f_k(x) dx \leq \int_P f(x) dx.$$

(b) Pour traiter la réciproque et l'inégalité inverse, commençons par le cas où P est fermé borné dans \mathbb{R}^n . Supposons $A = \lim A_k < +\infty$. On se propose alors de montrer que f est KH-intégrable sur P d'intégrale A . Fixons $\varepsilon > 0$. On peut choisir un entier k_0 tel que

$$A - \varepsilon \leq \int_P f_k(x) dx \leq A \quad \text{pour } k \geq k_0.$$

Pour chaque indice n , choisissons une jauge $\delta_k : P \rightarrow \mathbb{R}_+^*$ adaptée à f_k pour une tolérance d'erreur $2^{-k}\varepsilon$. Enfin, pour chaque $x \in P$, choisissons un indice $K(x) \geq k_0$ tel que

$$f(x) - \varepsilon \leq f_k(x) \leq f(x) \quad \text{pour } k \geq K(x).$$

On définit une jauge δ sur P par $\delta(x) = \delta_{K(x)}(x)$. Soit $D = \{(Q_i, x_i)\}_{0 \leq i < N}$ une subdivision pointée δ -fine de P . On peut écrire

$$\begin{aligned} |S_D(f) - A| &\leq \left| \sum_{0 \leq i < N} \text{vol}(Q_i) (f(x_i) - f_{K(x_i)}(x_i)) \right| \\ &\quad + \left| \sum_{0 \leq i < N} \left(\text{vol}(Q_i) f_{K(x_i)}(x_i) - \int_{Q_i} f_{K(x_i)}(x) dx \right) \right| \\ &\quad + \left| \sum_{0 \leq i < N} \int_{Q_i} f_{K(x_i)}(x) dx - A \right|. \end{aligned}$$

Par définition de $K(x)$, la première somme du membre de droite est majorée par $\varepsilon \sum \text{vol}(Q_i) = \varepsilon \text{vol}(P)$. Comme $K(x_i) \geq k_0$, on voit facilement que la troisième somme est majorée par

$$\left| \int_P f_{k_0}(x) dx - A \right| \leq \varepsilon.$$

En effet, grâce à la monotonie de la suite (f_k) , on a

$$\int_P f_p(x) dx \leq \sum_{0 \leq i < N} \int_{Q_i} f_{K(x_i)}(x) dx \leq \int_P f_q(x) dx$$

avec $p = \min(K(x_i))$, $q = \max(K(x_i))$ $q \geq p \geq k_0$. Reste la deuxième somme du membre de droite. Pour cela, on regroupe les indices j tels que $K(x_i)$ soit égal à un indice k donné. Le lemme de Henstock 1.1 (a) implique

$$\left| \sum_{K(x_i)=k} \left(f_k(x_i) \text{vol}(Q_i) - \int_{Q_i} f_k(x) dx \right) \right| \leq 2^{-k}\varepsilon.$$

En sommant sur toutes les valeurs de k , on voit que la deuxième somme est majorée par 2ε . Au total nous avons $|S_D(f) - A| \leq \varepsilon(\text{vol}(P) + 3)$, par conséquent f est KH-intégrable d'intégrale $A = \lim_{k \rightarrow +\infty} \int_P f_k(x) dx$. Il nous reste à traiter le cas d'un

pavé P non nécessairement borné. Quitte à remplacer f_k par $f_k - f_0$, on peut supposer $f_k \geq 0$ pour tout k (et donc $f \geq 0$). Dans ce cas, pour toute partie compacte pavable et bien remplie $R \subset P$, on a d'après ce qui précède et grâce à la positivité de f_k

$$\int_R f(x) dx = \lim_{k \rightarrow +\infty} \int_R f_k(x) dx \leq \lim_{k \rightarrow +\infty} \int_P f_k(x) dx.$$

Par conséquent, en faisant tendre R vers P et en utilisant le théorème de Hake 1.12, on voit que f est KH-intégrable sur P et que

$$\int_P f(x) dx = \lim_{R \in \mathcal{R}(P), R \rightarrow P} \int_R f(x) dx \leq \lim_{k \rightarrow +\infty} \int_P f_k(x) dx < +\infty.$$

Ceci termine la démonstration. □

(3.2) Remarque. On a bien entendu un résultat entièrement analogue pour les suites décroissantes de fonctions. On le déduit aussitôt en remplaçant (f_k) par $(-f_k)$.

4. Mesure de Lebesgue et ensembles négligeables

Nous commençons par la définition et les propriétés fondamentales de la mesure de Lebesgue dans \mathbb{R}^n .

(4.1) Théorème et définition. Si E est une partie de \mathbb{R}^n , on dit que E est intégrable si sa fonction caractéristique χ_E est KH-intégrable sur $P = \mathbb{R}^n$, et si c'est le cas, on définit la mesure de Lebesgue de E par

$$m(E) = \int_{\mathbb{R}^n} \chi_E(x) dx.$$

La mesure de Lebesgue jouit des propriétés suivantes :

- (a) Si E et F sont des parties intégrables, alors $E \cup F$, $E \cap F$, $E \setminus F$ sont intégrables. Si $E \subset F$, alors $m(E) \leq m(F)$.
- (b) Si (E_k) est une suite croissante de parties intégrables, $\bigcup E_k$ est intégrable si et seulement si la suite $m(E_k)$ est bornée, et alors $m(\bigcup E_k) = \lim m(E_k)$.
- (c) Si (E_k) est une suite décroissante de parties intégrables, alors $\bigcap E_k$ est intégrable et $m(\bigcap E_k) = \lim m(E_k)$.
- (d) Si $(E_k)_{k \in \mathbb{N}}$ est une famille de parties intégrables deux à deux disjointes, la réunion $\bigcup E_k$ est intégrable si et seulement si la série $\sum m(E_k)$ converge, et alors

$$m(\bigcup E_k) = \sum m(E_k),$$

Démonstration. (a) Il suffit d'appliquer la remarque 4.7, en observant que

$$\chi_{E \cup F} = \max(\chi_E, \chi_F), \quad \chi_{E \cap F} = \min(\chi_E, \chi_F), \quad \chi_{E \setminus F} = \chi_E - \chi_{E \cap F}.$$

Par ailleurs $E \subset F$ implique $\chi_E \leq \chi_F$.

(b) et (c) résultent du théorème de convergence monotone appliqué à la suite $f_k = \chi_{E_k}$, qui est croissante sous l'hypothèse (b) et décroissante sous l'hypothèse (c).

(d) On pose $F_k = E_0 \cup E_1 \cup \dots \cup E_k$. Alors F_k est intégrable d'après (a), et comme $\chi_{F_k} = \sum_{0 \leq i \leq k} \chi_{E_i}$ on a bien $m(F_k) = \sum_{0 \leq i \leq k} m(E_i)$. D'après (b) on obtient

$$m\left(\bigcup E_k\right) = \lim_{k \rightarrow +\infty} m(F_k) = \sum_{k=0}^{+\infty} m(E_k),$$

la suite $m(F_k)$ étant bornée si et seulement si la série converge. \square

(4.2) Définition. On dit

- (a) qu'une fonction $f : P \rightarrow \mathbb{R}$ est négligeable si $|f|$ est intégrable et $\int_P |f(x)| dx = 0$.
- (b) qu'une partie $E \subset \mathbb{R}^n$ est négligeable si sa fonction caractéristique χ_E est négligeable, autrement dit si $m(E) = 0$.

(4.3) Propriétés des fonctions et ensembles négligeables.

- (a) Si $f \geq 0$ est négligeable et si $|g| \leq f$, alors g est négligeable.
- (b) Si E est négligeable, toute partie $F \subset E$ est négligeable.
- (c) Toute réunion finie ou dénombrable $\bigcup_{k \geq 0} E_k$ de parties négligeables est négligeable.
- (d) Une fonction $f : P \rightarrow \mathbb{R}$ est négligeable si et seulement si $E_f = \{x \in P ; f(x) \neq 0\}$ est négligeable.

Démonstration. (a) résulte aussitôt par passage à la limite de la majoration des sommes de Riemann correspondantes $\sum \text{vol}(Q_i) |g(x_i)| \leq \sum \text{vol}(Q_i) f(x_i)$.

(b) Si $F \subset E$ on a $0 \leq \chi_F \leq \chi_E$, donc E négligeable $\Rightarrow F$ négligeable d'après (a).

(c) Posons $F_k = E_0 \cup E_1 \cup \dots \cup E_k$ et $F = \bigcup E_k$. Alors $0 \leq \chi_{F_k} \leq \chi_{E_0} + \chi_{E_1} + \dots + \chi_{E_k}$, donc χ_{F_k} est négligeable. Le théorème de convergence monotone montre que la limite croissante $\chi_F = \lim \chi_{F_k}$ est elle aussi négligeable.

(d) Supposons f négligeable, et soit $g_k(x) = \min(\chi_E(x), k|f(x)|)$. Alors (g_k) est une suite croissante de fonctions KH-intégrables telles que $0 \leq g_k \leq k|f|$ et $\lim g_k = \chi_E$. Elles vérifient donc $\int_P g_k(x) dx = 0$, et on obtient $\int_{\mathbb{R}^n} \chi_E(x) dx = \int_P \chi_E(x) dx = 0$ par le théorème de convergence monotone. Inversement si χ_E est négligeable, on considère $h_k(x) = \min(|f(x)|, k\chi_E(x))$ qui est une suite croissante de fonctions négligeables convergeant vers $|f|$ sur P , par suite f est négligeable. \square

Il résulte de ce qui précède que les ensembles et fonctions négligeables ne jouent aucun rôle dans la théorie de l'intégration. Par exemple, si $f, g : P \rightarrow \mathbb{R}$ diffèrent sur un ensemble négligeable, alors $f - g$ est négligeable et par suite l'intégrabilité de f équivaut à celle de g et dans ce cas $\int_P f(x) dx = \int_P g(x) dx$. De manière générale, on dit qu'une propriété $\mathcal{P}(x)$ dépendant d'un nombre réel x est vraie presque partout si l'ensemble E des x tels que $\mathcal{P}(x)$ ne soit pas vraie est négligeable. Ces résultats permettent de poser les définitions suivantes.

(4.4) Espace $L^1(P)$. L'ensemble $N(P)$ des fonctions négligeables est un sous-espace vectoriel de l'espace $\tilde{L}^1(P)$ des fonctions absolument intégrables. On note

$$L^1(P) = \tilde{L}^1(P)/N(P)$$

l'espace quotient. Les classes d'équivalence sont constituées de fonctions égales presque partout – ces classes seront encore notées comme s'il s'agissait de fonctions, en considérant qu'on s'autorise à changer éventuellement leurs valeurs sur un ensemble négligeable. Par définition de la semi-norme $\|\cdot\|_1$, nous avons $\|f\|_1 = 0$ si et seulement si $f \in N(P)$, c'est-à-dire $f = 0$ dans $L^1(P) = \tilde{L}^1(P)/N(P)$. Par conséquent $f \mapsto \|f\|_1$ définit une vraie norme sur $L^1(P)$. \square

Nous pouvons maintenant énoncer une version nettement renforcée du théorème de convergence monotone.

(4.5) Version forte du théorème de convergence monotone. Soit $f_k : P \rightarrow \mathbb{R}$ une suite de fonctions KH-intégrables. On suppose que $(f_k(x))$ est une suite croissante pour presque tout $x \in P$.

(a) Si $(f_k(x))$ converge presque partout vers une limite $f(x)$ où $f : P \rightarrow \mathbb{R}$ est une fonction KH-intégrable, alors

$$\lim_{k \rightarrow +\infty} \int_P f_k(x) dx = \int_P f(x) dx < +\infty.$$

(b) Inversement, si la limite $\lim_{k \rightarrow +\infty} \int_P f_k(x) dx$ est finie, alors l'ensemble E des réels $x \in P$ tels que $\lim_{k \rightarrow +\infty} f_k(x) = +\infty$ est négligeable. De plus, la fonction $f : P \rightarrow \mathbb{R}$ telle que

$$f(x) = \begin{cases} \lim_{k \rightarrow +\infty} f_k(x) & \text{si } x \notin E, \\ 0 & \text{si } x \in E \end{cases}$$

est KH-intégrable et satisfait (a).

Démonstration. (a) Soit E l'ensemble des points $x \in P$ tels que ou bien $(f_k(x))$ n'est pas une suite croissante, ou bien c'est une suite croissante mais $\lim f_k(x) = +\infty$, ou bien encore $f(x) \neq \lim f_k(x) < +\infty$. Par hypothèse, E est un ensemble négligeable (comme réunion finie d'ensembles négligeables). Quitte à redéfinir $f_k(x)$ et $f(x)$ comme étant égales à 0 en tout point de E , on a $f(x) = \lim_{k \rightarrow +\infty} f_k(x) < +\infty$ partout sur P et on peut alors appliquer le théorème de convergence monotone « ordinaire » 3.1.

(b) Quitte à redéfinir les f_k par 0 sur un ensemble négligeable, on peut supposer que la suite $(f_k(x))$ est croissante pour tout $x \in P$. En remplaçant f_k par $f_k - f_0$, on peut également supposer $f_k \geq 0$ pour tout k . Par hypothèse $0 \leq \int_P f_k(x) dx \leq M < +\infty$. Pour k et s entiers, $s > 0$, soit $E_{k,s}$ l'ensemble des $x \in P$ tels que $f_k(x) > s$. Nous avons $\chi_{E_{k,s}} = \lim_{N \rightarrow +\infty} \min(1, N(f_k - s)_+)$ comme limite croissante, et de plus $0 \leq \chi_{E_{k,s}} \leq \frac{1}{s} f_k$, donc $\chi_{E_{k,s}}$ est intégrable sur P et

$$\int_P \chi_{E_{k,s}}(x) dx \leq \frac{1}{s} \int_P f_k(x) dx \leq M/s.$$

Or $(E_{k,s})_{k \in \mathbb{N}}$ est une suite croissante d'ensembles dont la réunion est l'ensemble K_s des $x \in P$ tels que $\lim_{k \rightarrow +\infty} f_k(x) > s$. Ceci prouve que K_s est intégrable et que $m(K_s) \leq M/s$. Maintenant, l'ensemble E des réels $x \in P$ tels que $\lim_{k \rightarrow \infty} f_k(x) = +\infty$ est l'intersection décroissante $\bigcap_{s \geq 0} K_s$, par suite E est négligeable. Quitte à redéfinir $f_k(x) = 0$ sur E , nous pouvons appliquer le théorème de convergence monotone ordinaire pour conclure que $f = \lim f_k$ est KH-intégrable. \square

Désormais, on s'autorisera à écrire des intégrales dans lesquelles figurent des fonctions qui prennent les valeurs $+\infty$ ou $-\infty$, pourvu que cela soit sur un ensemble négligeable E – en fait, si on le souhaite, on pourra toujours redéfinir ces fonctions comme valant 0 ou toute autre valeur réelle en chaque point de E . Dans ce contexte, le théorème de convergence monotone peut se reformuler comme la possibilité de commuter intégration et passage à la limite monotone :

$$(4.6) \quad (f_k) \text{ monotone} \Rightarrow \int_P \lim_{k \rightarrow +\infty} f_k(x) dx = \lim_{k \rightarrow +\infty} \int_P f_k(x) dx,$$

dès lors que l'un des deux membres est fini.

5. Lemme de Fatou et théorème de convergence dominée

On déduit ici du théorème de convergence monotone plusieurs autres résultats fondamentaux de convergence. Le premier, dit lemme de Fatou, est très utile dans bien des situations.

(5.1) Lemme de Fatou. *Soient $f_k, g : P \rightarrow \mathbb{R}$ des fonctions KH-intégrables telles que $f_k \geq g$ presque partout. Si $\liminf_{k \rightarrow +\infty} \int_P f_k(x) dx < +\infty$, alors la fonction $f = \liminf_{k \rightarrow +\infty} f_k$ est finie presque partout et KH-intégrable, et on a*

$$\int_P \liminf_{k \rightarrow +\infty} f_k(x) dx \leq \liminf_{k \rightarrow +\infty} \int_P f_k(x) dx.$$

Démonstration. Quitte à remplacer f_k par $f_k - g$, on peut supposer $f_k \geq 0$ (et $g = 0$). De manière générale la \liminf d'une suite est obtenue comme une limite croissante

$$\liminf_{k \rightarrow +\infty} u_k = \lim_{k \rightarrow +\infty} \uparrow \inf_{i \in [k, +\infty[} u_i.$$

Or pour tout $k \geq 0$ nous avons

$$\int_P \inf_{i \in [k, +\infty[} f_i(x) dx \leq \inf_{i \in [k, +\infty[} \int_P f_i(x) dx$$

puisque l'intégrande $\varphi_k(x) = \inf_{i \in [k, +\infty[} f_i(x)$ du membre de gauche est majoré par f_i pour chaque $i \in [k, +\infty[$ (de plus φ_k est KH-intégrable comme limite décroissante de la suite $s \mapsto \varphi_{k,s}(x) = \min_{i \in [k, s]} f_i(x)$ quand $s \rightarrow +\infty$). Le lemme de Fatou

résulte maintenant du théorème de convergence monotone (4.6) appliqué à la suite croissante (φ_k) . \square

On a bien entendu l'énoncé symétrique, à savoir que si $f_k \leq h$ pour tout k avec h KH-intégrable, alors

$$(5.2) \quad \int_P \limsup_{k \rightarrow +\infty} f_k(x) dx \geq \limsup_{k \rightarrow +\infty} \int_P f_k(x) dx$$

pourvu que le second membre ne soit pas $-\infty$. En combinant (5.1) et (5.2), on obtient le

(5.3) Théorème de la convergence encadrée. Soient $f_k, g, h : P \rightarrow \mathbb{R}$ des fonctions KH-intégrables telles que $g \leq f_k \leq h$ presque partout. On suppose que $f(x) = \lim_{k \rightarrow +\infty} f_k(x)$ existe presque partout. Alors f est KH-intégrable sur P et

$$\lim_{k \rightarrow +\infty} \int_P f_k(x) dx = \int_P f(x) dx.$$

Démonstration. Le lemme de Fatou donne en effet

$$\limsup_{k \rightarrow +\infty} \int_P f_k(x) dx \leq \int_P \lim_{k \rightarrow +\infty} f_k(x) dx \leq \liminf_{k \rightarrow +\infty} \int_P f_k(x) dx$$

puisque $\lim_{k \rightarrow +\infty} f_k = \limsup_{k \rightarrow +\infty} f_k = \liminf_{k \rightarrow +\infty} f_k$ et que les membres de gauche et de droite sont encadrés par $\int_P g(x) dx$ et $\int_P h(x) dx$. \square

(5.4) Théorème de convergence dominée. C'est le cas particulier du théorème de convergence encadrée où la suite (f_k) vérifie la condition plus forte $|f_k| \leq g$ pour une certaine fonction $g \geq 0$ KH-intégrable. Dans ce cas toutes les fonctions f_k sont absolument intégrables et la limite $f = \lim f_k$ vérifie $\|f\|_1 = \lim \|f_k\|_1$.⁽¹⁸⁾

(5.5) Séries convergentes dans L^1 . Soit $f_k : P \rightarrow \mathbb{R}$ une suite de fonctions absolument intégrables telles que $\sum \|f_k\|_1 < +\infty$. Alors la série $\sum f_k$ converge dans $L^1(P)$, et la convergence a lieu également de manière ponctuelle, presque partout sur P .

Démonstration. Posons $S_k = |f_0| + |f_1| + \dots + |f_k|$. C'est une suite croissante de fonctions absolument intégrables telles que $\int_P S_k(x) dx \leq \sum \|f\|_1 < +\infty$. Le théorème de convergence monotone montre que la somme $S(x) = \sum |f_k(x)|$ converge presque partout. En particulier $\varphi(x) = \sum_{k \geq 0} f_k(x)$ existe presque partout comme somme d'une série absolument convergente, et le théorème de convergence dominée appliqué aux sommes partielles $\varphi_k = f_0 + f_1 + \dots + f_k$ entraîne que φ est absolument intégrable, du fait que $0 \leq |\varphi_k| \leq S$. Nous avons de plus $|\varphi - \varphi_k| \leq \sum_{i \in [k+1, +\infty[} |f_i|$ donc $\|\varphi - \varphi_k\|_1 \leq \sum_{i \geq k+1} \|f_i\|_1$, et par conséquent $\lim_{k \rightarrow +\infty} \|\varphi - \varphi_k\|_1 = 0$. \square

Comme conséquence immédiate, nous avons le

(18) En fait les deux théorèmes sont équivalents, puisque le théorème de la convergence encadrée se ramène au cas de fonctions ≥ 0 en observant que l'on a $0 \leq f_k - g \leq h - g$.

(5.6) Théorème. $L^1(P)$ est un espace de Banach (c'est-à-dire un espace normé complet). De plus, de toute suite de Cauchy (donc L^1 -convergente), on peut extraire une sous-suite convergeant ponctuellement presque partout sur P .

Démonstration. C'est une conséquence purement formelle de (5.5). Soit (u_k) une suite de Cauchy dans $L^1(P)$. Il existe une sous-suite (u_{k_i}) telle que $\|u_{k_{i+1}} - u_{k_i}\|_1 \leq 2^{-i}$. Ceci implique que la série $\sum_i (u_{k_{i+1}} - u_{k_i})$ converge normalement vers une somme S dans $L^1(P)$, et donc que la sous-suite (u_{k_i}) converge vers $u = S + u_{k_0}$ en norme L^1 et presque partout. Il en résulte que la suite (u_k) converge elle aussi vers u , en norme L^1 . \square

On peut tirer du théorème de convergence encadrée des propriétés très agréables de continuité et de dérivabilité sous le signe somme des intégrales dépendant de paramètres qui généralisent les résultats du Chapitre I, § 8.⁽¹⁹⁾

(5.7) Théorème. Soit $P \subset \mathbb{R}^n$ un pavé, T une partie de \mathbb{R}^d et t_0 un point de l'adhérence de T dans \mathbb{R}^d . Soit $f : P \times T \rightarrow \mathbb{R}$, $(x, t) \mapsto f(x, t)$ une fonction telle que

- (a) Pour tout $t \in T$, l'application $P \ni x \mapsto f(x, t)$ est KH-intégrable sur P ;
- (b) Il existe des fonctions $g, h : P \rightarrow \mathbb{R}$ KH-intégrables et un voisinage V de t_0 tel que $g(x) \leq f(x, t) \leq h(x)$ pour tout $t \in T \cap V$ et presque tout $x \in P$ (l'ensemble négligeable $N(t) \subset P$ correspondant peut dépendre de t).
- (c) Pour presque tout $x \in P$, l'application $T \ni t \mapsto f(x, t)$ possède une limite $\varphi(x)$ quand $t \rightarrow t_0$.

Alors φ est KH-intégrable sur P et

$$\lim_{T \ni t \rightarrow t_0} \int_P f(x, t) dx = \int_P \varphi(x) dx.$$

(5.8) Continuité sous le signe somme. Sous les hypothèses 5.7 (a, b), si $t_0 \in T$ et si $T \ni t \mapsto f(x, t)$ est continue en t_0 pour presque tout $x \in P$, on a

$$\lim_{T \ni t \rightarrow t_0} \int_P f(x, t) dx = \int_P f(x, t_0) dx,$$

c'est-à-dire que l'application $F(t) = \int_P f(x, t) dx$ est continue en t_0 .

Démonstration de (5.7). Il suffit de montrer qu'il y a convergence pour toute suite $t_k \in T$ tendant vers t_0 . Or par hypothèse, nous avons $\varphi_k(x) = f(x, t_k) \rightarrow \varphi(x)$ sauf sur un ensemble négligeable $E \subset P$, tandis que $g(x) \leq \varphi_k(x) \leq h(x)$ en dehors de $N(t_k)$. Il suffit d'appliquer le théorème de convergence encadrée, les hypothèses étant satisfaites en dehors de l'ensemble négligeable $E' = E \cup \bigcup N(t_k)$. \square

(5.9) Dérivation sous le signe somme. Soient $P \subset \mathbb{R}^n$ un pavé, $T \subset \mathbb{R}$ un intervalle, $t_0 \in T$ un point fixé et $f : P \times T \rightarrow \mathbb{R}$, $(x, t) \mapsto f(x, t)$ une fonction tels que

(19) Comme on va le voir, on peut le faire sous des conditions même plus générales que dans la théorie de Lebesgue, avec des intégrales non nécessairement absolument convergentes – on profite en cela de ce que la convergence encadrée est plus générale que la convergence dominée. Les preuves sont cependant identiques.

- (a) pour tout $t \in T$, l'application $P \ni x \mapsto f(x, t)$ est KH-intégrable sur P ;
- (b) pour presque tout $x \in P$, l'application $t \mapsto f(x, t)$ admet une dérivée partielle $\frac{\partial f}{\partial t}(x, t)$ sur T , et celle-ci est continue en $t_0 \in T$;
- (c) il existe un voisinage V de t_0 et des fonctions $g, h : P \rightarrow \mathbb{R}$ KH-intégrables telles que $g(x) \leq \frac{\partial f}{\partial t}(x, t) \leq h(x)$ pour tout $t \in T \cap V$ et presque tout $x \in P$ (l'ensemble négligeable $N(t) \subset P$ correspondant peut dépendre a priori de t).

Alors l'application $F(t) = \int_P f(x, t) dx$ est différentiable au point t_0 et on a

$$F'(t_0) = \int_P \frac{\partial f}{\partial t}(x, t_0) dx.$$

De plus si $t \mapsto \frac{\partial f}{\partial t}(x, t)$ est continue sur T pour presque tout $x \in P$, alors F est de classe C^1 sur T et la formule ci-dessus a lieu pour tout $t_0 \in T$.

Démonstration. En appliquant le théorème des accroissements finis à $t \mapsto f(x, t)$, on voit que

$$\frac{F(t) - F(t_0)}{t - t_0} = \int_P \frac{f(x, t) - f(x, t_0)}{t - t_0} dx = \int_P \frac{\partial f}{\partial t}(x, c_{t,x}) dx$$

pour un certain point $c = c_{t,x} \in]t_0, t[$. Soit $V =]t_0 - \varepsilon, t_0 + \varepsilon[$ tel que (c) ait lieu, et soit (t_k) une suite de points de V convergeant vers t_0 . L'hypothèse (c) est vérifiée pour tout $x \in P \setminus N'$ et $t \in \{t_k\}$, avec $N' = \bigcup N(t_k)$. Soit E un ensemble négligeable en dehors duquel (b) a lieu. Lorsque $x \in P \setminus (E \cup N')$ et $t = t_k$ nous avons

$$g(x) \leq \frac{\partial f}{\partial t}(x, c_{t_k,x}) \leq h(x) \quad \text{et} \quad \frac{\partial f}{\partial t}(x, c_{t_k,x}) \rightarrow \frac{\partial f}{\partial t}(x, t_0).$$

Le résultat découle alors de nouveau du théorème de convergence encadrée. \square

Pour un paramètre $t = (t_1, \dots, t_d) \in \mathbb{R}^d$, nous avons le résultat analogue suivant.

(5.10) Différentiabilité sous le signe somme. Soit $P \subset \mathbb{R}^n$ un pavé et $T \subset \mathbb{R}^d$ un ouvert. Soit $f : P \times T \rightarrow \mathbb{R}$, $(x, t) \mapsto f(x, t)$ une fonction telle que

- (a) pour tout $t \in T$, l'application $P \ni x \mapsto f(x, t)$ est KH-intégrable sur P ;
- (b) pour presque tout $x \in P$, l'application $t \mapsto f(x, t)$ admet des dérivées partielles $\frac{\partial f}{\partial t_j}(x, t)$ continues sur T ;
- (c) pour tout point $t_0 \in T$, il existe un voisinage V de t_0 et des fonctions $g_j, h_j : P \rightarrow \mathbb{R}$ KH-intégrables telles que $g_j(x) \leq \frac{\partial f}{\partial t_j}(x, t) \leq h_j(x)$ pour tout $t \in T \cap V$ et presque tout $x \in P$ (l'ensemble négligeable $N_j(t) \subset P$ correspondant peut dépendre a priori de t).

Alors l'application $F(t) = \int_P f(x, t) dx$ est différentiable sur T et on a

$$\frac{\partial F}{\partial t_j}(t) = \int_P \frac{\partial f}{\partial t_j}(x, t) dx.$$

6. Lemme de recouvrement de Vitali et différentiabilité presque partout des intégrales indéfinies

Nous commençons par un lemme de recouvrement très utile, puis nous en donnons quelques applications fondamentales. On définit *l'aspect* $\text{aspect}(P)$ d'un pavé P de \mathbb{R}^n d'intérieur non vide comme étant le rapport entre sa plus grande et sa plus petite arête.

(6.1) Lemme de recouvrement de Vitali. *Soit $\mathcal{V} = \{Q_\alpha\}_{\alpha \in A}$ une famille de pavés fermés contenus dans un pavé fermé borné P , et soit E une partie de P . On dit que \mathcal{V} est un recouvrement de Vitali de E si :*

- (a) *L'aspect des pavés Q_α reste encadré par deux constantes positives, c'est-à-dire qu'il existe une constante $C \geq 1$ telle que $\text{aspect}(Q_\alpha) \leq C$ pour tout α .*
- (b) *pour tout $x \in E$ et tout $\varepsilon > 0$, il existe un pavé $Q \in \mathcal{V}$ tel que $x \in Q$ et $\text{diam}(Q) < \varepsilon$.*

Alors :

- (i) *il existe une famille finie ou dénombrable Q_k de pavés deux à deux disjoints de \mathcal{V} et une partie négligeable N de P telles que $E \subset \bigcup_{k \geq 0} Q_k \cup N$;*
- (ii) *on peut de plus choisir N contenu dans une intersection $\bigcap_{s \geq 0} \bigcup_{k \geq s} Q'_k$ avec des pavés fermés $Q'_k \subset P$ tels que $\sum_{k \geq 0} m(Q'_k) < +\infty$.*

Démonstration. On définit par récurrence Q_0, Q_1, \dots, Q_k comme suit. On pose

$$F_0 = \emptyset, \quad F_k = Q_0 \cup Q_1 \cup \dots \cup Q_{k-1} \quad \text{si } k \geq 1,$$

et on considère $\mathcal{V}_{E,k} \subset \mathcal{V}$ la sous-famille des pavés $Q \in \mathcal{V}$ tels que Q contienne un point $x \in E$ et $Q \subset P \setminus F_k$. S'il existe un entier k tel que $E \setminus F_k = \emptyset$ pour un certain k , alors la famille finie (Q_0, \dots, Q_{k-1}) répond à la question et le lemme est démontré avec $N = \emptyset$. Sinon $E \setminus F_k \neq \emptyset$ pour tout $k \geq 0$ et alors la famille $\mathcal{V}_{E,k}$ est non vide : en effet, $x \in E \setminus F_k$ étant fixé, il existe des pavés $Q \in \mathcal{V}$ contenant x vérifiant $\text{diam}(Q) < \varepsilon = d(x, F_k)$, de sorte qu'on a nécessairement $Q \subset P \setminus F_k$. Soit μ_k le sup de la mesure $m(Q)$ de tous les pavés $Q \in \mathcal{V}_{E,k}$; on choisit $Q_k \in \mathcal{V}_{E,k}$ en sorte que $m(Q_k) \geq \frac{1}{2}\mu_k$. Les pavés Q_k sont bien disjoints par construction, et on a par conséquent $\sum m(Q_k) \leq m(P) < +\infty$, donc $\sum \mu_k \leq 2 \sum m(Q_k) < +\infty$.

Nous allons montrer que $N = E \setminus \bigcup_{k \geq 0} Q_k$ est négligeable. Soit $x \in N$. Pour tout entier s , on a $x \in E \setminus F_s$, et il existe un pavé $Q \in \mathcal{V}$ de diamètre $\text{diam}(Q) < d(x, F_s)$ contenant x . Par suite $Q \subset P \setminus F_s$. Or, si $Q \subset P \setminus F_k$, alors $Q \in \mathcal{V}_{E,k}$ et $m(Q) \leq \mu_k$ par définition de μ_k . Comme $\lim \mu_k = 0$ ceci ne peut se produire que pour un nombre fini d'indices k . Choisissons le plus grand indice k possible, de sorte que $k \geq s$, $Q \subset P \setminus F_k$ et $Q \cap F_{k+1} \neq \emptyset$. Ceci implique que $Q \cap Q_k \neq \emptyset$. Si a est la plus grande arête de Q et b_k la plus petite de Q_k nous avons

$$a^n \leq C^{n-1} m(Q), \quad m(Q) \leq \mu_k \leq 2 m(Q_k), \quad m(Q_k) \leq C^{n-1} b_k^n,$$

par conséquent que $a \leq C_1 b_k$ avec $C_1 = 2^{1/n} C^{2-2/n}$. Ceci entraîne que Q est contenu dans le pavé Q'_k de même centre que Q_k et homothétique dans le rapport $C_2 = 1 + 2C_1$, par suite $m(Q'_k) \leq C_3 m(Q_k)$ avec $C_3 = C_2^n = (1 + 2^{1+1/n} C^{2-2/n})^n$. Nous avons

par conséquent $N \subset \bigcup_{k \geq s} Q'_k$ et comme ceci est vrai pour tout entier s on voit que $N \subset \bigcap_{s \geq 0} \bigcup_{k \geq s} Q'_k$. Cependant

$$\sum_{k \geq 0} m(Q'_k) \leq C_3 \sum_{k \geq 0} m(Q_k) \leq C_3 m(P) < +\infty$$

et donc $m(\bigcup_{k \geq s} Q'_k) \leq \sum_{k \geq s} m(Q'_k) \rightarrow 0$ quand $s \rightarrow +\infty$. Par conséquent N est négligeable et (i) est démontré. Pour obtenir (ii), il suffit de remplacer éventuellement Q'_k par $Q'_k \cap P$ pour avoir à coup sûr $Q'_k \subset P$. \square

(6.2) Théorème. *Soit $E \subset \mathbb{R}^n$ une partie intégrable. Alors pour tout $\varepsilon > 0$ il existe une réunion de cubes ouverts $U = \bigcup_{k \geq 0} Q_k$ contenant E telle que*

$$m(U) \leq \sum_{k \geq 0} m(Q_k) \leq m(E) + \varepsilon.$$

Démonstration. (1) Commençons par le cas où E est borné, $E \subset P$ pavé fermé borné. Soit $\varepsilon > 0$ et δ une jauge ε -adaptée à $f = \chi_E$ sur P . On considère la famille de pavés fermés

$$\mathcal{V}_\delta = \{Q \text{ cube fermé} \subset P; \exists x \in E, Q \ni x \text{ et } \text{diam}(Q) \leq \delta(x)\}.$$

C'est un recouvrement de Vitali de E . Il existe par conséquent des cubes fermés disjoints $Q_k \in \mathcal{V}_\delta$ et un ensemble négligeable N tels que $E \subset \bigcup_{k \geq 0} Q_k \cup N$, avec de plus $N \subset \bigcup_{k \geq 0} Q'_k$, $\sum_{k \geq 0} m(Q'_k) \leq \varepsilon$. Soit $x_k \in E$ tel que $x_k \in Q_k$ et $\text{diam}(Q_k) \leq \delta(x_k)$.

Grâce au lemme 1.3, la famille finie d'intervalles pointés disjoints $(Q_k, x_k)_{0 \leq k \leq s}$ peut être complétée en une subdivision δ -fine D de P . Comme $x_k \in E$, on en déduit

$$\sum_{k=0}^s m(Q_k) = \sum_{k=0}^s \text{vol}(Q_k) \chi_E(x_k) \leq S_D(\chi_E) \leq \int_P \chi_E(x) dx + \varepsilon = m(E) + \varepsilon.$$

Quand $s \rightarrow +\infty$, il vient à la limite $\sum_{k \geq 0} m(Q_k) \leq m(E) + \varepsilon$, donc quitte à remplacer Q_k et Q'_k par des cubes ouverts $\tilde{Q}_k \supset Q_k$, $\tilde{Q}'_k \supset Q'_k$ tels que $m(\tilde{Q}_k) \leq m(Q_k) + 2^{-k-1}\varepsilon$, $m(\tilde{Q}'_k) \leq m(Q'_k) + 2^{-k-1}\varepsilon$, il vient

$$E \subset \bigcup_{k \geq 0} \tilde{Q}_k \cup \tilde{Q}'_k, \quad \sum_{k \geq 0} m(\tilde{Q}_k) + m(\tilde{Q}'_k) \leq m(E) + 4\varepsilon.$$

Quitte à remplacer ε par $\varepsilon/4$, le résultat est démontré dans le cas où E est borné.

(2) Dans le cas général $E \subset \mathbb{R}^n$, on écrit $E = \bigcup_{k \in \mathbb{Z}^n} E \cap P_k$ où $P_k = k + [0, 1]^n$, et on trouve pour chaque $k \in \mathbb{Z}^n$ un ouvert U_k contenant $E \cap P_k$ et des cubes ouverts $Q_{k,\ell}$ le recouvrant tels que

$$m(U_k) \leq \sum_{\ell} m(Q_{k,\ell}) \leq m(E \cap P_k) + 2^{-\Sigma(k_i+1)}\varepsilon.$$

Alors $U = \bigcup U_k$ contient E et on a $m(U) \leq \sum_{k,\ell} m(Q_{k,\ell}) \leq m(E) + \varepsilon$. \square

(6.3) Corollaire. Soit E une partie de \mathbb{R}^n . Les propriétés suivantes sont équivalentes.

- (a) E est négligeable.
- (b) Pour tout $\varepsilon > 0$, il existe une suite finie ou dénombrable de cubes (resp. pavés) ouverts $(Q_k)_{k \geq 0}$ tels que $E \subset \bigcup_{k \geq 0} Q_k$ et $\sum_{k \geq 0} m(Q_k) \leq \varepsilon$.
- (c) Pour tout $\varepsilon > 0$, il existe une suite finie ou dénombrable de cubes (resp. pavés) fermés $(Q_k)_{k \geq 0}$ tels que $E \subset \bigcup_{k \geq 0} Q_k$ et $\sum_{k \geq 0} m(Q_k) \leq \varepsilon$.

Démonstration. (a) \Rightarrow (b) grâce à 6.2, tandis que (b) \Rightarrow (c) est évident, et (c) \Rightarrow (a) en prenant une intersection dénombrable avec $\varepsilon = 1/s \rightarrow 0$, $s \in \mathbb{N}^*$. \square

(6.4) Théorème de Lebesgue-Denjoy. Soit $f : P \rightarrow \mathbb{R}$ une fonction KH-intégrable sur un pavé $P \subset \mathbb{R}^n$ et $\mathcal{V}_{P,C}$ l'ensemble des pavés $Q \subset P$ tels que $\text{aspect}(Q) \leq C$. Alors pour presque tout $x \in P$ on a

$$(*) \quad \lim_{\mathcal{V}_{P,C} \ni Q \ni x, \text{diam}(Q) \rightarrow 0} \frac{1}{\text{vol}(Q)} \int_Q f(t) dt = f(x).$$

Démonstration. Comme nous allons le voir, il s'agit d'une conséquence directe du lemme de Henstock, combiné avec le lemme de recouvrement de Vitali.

Soit E l'ensemble des points $x \in P$ pour lesquels la limite $(*)$ n'existe pas, ou bien vaut une certaine valeur $\ell(x) \neq f(x)$. Il s'agit de montrer que E est négligeable. Écrivons $E = \bigcup_{s>0} K_s$, où K_s est l'ensemble des points $x \in P$ tels que pour tout $\eta > 0$ il existe $Q \in \mathcal{V}_{P,C}$ tel que $Q \ni x$, $\text{diam}(Q) \leq \eta$ et

$$\left| f(x) - \frac{1}{\text{vol}(Q)} \int_Q f(t) dt \right| \geq \frac{1}{s}.$$

Fixons $\varepsilon > 0$ et δ une jauge ε -adaptée à f . On observe que la famille $\mathcal{V}_{P,C,s,\delta}$ des pavés pointés (Q, x) vérifiant la minoration précédente et tels que $\text{diam}(Q) \leq \delta(x)$ fournit un recouvrement de Vitali de K_s . Il existe par conséquent une famille finie ou dénombrable (Q_k, x_k) de pavés de $\mathcal{V}_{P,C,s,\delta}$ (qui dépend aussi de ε via δ) et un ensemble négligeable $N_{s,\varepsilon}$ tels que $K_s \subset R_{s,\varepsilon} := \bigcup_k Q_k \cup N_{s,\varepsilon}$. Le lemme de Henstock 1.1 (b) appliqué à la famille finie de pavés pointés $\{(Q_k, x_k)\}_{0 \leq k \leq m}$ donne

$$\sum_{0 \leq k \leq m} \text{vol}(Q_k) \left| f(x_k) - \frac{1}{\text{vol}(Q_k)} \int_{Q_k} f(t) dt \right| \leq 2\varepsilon$$

puisque cette famille est δ -fine par construction. On en déduit $\sum_{0 \leq k \leq m} \frac{1}{s} \text{vol}(Q_k) \leq 2\varepsilon$ pour tout m , donc $m(R_{s,\varepsilon}) = m(\bigcup_{k \geq 0} Q_k) \leq 2s\varepsilon$. En écrivant $K_s \subset \bigcap_{k>0} R_{s,1/k}$ on voit que l'ensemble K_s est négligeable, donc $E = \bigcup_{s>0} K_s$ l'est aussi. \square

(6.5) Corollaire. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction KH-intégrable et $F(x) = \int_a^x f(t) dt$ son intégrale indéfinie. Alors F est presque partout dérivable de dérivée $F'(x) = f(x)$.

Démonstration. C'est un corollaire immédiat du théorème précédent en choisissant des pavés Q de la forme $[x, x + h]$ (resp. $[x - h, x]$), de sorte que

$$\frac{1}{\text{vol}(Q)} \int_Q f(t) dt = \frac{F(x+h) - F(x)}{h}, \quad \text{resp.} \quad \frac{F(x) - F(x-h)}{h}. \quad \square$$

7. Ensembles et fonctions mesurables

Une fonction intégrable peut être extrêmement irrégulière du point de vue de la continuité (par exemple χ_Q est KH-intégrable sur \mathbb{R} , bien qu'elle soit discontinue en tout point). Une fonction intégrable doit tout de même satisfaire certaines propriétés très faibles de régularité locale, pour lesquelles les ensembles négligeables ne jouent aucun rôle. C'est précisément l'objet de la notion de mesurabilité.

(7.1) Définition. On dit qu'une partie $E \subset \mathbb{R}^n$ est mesurable si $E \cap P$ est intégrable pour tout pavé fermé borné $P \subset \mathbb{R}^n$.

Compte tenu des propriétés des ensembles intégrables (cf. (4.1)), on peut bien entendu se limiter à vérifier que l'intersection $E \cap [-k, k]^n$ est intégrable pour tout entier k . Le théorème suivant découle immédiatement de (4.1).

(7.2) Théorème. Toute réunion dénombrable $\bigcup E_k$, toute intersection dénombrable $\bigcap E_k$ de parties mesurables E_k est mesurable. Le complémentaire $\complement E$ d'une partie mesurable E est mesurable. Tout pavé, toute partie ouverte ou fermée $E \subset \mathbb{R}^n$ est mesurable.

La dernière assertion résulte du fait que les pavés sont trivialement mesurables, et du fait qu'une partie ouverte de \mathbb{R}^n est réunion finie ou dénombrable de pavés ouverts. On peut étendre la mesure de Lebesgue aux parties mesurables en posant

$$m(E) = \lim_{k \rightarrow +\infty} m(E \cap [-k, k]^n), \quad m(E) \in [0, +\infty].$$

Les parties intégrables sont alors exactement les parties mesurables de mesure $m(E) < +\infty$, et la propriété d'additivité de la mesure d'une réunion dénombrable de parties mesurables disjointes est encore valable lorsque les sommations sont prises dans $[0, +\infty]$. Nous avons la caractérisation suivante assez claire de la mesurabilité.

(7.3) Caractérisation des ensembles mesurables. Soit E une partie de \mathbb{R}^n . Les propriétés suivantes sont équivalentes.

- (a) E est mesurable.
- (b) Pour tout $\varepsilon > 0$ il existe une partie fermée F et une partie ouverte U telles que $F \subset E \subset U$ et $m(U \setminus F) \leq \varepsilon$.
- (c) Il existe F un F_σ (= réunion finie ou dénombrable de fermés) et G un G_δ (= intersection finie ou dénombrable d'ouverts) tels que $F \subset E \subset G$ et $m(G \setminus F) = 0$.
- (d) E peut s'écrire comme une réunion $E = F \cup N$ d'un F_σ et d'une partie négligeable.

Démonstration. (a) \Rightarrow (b). Commençons par le cas où E est borné, $E \subset P$ pavé fermé borné. Alors E et $E' = P \setminus E$ sont des parties intégrables. D'après le théorème (6.2) il existe des ouverts U et U' de \mathbb{R}^n tels que $E \subset U$, $m(U \setminus E) \leq \varepsilon/2$, et $E' \subset U'$, $m(U' \setminus E') \leq \varepsilon/2$. On pose $F = P \setminus U'$. Il vient $F \subset P \setminus E' = E \subset U$ et

$$m(U \setminus F) \subset m(U \setminus E) + m(E \setminus F) \leq m(U \setminus E) + m(U' \setminus E') \leq \varepsilon.$$

Dans le cas général, on pose $E_k = E \cap Q_k$, $Q_k = k + [0, 1]^n$, $k \in \mathbb{Z}^n$, et on trouve des parties fermées $F_k \subset Q_k$ et ouvertes $U_k \subset \mathbb{R}^n$ telles que $F_k \subset E_k \subset U_k$ et $m(U_k \setminus F_k) \leq 2^{-\Sigma(k_i+1)}\varepsilon$. Alors $F = \bigcup F_k$ et $U = \bigcup U_k$ répondent à la question puisque $U \setminus F \subset \bigcup (U_k \setminus F_k)$.

(b) \Rightarrow (c). Pour tout entier $s > 0$, on peut trouver un fermé F_s et un ouvert U_s tels que $F_s \subset E \subset U_s$ et $m(U_s \setminus F_s) < 1/s$. Alors $F = \bigcup F_s$ et $G = \bigcup U_s$ répondent à la question, puisque $G \setminus F \subset U_s \setminus F_s$ pour tout s .

(c) \Rightarrow (d) est évident, il suffit de poser $N = E \setminus F$ qui est contenu dans $G \setminus F$, donc négligeable.

(d) \Rightarrow (a) résulte de (7.2). □

(7.4) Théorème et définition. Soit $f : P \rightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ une fonction quelconque définie sur un pavé $P \subset \mathbb{R}^n$. Les propriétés suivantes sont équivalentes.

- (a) L'image réciproque $f^{-1}([-\infty, c[)$ de tout intervalle ouvert de $\overline{\mathbb{R}}$ est mesurable.
- (b) L'image réciproque $f^{-1}([-\infty, c])$ de tout intervalle fermé de $\overline{\mathbb{R}}$ est mesurable.
- (c) L'image réciproque $f^{-1}(J)$ de tout intervalle $J \subset \overline{\mathbb{R}}$ est mesurable, respectivement [(c') J ouvert], [(c'') J fermé],
- (d) L'image réciproque $f^{-1}(U)$ de tout ouvert $U \subset \overline{\mathbb{R}}$ est mesurable.

On dit alors que la fonction f est mesurable.

Démonstration. L'équivalence de (a) et (b) résulte des égalités

$$f^{-1}([-\infty, c]) = \bigcap_{n>0} f^{-1}([-\infty, c + 1/n]), \quad f^{-1}([-\infty, c[) = \bigcup_{n>0} f^{-1}([-\infty, c - 1/n])$$

si $c \in \mathbb{R}$. Le passage au cas d'intervalles J quelconques (c), (c'), (c'') se voit aisément en écrivant par exemple

$$f^{-1}(]c, d]) = f^{-1}([-\infty, d]) \cap \mathbb{C} f^{-1}([-\infty, c])$$

pour tous $c < d$ dans $\overline{\mathbb{R}}$. Enfin l'équivalence avec (d) résulte du fait que tout ouvert $U \subset \overline{\mathbb{R}}$ est réunion dénombrable d'intervalles ouverts disjoints Q_k , de sorte que $f^{-1}(U) = \bigcup f^{-1}(Q_k)$. □

Une conséquence immédiate de la définition est la suivante.

(7.5) Proposition. Soit $f = (f_1, \dots, f_m) : P \rightarrow E$ une famille de fonctions mesurables définies sur un pavé $P \subset \mathbb{R}^n$, à valeurs dans une partie $E \subset \mathbb{R}^m$. Si $g : E \rightarrow \overline{\mathbb{R}}$ est une fonction continue, alors $G = g(f_1, \dots, f_m)$ est mesurable.

Démonstration. Si U est un ouvert de $\overline{\mathbb{R}}$, alors $g^{-1}(U)$ est un ouvert de E , donc la trace $E \cap V$ d'un ouvert $V \subset \mathbb{R}^m$. On peut écrire V comme une réunion dénombrable de pavés ouverts $Q_k = Q_k^1 \times \dots \times Q_k^m$ et on obtient ainsi

$$G^{-1}(U) = \bigcup_k (f_1, \dots, f_m)^{-1}(Q_k) = \bigcup_k f_1^{-1}(Q_k^1) \cap \dots \cap f_m^{-1}(Q_k^m).$$

Ceci montre que G est mesurable. \square

(7.6) Corollaire. *Toute combinaison linéaire finie, tout produit de fonctions mesurables à valeurs réelles est encore mesurable. L'ensemble des fonctions mesurables $f : P \rightarrow \mathbb{R}$ a donc une structure d'algèbre.* \square

Par ailleurs, comme les ensembles négligeables sont mesurables, la mesurabilité n'est pas sensible au fait que la fonction f soit modifiée arbitrairement sur un ensemble négligeable. Un fait très utile est que la mesurabilité est préservée par passage à la limite dénombrable :

(7.7) Proposition. *Soit $f_k : P \rightarrow \overline{\mathbb{R}}$ une suite de fonctions mesurables. Alors $\limsup f_k$ et $\liminf f_k$ sont mesurables. Si (f_k) admet presque partout une limite f , alors f est mesurable.*

Démonstration. Posons $f = \limsup f_k$. Nous avons alors par définition $f(x) < c$ si $\exists s > 0, \exists N \geq 0, \forall k \geq N, f_k(x) < c - 1/s$, c'est-à-dire

$$f^{-1}([-\infty, c]) = \bigcup_{s>0} \bigcup_{N \geq 0} \bigcap_{k \geq N} f_k^{-1}([-\infty, c - 1/s]),$$

ce qui montre que f est mesurable. La preuve pour $\liminf f_k$ se déduit du cas de la \limsup en remplaçant (f_k) par $(-f_k)$. Si (f_k) admet une limite presque partout f , on conclut en écrivant par exemple $f = \limsup f_k$ presque partout. \square

(7.8) Proposition. *Toute fonction $f : P \rightarrow \mathbb{R}$ continue ou KH-intégrable sur un pavé $P \subset \mathbb{R}^n$ est mesurable.*

Démonstration. Il est clair que les fonctions continues sont mesurables, puisque l'image inverse d'un intervalle ouvert est une partie ouverte (donc mesurable) de P . Supposons maintenant f KH-intégrable. Soit

$$F(x_1, \dots, x_n; y_1, \dots, y_n) = \int_{x_1}^{y_1} \dots \int_{x_n}^{y_n} f(t_1, \dots, t_n) dt_1 \dots dt_n$$

l'« intégrale indéfinie » de f sur $P \times P$. On sait que F est continue sur $P \times P$ d'après la proposition 1.15, et que (par exemple)

$$f(x) = \lim_{h \rightarrow 0} \frac{1}{h^n} F(x_1, \dots, x_n; x_1 + h, \dots, x_n + h)$$

pour presque tout $x \in P$ (théorème 6.4). En prenant $h = 1/k \rightarrow 0$, on voit que f est égale presque partout à une limite d'une suite de fonctions continues, par conséquent f est mesurable. \square

(7.9) Proposition. *Toute fonction mesurable $f : P \rightarrow \overline{\mathbb{R}}$ peut s'écrire $f = \lim f_k$ où les fonctions f_k sont des fonctions intégrables étagées, c'est-à-dire des combinaisons linéaires finies $\sum_{1 \leq \ell \leq N_k} c_{k,\ell} \chi_{E_{k,\ell}}$ de fonctions caractéristiques d'ensembles intégrables.*

Si de plus $f \geq 0$, on peut choisir la suite (f_k) croissante et telle que $0 \leq f_k \leq f$.

Démonstration. Supposons d'abord P fermé borné. On prend alors

$$\begin{aligned} c_{k,\ell} &= (\ell - 1)2^{-k} - k, & E_{k,\ell} &= f^{-1}([c_{k,\ell}, c_{k,\ell} + 2^{-k}[), & \text{si } \ell &= 1, 2, \dots, k 2^{k+1}, \\ c_{k,\ell} &= k, & E_{k,\ell} &= f^{-1}([k, +\infty[), & \text{si } \ell &= k 2^{k+1} + 1, \\ c_{k,\ell} &= -k, & E_{k,\ell} &= f^{-1}([-\infty, -k[), & \text{si } \ell &= p_n = k 2^{k+1} + 2, \end{aligned}$$

de sorte que les ensembles $E_{k,\ell}$, $1 \leq \ell \leq N_k$ forment une partition mesurable de P . Il est évident par construction que la fonction $f_n = \sum c_{k,\ell} \chi_{E_{k,\ell}}$ vérifie $|f_k - f| \leq 2^{-k}$ là où $|f(x)| < k$, tandis que $|f_k(x)| = k$ avec le même signe que $f(x)$ là où $|f(x)| \geq k$. On a donc bien $\lim f_k = f$. Lorsque P est borné, les ensembles $E_{k,\ell}$ sont intégrables et la proposition est démontrée. Si P n'est pas borné, il suffit de remplacer $E_{k,\ell}$ par $E'_{k,\ell} = E_{k,\ell} \cap [-k, k]^n$ pour conclure la démonstration. En effet, il est facile de constater que $0 \leq f_k \leq f$ et que la suite (f_k) est croissante lorsque f est elle-même positive ou nulle. \square

Le résultat suivant montre que la mesurabilité est la propriété de régularité requise pour obtenir l'intégrabilité, lorsqu'on a un encadrement par des fonctions KH-intégrables.

(7.10) Proposition. *Soit $f, g, h : P \rightarrow \mathbb{R}$ des fonctions telles que $g \leq f \leq h$. On suppose que g et h sont KH-intégrables et que f est mesurable. Alors f est KH-intégrable.*

Démonstration. Quitte à remplacer f par $f - g$ et h par $h - g$, on peut supposer $0 \leq f \leq h$ avec h KH-intégrable. La proposition 7.9 implique que $f = \lim f_k$ avec des fonctions f_k étagées KH-intégrables, que l'on peut choisir ≥ 0 d'après la démonstration. Toutes ces fonctions sont donc absolument intégrables, et on peut écrire $f = \lim_{k \rightarrow +\infty} \min(f_k, h)$. On voit donc que f est KH-intégrable comme limite dominée de fonctions KH-intégrables. \square

(7.11) Espace $L^\infty(P)$. Soit $f : P \rightarrow \mathbb{R}$ (ou même plus généralement $f : P \rightarrow \overline{\mathbb{R}}$) une fonction mesurable. On définit le sup essentiel de f par

$$\begin{aligned} \sup_{x \in P} \text{ess } f(x) &= \min \{ M \in [-\infty, +\infty]; f(x) \leq M \text{ presque partout sur } P \} \\ &= \min \{ M \in [-\infty, +\infty]; f^{-1}(]M, +\infty]) \text{ est négligeable} \}. \end{aligned}$$

Il s'agit bien d'un minimum du fait que si M est la limite d'une suite M_k strictement décroissante pour laquelle $f^{-1}(]M_k, +\infty])$ est négligeable, alors $f^{-1}(]M, +\infty]) = \bigcup f^{-1}(]M_k, +\infty])$ est encore négligeable. De même, on définit l'inf essentiel de f par

$$\inf_{x \in P} \text{ess } f(x) = \max \{ \mu \in [-\infty, +\infty]; f^{-1}([-\infty, \mu]) \text{ est négligeable} \}.$$

On définit la norme L^∞ de f par

$$\|f\|_\infty = \max \left(\sup_{x \in P} \text{ess } f(x), \sup_{x \in P} \text{ess } (-f(x)) \right),$$

de sorte que $|f(x)| \leq M = \|f\|_\infty$ presque partout, M étant le plus petit majorant pour lequel ceci ait lieu.

Enfin, on définit l'espace $\tilde{L}^\infty(P)$ des fonctions dites *essentiellement bornées* comme l'ensemble des fonctions $f : P \rightarrow \mathbb{R}$ de norme $\|f\|_\infty$ finie, et $L^\infty(P) = \tilde{L}^\infty(P)/N(P)$ comme le quotient de $\tilde{L}^\infty(P)$ par l'espace des fonctions négligeables $N(P)$ – lequel est précisément donné par $N(P) = \{f \in \tilde{L}^\infty(P) ; \|f\|_\infty = 0\}$.

Il est clair que si $f \in L^\infty(P)$, on peut choisir un représentant de la fonction f tel qu'on ait $|f(x)| \leq \|f\|_\infty$ *partout*, et pas seulement presque partout [quitte à redéfinir $f(x) = 0$ sur l'ensemble négligeable dans lequel $|f(x)| > \|f\|_\infty$]. On voit alors que toute série normalement convergente dans $L^\infty(P)$ est convergente, de sorte que $L^\infty(P)$ est encore un espace de Banach.

(7.12) Inégalité de Hölder pour L^1 et L^∞ : pour tout couple de fonctions $f \in L^1(P)$ et $g \in L^\infty(P)$, on a

$$\left| \int_P f(x)g(x) dx \right| \leq \|f\|_1 \|g\|_\infty.$$

En effet, ceci résulte du fait que $|f(x)g(x)| \leq \|g\|_\infty |f(x)|$ presque partout sur P . \square

Le résultat suivant montre que les fonctions intégrables au sens de Kurzweil-Henstock doivent tout de même être absolument intégrables sur des ensembles assez gros, bien qu'il puisse subsister des endroits où on a affaire à une intégrale oscillante non absolument convergente. La démonstration exploite la mesurabilité.

(7.13) Théorème. Soit P un pavé fermé borné et $f : P \rightarrow \mathbb{R}$ une fonction intégrable au sens de Kurzweil-Henstock sur P . Alors

(a) Il existe une suite croissante de parties fermées $K_s \subset P$ telle que $\bigcup K_s = P$ et

$$\int_{K_s} |f(x)| dx = \int_P \chi_{K_s}(x) |f(x)| dx < +\infty.$$

En particulier, la mesure $m(P \setminus K_s)$ peut être prise arbitrairement petite.

(b) L'ensemble $U \subset P$ des points x_0 pour lesquels il existe un pavé ouvert Q de centre x_0 sur lequel $\int_Q |f(x)| dx < +\infty$ est un ouvert dense dans P .

Démonstration. (a) Fixons $\varepsilon_0 > 0$ et une jauge δ_0 ε_0 -adaptée à f . On pose

$$E_s = \{x \in P ; |f(x)| \leq s \text{ et } \delta_0(x) \geq 1/s\} \quad \text{et} \quad K_s = \overline{E_s}, \quad s \in \mathbb{N}^*.$$

Il est clair que $E_s \subset K_s$ sont des suites croissantes d'ensembles et que $\bigcup E_s = P$ (du fait que f est partout finie et δ_0 partout > 0); on a donc aussi $\bigcup K_s = P$. D'après le lemme de Henstock 1.1 (b), nous obtenons

$$(7.14) \quad \sum_{0 \leq j < N} \left| \text{vol}(Q_j) f(x_j) - \int_{Q_j} f(x) dx \right| \leq 2\varepsilon_0$$

pour toute subdivision pointée $D = \{(Q_j, x_j)\}_{0 \leq j < N}$ δ_0 -fine (complète ou partielle) de P . Soit A un réel positif assez grand et $\varepsilon > 0$ arbitrairement petit. La fonction $g(x) = \chi_{K_s}(x) \min(|f(x)|, A)$ est mesurable bornée, donc KH-intégrable, ce qui implique

$$\begin{aligned}
 \int_P \chi_{K_s}(x) \min(|f(x)|, A) dx &= \int_P g(x) dx = \limsup_{D, \text{KH}} S_D(g) \\
 &= \inf_{\delta} \sup_{D \text{ } \delta\text{-fine}} \sum_{0 \leq j < N} \chi_{K_s}(x_j) \min(|f(x_j)|, A) \text{vol}(Q_j) \\
 &\leq \inf_{\delta} \sup_{D \text{ } \delta\text{-fine}} \sum_{j, x_j \in K_s} \text{vol}(Q_j) |f(x_j)| \\
 (7.15) \quad &\leq \sup_{D \text{ } \delta_\varepsilon\text{-fine}} \sum_{j, x_j \in K_s} \left| \int_{Q_j} f(x) dx \right| + 2\varepsilon
 \end{aligned}$$

où δ_ε désigne une jauge ε -adaptée à f (pour cette dernière inégalité, on applique bien sûr encore le lemme de Henstock 1.1 (b)); on suppose de plus δ_ε choisie en sorte que $\delta_\varepsilon(x) < 1/s$ en tout point. Soit $D = \{(Q_j, x_j)\}_{0 \leq j < N}$ une subdivision δ_ε -fine arbitraire. On pose $Q_j = \prod_{1 \leq k \leq n} [a_{j,k}, b_{j,k}]$, $x_j = (x_{j,1}, \dots, x_{j,n})$, et on note J l'ensemble des indices $j = 0, 1, 2, \dots, N-1$ tels que $x_j \in K_s$. Enfin, on associe à chaque indice $j \in J$ l'élément $\alpha_j \in \{-1, 0, +1\}^n$ tel que

$$\alpha_{j,k} = \begin{cases} -1 & \text{si } x_{j,k} = a_{j,k}, \\ 0 & \text{si } x_{j,k} \in]a_{j,k}, b_{j,k}[, \\ +1 & \text{si } x_{j,k} = b_{j,k}. \end{cases}$$

Ceci fournit une partition $J = \bigcup J_\alpha$ en 3^n parties $J_\alpha = \{j \in J; \alpha_j = \alpha\}$. Fixons $\eta > 0$ assez petit. Pour $j \in J$, on a par définition $x_j \in K_s = \overline{E_s}$, donc on peut trouver un point $x'_j \in E_s$ tel que $d(x_j, x'_j) \leq \eta$. On définit d'autre part

$$Q'_j = \prod_{1 \leq k \leq n} [a'_{j,k}, b'_{j,k}]$$

avec

$$a'_{j,k} = \begin{cases} x'_{j,k} & \text{si } x_{j,k} = a_{j,k}, \\ a_{j,k} + \eta & \text{sinon,} \end{cases} \quad b'_{j,k} = \begin{cases} x'_{j,k} & \text{si } x_{j,k} = b_{j,k}, \\ b_{j,k} - \eta & \text{sinon,} \end{cases}$$

et on suppose de plus qu'on a pris η plus petit que la moitié $\frac{1}{2} \min(x_{j,k} - a_{j,k}, b_{j,k} - x_{j,k})$ du minimum des différences non nulles $x_{j,k} - a_{j,k}$ et $b_{j,k} - x_{j,k}$, et aussi plus petit que la moitié des distances des cubes Q_j, Q_ℓ non adjacents. Dans ces conditions, il est clair que $x'_j \in Q'_j$ et que les pavés Q'_j correspondant aux indices $j \in J_\alpha$ sont d'intérieurs disjoints pour chaque $\alpha \in \{-1, 0, +1\}^n$ fixé; en effet, il ne pourrait y avoir une intersection d'intérieur non vide que pour deux indices distincts $j, \ell \in J_\alpha$ tels que Q_j et Q_ℓ étaient adjacents relativement à une face $x_k = \text{Cte}$, pour laquelle on a par exemple $a'_{j,k} = x'_{j,k}$ et $b'_{\ell,k} = x'_{\ell,k}$ avec $x_{j,k} = a_{j,k} = x_{\ell,k} = b_{\ell,k}$, mais ceci est exclu

puisque alors $-1 = \alpha_{j,k} \neq \alpha_{\ell,k} = +1$. D'autre part, pour η assez petit, le théorème 1.3 (b) implique

$$(7.16) \quad \left| \int_{Q'_j} f(x) dx - \int_{Q_j} f(x) dx \right| \leq \varepsilon/N,$$

et comme $\text{diam } Q_j \leq \delta_\varepsilon(x_j) < 1/s$, on pourra aussi obtenir que $\text{diam } Q'_j < 1/s \leq \delta_0(x'_j)$ du fait que $x'_j \in E_s$. Ceci montre que $\{(Q'_j, x'_j)\}_{j \in J_\alpha}$ est une subdivision partielle δ_0 -fine de P . D'après (7.14) on en déduit

$$\sum_{j \in J_\alpha} \left| \int_{Q'_j} f(x) dx \right| \leq \sum_{j \in J_\alpha} \text{vol}(Q'_j) |f(x'_j)| + 2\varepsilon_0 \leq \text{vol}(P) s + 2\varepsilon_0,$$

puisque $x'_j \in E_s$ implique $|f(x'_j)| \leq s$. En prenant la somme sur $\alpha \in \{-1, 0, +1\}^n$, il vient d'après (7.16)

$$\sum_{j \in J} \left| \int_{Q_j} f(x) dx \right| \leq 3^n (\text{vol}(P) s + 2\varepsilon_0) + N\varepsilon/N,$$

et en combinant ceci avec (7.15) après être passé au sup sur toutes les subdivisions D δ_ε -fines on trouve

$$\int_P \chi_{K_s}(x) \min(|f(x)|, A) dx \leq 3^n (\text{vol}(P) s + 2\varepsilon_0) + 3\varepsilon.$$

On termine en appliquant le théorème de convergence monotone avec $A \rightarrow +\infty$. Comme ε peut être pris arbitrairement petit ceci donne

$$\int_P \chi_{K_s}(x) |f(x)| dx \leq 3^n (\text{vol}(P) s + 2\varepsilon_0) < +\infty.$$

Comme $\bigcup K_s = P$, on voit de plus que $\lim m(P \setminus K_s) = 0$.

(b) Il est clair par définition que U est ouvert. Pour tout pavé fermé borné $P' \subset P$ d'intérieur non vide, le théorème de Baire appliqué à $P' = \bigcup P' \cap K_s$ montre que l'une des parties $P' \cap K_s$ est d'intérieur non vide dans P' , donc $P' \cap U \neq \emptyset$. Ceci entraîne que U est un ouvert dense dans P . \square

On va voir, déjà dans le cas des fonctions d'une seule variable, que l'ouvert U du théorème 7.13 (b) peut être de mesure arbitrairement petite dans P . L'exemple suivant est dû à Jitan Lu et Peng Yee Lee [LL].

(7.17) Exemple. Soit $f_{a,b} : \mathbb{R} \rightarrow \mathbb{R}$ la fonction telle que

$$\begin{cases} f_{a,b}(x) = 0 & \text{si } x \in \mathbb{R} \setminus]a, b[, \\ f_{a,b}(x) = (x-a)^2(x-b)^2 \sin \frac{1}{(x-a)^2(x-b)^2} & \text{si } x \in]a, b[. \end{cases}$$

Il est facile de constater que $f_{a,b}$ est partout dérivable sur \mathbb{R} , mais que $f'_{a,b}$ n'est pas absolument intégrable au voisinage des points a ou b : le comportement est le même que

celui de la fonction $x \mapsto x^2 \sin(1/x^2)$ en 0, dont la dérivée $2x \sin(1/x^2) - (2/x) \cos(1/x^2)$ n'est pas absolument intégrable. On considère maintenant un ensemble de Cantor $K \subset [0, 1]$ de mesure $\prod_{k \geq 1} (1 - \varepsilon_k)$ en pratiquant des divisions itératives de $[0, 1]$ en trois intervalles de proportions respectives $(1 - \varepsilon_k)/2$, ε_k , $(1 - \varepsilon_k)/2$ et en omettant chaque fois l'intervalle central. On suppose ici $\sum \varepsilon_k \leq \varepsilon$, de sorte que $m(K) \geq 1 - \varepsilon$. Le complémentaire $[0, 1] \setminus K$ est formé d'une réunion dénombrable d'intervalles ouverts $]a_j, b_j[$. On pose $f(x) = 0$ sur K et $f(x) = f_{a_j, b_j}(x)$ sur $]a_j, b_j[$. Il est facile de voir que f est partout dérivable sur $[0, 1]$ avec une dérivée nulle sur K . En effet, la dérivabilité est évidente sur chaque intervalle $]a_j, b_j[$. Montrons que $f'(x_0) = 0$ en tout point $x_0 \in K$. Si $x \in]a_j, b_j[$ avec par exemple $a_j \geq x_0$, on a

$$|f(x)| \leq (b_j - a_j)^2 (x - a_j)^2 \leq (x - a_j)^2 \quad \text{tandis que } x - x_0 \geq x - a_j,$$

d'où $|f(x)/(x - x_0)| \leq |x - a_j| \leq |x - x_0|$. Le raisonnement est le même si $x \in]a_j, b_j[$ avec $b_j \leq x_0$, et si $x \in K$ on a $f(x) = 0$ par définition. En particulier f' est intégrable au sens de Henstock sur $[0, 1]$, mais ne peut être intégrable au sens de Lebesgue au voisinage d'aucun point de K car les points a_j, b_j s'y accumulent.

8. Densité des fonctions continues à support compact

Si $A \subset \mathbb{R}^n$ est une partie mesurable quelconque, on définit l'intégrale $\int_A f(x) dx$ d'une fonction $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ en considérant tout simplement l'intégrale $\int_{\mathbb{R}^n} \tilde{f}(x) dx$ du prolongement \tilde{f} de f à \mathbb{R}^n tel que $\tilde{f}(x) = 0$ sur $\mathbb{R}^n \setminus A$. Ceci permet de parler de l'intégrabilité ou de la mesurabilité de f ; on obtient ainsi un espace de Banach $L^1(A)$ en adaptant les résultats des sections 4 et 5 au cas des fonctions sur le pavé infini $P = \mathbb{R}^n$, nulles sur $\mathbb{R}^n \setminus A$.

Nous démontrons maintenant un résultat très important, à savoir la densité des fonctions continues à support compact dans l'espace de Banach $L^1(\Omega)$ pour tout ouvert Ω de \mathbb{R}^n .

(8.1) Densité des fonctions continues à support compact.

- (a) Soit $E \subset \mathbb{R}^n$ un ensemble mesurable. Pour tout $\varepsilon > 0$, il existe une application continue $g : \mathbb{R}^n \rightarrow [0, 1]$ telle que $g(x) = \chi_E(x)$ hors d'un ensemble ouvert V de mesure $m(V) \leq \varepsilon$.
- (b) Soit $f : \Omega \rightarrow \mathbb{R}$ une fonction mesurable sur un ouvert $\Omega \subset \mathbb{R}^n$. Il existe une suite $g_k : \mathbb{R}^n \rightarrow \mathbb{R}$ d'applications continues à support compact dans Ω telles que $g_k \rightarrow f$ presque partout.
- (c) Soit $f \in L^1(\Omega)$. Il existe une suite $g_k : \Omega \rightarrow \mathbb{R}$ d'applications continues à support compact dans Ω telles que $g_k \rightarrow f$ presque partout et $\|g_k - f\|_1 \rightarrow 0$.
- (d) L'espace $\mathcal{C}_c(\Omega)$ des fonctions continues à support compact dans Ω est dense dans $L^1(\Omega)$, et $L^1(\Omega)$ s'identifie au complété de $\mathcal{C}_c(\Omega)$ pour la norme $\|\cdot\|_1$.

Démonstration. (a) Grâce à 7.4, choisissons F un ensemble fermé et U un ensemble ouvert tels que $F \subset E \subset U$ et $m(U \setminus F) \leq \varepsilon$. On pose

$$g(x) = \frac{d(x, \mathbb{C}U)}{d(x, F) + d(x, \mathbb{C}U)}.$$

Il est clair que $0 \leq g(x) \leq 1$ et que g est continue (le dénominateur ne s'annule pas puisque F et $\complement U$ sont des fermés disjoints). De plus $g(x) = 1$ si $x \in F$ et $g(x) = 0$ si $x \in \complement U$, donc g répond à la question en posant $V = U \setminus F$.

(b) et (c). On se ramène d'abord au cas où $f \geq 0$: si le résultat est démontré dans ce cas, on écrit $f = f_+ - f_-$ et

$$f_+ = \lim g'_k, \quad f_- = \lim g''_k, \quad f = \lim g'_k - g''_k \quad \text{presque partout}$$

avec g'_k et g''_k continues à support compact. On supposera donc dans la suite que $f \geq 0$. Dans ce cas, la proposition 7.9 fournit une suite croissante de fonctions étagées $f_k = \sum_{1 \leq \ell \leq N_k} c_{k,\ell} \chi_{E_{k,\ell}}$ telles que $f = \lim f_k$ en tout point. Quitte à perdre éventuellement la convergence sur le bord $\partial\Omega$ (négligeable), on peut supposer que $E_{k,\ell} \subset \Omega_k$ pour une certaine suite strictement croissante d'ouverts relativement compacts $\Omega_k \subset \Omega$ tels que $\Omega = \bigcup \Omega_k$ (sinon on remplace tout simplement $E_{k,\ell}$ par $E_{k,\ell} \cap \Omega_k$). D'après (a) [avec $U \subset \Omega_k$, on peut trouver une fonction continue $g_{k,\ell} : \Omega \rightarrow \mathbb{R}$ telle que $g_{k,\ell} = \chi_{E_{k,\ell}}$ hors d'un ensemble ouvert $V_{k,\ell}$ de mesure $\leq 2^{-k}/N_k$ et à support dans $\overline{\Omega}_k$. Par suite $g_k = \sum_{1 \leq \ell \leq N_k} c_{k,\ell} g_{k,\ell}$ est continue à support compact dans $\overline{\Omega}_k \subset \Omega$ et coïncide avec f_k hors de $V_k = \bigcup_{\ell} V_{k,\ell}$, $m(V_k) \leq 2^{-k}$. En dehors de $W_s = \bigcup_{k > s} V_k$ qui est de mesure $\leq 2^{-s}$, il est clair que $g_k \rightarrow f$ puisque $g_k = f_k$ pour $k > s$, et par conséquent $g_k \rightarrow f$ hors de l'ensemble négligeable $N = \bigcap_s W_s$. Ceci démontre (b). Si en outre f est intégrable, alors le théorème de convergence dominée appliqué à la suite décroissante $f - f_k \rightarrow 0$ dominée par f montre que $\lim \|f_k - f\|_1 = 0$. En prenant de plus $m(V_{k,\ell}) \leq 2^{-k}/(1 + N_k |c_{k,\ell}|)$, il vient $\|g_{k,\ell} - \chi_{E_{k,\ell}}\|_1 \leq 2^{-k}/(1 + N_k |c_{k,\ell}|)$, donc $\|g_k - f_k\|_1 \leq 2^{-k}$ et (c) s'ensuit.

(d) résulte de (c), puisque nous savons que $L^1(\Omega)$ est complet, et de plus $\mathcal{C}_c(\Omega)$ s'identifie à un sous-espace de $L^1(\Omega)$ d'après 2.8 (a) (ou, du moins, son analogue en plusieurs variables, qui se démontre de la même manière). \square

Une première conséquence du théorème de densité des fonctions continues est la caractérisation suivante des fonctions mesurables.

(8.2) Caractérisation des fonctions mesurables (théorème de Lusin).

Soit $A \subset \mathbb{R}^n$ une partie mesurable et $f : A \rightarrow \overline{\mathbb{R}}$ une fonction quelconque. Il y a équivalence entre :

- (a) f est mesurable.
- (b) Pour tout $\varepsilon > 0$, il existe une partie ouverte $V \subset \mathbb{R}^n$ de mesure $m(V) \leq \varepsilon$ telle que la restriction $f|_{A \setminus V}$ soit continue.

Démonstration. Commençons par le sens « facile ».

(b) \Rightarrow (a). Pour k entier > 0 , fixons un ouvert V_k tel que $m(V_k) \leq 2^{-k}$ et $f|_{A \setminus V_k}$ continue. Quitte à remplacer V_k par $V'_k = \bigcup_{\ell > k} V_\ell$, on peut supposer que la suite V_k est décroissante. Posons alors $f_k(x) = 0$ si $x \in V_k$ et $f_k(x) = f(x)$ si $x \in A \setminus V_k$. Soit U un ouvert de $\overline{\mathbb{R}}$. Alors

$$\begin{aligned} f_k^{-1}(U) &= f^{-1}(U) \cap (A \setminus V_k) && \text{si } 0 \notin U, \\ f_k^{-1}(U) &= V_k \cup (f^{-1}(U) \cap (A \setminus V_k)) && \text{si } 0 \in U. \end{aligned}$$

Dans les deux cas $f_k^{-1}(U)$ est mesurable puisque $f^{-1}(U) \cap (A \setminus V_k)$ est une partie ouverte de $A \setminus V_k$ (et donc l'intersection d'un ouvert avec la partie mesurable $A \setminus V_k$). Ceci prouve que f_k est mesurable. Comme $f = \lim f_k$ hors de la partie négligeable $N = \bigcap V_k$, on voit que f est elle aussi mesurable.

(a) \Rightarrow (b). Quitte à étendre f par 0 sur $\mathbb{R}^n \setminus A$, on peut supposer $A = \mathbb{R}^n$. D'après 8.1 (b), il existe une suite d'applications continues $g_k : \mathbb{R}^n \rightarrow \mathbb{R}$ telles que $g_k \rightarrow f$ presque partout. Comme $\overline{\mathbb{R}}$ est homéomorphe à $[-1, 1]$ par l'application $x \mapsto x/(1+|x|)$, on peut tout aussi bien supposer que f et les g_k sont à valeurs dans $[-1, 1]$. Soit $N \subset \mathbb{R}^n$ un ensemble négligeable tel que $g_k(x)$ converge vers $f(x)$ pour tout $x \in \mathbb{R}^n \setminus N$. Soit $(P_s)_{s \geq 0}$ une suite croissante de pavés fermés bornés tels que $\bigcup P_s = \mathbb{R}^n$. Pour chaque paire d'entiers r, s , considérons

$$U_{r,s} = \{x \in P_s; \exists k, \ell \geq r, |g_k(x) - g_\ell(x)| > 2^{-s}\}.$$

C'est une partie ouverte de P_s , de plus la suite $(U_{r,s})_{r \geq 0}$ est décroissante et $\bigcap_{r \geq 0} U_{r,s} \subset N$ d'après le critère de Cauchy. Comme N est négligeable, nous avons $\lim_{r \rightarrow +\infty} m(U_{r,s}) \leq m(N) = 0$, donc il existe un indice $r(s)$ tel que $m(U_{r(s),s}) \leq 2^{-s}$. Il n'est pas restrictif de supposer que l'on choisit $r(s+1) > r(s)$ pour tout s . Soit V_q une partie ouverte de P contenant $N \cup \bigcup_{s \geq q} U_{r(s),s}$, telle que $m(V_q) \leq 2^{2-q}$. Pour $k \geq s \geq q$ et $x \in P_q \setminus V_q \subset P_s \setminus U_{r(s),s}$, nous avons $|g_{r(s)}(x) - g_{r(s+1)}(x)| \leq 2^{-s}$. Ceci entraîne que f est la limite uniforme des $g_{r(s)}$ sur toute partie compacte de $\mathbb{R}^n \setminus V_q$, donc $f|_{\mathbb{R}^n \setminus V_q}$ est continue. \square

Nous pouvons maintenant élucider le lien entre l'intégrabilité au sens de Kurzweil-Henstock et l'intégrabilité au sens de Lebesgue.

(8.3) Définition. Soit $P \subset \mathbb{R}^n$ un pavé. On dit que $f : P \rightarrow \mathbb{R}$ est intégrable au sens de Lebesgue si f est mesurable et si $|f|$ est KH-intégrable.

(8.4) Théorème. Soit $f : P \rightarrow \mathbb{R}$. Les propriétés suivantes sont équivalentes :

- (a) f est intégrable au sens de Lebesgue;
- (b) f est absolument KH-intégrable;
- (c) f est la limite presque partout d'une suite $(g_n)_{n \in \mathbb{N}}$ de fonctions continues vérifiant $\sup_{n \in \mathbb{N}} \|g\|_{L^1} < +\infty$.

Démonstration. L'implication (a) \Rightarrow (b) résulte de la Proposition 7.10 appliquée avec $g = -|f|$ et $h = |f|$, tandis que (b) \Rightarrow (c) est une conséquence de 8.1 (c). Enfin, sous l'hypothèse (c), le Corollaire 7.6 montre que f est mesurable, et le lemme de Fatou implique que $|f|$ est KH-intégrable avec

$$\int_P |f(x)| dx = \int_P \liminf_{n \rightarrow +\infty} |g_n(x)| dx \leq \liminf_{n \rightarrow +\infty} \int_P |g_n(x)| dx < +\infty.$$

Les théories de Kurzweil-Henstock et de Lebesgue se rejoignent donc complètement dans le cas des fonctions absolument intégrables. \square

(8.5) Remarque. Supposons $f : P \rightarrow \mathbb{R}$ KH-intégrable mais non Lebesgue intégrable, i.e. non absolument KH-intégrable. Alors on a

$$\lim_{n \rightarrow +\infty} \int_P \min(|f(x)|, n) dx = +\infty,$$

sinon le théorème de convergence monotone entraînerait que $|f|$ est KH-intégrable. Dans ce cas on conviendra d'écrire que

$$\int_P |f(x)| dx = +\infty.$$

On adoptera la même convention pour $\|f\|$ si $f : P \rightarrow E$ est KH-intégrable à valeurs dans un espace de Banach E avec $\|f\|$ mesurable et non KH-intégrable (on notera d'ailleurs que $\|f\|$ est automatiquement mesurable si le dual E^* est séparable car étant donné une suite φ_n de formes linéaires dense dans la boule unité de E^* on a $\|f\| = \sup_{n \in \mathbb{N}} \varphi_n \circ f$ et les $\varphi_n \circ f$ sont KH-intégrables). \square

Nous utilisons maintenant la densité des fonctions continues pour étendre les théorèmes fondamentaux déjà démontrés dans ce cadre. Pour cela, on utilise le fait que toute fonction $f \in L^1(\Omega)$ absolument intégrable sur un ouvert $\Omega \subset \mathbb{R}^n$ est limite en norme L^1 d'une suite de fonctions continues à support compact (g_k) telles que (disons) $\|g_k - g_{k-1}\|_1 \leq 2^{-k}$. On peut donc représenter f comme somme d'une série L^1 convergente $f = \sum u_k$ de fonctions continues à support compact, avec $u_0 = g_0$ et $u_k = g_k - g_{k-1}$. Enfin, on peut écrire

$$(8.6) \quad f = f' - f'' \quad \text{avec} \quad f' = \sum_k u_k^+, \quad f'' = \sum_k u_k^-, \quad \|u_k^+\|_1 + \|u_k^-\|_1 \leq 2^{-k} \quad \text{si } k \geq 1,$$

de sorte que f apparait comme une différence de limites croissantes L^1 convergentes de fonctions positives continues à support compact. L'égalité (8.6) est vraie seulement presque partout.

(8.7) Formule du changement de variable dans \mathbb{R}^n . Soit $\varphi : \Omega \rightarrow \Omega'$ une bijection de classe C^1 entre deux ouverts de \mathbb{R}^n , telle que le déterminant jacobien

$$\text{Jac } \varphi(x) = \det \left(\frac{\partial \varphi_i}{\partial x_j} \right)_{1 \leq i, j \leq n}$$

soit partout non nul. Alors pour toute fonction $f : \varphi(\Omega) = \Omega' \rightarrow \mathbb{R}$ dans $L^1(\Omega')$ on a

$$\boxed{\int_{\varphi(\Omega)} f(y) dy = \int_{\Omega} f(\varphi(x)) |\text{Jac } \varphi(x)| dx}$$

et l'intégrale du membre de droite est elle aussi absolument convergente sur Ω .

Démonstration. La première chose à voir (f étant peut-être seulement définie presque partout) est que l'image inverse $\varphi^{-1}(N)$ d'un ensemble négligeable dans Ω' est négligeable dans Ω . Or ceci résulte du lemme suivant et du fait que le difféomorphisme φ est une application localement lipschitzienne ainsi que son inverse φ^{-1} .

(8.8) Lemme. Soit $\psi : \Omega' \rightarrow \Omega$ une application localement lipschitzienne dans Ω' . Alors l'image $\psi(N)$ de tout ensemble négligeable N de Ω' est négligeable dans Ω .

En effet, quitte à écrire $\psi(N) = \bigcup \psi(N \cap K_j)$ pour une certaine suite croissante de parties compactes $K_j \subset \Omega'$, on peut supposer $N \subset K$ compact $\subset \Omega'$ et ψ lipschitzienne de rapport k . Pour tout ε , on peut alors recouvrir N par une famille dénombrable de cubes Q_α de diamètre d_α , tels que $\sum \text{vol}(Q_\alpha) = \sum d_\alpha^n \leq \varepsilon$. Comme $\psi(Q_\alpha)$ est contenu dans un cube Q'_α de diamètre $d'_\alpha \leq kd_\alpha$, on en déduit que $m(\psi(N)) \leq k^n \varepsilon$ pour tout $\varepsilon > 0$, ce qui montre que $\psi(N)$ est négligeable. \square

Pour terminer la démonstration, on utilise maintenant (8.6) et le fait que la formule est vraie pour u_k^+ et u_k^- d'après le théorème 3.6. On en déduit que la formule est vraie pour f' et f'' par le théorème de convergence monotone, puis qu'elle est vraie pour f par différence.⁽²⁰⁾ \square

(8.9) Théorème de Fubini. Soient $A \subset \mathbb{R}^n$ et $B \subset \mathbb{R}^m$ des parties mesurables et $f : A \times B \rightarrow \mathbb{R}$ une fonction absolument intégrable. Alors, en notant

$$(x, y) = (x_1, \dots, x_n; y_1, \dots, y_m) \in \mathbb{R}^n \times \mathbb{R}^m$$

les variables dans $A \times B$, on a

$$\boxed{\int_{A \times B} f(x, y) \, dx \, dy = \int_A \left(\int_B f(x, y) \, dy \right) dx.}$$

Ceci signifie implicitement :

- (a) que pour presque tout $x \in A$, la fonction $y \mapsto f(x, y)$ est dans $L^1(B)$.
- (b) que la fonction presque partout définie $x \mapsto \int_{y \in B} f(x, y) \, dy$ est dans $L^1(A)$.

[et on peut bien entendu échanger les rôles de A et B dans ce qui précède].

Démonstration. Quitte à étendre f par 0 hors de $A \times B$, on peut supposer que $A = \mathbb{R}^n$ et $B = \mathbb{R}^m$.

(1) On commence par démontrer le résultat lorsque $f = \lim g_k$ est la limite d'une suite croissante de fonctions continues $g_k \geq 0$ à support compact dans $\mathbb{R}^n \times \mathbb{R}^m$. Comme le résultat est vrai pour les g_k d'après le théorème 1.12, les fonctions

$$G_k(x) = \int_{y \in \mathbb{R}^m} g_k(x, y) \, dy$$

fournissent une suite croissante de fonctions continues à support compact dans \mathbb{R}^n telles

(20) On observera que si les résultats démontrés au chapitre II avaient été traités *après la preuve des théorèmes fondamentaux de convergence*, on aurait pu en simplifier notablement la démonstration : en effet, comme on vient de le voir, le résultat local valable a priori pour les fonctions continues à support compact dans l'image d'un petit pavé $\Omega'_P = \varphi(P^\circ)$ s'étend aux fonctions $f \in L^1(\Omega'_P)$ quelconques, et de là à un ouvert quelconque en prenant une partition dénombrable de Ω par des pavés. L'argument de partition continue de l'unité devient inutile, de même que le préambule destiné à introduire ce qu'est l'intégrale sur un ouvert – sachant que la fonction caractéristique d'un ouvert est une fonction mesurable.

que

$$\int_{\mathbb{R}^n} G_k(x) dx = \int_{\mathbb{R}^n \times \mathbb{R}^m} g_k(x, y) dx dy \leq \int_{\mathbb{R}^n \times \mathbb{R}^m} f(x, y) dx dy < +\infty.$$

D'après le théorème de convergence monotone 4.5, on voit que

$$F(x) := \lim_{k \rightarrow +\infty} G_k(x) = \lim_{k \rightarrow +\infty} \int_{\mathbb{R}^m} g_k(x, y) dy = \int_{\mathbb{R}^m} f(x, y) dy$$

est finie presque partout, puisque

$$\begin{aligned} \int_{\mathbb{R}^n} F(x) dx &= \lim_{k \rightarrow +\infty} \int_{\mathbb{R}^n} G_k(x) dx = \lim_{k \rightarrow +\infty} \int_{\mathbb{R}^n \times \mathbb{R}^m} g_k(x, y) dx dy \\ &= \int_{\mathbb{R}^n \times \mathbb{R}^m} f(x, y) dx dy < +\infty. \end{aligned}$$

Ceci démontre à la fois l'égalité cherchée et les observations complémentaires (a) et (b).

(2) Il résulte de (1) que la formule est vraie si $f = \chi_U$ est la fonction caractéristique d'un ouvert intégrable, puisqu'une telle fonction est la limite d'une suite croissante de fonctions continues à support compact dans U . Plus généralement, la formule est vraie si $f = \chi_A$ où $A = \bigcap U_k$ est l'intersection d'une suite décroissante d'ouverts intégrables (c'est-à-dire un G_δ intégrable) : il suffit pour cela d'appliquer le théorème de convergence monotone à la suite décroissante χ_{U_k} .

(3) Si $f = \chi_E$ est la fonction caractéristique d'un ensemble négligeable, les tranches $E \cap (\{x\} \times \mathbb{R}^m)$ sont négligeables pour presque $x \in \mathbb{R}$ et la formule est vraie (avec des intégrales presque toutes égales à 0). Pour le voir, on utilise (2) et 7.3 (a), qui implique l'existence d'un G_δ négligeable G tel que $E \subset G$.

(4) Comme conséquence de (3), le fait que f soit modifiée sur un ensemble négligeable de $\mathbb{R}^n \times \mathbb{R}^m$ ne change rien, puisque les intégrales $\int_{\mathbb{R}^m} f(x, y) dy$ ne sont alors modifiées que sur une partie négligeable de valeurs de $x \in \mathbb{R}^n$. On peut donc supposer f seulement définie presque partout dans $\mathbb{R}^n \times \mathbb{R}^m$. Cette dernière observation, combinée à (8.6) et à la partie (1) conclut la démonstration. \square

(8.10) Théorème de Tonelli. Soient $A \subset \mathbb{R}^n$ et $B \subset \mathbb{R}^m$ des parties mesurables et $f : A \times B \rightarrow \mathbb{R}$ une fonction mesurable. On suppose que

- (a) pour presque tout $x \in A$, la fonction $y \mapsto |f(x, y)|$ est intégrable sur B ;
- (b) l'intégrale itérée $\int_A \left(\int_B |f(x, y)| dy \right) dx$ est convergente.

Alors f est dans $L^1(A \times B)$ et on a

$$\boxed{\int_{A \times B} f(x, y) dx dy = \int_A \left(\int_B f(x, y) dy \right) dx.}$$

Démonstration. Quitte à écrire $f = f_+ - f_-$ où $f_+ = \max(f, 0)$, $f_- = \max(-f, 0)$, on voit que f_+ et f_- sont mesurables d'après la proposition 7.5 et on se ramène immédiatement au cas où $f \geq 0$.

On écrit maintenant $f = \lim f_k$ comme limite croissante des fonctions $f_k \geq 0$ telles que

$$f_k(x, y) = \min(f(x, y), k \chi_{[-k, k]^{n+m}}(x, y)).$$

Ces fonctions sont intégrables sur $A \times B$ puisque la fonction caractéristique $\chi_{[-k, k]^{n+m}}$ du pavé $[-k, k]^{n+m}$ l'est sur $\mathbb{R}^n \times \mathbb{R}^m$ tout entier). Le théorème de Fubini 8.9 donne

$$\int_{A \times B} f_k(x, y) dx dy = \int_A \left(\int_B f_k(x, y) dy \right) dx \leq \int_A \left(\int_B f(x, y) dy \right) dx < +\infty.$$

On en déduit l'intégrabilité de f sur $A \times B$ à l'aide du théorème de convergence monotone 4.5, et on applique maintenant de nouveau 8.9 pour obtenir l'égalité des intégrales. \square

D'après les théorèmes de Fubini 8.9 et de Tonelli 8.10, dans le cas particulier où $f = \chi_E$ est la fonction caractéristique d'un ensemble mesurable E dans $\mathbb{R}^n = A \times B = \mathbb{R}^{n-1} \times \mathbb{R}$, on déduit la formule suivante permettant le calcul d'aires et de volumes par récurrence sur la dimension.

(8.11) Calcul d'aires et de volumes par récurrence sur la dimension. Soit E une partie mesurable de \mathbb{R}^n . On note $a_E(x_n)$ l'aire (ou plus précisément la mesure $(n-1)$ -dimensionnelle) de la projection dans \mathbb{R}^{n-1} de la tranche $E \cap (\mathbb{R}^{n-1} \times \{x_n\})$. Alors cette tranche $E \cap (\mathbb{R}^{n-1} \times \{x_n\})$ est \mathbb{R}^{n-1} -mesurable pour presque tout $x_n \in \mathbb{R}$, la fonction $x_n \mapsto a_E(x_n)$ est mesurable sur \mathbb{R} et on a l'égalité

$$\text{vol}(E) = \int_{\mathbb{R}} a_E(x_n) dx_n$$

(que les intégrales soient finies ou non).

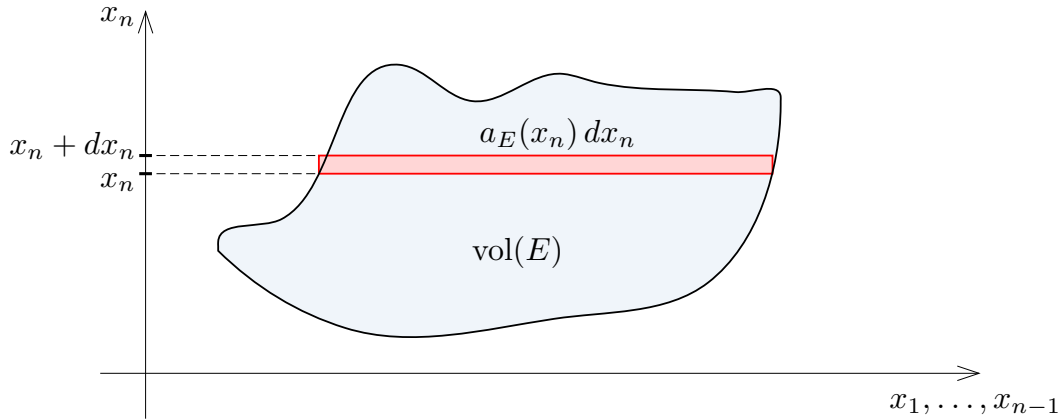


Fig. 24. Calcul de volume par tranchage suivant la direction x_n .

Démonstration. On se ramène au cas d'un ensemble intégrable en remplaçant E par $E_k = E \cap [-k, k]^n$, et dans cette situation la formule résulte de l'application du théorème de Fubini à $f_k = \chi_{E_k}$. Le cas général s'obtient grâce au théorème de convergence monotone 4.5 appliqué aux suites croissantes $\chi_{E_k} \rightarrow \chi_E$, et, pour presque tout $x_n \in \mathbb{R}$, $a_{E_k}(x_n) \rightarrow a_E(x_n)$. \square

9. Variation totale, fonctions à variation bornée

Nous commençons par une définition de la variation totale d'une fonction sur un intervalle, dans le sens de Kurzweil-Henstock.

(9.1) Définition. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction réelle et $Z \subset [a, b]$ un sous-ensemble quelconque. On définit la variation totale de f sur Z comme étant la *limsup*

$$\begin{aligned} \text{Var}(f, Z) &= \limsup_{\text{KH}, D} \sum_{j, x_j \in Z} |f(a_{j+1}) - f(a_j)| \\ &:= \inf_{\delta > 0} \sup_{D \text{ } \delta\text{-fine}} \sum_{0 \leq j < N, x_j \in Z} |f(a_{j+1}) - f(a_j)| \end{aligned}$$

étendue à toutes les subdivisions pointées $D = \{([a_j, a_{j+1}], x_j)\}_{0 \leq j < N}$. On a en général $\text{Var}(f, Z) \in [0, +\infty]$. La fonction f est dite à variation bornée si sa variation totale sur $[a, b]$ est finie, c'est-à-dire

$$\text{Var}(f, [a, b]) < +\infty.$$

Indiquons d'abord quelques résultats élémentaires concernant la variation totale.

(9.2) Proposition. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction réelle.

(a) Pour toutes parties $Z_1, Z_2 \subset [a, b]$, on a

$$\text{Var}(f, Z_1 \cup Z_2) \leq \text{Var}(f, Z_1) + \text{Var}(f, Z_2).$$

(b) Si Z est inclus dans un sous-intervalle $[c, d] \subset [a, b]$, on a

$$\begin{aligned} \text{Var}(f, Z) &= \text{Var}(f|_{[c, d]}, Z) \\ &\quad + \chi_{Z \cap]a, b]}(c) \limsup_{x \rightarrow c-0} |f(x) - f(c)| + \chi_{Z \cap [a, b[}(d) \limsup_{x \rightarrow d+0} |f(x) - f(d)|, \end{aligned}$$

de sorte que $\text{Var}(f, Z) = \text{Var}(f|_{[c, d]}, Z)$ pour f continue à gauche en c et à droite en d .

(c) Si $Z_1 \subset [a, c]$ et $Z_2 \subset [c, b]$ avec $c \in]a, b[$, alors

$$\text{Var}(f, Z_1 \cup Z_2) = \text{Var}(f|_{[a, c]}, Z_1) + \text{Var}(f|_{[c, b]}, Z_2).$$

en particulier

$$\text{Var}(f, [a, b]) = \text{Var}(f|_{[a, c]}, [a, c]) + \text{Var}(f|_{[c, b]}, [c, b]).$$

Démonstration. (a) résulte immédiatement du fait que

$$\sum_{\substack{0 \leq j < N \\ x_j \in Z_1 \cup Z_2}} |f(a_{j+1}) - f(a_j)| \leq \sum_{\substack{0 \leq j < N \\ x_j \in Z_1}} |f(a_{j+1}) - f(a_j)| + \sum_{\substack{0 \leq j < N \\ x_j \in Z_2}} |f(a_{j+1}) - f(a_j)|.$$

(b) Il suffit de contraindre la subdivision D à passer par les points intermédiaires $x_j = c$ et $x_k = d$ en prenant par exemple $\delta(x) < \min(|x - c|, |x - d|)$ pour $x \neq c, d$. Toute subdivision pointée δ -fine D de $[a, b]$ fait alors intervenir un certain intervalle pointé $([a_j, c], c)$ après subdivision éventuelle de $([a_j, a_{j+1}], c)$ en $([a_j, c], c)$ et $([c, a_{j+1}], c)$, ce qui rajoute à la limsup des sommes de Riemann étendues aux points $a_j \in Z$ un terme $\limsup_{x \rightarrow c-0} |f(x) - f(c)|$ si $c \in Z$ et $c > a$. Raisonnement identique pour $x_k = d$.

(c) Ici encore il suffit de forcer la subdivision D à contenir le point intermédiaire $x_j = c$ en prenant $\delta(x) < |x - c|$ pour $x \neq c$, et de subdiviser si nécessaire le sous-intervalle contenant $x_j = c$ pour avoir deux intervalles pointés $([a_{j-1}, c], c)$ et $([c, a_{j+1}], c)$. \square

(9.3) Proposition. *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction KH-intégrable. Alors la variation totale de l'intégrale indéfinie*

$$F(x) = \int_a^x f(t) dt, \quad x \in [a, b]$$

sur toute partie mesurable $Z \subset [a, b]$ est donnée par

$$\text{Var}(F, Z) = \int_a^b \chi_Z(x) |f(x)| dx = \int_Z |f(x)| dx \in [0, +\infty],$$

et elle est finie si et seulement si $\chi_Z |f|$ est KH-intégrable.

Démonstration. Fixons $\varepsilon > 0$ et une jauge δ_ε qui soit ε -adaptée pour f . D'après le lemme de Henstock, si $\delta \leq \delta_\varepsilon$ et si $D = \{([a_j, a_{j+1}], x_j)\}_{0 \leq j < N}$ est une subdivision pointée δ -fine, on a

$$\sum_{0 \leq j < N, x_j \in Z} \left| (a_{j+1} - a_j) f(x_j) - \int_{a_j}^{a_{j+1}} f(x) dx \right| \leq 2\varepsilon,$$

c'est-à-dire

$$\sum_{0 \leq j < N} \chi_Z(x_j) |(a_{j+1} - a_j) f(x_j) - (F(a_{j+1}) - F(a_j))| \leq 2\varepsilon.$$

On obtient par conséquent

$$\left| \sum_{0 \leq j < N, x_j \in Z} |F(a_{j+1}) - F(a_j)| - \sum_{0 \leq j < N} (a_{j+1} - a_j) \chi_Z(x_j) |f(x_j)| \right| \leq 2\varepsilon.$$

Si $\chi_Z |f|$ est KH-intégrable, il existe une jauge $\tilde{\delta}_\varepsilon$ telle que

$$\delta \leq \tilde{\delta}_\varepsilon \text{ et } D \text{ } \delta\text{-fine} \implies \left| \sum_{0 \leq j < N} (a_{j+1} - a_j) \chi_Z(x_j) |f(x_j)| - \int_Z |f(x)| dx \right| \leq \varepsilon,$$

par conséquent

$$\delta \leq \min(\delta_\varepsilon, \tilde{\delta}_\varepsilon) \text{ et } D \text{ } \delta\text{-fine} \implies \left| \sum_{0 \leq j < N, x_j \in Z} |F(a_{j+1}) - F(a_j)| - \int_Z |f(x)| dx \right| \leq 3\varepsilon.$$

Ceci implique bien par définition que

$$\text{Var}(F, Z) = \int_Z |f(x)| dx.$$

Si $\chi_Z|f|$ n'est pas KH-intégrable, sa mesurabilité et la Remarque 7.15 montrent qu'il existe une jauge $\tilde{\delta}_\varepsilon$ telle que

$$\delta \leq \tilde{\delta}_\varepsilon \text{ et } D \text{ } \delta\text{-fine} \implies \sum_{0 \leq j < N} (a_{j+1} - a_j) \chi_Z(x_j) |f(x_j)| \geq \frac{1}{\varepsilon}.$$

Les inégalités précédentes entraînent alors que $\text{Var}(F, Z) = +\infty$. □

(9.4) Corollaire. *Si F est l'intégrale indéfinie d'une fonction $f : [a, b] \rightarrow \mathbb{R}$ KH-intégrable, alors $\text{Var}(F, Z) = 0$ sur tout ensemble négligeable $Z \subset [a, b]$.*

Nous pouvons maintenant donner une caractérisation nécessaire et suffisante des fonctions qui sont les « primitives » de fonctions KH-intégrables.

(9.5) Théorème. *Soit $F : [a, b] \rightarrow \mathbb{R}$ une fonction. Il y a équivalence entre*

(a) *Il existe une fonction KH-intégrable $f : [a, b] \rightarrow \mathbb{R}$ et une constante $C \in \mathbb{R}$ telles que*

$$F(x) = \int_a^x f(t) dt + C \quad \text{sur } [a, b].$$

(b) *F est continue sur $[a, b]$, différentiable presque partout sur $[a, b]$, et si $Z \subset [a, b]$ est l'ensemble négligeable où F n'est pas différentiable, alors $\text{Var}(F, Z) = 0$.*

(c) *F est continue sur $[a, b]$, différentiable presque partout sur $[a, b]$, et $\text{Var}(F, Z) = 0$ pour tout ensemble négligeable $Z \subset [a, b]$.*

Démonstration. Les implications (a) \Rightarrow (b) et (a) \Rightarrow (c) résultent du Corollaire 6.5 et du Corollaire 9.4. L'implication (c) \Rightarrow (b) étant évidente, il reste à montrer que (b) \Rightarrow (a).

Supposons donc que la propriété (b) soit vraie, et définissons $f = F'$ sur $[a, b] \setminus Z$ et $f = 0$ sur Z . On va montrer que f est KH-intégrable et que

$$(9.6) \quad \int_a^b f(x) dx = F(b) - F(a).$$

En remplaçant b par un point intermédiaire quelconque de $[a, b]$, ceci démontrera que F est l'intégrale indéfinie de f à une constante près, et donc que la propriété (a) est satisfaite avec $C = F(a)$. Pour vérifier (9.6), on procède comme dans la démonstration du théorème fondamental de l'analyse. Pour cela, on commence par choisir une jauge δ_ε telle que

$$|y - x| \leq \delta_\varepsilon(x) \implies |F(y) - F(x) - (y - x)F'(x)| \leq \varepsilon|y - x|$$

pour $x \in [a, b] \setminus Z$, en utilisant l'hypothèse de différentiabilité de F presque partout. Pour $\delta \leq \delta_\varepsilon$ et pour une subdivision $D = \{([a_j, a_{j+1}], x_j)\}_{0 \leq j < N}$ δ -fine, on en déduit alors que

$$(9.7') \quad x_j \in [a, b] \setminus Z \implies |F(a_{j+1}) - F(a_j) - (a_{j+1} - a_j)F'(x_j)| \leq \varepsilon |a_{j+1} - a_j|.$$

L'hypothèse $\text{Var}(F, Z) = 0$ entraîne par ailleurs l'existence d'une jauge $\tilde{\delta}_\varepsilon$ telle que pour $\delta \leq \tilde{\delta}_\varepsilon$ et D δ -fine on ait

$$(9.7'') \quad \sum_{0 \leq j < N, x_j \in Z} |F(a_{j+1}) - F(a_j) - (a_{j+1} - a_j)f(x_j)| = \sum_{0 \leq j < N, x_j \in Z} |F(a_{j+1}) - F(a_j)| \leq \varepsilon,$$

(ceci en vertu du choix $f(x_j) = 0$ si $x_j \in Z$). En prenant $\delta \leq \min(\delta_\varepsilon, \tilde{\delta}_\varepsilon)$ et en ajoutant (9.7') et (9.7'') on trouve

$$\sum_{0 \leq j < N} |F(a_{j+1}) - F(a_j) - (a_{j+1} - a_j)f(x_j)| \leq \varepsilon(b - a) + \varepsilon,$$

d'où

$$|F(b) - F(a) - S_D(f)| \leq \varepsilon(b - a) + \varepsilon.$$

L'égalité (9.6) s'ensuit, et avec elle, le Théorème 9.5. \square

(9.8) La fonction de Cantor-Lebesgue. Considérons l'ensemble triadique de Cantor K , défini comme l'ensemble des nombres réels de l'intervalle $[0, 1]$ pouvant s'écrire en base 3 sous la forme

$$x = \sum_{n=1}^{+\infty} a_n 3^{-n}, \quad a_n = 0 \text{ ou } a_n = 2,$$

autrement dit l'ensemble des réels de $[0, 1]$ n'ayant en base 3 que les chiffres 0 ou 2. Il est clair par définition que $K = \bigcap K_N$ où K_N est la réunion des 2^N intervalles de la forme $[r, r + 3^{-N}]$, $r = \sum_{n=1}^N a_n 3^{-n}$ décrivant l'ensemble des nombres triadiques à N chiffres ne s'écrivant qu'avec des chiffres 0 et 2. Les ensembles K_N s'obtiennent itérativement en partant de l'intervalle $K_0 = [0, 1]$ et en otant le tiers médian de chacun des intervalles constituant K_{N-1} :



Fig. 25. Les ensembles triadiques K_N , $0 \leq N \leq 6$.

Il en résulte que K est une partie compacte ; de plus, si μ est la mesure de Lebesgue sur \mathbb{R} , on a $\mu(K_N) = 2^N \cdot 3^{-N} = (2/3)^N$, donc $\mu(K) = 0$. Puisque que K est fermé, il existe pour tout $x \in [0, 1]$ un plus grand élément $\xi \in K$ tel que $\xi \leq x$. Posons $\xi = \varphi(x) = \sum_{n=1}^{+\infty} a_n 3^{-n}$, $a_n = 0$ ou 2 , et $F(x) = \sum_{n=1}^{+\infty} a_n 2^{-n-1} \in [0, 1]$. Par définition, $x \mapsto \varphi(x)$ est croissante, ce qui signifie que $x \mapsto (a_n)$ est croissante

pour l'ordre lexicographique sur $\{0, 2\}^{\mathbb{N}}$, et donc $x \mapsto F(x)$ est également croissante. Le complémentaire $[0, 1] \setminus K$ est la réunion de la famille dénombrable d'intervalles ouverts $I_{r,N} =]r + 3^{-N}, r + 2 \times 3^{-N}[$, où $r = \sum_{n=1}^{N-1} a_n 3^{-n}$, $a_n = 0$ ou 2 , pour tous les $N \geq 1$ (ce sont les nombres dont le N -ième chiffre triadique est 1, tous les précédents étant 0 ou 2); sur chacun de ces intervalles $I_{r,N}$, on a $\varphi(x) = r + 3^{-N}$ et $F(x) = \sum_{n=0}^{N-1} a_n 2^{-n-1} + 2^{-N}$ est constante. On voit donc que $F'(x) = 0$ sur $[0, 1] \setminus K$, par conséquent F est différentiable presque partout et de dérivée nulle. On a par ailleurs $F(0) = \varphi(0) = 0$ et $F(1) = \varphi(1) = 1$. Montrons enfin que F est continue. Pour le voir, on observe que $F(x)$ varie de 0 ou de 2^{-N} sur chacun des 3^N intervalles divisant $[0, 1]$ en sous-intervalles de longueur 3^{-N} (la variation est nulle s'il s'agit d'un des intervalles $I_{r,N}$, elle vaut 2^{-N} sinon). Étant donnés deux réels distincts $x, x' \in [0, 1]$, le nombre de tels intervalles qui rencontrent $[x, x']$ est majoré par $2 + |x' - x|/3^{-N}$, d'où $|F(x') - F(x)| \leq 2^{-N}(2 + |x' - x|/3^{-N})$. En choisissant $N \geq 1$ tel que $3^{-N} < |x' - x| \leq 3^{-N+1}$, on en déduit

$$|F(x') - F(x)| \leq 5 \times 2^{-N} \leq 5 |x' - x|^{\log 2 / \log 3},$$

ce qui prouve la continuité de F .

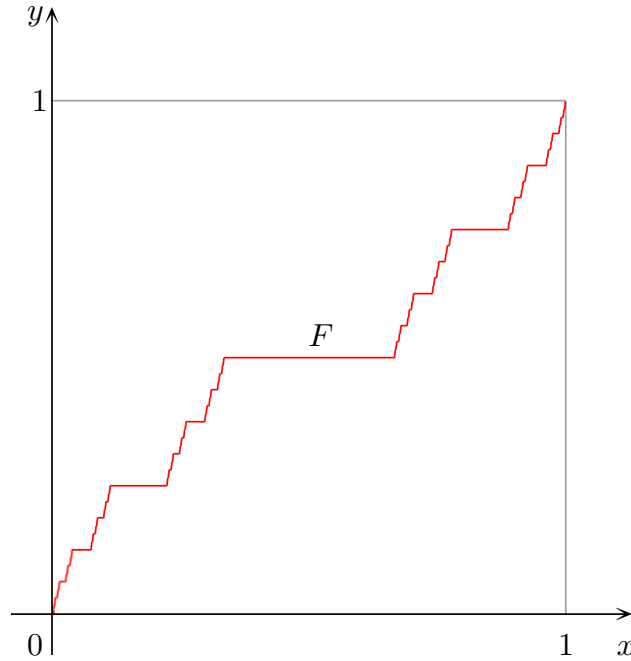


Fig. 26. L'escalier de Cantor-Lebesgue.

Comme F est croissante sur $[0, 1]$, il est immédiat de voir que la variation totale de F sur $[0, 1]$ est $\text{Var}([0, 1], F) = F(1) - F(0) = 1$. Par ailleurs, comme on l'a déjà observé, $Z := [0, 1] \setminus K$ est constitué d'une réunion dénombrable d'intervalles $]\alpha_p, \beta_p[$ sur lesquels F est constante. On en déduit que $\text{Var}(Z, F) = 0$: pour le voir, il suffit de prendre une jauge δ telle que l'on ait (disons) $\delta(\alpha_p) = \delta(\beta_p) = \varepsilon 2^{-p}$ et $\delta(x) \leq \min(x - \alpha_p, \beta_p - x)$ sur $]\alpha_p, \beta_p[$; on obtient alors $\text{Var}(Z, F) \leq C \varepsilon^{\log 2 / \log 3}$. La Proposition 9.2 (a) montre maintenant que $\text{Var}(K, F) = \text{Var}([0, 1], F) = 1$. Ceci donne un exemple pour lequel toute la variation de F est réalisée sur un ensemble négligeable, bien que F soit continue

et dérivable presque partout, et on a ici

$$\text{Var}([0, 1], F) = 1 \neq 0 = \int_0^1 F'(x) dx. \quad \square$$

10. Fonctions de puissance k-ième intégrable

L'objet de cette section est d'introduire des espaces $L^k(P)$ qui généralisent l'espace $L^1(P)$ introduit à la section 5. La présentation en est tout à fait classique.

(10.1) Définition. Soit P un pavé quelconque dans \mathbb{R}^n et $k \in]0, +\infty[$. On définit l'espace $\tilde{L}^k(P)$ comme étant l'ensemble des fonctions $f : P \rightarrow \mathbb{R}$ mesurables telles que $|f|^k$ soit KH-intégrable sur P .

On dit que $\tilde{L}^k(P)$ est l'espace des fonctions de puissance k-ième intégrable sur P . Si $f \in \tilde{L}^k(P)$, on définit la norme L^k de f comme étant

$$\|f\|_k = \left(\int_P |f(x)|^k dx \right)^{1/k}.$$

Nous avons besoin d'un lemme préparatoire fournissant un certain nombre d'inégalités élémentaires de convexité.

(10.2) Lemme. Soient $x, y \in \mathbb{R}_+$ des réels positifs ou nuls et $t > 0$.

Si $k \geq 1$ on a

$$(10.2 \text{ a}) \quad (x + y)^k \leq \alpha^{1-k} x^k + \beta^{1-k} y^k \quad \text{pour } \alpha, \beta > 0, \quad \alpha + \beta = 1,$$

$$(10.2 \text{ b}) \quad (x + y)^k \geq x^k + y^k,$$

$$(10.2 \text{ c}) \quad |x - y|^k \leq |x^k - y^k|,$$

$$(10.2 \text{ d}) \quad |x^k - y^k| \leq t^{-(k-1)} |x - y|^k + (k-1)t \max(x^k, y^k).$$

Si $0 < k \leq 1$ on a les inégalités analogues

$$(10.2 \text{ a}') \quad (x + y)^k \geq \alpha^{1-k} x^k + \beta^{1-k} y^k \quad \text{pour } \alpha, \beta > 0, \quad \alpha + \beta = 1,$$

$$(10.2 \text{ b}') \quad (x + y)^k \leq x^k + y^k,$$

$$(10.2 \text{ c}') \quad |x - y|^k \geq |x^k - y^k|,$$

$$(10.2 \text{ d}') \quad |x - y|^k \leq t^{-(1-k)} |x^k - y^k| + (1-k)t^k \max(x^k, y^k).$$

Démonstration. (a) résulte de la convexité de la fonction $x \mapsto x^k$ sur \mathbb{R}_+ (la dérivée seconde $k(k-1)x^{k-2}$ étant ≥ 0 sur $]0, +\infty[$ si $k \geq 1$), qui donne

$$(x + y)^k = (\alpha(\alpha^{-1}x) + \beta(\beta^{-1}y))^k \leq \alpha(\alpha^{-1}x)^k + \beta(\beta^{-1}y)^k.$$

De même, (a') est une conséquence du fait que la fonction est concave si $0 < k \leq 1$.

Les inégalités (b) et (b') résultent de la monotonie de la fonction $u \mapsto (u+y)^k - u^k$ sur l'intervalle $[0, x]$, qui se déduit du fait que la dérivée $k((u+y)^{k-1} - u^{k-1})$ est ≥ 0 et $k \geq 1$ (resp. ≤ 0 pour $k \leq 1$) et $u > 0$.

Les inégalités (c) et (c') se déduisent de (b) et (b') : en supposant par exemple $0 \leq y \leq x$, on trouve ainsi pour $k \geq 1$

$$x^k = (y + (x - y))^k \geq y^k + (x - y)^k.$$

Enfin, pour obtenir (d), on utilise l'inégalité des accroissements finis

$$|x^k - y^k| \leq k|x - y| \max(x^{k-1}, y^{k-1}) = k A^{1/k} B^{1-1/k}$$

avec $A = |x - y|^k$ et $B = \max(x^k, y^k)$, et on combine ceci avec l'inégalité de convexité

$$A^{1/k} B^{1-1/k} = \exp\left(\frac{1}{k} \ln(t^{-(k-1)} A) + \left(1 - \frac{1}{k}\right) \ln(tB)\right) \leq \frac{1}{k} t^{-(k-1)} A + \left(1 - \frac{1}{k}\right) tB.$$

Pour (d'), on pose $x' = x^k$, $y' = y^k$ et $k' = 1/k$, ce qui donne

$$|x - y| = |(x')^{k'} - (y')^{k'}| \leq k'|x' - y'| \max((x')^{k'-1}, (y')^{k'-1}),$$

d'où

$$|x - y|^k \leq \left(\frac{1}{k} |x^k - y^k| \max(x, y)^{1-k}\right)^k = A^k B^{1-k}$$

avec $A = \frac{1}{k} |x^k - y^k|$ et $B = \max(x^k, y^k)$. Il suffit alors de combiner cette inégalité avec l'inégalité de convexité

$$A^k B^{1-k} = \exp(k \ln(t^{-(1-k)} A) + (1-k) \ln(t^k B)) \leq k t^{-(1-k)} A + (1-k) t^k B. \quad \square$$

(10.3) Théorème. Soit P un pavé quelconque de \mathbb{R}^n .

- (a) L'ensemble $\tilde{L}^k(P)$ est un espace vectoriel, et $f \mapsto \|f\|_k$ est une semi-norme sur $\tilde{L}^k(P)$ si $k \geq 1$ (c'est seulement une « quasi semi-norme » si $0 < k < 1$).
- (b) L'ensemble des fonctions f de semi-norme $\|f\|_k$ nulle s'identifie à l'espace $N(P)$ des fonctions négligeables, et le quotient $L^k(P) = \tilde{L}^k(P)/N(P)$ est un espace complet pour tout $k > 0$ (et donc un espace de Banach si $k \geq 1$).

Démonstration. (a) Soient $f, g \in \tilde{L}^k(P)$. L'inégalité (10.2 b') implique

$$|f + g|^k \leq (|f| + |g|)^k \leq |f|^k + |g|^k \quad \text{si } 0 < k \leq 1,$$

tandis que pour $k \geq 1$ et $\alpha, \beta > 0$, $\alpha + \beta = 1$, (10.2 a) donne

$$(*) \quad |f + g|^k \leq \alpha^{1-k} |f|^k + \beta^{1-k} |g|^k.$$

La proposition 7.10 implique que $f + g \in \tilde{L}^k(P)$. Ceci montre que $\tilde{L}^k(P)$ est bien un espace vectoriel et que

$$\|f + g\|_k^k \leq \|f\|_k^k + \|g\|_k^k \quad \text{si } 0 < k \leq 1.$$

Dans ce cas, à défaut d'obtenir une semi-norme, on obtient donc une semi-distance invariante par translation en posant $d_k(f, g) = \|f - g\|_k^k$. L'élévation à la puissance $1/k$ conduit d'après (10.2 a) (appliqué avec $k' = 1/k$ et $\alpha = \beta = 1/2$) à la propriété dite de « quasi semi-norme »

$$\|f + g\|_k \leq 2^{1/k-1} (\|f\|_k + \|g\|_k).$$

L'ensemble des fonctions f de semi-norme $\|f\|_k = 0$ s'identifie à l'ensemble $N(P)$ des fonctions négligeables, et dans le cas où $\|f\|_k = \|g\|_k = 0$ on a donc

$$\|f + g\|_k = 0 = \|f\|_k + \|g\|_k.$$

Si $k \geq 1$ et $\|f\|_k + \|g\|_k > 0$, on peut appliquer l'inégalité (*) ci-dessus avec

$$\alpha = \frac{\|f\|_k}{\|f\|_k + \|g\|_k}, \quad \beta = \frac{\|g\|_k}{\|f\|_k + \|g\|_k}$$

pour obtenir

$$\begin{aligned} \|f + g\|_k^k &\leq \alpha^{1-k} \|f\|_k^k + \beta^{1-k} \|g\|_k^k \\ &\leq (\|f\|_k + \|g\|_k)^{k-1} \|f\|_k + (\|f\|_k + \|g\|_k)^{k-1} \|g\|_k, \end{aligned}$$

d'où $\|f + g\|_k \leq \|f\|_k + \|g\|_k$.

(b) Il est clair d'après ce qui précède que le quotient $L^k(P) = \tilde{L}^k(P)/N(P)$ est un espace métrisable séparé, normé ou quasi-normé suivant que $k \geq 1$ ou $0 < k < 1$. Pour montrer que c'est un espace complet, on considère les bijections inverses l'une de l'autre

$$\begin{aligned} \Phi : \tilde{L}^k(P) &\rightarrow \tilde{L}^1(P), & f &\mapsto f_+^k - f_-^k, \\ \Psi : \tilde{L}^1(P) &\rightarrow \tilde{L}^k(P), & f &\mapsto f_+^{1/k} - f_-^{1/k}. \end{aligned}$$

Comme $L^1(P)$ est complet (théorème 5.6), il suffit de voir que Φ et Ψ induisent des homéomorphismes entre $\tilde{L}^k(P)$ et $\tilde{L}^1(P)$ qui transforment suites de Cauchy en suites de Cauchy. Soient f, g dans $\tilde{L}^k(P)$. Il est tout d'abord évident par définition de Φ que

$$\|\Phi(f)\|_1 = \|f\|_k^k,$$

donc Φ, Ψ échangent les parties bornées. Lorsque $k \leq 1$, on a d'après (10.2 c')

$$\|f_+^k - g_+^k\|_1 = \int_P |f_+^k(x) - g_+^k(x)| dx \leq \int_P |f_+(x) - g_+(x)|^k dx \leq \|f - g\|_k^k,$$

(on utilise le fait trivial que $|x_+ - y_+| \leq |x - y|$), et on a une inégalité analogue pour les parties négatives, donc

$$\|\Phi(f) - \Phi(g)\|_1 \leq 2 \|f - g\|_k^k.$$

Lorsque $k \geq 1$, on utilise plutôt (10.2 d) qui donne

$$\|f_+^k - g_+^k\|_1 \leq t^{-(k-1)} \|f - g\|_k^k + (k-1)t \max(\|f\|_k, \|g\|_k)^k$$

pour tout $t > 0$. En prenant $t = \|f - g\|_k$, on trouve alors

$$\begin{aligned} \|f_+^k - g_+^k\|_1 &\leq \|f - g\|_k (1 + (k-1) \max(\|f\|_k, \|g\|_k)^k), \\ \|\Phi(f) - \Phi(g)\|_1 &\leq 2\|f - g\|_k (1 + (k-1) \max(\|f\|_k, \|g\|_k)^k). \end{aligned}$$

Dans le sens inverse, (10.2 c) implique

$$\begin{aligned} \|f_+ - g_+\|_k^k &\leq \|f_+^k - g_+^k\|_1 \leq \|\Phi(f) - \Phi(g)\|_1 \quad \text{si } k \geq 1, \text{ d'où} \\ \|f - g\|_k &\leq 2\|\Phi(f) - \Phi(g)\|_1^{1/k} \quad \text{si } k \geq 1. \end{aligned}$$

Si $k \leq 1$, l'utilisation de (10.2 d') avec $t = \|f_+^k - g_+^k\|_1$ fournit

$$\begin{aligned} \|f_+ - g_+\|_k^k &\leq \|f_+^k - g_+^k\|_1^k (1 + (1-k) \max(\|f\|_k, \|g\|_k)^k), \\ \|f - g\|_k &\leq 2^{1/k} \|\Phi(f) - \Phi(g)\|_1 (1 + (1-k) \max(\|f\|_k, \|g\|_k)^k)^{1/k}. \end{aligned}$$

La propriété de transformation des suites de Cauchy par Φ et $\Psi = \Phi^{-1}$ résulte directement de ces inégalités, et le théorème s'ensuit. \square

(10.4) Inégalité de Hölder. Soient $k, \ell > 1$ tels que $\frac{1}{k} + \frac{1}{\ell} = 1$ (« exposants conjugués »). Alors pour tout $f \in \tilde{L}^k(P)$ et tout $g \in \tilde{L}^\ell(P)$ le produit fg est dans $\tilde{L}^1(P)$ et on a

$$\left| \int_P f(x)g(x) dx \right| \leq \|f\|_k \|g\|_\ell.$$

Démonstration. La convexité de la fonction exponentielle donne

$$|fg| = \exp\left(\frac{1}{k} \ln(t^k |f|^k) + \frac{1}{\ell} \ln(t^{-\ell} |g|^\ell)\right) \leq \frac{1}{k} t^k |f|^k + \frac{1}{\ell} t^{-\ell} |g|^\ell$$

pour tout $t > 0$, et ceci montre déjà que $fg \in \tilde{L}^1(P)$ d'après la proposition 7.10. On obtient de plus

$$\left| \int_P f(x)g(x) dx \right| \leq \int_P |f(x)| |g(x)| dx \leq \frac{1}{k} t^k \|f\|_k^k + \frac{1}{\ell} t^{-\ell} \|g\|_\ell^\ell.$$

Si on prend $t = \|f\|_k^{-1/\ell} \|g\|_\ell^{1/k}$ (en supposant f, g non négligeables), on déduit bien l'inégalité de Hölder (10.4) cherchée de l'hypothèse $\frac{1}{k} + \frac{1}{\ell} = 1$, qui entraîne aussi $k - k/\ell = \ell - \ell/k = 1$. Si f ou g est négligeable, le membre de gauche est nul et l'inégalité est évidente. \square

(10.5) Remarque. On notera que l'inégalité de Hölder est également valide lorsque $k = 1$, si l'on définit l'exposant conjugué par $\ell = \infty$ dans ce cas. Ce n'est autre que l'inégalité déjà donnée en (7.12).

(10.6) Corollaire. L'espace $L^2(P) = \tilde{L}^2(P)/N(P)$ des fonctions de carré intégrable presque partout définies sur P a une structure d'espace de Hilbert.

Démonstration. Si $f, g \in \tilde{L}^2(P)$, alors $|fg| \leq \frac{1}{2}(f^2 + g^2) \in \tilde{L}^1(P)$, donc $fg \in \tilde{L}^1(P)$. On peut ainsi définir sur $L^2(P)$ un produit scalaire

$$\langle f, g \rangle = \int_P f(x)g(x) dx$$

tel que $\langle f, f \rangle = \|f\|_2^2$. Cette forme bilinéaire positive est bien non dégénérée puisque son noyau dans $\tilde{L}^2(P)$ est précisément $N(P)$, et que dans $L^2(P)$ on est passé au quotient. Le théorème 10.3 montre que $L^2(P)$ est un espace de Hilbert. On en tire en particulier l'inégalité de Cauchy-Schwarz $|\langle f, g \rangle| \leq \|f\|_2 \|g\|_2$, qui se réécrit sous forme intégrale

$$(10.7) \quad \left(\int_P f(x)g(x) dx \right)^2 \leq \int_P |f(x)|^2 dx \int_P |g(x)|^2 dx,$$

et qui est précisément le cas particulier de l'inégalité de Hölder pour $k = \ell = 2$. \square

11. Caractérisation de l'intégrabilité au sens de Riemann

Les résultats qui précèdent permettent de donner une caractérisation des fonctions intégrables au sens de Riemann.

(11.1) Théorème. *Soit P un pavé fermé borné de \mathbb{R}^n et $f : P \rightarrow \mathbb{R}$ une fonction quelconque. Alors f est intégrable au sens de Riemann sur P si et seulement si*

- (a) f est bornée sur P ;
- (b) l'ensemble N des points de discontinuité de f est une partie négligeable de P .

Démonstration. (1) Les conditions sont nécessaires.

Supposons f intégrable au sens de Riemann. Alors, pour $\varepsilon > 0$ donné, il existe $\delta > 0$ tel que pour toute subdivision $D = \{(Q_j, x_j)\}_{0 \leq j < N}$ δ -fine on ait $|S_D(f) - A| \leq \varepsilon$ où A est l'intégrale de f , ce qui exige que $S_D(f) = \sum \text{vol}(Q_j) f(x_j)$ reste borné. En fixant un indice j et en faisant varier $x_j \in Q_j$ (tous les autres points $x_k \in Q_k$, $k \neq j$, étant fixés), on voit que f doit être bornée sur Q_j . Par conséquent f doit être bornée sur P .

Si nous posons maintenant $M_j = \sup_{x_j \in Q_j} f(x_j)$ et $m_j = \inf_{x_j \in Q_j} f(x_j)$, nous voyons que $\sum M_j \text{vol}(Q_j) \leq A + \varepsilon$ et $\sum m_j \text{vol}(Q_j) \geq A - \varepsilon$ en prenant respectivement le sup et l'inf de $S_D(f)$. Par conséquent $\sum (M_j - m_j) \text{vol}(Q_j) \leq 2\varepsilon$. Soit J_0 l'ensemble des indices tels que $M_j - m_j \geq \sqrt{\varepsilon}$. On obtient $\sum_{j \in J_0} \sqrt{\varepsilon} \text{vol}(Q_j) \leq 2\varepsilon$, donc $\text{vol}(\bigcup_{j \in J_0} Q_j) \leq 2\sqrt{\varepsilon}$. Du fait que $\text{vol}(\partial Q_j) = 0$, il existe un voisinage ouvert

$$U_\varepsilon \quad \text{de} \quad \bigcup_{j \in J_0} Q_j \cup \bigcup_{j \notin J_0} \partial Q_j$$

tel que $\text{vol}(U_\varepsilon) \leq 4\sqrt{\varepsilon}$. Le complémentaire $P \setminus U_\varepsilon$ est contenu par construction dans la réunion des cubes ouverts Q_j° , $j \notin J_0$, donc tout point $x_0 \in P \setminus U_\varepsilon$ admet un voisinage sur lequel l'oscillation de f est majorée par $\sqrt{\varepsilon}$ (puisque $M_j - m_j < \sqrt{\varepsilon}$ pour $j \notin J_0$).

Considérons l'ensemble négligeable

$$S = \bigcup_{k \geq 1} \bigcap_{j \geq k} U_{1/j} \quad \text{et son complémentaire} \quad P \setminus S = \bigcap_{k \geq 1} \bigcup_{j \geq k} P \setminus U_{1/j}.$$

(On notera que $\text{vol}(\bigcap_{j \geq k} U_{1/j}) \leq 4/\sqrt{j}$ pour tout $j \geq k$, donc cet ensemble est bien négligeable). Alors, pour tout $x_0 \in P \setminus S$ et tout entier $k \geq 1$, il existe $j \geq k$ tel que $x_0 \in P \setminus U_{1/j}$, et par conséquent x_0 possède un voisinage sur lequel l'oscillation de f est majorée par $\sqrt{1/j} \leq \sqrt{1/k}$. Ceci implique que f est continue en x_0 et donc $N \subset S$.

(2) Réciproquement, supposons que $|f| \leq M$ et que l'ensemble N des points de discontinuité de f soit négligeable. Fixons $\varepsilon > 0$ et posons

$$S_k = \{x \in P; \exists \delta > 0, \sup_{\overline{B}(x, 2^{-k}-\delta)} f \geq \varepsilon\}.$$

Il est clair par définition que S_k est une suite décroissante et que S_k est une partie ouverte de P (si $x_0 \in S_k$ avec $\sup_{\overline{B}(x_0, 2^{-k}-\delta)} f \geq \varepsilon$, alors pour $x \in B(x_0, \delta/2)$ on a $\sup_{\overline{B}(x, 2^{-k}-\delta/2)} f \geq \sup_{\overline{B}(x_0, 2^{-k}-\delta)} f \geq \varepsilon$). Comme $N = \bigcap S_k$ est de mesure nulle, il existe un entier k tel que la mesure de Lebesgue de S_k satisfasse $\mu(S_k) \leq \varepsilon$.

Considérons un recouvrement fini du compact $P \setminus S_k$ par des pavés ouverts $(Q'_j)^\circ$ centrés en des points $x'_j \in P \setminus S_k$, de diagonale $< 2^{-k}$. On a $Q'_j \subset B(x'_j, 2^{-k} - \delta)$ pour δ assez petit, et comme $x'_j \notin S_k$, il vient

$$\sup_{B(x'_j, 2^{-k}-\delta)} f < \varepsilon \quad \text{et donc} \quad \sup_{Q'_j} f < \varepsilon.$$

On obtient ainsi un recouvrement $\bigcup (Q'_j)^\circ$ d'un voisinage de $P \setminus S_k$. On complète la réunion $\bigcup Q'_j$ pour recouvrir P tout entier en adjoignant un nombre fini de pavés $Q''_j \subset S_k$ dont les intérieurs $(Q''_j)^\circ$ sont deux à deux disjoints et contenus dans $P \setminus \bigcup Q'_j$.

Soit $P = \bigcup Q_\ell$ un découpage de P obtenu en redécoupant les pavés Q'_j et Q''_j de sorte que $Q_\ell \subset Q'_j$ ou $Q_\ell \subset Q''_j$ pour au moins un j . On pose

$$M_\ell = \sup_{Q_\ell} f, \quad m_\ell = \inf_{Q_\ell} f.$$

Pour $Q_\ell \subset Q'_j$ on a $M_\ell - m_\ell \leq \text{oscil}_{Q_\ell} f \leq \text{oscil}_{Q'_j} f < \varepsilon$, et d'autre part $\bigcup Q''_j \subset S_k$, donc

$$\begin{aligned} \sum_{\ell} \text{vol}(Q_\ell)(M_\ell - m_\ell) &\leq \sum_{Q_\ell \subset \bigcup Q'_j} \text{vol}(Q_j)(M_j - m_j) + \sum_{Q_\ell \subset \bigcup Q''_j} \text{vol}(Q_j)(M_j - m_j) \\ &\leq \varepsilon \sum_{Q_\ell \subset \bigcup Q'_j} \text{vol}(Q_j) + M \sum_{Q_\ell \subset \bigcup Q''_j} \text{vol}(Q_j) \\ &\leq \varepsilon \text{vol}(P) + M\mu(S_k) \leq (\text{vol}(P) + M)\varepsilon. \end{aligned}$$

Ceci démontre que f est intégrable au sens de Riemann. □

12. Fonctions approximativement continues et primitives

Soit Ω un ouvert de \mathbb{R}^n , E un espace métrique (ou topologique) ayant une base dénombrable d'ouverts.

(12.1) Définition. On dit qu'une application $f : \Omega \rightarrow E$ est

- (a) mesurable si pour tout ouvert V de E l'image réciproque $f^{-1}(V)$ est mesurable ;
- (b) approximativement continue en un point $x_0 \in \Omega$ (relativement à la mesure de Lebesgue μ) si elle est mesurable au voisinage de x_0 et si pour tout voisinage ouvert V de $f(x_0)$, la densité des points x proches de x_0 dont l'image $f(x)$ est en dehors de V est nulle, c'est-à-dire :

$$\limsup_{r \rightarrow 0_+} \frac{\mu(B(x_0, r) \setminus f^{-1}(V))}{\mu(B(x_0, r))} = 0.$$

- (c) approximativement continue sur Ω si elle est approximativement continue en tout point de Ω .

Il est facile de voir que la condition 12.1 (b) est indépendante de la norme choisie sur \mathbb{R}^n . Pour les fonctions d'une variable, la condition devient

$$(12.2) \quad \forall V \text{ ouvert } \ni f(x_0), \quad \limsup_{h \rightarrow 0_+} \frac{\mu([x_0 - h, x_0 + h] \setminus f^{-1}(V))}{2h} = 0.$$

L'énoncé suivant est à peu près évident.

(12.3) Proposition Soit $f : \Omega \rightarrow E$ une application approximativement continue en un point $x_0 \in \Omega$ et $g : E \rightarrow E'$ une application continue entre espaces topologiques. Alors $g \circ f$ est approximativement continue en x_0 .

Démonstration. En effet, si W est un voisinage de $g \circ f(x_0)$ alors $V = g^{-1}(W)$ est un voisinage de $f(x_0)$ et $(g \circ f)^{-1}(W) = f^{-1}(V)$. \square

L'intérêt principal des fonctions approximativement continues réside dans le résultat élémentaire suivant dû à A. Denjoy.

(12.4) Théorème (Denjoy [Dj2]). Soit I un intervalle de \mathbb{R} et $f : I \rightarrow \mathbb{R}$ une fonction mesurable bornée et approximativement continue en un point $x_0 \in I$. Alors l'intégrale indéfinie

$$F(x) = \int_a^x f(t) dt, \quad a \in I$$

est dérivable en x_0 et $F'(x_0) = f(x_0)$. En particulier, toute fonction $f : I \rightarrow \mathbb{R}$ localement bornée et approximativement continue possède une primitive.

Démonstration. Considérons un point $x = x_0 + h \in I$, $h \in \mathbb{R}$, et fixons $\varepsilon > 0$. En posant $M = \|f\|_\infty$ et $V =]f(x_0) - \varepsilon, f(x_0) + \varepsilon[$, on a alors

$$\frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) = \frac{1}{h} \int_{x_0}^{x_0+h} (f(t) - f(x_0)) dt,$$

et comme car $|f(t) - f(x_0)| \leq \varepsilon$ pour $t \in f^{-1}(V)$ et $|f(t) - f(x_0)| \leq 2M$ presque partout pour $t \notin f^{-1}(V)$, il vient

$$\begin{aligned} \left| \frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) \right| &\leq \frac{1}{|h|} \int_{]x_0, x_0+h[\cap f^{-1}(V)} |f(t) - f(x_0)| dt \\ &\quad + \frac{1}{|h|} \int_{]x_0, x_0+h[\setminus f^{-1}(V)} |f(t) - f(x_0)| dt \\ &\leq \varepsilon + \frac{2M \mu([x_0, x_0 + h[\setminus f^{-1}(V))}{|h|}. \end{aligned}$$

L'hypothèse sur la continuité approximative de f implique qu'il existe $\delta > 0$ tel que

$$\begin{aligned} 0 < |h| < \delta &\implies \frac{\mu([x_0 - h, x_0 + h[\setminus f^{-1}(V))}{2|h|} \leq \varepsilon \\ &\implies \left| \frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) \right| \leq \varepsilon(1 + 4M). \end{aligned}$$

Les résultats annoncés en découlent. \square

(12.5) Corollaire. *Une fonction approximativement continue $f : I \rightarrow \overline{\mathbb{R}}$ satisfait le théorème des valeurs intermédiaires.*

Démonstration. En effet, quitte à composer f avec un homéomorphisme $g : \overline{\mathbb{R}} \rightarrow [-1, 1]$, on peut supposer f bornée. Alors f admet une primitive F , et on sait que la dérivée $f = F'$ satisfait le théorème des valeurs intermédiaires. \square

Le théorème de Denjoy permet de construire des fonctions dont la dérivée est approximativement continue mais discontinue en de nombreux points.

(12.6) Proposition (Denjoy [Dj2]). *Soit $(r_n)_{n \geq n_0}$ une suite partout dense dans \mathbb{R} (par exemple la suite des rationnels ordonnés de manière quelconque), et $\alpha_n > 0$, $\beta_n > 1$, $n \geq n_0$, des suites de réels telles que*

$$\sum_{n=n_0}^{+\infty} \alpha_n < +\infty, \quad \sum_{n=n_0}^{+\infty} K^{-\beta_n} < +\infty$$

pour une certaine constante $K > 1$ (on peut prendre par exemple $\alpha_n = 2^{-n}$, $\beta_n = \ln n$, $n \geq n_0 = 3$, avec $K = e^2$). Alors

(a) *la fonction $u : \mathbb{R} \rightarrow [0, +\infty]$ telle que*

$$u(x) = \sum_{n=n_0}^{+\infty} \alpha_n |x - r_n|^{-1/\beta_n}$$

est approximativement continue sur \mathbb{R} , finie presque partout, mais infinie sur un G_δ dense de mesure nulle dans \mathbb{R} .

(b) la fonction $f(x) = e^{-u(x)} \cos u(x)$ est mesurable sur \mathbb{R} , bornée par 1 en valeur absolue, approximativement continue en tout point, et change de signe sur tout intervalle. En particulier, elle admet une primitive F qui n'est monotone sur aucun intervalle. De plus f est continue avec $f(x_0) = 0$ en tout point x_0 tel que $u(x_0) = +\infty$, et discontinue en tout point x_0 tel que $u(x_0) < +\infty$.

Démonstration. (a) Considérons les fonctions $u_n : \mathbb{R} \rightarrow [0, +\infty]$ définies par $u_n(x) = \alpha_n |x - r_n|^{-1/\beta_n}$. Elles sont continues, avec $u_n(r_n) = +\infty$ et $u_n(x) < +\infty$ si $x \neq r_n$, donc les sommes partielles

$$S_N = \sum_{n_0 \leq n \leq N} u_n$$

sont continues. Par conséquent, la limite croissante $u = \lim S_N$ est semi-continue inférieurement, c'est-à-dire que pour tout $\lambda < u(x_0)$ il existe un voisinage $]x_0 - \delta, x_0 + \delta[$ sur lequel $u(x) > \lambda$ (que $u(x_0)$ soit fini ou infini) – pour le voir, on choisit N tel que $S_N(x_0) > \lambda$, puis $\delta > 0$ tel que $S_N(x) > \lambda$ sur $]x_0 - \delta, x_0 + \delta[$, en utilisant la continuité de S_N .

Il suffira donc de montrer que u est approximativement semi-continue supérieurement, c'est-à-dire que pour tout $x_0 \in \mathbb{R}$ tel que $u(x_0) < +\infty$ et tout $\lambda > u(x_0)$, l'ensemble $u^{-1}(] \lambda, +\infty])$ est de densité nulle en x_0 .

Par hypothèse, la suite β_n doit vérifier $\lim K^{-\beta_n} = 0$, donc $\lim \beta_n = +\infty$, ce qui implique l'existence d'une borne inférieure $\beta = \inf \beta_n = \min \beta_n > 1$. Pour tout intervalle $[-A, A]$ et $r \geq 0$ on a

$$\begin{aligned} \int_{-A}^A |x - r|^{-1/\beta} dx &= \int_{-A-r}^{A-r} |x|^{-1/\beta} dx \\ &= \int_{-A}^A |x|^{-1/\beta} dx + \int_A^{A+r} (|x|^{-1/\beta} - |x-r|^{-1/\beta}) dx \\ &\leq \int_{-A}^A |x|^{-1/\beta} dx = 2(1 - 1/\beta)^{-1} A^{1-1/\beta}, \end{aligned}$$

et, par symétrie, on a la même inégalité pour $r \leq 0$. Ceci entraîne pour tout N

$$\int_{-A}^A S_N(x) dx \leq 2(1 - 1/\beta)^{-1} A^{1-1/\beta} \sum_{n=n_0}^N \alpha_n.$$

Le théorème de convergence monotone implique dans ces conditions que $u = \lim S_N$ est intégrable au sens de Lebesgue sur tout intervalle $[-A, A]$, avec

$$\int_{-A}^A u(x) dx \leq 2(1 - 1/\beta)^{-1} A^{1-1/\beta} \sum_{n=n_0}^{+\infty} \alpha_n.$$

En particulier, on a $u(x) < +\infty$ presque partout. En revanche, $u^{-1}(+\infty) = \bigcap_{\lambda \in \mathbb{N}} u^{-1}(] \lambda, +\infty])$ est un G_δ qui contient la suite $\{r_n\}$, donc c'est un G_δ dense, de mesure nulle d'après ce qui précède.

Fixons maintenant x_0 tel que $u(x_0) = \sum \alpha_n |x_0 - r_n|^{-1/\beta_n} < +\infty$ et un réel $\varepsilon > 0$. Ceci implique en particulier que $x_0 \notin \{r_n\}$. L'ensemble des points $x \in \mathbb{R}$ tel que $u_n(x) > K u_n(x_0)$ est l'intervalle J_n défini par $|x - r_n| < K^{-\beta_n} |x_0 - r_n|$. Pour que l'intervalle J_n rencontre un intervalle donné $]x_0 - h, x_0 + h[$, il faut que la distance séparant les centres des intervalles soit plus petite que la somme de leurs demi-longueurs, soit $|x_0 - r_n| < h + K^{-\beta_n} |x_0 - r_n|$, et donc $|x_0 - r_n| < (1 - K^{-1})^{-1} h$. Par conséquent, la longueur d'un tel intervalle J_n est majorée par

$$2K^{-\beta_n} |x_0 - r_n| < 2K^{-\beta_n} (1 - K^{-1})^{-1} h$$

et on en déduit

$$(12.7) \quad \frac{1}{2h} \mu \left(]x_0 - h, x_0 + h[\cap \bigcup_{n>N} J_n \right) \leq \sum_{n>N} K^{-\beta_n} (1 - K^{-1})^{-1}.$$

Comme la série $\sum K^{-\beta_n}$ converge par hypothèse, on peut fixer un entier N tel que $\sum_{n>N} K^{-\beta_n} (1 - K^{-1})^{-1} < \varepsilon$, et on peut supposer également par ailleurs que $\sum_{n>N} u_n(x_0) < \varepsilon/2K$. Pour $x \in]x_0 - h, x_0 + h[\setminus \bigcup_{n>N} J_n$ nous avons

$$\begin{aligned} u(x) &\leq \sum_{n_0 \leq n \leq N} u_n(x) + \sum_{n>N} u_n(x) \\ &\leq \sum_{n_0 \leq n \leq N} u_n(x) + K \sum_{n>N} u_n(x_0) \leq \sum_{n_0 \leq n \leq N} u_n(x) + \varepsilon/2. \end{aligned}$$

Par continuité de u_n , on peut choisir $\delta > 0$ assez petit en sorte que $h \leq \delta$ et $x \in]x_0 - h, x_0 + h[\Rightarrow u_n(x) < u_n(x_0) + \varepsilon/2N$, d'où $u(x) \leq u(x_0) + \varepsilon$ d'après l'inégalité précédente. Ceci montre que pour $h \leq \delta$, les points de $]x_0 - h, x_0 + h[$ en lesquels $u(x) > u(x_0) + \varepsilon$ sont contenus dans $]x_0 - h, x_0 + h[\cap \bigcup_{n>N} J_n$, et donc que leur densité est majorée par ε d'après (12.7) et le choix de N . On en conclut que la limite de cette densité est nulle, et donc que u est bien approximativement semi-continue supérieurement en x_0 .

(b) La fonction $g(x) = e^{-x} \cos x$ définie sur $[0, +\infty[$ admet un maximum qui vaut $g(0) = 1$, et un minimum égal à $g(3\pi/4) = m$ où $m = -\frac{\sqrt{2}}{2} e^{-3\pi/4}$. Elle se prolonge en une application continue $g : [0, +\infty] \rightarrow [m, 1]$ en posant $g(+\infty) = 0$. On en déduit que la composée $f(x) = g \circ u(x) = e^{-u(x)} \cos u(x)$ est approximativement continue sur \mathbb{R} tout entier, avec $\|f\|_\infty \leq 1$. Elle admet donc une primitive F . Comme $u(r_n) = +\infty$ et que u satisfait le théorème des valeurs intermédiaires, on voit que f prend tous les signes $+$, $-$, 0 au voisinage de chaque point r_n , et donc sur tout intervalle ouvert de \mathbb{R} . Il résulte de (a) que $f(x) = 0$ sur le G_δ dense de mesure nulle $u^{-1}(+\infty)$ [il s'agit de points où f est continue], en revanche f est discontinue en tout point x_0 où $u(x_0) < +\infty$ puisqu'il va y avoir des points r_n arbitrairement proches en lesquels $u(r_n) = +\infty$ et donc $f(x)$ va prendre toute valeur de l'intervalle $g([u(x_0), +\infty[)$ aussi près qu'on veut de x_0 . \square

Nous présentons maintenant un exemple montrant qu'une fonction dérivée peut avoir un comportement très pathologique au sens de Lebesgue, même lorsque la théorie de Kurzweil-Henstock permet d'intégrer une telle fonction sans difficulté.

(12.8) Proposition Soit $(r_n)_{n \geq n_0}$ une suite partout dense dans \mathbb{R} (par exemple la suite des rationnels ordonnés de manière quelconque), et $\alpha_n > 0$, $\beta_n > 0$, $n \geq n_0$, des suites de réels positifs telles que

$$\sum_{n \geq n_0} \beta_n < +\infty, \quad \sum_{n \geq n_0} \alpha_n \beta_n^{-1} < +\infty.$$

On considère les fonctions $v : \mathbb{R} \rightarrow \mathbb{R}$ et $f : \mathbb{R} \rightarrow \mathbb{R}$ telles que

$$v(x) = x^2 \sin \frac{1}{x^3} \quad \text{si } x \neq 0, \quad v(0) = 0,$$

$$f(x) = \sum_{n \geq n_0} \alpha_n v(\beta_n^{-1}(x - r_n)).$$

Alors :

- (a) f est une fonction continue bornée sur \mathbb{R} .
- (b) f est dérivable presque partout sur \mathbb{R} .
- (c) Si les suites (r_n) , (α_n) et (β_n) vérifient les conditions supplémentaires

$$\beta_p \leq \frac{1}{2} \min_{n < p} |r_n - r_p|, \quad \lim_{n \rightarrow \infty} \frac{\alpha_n \beta_n^2}{\beta_{n+1}} = +\infty,$$

la fonction f' n'est intégrable au sens de Lebesgue sur aucun intervalle de \mathbb{R} .⁽²¹⁾

On notera que si $(r_n)_{n \geq 1}$ est la suite des rationnels distincts énumérés en sorte que le dénominateur de r_n soit inférieur ou égal à n , alors $|r_n - r_p| \leq 1/np$ pour $n < p$, et par conséquent les suites $\beta_n = 2^{-5^n}$ et $\alpha_n = \beta_n^2 = 4^{-5^n}$ satisfont toutes les hypothèses.

Démonstration. L'inégalité $|\sin x| \leq |x|$ implique $|v(x)| \leq 1/|x|$, donc $|v(x)| \leq 1$ sur \mathbb{R} tout entier (cette majoration étant triviale par définition si $x \in [-1, 1]$). On observe également que v est partout dérivable sur \mathbb{R} avec $v'(0) = \lim_{x \rightarrow 0} v(x)/x = 0$ et

$$v'(x) = 2x \sin \frac{1}{x^3} - \frac{3}{x^2} \cos \frac{1}{x^3} \quad \text{si } x \neq 0,$$

par conséquent

$$(12.9) \quad |v'(x)| \leq \frac{5}{|x|^2} \quad \text{si } x \neq 0.$$

(a) *Continuité de f .* Puisque la convergence de $\sum \beta_n$ entraîne $\lim \beta_n = 0$, la convergence de $\sum \alpha_n \beta_n^{-1}$ implique celle de $\sum \alpha_n$. Par conséquent, du fait que v est

(21) D'après le théorème 7.13, on en déduit aussi que f' n'est alors KH-intégrable sur aucun intervalle. On notera que ceci se produit bien que f' soit somme d'une série presque partout convergente de fonctions KH-intégrables dont la série des intégrales indéfinies converge normalement, de sorte que f' définit tout de même une distribution dont la primitive est la fonction continue f .

continue bornée, nous voyons que $f(x) = \sum_{n \geq n_0} \alpha_n v(\beta_n^{-1}(x - r_n))$ est continue bornée comme somme d'une série uniformément convergente de telles fonctions.

(b) *Différentiabilité de f presque partout.* Considérons l'ensemble

$$E = \bigcap_{N \geq n_0} \bigcup_{n \geq N}]r_n - \beta_n, r_n + \beta_n[,$$

qui est négligeable en vertu de l'hypothèse $\sum \beta_n < +\infty$. Pour $x \in \mathbb{R} \setminus E$, il existe un indice $N(x)$ tel $x \notin \bigcup_{n \geq N(x)}]r_n - \beta_n, r_n + \beta_n[$. Ceci implique $|v'(\beta_n^{-1}(x - r_n))| \leq 5$ grâce à (12.9), donc la série des dérivées terme à terme

$$g(x) = \sum_{n \geq n_0} \alpha_n \beta_n^{-1} v'(\beta_n^{-1}(x - r_n))$$

converge absolument au point x d'après l'hypothèse $\sum \alpha_n \beta_n^{-1} < +\infty$. Choisissons un point $x + h \in \mathbb{R}$ et, pour $n \geq N(x)$, distinguons suivant que $h \geq \beta_n/2$ ou $h < \beta_n/2$. Si $h \geq \beta_n/2$, alors

$$\left| \frac{1}{h} \left(v(\beta_n^{-1}(x + h - r_n)) - v(\beta_n^{-1}(x - r_n)) \right) \right| \leq \frac{2}{|h|} \leq 4\beta_n^{-1}.$$

Si au contraire $h < \beta_n/2$, alors comme $x \notin]r_n - \beta_n, r_n + \beta_n[$, on voit que $[x, x + h]$ est contenu dans $\mathbb{C}]r_n - \beta_n/2, r_n + \beta_n/2[$, et la dérivée de $t \mapsto v(\beta_n^{-1}(t - r_n))$ y est majorée par $20\beta_n^{-1}$ grâce à (12.9). On a donc cette fois

$$\left| \frac{1}{h} \left(v(\beta_n^{-1}(x + h - r_n)) - v(\beta_n^{-1}(x - r_n)) \right) \right| \leq 20\beta_n^{-1}.$$

La contribution des taux d'accroissement des termes d'indice $n \geq N(x)$ dans la série définissant f est donc majorée par $20 \sum_{n \geq N(x)} \alpha_n \beta_n^{-1}$ tandis que celle des dérivées $v'(\beta_n^{-1}(x - r_n))$ d'indice $n \geq N(x)$ dans la série $g(x)$ est majorée par $5 \sum_{n \geq N(x)} \alpha_n \beta_n^{-1}$. On peut toujours choisir $N(x)$ assez grand pour que $\sum_{n \geq N(x)} \alpha_n \beta_n^{-1} \leq \varepsilon$. On voit alors que la différence

$$\frac{f(x + h) - f(x)}{h} - g(x)$$

diffère de la somme finie

$$\sum_{n_0 \leq n < N(x)} \alpha_n \frac{1}{h} \left(v(\beta_n^{-1}(x + h - r_n)) - v(\beta_n^{-1}(x - r_n)) \right) - \alpha_n \beta_n^{-1} v'(\beta_n^{-1}(x - r_n))$$

d'au plus 25ε . Comme v est partout dérivable, cette somme finie tend vers 0 quand h tend vers 0. On en conclut que $f'(x) = g(x)$ pour $x \in \mathbb{R} \setminus E$.

(c) *Non intégrabilité de f' au sens de Lebesgue.* Pour simplifier les notations, posons

$$v_n(x) = v(\beta_n^{-1}(x - r_n)).$$

On considère l'intégrale de $|f'|$ sur l'ensemble compact

$$K = [r_p - \beta_p, r_p + \beta_p] \setminus \bigcup_{n > p}]r_n - \beta_n, r_n + \beta_n[.$$

Comme $f'(x) = g(x) = \sum \alpha_n v'_n(x)$ presque partout, nous avons

$$(12.10) \quad \int_K |f'(x)| dx \geq \int_K \alpha_p |v'_p(x)| dx - \sum_{n \neq p} \int_K \alpha_n |v'_n(x)| dx.$$

L'hypothèse (d) entraîne $\beta_{n+1} \ll \alpha_n \beta_n^2 \ll \frac{1}{2} \beta_n$ pour n assez grand, donc la suite (β_n) est décroissante à partir d'un certain rang et à décroissance au moins exponentiellement rapide. Compte tenu de l'hypothèse $\beta_p \leq \frac{1}{2} \min_{n < p} |r_n - r_p|$, nous voyons que K est disjoint de l'intervalle $]r_n - \beta_p, r_n + \beta_p[$, $n < p$. Or la majoration $|v'(x)| \leq 5/x^2$ implique $\int_{[\gamma, +\infty[} |v'(x)| dx \leq 5\gamma^{-1}$ pour tout $\gamma > 0$, donc

$$\int_{|x - r_n| \geq \gamma} |v'_n(x)| dx = \int_{|x| \geq \gamma} \beta_n^{-1} |v'(\beta_n^{-1}x)| dx = \int_{|x| \geq \gamma/\beta_n} |v'(x)| dx \leq 10 \beta_n / \gamma.$$

Cette dernière majoration appliquée à $\gamma = \beta_p$ pour $n < p$ (resp. $\gamma = \beta_n$ pour $n > p$) donne

$$(12.11) \quad \int_K \alpha_n |v'_n(x)| dx \leq \begin{cases} 10 \alpha_n \beta_n / \beta_p & \text{pour } n < p, \\ 10 \alpha_n & \text{pour } n > p. \end{cases}$$

Or, en posant $\gamma = \sum_{n > p} \beta_n \leq 2\beta_{p+1}$ et en faisant le changement de variable $t = \beta_p^{-1}(x - r_p)$, on voit que

$$\int_K |v'_p(x)| dx = \int_{[-1, 1] \setminus E} |v'(t)| dt$$

où E est une réunion d'intervalles de mesure $\leq 2\gamma/\beta_p$. Comme $\int_{\delta}^1 |v'(t)| dt \geq c\delta^{-1}$ pour tout $\delta \leq 1/2$ et pour une certaine constante $c > 0$, on peut utiliser le fait que $[-1, 1] \supset [-1, -\lambda\gamma/\beta_p] \cup [\lambda\gamma/\beta_p, 1]$, la parité de v' , la majoration $|v'(t)| \leq 5/t^2$ et la décroissance de $t \mapsto 5/t^2$ pour en déduire

$$\int_K |v'_p(x)| dx \geq 2 \int_{\lambda\gamma/\beta_p}^1 |v'(t)| dt - \int_{\lambda\gamma/\beta_p}^{(\lambda+2)\gamma/\beta_p} \frac{5}{t^2} dt \geq \frac{2c\beta_p}{\lambda\gamma} - \frac{10}{\lambda(\lambda+2)} \frac{\beta_p}{\gamma}$$

pour tout $\lambda > 0$ et tout $p \geq p_0(\lambda)$ assez grand. En choisissant λ assez grand pour que $10/(\lambda+2) \leq c$, on obtient

$$(12.12) \quad \int_K |v'_p(x)| dx \geq \frac{c}{\lambda} \frac{\beta_p}{\gamma} \geq c_1 \frac{\beta_p}{\beta_{p+1}}$$

avec $c_1 = c/2\lambda$. En combinant (12.10), (12.11), (12.12) il vient

$$\int_K |f'(x)| dx \geq c_1 \frac{\alpha_p \beta_p}{\beta_{p+1}} - 10 \sum_{n < p} \frac{\alpha_n \beta_n}{\beta_p} - 10 \sum_{n > p} \alpha_n \geq c_1 \frac{\alpha_p \beta_p}{\beta_{p+1}} - \frac{c_2}{\beta_p} - c_3 \geq \frac{c_1}{2} \frac{\alpha_p \beta_p}{\beta_{p+1}}$$

pour p assez grand, sous l'hypothèse $\lim_{p \rightarrow \infty} \frac{\alpha_p \beta_p^2}{\beta_{p+1}} = +\infty$. Ceci entraîne en particulier que

$$\int_{[r_p - \beta_p, r_p + \beta_p]} |f'(x)| dx \geq \frac{c_1}{2} \frac{\alpha_p \beta_p}{\beta_{p+1}} \rightarrow +\infty.$$

Grâce à la densité de la suite (r_p) et au fait que $\beta_p \rightarrow 0$, on en déduit que $|f'|$ ne peut-être intégrable au sens de Lebesgue sur aucun intervalle de \mathbb{R} . \square

Chapitre IV

Compléments historiques

Ces compléments historiques ont été compilés et rédigés par Bernard Ycart à l'intention des étudiants de L1 de Grenoble (cours en ligne).

1. Archimède et la quadrature de la parabole

Voici une traduction de la préface d'un texte d'Archimède (287–212 av. JC), à propos de l'aire d'un arc de parabole.

« Archimède, à Dositheus, salut.

Quand j'ai appris que Conon, qui était mon ami, était mort, mais que vous connaissiez Conon, et étiez également versé en géométrie, tandis que je pleurais la perte non seulement d'un ami, mais aussi d'un admirable mathématicien, je me fixai la tâche de vous communiquer, comme j'avais l'intention de le faire pour Conon, un certain théorème géométrique qui n'avait pas été cherché avant, mais a maintenant été cherché par moi, et que j'ai découvert par des moyens mécaniques, et ensuite montré par des moyens géométriques. Certains des géomètres anciens ont essayé de montrer qu'il est possible de trouver une surface carrée égale à un cercle donné, et à un arc donné d'un cercle ; et après cela, ils ont tenté de calculer la surface bornée par une section de cône [ellipse] et une ligne droite, en supposant des lemmes si peu aisés à concevoir, qu'il fut reconnu par la plupart que le problème n'était pas résolu. Mais à ma connaissance, aucun de mes prédécesseurs n'a tenté de calculer la surface de l'arc borné par une ligne droite et la section à angle droit d'un cône [parabole], problème dont j'ai découvert la solution. Car il est montré ici que tout arc borné par une ligne droite et la section à angle droit d'un cône est les quatre tiers du triangle qui a la même base et la même hauteur que l'arc. Pour la démonstration de cette propriété, le lemme suivant est supposé : l'excès par lequel la plus grande de deux aires inégales excède la plus petite peut, si on l'ajoute à lui-même, être rendu plus grand que toute surface finie. Les géomètres anciens ont aussi utilisé ce lemme ; car c'est en utilisant ce même lemme qu'ils ont montré que les cercles sont entre eux dans le même rapport que le carré de leurs diamètres, que les sphères sont entre elles dans le même rapport que le cube de leurs diamètres, et de plus que toute pyramide est le tiers du prisme qui a la même base que la pyramide et la même hauteur ; aussi, que chaque cône est le tiers du cylindre ayant la même base que le cône et même hauteur fut prouvé en supposant un lemme similaire à celui cité plus haut. Donc, comme mon travail maintenant publié a satisfait le même test que les propositions citées, j'ai rédigé la démonstration, et je vous l'envoie, telle qu'elle a été trouvée par le moyen de la mécanique, et ensuite aussi comme elle est prouvée par la géométrie. »

En termes modernes, le résultat démontré par Archimède est le suivant (voir Fig. 27).

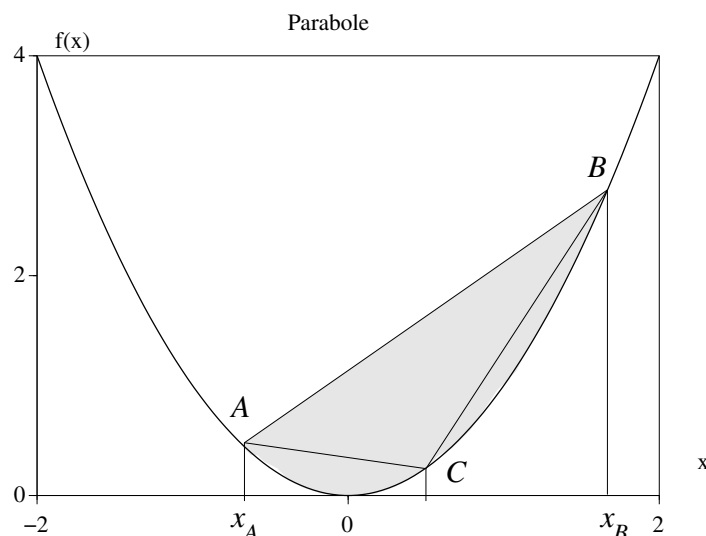


Fig. 27. *Théorème d'Archimède.* La surface d'un arc de parabole est égale aux quatre tiers du triangle inscrit.

(1.1) Proposition. *Considérons la courbe d'équation $y = x^2$. Soient A et B deux points de cette courbe. Soit C le point tel que la tangente en ce point soit parallèle à la droite passant par A et B . Alors la surface délimitée par la courbe et le segment $[A, B]$ est égale à la surface du triangle ABC , multipliée par $4/3$.*

Avec la possibilité de calculer l'aire contenue sous une courbe comme une intégrale, le problème est relativement facile. Nous laissons au lecteur les calculs qui montrent que si x_A est l'abscisse du point A et x_B l'abscisse du point B , alors le point C a pour abscisse $(x_A + x_B)/2$, la surface de l'arc de parabole est $(x_B - x_A)^3/6$ et celle du triangle ABC est $(x_B - x_A)^3/8$.

Mais Archimède ne connaissait pas les intégrales, pas même les fonctions. Remarquez qu'il n'exprime pas l'aire de l'arc de parabole comme un nombre, fonction de x_A et x_B , comme nous le faisons : son théorème énonce un rapport de surfaces. C'est aussi ainsi qu'il rappelle dans son introduction les résultats d'aires et de volumes connus avant lui.

Les arguments de mécanique auxquels Archimède fait référence sont des calculs de longueurs de leviers qu'il imagine pour équilibrer deux surfaces différentes (« donnez-moi un point d'appui et je soulèverai le monde »). Son découpage de la surface à calculer en surfaces approchées de plus en plus petites est assez remarquable. Il faudra attendre une dizaine de siècles avant qu'on fasse mieux.

2. La famille ibn Qurra.

L'idée de calculer une aire ou un volume en les découpant en petits morceaux (méthode d'exhaustion) était présente, bien avant Archimède, chez les mathématiciens grecs du

cinquième siècle avant notre ère. Elle fut reprise et raffinée, tout au long du moyen-âge par les mathématiciens arabes ; parmi eux, la famille des ibn Qurra.

Abul Hassan Thabit ibn Qurra ibn Marwan al-Sabi al-Harrani (836-901) était le père de Sinan ibn Thabit ibn Qurra (880-843), un médecin lui-même père de Ibrahim ibn Sinan ibn Thabit ibn Qurra (908-946), mathématicien et astronome comme son grand-père.

Les calculs de surface de Thabit Ibn Qurra étaient basés sur un encadrement par des sommes de rectangles majorantes et minorantes. A l'occasion de la quadrature de la parabole, il fut le premier à avoir l'idée de diviser l'intervalle d'intégration en segments inégaux. Ce que son petit-fils Ibrahim dit dans son « traité sur la quadrature de la parabole », donne une bonne idée de l'intérêt suscité par les problèmes de quadrature, et de la compétition entre les mathématiciens de ce temps.

« J'ai composé un travail sur la mesure de la parabole, dans un traité séparé. Mon grand-père avait résolu la mesure de la parabole, mais un géomètre m'a appris que al-Mahani avait une solution de ce problème, qu'il m'a communiquée, qui est plus facile que la solution de mon grand-père, et personne parmi nous n'avait une solution meilleure que celle-là. Mon grand-père avait résolu le problème en 20 propositions, il utilisait de nombreux lemmes préliminaires parmi ces propositions, et il démontrait la quadrature de la parabole par la méthode de contradiction. Al-Mahani utilisait aussi des lemmes sur les nombres pour sa démonstration, et ensuite il prouvait le résultat par la méthode de contradiction en cinq ou six propositions, de manière longue. Alors je l'ai démontré en trois propositions géométriques, sans théorème préliminaire sur les nombres. J'ai démontré la mesure de la parabole elle-même par la méthode de la preuve directe, et je n'ai pas eu besoin de la méthode de contradiction. »

La méthode d'Ibrahim était effectivement plus simple que celles de tous ses prédécesseurs, et elle ne sera surpassée qu'après la découverte du calcul intégral, par Newton et Leibniz, au 17ème siècle.

3. Calcul numérique des intégrales

Comment fait-on pour calculer numériquement une intégrale sur ordinateur ? Deux cas se présentent, selon que l'on ne connaît que certaines valeurs de la fonction (typiquement, par une table de données expérimentales), ou bien que l'on peut l'évaluer en n'importe quel point. Comme nous allons le voir, la deuxième situation est beaucoup plus favorable.

Supposons que l'on dispose de $n+1$ abscisses x_0, \dots, x_n et de $n+1$ ordonnées y_0, \dots, y_n . L'idée la plus naïve est de calculer la somme des aires de rectangles basés sur les intervalles $[x_i, x_{i+1}]$, et de hauteur y_i ou bien y_{i+1} : ce sont les *méthodes des rectangles*, à gauche, ou à droite.

$$R_g = \sum_{j=0}^{n-1} y_j(x_{j+1} - x_j) \quad \text{et} \quad R_d = \sum_{j=0}^{n-1} y_{j+1}(x_{j+1} - x_j)$$

Si on ne sait absolument rien de la fonction à intégrer, il n'y a pas de raison d'aller plus loin. Mais si on imagine un modèle, dans lequel les ordonnées y_i sont les évaluations en

x_i d'une fonction continue f , alors on obtient une meilleure précision par la *méthode des trapèzes*, qui consiste simplement à prendre la demi-somme de R_g et R_d .

$$T = \frac{1}{2} \sum_{j=0}^{n-1} (y_j + y_{j+1})(x_{j+1} - x_j)$$

Si on imagine un modèle où la fonction à intégrer f est encore plus lisse, on gagnera en interpolant par une parabole sur des triplets de points successifs. C'est la *méthode de Simpson*.

Pour donner une idée de la précision atteinte, supposons que nous disposions d'une table de 101 valeurs régulièrement espacées de la fonction sinus sur l'intervalle $[0, \pi/2]$. L'intégrale de $\sin(x)$ sur $[0, \pi/2]$ vaut 1. Voici ce que donnent la méthode des rectangles à droite et à gauche, la méthode des trapèzes, et la méthode de Simpson.

$$R_g = 0.9921255, \quad R_d = 1.0078334, \quad T = 0.9999794, \quad S = 1.0000000003$$

Quand une fonction est donnée par une expression qui permet de la calculer en n'importe quel point, il serait particulièrement inefficace de l'évaluer en 101 points régulièrement espacés, pour se ramener au cas précédent. On utilise plutôt les *méthodes de quadrature de Gauss*. Voici comment fonctionne la plus simple, celle de Gauss-Legendre. Soit une fonction f , à intégrer sur l'intervalle $[a, b]$. Quitte à effectuer un changement de variable affine, on peut se ramener au cas où $a = -1$ et $b = 1$. Les points auxquels on doit évaluer la fonction sont les racines des *polynômes de Legendre*. On peut définir ces polynômes par récurrence, par une équation différentielle, ou bien comme la dérivée n -ième d'un polynôme de degré $2n$.

$$P_n(x) = \frac{1}{2^n(n!)} \left((x^2 - 1)^n \right)^{(n)}$$

Les racines de P_n sont dans l'intervalle $[-1, 1]$, et réparties de façon symétrique par rapport à l'origine. Comme P_n est de degré n , il a n racines : notons-les x_1, \dots, x_n . À la racine x_i , on associe le « poids » ω_i :

$$\omega_i = \frac{2}{(1 - x_i)^2 (P'_n(x_i))^2}$$

On calcule ensuite la somme :

$$L_n = \sum_{i=1}^n \omega_i f(x_i)$$

Les racines des polynômes de Legendre, ainsi que les poids qui leur sont associés sont connus depuis longtemps, et inclus dans les bibliothèques de codes des langages de calcul scientifique : le calcul de L_n est donc extrêmement rapide. Vous apprendrez plus tard les raisons mathématiques pour lesquelles ce calcul donne en général une valeur très proche de l'intégrale de f . Les résultats sont spectaculaires. Voici pour la même fonction sinus entre 0 et $\pi/2$, ce que donne la méthode de Gauss-Legendre pour $n \leq 5$.

$$L_2 - 1 = -1.5 \cdot 10^{-3}, \quad L_3 - 1 = 8.1 \cdot 10^{-6}, \quad L_4 - 1 = -2.3 \cdot 10^{-8}, \quad L_5 - 1 = 3.9 \cdot 10^{-11}$$

Ainsi, en évaluant la fonction sinus en 5 points seulement, on calcule son intégrale avec une précision de l'ordre de 10^{-11} : impressionnant non ?

4. Les différentes intégrales.

L'idée qu'une intégrale pouvait être calculée comme une somme de petits rectangles était depuis longtemps bien admise. Voici ce qu'en dit Pascal (1623-1662).

« On n'entend autre chose par somme des ordonnées d'un cercle sinon la somme d'un nombre indéfini de rectangles faits de chaque ordonnée avec chacune des petites portions égales du diamètre, dont la somme [...] ne diffère de l'espace d'un demi-cercle que d'une quantité moindre qu'aucune autre donnée. »

Il faudra attendre encore deux siècles après Pascal, avant qu'on ne donne un sens rigoureux à « nombre indéfini de rectangles » et « moindre qu'aucune autre donnée » : celui d'une limite pour des subdivisions dont le pas tend vers 0.

Vers 1670, Newton et Leibniz firent simultanément la découverte fondamentale, que l'intégration et la dérivation étaient des opérations inverses l'une de l'autre. Pour calculer une aire, il suffisait désormais de connaître une primitive de la fonction qui la délimitait. Ceci fit quelque peu passer au second plan la méthode d'intégration des grecs et des arabes, consistant à découper l'aire à calculer en petits rectangles. Mais jusqu'au 19^{ème} siècle, personne ne s'était posé la question de définir la convergence d'une somme d'aires de petits rectangles, vers une intégrale.

En 1823, Cauchy fut le premier à tenter une définition rigoureuse, pour l'intégrale d'une fonction continue sur un segment. Malheureusement, la définition de la continuité qu'il donnait, confondait continuité et continuité uniforme, deux notions qui ne seront distinguées que bien après lui. Plus grave, Cauchy énonçait et « démontrait » un théorème faux, selon lequel l'intégrale de la limite d'une suite de fonctions serait toujours la limite des intégrales des fonctions de la suite.

Tout le monde était d'accord depuis longtemps sur l'intégrale des fonctions en escalier. A partir de là, il semblerait qu'il suffise de dire qu'une fonction est limite en tout point de fonctions en escalier, et de passer à la limite sur les intégrales. Ce n'est malheureusement pas si simple. L'exemple suivant aidera à comprendre pourquoi. Pour tout entier n , définissons la fonction f_n , de $[0, 1]$ dans \mathbb{R} , qui à x associe :

$$f_n(x) = \begin{cases} n & \text{si } x \in [0, 1/n] \\ 0 & \text{si } x \in]1/n, 1] \end{cases}$$

Pour tout n , f_n est une fonction en escalier, dont l'intégrale vaut 1. Pourtant, pour tout $x \in]0, 1]$, la suite $f_n(x)$ tend vers 0 (elle est nulle à partir d'un certain rang). Or la fonction nulle est d'intégrale nulle : n'en déplaise à Cauchy, l'intégrale de la limite n'est pas toujours la limite des intégrales !

En 1854, Riemann proposa un raffinement des définitions de Cauchy, qu'il rendait rigoureuses, tout en les adaptant à des fonctions qui n'étaient plus nécessairement continues. L'idée d'approcher l'intégrale par une subdivision de l'intervalle d'intégration restait la même.

En 1902, Lebesgue proposa une nouvelle approche : au lieu de subdiviser l'intervalle d'intégration, il subdivisait l'intervalle des valeurs de la fonction. La théorie de Lebesgue, si elle est plus difficile à comprendre, présente de nombreux avantages sur celle

de Riemann : les théorèmes de convergence sont plus puissants, et plus de fonctions peuvent être intégrées. L'intégrale de Lebesgue est généralement enseignée en troisième ou quatrième année d'université.

Cependant, la construction de Lebesgue n'était toujours pas parfaitement satisfaisante. A peu près dix ans après Lebesgue, Arnaud Denjoy et Oskar Perron, proposèrent chacun une variante plus générale, mais plus compliquée. Il fallut attendre la fin des années cinquante pour que Ralph Henstock et Jaroslav Kurzweil se rendent compte que non seulement les intégrales de Denjoy et Perron étaient équivalentes, mais que l'on pouvait donner de leur théorie une version beaucoup plus simple : celle qui a été présentée dans ce cours.

Entre temps, bien d'autres mathématiciens ont laissé leur nom à une définition d'intégrale, comme Darboux, Stieltjes, Daniell, Radon, Itô, etc.

5. La mésaventure de Chasles

Michel Chasles (1793-1880) avait des relations : il était titulaire d'une chaire de géométrie à la Sorbonne créée tout spécialement pour lui, et membre de l'Académie des Sciences. Il aimait l'histoire et collectionnait les documents anciens. Il était aussi d'un patriotisme fervent... et d'une naïveté désarmante.

Un jour de 1861, il reçut un certain Vrain-Lucas, qui se disait dépositaire d'un lot de vieux papiers, sur lequel il sollicitait l'avis éclairé du grand académicien. Celui-ci, émerveillé, reconnut aussitôt dans les premiers échantillons fournis, un inestimable trésor pour notre patrimoine national : des lettres du grand Pascal, dont certaines établissaient clairement que leur auteur avait découvert l'attraction universelle avant Newton : quelle fierté ! Pendant des années, Chasles acheta tout ce que Vrain-Lucas pouvait lui fournir : en tout, pas moins de vingt sept mille documents, payés cent quarante mille francs, une fortune pour l'époque. Tant pis si quelques esprits chagrins faisaient remarquer que dans les lettres de Pascal apparaissaient des données astronomiques recueillies bien après sa mort. Tant pis si dans une lettre, Galilée se plaignait de sa vue qui devenait mauvaise, alors qu'il l'avait perdue depuis quatre ans : qu'à cela ne tienne, Chasles exhibait aussitôt d'autres lettres fournies par Vrain-Lucas, qui comme par hasard répondaient exactement à ses détracteurs. A ceux qui s'étonnaient que Jules César, Socrate, Cicéron, Hérode, Vercingétorix, aient tous écrit dans le même vieux français fantaisiste, Vrain-Lucas répondait qu'un savant du temps de Charlemagne avait rassemblé cette collection de lettres anciennes, qu'il avait déposé dans une abbaye. Sept siècles plus tard, Rabelais avait traduit les lettres, et Chasles était l'heureux acheteur des copies autographes de Rabelais.

Le procès pour fraude de Vrain-Lucas eut lieu en 1870. Chasles vint lui-même exposer au tribunal comment il avait enfin découvert la supercherie : il avait fait surveiller Vrain-Lucas qui tardait à livrer trois mille nouvelles lettres qu'il avait promises (il fallait bien le temps de les écrire !). Pendant les huit ans que dura l'affaire, Chasles n'avait apparemment jamais eu le moindre doute, aveuglé qu'il était sans doute, par la fierté d'être celui qui conserverait à la France ces témoignages inestimables de son passé glorieux. Voici un extrait d'une *lettre d'Alexandre le Grand à Aristote* (sic), qui fit éclater de rire l'auditoire du tribunal.

«... Quant à ce que m'avez mandé d'aller faire un voyage au pays des Gaules, afin d'y apprendre la science des druides, non seulement vous le permettez, mais vous y engagez pour le bien de mon peuple, car vous n'ignorez pas l'estime que je fais d'icelle nation que je considère comme étant ce qui porte la lumière dans le monde. »

Références bibliographiques

- [Ba1] Robert G. Bartle. *A Modern Theory of Integration*, Graduate Studies in Mathematics (American Mathematical Society, Providence, RI), vol. 32 (2001) [suitable for advanced undergraduates].
- [Ba2] Robert G. Bartle. *Return to the Riemann integral*, American Mathematical Monthly **103** (1996) 625–632 [brief and easy to read].
- [BS] Robert G. Bartle and Donald R. Sherbert. *Introduction to Real Analysis*, Third Edition, John Wiley & Sons (2000) [suitable for advanced undergraduates].
- [Br] Jean-Yves Briend. *Intégration I, Licence de Mathématiques*, (page web)
<http://www.cmi.univ-mrs.fr/~briend/cours/cours.php>.
- [Ch] Eric Charpentier. *L'intégrale de Riemann complète*, (page web)
<http://www.math.u-bordeaux.fr/~echarpen/>.
- [Dj1] Arnaud Denjoy. *Une extension de l'intégrale de M. Lebesgue*, Comptes Rendus Acad. Sci. Paris, **154** (1912) 859–862.
- [Dj2] Arnaud Denjoy. *Sur les fonctions dérivées sommables*, Bulletin Soc. Math. France **43** (1915) 161–248.
- [DS] John DePree and Charles Swartz. *Introduction to Real Analysis*, John Wiley & Sons, New York (1988) [advanced undergraduate students, very readable].
- [Go1] Russell Gordon. *Integrals of Lebesgue, Denjoy, Perron, and Henstock*, Amer. Math. Soc. (1994) [comprehensive introduction to the subject].
- [Go2] Russell Gordon. *Historical overview of Riemann and Henstock integrals*, Real Analysis Exchange **22** (1) (1996/7) 23–33.
- [Go3] Russell Gordon. *The use of tagged partitions in elementary real analysis*, Amer. Math. Monthly **105** (1998) 107–117.
- [He] Ralph Henstock. *The efficiency of convergence factors for functions of a continuous real variable*, J. London Math. Soc., **30** (1955) 273–286.
- [Ku1] Jaroslav Kurzweil. *Generalized ordinary differential equations and continuous dependence on a parameter*, Czech. Math. J. **7** (1957) 418–446.
- [Ku2] Jaroslav Kurzweil. *Henstock-Kurzweil integration: its relation to topological vector spaces*, World Scientific Publ. Co. (2000) [for advanced specialists in integration theory].
- [LL] Jitan Lu and Peng-Yee Lee. *On singularity of Henstock integrable functions*, Real Analysis Exchange **25** (1999), 795–798.
- [Ma] Jean Mawhin. *Analyse. Fondements, Techniques, Évolutions*, De Broeck Université, Bruxelles, deuxième édition, 1997 (première édition 1992).
- [ML] Robert M. McLeod. *The Generalized Riemann Integral*, Carus Monograph (Mathematical Association of America, Washington, DC), vol. 20 (1980) [niveau introductif].
- [Mu] Pat Muldowney. *A general theory of integration in function spaces, including Wiener and Feynman integration*, Pitman Research Notes in Math. Series, John Wiley & Sons, New York (1987).
- [Pe] Oskar Perron. *Über den integralbegriff*, Sitzber. Heidelberg Akad. Wiss., Math.-Naturw. Klasse Abt. A, **16** (1914) 1–16.
- [RW] Ouvrage collectif, *Mathématiques - Tout-en-un pour la Licence - Niveau L2, Cours et exercices corrigés*, Collection dirigée par Jean-Pierre Ramis et André Warusfel, Dunod, 2007.
- [Sc] Eric Schechter. *An Introduction to the gauge integral (also known as the generalized Riemann integral, the Henstock integral, the Kurzweil integral, the Henstock-Kurzweil integral, the HK-integral, the Denjoy-Perron integral, etc.)*, (web page)
<http://www.math.vanderbilt.edu/~schectex/ccc/gauge/>.
- [Sw] Charles Swartz. *Introduction to Gauge Integrals*, World Scientific Publ. Co., (2001) [suitable for advanced undergraduates].

- [Ye] Lee Peng Yee. *Lanzhou Lectures on Henstock Integration*, World Scientific Publ. Co., Series in Real Analysis **2** (1989) [exposition of research-level material].
- [YV] Lee Peng Yee and Rudolf Výborný. *The Integral. An Easy Approach after Kurzweil and Henstock*, Cambridge University Press, Cambridge, United Kingdom (2000) [a combined textbook/research monograph, suitable for advanced undergraduates].