Polytech' Paris-Sud
PeiP1

2010/2011

Notes de cours

Maths 1 : DL et intégrales

Filippo SANTAMBROGIO

Ce texte remplace et complète les chapitres 5, 6 et 10 du Poly "officiel"

Chapitre 1

Dérivées et développements limités

Nous commençons ce chapitre en considérant certaines conséquences que le théorème des accroissements finis a en terme de dévéloppement des fonctions dérivables. Nous avons déjà vu ses conséquences en terme de comportement croissant, constant etc.

Une conséquence importante du théorème des acroissements finis est le fait qu'on peut faire le développement suivant : si f est une fonction dérivable en tout point d'un intervalle A et $x, x_0 \in A$, alors on a

$$f(x) = f(x_0) + f'(\xi)(x - x_0),$$

où ξ est un point de l'intervalle $]x_0, x[$ ou $]x, x_0[$ (ce qui dépende de $x > x_0$ ou $x_0 > x$). Celui-ci est un développement exacte de la fonction f, qui est continue car dérivable, et qui donne une idée de l'écart entre $f(x)$ et $f(x_0)$. Il s'agit en fait d'un développement qui est un peu complémentaire par rapport au suivant :

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \varepsilon(x)(x - x_0), \tag{1.1}$$

où $\varepsilon(x)$ est une quantité qui tend vers 0 lors que $x \to x_0$ (en fait, si l'on définit ε grâce à la rélation ci-dessus, qui en donne la valeur en tout point $x \neq x_0$,et on rajoute $\varepsilon(x_0) = 0$, on peut dire qu'on obtient une fonction continue). Ce dernier. développement est obtenible à partir de la défintion de dérivée, car si

$$\frac{f(x) - f(x_0)}{x - x_0} \to f'(x_0),$$

on peut dire

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) + \varepsilon(x)$$

et après réconstruire le dévéloppement souhaité.

La quantité $\varepsilon(x)(x - x_0)$ est aussi souvent notée par le symbol $o(x - x_0)$. Il est en fait commun d'écrire $o(g(x))$ si g est une fonction qui converge à zéro lorsque $x \to x_0$ et alors dans ce cas là dire $f(x) = o(g(x))$ ou "f est un petit o de g" signifie

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0.$$

Ceci équivaut, dans la notation de tout à l'heure, $f(x) = \varepsilon(x)g(x)$. Typiquement on utilise la notation du petit o avec des puissances de $x - x_0$. Par exemple on peut dire que $x^2 + x^4$ est un petit o de x pour $x \to 0$, mais non pas qu'elle est un petit o de x^2 (car la limite du ratio n'est pas nulle mais 1).

Les deux dévéloppements ont des avantages et désavantages : le premier est exacte, mais les termes qui sont connus (si l'on connait le comportement de f en x_0, c'est-à-dire $f(x_0)$ et $f'(x_0)$) s'arrêtent tout de suite et il n'y a que la partie $f(x_0)$; par contre, le deuxième utilise une fonction ε dont on sait seulement qu'elle converge vers zéro, mais il a l'avantage d'arriver un peu plus loin avec les termes connus, car il donne déjà une approximation avec une droite (la fonction $x \mapsto f(x_0) + f'(x_0)(x - x_0)$). Le but de ce genre de dévéloppement est en fait d'approcher une fonction autour d'un point x_0 en connaissant les valeurs de f et de sa dérivée. On verra dans les sections suivantes que on arrivera à faire beaucoup mieux en utilisant aussi ses dérivées d'ordre supérieur.

1.1 Dérivées d'ordre supérieur

Une question naturelle qu'on se pose quand on a une fonction f dérivable en tout point est celle de considérer la fonction $x \mapsto f'(x)$ est se demander : est-elle continue ? est-elle dérivable ?

Il est assez naturel de se convaincre que f' n'est pas forcement une fonction dérivable et l'exemple le plus facile est le suivant :

Exemple 1.1.1. Considérons $f(x) = x|x|$, qui corréspond à deux paraboles différentes, l'une avec concavité en haut, l'autre en bas, qui se joignent en 0 (car $f(x) = x^2$ pour $x \geq 0$ et $f(x) = -x^2$ pour $x \leq 0$). Il est facile de calculer $f'(x)$ (et de vérifier qu'elle existe) pour $x \neq 0$, car localement la fonction coïncide avec une parabole connue. Concernant 0, on peut calculer à la main la limite

$$f'(0) = \lim_{x\to 0} \frac{x|x| - 0}{x - 0} = \lim_{x\to 0} |x| = 0$$

et finalement on obtient l'expression générale $f'(x) = 2|x|$. Cette fonction n'est pas dérivable en 0.

Il est moins facile de comprendre si f' est forcement continue ou non.

En fait il y a ce résultat, dû à Darboux, qui montre que f' a une propriété typique des fonctions continues.

Théorème 1.1.1 (Darboux)**.** *Soit A un intervalle et $f : A \to \mathbb{R}$ une fonction dérivable en tout point : alors, si $x_0, x_1 \in A$, pour toute valeur l intermediaire entre $f'(x_0)$ et $f'(x_1)$, il existe ξ entre x_0 et x_1 tel que $f'(\xi) = l$.*

Démonstration. Par praticité, on va supposer $l = 0$, $x_0 < x_1$ et $f'(x_0) < 0 < f'(x_1)$. Ceci ne réduit pas la généralité du théorème, car il suffit de remplacr f par la fonction $x \mapsto f(x) - lx$ et, le cas écheant, de changer de signe a f ou de faire un changement de variable $x =\mapsto -x$.

Considérons donc la fonction f sur l'intervalle $[x_0, x_1]$: comme elle est dérivable, elle est continue aussi et, l'intervalle étant fermé et borné, elle admet minimum sur cet intervalle. Soit ξ ce point de minimum : on montrera $\xi \neq x_0$ et $\xi \neq x_1$, ce qui entrainera qu'il s'agit d'un point intérieur, et donc $f'(\xi) = 0$, ce qui conclut la preuve.

Pour montrer $\xi \neq x_0$ il suffit de remarquer $f'(x_0) < 0$, ce qui empèche au point x_0 d'être un point de minimum sur l'intervalle $[x_0, x_1]$, car il ne respecte pas la condition nécessaire pour les points du bord. En fait, si la dérivée au point initial est négative, le point n'est pas un minimum car juste à sa droite on a des valeurs plus petites. Pareillement on a $\xi \neq x_1$. $\square$

Une autre propriété typique des fonctions continues qui est satisfaite par la dérivée est la suivante.

Théorème 1.1.2. *Soit A un intervalle et $f : A \to \mathbb{R}$ une fonction dérivable sur $A \setminus \{x_0\}$. Supposons*

$$\lim_{x\to x_0} f'(x) = l \in \mathbb{R}.$$

Alors $f'(x_0)$ existe et est égale à l. En particulier, si f est déjà dérivable partout en A et $\lim_{x\to x_0} f'(x)$ existe alors f' est continue au point x_0.

Démonstration. On veut démontrer

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = l.$$

Pour cela on utilise le fait suivant : pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que $|f' - l| < \varepsilon$ sur $]x_0 - \delta, x_0 + \delta[$ (grâce à l'hypothèse sur la limite des dérivées au point x_0). De plus, on utilise le théorème des accroissements finis et on voit que, si $x \in]x_0 - \delta, x_0 + \delta[$ alors

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\xi) \in]l - \varepsilon, l + \varepsilon[,$$

car ξ aussi appartient à $]x_0 - \delta, x_0 + \delta[$ (comme il est un point intermédiaire entre x_0 et x). Ceci montre que la limite des taux d'accroissements est égale à l et que donc f est dérivable au point x_0 et $f'(x_0) = l$. La même démonstration montre la continuité de f' sous l'hypothèse de l'existence de la limite. $\square$

Observation 1.1.1. Le même enoncé peut s'étendre au cas d'une limite à gauche ou à droite seulement, avec égalité avec la dérivée gauche (ou droite).

Observation 1.1.2. Il est important remarquer la différence qu'il y a en théorie entre la limite des taux d'accroissement et la limite des dérivées. Si on nous demande de démontrer qu'une fonction f donnée est dérivable au point x_0 on devrait considérere les taux d'accroissement $(f(x) - f(x_0))/(x - x_0)$ et leur limite pour $x \to x_0$, et non pas la limite $\lim_{x\to x_0} f'(x)$. Pareillement, pour démontrer qu'elle n'est pas dérivable il faut démontrer que la limite des taux d'accroissement n'existe pas, et non pas celle des dérivées. L'exemple d'en bas montre aussi que cette deuxième limite peut ne pas exister alors que le première existe. Pourtant, le résultat du Théorème 1.1.2 nous aide si la limite des dérivées existe. De plus, si les limites droite et gauche des dérivées existent et sont différentes, la fonction ne sera pas dérivable (car la limite droite et la limite gauche des taux d'accroissement seront différentes).

Pourtant, même si ces propriété sont typiques des fonctions continues, il n'est pas vrai que la dérivée d'une fonction dérivable est continue, et on peut le voir de l'exemple suivant.

Exemple 1.1.2. Considérons la fonction f donnée par

$$f(x) = \begin{cases} x^2 \sin 1/x & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases}$$

Cette fonction est dérivable en tout point : elle est dérivable hors de 0 car composée et produit de fonctions dérivables, et pour vérifier la dérivabilité en 0 il suffite de calculer la limite des taux d'accroissement, en obtenant

$$f'(0) = \lim_{x \to 0} \frac{x^2 \sin 1/x - 0}{x - 0} = \lim_{x \to 0} x \sin 1/x = 0,$$

où la dernière limite peut être caluclée grâce à $-x \leq x \sin 1/x \leq x$. On a donc

$$f'(x) = \begin{cases} 2x \sin 1/x + \cos 1/x & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases}$$

On peut vérifier que cette expression n'est pas continue en $x = 0$, car il n'existe pas la limite $\lim_{x \to 0} 2x \sin 1/x + \cos 1/x$, le premier terme convergeant à zéro, mais le deuxième n'ayant pas de limite.

En vue des exemples ci-dessus, il est naturel introduire la notion de fonctions C^k, c'est-à-dire les fonction dérivables avec continuité k fois. Qu'est-ce qu'il signifie ? une fonction C^1 est une fonction qui est dérivable et telle que sa dérivée est une fonction continue. Une fonction C^2 est une fonction telle qu'on puisse faire cette opération deux fois, c'est à dire que elle est dérivable et sa dérivée est non seulement continue, mais dérivable aussi avec dérivée continue. On dit aussi C^0 a propos des fonctions continue. On peut résumer par récurrence tout ça en cette définition.

Définition 1.1.1. On dit qu'une fonction $f : A \to \mathbb{R}$ est $C^0(A)$ si elle est continue. Pour tout $k \geq 1$ on dit qu'une fonction $f : A \to \mathbb{R}$ est $C^k(A)$ si elle est dérivable et la fonction f' est $C^{k-1}(A)$.

Il n'est pas difficile établir le suivant :

Théorème 1.1.3. *Soit $k \geq 1$. Si f et g sont deux fonctions $C^k(A)$ alors $f + g$, fg sont aussi $C^k(A)$, ainsi que $1/f$ si en plus $f \neq 0$ sur A. Si g est une fonction $C^k(f(A))$ (C^k sur l'ensemble $f(A)$, qui est un intervalle si A est un intervalle) alors $g \circ f \in C^k(A)$. Si en plus $f' \neq 0$ sur A alors f^{-1} est $C^k(f(A))$ (f^{-1} existe car $f' \neq 0$ et donc, grâce au théorème de Darboux, f' est soit toujours positive soit toujours négative, donc f est strictement monotone et donc inversible).*

Démonstration. Il suffit de tout démontrer par récurrence. On sait que le résultat est vrai pour $k = 1$. Maintenant on le suppose vrai pour $k = n$ et on le montre pour $k = n + 1$. Prenons f, $g \in C^{n+1}(A)$ (donc f', $g' \in C^n(A)$) et considérons la somme et le produit. Comme on veut montrer que $f + g$, fg et $1/f$ appartiennent à $C^{n+1}(A)$ il nous suffit de montrer que leurs dérivées appartiennent à $C^n(A)$. Les dérivées qu'on regarde sont $f' + g'$, $fg' + f'g$ et $-f'/f^2$ qui sont obtenues comme sommes, produits et ratios de fonctions C^n. Donc, comme le résultat est supposé vrai pour $k = n$, on en déduit que $f' + g'$, $fg' + f'g$ et $-f'/f^2$ sont C^n et donc $f + g$ et fg sont C^{n+1}.
L'idée est la même en ce qui concerne la composition : prenons $f \in C^{n+1}(A)$ et $g \in C^{n+1}(f(A))$. La dérivée de la composée $g \circ f$ est donnée par $g' \circ f \cdot f'$ qui est un produit d'une fonction $C^n(A)$ (la fonction f', par hypothèse sur f) et d'une fonction qui est la composition de deux fonctions C^n. en appliquant le résultat dans le cas $k = n$ on trouve que $(g \circ f)'$ est C^n et donc $g \circ f$ est C^{n+1}.
Pour terminer il reste le cas de l'inverse. Si $f \in C^{n+1}(A)$ on calcule $(f^{-1})'$. On a $1/(f' \circ f^{-1})$. La fonction $f' \circ f^{-1}$ est la composition de deux fonctions C^n et est donc C^n. Là aussi on a montré, en utilisant les résultats pour $k = n$, que f^{-1} est C^{n+1}. □

Observation 1.1.3. On peut dire qu'une fonction est continue en un point x_0, on peut dire qu'elle est dérivable en ce point, mais si l'on dit qu'elle est C^1 en x_0 alors on veut en fait dire qu'elle est au moins dérivable hors de x_0 aussi et que sa dérivée est continue en x_0. Pareillement, dire qu'une fonction est C^k en un point donné implique parler de ses dérivées hors de x_0 aussi.

Sous le côté notations, on appelle dérivée seconde d'une fonction f, et on la note par f'', la dérivée de sa dérivée ($f'' = (f')'$) et plus en général on parle de dérivée $k-$ième. On écris en général f, f', f'', f''' mais après on utilise la notation $f^{(k)}$ pour la dérivée $k-$ième.

Quand on parle de combien de fois une fonction est dérivable avec continuité on appelle ça le dégré de régularité de cette fonction. Une fonction est usuellement dite régulière si elle est suffisamment dérivable pour ce qu'on veut y faire avec. Parfois on dit régulière pour dire C^∞, ce qui est introduit dans la définition suivante.

Définition 1.1.2. Une fonction f est dite $C^\infty(A)$ si $f \in C^k(A)$ pour tout $k \geq 0$.

1.2 Formule de Taylor et dévéloppements limités

On considère ici une fonction f suffisamment régulière sur A et on donne des formules de développement beaucoup plus fines de celles qu'on a donné avant. Il s'agit en général d'approcher une fonction f par des fonctions polynomiales, en utilisant les valeurs de f et de ses dérivées dans un point donné. Ce genre de développement est connu comme développement de Taylor. On présente ici deux différentes formules, qui donnent le même résultat (le même polynôme) mais l'expression du reste est différente (et les hypothèses de régularité aussi). On donnera plus tard une définition plus générale de qu'est-ce que c'est qu'un développement limité (ou DL, en général il s'agit de remplacer une fonction par un nombre limité d'expressions plus faciles, des monômes) et on verra que les dveloppements de Taylor sont des DL.

On commence par le premier, le développement de Taylor-Lagrange.

Théorème 1.2.1. *Soit $f : A \to \mathbb{R}$ une fonction $C^n(A)$ qui admet dérivées $n+1-$ième en tout point de A et x_0 un point à l'intérieur de A. Alors pour tout $x \in A$ il existe un point ξ entre x_0 et x (c'est-à-dire $\xi \in]x_0, x[$ si $x_0 < x$ ou $\xi \in]x, x_0[$ si $x < x_0$) tel que*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x - x_0)^n + \frac{1}{(n+1)!}f^{(n+1)}(\xi)(x - x_0)^{n+1}.$$

Démonstration. (HORS PROGRAMME) La démonstration qu'on va voir est un exemple de démonstration magique où l'on arrive à la thèse en considérant des fonctions qui n'ont pas grande chose à voir et en leur appliquant d'autres théorèmes, jusqu'à en obtenir une formule qui, apparamment par hasard, est utilisable dans le cadre qui nous intéresse. Dans ce cas on considérera la fonction

$$\phi(t) = f(x) - f(t) - \sum_{k=1}^{n} \frac{(x-t)^k}{k!} f^{(k)}(t) - A\frac{(x-t)^{n+1}}{(n+1)!},$$

où A est une constante à choisir pour faire en sorte d'avoir $\phi(x) = \phi(x_0)$ et appliquer le théorème de Rolle à ϕ sur l'intervalle $[x_0, x]$.

Remarquons d'abord que $\phi(x) = 0$ et $\phi(x_0) = R - A(x - x_0)^{n+1}/(n+1)!$, où R est le reste du développement de Taylor, c'est-à-dire

$$R = f(x) - f(x_0) - \sum_{k=1}^{n} \frac{(x-x_0)^k}{k!} f^{(k)}(x_0).$$

Notre but sera exactement de montrer $R = f^{(n+1)}(\xi)(x - x_0)^{n+1}/(n+1)!$ pour un ξ entre x et x_0.

Supposons maintenant qu'on ait choisi A de façon à avoir $\phi(x_0) = 0$, c'est-à-dire

$$R = A\frac{(x-x_0)^{n+1}}{(n+1)!}. \tag{1.2}$$

On peut bien appliquer le théorème de Rolle à la fonction ϕ car elle est continue et dérivable partout à l'intérieur de l'intervalle $[x_0, x]$ par conséquence du fait que $f \in C^n(A)$ (donc toutes les fonctions qui apparaissent dans la définition de ϕ sont continues) et que $f^{(n)}$ est dérivable (et donc toutes les fonctions qui apparaissent sont dérivables aussi).

Si l'on calcule la dérivée de ϕ on trouve un effet téléscopique :

$$\begin{aligned}
\phi'(t) &= -f'(t) + \sum_{k=1}^{n} \frac{(x-t)^{k-1}}{(k-1)!} f^{(k)}(t) - \sum_{k=1}^{n} \frac{(x-t)^k}{k!} f^{(k+1)}(t) + A\frac{(x-t)^n}{(n)!} \\
&= -\frac{(x-t)^n}{n!} f^{(n+1)}(t) + A\frac{(x-t)^n}{(n)!}.
\end{aligned}$$

Le théorème de Rolle nous dit qu'il existe un point ξ dans l'intervalle entre x_0 et x avec $\phi'(\xi) = 0$. Ceci implique, grâce à la formule pour ϕ', qu'on a $f^{(n+1)}(\xi) = A$.

En remplaceant A par $f^{(n+1)}(\xi)$ dans (1.2), on trouve justement

$$R = f^{(n+1)}(\xi)\frac{(x-x_0)^{n+1}}{(n+1)!},$$

ce qui était notre but. □

Comme dans le cas des accroissements finis, on peut obtenir un autre développement, plus puissant sous le point de vue hypothèses de régularité-degré du développement, mais dont le reste a une formulation moins explicite.

Théorème 1.2.2. *Soit $f : A \to \mathbb{R}$ une fonction $C^{n-1}(A)$ qui admet une dérivée $n-$ième au point $x_0 \in A$. Alors on a*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x - x_0)^n + o((x - x_0)^n).$$

Démonstration. Dans cette démonstration on va par contre voir une autre technique de preuve très typique, et notamment celle de démonstration par récurrence. On remarque que si $n = 1$ le théorème dit seulement que toute fonction dérivable admet le développeent $f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)$. Ceci descend directement de la définition de dérivée. Or, supposons le théorème vrai pour un certain rang n et démontrons-le pour $n + 1$. Prenons f qui satisfait les hypothèses de régularité du théorème avec $n+1$: il est évident que alors f' satisfait les hypothèses du théorème avec n. Appliquons donc le résultat à f' : on trouve

$$f'(x) = \sum_{k=1}^{n} \frac{f^{(k)}(x_0)}{(k-1)!}(x - x_0)^{k-1} + o((x - x_0)^n). \tag{1.3}$$

On écrit $R(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$ et on remarque que (1.3) peut s'écrire sous la forme

$$R'(x) = \varepsilon(x)(x - x_0)^n,$$

ε étant une fonction qui tend vers zéro lorsque $x \to x_0$. Ceci signifie que pour tout $\eta > 0$ il existe $\delta > 0$ tel qu'on a

$$|R'(x)| < \eta|x - x_0|^n \text{ pour tout } x \in (x_0 - \delta, x_0 + \delta).$$

L'enoncé du théorème revient à montrer $R(x) = o((x - x_0)^{n+1})$. Pour cela on remarque $R(0) = 0$ et on écrit

$$|R(x)| = |R'(\xi)||x - x_0| \le \eta|\xi - x_0|^n|x - x_0| \le \eta|x - x_0|^{n+1},$$

où l'inégalité est valable si $x \in (x_0 - \delta, x_0 + \delta)$, car dans ce cas là ξ aussi appartient au même intervalle. Ceci montre exactement

$$\lim_{x \to x_0} \frac{R(x)}{(x - x_0)^{n+1}} = 0,$$

et donc $R(x) = o((x - x_0)^{n+1})$. □

Observation 1.2.1. Il faut remarquer que la formule de Taylor-Young devient beaucoup plus facile à démontrer, et on peut la déduire de celle de Taylor-Lagrange, si on met l'hypothèse $f \in C^n(A)$ (ce qui est deux fois plus restricitve : parce que l'on demande une dérivée $n-$ième partout, pas seulement en x_0, et parce qu'on la veut continue).

En fait, il suffit d'écrire la formule de Taylor-Lagrange à l'ordre $n-1$ et remarquer que $(f^{(n)}(\xi) - f^{(n)}(x_0))(x - x_0)^n = o((x - x_0)^n)$.

Plus en général on appelle développement limité au point x_0 d'une fonction f à l'ordre n un polynôme P de dégré n tel que $f(x) = P(x) + o((x - x_0)^n)$. Ce qu'on vient de montrer est que, si la fonction f est C^n (C^{n-1} avec dérivées $n-$ièmes partout étant suffisant) alors il existe en tout point un dévéloppement limité et les coefficients du développement sont calculés d'après les valeurs des dérivées au point x_0. Il se peut que d'autres fonctions, qui ne sont pas assez régulières, admettent quand même l'existence d'un développement limité d'ordre supérieur à ce que l'on aurait pu s'attendre.

En tout cas, si une fonction admet un développement limité, ce developpement est unique parmi les polynômes du même dégré.

Théorème 1.2.3. *Soit $f : A \to \mathbb{R}$ une fonction qui admet deux dévéloppements limités du même ordre n et au même point x_0, de la forme*

$$f(x) = P(x) + o((x - x_0)^n) = Q(x) + o((x - x_0)^n),$$

avec $\deg P$, $\deg Q \le n$. *Alors* $P = Q$.

Démonstration. On déduit facilement du fait que P et Q sont des développements limités d'une même fonction f que

$$P(x) - Q(x) = o((x - x_0)^n).$$

On sait bien que tout polynôme peut être écrit en terme de puissances de $x - x_0$ à la place des puissances de x (il suffit de vérifier que 1, $x - x_0$, $(x - x_0)^2, \ldots, (x - x_0)^n$ forment une base de l'espace des polynômes de dégré inférieur ou égale à n, où de démontrer ce résultat par récurrence sur n). Donc $P - Q$ peut s'écrire sous la forme

$$P(x) - Q(x) = A(x - x_0)^k + \sum_{i=k+1}^{n} A_i(x - x_0)^i,$$

pour un coefficient $A \neq 0$ (en choisissant pour k le premier dégré avec coefficient non nul dans cette déomposition : si ce dégré n'existe pas c'est par ce que $P - Q$ est le polynôme nul, et dans ce cas là on a obtenu $P = Q$). Donc on a

$$\lim_{x \to x_0} \frac{A(x - x_0)^k + \sum_{i=k+1}^{n} A_i(x - x_0)^i}{(x - x_0)^n} = 0.$$

Pourtant, dans cette limite, la seule partie qui compte est

$$\lim_{x \to x_0} \frac{A(x - x_0)^k}{(x - x_0)^n},$$

ce qui donne comme résultat soit $\pm\infty$ si $k < n$ soit A si $k = n$. Comme cette limite ne donne jamais zéro, on a une contradiction et donc $P = Q$. □

Ce résultat d'unicité des DL a plusieurs conséquences. Une première conséquence que l'on voit concerne les fonctions paires et impaires.

Proposition 1.2.4. *Soit $A =]-a, a[$ un ouvert symétrique de $\mathbb{R}$ qui inclut 0. Supposons que $f : A \to \mathbb{R}$ soit paire (c'est-à-dire $f(x) = f(-x)$ pour tout $x \in A$). Alors tout développement limité de f en zéro est paire aussi (et donc toute puissance d'exposant impaire a coefficient nul).*

Si par contre on suppose que f soit impaire (c'est-à-dire $f(x) = -f(-x)$ pour tout $x \in A$) on peut dire alors que ses développements sont impaires aussi.

Démonstration. Soit P un polynôme de dégré n tel que $f(x) = P(x) + o(x^n)$ et f paire. On a donc

$$f(x) = f(-x) = P(-x) + o((-x)^n) = P(-x) + o(x^n).$$

Le polynôme $x \mapsto P(-x)$ est donc un autre DL du même ordre de f autour de zéro. Par conséquent $P(x) = P(-x)$ et donc P est paire.

Si f est impaire on obtient

$$f(x) = -f(-x) = -P(-x) - o((-x)^n) = -P(-x) + o(x^n),$$

d'où l'on tire $P(x) = -P(-x)$. □

1.3 Exemples et applications

On va se rendre compte que l'unicité du DL est à la base de toute méthode de calcul du DL différente de la simple dérivation itérée. On en donnera tout de suite un exemple, mais cet exemple sera précedé par quelques définitions concernant l'équivalence et la négligeabilité des fonctions tendant vers zéro.

Définition 1.3.1. Soient f, $g : A \to \mathbb{R}$ deux fonctions définies au voisinage du point $x_0 \in A$ et telles que

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0,$$

on dit alors que f et g sont équivalentes si

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = l \in \mathbb{R} \setminus \{0\},$$

et on dit que f est negligeable devant g si

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = l \in \mathbb{R} \setminus \{0\}.$$

Si f est équivalente à g on écrit $f(x) = O(g(x))$ et si f est négligeable devant à g on écrit $f(x) = o(f(x))$. La relation d'équivalence étant une relation d'équivalence, elle est symétrique et donc $f(x) = O(g(x))$ équivaut à $g(x) = O(f(x))$.

Observation 1.3.1. On dit parfois que deux fonctions sont équivalentes en utilisant une définition beaucoup plus stricte, c'est-à-dire que la limite fasse 1 à la place de n'importe quel nombre fini et non nul. Aussi, on a parfois envie d'étendre note définition au cas où la limite n'existe pas : dans ce cas là on dit que les deux fonctions sont équivalentes si le quotient est localement borné entre deux constantes c et C avec $0 < c < C < +\infty$ (ou $0 > c > C > -\infty$).

Non seulement, souvent par abus de notation on dit $f(x) = O(g(x))$ pour dire "f est au pire de l'ordre de g", c'est-à-dire "soit $f(x) = O(g(x))$, soit $f(x) = o(g(x))$". C'est un peu comme quand on dit "dénombrable" pour dire "soit dénombrable soit fini".

On connait bien par exemple les limites suivantes, qu'on peut démontrer par voie de considérations géometriques et de manipulations algébriques :

$$\lim_{x\to 0} \frac{\sin x}{x} = 1 \quad ; \quad \lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2},$$

ce qui nous donne les équivalences entre $\sin x$ et x et entre $1 - \cos x$ et x^2 au voisinage de 0.

Ce qui nous intéresse en vue des DL est le suivant :

Observation 1.3.2. Si f et g sont équivalentes au voisinage d'un certain point x_0, les écritures $o(f(x))$ et $o(g(x))$ (eu voisinage du même point x_0) coïncident.

1.3.1 Calculs de DL et de limites

En utilisant tout ça, on peut voir les exemples suivants.

Exemple 1.3.1. Considérons la fonction

$$f(x) = \sin(2\sin(x/2))$$

et cherchons-en le DL jusqu'à l'ordre 4 en 0. On commence par le sinus à l'exterieur et on se souvient du DL $\sin y = y - y^3/3 + o(y^4)$. On va choisir après $y = 2\sin(x/2)$. Pourquoi ? pour réaliser la situation de la fonction qu'on veut développer. Pourquoi a-t on pris le DL du sinus jusqu'à l'ordre 4 ? car le y qu'on va utiliser sera équivalent à x et donc si l'on cherche un reste $o(x^4)$ on veut un reste du genre $o(y^4)$. On a donc

$$\sin(2\sin(x/2)) = 2\sin(x/2) - \frac{1}{3}\left(2\sin(x/2)\right)^3 + o((2\sin(x/2))^4).$$

On peut remplacer $o((2\sin(x/2))^4)$ par $o(x^4)$ et les expressions avant par leurs DL. Celui de $2\sin(x/2)$ est facile, mais pour le développement de son cube il y a plusieurs manières de procéder. On pourrait développer le sinus jusqu'à l'ordre 4 et puis en prendre le cube. Ceci va marcher sûrement : il est toujours suffisant de développer les fonctions qu'on a au dedans jusqu'au même ordre qu'on souhaite réaliser pour la fonction composée. Pourtant, on peut faire mieux. Dans l'expression

$$\left(2\sin(x/2)\right)^3 = (2\sin(x/2)) \cdot (2\sin(x/2)) \cdot (2\sin(x/2)) = (x + o(x^2)) \cdot (x + o(x^2)) \cdot (x + o(x^2))$$

on a un produit avec un terme x^3 et d'autres termes qui ont au moins deux fois x et une fois $o(x^2)$. Comme $x \cdot x \cdot o(x^2) = o(x^4)$, tous ces termes peuvent être negligés dans notre développement et mis ensemble au $o(x^4)$ à la fin. On a donc

$$f(x) = \sin(2\sin(x/2)) = x - \frac{2}{6}\left(\frac{x}{2}\right)^3 + o(x^4) - \frac{1}{6}x + o(x^4) = x - \frac{5}{24}x^3 + o(x^4).$$

Exemple 1.3.2. Considérons la fonction

$$f(x) = \ln(\cos x + x^3)$$

et cherchons-en le DL jusqu'à l'ordre 5 en 0. On écrit ça comme

$$f(x) = \ln(1 + (\cos x + x^3 - 1)) = \ln(1 + y) \text{ avec } y = g(x) = \cos x + x^3 - 1.$$

Comme ça on a une fonction g telle que $g(0) = 0$ et on la compose avec la fonction $\ln(1 + y)$, dont on connaît bien le développement, et notamment on a

$$\ln(1 + y) = y - \frac{y^2}{2} + o(y^2) = y - \frac{y^2}{2} + \frac{y^3}{3} + o(y^3) = y - \frac{y^2}{2} + O(y^3).$$

On utilisera la dernière version. Pourquoi s'arrêt-on au reste $O(y^3)$? Car c'est la fonction g qui sera de l'ordre de x^2, et précisément

$$g(x) = -\frac{x^2}{2} + x^3 + \frac{x^4}{24} + o(x^5).$$

On a donc $g(x) = O(x^2)$ (pour calculer la puissance dont une fonction donnée est grand O il suffit de trouver le développement de Taylor et après regarder le premier terme non nul). Or, si $y = g(x) = O(x^2)$, on a $O(y^3) = O(x^6) = o(x^5)$. Il suffit donc de s'arrêter à ce développement là dans le logarithme. Évidemment pour choisir l'ordre où s'arrêter dans la fonction à l'extérieur il faut regarder l'ordre d'équivalence de la fonction à l'intérieur de la composition. En général il est quand même toujours suffisant développer la fonction à l'extérieur aussi jusqu'à l'ordre final souhaité (dans ce cas, 5), mais parfois on peut économiser des calculs. Pareil pour la fonction à l'intérieur : il est toujours suffisant d'arriver jusqu'à l'ordre souhaité, mais parfois on peut faire mieux. Notamment, si dans le développement de la fonction à l'extérieur il y a une partie y (comme dans ce là), on y peut rien faire : on est forcé à développer jusqu'au but car de toute façon on devra copier le développement de $y = g(x)$ dans le résultat final. Si par contre on commence par y^2 et on est dans le cas $g(0) = 0$ (donc g commence par x) on se retrouvera toujours à calculer des termes au moins $g^2(x)$ avec g qui commence par x : chaque terme de g sera multiplié dans le carré par x au moins et on peut donc s'arrêter un ordre avant. C'est un peu ce qu'on va voir pour calculer $g^2(x)$ et son développement dans ce cas-ci. On a

$$g^2(x) = \left(-\frac{x^2}{2} + x^3 + \frac{x^4}{24} + o(x^5)\right)^2.$$

On commence à multiplier chaque terme fois tous les autres mais on néglige tout terme qui dépasse l'ordre 5 (par exemple x^3 fois x^3 ou x^4 fois x^2 ou les petits o). On obtient

$$g^2(x) = \frac{x^4}{4} - x^5 + o(x^5)$$

et on s'aperçoit aussi que les termes en x^4 et $o(x^5)$ n'ont joué aucun rôle (car, en étant obligé de les multiplier toujours fois x^2 au moins, on dépassait sûrement l'ordre 5). Plus en général, si je dois calculer le DL à l'ordre n d'un produit du genre $g_1(x)g_2(x)$ avec $g_1(x) = O(x^k)$ et $g_2(x) = O(x^h)$ on peut se contenter de développer g_1 jusqu'à l'ordre $n - h$ (car on multipliera toujours fois x^h au moins) et g_2 jusqu'à $n - k$ (car on multipliera toujours fois x^k au moins). Quand on met tout ça ensemble on obtient

$$f(x) = \ln(\cos x + x^3) = -\frac{x^2}{2} + x^3 + \frac{x^4}{24} + o(x^5) - \frac{1}{2}\left(\frac{x^4}{4} - x^5 + o(x^5)\right) = -\frac{x^2}{2} + x^3 - \frac{x^4}{12} + \frac{x^5}{2} + o(x^5).$$

Un autre exemple intéressant est le suivant.

Exemple 1.3.3. Considérons la fonction

$$f(x) = \tan x$$

et cherchons-en le DL jusqu'à l'ordre 3 en 0. Mais on veut le trouver rapidement, et sans regarder les tableaux de développements connus. Cependant, on admet de connaître les développements du sinus et du cosinus, qui sont plus simples à s'en souvenir. Non seulement, on admet l'existence du DL à l'ordre 3 de la tangente, car on sait qu'elle est une fonction $C^\infty(]-\pi/2, \pi/2[)$.

On écrit donc $\tan x \cos x = \sin x$. On suppose $\tan x = A + Bx + Cx^2 + Dx^3 + o(x^3)$. Écrire les trois développements à l'ordre 3 donnera quatre équations qui permettront de déduire les valeurs des snstantes inconnus A, B, C et D. Cependant, faut pas exagérer : on sait bien que A vaut 0 car le premier terme du DL en un point x_0 est toujours la valeur $f(x_0)$ et dans ce cas là on $\tan 0 = 0$. En plus, on connaît la valeur de B aussi, car le coefficient du terme linéaire coïncide toujours avec $f'(x_0)$. Ceci peut être vu en écrivant le développement (1.1) et en utilisant l'unicité du DL. Et comme on sait calculer la dérivée de la tangente en 0 est elle vaut 1, on a

$$\left(x + Cx^2 + Dx^3 + o(x^3)\right)\left(1 - \frac{1}{2}x^2 + o(x^3)\right) = x - \frac{1}{6}x^3 + o(x^3).$$

En développant les produits (et en négligeant les termes en $o(x^3)$) on trouve

$$x + Cx^2 + Dx^3 - \frac{1}{2}x^3 + o(x^3) = x - \frac{1}{6}x^3 + o(x^3).$$

Ceci signifie que la même fonction, le sinus, admet deux développements limités en 0, et donc ils coïncident. Ceci implique $C = 0$ et $D - 1/2 = -1/6$, ce qui donne $C = 1/3$. On a donc

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

On peut voir ce qu'il se passe avec les fonctions réciproques.

Exemple 1.3.4. Cherchons les DL jusqu'à l'ordre 3 des fonctions $f(x) = \arcsin x$ et $g(x) = \arctan x$. On sait $\sin(\arcsin x) = x$. Prenons les DL limités à l'ordre 3 de toute expression. Évidemment on ne connaît pas encore celui de l'arcsinus, mais on peut savoir qu'il est de la forme $x + Ax^3 + o(x^3)$. Ceci on le sait car la fonction arcsinus vaut zéro en zéro, sa dérivée vaut 1, et il s'agit d'une fonction impaire (donc il n'y a pas de partie de dégré 2). On a donc

$$\arcsin x - \frac{1}{6}(\arcsin x)^3 + o((\arcsin x)^3) = x; \quad x + Ax^3 + o(x^3) - \frac{1}{6}\left(x + Ax^3 + o(x^3)\right)^3 + o((\arcsin x)^3) = x.$$

Grâce au fait que $\arcsin x$ est équivalent à x au voisinage de 0 (ce qui est une conséquence du fait que la fonction vaut zéro en zéro mais sa dérivée non), on peut remplacer le $o((\arcsin x)^3)$ par $o(x^3)$. En plus, en développant et en négligeant les termes d'ordre trop élevé, on obtient

$$x + Ax^3 - \frac{1}{6}x^3 + o(x^3) = x,$$

d'où l'on tire $A = 1/6$.

De manière toute à fait équivalente on suppose $\arctan x = x + Bx^3 + o(x^3)$ et on trouve $B = -1/3$.

On applique tout ça au calcul d'une limite.

Exemple 1.3.5. On veut calculer la limite

$$\lim_{x\to 0} \frac{\tan x - \arctan x}{\sin x - \arcsin x}.$$

On remarque que dénominateur et numérateur, les deux, valent 0 et qu'on a donc une condition d'indétermination 0/0. Ceci correspond à faire un développement à l'ordre 0, c'est-à-dire à ne regarder que les valeurs des fonctions, mais nous on peut faire mieux. On regarde le développement à l'ordre 1 et là aussi on voit que, comme toutes les fonctions $\tan x, \arctan x, \sin x, \arcsin x$ sont de la forme $x + o(x)$, on tombe encore sur une expression qui ne nous aide pas. On a en effet

$$\lim_{x\to 0} \frac{o(x)}{o(x)},$$

ce qu'on n'est pas capables de traiter car en ce qui concerne $o(x)$ on sait seulement $\lim_{x\to 0} o(x)/x = 0$ mais on ne sait pas diviser un $o(x)$ par un autre. Il faut donc continuer avec les développements jusqu'à un résultat meilleur. On le trouve à l'ordre 3, où l'on a

$$\frac{\tan x - \arctan x}{\sin x - \arcsin x} = \frac{x - \frac{1}{3}x^3 + o(x^3) - x - \frac{1}{3}x^3 + o(x^3)}{x - \frac{1}{6}x^3 + o(x^3) - x - \frac{1}{6}x^3 + o(x^3)} = \frac{-\frac{2}{3}x^3 + o(x^3)}{-\frac{1}{3}x^3 + o(x^3)} = \frac{-\frac{2}{3} + \frac{o(x^3)}{x^3}}{-\frac{1}{3} + \frac{o(x^3)}{x^3}}.$$

Maintenant on peut calculer les limites des nouveaux dénominateur et numérateur séparément et on obtient

$$\lim_{x\to 0} \frac{\tan x - \arctan x}{\sin x - \arcsin x} = -2.$$

En général on développe séparément dénominateur et numérateur jusqu'au premier ordre qui ne donne pas un résultat banal (ici, l'ordre 3). Après on regarde les exposants qu'on a obtenus : par exemple

$$\lim_{x\to 0} \frac{x^5 + o(x^5)}{x^3 + o(x^3)} = \lim_{x\to 0} x^2 \frac{1 + \frac{o(x^5)}{x^5}}{1 + \frac{o(x^3)}{x^3}} = 0,$$

alors que, avec la même méthode (toujours diviser par la puissance qui apparaît dans chaque terme de la fraction) on a

$$\lim_{x\to 0} \frac{x^3 + o(x^3)}{x^5 + o(x^5)} = +\infty; \quad \lim_{x\to 0^\pm} \frac{x^3 + o(x^3)}{x^4 + o(x^4)} = \pm\infty; \quad \lim_{x\to 0} \frac{Ax^n + o(x^n)}{Bx^n + o(x^n)} = \frac{A}{B}.$$

1.3.2 Minima, maxima et DL

Les calculs des limites ne sont pas les seules applications possibles et intéressantes des DL. Les développements limités sont aussi très importants en vue de la détection des points des minima et maxima locaux. On a en fait le théorème suivant, qu'on donne dans le cas d'une fonction C^∞.

Théorème 1.3.1. *Soit* $f :]a,b[\to \mathbb{R}$ *une fonction* $C^\infty(]a,b[)$ *et* $x_0 \in]a,b[$. *Soit* $n = \min\{k \geq 1 : f^{(k)}(x_0) \neq 0\}$. *alors :*

– *si* n *est impaire* x_0 *n'est ni un minimum local ni un maximum local ;*

– *si n est paire et $f^{(n)}(x_0) > 0$ alors x_0 est un point de minimum local ;*
– *si n est paire et $f^{(n)}(x_0) < 0$ alors x_0 est un point de maximum local.*

Démonstration. Grâce au développement de Taylor on a

$$f(x) = f(x_0) + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + o((x - x_0)^n)$$

et donc

$$f(x) - f(x_0) = (x - x_0)^n \left(\frac{f^{(n)}(x_0)}{n!} + \frac{o((x - x_0)^n)}{(x - x_0)^n} \right).$$

Le signe à droite est donné par le signe de $(x - x_0)^n$ fois le signe de la parenthèse, qui ne dépende localement que du signe de $f^{(n)}(x_0)$, car $f^{(n)}(x_0) \neq 0$ et $\lim_{x \to x_0} o((x - x_0)^n)/(x - x_0)^n = 0$.

Dans le cas n impair le signe du premier facteur varie selon la position de x par rapport à x_0 alors que l'autre ne varie pas. Par conséquent, au voisinage de x_0, la quantité $f(x) - f(x_0)$ est d'un côté positive et de l'autre négative, ce qui implique que x_0 ne peut être ni un point de minimum local (car dans ce cas là on aurait eu $f(x) - f(x_0) \geq 0$), ni de maximum local (car on aurait eu $f(x) - f(x_0) \leq 0$).

Dans le cas n pair, par contre, le premier facteur est e tout cas positif. Ceci implique que $f(x) - f(x_0)$ a le même signe de la parenthèse, et donc on a un minimum ($f(x) - f(x_0) \geq 0$) dans le cas $f^{(n)}(x_0) > 0$ et un maximum ($f(x) - f(x_0) \leq 0$) dans le cas $f^{(n)}(x_0) < 0$. ☐

Dans ce théorème on n'a pas supposé $\{k \geq 1 : f^{(k)}(x_0) \neq 0\} \neq \emptyset$ simplement car l'énoncé même du théorème portait sur n, son minimum : si cet ensemble est vide alors n n'existe pas et donc il n'y a rien à démontrer. Il y a aussi un vice-versa un peu plus faible. cette version porte directement sur x_0 et déduit quelque chose sur n, et c'est pour ça qu'on suppose, pour clarifier, que n existe (c'est -à-dire que l'ensemble dont on prend le minimum soit non vide).

Théorème 1.3.2. *Soit $f :]a, b[\to \mathbb{R}$ une fonction $C^\infty(]a, b[)$ et $x_0 \in]a, b[$. Supposons $\{k \geq 1 : f^{(k)}(x_0) \neq 0\} \neq \emptyset$ et soit $n = \min\{k \geq 1 : f^{(k)}(x_0) \neq 0\}$. Alors :*
– *si x_0 est un point de minimum local alors n est paire x_0 et $f^{(n)}(x_0) > 0$;*
– *si x_0 est un point de maximum local alors n est paire x_0 et $f^{(n)}(x_0) < 0$.*

Démonstration. Il suffit d'utiliser le théorème de tout à l'heure pour en déduire qu'il n'est pas possible que n soit impaire, ni que n soit paire mais que la dérivée ait le mauvais signe : dans ce cas là on aurait un point qui est au même temps maximum local et minimum local, mais cela implique que la fonction est localement constante, c'est à dire que toute dérivée s'annule en x_0. Et cela est en contradiction avec les hypothèses. ☐

Ces théorèmes correspondent un peu à ce qu'on a dans le cas de la dérivée première (si un point est un extremum local alors la dérivée vaut zéro, et cela peut s'étendre à toute dérivée d'ordre impair, dans un certain sens) et aux théorèmes bien connus sur la dérivée seconde (si elle est strictement positive le point réalise un minimum local, si le point réalise un minimum locale alors elle est ≥ 0 : ici l'inégalité est stricte mail reste quand même la possibilité que les dérivées s'annule toutes, ce qui correspond au cas d'égalité).

Or, dans notre intuition on ne connaît pas trop de fonctions qui ont toutes les dérivées nulles en un point donné et qui ne sont pas la fonction nulle. Par exemple on sait que cela ne peut pas être vrai pour un polynôme, car un polynôme coïncide avec son développement de Taylor (c'est-à-dire le DL d'ordre n d'un polynôme d'ordre n a reste nul, et cela peut être vérifié en utilisant l'unicité du *DL*). Pourtant, il y a l'exemple suivant.

Exemple 1.3.6. Considérons la fonction

$$f(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ e^{-1/x} & \text{si } x > 0. \end{cases}$$

Cette fonction est une fonction $C^\infty(\mathbb{R})$ et $f^{(k)}(0) = 0$ pour tout $k \in \mathbb{N}$.

Pour s'en rendre compte, on commence par voir que $\lim_{x \to 0^+} f(x) = 0$ car $1/x \to +\infty$ et donc $e^{-1/x} \to e^{-\infty} = 0$. Ceci montre la continuité. Pour les dérivées, il suffit de calculer la dérivée à droite et on a

$$f'(x) = \frac{e^{-1/x}}{x^2} \quad \text{et} \quad \lim_{x \to 0^+} \frac{e^{-1/x}}{x^2} = 0,$$

car cette dernière limite est une forme indéterminée $0/0$ mais le zéro du numérateur l'emporte sur celui du dénominateur grâce à son comportement exponentiel.

Plus en général, on peut montrer que la dérivée $k-$ième de f est de la forme

$$f^{(k)}(x) = \frac{P(x)}{Q(x)} e^{-1/x}$$

pour des polynômes P et Q (ceci peut être montré par récurrence). Par conséquent, on a toujours

$$\lim_{x \to 0^+} f^{(k)}(x) = 0,$$

ce qui montre au même temps que la fonction est C^k (car de l'autre côté toute dérivée est nulle, la fonction étant constante) et que $f^{(k)}(0) = 0$.

Chapitre 2

Intégrales

2.1 Fonctions Intégrables

Supposons qu'une voiture roule à une certaine vitesse v le long d'une autoroute au départ de Paris et qu'on veuille savoir à quelle distance de Paris elle se trouver après un temps T : la réponse est évidemment vT. Supposons maintenant que la voiture n'ait pas une vitesse constante mais qui elle change de temps en temps sa vitesse et que, par exemple, elle ait une vitesse v_1 pendant un certain temps T_1, puis une nouvelle vitesse v_2 pendant un temps T_2 (c'est-à-dire entre l'instant T_1 et l'instant $T_1 + T_2$) et enfin une vitesse v_3 à partir de l'instant $T_1 + T_2$: si l'on veut calculer la distance où elle se trouvera après un temps t la réponse sera

$$v_1 t \text{ si } t \leq T_1; \quad v_1 T_1 + v_2(t - T_1) \text{ si } T_1 \leq t \leq T_1 + T_2; \quad v_1 T_1 + v_2 T_2 + v_3(t - (T_1 + T_2)) \text{ si } T_1 + T_2 \leq t.$$

Et si la vitesse changeait continuellement ? l'idée est que pour calculer la réponse dans le cas d'une vitesse qui change de temps en temps il faut faire une somme de résultats du type “vitesse fois temps" et que pour une vitesse en variation perpétuelle il faudra une nouvelle technique, ce qui donnera lieu aux intégrales. Pourtant, elle ne sera pas si nouvelle que ça, comme en fait sa définition sera basée justement sur l'idée d'approcher une vitesse en variation par des cas où elle varie de temps en temps (mais très souvent, pour réaliser une meilleure approximation).

On commence donc par la définition de “fonction qui varie seulement de temps en temps".

Définition 2.1.1. On appelle fonction étagée toute fonction $f : [a, b] \to \mathbb{R}$ telle qu'il existe un nombre fini de points $(t_i)_{i=1,\dots,n}$ avec $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$ tels que f est constante sur tout intervalle de la forme $I_i =]t_i, t_{i+1}[$ avec $i < n$. On dit que les points t_i donnent lieu à une partition adaptée à la fonction étagée f. On écrira $\mathcal{E}([a, b])$ pour l'ensemble des fonctions étagées sur l'intervalle $[a, b]$.

Une fonction étagée sera donc de la forme

$$f(x) = \begin{cases} k_0 & \text{si } x = a, \\ m_1 & \text{si } a = t_0 < x < t_1, \\ k_1 & \text{si } x = t_1, \\ m_2 & \text{si } t_1 < x < t_2, \\ k_2 & \text{si } x = t_2, \\ \dots & \\ k_{n-1} & \text{si } x = t_{n-1}, \\ m_n & \text{si } t_{n-1} < x < t_n = b, \\ k_n & \text{si } x = b. \end{cases} \tag{2.1}$$

Souvent on considère des intervalles semi-fermés genre $[t_i, t_{i+1}[$ de façon à avoir $k_i = m_{i+1}$ mais ce qu'on veut souligner par cette définition est que la valeur de f aux points de séparation entre les intervalles n'est pas importante.

Évidemment toute fonction étagée f admet, par définition, au moins une partition adaptée mais elle en admet en général plusieurs (tout intervalle où la fonction est constante peut en fait être ultérieurement séparée, tout en gardant son caractère adapté).

En suivant l'esprit des relations qu'on a présentés entre vitesse et position, on peut définir l'intégrale d'une fonction étagée :

Définition 2.1.2. Si $f : [a,b] \to \mathbb{R}$ est une fonction étagée de la forme (2.1) on définit l'intégrale de f sur l'intervalle $[a,b]$ la quantité

$$\int_a^b f(t)dt := \sum_{i=1}^n m_i(t_i - t_{i-1}).$$

Cette quantité est bien définie (elle ne dépende pas de la partition adaptée choisie pour représenter f).

Il faut quand même démontrer que l'intégrale d'une fonction étagée est bien définie.

Démonstration. Considérons deux partitions différentes $T = ((t_i)_i)$ et $S = ((s_j)_j)$ et supposons que la deuxième soit plus fine que la première : cela signifie que tout intervalle du type $]t_i, t_{i+1}[$ est divisé en plusieurs intervalles du type $]s_j, s_{j+1}[$ pour $j = j_i^-, \dots, j_i^+$. La fonction f valant m_i sur tout le gros intervalle on obtiendrait un terme $m_i(t_{i+1} - t_i)$ dans l'expression de l'intégrale due à la première partition, égalisée par un terme $\sum_{j=j_i^-}^{j_i^+} m_i(s_{j+1} - s_j)$. Ceci montre que la valeur de l'intégrale ne change pas si l'on passe d'une partition à une partition plus fine. Comme deux partitions distinctes admettent toujours une partition qui est plus fine des deux (il suffit de prendre la partition composée par la réunion des points des deux partitions), on en déduit que la valeur de l'intégrale est indépendante de la partition adaptée choisie. □

Il est facile vérifier la propriété suivante :

Proposition 2.1.1. *Soient f et g deux fonctions en $\mathcal{E}([a,b])$: on a alors*

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

Démonstration. En passant à une partition plus fine, on peut supposer d'écrire f sous la forme (2.1) et g sous une forme analogue pour des coefficients m_i' qui remplacent les m_i et la même partition $(t_i)_i$. On a donc

$$\int_a^b (f+g)(t)dt = \sum_{i=1}^n (m_i + m_i')(t_i - t_{i-1}) = \sum_{i=1}^n m_i(t_i - t_{i-1}) + \sum_{i=1}^n m_i'(t_i - t_{i-1}) = \int_a^b f(t)dt + \int_a^b g(t)dt.$$

□

Le but de la théorie de l'intégration (ce type d'intégration s'appelant notamment "Intégration de Riemann" est de donner un sens à l'expression $\int_a^b f(t)dt$ quand f n'est pas n'est pas étagée. On s'occupera de le faire dans les prochaines sections pour les fonctions continues, par une procédure d'approximation supérieure et inférieure.

2.1.1 L'intégrabilité des fonctions continues et l'uniforme continuité

Avant de nous occuper de définir une intégrale pour les fonctions continues, nous avons besoin d'étudier des propriétés supplémentaires des fonctions continues sur un intervalle $[a,b]$.

Définition 2.1.3. Une fonction $f : A \to \mathbb{R}$ est dite uniformement continue si pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que tout couple de points $(x,y) \in A \times A$ satisfaisant $|x-y| < \delta$ satisfait aussi $|f(x) - f(y)| < \varepsilon$.

C'est quoi la différence entre la définition de continuité et d'uniforme continuité ? d'abord la continuité est une propriété qui concerne une fonction f et un point x_0, alors que l'uniforme continuité ne concerne que f. De plus, comme dans la définition de continuité le point x_0 a été fixé, il a le droit d'intervenir dans le choix du δ, alors qu'ici δ ne doit dépendre que de ε. Évidemment si une fonction est uniformément continue, alors elle est continue sur A (la notion d'uniforme continuité est plus forte que celle de continuité en tout point de A). Le contraire n'est pas vrai, comme on peut voir de l'exemple suivant.

Exemple 2.1.1. La fonction $f : \mathbb{R} \to \mathbb{R}$ donnée par $f(x) = x^2$ est continue mais elle n'est pas uniformément continue.

Pour voir ceci, il suffit d'estimer les variations de f entre un point x et un point $x+h$ (prenons par simplicité $x, h > 0$) : $f(x+h) = x^2 + 2xh + h^2 \geq f(x) + 2xh \geq f(x)$. Si l'on suppose par l'absurde que f ést uniformément continue on aurait $f(x+h) - f(x) \leq \varepsilon$ pour tout $h < \delta$. Mais la quantité $2xh$ peut être rendue aussi grande qu'on veut, si x est très grand, et va sûrement dépasser ε. Par contre, la même démonstration n'aurait pas marché si on se restreignait à un ensemble borné de valeurs de x.

Sinon, il existe pas mal de classes de fonctions qui sont uniformément continues et il s'agit en général de classes de fonctions où la continuité uniforme a été mieux "quantitifiée". Par exemple, toute fonction Lipschitzienne est uniformément continue car, si L est la constante de Lipschitz d'une fonction f, on satisfait toujours la définition d'uniforme continuité en choisissant $\delta = \varepsilon/L$ (nous rappelons ici qu'une fonctions Lipschitzienne est une fonction satisfaisant $|f(x) - f(y)| \leq L|x - y|$ et que la meilleure constante L qui fait marcher cette inégalité pour tout x et y est dite Constante de Lipschitz).. Une classe plus générale est la classe des fonctions Hölderiennes (là aussi Hölder est le mec qui a - peut-être - étudié en premier les propriétés de ces fonctions).

Définition 2.1.4. Si $\alpha \in]0,1]$ une fonction $f : A \to \mathbb{R}$ est dite Hölderienne d'exposant α (ou $\alpha-$Hölderienne) si il existe une constante C telle que

$$|f(x) - f(y)| \leq C|x - y|^\alpha \text{ pour tout couple } (x, y) \in A \times A.$$

Évidemment dans le cas $\alpha = 1$ on retombe sur la définition des fonctions Lipschitziennes.

Un exemple de fonction Hölderienne qui n'est pas Lipschitzienne est le suivant.

Exemple 2.1.2. La fonction $f : [0,1] \to \mathbb{R}$ donnée par $f(x) = x^\alpha$ est Hölderienne d'exposant α mais elle n'est pas Lipschitzienne.

Pour voir ceci, il suffit d'utiliser le fait que la fonction f, en tant que fonction concave, est sous-additive, ce qui signifie $f(x + y) \leq f(x) + f(y)$: on a donc, si $x < y$

$$f(x) = x^\alpha < f(y) = y^\alpha \leq f(y - x) + f(x) = |y - x|^\alpha + f(x),$$

ce qui implique $|f(y) - f(x)| \leq |y - x|^\alpha$.

Pour voir qu'elle n'est pas Lipschitzienne il suffit de voir que sa dérivée n'est pas bornée. La fonction f est en fait dérivable sur $]0,1[$ mais sa dérivée est donnée par $\alpha x^{\alpha-1}$ et on a donc $\lim_{x\to 0^+} f'(x) = +\infty$, ce qui implique que f' n'est pas bornée (et donc f n'est pas Lipschitzienne sur $]0,1[$, ni a fortiori sur $[0,1]$.

Observation 2.1.1. Pourquoi parle-t-on de fonctions Höldériennes pour $\alpha \leq 1$ et pas pour $\alpha > 1$? Simplement car les fonctions Höldériennes pour $\alpha > 1$ sont triviales : toute fonction $f : I \to \mathbb{R}$ (I étant un intervalle de $\mathbb{R}$) satisfaisant $|f(x) - f(y)| \leq C|x - y|^\alpha$ pour un certain $\alpha > 1$ est en fait constante ! ! On peut le voir facilement en calculant sa dérivée :

$$|f'(x)| = \lim_{y\to x} \frac{|f(x) - f(y)|}{|x - y|} \leq \lim_{y\to x} \frac{C|x - y|^\alpha}{|x - y|} = 0,$$

ce qui démontre que la fonction est partout dérivable et la dérivée est partout nulle.

Pour une fonction uniformément continue, il est intéressant de définir ce qu'on appelle son "module de continuité".

Définition 2.1.5. Pour toute fonction $f : A \to \mathbb{R}$ on appelle module de continuité de f la fonction $\omega_f : \mathbb{R}^+ \to \mathbb{R}^+$ donnée par

$$\omega_f(a) := \sup \{|f(x) - f(y)| \ : \ x, y \in A, \ |x - y| \leq a\}.$$

Le fait que f soit uniformément continue équivaut à $\lim_{a\to 0} \omega_f(a) = 0$.

On peut facilement vérifier que le module de continuité d'une fonction Lipschitzienne f satisfait $\omega_f(a) \leq Ca$, où C est la Constante de Lipschitz de f, et que si f est Höldérienne on a par contre $\omega_f(a) \leq Ca^\alpha$.

Comme on a remarqué dans l'exemple, l'uniforme continuité est plus difficile sur $\mathbb{R}$ mais plus facile sur les ensembles bornés. En particulier, on a le théorème suivant (Théorème de Heine) :

Théorème 2.1.2. *Soient $A = [a,b]$ un intervalle fermé et borné de $\mathbb{R}$ et $f : A \to \mathbb{R}$ une fonction continue. Alors f est uniformément continue sur A.*

Démonstration. (HORS PROGRAMME) Supposons par l'absurde que f ne soit pas uniformément continue. Alors il existe $\varepsilon > 0$ tel que, pour tout $\delta > 0$, il y a deux points $x, y \in A$ avec $|f(x) - f(y)| \geq \varepsilon$ et $|x - y| < \delta$. En prenant $\delta = 1/n$ on peut appeler x_n et y_n les points correspondants. On a $|x_n - y_n| < 1/n \to 0$ et $|f(x_n) - f(y_n)| \geq \varepsilon$. Comme A est fermé et borné, il existe une sous-suite $(x_{n_k})_k$ convergente. Soit x sa limite. La correspondante sous-suite $(y_{n_k})_k$ converge à la même limite x car on a $|x_{n_k} - y_{n_k}| \to 0$. En passant à la limite dans l'inégalité $|f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon$, comme f est continue, on trouve $0 = |f(x) - f(x)| \geq \varepsilon > 0$, ce qui est une contradiction. □

Une autre propriété importante des fonctions uniformément continues est la suivante, qu'on écrit et on démontre par complétude :

Théorème 2.1.3. *Soit $f :]a,b[\to \mathbb{R}$ une fonction uniformément continue. Alors il existe une unique extension continue g de f à l'intervalle fermé $[a,b]$, c'est-à-dire une fonction continue $g : [a,b] \to \mathbb{R}$ telle que $g(x) = f(x)$ pour tout $x \in]a,b[$.*

Démonstration. (HORS PROGRAMME) Il suffit de définir g sur les points a et b et vérifier la continuité. Pour faire cela, prenons une suite $(x_n)_n \subset [a,b]$ convergeante vers a et considérons $(f(x_n))_n$. Si l'on fixe $\varepsilon > 0$, grâce à l'uniforme continuité de f, il existe $\delta > 0$ tel que $|x_n - x_m| < \delta$ entraîne $|f(x_n) - f(x_m)| < \varepsilon$. Or, la suite $(x_n)_n$ est de Cauchy, et donc l'inégalité $|x_n - x_m| < \delta$ est vraie pour $n, m \geq N_0$. Ceci implique que la suite $(f(x_n))_n$ satisfait l'inégalité $|f(x_n) - f(x_m)| < \varepsilon$ à partir du même rang, et donc $(f(x_n))_n$ est de Cauchy aussi. Elle admet donc une limite et on posera $g(a) = \lim_{n\to\infty} f(x_n)$. Celle-ci est une bonne définition, c'est-à-dire le résultat ne dépend pas de la suite que l'on a choisi. En effet, si $(y_n)_n$ était une suite différente on aurait eu en fait $|x_n - y_n| < \delta$ à partir d'un certain rang (car x_n et y_n ont la même limite a) et ceci impliquerait $|f(x_n) - f(y_n)| < \varepsilon$. Cette inégalité passant à la limite, elle implique que la différence des deux limites ne dépasse pas ε. Grâce à l'arbitrarieté de ε les deux limites doivent coïncider. En définissant comme ça $g(a)$ (et en utilisant une définition analogue pour $g(b)$ et en laissant $g(x) = f(x)$ pour tout $x \in]a,b[$) on trouve une fonction g. Elle coïncide avec f à l'intérieur de l'intervalle ouvert, donc elle est continue (comme la continuité peut être vérifiée localement). De plus, on vient d'imposer que pour toute suite $x_n \to a$ on a $\lim_n f(x_n) = \lim_n g(x_n) = g(a)$, ce qui donne la continuité en a (et en b c'est pareil). De plus, il est évident que tout extension continue de f doit coïncider avec cette fonction g qu'on vient de définir sur a et b comme il faut imposer la continuité en ces deux points-là et donc cette extension est unique. $\square$

Observation 2.1.2. La même démonstration peut marcher si f est définie sur un ensemble $A \subset \mathbb{R}$ quelconque, pas forcement un intervalle. Elle garantie l'existence d'une unique extension de f à l'ensemble des points qui peuvent être atteints comme limite d'une suite de points de A. Cet ensemble s'appelle l'adhérence de A.

Corollaire 2.1.4. *Toute fonction $f :]a,b[\to \mathbb{R}$ uniformément continue sur un intervalle borné est bornée.*

Démonstration. Il suffit de prendre l'extension g de f à $[a,b]$, qui sera une fonction continue sur un intervalle fermé et borné. Cette extension admettra donc un maximum et un minimum et sera donc bornée. La fonction f résulte bornée comme restriction à un sous-ensemble plus petit d'une fonction bornée sur un ensemble plus grand. $\square$

Observation 2.1.3. La démonstration qu'une fonction uniformément continue est bornée sur un intervalle bornée pourrait se faire sans passer à l'extension continue sur un compacte. Il suffirait de dire que, en prenant $\varepsilon = 1$, il existe un $\delta > 0$ tel que $|x - y| < \delta$ entraîne $|f(x) - f(y)| < 1$ et puis prendre une partition $t_0, t_1, \ldots t_n$ de l'intervalle telle que $|t_{i+1} - t_i| < \delta$. Si on pose $M = \max_i |f(t_i)|$ on trouve certainement $|f(x)| \leq M + 1$ pour tout x, car tout x est à distance inférieure à δ de l'un des points t_i et donc $f(x) \leq |f(t_i)| + 1$.

Après cette parenthèse sur les fonctions uniformément continues, on peut profiter de cette notion pour revenir aux intégrales.

Définition 2.1.6. Si $T = ((t_i)_i)$ est une partition $a = t_0 < t_1, \cdots < t_{n-1} < t_n = b$ est une partition d'un intervalle $[a,b]$, nous appelons taille de cette partition le nombre $\max_{i=0,\ldots,n-1}(t_{i+1} - t_i)$, c'est-à-dire la taille maximale des intervalles déterminés par cette partition, et nous écrivons ce nombre $taille(T)$.

Définition 2.1.7. Pour définir l'intégrale $\int_a^b f(t)dt$ d'une fonction continue $f : [a,b] \to \mathbb{R}$ nous procédons de la manière suivante : prenons une suite T_k de partitions, chacune avec un nombre n_k de points (donc $T_k = ((t_{i,k})_i)$ avec $a = t_{0,k} < t_{1,k}, \cdots < t_{n-1,k} < t_{n,k} = b$) ; supposons que pour tout k, la partition T_{k+1} soit plus fine que T_k (c'est-à-dire, on obtient T_{k+1} en rajoutant des points à T_k) ; supposons également que $\lim_k taille(T_k) = 0$; définissions deux suites de fonctions étagées g_k^- et g_k^+ par

$$g_k^-(x) = \min\{f(x) \,:\, x \in [t_{i,k}, t_{i+1,k}]\}\ g_k^+(x) = \max\{f(x) \,:\, x \in [t_{i,k}, t_{i+1,k}]\} \quad \text{pour } x \in]t_{i,k}, t_{i+1,k}[,$$

(comme toujours la valeur de $g_k^{\pm}$ aux points $t_{i,k}$ de leur partition adaptée n'est pas importante) ; définissons ensuite deux suites de valeurs I_k^- et I_k^+ par $I_k^{\pm} = \int_a^b g_k^{\pm}(t)dt$; ces suites étant adjacentes, elles ont une limite commune I qui sera prise comme valeur de $\int_a^b f(t)dt$.

Pour que cette défintion soit valable, il faut vérifier deux choses : que les suites $I_k^{\pm}$ soient vraiment adjacentes, et que leur limite ne dépende pas de la suite de partition effectivement choisie. C'est ce qu'on va faire dans le théorème suivant.

Théorème 2.1.5. *Soit $f : [a,b] \to \mathbb{R}$ une fonction continue et T_k une suite de partitions plus fines chacune que la précédente, avec $\lim_k taille(T_k) = 0$, comme dans la définition ci-dessus. Alors les suites I_k^- et I_k^+ sont adjacentes et ont une limite commune I. De plus, cette limite ne dépend pas de la suite T_k chosie.*

Démonstration. Nous commençons par remarquer que, grâce au fait que T_{k+1} est plus fine que T_k, alors on a $g^-_{k+1} \geq g^-_k$ ainsi que $g^+_{k+1} \leq g^+_k$. Ceci parce que, en chaque point, g^-_{k+1} est définie en prenant le minimum sur un intervalle plus petit que dans la définition de g^-_k, ce qui en fait augmenter la valeur. De même, g^+_{k+1} est un maximum sur un intervalle plus petit que dans la définition de g^+_k, ce qui fait au contraire baisser le résultat. Par conséquent, la suite I^-_k est croissante et la suite I^+_k est décroissante. Évidemment on a aussi $I^-_k \leq I^+_k$, ce qui laisse à démontrer seulement $\lim_k I^+_k - I^-_k = 0$ pour avoir des suites adjacentes.

Or, grâce au Théorème 2.1.2 la fonction f est uniformément continue aussi. Calculons la différence

$$I^+_k - I^-_k = \int_a^b (g^+_k(t) - g^-_k(t))dt = \sum_{i=1}^{n_k-1} (t_{i+1,k} - t_{i,k})(\max\{f(x) : x \in [t_{i,k}, t_{i+1,k}]\} - \min\{f(x) : x \in [t_{i,k}, t_{i+1,k}]\}).$$

Or, la quantité $\max\{f(x) : x \in [t_{i,k}, t_{i+1,k}]\} - \min\{f(x) : x \in [t_{i,k}, t_{i+1,k}]\}$ (l'oscillation de f sur $[t_{i,k}, t_{i+1,k}]$), est évidemment inférieure ou égale à $\omega(t_{i+1,k} - t_{i,k})$, où ω est le module de continuité de f. On a donc, a fortiori, $\max\{f(x) : x \in [t_{i,k}, t_{i+1,k}]\} - \min\{f(x) : x \in [t_{i,k}, t_{i+1,k}]\} \leq \omega(taille(T_k))$. On retrouve alors

$$I^+_k - I^-_k \leq \sum_{i=1}^{n_k-1} (t_{i+1,k} - t_{i,k})\omega(taille(T_k)) = (b-a)\omega(taille(T_k)).$$

Comme $\lim_k taille(T_k) = 0$ est $\lim_{a\to 0} \omega(a) = 0$, on a $\lim_k I^+_k - I^-_k = 0$ et les deux suites sont donc adjacentes.

Considérons maintenant deux suites de partitions T_k et S_k. On veut démontrer que la limite I qu'on obtient est la même. Considérons une troisième suite de partitions $R_k = T_k \cup S_k$ (pour chaque k, nous prénons tous les points de T_k ainsi que tous les points de S_k, elle est donc une partition plus fine des deux, et on a encore $\lim_k taille(R_k) = 0$). Écrivons $I(T)^\pm_k$ et $I(R)^\pm_k$ pour les suites d'intégrales faites avec les partitions T_k et R_k, respectivement. On a, pour la même raison que R_k est plus fine que T_k,

$$I(T)^-_k \leq I(R)^-_k \leq I(R)^+_k \leq I(T)^+_k.$$

Par le théorème des gendarmes, les limites de $I(T)^\pm_k$ et $I(R)^\pm_k$ doivent être les mêmes. Ceci donne l'égalité entre ce qu'on obtient avec T_k et avec R_k mais, de même, on peut l'appliquer à R_k et S_k. Finalement les limites obtenues avec T_k et S_k sont les mêmes. □

Il n'est pas difficile, ensuite, d'étendre la définition d'intégrale aux fonctions continues par morceaux. Imaginons d'avoir une fonction $f : [a,b] \to \mathbb{R}$ ainsi que des fonctions continues $f^j : I_j \to \mathbb{R}$, les intervalles I_j étant de la forme $[c_j, c_{j+1}]$ où $C = ((c_j)_j)$ est une partition de $[a,b]$ et supposons $f = f^j$ sur $]c_j, c_{j+1}[$. On peut alors définir $\int_a^b f(t)dt$ en considérant des partitions T_k qui passent chacune par tous les points c_j, en rendant de dette amnière "indépendant" ce que f fait sur chacun des intervalles I_j.

Il est utile de remarquer que toute fonction étagée est également continue par morceaux.

2.2 Le théorème fondamental du calcul et l'intégration par parties

On verra dans cette section l'instrument principal pour le calcul des intégrales, qui dans la pratique ne passe pas du tout par la définition via les fonctions étagées, ni par les limites avec les partitions. Cet outil qu'on verra s'appelle à bonne raison "théorème fondamental du calcul intégral" et rélie les concepts d'intégrale et de dérivée.

Pour en parler on a besoin d'abord de remarquer un petit détail assez facile et connu (on m'a dit comme ça en TD) sous le nom de "formule de Chasles".

Lemme 2.2.1. *Si f est une fonction continue ou continue par morceaux sur un intervalle $[a,b] = [a,c] \cup [c,b]$ (c étant un point intermédiaire entre a et b), alors on a*

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

Démonstration. Il faut d'abord remarquer que cette égalité, plutôt triviale, est encore plus triviale quand on parle de fonctions étagées, car on a à faire avec une égalité des sommes. Si f est une fonction étagée et $(t_i)_i$ est une partition qui lui est adaptée, on peut supposer que c est l'un des points séparateurs de la partition (quitte à remplacer la partition par une partition plus fine, où l'on rajoute le point c). Supposons donc $t_k = c$ avec $0 < k < n$. On a

$$\int_a^b f(t)dt = \sum_{i=0}^{n} m_i(t_{i+1} - t_i) = \sum_{i=0}^{k-1} m_i(t_{i+1} - t_i) + \sum_{i=k}^{n-1} m_i(t_{i+1} - t_i) = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

Une fois cette égalité établie pour les fonction étagées, on l'applique aux fonctions $g_k^{\pm}$ de la définition d'intégrale pour f (en faisant attention à mettre c dans toutes les partitions), et le résultat découle pour f aussi. □

On est prêt à montrer le suivant :

Théorème 2.2.2 (Théorème fondamental du calcul intégral)**.** *Soit $f : [a, b] \to \mathbb{R}$ une fonction continue. Alors la fonction $F : [a, b] \to \mathbb{R}$ définie par*

$$F(x) = \int_a^x f(t)dt$$

est dérivable et $F'(x) = f(x)$ pour tout $x \in [a, b]$ (en considérant la dérivée droite et gauche seulement pour les points a et b, respectivement).

De même, si f est seulement continue par morceaux sur $[a, b]$ mais elle est continue en un point $x_0 \in [a, b]$ alors F est dérivable au point x_0 et $F'(x_0) = f(x_0)$.

Démonstration. Si l'on remarque que toute fonction continue est continue par morceaux, il suffit de démontrer la deuxième partie de l'énoncé (le fait que la continuité en un point implique la dérivabilité de F au même point et l'égalité $F' = f$ à ce point-là). Quand f est continue en tout point, appliquer le résultat de la deuxième partie donne une démonstration de la première.

Considérons donc f continue au point x_0. Si l'on fixe $\varepsilon > 0$ et on applique la définition de continuité on trouve un $\delta > 0$ tel que $|t - x_0| < \delta$ implique $|f(t) - f(x_0)| < \varepsilon$. Prenons maintenant $x \in]x_0, x_0 + \delta[$: on a

$$F(x) = \int_a^x f(t)dt = \int_a^{x_0} f(t)dt + \int_{x_0}^x f(t)dt = F(x_0) + \int_{x_0}^x f(t)dt \leq F(x_0) + (f(x_0) + \varepsilon)(x - x_0).$$

La dernière inégalité vient du fait que, si $x \in]x_0, x_0 + \delta[$ alors tous les points $t \in]x_0, x[$ satisfont $|t - x_0| < \delta$ et donc $f(t) \leq f(x_0) + \varepsilon$. On déduit donc

$$\frac{F(x) - F(x_0)}{x - x_0} \leq f(x_0) + \varepsilon \text{ pour tout } x \in]x_0, x_0 + \delta[.$$

On peut obtenir aussi

$$\frac{F(x) - F(x_0)}{x - x_0} \geq f(x_0) - \varepsilon \text{ pour tout } x \in]x_0, x_0 + \delta[$$

si l'on considère que les points $t \in]x_0, x[$ satisfont $f(t) \geq f(x_0) - \varepsilon$ aussi. Ceci montre, par la définition de limite, que l'on a

$$\lim_{x \to x_0^+} \frac{F(x) - F(x_0)}{x - x_0} = f(x_0).$$

La limite de l'autre côté n'est pas différente si l'on considère que l'on a, pour $x < x_0$,

$$F(x) = \int_a^x f(t)dt = \int_a^{x_0} f(t)dt - \int_x^{x_0} f(t)dt = F(x_0) - \int_x^{x_0} f(t)dt$$

et on continue avec les inégalités usuelles. □

Définition 2.2.1. Toute fonction dérivable F ayant la propriété $F' = f$ est dite une primitive de la fonction f.

Ce qu'on vient de montrer a la double utilité de montrer que toute fonction continue admet une primitive (et alors elle admet une quantité infinie de primitives, $F + c$ étant elle aussi un primitive de f si F l'est et c est une constante) et que pour calculer des intégrales il suffit de calculer des primitives. En effet, pour calculer $\int_a^b f(t)dt$, il suffit de détecter une primitive F de f : cette primitive ne coïncidera pas forcement avec la fonction G donnée par $G(x) = \int_a^x f(t)dt$, mais les dérivées des deux fonctions seront les mêmes. Alors on pourra dire que leur différence sera une constante (comme la différence aura dérivée nulle). On aura alors $F = G + c$ et on pourra calculer la constante c en utilisant $G(a) = 0$. On aura donc $G(x) = F(x) - F(a)$ (la constante $-F(a)$ étant la seule compatible avec $G(a) = 0$). En prenant $x = b$ on aura donc

$$\int_a^b f(t)dt = F(b) - F(a), \text{ pour n'importe quelle primitive } F \text{ de } f,$$

ce qui représente la méthode plus courante de calcul d'une intégrale.

Pour plein de fonction en effet on est capables de déviner des primitives. Cela est plutôt facile pour toute fonction puissance : la primitive de $x \mapsto x^k$ est $x^{k+1}/(k + 1)$, pour tout $k \neq -1$ (si $k = -1$ on ne peut pas utiliser la même formule, mais on s'aperçoit assez rapidement qu'on trouve un logarithme comme primitive de $1/x$).

Pourtant, le calcul d'une primitive ne suit pas les mêmes règles schématiques du calcul d'une dérivée et toute fonction n'est pas facile à intégrer (c'est-à-dire en trouver une primitive). Non seulement, il y a aussi des fonction “élémentaires" dont la primitive existe (car elles sont des fonctions continues) mais elle n'est pas de la même forme (élémentaire, terme qui est utilisé pour indiquer toute fonction obtenue par composition, produit, somme, quotient et parfois inversion des fonctions usuels : polynômes, exponentielles, logarithme, sinus, cosinus...). C'est le cas par exemple de la célèbre Gaussienne $f(t) = e^{-t^2}$. Par contre une fonction apparemment plus compliquée comme $f(t) = te^{-t^2}$ admet une primitive calculable explicitement, donnée par $F(x) = -e^{-x^2}/2$.

C'est pour cette raison qu'on est intéressée à des méthodes de calcul des intégrales qui nous aident soit à trouver des primitives, soit à trouver directement les valeurs d'expressions de la forme $\int_a^b f(t)dt$. Ces expressions ne sont pas toujours calculées par voie de la recherche d'une primitive, car parfois des considérations de symétrie (supposons que f est impaire et $a = -b$, on obtient sûrement une intégrale nulle, quoi que ce soit l'expression de f) peuvent aider, mais il est vrai que souvent la même méthode de calcul de l'intégrale serait capable de donner le résultat général (c'est à dire pour tout b, et donc la primitive).

Une méthode qui est souvent utile est celle d'intégration par parties, ce qui découle directement du théorème fondamental du calcul

Proposition 2.2.3 (Intégration par parties)**.** *Soient f et g deux fonctions de classe C^1 sur $[a,b]$. On alors*

$$\int_a^b f(t)g'(t)dt = -\int_a^b f'(t)g(t)dt + f(b)g(b) - f(a)g(a).$$

Démonstration. Prenons la fonction $F = fg$. Sa dérivée est $F' = f'g + fg'$. On peut dire alors que F est la primitive de sa dérivée et on a donc

$$\int_a^b (f'(t)g(t) + f(t)g'(t))\, dt = F(b) - F(a) = f(b)g(b) - f(a)g(a).$$

En balançant le terme en $f'g$ de l'autre côté on obtient la thèse. □

L'utilité de l'intégration par parties est variée. Elle donne une méthode de calcul des intégrales (et des primitives) qui est applicable si l'on a un produit de deux fonctions (f et g') dont une est facile à dériver, l'autre à intégrer (celle qui doit jouer le rôle de g'), et le produit de leur dérivée et primitive respectivement est facile à intégrer (plus facile que le produit originale).

D'autre côté l'intégration par parties est une méthode essentielle dans des problèmes avancés de mathématiques en cas de manque de régularité. Si g est une fonction régulière mais f non, on peut remplacer l'expression $\int f'g$ par une expression contenant $\int fg'$ et les valeurs de f et g sur le bord. De plus, on peut vérifier (ce n'est pas trivial mais pas trop difficile non plus) qu'il vaut la proposition suivante : si u est une fonction telle que

$$\int_a^b f(t)g'(t)dt = -\int_a^b u(t)g(t)dt + f(b)g(b) - f(a)g(a)$$

pour toute fonction $g \in C^1([a,b])$, alors $u = f'$. Ceci peut être utilisé comme caractérisation de la dérivée d'une fonction (et donne lieu à une définition de dérivée pour des fonctions qui ne sont pas dérivable, mais vous verrez peut-être cela dans quelques ans)

2.3 Méthodes de calcul de primitives et intégrales

On verra dans cette section, par voie de nombreux exemples, l'application des méthodes les plus connues et utiles pour le calcul des intégrales.

2.3.1 Intégrales sur des intervalles spéciaux (symétrie, périodicité)

Exemple 2.3.1. Calculer $\int_0^{2\pi} \sin^2(x)dx$.

Si l'on trace le graphe et on regarde le sous-graphe de $x \mapsto \sin^2(x)$ on s'aperçoit d'une particularité : il est inscrit dans le rectangle $[0,2\pi] \times [0,1]$ et apparemment la partie qui lui manque pour remplir le rectangle a la même forme renversée et décalée. Ceci suggère que l'aire manquante soit égale à l'aire du sous-graphe et que donc l'aire du sous-graphe soit la moitié de celle du rectangle. Ceci donnerait $\int_0^{2\pi} \sin^2(x)dx = \pi$.

Essayons de formaliser ça avec des calculs. La partie manquante correspond à la fonction différence entre 1 et $\sin^2(x)$, c'est-à-dire le cosinus carré. On a

$$\int_0^{2\pi} \sin^2(x)dx + \int_0^{2\pi} \cos^2(x)dx = \int_0^{2\pi} 1dx = 2\pi.$$

Si l'on montre que l'intégrale du sinus au carré et du cosinus au carré donnent le même résultat (non pas leurs fonctions primitives, seulement leurs intégrales sur cet intervalle) on a terminé, car on aura $2\int_0^{2\pi} \sin^2(x)dx = \int_0^{2\pi} 1dx = 2\pi$.

On remarque $\cos(x - \pi/2) = \cos(\pi/2 - x) = \sin(x)$, et donc $\cos(x - \pi/2)^2 = \sin^2(x)$. Donc les deux fonctions sont la même, à un décalage de $\pi/2$ près. Ceci dit

$$\begin{aligned}\int_0^{2\pi} \sin^2(x)dx &= \int_0^{2\pi} \cos(x-\pi/2)^2 dx = \int_{-\pi/2}^{2\pi-\pi/2} \cos^2(x)dx \\ &= \int_{-\pi/2}^{0} \cos^2(x)dx + \int_0^{2\pi-\pi/2} \cos^2(x)dx = \int_{2\pi-\pi/2}^{2\pi} \cos^2(x)dx + \int_0^{2\pi-\pi/2} \cos^2(x)dx = \int_0^{2\pi} \cos^2(x)dx.\end{aligned}$$

On en conclut que l'intégalre du sinus au carré sur cet intervalle est la moitié de celui de la fonction constante 1. On a utilisé la périodicité pour transformer $\int_{-\pi/2}^0 \cos^2(x)dx$ en $\int_{2\pi-\pi/2}^{2\pi} \cos^2(x)dx$.

Exemple 2.3.2. Calculer $\int_{-1}^1 \sin(x)e^{x^2} \log(1 + |x|)\cos(x^3)dx$.

On remarque tout de suite qu'essayer de calculer une primitive de la fonction intégrande est un défi sans espoir de réussite : on n'est pas trop capables de travailler avec des produits ni de mélanger fonctions trigonométrique et logarithme.

Mais, il y a une particularité de la fonction dans l'intégrale qui nous pourra aider : elle est impaire ! pourquoi ? car si l'on change le signe de x on change le signe à $\sin(x)$ mais on le laisse inchangé dans tous les autres factors des produits, comme soit ils contiennent $|x|$ ou x^2, soit ils sont des fonctions paires (comme le cosinus, où l'on change le signe du cube à l'intérieur).

Et alors, c'est combien l'intégrale d'une fonction impaire sur cet intervalle ? l'intégrale est nulle comme l'intervalle est symétrique (si on avait dû l'intégrer sur $[-1, 2]$ on n'aurait pas pu le dire). Ceci est un résultat général qui peut se démontrer (sans passer par le changement des variables, ce qui simplifierait beaucoup les choses) comme ça : démontrons que sur un intervalle symétrique $\int -M^M f(t)dt = \int -M^M f(-t)dt$ (cette propriété peut être obtenue à partir des fonctions étagées) ; après on utilise le fait que f est impaire pour dire $\int -M^M f(t)dt = -\int -M^M f(-t)dt$, et les deux égalités ensemble disent que les deux intégrales sont nulles.

Qu'est-ce qu'on aurait pu dire à propose de l'intégrale dune fonction paire, par contre ? Ici ce qu'on peut montrer est $\int 0^M f(t)dt = \int -M^0 f(t)dt$ (toujours en héritant la propriété des fonction étagées). Et donc $\int -M^M f(t)dt = 2\int 0^M f(t)dt$. Mais ceci ne nous aide pas à calculer la valeur…

2.3.2 Intégration par partie : récurrence et ruses spéciales

En partant du cas du sinus au carré, on se concentrera ici sur le calcul de quelques primitives et de quelques intégrales sur des intervalles qui ne présentent pas de symétrie spéciales, où l'instrument principale est l'intégration par partie. Pourtant, il s'agira d'intégration qui ne seront pas tout de suite achevée par une simple intégration par partie, mais il faudra faire quelque chose en plus.

Exemple 2.3.3. Calculer $\int_a^b \sin^2(x)dx$ et déterminer une primitive de la fonction intégrande.

Comme on a un produit $(\sin(x) \cdot \sin(x))$ dans la fonction intégrande, on peut essayer une intégration par partie. On posera donc $f(x) = \sin(x)$ et $g'(x) = \sin(x)$, ce qui entraîne $f'(x) = \cos(x)$ et $g(x) = -\cos(x)$. On a donc

$$\begin{aligned}\int_a^b \sin^2(x)dx &= \int_a^b f(x)g'(x)dx = -\int_a^b f'(x)g(x)dx + [fg]_a^b \\ &= \int_a^b \cos^2(x)dx - \sin(b)\cos(b) + \sin(a)\cos(a).\end{aligned}$$

Il faut donc calculer l'intégrale du cosinus au carré. Apparemment on n'a pas gagné grande chose car la difficulté de cette intégrale sera la même de l'autre. On peut en fait voir que la difficulté est la même si l'on écrit $\cos^2(x) = 1 - \sin^2(x)$. On trouve donc

$$\int_a^b \sin^2(x)dx = \int_a^b 1dx - \int_a^b \sin^2(x)dx - \sin(b)\cos(b) + \sin(a)\cos(a).$$

Apparemment c'est encore pire, car pour calculer l'intégrale du sinus au carré on a justement trouvé une expression qui reprend la même valeur. Mais, les coefficients de cette intégrale d'une côté et de l'autre de l'égalité ne sont pas les mêmes, et on peut donc balancer la deuxième intégrale du côté gauche de l'égalité en obtenant

$$2\int_a^b \sin^2(x)dx = \int_a^b 1dx - \sin(b)\cos(b) + \sin(a)\cos(a) = b - a - \sin(b)\cos(b) + \sin(a)\cos(a).$$

Ceci donne le résultat final

$$\int_a^b \sin^2(x)dx = \frac{b-a}{2} - \frac{1}{2}\sin(b)\cos(b) + \frac{1}{2}\sin(a)\cos(a).$$

C'est qui la primitive associée à cette intégration ? si on considère le Théorème fondamental du calcul on sait que l'on trouve une primitive d'une fonction f si on calcule $F(x) = \int_a^x f(t)dt$. Il suffit donc prendre l'expression qu'on a calculée et remplacer b par la variable x. Non seulement, on a même le droit d'oublier tout ce qui ne concerne que a et non pas b, car il sera une constante par rapport à cette variable. On a donc

$$F(x) = \frac{1}{2}x - \frac{1}{2}\sin(x)\cos(x).$$

Comment peut-on vérifier que cela soit bien la primitive ? en la dérivant ! est-il obligatoire de le faire ? non, si on l'a trouvé grâce à une méthode existante et rigoureuse, qu'on n'a pas inventée. Cependant, on peut le faire quand on n'est pas sûr de notre résultat, et on peut le faire aussi chaque fois qu'on ne veut pas détailler les passages intermédiaires pour arriver au résultat même. Ceci pourrait être le cas soit quand ils sont trop longs, soit quand on doute de leur rigueur ou quand on est sûr de l'absence de rigueur (disons, quand on a triché). Donc on utilisera dans ce cas nos calculs comme une aide à deviner une candidate primitive, et son être primitive sera montré en dérivant.

Ici on a

$$F'(x) = \frac{1}{2} - \frac{1}{2}\cos^2(x) + \frac{1}{2}\sin^2(x) = \frac{1}{2}\sin^2(x) + \frac{1}{2}\sin^2(x) = \sin^2(x).$$

Exemple 2.3.4. Calculer $\int_a^b \sin^3(x)dx$ et déterminer une primitive de la fonction intégrande.

Ici aussi on a un produit, de la forme $\sin^2(x)\cdot\sin(x)$, par exemple. Il faut choisir qui jouera le rôle de f et qui de g'. Le rôle plus délicat est celui de g', car il faudra intégrer cette fonction. Normalement on est capable d'intégrer les deux, maintenant, mais c'est quand même vrai qu'intégrer le sinus plutôt que son carré est plus simple, alors que pour le dériver la différence n'est pas trop grande. Donc on pose $f(x) = \sin^2(x)$ et $g'(x) = \sin(x)$. On a $f'(x) = 2\sin(x)\cos(x)$ et $g(x) = -\cos(x)$. Donc

$$\int_a^b \sin^3(x)dx = \int_a^b 2\sin(x)\cos^2(x)dx - \sin^2(b)\cos(b) + \sin^2(a)\cos(a).$$

Il faut intégrer $\sin(x)\cos^2(x)$. On peut utiliser $\cos^2(x) = 1 - \sin^2(x)$ et obtenir

$$\int_a^b \sin^3(x)dx = 2\int_a^b \sin(x)dx - 2\int_a^b \sin^3(x)dx - \sin^2(b)\cos(b) + \sin^2(a)\cos(a).$$

Comme tout à l'heure, on peut ramener le terme en sinus au carré du côté droit sur le côté gauche, en trouvant

$$\begin{aligned} 3\int_a^b \sin^3(x)dx &= 2\int_a^b \sin(x)dx - \sin^2(b)\cos(b) + \sin^2(a)\cos(a) \\ &= -2\cos(b) + 2\cos(a) - \sin^2(b)\cos(b) + \sin^2(a)\cos(a) \end{aligned}$$

et finalement

$$\int_a^b \sin^3(x)dx = -\frac{2}{3}\cos(b) + \frac{2}{3}\cos(a) - \frac{1}{3}\sin^2(b)\cos(b) + \frac{1}{3}\sin^2(a)\cos(a).$$

On vérifie dans ce cas aussi, en posant $F(x) = -\frac{2}{3}\cos(x) - \frac{1}{3}\sin^2(x)\cos(x)$.

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

Le cas général peut être approché par récurrence

Exemple 2.3.5. Donner une formule récursive pour $\int_a^b \sin^n(x)dx$.

On appellera $I_n(a,b)$ cette intégrale. Il s'agit de l'intégrale d'un produit, de la forme $\sin^{n-1}(x)\cdot\sin(x)$, par exemple. Le seul choix raisonnable est $f(x) = \sin^{n-1}(x)$ et $g'(x) = \sin(x)$. On a donc $f'(x) = (n-1)\sin^{n-2}(x)\cos(x)$ et $g(x) = -\cos(x)$. Donc

$$\int_a^b \sin^n(x)dx = \int_a^b (n-1)\sin^{n-2}(x)\cos^2(x)dx - \sin^{n-1}(b)\cos(b) + \sin^{n-1}(a)\cos(a).$$

Pour le terme $\sin^{n-2}(x)\cos^2(x)$ on peut utiliser l'astuce habituelle et obtenir

$$\int_a^b \sin^n(x)dx = \int_a^b (n-1)\sin^{n-2}(x)dx - \int_a^b (n-1)\sin^n(x)dx - \sin^{n-1}(b)\cos(b) + \sin^{n-1}(a)\cos(a).$$

Ceci signifie

$$I_n(a,b) = (n-1)I_{n-1}(a,b) - (n-1)I_n(a,b) - \sin^{n-1}(b)\cos(b) + \sin^{n-1}(a)\cos(a),$$

d'où

$$I_n(a,b) = \frac{n-1}{n}I_{n-1}(a,b) - \frac{1}{n}\sin^{n-1}(b)\cos(b) + \frac{1}{n}\sin^{n-1}(a)\cos(a).$$

Le même résultat peut être utilisé pour la recherche d'une primitive, en obtenant

$$F_n(x) = \frac{n-1}{n}F_{n-1}(x) - \frac{1}{n}\sin^{n-1}(x)\cos(x).$$

2.3.3 Des cas simples

Ici on regarde deux exemples qui ont la seule propriété en commun de ne pas demander trop de travail : une intégration par partie qui n'a pas besoin de faire de la récurrence ou de ramener un terme de l'autre côté et une fonction composée dont on peut deviner la primitive (et qui nous sera utile pour introduire les changements de variables).

Exemple 2.3.6. Calculer $\int_a^b x\sin(x)dx$ et déterminer une primitive de la fonction intégrande.

On a un produit, on a envie de faire une intégration par partie. On veut transformer cette intégrale en une intégrale plus simple. Ceci est possible car la fonction x se simplifie quand on la dérive alors que la fonction $\sin(x)$ ne se complique pas. Posons $f(x) = x$ et $g'(x) = \sin(x)$. On a $f'(x) = 1$ et $g(x) = -\cos(x)$. Donc

$$\int_a^b x\sin(x)dx = \int_a^b \cos(x)dx - b\cos(b) + a\cos(a) = \sin(b) - \sin(a) - b\cos(b) + a\cos(a).$$

On vérifie dans ce cas aussi, en posant $F(x) = \sin(x) - x\cos(x)$.

$$F'(x) = \cos(x) - \cos(x) + x\sin(x) = x\sin(x).$$

Exemple 2.3.7. Calculer $\int_a^b \frac{\ln^k(x)}{x}dx$ et déterminer une primitive de la fonction intégrande (avec $b > a > 0$).

Ici le gros avantage est qu'on a le produit d'une fonction du logarithme fois la fonction $1/x$, qui en est la dérivée. Devinons la primitive : si l'on prend une puissance du logarithme en dérivant on trouve bien une puissance du logarithme divisée par x. Il est facile de voir que l'exposant à prendre sera un en plus par rapport à l'exposant dans l'intégrale. Le coefficient aussi peut être deviné. On pose $F_k(x) = \ln^{k+1}(x)/(k+1)$. On a bien $F_k'(x) = \ln^k(x)/x$. Bien sûr, ce calcul ne marche pas si $k = -1$. Ceci est normal, on ne peut pas calculer la primitive d'une puissance avec exposant -1 en rajoutant 1 à l'exposant, on le sait. On sait aussi que dans ce cas on trouverait un logarithme. On essaie donc $F_{-1}(x) = \ln(\ln(x))$, ce qui donne bien $F_{-1}'(x) = 1/x\ln(x)$.

Après avoir trouvé les primitives, on peut facilement dire

$$\int_a^b \frac{\ln^k(x)}{x}dx = \begin{cases} \frac{\ln^{k+1}(b)-\ln^{k+1}(a)}{k+1}, & \text{si } k \neq -1, \\ \ln(\ln(b)) - \ln(\ln(a)) & \text{si } k = -1. \end{cases}$$

2.3.4 Changement de variable d'intégration

L'exemple 2.3.7 nous donne une idée intéressante : si dans une expression à intégrer il y a le logarithme qui apparait souvent, mais il y a aussi la dérivée du logarithme, alors on peut oublier la partie avec la dérivée, regarde la fonction qu'on a composée avec le logarithme, en prendre la primitive, et y mettre au dedans le logarithme dans le résultat. Ceci se peut généraliser.

En utilisant la convention $\int_a^b f(t)dt := -\int_b^a f(t)dt$ si $a > b$, on a :

Théorème 2.3.1. *Soit $f : [a, b] \to \mathbb{R}$ une fonction continue et $g : [c, d] \to [a, b]$ une bijection C^1. Alors on a*

$$\int_a^b f(t)dt = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(s))g'(s)ds = \int_c^d f(g(s))|g'(s)|ds.$$

Démonstration. Soit F une primitive de f, qui existe grâce au théorème fondamental du calcul (f étant continue). Considérons $F \circ g$: est-elle primitive de quelque chose ? oui, car elle est C^1 (comme composée de fonctions C^1) et elle est donc une primitive de sa dérivée. Sa dérivée est bien $s \mapsto f(g(s))g'(s)$. On a donc, pour tout x et y,

$$\int_x^y f(g(s))g'(s)ds = F(g(y)) - F(g(x)).$$

Si l'on prend $x = g^{-1}(a)$ et $y = g^{-1}(b)$ on trouve

$$\int_{g^{-1}(a)}^{g^{-1}(b)} f(g(s))g'(s)ds = F(b) - F(a) = \int_a^b f(t)dt,$$

où la dernière égalité vient du Théorème Fondamental du Calcul et de ses conséquences. Ceci montre la première partie de la thèse. Pour conclure et voir

$$\int_{g^{-1}(a)}^{g^{-1}(b)} f(g(s))g'(s)ds = \int_c^d f(g(s))|g'(s)|ds$$

il suffit de distinguer deux cas : comme g est une bijection continue (elle est C^1), elle est strictement monotone ; soit elle est croissante (et dans ce cas on a $g(c) = a$ et $g(d) = b$, car sinon on ne peut pas réaliser une bijection, et $g' \geq 0$), soit elle est décroissante (avec $g(c) = b$ et $g(d) = a$ et $g' \leq 0$). Dans le premier cas l'égalité est immédiate, dans le deuxième il y a un double changement de signe, car l'intervalle résulte renversé mais $|g'| = -g'$. □

Ce théorème peut être utilisé dans deux directions : soit on doit intégrer une fonction composée $f(g(t))$ multipliée fois g' (c'est le cas de notre dernier exemple), et alors on passera à l'intégrale de f tout court, soit on a une fonction $f(t)$ et pour quelques raisons on soupçonne que remplacer t avec $g(s)$ pourrait être malin (il faut que dans le produit $f(g(s))g'(s)$ il y ait quelque simplification, sinon cela ne vaut pas le coup).

Ce dernier cas sera développé dans l'exemple suivant.

Exemple 2.3.8. Calculer $\int_a^b \sqrt{1 - x^2}dx$ et déterminer une primitive de la fonction intégrande (avec $-1 \leq a < b \leq 1$). On fera le changement de variable suivant : $x = \sin(y)$. La fonction $y \sin(y)$ est une bijection croissante de $[-\pi/2, \pi/2]$ vers $[-1, 1]$ et sa restriction à l'intervalle $[\arcsin(a), \arcsin(b)]$ le sera vers $[a, b]$. On appliquera le théorème 2.3.1 avec $f(t) = \sqrt{1 - t^2}$ et $g(s) = \sin(s)$. Il faut remplacer x avec $\sin(y)$, dx avec $\cos(y)dy$ et les bornes d'intégration. On a

$$\int_a^b \sqrt{1 - x^2}dx = \int_{\arcsin(a)}^{\arcsin(b)} \sqrt{1 - \sin^2(y)}\cos(y)dy = \int_{\arcsin(a)}^{\arcsin(b)} \cos^2(y)dy$$

(on a utilisé $\cos^2(y) = 1 - \sin^2(y)$ mais $\cos(y) \geq 0$ sur $[-\pi/2, \pi/2]$ aussi). Il faut maintenant calculer la primitive de $\cos^2(y)$ mais cela on sait le faire. De la relation $\cos^2(y) = 1 - \sin^2(y)$ on tire que la primitive F du cosinus au carré est donnée par

$$F(x) = x - \frac{1}{2}x + \frac{1}{2}\sin(x)\cos(x) = \frac{1}{2}x + \frac{1}{2}\sin(x)\cos(x).$$

On a donc

$$\int_a^b \sqrt{1 - x^2}dx = \frac{1}{2}(\arcsin(b) - \arcsin(a)) + \frac{1}{2}\sin(\arcsin(b))\cos(\arcsin(b)) - \frac{1}{2}\sin(\arcsin(a))\cos(\arcsin(a)).$$

On remarque $\sin(\arcsin(x)) = x$ et $\cos(\arcsin(x)) = \sqrt{1 - \sin^2(\arcsin(x))} = \sqrt{1 - x^2}$ et on continue :

$$\int_a^b \sqrt{1 - x^2}dx = \frac{1}{2}(\arcsin(b) - \arcsin(a)) + \frac{1}{2}b\sqrt{1 - b^2} - \frac{1}{2}a\sqrt{1 - a^2}.$$

On peut déduire que une primitive de $f(x) = \sqrt{1 - x^2}$ est donnée par $F(x) = \frac{1}{2}\arcsin(x) + \frac{1}{2}x\sqrt{1 - x^2}$. Ce résultat peut être vérifié en utilisant $(\arcsin)'(x) = (1 - x^2)^{-1/2}$, ce qui est une conséquence de la formule pour le calcul de la dérivée d'une fonction réciproque.

Dans l'exemple précédent on a directement cherché à calculer le résultat d'une intégrale sur un intervalle $[a, b]$ et on a bien fait attention à changer les bornes lors du changement de la variable. Il faut faire attention quand on calcule une primitive aussi. Une écriture du type

$$\int \sqrt{1-x^2}dx = \int \cos^2(y)dy$$

pourrait engendrer de la confusion et suggérer que la primitive de la fonction $\sqrt{1-x^2}$ coïncide avec la primitive de la fonction $\cos^2$. Ceci n'est pas vrai, car il faut composer cette dernière primitive avec la fonction réciproque du changement de variable, l'arcsinus dans ce cas.

On termine cette section par quelques remarques sur des fonctions moins connues que le sinus et le cosinus mais qui peuvent être aussi utiles dans ce genre d'intégrales.

On définit les fonctions sinus hyperbolique et cosinus hyperbolique par

$$\sinh(x) = \frac{e^x - e^{-x}}{2}; \quad \cosh(x) = \frac{e^x + e^{-x}}{2}.$$

Ces fonctions, qui sont définis tout à fait différemment (sauf si l'on considère l'expression complexe du sinus et du cosinus, $\sin(x) = (e^{ix} - e^{-ix})/2$ et $\cos(x) = (e^{ix} + e^{-ix})/2$), ont beaucoup de propriétés similaires aux fonctions trigonométriques, qui peuvent être vérifiées à la main.

On a notamment

$$\sinh' = \cosh; \quad \cosh' = \sinh;$$
$$\cosh^2(x) - \sinh^2(x) = 1;$$
$$\cosh(0) = 1, \quad \sinh(0) = 0;$$

ainsi que des formules d'addition similaires à celles trigonométriques. Dans tous les cas, il y a des différences de signes.

Un changement de variable $x = \sinh(y)$ est très utile pour le calcul d'une primitive de $\sqrt{1+x^2}$ ainsi comme $x = \cosh(y)$ pour $\sqrt{x^2-1}$ (remarquer la différence de signe par rapport à l'exemple précédent). Évidemment, dans le calcul explicite des primitives on trouvera les fonctions réciproques du sinus et du cosinus hyperboliques, qui peuvent être obtenues en résolvant les équations

$$\frac{e^x - e^{-x}}{2} = y \quad \text{et} \quad \frac{e^x + e^{-x}}{2} = y,$$

en mettant $X = e^x$, résolvant une équation de seconde dégré, puis prenant le logarithme.

2.3.5 Fonctions rationnelles

Exemple 2.3.9. Calculer $\int_a^b \frac{x+1}{x^2-3x+2}dx$ et déterminer une primitive de la fonction intégrande (il faut se placer sur un intervalle qui ne touche pas les zéros de dénominateur : par souci de simplicité supposons ici $2 < a < b$).

Pour trouver cette primitive on réécrira la fonction intégrande de manière plus simple. Il est en effet possible trouver des constantes A et B telles que

$$\frac{x+1}{x^2-3x+2} = \frac{A}{x-1} + \frac{B}{x-2}.$$

On verra après quelles sont les critères pour exprimer une fraction du type $P(x)/Q(x)$ comme somme de fractions plus simples, et ici on se contentera de trouver explicitement ces deux constantes en résolvant un système. Pour que l'égalité qu'on cherche soit vérifiée il faut imposer

$$\frac{x+1}{x^2-3x+2} = \frac{A(x-2) + B(x-1)}{x^2-3x+2},$$

ce qui équivaut, en imposant l'égalités des coefficients de x et des constantes, à

$$\begin{cases} 1 = A + B, \\ 1 = -2A - B. \end{cases}$$

La solution de ce système est $A = -2$, $B = 3$. On a donc

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

et donc la primitive cherchée est donnée par

$$F(x) = -2\ln(x-1) + 3\ln(x-2).$$

Exemple 2.3.10. Calculer $\int_a^b \frac{x^4-2}{x^3-1}dx$ et déterminer une primitive de la fonction intégrande (là aussi, par souci de simplicité, on supposera $1 < a < b$).

On cherchera de faire une décomposition de la fraction intégrande qui s'inspire à celle de l'exemple précedent. Pour commencer, cependant, on s'aperçoit que, comme le numérateur a un degré plus élevé que le dénominateur, il sera possible de sortir des termes de la fraction. On a en effet $x^4 - 2 = x(x^3 - 1) + x - 2$ et donc

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

On est bien capable de calculer une primitive de x et ce qui nous reste à faire est calculer une primitive de la fraction qu'on a obtenu. L'avantage est qu'on a réduit le degré du numérateur. Ceci peut être toujours fait si l'on a $P(x)/Q(x)$ et $\deg P \geq \deg Q$: grâce à $P = P_1Q + P_2$ (division euclidienne des polynômes avec un reste) on arrive à des fractions avec $\deg P_2 < \deg Q$.

Tout à l'heure on avait décomposé la fraction comme une somme de termes plus simples dont les dénominateurs étaient les facteurs du dénominateur initial. Ici, si l'on veut faire une décomposition en facteurs de degré 1, il faudrait passer aux complexes et utiliser $x^3 - 1 = (x-1)(x+1/2-i\sqrt{3}/2)(x+1/2+i\sqrt{3}/2)$. Pourtant, si le but est prendre des primitives plus faciles et passer notamment aux logarithmes, ceci n'est pas très utile car $\ln(x + 1/2 - i\sqrt{3}/2)$ ne signifie rien ! On est donc forcé à garder la décomposition réelle $(x^3 - 1) = (x - 1)(x^2 + x + 1)$. Il ne sera pas possible alors d'espérer de trouver une décomposition du type

$$\frac{x - 2}{x^3 - 1} = \frac{A}{x - 1} + Bx^2 + x + 1$$

car le résultat à droite ne dépende que de deux paramètres, alors que à droite on aurait pu avoir a priori n'importe quelle fraction P_2/Q avec $\deg P_2 < 3$ (un résultat qui dépend donc de trois paramètres, les trois coefficients de P_2).

La décomposition qui sera vrai sera par contre la suivante :

$$\frac{x - 2}{x^3 - 1} = \frac{A}{x - 1} + \frac{Bx + C}{x^2 + x + 1}.$$

Pour calculer les coefficients A, B et C (et au même temps démontrer leur existence, car on n'a pas évoqué des théorèmes généraux qui l'assurent) il faut imposer

$$x - 2 = A(x^2 + x + 1) + (Bx + C)(x - 1),$$

ce qui signifie

$$\begin{cases} 0 = A + B, \\ 1 = A - B + C \\ -2 = A - C. \end{cases}$$

Ici aussi la solution est simple (il suffit de trouver A en sommant les trois lignes) : $A = -1/3$, $B = 1/3$ et $C = 5/3$. On ne détaillera pas tous les calculs ici mais on veut seulement montrer comment on peut intégrer les fonctions qui apparaissent. Il n'y a pas de problèmes à intégrer $A/(x - 1)$, qui donnera lieu à un logarithme. Concernant la partie $(Bx + C)/(x^2 + x + 1)$, on commence en écrivant le numérateur comme $\lambda(2x + 1) + \mu$ (ce qui est possible en prenant $\lambda = B/2$ et $\mu = C - B/2$). Pourquoi ? parce que la première partie est facile à intégrer. On a en effet la fraction $(2x+1)/(x^2+x+1)$, qui a la propriété que le numérateur est la dérivée du dénominateur. Chaque fois qu'on a u'/u la primitive est $\ln(u)$ et ceci est facile à vérifier en dérivant. Ici on aurait donc $\lambda \ln(x^2+x+1)$. Il nous reste à trouver une primitive de $1/(x^2 + x + 1)$.Plus en général on vourrait trouver une primitive de toute fraction du type

$$\frac{1}{\text{polyôme de degré 2 sans racines réelles}}.$$

Le cas typique est celui de la primitive$1/(1 + x^2)$, qui est donnée par la fonction $\arctan(x)$. Tout autre cas peut se ramener à celui-ci, grâce à une écriture du type

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

et plus en général

$$Q(x) = c_1(x - x_0)^2 + c_2.$$

Il est toujours possible écrire un polynôme de degré deux sous cette forme (on trouve c_1 d'après le coefficient de x^2, puis x_0 pour faire en sorte que le double produit du carré égalise le terme en x, puis c_2 par différence). Dans le cas d'un polynôme sans racines réelles on aura $c_1c_2 > 0$. Ceci dit que, à un coefficient près, on peut se ramener à x^2+1. Ici on a

$$\frac{1}{x^2+x+1} = \frac{4}{3}\frac{1}{1+\left(\frac{2}{\sqrt{3}}(x+\frac{1}{2})\right)^2}$$

et une primitive sera donc donnée par

$$F(x) = \frac{4}{3}\frac{\sqrt{3}}{2}\arctan\left(\frac{2}{\sqrt{3}}(x+\frac{1}{2})\right).$$

En composant tous ces résultats on peut trouver une primitive de la fonction donnée et en calculer l'intégrale. Dans le résultat il y aura un polynôme (qui vient de la partie qui est sortie de la fraction), un logarithme de $x-1$, un logarithme de x^2+x+1, une arcotangente.

On verra un dernier exemple et on donnera en suite une recette pour intégrer toute fonction rationnelle.

Exemple 2.3.11. Calculer $\int_a^b \frac{1}{(x^3-1)^2}dx$ et déterminer une primitive de la fonction intégrande. Il est encore sous-entendu que l'intervalle ne touche pas les zéros de la fonction. De plus, on met 1 au numérateur par simplicité car on a vu que la technique à utiliser dépend fortement du dénominateur et de sa factorisation, et le numérateur n'influence que les constantes qui apparaissent dans les fractions plus simples qu'on va trouver.

Ici le problème est que la factorisation du dénominateur ademt des facteurs multiples. En effet on a

$$(x^3-1)^2 = (x-1)^2(x^2+x+1)^2 = (x-1)(x-1)(x^2+x+1)(x^2+x+1).$$

Si on cherchait d'écrire

$$\frac{1}{(x^3-1)^2} = \frac{A}{x-1} + \frac{B}{x-1} + \frac{Cx+D}{x^2+x+1} + \frac{Ex+F}{x^2+x+1},$$

on s'apercevrait que le résultat à droite ne dépend que de $A+B$, $C+E$ et $D+F$, donc d'un nombre trop petit de variables. O n ne pourra pas s'en sortir comme ça, par conséquent.

Dans le cas de facteurs multiples la décomposition est un peu plus subtile. On peut écrire

$$\frac{1}{(x^3-1)^2} = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{Cx+D}{x^2+x+1} + \frac{Ex+F}{(x^2+x+1)^2},$$

c'est-à-dire on obtient une décomposition avec des fractions dont les dénominateur ont les facteurs du dénominateur original $Q(x)$, à une puissance qui va de un jusqu'à la puissance qu'on voit dans Q, et dont le numérateur est une constante s'il s'agit d'un facteur de premier degré et un polynôme de degré un si le facteur est de degré deux.

En résolvant le système on peut voir qu'il admet une solution unique et la trouver. On s'intéresse maintenant à montrer que toute fraction qu'on a écrite est facile à intégrer. C'est évidemment le cas de $1/(x-1)$, qui donne lieu à $\ln(x-1)$, et de $1/(x-1)^2$, qui donne lieu à $-1/(x-1)$. On a déjà vu comment traiter $(Ax+B)/(x^2+x+1)$, en le décomposant en une partie avec $(2x+1)/(x^2+x+1)$, qui donne $\ln(x^2+x+1)$, et une avec $1/(x^2+x+1)$, qui donne un résultat du type $c_1\arctan(c_2(x-x_0))$. Il nous reste la dernière fraction. Là aussi on peut dire $Ex+F = \lambda(2x+1)+\mu$ (toujours grâce à la division de polynômes). La partie avec $(2x+1)/(x^2+x+1)^2$ donne tout simplement $-1/(x^2+x+1)$. Pour l'autre il faut faire le même changement de variable qui passe par l'écriture de x^2+x+1 comme $c_1+c_2(x-x_0)^2$ et qui nous ramène à $1/(1+x^2)^2$ et se référer à l'exemple suivant.

Exemple 2.3.12. Donner une formule récursive pour $\int_a^b \frac{1}{(1+x^2)^n}dx$.

On appellera $I_n(a,b)$ cette intégrale, qu'on sait calculer pour $n=0$ et 1. On cherchera une relation entre I_n et I_{n+1}. On a

$$\int_a^b \frac{1}{(1+x^2)^n}dx = \int_a^b \frac{1+x^2}{(1+x^2)^{n+1}}dx = I_{n+1}(a,b) + \int_a^b x\frac{x}{(1+x^2)^{n+1}}dx.$$

On intégrera par partie la dernière expression, en prenant $f(x)=x$ et $g'(x) = x/(1+x^2)^{n+1}$, ce qui donne $f'(x)=1$ et $g(x) = -(1+x^2)^{-n}/(2n)$. On a donc

$$\int_a^b x\frac{x}{(1+x^2)^{n+1}}dx = \frac{1}{2n}\int_a^b \frac{1}{(1+x^2)^n}dx - \frac{b}{2n(1+b^2)^n} + \frac{a}{2n(1+a^2)^n} = \frac{I_n(a,b)}{2n} - \frac{b}{2n(1+b^2)^n} + \frac{a}{2n(1+a^2)^n}.$$

Ceci nous dit

$$= I_{n+1}(a,b) = I_n(a,b)(1-\frac{1}{2n}) + \frac{b}{2n(1+b^2)^n} - \frac{a}{2n(1+a^2)^n}.$$

On peut résumer la recette qu'on a construit avec les derniers exemples.

Recette pour l'intégration d'une fonction rationnelle $P(x)/Q(x)$ **:**

– diviser P par Q en obtenant $P = P_1 Q + P_2$ et donc $P(x)/Q(x) = P_1(x) + P_2(x)/Q(x)$;
– P_1 est facile à intégrer ;
– factoriser Q sous la forme $Q = Q_1^{\alpha_1} \cdot \dots \cdot Q_m^{\alpha_m} Q_{m+1}^{\alpha_{m+1}} \cdot \dots \cdot Q_k^{\alpha_k}$, où les Q_i avec $i \leq m$ sont des polynômes de degré un et les Q_i avec $i \geq m+1$ sont des polynômes de degré deux sans racines réelles ; les correspondants α_i leurs exposants dans la factorisation ;
– écrire

$$\begin{aligned} \frac{P_2(x)}{Q(x)} \quad = \quad & \frac{A_{1,1}}{Q_1(x)} + \dots + \frac{A_{1,\alpha_1}}{Q_1(x)^{\alpha_1}} + \dots + \frac{A_{m,1}}{Q_m(x)} + \dots + \frac{A_{m,\alpha_m}}{Q_m(x)^{\alpha_m}} \\ & + \frac{A_{m+1,1} + B_{m+1,1}x}{Q_{m+1}(x)} + \dots + \frac{A_{m+1,\alpha_{m+1}} + B_{m+1,\alpha_{m+1}}x}{Q_{m+1}(x)^{\alpha_{m+1}}} + \dots \\ & + \frac{A_{k,1} + B_{k,1}x}{Q_k(x)} + \dots + \frac{A_{k,\alpha_k} + B_{k,\alpha_k}x}{Q_k(x)^{\alpha_k}}, \end{aligned}$$

  c'est-à-dire en mettant au dénominateur tous les facteurs Q_i et en les prenant à une puissance entre un et leur α_i, et au numérateur soit des constantes (si $\deg Q_i = 1$), soit des polynômes de degré un (si $\deg Q_i = 2$) ;
– tous les termes du type A/Q_i^β avec $\deg Q_i = 1$ sont faciles à intégrer et donnent soit un logarithme ($\beta = 1$), soit une puissance négative de Q_i ;
– tous les termes du type $(A + Bx)/Q_i^\beta$ avec $\deg Q_i = 2$ se décomposent en $\lambda Q_i'/Q_i^\beta + \mu/Q_i^\beta$.
– tous les termes du type $\lambda Q_i'/Q_i^\beta$ sont faciles à intégrer et donnent soit $\ln(Q_i)$ (si $\beta = 1$), soit une puissance négative de Q_i ;
– pour les termes μ/Q_i^β on utilise $Q_i(x) = c_1 + c_2(x - x_0)^2$ pour se ramener à $1/(1+x^2)^\beta$ à des coefficients près, et puis on utilise le résultat de récurrence de l'exemple 2.3.12

2.3.6 Fonctions rationnelles de fonctions trigonométriques

On montre maintenant une technique pour transformer pas mal d'intégrales avec les sinus et les cosinus en d'intégrales de fonctions rationnelles.

Exemple 2.3.13. Calculer $\int_a^b \frac{\sin^3(x) - \cos^4(x)}{\sin(x)\cos(x) + 2\cos^3(x)} dx$ et déterminer une primitive de la fonction intégrande.

L'idée pour calculer cette intégrale (qu'on ne développera pas en détails) et la suivante : le sinus et le cosinus peuvent s'exprimer assez bien à l'aide d'une quantité magique commune, qui est $t = \tan(x/2)$, et cela donne lieu à un changement de variable assez puissant.

On a

$$\sin(x) = 2\sin(x/2)\cos(x/2) = 2\tan(x/2)\cos^2(x/2) = \frac{2t}{1+t^2},$$

grâce à l'égalité $\cos^2 = 1/(1 + \tan^2)$. De même, on a

$$\sin(x) = \cos^2(x/2) - \sin^2(x/2) = \cos^2(x/2)\left(1 - \tan^2(x/2)\right) = \frac{1-t^2}{1+t^2},$$

Faisons donc le changement de variable $t = \tan(x/2)$, qui implique $x = 2\arctan(t)$. Il faut remplacer tous les sinus et les cosinus par les expressions qu'on a trouvées, qui sont des fonctions rationnelles de t. De plus, il faut remplacer dx par $2dt/(1+t^2)$, grâce à la formule de changement de variable qui fait apparaître une dérivée. Celle-ci aussi est une fonction rationnelle de t. Dans notre cas on se ramerait à une intégrale du type

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

Comme $1 + t^2$ apparaît au dénominateur un peu partout, il est possible de le simplifier dans pas mal d'endroits. Attention : on pourrait se demander comment peut-on obtenir des sinus et des cosinus quand on calcule cette primitive et on la rédérive, si dans les formules pour les fonction rationnelles ils n'apparaissent jamais. . .la réponse est évidemment qu'il faut faire gaffe aux bornes, car il faut remplacer a par $\tan(a/2)$ et b par $\tan(b/2)$.

2.4 Applications des intégrales aux développements limités

On termine le long discours sur les intégrales avec deux applications aux DL. La première sort de la question suivante : on connaît les DL de la fonction $1/(1+x^2)$

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots + (-1)^n x^{2n} + o(x^{2n+1}).$$

Peut-on déduire d'ici les DL de sa primitive, la fonction arcotangente ? On imagine trouver

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + o(x^{2n+2}).$$

La réponse est oui et elle est assez générale.

Théorème 2.4.1. *Soit $f : A \to \mathbb{R}$ une fonction qui admet un développement limité d'ordre n au point $x_0 \in A$: $f(x) = p(x) + o((x-x_0)^n)$ avec p un polynôme de degré n. Soient F et P des primitives de f et p, respectivement, avec $P(x_0) = 0$ (c'est-à-dire on a forcement $P(x) = \int_{x_0}^x p(t)dt$). Alors on a le DL suivant*

$$F(x) = F(x_0) + P(x) + o((x-x_0)^{n+1}).$$

Démonstration. On veut montrer $\lim_{x\to x_0} \frac{F(x)-F(x_0)-P(x)}{(x-x_0)^{n+1}} = 0$. Soit $H(x) = F(x) - F(x_0) - P(x)$. On sait $H(x_0) = 0$ et $H'(x) = f(x) - p(x)$. Comme on a $f(x) - p(x) = o((x-x_0)^n)$, on peut dire que pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que, si $|y - x_0| < \delta$, alors $|H'(y)| < \varepsilon |y - x_0|^n$. Prenons maintenant $x \in]x_0 - \delta, x_0 + \delta[$. Grâce au théorème des accroissements finis on a $|H(x)| = |H(x) - H(x_0)| = |x - x_0||H'(\xi)|$ avec $\xi \in]x_0, x[\subset]x_0 - \delta, x_0 + \delta[$. Donc $|H'(\xi)| \le \varepsilon |\xi - x_0|^n \le \varepsilon |x - x_0|^n$ et finalement $|H(x)| \le \varepsilon |x - x_0|^{n+1}$. Ceci montre exactement que la quantité $H(x)/(x-x_0)^{n+1}$ tend vers zéro lorsque $x \to x_0$ et conclut la preuve. □

Dernière application des intégrales, la fameuse Formule de Taylor avec reste intégrale, qui est celle, parmi les Formules de Taylor, qui donne l'expression du reste la plus précise. Dans la formule de Taylor-Young on avait en fait un terme en $o((x-x_0)^n)$, qui n'est guère précis mais donne seulement des informations sur son comportement local autour de x_0 ; la formule de Taylor-Lagrange donnait une vraie égalité mais avec un terme $f^{n+1}(\xi)(x - x_0)^{n+1}/(n+1)!$, où le point ξ n'est pas precisé ; cette formule donnera par contre une égalité explicite. Par contre, elle demandera plus de régularité pour pouvoir l'appliquer.

Théorème 2.4.2. *Soit $f \in C^{n+1}(A)$ et $x_0 \in A$. Alors pour tout $x \in A$ on a*

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f''(x_0)(x-x_0)^2 + \cdots + \frac{1}{n!} f^{(n)}(x_0)(x-x_0)^n + \frac{1}{n!}\int_{x_0}^x (x-t)^n f^{(n+1)}(t)dt.$$

Démonstration. Le théorème sera démontré par récurrence sur n. Si $n = 0$ il faut démontrer que pour toute fonction C^1 on a $f(x) = f(x_0) + \int_{x_0}^x f'(t)dt$ et ceci est vrai comme conséquence du théorème fondamental du calcul. Le fait que f' soit continue est justment la bonne hypothèse pour l'appliquer.

Supposons que le résultat soit vrai pour un certain rang n. Passons au rang suivant : il faut prendre $f \in C^{n+2}(A)$. Mais alors $f \in C^{n+1}(A)$ et on peut appliquer le résultat au rang n. On obtient bien

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f''(x_0)(x-x_0)^2 + \cdots + \frac{1}{n!} f^{(n)}(x_0)(x-x_0)^n + \frac{1}{n!}\int_{x_0}^x (x-t)^n f^{(n+1)}(t)dt.$$

Dans la dernière intégrale il y a un produit qu'on peut traiter grâce à une intégration par partie. Soit $h(t) = f^{(n+1)}(t)$ et $g'(t) = (x-t)^n$. Grâce à l'hypothèse $f \in C^{n+2}(A)$ la fonction h est C^1 et on a $h'(t) = f^{(n+2)}(t)$, ainsi que $g(t) = -(x-t)^{n+1}/(n+1)$.

$$\begin{aligned}
\int_{x_0}^x (x-t)^n f^{(n+1)}(t)dt &= \int_{x_0}^x \frac{(x-t)^{n+1}}{n+1} f^{(n+2)}(t)dt - \left[f^{(n+1)}(t)\frac{(x-t)^{n+1}}{(n+1)} \right]_{x_0}^x \\
&= \int_{x_0}^x \frac{(x-t)^{n+1}}{n+1} f^{(n+2)}(t)dt + f^{(n+1)}(x_0)\frac{(x-x_0)^{n+1}}{(n+1)}.
\end{aligned}$$

Par conséquent on peut transformer l'égalité d'en haut en

$$\begin{aligned}
f(x) &= f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f''(x_0)(x-x_0)^2 + \cdots + \frac{1}{n!} f^{(n)}(x_0)(x-x_0)^n \\
&\quad + \frac{1}{(n+1)!} f^{(n+1)}(x_0)(x-x_0)^{n+1} + \frac{1}{(n+1)!}\int_{x_0}^x (x-t)^{n+1} f^{(n+2)}(t)dt. \quad \square
\end{aligned}$$