Polytech'Paris-Sud
PeiP1

2011/2012

**Notes de cours**

# *Mathématiques, Semestre S1*

Filippo SANTAMBROGIO

# Table des matières

# Avertissements

Ces notes sont de support pour le cours de mathématiques du premier semestre de la première année des élèves ingénieurs de Polytech'Paris-Sud.

Dans ce semestre, les cours de maths s'articulent en deux parties. Une première partie, bizarrement appelée "MathsInfo" dans l'emploi du temps, porte sur la logique, les démonstrations, les ensembles et les fonctions dans leur généralité, comme introduction à la rigueur mathématique abstraite. Pour cette partie, un poly préparé exprès par S. Lelièvre et P. Pansu est disponible. La deuxième partie ("Maths" dans l'emploi du temps) porte sur les bases de l'analyse mathématiques des fonctions d'une variable (limites, continuité, dérivées, intégrales...) avec quelques ouvertures aux deux variables vers la fin. Pour cette 2e partie, on utilisait traditionnellement le poly préparé par J.-C. Léger et F. Menous pour la fac de sciences. Or, ce poly ne s'adaptait pas aux finalités appliquées d'un cours pour ingénieurs, et s'appuyait aussi sur certaines notions qui nous sont, malheureusement, interdites dans ce cours (c'est notamment le cas des suites, que vous allez voir au 2e semestre, et du langage du inf et du sup, bornes inférieure et supérieure, pour ceux qui connaissent...). Il était donc opportun de refaire un nouveau poly spécifique à ce cours, et ce poly veut répondre à cette exigence. Tant qu'à faire, j'en ai également profité pour modifier certaines définitions ou approches que je préférais changer (il y a, de temps en temps, en maths, des approches différentes, qu'est-ce qu'on donne comme définition et qu'est-ce que l'on démontre ensuite...quelle est la définition meilleure...et on n'est pas tout le temps d'accord !).

Or, quels sont les avertissements ? surtout que ce poly est en construction et que pendant l'année je pourrais trouver des détails à changer sur les parties que vous avez déjà reçues. Ces modifications seront mises en ligne et seront, évidemment, utiles aux étudiants des promotions futures. En particulier il pourrait toujours y avoir des fautes de français, et je m'en excuse. Et, plus que des fautes, des expressions qui me paraissent tout à fait claires mais que tout français trouverait incompréhensible. N'hésitez pas à me les signaler, par e-mail ou en fin de cours. Sinon, le poly se base surtout sur l'ancien poly de Léger et Menous et sur des polys que j'avais rédigés pour des cours que j'avais donnés ailleurs, des cours qui partageaient avec celui-ci un esprit plus appliqué. Il peut toujours y avoir des erreurs dus au copie-coller mais surtout des problèmes d'harmonisation entre les différentes sources (notations ou noms des variables différents...) qui devraient s'améliorer au fur et à mesure des corrections.

Un dernier mot sur les preuves : bien que le but soit d'apprendre des notions pour les appliquer, la rigueur des mathématiques reste néanmoins très importante et il est également important de savoir pourquoi certaines choses sont vraies. Voilà pourquoi il est important de voir les preuves. Mais tous les théorèmes ne trouvent pas de preuve dans ce poly et dans ce cours : si vous trouvez un théorème (ou proposition, lemme, corollaire...) sans preuve, c'est que sa preuve nous demanderait un langage ou des outils qu'on n'a pas, donc on va la zapper tout en en gardant l'énoncé, qu'on pourra utiliser. Il y a également des théorèmes dont la preuve est écrite dans le poly mais qu'on ne détaillera pas en classe : si c'est sûr dès le début du cours qu'une certaine preuve sera omise j'ai indiquée HP (= hors programme) ; d'autres preuves pourraient être faites ou pas, d'après le temps disponible. Dans ce cas, quand le type de raisonnement est similaire à d'autres et le niveau de difficulté aussi, on pourrait par exemple demander à les reconstruire par exercice (mais, ne vous inquiétez pas, un contrôle ne pourra pas se baser que sur ça).

# Chapitre 1

# Les fonctions dans $\mathbb{R}$ et leurs limites

## 1.1 Fonctions réelles d'une variable réelle

Nous rappelons ici certaines des notions qu'on a déjà vu à propos du concept de fonction en général, celles qui sont plus spécifiques aux fonctions définies sur des ensembles de nombres réels et à valeurs nombres réels, et on précisera les notions dont on a besoin dans le cadre de $\mathbb{R}$.

### 1.1.1 Qu'est-ce que c'est ?

Commençons par une définition informelle : nous avons à disposition deux ensembles de nombres réels, deux parties de $\mathbb{R}$, l'ensemble de tous les nombres réels, $D$ et $A$. Une fonction, $f$, de $D$ vers $A$ est un « objet mathématique » qui, à tout nombre, tout élément de $D$ associe ou fait correspondre un unique élément de $A$.

1. $D$ est appelé l'**ensemble de départ**, la **source** ou le **domaine de définition** de la fonction $f$
2. $A$ est appelé l'**ensemble d'arrivée** ou le **but** de la fonction $f$
3. Si $x$ est un élément de $D$, on note $f(x)$ l'élément de $A$ associé à $x$ par la fonction $f$. Cet élément $f(x)$ est appelé l'**image** de $x$ par $f$.
4. Si on écrit « Soit une fonction $f : D \to A$... », on entend par là « Considérons une fonction $f$, dont l'ensemble de départ est $D$ et l'ensemble d'arrivée est $A$ », et, à moins que ce ne soit précisé ultérieurement, cette fonction n'a aucune propriété particulière hormis le fait d'être un objet qui satisfait à notre définition informelle.

La fonction $f$ est « de variable réelle » car le domaine dans lequel varie son argument, sa source, est une partie de $\mathbb{R}$. Elle est « à valeurs réelles » car son but est une partie de $\mathbb{R}$.

– Si on écrit « Soit la fonction $f : [0,1] \to \mathbb{R}$ définie par $f(x) = 2x^2 - 1$ pour tout $x \in [0,1]$ », on considère un objet mathématique uniquement défini : **la** fonction $f$, dont l'ensemble de départ est l'intervalle $[0,1]$, l'ensemble d'arrivée est $\mathbb{R}$ tout entier, qui associe à tout élément $x$ de $[0,1]$ l'unique nombre réel calculé par la formule ci-avant.
– On doit souligner que si $f : D \to A$ est une fonction, si $x$ est un nombre n'appartenant pas à $D$ alors $f(x)$ n'est pas défini.
– Dans l'exemple précédent $f(2)$ n'a donc pas de sens même si $2 \times 2^2 - 1$ en a un.

### Graphes

Plaçons-nous dans le contexte précédent. Soit $f : D \to A$ une fonction réelle de variable réelle. Son graphe est la partie $G$ de l'ensemble-produit [1] $D \times A$ définie par

$$G = \{(x, f(x)), x \in D\} = \{(x, y) \in D \times A,\ y = f(x)\}$$

$G$ est donc l'ensemble de **tous** les couples possibles formés d'une valeur $x$ prise dans $D$ et de son image $f(x)$.

$G$ est une partie de l'ensemble $D \times A$ qui est lui-même une partie de $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$. On représente graphiquement (une partie de) $\mathbb{R}^2$ sur une feuille quadrillée de la façon que vous connaissez bien : le couple de réels $(x, y)$ est représenté par le point de coordonnées $(x, y)$ relativement au quadrillage.

L'usage veut que l'axe des abscisses (coordonnée $x$) soit représenté horizontalement, orienté de la gauche vers la droite et que l'axe des ordonnées (coordonnée $y$) soit représenté verticalement, orienté du bas vers le haut. Certaines situations peuvent forcer à adopter d'autres conventions.

En retournant notre manche, on peut maintenant préciser l'« objet mathématique » dont nous parlions dans la définition informelle de fonction.

**Définition 1.1.1.** Soient $D$ et $A$ deux parties de $\mathbb{R}$, une fonction $f$ de $D$ vers $A$ est une partie $G$ de $D \times A$ ayant la propriété suivante :

$$\text{Pour tout } x \in D, \text{ il existe un unique } y \in A \text{ tel que } (x, y) \in G.$$

L'astuce réside en ceci, qu'étant donnée une telle partie $G$, on peut définir, pour $x \in D$, $f(x)$ comme étant l'unique $y \in A$ tel que $(x, y) \in G$. On a alors défini sans ambiguïté une fonction $f : D \to A$. Il s'avère a posteriori que $G$ est le graphe de $f$.

Ce que dit la définition, c'est très exactement qu'une fonction $f$, c'est son graphe. L'usage montre cependant qu'utiliser le graphe de la fonction est assez malcommode alors que la notation $f(x)$ est très parlante.

### Composition

L'opération la plus importante que l'on puisse réaliser avec deux fonctions est leur composition : On dispose de deux fonctions $f : A \to B$ et $g : C \to D$.

Si pour tout élément $x$ de $A$, $f(x)$ (qui est élément de $B$ a coup sûr) appartient à $C$, on peut alors calculer $g(f(x))$. On peut donc associer à tout $x \in A$, $g(f(x)) \in D$ et donc définir une nouvelle fonction, notée $g \circ f$, dont la source est $A$ et le but est $D$.

En résumé, si $f : A \to B$ et $g : C \to D$ sont deux fonctions, la composée $g \circ f : A \to D$ est définie par

$$(g \circ f)(x) = g(f(x)), x \in A$$

Cette définition n'a de sens que si $f(x) \in C$, la source de $g$, pour tout $x \in A$, la source de $f$.

Dès que l'on écrit une formule imbriquant des fonctions élémentaires, on est en train de faire un certain nombre de compositions.

## 1.1.2 Définir une fonction par des formules

La façon la plus usuelle de définir une fonction d'une variable, disons $x$, est d'utiliser un dictionnaire de fonctions élémentaires (exp, ln, cos, etc.), les opérations usuelles de l'arithmétique,

1. $D \times A$ est l'ensemble de tous les couples $(x, y)$ pouvant être formés avec une valeur $x$ dans $D$ et une valeur $y$ dans $A$.

et de combiner tout ceci en une *formule* comportant la variable $x$ et pouvant être calculée pour certaines valeurs de cette variable.

*Exemples et remarques* 1.1.1.

1. La locution

$$\text{Soit } f : [-1,1] \to \mathbb{R} \text{ la fonction définie par } f(x) = \exp(x^2+3) \text{ pour tout } x \in [-1,1]$$

   définit parfaitement la fonction $f$. On a $f(0) = e^3$, $f(1) = e^4$, etc. Par contre $f(2)$ n'est pas défini même si la formule définissant $f$ a un sens lorsque $x = 2$ : on a imposé que le domaine de définition de $f$ soit l'intervalle $[-1,1]$.

   La formule doit, d'une part, être syntaxiquement correcte et d'autre part, on doit être en mesure d'identifier les valeurs de la variable $x$ pour lesquelles le calcul ne peut aboutir. Les raisons pour lesquelles ce calcul ne peut aboutir sont par exemple, une division par 0, la prise du logarithme d'un nombre négatif...

2. C'est ce type de problème qui nous a amenés, plus haut, à donner une condition sur les fonctions $f : A \to B$ et $g : C \to D$ pour que la composée $g \circ f$ soit bien définie sur tout $A$. Déterminer l'ensemble de définition d'une telle formule, c'est identifier les valeurs de la variable $x$ pour lesquelles le calcul aboutira. C'est aussi, en un certain sens, identifier le domaine source de la fonction que nous sommes en train de définir.

   Ce type de question mène souvent à la résolution, en cascade, d'une suite d'équations et/ou d'inéquations en se basant sur le principe que $f(g(x))$ n'est défini pour une certaine valeur de $x$ qu'à partir du moment où, à la fois, $y = g(x)$ est bien défini et $f(y)$ est bien défini.

3. On veut définir une fonction $g$ à valeurs réelles sur un certain domaine $D$ de $\mathbb{R}$ par la formule

$$g(t) = \sqrt{-\ln(t^2-1)}, \text{ pour tout } t \in D$$

   La question pertinente est « Quel est le plus grand domaine $D$ possible sur lequel on peut définir cette fonction $g$ ? »

   Pour que $g(t)$ soit défini, il faut (et il suffit) que

$$\left\{ \begin{array}{ll} & \text{(a) } t^2-1>0 \text{ car le domaine de définition de } \ln \text{ est } ]0,+\infty[. \\ \text{et} & \\ & \text{(b) } -\ln(t^2-1) \geq 0 \text{ car le domaine de définition de } \sqrt{\ } \text{ est } [0,+\infty[. \end{array} \right.$$

   On tombe donc sur un système d'inéquations d'inconnue $t \in \mathbb{R}$ qu'il faut résoudre. Ce système est équivalent à

$$\left\{ \begin{array}{ll} & \text{(a) } t>1 \text{ ou } t<-1 \\ \text{et} & \\ & \text{(b) } t^2-1 \leq 1 \text{ car } \ln X \leq 0 \text{ si et seulement si } X \leq 1 \end{array} \right.$$

   c'est-à-dire

$$\left\{ \begin{array}{ll} & \text{(a) } t>1 \text{ ou } t<-1 \\ \text{et} & \\ & \text{(b) } -\sqrt{2} \leq t \leq \sqrt{2} \end{array} \right.$$

   Pour résumer, on a donc $D = \left[-\sqrt{2}, -1\right[ \cup \left]1, \sqrt{2}\right]$.

4. On peut utiliser des définitions par morceaux. On pose la même question que précédemment en voulant définir $g$ sur un domaine $D$ de $\mathbb{R}$ par la formule, pour tout $t \in D$,

$$g(t) = \left\{ \begin{array}{ll} \sqrt{-\ln(t^2-1)} & \text{si } t>0 \\ \tan t & \text{si } t \in ]-\pi,0[ \end{array} \right.$$

   Pour que $g(t)$ soit défini, il faut (et il suffit) que

$$\left\{ \begin{array}{ll} & \text{(a) soit } t>0 \text{ et } \sqrt{-\ln(t^2-1)} \text{ est bien défini} \\ \text{et} & \\ & \text{(b) soit } t \in ]-\pi,0[ \text{ et } \tan t \text{ est bien défini} \end{array} \right.$$

c'est-à-dire

$$\left\{\begin{array}{l} \text{(a) soit } t>0 \text{ et } t\in\left[-\sqrt{2},-1\right[\cup\left]1,\sqrt{2}\right] \text{ (on se sert ici de l'exemple précédent)} \\ \text{et} \\ \text{(b) Soit } t\in\left]-\pi,0\right[ \text{ et } t\neq-\dfrac{\pi}{2} \text{ (quel est le domaine de définition de la tangente ?)} \end{array}\right.$$

En résumé, le domaine $D$ cherché est

$$D=\left]-\pi,-\frac{\pi}{2}\right[\cup\left]\frac{\pi}{2},0\right[\cup\left]1,\sqrt{2}\right]$$

**Les fonctions usuelles plus utiles**

Nous rappelons ici le rôle joué par certaines fonctions (ou formules) qu'on rencontrera souvent, et qu'on a déjà rencontrées dans ce chapitre.

- **Les polynomes :** c'est peut-être les expressions qu'on connait le mieux, celles où l'on prend la variable $x$ et on en calcule une puissance entière $x^n = x\times x\times\cdots\times x$ (où le produit a $n$ facteurs, tous égaux à $x$) ; si ensuite on prend plusieurs de ces expressions, on les multiplies fois des coefficients et on les ajoute, on trouve un polynome, de la forme, $a_0+a_1x+a_2x^2+\cdots+a_nx^n$ (les coefficients $a_i$ étant des nombres réels).
- **Les racines et les puissances rationnelles :** que signifie-t-il $x^{1/n}$ ? il signifie qu'on prend le nombre tel que, élevé à la puissance $n$, on retrouve $x$, c'est-à-dire $\sqrt[n]{x}$, pour respecter les propriétés des puissance et faire en sorte que $(x^{1/n})^n = x$. Ce nombre n'est pas bien défini tout le temps : si $n$ est impaire aucun problème, pour tout $x$ il existe unique un nombre $y$ tel que $y^n = x$, mais si $n$ est paire 1) cela marche juste pour $x\geq 0$ 2) pour tout $x>0$ il y a même deux nombres (de signe opposé) avec cette propriété. On choisit alors, par convention, d'indiquer par $\sqrt[n]{x}$ celui qui est non-négatif, et cette expression n'a un sens que si $x\geq 0$. Et que signifie-t-il $x^{p/q}$ (puissances rationnelles) ? il signifie juste $\sqrt[q]{x^p}$. Pour éviter des problèmes de définition on préfère dire que cela n'est défini que pour $x>0$. Pourquoi ? parce qu'on voudrait bien que $x^{1/3}$ désigne la même quantité que $x^{2/6}$. Or, la deuxième, s'agissant d'une racine sixième (et 6 étant paire) donne toujours un résultat positif, et elle est même définie si $x<0$ parce que de toute manière on prendra la racine sixième de $x^2$. Ceci ne colle pas avec le comportement de $x^{1/3}$, mais il n'y a pas d'ambiguité si on s limite à $x\geq 0$. Et pourquoi pas $x=0$ ? parce que ça poserait des problèmes pour $p<0$ (et si on considère que $p/q=(-p)/(-q)$, ça poserait toujours des problèmes).
- **Les exponentielles :** Tout d'abord parlons de $x^\alpha$, avec $\alpha\in\mathbb{R}$. considérons $x>0$ : alors on peut toujours définir $x^r$ pour tout $r$ rationnel. Par une procédure de limite, qu'on ne définit pas ici, en prenant des rationnels $r$ qui approchent $\alpha$, on peut définir $x^\alpha$ aussi (comme la limite des puissances rationnelles quand les exposants ont $\alpha$ comme limite). Cela nous permet de définir toutes les puissances avec base positive et exposant quelconque. Alors on peut également mettre le $x$ en haut et considérer $a^x$ (avec $a>0$). Parmi toutes les fonctions de ce type $2^x$, $3^x$, $0.001^x$...il y en a une spéciale : c'est $e^x$, qui est souvent appelée LA fonction exponentielle et on écrit aussi $\exp(x)$. Ce nombre $e=2,71828183\ldots$ est un nombre irrationnel qui a plusieurs propriétés, dont une est le fait que la dérivée (on verra ensuite de quoi l'on parle) de la fonction $e^x$ est elle-même[2], c'est-à-dire c'est encore la fonction $e^x$. Ce même nombre est aussi la limite (là aussi, on en parlera dans pas longtemps) de $(1+x)^{1/x}$ lorsque $x>0$ approche 0. La fonction $e^x$ est strictement croissante (ce qui est le cas de toute fonction $a^x$ pour $a>1$, parce qu'on prend un nombre plus grand que 1 et on l'élève à $x$, donc si $x$ augmente le résultat augmente, au contraire de ce qui se passe pour des $a$ petits) et elle est donc inversible. Son inverse s'appelle le logarithme mais on en parlera ensuite.

2. et cela est à la base de plein de blagues matheuses sur les fêtes des fonctions ou similaires, que vous connaissez sans doute

– **Sinus et cosinus :** Ces deux fonctions ont une origine géométrique : prenez un cercle de rayon 1 centré en $(0,0)$ dans le plan ; bougez dans le sens contraire des aiguilles d'une montre en parcourant une longueur $x$ à partir du point $(1,0)$ (par exemple, si vous prenez $x < 2\pi$ vous n'avez pas terminé un tour complet, pour $x > 2\pi$ vous pouvez éventuellement faire plusieurs tours) ; regardez où vous arrivez et appelez $\cos x$ (cosinus de $x$) l'abscisse du point où vous êtes arrivées et $\sin x$ (sinus de $x$) son ordonnée. Ceci est une définition géométrique, qui montre bien que ces fonctions sont périodiques (si vous rajoutez $2\pi$ à $x$ vous ne faites que rajouter un tour, ce qui ne change pas votre point d'arrivée). Elle cache d'autres propriétés, dont probablement la plus importante est le fait que, si $x$ tend vers 0, alors $\sin x/x$ tend vers 1 : on peut s'en convaincre aussi par cette définition, si nous considérons que pour $x$ très petit le déplacement qu'on fait est petit mais surtout presque vertical ; il est donc vrai que la distance parcourue et l'ordonnée d'arrivée sont presque la même chose.
– **D'autres fonctions trigonométriques :** Avec sinus et cosinus on peut construire ensuite plein d'autres fonctions, dont la plus importante est la tangente : $\tan x := \sin x/\cos x$. Seul problème : diviser par 0. La tangente sera donc définie sur les points où $\cos x \neq 0$ seuls, c'est-à-dire en tout point sauf ceux de la forme $\pi/2 + k\pi$ avec $k$ entier (parce qu'avec $\pi/2$, i.e. un quart de tour, on arrive avec abscisse nulle, et cela reste vrai en rajoutant un nombre entier de demi-cercles).

D'autres fonctions, et notamment les fonctions réciproques des fonctions trigonométriques, et ben sûr le logarithme, seront importantes aussi, mais on va les présenter quand on parlera un peu plus en détails des fonctions réciproques.

## 1.2 Limites finies en un point $x_0$ de $\mathbb{R}$

La notion de limite est à la base de l'analyse. Nous allons rappeler les résultats vus en Première et Terminale et préciser quelques notions.

### 1.2.1 Voisinages, ouverts, adhérence, fermés

On commence par une définition essentielle pour les limites.

**Définition 1.2.1.** Un ensemble $V \subset \mathbb{R}$ est un voisinage d'un point $x_0 \in \mathbb{R}$ s'il existe un intervalle du type $]x_0 - h, x_0 + h[$, avec $h > 0$, contenu dans $V$.

*Exemples et remarques* 1.2.1. 1. Un intervalle $]a, b[$ est voisinage de chacun de ses points
2. $]0, 1]$ est voisinage de tout $x_0 \in ]0, 1[$. Il n'est pas voisinage de 1.
3. Soit $A = \mathbb{R}^*$. $A$ est voisinage de $x_0$ si $x_0 \neq 0$. $A$ n'est pas voisinage de 0.

La notion de voisinage est utile pour définir ce qu'un point adhérent à un ensemble.

**Définition 1.2.2.** Soit $A \subset \mathbb{R}$ un ensemble quelconque et $x_0 \in \mathbb{R}$. On dit que $x_0$ est adhérent à $A$ si pour tout voisinage $V$ de $x_0$ il y a un point $y \in A \cap V$.

*Exemples et remarques* 1.2.2. 1. Tout point $x_0 \in A$ est adhérent à $A$ (il suffit de prendre $y = x_0$).
2. Être adhérent à $A$ ne signifie pas forcément appartenir à $A$, mais plutôt être "collé" à $a$. Notamment les bornes d'un intervalle ouvert sont adhérentes à ce même intervalle.
3. Les points d'adhérence de $\mathbb{Q}$ sont tous les points de $\mathbb{R}$.

Tant qu'à faire, on peut également donner la définition d'ensemble ouvert et d'ensemble fermé, deux définitions qui vont être importantes à plusieurs reprises.

**Définition 1.2.3.** Soit $A$ un sous-ensemble de $\mathbb{R}$. On dit que $A$ est *ouvert* si il est un voisinage de tous ses points : autrement, dit, si pour tout $x \in A$ il existe un rayon $r > 0$ tel que $]x-r, x+r[= \{y \in \mathbb{R} : |x-y| < r\} \subset A$.

On dit que $A$ est *fermé* si son complémentaire est ouvert.

*Exemples et remarques* 1.2.3.

1. La définition d'ensemble ouvert généralise la notion d'intervalle ouvert ("sans ses bornes") à des ensembles qui ne sont pas des intervalles). Une réunion de plusieurs intervalles ouverts est ouverte, par exemple, et cela vaut aussi pour les réunions infinies.
2. L'ensemble $A = ]0, \infty[$ n'est pas fermé car son complémentaire n'est pas voisinage de $x_0 = 0$. Par contre l'ensemble $A = [0, \infty[$ et plus en général tout intervalle ou demi-droite qui inclut ses points extremaux est fermé.
3. Il ne faut pas penser que tout ensemble est soit ouvert soit fermé : par exemple l'ensemble $[0, 1[$ n'est ni ouvert ni fermé et l'ensemble vide est en même temps ouvert et fermé (et $\mathbb{R}$ aussi).

La notion d'ensemble fermé est en effet liée à la notion de points d'adhérence. En effet, notons $\bar{A}$ l'ensemble des points d'adhérence de $A$ :

$$\bar{A} = \{x \in \mathbb{R} : \forall V \text{ voisinage de } x \text{ on a } V \cap A \neq \emptyset\}.$$

Nous avons alors

**Proposition 1.2.1.** *Pour tout $A \subset \mathbb{R}$, l'ensemble $\bar{A}$ est le plus petit ensemble fermé contenant $A$. En particulier, un ensemble $A$ est lui même fermé si et seulement si $\bar{A} \subset A$, c'est-à-dire si et seulement s'il contient tous ses points d'adhérence.*

*Démonstration.* **HP** Il faut d'abord montrer que $\bar{A}$ est fermé. Prenons son complémentaire : c'est l'ensemble des points $x$ tels qu'il existe un voisinage $V$ de $x$ avec $V \cap A = \emptyset$. Or, par définition de voisinage, il y a un $h > 0$ tel que $]x-h, x+h[\subset V$ et $V$ est également un voisinage de tout point $y \in ]x-h, x+h[$. Donc $]x-h, x+h[$ est inclus dans le complémentaire de $\bar{A}$ aussi, qui est donc ouvert (parce que pour tout $x \in \bar{A}$ il y a $h > 0$ tel que $]x-h, x+h[$ est inclus dans ce complémentaire.

Il faut ensuite montrer que c'est le plus petit fermé contenant $A$. Soit alors $F$ un fermé contenant $A$. Prenons $x \in F^c$ (dans le complémentaire de $F$). $F^c$ est ouvert donc c'est un voisinage de $x$. Mais $F^c \cap F = \emptyset$ donc, a fortiori $F^c \cap A = \emptyset$ (car $A \subset F$). Alors $F^c$ est un voisinage de $x$ disjoint de $A$, donc $x \notin \bar{A}$. Cela montre que tout point de $\bar{A}$ doit appartenir à $F$, et donc $\bar{A} \subset F$. Mais alors c'est vrai que $\bar{A}$ est le plus petit fermé contenant $A$.

La deuxième partie suit facilement : $A$ est fermé si et seulement si il coïncide avec le plus petit fermé qui le contient, donc avec $\bar{A}$. Mais comme il est toujours vrai que $A \subset \bar{A}$ on peut également dire que $A$ est fermé si et seulement si $\bar{A} \subset A$. □

### 1.2.2 Définition de limite

**Définition 1.2.4.** Soit $\ell \in \mathbb{R}$, $f : D \to \mathbb{R}$ et $x_0$ un point d'adhérence de $D$. On dit que $f$ admet la limite $\ell$ en $x_0$, ou que la limite de $f(x)$ lorsque $x$ tend vers $x_0$ (sous-entendu "et $x \in D$") est $\ell$ si

Pour tout voisinage $V$ de $\ell$,
il existe un voisinage $U$ de $x_0$ tel que,
pour tout $x \in U \cap D$,
$f(x) \in V$.

On note ce fait $\lim_{x\to x_0} f(x)=\ell$ ou $f(x)\to\ell$ lorsque $x\to x_0$, ou $f(x)\overset{x\to x_0}{\to}\ell$.

Si on traduit les définitions de voisinage à l'aide des intervalles, on arrive à la définition suivante, équivalente :

Pour tout $\varepsilon>0$, il existe un $\delta>0$ tel que pour tout $x\in D\cap ]x_0-\delta, x_0+\delta[$ on a $f(x)\in ]\ell-\varepsilon, \ell+\varepsilon[$,

ce qui peut être encore réécrit comme

Pour tout $\varepsilon>0$, il existe un $\delta>0$ tel que pour tout $x\in D$ satisfaisant $|x-x_0|<\delta$ on a $|f(x)-\ell|<\varepsilon$.

Il est souvent utile de restreindre, dans la définition de limite, la fonction $f$ à un ensemble de définition plus petit. Etant donné un ensemble $B$ on peut définir la limite $\lim_{x\to x_0, x\in B} f(x)$ en demandant la condition

Pour tout voisinage $V$ de $\ell$,
il existe un voisinage $U$ de $x_0$ tel que,
pour tout $x\in U\cap D\cap B$,
$f(x)\in V$.

Ceci revient à changer le domaine de définition de $f$ en prenant $D\cap B$ au lieu de $D$ (ce qui est toujours possible, de *restreindre* le domaine). Ceci est particulièrement utile quand on veut par exemple

- exclure le point $x_0$ (prendre $B=\mathbb{R}\setminus\{x_0\}$), pour ignorer la valeur exacte prise par la fonction en ce point, et ne considérer que le comportement de $f(x)$ lorsque $x$ est proche de $x_0$ en en étant distinct [3] Dans ce cas on écrit $\lim_{x\to x_0,\, x\neq x_0} f(x)=\ell$.
  - il est utile de remarquer que, souvent, le point $x_0$ est automatiquement exclu de la limite parce qu'il n'appartient pas au domaine $D$ : il est en effet beaucoup plus intéressant déterminer $\lim_{x\to 0}\frac{\sin x}{x}$, où la fraction n'est pas définie en 0, que de calculer $\lim_{x\to 0}\sin x+(3x+2)^4-e^x$, limite d'une expression parfaitement définie en 0, qui donne $\sin 0+4-e^0=3$, sans aucune surprise.
  - il est également utile de remarquer que, si $x_0\in D$ et qu'on ne l'exclut pas, alors la seule valeur possible de la limite $\ell$ est bien $f(x_0)$ (**exercice !!**).
  - cependant, comme on le verra dans la partie dédiée à la continuité, la première question qu'on se pose est justement "la limite pour $x\to x_0$ est égale ou non à $f(x_0)$ ?", question qu'on peut se poser en retirant ou non le point $x_0$. En effet (le démontrer par **exercice !!**) on peut vérifier que

$$\lim_{x\to x_0} f(x)=\ell \Leftrightarrow \lim_{x\to x_0,\, x\neq x_0} f(x)=\ell \text{ ET } f(x_0)=\ell,$$

    et donc pour vérifier que la limite fasse bien $f(x_0)$ on peut se contenter de regarder la limite avec $B=\mathbb{R}\setminus\{x_0\}$.
- considérer la limite d'un seul côté (limite gauche ou droite) ; ceci revient à prendre $B=]x_0,+\infty[$ ou $B=]-\infty, x_0[$. On écrit alors $\lim_{x\to x_0,\, x>x_0} f(x)=\ell$ (mais parfois on écrit également $\lim_{x\to x_0^+} f(x)=\ell$) ou $\lim_{x\to x_0,\, x<x_0} f(x)=\ell$ ( $\lim_{x\to x_0^-} f(x)=\ell$). En effet la définition de limite s'applique lorsqu'on travaille sur une fonction $f$ et que l'on s'intéresse à la fois à ce qui se passe à droite et à gauche de $x_0$, alors que les limites droites et gauches nous permettent de séparer les deux côtés.

3. Dans certains textes celle-ci est la définition de "limite en $x_0$", c'est-à-dire celle qui exclut automatiquement le point $x_0$. Ce n'est qu'une convention, ce qui est important est de clarifier toujours quels sont les points qu'on prend en considération, et ici nous choisissons de considérer le point $x_0$ aussi. Il y a des avantages et des désavantages dans ce choix, bien entendu.

  – Par exemple $f : \mathbb{R}^+ \to \mathbb{R}$, $x \mapsto f(x) = \frac{x}{|x|}$ n'a pas de limite en 0 mais exhibe une certaine régularité séparément à droite et à gauche de 0. En effet, les deux limites à gauche et à droite existent, la limite globale non.
  – $g : \left]-3, +\infty\right[ \to \mathbb{R}$, $x \mapsto g(x) = \sqrt{3+x}$ n'est définie qu'à droite de $-3$ mais $g(x)$ a un comportement régulier lorsque $x$ s'approche de $-3$ par la droite. Dans ce cas calculer la limite à droite ou la limite globale, du fait de l'intersection avec $D$, revient au même, mais souvent on préfère préciser qu'on ne parle que d'une limite en s'approchant par la droite.
– Au niveau de notation, nous omettons en général d'indiquer l'ensemble sur lequel nous prenons la limite, s'il coïncide juste avec le domaine de la fonction $f$, sauf si ce domaine n'est pas claire à priori. On ne verra donc pas d'écritures $\lim_{x\to x_0, x\in D} f(x) = \ell$. Il peut cependant y avoir des exceptions si le domaine $D$ est inattendu, n'a pas été introduit auparavant et n'est pas évident à deviner ; par exemple : si on considère la fonction $f(x) = \sin x$ sur les points $x \in \mathbb{Q}$ uniquement, alors probablement on écrira $\lim_{x\to x_0, x\in\mathbb{Q}} \sin x$, juste pour préciser au lecteur le domaine.

Les notions de limites droite et gauche peuvent être utile lorsqu'il s'agit de traiter le cas d'une fonction définie par une alternative du fait du résultat suivant

**Proposition 1.2.2.** *Soit $f$ une fonction réelle de variable réelle définie sur $D$ et $x_0 \in \bar{D}$. Sont équivalentes les propositions suivantes*

1. *$\ell$ est limite de $f(x)$ lorsque $x \to x_0$, $x \neq x_0$*
2. *$\ell$ est la limite à droite de $f(x)$ lorsque $x \to x_0$, mais aussi la limite à gauche de $f(x)$ lorsque $x \to x_0$.*

*Egalement, les propositions suivantes sont équivalentes*

1. *$\ell$ est limite de $f(x)$ lorsque $x \to x_0$*
2. *$\ell$ est la limite à droite de $f(x)$ lorsque $x \to x_0$, mais aussi la limite à gauche de $f(x)$ lorsque $x \to x_0$, et $f(x_0) = \ell$.*

**DL à l'ordre** 0

**Proposition 1.2.3.** *Étant donnée $f : D \to \mathbb{R}$ une fonction définie sur $D$, $x_0 \in \bar{D}$ et $\ell \in \mathbb{R}$, $f$ tend vers $\ell$ en $x_0$ si et seulement si, il existe une fonction $\epsilon$, définie au voisinage de 0, $\lim_{h\to 0} \epsilon(h) = 0$ telle que*

$$\textit{Pour tout } x \textit{ au voisinage de } x_0,\ f(x) = \ell + \epsilon(x - x_0)$$

Le fait d'écrire, pour $x$ voisin de $x_0$, que $f(x) = \ell + \epsilon(x - x_0)$ s'appelle effectuer un Développement Limité de $f$ à l'ordre 0 en $x_0$.

Nous verrons, au chapitre **??**, ce qu'est un DL de $f$ en $x_0$ à un certain ordre $n \in \mathbb{N}$.

## 1.3 Propriétés

### 1.3.1 Unicité, gendarmes

**Unicité**

**Proposition 1.3.1.** *Soit $x_0 \in \bar{D}$, $\ell, \ell' \in \mathbb{R}$ et $f : D \to \mathbb{R}$ une fonction définiesur un ensemble dont $x_0$ est point d'adhérence. Si*

$$\lim_{x\to x_0} f(x) = \ell \textit{ et } \lim_{x\to x_0} f(x) = \ell'$$

*alors*

$$\ell = \ell'$$

*Démonstration.* Supposons que $\ell \neq \ell'$, on peut alors trouver $V$ et $V'$ des voisinages respectivement de $\ell$ et $\ell'$ tels que $V$ et $V'$ sont deux parties disjointes (par exemple il suffit de prendre deux intervalles du type $]l-\varepsilon, l+\varepsilon[$ et $]l'-\varepsilon, l'+\varepsilon[$ avec $\varepsilon < |l-l'|/2$).

De la définition de $f(x) \to \ell$ lorsque $x \to x_0$, on tire l'existence d'un voisinage $U$ de $x_0$ tel que si $x \in U \cap D$, alors $f(x) \in V$ et, de même, de la définition de $f(x) \to \ell'$ lorsque $x \to x_0$, on tire l'existence d'un voisinage $U'$ de $x_0$ tel que si $x \in U' \cap D$, $x \neq x_0$, alors $f(x) \in V'$.

Puisque $U \cap U'$ est encore un voisinage de $x_0$ (il contient le plus petit des deux intervalles centrés en $x_0$ qui étaient contenus en $U$ ou $U'$), du fait que $x_0$ est adhérent à $D$ on trouve l'existence d'un point $y \in U \cap U' \cap D$.

La contradiction provient du fait que son image $f(y)$ est donc à la fois dans $V$ et $V'$, ce qui est impossible. □

**Théorème des gendarmes**

**Proposition 1.3.2.** *Soit $x_0 \in \mathbb{R}$, $f$, $g$, et $h$ trois fonctions à valeurs réelles définies sur $D$, et $x_0 \in \bar{D}$, telles que*

$$\text{pour tout } x \in D \text{ , } f(x) \leq g(x) \leq h(x).$$

*Alors*

1. *Si $f$, $g$, $h$ ont pour limites respectives $\ell$, $m$, $n$ en $x_0$, on a*
$$\ell \leq m \leq n$$
2. *Si $f$ et $h$ ont même limite $\ell$ en $x_0$, $g$ admet $\ell$ comme limite en $x_0$.*

La preuve de 1. est basée sur le même raisonnement que la preuve d'unicité. Si on suppose que $m < \ell$, on prend $\varepsilon$ tel que $m+\varepsilon < l-\varepsilon$ ; ensuite on peut trouver un point $y$ dans un certain voisinage de $x_0$ tel que $f(y) > l-\varepsilon > m+\varepsilon > g(y)$, ce qui apporte une contradiction.

La preuve de 2. n'est pas différente, mais le point important est le fait que l'existence de la limite de $g$ n'est pas supposée dans les hypothèses, elle est déduite du comportement de $f$ et $h$.

*Exemples et remarques* 1.3.1.

1. Cette proposition fournit la **méthode** fondamentale pour montrer que $f(x) \to \ell$ lorsque $x \to x_0$. Il suffit de trouver une majoration du type
$$|f(x_0+h) - \ell| \leq \epsilon(h)$$
pour $h$ voisin de 0 où $\epsilon$ est une fonction de limite 0 en 0. On peut en général prendre une telle fonction $\epsilon$ dans une liste de fonctions de référence : par exemple $\epsilon(h) = C|h|^\alpha$ avec $\alpha > 0$, $C$ une constante positive.
2. Par exemple, montrons que la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2 + 2x + 3$ a pour limite $f(1) = 6$ en $x_0 = 1$. On a
$$\begin{aligned} f(1+h) - f(1) &= (1+h)^2 + 2(1+h) - 1^2 - 2 \\ &= h(2+h) + 2.h = h(4+h) \end{aligned}$$
On travaille pour $h$ voisin de 0, on peut donc supposer que $|h| \leq 1$. On a donc, pour ces $h$,
$$|f(1+h) - f(1)| \leq 5|h|$$
Comme l'expression à droite de cette inégalité définit une fonction $\epsilon(h)$ de limite 0 en 0, on a donc démontré que
$$\lim_{x \to 1} f(x) = f(1)$$

#### Fonctions de références

Les fonctions suivantes sont des exemples fondamentaux de fonctions de limites nulle en 0.

1. Soit $\alpha > 0$. La fonction $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) = |x|^\alpha$ a pour limite 0 en 0.
2. La fonction $g : \mathbb{R}^* \to \mathbb{R}$, $x \mapsto x.(\ln |x|)$ a pour limite 0 en 0.
3. Soit $\alpha > 0$, $\beta \in \mathbb{R}$. La fonction $h : \mathbb{R}^* \to \mathbb{R}$, $x \mapsto |x|^\alpha.|\ln |x||^\beta$ a pour limite 0 en 0.

### 1.3.2 Quelques résultats utiles

#### $f$ bornée localement

**Proposition 1.3.3.** *Si $f$ admet une limite $\ell \in \mathbb{R}$ en $x_0$ alors il existe un voisinage $U$ de $x_0$ tel que $f$ est bornée sur $U \cap D$ .*

Il suffit pour cela de considérer le voisinage $V = ]\ell - 1, \ell + 1[$ de $\ell$ et d'appliquer la définition de limite.

#### Positivité sur un voisinage

**Proposition 1.3.4.** *Si $f$ admet une limite $\ell > 0$ en $x_0$ alors $f(x) > \frac{1}{2}\ell > 0$ pour tout $x \in U \cap D$ ($U$ étant un certain voisinage de $x_0$). En particulier $f$ reste positive sur $U \cap D$.*

Il suffit pour cela de considérer le voisinage $V = ]\ell/2, 3\ell/2[$ de $\ell$ et d'appliquer la définition de limite.

### 1.3.3 Opérations sur les limites

En général, on connaît un certain nombre de limites de fonctions « simples » : polynômes, exponentielles, etc... et, comme les fonctions que nous étudions sont pour la plupart obtenues grâce à des sommes, produits, quotients, etc... de ces fonctions simples, on veut pouvoir calculer des limites grâce à des moyens « opératoires ».

Dans tous les énoncés suivants, que nous admettons, $f$, $g$ sont des fonctions définies sur un domaine dont $x_0$ est un point d'adhérence.

#### Somme

Si

$$\lim_{x \to x_0} f(x) = \ell \text{ et } \lim_{x \to x_0} g(x) = m$$

alors

$$\lim_{x \to x_0} (f + g)(x) = \ell + m$$

#### Produit

Si

$$\lim_{x \to x_0} f(x) = \ell \text{ et } \lim_{x \to x_0} g(x) = m$$

alors

$$\lim_{x \to x_0} (f.g)(x) = \ell.m$$

**Quotient**

Si

$$\lim_{x\to x_0} f(x) = \ell \text{ et } \lim_{x\to x_0} g(x) = m$$

et $m \neq 0$ alors, la fonction $f/g$ est définie sur un certain voisinage de $x_0$ (plus précisément, sur l'intersection de ce voisinage avec les domaines de définition de $f$ et $g$) et l'on a

$$\lim_{x\to x_0} (f/g)(x) = \ell/m$$

**Composition**

**Proposition 1.3.5.** *Soient $f : D \to \mathbb{R}$, $g : D' \to \mathbb{R}$ deux fonctions réelles d'une variable réelle, définies sur $D$ et $D'$, respectivement, et $x_0$, $y_0$ deux réels. Supposons que $x_0$ est un point d'adhérence de $D$ et que $D'$ est un voisinage de $D'$.*

1. *Si* $\lim_{x\to x_0} f(x) = y_0$,
2. *si* $\lim_{y\to y_0} g(y) = \ell$,

*alors $g \circ f$ est définie sur l'intersection d'un voisinage de $x_0$ avec $D$ et*

$$\lim_{x\to x_0} (g \circ f)(x) = \ell.$$

*Démonstration.* On considère $h = g \circ f$.

– La première question concerne la bonne définition de $h$ au voisinage de $x_0$. Pour que $h$ soit bien définie en un point $x$ il est nécessaire et suffisant que $f(x) \in D'$. Or, nous savons que $D'$ est un voisinage de $y_0$ et que la limite de $f$ en $x_0$ est bien $y_0$. Par définition de limite, il existe un voisinage $U$ de $x_0$ tel que pour tout $x \in U \cap D$ on a $f(x) \in D'$. Ceci montre que $h$ est bien définie sur $U \cap D$.
– Le même type de raisonnement nous donne la limite cherchée. Soit $W$ un voisinage de $\ell$ : comme la limite de $g$ en $y_0$ est $\ell$, on sait qu'il existe un voisinage $V$ de $y_0$ tel que pour tout $y \in V$, $g(y) \in W$. D'autre part, à ce voisinage $V$ correspond, d'après la définition de $\lim_{x\to x_0} f(x) = y_0$, un voisinage $U$ de $x_0$ tel que pour tout $x \in U \cap D$ on a $f(x) \in V$. Ceci implique $h(x) = g(f(x)) \in W$.
– Conclusion : étant donné $W$ voisinage de $\ell$, on a donc trouvé un voisinage $U$ de $x_0$ tel que pour tout $x \in U \cap D$, $h(x) := g(f(x))$ est dans $W$ : c'est exactement la définition du fait que la limite de $h(x)$ lorsque $x \to x_0$ est $\ell$. □

Cela revient, dans la pratique courante, à effectuer un **changement de variable**, $y = f(x)$.

**Exercice :** On rappelle que $\lim_{x\to 0} \frac{\sin x}{x} = 1$. En déduire, en utilisant les formules trigonométriques d'angle double et les règles opératoires sur les limites que

$$\lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}.$$

## 1.4 Limites infinies et en l'infini

### 1.4.1 Voisinages de l'infini

**Définition 1.4.1.** Soit $V$ une partie de $\mathbb{R}$, on dit que

1. $V$ est un voisinage de $+\infty$ si $V$ contient un intervalle du type $]a, +\infty[$, pour un certain $a \in \mathbb{R}$.

2. $V$ est un voisinage de $-\infty$ si $V$ contient un intervalle du type $]-\infty, a[$, pour un certain $a \in \mathbb{R}$.

Pour pouvoir définir les limites en $\pm\infty$ in sera alors nécessaire de dire quand “$+\infty$ (ou $-\infty$) est adhérent” à un ensemble $A$. Or, en appliquant la définition d’adhérence avec les vosinages, on trouve que $+\infty$ est adhérent à $A$ si et seulement si $A$ n’est pas borné supérieurement (et donc “pénètre” dans toute demi-droite $]a, +\infty[$) et que $-\infty$ est adhérent à $A$ si et seulement si $A$ n’est pas borné inférieurement (et donc “pénètre” dans toute demi-droite $]-\infty, a[$).

### 1.4.2 Définition et opérations

**Définition 1.4.2.** Soit $f : D \to \mathbb{R}$ définie sur un domaine $D$ non borné supérieurement, $\ell$ un réel. On dit que la limite $f$ en $+\infty$ est $\ell$ si

Pour tout voisinage $V$ de $\ell$,
il existe un voisinage $U$ de $+\infty$ tel que,
pour tout $x$,
si $x \in U \cap D$, $f(x) \in V$.

On note ceci par $\lim_{x\to+\infty} f(x) = \ell$ ou $f(x) \to \ell$ lorsque $x \to +\infty$.

*Exemples et remarques* 1.4.1.
1. On a évidemment une définition analogue pour la situation en $-\infty$.
2. On a encore unicité de la limite, sous réserve d’existence.
3. La définition de $f(x) \to 0$ lorsque $x \to +\infty$ s’écrit donc, au cas où $f$ est définie sur un voisinage de $+\infty$,

$$\text{Pour tout } \epsilon > 0, \text{ il existe } a \in \mathbb{R} \text{ tel que pour tout } x \in ]a, +\infty[,\ |f(x)| < \epsilon.$$

#### Fonctions de référence

Les fonctions suivantes sont des exemples fondamentaux de fonctions de limites nulle en $+\infty$.

1. Soit $\alpha > 0$, la fonction $f : \mathbb{R}^* \to \mathbb{R}$, $x \mapsto \frac{1}{|x|^\alpha} = |x|^{-\alpha}$ a pour limite 0 en $\pm\infty$.
2. Soit $\alpha > 0$, la fonction $f : \mathbb{R}^* \to \mathbb{R}$, $x \mapsto \frac{\ln|x|}{|x|^\alpha} = |x|^{-\alpha} \ln|x|$ a pour limite 0 en $\pm\infty$.
3. Soit $\alpha \in \mathbb{R}$, $\beta > 0$, la fonction $f : \mathbb{R}^* \to \mathbb{R}$, $x \mapsto |x|^\alpha e^{-|x|^\beta}$ a pour limite 0 en $\pm\infty$.

#### Opérations

Concernant les opérations et les limites en $+\infty$, les résultats vus sur les limites usuelles restent valables en remplaçant systématiquement $\lim_{x\to x_0}$ par $\lim_{x\to+\infty}$.

### 1.4.3 Limites infinies

Il s’agit, dans ce paragraphe, d’exprimer le fait que la limite d’une fonction en un point $x_0$, à gauche où à droite ou en $\pm\infty$ puisse être l’un des deux infinis.

Nous ne donnons que l’une des définitions, le schéma général de telles définitions devant être maintenant familier :

**Définition 1.4.3.** Soit $f : D \to \mathbb{R}$ une fonction définie sur $D$ et $x_0 \in \bar{D}$ un point d’adhérence de $D$ ;on dit que $f$ admet $+\infty$ comme limite en $x_0$ si

Pour tout voisinage $V$ de $+\infty$,
il existe un voisinage $U$ de $x_0$ tel que
pour tout $x \in U \cap D$,
$f(x) \in V$.

De manière similaire on peut définir des limites droites ou gauches en $x_0$ ayant valeur $\pm\infty$, ou exclure le point $x_0$...

Ce qui est important est le critère suivant, qui résume tous les cas :

**Définition générale pour les limites :** si on a une fonction $f : D \to \mathbb{R}$ et $x_0, \ell \in \mathbb{R} \cup \{\pm\infty\}$ ($x_0$ étant adhérent à $D$) on dit que la limite de $f$ lorsque $x \to x_0$ est $\ell$ si pour tout voisinage $V$ de $\ell$ il existe un voisinage $U$ de $x_0$ tel que, pour tout $x \in U \cap D$ on a $f(x) \in V$.

### Fonctions de référence

Les fonctions suivantes sont des exemples fondamentaux de fonctions de limites infinies. Les exemples ci-dessous relèvent de ce que l'on appelle les « croissances comparées », du logarithme, de l'exponentielle et des fonctions puissances, voir la partie **??** du chapitre **??** pour des justifications de ceci.

1. Soit $\alpha > 0$, la fonction $f : \mathbb{R}^* \to \mathbb{R}$, $x \mapsto \frac{1}{|x|^\alpha} = |x|^{-\alpha}$ a pour limite $+\infty$ en 0.
2. La fonction $f : \mathbb{R}^* \to \mathbb{R}$, $x \mapsto \frac{1}{x}$ a pour limite $+\infty$ en $0^+$ et $-\infty$ en $0^-$.
3. La fonction $f : \mathbb{R}^* \to \mathbb{R}$, $x \mapsto \ln|x|$ a pour limite $-\infty$ en 0 et $+\infty$ en $\pm\infty$.
4. Soit $\alpha \in \mathbb{R}$, $\beta > 0$, la fonction $f : \mathbb{R}^* \to \mathbb{R}$, $x \mapsto |x|^\alpha e^{+|x|^\beta}$ a pour limite $+\infty$ en $\pm\infty$.

**Proposition 1.4.1** (Composition et limites infinies)**.** *Si* $\lim_{y\to+\infty} g(y) = \ell$ *où* $\ell$ *est soit un nombre, soit l'un des symboles* $-\infty$ *ou* $+\infty$*, si* $\lim_{x\to x_0} f(x) = +\infty$*, alors*

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

.

*Exemples et remarques* 1.4.2.

1. On a le même type d'énoncé avec $g(y) \to \ell$ lorsque $y \to -\infty$, $f(x) \to -\infty$ lorsque $x \to x_0$.
2. On peut remplacer les « $x \to x_0$ » par des limites lorsque $x \to \pm\infty$ ou lorsque $x \to x_0^\pm$.
3. La preuve de cette proposition est sur le même canevas que la preuve de la composition pour les limites finies, proposition 1.3.5.

### Quelques remarques finales concernant les opérations

On a vu les règles opératoires concernant les limites finies et leur comportement vis à vis des opérations. Quand l'une des limites est infinie ou quand on ne tombe pas dans l'un des cas décrit, on tombe sur ce que l'on appelle **une forme indéterminée**. Il s'agit alors de travailler un peu plus pour « lever » cette indétermination.

Quelques exemples de formes indéterminées (lorsque $x \to x_0$) :

1. $f(x) \to +\infty$, $g(x) \to -\infty$, $f(x) + g(x) \to ??$
2. $f(x) \to \pm\infty$, $g(x) \to \pm\infty$, $f(x) + g(x) \to ??$, $f(x)/g(x) \to ??$
3. $f(x) \to 0$, $g(x) \to 0$, $f(x)/g(x) \to ??$
4. $f(x) \to 0$, $g(x) \to +\infty$, $f(x).g(x) \to ??$

Il y a par contre un certain nombre de règles combinant limites finies et infinies qui sont à connaître

1. $f(x) \to +\infty$, $g(x) \to +\infty$, $f(x) + g(x) \to +\infty$, $f(x).g(x) \to +\infty$
2. $f(x) \to -\infty$, $g(x) \to -\infty$, $f(x) + g(x) \to -\infty$, $f(x).g(x) \to +\infty$
3. $f(x) \to \ell$, $g(x) \to +\infty$, $f(x) + g(x) \to +\infty$,
   (a) Si $\ell > 0$, $f(x).g(x) \to +\infty$,
   (b) si $\ell < 0$, $f(x).g(x) \to -\infty$,
   (c) si $\ell = 0$, indétermination.
4. $f(x) \to \ell$, $g(x) \to \pm\infty$, $f(x)/g(x) \to 0$
5. $f(x) \to \ell > 0$, $g(x) \to 0^+$ (cela signifie que $g$ reste positive au voisinage du point considéré), $f(x)/g(x) \to +\infty$

Voici, sur un certain nombre d'exemples et de techniques pour lever ces indéterminations

1. $0/0$, $\lim_{x\to 0} \frac{\sin x}{x} = 1$, c'est à connaître et une preuve de ceci ne peut reposer que sur la définition même de la fonction sinus (si on utilise la définition géométrique qu'on a donnée, cela revient à vérifier certaines propriétés géométriques du cercle).
2. $+\infty/+\infty$. $\lim_{x\to+\infty} \frac{7x^2-3x+2}{3x^2+5} = \frac{7}{3}$. Ici numérateur et dénominateur tendent vers $+\infty$ (pourquoi est-ce vrai du numérateur ?). La technique consiste à factoriser en haut et en bas la plus grande puissance, qui est le terme dominant lorsque $x \to \infty$ dans une expression polynomiale.

   On a donc, pour $x$ assez grand,

   $$\frac{7x^2 - 3x + 2}{3x^2 + 5} = \frac{x^2}{x^2} . \frac{7 - \frac{3}{x} + \frac{2}{x^2}}{3 + \frac{5}{x^2}}$$

   Les règles opératoires indiquent qu'alors la deuxième fraction tend vers $\frac{7}{3}$.
3. $+\infty - \infty$. $\lim_{x\to+\infty}(\sqrt{x^2 + x + 1} - x) = \frac{1}{2}$ (expression conjuguée, puis technique précédente)
4. $0 \times \infty$. $(3x^2 - x + 2) \sin \frac{1}{x^2} \to 3$ lorsque $x \to +\infty$
5. Les techniques de développements limités du chapitre **??** seront fondamentales pour la levée des indéterminations.

# Chapitre 2

# Continuité

## 2.1 Définition et premières propriétés des fonctions continues

**Définition 2.1.1.** Soit $D$ un sous-ensemble de $\mathbb{R}$. Une fonction $f : D \to \mathbb{R}$ est dite continue au point $x_0 \in D$ si la limite $\lim_{x \to x_0,\, x \in D} f(x)$ existe et est égale à la valeur $f(x_0)$.

De façon équivalente, on peut dire que $f$ est continue au point $x_0$ si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que tout $x \in D$ avec $|x_0 - x| < \delta$ satisfait aussi $|f(x_0) - f(x)| < \varepsilon$ (cela vient de la définition de limite).

Une fonction est dite continue si elle est continue en tout point de son domaine de définition $D$.

*Exemples et remarques* 2.1.1.

1. Attention qu'ici il est nécessaire que $x_0$ appartienne à $D$ et pas seulement à $\bar{D}$ (parce qu'il faut comparer la limite à la valeur $f(x_0)$).
2. On peut remarquer que d'après cette définition toute fonction $f$ est continue en tout point isolé de son ensemble de définition. Si par exemple $D = [0, 1] \cup \{2\}$ la fonction $f$ est sûrement continue au point 2 car pour tout $\varepsilon > 0$ il suffit de choisir $\delta = 1/2$ : de cette façon le seul point $x \in D$ avec $|x - 2| < \delta = 1/2$ sera le point 2 lui-même et la condition $|f(x) - f(2)| < \varepsilon$ sera verifiée car $f(x) = f(2)$ (une conséquence de $x = 2$) !
3. Les fonctions ln, exp, $x \mapsto |x|^\alpha$, les fonctions trigonométriques usuelles sont continues en tout point de leurs ensembles de définition respectifs.
4. De ce qui a été dit sur les limites, on a par exemple que toute fonction polynomiale définie sur $\mathbb{R}$ est continue en tout $x_0 \in \mathbb{R}$.

Lorsque la fonction $f$ n'est pas définie en $x_0$, on a dit qu'on ne peut pas parler de continuité de $f$ en $x_0$. Cependant, par un mécanisme de prolongement, on peut parfois obtenir une « variante » de $f$, une **extension** de $f$, qui est continue en $x_0$.

**Proposition 2.1.1.** *Soit $f : D \to \mathbb{R}$, et $x_0 \in \bar{D} \setminus D$. Sont équivalentes les assertions suivantes*

1. *Il existe une fonction $g : D \cup \{x_0\} \to \mathbb{R}$, continue en $x_0$, égale à $f$ sur $D$.*
2. $\lim_{x \to x_0} f(x) = \ell \in \mathbb{R}$ *existe (et est réelle).*

*Dans ce cas, la fonction $g$ est unique, elle est définie par*

$$g(x) = f(x), \text{ si } x \in D,\ g(x_0) = \ell = \lim_{x \to x_0} f(x).$$

*$g$ est appelée le* **prolongement par continuité** *de $f$ au point $x_0$.*

*Exemples et remarques* 2.1.2.

1. La fonction « sinus cardinal », utile en traitement du signal, est définie, pour $x \in \mathbb{R}$ par $sinc(x) = \frac{\sin x}{x}$ si $x \neq 0$, $sinc(0) = 1$. Il s'agit du **prolongement par continuité en** 0 de la fonction définie par $x \in \mathbb{R} \setminus \{0\} \mapsto \frac{\sin x}{x}$.

2. $f(x) = x \sin \frac{1}{x}$ est prolongeable par continuité en 0 car $f(x) \to 0$ lorsque $x \to 0$, $x \neq 0$.
3. On entend parfois dire que la fonction $f(x) = 1/x$ est discontinue, ou qu'elle a une discontinuité en 0. Cela est faux : elle est une belle fonction qui est continue en tout point de son domaine $D = \mathbb{R} \setminus \{0\}$ ; le point 0 n'appartient pas au domaine. Par contre, son problème est que la limite en 0 n'existe pas (elle a une limite droite $+\infty$ et gauche $-\infty$, ce qui donne un double problème, que les deux limites ne sont pas finies, et qu'elles sont différentes), et cela empêche d'en faire un prolongement par continuité en 0.

On peut également définir la continuité à droite et à gauche en un point $x_0$.

**Définition 2.1.2.** Soit $f : D \to \mathbb{R}$ une fonction définie sur un domaine comprenant $x_0 \in D$. On dit que $f$ est **continue à droite** en $x_0$ si

$$\lim_{\substack{x \to x_0 \\ x > x_0}} f(x) = f(x_0).$$

On a une définition semblable concernant la continuité à gauche.

*Exemples et remarques* 2.1.3. 1. Soit $H$ définie sur $\mathbb{R}$ par $H(x) = 1$ si $x \geq 0$, $H(x) = 0$ si $x < 0$. $H$ est continue à droite, elle n'est pas continue à gauche.
2. Soit $\widetilde{H}$ définie sur $\mathbb{R}$ par $\widetilde{H}(x) = 1$ si $x > 0$, $\widetilde{H}(x) = 0$ si $x \leq 0$. $\widetilde{H}$ est continue à gauche, elle n'est pas continue à droite.

**Proposition 2.1.2.** *Une fonction est continue en un point $x_0$ si et seulement elle y est continue à gauche et à droite.*

**Exercice :** étudier, suivant les valeurs du paramètre réel $a$, la continuité en 0 de la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = \begin{cases} \frac{\sin 2x}{x} & \text{si } x < 0 \\ x + a & \text{si } x \geq 0 \end{cases}$$

*Indication :* La limite à gauche en 0 de cette fonction est 2, sa limite à droite y est $a$, qui est par ailleurs sa valeur en 0, $f$ est donc continue en 0 si et seulement si $a = 2$.

### 2.1.1 Quelques théorèmes utiles

On peut reformuler pour les fonctions continues en un point $x_0$, les résultats énoncés lors de l'étude des limites.

**Proposition 2.1.3.** *Si $f$ est continue en $x_0$ alors $f$ est bornée sur un voisinage de $x_0$.*

**Proposition 2.1.4.** *Si $f$ est continue en $x_0$ et $f(x_0) > 0$, alors $f(x) > \frac{1}{2} f(x_0)$ pour tout $x$ dans un certain voisinage de $x_0$, et elle est donc positive sur ce même voisinage.*

**Proposition 2.1.5.** *Si $f$ et $g$ sont continues en $x_0$, il en est de même de $f + g$ et $f.g$ ; et, pourvu que $g(x_0) \neq 0$, $f/g$ est définie au voisinage de $x_0$ et est continue en $x_0$.*

**Proposition 2.1.6.** *Si $f : D \to \mathbb{R}$ est continue en $x_0 \in D$, $y_0 = f(x_0)$, $g : D' \to \mathbb{R}$ est continue en $y_0$ et $D'$ est un voisinage de $y_0$, alors $h = g \circ f$ est bien définie sur un voisinage de $x_0$ et est continue en $x_0$.*

La preuve est la même que pour la composition des limites.

*Démonstration.* La preuve est la même que pour la composition des limites. Néanmoins, on peut en donner une autre, basée sur le DL à l'ordre 0, parce que nous utiliserons la même technique de preuve lors de la proposition 3.1.5 pour prouver l'énoncé concernant la dérivée d'une fonction composée.

La continuité de $g$ en $y_0 = f(x_0)$ implique l'existence d'un DL d'ordre 0 de $g$ en $y_0$, i.e il existe une fonction $\eta$, $\eta(0) = 0$, continue en 0 telle que, pour tout $y$ dans un voisinage $V$ de $y_0$, on a

$$g(y) = g(y_0) + \eta(y - y_0).$$

De même, la continuité de $f$ en $x_0$ implique l'existence d'un DL d'ordre 0 de $f$ en $x_0$, i.e il existe une fonction $\epsilon$, $\epsilon(0) = 0$, continue en 0 telle que, pour tout $x$ dans un voisinage $U$ de $x_0$, on a

$$f(x) = f(x_0) + \epsilon(x - x_0).$$

On peut par ailleurs supposer que si $x \in U,\ f(x) \in V$ quitte à restreindre l'ensemble $U$.

On peut donc substituer $f(x)$ à $y$ dans la première égalité pour obtenir que pour tout $x \in U$,

$$h(x) = g(y_0) + \eta(f(x_0) + \epsilon(x - x_0) - y_0) = h(x_0) + (\eta \circ \epsilon)(x - x_0).$$

La fonction $\eta \circ \epsilon$ est nulle en 0 et est continue en 0 (voir à ce propos l'énoncé de composition de limites), on a donc obtenu un DL de $h$ en $x_0$ à l'ordre 0 qui prouve que $h$ est continue en $x_0$. $\square$

Nous terminons cette partie avec quelques remarques sur la notion de continuité sur un ensemble ou sur un intervalle.

En effet, quand on a une fonction $f : D \to \mathbb{R}$ et un sous-ensemble $B \subset D$, il ne faut pas confondre, au vu des définitions qu'on a données, "$f$ est continue en tout point de $B$" (en considérant $f$ comme une fonction sur $D$ mais en regardant sa continuité sur certains des points de $D$, ceux qui appartiennent à $B$" et "la restriction de $f$ à $B$ est continue". Par exemple, si $f : [0, 3] \to \mathbb{R}$ est définie par $f(x) = 1$ si $x \in [1, 2]$ et $f(x) = 0$ si $x \notin [1, 2]$, en prenant $D = [0, 3]$ et $B = [1, 2]$ on est dans la situation suivante : tout point de $B$ n'est pas un point de continuité ($f$ n'est pas continue en 1, ni en 2) ; pourtant si on restreint $f$ à $B$ on obtient la fonction constante 1, qui est donc continue.

En particulier, on peut remarquer que, quand le sous-ensemble $B$ est un intervalle $[a, b] \subset D$, la continuité de la restriction équivaut à la continuité de $f$ (considérée comme une fonction sur $D$) en tout point de $]a, b[$ accompagnée de la continuité à droite en $a$ et à gauche en $b$. Ce qui est différent de la continuité en tout point de $[a, b]$.

**Proposition 2.1.7** (Recollement)**.** *Soient $I$ et $J$ deux intervalles fermés ayant un seul point commun $a$. Soit $f : I \cup J \to \mathbb{R}$ une fonction dont la restriction à $I$, ainsi que la restriction à $J$, sont continues. Alors $f$ est continue sur l'intervalle $I \cup J$.*

Il suffit de voir que dans cette situation, comme $I$ et $J$ sont non triviaux, $a$ est un point intérieur à l'intervalle $I \cup J$. Par ailleurs, les limites à gauche et à droite de $f$ en $a$ coincident toutes deux avec $f(a)$. En conclusion $f$ admet $f(a)$ pour limite en $a$.

**Proposition 2.1.8.** *Si $I$ est un intervalle de $\mathbb{R}$ non trivial, $f, g$, fonctions à valeurs réelles sont continues sur $I$ alors*

1. *$f + g$ est continue sur $I$,*
2. *$f.g$ est continue sur $I$,*
3. *si $g$ ne s'annule pas sur $I$, $f/g$ est continue sur $I$.*

**Proposition 2.1.9.** *Soient $I$ et $J$ deux intervalles de $\mathbb{R}$. Si $f$, à valeurs réelles, continue sur $I$, $g$, à valeurs réelles, continue sur $J$. Si, pour tout $x \in I$, $f(x) \in J$, la composée $h = g \circ f$ est continue sur $I$.*

## 2.2 Les gros théorèmes sur les fonctions continues

Nous présentons ici deux théorèmes importants sur les fonctions continues qu'il nous est, hélas, interdit de démontrer...

### 2.2.1 Le théorème des valeurs intermédiaires (TVI)

Ce théorème montre l'une des propriétés plus importantes et intuitives des fonctions continues, qui colle bien avec l'idée "une fonction est continue si on peut tracer son graphe sans lever le crayon". En particulier, si on prend la feuille où l'on trace le graphe et on la sépare en deux parties par une droite horizontale (parallèle aux abscisses), pour passer d'une région à l'autre il faudra bien traverser (et donc croiser) la droite elle même.

Nous allons en donner plusieurs versions.

**Théorème 2.2.1** (TVI 1)**.** *Soit $g$ une fonction, à valeurs réelles, continue sur un intervalle $I$ de $\mathbb{R}$, $a, b \in I$. Si $g(a)$ et $g(b)$ sont non nuls et de signes opposés, il existe alors $c$,* **strictement entre**[1] *$a$ et $b$ tel que $g(c) = 0$.*

Ce théorème correspond à l'énoncé qu'on vient de donner en termes de crayons et graphes, si on considère la droite $\{y = 0\}$. Il a des applications très puissantes, si on construit bien la fonction auquel on veut l'appliquer.

**Exercice :** Montrer qu'il existe un nombre $c \in \left]0, \frac{\pi}{2}\right[$ tel que $c = \cos^2 c$.

On peut évidemment donner une version plus générale.

**Théorème 2.2.2** (TVI 2)**.** *Soit $I$ un intervalle de $\mathbb{R}$, $f$ à valeurs réelles, continue sur $I$, $a$, $b \in I$. Si $y$ est entre $f(a)$ et $f(b)$, il existe alors $x$ entre $a$ et $b$ tel que $f(x) = y$.*

*Démonstration.* Soit $y$ un nombre entre $f(a)$ et $f(b)$. Il s'agit de se mettre en position pour appliquer le théorème 2.2.1. Il suffit de considérer la fonction $g$ définie par $g(t) = f(t) - y$ pour $t \in I$ car une solution de l'équation $g(x) = 0$, $x \in I$ est une solution de $f(x) = y$, $x \in I$ et réciproquement.

Soit donc, si $f(a) \leq y$ et $f(b) \geq y$ (l'autre cas se traite de manière similaire),

$$g(t) = f(t) - y$$

cette fonction est continue sur $I$, $g(a) = f(a) - y \leq 0$, $g(b) = f(b) - y \geq 0$ et donc il existe $c$ entre $a$ et $b$ tel que $g(c) = 0$ et donc $f(c) = y$. $\square$

Ainsi qu'une version plus abstraite, qui porte sur la définition d'intervalle (un concept qu'on a toujours utilisé sans jamais bien se poser la question). Donc, tout d'abord, qu'est-ce qu'un intervalle ? on peut dire qu'un intervalle est un ensemble de la forme $[a, b]$, $[a, b[$, $]a, b]$, $]a, b[$, $[a, +\infty[$, $]a, +\infty[$, $]-\infty, b]$, $]-\infty, b[$, $]-\infty, +\infty[\,(= \mathbb{R})$, $\{a\}\,(= [a, a])$ (tout ça pour $a, b \in \mathbb{R}$ ; et éventuellement on peut y mettre aussi $\emptyset$). En gros, un intervalle est quelque chose qui a un début et une fin, ses bornes, et qui comprend tout ce qu'il y a au milieu (peu importe, par contre, s'il inclut les bornes). Surtout il n'a pas de trous. Mais alors on pourrait donner une définition qui se base sur ça, plutôt que faire la liste de tous les types d'intervalles possibles.

**Définition 2.2.1.** Un intervalle est une partie $A$ de $\mathbb{R}$ possédant la propriété suivante,

« Pour toute paire de points $y_1$ et $y_2$ dans $A$, si $y$ est un réel entre $y_1$ et $y_2$ alors $y \in A$ »

alors $A$ est un intervalle de $\mathbb{R}$.

1. $c$ est entre $a$ et $b$ si $a \leq c \leq b$ ou $a \geq c \geq b$, il est strictement entre $a$ et $b$ si $a < c < b$ ou $a > c > b$

Il est possible de vérifier que cette définition définit exactement les types d'intervalles que l'on connait (la difficulté étant de démontrer qu'un ensemble $A$ avec cette propriété doit forcément avoir une de ces formes, et en particulier de trouver - ou de démontrer qu'elles existent - ses bornes).

Maintenant on peut dire

**Théorème 2.2.3** (TVI 3)**.** *Soit $I$ un intervalle quelconque de $\mathbb{R}$, $f$ une fonction à valeurs réelles, continue sur $I$. L'ensemble $f(I) = \{f(x), x \in I\}$ est un intervalle de $\mathbb{R}$.*

*Démonstration.* Ce théorème est simple à montrer si on utilise la caractérisation qu'on vient de donner des intervalles, ainsi que la version TVI 2. Soit $A = f(I) = \{f(x), x \in I\}$. Soient $y_1$ et $y_2$ dans $A$ et $y$ un réel quelconque entre $y_1$ et $y_2$. On peut alors trouver deux points $a$ et $b$ dans $I$ tels que $f(a) = y_1$ et $f(b) = y_2$. Le théorème 2.2.2, appliqué à l'intervalle fermé $J$ d'extrémités $a$ et $b$, affirme que, puisque $f$ est continue sur $J$ et que $y$ est entre $f(a)$ et $f(b)$, il existe $x$ entre $a$ et $b$ tel que $f(x) = y$. Ce nombre $x$ est dans $I$ car $a$ et $b$ y sont et donc cela montre que $y \in f(I) = A$.

On a donc montré que $f(I)$ est un intervalle.

□

*Exemples et remarques* 2.2.1.
1. Cela signifie que lorsque $x$ décrit $I$, $f(x)$ décrit exactement un intervalle....
2. Aucune précision n'est apportée quant au « type » de l'intervalle image. Sauf dans le cas très particulier où l'intervalle $I$ est de la forme $[a, b]$, le type de l'intervalle image est indépendant du type de l'intervalle de départ. Par exemple, si $f : I = ]-1, 1[ \to \mathbb{R}$ est définie par $f(x) = x^2$, on a $f(I) = [0, 1[$.

### 2.2.2 Maxima et minima

On veut arriver ici à un théorème qui parle de l'existence des minima (ou de maxima) d'une fonction $f$ sur son domaine $D$, sous des hypothèses sur $f$ et $D$. Ceci est très important parce que les problèmes de recherche des minima et de maxima sont des problèmes d'optimisation qui ont plein d'applications (par exemple minimisation des coûts, de la dispersion d'énergie... maximisation de la rentabilité...). Et, même si cela a un parfum beaucoup trop théorique, savoir si cet optimum va exister ou pas est crucial.

Tout d'abord il faut clarifier ce que le minimum d'un ensemble. Prenons une part $A$ de $\mathbb{R}$. Ici le fait que l'on parle d'un ensemble de nombre réels (et non de nombres complexes, de points dans l'espace, d'animaux...) est important puisqu'on a besoin de les comparer avec un ordre. Un nombre $m$ est dit *minimum* de $A$ si $m \in A$ et pour tout $x \in A$ on a $x \geq m$. Pareillement, un nombre $M$ est dit *maximum* de $A$ si $M \in A$ et pour tout $x \in A$ on a $x \leq M$. Le minimum, s'il existe, est évidemment unique (et le maximum aussi) : s'il y en avait deux, $m$ et $m'$, on devrait avoir $m \leq m'$ et $m' \leq m$, c'est-à-dire $m = m'$.

Le minimum n'existe pas toujours (et le maximum non plus). Il existe si $A$ est fini, parce qu'il suffit d'ordonner les éléments de $A$ et voir qui tombe en première place. Mais si $A$ est l'ensemble des réciproques des nombres entiers positifs, $A = \{1/n,\ n \in \mathbb{N}, n \geq 1\}$, il n'y a pas de minimum (ça serait comme si on cherchait le nombre entier positif le plus grand). Quand on a une infinité de valeurs il se peut qu'aucune ne soit plus petite que toute les autres. Ici les valeurs se concentrent vers la valeur 0, qui satisfait bien "pour tout $x \in A$ on a $x \geq 0$", mais $0 \notin A$.

Et qu'est-ce que le minimum d'une fonction ? une fonction n'est pas vraiment un ensemble...mais on peut considérer l'ensemble de ses valeurs. Quand on a $f : D \to \mathbb{R}$ on cherche le minimum de $f(D) = \{f(x) : x \in D\}$. Là aussi, dès que $D$ est infini (ce qui est tout le temps le cas quand on a à faire à des intervalles non triviaux, par exemple) on risque de pas avoir un minimum.

Mais on va voir ici que, si $f$ est continue et $D$ satisfait deux hypothèses, alors un théorème dû à Weierstrass nous garantit l'existence du minimum et du maximum. Attention : quand on parle du minimum d'une fonction on parle encore de sa valeur minimale, et pas du point où elle est réalisée. Si la valeur minimale est sûrement unique, les points où elle est réalisée pourraient être plusieurs. C'est le cas par exemple de la fonction $f(x) = x^2(1-x)^2$ qui est minimisée en 0 et 1, ainsi que de toute fonction constante. . .

Revenons donc aux hypothèses dont on a besoin. La première est que $f$ soit continue sur $D$. La deuxième que $D$ soit fermé (on a déjà vu ce qu'il siginife). La troisième est plus simple à définir (par rapport à la notion d'ensemble fermé).

**Définition 2.2.2.** Soit $D$ un sous-ensemble de $\mathbb{R}$. On dit que $D$ est *borné* si il est inclus dans un intervalle $[-R, R]$. Dutrement dit, $D$ est borné s'il existe $R$ tel que pour tout $x \in D$ on a $|x| \leq R$.

Évidemment ce n'est pas la même chose de dire "il existe $R$ tel que pour tout $x \in D$ on ait $|x| \leq R$" et "pour tout $x \in D$ il existe $R$ tel que l'on a $|x| \leq R$. Dans ce deuxième cas la quantité $R$ peut dépendre de $x$ et donc cette propriété est toujours vérifiée, car pour tout $x$ on peut choisir par exemple $R = 1 + |x|$ et réaliser l'inégalité.

On donnera aussi cette définition.

**Définition 2.2.3.** Un sous-ensemble de $\mathbb{R}$ est dit compact s'il est fermé et borné.

En particulier, tout intervalle de la forme $[a, b]$ est compact (il s'agit du type principal d'ensemble compact, mais évidemment on peut vérifier que la réunion de deux intervalles comme ça est aussi compacte etc.).

On peut maintenant énoncer le bien connu théorème d'existence des minima et maxima.

**Théorème 2.2.4** (Weierstrass)**.** *Soit $D$ un sous-ensemble compacte de $\mathbb{R}$ et $f : D \to \mathbb{R}$ une fonction continue. Il existe alors un point $x_0 \in D$ tel que $f(x_0) = \min\{f(x) : x \in D\}$ (et, symétriquement, il existe un point $x^0 \in D$ tel que $f(x^0) = \max\{f(x) : x \in D\}$).*

*Observation* 2.2.2. Une conséquence de ce résultat, si on l'utilise avec le TVI3, est que l'image par une fonction continue d'un intervalle fermé borné $[a, b]$ est un intervalle fermé borné $[f(x_0), f(x^0)]$. Que l'image d'un intervalle soit un intervalle est une conséquence des théorème des valeurs intermédiaires (théorème des zéros), alors que le fait qu'il soit borné et que ses bornes appartiennent à l'image sont des conséquences de l'existence du minimum et du maximum que l'on vient de montrer.

Il est utile de voir quelques exemples de non-existence du minimum ou maximum pour comprendre l'importance des hypothèses.

*Exemples et remarques* 2.2.3.

1. Soit $D = [-1, 1]$ et $f$ la fonction donnée par

$$f(x) = \begin{cases} |x| & \text{si } x \neq 0 \\ 2 & \text{si } x = 0. \end{cases}$$

   Cette fonction n'admet pas de minimum sur l'ensemble $D$, parce que $f(x) > 0$ pour tout $x \in D$ mais les valeurs prises sont aussi proches de 0 qu'on veut. On voudrait dire alors que le minimum est 0, mais cette valeur n'est prise nulle part. En effet, il s'agit d'une fonction qui est continue en tout point de $D \setminus \{0\}$ mais pas en 0. Et c'est justement au point 0 que cette fonction "voudrait" prendre la valeur nulle. . .Par contre la même fonction admet un maximum, et ce maximum est 2, car on peut bien voir $2 = f(0) \geq f(x)$ pour tout $x$.

2. Soit $D = ]0, 1]$ et $f$ la fonction donnée par $f(x) = 1 - x$. Cette fonction n'admet pas de maximum sur l'ensemble $D$, parce qu'elle prend des valeurs très proches de 1 mais,

évidemment, $f(x) < 1$ pour tout $x \in D$. En effet, ici le problème n'est pas posé par $f$ (qui est bien continue), mais par $D$, qui est borné mais pas fermé. Il est intéressant de remarquer que, en remplaçant $1 - x$ par $1/x$, on aurait même pu construire une fonction qui non seulement n'admet pas de maximum, mais elle n'est pas non plus bornée supérieurement. Ceci n'aurait pas évidemment été possible si l'on avait voulu utiliser une fonction continue sur un fermé borné.

3. Si $D = ]0, 1]$ et $f(x) = (1 - x)\sin\frac{1}{x}$, $f$ n'a ni minimum ni maximum sur $D$. Lorsque $x$ s'approche de 0, $f(x)$ s'approche d'aussi près que l'on veut de $\pm 1$ sans jamais atteindre ces valeurs.

Nous terminons cette section en montrant que l'existence du minimum (ou du maximum) reste vraie même en étant un peu plus souple sur la nature de l'ensemble $D$, à condition de demander à la fonction $f$ de l'aider un peu. On fera l'exemple du cas $D = \mathbb{R}$.

**Théorème 2.2.5.** *Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue. Supposons que $f$ tende vers $+\infty$ des deux côtés, c'est-à-dire*

$$\lim_{x \to +\infty} f(x) = \lim_{x \to -\infty} f(x) = +\infty.$$

*Il existe alors un point $x_0 \in \mathbb{R}$ tel que $f(x_0) = \min\{f(x) : x \in \mathbb{R}\}$.*

*Démonstration.* Prenons une valeur quelconque réalisée par la fonction $f$, par exemple $f(0)$. Il est évident que le minimum $m$, si il existe, va vérifier la condition $m \leq f(0)$. Tous les points $x$ avec $f(x) > f(x_0)$ n'influencent pas l'existence et la nature du minimum.

Par définition de limite infinie, on sait que pour tout $K \in \mathbb{R}$ il existe un nombre (réel) $b$ tel que $x > b$ entraine $f(x) > K$ (ceci car $\lim_{x \to +\infty} f(x) = +\infty$) ainsi qu'un nombre $a$ tel que $x < a$ entraine $f(x) > K$ (car $\lim_{x \to -\infty} f(x) = +\infty$).

Prenons $K = f(0)$. Grâce à ce qu'on a dit avant, on peut donc se restreindre à l'intervalle $[a, b]$ (qui, lui, est fermé et borné), parce que tous les points à l'extérieur de cet intervalle ne sont pas pris en compte dans la recherche d'un minimum (parce qu'ils satisfont $f(x) > f(0)$). Et sur cet intervalle là il existe, grâce au théorème précedant, un point $x_0 \in [a, b] \subset \mathbb{R}$ tel que $f(x_0) \leq f(x)$ pour tout $x \in X$. Comme en plus on a $f(x) \geq K = f(0) \geq f(x_0)$ pour tout $x \notin [a, b]$, on a finalement $f(x) \geq f(x_0)$ pour tout $x \in \mathbb{R}$, ce qui montre que $x_0$ est un point de minimum. $\square$

Il est évident qu'on pourrait énoncer un théorème similaire pour l'existence du maximum, en mettant comme hypothèse que $\lim_{x \to +\infty} f(x) = \lim_{x \to -\infty} f(x) = -\infty$.

## 2.3 Les fonctions réciproques et leur continuité

On peut se demander comment les notions de fonction bijective et de fonction réciproque interagissent avec la notion de continuité. Tout d'abord nous rappelons de qui l'on parle

### 2.3.1 Bijections et fonctions réciproques

**Définition 2.3.1.** Soit $f : D \to \mathbb{R}$ une fonction réelle de variable réelle, $A \subset D$ et $B \subset \mathbb{R}$. On dit que $f$ établit une bijection de $A$ dans (sur) $B$ si pour tout $y \in B$, il existe un unique $x \in A$ tel que $y = f(x)$.

Autrement dit, $f$ établit une bijection de $A$ sur $B$ si pour tout $y$ dans $B$, l'équation $f(x) = y$ d'inconnue $x$ dans $A$ admet une unique solution.

On dit de façon plus classique que $f : A \to B$ **est une bijection**.

*Exemples et remarques* 2.3.1. Soit $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$. Pour que l'équation $y = f(x)$ d'inconnue $x \in \mathbb{R}$ admette **au moins** une solution, il faut que $y \geq 0$. Si $y > 0$, il est bien connu que l'équation en question admet deux solutions distinctes (non nulles et opposées). Pour que $f$ établisse une bijection d'une partie $A$ de $\mathbb{R}$ sur $\mathbb{R}^+$, il nous faut une unique solution. On choisit en général la solution positive (que l'on note $\sqrt{y}$) et, dans notre langage, cela se traduit par le fait que $f$ établit une bijection de $\mathbb{R}^+$ sur $\mathbb{R}^+$.

On aurait pu choisir de prendre systématiquement la solution négative, on a aussi que $f$ établit une bijection de $\mathbb{R}^-$ sur $\mathbb{R}^+$.

Si $f$ établit une bijection de $A$ sur $B$, alors, a tout $y \in B$, on peut faire correspondre $g(y) \in A$, l'unique solution de l'équation $y = f(x)$ d'inconnue $x \in A$. Cela définit une fonction $g : B \to A$ appelée **fonction réciproque** de $f$ vue comme application de $A \to B$.

On note cette application $g = f^{-1}$ pour marquer son lien avec la fonction $f$. Remarquons que cette fonction $f^{-1}$ dépend crucialement des ensembles $A$ et $B$ entre lesquels $f$ établit une bijection.

*Exemples et remarques* 2.3.2.

1. Il ne faut surtout pas confondre $f^{-1}$ avec la fonction $\frac{1}{f}$ ! Le vocable **fonction inverse** de $f$ est parfois utilisé. Suivant le contexte, cela peut désigner $f^{-1}$, la bijection réciproque de $f$ ou $\frac{1}{f}$, la composée de l'inversion, $x \mapsto \frac{1}{x}$ et de $f$.
2. $f^{-1}(y)$ est l'unique $x$ de $A$ vérifiant $f(x) = y$. On a donc
   (a) $f(f^{-1}(y)) = y$ pour tout $y$ dans $B$ ($x = f^{-1}(y)$ vérifie que $f(x) = y$)
   (b) $f^{-1}(f(x)) = x$ pour tout $x$ dans $A$ (Si $z = f^{-1}(f(x))$, on a $f(z) = f(x)$ et $f(z) \in B$. Comme, il y a **unicité**, on en tire que $z = x$.)
3. Dans ce contexte, la fonction $f^{-1} : B \to A$ est une bijection de $B$ sur $A$ et sa fonction réciproque est $(f^{-1})^{-1} = f$ avec $f : A \to B$.

Le graphe $G_g$ d'une fonction $f$ de variable réelle à valeurs réelles est le sous-ensemble du plan formé par les points $(x, f(x))$ lorsque $x$ décrit le domaine de définition de $f$.

Si $f : A \to B$ est une bijection, alors un couple $(x, y)$ est dans $G_g$ si et seulement si $x \in A$, $y \in B$, $y = f(x)$.

Cela est clairement équivalent au fait que le couple $(y, x)$ est dans le graphe de $f^{-1}$.

Supposons que le plan est rapporté à un repère orthonormé alors les graphes de $f$ et $f^{-1}$, plus précisément, les courbes d'équation $y = f(x), x \in A$ et $y = f^{-1}(x), x \in B$, sont symétriques l'un de l'autre dans la symétrie orthogonale d'axe la première bissectrice du plan, i.e la droite d'équation $y = x$.

Remarquons par ailleurs que les courbes d'équation $y = f(x), x \in A, y \in B$ et $x = f^{-1}(y), x \in A, y \in B$ sont égales.

### 2.3.2 Injectivité, monotonie et continuité

La notion de fonction bijective est en effet la réunion de deux concepts, celui de fonction injective (une même valeur n'est jamais prise plus qu'une fois) et celui de fonction surjective (toute valeur de l'espace d'arrivée est prise au moins une fois). Ce deuxième concept dépend fortement de ce que l'on choisit comme espace d'arrivée (en prenant un ensemble plus grands la surjectivité disparaîtra). Nous allons ici nous occuper un instant de l'injectivité, qui est plus intrinsèque.

Quand on parle des fonctions réelles d'une variable réelle il y a une manière simple de garantir l'injectivité : toute fonction strictement monotone est forcément injective, parce qu'elle ne peut pas revenir sur la même valeur une deuxième fois (à cause de la monotonie), ni y rester pendant plus qu'un instant seul (à cause de la *stricte* monotonie). Pourtant toute fonction injective n'est pas monotone, comme on peut le voir avec l'exemple suivant.

*Exemple* 2.3.3. Considérons la fonction

$$f : [0,2[ \to [0,2[ \quad \text{donnée par } f(x) = \begin{cases} x & \text{si } x \leq 1; \\ 3-x & \text{si } x > 1. \end{cases}$$

Cette fonction est bien injective (elle prend les valeurs de $[0,1]$ sur $[0,1]$ et celles de $]1,2[$ sur $]1,2[$, et sur chacun de ces intervalles elle est strictement monotone), même bijective, mais elle n'est pas du tout monotone !

Cependant, la situation est plus claire quand on a la continuité, puisqu'on a la proposition suivante.

**Proposition 2.3.1.** *Soit $f : I \to \mathbb{R}$ une fonction continue définie sur un intervalle $I$. Alors $f$ est injective si et seulement si elle est strictement monotone.*

*Démonstration.* C'est évident que, si $f$ est strictement monotone, alors elle est injective. Il faut prouver la réciproque. Supposons par l'absurde que $f$ n'est pas strictement monotone. Cela signifie que, soit elle n'est pas monotone, soit elle l'est mais pas strictement. Ce deuxième cas est évidemment impossible à cause de l'injectivité, parce qu'une fonction monotone non strictement est une fonction qui prend la même valeur en deux points distincts. Alors $f$ n'est pas monotone, c qui signifie qu'il y a trois points $x_1 < x_2 < x_3 \in I$ tels que le sens de variations de $f$ entre $y_1 = f(x_1)$ et $y_2 = f(x_2)$ est différent de celui entre $y_2 = f(x_2)$ et $y_3 = f(x_3)$. Supposons par exemple (l'autre cas est analogue) que $y_1 > y_2$ mais $y_2 < y_3$. De plus on sait (par l'injectivité) que $y_1 \neq y_3$. Supposons par praticité que $y_1 < y_3$ (là aussi, l'autre cas est analogue). La valeur $y_1$ se trouve alors strictement entre $y_2$ et $y_3$ et, en appliquant le TVI 2, on trouve qu'il y a un point $\bar{x}$ strictement entre $x_2$ et $x_3$ tel que $f(\bar{x}) = y_1 = f(x_1)$. Ceci contredit l'injectivité et la thèse est bien prouvée. □

Le fait de savoir que, sur les intervalles, toutes les fonctions injectives sont monotones nous permet de déterminer facilement l'intervalle image $f(I)$, étant donnée une fonction $f : I \to \mathbb{R}$ injective et continue (et donc monotone). Ceci est indispensable pour trouver l'intervalle $J$ tel que $f : I \to J$ est une bijection (et donc pour pouvoir définir $f^{-1} : J \to I$).

Le critère est simple : soit $I$ un intervalle de $\mathbb{R}$, $f : I \to \mathbb{R}$ une fonction continue strictement monotone. Si $a$ et $b$ sont les bornes de $I$ ($a$ et $b$ pouvant être des nombres réels ou les symboles $\pm\infty$) alors $J = f(I)$ est un intervalle dont les bornes sont $\lim_{x\to a} f(x)$ et $\lim_{x\to b} f(x)$. La nature de l'intervalle $J$ en la borne $\lim_{x\to a} f(x)$ (fermé ou ouvert) est la même que celle de l'intervalle $I$ en $a$.

*Exemples et remarques* 2.3.4.
1. Soit $f : ]0,1] \to \mathbb{R}$ définie par $f(x) = \frac{1}{x}$. La fonction $f$ est continue, strictement décroissante sur $I = ]0,1]$. On a $\lim_{x\to 0} f(x) = +\infty$ et $\lim_{x\to 1} f(x) = 1$. On en déduit que $J = f(I) = [1,+\infty[$.
2. La fonction tangente est continue, strictement croissante sur l'intervalle $I = \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$. On a $\lim_{x\to\pm\frac{\pi}{2}} \tan x = \pm\infty$, on a donc $\tan(I) = \mathbb{R}$.

Mais venons maintenant à un théorème assurant la continuité.

**Théorème 2.3.2.** *Soit $f$ une fonction continue, strictement monotone sur un intervalle $I$ alors*
*(i) $J = f(I)$ est intervalle de $\mathbb{R}$*
*(ii) $f$ établit une bijection de $I$ sur $J$*
*(iii) La fonction réciproque $f^{-1} : J \to I$ est continue, strictement monotone sur $J$, de même sens de variation que $f$.*

*Démonstration.* (i) C'est une des versions que nous avons donnée du théorème des valeurs intermédiaires, le théorème 2.2.3.

(ii) Soit $y \in J$ fixé. Par définition même de l'ensemble $J = f(I)$, l'équation $y = f(x)$ d'inconnue $x \in I$ admet une solution. Rappelons que $f(I)$ est l'ensemble de **toutes** les images possibles des éléments de $I$ par la fonction $f$.
Cette équation ne peut admettre qu'une solution car $f$ est **strictement** monotone. En effet, si l'on dispose de deux solutions distinctes $x_1 < x_2$, on a alors $f(x_1) = y = f(x_2)$, ce qui est impossible car $f$ étant strictement monotone, on a soit $f(x_1) < f(x_2)$, soit $f(x_1) > f(x_2)$.

(iii) Supposons que $f$ est strictement croissante pour fixer les idées. Soient $y_1 < y_2 \in J$. De deux choses l'une, soit $f^{-1}(y_1) \geq f^{-1}(y_2)$, soit $f^{-1}(y_1) < f^{-1}(y_2)$. Si le premier cas a lieu, comme $f$ est croissante, on alors que
$$f(f^{-1}(y_1)) \geq f(f^{-1}(y_2))$$
et donc $y_1 \geq y_2$, ce qui est contraire à l'hypothèse sur $y_1$ et $y_2$, c'est donc le deuxième cas qui est vérifié.
Donc, pour $y_1 < y_2$, $f^{-1}(y_1) < f^{-1}(y_2)$, ce qui montre que $f$ est strictement croissante.

(iii bis) En ce qui concerne la continuité de $f^{-1}$. Il suffit de montrer que $f$est continue en tout point $y_0$ de $J$. Il existe $x_0 \in I$ tel que $f(x_0) = y_0$, ce qui signifie $x_0 = f^{-1}(y_0)$. Supposons par simplicité que $y_0$ est à l'intérieur de $J$, ce qui implique, par monotonie stricte, que $x_0$ est à l'intérieur de $I$. Prenons maintenant un voisinage $V = ]x_0 - \delta, x_0 + \delta[$ de $x_0$, avec $\delta > 0$ suffisamment petit pour que $]x_0 - \delta, x_0 + \delta[ \subset I$. Prenons les points $y' := f(x_0 - \delta)$ et $y'' := f(x_0 + \delta)$. On a $y' < y_0 < y''$ et l'intervalle $U := ]y', y''[$ est bien un voisinage de $y_0$. Or, par monotonie, de $f^{-1}(y') = x_0 - \delta$ et $f^{-1}(y'') = x_0 + \delta$ on tire facilement que pour tout $y \in ]y', y''[$ on a $f^{-1}(y') \in ]x_0 - \delta, x_0 + \delta[= V$. Ceci montre la continuité de $f^{-1}$ en $y_0$. Si $y_0$ est une borne de l'intervalle $J$ on prendra alors l'intervalle $[x_0, x_0 + \delta[$ (ou $]x_0 - \delta, x_0]$) et l'intervalle image $[y', y''[$, avec $y' = y_0$ (ou $]y', y'']$, avec $y'' = y_0$) sera bien l'intersection d'un voisinage de $y_0$ avec $J$. Dans tous les cas on a démontré la continuité de $f^{-1}$ en $y_0$. $\square$

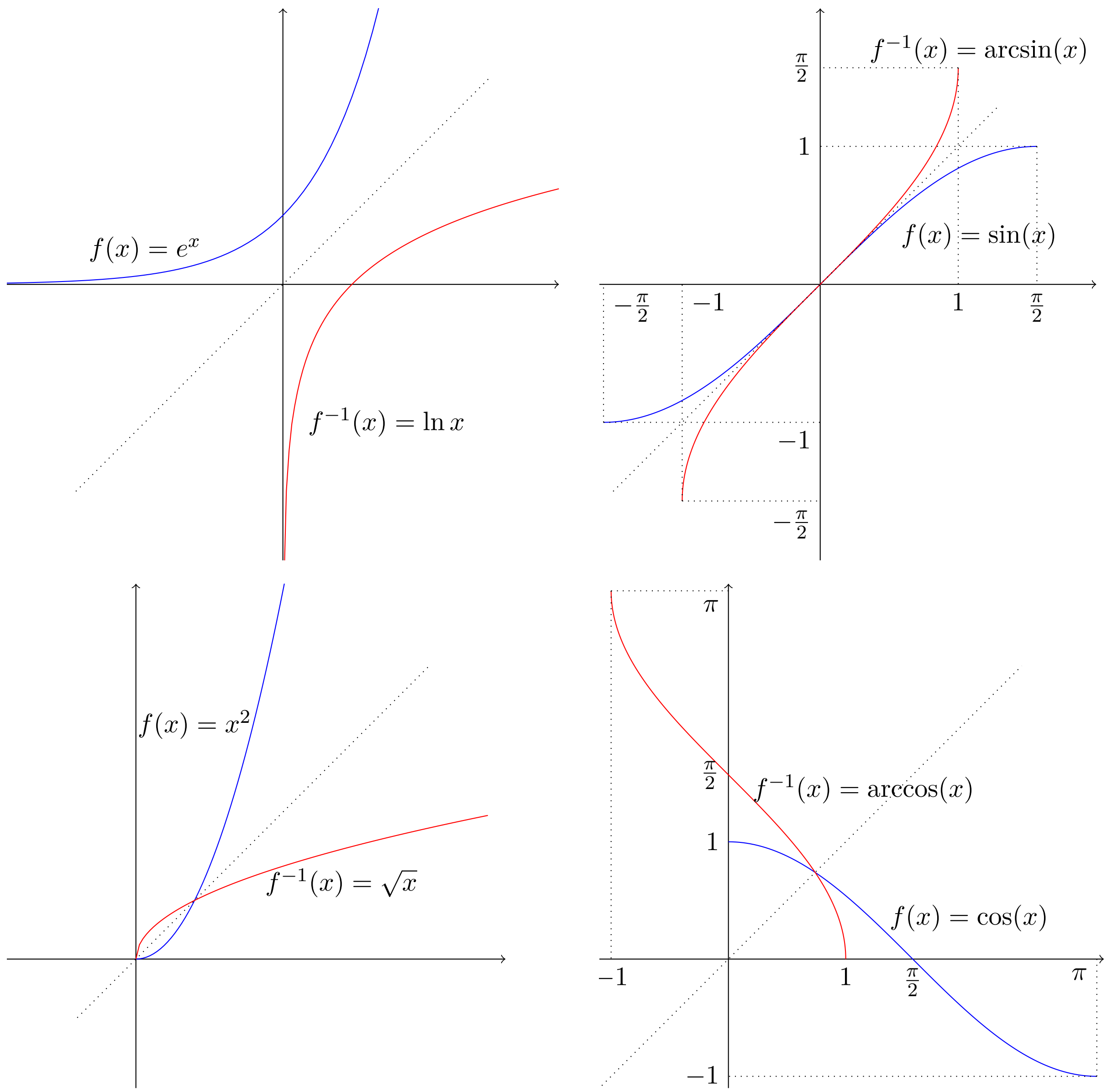
### 2.3.3 Fonction réciproques fondamentales

Nous terminons ce chapitre avec une liste de fonctions bijective dont la réciproque est importante et à connaître.

- **Les racines $n-$ièmes.** Est-ce que la fonction $x \mapsto x^n$ est bijective ? tout d'abord, est-elle injective ? il faut distinguer deux cas : si $n$ est paire, la réponse est non, puisque la fonction est paire et $(-x)^n = x^n$ ; si $n$ est impaire il s'agit d'une fonction strictement croissante et donc injective.
  Si on reste sur le cas $n$ impaire, puisque $\lim_{x\to\pm\infty} x^n = \pm\infty$, la fonction $x \mapsto x^n$ est aussi surjective sur $\mathbb{R}$ et elle est donc une bijection de $\mathbb{R}$ dans $\mathbb{R}$. Sa réciproque est la fonction *racine $n-$ième* $y \mapsto \sqrt[n]{y}$.
  Dans le cas $n$ paire il faut restreindre la fonction $f(x) = x^n$ à un intervalle où elle est injective. Par exemple à la demi-droite des nombres positifs (la demi-droite négatve conviendrait aussi, et d'autres intervalles inclus dans ces demi-droites, mais nous cherchons, si possible à choisir l'intervalle le plus grand sur lequel elle soit injective). Si $I = [0, +\infty]$, qui est $f(I)$ ? là aussi on considère les limites et on trouve $J = f(I) = [0, +\infty]$. $f$ est alors une bijection de $[0, +\infty]$ dans $[0, +\infty]$ et sa réciproque est aussi appelé *racine $n-$ième*. Seule remarque : par convention, on a choisi de prendre la racine positive. C'est ce qu'on pense quand on dit $\sqrt{4} = 2$ (et pas $-2$).
- **Le logarithme.** La fonction exponentielle $f(x) = e^x$, $f : \mathbb{R} \to ]0, +\infty[$ est strictement croissante et est une bijection de $\mathbb{R}$ dans $]0, +\infty[$. Sa réciproque s'appelle logarithme naturel, est indiquée ln, et est une fonction strictement croissante définie sur $]0, +\infty[$ à valeurs dans tout $\mathbb{R}$.
- **Arcsin et arccos.** Les fonctions sinus et cosinus sont-elles injectives ? pas du tout, puisqu'elles sont périodiques et elles reprennent les mêmes valeurs une infinité de fois. Mais on peut toujours les restreindre à des intervalles où elles sont monotones. On remarque alors que $\sin : [-\pi/2, \pi/2] \to [-1, 1]$ est strictement croissante et que $\cos : [0, \pi] \to [-1, 1]$ est stric-

tement décroissante. Les applications réciproques, définies sur ces intervalles, sont appelées respectivement arcsinus, $\arcsin : [-1,1] \to [-\pi/2, \pi/2]$, et arccosinus, $\arccos : [-1,1] \to [0,\pi]$.

- **Arctangente.** La fonction tangente, définie par $\tan x = \sin x / \cos x$, est définie là où le cosinus ne s'annule pas, c'est-à-dire sur la réunion d'une infinité d'intervalles bornés par les points du type $\pi/2 + k\pi$. Elle est périodique de période $\pi$ donc on peut la considérer sur l'un de ces intervalles sans perdre d'information. Sur $]-\pi/2, \pi/2[$ elle est strictement croissant et on a déjà vu que son image est tout $\mathbb{R}$. Elle admet alors une application réciproque appelée arctangente, définie sur $\mathbb{R}$ et à valeurs dans $]-\pi/2, \pi/2[$. Cette fonction $\arctan : \mathbb{R} \to ]-\pi/2, \pi/2[$ est souvente utile en tant que fonction strictement croissante et bornée.



# Chapitre 3

# Dérivées et fonctions dérivables

## 3.1 Dérivée en un point et interprétation géométrique

Dans la suite $f$ est toujours une fonction définie au voisinage de $x_0$ (ce qui signifie qu'elle est définie sur un intervalle qui est un voisinage de $x_0$).

**Définition 3.1.1.** Soit $x_0$ un réel, $f$ une fonction définie au voisinage de $x_0$.

1. La fonction $\theta$ définie par $\theta(x) = \frac{f(x)-f(x_0)}{x-x_0}$ est bien définie au voisinage de $x_0$ sauf en $x_0$. $\theta(x)$ est le **taux d'accroissement** de $f$ entre $x$ et $x_0$.
2. On dit que $f$ est dérivable en $x_0$ lorsque le taux d'accroissement $\theta(x)$ de $f$ entre $x$ et $x_0$ admet une limite lorsque $x \to x_0$. Cette limite s'appelle la dérivée (ou le nombre dérivé) de $f$ en $x_0$ et est noté $f'(x_0)$. On a donc, lorsque cela à un sens,
$$f'(x_0) := \lim_{x \to x_0} \frac{f(x) - f(x_0}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

*Exemples et remarques* 3.1.1.
1. $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$ admet un nombre dérivé en tout $x_0 \in \mathbb{R}$. On a $f'(x_0) = 2x_0$.
2. $g : \mathbb{R}^+ \to \mathbb{R}$ définie par $g(x) = \sqrt{x}$ admet un nombre dérivé en $x_0$ pour tout $x_0 > 0$. On a
$$g'(x_0) = \frac{1}{2\sqrt{x_0}}$$
3. $h : \mathbb{R} \to \mathbb{R}$ définie par $h(x) = x^2 \sin \frac{1}{x}$ si $x \neq 0$ et $h(0) = 0$ est dérivable en $x_0 = 0$ avec $h'(0) = 0$ (Elle est d'ailleurs dérivable en tout $x_0 \in \mathbb{R}$)
4. C'est par praticité qu'on se restreint, ici dans le chapitre sur la dérivabilité, aux fonctions définies sur des intervalles et, sauf quand on parlera de dérivabilité à gauche et à droite, au voisinage du point où l'on calcule la dérivée. Ceci parce que beaucoup de théorèmes portant sur les dérivées demandent à être sur un intervalle et parce que "on a besoin d'espace autour du point" pour calculer une dérivée.

### Interprétation géométrique

En considérant le graphe de $f$ au voisinage de $x_0$, pour $x \neq x_0$ la **corde** au graphe passant par les points $M_0$ et $M$ de coordonnées $(x_0, f(x_0))$ et $(x, f(x))$ a pour équation (d'inconnues $X, Y \in \mathbb{R}$)
$$Y - f(x_0) = \theta(x).(X - x_0)$$
Lorsque $x$ tend vers $x_0$, $\theta(x)$ tend vers $f'(x_0)$, ce qui signifie géométriquement que la corde se « rapproche » de la droite d'équation
$$Y - f(x_0) = f'(x_0).(X - x_0)$$

Cette droite est appelée la tangente au graphe de $f$ en $x_0$.

Si le taux d'accroissement d'une fonction $f$ en un point $x_0$ admet une limite infinie lorsque $x$ tend vers $x_0$. On dit que le graphe de $f$ admet une tangente verticale. Attention ! Dans un tel cas, la fonction $f$ n'est pas dérivable en $x_0$

Un exemple typique est le cas fonction $x \mapsto x^{\frac{1}{3}}$ en 0. Nous invitons le lecteur à tracer l'allure de son graphe au voisinage de 0.

### DL d'ordre 1

De la définition même de limite, on déduit la

**Proposition 3.1.1.** *Soit $x_0$ un réel, $f$ une fonction définie au voisinage de $x_0$. $f$ est dérivable en $x_0$ si et seulement si il existe un nombre $\ell$ et une fonction $\epsilon$, définie au voisinage de 0 tels que*
*(i) $\epsilon$ est continue en 0, $\epsilon(0) = 0$*
*(ii) pour tout $x$ voisin de $x_0$,*

$$f(x) = f(x_0) + \ell.(x - x_0) + (x - x_0)\epsilon(x - x_0)$$

*Dans ce cas, $\ell = f'(x_0)$.*

L'écriture du point (ii) s'appelle le **développement limité d'ordre 1** de $f$ au voisinage de $x_0$.

### Dérivabilité et continuité

**Corollaire 3.1.2.** *Si $f$ est dérivable en $x_0$, elle y est continue.*

*Exemples et remarques* 3.1.2.
1. La réciproque est fausse comme le montre l'exemple de la fonction valeur absolue en 0.
2. On peut construire une fonction, continue sur $\mathbb{R}$, qui n'est dérivable en aucun point

### Dérivabilité à droite et à gauche

**Définition 3.1.2.** Soit $x_0 \in \mathbb{R}$, $f$ une fonction définie dans un voisinage à droite de $x_0$. On dit que $f$ est dérivable à droite en $x_0$, de nombre dérivé $f'_d(x_0)$ si $\frac{f(x)-f(x_0)}{x-x_0}$ admet une limite lorsque $x$ tend vers $x_0$ à droite. Dans ce cas

$$f'_d(x_0) = \lim_{\substack{x \to x_0 \\ x > x_0}} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{\substack{h \to 0 \\ h > 0}} \frac{f(x_0 + h) - f(x_0)}{h}$$

On a une définition similaire pour la dérivabilité à gauche et $f'_g(x_0)$, le nombre dérivé à gauche.

**Proposition 3.1.3.** *Soit $f$ définie au voisinage du point $x_0$. $f$ est dérivable en $x_0$ si et seulement si elle y est dérivable à gauche et à droite et*

$$f'_d(x_0) = f'_g(x_0)$$

**Exercice :** étude de la dérivabilité en $-1$ à droite et en 0 de la fonction $f(x) = \sqrt{x^2 + x^3}$, définie sur $[-1, +\infty[$.

#### Opérations algébriques

**Proposition 3.1.4.** *Soit $x_0$ un réel, $f$ et $g$ deux fonctions définies au voisinage de $x_0$, dérivables en $x_0$.*

1. *$f+g$ est dérivable en $x_0$ et*
$$(f+g)'(x_0) = f'(x_0) + g'(x_0)$$
2. *(Règle de Leibniz) $f.g$ est dérivable en $x_0$ et*
$$(f.g)'(x_0) = f'(x_0).g(x_0) + f(x_0).g'(x_0).$$
*En particulier, ceci s'applique si $f$ est une fonction constante, $f(x) = C$, pour donner que $C.g$ est dérivable en $x_0$ avec*
$$(C.g)'(x_0) = C.g'(x_0)$$
3. *Si $g(x_0) \neq 0$, les fonctions $\frac{1}{g}$ et $\frac{f}{g}$ sont définies au voisinage de $x_0$, sont dérivables en $x_0$ avec*
$$\left(\frac{1}{g}\right)'(x_0) = -\frac{g'(x_0)}{g(x_0)^2} \ \text{et} \ \left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0).g(x_0) - f(x_0).g'(x_0)}{g(x_0)^2}$$

#### Composition et fonctions réciproques

**Proposition 3.1.5** (Règle de la chaîne)**.** *Soit $x_0 \in \mathbb{R}$, $f$ une fonction définie au voisinage de $x_0$, dérivable en $x_0$. Soit $y_0 = f(x_0)$, $g$ une fonction définie au voisinage de $y_0$, dérivable en $y_0$.*

*La fonction $h = g \circ f$ est définie au voisinage de $x_0$, y est dérivable et*
$$h'(x_0) = (g \circ f)'(x_0) = f'(x_0).g'(y_0) = f'(x_0).g'(f(x_0))$$
*En notation physicienne, en posant $y = f(x)$,*
$$\frac{dh}{dx}(x_0) = \frac{dg}{dy}(y_0)\frac{dy}{dx}(x_0)$$

*Démonstration.* Le fait que $h$ soit bien définie au voisinage de $x_0$ est une conséquence de l'énoncé similaire sur la continuité.

D'après la proposition 3.1.1, il existe deux fonctions $\epsilon$ et $\eta$, définies au voisinage de 0 telles que pour tout $x$ voisin de $x_0$, tout $y$ voisin de $y_0 = f(x_0)$,
$$\begin{aligned} f(x) &= f(x_0) + (x - x_0)f'(x_0) + (x - x_0)\epsilon(x - x_0) \\ g(y) &= g(y_0) + (y - y_0)g'(y_0) + (y - y_0)\eta(y - y_0) \end{aligned}$$

$y = f(x)$ est suffisamment voisin de $y_0$ pour que l'on puisse écrire, par substitution de la première égalité dans la seconde que
$$\begin{aligned} g(f(x)) &= g(f(x_0)) + \left((x - x_0)f'(x_0) + (x - x_0)\epsilon(x - x_0)\right) g'(y_0) + \\ &\quad + \left((x - x_0)f'(x_0) + (x - x_0)\epsilon(x - x_0)\right) \eta \left((x - x_0)f'(x_0) + (x - x_0)\epsilon(x - x_0)\right) \\ &= g(f(x_0)) + (x - x_0)f'(x_0).g'(y_0) + (x - x_0)\epsilon'(x - x_0) \end{aligned}$$

On a ici posé
$$\epsilon'(h) = \epsilon(h)g'(y_0) + \left(f'(x_0) + \epsilon(h)\right) \eta \left(h.f'(x_0) + h\epsilon(h)\right)$$
qui est clairement une fonction de limite 0 en 0. On a donc prouvé l'existence d'un DL d'ordre 1 de $h$ au voisinage de $x_0$, i.e la dérivabilité de $h$ en $x_0$ et le fait que
$$h'(x_0) = f'(x_0).g'(y_0) \quad \square$$

Maintenant que l'on connaît ce qui se passe pour les compositions, on peut se poser la question sur les fonctions réciproques.

Une première remarque est la suivante :

*Observation* 3.1.3. Si $f : I \to J$ est strictement monotone et dérivable, en prenant $y_0 = f(x_0)$, la seule valeur possible de $(f^{-1})'(y_0)$ est $\frac{1}{f'(x_0)}$ (sous condition, évidemment, que $f'(x_0) \neq 0$). Pourquoi ? prenons l'identité

$$f^{-1}(f(x)) = x$$

et dérivons à gauche et à droite. On trouve

$$(f^{-1})'(f(x)) \cdot f'(x) = 1.$$

Si on calcule en $x = x_0$, on trouve $(f^{-1})'(y_0) \cdot f'(x_0) = 1$. Ceci donnerait une formule pour $(f^{-1})'$, mais le seul problème est qu'on ne peut pas l'établir si on ne démontre pas au préalable que $f^{-1}$ est bien dérivable. Ce calcul en effet ne le montre pas, et donne juste la seule valeur possible de la dérivée, pourvu que la fonction réciproque soit dérivable.

Par contre il montre également que $f^{-1}$ ne peut pas être dérivable en $y_0$ si $f'(x_0) = 0$ (le produite des deux dérivées ne pourrait pas valoir 1).

Démontrons alors la dérivabilité.

**Théorème 3.1.6.** *Soit $f$ une fonction continue, strictement monotone sur un intervalle $I$. Soit $J = f(I)$ et $f^{-1} : J \to I$ sa fonction réciproque. Alors, Si $x_0 \in I$, $f'(x_0) \neq 0$ alors $f^{-1}$ est dérivable en $y_0 = f(x_0) \in J$ et*

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$

*Démonstration.* **(HP)**

Comme $f$ est dérivable en $x_0$, il existe une fonction $\epsilon$ de limite nulle en 0 telle que

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + (x - x_0)\epsilon(x - x_0) = y_0 + (x - x_0)(f'(x_0) + \epsilon(x - x_0))$$

Soit $y \in J$, $y = f(x)$ pour un unique $x \in I : x = f^{-1}(y)$. En substituant $y = f(x)$ dans le développement limité précédent, on a

$$y - y_0 = (f^{-1}(y) - f^{-1}(y_0))(f'(x_0) + \epsilon(f^{-1}(y) - f^{-1}(y_0)))$$

Comme $f'(x_0) \neq 0$, lorsque $y \to y_0$, $\frac{1}{f'(x_0)}\epsilon(f^{-1}(y) - f^{-1}(y_0)) \to 0$ car $f^{-1}$ est continue en $y_0$ et donc $f^{-1}(y) - f^{-1}(y_0) \to 0$. Il existe donc une fonction $\eta$ définie au voisinage de 0 et de limite nulle en 0 telle que $\eta(y - y_0) = \frac{1}{f'(x_0)}\epsilon(f^{-1}(y) - f^{-1}(y_0))$ et

$$f'(x_0) + \epsilon(f^{-1}(y) - f^{-1}(y_0)) = f'(x_0)(1 + \eta(y - y_0)).$$

Par ailleurs, $1 + \eta(y - y_0)$ a pour limite 1 lorsque $y \to y_0$ et donc pour tout $y$ suffisamment proche de $y_0$,

$$f^{-1}(y) - f^{-1}(y_0) = (y - y_0)\frac{1}{f'(x_0)}\frac{1}{1 + \eta(y - y_0)} = (y - y_0)\frac{1}{f'(x_0)} + (y - y_0)\frac{1}{f'(x_0)}\left(\frac{1}{1 + \eta(y - y_0)} - 1\right)$$

Comme $\left(\frac{1}{1+\eta(y-y_0)} - 1\right)$ tend vers 0 lorsque $y \to y_0$, on a donc obtenu un développement limité d'ordre 1 de $f^{-1}$ au voisinage de $y_0$. Ceci démontre que $f^{-1}$ est dérivable en $y_0$ et nous donne la valeur de $(f^{-1})'(y_0)$ annoncée. □

*Exemple* 3.1.4. **Les dérivées des fonctions trigonométriques inverses :** rappelons que

$$(\sin)'(x) = \cos x; \quad (\cos')(x) = -\sin x; \quad (\tan)'(x) = \left(\frac{\sin}{\cos}\right)'(x) = \frac{\sin^2 x + \cos^2 x}{\cos^2 x} = 1 + \tan^2 x.$$

Calculons alors

$$(\arcsin)'(x) = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-x^2}},$$

ceci parce que : que signifie-t-il $\cos(\arcsin x)$ ? il signifie trouver la valeur du cosinus du point $y$ dont le sinus vaut $x$, mais la relation $\cos^2 y + \sin^2 y = 1$ nous donne que cette valeur doit forcément être $\pm\sqrt{1-x^2}$ ; Pour déterminer son signe on utilise le fait que la définition d'arcsinus nous donne $\arcsin x \in -\pi/2, \pi/2$ et donc le cosinus doit être non négatif. De manière analogue on a

$$(\arccos)'(x) = \frac{1}{-\sin(\arcsin x)} = -\frac{1}{\sqrt{1-x^2}},$$

et enfin

$$(\arctan)'(x) = \frac{1}{1+\tan^2(\arctan x)} = \frac{1}{1+x^2}.$$

## 3.2 Fonctions dérivables sur un intervalle

**Définition 3.2.1.** Soit $f : D \to \mathbb{R}$ une fonction de variable réelle, à valeurs réelles. Soit $I$ un intervalle non trivial contenu dans $D$. Si la fonction $f$ est dérivable en tout point de $I$ (à droite ou à gauche si l'on s'intéresse à une borne de $I$ contenue dans $I$), on dit que $f$ est dérivable sur $I$. Dans ce cas, la fonction $f'$ définie sur $I$ par

$$f'(x) = \text{le nombre dérivé de } f \text{ en } x$$

est appelée **la fonction dérivée** de $f$ sur $I$

*Exemples et remarques* 3.2.1. 1. Les fonctions les plus simples sont les fonctions constantes sur $\mathbb{R}$. Celles-ci sont dérivables sur $\mathbb{R}$, leur dérivée est la fonction nulle.

2. La fonction $\sqrt{.}$ est une fonction dérivable sur $]0, +\infty[$, sa dérivée est la fonction $x \in\, ]0, +\infty[\, \mapsto \frac{1}{2\sqrt{x}}$.
3. La fonction $\sqrt{.}$ n'est pas dérivable sur son intervalle de définition car elle n'est pas dérivable à droite en 0.
4. exp, sin, cos sont dérivables sur $\mathbb{R}$. Leurs fonction dérivées sont respectivement exp, cos et $-\sin$.

**Proposition 3.2.1** (Recollement)**.** *Soient $I$ et $J$ deux intervalles fermés ayant pour seul point commun le point $a$. Soit $f$ une fonction de classe $C^1$ sur $I$ et sur $J$ alors,* **pourvu que** $f'_g(a) = f'_d(a)$ *la fonction $f$ est de classe $C^1$ sur l'intervalle $I \cup J$.*

Les résultats concernant les opérations sur les fonctions et la dérivabilité en un point $a$ s'étendent naturellement en des résultats sur les fonctions dérivées

**Proposition 3.2.2.** *1. Soit $f : I \to \mathbb{R}$ une fonction dérivable alors*
*(i) Si $\lambda \in \mathbb{R}$, $\lambda.f$ est dérivable sur $I$ et*

$$(\lambda.f)' = \lambda.f'$$

| $f(x)$ | $f'(x)$ | commentaires |
|---|---|---|
| $x^n$ | $nx^{n-1}$ | pour $x \in \mathbb{R}$ si $n \in \mathbb{N}$ |
| $x^n$ | $nx^{n-1}$ | pour $x \neq 0$ si $n \in \mathbb{Z}$ |
| $x^\alpha$ | $\alpha x^{\alpha-1}$ | pour $x > 0$ si $\alpha \in \mathbb{R}$ |
| $e^x$ | $e^x$ | pour $x \in \mathbb{R}$ |
| $\ln x$ | $\frac{1}{x}$ | pour $x > 0$ |
| $\ln \|x\|$ | $\frac{1}{x}$ | pour $x \neq 0$ |
| $\sin x$ | $\cos x$ | pour $x \in \mathbb{R}$ |
| $\cos x$ | $-\sin x$ | pour $x \in \mathbb{R}$ |
| $\tan x$ | $\frac{1}{\cos^2 x} = 1 + \tan^2 x$ | pour $x \notin \{\frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}\}$ |
| $\arcsin x$ | $\frac{1}{\sqrt{1-x^2}}$ | pour $-1 < x < 1$ |
| $\arccos x$ | $-\frac{1}{\sqrt{1-x^2}}$ | pour $-1 < x < 1$ |
| $\arctan x$ | $\frac{1}{1+x^2}$ | pour $x \in \mathbb{R}$ |

TABLE 3.1 – Les dérivées classiques

*(ii) Si $f$ ne s'annule pas sur $I$, $\frac{1}{f}$ est dérivable sur $I$ et*

$$\left(\frac{1}{f}\right)' = -\frac{f'}{f^2}$$

*2. Soit de plus $g$ dérivable sur $I$*

*(i) Les fonctions $f+g$ et $f.g$ sont dérivables sur $I$ et, sur $I$,*

$$(f+g)' = f' + g' \text{ et } (f.g)' = f.g' + f'.g$$

*(ii) Si $f$ ne s'annule pas sur $I$, alors $\frac{g}{f}$ est dérivable sur $I$ et*

$$\left(\frac{g}{f}\right)' = \frac{g'.f - g.f'}{f^2}$$

*3. Soit $J$ un intervalle et $g$ dérivable sur $J$. Si pour tout $x \in I$, $f(x) \in J$ alors la fonction $g \circ f$ est dérivable sur $I$ et*

$$(g \circ f)' = (g' \circ f).f'$$

*4. Si $f$ : est strictement monotone et $f' \neq 0$ sur $I$ alors $f^{-1} : J \to \mathbb{R}$ (avec $J = f(I)$) est dérivable sur $J$ et, sur $J$, on a*

$$(f^{-1})' = \frac{1}{f'} \circ f^{-1}.$$

## 3.3 Extrema et points critiques

**Définition 3.3.1.** Une fonction $f : D \to \mathbb{R}$ admet un **maximum local** en $a \in D$ s'il existe un voisinage $V$ de $a$ tel que pour tout $x \in V \cap D$, $f(x) \leq f(a)$.

*Exemples et remarques* 3.3.1. On a une définition semblable pour un minimum local. Si $f$ admet en $a$ un maximum ou un minimum local, on dit que $f$ admet un extremum local.

Si $f$ admet un **maximum global** en $a$, i.e pour tout $x \in D$, $f(x) \leq a$, elle y admet *a fortiori* un maximum local.

**Proposition 3.3.1.** *Soit $f$ une fonction à valeurs réelles, dérivable sur un intervalle $I$, et $a$ un point intérieur à $I$ (i.e $a$ n'est pas une borne de $I$).*

*Si $f$ admet un extremum local en $a$ alors $f'(a) = 0$.*

*Démonstration.* Supposons que $f'(a) \neq 0$, nous pouvons écrire le DL d'ordre 1 de $f$ en $a$ : il existe une fonction $\epsilon$, continue en 0, $\epsilon(0) = 0$, telle que pour tout $x$ voisin de $a$,

$$f(x) - f(a) = f'(a).(x-a)\,(1 + \epsilon(x-a))$$

Sur un certain voisinage de $a$, la quantité $1 + \epsilon(x-a)$ est strictement positive et donc, sur ce voisinage, $f(x) - f(a)$ est du même signe que $f'(a).(x-a)$. De part et d'autre de $a$ (on se sert **là** du fait que $a$ est intérieur à $I$), cette quantité prend des valeurs $> 0$ et $< 0$, ce qui est impossible si $f$ admet un extremum local en $a$. $\square$

**Critère de recherche de minima et maxima absolus :**

Soit $f : [a,b] \to \mathbb{R}$ une fonction continue, dérivable en tout point de $]a,b[\backslash S$, où $S$ est en ensemble fini et, si possible, pas nombreux. On peut donc composer une liste de points de $[a,b]$ qui sont candidats à être minima et/ou maxima, et qui est composée de :

– les points de bord $a$ et $b$, car là le critère du théorème 3.3.1 n'est pas applicable ;
– les points de $S$ car là non plus on ne peut l'appliquer ;

– tout point $x_0 \in ]a,b[\setminus S$ qui satisfasse $f'(x_0) = 0$ (ce qui est souvent réalisé par un petit nombre de points).

Le théorème de Weierstrass nous garantit que le minimum et le maximum existent et en plus on est sûr qu'on les trouvera parmi les points qu'on a listé. Il suffit donc de calculer les valeurs $f(x)$ corréspondant à tous les points $x$ de la liste et on trouvera le (ou les) point(s) de minimum en prenant ceux qui ont les valeurs les moins élevées et les points de maximums en prenant ceux qui ont les valeurs le plus élevées.

Évidemment plein d'autres critères similaires peuvent être bâtis pour d'autres situations, par exemple pour une fonction satisfaisant aux conditions du Théorème 2.2.5.

*Observation* 3.3.2. Quelle est la différence entre cette approche et celle par tableau de variations (qui se base sur le lien entre le signe de $f'$ et le sens de variations de $f$, ce que l'on verra dans la prochaine section) ? pas grande chose. Ici on ne calcule pas les signes de la dérivée (et donc on économise du temps) mais on se retrouve à regarder comme candidats ces points aussi qui ont dérivée nulle (ou où $f$ n'est pas dérivable) mais qui ne peuvent pas vraiment minimiser à cause des signes de la dérivée juste avant et juste après (et donc on perd un petit peu plus de temps à calculer les valeurs de $f$ sur ces points là). Si on cherche et le maximum et le minimum le tableau des variations nous force à calculer les signes et en plus souvent ne nous fait pas économiser beaucoup d'évaluations de $f$ sur les points qu'on a trouvé, car la plus part d'entre eux sont soit candidats à minimiser soit à maximiser (sauf ceux qui sont du type $x^3$, où la dérivée s'annule sans changer de signe, oeuceux qui sont de type $2x + |x|$, où la fonction n'est pas dérivable mais dérivée gauche et droite ont le même signe). Par contre, si l'on cherche seulement l'une des deux bornes (min ou max), alors p-e le tableau de variations peut être rentable.
Mais c'est des petites différences, c'est à peu près la même idée. Un avantage de cette approche est par contre qu'on pourra l'appliquer en dimension plus grande parce qu'une fonction définie sur le plan n'a pas un "sens de variations".
Dernière chose : imaginons que $f$ n'est pas continue, mais que quelqu'un nous garantisse qu'il existe un maximum (ou un minimum). Après, si l'on utilise un tableau de variations, il faut être trèèès attentifs, car aux points de discontinuité il y a typiquement un saut. La fonction $f(x) = x$ sur $x \geq 0$ et $f(x) = x + 1$ sur $x < 0$ a une dérivée positive avant et après 0 mais 0 est quand même un candidat minimum (à cause du saut en bas entre $0^-$ et $0^+$).

## 3.4 Rolle, TAF et applications

On présente ici quelques théorèmes importants sur les fonctions dérivables qui seront répris dans la suite.

**Théorème 3.4.1** (Rolle)**.** *Soit $f : [a,b] \to \mathbb{R}$ une fonction continue sur l'intervalle fermé $[a,b]$ et dérivable en tout point de l'intervalle ouvert $]a,b[\subset [a,b]$. Supposons $f(a) = f(b)$. Il existe alors un point $\xi \in ]a,b[$ tel que l'on ait $f'(\xi) = 0$.*

*Démonstration.* La fonction $f$ étant continue et l'intervalle $[a,b]$ fermé et borné, d'après le théorème de Weierstrass on sait que $f$ admet un maximum et un minimum sur cet intervalle. Soient $m$ et $M$ les valeurs du minimum et du maximum, respectivement, et $K$ la valeur commune de $f(a)$ et $f(b)$. On a forcement $m \leq K \leq M$. Si toute les inégalités sont des égalités alors on a $m = M$ et la fonction est constante. Mais une fonction constante admet partout dérivée nulle et tout point $\xi \in ]a,b[$ satisfait la condition. Supposons donc qu'on ait soit $m < K = f(a) = f(b)$ soit l'autre inégalité. Dans ce premier cas, il signifie qu'il existe un point de minimum $x_0 \in [a,b]$ et que $f(x_0) < f(a) = f(b)$. Celci implique en particulier que $x_0$ est distinct de $a$ et de $b$. Il est donc à l'intérieur et on sait que si un point à l'intérieur réalise le minimum (un minimum local aurait été suffisant) d'une fonction dérivable, la dérivée en ce point là doit forcement être nulle.

Il suffit donc de choisir $\xi = x_0$. Pareillement, si c'est le maximum $M$ qui dépasse strictement $K$, on en déduit l'existence d'un point intérieur de maximum et ce point aura dérivée nulle. $\square$

*Observation* 3.4.1. Vous pouvez remarquer qu'on a vraiment utilisé toutes les hypothèses : la dérivabilité en tout point intérieur est à demander car on ne sait pas où le minimum tombera, et là où il tombe on a besoin de dire que la dérivée vaut zéro ; la continuité d'autre côté est demandée pour assurer l'existence d'un minimum et un maximum, et on en a besoin sur un ensemble fermé borné. Non seulement, il est facile de se rendre compte que sans la continuité en $a$ et $b$ le théorème n'aurait aucune chance d'être vrai : si l'on admet que les valeurs en $a$ et $b$ puissent différer dess limites de $f(x)$ lors que $x$ tend vers $a$ et $b$ repsectivement, on a beau à supposer que les deux valeurs $f(a)$ et $f(b)$ sont égales, mais cela serait toujours une hypotèse inutile car elle n'implique rien au niveau du comportement intérieur de la fonction. Et évidemment sans $f(a) = f(b)$ le théorème est faux (par exemple prenez $f(x) = x$).

**Théorème 3.4.2** (Acroissements finis)**.** *Soit $f : [a, b] \to \mathbb{R}$ une fonction continue sur l'intervalle fermé $[a, b]$ et dérivable en tout point de l'intervalle ouvert $]a, b[\subset [a, b]$. Il existe alors un point $\xi \in ]a, b[$ tel que l'on ait*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

*Démonstration.* Soit $L = (f(b) - f(a))/(b - a)$ et considérons la fonction $g$ donnée par

$$g(x) = f(x) - Lx.$$

Cette fonction garde les mêmes propriétés de continuité et dérivabilité de la fonction $f$ car on y a rajouté tout simplement une fonction linéaire, qui est elle même continue et dérivable en tout point. Montrons que $g(a) = g(b)$, de façon à pouvoir appliquer le théorème de Rolle à la fonction $g$. On a

$$g(b) - g(a) = f(b) - f(a) - L(b - a) = f(b) - f(a) - \frac{f(b) - f(a)}{b - a}(b - a) = 0.$$

Ceci montre $g(b) = g(a)$. On peut donc appliquer le théorème de Rolle à la fonction $g$ et on obtient qu'il existe un point $\xi \in ]a, b[$ tel que

$$0 = g'(\xi) = f'(\xi) - L.$$

Grâce au choix de $L$ qu'on a fait, on vient de montrer la thèse. $\square$

Remarque inutile : au délà des Alpes ce théorème s'appelle Théorème de Lagrange. Indépendemment de son nom, il a plusieurs conséquences intéressantes.

La première qu'on va voir concerne le sens de variations d'une fonction.

**Définition 3.4.1.** Soit $D \subset \mathbb{R}$ et $f$ une fonction à valeurs réelles définie sur $D$.

1. On dit que $f$ est **croissante** sur $D$ si, pour tous $x, y \in D$, $x < y \Rightarrow f(x) \leq f(y)$
2. On dit que $f$ est **décroissante** sur $D$ si, pour tous $x, y \in D$, $x < y \Rightarrow f(x) \geq f(y)$
3. On dit que $f$ est **strictement croissante** sur $D$ si, pour tous $x, y \in D$, $x < y \Rightarrow f(x) < f(y)$
4. On dit que $f$ est **strictement décroissante** sur $D$ si, pour tous $x, y \in D$, $x < y \Rightarrow f(x) > f(y)$

**Théorème 3.4.3.** *Si $I$ est un intervalle de $\mathbb{R}$ et si $f$ est une fonction dérivable sur $I$ telle que $f' > 0$ sur $I$ alors $f$ est strictement croissante sur $I$.*

*Démonstration.* Soient $x < y$ deux points de $I$. La fonction $f$ est dérivable sur $]x, y[$ et continue sur $[x, y]$. On peut alors appliquer le théorème 3.4.2 (TAF) pour obtenir que, pour un certain $\xi \in I$,

$$\frac{f(y) - f(x)}{y - x} = f'(\xi) > 0$$

On a donc clairement que $f(y) > f(x)$. □

*Exemples et remarques* 3.4.2. 1. Si $f' < 0$ sur $I$, $f$ y est strictement décroissante.

2. Si $f'$ est seulement positive sur $I$ (i.e, elle peut s'annuler) alors $f$ est croissante sur $I$. Réciproquement, si $f$ est croissante sur un intervalle $I$, sa dérivée y est $\geq 0$.

3. Le théorème 3.4.3 est utilisé constamment lorsque l'on construit le **tableau de variations** de $f$. C'est lui qui justifie le passage de la ligne indiquant le signe de $f'$ à la ligne indiquant le sens de variation de $f$.

4. Lorsque l'on construit le tableau de variations de $f$, on cherche les intervalles sur lequel sa dérivée est $> 0$, ce que l'on marque dans le tableau de signe de $f'$ par un $+$. De même, on marque par un $-$ les intervalles sur lesquels $f' < 0$. Il s'agit donc de résoudre les deux inéquations d'inconnue $x$ : $f'(x) < 0$ et $f'(x) > 0$. Dans la pratique et dans la plupart des cas, la fonction $f$ est dérivable et sa dérivée $f'$ est une jolie fonction continue aussi : donc, pour déterminer ces zones, d'après le TVI, le théorème 2.2.1 [1], il suffit de résoudre l'équation $f'(x) = 0$. On sait alors que $f$ est de signe constant sur les intervalles délimités par les solutions de cette équation. On peut alors trouver ce signe en calculant, par exemple, $f(x)$ pour un $x$ bien choisi de l'intervalle ou par tout autre argument du même calibre.

*Exercice* 3.4.3. Donner un exemple de fonction $f$ dérivable sur un intervalle $I = [a, b]$ telle que
(i) $f$ est strictement croissante sur $I$,
(ii) $f$ s'annule en au moins un point de $I$.

**Théorème 3.4.4.** *Si $f$ est dérivable sur un intervalle $I$ et sa dérivée y est nulle alors $f$ est constante sur $I$.*

*Exemples et remarques* 3.4.4. 1. Ce théorème implique l'égalité, à une constante additive près, des primitives sur un intervalle $I$ d'une fonction $f$ continue sur $I$, voir le chapitres sur les intégrales

2. Il est aussi à la base de résultats d'unicité concernant les équations différentielles.

**Définition 3.4.2.** Une fonction $f : A \to \mathbb{R}$ est dite Lipschitzienne si il existe une constante $L$ (appelée "constante de Lipschitz", Lipschitz étant le mec qui a, peut-être, introduit cette notion) telle qu'on ait

$$|f(x) - f(y)| \leq L|x - y| \ \text{ pour tout } x, y \in A.$$

Souvent on appelle constante de Lipschitz la plus petite constante $L$ qui permet de réaliser l'inégalité (car, évidemment, si $L$ est une constante qui satisfait l'inégalité ci-dessus, tout $L' \geq L$ la satisfait aussi). On définit donc

$$Lip(f) := \min\{L \geq 0 : |f(x) - f(y)| \leq L|x - y| \ \text{ pour tout } x, y \in A\}.$$

Il est facile de voir que toute fonction Lipschitz est continue (il suffit toujours de choisir $\delta = \varepsilon/L$). Par contre, elle n'est pas forcement dérivable : il suffit de penser à a fonction $f(x) = |x|$, qui admet 1 comme constante de Lipschitz mais n'est pas dérivabe au point 0. Il est intéressant de remarquer que la non-dérivabilité vient ici de le non-existence de la limite des taux d'acroissement (limite différente à droite et à gauche), même si ils restent quand même bornés (contrairement

1. l'hypothèse de continuité de la dérivée est en fait inutile, on peut en effet montrer, c'est le théorème de Darboux, que $f'$ verifie toujours le TVI

au cas par exemple de la fonction $f(x) = x^{1/3}$, qui admet la limite des taux d'acroissement en 0, mais elle n'est pas finie, les taux n'étant pas bornés).

La notion de fonction Lipschitzienne dévient très facile dans le cas des fonctions dérivables.

**Proposition 3.4.5.** *Soit $f : [a,b] \to \mathbb{R}$ une fonction dérivable en tout point de $]a,b[$. Les deux faits suivants sont équivalents :*
– *la dérivée de $f$ est bornée sur $]a,b[$,*
– *$f$ est Lipschitzienne.*
*De plus, si la dérivée est bornée et on a $|f'(x)| \leq L$ pour tout $x \in ]a,b[$, alors $f$ admet la même valeur $L$ comme constante de Lipschitz*[2] *et si $f$ admet une constante de Lipschitz, la même constante est une borne supérieure pour $|f'(x)|$.*

*Démonstration.* Démontrons que une borne $L$ sur la dérivée entraîne la nature Lipschitz de la fonction $f$ avec constante de Lipschitz $L$. Prenons $x$, $y \in ]a,b[$ et supposons par simplicité $x < y$ (la condition de Lipschitz étant tout à fait symmétrique en $x$ et $y$). On peut appliquer le théorème des acroissements finis à la fonction $f$ sur $[x,y]$ et on obtient

$$\frac{f(y) - f(x)}{y - x} = f'(\xi),$$

ce qui implique

$$|f(x) - f(y)| = |f'(\xi)||x - y| \leq L|x - y|,$$

si $L$ est une constante qui borne la dérivée, c'est-à-dire telle que $|f'(\xi)| \leq L$ pour tout $\xi \in ]a,b[$.

Maintenant supposons que $f$ est Lipschitzienne avec constante de Lipschitz $L$ et démontrons $|f'(x_0)| \leq L$ pour tout $x_0 \in ]a,b[$. En mettant la valeur absolue partout dans la limite qui donne la définition de dérivée, on trouve

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

Pourtant on a $|f(x) - f(x_0)|/|x - x_0| \leq L$ et donc on en déduit $|f'(x_0)| \leq L$ (celle-ci est la partie facile de l'équivalence, alors que pour l'autre il est nécessaire d'utiliser le théorème des acroissements finis). □

.

2. Ce fait s'appelle parfois *inégalité des accroissements finis*

# Chapitre 4

# 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 $\xi$ 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 :

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \varepsilon(x)(x - x_0), \tag{4.1}$$

où $\varepsilon(x)$ est une quantité qui tend vers 0 lors que $x \to x_0$ (en fait, si l'on définit $\varepsilon$ 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

$\varepsilon$ 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.

## 4.1 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* 4.1.1. 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 4.1.1** (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 $\xi$ entre $x_0$ et $x_1$ tel que $f'(\xi) = l$.*

*Démonstration.* **(H.P.)** 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 remplacer $f$ par la fonction $x \mapsto f(x) - lx$ et, le cas échéant, 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 $\xi$ 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 empêche 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 $\xi \neq x_1$. □

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

**Théorème 4.1.2.** *Soit $A$ un intervalle et $f : A \to \mathbb{R}$ une fonction dérivable sur $A \setminus \{x_0\}$. Supposons*

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

*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 $\xi$ 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. □

*Observation* 4.1.2. 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). En particulier, cela a la conséquence suivante : si $f : [a, b] \to \mathbb{R}$ est une fonction dérivable en tout point sauf, peut-être en $x_0 \in ]a, b[$, et si les limites

$$\lim_{x\to x_0^+} f'(x) = l^+, \quad \lim_{x\to x_0^-} f'(x) = l^-$$

existent, alors il y a deux cas

– soit $l^+ = l^-$ et alors la limite $\lim_{x\to x_0,\, x\neq x_0} f'(x)$ existe, on applique le théorème précédent et on trouve que $f$ est également dérivable en $x_0$ et la dérivée y vaut $l^+ = l^-$ ;
– soit $l^+ \neq l^-$ et alors ce n'est que $f'_d(x_0)$ et $f'_g(x_0)$ qui existent, mais comme elles sont différentes, $f'(x_0)$ n'existe pas.

*Observation* 4.1.3. 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 4.1.2 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* 4.1.4. 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 4.1.1.** 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 4.1.3.** *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 $f' \neq 0$ e, grâce à $k \geq 1$, $f \in C^k$ implique la continuité de $f'$, qui sera donc 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* 4.1.5. 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 $C^\infty$, ce qui est introduit dans la définition suivante.

**Définition 4.1.2.** Une fonction $f$ est dite $C^\infty(A)$ si $f \in C^k(A)$ pour tout $k \geq 0$.

## 4.2 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 développements de Taylor sont des DL.

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

**Théorème 4.2.1.** *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 $\xi$ 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.* **(H.P.)** 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, apparemment 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 $\phi$ 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 $\xi$ entre $x$ et $x_0$.

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

$$R = A\frac{(x-x_0)^{n+1}}{(n+1)!}. \tag{4.2}$$

On peut bien appliquer le théorème de Rolle à la fonction $\phi$ 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 $\phi$ 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 $\phi$ 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 $\xi$ dans l'intervalle entre $x_0$ et $x$ avec $\phi'(\xi) = 0$. Ceci implique, grâce à la formule pour $\phi'$, qu'on a $f^{(n+1)}(\xi) = A$.

En remplaceant $A$ par $f^{(n+1)}(\xi)$ dans (4.2), 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-dégré du développement, mais dont le reste a une formulation moins explicite.

**Théorème 4.2.2.** *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

$$f'(x) = \sum_{k=1}^{n} \frac{f^{(k)}(x_0)}{(k-1)!}(x - x_0)^{k-1} + o((x - x_0)^n). \tag{4.3}$$

On écrit $R(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$ et on remarque que (4.3) peut s'écrire sous la forme

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

$\varepsilon$ 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

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

où l'inégalité est valable si $x \in (x_0 - \delta, x_0 + \delta)$, car dans ce cas là $\xi$ 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* 4.2.1. 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 étant suffisant) alors il existe en tout point un développement 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 développement est unique parmi les polynômes du même dégré.

**Théorème 4.2.3.** *Soit $f : A \to \mathbb{R}$ une fonction qui admet deux développements 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* $\deg P, \deg Q \leq n$. *Alors $P = Q$.*

*Démonstration.* On déduit facilement du fait que $P$ et $Q$ sont des dévéloppements 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 $1$, $x - x_0$, $(x-x_0)^2, \dots, (x-x_0)^n$ 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 $P = Q$. $\square$

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

**Proposition 4.2.4.** *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 le coefficient devant toute puissance d'exposant impaire est 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 $P(x) = -P(-x)$. □

## 4.3 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 iterée. On en donnera tout de suite un exemple, mais cet exemple sera précédé par quelques définitions concernant l'équivalence et la négligeabilité des fonctions tendant vers zéro.

**Définition 4.3.1.** 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* 4.3.1. 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".

Comme toujours, ce qui est important est d'être clair, si vous utilisez l'une de ces expressions, en faisant comprendre, si besoin y a, quelle convention vous choisissez, et de ne pas se bouleverser si trouvez une convention différente sur un livre ou un texte que vous lisez.

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* 4.3.2. 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.

### 4.3.1 Calculs de DL et de limites

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

*Exemple* 4.3.3. 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

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

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

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

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* 4.3.4. 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* 4.3.5. 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 (4.1) 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* 4.3.6. 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* 4.3.7. 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

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

### 4.3.2 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 $C^\infty$.

**Théorème 4.3.1.** *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 4.3.2.** *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 $\geq 0$ : 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* 4.3.8. 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 5

# Intégrales

## 5.1 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

$$v_1 t \text{ si } t \leq T_1; \quad v_1T_1+v_2(t-T_1) \text{ si } T_1 \leq t \leq T_1+T_2; \quad v_1T_1+v_2T_2+v_3(t-(T_1+T_2)) \text{ si } T_1+T_2 \leq t.$$

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 5.1.1.** 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

$$f(x) = \begin{cases} k_0 & \text{si } x = a, \\ m_1 & \text{si } a = t_0 < x < t_1, \\ k_1 & \text{si } x = t_1, \\ m_2 & \text{si } t_1 < x < t_2, \\ k_2 & \text{si } x = t_2, \\ \dots & \\ k_{n-1} & \text{si } x = t_{n-1}, \\ m_n & \text{si } t_{n-1} < x < t_n = b, \\ k_n & \text{si } x = b. \end{cases} \tag{5.1}$$

Souvent on considère des intervalles semi-fermés du 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. En fait, dans les intégrales, la valeur en un point seul ne compte jamais.

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 5.1.2.** Si $f : [a, b] \to \mathbb{R}$ est une fonction étagée de la forme (5.1) 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 $f$). Quand il n'y aura pas d'ambiguïté on pourra omettre l'indication de la variable muette $t$ et/ou de l'intervalle d'intégration, en écrivant par exemple $\int f$.

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 5.1.1.** *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 (5.1) 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.

**Définition 5.1.3.** Si $f : A \to \mathbb{R}$ est une fonction bornée et $I \in \mathbb{R}$ on dit que l'intégrale de $f$ sur $A$ vaut $I$ si la propriété suivante est satisfaite : pour tout $\varepsilon > 0$ il existe une fonction étagée $g^+ \geq f$ et une fonction étagée $g^- \leq f$ telles que

$$\int_A g^+(t)dt < I + \varepsilon \quad et \quad \int_A g^-(t)dt > I - \varepsilon.$$

Une fonction est dite intégrable si il existe une valeur $I \in \mathbb{R}$ qui satisfait cette définition.

Pourquoi avons-nous supposé que $f$ est bornée ? parce qu'autrement les ensembles des $g \in \mathcal{E}(A);\ g \geq f$ et des $g \in \mathcal{E}(A);\ g \leq f$ pourraient être vides.

Si vous y réfléchissez-bien, cela rappelle un peu la définition de limite. On va donc donner un théorème d'unicité aussi.

**Proposition 5.1.2.** *Si $f : A \to \mathbb{R}$ est intégrable, alors la valeur de son intégrale est unique.*

*Démonstration.* Il faut démontrer qu'on ne peut pas avoir deux valeurs différentes $I < J$ qui satisfont en même temps la définition d'intégrale. Pour le faire, prenons $\varepsilon < (J - I)/2$. Il existe alors $g^+ \in \mathcal{E}(A)$ telle que $g^+ \geq f$ et $\int_A g^+(t)dt < I + \varepsilon$. Il existe aussi $g^- \in \mathcal{E}(A)$ telle que $g^- \leq f$ et $\int_A g^-(t)dt > J + \varepsilon$. Or, puisque $J - \varepsilon > I + \varepsilon$, on trouve $\int_A g^+(t)dt < \int_A g^-(t)dt$. Pourtant on a $g^- \leq f \leq g^+$ et on sait que l'intégrale préserve l'ordre. On a donc une contradiction. □

Les fonctions intégrables d'après cette définition sont souvent dites fonction intégrables selon Riemmann ou Riemann-intégrables pour les distinguer des fonction intégrables d'après une autre définition, plus générale, donnée plus tard par Lebesgue. On notera $\mathcal{R}(A)$ l'ensemble des fonctions intégrables selon Riemann sur l'interval $A$.

On remarquera aussi la caractérisation suivante :

**Proposition 5.1.3.** *Une fonction bornée $f : A \to \mathbb{R}$ est intégrable si est seulement si, pour tout $\varepsilon > 0$, il existe une fonction étagée $g^+ \geq f$ et une fonction étagée $g^- \leq f$, définies sur une même partition adaptée $(t_i)_{i=1,\dots,n}$ par $g^+ = m_i^+$ sur $]t_i, t_i + 1[$ et $g^- = m_i^-$ sur $]t_i, t_i + 1[$, telles que*

$$\sum_{i=0}^{n-1} (t_{i+1} - t_i)(m_i^+ - m_i^-) < \varepsilon.$$

*Démonstration.* Démontrer que si $f$ est intégrable alors des fonctions $g^\pm$ comme ça existent n'est pas dur. Le fait qu'elles soient définies sur une même partition adaptée n'est pas restrictif (il suffit de rajouter des points aux partitions, si nécessaire). La condition qu'on cherche à vérifier est

$$\int_A g^+(t)dt - \int_A g^-(t)dt < \varepsilon,$$

ce qui est évidemment vérifié si on prend les fonctions correspondantes à $\varepsilon/2$ dans la définition d'intégrale.

Il est par contre très délicat de démontrer la réciproque. Pourquoi ? parce que pour démontrer qu'une fonction $f$ admet une intégrale (et elle est donc intégrable), il faudrait produire la valeur du nombre $I$. Il faudrait en effet prendre quelque chose du type "le minimum parmi les valeurs qui sont plus grandes que toutes les intégrales des fonctions étagées inférieures ou égales à $f$". Mais il faudrait démontrer que ce minimum existe. Cela repose sur des propriétés des nombre réelles qu'on ne peut pas approfondir, on va donc y croire. □

Ce critère d'intégrabilité nous dit que les fonctions intégrables sont celles qu'on peut encadrer par des fonctions étagées tout en obtenant une différence très petite entre les deux intégrales. Ceci demande à avoir soit des petites valeurs de $t_{i+1} - t_i$, soit des petites valeurs de $m_i^+ - m_i^-$.

On peut donc dire que les fonctions intégrables sont les fonctions qui ont une petite oscillation, si l'on considère la possibilité de découper $[a, b]$ en sous-intervalles pour réduire les oscillations et on pondère ces oscillations avec les longueurs des correspondants intervalles. Une fonction est intégrable donc si elle oscille peu ou si ses grandes oscillations sont concentrées sur des petits intervalles.

Il est facile vérifier la propriété suivante :

**Proposition 5.1.4.** *L'espace $\mathcal{R}(A)$ est un espace vectoriel et, si $f, g \in \mathcal{R}(A)$ et $\lambda \in \mathbb{R}$, alors $\int_A (f+g)(t)dt = \int_A f(t)dt + \int_A g(t)dt$ et $\int_A (\lambda f)(t)dt = \lambda \int_A f(t)dt$. De plus, si $\phi : [c, d] \to \mathbb{R}$ est une fonction Lipschitzienne et $f(A) \subset [c, d]$ alors $\phi \circ f$ est intégrable aussi. Pour terminer, $fg$ aussi est intégrable*

*Démonstration.* Commençons par démontrer que $f + g$ est intégrable, si $f$ et $g$ le sont. Évidemment elle est bornée comme somme de fonctions bornées. Considérons en plus des fonctions étagées $h^\pm$ et $k^\pm$, toutes définies sur une même partition adaptée avec valeurs $h_i^\pm$ et $k_i^\pm$ sur $]t_i, t_i + 1[$, respectivement, et telles que

$$h^+ \geq f,\ h^- \leq f,\ k^+ \geq g \ \text{ et } k^- \leq g,$$
$$\sum_{i=0}^{n-1} (t_{i+1} - t_i)(h_i^+ - h_i^-),\ \sum_{i=0}^{n-1} (t_{i+1} - t_i)(k_i^+ - k_i^-) < \varepsilon/2.$$

Si l'on considère les fonctions étagées $u^\pm = h^\pm + k^\pm$ on a justement $u^+ \geq f + g \geq u^-$ et $u^\pm = h_i^\pm + k_i^\pm$ sur $]t_i, t_i + 1[$, tout en ayant aussi

$$\sum_{i=0}^{n-1} (t_{i+1} - t_i)((h_i^+ + k_i^+) - (h_i^- + k_i^-)) < \varepsilon.$$

Ceci montre que $f + g$ est intégrable grâce à la Proposition (5.1.3).

Pour démontrer que $\int f + g = \int f + \int g$ il suffit de prendre $I$ et $J$ ; intégrales de $f$ et $g$, et des fonctions étagées $h^\pm$ et $k^\pm$ telles que

$$h^+ \geq f,\ h^- \leq f,\ k^+ \geq g \ \text{ et } k^- \leq g,$$
$$I - \frac{\varepsilon}{2} < \int h^-(t)dt < \int h^+(t)dt < I + \frac{\varepsilon}{2}\ ; \quad J - \frac{\varepsilon}{2} < \int k^-(t)dt < \int k^+(t)dt < J + \frac{\varepsilon}{2}.$$

Ceci nous montre facilement que les fonctions étagées satisfont $h^+ + k^+ \geq f + g \geq h^- + k^-$ et $\int (h^+ + k^+)(t)dt < I + J + \varepsilon$ et $\int (h^- + k^-)(t)dt > I + J - \varepsilon$, ce qui nous dit $\int (f + g)(t)dt = I + J = \int f(t)dt + \int g(t)dt$ (remarquer comment cette preuve est similaire à celle correspondante concernant les limites).

Pour montrer que si $f$ est intégrable et $\lambda \in \mathbb{R}$ alors $\lambda f$ est intégrable on peut voir ça comme un cas particulier de la composition avec une fonction Lipschitzienne qu'on verra plus en bas. Par contre, dans le cas général d'une fonction Lipschitzienne il n'est pas possible de déterminer la valeur de l'intégrale. Ici il est au contraire simple de démontrer qu'elle vaut $\lambda \int f$ et la démonstration procède comme celle pour le cas de la somme

On veut montrer aussi l'intégrabilité de $\phi \circ f$ et $fg$. Pour la première, prenons $g^\pm$ fonctions étagée définies usuellement avec $g^+ \geq f \geq g^-$ et

$$\sum_{i=0}^{n-1} (t_{i+1} - t_i)(m_i^+ - m_i^-) < \varepsilon/L,$$

où $L$ est la constante de Lipschitz de $\phi$ sur l'intervalle $[c, d]$. On peut aussi supposer que, comme $f(A) \subset [c, d]$, les coefficients $m_i^\pm$ ont été choisis dans l'intervalle $[c, d]$. Posons

$$h_i^+ := \max\{\phi(s) \,:\, s \in [m_i^-, m_i^+]\}; \quad h_i^- := \min\{\phi(s) \,:\, s \in [m_i^-, m_i^+]\}.$$

Ces minima et maxima existent, parce que $\phi$ est une fonction continue et on regarde les intervalles compacts $[m_i^-, m_i^+]$. Comme $f(t) \in [m_i^-, m_i^+]$ pour $t \in ]t_i, t_{i+1}[$, alors les fonctions $h^\pm$ définies comme étant $h_i^\pm$ sur $]t_i, t_{i+1}[$ majorent et minorent $\phi \circ f$, respectivement. De plus, grâce à la Lipschitziennité sur $[c, d]$, on a $h_i^+ - h_i^- \leq L(m_i^+ - m_i^-)$. En particulier

$$\sum_{i=0}^{n-1} (t_{i+1} - t_i)(h_i^+ - h_i^-) < \varepsilon,$$

ce qui montre l'intégrabilité de $\phi \circ f$.

Pour le produit $fg$ on utilisera un truc malin : on note que

$$fg = \frac{(f+g)^2 - f^2 - g^2}{2}$$

et que le carré d'une fonction intégrable est intégrable. Pourquoi ? parce que si $f$ est une fonction intégrable qui prend ses valeurs dans $[c, d]$ (comme elle est bornée, ses valeurs sont toujours inclues dans un intervalle borné comme ça), alors on peut appliquer le critère précédent, comme la fonction $x \mapsto x^2$ est Lipschitzienne de constante $2\max\{|c|, |d|\}$ sur cet intervalle (il s'agit d'une fonction $C^1(\mathbb{R})$, donc sa dérivée est bornée sur tout intervalle borné). En appliquant ceci à $f$, $g$ et $f + g$, puis en faisant des sommes et la multiplication par $1/2$, on trouve le résultat. $\square$

On termine la section par un exemple bien connu de fonction non-intégrable, qui est connue comme "fonction de Dirichlet".

*Exemple* 5.1.1. Soit $f$ la fonction donnée par

$$f(x) = \begin{cases} 1 & \text{si } x \in \mathbb{Q}, \\ 0 & \text{si } x \notin \mathbb{Q}. \end{cases}$$

Bien que bornée, la fonction $f$ n'est pas intégrable sur aucun intervalle $[a, b]$. En effet, la fonction constante 1 est une fonction étagée qui majore $f$ si comme la fonction constante 0 est une fonction étagée qui la minore. De plus, toute fonction étagée qui majore $f$ doit valoir au moins 1 à l'intérieur de tout intervalle $]t_i, t_{i+1}[$ de sa partition adaptée, car cet intervalle inclut forcement des points rationnels où $f$ vaut 1. Symétriquement, toute fonction étagée minorante ne peut pas dépasser la valeur 0 sur les mêmes intervalles car ces intervalles contiennent tous des point irrationnels où $f$ vaut 0. Ceci montre que l'intégrale des fonctions étagées minorantes vaut au maximum 0 et celle des majorantes au minimum 1. Il n'est donc pas possible de satisfaire la condition $\int g^+ - \int g^- < \varepsilon$.

## 5.2 Des classes de fonctions intégrables

### 5.2.1 Fonctions monotones

Dans cette section on va démontrer que des classes suffisamment amples de fonctions sont en fait des fonctions intégrables. On parlera notamment de fonctions continues et de fonctions monotone. Le cas le plus simple est celui des fonctions monotone. Nous rappelons ici que, bien qu'on ait mis en évidence au Chapitre 3 le lien entre signe de la dérivée et monotonie, il y a des fonctions monotones non dérivables. Et même non continues. Parce que si on se limitait aux fonctions dérivables avec dérivée de signe constante, ça serait des fonctions dérivables, donc continues, et donc pas la peine de distinguer le cas des monotones de celui des continues.

**Théorème 5.2.1.** *Si $f : [a, b] \to \mathbb{R}$ est monotone, alors $f$ est intégrable.*

*Démonstration.* Supposons par simplicité que $f$ soit croissante (le cas décroissante étant tout à fait symétrique) Fixons $n$ et prenons la partition $(t_i)_{i=0,\dots,n}$ définie par $t_i = a + (b-a)/n$ (une partition uniforme qui divise $[a,b]$ en $n$ intervalles de même taille). On appliquera le critère 5.1.3. Comme $f$ est croissante sur les intervalles $t_i, t_{i+1}[$ on peut choisir deux fonctions étagées $g_n^+$ et $g_n^-$, majorante et minorante de $f$, définies par

$$g_n^+(x) = f(t_{i+1}), \quad g_n^-(x) = f(t_i) \text{ si } x \in ]t_i, t_{i+1}[,$$

sur les mêmes intervalles de ces partitions uniformes $(t_0, t_1, \dots, t_n)$.

Faisons le calcul de $\int_a^b g_n^+(t)dt - \int_a^b g_n^-(t)dt = \int_a^b (g_n^+ - g_n^-)(t)dt$ :

$$\int_a^b (g_n^+ - g_n^-)(t)dt = \sum_{i=0}^{n-1} \frac{b-a}{n}(f(t_{i+1}) - f(t_i)) = \frac{b-a}{n}(f(b) - f(a)).$$

Cette dernière quantité peut devenir aussi petite que l'on veut si $n$ est très grand. Il suffit alors de choisir $n$ suffisamment grand pour que cela soit plus petit qu'$\varepsilon$ et appliquer donc le critère 5.1.3. □

L'idée clé de cette preuve est que pour une fonction monotone on obtient un effet téléscopique de termes qui s'effacent si l'on prend des sous-intervalles de la même longueur. Il faut remarquer qu'en général il n'est pas possible de démontrer l'intégrabilité d'une fonction en se restreignant aux partitions régulières (celles où tous les intervalles qui ont la même longueur). Le fait que pour les fonctions monotones soit possible est une propriété en plus de cette classe de fonctions et le fait que l'intégrale soit obtenue comme limite des sommes des valeurs de la fonction dans des points spéciaux aussi.

Les mêmes propriétés sont vraies pour les fonctions continues mais sont moins évidentes. Ceci est dû au fait qu'on n'a pas encore introduit une notion importante reliée à la continuité, qui est celle de continuité uniforme.

### 5.2.2 L'intégrabilité des fonctions continues et l'uniforme continuité

**Définition 5.2.1.** 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 $\delta$, alors qu'ici $\delta$ ne doit dépendre que de $\varepsilon$. 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* 5.2.1. 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 $\varepsilon$. Par contre, la même démonstration n'aurait pas marché si on se restreignait à un ensemble borné de valeurs de $x$.

En revanche, 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 $\delta = \varepsilon/L$. Une classe plus générale est la classe des fonctions Hölderiennes (là aussi Hölder est le mec qui a - peut-être - étudié le premier les propriétés de ces fonctions).

**Définition 5.2.2.** Si $\alpha \in ]0,1]$ une fonction $f : A \to \mathbb{R}$ est dite Hölderienne d'exposant $\alpha$ (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* 5.2.2. La fonction $f : [0,1] \to \mathbb{R}$ donnée par $f(x) = \sqrt{x}$ est Hölderienne d'exposant $\frac{1}{2}$ mais elle n'est pas Lipschitzienne.

Pour voir qu'elle Hölderienne, prenons $x < y$, puis considérons

$$(\sqrt{x} + \sqrt{y-x})^2 = x + y - x + 2\sqrt{x}\sqrt{y-x} \geq y.$$

En prenant la racine à droite et à gauche on obtient

$$\sqrt{x} + \sqrt{y-x} \geq \sqrt{y} \Rightarrow |f(y) - f(x)| = \sqrt{y} - \sqrt{x} \leq |y-x|^{1/2}.$$

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 $\frac{1}{2\sqrt{x}}$ 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* 5.2.3. Pourquoi parle-t-on de fonctions Höldériennes pour $\alpha \leq 1$ et non 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.

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 les théorème suivants, dont on ne voit pas es preuves.

**Théorème 5.2.2.** *Soient $A$ un sous-ensemble fermé et borné de $\mathbb{R}$ et $f : A \to \mathbb{R}$ une fonction continue. Alors $f$ est uniformément continue sur $A$.*

Une autre propriété importante des fonctions uniformément continues est la suivante :

**Théorème 5.2.3.** *Soit $f : A \to \mathbb{R}$ une fonction uniformément continue. Alors il existe une unique extension continue $g$ de $f$ à l'adhérence $\bar{A}$ de l'ensemble $A$, c'est-à-dire une fonction continue $g : \bar{A} \to \mathbb{R}$ telle que $g(x) = f(x)$ pour tout $x \in A$.*

**Corollaire 5.2.4.** *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* 5.2.4. 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 $t_0, t_1, \dots t_n$ 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 $\delta$ 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.

**Théorème 5.2.5.** *Soit $f : [a, b] \to \mathbb{R}$ une fonction continue. Alors $f$ est intégrable.*

*Démonstration.* Grâce au Théorème 5.2.2 la fonction $f$ est uniformément continue aussi. Fixons $\varepsilon > 0$ et appliquons la définition d'uniforme continuité à $\varepsilon/(b-a)$ en trouvant un $\delta$. On divisera l'intervalle $[a, b]$ en $n$ sous-intervalles de la même longueur $(b-a)/n$ en appelant $t_i = a+i(b-a)/n$ les points de séparation. Si on choisit $n$ suffisamment grand on aura $(b - a)/n < \delta$.

Comme $f$ est continue elle admet un minimum et un maximum sur chacun des intervalles fermés $[t_i, t_{i+1}]$ et on peut donc définir deux fonctions étagées $g^+$ et $g^-$, majorante et minorante de $f$ comme suit

$$g^+(x) = \max\{f(t) \,:\, t \in [t_i, t_{i+1}]\}, \quad g^-(x) = \min\{f(t) \,:\, t \in [t_i, t_{i+1}]\} \quad \textit{pour} x \in ]t_i, t_{i+1}[.$$

Remarquons que, à cause de leur définition et du choix de $\delta$, les fonctions $g^+$ et $g^-$ satisfont $g^+(x) - g^-(x) \leq \varepsilon/(b - a)$. En effet, on a $\max\{f(t) \,:\, t \in [t_i, t_{i+1}]\} = f(t_M)$ et $\min\{f(t) \,:\, t \in [t_i, t_{i+1}]\} = f(t_m)$ et $|t_M - t_m| \leq |t_{i+1} - t_i| < \delta$ et $\delta$ avait été choisi de manière à garantir que $|t - s| < \delta \Rightarrow |f(t) - f(s)| < \varepsilon/(b - a)$.

Faisons le calcul de $\int_a^b g^+(t)dt - \int_a^b g^-(t)dt = \int_a^b (g^+ - g^-)(t)dt$ :

$$\int_a^b (g^+ - g^-)(t)dt \leq \sum_{i=0}^{n-1} \frac{b-a}{n} \frac{\varepsilon}{b-a} = n\frac{b-a}{n} \frac{\varepsilon}{b-a} = \varepsilon.$$

Ceci démontre que les intégrales des deux fonction étagées qu'on a choisies, qui majorent et minorent $f$, diffèrent d'une quantité plus petite que $\varepsilon$, et démontre l'intégrabilité. $\square$

Ceci montre que les fonctions continues aussi, comme les fonctions monotones, non seulement sont intégrables, mais elles ont aussi la propriété que leurs intégrales peuvent être calculées par partitions uniformes.

## 5.3 Le théorème fondamental du calcul et l'IPP

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 relie les concepts d'intégrale et de dérivée.

Pour en parler on a besoin d'abord de remarquer un petit détail assez facile et connu sous le nom de "formule de Chasles".

**Lemme 5.3.1.** *Si $f$ est une fonction intégrable 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 $g^\pm$ de la définition d'intégrale pour $f$ et le résultat découle pour $f$ aussi. $\square$

On est prêt à montrer le suivant :

**Théorème 5.3.2** (Théorème fondamental du calcul intégral)**.** *Soit $f : [a,b] \to \mathbb{R}$ une fonction intégrable, continue au point $x_0$. Alors la fonction $F : [a,b] \to \mathbb{R}$ définie par*

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

*est dérivable au point $x_0$ et $F'(x_0) = f(x_0)$. Si $f$ est continue en tout point alors $F \in C^1$ et $F' = f$.*

*Démonstration.* Si l'on remarque que toute fonction continue est intégrable, il suffit de démontrer la première 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, il suffira d'appliquer le même raisonnement en tout point.

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

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

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 $f(t) \le f(x_0) + \varepsilon$. On déduit donc

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

On peut obtenir aussi

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

si l'on considère que les points $t \in ]x_0, x[$ satisfont $f(t) \ge f(x_0) - \varepsilon$ 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 5.3.1.** 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 5.3.3** (Intégration par parties, ou IPP). *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. $\square$

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.

## 5.4 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.

### 5.4.1 Intégrales sur des intervalles spéciaux (symétrie, périodicité)

*Exemple* 5.4.1. 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

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

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* 5.4.2. 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…

### 5.4.2 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* 5.4.3. 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'un 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 $a$, car il sera une constante par rapport à la 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* 5.4.4. 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* 5.4.5. 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

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

### 5.4.3 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* 5.4.6. 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* 5.4.7. 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

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

### 5.4.4 Changement de variable d'intégration

L'exemple 5.4.7 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 5.4.1.** *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* 5.4.8. 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 $y \mapsto \sin(y)$ 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 5.4.1 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

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

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

### 5.4.5 Fonctions rationnelles

*Exemple* 5.4.9. 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* 5.4.10. 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 $P = P_1 Q + P_2$ (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

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

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 vraie 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 $c_1 c_2 > 0$. 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* 5.4.11. 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 admet 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. Du coup, on ne pourra pas s'en sortir comme ça.

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* 5.4.12. 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 $P = P_1Q + P_2$ 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 $i \le m$ sont des polynômes de degré un et les $Q_i$ avec $i \ge m+1$ sont des polynômes de degré deux sans racines réelles ; les correspondants $\alpha_i$ leurs exposants dans la factorisation ;
- écrire
$$\begin{aligned} \frac{P_2(x)}{Q(x)} &= \frac{A_{1,1}}{Q_1(x)} + \dots + \frac{A_{1,\alpha_1}}{Q_1(x)^{\alpha_1}} + \dots + \frac{A_{m,1}}{Q_m(x)} + \dots + \frac{A_{m,\alpha_m}}{Q_m(x)^{\alpha_m}} \\ &\quad + \frac{A_{m+1,1} + B_{m+1,1}x}{Q_{m+1}(x)} + \dots + \frac{A_{m+1,\alpha_{m+1}} + B_{m+1,\alpha_{m+1}}x}{Q_{m+1}(x)^{\alpha_{m+1}}} + \dots \\ &\quad + \frac{A_{k,1} + B_{k,1}x}{Q_k(x)} + \dots + \frac{A_{k,\alpha_k} + B_{k,\alpha_k}x}{Q_k(x)^{\alpha_k}}, \end{aligned}$$
  c'est-à-dire en mettant au dénominateur tous les facteurs $Q_i$ et en les prenant à une puissance entre un et leur $\alpha_i$, 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 $A/Q_i^\beta$ 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 $\mu/Q_i^\beta$ 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 5.4.12

### 5.4.6 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* 5.4.13. 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 $\tan(b/2)$.

## 5.5 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 5.5.1.** *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 $|H'(\xi)| \leq \varepsilon|\xi - x_0|^n \leq \varepsilon|x - x_0|^n$ et finalement $|H(x)| \leq \varepsilon|x - x_0|^{n+1}$. 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 $\xi$ n'est pas precisé ; cette formule donnera par contre une égalité explicite. Par contre, elle demandera plus de régularité pour pouvoir l'appliquer.

**Théorème 5.5.2.** *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

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

# Chapitre 6

# Courbes paramétrées planes

## 6.1 Généralités

On se place dans le plan muni d'un repère $(O, \vec{i}, \vec{j})$ orthonormé. Ce plan est alors identifié à $\mathbb{R}^2$, l'ensemble des couples de réels, via le mécanisme des coordonnées cartésiennes.

**Définition 6.1.1.** – Un **paramétrage de courbe plane** ou, plus simplement, une **courbe paramétrée plane** est la donnée de deux fonctions réelles de variable réelle $x$ et $y$ définies, continues sur un même intervalle non trivial $I$ de $\mathbb{R}$. C'est aussi la donnée d'une fonction continue [1] $\gamma : I \to \mathbb{R}^2$, $\gamma(t) = (x(t), y(t))$.

– A chaque $t \in I$, on associe le point $M(t)$ de coordonnées $\gamma(t) = (x(t), y(t))$. L'ensemble des points $M(t)$ lorsque $t$ décrit $I$

$$\Gamma = \{(x(t), y(t)),\ t \in I\}$$

est appelé la **courbe plane** [2] associée à ce paramétrage.

*Exemples et remarques* 6.1.1.

1. La droite. Donner un paramétrage de la droite $D$ passant par le point $A = (1, 1)$, de vecteur directeur $u = (-2, 3)$.
   Donner un autre paramétrage de cette même droite.
2. Le cercle. Même exercice avec le cercle de centre 0 de rayon 1. En utilisant le fait que le cercle de centre $C = (-1, 2)$, de rayon 3 est l'image par une homothétie de centre $O$ suivie d'une translation du cercle précédent, déterminer un paramétrage de ce cercle.
3. Le graphe d'une fonction continue. Soit $f : I \to \mathbb{R}$ une fonction continue, son graphe $\Gamma$ est la courbe paramétrée dont un paramétrage est
$$x(t) = t,\ y(t) = f(x(t)),\ t \in I$$
4. Si $\Gamma$ est la courbe paramétrée par $(x(t), y(t))$, $t \in I \subset \mathbb{R}$ et si $\phi$ est une bijection continue d'un intervalle $J \subset \mathbb{R}$ sur $I$ alors $\Gamma$ est aussi la courbe paramétrée par $(x(\phi(s), y(\phi(s)),\ s \in J$.

### 6.1.1 Vecteur vitesse et tangente

Dans les exemples qui vont nous occuper, les fonctions $x$ et $y$ seront, la plupart du temps, des fonctions régulières (i.e. de classe $C^1$) de $t \in I$.

Dans une interprétation physique usuelle des courbes paramétrées, la variable $t$ représente le temps et $x(t), y(t)$ sont les coordonnées d'un point mobile $M(t)$.

1. Une fonction à valeurs $\mathbb{R}^2$ est continue si chacune des fonctions coordonnées l'est
2. ou **courbe géométrique** ou **support de la courbe paramétrée**

Si $t \neq t_0$, la vitesse moyenne du mobile entre les instants $t$ et $t_0$ est donnée par le vecteur $\frac{1}{t-t_0}\overrightarrow{M(t_0)M(t)}$ qui a $(\frac{x(t)-x(t_0)}{t-t_0}, \frac{y(t)-y(t_0)}{t-t_0})$ pour coordonnées dans la base $(\vec{i}, \vec{j})$.

Lorsque $t \to t_0$, ce vecteur vitesse moyenne tend, au sens ou chacune de ses coordonnées tend, vers $\vec{v}(t_0) := (x'(t_0), y'(t_0))$. Ce vecteur est appelée le vecteur **vitesse instantanée** ou **vecteur vitesse** de $M$ à l'instant $t_0$.

Lorsque $\vec{v}(t_0) \neq 0$, la droite passant par le point $M(t_0)$, de vecteur directeur $\vec{v}(t_0)$, donnée paramétriquement par

$$x_T(s) = x(t_0) + (s - t_0)x'(t_0),\ y_T(s) = y(t_0) + (s - t_0)y'(t_0),$$

est appelée la **tangente** à $\Gamma$ au point $M(t_0)$, à l'instant $t_0$.

Une équation cartésienne de la tangente à la courbe paramétrée à l'instant $t_0$ est donc

$$(X - x(t_0)).y'(t_0) - (Y - y(t_0)).x'(t_0) = 0 \text{ d'inconnue } (X, Y) \in \mathbb{R}^2$$

Dans le cas du paramétrage naturel d'un graphe de fonction, on retrouve la notion de tangente à un graphe déjà rencontrée.

*Exemples et remarques* 6.1.2.

1. Si $\phi$ est une bijection $C^1$ d'un intervalle $J \subset \mathbb{R}$ sur $I \subset \mathbb{R}$, $(x(t), y(t))$, $t \in I \subset \mathbb{R}$ et $(\tilde{x}(s), \tilde{y}(s)) = (x(\phi(s)), y(\phi(s)))$, $s \in J$ sont alors deux paramétrages d'une même courbe $\Gamma$.

   On a alors, par le théorème de dérivation des fonctions composées, si $t_0 = \phi(s_0)$, que

   $$\vec{v}(t_0) = \phi'(s).\vec{\tilde{v}}(s_0).$$

   Si $\phi'(s_0) \neq 0$, la tangente à $\Gamma$ en $M(t_0)$ à l'instant $t_0$ est alors égale à la tangente à $\Gamma$ en $\tilde{M}(s_0)$ à l'instant $s_0$. Cela donne un début de définition intrinsèque, i.e. indépendant du paramétrage choisi, de la tangente à une courbe paramétrée.

2. Il se peut que $\vec{v}(t_0) = 0$ mais que le support de la courbe admette quand même une tangente au point $t = t_0$. Juste pour faire un exemple, cela est le cas quand la limite de $\frac{1}{(t-t_0)^2}\overrightarrow{M(t_0)M(t)}$ lorsque $t \to t_0$ existe et est non nulle. Par extension, on dira aussi que cette limite est un vecteur tangent à la courbe à l'instant $t_0$.

### 6.1.2 Distance parcourue entre deux instants

**Définition 6.1.2.** Soit $\gamma(t) = (x(t), y(t)), t \in I$ une courbe paramétrée de classe $C^1$ sur un intervalle $I$. Soit $v(t) = \|\gamma'(t)\| = \sqrt{x'^2(t) + y'^2(t)}$. Soient $t_0, t_1 \in I$, $t_0 \leq t_1$. On définit la **distance** parcourue entre les instants $t_0$ et $t_1$ par la formule

$$D(t_0, t_1) = \int_{t_0}^{t_1} v(t)\ dt$$

*Exemples et remarques* 6.1.3.

1. La quantité $v(t)$ est la longueur du vecteur vitesse à l'instant $t$. C'est ce que l'on appelle la **vitesse instantanée**. Si on imagine une voiture suivant la courbe $\gamma(t), t \in I$, $v(t)$ est ce que l'on peut lire sur le compteur de vitesse de cette voiture. D'une certaine manière, on "oublie" l'information de direction contenue dans le vecteur vitesse $\vec{v}(t)$. Si l'on pense qu'une intégrale est un genre de somme, la quantité $\frac{1}{t_1-t_0}\int_{t_0}^{t_1} v(t)\ dt$ est en quelque sorte la "moyenne" des vitesses instantanées sur l'intervalle de temps $[t_0, t_1]$ : il s'agit donc de la vitesse moyenne sur cet intervalle de temps, i.e la distance parcourue par le mobile entre les instants $t_0$ et $t_1$ divisée par la durée $t_0 - t_1$ de cet intervalle de temps.

2. Prenons l'exemple d'un segment géométrique $L = [AB]$ où $A$ a pour coordonnées $(x_A, y_A)$, $B$ a pour coordonnées $(x_B, y_B)$. Ce segment est naturellement paramétré par

$$\gamma(t) = ((1-t)x_A + tx_B, (1-t)y_A + ty_B),\ t \in [0,1]$$

On a

$$\vec{v}(t) = (x_B - x_A, y_B - y_A),\ v(t) = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2} = AB$$

D'après la définition, la distance parcourue entre les instants $t_0 = 0$ et $t_1 = 1$ est donc

$$D(0,1) = \int_0^1 \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}\ dt = AB$$

Dans ce cas, on voit donc que la distance parcourue entre ces deux instants est exactement la longueur du segment, ce qui est finalement rassurant !!

Un calcul similaire montre que la distance parcourue entre l'instant 0 et l'instant $t \in [0,1]$ quelconque est $D(0,t) = t.AB$, ce qui est exactement la longueur du segment $A\gamma(t)$.

3. Traitons maintenant l'exemple d'un cercle ou d'une partie de cercle. Nous allons voir que les distances calculées à l'aide de deux paramétrages distincts ont au bout du compte la même valeur

   (a) Paramétrage trigonométrique du cercle. Soit $\mathcal{C}$ le cercle de centre $(0,0)$, de rayon 1. Soit $\gamma_0(t) = (x(t), y(t)) = (\cos t, \sin t)$, $t \in [0, 2\pi]$ le paramétrage usuel de ce cercle. On a

   $$\vec{v_0}(t) = (-\sin t, \cos t),\ v_0(t) = \sqrt{\cos^2(t) + \sin^2(t)} = 1$$

   La distance parcourue par le mobile $\gamma_0(t)$ entre les instants 0 et $\frac{\pi}{4}$ est donc

   $$D_0(0, \frac{\pi}{4}) = \int_0^{\frac{\pi}{4}} 1\ dt = \frac{\pi}{4}$$

   La trajectoire de $\gamma_0(t)$ sur l'intervalle $[0, \frac{\pi}{4}]$ est un huitième du cercle $\mathcal{C}$, sa longueur au sens usuel du terme est donc $\frac{\pi}{4}$, ce qui coincide bien avec la distance parcourue.

   (b) Posons maintenant $\gamma_1(t) = (\cos t^2, \sin t^2)$, $t \in [0, \sqrt{2\pi}]$. Il s'agit d'un autre paramétrage du même cercle $\mathcal{C}$. On a

   $$\vec{v_1}(t) = (-2t \sin t^2, 2t \cos t^2),\ v_1(t) = \sqrt{4t^2 \cos^2(t^2) + 4t^2 \sin^2(t^2)} = 2t$$

   La distance parcourue par le mobile $\gamma_1(t)$ entre les instants 0 et $\sqrt{\frac{\pi}{4}}$ est donc

   $$D_1(0, \sqrt{\frac{\pi}{4}}) = \int_0^{\sqrt{\frac{\pi}{4}}} 2t\ dt = \left[t^2\right]_0^{\sqrt{\frac{\pi}{4}}} = \frac{\pi}{4}$$

   La trajectoire de $\gamma_1(t)$ sur l'intervalle $[0, \sqrt{\frac{\pi}{4}}]$ est le même huitième du cercle $\mathcal{C}$ que précédemment et la encore, sa longueur au sens usuel du terme coincide bien avec la distance parcourue.

La proposition suivante indique, à la manière de l'énoncé similaire sur les tangentes, que la distance parcourue sur un arc paramétré est par nature liée à la trajectoire et non à un paramétrage particulier.

**Proposition 6.1.1.** *Soient $I$ et $J$ sont deux intervalles non triviaux de $\mathbb{R}$, $\sigma : I \to J$, une bijection croissante de classe $C^1$, $\gamma_0 : I \to \mathbb{R}^2$ et $\gamma_1 : J \to \mathbb{R}^2$ sont deux arcs paramétrés de classe $C^1$, de vitesses instantanées respectives $v_0 : I \to \mathbb{R}$ et $v_1 : J \to \mathbb{R}$,*

*Si pour tout $t \in I$, $\gamma_1(\sigma(t)) = \gamma_0(t)$ alors, pour $t_0 < t_1$, $t_0, t_1 \in I$, en posant $\sigma(t_0) = s_0 < s_1 = \sigma(t_1)$, on a*

$$D_0(t_0, t_1) = \int_{t_0}^{t_1} v_0(t)\ dt = \int_{s_0}^{s_1} v_1(s)\ ds = D_1(s_0, s_1)$$

*Démonstration.* Comme pour tout $t \in I$, $\gamma_1(\sigma(t)) = \gamma_0(t)$, en dérivant, on obtient que pour tout $t \in I$, $\sigma'(t).\vec{v}_1(\sigma(t)) = \vec{v}_0(t)$. Comme $\sigma$ est croissante, $\sigma'(t) \geq 0$ et donc, pour tout $t \in I$,

$$\sigma'(t).v_1(\sigma(t)) = v_0(t)$$

On a donc, par le théorème du changement de variable, Théorème **??**, dans les intégrales que

$$\begin{aligned} D_0(t_0, t_1) &= \int_{t_0}^{t_1} v_0(t)\ dt \\ &= \int_{t_0}^{t_1} \sigma'(t).v_1(\sigma(t))\ dt \\ &\overset{s=\sigma(t)}{=} \int_{s_0}^{s_1} v_1(s)\ ds = D_1(s_0, s_1) \end{aligned}$$

□

### 6.1.3 Exemples de tracés - Coordonnées cartésiennes

**Schéma général** Le but est de tracer une courbe paramétrée donnée par le paramétrage $x(t), y(t)$, $t \in I$.

- On regarde le domaine de définition, la continuité, la dérivabilité des fonctions $x$ et $y$
- On regarde si des symétries communes aux fonctions $x$ et $y$ ne permettent pas de réduire le domaine d'étude, quitte à compléter la courbe par des symétries centrales, axiales, des translations...
- On étudie les variations de $x$ et $y$ que l'on regroupe dans un tableau de variations commun à ces deux fonctions avec repérage des limites et de valeurs particulières du paramètre $t$.
- On repère des points particuliers de la courbe et l'on calcule si besoin les tangentes à ces points : intersections avec les axes de coordonnées, points où les tangentes sont parallèles aux axes, points multiples [3]
- Une classe importante de points particuliers sont les points « singuliers » du paramétrage, où la vitesse s'annule. On verra sur des exemples comment étudier la courbe au voisinage de tels points.
- recherche de droites asymptotes
- On trace la courbe en respectant ces contraintes

**Un exemple : $x(t) = 3t^2 - 2$, $y(t) = 3t - t^3$, $t \in \mathbb{R}$.**

1. Les fonctions $x$ et $y$ étant polynomiales, elles sont clairement de classe $C^\infty$ sur $\mathbb{R}$.
2. Symétries : on remarque que pour $t \in \mathbb{R}$, $x(-t) = x(t)$ et $y(-t) = -y(t)$. Cela implique que la courbe est symétrique par rapport à l'axe des abscisses et que, pour la tracer, il suffit de la tracer pour $t \geq 0$, le reste de la courbe s'obtenant par la symétrie axiale.
3. Limites : on a
$$\lim_{t \to \pm\infty} x(t) = \infty \text{ et } \lim_{t \to \pm\infty} y(t) = - \pm \infty$$
4. Dérivées et variations : on a
$$x'(t) = 6t \text{ et } y'(t) = 3(1 - t^2)$$
On obtient donc le tableau de variations suivant sur $[0, +\infty[$

3. Ce sont les points de la courbe géométrique en lesquels on passe au moins deux fois, à deux instants différents

| $t$ | $0$ | | $1$ | | $+\infty$ |
|---|---|---|---|---|---|
| $x'(t)$ | $0$ | $+$ | $6$ | $+$ | |
| $y'(t)$ | $3$ | $+$ | $0$ | $-$ | |
| $x(t)$ | $-2$ | $\nearrow$ | $1$ | $\nearrow$ | $+\infty$ |
| $y(t)$ | $0$ | $\nearrow$ | $2$ | $\searrow$ | $-\infty$ |

Le graphe admet donc une tangente verticale au point $M_0 = (0,0)$ à l'instant $t = 0$ et une tangente horizontale au point $M_1 = (1,2)$ à l'instant $t = 1$.

5. Recherche de points doubles : résolvons le système de deux équations à deux inconnues $t, s \in \mathbb{R}$, $t \neq s$

$$\begin{cases} x(t) = x(s) \\ y(t) = y(s) \end{cases}$$

Les solutions de ce système telles que $t \neq s$ correspondent à un point de la courbe atteint à deux instants distincts. Le système en question est équivalent succesivement à chacun des systèmes suivants

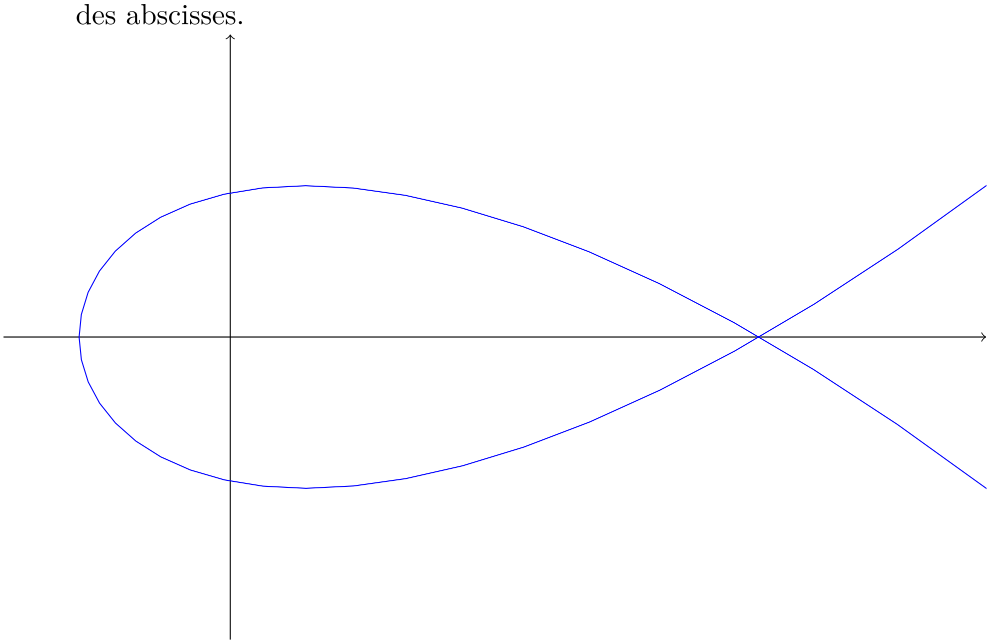
$$\begin{cases} t^2 - s^2 &= 0 \\ 3t - t^3 - (3s - s^3) &= 0 \end{cases}, \quad \begin{cases} s + t &= 0 \\ 3 - (t^2 + st + s^2) &= 0 \end{cases},$$

$$\begin{cases} s &= -t \\ 3 - (t^2 - t^2 + t^2) &= 0 \end{cases}, \quad \begin{cases} s &= -t \\ t &= \pm\sqrt{3} \end{cases}$$

Il y a donc un « point double » sur la courbe, atteint aux instants $t = \pm\sqrt{3}$, il s'agit du point $M_{\sqrt{3}} = (7,0)$. La tangente $T$ en ce point à l'instant $t = \sqrt{3}$ a pour équation paramétrique

$$\begin{cases} x_T(t) = x(\sqrt{3}) + tx'(\sqrt{3}) &= 7 + 6\sqrt{3}.t \\ y_T(t) = y(\sqrt{3}) + ty'(\sqrt{3}) &= -6t \end{cases}$$

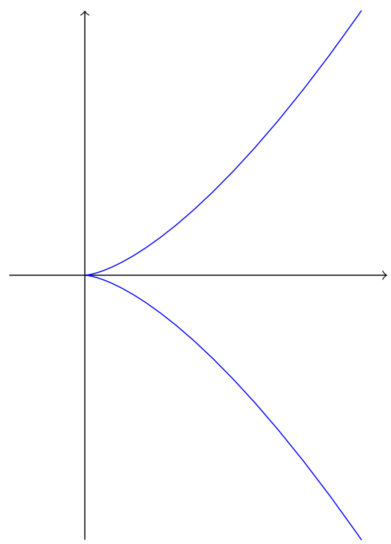
La tangente $S$ en ce point à l'instant $t = -\sqrt{3}$ est la symétrique de $T$ par rapport à l'axe des abscisses.



### 6.1.4 $x(t) = t^2$, $y(t) = t^3$, $t \in \mathbb{R}$.

1. On a $x(-t) = x(t)$, $y(-t) = -y(t)$, la courbe présente donc une symétrie axiale par rapport à l'axe des abscisses et il suffit de l'étudier sur $[0, +\infty[$ puis de faire agir cette symétrie.
2. $x'(t) = 2t$ et $y'(t) = 3t^2$. Sur $[0, +\infty[$ les deux fonctions $x$ et $y$ sont donc strictement croissantes. La vitesse $v(t)$ est nulle si et seulement si $t = 0$.
3. (La remarque qui tue). On a, pour tout $t \in \mathbb{R}$, $t = \sqrt[3]{y(t)}$ et donc $x(t) = \sqrt[3]{y(t)}^2$. Cela signifie, par exemple, que la courbe $\Gamma$, restreinte aux instants $t \geq 0$ est le graphe de la fonction $y = x^{\frac{3}{2}}$. Ce graphe est bien connu ! On note par exemple la présence d'une demi-tangente verticale en $x = 0, y = 0$.

L'étude de ce type de courbe, $x(t) = t^p$, $y(t) = t^q, t$ voisin de 0 avec $p$ et $q$ deux entiers naturels est très importante pour l'étude locale des courbes paramétrées au voisinage d'un **point singulier**. Un point singulier d'une courbe paramétrée est un point atteint à un instant $t_0$ tel que $v(t_0) = 0$. il s'agit donc de ces instants pour lesquels nous ne sommes pas capables de déterminer une droite tangente géométrique.



## 6.2 Courbes paramétrées en coordonnées polaires

### 6.2.1 Courbes en polaires : généralités

#### Passage d'un système de coordonnées à l'autre

Considérons deux fonctions continues $r(t)$ et $\theta(t)$ d'une variable réelle $t$, définies sur un même intervalle $I$ de $\mathbb{R}$.

La courbe paramétrée définie en coordonnées polaires (où $\rho$ représente la distance à l'origine et $\theta$ l'angle, pris en sens trigonométrique, entre la direction du segment qui joint l'origine au point et l'axe horizontal orienté vers la droite) par le système paramétrique

$$\rho = r(t),\ \theta = \theta(t),\ t \in I$$

est par définition la courbe paramétrée définie en coordonnées cartésiennes par le système paramétrique

$$x(t) = r(t)\cos\theta(t),\ y(t) = r(t)\sin\theta(t),\ t \in I$$

Remarquons que si $r(t)$ et $\theta(t)$ sont des fonctions régulières (par exemple de classe $C^1$), sur $I$ alors $x(t)$ et $y(t)$ sont des fonctions ayant au moins le même degré de régularité (dans l'exemple, elles seront au moins de classe $C^1$).

Réciproquement, étant donnée une courbe paramétrée donnée en coordonnées cartésiennes par le système paramétrique $x(t), y(t)$, $t \in I \subset \mathbb{R}$, on peut en donner une définition polaire par la correspondance coordonnées cartésiennes–coordonnées polaires décrite dans la partie précédente.

Le fait que, par exemple, $r(t) = \sqrt{x^2(t) + y^2(t)}$, montre qu'en général, on ne peut espérer que $r(t)$ soit automatiquement de classe $C^1$ si $x(t)$ et $y(t)$ le sont. Les instants $t \in D$ où $x(t) = y(t) = 0$ posent en effet problème lors de la dérivation.

Le problème de la régularité est encore plus prégnant lors de la détermination de l'angle $\theta(t)$. D'une part, lorsque $r(t) = 0$, l'angle n'est pas défini et d'autre part, si la courbe fait « plus d'un tour autour de 0 », la fonction $\theta(t)$ risque d'exhiber une discontinuité due au problème du choix de l'angle (parce que $2\pi$ et 0 sont le même angle).

### Expression de la vitesse

Posons quelques notations couramment utilisées en physique. Si $\theta \in \mathbb{R}$, on pose [4]

$$u_r = (\cos\theta, \sin\theta),\ u_\theta = (-\sin\theta, \cos\theta)$$

$(u_r, u_\theta)$ forme une base orthonormée et

le vecteur $\overrightarrow{OM}$ est donc égal à $ru_r$. En dérivant et en utilisant les règles de dérivation d'un produit et d'une composée, on obtient que

$$v = r'.u_r + r\theta'.u_\theta$$

Le carré de la vitesse est donc

$$|v|^2 = r'^2 + r^2\theta'^2$$

### Le rayon dépend de l'angle

Nous allons concentrer nos efforts dans la suite sur les courbes paramétrées données par une équation de la forme

$$\rho = r(\theta),\ \theta \in I$$

où $r$ est une fonction régulière définie sur un intervalle $I$ de $\mathbb{R}$. Par ceci, on entend les courbes définies en coordonnées polaires par le système paramétrique

$$\rho = r(t),\ \theta = t,\ t \in I$$

L'exemple le plus simple est donné par le cercle de centre 0, de rayon $r_0$. Il est défini par l'équation paramétrique

$$r = r_0,\ \theta \in \mathbb{R}$$

.

En restreignant le domaine de variation de $\theta$, on obtient une partie du cercle. Par exemple, si $\theta_0 < \theta_1$, la courbe définie par l'équation paramétrique

$$r = r_0,\ \theta \in [\theta_0, \theta_1]$$

4. les $r$, $\theta$ en indice n'ont rien à voir avec les fonctions ou variables portant ces noms, les vecteurs $u_r$ et $u_\theta$ dépendent seulement de l'angle $\theta$

est l'arc du cercle de centre 0, de rayon $r_0$ constitué des points dont l'argument varie entre $\theta_0$ et $\theta_1$.

Les droites ne passant pas par 0 ont une équation polaire du type

$$r = \frac{a}{\cos(\theta - \theta_0)}$$

Si l'on note $H$ le projeté orthogonal de $O$ sur une telle droite. On a $OH = a$ et l'angle $(\vec{i}, \overrightarrow{OH})$ est $\theta_0$.

### 6.2.2 Un exemple : la cardioïde : $r = 2(1 - \cos\theta)$.

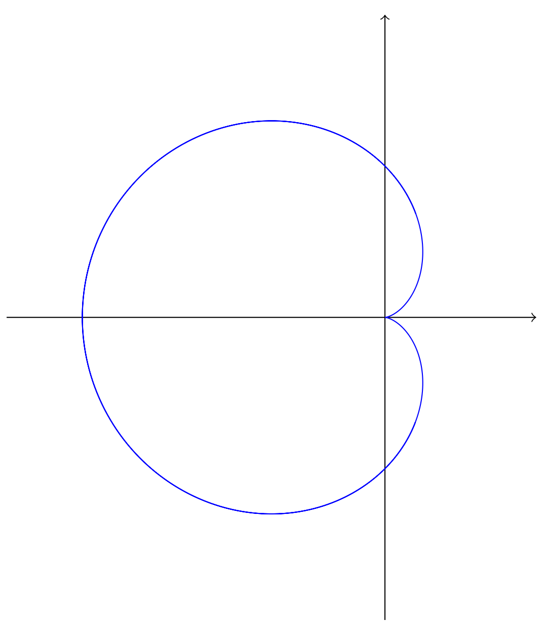
1. $r$ est une fonction de classe $C^\infty$ sur $\mathbb{R}$. Elle s'annule exactement aux point $\theta$ tels qu'il existe une entier relatif $k$ tel que $\theta = k2\pi$.
2. Symétries. On a $r(\theta + 2\pi) = r(\theta)$. Il suffit donc d'étudier la courbe sur un intervalle de longueur $2\pi$ pour obtenir toute la courbe géométrique. On a par ailleurs $r(-\theta) = r(\theta)$, ce qui démontre que la courbe est symétrique par rapport à l'axe des abscisses. Il suffit d'étudier $r$ sur l'intervalle $[0, \pi]$ pour reconstituer la courbe en faisant agir la symétrie.
3. On a $r'(\theta) = 2\sin\theta$. $r'$ est donc strictement positif sur $]0, \pi[$, $r$ est donc strictement croissant sur $[0, \pi]$, $r$ y croit de 0 jusque 4. Le carré de la vitesse est $r(\theta)^2 + r'(\theta)^2 = 8(1 - \cos\theta)$. La vitesse s'annule donc uniquement lorsque $\theta = 0$.
4. Pour étudier la situation au voisinage de $\theta$, revenons en coordonnées cartésiennes et effectuons un développement limité de chacune des coordonnées : on a

$$\begin{aligned} x(\theta) &= 2(1 - \cos\theta)\cos\theta = \theta^2 + o(\theta^2) \\ y(\theta) &= 2\sin\theta(1 - \cos\theta) = \sin 2\theta = \theta^3 + o(\theta^3) \end{aligned}$$

On admet qu'au voisinage de $\theta = 0$, la courbe « ressemble » [5] à la courbe donnée par les premiers termes non triviaux de chacun des développements limités :

$$\begin{aligned} x(\theta) &= \theta^2 \\ y(\theta) &= \theta^3 \end{aligned}$$

Cette courbe a été étudiée dans la partie précédente. Elle présente en 0 une demi-tangente horizontale et un point de rebroussement.



5. par exemple au niveau du positionnement de la courbe par rapport à sa tangente ou demi-tangente

# Chapitre 7

# Fonctions réelles de deux variables

Dans ce chapitre aussi, le plan est rapporté à un repère orthonormé $(O, \vec{i}, \vec{j})$. On l'identifie par ce moyen à $\mathbb{R}^2$.

## 7.1 Généralités

### 7.1.1 Fonctions, ensemble de définition

Soit $D$ une partie de $\mathbb{R}^2$. Une fonction $f : D \to \mathbb{R}$ est un moyen d'associer à tout couple de réels $(x, y)$ de l'ensemble de définition $D$ de $f$ un nombre réel unique noté $f(x, y)$.

*Exemples et remarques* 7.1.1.

1. Les fonctions constantes. Soit $a$ un réel fixé. La formule $f(x, y) = a$ définit une fonction $f : \mathbb{R}^2 \to \mathbb{R}$.
2. Les fonctions affines. Soient $a, b, c$ trois réels fixés, la formule $f(x, y) = a.x + b.y + c$ définit une fonction $f : \mathbb{R}^2 \to \mathbb{R}$. Les fonctions de cette forme sont appelées fonctions affines.
3. Les fonctions quadratiques. Soit $a_{ij}$ une famille de nombres réels fixés. La formule
$$f(x, y) = a_{20}x^2 + a_{02}y^2 + a_{11}xy + a_{10}x + a_{01}y + a_{00}$$
définit là encore une fonction $f : \mathbb{R}^2 \to \mathbb{R}$.
4. Fonctions polynomiales. Elles sont définies par une formule du type
$$f(x, y) = \sum_{i,j \geq 0}^{\text{finie}} a_{ij} x^i . y^j$$
5. Fractions rationnelles : elles sont définies comme étant le quotient de deux fonctions polynomiales, par une formule du type
$$f(x, y) = \frac{P(x, y)}{Q(x, y)}$$
Cette formule ne définit une fonction $f : D \to \mathbb{R}$ que sur une partie du plan $D$ sur laquelle $Q$ ne s'annule pas, par exemple
$$D = \{(x, y) \in \mathbb{R}^2,\ Q(x, y) \neq 0\} = \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2,\ Q(x, y) = 0\}$$
   (a) Supposons que $Q(x, y) = 2x + 3y - 2$. Le domaine de définition naturel de la fonction $f$ définie par la formule
$$f(x, y) = \frac{x^2 + 3xy - 4}{2x + 3y - 2}$$

est l'ensemble $D$ des points $(x, y)$ tels que $2x + 3y - 2 \neq 0$. C'est la réunion des deux demi-plans ouverts

$$P_+ = \{(x, y) \in \mathbb{R}^2,\ 2x + 3y - 2 > 0\} \text{ et } P_- = \{(x, y) \in \mathbb{R}^2,\ 2x + 3y - 2 < 0\}$$

(b) Supposons que $Q(x, y) = x^2 + y^2 - 1$. La formule

$$f(x, y) = \frac{x^2 - 17xy + 12}{x^2 + y^2 - 1}$$

définit une fonction $f : D \to \mathbb{R}$ avec

$$D = \{(x, y) \in \mathbb{R}^2,\ Q(x, y) \neq 0\} = \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2,\ Q(x, y) = 0\}$$

Géométriquement $D$ est le plan, privé du cercle de centre 0 et de rayon 1. Ce cercle est exactement l'ensemble

$$C = \{(x, y) \in \mathbb{R}^2,\ x^2 + y^2 - 1 = 0\}$$

6. On peut évidemment faire intervenir d'autres fonctions élémentaires d'une variable réelle.
   (a) La formule $f(x, y) = \ln(2x^2 - 4x + 2y^2 - 6)$ définit une fonction $f : D \to \mathbb{R}$ si l'on prend
   $$D = \{(x, y) \in \mathbb{R}^2,\ 2x^2 - 4x + 2y^2 > 0\}$$
   Comme $2x^2 - 4x + 2y^2 - 6 = 2(x-1)^2 + y^2 - 8$, l'ensemble $D$ est donc géométriquement la partie à l'extérieur du boule de centre $(1, 0)$, de rayon 2.
   (b) La formule $f(x, y) = \arcsin(x + 2y - 3)$ définit une fonction $f : D \to \mathbb{R}$ sur le domaine $D$ du plan définit par
   $$D = \{(x, y) \in \mathbb{R}^2,\ -1 \leq x + 2y - 3 \leq 1\}$$
   Il s'agit de la partie de plan comprise, au sens large, entre les deux droites d'équation $x + 2y = 2$ et $x + 2y = 4$.

On voit de ces exemples que la détermination du domaine de définition d'une fonction de deux variables réelles associée à une formule $f(x, y) = ...$ nécessite la détermination de parties du plan définies par des équations ou des inéquations cartésiennes du type

$$\{(x, y) \in \mathbb{R}^2, g(x, y) \text{ est défini et } g(x, y) = 0\} \text{ ou } \{(x, y) \in \mathbb{R}^2, g(x, y) \text{ est défini et } g(x, y) > 0\}$$

pour certaines (autres) fonctions de deux variables $g$.

### 7.1.2 Compositions

Soit $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables réelles à valeurs réelles. Avec les types de fonctions que nous avons déjà rencontrés, nous pouvons envisager les compositions suivantes :

1. Composition à gauche avec une fonction réelle de variable réelle. Soit $g : I \to \mathbb{R}$ une telle fonction définie sur $I$ une partie de $\mathbb{R}$. La formule $(g \circ f)(x, y) = g(f(x, y))$ définit la fonction $g \circ f$ sur l'ensemble
   $$\{(x, y) \in D,\ f(x, y) \in I\}$$
   La fonction $g \circ f$ est définie sur $D$ à partir du moment où $f(x, y) \in I$ pour toute valeur de $(x, y) \in D$.
2. Composition à droite avec un arc paramétré. Soit $g : I \to \mathbb{R}^2$ une fonction définie sur une certaine partie $I$ de $\mathbb{R}$. La formule $(f \circ g)(t) = f(g(t))$ définit la fonction $f \circ g$ sur la partie de $\mathbb{R}$
   $$\{t \in I,\ g(t) \in D\}$$
   La fonction $f \circ g$ est définie sur tout $I$ à partir du moment où $g(t) \in D$ pour tout $t \in I$.

### 7.1.3 Représentations graphiques

#### Graphe

**Définition 7.1.1.** Le graphe d'une fonction $f : D \to \mathbb{R}^2$ est la partie de $\mathbb{R}^3$ définie par

$$\{(x, y, f(x, y)),\ (x, y) \in D\} = \{(x, y, z) \in \mathbb{R}^3,\ (x, y) \in D \text{ et } z = f(x, y)\}$$

On peut représenter le graphe d'une telle fonction $f$ comme une surface dans l'espace. On n'insistera pas ce semestre sur ce type de représentation.

#### Lignes de niveaux

**Définition 7.1.2.** Si $f : D \to \mathbb{R}$ est une fonction de deux variables réelles, $t \in \mathbb{R}$, la ligne de niveau $t$ est l'ensemble

$$L_t = \{(x, y) \in D,\ f(x, y) = t\}$$

On peut représenter efficacement une fonction de deux variables en dessinant schématiquement quelques lignes de niveaux pour des valeurs de $t$ bien choisies.

## 7.2 Limites et continuité

### 7.2.1 Boules, voisinages, parties ouvertes et fermées

Nous allons définir la notion de limite de $f(x, y)$ lorsque le couple $z = (x, y)$ tend vers un certain couple fixé $z_0 = (x_0, y_0)$. Pour cela nous allons copier la définition que nous avions en une variable réelle.

Les boules ouvertes $B(z_0, \epsilon)$ de centre $z_0$, de rayon $\epsilon$ vont jouer en deux variable un rôle identique à celui que jouait en une variable l'intervalle ouvert $]x_0 - \epsilon, x_0 + \epsilon[$, centré en $x_0$, de rayon $\epsilon$.

Si $z = (x, y)$, on pose $|z| = \sqrt{x^2 + y^2}$, la distance de $z$ à 0. On lira $|z|$, le **module** de $z$, ou sa **norme**. On remarque que si $z = (x, y)$, $z_0 = (x_0, y_0)$, alors

$$|z - z_0| = \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

est la distance entre $z$ et $z_0$.

Le boule ouverte $B(z_0, \epsilon)$ de centre $z_0 = (x_0, y_0)$, de rayon $\epsilon$ est

$$B(z_0, \epsilon) = \{z = (x, y) \in \mathbb{R}^2,\ |z - z_0| < \epsilon\}$$

Soit $z_0$ un point, une partie $V$ de $\mathbb{R}^2$ est dite **voisinage de** $z_0$ s'il existe $\epsilon > 0$ tel que

$$B(z_0, \epsilon) \subset V$$

*Exemples et remarques* 7.2.1.

1. Un demi-plan ouvert $P$ est voisinage de chacun de ses points. Un tel demi-plan est défini par trois nombres $a, b, c$, $(a, b) \neq (0, 0)$ par

   $$P = \{(x, y) \in \mathbb{R}^2,\ a.x + b.y + c < 0\}$$

2. Une boule ouverte est voisinage de chacun de ses points.
3. Si $V$ est l'intersection d'un nombre fini de demi-plans ouverts ou de boules ouverts alors $V$ est voisinage de chacun de ses points.

Nous reprenons maintenant les définitions données au Chapitre 1.

**Définition 7.2.1.** Soit $A \subset \mathbb{R}^2$ un ensemble quelconque et $z_0 \in \mathbb{R}^2$. On dit que $z_0$ est adhérent à $A$ si pour tout voisinage $V$ de $z_0$ il y a un point $z \in A \cap V$.

**Définition 7.2.2.** Soit $A$ un sous-ensemble de $\mathbb{R}^2$. On dit que $A$ est *ouvert* si il est un voisinage de tous ses points : autrement, dit, si pour tout $z \in A$ il existe un rayon $r > 0$ tel que $B(z, r) \subset A$.

On dit que $A$ est *fermé* si son complémentaire est ouvert.

On notera encore $\bar{A}$ l'ensemble des points d'adhérence de $A$ :

$$\bar{A} = \{z \in \mathbb{R}^2 \,:\, \forall V \text{ voisinage de } z \text{ on a } V \cap A \neq \emptyset\}.$$

### 7.2.2 Définition

**Définition 7.2.3.** Soit $z_0 \in \mathbb{R}^2$, $f : D \to \mathbb{R}$ une fonction de deux variables, $\ell \in \mathbb{R}$. On suppose $z_0 \in \bar{D}$.

1. On dit que $f$ a pour limite $\ell$ en $z_0$, ce que l'on note
$$\lim_{z \to z_0} f(z) = \ell$$
si, pour tout voisinage $U$ de $\ell$ dans $\mathbb{R}$, il existe un voisinage $V$ de $z_0$, tel que si $z \in V \cap D$, alors $f(z) \in U$.
2. Si $f$ est définie en $z_0$, on dit que $f$ est continue en $z_0$ si
$$\lim_{z \to z_0} f(z) = f(z_0).$$
3. $f$ est dite continue si elle est continue en tout point de son domaine $D$.

**Proposition 7.2.1** (DL d'ordre 0)**.** *Dans le même contexte que précédemment, $f$ admet pour limite $\ell$ en $z_0$ si et seulement si il existe une fonction $\epsilon$, définie au voisinage de $(0, 0)$, de limite $0$ en $0$ telle que, pour tout $z$ voisin de $z_0$,*

$$f(z) = \ell + \epsilon(z - z_0)$$

**Remarques**

1. On a, comme dans le cas d'une variable, unicité de la limite de $f$ en $z_0$ sous réserve que cette limite existe.
2. Si $f$ admet une limite en un point $z_0$, la fonction $f$ est bornée sur un certain voisinage de $z_0$.
3. Pour prouver qu'une fonction $\epsilon(h, k)$ a pour limite $0$ en $0$, il suffit de démontrer une majoration du type
$$|\epsilon(h, k)| \leq \eta(\sqrt{h^2 + k^2})$$
pour une certaine fonction $\eta$ d'une variable réelle, de limite $0$ en $0$.
4. Dans le cas de deux variables, nous n'avons pas de notion de limite à droite ou à gauche en un point car cela n'a pas grand sens. On peut en revanche définir la limite en un point suivant une partie $A$ de $\mathbb{R}^2$ comme suit.

**Définition 7.2.4.** Soit $f : D \to \mathbb{R}$ une fonction, $z_0 \in \mathbb{R}^2$, $A \subset D$, $\ell \in \mathbb{R}$. On suppose $z_0 \in \overline{(D \cap A)}$. On dit que $f$ a pour limite $\ell$ en $z_0$ **en suivant** $A$, ce que l'on note

$$\lim_{\substack{z \to z_0 \\ z \in A}} f(z) = \ell \quad \text{ou} \quad f(z) \overset{\substack{z \to z_0 \\ z \in A}}{\to} \ell$$

si, pour tout voisinage $U$ de $\ell$ dans $\mathbb{R}$, il existe un voisinage $V$ de $z_0$, tel que si $z \in V \cap A \cap D$, alors $f(z) \in U$.

*Exemples et remarques* 7.2.2. 1. Les demi-plans ouverts, les boules ouverts sont des parties ouvertes.

2. Une intersection d'un nombre fini de parties ouvertes est ouverte.
3. Une réunion de parties ouvertes est ouverte.
4. Pour démontrer qu'une partie $D$ de $\mathbb{R}^2$ est ouverte, il suffit de trouver une fonction continue $g : \mathbb{R}^2 \to \mathbb{R}$ telle que
$$D = \{(x, y) \in \mathbb{R}^2,\ g(x, y) > 0\}.$$
Disques et demi-plans ouverts rentrent dans ce cadre.

### 7.2.3 Opérations

Les règles sur les opérations et les limites dans ce cadre sont les mêmes que les règles que nous avons énoncées pour le cas des fonctions d'une variable réelle. Nous nous focalisons sur le cas des fonctions continues sur un ensemble ouvert.

#### Opérations algébriques

**Proposition 7.2.2.** *Soient $D$ un domaine ouvert, $f, g : D \to \mathbb{R}$, deux fonctions continues alors*

1. *$f + g$, $f.g$ sont continues sur $D$.*
2. *l'ensemble $D' = \{(x, y) \in D,\ g(x, y) \neq 0\}$ est ouvert et la fonction $f/g : D' \to \mathbb{R}$ définie par la formule $(f/g)(x) = \frac{f(x)}{g(x)}$ est continue sur $D'$.*

*Exemples et remarques* 7.2.3. 1. Les fonctions polynomiales sont continues sur $\mathbb{R}^2$.

2. Une fraction rationnelle est continue sur son domaine de définition, qui est ouvert.

#### Composition

Là encore, les résultats d'une variable réelle se transposent : la continuité se comporte bien vis à vis des compositions

**Proposition 7.2.3.** *Soit $D$ une partie ouverte de $\mathbb{R}$, $f : D \to \mathbb{R}$ une fonction continue sur $D$.*

1. *Composition à gauche. Soit $g : I \to \mathbb{R}$ avec $I$ un intervalle de $\mathbb{R}$ une fonction continue sur $I$. Si $f(x, y) \in I$ pour tout $(x, y) \in D$, la fonction $g \circ f$ est une fonction définie, continue sur $D$.*
2. *Composition à droite. Soit $g : I \to \mathbb{R}^2$ avec $I$ un intervalle de $\mathbb{R}$ une fonction continue sur $I$. Si $g(t) \in D$ pour tout $t \in I$, la fonction $f \circ g$ est une fonction définie, continue sur $I$.*

## 7.3 Dérivabilité

### 7.3.1 DL d'ordre 1

En une variable réelle, la notion de développement limité d'ordre 1 en un point donné est intimement liée à la dérivabilité de la fonction en ce point. Le but est de comparer, localement, au voisinage du point considéré la fonction avec une fonction affine. Il en est de même en deux variables.

**Définition 7.3.1.** Soit $f : D \to \mathbb{R}$, $z_0 = (x_0, y_0)$, $D$ voisinage de $z_0$. On dit que $f$ admet un DL d'ordre 1 en $z_0$, ou de façon plus courte, que $f$ est **différentiable en** $z_0$, s'il existe $a, b \in \mathbb{R}$, une fonction $\epsilon$, définie au voisinage de $0 = (0,0)$, de limite 0 en 0 telle que, pour $z = (x, y)$, voisin de $z_0$, on a

$$\begin{aligned} f(x,y) &= f(x_0,y_0) + a(x-x_0) + b.(y-y_0) + |(x-x_0, y-y_0)|\epsilon(x-x_0, y-y_0) \\ f(z) &= f(z_0) + < \begin{pmatrix} a \\ b \end{pmatrix}, z - z_0 > + |z-z_0|\epsilon(z-z_0) \end{aligned}$$

ou encore, telle que pour tout $(h, k)$ suffisamment voisin de $(0, 0)$, on a

$$\begin{aligned} f(x_0+h, y_0+k) &= f(x_0,y_0) + a.h + b.k + |(h,k)|\epsilon(h,k) \\ f(x_0+h, y_0+k) &= f(x_0,y_0) + < \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} h \\ k \end{pmatrix} > + |(h,k)|\epsilon(h,k). \end{aligned}$$

**Notation :** on peut écrire indifféremment $|(h, k)|\epsilon(h, k)$ ou $o(\sqrt{h^2 + k^2})$. Plus en général, dans les DL des fonctions de deux variables, on utilisera $o(\sqrt{(x-x_0)^2 + (y-y_0)^2}^n)$.

Si $f$ admet un DL d'ordre 1 en $z_0$, le vecteur $\begin{pmatrix} a \\ b \end{pmatrix}$ s'appelle le gradient de la fonction $f$ en $z_0$.
On le note $\nabla f(z_0)$, à lire « nabla » de $f$ ou gradient de $f$ en $z_0$.

*Exemples et remarques* 7.3.1.

1. Fonctions affines. Si $f(x, y) = a.x + b.y + c$ alors $f$ admet un DL d'ordre 1 en tout $z_0 = (x_0, y_0) \in \mathbb{R}^2$. On a

$$f(x_0 + k, y_0 + h) = f(x_0, y_0) + a(x - x_0) + b(y - y_0)$$

Le gradient de $f$ en $z_0$ est donc le vecteur indépendant de $z_0$

$$\nabla f(z_0) = \begin{pmatrix} a \\ b \end{pmatrix}$$

2. Fonctions quadratiques. Si $f(x, y) = a_{20}x^2 + a_{11}xy + a_{02}y^2$, on a

$$\begin{aligned} f(x_0+h, y_0+k) - f(x_0,y_0) &= 2a_{20}x_0.h + a_{11}x_0.k + a_{11}y_0.h + 2a_{02}y_0.k + a_{20}h^2 + a_{11}hk + a_{02}k^2 \\ &= 2a_{20}x_0.h + a_{11}x_0.k + a_{11}y_0.h + 2a_{02}y_0.k + |(h,k)|\epsilon(h,k) \end{aligned}$$

car $|a_{20}h^2 + a_{11}hk + a_{02}k^2| \leq C(h^2 + k^2)$ et donc $a_{20}h^2 + a_{11}hk + a_{02}k^2 = \sqrt{h^2 + k^2}\epsilon(h, k)$ où $\epsilon$ est une fonction de limite nulle en 0. Cela démontre que $f$ admet un DL d'ordre 1 en tout $z_0 = (x_0, y_0)$ et que le gradient de $f$ en $z_0$ est

$$\nabla f(x_0, y_0) = \begin{pmatrix} 2a_{20}x_0 + a_{11}y_0 \\ 2a_{02}y_0 + a_{11}x_0 \end{pmatrix}$$

Nous verrons un peu plus loin comment retrouver facilement cette formule grâce au mécanisme des dérivées partielles.

Il est important de remarquer que, comme dans le cas des fonctions dérivables en une variable, les fonctions différentiables sont continues.

**Proposition 7.3.1.** *Soit $f : D \to \mathbb{R}$, $z_0 = (x_0, y_0)$, $D$ voisinage de $z_0$. Si $f$ est* **différentiable en** $z_0$*, alors $f$ est continue en $z_0$.*

*Démonstration.* Il suffit de remarquer que

$$\lim_{(h,k)\to(0,0)} a.h + b.k + |(h,k)|\epsilon(h,k) = 0,$$

ce qui montre que le DL à l'ordre 1 donne également un DL à l'ordre 0 pour la fonction $f$. □

### 7.3.2 Dérivées partielles

Une façon naturelle d'analyser une fonction de deux variables consiste à « geler » une variable et à étudier la fonction d'une variable réelle obtenue en laissant libre l'autre variable.

**Définition 7.3.2** (Applications partielles)**.** Soit $D \subset \mathbb{R}^2$, $f : D \to \mathbb{R}$ une fonction, $(x_0, y_0) \in D$

1. La première application partielle $f_{|y=y_0}$ est définie par la formule $f_{|y=y_0}(x) = f(x, y_0)$. Si $D$ est voisinage de $(x_0, y_0)$ alors le domaine de définition de $f_{|y=y_0}$ est voisinage, dans $\mathbb{R}$, de $x_0$.
2. La deuxième application partielle $f_{|x=x_0}$ est définie par la formule $f_{|x=x_0}(y) = f(x_0, y)$. Si $D$ est voisinage de $(x_0, y_0)$ alors le domaine de définition de $f_{|x=x_0}$ est voisinage, dans $\mathbb{R}$ de $y_0$.

**Remarque**. Il s'agit en fait de regarder les restrictions de $f$ sur les droites respectivement d'équation $x = x_0$ et $y = y_0$. On a pour cela recours à un paramétrage de chacune de ces droites.

**Définition 7.3.3.** Soit $D$ une partie ouverte de $\mathbb{R}^2$, $f : D \to \mathbb{R}$ une fonction, $(x_0, y_0) \in D$

– Si la première application partielle $f_{|y=y_0}$ est dérivable en $x_0$, on appelle dérivée partielle [1] de $f$ par rapport à la première variable en $(x_0, y_0)$ le nombre

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{df_{|y=y_0}}{dx}(x_0)$$

– Si la deuxième application partielle $f_{|x=x_0}$ est dérivable en $y_0$, on appelle dérivée partielle de $f$ par rapport à la deuxième variable en $(x_0, y_0)$ le nombre

$$\frac{\partial f}{\partial y}(x_0, y_0) = \frac{df_{|x=x_0}}{dy}(y_0)$$

*Exemples et remarques* 7.3.2. 1. Soit $f$ définie sur $\mathbb{R}^2$ par $f(x, y) = x^2 + 3xy - y^2$. Soit $(x_0, y_0) \in \mathbb{R}^2$. On a

$$\begin{aligned} f_{y=y_0}(x) &= x^2 + 3y_0.x - y_0^2 \\ f_{x=x_0}(y) &= x_0^2 + 3y.x_0 - y^2 \end{aligned}$$

En dérivant la première égalité, ($y_0$ est un paramètre dans ce cas), on obtient

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{df_{|y=y_0}}{dx}(x_0) = 2x_0 + 3y_0$$

Pour la deuxième égalité, on obtient

$$\frac{\partial f}{\partial y}(x_0, y_0) = \frac{df_{|x=x_0}}{dy}(y_0) = 3x_0 - 2y_0$$

Dans la pratique, on accélère l'écriture en omettant les indices $_0$. Pour obtenir une expression de $\frac{\partial f(x,y)}{\partial x}$, on dérive l'expression de $f(x, y)$ en considérant que $y$ est une constante.

---

1. Dans cette partie, conformément à l'usage en physique et contrairement à l'usage que nous nous sommes fixés jusqu'à présent, les deux variables d'une fonction $f$ vont porter des noms privilégiés. Dans la plupart des exemples suivants, la première variable s'appelera '$x$', la deuxième '$y$'.

Si on définit une fonction $g$ par une formule du type $g(u, v) = ...$ pour $(u, v)$, on considérera que sa première variable s'appelle $u$ et la deuxième s'appelle $v$, on laisse au lecteur le soin d'interpréter correctement les expressions

$$\frac{\partial g}{\partial u} \text{ et } \frac{\partial g}{\partial v}$$

Si nous avions voulu nous conformer à notre usage que les arguments d'une fonction ne portent pas de nom privilégiés, il faudrait, pour noter respectivement les première et deuxième dérivées partielles de $f$, écrire

$$\partial_1 f \text{ et } \partial_2 f$$

Cette notation n'est pas toujours très parlante aussi nous ne l'utiliserons pas.

2. Si $f(x,y) = \ln(1-x^2-2y^2)$ définie sur l'ellipse remplie ouverte $D = \{(x,y) \in \mathbb{R}^2, x^2+2y^2 < 1\}$, on a
$$\frac{\partial f}{\partial x}(x,y) = \frac{-2x}{1-x^2-2y^2}, \ \frac{\partial f}{\partial y}(x,y) = \frac{-4y}{1-x^2-2y^2}$$

**Proposition 7.3.2.** *Soit $f : D \to \mathbb{R}$, $D$ un ouvert de $\mathbb{R}^2$, $z_0 \in D$. Si $f$ est différentiable en $z_0$ alors les deux dérivées partielles de $f$ en $(x_0, y_0)$ existent et*
$$\nabla f(x_0,y_0) = \begin{pmatrix} \frac{\partial f}{\partial x}(x_0,y_0) \\ \frac{\partial f}{\partial y}(x_0,y_0) \end{pmatrix}$$

*Démonstration.* On ne traite que le premier cas. On a
$$\begin{aligned} f_{|y=y_0}(x_0+h) &= f(x_0+h, y_0) \\ &= f(x_0,y_0) + < \nabla f(z_0), \begin{pmatrix} h \\ 0 \end{pmatrix} > + |h|\epsilon(h,0) \\ &= f(x_0,y_0) + h < \nabla f(z_0), \begin{pmatrix} 1 \\ 0 \end{pmatrix} > + |h|\epsilon(h,0) \end{aligned}$$

Cette expression fournit clairement un développement limité d'ordre 1 de $f_{|y=y_0}$ au voisinage de $x_0$. $f_{|y=y_0}$ est donc dérivable en $x_0$ et
$$\frac{df_{|y=y_0}}{dx}(x_0) = \frac{\partial f}{\partial x}(x_0,y_0) =< \nabla f(z_0), \begin{pmatrix} 1 \\ 0 \end{pmatrix} >$$

De la même façon, $f_{|x=x_0}$ est dérivable en $y_0$ et
$$\frac{df_{|x=x_0}}{dy}(y_0) = \frac{\partial f}{\partial y}(x_0,y_0) =< \nabla f(z_0), \begin{pmatrix} 0 \\ 1 \end{pmatrix} > \quad \square$$

Il est important de remarquer que si $f$ est différentiable alors elle admet des dérivées partielles, mais que la réciproque n'est pas vraie.

*Exemple* 7.3.3. Considérons la fonction
$$f(x,y) = \begin{cases} 0 & \text{si } xy = 0, \\ 1 & \text{si } xy \neq 0. \end{cases}$$

Les applications partielles en $(x_0,y_0) = (0,0)$ coïncident avec la fonction nulle, et donc $\frac{\partial f}{\partial x}(0,0)$ et $\frac{\partial f}{\partial y}(0,0)$ existent et valent 0. Pourtant, $f$ n'est pas différentiable en $(0,0)$, puisqu'elle n'y est pas continue.

Par contre, il est toujours vrai que, en cas d'existence des dérivées partielles, le seul gradient possible de la fonction $f$ est le vecteur composé par ses dérivées partielles.

**Attention aux limites partielles !** Nous pourrions être tentés d'étudier les fonctions de deux variables en se basant uniquement sur les applications partielles, pour se réduire au cas (plus simple) de fonctions d'une variable. Cela est très dangereux !

Nous venons de voir que les dérivées partielles ne sont pas suffisantes à garantir l'existence d'un gradient. Plus en général,
$$\lim_{x \to x_0} f_{|y=y_0}(x) = \lim_{y \to y_0} f_{|x=x_0}(y) = \ell$$

n'implique pas $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = \ell$. Pour s'en convaincre, il suffit de regarder l'exemple 7.3.3.

Non seulement, les limites sur *toutes* les droites passant par $(x_0, y_0)$ ne sont pas suffisantes non plus. Considérons

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

Soit $A$ une droite passante par $(0,0)$ (c'est-à-dire que soit $A = \{(x,y) : x = 0\}$, soit $A = \{(x,y) : y = kx\}$ pour une certaine valeur $k \in \mathbb{R}$). Il est facile de vérifier que

$$\lim_{(x,y)\to(0,0),(x,y)\in A} f(x,y) = 0.$$

Cela est vrai car toute droite n'intersecte l'ensemble $\{(x,y) : f(x,y) = 1\}$ qu'une seule fois et donc, au voisinage de $(0,0)$, $f$ est nulle sur $A$. Pourtant, la limite $\lim_{(x,y)\to(0,0)} f(x,y)$ n'existe pas parce que dans tout voisinage de $(0,0)$ la fonction $f$ prend la valeur 0 aussi bien que la valeur 1.

L'étude des applications partielles peut au contraire être utile pour nier la continuité d'un fonction, sa différentiabilité, l'existence d'une limite...

### 7.3.3 Fonctions de classe $C^1$ sur un ouvert du plan

Nous introduisons maintenant une classe de fonctions définies sur une partie ouverte $D$ de $\mathbb{R}^2$ dont la différentiabilité en chaque point de $D$ sera facilement vérifiable.

**Définition 7.3.4.** Soit $D$ une partie ouverte de $D$, on dit que $f : D \to \mathbb{R}$ est de classe $C^1$ sur $D$ (ce que l'on note $f \in C^1(D)$) si

1. $f$ est continue sur $D$
2. Les deux dérivées partielles de $f$ existent en chaque point de $D$
3. Ces fonctions dérivées partielles, $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$, sont continues sur $D$.

Ce qui est important (et au même temps étonnant, vu les recommandations de toute à l'heure) est le résultat suivant.

**Proposition 7.3.3.** *Si $f : D \to \mathbb{R}$ est de classe $C^1$ alors elle est différentiable en tout point de $D$.*

Ensuite, il y a évidemment tous les résultats usuels sur les fonctions de classe $C^1$ et sur leur comportement vis à vis des opérations algébriques, plus précisément

**Proposition 7.3.4.** *Soient $f$ et $g$ deux fonctions de classe $C^1$ sur une partie ouverte $D$ de $\mathbb{R}^2$*

1. *$f + g$ et $f.g$ sont de classe $C^1$ sur $D$*
2. *Si $g$ ne s'annule pas sur $D$, $f/g$ est une fonction de classe $C^1$ sur $D$*

*On a de plus les formules, pour la dérivation partielle par rapport à la première variable*

$$\frac{\partial(f+g)}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}, \quad \frac{\partial(f.g)}{\partial x} = \frac{\partial f}{\partial x}.g + f.\frac{\partial g}{\partial x}, \quad \frac{\partial(f/g)}{\partial x} = \frac{1}{g^2}\left(g\frac{\partial f}{\partial x} - f\frac{\partial g}{\partial x}\right).$$

*Les formules correspondantes pour la dérivation par rapport à la deuxième variable étant similaires.*

### 7.3.4 Dérivées d'ordre 2 et DL d'ordre 2

**Définition 7.3.5.** Soit $D$ une partie ouverte de $D$, on dit que $f : D \to \mathbb{R}$ est de classe $C^2$ sur $D$ (ce que l'on note $f \in C^2(D)$) si

1. $f$ est de classe $C^1$ sur $D$
2. Les deux dérivées partielles de $f$ sont de classe $C^1$ sur $D$.
3. Les dérivées partielles secondes sont définies par

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial \frac{\partial f}{\partial x}}{\partial x}, \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial \frac{\partial f}{\partial x}}{\partial y}, \frac{\partial^2 f}{\partial y^2} = \frac{\partial \frac{\partial f}{\partial y}}{\partial y}, \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial \frac{\partial f}{\partial y}}{\partial x}$$

Ces quatre fonctions sont continues sur $D$.

Une définition similaire existe pour les fonction $C^k$, à savoir on demande à ce que les dérivées partielles existe et soient $C^{k-1}$.

Les fonctions de classe $C^2$ sur un domaine ouvert se comportent bien vis à vis des opérations algébriques (on ne précisera pas les détails). Le phénomène suivant est assez étonnant

**Théorème 7.3.5** (Schwarz)**.** *Si $f$ est de classe $C^2$ sur $D$, les dérivées partielles secondes $\frac{\partial^2 f}{\partial x \partial y}$ et $\frac{\partial^2 f}{\partial y \partial x}$ sont égales sur $D$.*

On dispose enfin d'une formule de Taylor à l'ordre 2.

**Théorème 7.3.6.** *Soit $f$ une fonction de classe $C^2$ sur $D$, $(x_0, y_0) \in D$, il existe alors une fonction $\epsilon$ de limite nulle en $(0,0)$ telle que, pour $(h,k)$ voisin de $(0,0)$*

$$\begin{aligned} f(x_0+h, y_0+k) &= \underbrace{f(x_0,y_0)}_{\text{partie cste.}} + \\ &+ \underbrace{h.\frac{\partial f}{\partial x}(x_0,y_0) + k.\frac{\partial f}{\partial y}(x_0,y_0) +}_{\text{partie linéaire}} \\ &+ \underbrace{\frac{1}{2}\left(h^2.\frac{\partial^2 f}{\partial x^2}(x_0,y_0) + k^2.\frac{\partial^2 f}{\partial y^2}(x_0,y_0) + 2hk.\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)\right) +}_{\text{partie quadratique}} \\ &+ \underbrace{(h^2+k^2)\epsilon(h,k)}_{\text{reste}} \end{aligned}$$

### 7.3.5 Dérivation des fonctions composées

Nous avons traité le cas des compositions à gauche par une fonction de variable réelle. Nous examinons maintenant le cas d'une composition à droite par un arc paramétré.

#### Formules

**Théorème 7.3.7.** *Soit $f : D \to \mathbb{R}$ une fonction de classe $C^1$ sur $D$, une partie ouverte de $\mathbb{R}^2$. Soit $\gamma : I \to \mathbb{R}^2$, $\gamma(t) = (x(t), y(t))$ un arc paramétré de classe $C^1$ sur l'intervalle $I$. Si $\gamma(t) \in D$ pour tout $t \in I$, la fonction $f \circ \gamma$ est alors de classe $C^1$ sur $I$ et l'on a, pour $t \in I$*

$$\begin{aligned} (f \circ \gamma)'(t) &= < \nabla f(\gamma(t))), \gamma'(t) > \\ &= \frac{\partial f}{\partial x}(x(t), y(t)).x'(t) + \frac{\partial f}{\partial y}(x(t), y(t)).y'(t) \end{aligned}$$

*Exemples et remarques* 7.3.4. Si $f$ et $\gamma$ sont de classe $C^2$ alors $f \circ \gamma$ est de classe $C^2$. La formule précédente montre qu'en effet, $(f \circ \gamma)'$ est de classe $C^1$, du fait du même théorème. La formule pour $(f \circ \gamma)''$ est donc

$$(f \circ \gamma)'' = \frac{\partial f}{\partial x}.x'' + \frac{\partial f}{\partial y}.y'' + \frac{\partial^2 f}{\partial x^2}.x'^2 + \frac{\partial^2 f}{\partial y^2}.y'^2 + 2\frac{\partial^2 f}{\partial x \partial y}.x'.y'$$

*Démonstration.* On fixe $t_0 \in I$. On va chercher à faire un DL d'ordre 1 de $f \circ \gamma$ au voisinage de $t_0$. L'hypothèse donne que, $\gamma$ admet un DL d'ordre 1 en $t_0$. On a donc

$$\begin{aligned} x(t_0+\tau) &= x(t_0) + \tau.x'(t_0) + o(\tau) \\ y(t_0+\tau) &= y(t_0) + \tau.y'(t_0) + o(\tau) \\ \gamma(t_0+\tau) &= \gamma(t_0) + \tau.\gamma'(t_0) + o(\tau) \end{aligned}$$

Par ailleurs, $f$ est différentiable en $z_0 = (x_0, y_0) = \gamma(t_0)$ et donc, il existe $\eta$, une fonction de limite nulle en 0 telle que, pour $(h, k)$ voisin de 0,

$$f(x_0+h, y_0+k)) = f(x_0, y_0) + < \nabla f(x_0, y_0), \begin{pmatrix} h \\ k \end{pmatrix} > + o(|(h,k)|)$$

En substituant

$$\begin{aligned} h &= \tau.x'(t_0) + o(\tau) = x(t_0+\tau) - x_0 \\ k &= \tau.y'(t_0) + o(\tau) = y(t_0+\tau) - y_0 \end{aligned}$$

dans l'égalité précédente ( cette substitution est légitime car, lorsque $\tau$ est suffisamment voisin de 0, $(h, k)$ est suffisamment voisin de $(0, 0)$), on obtient

$$f(x(t_0+\tau), y(t_0+\tau)) = f(x_0, y_0) + \tau. < \nabla f(x_0, y_0), \gamma'(t_0) > + o(\tau)$$

où $\tilde{\eta}$ tend vers 0 en 0. Ceci montre que $f \circ \gamma$ est dérivable en $t_0$ et la dérivée a la forme annoncée. La formule de cette dérivée montre qu'elle est continue par les théorèmes continuité de composées et de produits. □

**Gradient et tangentes aux lignes de niveau**

Ce que montre le théorème précédent, c'est que si $\gamma(t), t \in I$ est une courbe paramétrée contenue dans une ligne de niveau de $f$ (et donc *a fortiori* si $\gamma$ est un paramétrage d'un morceau de ligne de niveau de $f$), on a alors

$$< \nabla f(\gamma(t)), \gamma'(t) >= 0$$

En d'autres termes, $\nabla f(\gamma(t))$ est orthogonal à la tangente en $t$ à la courbe $\gamma$.

### 7.3.6 Le théorème des fonctions implicites

Dans la section précédente, on a vu que si $\nabla f(x_0, y_0) \neq 0$ et si la ligne de niveau de $f$ passant par $(x_0, y_0)$ est une courbe paramétrée admettant une tangente en ce point alors l'équation de cette tangente est

$$(X - x_0)\frac{\partial f}{\partial x}(x_0, y_0) + (Y - y_0)\frac{\partial f}{\partial y}(x_0, y_0) = 0$$

Le théorème suivant, que nous admettrons, bien qu'il soit démontrable avec les outils dont nous disposons, nous dispense du deuxième type d'hypothèse.

**Théorème 7.3.8** (Fonctions implicites)**.** *Soit $f : D \to \mathbb{R}$ une fonction de classe $C^1$ sur $D$. $(x_0, y_0) \in D$. Si $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$, alors il existe un intervalle $U$, voisinage de $x_0$ dans $\mathbb{R}$, un intervalle $V$, voisinage de $y_0$ dans $\mathbb{R}$, une fonction $\phi : U \to V$ de classe $C^1$ sur $U$ telle que $L_{f(x_0,y_0)} \cap (U \times V)$ est le graphe de la fonction $\phi$.*

*Par ailleurs, si $x \in U$, on a*

$$\phi'(x) = -\frac{\frac{\partial f}{\partial x}(x, \phi(x))}{\frac{\partial f}{\partial y}(x, \phi(x))}$$

*Exemples et remarques* 7.3.5.

1. On a, par définition de $L_{f(x_0,y_0)}$ que $x \in U, y \in V$ et $f(x, y) = f(x_0, y_0)$ équivaut à $y = \phi(x)$. En particulier $y_0 = \phi(x_0)$.
2. Une fois que l'on sait que le graphe de $\phi$ (une fonction de classe $C^1$) est contenu dans la ligne de niveau, la formule de la dérivée est claire : il suffit de dériver par rapport à $x$ la relation
$$f(x, \phi(x)) = \text{Cste}$$
3. On a un énoncé analogue en supposant que $\frac{\partial f}{\partial x}(x_0, y_0) \neq 0$. La ligne de niveau passant par $(x_0, y_0)$ est alors localement le graphe $x = \psi(y)$ d'une certaine fonction $\psi$.
4. Si $\nabla f(x_0, y_0) \neq 0$ alors l'une des deux dérivées partielles est non nulle et, dans tous les cas, $L_{f(x_0,y_0)}$ est au voisinage de $(x_0, y_0)$ une courbe géométrique admettant un paramétrage de classe $C^1$, régulier [2]
5. Si $f$ est de classe $C^2$, la fonction $\phi$ est de classe $C^2$ comme le montre la formule.

## 7.4 Recherche d'extrema locaux et globaux

Comme pour les fonctions d'une variable, on dispose du critère d'annulation de la dérivée pour repérer les extrema locaux d'une fonction de deux variables.

### 7.4.1 Extrema locaux et points critiques

**Définition 7.4.1.** Soit $f : D \to \mathbb{R}$ une fonction définie sur une partie $D$ de $\mathbb{R}^2$. Soit $(x_0, y_0) \in D$.

1. On dit que $f$ admet un **maximum local** en $z_0 = (x_0, y_0)$ s'il existe $B(z_0, \epsilon)$ une boule ouvert centré en $z_0$ tel que pour tout $z \in B(z_0, \epsilon) \cap D$, $f(z) \leq f(z_0)$. La valeur de ce maximum local est $f(z_0)$.
2. $f$ admet un **minimum local** en $z_0 = (x_0, y_0)$ s'il existe $B(z_0, \epsilon)$ une boule ouverte centré en $z_0$ tel que pour tout $z \in B(z_0, \epsilon) \cap D$, $f(z) \geq f(z_0)$. La valeur de ce minimum local est $f(z_0)$.
3. $f$ admet un **extremum local** en $z_0 = (x_0, y_0)$ si elle y admet un maximum local ou un minimum local.

**Théorème 7.4.1.** *Soit $f : D \to \mathbb{R}^2$ une fonction de classe $C^1$ sur $D$. Si $f$ admet un extremum local en $z_0 = (x_0, y_0)$, et si $z_0$ est à l'intérieur de $D$ (c'est-à-dire qu'il existe $\varepsilon > 0$ tel que $B(z_0, \varepsilon) \subset D$) on a alors*

$$\nabla f(z_0) = 0$$

*Démonstration.* La application partielle $f_{|y=y_0}$ est définie, de classe $C^1$ sur un certain voisinage de $x_0$. Sa dérivée s'annule donc en $x_0$ et l'on a donc

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0$$

2. i.e. sans point singulier

Le même raisonnement implique que

$$\frac{\partial f}{\partial y}(x_0, y_0) = 0$$

et finalement, le vecteur $\nabla f(z_0)$ est nul. □

*Exemples et remarques* 7.4.1. Comme en une variable il s'agit d'une condition nécessaire.

Un point $z_0$ tel que $\nabla f(z_0) = 0$ s'appelle un **point critique** de la fonction $f$. Ce que dit le théorème c'est que les extrema locaux d'une fonction de classe $C^1$ sur $D$ sont à rechercher en examinant les points critiques de cette fonction. Une étude supplémentaire au voisinage du point critique est alors nécessaire pour affirmer le caractère de maximum ou de minimum local.

### 7.4.2 Conditions suffisantes à l'ordre 2

Pour les fonctions d'une variable réelle, nous avons à notre disposition le critère suivant (un cas particulier du Théorème 4.3.1) pour conclure à l'extrémalité d'un point critique.

**Proposition 7.4.2.** *Soit $I$ un intervalle ouvert de $\mathbb{R}$, $f : I \to \mathbb{R}$ une fonction de classe $C^2$ sur $I$ et $x_0$ un point critique de $f$ : $f'(x_0) = 0$*

1. *Si $f''(x_0) > 0$, la fonction $f$ admet un minimum local en $x_0$.*
2. *Si $f''(x_0) < 0$, la fonction $f$ admet un maximum local en $x_0$.*

Pour les fonctions de deux variables, on a, par une méthode similaire le résultat suivant

**Théorème 7.4.3.** *Soit $D$ une partie ouverte de $\mathbb{R}^2$, $f : D \to \mathbb{R}$ une fonction de classe $C^2$ sur $D$ et $z_0 = (x_0, y_0)$ un point critique de $f$ : $\nabla f(z_0) = 0$*

1. *Si $\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 - \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} < 0$ et $\frac{\partial^2 f}{\partial x^2} > 0$, la fonction $f$ admet un minimum local en $z_0$.*
2. *Si $\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 - \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} < 0$ et $\frac{\partial^2 f}{\partial x^2} < 0$, la fonction $f$ admet un maximum local en $z_0$.*
3. *Si $\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 - \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} > 0$, la fonction $f$ n'admet pas d'extremum local en $z_0$.*

*Démonstration.* La preuve est basée sur un développement de Taylor de $f$ en $z_0$ à l'ordre 2.

Écrivons le développement de Taylor à l'ordre 2 de $f$ en $z_0 = (x_0, y_0)$.

$$f(x_0 + h, y_0 + k) - f(z_0) = \left(a.h^2 + 2b.hk + c.k^2\right) + o(h^2 + k^2)$$

pour tout $(h, k)$ voisin de 0, avec

$$a = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x_0, y_0),\ b = \frac{1}{2}\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0),\ c = \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(x_0, y_0).$$

Réécrivons cela en coordonnées polaires avec $h = r\cos\theta$, $k = r\sin\theta$, d'où $r^2 = h^2 + k^2$. Nous avons alors

$$f(x_0 + h, y_0 + k) - f(x_0) = r^2\left(a\cos^2\theta + 2b\cos\theta\sin\theta + c\sin^2\theta\right) + o(r^2) = r^2 g(\theta) + o(r^2)$$

Grâce au Lemme 7.4.4 nous avons l'un des cas suivants :

- si $\Delta < 0$ et $a > 0$ on a $g(\theta) \geq a_0 > 0$, ce qui entraine $f(x_0+h, y_0+k) - f(z_0) \geq r^2(a_0 + o(r^2)/r^2)$, et, pour $r$ suffisamment petit, $r^2(a_0 + o(r^2)/r^2) > 0$. Finalement cela démontre que $z_0$ est un minimum local ;
- si $\Delta < 0$ et $a < 0$ on a $g(\theta) \leq a_1 < 0$, ce qui entraine $f(x_0+h, y_0+k) - f(z_0) \geq r^2(a_1 + o(r^2)/r^2)$, et, pour $r$ suffisamment petit, $r^2(a_1 + o(r^2)/r^2) < 0$. Finalement $z_0$ est un maximum local ;

– ou alors, si $\Delta > 0$, la fonction $g(\theta)$ prend des valeurs positives aussi bien que des valeurs négatives. En fixant $\theta_0$ tel que $g(\theta_0) < 0$ et en regardant $f$ sur la droite passant par $z_0$ d'angle $\theta_0$ nous avons un maximum. D'autre part, en fixant $\theta_1$ tel que $g(\theta_1) > 0$ et en regardant $f$ sur la droite passant par $z_0$ d'angle $\theta_1$ nous avons un minimum. Par conséquent, $z_0$ n'est ni un minimum ni un maximum local de $f$.
– si $\Delta = 0$ nous ne pouvons pas déterminer la nature du point $z_0$. □

**Lemme 7.4.4.** *Etant donnés $a, b, c \in \mathbb{R}$, considérons la fonction $g : [0, 2\pi] \to \mathbb{R}$ donnée par*

$$g(\theta) = a\cos^2\theta + 2b\cos\theta\sin\theta + c\sin^2\theta$$

*et posons $\Delta = b^2 - ac$. Alors, si $\Delta < 0$ et $a > 0$, il existe un nombre positif $a_0 > 0$ tel que $g(\theta) \geq a_0$ ; si $\Delta < 0$ et $a < 0$, il existe un nombre négatif $a_1 < 0$ tel que $g(\theta) \leq a_1$ ; si $\Delta > 0$ la fonction $g$ prend des valeurs positives aussi bien que des valeurs négatives.*

*Démonstration.* Commençons par $\Delta < 0$ et $a > 0$. La fonction $g$ étant continue et l'intervalle $[0, 2\pi]$ compact, pour démontrer que $g$ est minorée par $a_0 > 0$ il suffit de démontrer que son minimum, qui existe, est positif. Or, on a $g(0) = g(\pi) = a > 0$. Il suffit de prouver que $g$ ne prend jamais la valeur 0. Cela démontrerait que le minimum doit être positif, car, si $g$ prenait des valeurs négatives, par le TVI, elle prendrait aussi la valeur nulle. Or, si on considère $\theta \neq 0, \pi$, on peut diviser par $\sin\theta$ et on aurait

$$g(\theta) = 0 \Rightarrow aX^2 + 2bX + c = 0, \text{ avec } X = \frac{\cos\theta}{\sin\theta},$$

mais la condition $\Delta < 0$ est bien celle qui garantit que le polynôme $aX^2 + 2bX + c$ n'a pas de racines réelles.

De manière similaire, si $\Delta < 0$ et $a < 0$, alors $g(0) = g(\pi) = a < 0$ et le même raisonnement montre $g(\theta) < 0$ pour tout $\theta$. En considérant le maximum de $g$ on trouve la valeur $a_1 < 0$.

Si $\Delta > 0$ on sait que le polynôme $aX^2 + 2bX + c$ admet une racine réelle $X_0$ et prend des valeurs positives ainsi que des valeurs négatives (il s'agit d'une simple parabole), disons $aX_1^2 + 2bX_1 + c > 0$ et $aX_2^2 + 2bX_2 + c < 0$. en correspondance de $X_1$ et $X_2$ il y a des valeurs $\theta_1$ et $\theta_2$ telles que $X_1 = \cos\theta_1 / \sin\theta_1$ et $X_2 = \cos\theta_2 / \sin\theta_2$, et on a $g(\theta_1) > 0$ ainsi que $g(\theta_2) < 0$. □

### 7.4.3 Extrema absolus

Dans cette section finale et courte, nous allons reprendre des idées qu'on a déjà vues dans le cas des fonctions d'une variable. Tout d'abord cette définition avec le théorème qui en suit.

**Définition 7.4.2.** Un sous-ensemble $D$ de $\mathbb{R}^2$ est dit borné s'il existe un rayon $R$ tel que $D \subset B(0, R)$. Un sous-ensemble de $\mathbb{R}^2$ est dit compact s'il est fermé et borné.

On peut maintenant énoncer le bien connu théorème d'existence des minima et maxima dans sa version multidimensionnelle.

**Théorème 7.4.5** (Weierstrass 2D)**.** *Soit $D$ un sous-ensemble compacte de $\mathbb{R}^2$ et $f : D \to \mathbb{R}$ une fonction continue. Il existe alors un point $z_0 \in D$ tel que $f(z_0) = \min\{f(z) : z \in D\}$ (et, symétriquement, il existe un point $z^0 \in D$ tel que $f(z^0) = \max\{f(z) : z \in D\}$).*

Le lecteur pourra facilement faire des exemples inspirés des exemples dans $\mathbb{R}$ qui montent l'utilité de chaque hypothèse de ce théorème.

Aussi, on pourrait faire un analogue du Théorème 2.2.5, mais on ne le fera pas parce qu'on n'a pas donné une définition de limite en l'infini dans $\mathbb{R}^2$.

**Critère de recherche de minima et maxima absolus :**

Soit $D$ un domaine dans le plan délimitée par une courbe $F$ (la frontière de $D$). On suppose (le théorème des fonctions implicites peut nous aider à le garantir) que $F$ est en effet une courbe paramétrée régulière. Soit, de plus, $f : D \to \mathbb{R}$ une fonction continue, différentiable en tout point de $D \setminus (S \cup F)$, où $S$ est en ensemble fini et, si possible, petit .On peut donc dresser une liste de points de $D$ qui sont candidats à être minima et/ou maxima, et qui est composée de :

– tout point $z_0 \in D \setminus (S \cup F)$ qui satisfasse $\nabla f(z_0) = 0$ (ce qui est souvent réalisé par un petit nombre de points, puisqu'il s'agit d'annuler les deux dérivées partielles, c'est-à-dire résoudre un système avec deux équations et deux inconnues) ;
– les points de $S$ car le critère du gradient ne peut pas s'appliquer à ces points ;
– les points de la frontière $F$ car là non plus on ne peut l'appliquer.

Les points de la dernière catégorie sont a priori infinis, donc trop nombreux pour les considérer tous. Cependant, si $F = \{\gamma(t), t \in [a, b]\}$ ($F$ est l'image d'une courbe paramétrée, ou, de même, si il est la réunion d'un nombre fini de courbes de ce type, ce qui est le cas par exemple pour un polygone), il suffit de considérer la fonction d'une variable $t \mapsto f(\gamma(t))$ et de dresser une liste de candidats à la minimisation relative à cette fonction (comme dans le chapitre 3). Cela réduit normalement la liste à un nombre fini de points.

Ensuite, le théorème de Weierstrass en deux dimensions nous garantit que le minimum et le maximum existent et en plus on est sûr qu'on les trouvera parmi les points qu'on a listé. Il suffit donc de calculer les valeurs $f(z)$ correspondant à tous les points $z$ de la liste et on trouvera le (ou les) point(s) de minimum en prenant ceux qui ont les valeurs les moins élevées et les points de maximums en prenant ceux qui ont les valeurs le plus élevées.

On fera juste remarquer que dans ce cadre bidimensionnel, une approche par tableau de variations n'aurait pas eu un sens.