

Notes de cours de mathématiques

Table des matières

6 Fonctions : dérivabilité	12
6.1 Définition, calcul pratique	12
6.1.1 Définition, interprétation géométrique	12
6.1.2 Propriétés	12
6.1.3 Dérivées et opérations	13
6.2 Résultats	13
6.2.1 Monotonie	13
6.2.2 Extrema	14
7 Intégration	15
7.1 Primitives	15
7.2 Intégrales	15
7.3 Interprétation géométrique	16
7.4 Quelques techniques de calcul	16
7.4.1 Changement de variable	16
7.4.2 Intégration par parties	16
7.5 Intégration numérique	17
7.5.1 Méthodes des rectangles	17
7.5.2 Méthodes des trapèzes	17
8 Équations Différentielles	18
8.1 Équations différentielles linéaires d'ordre 1	18
8.1.1 Définitions	18
8.1.2 Méthode de résolution	18
8.2 Résolution numérique des équations linéaires d'ordre 1	20
1 Probabilités	1
1.1 Quelques rappels	1
1.2 Probabilités conditionnelles	1
1.2.1 Indépendance	1
1.2.2 La formule des probabilités totales	2
1.3 Variables aléatoires discrètes	2
1.3.1 Fonction de répartition	2
1.3.2 Loi de probabilité	2
1.3.3 Moyenne et Variance	2
1.3.4 Combinaison linéaire de variables aléatoires indépendantes	3
1.3.5 Lois usuelles discrètes	3
2 Initiation aux tests du χ^2	5
3 Fonctions : Généralités	6
3.1 Domaine de définition	6
3.2 Composition des fonctions f et g	6
3.3 Représentation graphique	6
3.4 Sens de variation	6
3.5 Fonction réciproque	7
4 Deux fonctions classiques : exp et ln	8
4.1 L'exponentielle	8
4.1.1 Propriétés classiques	8
4.1.2 Graphes	8
4.2 Le Logarithme Népérien	8
4.2.1 Propriétés classiques	8
4.2.2 Graphique	8
4.3 Relations exp / ln	9
5 Fonctions : limites et continuité	10
5.1 Limites d'une fonction	10
5.1.1 Limites classiques	10
5.1.2 Comparaison des fonctions ln, exp et puissance	10
5.1.3 Propriétés	10
5.2 Continuité	11

1 Probabilités

1.1 Quelques rappels

Une expérience aléatoire est une expérience dont le résultat est soumis au hasard. On lui associe l'ensemble **fondamental** Ω formé de tous les résultats possibles de l'expérience aléatoire. L'ensemble Ω peut être fini ou infini.

Un événement A est une partie (sous-ensemble) de Ω ($A \subset \Omega$). Un événement élémentaire est un élément ω de Ω ($\omega \in \Omega$). Notons les deux événements particuliers : Ω est l'événement certain et \emptyset est l'événement impossible.

Si $A \subset \Omega$ et $B \subset \Omega$ sont des événements, on note

- $A \cup B$: l'événement A ou B est réalisé
- $A \cap B$: les événements A et B sont réalisés
- \overline{A} : l'événement A n'est pas réalisé

Il existe une définition mathématique de la notion de probabilité que nous ne donnerons pas. Intuitivement, la probabilité d'un événement A (notée ici $Pr(A)$) est la mesure de la place qu'il occupe dans Ω . Il est alors clair que $Pr(A) \in [0 ; 1]$.

On a bien entendu $Pr(\emptyset) = 0$ et $Pr(\Omega) = 1$ mais attention, si $Pr(A) = 0$ alors A n'est pas forcément égal à \emptyset (c'est-à-dire, A n'est pas impossible).

On a les propriétés

- $A \subset B \implies Pr(A) \leq Pr(B)$
- $Pr(\overline{A}) = 1 - Pr(A)$
- $Pr(A) = Pr(A \cap B) + Pr(A \cap \overline{B})$
- $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$
- $A \cap B = \emptyset$ (A et B incompatibles) $\implies Pr(A \cup B) = Pr(A) + Pr(B)$

Remarque 1 Si Ω est un ensemble fini et si tous les événements élémentaires ont la même probabilité ($1/|Card(\Omega)|$) on a alors, pour tout événement A :

$$Pr(A) = \frac{Card(A)}{Card(\Omega)} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

1.2 Probabilités conditionnelles

On considère deux événements A et B . La probabilité conditionnelle de A sachant B est définie par :

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)}$$

On a les propriétés

$$Pr(B|B) = 1, \quad Pr(A|B) + Pr(\overline{A}|B) = 1 \text{ et } Pr(B|A) = \frac{Pr(B)}{Pr(A)}.Pr(A|B).$$

1.2.1 Indépendance

Les événements A et B sont dits **indépendants** si $Pr(A \cap B) = Pr(A) \times Pr(B)$. On a la propriété :

$$A \text{ et } B \text{ indépendants} \iff Pr(A|B) = P(A).$$

1.2.2 La formule des probabilités totales

On considère des événements B_1, B_2, \dots, B_k . S'ils forment une partition de Ω (c'est-à-dire, s'ils sont deux à deux disjoints et leur réunion est égale à Ω) on dit qu'ils forment un **système de probabilité totale** ou un **système complet d'événements**. On a alors la formule :

$$Pr(A) = Pr(A|B_1).Pr(B_1) + Pr(A|B_2).Pr(B_2) + \dots + Pr(A|\overline{B}).Pr(\overline{B}).$$

Cas particulier : $Pr(A) = Pr(A|B).Pr(B) + Pr(A|\overline{B}).Pr(\overline{B})$.

1.3 Variables aléatoires discrètes

On fait une expérience aléatoire menant à un ensemble fondamental Ω . Intuitivement, une variable aléatoire X est une fonction $X : \Omega \rightarrow \mathbb{R}$. Là encore, il existe une définition mathématique plus correcte qu'on ne donnera pas ici.

Une variable aléatoire X est dite **discrète** si l'ensemble des valeurs qu'elle peut prendre est un sous-ensemble fini ou infini dénombrable de \mathbb{R} . Si par contre, l'ensemble de ses valeurs est un intervalle de \mathbb{R} , on dit qu'elle est **continue**.
Dans tout ce qui suit, on ne s'intéresse qu'aux variables aléatoires discrètes même si les définitions données peuvent « facilement » s'adapter au cas continu.

1.3.1 Fonction de répartition

C'est la fonction $F : \mathbb{R} \rightarrow \mathbb{R}$ définie par $F(x) = Pr(X \leq x)$. Elle a quelques propriétés :

- F est une fonction en escalier,
- $Pr(a < X \leq b) = F(b) - F(a)$,
- F est croissante,
- $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$,
- $Pr(X > a) = 1 - Pr(X \leq a) = 1 - F(a)$.

1.3.2 Loi de probabilité

Si X est une variable aléatoire discrète avec pour ensemble de valeurs $\{x_1, x_2, \dots\}$, sa loi de probabilité est définie par tous les $p_i = Pr(X = x_i)$. Bien sûr, on a $\sum_i p_i = 1$. En fait, donner la loi de probabilité de X revient à donner l'ensemble de ses valeurs x_i ainsi que les probabilités correspondantes $Pr(X = x_i)$.

1.3.3 Moyenne et Variance

Comme dans le cas des statistiques, on peut définir moyenne (ou plutôt **espérance mathématique**) et variance d'une variable aléatoire X .

Si x_1, x_2, x_3, \dots sont les valeurs de X , alors l'espérance mathématique et la variance de X sont respectivement

$$\begin{aligned} E(X) &= \sum_i x_i Pr(X = x_i) \\ Var(X) &= \sum_i [(x_i - E(X))^2 Pr(X = x_i)] \\ &= (\sum_i x_i^2 Pr(X = x_i)) - (\sum_i x_i Pr(X = x_i))^2 \\ &= E(X^2) - E(X)^2. \end{aligned}$$

1.3.4 Combinaison linéaire de variables aléatoires indépendantes

On dit que deux variables aléatoires discrètes X et Y sont indépendantes si pour tous réels x et y , on a

$$Pr(X = x \text{ et } Y = y) = Pr(X = x) \times Pr(Y = y).$$

Dans ce cas, si α et β sont deux réels, on a

$$\begin{aligned} E(\alpha X + \beta Y) &= \alpha E(X) + \beta E(Y), \\ Var(\alpha X + \beta Y) &= \alpha^2 Var(X) + \beta^2 Var(Y). \end{aligned}$$

Notons qu'en fait, la formule précédente sur la moyenne est vraie même si les variables X et Y ne sont pas indépendantes. Par contre, la formule sur les variances nécessite l'hypothèse d'indépendance.

1.3.5 Lois usuelles discrètes

On se contente de donner ici les lois les plus usuelles avec leurs paramètres.

Loi de Bernoulli $B(p)$ Une variable aléatoire X suit une loi de Bernoulli de paramètre $p \in [0, 1]$ si elle ne prend que deux valeurs : 0 et 1, avec $Pr(X = 1) = p$. La loi de probabilité de X est alors

$$Pr(X = 1) = p \quad \text{et} \quad Pr(X = 0) = 1 - p,$$

et on a $E(X) = p$ et $Var(X) = p(1 - p)$.

Loi Binomiale $B(n, p)$ On fait de façon indépendante n fois la même expérience de Bernoulli $B(p)$. Plus précisément, on répète n fois, de manière indépendante, la même expérience aléatoire qui n'a que deux issues possibles : les événements A et \bar{A} , avec $Pr(A) = p$. La variable X correspond alors au nombre de fois où le résultat de l'expérience a été A . Les valeurs de X sont donc $\{0, 1, 2, \dots, n\}$, sa loi est

$$Pr(X = k) = C_n^k p^k (1 - p)^{n-k} \quad \text{où} \quad C_n^k = \frac{n!}{k!(n - k)!}.$$

et on a $E(X) = np$ et $Var(X) = np(1 - p)$.

Rappel : $n! = n(n - 1)(n - 2)\dots 1$ avec $0! = 1$.

Loi de Poisson $\mathcal{P}(\lambda)$ On l'appelle parfois la loi des « événements rares ». Une variable aléatoire X suit une loi de Poisson $\mathcal{P}(\lambda)$ de paramètre $\lambda > 0$ si ses valeurs sont $\{0, 1, 2, 3, \dots\}$, c'est-à-dire l'ensemble des entiers naturels \mathbb{N} et sa loi de probabilité :

$$Pr(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Notons que l'ensemble des valeurs de X n'est pas fini mais est infini dénombrable.

Un calcul (pas forcément évident) donne alors $E(X) = Var(X) = \lambda$.

Remarque 2 Si X suit une loi $\mathcal{P}(\lambda)$ et Y une loi $\mathcal{P}(\mu)$ et si X et Y sont indépendantes alors, la somme $X + Y$ suit une loi de Poisson $\mathcal{P}(\lambda + \mu)$.

Approximation d'une loi Binomiale par une loi de Poisson On suppose que X suit une loi binomiale $\mathcal{B}(n, p)$. Si, par exemple, $n \geq 30$ (n grand), $p \leq 0.1$ (p petit) et $np \leq 10$ alors on peut supposer que X suit en fait approximativement une loi de Poisson $\mathcal{P}(\lambda)$ avec $\lambda = np$. Les conditions de validité de cette approximation ne sont ni standard ni forcément « optimales »; tout dépend de la finesse que l'on souhaite donner à cette approximation.

2 Initiation aux tests du χ^2

On considère une population dans laquelle chaque individu peut être classé dans une catégorie parmi k . On note C_1, \dots, C_k ces catégories, supposées disjointes.

On se donne des proportions p_1, p_2, \dots, p_k (avec $\sum p_j = 1$) et on se demande alors si on peut considérer que les proportions des catégories C_1, \dots, C_k dans la population sont p_1, p_2, \dots, p_k . On prend un échantillon et on observe la répartition des individus p_1, p_2, \dots, p_k dans les diverses catégories

On dispose d'un échantillon de taille n dans lequel on compte n_1 individus de la 1^{ere} catégorie, n_2 de la 2^{de} catégorie, ..., n_k de la k^{th} catégorie. La question est alors : les différences entre les proportions $n_1/n, \dots, n_k/n$ et des valeurs théoriques données p_1, p_2, \dots, p_k sont-elles dues à des fluctuations du hasard (l'échantillon appartient à une population distribuée avec les proportions p_1, p_2, \dots, p_k) ou bien ces différences sont-elles le reflet d'un vrai phénomène qui crée une différence entre la réalité et les proportions proposées (les proportions données p_1, p_2, \dots, p_k ne sont pas bonnes).

Pour chaque catégorie C_j on définit :

- l'**effectif observé** : c'est le nombre d'individus de l'échantillon appartenant à la catégorie C_j , c'est-à-dire n_j
 - l'**effectif théorique** : c'est le nombre d'individus appartenant à la catégorie C_j qu'on s'attend à trouver théoriquement dans un échantillon de taille n , c'est-à-dire $n_{\text{th},j}$.
- On calcule alors

$$Q = \sum \frac{(\text{effectif observé} - \text{effectif théorique})^2}{\text{effectif théorique}} = \sum_{j=1}^k \frac{(n_j - np_j)^2}{np_j}$$

Comme il y a k catégories, le **degré de liberté** du test sera égal à $k - 1$ (nombre de catégories moins un) et le **seul critère** correspondant z_{k-1} sera lu dans la table numérique ci-dessous (d représente le degré de liberté et z_d le seuil critique correspondant) :

d	1	2	3	4	5	6	7	8	9	10
z_d	3.841	5.991	7.815	9.488	11.070	12.592	14.067	15.507	16.919	18.307

Sous les conditions $n \geq 30$ et $np_j \geq 5$ pour tout j (les effectifs théoriques sont tous supérieurs à 5), on a alors la conclusion

- si $Q > z_{k-1}$, la différence entre les proportions théoriques et les proportions trouvées sur l'échantillon est trop importante, on peut dire que les proportions p_1, p_2, \dots, p_k ne sont pas bonnes et ne correspondent pas à la réalité.
- si $Q \leq z_{k-1}$, alors on peut estimer que les proportions théoriques données sont à peu près correctes.

Remarque 3 Quelle que soit la conclusion du test, elle est peut-être fausse ! En effet, le seul critère z_{k-1} qui permet de créer une frontière artificielle est accompagné d'une probabilité égale à 0,05 de se tromper. On peut minimiser ce risque d'erreur si on veut mais cela peut éventuellement créer d'autres problèmes. Il faut accepter ce risque (correct) de se tromper. Pourquoi parle-t-on de test du χ^2 ? En fait, on peut montrer que la variable aléatoire qui correspond à la formule de Q suit ce qu'on appelle une loi du χ^2 à $k - 1$ degrés de liberté. On a donné ici une version très simplifiée, voire grossière, des tests du χ^2 . Les étudiants désireux d'en savoir plus peuvent consulter les cours classiques de biostatistique.

3 Fonctions : Généralités

3.1 Domaine de définition

Une fonction f est habituellement donnée par une expression $y = f(x)$ et, par définition, le **domaine de définition** (noté \mathcal{D}_f) de f est constitué de l'ensemble des nombres réels x pour lesquels l'expression $f(x)$ a un sens. On pourra alors écrire

$$\begin{aligned} f : \quad I &\longrightarrow \mathbb{R} \\ x &\mapsto f(x) \end{aligned}$$

dès que I est un sous-ensemble de \mathcal{D}_f .

3.2 Composition des fonctions f et g

Par définition, $f \circ g : x \mapsto (f \circ g)(x) := f(g(x))$.

En particulier, x est dans $\mathcal{D}(f \circ g)$ si l'expression $g(x)$ a un sens (autrement dit si x est dans $\mathcal{D}g$) et si l'expression $f(g(x))$ a aussi un sens donc si $g(x)$ est dans $\mathcal{D}f$.

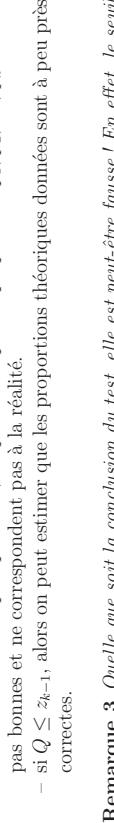
Exemple : Soient $f(x) = \frac{1}{x}$ et $g(x) = \sqrt{x-2}$. Alors $f \circ g(x) = f(g(x)) = \frac{1}{g(x)} = \frac{1}{\sqrt{x-2}}$. De plus, $f(x)$ n'a de sens que si $x \neq 0$ et comme la racine carrée n'est définie que pour des nombres positifs ou nuls, $g(x)$ n'a de sens que si $x - 2 \geq 0$ ($\mathcal{D}f = \mathbb{R}^*$ et $\mathcal{D}g = [2, +\infty)$). On en déduit que $f \circ g(x)$ a un sens si $x \geq 2$ et $\sqrt{x-2} \neq 0$ ($\mathcal{D}(f \circ g) = [2, +\infty)$.

3.3 Représentation graphique

- La fonction f est périodique de période T ou T -périodique si $f(x+T) = f(x)$ pour tout x de $\mathcal{D}f$; il suffit alors de connaître f sur un intervalle de longueur T pour la connaître sur $\mathcal{D}f$ entier.
- La fonction f est impaire si $f(-x) = -f(x)$ pour tout x de $\mathcal{D}f$; $(0, 0)$ est alors centre de symétrie du graphe de f .
- La fonction f est paire si $f(-x) = f(x)$ pour tout x de $\mathcal{D}f$; l'axe des y est alors axe de symétrie du graphe de f .

Exemples :

- f définie par $f(x) = \sin(2x+a)$ est π -périodique.
- f définie par $f(x) = x^3 - 4x$ est impaire.
- f définie par $f(x) = x^4 - 4x^2 - 1$ est paire.



3.4 Sens de variation

- f est croissante (resp. strictement croissante) sur I si

$$\forall x, y \in I, \quad x > y \Rightarrow f(x) \geq f(y) \quad (\text{resp. } f(x) > f(y)).$$

- f est décroissante (resp. strictement décroissante) sur I si

$$\forall x, y \in I, x > y \Rightarrow f(x) \leq f(y) \text{ (resp. } f(x) < f(y)).$$

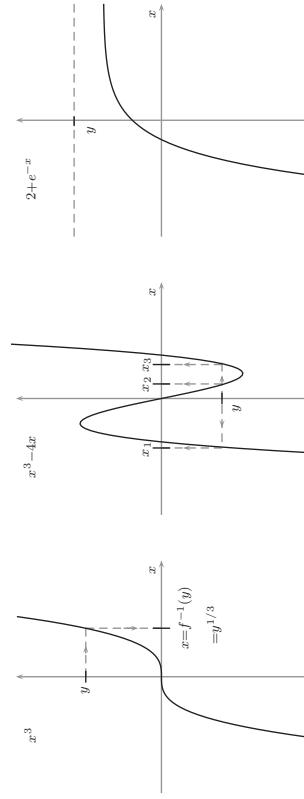
- f est monotone (resp. strictement monotone) sur I si elle est croissante sur I ou décroissante sur I (resp. strictement croissante sur I ou strictement décroissante sur I).

3.5 Fonction réciproque

L'application $f : I \rightarrow J$ est une bijection de I sur J si pour tout y de J , il existe un unique x dans I tel que $f(x) = y$. Dans ce cas, on peut définir la fonction réciproque de f , notée f^{-1} , par $f^{-1}(y) = x$ si $f(x) = y$; f^{-1} est alors une bijection de J sur I et $(f \circ f^{-1})(y) = y$ pour tout y de J , $(f^{-1} \circ f)(x) = x$ pour tout x de I .

Exemples :

- $\ln :]0, +\infty[\rightarrow \mathbb{R}$ est la fonction réciproque de $\exp : \mathbb{R} \rightarrow]0, +\infty[$.
- La fonction $f(x) = x^3$ est bijective de \mathbb{R} dans \mathbb{R} .
- La fonction $f(x) = x^3 - 4x$ n'est pas bijective de \mathbb{R} dans \mathbb{R} (pour certaines valeurs y dans \mathbb{R} , il existe plusieurs réels x tels que $f(x) = y$).
- La fonction $f(x) = 2 + e^{-x}$ n'est pas bijective de \mathbb{R} dans \mathbb{R} (pour certaines valeurs y dans \mathbb{R} , il n'existe pas de réel x tels que $f(x) = y$).



4 Deux fonctions classiques : \exp et \ln

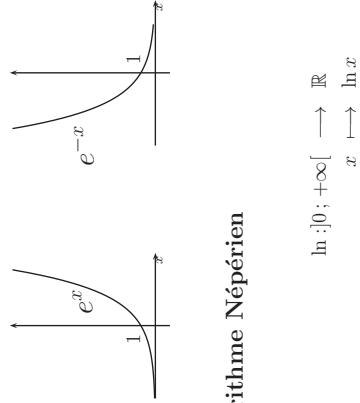
4.1 L'exponentielle

$$\begin{aligned} \exp : \mathbb{R} &\longrightarrow \mathbb{R}_+^* \\ x &\longmapsto e^x \quad \text{avec } e \approx 2,718\dots \end{aligned}$$

4.1.1 Propriétés classiques

- formules usuelles : $e^{a+b} = e^a e^b$; $e^0 = 1$; $e^{-a} = \frac{1}{e^a}$; $e^{ab} = (e^a)^b$.
- sens de variation : l'exponentielle est strictement croissante.
- limites : $\lim_{x \rightarrow -\infty} e^x = 0$ et $\lim_{x \rightarrow +\infty} e^x = +\infty$.

4.1.2 Graphes



4.2 Le Logarithme Népérien

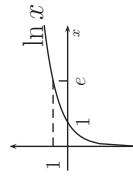
$$\begin{aligned} \ln :]0 ; +\infty[&\longrightarrow \mathbb{R} \\ x &\longmapsto \ln x \end{aligned}$$

4.2.1 Propriétés classiques

Si $a > 0, b > 0$ et α réel,

- formules usuelles : $\ln(ab) = \ln a + \ln b$; $\ln(a^\alpha) = \alpha \ln a$; $\ln 1 = 0$; $\ln(\frac{1}{a}) = -\ln a$; $\ln(\frac{a}{b}) = \ln a - \ln b$.
- sens de variation : \ln est strictement croissante.
- limites : $\lim_{x \rightarrow 0^+} \ln x = -\infty$ et $\lim_{x \rightarrow +\infty} \ln x = +\infty$.

4.2.2 Graphe



4.3 Relations \exp / \ln

Les fonctions \ln et \exp sont réciproques l'une de l'autre :

$$\begin{aligned}\ln(e^x) &= x && \text{pour tout } x \in \mathbb{R}, \\ \exp(\ln x) &= x && \text{pour tout } x > 0.\end{aligned}$$

On peut aussi l'exprimer par

$$y = e^x \iff x = \ln y \quad \text{pour tout } x \in \mathbb{R} \text{ et } y \in \mathbb{R}_+^*.$$

Remarque 4 Les fonctions \log_a .

En biologie, chimie ou physique, on utilise parfois des logarithmes \log_a de base a où a est un réel strictement positif (par exemple $a = 10$). Ce logarithme est défini par la relation suivante :

$$y = a^x \iff x = \log_a y.$$

Bien sûr, la fonction \ln est le logarithme \log_e de base e . Il est assez simple de vérifier la relation

$$\log_a y = \frac{\ln y}{\ln a}$$

pour tout $y > 0$.

5 Fonctions : limites et continuité

5.1 Limites d'une fonction

Dans cette section, on ne donnera pas les définitions « mathématiques » des notions de limites mais seulement les limites usuelles et les règles de calcul. On rappelle juste que rechercher la limite d'une fonction, c'est déterminer si cette fonction s'approche d'une valeur particulière lorsque la variable s'approche d'une valeur donnée (finie ou infinie).

5.1.1 Limites classiques

$$\begin{aligned}\lim_{x \xrightarrow{a_0} x} \frac{\sin x}{x} &= 1 & \lim_{x \xrightarrow{a_0} x} \frac{\ln(1+x)}{x} &= 1 & \lim_{x \xrightarrow{a_0} x} \frac{e^x - 1}{x} &= 1 \\ \lim_{x \xrightarrow{0^+} x} x \ln x &= 0 & \lim_{x \xrightarrow{+\infty} x} \frac{\ln x}{x} &= 0 & \lim_{x \xrightarrow{+\infty} x} \frac{e^x}{x} &= +\infty \\ \lim_{x \xrightarrow{+\infty} x} x e^x &= 0 & \lim_{x \xrightarrow{+\infty} x} \frac{x e^x}{x} &= +\infty & \lim_{x \xrightarrow{+\infty} x} x e^x &= 0\end{aligned}$$

5.1.2 Comparaison des fonctions \ln , \exp et puissance

Pour tout a réel et $x > 0$, on définit les fonctions puissances $x^a := e^{a \ln x}$.

Règles générales de comparaisons :

$$\begin{aligned}&\forall a \in \mathbb{R}, \forall b > 0, \lim_{x \rightarrow 0^+} (-\ln x)^a x^b = 0 \\ &\forall a \in \mathbb{R}, \forall b > 0, \lim_{x \rightarrow +\infty} \frac{(\ln x)^a}{x^b} = 0 \\ &\forall a > 0, \forall b \in \mathbb{R}, \lim_{x \rightarrow +\infty} \frac{e^{ax}}{x^b} = +\infty \\ &\forall a > 0, \forall b \in \mathbb{R}, \lim_{x \rightarrow +\infty} e^{ax} x^b = 0\end{aligned}$$

Graphes de $x \mapsto x^a$:

- $x^a, a < 0$: courbes passant par l'origine, croissantes pour $x < 0$ et décroissantes pour $x > 0$.
- $x^1 = x$: droite passant par l'origine et l'unité.
- $x^a, a > 1$: courbes passant par l'origine, décroissantes pour $x < 0$ et croissantes pour $x > 0$.

On résume souvent cela en : « Dans le cas d'une forme indéterminée, les puissances l'emportent sur les logarithmes et l'exponentielle l'emporte sur les puissances ».

5.1.3 Propriétés

Théorème des gendarmes Soient f , g et h des fonctions définies autour de x_0 (sauf, éventuellement, en x_0) et telles que $f(x) \leq h(x) \leq g(x)$, pour tout x de ce voisinage ($x \neq x_0$). Si f et g admettent la même limite l en x_0 , alors l est aussi la limite de h en x_0 .

Opérations sur les limites On peut résumer les différents comportements par les tableaux suivants :

6 Fonctions : dérivabilité

6.1 Définition, calcul pratique

6.1.1 Définition, interprétation géométrique

Un des intérêts de la notion de dérivée est de pouvoir approcher localement le graphe d'une fonction f par une droite. Soit $I = [a, b]$ un intervalle ouvert, $x_0 \in I$ et $f : I \rightarrow \mathbb{R}$. Pour tout x de I , $x \neq x_0$, le rapport

$$t_{x,x_0} = \frac{f(x) - f(x_0)}{x - x_0}$$

est appelé **taux d'accroissement** de f entre x et x_0 . C'est la pente de la corde entre les points $(x_0, f(x_0))$ et $(x, f(x))$. Plus x est proche de x_0 , plus la corde approche bien la courbe.

$\lim_{x \rightarrow x_0} f(x)$	$-\infty$	$l \in \mathbb{R}$	$+\infty$
$\lim_{x \rightarrow x_0} g(x)$	$-\infty$	$-\infty$	IND
$\lim_{x \rightarrow x_0} (f + g)(x)$	$l' \in \mathbb{R}$	$l + l'$	$+\infty$
$\lim_{x \rightarrow x_0} \frac{1}{g(x)}$	$+\infty$	IND	$+\infty$

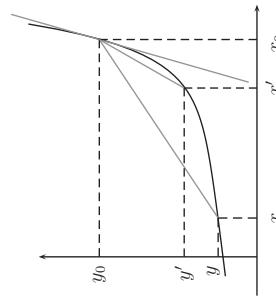
$\lim_{x \rightarrow x_0} f(x)$	$-\infty$	$l < 0$	0	$l > 0$	$+\infty$
$\lim_{x \rightarrow x_0} g(x)$	$+\infty$	$+ \infty$	IND	$-\infty$	$-\infty$
$\lim_{x \rightarrow x_0} (f \cdot g)(x)$	$l' < 0$	$+ \infty$	ll'	0	$-\infty$
$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$	0	IND	0	0	IND
$\lim_{x \rightarrow x_0} \frac{1}{f(x)}$	$-\infty$	ll'	0	ll'	$+\infty$
$\lim_{x \rightarrow x_0} \frac{1}{g(x)}$	$+\infty$	$- \infty$	IND	$+\infty$	$+\infty$

$\lim_{x \rightarrow x_0} f(x)$	$l \in \mathbb{R}^*$	0^+	0^-	$\pm \infty$	0
$\lim_{x \rightarrow x_0} \left(\frac{1}{f}\right)(x)$	$\frac{1}{l}$	$\frac{1}{l}$	$+\infty$	$-\infty$	0

Remarque 5 • IND signifie qu'il s'agit d'une forme indéterminée. On ne peut pas conclure

directement. On doit transformer l'expression pour éliminer la forme indéterminée. Il existe d'autres formes indéterminées dont il faut se méfier, comme « $1^{+\infty}$ ». Par exemple, lorsque x tend vers $+\infty$, l'expression $(1 + \frac{1}{x})^x$ ne tend pas vers 1 mais vers e .

- On a donné les règles de calculs pour des limites en x_0 réel mais elles restent valables pour des limites en $\pm\infty$.



5.2 Continuité

Soit f une fonction et $x_0 \in \mathcal{D}_f$; on suppose que \mathcal{D}_f contient un voisinage de x_0 (i.e., on suppose qu'il existe un intervalle $[a, b]$ inclus dans \mathcal{D}_f et contenant x_0).

- On a donné les règles de calculs pour des limites en x_0 si les limites restent valables pour des limites en $\pm\infty$.

Proposition 1 1. Si f et g sont continues en x_0 , alors $f + g$ et fg sont aussi continues en x_0 ; si, de plus, $g(x_0) \neq 0$, on a également f/g continue en x_0 .

2. Si f est continue en x_0 et g en $f(x_0)$, alors $g \circ f$ est continue en x_0 .

Définition 1 f est continue en x_0 si $\lim_{x \rightarrow x_0} f(x) = f(x_0)$.

Proposition 1 1. Si f et g sont continues en x_0 , alors $f + g$ et fg sont aussi continues en x_0 ; si, de plus, $g(x_0) \neq 0$, on a également f/g continue en x_0 .

2. Si f est continue en x_0 et g en $f(x_0)$, alors $g \circ f$ est continue en x_0 .

Définition 2 f est continue sur l'intervalle I de \mathbb{R} si et seulement si f est continue en tout point de I .

Remarque 6 Prolongement par continuité : si f est une fonction non définie en $a \in \mathbb{R}$, mais définie « autour » de a (i.e. si f est définie sur un intervalle ouvert contenant a sauf en a) et si $\lim_{x \rightarrow a} f(x)$ existe et est finie, on peut prolonger f par continuité en a en posant

$$f(a) = \lim_{x \xrightarrow{x \neq a} a} f(x).$$

Parfois, on voudra distinguer f (non définie en a) et son prolongement (dont

$$f(a) = \lim_{x \xrightarrow{x \neq a} a} f(x),$$

le domaine de définition est $\mathcal{D} \cup \{a\}$).

Par exemple, f définie par $f(x) = x \cos(\frac{1}{x})$ est continue sur \mathbb{R}^* et la fonction g définie par $g(x) = f(x)$ si $x \in \mathbb{R}^*$ et $g(0) = 0$ est continue sur \mathbb{R} ; g est le prolongement par continuité de f à \mathbb{R} tout entier.

Définition 3 Soit $I =]a, b[$ un intervalle ouvert, $x_0 \in I$ et $f : I \rightarrow \mathbb{R}$. On dit que f est dérivable en x_0 si $\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$ existe. On appelle alors $f'(x_0)$ cette limite. Si f est dérivable en tout point de I , on dit que f est dérivable sur I et on note $f' : x \mapsto f'(x)$ sa dérivée. Autrement dit, f est dérivable en x_0 si l'application

$$t_{x_0} : I - \{x_0\} \rightarrow \mathbb{R}$$

$$x \mapsto t_{x_0}(x) = \frac{f(x) - f(x_0)}{x - x_0}$$

admet une limite en x_0 (et cette limite est $f'(x_0)$).

Ainsi $f'(x_0)$ s'interprète naturellement comme la pente de la tangente à la courbe de f en x_0 . Plus précisément, si f est dérivable en x_0 , alors la courbe représentative de f admet une tangente en $(x_0, f(x_0))$ et l'équation de cette tangente est

$$y = f'(x_0)(x - x_0) + f(x_0).$$

Le tableau suivant donne les dérivées usuelles :

$f(x)$	x^a	$a \in \mathbb{R}$	$\ln(x)$	e^x	$\sin(x)$	$\cos(x)$	$\tan(x)$
$f'(x)$	ax^{a-1}		$\frac{1}{x}$	e^x	$\cos(x)$	$-\sin(x)$	$\frac{1}{\cos^2(x)}$

On peut énoncer quelques autres propriétés :

1. La dérivée d'une fonction constante est la fonction nulle.

2. On peut avoir une inclusion stricte $\mathcal{D}' \subset \mathcal{D}f$ (exemple : $f(x) = \sqrt{x}$).

3. Ne pas oublier !!! : si f est dérivable en x_0 , alors f est continue en x_0 mais la réciproque est fausse (exemple : $|x|$ en 0).

4. On définit également $f'_d(x_0)$ et $f'_g(x_0)$, les dérivées à droite et à gauche en x_0 . Bien entendu, $f'_d(x_0)$ et $f'_g(x_0)$ peuvent exister et être différentes ; on a en fait :

$$f'(x_0) \text{ existe} \iff \begin{cases} i) f'_d(x_0) \text{ et } f'_g(x_0) \text{ existent} \\ \text{et} \\ ii) f'_d(x_0) = f'_g(x_0) \end{cases}$$

Dans ce cas, on a alors $f'(x_0) = f'_d(x_0) = f'_g(x_0)$.

6.1.3 Dérivées et opérations

On suppose que f et g sont des fonctions dériviales en x_0 . Alors $f + g$ et fg sont dériviales en x_0 avec

$$(f + g)'(x_0) = f'(x_0) + g'(x_0) \text{ et } (fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

Si, de plus, on a $g(x_0) \neq 0$, alors $\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$ est aussi dérivable en x_0 avec

$$(g \circ f)'(x_0) = f'(x_0) \times g(f(x_0)).$$

En utilisant la formule de dérivation d'une fonction composée on peut démontrer les formules suivantes :

$$(f^a)' = af' \times f^{a-1}, \quad (\exp f)' = f' \times \exp f, \quad (\ln f)' = \frac{f'}{f}.$$

De même, si f est bijective et dérivable, de dérivée non nulle, alors f^{-1} est également dérivable et on a

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

6.2 Résultats

6.2.1 Monotonie

Si f est dérivable sur $I =]a, b[$. Alors :

- f est croissante sur I si et seulement si $f' \geq 0$ sur I .
- f strictement croissante sur I si et seulement si $f' > 0$ sur I et ne s'annule qu'en des points isolés de I .
- f est décroissante sur I si et seulement si $f' \leq 0$ sur I .
- f strictement décroissante sur I si et seulement si $f' < 0$ sur I et ne s'annule qu'en des points isolés de I .

6.2.2 Extrêmes

Définition 4 Soit $f : I \rightarrow \mathbb{R}$ une fonction définie sur un intervalle ouvert I .

- x_0 est un maximum local de f si il existe un intervalle ouvert $J \subset I$ tel que $f(x) \leq f(x_0)$ pour tout x de J .
- Si on remplace \leq par \geq dans la définition précédente, on obtient la définition d'un minimum local.

- Un extrême désigne soit un maximum, soit un minimum.

Théorème 1 Soit $f : I =]a, b[\rightarrow \mathbb{R}$, $x_0 \in I$. On suppose que f est dérivable en x_0 . Si f admet un extrême local en x_0 , alors $f'(x_0) = 0$.

Remarque 7 1. La "réciproque" est fausse. On peut avoir $f'(x_0) = 0$ en un point x_0 où f est dérivable sans que ce point soit un extrême (exemple : $f(x) = x^3$ et $x_0 = 0$).
2. Ainsi, lorsque f est dérivable sur un intervalle I ouvert, les extrema de f sont à chercher parmi les points critiques (là où la dérivée s'annule).

3. Attention!!! Ce qui précède ne vaut que pour $]a, b[$ (intervalle ouvert) ; si l'on cherche les extrema d'une fonction f sur un intervalle fermé $[a, b]$, il faut également regarder ce qui se passe aux bornes de l'intervalle (en a et en b). Par exemple, le maximum de f définie par $f(x) = x$ sur $[0, 1]$ est en 1 bien que $f'(1) = 1 \neq 0$.

7 Intégration

Dans ce chapitre, f désigne une fonction (au moins) continue de $[a, b]$ dans \mathbb{R} .

7.1 Primitives

Définition 5 Une primitive de f est une fonction dérivable $F : [a, b] \rightarrow \mathbb{R}$ telle que $F' = f$.
On note souvent $F = \int f$ ou bien $F(x) = \int_a^x f(t)dt$.

Les primitives de f sont définies à une constante additive près. Si F est une primitive de f et λ un réel, alors $F + \lambda$ est aussi une primitive de f . Réciproquement, si F_1 et F_2 sont deux primitives de f alors $F_1 - F_2$ est une constante.

Le tableau suivant donne les primitives usuelles :

$f(x)$	1	$x^a, a \neq -1$	$-1/x$	e^x	$\cos(x)$	$\sin(x)$
$F(x)$	x	$\frac{x^{a+1}}{a+1}$	$\ln x $	e^x	$\sin(x)$	$-\cos(x)$

Comme pour la dérivation, l'opération de « primitive » est linéaire :

Proposition 2 Si F (resp. G) est une primitive de f (resp. g) et λ et μ deux réels, alors $\lambda F + \mu G$ est une primitive de $\lambda f + \mu g$.

7.2 Intégrales

Définition 6 L'intégrale de f sur $[a, b]$ est le nombre réel

$$\int_a^b f(t)dt = [F(t)]_a^b := F(b) - F(a),$$

où F est une primitive de f (cette définition ne dépend pas du choix de la primitive F).

Proposition 3 • $\int_a^b (\lambda f(t) + \mu g(t))dt = \lambda \int_a^b f(t)dt + \mu \int_a^b g(t)dt$.

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

- $\int_a^a f(t)dt = 0$.

- $\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt$ (relation de Chasles).

- Si $f(x) \geq 0$ pour tout x dans $[a ; b]$ alors $\int_a^b f(t)dt \geq 0$.

- La valeur moyenne de la fonction f sur l'intervalle $[a, b]$ est $\frac{1}{b-a} \int_a^b f(t)dt$.

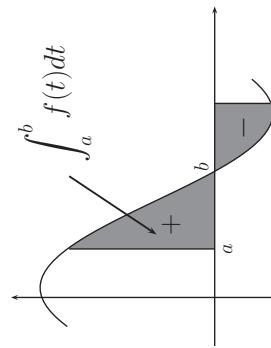
Définition 7 Si c est un réel dans $[a, b]$, la fonction

$$\begin{aligned} [a, b] &\rightarrow \mathbb{R} \\ x &\mapsto \int_c^x f(t)dt \end{aligned}$$

est la primitive de f qui s'annule en c .

7.3 Interprétation géométrique

Le graphe de la fonction f sur $[a, b]$ est par exemple



Le réel $\int_a^b f(t)dt$ est la somme algébrique des aires de la partie du plan délimitée par les droites $x = a$, $x = b$, $y = 0$ et la courbe de f ; ces aires étant comptées positivement si f est positive et négativement si f est négative.

7.4 Quelques techniques de calcul

7.4.1 Changement de variable

La formule donnant la dérivée d'une composition de deux fonctions (voir chapitre sur la dérivation) donne

$$\int_a^x f'(t) \times \varphi'(t)dt = f(\varphi(x)).$$

On en déduit alors

$$\begin{aligned} \int_x^y e^{\varphi(x)}\varphi'(x)dx &= e^{\varphi(y)} \\ \int_x^y \varphi(x)^\alpha \varphi'(x)dx &= \frac{\varphi(y)^{\alpha+1} - \varphi(x)^{\alpha+1}}{\alpha+1} \quad \text{avec } \alpha \neq -1 \\ \int_x^y \frac{\varphi'(x)}{\varphi(x)}dx &= \ln(|\varphi(x)|) \\ \int_x^y \sin(\varphi(x))\varphi'(x)dx &= -\cos(\varphi(x)) \\ \int_x^y \cos(\varphi(x))\varphi'(x)dx &= \sin(\varphi(x)) \end{aligned}$$

7.4.2 Intégration par parties

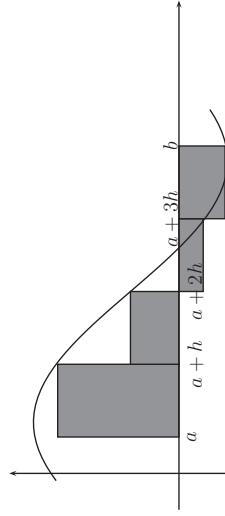
En utilisant la formule de dérivation d'un produit de deux fonctions, on obtient facilement

$$\int_a^b f(t)g'(t)dt = [f(t)g(t)]_a^b - \int_a^b f'(t)g(t)dt.$$

7.5 Intégration numérique

Dans certains cas, on ne sait pas calculer explicitement une primitive de la fonction, ou bien on ne dispose pas de l'expression de la fonction mais seulement de valeurs expérimentales en des points discrets, mais on souhaite quand même calculer une valeur approchée de l'intégrale de la fonction sur un intervalle $[a, b]$ donné. Dans ce cas, on approche la fonction sur des sous-intervalles de $[a, b]$ par une fonction constante sur une fonction affine.

7.5.1 Méthodes des rectangles

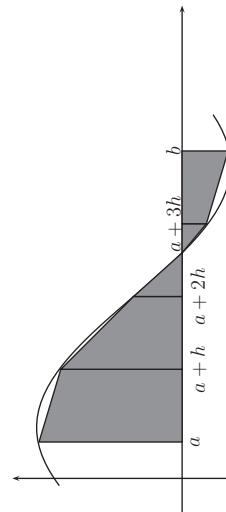


Sur chaque intervalle, l'aire à calculer est celle d'un rectangle. Dans le cas d'une subdivision régulière de l'intervalle en n morceaux, on obtient :

$$\int_a^b f(t) dt \simeq h \sum_{k=1}^n f(a + kh) \text{ avec } h = \frac{b-a}{n}.$$

Dans notre cas, on a choisi arbitrairement d'utiliser la valeur de la fonction à droite de la subdivision. On a des formules équivalentes en utilisant la valeur de la fonction à gauche de l'intervalle, ou au milieu de l'intervalle.

7.5.2 Méthodes des trapèzes



Sur chaque intervalle, l'aire à calculer est celle d'un trapèze. Dans le cas d'une subdivision régulière de l'intervalle en n morceaux, on obtient :

$$\int_a^b f(t) dt \simeq h \sum_{k=0}^{n-1} \left(\frac{f(a + kh) + f(a + (k+1)h)}{2} \right) \text{ avec } h = \frac{b-a}{n}.$$

8 Equations Différentielles

Une équation différentielle d'ordre k en la fonction f est une équation faisant intervenir $t, f(t)$ et les dérivées successives de f au point t , jusqu'à la dérivée k -ème :

$$\mathcal{F}(t, f(t), f'(t), f''(t), \dots, f^{(k)}(t)) = 0. \quad (1)$$

Exemples :

- $t^3 f'''(t) + f'(t) \cos(f(t)) = e^t$,
- $\frac{f'(t)}{f(t)^2} = \ln t$.

Résoudre (1) consiste à trouver toutes les fonctions f qui vérifient (1) sur un certain intervalle I .

8.1 Équations différentielles linéaires d'ordre 1

8.1.1 Définitions

Ce sont les équations différentielles du type

$$a(t)f'(t) + b(t)f(t) = g(t) \quad (2)$$

où a, b et g sont des fonctions continues sur un intervalle I .

On suppose que la fonction a ne s'annule pas sur I . Sinon, on se place sur un sous-intervalle de I sur lequel la fonction a ne s'annule pas.

Si la fonction g est la fonction identiquement nulle, on dit que l'équation (2) est homogène.

Si la fonction g n'est pas la fonction identiquement nulle, on appelle équation différentielle homogène associée à (2) l'équation différentielle suivante :

$$a(t)f'(t) + b(t)f(t) = 0. \quad (3)$$

Le résultat suivant assure que l'équation (2) admet des solutions :

Théorème 2 (Cauchy-Lipschitz) *On suppose que les fonctions a, b et g sont continues sur I et que a ne s'annule pas sur I . Si $t_0 \in I$ et γ est réel, il existe une unique fonction f qui vérifie (2) pour tout t dans I et telle que $f(t_0) = \gamma$.*

La condition $f(t_0) = \gamma$ s'appelle une condition initiale.

8.1.2 Méthode de résolution

La résolution de (2) se fait en deux ou trois étapes :

1. On commence par résoudre l'équation homogène associée à E (3).

Remarque 8 Important La fonction nulle $N(t) = 0$ pour tout t est une solution de (3). L'unicité dans le théorème de Cauchy-Lipschitz dit alors que si $h(t)$ est une solution de (3) telle qu'il existe $\tau \in \mathbb{R}$ vérifiant $h(\tau) = 0$ alors forcément on a $h = N$. Autrement dit, la seule solution de (3) qui puisse s'annuler est la fonction nulle (constamment égale à 0).

Cherchons donc une solution non nulle f de (3). Elle ne s'annule pas donc on peut diviser par f :

$$\begin{aligned} a(t)f'(t) + b(t)f(t) = 0 &\Leftrightarrow \frac{a(t)f'(t)}{f(t)} = -b(t) \\ &\Leftrightarrow \frac{f'(t)}{f(t)} = -\frac{b(t)}{a(t)} \\ &\Leftrightarrow \ln|f(t)| = -\int \frac{b(t)}{a(t)} dt + C, \quad C \in \mathbb{R} \end{aligned}$$

$$\begin{aligned} &\Leftrightarrow |f(t)| = e^C e^{-\int \frac{b(t)}{a(t)} dt}, \quad C \in \mathbb{R}^* \\ &\Leftrightarrow f(t) = K e^{-\int \frac{b(t)}{a(t)} dt}, \quad K \in \mathbb{R}^*. \end{aligned}$$

On en déduit le résultat suivant

Théorème 3 *Les solutions de (3) sont toutes les fonctions du type $h(t) = K e^{\lambda t}$, où $K \in \mathbb{R}$ est une constante et λ est une primitive de $-\frac{b}{a}$.*

Remarque 9 *Si les coefficients a et b sont des constantes, les solutions de (3) sont définies sur \mathbb{R} tout entier.*

2. On résout l'équation complète.

Théorème 4 *La solution générale de (2) est donnée par :*

$$f = \text{Solution de (3)} + \text{Solution particulière de (2),}$$

c'est-à-dire que les solutions de (2) sont toutes les fonctions du type

$$f(t) = K e^{\lambda t} + f_{part}(t),$$

avec $K \in \mathbb{R}$ et f_{part} une solution particulière de (2).

La recherche de cette solution particulière f_{part} est souvent guidée par le problème concrètement lié à l'équation différentielle.

Une méthode permet souvent de trouver des solutions particulières de (2) en utilisant la forme des solutions de (3). Il s'agit de la méthode dite **méthode de la variation de la constante**. Son principe est décrit ci-dessous.

On a résolu l'équation homogène (3) dont les solutions sont $h(t) = K e^{\lambda t}$ où K est une constante et λ une primitive de $-b/a$. On cherche alors les solutions de (2) sous la forme $f(t) = C(t)e^{\lambda t}$ où maintenant C représente une **fonction** de t et non plus une constante. On a donc $f'(t) = C'(t)e^{\lambda t} + \lambda(t)C(t)e^{\lambda t}$. En reportant dans (2) on voit que :

$$\begin{aligned} f \text{ est solution de (2)} &\Leftrightarrow a(t)\left(C'(t)e^{\lambda t} + \lambda(t)C(t)e^{\lambda t}\right) + b(t)C(t)e^{\lambda t} = g(t) \\ &\Leftrightarrow a(t)\left(C'(t)e^{\lambda t} - \frac{b(t)}{a(t)}C(t)e^{\lambda t}\right) + b(t)C(t)e^{\lambda t} = g(t) \\ &\Leftrightarrow C'(t) = \frac{g(t)}{e^{\lambda t}a(t)}. \end{aligned}$$

Il ne reste plus qu'à intégrer ... pour trouver la fonction C .
Les solutions de (2) seront les fonctions

$$f(t) = (C(t) + K)e^{\lambda t}, \quad K \in \mathbb{R}.$$

Remarque 10 *Il ne faut pas croire que cette méthode de variation de la constante soit une méthode miracle. En effet, si on essaie de résoudre avec cette méthode l'équation $f'(t) + 2tf(t) = 1$, on est ramené à trouver l'expression de la fonction $K(t)$ vérifiant $K'(t) = e^2$, ce qui n'a rien d'évident ...*

A ce stade, on voit qu'il y a une infinité de solutions : à chaque constante C correspond une solution.

3. Si on dispose d'une condition initiale du type $f(t_0) = \gamma$, alors on peut déterminer la constante C et la solution devient unique.

8.2 Résolution numérique des équations linéaires d'ordre 1

Dans certains cas, on ne peut pas résoudre explicitement les équations différentielles (on ne sait pas calculer les primitives nécessaires). Dans ce cas, il existe des méthodes de résolution numérique.

L'une d'entre elles est la méthode d'Euler explicite. L'idée est, à partir des conditions initiales, d'approcher, à chaque pas de temps, la courbe de la solution par sa tangente. Autrement dit, si on connaît $f(t)$, on estime que $f(t + \delta t) \approx f(t) + f'(t)\delta t$. Comme f est solution d'une équation différentielle d'ordre 1, il ne reste plus qu'à exprimer, grâce à cette équation, $f'(t)$ en fonction de t et $f(t)$.

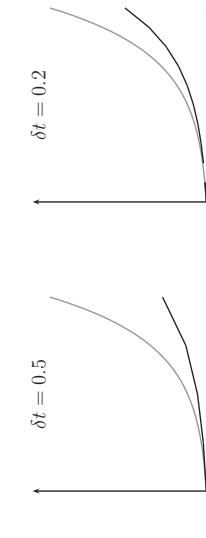
Exemple : Résolution numérique de l'équation différentielle $f'(t) - 2f(t) = 4$ avec la condition initiale $f(0) = 1$.

On note $t_n = t_0 + n\delta t$ et y_n une approximation de $f(t_n)$.

On part de $t_0 = 0$, $y_0 = f(t_0) = 1$. D'après l'équation différentielle, on a $f'(t) = 4 + 2f(t)$. On a donc les relations de récurrence $t_n = t_{n-1} + \delta t$ et

$$\begin{aligned} y_n &\simeq f(t_n) \\ &= f(t_{n-1} + \delta t) \\ &\simeq f(t_{n-1}) + f'(t_{n-1})\delta t \\ &= f(t_{n-1}) + \left(4 + 2f(t_{n-1})\right)\delta t \\ &\simeq y_{n-1} + \left(4 + 2y_{n-1}\right)\delta t \\ &= (1 + 2\delta t)y_{n-1} + 4\delta t. \end{aligned}$$

Graphiquement, cela donne



Plus le pas est petit, meilleure est l'approximation, mais cela augmente considérablement le temps de calcul. De plus, on s'éloigne toujours de plus en plus de la solution théorique car, sauf au premier pas, on ne prend pas la bonne valeur pour la dérivée. Dans notre exemple, à chaque itération, on a tendance à sous-estimer la pente de la courbe.