UNIVERSITÉ
D'EVRY-VAL-D'ESSONNE

L2 ÉCONOMIE & GESTION 2010-11

COURS DE MÉTHODES MATHÉMATIQUES 3

Alexandre VIDAL

Dernière modification : 11 janvier 2011

Table des matières

I Généralités et rappels sur les fonctions

I.1 Définition

Définition 1 : Une **fonction** f est la donnée :
- d'un ensemble de départ E,
- d'un ensemble d'arrivée F,
- et d'un ensemble G de couples $(x, y) \in E \times F$, appelé **graphe**, tel que si $(x, y_1) \in G$ et $(x, y_2) \in G$ alors $y_1 = y_2$.

Le **domaine de définition** de la fonction f, noté D_f, est défini par :

$$D_f = \{x \in E | \exists y \in F, (x, y) \in G\} \subset E.$$

Pour $x \in D_f$, l'**image** de x par f est l'unique $f(x) \in F$ tel que $(x, f(x)) \in G$. On note :

$$\begin{array}{rcl} f : D_f & \to & F \\ x & \to & f(x) \end{array} \quad \text{ou} \quad \begin{array}{rcl} f : D_f \subset E & \to & F \\ x & \to & f(x) \end{array}$$

Pour $y \in F$, tout élément $x \in D_f$ tel que $y = f(x)$ est appelé un **antécédent** de y (par f). Pour $U \subset E$, l'image de U par f est définie par $f(U) = \{f(x) | x \in D_f \cap U\} \subset F$.

Exemple 1 :

a) Soit la fonction f définie par l'ensemble de départ $E = \{a; b; c; d\}$, l'ensemble d'arrivée $F = \{rouge; vert; bleu; jaune\}$ et le graphe

$$G = \{(a, rouge); (b, bleu); (c, rouge)\}.$$

On vérifie que chaque élément de E est en relation avec au plus un élément de F. Le domaine de définition de la fonction est $\{a, b, c\}$. De plus, $f(E) = \{rouge, bleu\}$.
Représentation du graphe de la fonction :

jaune				
bleu		$(b, bleu)$		
vert				
rouge	$(a, rouge)$		$(c, rouge)$	
	a	b	c	d

On remarque que *rouge* admet plusieurs antécédents par cette fonction : a et c.

b) Avec E et F définis comme précédemment, la donnée :

$$E, \quad F, \quad G = \{(a, rouge); (b, bleu); (c, vert); (a, jaune)\} \subset E \times F$$

ne définit pas une fonction (à supposer, bien sûr, que $rouge \neq jaune$).

c) Soient $E = \mathbb{R}$, $F = \mathbb{R}$ et le graphe G défini par la relation $f(x) = 1 + \frac{1}{x}$, i.e.

$$G = \left\{ \left(x, 1 + \frac{1}{x} \right) | x \in \mathbb{R} \text{ et } 1 + \frac{1}{x} \in \mathbb{R} \right\}.$$

La donnée E,F,G définit bien une fonction f puisque, si $(x, y_1) \in G$ et $(x, y_2) \in G$, alors $y_1 = y_2 = 1 + \frac{1}{x} \in \mathbb{R}$ et $x \neq 0$. Le domaine de définition de f est $\mathbb{R}^*$. On a :

$$\begin{array}{rcl} f : \mathbb{R}^* \subset \mathbb{R} & \to & \mathbb{R} \\ x & \to & 1 + \frac{1}{x} \end{array}$$

De plus, $f(E) = \mathbb{R} \backslash \{1\} \subset F$. Voir Fig. 1, panel gauche.

d) Soient $E = \mathbb{R}$, $F = \mathbb{R}_+$ et le graphe défini par la relation $g(x) = 1 + \frac{1}{x}$, i.e.

$$G = \left\{ \left(x, 1 + \frac{1}{x} \right) | x \in \mathbb{R} \text{ et } 1 + \frac{1}{x} \in \mathbb{R}_+ \right\}.$$

La donnée E,F,G définit bien une fonction g puisque, si $(x, y_1) \in G$ et $(x, y_2) \in G$, alors $y_1 = y_2 = 1 + \frac{1}{x} \in \mathbb{R}_+$ et $x > 0$ et $1 + \frac{1}{x} \geq 0$. Le domaine de définition de f est $]-\infty, -1] \cup]0, +\infty[$. On note alors :

$$\begin{array}{rcl} f :]-\infty, -1] \cup]0, +\infty[\subset \mathbb{R} & \to & \mathbb{R} \\ x & \to & 1 + \frac{1}{x} \end{array}$$

De plus, $f(E) = \mathbb{R}_+ \backslash \{1\} \subset F$. Voir Fig. 1, panel droit.

Pour une fonction f d'une variable réelle à valeurs réelles, on représente le graphe de f dans le plan $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ muni d'un repère orthogonal (ou orthonormé) (O, x, y). Les deux derniers exemples sont illustrés par Fig. 1.

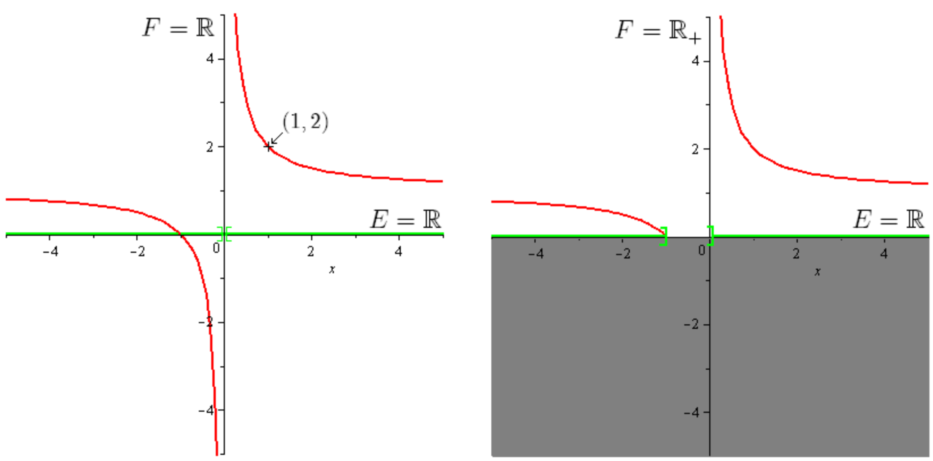

FIGURE 1 – Les graphes des fonctions représentées à gauche et à droite sont tous les deux constitués de points de la forme $\left(x, 1 + \frac{1}{x}\right)$. L'ensemble de départ est $\mathbb{R}$ dans les deux cas. Mais, la fonction de gauche (Exemple 1 c))est donnée avec l'ensemble d'arrivée $\mathbb{R}$ et celle de droite (Exemple 1 d)) avec l'ensemble d'arrivée $\mathbb{R}_+$. Dans ce dernier cas, le graphe $G \subset \mathbb{R} \times \mathbb{R}_+$ ne peut contenir de point de la partie grise. En conséquence, les domaines de définition des deux fonctions (en vert) sont différents : respectivement $\mathbb{R}^*$ et $]-\infty, -1] \cup]0, +\infty[$.

Les exemples précédents mettent en lumière l'importance des ensembles de départ et d'arrivée dans la définition d'une fonction : ces derniers ont une influence directe sur le domaine de définition. Ainsi, une expression numérique P_x ne suffit pas à définir une fonction.

I.2 Restriction, injectivité, surjectivité, bijectivité

Définition 2 : Soit

$$\begin{array}{rcl} f : D_f \subset E & \to & F \\ x & \to & f(x) \end{array}$$

et deux ensembles $A \subset E$ et $B \subset F$. On définit la **restriction** de f à A (au départ) et B à l'arrivée par :

$$\begin{array}{rcl} f_{|A}^{|B} : (D_f \cap A) \subset A & \to & B \\ x & \to & f(x) \end{array}$$

i.e. $f_{|A}^{|B}$ est donnée par l'ensemble départ A, l'ensemble d'arrivée B et le graphe

$$G = \{(x, f(x)) | x \in (D_f \cap A) \text{ et } f(x) \in B\}.$$

Exemple 2 : Dans l'Exemple 1, la fonction g définie au d) est la restriction de la fonction f à l'ensemble d'arrivée $\mathbb{R}_+$. Voir leurs graphes sur la Fig. 1.

Définition 3 : Soit

$$\begin{array}{rcl} f : D_f \subset E & \to & F \\ x & \to & f(x) \end{array}$$

– f est dite injective (de E dans F) si

$$D_f = E \text{ et } \forall (x_1, x_2) \in E \times E, f(x_1) = f(x_2) \Longrightarrow x_1 = x_2.$$

Cette propriété signifie que f est définie sur tout E et qu'un élément de l'ensemble d'arrivée admet **au plus** un antécédent par f.

– f est dite surjective (de E sur F) si $f(E) = F$, soit encore :

$$\forall y \in F, \exists x \in E, f(x) = y.$$

Cette propriété signifie que tout élément de l'ensemble d'arrivée admet **au moins** un antécédent par f.

– f est dite bijective (entre E et F) si elle est injective et surjective, i.e :

$$D_f = E \text{ et } \forall y \in F, \exists! x \in E, f(x) = y$$

($\exists!$... signifie "il existe un unique...").
Cette propriété signifie que f est définie sur tout E et que tout élément de F admet

exactement un antécédent dans E. On dit également que f est une bijection (entre E et F). Dans ce cas, on définit la <u>**bijection réciproque**</u> de f :

$$\begin{array}{rcl} f^{-1}: F & \to & E \\ y & \to & f^{-1}(y) \end{array}$$

où $f^{-1}(y)$ est l'unique antécédent de y par f. On a les propriétés suivantes :

$$\forall x \in E, f^{-1}(f(x)) = x$$
$$\forall y \in F, f(f^{-1}(y)) = y$$

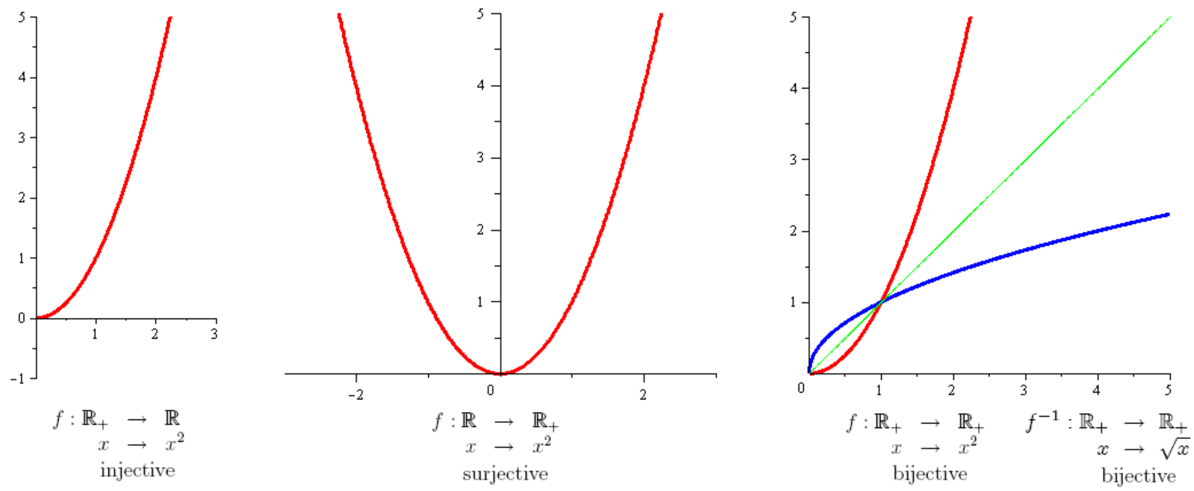

FIGURE 2 – Propriétés de $x \to x^2$ selon l'ensemble de départ et l'ensemble d'arrivée choisis.

Exemple 3 : On choisit $E = D_f$ pour l'ensemble de départ des fonctions suivantes. Les graphes des fonctions suivantes sont représentées sur la Fig. 2.

– La fonction
$$\begin{array}{rcl} f: \mathbb{R}_+ & \to & \mathbb{R} \\ x & \to & x^2 \end{array}$$
est injective . En effet si $x_1 \in \mathbb{R}_+$ et $x_2 \in \mathbb{R}_+$ sont tels que $x_1^2 = x_2^2$, alors $x_1 = x_2$. Attention : ceci est vrai parce que x_1 et x_2 sont positifs ($D_f = \mathbb{R}_+$).

– La fonction
$$\begin{array}{rcl} f: \mathbb{R} & \to & \mathbb{R}_+ \\ x & \to & x^2 \end{array}$$
est surjective. En effet pour tout élément $y \in \mathbb{R}_+$, il existe au moins un élément $x \in \mathbb{R}$ tel que $y = x^2$.

– La fonction
$$\begin{array}{rcl} f: \mathbb{R}_+ & \to & \mathbb{R}_+ \\ x & \to & x^2 \end{array}$$

est bijective. En effet, $E = D_f$ et pour tout élément $y \in \mathbb{R}_+$, il existe un unique élément $x \in \mathbb{R}_+$ tel que $y = x^2$.

Remarque 1 : Si G est le graphe d'une bijection f entre E et F, le graphe de sa bijection réciproque f^{-1} est

$$\{(y,x) \in F \times E | (x,y) \in G\}$$

Ainsi, si E et F sont des parties de $\mathbb{R}$ (f une fonction d'une variable réelle à valeurs réelles), on remarque donc que le graphe de f^{-1} est le symétrique du graphe de f par rapport à la première bissectrice (droite d'équation $y = x$). Voir Exemple 4 et Fig. 2.

Définition 4 : Soit

$$\begin{array}{rcl} f : D_f \subset E & \to & F \\ x & \to & f(x) \end{array}$$

et deux ensembles $A \subset E$ et $B \subset F$. On dit que f **induit** une bijection entre A et B si la restriction $\bar{f} = f_{|A}^{|B}$ est une fonction bijective (entre A et B).
Dans ce cas, on peut définir la bijection réciproque $\bar{f}^{-1}$ de cette restriction en suivant la Définition 3.

Exemple 4 : La fonction

$$\begin{array}{rcl} f : \mathbb{R} & \to & \mathbb{R} \\ x & \to & x^2 \end{array}$$

induit une bijection entre $\mathbb{R}_+$ et $\mathbb{R}_+$ puisque $f_{|\mathbb{R}_+}^{|\mathbb{R}_+}$ est une bijection. La bijection réciproque est donc bien définie : il s'agit de la fonction "racine carrée" définie de $\mathbb{R}_+$ dans $\mathbb{R}_+$.

I.3 Rappels sur les fonctions trigonométriques

On rappelle les fonctions trigonométriques classiques et leur graphes. Voir Fig. 3.
Les fonctions sin et cos sont définies sur tout $\mathbb{R}$ à valeurs dans $[-1,1]$ et 2π-périodiques. Comme

$$\cos x = 0 \Longleftrightarrow x = \frac{\pi}{2} + k\pi, \text{ avec } k \in \mathbb{Z},$$

la fonction π-périodique $\tan = \frac{\sin}{\cos}$ est définie sur

$$D_{\tan} = \mathbb{R} \backslash \left\{ \frac{\pi}{2} + k\pi | k \in \mathbb{Z} \right\}.$$

Formulaire :

$$
\begin{array}{lll}
\cos(-a) = \cos a, & \sin(-a) = -\sin a, & \tan(-a) = -\tan a, \\
\cos(\pi - a) = -\cos a, & \sin(\pi - a) = \sin a, & \tan(\pi - a) = -\tan a, \\
\cos(\pi + a) = -\cos a, & \sin(\pi + a) = -\sin a, & \tan(\pi + a) = \tan a, \\
\cos(\frac{\pi}{2} - a) = \sin a, & \sin(\frac{\pi}{2} - a) = \cos a, & \tan(\frac{\pi}{2} - a) = \text{cotan } a = \frac{1}{\tan a}, \\
\cos(\frac{\pi}{2} + a) = -\sin a, & \sin(\frac{\pi}{2} + a) = \cos a, & \tan(\frac{\pi}{2} + a) = -\text{cotan } a = -\frac{1}{\tan a}.
\end{array}
$$

Propriété fondamentale : $\forall a \in \mathbb{R}, \cos^2 a + \sin^2 a = 1.$

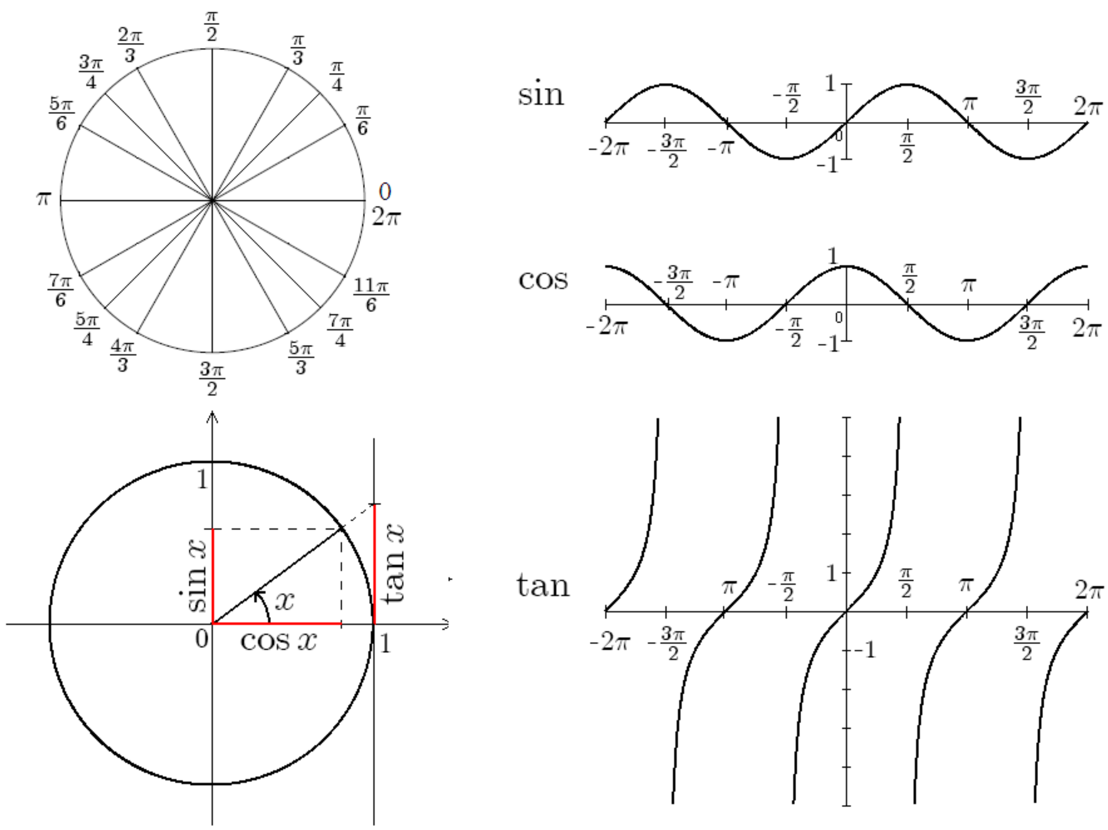

FIGURE 3 – Définition géométrique et graphe des fonctions trigonométriques sin, cos et tan.

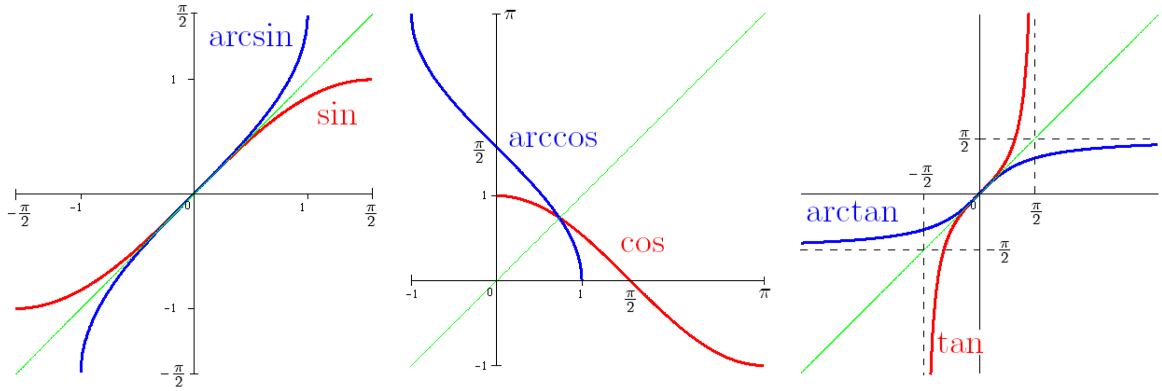

FIGURE 4 – Graphes des restrictions bijectives des fonctions sin, cos et tan et des bijections réciproques arcsin, arccos et arctan.

Les fonctions trigonométriques classiques induisent des bijections (voir Fig. 3 et Fig. 4) :
- sin induit une bijection entre $[-\frac{\pi}{2}, \frac{\pi}{2}]$ et $[-1, 1]$,
- cos induit une bijection entre $[0, \pi]$ et $[-1, 1]$,
- tan induit une bijection entre $[-\frac{\pi}{2}, \frac{\pi}{2}]$ et $\mathbb{R}$.

Les bijections réciproques sont respectivement notées :
- arcsin : $[-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$,
- arccos : $[-1, 1] \to [0, \pi]$,
- arctan : $\mathbb{R} \to [-\frac{\pi}{2}, \frac{\pi}{2}]$.

II Fonction d'une variable réelle

On s'intéresse dans ce chapitre aux fonctions d'une variable réelle (ensemble de départ inclus dans $\mathbb{R}$) à valeurs réelles (ensemble d'arrivée inclus dans $\mathbb{R}$).

II.1 Limite

Soient f une fonction réelle à valeurs réelles et $x_0 \in \mathbb{R}$. Les définitions suivantes sont illustrées par Fig. 5.

Définition 5 : 1) On suppose qu'il existe $\alpha > 0$ tel que $]x_0 - \alpha, x_0[\cup]x_0, x_0 + \alpha[\subset D_f$. Alors f admet une limite $l \in \mathbb{R}$ en x_0 si :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in D_f \backslash \{x_0\}, |x - x_0| < \eta \Longrightarrow |f(x) - l| < \varepsilon.$$

Cette propriété peut également s'écrire :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in D_f \backslash \{x_0\}, x_0 - \eta < x < x_0 + \eta \Longrightarrow l - \varepsilon < f(x) < l + \varepsilon.$$

Dans ce cas, cette limite est unique et on note :

$$\lim_{x \to x_0} f(x) = l \text{ ou } \lim_{x_0} f = l.$$

2) On suppose qu'il existe $\alpha > 0$ tel que $]x_0, x_0 + \alpha[\subset D_f$. f admet une limite $l_d \in \mathbb{R}$ à droite de x_0 si :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in D_f, x_0 < x < x_0 + \eta \Longrightarrow |f(x) - l_d| < \varepsilon,$$

Dans ce cas, cette limite est unique et on note :

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

3) On suppose qu'il existe $\alpha > 0$ tel que $]x_0 - \alpha, x_0[\subset D_f$. f admet une limite $l_g \in \mathbb{R}$ à gauche de x_0 si :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in D_f, x_0 - \eta < x < x_0 \Longrightarrow |f(x) - l_g| < \varepsilon,$$

Dans ce cas, cette limite est unique et on note :

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

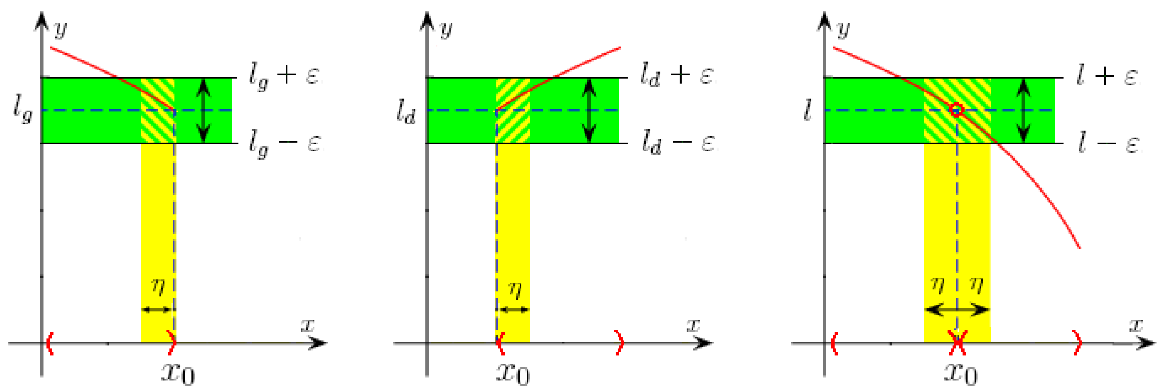

FIGURE 5 – Limites à gauche de x_0 (panel gauche), à droite de x_0 (panel du milieu), limite en x_0 (panel droit). Le graphe de la fonction f est représenté en rouge. Le domaine de définition de f contient respectivement $]x_0 - \alpha, x_0[$ (panel gauche), $]x_0, x_0 + \alpha[$ (panel du milieu), $]x_0 - \alpha, x_0[\cup]x_0, x_0 + \alpha[$ (panel droit) avec $\alpha > 0$. Dans chacun des cas, quelque soit la largeur de la bande verte (2ε), on peut réduire la bande jaune (η ou 2η) tel que le graphe de f au dessus de la bande jaune soit inclus dans la bande verte.

Remarque 2 :

- La valeur de la fonction en x_0 n'a aucune importance pour la notion de limite : la fonction peut même ne pas être définie en x_0.
- La limite (resp. à droite, à gauche) d'une fonction en x_0 dépend de son comportement local autour (resp. à droite, à gauche) de ce point : la définition exprime en termes mathématiques la propriété "on peut rendre $f(x)$ aussi proche qu'on veut de l en choisissant x assez proche de x_0 (resp. avec $x > x_0$, avec $x < x_0$)."

– Attention : l'existence d'une limite à droite et à gauche de x_0 ne signifie pas qu'elles sont égales.
– Supposons que f admet une limite l_d à droite de x_0 et une limite l_g à gauche de x_0. Alors f admet une limite en x_0 si et seulement si $l_d = l_g$. Dans ce cas, $\lim_{x \to x_0} f(x) = l_g = l_d$.

Exercice 1 : Montrer que, si elle existe, la limite de f à droite (resp. à gauche) de x_0 est unique.
Montrer que, si la limite de f (au sens du 1)) en x_0 existe, alors elle est unique.
Nota : pour montrer ce type de résultat d'unicité, on suppose que deux réels l_1 et l_2 vérifie la définition et on montre que $l_1 = l_2$.

Exemple 5 : Soient la fonction $f : x \to 1 + \sqrt{|x|}$ définie sur $\mathbb{R}$ et $x_0 = 0$. Pour montrer l'existence des limites de f à droite et à gauche de 0, on écrit :

$$\forall x \geq 0, f(x) = 1 + \sqrt{x}, \qquad \forall x \leq 0, f(x) = 1 + \sqrt{-x}.$$

Soit $\varepsilon > 0$. Posons $\eta = \varepsilon^2$. Alors, montrons que la limite de f à droite de 0 est 1 :

$$0 < x < \eta \Longrightarrow |f(x) - 1| = |1 + \sqrt{x} - 1| = \sqrt{x} < \sqrt{\eta} = \varepsilon.$$

De même la limite à gauche de 0 est 1 puisque :

$$-\eta < x < 0 \Longrightarrow 0 < -x < \eta \Longrightarrow |f(x) - 1| = |1 + \sqrt{-x} - 1| = \sqrt{-x} < \sqrt{\eta} = \varepsilon.$$

Ceci prouve (voir Fig. 6) que

$$\lim_{x \to 0} f(x) = 1.$$

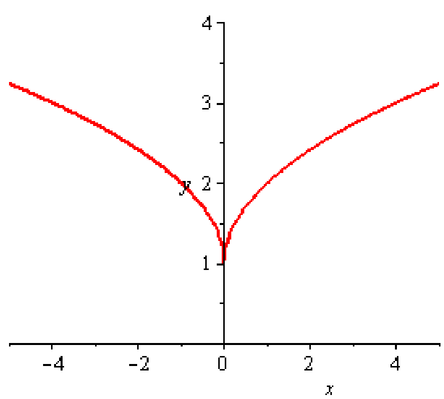

FIGURE 6 – Graphe de la fonction $f : x \to 1 + \sqrt{|x|}$ définie sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$. Les limites à gauche et à droite de 0 valent 1. On en déduit que f admet la limite 1 en 0.

Définition 6 : On suppose qu'il existe $\alpha > 0$ tel que $]x_0-\alpha, x_0[\cup]x_0, x_0+\alpha[\subset D_f$. Alors f tend vers $+\infty$ en x_0 si :

$$\forall M > 0, \exists \eta > 0, \forall x \in I\backslash\{x_0\}, |x-x_0| < \eta \Longrightarrow f(x) > M.$$

On note alors :

$$\lim_{x\to x_0} f(x) = +\infty \text{ ou } \lim_{x_0} f = +\infty.$$

f tend vers $-\infty$ en x_0 si $-f$ tend vers $+\infty$ en x_0 (Exercice : écrire la définition en utilisant les quantificateurs comme ci-dessus). On déduit les notions de limite infinie à droite (resp. à gauche) de x_0 en réduisant les intervalles comme dans la définition précédente.

Définition 7 : On suppose qu'il existe $a \in \mathbb{R}$ tel que $]a, +\infty[\subset D_f$. Alors f admet une limite $l \in \mathbb{R}$ en $+\infty$ si :

$$\forall \varepsilon > 0, \exists u > a, \forall x > u, |f(x) - l| < \varepsilon.$$

On note alors :

$$\lim_{x\to+\infty} f(x) = l \text{ ou } \lim_{+\infty} f = l.$$

On suppose $]-\infty, a[\subset D_f$. f tend vers l en $-\infty$ si $x \to f(-x)$ tend vers l en $+\infty$ (Exercice : écrire la définition en utilisant les quantificateurs comme ci-dessus).

Exercice 2 : Ecrire les définitions d'une fonction f tendant vers $+\infty$ en $+\infty$, vers $+\infty$ en $-\infty$, vers $-\infty$ en $+\infty$, vers $-\infty$ en $-\infty$.

La limite d'une expression numérique réelle P_x dépendant de x est la limite d'une fonction $x \to P_x$ définie sur un intervalle qui convient à l'application de la définition correspondante ci-dessus. Le caractère local de la notion de limite assure que la valeur de la limite ne dépend pas du domaine de définition (essentiellement caractérisé par l'ensemble de départ) choisi pour la fonction $x \to P_x$ tant qu'il contient une partie convenable à la vérification de l'existence de la limite.

II.2 Continuité

Dans toute la suite de ce chapitre, I désigne un intervalle de $\mathbb{R}$ de bornes a et b avec $-\infty \leq a < b \leq +\infty$ (I non vide et non réduit à un point) et f une fonction définie sur $I \subset D_f$ à valeurs réelles :

$$\begin{array}{rcl} f : I \subset D_f \subset \mathbb{R} & \to & \mathbb{R} \\ x & \to & f(x). \end{array}$$

Les notions de voisinage, de point intérieur et de point au bord d'un ensemble sont très intuitives. En voici les définitions mathématiques pour les points et parties de $\mathbb{R}$.

Définition 8 : On appelle **voisinage** d'un point $x_0 \in \mathbb{R}$ une partie (sous-ensemble) de $\mathbb{R}$ contenant un intervalle $]x_0 - \alpha, x_0 + \alpha[$ où $\alpha > 0$.
En particulier, les intervalles de $]x_0 - \alpha, x_0 + \alpha[$ où $\alpha > 0$ sont des voisinages de x_0.

Définition 9 : Soient $A \subset \mathbb{R}$ et $x_0 \in \mathbb{R}$.

- x_0 est un **point à l'intérieur** de A si

$$\exists \alpha > 0,]x_0 - \alpha, x_0 + \alpha[\subset A.$$

  i.e. s'il existe un voisinage de x_0 inclus dans A.
- x_0 est un **point au bord** de A si

$$\forall \alpha > 0, (\exists a \in]x_0 - \alpha, x_0 + \alpha[\setminus \{x_0\}, a \in A) \text{ et } (\exists \in]x_0 - \alpha, x_0 + \alpha[, b \notin A)$$

  i.e. tout voisinage de x_0 contient au moins un point de A différent de x_0 et un point en dehors de A.

Exemple 6 :

- 0 est un point à l'intérieur de $]-1, 1[$ (également de tout intervalle (α, β) avec $\alpha < 0$ et $\beta > 0$) mais est un point au bord de $[0, 1[$.
- Un point peut être au bord d'un ensemble sans y appartenir. 0 est un point au bord de]0,1[puisque tout intervalle $] - \alpha, \alpha[$ où $\alpha > 0$ contient $\min(\alpha/2, 1/2) \in]0, 1[$ et contient également $-\alpha/2 \notin]0, 1[$.

Définition 10 : Soit $x_0 \in I$.
1) f est dite continue à gauche en $x_0 > a$ si f admet une limite à gauche de x_0 et :

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

2) f est dite continue à droite en $x_0 < b$ si f admet une limite à droite de x_0 et :

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

3) f est dite **continue** en $x_0 \in]a, b[$ si f est continue à droite et à gauche en x_0, i.e.

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

On revenant à la définition de limite, on obtient : f est continue en x_0 si et seulement si

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in I, |x - x_0| < \eta \Longrightarrow |f(x) - f(x_0)| < \varepsilon.$$

Cette propriété peut également s'écrire :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in I, x \in]x_0 - \alpha, x_0 + \alpha[\Longrightarrow f(x) \in]f(x_0) - \varepsilon, f(x_0) + \varepsilon[.$$

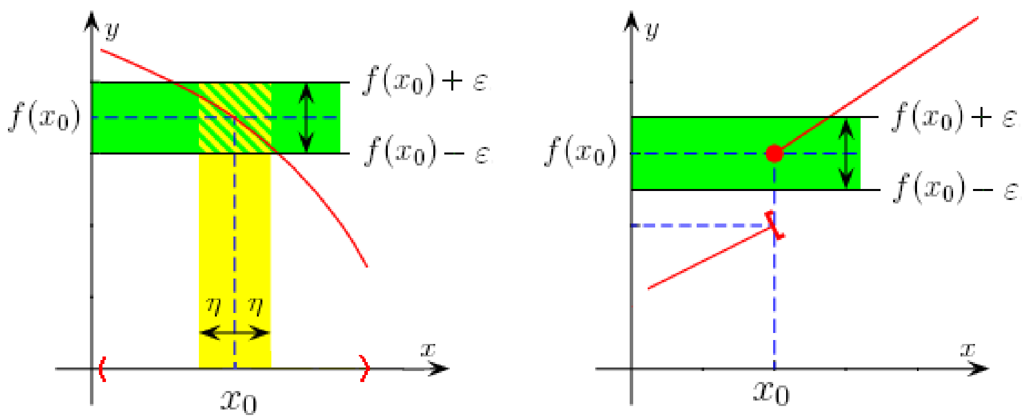

FIGURE 7 – Panel gauche : la fonction f dont le graphe est représenté en rouge est continue en x_0. Panel droit : la fonction f dont le graphe est représenté en rouge est discontinue en x_0. En effet, en choisissant ε assez petit (comme sur la figure), $\forall \eta > 0, \exists x \in]x_0 - \eta, x_0 + \eta[, f(x) \notin]f(x_0) - \varepsilon, f(x_0) + \varepsilon[$. Par exemple, on remarque $f(x_0 - \eta/2) < f(x_0) - \varepsilon$. Ceci exprime la négation de la continuité en x_0. Par contre, elle est continue à droite en x_0 mais pas à gauche en x_0.

Remarque 3 :

- f est continue en $x_0 \in]a, b[$ ssi elle admet une limite en x_0 qui vaut $f(x_0)$:

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

- Si $x_0 = a \in I$ est au bord de D_f, on identifie continuité et continuité à droite en x_0.
- Si $x_0 = b \in I$ est au bord de D_f, on identifie continuité et continuité à gauche en x_0.

Définition 11 :
1) f est dite continue sur $J \subset D_f$ si elle est continue en tout point de J.
2) Le domaine de continuité de la fonction f est la plus grande partie (pour l'inclusion) de D_f sur laquelle la fonction est continue.

II.3 Dérivée en un point

On rappelle que, dans toute la suite, I désigne un intervalle de $\mathbb{R}$ de bornes a et b avec $-\infty \leq a < b \leq +\infty$ et f une fonction définie sur $I \subset D_f$ à valeurs réelles.

Définition 12 : Soit $x_0 \in I$.

1) f est **dérivable à droite** en x_0 si la limite suivante existe et est finie :

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

Cette limite, notée $f'_d(x_0)$, est appelée **dérivée à droite** en x_0.

2) f est **dérivable à gauche** en x_0 si la limite suivante existe et est finie :

$$\lim_{\substack{x\to x_0\\ x<x_0}} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{\substack{h\to 0\\ h<0}} \frac{f(x_0+h)-f(x_0)}{h}.$$

Cette limite, notée $f'_g(x_0)$, est appelée **dérivée à gauche** en x_0.

3) f est **dérivable** en x_0 si elle est dérivable à droite et à gauche en x_0 et $f'_d(x_0) = f'_g(x_0)$, i.e. si la limite suivante existe et est finie :

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}.$$

Cette limite, notée $f'(x_0)$ ou $\frac{df}{dx}(x_0)$, est appelée **dérivée** de f en x_0.

La dérivée de f en x_0 est la limite de la pente de la droite passant par $(x_0, f(x_0))$ et $(x, f(x)) = (x_0+h, f(x_0+h))$ quand x tend vers x_0, i.e. $h \to 0$. Voir Fig. 8.

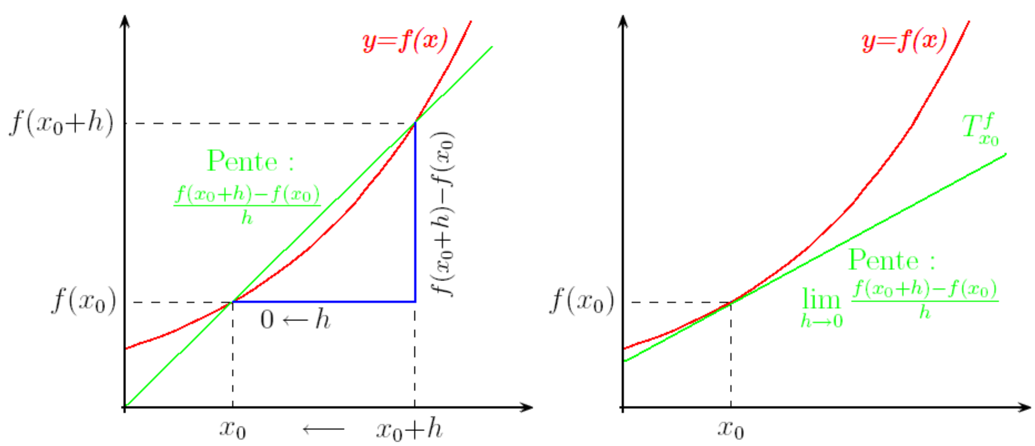

FIGURE 8 – Construction géométrique de la dérivée de f en x_0. Lorsque cette limite existe, la pente de la corde (droite verte du panel gauche) reliant $(x_0, f(x_0))$ et $(x_0+h, f(x_0+h))$ tend vers la pente de la tangente (droite verte du panel droit) : cette limite est la dérivée de f en x_0.

Si f est dérivable en x_0 alors elle est continue en x_0. Autrement dit, la continuité de f en x_0 est une **condition nécessaire** à la dérivabilité de f en x_0. Ainsi, une fonction

qui n'est pas continue en x_0 ne peut pas être dérivable en x_0.
Ceci apparait directement dans le calcul de la limite introduite dans la définition précédente : Supposons que f est dérivable en x_0, alors

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \text{ existe.}$$

Comme $\lim\limits_{x \to x_0} (x - x_0) = 0$, on doit nécéssairement avoir $\lim\limits_{x \to x_0} f(x) - f(x_0) = 0$ (dans le cas contraire, on aurait immédiatement la non-existence de la limite du taux d'accroissement ou une limite ∞). On en déduit $\lim\limits_{x \to x_0} f(x) = f(x_0)$ et f continue en x_0.

Attention : La continuité est une condition nécessaire mais certainement pas suffisante. $x \to \sqrt{|x|}$ est continue en 0 mais n'est dérivable ni à droite ni à gauche en 0.

II.4 Tangente

Définition 13 : Vecteur normal à une droite.
Un vecteur u du plan est dit **normal** à une droite Δ du plan muni d'un repère orthogonal (O, x, y), si une droite de direction $\vec{u}$ (et donc toute droite de direction $\vec{u}$) est perpendiculaire à Δ.

Proposition 1 : La droite Δ passant par le point $(x_0, y_0) \in \mathbb{R}^2$ dont $\vec{u} = (a, b) \in \mathbb{R}^2$ est un vecteur normal admet pour équation cartésienne :

$$\Delta : a(x - x_0) + b(y - y_0) = 0.$$

Définition 14 : Tangente à un graphe.
La droite de meilleure approximation du graphe de f au voisinage de x_0, si elle existe, est appellé la **tangente** au graphe de f en $(x_0, f(x_0))$. Voir Fig. 8, panel droit.

Proposition 2 : Si f est dérivable en x_0, alors la tangente $T^f_{x_0}$ au graphe de f en $(x_0, f(x_0))$ est bien définie, $(1, f'(x_0))$ en est un vecteur directeur et $(f'(x_0), -1)$ un vecteur normal. Elle admet pour équation :

$$T^f_{x_0} : y = f(x_0) + (x - x_0) f'(x_0), \quad x \in \mathbb{R}.$$

Si f est dérivable à droite (resp. à gauche) en x_0, la **demi-tangente à droite (**resp. **à gauche)** au graphe de f en $(x_0, f(x_0))$ est, par définition, la demi-droite d'équation :

$$y = f(x_0) + (x - x_0) f'_d(x_0), \quad x \geq x_0$$
$$(\text{resp. } y = f(x_0) + (x - x_0) f'_g(x_0), \quad x \leq x_0).$$

Exemple 7 : Considérons la fonction (voir Fig. 9) :

$$\begin{array}{rccc} f: & \mathbb{R} & \to & \mathbb{R} \\ & x & \to & \begin{cases} \frac{x}{2}\cos(x) & \text{si } x \geq 0 \\ x^2 - x & \text{si } x < 0 \end{cases} \end{array}$$

On peut vérifier que :

$$f'_g(0) = -2 \text{ et } f'_d(0) = \frac{1}{2}.$$

On en déduit que f n'est pas dérivable en 0 mais son graphe admet des demi-tangentes à gauche et à droite :

$$\begin{aligned} y &= -\frac{x}{2}, \quad x \leq 0, \\ y &= 2x, \quad x \geq 0. \end{aligned}$$

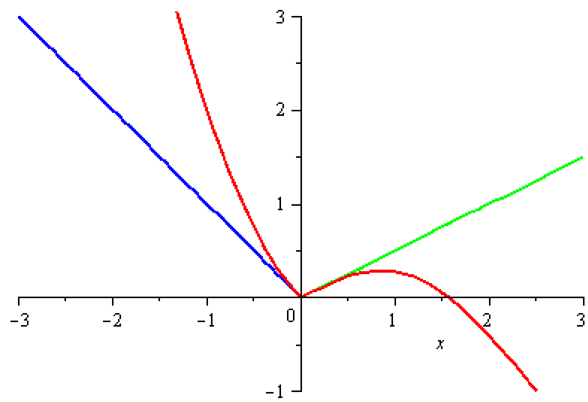

FIGURE 9 – Graphe (en rouge) de la fonction f définie dans l'Exemple 7 et ses demi-tangentes à gauche (en bleu) et à droite (en vert) en 0.

II.5 Fonction dérivée

Définition 15 :
La fonction f est dite dérivable sur $J \subset D_f$ si elle est dérivable en tout point de J.
Le **domaine de dérivabilité** de f est l'ensemble des points x_0 de D_f où :
- f est dérivable si x_0 est à l'intérieur de D_f,
- f est dérivable à droite ou à gauche si x_0 est au bord de D_f.

La **fonction dérivée**, appelée aussi simplement **la dérivée**, de f est alors définie par :

$$\begin{aligned} f' : J &\rightarrow \mathbb{R} \\ x &\rightarrow f'(x). \end{aligned}$$

La fonction dérivée de $f : x \rightarrow f(x)$ peut également être notée $\frac{df}{dx}$ et $f'(x_0) = \frac{df}{dx}(x_0)$.

Dérivée des fonctions usuelles : on note respectivement D_f et D'_f les domaines de définition et de dérivabilité de la fonction f.

f	D_f	f'	D'_f
$x \to cst \in \mathbb{R}$	$\mathbb{R}$	$x \to 0$	$\mathbb{R}$
$x \to x^n, n \in \mathbb{N}^*$	$\mathbb{R}$	$x \to nx^{n-1}$	$\mathbb{R}$
$x \to x^{-n}, n \in \mathbb{N}^*$	$\mathbb{R}^*$	$x \to -nx^{-n-1}$	$\mathbb{R}^*$
$x \to x^\alpha, \alpha > 1$	$\mathbb{R}_+$	$x \to \alpha x^{\alpha-1}$	$\mathbb{R}_+$
$x \to x^\alpha, 0 < \alpha < 1$	$\mathbb{R}_+$	$x \to \alpha x^{\alpha-1}$	$\mathbb{R}_+^*$
$x \to x^\alpha, \alpha < 0$	$\mathbb{R}_+^*$	$x \to \alpha x^{\alpha-1}$	$\mathbb{R}_+^*$
$x \to \ln x$	$\mathbb{R}_+^*$	$x \to \frac{1}{x}$	$\mathbb{R}_+^*$
$x \to e^x$	$\mathbb{R}$	$x \to e^x$	$\mathbb{R}$
$x \to \sin x$	$\mathbb{R}$	$x \to \cos x$	$\mathbb{R}$
$x \to \cos x$	$\mathbb{R}$	$x \to -\sin x$	$\mathbb{R}$
$x \to \tan x$	$\mathbb{R}\backslash\left\{\frac{\pi}{2} + k\pi \vert k \in \mathbb{Z}\right\}$	$x \to 1 + \tan^2 x$	$\mathbb{R}\backslash\left\{\frac{\pi}{2} + k\pi \vert k \in \mathbb{Z}\right\}$
$x \to \arcsin x$	$[-1, 1]$	$x \to \frac{1}{\sqrt{1-x^2}}$	$]-1, 1[$
$x \to \arccos x$	$[-1, 1]$	$x \to -\frac{1}{\sqrt{1-x^2}}$	$]-1, 1[$
$x \to \arctan x$	$\mathbb{R}$	$x \to \frac{1}{1+x^2}$	$\mathbb{R}$
$x \to \sinh x = \frac{1}{2}(e^x - e^{-x})$	$\mathbb{R}$	$x \to \cosh x$	$\mathbb{R}$
$x \to \cosh x = \frac{1}{2}(e^x + e^{-x})$	$\mathbb{R}$	$x \to \sinh x$	$\mathbb{R}$
$x \to \tanh x = \frac{\sinh x}{\cosh x}$	$\mathbb{R}$	$x \to 1 - \tanh^2 x$	$\mathbb{R}$

Proposition 3 : Dérivée d'une somme, d'un produit, d'un quotient.
Soient f et g dérivables sur un intervalle I de $\mathbb{R}$ et $\lambda \in \mathbb{R}$. Alors $f + g$, λf, fg sont dérivables sur I et :

$$\begin{aligned} (f+g)' &= f' + g', \\ (\lambda f)' &= \lambda f', \\ (fg)' &= f'g + fg'. \end{aligned}$$

Les fonctions $\frac{1}{g}$ et $\frac{f}{g}$ sont dérivables sur $I\backslash\{x \in I | g(x) = 0\}$ (i.e. partout où g ne s'annule pas, soit encore sur les domaines de définition de $\frac{1}{g}$ et $\frac{f}{g}$) :

$$\begin{aligned} \left(\frac{1}{g}\right)' &= -\frac{g'}{g^2}, \\ \left(\frac{f}{g}\right)' &= \frac{f'g - fg'}{g^2}. \end{aligned}$$

Proposition 4 : Dérivée d'une fonction composée.
Si f est dérivable sur I et g est une fonction dérivable sur $f(I)$, alors la fonction composée $g \circ f$ est dérivable sur I et :

$$(g \circ f)' = f'.(g' \circ f) \text{ i.e. } \forall x \in I, (g \circ f)'(x) = f'(x).g'(f(x)).$$

Proposition 5 : Dérivée d'une bijection réciproque.
Soit f une fonction continue sur un intervalle I et bijective de I dans $J = f(I)$. Si f est dérivable en x et $f'(x) \neq 0$, alors la bijection réciproque $f^{-1} : J \to I$ est dérivable en $y = f(x)$ et :

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}.$$

Ainsi, le domaine de dérivabilité de f^{-1} est

$$D'_{f^{-1}} = \left\{y \in J | f'(f^{-1}(y)) \neq 0\right\}$$

et sa dérivée est :

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

Exemple 8 : illustré par Fig. 10. La fonction :

$$\begin{array}{rcl} f : [0, +\infty[& \to & [-2, +\infty[\\ x & \to & (x-1)^3 - 1 \end{array}$$

est bijective. La bijection réciproque est donnée par :

$$\begin{array}{rcl} f^{-1} : [-2, +\infty[& \to & [0, +\infty[\\ x & \to & 1 + \mathrm{sign}(x+1).|x+1|^{1/3} = \left\{ \begin{array}{ll} 1 + (x+1)^{1/3} & \text{si } x \geq -1 \\ 1 - (-x-1)^{1/3} & \text{si } -2 \leq x \leq -1 \end{array} \right. \end{array}$$

La fonction f est dérivable sur $[0, +\infty[$ et sa fonction dérivée est donnée par :

$$\begin{array}{rcl} f' : [0, +\infty[& \to & [-2, +\infty[\\ x & \to & 3(x-1)^2. \end{array}$$

On a : $f'(x) = 0 \Longleftrightarrow x = 1$. Ainsi, f^{-1} est dérivable sur $f([0, +\infty[) \backslash \{f(1)\}$ où $f(1) = -1$, soit $[-2, +\infty[\backslash\{-1\}$. Exercice : calculer cette dérivée.

Définition 16 : Dérivées successives.

1) Définition par récurrence : la fonction dérivée de f', appelée **dérivée seconde** de f, est notée f''ou $f^{(2)}$. Pour k entier positif, la **dérivée k-ème** de f, notée $f^{(k)}$, est la dérivée de la dérivée $(k-1)$-ème de f : $f^{(k)} = \left(f^{(k-1)}\right)'$.
On note également $f = f^{(0)}$.

2) La fonction f est dite de **classe C^k**, $k \in \mathbb{N}$, en x_0 (resp. sur $J \subset D_f$) si $f^{(k)}$ est définie et continue en x_0 (resp. sur J). Si f est de classe C^k alors elle est de classe C^p pour tout $p \leq k$.

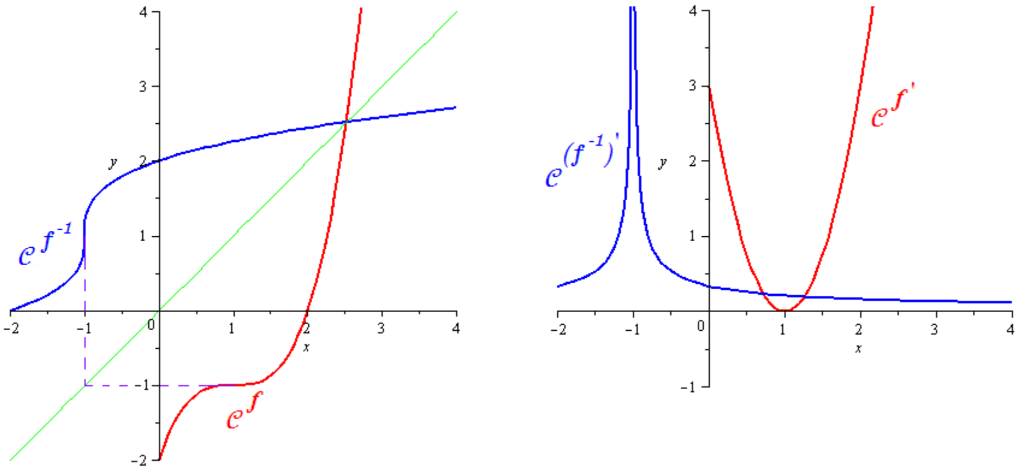

FIGURE 10 – La fonction f est donnée par l'Exemple 8. Panel gauche : graphes des bijections f (en rouge) et f^{-1} (en bleu). Les deux graphes sont symétriques par rapport à la première bissectrice du plan (en vert). Panel droit : graphes des fonctions dérivées de f (en rouge) et de f^{-1} (en bleu). La dérivée de f s'annule uniquement en 1. On en déduit que f^{-1} est dérivable sur $f(D_f)\backslash\{f(1)\}$ avec $f(1) = -1$.

II.6 Sens de variations et extrema

Définition 17 : Soit f une fonction d'une variable réelle à valeurs réelles et un intervalle $I \subset D_f$.

- f est dite croissante sur I si

$$\forall (x, y) \in I \times I, x < y \Longrightarrow f(x) \leq f(y).$$

- f est dite décroissante sur I si

$$\forall (x, y) \in I \times I, x < y \Longrightarrow f(x) \geq f(y).$$

- f est dite strictement croissante sur I si

$$\forall (x, y) \in I \times I, x < y \Longrightarrow f(x) < f(y)$$

.

- f est dite strictement décroissante sur I si

$$\forall (x, y) \in I \times I, x < y \Longrightarrow f(x) > f(y).$$

f est dite monotone sur I si elle est croissante sur I ou décroissante sur I (sans préciser lequel) et strictement monotone sur I si elle est strictement croissante sur I ou strictement décroissante I (sans préciser lequel).

Proposition 6 : Soit f une fonction d'une variable réelle à valeurs réelles et un intervalle $I \subset D_f$. Si f est strictement monotone sur I alors f induit une bijection de I sur $f(I)$.

Exercice 3 : Démontrer ce résultat (ce n'est pas difficile : il suffit de revenir aux définitions).

Proposition 7 : Soit f une fonction dérivable sur un intervalle I de $\mathbb{R}$. Alors :
- f est croissante sur I si et seulement si $f' \geq 0$ sur I.
- f est décroissante sur I si et seulement si $f' \leq 0$ sur I.
- Si $f' > 0$ sur I alors f est strictement croissante sur I.
- Si $f' < 0$ sur I alors f est strictement décroissante sur I.

Remarque 4 : Les deux derniers points ne sont pas des équivalences. En effet, une fonction peut être strictement monotone sur un intervalle tout en ayant une dérivée nulle en un point de cet intervalle. Par exemple : $x \to x^3$ est strictement croissante sur $\mathbb{R}$ mais sa dérivée $x \to 3x^2$ s'annule en 0.

Définition 18 : **Extrema d'une fonction.**
f admet un minimum (resp. maximum) local en x_0 s'il existe un voisinage V de x_0 tel que pour tout $x \in V \cap D_f$, $f(x) \geq f(x_0)$ (resp. $f(x) \leq f(x_0)$).
f admet un minimum (resp. maximum) global en x_0 si pour tout $x \in D_f$, $f(x) \geq f(x_0)$ (resp. $f(x) \leq f(x_0)$).
Les extrema (respectivement globaux ou locaux) d'une fonction sont les minima et les maxima de cette fonction (resp. globaux ou locaux).

Théorème 1 :
1) Si f est dérivable en en un point x_0 intérieur à D_f, et admet un extremum local en x_0, alors $f'(x_0) = 0$.
2) Si f est deux fois dérivable en un point x_0 intérieur à D_f et telle que :
- $f'(x_0) = 0$,
- $f''(x_0) \neq 0$,

alors f admet un extremum local en x_0. Cet extremum est :
- un maximum si $f''(x_0) < 0$,
- un minimum si $f''(x_0) > 0$.

Remarque 5 :
- le 1) concerne les points de l'intervalle $]a, b[$ ouvert mais pas les bornes a et b. Si a (resp. b) est un réel et appartient à l'intervalle de définition I, la fonction f peut admettre un extremum local (ou même global) en ce point sans que sa dérivée s'annule. Par exemple :

$$\begin{aligned} f : [0,1] &\rightarrow \mathbb{R} \\ x &\rightarrow x. \end{aligned}$$

  Puisque, pour tout $x \in [0,1]$, $f(x) \geq 0$, f admet un extremum (global) en 0 mais sa dérivée ne s'annule pas en 0.
- le 1) donne une condition suffisante mais pas nécessaire. Une fonction peut tout à fait avoir une dérivée nulle en un point et ne pas admettre d'extremum local. Toujours le même exemple : $x \rightarrow x^3$

II.7 Développements limités

Définition 19 : Soit f une fonction définie au voisinage de x_0, i.e. sur $]x_0 - \alpha, x_0 + \alpha[$, $\alpha > 0$. f admet un **développement limité à l'ordre** $n \in \mathbb{N}$ **en** x_0, noté $DL_n^f(x_0)$, s'il existe $n+1$ réels a_i, $i \in \mathbb{N} \cap [0, n] = \{0, 1, ..., n\}$, et une fonction ε définie sur $]-\alpha, \alpha[$ tels que $\lim\limits_{0} \varepsilon = 0$ et :

$$\forall h \in]-\alpha, \alpha[, f(x_0 + h) = \sum_{i=0}^{n} a_i h^i + h^n \varepsilon(h) \quad (1)$$

qui peut encore s'écrire, en posant $h = x - x_0$:

$$\forall x \in]x_0 - \alpha, x_0 + \alpha[, f(x) = \sum_{i=0}^{n} a_i (x - x_0)^i + (x - x_0)^n \varepsilon(x - x_0).$$

Le $DL_n^f(x_0)$ est une fonction définie sur un voisinage de x_0 **égale** à la restriction de la fonction f à ce voisinage.

Attention : Le reste est indispensable dans l'écriture du DL. D'une part, la partie polynomiale (également appelée partie principale) ne donne qu'une **approximation** de la fonction f au voisinage de x_0 par un polynôme de degré n (voir Fig. 11). Sans le reste, il n'y a donc pas égalité avec la fonction f. D'autre part, on peut lire l'ordre n du DL directement sur la forme du reste : $h^n \varepsilon(h)$, avec $\varepsilon \underset{0}{\rightarrow} 0$.

Proposition 8 : S'il existe, le $DL_n^f(x_0)$ d'une fonction f est unique, i.e. il existe un unique n-uplet $(a_i)_{1 \leq i \leq n}$ vérifiant (1). En outre, pour $k \leq n$, le $DL_k^f(x_0)$ s'obtient en tronquant le $DL_n^f(x_0)$:

$$f(x_0 + h) \underset{h \to 0}{=} \sum_{i=0}^{k} a_i h^i + h^k \varepsilon(h).$$

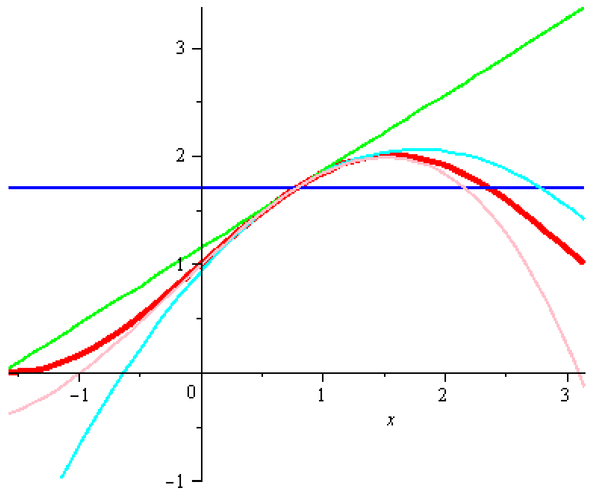

FIGURE 11 – Graphes de la fonction $x \to 1 + \sin(x)$ (en rouge) et de ses approximations au voisinage de $\pi/4$ par les parties principales des $DL_i^f(\pi/4)$ pour i=0 (bleu foncé), i=1 (vert), i=2 (bleu ciel), i=3 (rose).

Les fonctions ε introduites dans le reste ne sont pas les mêmes, mais ont une limite nulle en 0. Par abus de notation, dans le contexte d'un DL, on peut utiliser le même nom de fonction.

Proposition 9 : DL_0 et DL_1.
1) Soit f une fonction définie au voisinage de x_0. f admet un $DL_0^f(x_0)$ si et seulement si f admet une limite en x_0. Le $DL_0^f(x_0)$ s'écrit alors :

$$f(x_0 + h) = a_0 + \varepsilon(h), \quad \varepsilon \underset{0}{\to} 0 \quad \text{avec } a_0 = \lim_{x_0} f.$$

2) Soit f une fonction continue en x_0. f admet un $DL_1^f(x_0)$ si et seulement si f est dérivable en x_0. Le $DL_1^f(x_0)$ de f s'écrit alors :

$$f(x_0 + h) = f(x_0) + hf'(x_0) + h\varepsilon(h), \quad \varepsilon \underset{0}{\to} 0.$$

Remarque 6 :
- Si f n'est pas continue en x_0 mais admet une limite en x_0, le terme constant qui constitue la partie principale du $DL_0^f(x_0)$ est égal à la limite de f en x_0 et non $f(x_0)$.
- La partie principale du $DL_1^f(x_0)$ est l'approximation au permier ordre de f au voisinage de x_0, i.e. l'approximation par une fonction affine. Ainsi, elle permet d'obtenir directement l'équation de la tangente au graphe de f en x_0 (voir Fig. 11) :

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

DL des fonctions usuelles en 0. n représente un entier naturel quelconque :

e^x	$\underset{x\to 0}{=}$	$1+x+\frac{x^2}{2}+\frac{x^3}{3!}+...+\frac{x^n}{n!}+x^n\varepsilon(x)$	$=\sum_{i=0}^{n}\frac{x^i}{i!}+x^n\varepsilon(x)$,
$\sinh x$	$\underset{x\to 0}{=}$	$x+\frac{x^3}{3!}+...+\frac{x^{2n+1}}{(2n+1)!}+x^{2n+1}\varepsilon(x)$	$=\sum_{i=0}^{n}\frac{x^{2i+1}}{(2i+1)!}+x^{2n+1}\varepsilon(x)$,
$\cosh x$	$\underset{x\to 0}{=}$	$1+\frac{x^2}{2}+...+\frac{x^{2n}}{(2n)!}+x^{2n}\varepsilon(x)$	$=\sum_{i=0}^{n}\frac{x^{2i}}{(2i)!}+x^{2n}\varepsilon(x)$,
$\sin x$	$\underset{x\to 0}{=}$	$x-\frac{x^3}{3!}+...+\frac{(-1)^n x^{2n+1}}{(2n+1)!}+x^{2n+1}\varepsilon(x)$	$=\sum_{i=0}^{n}\frac{(-1)^i x^{2i+1}}{(2i+1)!}+x^{2n+1}\varepsilon(x)$,
$\cos x$	$\underset{x\to 0}{=}$	$1-\frac{x^2}{2}+...+\frac{(-1)^n x^{2n}}{(2n)!}+x^{2n}\varepsilon(x)$	$=\sum_{i=0}^{n}\frac{(-1)^i x^{2i}}{(2i)!}+x^{2n}\varepsilon(x)$,
$\ln(1+x)$	$\underset{x\to 0}{=}$	$x-\frac{x^2}{2}+\frac{x^3}{3}+...+\frac{(-1)^{n+1}x^n}{n!}+x^n\varepsilon(x)$	$=\sum_{i=1}^{n}\frac{(-1)^{i+1}x^i}{i}+x^n\varepsilon(x)$,
$(1+x)^\alpha$	$\underset{x\to 0}{=}$	$1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+...+\frac{\alpha(\alpha-1)...(\alpha-n+1)}{n!}x^n+x^n\varepsilon(x)$ $=1+\sum_{i=1}^{n}\left[\frac{x^i}{i!}\prod_{k=0}^{i-1}(\alpha-k)\right]+x^n\varepsilon(x)$.	

Les DL du tableau ci-dessus sont donnés en 0. On peut en déduire le calcul des DL en un point quelconque x_0 en utilisant un changement de variable $x = x_0+h$, i.e. $h = x-x_0$. Ainsi, $x \to x_0$ quand $h \to 0$. On utilise alors les propriétés et formules spécifiques à chaque fonction pour faire apparaître les DL connus en 0.

Exemple 9 :

– Le $DL_n(x_0)$ de $x \to e^x$ s'obtient en écrivant :

$$e^x = e^{x_0+h} = e^{x_0}e^h \underset{h\to 0}{=} e^{x_0}\left(\sum_{i=0}^{n}\frac{h^i}{i!}\right)+h^n\varepsilon(h)$$

On peut se contenter de cette écriture puisque la partie polynomiale d'un DL en x_0 s'écrit en utilisant des monômes de $(x-x_0) = h$. On peut aussi remplacer h par $x-x_0$ mais en aucun cas on ne développe.

– Le $DL_n(x_0)$ (où $x_0 > 0$) de $x \to \ln x$ s'obtient en écrivant :

$$\ln(x) = \ln(x_0+h) = \ln\left[x_0\left(1+\frac{h}{x_0}\right)\right] = \ln x_0 + \ln\left(1+\frac{h}{x_0}\right)$$
$$\underset{h\to 0}{=} \ln x_0 + \sum_{i=0}^{n}\frac{(-1)^{i+1}}{i}\left(\frac{h}{x_0}\right)^i + h^n\varepsilon(h)$$

II.8 Formules de Taylor

Théorème 2 : Taylor-Lagrange.
Si f est de classe C^n sur un intervalle $[a,b]$ admettant une dérivée n-ème sur $]a,b[$. Alors :

$$\begin{aligned}\exists c \in]a,b[, f(b) &= \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(c) \\ &= f(a) + (b-a)f'(a) + ... + \frac{(b-a)^n}{n!} f^{(n)}(a) + \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(c).\end{aligned}$$

En posant $b = a+h$, la propriété précédente s'écrit :

$$\exists \theta]0,1[, f(a+h) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} h^k + \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(a+\theta h).$$

La formule de Taylor-Lagrange est un résultat global : a et b peuvent être aussi éloignés qu'on veut. Cependant, la seule propriété connue sur le réel c est $c \in]a,b[$. Ainsi, plus a et b sont éloignés, moins on a de contrôle sur le reste. On utilise souvent cette formule quand on peut borner la dérivée $f^{(n+1)}$ sur $]a,b[$.

Théorème 3 : Taylor-Young.
Si une fonction f admet une dérivée n-ème en x_0, alors elle admet un $DL_n^f(0)$ qui s'écrit :

$$\begin{aligned}f(x_0+h) &= \sum_{i=0}^{n} \frac{f^{(i)}(x_0)}{i!} h^i + h^n \varepsilon(h) \\ &= f(x_0) + f'(x_0)h + \frac{f''(x_0)}{2!} h^2 + ... + \frac{f^{(n)}(x_0)}{n!} h^n + h^n \varepsilon(h), \qquad \varepsilon \underset{0}{\to} 0.\end{aligned}$$

La formule de Taylor-Young est un résultat local. Elle permet de connaître les coefficients du DL_n pour la plupart des fonctions qui en admettent un (celles n fois dérivables au point considéré).

III Fonctions de deux variables réelles

Dans cette section, f désigne une fonction de deux variables réelles à valeurs réelles (ensemble de départ $E \subset \mathbb{R}^2$, ensemble d'arrivée $F \subset \mathbb{R}$). En revenant à la définition d'une "fonction", on obtient immédiatement que le graphe G de f est une partie de $E \times F \subset \mathbb{R}^3$. Le domaine de définition D_f de f est la partie de $\mathbb{R}^2$ formée des $(x,y) \in E \subset \mathbb{R}^2$ tels qu'il existe $z \in F \subset \mathbb{R}$ tel que $(x,y,z) \in G$. Cet élement z (unique pour chaque $(x,y) \in D_f$) est noté $f(x,y) \in F$.

Exemple 10 : Quelques types de fonctions de deux variables essentielles en Economie (fonctions de production ou d'utilité) :

Fonction de type linéaire :	$f(x,y) = ax + by,$	$(a,b) \in \mathbb{R}^2.$
Fonction de type Cobb-Douglas :	$f(x,y) = ax^b y^c,$	$(a,b,c) \in \mathbb{R}^3.$
Fonction de type Leontieff :	$f(x,y) = \min(ax, by),$	$(a,b) \in \mathbb{R}^2.$
Fonction à élasticité de substitution constante :	$f(x,y) = k(cx^{-a} + dy^{-b})^{b/a},$	$a \in \mathbb{R}*, (b,c,d) \in \mathbb{R}^3.$

III.1 Lignes de niveau

Définition 20 : La **ligne de niveau** ou **courbe de niveau** $c \in \mathbb{R}$ de f, notée $\mathcal{L}_c^f$, est l'ensemble

$$\{(x,y) \in D_f | f(x,y) = c\}.$$

Il s'agit d'une partie de $\mathbb{R}^2$ qui peut être représentée dans le plan.
Les lignes de niveau d'une fonction de production sont appelées les **isoquantes**, celles d'une fonction d'utilité sont appelées les **courbes d'indifférence**.

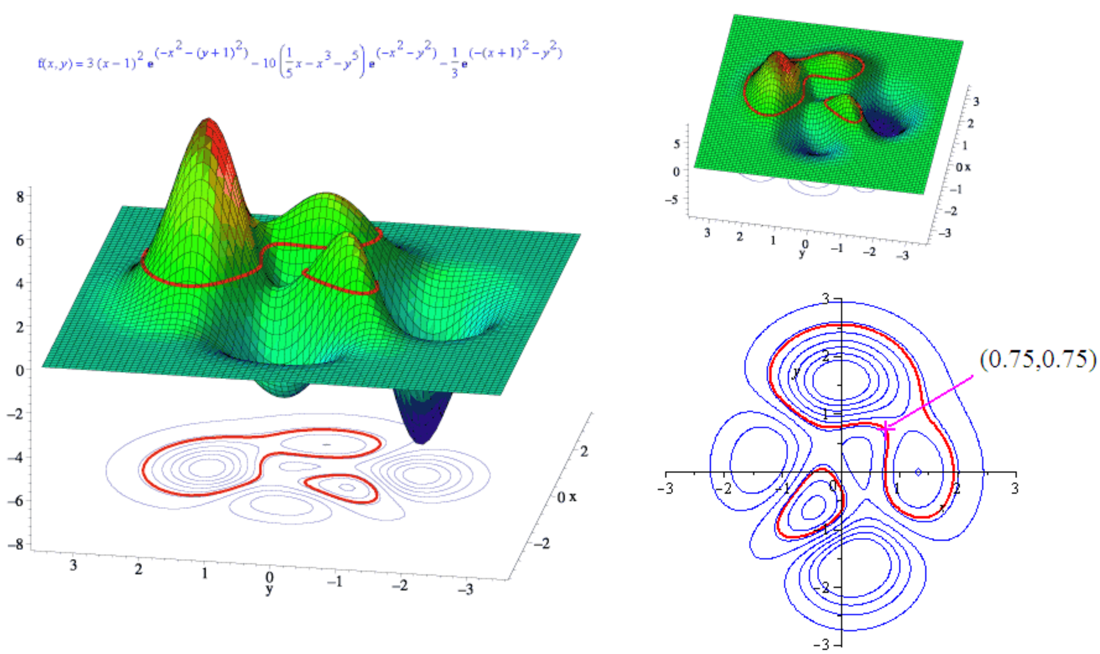

FIGURE 12 – Graphe d'une fonction f de deux variables (représentation sous deux angles différents) et construction de la ligne de niveau $c = f(0.75, 0.75)$: on construit l'intersection du graphe de f avec le plan $z = c$ puis on la projette selon l'axe (O, z) sur le plan (O, x, y) : cette projection (en rouge) est la ligne de niveau c de f et, en particulier, on retrouve sur $(0.75, 0.75)$ sur cette courbe. En bleu, d'autres lignes de niveau de la fonction f.

La représentation du graphe de la fonction f est une nappe $\mathcal{N}$ dans $\mathbb{R}^3$. Considérons le plan d'équation $z = c \in \mathbb{R}$ (plan horizontal dans (O, x, y, z)). L'intersection de $\mathcal{N}$ avec ce plan est donc $\{(x, y, c)|f(x, y) = c\}$. La projection (selon l'axe (O, z)) de cette intersection sur le plan (O, x, y) est donc $\{(x, y)|f(x, y) = c\}$: il s'agit de la ligne de niveau c de f. Voir Fig. 12 et 13.

III.2 Formes des lignes de niveau

La figure 13 donne des exemples de formes de lignes de niveau pour les types présentés dans l'Exemple 10 :

- a) fonction de type linéaire.
- b) fonction de type Léontieff.
- c) et d) fonction de type Cobb-Douglas.
- e) et f) fonction à élasticité de substitution constante.

III.3 Transformations monotones

Théorème 4 : Transformation strictement monotone.
Soit u une fonction d'une variable réelle à valeurs réelles strictement monotone sur $f(D_f)$. Alors, l'ensemble des courbes de niveau de f (quand c parcourt $\mathbb{R}$) est égal à l'ensemble des courbes de niveau de $u \circ f$:

$$\left\{\mathcal{L}_c^f \,|c \in \mathbb{R}\right\} = \left\{\mathcal{L}_c^{u\circ f} \,|c \in \mathbb{R}\right\}.$$

Plus précisément, u induit une bijection entre $f(D_f)$ et $(u \circ f)(D_f)$: notons u^{-1} la bijection réciproque. Alors :

$$\begin{array}{llll} \forall c \in f(D_f), & \mathcal{L}_c^f = \mathcal{L}_{u(c)}^{u\circ f}, & \forall c \in \mathbb{R}\backslash f(D_f), & \mathcal{L}_c^f = \emptyset, \\ \forall c' \in \mathbb{R}, & \mathcal{L}_{c'}^{u\circ f} = \mathcal{L}_{u^{-1}(c')}^{f}, & \forall c' \in \mathbb{R}\backslash (u\circ f)(D_f), & \mathcal{L}_c^{u\circ f} = \emptyset. \end{array}$$

Les fonctions f et $u \circ f$ sont dites <u>**transformations monotones l'une de l'autre**</u>.

Exemple 11 : Soient

$$\begin{array}{rcl} f : \left(\mathbb{R}_+^*\right)^2 & \to & \mathbb{R} \\ (x, y) & \to & \sqrt{xy} \end{array} \quad \text{et} \quad \begin{array}{rcl} g : \left(\mathbb{R}_+^*\right)^2 & \to & \mathbb{R} \\ (x, y) & \to & \ln x + \ln y \end{array}$$

Montrons que f et g sont transformations monotones l'une de l'autre. Tout d'abord, elles ont le même domaine de définition $\left(\mathbb{R}_+^*\right)^2$. Considérons ensuite la fonction

$$\begin{array}{rcl} u : \mathbb{R} & \to & \mathbb{R} \\ z & \to & 2\ln z \end{array}$$

u est strictement monotone sur $f\left(\left(\mathbb{R}_+^*\right)^2\right) = \mathbb{R}$ et :

$$\forall (x, y) \in D_f, (u \circ f)(x, y) = 2 \ln \sqrt{xy} = \ln (xy) = \ln x + \ln y = g(x, y).$$

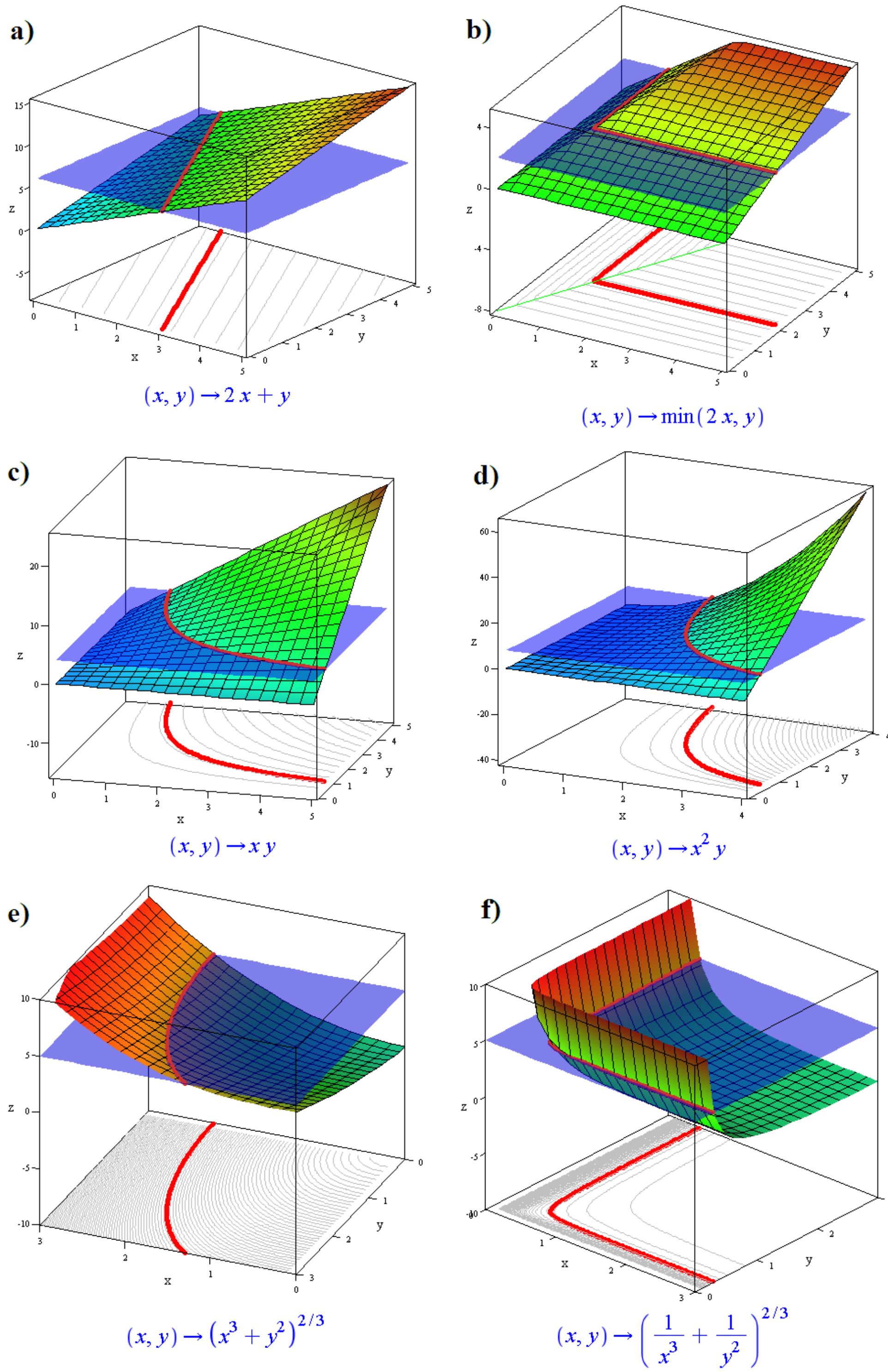

FIGURE 13 – Exemples de lignes de niveau. Dans chaque panel, la nappe est le graphe de la fonction, le plan bleu horizontal admet pour équation $z = c$, la courbe dans le plan (O, x, y) est la ligne de niveau c de f (projection de l'intersection entre le graphe de la fonction et le plan bleu), les courbes grises sont des lignes de niveaux différents.

On peut également remarquer que la fonction u est bijective et donc :

$$\forall (x,y) \in D_g = D_f, (u^{-1} \circ g)(x,y) = f(x,y)$$

Définition 21 : Biens désirables, indésirables, neutres.
Soit $f : (x,y) \to f(x,y)$ une fonction d'utilité. Si, pour $y = y_0$ (resp. $x = x_0$) fixé, la fonction $f_{y_0} : x \to f(x, y_0)$ (resp. $f_{x_0} : x \to f(x_0, y)$) est croissante, le bien x (resp. y) est appelé "**bien désirable**" ; si elle est décroissante, le bien x (resp. y) est appelé "**bien indésirable**" ; si elle est constante, le bien x (resp. y) est appelé "**bien neutre**".

Définition 22 : Point de saturation.
Soit $f : (x,y) \to f(x,y)$ une fonction d'utilité. (x_0, y_0) est appelé "**point de saturation**" de la fonction f si f admet un extremum local en (x_0, y_0) (cf. Section VI).

III.4 Rendement d'échelle, productivité moyenne

Définition 23 : Facteur d'échelle, Productivité moyenne.
Une fonction de production $Q : (x,y) \to Q(x,y)$ admet un **facteur d'échelle**
- **constant** si $Q(kx, ky) = kQ(x,y)$ pour tout $k > 1$,
- **croissant** si $Q(kx, ky) > kQ(x,y)$ pour tout $k > 1$,
- **décroissant** si $Q(kx, ky) < kQ(x,y)$ pour tout $k > 1$.

La **productivité apparente moyenne** du premier (resp. deuxième) facteur est définie par :

$$(x,y) \to \frac{Q(x,y)}{x} \quad \text{(resp. } (x,y) \to \frac{Q(x,y)}{y}\text{)}.$$

IV Fonctions de n variables réelles

Dans tout le chapitre IV, $n \in \mathbb{N} \backslash \{0,1\}$ et f désigne une fonction de n variables réelles à valeurs réelles : ensemble de départ $E \subset \mathbb{R}^n$, ensemble d'arrivée $F \subset \mathbb{R}$. En revenant à la définition de "fonction", on obtient immédiatement que le graphe G de f est une partie de $\mathbb{R}^{n+1}$. Le domaine de définition D_f de f est la partie de $\mathbb{R}^n$ formée des n-uplets $\mathbf{x} = (x_1, ..., x_n) \in E \subset \mathbb{R}^n$ tels qu'il existe $z \in F \subset \mathbb{R}$ tel que $(x_1, ..., x_n, z) \in G$. Cet élement z (unique) est noté $f(\mathbf{x}) = f(x_1, ..., x_n) \in F$. On note donc :

$$f : \begin{array}{ccc} D_f \subset E \subset \mathbb{R}^n & \to & F \subset \mathbb{R} \\ \mathbf{x} = (x_1, ..., x_n) & \to & f(\mathbf{x}) = f(x_1, ..., x_n) \end{array}.$$

Pour fixer les idées, nous donnons, à la suite de chaque définition, proposition et théorème, les constructions et les formules pour une fonction de deux variables :

$$f : \begin{array}{ccc} D_f \subset \mathbb{R}^2 & \to & F \subset \mathbb{R} \\ (x,y) & \to & f(x,y) \end{array}$$

Notation : Dans la suite, on note $\mathbf{0}_n = (0, ..., 0) \in \mathbb{R}^n$.

IV.1 Dérivées partielles d'ordre 1

Notation : Soit $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_n) \in \mathbb{R}^n$ fixé. On note alors, pour tout $i \in \mathbb{N} \cap [1, n] = \{1, 2, ..., n\}$:

- $\bar{\mathbf{x}}_{\setminus i} = (\bar{x}_1, ..., \bar{x}_{i-1}, \bar{x}_{i+1}, ..., \bar{x}_n) \in \mathbb{R}^{n-1}$
- $\varphi_{i,\bar{\mathbf{x}}_{\setminus i}}$ la fonction d'une variable réelle à valeurs réelles :

$$\begin{array}{rccl} \varphi_{i,\bar{\mathbf{x}}_{\setminus i}}: & D_{\varphi_{i,\bar{\mathbf{x}}_{\setminus i}}} & \to & \mathbb{R} \\ & x_i & \to & f(\bar{x}_1, ..., \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, ..., \bar{x}_n) \end{array},$$

définie sur :

$$D_{\varphi_{i,\bar{\mathbf{x}}_{\setminus i}}} = \{x_i \in \mathbb{R} | (\bar{x}_1, ..., \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, ..., \bar{x}_n) \in D_f\}$$

$\varphi_{i,\bar{\mathbf{x}}_{\setminus i}}$ est la fonction qui, à la i-ème variable x_i, associe la valeur de la fonction f en considérant chacune des variables x_k, $k \neq i$, égale à $\bar{x}_k$ fixé.

Pour $n = 2$: f est une fonction de deux variables $f : (x, y) \to f(x, y)$. Soit $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ fixé. On construit alors les deux fonctions :

$$\begin{array}{rccl} \varphi_{1,\bar{y}}: & D_{\varphi_{1,\bar{y}}} & \to & \mathbb{R} \\ & x & \to & f(x, \bar{y}) \end{array} \quad \text{et} \quad \begin{array}{rccl} \varphi_{2,\bar{x}}: & D_{\varphi_{2,\bar{x}}} & \to & \mathbb{R} \\ & y & \to & f(\bar{x}, y) \end{array}.$$

Voir Fig. 14 pour un exemple de construction des fonctions $\varphi_{2,\bar{x}}$ et $\varphi_{1,\bar{y}}$.
Pour des raisons de légèreté de notation, nous ne précisons pas les domaines de définition des fonctions par la suite. Un bon exercice peut consister, au moins pour des fonctions de deux variables, à écrire ces domaines de définition.

Définition 24 : Dérivées partielles, vecteur gradient, différentielle.

- Pour $i \in \mathbb{N} \cap [1, n]$, la **dérivée partielle de f par rapport à x_i** est la fonction de n variables :
$$\frac{\partial f}{\partial x_i} : \bar{\mathbf{x}} = (\bar{x}_1, ..., \bar{x}_n) \to \varphi'_{i,\bar{\mathbf{x}}_{\setminus i}}(\bar{x}_i).$$
Elle est parfois notée $\partial_i f$.
- Si les dérivées partielles de f sont définies en $\bar{\mathbf{x}}$, on définit le **vecteur gradient** de f en $\bar{\mathbf{x}}$:
$$\vec{\nabla} f(\bar{\mathbf{x}}) = \left(\frac{\partial f}{\partial x_1}(\bar{\mathbf{x}}), ..., \frac{\partial f}{\partial x_n}(\bar{\mathbf{x}}) \right) \in \mathbb{R}^n.$$
parfois noté $\overrightarrow{grad} f(\bar{\mathbf{x}})$. ∇ se prononce "nabla".
- La **différentielle** de la fonction f au point $\bar{\mathbf{x}}$ est la fonction de n variables définie sur tout $\mathbb{R}^n$:
$$\begin{array}{rccl} Df_{\bar{\mathbf{x}}}: & \mathbb{R}^n & \to & \mathbb{R} \\ & \mathbf{h} = (h_1, ..., h_n) & \to & \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}(\bar{\mathbf{x}}).h_i \right) = Df_{\bar{\mathbf{x}}}(\mathbf{h}) \end{array}.$$

Pour $n = 2$: f est une fonction de deux variables $f : (x, y) \to f(x, y)$.

- Les dérivées partielles de f sont :

$$\frac{\partial f}{\partial x} : (x, y) \to \varphi'_{1,y}(x) \text{ et } \frac{\partial f}{\partial y} : (x, y) \to \varphi'_{2,x}(y),$$

- Le vecteur gradient de f en $(\bar{x}, \bar{y})$ est :

$$\vec{\nabla} f(\bar{x}, \bar{y}) = \left(\frac{\partial f}{\partial x}(\bar{x}, \bar{y}), \frac{\partial f}{\partial y}(\bar{x}, \bar{y}) \right) \in \mathbb{R}^2,$$

- La différentielle de f en $(\bar{x}, \bar{y})$ est :

$$Df_{(\bar{x},\bar{y})} : \begin{array}{rcl} \mathbb{R}^n & \to & \mathbb{R} \\ (h, k) & \to & \frac{\partial f}{\partial x}(\bar{x}, \bar{y}).h + \frac{\partial f}{\partial y}(\bar{x}, \bar{y}).k = Df_{(\bar{x},\bar{y})}(h, k) \end{array}.$$

Remarque 7 : Pour le calcul pratique de la dérivée partielle d'une fonction f par rapport à sa i-ème variable x_i, c'est-à-dire $\frac{\partial f}{\partial x_i}$, il suffit de dériver l'expression définissant f en la considérant comme une fonction de la seule variable x_i, tous les autres x_j étant considérés comme des constantes.

Exemple 12 : Considérons la fonction de 3 variables :

$$f : \begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R} \\ (x, y, z) & \to & x\mathrm{e}^{2y+3z} \end{array}.$$

Pour calculer $\frac{\partial f}{\partial x}$, on dérive la fonction :

$$\varphi_{1,(y,z)} : \begin{array}{rcl} \mathbb{R} & \to & \mathbb{R} \\ x & \to & x\mathrm{e}^{2y+3z} \end{array}$$

où y et z sont considérés comme des paramètres constants. On obtient :

$$\frac{\partial f}{\partial x} : \begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R} \\ (x, y, z) & \to & \mathrm{e}^{2y+3z} \end{array}$$

De même on obtient, à l'aide des dérivées des fonctions $\varphi_{2,(x,z)}$ et $\varphi_{3,(x,y)}$ (qu'on pourra expliciter à titre d'exercice) :

$$\frac{\partial f}{\partial y} : \begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R} \\ (x, y, z) & \to & 2x\mathrm{e}^{2y+3z} \end{array} \text{ et } \frac{\partial f}{\partial z} : \begin{array}{rcl} \mathbb{R}^3 & \to & \mathbb{R} \\ (x, y, z) & \to & 3x\mathrm{e}^{2y+3z} \end{array}$$

On obtient la première en "dérivant" (plus exactement, en "différenciant") $x \exp(2y + 3z)$ par rapport à y en considérant x et z constants, et la seconde en différenciant par rapport

à z en considérant x et y constants.
Le vecteur gradient en un point $(\bar{x}, \bar{y}, \bar{z})$ est donc :

$$\vec{\nabla} f(\bar{x}, \bar{y}, \bar{z}) = \left(\mathrm{e}^{2\bar{y}+3\bar{z}}; 2\bar{x}\mathrm{e}^{2\bar{y}+3\bar{z}}; 3\bar{x}\mathrm{e}^{2\bar{y}+3\bar{z}}\right)$$

La différentielle en ce même point est la fonction :

$$Df_{(\bar{x},\bar{y},\bar{z})} : \begin{array}{rcl} \mathbb{R}^n & \to & \mathbb{R} \\ (h_1, h_2, h_3) & \to & \mathrm{e}^{2\bar{y}+3\bar{z}} h_1 + 2\bar{x}\mathrm{e}^{2\bar{y}+3\bar{z}} h_2 + 3\bar{x}\mathrm{e}^{2\bar{y}+3\bar{z}} h_3 \end{array}.$$

La Fig. 14 montre la construction des dérivées partielles d'une fonction de deux variables $f : (x, y) \to f(x, y)$ en un point $(\bar{x}, \bar{y})$ et leur interprétation géométrique respectivement dans les plans $y = \bar{y}$ et $x = \bar{x}$.

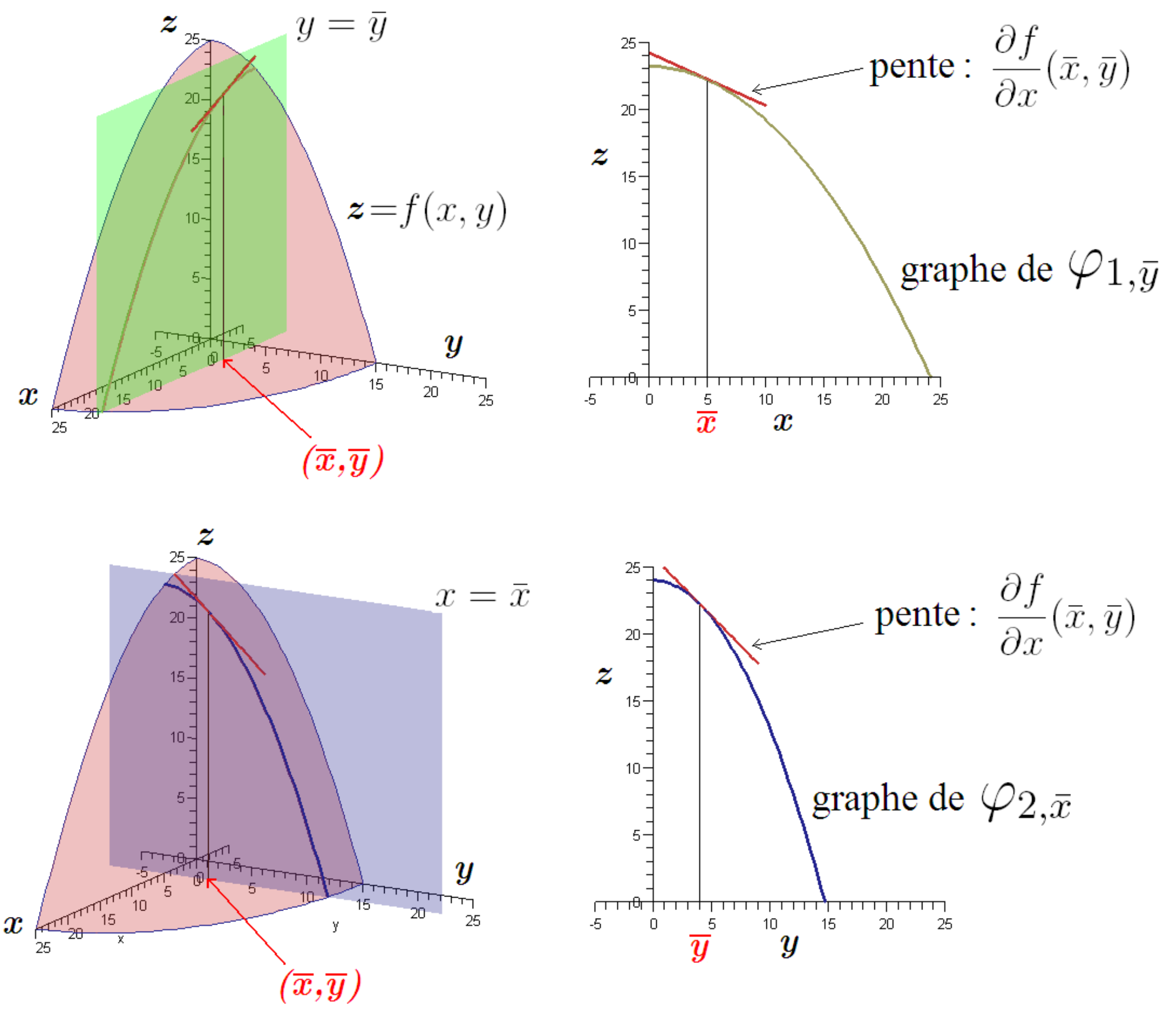

Figure 14 – Construction des dérivées partielles d'une fonction f de 2 variables en un point $(\bar{x}, \bar{y})$ comme les dérivées des fonctions $\varphi_{1,\bar{y}}$ en $\bar{x}$ et $\varphi_{2,\bar{x}}$ en $\bar{y}$.

Le vecteur gradient en un point est une généralisation, pour les fonctions de plusieurs variables, du nombre dérivé en un point pour les fonctions d'une variable. Ceci est montré immédiatement en posant $n = 1$ dans la définition : le vecteur gradient d'une fonction d'une variable en un point admet une unique composante qui est la dérivée de la fonction en ce point (dérivée partielle par rapport à l'unique variable). Il est intéressant de noter que le vecteur gradient en un point d'une fonction de n variables est un vecteur de $\mathbb{R}^n$ qui indique la direction de la plus grande pente (positive) du graphe de f en ce point. La norme du vecteur gradient est proportionnelle à la variation de f dans cette direction : plus celle-ci est grande, plus la norme du vecteur est grande.

On peut également interpréter la différentielle comme une généralisation aux fonctions de plusieurs variables de l'application linéaire tangente (obtenue grâce à la formule de Taylor-Young à l'ordre 1) associée à une fonction f d'une variable réelle : $h \to f'(\bar{x}).h$.

Définition 25 : Un point $\bar{\mathbf{x}} \in \mathbb{R}^n$ est dit **point singulier** ou **point stationnaire** de f si $\vec{\nabla} f(\bar{\mathbf{x}}) = \mathbf{0}_n$, i.e.

$$\forall i \in \mathbb{N} \cap [1, n], \frac{\partial f}{\partial x_i}(\bar{\mathbf{x}}) = 0.$$

On a alors directement :

$$\begin{array}{rccc} Df_{\bar{\mathbf{x}}}: & \mathbb{R}^n & \to & \mathbb{R} \\ & \mathbf{h} & \to & 0 \end{array}.$$

Proposition 10 : Dérivation en chaîne.
Soit g une fonction de m variables réelles à valeurs réelles :

$$g : (u_1, ..., u_m) \to g(u_1, ..., u_m).$$

Soient u_i, $i \in \mathbb{N} \cap [1, m]$, m fonctions de n variables à valeurs dans $\mathbb{R}$:

$$\begin{array}{rccc} u_i : & \mathbb{R}^n & \to & \mathbb{R} \\ & \mathbf{x} = (x_1, ..., x_n) & \to & u_i(\mathbf{x}) = u_i(x_1, ..., x_n) \end{array}.$$

Si la fonction f de n variables est donnée par :

$$f : \mathbf{x} = (x_1, ..., x_n) \to g(u_1(\mathbf{x}), ..., u_m(\mathbf{x})) = g(u_1(x_1, ..., x_n), ..., u_m(x_1, ..., x_n)).$$

Alors, sous réserve d'existence des dérivées partielles utilisées, pour tout $k \in \mathbb{N} \cap [1, n]$:

$$\frac{\partial f}{\partial x_k}(\mathbf{x}) = \sum_{i=1}^{m} \left(\frac{\partial g}{\partial u_i}(u_1(\mathbf{x}), ..., u_m(\mathbf{x})). \frac{\partial u_i}{\partial x_k}(\mathbf{x}) \right). \tag{2}$$

En abrégé :

$$\boxed{\frac{\partial f}{\partial x_k} = \sum_{i=1}^{m} \left(\frac{\partial g}{\partial u_i}. \frac{\partial u_i}{\partial x_k} \right).}$$

Remarque 8 : Pour $n = 1$, $f : x \to g(u_1(x), ..., u_m(x))$, (2) s'écrit :

$$f'(x) = \sum_{i=1}^{m} \left(u_i'(x) \frac{\partial g}{\partial u_i}(u_1(x), ..., u_m(x)) \right).$$

Pour $m = 1$, $f : \mathbf{x} = (x_1, ..., x_n) \to g(u(\mathbf{x})) = g(u(x_1, ..., x_n))$, (2) s'écrit :

$$\forall k \in \mathbb{N} \cap [1, n], \frac{\partial f}{\partial x_k}(\mathbf{x}) = g'(u(\mathbf{x})).\frac{\partial u}{\partial x_k}(\mathbf{x}).$$

Définition 26 : Hyperplan et vecteur normal.
Un hyperplan de $\mathbb{R}^n$ est un sous-espace affine de dimension $n-1$ de $\mathbb{R}^n$, i.e. la translation d'un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $n - 1$.
L'hyperplan $\mathcal{P}$ de $\mathbb{R}^m$ passant par $(\bar{x}_1, ..., \bar{x}_m)$ dont $\vec{u} = (a_1, ..., a_m)$ est un vecteur normal est l'ensemble des points $(x_1, ..., x_m)$ vérifiant l'équation :

$$\mathcal{P} : \sum_{i=1}^{m} a_i(x_i - \bar{x}_i) = 0.$$

Définition 27 : L'hyperplan de meilleure approximation du graphe de f au voisinage de $\bar{\mathbf{x}}$, s'il existe, est appellé l'**hyperplan tangent** au graphe de f en $(\bar{\mathbf{x}}, f(\bar{\mathbf{x}}))$.

Proposition 11 : Si les dérivées partielles de f sont définies en $\bar{\mathbf{x}} = (\bar{x}_1, ..., \bar{x}_n)$, le vecteur :

$$\left(\frac{\partial f}{\partial x_1}(\bar{\mathbf{x}}), ..., \frac{\partial f}{\partial x_n}(\bar{\mathbf{x}}), -1 \right) \in \mathbb{R}^{n+1}$$

est normal à l'hyperplan tangent $\mathcal{P}^f_{\bar{\mathbf{x}}}$ au graphe de f en $(\bar{\mathbf{x}}, f(\bar{\mathbf{x}}))$. Il admet pour équation :

$$\mathcal{P}^f_{\bar{\mathbf{x}}} : z - f(\bar{\mathbf{x}}) = \sum_{i=1}^{n} \left[\frac{\partial f}{\partial x_i}(\bar{\mathbf{x}}).(x_i - \bar{x}_i) \right], \quad (x_1, ..., x_n) \in \mathbb{R}^n.$$

Pour $n = 2$, $f : (x, y) \to f(x, y)$, sous réserve d'existence des dérivées partielles, le vecteur

$$\left(\frac{\partial f}{\partial x}(\bar{x}, \bar{y}), \frac{\partial f}{\partial y}(\bar{x}, \bar{y}), -1 \right) \in \mathbb{R}^3$$

est un vecteur normal au plan tangent au graphe de f en $(\bar{x}, \bar{y}, f(\bar{x}, \bar{y}))$. Ce plan admet donc pour équation cartésienne :

$$\mathcal{P}^f_{(\bar{x},\bar{y})} : z - f(\bar{x}, \bar{y}) = \frac{\partial f}{\partial x}(\bar{x}, \bar{y}).(x - \bar{x}) + \frac{\partial f}{\partial y}(\bar{x}, \bar{y}).(y - \bar{y}), \quad (x, y) \in \mathbb{R}^2.$$

Proposition 12 : Tangente à une ligne de niveau.
Soit f une fonction de deux variables dont les dérivées partielles sont définies au point $(\bar{x}, \bar{y}) \in \mathcal{L}^f_c$, où $\mathcal{L}^f_c$ désigne la ligne de niveau c de f. Supposons, $\vec{\nabla} f(\bar{x}, \bar{y}) \neq (0, 0)$, c'est-à-dire, $(\bar{x}, \bar{y})$ n'est pas un point singulier de f. Alors :

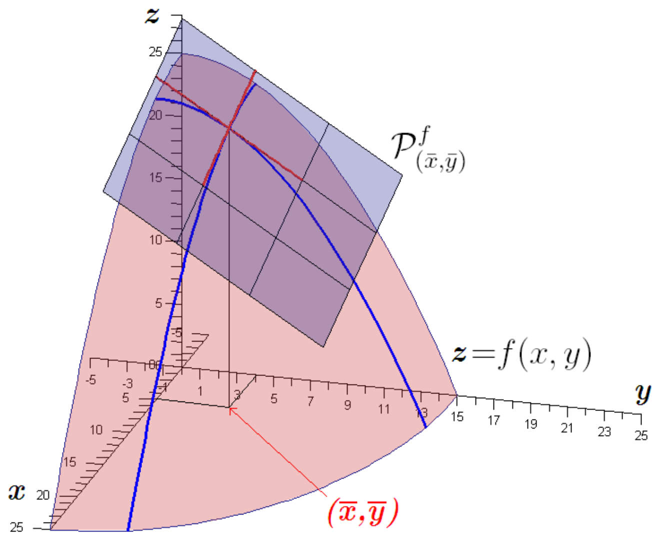

FIGURE 15 – Construction du plan $\mathcal{P}^f_{(\bar{x},\bar{y})}$ tangent au graphe de f au point $(\bar{x}, \bar{y}, f(\bar{x}, \bar{y}))$.

- la tangente à $\mathcal{L}^f_c$ en $(\bar{x}, \bar{y})$ est bien définie : il s'agit de la droite intersection entre le plan tangent au graphe de f en $(\bar{x}, \bar{y}, c)$ et le plan $z = c$;
- elle admet pour équation :

$$\frac{\partial f}{\partial x}(\bar{x}, \bar{y}).(x - \bar{x}) + \frac{\partial f}{\partial y}(\bar{x}, \bar{y}).(y - \bar{y}) = 0;$$

- le vecteur gradient est normal à cette tangente et il est dirigé vers les lignes de niveaux supérieurs (à c).

IV.2 Dérivées partielles d'ordre 2

Définition 28 : Sous réserve d'existence des dérivées partielles des dérivées partielles de f, on définit les n^2 **dérivées partielles d'ordre** 2 par les fonctions de n variables :

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right), \quad (i, j) \in (\mathbb{N} \cap [1, n]) \times (\mathbb{N} \cap [1, n]),$$

parfois noté $\partial_j \partial_i f$. On note :

$$\frac{\partial^2 f}{\partial x_i^2} = \frac{\partial^2 f}{\partial x_i \partial x_i} \text{parfois noté } \partial_i^2 f.$$

On définit alors la **matrice Hessienne** de f en $\bar{\mathbf{x}} \in \mathbb{R}^n$ où toutes ses dérivées partielles d'ordre 2 sont bien définies, comme la matrice $n \times n$ réelle suivante :

$$H^f(\bar{\mathbf{x}}) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{\mathbf{x}})\right)_{1 \leq i,j \leq n}.$$

Pour $n = 2$, $f : (x, y) \to f(x, y)$, sous réserve d'existence des dérivées partielles des dérivées partielles, les 4 dérivées partielles d'ordre 2 sont les fonctions de deux variables :

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right).$$

La matrice Hessienne de f en un point $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ où ces 4 dérivées partielles d'ordre 2 sont bien définies est la matrice 2×2 réelle :

$$H^f(\bar{x}, \bar{y}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(\bar{x}, \bar{y}) & \frac{\partial^2 f}{\partial y \partial x}(\bar{x}, \bar{y}) \\ \frac{\partial^2 f}{\partial x \partial y}(\bar{x}, \bar{y}) & \frac{\partial^2 f}{\partial y^2}(\bar{x}, \bar{y}) \end{pmatrix}.$$

IV.3 Approximations de Taylor

Proposition 13 : Approximation de Taylor à l'ordre 1.
Si toutes les dérivées partielles de f sont définies au point $\bar{\mathbf{x}} = (\bar{x}_1, ..., \bar{x}_n)$, alors pour $\mathbf{h} = (h_1, ..., h_n) \in \mathbb{R}^n$ où les h_i sont considérés "petits", on appelle **approximation de Taylor à l'ordre 1**, l'approximation suivante :

$$\boxed{f(\bar{\mathbf{x}} + \mathbf{h}) \approx f(\bar{\mathbf{x}}) + Df_{\bar{\mathbf{x}}}(\mathbf{h}),}$$

i.e. :

$$\boxed{f(\bar{x}_1 + h_1, ..., \bar{x}_n + h_n) \approx f(\bar{x}_1, ..., \bar{x}_n) + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}(\bar{x}_1, ..., \bar{x}_n).h_i\right).}$$

Remarque 9 : Il ne s'agit, dans cet énoncé, que d'une approximation. Dans ce cours, on ne définit pas de reste et cette proposition ne donne pas de maîtrise de l'erreur. Il en est de même pour la proposition suivante.

Proposition 14 : Approximation de Taylor à l'ordre 2.
Si toutes les dérivées partielles d'ordre 2 de f sont définies au point $\bar{\mathbf{x}} = (\bar{x}_1, ..., \bar{x}_n)$, alors pour $\mathbf{h} = (h_1, ..., h_n) \in \mathbb{R}^n$ où les h_i sont considérés "petits", on appelle **approximation de Taylor à l'ordre 2**, l'approximation suivante :

$$\boxed{f(\bar{\mathbf{x}} + \mathbf{h}) \approx f(\bar{\mathbf{x}}) + Df_{\bar{\mathbf{x}}}(\mathbf{h}) + \frac{1}{2} \sum_{1 \le i,j \le n} \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{x}_1, ..., \bar{x}_n).h_i.h_j\right).}$$

Pour $n = 2$, les approximations de Taylor à l'ordre 1 et 2 s'écrivent respectivement, pour h et k "petits" :

$$f(\bar{x} + h, \bar{y} + k) \approx f(\bar{x}, \bar{y}) + Df_{(\bar{x}, \bar{y})}(h, k) = f(\bar{x}, \bar{y}) + \frac{\partial f}{\partial x}(\bar{x}, \bar{y}).h + \frac{\partial f}{\partial y}(\bar{x}, \bar{y}).k,$$

$$\begin{aligned} f(\bar{x} + h, \bar{y} + k) \approx {} & f(\bar{x}, \bar{y}) + \frac{\partial f}{\partial x}(\bar{x}, \bar{y}).h + \frac{\partial f}{\partial y}(\bar{x}, \bar{y}).k \\ & + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(\bar{x}, \bar{y}).h^2 + \frac{\partial^2 f}{\partial x \partial y}(\bar{x}, \bar{y}).hk + \frac{\partial^2 f}{\partial y \partial x}(\bar{x}, \bar{y}).hk + \frac{\partial^2 f}{\partial y^2}(\bar{x}, \bar{y}).k^2\right). \end{aligned}$$

IV.4 Continuité des fonctions de n variables et classe C^k

Note : Les Définitions 29 à 32 et les Propositions 15 et 16 sont données à titre informatif. Dans le cadre des TD's et des examens de cette année, nous considérerons essentiellement des fonctions vérifiant les propriétés de la Proposition 17. Il suffira de vérifier que les fonctions sont différentiables au moins 3 fois pour pouvoir appliquer le Théorème de Schwarz puis les résultats des chapitres suivants. La Proposition 17 et le Théorème de Schwarz sont donc des résultats cruciaux à connaître par coeur.

Définition 29 : La **boule ouverte** de $\mathbb{R}^n$ de centre $\bar{\mathbf{x}} = (\bar{x}_1, ..., \bar{x}_n)$ et de rayon r est définie par :

$$B(\bar{\mathbf{x}}, r) = \left\{ \mathbf{x} = (x_1, ..., x_n) \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^{n} |x_i - \bar{x}_i|^2 < r^2 \right\}.$$

Un **voisinage** $V_{\bar{\mathbf{x}}}$ de $\bar{\mathbf{x}}$ est un ensemble contenant une boule ouverte non vide de centre $\bar{\mathbf{x}}$, i.e. il existe $r > 0$ tel que $B(\bar{\mathbf{x}}, r) \subset V_{\bar{\mathbf{x}}}$.

Exemple 13 : La boule ouverte de $\mathbb{R}$ de centre $\bar{x}$ et de rayon α est l'intervalle ouvert $]\bar{x} - \alpha, \bar{x} + \alpha[$. La boule ouverte de $\mathbb{R}^2$ de centre $(\bar{x}, \bar{y})$ et de rayon r est le disque ouvert (sans le cercle le bordant) de centre $(\bar{x}, \bar{y})$ et de rayon r.

Définition 30 : Un point $\bar{\mathbf{x}}$ est dit **point intérieur** d'une partie U de $\mathbb{R}^n$ s'il existe un voisinage de $\bar{\mathbf{x}}$ (donc une boule ouverte non vide de centre $\bar{\mathbf{x}}$) inclus(e) dans U.
La partie U de $\mathbb{R}^n$ est un **ouvert** de $\mathbb{R}^n$ si tout point de U est intérieur à U (i.e. les points de la frontière de U n'appartiennent pas à U).

Dans la fin de cette sous-section, on considère une fonction f de n variables réelles à valeurs réelles définie sur un ouvert $U \subset \mathbb{R}^n$:

$$f : \begin{array}{rcl} U & \to & F \subset \mathbb{R} \\ \mathbf{x} = (x_1, ..., x_n) & \to & f(\mathbf{x}) = f(x_1, ..., x_n) \end{array}.$$

Définition 31 : Continuité d'une fonction de n variables.
f est dite continue en $\bar{\mathbf{x}} \in U$ si :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall \mathbf{x} \in U \cap B(\bar{\mathbf{x}}, \eta), |f(\mathbf{x}) - f(\bar{\mathbf{x}})| < \varepsilon.$$

f est dite continue sur $V \subset U$ si elle est continue en tout point de V.

Définition 32 : Fonction de classe C^k.
La fonction f est **de classe C^k** sur U, $k \in \mathbb{N}$, si pour tout vecteur $\mathbf{u} = (u_1, ..., u_n) \in \mathbb{R}^n$, pour tout $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{R}^n$, la fonction $t \to f(\mathbf{x} + t\mathbf{u}) = f(x_1 + tu_1, ..., x_n + tu_n)$ est de classe C^k en 0 (au sens des fonctions d'une variable, cf. II.3 Def. 8).
Si f est de classe C^k alors elle est de classe C^p pour tout $p \leq k$.

Proposition 15 : Condition nécessaire et suffisante pour la classe C^1.
f est de classe C^1 sur U si et seulement si toutes les dérivées partielles d'ordre 1 sont définies et continues sur U.

Proposition 16 : Condition nécessaire pour la classe C^2.
Si f est de classe C^2 sur U alors toutes les dérivées partielles d'ordre 2 sont définies et continues sur U.

Proposition 17 : Condition suffisante pour la classe C^2.
Si toutes les dérivées partielles d'ordre 2 de f sont définies et continues sur U et admettent des dérivées partielles par rapport à chacune des variables x_i définies sur U, alors f est de classe C^2 sur U.

Théorème 5 : Théorème de Schwarz
Si f est de classe C^2 sur U, pour tout $i \in \mathbb{N} \cap [1, n]$, pour tout $j \in \mathbb{N} \cap [1, n]$:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} \text{ sur } U.$$

Ainsi, pour tout $\mathbf{x} \in U$, la matrice Hessienne de f en $\mathbf{x}$ est une matrice symétrique réelle.

V Formes quadratiques

Définition 33 : Forme quadratique et matrice associée.
Une **forme quadratique** à n variables est une fonction de la forme

$$q : \begin{array}{rcl} \mathbb{R}^n & \to & \mathbb{R} \\ \mathbf{x} = (x_1, ..., x_n) & \to & q(\mathbf{x}) = \sum\limits_{1 \leq i,j \leq n} a_{i,j} x_i x_j \end{array}.$$

avec $a_{i,j} \in \mathbb{R}$ et, pour tout (i, j), $a_{i,j} = a_{j,i}$.
La **matrice associée** à q est la matrice symétrique $M_q = (a_{i,j})_{1 \leq i,j \leq n}$.

Remarque 10 : 1) Pour toute forme quadratique q à n variables, $q(\mathbf{0}_n) = 0$.
2) Il existe une unique forme quadratique à n variables, appelée **forme quadratique nulle**, vérifiant :

$$\forall \mathbf{x} \in \mathbb{R}^n, q(\mathbf{x}) = 0.$$

Les coefficients $a_{i,j}$ de cette forme sont tous nuls. Par conséquent, la matrice associée est la matrice nulle $n \times n$.

Définition 34 : Types des formes quadratiques et des matrices symétriques.
La forme quadratique q (ou la matrice associée) est dite :

- **définie positive** ("DP" ou "de type I") si :

$$\forall \mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}_n\}, q(\mathbf{x}) > 0,$$

- **définie négative** ("DN" ou "de type II") si :

$$\forall \mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}_n\}, q(\mathbf{x}) < 0,$$

– **semi-définie positive** ("SDP" ou "de type III") si elle n'est pas DP et vérifie

$$\forall \mathbf{x} \in \mathbb{R}^n, q(\mathbf{x}) \geq 0,$$

– **semi-définie négative** ("SDN" ou "de type IV") si elle n'est pas DN et vérifie

$$\forall \mathbf{x} \in \mathbb{R}^n, q(\mathbf{x}) \leq 0,$$

– **indéfinie** ("IND" ou "de type V") si elle n'est d'aucun des types précédents, i.e. :

$$\text{il existe } \mathbf{u} \in \mathbb{R}^n \text{ et } \mathbf{v} \in \mathbb{R}^n \text{ tels que } q(\mathbf{u}) > 0 \text{ et } q(\mathbf{v}) < 0.$$

Toute forme quadratique non identiquement nulle admet un unique type.
Seule la forme quadratique nulle peut être considérée comme étant à la fois SDP et SDN.

Proposition 18 : Classification des formes quadratiques à deux variables.
Soit la forme quadratique $q(x, y) = ax^2 + by^2 + 2cxy$ et sa matrice associée :

$$M_q = \begin{pmatrix} a & c \\ c & b \end{pmatrix}.$$

On note $\Delta = \det M_q = ab - c^2$ le **discriminant** de q (ou de M_q). Alors, le type de q est donné par le tableau suivant :

Type de q	$a > 0$ et $b > 0$	$a < 0$ et $b < 0$	$ab = 0$ et $a + b > 0$	$ab = 0$ et $a + b < 0$	$ab < 0$
$\Delta > 0$	DP type I	DN type II	impossible	impossible	impossible
$\Delta = 0$	SDP type III	SDN type IV	SDP type III	SDN type IV	impossible
$\Delta < 0$	IND type V				

Remarques :

1. La condition "$a > 0$ et $b > 0$" est équivalente à "$ab > 0$ et $a + b > 0$".
De même, "$a < 0$ et $b < 0$" est équivalente à "$ab > 0$ et $a + b < 0$".
2. Δ discrimine les cas "défini" ($\Delta > 0$), "semi-défini" ($\Delta = 0$) et "indéfini" ($\Delta < 0$).
Pour $\Delta \geq 0$, le signe de $a + b$ donne la "positivité" ou la "négativité".
3. Si $a = 0$ ou $b = 0$ (i.e. $ab = 0$),
$\Delta = -c^2 \leq 0$: q ne peut pas être "définie" (i.e. de type I ou II).
   – Si $\Delta = 0$, elle est semi-définie (positive ou négative selon le signe de $a + b$, i.e. du coeficient non nul).
   – Si $\Delta < 0$, elle est indéfinie.
4. Si $ab < 0$, $\Delta = ab - c^2$ est toujours strictement négatif et q est donc toujours indéfinie.

La Fig. 16 montre les graphes de formes quadratiques de deux variables des 5 types différents.

La classification de la Proposition 18 est à connaître par cœur.
La classification de la Proposition 19 sera rappelée, si besoin est, dans les énoncés d'examens.

Proposition 19 : Classification des formes quadratiques à trois variables.
Soit la forme quadratique à trois variables et la matrice symétrique associée :

$$q(x,y,z) = ax^2 + by^2 + cz^2 + 2d_1xy + 2d_2xz + 2d_3yz, \quad M = \begin{pmatrix} a & d_1 & d_2 \\ d_1 & b & d_3 \\ d_2 & d_3 & c \end{pmatrix}.$$

– M est **définie positive** ("DP" ou "de type I") si et seulement si

$$a > 0, \; \begin{vmatrix} a & d_1 \\ d_1 & b \end{vmatrix} > 0, \; \det M > 0.$$

– M est **définie négative** ("DN" ou "de type II") si et seulement si

$$a < 0, \; \begin{vmatrix} a & d_1 \\ d_1 & b \end{vmatrix} > 0, \; \det M < 0.$$

– M est **semi-définie positive** ("SDP" ou "de type III") si et seulement si

$$a, b, c \geq 0; \; \begin{vmatrix} a & d_1 \\ d_1 & b \end{vmatrix}, \; \begin{vmatrix} a & d_2 \\ d_2 & c \end{vmatrix} \text{ et } \begin{vmatrix} b & d_3 \\ d_3 & c \end{vmatrix} \geq 0; \; \det M = 0.$$

– M est **semi-définie négative** ("SDN" ou "de type IV") si et seulement si

$$a, b, c \leq 0; \; \begin{vmatrix} a & d_1 \\ d_1 & b \end{vmatrix}, \; \begin{vmatrix} a & d_2 \\ d_2 & c \end{vmatrix} \text{ et } \begin{vmatrix} b & d_3 \\ d_3 & c \end{vmatrix} \geq 0; \; \det M = 0.$$

– M est **indéfinie** ("IND" ou "de type V") si et seulement si l'un des neuf réels

$$ab, \; ac, \; bc, \; a \det M, \; b \det M, \; c \det M, \; \begin{vmatrix} a & d_1 \\ d_1 & b \end{vmatrix}, \; \begin{vmatrix} a & d_2 \\ d_2 & c \end{vmatrix}, \; \begin{vmatrix} b & d_3 \\ d_3 & c \end{vmatrix}$$

est strictement négatif.

Définition 35 : Mineurs diagonaux principaux.
Pour une matrice $M = (a_{i,j})_{1 \leq i,j \leq n}$, on appelle **mineurs diagonaux principaux** les n déterminants :

$$\Delta_k = \det\left[(a_{i,j})_{1 \leq i,j \leq k}\right], \quad k \in \mathbb{N} \cap [1, n] = \{1, 2, ..., n\}.$$

Δ_k est appelé mineur diagonal principal d'ordre k : il s'agit du déterminant obtenu à partir de la sous-matrice de M constituée des k premières lignes et k premières colonnes.

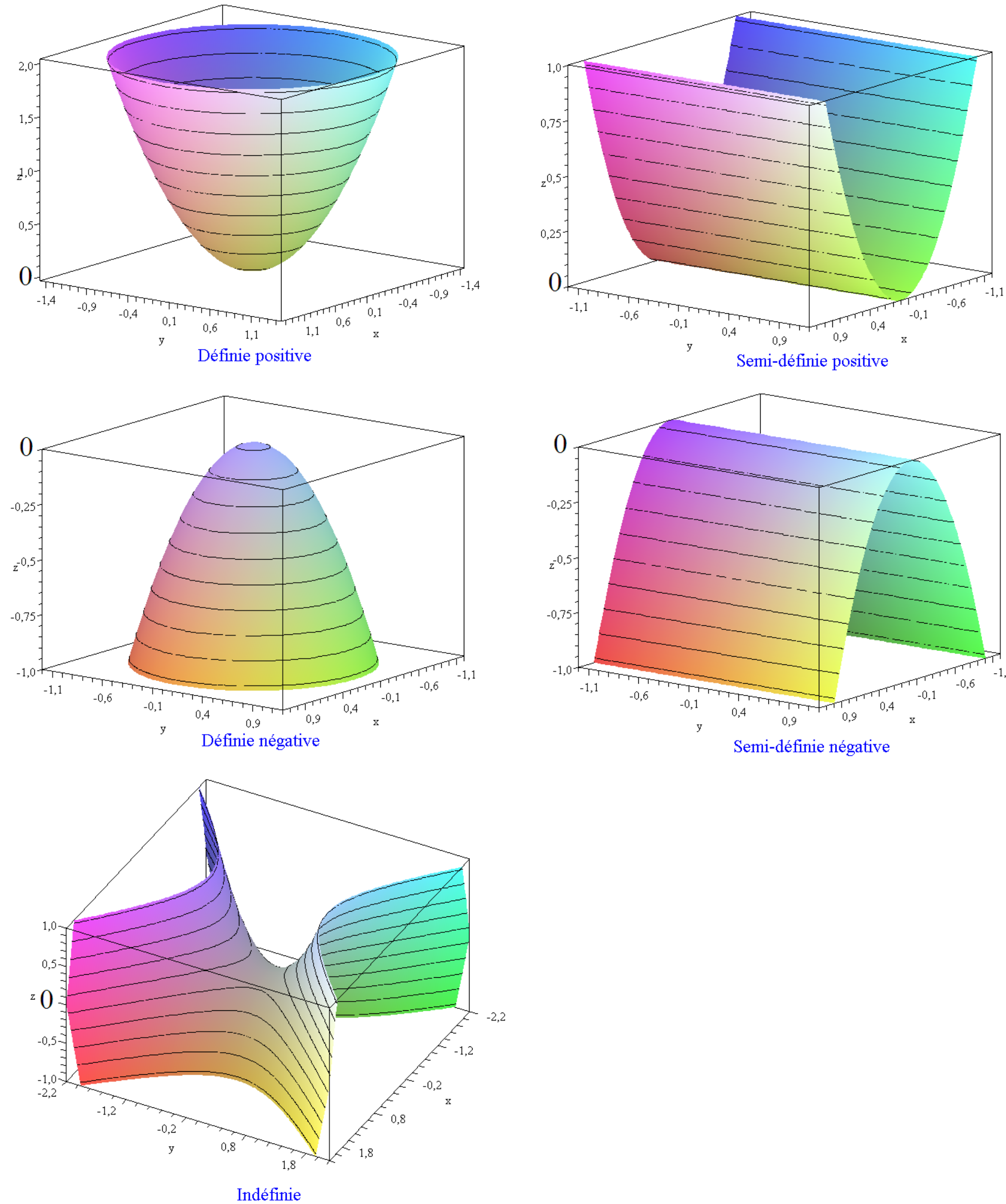

FIGURE 16 – Graphes de formes quadratiques de 2 variables illustrant les 5 types différents.

Théorème 6 : Critère de Sylvester.
Soit M une matrice symétrique et $q : \mathbb{R}^n \to \mathbb{R}$ la forme quadratique associée.

- q est définie positive ("DP" ou "de type I") si et seulement si tous ses mineurs diagonaux principaux sont strictement positifs, i.e. :
$$\forall k \in \mathbb{N} \cap [1, n], \Delta_k > 0.$$
- q est définie négative ("DN ou de "type II") si et seulement si tous ses mineurs diagonaux principaux sont non nuls et ont des signes alternés, le premier étant négatif, i.e. :
$$\forall k \in \mathbb{N} \cap [1, n], (-1)^k \Delta_k > 0 \qquad (i.e.\ \Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, ...).$$
- Si, pour tout $k \in \mathbb{N} \cap [1, n]$, $\Delta_k \neq 0$ et la suite $(\Delta_i)_i$ ne respecte aucune des deux structures de signes précédentes (i.e. soit il existe k pair tel que $\Delta_k < 0$, soit il existe k et l impairs tels que $\Delta_k \Delta_l < 0$), alors q est indéfinie ("IND" ou "de type V").

VI Optimisation

VI.1 Convexité

Définition 36 : Une partie U de $\mathbb{R}^n$ est dite **<u>convexe</u>** si

$$\forall \mathbf{x} = (x_1, ..., x_n) \in U, \forall \mathbf{y} = (y_1, ..., y_n) \in U, \forall \lambda \in [0, 1],$$
$$\lambda \mathbf{x} + (1 - \lambda)\mathbf{y} = (\lambda x_1 + (1 - \lambda) y_1, ..., \lambda x_n + (1 - \lambda) y_n) \in U$$

En d'autres termes, tout segment reliant deux points de U est entièrement inclus dans U (voir Fig. 17 a) et b)). Par définition, les parties convexes de $\mathbb{R}$ sont les intervalles de $\mathbb{R}$.

Définition 37 : Une partie U de $\mathbb{R}^n$ est dite **<u>strictement convexe</u>** si

$$\forall \mathbf{x} = (x_1, ..., x_n) \in U, \forall \mathbf{y} = (y_1, ..., y_n) \in U, \forall \lambda \in]0, 1[,$$
$$\mathbf{x} \neq \mathbf{y} \Longrightarrow \lambda \mathbf{x} + (1 - \lambda)\mathbf{y} = (\lambda x_1 + (1 - \lambda) y_1, ..., \lambda x_n + (1 - \lambda) y_n) \in \mathring{U}$$

où $\mathring{U}$ est l'intérieur de U, i.e. l'ensemble des points intérieurs à U (voir Fig. 17 a)). .

Définition 38 : Fonctions convexes, fonctions concaves.
Soient U une partie convexe de $\mathbb{R}^n$ et f une fonction de n variables réelles à valeurs réelles définie sur U. La fonction f est dite **<u>convexe</u>** sur U si

$$\forall \mathbf{x} = (x_1, ..., x_n) \in U, \forall \mathbf{y} = (y_1, ..., y_n) \in U, \forall \lambda \in [0, 1],$$
$$f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y}),$$
$$\text{i.e. } f(\lambda x_1 + (1 - \lambda) y_1, ..., \lambda x_n + (1 - \lambda) y_n) \leq \lambda f(x_1, ..., x_n) + (1 - \lambda) f(y_1, ..., y_n).$$

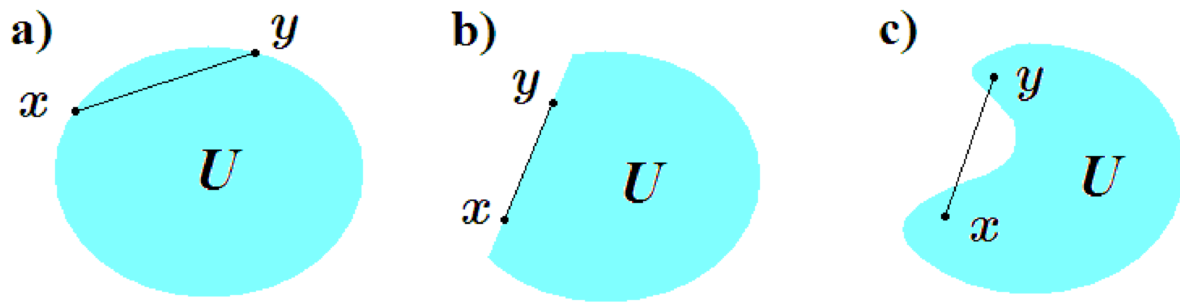

FIGURE 17 – Exemples de parties U présentant différentes propriétés de convexité. a) U est strictement convexe. b) U est convexe sans être strictement convexe. c) U est non convexe.

Autrement dit, le segment de $\mathbb{R}^{n+1}$ reliant deux points du graphe de f est au-dessus du graphe de f. Typiquement, une forme quadratique DP ou SDP est convexe.
La fonction f est dite **strictement convexe** sur U si

$$\forall \mathbf{x} = (x_1, ..., x_n) \in U, \forall \mathbf{y} = (y_1, ..., y_n) \in U, \forall \lambda \in]0, 1[,$$
$$\mathbf{x} \neq \mathbf{y} \Longrightarrow f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) < \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})$$

Autrement dit, le segment ouvert reliant deux points du graphe de f est strictement au-dessus du graphe de f. Typiquement, une forme quadratique DP est strictement convexe.
La fonction f est dite **concave** sur U (resp. **strictement concave**) si $-f$ est convexe (resp. strictement convexe) sur U.

Remarque 11 : Si une fonction définie sur un intervalle I est convexe (ou concave), alors elle est continue sur I.

Rappels : Les deux propositions suivantes concernent des fonctions d'une variable réelle et sont des résultats déjà mentionné en Méthodes Mathématiques 1.

Proposition 20 : Convexité pour les fonctions d'une variable.
Soit f une fonction définie et continue sur un intervalle I de $\mathbb{R}$.

- Si f est dérivable à droite et à gauche en tout point de I, alors f est convexe ssi

$$\forall (x, y) \in I^2, x < y \Longrightarrow f'_g(x) \leq f'_d(x) \leq f'_g(y) \leq f'_d(y).$$

- Si f est dérivable sur I, alors f est convexe ssi f' est croissante sur I.
- Si f est deux fois dérivable sur I, alors f est convexe ssi $f'' \geq 0$ sur I.

Proposition 21 : Stricte convexité pour les fonctions d'une variable.
Soit f une fonction définie sur un intervalle I de $\mathbb{R}$.

- Si f est dérivable à droite et à gauche en tout point et

$$\forall (x, y) \in I^2, x < y \Longrightarrow f'_g(x) \leq f'_d(x) < f'_g(y) \leq f'_d(y).$$

alors f est strictement convexe sur I.

– Si f est dérivable sur I et f' est strictement croissante sur I alors f est strictement convexe sur I.
– Si f est deux fois dérivable sur I et $f'' > 0$ sur I alors f est convexe sur I.

Remarque 12 : Pour la stricte convexité, il ne s'agit pas d'équivalence. En effet, f peut être strictement convexe même si f'' s'annule en un point isolé.

Proposition 22 : Caractérisation de la convexité et de la concavité pour les fonctions à n variables.
Soient U une partie convexe de $\mathbb{R}^n$ et f une fonction de n variables réelles à valeurs réelles de classe C^2 sur U. On note $H^f(\mathbf{x})$ la matrice Hessienne de f évaluée au point $\mathbf{x} \in U$. D'après le théorème de Schwarz, $H^f(\mathbf{x})$ est symétrique. Alors :
– f est convexe sur U si et seulement si, en tout point $\mathbf{x} \in U$, $H^f(\mathbf{x})$ est positive (i.e. DP ou SDP),
– f est concave sur U si et seulement si, en tout point $\mathbf{x} \in U$, $H^f(\mathbf{x})$ est négative (i.e. DN ou SDN).

Proposition 23 : Stricte convexité pour les fonctions à n variables.
Sous les mêmes hypothèses que la proposition précédente :
– si, en tout $\mathbf{x} \in U$, $H^f(\mathbf{x})$ est définie positive, f est strictement convexe sur U,
– si, en tout $\mathbf{x} \in U$, $H^f(\mathbf{x})$ est définie négative, f est strictement concave sur U.

VI.2 Optimisation sans contrainte

Position du problème : déterminer le maximum ou le minimum d'une fonction
$f : D_f \subset \mathbb{R}^n \to \mathbb{R}$.

Définition 39 : La fonction f atteint en un point $\mathbf{x}^* \in D_f$ un
– **maximum global** si $f(\mathbf{x}^*) \geq f(\mathbf{x})$ pour tout $\mathbf{x} \in D_f$,
– **minimum global** si $f(\mathbf{x}^*) \leq f(\mathbf{x})$ pour tout $\mathbf{x} \in D_f$,
– **maximum global strict** si $f(\mathbf{x}^*) > f(\mathbf{x})$ pour tout $\mathbf{x} \in D_f \backslash \{\mathbf{x}^*\}$,
– **minimum global strict** si $f(\mathbf{x}^*) < f(\mathbf{x})$ pour tout $x \in D_f \backslash \{\mathbf{x}^*\}$,
– **maximum local** s'il existe un voisinage $V_{\mathbf{x}^*}$ de $\mathbf{x}^*$ tel que $f(\mathbf{x}^*) \geq f(\mathbf{x})$ pour tout $\mathbf{x} \in V_{\mathbf{x}^*} \cap D_f$,
– **minimum local** s'il existe un voisinage $V_{\mathbf{x}^*}$ de $\mathbf{x}^*$ tel que $f(\mathbf{x}^*) \leq f(\mathbf{x})$ pour tout $\mathbf{x} \in V_{\mathbf{x}^*} \cap D_f$,
– **maximum local strict** s'il existe un voisinage $V_{\mathbf{x}^*}$ de $\mathbf{x}^*$ tel que $f(\mathbf{x}^*) > f(\mathbf{x})$ pour tout $\mathbf{x} \in (V_{\mathbf{x}^*} \cap D_f) \backslash \{\mathbf{x}^*\}$,
– **minimum local strict** s'il existe un voisinage $V_{\mathbf{x}^*}$ de $\mathbf{x}^*$ tel que $f(\mathbf{x}^*) < f(\mathbf{x})$ pour tout $\mathbf{x} \in (V_{\mathbf{x}^*} \cap D_f) \backslash \{\mathbf{x}^*\}$.

Remarque 13 : Un extremum (i.e. minimum ou maximum) global est un extremum local. Il suffit de prendre $V_{\mathbf{x}^*} = \mathbb{R}^n$.

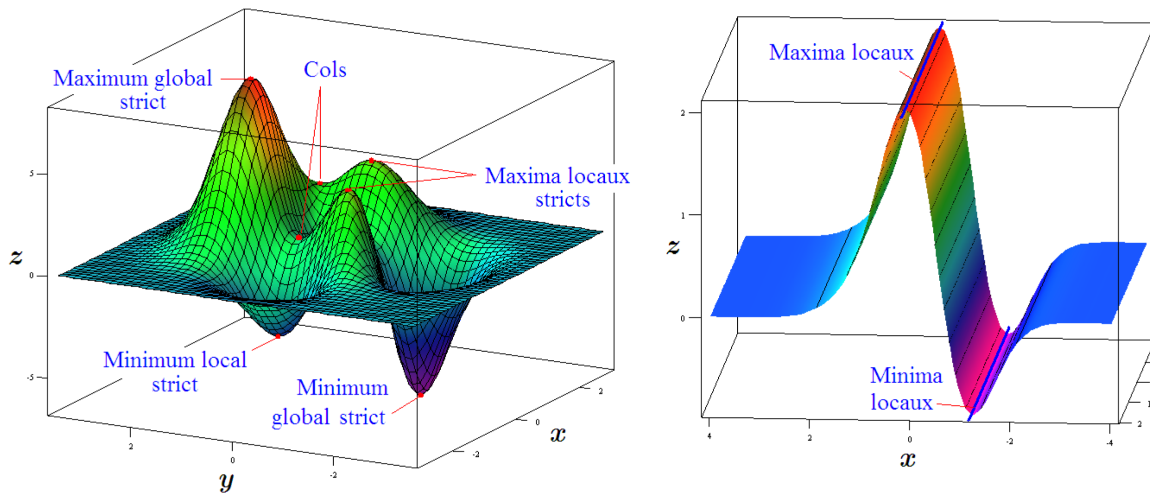

FIGURE 18 – Graphes de deux fonctions et nature de leurs extrema.

On remarque que si une fonction f de classe C^1 admet un extremum en un point $\bar{\mathbf{x}}$ intérieur à D_f, alors l'hyperplan tangent au graphe de f est donné par l'équation $z = f(\bar{\mathbf{x}}$ (pour une fonction de deux variables, ce plan tangent est horizontal). En un tel point, on a $\vec{\nabla} f(\bar{\mathbf{x}}) = \mathbf{0}_n$. Plus précisément :

Théorème 7 : Condition nécessaire du premier ordre.
Si f est de classe C^1 sur $\mathring{D}_f$ (ensemble des points intérieurs de D_f) et admet un extremum local en $\mathbf{x}^* \in D_f$ alors :
- soit $\mathbf{x}^*$ n'est pas un point intérieur de D_f,
- soit $\mathbf{x}^*$ est un point intérieur de D_f singulier pour f, i.e. $Df_{\mathbf{x}^*}$ est la forme linéaire identiquement nulle ou encore :

$$\vec{\nabla} f(\bar{\mathbf{x}}) = \mathbf{0}_n \text{ ou encore } \forall i \in \mathbb{N} \cap [1, n], \frac{\partial f}{\partial x_i}(\mathbf{x}^*) = 0$$

Remarque 14 : Il faut remarquer qu'une fonction admettant un minimum (resp. maximum) local strict est localement strictement convexe (resp. concave) près de ce point.
En se référant à la Proposition 23, ceci explique les similitudes entre :
- la forme du graphe près d'un minimum strict et le cas “Définie positive” de Fig. 16,
- la forme du graphe près d'un maximum strict et le cas “Définie négative” de Fig. 16.

Or, grâce à l'approximation de Taylor à l'ordre 2, il existe une forme quadratique associée à la fonction f qui décrit le comportement local au voisinage d'un point singulier : il s'agit de la forme quadratique associée à la matrice Hessienne de f en ce point. Les propriétés sous-jacentes à cette remarque sont précisées par le théorème suivant.

Théorème 8 : Condition suffisante du second ordre pour un extremum local. On suppose f de classe C^2 sur $\mathring{D}_f$ et $\mathbf{x}^* \in \mathring{D}_f$ un point singulier de f (i.e. la condition nécessaire du premier ordre est satisfaite). On considère la matrice Hessienne de f au point $\mathbf{x}^*$:

$$H^f(\mathbf{x}^*) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}^*) \right)_{1 \leq i,j \leq n}$$

$H^f(\mathbf{x}^*)$ est symétrique d'après le théorème de Schwarz et donc associée à une forme quadratique. Alors :
- Si $H^f(\mathbf{x}^*)$ est définie positive, alors f admet un minimum local strict en $\mathbf{x}^*$,
- Si $H^f(\mathbf{x}^*)$ est définie négative, alors f admet un maximum local strict en $\mathbf{x}^*$,
- Si $H^f(\mathbf{x}^*)$ est indéfinie, alors f n'admet pas d'extremum en $\mathbf{x}^*$: $\mathbf{x}^*$ est appelé **point col** ou **point selle**.

Il est important de noter que la propriété $H^f(\mathbf{x}^*)$ est SDP ou SDN ne permet pas de conclure sur la nature du point singulier.

Méthode de résolution du problème sans contrainte :

$$\min_{D_f} f \text{ ou } \max_{D_f} f.$$

1. On détermine les points vérifiant la condition nécessaire du premier ordre : E_1 l'ensemble des points $\mathbf{x}$ intérieurs à D_f tels que $\vec{\nabla} f(\bar{\mathbf{x}}) = \mathbf{0}_n$ (singuliers) et E_2 l'ensemble des points non intérieurs à D_f.
2. On calcule la matrice Hessienne de f en chaque point de E_1 : pour les points où elle est DP, DN ou IND, on peut conclure, pour les autres points de E_1 et les points de E_2, on effectue une étude locale.
3. On compare les valeurs des extrema locaux pour en déduire les extrema globaux.

VI.3 Optimisation sous contrainte

Position du problème général : Soit U un ouvert de $\mathbb{R}^n$. Soient $g_1, ..., g_p$ et $h_1, ..., h_q$, $p + q$ fonctions de classe C^2 définies sur U à valeurs réelles. Soit f une fonction de classe C^2 définie sur U à valeurs réelles. Trouver le minimum ou le maximum de f sur U sous les contraintes :

$$\forall i \in \mathbb{N} \cap [1, p], g_i(\mathbf{x}) \geq 0, \tag{3}$$
$$\forall j \in \mathbb{N} \cap [1, q], h_j(\mathbf{x}) = 0. \tag{4}$$

- La fonction f est appelée la **fonction objectif**,
- les g_i, $i \in \mathbb{N} \cap [1, p]$, les **contraintes en inéquations**,
- les h_j, $j \in \mathbb{N} \cap [1, q]$, les **contraintes en équations**
- l'ensemble des $\mathbf{x} \in \mathbb{R}^n$ vérifiant (3) et (4) est appelé le **domaine des réalisables**.

Dans cette section, nous nous limiterons à une seule contrainte en équation, i.e. $p = 0$ et $q = 1$. Ainsi, on considère h de classe C^2 sur U à valeurs réelles et f de classe C^2 sur U à valeurs réelles. On pose le problème :

$$\min_{\{\mathbf{x}\in U | h(\mathbf{x})=0\}} f \quad \text{ou} \quad \max_{\{\mathbf{x}\in U | h(\mathbf{x})=0\}} f \tag{5}$$

Définition 40 : Lagrangien.
On appelle **Lagrangien du problème** (5) la fonction à $n+1$ variables

$$L(\mathbf{x}, \lambda) = L(x_1, ..., x_n, \lambda) = f(\mathbf{x}) - \lambda h(\mathbf{x}). \tag{6}$$

Proposition 24 : Condition nécessaire du premier ordre.
Si $\mathbf{x}^* = (x_1^*, ..., x_2^*)$ est une solution du problème (5) telle que

$$\vec{\nabla} h(\mathbf{x}^*) \neq \mathbf{0}_n \quad \textbf{(Condition de qualification)}$$

alors il existe λ^* tel que $(\mathbf{x}^*, \lambda^*) = (x_1^*, ..., x_2^*, \lambda^*)$ soit un point stationnaire du Lagrangien, i.e. :

$$\vec{\nabla} L(\mathbf{x}^*, \lambda^*) = \mathbf{0}_n$$

ou encore :

$$\forall i \in \mathbb{N} \cap [1, n], \frac{\partial f}{\partial x_i}(\mathbf{x}^*) - \lambda^* \frac{\partial h}{\partial x_i}(\mathbf{x}^*) = 0, \text{ et } h(\mathbf{x}^*) = 0.$$

Définition 41 : Matrice Hessienne bordée.
La **matrice Hessienne bordée** du Lagrangien (6) en $(\mathbf{x}^*, \lambda^*)$ est la matrice $(n+1) \times (n+1)$, symétrique si f et h sont de classe C^2 en $\mathbf{x}^*$, définie par :

$$\bar{H}_h^f(\mathbf{x}^*, \lambda^*) = \begin{pmatrix} 0 & \frac{\partial h}{\partial x_1}(\mathbf{x}^*) & \dots & \frac{\partial h}{\partial x_n}(\mathbf{x}^*) \\ \frac{\partial h}{\partial x_1}(\mathbf{x}^*) & \frac{\partial^2 L}{\partial x_1^2}(\mathbf{x}^*, \lambda^*) & \dots & \frac{\partial^2 L}{\partial x_1 \partial x_n}(\mathbf{x}^*, \lambda^*) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h}{\partial x_n}(\mathbf{x}^*) & \frac{\partial^2 L}{\partial x_n \partial x_1}(\mathbf{x}^*, \lambda^*) & \dots & \frac{\partial^2 L}{\partial x_n^2}(\mathbf{x}^*, \lambda^*) \end{pmatrix}$$

Théorème 9 : Condition suffisante d'existence d'un extremum.
Soit $(\mathbf{x}^*, \lambda^*) = (x_1^*, ..., x_2^*, \lambda^*)$ un point stationnaire du Lagrangien. Alors,
- si les $n-1$ derniers mineurs diagonaux principaux Δ_k pour $k \in \mathbb{N} \cap [3, n+1]$ de $\bar{H}_h^f(\mathbf{x}^*, \lambda^*)$ sont tous strictement négatifs, alors $\mathbf{x}^*$ est un minimum local de f sous la contrainte $h(\mathbf{x}) = 0$.
- si les $n-1$ derniers mineurs diagonaux principaux Δ_k pour $k \in \mathbb{N} \cap [3, n+1]$ de $\bar{H}_h^f(\mathbf{x}^*, \lambda^*)$ sont alternativement > 0 et < 0, le dernier d'entre eux Δ_{n+1} étant du signe de $(-1)^n$ (i.e. $sign(\Delta_k) = sign((-1)^{k+1})$), alors $\mathbf{x}^*$ est un maximum local de f sous la contrainte $h(\mathbf{x}) = 0$.

Pour une fonction de deux variables $f : (x, y) \rightarrow f(x, y)$, le seul mineur qui nous intéresse est le déterminant de la matrice Hessienne bordée en (x^*, y^*, λ^*) :

$$\bar{H}_h^f(x^*, y^*, \lambda^*) = \begin{pmatrix} 0 & \frac{\partial h}{\partial x}(x^*, y^*) & \frac{\partial h}{\partial y}(x^*, y^*) \\ \\ \frac{\partial h}{\partial x}(x^*, y^*) & \frac{\partial^2 L}{\partial x^2}(x^*, y^*, \lambda^*) & \frac{\partial^2 L}{\partial x \partial y}(x^*, y^*, \lambda^*) \\ \\ \frac{\partial h}{\partial y}(x^*, y^*) & \frac{\partial^2 L}{\partial y \partial x}(x^*, y^*, \lambda^*) & \frac{\partial^2 L}{\partial y^2}(x^*, y^*, \lambda^*) \end{pmatrix}$$

Index

G

H

I

L

M

O

P

R

S

T

V