

R.Rhodes

**ANALYSE RÉELLE, OPTIMISATION LIBRE ET
SOUS CONTRAINTE**

Responsable : R. Rhodes
UE 13 DEGEAD
Année 2013-2014

R.Rhodes

Université Paris-Dauphine.

Url : <http://www.ceremade.dauphine.fr/~rhodes>

ANALYSE RÉELLE, OPTIMISATION LIBRE ET SOUS CONTRAINTE

R.Rhodes

**Responsable : R. Rhodes
UE 13 DEGEAD
Année 2013-2014**

TABLE DES MATIÈRES

1. Fonctions	1
1.1. Intervalles de \mathbb{R}	1
1.2. Voisinages	2
1.3. Définitions sur les fonctions	2
1.4. Opérations sur les fonctions	3
1.5. Exemples	3
1.6. Fonctions et relation d'ordre sur \mathbb{R}	4
2. Limites	7
2.1. Définitions	7
2.2. Détermination des limites	8
3. Fonctions continues	13
3.1. Continuité en un point	13
3.2. Continuité sur un intervalle	14
3.3. Propriétés fondamentales	15
4. Fonctions dérivables	17
4.1. Dérivabilité	17
4.2. Opérations sur les fonctions dérivables	19
4.3. Applications au sens de variations des fonctions	20
4.4. Primitive	20
5. Propriétés graphiques des fonctions	21
5.1. Symétries	21
5.2. Fonctions convexes et concaves	21
5.3. Branches infinies	23
5.4. Représentation graphique d'une fonction	23
6. Fonctions bijectives et réciproques	25
6.1. Définitions	25
6.2. Caractérisation des fonctions bijectives continues	25
6.3. Calcul de la dérivée d'une fonction réciproque	26
7. Différentielle et approximation affine	27
7.1. Variation absolue et différentielle	27
7.2. Approximation affine	27
7.3. Notation différentielle	29
8. Calculs approchés des variations	31
8.1. Fonction moyenne	31
8.2. Fonction marginale et variation absolue	31
8.3. Variation relative et élasticité	32
8.4. Récapitulatif	33
9. Formule de Taylor	35
9.1. Polynôme de Taylor	35
9.2. Formule de Taylor-Young à l'ordre n	36
10. Extrema des fonctions d'une variable	37

10.1. Définitions	37
10.2. Fonction dérivable sur un intervalle ouvert	38
10.3. Fonction dérivable sur un intervalle fermé borné	39
10.4. Fonction dérivable par morceaux sur un intervalle	40
11. Géométrie de \mathbb{R}^2 et \mathbb{R}^3	41
11.1. Les ensembles \mathbb{R}^2 et \mathbb{R}^3	41
11.2. Points et vecteurs	41
11.3. Produit scalaire, norme et distance associée	44
11.4. Produit cartésien de deux intervalles	46
11.5. Droite dans \mathbb{R}^2 et dans \mathbb{R}^3	46
11.6. Plan dans \mathbb{R}^3	49
11.7. Cercles et sphères	50
12. Topologie de l'ensemble \mathbb{R}^2	51
12.1. Boules	51
12.2. Frontière d'un sous-ensemble	52
12.3. Ouverts, fermés	52
12.4. Sous ensembles bornés - Compacts	53
13. Parties convexes de \mathbb{R}^2	55
13.1. Segment dans \mathbb{R}^2	55
13.2. Sous-ensembles convexes de \mathbb{R} et \mathbb{R}^2	55
14. Fonctions réelles de deux variables réelles	59
14.1. Définitions	59
14.2. Exemples fondamentaux	60
14.3. Relation d'ordre	60
14.4. Représentations graphiques	60
14.5. Opérations algébriques	61
15. Fonctions continues	63
15.1. Définitions	63
15.2. Opérations algébriques et composition	64
16. Dérivées partielles du premier ordre	67
16.1. Définitions	67
16.2. Opérations sur les dérivées partielles	68
16.3. Fonctions de classe \mathcal{C}^1	69
17. Développement limité d'ordre 1 et Différentielle	71
17.1. Variation absolue et différentielle	71
17.2. Développement limité d'ordre 1	72
17.3. Notation différentielle	74
18. Calculs approchés de variations	77
18.1. Fonctions marginales et variation absolue	77
18.2. Variation relative et élasticités	77
19. Dérivées partielles du deuxième ordre	81
19.1. Fonctions dérivées partielles secondes	81
19.2. Matrice Hessienne	81
19.3. Fonctions de classe \mathcal{C}^2	82
19.4. Opérations algébriques et composition	82
20. Développement limité d'ordre 2	83
20.1. Définitions	83
20.2. Forme quadratique associée à la matrice hessienne	84
20.3. Position du plan tangent par rapport à la surface représentative	85
21. Fonctions convexes ou concaves de deux variables	87
21.1. Définitions pour les fonctions de classe \mathcal{C}^1	87
21.2. Critère pour les fonctions de classe \mathcal{C}^2	88
21.3. Propriétés des fonctions convexes ou concaves	89
21.4. Convexité ou concavité locale	91

22. Généralités sur les extrema	93
22.1. Définitions	93
22.2. Simplification d'un problème d'extremum	95
23. Extrema libres pour deux variables	97
23.1. Conditions nécessaires ou conditions du premier ordre	97
23.2. Nature des points critiques	98
23.3. Quelques conseils pour résoudre un système non linéaire	101
24. Propriétés du gradient	103
24.1. Équations des tangentes aux courbes de niveau	103
24.2. Direction d'accroissement optimal d'une fonction	103
25. Extrema liés des fonctions de deux variables	105
25.1. Cas d'une liaison explicite	105
25.2. Conditions nécessaires ou conditions du premier ordre	106
25.3. Optimisation sur un compact	108
25.4. Conditions suffisantes	110

CHAPITRE 1

FONCTIONS

1.1. Intervalles de \mathbb{R}

Les intervalles jouent un rôle primordial en analyse car ce sont les parties de \mathbb{R} *sans trou*.

Définition 1.1. — Un intervalle de \mathbb{R} est une partie de \mathbb{R} qui contient tout nombre réel compris entre deux de ses éléments.

Ainsi une partie I est un intervalle de \mathbb{R} si et seulement si pour tout a et b deux éléments de I avec $a \leq b$, alors tout réel x vérifiant $a \leq x \leq b$ est encore un élément de I .

Proposition 1.2. — Il existe dix formes possibles pour un intervalle de \mathbb{R} :

1. \emptyset
2. $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ avec $a \leq b$
3. $]a, b[= \{x \in \mathbb{R} \mid a < x < b\}$ avec $a < b$
4. $[a, b[= \{x \in \mathbb{R} \mid a \leq x < b\}$ avec $a < b$
5. $]a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}$ avec $a < b$
6. $]a, +\infty[= \{x \in \mathbb{R} \mid a < x\}$
7. $[a, +\infty[= \{x \in \mathbb{R} \mid a \leq x\}$
8. $] - \infty, b] = \{x \in \mathbb{R} \mid x \leq b\}$
9. $] - \infty, b[= \{x \in \mathbb{R} \mid x < b\}$
10. $] - \infty, +\infty[= \mathbb{R}$.

Définition 1.3. — Un intervalle est fermé s'il contient ses éventuelles extrémités finies ; un intervalle est ouvert s'il ne contient aucune de ses éventuelles extrémités finies.

Exemple 1.4. — Les intervalles du type $[a, b]$, $[a, +\infty[$ et $] - \infty, b]$ sont fermés. Les intervalles du type $]a, b[$, $]a, +\infty[$ et $] - \infty, b[$ sont ouverts. Les intervalles du type $[a, b[$ et $]a, b]$ ne sont ni ouverts ni fermés.

Remarque 1.5. — 1. Si $a = b$ alors $[a, a] = \{a\}$. Ainsi, un singleton (*i.e.* un sous-ensemble ne contenant qu'un seul élément) est un intervalle fermé.

2. L'ensemble vide \emptyset et \mathbb{R} sont les deux seuls intervalles qui soient à la fois ouverts et fermés.

Définition 1.6. — On note $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ l'ensemble des réels auquel on adjoint les deux symboles $-\infty$ et $+\infty$. Cet ensemble est appelé *la droite numérique achevée*.

1.2. Voisinages

Définition 1.7. — Soit a un nombre réel et soit V est une partie de \mathbb{R} .

1. V est un *voisinage* de a s'il existe $\alpha > 0$ tel que l'intervalle ouvert $]a - \alpha, a + \alpha[$ soit inclus dans V .
2. V est un *voisinage à droite* de a s'il existe $\alpha > 0$ tel que l'intervalle $[a, a + \alpha[$ soit inclus dans V .
3. V est un *voisinage à gauche* de a s'il existe $\alpha > 0$ tel que l'intervalle $]a - \alpha, a]$ soit inclus dans V .
4. V est un *voisinage* de $+\infty$ ($-\infty$) s'il existe M tel que l'intervalle $]M, +\infty[$ (resp. $] - \infty, M]$) soit inclus dans V .
5. Si V est un voisinage de a alors $V \setminus \{a\}$ est appelé *voisinage épointé* de a . On définit de même les concepts de voisinage à droite épointé de a et de voisinage à gauche épointé de a .

Exemple 1.8. — L'ensemble $]1, 3[$ est un voisinage de 2, mais ce n'est pas un voisinage de 1, ni de 3. Par contre c'est un voisinage à gauche épointé de 3, ainsi qu'un voisinage à droite épointé de 1. L'ensemble $[1, 2[$ est un voisinage à droite de 1 et $]1, 2]$ est un voisinage à gauche de 2.

Notation 1.9. — Pour $a \in \overline{\mathbb{R}}$, on notera $\mathcal{V}(a)$ l'ensemble des voisinages de a .

1.3. Définitions sur les fonctions

1.3.1. Fonction numérique. —

Définition 1.10. — On appelle *fonction numérique* ou *fonction réelle d'une variable réelle* toute application f d'une partie \mathcal{D}_f de \mathbb{R} dans \mathbb{R} , notée

$$\begin{aligned} f : \mathcal{D}_f &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) \end{aligned}$$

L'ensemble \mathcal{D}_f est appelé *domaine de définition* de f .

Le domaine de définition est une donnée de la fonction. Ainsi les deux fonctions suivantes sont différentes :

$$\begin{array}{ccc} f : [0, +\infty[&\longrightarrow \mathbb{R} & \text{et} & g : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto x^2 & & x &\longmapsto x^2 \end{array}$$

Si dans la définition de la fonction, le domaine de définition n'est pas donné de façon explicite, alors on choisit de manière arbitraire l'ensemble de définition comme étant le plus grand sous-ensemble de \mathbb{R} sur lequel l'expression donnant $f(x)$ a mathématiquement un sens. En principe, nous n'étudierons que les fonctions sur un domaine de définition qui sera soit un intervalle, soit une réunion d'intervalles non réduits à un point.

Exemple 1.11. — On pourra dire indifféremment :

1. considérons la fonction f définie sur $]0, +\infty[$ par $f(x) = x^x$.
2. considérons la fonction f définie par $f(x) = \exp\{x \ln(x)\}$.

Il n'y a pas d'ambiguïté : dans les deux cas l'ensemble de définition est $\mathcal{D}_f =]0, +\infty[$. Maintenant si on dit : considérons la fonction g définie $g(x) = x^x$, alors apparaît une difficulté pour préciser le domaine de définition. En effet, on peut remarquer que $g(-2)$ a un sens et vaut $1/4$, mais on ne sait pas quel sens donner à $g(-\pi)$. Le plus grand sous-ensemble de \mathbb{R} sur lequel f a un sens est un ensemble avec des trous. C'est pourquoi on se limitera à l'intervalle $]0, +\infty[$.

Remarque 1.12. — Le nom donné à la variable n'a aucune importance. La fonction exponentielle définie sur \mathbb{R} peut s'écrire $x \mapsto \exp(x)$ ou $y \mapsto \exp(y)$. Dans les applications économiques, la variable est parfois notée :

- p lorsqu'elle représente le prix,
- q lorsqu'elle représente une quantité,
- L lorsqu'elle représente le travail (Labour en anglais).

Définition 1.13. — On dira qu'une fonction f est *définie sur un voisinage de a sauf éventuellement en a* pour traduire les deux cas suivants :

- \mathcal{D}_f est un voisinage de a et $a \in \mathcal{D}_f$,
- \mathcal{D}_f est un voisinage épointé de a et $a \notin \mathcal{D}_f$.

Les voisinages pourront être des voisinages à droite ou à gauche de a .

1.3.2. Image et graphe. —

Définition 1.14. — Soit $f : \mathcal{D}_f \rightarrow \mathbb{R}$, si E est une partie de \mathcal{D}_f alors on appelle *image* de E par f le sous-ensemble de \mathbb{R} , noté $f(E)$, défini par

$$f(E) = \{f(x) \mid x \in E\}.$$

On appelle *graphe* de f le sous-ensemble \mathcal{G}_f de \mathbb{R}^2 défini par

$$\mathcal{G}_f = \{(x, y) \in \mathbb{R}^2 \mid x \in \mathcal{D}_f \text{ et } y = f(x)\} = \{(x, f(x)) \mid x \in \mathcal{D}_f\}.$$

On peut représenter le graphe de f dans un plan muni d'un repère orthonormé en associant à tout couple $(x, f(x))$ le point d'abscisse x et d'ordonnée $y = f(x)$. On obtient alors la courbe représentative \mathcal{C}_f de f et on dit que $y = f(x)$ est l'équation de cette courbe. Pour simplifier on dira parfois graphe de f au lieu de courbe représentative de f .

1.4. Opérations sur les fonctions

À partir de deux fonctions f et g et d'un réel λ on peut construire d'autres fonctions à l'aide d'opérations algébriques et de la composition.

Définition 1.15. — Soient f et g deux fonctions définies respectivement sur \mathcal{D}_f et \mathcal{D}_g et soit λ un réel. On note

1. λf la fonction définie sur \mathcal{D}_f par $(\lambda f)(x) = \lambda f(x)$,
2. $f + g$ la fonction définie sur $\mathcal{D}_f \cap \mathcal{D}_g$ par $(f + g)(x) = f(x) + g(x)$,
3. $f g$ la fonction définie sur $\mathcal{D}_f \cap \mathcal{D}_g$ par $(f g)(x) = f(x)g(x)$,
4. f/g la fonction définie sur $\mathcal{D}_f \cap \mathcal{D}_{g^*}$ par $(f/g)(x) = f(x)/g(x)$ où

$$\mathcal{D}_{g^*} = \{x \in \mathcal{D}_g \mid g(x) \neq 0\},$$

5. $g \circ f$ la fonction définie sur $\mathcal{D}_{g \circ f}$ par $(g \circ f)(x) = g[f(x)]$ où

$$\mathcal{D}_{g \circ f} = \{x \in \mathcal{D}_f \mid f(x) \in \mathcal{D}_g\}.$$

Les ensembles $\mathcal{D}_f \cap \mathcal{D}_g$, $\mathcal{D}_f \cap \mathcal{D}_{g^*}$ et $\mathcal{D}_{g \circ f}$ ne sont pas toujours des intervalles ou des réunions d'intervalles non réduits à un point. On se ramènera donc à un sous-ensemble E si cela est nécessaire.

Exemple 1.16. — Un exemple très utilisé de la composition est la fonction $|f|$ définie sur \mathcal{D}_f par $|f|(x) = |f(x)|$.

1.5. Exemples

1.5.1. Fonctions économiques. — Voici quelques exemples qui apparaissent en Économie :

- une fonction de production lie la quantité q du facteur de production utilisé à la quantité $y = f(q)$ produite du bien fabriqué,
- une fonction de coût total associe, à chaque quantité x d'un facteur de production, le coût total $C(x)$ lié à sa production,
- une fonction de demande exprime la relation qui existe entre le prix unitaire p d'un bien sur un marché et la quantité $x = f(p)$ demandée de ce bien,
- une fonction d'utilité associe, à chaque quantité consommée x d'un bien, le niveau de satisfaction, souvent noté $u(x)$ du consommateur considéré.

Remarque 1.17. — Dans de nombreuses applications économiques le domaine de définition d'une fonction f sera un sous-ensemble de $[0, +\infty[$.

1.5.2. Fonctions polynômes et fractions rationnelles. —

Définition 1.18. — S'il existe des réels a_0, a_1, \dots, a_n avec $a_n \neq 0$ tels que

$$\forall x \in \mathbb{R}, \quad f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

alors la fonction f définie sur \mathbb{R} est appelée *fonction polynôme (ou polynômiale) de degré n* . Une fonction polynôme de degré 1 est appelée *fonction affine* et un quotient de deux fonctions polynômes est appelé *fraction rationnelle*.



Dans la définition d'une fonction polynôme, les exposants sont des nombres entiers positifs. Ainsi la fonction définie par $f(x) = 2x^{1/2} + 3x^{1/3}$ n'est pas une fonction polynôme.

Si f est une fonction affine alors il existe deux réels a et b tels que

$$\forall x \in \mathbb{R}, \quad f(x) = ax + b.$$

La courbe représentative de f est une droite D . Pour tracer cette droite il suffit de déterminer deux points de D . Si $M_1 = (x_1, y_1)$ et $M_2 = (x_2, y_2)$ sont deux points de D alors l'accroissement

$$\frac{y_2 - y_1}{x_2 - x_1}$$

est indépendant des points M_1 et M_2 et vaut a . Ce nombre s'appelle la *pente* de la droite D . Le point $B = (0, b)$ appartient à la droite D et s'appelle *ordonnée à l'origine*.

Exemple 1.19. — Considérons un consommateur qui consacre tout son revenu $R > 0$ à l'achat des quantités x et y de deux biens appelés X et Y . Si on note p le prix d'une unité de bien X et q le prix d'une unité de bien Y alors le budget du consommateur est l'ensemble

$$\mathcal{B}(p, q) = \{(x, y) \in [0, +\infty[^2 \mid px + qy \leq R\}.$$

Si on suppose que le consommateur dépense tout son revenu, alors l'ensemble des quantités (x, y) qu'il peut consommer est donné par l'équation $px + qy = R$. C'est l'équation d'une droite appelée droite budgétaire.

1.6. Fonctions et relation d'ordre sur \mathbb{R}

1.6.1. Fonctions bornées. —

Définition 1.20. — Soit E une partie non vide de \mathbb{R} et soit f une fonction définie sur E . On dit que f est *majorée* sur E s'il existe un réel A tel que pour tout x de E , on a $f(x) \leq A$, *i.e.*

$$(1.1) \quad \exists A \in \mathbb{R}, \quad \forall x \in E, \quad f(x) \leq A.$$

Chaque réel A vérifiant (1.1) est appelé un *majorant* de f ;

f est *minorée* sur E s'il existe un réel B tel que pour tout x de E , on a $f(x) \geq B$, *i.e.*

$$(1.2) \quad \exists B \in \mathbb{R}, \quad \forall x \in E, \quad f(x) \geq B.$$

Chaque réel B vérifiant (1.2) est appelé un *minorant* de f ;

f est *bornée* si elle est à la fois minorée et majorée.

Remarque 1.21. — Il est utile de retenir la caractérisation suivante : la fonction f est bornée si et seulement si

$$\exists M > 0, \quad \forall x \in E, \quad |f(x)| \leq M.$$

En effet, la formule précédente est équivalente à :

$$\exists M > 0, \quad \forall x \in E, \quad -M \leq f(x) \leq M.$$

Définition 1.22. — S'il existe $\bar{x} \in E$ tel que $f(\bar{x})$ est un majorant de f , alors on dit que f admet un *maximum* sur E atteint en \bar{x} . De même s'il existe $\underline{x} \in E$ tel que $f(\underline{x})$ est un minorant de f , alors on dit que f admet un *minimum* sur E atteint en \underline{x} .

Remarque 1.23. — L'hypothèse fondamentale de la microéconomie moderne est que les agents économiques (consommateur, producteur, investisseur, état) ont un comportement rationnel. Par exemple le producteur cherche à maximiser son profit, le consommateur cherche à maximiser son utilité et l'investisseur cherche à minimiser les risques. Une problématique récurrente en économie consistera à maximiser ou minimiser des fonctions.

1.6.2. Fonctions monotones. —

Définition 1.24. — Soit E une partie non vide de \mathbb{R} et soit f une fonction définie sur E . On dit que

– f est *croissante* sur E si

$$\forall (x_1, x_2) \in E^2, \quad x_1 < x_2 \implies f(x_1) \leq f(x_2),$$

– f est *décroissante* sur E si

$$\forall (x_1, x_2) \in E^2, \quad x_1 < x_2 \implies f(x_1) \geq f(x_2),$$

– f est *strictement croissante* sur E si

$$\forall (x_1, x_2) \in E^2, \quad x_1 < x_2 \implies f(x_1) < f(x_2),$$

– f est *strictement décroissante* sur E si

$$\forall (x_1, x_2) \in E^2, \quad x_1 < x_2 \implies f(x_1) > f(x_2),$$

– f est *monotone* sur E si f est décroissante sur E ou croissante sur E ,

– f est *strictement monotone* sur E si f est strictement décroissante sur E ou strictement croissante sur E .

Une fonction f n'est pas forcément monotone sur \mathcal{D}_f . Cependant on peut partager \mathcal{D}_f en intervalles sur lesquels f est monotone. Cela s'appelle *étudier les variations de f* .

Exemple 1.25. — En Économie, les fonctions de coût de production sont des fonctions croissantes de la quantité produite. Les fonctions de demande d'un consommateur sont généralement des fonctions décroissantes du prix.

1.6.3. Comparaison des fonctions. —

Définition 1.26. — Soient f et g deux fonctions et soit E une partie de $\mathcal{D}_f \cap \mathcal{D}_g$. On dit que

– $f = g$ sur E si pour tout $x \in E$, on a $f(x) = g(x)$,

– $f \leq g$ sur E si pour tout $x \in E$, on a $f(x) \leq g(x)$,

– $f < g$ sur E si pour tout $x \in E$, on a $f(x) < g(x)$.

Remarque 1.27. — Lorsque la fonction g est nulle, on obtient la définition d'une fonction négative ou strictement négative. Lorsque la fonction f est nulle, on obtient la définition d'une fonction positive ou strictement positive.

CHAPITRE 2

LIMITES

2.1. Définitions

2.1.1. Condition préalable. — On fixe $a \in \overline{\mathbb{R}}$, i.e. a est un nombre réel fini ou $a \in \{+\infty, -\infty\}$. Considérons une fonction f définie sur \mathcal{D}_f . On supposera dorénavant que \mathcal{D}_f est un intervalle ou une réunion d'intervalles non réduits à un point. On s'intéresse à la question suivante :

Que devient $f(x)$ lorsque x se rapproche de a tout en restant dans \mathcal{D}_f ?

Pour que la condition *x se rapproche de a tout en restant dans \mathcal{D}_f* ait un sens, nous énoncerons donc la *condition préalable* suivante :

- ou bien \mathcal{D}_f contient un voisinage épointé de a ;
- ou bien \mathcal{D}_f contient un voisinage à droite (ou à gauche) épointé de a .

La fonction f peut être définie en a ou ne pas être définie en a .

2.1.2. Définition. —

Exemple 2.1. — Soit f la fonction définie sur $\mathcal{D}_f = \mathbb{R} \setminus \{1\}$ par

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

On cherche la limite de f en $a = 1$. La condition préalable est satisfaite puisque \mathcal{D}_f est un voisinage épointé de 1. Notons que $f(x) = x + 1$, ainsi $f(x)$ se rapproche aussi près qu'on le souhaite de 2 pourvu que x soit assez proche de 1. On dira que f admet 2 comme limite lorsque x tend vers 1.

Pour donner un sens rigoureux à l'expression : *$f(x)$ se rapproche aussi près qu'on le souhaite de ℓ pourvu que x soit assez proche de a* , on considère la définition suivante.

Définition 2.2. — Soit $\ell \in \overline{\mathbb{R}}$, sous la condition préalable pour $a \in \overline{\mathbb{R}}$, on dit que f admet ℓ comme limite en a , si tout voisinage de ℓ contient l'image par f d'un voisinage (assez petit) épointé de a , c'est-à-dire, pour tout voisinage V de ℓ , il est toujours possible de trouver un voisinage épointé U de a , tel que $f(U \cap \mathcal{D}_f)$ soit inclus dans V .

Proposition 2.3. — Soit $\ell \in \overline{\mathbb{R}}$, sous la condition préalable pour $a \in \overline{\mathbb{R}}$, la fonction f admet ℓ comme limite en a si et seulement si :

- Si $\ell \in \mathbb{R}$ et $a \in \mathbb{R}$,

$$\forall \varepsilon > 0, \quad \exists \alpha > 0, \quad f(]a - \alpha, a + \alpha[\cap \mathcal{D}_f) \subset]\ell - \varepsilon, \ell + \varepsilon[,$$

- Si $\ell \in \mathbb{R}$ et $a = +\infty$,

$$\forall \varepsilon > 0, \quad \exists B > 0, \quad f(]B, +\infty[\cap \mathcal{D}_f) \subset]\ell - \varepsilon, \ell + \varepsilon[,$$

- Si $\ell = +\infty$ et $a \in \mathbb{R}$,

$$\forall A > 0, \quad \exists \alpha > 0, \quad f(]a - \alpha, a + \alpha[\cap \mathcal{D}_f) \subset]A, +\infty[,$$

- Si $\ell = +\infty$ et $a = +\infty$,

$$\forall A > 0, \quad \exists B > 0, \quad f(]B, +\infty[\cap \mathcal{D}_f) \subset]A, +\infty[.$$

2.1.3. Propriétés de la limite. —

Proposition 2.4. — Si une fonction f admet une limite $\ell \in \overline{\mathbb{R}}$ en $a \in \overline{\mathbb{R}}$ alors cette limite est unique.

Lorsque f admet ℓ comme limite en a alors on note

$$\lim_{x \rightarrow a} f(x) = \ell.$$

On ne doit pas écrire $\lim_{x \rightarrow a} f(x)$ avant d'avoir justifié l'existence de cette limite.



Si une fonction possède une limite ℓ finie (i.e. $\ell \in \mathbb{R}$) en a et si a appartient à \mathcal{D}_f , alors on n'a pas forcément $\ell = f(a)$. Par exemple on considère f la fonction définie sur \mathbb{R} par $f(x) = -1$ si $x \neq 0$ et $f(0) = 1$. Alors f admet -1 comme limite en 0 mais $-1 \neq f(0) = 1$.

2.1.4. Limite à droite, limite à gauche. — Lorsqu'une fonction f n'a pas de limite lorsque x tend vers a , on peut se limiter à étudier le comportement de $f(x)$ lorsque x tend vers a à droite de a ($x > a$) ou à gauche de a ($x < a$).

Définition 2.5. — Soit $\ell \in \overline{\mathbb{R}}$, sous la condition préalable pour $a \in \mathbb{R}$, on dit que f admet ℓ comme limite à droite en a , si tout voisinage de ℓ contient l'image par f d'un voisinage (assez petit) épointé à droite de a . C'est-à-dire, pour tout voisinage V de ℓ , il est toujours possible de trouver un voisinage épointé U de a , tel que $f(U \cap]a, +\infty[\cap \mathcal{D}_f)$ est inclus dans V .

Remarque 2.6. — En remplaçant $]a, +\infty[$ par $] - \infty, a[$ dans la définition précédente, on obtient la notion de limite à gauche.

Proposition 2.7. — Soit $\ell \in \overline{\mathbb{R}}$, sous la condition préalable pour $a \in \mathbb{R}$, la fonction f admet ℓ comme limite à droite en a si et seulement si :

– Si $\ell \in \mathbb{R}$,

$$\forall \varepsilon > 0, \quad \exists \alpha > 0, \quad f(]a, a + \alpha[\cap \mathcal{D}_f) \subset]\ell - \varepsilon, \ell + \varepsilon[.$$

– Si $\ell = +\infty$ et $a \in \mathbb{R}$,

$$\forall A > 0, \quad \exists \alpha > 0, \quad f(]a, a + \alpha[\cap \mathcal{D}_f) \subset]A, +\infty[.$$

Proposition 2.8. — Si une fonction f admet une limite $\ell \in \overline{\mathbb{R}}$ à gauche (respectivement, à droite) en $a \in \overline{\mathbb{R}}$ alors cette limite est unique.

Lorsque f admet ℓ comme limite à gauche en a (respectivement, à droite) alors on note

$$\lim_{x \rightarrow a^-} f(x) = \ell \quad (\text{respectivement, } \lim_{x \rightarrow a^+} f(x) = \ell).$$

Exemple 2.9. —

1. Soit f la fonction définie sur \mathbb{R}^* par $f(x) = |x|/x$. La fonction admet -1 comme limite à gauche en 0 et 1 comme limite à droite en 0 .
2. Soit f la fonction définie sur \mathbb{R}^* par $f(x) = x \exp\{1/x\}$. On peut vérifier que

$$\lim_{x \rightarrow 0^-} f(x) = 0 \quad \text{et} \quad \lim_{x \rightarrow 0^+} f(x) = +\infty.$$

Proposition 2.10. — Soit $\ell \in \mathbb{R}$, sous la condition préalable pour $a \in \mathbb{R}$, la fonction f admet ℓ comme limite en a si et seulement si f admet ℓ comme limite à droite et à gauche en a , i.e.

$$\lim_{x \rightarrow a} f(x) = \ell \iff \lim_{x \rightarrow a^-} f(x) = \lim_{x \rightarrow a^+} f(x) = \ell.$$

2.2. Détermination des limites

En pratique on n'utilisera pas la définition de la limite. En effet, la connaissance des limites des fonctions *usuelles* permettra avec les opérations sur les limites de déterminer les limites des fonctions plus élaborées.

2.2.1. Opérations algébriques sur les limites. —

Théorème 2.11. — Soient f et g deux fonctions sous la condition préalable pour $a \in \overline{\mathbb{R}}$ et admettant des limites respectives ℓ et ℓ' (dans $\overline{\mathbb{R}}$) dans les mêmes conditions (en a , à droite en a ou à gauche en a). Alors les fonctions : $f + g$, λf , fg et $1/f$ ont dans les cas précisés ci-après une limite dont la valeur est indiquée dans le tableau correspondant :

– Pour $f + g$ on a le tableau

$f + g$	$\ell \in \mathbb{R}$	$+\infty$	$-\infty$
$\ell' \in \mathbb{R}$	$\ell + \ell'$	$+\infty$	$-\infty$
$+\infty$	$+\infty$	$+\infty$?
$-\infty$	$-\infty$?	$-\infty$

– Pour λf on a le tableau

λf	$\ell \in \mathbb{R}$	$+\infty$	$-\infty$
$\lambda = 0$	0	0	0
$\lambda > 0$	$\lambda \ell$	$+\infty$	$-\infty$
$\lambda < 0$	$\lambda \ell$	$-\infty$	$+\infty$

– Pour fg on a le tableau

fg	0	$\ell < 0$	$\ell > 0$	$+\infty$	$-\infty$
0	0	0	0	?	?
$\ell' < 0$	0	$\ell \ell'$	$\ell \ell'$	$-\infty$	$+\infty$
$\ell' > 0$	0	$\ell \ell'$	$\ell \ell'$	$+\infty$	$-\infty$
$+\infty$?	$-\infty$	$+\infty$	$+\infty$	$-\infty$
$-\infty$?	$+\infty$	$-\infty$	$-\infty$	$+\infty$

– Pour $1/f$ on a le tableau

f	$\ell \neq 0$	$\ell = 0$	$+\infty$	$-\infty$
$1/f$	$1/\ell$?	0	0



Tous les problèmes ne sont pas réglés. En effet, les cases marquées d'un point d'interrogation correspondent à des formes indéterminées, qui nécessitent une étude spécifique dans chaque cas.

Rappelons ce résultat très utile.

Proposition 2.12. — Toute fonction polynôme admet en $\pm\infty$ la même limite que son terme de plus haut degré. Toute fraction rationnelle admet en $\pm\infty$ la même limite que le quotient des termes de plus haut degré du numérateur et du dénominateur.

2.2.2. Composition des limites. —

Théorème 2.13. — Soient f et g deux fonctions définies sur \mathcal{D}_f et \mathcal{D}_g et soient $(a, \ell, k) \in \overline{\mathbb{R}}^3$. On suppose que :

- (i) la fonction $g \circ f$ satisfait la condition préalable en a ,
 - (ii) la fonction f admet ℓ comme limite en a ,
 - (iii) la fonction g admet k comme limite en ℓ ,
- alors la fonction $g \circ f$ admet k comme limite en a .

On réécrit ce théorème sous la forme

$$\lim_{x \rightarrow a} f(x) = \ell \quad \text{et} \quad \lim_{y \rightarrow \ell} g(y) = k \quad \implies \quad \lim_{x \rightarrow a} g \circ f(x) = k.$$

Ce théorème permet par des changements de variables de se ramener à des limites connues.

Exemple 2.14. — Soit h la fonction définie sur \mathbb{R} par $h(x) = x^4 \exp\{-x^2\}$. On cherche à déterminer (forme indéterminée) si h admet une limite en $+\infty$. On considère f et g les deux fonctions définies sur \mathbb{R} par $f(x) = x^2$

et $g(u) = u^2 \exp\{-u\}$. On a alors $h = g \circ f$ (ce qui revient à poser $u = x^2$). On sait par le théorème des croissances comparées que

$$\lim_{x \rightarrow +\infty} f(x) = +\infty \quad \text{et} \quad \lim_{u \rightarrow +\infty} g(u) = 0.$$

En appliquant le théorème de composition, on obtient alors que h admet 0 comme limite en $+\infty$.

Une application particulière de la composition donne la proposition suivante.

Proposition 2.15. — Si f admet $\ell \in \overline{\mathbb{R}}$ comme limite en $a \in \overline{\mathbb{R}}$ alors $|f|$ admet $|\ell|$ comme limite en a (par convention $|\infty| = +\infty$), i.e.

$$\lim_{x \rightarrow a} f(x) = \ell \quad \implies \quad \lim_{x \rightarrow a} |f(x)| = |\ell|.$$

2.2.3. Limites des fonctions monotones. —

Proposition 2.16. — Soient $a \in \mathbb{R} \cup \{+\infty\}$ et $b \in \mathbb{R} \cup \{-\infty\}$.

- (i) Si f est une fonction croissante non majorée sur $]a, b[$ alors $\lim_{x \rightarrow b^-} f(x) = +\infty$.
- (ii) Si f est une fonction croissante non minorée sur $]a, b[$ alors $\lim_{x \rightarrow a^+} f(x) = -\infty$.
- (i) Si f est une fonction décroissante non minorée sur $]a, b[$ alors $\lim_{x \rightarrow b^-} f(x) = -\infty$.
- (ii) Si f est une fonction décroissante non majorée sur $]a, b[$ alors $\lim_{x \rightarrow a^+} f(x) = +\infty$.

2.2.4. Passage à la limite dans les inégalités. — Voici un résultat très utilisé pour calculer des limites. Il est connu sous le nom de "théorème des gendarmes".

Théorème 2.17 (Théorème des gendarmes). — Soit $a \in \overline{\mathbb{R}}$ et soient f, g et h trois fonctions dont le domaine de définition contient un voisinage épointé V de a et vérifiant

$$\forall x \in V, \quad f(x) \leq h(x) \leq g(x).$$

Si f et g ont la même limite finie $\ell \in \mathbb{R}$ lorsque x tend vers a alors h admet ℓ comme limite en a , i.e.

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = \ell \quad \implies \quad \lim_{x \rightarrow a} h(x) = \ell.$$

Remarque 2.18. — Pour montrer que $\lim_{x \rightarrow a} f(x) = \ell$, on peut considérer la fonction $h(x) = |f(x) - \ell|$. Cette fonction est trivialement minorée par 0. Il suffit donc de majorer h par une fonction g dont on sait que $\lim_{x \rightarrow a} g(x) = 0$.

Les deux propositions suivantes fournissent des résultats moins performants que le théorème des gendarmes, mais donnent des indications utiles.

Proposition 2.19. — Soit $a \in \overline{\mathbb{R}}$ et soient f, g deux fonctions dont l'ensemble de définition contient un voisinage épointé V de a tel que

$$\forall x \in V, \quad f(x) \geq g(x).$$

Alors

- (i) si f et g admettent une limite finie en a alors $\lim_{x \rightarrow a} f(x) \geq \lim_{x \rightarrow a} g(x)$,
- (ii) si f tend vers $-\infty$ alors g aussi, et si g tend vers $+\infty$ alors f aussi.



Considérons f et g définies sur \mathbb{R} par $f(x) = 0$ et $g(x) = \exp\{-x\}$. On a $f(x) < g(x)$ pour tout $x \in \mathbb{R}$ mais f et g admettent la même limite en $+\infty$. Ainsi en général, par passage à la limite, les inégalités strictes deviennent des inégalités larges.

Proposition 2.20. — Soit f une fonction dont le domaine de définition contient un voisinage épointé V de $a \in \overline{\mathbb{R}}$. Si f admet une limite $\ell \in \overline{\mathbb{R}}$ en a alors

- (i) si f est de signe constant sur V alors ℓ est de même signe, par exemple

$$\forall x \in V, \quad f(x) \geq 0 \quad \implies \quad \ell \geq 0,$$

- (ii) si ℓ est non nulle, il existe un voisinage épointé W de a tel que f est non nulle et garde le même signe que ℓ , par exemple

$$\ell > 0 \quad \implies \quad \exists W \in \mathcal{V}(a), \quad \forall x \in W, \quad f(x) > 0.$$



Dans le (i) de la proposition précédente, si on a $f(x) > 0$, la conclusion reste en général $\ell \geq 0$. Considérons par exemple la fonction f définie sur $]0, +\infty[$ par $f(x) = 1/x$. On a $f(x) > 0$ pour tout $x > 0$ mais $\lim_{x \rightarrow +\infty} f(x) = 0$.

Méthode 2.21. — Pour rechercher une limite :

1. Connaître toutes les limites usuelles, y compris les croissances comparées des fonctions \ln , \exp et puissance.
2. Essayer les opérations algébriques sur les limites, la composition des limites (changement de variables), le théorème d'encadrement, passage à la limite dans les inégalités, limites des fonctions monotones.
3. Pour les cas d'indétermination : passer au logarithme (pour les indéterminations exponentielles), mettre en facteur le terme prépondérant, multiplier par la quantité conjuguée.

CHAPITRE 3

FONCTIONS CONTINUES

Les expressions de la vie courante : *évolution continue*, *politique continue*, évoquent l'absence de modifications brutales instantanées. Dans ce chapitre nous allons donner des définitions précises de ce concept qui est fondamental en analyse.

3.1. Continuité en un point

Exemple 3.1. — Considérons la fonction f définie sur \mathbb{R} par :

$$f(x) = \begin{cases} x^2 + 1 & \text{si } x \geq 0 \\ -x - 1 & \text{si } x < 0. \end{cases}$$

On constate un saut dans les valeurs prises par $f(x)$ lorsque x passe par 0. En effet on a

$$\lim_{x \rightarrow 0^-} f(x) = -1 \quad \text{et} \quad \lim_{x \rightarrow 0^+} f(x) = 1.$$

Autrement dit, des petites variations de x au voisinage de 0 ne conduisent pas à des petites variations de $f(x)$.

En utilisant le concept de limite, ceci conduit à poser les définitions suivantes.

Définition 3.2. — Soit a un réel et soit f une fonction définie sur \mathcal{D}_f qui contient un voisinage de a (éventuellement à droite ou à gauche). La fonction f est *continue en a* si la limite en a existe et vaut $f(a)$, *i.e.*

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

La fonction f est *continue à droite en a* si la limite à droite en a existe et vaut $f(a)$, *i.e.*

$$\lim_{x \rightarrow a^+} f(x) = f(a).$$

La fonction f est *continue à gauche en a* si la limite à gauche en a existe et vaut $f(a)$, *i.e.*

$$\lim_{x \rightarrow a^-} f(x) = f(a).$$

Remarque 3.3. —

1. Pour qu'une fonction soit continue (éventuellement à droite ou à gauche) en a il faut qu'elle soit définie en a , *i.e.* $a \in \mathcal{D}_f$.
2. Lorsque f est définie sur un voisinage de a , on a l'équivalence : f est continue en a si et seulement si elle est continue à droite et à gauche de a .

Exemple 3.4. — La fonction *partie entière*, notée \mathbb{E} , est définie sur \mathbb{R} par $\mathbb{E}(x)$ est le seul entier $k \in \mathbb{Z}$ tel que $k \leq x < k + 1$. On peut vérifier que \mathbb{E} est continue en tout point $a \notin \mathbb{Z}$ et que \mathbb{E} est continue à droite et discontinue à gauche en chaque entier relatif $k \in \mathbb{Z}$.

3.1.1. Prolongement par continuité. — Soit f une fonction définie sur \mathcal{D}_f qui contient un voisinage épointé de a . Si la limite en a existe et vaut $\ell \in \mathbb{R}$ alors on considère la fonction g définie sur $\mathcal{D}_g = \mathcal{D}_f \cup \{a\}$ par

$$g(x) = \begin{cases} f(x) & \text{si } x \in \mathcal{D}_f \\ \ell & \text{si } x = a. \end{cases}$$

Par construction, \mathcal{D}_g contient un voisinage de a et g est continue en a . On dit que l'on a *prolongé f par continuité en a* . On peut définir de même le prolongement par continuité de f à droite de a ou à gauche de a . On garde souvent la même notation pour la fonction et son prolongement par continuité.

Exemple 3.5. —

1. Soit f définie sur \mathbb{R}^* par $f(x) = x \ln |x|$. On prolonge f sur \mathbb{R} en posant $f(0) = 0$.
2. Soit f définie sur \mathbb{R}^* par $f(x) = x \ln x$ si $x > 0$ et $f(x) = 1/x$ si $x < 0$. On prolonge f par continuité à droite en 0 en posant $f(0) = 0$.
3. Soit f définie sur \mathbb{R}^* par $f(x) = 1/x$. On ne peut pas prolonger f par continuité en 0 puisque les limites à droite et à gauche ne sont pas réelles.

Lorsqu'un prolongement par continuité est possible, on l'opère systématiquement.

3.1.2. Opérations sur les fonctions continues en un point. —

Proposition 3.6. — Soient λ un réel et f, g deux fonctions continues en un point a . Alors les fonctions $\lambda f, f + g, f - g$ et fg sont continues en a . Si $g(a) \neq 0$ alors la fonction f/g est continue en a .

Proposition 3.7. — Si f est continue en a et g est continue en $f(a)$ alors la fonction $g \circ f$ est continue en a .

3.2. Continuité sur un intervalle

En pratique la notion de continuité la plus intéressante et la plus utile est la suivante.

Définition 3.8. —

- Si I est un intervalle ouvert de \mathbb{R} , on dit que f est dite *continue sur I* si f est continue en tout point de I .
- f est continue sur $[a, b]$ si f est continue sur $]a, b[$, continue à droite en a et continue à gauche en b .
- f est continue sur $]a, b]$ si f est continue sur $]a, b[$ et continue à gauche en b . Cette définition reste valable si $a = -\infty$.
- f est continue sur $[a, b[$ si f est continue sur $]a, b[$ et continue à droite en a . Cette définition reste valable si $b = +\infty$.

Remarque 3.9. —

1. La courbe représentative d'une fonction continue sur un intervalle se trace sans lever le crayon.
2. Si I et J sont deux intervalles ouverts non vides, et si f est continue sur I et sur J , alors est continue sur tout point de la réunion $I \cup J$. On dit que f est continue sur $I \cup J$.



Lorsque l'intervalle n'est pas ouvert, les définitions ci-dessus peuvent entraîner quelques bizarreries. Par exemple la fonction partie entière est continue sur $[1, 2[$ et sur $[2, 3[$ mais elle n'est pas continue en 2. Donc elle n'est pas continue sur l'intervalle ouvert $]1, 3[$.

Notation 3.10. — Soit I un intervalle non vide de \mathbb{R} , on note $\mathcal{C}^0(I)$ l'ensemble des fonctions continues sur I . Si $f \in \mathcal{C}^0(I)$ alors f est dite de classe \mathcal{C}^0 sur I .

La plupart des fonctions qui apparaissent dans les modèles économiques (coût de production, utilité, demande) sont continues.

3.2.1. Opérations sur les fonctions continues sur un intervalle. —

Théorème 3.11. — Soient I un intervalle de \mathbb{R} , λ un réel et f et g deux fonctions continues sur I , alors

- (i) $\lambda f, f + g, f - g$ et fg sont continues sur I ,
- (ii) si de plus $g(x) \neq 0$ pour chaque $x \in I$, alors f/g est continue sur I .

En application de ce théorème on obtient que toutes les fonctions polynômes sont continues sur \mathbb{R} . De même toute fraction rationnelle $R = P/Q$ avec P et Q deux fonctions polynômes, est continue sur tout intervalle ne contenant pas les racines de Q .

Théorème 3.12. — Soient I et J deux intervalles de \mathbb{R} . Si les conditions suivantes sont satisfaites :

- f est continue sur I ,
- g est continue sur J ,
- $f(I) \subset J$,

alors la fonction composée $g \circ f$ est continue sur I .

Ce sont les deux théorèmes précédents qui vont nous permettre de justifier qu'une fonction construite à partir des fonctions usuelles (somme, produit, quotient et composée) est continue sur un intervalle. Par exemple, si f est une fonction continue sur un intervalle I alors $|f|$ est continue sur I .

3.3. Propriétés fondamentales

3.3.1. Fonction continue sur un intervalle. —

Théorème 3.13 (Théorème des valeurs intermédiaires). — Soient f une fonction continue sur un intervalle I et $\alpha < \beta$ deux réels de I . Toutes les valeurs comprises strictement entre $f(\alpha)$ et $f(\beta)$ sont atteintes par f en au moins un point de $] \alpha, \beta [$. Par exemple si $f(\alpha) < f(\beta)$ on a

$$\forall y \in]f(\alpha), f(\beta)[, \quad \exists x \in]\alpha, \beta[, \quad y = f(x).$$

Corollaire 3.14. — Si f est une fonction continue sur un intervalle I , alors $f(I)$ est un intervalle.

Les intervalles I et $f(I)$ ne sont pas nécessairement de même nature. Par exemple :

1. Si $f(x) = x^2$ et $I =] - 1, 1 [$ alors $f(I) = [0, 1 [$.
2. Si $g(x) = \sin x$ et $I = \mathbb{R}$ alors $g(I) = [-1, 1]$.



En général, si $I =]a, b [$ l'intervalle $f(I)$ n'est pas d'extrémités $f(a)$ et $f(b)$. Par contre si f est une fonction continue croissante sur I alors $f(I) = [f(a), f(b)[$.

Corollaire 3.15. — Une fonction continue sur un intervalle I ne peut changer de signe sur I qu'en s'annulant.

Autrement dit, s'il existe deux points a et b de I tels que $f(a)f(b) < 0$, alors il existe $c \in]a, b [$ tel que $f(c) = 0$. Le point c vérifiant cette propriété n'est pas forcément unique.

3.3.2. Fonctions continues sur un segment. —

Théorème 3.16. — Si f une fonction continue sur un segment $[a, b]$ alors l'image $f([a, b])$ est un segment $[m, M]$. On en déduit que f est bornée sur $[a, b]$ et qu'elle atteint ses bornes, i.e.

$$\exists(\alpha, \beta) \in [a, b]^2, \quad \forall x \in [a, b], \quad m = f(\alpha) \leq f(x) \leq f(\beta) = M.$$

Ce théorème est fondamental, c'est le premier théorème d'optimisation que nous voyons. En d'autres termes, il nous montre qu'une fonction continue sur un ensemble compact admet un maximum et un minimum et que ceux-ci sont atteints. Nous verrons qu'il est utile dans certain cas (on pourra consulter l'exemple 10.5 page 37). De plus nous verrons le même théorème appliqué à des fonctions de 2 variables (on pourra consulter le paragraphe 25.3 page 108).

CHAPITRE 4

FONCTIONS DÉRIVABLES

Dans les relations entre variables économiques, la question centrale est : dans quelle mesure la variation d'une variable affecte-t-elle les autres ? Lorsque ces relations sont décrites par des fonctions l'effet de la variation d'une variable sur l'autre s'obtient grâce à la notion de dérivée.

4.1. Dérivabilité

4.1.1. Dérivée en un point. —

Définition 4.1. — Soit a un réel et soit f une fonction définie sur une partie \mathcal{D}_f qui contient un voisinage de a . On appelle *taux d'accroissement de f en a* , noté $\theta_a(x)$, la fonction quotient définie par

$$\forall x \in \mathcal{D}_f \setminus \{a\}, \quad \theta_a(x) = \frac{f(x) - f(a)}{x - a}.$$

L'exemple classique pour illustrer cette notion est le suivant : si $x > a$ et si $f(x) - f(a)$ représente la distance parcourue entre les instants $t = a$ et $t = x$, le taux $\theta_a(x)$ représente la vitesse moyenne entre ces deux instants, c'est-à-dire la distance parcourue en moyenne par unité de temps. Si f est affine, cette vitesse moyenne est constante : c'est la pente. Si f n'est pas une fonction affine, cette vitesse moyenne dépend de x et de a . D'autre part, le compteur de la voiture affiche une autre information : la vitesse instantanée, c'est-à-dire la vitesse moyenne entre deux instants très proches. Cette vitesse instantanée est en fait une limite. Il ne reste plus qu'à formaliser rigoureusement ces notions.

Définition 4.2 (Dérivée en un point). — Soient a un réel et f une fonction définie sur \mathcal{D}_f . Si \mathcal{D}_f contient un voisinage de a , on dit que la fonction f est *dérivable en a* si la fonction θ_a admet une limite finie en a . Cette limite est notée $f'(a)$, *i.e.*

$$f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}.$$

Définition 4.3 (Dérivées à droite et à gauche). —

– Si \mathcal{D}_f contient un voisinage à gauche de a , on dit que la fonction f est *dérivable à gauche de a* si la fonction θ_a admet une limite finie à gauche de a . Cette limite est notée $f'_g(a)$ ou $f'(a^-)$, *i.e.*

$$f'_g(a) = f'(a^-) = \lim_{x \rightarrow a^-} \frac{f(x) - f(a)}{x - a}.$$

– Si \mathcal{D}_f contient un voisinage à droite de a , on dit que la fonction f est *dérivable à droite de a* si la fonction θ_a admet une limite finie à droite de a . Cette limite est notée $f'_d(a)$ ou $f'(a^+)$, *i.e.*

$$f'_d(a) = f'(a^+) = \lim_{x \rightarrow a^+} \frac{f(x) - f(a)}{x - a}.$$

Remarque 4.4. —

1. Lorsque f est dérivable en a , le nombre $f'(a)$ est appelé le nombre dérivé de f en a .
2. Si $f'(a)$ existe, en posant $x = a + h$, on peut écrire

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

En Économie, l'accroissement h est souvent noté Δx .

3. On a l'équivalence suivante : f est dérivable en a si et seulement si f est dérivable à gauche et à droite de a et $f'_g(a) = f'_d(a)$. On a alors $f'(a) = f'_g(a) = f'_d(a)$.

Exemple 4.5. —

1. Soit f définie sur \mathbb{R} par $f(x) = x^2$ et soit $a \in \mathbb{R}$. On a

$$\forall h \neq 0, \quad \theta_a(h) = \frac{f(a+h) - f(a)}{h} = \frac{(a+h)^2 - a^2}{h} = 2a + h$$

donc f est dérivable en a et $f'(a) = 2a$.

2. Soit g définie sur \mathbb{R} par $g(x) = |x|$. Pour $a = 0$, on a

$$\forall h \neq 0, \quad \theta_0(h) = \frac{g(h) - g(0)}{h} = \frac{|h|}{h}.$$

Ainsi θ_0 est dérivable à droite et à gauche en 0, mais $g'_d(0) = 1$ et $g'_g(0) = -1$. Comme les dérivées à droite et à gauche ne coïncident pas, la fonction g n'est pas dérivable en 0.

On admet le résultat suivant.

Proposition 4.6. — Si f est dérivable en a alors f est continue en a .

Cette proposition affirme que la continuité de f en a est une condition nécessaire à la dérivabilité de f en a . En particulier, si f n'est pas continue en a alors f n'est pas dérivable en a .



La réciproque de la proposition précédente est fautive. Une fonction peut être continue en un point sans être dérivable en ce même point. Par exemple la fonction f définie sur \mathbb{R} par $f(x) = |x|$, est continue en 0 mais pas dérivable en 0.

Soit \mathcal{G} le graphe d'une fonction f dans un plan rapporté à un repère orthonormé. Soit A le point du plan de coordonnées $(a, f(a))$ et soit M un point de coordonnées $(x, f(x))$ avec $a \neq x \in \mathcal{D}_f$. On suppose que f est continue en a . On considère \mathcal{D} la droite passant par les points A et M . La pente de cette droite est $\theta_a(x)$. On dit que \mathcal{G} admet une tangente en a si $\theta_a(x)$ admet une limite $\ell \in \overline{\mathbb{R}}$ lorsque x tend vers a . La tangente à \mathcal{G} en a est alors la droite passant par A et de pente ℓ . Lorsque $\ell \in \{+\infty, -\infty\}$, la tangente est dite verticale.

On déduit de la définition de la dérivabilité, les équivalences suivantes.

Proposition 4.7 (Dérivabilité et tangente). — Une fonction f est dérivable en a si et seulement si la graphe de f admet une tangente non verticale en $A = (a, f(a))$ d'équation

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

4.1.2. Fonction dérivée – Dérivée sur un intervalle. —

Définition 4.8. — Soit I un intervalle de \mathbb{R} et soit f une fonction définie sur \mathcal{D}_f contenant I . La fonction f est dite *dérivable sur I* si f est dérivable en tout point à l'intérieur de I , dérivable à droite en l'extrémité inférieure (si elle est finie) de I et dérivable à gauche en l'extrémité supérieure (si elle est finie) de I .

On appelle alors *fonction dérivée* de f sur I , la fonction f' définie sur I par $f' : x \mapsto f'(x)$.

Il est facile de vérifier que la dérivée d'une fonction constante sur un intervalle est la fonction nulle.

4.1.3. Dérivées successives. —

Définition 4.9. — Soit I un intervalle de \mathbb{R} et f une fonction dérivable sur I . Si f' est elle-même dérivable sur I , on note f'' sa fonction dérivée et on l'appelle *dérivée seconde* de f sur I .

Ce procédé peut être réitéré si f'' est encore dérivable. Par récurrence on obtient ainsi la *dérivée n -ième* ou *dérivée d'ordre n* de f , notée $f^{(n)}$ et définie par

$$f^{(0)} = f \quad \text{et} \quad \forall n \in \mathbb{N}^*, \quad f^{(n)} = \left[f^{(n-1)} \right]'$$

Notation 4.10. — La fonction $f^{(n)}$ se note aussi $\frac{d^n f}{dx^n}$.

Exemple 4.11. — Soit f la fonction définie sur \mathbb{R} par $f(x) = \exp\{-x^2\}$. En utilisant le théorème sur la composition, on démontre que f est dérivable sur \mathbb{R} avec

$$f'(x) = -2x \exp\{-x^2\}.$$

Une nouvelle fois en utilisant les théorèmes sur les opérations algébriques et la composition, on démontre que f' est dérivable et que

$$f''(x) = (4x^2 - 2) \exp\{-x^2\}.$$

De même on montre que f est trois fois dérivable avec

$$f^{(3)}(x) = (-8x^3 + 12x) \exp\{-x^2\}.$$

Définition 4.12. — Soit f une fonction définie sur I et soit $k \in \mathbb{N}^*$. On dit que

- f est de classe \mathcal{C}^k sur I si f admet des dérivées successives jusqu'à l'ordre k et si la dérivée k -ième est continue sur I ,
- f est de classe \mathcal{C}^∞ sur I si f admet des dérivées à tous les ordres (ces dérivées sont alors continues sur I).

Pour chaque $k \in \mathbb{N} \cup \{\infty\}$, on note $\mathcal{C}^k(I)$ l'ensemble des fonctions de classe \mathcal{C}^k sur I .

Il est très utile de savoir qu'une fonction de classe \mathcal{C}^1 est une fonction dont la courbe est *lisse*, c'est-à-dire sans point anguleux.

4.2. Opérations sur les fonctions dérivables

4.2.1. Opérations algébriques. — Les résultats suivants découlent de la définition de la dérivée et des propriétés des limites.

Théorème 4.13. — Soient I un intervalle de \mathbb{R} , u et v deux fonctions dérivables sur I et $\lambda \in \mathbb{R}$.

– Les fonctions $u + v$, uv et λu sont dérivables sur I avec

$$(u + v)' = u' + v', \quad (uv)' = u'v + uv' \quad \text{et} \quad (\lambda u)' = \lambda u'.$$

– Si pour chaque $x \in I$, $v(x) \neq 0$ alors la fonction $1/v$ est dérivable sur I avec

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

– Si pour chaque $x \in I$, $v(x) \neq 0$ alors la fonction u/v est dérivable sur I avec

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}.$$

Toutes les fonctions usuelles sont dérivables sur leur domaine de définition, sauf certaines fonctions puissances qui sont définies (et même continues) en 0 mais ne sont pas dérivables en 0. Voici un tableau des dérivées usuelles.

Fonction	Domaine de dérivabilité	Dérivée
e^x	\mathbb{R}	e^x
$\sin x$	\mathbb{R}	$\cos x$
$\cos x$	\mathbb{R}	$-\sin x$
$a^x = e^{x \ln a}$, $a > 0$	\mathbb{R}	$(\ln a)a^x$
x^α , $\alpha \in \mathbb{R}$	$\begin{cases} \mathbb{R} & \text{si } \alpha \in \mathbb{N} \\ \mathbb{R}^{+*} & \text{si } \alpha \in \mathbb{R} \setminus \mathbb{Z} \\ \mathbb{R}^* & \text{si } -\alpha \in \mathbb{N}^* \end{cases}$	$\alpha x^{\alpha-1}$
$\ln x$	$]0, +\infty[$	$\frac{1}{x}$
$\tan x$	$x \neq \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$	$1 + \tan^2 x = \frac{1}{\cos^2 x}$

4.2.2. Composition de deux fonctions dérivables.

Théorème 4.14. — Soit u une fonction dérivable en a et soit f une fonction dérivable en $u(a)$. La fonction $f \circ u$ est dérivable en a et l'on a

$$(f \circ u)'(a) = (f' \circ u)(a) \times u'(a) = f'[u(a)] \times u'(a).$$

Si les propriétés suivantes sont satisfaites

- la fonction u est dérivable sur un intervalle I ,
- la fonction f est dérivable sur un intervalle J ,
- $u(I) \subset J$,

alors la fonction $f \circ u$ est dérivable sur I avec

$$(f \circ u)' = (f' \circ u) \times u'.$$

Remarque 4.15. — Les théorèmes précédents sont encore valables en remplaçant *dérivable* par *de classe \mathcal{C}^k* . En pratique ce sont ces théorèmes qui permettront de justifier qu'une fonction construite à l'aide des fonctions usuelles (somme, produit, quotient et composée) est dérivable (ou de classe \mathcal{C}^k) sur un intervalle I . Ainsi toute

fonction polynôme est de classe C^∞ sur \mathbb{R} et toute fraction rationnelle est de classe C^∞ sur son domaine de définition (réunion d'intervalles ouverts disjoints).

4.3. Applications au sens de variations des fonctions

Le signe de la dérivée première d'une fonction fournit des informations sur son sens de variation.

Théorème 4.16. — Soit I un **intervalle** de \mathbb{R} , on note $\overset{\circ}{I}$ l'intervalle I privé de ses extrémités finies éventuelles. L'intervalle $\overset{\circ}{I}$ est un intervalle ouvert. Si f est une fonction continue sur I , dérivable sur $\overset{\circ}{I}$ alors

- (i) f est constante sur I si et seulement si $f'(x) = 0$ pour chaque $x \in \overset{\circ}{I}$,
- (ii) f est croissante sur I si et seulement si $f'(x) \geq 0$ pour chaque $x \in \overset{\circ}{I}$,
- (iii) f est décroissante sur I si et seulement si $f'(x) \leq 0$ pour chaque $x \in \overset{\circ}{I}$.

Le théorème suivant donne des conditions suffisantes pour justifier qu'une fonction est strictement croissante.

Théorème 4.17. — Avec les mêmes hypothèses que le théorème précédent,

- (iv) si $f'(x) > 0$ pour chaque $x \in \overset{\circ}{I}$ alors f est strictement croissante sur I ,
- (v) si $f'(x) < 0$ pour chaque $x \in \overset{\circ}{I}$ alors f est strictement décroissante sur I ,
- (vi) si $f'(x) \geq 0$ pour chaque $x \in \overset{\circ}{I}$ et f' ne s'annule qu'en un nombre fini de points alors f est strictement croissante sur I ,
- (vii) si $f'(x) \leq 0$ pour chaque $x \in \overset{\circ}{I}$ et f' ne s'annule qu'en un nombre fini de points alors f est strictement décroissante sur I .



Les deux théorèmes précédents ne s'appliquent que sur un intervalle.

Exemple 4.18. —

- Soit f la fonction définie sur \mathbb{R} par $f(x) = 2$ si $x \leq 0$ et $f(x) = 0$ si $x > 0$. La fonction f est dérivable sur \mathbb{R}^* avec $f'(x) = 0$ pour tout $x \in \mathbb{R}^*$ et pourtant f n'est pas constante sur \mathbb{R}^* .
- Soit f la fonction définie sur \mathbb{R}^* par $f(x) = -1/x$. La fonction f est dérivable et sa dérivée est strictement positive sur \mathbb{R}^* . Pourtant f n'est pas croissante puisque $f(-1) = 1 > f(1) = -1$. Par contre f est strictement croissante sur $] -\infty, 0[$ et strictement croissante sur $]0, \infty[$.
- La réciproque de l'assertion (iv) est fausse : prenons par exemple f la fonction définie sur \mathbb{R} par $f(x) = x^3$. Cette fonction est strictement croissante sur \mathbb{R} et pourtant $f'(0) = 0$.

4.4. Primitive

Définition 4.19. — Soit f une fonction définie sur un intervalle I . On dit que f admet une *primitive* F sur I si F est une fonction définie et dérivable sur I avec $F'(x) = f(x)$ pour chaque $x \in I$.

Proposition 4.20. — Si la fonction f admet une primitive F sur l'intervalle I , alors

- elle en possède une infinité de la forme $F + c$ où c est une fonction constante sur I ,
- elle en possède une seule qui prenne une valeur donnée en un point a de I .

Théorème 4.21. — [Théorème de Darboux] Toute fonction continue sur un intervalle I admet une primitive sur I .

Exemple 4.22. — La fonction f définie sur $]0, +\infty[$ par $f(x) = 1/x$ possède une primitive sur cet intervalle qui s'annule en $x = 1$. On l'appelle logarithme népérien et on la note \ln .

CHAPITRE 5

PROPRIÉTÉS GRAPHIQUES DES FONCTIONS

5.1. Symétries

Définition 5.1. — Une fonction f est dite

- *paire* si pour chaque $x \in \mathcal{D}_f$, $-x \in \mathcal{D}_f$ et $f(-x) = f(x)$,
- *impaire* si pour chaque $x \in \mathcal{D}_f$, $-x \in \mathcal{D}_f$ et $f(-x) = -f(x)$.

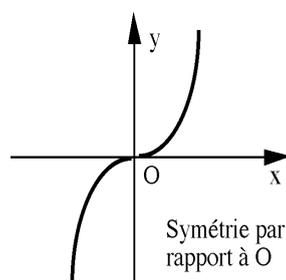
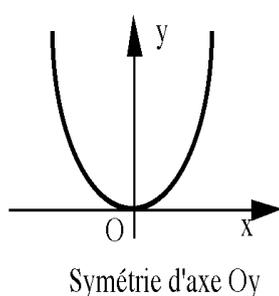


FIGURE 5.1. Exemples d'une fonction paire et d'une fonction impaire

Dans les deux cas, on construit la courbe représentative de f pour $x \in \mathcal{D}_f \cap [0, +\infty[$. Si f est paire, on complète ensuite la représentation graphique de f par une symétrie d'axe Oy . Si f est impaire, on complète ensuite la représentation graphique de f par une symétrie par rapport à l'origine.

5.2. Fonctions convexes et concaves

Nous allons voir dans ce paragraphe comment le signe de la dérivée seconde d'une fonction apporte des indications supplémentaires sur la forme de son graphe. Nous nous limiterons aux fonctions qui sont au moins de classe \mathcal{C}^1 et nous ne donnerons donc pas la définition des fonctions convexes ou concaves dans le cas le plus général.

5.2.1. Définition pour les fonctions de classe \mathcal{C}^1 . —

Définition 5.2. — Soient I un intervalle de \mathbb{R} et f une fonction de classe \mathcal{C}^1 sur I . On note $\mathcal{G}_f(I)$ la portion de la courbe représentative de f correspondant à $x \in I$.

– La fonction f est *convexe sur I* si $\mathcal{G}_f(I)$ est située au-dessus de toutes ses tangentes, *i.e.*

$$\forall x \in I, \quad \forall a \in I, \quad f(x) \geq f(a) + (x - a)f'(a).$$

– La fonction f est *concave sur I* si $\mathcal{G}_f(I)$ est située en-dessous de toutes ses tangentes, *i.e.*

$$\forall x \in I, \quad \forall a \in I, \quad f(x) \leq f(a) + (x - a)f'(a).$$

Remarque 5.3. — 1. La figure suivante illustre les cas d'une fonction convexe et d'une fonction concave.

2. La fonction f est concave sur I si et seulement si $-f$ est convexe sur I .

3. Si dans les définitions précédentes, les inégalités sont strictes pour $x \neq a$, on parle alors de fonctions strictement convexe sur I ou strictement concave sur I .

4. Si \mathcal{D}_f est un intervalle et si f est convexe (resp. concave) sur \mathcal{D}_f , on peut dire pour simplifier que f est convexe (resp. concave). On sous-entend l'expression "sur \mathcal{D}_f ".

5. On peut parler d'une fonction convexe seulement sur un intervalle, nous verrons les raisons dans le chapitre 21.

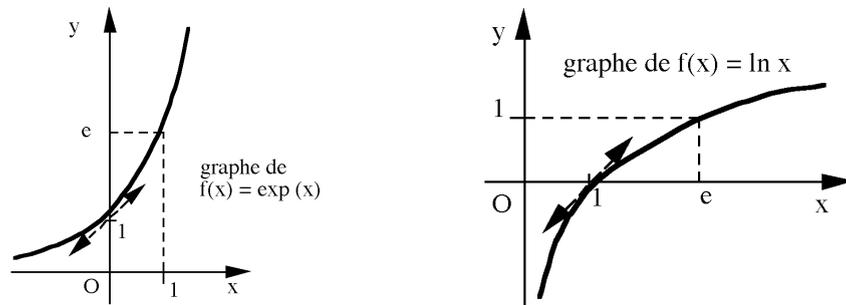


FIGURE 5.2. $x \mapsto e^x$ est convexe et $x \mapsto \ln x$ est concave



Même si \mathcal{D}_f est un intervalle, la fonction f peut être convexe (concave) sur un intervalle $I \subset \mathcal{D}_f$ sans être convexe (resp. concave) sur \mathcal{D}_f .

5.2.2. Caractérisation pour les fonctions de classe \mathcal{C}^2 . — En comparant les pentes des tangentes de la courbe représentative d'une fonction convexe, on constate que la pente de la tangente en un point d'abscisse x est une fonction croissante de x . La fonction dérivée est donc croissante. Si f est de classe \mathcal{C}^2 , alors f'' est positive. On admettra la réciproque.

Théorème 5.4. — Soit I un intervalle de \mathbb{R} et soit f une fonction de classe \mathcal{C}^2 sur I . Les assertions suivantes sont équivalentes.

- (i) f est convexe sur I ,
- (ii) f' est croissante sur I ,
- (iii) $f''(x) \geq 0$ pour chaque $x \in I$.

Si on remplace “convexe” par “concave”, on a le même théorème en remplaçant “croissant” par “décroissant” et en remplaçant “ $f''(x) \geq 0$ ” par “ $f''(x) \leq 0$ ”. Ce théorème donne un moyen très simple pour étudier la convexité ou la concavité d'une fonction de classe \mathcal{C}^2 .

Exemple 5.5. —

1. Les fonctions $x \mapsto x^2$ et $x \mapsto e^x$ sont convexes sur \mathbb{R} .
2. La fonction $x \mapsto \ln x$ est concave sur $]0, +\infty[$.
3. La fonction affine $x \mapsto ax + b$ est concave et convexe sur \mathbb{R} .
4. La fonction $x \mapsto x^3$ est convexe sur $[0, +\infty[$ et concave sur $] -\infty, 0]$.

5.2.3. Application au graphe des fonctions monotones. — Lorsqu'une fonction monotone est aussi convexe ou concave, cela donne les quatre cas suivants qui seront fréquemment utilisés pour modéliser des fonctions économiques selon leurs propriétés.

1. Si une fonction est concave croissante, cela signifie qu'elle croît de plus en plus lentement. C'est-à-dire les valeurs prises par la fonction augmentent beaucoup moins rapidement que les valeurs prises par la variable.
2. Si une fonction est convexe croissante, cela signifie qu'elle croît de plus en plus rapidement. C'est-à-dire les valeurs prises par la fonction augmentent beaucoup plus rapidement que les valeurs prises par la variable.
3. Si une fonction est concave décroissante, cela signifie qu'elle décroît de plus en plus rapidement. C'est-à-dire les valeurs prises par la fonction diminuent beaucoup plus rapidement que les valeurs prises par la variable.
4. Si une fonction est convexe décroissante, cela signifie qu'elle décroît de plus en plus lentement. C'est-à-dire les valeurs prises par la fonction diminuent beaucoup moins rapidement que les valeurs prises par la variable.

5.3. Branches infinies

Les branches infinies d'un graphe correspondent à la forme du graphe pour les très grandes valeurs positives ou les très petites valeurs négatives de x ou de $f(x)$. Lorsque la distance entre la courbe représentative d'une fonction et une droite est de plus en plus faible, on dit que *la droite est asymptote à la courbe*.

1er cas : Si pour un réel a , on a

$$\lim_{x \rightarrow a} f(x) = \pm\infty \quad \text{ou} \quad \lim_{x \rightarrow a^-} f(x) = \pm\infty \quad \text{ou} \quad \lim_{x \rightarrow a^+} f(x) = \pm\infty$$

alors la droite verticale d'équation $x = a$ est asymptote à la courbe.

2ème cas : Si pour un réel b , on a

$$\lim_{x \rightarrow +\infty} f(x) = b \quad \text{ou} \quad \lim_{x \rightarrow -\infty} f(x) = b$$

alors la droite horizontale d'équation $y = b$ est asymptote à la courbe.

3ème cas : Si

$$\lim_{x \rightarrow +\infty} f(x) = \pm\infty \quad \text{ou} \quad \lim_{x \rightarrow -\infty} f(x) = \pm\infty$$

alors on poursuit l'étude en vérifiant s'il existe $a \in \overline{\mathbb{R}}$ tel que

$$\lim_{x \rightarrow \pm\infty} \frac{f(x)}{x} = a.$$

Dans l'affirmative, trois sous-cas sont à envisager :

3.a : si $a = 0$ alors f admet une branche parabolique de direction Ox ,

3.b : si $a \in \{+\infty, -\infty\}$ alors f admet une branche parabolique de direction Oy ,

3.c : si a est fini non nul, on poursuit l'étude en vérifiant s'il existe $b \in \overline{\mathbb{R}}$ tel que

$$\lim_{x \rightarrow \pm\infty} [f(x) - ax] = b.$$

Dans l'affirmative, deux sous-cas sont à envisager :

3.c.i : si $b \in \{+\infty, -\infty\}$ alors f admet une branche parabolique de direction $y = ax$,

3.c.ii : si $b \in \mathbb{R}$ alors la droite $y = ax + b$ est asymptote à la courbe. Dans ce dernier cas, le signe de $f(x) - ax - b$ au voisinage de $\pm\infty$ donne la position de la courbe par rapport à son asymptote.

Exemple 5.6. —

1. La fonction f définie $f(x) = x^2$ donne un exemple d'une branche parabolique de direction Oy .
2. La fonction f définie par $f(x) = \sqrt{x}$ donne un exemple d'une branche parabolique de direction Ox .

5.4. Représentation graphique d'une fonction

Voici les différentes étapes d'une étude de fonction.

1. Si le domaine de définition \mathcal{D}_f n'est pas indiqué, on détermine le plus grand sous-ensemble de \mathbb{R} sur lequel f est définie. Par des considérations telles que la parité, on détermine un sous-ensemble $\mathcal{D}_e \subset \mathcal{D}_f$ appelé domaine d'étude, sur lequel il suffit d'étudier f .
2. On recherche les limites aux bornes de l'ensemble d'étude en utilisant les limites de référence (fonctions usuelles) et les opérations sur les limites.
3. On justifie la continuité et la dérivabilité en utilisant les théorèmes sur somme, produit, quotient et composée de fonctions continues et dérivables.
4. On détermine le sens de variations en étudiant le signe de la dérivée $f'(x)$ pour $x \in \mathcal{D}_e$. Pour cela, il peut être nécessaire de calculer f'' ou d'utiliser une fonction auxiliaire pour avoir les variations de f' et en déduire le signe de f' . Si f' est continue alors sur chaque sous-intervalle où f' ne s'annule pas, le signe de f' est constant. Dans les cas non évidents, il suffit de déterminer le signe de $f'(a)$ pour un point particulier (judicieusement choisi) a du sous-intervalle. On dresse le tableau de variations.
5. On étudie les branches infinies.
6. On trace un repère orthonormé. On place les points remarquables figurant dans le tableau de variations avec leur tangente. On place les droites asymptotes éventuelles et on complète la représentation graphique par des symétries si nécessaire.

CHAPITRE 6

FONCTIONS BIJECTIVES ET RÉCIPROQUES

La notion de fonction réciproque est utile, non seulement pour définir les fonctions exponentielles et les fonctions puissances d'exposant fractionnaire ou négatif, mais aussi parce qu'elle joue un rôle dans certains modèles économiques.

6.1. Définitions

Définition 6.1. — Soient I un intervalle de \mathbb{R} et f une fonction définie sur I . On dit que f est *bijection de I sur $f(I)$* ou est *une bijection de I sur $f(I)$* si, pour chaque $y \in f(I)$, l'équation $y = f(x)$ d'inconnue x admet toujours une unique solution $x \in I$.

Si f est bijective de I sur $f(I)$, on peut définir l'*application réciproque* de f , notée f^{-1} , définie sur $f(I)$ par : pour chaque $y \in f(I)$, $f^{-1}(y)$ est l'unique solution dans I de l'équation $y = f(x)$.

Exemple 6.2. — La figure 6.1 représente les graphes d'une fonction bijective et d'une fonction non bijective.

Remarque 6.3. — 1. Si $(x, y) \in I \times f(I)$ alors

$$y = f(x) \iff x = f^{-1}(y).$$

2. Si on note Id_I (resp. $\text{Id}_{f(I)}$) la fonction définie sur I (resp. sur $f(I)$) par $x \mapsto x$ alors

$$f^{-1} \circ f = \text{Id}_I \quad \text{et} \quad f \circ f^{-1} = \text{Id}_{f(I)}.$$

6.2. Caractérisation des fonctions bijectives continues

La figure 6.1 montre que pour qu'une fonction f soit bijective de I sur $f(I)$, elle doit nécessairement être au moins strictement monotone sur I . Nous admettrons que cette condition est aussi suffisante. Mais $f(I)$ n'est pas forcément un intervalle et c'est pourquoi nous nous limiterons aux fonctions strictement monotones et continues.

Théorème 6.4. — Si f une fonction définie sur un intervalle I , continue et strictement monotone sur I , alors f est une application bijective de I sur $f(I)$. En outre,

- l'application réciproque f^{-1} est continue et monotone sur $f(I)$ avec le même sens de monotonie que f ,
- dans un repère orthonormé, les courbes représentatives de f et f^{-1} sont symétriques par rapport à la première bissectrice.

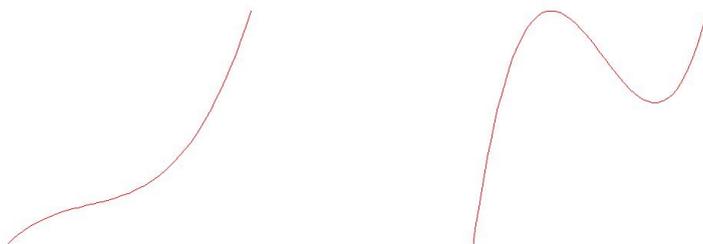


FIGURE 6.1. Une fonction bijective et une fonction non bijective

6.3. Calcul de la dérivée d'une fonction réciproque

Proposition 6.5. — Soit f une fonction continue, strictement monotone de l'intervalle I sur $f(I)$ et dérivable sur I . Soit $x \in I$ tel que $f'(x) \neq 0$ alors f^{-1} est dérivable en $y = f(x)$ et on a

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

Preuve. — On a vu que $f \circ f^{-1} = \text{Id}_{f(I)}$, i.e.

$$\forall z \in f(I), \quad f[f^{-1}(z)] = z.$$

Si on admet la dérivabilité de f^{-1} , on peut appliquer le théorème de dérivation d'une composée et on obtient

$$\forall z \in f(I), \quad f'[f^{-1}(z)] \times (f^{-1})'(z) = 1.$$

□

CHAPITRE 7

DIFFÉRENTIELLE ET APPROXIMATION AFFINE

7.1. Variation absolue et différentielle

Définition 7.1. — Soit f une fonction définie sur un voisinage d'un réel a . On appelle *accroissement* ou *variation absolue de la fonction f lorsque la variable passe de a à $a + h$* la valeur de la différence

$$\Delta f_a(h) = f(a + h) - f(a).$$

L'accroissement est défini pour tout réel h tel que $f(a + h)$ existe.

Exemple 7.2. — Soit f définie sur \mathbb{R} par $f(x) = x^2$. Pour $a = 2$, on a $\Delta f_2(h) = (2 + h)^2 - 2^2 = 4h + h^2$. Pour $a = -1$, on a $\Delta f_{-1}(h) = (-1 + h)^2 - 1 = -2h + h^2$.

Par analogie de notation, la variation de la variable est parfois notée Δx au lieu de h .

Définition 7.3. — Soit f une fonction définie, continue et dérivable sur un voisinage de $a \in \mathbb{R}$. On appelle *différentielle de f au point a* , l'application **linéaire** df_a définie sur \mathbb{R} par

$$df_a : h \mapsto df_a(h) = f'(a)h.$$

Exemple 7.4. — Soit f définie sur \mathbb{R} par $f(x) = x^2$. Pour $a = 2$ on a $df_2 : h \mapsto 4h$ et pour $a = -1$ on a $df_{-1} : h \mapsto -2h$.

7.2. Approximation affine

Si f est affine, par exemple $f(x) = \alpha x + \beta$ avec $\alpha, \beta \in \mathbb{R}$, alors pour tout $a \in \mathbb{R}$

$$\Delta f_a(h) = \alpha h = df_a(h), \quad \forall h \in \mathbb{R}.$$

Mais cette égalité n'est plus vraie pour une fonction quelconque.

Exemple 7.5. — Soit f définie sur \mathbb{R} par $f(x) = x^2$ et soit a un réel quelconque. Comparons $\Delta f_a(h)$ et $df_a(h)$:

$$\Delta f_a(h) = 2ah + h^2, \quad df_a(h) = 2ah \quad \text{donc} \quad \Delta f_a(h) - df_a(h) = h^2.$$

Si on choisit $a = 10$ et $h = 0,1$ on obtient

$$\Delta f_{10}(0,1) = 2,01 \quad \text{et} \quad df_{10}(0,1) = 2.$$

On constate que $df_{10}(0,1)$ est une valeur approchée de $\Delta f_{10}(0,1)$.

7.2.1. Développement limité d'ordre 1. —

Proposition 7.6. — Soit f une fonction définie et continue sur un voisinage de $a \in \mathbb{R}$. La fonction f est dérivable si et seulement si il existe une fonction ε définie sur un voisinage V de 0 et continue en 0 telle que $\varepsilon(0) = 0$ et pour tout réel $x \in \mathbb{R}$

$$x - a \in V \quad \implies \quad f(x) = f(a) + (x - a)f'(a) + (x - a)\varepsilon(x - a).$$

Comme la fonction ε est continue en 0, on a

$$\lim_{x \rightarrow a} \varepsilon(x - a) = \varepsilon(0) = 0.$$

Preuve. — Supposons que f est définie sur un voisinage U de a et dérivable en a . On pose $V = \{h \in \mathbb{R} : a+h \in U\}$. Comme U est un voisinage de a , l'ensemble V est un voisinage de 0. On considère la fonction ε définie sur V par

$$\forall h \in V, \quad \varepsilon(h) = \begin{cases} \frac{f(a+h)-f(a)}{h} - f'(a) & \text{si } h \neq 0 \\ 0 & \text{si } h = 0. \end{cases}$$

Par définition de $f'(a)$ on a $\lim_{h \rightarrow 0} \varepsilon(h) = 0$ et d'après la définition de la fonction ε , on a

$$\forall x \in U, \quad f(x) = f(a) + (x-a)f'(a) + (x-a)\varepsilon(x-a).$$

La réciproque est évidente. □

Remarque 7.7. — En posant $x = a + h$, on peut reformuler la proposition 7.6 de la façon suivante :

$$f(a+h) = f(a) + f'(a)h + h\varepsilon(h)$$

avec $\lim_{h \rightarrow 0} \varepsilon(h) = \varepsilon(0) = 0$.

7.2.2. Approximation affine. —

Définition 7.8. — Soit f une fonction définie sur un voisinage de a telle que $f'(a)$ existe. On appelle *approximation affine de f au voisinage de a* , notée \widehat{f}_a , la fonction affine définie sur \mathbb{R} par

$$\forall x \in \mathbb{R}, \quad \widehat{f}_a(x) = f(a) + (x-a)f'(a).$$

La fonction \widehat{f}_a est de la forme $x \mapsto \alpha x + \beta$ avec $\alpha = f'(a)$ et $\beta = f(a) - af'(a)$. Réaliser une approximation affine de f au voisinage de a revient à remplacer la courbe par sa tangente au point d'abscisse $x = a$. On dit aussi qu'on linéarise la fonction f .

Remarque 7.9. —

1. L'approximation affine dépend du point a choisi.
2. Il n'y a pas de notation standard. Nous proposons \widehat{f}_a qui est simple à utiliser.
3. Dans l'expression de $\widehat{f}_a(x)$ il est préférable de laisser le terme en $(x-a)$ sans le développer.
4. On peut poser $h = x - a$ pour obtenir $\widehat{f}_a(a+h) = f(a) + f'(a)h$.

Exemple 7.10. — (i) L'approximation affine de la fonction \exp au voisinage de 0 est

$$\widehat{\exp}_0(x) = 1 + x.$$

(ii) L'approximation affine de la fonction \ln au voisinage de 1 est $\widehat{\ln}_1(x) = x - 1$.

7.2.3. Calculs approchés. —

Théorème 7.11. — Soit f une fonction définie au voisinage de $a \in \mathbb{R}$, dérivable en a . Pour h suffisamment petit, $\Delta f_a(h) \simeq df_a(h)$ c'est-à-dire $f(a+h) - f(a) \simeq f'(a)h$ ou encore $f(a+h) \simeq \widehat{f}_a(a+h)$.

Preuve. — D'après la proposition précédente, $f(a+h) - f(a) = f'(a)h + h\varepsilon(h)$. Le terme $h\varepsilon(h)$ est négligeable devant $f'(a)h$, d'où la valeur approchée. □

Remarque 7.12. — Le signe \simeq signifie que l'un des termes est une approximation de l'autre. Par le théorème précédent, $\Delta f_a(h)$ est une approximation de $df_a(h)$. Cette notation sera utilisée dans de nombreux chapitres.

Exemple 7.13. — Calcul approché de $\exp(0.002)$: la fonction f définie sur \mathbb{R} par $f(x) = \exp(x)$ est dérivable sur \mathbb{R} . En posant $a = 0$ et $h = 0.002$ on a

$$\Delta f_0(0.002) = e^{0.002} - e^0 \simeq e^0 \times (0.002) \quad \text{d'où} \quad e^{0.002} \simeq 1.002.$$

La valeur obtenue par une calculatrice est $e^{0.002} = 1.002002001$.

7.3. Notation différentielle

7.3.1. Version des mathématiciens. — Considérons l'application identité J de \mathbb{R} dans \mathbb{R} définie par $J : x \mapsto x$. Soit x un réel fixé, pour tout réel h

$$dJ_x(h) = J'(x)h = 1h = h.$$

On constate que le résultat obtenu ne dépend pas de x , ce qui va permettre de simplifier la notation.

Comme $J(x) = x$, on va faire l'abus de langage, excessif mais commode, qui consiste à confondre une fonction f avec sa valeur $f(x)$ en un point, et prendre comme notation dx pour représenter la différentielle dJ_x . Ainsi pour tout réel h , $dx(h) = h$.

Maintenant, la différentielle d'une fonction f en un point $x \in \mathbb{R}$ s'écrit

$$\forall h \in \mathbb{R}, \quad df_x(h) = f'(x)h = f'(x)dx(h).$$

Cette relation est valable pour tout réel h , on en déduit l'égalité entre fonctions

$$df_x = f'(x)dx.$$

Le point x à partir duquel s'effectue la variation étant indiqué par $f'(x)$, on peut l'omettre dans le terme de gauche, et on écrit seulement

$$df = f'(x)dx.$$

Exemple 7.14. — Soit f la fonction définie sur $]0, +\infty[$ par $f(x) = e^{-x} \ln x$. La différentielle est $df = e^{-x}(1/x - \ln x)dx$.

7.3.2. Utilisation faite par les économistes. — Soit f une fonction, la variation de f pour un accroissement Δx de x est égale à $\Delta f_x(\Delta x) = f(x + \Delta x) - f(x)$. Les économistes écrivent simplement Δf au lieu de $\Delta f_x(\Delta x)$. D'après le théorème 7.11 on déduit l'approximation

$$\Delta f \simeq f'(x)\Delta x.$$

Cette approximation sera d'autant meilleure que l'accroissement Δx sera plus petit. Lorsque Δx est suffisamment petit pour que l'erreur soit négligeable, les économistes écrivent dx au lieu de Δx et df au lieu de Δf . Ils écrivent alors $df \simeq f'(x)dx$. Autrement dit, les économistes utilisent la formule $df \simeq f'(x)dx$ en considérant que df et dx représentent des petits accroissements, c'est-à-dire des nombres et non des fonctions.

Notation 7.15. — Lorsqu'on pose $y = f(x)$, on écrit Δy à la place de Δf et dy au lieu de df .

D'après ce qui précède, on est tenté d'écrire

$$f'(x) = \frac{df}{dx}.$$

Il faut toutefois remarquer que cette notation $\frac{df}{dx}$ appelée *notation différentielle de la dérivée* rappelle bien la définition de la dérivée comme limite du taux d'accroissement, mais quoique commode, elle est imprécise puisque le point où l'on prend la dérivée reste sous-entendu. Cette notation se révélera néanmoins pratique pour le calcul des dérivées de fonctions composées, encore appelée la dérivation en chaîne.

7.3.3. Opérations sur les différentielles. — La différentielle d'une fonction en un point étant caractérisée par la dérivée de la fonction en ce point, les règles de calcul sur les différentielles se déduisent naturellement des règles de calcul sur les dérivées.

Proposition 7.16 (Règles de calcul sur les différentielles). —

$$d(\lambda f) = \lambda df, \quad d(f + g) = df + dg, \quad d(fg) = gdf + f dg$$

et

$$d(f/g) = \frac{gdf - f dg}{g^2} \quad (g \neq 0).$$

Considérons une fonction $f : u \mapsto f(u)$, alors

$$(7.1) \quad df = f'(u)du.$$

Supposons que u ne soit plus une variable mais une fonction $u : x \mapsto u(x)$ d'une autre variable x , alors

$$(7.2) \quad du = u'(x)dx.$$

Posons $F(x) = f[u(x)]$, d'après le théorème de la dérivée de la composée d'une fonction, on a

$$(7.3) \quad df = F'(x)dx = (f \circ u)'(x)dx = f'(u(x))u'(x)dx.$$

En utilisant les relations (7.1), (7.2) et (7.3), on obtient

$$dF = f'(u)du = df.$$

En pratique on peut utiliser ce résultat pour calculer la dérivée d'une fonction composée en utilisant la différentielle. Si $F(x) = f \circ u(x)$ en utilisant la notation différentielle pour les dérivées, on obtient la règle de dérivation en chaîne

$$\frac{dF}{dx} = \frac{df}{du} \times \frac{du}{dx}.$$

Exemple 7.17. — Soit F la fonction définie sur $]0, +\infty[$ par $F(x) = \exp\{-\ln x\}$. Soit f la fonction définie sur \mathbb{R} par $f(u) = \exp\{u\}$, alors $F(x) = f(u)$ avec $u = \ln x$. Donc $dF = df$ c'est-à-dire

$$F'(x)dx = f'(u)du = f'(u)u'(x)dx = -e^u \frac{1}{x} dx = -e^{-\ln x} \frac{1}{x} dx$$

d'où par identification

$$F'(x) = -\frac{1}{x} \exp\{-\ln x\}.$$

Remarque 7.18. — La notation différentielle est très pratique pour calculer la dérivée d'une fonction réciproque si on se permet quelques abus de notations. Si $y = f(x)$, on écrira $f'(x) = dy/dx$ au lieu de df/dx . Si $x = f^{-1}(y)$, on écrira $(f^{-1})'(y) = dx/dy$ au lieu de df^{-1}/dy . Pour mémoriser la formule donnant $(f^{-1})'(y)$ il suffira d'écrire

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}.$$

CHAPITRE 8

CALCULS APPROCHÉS DES VARIATIONS

Un des problèmes fondamentaux en Économie est la mesure des effets de la variation d'une variable économique sur une autre. Nous allons voir dans ce chapitre comment la notion de dérivée permet de définir les outils nécessaires pour effectuer ces mesures. Il s'agit en fait d'interpréter les résultats du chapitre précédent.

8.1. Fonction moyenne

Définition 8.1. — Soit f une fonction économique positive (production, coût, utilité, demande) dépendant de la variable $x \in]0, +\infty[$ (quantité, prix). La *fonction moyenne*, notée f_M , est définie par

$$\forall x \in]0, +\infty[, \quad f_M(x) = \frac{f(x)}{x}.$$

Remarque 8.2. — Si f est une fonction de coût, alors le coût moyen représente le coût de production d'une unité lorsqu'on a produit x unités.

Exemple 8.3. — Pour un chef d'entreprise le coût total de fabrication d'un produit n'est pas proportionnel au nombre d'unités fabriquées car certains coûts sont fixes (loyer des ateliers, achat de machines). Supposons qu'il fabrique 100000 unités pour un coût total de un million d'euros et que s'il en fabrique 120000 cela lui revienne à 1 140 000 euros. Le coût moyen pour 100000 unités est de 10 euros et le coût moyen pour 120000 unités est de 9,5 euros.

Cet exemple montre l'intérêt de calculer ce qui se passe "à la marge", c'est-à-dire le supplément de coût pour la production de la 100001^{ème} unité ou de la 120001^{ème}. Ceci fait l'objet du paragraphe suivant.

8.2. Fonction marginale et variation absolue

8.2.1. Fonction marginale. —

Définition 8.4. — Soit f une fonction économique définie, positive et dérivable sur $]0, +\infty[$. La *fonction marginale*, notée f_m , est définie sur $]0, +\infty[$ par

$$\forall x \in]0, +\infty[, \quad f_m(x) = f'(x).$$

Les fonctions marginales sont des exemples de fonctions économiques pouvant prendre des valeurs négatives. En appliquant le développement limité d'ordre 1 de f en a pour $h = 1$, on a

$$\Delta f_a(1) = f(a+1) - f(a) \simeq f'(a) = f_m(a).$$

Autrement dit, $f_m(a)$ est une valeur approchée de l'accroissement de f lorsque la variable x augmente d'une unité à partir de a . Elle mesure la sensibilité de la fonction f aux variations de x . La fonction marginale est ainsi appelée car elle traduit les variations "à la marge".

On pourrait contester le calcul approché que nous venons de faire, sous le prétexte qu'une unité supplémentaire n'est pas une variation "très petite". En fait, on suppose que la valeur a est suffisamment grande pour admettre qu'une unité supplémentaire est une petite variation.

Exemple 8.5. — La fonction de production d'une entreprise utilisant le travail comme seul facteur est donnée par $f(x) = (1/2)\sqrt{x}$, pour $x \geq 0$. Supposons que la firme utilise $a = 900$ unités de travail. Sa production est donc de 15 unités. Quelle sera l'augmentation de la production si la firme utilise une unité supplémentaire de travail? En toute rigueur, on devrait prendre $f(901) - f(900) = 0,0083310198\dots$. Mais on peut donner une estimation

de la production additionnelle obtenue avec une unité supplémentaire de travail en raisonnant comme suit. La fonction marginale est

$$f_m(x) = f'(x) = \frac{1}{4\sqrt{x}} \quad \text{donc} \quad f_m(900) = \frac{1}{120} \simeq 0,00833333$$

est une bonne mesure de la production additionnelle cherchée et il est plus facile de travailler avec le seul terme $f_m(900)$ qu'avec la différence $f(901) - f(900)$.

8.2.2. Variation absolue. —

Proposition 8.6. — Soit f une fonction économique (production, coût, utilité, demande) définie sur $]0, +\infty[$ à valeurs positives, dérivable et dépendant de la variable x (quantité, prix). Si la variable x varie de a à $a + \Delta x$, alors la variation absolue de f est égale au produit de la fonction marginale (calculée au point considéré a) par l'accroissement de la variable, i.e.

$$\Delta f_a(\Delta x) = f(a + \Delta x) - f(a) \simeq f_m(a)\Delta x.$$

Exemple 8.7. — Supposons que la firme de l'exemple 8.5 diminue sa force de travail en passant de $a = 900$ unités à 891 unités. On cherche à donner une estimation de la variation de la production qui en résulte. Les données sont :

$$f(x) = \frac{1}{2}\sqrt{x}, \quad f'(x) = f_m(x) = \frac{1}{4\sqrt{x}}, \quad f_m(900) = \frac{1}{120} \quad \text{et} \quad \Delta x = -9.$$

On en déduit que

$$\Delta f_{900}(-9) = f(891) - f(900) \simeq f_m(900) \times (-9) \simeq 0,075,$$

la production diminue d'environ 0,075 unités. On peut dire que la nouvelle production sera approximativement égale à

$$f(891) \simeq f(900) + f_m(900) \times (-9) = 14,925.$$

La méthode précédente permet de faire des estimations rapides avec des calculs assez simples.

La fonction marginale et la variation absolue dépendent des unités choisies. Pour éviter cet inconvénient, on introduit la notion de variation relative.

8.3. Variation relative et élasticité

La variation relative est le quotient de la variation absolue avec la valeur initiale. Dans la situation évoquée dans l'exemple précédent, la variation relative de la production quand x passe de 900 à 891 sera

$$\frac{f(891) - f(900)}{f(900)}.$$

Comme le numérateur et le dénominateur sont mesurés avec la même unité, les unités apparaissent dans le processus de division et le résultat s'exprime en pourcentage.

Plus généralement, soit f une fonction économique définie sur $]0, +\infty[$, positive et dérivable, dépendant de la variable x . Lorsque x varie de a à $a + \Delta x$ et si $f(a) \neq 0$, la *variation relative* de f est

$$\frac{f(a + \Delta x) - f(a)}{f(a)}.$$

On reconnaît au numérateur la variation absolue de f . En utilisant une valeur approchée de cette variation absolue, on obtient :

$$\frac{f(a + \Delta x) - f(a)}{f(a)} \simeq \frac{f'(a)}{f(a)} \Delta x.$$

Si de plus $a \neq 0$, on peut réécrire cette relation comme suit :

$$\frac{f(a + \Delta x) - f(a)}{f(a)} \simeq f'(a) \frac{a}{f(a)} \frac{\Delta x}{a}.$$

On met ainsi en évidence dans l'expression obtenue un facteur remarquable.

Définition 8.8. — Soit f une fonction économique définie sur $]0, +\infty[$, strictement positive et dérivable, dépendant de la variable x . On appelle *élasticité de f* la fonction définie par :

$$\forall x \in]0, +\infty[, \quad e_f(x) = f'(x) \frac{x}{f(x)} = x \times (\ln \circ f)'(x).$$

L'élasticité permet de calculer une valeur approchée de la variation relative de f . Lorsque x passe de a à $a + \Delta x$, on a

$$\underbrace{\frac{f(a + \Delta x) - f(a)}{f(a)}}_{\text{Variation relative de } f} \simeq e_f(a) \times \underbrace{\frac{\Delta x}{a}}_{\text{Variation relative de } x}.$$

On peut dire que la variation relative est pratiquement égale au produit de l'élasticité par la variation relative de la variable. Cela donne le même genre de calcul pour les variations relatives et les variations absolues : l'élasticité remplace la fonction marginale, la variation relative remplace la variation absolue.

Remarque 8.9. —

1. La valeur de l'élasticité dépend du point considéré mais ne dépend pas des unités choisies.
2. Le nombre $e_f(a)$ donne une valeur approchée de la variation relative de f lorsque x augmente de 1% à partir de $x = a$.
- 3.

$$e_f(a) = \lim_{\Delta x \rightarrow 0} \frac{f(a + \Delta x) - f(a)}{\Delta x} \times \frac{a}{f(a)} = \lim_{\Delta x \rightarrow 0} \frac{\frac{f(a + \Delta x) - f(a)}{\Delta x}}{\frac{f(a)}{a}}.$$

4.

$$\forall x \in]0, +\infty[, \quad e_f(x) = \frac{f_m(x)}{f_M(x)}.$$

5. Si l'élasticité est proche de 0 cela signifie que la fonction f est plutôt insensible aux variations de x .
6. Dans certains cas, lorsque l'expression de la fonction f s'y prête, le calcul de l'élasticité est plus simple si on passe par la dérivée de $\ln f$.

Exemple 8.10. —

1. Reprenons l'exemple 8.5 et donnons une valeur approchée de la variation relative de la production. Pour tout $x > 0$, $\ln[f(x)] = -\ln 2 + (1/2) \ln x$ donc $(\ln \circ f)'(x) = 1/(2x)$. Alors pour $a = 900$,

$$e_f(900) = 900 \times \frac{1}{2 \times 900} = 0,5 \quad \text{et} \quad \frac{f(891) - f(900)}{f(900)} \simeq 0,5 \times \frac{-9}{900}.$$

La production diminue donc d'environ 0,005%.

2. Dire que l'élasticité de la demande par rapport au prix est égale à -2 lorsque le prix unitaire est fixé à 10 euros, signifie que si le prix augmente de 1%, la demande diminuera de 2%.

8.4. Récapitulatif

Résumons dans les deux tableaux suivants les notions définies dans ce chapitre et indispensables en Économie. On considère une fonction économique f définie sur $]0, +\infty[$, strictement positive et dérivable en a .

	valeur exacte	valeur approchée
variation absolue	$\Delta_a f(\Delta x) = f(a + \Delta x) - f(a)$	$f'(a)\Delta x$
variation relative	$\frac{f(a + \Delta x) - f(a)}{f(a)}$	$e_f(a) \times \frac{\Delta x}{a}$

	taux moyen	taux instantané
taux d'accroissement absolu	$\frac{f(a + \Delta x) - f(a)}{\Delta x}$	$f'(a)$
taux d'accroissement relatif	$\frac{f(a + \Delta x) - f(a)}{\Delta x} \times \frac{1}{f(a)}$	$\frac{f'(a)}{f(a)} = (\ln \circ f)'(a)$
taux de variation relative	$\frac{f(a + \Delta x) - f(a)}{f(a)} / \frac{\Delta x}{a}$	$e_f(a) = f'(a) \times \frac{a}{f(a)}$

CHAPITRE 9

FORMULE DE TAYLOR

Dans tout ce chapitre n est un entier naturel non nul.

9.1. Polynôme de Taylor

9.1.1. Problème posé. — Les fonctions polynômes sont parmi les plus simples à étudier et à calculer en un point x donné, car tous les calculs peuvent se faire à la main. D'où l'idée d'approcher une fonction f quelconque définie sur un intervalle I contenant a , par un polynôme, et ceci pour en simplifier l'étude au voisinage du point a .

Nous avons déjà vu une première illustration de cette idée. En effet si f est une fonction de classe \mathcal{C}^1 sur un intervalle ouvert I contenant a , alors f est dérivable en a et on peut écrire le développement limité d'ordre 1 : il existe une fonction ε définie sur \mathbb{R} , continue en 0 avec $\varepsilon(0) = 0$ et telle que

$$\forall x \in I, \quad f(x) = f(a) + (x - a)f'(a) + (x - a)\varepsilon(x - a).$$

L'expression $f(a) + (x - a)f'(a)$ est un polynôme de degré 1 en $(x - a)$ qui donne une approximation de f , dite approximation affine.

Pour améliorer l'approximation de f , on souhaite approcher f au voisinage de a par un polynôme de degré n de la forme

$$P(x) = \lambda_0 + \lambda_1(x - a) + \lambda_2(x - a)^2 + \dots + \lambda_n(x - a)^n.$$

9.1.2. Étude d'un cas particulier. — Les coefficients du polynôme recherché vont dépendre de la fonction f considérée et du point a choisi. Pour les déterminer on va considérer le cas particulier où f est déjà un polynôme de degré n . Supposons que

$$f(x) = \lambda_0 + \lambda_1(x - a) + \lambda_2(x - a)^2 + \dots + \lambda_n(x - a)^n.$$

– Calcul de λ_0 : on constate que $\lambda_0 = f(a)$.

– Calcul de λ_1 : dérivons f , on obtient

$$f'(x) = \lambda_1 + 2\lambda_2(x - a) + 3\lambda_3(x - a)^2 + \dots + n\lambda_n(x - a)^{n-1} \quad \text{et} \quad \lambda_1 = f'(a).$$

– Calcul de λ_2 : dérivons f' , on obtient

$$f''(x) = 2\lambda_2 + 3 \times 2\lambda_3(x - a) + \dots + n \times (n - 1)\lambda_n(x - a)^{n-2} \quad \text{et} \quad \lambda_2 = \frac{1}{2}f''(a).$$

– Par dérivations successives, on vérifie que

$$\forall k \in \{0, 1, \dots, n\}, \quad \lambda_k = \frac{1}{k!}f^{(k)}(a).$$

On peut donc écrire

$$f(x) = f(a) + (x - a)f'(a) + (x - a)^2 \frac{f''(a)}{2!} + \dots + (x - a)^n \frac{f^{(n)}(a)}{n!}.$$

9.1.3. Cas général. —

Définition 9.1. — Soit f une fonction définie sur un intervalle I contenant a . On suppose que f est n fois dérivable en a . On appelle *polynôme de Taylor de f à l'ordre n au point a* le polynôme $P_n(f, a)$ défini par

$$\forall h \in \mathbb{R}, \quad P_n(f, a)(h) = f(a) + hf'(a) + h^2 \frac{f''(a)}{2!} + \dots + h^n \frac{f^{(n)}(a)}{n!}.$$

Ce polynôme dépend de f , de a et de n .

Exemple 9.2. — Soit f la fonction définie sur \mathbb{R} par $f(x) = \exp\{x\}$ et $a = 0$, alors

$$P_2(f, 0)(h) = 1 + h + \frac{h^2}{2} \quad \text{et} \quad P_n(f, 0)(h) = \sum_{k=0}^n \frac{h^k}{k!}.$$

Si f n'est pas un polynôme de degré n , la différence $f(a+h) - P_n(f, a)(h)$ n'est pas nulle. Si on fait tracer par une calculatrice les courbes représentatives de $h \mapsto f(a+h)$, de $h \mapsto P_2(f, a)$ et de $h \mapsto P_3(f, a)$, on voit que ces courbes ne coïncident pas, mais qu'elles sont presque confondues au voisinage de a . De plus on constate que la courbe représentative du polynôme de Taylor à l'ordre 3 "colle" mieux à la courbe représentative de f que celle du polynôme d'ordre 2.

9.2. Formule de Taylor-Young à l'ordre n

La formule de Taylor-Young, que nous admettrons, précise dans quelles conditions on peut approcher une fonction f par son polynôme de Taylor à l'ordre n en un point a . Cette formule permet d'évaluer la différence $f(a+h) - P_n(f, a)(h)$ qu'on appellera le reste.

Théorème 9.3. — Soit f une fonction de classe C^n sur un intervalle **ouvert** I contenant a . Il existe une fonction ε définie et continue sur un voisinage de 0 avec $\varepsilon(0) = 0$ et telle que pour tout réel h ,

$$a+h \in I \quad \implies \quad f(a+h) = f(a) + hf'(a) + h^2 \frac{f''(a)}{2!} + \dots + h^n \frac{f^{(n)}(a)}{n!} + h^n \varepsilon(h).$$

Remarque 9.4. —

1. Cette formule s'appelle aussi le *développement limité d'ordre n de f en a* .
2. Le terme $h^n \varepsilon(h)$ est appelé le *reste de Young à l'ordre n* . En particulier

$$\forall h \in \mathbb{R}, \quad a+h \in I \implies f(a+h) = P_n(f, a)(h) + h^n \varepsilon(h),$$

c'est-à-dire la fonction $h \mapsto f(a+h)$ s'écrit comme un polynôme en h et un terme complémentaire (correctif).

3. Comme la fonction ε est continue en 0, on a $\lim_{h \rightarrow 0} \varepsilon(h) = 0$. Ainsi quand h tend vers 0, le reste $f(a+h) - P_n(f, a)(h)$ tend vers 0 "plus vite" que h^n .
4. C'est une formule locale : on ne sait rien de la fonction $x \mapsto \varepsilon(x)$, sinon qu'elle tend vers 0 lorsque x tend vers 0. On ne peut exploiter cette formule que pour x "voisin" de a .
5. En posant $h = x - a$ on peut aussi écrire la formule sous la forme

$$\forall x \in I, \quad f(x) = f(a) + (x-a)f'(a) + (x-a)^2 \frac{f''(a)}{2!} + \dots + (x-a)^n \frac{f^{(n)}(a)}{n!} + (x-a)^n \varepsilon(x-a).$$

Dans cette expression, on ne développe pas les puissances de $(x-a)$ en puissances de x car c'est $x-a$ la quantité pertinente.

6. Pour $n = 0$ la formule donne $f(a+h) = f(a) + \varepsilon(h)$ ce qui traduit la continuité de f en a .

CHAPITRE 10

EXTREMA DES FONCTIONS D'UNE VARIABLE

Dans le cas de la dimension 1, la recherche des extrema peut se résoudre à l'aide d'une *étude de fonction* et d'un *tableau de variations*. Il faut juste faire un peu attention aux bords de l'intervalle. Par contre il est important de bien comprendre les notions et les définitions étudiées dans ce chapitre pour aborder le cas plus compliqué de la dimension 2 traité dans le chapitre 23.

10.1. Définitions

Définition 10.1 (Extremum global). — Soit f une fonction définie sur \mathcal{D}_f et soit E un sous-ensemble de \mathcal{D}_f . On dit que f présente un *extremum global* sur E en $a \in E$ si pour tout $x \in E$, le terme $f(x) - f(a)$ garde un signe constant. Si ce signe est négatif alors $f(a)$ est un *maximum global* de f sur E si ce signe est positif alors $f(a)$ est un *minimum global* de f sur E .

Exemple 10.2. — Soit f la fonction définie sur \mathbb{R} par $f(x) = x^2$, $f(0)$ est un minimum global de f sur \mathbb{R} .

Définition 10.3 (Extremum local). — Soit f une fonction définie sur \mathcal{D}_f et soit E un sous-ensemble de \mathcal{D}_f . On dit que f présente un *extremum local* sur E en $a \in E$ si il existe $r > 0$ tel que pour tout $x \in]a - r, a + r[\cap E$, le terme $f(x) - f(a)$ garde un signe constant. Si ce signe est négatif alors $f(a)$ est un *maximum local* de f sur E si ce signe est positif alors $f(a)$ est un *minimum local* de f sur E .

- Remarque 10.4.** —
1. Il faut distinguer la valeur de l'extremum $f(a)$ du point a où l'extremum est atteint. Autrement dit, a n'est pas un extremum, mais a donne un extremum pour f . C'est $f(a)$ qui est un extremum.
 2. Lorsque dans les définitions précédentes les inégalités sont strictes pour $x \neq a$, on parle d'extremum strict, ce qui signifie que l'extremum est atteint en un seul point.
 3. Tout extremum global sur E est un extremum local sur E . La réciproque est fausse.

Exemple 10.5. — Toutes les définitions sont illustrées par la figure 10.1 :

– Sur $] - 1, 1, 5[$

- f présente en $x = -0.5$ un minimum global de valeur 0,
- f présente en $x = 0.3$ un maximum local de valeur 2.
- f présente en $x = 1$ un minimum local de valeur 1.

Notons que sur l'intervalle $] - 1, 1, 5[$, f présente un minimum global mais pas de maximum global. L'intervalle $] - 1, 1, 5[$ n'est pas compact, le théorème 3.16 page 15 n'est pas mis en défaut.

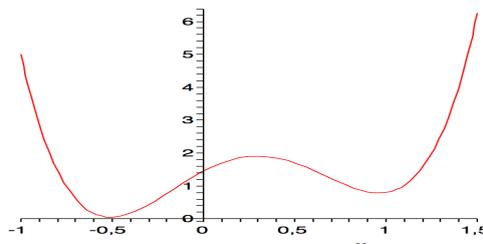


FIGURE 10.1. Un exemple en dimension 1

- Sur $[-1, 1.5]$ Les quatre résultats précédents restent valables, mais on a en plus
 - f admet un maximum local (mais non global) en $x = -1$ de valeur 5,
 - f admet un maximum global en $x = 1.5$ de valeur 6.

Sur l'intervalle $[-1, 1.5]$, f admet bien un maximum global et un minimum global ce qui illustre bien le théorème 3.16 page 15 (f étant bien continue sur un ensemble compact). Nous pouvons remarquer que, sur cet exemple, les extrema sont atteints en des points où la dérivée ne s'annule pas. On pourra comparer cette observation avec le théorème 10.6 ci-dessous.

10.2. Fonction dérivable sur un intervalle ouvert

On suppose que I est un intervalle **ouvert** non vide.

10.2.1. Condition nécessaire ou condition du 1er ordre. —

Théorème 10.6. — Soit f une fonction définie sur \mathcal{D}_f contenant un intervalle ouvert I tel que f est dérivable sur I . Soit $a \in I$, si f présente un extremum local ou global sur I en a alors $f'(a) = 0$.

Preuve. — Tout extremum global de f sur I étant d'abord un extremum local sur I , montrons ce théorème dans le cas où f admet en a un extremum local. Pour fixer les idées, choisissons par exemple le cas où f admet un maximum local en a . Alors il existe $r > 0$ tel que

$$\forall x \in]a - r, a + r[, \quad f(x) \leq f(a).$$

Ainsi pour tout $h \in]-r, r[$, $f(a + h) - f(a) \leq 0$. Comme f est dérivable en a ,

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

On note $\theta_a(h) = [f(a + h) - f(a)]/h$ le taux de variation. Si $h > 0$ alors $\theta_a(h) \leq 0$ et en passant à la limite $f'(a) \leq 0$. Si $h < 0$ alors $\theta_a(h) \geq 0$ et en passant à la limite $f'(a) \geq 0$. D'où nécessairement $f'(a) = 0$. \square

Si $f'(a) = 0$ alors la tangente au graphe en $(a, f(a))$ est horizontale. Ce théorème fournit seulement une **condition nécessaire**, car on peut avoir l'une des trois situations décrites dans la figure 10.2.

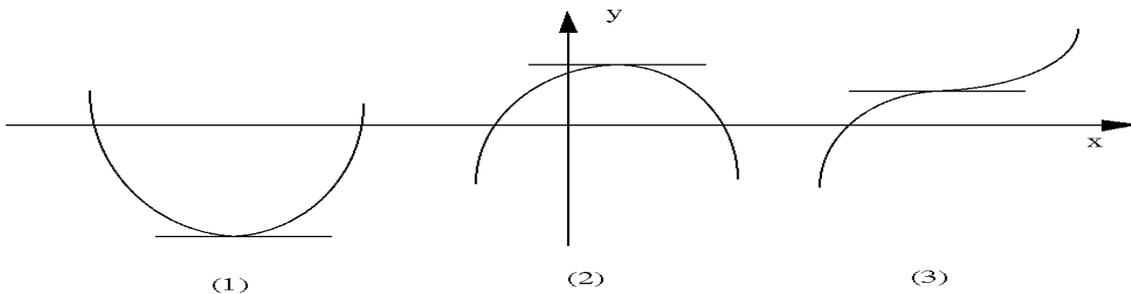


FIGURE 10.2. Nature possible des points critiques : (1) minimum local, (2) maximum local, (3) point d'inflexion

Définition 10.7. — On appelle *point critique* ou *point stationnaire* de f tout point $a \in \mathcal{D}_f$ tel que f est dérivable en a et $f'(a) = 0$.



Si f est dérivable sur I ouvert, tout extremum de f est atteint en un point critique de f mais la réciproque est fautive. Un point critique est seulement un point candidat.

10.2.2. Nature des points critiques. —

$$\Delta f_a(h) = f(a + h) - f(a)$$

Soit a un point critique de f , on étudie le signe de la différence pour h dans un voisinage de a . Si le signe de $\Delta f_a(h)$ est facile à étudier (par exemple pour les fonctions polynômes de faible degré) on peut faire une étude directe. Mais en général ce n'est pas le cas et on utilise plutôt l'un des trois théorèmes qui suivent.

Supposons que f est convexe sur I , on a alors

$$\forall x \in I, \quad f(x) - f(a) \geq (x - a)f'(a).$$

Si $f'(a) = 0$ on en déduit que pour tout $x \in I$, $f(x) - f(a) \geq 0$ donc $f(a)$ est un minimum global sur I . Pour une fonction concave sur I , on aurait trouvé un maximum global sur I .

Théorème 10.8 (Extrema d'une fonction convexe ou concave). — Soit f définie sur \mathcal{D}_f contenant un intervalle ouvert I tel que f soit de classe \mathcal{C}^1 sur I . Soit $a \in I$ un point critique de f , i.e. $f'(a) = 0$,

- si f est convexe sur I alors f admet un minimum global sur I en a ,
- si f est concave sur I alors f admet un maximum global sur I en a .

En étudiant le signe de la dérivée, on obtient une condition suffisante d'extremum local.

Théorème 10.9. — Soit f définie sur \mathcal{D}_f contenant un intervalle ouvert I tel que f soit de classe \mathcal{C}^1 sur I . Soit $a \in I$ un point critique de f , i.e. $f'(a) = 0$,

- si, au voisinage de a , f' change de signe du $-$ au $+$ alors f admet un minimum local en a ,
- si, au voisinage de a , f' change de signe du $+$ au $-$ alors f admet un maximum local en a .

Pour éviter l'étude, pas toujours facile, du signe de la dérivée, on peut utiliser le théorème suivant.

Théorème 10.10. — Soit f définie sur \mathcal{D}_f contenant un intervalle ouvert I tel que f est de classe \mathcal{C}^∞ sur I . Soit $a \in I$ un point critique de f et soit $p \in \mathbb{N}$ tel que $f^{(p)}(a)$ est la première dérivée non nulle au point a . Alors

- (i) si p est impair alors f n'admet pas d'extremum en a ,
- (ii) si p est pair et $f^{(p)}(a) < 0$ alors f admet un maximum local en a ,
- (iii) si p est pair et $f^{(p)}(a) > 0$ alors f admet un minimum local en a .

En général on applique ce théorème avec $p = 2$ ou $p = 3$.

Preuve. — Écrivons la formule de Taylor-Young à l'ordre p au voisinage de a , pour tout $h \in \mathbb{R}$

$$a + h \in I \implies f(a + h) = f(a) + h^p \frac{f^{(p)}(a)}{p!} + h^p \varepsilon(h)$$

où ε est une fonction continue en 0 avec $\varepsilon(0) = 0$. En particulier

$$\Delta f_a(h) = f(a + h) - f(a) = h^p \left[\frac{f^{(p)}(a)}{p!} + \varepsilon(h) \right].$$

- Si p est pair, $\Delta f_a(h)$ est de signe constant, celui de $f^{(p)}(a)$. On a donc un extremum local en a .
- Si p est impair, $\Delta f_a(h)$ change de signe, comme h^p , et a ne donne pas un extremum.

□

Exemple 10.11. —

1. Soit f la fonction définie sur \mathbb{R} par $f(x) = x^4 + 1$. On pose $I = \mathcal{D}_f = \mathbb{R}$. On a $f'(x) = 4x^3$ et donc $x = 0$ est le seul point critique. Comme $f(x) \geq 1 = f(0)$ pour chaque $x \in \mathbb{R}$, on peut conclure que f admet sur \mathbb{R} un minimum global en $x = 0$ de valeur 1.
2. Soit f la fonction définie sur \mathbb{R} par $f(x) = x^3 - 3x + 2$. On pose $I = \mathcal{D}_f = \mathbb{R}$. On a $f'(x) = 3x^2 - 3$ et l'ensemble des points critiques est $\{-1, 1\}$. Comme $f''(-1) = -6 < 0$, la fonction f admet en $x = -1$ un maximum local de valeur $f(-1) = 4$. Comme $f''(1) = 6 > 0$, la fonction f admet en $x = 1$ un minimum local de valeur $f(1) = 0$. On a $f(3) = 20 > 4$, ainsi 4 n'est pas un maximum global. On a $f(-3) = -16 < 0$, ainsi 0 n'est pas un minimum global.
3. Soit f la fonction définie sur $]0, +\infty[$ par $f(x) = 2x - x \ln x$. Pour tout $x > 0$, $f'(x) = 2 - \ln x - 1$. Il y a un seul point critique $x = e$. Or pour tout $x > 0$, $f''(x) = -1/x < 0$, donc f est concave sur $]0, +\infty[$ et admet un maximum global (et donc local) sur $]0, +\infty[$ en $x = e$ de valeur $f(e) = e$.

10.3. Fonction dérivable sur un intervalle fermé borné

D'après le Théorème 3.16, si une fonction f est définie et continue sur un intervalle fermé borné (segment), alors elle atteint ses bornes, c'est-à-dire il existe au moins un point $\underline{x} \in I$ qui donne un minimum global pour f sur I et il existe au moins un point $\bar{x} \in I$ qui donne un maximum global pour f sur I .

Pour trouver les extrema de f sur I , il suffit d'étudier la fonction sur l'intervalle $\overset{\circ}{I}$ (intervalle I privé de ses extrémités) qui est un intervalle ouvert, puis de comparer les valeurs des extrema éventuels avec les valeurs aux bornes de I .

Exemple 10.12. — On cherche les extrema sur $I = [-1, 2]$ de la fonction f définie par $f(x) = x^2$.

- Sur $] - 1, 2[$ la fonction f' s'annule en $x = 0$, elle est convexe donc f admet un minimum global sur $] - 1, 2[$ en $x = 0$.
 - On calcule $f(0) = 0$, $f(-1) = 1$ et $f(2) = 4$.
 - Par suite, sur $I = [-1, 2]$ la fonction f admet en $x = 2$ un maximum global et en $x = 0$ un minimum global. De plus f admet en $x = -1$ un maximum local de valeur 1.
- Il faut bien remarquer que pour $x = 2$ on a $f'(2) \neq 0$ et pour $x = -1$, on a $f'(-1) = -2 \neq 0$. Ce ne sont pas des points critiques, seul $x = 0$ est un point critique.

10.4. Fonction dérivable par morceaux sur un intervalle

Soit f une fonction continue définie sur \mathcal{D}_f contenant un intervalle I . On suppose que f est dérivable sur I sauf en un nombre fini de points de I . On considère tous les sous-intervalles ouverts de I sur lesquels f est dérivable.

1. On recherche alors les extrema de f sur ces intervalles (s'il en existe) en appliquant les méthodes du paragraphe 10.2.
2. Ensuite on fait une étude aux points de raccordement. La fonction f n'est en général pas dérivable en ces points de raccordement, mais elle est dérivable à gauche et à droite de ces points. En se servant du tableau de variations au voisinage de ces points, on constate que f y admet un extremum si la dérivée change de signe (du $+$ au $-$ pour un maximum, et du $-$ au $+$ pour un minimum).
3. On termine par une étude aux bornes de I , si celles-ci sont finies et appartiennent à I .

Exemple 10.13. — Soit f la fonction définie sur \mathbb{R} par $f(x) = |x(x - 1)|$, c'est-à-dire

$$f(x) = \begin{cases} x(1 - x) & \text{si } x \in [0, 1] \\ x(x - 1) & \text{si } x \in] - \infty, 0[\cup] 1, +\infty[. \end{cases}$$

On cherche les extrema de f sur $I = [1, +\infty[$. Attention : I n'est pas un segment !

- On commence par chercher les extrema locaux de f sur chacun des intervalles $]0, 1[$, $] - 1, 0[$ et $]1, +\infty[$. On trouve que f admet un maximum local en $x = 1/2$ de valeur $1/4$.
- Pour les points $x = 0$ et $x = 1$, la dérivée change de signe. On vérifie que ces points donnent des minima a priori locaux. Or $f(1) = f(0) = 0$ et f est positive sur \mathbb{R} , donc ces minima sont globaux (notons que ces points ne sont pas des points critiques).
- Pour $x = -1$, on a un maximum local sur I , car la fonction est décroissante sur $[-1, 0]$.

CHAPITRE 11

GÉOMÉTRIE DE \mathbb{R}^2 ET \mathbb{R}^3

11.1. Les ensembles \mathbb{R}^2 et \mathbb{R}^3

11.1.1. Définitions. — Ces deux ensembles sont déjà connus du lecteur. Rappelons-en les définitions :

Définition 11.1. — \mathbb{R}^2 est l'ensemble formé de tous les couples de réels :

$$\mathbb{R}^2 = \{(x, y) \mid x \in \mathbb{R}, y \in \mathbb{R}\}.$$

Définition 11.2. — \mathbb{R}^3 est l'ensemble formé de tous les triplets de réels :

$$\mathbb{R}^3 = \{(x, y, z) \mid x \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{R}\}.$$

Nous avons adopté les notations les plus courantes : (x, y) pour un élément de \mathbb{R}^2 , (x, y, z) pour un élément de \mathbb{R}^3 . On pourra aussi noter ces éléments respectivement (x_1, x_2) et (x_1, x_2, x_3) .

Les réels seront appelés *scalaires* pour les distinguer des éléments de \mathbb{R}^2 ou \mathbb{R}^3 .

Notation 11.3. — L'élément nul de \mathbb{R}^2 est $(0, 0)$. Celui de \mathbb{R}^3 est $(0, 0, 0)$. On note aussi 0 l'élément nul de \mathbb{R}^n ($n = 2$ ou 3).

Exemple 11.4. — Un consommateur dispose de deux biens : X_1 et X_2 . On suppose une unité choisie pour chaque bien et on note, pour $i = 1$ ou 2 , x_i ($x_i \geq 0$) la quantité consommée du bien X_i . Le couple (x_1, x_2) est appelé "panier de biens".

Proposition 11.5 (Égalité de deux couples ou de deux triplets de réels)

$$\begin{aligned} (x_1, x_2) &= (y_1, y_2) && \text{si et seulement si } x_1 = y_1 \text{ et } x_2 = y_2 \\ (x_1, x_2, x_3) &= (y_1, y_2, y_3) && \text{si et seulement si } x_1 = y_1, x_2 = y_2 \text{ et } x_3 = y_3 \end{aligned}$$

Remarque 11.6. — Les couples et les triplets sont ordonnés. On a $(1, 2, 3) \neq (3, 2, 1)$.

11.1.2. Opérations. — Définissons les opérations dans \mathbb{R}^3 (pour \mathbb{R}^2 il suffira de supprimer les termes d'indice 3).

Définition 11.7. — Soient $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, $y = (y_1, y_2, y_3) \in \mathbb{R}^3$ et $\lambda \in \mathbb{R}$. On définit les opérations suivantes :

- *Addition* : $x + y = (x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$
 - *Multiplication par un scalaire* : $\lambda x = \lambda(x_1, x_2, x_3) = (\lambda x_1, \lambda x_2, \lambda x_3)$
 - *Opposé* : $-x = (-1)(x_1, x_2, x_3) = (-x_1, -x_2, -x_3)$
 - *Soustraction* : $x - y = x + (-y) = (x_1 - y_1, x_2 - y_2, x_3 - y_3)$
- $x + y$ s'appelle la *somme* de x et y et $x - y$ s'appelle la *différence* de x et y .

11.2. Points et vecteurs

11.2.1. Représentation géométrique de \mathbb{R}^2 . — Une feuille de papier ne nous permet de faire correctement que la représentation graphique de \mathbb{R}^2 . Soit un repère orthonormé (O, \vec{i}, \vec{j}) du plan. Nous allons voir deux interprétations possibles pour représenter les éléments de \mathbb{R}^2 .

11.2.1.1. Ensemble de points. —

Définition 11.8. — Dans le repère, à tout couple (x, y) de réels correspond un point M du plan et un seul : x est son *abscisse*, y son *ordonnée*. On note $M = (x, y)$ le point ainsi repéré et x et y sont appelées *coordonnées* de M dans le repère choisi.

Un point représente une *position* particulière dans le plan. Les coordonnées du point O sont $(0, 0)$.

11.2.1.2. Ensemble de vecteurs. — On peut également interpréter un couple (x, y) de réels comme un *déplacement* de x unités (vers la droite si $x > 0$, vers la gauche si $x < 0$), de y unités (vers le haut si $y > 0$, vers le bas si $y < 0$).

Pour figurer un tel déplacement, il faut un point de départ, appelé *origine*, qui marque la position initiale, et un point d'arrivée, appelé *extrémité*, qui marque la position finale après déplacement.

Les déplacements sont figurés par des flèches. Un déplacement est aussi appelé *vecteur*, x et y sont les *coordonnées* du vecteur.

Nous noterons \vec{AB} un vecteur ou déplacement dont l'origine est le point A et l'extrémité le point B (figure 11.1).

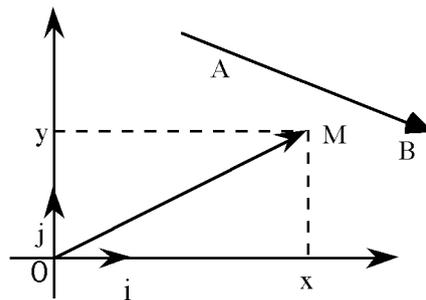


FIGURE 11.1. Représentation du vecteur \vec{AB} dans un repère orthonormé

Si $A = (x_A, y_A)$ et $B = (x_B, y_B)$, le vecteur ou déplacement a pour coordonnées $(x_B - x_A, y_B - y_A)$ puisque pour aller de A à B , il faut se déplacer de $x_B - x_A$ unités horizontalement et $y_B - y_A$ unités verticalement.

Un même déplacement peut être représenté par plusieurs couples de points : dans la figure 11.2, les trois flèches représentent le même déplacement, mais les positions initiales sont différentes. Nous venons de voir qu'il y avait une infinité de représentants pour un même vecteur ou déplacement car on peut choisir une origine quelconque. Remarquons qu'un couple (x, y) a comme représentant privilégié le vecteur où le point M a comme coordonnées (x, y) (figure 11.2).

Si on considère les couples de \mathbb{R}^2 comme des déplacements (ou vecteurs) on peut :

- faire successivement deux déplacements, le deuxième ayant pour origine l'extrémité du premier (c'est la somme),
- changer le sens d'un déplacement (multiplication par -1),
- réduire ou augmenter un déplacement (multiplication par un réel λ).

Autrement dit, les opérations définies dans le paragraphe 11.1.2 ont une signification concrète si on considère les éléments de \mathbb{R}^2 comme des vecteurs.

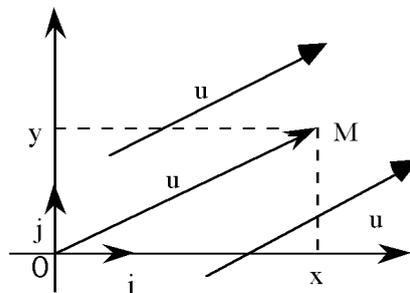


FIGURE 11.2. Les trois flèches représentent le même déplacement, mais les positions initiales sont différentes

En choisissant de représenter tous les couples de \mathbb{R}^2 par des vecteurs d'origine O , les opérations définies dans le paragraphe précédent se représentent géométriquement (voir les figures 11.3 et 11.4).

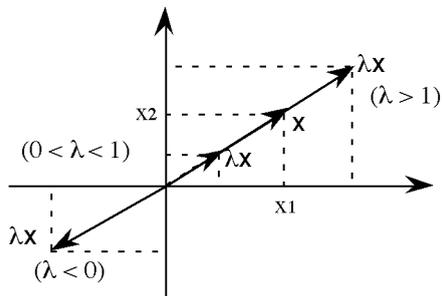


FIGURE 11.3. Multiplication d'un vecteur par un réel. On a distingué les cas : $\lambda < 0$, $0 < \lambda < 1$ et $\lambda > 1$

11.2.2. Cas de \mathbb{R}^3 . — Ce qui a été vu dans le paragraphe précédent peut se généraliser à \mathbb{R}^3 en prenant un repère orthonormé dans l'espace. Un triplet correspondra soit à un point, soit à un vecteur. La troisième coordonnée s'appelle la "côte".

11.2.3. Relation entre points et vecteurs. — Selon le contexte on considérera un élément de \mathbb{R}^2 ou \mathbb{R}^3 , soit comme un point, soit comme un vecteur. Si on le considère comme un point, on le notera par une lettre **majuscule** (par exemple M). Si on le considère comme un vecteur, il sera noté par une lettre **minuscule cursive** (par exemple v).

Nous raisonnerons dans la suite de ce paragraphe avec $n = 3$. Pour $n = 2$, il suffit de supprimer la troisième coordonnée.

À tout couple de point (A, B) où $A = (x_A, y_A, z_A)$ et $B = (x_B, y_B, z_B)$, on peut associer le vecteur défini par

$$\vec{AB} = (x_B - x_A, y_B - y_A, z_B - z_A).$$

On constate alors, d'après la définition des opérations, que l'on peut écrire :

$$\vec{AB} = B - A \quad \text{ou encore} \quad B = A + \vec{AB}.$$

Remarque 11.9. — Il faut bien comprendre que la "différence de deux points" est un vecteur !

Réciproquement, à tout vecteur $u = (x, y, z)$ on peut associer une infinité de couple de points : si $O = (0, 0, 0)$ et $M = (x, y, z)$ on a $u = \vec{OM}$, c'est-à-dire M, u et \vec{OM} représentent le même triplet.

Les propriétés qui suivent sont déjà connues du lecteur et se justifient simplement en utilisant la définition des opérations (voir page 41) :

Proposition 11.10. — Soit $n = 2$ ou $n = 3$ et soient trois points A, B et C de \mathbb{R}^n , alors

- (i) $\vec{BA} = -\vec{AB}$
- (ii) $\vec{AB} = 0 \iff A = B$
- (iii) Relation de Chasles : $\vec{AB} + \vec{BC} = \vec{AC}$

Preuve. — Raisonons pour $n = 3$. Posons $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$ et $C = (x_C, y_C, z_C)$.

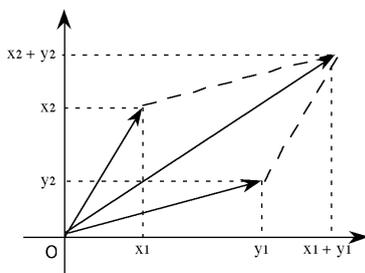


FIGURE 11.4. Somme de deux vecteurs

- *Preuve de (i).* Par définition, $\overrightarrow{AB} = (x_B - x_A, y_B - y_A, z_B - z_A)$ et $\overrightarrow{BA} = (x_A - x_B, y_A - y_B, z_A - z_B) = -(x_B - x_A, y_B - y_A, z_B - z_A)$, d'où le résultat.
- *Preuve de (ii).* Supposons $\overrightarrow{AB} = (x_B - x_A, y_B - y_A, z_B - z_A) = (0, 0, 0)$. Alors avec la définition de l'égalité de deux triplets : $x_B - x_A = 0$, $y_B - y_A = 0$ et $z_B - z_A = 0$. Ce qui s'écrit aussi : $x_B = x_A$, $y_B = y_A$ et $z_B = z_A$, c'est-à-dire $A = B$. La réciproque est évidente.
- *Preuve de (iii).* On a $\overrightarrow{AB} = (x_B - x_A, y_B - y_A, z_B - z_A)$ et $\overrightarrow{BC} = (x_C - x_B, y_C - y_B, z_C - z_B)$. Donc $\overrightarrow{AB} + \overrightarrow{BC} = (x_C - x_A, y_C - y_A, z_C - z_A) = \overrightarrow{AC}$.

□

11.2.4. Interprétation. — Cette **distinction point - vecteur** a une signification concrète comme le montre l'exemple suivant.

Exemple 11.11. — Un agent économique possède, à l'année initiale, un avoir de 10 000\$ et de 10 000€. Si on prend ces quantités comme unités de compte, la position initiale de l'agent est le couple $(1, 1)$. Ce couple correspond à une situation instantanée, donc à un point : $A = (1, 1)$.

Supposons que son avoir en dollars augmente au rythme de 10 000 \$ par an et son avoir en euros de 20 000€ par an. Dans la même unité, cette évolution se traduit par le couple $(1, 2)$. Or les variations s'ajoutent entre elles, il est donc naturel de les considérer comme des vecteurs et si on pose $u = (1, 2)$, au bout de n années, la position de l'agent sera B , définie par $\overrightarrow{AB} = nu = B - A$.

Comme on peut écrire $B = A + \overrightarrow{AB}$, on a finalement :

$$B = A + nu = (1, 1) + n(1, 2).$$

Les positions initiale et finale correspondent à des situations instantanées donc sont représentées par des points A et B , mais l'évolution de la situation de l'agent est représentée par un vecteur.

11.3. Produit scalaire, norme et distance associée

Dans \mathbb{R} la longueur du segment $[x, y]$ est $|x - y|$. C'est aussi la distance de x à y .

Comment étendre cette notion à \mathbb{R}^2 ?

Cette question nous amène à définir la notion de norme. Reprenons les notations utilisées dans le paragraphe 11.2.1. Par le théorème de Pythagore la longueur du vecteur \overrightarrow{OM} est donnée par $\sqrt{x^2 + y^2}$. La distance de $A = (x_A, y_A)$ à $B = (x_B, y_B)$ est

$$d(A, B) = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}.$$

Il est facile de généraliser ces définitions pour $n = 3$. Dans les paragraphes 11.3.1 et 11.3.2 (produit scalaire et norme), les couples et triplets sont considérés comme des vecteurs. Dans le paragraphe 11.3.3 (distance) ils sont considérés comme des points.

11.3.1. Produit scalaire euclidien. —

Définition 11.12. — Pour $n = 3$ et pour tout $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ et tout $y = (y_1, y_2, y_3) \in \mathbb{R}^3$ on appelle *produit scalaire de x et y* le nombre **réel**, noté $\langle x, y \rangle$ défini par :

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

Les vecteurs x et y sont dits *orthogonaux* si $\langle x, y \rangle = 0$.

Pour $n = 2$ il suffit de supprimer $x_3 y_3$.

Le produit scalaire est ainsi nommé car il associe à deux vecteurs de \mathbb{R}^2 (ou de \mathbb{R}^3) un nombre réel (appelé aussi un scalaire).

Exemple 11.13. — Reprenons la suite de l'exemple du consommateur. Si on note, pour $i = 1$ ou 2 , p_i le prix unitaire du bien X_i , la dépense du consommateur est donc :

$$D(x_1, x_2) = p_1 x_1 + p_2 x_2.$$

En notant $x = (x_1, x_2)$ et $p = (p_1, p_2)$, on peut écrire $D(x) = \langle p, x \rangle$. p est appelé *vecteur prix*.

11.3.2. Norme euclidienne. —

Définition 11.14. — Pour tout x dans \mathbb{R}^n ($n = 2$ ou 3), on appelle *norme euclidienne* de x le nombre réel noté $\|x\|$ défini par :

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

- Pour $n = 2$ si $x = (x_1, x_2)$ alors $\|x\| = \sqrt{x_1^2 + x_2^2}$.
- Pour $n = 3$ si $x = (x_1, x_2, x_3)$ alors $\|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$.

Remarque 11.15. — $\| -x \| = \|x\|$ et si $n = 1$ on retrouve : $\|x\| = |x| = \sqrt{x^2}$.

Proposition 11.16 (Inégalité de Cauchy-Schwarz). — Soit $n = 2$ ou 3 . Pour tout x et tout y dans \mathbb{R}^n , on a

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Cas où l'inégalité est une égalité :

$$|\langle x, y \rangle| = \|x\| \|y\| \iff \begin{cases} \text{soit } x = 0, & \text{soit } y = 0, \\ \text{soit } x \neq 0, y \neq 0 & \text{et } \exists \lambda \in \mathbb{R}^*, y = \lambda x. \end{cases}$$

L'inégalité de Cauchy-Schwarz est automatiquement vérifiée si l'un des vecteurs est nul. Dans ce cas, c'est une égalité : $0 = 0$.

Il reste donc à démontrer les résultats de la proposition 11.16 pour $x \neq 0$ et $y \neq 0$.

Proposition 11.17 (Propriétés de la norme ($n = 2$ ou $n = 3$)). — Pour tout x et tout y dans \mathbb{R}^n , on a

- (i) $\|x\| \geq 0$.
- (ii) $\|x\| = 0 \iff x = 0$.
- (iii) Pour tout $\lambda \in \mathbb{R}$, $\|\lambda x\| = |\lambda| \|x\|$.
- (iv) Inégalité triangulaire : $\|x + y\| \leq \|x\| + \|y\|$.
- (v) $|\|x\| - \|y\|| \leq \|x \pm y\| \leq \|x\| + \|y\|$.

On constate que la norme euclidienne possède exactement les mêmes propriétés que la valeur absolue, ce qui est naturel puisque la valeur absolue n'est autre que la norme euclidienne pour $n = 1$.

Ces propriétés seront démontrées un l'exercice.

11.3.3. Distance euclidienne. —

Définition 11.18. — Pour tous points $A \in \mathbb{R}^n$ et $B \in \mathbb{R}^n$, ($n = 2$ ou $n = 3$) on appelle *distance euclidienne* du point A au point B le réel, noté $d(A, B)$, défini par :

$$d(A, B) = \|\overrightarrow{AB}\| = \|B - A\|.$$

- Pour $n = 2$ si $A = (a_1, a_2)$ et $B = (b_1, b_2)$ alors

$$d(A, B) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2}.$$

- Pour $n = 3$ si $A = (a_1, a_2, a_3)$ et $B = (b_1, b_2, b_3)$ alors

$$d(A, B) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2}.$$

Remarque 11.19. — Si $n = 1$, pour $A = a$ et $B = b$ on retrouve $d(A, B) = |b - a|$.

Proposition 11.20 (Propriétés de la distance euclidienne). — Pour tous points A, B et C dans \mathbb{R}^n , on a

- (i) Symétrie : $d(A, B) = d(B, A)$
- (ii) $d(A, B) = 0 \iff A = B$
- (iii) Inégalité triangulaire : $d(A, B) \leq d(A, C) + d(C, B)$

Preuve. — Il suffit d'utiliser la définition et les propriétés de la norme euclidienne. Pour A, B et C dans \mathbb{R}^n , on peut écrire :

- (i) $d(A, B) = \|\overrightarrow{AB}\| = \|-\overrightarrow{BA}\| = \|\overrightarrow{BA}\| = d(B, A)$
- (ii) $d(A, B) = 0 \Leftrightarrow \|\overrightarrow{AB}\| = 0 \Leftrightarrow \overrightarrow{AB} = \mathbf{0} \Leftrightarrow A = B$ d'après (ii) de la proposition 11.17 et (iv) de la proposition 11.10
- (iii) $d(A, B) = \|AB\| = \|AC + CB\| \leq \|\overrightarrow{AC}\| + \|\overrightarrow{CB}\|$ d'après (4) de la proposition 11.17. On en déduit l'inégalité triangulaire.

□

Dans la suite de ce chapitre, nous allons définir quelques sous-ensembles remarquables de \mathbb{R}^2 et de \mathbb{R}^3 .

11.4. Produit cartésien de deux intervalles

Définition 11.21. — Soient I et J deux intervalles quelconques de \mathbb{R} . On appelle *produit cartésien de I et J* , le sous-ensemble de \mathbb{R}^2 , noté $I \times J$, défini par :

$$I \times J = \{(x, y) \in \mathbb{R}^2 \mid x \in I \text{ et } y \in J\}.$$

En particulier, rappelons que $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} =]-\infty, +\infty[\times]-\infty, +\infty[$.

Quelques exemples sont représentés sur la figure 11.5.

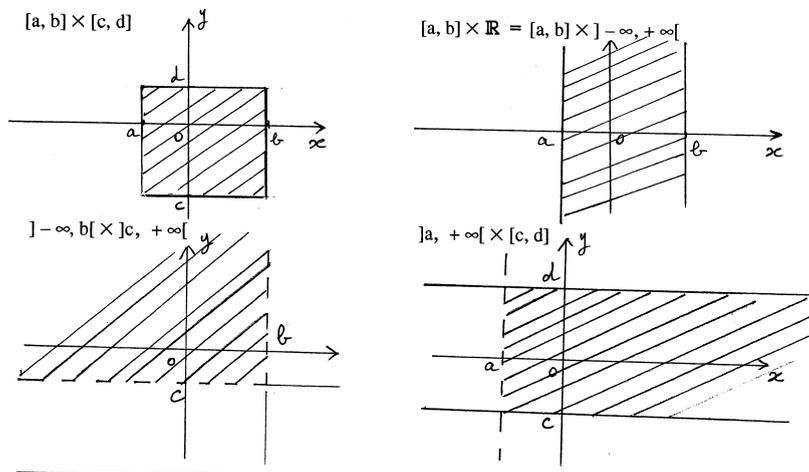


FIGURE 11.5. Exemples de produits cartésiens : $a, b, c, d \in \mathbb{R}$

11.5. Droite dans \mathbb{R}^2 et dans \mathbb{R}^3

11.5.1. Définition. —

Définition 11.22. — Soient A un point de \mathbb{R}^n ($n = 2$ ou $n = 3$) et soit v un vecteur non nul de \mathbb{R}^n . On appelle *droite passant par A et de vecteur directeur v* le sous-ensemble de \mathbb{R}^n , défini par :

$$\mathcal{D}(A, v) = \{M \in \mathbb{R}^n \mid \exists t \in \mathbb{R}, \overrightarrow{AM} = tv\} = \{M \in \mathbb{R}^n \mid \exists t \in \mathbb{R}, M = A + tv\}.$$

Remarque 11.23. — La notation $\mathcal{D}(A, v)$ n'est pas standard, mais pratique et suffisamment explicite.

Remarque 11.24. — Pour définir une droite il faut donc un point A et un vecteur non nul. Le point A appartient bien à la droite (avec le choix $t = 0$). C'est pourquoi on dit que *la droite passe par A* . Une droite contient une infinité de points (car t décrit \mathbb{R}), donc au moins deux points distincts.

Remarque 11.25. — Il n'y a **pas unicité** du vecteur directeur. Tout vecteur w de la forme $w = \lambda v$ avec $\lambda \neq 0$, c'est-à-dire proportionnel (on dit aussi colinéaire) à v , est encore un vecteur directeur de $\mathcal{D}(A, v)$. En effet, pour tout point M de $\mathcal{D}(A, v)$ on peut écrire $\overrightarrow{AM} = tv = \alpha w$ avec $\alpha = t/\lambda$. En particulier, si B est un point de $\mathcal{D}(A, v)$, distinct de A , il existe un réel t tel que $\overrightarrow{AB} = tv$ et alors $\overrightarrow{BA} = -tv$. Par suite les vecteurs \overrightarrow{AB} et \overrightarrow{BA} sont des vecteurs directeurs de $\mathcal{D}(A, v)$.

11.5.1.1. *Droite définie par deux points distincts.* — Soient A et B deux points distincts de \mathbb{R}^n . On a vu que si une droite contient ces deux points, le vecteur \overrightarrow{AB} est un vecteur directeur de la droite considérée.

Montrons que les droites $\mathcal{D}(A, \overrightarrow{AB})$ et $\mathcal{D}(A, \overrightarrow{BA})$ sont confondues. Si $M \in \mathcal{D}(A, \overrightarrow{AB})$ alors il existe $t \in \mathbb{R}$, tel que $\overrightarrow{AM} = t \overrightarrow{AB}$. Or $\overrightarrow{BM} = \overrightarrow{BA} + \overrightarrow{AM} = (t-1) \overrightarrow{AB}$, c'est-à-dire $M \in \mathcal{D}(B, \overrightarrow{BA})$. On montre l'inclusion dans l'autre sens de la même façon.

Autrement dit, deux points distincts définissent une droite et une seule.

Proposition 11.26. — *Si A et B sont deux points distincts, la droite passant par les points A et B est le sous-ensemble défini par :*

$$\mathcal{D}(A, B) = \{M \in \mathbb{R}^n \mid \exists t \in \mathbb{R}, \overrightarrow{AM} = t \overrightarrow{AB}\}.$$

La notation $\mathcal{D}(A, B)$ n'est pas standard mais pratique. On peut échanger les rôles joués par A et B .

En conclusion, une droite de \mathbb{R}^n ($n = 2$ ou $n = 3$) peut être définie par un point et un vecteur directeur ou par deux points distincts.

11.5.2. Équations d'une droite. — *Écrire les équations d'une droite, c'est exprimer les coordonnées d'un point quelconque de la droite en fonction des coordonnées, "des données", de la droite considérée dans un repère du plan.*

11.5.2.1. *Équations paramétriques d'une droite définie par un point et un vecteur directeur.* — Il suffit d'appliquer la définition. Supposons $n = 3$ et soient $A = (\alpha, \beta, \gamma)$ et $v = (v_1, v_2, v_3)$. Si $M \in \mathcal{D}(A, v)$ on a $\overrightarrow{AM} = tv$. En traduisant cette relation vectorielle, les coordonnées (x, y, z) d'un point M de $\mathcal{D}(A, v)$ satisfont :

$$\begin{cases} x - \alpha = t v_1 \\ y - \beta = t v_2 \\ z - \gamma = t v_3, \quad t \in \mathbb{R}. \end{cases}$$

Pour $n = 2$ on a seulement deux équations.

11.5.2.2. *Équation paramétrique d'une droite définie par deux points.* — Considérons deux points distincts A et B . Soit M un point de la droite joignant ces deux points. Or, $M \in \mathcal{D}(A, \overrightarrow{AB})$ si et seulement si il existe un réel t tel que $\overrightarrow{AM} = t \overrightarrow{AB}$. Cette égalité peut s'écrire $M - A = t(B - A)$ ou encore $M = (1-t)A + tB$ avec $t \in \mathbb{R}$. Mais, on a aussi $M \in \mathcal{D}(B, \overrightarrow{BA})$. On aurait ainsi obtenu $M = tA + (1-t)B$ ($t \in \mathbb{R}$) ce qui confirme le rôle symétrique joué par A et B . Ceci conduit à la propriété suivante :

Proposition 11.27. — *Dans \mathbb{R}^n , $n = 2$ ou 3 , la droite passant par les deux points A et B ($A \neq B$) est l'ensemble des points M pouvant s'écrire sous la forme :*

$$M = tA + (1-t)B \quad \text{ou} \quad M = tB + (1-t)A \quad \text{avec } t \in \mathbb{R}.$$

Les rôles de A et B sont symétriques.

On obtient ainsi une *équation paramétrique vectorielle* qui, en passant aux coordonnées, se traduit par deux équations pour $n = 2$ et par trois équations pour $n = 3$.

11.5.2.3. *Équation cartésienne d'une droite dans \mathbb{R}^2 .* — Les résultats de ce paragraphe **ne sont valables que pour** $n = 2$ et devraient évoquer quelques souvenirs...

Proposition 11.28. — *\mathcal{D} est une droite de \mathbb{R}^2 si et seulement si il existe trois réels a, b et c avec $(a, b) \neq (0, 0)$ tels que $\mathcal{D} = \{(x, y) \in \mathbb{R}^2 \mid ax + by + c = 0\}$.*

Définition 11.29. — " $ax + by + c = 0$ " est appelée *équation cartésienne* de \mathcal{D} . Le vecteur $(-b, a)$ est un vecteur directeur de \mathcal{D} .

Preuve de la proposition 11.28. Soit \mathcal{D} une droite de \mathbb{R}^2 . Elle est définie par un point $A = (\alpha, \beta)$ et un vecteur directeur $v = (v_1, v_2) \neq (0, 0)$. Soit $M = (x, y)$ un point de $\mathcal{D}(A, v)$. Par définition, il existe un réel t tel que $\overrightarrow{AM} = (x - \alpha, y - \beta) = (t v_1, t v_2)$. On a donc le système

$$(11.1) \quad \begin{cases} x - \alpha = t v_1, \\ y - \beta = t v_2. \end{cases}$$

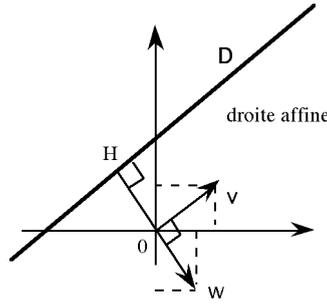


FIGURE 11.6. Droite affine

On suppose d'abord que $v_1 \neq 0$ et $v_2 \neq 0$. En multipliant la première égalité par v_2 , la deuxième par $-v_1$ et en ajoutant les deux on obtient :

$$(11.2) \quad v_2(x - \alpha) - v_1(y - \beta) = 0$$

c'est-à-dire :

$$v_2x - v_1y + (\beta v_1 - \alpha v_2) = 0.$$

Il suffit alors de poser : $a = v_2$, $b = -v_1$, $c = (\beta v_1 - \alpha v_2)$, ce qui donne $v = (-b, a)$, pour montrer que les coordonnées du point M vérifient $ax + by + c = 0$.

Ce raisonnement reste valable si $v_1 = 0$ ou si $v_2 = 0$. En effet,

- si $v_1 = 0$ et $v_2 \neq 0$, la première équation dans le système (11.1) donne $x = \alpha$, ainsi que l'équation (11.2).
- si $v_2 = 0$ et $v_1 \neq 0$, la deuxième équation dans le système (11.1) donne $y = \beta$, ainsi que l'équation (11.2).

Finalement, on a montré que si $M = (x, y) \in \mathcal{D}(A, v)$, $v \neq 0$, alors il existe trois réels a , b et c avec $a \neq 0$ ou $b \neq 0$ ne dépendant que de A et v , tels que $ax + by + c = 0$.

Réciproquement, soient trois réels a , b et c avec $a \neq 0$ ou $b \neq 0$. Considérons l'ensemble $E = \{M = (x, y) \in \mathbb{R}^2 \mid ax + by + c = 0\}$. L'ensemble E n'est pas vide : si $a \neq 0$, le point $(-c/a, 0) \in E$ et si $a = 0$ mais $b \neq 0$ le point $(0, -c/b) \in E$. Soit donc un point quelconque $A = (\alpha, \beta)$ de E , alors

$$(11.3) \quad a\alpha + b\beta + c = 0.$$

Considérons un autre point $M = (x, y)$ de E , alors

$$(11.4) \quad ax + by + c = 0.$$

En soustrayant (11.3) de (11.4) on obtient :

$$(11.5) \quad M = (x, y) \in E \Leftrightarrow a(x - \alpha) + b(y - \beta) = 0.$$

Cette dernière équation admet une infinité de solutions en $(x - \alpha, y - \beta)$: tout couple de la forme

$$(11.6) \quad x - \alpha = -tb,$$

$$(11.7) \quad y - \beta = ta$$

avec $t \in \mathbb{R}$ convient. Si on pose $v = (-b, a)$, les deux égalités (11.6) et (11.7) sont équivalentes à $\overrightarrow{AM} = tv$. Donc l'ensemble E est confondu avec la droite $\mathcal{D}(A, v)$. \square

Remarque 11.30. — L'équation cartésienne n'est pas unique. Les équations : $x + 2y + 3 = 0$ et $4x + 8y + 12 = 0$ définissent la même droite.

Dans le plan, on peut aussi définir une droite par un point et un vecteur orthogonal :

Proposition 11.31 ($n = 2$). — Soient $w = (a, b)$ un vecteur non nul de \mathbb{R}^2 et A un point de \mathbb{R}^2 . Alors l'ensemble

$$\mathcal{D} = \{M \in \mathbb{R}^2 \mid \langle \overrightarrow{AM}, w \rangle = 0\}$$

est la droite passant par A et de vecteur directeur $(-b, a)$.

Réciproquement, si \mathcal{D} est une droite de \mathbb{R}^2 , il existe un vecteur $w \in \mathbb{R}^2$ non nul tel que $\forall A \in \mathcal{D}, \mathcal{D} = \{M \in \mathbb{R}^2 \mid \langle \overrightarrow{AM}, w \rangle = 0\}$.

Définition 11.32. — On dit que \mathcal{D} est orthogonale au vecteur w ou que w est orthogonal à \mathcal{D} . Avec les notations de la proposition 11.28 le vecteur $w = (a, b)$ convient.

Preuve. — Soient $w = (a, b)$ un vecteur non nul de \mathbb{R}^2 et $A = (\alpha, \beta)$ un point de \mathbb{R}^2 . Posons $E = \{M \in \mathbb{R}^2 \mid \langle \overrightarrow{AM}, w \rangle = 0\}$. Alors $M = (x, y) \in E$ si et seulement si $a(x - \alpha) + b(y - \beta) = 0$. Or, d'après la proposition 11.28, si on pose $v = (-b, a)$, l'égalité à droite de l'équivalence est une équation cartésienne de la droite $\mathcal{D}(A, v)$. On a donc bien que $M = (x, y) \in E$ si et seulement si $M = (x, y) \in \mathcal{D}(A, v)$.

Réciproquement, soit \mathcal{D} une droite de \mathbb{R}^2 d'équation cartésienne $ax + by + c = 0$. Soient $A = (\alpha, \beta)$ et $M = (x, y)$ deux points distincts de \mathcal{D} . D'après l'égalité (11.5) dans la réciproque de proposition 11.28, on a $a(x - \alpha) + b(y - \beta) = 0$. Ceci signifie que le produit scalaire des vecteurs \overrightarrow{AM} et $w = (a, b)$ est nul. \square

Pour trouver une équation cartésienne d'une droite de \mathbb{R}^2 il est pratique d'utiliser la notion suivante

11.5.2.4. *Déterminant de deux vecteurs de \mathbb{R}^2 .* —

Définition 11.33. — Soient $x = (x_1, x_2)$ et $y = (y_1, y_2)$ deux vecteurs de \mathbb{R}^2 . On appelle *déterminant de x et y* le nombre réel défini par :

$$\det(x, y) = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1.$$

Ce déterminant est nul si et seulement si les deux vecteurs sont colinéaires.

Méthode 11.34. — Grâce à l'égalité (11.2) dans la démonstration de la proposition 11.28 on constate que pour trouver une équation cartésienne de la droite $\mathcal{D}(A, v)$ il suffit d'écrire que le déterminant $\det(\overrightarrow{AM}, v)$ est nul.

De même pour trouver une équation cartésienne de la droite \mathcal{D} passant par les points A et B , il suffit d'écrire que le déterminant $\det(\overrightarrow{AM}, \overrightarrow{AB})$ est nul.

11.5.3. *Demi-plans dans \mathbb{R}^2 .* —

Définition 11.35. — Une droite de \mathbb{R}^2 d'équation $ax + by + c = 0$ définit quatre régions du plan appelées *demi-plans* :

$$\begin{aligned} E_1 &= \{(x, y) \in \mathbb{R}^2 \mid ax + by + c > 0\} & E_2 &= \{(x, y) \in \mathbb{R}^2 \mid ax + by + c \geq 0\} \\ E_3 &= \{(x, y) \in \mathbb{R}^2 \mid ax + by + c < 0\} & E_4 &= \{(x, y) \in \mathbb{R}^2 \mid ax + by + c \leq 0\} \end{aligned}$$

11.6. *Plan dans \mathbb{R}^3*

11.6.1. *Définition.* —

Définition 11.36. — On appelle *plan \mathcal{P}* de \mathbb{R}^3 tout sous-ensemble pour lequel il existe quatre réels a, b, c, d avec $(a, b, c) \neq (0, 0, 0)$ vérifiant :

$$\mathcal{P} = \{(x, y, z) \in \mathbb{R}^3 \mid ax + by + cz + d = 0\}.$$

" $ax + by + cz + d = 0$ " est une *équation cartésienne* de \mathcal{P} .

Cette définition généralise la propriété 11.28 vue au paragraphe précédent pour une droite du plan :

- (i) Un plan contient une infinité de points.
- (ii) On admettra que trois points distincts non alignés définissent un plan et un seul.

11.6.2. *Plan orthogonal à un vecteur.* —

Proposition 11.37 ($n = 3$). — (i) Si \mathcal{P} est un plan de \mathbb{R}^3 alors il existe un vecteur non nul w de \mathbb{R}^3 tel que pour tout point A de \mathcal{P}

$$\mathcal{P} = \{M \in \mathbb{R}^3 \mid \langle \overrightarrow{AM}, w \rangle = 0\}.$$

Avec les notations de la définition $w = (a, b, c)$ convient. On dit que le vecteur w est orthogonal à \mathcal{P} ou que le plan \mathcal{P} est orthogonal à w .

- (ii) Réciproquement, si w est un vecteur non nul de \mathbb{R}^3 et si A est un point de \mathbb{R}^3 , alors l'ensemble

$$\mathcal{E} = \{M \in \mathbb{R}^3 \mid \langle \overrightarrow{AM}, w \rangle = 0\}$$

est un plan. C'est le plan passant par A et orthogonal à w .

Les démonstrations de cette proposition sont analogues à celles de la proposition 11.31.

11.7. Cercles et sphères

Soient A un point de \mathbb{R}^n ($n = 2$ ou $n = 3$) et r un réel positif ou nul fixés. On considère dans ce paragraphe l'ensemble des points M de \mathbb{R}^n situés à la distance r du point A . On a donc : $d(A, M) = r$, ou de manière équivalente (puisque la distance est toujours positive ou nulle), $d(A, M)^2 = r^2$.

Avec la définition de la distance euclidienne d on obtient les définitions suivantes

Définition 11.38 ($n = 2$). — On appelle *cercle de centre* $A = (a, b)$ et de rayon $r \geq 0$, l'ensemble des points M de \mathbb{R}^2 défini par :

$$\mathcal{C}(A, r) = \mathcal{C}((a, b), r) = \{M = (x, y) \in \mathbb{R}^2 \mid (x - a)^2 + (y - b)^2 = r^2\}.$$

Définition 11.39 ($n = 3$). — On appelle *sphère de centre* $A = (a, b, c)$ et de rayon $r \geq 0$ l'ensemble des points M de \mathbb{R}^3 défini par :

$$\mathcal{S}(A, r) = \mathcal{S}((a, b, c), r) = \{M = (x, y, z) \in \mathbb{R}^3 \mid (x - a)^2 + (y - b)^2 + (z - c)^2 = r^2\}.$$

Remarque 11.40. — Si $r = 0$, ces ensembles sont réduits à un seul point : leur centre A . Si $r < 0$ ces ensembles sont vides.

Dans la suite du cours, il faudra savoir reconnaître l'équation d'un cercle ou d'une sphère.

Exemple 11.41. — L'équation $x^2 + y^2 - 2x + 4y - 4 = 0$ s'écrit aussi $(x - 1)^2 + (y + 2)^2 = 9$. C'est donc l'équation du cercle de centre $(1, -2)$ et de rayon 3.

CHAPITRE 12

TOPOLOGIE DE L'ENSEMBLE \mathbb{R}^2

Comment généraliser à \mathbb{R}^2 les notions d'intervalle ouvert, d'intervalle fermé, de voisinage d'un point ?

Dans ce chapitre nous allons, grâce aux notions de norme et de distance, construire les ensembles qui vont nous permettre de faire de l'analyse dans \mathbb{R}^2 .

Dans ce qui suit, un couple de réels, élément de \mathbb{R}^2 , sera plutôt considéré comme un point.

12.1. Boules

Soit a un réel. Dans \mathbb{R} , l'intervalle $I =]a - \alpha, a + \alpha[$ ($\alpha > 0$) est un intervalle ouvert de centre a . Remarquons que l'on peut écrire :

$$I =]a - \alpha, a + \alpha[= \{x \in \mathbb{R} \mid |x - a| < \alpha\}.$$

Or nous avons vu que, dans \mathbb{R}^2 , la norme euclidienne remplaçait la valeur absolue, ce qui permet de généraliser la notion d'intervalle de centre a aux dimensions supérieures.

12.1.1. Définitions. —

Définition 12.1. — Soient r un réel positif et $A \in \mathbb{R}^2$.

On appelle *boule ouverte de centre A et de rayon r* ($r \geq 0$), notée $\mathcal{B}(A, r)$, le sous-ensemble :

$$\mathcal{B}(A, r) = \{M \in \mathbb{R}^2 \mid d(A, M) < r\} = \{M \in \mathbb{R}^2 \mid \|\vec{AM}\| = \|M - A\| < r\}.$$

On appelle *boule fermée de centre A et de rayon r* ($r \geq 0$), notée $\overline{\mathcal{B}}(A, r)$, le sous-ensemble :

$$\overline{\mathcal{B}}(A, r) = \{M \in \mathbb{R}^2 \mid d(A, M) \leq r\} = \{M \in \mathbb{R}^2 \mid \|\vec{AM}\| = \|M - A\| \leq r\}.$$

Remarque 12.2. — (i) si $r = 0$ alors $\mathcal{B}(A, 0)$ est vide et $\overline{\mathcal{B}}(A, 0) = \{A\}$.

(ii) Dans \mathbb{R} , si a est un réel, et $r > 0$, on retrouve bien :

- la boule ouverte $\mathcal{B}(a, r)$ est l'intervalle ouvert $]a - r, a + r[$.
- la boule fermée $\overline{\mathcal{B}}(a, r)$ est l'intervalle fermé $[a - r, a + r]$.

Représentation graphique. — Soient $A = (a, b)$, $M = (x, y)$ et $r > 0$. On a :

$$d(A, M) = \sqrt{(x - a)^2 + (y - b)^2}$$

donc

$$M \in \overline{\mathcal{B}}(A, r) \iff (x - a)^2 + (y - b)^2 \leq r^2.$$

C'est-à-dire : M appartient au *disque* de centre A et de rayon r , cercle compris. La boule ouverte $\mathcal{B}(A, r)$ est l'intérieur du disque : on exclut les points du cercle qui vérifient $(x - a)^2 + (y - b)^2 = r^2$.

Conclusion : Une boule de \mathbb{R}^2 est représentée par un disque limité par un cercle. Si la boule est ouverte les points du cercle sont exclus. Si la boule est fermée les points du cercle sont compris. Le cercle est la *frontière* du disque.

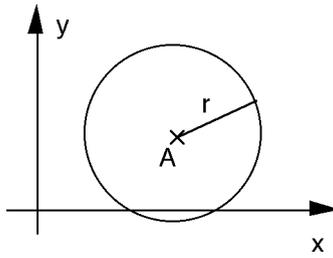


FIGURE 12.1. Disque du plan

12.2. Frontière d'un sous-ensemble

Rappel : Si \mathcal{D} est une partie de \mathbb{R}^2 , on appelle complémentaire de \mathcal{D} , le sous-ensemble noté \mathcal{D}^c défini par :

$$\mathcal{D}^c = \mathbb{R}^2 \setminus \mathcal{D} = \{(x, y) \in \mathbb{R}^2 \mid (x, y) \notin \mathcal{D}\}.$$

Définition 12.3. — Soit \mathcal{D} un sous-ensemble de \mathbb{R}^2 , $\mathcal{D} \neq \mathbb{R}^2$. On dit que le point A est une *point-frontière* de \mathcal{D} si toute boule ouverte de centre A , de rayon non nul, rencontre à la fois \mathcal{D} et son complémentaire. La *frontière* de \mathcal{D} , notée $\mathcal{F}r(\mathcal{D})$, est l'ensemble de ses points-frontière.

Le complémentaire de \mathbb{R}^2 est l'ensemble vide. Par suite, pour tout A de \mathbb{R}^2 , toute boule ouverte de centre A , de rayon non nul, ne peut avoir de point commun avec le complémentaire de \mathbb{R}^2 . Autrement dit, \mathbb{R}^2 n'a pas de frontière.

Deux sous-ensembles complémentaires, différents de \mathbb{R}^2 et de l'ensemble vide \emptyset , ont la même frontière. Dans la figure 12.2, les points B_1 et B_2 sont des points-frontière de \mathcal{D} . B_1 appartient à \mathcal{D} mais pas B_2 .

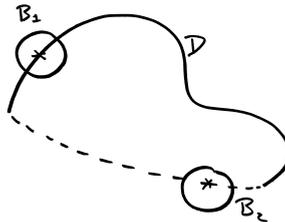


FIGURE 12.2. Frontière d'un ensemble

12.3. Ouverts, fermés

12.3.1. Définitions. —

Définition 12.4. — Soit \mathcal{D} une partie de \mathbb{R}^2 , $\mathcal{D} \neq \mathbb{R}^2$.

\mathcal{D} est une *partie ouverte* ou un *ouvert* si \mathcal{D} est vide ou si \mathcal{D} ne contient aucun point de sa frontière.

\mathcal{D} est une *partie fermée* ou un *fermé* si son complémentaire est un ouvert, donc si \mathcal{D} contient tous les points de sa frontière.

Remarque 12.5. — Si \mathcal{D} est un ouvert, pour tout point A de \mathcal{D} , il existe une boule ouverte de centre A et de rayon r ($r > 0$) contenue dans \mathcal{D} :

$$\mathcal{D} \text{ est un ouvert non vide} \iff \forall A \in \mathcal{D}, \exists r > 0, \mathcal{B}(A, r) \subset \mathcal{D}.$$

En particulier, si A est un point d'un ouvert \mathcal{D} , alors tous les points très "proches" de A (au voisinage de A) appartiennent encore à \mathcal{D} . On peut toujours se déplacer d'une faible distance à partir de A , dans toutes les directions, tout en restant dans \mathcal{D} .



De nombreuses parties de \mathbb{R}^2 ne sont ni ouvertes ni fermées, comme l'illustre l'exemple 12.6 ci-dessous.

Exemple 12.6. — Soit $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 4 \text{ et } x + y > 0\}$. Il suffit de faire un dessin pour constater que \mathcal{D} contient certains points de sa frontière, par exemple $(2, 0)$, donc \mathcal{D} n'est pas un ouvert. Mais \mathcal{D} ne contient pas tous ses points-frontière (par exemple $(0, 0)$) donc \mathcal{D} n'est pas fermé non plus.

12.3.2. Intérieur d'un ensemble. —

Définition 12.7. — Soit \mathcal{D} une partie quelconque de \mathbb{R}^2 . Un point A de \mathcal{D} est dit *point intérieur* de \mathcal{D} s'il existe une boule ouverte de centre A et de rayon r ($r > 0$) contenue dans \mathcal{D} .
On appelle *intérieur de \mathcal{D}* l'ensemble de tous les points intérieurs de \mathcal{D} . On le note $\overset{\circ}{\mathcal{D}}$.

Soit $\mathcal{D} \neq \mathbb{R}^2$. Un point A de \mathcal{D} n'est pas un point intérieur de \mathcal{D} , si toute boule ouverte de centre A et de rayon r ($r > 0$) n'est pas entièrement contenue dans \mathcal{D} , c'est-à-dire si son intersection avec le complémentaire de \mathcal{D} est non vide.

Autrement dit : les points intérieurs de \mathcal{D} sont les points de \mathcal{D} qui ne sont pas sur la frontière.

On montre, et nous l'admettrons, que l'ensemble $\overset{\circ}{\mathcal{D}}$ est un ouvert. C'est le plus grand ouvert contenu dans \mathcal{D} . Par suite, \mathcal{D} est un ouvert si et seulement si $\mathcal{D} = \overset{\circ}{\mathcal{D}}$.

Exemple 12.8. — La boule ouverte $\mathcal{B}(A, r)$ est l'intérieur de la boule fermée $\overline{\mathcal{B}}(A, r)$.

12.3.3. Catalogue. — Notons que dans les exercices la nature topologique des ensembles sera toujours précisée.

Le catalogue suivant n'est pas à retenir mais aidera à comprendre la nature topologique des exemples rencontrés (voir exercices).

- (i) \mathbb{R}^2 et \emptyset sont les seuls ensembles ouverts et fermés.
- (ii) Une boule ouverte est une partie ouverte.
- (iii) Une boule fermée est une partie fermée.
- (iv) Une partie réduite à un point est fermée car on peut écrire $\{A\} = \overline{\mathcal{B}}(A, 0)$.
- (v) \mathbb{R}^2 privé d'un point est un ouvert (puisque son complémentaire est fermé).
- (vi) Toute partie, définie à l'aide de fonction(s) continue(s) (par exemple un ou des polynôme(s)) et par une ou des inégalités strictes est une partie ouverte.
- (vii) Toute partie, définie à l'aide de fonction(s) continue(s) et par une ou des inégalités larges (ou des égalités) est une partie fermée. En particulier :
 - Tout demi-plan de \mathbb{R}^2 défini par une inégalité stricte est un ouvert.
 - Tout demi-plan de \mathbb{R}^2 défini par une inégalité large est un fermé.
 - Toute droite de \mathbb{R}^2 est un fermé.
 - Tout cercle de \mathbb{R}^2 est un fermé.
- (viii) Le produit cartésien de deux intervalles ouverts (resp. fermés) est un ouvert de \mathbb{R}^2 (resp. fermé).

12.4. Sous ensembles bornés - Compacts

12.4.1. Sous-ensembles bornés. —

Définition 12.9. — Une partie \mathcal{D} de \mathbb{R}^2 est une *partie bornée* ou un *sous-ensemble borné* si l'une des deux conditions équivalentes suivantes est réalisée :

- (i) On peut inclure \mathcal{D} dans une boule (fermée ou ouverte)
- (ii) Il existe $m_1 > 0$ et $m_2 > 0$, tels que pour tout $M = (x, y) \in \mathcal{D}$, $|x| \leq m_1$ et $|y| \leq m_2$.

Preuve. — On va montrer que (i) implique (ii) puis que (ii) implique (i).

- (i) \Rightarrow (ii) Supposons donc $\mathcal{D} \subset \overline{\mathcal{B}}(A, r)$ (le raisonnement est le même avec la boule ouverte). On pose $A = (a, b)$. Alors pour tout M dans \mathcal{D} , $d(A, M) \leq r$, c'est-à-dire

$$\sqrt{(x-a)^2 + (y-b)^2} \leq r.$$

Or $(x-a)^2 \leq (x-a)^2 + (y-b)^2$ puisqu'un carré est toujours positif ou nul, donc

$$|x-a| = \sqrt{(x-a)^2} \leq \sqrt{(x-a)^2 + (y-b)^2} \leq r$$

puisque la racine carrée est croissante sur \mathbb{R}^+ . Comme

$$|x-a| \geq ||x| - |a|| \geq |x| - |a|,$$

- on en déduit que $|x| \leq |a| + r$. On trouve donc la borne sur $|x|$ dans (ii) en posant $m_1 = |a| + r$. De même, en partant de $|y - b| \leq r$, on trouve que $|y| \leq m_2$ en posant $m_2 = r + |b|$.
- (ii) \Rightarrow (i) Supposons qu'il existe $m_1 > 0$ et $m_2 > 0$ tels que pour tout $M = (x, y) \in \mathcal{D}$, $|x| \leq m_1$ et $|y| \leq m_2$. Posons $m = \max(m_1, m_2)$. Alors

$$x^2 \leq m^2 \quad \text{et} \quad y^2 \leq m^2$$

ce qui donne

$$x^2 + y^2 \leq 2m^2.$$

En passant aux racines carrées, on obtient : $\sqrt{x^2 + y^2} \leq \sqrt{2}m$ avec $m > 0$. Donc \mathcal{D} est inclus dans la boule de centre O et de rayon $\sqrt{2}m$. □

Exemple 12.10. — En particulier toute boule (ouverte ou fermée) ainsi que sa frontière (cercle) sont des parties bornées puisque ces sous-ensembles sont inclus dans la boule fermée de même rayon.

Remarque 12.11. — D'après (ii) une partie \mathcal{D} est bornée si et seulement si les deux coordonnées des points de \mathcal{D} sont bornées dans \mathbb{R} (donc on peut inclure \mathcal{D} dans un rectangle).

Exemple 12.12. — $E = [1, 2] \times]-1, 5]$ est un borné de \mathbb{R}^2 car pour tout point $M = (x, y) \in E$, on a $|x| \leq 2$ et $|y| \leq 5$.

Dans les exercices pour montrer qu'un ensemble n'est pas borné, on s'appuiera sur le résultat suivant :

Proposition 12.13. — Une demi-droite (donc une droite) du plan n'est pas un ensemble borné. En particulier, tout sous-ensemble de \mathbb{R}^2 qui contient une demi-droite n'est pas borné.

Preuve. — On considère la demi-droite \mathcal{D} d'extrémité le point $M_0 = (x_0, y_0)$ et de vecteur directeur $v = (a, b)$ avec $(a, b) \neq (0, 0)$. \mathcal{D} est alors l'ensemble des points $M = M_0 + tv$ avec t décrivant \mathbb{R}^+ . Donc tous les points de coordonnées (x_n, y_n) avec $n \in \mathbb{N}$ de la forme

$$\begin{cases} x_n &= x_0 + an, \\ y_n &= y_0 + bn, \end{cases}$$

appartiennent à la demi-droite. La demi-droite n'est pas bornée car l'une au moins des deux suites de coordonnées est une suite non bornée (tendant vers l'infini lorsque n tend vers l'infini). □

12.4.2. Sous-ensembles compacts. —

Définition 12.14. — Une partie \mathcal{D} de \mathbb{R}^2 est une partie compacte ou un compact de \mathbb{R}^2 si \mathcal{D} est un ensemble fermé et borné.

Proposition 12.15. — Toute boule fermée et tout cercle sont des compacts de \mathbb{R}^2 .

Preuve. — En effet, d'une part, d'après l'exemple 12.10 les boules fermées et les cercles sont des ensembles bornés. D'autre part, d'après le catalogue, toute boule fermée est un fermé et tout cercle est un fermé. □

CHAPITRE 13

PARTIES CONVEXES DE \mathbb{R}^2

Nous verrons au chapitre 21 la définition des fonctions convexes ou concaves. Ces fonctions sont très importantes en économie. Le domaine de définition d'une telle fonction ne peut pas être une partie quelconque de \mathbb{R}^2 mais devra vérifier les propriétés introduites dans ce chapitre.

13.1. Segment dans \mathbb{R}^2

Considérons la droite passant par deux points distincts A et B de \mathbb{R}^2 . Pour tout point M de la droite $(AB) = \mathcal{D}(A, B)$ (voir Proposition 11.27 page 47) il existe un réel t tel que

$$(13.1) \quad M = A + t\overrightarrow{AB} = (1-t)A + tB.$$

Le segment joignant les points A et B est l'ensemble des points de la droite (AB) qui sont compris entre A et B . On le note $[A, B]$. On a d'ailleurs $[A, B] = [B, A]$.

Si $t = 0$ dans l'équation (13.1), on trouve le point $M = A$, si $t = 1$ on trouve le point $M = B$. Donc si t est compris entre 0 et 1, le point M sera entre A et B . Ceci nous conduit à la définition suivante :

Définition 13.1. — Soient A et B deux points de \mathbb{R}^2 . On appelle *segment* $[A, B]$ l'ensemble des points M de \mathbb{R}^2 pouvant s'écrire sous la forme : $M = (1-t)A + tB$ avec $t \in [0, 1]$.

Comme $[A, B] = [B, A]$, on peut échanger les rôles joués par A et B .

13.2. Sous-ensembles convexes de \mathbb{R} et \mathbb{R}^2

Dans la figure 13.1, l'ensemble \mathcal{D}_1 sera dit convexe par contre l'ensemble \mathcal{D}_2 n'est pas convexe.

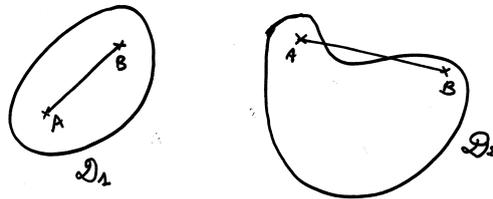


FIGURE 13.1. \mathcal{D}_1 est convexe, \mathcal{D}_2 ne l'est pas

Définition 13.2. — Un sous-ensemble \mathcal{C} de \mathbb{R}^n ($n = 1$ ou $n = 2$) est dit *convexe* s'il est vide ou si pour tout couple (A, B) de points de \mathcal{C} le segment $[A, B]$ est entièrement contenu dans \mathcal{C} . \mathcal{C} est-à-dire :

$$\forall A \in \mathcal{C}, \forall B \in \mathcal{C}, \forall t \in [0, 1], \quad (1-t)A + tB \in \mathcal{C}.$$

Remarque 13.3. — Si $A = B$, pour tout $t \in [0, 1]$, $tA + (1-t)A = A$. Donc $[A, A] = \{A\}$. Un sous-ensemble réduit à un point est donc convexe.

Méthode 13.4. — Pour montrer qu'un sous-ensemble \mathcal{D} de \mathbb{R}^2 n'est pas convexe on exhibe deux points A et B de \mathcal{D} dont le milieu $\frac{A+B}{2} = \frac{1}{2}A + \frac{1}{2}B$ n'est pas dans \mathcal{D} .

La proposition suivante liste les ensembles convexes à connaître.

Proposition 13.5 (Exemples d'ensembles convexes). — (i) \mathbb{R} et \mathbb{R}^2 sont convexes.

(ii) Les seuls convexes de \mathbb{R} sont les intervalles.

(iii) Toute droite ou tout segment de \mathbb{R}^2 est convexe.

(iv) Toute boule de \mathbb{R}^2 (ouverte ou fermée) est convexe.

(v) Tout sous-ensemble réduit à un point est convexe.

(vi) Tout demi-plan de \mathbb{R}^2 est convexe.

(vii) Si I et J sont deux intervalles (non vides) de \mathbb{R} alors leur produit cartésien $I \times J$ est un convexe de \mathbb{R}^2 . En particulier les rectangles sont des ensembles convexes.

Preuve. — (i) et (iii) sont évidents avec la définition.

Preuve de (ii). D'après la propriété d'un intervalle donnée page 1, on a l'équivalence :

$$I \text{ est intervalle non vide de } \mathbb{R} \Leftrightarrow (\forall a \in I, \forall b \in I, a \leq b \Rightarrow [a, b] \subset I)$$

L'implication \Rightarrow signifie qu'un intervalle est un convexe de \mathbb{R} .

L'implication \Leftarrow signifie qu'un convexe de \mathbb{R} est un intervalle.

C'est à cause de cette propriété des intervalles de \mathbb{R} , que nous avons supposé que les fonctions convexes ou concaves d'une variable étaient toujours définies sur un intervalle.

Preuve de (iv). Raisonnons pour la boule fermée. Soit \mathcal{B} la boule fermée de centre A et de rayon r . Si $r = 0$, $\mathcal{B} = \{A\}$ est convexe d'après (iv). Si $r > 0$, soient $C \in \mathcal{B}$, $D \in \mathcal{B}$ et $0 < t < 1$. Montrons que le point $M = tD + (1-t)C$ appartient à \mathcal{B} . Pour cela, il faut justifier que $d(A, M) = \|\overrightarrow{AM}\| \leq r$. Or, $\overrightarrow{AM} = t\overrightarrow{AD} + (1-t)\overrightarrow{AC}$, donc en utilisant les propriétés de la norme (page 45) on obtient :

$$\|\overrightarrow{AM}\| = \|t\overrightarrow{AD} + (1-t)\overrightarrow{AC}\| \leq t\|\overrightarrow{AD}\| + (1-t)\|\overrightarrow{AC}\| \leq tr + (1-t)r = r.$$

D'où le résultat (on a utilisé $t \geq 0$ et $1-t \geq 0$).

(v) a été traité dans la remarque 13.3. On peut aussi voir un point comme une boule fermée de rayon nul.

Preuve de (vi). Contentons-nous de montrer que le demi-plan \mathcal{E}_1 défini page 49 est convexe. Soient $Y = (y_1, y_2)$ et $Z = (z_1, z_2)$ deux points de \mathcal{E}_1 . Considérons le point M défini par : $M = (x_1, x_2) = tY + (1-t)Z$ avec $0 < t < 1$ et montrons que M appartient à \mathcal{E}_1 , c'est-à-dire montrons que $ax_1 + bx_2 + c > 0$. Pour $i = 1$ ou 2 , on a $x_i = ty_i + (1-t)z_i$ et donc

$$ax_1 + bx_2 + c = a(ty_1 + (1-t)z_1) + b(ty_2 + (1-t)z_2) + c.$$

Comme $c = tc + (1-t)c$, on peut finalement écrire :

$$ax_1 + bx_2 + c = t(ay_1 + by_2 + c) + (1-t)(az_1 + bz_2 + c).$$

Or tous les termes du second membre sont strictement positifs. Donc $M \in \mathcal{E}_1$.

Preuve de (viii). Elle sera faite en exercice. □

À partir de ces ensembles convexes de référence on peut construire d'autres ensembles convexes plus complexes par la propriété suivante :

Proposition 13.6 (Intersection d'ensembles convexes). — L'intersection d'un nombre fini d'ensembles convexes est un ensemble convexe.

Remarque 13.7. — À l'aide de la proposition ci-dessus, on montre que les polygones du plan (triangles, carrés, rectangles, losanges...) sont des ensembles convexes.

Preuve. — Il suffit de montrer que l'intersection de deux convexes est un convexe. Soient \mathcal{C}_1 et \mathcal{C}_2 deux sous-ensembles convexes de \mathbb{R}^2 . Si $\mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset$ alors l'intersection est convexe. Supposons maintenant que $\mathcal{C}_1 \cap \mathcal{C}_2 \neq \emptyset$. Considérons alors deux points A et B de $\mathcal{C}_1 \cap \mathcal{C}_2$. Montrons que pour tout $t \in [0, 1]$ le point $M = tA + (1-t)B$ appartient à $\mathcal{C}_1 \cap \mathcal{C}_2$. Puisque A et B sont dans \mathcal{C}_1 , $M \in \mathcal{C}_1$ puisque \mathcal{C}_1 est convexe. Avec le même raisonnement on prouve que $M \in \mathcal{C}_2$ donc M est bien dans l'intersection. □



Comme nous pouvons le remarquer sur la figure 13.2, la réunion de deux ensembles convexes n'est pas nécessairement un ensemble convexe.

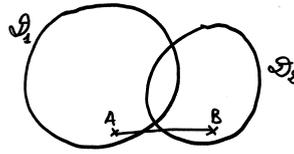


FIGURE 13.2. $\mathcal{D}_1 \cup \mathcal{D}_2$ n'est pas convexe

Finalement, le lien avec les fonctions convexes sur un intervalle est le suivant :

Définition 13.8 (Épigraphe d'une fonction). — Soit f une fonction définie sur un intervalle I de \mathbb{R} à valeurs dans \mathbb{R} . On appelle *épigraphe de f* , et on le note $\text{epi}(f)$ le sous-ensemble de \mathbb{R}^2 défini par :

$$\text{epi}(f) = \{(x, y) \in I \times \mathbb{R} \mid y \geq f(x)\}.$$

Proposition 13.9 (Épigraphe d'une fonction convexe). — Soit f une fonction définie sur un intervalle I de \mathbb{R} à valeurs dans \mathbb{R} . Alors f est une fonction convexe sur I si et seulement si son épigraphe est un sous-ensemble convexe de \mathbb{R}^2 .

Exemple 13.10. — L'ensemble $\{(x, y) \in \mathbb{R}^2 \mid y - x^2 \geq 0\}$ est un ensemble convexe du plan.

Il faudra parfois revenir à la définition pour montrer qu'un sous-ensemble de \mathbb{R}^2 est convexe. On pourra ainsi montrer à titre d'exercice que l'ensemble

$$\{(x, y) \in \mathbb{R}^2 \mid x - y^2 \geq 0\}$$

est convexe.

CHAPITRE 14

FONCTIONS RÉELLES DE DEUX VARIABLES RÉELLES

Dans tout ce chapitre, nous considérons plutôt les couples de \mathbb{R}^2 comme des points.

14.1. Définitions

14.1.1. Fonctions de deux variables. —

Définition 14.1. — On appelle *fonction réelle de deux variables réelles*, toute application f définie sur une partie \mathcal{D}_f de \mathbb{R}^2 , qui à tout couple (x, y) de \mathcal{D}_f associe **un nombre réel unique** noté $f(x, y)$. \mathcal{D}_f s'appelle *le domaine de définition* de la fonction f .

Si le domaine de définition \mathcal{D}_f n'est pas précisé, c'est le plus grand sous-ensemble de \mathbb{R}^2 sur lequel l'expression donnant $f(x, y)$ a un sens. Parfois, on ne prend pas pour \mathcal{D}_f le sous-ensemble précédent. Dans ce cas, le domaine de définition \mathcal{D}_f est donné de façon explicite avec la définition de la fonction. En particulier, dans des problèmes de nature économique, on supposera que les variables x et y sont positives ou nulles.

Exemple 14.2. — $f_1(x, y) = \frac{1}{x - y}$, $\mathcal{D}_{f_1} = \{(x, y) \in \mathbb{R}^2 \mid x \neq y\}$.

Exemple 14.3. — $f_2(x, y) = \sqrt{1 - x^2 - y^2}$, $\mathcal{D}_{f_2} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$.

Fonctions partielles. —

Définition 14.4. — Soit f une fonction définie sur \mathcal{D}_f et soit $A = (a, b) \in \mathcal{D}_f$. On associe à f deux fonctions d'une variable, appelées *fonctions partielles de f en A* , définies par :

$$f_A^1(x) = f(x, b) \text{ pour tout réel } x \text{ tel que } (x, b) \in \mathcal{D}_f$$

$$f_A^2(y) = f(a, y) \text{ pour tout réel } y \text{ tel que } (a, y) \in \mathcal{D}_f.$$

Dans chaque fonction partielle de f en A on a fixé une des variables.



Les fonctions partielles de f dépendent du point A considéré et donnent une information incomplète sur la fonction f .

Exemple 14.5. — $f(x, y) = x^2 + y$ et considérons les points $A = (1, 2)$ et $B = (-1, -2)$. Alors

$$f_A^1(x) = x^2 + 2, \quad f_A^2(y) = 1 + y, \quad f_B^1(x) = x^2 - 2, \quad f_B^2(y) = 1 + y.$$

14.1.2. Graphe d'une fonction de deux variables. —

Définition 14.6. — Soit f une fonction définie sur \mathcal{D}_f .

On appelle *graphe de f* et on le note \mathcal{G}_f le sous-ensemble de \mathbb{R}^3 défini par :

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

Si $\mathcal{A} \subset \mathcal{D}_f$, on appelle *image de \mathcal{A}* le sous-ensemble de \mathbb{R} défini par

$$f(\mathcal{A}) = \{f(x, y) \mid (x, y) \in \mathcal{A}\}.$$

Exemple 14.7. — $f(x, y) = 4$, $\mathcal{D}_f = \mathbb{R}^2$, \mathcal{G}_f est le plan horizontal d'équation $z = 4$.

Suite de l'exemple 14.3 : \mathcal{G}_{f_2} est la demi-sphère supérieure de centre O et de rayon 1.

14.2. Exemples fondamentaux

14.2.1. Exemples économiques. — Les exemples économiques, vus page 3 du Chapitre 1, se généralisent en considérant selon le cas : deux facteurs de production (fonction de production ou de coût ...), deux biens (fonction d'utilité...), deux prix (fonction de demande). Ces modèles sont d'ailleurs plus réalistes que ceux à une seule variable.

14.2.2. Exemples mathématiques. —

14.2.2.1. *Fonction affine.* —

Définition 14.8. — On appelle *fonction affine* toute application de \mathbb{R}^2 dans \mathbb{R} de la forme : $f(x, y) = ax + by + c$ où a, b et c sont trois réels fixés. Si $c = 0$, la fonction est dite *linéaire*.

Alors $\mathcal{D}_f = \mathbb{R}^2$ et \mathcal{G}_f est le plan de \mathbb{R}^3 d'équation : $z = ax + by + c$.

14.2.2.2. *Projections.* —

Définition 14.9. — On appelle *projection* de \mathbb{R}^2 sur \mathbb{R} , les deux applications définies par :

$$\pi_1 : (x, y) \mapsto x, \quad \pi_2 : (x, y) \mapsto y.$$

Les projections sont des fonctions linéaires donc affines.

14.2.2.3. *Fonctions polynômes.* — Ces fonctions contiennent exclusivement des expressions de la forme $ax^n, by^m, cx^p y^q$ avec n, m, p, q entiers positifs ou nuls, a, b, c réels fixés et des constantes.

Exemple 14.10. — La fonction $f(x, y) = 3x^4 - 5x^3y^2 + y^6 + x - 7y + 10$ est un polynôme de degré 6. Une fonction affine est une fonction polynôme de degré 1. La fonction $f(x, y) = xy - \sqrt{x} + 3$ n'est pas un polynôme.

14.3. Relation d'ordre

14.3.1. Sur les images $f(X)$. — Si f est une fonction de deux variables x et y , et si on pose $X = (x, y)$, les images $f(X)$ sont des réels. Donc en conservant la relation d'ordre sur \mathbb{R} , les notions de fonction majorée, minorée, bornée subsistent en remplaçant dans les définitions vues pour les fonctions d'une variable (voir Chapitre 1 page 4) " x " réel par " X " point de \mathbb{R}^2 .

On peut de même définir l'égalité $f = g$ ou les inégalités :

$$f \leq g, \quad f < g, \quad f \geq g, \quad f > g, \quad f \geq 0, \quad f \leq 0, \quad f > 0, \quad f < 0.$$

14.3.2. Sur les variables X . — Il n'y a pas de relation d'ordre semblable à celle définie dans \mathbb{R} entre les points de \mathbb{R}^2 . On peut éventuellement classer les n -uplets comme des "mots" (ordre lexicographique) en disant que $(1, 2) \leq (2, -1)$, mais cet ordre ne respecte pas l'ordre sur les "composantes". Donc les notions de fonctions monotones disparaissent.

Par suite, la méthode d'étude d'une fonction d'une variable utilisant un tableau de variations ne se généralise pas aux fonctions de deux variables. On peut cependant étudier la monotonie des applications partielles qui sont des fonctions d'une seule variable.

14.4. Représentations graphiques

14.4.1. Du domaine de définition. — Le domaine de définition \mathcal{D}_f d'une fonction f de deux variables est un sous-ensemble de \mathbb{R}^2 . Pour le représenter "géométriquement" (en choisissant un repère orthonormé du plan) nous aurons souvent à effectuer des "partages du plan" en utilisant la propriété suivante :

Considérons l'équation $g(x, y) = c$ (c constante réelle). Alors le plan est divisé en trois zones :

- une courbe pour laquelle tous les points $M = (x, y)$ vérifient $g(x, y) = c$,
- une zone du plan dans laquelle tous les points $M = (x, y)$ vérifient $g(x, y) > c$,
- une zone du plan dans laquelle tous les points $M = (x, y)$ vérifient $g(x, y) < c$.

Les deux zones sont délimitées par la courbe. Pour savoir dans quelle zone on se trouve, il suffit de prendre un point particulier $M_0 = (x_0, y_0)$ et de comparer $g(x_0, y_0)$ avec c .

Ceci est illustré par les exemples ci-dessus

Exemple 14.11. — $\mathcal{D}_{f_1} = \{(x, y) \in \mathbb{R}^2 \mid x \neq y\}$ est le plan privé de la droite d'équation $y = x$.
 $\mathcal{D}_{f_2} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ est le disque fermé de centre O et de rayon 1.

Du graphe. — Considérons un repère orthonormé de \mathbb{R}^3 d'origine O et d'axes Ox , Oy et Oz . Si la fonction f est "suffisamment régulière" (c'est-à-dire continue— voir Chapitre 15), le graphe de f est représenté par une surface de \mathbb{R}^3 . C'est l'équivalent de la courbe représentative pour les fonctions d'une seule variable réelle. Comme il est souvent difficile d'avoir une idée de l'allure de la surface (pas de méthode systématique d'étude d'après le paragraphe 14.3.2) on étudie ses différentes intersections par des familles de plans parallèles, ce qui permet de mieux la "visualiser".

Des logiciels comme "Maple" ou "Mathematica" donnent une représentation en trois dimensions de ces graphes (voir la figure 14.4.1.2 page 62).

14.4.1.1. *Courbes de niveau.* — Si les plans évoqués ci-dessus sont des plans horizontaux, d'équation $z = k$ où $k \in \mathbb{R}$ (on dit "de côte" k), les projections, sur le plan xOy , des courbes obtenues sont appelées *courbes de niveau*.

Définition 14.12. — Pour tout réel k , on appelle *courbe de niveau k* de la fonction f , l'ensemble \mathcal{C}_k des points $M = (x, y)$ défini par :

$$\mathcal{C}_k = \{(x, y) \in \mathbb{R}^2 \mid (x, y) \in \mathcal{D}_f \text{ et } f(x, y) = k\}.$$

Remarque 14.13. — (i) Une courbe de niveau peut être vide (voir le cas $k < 0$ dans l'Exercice 14.15 page 61).

(ii) D'après la définition, **une courbe de niveau est toujours incluse dans \mathcal{D}_f** .

(iii) Deux courbes de niveaux différents ne se coupent jamais.

(iv) Lorsqu'on sait représenter graphiquement les courbes de niveau (ce n'est pas toujours le cas), la représentation sur un même graphique des différentes courbes de niveau donne une idée "du relief" de la surface représentative de f . Les courbes de niveau sont utilisées sur les cartes géographiques, mais aussi en Économie : courbe d'indifférence pour une fonction d'utilité, isoquantes pour une fonction de coût, etc...

14.4.1.2. *Coupes verticales.* — Au lieu de fixer z , on peut fixer une des deux autres variables, par exemple $x = x_0$. On obtient ainsi une fonction d'une seule variable $g(y) = f(x_0, y)$ dont le graphe (ou courbe représentative) est la section de la surface représentative de \mathcal{G}_f par le plan d'équation $x = x_0$. De même en fixant $y = y_0$. Ces sections verticales ne sont rien d'autre que les graphes des fonctions partielles. Ces différentes "coupes" de \mathcal{G}_f peuvent aider à la reconstitution spatiale du graphe de f : c'est le principe du scanner en imagerie médicale.

Remarque 14.14. — Une droite verticale (parallèle à Oz) coupe la surface représentative de f en au plus un seul point, car pour tout (x, y) la valeur de $f(x, y)$ est unique.

Exemple 14.15. — Soit $f(x, y) = x^2 + y^2$. Alors $\mathcal{D}_f = \mathbb{R}^2$. Les courbes de niveau sont données par :

$$\mathcal{C}_k = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = k\}.$$

– Si $k < 0$, alors \mathcal{C}_k est vide.

– Si $k = 0$, alors \mathcal{C}_k est réduite à l'origine : $\mathcal{C}_k = \{O\}$.

– Si $k > 0$, alors \mathcal{C}_k est un cercle de centre O et de rayon \sqrt{k} .

Fixons $y = y_0$, et considérons la fonction partielle définie par : $x \mapsto f(x, y_0) = x^2 + y_0^2$. La section de \mathcal{G}_f par le plan d'équation $y = y_0$ est la parabole d'équation $z = x^2 + y_0^2$. C'est la courbe représentative de la fonction partielle dans le plan xOz .

Fixons $x = x_0$, et considérons la fonction partielle définie par : $y \mapsto f(x_0, y) = y^2 + x_0^2$. La section de \mathcal{G}_f par le plan d'équation $x = x_0$ est la parabole d'équation $z = y^2 + x_0^2$. C'est la courbe représentative de la fonction partielle dans le plan yOz .

La surface représentative de \mathcal{G}_f s'appelle un parabololoïde de révolution autour de l'axe Oz .

14.5. Opérations algébriques

On définit les opérations algébriques (addition, soustraction, multiplication, division) comme pour les fonctions d'une variable. Toutes les définitions vues pour les fonctions d'une variable (voir Chapitre 1, paragraphe 1.4, page 3) se généralisent en remplaçant x réel par X point de \mathbb{R}^2 . Les domaines de définition des fonctions étudiées sont alors des sous-ensembles de \mathbb{R}^2 .

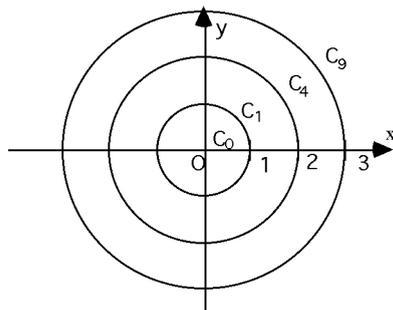
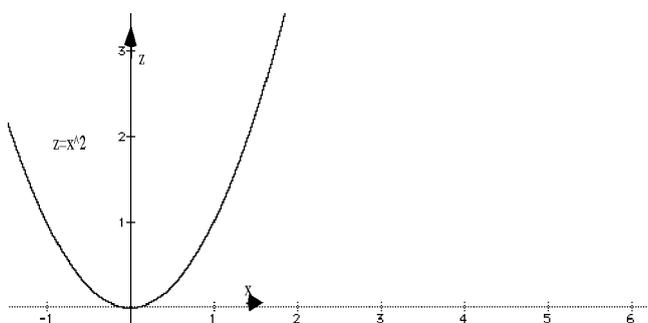
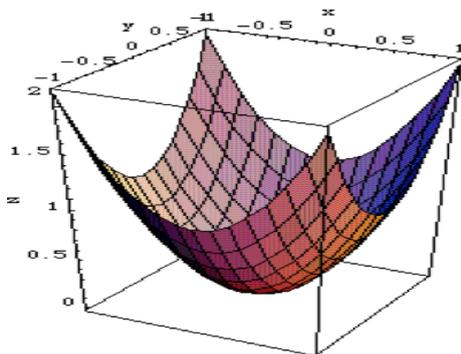


FIGURE 14.1. Courbes de niveau

FIGURE 14.2. Section verticale du graphe de f FIGURE 14.3. Graphe de f

CHAPITRE 15

FONCTIONS CONTINUES

L'objectif de ce chapitre est de donner la définition d'une fonction continue sur un sous-ensemble de \mathbb{R}^2 . Dans les théorèmes des chapitres suivants nous considérerons généralement des fonctions continues sur un ouvert. En optimisation, nous étudierons aussi des fonctions continues sur un fermé.

15.1. Définitions

15.1.1. Continuité en un point. — Nous voulons traduire la même propriété que celle vue pour les fonctions d'une variable (Chapitre 3), à savoir qu'une fonction définie en un point A , est continue en A , si les images des points de \mathcal{D}_f proches de A restent proches de $f(A)$. Autrement dit, f est continue en A si des petites variations de X autour de A entraînent des petites variations de $f(X)$ autour de $f(A)$.

Exemple 15.1. — Soit α un réel quelconque, on considère la fonction constante telle que $f(X) = \alpha$ pour tout $X \in \mathbb{R}^2$. Pour tout $A \in \mathbb{R}^2$ et tout $X \in \mathbb{R}^2$, $f(X) - f(A) = 0$. Donc f est continue en tout point A de \mathbb{R}^2 .

Exemple 15.2 (Fonction définie en $(0,0)$ non continue en $(0,0)$). — Soit f la fonction définie sur \mathbb{R}^2 par :

$$\begin{cases} f(x, y) = \frac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0), \\ f(0, 0) = 0. \end{cases}$$

Montrons que f n'est pas continue en $(0,0)$. Remarquons que $f(x, x) = \frac{1}{2}$ si $x \neq 0$. En particulier $f(10^{-10}, 10^{-10}) = \frac{1}{2}$, pourtant $(10^{-10}, 10^{-10})$ est très proche de $(0,0)$. Mais on ne peut pas rendre $f(10^{-10}, 10^{-10})$ aussi proche de $f(0,0) = 0$ que l'on voudrait.

Toutes les fonctions usuelles d'une variable sont continues sur leur domaine de définition. Or, les fonctions de deux variables sont définies à l'aide des fonctions d'une variable et des opérations algébriques et/ou de la composition. C'est pourquoi, nous ne donnerons pas de définition rigoureuse de la continuité d'une fonction de deux variables, mais nous indiquerons les opérations qui permettent de fabriquer des fonctions continues de deux variables à partir de fonctions continues d'une variable (voir les théorèmes 15.6, 15.7, 15.11, 15.13).

Remarque 15.3 (Propriétés importantes pour la suite du cours). — Si f est continue en A et si $f(A) = 0$, alors la valeur absolue $|f(X)|$ reste très petite si X reste "proche de A ", c'est-à-dire si X reste dans une boule ouverte de centre A et de rayon α "très petit".

Si f est définie sur \mathbb{R}^2 et continue en A avec $f(A) > 0$, alors $f(X)$ ne peut pas brutalement prendre des valeurs négatives ou nulles : $f(X)$ reste strictement positif si X reste dans une boule ouverte de centre A et de rayon $\alpha > 0$, "suffisamment petit". On obtient un résultat analogue si f est définie sur \mathbb{R}^2 et continue en A avec $f(A) < 0$.

Nous admettrons le théorème suivant qui traduit ce qui vient d'être dit :

Théorème 15.4. — Soient f une fonction définie sur \mathbb{R}^2 tout entier et A un point de \mathbb{R}^2 .

- (i) (f continue en A et $f(A) > 0$) \Rightarrow ($\exists \alpha > 0$ tel que $\forall X \in \mathcal{B}(A, \alpha)$, $f(X) > 0$).
- (ii) (f continue en A et $f(A) < 0$) \Rightarrow ($\exists \alpha > 0$ tel que $\forall X \in \mathcal{B}(A, \alpha)$, $f(X) < 0$).

Continuité sur un sous-ensemble de \mathcal{D}_f . —

Définition 15.5. — Soit f une fonction définie sur $\mathcal{D}_f \subset \mathbb{R}^2$, f est continue sur \mathcal{D}_f si f est continue en tout point de \mathcal{D}_f .

Si f est continue sur \mathcal{D}_f tout entier, alors f est continue sur toute partie \mathcal{E} (ouverte, fermée, quelconque) incluse dans \mathcal{D}_f .

Mais on peut avoir une situation plus délicate : une fonction f peut ne pas être continue sur \mathcal{D}_f et on peut vouloir l'étudier seulement sur une partie \mathcal{E} strictement incluse dans \mathcal{D}_f . On étudiera alors la fonction g définie sur \mathcal{E} par : $\forall x \in \mathcal{E}, g(x) = f(x)$. La fonction g s'appelle la *restriction de f à \mathcal{E}* . On dira que f est continue sur \mathcal{E} si g est continue en tout point de \mathcal{E} .

Théorème 15.6 (Exemples fondamentaux). — La norme euclidienne et les 2 projections $\pi_1 : (x, y) \mapsto x$ et $\pi_2 : (x, y) \mapsto y$ sont continues sur \mathbb{R}^2 tout entier.

Nous admettrons le théorème 15.6 mais nous donnons deux indications pour mieux le comprendre.

- Posons $f(X) = \|X\|$, alors pour tout X et tout A dans \mathbb{R}^2 , $|f(X) - f(A)| \leq \|X - A\|$ (voir (v) dans la proposition 11.17 page 45). En d'autres termes, si la norme $\|X - A\|$ est petite (c'est-à-dire X "proche" de A), alors la différence $f(X) - f(A)$ l'est aussi.
- Posons $X = (x, y)$ et $A = (a, b)$, alors $\pi_1(X) = x$, $\pi_1(A) = a$ et $|\pi_1(X) - \pi_1(A)| = |x - a|$. Or, $(x - a)^2 \leq (x - a)^2 + (y - b)^2$. Donc en passant aux racines carrées (les deux membres de l'inégalité sont des nombres positifs ou nuls), on obtient $|x - a| \leq \|X - A\|$. Par suite :

$$|\pi_1(X) - \pi_1(A)| \leq \|X - A\|,$$

et on a la même situation que ci-dessus. Le raisonnement est exactement le même pour la projection π_2 .

15.2. Opérations algébriques et composition

Dans les théorèmes qui suivent, on envisage seulement la continuité sur un sous-ensemble car en pratique c'est la plus utilisée. Nous admettrons les théorèmes 15.7 et 15.11.

15.2.1. Opérations algébriques. —

Théorème 15.7 (Exemples fondamentaux). — Soient f et g deux fonctions définies et continues sur une partie \mathcal{D} de \mathbb{R}^2 et soit λ un réel quelconque, alors

- (i) $f + g$, $f - g$, λf et $f g$ sont continues sur \mathcal{D} .
- (ii) Si g ne s'annule pas sur \mathcal{D} , alors f/g est continue sur \mathcal{D} .

On sait déjà que toute fonction constante est continue sur \mathbb{R}^2 (voir l'exemple 15.1). On déduit des théorèmes 15.6 et 15.7 les résultats suivants que nous utiliserons régulièrement :

Corollaire 15.8. — (i) Toute fonction polynôme de deux variables, en particulier toute fonction affine, est continue sur \mathbb{R}^2 .
(ii) Toute fraction rationnelle, quotient de polynômes, est une fonction continue en tout point où le dénominateur ne s'annule pas.

En effet, une fonction polynôme s'obtient comme somme et produit de projections comme le montre l'exemple ci-dessous.

Exemple 15.9. — Soit $f(x, y) = x^2 y^3 - 2x + 3y + 4$, alors $f = \pi_1^2 \pi_2^3 - 2\pi_1 + 3\pi_2 + g$ où $g(x, y) = 4$.

15.2.2. Composition à gauche $h \circ g$. — Soit g une fonction de deux variables définie sur \mathcal{D}_g . On peut seulement composer g à gauche par une fonction h d'une seule variable :

$$\mathbb{R}^2 \supset \mathcal{E} \xrightarrow{g} \mathbb{R} \xrightarrow{h} \mathbb{R}.$$

On suppose que la **composition est possible** : c'est-à-dire on se place sur un sous-ensemble \mathcal{E} de \mathcal{D}_g tel que $g(x, y) \in \mathcal{D}_h$ pour tout $(x, y) \in \mathcal{E}$. Ce cas se retrouve très régulièrement lorsqu'on fabrique une fonction de deux variables en la construisant à l'aide des fonctions usuelles d'une seule variable.

Exemple 15.10. — Soit la fonction $f(x, y) = \ln(xy)$ alors $f = h \circ g$ avec $g(x, y) = xy$ et $h(u) = \ln(u)$. $\mathcal{D}_g = \mathbb{R}^2$ mais $\mathcal{D}_h =]0, +\infty[$. Donc la composition est possible sur

$$\mathcal{E} = \{(x, y) \in \mathbb{R}^2 \mid xy > 0\}.$$

Théorème 15.11 (Continuité de la composition à gauche). — Soit $f = h \circ g$ définie sur un sous-ensemble non vide \mathcal{E} de \mathbb{R}^2 . Si g est une fonction continue sur \mathcal{E} et si h est une fonction d'une variable réelle continue sur $g(\mathcal{E}) \subset \mathbb{R}$ alors f est continue sur \mathcal{E} .

On résume le théorème 15.11 par : “la composée de deux fonctions continues est continue”.

Exemple 15.12 (Suite de l'exemple 15.10). — (i) g est une fonction polynôme donc continue sur tout \mathbb{R}^2 et en particulier sur \mathcal{E} .

(ii) De plus

$$(x, y) \in \mathcal{E} \iff g(x, y) = xy > 0,$$

donc $g(\mathcal{E}) =]0, +\infty[$.

(iii) h est continue sur $]0, +\infty[= g(\mathcal{E})$.

Par suite, par application du théorème 15.11, f est continue sur \mathcal{E} .

Ce sont essentiellement les théorèmes 15.7 et 15.11 qui seront utilisés pour justifier la continuité d'une fonction de deux variables. (Voir aussi le lemme d'extension au paragraphe suivant).

15.2.3. Fonctions à variables séparées. —

Lemme 15.13 (Lemme d'extension). — Soit g une fonction d'une seule variable définie et continue sur un intervalle J de \mathbb{R} . Alors

(i) la fonction g_1 définie par : pour tout $(x, y) \in J \times \mathbb{R}$, $g_1(x, y) = g(x)$ est continue sur $J \times \mathbb{R}$,

(ii) la fonction g_2 définie par : pour tout $(x, y) \in \mathbb{R} \times J$, $g_2(x, y) = g(y)$ est continue sur $\mathbb{R} \times J$.

Preuve. — En utilisant les projections π_1 et π_2 , on peut écrire : $g_1 = g \circ \pi_1$ et $g_2 = g \circ \pi_2$. Le résultat s'obtient en appliquant le théorème 15.11. □

Exemple 15.14. — Soit la fonction $f(x) = \sqrt{x}$. La fonction f est continue sur $[0, +\infty[$. Le théorème précédent permet d'en déduire que la fonction g_1 , définie par $g_1(x, y) = \sqrt{x}$, est continue sur $[0, +\infty[\times \mathbb{R}$. De même, la fonction g_2 définie par $g_2(x, y) = \sqrt{y}$ est continue sur $\mathbb{R} \times [0, +\infty[$.

On déduit des théorèmes 15.6, 15.7, 15.11 et 15.7 que toute fonction de deux variables construite à partir de fonctions continues d'une seule variable en utilisant les opérations algébriques et/ou la composition (à gauche) est continue sur son domaine de définition.

Exemple 15.15. — Soit la fonction f définie par $f(x, y) = \frac{\sin(y)}{\sqrt{x}}$. Son domaine de définition est $\mathcal{D}_f = [0, +\infty[\times \mathbb{R}$. Montrons que f est continue sur $]0, +\infty[\times \mathbb{R}$. En effet, soit $g_1(x, y) = \sin(y)$ et $g_2(x, y) = \sqrt{x}$. La fonction $y \mapsto \sin(y)$ est continue sur \mathbb{R} . Par le lemme d'extension 15.13, g_1 continue sur $\mathbb{R} \times \mathbb{R}$ donc sur $]0, +\infty[\times \mathbb{R}$. Par l'exercice précédent g_2 est continue sur $[0, +\infty[\times \mathbb{R}$ donc sur $]0, +\infty[\times \mathbb{R}$. En outre, g_2 ne s'annule pas sur $]0, +\infty[\times \mathbb{R}$. Or f est le quotient de deux fonctions continues sur $]0, +\infty[\times \mathbb{R}$ dont le dénominateur ne s'annule pas sur $]0, +\infty[\times \mathbb{R}$, donc f est continue sur $]0, +\infty[\times \mathbb{R}$.

CHAPITRE 16

DÉRIVÉES PARTIELLES DU PREMIER ORDRE

Dans tout ce chapitre on fait l'hypothèse suivante :

(\mathcal{H}) Soit f une fonction de deux variables, définie sur \mathcal{D}_f . Soit \mathcal{U} un ouvert de \mathbb{R}^2 contenu dans \mathcal{D}_f . (On peut avoir \mathcal{U} strictement inclus dans \mathcal{D}_f ou $\mathcal{U} = \mathcal{D}_f$).

16.1. Définitions

Commençons par l'exemple suivant :

Exemple 16.1. — Soient $f(x, y) = x^2 y^3$ et $M_0 = (x_0, y_0)$. La fonction f est définie sur \mathbb{R}^2 . Considérons les fonctions partielles de f au point M_0

$$f_{M_0}^1(x) = x^2 y_0^3 := g(x) \quad \text{et} \quad f_{M_0}^2(y) = x_0^2 y^3 := h(y).$$

Les fonctions g et h sont des polynômes donc dérivables sur \mathbb{R} et on a :

$$g'(x) = 2x y_0^3, \quad h'(y) = 3x_0^2 y^2, \quad g'(x_0) = 2x_0 y_0^3, \quad h'(y_0) = 3x_0^2 y_0^2.$$

Le nombre réel $g'(x_0)$ s'appelle la dérivée partielle première de f par rapport à x en (x_0, y_0) . Le nombre réel $h'(y_0)$ s'appelle la dérivée partielle première de f par rapport à y en (x_0, y_0) .

Dérivées partielles premières en un point. — Généralisons ce que nous venons de voir dans l'exemple 16.1. Plaçons-nous sous l'hypothèse (\mathcal{H}). Soit $M_0 = (x_0, y_0) \in \mathcal{U}$. Considérons la première fonction partielle $f_{M_0}^1(x) = f(x, y_0)$. Cette fonction ne dépend que de la variable x . Si cette fonction est dérivable en x_0 , on dit que f admet une dérivée partielle première par rapport à x , au point (x_0, y_0) . On définit de même la dérivée partielle première de f par rapport à la deuxième variable y .

Définition 16.2. — Sous l'hypothèse (\mathcal{H}) et si $M_0 = (x_0, y_0) \in \mathcal{U}$, on définit les dérivées partielles premières de f au point M_0 par :

$$\begin{aligned} \frac{\partial f}{\partial x}(x_0, y_0) &= \frac{\partial f}{\partial x}(M_0) := \lim_{x \rightarrow x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \\ \frac{\partial f}{\partial y}(x_0, y_0) &= \frac{\partial f}{\partial y}(M_0) := \lim_{y \rightarrow y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0} = \lim_{h \rightarrow 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} \end{aligned}$$

Autres notations possibles :

$$\frac{\partial f}{\partial x}(x_0, y_0) = f'_x(x_0, y_0), \quad \frac{\partial f}{\partial y}(x_0, y_0) = f'_y(x_0, y_0).$$

Notations de Monge : $p = \frac{\partial f}{\partial x}(x_0, y_0), \quad q = \frac{\partial f}{\partial y}(x_0, y_0).$

Les notations (p, q) de Monge sont pratiques mais ambiguës car elles désignent les dérivées partielles premières en un point sans préciser explicitement le point concerné.

16.1.1. Vecteur gradient. —

Définition 16.3. — Si les deux dérivées partielles premières de f existent au point $M_0 = (x_0, y_0)$, on appelle *gradient de f au point M_0* le vecteur de \mathbb{R}^2 défini par :

$$\text{grad}f(M_0) = \text{grad}f(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right).$$

$\text{grad}f(x_0, y_0)$ se note aussi $\nabla f(x_0, y_0)$ ou $\nabla f(M_0)$. Le symbole " ∇ " se prononce "nabla".

Fonctions dérivées partielles premières. —

Définition 16.4. — Sous l'hypothèse (\mathcal{H}) et si les deux dérivées partielles premières de f existent en tout point (x, y) de \mathcal{U} , on définit les applications suivantes de \mathcal{U} dans \mathbb{R} :

- la fonction dérivée partielle première par rapport à x : $\frac{\partial f}{\partial x} : (x, y) \mapsto \frac{\partial f}{\partial x}(x, y)$,
- la fonction dérivée partielle première par rapport à y : $\frac{\partial f}{\partial y} : (x, y) \mapsto \frac{\partial f}{\partial y}(x, y)$.

On se place sur un ouvert car la notion de dérivées partielles en un point frontière pose problème.

Exemple 16.5 (Méthode pratique). — Soient $f(x, y) = y e^{x^2}$ et $M_0 = (1, 1)$ avec $\mathcal{D}_f = \mathbb{R}^2 = \mathcal{U}$. On dérive f par rapport à x en considérant y comme une constante, et on trouve :

$$\frac{\partial f}{\partial x}(x, y) = 2xy e^{x^2}.$$

De même, on dérive f par rapport à y en considérant x comme une constante, et on trouve :

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

Donc

$$\frac{\partial f}{\partial x}(1, 1) = 2e, \quad \frac{\partial f}{\partial y}(1, 1) = e \quad \text{et} \quad \nabla f(1, 1) = (2e, e).$$

 De même qu'on note $f(x, y) = y e^{x^2}$ et non $f = y e^{x^2}$, il faut écrire comme nous l'avons fait $\frac{\partial f}{\partial x}(x, y) = 2xy e^{x^2}$ et non $\frac{\partial f}{\partial x} = 2xy e^{x^2}$. De même pour $\frac{\partial f}{\partial y}(x, y)$.

Remarque 16.6. — Dans la notation $\frac{\partial f}{\partial x}(x, y)$ les deux " x " ne jouent pas le même rôle. Le symbole " ∂x " constitue un "tout" et signifie qu'on a dérivé par rapport à la première variable. On écrit $\frac{\partial f}{\partial x}(x, y)$ mais pas $\frac{\partial f}{\partial 1}(x, y)$.

16.2. Opérations sur les dérivées partielles

16.2.1. Opérations algébriques. — Les dérivées partielles sont des dérivées de fonctions d'une seule variable, donc les formules sur la dérivée d'une somme, d'un produit, d'un quotient vues pour les fonctions d'une seule variable (voir Chapitre 4 page 19) se transposent aux dérivées partielles.

Composition à gauche. — Soit f une fonction admettant des dérivées partielles du premier ordre en tout point (x, y) d'un ouvert \mathcal{U} de \mathbb{R}^2 . Soit g une fonction définie et dérivable sur un intervalle ouvert I de \mathbb{R} , telle que $f(\mathcal{U}) \subset I$. Posons $\varphi = g \circ f$. Alors φ est une fonction de deux variables définie sur \mathcal{U} . Il s'agit de calculer les dérivées partielles de φ en fonction de celles de f et de la dérivée de g .

Nous n'avons besoin d'aucun nouveau théorème dans ce cas, car nous pouvons calculer les dérivées partielles de φ en utilisant la dérivation des fonctions composées d'une variable, en considérant successivement que y est constante puis que x est constante :

$$(16.1) \quad \frac{\partial \varphi}{\partial x}(x, y) = g'(f(x, y)) \frac{\partial f}{\partial x}(x, y)$$

$$(16.2) \quad \frac{\partial \varphi}{\partial y}(x, y) = g'(f(x, y)) \frac{\partial f}{\partial y}(x, y)$$

16.3. Fonctions de classe C^1

16.3.1. Définition. —

Définition 16.7. — Sous l'hypothèse (\mathcal{H}) , on dit que f est de classe C^1 sur \mathcal{U} , et on note $f \in C^1(\mathcal{U})$, si les deux fonctions dérivées partielles premières de f existent et sont continues sur \mathcal{U} .

Remarque 16.8. — Il est très utile de retenir qu'une fonction de classe C^1 est une fonction dont la surface représentative est "lisse" c'est-à-dire sans point anguleux (voir aussi le corollaire 17.8 page 72 du Chapitre 17).

16.3.2. Opérations algébriques. — Les résultats vus pour les opérations algébriques sur les fonctions d'une seule variable de classe C^1 restent valables. Nous utiliserons dans pratiquement tous les exercices le théorème suivant :

Théorème 16.9. — Soient f et g deux fonctions de classe C^1 sur un ouvert \mathcal{U} de \mathbb{R}^2 et soit λ un réel quelconque, alors les fonctions

- (i) $f + g, f - g, \lambda f$ et fg sont de classe C^1 sur \mathcal{U} .
- (ii) Si g ne s'annule pas sur \mathcal{U} , alors f/g est de classe C^1 sur \mathcal{U} .

Le résultat suivant sera fréquemment utilisé. C'est une conséquence directe du théorème précédent :

Corollaire 16.10 (Exemples importants). — (i) Les fonctions polynômes de deux variables, donc en particulier les fonctions affines et les projections, sont de classe C^1 sur \mathbb{R}^2 .

- (ii) Les fractions rationnelles (c'est-à-dire les quotients de deux polynômes) sont de classe C^1 sur le sous-ensemble de \mathbb{R}^2 où leur dénominateur ne s'annule pas.

Composition à gauche. —

Théorème 16.11. — Soit f une fonction de classe C^1 sur un ouvert \mathcal{U} de \mathbb{R}^2 et soit g une fonction de classe C^1 sur un intervalle I de \mathbb{R} telles que $f(\mathcal{U}) \subset I$, alors la fonction composée $g \circ f$ est de classe C^1 sur \mathcal{U} .

Preuve. — On pose $\varphi = g \circ f$, et on utilise les formules (16.1) et (16.2). On peut donc écrire les égalités suivantes entre fonctions :

$$\frac{\partial \varphi}{\partial x} = (g' \circ f) \frac{\partial f}{\partial x} \quad \text{et} \quad \frac{\partial \varphi}{\partial y} = (g' \circ f) \frac{\partial f}{\partial y}.$$

Si f est de classe C^1 sur \mathcal{U} et g de classe C^1 sur I , alors la fonction $g' \circ f$ est continue sur \mathcal{U} , comme fonction composée (à gauche) de fonctions continues. D'autre part, les fonctions dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont continues sur \mathcal{U} . Par suite, les deux fonctions dérivées partielles premières de φ sont continues sur \mathcal{U} comme produit de fonctions continues sur \mathcal{U} , et on en déduit que la fonction φ est de classe C^1 sur \mathcal{U} . \square

Le corollaire suivant permet, en utilisant en plus le théorème 16.11 de démontrer qu'une fonction f , fabriquée à l'aide de sommes, produits ou quotients de **fonctions usuelles d'une variable de classe C^1** , est de classe C^1 . C'est l'analogie du théorème 15.13 page 65 où on a remplacé "continue" par "de classe C^1 ".

Lemme 16.12 (Lemme d'extension). — Soit g une fonction d'une seule variable définie et de classe C^1 sur un intervalle ouvert J de \mathbb{R} . Alors

- (i) la fonction g_1 définie pour tout $(x, y) \in J \times \mathbb{R}$ par, $g_1(x, y) = g(x)$ est de classe C^1 sur $J \times \mathbb{R}$,
- (ii) la fonction g_2 définie pour tout $(x, y) \in \mathbb{R} \times J$ par $g_2(x, y) = g(y)$ est de classe C^1 sur $\mathbb{R} \times J$.

Preuve. — Les projections sont de classe C^1 sur \mathbb{R}^2 puisque ce sont des fonctions affines. Il suffit alors d'écrire : $g_1 = g \circ \pi_1$ et $g_2 = g \circ \pi_2$ et d'appliquer le théorème 16.11. \square

Exemple 16.13. — Le raisonnement qui suit est un modèle à adapter à chaque fonction rencontrée. La recette est de bien décomposer la fonction à étudier.

Soit la fonction

$$f(x, y) = x^y = e^{y \ln x}.$$

Alors

$$\mathcal{D}_f = \{(x, y) \in \mathbb{R}^2 \mid x > 0\} =]0, +\infty[\times \mathbb{R}$$

est un ouvert comme produit cartésien d'intervalles ouverts.

Posons $g(x) = \ln(x)$, $h(y) = y$, $g_1(x, y) = \ln(x)$, $h_2(x, y) = y$, $\varphi(x, y) = y \ln(x)$ et $F(u) = e^u$.

La fonction g est de classe \mathcal{C}^1 sur $]0, +\infty[$ et la fonction h est de classe \mathcal{C}^1 sur \mathbb{R} ; ce sont des fonctions usuelles d'une variable. On en déduit, par application du lemme d'extension, que :

- la fonction g_1 est de classe \mathcal{C}^1 sur $]0, +\infty[\times \mathbb{R} = \mathcal{D}_f$,
- la fonction h_2 est de classe \mathcal{C}^1 sur \mathbb{R}^2 donc sur \mathcal{D}_f .

De plus, $\varphi = g_1 h_2$ donc φ est de classe \mathcal{C}^1 sur \mathcal{D}_f comme produit de deux fonctions de classe \mathcal{C}^1 sur \mathcal{D}_f .

Enfin, on constate que $f = F \circ \varphi$. La fonction exponentielle est de classe \mathcal{C}^1 sur \mathbb{R} et la fonction φ est de classe \mathcal{C}^1 sur \mathcal{D}_f . En outre $\varphi(\mathcal{D}_f) \subset \mathbb{R}$, on peut donc appliquer le corollaire 16.11, et en déduire que f est de classe \mathcal{C}^1 sur \mathcal{D}_f .

CHAPITRE 17

DÉVELOPPEMENT LIMITÉ D'ORDRE 1 ET DIFFÉRENTIELLE

On généralise dans ce chapitre les notions vues au Chapitre 7 pour les fonctions d'une variable.

Dans tout ce chapitre on fait l'hypothèse :

(\mathcal{H}) Soit f une fonction de deux variables, définie sur \mathcal{D}_f . Soit \mathcal{U} un ouvert de \mathbb{R}^2 contenu dans \mathcal{D}_f . (On peut avoir \mathcal{U} strictement inclus dans \mathcal{D}_f ou $\mathcal{U} = \mathcal{D}_f$).

17.1. Variation absolue et différentielle

17.1.1. Variation absolue ou accroissement. —

Définition 17.1. — Soit f une fonction définie sur un sous-ensemble \mathcal{U} de \mathbb{R}^2 et soit $M_0 = (x_0, y_0) \in \mathcal{U}$. On appelle *accroissement ou variation absolue de f* lorsque les valeurs des variables passent de x_0 à $x_0 + h$ et de y_0 à $y_0 + k$ la valeur de la différence :

$$\Delta f_{M_0}(h, k) = f(x_0 + h, y_0 + k) - f(x_0, y_0).$$

h et k représentent les accroissements (positifs ou négatifs) des variables (voir les exemples 17.3 et 17.4). Ces accroissements sont souvent notés Δx et Δy .

Le couple (h, k) est à considérer comme un vecteur car il représente des variations (voir l'exemple 11.11 page 44).

17.1.2. Différentielle. —

Définition 17.2. — Soit une fonction f de classe \mathcal{C}^1 sur un ouvert \mathcal{U} de \mathbb{R}^2 et soit $M_0 = (x_0, y_0) \in \mathcal{U}$. On appelle *différentielle de f au point M_0* l'application **linéaire** de \mathbb{R}^2 dans \mathbb{R} définie par :

$$df_{M_0} : (h, k) \mapsto h \frac{\partial f}{\partial x}(M_0) + k \frac{\partial f}{\partial y}(M_0) = df_{M_0}(h, k).$$

En notant \vec{h} le vecteur (h, k) on peut écrire aussi :

$$df_{M_0}(h, k) = \langle \nabla f(M_0), \vec{h} \rangle.$$

Exemple 17.3 (Exemple fondamental). — Si f est une fonction affine on a l'égalité :

$$\forall M_0 \in \mathbb{R}^2, \forall (h, k) \in \mathbb{R}^2, \Delta f_{M_0}(h, k) = df_{M_0}(h, k).$$

En effet, supposons : $f(x, y) = ax + by + c$ et $M_0 = (x_0, y_0)$. Alors :

$$\Delta f_{M_0}(h, k) = a(x_0 + h) + b(y_0 + k) + c - (ax_0 + by_0 + c) = ah + bk.$$

Comme le gradient de f en tout point vaut (a, b) on en déduit :

$$df_{M_0}(h, k) = ah + bk.$$

Ceci n'est plus vrai dans le cas général, comme le montre l'exemple suivant.

Exemple 17.4. — Soient $f(x, y) = x^2 y$, $\mathcal{D}_f = \mathbb{R}^2 = \mathcal{U}$ et $M_0 = (x_0, y_0)$. Alors :

$$\Delta f_{M_0}(h, k) = (x_0 + h)^2 (y_0 + k) - x_0^2 y_0 = x_0^2 k + 2x_0 y_0 h + 2x_0 h k + y_0 h^2 + h^2 k.$$

et

$$df_{M_0}(h, k) = 2x_0 y_0 h + x_0^2 k.$$

On choisit $x_0 = y_0 = 10$, $h = 0.1$ et $k = 0.2$, de sorte que

$$\Delta f_{M_0}(0.1, 0.2) = (10.1)^2 \times 10.2 - 1000 = 40.502$$

et $df_{M_0}(0.1, 0.2) = 200 \times 0.1 + 100 \times 0.2 = 40$. On constate cependant que $\Delta f_{M_0}(0.1, 0.2) \simeq df_{M_0}(0.1, 0.2)$.

17.2. Développement limité d'ordre 1

17.2.1. Existence. — Pour généraliser à deux variables le résultat vu pour une variable (Chapitre 9, page 36), on peut “deviner” les modifications à faire. Il y a deux dérivées partielles au lieu d'une seule dérivée et deux accroissements. Comme on ne veut pas faire de jaloux, le résultat suivant, que nous admettrons, semble naturel :

Théorème 17.5. — Soit f une fonction de classe \mathcal{C}^1 sur un ouvert \mathcal{U} de \mathbb{R}^2 et soit $M_0 = (x_0, y_0) \in \mathcal{U}$. Alors il existe une fonction ε , continue sur une boule ouverte \mathcal{B} de centre $(0, 0)$ telle que $\varepsilon(0, 0) = 0$ et telle que pour tout $(h, k) \in \mathcal{B}$ on a :

$$\begin{aligned} f(x_0 + h, y_0 + k) &= f(x_0, y_0) + df_{M_0}(h, k) + \|(h, k)\| \varepsilon(h, k) \\ &= f(x_0, y_0) + h \frac{\partial f}{\partial x}(x_0, y_0) + k \frac{\partial f}{\partial y}(x_0, y_0) + \sqrt{h^2 + k^2} \varepsilon(h, k). \end{aligned}$$

On dit que f admet un développement limité (ou DL) d'ordre 1 au voisinage de M_0 .

La fonction ε est inconnue, on sait seulement qu'elle est continue en $(0, 0)$ avec $\varepsilon(0, 0) = 0$. Donc $\varepsilon(h, k)$ est “petite” si $\|(h, k)\|$ est “petite”.

Le DL d'ordre 1 est donc une **formule locale** qui ne s'utilise qu'au voisinage d'un point, c'est-à-dire pour des valeurs petites de $|h|$ et $|k|$. Il est inutile de préciser la boule ouverte \mathcal{B} .

La quantité $\sqrt{h^2 + k^2} \varepsilon(h, k)$ s'appelle *le reste* du DL d'ordre 1.

Exemple 17.6. — Reprenons l'exemple 17.4 page 71. La fonction f est de classe \mathcal{C}^1 sur \mathbb{R}^2 (les dérivées partielles sont continues sur \mathbb{R}^2 comme produit de fonctions continues sur \mathbb{R}^2). Il existe une boule ouverte \mathcal{B} de centre $(0, 0)$ telle que :

$$\forall (h, k) \in \mathcal{B}, \quad f(1+h, 1+k) = e + 2eh + ek + \sqrt{h^2 + k^2} \varepsilon(h, k)$$

avec ε continue sur \mathcal{B} et $\varepsilon(0, 0) = 0$.

Nous allons déduire du théorème 17.5 une propriété fondamentale pour les fonctions de classe \mathcal{C}^1 .

Nous savons que, pour une fonction d'une variable, l'existence de la dérivée en un point assure la continuité en ce point. Ce résultat n'est plus vrai pour les fonctions de deux variables comme l'illustre l'exemple suivant :

Exemple 17.7. — La fonction f définie sur \mathbb{R}^2 par :

$$\begin{cases} f(x, y) = \frac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0), \\ f(0, 0) = 0. \end{cases}$$

admet des dérivées partielles en $(0, 0)$:

$$\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0.$$

Mais f n'est pas continue en $(0, 0)$ (voir l'exemple 15.2 page 63).

Corollaire 17.8. — Si f admet des dérivées partielles premières **continues** sur un ouvert \mathcal{U} contenant M_0 alors f est continue au point M_0 .

Plus généralement, si f est de classe \mathcal{C}^1 sur un ouvert \mathcal{U} alors f est continue sur \mathcal{U} .

On admettra ce résultat, mais il est facile de voir que $f(x_0 + h, y_0 + k)$ reste très proche de $f(x_0, y_0)$ pour des valeurs petites de $|h|$ et $|k|$ car la valeur de $\varepsilon(h, k)$ reste proche de zéro si h et k sont proches de 0.



L'hypothèse du corollaire 17.8 est cruciale : d'après l'exemple 17.7, ce théorème n'est pas valable si on suppose seulement l'existence des dérivées partielles.

17.2.2. Approximation affine. — La partie "régulière", c'est-à-dire l'expression sans le reste, du DL d'ordre 1 d'une fonction f en un point M_0 est une fonction affine en h et k , d'où la définition :

Définition 17.9. — Soit f une fonction de classe \mathcal{C}^1 sur un ouvert \mathcal{U} de \mathbb{R}^2 et soit $M_0 = (x_0, y_0) \in \mathcal{U}$. On appelle *approximation affine de f au voisinage du point M_0* la fonction affine notée \widehat{f}_{M_0} définie sur tout \mathbb{R}^2 par :

$$\begin{aligned} \forall (h, k) \in \mathbb{R}^2, \widehat{f}_{M_0}(x_0 + h, y_0 + k) &= f(x_0, y_0) + h \frac{\partial f}{\partial x}(x_0, y_0) + k \frac{\partial f}{\partial y}(x_0, y_0) \\ &= f(x_0, y_0) + df_{M_0}(h, k). \end{aligned}$$

Remarque 17.10. — 1. L'approximation affine dépend du point M_0 choisi.

2. Il n'y a pas de notation standard. Nous proposons \widehat{f}_{M_0} qui est simple à utiliser.

3. Si on pose $x = x_0 + h$, $y = y_0 + k$ dans les écritures ci-dessus, on obtient

$$\widehat{f}_{M_0}(x, y) = f(x_0, y_0) + (x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0).$$



Dans l'expression de l'approximation affine \widehat{f}_{M_0} il ne faut pas développer les termes en $(x - x_0)$ et $(y - y_0)$.

Exemple 17.11 (suite de l'exemple 17.4). — Soient $A = (2, -1)$ et $B = (1, 1)$ alors

$$\widehat{f}_A(x, y) = -4 - 4(x - 2) + 4(y + 1) \quad \widehat{f}_B(x, y) = 1 + 2(x - 1) + (y - 1).$$

17.2.3. Calculs approchés. — On peut, avec les définitions du paragraphe 17.1, traduire le théorème 17.5 par :

Corollaire 17.12. — Si f admet des dérivées partielles premières continues sur un ouvert contenant M_0 non simultanément nulles en M_0 alors pour $|h|$ et $|k|$ "petits" $\Delta f_{M_0}(h, k) \simeq df_{M_0}(h, k)$. Autrement dit, la différentielle fournit une valeur approchée (au premier ordre) de l'accroissement.

Les calculs approchés se font en négligeant le reste du développement limité ce qui revient à remplacer la fonction f par son approximation affine au voisinage du point considéré. On a donc

$$f(x_0 + h, y_0 + k) \simeq \widehat{f}_{M_0}(x_0 + h, y_0 + k) \quad \text{ou encore} \quad f(x, y) \simeq \widehat{f}_{M_0}(x, y).$$

Exemple 17.13. — Calcul approché de $(1.01)^{0.98}$. On choisit $f(x, y) = x^y = e^{y \ln x}$ et $M_0 = (1, 1)$. La fonction f est de classe \mathcal{C}^1 sur $]0, +\infty[\times \mathbb{R}$ d'après l'exemple 16.13 page 69 et $f(1, 1) = 1$. En outre,

$$\frac{\partial f}{\partial x}(x, y) = y x^{y-1}, \quad \frac{\partial f}{\partial y}(x, y) = x^y \ln x, \quad \frac{\partial f}{\partial x}(1, 1) = 1, \quad \frac{\partial f}{\partial y}(1, 1) = 0.$$

Alors $f(1.01, 0.98) \simeq \widehat{f}_{(1,1)}(1+0.01, 1-0.02)$ et $\widehat{f}_{(1,1)}(1+0.01, 1-0.02) = 1+0.01 \times 1 = 1.01$. Donc $f(1.01, 0.98) \simeq 1.01$. La valeur obtenue par une calculatrice est $(1.01)^{0.98} = 1.009799023$.

17.2.4. Interprétation géométrique. — Le graphe de l'approximation affine de f au point $M_0 = (x_0, y_0)$ est un plan π_0 passant par le point $P_0 = (x_0, y_0, f(x_0, y_0))$ ayant pour équation $z = \widehat{f}_{M_0}(x, y)$.

Soit \mathcal{S} la surface représentative du graphe de f . Si f est de classe \mathcal{C}^1 sur un ouvert \mathcal{U} contenant M_0 , on montre que toutes les tangentes en P_0 aux différentes courbes tracées sur \mathcal{S} et passant par P_0 , sont justement dans le plan π_0 appelé plan tangent à \mathcal{S} en P_0 .

Théorème 17.14 (Équation du plan tangent). — Soit $f : \mathcal{U} \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 sur \mathcal{U} ouvert et soit $M_0 = (x_0, y_0) \in \mathcal{U}$. Si on note $P_0 = (x_0, y_0, f(x_0, y_0))$, l'équation du plan tangent (π_0) en P_0 à la surface représentative \mathcal{S} du graphe de f est : $z = \widehat{f}_{M_0}(x, y)$, ce qui s'écrit aussi :

$$z - f(x_0, y_0) = (x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0).$$

Le plan π_0 est orthogonal au vecteur $\left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1 \right)$ et passe par P_0 .

Remarque 17.15. — Ce résultat généralise la formule $y - f(x_0) = f'(x_0)(x - x_0)$ qui est l'équation de la tangente au point $(x_0, f(x_0))$ pour les fonctions d'une variable.

Exemple 17.16. — Soient $f(x, y) = x^2 + xy + y^2 - 2x + 1$ et $M_0 = (2, 1)$. L'équation de \mathcal{S} est :

$$z = x^2 + xy + y^2 - 2x + 1.$$

On a :

$$f(2, 1) = 4, \quad \frac{\partial f}{\partial x}(x, y) = 2x + y - 2, \quad \frac{\partial f}{\partial y}(x, y) = x + 2y, \quad \frac{\partial f}{\partial x}(2, 1) = 3, \quad \frac{\partial f}{\partial y}(2, 1) = 4.$$

Donc l'équation du plan tangent π_0 à la surface \mathcal{S} au point $(2, 1, f(2, 1))$ est :

$$z - 4 = 3(x - 2) + 4(y - 1).$$

Il n'est pas utile de développer !

Remarques sur les notions introduites dans les paragraphes 17.1 et 17.2 :

- (i) $\Delta \mathbf{f}_{M_0}(\mathbf{h}, \mathbf{k}) = f(x_0 + h, y_0 + k) - f(x_0, y_0)$ mesure l'accroissement de la fonction lorsque l'on passe du point $M_0 = (x_0, y_0)$ au point $M = (x_0 + h, y_0 + k)$ en restant sur la surface représentative \mathcal{S} de f , tandis que $df_{M_0}(h, k) = h f'_x(x_0, y_0) + k f'_y(x_0, y_0)$ mesure l'accroissement obtenu lorsque l'on passe du point M_0 au point M en restant sur le plan tangent π_0 en P_0 à \mathcal{S} .
- (ii) Remplacer f par son approximation affine au voisinage d'un point M_0 , revient à remplacer la surface représentative de f au voisinage de P_0 par le plan tangent à \mathcal{S} en P_0 . On dit aussi que l'on "linéarise" la fonction f au voisinage de M_0 .

17.3. Notation différentielle

17.3.1. Différentielle et accroissement des projections. — Considérons les projections $\pi_1 : (x, y) \mapsto x$ et $\pi_2 : (x, y) \mapsto y$. Calculons les différentielles de ces projections :

$$\forall M_0 \in \mathbb{R}^2, \forall (h, k) \in \mathbb{R}^2, d\pi_{1M_0}(h, k) = h \quad \text{et} \quad d\pi_{2M_0}(h, k) = k.$$

On constate que les résultats obtenus sont **indépendants** du point M_0 . On peut donc supprimer la référence au point M_0 et noter plus simplement $d\pi_1(h, k) = h$ et $d\pi_2(h, k) = k$.

Comme dans le cas des fonctions d'une variable (voir page 29), on va se permettre l'abus de langage commode, qui consiste à confondre les projections avec leurs images, et prendre comme notations pratiques, dx et dy au lieu de $d\pi_1$ et $d\pi_2$. Donc :

$$(17.1) \quad \forall (h, k) \in \mathbb{R}^2, h = dx(h, k), \quad k = dy(h, k).$$

Dans ce qui suit, f est une fonction de classe \mathcal{C}^1 sur un ouvert \mathcal{U} et $M_0 = (x_0, y_0) \in \mathcal{U}$.

17.3.2. Différentielle totale. — La différentielle de f au point M_0 peut s'écrire en utilisant l'équation (17.1) :

$$df_{M_0}(h, k) = \frac{\partial f}{\partial x}(x_0, y_0) dx(h, k) + \frac{\partial f}{\partial y}(x_0, y_0) dy(h, k).$$

Comme cette égalité est valable pour tout $(h, k) \in \mathbb{R}^2$, on en déduit une égalité entre fonctions :

$$(17.2) \quad df_{M_0} = \frac{\partial f}{\partial x}(x_0, y_0) dx + \frac{\partial f}{\partial y}(x_0, y_0) dy.$$

L'égalité (17.2) étant valable pour tout point $M_0 = (x_0, y_0)$ de \mathcal{U} , on omet le point M_0 pour simplifier, ce qui donne une **notation symbolique** :

$$(17.3) \quad df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy.$$

Cette notation s'appelle *la différentielle totale de f* et exprime la relation, facile à retenir, qui relie les quantités df , dx et dy . Elle ne signifie rien d'autre que "pour avoir la valeur de $df_{M_0}(h, k)$, on remplace dans (17.3) dx par h , dy par k et on prend les dérivées partielles au point M_0 ".

Exemple 17.17. — Soit $f(x, y) = x^{1/3} y^{1/2}$ alors $df = (\frac{1}{3} x^{-2/3} y^{1/2}) dx + (\frac{1}{2} x^{1/3} y^{-1/2}) dy$.

17.3.3. Utilisation faite par les économistes. — En notant $\Delta x = h$ et $\Delta y = k$ les accroissements des variables et en utilisant le corollaire 17.12, on peut écrire :

$$(17.4) \quad \Delta f_{M_0}(\Delta x, \Delta y) \simeq \frac{\partial f}{\partial x}(x_0, y_0) \Delta x + \frac{\partial f}{\partial y}(x_0, y_0) \Delta y.$$

Les économistes utilisent la notation moins rigoureuse Δf_{M_0} pour $\Delta f_{M_0}(\Delta x, \Delta y)$. Comme l'égalité (17.4) est valable pour tout point $M_0 = (x_0, y_0)$ de \mathcal{U} , on omet le point M_0 pour simplifier, ce qui donne :

$$(17.5) \quad \Delta f_{M_0} \simeq \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y.$$

On constate une certaine similitude d'écriture entre (17.3) et (17.5), c'est pourquoi, lorsque les nombres Δx et Δy sont suffisamment petits pour que l'erreur soit négligeable, voir quasi nulle, les économistes écrivent dx au lieu de Δx , dy au lieu de Δy et df au lieu de Δf et retrouvent ainsi (17.3).

En d'autres termes, les économistes utilisent directement la formule (17.3) en considérant que df , dx et dy représentent des petits accroissements, c'est-à-dire des nombres et non des fonctions.

Remarque 17.18. — Si on écrit $z = f(x, y)$ on écrit plutôt Δz à la place de Δf et dz au lieu de df .

17.3.4. Opérations algébriques sur les différentielles totales. — La proposition 7.16 page 29 du chapitre 7 reste entièrement valable pour les différentielles de fonctions de deux variables. Ces formules permettent un calcul mécanique sur les différentielles.

CHAPITRE 18

CALCULS APPROCHÉS DE VARIATIONS

Nous allons dans ce chapitre généraliser les définitions vues au Chapitre 8 pour les fonctions d'une variable. On notera

$$\mathcal{D} = \{(x, y) \in \mathbb{R}^2 \mid x > 0 \text{ et } y > 0\}.$$

18.1. Fonctions marginales et variation absolue

Définition 18.1. — Soit f une fonction économique (production, coût, utilité, demande ...) dépendant des variables x et y (quantités, prix ...). On suppose que f est de classe \mathcal{C}^1 sur \mathcal{D} . **Les fonctions marginales** par rapport à x et y sont respectivement $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$.

Interprétation : Soit $M_0 = (x_0, y_0)$ un point où f est définie. En utilisant le corollaire 17.12 de la page 73 on peut écrire :

$$\Delta f_{M_0}(1, 0) = f(x_0 + 1, y_0 + 0) - f(x_0, y_0) \simeq \frac{\partial f}{\partial x}(x_0, y_0) \times 1$$

$$\Delta f_{M_0}(0, 1) = f(x_0 + 0, y_0 + 1) - f(x_0, y_0) \simeq \frac{\partial f}{\partial y}(x_0, y_0) \times 1.$$

Ce qui signifie que la valeur des fonctions marginales au point M_0 donne une valeur approchée de l'accroissement de f , lorsque l'une des variables augmente d'une unité à partir du point M_0 , l'autre variable restant constante.

Notons Δx et Δy les accroissements des variables x et y à partir du point M_0 . L'accroissement de f s'appelle aussi *variation absolue* : $\Delta f_{M_0}(\Delta x, \Delta y) = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$.

En utilisant toujours le corollaire 17.12 de la page 73 on peut écrire (voir l'exemple 17.3 page 71) :

Si $\nabla f(x_0, y_0) \neq (0, 0)$, on a alors l'approximation suivante

$$\underbrace{\Delta f_{M_0}(\Delta x, \Delta y)}_{\text{variation absolue de } f} \simeq \frac{\partial f}{\partial x}(x_0, y_0) \underbrace{\Delta x}_{\text{variation absolue de } x} + \frac{\partial f}{\partial y}(x_0, y_0) \underbrace{\Delta y}_{\text{variation absolue de } y}.$$

Autrement dit, la variation absolue est pratiquement égale à la somme des produits des fonctions marginales (calculées au point considéré) par les accroissements des variables.

18.2. Variation relative et élasticités

Définition 18.2. — Sous les hypothèses de la définition précédente, si f est strictement positive, on appelle *élasticité de f par rapport à x* et *élasticité de f par rapport à y* les fonctions $e_{f/x}$ et $e_{f/y}$ définies sur \mathcal{D} par :

$$e_{f/x}(x, y) = \frac{\partial f}{\partial x}(x, y) \frac{x}{f(x, y)} \quad e_{f/y}(x, y) = \frac{\partial f}{\partial y}(x, y) \frac{y}{f(x, y)}.$$

Si l'expression de $\ln(f)$ est simple, la différentielle de la fonction $g = \ln(f)$ permet de trouver facilement les élasticités car avec la notation différentielle on peut écrire :

$$(18.1) \quad d(\ln f) = \frac{df}{f} = \frac{\frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy}{f} = e_{f/x} \frac{dx}{x} + e_{f/y} \frac{dy}{y}.$$

Exemple 18.3 (Important en Économie). — Si $f(x, y) = x^\alpha y^\beta$ alors

$$g(x, y) = \ln(f(x, y)) = \alpha \ln(x) + \beta \ln(y)$$

et $dg = \alpha \frac{dx}{x} + \beta \frac{dy}{y}$, donc $e_{f/x} = \alpha$ et $e_{f/y} = \beta$.

Interprétation : Soit $M_0 = (x_0, y_0)$ un point de \mathcal{D} donc $x_0 \neq 0$ et $f(M_0) \neq 0$. Si $\Delta x \neq 0$, d'après l'estimation de la variation absolue vue au paragraphe 18.1, on peut écrire :

$$\frac{\partial f}{\partial x}(M_0) \simeq \frac{\Delta f_{M_0}(\Delta x, 0)}{\Delta x}$$

donc

$$e_{f/x}(M_0) \simeq \frac{\Delta f_{M_0}(\Delta x, 0)}{f(M_0)} \bigg/ \frac{\Delta x}{x_0}$$

est le quotient des variations relatives de f , $\frac{\Delta f_{M_0}(\Delta x, 0)}{f(M_0)}$, et de x , $\frac{\Delta x}{x_0}$. En particulier $e_{f/x}(M_0)$ donne une valeur approchée de la variation relative de f à partir du point M_0 si $\frac{\Delta x}{x_0} = 1\%$ et si y ne varie pas.

L'élasticité ne dépend pas des unités choisies.

Application au calcul des variations relatives. Notons

$$\frac{\Delta f}{f(x_0, y_0)} = \frac{f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)}{f(x_0, y_0)}.$$

Comme on peut écrire :

$$\frac{\partial f}{\partial x}(x_0, y_0) = e_{f/x}(x_0, y_0) \frac{f(x_0, y_0)}{x_0} \quad \text{et} \quad \frac{\partial f}{\partial y}(x_0, y_0) = e_{f/y}(x_0, y_0) \frac{f(x_0, y_0)}{y_0}$$

on en déduit, en divisant la relation (18.1) par $f(x_0, y_0) \neq 0$:

$$\underbrace{\frac{\Delta f}{f(x_0, y_0)}}_{\text{variation relative de } f} \simeq e_{f/x}(x_0, y_0) \underbrace{\frac{\Delta x}{x_0}}_{\text{variation relative de } x} + e_{f/y}(x_0, y_0) \underbrace{\frac{\Delta y}{y_0}}_{\text{variation relative de } y}.$$

On peut dire que la variation relative est pratiquement égale à la somme des produits des élasticité par les variations relatives respectives des deux variables. Cela donne des formules similaires pour les variations relatives et les variations absolues : les élasticité remplacent les fonctions marginales, les variations relatives remplacent les variations absolues. La relation (18.1) permet de retrouver très facilement l'expression encadrée.

Exemple 18.4. — Soit la fonction de Cobb-Douglas définie sur \mathcal{D} par $f(x, y) = 4x^{3/4}y^{1/4}$ et soit $(x_0, y_0) = (10000, 625)$. Calculer une valeur approchée de la variation absolue et de la variation relative de f lorsque x passe de 10000 à 10010 et y de 625 à 623.

\mathcal{D} est ouvert et la fonction f est bien de classe \mathcal{C}^1 sur \mathcal{D} d'après le corollaire 16.12 page 69. On commence par calculer les dérivées partielles premières :

$$\frac{\partial f}{\partial x}(x, y) = 3x^{-1/4}y^{1/4} \quad \text{et} \quad \frac{\partial f}{\partial y}(x, y) = x^{3/4}y^{-3/4},$$

donc

$$\frac{\partial f}{\partial x}(10000, 625) = 1.5 \quad \text{et} \quad \frac{\partial f}{\partial y}(10000, 625) = 8$$

et $f(10000, 625) = 20000$ (ces calculs se font sans calculatrice). On peut répondre à la première question :

$$f(10010, 623) - f(10000, 625) \simeq 1.5 \times 10 + 8 \times (-2) = -1.$$

Pour la variation relative, on calcule les élasticité :

$$e_{f/x}(10000, 625) = \frac{(1.5) \times 10000}{20000} = 0.75$$

$$e_{f/y}(10000, 625) = \frac{8 \times 625}{20000} = 0.25.$$

On en déduit :

$$\frac{f(10010, 623) - f(10000, 625)}{f(10000, 625)} \simeq 0.75 \times \frac{10}{10000} + 0.25 \times \frac{-2}{625} \simeq -0.005\%.$$

Ces calculs approchés donnent des estimations rapides avec des calculs assez simples. On peut toujours prendre sa calculette pour calculer les valeurs exactes!!! Dans cet exercice les élasticités sont constantes (on peut le vérifier). Cela est dû au choix particulier de la fonction f . Il suffit de refaire le calcul des élasticités en partant de la différentielle de $\ln f$ (voir (18.1) page 77).



*La variation relative s'exprime **toujours** en pourcentages.*

CHAPITRE 19

DÉRIVÉES PARTIELLES DU DEUXIÈME ORDRE

Dans tout ce chapitre on fait l'hypothèse :

(\mathcal{H}_1) f est une fonction de deux variables, définie sur un ouvert \mathcal{U} de \mathbb{R}^2 contenu dans \mathcal{D}_f et f admet des dérivées partielles du premier ordre en tout point de \mathcal{U} .

19.1. Fonctions dérivées partielles secondes

On peut se demander si les fonctions $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ admettent à leur tour des fonctions dérivées partielles. Si c'est le cas, on définit ainsi les fonctions dérivées partielles secondes.

Définition 19.1. — Sous l'hypothèse (\mathcal{H}_1), si les fonctions dérivées partielles premières $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ admettent des fonctions dérivées partielles sur \mathcal{U} ces dérivées partielles sont appelées *fonctions dérivées partielles secondes de f sur \mathcal{U}* , et notées :

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) = (f'_x)'_x = f''_{xx} & \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = (f'_x)'_y = f''_{xy} \\ \frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) = (f'_y)'_y = f''_{yy} & \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = (f'_y)'_x = f''_{yx}. \end{aligned}$$

Nous utiliserons, en général, les premières notations, appelées notations de Lagrange.

Si les dérivées partielles secondes de f existent en tout point M de \mathcal{U} , à partir d'une fonction f de deux variables, on peut donc définir quatre fonctions dérivées partielles secondes.

Exemple 19.2 (Suite de l'exemple 16.5 page 68). — On trouve :

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 2y e^{x^2} + 4x^2 y e^{x^2}, \quad \frac{\partial^2 f}{\partial y \partial x}(x, y) = 2x e^{x^2}, \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = 2x e^{x^2} \quad \text{et} \quad \frac{\partial^2 f}{\partial y^2}(x, y) = 0.$$

L'égalité entre les dérivées partielles secondes "croisées" est vraie sous certaines hypothèses (voir le théorème 19.6 page 82).



Avec la notation $\frac{\partial^2 f}{\partial y \partial x}$ les dérivations s'écrivent de droite à gauche (comme pour la composition) c'est-à-dire que l'on dérive d'abord par rapport à x , puis par rapport à y . Dans les autres notations les dérivations s'écrivent de gauche à droite.

19.2. Matrice Hessienne

Une matrice désigne tout simplement un tableau de nombres.

Définition 19.3. — Si f admet des dérivées partielles secondes au point $M_0 = (x_0, y_0)$ on appelle *matrice hessienne de f au point $M_0 = (x_0, y_0)$* la matrice :

$$D^2 f(M_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(M_0) & \frac{\partial^2 f}{\partial y \partial x}(M_0) \\ \frac{\partial^2 f}{\partial x \partial y}(M_0) & \frac{\partial^2 f}{\partial y^2}(M_0) \end{pmatrix}$$

La première ligne (resp. la deuxième) correspond à une première dérivation par rapport à la première (resp. la deuxième) variable.

La première colonne (resp. la deuxième) correspond à une deuxième dérivation par rapport à la première (resp. la deuxième) variable.

On obtient pour l'exemple 19.2 de la page 81 :

$$D^2 f(x, y) = \begin{pmatrix} 2ye^{x^2} + 4x^2ye^{x^2} & 2xe^{x^2} \\ 2xe^{x^2} & 0 \end{pmatrix} \quad \text{et} \quad D^2 f(1, 1) = \begin{pmatrix} 6e & 2e \\ 2e & 0 \end{pmatrix}$$

19.3. Fonctions de classe \mathcal{C}^2

Définition 19.4. — Sous l'hypothèse (H1), on dit que f est de classe \mathcal{C}^2 sur \mathcal{U} si f admet des fonctions dérivées partielles secondes sur \mathcal{U} et si toutes les fonctions dérivées partielles secondes de f sont continues sur \mathcal{U} .

Remarque 19.5. — Si f est de classe \mathcal{C}^2 sur \mathcal{U} alors les fonctions dérivées partielles premières de f sont de classe \mathcal{C}^1 sur \mathcal{U} donc continues. Par suite f est de classe \mathcal{C}^1 sur \mathcal{U} et f est continue sur \mathcal{U} .

Les fonctions de classe \mathcal{C}^2 vérifient l'égalité entre les dérivées partielles secondes "croisées". Elles jouent un rôle important dans la suite du cours.

Théorème 19.6 (Théorème de Schwarz). — Si f est de classe \mathcal{C}^2 sur un ouvert \mathcal{U} de \mathbb{R}^2 alors on a sur \mathcal{U} :

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

Autrement dit : si f est de classe \mathcal{C}^2 sur un ouvert \mathcal{U} de \mathbb{R}^2 on peut intervertir l'ordre des dérivations. Ceci justifiera l'utilisation des "notations de Monge" dans le chapitre suivant.

19.4. Opérations algébriques et composition

Les résultats du paragraphe 16.3 (les opérations algébriques, la composition à gauche et le lemme d'extension) restent vrais si l'on remplace "de classe \mathcal{C}^1 " par "de classe \mathcal{C}^2 ". En particulier, les fonctions polynômes et les fractions rationnelles de deux variables sont de classe \mathcal{C}^2 sur leur domaine de définition.

CHAPITRE 20

DÉVELOPPEMENT LIMITÉ D'ORDRE 2

20.1. Définitions

Comme pour le DL d'ordre 1, on peut deviner les modifications à faire pour généraliser à deux variables, le résultat vu pour une variable (voir le chapitre 9). Il y a quatre dérivées partielles secondes et quatre combinaisons possibles des accroissements h et k des variables. Or, pour une fonction de classe \mathcal{C}^2 les dérivées partielles secondes croisées sont égales.

Nous admettrons le résultat suivant :

Théorème 20.1 (Formule de Taylor-Young à l'ordre 2 ($n = 2$)). — Soient f une fonction de classe \mathcal{C}^2 sur \mathcal{U} ouvert de \mathbb{R}^2 et $M_0 = (x_0, y_0) \in \mathcal{U}$. Alors il existe une fonction ε , continue sur une boule ouverte \mathcal{B} de centre $(0, 0)$, telle que : $\varepsilon(0, 0) = 0$ et

$$\forall (h, k) \in \mathcal{B}, f(x_0 + h, y_0 + k) = f(x_0, y_0) + h \frac{\partial f}{\partial x}(x_0, y_0) + k \frac{\partial f}{\partial y}(x_0, y_0) + \frac{1}{2} \left(h^2 \frac{\partial^2 f}{\partial x^2}(x_0, y_0) + k^2 \frac{\partial^2 f}{\partial y^2}(x_0, y_0) + 2hk \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \right) + (h^2 + k^2)\varepsilon(h, k).$$

On dit que f admet un développement limité d'ordre 2 (DL d'ordre 2) au voisinage de M_0 .

Remarque 20.2. — La fonction ε est inconnue, on sait seulement qu'elle est continue en $(0, 0)$ et vérifie $\varepsilon(0, 0) = 0$. Le DL d'ordre 2 est donc une formule locale qui ne s'utilise qu'au voisinage d'un point, c'est-à-dire pour des valeurs petites de $|h|$ et $|k|$. Il est inutile de préciser la boule ouverte \mathcal{B} .
 – La quantité $(h^2 + k^2)\varepsilon(h, k)$ s'appelle *le reste* du DL d'ordre 2.

Exemple 20.3. — Soit la fonction $f(x, y) = e^{x+2y}$, on a $D_f = \mathbb{R}^2$ et posons $M_0 = (2, -1)$ (f est de classe \mathcal{C}^2 sur \mathbb{R}^2).

On a : $f(2, -1) = 1$, $\nabla f(x, y) = (e^{x+2y}, 2e^{x+2y})$, donc $\nabla f(2, -1) = (1, 2)$ et

$$D^2 f(x, y) = \begin{pmatrix} e^{x+2y} & 2e^{x+2y} \\ 2e^{x+2y} & 4e^{x+2y} \end{pmatrix}, \quad D^2 f(2, -1) = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}.$$

Donc le développement limité de f au voisinage du point $(2, -1)$ s'écrit pour tous h et k assez petits

$$f(2 + h, -1 + k) = 1 + h + 2k + \frac{1}{2}(h^2 + 2hk + 4k^2) + (h^2 + k^2)\varepsilon(h, k)$$

Notation. — Nous avons vu au chapitre 17 que :

$$h \frac{\partial f}{\partial x}(x_0, y_0) + k \frac{\partial f}{\partial y}(x_0, y_0) = df_{M_0}(h, k) = \langle \nabla f(M_0), (h, k) \rangle$$

Par analogie on note :

$$d^2 f_{M_0}(h, k) = h^2 \frac{\partial^2 f}{\partial x^2}(x_0, y_0) + 2hk \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) + k^2 \frac{\partial^2 f}{\partial y^2}(x_0, y_0)$$

et $d^2 f_{M_0}$ s'appelle *la forme quadratique associée à la matrice hessienne de f au point M_0* . Le théorème 20.1 s'écrit alors en posant : $H = (h, k)$ et $M_0 + H = (x_0 + h, y_0 + k)$:

$$f(M_0 + H) = f(M_0) + df_{M_0}(H) + \frac{1}{2}d^2 f_{M_0}(H) + \|H\|^2 \varepsilon(H).$$

Ce qui donne une formule très proche du cas à une variable (voir le chapitre 9 page 35) : df_{M_0} remplace $f'(x_0)$ et $d^2 f_{M_0}$ remplace $f''(x_0)$.

Calculons $d^2 f_{M_0}(h, k)$ pour une fonction f de classe \mathcal{C}^2 .

La méthode la plus simple consiste à partir de la matrice hessienne au point M_0 et à écrire :

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(M_0) & \frac{\partial^2 f}{\partial x \partial y}(M_0) \\ \frac{\partial^2 f}{\partial x \partial y}(M_0) & \frac{\partial^2 f}{\partial y^2}(M_0) \end{pmatrix} = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

on a alors

$$d^2 f_{M_0}(h, k) = rh^2 + 2shk + tk^2.$$

Remarquons que nous pouvons aussi écrire la forme quadratique grâce au produit scalaire

$$d^2 f_{M_0}(H) = \langle H, D^2 f(M_0)H \rangle.$$

20.2. Forme quadratique associée à la matrice hessienne

20.2.1. Définitions. — Avec les notations de Monge, r , s et t sont des nombres réels. Ainsi la fonction $d^2 f_{M_0}$ est une fonction polynôme des deux variables h et k ne contenant que des monômes de degré 2.

On peut constater que pour tous réels h , k et λ on a : $d^2 f_{M_0}(\lambda h, \lambda k) = \lambda^2 d^2 f_{M_0}(h, k)$. Nous parlons de *forme quadratique* associée à la matrice hessienne au point M_0 . Remarquons que l'on a toujours $d^2 f_{M_0}(0, 0) = 0$.

Dans la suite du cours il sera utile de connaître le signe du nombre réel $d^2 f_{M_0}(h, k)$. Pour cela on utilisera l'outil suivant :

Définition 20.4. — Soit f une fonction de classe \mathcal{C}^2 définie sur \mathcal{U} ouvert de \mathbb{R}^2 et soit M_0 un point de \mathcal{U} . En utilisant r , s , t pour les dérivées partielles secondes en M_0 on appelle *déterminant de la matrice hessienne* de f en M_0 le réel défini par :

$$\begin{vmatrix} r & s \\ s & t \end{vmatrix} = rt - s^2$$

20.2.2. Signe de la forme quadratique associée à la matrice hessienne. —

Théorème 20.5. — Soient r , s et t trois réels et $d^2 f_{M_0}$ la forme quadratique associée.

1. Si $rt - s^2 > 0$ et si
 - $r > 0$ ou $t > 0$ alors $\forall (h, k) \neq (0, 0)$, $d^2 f_{M_0}(h, k) > 0$
 - $r < 0$ ou $t < 0$ alors $\forall (h, k) \neq (0, 0)$, $d^2 f_{M_0}(h, k) < 0$
2. Si $rt - s^2 \geq 0$ et si
 - $r > 0$ ou $t > 0$ alors $\forall (h, k) \in \mathbb{R}^2$, $d^2 f_{M_0}(h, k) \geq 0$
 - $r < 0$ ou $t < 0$ alors $\forall (h, k) \in \mathbb{R}^2$, $d^2 f_{M_0}(h, k) \leq 0$
3. Si $rt - s^2 < 0$ alors la forme quadratique $d^2 f_{M_0}$ change de signe.

Preuve. — Avec les notations de Monge : $d^2 f_{M_0}(h, k) = rh^2 + 2shk + tk^2$. On suppose que r , s et t ne sont pas tous nuls sinon $d^2 f_{M_0}(h, k) = 0$. Si $rt - s^2 > 0$ alors $rt > s^2 \geq 0$ et donc r et t sont du même signe et non nuls. Si $rt - s^2 = 0$ alors on a ($s \neq 0$, $rt = s^2$ donc $rt > 0$) ou ($r = s = 0$ et $s \neq 0$) ou ($s = t = 0$ et $r \neq 0$). Étudions le signe de $d^2 f_{M_0}(h, k)$ en distinguant différents cas.

– Supposons $r \neq 0$, on peut écrire

$$d^2 f_{M_0}(h, k) = r \left[\left(h + \frac{s}{r} k \right)^2 + k^2 \frac{rt - s^2}{r^2} \right].$$

- Si $rt - s^2 > 0$ alors on a : $d^2 f_{M_0}(h, k) \geq 0$ si $r > 0$, $d^2 f_{M_0}(h, k) \leq 0$ si $r < 0$ et enfin $d^2 f_{M_0}(h, k) = 0$ seulement pour $h = k = 0$.
- Si $rt - s^2 = 0$ alors on a : $d^2 f_{M_0}(h, k) \geq 0$ si $r > 0$, $d^2 f_{M_0}(h, k) \leq 0$ si $r < 0$ et enfin $d^2 f_{M_0}(h, k) = 0$ pour les couples (h, k) vérifiant $h + \frac{s}{r} k = 0$.
- Si $rt - s^2 < 0$ la forme quadratique change de signe car $d^2 f_{M_0}(1, 0) = r$ et $d^2 f_{M_0}\left(-\frac{s}{r}, 1\right) = \frac{rt - s^2}{r}$ qui est du signe de $-r$.

– Supposons $r = 0$ et $t \neq 0$ alors $rt - s^2 = -s^2 \leq 0$. On peut écrire

$$d^2 f_{M_0}(h, k) = 2shk + tk^2 = t \left(k + \frac{s}{t} h \right)^2 - h^2 \frac{s^2}{t}.$$

- Si $rt - s^2 < 0$ alors $s \neq 0$. On constate que $d^2f_{M_0}(0, 1) = t$ et $d^2f_{M_0}(1, -s/t) = -s^2/t$, expression du signe de $-t$. Donc $d^2f_{M_0}$ change de signe.
 - Si $rt - s^2 = 0$ alors $s = 0$. On constate que $d^2f_{M_0}(h, k) = tk^2$ dont $d^2f_{M_0}(h, k) \geq 0$ si $t \geq 0$ et $d^2f_{M_0}(h, k) \leq 0$ si $t \leq 0$
 - Supposons que $r = t = 0$ et $s \neq 0$. Alors $rt - s^2 = -s^2 < 0$. On a $d^2f_{M_0}(h, k) = 2shk$. On remarque que $d^2f_{M_0}(1, 1) = 2s$ et $d^2f_{M_0}(1, -1) = -2s$ donc la forme quadratique $d^2f_{M_0}$ change de signe.
- On peut échanger les rôles de s et t dans les raisonnements précédents. \square

Remarque 20.6. — La démonstration précédente repose sur la méthode dite de décomposition en carrés d'une forme quadratique. Cette méthode peut être reproduite "à la main" sur chaque exemple si on ne souhaite pas apprendre par coeur les critères donnant le signe d'une forme quadratique.

20.3. Position du plan tangent par rapport à la surface représentative

D'après le théorème 20.1, si f est une fonction de classe \mathcal{C}^2 sur \mathcal{U} ouvert de \mathbb{R}^2 et $M_0 = (x_0, y_0) \in \mathcal{U}$, alors pour des valeurs suffisamment petites de $|h|$ et $|k|$, on peut écrire le développement limité :

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + df_{M_0}(h, k) + \frac{1}{2}d^2f_{M_0}(h, k) + (h^2 + k^2)\varepsilon(h, k).$$

Rappelons que l'équation cartésienne du plan tangent en $P_0 = (x_0, y_0, f(x_0, y_0))$ à la surface représentative de f est : $z = f(x_0, y_0) + df_{M_0}(h, k)$ où $h = x - x_0$, $k = y - y_0$.

Par suite, la position du plan tangent par rapport à la surface représentative de f est donnée par le signe de :

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) - df_{M_0}(h, k) = \frac{1}{2}d^2f_{M_0}(h, k) + (h^2 + k^2)\varepsilon(h, k).$$

Si $d^2f_{M_0}(h, k) \neq 0$, pour $(h, k) \neq (0, 0)$, comme le reste $(h^2 + k^2)\varepsilon(h, k)$ est négligeable devant le terme d'ordre 2, la position du plan tangent par rapport à la surface représentative de f est donnée par le signe de $d^2f_{M_0}(h, k)$ pour h et k assez petits.

En appliquant le théorème 20.5 on en déduit

Proposition 20.7. — Soit f une fonction de classe \mathcal{C}^2 sur \mathcal{U} ouvert de \mathbb{R}^2 et soit $M_0 = (x_0, y_0) \in \mathcal{U}$. Notons π_0 le plan tangent en $P_0 = (x_0, y_0, f(x_0, y_0))$ à la surface représentative de f . Alors, en utilisant r, s, t pour les dérivées partielles secondes en M_0 on a

1. Si $rt - s^2 > 0$ et ($r > 0$ ou $t > 0$) alors au voisinage de M_0 , la surface représentative de f est située au-dessus du plan tangent π_0
2. Si $rt - s^2 > 0$ et ($r < 0$ ou $t < 0$) alors au voisinage de M_0 , la surface représentative de f est située en-dessous du plan tangent π_0 .
3. Si $rt - s^2 < 0$ alors au voisinage de M_0 , la surface représentative de f traverse le plan tangent π_0 .



Les inégalités sont strictes : si $rt - s^2 = 0$ on ne peut rien dire, le reste du développement limité n'est plus négligeable devant la forme quadratique.

Exemple 20.8. — Soit la fonction $f(x, y) = x^3 + xy^2$. La fonction f est un polynôme donc de classe \mathcal{C}^2 sur \mathbb{R}^2 . On a

$$\nabla f(x, y) = (3x^2 + y^2, 2xy), \quad r = 6x, \quad t = 2x, \quad s = 2y.$$

Donc $rt - s^2 = 12x^2 - 4y^2$.

- Pour $M_1 = (2, 1)$ on a $rt - s^2 = 48 - 4 > 0$ et $r = 12 > 0$ donc $d^2f_{M_1}(h, k) \geq 0$. Au voisinage de M_1 , la surface représentative de f est située au-dessus du plan tangent associé à M_1 .
- Pour $M_2 = (-3, 2)$ on a $rt - s^2 = 12 \times 9 - 16 > 0$ et $r = -18 < 0$. Donc $d^2f_{M_2}(h, k) \leq 0$. Au voisinage de M_2 , la surface représentative de f est située au-dessous du plan tangent associé à M_2 .
- Pour $M_3 = (1, 2)$ on a $rt - s^2 = 12 - 16 < 0$. Le plan tangent associé à M_3 traverse la surface représentative de f .

CHAPITRE 21

FONCTIONS CONVEXES OU CONCAVES DE DEUX VARIABLES

Les notions de convexité ou de concavité jouent un rôle très important en Économie et en particulier dans les problèmes d'optimisation. Nous avons déjà rencontré ces notions pour les fonctions d'une variable. Il nous reste donc à les généraliser à deux variables. Pour simplifier l'exposé, nous ne donnerons pas la définition la plus générale pour des fonctions quelconques. Nous nous limiterons aux fonctions de classe \mathcal{C}^1 et éventuellement de classe \mathcal{C}^2 car ce sont celles que nous utiliserons en pratique.

21.1. Définitions pour les fonctions de classe \mathcal{C}^1

On avait vu que les fonctions convexes ou concaves d'une variable devaient être définies sur un intervalle, car les intervalles sont les seuls sous-ensembles convexes de \mathbb{R} . D'autre part, la définition de telles fonctions faisait intervenir les tangentes à la courbe représentative. Il est très facile d'étendre la définition à deux variables en remplaçant un intervalle de \mathbb{R} par un sous-ensemble convexe de \mathbb{R}^2 , et les tangentes à la courbe représentative par les plans tangents à la surface représentative, ce qui donne :

Définition 21.1. — Soient f une fonction de classe \mathcal{C}^1 sur un ouvert \mathcal{U} de \mathbb{R}^2 et C un **ensemble convexe** de \mathbb{R}^2 inclus dans \mathcal{U} .

On note $\mathcal{G}_f(C)$ la portion de la surface représentative de f correspondant à $(x, y) \in C$.

– f est *convexe* sur C si $\mathcal{G}_f(C)$ est située **au-dessus** de tous ses plans tangents. De façon équivalente, f est concave sur C si l'on a

$$\forall (x, y) \in C, \forall (x_0, y_0) \in C, f(x, y) \geq f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).$$

– f *concave* sur C si $\mathcal{G}_f(C)$ est située **au-dessous** de tous ses plans tangents. De façon équivalente, f est convexe sur C si l'on a

$$\forall (x, y) \in C, \forall (x_0, y_0) \in C, f(x, y) \leq f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).$$

Remarque 21.2. — 1. Nous voyons ici une définition pour une fonction de classe \mathcal{C}^1 mais il existe une définition plus générale, on dit que f est convexe sur C si pour tout $A, B \in C$

$$\forall t \in [0, 1], f(tA + (1 - t)B) \leq tf(A) + (1 - t)f(B).$$

Cette définition nous fait comprendre qu'il est nécessaire que pour tout $A, B \in C$ et $t \in [0, 1]$ on ait $tA + (1 - t)B \in C$, ce qui implique que l'ensemble C doit être un ensemble convexe (voir la définition 13.2 page 55).

- f est concave sur C si et seulement si $-f$ est convexe sur C .
- Si dans les définitions précédentes les inégalités sont strictes pour $(x, y) \neq (x_0, y_0)$, on parle alors de fonctions strictement convexes sur C ou strictement concaves sur C .
- Il est important de mentionner l'ensemble convexe sur lequel une fonction est convexe. Attention, même si \mathcal{D}_f est un ouvert convexe de \mathbb{R}^2 , f peut être convexe (resp. concave) sur un sous-ensemble convexe $C \subset \mathcal{D}_f$ sans être convexe (resp. concave) sur \mathcal{D}_f .

Dans les modèles économiques on suppose souvent (mais pas toujours) que les fonctions d'utilité et les fonctions de production sont des fonctions concaves.

Exemple 21.3. — La fonction affine $f(x, y) = ax + by + c$ est à la fois convexe **et** concave sur \mathbb{R}^2 . En effet, sa surface représentative est le plan d'équation : $z = ax + by + c$ qui est confondu avec tous ses plans tangents.

Donc les fonctions affines vérifient à la fois l'inégalité de convexité et l'inégalité de concavité ce qui donne **pour toute fonction affine** f :

$$\forall (x, y) \in \mathbb{R}^2, \forall (x_0, y_0) \in \mathbb{R}^2, f(x, y) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).$$

21.2. Critère pour les fonctions de classe \mathcal{C}^2

La surface représentative d'une fonction de deux variables n'est pas aisée à construire, donc la définition n'est pas aussi "visuelle" que pour les fonctions d'une variable (la définition d'une tangente étant assez simple). D'autre part, la définition n'est pas très maniable puisqu'elle demande de démontrer une inégalité qui dépend de quatre variables (x, y, x_0, y_0) . Il nous faut donc un critère plus simple à utiliser. On caractérisait les fonctions convexes ou concaves d'une variable sur un intervalle à l'aide du signe de la dérivée seconde sur cet intervalle, **les inégalités étant larges** (voir le théorème 5.4 page 22). Pour les fonctions de deux variables, le signe de la dérivée seconde est remplacé par le signe de la forme quadratique associée à la matrice hessienne (voir la définition page 84) en un point quelconque du sous-ensemble convexe C .

Appliquons le théorème 20.5 page 84 pour tout point M de C . Nous obtenons le théorème suivant :

Théorème 21.4. — Soient f une fonction de classe \mathcal{C}^2 sur un ouvert \mathcal{U} et C un ensemble convexe de \mathbb{R}^2 inclus dans \mathcal{U} . Pour tout $(x, y) \in C$, notons

$$D^2f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial x \partial y}(x, y) \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) & \frac{\partial^2 f}{\partial y^2}(x, y) \end{pmatrix} = \begin{pmatrix} r & s \\ s & t \end{pmatrix}.$$

Alors :

– Si

$$\forall (x, y) \in C, rt - s^2 \geq 0 \text{ et } (r \geq 0 \text{ et } t \geq 0)$$

alors f est convexe sur C

– Si

$$\forall (x, y) \in C, rt - s^2 \geq 0 \text{ et } (r \leq 0 \text{ et } t \leq 0)$$

alors f est concave sur C

– S'il existe $(x, y) \in C$ tel que $rt - s^2 < 0$ au point (x, y) alors f n'est ni convexe, ni concave sur C .

Remarque 21.5. — Si l'on a $\forall (x, y) \in C, r = 0, s = 0$ et $t = 0$ alors f est affine donc convexe et concave.



Dans le théorème 21.4 les notations r, s et t désignent les valeurs des dérivées partielles secondes de f en un point (x, y) quelconque de C donc l'expression $rt - s^2$ **dépend en général des variables x et y** et par suite l'étude du signe n'est pas toujours aisée.

Exemple 21.6. — Soit la fonction $f(x, y) = \frac{1}{xy}$ sur l'ensemble $\mathcal{D}_f = \{(x, y) \in \mathbb{R}^2 \mid xy \neq 0\}$ qui est ouvert. La fonction f est une fraction rationnelle donc de classe \mathcal{C}^2 sur \mathcal{D}_f qui est un ensemble ouvert. On a

$$r = \frac{2}{x^3y}, \quad s = \frac{1}{x^2y^2} \quad \text{et} \quad t = \frac{2}{xy^3},$$

et donc $rt - s^2 = 3/(x^4y^4) > 0$.

Attention, dans cet exemple \mathcal{D}_f n'est pas convexe. Considérons les quatre sous-ensembles suivants de \mathcal{D}_f qui sont convexes (comme intersections de demi-plans) :

$$C_1 = \{(x, y) \in \mathbb{R}^2 \mid x > 0 \text{ et } y > 0\} \quad \text{et} \quad C_2 = \{(x, y) \in \mathbb{R}^2 \mid x < 0 \text{ et } y > 0\}$$

$$C_3 = \{(x, y) \in \mathbb{R}^2 \mid x < 0 \text{ et } y < 0\} \quad \text{et} \quad C_4 = \{(x, y) \in \mathbb{R}^2 \mid x > 0 \text{ et } y < 0\}$$

On déduit du théorème 21.4 que f est convexe sur C_1 et sur C_3 , et que f est concave sur C_2 et C_4 .

Cas particulier des polynômes de degré 2 :

Si $f(x, y) = ax^2 + bxy + cy^2 + dx + ey + h$ alors $\nabla f(x, y) = (2ax + by + d, bx + 2cy + e)$ et

$$D^2f(x, y) = \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix}$$

ce qui donne : $rt - s^2 = 4ac - b^2$. Les dérivées partielles secondes sont constantes (indépendantes de x et y) ce qui donne le corollaire 21.7 bien commode à utiliser :

Corollaire 21.7. — Si f est donnée par $f(x, y) = ax^2 + bxy + cy^2 + dx + ey + h$ avec a, b, c, d, e et h des nombres réels.

- Si $4ac - b^2 \geq 0$ et ($a > 0$ ou $c > 0$) alors f est convexe sur \mathbb{R}^2 .
- Si $4ac - b^2 \geq 0$ et ($a < 0$ ou $c < 0$) alors f est concave sur \mathbb{R}^2 .
- Si $4ac - b^2 < 0$ alors f n'est ni concave ni convexe.

Exemple 21.8. — Appliquons le corollaire 21.7 aux fonctions suivantes.

- Si $f(x, y) = 2x^2 + 3xy + 2y^2$ alors $4ac - b^2 = 16 - 9 > 0$ et $a > 0$. La fonction f est donc convexe sur \mathbb{R}^2 .
- Si $g(x, y) = -3x^2 + 3xy - y^2$ alors $4ac - b^2 = 12 - 9 > 0$ et $a < 0$. La fonction g est donc concave sur \mathbb{R}^2 .
- Si $h(x, y) = 2x^2 + 3xy + y^2$ alors $4ac - b^2 = 8 - 9 < 0$. Ainsi h n'est ni convexe, ni concave sur \mathbb{R}^2 .

21.3. Propriétés des fonctions convexes ou concaves

Le critère du signe de $rt - s^2$ n'est pas aussi pratique qu'il en a l'air car cette expression dépend des deux variables x et y et l'étude de son signe n'est pas toujours aisée. Dans certains cas on pourra simplifier les raisonnements grâce aux théorèmes qui suivent.

Conformément à la définition choisie toutes les fonctions convexes et concaves seront supposées de classe \mathcal{C}^1 .

21.3.1. Combinaison linéaire à coefficients positifs. —

Proposition 21.9. — Si f et g sont convexes (resp. concaves) sur $C \subset \mathbb{R}^2$, si α et β sont des réels positifs alors $\alpha f + \beta g$ est convexe (resp. concave) sur C .

– Toute combinaison linéaire à coefficients positifs d'un nombre fini de fonctions convexes (resp. concaves) sur C est convexe (resp. concave) sur C .

Preuve. — Nous allons nous contenter de démontrer cette proposition pour deux fonctions f et g convexes sur $C \subset \mathbb{R}^2$. En appliquant la définition, on a pour f et g :

$$\forall (x, y) \in C, \forall (x_0, y_0) \in C, f(x, y) \geq f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

et

$$\forall (x, y) \in C, \forall (x_0, y_0) \in C, g(x, y) \geq g(x_0, y_0) + \frac{\partial g}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial g}{\partial y}(x_0, y_0)(y - y_0)$$

Multiplions la première inégalité par $\alpha > 0$ et la seconde par $\beta > 0$ (les inégalités conservent le même sens car les coefficients sont positifs) et ajoutons-les membre à membre ce qui donne :

$$\forall (x, y) \in C, \forall (x_0, y_0) \in C, \alpha f(x, y) + \beta g(x, y) \geq \alpha f(x_0, y_0) + \beta g(x_0, y_0) + \left(\alpha \frac{\partial f}{\partial x}(x_0, y_0) + \beta \frac{\partial g}{\partial x}(x_0, y_0) \right) (x - x_0) + \left(\alpha \frac{\partial f}{\partial y}(x_0, y_0) + \beta \frac{\partial g}{\partial y}(x_0, y_0) \right) (y - y_0).$$

Posons $h = \alpha f + \beta g$, l'inégalité ci-dessus s'écrit alors :

$$\forall (x, y) \in C, \forall (x_0, y_0) \in C, h(x, y) \geq h(x_0, y_0) + \frac{\partial h}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial h}{\partial y}(x_0, y_0)(y - y_0),$$

ce qui signifie que la fonction h est convexe sur C .

Dans le cas de fonctions concaves la démonstration est la même. C'est seulement le sens des inégalités qui change. \square

Exemple 21.10. — Soient la fonction $h(x, y) = \frac{3}{xy} + 5(2x^2 + 3xy + 2y^2)$ et l'ensemble

$$C = \{(x, y) \in \mathbb{R}^2 \mid x > 0 \text{ et } y > 0\}.$$

La fonction $f(x, y) = 1/(xy)$ est convexe sur C convexe (voir l'exemple 21.6). La fonction $g(x, y) = 2x^2 + 3xy + 2y^2$ est convexe sur \mathbb{R}^2 (voir l'exemple 21.8) donc sur C . Comme $h = 3f + 5g$, ($3 > 0$, $5 > 0$) on déduit de la proposition 21.9, que h est convexe sur C .

21.3.2. Somme de fonctions d'une variable. —

Lemme 21.11 (Lemme d'extension). — Si f est convexe (resp. concave) sur un intervalle I de \mathbb{R} alors
 – la fonction g définie par : $g(x, y) = f(x)$ est convexe (resp. concave) sur $I \times \mathbb{R} \subset \mathbb{R}^2$
 – la fonction h définie par : $h(x, y) = f(y)$ est convexe (resp. concave) sur $\mathbb{R} \times I \subset \mathbb{R}^2$.

Preuve. — Supposons f convexe sur I . Alors par définition de la convexité :

$$(21.1) \quad \forall x \in I, \forall x_0 \in I, f(x) \geq f(x_0) + (x - x_0)f'(x_0).$$

Or, par définition de g , on a : $g(x, y) = f(x)$, $g(x_0, y_0) = f(x_0)$, $\frac{\partial g}{\partial x}(x_0, y_0) = f'(x_0)$ et enfin $\frac{\partial g}{\partial y}(x_0, y_0) = 0$. L'ensemble $I \times \mathbb{R}$ est un sous-ensemble convexe de \mathbb{R}^2 comme produit cartésien de deux intervalles. En utilisant les égalités, l'inégalité 21.1 s'écrit :

$$\forall (x, y) \in I \times \mathbb{R}, \forall (x_0, y_0) \in \times \mathbb{R}, g(x, y) \geq g(x_0, y_0) + \frac{\partial g}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial g}{\partial y}(x_0, y_0)(y - y_0)$$

Ce qui signifie que g est convexe sur $I \times \mathbb{R}$. On montre de la même façon que h est convexe en permutant les rôles de x et y .

Pour les fonctions concaves il suffit de changer le sens des inégalités. □

Exemple 21.12. — La fonction $g(x, y) = e^x$ est convexe sur \mathbb{R}^2 . En effet la fonction f définie par $f(x) = e^x$ est convexe sur \mathbb{R} . De même la fonction $h(x, y) = \ln(y)$, h est concave sur $\mathbb{R} \times]0, +\infty[$ car la fonction f définie par $f(y) = \ln(y)$ est concave sur $]0, +\infty[$.

Corollaire 21.13. — Soient f et g deux fonctions d'une variable. On suppose que f est convexe (resp. concave) sur I intervalle de \mathbb{R} , et que g convexe (resp. concave) sur J intervalle de \mathbb{R} . Alors la fonction h définie par : $h(x, y) = f(x) + g(y)$ est convexe (resp. concave) sur le sous-ensemble $I \times J$ (qui est convexe).

Preuve. — Supposons f et g convexes et posons : $u(x, y) = f(x)$ et $v(x, y) = g(y)$. On déduit du lemme 21.11 que u est convexe sur $I \times \mathbb{R}$ et que v est convexe sur $\mathbb{R} \times J$.

Or $I \times J$ est un convexe (produit cartésien de deux intervalles) inclus à la fois dans $I \times \mathbb{R}$ et $\mathbb{R} \times J$. Donc u et v sont convexes sur $I \times J$ et par suite $h = u + v$ est convexe sur $I \times J$ d'après la proposition 21.9.

On applique le même raisonnement pour les fonctions concaves. □

Exemple 21.14. — La fonction $f(x, y) = e^x + y^2$ est convexe sur \mathbb{R}^2 , et la fonction $f(x, y) = \ln(x) - 1/y$ est concave sur $C =]0, +\infty[\times]0, +\infty[$.

21.3.3. Composition. — Notons que les propositions 21.15 et 21.17 seront admises, elles sont néanmoins fondamentales.

Proposition 21.15 (Fonction affine composée par une fonction convexe ou concave)
 Soient h une fonction affine sur \mathbb{R}^2 et f une fonction convexe (resp. concave) sur $I \subset \mathbb{R}$.
 Alors $f \circ h$ est convexe (resp. concave) sur $C = \{(x, y) \in \mathbb{R}^2 \mid h(x, y) \in I\} \subset \mathbb{R}^2$.

Exemple 21.16. — La fonction $f(x, y) = \ln(2x - 3y + 4)$ est concave sur $C = \{(x, y) \in \mathbb{R}^2 \mid 2x - 3y + 4 > 0\}$.

Proposition 21.17 (Composition par une fonction convexe ou concave croissante)
 Soit g convexe (resp. concave) sur C convexe de \mathbb{R}^2 . Soit $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ convexe (resp. concave) **croissante** sur I , telle que $f \circ g$ soit définie sur C .
 Alors $f \circ g$ est convexe (resp. concave) sur C .

Exemple 21.18. — Soit $h(x, y) = (x^2 + y^2) + \exp(x^2 + y^2)$. Montrons que h est convexe sur \mathbb{R}^2 .

En effet posons $g(x, y) = x^2 + y^2$ et $f(u) = u + \exp(u)$. La fonction g est convexe sur \mathbb{R}^2 comme somme de deux fonctions d'une variable convexes sur \mathbb{R} . La fonction f est convexe sur \mathbb{R} car : $\forall u \in \mathbb{R}, f'(u) = 1 + \exp(u) > 0$ et $f''(u) = \exp(u) > 0$.

Ainsi $h = f \circ g$ est convexe sur \mathbb{R}^2 par application de la proposition 21.17

Remarque 21.19. — Attention à l'application de la proposition 21.17 : la composée de deux fonctions convexes (resp. concaves) n'est pas toujours une fonction convexe ou concave. Il faut une hypothèse supplémentaire pour la deuxième fonction, celle d'une variable, à savoir que celle-ci doit être **croissante**.

En choisissant des fonctions particulières pour la fonction f d'une variable vérifiant les hypothèses de la proposition 21.17, on obtient les critères suivants très commodes à utiliser.

Corollaire 21.20. — Soit C un ensemble convexe de \mathbb{R}^2 :

1. Si g convexe sur C alors e^g convexe sur C .
2. Si $g > 0$ et si $\ln(g)$ convexe sur C alors g convexe sur C .
3. Si $g > 0$, $h > 0$, $\ln(g)$ et $\ln(h)$ convexes sur C alors gh est convexe sur C .
4. Si $g > 0$ et si g convexe sur C alors g^2 est convexe sur C .

Preuve. — Pour la première assertion, on choisit $f(u) = \exp(u)$, la fonction f est croissante et convexe sur \mathbb{R} , et par suite $f \circ g = e^g$ est convexe sur C d'après la proposition 21.17.

Pour prouver la seconde assertion, on applique le premier résultat à la fonction $\ln(g)$ d'où $e^{\ln(g)} = g$ est convexe sur C .

La troisième assertion se prouve de la façon suivante : la fonction $\ln(g) + \ln(h) = \ln(g \times h)$ est convexe sur C d'après la proposition 21.9 et on applique la seconde assertion à la fonction $\ln(g \times h)$.

Pour la dernière, on choisit $f(u) = u^2$, f est convexe sur \mathbb{R} et croissante sur $]0, +\infty[$. Comme $g > 0$, on applique la proposition 21.17 à $f \circ g = g^2$. \square

Exemple 21.21. — Soit la fonction $f_1(x, y) = e^{2x^2+3xy+2y^2}$, posons $g(x, y) = 2x^2 + 3xy + 2y^2$. La fonction g est convexe sur \mathbb{R}^2 (voir la première fonction de l'exemple 21.8). Donc en appliquant le corollaire 21.20, f_1 est convexe sur \mathbb{R}^2 .

— Soit maintenant $f_2(x, y) = (x^2 + y^2 + 1)^2$ posons $g(x, y) = x^2 + y^2 + 1$ alors g est convexe sur \mathbb{R}^2 (appliquer le corollaire 21.7) et strictement positive. Donc en appliquant la quatrième assertion du corollaire 21.20 on en déduit que f_2 est convexe sur \mathbb{R}^2 .

— Prenons $f_3(x, y) = \frac{1}{xy}$, posons $g(x, y) = 1/x$ et $h(x, y) = 1/y$. Les fonctions g et h sont strictement positives sur $C_1 = \{(x, y) \in \mathbb{R}^2 \mid x > 0 \text{ et } y > 0\}$.

On a $\ln(g(x, y)) = -\ln(x)$ et $\ln(h(x, y)) = -\ln(y)$. Ces deux dernières fonctions sont convexes sur C_1 (appliquer le lemme 21.11). Alors on en déduit, en appliquant la troisième assertion du corollaire 21.20, que f_3 est convexe sur C_1 (voir l'exemple 21.6).

21.4. Convexité ou concavité locale

Soient f une fonction définie sur $\mathcal{D}_f \subset \mathbb{R}^2$ et A un point intérieur à \mathcal{D}_f . Si on veut seulement étudier le comportement de la fonction f au voisinage du point A , on l'étudie donc sur une boule ouverte de centre A . Or une boule est un sous-ensemble convexe de \mathbb{R}^2 et on peut donc étudier la convexité ou la concavité de f sur une telle boule. Ceci nous conduit à la notion de convexité ou de concavité locale.

21.4.1. Définitions. —

Définition 21.22. — Soient f une fonction définie sur $\mathcal{D}_f \subset \mathbb{R}^2$ et A un point intérieur à \mathcal{D}_f .

- On dit que f est *localement convexe* au voisinage du point A s'il existe une boule ouverte $\mathcal{B}(A, \alpha) \subset \mathcal{D}_f$ avec $\alpha > 0$ telle que f est convexe sur $\mathcal{B}(A, \alpha)$.
- On dit que f est *localement concave* au voisinage du point A s'il existe une boule ouverte $\mathcal{B}(A, \alpha) \subset \mathcal{D}_f$ avec $\alpha > 0$ telle que f est concave sur $\mathcal{B}(A, \alpha)$.

21.4.2. Condition suffisante. — Si f est une fonction de classe \mathcal{C}^2 sur un ouvert \mathcal{U} contenant A , si r, s, t désignent les fonctions dérivées partielles secondes de f , la fonction $rt - s^2$ est continue sur \mathcal{U} , comme somme et produit de fonctions qui sont continues sur \mathcal{U} . Si la valeur de $rt - s^2$ calculée au point A est non nulle, on en déduit qu'elle reste de signe constant sur une boule ouverte de centre A .

Notons $(rt - s^2)(A)$ la valeur de $rt - s^2$ calculée au point A .

- Si $(rt - s^2)(A) > 0$, alors l'expression $(rt - s^2)$ reste strictement positive sur une boule ouverte de centre A . On peut alors conclure à la convexité ou à la concavité locale de f au voisinage de A .
- Si $(s^2 - rt)(A) = 0$ on ne peut rien dire. Par suite, en utilisant le théorème 20.5 page 84 on obtient le théorème suivant qui ne donne qu'une condition suffisante.

Théorème 21.23. — Soit f une fonction de classe \mathcal{C}^2 définie sur un ouvert \mathcal{U} contenant A . Posons

$$D^2f(A) = \begin{pmatrix} r & s \\ s & t \end{pmatrix},$$

où r , s et t représentent les dérivées partielles calculées au point A .

– Si $rt - s^2 > 0$ et ($r > 0$ ou $t > 0$) alors f est localement convexe au voisinage du point A .

– Si $rt - s^2 > 0$ et ($r < 0$ ou $t < 0$) alors f est localement concave au voisinage du point A .

– Si $rt - s^2 < 0$ alors f n'est ni localement convexe, ni localement concave, au voisinage du point A .

Dans ce théorème les **inégalités sont strictes et les équivalences sont remplacées par des implications**. Si $rt - s^2 = 0$ on ne peut rien dire.

Remarque 21.24. — Les polynômes de degré 2 ont des dérivées partielles secondes constantes (indépendantes du point considéré). Pour ces fonctions la convexité ou concavité locale est donc équivalente à la convexité ou concavité sur \mathbb{R}^2 (voir le corollaire 21.7 page 89).

Exemple 21.25. — Reprenons l'exemple 21.6 page 88 avec $f(x, y) = x^3 + xy^2$. On a $rt - s^2 = 12x^2 - 4y^2$.

– Pour $M_1 = (2, 1)$ on a $rt - s^2 = 48 - 4 > 0$ et $r = 12 > 0$ (ou $t = 4 > 0$). La fonction f est donc localement convexe au voisinage du point M_1 .

– Pour $M_2 = (-3, 2)$ on a $rt - s^2 = 12 \times 9 - 16 > 0$ et $r = -18 < 0$ (ou $t = -6 < 0$) La fonction f est donc localement concave au voisinage du point M_2 .

– Pour $M_3 = (1, 2)$ on a $rt - s^2 = 12 - 16 < 0$. La fonction f n'est ni localement convexe, ni localement concave au voisinage du point M_3 .

CHAPITRE 22

GÉNÉRALITÉS SUR LES EXTREMA

Dans ce chapitre n désigne un entier égal à 1 ou 2. Lorsque $n = 1$, X et \bar{X} désigneront tout simplement des réels x et \bar{x} et pour $n = 2$, X et \bar{X} désigneront des couples de réels : $X = (x, y)$ et $\bar{X} = (\bar{x}, \bar{y})$.

22.1. Définitions

22.1.1. Extrema globaux sur un sous-ensemble. —

Définition 22.1. — Soit f une fonction définie sur $\mathcal{D}_f \subset \mathbb{R}^n$ avec ($n = 1$ ou 2), à valeurs dans \mathbb{R} et soit $E \subset \mathcal{D}_f$.

On dit que f présente un *extremum global sur E en \bar{X}* , si :

$$\forall X \in E, f(X) - f(\bar{X})$$

garde un signe constant.

- Si $f(X) - f(\bar{X}) \leq 0$ alors $f(\bar{X})$ est un *maximum global de f sur E* . On dit aussi que f présente en \bar{X} un maximum global sur E .
- Si $f(X) - f(\bar{X}) \geq 0$ alors $f(\bar{X})$ est un *minimum global de f sur E* . On dit aussi que f présente en \bar{X} un minimum global sur E .

Exemple 22.2. — Soit la fonction $f(x, y) = x^2 + y^2$, on a $\mathcal{D}_f = \mathbb{R}^2 = E$. On a de façon évidente :

$$\forall (x, y) \in \mathbb{R}^2, f(x, y) \geq 0.$$

Or $f(0, 0) = 0$, donc $f(0, 0)$ est un minimum global de f sur \mathbb{R}^2 .

Exemple 22.3. — Soit maintenant $f(x, y) = -x^2 + y$ avec $\mathcal{D}_f = \mathbb{R}^2$ et soit $E = \{(x, y) \in \mathbb{R}^2 \mid y \leq 0\}$. On a dans ce cas :

$$\forall (x, y) \in E, f(x, y) \leq 0.$$

Or $f(0, 0) = 0$, ainsi $f(0, 0)$ est un maximum global de f sur E .

Attention : $f(0, 0)$ n'est pas un maximum global de f sur \mathbb{R}^2 car $f(1, 2) = 1 > f(0, 0)$.

22.1.2. Extrema locaux sur un sous-ensemble. —

Définition 22.4. — Soit f une fonction définie sur $\mathcal{D}_f \subset \mathbb{R}^n$ ($n = 1$ ou 2), à valeurs dans \mathbb{R} et soit $E \subset \mathcal{D}_f$. On dit que f présente un *extremum local ou relatif sur E en $\bar{X} \in E$* , s'il existe $r > 0$, tel que :

$$\forall X \in \mathcal{B}(\bar{X}, r) \cap E, f(X) - f(\bar{X})$$

garde un signe constant.

- Si $f(X) - f(\bar{X}) \leq 0$ alors $f(\bar{X})$ est un *maximum local (ou relatif) de f sur E* . On dit aussi que f présente en \bar{X} un maximum local sur E .
- Si $f(X) - f(\bar{X}) \geq 0$ alors $f(\bar{X})$ est un *minimum local (ou relatif) de f sur E* . On dit aussi que f présente en \bar{X} un minimum local sur E .

Remarque 22.5. — Si la propriété précédente est vérifiée sur $\mathcal{B}(\bar{X}, r)$ alors elle reste vraie sur toute boule ouverte centrée en \bar{X} et de rayon inférieur à r .

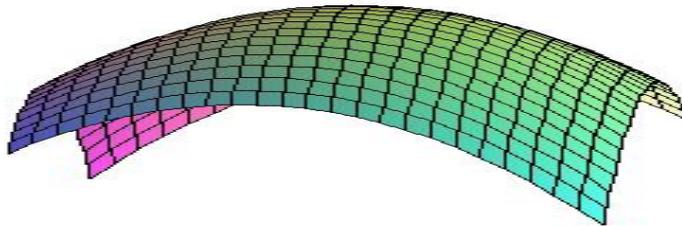


FIGURE 22.1. Un maximum global

Cas particulier où E est un ouvert (non vide) :

Soit E un ensemble ouvert non vide. Par définition, si $\bar{X} \in E$, il existe $r_1 > 0$ tel que $\mathcal{B}(\bar{X}, r_1) \subset E$ (voir la définition d'un ouvert page 52).

D'autre part, si f présente un extremum local sur E en $\bar{X} \in E$, il existe $r_2 > 0$ tel que

$$\forall X \in \mathcal{B}(\bar{X}, r_2) \cap E, f(X) - f(\bar{X})$$

garde un signe constant.

Par suite, si on pose $r = \min(r_1, r_2)$, alors $\mathcal{B}(\bar{X}, r) \subset \mathcal{B}(\bar{X}, r_1) \subset E$ et $\mathcal{B}(\bar{X}, r) \subset \mathcal{B}(\bar{X}, r_2)$ d'où

$$\forall X \in \mathcal{B}(\bar{X}, r), f(X) - f(\bar{X})$$

garde un signe constant.

On peut donc écrire la définition sous la forme équivalente suivante :

Définition 22.6. — Soit f une fonction définie sur $\mathcal{D}_f \subset \mathbb{R}^n$ ($n = 1$ ou 2), à valeurs dans \mathbb{R} et soit $E \subset \mathcal{D}_f$ un ensemble ouvert.

On dit que f présente un extremum local ou relatif sur E en $\bar{X} \in E$, s'il existe $r > 0$, tel que $\mathcal{B}(\bar{X}, r) \subset E$:

$$\forall X \in \mathcal{B}(\bar{X}, r), f(X) - f(\bar{X})$$

garde un signe constant.

- Si $f(X) - f(\bar{X}) \leq 0$ alors $f(\bar{X})$ est un maximum local (ou relatif) de f sur E . On dit aussi que f présente en \bar{X} un maximum local sur E .
- Si $f(X) - f(\bar{X}) \geq 0$ alors $f(\bar{X})$ est un minimum local (ou relatif) de f sur E . On dit aussi que f présente en \bar{X} un minimum local sur E .

En pratique :

- Pour $n = 1$, $\mathcal{B}(\bar{x}, r) =]\bar{x} - r, \bar{x} + r[$ et posons $x - \bar{x} = H$ avec $|H| < r$ (H est un réel).
- Pour $n = 2$, posons $H = X - \bar{X} = (h, k)$ (H est un vecteur).

Dire que $X \in \mathcal{B}(\bar{X}, r)$, revient à dire que $\|H\| < r$. Or, si $|h| < \frac{r}{\sqrt{2}}$ et $|k| < \frac{r}{\sqrt{2}}$ alors $\|H\| < r$.

Avec ces remarques on peut traduire, à nouveau, la définition par :

Définition 22.7. — La fonction f présente un extremum local en \bar{X} sur E (**ensemble ouvert**), si pour $H = (h, k)$, $f(\bar{X} + H) - f(\bar{X})$ garde un signe constant lorsque $\bar{X} + H \in E$ avec $\|H\|$ suffisamment petit ($|h|$ et $|k|$ suffisamment petits dans le cadre de la dimension 2).

22.1.3. Commentaires sur les définitions précédentes. —

1. Il faut distinguer la *valeur de l'extremum* qui est $f(\bar{X})$, du point où l'extremum est atteint \bar{X} . Autrement dit \bar{X} n'est pas un extremum, mais donne un extremum pour f .
2. Lorsque dans les définitions précédentes les inégalités sont strictes pour $X \neq \bar{X}$ on parle d'extremum strict ce qui signifie que l'extremum est atteint en un seul point.
3. Tout extremum global sur E est un extremum local sur E . La réciproque est bien entendu fausse.

On pourra voir, dans le cas de la dimension 1, l'exemple 10.5 page 37 et la figure associée 10.2 proposé dans le chapitre 10. Toutes les définitions et subtilités y sont illustrées. La figure 22.1 représente le graphe d'une fonction en dimension 2 présentant un maximum global.

22.2. Simplification d'un problème d'extremum

Théorème 22.8. — Soient $f : \mathcal{D}_f \subset \mathbb{R}^n \rightarrow \mathbb{R}$, ($n = 1$ ou 2) et $g : I \subset \mathbb{R} \rightarrow \mathbb{R}$. On suppose f continue sur \mathcal{D}_f et g continue et strictement croissante sur I , avec $f(\mathcal{D}_f) \subset I \subset \mathbb{R}$. Alors f et $g \circ f$ admettent sur \mathcal{D}_f des extrema de même nature (minima, maxima, local, global) aux mêmes points.

Preuve. — Nous raisonnerons dans le cas d'un minimum global en \bar{X} sur E , où $E \subset \mathcal{D}_f$.

– Supposons que f admette un minimum global en \bar{X} sur E . Alors :

$$\forall X \in E, f(X) \geq f(\bar{X}).$$

Par hypothèse, $\forall X \in E$, on a $f(X) \in I$, puisque $E \subset \mathcal{D}_f$. Composons l'inégalité par g . Comme g est croissante, le sens de l'inégalité est conservé et on obtient :

$$\forall X \in E, g(f(X)) \geq g(f(\bar{X}))$$

donc $g \circ f$ admet un minimum global en \bar{X} sur E .

– Supposons maintenant que $g \circ f$ admette un minimum global en sur E . Alors :

$$\forall X \in E, g \circ f(X) \geq g \circ f(\bar{X}).$$

Mais comme g est continue et strictement croissante on en déduit que g est une fonction bijective de I sur $g(I)$ et g^{-1} est continue et strictement croissante sur $g(I)$. En composant l'inégalité à gauche par g^{-1} on en déduit :

$$\forall X \in E, f(X) \geq f(\bar{X})$$

donc f admet un minimum global en sur E . □

Application : Si $f > 0$ sur E , alors f et $\ln(f)$ ont sur E des extrema de même nature aux mêmes points.

Exemple 22.9. — Soit la fonction $f(x, y) = x^\alpha y^\beta$, on a $\mathcal{D}_f = \{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0\}$ et pour $\forall (x, y) \in \mathcal{D}_f$, $f(x, y) > 0$. On peut utiliser $g(x, y) = \ln f(x, y) = \alpha \ln(x) + \beta \ln(y)$. Au lieu de rechercher les extrema de f sur \mathcal{D}_f , on pourra rechercher les extrema de g sur \mathcal{D}_f . Les calculs seront plus simples.

CHAPITRE 23

EXTREMA LIBRES POUR DEUX VARIABLES

Dans tout ce chapitre on fait l'hypothèse suivante :

(\mathcal{H}_1) Soit f une fonction définie et continue sur $\mathcal{D}_f \subset \mathbb{R}^2$ et à valeurs dans \mathbb{R} . On considère un ouvert \mathcal{U} , $\mathcal{U} \subset \mathcal{D}_f$.

Problème \mathcal{P}_1 : Déterminer les extrema de f sur \mathcal{U} .

Autrement dit, nous étudions les extrema, s'ils existent, des fonctions de deux variables sur un ouvert. Il s'agit de l'optimisation dite **sans contrainte ou libre**. Nous ne chercherons pas à résoudre le problème posé dans le cas le plus général où les fonctions sont seulement supposées continues.

Nous nous limiterons aux fonctions "assez régulières", qui admettent au moins des dérivées partielles du premier ordre sur \mathcal{U} . En pratique les fonctions étudiées seront en général de classe C^1 ou même C^2 sur \mathcal{U} .

23.1. Conditions nécessaires ou conditions du premier ordre

Théorème 23.1. — Sous l'hypothèse \mathcal{H}_1 , on suppose que f admet en tout point de \mathcal{U} des dérivées partielles du premier ordre.

Si f présente un extremum (local ou global) sur l'ouvert \mathcal{U} en (\bar{x}, \bar{y}) alors :

$$\nabla f(\bar{x}, \bar{y}) = (0, 0) \iff \frac{\partial f}{\partial x}(\bar{x}, \bar{y}) = \frac{\partial f}{\partial y}(\bar{x}, \bar{y}) = 0,$$

où ∇f est le gradient de f

Preuve. — Rappelons qu'un extremum global sur un ouvert est d'abord local (voir page 93). Supposons donc que f présente un extremum local en $(\bar{x}, \bar{y}) \in \mathcal{U}$. Comme \mathcal{U} est ouvert, il existe $r > 0$, tel $\mathcal{B}((\bar{x}, \bar{y}), r) \subset \mathcal{U}$. Donc la fonction f est définie au moins sur $\mathcal{B}((\bar{x}, \bar{y}), r)$.

Si f admet un extremum local en (\bar{x}, \bar{y}) , alors l'application partielle : $x \mapsto f(x, \bar{y})$ admet un extremum local en \bar{x} et l'application partielle : $y \mapsto f(\bar{x}, y)$ admet un extremum local en \bar{y} .

La première application partielle est définie au moins sur l'intervalle ouvert $]\bar{x} - r, \bar{x} + r[$ et on en déduit que la première dérivée partielle s'annule en \bar{x} ainsi

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

Par le même raisonnement on montre que la seconde dérivée partielle s'annule en \bar{y}

$$\frac{\partial f}{\partial y}(\bar{x}, \bar{y}) = 0,$$

d'où le résultat. □

Définition 23.2. — On appelle *point critique* ou *point stationnaire* de f tout point où le gradient est nul.

Exemple 23.3. — Soit $f(x, y) = x^2 + y^2$. La fonction f est de classe C^1 sur \mathbb{R}^2 (voir le graphe de f page 62). On a

$$\nabla f(x, y) = (2x, 2y) = (0, 0) \iff x = y = 0.$$

Or $\forall (x, y) \in \mathbb{R}^2, f(x, y) \geq 0$, on en déduit que $(0, 0)$ donne un minimum (global) pour f sur \mathbb{R}^2 .

Exemple 23.4. — Soit l'ensemble $U = \{(x, y) \in \mathbb{R}^2 \mid 1 - x^2 - y^2 > 0\}$. La fonction

$$f(x, y) = \sqrt{1 - x^2 - y^2}$$

est de classe C^1 sur \mathcal{U} (à justifier).

On a

$$\nabla f(x, y) = \left(\frac{-x}{\sqrt{1 - x^2 - y^2}}, \frac{-y}{\sqrt{1 - x^2 - y^2}} \right) = (0, 0) \iff x = y = 0$$

et $f(0, 0) = 1$. Or, $\forall (x, y) \in U$, $f(x, y) \leq 1$, on en déduit que $(0, 0)$ donne un maximum (global) pour f sur \mathcal{U} .

Exemple 23.5. — Soit la fonction $f(x, y) = x^2 - y^2$. La fonction f est de classe C^1 sur \mathbb{R}^2 (polynôme).

On a $\nabla f(x, y) = (2x, -2y) = (0, 0) \iff x = y = 0$. $f(0, 0) = 0$. Pour $x \neq 0$ on a $f(x, 0) = x^2 > 0$, et pour $y \neq 0$ on a $f(0, y) = -y^2 < 0$. Suivant les directions prises pour aller vers $(0, 0)$, on rencontre en ce point soit un minimum, soit un maximum pour f .

La fonction f n'admet ni maximum, ni un minimum en $(0, 0)$. La condition du théorème 23.1 n'est donc pas suffisante. On pourra voir sur la figure 23.1 le graphique de la fonction f autour du point $(0, 0)$.

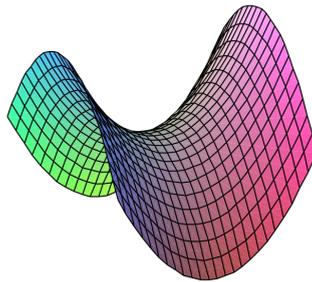


FIGURE 23.1. Un point col

Définition 23.6. — Si (\bar{x}, \bar{y}) est un point critique qui ne donne ni un maximum local, ni un minimum local, pour f sur \mathcal{U} , on dit que f présente un *point selle* ou *point col* en (\bar{x}, \bar{y}) .

Remarque 23.7. — L'hypothèse \mathcal{U} ouvert de \mathbb{R}^2 est indispensable comme le montre l'exemple suivant. Soit la fonction $f(x, y) = x + y$ et soit $U = [0, 1] \times [0, 1]$. Alors $\forall (x, y) \in U$, $f(x, y) \leq 2$ et $f(1, 1) = 2$. Ainsi f admet en $(1, 1)$ un maximum global sur \mathcal{U} . Or, les dérivées partielles de f ne sont pas nulles en $(1, 1)$.

23.2. Nature des points critiques

Comme le montre l'exemple 23.5, les points critiques ne donnent pas forcément un extremum pour f . Ce sont seulement des points candidats et ce sont d'ailleurs les seuls, si f est de classe C^1 sur \mathcal{U} ouvert.

Principe : soit $(\bar{x}, \bar{y}) \in \mathcal{U}$, point critique de f et soit le vecteur (h, k) tel que $(\bar{x} + h, \bar{y} + k) \in \mathcal{U}$. Pour pouvoir conclure que (\bar{x}, \bar{y}) donne bien un extremum pour f sur \mathcal{U} et préciser s'il s'agit d'un maximum ou d'un minimum il faut :

- Si $\forall (h, k)$ tel que $(\bar{x} + h, \bar{y} + k) \in \mathcal{U}$, $\Delta f_{(\bar{x}, \bar{y})}(h, k)$ est de signe constant alors (\bar{x}, \bar{y}) donne un extremum global de f sur \mathcal{U} .
- Si lorsque la norme de (h, k) est petite, c'est-à-dire lorsque h et k sont assez petits, $\Delta f_{(\bar{x}, \bar{y})}(h, k)$ est de signe constant alors (\bar{x}, \bar{y}) donne un extremum local de f sur \mathcal{U} .
- Dans le cas contraire, c'est-à-dire lorsque $\Delta f_{(\bar{x}, \bar{y})}(h, k)$ change de signe, même pour h et k petits, alors (\bar{x}, \bar{y}) donne un point col. On pourra de nouveau consulter la figure 23.1.

L'étude du signe de la différence $\Delta f_{(\bar{x}, \bar{y})}(h, k)$ peut se faire directement dans certains cas simples (voir l'exemple 23.3) mais ce n'est pas en général le cas, et on utilise alors les théorèmes 23.10 ou 23.11 qui sont énoncés dans les pages suivantes.

23.2.1. Étude directe. — Dans certains problèmes, il est inutile de faire appel à des résultats savants, le simple bon sens suffit.

Exemple 23.8. — Soit la fonction $f(x, y) = 10 + (x - y)^4 + (y - 1)^4$. La fonction f est de classe \mathcal{C}^1 sur \mathbb{R}^2 (c'est un polynôme).

Recherchons les points critiques :

$$\nabla f(x, y) = (4(x - y)^3, -4(x - y)^3 + 4(y - 1)^3) = (0, 0) \iff (x, y) = (1, 1).$$

Or $f(1, 1) = 10$ et il est clair que $\forall (x, y) \in \mathbb{R}^2$, on a $f(x, y) \geq 10$. Par suite $(1, 1)$ donne un minimum global pour f sur \mathbb{R}^2 .

Exemple 23.9. — Soit maintenant $f(x, y) = x^4 - y^2$. La fonction f est de classe \mathcal{C}^1 sur \mathbb{R}^2 (c'est un polynôme), $(0, 0)$ est le seul point critique et

$$\Delta f_{(0,0)}(h, k) = f(0 + h, 0 + k) - f(0, 0) = h^4 - k^2.$$

Supposons $|h|$ et $|k|$ suffisamment petits. Si $k = 0$ et $h \neq 0$, alors $\Delta f_{(0,0)}(h, k) = h^4 > 0$, si $h = 0$ et $k \neq 0$, alors $\Delta f_{(0,0)}(h, k) = -k^2 < 0$. Donc $\Delta f_{(0,0)}(h, k)$ change de signe sur tout voisinage de $(0, 0)$. Ainsi le point $(0, 0)$ donne un point col (voir la figure 23.1).

23.2.2. Condition suffisante d'extremum global sur \mathcal{U} ouvert convexe. —

Théorème 23.10 (Extrema des fonctions convexes et concaves). — Soient f une fonction définie sur $\mathcal{D}_f \subset \mathbb{R}^2$ et $\mathcal{U} \subset \mathcal{D}_f$ un ensemble ouvert convexe. On suppose que f est de classe \mathcal{C}^1 sur \mathcal{U} . Soit $(\bar{x}, \bar{y}) \in \mathcal{U}$.

- Si $\nabla f(\bar{x}, \bar{y}) = (0, 0)$ et f convexe sur \mathcal{U} alors f présente en (\bar{x}, \bar{y}) un minimum global \mathcal{U} .
- Si $\nabla f(\bar{x}, \bar{y}) = (0, 0)$ et f concave sur \mathcal{U} alors f présente en (\bar{x}, \bar{y}) un maximum global \mathcal{U} .

Preuve. — Supposons que f soit convexe sur \mathcal{U} . D'après la définition de la convexité, le graphe de f est donc au-dessus du plan tangent passant par (\bar{x}, \bar{y}) . Ainsi on obtient pour tout $(x, y) \in \mathcal{U}$,

$$f(x, y) \geq f(\bar{x}, \bar{y}) + \langle \nabla f(\bar{x}, \bar{y}), (x - \bar{x}, y - \bar{y}) \rangle,$$

ce qui donne, puisque (\bar{x}, \bar{y}) est un point critique de f , $f(x, y) \geq f(\bar{x}, \bar{y})$. Donc f présente en (\bar{x}, \bar{y}) un minimum global sur \mathcal{U} .

Avec f concave les inégalités changent de sens et on trouve un maximum global. □

23.2.3. Condition suffisante d'extremum local ou condition du second ordre. —

Théorème 23.11. — Soit f une fonction de classe \mathcal{C}^2 définie sur un \mathcal{U} ouvert de \mathbb{R}^2 et soit $(\bar{x}, \bar{y}) \in \mathcal{U}$ tel que $\nabla f(\bar{x}, \bar{y}) = (0, 0)$. Notons la matrice hessienne de f au point (\bar{x}, \bar{y}) :

$$D^2 f(\bar{x}, \bar{y}) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

- Si $rt - s^2 > 0$ et si
 - $r > 0$ ou $t > 0$, alors f présente au point (\bar{x}, \bar{y}) un minimum local sur \mathcal{U}
 - $r < 0$ ou $t < 0$, alors f présente au point (\bar{x}, \bar{y}) un maximum local sur \mathcal{U} .
- Si $rt - s^2 < 0$ alors (\bar{x}, \bar{y}) est un point col pour f .

Dans le théorème 23.11 les dérivées partielles secondes sont calculées au point critique. Ce sont donc des nombres réels.

Preuve. — Si $rt - s^2 > 0$ alors la fonction f est localement convexe ou localement concave au voisinage du point (\bar{x}, \bar{y}) , on en déduit qu'elle est convexe ou concave sur une boule ouverte de centre (\bar{x}, \bar{y}) et, en appliquant le théorème 23.10, f admet un extremum global sur cette boule ouverte. Cet extremum est alors au moins un extremum local pour f sur \mathcal{U} , d'où les deux premiers résultats.

Cet extremum est peut-être global sur \mathcal{U} mais il faudra d'autres arguments pour le justifier.

Pour justifier le troisième résultat, écrivons la formule de Taylor-Young à l'ordre 2 (ou un DL d'ordre 2) au voisinage du point critique. Comme le terme du premier ordre est nul on obtient :

$$f(\bar{x} + h, \bar{y} + k) = f(\bar{x}, \bar{y}) + \frac{1}{2}(rh^2 + 2shk + tk^2) + (h^2 + k^2)\varepsilon(h, k)$$

c'est-à-dire : $f(\bar{x} + h, \bar{y} + k) - f(\bar{x}, \bar{y}) = d^2 f_{(\bar{x}, \bar{y})}(h, k) + \text{reste}$.

Si $rt - s^2 < 0$, alors r, s, t ne sont pas nuls tous les trois. On suppose $(h, k) \neq (0, 0)$. Si $d^2 f_{(\bar{x}, \bar{y})}(h, k) \neq 0$ et si les valeurs de $|h|$ et $|k|$ sont suffisamment petites, le reste est négligeable devant le terme du deuxième ordre et

le signe de la différence $f(\bar{x} + h, \bar{y} + k) - f(\bar{x}, \bar{y})$ est celui de $d^2 f_{(\bar{x}, \bar{y})}(h, k)$. D'après le théorème 20.5 page 84, comme $rt - s^2 < 0$, $d^2 f_{(\bar{x}, \bar{y})}(h, k)$ change de signe. Il existe donc (h_1, k_1) et (h_2, k_2) tels que $d^2 f_{(\bar{x}, \bar{y})}(h_1, k_1) > 0$ et $d^2 f_{(\bar{x}, \bar{y})}(h_2, k_2) < 0$.

Or, $\forall \alpha > 0$, $d^2 f_{(\bar{x}, \bar{y})}(\alpha h_1, \alpha k_1) = \alpha^2 d^2 f_{(\bar{x}, \bar{y})}(h_1, k_1)$ et $d^2 f_{(\bar{x}, \bar{y})}(\alpha h_2, \alpha k_2) = \alpha^2 d^2 f_{(\bar{x}, \bar{y})}(h_2, k_2)$. Dans toute boule ouverte de centre (\bar{x}, \bar{y}) il y a des points de la forme $(\bar{x} + \alpha h_1, \bar{y} + \alpha k_1)$ et de la forme $(\bar{x} + \alpha h_2, \bar{y} + \alpha k_2)$ donc $d^2 f_{(\bar{x}, \bar{y})}$ change de signe sur toute boule ouverte de centre (\bar{x}, \bar{y}) même de rayon suffisamment petit.

Par suite (\bar{x}, \bar{y}) donne un point col pour f sur \mathcal{U} . □

Remarque 23.12. — Voici quelques points importants :

1. On déduit du théorème 23.11 que, si une fonction est localement convexe au voisinage d'un point critique alors elle admet un minimum local en ce point, et si une fonction est localement concave au voisinage d'un point critique alors elle admet un maximum local en ce point.
2. La surface représentative d'une fonction localement convexe au voisinage du point (\bar{x}, \bar{y}) a, au voisinage de ce point, l'allure du graphe de $f(x, y) = x^2 + y^2$ au voisinage de $(0, 0)$, voir figure 14.4.1.2 page 62.
3. La surface représentative d'une fonction localement concave au voisinage du point (\bar{x}, \bar{y}) a, au voisinage de ce point, l'allure du graphe de $f(x, y) = -x^2 - y^2$ au voisinage de $(0, 0)$ (renverser la figure 14.4.1.2 page 62).
4. La surface représentative d'une fonction qui admet un point col au point (\bar{x}, \bar{y}) a, au voisinage de ce point, l'allure du graphe de $f(x, y) = x^2 - y^2$ au voisinage de $(0, 0)$, voir figure 23.1 page 98.

Exemple 23.13. — Nous cherchons à optimiser la fonction $f(x, y) = x^3 + y^3 - 3xy$ sur \mathbb{R}^2 . La fonction f est de classe C^2 sur \mathbb{R}^2 (c'est un polynôme).

Recherche des points critiques :

On a

$$\nabla f(x, y) = (3x^2 - 3y, 3y^2 - 3x) = (0, 0) \iff (x = y = 0) \text{ ou } (x = y = 1)$$

On obtient donc deux points critiques : $A = (0, 0)$ avec $f(A) = 0$ et $B = (1, 1)$ avec $f(B) = -1$.

Nature des points critiques : On obtient $r = 6x$, $t = 6y$ et $s = -3$.

Pour le point $A = (0, 0)$ on a $rt - s^2 = -9 < 0$, ce point donne un point col ou point selle.

Pour le point $B = (1, 1)$ on a $rt - s^2 = 36 - 9 = 27 > 0$ et $r = 6 > 0$, ce point donne un minimum local sur \mathbb{R}^2 . Ce minimum n'est pas global sur \mathbb{R}^2 car $f(-1, -1) = -5 < f(B)$.

Remarque 23.14. — Le théorème 23.11 ne donne aucune conclusion lorsque $rt - s^2 = 0$ car, dans ce cas, tout peut se produire : maximum, minimum, point col. Dans une telle situation :

1. On écrit le DL d'ordre 2 au point critique (\bar{x}, \bar{y}) : $f(\bar{x} + h, \bar{y} + k) - f(\bar{x}, \bar{y}) = d^2 f_{(\bar{x}, \bar{y})}(h, k) + \text{reste}$.
2. On suppose les valeurs de $|h|$ et $|k|$ suffisamment petites.
 - (a) pour les valeurs de h et k telles que $d^2 f_{(\bar{x}, \bar{y})}(h, k) \neq 0$, on néglige le reste et le signe de $f(\bar{x} + h, \bar{y} + k) - f(\bar{x}, \bar{y})$ est celui de $d^2 f_{(\bar{x}, \bar{y})}(h, k)$.
 - (b) pour les valeurs de h et k telles que $d^2 f_{(\bar{x}, \bar{y})}(h, k) = 0$ on étudie directement le signe de la différence $f(\bar{x} + h, \bar{y} + k) - f(\bar{x}, \bar{y})$.
3. On conclut en comparant les résultats des cas (a) et (b).

Exemple 23.15. — Dans l'exemple 23.9 on aurait pu aussi raisonner comme suit.

La fonction f est de classe C^2 sur \mathbb{R}^2 (c'est un polynôme) et $O = (0, 0)$ est le seul point critique.

On trouve $r = 12x^2$, $s = 0$ et $t = -2$.

Pour le point $(0, 0)$ on a $rt - s^2 = 0$. Le théorème 23.11 ne permet pas de conclure.

1. Écrivons le DL d'ordre 2 au point critique $(0, 0)$: $f(0 + h, 0 + k) - f(0, 0) = -k^2 + \text{reste}$
2. Supposons les valeurs de $|h|$ et $|k|$ suffisamment petites.
3. Si $k \neq 0$, on peut négliger le reste et écrire $f(0 + h, 0 + k) - f(0, 0) \simeq -k^2$. La différence est donc négative dans ce cas.
4. Si $k = 0$, $f(0 + h, 0) - f(0, 0) = \text{reste}$ et il faudrait étudier le signe du reste qui est inconnu. On revient alors à l'étude directe pour ce cas particulier : $f(0 + h, 0) - f(0, 0) = h^4 \geq 0$. Dans ce cas la différence est positive.
5. Sur tout voisinage de $(0, 0)$ la différence $f(0 + h, 0 + k) - f(0, 0)$ change de signe et par suite f admet un point col en $(0, 0)$.

23.3. Quelques conseils pour la résolution du système des conditions nécessaires du premier ordre

Il faut, pour trouver les points critiques résoudre le système suivant

$$(23.1) \quad \begin{cases} (x, y) \in \mathcal{U} \\ \frac{\partial f}{\partial x}(x, y) = 0 \\ \frac{\partial f}{\partial y}(x, y) = 0. \end{cases}$$

Le système (23.1) est un système de deux équations à deux inconnues x et y , qui n'est pas toujours facile à résoudre. Dans la plupart des exemples il est non-linéaire et contrairement aux systèmes linéaires qui ont généralement une seule solution, le système (23.1) a souvent plusieurs solutions comme on peut le constater dans les exemples. Une des difficultés de la résolution est qu'il ne faut ni perdre des solutions, ni en ajouter.

Remarque 23.16. — La condition $(x, y) \in \mathcal{U}$ n'est habituellement pas indiquée dans le système (23.1). Elle est inutile si $\mathcal{U} = \mathbb{R}^2$, mais, dans le cas contraire, sa présence évitera de prendre en compte des solutions qui n'appartiennent pas à \mathcal{U} .

Voici quelques méthodes dont les deux premières ont déjà été utilisées dans le cadre des systèmes linéaires. Ces méthodes sont *a priori* bien connues.

1. *Substitution* : L'idée est très simple mais pas toujours applicable. Si on le peut, on exprime l'une des inconnues en fonction de l'autre en utilisant l'une des deux équations et on remplace dans l'équation non utilisée. On est alors ramené à résoudre une équation à une variable potentiellement plus simple.
2. *Combinaisons linéaires des équations* : L'idée est de multiplier une ligne par un coefficient a et l'autre par b et ensuite additionner les deux lignes. On obtient ainsi de l'information supplémentaire.
3. *Multiplication d'une équation par x ou y* : C'est la même méthode que précédemment mais au lieu de multiplier les équations par un nombre réel on multiplie par une fonction. Par exemple x ou alors y ...

Conclusion : Toutes les méthodes algébriques sont bonnes pour trouver le résultat. A priori les systèmes proposés ici ont une solution en quelques lignes seulement, à condition de trouver la méthode.

CHAPITRE 24

PROPRIÉTÉS DU GRADIENT

24.1. Équations des tangentes aux courbes de niveau

Exemple 24.1. — Soit la fonction $f(x, y) = xy$, f est un polynôme donc de classe \mathcal{C}^1 sur \mathbb{R}^2 (rappelons que c'est bien un ensemble ouvert). Si $k \neq 0$, la courbe C_k de niveau k a pour équation : $xy = k$ (donc $xy \neq 0$) ce qui s'écrit aussi : $y = k/x$.

Soit $M_0 = (x_0, y_0) \in C_k$. Alors $k = x_0 y_0$ avec x_0 et y_0 non nuls.

La tangente à la courbe C_k au point M_0 a pour équation :

$$y - y_0 = -\frac{y_0}{x_0}(x - x_0).$$

On peut simplifier par x_0 et écrire l'équation précédente sous la forme :

$$(24.1) \quad y_0(x - x_0) + x_0(y - y_0) = 0$$

D'autre part on constate que $(y_0, x_0) = \nabla f(x_0, y_0)$. Donc $\nabla f(M_0) \neq 0$. Donc si $M = (x, y)$ est un point de la tangente à la courbe C_k au point M_0 on a : $\overrightarrow{M_0 M} = (x - x_0, y - y_0)$. L'égalité (24.1) signifie que les vecteurs $\overrightarrow{M_0 M}$ et $\nabla f(M_0)$ sont orthogonaux.

Le résultat vérifié sur l'exemple (24.1) est général d'où le théorème suivant que nous admettrons.

Théorème 24.2. — Soit f une fonction de classe \mathcal{C}^1 sur un ouvert \mathcal{U} contenant le point $M_0 = (x_0, y_0)$. Le gradient de f au point M_0 est orthogonal à la tangente en M_0 à la courbe de niveau passant par M_0 .

On en déduit le corollaire suivant :

Corollaire 24.3. — Soit f une fonction de classe \mathcal{C}^1 sur un ouvert \mathcal{U} .

On considère la courbe de niveau k , notée C_k , que l'on suppose non vide et soit $M_0 = (x_0, y_0)$ un point de C_k . Si $\nabla f(M_0) \neq 0$ alors l'équation de la tangente à la courbe C_k au point M_0 s'écrit :

$$\frac{\partial f}{\partial x}(M_0)(x - x_0) + \frac{\partial f}{\partial y}(M_0)(y - y_0) = 0.$$

Remarque 24.4. — Si $\frac{\partial f}{\partial y}(M_0) = 0$ alors la tangente à la courbe C_k au point M_0 est une droite verticale d'équation $x = x_0$.

En effet, si $\frac{\partial f}{\partial y}(M_0) = 0$ l'équation s'écrit $(x - x_0) = 0$. Notons que dans ce cas on a toujours $\frac{\partial f}{\partial x}(M_0) \neq 0$.

24.2. Direction d'accroissement optimal d'une fonction

Proposition 24.5. — Soit f une fonction de classe \mathcal{C}^1 sur un ouvert \mathcal{U} et soit $M_0 = (x_0, y_0)$ un point de \mathcal{U} . Considérons $H = (h, k)$ un vecteur de norme constante égale à r ($r > 0$).

Alors l'accroissement

$$\Delta f_{M_0}(h, k) = f(x_0 + h, y_0 + k) - f(x_0, y_0)$$

sera le plus grand possible si $(h, k) = \lambda \nabla f(M_0)$ avec $\lambda > 0$.

On dit que le gradient est dirigé vers les f croissants.

Autrement dit, si on considère toutes les valeurs prises par $\Delta f_{M_0}(h, k)$ pour $\|(h, k)\|$ constante, la valeur la plus grande sera obtenue lorsque le vecteur (h, k) sera proportionnel au gradient avec un coefficient positif de proportionnalité. Cette propriété est utilisée dans les algorithmes de recherche d'extrema.

Preuve. — Soit $H = (h, k)$ avec $\|H\| = r$ (donc $r > 0$). On se place en un point $M_0 = (x_0, y_0)$ tel que $\nabla f(M_0) \neq 0$ et l'on va faire varier x de x_0 à $x_0 + h$ et y de y_0 à $y_0 + k$.

On sait (voir le corollaire 17.12 page 73) que pour r assez petit on a

$$\Delta f_{M_0}(h, k) = f(x_0 + h, y_0 + k) - f(x_0, y_0) \simeq df_{M_0}(h, k).$$

Ceci donne, en utilisant le produit scalaire :

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) \simeq \langle \nabla f(M_0), (h, k) \rangle.$$

Utilisons l'inégalité de Cauchy-Schwarz. On a :

$$|\langle \nabla f(M_0), (h, k) \rangle| \leq \|\nabla f(M_0)\| \|(h, k)\|.$$

Comme $\|(h, k)\| = r$, on obtient : $\langle \nabla f(M_0), (h, k) \rangle \leq r \|\nabla f(M_0)\|$. La valeur absolue

$$|\langle \nabla f(M_0), (h, k) \rangle|$$

sera la plus grande si $|\langle \nabla f(M_0), (h, k) \rangle| = r \|\nabla f(M_0)\|$.

Dans ce cas d'égalité, voir page proposition 11.16 page 45, puisque les deux vecteurs $\nabla f(M_0)$ et (h, k) sont non nuls, il existe un réel $\lambda \neq 0$ tel que $H = \lambda \nabla f(M_0)$ c'est-à-dire $h = \lambda \frac{\partial f}{\partial x}(M_0)$ et $k = \lambda \frac{\partial f}{\partial y}(M_0)$.

Nous obtenons dans ce cas :

$$\langle \nabla f(M_0), (h, k) \rangle = h \frac{\partial f}{\partial x}(M_0) + k \frac{\partial f}{\partial y}(M_0) = \lambda \left(\left(\frac{\partial f}{\partial x}(M_0) \right)^2 + \left(\frac{\partial f}{\partial y}(M_0) \right)^2 \right) = \lambda r^2,$$

donc $\Delta f_{M_0} \simeq \lambda r^2$.

On peut résumer ce qui précède de la façon suivante : pour tout accroissement (h, k) de norme r (petit), on atteint une courbe de niveau plus élevé si $(h, k) = \lambda \nabla f(M_0)$ avec $\lambda > 0$. \square

Le raisonnement précédent suppose que r est "petit". En pratique si r est grand, on fait une succession de (petites) variations proportionnelles au gradient obtenu à chaque étape.

CHAPITRE 25

EXTREMA LIÉS DES FONCTIONS DE DEUX VARIABLES

Dans tout ce chapitre on fait l'hypothèse suivante :

(\mathcal{H}_2) Soit f une fonction définie et continue sur $\mathcal{D}_f \subset \mathbb{R}^2$ et à valeurs dans \mathbb{R} . On considère un **ouvert** \mathcal{U} , $\mathcal{U} \subset \mathcal{D}_f$.
Soit g une fonction définie sur $D_g \subset \mathbb{R}^2$ et à valeurs dans \mathbb{R} et soit $\mathcal{U} \subset D_g$. On suppose que f et g sont au moins de classe \mathcal{C}^1 sur \mathcal{U} .

Problème \mathcal{P}_2 : Déterminer les extrema de f sur \mathcal{U} lorsque $g(x, y) = 0$.

Posons $E = \{(x, y) \in \mathcal{U} \mid g(x, y) = 0\}$. Cet ensemble **n'est pas un ouvert**. Le problème (\mathcal{P}_2) revient à rechercher les extrema de f sur l'ensemble E .

Autrement dit, nous étudions dans ce chapitre l'optimisation des fonctions de deux variables lorsque ces variables sont liées par une contrainte de type égalité, appelée aussi "liaison".

25.1. Cas d'une liaison explicite

On dit que la **liaison est explicite** si on peut exprimer une des variables en fonction de l'autre. On se ramène alors à un problème d'extremum sans contrainte à **une** variable (voir chapitre 10).

Exemple 25.1. — $f(x, y) = x^2 + y$. Optimiser f sous la contrainte $g(x, y) = x^3 + y = 0$.

Les fonctions f et g sont des fonctions polynômes donc de classe \mathcal{C}^2 sur \mathbb{R}^2 .

La contrainte s'écrit sous la forme équivalente : $y = -x^3$. Ainsi nous posons de façon naturelle

$$F(x) = f(x, -x^3) = x^2 - x^3.$$

Optimiser f sous la contrainte $g(x, y) = 0$ revient alors à optimiser F sur \mathbb{R} . Nous sommes ramenés à un problème à une variable. Or $F'(x) = 2x - 3x^2$. La fonction F admet comme points critiques $x = 0$ et $x = 2/3$. On a $F''(x) = 2 - 6x$, donc $F''(0) = 2 > 0$, $F''(2/3) = -2 < 0$. Par suite F admet sur \mathbb{R} un minimum local en 0 et un maximum local en $2/3$.

On en déduit que la fonction f admet sous la contrainte $x^3 + y = 0$, un minimum local en $(0, 0)$ de valeur $f(0, 0) = 0$ et un maximum local en $(2/3, -8/27)$ de valeur $f(2/3, -8/27) = 4/27$.

Traçons les courbes de niveau de f passant par chacun des extrema et la courbe représentative G de la contrainte. La courbe C_k de niveau k , de f , est la parabole d'équation : $y = -x^2 + k$ (voir figure 25.1).

Géométriquement, pour rechercher le maximum de f sous la contrainte, il nous faut trouver la courbe de niveau k le plus élevé possible rencontrant la courbe G . Mais cette courbe de niveau ne peut pas couper la courbe G , car s'il en était ainsi, il existerait d'autres courbes de niveau, de valeur plus élevée, coupant également la courbe G . En d'autres termes, la courbe de niveau maximale doit rencontrer la courbe G sans la couper, ce qui revient à dire qu'elle doit être tangente à la courbe G au point correspondant au maximum. Donc, en ce point, les tangentes aux deux courbes sont confondues. **De même pour rechercher le minimum** de f sous la contrainte, il nous faut trouver la courbe de niveau k , le moins élevé possible, tangente à la courbe G .

Vérifions par le calcul les affirmations précédentes :

- La tangente à la courbe G au point (a, b) a pour équation : $y - b = -3a^2(x - a)$.
- La tangente à la courbe de niveau C_k au point (a, b) a pour équation : $2a(x - a) + (y - b) = 0$.
- Au point $(0, 0)$ ces deux équations sont confondues et s'écrivent $y = 0$.

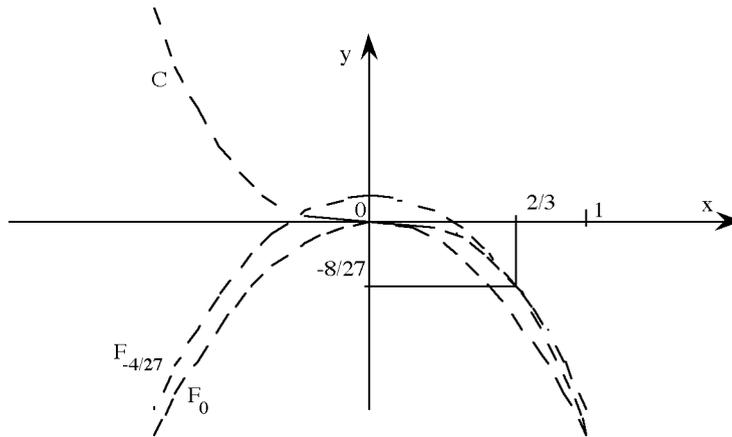


FIGURE 25.1

– Au point $(2/3, -8/27)$ ces deux équations sont confondues et s'écrivent : $y + 8/27 = -4/3(x - 2/3)$.

25.2. Conditions nécessaires ou conditions du premier ordre

Nous allons admettre que le résultat géométrique vu dans l'exemple précédent est général : si (\bar{x}, \bar{y}) est solution du problème (\mathcal{P}_2) et si $m = f(\bar{x}, \bar{y})$, alors la courbe G représentative de la contrainte et la courbe de niveau C_m ($k = m$) de f ont la même tangente au point (\bar{x}, \bar{y}) . Ce résultat ne s'applique que si on peut écrire l'équation de la tangente à la courbe G au point (\bar{x}, \bar{y}) .

Cela suppose d'après le corollaire 24.3 page 103 que **les dérivées partielles de g ne sont pas toutes les deux nulles en (\bar{x}, \bar{y})** . Supposons donc que ce soit le cas.

Nous avons vu que le **vecteur gradient de f au point (\bar{x}, \bar{y}) était orthogonal à la tangente** à la courbe de niveau passant par ce point.

De même, le vecteur gradient de g au point (\bar{x}, \bar{y}) est orthogonal à la tangente à la courbe G , puisque (\bar{x}, \bar{y}) appartient à G , courbe de niveau 0 de g .

Comme les deux tangentes au point (\bar{x}, \bar{y}) sont confondues, on en déduit que les deux vecteurs gradients au point (\bar{x}, \bar{y}) sont proportionnels : il existe donc un réel λ tel que :

$$\frac{\partial f}{\partial x}(\bar{x}, \bar{y}) = \lambda \frac{\partial g}{\partial x}(\bar{x}, \bar{y}) \text{ et } \frac{\partial f}{\partial y}(\bar{x}, \bar{y}) = \lambda \frac{\partial g}{\partial y}(\bar{x}, \bar{y}),$$

ou de façon équivalente

$$\nabla f(\bar{x}, \bar{y}) = \lambda \nabla g(\bar{x}, \bar{y}).$$

Nous pouvons ainsi énoncer le théorème fondamental suivant qui tient compte des deux cas :

Théorème 25.2 (Théorème des extréma liés). — Soient f et g deux fonctions de classe \mathcal{C}^1 sur un ouvert $\mathcal{U} \subset \mathbb{R}^2$.

Si (\bar{x}, \bar{y}) est un extremum (local ou global) pour f sur \mathcal{U} et sous la contrainte $g(x, y) = 0$, alors il vérifie l'une des deux conditions suivantes :

1. $g(\bar{x}, \bar{y}) = 0$ et $\nabla g(\bar{x}, \bar{y}) = 0$
2. $g(\bar{x}, \bar{y}) = 0$, $\nabla g(\bar{x}, \bar{y}) \neq 0$, et $\exists \lambda \in \mathbb{R}$, $\nabla f(\bar{x}, \bar{y}) - \lambda \nabla g(\bar{x}, \bar{y}) = 0$

On appelle **Lagrangien** du problème (\mathcal{P}_2) la fonction \mathcal{L} de deux variables définie sur \mathcal{U} par : $\mathcal{L}(x, y) = f(x, y) - \lambda g(x, y)$

On peut, sans aucun problème, prendre $\mathcal{L}(x, y) = f(x, y) + \lambda g(x, y)$. Les valeurs trouvées pour λ seront tout simplement opposées.

Les économistes préfèrent la première forme à cause d'une interprétation économique de λ .

Bien entendu les fonctions \mathcal{L} et f coïncident sur le sous-ensemble

$$E = \{(x, y) \in \mathcal{U} \mid g(x, y) = 0\}.$$

Définition 25.3. — On suppose que (\bar{x}, \bar{y}) vérifie l'une des conditions du théorème 25.2.

- Dans le premier cas on dit que (\bar{x}, \bar{y}) est **un point critique de deuxième espèce** du problème (\mathcal{P}_2) ou que **la contrainte n'est pas qualifiée** en (\bar{x}, \bar{y}) .
 - Dans le deuxième cas on dit que (\bar{x}, \bar{y}) est un **point critique de première espèce** du problème (\mathcal{P}_2) ou que **la contrainte est qualifiée** en (\bar{x}, \bar{y}) .
- Dans ce cas, λ est le *multiplicateur de Lagrange* associé au point (\bar{x}, \bar{y}) .

Remarque 25.4. — – Les points critiques vérifient toujours $g(\bar{x}, \bar{y}) = 0$.

- Les points critiques de deuxième espèce sont en plus des points critiques pour la fonction g associée à la contrainte.
- Les points critiques de première espèce sont les points critiques du Lagrangien et ne sont pas des points critiques de g .

Mise en œuvre pratique du théorème 25.2.

1. On recherche s'il existe **des points critiques de seconde espèce**. Ces points vérifient le système ci-dessous à deux inconnues x et y .

$$\left\{ \begin{array}{l} (x, y) \in \mathcal{U}, \\ g(x, y) = 0, \\ \frac{\partial g}{\partial x}(x, y) = 0, \\ \frac{\partial g}{\partial y}(x, y) = 0. \end{array} \right.$$

En général ce système n'a pas de solution car il y a plus d'équations que d'inconnues.

2. On recherche **les points critiques de première espèce**. On forme alors le **Lagrangien** du problème : $\mathcal{L}(x, y) = f(x, y) - \lambda g(x, y)$ (ou $+\lambda$ suivant la convention) Ces points ne sont pas des points critiques de seconde espèce et vérifient le système ci-dessous de trois équations à trois inconnues x , y et λ :

$$\left\{ \begin{array}{l} (x, y) \in \mathcal{U}, \\ \frac{\partial \mathcal{L}}{\partial x}(x, y) = 0, \\ \frac{\partial \mathcal{L}}{\partial y}(x, y) = 0, \\ g(x, y) = 0, \\ \nabla g(x, y) \neq 0. \end{array} \right.$$

Remarque 25.5. — 1. On considère parfois le Lagrangien comme une fonction de trois variables (x, y, λ) . Dans ce cas l'équation $g(x, y) = 0$ s'écrit $\frac{\partial \mathcal{L}}{\partial \lambda}(x, y, \lambda) = 0$. Nous ne le ferons pas dans ce cours.

2. La valeur de λ peut être nulle : cela signifie qu'un point critique de f (candidat éventuel pour l'optimisation sans contrainte) vérifie la contrainte.
3. Pour la résolution des systèmes on peut utiliser les conseils vus au paragraphe 23.3 du chapitre 23 qui se généralisent à trois variables.

Exemple 25.6. — Le but est d'optimiser $f(x, y) = x$ sur \mathbb{R}^2 et sous la contrainte $g(x, y) = x^2 + y^2 - 2 = 0$. Les fonctions f et g sont des fonctions polynômes donc de classe \mathcal{C}^2 sur \mathbb{R}^2 .

→ **Recherchons les points critiques de seconde espèce.** Ces points vérifient :

$x^2 + y^2 - 2 = 0$ et $\nabla g(x, y) = (2x, 2y) = (0, 0)$. Or seul le point $(0, 0)$ annule le gradient de g et comme $g(0, 0) \neq 0$, on en déduit qu'il n'y a pas de point critique de seconde espèce.

→ **Recherchons les points critiques de première espèce.**

Toutes les solutions du problème seront donc des points critiques de première espèce. Le Lagrangien du problème est :

$$\mathcal{L}(x, y) = x - \lambda(x^2 + y^2 - 2).$$

Le système des conditions nécessaires s'écrit :

$$\begin{cases} (x, y) \in \mathbb{R}^2, \\ 1 - 2\lambda x = 0 \\ 2\lambda y = 0 \\ x^2 + y^2 = 0. \end{cases}$$

Remarquons que la dernière équation $\nabla g(x, y) \neq 0$ n'a aucun sens ici car nous savons qu'il n'y a pas de point critique de seconde espèce.

La deuxième équation donne $\lambda = 0$ ou $y = 0$. Si $\lambda = 0$, la première équation est impossible ainsi on a forcément $y = 0$. On reporte donc dans la dernière équation et on obtient $x = \sqrt{2}$ ou $x = -\sqrt{2}$. La première équation donne alors $\lambda = 1/(2x)$.

On trouve finalement deux points critiques de première espèce :

$$\begin{cases} A = (\sqrt{2}, 0) \text{ avec } \lambda = 1/(2\sqrt{2}), \\ B = (-\sqrt{2}, 0) \text{ avec } \lambda = -1/(2\sqrt{2}). \end{cases}$$



Il ne faut jamais oublier de donner le coefficient de Lagrange λ dans un point critique de première espèce.

On verra aux paragraphes suivants comment préciser la nature des points trouvés. Mais on peut faire ici une résolution géométrique et voir que A donne un maximum global pour f sous la contrainte et B un minimum global sous la contrainte.

Exemple 25.7. — Nous cherchons à optimiser $f(x, y) = xy$ sur \mathbb{R}^2 sous la contrainte $g(x, y) = x^2 + y^2 + 3xy = 0$. Les fonctions f et g sont des fonctions polynômes donc de classe \mathcal{C}^2 sur \mathbb{R}^2 .

→ **Recherchons les points critiques de seconde espèce.** Ces points vérifient :

$$\begin{cases} x^2 + y^2 + 3xy = 0 \\ \nabla g(x, y) = (2x + 3y, 2y + 3x) = (0, 0). \end{cases}$$

Or seul le point $(0, 0)$ annule le gradient de g et comme $g(0, 0) = 0$, on en déduit qu'il y a un point critique de seconde espèce.

Nous pouvons déjà remarquer que la contrainte implique : $xy = -(x^2 + y^2)/3 \leq 0$. Au point $(0, 0)$ on a $f(0, 0) = 0$ et puisque pour tout couple (x, y) vérifiant la contrainte on a $f(x, y) \leq f(0, 0)$, le point $(0, 0)$ donne un maximum global pour f sous la contrainte.

→ **Recherchons les points critiques de première espèce**

Toutes les solutions du problème, différentes de $(0, 0)$, seront donc des points critiques de première espèce. Le Lagrangien du problème est : $\mathcal{L}(x, y) = xy - \lambda(x^2 + y^2 + 3xy)$. Nous cherchons à résoudre le problème suivant pour $(x, y) \in \mathbb{R}^2$ avec $(x, y) \neq (0, 0)$

$$\begin{cases} y - \lambda(2x + 3y) = 0 \\ x - \lambda(2y + 3x) = 0 \\ x^2 + y^2 + 3xy = 0 \end{cases}$$

On résout ce système en remplaçant la première équation par la différence entre les deux premières. On trouve qu'il n'y a pas de solution à part $(0, 0)$. Le problème est donc résolu.

25.3. Optimisation sur un compact

On a un théorème analogue à celui des fonctions continues d'une variable sur un intervalle fermé borné. Le théorème suivant sera admis. Ce théorème est l'analogue en dimension 2 du théorème 3.16 page 15.

Théorème 25.8 (Un théorème fondamental). — *Toute fonction continue sur une partie E fermée bornée (compacte) de \mathbb{R}^2 et à valeurs dans \mathbb{R} , admet un maximum global et un minimum global sur E , c'est-à-dire : $\exists M_1 \in E$ et $\exists M_2 \in E$ tels que :*

$$\forall X \in E, f(M_1) \leq f(X) \leq f(M_2).$$

Dans le cadre du problème (P_2) , la fonction f envisagée est au moins de classe C^1 sur \mathcal{U} , donc continue sur \mathcal{U} , donc continue sur E inclus dans \mathcal{U} . Elle vérifie l'hypothèse du théorème 25.8. Par suite si E est compact on est sûr que les extrema globaux existent.

Méthode pour résoudre un problème d'optimisation sur un fermé, borné :

1. Justifier que E est fermé borné et que f est continue sur E
2. On cherche, en premier lieu, les extrema de la fonction sur "l'intérieur" de E ce qui est un problème d'optimisation sur un ouvert.
3. On recherche les extrema de la fonction sur "la frontière" de E , donc lorsque les variables sont liées par des contraintes d'égalités.
4. On compare les différentes valeurs prises par f aux différents points candidats trouvés ce qui permet de déterminer le ou les points qui donnent le maximum global et le minimum global de la fonction sur E .

Reprenons l'exemple 25.6. Posons $E = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 - 2 = 0\}$. L'ensemble E est un cercle donc E est un ensemble fermé et borné. On a vu que seuls les points $A = (\sqrt{2}, 0)$ avec $\lambda = 1/(2\sqrt{2})$ et $B = (-\sqrt{2}, 0)$ avec $\lambda = -1/(2\sqrt{2})$ sont des points critiques de première espèce et dont pouvaient donner des extrema pour f sous la contrainte $g(x, y) = 0$.

Or $f(A) = \sqrt{2}$ et $f(B) = -\sqrt{2}$. On déduit donc du théorème 25.8 page 108 que le maximum global sous la contrainte est atteint en A et le minimum global en B .

Exemple 25.9. — Le but maintenant est d'optimiser $f(x, y) = xy$ sur E où

$$E = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = x^2 + y^2 - 2 \leq 0\}.$$

Les fonctions f et g sont des fonctions polynômes donc de classe C^2 sur \mathbb{R}^2 . L'ensemble E est la boule fermée de centre $(0, 0)$ et de rayon $\sqrt{2}$ donc E est un ensemble compact.

L'intérieur de E est

$$\overset{\circ}{E} = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = x^2 + y^2 < 2\}.$$

Le seul point critique de f sur $\overset{\circ}{E}$, qui est un ouvert, est $(0, 0)$. Or ce point donne un point col pour f ($r = t = 0$ et $s = 1$). Il n'y a pas donc pas d'extremum sur l'intérieur de E .

On recherche maintenant des extrema sur la frontière de E . On a

$$Fr(E) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 - 2 = 0\}.$$

On a $\nabla g(x, y) = (2x, 2y)$ nul en $(0, 0)$, point qui ne vérifie pas la contrainte donc pas de point critique de deuxième espèce.

Le Lagrangien s'écrit : $\mathcal{L}(x, y) = xy - \lambda(x^2 + y^2 - 2)$. Le système des conditions nécessaires est :

$$\begin{cases} y - 2\lambda x = 0 \\ x - 2\lambda y = 0 \\ x^2 + y^2 = 2 \end{cases} \iff \begin{cases} y = 2\lambda x \\ x(1 - 4\lambda^2) = 0 \\ x^2 + y^2 = 2 \end{cases}$$

La deuxième équation donne $x = 0$ ou $1 - 4\lambda^2 = 0$. La solution $x = 0$ entraîne $y = 0$ mais $(0, 0)$ ne vérifie pas la contrainte. Seule la solution $\lambda = 1/2$ convient. Avec la contrainte elle conduit aux solutions :

$$\begin{cases} A = (1, 1), & \lambda = 1/2 \\ B = (-1, -1), & \lambda = 1/2 \\ C = (1, -1), & \lambda = -1/2 \\ D = (-1, 1), & \lambda = -1/2 \end{cases}$$

Or, comme la fonction f est continue sur \mathbb{R}^2 , donc sur E , on peut appliquer le théorème 25.8 : $f(A) = f(B) = 1$ et $f(C) = f(D) = -1$. Par suite A et B donnent un maximum global pour f sous la contrainte, C et D donnent un minimum global. Ici les extrema ne sont pas stricts car ils sont atteints en deux points différents.

Remarque 25.10. — Si, dans un problème d'optimisation sur un compact, il y a d'autres points candidats, les conditions suffisantes pour ces points s'étudient comme dans le paragraphe suivant. Ces points ne pourront, au mieux, que donner des extrema locaux.

25.4. Conditions suffisantes

25.4.1. Conditions suffisantes sur le Lagrangien. — Ce paragraphe n'est valable que pour les **points critiques de première espèce** qui sont donc des points critiques du Lagrangien. Nous allons voir tout l'intérêt de la notion de Lagrangien.

25.4.1.1. *Théorème fondamental.* —

Théorème 25.11. — Soit (\bar{x}, \bar{y}) un point critique de première espèce du problème (\mathcal{P}_2) , soit λ le multiplicateur de Lagrange associé et soit $\mathcal{L}(x, y) = f(x, y) - \lambda g(x, y)$.

Si (\bar{x}, \bar{y}) est un extremum global (resp. local) du Lagrangien \mathcal{L} sur \mathcal{U} alors (\bar{x}, \bar{y}) est aussi un extremum global (resp. local) de même nature (max, min) de f sur \mathcal{U} sous la contrainte $g(x, y) = 0$.

Preuve. — Plaçons nous dans le cas d'un maximum global sur \mathcal{U} .

Soit $\mathcal{L}(x, y) = f(x, y) - \lambda g(x, y)$ puisque (\bar{x}, \bar{y}) est un maximum pour \mathcal{L} , on a

$$\forall (x, y) \in \mathcal{U}, \mathcal{L}(x, y) \leq \mathcal{L}(\bar{x}, \bar{y}),$$

ce qui s'écrit par définition de \mathcal{L} :

$$f(x, y) - \lambda g(x, y) \leq f(\bar{x}, \bar{y}) - \lambda g(\bar{x}, \bar{y}).$$

L'inégalité est vraie pour tout (x, y) dans \mathcal{U} , donc *a fortiori* pour tout (x, y) dans \mathcal{U} vérifiant la contrainte. Si $g(x, y) = g(\bar{x}, \bar{y}) = 0$, on tire de l'inégalité ci-dessus : $f(x, y) \leq f(\bar{x}, \bar{y})$.

Ainsi (\bar{x}, \bar{y}) donne bien un maximum global pour f sur \mathcal{U} sous la contrainte $g(x, y) = 0$. □

Remarque 25.12. — Le théorème 25.11 est fondamental. Il montre tout l'intérêt de la notion de Lagrangien. Les points critiques du Lagrangien sont des points candidats pour le problème (\mathcal{P}_2) . Pour étudier la nature de ces points critiques on peut utiliser les théorèmes 23.10 page 99 et 23.11 page 99 du chapitre 23 en remplaçant la fonction f par \mathcal{L} avec λ fixé à sa valeur. Autrement dit, si le point (\bar{x}, \bar{y}) est un extremum libre du Lagrangien, c'est un extremum de même nature (max, min, local ou global) de f sous la contrainte.

 *La condition du théorème 25.11 est seulement suffisante, mais non nécessaire pour avoir un extremum pour f sous la contrainte (voir l'exemple 25.16 page 111). En d'autres termes si le point critique de première espèce n'est pas un extremum pour le Lagrangien on ne peut rien conclure du théorème 25.11. Il faudra utiliser les méthodes des paragraphes suivants*

25.4.1.2. *Conditions suffisantes d'extremum global.* — Si le Lagrangien est convexe ou concave sur \mathcal{U} , on déduit du théorème 25.11 le corollaire suivant :

Corollaire 25.13. — Soient (\bar{x}, \bar{y}) un point critique de première espèce du problème (\mathcal{P}_2) , \mathcal{U} un ensemble convexe, et λ le multiplicateur associé. Posons le lagrangien $\mathcal{L}(x, y) = f(x, y) - \lambda g(x, y)$.

1. Si \mathcal{L} est convexe sur \mathcal{U} alors (\bar{x}, \bar{y}) donne un minimum global du problème (\mathcal{P}_2) .
2. Si \mathcal{L} est concave sur \mathcal{U} alors (\bar{x}, \bar{y}) donne un maximum global du problème (\mathcal{P}_2) .

25.4.1.3. *Conditions suffisantes d'extremum local.* — Si le Lagrangien est seulement localement convexe ou concave au voisinage du point (\bar{x}, \bar{y}) on déduit du théorème 25.11 le corollaire suivant :

Corollaire 25.14. — Soit (\bar{x}, \bar{y}) un point critique de première espèce du problème (\mathcal{P}_2) et soit λ le multiplicateur associé. Posons le lagrangien $\mathcal{L}(x, y) = f(x, y) - \lambda g(x, y)$.

On note

$$D^2\mathcal{L}(\bar{x}, \bar{y}) = \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{y}) & \frac{\partial^2 \mathcal{L}}{\partial x \partial y}(\bar{x}, \bar{y}) \\ \frac{\partial^2 \mathcal{L}}{\partial x \partial y}(\bar{x}, \bar{y}) & \frac{\partial^2 \mathcal{L}}{\partial y^2}(\bar{x}, \bar{y}) \end{pmatrix} = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

1. Si $(rt - s^2 > 0$ et $(r > 0$ ou $t > 0))$ alors (\bar{x}, \bar{y}) donne un minimum local pour f sous la contrainte du problème (\mathcal{P}_2) .
2. Si $(rt - s^2 > 0$ et $(r < 0$ ou $t < 0))$ alors (\bar{x}, \bar{y}) donne un maximum local pour f sous la contrainte du problème (\mathcal{P}_2) .

Remarque 25.15. — Les corollaires 25.13 et 25.14 ne permettent pas de conclure sur la nature du point (\bar{x}, \bar{y}) dans chacun des deux cas suivants :

1. Le lagrangien \mathcal{L} n'est ni convexe ni concave sur \mathcal{U} .
2. (\bar{x}, \bar{y}) est un point col du lagrangien.

Exemple 25.16. — Soit à résoudre le problème : Optimiser $f(x, y) = -x^2 + 4y^2$ sous la contrainte $g(x, y) = x + 2y^2 - 2 = 0$.

Les fonctions f et g sont des fonctions polynômes donc de classe \mathcal{C}^2 sur \mathbb{R}^2 .

Résolvons le problème de deux façons différentes pour comparer les résultats.

1. En explicitant la contrainte.

Celle-ci peut s'écrire de la façon suivante $x = 2 - 2y^2$. On est alors ramené à optimiser la fonction d'une seule variable : $\varphi(y) = 4y^2 - (2 - 2y^2)^2$. Cette fonction admet un minimum local en $y = 0$ et un maximum global atteint pour les deux points $y = \sqrt{3/2}$ et $y = -\sqrt{3/2}$. Grâce à un tableau de variations on peut voir que le maximum est global.

On en déduit que le point $(2, 0)$ donne un minimum local pour f sous la contrainte et les points $(-1, \sqrt{3/2})$ et $(-1, -\sqrt{3/2})$ donnent un maximum global pour f sous la contrainte.

2. En utilisant le Lagrangien

Comme le gradient de la contrainte ne s'annule jamais, les seuls points candidats sont les points critiques du lagrangien $\mathcal{L}(x, y) = -x^2 + 4y^2 - \lambda(x + 2y^2 - 2)$.

On retrouve : $(2, 0)$ avec $\lambda = -4$, et $(-1, \sqrt{3/2})$ et $(-1, -\sqrt{3/2})$ avec $\lambda = 2$.

– Pour $\lambda = 2$: on a $\mathcal{L}(x, y) = -x^2 - 2x + 4 = h(x)$. Cette fonction h est une fonction concave sur \mathbb{R} , donc \mathcal{L} est une fonction concave sur \mathbb{R}^2 en application de la proposition 21.11 page 90. On déduit du corollaire 25.13 que les deux points $(-1, \sqrt{3/2})$ et $(-1, -\sqrt{3/2})$ donnent un maximum global, de valeur 5, pour f sous la contrainte.

– Pour $\lambda = -4$: on a $\mathcal{L}(x, y) = -x^2 + 12y^2 + 4x - 8$. Le lagrangien \mathcal{L} n'est ni convexe, ni concave sur \mathbb{R}^2 . On peut même s'apercevoir que le lagrangien n'est pas localement convexe ou concave au voisinage du point $(2, 0)$.

Pour l'instant nous ne pouvons pas conclure pour ce point (voir la suite dans l'exercice 25.18 page 112 et dans l'exercice 25.19 page 112). Nous avons vu avec la première méthode que ce point donnait un minimum local pour f sous la contrainte.

25.4.1.4. *Linéarisation de la contrainte.* — Le résultat du corollaire 25.14 revient à étudier le signe de la forme quadratique $d^2\mathcal{L}(h, k)$ associée à la matrice hessienne du lagrangien au point (\bar{x}, \bar{y}) (voir théorème 20.5 page 84) et ceci pour tous les couples (h, k) de \mathbb{R}^2 .

On peut assouplir cette condition en se limitant aux couples (h, k) tels que $g(\bar{x} + h, \bar{y} + k) = 0$.

Cette dernière égalité doit permettre de trouver une relation entre h et k , mais elle n'est pas toujours facile à exploiter.

Pour un point critique de première espèce on a $\nabla g(\bar{x}, \bar{y}) \neq (0, 0)$. Nous admettrons que dans ce cas la contrainte peut être remplacée par son approximation affine au voisinage du point (\bar{x}, \bar{y}) :

$$g(\bar{x} + h, \bar{y} + k) \simeq g(\bar{x}, \bar{y}) + h \frac{\partial g}{\partial x}(\bar{x}, \bar{y}) + k \frac{\partial g}{\partial y}(\bar{x}, \bar{y})$$

Comme $g(\bar{x} + h, \bar{y} + k) = g(\bar{x}, \bar{y}) = 0$, on en déduit :

$$h \frac{\partial g}{\partial x}(\bar{x}, \bar{y}) + k \frac{\partial g}{\partial y}(\bar{x}, \bar{y}) = 0$$

ce qui peut aussi s'écrire avec un produit scalaire : $\langle (h, k), \nabla g(\bar{x}, \bar{y}) \rangle = 0$ et permet d'obtenir facilement une relation entre h et k .

On peut alors énoncer le théorème suivant que nous admettrons :

Théorème 25.17. — Soit (\bar{x}, \bar{y}) un point critique de première espèce du problème (\mathcal{P}_2) et soit λ le multiplicateur de Lagrange associé. Soit le lagrangien $\mathcal{L}(x, y) = f(x, y) - \lambda g(x, y)$.

On note

$$D^2\mathcal{L}(\bar{x}, \bar{y}) = \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{y}) & \frac{\partial^2 \mathcal{L}}{\partial x \partial y}(\bar{x}, \bar{y}) \\ \frac{\partial^2 \mathcal{L}}{\partial x \partial y}(\bar{x}, \bar{y}) & \frac{\partial^2 \mathcal{L}}{\partial y^2}(\bar{x}, \bar{y}) \end{pmatrix} = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

et

$$d^2\mathcal{L}_{(\bar{x}, \bar{y})}(h, k) = rh^2 + 2shk + tk^2.$$

1. Si $\forall (h, k) \neq (0, 0)$ tel que $\langle (h, k), \nabla g(\bar{x}, \bar{y}) \rangle = 0$, on a $d^2\mathcal{L}_{(\bar{x}, \bar{y})}(h, k) > 0$ alors le point (\bar{x}, \bar{y}) donne un minimum local pour f sous la contrainte.
2. Si $\forall (h, k) \neq (0, 0)$ tel que $\langle (h, k), \nabla g(\bar{x}, \bar{y}) \rangle = 0$, on a $d^2\mathcal{L}_{(\bar{x}, \bar{y})}(h, k) < 0$ alors le point (\bar{x}, \bar{y}) donne un maximum local pour f sous la contrainte.

Exemple 25.18. — [Suite de l'exemple 25.16]

Revenons à l'étude du point $(2, 0)$. Pour ce point critique on a $\mathcal{L}(x, y) = -x^2 + 12y^2 + 4x - 8$, donc

$$D^2\mathcal{L}(2, 0) = \begin{pmatrix} -2 & 0 \\ 0 & 24 \end{pmatrix}$$

et $d^2\mathcal{L}(h, k) = -2h^2 + 24k^2$. Or $\nabla g(x, y) = (1, 2y)$ donc $\langle (h, k), \nabla g(2, 0) \rangle = \langle (h, k), (1, 0) \rangle = h$.

Si ce produit scalaire est nul on a $h = 0$. Pour $h = 0$, on a $d^2\mathcal{L}(h, k) = 24k^2 > 0$ pour $k \neq 0$, c'est-à-dire pour $(h, k) \neq (0, 0)$, donc le point $(2, 0)$ donne un minimum local pour f sous la contrainte.

25.4.2. Revenir à l'étude directe. — Si aucune des méthodes précédentes n'a permis d'aboutir, si vraiment rien ne marche, en particulier :

- Si $D^2\mathcal{L}(\bar{x}, \bar{y})$ est la matrice nulle.
- Si le Lagrangien n'est ni localement convexe ni localement concave au voisinage de (\bar{x}, \bar{y}) , c'est-à-dire si (\bar{x}, \bar{y}) est un point col du lagrangien.
- Si (\bar{x}, \bar{y}) un point **critique de deuxième espèce**.
- Si les dérivées secondes du lagrangien sont difficiles à calculer

Il faut revenir au principe général (voir ci-dessous).

Principe : Soient $\bar{X} = (\bar{x}, \bar{y}) \in \mathcal{U}$ un point critique de première ou de deuxième espèce et $H = (h, k)$ tel que $\bar{X} + H = (\bar{x} + h, \bar{y} + k) \in \mathcal{U}$.

Pour pouvoir conclure que donne \bar{X} bien un extremum pour f sur \mathcal{U} sous la contrainte, et préciser s'il s'agit d'un maximum ou d'un minimum il faut étudier le signe de

$$\Delta f_{\bar{X}}(H) = f(\bar{x} + h, \bar{y} + k) - f(\bar{x}, \bar{y}).$$

- Si pour tout H tel que $g(\bar{X} + H) = 0$, $\Delta f_{\bar{X}}(H)$ est de signe constant alors \bar{X} donne un extremum global de f sur \mathcal{U} sous la contrainte $g(x, y) = 0$.
- Si lorsque la norme de H est suffisamment petite et $g(\bar{X} + H) = 0$, $\Delta f_{\bar{X}}(H)$ est de signe constant alors \bar{X} donne un extremum local de f sur \mathcal{U} sous la contrainte $g(x, y) = 0$.
- Sinon (\bar{x}, \bar{y}) n'est pas solution du problème (\mathcal{P}_2) .

Exemple 25.19. — [Suite-bis de l'exemple 25.16] Reprenons une dernière fois le cas du point $(2, 0)$ dans l'exemple 25.16. Considérons le point $(2+h, 0+k)$. On doit avoir $g(2+h, 0+k) = 0$, c'est-à-dire : $(2+h) + 2k^2 - 2 = 0$ d'où on tire : $h = -2k^2$.

Étudions alors le signe de :

$$f(2+h, 0+k) - f(2, 0) = -(2+h)^2 + 4k^2 - (-4) = -4h - h^2 + 4k^2.$$

En utilisant la relation entre h et k on obtient :

$$f(2+h, 0+k) - f(2, 0) = 12k^2 - 4k^4 = 4k^2(3 - k^2).$$

Or $4k^2 \geq 0$ et $(3 - k^2) > 0$ si $|k| < 1$. Par suite si $|k|$ est petit on a : $f(2+h, 0+k) - f(2, 0) \geq 0$, ce qui signifie que le point $(2, 0)$ donne un minimum local pour f sous la contrainte.

La méthode de l'exercice précédent n'est utilisable que pour des fonctions "simples" !!