Cours FPV - Semaine 3 : Recherche d'Extrema et Formes Différentielles

Frédéric Messine

1 Introduction

Dans ce chapitre, nous allons étudier une application de la dérivation des fonctions de plusieurs variables à variables numériques : la recherche d'extrema. Nous étudierons ensuite plus précisément les formes différentielles qui seront utilisées dans un prochain cours sur l'intégration.

Les notions vu dans le chapitre précédent vont être nécessaires pour permettre de caractériser les extrema (cad maximum et/ou minimum) d'une fonction à valeur réelle, ainsi que pour définir et étudier les formes différentielles.

Ce chapitre se décompose naturellement en deux parties : la recherche d'extrema et les formes différentielles qui sont deux applications de la dérivation de fonctions de plusieurs variables. Ces deux parties de ce chapitre 3 sont complètements autonomes et disjointes et peuvent être étudiées séparéments.

2 Recherche d'extrema

Dans cette partie, nous allons étudier des problèmes d'optimisation de fonctions numériques; et plus particulièrement de fonctions continues en général C^2 de $\mathbb{R}^2 \rightarrow \mathbb{R}$.

Il est primordial de ne considérer que des fonctions à valeurs réelles, car en optimisation on va rechercher le minimum ou le maximum d'une fonction sur un ensemble. Or l'on va avoir besoin de comparer les valeurs de cette fonction en certains points afin de pouvoir dire qu'un certain point M donne à f sa plus petite valeur.

En effet, seulement l'ensemble $\mathbb{R}$ est ordonné, et $\mathbb{R}^n$ quand $n > 1$ ne l'est plus. Pour s'en rendre compte il suffit de prendre quelques exemples; comparaison dans $\mathbb{R}$: $2 < 3.5$,$2 = 2$,$4 \geq 1.5$. Cela ne pose pas de problème particulier, par contre sur $\mathbb{R}^2$ que peut on dire de $(2,1)$ comparé à $(10,-4)$ lequel des deux est le plus grand. On peut bien sûr prendre la norme et dire que $||(2,1)||_2 = \sqrt{5} < ||(10,-4)||_2 = \sqrt{116}$, mais le fait de prendre la norme nous ramène dans $\mathbb{R}$ et l'on étudie ainsi $||f||$ qui est à valeur dans $\mathbb{R}$ et pas f qui dans ce cas serait à valeur dans $\mathbb{R}^2$.

Le cas de l'optimisation multicritère, cad de fonctions à valeurs vectorielles, n'est pas considéré dans ce cours.

Définition 1 *Soit A un ouvert de $\mathbb{R}^n$ et soit $P = (p_1,\cdots,p_n)$ un point de A.*

Une fonction f définie sur A et à valeurs réelles, admet au point P un minimum local, resp. un maximum local, si et seulement si

$$\exists \eta > 0, \forall Q \in P, ||P - Q|| < \eta \Longrightarrow f(P) \leq f(Q), \textit{ resp. } f(Q) \leq f(P)$$

cad, qu'il existe une boule (un voisinage) autour du point P tel que tous les éléments de cette boule ont des valeurs par f supérieures, resp. inférieures, à $f(P)$.

On emploie le terme local car on a restreint la définition à un certain voisinage dans une boule autour de P.

Définition 2 *Soit A un ouvert de $\mathbb{R}^n$ et soit $P = (p_1, \cdots, p_n)$ un point de A.*

Une fonction f définie sur A et à valeurs réelles, admet au point P son minimum (global), resp. un maximum (global), si et seulement si

$$\forall Q \in A, \textit{ on a } f(P) \leq f(Q), \textit{ resp. } f(Q) \leq f(P)$$

Dans ce cas, on a le minimum absolu (on dit souvent global par opposition à local) de f sur A. On n'utilise pas dans cette définition de boule (de voisinage) autour de P,et donc P est la solution des problèmes d'optimisation suivants :

$$f(P) = \min_{P \in A} f(x), \text{ resp. } f(P) = \max_{P \in A} f(x).$$

Remarque 1 *On va voir dans ce qui suit, que ce n'est facile de caractériser par des propriétés mathématiques ce fameux optimum global -on emploie le terme optimum pour parler indifféremment de minimum ou de maximum- par contre, on en aura pour caractériser les optima locaux.*

Cette partie est succinte et elle a pour but de vous faire découvrir l'optimisation de manière très intuitive.

2.1 Propriétés avec la dérivée première

Les optima, appelés aussi extrema, vérifie la propriété suivante :

Théorème 1 *Soit f une fonction numérique de plusieurs variables si P est un extrema (ou optima, soit un minimum ou un maximum), alors*

$$\nabla f(P) = 0$$

cad le vecteur gradient constitué de toutes les dérivées patielles s'annule en P.

Remarque 2 *En fait, la condition (d'Euler) du gradient qui s'annule en un point, veut dire que ce point est stationnaire, donc c'est soit*

1. *un minimum local,*
2. *un maximum local,*
3. *mais aussi d'autres points stationnaires.*
   Par exemple prenons $f(x) = x^3$ la dérivée s'annule en 0 mais 0 n'est ni un minimum local ni un maximum local.

On va maintenant donner quelques propriétés supplémentaires pour établir si ces points stationnaires sont des extrema et donc des minima ou des maxima.

2.2 Propriétés avec la dérivée première

Regardons ce qu'il se passe au voisinage ou les dérivées premières s'annulent. Ecrivons la formule de Taylor à l'ordre 2 en P :

$$f(P+h) = f(P) + h.\nabla f(P) + \frac{1}{2}h.(Hf(P).h) + ||h||^2\epsilon(h)$$

C'est une notation matricielle, où $x \in I\!R^n$, $h \in I\!R^n$, $\nabla f(x) \in I\!R^n$, et $Hf(x)$ représente la matrice Hessienne de f, c'est à dire la matrice Jacobienne de la fonction à valeur vectorielle $\nabla f(x)$. Cela représente la dérivée seconde de f en x. L'opération "." est le produit scalaire.

La matrice Hessienne est symétrique d'après Schwartz car ces termes sont de la forme :

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}, \text{ d'après le théorème de Schwartz, cf chapitre 2.}$$

Pour se représenter les idées regardons ce qu'il se passe quand $n = 2$:

$$\begin{aligned} f(p_1+h_1,p_2+h_2) &= f(p_1,p_2) + h_1\frac{\partial f(p_1,p_2)}{\partial x} + h_2\frac{\partial f(p_1,p_2)}{\partial y} + \frac{1}{2}\left[h_1^2\frac{\partial^2 f(p_1,p_2)}{\partial x^2}\right. \\ &\quad \left. +2h_1h_2\frac{\partial^2 f(p_1,p_2)}{\partial x \partial y} + h_2^2\frac{\partial^2 f(p_1,p_2)}{\partial y^2}\right] + ||(h_1,h_2)||^2\epsilon(h_1,h_2). \end{aligned}$$

Nous ne verrons ici les propriétés que dans le cas de deux variables.

En utilisant les formules de Monge :

$$\begin{aligned} r &= \frac{\partial^2 f}{\partial x^2}(x,y) \\ s &= \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ t &= \frac{\partial^2 f}{\partial y^2}(x,y) \\ p &= \frac{\partial f}{\partial x}(x,y) \\ q &= \frac{\partial f}{\partial y}(x,y) \end{aligned}$$

Seuls r,s,t sont utilisés ici, car au point P la dérivée va s'annuler et donc : $p = q = 0$.

$$r\left(\frac{h_1}{h_2}\right)^2 + 2\frac{h_1}{h_2}s + t,$$

est un trinôme du second degré en $\frac{h_1}{h_2}$.

Si ces dérivées secondes ne sont pas toutes nulles en P, $f(p_1+h_1,p_2+h_2) - f(p_1,p_2)$ sera du même signe que ce trinôme.

Donc autour du point P (on dit au voisinage de P en topologie générale), on va savoir si la fonction est toujours supérieure (minimum local) ou toujours inférieure (maximum local) ou ni l'un ni l'autre (ce n'est pas un extremum).

Etudions le signe de ce trinôme et donc il vient les propriétés supplémentaires suivantes :

Proposition 1 *On pose $\Delta' = (s^2 - rt)$ c'est à dire $\Delta = 4\Delta'$. Δ et Δ' sont donc du même signe.*

- *Si $\Delta' < 0$ alors le trinôme est du même signe que r, et donc on a que :*
  - *Si $r > 0$ alors $f(p_1+h_1,p_2+h_2) > f(p_1,p_2)$ et donc P est un* **minimum local**.
  - *Si $r < 0$ alors $f(p_1+h_1,p_2+h_2) < f(p_1,p_2)$ et donc P est un* **maximum local**.
- *Si $\Delta' > 0$ alors le trinôme ne garde pas de signe constant et donc P n'est pas un extremum.*
- *Si $\Delta' = 0$ alors on ne peut rien conclure et il faut essayer Taylor à l'ordre 3, ce qui s'avère fastidieux et qui peut aussi être limité (comme exemple considérez $f(x) = x^4$).*

Exemple 1 *Cherchons les extrema de f définie par :*

$$f(x,y) = x^3 + y^3 - 3xy.$$

Calculons les termes de Monge :

$$\begin{aligned}
p &= \frac{\partial f}{\partial x}(x,y) = 3(x^2 - y) \\
q &= \frac{\partial f}{\partial y}(x,y) = 3(y^2 - x) \\
r &= \frac{\partial^2 f}{\partial x^2}(x,y) = 6x \\
s &= \frac{\partial^2 f}{\partial x \partial y}(x,y) = -3 \\
t &= \frac{\partial^2 f}{\partial y^2}(x,y) = 6y
\end{aligned}$$

Pour annuler le gradient, il faut que $p = q = 0$, soit $3(x^2 - y) = 0$ et $3(y^2 - x) = 0$. On a donc deux solutions : soient $P = (0,0)$ et $Q = (1,1)$.

Regardons ce qu'il se passe au voisinage de $P = (0,0)$:

$$\begin{aligned}
f(h_1,h_2) - f(0,0) &= \frac{1}{2}\left[h_1^2\frac{\partial^2 f}{\partial x^2}(0,0) + 2h_1h_2\frac{\partial^2 f}{\partial x \partial y}(0,0) + h_2^2\frac{\partial^2 f}{\partial y^2}(0,0)\right] \\
&\quad +\epsilon(h_1,h_2)||(h_1,h_2)||^2 \\
&= -3h_1h_2 + \epsilon(h_1,h_2)||(h_1,h_2)||^2.
\end{aligned}$$

On a le discriminant $\Delta' = (s^2 - rt) = 9 > 0$, le trinôme ne garde donc pas de signe constant en ce point P. P n'est pas un extremum.

Regardons ce qu'il se passe au voisinage de $Q = (1,1)$:

$$\begin{aligned}
f(1+h_1,1+h_2) - f(1,1) &= \frac{1}{2}\left[h_1^2\frac{\partial^2 f}{\partial x^2}(1,1) + 2h_1h_2\frac{\partial^2 f}{\partial x \partial y}(1,1) + h_2^2\frac{\partial^2 f}{\partial y^2}(1,1)\right] \\
&\quad +\epsilon(h_1,h_2)||(h_1,h_2)||^2 \\
&= 3(h_1^2 - h_1h_2 + h_2^2) + \epsilon(h_1,h_2)||(h_1,h_2)||^2.
\end{aligned}$$

On a le discriminant $\Delta' = (s^2 - rt) = 9 - 6*6 < 0$, *et* $r = 6 > 0$, Q *est donc un minimum local.*

2.3 Exemples de fonctions à optimiser

Cette partie, vous donne quelques exemples classiques d'allure des fonctions que l'on est amené à considérer en optimisation.

- Fonctions convexes comme par exemple $f(x,y) = x^2 + y^2$. Elles ont l'allure de paraboloïdes :

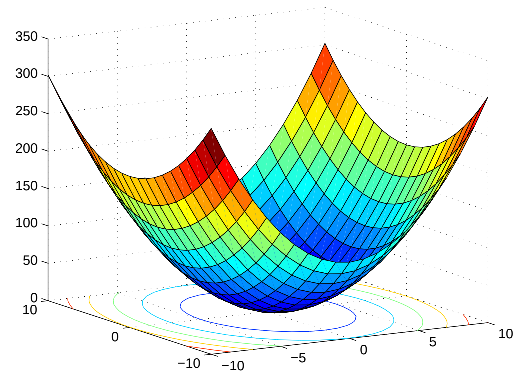

FIG. 1 – *Représentation graphique d'une fonction convexe.*

Il existe un seul minimum local et un seul point de gradient nul et c'est un minimum.

- Fonctions concaves comme par exemple $f(x,y) = -x^2 - y^2$. Elles ont l'allure de paraboloïdes renversées, cf Fig. 2.

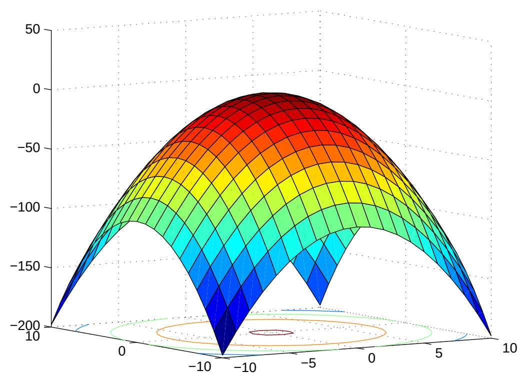

FIG. 2 – *Représentation graphique d'une fonction concave.*

Il existe un seul maximum local et un seul point de gradient nul et c'est un maximum.

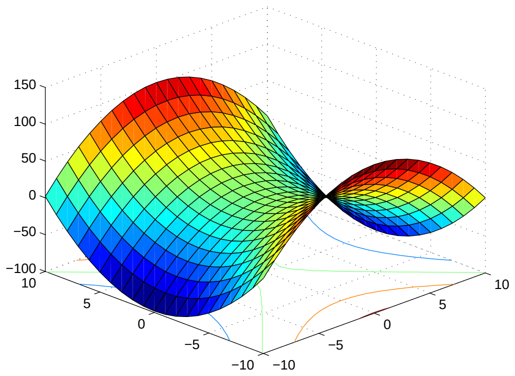

FIG. 3 – *Représentation graphique d'une fonction selle.*

- Fonctions selles comme par exemple $f(x,y) = x^2 - y^2$. Elles ont l'allure d'une "selle d'équitation", cf Fig. 3.
  Il existe un seul point de gradient nul mais ce n'est ni un minimum ni un maximum. Ces fonctions ont un intérêt en optimisation avec des contraintes.
- Exemple de cas où $\Delta' = 0$. Fonctions de la forme $(x+y)^3$ mais aussi $(x+y)^4$, cf Fig. 4. et cf Fig. 5.

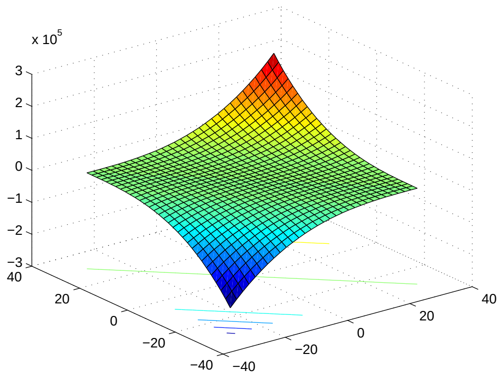

FIG. 4 – *Représentation graphique d'une fonction avec solution sur $y = x$.*

  Pour ces fonctions, on a une infinité de points stationnaires, sur la droite $y = -x$ qui sont des minima pour la deuxième fonction. Ceci est, en général, difficile à caractériser.
- Exemple de cas de fonctions avec plusieurs extrema (problèmes difficiles d'optimisation globale). Par exemple, des fonctions de la forme $sin(xy^2) - cos(y/2)$, cf Fig. 6.

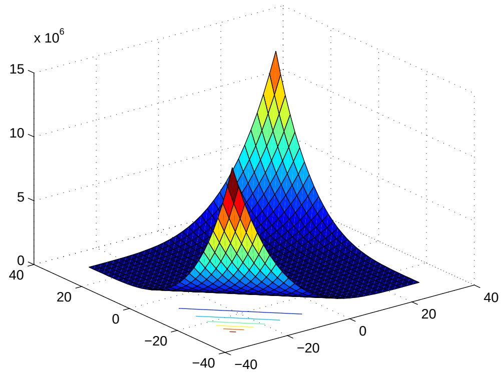

Fig. 5 – *Représentation graphique d'une fonction avec solution sur* $y = x$*.*

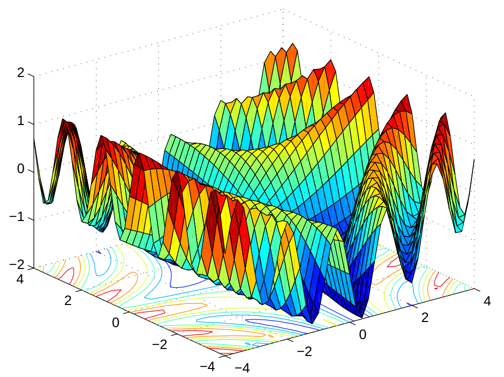

Fig. 6 – *Représentation graphique d'une fonction quelconque.*

3 Formes Différentielles

On a vu que si f est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, alors ce que l'on a noté df; la différentielle de f, est une application de $\mathbb{R}^n$ vers l'ensemble des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}$, noté $L(\mathbb{R}^n,\mathbb{R})$.

$$f : \mathbb{R}^n \to \mathbb{R}, \text{ alors } df : \mathbb{R}^n \to L(\mathbb{R}^n,\mathbb{R})$$

Soit P un point de $\mathbb{R}^n$ alors $df(P)$ est une application linéaire, mais df n'en est pas.

En physique, il est souvent aisé de décrire des phénomènes à partir de petits changements de paramètres. En effet, on suppose souvent au niveau macroscopique, après avoir déterminés les paramètres -nos variables- qui régissent le sytème considéré, que :
- ces paramètres varient de manière continus,
- une petite variation du système peut-être décrite à partir d'une petite variation des paramètres.

Ceci n'est plus vrai en physique atomique.

Ainsi, on étudie souvent des problèmes de la forme :

$$P(x,y)dx + Q(x,y)dy, \text{ ou } P(x,y,z)dx + Q(x,y,z)dy + R(x,y,z)dz$$

Ceci sont deux formes différentielles dans le plan et dans l'espace des réels.

3.1 Définition

Définition 3 *On appelle forme différentielle toute application de $\mathbb{R}^n$ ou d'une partie de $\mathbb{R}^n$ vers $L(\mathbb{R}^n,\mathbb{R})$. En utilisant les notations $dx_1,dx_2,\cdots,dx_n$, une forme différentielle s'écrit :*

$$A_1(x_1,\cdots,x_n)dx_1 + A_2(x_1,\cdots,x_n)dx_2 + \cdots + A_n(x_1,\cdots,x_n)dx_n$$

C'est sous cette forme que l'on rencontre le plus souvent les formes différentielles et c'est sous cette forme que nous les étudieront dans ce chapitre.

Remarque 3 *On dit forme différentielle et non différentielle, car ceci a seulement l'apparence d'une différentielle (qui s'écrit comme ceci aussi). Cependant ce n'est pas forcément une différentielle de fonction et c'est là le point important qui génére cette différence d'appellation.*

3.2 Différentielles et formes différentielles

Il est essentiel de faire la différence entre différentielle et forme différentielle car il y a des formes différentielles qui ne sont pas issues de fonctions dont elles seraient la différentielle.

Nous disposons de trois termes pour définir une forme différentielle :
1. une différentielle de fonction,
2. une forme différentielle exacte (encore appelé différentielle exacte),
3. une forme différentielle (quelconque).

Le deuxième terme sert à dire que pour cette forme différentielle, il existe une fonction dont c'est la differentielle même si l'on ne connait pas cette fonction.

Remarque 4 *On fait la différence entre les termes fonctions différentielles et formes différentielles exactes car, dans le premier cas, on s'est donné une fonction et on a calculé sa différentielle, alors que dans le deuxième cas, on est directement partie de la forme différentielle et l'on a découvert qu'il existait une fonction dont c'était la différentielle.*

3.3 Formes différentielles exactes

Définition 4 *Soit ω une forme différentielle, définie sur $A \subseteq I\!R^n$:*

$$\omega = A_1(x_1,\cdots,x_n)dx_1 + A_2(x_1,\cdots,x_n)dx_2 + \cdots + A_n(x_1,\cdots,x_n)dx_n$$

cette forme est dite exacte si et seulement si il existe une fonction f différentiable sur $A \subseteq I\!R^n$ telle que

$$\omega = df$$

On dit que f est une primitive de ω. En physique, on dit aussi que ω dérive du potentiel f.

Comme il est dans le cas général souvent difficile de chercher des primitives, nous allons chercher des propriétés qui vont permettre de statuer sur le fait qu'une forme différentielle soit exacte ou non sans avoir à calculer de primitive; ceci dans le cas de deux puis de trois variables qui généralement suffisent en physique.

Théorème 2 *Soit une forme différentielle de deux variables :*

$$\omega = P(x,y)dx + Q(x,y)dy$$

ω est une forme différentielle exacte sur $A \subseteq I\!R^2$ si et seulement si

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

Théorème 3 *Soit une forme différentielle de trois variables :*

$$\omega = P(x,y,z)dx + Q(x,y,z)dy + R(x,y,z)dz$$

ω est une forme différentielle exacte sur $A \subseteq I\!R^3$ si et seulement si

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y} \ \ et \ \ \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}.$$

Remarque 5 *Sans donner les démonstrations ces deux théorèmes, on peut s'apercevoir que l'on va utiliser le théorème de Schwartz dans la demonstration. En effet, si il exite f tel que $\omega = df$, soit,*

$$df = P(x,y)dx + Q(x,y)dy$$

en fait $\dfrac{\partial f}{\partial x}(x,y) = P(x,y)$ *et* $\dfrac{\partial f}{\partial y}(x,y) = Q(x,y)$*, d'après le théorème de Schwartz on a que*

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(x,y)$$

et donc on a bien que :

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

Ces théorèmes sont dus à Poincarré et sont aussi connus sous le nom de théorèmes de Poincarré.

3.4 Retrouver une fonction à partir de sa forme différentielle exacte

Soit ω une forme différentielle exacte définie sur $A \subseteq I\!R^3$:

$$\omega = P(x,y,z)dx + Q(x,y,z)dy + R(x,y,z)dz \text{ et } \omega = df$$

En fait, l'on doit résoudre le système suivant :

$$\begin{cases} \frac{\partial f}{\partial x}(x,y,z) = P(x,y,z) \\ \frac{\partial f}{\partial y}(x,y,z) = Q(x,y,z) \\ \frac{\partial f}{\partial z}(x,y,z) = R(x,y,z) \end{cases}$$

Si ω est une forme différentielle exacte alors on est sûr que ce sytème a une solution (on peut même rajouter que cette solution sera unique à une constante près).

On est donc ramené à résoudre des systèmes d'équations aux dérivées partielles ce qui n'est pas toujours évident à effectuer.

Exemple 2 *Prenons un premier exemple de deux variables :*

$$\frac{dx}{x} - \frac{dy}{y}$$

c'est bien une forme différentielle exacte car d'après le premier théorème de Poincarré :

$$\frac{\partial \frac{1}{x}}{\partial y} = 0 \ \textit{et} \ \frac{\partial - \frac{1}{y}}{\partial x} = 0$$

D'où l'on est sûr d'avoir une solution au problème suivant (on n'est cependant pas sûr d'être toujours capable de la trouver),

$$\begin{cases} \frac{\partial f}{\partial x}(x,y) = \frac{1}{x} \\ \frac{\partial f}{\partial y}(x,y) = -\frac{1}{y} \end{cases}$$

Et donc, en considérant la première équation on doit avoir que f s'écrit :

$$f(x,y) = \ln x + C(y)$$

En considérant maintenant la deuxième équation on a que :

$$f(x,y) = \ln x - \ln y + K$$

où $K \in I\!R$ est une constante.

Exemple 3 *Considérons maintenant un cas plus complexe à trois variables :*

$$\omega = \left[\frac{1}{(x+y)z} + 3x\right] dx + \frac{1}{(x+y)z} dy - \frac{\ln(x+y)}{z^2} dz$$

C'est bien une différentielle exacte mais on ne va pas le montrer (montrez le à titre d'exercice).

Supposons donc que cette forme différentielle est exacte. Nous avons donc le système suivant à résoudre :

$$\begin{cases} \frac{\partial f}{\partial x}(x,y,z) = \frac{1}{(x+y)z} + 3x \\ \frac{\partial f}{\partial y}(x,y,z) = \frac{1}{(x+y)z} \\ \frac{\partial f}{\partial z}(x,y,z) = \frac{-\ln(x+y)}{z^2} \end{cases}$$

Si l'on considère la troisième équation, on obtient que :

$$f(x,y,z) = \frac{\ln(x+y)}{z} + C(x,y)$$

En considérant cette relation, on retrouve le système réduit suivant :

$$\begin{cases} \frac{\partial f}{\partial x}(x,y,z) = \frac{1}{(x+y)z} + 3x = \frac{1}{(x+y)z} + \frac{\partial C}{\partial x}(x,y) \\ \frac{\partial f}{\partial y}(x,y,z) = \frac{1}{(x+y)z} = \frac{1}{(x+y)z} + \frac{\partial C}{\partial y}(x,y) \end{cases}$$

d'où l'on tire que :

$$\begin{cases} \frac{\partial C}{\partial x}(x,y) = 3x \\ \frac{\partial C}{\partial y}(x,y) = 0 \end{cases}$$

On a donc que la fonction C ne dépend pas de la variable y, on peut donc écrire seulement $C(x) = C(x,y)$, et on en déduit que $C(x) = \frac{3x^2}{2} + K$.

Finallement, on obtient

$$f = \frac{\ln(x+y)}{z} + \frac{3x^2}{2} + K$$

Comme on a obtenu une fonction telle que $df = \omega$, ω est bien une forme différentiable exacte.

Remarque 6 *Vous n'êtes pas obliger comme dans l'exercice précédent de regarder au préalable si la forme différentielle considérée est exacte ou non. Cependant, si c'est une forme différentielle qui n'est pas exacte, on ne pourra pas résoudre le système correspondant. Donc si vous essayer de le résoudre quand même, vous n'y arriverez pas et j'espère que vous vous poserez la question de savoir si vous pouvez démontrer que la forme différentielle considérée n'est pas exacte.*

A titre d'exercice, essayer de résoudre :

$$xydx + x^2dy$$

4 Conclusion

Nous avons vu dans ce chapitre deux applications de la différentiabilité de fonctions de plusieurs variables : l'optimisation et les formes différentielles. Les formes différentielles sont encore des objets abstraits mais qui interviennent dans toutes les intégrales. Il est donc important de savoir les manipuler pour le cours sur l'intégration.

Le dernier chapitre est très court, il porte sur les fonctions implicites. C'est à dire des fonctions de plusieurs variables qui ne sont pas explicitement connues comme dans tous les exercices que nous venons de faire. On va voir que même si l'on ne connait pas la fonction explicitement, l'on va pouvoir calculer les dérivées de ces fonctions en certains points; ceci bien sûr sous certaines hypothèses.