

Notes de cours
Analyse à plusieurs variables

Cyrille CHENAVIER

ENSIL ENSCI

FIMI2A - semestre 2

1	Qu'est-ce que le calcul vectoriel?	3
1.1	Motivations	3
1.2	Contenu du cours	5
1.3	Rappels	5
1.4	Fiche d'exercices 1	11
2	Fonctions de plusieurs variables	12
2.1	Représentations des champs scalaires et vectoriels	12
2.2	Dérivées partielles et directionnelles	14
2.3	Règle de dérivation des fonctions composées	16
2.4	Fiche d'exercices 2	20
3	Optimisation	22
3.1	Gradient et points critiques	22
3.2	Hessienne et extrema locaux	26
3.3	Multiplicateurs de Lagrange	29
3.4	Fiche d'exercices 3	31
4	Opérateurs différentiels et intégration	33
4.1	Opérateurs différentiels classiques	33
4.2	Intégrales curviligne, de surface et volumique	35
4.3	Théorème de Stokes et généralisations	39
4.4	Fiche d'exercices 4	43

Chapitre 1 :

Qu'est-ce que le calcul vectoriel ?

Dans ce chapitre, on présente les motivations et objectifs de ce cours ainsi que les pré-requis attendus pour pouvoir le suivre. Tout au long de ce cours, on se concentre principalement sur les aspects géométriques de l'analyse à plusieurs variables, mais on insiste également sur les liens étroits entre la géométrie et la physique. Ces liens sont esquissés dans la section 1.1 et sont rappelés par des exemples illustratifs tout au long du cours. De plus, plusieurs exercices proposés à la fin des chapitres portent sur des problèmes motivés par la physique. Enfin, bien que l'analyse à plusieurs variables puisse être développée sur des espaces géométriques courbés, on n'aborde uniquement le cas des espaces plats ici, si bien que la plupart des notions introduites et des calculs effectués le sont dans le langage vectoriel. Pour cette raison, à l'acceptation de la section 1.1, les expressions "analyse à plusieurs variables" et "calcul vectoriel" sont considérées comme étant synonymes.

1.1. Motivations

L'analyse à plusieurs variables porte sur l'étude des variations d'une ou plusieurs quantités scalaires exprimées en fonction d'une ou plusieurs autres quantités scalaires. Les quantités de ce deuxième groupe sont appelées *variables* et on dit que les quantités du premier groupe sont des *fonctions* de ces variables. Le cas d'une fonction à une seule variable est bien connu : leur étude est le domaine de l'analyse réelle. Deux autres cas particuliers apparaissent naturellement : plusieurs fonctions d'une seule variable et une seule fonction de plusieurs variables. Par exemple, le mouvement d'une particule dans l'espace au cours du temps peut être décrit à l'aide de 3 fonctions coordonnées dépendant de la seule variable temporelle. Par ailleurs, la température dans une pièce contenant un radiateur est décrite par une seule fonction de 2 variables : une coordonnée spatiale, car la température dépend de la distance par rapport au radiateur, et la coordonnée temporelle, car le radiateur est généralement en marche uniquement en hiver et avec une intensité variable au cours de la journée. Enfin, un exemple de plusieurs fonctions de plusieurs variables : la mesure de la vitesse d'un vent à la surface de la terre est décrite par deux fonctions indiquant la direction et la vitesse du vent et ces fonctions dépendent à la fois de la position sur le globe et du moment où la vitesse est mesurée, ce qui nécessite 2 variables spatiales et la variable temporelle.

Les exemples précédents illustrent que les fonctions de plusieurs variables offrent un formalisme pour mesurer certains attributs d'un espace géométrique ou physique, les variables étant des coordonnées permettant de se situer dans cet espace. Il n'est donc pas surprenant que les fonctions de plusieurs variables jouent un rôle central en géométrie et en physique. En géométrie, les espaces étudiés peuvent être des courbes (1 coordonnée), des surfaces (2 coordonnées), des volumes (3 coordonnées) et leurs analogues en dimensions supérieures, appelés *variétés* (nombre quelconque de coordonnées). Les fonctions sur les variétés permettent par exemple de mesurer des distances, des angles, des volumes ou des éléments de courbure. Par ailleurs, de nombreux espaces physiques sont décrits mathématiquement

comme étant des espaces géométriques, *i.e.*, des variétés, admettant certaines propriétés induites par la physique. Par exemple, en régime non statique, *i.e.*, quand les mesures physiques dépendent du temps, l'espace physique est muni de coordonnées spatiales et d'une coordonnée temporelle, un tel espace est appelé *espace-temps* et sa géométrie dépend de la physique. En effet, à l'échelle macroscopique il est plat en l'absence de matière et est courbé au voisinage de corps célestes massifs. L'utilisation des fonctions de plusieurs variables en physique découle de cette modélisation géométrique : les mesures physiques sont des mesures géométriques sur la variété représentant le problème étudié.

En géométrie comme en physique, on parle généralement de *champs* plutôt que de fonctions de plusieurs variables, ces champs étant de différentes natures. Par exemple, la mesure d'une température, d'une pression ou d'une concentration dans une pièce ou un fluide s'expriment avec un seul nombre, appelé scalaire, si bien qu'on parle de champs scalaires. Les mesures de vitesses, d'accélération ou de forces s'expriment avec des vecteurs et on parle alors de champs de vecteurs. Plus généralement, on peut s'intéresser à des mesures représentées par des tableaux à plusieurs dimensions, tels que les matrices, appelés tenseurs, et on parle alors de champs de tenseurs. Par exemple, les règles pour mesurer les distances, la courbure ou les volumes en géométrie sont des champs de tenseurs. Des champs de tenseurs classiques en physique sont le tenseur électromagnétique, le tenseur d'impulsion-énergie ou le tenseur de déformation. On remarque qu'il est naturel de parler de champs : de la même façon qu'un champ de tulipes contient en tout point une tulipe, un champ scalaire/vectorel/tensoriel contient en tout point un scalaire/vecteur/tenseur.

Le langage géométrique permet d'exprimer certaines lois de la physique à travers des équations vérifiées par des champs. Les méthodes de résolution et d'analyse qualitative de ces équations permettent alors de déduire de nombreuses propriétés physiques. Par exemple, la deuxième loi de Newton est exprimée sous forme d'une équation vectorielle reliant forces et accélérations, et une résolution de cette équation permet de décrire le chemin d'une particule dans un champ de forces. Voici d'autres exemples d'équations de plusieurs variables en physique :

- les équations de Maxwell en électromagnétisme,
- l'équation de transport décrit le mouvement d'un champ scalaire induit par un fluide,
- l'équation de chaleur régit la diffusion de la chaleur dans un milieu conducteur,
- l'équation d'onde exprime la tendance d'un point à revenir en position d'équilibre,
- l'équation de Laplace est satisfaite par le potentiel électrique ou gravitationnel en l'absence de charges ou dans le vide,
- les équations d'Euler-Lagrange en mécanique Lagrangienne caractérisent le chemin minimisant la fonctionnelle d'action,
- les équations de la relativité générale relient la courbure de l'espace-temps à la présence de matière et d'énergie,
- on peut encore citer l'équation de Navier-Stokes en mécanique des fluides et l'équation de Schrödinger en mécanique quantique non relativiste.

Dans ce cours, on évoque seulement une partie de ces équations, puisqu'on se limite aux champs scalaires et vectoriels sur des espaces géométriques ou physiques plats. Les études des champs de tenseurs et des espaces courbés nécessitent en effet des pré-requis d'algèbre et de géométrie introduits en troisième année de licence et en master de mathématiques, comme la topologie générale, le calcul différentiel et intégral ou l'algèbre multi-linéaire.

1.2. Contenu du cours

Le cours comprend 4 parties.

- Dans la section 1.3, on rappelle les pré-requis nécessaires pour aborder les chapitres suivants.
- Dans le chapitre 2, on introduit les fonctions de plusieurs variables et les notions de *dérivées partielles* et *dérivées directionnelles*. On énonce également une formule centrale de l'analyse à plusieurs variables qu'est la règle de *dérivation des fonctions composées*.
- Le chapitre 3 porte sur l'*optimisation* de fonctions de plusieurs variables, *i.e.*, la recherche de points maximisant et minimisant une telle fonction. On présente deux méthodes, selon que la fonction à optimiser soit avec ou sans *contrainte*.
- Dans le chapitre 4, on introduit et analyse les propriétés géométriques de nouveaux opérateurs classiques (*divergence*, *rotationnel* et *Laplacien*) et leurs liens avec les *intégrales multiples*. En particulier, on énonce la *formule de Stokes* et ses généralisations.

De plus, on joint des sujets d'exercices correspondant à la fin des chapitres 1, 2, 3 et 4. On ajoute également une fiche d'exercices sur les systèmes d'*équations aux dérivées partielles* à la fin du cours.

1.3. Rappels

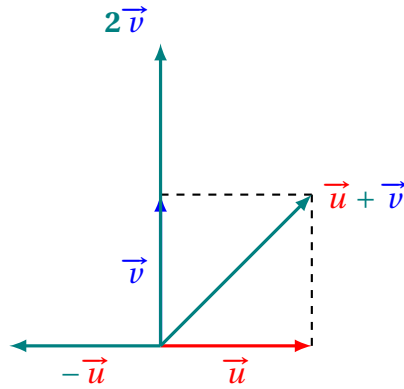
Vecteurs. Un vecteur dans le plan ou dans l'espace est un objet mathématique possédant 2 attributs : une direction et une longueur. Un tel vecteur est représenté par une flèche dans le plan ou dans l'espace, deux flèches ayant même direction et même longueur étant supposées représentées le même vecteur. On note les vecteurs par des lettres surmontées d'une flèche, *e.g.*, \vec{u} ou \vec{v} . Dans un contexte purement mathématique, la lettre sous la flèche est généralement en minuscule, mais dans un contexte physique, elles peuvent aussi bien être en minuscule, comme \vec{a} pour l'accélération d'un corps, ou en majuscules, comme \vec{F} pour la force résultante subie par un corps.

On peut effectuer trois opérations avec les vecteurs, chacune d'elle ayant une illustration physique simple.

- On peut ajouter deux vecteurs grâce à la relation de Chasles. Cette addition correspond au principe de superposition en physique, *e.g.*, lorsqu'un corps est soumis à deux forces \vec{F}_1 et \vec{F}_2 , alors la force résultante sur ce corps est $\vec{F}_1 + \vec{F}_2$.
- On peut prendre l'opposée d'un vecteur en inversant sa direction et en conservant la même longueur. L'opposé permet de formaliser la troisième loi de Newton : si un corps A exerce une force \vec{F} sur un corps B , alors le corps B exerce la force $-\vec{F}$ sur A .
- On peut enfin multiplier un vecteur par un scalaire, en multipliant la longueur du vecteur par ce scalaire et en gardant ou inversant sa direction, selon que le scalaire soit positif ou négatif. L'opération de multiplication d'un vecteur par un scalaire se retrouve dans le membre droit la deuxième loi de Newton $\vec{F} = m\vec{a}$.

Enfin, il existe un vecteur, appelé *vecteur nul* et noté $\vec{0}$, qui a la propriété que pour tout autre vecteur \vec{v} , on a $\vec{v} + \vec{0} = \vec{v}$. Ce vecteur n'a pas de direction spécifique et a une longueur nulle.

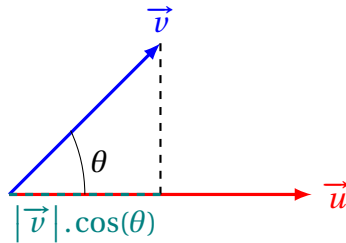
Sur le dessin suivant, on considère deux vecteurs \vec{u} et \vec{v} et on forme leur somme, on prend l'opposée de \vec{u} et le produit $2\vec{v}$:



Produits scalaire et vectoriel. Il existe deux façons de multiplier des vecteurs : le produit scalaire construit un nombre ou scalaire et produit vectoriel un vecteur. La valeur absolue du produit scalaire $\vec{u} \cdot \vec{v}$ des deux vecteurs \vec{u} et \vec{v} est le produit $|\vec{u}| \cdot |\vec{v}_h|$, où \vec{v}_h est le vecteur obtenu par projection orthogonale de \vec{v} sur \vec{u} , et $|\cdot|$ désigne l'opérateur prenant la longueur d'un vecteur. Le signe de $\vec{u} \cdot \vec{v}$ dépend de l'orientation de \vec{v}_h par rapport à \vec{u} : positif s'ils pointent dans le même sens et négatif sinon. Le produit scalaire peut également être défini par une formule close :

$$\vec{u} \cdot \vec{v} = |\vec{u}| \cdot |\vec{v}| \cdot \cos(\theta),$$

où θ est l'angle formé par les vecteurs \vec{u} et \vec{v} . On remarque en particulier que le produit scalaire est commutatif, *i.e.*, on a toujours l'égalité $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$. Sur le dessin suivant, on illustre ces différentes considérations :



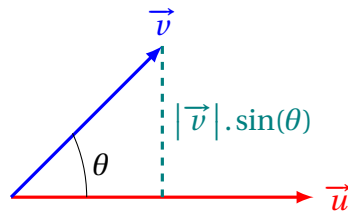
Le produit vectoriel $\vec{u} \times \vec{v}$ est le vecteur orthogonal au plan dirigé par \vec{u} et \vec{v} , son orientation est déterminée par la règle de la main droite et sa longueur est l'aire du parallélogramme défini par \vec{u} et \vec{v} . En particulier, le produit vectoriel est anti-commutatif, *i.e.*, on a toujours l'égalité $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$. En notant \vec{n} le vecteur normal unitaire obtenu par la règle de la main droite, les dessins suivants, tirés de la page anglaise de Wikipédia sur le produit vectoriel ("cross product"), illustrent ces considérations :



Le produit vectoriel peut également être défini par la formule close suivante :

$$\vec{u} \times \vec{v} = |\vec{u}| \cdot |\vec{v}| \cdot \sin(\theta) \cdot \vec{n},$$

où θ est l'angle direct, *i.e.*, compris entre 0 et π , formé par \vec{u} et \vec{v} . En effet, l'aire d'un parallélogramme étant le produit de la base par la hauteur du parallélogramme est bien égale à $|\vec{u}| \cdot |\vec{v}| \cdot \sin(\theta)$ comme on peut le remarquer sur la figure suivante :

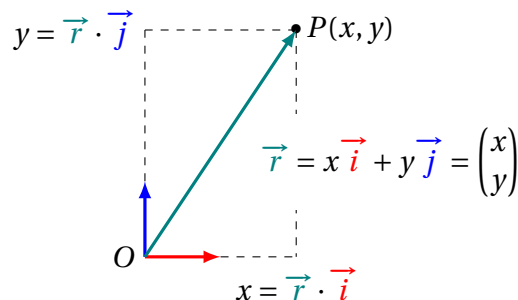


Vecteur position et coordonnées cartésiennes. Dans le plan et dans l'espace, tout point peut être repéré par son vecteur position. Pour cela, on choisit un point d'origine noté O dans le plan ou dans l'espace, et le vecteur position d'un point P du plan ou de l'espace est le vecteur \vec{OP} allant de O vers P . Le vecteur position est noté $\vec{r}(P)$, ou plus simplement \vec{r} quand il n'y a pas d'ambiguïté.

En plus d'une origine, on fixe un repère cartésien permettant de repérer un point par ses coordonnées notées (x, y) dans le plan et (x, y, z) dans l'espace. Les vecteurs dirigeant l'axe des x , des y et des z sont généralement notés \vec{i} , \vec{j} et \vec{k} . Chaque point P est repéré de façon unique par ses coordonnées définies par $x = \vec{r} \cdot \vec{i}$, $y = \vec{r} \cdot \vec{j}$ et $z = \vec{r} \cdot \vec{k}$. On identifie un point avec ses coordonnées, et on écrit simplement $P(x, y, z)$ à la place de $P = (x, y, z)$. Afin de différencier les coordonnées d'un point $P(x, y, z)$ et les composantes du vecteur position $\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$ correspondant, celui-ci pourra être noté sous forme d'une matrice colonne :

$$P(x, y, z) \Leftrightarrow \vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Afin d'avoir un dessin lisible, on illustre ces considérations dans le plan plutôt que dans l'espace :



Les coordonnées cartésiennes permettent de calculer les produits scalaires et vectoriels de deux vecteurs $\vec{u} = x\vec{i} + y\vec{j} + z\vec{k}$ et $\vec{v} = x'\vec{i} + y'\vec{j} + z'\vec{k}$ par

$$\vec{u} \cdot \vec{v} = xx' + yy' + zz' \quad \text{et} \quad \vec{u} \times \vec{v} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x & y & z \\ x' & y' & z' \end{vmatrix} = (yz' - zy')\vec{i} + (zx' - xz')\vec{j} + (xy' - yx')\vec{k},$$

où l'expression du milieu est un déterminant. On rappelle qu'un déterminant de taille 2×2 se calcule en faisant le produit croisé, *i.e.*, la différence du produit des termes diagonaux par les termes anti-diagonaux :

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$$

et qu'un déterminant 3×3 se calcule par la formule

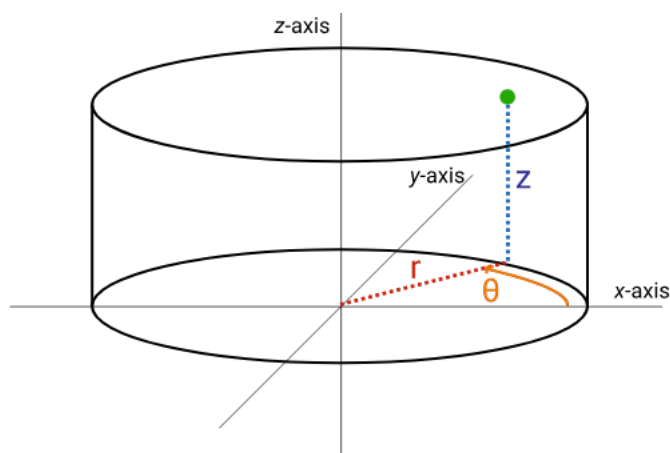
$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}.$$

Il existe une autre façon de calculer les coordonnées de $\vec{u} \times \vec{v}$. Pour cela, on écrit les deux vecteurs sous forme matricielle, et le coefficient d'une ligne du produit vectoriel est obtenu en faisant le produit croisé des deux autres lignes de \vec{u} et \vec{v} . Afin de calculer la deuxième ligne du produit vectoriel, on doit réécrire les premières lignes de \vec{u} et \vec{v} en-dessous de la troisième :

$$\begin{pmatrix} x \\ y \\ z \\ \mathbf{x} \end{pmatrix} \times \begin{pmatrix} x' \\ y' \\ z' \\ \mathbf{x}' \end{pmatrix} = \begin{pmatrix} yz' - zy' \\ zx' - xz' \\ xy' - yx' \end{pmatrix}.$$

Dans ce paragraphe, on a vu qu'on peut identifier un point indifféremment par ses coordonnées cartésiennes ou par les composantes de son vecteur position selon les vecteurs \vec{i} , \vec{j} et \vec{k} . Cette situation est possible car le plan et l'espace sont plats, *i.e.*, deux points quelconques peuvent être reliés par un vecteur. Dans le cas des espaces courbés, il existe toujours des coordonnées permettant de repérer un point, mais il n'y a cette fois-ci pas de notion satisfaisante de vecteur position. Pour s'en convaincre, il suffit de considérer un point sur la sphère et de multiplier le vecteur qui le positionne dans l'espace par un scalaire pour constater que le résultat obtenu "sort" de la sphère. Ainsi, toutes les notions introduites dans la suite utilisant les vecteurs ne peuvent pas être étendues telles quelles aux espaces courbés. Il est cependant parfois possible d'obtenir de telles extensions, mais cela nécessite d'introduire d'autres outils que les vecteurs.

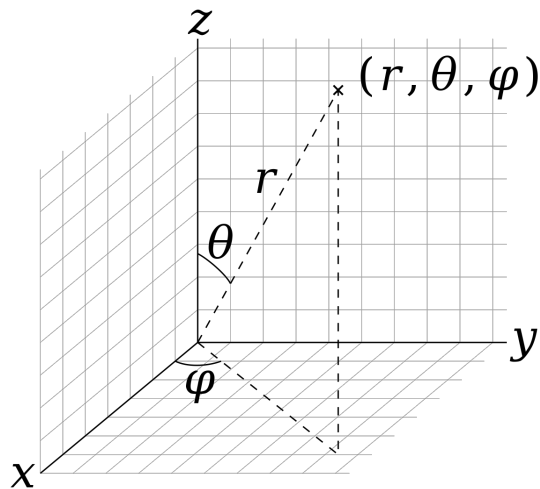
Coordonnées cylindro-polaires et sphériques. Un point dans le plan ou dans l'espace peut être repéré par d'autres coordonnées que les coordonnées cartésiennes. Les coordonnées polaires (dans le plan) et cylindriques (dans l'espace) permettent de repérer un point par sa distance r à l'axe des z , l'angle θ par rapport à l'axe des x et sa hauteur z :



Le passage des coordonnées cylindro-polaires au cartésiennes se fait par les formules suivantes :

$$x = r \cos(\theta), \quad y = r \sin(\theta), \quad z = z.$$

En coordonnées polaires, on repère un point par l'angle θ par rapport à l'axe des z , avec $0 \leq \theta \leq \pi$, l'angle φ par rapport à l'axe des x et la distance r à l'origine :



Le passage des coordonnées sphériques au cartésiennes se fait par les formules suivantes :

$$x = r \sin(\theta) \cos(\varphi), \quad y = r \sin(\theta) \sin(\varphi), \quad z = r \cos(\theta).$$

Équations paramétriques et implicites. Les premiers exemples de fonctions de plusieurs variables sont les équations paramétriques de courbes et de surfaces dans l'espace : il s'agit de fonctions vectorielles qui à 1 ou 2 paramètres, notés généralement s et t , associent un point de l'espace. Un tel point étant repéré par un vecteur position, on note les équations paramétriques par $\vec{r}(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}$ pour une courbe et $\vec{r}(s, t) = x(s, t)\vec{i} + y(s, t)\vec{j} + z(s, t)\vec{k}$ pour une surface, ou avec la notation matricielle

$$\vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \quad \text{et} \quad \vec{r}(s, t) = \begin{pmatrix} x(s, t) \\ y(s, t) \\ z(s, t) \end{pmatrix}.$$

Les courbes paramétrées permettent notamment de décrire le mouvement d'une particule dans l'espace. De façon alternative, une courbe ou surface peut être représentée par des équations implicites. Une courbe nécessite deux équations $f(x, y, z) = 0$ et $g(x, y, z) = 0$ alors qu'une surface nécessite une seule équation $h(x, y, z) = 0$. La forme d'équation utilisée dépend de la nature du problème que l'on cherche à résoudre :

- les équations paramétriques permettent, en faisant varier les paramètres t ou (s, t) , de trouver les points de la courbe ou surface correspondante ; c'est par exemple cette méthode qui est utilisée pour tracer des courbes ou surfaces sur des ordinateurs,
- les équations implicite permettent de tester si un point (x_0, y_0, z_0) appartient à la courbe ou surface en substituant les coordonnées du point dans les équations et en vérifiant si le résultat obtenu vaut 0 ou pas.

Pour se rappeler quel nombre de paramètres ou quel nombre d'équations implicites sont nécessaires pour décrire une courbe ou une surface, il faut raisonner en nombre de degrés de liberté et se rappeler que l'espace contient 3 degrés de liberté :

- une courbe étant un objet géométrique de dimension 1, elle a 1 degré de liberté, si bien que son équation paramétrique contient 1 paramètre et qu'il faut 2 équations implicites pour supprimer 2 des 3 paramètres de l'espace,
- une surface étant un objet géométrique de dimension 2, elle a 2 degrés de liberté, si bien que son équation paramétrique contient 2 paramètres et qu'il faut 1 équation implicite pour supprimer 1 des 3 paramètres de l'espace.

Droites et plans de l'espace. Les exemples les plus simples de courbes et surfaces de l'espace sont les droites et les plans.

Une droite \mathcal{D} étant une courbe, elle est décrite par 1 seule paramètre et par 2 équations implicites. Pour caractériser \mathcal{D} , il faut se donner 1 point $P(x_0, y_0, z_0) \in \mathcal{D}$, de vecteur position $\vec{r}_0 = x_0 \vec{i} + y_0 \vec{j} + z_0 \vec{k}$, et un vecteur directeur $\vec{u} = a \vec{i} + b \vec{j} + z \vec{k}$. Les points de la droite sont donc les points de vecteur position \vec{r} tels que $\vec{r} - \vec{r}_0$ est colinéaire à \vec{u} , *i.e.*, tels qu'il existe $t \in \mathbb{R}$ tel que $\vec{r} = \vec{r}_0 + t \vec{u}$. Ainsi, une équation paramétrique de \mathcal{D} est

$$\vec{r}(t) = (x_0 + at) \vec{i} + (y_0 + bt) \vec{j} + (z_0 + ct) \vec{k} = \begin{pmatrix} x_0 + at \\ y_0 + bt \\ z_0 + ct \end{pmatrix}.$$

Un point $Q(x, y, z)$ appartient à \mathcal{D} si et seulement s'il existe un t tel que $x = x_0 + at$, $y = y_0 + bt$ et $z = z_0 + ct$. En éliminant t ces équations, on trouve les équations implicites :

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

Attention : dans le cas où un des coefficients a , b ou c vaut 0, il faut comprendre que parmi les équations implicites, l'une stipule que le numérateur correspondant vaut 0. Par exemple, si \mathcal{D} passe par $P(1, 3, 0)$ et est dirigée par $\vec{u} = \vec{i} + 2 \vec{k}$, les équations implicites de \mathcal{D} sont

$$y - 3 = 0 \quad \text{et} \quad x - 1 = \frac{z}{2}.$$

Un plan \mathcal{P} étant une surface, il est décrit par un 2 paramètres et par 1 équation implicite. Pour caractériser \mathcal{P} , il faut se donner 1 point $P(x_0, y_0, z_0) \in \mathcal{P}$, de vecteur position $\vec{r}_0 = x_0 \vec{i} + y_0 \vec{j} + z_0 \vec{k}$, et un vecteur normal $\vec{n} = a \vec{i} + b \vec{j} + z \vec{k}$. Alors, un point $Q(x, y, z)$ de vecteur position $\vec{r} = x \vec{i} + y \vec{j} + z \vec{k}$ appartient au plan si et seulement si le vecteur $\vec{r} - \vec{r}_0$ est orthogonal à \vec{n} , *i.e.*, si et seulement si $(\vec{r} - \vec{r}_0) \cdot \vec{n}$ s'annule. On en déduit donc qu'une équation implicite de \mathcal{P} est $(x - x_0)a + (y - y_0)b + (z - z_0)c = 0$, qui s'écrit de la forme

$$ax + by + cz = ax_0 + by_0 + cz_0.$$

En pratique, on procède de la façon suivante pour déterminer une équation implicite : on la cherche sous la forme $ax + by + cz = d$, où d est déterminé de telle sorte que (x_0, y_0, z_0) vérifie cette équation. Pour déterminer une équation paramétrique de \mathcal{P} , il suffit de constater que $Q(x, y, z) \in \mathcal{P}$ si et seulement si le vecteur $\vec{r} - \vec{r}_0$ est une combinaison linéaire de deux vecteurs orthogonaux à \vec{n}

$$\vec{u} = \alpha_1 \vec{i} + \beta_1 \vec{j} + \gamma_1 \vec{k} \quad \text{et} \quad \vec{v} = \alpha_2 \vec{i} + \beta_2 \vec{j} + \gamma_2 \vec{k},$$

i.e., si et seulement si il existe des réels s et t tels que $\vec{r} - \vec{r}_0 = s \vec{u} + t \vec{v}$. On en déduit donc une équation paramétrique de \mathcal{P} :

$$\vec{r}(s, t) = s \vec{u} + t \vec{v} + \vec{r}_0 = \begin{pmatrix} x_0 + \alpha_1 s + \alpha_2 t \\ y_0 + \beta_1 s + \beta_2 t \\ z_0 + \gamma_1 s + \gamma_2 t \end{pmatrix}.$$

1.4. Fiche d'exercices 1

Exercice 1. Soit la droite \mathcal{D} passant par les points $(0, -1, 0)$ et $(1, 1, 2)$.

1. Donner une représentation graphique de \mathcal{D} .
2. Déterminer des équations paramétrique et implicites de \mathcal{D} .
3. Montrer que \mathcal{D} passe par le point $P(3, 5, 6)$.
4. Déterminer l'intersection de \mathcal{D} avec le plan d'équation $y = 0$.

Exercice 2. Pour chacun des plans suivants, donner une représentation graphique de celui-ci et en déterminer des équations implicite et paramétrique.

1. Le plan \mathcal{P}_1 passant par le point $(0, 0, 2)$ et perpendiculaire au vecteur \vec{k} .
2. Le plan \mathcal{P}_2 passant par les points $(1, 0, 0)$, $(0, 1, 0)$ et $(0, 0, 1)$.

Exercice 3. On considère la courbe \mathcal{C} du plan paramétrée par $\vec{r}(t) = t^2 \vec{i} + t^3 \vec{j}$.

1. Tracer les points de \mathcal{C} correspondant aux valeurs de t suivantes : $0, \pm 1, \pm 1,5$ et ± 2 .
2. Déterminer une équation implicite de \mathcal{C} . En déduire que le point $(3, \sqrt{27})$ appartient à \mathcal{C} .

Exercice 4. On considère la surface \mathcal{S} de l'espace paramétrée par $\vec{r}(s, t) = s \vec{i} + t \vec{j} + (s^2 + t^2) \vec{k}$.

1. Justifier que la surface \mathcal{S} si situe au-dessus du plan d'équation $z = 0$.
2. Soit $z_0 \geq 0$. Montrer que l'intersection de \mathcal{S} avec le plan d'équation $z = z_0$ est un cercle dont on précisera le centre et le rayon.
3. En déduire une représentation graphique de \mathcal{S} .
4. Déterminer une équation implicite de \mathcal{S} . En déduire si le point $(1, 1, 1)$ appartient à \mathcal{S} .

Exercice 5. Un projectile est émis depuis le sol à une vitesse v et un angle α par rapport au sol. On considère des coordonnées (x, y) , où x désigne la distance au sol par rapport à la position initiale du projectile et y l'altitude. En supposant que la seule force active soit la gravité, la trajectoire du projectile est décrite par le vecteur position

$$\vec{r}(t) = \begin{pmatrix} v \cos(\alpha) t \\ -\frac{1}{2} g t^2 + v \sin(\alpha) t \end{pmatrix}.$$

1. Calculer une équation implicite désignant le mouvement du projectile.
2. On suppose que v^2 mesurée en $m^2.s^{-2}$ est égale à g mesurée en $m.s^{-2}$, et on suppose que $\alpha = 30^\circ$. Le projectile passe-t-il par le point $(1/2, 1/2)$? Au bout de combien de temps et à quelle distance de la position initiale le projectile retombe-t-il au sol?

Chapitre 2 :

Fonctions de plusieurs variables

Dans ce chapitre, on introduit les fonctions de plusieurs variables et on s'attache à représenter deux types de telles fonctions : les champs scalaires et les champs de vecteurs. On introduit ensuite les dérivées partielles et directionnelles, qui permettent d'analyser les variations du premier ordre d'une fonction scalaire dans une direction donnée. Enfin, on présente la règle de dérivation des fonctions composées, dont on illustre le rôle pour calculer effectivement les dérivées directionnelles. On illustre également le rôle de cette règle en physique grâce à deux exemples : la dérivée particulaire et l'existence de solutions de l'équation de la chaleur sous forme d'exponentielle.

En toute généralité, une fonction de plusieurs variables est définie sur un domaine $D \subseteq \mathbb{R}^n$ et est à valeurs dans \mathbb{R}^m , où n et m sont des entiers. On fixe quelques conventions et points de terminologie utilisés dans le reste de ce cours :

- dans les exemples et sujets d'exercices, n vaut 2, 3 ou 4, et les variables sont notées (x, y) ou (t, x) quand $n = 2$, x, y, z quand $n = 3$ et t, x, y, z quand $n = 4$, avec la convention que t est la coordonnée temporelle et x, y, z sont des coordonnées spatiales,
- dans les parties théoriques du cours, l'entier n peut être quelconque et dans ce cas, les variables sont notées x_1, \dots, x_n ,
- lorsque l'entier m est supérieur ou égal à 2, on dit que la fonction est une *fonction vectorielle*, et dans le cas $m = 1$, on dit que la fonction est une *fonction scalaire*, un *champ scalaire*, voire tout simplement une *fonction*,
- lorsque $n = m$, la fonction est appelée un *champ de vecteurs*¹,
- l'ensemble D est appelé le *domaine de définition* de la fonction; la plupart du temps, le domaine est \mathbb{R}^n tout entier, même si on considère quelques exemples où le domaine est \mathbb{R}^n privé de l'origine.

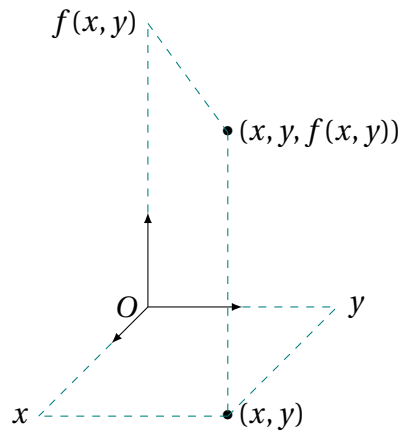
2.1. Représentations des champs scalaires et vectoriels

On commence par introduire deux façons de représenter une fonction scalaire² et une façon de représenter les champs de vecteurs. Pour des raisons de clarté, on se limite au cas de 2 variables.

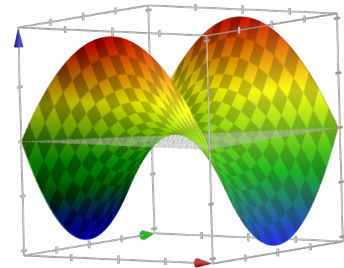
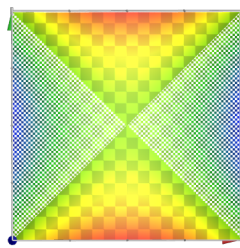
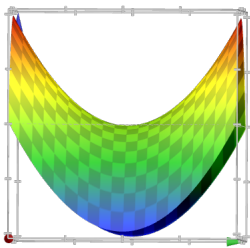
Graphe d'un champ scalaire. Soit $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto f(x, y)$ une fonction de 2 variables. Son graphe est la surface d'équation paramétrique $\vec{r}(x, y) = x \vec{i} + y \vec{j} + f(x, y) \vec{k}$. Cela signifie qu'au-dessus de chaque (x, y) du plan, se trouve un point à la hauteur $f(x, y)$:

1. Il y a ici un abus de langage : les champs de vecteurs sont des objets plus compliqués pouvant être représentés par des fonctions de \mathbb{R}^n dans \mathbb{R}^n , alors qu'une telle fonction est en fait une transformation de \mathbb{R}^n . Cependant, les fonctions de \mathbb{R}^n dans lui-même qui interviennent dans ce cours sont implicitement supposées représenter des champs de vecteurs.

2. Il existe cependant des méthodes alternatives, qu'on n'aborde pas dans ce cours.



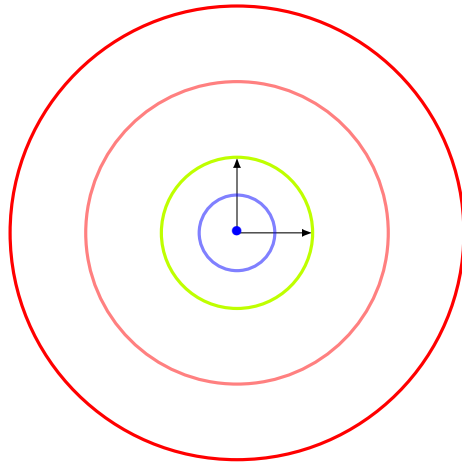
Le *graphe* ou *surface représentative* de f , est obtenu(e) en reliant tous ces points. Pour obtenir une visualisation animée du graphe de f , on entre l'expression de la fonction dans google. Par exemple, si on tape " $y^2 - x^2$ ", on obtient une image animée du graphe de la fonction définie par $f(x, y) = y^2 - x^2$. On représente ici quelques pauses sur cette image : la première est une vue "de côté" (l'axe des x est orthogonal à l'écran et celui des y est l'axe horizontal), la deuxième une vue de dessus et la troisième permet d'avoir une vision plus globale de la surface :



Courbes de niveaux d'un champ scalaire. Une deuxième façon de représenter une fonction scalaire f est par ses courbes de niveaux. Il s'agit de courbes du plan d'équations implicites $f(x, y) = c$, où c est une constante. En pratique, on peut utiliser un code couleur avec la convention suivantes :

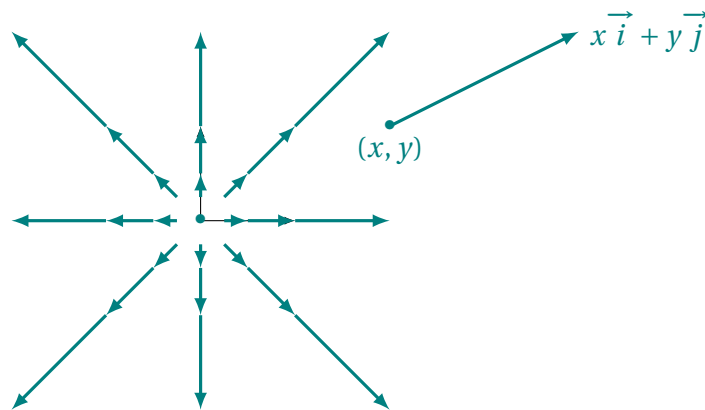
- les courbes correspondant à une valeur c positive sont en rouge, et plus la valeur de c est élevée, plus le rouge est prononcé,
- courbe correspondant à $c = 0$ est représentée dans une couleur type jaune ou vert clair,
- les courbes correspondant à une valeur c négative sont en bleu, et plus la valeur de c est petite, plus le bleu est prononcé.

Parfois, la fonction f ne change pas de signe, *e.g.*, s'il s'agit d'une exponentielle. Dans ce cas, la convention est d'utiliser 1 au lieu de 0 comme valeur pivot. De plus, on voit qu'un champ scalaire peut être représenté comme une coloration du plan, ce qui correspond à l'idée de champ mentionnée dans le chapitre précédent : au-dessus de chaque point du champ, se trouve une couleur représentant un scalaire! On illustre cela en traçant quelques courbes de niveaux de la fonction définie par $f(x, y) = x^2 + y^2$:



Les équations $x^2 + y^2 = c$ n'ont de solution que quand $c \geq 0$, et dans ce cas il s'agit d'un cercle centré en l'origine et de rayon \sqrt{c} . En particulier, pour $c = 0$, la courbe est en fait un point.

Champs de vecteurs. On s'intéresse maintenant au cas $n = m = 2$, *i.e.*, aux champs de vecteurs du plan. Une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ s'écrit sous la forme $f(x, y) = (f_1(x, y), f_2(x, y))$, où f_1 et f_2 sont des fonctions scalaires. La fonction f peut être vue comme la fonction qui à (x, y) associe le vecteur $f_1(x, y) \vec{i} + f_2(x, y) \vec{j}$. Pour la représenter, il suffit d'attacher en chaque (x, y) le vecteur $f_1(x, y) \vec{i} + f_2(x, y) \vec{j}$: on retrouve bien l'idée de champ puisqu'en tout point du plan se trouve un vecteur. On illustre cette représentation par le *champ d'Euler* défini par $f(x, y) = (x, y)$:



2.2. Dérivées partielles et directionnelles

Dans cette section, on emploie la convention annoncée en préambule de chapitre : dans toutes les définitions générales, on considère des fonctions dépendant des variables x_1, \dots, x_n .

Dérivées premières. La *dérivée partielle* d'une fonction scalaire f en un point $p(x_1, \dots, x_n)$ par rapport à la variable x_i est la limite du taux d'accroissement relativement à x_i de f en p :

$$\frac{\partial f}{\partial x_i}(p) = \lim_{dx_i \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + dx_i, x_{i+1}, \dots, x_n) - f(p)}{dx_i}.$$

On remarque qu'en fait une dérivée partielle est une dérivée ordinaire où toutes les variables à l'exception de celle par rapport laquelle la fonction est dérivée sont fixées. Ainsi, pour dériver une fonction par rapport à une variable, on dérive comme si toutes les autres variables sont des constantes. Les dérivées

partielles vérifient les règles de linéarité et de produit. Par exemple, si f et g sont deux fonctions et λ est un scalaire, on a les égalités fonctionnelles :

$$\frac{\partial(f + \lambda g)}{\partial x_i} = \frac{\partial f}{\partial x_i} + \lambda \frac{\partial g}{\partial x_i} \quad \text{et} \quad \frac{\partial(fg)}{\partial x_i} = f \frac{\partial g}{\partial x_i} + \frac{\partial f}{\partial x_i} g.$$

Par exemple, si f est la fonction de 2 variables x, y définie par $f(x, y) = 2x^2y + xy^2 - x^2$, on a :

$$\frac{\partial f}{\partial x}(x, y) = 4xy + y^2 - 2x \quad \text{et} \quad \frac{\partial f}{\partial y}(x, y) = 2x^2 + 2xy.$$

Dérivées secondes. Les dérivées partielles définissant des nouvelles fonctions, elles peuvent elles-mêmes être dérivées, et ces nouvelles dérivées sont appelées les *dérivées secondes*. On les note :

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) \quad \text{et} \quad \frac{\partial f}{\partial x_i^2} = \frac{\partial f}{\partial x_i \partial x_i}.$$

Les dérivées correspondant à $i \neq j$, sont parfois appelées *dérivées croisées* ou *dérivées mixtes*. Le *théorème de Schwarz* stipule que lorsque la fonction f est suffisamment régulière, ce qui est le cas tout au long de ce cours, les dérivées croisées commutent :

$$\forall 1 \leq i, j \leq n : \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Bien sûr, il est possible de dériver les dérivées secondes, puis leurs dérivées, etc. Cependant, dans ce cours, on ne s'intéresse qu'aux dérivées premières et secondes.

En reprenant l'exemple précédent $f(x, y) = 2x^2y + 2xy^2 - x^2$, on a

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 4y - 2, \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = 4x + 2y = \frac{\partial^2 f}{\partial y \partial x}(x, y), \quad \frac{\partial^2 f}{\partial y^2}(x, y) = 2x.$$

Dérivées directionnelles. Les dérivées partielles premières d'une fonction f permettent de mesurer les variations de f selon une de ses coordonnées, ou de façon analogue, selon l'un des axes de coordonnées de \mathbb{R}^n . On peut plus généralement s'intéresser à la variation de f dans une direction donnée, au moyen de la *dérivée directionnelle*. Une direction étant déterminée par un vecteur \vec{v} , la dérivée de la fonction f dans la direction \vec{v} en un point p est définie par :

$$\frac{\partial f}{\partial \vec{v}}(p) = \lim_{h \rightarrow 0} \frac{f(p + h\vec{v}) - f(p)}{h}. \quad (2.1)$$

On finit cette section par quelques remarques sur les dérivées directionnelles :

- la méthode pratique pour calculer les dérivées directionnelle est présentée dans la prochaine section,
- selon certaines conventions, le vecteur \vec{v} doit être unitaire, afin de mesurer la variation de f par unité de variation,
- la dérivée partielle par rapport à la variable x_i n'est rien d'autre que la dérivée dans la direction du i ème vecteur de base,
- la notation $p + h\vec{v}$ est abusive : il faudrait en fait considérer le points de vecteur position $\vec{r}(p) + h\vec{v}$.

2.3. Règle de dérivation des fonctions composées

On présente maintenant une formule appelée *règle de dérivation des fonctions composées*, ou *chain rule* en anglais. Avant de donner l'expression générale de cette formule, on présente plusieurs cas particuliers.

Cas des fonctions scalaires d'une variable. On considère deux fonctions f et g d'une seule variable t et on souhaite dériver la composée $f \circ g$. Une petite variation dt de t induit une petite variation $dg = g'(t)dt$ et une petite variation de g induit une petite variation $df = f'(g)dg$. En combinant ces deux expressions, on en déduit la petite variation de la composée $d(f \circ g) = f'(g(t))d(g(t)) = f'(g(t))g'(t)dt$. En divisant par dt , on en déduit la formule de la chain rule pour les fonctions scalaires d'une variable :

$$(f \circ g)'(t) = f'(g(t))g'(t).$$

Dérivée totale d'une fonction scalaire. On considère maintenant une composée $\mathbb{R} \rightarrow \mathbb{R}^n \rightarrow \mathbb{R}$. Intuitivement :

- la fonction $\mathbb{R}^n \rightarrow \mathbb{R}$ mesure une quantité sur \mathbb{R}^n , on la note $(x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n)$,
- la fonction $\mathbb{R} \rightarrow \mathbb{R}^n$ définit une trajectoire dans \mathbb{R}^n , on la note $t \mapsto (x_1(t), \dots, x_n(t))$,
- la composée des deux fonctions $t \mapsto f(x_1(t), \dots, x_n(t))$ permet de mesurer la quantité de f le long de la trajectoire.

En supposant qu'une seule des variables, notée x_i , admet une petite variation et que les autres sont fixés, f admet une petite variation $df = \frac{\partial f}{\partial x_i}(x_1, \dots, x_n)dx_i$. Il est de plus naturel d'admettre le principe de superposition dans ce contexte : lorsqu'on autorise des petites variations de toutes les variables x_1, \dots, x_n simultanément, on ajoute les petites variations, *i.e.*,

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_1, \dots, x_n)dx_i = \frac{\partial f}{\partial x_1}(x_1, \dots, x_n)dx_1 + \dots + \frac{\partial f}{\partial x_n}(x_1, \dots, x_n)dx_n.$$

De plus, une petite variation de t induit une petite variation $(dx_1, \dots, dx_n) = (\dot{x}_1(t)dt, \dots, \dot{x}_n(t)dt)$ de la courbe³. On en déduit qu'une petite variation de t induit une petite variation :

$$d(f \circ (x_1, \dots, x_n)) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_1(t), \dots, x_n(t))\dot{x}_i(t)dt.$$

On en déduit donc l'expression de la chain rule dans ce contexte :

$$\frac{d}{dt}(f(x_1(t), \dots, x_n(t))) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_1(t), \dots, x_n(t))\dot{x}_i(t). \quad (2.2)$$

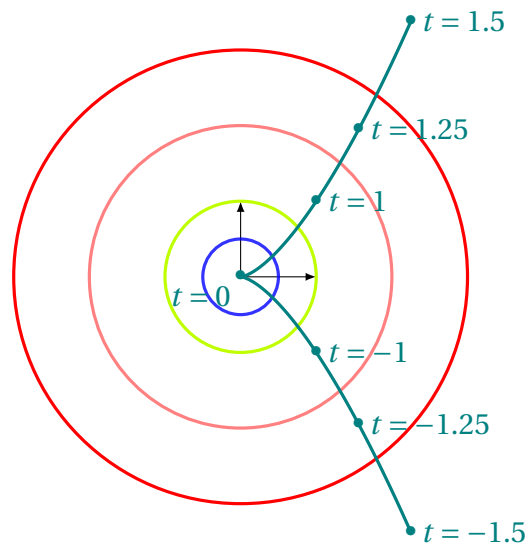
En physique, la dérivée (2.2) est également appelée *dérivée totale*.

On illustre cette forme de la chain rule en prenant $n = 2$, $f(x, y) = x^2 + y^2$ et $(x(t), y(t)) = (t^2, t^3)$. Alors, la dérivée totale de f le long de la trajectoire se calcule grâce à la formule (2.2) :

$$\begin{aligned} \frac{d}{dt}(f(x(t), y(t))) &= \frac{\partial f}{\partial x}(x(t), y(t))\dot{x}(t) + \frac{\partial f}{\partial y}(x(t), y(t))\dot{y}(t) \\ &= 2x(t)\dot{x}(t) + 2y(t)\dot{y}(t) \\ &= 2t^2(2t) + 2t^3(3t^2) \\ &= 6t^5 + 4t^3. \end{aligned}$$

3. On note cette fois-ci les dérivées avec des points plutôt que de prime pour coller à l'intuition physique : dériver par rapport à t revient à prendre une vitesse.

On peut vérifier la cohérence de ce résultat puisque $f(x(t), y(t)) = (t^2)^2 + (t^3)^2 = t^4 + t^6$. On peut visualiser ce résultat en commençant par représenter la fonction f par ses courbes de niveaux, la courbe dans le plan et en graduant celle-ci :



Alors, la dérivée totale de f indique combien de courbes de niveaux sont traversées par unité infinitésimale de temps.

Forme générale. Il reste à énoncer la forme générale de la chain rule pour une fonction à valeurs vectorielles $g : \mathbb{R}^m \rightarrow \mathbb{R}^n$, dont on a écrit g_i la i ème fonction coordonnée. Intuitivement, si $m = 2$, g peut être pensée comme une paramétrisation d'une surface dans \mathbb{R}^n , d'un volume si $m = 3$, etc. En notant y_i , avec $1 \leq i \leq m$, les coordonnées dans \mathbb{R}^m , on peut dériver la composée $f \circ g$ par rapport à chacune de celles-ci. On a alors pour tout $p \in \mathbb{R}^m$:

$$\frac{\partial (f \circ g)}{\partial y_i}(p) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(g(p)) \frac{\partial g_j}{\partial y_i}(p).$$

On illustre maintenant l'utilité de la chain rule avec le calcul de la dérivée directionnelle et avec deux exemples issus de la physique.

Calcul pratique de la dérivée directionnelle. On commence par considérer une fonction scalaire f dépendant de trois variables x, y, z , un vecteur $\vec{v} = a\vec{i} + b\vec{j} + c\vec{k}$ et un point p . D'après la formule (2.1), la dérivée dans la direction \vec{v} de f en p est la dérivée de la composée de f avec la fonction d'une variable définie par $h \mapsto p + (ah, bh, ch)$. D'après la forme (2.2) de la chain rule, on a donc

$$\frac{\partial f}{\partial \vec{v}}(p) = \frac{\partial f}{\partial x}(p)a + \frac{\partial f}{\partial y}(p)b + \frac{\partial f}{\partial z}(p)c.$$

On remarque que cette quantité est le produit scalaire de \vec{v} par le vecteur $\overrightarrow{\text{grad}}(f)(p)$ dont les composantes sont les dérivées partielles de f en p :

$$\overrightarrow{\text{grad}}(f)(p) = \begin{pmatrix} \frac{\partial f}{\partial x}(p) \\ \frac{\partial f}{\partial y}(p) \\ \frac{\partial f}{\partial z}(p) \end{pmatrix}.$$

Ce vecteur dépendant de p , on obtient en fait le champ de vecteurs, appelé *gradient* de f :

$$\overrightarrow{\text{grad}}(f) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix}.$$

Le gradient de f est parfois noté $\vec{\nabla} f$, puisqu'il s'agit de l'opérateur *nabla*

$$\vec{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}$$

appliqué à la fonction f . On retrouve l'opérateur ∇ dans le chapitre 4, au moment où on parle des autres opérateurs différentiels classiques : *rotationnel*, *divergence* et *Laplacien*.

Attention : les composantes du gradient sont les dérivées partielles en coordonnées cartésiennes uniquement.

Bien sûr, pour une fonction d'un nombre n de variables, la dérivée directionnelle se calcule selon le même principe : il s'agit du produit scalaire de l'opérateur gradient par le vecteur de la direction considérée. Les propriétés, la signification géométrique et les expressions en coordonnées cylindriques et sphériques du gradient sont présentées dans le chapitre 3.

On illustre le calcul de la dérivée directionnelle par un exemple : on considère la fonction de 3 variables définie par $f(x, y, z) = 2xyz + x - z$ et $\vec{v} = \vec{i} - \vec{j} + 2\vec{k}$. Le gradient de f est le champ de vecteur défini par

$$\vec{\text{grad}}(f)(x, y, z) = \begin{pmatrix} 2yz + 1 \\ 2xz \\ 2xy - 1 \end{pmatrix},$$

de sorte que

$$\frac{\partial f}{\partial \vec{v}}(x, y, z) = 2yz + 1 - 2xz + 2(2xy - 1) = 4xy - 2xz + 2yz - 1.$$

Dérivée particulière. En mécanique des fluides, la dérivée totale (2.2) est appelée *dérivée particulière*. On considère une quantité scalaire φ associée à une particule fluide, *e.g.*, sa température ou sa masse volumique, dépendant à la fois de la position dans le fluide mais également du temps. En effet, on peut imaginer que le fluide est chauffé par moments et donc que la température dépend explicitement du temps. On se fixe un système de coordonnées cartésiennes et une particule dont le mouvement admet pour équation paramétrique $\vec{r}(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}$. Alors, la variation de φ au cours du temps est décrite par la dérivée particulière, notée avec un D majuscule droit :

$$\frac{D\varphi}{Dt}(t) = \frac{\partial\varphi}{\partial t}(t, x(t), y(t), z(t)) + \frac{\partial\varphi}{\partial x}(t, x(t), y(t), z(t))\dot{x}(t) + \frac{\partial\varphi}{\partial y}(t, x(t), y(t), z(t))\dot{y}(t) + \frac{\partial\varphi}{\partial z}(t, x(t), y(t), z(t))\dot{z}(t).$$

En physique, cet opérateur de dérivation est généralement noté de la façon suivante :

$$\frac{D\varphi}{Dt} = \frac{\partial\varphi}{\partial t} + (\vec{v} \cdot \vec{\nabla})\varphi.$$

Le vecteur $\vec{v} = \dot{x}(t)\vec{i} + \dot{y}(t)\vec{j} + \dot{z}(t)\vec{k}$ désigne la vitesse de la particule, l'opérateur $\vec{\nabla}$ désigne l'opérateur nabla **relativement aux variables spatiales uniquement** et $\vec{v} \cdot \vec{\nabla}$ désigne l'opérateur produit scalaire de \vec{v} par le gradient du champ donné en argument. L'intérêt d'introduire une notation spécifique pour l'opérateur $\vec{v} \cdot \vec{\nabla}$ est qu'il peut également être appliquée à un champ de vecteurs. Si par exemple, on considère un champ de vitesses $\vec{u} = u_x\vec{i} + u_y\vec{j} + u_z\vec{k}$, on a

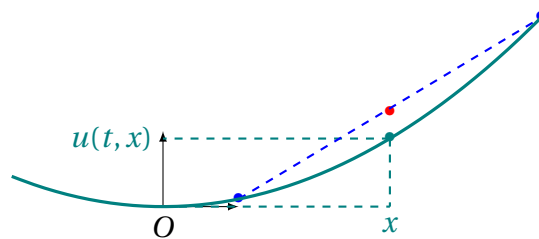
$$(\vec{v} \cdot \vec{\nabla})\vec{u} = \begin{pmatrix} \vec{v} \cdot \vec{\text{grad}}(u_x) \\ \vec{v} \cdot \vec{\text{grad}}(u_y) \\ \vec{v} \cdot \vec{\text{grad}}(u_z) \end{pmatrix}.$$

Équation de la chaleur. La diffusion de la chaleur au cours du temps dans une barre uni-dimensionnelle est régie par l'équation aux dérivées partielles

$$\frac{\partial u}{\partial t}(t, x) = \alpha \frac{\partial^2 u}{\partial x^2}(t, x),$$

où α est le coefficient de diffusivité thermique du milieu dans lequel se propage la chaleur, t désigne le temps et x est une coordonnée spatiale permettant de se repérer sur la barre.

Bien qu'on puisse déduire l'équation de la chaleur de la loi de Fick, on va se contenter de comprendre ce qu'elle signifie. Pour cela, on rappelle d'abord que le signe d'une dérivée seconde indique si une courbe se trouve au-dessus (signe positif) ou en-dessous (signe négatif) de sa tangente. On fixe un instant t et une position x dans la barre tels que $\frac{\partial^2 u}{\partial x^2}(t, x) > 0$. Sur le dessin suivant, on trace la température à l'instant t au voisinage de x (l'axe vertical repère la position sur la barre et l'axe horizontal la température à l'instant t en fonction de la position) :



En bleu, on a représenté deux points situés en positions $x - \varepsilon$ et $x + \varepsilon$ sur la barre. On remarque que la température moyenne de la barre (point rouge) autour de x est supérieure à la valeur en x . Ainsi, il est naturel de considérer que dans cette situation, la température en x a tendance à augmenter dans le futur, ce qui signifie que $\frac{\partial u}{\partial t}(t, x) > 0$. De même, si $\frac{\partial^2 u}{\partial x^2}(t, x) < 0$, il est naturel d'avoir $\frac{\partial u}{\partial t}(t, x) < 0$. Enfin, plus le milieu est conducteur, *i.e.*, plus α est grand, plus on s'attend à ce que la température augmente rapidement. L'équation de la chaleur rend bien compte de toutes ces considérations.

On cherche maintenant une solution de l'équation de la chaleur sous forme d'exponentielle. Étant données deux constantes a et b , on considère la fonction définie par

$$u(t, x) = e^{ax+bt}.$$

Cette fonction est la composée des deux fonctions définies par $f(t, x) = ax + bt$ et $g(x) = e^x$. D'après la chain rule, on a

$$\frac{\partial u}{\partial t}(t, x) = g'(ax + bt) \cdot \frac{\partial f}{\partial t}(t, x) = e^{ax+bt} b \quad \text{et} \quad \frac{\partial u}{\partial x}(t, x) = g'(ax + bt) \cdot \frac{\partial f}{\partial x}(t, x) = e^{ax+bt} a.$$

En appliquant de nouveau la chain rule, on a

$$\frac{\partial^2 u}{\partial x^2}(t, x) = e^{ax+bt} a^2.$$

On a donc

$$\frac{\partial u}{\partial t}(t, x) = \frac{b}{a^2} \cdot \frac{\partial^2 u}{\partial x^2}(t, x),$$

de sorte que u est solution de l'équation de la chaleur pour $\alpha = b/a^2$. La solution proposée ici n'est pas une solution physiquement valide, sinon cela signifierait que la température de la barre augmenterait indéfiniment lorsque le temps s'écoule : en effet, $\alpha = b/a^2$ et $\alpha > 0$ impliquent $b > 0$. Pour étudier l'existence de solutions physiquement valides, il faut faire appel aux *séries de Fourier* qui ne sont pas au programme de ce cours.

2.4. Fiche d'exercices 2

Exercice 1. Représenter les courbes de niveaux des fonctions définies par les expressions suivantes :

$$f(x, y) = y - 2, \quad g(x, y) = x, \quad h(x, y) = e^{y-x^2}, \quad k(x, y) = (xy)^2.$$

Exercice 2. Apparier la fonction avec sa représentation graphique. Justifier.

1. $f(x, y) = |x| + |y|$

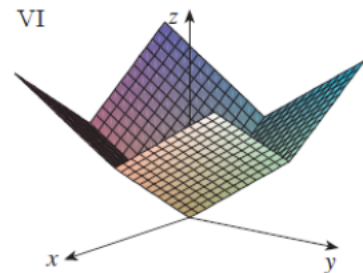
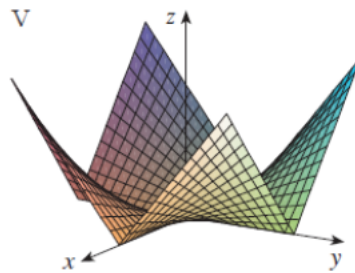
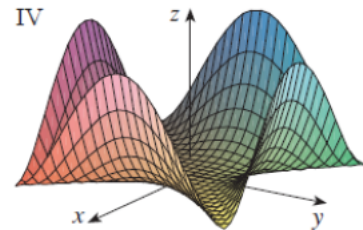
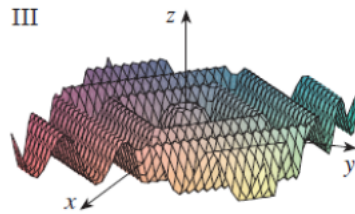
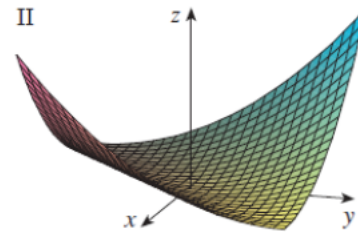
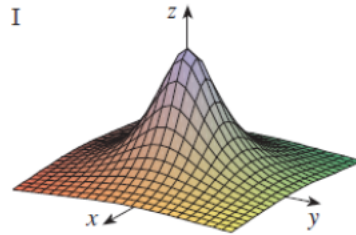
2. $f(x, y) = |xy|$

3. $f(x, y) = \frac{1}{1+x^2+y^2}$

4. $f(x, y) = (x^2 - y^2)^2$

5. $f(x, y) = (x - y)^2$

6. $f(x, y) = \sin(|x| + |y|)$



Exercice 3. On considère la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = x \cos(y) + 2x^2y + e^y$.

1. Calculer les dérivées partielles premières de f .
2. Calculer les dérivées de f dans les directions $\vec{u} = \vec{i} + \vec{j}$ et $\vec{v} = 2\vec{i} + 3\vec{j}$.
3. Calculer les dérivées secondes de f . Vérifier que les dérivées croisées sont égales.

Exercice 4. On considère la fonction $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ définie par $f(x, y, z) = x^2 + y^2 + z^2$.

1. Soit $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ la fonction vectorielle définie par $\sigma(s, t) = (s, t, s^2 + t^2)$. Calculer les dérivées partielles de la fonction $f \circ \sigma : \mathbb{R}^2 \rightarrow \mathbb{R}$.
2. Calculer la dérivée totale de f le long de la trajectoire définie par $h(t) = (\cos(t), \sin(t), t)$. Comparer à la dérivée de f en $h(t)$ dans la direction du vecteur $\vec{v}(t) = \dot{h}(t)$.

Exercice 5. Étant données une constante $k \in \mathbb{R}$ et une fonction $Q : \mathbb{R} \rightarrow \mathbb{R}$, on considère la fonction de 4 variables définie par :

$$\forall (t, x, y, z) \in \mathbb{R}^4 : V(t, x, y, z) = \frac{kQ(t)}{\sqrt{x^2 + y^2 + z^2}}.$$

1. Calculer les dérivées partielles de la fonction V .
2. Calculer la dérivée totale de V selon la trajectoire d'équation paramétrique $\vec{r}(t) = (x(t), y(t), z(t))$.

Exercice 6. Le but de cet exercice est de déduire et résoudre l'équation de transport en une dimension spatiale. Soient un fluide se déplaçant de façon rectiligne à une vitesse constante c , et $f(x)$ une quantité scalaire dans ce fluide⁴. La quantité f se déplaçant avec le fluide, on introduit $u(t, x)$ la valeur de f à l'instant t et la position x .

1. Quel est le comportement de u le long des droites d'équation $x = x_0 + ct$, où x_0 est une constante? En déduire l'équation de transport :

$$\frac{\partial u}{\partial t}(t, x) + c \frac{\partial u}{\partial x}(t, x) = 0.$$

2. Quelle est la signification géométrique de l'équation de transport?
3. Donner la solution de l'équation de transport.

Exercice 7. Dans cet exercice on considère une surface \mathcal{S} de l'espace définie par l'équation $f(x, y, z) = 0$, où f est une fonction. On admet que pour la plupart des points de \mathcal{S} , au moins une des coordonnées peut s'exprimer localement comme fonction des deux autres.

1. Dans cette question, on suppose que f est la fonction définie par $f(x, y, z) = x^2 + y^2 + z^2 - 1$.
 - a). Quelle est la surface \mathcal{S} d'équation $f(x, y, z) = 0$?
 - b). Sous quelle condition sur un point (x, y, z) de \mathcal{S} , la coordonnée z peut s'exprimer comme fonction des deux autres coordonnées? Donner la forme explicite de cette fonction.
 - c). Répondre aux questions du point b). en considérant les coordonnées x puis y à la place de z .
2. Sur \mathcal{S} , exprimer dx en fonction de dy et dz puis dz en fonction de dx et dy .
3. Quelle relation vérifient entre elles dx et dy sur \mathcal{S} ? En déduire qu'en tout point (x, y, z) de \mathcal{S} , on a la formule, connue sous le nom de *triple product rule* en anglais, suivante :

$$\frac{\partial x}{\partial y}(y, z) \times \frac{\partial y}{\partial z}(x, z) \times \frac{\partial z}{\partial x}(x, y) = -1.$$

4. On rappelle la loi des gaz parfaits : $PV = nRT$, où P est la pression, V le volume du gaz, T la température absolue, n la quantité de matière et R la constante universelle des gaz parfaits.
 - a). À n fixé, exprimer la loi des gaz parfaits sous forme d'une équation de 3 variables.
 - b). Retrouver la triple product rule par un calcul direct dans cette situation.

4. Par exemple, f peut représenter une concentration, une pression ou la hauteur d'une vague.

Chapitre 3 :

Optimisation

Ce chapitre porte sur l'optimisation d'une fonction scalaire, *i.e.*, la recherche des points où cette fonction admet une valeur qui maximise ou minimise localement la fonction. Il existe deux formes d'optimisation : avec ou sans contrainte. Pour les fonctions sans contrainte, la recherche d'extrema, (maxima ou de minima), se fait en deux étapes : on localise d'abord les *points critiques* d'une fonction, à savoir les points qui annulent le gradient, et ensuite on classe ces points critiques grâce à un nouvel outil appelé *matrice Hessienne*. Lorsque la fonction est contrainte, la recherche d'extrema respectant ces contraintes se ramène à la résolution d'un système d'équation. Cette méthode de recherche s'appelle *méthode des multiplicateurs de Lagrange*.

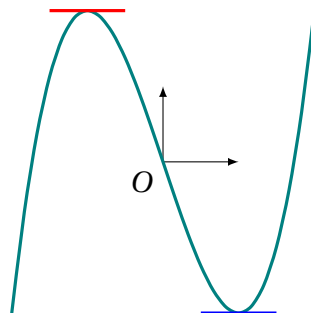
Certaines généralités présentées dans ce chapitre sont valides pour une fonction d'un nombre quelconque de variables. Cependant, lorsqu'on présente le critère de la Hessienne et la méthode des multiplicateurs de Lagrange, on se limite aux fonctions de deux variables. Ainsi, dans tout ce chapitre, lorsque aucune précision n'est donnée, f désigne une fonction de n variables x_1, \dots, x_n , et on mentionne explicitement les endroits où on se limite au cas $n = 2$.

3.1. Gradient et points critiques

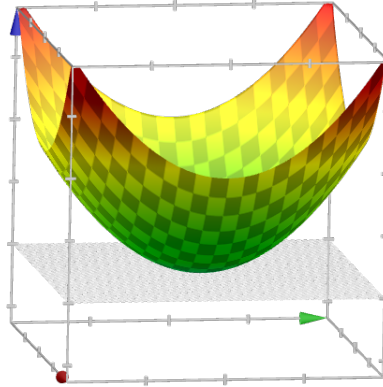
Extrema locaux. On dit que f admet un *extremum local* au point $p \in \mathbb{R}^n$, si au voisinage de p , l'une des deux conditions suivantes est vérifiée :

- les valeurs de f sont inférieures à $f(p)$; dans ce cas p est appelé un *maximum local*,
- les valeurs de f sont supérieures à $f(p)$; dans ce cas p est appelé un *minimum local*.

Géométriquement, cela signifie qu'au voisinage de p , le graphe de f se trouve en-dessous (si p est un maximum local) ou au-dessus (si p est un minimum local) de $f(p)$. Par exemple pour $n = 1$, la fonction suivante admet un maximum local en le point tangent au trait rouge et un minimum local en le point tangent au trait bleu :



On remarque qu'aucun de ces extrema n'est un extremum global puisque la fonction tend vers $\pm\infty$ quand $x \rightarrow \pm\infty$. Sur le dessin suivant on donne un exemple en deux variables, où la fonction admet un minium local (même global) au point tangent au plan gris :



Propriétés du gradient. Pour les fonctions d'une variables, *i.e.*, quand $n = 1$, on sait que les extrema locaux sont situés en les points où la dérivée s'annule. Pour étendre ce résultat au cas $n \geq 2$, on a besoin d'analyser les propriétés géométriques du gradient, dont on rappelle l'expression en coordonnées cartésiennes :

$$\overrightarrow{\text{grad}}(f)(p) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial f}{\partial x_n}(p) \end{pmatrix}.$$

Le gradient vérifie les propriétés géométriques fondamentales suivantes :

- il indique la direction dans laquelle la fonction croit le plus rapidement,
- plus sa norme est élevée, plus la fonction croit rapidement dans sa direction,
- il est orthogonal aux ensembles de niveaux de f .

On commence par montrer les deux premiers points : en chaque point p , la direction dans laquelle f grandit le plus vite est déterminée par un vecteur \vec{v} qui maximise la dérivée de f en p dans la direction \vec{v} . Or, on rappelle que cette dérivée directionnelle est égale à $(\overrightarrow{\text{grad}}(f)(p)) \cdot \vec{v}$, de sorte qu'elle est maximisée quand $\vec{v} = \overrightarrow{\text{grad}}(f)(p)$, ce qui montre le premier point, et dans ce cas on a :

$$\frac{\partial f}{\partial \vec{v}}(p) = \left| \overrightarrow{\text{grad}}(f)(p) \right|^2,$$

ce qui montre le deuxième point. Pour le dernier point, on considère un ensemble de niveau N de f , *i.e.*, tel que la restriction de f à N est constante, et on se donne un vecteur \vec{v} tangent à N en $p \in N$. Alors, puisque f est constante sur N , elle ne varie pas dans la direction de \vec{v} , ce qui signifie que

$$\frac{\partial f}{\partial \vec{v}}(p) = (\overrightarrow{\text{grad}}(f)(p)) \cdot \vec{v} = 0.$$

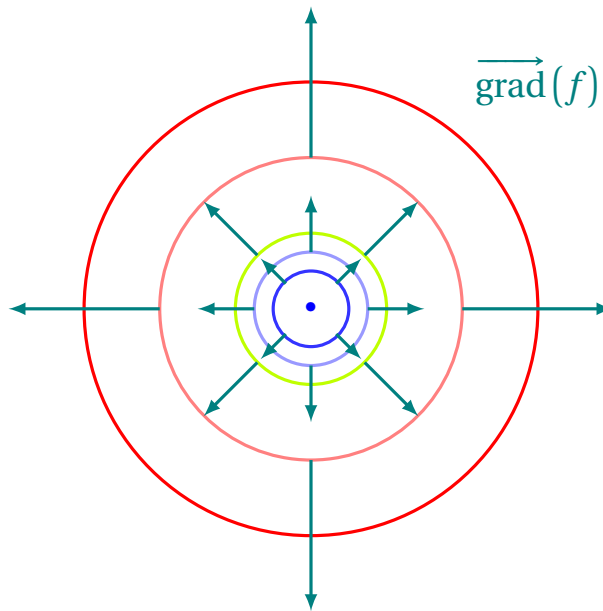
Ainsi, $\overrightarrow{\text{grad}}(f)(p)$ est orthogonal à tous les vecteurs tangents aux ensembles de niveau, ce qui montre le dernier point. Intuitivement, ces trois points donnent des informations cohérentes : ils signifient que le gradient s'"échappe" d'une courbe de niveau de f pour en rejoindre une autre située au dessus, et cette échappée est d'autant plus rapide que la fonction varie rapidement. Ce point de vue justifie

également la terminologie : les ensembles de niveaux peuvent être vus comme une graduation de l'espace et cette graduation étant orthogonale au gradient, elle suit la direction indiquée par celui-ci. En ce sens, le gradient "gradue" l'espace.

On considère par exemple la fonction de deux variables définie par $f(x, y) = x^2 + y^2$. Alors en tout point $p(x, y)$ du plan, le gradient

$$\overrightarrow{\text{grad}}(f)(p) = \begin{pmatrix} 2x \\ 2y \end{pmatrix},$$

est un vecteur radial¹ et donc orthogonal aux cercles centrés en l'origine passant par p , ces cercles étant précisément les courbes de niveaux de f :



Gradient en coordonnées non cartésiennes. On rappelle qu'en coordonnées non cartésiennes, les composantes du gradient de f ne sont pas les dérivées partielles de f . En fait, celles-ci définissent la *différentielle* df , qui est un vecteur ligne (également appelé *covecteur*), et qui agit sur les vecteurs par produit matriciel; le gradient étant défini comme le vecteur qui transforme ce produit matriciel en produit scalaire. Plus précisément, des petites variations dx , dy et dz induisent une petite variation df , représentée par un covecteur :

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix}.$$

Alors, étant donné un point p , le gradient est défini par la formule suivante, vraie pour tout vecteur de l'espace $\vec{v} = a\vec{i} + b\vec{j} + c\vec{k}$:

$$\overrightarrow{\text{grad}}(f)(p) \cdot \vec{v} = df(p)(\vec{v}) = \begin{pmatrix} \frac{\partial f}{\partial x}(p) & \frac{\partial f}{\partial y}(p) & \frac{\partial f}{\partial z}(p) \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \frac{\partial f}{\partial x}(p)a + \frac{\partial f}{\partial y}(p)b + \frac{\partial f}{\partial z}(p)c.$$

En utilisant que

$$\overrightarrow{\text{grad}}(f)(p) = \left(\overrightarrow{\text{grad}}(f)(p) \cdot \vec{i} \right) \vec{i} + \left(\overrightarrow{\text{grad}}(f)(p) \cdot \vec{j} \right) \vec{j} + \left(\overrightarrow{\text{grad}}(f)(p) \cdot \vec{k} \right) \vec{k},$$

1. Il s'agit en fait du champ d'Euler multiplié par 2.

on retrouve que les composantes du gradient en coordonnées cartésiennes sont les dérivées partielles. Pour calculer les composantes du gradient en coordonnées cylindriques et sphériques, on introduit d'abord le vecteur déplacement infinitésimal :

$$\begin{aligned}\vec{dr} &= dx \vec{i} + dy \vec{j} + dz \vec{k} && \text{(en coordonnées cartésiennes)} \\ &= dr \vec{e}_r + r d\theta \vec{e}_\theta + dz \vec{e}_z && \text{(en coordonnées cylindriques)} \\ &= dr \vec{e}_r + r d\theta \vec{e}_\theta + r \sin(\theta) d\varphi \vec{e}_\varphi && \text{(en coordonnées sphériques)}\end{aligned}$$

de sorte qu'en tout point p , on a la relation

$$df(p) = \left(\overrightarrow{\text{grad}}(f)(p) \right) \cdot \vec{dr}. \quad (3.1)$$

En coordonnées cylindriques, l'équation (3.1) s'écrit :

$$\begin{aligned}\frac{\partial f}{\partial r}(p) dr + \frac{\partial f}{\partial \theta}(p) d\theta + \frac{\partial f}{\partial z}(p) dz \\ = \left(\left(\overrightarrow{\text{grad}}(f)(p) \right) \cdot \vec{e}_r \right) dr + \left(r \left(\overrightarrow{\text{grad}}(f)(p) \right) \cdot \vec{e}_\theta \right) d\theta + \left(\left(\overrightarrow{\text{grad}}(f)(p) \right) \cdot \vec{e}_z \right) dz.\end{aligned}$$

En identifiant les termes, on trouve donc l'expression du gradient en coordonnées cylindriques :

$$\overrightarrow{\text{grad}}(f)(r, \theta, z) = \frac{\partial f}{\partial r}(r, \theta, z) \vec{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta}(r, \theta, z) \vec{e}_\theta + \frac{\partial f}{\partial z}(r, \theta, z) \vec{e}_z.$$

Maintenant, on écrit (3.1) et coordonnées sphériques :

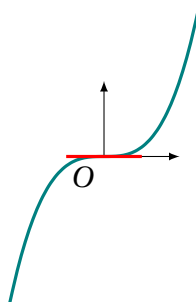
$$\begin{aligned}\frac{\partial f}{\partial r}(p) dr + \frac{\partial f}{\partial \theta}(p) d\theta + \frac{\partial f}{\partial \varphi}(p) d\varphi \\ = \left(\left(\overrightarrow{\text{grad}}(f)(p) \right) \cdot \vec{e}_r \right) dr + \left(r \left(\overrightarrow{\text{grad}}(f)(p) \right) \cdot \vec{e}_\theta \right) d\theta + \left(r \sin(\theta) \left(\overrightarrow{\text{grad}}(f)(p) \right) \cdot \vec{e}_\varphi \right) d\varphi.\end{aligned}$$

En identifiant les termes, on trouve donc l'expression du gradient en coordonnées sphériques :

$$\overrightarrow{\text{grad}}(f)(r, \theta, \varphi) = \frac{\partial f}{\partial r}(r, \theta, \varphi) \vec{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta}(r, \theta, \varphi) \vec{e}_\theta + \frac{1}{r \sin(\theta)} \frac{\partial f}{\partial \varphi}(r, \theta, \varphi) \vec{e}_\varphi.$$

Points critiques. D'après les propriétés du gradient, celui-ci s'annule en les maxima locaux de la fonction, car on ces points on se trouve au sommet d'un ensemble de niveau, dont le gradient ne peut pas s'échapper. En raisonnant de même avec l'opposée de la fonction, on déduit que le gradient s'annule également en ses minima locaux.

Un point qui annule le gradient de f , s'appelle un *point critique* de f . D'après paragraphe précédent, tout extremum local de f est un point critique de f . Il est a noter que la réciproque est fausse, e.g., la fonction d'une variable définie par $f(x) = x^3$ a une dérivée, et donc un gradient, qui s'annule en $x = 0$, alors que ce point n'est pas un extremum local :

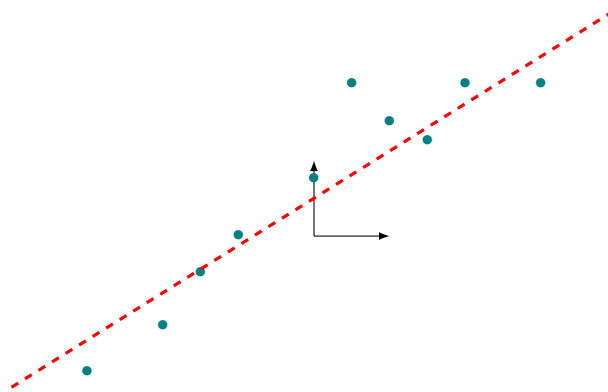


3.2. Hessienne et extrema locaux

En général, la recherche de points critiques, qui se ramène à la résolution d'une équation vectorielle, n'est pas suffisante pour déterminer les extrema d'une fonction, puisqu'on a vu qu'il existe des points critiques qui ne sont pas des extrema. Il faut donc des outils plus précis pour permettre de conclure lorsque ces points critiques sont des extrema, et si oui, de quelle nature. On a pour cela besoin de la Hessienne.

Méthode des moindres carrés. Avant d'introduire la Hessienne, on mentionne qu'il y a un cas particulier où la recherche d'extrema se ramène à celle de points critiques : lorsqu'on sait *a priori* que la fonction admet un extremum et que de plus, elle admet un unique point critique. On illustre ce cas particulier par la *méthode des moindres carrés*.

On se donne une population ainsi que deux caractéristiques de ses individus x et y , et on cherche à savoir comment la caractéristique x influence la caractéristique y . Par exemple, si on souhaite savoir comment l'âge influence la tension artérielle d'un individu, alors x désigne l'âge des individus et y la tension artérielle. Pour avoir une idée de l'influence de x sur y , on peut choisir plusieurs individus et pour chacun d'eux mesurer leur attribut x , supposé connu de façon exacte, et mesurer l'attribut y , cette mesure pouvant être bruitée. De cette façon, on obtient un nuage de points $(x_1, y_1), \dots, (x_n, y_n)$:



Parfois, il se peut que le tirage s'apparente à une droite (telle que celle en rouge sur le dessin) d'équation $y = ax + b$. Les points ne sont pas nécessairement sur la droite car les mesures peuvent être bruitées, et le but est de trouver les coefficients a et b définissant la droite qui approche le plus fidèlement les mesures observées. Pour cela on introduit la fonction de deux variables (a, b) définie par

$$f(a, b) = \sum_{i=1}^n (ax_i + b - y_i)^2,$$

et mesurant la distance entre les valeurs observées et la relation théorique $y = ax + b$. Le couple (a, b) correspondant à la droite qui approxime le mieux les données est celle correspondant à une valeur minimale de f . On cherche donc les points critiques de f en résolvant l'équation $\overrightarrow{\text{grad}}(f)(a, b) = \overrightarrow{0}$:

$$\overrightarrow{\text{grad}}(f)(a, b) = \begin{pmatrix} \frac{\partial f}{\partial a}(a, b) \\ \frac{\partial f}{\partial b}(a, b) \end{pmatrix} = \begin{pmatrix} 2x_1(ax_1 + b - y_1) + \dots + 2x_n(ax_n + b - y_n) \\ 2(ax_1 + b - y_1) + \dots + 2(ax_n + b - y_n) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

De ces deux équations, on obtient

$$a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i \quad \text{et} \quad a \sum_{i=1}^n x_i + nb = \sum_{i=1}^n y_i.$$

En multipliant la première équation par n , la deuxième par $(x_1 + \dots + x_n)$ et en faisant la différence de ces deux expressions, on obtient :

$$a = \frac{n(x_1 y_1 + \dots + x_n y_n) - (x_1 + \dots + x_n)(y_1 + \dots + y_n)}{n(x_1^2 + \dots + x_n^2) - (x_1 + \dots + x_n)^2}.$$

En multipliant la première équation par $(x_1 + \dots + x_n)$, la deuxième par $x_1^2 + \dots + x_n^2$ et en faisant la différence, on trouve

$$b = \frac{(x_1 + \dots + x_n)(x_1 y_1 + \dots + x_n y_n) - (x_1^2 + \dots + x_n^2)(y_1 + \dots + y_n)}{(x_1 + \dots + x_n)^2 - (x_1^2 + \dots + x_n^2)n}.$$

La fonction f admet donc un unique point critique, qui est nécessairement un minimum si y est effectivement une fonction affine de x . De plus, on remarque que les équations sur a et b ne font intervenir que les x_i et les y_i , si bien que a et b peuvent être calculés à partir de l'échantillon utilisé pour faire cette expérience.

Matrice Hessienne. Le critère de classification des points critiques utilisant la Hessienne est basé sur le *développement de Taylor* à l'ordre 2. Au voisinage de tout point p , la fonction f peut être approximée de la façon suivante :

$$f(p + \vec{dr}) \simeq f(p) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p) dx_i + \frac{1}{2} \sum_{1 \leq i, j \leq n} \frac{\partial^2 f}{\partial x_i \partial x_j}(p) dx_i dx_j,$$

où \vec{dr} désigne le vecteur déplacement infinitésimal :

$$\vec{dr} = \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix}.$$

L'approximation de f par son développement de Taylor à l'ordre 2 se réécrit :

$$f(p + \vec{dr}) \simeq f(p) + \left(\overrightarrow{\text{grad}}(f)(p) \right) \cdot \vec{dr} + \vec{dr}^T H(f)(p) \vec{dr},$$

où $H(f)$ est la *matrice Hessienne* de f , définie en tout point p par

$$H(f)(p) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(p) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(p) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(p) \end{pmatrix}.$$

Ainsi, en un point critique p de f , le signe de $f(p + \vec{dr}) - f(p)$ est déterminé par le signe de $\vec{dr}^T H(f)(p) \vec{dr}$. En particulier, si cette quantité est toujours négative ou toujours positive, alors p est un maximum local ou un minimum local, respectivement. Dans le prochain paragraphe, on présente un critère permettant de connaître ce signe dans le cas où $n = 2$.

Classification des points critiques en 2 variables. On suppose que f dépend de 2 variables x, y , et on se place en un point critique p de f . On note le déterminant de la Hessienne par 2 :

$$K(f)(p) = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(p) & \frac{\partial^2 f}{\partial x \partial y}(p) \\ \frac{\partial^2 f}{\partial x \partial y}(p) & \frac{\partial^2 f}{\partial y^2}(p) \end{vmatrix} = \frac{\partial^2 f}{\partial x^2}(p) \frac{\partial^2 f}{\partial y^2}(p) - \left(\frac{\partial^2 f}{\partial x \partial y}(p) \right)^2.$$

On a alors :

2. Cette notation est issue de la géométrie, K désignant la *courbure Gaussienne* de la surface définie par le graphe de f .

- si $K(f)(p) < 0$, alors f n'admet pas d'extremum en p ,
- si $K(f)(p) > 0$ et $\frac{\partial^2 f}{\partial x^2}(p) > 0$, alors f admet un minimum local en p ,
- si $K(f)(p) > 0$ et $\frac{\partial^2 f}{\partial x^2}(p) < 0$, alors f admet un maximum local en p ,
- si $K(f)(p) = 0$, alors on ne peut pas conclure.

Dans le premier cas, le point p est appelé *point selle* de f . On emploie cette terminologie car l'exemple typique de point selle est, comme on va le voir, l'origine pour la fonction définie par $f(x, y) = y^2 - x^2$. Or, on a tracé le graphe de cette fonction dans la section 2.1 du chapitre 2, et on peut observer que celui-ci s'apparente en effet à une selle de cheval.

Voici résumée la méthode de classification des points critiques d'une fonction f de 2 variables x, y :

- on cherche les points critiques de f en résolvant le système $\overrightarrow{\text{grad}}(f)(x, y) = \overrightarrow{0}$,
- on calcule la forme générale de $H(f)(x, y)$,
- pour chaque point critique, on calcule $K(f)(x, y)$ à partir de la forme générale de $H(f)(x, y)$,
- on classe p en fonction du critère présenté plus haut.

Exemple de la selle de cheval. Comme annoncé plus haut, on illustre la méthode présentée par la fonction définie par $f(x, y) = y^2 - x^2$. On calcule le gradient de f :

$$\overrightarrow{\text{grad}}(f)(x, y) = \begin{pmatrix} -2x \\ 2y \end{pmatrix}.$$

L'équation $\overrightarrow{\text{grad}}(f)(x, y) = \overrightarrow{0}$ admet une unique solution $p(0, 0)$. On calcule la Hessienne de f en tout (x, y) :

$$H(f)(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Ainsi, $K(f)$ est la fonction constante égale à -4 . Le point p est donc un point selle de f .

Exemple d'une fonction admettant plusieurs points critiques. On considère la fonction de deux variables définie par $f(x, y) = x^3 + y^2 - 4xy + 5x$. On commence par calculer son gradient :

$$\overrightarrow{\text{grad}}(f)(x, y) = \begin{pmatrix} 3x^2 - 4y + 5 \\ 2y - 4x \end{pmatrix}.$$

Le calcul des points critiques se ramène donc à la résolution du système

$$3x^2 - 4y + 5 = 0 \quad \text{et} \quad y - 2x = 0.$$

En substituant $y = 2x$ dans la première équation, on obtient

$$3x^2 - 8x + 5 = 3(x - 1)(x - 5/3) = 0.$$

On en déduit les deux points critiques : $p_1(1, 2)$ et $p_2(5/3, 10/3)$. On calcule maintenant la forme générale de la Hessienne de f :

$$H(f)(x, y) = \begin{pmatrix} 6x & -4 \\ -4 & 2 \end{pmatrix}.$$

On a $K(f)(x, y) = 12x - 16$. Ainsi, $K(f)(p_2) < 0$ de sorte que p_2 est un point selle de f . De plus, $K(f)(p_1) > 0$ et comme $\frac{\partial^2 f}{\partial x^2}(p_1) > 0$, on en déduit que p_1 est un minimum local.

3.3. Multiplicateurs de Lagrange

On a vu dans le chapitre précédent comment optimiser une fonction en recherchant ses extrema. La question est maintenant d'optimiser une fonction mais en respectant certaines contraintes. Les outils qu'on utilise pour cela sont les *multiplicateurs de Lagrange*.

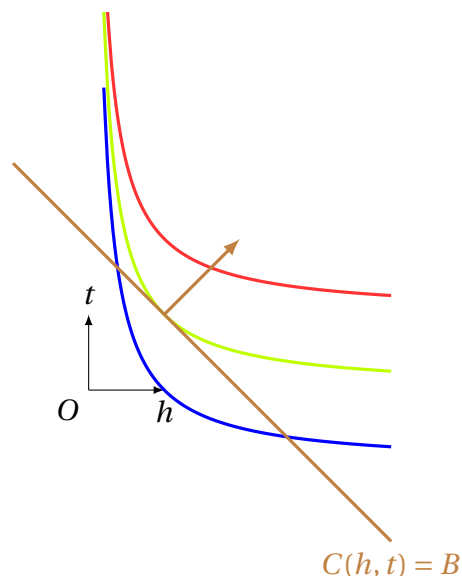
Exemple de motivation. On considère un petit exemple³ illustrant l'utilisation des multiplicateurs de Lagrange. Une usine produit des objets en acier, et son revenu dépend de deux paramètres principaux :

- le nombre total d'heures travaillées h par les employés,
- le nombre total de tonnes d'acier achetées t par l'usine pour produire les objets.

On suppose que le revenu de l'usine est une fonction R des variables h, t . De plus, les contraintes budgétaires de l'usine sont les suivantes :

- une heure de travail revient à $H\text{€}$,
- une tonne d'acier coûte $T\text{€}$,
- le budget total de l'usine est de $B\text{€}$.

Les directeurs de l'usine souhaitant maximiser leur gain, ils souhaitent choisir de façon appropriée le nombre de tonnes d'acier à acheter et le nombre d'heures de travail des employés tout en respectant le budget. Il s'agit donc de maximiser la fonction R sous la contrainte que $H.h + T.t = B$. Dit autrement, en définissant la fonction de contrainte par $C(h, t) = H.h + T.t$, la contrainte se formule par $C(h, t) = B$. Géométriquement, on trace la courbe de niveaux B de la fonction de contrainte C dans le plan h, t et on représente la fonction revenu R par ses courbes de niveaux :



Il est à noter que plus les courbes de niveaux de R sont en haut à droite, plus elles correspondent à un revenu élevé, car elles correspondent à un nombre d'heures et de tonnes d'acier achetés élevés. C'est pour cette raison que la courbe haute est représentée en rouge et la basse en bleu. En particulier, la courbe verte, tangente à la droite représentant la contrainte, est celle correspondant au revenu le plus élevé qu'on peut atteindre en respectant la contrainte. La condition que cette courbe est tangente à la courbe de contrainte signifie que les gradients de la fonction de revenu R et de la contrainte C sont colinéaires, pointant dans une direction représentée par vecteur marron. En particulier, il existe un scalaire λ , appelé *multiplicateur de Lagrange*, tel que

$$\overrightarrow{\text{grad}}(R)(p) = \lambda \times \overrightarrow{\text{grad}}(C)(p).$$

3. Cet exemple est repris de la chaîne youtube *Khan academy*.

La méthode des multiplicateurs de Lagrange consiste à trouver les coordonnées (h, t) de p en résolvant le système de 3 inconnues h, t, λ et 3 équations, dont 2 viennent de la colinéarité des gradients et la troisième provient de la contrainte.

Multiplicateur de Lagrange pour une contrainte. Comme au début du chapitre, on considère une fonction f des variables x_1, \dots, x_n , et on résume la méthode des multiplicateurs de Lagrange pour optimiser f sous une contrainte :

- on formule la contrainte par l'équation $C(x_1, \dots, x_n) = c$, où c est une constante explicite,
- on introduit une nouvelle variable λ , appelée *multiplicateur de Lagrange* et on considère le système des $n + 1$ inconnues x_1, \dots, x_n, λ , et $n + 1$ équations

$$\overrightarrow{\text{grad}}(f)(x_1, \dots, x_n) = \lambda \times \overrightarrow{\text{grad}}(C)(x_1, \dots, x_n) \quad \text{et} \quad C(x_1, \dots, x_n) = c,$$

- en résolvant le système précédent, on trouve une solution $(x_1, \dots, x_n, \lambda)$,
- la valeur optimisant f sous la contrainte C est le point (x_1, \dots, x_n) correspondant aux n premières coordonnées de la solution calculée au point précédent.

Bien sûr, cette méthode se généralise aux cas de plusieurs contraintes en considérant plusieurs multiplicateurs de Lagrange. Cependant, les seuls exemples abordés dans ce cours concernent le cas d'une seule contrainte.

Exemple numérique. On reprend l'exemple précédent avec

$$R(h, t) = 100h^{\frac{2}{3}}t^{\frac{1}{3}}, \quad H = 20, \quad T = 2000 \quad \text{et} \quad B = 20000.$$

On commence par définir la fonction de contrainte par $C(h, t) = 20h + 2000t$, si bien que la contrainte se traduit par $C(h, t) = 20000$. On écrit maintenant la condition de colinéarité $\overrightarrow{\text{grad}}(f)(h, t) = \lambda \overrightarrow{\text{grad}}(C)(h, t)$:

$$\begin{pmatrix} \frac{\partial R}{\partial h}(h, t) \\ \frac{\partial R}{\partial t}(h, t) \end{pmatrix} = \lambda \begin{pmatrix} \frac{\partial C}{\partial h}(h, t) \\ \frac{\partial C}{\partial t}(h, t) \end{pmatrix} \Leftrightarrow \begin{pmatrix} \frac{200}{3}h^{-\frac{1}{3}}t^{\frac{1}{3}} \\ \frac{100}{3}h^{\frac{2}{3}}t^{-\frac{2}{3}} \end{pmatrix} = \lambda \begin{pmatrix} 20 \\ 2000 \end{pmatrix}$$

On résout donc le systèmes d'inconnues h, t, λ et aux 3 équations :

$$\begin{cases} \frac{200}{3}h^{-\frac{1}{3}}t^{\frac{1}{3}} = 20\lambda \\ \frac{100}{3}h^{\frac{2}{3}}t^{-\frac{2}{3}} = 2000\lambda \\ 20h + 2000t = 20000 \end{cases} \Leftrightarrow \begin{cases} 10\left(\frac{t}{h}\right)^{\frac{1}{3}} = 3\lambda \\ \left(\frac{h}{t}\right)^{\frac{2}{3}} = 60\lambda \\ h + 100t = 1000 \end{cases}$$

En multipliant le carré de la première équation par la deuxième équation, on a $10^2 = (3\lambda)^2 \times 60\lambda$, i.e.,

$$\lambda^3 = \frac{100}{9 \times 60} \Leftrightarrow \lambda = \left(\frac{5}{27}\right)^{\frac{1}{3}}.$$

En reportant cette quantité dans la première équation puis en élevant le tout au cube, on a

$$10^3 \frac{t}{h} = \frac{3^3 \times 5}{27} \Leftrightarrow t = \frac{h}{200}.$$

En reportant cette expression dans la dernière équation on trouve

$$h + \frac{h}{2} = 1000 \Leftrightarrow h = \frac{2000}{3}.$$

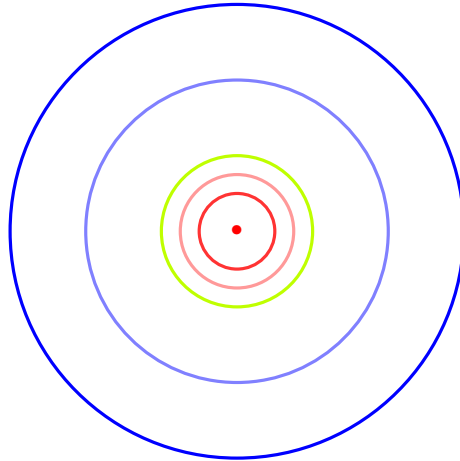
Donc finalement, on trouve

$$t = \frac{10}{3}.$$

En déduit que l'entreprise doit acheter $10/3$ tonnes et faire travailler ses employé pendant $2000/3$ heures pour maximiser son gain à partir d'un budget initial de 20000€ .

3.4. Fiche d'exercices 3

Exercice 1. On considère une fonction de 2 variables dont on représente certaines courbes de niveaux :



1. Dessiner approximativement le gradient en certains points. Quel est sa valeur en le point au centre?
2. Parmi les fonctions de l'exercice 2 de la deuxième fiche d'exercices, laquelle est la plus susceptible d'avoir les courbes de niveaux représentées dans cet exercice?
3. Calculer le gradient de la fonction trouvée à la question précédente. Le résultat obtenu est-il cohérent avec les dessins de la question 1).?

Exercice 2. On considère les fonctions de \mathbb{R}^2 dans \mathbb{R} définies par pour tout $(x, y) \in \mathbb{R}^2$:

$$f(x, y) = (x^2 + y^2)^2 - 2x^2 + 2y^2, \quad g(x, y) = x^3 + y^3 - 3xy, \quad h(x, y) = (x^2 + y^2)^2 - 8xy.$$

Pour chacune de ces fonctions, déterminer les points critiques ainsi que leur nature (extrema, points-selle).

Exercice 3. Trouver les valeurs maximal et minimal de la fonction définie par $f(x, y, z) = xyz$ sur le plan d'équation $x + y + z = 1$.

Exercice 4. On désire fabriquer une boîte ayant la forme d'un parallélépipède rectangle, sans couvercle sur le dessus. Le volume de cette boîte doit être égal à $0,5 \text{ m}^3$ et pour optimiser la quantité de matière utilisée, on désire que la somme des aires des faces soit aussi petite que possible.

1. En utilisant la contrainte sur le volume de la boîte, modéliser la quantité à minimiser par une fonction de deux variables. En déduire les dimensions de la boîte minimisant sa surface.
2. Retrouver le résultat de la question précédente grâce à la méthode des multiplicateurs de Lagrange.

Exercice 5. Le but de cet exercice est de comprendre la signification de la valeur λ qu'on obtient quand on résout les équations des multiplicateurs de Lagrange pour les fonctions sous contrainte. On considère une fonction de 2 variables x, y , ainsi qu'une contrainte $C(x, y) = c$. On note (x^*, y^*, λ^*) la solution du système défini par la colinéarité des gradients de f et C et par la contrainte C elle-même, et on note $f^* = f(x^*, y^*)$ la valeur optimale de f . Enfin, on introduit la fonction \mathcal{L} des variables x, y, λ , définie par

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda(C(x, y) - c).$$

1. Exprimer grâce à \mathcal{L} les équations satisfaites par (x^*, y^*, λ^*) .
2. Quelle est la valeur de $\mathcal{L}(x^*, y^*, \lambda^*)$?
3. On suppose maintenant que c est une variable, si bien que \mathcal{L} dépend des 4 variables x, y, λ, c et que les quantités f^*, x^*, y^*, λ^* sont des fonctions de c .
 - a). Exprimer les relations des questions 1) et 2) dans ce nouveau contexte.
 - b). Montrer que la dérivée totale de \mathcal{L} est égale à λ^* . En déduire la signification de $\lambda^*(c)$.
 - c). On suppose que f représente une fonction revenue et que C représente une contrainte budgétaire. Si pour une valeur c donnée, on trouve $\lambda^*(c) = (5/27)^{1/3}$, est-ce qu'on a intérêt à augmenter le budget initial pour maximiser ses gains?

Chapitre 4 :

Opérateurs différentiels et intégration

Ce chapitre porte sur les opérateurs différentiels, qui sont des outils permettant d'analyser les propriétés géométriques des champs scalaires ou vectoriels. On introduit également la notion d'intégrale curviligne, de surface et de volume. Enfin, on rappelle la formule de Stokes et ses généralisations, qui font le lien entre les opérateurs différentiels et l'intégration.

4.1. Opérateurs différentiels classiques

Les opérateurs différentiels sont des outils permettant d'analyser les propriétés géométriques des champs scalaires ou vectoriels, par exemple. On a déjà croisé un exemple d'opérateur : le gradient, qui permet de localiser les directions dans lesquelles un champ scalaire varie le plus rapidement. Dans cette section on introduit trois autres opérateurs classiques : le *rotationnel*, la *divergence* et le *Laplacien*. Pour définir ces opérateurs, on a besoin de l'opérateur nabla, dont on rappelle la définition en coordonnées cartésiennes :

$$\vec{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}.$$

On a insisté depuis le début de ce cours sur l'intérêt de modèles géométriques en physique. Dès lors, il n'est pas étonnant que les opérateurs différentiels jouent un rôle fondamental en physique : la divergence et le rotationnel permettent de formuler les versions locales des équations de Maxwell, quand au gradient, il permet d'exprimer des potentiels (électrique ou gravitationnel).

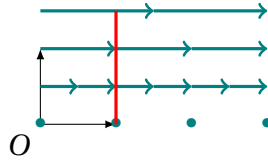
Rotationnel d'un champ de vecteurs. Le rotationnel d'un champ de vecteur permet d'indiquer une tendance qu'a ce champ à faire tourner une particule se trouvant à l'intérieur de celui-ci. Il s'agit d'une quantité vectorielle, dont la direction indique le sens dans lequel tourne la particule en respectant la règle de la main droite : vers le haut si la particule tourne dans le sens trigonométrique et vers le bas sinon. La norme du rotationnel indique quand à elle la vitesse de la rotation : plus la norme est élevée, plus la particule tourne vite. Étant donnée un champ de vecteurs \vec{v} défini en **coordonnées cartésiennes** par

$$\vec{v}(x, y, z) = v_x(x, y, z) \vec{i} + v_y(x, y, z) \vec{j} + v_z(x, y, z) \vec{k},$$

le rotationnel de \vec{v} est le champ de vecteurs

$$\vec{\text{rot}}(\vec{v})(x, y, z) = \vec{\nabla} \times \vec{v} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \times \begin{pmatrix} v_x(x, y, z) \\ v_y(x, y, z) \\ v_z(x, y, z) \end{pmatrix} = \begin{pmatrix} \frac{\partial v_z}{\partial y}(x, y, z) - \frac{\partial v_y}{\partial z}(x, y, z) \\ \frac{\partial v_x}{\partial z}(x, y, z) - \frac{\partial v_z}{\partial x}(x, y, z) \\ \frac{\partial v_y}{\partial x}(x, y, z) - \frac{\partial v_x}{\partial y}(x, y, z) \end{pmatrix}.$$

On considère par exemple le champ de vecteurs défini par $\vec{v}(x, y, z) = y \vec{i}$ et une particule représentée en rouge dans ce champ :



Alors, le champ étant plus fort au sommet qu'à la base de la particule, celle-ci tournera dans le sens opposé du sens trigonométrique, et donc la direction de rotation sera $-\vec{k}$. Or, en calculant $\text{rot}(\vec{v})$, on obtient

$$\text{rot}(\vec{v})(x, y, z) = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \times \begin{pmatrix} y \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} = -\vec{k}.$$

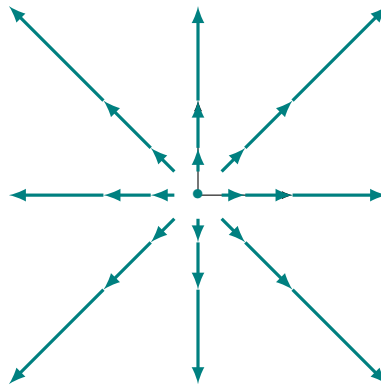
Divergence d'un champ de vecteurs. Un champ de vecteurs peut être pensé comme étant la vitesse d'un fluide. La divergence d'un champ de vecteurs est un champ scalaire permettant d'indiquer en chaque point du fluide si la matière s'échappe, ou diverge, de ce point (dans ce cas, la divergence est positive) ou au contraire si la matière converge en ce point (dans ce cas la divergence est négative). Étant donné un champ de vecteurs \vec{v} défini en **coordonnées cartésiennes** par

$$\vec{v}(x, y, z) = v_x(x, y, z) \vec{i} + v_y(x, y, z) \vec{j} + v_z(x, y, z) \vec{k},$$

la divergence de \vec{v} est définie par

$$\text{div}(\vec{v})(x, y, z) = \vec{\nabla} \cdot \vec{v} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \cdot \begin{pmatrix} v_x(x, y, z) \\ v_y(x, y, z) \\ v_z(x, y, z) \end{pmatrix} = \frac{\partial v_x}{\partial x}(x, y, z) + \frac{\partial v_y}{\partial y}(x, y, z) + \frac{\partial v_z}{\partial z}(x, y, z).$$

On considère par exemple le champ d'Euler défini par $\vec{v}(x, y, z) = x \vec{i} + y \vec{j} + z \vec{k}$:



En tout point, la norme des lignes de champs entrantes est inférieure à celle des lignes de champs sortantes, si bien qu'une plus grande quantité de fluide a tendance à sortir d'un point qu'à y entrer. On s'attend donc à ce que la divergence de \vec{v} soit positive, ce qu'on peut vérifier puisque $\text{div}(\vec{v})(x, y, z) = 3$.

Laplacien d'un champ scalaire. Le Laplacien d'un champ scalaire est une quantité scalaire permettant de repérer des creux ou des bosses dans le graphe de la fonction f . Par définition, le Laplacien d'un champ scalaire f est la divergence de son gradient :

$$\Delta f(x, y, z) = \text{div}(\overrightarrow{\text{grad}}(f))(x, y, z) = \frac{\partial^2 f}{\partial x^2}(x, y, z) + \frac{\partial^2 f}{\partial y^2}(x, y, z) + \frac{\partial^2 f}{\partial z^2}(x, y, z).$$

Il s'agit donc d'une quantité indiquant si le gradient de f converge ou diverge d'un point, le gradient pointant dans la direction où la fonction f croit le plus rapidement. Ainsi, dans un creux du graphe de f , le gradient diverge et donc le Laplacien est positif, alors qu'au sommet d'une bosse, le gradient converge et donc le Laplacien est négatif.

Équations de Maxwell. Les opérateurs différentiels permettent d'exprimer de nombreux phénomènes physiques. L'exemple le plus classique est donné par les *équations de Maxwell*.

- La *loi de Gauss* stipule qu'une densité de charge ρ est la source d'un champ électrique divergeant :

$$\operatorname{div}(\vec{E}) = \frac{\rho}{\epsilon_0}.$$

- La *loi de Faraday* stipule que la variation d'un champ magnétique cause une rotation du champ électrique :

$$\operatorname{rot}(\vec{E}) = -\frac{\partial \vec{B}}{\partial t}.$$

- la *loi de Gauss magnétique* stipule qu'il n'existe pas de monopole magnétique :

$$\operatorname{div}(\vec{B}) = 0.$$

- La *loi d'Ampère-Maxwell* stipule qu'un courant ou la variation temporelle du champ électrique sont la source d'un champ magnétique :

$$\operatorname{rot}(\vec{B}) = \mu_0 \vec{j} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}.$$

Potentiels scalaires Certains champs de vecteurs peuvent dériver d'un *potentiel*, celui-ci pouvant être scalaire ou vectoriel. Étant donné un champ de vecteurs \vec{v} , un potentiel scalaire (resp., vectoriel) de \vec{v} est un champ scalaire f (resp., vectoriel \vec{a}) tel que $\vec{v} = \overrightarrow{\operatorname{grad}}(f)$ (resp., $\vec{v} = \overrightarrow{\operatorname{rot}}(\vec{a})$). Dans les deux cas, on dit que \vec{v} *dérive d'un potentiel*. Comme on le verra en TD, tous les champs de vecteurs ne dérivent pas d'un potentiel.

Deux exemples de potentiels scalaires sont les potentiels gravitationnel V_G et électrique V_E , dont dérivent respectivement les champs gravitationnel \vec{g} et électrique \vec{E} : $\vec{g} = -\overrightarrow{\operatorname{grad}}(V_G)$ et $\vec{E} = -\overrightarrow{\operatorname{grad}}(V_E)$. La terminologie "potentiel" provient du fait que V_G et V_E sont des multiples de l'énergie potentielle : V_G est l'énergie potentielle gravitationnelle par unité de masse et V_E est l'énergie potentielle électrique par charge. Comme on le verra dans la section 4.3, lorsque les champs gravitationnel ou électrique dérivent d'un potentiel, il y a conservation de l'énergie. Pour cette raison, on dit qu'un champ de vecteurs dérivant d'un potentiel est conservatif.

4.2. Intégrales curviligne, de surface et volumique

Un champ scalaire ou vectoriel peut être intégré le long d'une courbe, d'une surface ou d'un volume ; on parle respectivement d'intégrale *curviligne*, *de surface* ou *volumique*.

Intégrale curviligne d'un champ scalaire. On commence par définir l'intégrale curviligne d'un champ scalaire f de \mathbb{R}^2 le long d'une courbe \mathcal{C} . Par définition, il s'agit de la somme des valeurs prises par f le long de \mathcal{C} , ce qui revient à prendre l'aire de la surface comprise entre \mathcal{C} et le graphe de f . On note cette intégrale :

$$I(f) = \int_{\mathcal{C}} f(\vec{r}) ds,$$

où ds désigne une longueur infinitésimale le long de \mathcal{C} . En pratique, pour calculer $I(f)$, on fixe une paramétrisation $\vec{r}(t) = x(t)\vec{i} + y(t)\vec{j}$, avec $t_i \leq t \leq t_f$, de \mathcal{C} . L'élément de longueur infinitésimal ds est

égal à $\left| \dot{\vec{r}}(t) \right| dt = \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt$, et $I(f)$ se calcule par une intégrale au sens usuel :

$$I(f) = \int_{t_i}^{t_f} f(\vec{r}(t)) \left| \dot{\vec{r}}(t) \right| dt = \int_{t_i}^{t_f} f(x(t), y(t)) \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt.$$

Bien sûr, cette définition s'étend pour une courbe dans \mathbb{R}^3 .

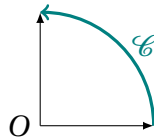
Lorsque f est le champ scalaire constant égal à 1, l'intégrale $I(1)$ est la longueur de \mathcal{C} . Par exemple, soit le cercle $\mathcal{C}(0, R)$ de centre O et de rayon R , paramétré en coordonnées polaires

$$\vec{r}(\theta) = R \cos(\theta) \vec{i} + R \sin(\theta) \vec{j},$$

avec $0 \leq \theta \leq 2\pi$. Alors la longueur de $\mathcal{C}(O, R)$ est égale à

$$\int_0^{2\pi} \sqrt{R^2 \sin^2(\theta) + R^2 \cos^2(\theta)} d\theta = \int_0^{2\pi} R d\theta = 2\pi R.$$

On souhaite maintenant calculer l'intégrale du champ scalaire défini par $f(x, y) = xy$ le long de l'arc de cercle \mathcal{C} joignant les points $(1, 0)$ et $(0, 1)$:



La courbe \mathcal{C} est paramétrée par $\vec{r}(\theta) = \cos(\theta) \vec{i} + \sin(\theta) \vec{j}$, avec $0 \leq \theta \leq \pi/2$. L'élément de longueur de cette paramétrisation est $\sqrt{(-\sin(\theta))^2 + \cos^2(\theta)} d\theta = d\theta$, de sorte qu'on a

$$\int_{\mathcal{C}} f(\vec{r}) ds = \int_0^{\pi/2} \cos(\theta) \sin(\theta) d\theta = \left[\frac{\sin^2(\theta)}{2} \right]_0^{\pi/2} = \frac{1}{2}.$$

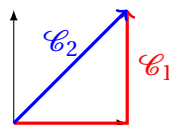
Intégrale curviligne d'un champ de vecteurs. On définit l'intégrale curviligne d'un champ de vecteurs \vec{v} sur \mathbb{R}^3 le long d'une courbe \mathcal{C} :

$$I(\vec{v}) = \int_{\mathcal{C}} \vec{v} \cdot d\vec{r} = \int_{\mathcal{C}} \vec{v} \cdot \vec{t} ds,$$

où \vec{t} est un vecteur tangent unitaire à \mathcal{C} . En physique, l'intégrale $I(\vec{v})$ s'appelle la *circulation* de \vec{v} le long de \mathcal{C} . En pratique, pour calculer $I(\vec{v})$, on se donne une paramétrisation $\vec{r}(t)$, avec $t_1 \leq t \leq t_f$, de \mathcal{C} , et on calcule l'intégrale usuelle :

$$I(\vec{v}) = \int_{\mathcal{C}} \vec{v}(\vec{r}(t)) \cdot d\vec{r} = \int_{t_0}^{t_f} \vec{v}(\vec{r}(t)) \cdot \frac{d\vec{r}}{dt} dt.$$

Par exemple, on cherche à calculer la circulation du champ d'Euler défini par $\vec{v} = x \vec{i} + y \vec{j} + z \vec{k}$ le long des courbes planes



On paramétrise les courbes \mathcal{C}_1 et \mathcal{C}_2 par

$$\vec{r}_1(t) = \begin{cases} t \vec{i}, & \text{si } 0 \leq t \leq 1 \\ (t-1) \vec{j}, & \text{si } 1 \leq t \leq 2 \end{cases} \quad \text{et} \quad \vec{r}_2(t) = t \vec{i} + t \vec{j}, \quad 0 \leq t \leq 1.$$

On a donc

$$\begin{aligned} \int_{\mathcal{C}_1} \vec{v} \cdot d\vec{r}_1 &= \int_0^1 t \vec{i} \cdot \frac{d\vec{r}_1}{dt} dt + \int_1^2 (t-1) \vec{j} \cdot \frac{d\vec{r}_1}{dt} dt = \int_0^1 t \vec{i} \cdot \vec{i} dt + \int_1^2 (t-1) \vec{j} \cdot \vec{j} dt \\ &= \left[\frac{t^2}{2} \right]_0^1 + \left[\frac{t^2}{2} - t \right]_1^2 = \frac{1}{2} + \left(2 - 2 - \frac{1}{2} + 1 \right) = 1, \end{aligned}$$

et

$$\int_{\mathcal{C}_2} \vec{v} \cdot d\vec{r}_2 = \int_0^1 (t \vec{i} + t \vec{j}) \cdot \frac{d\vec{r}_2}{dt} dt = \int_0^1 (t \vec{i} + t \vec{j}) \cdot (\vec{i} + \vec{j}) dt = \int_0^1 2t dt = 1.$$

Comme on verra dans la section 4.3, le fait qu'on retrouve la même valeur pour les deux intégrales est dû au fait que \mathcal{C}_1 et \mathcal{C}_2 ont les mêmes extrémités.

Intégrales de surface d'un champ scalaire. L'intégrale d'un champ scalaire f de \mathbb{R}^3 le long d'une surface \mathcal{S} de \mathbb{R}^3 est la somme que prend f le long de \mathcal{S} , ce qui revient à prendre le volume compris entre \mathcal{S} et le graphe de f . On note cette intégrale

$$I(f) = \iint_{\mathcal{S}} f(\vec{r}) dS.$$

où dS désigne un élément de surface infinitésimale de \mathcal{S} . En pratique, pour calculer $I(f)$, on fixe une paramétrisation $\vec{r}(s, t) = x(s, t) \vec{i} + y(s, t) \vec{j} + z(s, t) \vec{k}$, avec $s_i \leq s \leq s_f$ et $t_i \leq t \leq t_f$, de \mathcal{S} . En toute généralité, les valeurs initiales et finales de s et t pourraient dépendre de t ou s , respectivement. Cependant, mise à part pour le calcul de l'aire algébrique décrite par une fonction d'une seule variable que l'on détaille ci-dessous, dans ce cours on ne considère que des valeurs constantes pour les extrémités de s et t . Une fois la paramétrisation de S fixée, son élément de surface infinitésimal dS est égal à

$$dS = \left| \frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t} \right| ds dt,$$

et $I(f)$ se calcule par une intégrale double au sens usuel :

$$I(f) = \int_{s_i}^{s_f} \int_{t_i}^{t_f} f(\vec{r}(s, t)) \left| \frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t} \right| ds dt.$$

Lorsque f est le champ scalaire constant égal à 1, l'intégrale $I(1)$ est l'aire de la surface de \mathcal{S} . Par exemple, soit $\mathcal{S}(O, R)$ la sphère centrée en 0 de rayon R , paramétrée par

$$\vec{r}(\theta, \varphi) = \sin(\theta) \cos(\varphi) \vec{i} + R \sin(\theta) \sin(\varphi) \vec{j} + R \cos(\theta) \vec{k},$$

avec $0 \leq \theta \leq \pi$ et $0 \leq \varphi \leq 2\pi$. On commence par calculer l'élément de surface de cette paramétrisation :

$$\begin{aligned} dS &= \left| \begin{pmatrix} R \cos(\theta) \cos(\varphi) \\ R \cos(\theta) \sin(\varphi) \\ -R \sin(\theta) \end{pmatrix} \times \begin{pmatrix} -R \sin(\theta) \sin(\varphi) \\ R \sin(\theta) \cos(\varphi) \\ 0 \end{pmatrix} \right| d\theta d\varphi \\ &= R^2 \times \left| \begin{pmatrix} \sin^2(\theta) \cos(\varphi) \\ \sin^2(\theta) \sin(\varphi) \\ \cos(\theta) \sin(\theta) \end{pmatrix} \right| d\theta d\varphi = R^2 \sin(\theta) d\theta d\varphi \end{aligned}$$

La dernière égalité provient du fait que la fonction \sin est positive sur $[0, \pi]$. L'aire de $\mathcal{S}(O, R)$ est égale à

$$\int_0^{2\pi} \int_0^\pi R^2 \sin(\theta) d\theta d\varphi = R^2 \times \int_0^{2\pi} [-\cos(\theta)]_0^\pi d\varphi = 2R^2 \times \int_0^{2\pi} d\varphi = 4\pi R^2.$$

Un autre exemple est donné par l'aire définie par une fonction positive d'une seule variable f , définie sur un intervalle $[a, b]$. La surface \mathcal{S} considérée est la portion du plan se trouvant sous la courbe représentative de f , paramétrée par $\vec{r}(x, y) = x\vec{i} + y\vec{j}$, avec $a \leq x \leq b$ et $0 \leq y \leq f(x)$ ¹. L'élément de surface est égal à $dS = |\vec{i} \times \vec{j}| dx dy = |\vec{k}| dx dy = dx dy = dy dx$, de sorte que l'aire de \mathcal{S} est

$$\int_a^b \int_0^{f(x)} dy dx = \int_a^b f(x) dx.$$

On retrouve donc que l'intégrale d'une fonction positive d'une seule variable f est égale à l'aire se trouvant sous sa courbe représentative. Si la fonction f est négative, on fixe la même paramétrisation, mais avec $f(x) \leq y \leq 0$, et on trouve que l'aire au-dessus de la courbe représentative de f est égale à

$$\int_a^b \int_{f(x)}^0 dy dx = - \int_a^b f(x) dx.$$

Si la fonction f change de signe, on fixe une paramétrisation pour la partie positive de f et une autre pour la partie négative de f et on retrouve que l'intégrale de f est son aire algébrique.

On souhaite maintenant calculer l'intégrale du champ scalaire défini par $f(x, y, z) = xyz$ le long de la surface \mathcal{S} d'équation $x^2 + y^2 = 2$, avec $0 \leq x, y \leq 1$ et $0 \leq z \leq 3$. On paramétrise \mathcal{S} en coordonnées cylindriques : $\vec{r}(\theta, z) = 2 \cos(\theta)\vec{i} + 2 \sin(\theta)\vec{j} + z\vec{k}$, avec $0 \leq \theta \leq \pi/2$ et $0 \leq z \leq 3$. On a l'élément de surface

$$dS = \left| \begin{pmatrix} -2 \sin(\theta) \\ 2 \cos(\theta) \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right| d\theta dz = \left| \begin{pmatrix} 2 \cos(\theta) \\ 2 \sin(\theta) \\ 0 \end{pmatrix} \right| d\theta dz = 2 d\theta dz.$$

On a donc

$$\begin{aligned} I(f) &= \int_0^3 \int_0^{\pi/2} 8z \cos(\theta) \sin(\theta) d\theta dz = \int_0^3 8z \left[\frac{\sin^2(\theta)}{2} \right]_0^{\pi/2} dz \\ &= \int_0^3 4z dz = 2 [z^2]_0^3 = 18. \end{aligned}$$

Intégrale de surface d'un champ de vecteurs. On définit l'intégrale d'un champ de vecteurs \vec{v} sur \mathbb{R}^3 le long d'une surface \mathcal{S} :

$$I(\vec{v}) = \int_{\mathcal{S}} \vec{v} \cdot d\vec{S} = \int_{\mathcal{S}} \vec{v} \cdot \vec{n} dS,$$

où \vec{n} est un vecteur normal unitaire à \mathcal{S} , qui l'oriente selon la règle de main droite. En physique, l'intégrale $I(\vec{v})$ s'appelle le *flux* de \vec{v} à travers \mathcal{S} . En pratique, pour calculer $I(\vec{v})$, on se donne une paramétrisation $\vec{r}(s, t)$, avec $s_i \leq s \leq s_f$ et $t_1 \leq t \leq t_f$, de \mathcal{S} , et on calcule l'intégrale usuelle :

$$I(\vec{v}) = \iint_{\mathcal{S}} \vec{v}(\vec{r}(s, t)) \cdot d\vec{S} = \int_{t_1}^{t_f} \int_{s_i}^{s_f} \vec{v}(\vec{r}(s, t)) \cdot \left(\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t} \right) ds dt.$$

On remarque que l'ordre dans lequel on effectue le produit vectoriel oriente la surface \mathcal{S} .

Par exemple, on cherche à calculer le flux du champ d'Euler défini par $\vec{v} = x\vec{i} + y\vec{j} + z\vec{k}$ à travers le cube unité, orienté vers l'extérieur. On paramétrise les 6 faces du cube, par exemple en $x = \pm 1/2$, on a $\vec{r}(s, t) = \pm 1/2\vec{i} + s\vec{j} + t\vec{k}$, avec $-1/2 \leq s, t \leq 1/2$. Sur cette face, le vecteur normal est $\pm \vec{j} \times \vec{k} = \pm \vec{i}$ (le signe \pm provenant de l'orientation du cube). Sur ces faces, $\vec{v} \cdot \vec{n}$ est égal à $1/2$ et donc le flux de \vec{v} à travers cette surface est la moitié de son aire, *i.e.*, $1/2$. En procédant de même pour les faces où $y = \pm 1/2$ et $z = \pm 1/2$, on trouve que le flux total de \vec{v} à travers \mathcal{S} est égal à $6 \times 1/2 = 3$.

1. Il s'agit, comme annoncé, de l'exemple où les bornes de la paramétrisation ne sont pas constantes.

Intégrale volumique d'un champ scalaire. L'intégrale volumique d'un champ scalaire f de \mathbb{R}^3 le long d'un volume \mathcal{V} est la somme des valeurs prises par f le long de \mathcal{V} . On note cette intégrale :

$$I(f) = \iiint_{\mathcal{V}} f(\vec{r}) dV,$$

où dV désigne un volume infinitésimal le long de \mathcal{V} . En pratique, on commence par paramétriser \mathcal{V} par un système de coordonnées, cartésiennes, cylindriques ou sphériques. Ensuite, en coordonnées cartésiennes, l'élément de volume est $dV = dx dy dz$ et les autres éléments de volumes se déduisent par un calcul de déterminant, celui-ci indiquant par quel valeur est multiplié un volume élémentaire par changement de coordonnées :

- en coordonnées cylindriques, on a $x = r \cos(\theta)$, $y = r \sin(\theta)$ et $z = z$, de sorte que

$$dV = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial z} \end{vmatrix} dr d\theta dz = \begin{vmatrix} \cos(\theta) & -r \sin(\theta) & 0 \\ \sin(\theta) & r \cos(\theta) & 0 \\ 0 & 0 & 1 \end{vmatrix} dr d\theta dz = r dr d\theta dz.$$

- en coordonnées sphériques, on a $x = r \sin(\theta) \cos(\varphi)$, $y = r \sin(\theta) \sin(\varphi)$ et $z = r \cos(\theta)$, de sorte que

$$\begin{aligned} dV &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \varphi} \end{vmatrix} dr d\theta d\varphi = \begin{vmatrix} \sin(\theta) \cos(\varphi) & r \cos(\theta) \cos(\varphi) & -r \sin(\theta) \sin(\varphi) \\ \sin(\theta) \sin(\varphi) & r \cos(\theta) \sin(\varphi) & r \sin(\theta) \cos(\varphi) \\ \cos(\theta) & -r \sin(\theta) & 0 \end{vmatrix} dr d\theta d\varphi \\ &= \cos(\theta) (r^2 \cos(\theta) \sin(\theta)) + r \sin(\theta) (r \sin^2(\theta)) dr d\theta d\varphi = r^2 \sin(\theta) dr d\theta d\varphi. \end{aligned}$$

Dans le cas où f est la fonction constante égale à 1, l'intégrale $I(1)$ est le volume de \mathcal{V} . Par exemple, soit la boule centrée en O de rayon R , paramétrée en coordonnées sphériques. Son volume est égal à

$$\begin{aligned} \int_0^{2\pi} \int_0^\pi \int_0^R r^2 \sin(\theta) dr d\theta d\varphi &= \int_0^{2\pi} \int_0^\pi \sin(\theta) \left[\frac{r^3}{3} \right]_0^R d\theta d\varphi = \frac{R^3}{3} \times \int_0^{2\pi} [-\cos(\theta)]_0^\pi d\varphi \\ &= \frac{2R^3}{3} \times \int_0^{2\pi} d\varphi = \frac{2R^3}{3} \times 2\pi = \frac{4\pi R^3}{3}. \end{aligned}$$

On souhaite calculer l'intégrale du champ scalaire défini par $f(x, y, z) = xyz$ le long de l'intérieur du quart de cylindre d'équation $x^2 + y^2 = R^2$, compris entre les plans d'équations $x = 0$, $y = 0$, $z = 0$ et $z = 1$. On paramétrise le cylindre en coordonnées cylindriques, avec $0 \leq r \leq R$, $0 \leq \theta \leq \pi/2$ et $0 \leq z \leq 1$. On a donc

$$\begin{aligned} I(f) &= \int_0^1 \int_0^{\pi/2} \int_0^R r^2 \cos(\theta) \sin(\theta) z r dr d\theta dz = \int_0^1 \int_0^{\pi/2} \cos(\theta) \sin(\theta) z \left[\frac{r^3}{3} \right]_0^R d\theta dz \\ &= \frac{R^3}{3} \times \int_0^1 z \left[\frac{\sin^2(\theta)}{2} \right]_0^{\pi/2} dz = \frac{R^3}{6} \times \left[\frac{z^2}{2} \right]_0^1 = \frac{R^3}{12}. \end{aligned}$$

4.3. Théorème de Stokes et généralisations

Le théorème fondamentale du calcul intégral stipule que l'intégrale de la dérivée d'une fonction d'une variable est la différence des valeurs prises par la fonction sur le bord de l'intervalle d'intégration :

$$\int_a^b f'(t) dt = f(b) - f(a).$$

Le *théorème de Stokes* et ses adaptations, que sont le *théorème du gradient* et le *théorème de la divergence*, permettent également d'exprimer une intégrale curviligne, de surface ou volumique, d'un champ scalaire ou vectoriel grâce aux valeurs que prend ce champ sur le bord de son domaine d'intégration. Outre son intérêt calculatoire, le théorème de Stokes permet de faire le pont entre formes locales et globales d'équations de la physique, comme celles équations de Maxwell.

Le théorème de Stokes admet également une formulation générale pour les espaces courbés : il permet d'interpréter la *différentielle extérieure* comme un opérateur de bord. Il se trouve qu'en spécifiant cette formulation générale aux courbes, surfaces et volumes de l'espace, on obtient respectivement les théorèmes du gradient, de Stokes tel que présenté dans ce chapitre et de la divergence.

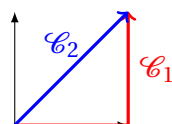
Théorème du gradient. Si f est un champ scalaire, alors l'intégrale de son gradient le long d'une courbe \mathcal{C} reliant les points P_i et P_f est la différence des valeurs prises par f aux extrémités de la courbe :

$$\int_{\mathcal{C}} \overrightarrow{\text{grad}}(f) \cdot \overrightarrow{dr} = f(P_f) - f(P_i).$$

Cela découle de l'équation 3.1 :

$$df = \overrightarrow{\text{grad}}(f) \cdot \overrightarrow{dr} \Rightarrow \int_{\mathcal{C}} \overrightarrow{\text{grad}}(f) \cdot \overrightarrow{dr} = \int_{\mathcal{C}} df = f(P_f) - f(P_i).$$

Une autre façon d'énoncer le théorème du gradient est que la circulation d'un champ de vecteurs conservatif $\vec{v} = \overrightarrow{\text{grad}}(f)$, ne dépend que des valeurs prises par le potentiel scalaire f aux extrémités de la courbe et pas de la courbe elle-même. Si on reprend l'exemple de la section précédente, on a vu que la circulation du champ d'Euler défini par $\vec{v} = x\vec{i} + y\vec{j} + z\vec{k}$ est égal à 1 le long des courbes

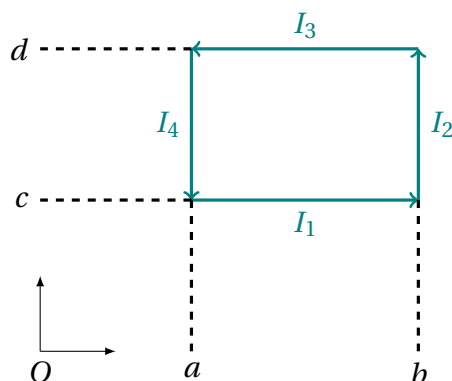


Or, le champ d'Euler dérive du potentiel scalaire défini par $f(x, y, z) = 1/2(x^2 + y^2 + z^2)$, qui prend bien les valeurs 1 et 0 aux extrémités $(1, 1, 1)$ et $(0, 0, 0)$ des courbes \mathcal{C}_1 et \mathcal{C}_2 .

On note enfin que le théorème du gradient généralise le théorème fondamentale du calcul intégral. En effet, soit une fonction d'une variable f , vue comme un champ scalaire de \mathbb{R}^2 , invariant selon la variable y . D'après l'expression du gradient en coordonnées cartésiennes, on a $\overrightarrow{\text{grad}}(f)(x, y) = f'(x)\vec{i}$. De plus, un intervalle $[a, b]$ est paramétré par $\vec{r}(t) = (tb + (1-t)a)\vec{i}$, avec $0 \leq t \leq 1$, de sorte que $\overrightarrow{dr} = (b-a)dt\vec{i}$ le long d'un tel intervalle. D'après le théorème du gradient, et en effectuant le changement de variable $x = tb + (1-t)a$, on a

$$f(b) - f(a) = \int_0^1 f'((tb + (1-t)a)(b-a)dt = \int_a^b f'(x)dx.$$

Théorème de Stokes. On commence par énoncer et démontrer le théorème de Green, qui n'est rien d'autre que le théorème de Stokes dans le plan. On considère un champ de vecteurs dans le plan défini par $\vec{v}(x, y) = v_x(x, y)\vec{i} + v_y(x, y)\vec{j}$ ainsi qu'un rectangle dans le plan



En notant le rectangle \mathcal{C} et la surface qu'il entoure \mathcal{S} , on a

$$\begin{aligned} \iint_{\mathcal{S}} \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) dS &= \int_c^d \int_a^b \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) dx dy = \int_c^d [v_y(x, y)]_a^b dy - \int_a^b [v_x(x, y)]_c^d dx \\ &= \int_c^d (v_y(b, y) - v_y(a, y)) dy - \int_a^b (v_x(x, d) - v_x(x, c)) dy \\ &= I_2 - I_4 - I_3 + I_1 = \oint_{\mathcal{C}} (v_x dx + v_y dy). \end{aligned}$$

On considère maintenant un champ de vecteurs \vec{v} et une surface \mathcal{S} entourée par une courbe fermée \mathcal{C} . On décompose la surface \mathcal{S} en petits rectangles et on applique la formule précédente à chacun de ces petits rectangles. En remarquant que le long de segments communs à deux rectangles, les circulations de \vec{v} sont en sens opposés et donc se compensent, on obtient la formule de Green :

$$\iint_{\mathcal{S}} \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) dS = \oint_{\mathcal{C}} (v_x dx + v_y dy).$$

Pour obtenir la formule de Stokes, il suffit de considérer la formule précédente dans l'espace : \vec{v} est un champ de vecteur dont la composante selon \vec{k} est nulle et la surface \mathcal{S} vit dans le plan d'équation $z = 0$ de l'espace. L'élément de surface orientée de \mathcal{S} est égal à $dS d\vec{k}$, puisque \vec{k} est un vecteur normal unitaire à \mathcal{S} . De plus, puisque \vec{v} s'annule selon l'axe des z , on a $\vec{v} \cdot d\vec{r} = v_x dx + v_y dy$ et

$$\text{rot}(\vec{v}) = \vec{\nabla} \times \begin{pmatrix} u_x \\ u_y \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \end{pmatrix},$$

de sorte que la formule de Green se réécrit :

$$\iint_{\mathcal{S}} \text{rot}(\vec{v}) \cdot d\vec{S} = \oint_{\mathcal{C}} \vec{v} \cdot d\vec{r}.$$

Il s'agit de la forme générale de la formule de Stokes : si \mathcal{S} est une surface de l'espace délimitée par une courbe fermée \mathcal{C} , alors le flux du rotationnel d'un champ de vecteurs à travers \mathcal{S} est égal à la circulation de ce champ de vecteurs le long de \mathcal{C} .

Important : la surface \mathcal{S} et la courbe \mathcal{C} doivent être orientées de façon cohérentes pour que la formule de Stokes soit valide. Le vecteur normal unitaire \vec{n} à \mathcal{S} est orienté selon la règle de la main droite en fonction de l'orientation de \mathcal{C} . Dans l'exemple du rectangle, celui-ci était parcouru dans le sens trigonométrique, si bien que le vecteur normal unitaire est bien \vec{k} et non $-\vec{k}$.

Théorème de la divergence. Étant donné un volume \mathcal{V} , entourée par une surface fermée \mathcal{S} et un champ de vecteurs \vec{v} de l'espace, le théorème de la divergence, également connu sous le nom de *formule d'Ostrogradski* stipule que le flux de \vec{v} à travers \mathcal{S} , **orientée vers l'extérieur**, est égal à l'intégrale de sa divergence à travers \mathcal{V} :

$$\iiint_{\mathcal{V}} \text{div}(\vec{v}) dV = \iint_{\mathcal{S}} \vec{v} \cdot d\vec{S}.$$

Comme pour le théorème de Green, il suffit de montrer ce théorème dans le cas particulier où \mathcal{V} est un parallélépipède carré : pour montrer la formule générale, il suffit de décomposer \mathcal{V} en petits cubes et de constater que les flux de \vec{v} à travers des surfaces communes se font en sens opposés et se compensent.

On considère donc un un parallélépipède \mathcal{V} , défini par les inéquations $0 \leq x \leq a$, $0 \leq y \leq b$ et $0 \leq z \leq c$, dont on note \mathcal{S} son contour, et un champ de vecteurs \vec{v} , de coordonnées v_x , v_y et v_z . On a

$$\begin{aligned} \iiint_{\mathcal{V}} \operatorname{div}(\vec{v}) dV &= \int_0^c \int_0^b \int_0^a \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \right) dx dy dz \\ &= \int_0^c \int_0^b [v_x(x, y, z)]_0^a dy dz + \int_0^c \int_0^a [v_y(x, y, z)]_0^b dx dz + \int_0^b \int_0^c [v_z(x, y, z)]_0^c dx dy \\ &= \int_0^c \int_0^b (v_x(a, y, z) - v_x(0, y, z)) dy dz + \int_0^c \int_0^a (v_y(x, b, z) - v_y(x, 0, z)) dx dz + \\ &\quad \int_0^b \int_0^c (v_z(x, y, c) - v_z(x, y, 0)) dx dy. \end{aligned}$$

On remarque que la première, la deuxième et le troisième intégrale sont les flux de \vec{v} à travers les surfaces d'équations $x = a$ et $x = 0$, $y = b$ et $y = 0$ et enfin $z = c$ et $z = 0$, respectivement, toutes ses surfaces étant orientées vers l'extérieur, ce qui montre le résultat voulu.

Dans un fluide, on considère deux quantités que sont la concentration ρ et la vitesse \vec{v} du fluide. Chacune de ces grandeurs dépend non seulement de la position dans le fluide, mais également du temps, si bien qu'il s'agit de fonctions dépendant de 4 variables t, x, y, z . On considère un volume infinitésimal \mathcal{V} , entouré par une surface \mathcal{S} . La masse de fluide à l'intérieur de \mathcal{V} évolue en fonction du flux à travers \mathcal{S} :

$$\frac{dM}{dt} = - \oiint_{\mathcal{S}} \rho \vec{v} \cdot \vec{dS}.$$

Le signe provient du fait que plus le flux à travers \mathcal{S} est important, plus la quantité de fluide dans \mathcal{V} diminue. De plus, la présence de ρ dans l'intégrale garantie que l'unité de celle-ci est bien une masse par unité de temps. De plus, la quantité de masse totale M est l'intégrale de la densité dans le volume :

$$M = \iiint_{\mathcal{V}} \rho dV,$$

de sorte qu'on a

$$\frac{dM}{dt} = \iiint_{\mathcal{V}} \frac{\partial \rho}{\partial t} dV.$$

En appliquant le théorème de la divergence, on obtient donc :

$$\iiint_{\mathcal{V}} \frac{\partial \rho}{\partial t} dV = - \oiint_{\mathcal{S}} \rho \vec{v} \cdot \vec{dS} = - \iiint_{\mathcal{V}} \operatorname{div}(\rho \vec{v}) dV \Rightarrow \iiint_{\mathcal{V}} \left(\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \vec{v}) \right) dV = 0.$$

La dernière équation étant vraie pour tout volume infinitésimal, on en déduit l'équation de continuité :

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \vec{v}) = 0.$$

Dans un fluide incompressible, la concentration ρ est constante, de sorte que l'équation de continuité devient $\operatorname{div}(\vec{v}) = 0$, ce qui signifie que la quantité de fluide entrant dans n'importe quel volume infinitésimal est égale à la quantité de volume sortant.

4.4. Fiche d'exercices 4

Exercice 1. Soient f et \vec{v} des champs scalaire et vectoriel de \mathbb{R}^3 , respectivement.

1. Montrer que

$$\operatorname{rot}(\overrightarrow{\operatorname{grad}}(f)) = \vec{0} \quad \text{et} \quad \operatorname{div}(\overrightarrow{\operatorname{rot}}(\vec{v})) = 0.$$

2. En déduire que si \vec{v} est conservatif, alors $\overrightarrow{\operatorname{rot}}(\vec{v}) = \vec{0}$.

3. On considère les champs de vecteurs \vec{u} et \vec{v} définis par

$$\vec{u}(x, y, z) = \begin{pmatrix} x^2(1+y^3) \\ y^2(1+x^3) \\ 0 \end{pmatrix} \quad \text{et} \quad \vec{v} = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}.$$

Ces champs de vecteurs sont-ils conservatifs? Si c'est le cas, expliciter un potentiel scalaire.

Exercice 2. On considère les champs scalaires f sur \mathbb{R}^2 et g, h sur \mathbb{R}^3 définis par

$$f(x, y) = x^2 + y^2, \quad g(x, y, z) = e^{z^2}, \quad \text{et} \quad h(x, y, z) = xyz.$$

1. Calculer l'intégrale de f le long du triangle passant par les sommets $(0, 0)$, $(1, 0)$, $(1, 1)$.

2. Calculer l'intégrale de g le long de la surface entourant le cône d'équation $z^2 = x^2 + y^2$ et bornée par le plan d'équation $z = 1$.

3. Calculer l'intégrale de h le long du quart de la boule unité, défini comme étant l'intersection de la boule unité avec les demi-espaces d'équations $0 \leq x$, $0 \leq y$ et $0 \leq z$.

Exercice 3. On considère les deux champs de vecteurs de \mathbb{R}^3 définis par $\vec{u}(x, y, z) = -y^2 \vec{i} + x \vec{j} + z^2 \vec{k}$ et $\vec{v} = (x + y^2) \vec{i} + x^2 \vec{j} + z \vec{k}$.

1. Soit la courbe \mathcal{C} de \mathbb{R}^3 définie par les équations $y + z = 2$ et $x^2 + y^2 = 1$.

a). Représenter \mathcal{C} sur un dessin. Quelle propriété remarquable a la courbe \mathcal{C} ?

b). Calculer la circulation de \vec{u} le long de \mathcal{C} . On rappelle pour cela que $2 \cos^2(\theta) = 1 + \cos(2\theta)$.

c). Retrouver le résultat de la question précédente en utilisant la formule de Stokes.

2. Soit la surface \mathcal{S} d'équation $x^2 + y^2 = 1$ comprise entre les plans d'équations $z = \pm 1$.

a). Représenter \mathcal{S} sur un dessin. Quelle propriété remarquable a la surface \mathcal{S} ?

b). Calculer flux de \vec{v} à travers \mathcal{S} , orientée vers l'extérieur.

c). Retrouver le résultat de la question précédente en utilisant la formule de Stokes.

Exercice 4. On rappelle que le travail d'une force nette \vec{F} s'exerçant sur un point mobile est égale à la différence d'énergie cinétique le long de la trajectoire C décrite par le point mobile :

$$\int_C \vec{F} \cdot d\vec{r} = T_f - T_i.$$

1. Montrer que si \vec{F} dérive d'un potentiel, alors la quantité totale d'énergie est conservée.

2. Calculer l'énergie potentielle d'une masse m en chute libre dans un champ gravitationnel induit par une masse M . On rappelle pour cela la formule du gradient en coordonnées sphériques :

$$\overrightarrow{\text{grad}}(f)(r, \theta, \varphi) = \frac{\partial f}{\partial r}(r, \theta, \varphi) \vec{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta}(r, \theta, \varphi) \vec{e}_\theta + \frac{1}{r \sin(\theta)} \frac{\partial f}{\partial \varphi}(r, \theta, \varphi) \vec{e}_\varphi.$$

3. On se donne un point mobile situé dans le champ gravitationnel d'un astre et circulant à la *vitesse d'évasion* de celui-ci, *i.e.*, la vitesse minimale que doit atteindre ce point pour échapper à l'attraction gravitationnelle de l'astre.
- Justifier qu'après un temps suffisamment long, on peut supposer que les énergies cinétiques et potentielles du mobile sont nulles. En déduire la formule de la vitesse d'évasion.
 - Quelle vitesse faut-il imposer à un mobile situé sur terre pour qu'il s'évade de l'attraction terrestre? On rappelle que le rayon et la masse de la terre sont respectivement 6370 km et $5,97 \cdot 10^{24} \text{ kg}$, et que $G = 6,67 \cdot 10^{-11} \text{ m}^3 \cdot \text{kg}^{-1} \cdot \text{s}^{-2}$.

Exercice 5. On rappelle les équations de Maxwell sous leurs formes différentielles :

$$\text{div}(\vec{E}) = \frac{\rho}{\epsilon_0}, \quad \overrightarrow{\text{rot}}(\vec{E}) = -\frac{\partial \vec{B}}{\partial t}, \quad \text{div}(\vec{B}) = 0, \quad \overrightarrow{\text{rot}}(\vec{B}) = \mu_0 \vec{j} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}.$$

- Montrer qu'en l'absence de charge, un potentiel électrique V vérifie l'équation de Laplace $\Delta V = 0$.
- Montrer les formes intégrales des équations de Maxwell :

$$\begin{aligned} \oiint_{\mathcal{S}} \vec{E} \cdot \vec{dS} &= \frac{Q}{\epsilon_0}, & \oint_{\mathcal{C}} \vec{E} \cdot \vec{dr} &= -\frac{d}{dt} \left(\iint_{\mathcal{S}} \vec{B} \cdot \vec{dS} \right), \\ \oiint_{\mathcal{S}} \vec{B} \cdot \vec{dS} &= 0, & \oint_{\mathcal{C}} \vec{B} \cdot \vec{dr} &= \mu_0 \left(I + \epsilon_0 \frac{d}{dt} \left(\int_{\mathcal{S}} \vec{E} \cdot \vec{dS} \right) \right), \end{aligned}$$

où \mathcal{C} et \mathcal{S} sont des courbes/surfaces (fermées) quelconques, Q est la charge totale à l'intérieur du volume limité par \mathcal{S} et I est le flux total du courant à travers la surface \mathcal{S} .

- On cherche maintenant à démontrer l'existence d'ondes électro-magnétiques.
 - Montrer que pour tout champ de vecteur \vec{v} , on a

$$\overrightarrow{\text{rot}}(\overrightarrow{\text{rot}}(\vec{v})) = \overrightarrow{\text{grad}}(\text{div}(\vec{v})) - \Delta \vec{v}.$$

En déduire que dans le vide $\overrightarrow{\text{rot}}(\overrightarrow{\text{rot}}(\vec{E})) = -\Delta \vec{E}$.

- Montrer que

$$\overrightarrow{\text{rot}}(\overrightarrow{\text{rot}}(\vec{E})) = -\mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2}.$$

En déduire l'existence d'ondes électro-magnétiques dans le vide.