



MATHÉMATIQUES GÉNÉRALES DEUXIÈME ANNÉE

Version enseignant

Mourad ABOUZAIID

**ISABTP 2^o année
2020-2021**

Dernière compilation : 20 août 2020

Table des matières

1	Équations différentielles	5
	Introduction	5
1.1	Classification des équations différentielles	6
1.1.1	Définitions	6
1.1.2	Équations linéaires	7
1.1.3	Équations autonomes	7
1.1.4	Systèmes différentiels	7
1.2	Équations différentielles linéaires	8
1.2.1	Théorème de Cauchy-Lipschitz	8
1.2.2	Structure de l'ensemble des solutions	9
1.2.3	Équations différentielles linéaires à coefficients constants	11
1.2.4	Équations linéaires à coefficients non constants (ordre 1)	18
1.3	Équations différentielles non linéaires (ordre 1)	24
1.3.1	Équations séparables	25
1.3.2	Changement de variable	25
1.4	Étude qualitative d'équations différentielles autonomes (ordre 1)	27
1.4.1	Principes généraux	27
1.4.2	Un exemple	27
1.4.3	Classification des points fixes (en dimension 1)	30
2	Espaces vectoriels et applications linéaires	33
	Introduction	33
2.1	Le plan et l'espace euclidiens	34
2.1.1	Qu'est-ce qu'un vecteur ?	34
2.1.2	Points et vecteurs	35
2.1.3	Repères et coordonnées	36
2.1.4	Coordonnées et opérations	37
2.2	Définitions et Exemples	38
2.2.1	Loi de composition interne	38
2.2.2	Multiplication externe	39
2.2.3	Espace vectoriel	40
2.2.4	Sous espaces vectoriels (SEV)	41
2.2.5	Sommes de sous espaces vectoriels	44
2.3	Familles de vecteurs	45
2.3.1	Familles génératrices	46
2.3.2	Familles libres	48

2.3.3	Bases et coordonnées	51
2.3.4	Espaces vectoriels de dimension finie	52
2.3.5	Matrice d'une famille de vecteurs	56
2.4	Applications linéaires	59
2.4.1	Définitions	59
2.4.2	Image et noyau d'une application linéaire	62
2.4.3	Injectivité, surjectivité des applications linéaires	64
2.4.4	Applications linéaires en dimension finie	65
2.5	Diagonalisation des endomorphismes et des matrices carrées	73
2.5.1	Définitions	73
2.5.2	Diagonalisation pratique	76
2.5.3	Diagonalisation des matrices symétriques	79
3	Coniques, quadriques, formes quadratiques	83
	Introduction	83
3.1	Définition géométrique 2D	85
3.2	Définition analytique	88
3.2.1	Une première équation	88
3.2.2	Coniques analytiques	98
3.2.3	Réduction d'une équation	101
3.3	Généralisation aux dimensions supérieures	105
3.3.1	Formes quadratiques de \mathbb{R}^n	105
3.3.2	Réduction matricielle des formes quadratiques	105
3.3.3	Quadriques de l'espace	106
4	Séries numériques, séries entières	109
	Introduction	109
4.1	Définitions	110
4.1.1	Sommes partielles	110
4.1.2	Convergence	110
4.2	Deux classiques	113
4.2.1	Les séries géométriques	113
4.2.2	Les séries télescopiques	114
4.3	Opérations sur l'ensemble des séries	115
4.4	Les séries alternées	116
4.5	Séries à termes positifs	117
4.5.1	Définition	117
4.5.2	Critères de comparaison	118
4.5.3	Séries de Riemann	121
4.5.4	Critères de D'Alembert et de Cauchy	124
4.6	Séries entières	125
4.6.1	Définitions	125
4.6.2	Rayon de convergence	125
4.6.3	Opérations sur les séries entières	128
4.6.4	Régularité d'une série entière	129
4.6.5	Développement en séries entières	130

4.6.6	Application aux équations différentielles	132
5	Fonctions de plusieurs variables	135
	Introduction	135
5.1	Définitions	136
5.1.1	Graphe d'une fonction	137
5.1.2	Isoclines	138
5.2	Fonctions et dérivées partielles	139
5.2.1	Fonctions partielles	140
5.2.2	Dérivées partielles	141
5.3	Plan tangent et différentielle	143
5.3.1	Définition	143
5.3.2	Équation du plan tangent	143
5.3.3	Différentielle	144
5.3.4	Développements limités	145
5.3.5	Recherche d'extrema	148
5.4	Champs de vecteurs	152
5.4.1	Définitions	152
5.4.2	Champs de gradient	154
5.4.3	Variations d'un champs de vecteurs	157
5.4.4	Changements de variables de \mathbb{R}^n	161
5.5	Équations aux dérivées partielles	163
6	Compléments d'intégration	167
6.1	Intégrales multiples	167
6.1.1	Motivation	167
6.1.2	Intégration sur un rectangle	170
6.1.3	Domaines variés	172
6.1.4	Interprétation géométrique (2D)	174
6.1.5	Propriétés du calcul intégral	175
6.1.6	Changement de variables en dimensions supérieures	176
6.2	Intégrales curvilignes, intégrales de surfaces	180
6.2.1	Paramétrisation	181
6.2.2	Intégrales curvilignes	183
6.2.3	Intégrales de surfaces	190

Chapitre 1

Équations différentielles

Introduction

Les équations différentielles sont l'outil principal permettant de modéliser le mouvement. En effet, une équation différentielle est une équation liant une fonction et ses dérivées. Or les lois de la physique sur le mouvement donnent l'accélération (i.e. la dérivée seconde de la position) d'un objet, d'une particule,... en fonction des forces qui agissent dessus. Ces forces sont souvent fonctions de la vitesse de l'objet (forces de frottement) et/ou de sa position (tension d'un ressort, champs magnétique, répartition de charges,...). Si l'on parvient à résoudre de façon exacte les équations ainsi obtenues, on peut connaître à chaque instant la position de l'objet que l'on étudie.

On peut ainsi modéliser le mouvement d'un pendule, l'évolution d'une population, la déformation d'une poutre, d'un mur, le mouvement des vagues, l'évolution d'une réaction chimique et de façon générale tout système (évolutif ou non) dont les différents états peuvent être donnés par des variables quantifiables.

L'objet de ce cours est de présenter différentes méthodes de résolutions.

Notons qu'il existe en réalité très peu d'équations différentielles que l'on sait résoudre de façon exactes. Elles sont en général issues de modèles simples, qui donnent une première idée du comportement de l'objet étudié. Ces équations sont donc triées en différentes classes. Chaque méthode étudiée ne s'appliquera qu'à certains type d'équations.

Si l'on souhaite affiner les modèles on obtient alors en général des équations que l'on ne sait pas résoudre. On verra en fin de chapitre que pour ces équations, il existe différentes méthodes d'études permettant d'obtenir

- de l'information qualitative,
- des solutions approchées (par approximation numérique).

On verra enfin que l'on peut s'inspirer de ces méthodes pour résoudre des problèmes plus complexes, portant notamment sur des fonctions de plusieurs variables. Les fonctions de plusieurs variables, et les équations différentielles associées (appelées Équations aux Dérivées Partielles ou EDP) sont au cœur de tous les modèles théoriques tentant

d'expliquer le monde qui nous entoure (diffusion de chaleur, mécanique des fluides, etc).

1.1 Classification des équations différentielles

1.1.1 Définitions

Une équation différentielle est une équation dont l'inconnue est une fonction réelle (y) d'une variable réelle (t) et liant cette fonction et ses dérivées. Formellement, une équation différentielle est donc de la forme

$$(E) : \Phi(y, y', \dots, y^{(n-1)}, y^{(n)}, t) = f(t).$$

où Φ est une fonction connue explicitant le lien entre la fonction inconnue et ses dérivées.

Définitions

- L'entier n est appelé *l'ordre* de l'équation (E).
- La fonction f (connue) est appelée *second membre* de l'équation (E). Si cette fonction est nulle, l'équation est dite *homogène*.

Notes :

- En règle générale, les équations différentielles issues des modèles physiques sont d'ordre 1 ou 2. Dans certains modèles peuvent cependant apparaître des équations d'ordres plus grand.
- Du point de vue de la modélisation, les équations homogènes permettent de modéliser des systèmes libres, i.e. sans intervention extérieure ; le second membre f (quand il est non nul) traduisant une intervention extérieure.

1.1.2 Équations linéaires

Définition

- Une équation différentielle est dite *linéaire* si la fonction Φ associée est *linéaire en y* :

$$(E) : a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = f(t)$$

$$\iff \sum_{k=0}^n a_k(t)y^{(k)} = f(t)$$

- Les fonctions a_k (connues) sont appelées *coefficients de l'équation (E)*. Si ces fonctions sont des constantes, l'équation différentielle linéaire (E) est dite à *coefficients constants*.

Note : les équations différentielles linéaires sont principalement les seules équations différentielles pour lesquelles il existe des méthodes systématiques de résolution, notamment dans le cas des équations linéaires à coefficients constants.

1.1.3 Équations autonomes

Définition

Une équation différentielle (E) est dite *autonome* si elle est *homogène* et si la fonction Φ ne dépend pas de t .

Exemples :

- Les équations différentielles linéaires *homogènes* à coefficients constants sont des équations autonomes.
- L'équation $y' = y^2$ est autonome non linéaire. Nous verrons dans la suite qu'il existe des protocoles d'étude qualitative pour ce type d'équations différentielles.

1.1.4 Systèmes différentiels

À l'image des systèmes d'équations algébriques, on appelle *système différentiel* un ensemble de p équations différentielles portant sur p fonctions inconnues y_1, \dots, y_p d'une même variable t :

$$(S) : \begin{cases} \Phi_1(y_1, y_1', \dots, y_1^{(n)}, \dots, y_p, y_p', \dots, y_p^{(n)}, t) = f_1(t) \\ \Phi_2(y_1, y_1', \dots, y_1^{(n)}, \dots, y_p, y_p', \dots, y_p^{(n)}, t) = f_2(t) \\ \vdots \\ \Phi_p(y_1, y_1', \dots, y_1^{(n)}, \dots, y_p, y_p', \dots, y_p^{(n)}, t) = f_p(t) \end{cases}$$

Il est alors possible de classer les systèmes différentiels en fonction des caractéristiques de chacune des fonctions Φ_k ci-dessus. On verra en outre qu'il existe en lien fort entre équations d'ordres n et systèmes différentiels.

1.2 Équations différentielles linéaires

1.2.1 Théorème de Cauchy-Lipschitz

Résoudre l'équation différentielle du type

$$(E) : \sum_{k=0}^n a_k(t)y^{(k)} = f(t)$$

consiste à déterminer l'ensemble des fonctions y qui vérifient (E) .

Afin de construire des méthodes permettant de déterminer toutes ces solutions, on commence par étudier la structure de l'ensemble des solutions.

Dans un premier temps, on peut montrer que, en règle générale, une équation du type (E) admet une infinité de solutions. D'un point de vue pratique, cela signifie que l'on peut (ou doit) ajouter des contraintes supplémentaires au problème afin d'obtenir un problème *clos*, i.e. un problème admettant une et une seule solution. Ces contraintes supplémentaires peuvent se donner sous deux formes :

— Des *conditions initiales* : on impose la valeur de y et certaines de ses dérivées en une valeur t_0 fixée :

$$y(t_0) = y_0, \quad y'(t_0) = y_1, \quad \dots$$

— Des *conditions limites* : on impose la valeur de y en différents points du domaine d'étude :

$$y(t_0) = y_0, \quad y(t_1) = y_1, \quad \dots$$

On peut alors montrer que l'ordre de l'équation (E) correspond exactement au nombre de contraintes supplémentaires que l'on doit imposer pour obtenir un problème clos. Précisément,

Définition (Problème de Cauchy)

On appelle *Problème de Cauchy* tout problème différentiel constitué de

- une équation différentielle d'ordre n ,
- n conditions limites et/ou initiales

$$(\mathcal{P}) : \begin{cases} \Phi(y, y', \dots, y^{(n)}, t) = f(t) & (E) \\ y(t_0) = y_0, y'(t_0) = y_1, \dots, y^{(n-1)}(t_0) = y_{n-1} & (C.I.) \end{cases}$$

ou

$$(\mathcal{P}) : \begin{cases} \Phi(y, y', \dots, y^{(n)}, t) = f(t) & (E) \\ y(t_0) = y_0, y(t_1) = y_1, \dots, y(t_{n-1}) = y_{n-1} & (C.I.) \end{cases}$$

Par ailleurs, un problème de Cauchy sera dit linéaire si l'équation (E) est linéaire.

On peut alors énoncer le théorème de Cauchy-Lipschitz suivant :

Théorème (Cauchy-Lipschitz)

Tout problème de Cauchy linéaire admet une et une seule solution, définie sur un intervalle (ouvert) contenant l'ensemble des valeurs de t imposées par les conditions supplémentaires.

En pratique, ce théorème est une traduction mathématique de la notion de *modèle déterministe* : si l'on connaît l'état d'un système dynamique à un instant donné (les conditions initiales) et que l'on connaît ses règles d'évolution (l'équation), alors on peut connaître l'état du système à chaque instant, donné par l'unique solution du problème de Cauchy {Équation + (C.I.) ou (C.L.)}.

Par ailleurs, à l'ordre 1, ce résultat s'interprète d'un point de vue géométrique de la façon suivante : si (E) est une équation différentielle linéaire d'ordre 1, en tout point (t_0, y_0) du plan (tOy) passe une et une seule courbe associée à une solution de (E) . Autrement dit, l'ensemble des courbes représentant les solutions de (E) ne se croisent jamais et recouvrent à elles toutes l'ensemble du plan (tOy) .

Enfin, notons que ce résultat, énoncé pour les équations linéaires, se généralise à d'autres classes d'équations différentielles.

1.2.2 Structure de l'ensemble des solutions

Les équations différentielles linéaires forment l'une des rares classes d'équations différentielles pour lesquelles il existe une méthode de résolution systématique. Cette méthode est basée sur une étude approfondie de *la structure* de l'ensemble des solutions d'une telle équation.

Intuitivement, cette structure est issue du *principe de superposition* :

— Soit (H) une équation différentielle linéaire à coefficients constants et *homogène* :

$$(H) : a_n(t)y^{(n)} + \dots + a_1(t)y' + a_0(t)y = 0 \quad \text{ou} \quad \sum_{k=0}^n a_k(t)y^{(k)} = 0$$

Si y_1 et y_2 sont deux solutions de (H) , alors

- la somme $y_1 + y_2$ est encore une solution de (H) ,
- pour tout $\lambda \in \mathbb{R}$, le produit λy_1 est encore une solution de (H) .

— Soit (E) une équation différentielle linéaire à coefficients constants et non homogène :

$$(E) : a_n(t)y^{(n)} + \dots + a_1(t)y' + a_0(t)y = f(t) \Leftrightarrow \sum_{k=0}^n a_k(t)y^{(k)} = f(t)$$

Si y_p est une solution de (E) , pour toute solution y_h de l'équation *homogène* associée à (E) , alors la somme $y_p + y_h$ est encore une solution de (E) .

Exercice : à démontrer

Formellement, cela permet d'établir une analogie entre l'ensemble des solutions d'une équation différentielle linéaire d'ordre 1 ou 2 et les droites du plan ou les plans de l'espace, qui produit une méthode de résolution qui se généralise aux ordres supérieurs. Précisément, on peut montrer que

— Si (H) est une équation différentielle linéaire *homogène* d'ordre n , il existe un nombre n de solutions y_1, \dots, y_n dites *de base* à partir desquelles on peut exprimer l'ensemble des solutions Σ_h de (H) sous forme de *combinaisons linéaires* :

$$\Sigma_h = \{\lambda_1 y_1 + \dots + \lambda_n y_n, \quad \lambda_i \in \mathbb{R}\}$$

— Si (E) est une équation différentielle linéaire *non homogène*, en notant y_p une solution de (E) , l'ensemble des solutions Σ de (E) est

$$\Sigma = \{y_h + y_p, \quad y_h \in \Sigma_h\}$$

En pratique, pour une équation différentielle linéaire *homogène* d'ordre n , ce résultat réduit l'étude à la recherche de n solutions de base. On va voir comment construire ces n solutions dans le cas des équations linéaires à *coefficients constants* en s'appuyant sur la notion de polynôme.

Dans le cas non homogène, on commence par résoudre l'équation homogène associée

(en trouvant n solutions de base) puis il reste à déterminer *une* solution de l'équation complète. En additionnant cette dernière aux solutions trouvées pour l'équation homogène, on obtient l'ensemble des solutions de l'équation étudiée. En pratique, cette seconde étape est souvent la plus complexe, bien que le problème physique de laquelle est issue l'équation permette parfois d'obtenir une solution particulière "évidente".

Dans tous les cas, l'ensemble des solutions s'exprime alors en fonction de n constantes $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$. Si l'équation étudiée est issue d'un problème de Cauchy, l'unique solution cherchée est alors obtenue en injectant les conditions supplémentaires dans la forme générale obtenue. En pratique, on obtient ainsi un système de n équations algébriques portant sur les n constantes $\lambda_1, \dots, \lambda_n$. Le théorème de Cauchy-Lipschitz assure que ce système est inversible (i.e. admet une unique solution) et sa résolution donne alors les constantes associées à l'unique solution cherchée.

1.2.3 Équations différentielles linéaires à coefficients constants

Équations homogènes

Équations d'ordre 1. Une équation différentielle linéaire homogène d'ordre 1 à coefficients constants est une équation de la forme

$$(H_1) : \alpha y' + \beta y = 0, \quad \alpha, \beta \in \mathbb{R}, \quad \alpha \neq 0$$

Quitte à diviser par α (qui est non nul), on peut se ramener à une équation de la forme

$$(H_1) \iff y' = ay$$

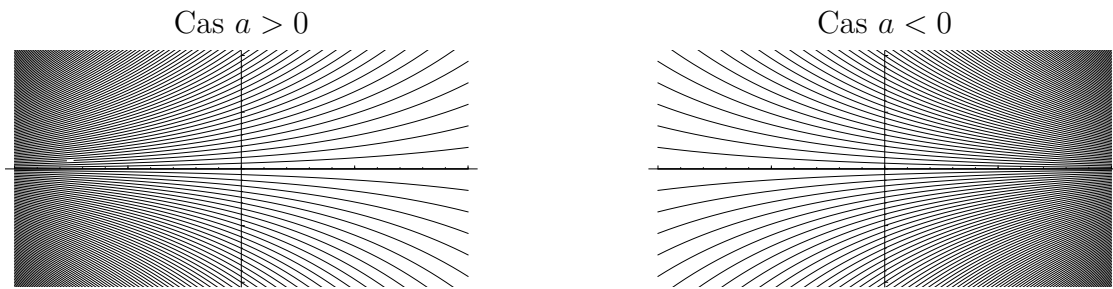
Il y a de nombreuses façons de résoudre ce type d'équations. La théorie évoquée plus haut permet par exemple d'affirmer que l'ensemble des solutions de l'équation (H_1) peut s'exprimer à partir d'une seule solution de base. Or on peut montrer rapidement que la fonction

$$y_1 : t \mapsto e^{at}$$

vérifie l'équation (H_1) (à faire). De ce fait, l'ensemble des solutions de (H_1) est

$$\Sigma_1 = \{\lambda \cdot y_1, \lambda \in \mathbb{R}\} = \{t \mapsto \lambda \cdot e^{at}, \lambda \in \mathbb{R}\}$$

D'un point de vue géométrique, on constate que, comme attendu, l'ensemble des courbes associées à chacune de ces solutions ne se rencontrent pas et recouvrent à elles toutes l'ensemble des points du plan :

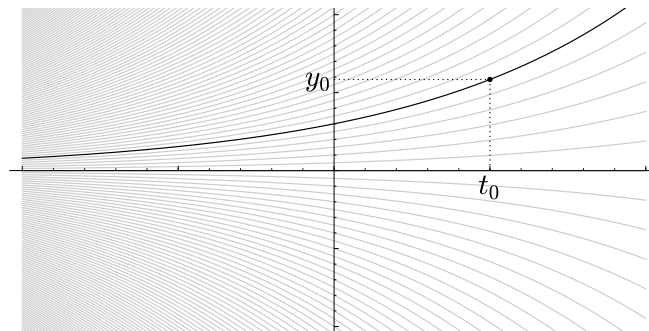


D'autre part, s'il on ajoute à (H_1) une condition supplémentaire du type $y(t_0) = y_0$ donné, on obtient un problème de Cauchy. La constante λ correspondant à l'unique solution de ce problème de Cauchy est obtenue en injectant cette condition supplémentaire dans la forme générale : puisque la fonction cherchée est une solution de (H_1) , elle est de la forme $y : t \mapsto \lambda e^{at}$. Mais alors

$$\begin{aligned} y(t_0) = y_0 &\iff \lambda e^{at_0} = y_0 \\ &\iff \lambda = y_0 e^{-at_0} \end{aligned}$$

la solution cherchée est donc

$$y : t \mapsto y_0 \cdot e^{a(t-t_0)}$$



Équations d'ordre 2. Une équation différentielle linéaire homogène, d'ordre 2, à coefficients constants est une équation de la forme

$$(H_2) : ay'' + by' + cy = 0, \quad a, b, c \in \mathbb{R}$$

La théorie relative à la structure de l'ensemble des solutions d'une équation de ce type indique qu'il suffit de déterminer deux solutions indépendantes y_1 et y_2 pour cette équation et qu'alors, l'ensemble des solutions de (H_2) pourra s'exprimer à l'aide de combinaisons linéaires de ces deux solutions de base. Or il est possible de construire ces deux solutions de base à l'aide du *polynôme caractéristique* de l'équation (H_2) , construit à partir des coefficients de (H_2) . Précisément,

Définition

On appelle *polynôme caractéristique* de l'équation (H_2) le polynôme

$$P(X) = aX^2 + bX + c$$

Il est alors possible de construire des solutions de l'équation (H_2) à partir des racines de $P(X)$. Précisément,

Propriété

Si $r \in \mathbb{C}$ est une racine de $P(X)$ (i.e. si $P(r) = 0$), alors la fonction

$$t \mapsto e^{rt}$$

est une solution de (H_2) .

Exercice : à démontrer

Ainsi, si $P(X)$ admet deux racines réelles r_1 et r_2 (i.e. si son discriminant $\Delta = b^2 - 4ac$ est strictement positif), chacune de ces racines permet de construire une solution de (H_2) et l'ensemble des solutions de (H_2) est

$$\Sigma_2 = \{y : t \mapsto \lambda e^{r_1 t} + \mu e^{r_2 t}, \lambda, \mu \in \mathbb{R}\}$$

Par ailleurs, si le discriminant Δ est strictement négatif, le polynôme $P(X)$ admet encore deux racines distinctes, mais ce sont deux racines complexes conjuguées $r_1 \in \mathbb{C}$ et $r_2 = \bar{r}_1$.

Dans ce cas, il est alors possible de décrire l'ensemble des solutions de (H_2) à l'aide d'exponentielles (complexes), avec une subtilité :

$$\Sigma_2 = \{y : t \mapsto \lambda e^{r_1 t} + \mu e^{r_2 t}, \lambda \in \mathbb{C}, \mu = \bar{\lambda}\}$$

Cependant, il est possible d'éviter le passage en complexe à l'aide des fonctions trigonométriques. Précisément, on peut montrer que

Si le polynôme $P(X)$ admet deux racines complexes conjuguées

$$r_1 = \alpha + i\beta \quad \text{et} \quad r_2 = \bar{r}_1 = \alpha - i\beta$$

les fonctions (réelles)

$$y_1 : t \mapsto e^{\alpha t} \cos(\beta t) \quad \text{et} \quad y_2 : t \mapsto e^{\alpha t} \sin(\beta t)$$

sont des solutions de (H_2) .

Exercice : à démontrer.

Ces deux solutions étant indépendantes, l'ensemble des solutions de (H_2) s'écrit alors

$$\Sigma_2 = \{y : t \mapsto e^{\alpha t}(\lambda \cos(\beta t) + \mu \sin(\beta t)), \lambda, \mu \in \mathbb{R}\}$$

Enfin, si le discriminant Δ est nul, cela signifie que le polynôme $P(X)$ admet une unique racine réelle r_0 , qualifiée de racine double. Cette racine permet encore de construire une solution de (H_2) , mais il nous manque alors une solution de base. On peut cependant montrer que, dans ce cas, la fonction

$$t \mapsto t.e^{r_0 t}$$

est également une solution de (H_2) (exercice : à démontrer). On a ainsi nos deux solutions de base, et l'ensemble des solutions de (H_2) est alors

$$\Sigma_2 = \{y : t \mapsto \lambda e^{r_0 t} + \mu t.e^{r_0 t}, \lambda, \mu \in \mathbb{R}\}$$

Notes :

- D'un point de vue qualitatif, on peut distinguer deux situations différentes pour les équations d'ordre 2 :
 - Si le discriminant de l'équation est positif ou nul, les solutions ont un comportement exponentiel.
 - Si le discriminant de l'équation est strictement négatif, les solutions ont un comportement oscillant. Dans ce cas, la partie imaginaire (β) des racines complexes du polynôme caractéristique correspond à la pulsation de l'oscillation et la partie réelle (α) produit un terme d'amortissement (si $\alpha < 0$) en d'entraînement (si $\alpha > 0$).
- Comme pour les équations d'ordre 1, si l'équation (H_2) étudiée est issue d'un problème de Cauchy contenant également deux conditions initiales et/ou limites, l'introduction de ces conditions supplémentaires dans la forme générale des solu-

tions produit un système de deux équations à deux inconnues (les coefficients λ et μ) admettant une unique solution permettant d'obtenir l'unique solution au problème de Cauchy étudié.

Équations d'ordre n . De façon générale, il est possible d'étendre cette méthode de résolution aux équations différentielles linéaires homogènes à coefficients constants de *n'importe quel ordre n* . Précisément, si (H) est une équation de ce type :

$$(H) : a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = 0, \quad a_i \in \mathbb{R},$$

on peut, à partir des coefficients a_i de (H) construire le polynôme caractéristique de (H) :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0 = \sum_{k=0}^n a_k X^k$$

C'est un polynôme degré n admettant $p \leq n$ racines distinctes r_1, r_2, \dots, r_p et une forme factorisée du type

$$P(X) = a_n (X - r_1)^{\alpha_1} (X - r_2)^{\alpha_2} \dots (X - r_p)^{\alpha_p}$$

où les entiers $\alpha_1, \dots, \alpha_p$ donnent les *ordres de multiplicité* des racines associées et vérifient

$$\sum_{k=1}^p \alpha_k = n$$

Chaque racine r d'ordre de multiplicité α produit alors α solutions distinctes :

$$[t \mapsto e^{rt}], [t \mapsto te^{rt}], \dots, [t \mapsto t^{\alpha-1} e^{rt}]$$

et l'on obtient ainsi les n solutions de base.

Note : si $P(X)$ admet des racines complexes, alors celles ci sont conjuguées deux à deux. Pour deux racines complexes conjuguées r et \bar{r} ayant le même ordre de multiplicité α , on peut alors remplacer les 2α solutions complexes par des fonctions trigonométriques, multipliée par des puissances de t .

Le cas non homogène

Soit

$$(E) : \sum_{k=0}^n a_k y^{(k)} = f(t)$$

une équation différentielle linéaire à coefficients constants non homogène. D'après la théorie évoquée plus haut, la résolution de (E) se fait en deux étapes :

1. Déterminer l'ensemble Σ_h des solutions de l'équation homogène

$$(H) : \sum_{k=0}^n a_k y^{(k)} = 0$$

en appliquant le protocole du paragraphe ci-dessus.

2. Déterminer UNE solution particulière y_p de (E) .

Comme cela a été évoqué, l'ensemble Σ des solutions de (E) est alors

$$\Sigma = \{y_h + y_p, \quad y_h \in \Sigma_h\}$$

Notons que cette seconde étape est souvent la plus compliquée à résoudre. Cependant, il n'est pas rare que le contexte physique duquel est en général issu l'équation à résoudre produise une solution évidente.

Par ailleurs, il est également possible de s'inspirer du second membre $f(t)$ pour chercher cette solution "évidente". Précisément, dans chacun des cas énoncés ci-dessous, on cherche une solution ayant la même forme que le second membre. En introduisant des paramètres dans la forme en question, on peut calculer les dérivées successives de la fonction cherchée et injecter la forme choisie dans l'équation. Par identification, on obtient alors des conditions sur les paramètres choisis. Si l'on peut satisfaire l'ensemble des ces conditions, on obtient une solution particulière y_p . Sinon, c'est que l'équation étudiée n'a pas de solution sous la forme cherchée. Ainsi,

- Si la fonction f est une constante, on peut chercher une solution particulière sous la forme $y_p(t) = k$. L'équation (E) donne alors des conditions sur la constante k .
- Si la fonction f est un polynôme, on cherche une solution y_p sous la forme d'un polynôme de même degré. L'équation (E) donne alors des conditions sur les coefficients de ce polynôme.
- Si la fonction f est de la forme $f(t) = P(t)e^{at}$, on cherche y_p sous la forme $y_p(t) = Q(t)e^{at}$ où $Q(t)$ est un polynôme de même degré que $P(t)$. L'équation (E) donne là encore des conditions sur les coefficients du polynôme $Q(t)$.
- Si la fonction f est de la forme $f(t) = \alpha \cos(\omega t) + \beta \sin(\omega t)$, on cherche y_p sous la forme $y_p(t) = a \cos(\omega t) + b \sin(\omega t)$. L'équation donne alors des conditions sur les coefficients a et b .

Notes :

- Comme cela a été évoqué plus haut, il n'est pas certain que cette méthode marche à tous les coups. Cependant, si cela ne donne rien, on peut recommencer en multipliant la forme choisie par t (ce qui revient à augmenter le degré du polynôme cherché lorsqu'un polynôme intervient).
- Si le second membre f est une somme de fonctions du type de celles évoquées ci-dessus, à l'aide du principe de superposition, il est possible de traiter chaque terme de cette somme séparément. Ainsi, si par exemple $f(t) = 1 + e^{2t}$, on peut traiter l'équation avec pour second membre $f_1(t) = 1$, puis traiter l'équation avec pour second membre $f_2(t) = e^{2t}$. On obtient ainsi deux solutions particulières y_{p1} et y_{p2} et on peut montrer que la somme $y_p = y_{p1} + y_{p2}$ est une solution particulière pour l'équation de départ.

Enfin, à l'ordre 1, quand l'instinct ne nous donne rien, on peut appliquer la méthode de *variation de la constante* : la résolution d'une équation linéaire homogène d'ordre 1 produit une constante. On a vu par exemple que les solutions de l'équation $y' = ay$ sont

$$y(t) = \lambda e^{at}, \quad \lambda \in \mathbb{R}.$$

Si l'on souhaite résoudre une équation de la forme $y' - ay = f(t)$ où $f(t)$ est connue (non nulle), on peut alors chercher une solution particulière de cette équation sous la forme

$$y_p(t) = \lambda(t)e^{at}.$$

(On fait varier la constante issue de la résolution de l'équation homogène associée). Dans ce cas, on a

$$y_p'(t) = \lambda'(t)e^{at} + \lambda(t).ae^{at}$$

et en injectant y_p dans l'équation, on obtient :

$$\begin{aligned} y_p'(t) - ay_p(t) = f(t) &\Leftrightarrow (\lambda'(t) + \cancel{a\lambda(t)})e^{at} - \cancel{a\lambda(t)}e^{at} = f(t) \\ &\Leftrightarrow \lambda'(t) = f(t)e^{-at}. \end{aligned}$$

Si l'on peut trouver une primitive à la fonction $t \mapsto f(t)e^{-at}$, on trouve $\lambda(t)$ puis une solution particulière à l'équation étudiée.

Pour finir, notons que si l'équation étudiée (E) est issue d'un problème de Cauchy, il faut prendre soin de terminer complètement la résolution de l'équation (E) avant d'introduire les conditions supplémentaires.

Petits systèmes différentiels linéaires

Les méthodes ci-dessus s'adaptent également à la résolution de “petits” systèmes différentiels linéaires dont chaque équation est linéaire à coefficients constants. À l'aide de quelques manipulations élémentaires des équations d'un tel système, on peut en effet, pour un système d'ordre 1 de n équations portant sur n fonctions inconnues, faire apparaître, une équation d'ordre n portant sur l'une des n inconnues. Cette équation hérite de l'ensemble des propriétés du système de départ et peut donc être résolue à l'aide des méthodes évoquées plus haut. Une fois cette équation résolue, on se base sur les équations du système pour en déduire la forme des autres fonctions inconnues.

Exemple : soit

$$(S) : \begin{cases} x' = 2x + y & (E_1) \\ y' = -x + y & (E_2) \end{cases}$$

1. À l'aide de la première équation, exprimer les fonctions y et y' en fonction de x et de ses dérivées.
2. Montrer, à l'aide de la question précédente que x vérifie l'équation

$$(E) : x'' - 4x' + 5x = 0$$

3. En déduire $x(t)$ puis $y(t)$ sous forme explicite.

Note : on verra qu'il existe une correspondance forte entre systèmes différentiels de n équations d'ordre 1 et équations d'ordre n . On verra également qu'à l'aide du calcul matriciel, il est possible de mettre en place des méthodes d'étude efficaces pour les systèmes différentiels. Ces méthodes sont tellement efficaces qu'à l'inverse de ce que l'on vient de faire dans l'exemple ci-dessus, on préférera transformer une équation d'ordre n en un système de n équations d'ordre 1 pour pouvoir appliquer les appliquer.

1.2.4 Équations linéaires à coefficients non constants (ordre 1)

Concernant les équations linéaires à coefficients non constants, bien que les propriétés théoriques évoquées plus haut concernant la structure de l'ensemble des solutions reste valable, la résolution pratique en est autrement plus complexe, en particulier parce que, dans ce cas, la notion de polynôme caractéristique n'a plus de sens. De ce fait, nous nous contenterons ici d'étudier les équations *d'ordre 1*.

Par ailleurs, on verra que, même dans ce cas simple, la méthode générale de résolution fait apparaître la notion de valeurs interdites pour le paramètre t . Contrairement au cas précédent, où l'ensemble des solutions obtenues sont définies sur \mathbb{R} tout entier, il faudra alors composer avec la notion d'*intervalles de résolution*.

Ainsi, soit

$$(E) : \alpha(t)y' + \beta(t)y = f(t)$$

une équation différentielle linéaire d'ordre 1, à coefficients variables.

Puisqu'il s'agit encore d'une équation linéaire, la démarche générale reste la même que dans le cas précédent : pour déterminer l'ensemble des solutions de (E) , on commence par résoudre l'équation homogène associée, puis on cherche une solution particulière de l'équation complète.

Cas homogène

Considérons l'équation homogène (H) associée à (E) :

$$(H) : \alpha(t)y' + \beta(t)y = 0$$

En posant

$$a(t) = -\frac{\beta(t)}{\alpha(t)}$$

on a

$$(H) \iff y' = a(t)y$$

En divisant cette équation par y , on obtient une égalité entre deux quantités que l'on peut intégrer. Précisément :

$$\begin{aligned} (H) &\iff \frac{y'}{y} = a(t) \\ &\iff \ln |y(t)| = A(t) + k && \text{où } A \text{ est une primitive de } a \text{ et } k \in \mathbb{R} \\ &\iff |y(t)| = e^{A(t)+k} \\ &\iff y(t) = \lambda \cdot e^{A(t)}, && \lambda = \pm e^k \in \mathbb{R} \end{aligned}$$

En résumé, l'ensemble des solutions de $(H) : \alpha(t)y' + \beta(t)y = 0$ est

$$\Sigma_h = \{y_h : t \mapsto \lambda e^{A(t)}, \lambda \in \mathbb{R}\}$$

où A est une primitive de la fonction $a : t \mapsto -\frac{\beta(t)}{\alpha(t)}$.

Note : on constate ici que, comme attendu, les solutions de (H) sont tous les multiples d'une solution de base $y_1 : t \mapsto e^{A(t)}$.

Exemple : Soit

$$(H) : y' + 2ty = 0$$

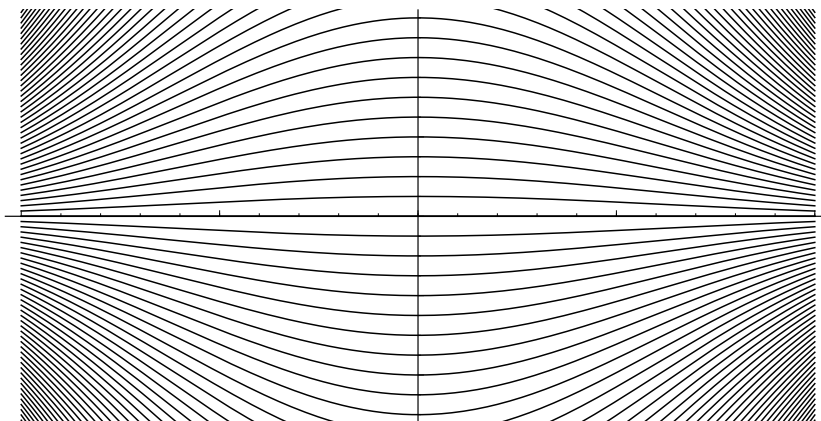
On a ici

$$a(t) = -2t \Rightarrow A(t) = -t^2$$

L'ensemble des solutions de (H) est alors

$$\Sigma_H = \{y_h : t \mapsto \lambda \cdot e^{-t^2}, \quad \lambda \in \mathbb{R}\}$$

D'un point de vue géométrique, ces solutions sont représentées par une famille de courbes du plan (tOy) qui ne se croisent pas et qui, à elles toutes, recouvrent l'ensemble du plan :



Cas non homogène

Pour résoudre l'équation complète (E) , on doit alors déterminer UNE solution y_p de (E) . Là encore, le plus simple consiste à déterminer une solution "évidente" en se basant, soit sur le problème physique ayant produit l'équation, soit en s'appuyant sur le second membre $f(t)$. Cependant, les coefficients α et β dépendant de t , cela peut rendre cette recherche plus compliquée que dans le cas constant. Dans ce cas, on doit alors se tourner vers la méthode dite de *variation de la constante* qui, comme dans le cas constant, consiste à faire varier la constante λ obtenue à l'étape précédente. Précisément, on cherche une solution particulière y_p sous la forme

$$y_p : t \mapsto \lambda(t)e^{A(t)}$$

On a alors

$$y_p'(t) = \lambda'(t) \cdot e^{A(t)} + \lambda(t) \cdot a(t)e^{A(t)}$$

et en injectant cette forme dans l'équation $(E) : y' - a(t)y = f(t)$, on a

$$\begin{aligned}
 y_p'(t) - a(t)y_p(t) = f(t) &\Leftrightarrow \lambda'(t).e^{A(t)} + \cancel{\lambda(t).a(t)e^{A(t)}} - \cancel{a(t)\lambda(t)e^{at}} = f(t) \\
 &\Leftrightarrow \lambda'(t) = f(t)e^{-A(t)}.
 \end{aligned}$$

Si l'on connaît une primitive de la fonction $t \mapsto f(t)e^{-A(t)}$, on obtient une fonction λ de t puis une solution particulière y_p .

Exemple : soit

$$(E) : y' + 2ty = e^{-t^2}$$

D'après l'étude précédente, les solutions de l'équation homogène associée à (E) sont

$$y_h : t \mapsto \lambda.e^{-t^2}, \quad \lambda \in \mathbb{R}$$

On cherche donc une solution particulière sous la forme

$$y_p : t \mapsto \lambda(t)e^{-t^2}$$

Or, puisque $f(t) = e^{-t^2}$, d'après le calcul ci-dessus, y_p est une solution de (E) si et seulement si

$$\lambda'(t) = f(t)e^{t^2} = e^{-t^2}.e^{t^2} = 1$$

Or, une primitive de la fonction $t \mapsto 1$ étant

$$\lambda(t) = t$$

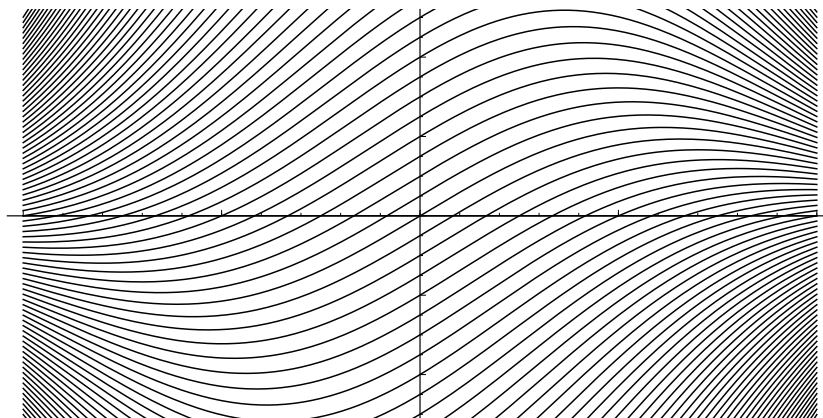
on a

$$y_p(t) = t.e^{-t^2}$$

Autrement dit, la fonction $t \mapsto te^{-t^2}$ est une solution de (E) (on peut le vérifier) et l'ensemble des solutions de (E) est

$$\Sigma = \left\{ y : t \mapsto \lambda e^{-t^2} + te^{-t^2}, \quad \lambda \in \mathbb{R} \right\}$$

Graphiquement, l'ensemble de ces solutions est encore associé à une famille de courbes du plan (tOy) qui ne se croisent pas et qui, à elles toutes, recouvrent l'ensemble du plan.



Intervalles de résolution

D'un point de vue théorique, la méthode de résolution employée ci-dessus peut imposer quelques restrictions sur les valeurs du paramètre t pour lesquelles elle est valable. En effet, dès la première étape de résolution, on divise l'égalité étudiée par le coefficient $\alpha(t)$ de y' . L'ensemble de la méthode n'est alors valable que pour les valeurs de t pour lesquelles ce coefficient est non nul. Ainsi,

Définition

On appelle *valeurs interdites* de l'équation

$$(E) : \alpha(t)y' + \beta(t)y = f(t)$$

les solutions de l'équation

$$\alpha(t) = 0$$

Par ailleurs, la théorie générale des équations différentielles nous oblige à résoudre toute équation différentielle sur UN intervalle. Ainsi, les valeurs interdites obtenues par la résolution de l'équation $\alpha(t) = 0$ imposent un découpage de l'axe réel en une famille d'intervalles : les intervalles de résolution de (E) .

Une fois que l'on connaît l'ensemble de ces intervalles, on doit alors résoudre (E) sur *chacun de ces intervalles*.

Exemples :

1.

$$(E) : y' + 2ty = e^{-t^2}$$

Ici, $\alpha(t) = 1$. Il n'y a donc pas de valeur interdite et toutes les solutions de (E) sont définies sur \mathbb{R} .

2.

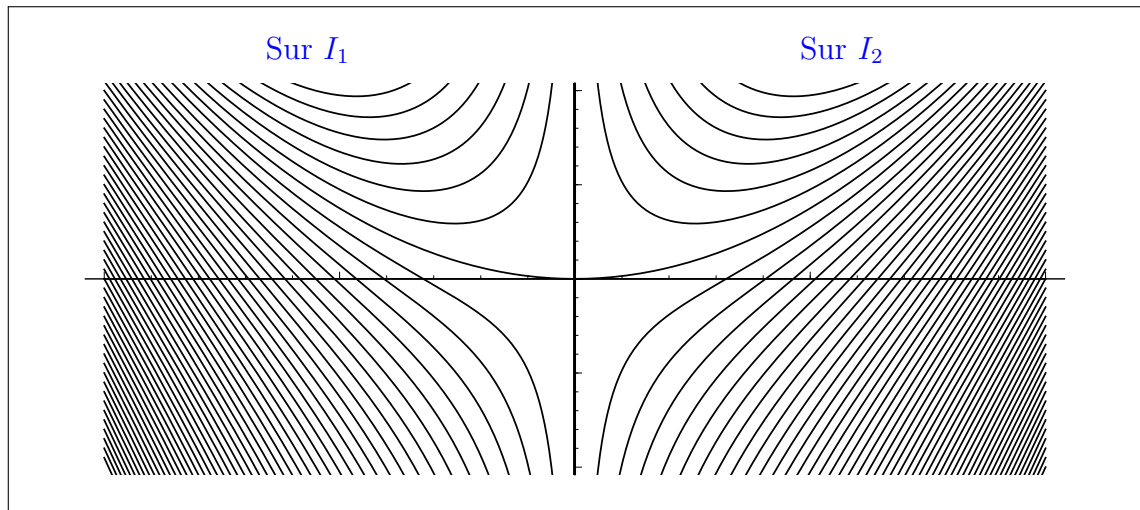
$$(E) : ty' + y = 3t^2$$

Ici, $\alpha(t) = t$. L'équation (E) admet donc une unique valeur interdite $t = 0$ et les intervalles de résolution sont

$$I_1 =]-\infty, 0[= \mathbb{R}_*^- \quad \text{et} \quad I_2 =]0, +\infty[= \mathbb{R}_*^+$$

Après résolution, on trouve donc deux ensembles de solutions :

$$\begin{aligned} \text{— sur } I_1 : \Sigma_1 &= \left\{ y : t \mapsto \frac{\lambda}{t} + 3t^2, \lambda \in \mathbb{R} \right\} \\ \text{— sur } I_2 : \Sigma_2 &= \left\{ y : t \mapsto \frac{\mu}{t} + 3t^2, \mu \in \mathbb{R} \right\} \end{aligned}$$



3.

$$(E) : (1-t)ty' + (1-t)y = 1$$

Ici, on a $\alpha(t) = (1-t)t$. Cette équation admet donc deux valeurs interdites :

$$t_1 = 0 \quad \text{et} \quad t_2 = 1$$

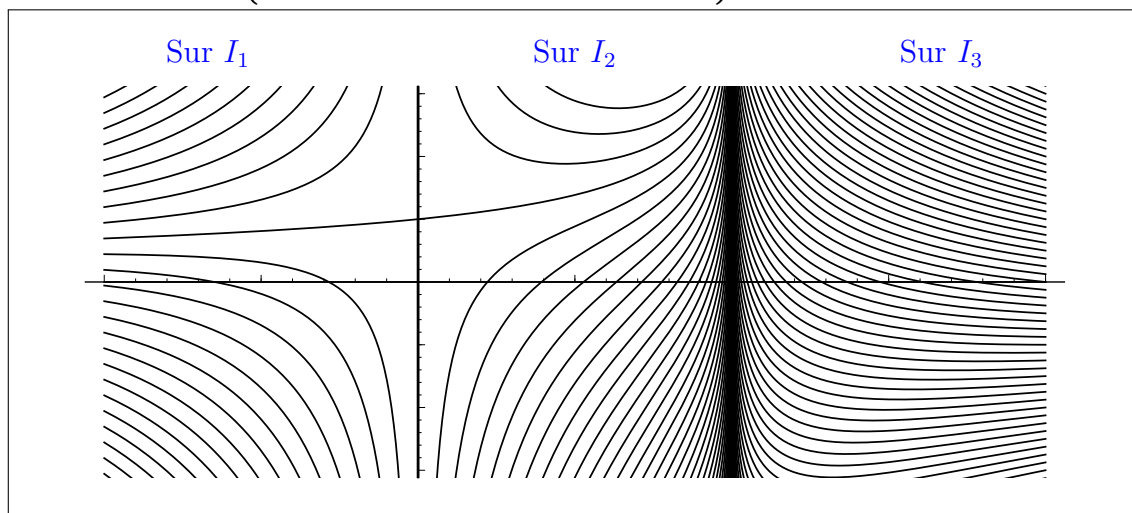
Les intervalles de résolution de (E) sont donc

$$I_1 =]-\infty, 0[, \quad I_2 =]0, 1[, \quad I_3 =]1, +\infty[$$

Après résolution, on obtient donc trois ensembles de solutions :

$$\begin{aligned} \text{— sur } I_1 : \Sigma_1 &= \left\{ y : t \mapsto \frac{\lambda_1}{t} - \frac{\ln(1-t)}{t}, \lambda_1 \in \mathbb{R} \right\} \\ \text{— sur } I_2 : \Sigma_2 &= \left\{ y : t \mapsto \frac{\lambda_2}{t} - \frac{\ln(1-t)}{t}, \lambda_2 \in \mathbb{R} \right\} \end{aligned}$$

$$\text{— sur } I_3 : \Sigma_3 = \left\{ y : t \mapsto \frac{\lambda_3}{t} - \frac{\ln(t-1)}{t}, \lambda_3 \in \mathbb{R} \right\}$$



Condition initiale et prolongement

Si l'équation étudiée est issue d'un problème de Cauchy, complété d'une condition initiale de la forme $y(t_0) = y_0$, on peut montrer que, là encore, le problème complet admet une et une seule solution. D'autre part, cette unique solution est définie sur l'unique intervalle de résolution qui contient la valeur t_0 du paramètre.

Par ailleurs, il est possible que, parmi toutes les solutions de (E) sur un intervalle donné, certaines "traversent" les valeurs interdites et se prolongent à l'un des intervalles adjacents en se "recollant" à l'une des solutions définies sur cet intervalle adjacent. Ainsi, dans le deuxième exemple donné plus haut, on voit que l'une des solutions définies sur I_1 se prolonge à l'intervalle I_2 en se raccordant à l'une des solutions de celui-ci, donnant ainsi une solution de (E) définie sur \mathbb{R} tout entier.

Dans le troisième exemple, on constate également que l'une des solutions de l'intervalle I_1 se prolonge en une solution définie sur $] -\infty, 1[$. Cependant, on peut montrer qu'aucune des solutions de (E) ne "traverse" la valeur interdite $t = 1$.

On verra comment, en pratique, on peut étudier d'éventuels prolongements aux intervalles voisins en étudiant les limites aux bords des intervalles de résolution.

1.3 Équations différentielles non linéaires (ordre 1)

Bien que, comme cela a déjà été évoqué, il n'existe pas de méthode générale pour résoudre les équations différentielles non linéaires, il existe tout de même certaines classes d'équations pour lesquelles on peut obtenir de l'information sur l'ensemble des solutions. Ainsi, nous verrons dans un premier temps certains exemples d'équations non linéaires pour lesquelles on peut obtenir l'ensemble des solutions sous forme explicite.

Nous verrons ensuite comment, dans le cas des équations autonomes, on peut effectuer une étude qualitative de ces équations et obtenir de l'information qualitative pour l'ensemble de leurs solutions, même si l'on ne peut accéder à leur forme explicite.

1.3.1 Équations séparables

Une équation séparable est une équation que l'on peut mettre sous la forme

$$y' = f(t)g(y) \quad \text{ou encore} \quad y'h(y) = f(t)$$

(avec $h = \frac{1}{g}$).

Si l'on connaît une primitive de h et une primitive de f , on peut intégrer l'équation :

$$H(y(t)) = F(t) + k$$

d'où on tire $y(t)$.

Note : cette méthode est loin de s'appliquer à toutes les équations séparables. Il faut tout d'abord pouvoir intégrer les fonctions f et h puis il faut pouvoir extraire $y(t)$ de l'équation $H(y(t)) = F(t) + k$.

Exemples :

- $y' = e^{t-y}$.
- $y' = \frac{1}{y(1+t^2)}$.

1.3.2 Changement de variable

Pour certaines équations, on peut changer de variable y , le but étant de transformer l'équation en une équation que l'on sait résoudre (linéaire, par exemple). En effectuant le changement de variable inverse sur les solutions trouvées, on obtient les solutions de la première équation.

Exemples : Pour chacune des équation ci-dessous, poser le changement de variable proposer et déterminer l'équation linéaire associée.

- *Une équation de Bernoulli :*

$$(B_1) : y' + y = y^2$$

Poser $z = \frac{1}{y}$.

On a $y = \frac{1}{z}$ donc $y' = -\frac{z'}{z^2}$. D'où

$$(B_1) \Leftrightarrow -\frac{z'}{z^2} + \frac{1}{z} = \frac{1}{z^2} \Leftrightarrow -z' + z = 1$$

— Une équation de Riccati :

$$(R) : y' = (y - 1)(ty - y - t)$$

Poser $z = \frac{1}{y-1}$, soit $y = 1 + \frac{1}{z}$.

— Un dernier exemple :

$$(E) : t^2 + y^2 - tyy' = 0$$

Poser $z = y/t$.

1.4 Étude qualitative d'équations différentielles autonomes (ordre 1)

1.4.1 Principes généraux

On rappelle qu'une équation différentielle autonome est une équation homogène dont la fonction Φ qui la définit ne dépend pas de t . On va voir pour finir comment étudier d'un point de vue qualitatif les équations autonomes que l'on peut présenter sous la forme

$$(E) : y' = F(y)$$

Précisément, on va voir comment une étude théorique de la fonction F ci-dessus permet de déterminer certaines solutions particulières de (E) (les solutions constantes) et permet de déterminer le comportement (sens de variation, limite) des autres solutions.

On supposera dans la suite que, là encore, l'ensemble des solutions de (E) peut être représentée par une famille de courbes du plan que ne se coupent pas et qui, à elles toutes, recouvrent l'ensemble du plan. C'est en particulier le cas si la fonction F est continue.

Précisément, on peut montrer que les solutions constantes de l'équation (E) sont obtenues en résolvant l'équation

$$F(x) = 0$$

En effet, si a est une constante, solution de l'équation ci-dessus, la fonction constante $t \mapsto a$ est une solution de (E) (à vérifier).

Ces solutions sont appelées *points fixes* de (E) (ou de F) et jouent un rôle fondamental dans l'évolution de toute solution de (E) .

Par ailleurs, si l'on parvient à dresser le tableau de signe de la fonction F (dans lequel apparaissent ses points fixes), on peut connaître les variations d'une solution y en fonction de sa position (i.e. de sa valeur $y(t)$).

Enfin, le graphe de la fonction F (qui nous donne y' en fonction de y) nous permet de tracer l'allure générale des solutions (dans un repère $(t, y(t))$).

1.4.2 Un exemple

Étudions un exemple issu de la dynamique des populations. On considère une population donnée et on note $y(t)$ le nombre d'individus à l'instant t . Certains modèles d'évolution de cette population sont basés sur le problème de Cauchy suivant :

$$(P) : \begin{cases} y' = y(1 - y) & (E) \\ y(0) = y_0 & (C.I.) \end{cases}$$

La valeur y_0 donnant le nombre d'individus à $t = 0$.

La fonction F définissant cette équation est

$$F : x \mapsto x(1 - x)$$

une rapide étude de cette fonction donne

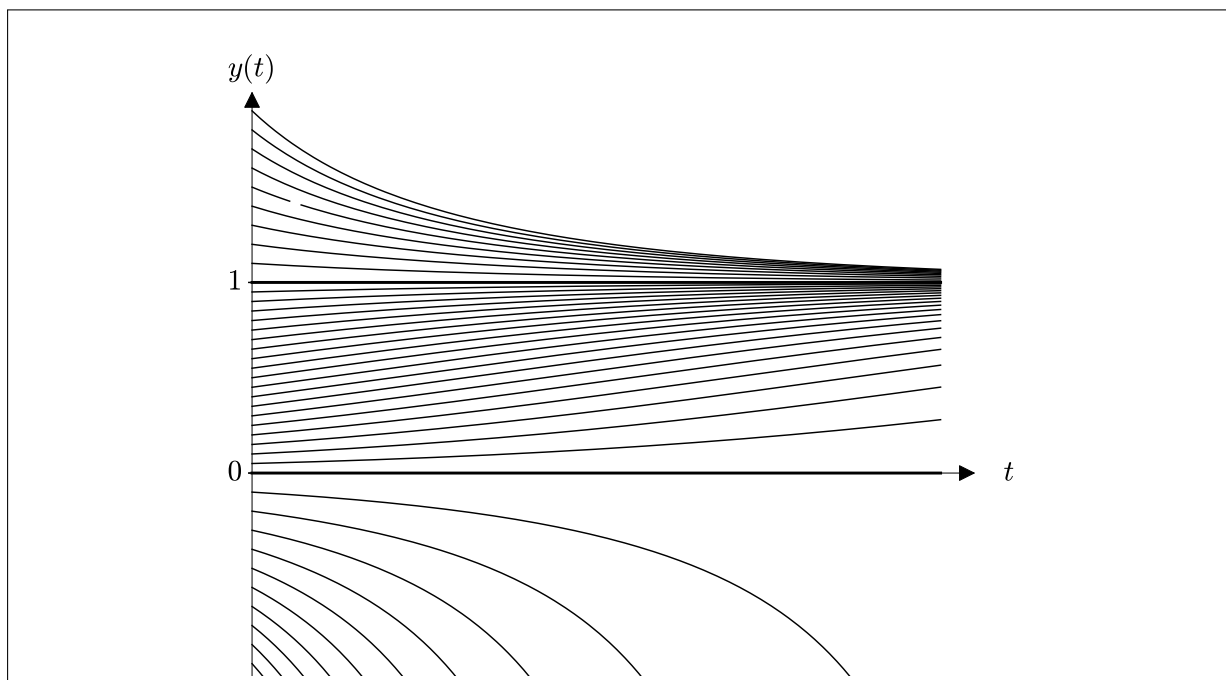
— Ses points fixes

$$a = 0 \quad \text{et} \quad b = 1$$

— Son tableau de signe est

x	$-\infty$	0	1	$+\infty$
$F(x)$	$-$	0	$+$	0
				$-$

À partir de ces informations, on va voir qu'il est possible de comprendre le comportement global de l'ensemble des solutions de (E) en distinguant différentes valeurs de la condition initiale y_0 . On pourra en particulier représenter ces résultats dans le repère (tOy) suivant :



On peut ainsi dans un premier temps faire apparaître les deux solutions constantes de (E) :

$$t \longmapsto 0 \quad \text{et} \quad t \longmapsto 1$$

correspondant respectivement aux conditions initiales

$$y_0 = 0 \quad \text{et} \quad y_0 = 1$$

Dans un second temps, étudions la solution de (E) associée à la condition initiale $y_0 = \frac{1}{2}$. Du tableau de signe de la fonction F , on déduit le signe de la dérivée de y à $t = 0$:

$$y'(0) = F(y(0)) = F\left(\frac{1}{2}\right) > 0$$

La solution étudiée est donc croissante à $t = 0$. Elle restera en outre croissante tant qu'elle ne dépassera pas la valeur 1. Mais puisque les différentes solutions de (E) ne peuvent se croiser, la solution étudiée restera inférieure à 1, donc croissante. On en déduit qu'elle croît et admet $y = 1$ comme asymptote horizontale.

De plus, ce raisonnement effectué pour $y_0 = \frac{1}{2}$ reste valable pour toute valeur de $y_0 \in]0, 1[$. On obtient donc une famille de courbes entre 0 et 1.

Dans un troisième temps, étudions la solution de (E) associée à $y_0 = \frac{3}{2}$. Du tableau de signe de F , on déduit que

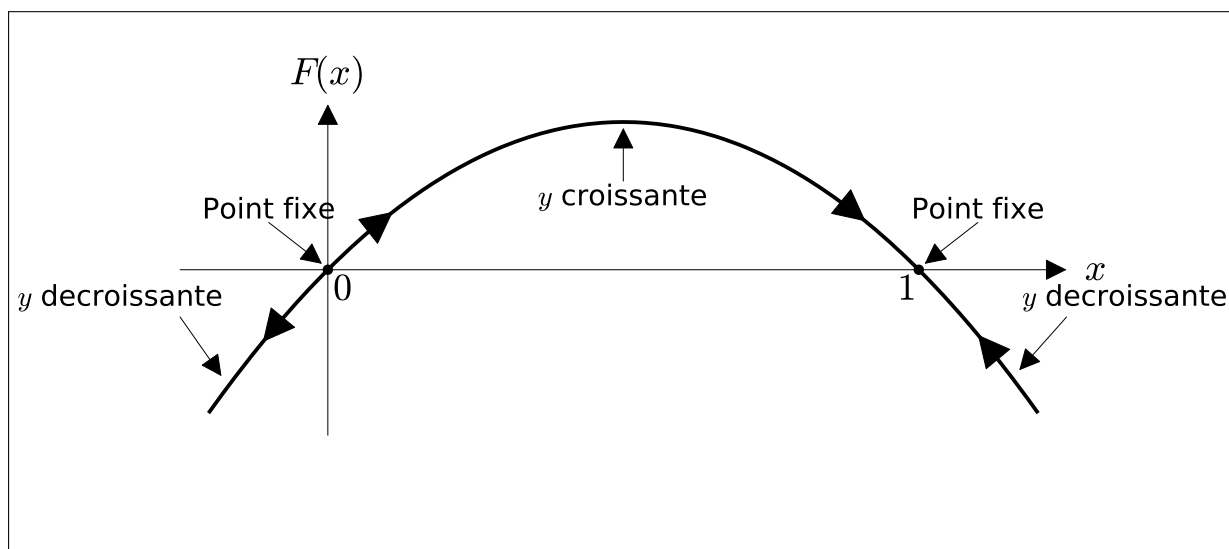
$$y'(0) = F(y(0)) = F\left(\frac{3}{2}\right) < 0$$

La solution étudiée ici est donc décroissante à $t = 0$. Comme sa courbe reste toute au dessus de la solution constante $y = 1$, cette solution reste décroissante et admet la droite d'équation $y = 1$ pour asymptote horizontale.

De plus, ce raisonnement reste valable pour tout $y_0 > 1$ et on obtient une famille de courbes au dessus de la droite d'équation $y = 1$.

Enfin, bien que cela n'ait de sens physique, on peut également étudier le cas $y_0 < 0$. Avec un raisonnement analogue aux deux premiers, on montre que toute solution associée à une condition initiale $y_0 < 0$ est décroissante et l'on obtient une dernière famille de courbes pour les cas $y_0 < 0$.

L'ensemble de cette étude peut être résumé sur le graphe de la fonction F :



Note : Une fois cette étude terminée, on peut l’interpréter en termes d’évolution de population. Par ailleurs, l’équation étudiée est en réalité une équation de Bernoulli. Il est possible d’obtenir ses solutions sous forme explicite à l’aide du changement de variables $z = \frac{1}{y}$. Il est laissé au lecteur le soin de résoudre cette équation de façon exacte et de vérifier que les solutions obtenues ont bien les propriétés établies lors de cette étude qualitative.

1.4.3 Classification des points fixes (en dimension 1)

Dans l’exemple ci-dessus, on a pu noter que les deux points fixes obtenus par la résolution de l’équation $F(x) = 0$ ne jouent pas le même rôle dans le système modélisé. On peut résumer cette différence de la façon suivante :

- Le point fixe $a = 0$ “repousse” les solutions. Il est dit *instable*.
- Le point fixe $b = 1$ “attire” les solutions. Il est dit *stable*.

Physiquement, cette notion de stabilité permet de prendre en compte de petites perturbations du système. Ainsi, concernant le point fixe $b = 1$, la théorie nous dit que si une solution y de l’équation étudiée prend pour valeur 1 à $t = 0$, alors elle reste constante égale à 1. Cependant, si l’on perturbe légèrement le système étudié stationnaire en 1, alors la fonction y retourne cet état stationnaire. A l’inverse, si le système est à la position d’équilibre $a = 0$, toute petite perturbation le fait sortir de cette position et le système s’en éloigne.

Formellement, on peut déterminer la nature d’un point fixe (stable ou instable) en étudiant directement la fonction F . Précisément, soit a un point fixe de la fonction F .

- Le point fixe $y = a$ est stable si $F'(a) < 0$.
- Le point fixe $y = a$ est instable si $F'(a) > 0$.

Dans le cas où $F'(a) = 0$, il faut étudier le signe de $F''(a)$ pour conclure. Précisément,

- Si $F''(a) > 0$, la courbe de F autour de a est convexe. Le point fixe a est donc stable si l'on déplace y vers la gauche et instable si l'on déplace y vers la droite. Il est dit *semi-stable à gauche*.
- Si $F''(a) < 0$, la courbe de F autour de a est concave. Le point fixe a alors *semi-stable à droite*.

Si $F''(a) = 0$, on étudie alors $F^{(3)}(a)$.

Remarque : dans la réalité, les points fixes semi-stables n'existent pas. En effet, ils correspondent à un état du système qui disparaît si le système subit une toute petite variation (ce qui ne peut manquer de se produire). On a alors disparition du point fixe ou apparition d'un couple de points fixes stable/instable. Les points fixes semi-stables sont donc dits structurellement instables.

Chapitre 2

Espaces vectoriels et applications linéaires

Introduction

L'algèbre linéaire rassemble l'ensemble des outils permettant d'étudier, de manipuler et d'utiliser la notion de vecteur.

L'utilisation la plus courante des vecteurs se fait en géométrie car la notion de repère et coordonnées permet de représenter les différentes positions du plan ou de l'espace sous forme numérique.

En effet, à l'aide de la notion de coordonnées, on peut représenter des points, des contours, des volumes, mais également des positions relatives, des déplacements...

Exemples :

- Si l'on se place dans un repère orthonormé du plan ou de l'espace, on peut définir *le produit scalaire* qui permet de déterminer si deux vecteurs donnés sont orthogonaux ou non à partir de leurs coordonnées.
- De même, *le produit vectoriel* permet de construire, à partir de deux vecteurs, un troisième vecteur orthogonal aux deux premiers, toujours en s'aidant des coordonnées.

Nous verrons que l'algèbre linéaire offre de nombreux outils de ce type (et d'autres bien plus élaborés) pour l'étude du plan ou de l'espace basé sur méthodes analytiques faisant intervenir les coordonnées.

Nous verrons en particulier que les matrices jouent un rôle fondamental en algèbre linéaire et qu'elles permettent en particulier de ramener la plupart des problèmes à la résolution d'un système linéaire.

Cependant, le développement et l'utilisation de ces outils est basée sur une théorisation poussée de les notions de vecteurs et de coordonnées.

Pour bien en comprendre le fonctionnement et les différentes applications, il est donc nécessaire de connaître et comprendre les détails toutes ces notions.

Après quelques rappels sur les vecteurs, nous verrons comment mettre en place le formalisme de l'algèbre linéaire.

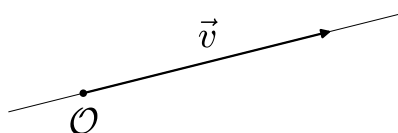
Nous verrons ensuite quelles utilisation on peut faire de ce formalisme non seulement en algèbre linéaire, mais également dans de nombreux autres domaines de l'analyse et de l'algèbre.

2.1 Le plan et l'espace euclidiens

2.1.1 Qu'est-ce qu'un vecteur ?

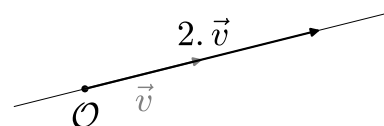
Dans le plan \mathcal{P} ou dans l'espace \mathcal{E} , muni d'un point d'origine \mathcal{O} , un vecteur, c'est :

- une direction,
- un sens,
- une longueur.

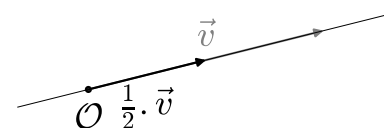


Par ailleurs, il existe plusieurs opération que l'on peut faire à partir des vecteurs du plan. Ainsi, étant donné un vecteur du plan (ou de l'espace), on peut

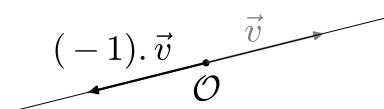
— l'agrandir en le multipliant par une constante > 1 ,



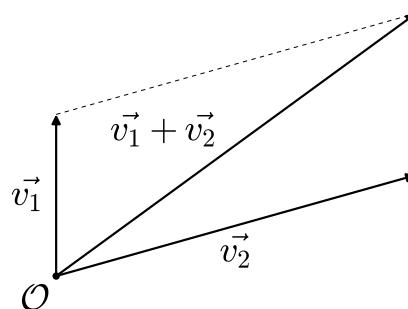
— le retressir en le multipliant par une constante < 1 ,



— le retourner en le multipliant par -1 .



D'autre part, étant donnés deux vecteurs du plan (ou de l'espace), on peut les additionner pour construire un troisième vecteur du plan, en les mettant bout à bout.



Formellement, on peut résumer l'ensemble de ces transformations en deux opérations sur l'ensemble des vecteurs :

— L'addition des vecteurs (“+”) :

$$(\vec{v}_1, \vec{v}_2) \mapsto \vec{v}_1 + \vec{v}_2$$

— La multiplication par un scalaire $\lambda \in \mathbb{R}$ (“.”) :

$$\vec{v} \mapsto \lambda \cdot \vec{v}$$

Note : il est possible de rassembler l'ensemble de ces transformations en une seule opération, appelée *combinaison linéaire*. Précisément, étant donnés deux vecteurs \vec{u} et \vec{v} et deux scalaires $\lambda, \mu \in \mathbb{R}$, on appelle *combinaison linéaire* des vecteurs \vec{u} et \vec{v} , de coefficients λ et μ le vecteur

$$\vec{w} = \lambda \cdot \vec{u} + \mu \cdot \vec{v}$$

2.1.2 Points et vecteurs

Comme cela a été évoqué plus haut, le plan et l'espace sont les ensembles les plus naturellement attachés à la notion d'espace vectoriels. Constitués au départ d'un ensemble de points, on va voir comment, via la notion de vecteurs évoquée plus haut, il est possible de définir *une structure* sur ces ensembles, permettant de mettre en place des outils d'étude efficaces pour les problèmes de géométrie.

Précisément, la première étape consiste à choisir un point \mathcal{O} arbitrairement. Ce sera *le point d'origine* de notre ensemble.

À partir de ce point, il est alors possible d'associer n'importe quel point M du plan ou de l'espace le vecteur $\overrightarrow{\mathcal{O}M}$ et d'identifier ainsi l'ensemble des points du plan ou de l'espace aux vecteurs ayant pour origine le point \mathcal{O} et d'étendre ainsi les opérations sur les vecteurs aux points du plan ou de l'espace.

Précisément,

- Pour tout point M du plan ou de l'espace et pour tout réel $k \in \mathbb{R}$, on notera $k.M$ le point associé au vecteur $k \cdot \overrightarrow{\mathcal{O}M}$.
- Pour tous points M_1 et M_2 du plan ou de l'espace, on notera $M_1 + M_2$ le point associé au vecteur $\overrightarrow{\mathcal{O}M_1} + \overrightarrow{\mathcal{O}M_2}$.

2.1.3 Repères et coordonnées

C'est grâce à la correspondance points - vecteurs évoquée plus haut que l'on peut mettre en place la notion de repère qui, en géométrie, permet de transformer tout problème géométrique en un problème analytique.

Précisément, dans le plan muni d'un point d'origine \mathcal{O} , on fixe deux vecteurs \vec{i} et \vec{j} (le couple de vecteur choisi est appelé *base du repère*). On note alors $\mathcal{R} = (\mathcal{O} ; \vec{i}, \vec{j})$ le repère ainsi constitué et il devient possible d'exprimer n'importe quel vecteur \vec{v} du plan (et donc n'importe quel point M du plan) en fonction des vecteurs \vec{i} et \vec{j} , à l'aide des deux opérations "+" et ".", i.e. sous forme de combinaisons linéaires des vecteurs \vec{i} et \vec{j} .

Précisément,

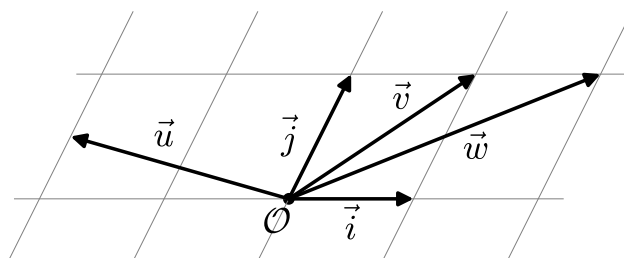
Si $\mathcal{R} = (\mathcal{O} ; \vec{i}, \vec{j})$ est un repère du plan, pour tout vecteur \vec{u} du plan, il existe un unique couple $(x, y) \in \mathbb{R}^2$ tel que

$$\vec{u} = x \cdot \vec{i} + y \cdot \vec{j}$$

Le couple (x, y) forme *les coordonnées de \vec{u} dans le repère \mathcal{R}* et l'on note

$$\vec{u} = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{R}}$$

Exemples : Exprimer en fonction de \vec{i} et \vec{j} les trois vecteurs \vec{u} , \vec{v} et \vec{w} ci dessous :



$$\vec{u} = -2\vec{i} + \frac{1}{2}\vec{j} = \begin{pmatrix} -2 \\ \frac{1}{2} \end{pmatrix}_{\mathcal{R}}, \quad \vec{v} = \vec{i} + \vec{j} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\mathcal{R}}, \quad \vec{w} = 2\vec{i} + \vec{j} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}_{\mathcal{R}}$$

Notes :

- Via la dualité points - vecteurs évoquée plus haut, il est possible d'étendre la notion de coordonnées aux points du plan. Il devient alors possible de représenter la position d'un point M du plan (par rapport à l'origine \mathcal{O}) à l'aide de nombres.

- Pour construire un repère du plan, on peut prendre n'importe quel couple de vecteurs, à une seule condition près : les deux vecteurs \vec{i} et \vec{j} d'un repère ne peuvent être pris colinéaires. En effet, si \vec{i} et \vec{j} sont colinéaires, seuls les vecteurs de la droite donnant la direction commune à \vec{i} et \vec{j} s'expriment sous la forme

$$\lambda \cdot \vec{i} + \mu \cdot \vec{j}, \quad \lambda, \mu \in \mathbb{R}$$

Exercice : à démontrer. On rappelle que deux vecteurs \vec{i} et \vec{j} sont colinéaires si et seulement si

$$\exists k \in \mathbb{R} / \vec{i} = k \cdot \vec{j}$$

De même, on appelle *repère de l'espace* \mathcal{E} la donnée d'un point d'origine \mathcal{O} et de trois vecteurs \vec{i} , \vec{j} , \vec{k} à partir desquels il est possible d'exprimer l'ensemble des points de l'espace sous forme de combinaisons linéaires : pour tout vecteur $\vec{v} \in \mathcal{E}$, il existe un unique triplet $(x, y, z) \in \mathbb{R}^3$ tel que

$$\vec{v} = x \cdot \vec{i} + y \cdot \vec{j} + z \cdot \vec{k}$$

Le triplet (x, y, z) forme les *coordonnées* du vecteur \vec{v} dans le repère $\mathcal{R} = (\mathcal{O}; \vec{i}, \vec{j}, \vec{k})$ et l'on note

$$\vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\mathcal{R}_2}$$

Note : comme dans le plan, il faut bien choisir les vecteurs de l'espace si l'on veut pouvoir atteindre tous les vecteurs de l'espace. Précisément, pour former un repère de l'espace, trois vecteurs \vec{i} , \vec{j} et \vec{k} ils ne doivent pas être co-planaires. En effet, s'ils sont dans un même plan, les seuls vecteurs que l'on pourra exprimer en fonction de \vec{i} , \vec{j} et \vec{k} sont les vecteur du plan les rassemblant.

2.1.4 Coordonnées et opérations

Outre la représentation d'une position dans le plan ou dans l'espace à l'aide de valeurs numériques, la notion de coordonnées (intimement liée à la notion de repère) permet également de traduire l'ensemble des problèmes géométriques (positions relatives, déplacements, etc) en équations numériques parce qu'elle (la notion de coordonnées) est compatible avec les deux opérations définies sur l'ensemble des vecteurs. Précisément, si l'on se fixe un repère $\mathcal{R} = (\mathcal{O}; \vec{i}, \vec{j})$ du plan, alors

Propriétés

- Soit $\vec{v} = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{R}}$ un vecteur du plan.

$$\forall \lambda \in \mathbb{R}, \quad \lambda \cdot \vec{v} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}_{\mathcal{R}}$$

- Soient $\vec{u} = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{R}}$ et $\vec{v} = \begin{pmatrix} x' \\ y' \end{pmatrix}_{\mathcal{R}}$ deux vecteurs du plans.

$$\vec{u} + \vec{v} = \begin{pmatrix} x + x' \\ y + y' \end{pmatrix}_{\mathcal{R}}$$

Exercice : démontrer ces résultats à l'aide de la définition de la notion de coordonnées.

Note : en ajoutant une troisième coordonnée, il est naturellement possible d'étendre ces notions aux points (et vecteurs) de l'espace.

2.2 Définitions et Exemples

Le premier objectif de ce cours est de formaliser et étudier les objets que l'on vient de voir et d'établir certaines propriétés générales. En effet, une fois que l'on a établi l'ensemble des propriétés qui font la spécificité de l'ensemble des points du plan ou de l'espace, on pourra étendre ces notions à tout ensemble d'objet mathématique partageant ces mêmes propriétés. Il devient alors possible de faire de la géométrie dans ces ensembles à l'image de ce que l'on sait faire dans le plan ou dans l'espace.

2.2.1 Loi de composition interne

Commençons par formaliser l'addition de deux vecteurs. De façon formelle, l'addition est une opération qui combine deux vecteurs du plan (ou de l'espace) pour en construire un troisième. C'est ce que l'on appelle une loi de composition interne (ici). On peut en outre remarquer que l'addition de deux vecteurs que l'on vient de voir possède quelques propriétés élémentaires. Ainsi, en notant \mathcal{P} l'ensemble des vecteurs du plan ou de l'espace munis d'un point d'origine \mathcal{O} , on a :

Propriétés

- L'addition des vecteurs est commutative :

$$\forall \vec{u}, \vec{v}, \quad \vec{u} + \vec{v} = \vec{v} + \vec{u}$$

- L'addition des vecteurs est associative :

$$\forall \vec{u}, \vec{v}, \vec{w}, (\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$$

- L'addition des vecteurs possède un élément neutre : le vecteur nul $\vec{0}$ (ou le point \mathcal{O}).

$$\forall \vec{u}, \vec{u} + \vec{0} = \vec{u}$$

- Tout vecteur \vec{v} possède un inverse pour l'addition :

$$\forall \vec{v}, \exists \vec{v}' / \vec{v} + \vec{v}' = \vec{0}$$

Ce vecteur \vec{v}' est alors noté $-\vec{v}$.

Exercice : représenter géométriquement ces différentes propriétés.

2.2.2 Multiplication externe

De la même façon, on peut formaliser la multiplication d'un vecteur par un réel. Cette opération est une multiplication externe, puisque l'on utilise un élément qui n'est pas dans notre ensemble (le réel λ). D'autre part, la multiplication externe que l'on a vu possède elle aussi quelques propriétés :

Propriétés

- La multiplication externe est distributive sur l'addition des vecteurs :

$$\forall \vec{u}, \vec{v} \in \mathcal{P}, \forall \lambda \in \mathbb{R}, \lambda \cdot (\vec{u} + \vec{v}) = \lambda \cdot \vec{u} + \lambda \cdot \vec{v}$$

- La multiplication externe est distributive par rapport à l'addition des réels :

$$\forall u \in \mathcal{P}, \forall \lambda, \mu \in \mathbb{R}, (\lambda + \mu) \cdot \vec{u} = \lambda \cdot \vec{u} + \mu \cdot \vec{u}$$

- La multiplication externe est associative :

$$\forall u \in \mathcal{P}, \forall \lambda, \mu \in \mathbb{R}, (\lambda \mu) \cdot \vec{u} = \lambda \cdot (\mu \cdot \vec{u})$$

- Le réel 1 est l'élément neutre pour la multiplication extérieure :

$$\forall v \in \mathcal{P}, 1 \cdot \vec{v} = \vec{v}$$

Note : notons également que la multiplication externe des vecteurs possède les deux propriétés suivantes : pour tout vecteur \vec{v} , on a

$$0 \cdot \vec{v} = \vec{0} \quad \text{et} \quad (-1) \cdot \vec{v} = -\vec{v}$$

2.2.3 Espace vectoriel

Ce sont ces deux opérations “+” et “.” et leurs propriétés qui font du plan et de l’espace des espaces vectoriels. Précisément,

Définition (Espace vectoriel)

Soit E un ensemble. Cet ensemble est un espace vectoriel si l’on peut définir, pour les éléments de E

- d’une loi commutative, associative et possédant un élément neutre (note $\vec{0}_E$) et telle que tout élément de E soit inversible pour cette loi,
- d’une multiplication externe par des réels qui soit distributive (aux sens où on l’a vu plus haut), associative et possédant également un élément neutre.

Définition (Vecteurs)

On appelle *vecteur* tout élément d’un espace vectoriel.

Définition (Combinaisons linéaires)

Soit $(E, +, \cdot)$ un espace vectoriel. On appelle *combinaison linéaire* de vecteurs de E tout vecteur $\vec{v} \in E$ de la forme

$$\vec{v} = \sum_{j=1}^n k_j \cdot \vec{v}_j = k_1 \cdot \vec{v}_1 + k_2 \cdot \vec{v}_2 + \dots + k_n \cdot \vec{v}_n$$

où les \vec{v}_i sont des vecteurs de E et les k_i sont des réels.

Exemples d'espaces vectoriels.

1. Le plan (resp. l'espace) muni d'un point d'origine et des opérations “+” et “.” évoquées plus haut est un espace vectoriel.
2. La droite munie d'un point d'origine et des mêmes opérations est un espace vectoriel.
3. L'ensemble des polynômes muni des opérations naturelles d'addition et de multiplication par un réel est un espace vectoriel.
4. L'ensemble des matrices d'une même taille, muni des opérations naturelles d'addition et de multiplication par un réel est un espace vectoriel.
5. L'ensemble \mathbb{R}^2 des couples de réels, muni des opérations suivantes est un espace vectoriel :

$$(x, y) + (x', y') = (x + x', y + y')$$

et

$$k.(x, y) = (kx, ky).$$

6. De façon plus général, pour un entier n donné, l'ensemble \mathbb{R}^n des n -uplets à coordonnées dans \mathbb{R} , muni des mêmes opérations (composante pas composante) est un espace vectoriel.
7. L'ensemble des solutions d'une équation différentielle linéaire homogène, muni des opérations adéquates définies sur l'ensemble des fonctions est un espace vectoriel.
8. L'ensemble

$$[0, 1]^2 = \{(x, y) \in \mathbb{R}^2 / 0 \leq x \leq 1 \text{ et } 0 \leq y \leq 1\}$$

muni des opérations usuelles sur les couples de nombres réels n'est pas un espace vectoriel.

9. L'ensemble \mathbb{N} muni des opérations naturelles “+” et “×” n'est pas un espace vectoriel.

2.2.4 Sous espaces vectoriels (SEV)

Plaçons dans un espace vectoriel E , muni d'une loi “+” et d'une multiplication extérieure “.” (par exemple, le plan ou l'espace dans lequel on fixe un point d'origine \mathcal{O}). La notion de sous espace vectorielle fait référence aux sous ensembles de E . Précisément,

Définition (Sous espace vectoriel)

Étant donné un espace vectoriel $(E, +, .)$, on appelle *sous espace vectoriel de E* tout sous ensemble $\mathcal{S} \subset E$ qui, muni des opérations “+” et “.”, est lui-même un espace vectoriel.

Dans ce cadre, on est alors souvent appelés à déterminer si un sous ensemble donné $\mathcal{S} \subset E$ est ou non un sous espace vectoriel de E . Pour répondre à cette question, il est évidemment possible de s'appuyer sur la définition d'un espace vectoriel énoncée plus haut.

Cependant, le fait que l'on travaille déjà dans un espace vectoriel E permet de répondre plus simplement à cette question. Précisément,

Théorème

Étant donné un espace vectoriel E , un sous ensemble \mathcal{S} de E est un sous espace vectoriel de E si et seulement si

1. L'ensemble \mathcal{S} est non vide.

2. \mathcal{S} est fermé pour la loi $+$:

$$\forall \vec{u}, \vec{v} \in \mathcal{S}, \quad \vec{u} + \vec{v} \in \mathcal{S}$$

3. \mathcal{S} est fermé pour la loi \cdot :

$$\forall k \in \mathbb{R}, \forall \vec{u} \in \mathcal{S}, \quad k \cdot \vec{u} \in \mathcal{S}$$

Notes :

— les deux derniers points peuvent être rassemblés dans la propriété suivante : pour tous $k, l \in \mathbb{R}$ et pour tous $\vec{u}, \vec{v} \in \mathcal{S}$, on a

$$k \cdot \vec{u} + l \cdot \vec{v} \in \mathcal{S}.$$

— Le premier point est essentiel car l'ensemble vide vérifie les deux points suivant et n'est pas un espace vectoriel. Il faut donc commencer par montrer que le sous ensemble \mathcal{S} que l'on étudie possède au moins un élément. Cependant, on peut montrer que tout sous espace vectoriel de E contient le vecteur nul $\vec{0}_E$. On peut donc vérifier le premier point en testant le vecteur nul de E . Si celui-ci appartient à \mathcal{S} , alors \mathcal{S} est non vide ; sinon, \mathcal{S} ne peut être un espace vectoriel et il n'est pas nécessaire de tester les autres points.

Exemples de SEV

1. Dans tout espace vectoriel E , le singleton $\{\vec{0}_E\}$ est un sous espace vectoriel de E .
2. Tout espace vectoriel E est un sous-espace vectoriel de lui même.
3. Dans le plan muni d'un point d'origine \mathcal{O} , toute droite passant par \mathcal{O} est un sous espace vectoriel du plan.

Exercice : à démontrer en notant que toute droite D du plan peut être décrit par

$$D = \{\lambda \cdot \vec{v}, \lambda \in \mathbb{R}\}$$

où \vec{v} est un vecteur directeur de D .

4. Une droite du plan ne passant pas par l'origine n'est pas un sous espace vectoriel du plan. Pourquoi ?
5. De façon générale, les sous espaces vectoriels du plan muni d'un point d'origine \mathcal{O} sont :
 - le point \mathcal{O} ,
 - les droites passant par \mathcal{O} ,
 - le plan tout entier.
6. Dans l'espace muni d'un point \mathcal{O} , les sous espaces vectoriels sont
 - le point \mathcal{O} ,
 - les droites passant par \mathcal{O} ,
 - les plans passant par \mathcal{O} ,
 - l'espace tout entier.

7. Dans \mathbb{R}^3 l'ensemble

$$\mathcal{S} = \{(a, b, 0), a, b \in \mathbb{R}\}$$

est un sous espace vectoriel de \mathbb{R}^3 (à démontrer). Par contre, l'ensemble

$$\mathcal{S}' = \{(a, b, 1), a, b \in \mathbb{R}\}$$

n'est pas un SEV de \mathbb{R}^3 . Pourquoi ?

8. L'ensemble des solutions d'une équation *linéaire et homogène* à n inconnues

$$(E) : a_1.x_1 + a_2.x_2 + \dots + a_n.x_n = 0$$

est un SEV de \mathbb{R}^n .

Note : si $n = 2$ ou $n = 3$, cet exemple est à rapprocher des premiers exemples. En effet, si $n = 2$, l'ensemble des solutions de l'équation $ax + by = 0$ peut être représenté par une droite du plan muni d'un repère, passant par l'origine du repère. De même, l'ensemble des solutions d'une équation de la forme $ax + by + cz = 0$ peut être représenté par un plan de l'espace muni d'un repère, passant par l'origine du repère.

9. L'ensemble des solutions d'un système linéaire homogène portant sur n inconnues est un sous espace vectoriel de \mathbb{R}^n . Là encore, si $n = 2$ ou $n = 3$, une approche géométrique permet de rapprocher cet exemple des premiers exemples. D'autre part, l'approche matricielle permet, elle, de démontrer cette propriété quelque soit

le nombre d'inconnues du système.

10. L'ensemble des polynôme d'un même degré fixé d n'est pas un sous espace vectoriel de l'ensemble des polynômes (pourquoi?). Cependant, on verra en TD quels sont les sous espace vectoriels classiques de l'espace vectoriel des polynômes.

2.2.5 Sommes de sous espaces vectoriels

Soit E un espace vectoriel et soient \mathcal{S}_1 et \mathcal{S}_2 deux sous espaces vectoriels de E . À partir de ces deux sous espaces et à l'aide de la loi “+” que l'on a défini sur E , on peut construire un nouveau sous ensemble de E : la somme $\mathcal{S}_1 + \mathcal{S}_2$ contenant tous les vecteurs $\vec{v} \in E$ qui s'écrivent comme la somme d'un élément de \mathcal{S}_1 et 'un élément de \mathcal{S}_2 :

$$\mathcal{S}_1 + \mathcal{S}_2 = \{v = s_1 + s_2, s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2\}$$

Ainsi défini, l'ensemble $\mathcal{S}_1 + \mathcal{S}_2$ est clairement un sous ensemble de E (pourquoi?) mais on peut également montrer que

Proposition

Soit E un espace vectoriel. La somme de deux sous espaces vectoriels de E est encore un sous espace vectoriel E .

Exercice : à démontrer en s'appuyant sur le théorème de la section 2.2.4, sur la définition de la somme $\mathcal{S}_1 + \mathcal{S}_2$ et sur le fait que \mathcal{S}_1 et \mathcal{S}_2 sont des espaces vectoriels.

Exemples :

- Dans le plan, la somme de deux droites linéaires distinctes est égale au plan tout entier.
- Dans l'espace, la somme de deux droites linéaires distinctes est égale à l'unique plan contenant ces deux droites.
- Dans l'espace, la somme d'un plan linéaire et d'une linéaire n'appartenant pas à ce plan et est égale à l'espace tout entier. Si la droite est dans le plan, la somme est égale au plan.
- Dans l'espace, la somme de deux plans linéaires distincts est encore égale à l'espace tout entier.

Enfin, en s'appuyant sur les exemples ci-dessus, on peut distinguer différents cas lorsque l'on additionne des sous espaces vectoriels :

Définition

Soient E un espace vectoriel et \mathcal{S}_1 et \mathcal{S}_2 deux sous espaces vectoriels de E .

- La somme $\mathcal{S}_1 + \mathcal{S}_2$ sera dite directe si

$$\mathcal{S}_1 \cap \mathcal{S}_2 = \{0_E\}.$$

On note alors $\mathcal{S}_1 \oplus \mathcal{S}_2$.

Note : l'intersection de deux sous espaces vectoriels contient toujours le vecteur nul. La somme est donc directe si cette intersection est minimale.

- Les sous espaces vectoriels \mathcal{S}_1 et \mathcal{S}_2 sont dit *supplémentaires* dans E si ils sont en somme directe et si

$$E = \mathcal{S}_1 \oplus \mathcal{S}_2$$

Et l'on peut énoncer la proposition suivante :

Proposition

Si \mathcal{S}_1 et \mathcal{S}_2 sont deux sous espaces vectoriels supplémentaires dans E , alors tout vecteur $\vec{v} \in E$ s'écrit de façon unique sous la forme

$$v = s_1 + s_2, \quad s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2$$

Exemples :

- Dans l'espace, deux droite linéaires distinctes sont en somme directe.
- Dans le plan, deux droites linéaires distinctes sont supplémentaires.
- Dans l'espace, un plan passant par \mathcal{O} et une droite passant par \mathcal{O} , extérieure à ce plan sont supplémentaires.
- Dans l'espace, deux plan ne sont jamais en somme directe.

2.3 Familles de vecteurs

Dans cette section, nous allons revenir en détails sur la notion de repère que l'on a vu plus haut. Intuitivement, il s'agit d'étudier le fait que dans un espace vectoriel, il est possible de trouver une famille de vecteurs (souvent en nombre fini) qui nous permette de reconstruire tous les vecteurs de l'espace via les combinaisons linéaires.

Pour définir proprement la notion de repère, il faut étudier en détails les différents types de familles que l'on construit en "piochant" quelques vecteurs dans même espace vectoriel.

Enfin, une fois établie la définition d'un repère, nous verrons comment définir proprement la notion de *dimension* qui donne une mesure de la taille d'un espace vectoriel.

2.3.1 Familles génératrices

Définitions (sous espace engendré et famille génératrice)

1. Soient $(E, +, \cdot)$ un espace vectoriel et $\mathcal{F} = \{v_1, \dots, v_n\}$ une famille de n vecteurs de E . On appelle *sous espace engendré par la famille \mathcal{F}* le sous ensemble $\text{Vect}(\mathcal{F})$ de E formé de l'ensemble des vecteurs de E obtenus par combinaisons linéaires des vecteurs de \mathcal{F} :

$$\text{Vect}(\mathcal{F}) = \{\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \dots + \lambda_n \cdot v_n, \lambda_k \in \mathbb{R}\} = \left\{ \sum_{k=1}^n \lambda_k \cdot v_k, \lambda_k \in \mathbb{R} \right\}$$

2. Soit $(E, +, \cdot)$ un espace vectoriel. On appelle *famille génératrice de E* toute famille \mathcal{F} de vecteurs de E telle que

$$\text{Vect}(\mathcal{F}) = E$$

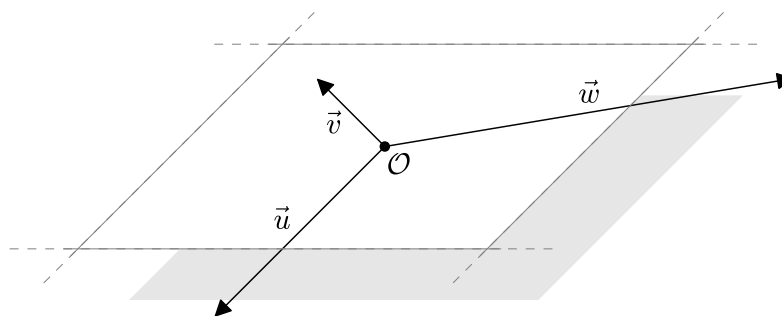
Exemples :

- Dans le plan muni d'un point d'origine \mathcal{O} , toute famille de vecteurs contenant au moins deux vecteurs non colinéaires engendre le plan.

- Dans ce même plan, tout vecteur $\vec{v} \neq \vec{0}$ engendre la droite

$$D_{\vec{v}} = \{\lambda \cdot \vec{v}, \lambda \in \mathbb{R}\}$$

- Sur le dessin ci-dessous, les trois vecteurs \vec{u} , \vec{v} et \vec{w} sont coplanaires. La famille $\mathcal{F} = \{\vec{u}, \vec{v}, \vec{w}\}$ engendrent le plan \mathcal{P} les contenant tous les trois, mais n'est pas une famille génératrice de l'espace.



- Dans \mathbb{R}^2 ,
- le sous espace engendré par le vecteur $(1, -1)$ est

$$\text{Vect}(\{(1, -1)\}) = \{\lambda \cdot (1, -1), \lambda \in \mathbb{R}\} = \{(\lambda, -\lambda), \lambda \in \mathbb{R}\}$$

- le sous espace engendré par la famille $\mathcal{F} = \{(1, 0), (1, 1)\}$ est

$$\text{Vect}(\{(1, 0), (1, 1)\}) = \{\lambda \cdot (1, 0) + \mu \cdot (1, 1), \lambda, \mu \in \mathbb{R}\} = \mathbb{R}^2$$

- le sous espace engendré par la famille $\mathcal{F} = \{(1, 2), (2, 4)\}$ est

$$\text{Vect}(\{(1, 2), (2, 4)\}) = \{(\lambda, 2\lambda), \lambda \in \mathbb{R}\}$$

- Dans tout espace vectoriel E , le sous espace engendré par le vecteur nul est

$$\text{Vect}(\{0_E\}) = \{0_E\}$$

Par ailleurs, on peut montrer que pour toute famille de vecteur \mathcal{F} , l'ensemble $\text{Vect}(\mathcal{F})$ est un sous espace vectoriel de E . Précisément,

Proposition

Soient $(E, +, \cdot)$ un espace vectoriel.

Pour toute famille \mathcal{F} de vecteurs de E , le sous ensemble $\text{Vect}(\mathcal{F})$ est un sous espace vectoriel de E .

Par ailleurs, tout sous espace vectoriel de E contenant l'ensemble des vecteurs de la famille \mathcal{F} contient également $\text{Vect}(\mathcal{F})$.

Exercice : démontrer les deux points ci-dessus.

On peut enfin énoncer la proposition suivante :

Proposition

Soient $(E, +, \cdot)$ un espace vectoriel. Toute famille de vecteurs de E contenant une famille génératrice de E engendre E .

Exercice : à démontrer.

2.3.2 Familles libres

Il est clair que la notion de famille génératrice que l'on vient de voir est essentiel dans la notion de repère. Pour qu'une famille de vecteurs de E forme un repère de E , elle doit engendrer E tout entier. Cependant, la notion de repère est plus fine que cela. Précisément, la dernière proposition énoncée ci-dessus traduit le fait que, si l'on part d'une famille génératrice de E (i.e. une famille de vecteur de E dans laquelle toutes les directions de E sont représentées), tout vecteur que l'on ajoutera à notre famille génératrice n'apportera pas d'information supplémentaire. Cette redondance d'information indique que la famille augmentée n'est pas optimale et cette famille augmentée ne peut être la base d'un repère de E .

Formellement, cette redondance est caractérisée par une *la dépendance linéaire* présente au sein de la famille. Précisément,

Définition (famille liée, famille libre)

Soient E un espace vectoriel et $\mathcal{F} = \{v_1, \dots, v_n\}$ une famille de vecteurs de E . La famille \mathcal{F} sera dite *liée* (ou *linéairement dépendante*) si l'un vecteur de \mathcal{F} peut s'écrire comme combinaison linéaire des autres.

Par ailleurs, une famille de vecteurs qui n'est pas liée sera dite libre (ou linéairement indépendante).

Intuitivement, une famille est libre si chaque vecteur apporte une "direction" supplémentaire. Précisément, si $\mathcal{F} = \{\vec{v}_1, \dots, \vec{v}_n\}$ est une famille liée, l'un des vecteurs (par exemple v_1) peut s'exprimer en fonction des autres :

$$\exists \lambda_2, \dots, \lambda_n \in \mathbb{R} / v_1 = \lambda_2 \cdot v_2 + \dots + \lambda_n \cdot v_n = \sum_{k=2}^n \lambda_k \cdot v_k$$

Mais alors, tout vecteur de E qui s'écrit comme combinaison linéaire des vecteurs de \mathcal{F} peut s'exprimer *sans* le vecteur v_1 :

$$\begin{aligned}
 v &= \mu_1 \cdot v_1 + \mu_2 \cdot v_2 + \dots + \mu_n \cdot v_n = \sum_{k=1}^n \mu_k \cdot v_k \\
 &= \mu_1 \cdot (\lambda_2 \cdot v_2 + \dots + \lambda_n \cdot v_n) + \mu_2 \cdot v_2 + \dots + \mu_n \cdot v_n = \mu_1 \cdot \sum_{k=2}^n \lambda_k \cdot v_k + \sum_{k=2}^n \mu_k \cdot v_k \\
 &= (\mu_1 \lambda_2 + \mu_2) v_2 + \dots + (\mu_1 \lambda_n + \mu_n) v_n = \sum_{k=2}^n (\mu_1 \lambda_k + \mu_k) \cdot v_k
 \end{aligned}$$

Autrement dit, si le vecteur v_1 est lié aux autres vecteurs de \mathcal{F} , on a

$$\text{Vect}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n) = \text{Vect}(\vec{v}_2, \dots, \vec{v}_n).$$

Du point de vue de la construction d'un repère de E , les familles liées ont un inconvénient majeur : si \mathcal{F} est une famille liée qui engendre E , tout vecteur de E s'écrit bien comme une combinaison linéaire des vecteurs de \mathcal{F} , mais cette écriture n'est pas unique !

Autrement dit, si l'on prend pour repère une famille de vecteurs liés, tout vecteur pourra être représenté par plusieurs jeux de coordonnées distincts, ce qui n'est pas acceptable.

Si l'on souhaite construire un repère de E , il est donc nécessaire de pouvoir déterminer par le calcul si une famille donnée est liée ou non. On peut pour cela se baser sur la proposition suivante :

Proposition

Soit E un espace vectoriel. Une famille $\mathcal{F} = \{v_1, \dots, v_n\}$ de vecteurs de E est liée si et seulement si il existe des réels $\lambda_1, \dots, \lambda_n$ non tous nuls tels que

$$\sum_{k=1}^n \lambda_k \cdot v_k = 0_E$$

ou plus précisément sur sa réciproque :

Proposition

Soient E un espace vectoriel et $\mathcal{F} = \{v_1, \dots, v_n\}$ une famille de vecteurs de E . La famille \mathcal{F} est libre si et seulement si *la seule* façon d'exprimer le vecteur nul de E comme combinaison linéaire des vecteurs de \mathcal{F} est de prendre tous les coefficients nuls. Autrement dit, \mathcal{F} est libre si

$$\sum_{k=1}^n \lambda_k \cdot v_k = 0_E \quad \Rightarrow \quad \lambda_1 = \lambda_2 = \dots = \lambda_n = 0$$

Exemples :

1. Dans tout espace vectoriel E ,
 - Toute famille constituée d'un seul vecteur non nul forme une famille libre.
 - Deux vecteurs non proportionnels forment une famille liée si et seulement s'ils sont colinéaires.
 - Toute famille contenant le vecteur nul est liée.
2. Dans le plan muni d'un point d'origine, toute famille de trois vecteurs ou plus est liée.
3. Dans l'espace muni d'un point d'origine, toute famille de quatre vecteurs ou plus est liée, mais trois vecteurs non coplanaires forment une famille libre.
4. Dans \mathbb{R}^3 , la famille formée des vecteurs

$$v_1 = (1, 0, 0), \quad v_2 = (1, 1, 0), \quad v_3 = (0, 0, 1)$$

est libre.

$$\alpha \cdot v_1 + \beta \cdot v_2 + \gamma \cdot v_3 = 0_{\mathbb{R}^3} = (0, 0, 0)$$

Via la définition des opérations “+” et “.” qui font de \mathbb{R}^3 un espace vectoriel, on peut déterminer explicitement le vecteur $\alpha \cdot v_1 + \beta \cdot v_2 + \gamma \cdot v_3$:

$$\begin{aligned} \alpha \cdot v_1 + \beta \cdot v_2 + \gamma \cdot v_3 &= \alpha \cdot (1, 0, 0) + \beta \cdot (1, 1, 0) + \gamma \cdot (0, 0, 1) \\ &= (\alpha, 0, 0) + (\beta, \beta, 0) + (0, 0, \gamma) \\ &= (\alpha + \beta, \beta, \gamma) \end{aligned}$$

Mais alors, par identification,

$$\begin{aligned}
 \alpha.v_1 + \beta.v_2 + \gamma.v_3 = (0, 0, 0) &\Leftrightarrow (\alpha + \beta, \beta, \gamma) = (0, 0, 0) \\
 &\Leftrightarrow \begin{cases} \alpha + \beta = 0 \\ \beta = 0 \\ \gamma = 0 \end{cases} \\
 &\Leftrightarrow \alpha = \beta = \gamma = 0
 \end{aligned}$$

Note : on a maintenant tous les outils pour démontrer la propriété recherchée : si \mathcal{F} est une famille libre d'éléments de E , alors toute combinaison linéaire des vecteurs de \mathcal{F} est unique : si $\mathcal{F} = \{v_1, \dots, v_n\}$ est libre, alors

$$v = \lambda_1.v_1 + \dots + \lambda_n.v_n = \mu_1.v_1 + \dots + \mu_n.v_n \Rightarrow \lambda_i = \mu_i \quad \forall i \in \{1, \dots, n\}$$

Exercice : à faire à l'aide de la définition de la notion de famille libre.

Pour finir, on peut énoncer une dernière propriété concernant les familles libres :

Proposition

Soit E un espace vectoriel. Toute famille *contenue dans une famille libre de E* est libre.

2.3.3 Bases et coordonnées

Définition (base d'un espace vectoriel)

Soit E un espace vectoriel. On appelle *base de E* toute famille libre de E qui engendre E .

Cette notion de base permet de généraliser la notion de repère à n'importe quel espace vectoriel. En effet, si \mathcal{B} est une base de E , alors

- Puisque \mathcal{B} engendre E , tout vecteur de E peut s'écrire comme une combinaison linéaire des vecteurs de \mathcal{B} .
- Puisque \mathcal{B} est libre, cette écriture est unique.

Ainsi, chaque vecteur $v \in E$ est parfaitement déterminé par les coefficients de l'unique combinaison linéaire des vecteurs de \mathcal{B} qui lui correspond. Ces coefficients sont *les coordonnées de v dans la base \mathcal{B}* .

Exemples

1. Dans le plan muni d'un point d'origine, tout couple de deux vecteurs non colinéaires forme une base du plan.
2. Dans l'espace muni d'un point d'origine, toute famille de trois vecteurs non coplanaires est une base de l'espace.
3. La famille $\mathcal{B} = \{(1, 0), (1, 1)\}$ est une base de \mathbb{R}^2 . Ainsi, le vecteur $(3, 2)$ s'écrit

$$(3, 2) = (1, 0) + 2 \cdot (1, 1)$$

Les coordonnées de ce vecteur dans la base \mathcal{B} sont donc

$$(3, 2) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}_{\mathcal{B}}$$

De façon générale, on peut montrer que les coordonnées d'un vecteur (x, y) de \mathbb{R}^2 dans la base \mathcal{B} sont

$$(x, y) = \begin{pmatrix} x - y \\ y \end{pmatrix}_{\mathcal{B}}$$

4. Soit n un entier fixé. Pour tout $i \in \{1, \dots, n\}$, on note \vec{e}_i le vecteur de \mathbb{R}^n dont toutes les coordonnées sont nulles, sauf la i -ième.

$$\vec{e}_i = (0, \dots, 0, \underset{\substack{\uparrow \\ i\text{ième}}}{1}, 0, \dots, 0).$$

La famille $\mathcal{B}_n = \{\vec{e}_1, \dots, \vec{e}_n\}$ est alors une base de \mathbb{R}^n , appelée la base canonique de \mathbb{R}^n car les coordonnées de n'importe quel vecteur (x_1, \dots, x_n) de \mathbb{R}^n dans la base \mathcal{B}_n sont

$$(x_1, \dots, x_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}_n}$$

2.3.4 Espaces vectoriels de dimension finie**Définition**

Outre la notion de repère, la notion de base définie ci-dessus met en évidence la notion de dimension. Précisément,

Définition (espace vectoriel de dimension finie)

Un espace vectoriel E est dit *de dimension finie* si il est engendré par une famille finie de vecteurs.

Théorème/Définition (dimension d'un espace vectoriel)

Soit E un espace vectoriel de dimension finie. Il existe un entier $d \geq 0$ tel que

1. toute famille génératrice de E possède au moins d vecteurs,
2. toute famille libre de E possède au plus d vecteurs,
3. toute base de E est composée de d vecteurs exactement.

Cet entier d est appelé la dimension de E . On note $\dim(E)$.

Exemples :

- Le plan muni d'un point d'origine est de dimension 2.
- L'espace muni d'un point d'origine est de dimension 3.
- L'espace \mathbb{R}^2 est de dimension 2.
- De façon générale, l'espace \mathbb{R}^n est de dimension n .
- L'ensemble des solutions d'une équation différentielle d'ordre n , linéaire et homogène est de dimension n .

On peut également établir les résultats suivants :

- L'ensemble $\mathcal{M}_{np}(\mathbb{R})$ des matrices à n lignes et p colonnes est un espace vectoriel de dimension $n \times p$.
- L'ensemble $\mathbb{R}_n[X]$ des polynômes de degré inférieur ou égal à n est de dimension $n + 1$.
- L'ensemble $\mathcal{F}(\mathbb{R}, \mathbb{R})$ des fonctions de \mathbb{R} dans \mathbb{R} n'est pas un espace vectoriel de dimension finie. On dit qu'il est de dimension infinie.
- L'ensemble $\mathbb{R}[X]$ des polynômes (de degré quelconque) est également un espace de dimension infinie.

Exercice : conjecturer des bases possibles pour les espaces vectoriels $\mathcal{M}_{n,p}(\mathbb{R})$ et $\mathbb{R}[X]$.

Sous espaces vectoriels en dimension finie

De façon générale, la structure d'un espace vectoriel E influence directement la structure de tous les sous espaces vectoriels qu'il contient. C'est en particulier le cas pour les espaces vectoriels de dimension finie. Précisément,

Proposition

Soit E un espace vectoriel de dimension finie. Pour tout sous espace vectoriel $\mathcal{S} \subset E$,

1. \mathcal{S} est un espace vectoriel de dimension finie et

$$\dim(\mathcal{S}) \leq \dim(E)$$

2. Si $\dim(\mathcal{S}) = \dim(E)$, alors $\mathcal{S} = E$.

Par ailleurs, si \mathcal{S}_1 et \mathcal{S}_2 sont deux sous espaces vectoriels de E , on peut également relier la dimension du sous espace $\mathcal{S}_1 + \mathcal{S}_2$ à celles de \mathcal{S}_1 et \mathcal{S}_2 . Précisément,

Proposition

Soient E un espace vectoriel et \mathcal{S}_1 et \mathcal{S}_2 deux sous espaces vectoriels de E de dimensions finies. On a

$$\dim(\mathcal{S}_1 + \mathcal{S}_2) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) - \dim(\mathcal{S}_1 \cap \mathcal{S}_2)$$

En particulier, si \mathcal{S}_1 et \mathcal{S}_2 sont en somme directe, alors

$$\dim(\mathcal{S}_1 \oplus \mathcal{S}_2) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2)$$

Note : cette dernière proposition permet également de montrer que, dans un espace vectoriel de dimension finie, deux sous espaces vectoriels “trop gros” ne peuvent pas être en somme directe. Précisément, si $\dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) > \dim(E)$, alors $\mathcal{S}_1 + \mathcal{S}_2$ contient au moins une droite de E .

Rang d'une famille de vecteurs

Dans tout espace vectoriel E , il est également possible d'établir un lien fort entre une famille (finie) de vecteurs de E et la dimension du sous espace qu'elle engendre. Précisément,

Théorème/Définition (rang d'une famille de vecteurs)

Soient E un espace vectoriel et $\mathcal{F} = \{v_1, \dots, v_n\}$ une famille finie de vecteurs de E . L'espace vectoriel $\text{Vect}(\mathcal{F})$ est de dimension finie, appelé **le rang de \mathcal{F}** , noté $\text{rg}(\mathcal{F})$:

$$\text{rg}(\mathcal{F}) = \dim(\text{Vect}(\mathcal{F}))$$

On a en outre les propriétés suivantes :

Proposition

Soit E un espace vectoriel et $\mathcal{F} = \{v_1, \dots, v_n\}$ une famille de n vecteurs de E . Alors

1. Le rang de la famille \mathcal{F} vérifie $\text{rg}(\mathcal{F}) \leq n$
2. La famille \mathcal{F} est *libre* si et seulement si $\text{rg}(\mathcal{F}) = n$
3. La famille \mathcal{F} est *liée* si et seulement si $\text{rg}(\mathcal{F}) < n$

Théorème de la base incomplète

Les propositions énoncées ci-dessus offrent des outils permettant de construire rapidement des bases pour des espaces vectoriels dont on connaît la dimension. En particulier, puisque le plan muni d'un repère est de dimension 2, toute famille de deux vecteurs non colinéaires est une base. De même, dans l'espace muni d'un point d'origine, toute famille de trois vecteurs non coplanaires est une base de l'espace.

De façon générale, dans un espace vectoriel de dimension d , toute famille libre de d vecteurs constitue une base.

De même, on peut montrer que, dans un espace vectoriel de dimension finie d , toute famille de $d + 1$ vecteurs ou plus est nécessairement liée.

Enfin, d'un point de vue plus formel, on peut énoncer le résultat suivant :

Théorème (de la base incomplète)

Soit E un espace vectoriel de dimension finie.

1. Toute famille libre peut être complétée en une base de E .
2. De toute famille génératrice de E , on peut extraire une base de E .

2.3.5 Matrice d'une famille de vecteurs

Définition

Soit E un espace vectoriel de dimension finie $n \in \mathbb{N}^*$ et $\mathcal{B} = \{e_1, \dots, e_n\}$ une base de E . Par construction, tout vecteur v de E peut être représenté par ses coordonnées dans la base \mathcal{B} :

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}} \in \mathbb{R}^n$$

Via ces notions, il est alors possible de représenter toute famille finie de vecteurs de E par une matrice. Précisément,

Définition (Matrice d'une famille de vecteurs)

Soient E un espace vectoriel de dimension $n \in \mathbb{N}^*$, $\mathcal{B} = \{e_1, \dots, e_n\}$ une base de E et $\mathcal{F} = \{v_1, \dots, v_p\}$ une famille de p vecteurs de E . On appelle *matrice de \mathcal{F} dans la base \mathcal{B}* la matrice à n lignes et p colonnes définie par

$$\text{mat}_{\mathcal{B}}(\mathcal{F}) = \begin{pmatrix} v_1 & v_2 & \cdots & v_p \\ \downarrow & \downarrow & & \downarrow \\ a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix} \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{matrix} \in \mathcal{M}_{np}(\mathbb{R})$$

Autrement dit, la matrice de \mathcal{F} dans \mathcal{B} est la matrice dont les colonnes donnent les coordonnées des vecteurs de \mathcal{F} dans la base \mathcal{B} .

Exemple : dans \mathbb{R}^3 , on considère la base $\mathcal{B} = \{e_1, e_2, e_3\}$ définie par

$$e_1 = (2, 0, 0), \quad e_2 = (0, 1, 0), \quad e_3 = (0, 0, 1)$$

En notant $\mathcal{F} = \{v_1, v_2\}$ où

$$v_1 = (2, 4, 4) \quad \text{et} \quad v_2 = (4, 1, 1)$$

on a

$$v_1 = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix}_{\mathcal{B}} \quad \text{et} \quad v_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}_{\mathcal{B}}$$

et

$$\text{mat}_{\mathcal{B}}(\mathcal{F}) = \begin{pmatrix} 1 & 2 \\ 3 & 1 \\ 4 & 1 \end{pmatrix}$$

Note : la matrice de \mathcal{F} dans la base \mathcal{B} est clairement dépendante de la base \mathcal{B} choisie. Si l'on change de base, il est clair que les coordonnées de chacun des vecteurs de \mathcal{F} changent et donc les coefficients de la matrice changent également. On verra dans la suite qu'il est cependant possible de relier les coordonnées d'un même vecteur exprimé dans deux bases différentes.

Déterminant et famille libre

L'étude d'une matrice M représentant une famille \mathcal{F} dans une base \mathcal{B} permet d'exploiter les propriétés du calcul matriciel pour déterminer certaines propriétés de la famille \mathcal{F} étudiée. On peut en particulier montrer la proposition suivante :

Proposition

Soient E un espace vectoriel de dimension $n \in \mathbb{N}^*$ et \mathcal{B} une base de E . Toute famille \mathcal{F} de n vecteurs de E est libre si et seulement si [la matrice \$\text{mat}_{\mathcal{B}}\(\mathcal{F}\)\$ est inversible](#).

Cette proposition n'a de sens que si la matrice étudiée est carrée, i.e. si le cardinal de la famille \mathcal{F} est égal à la dimension de l'espace E . Il est à noter qu'alors, si la matrice $\text{mat}_{\mathcal{B}}(\mathcal{F})$ est inversible, la famille \mathcal{F} est une (autre) base de E (pourquoi?).

Notons que, de façon plus générale, on peut énoncer la proposition suivante :

Proposition

Soient E un espace vectoriel, \mathcal{B} une base de E et \mathcal{F} une famille de p vecteurs de E (avec $p \leq n$). La famille \mathcal{F} est libre si et seulement si [on peut extraire de \$\text{mat}_{\mathcal{B}}\mathcal{F}\$ une matrice carrée à \$p\$ lignes inversible](#).

Plus généralement, le rang d'une famille \mathcal{F} de vecteurs de E est égal [au nombre de lignes de la plus grande matrice inversible que l'on peut extraire de \$\text{mat}_{\mathcal{B}}\(\mathcal{F}\)\$](#) .

Exemple : montrer que la famille $\mathcal{F} = \{v_1, v_2\}$ de \mathbb{R}^3 étudiée ci-dessus est libre en étudiant sa matrice $\text{mat}_{\mathcal{B}}(\mathcal{F})$.

Matrice de passage et changement de base

Définition (matrice de passage)

Soit E un espace vectoriel de dimension finie et $\mathcal{B}_1 = \{e_1, \dots, e_n\}$ et $\mathcal{B}_2 = \{\varepsilon_1, \dots, \varepsilon_n\}$ deux bases de E . On appelle *matrice de passage de \mathcal{B}_1 à \mathcal{B}_2* la matrice

$$P_{\mathcal{B}_1, \mathcal{B}_2} = \text{mat}_{\mathcal{B}_1}(\mathcal{B}_2) = \begin{pmatrix} \varepsilon_1 & \cdots & \varepsilon_n \\ \downarrow & & \downarrow \\ p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix} \begin{matrix} e_1 \\ \vdots \\ e_n \end{matrix} \in \mathcal{M}_n(\mathbb{R})$$

i.e. la matrice dont les colonnes donnent les coordonnées des vecteurs de \mathcal{B}_2 exprimés dans la base \mathcal{B}_1 .

Note : la matrice de passage entre deux bases de E est toujours inversible (pourquoi). On peut en outre montrer que

$$(P_{\mathcal{B}_1, \mathcal{B}_2})^{-1} = P_{\mathcal{B}_2, \mathcal{B}_1}$$

Par ailleurs, comme cela a été évoqué, un espace vectoriel de dimension finie possède une infinité de bases et les coordonnées d'un même vecteur changent à chaque fois que l'on change de base. Cependant, on peut montrer que la matrice de passage de $P_{\mathcal{B}_1, \mathcal{B}_2}$ permet de faire le lien entre les coordonnées d'un même vecteur respectivement dans les bases \mathcal{B}_1 et \mathcal{B}_2 . Précisément,

Théorème (Changement de base pour un vecteur)

Soit E un espace vectoriel de dimension n et \mathcal{B}_1 et \mathcal{B}_2 deux bases de E . Pour tout vecteur $v \in E$, les vecteurs colonnes

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}_1} \quad \text{et} \quad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}_{\mathcal{B}_2}$$

donnant respectivement les coordonnées de v dans \mathcal{B}_1 et \mathcal{B}_2 sont liés par l'équation

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}_1} = P_{\mathcal{B}_1, \mathcal{B}_2} \times \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}_{\mathcal{B}_2} \iff X = PX'$$

Note : par définition du produit de matrice, cette formule s'applique également aux matrices d'une famille \mathcal{F} de vecteurs de E . Précisément, si \mathcal{B}_1 et \mathcal{B}_2 sont deux bases de E , alors pour toute famille finie \mathcal{F} de vecteurs de E , on a

$$\text{mat}_{\mathcal{B}_1}(\mathcal{F}) = P_{\mathcal{B}_1, \mathcal{B}_2} \times \text{mat}_{\mathcal{B}_2}(\mathcal{F})$$

2.4 Applications linéaires

2.4.1 Définitions

De façon générale, la notion d'espace vectoriel permet de définir une structure particulière sur certains ensembles. Une fois cette structure mise en place, il devient naturel d'étudier les fonctions que l'on peut définir entre deux espaces vectoriels.

Bien qu'il soit formellement possible de construire n'importe quel type de fonction entre deux espaces vectoriels, il est tout aussi naturel de se concentrer sur les fonctions qui *respectent* la structure des espaces vectoriels impliqués : les applications *linéaires*.

D'un point de vue géométrique (i.e. si les espaces vectoriels impliqués sont le plan et/ou l'espace), ces fonctions permettent en particulier de déplacer les points du plan et/ou de l'espace et ainsi de faire bouger points par points différentes figures. Le caractère linéaire des applications que l'on considère permet alors de s'assurer que les figures que l'on déplace ou modifie ne soient pas trop déformées pendant le voyage.

Définitions (morphisme d'espaces vectoriels)

Soient E et F deux espaces vectoriel.

1. On appelle *morphisme de E dans F* , ou application *linéaire* de E dans F toute application $f : E \rightarrow F$ telle que

$$\forall u, v \in E, \forall \lambda, \mu \in \mathbb{R}, \quad f(\lambda.u + \mu.v) = \lambda.f(u) + \mu.f(v)$$

On note alors $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F .

2. On appelle *isomorphisme* de E dans F toute application *bijective* de $\mathcal{L}(E, F)$.
3. On appelle *endomorphisme de E* toute application linéaire de E dans E .
L'ensemble des endomorphismes de E est noté $\mathcal{L}(E)$.
4. On appelle *automorphisme* de E tout endomorphisme *bijectif* de E .

Exemples : de façon plus intuitive, les applications linéaires sont les applications qui

respectent la somme et la multiplication extérieure. Ainsi, si E et F sont deux espaces vectoriel et $f \in \mathcal{L}(E, F)$, alors

$$\forall u, v \in E, \quad f(u + v) = f(u) + f(v)$$

$$\forall u \in E, \forall \lambda \in \mathbb{R}, \quad f(\lambda.u) = \lambda.f(u)$$

Ainsi,

1. Les endomorphismes de $\mathcal{L}(\mathbb{R})$ sont les applications

$$\begin{aligned} f_a &: \mathbb{R} \longrightarrow \mathbb{R} \\ x &\longmapsto ax \end{aligned}$$

pour $a \in \mathbb{R}$ fixé.

2. Les applications

$$\begin{aligned} \varphi &: \mathbb{R} \longrightarrow \mathbb{R} & \psi &: \mathbb{R} \longrightarrow \mathbb{R} \\ x &\longmapsto 2x + 1 & x &\longmapsto x^2 \end{aligned}$$

ne sont pas linéaires.

3. Toute application du type

$$\begin{aligned} g &: \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (ax + by, cx + dy) \end{aligned}$$

est une application linéaire.

4. Dans le plan et l'espace, les rotations, les projections, les symétries et les homothéties sont des applications linéaires.
5. Dans l'espace vectoriel des fonctions, la dérivation est un morphisme sur l'espace vectoriel des fonctions dérivables.

6. Soient m, n, p trois entiers fixés. Pour toute matrice $A \in \mathcal{M}_{mn}(\mathbb{R})$. L'application

$$\begin{aligned} F_A &: \mathcal{M}_{np}(\mathbb{R}) \longrightarrow \mathcal{M}_{mp}(\mathbb{R}) \\ M &\longmapsto A \times M \end{aligned}$$

est une application linéaire.

7. Soient n et p deux entiers fixés. La transposition $M \mapsto {}^t M$ appartient à l'ensemble

$$\mathcal{L}(\mathcal{M}_{np}(\mathbb{R}), \mathcal{M}_{pn}(\mathbb{R}))$$

8. Pour $n \in \mathbb{N}^*$ fixé, la trace $M \mapsto \text{tr}(M)$ est une application linéaire de $\mathcal{M}_n(\mathbb{R})$ dans \mathbb{R} .

Mais le déterminant $M \mapsto \det(M)$ n'est pas linéaire.

Note : l'une des propriétés emblématique des applications linéaires est que toute application linéaire envoie le vecteur nul sur le vecteur nul :

Proposition

Soient E et F deux espaces vectoriels. Pour toute application $f \in \mathcal{L}(E, F)$,

$$f(0_E) = 0_F$$

Exercice : à démontrer.

Par ailleurs, l'ensemble $\mathcal{L}(E, F)$ possède, lui aussi certaines propriétés :

- Pour tous espaces vectoriels E et F , l'espace $\mathcal{L}(E, F)$ peut être muni d'une structure d'espace vectoriel via les opérations “+” et “.” héritées de F .
- Sur $\mathcal{L}(E)$, la composition \circ est une loi de composition interne associative et distributive sur l'addition. Autrement dit,

— Si f et g sont deux endomorphismes de E , la composée $g \circ f$ est encore un endomorphisme de E .

— Pour tous endomorphismes f, g, h de E , on a

$$(f \circ g) \circ h = f \circ (g \circ h)$$

et

$$f \circ (g + h) = f \circ g + f \circ h \quad \text{et} \quad (f + g) \circ h = f \circ h + g \circ h$$

Exercice : à démontrer.

ATTENTION : même pour les applications linéaires, l'opération de composition \circ n'est pas commutative !

2.4.2 Image et noyau d'une application linéaire

Image d'une application linéaire

Soient E et F deux espaces vectoriels et $f \in \mathcal{L}(E, F)$. Le fait que f respecte les combinaisons linéaires entraîne entre autre que f respecte également les sous espaces vectoriels de E . Ainsi, on peut montrer que toute droite de E est (presque toujours) envoyée sur une droite de F . En effet, si D est une droite de E dirigée par un vecteur $v_0 \in E$ non nul, on a

$$D = \{\lambda.v_0, \lambda \in \mathbb{R}\}$$

L'image de cette droite D par l'application f est alors

$$f(D) = \{f(v), v \in D\} = \{f(\lambda.v_0), \lambda \in \mathbb{R}\} = \{\lambda.f(v_0), \lambda \in \mathbb{R}\}$$

Autrement dit, l'image $f(D)$ de la droite D par f contient l'ensemble des vecteurs de F colinéaires au vecteur $f(v_0) \in F$. C'est donc une droite de F si le vecteur $f(v_0)$ est non nul. Sinon, $f(D) = \{0_F\}$.

On peut alors généraliser cette notion via la notion d'*image d'une application linéaire* :

Définition (image d'une application linéaire)

Soient E et F deux espaces vectoriels et f un morphisme de E dans F . On appelle *image de f* le sous ensemble de F défini par

$$\text{Im}(f) = \{v \in F / \exists u \in E / v = f(u)\} = \{f(u), u \in E\}$$

Le fait (évoqué plus haut) que l'image d'une droite soit une droite ou un point n'est qu'un cas particulier de la propriété suivante :

Proposition

Soient E et F deux espaces vectoriels et $f \in \mathcal{L}(E, F)$. L'image $\text{Im}(f)$ de f est un sous espace vectoriel de F .

Preuve : par définition, l'ensemble $\text{Im}(f)$ est un sous ensemble de F . Pour démontrer la proposition précédente, on doit donc établir que

1. $\text{Im}(f)$ est non vide.
On va montrer pour cela qu'il contient le vecteur nul 0_F .
2. $\text{Im}(f)$ est stable par combinaisons linéaires.

1. Puisque $f(0_E) = 0_F$, le vecteur 0_F est l'image par f d'un vecteur de E .
Donc $0_F \in \text{Im}(f)$ et $\text{Im}(f) \neq \emptyset$.

2. Soient v_1 et v_2 deux vecteurs de $\text{Im}(f)$ et $\lambda \in \mathbb{R}$.
Montrons que $\lambda.v_1 + v_2$ est un vecteur de $\text{Im}(f)$:

Puisque v_1 et v_2 sont dans $\text{Im}(f)$, il existe u_1 et u_2 dans E tels que

$$v_1 = f(u_1) \quad \text{et} \quad v_2 = f(u_2)$$

Mais alors

$$\lambda.v_1 + v_2 = \lambda.f(u_1) + f(u_2) = f(\lambda.u_1 + u_2)$$

Autrement dit, le vecteur $\lambda.v_1 + v_2$ est l'image par f du vecteur $\lambda.u_1 + u_2$. Il appartient donc bien à l'ensemble $\text{Im}(f)$ et $\text{Im}(f)$ est stable par combinaisons linéaires.

□

Noyau

Une autre propriété fondamentale des applications linéaires porte sur l'ensemble des vecteurs dont l'image est nulle. Ainsi,

Définition (noyau d'une application linéaire)

Soient E et F deux espaces vectoriels et $f \in \mathcal{L}(E, F)$. On appelle noyau de f l'ensemble $\text{Ker}(f)$ défini par

$$\text{Ker}(f) = \{u \in E / f(u) = 0_F\}$$

Comme pour l'image $\text{Im}(f)$, on peut alors montrer que

Proposition

Soient E et F deux espaces vectoriels et $f \in \mathcal{L}(E, F)$. Le noyau de f est un sous espace vectoriel de E .

Preuve :

1. Puisque $f(0_E) = 0_F$, le vecteur 0_E est bien un élément du noyau de f et $\text{Ker}(f)$ est non vide.

2. Soient u_1 et u_2 deux vecteurs de $\text{Ker}(f)$ et soit $\lambda \in \mathbb{R}$.
Montrons que le vecteur $\lambda.u_1 + u_2$ est encore un vecteur de $\text{Ker}(f)$:

$$\begin{aligned} f(\lambda.u_1 + u_2) &= \lambda.f(u_1) + f(u_2) \\ &= \lambda.0_F + 0_F = 0_F \end{aligned}$$

Donc $\lambda.u_1 + u_2 \in \text{Ker}(f)$ et $\text{Ker}(f)$ est stable par combinaisons linéaires. □

2.4.3 Injectivité, surjectivité des applications linéaires

Rappels

Soient E et F deux ensembles.

1. Une fonction $f : E \rightarrow F$ est dite *injective* si deux éléments distincts de E sont toujours envoyés sur deux éléments distincts de F :

$$\forall u, u' \in E, (u \neq u') \Rightarrow (f(u) \neq f(u'))$$

ou

$$\forall u, u' \in E, (f(u) = f(u')) \Rightarrow (u = u').$$

2. Une fonction $f : E \rightarrow F$ est dite *surjective* si chaque élément de E peut s'écrire comme l'image d'un élément de E :

$$\forall v \in F, \exists u \in E / v = f(u).$$

3. Une fonction f qui est à la fois injective et surjective sera dite *bijjective*. Il existe alors une fonction $f^{-1} : F \rightarrow E$, appelée bijection réciproque de f et vérifiant

$$\forall u \in E, f^{-1} \circ f(u) = u,$$

$$\forall v \in F, f \circ f^{-1}(v) = v$$

Cas des applications linéaires

Concernant les applications linéaires, on peut montrer que le caractère injectif et/ou surjectif est directement relié à la notion de noyau et d'image. Précisément,

Proposition

Soient E et F deux espaces vectoriels et $f \in \mathcal{L}(E, F)$. L'application f est *surjective* si et seulement si $\text{Im}(f) = F$.

Preuve : c'est évident d'après la définition d'une application surjective. Cette proposition est d'ailleurs valable pour toutes les applications (pas seulement les applications linéaires). Cependant, dans le cadre de l'algèbre linéaire, le fait que $\text{Im}(f)$ soit un espace vectoriel aura des conséquences particulières concernant les morphismes surjectifs.

Par ailleurs,

Proposition

Soient E et F deux espaces vectoriels et $f \in \mathcal{L}(E, F)$. L'application f est *injective* si et seulement si $\text{Ker}(f) = \{0_E\}$.

Preuve : cette proposition est, elle, spécifique aux applications linéaires. En effet, si $\text{Ker}(f) = \{0_E\}$, cela signifie que seul le vecteur nul de E est envoyé sur le vecteur nul de F . Il est clair que si cela n'est pas vérifié, l'application f n'est pas injective. Le résultat important de cette proposition est donc le suivant : si 0_F n'admet qu'un antécédent par f , alors tout vecteur de F admet au plus un antécédent par f . En effet, supposons que $\text{Ker}(f) = \{0_E\}$ et considérons deux vecteurs u et u' de E tels $f(u) = f(u')$. On a alors

$$f(u) = f(u') \iff f(u) - f(u') = 0_F$$

$$\iff f(u - u') = 0_F$$

$$\iff u - u' \in \text{Ker}(f)$$

Mais puisque $\text{Ker}(f) = \{0_E\}$, on a

$$u - u' = 0_E \quad \text{d'où} \quad u = u'$$

□

2.4.4 Applications linéaires en dimension finie

Les résultats énoncés ci-dessus sont des résultats généraux, valables dans tous les espaces vectoriels. Dans le cas spécifique des espaces vectoriels de dimension finie, on peut définir quelques objets supplémentaires énoncer quelques résultats spécifiques au cas de la dimension finie. Ces résultats supplémentaires sont en particulier liés à l'existence de *bases* pour les espaces vectoriels de dimension finie.

Matrice d'une application linéaire

Soient E et F deux espaces vectoriels *de dimensions finies* respectives p et n et soient

$$\mathcal{B}_E = \{e_1, \dots, e_p\} \quad \text{et} \quad \mathcal{B}_F = \{f_1, \dots, f_n\}$$

deux bases, respectivement de E et F . L'une des caractéristiques fondamentales des applications linéaires de E dans F est que pour connaître l'image $f(u)$ d'un vecteur quelconque de E , il suffit de connaître les images par f des vecteurs de \mathcal{B}_E et les coordonnées de

u dans \mathcal{B}_E . En effet, si l'on connaît les coordonnées $u = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}_{\mathcal{B}_E}$ de u dans \mathcal{B}_E , on a alors

$$u = x_1 \cdot e_1 + \dots + x_p \cdot e_p = \sum_{j=1}^p x_j \cdot e_j$$

Mais alors,

$$\begin{aligned} f(u) &= f(x_1 \cdot e_1 + \dots + x_p \cdot e_p) = f\left(\sum_{j=1}^p x_j \cdot e_j\right) \\ &= x_1 \cdot f(e_1) + \dots + x_p \cdot f(e_p) = \sum_{j=1}^p x_j \cdot f(e_j) \end{aligned}$$

Autrement dit, toute l'information portant sur la fonction f se trouve donc dans les images $f(e_j)$ des vecteurs de la base \mathcal{B}_E . Or puisque ces images $f(e_j)$ sont des vecteurs de E , ils ont des coordonnées a_{ij} dans la base \mathcal{B}_E :

$$\begin{aligned} f(e_1) &= a_{11} \cdot f_1 + a_{21} \cdot f_2 + \dots + a_{n1} \cdot f_n = \sum_{i=1}^n a_{i1} \cdot f_i = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}_{\mathcal{B}_F} \\ f(e_2) &= a_{12} \cdot f_1 + a_{22} \cdot f_2 + \dots + a_{n2} \cdot f_n = \sum_{i=1}^n a_{i2} \cdot f_i = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}_{\mathcal{B}_F} \\ \vdots & \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\ f(e_p) &= a_{1p} \cdot f_1 + a_{2p} \cdot f_2 + \dots + a_{np} \cdot f_n = \sum_{i=1}^n a_{ip} \cdot f_i = \begin{pmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{np} \end{pmatrix}_{\mathcal{B}_F} \end{aligned}$$

En rassemblant toutes ces coordonnées dans une matrice, on obtient *la matrice de f dans les bases \mathcal{B}_E et \mathcal{B}_F* :

$$\text{mat}_{\mathcal{B}_E, \mathcal{B}_F}(f) = \begin{pmatrix} f(e_1) & \cdots & f(e_p) \\ \downarrow & & \downarrow \\ a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix} \begin{matrix} f_1 \\ \vdots \\ f_n \end{matrix} \in \mathcal{M}_{np}(\mathbb{R})$$

Note : si f est un endomorphisme de E (i.e. si $E = F$), on prendra la même base \mathcal{B}_E au départ et à l'arrivée. La matrice associée, alors appelée *la matrice de f dans \mathcal{B}_E* , est en particulier une matrice carrée.

Par ailleurs, si l'on connaît la matrice d'une application $f : E \rightarrow F$ dans deux bases \mathcal{B}_E et \mathcal{B}_F , il est alors possible de calculer l'image par f de n'importe quel vecteur de E à l'aide d'un produit matriciel. Précisément, soit $u = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}_{\mathcal{B}_E}$. L'image $v = f(u)$ est un

vecteur de F . En notant $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{\mathcal{B}_F}$ ses coordonnées dans la base \mathcal{B}_F , on peut montrer que

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{\mathcal{B}_F} = \text{mat}_{\mathcal{B}_E, \mathcal{B}_F}(f) \times \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}_{\mathcal{B}_E}$$

Note : cette propriété s'étend bien entendu aux applications linéaires composées. Précisément, si f et g sont deux endomorphismes d'un même espace vectoriel E , alors pour toute base \mathcal{B} de E , la matrice de $g \circ f$ dans E vérifie

$$\text{mat}_{\mathcal{B}}(g \circ f) = \text{mat}_{\mathcal{B}}(g) \times \text{mat}_{\mathcal{B}}(f)$$

En particulier, pour tout endomorphisme $f \in \mathcal{L}(E)$, les matrices des composés successifs de f dans une base \mathcal{B} de E sont données par les puissances successives de la matrice $\text{mat}_{\mathcal{B}}(f)$:

$$\forall n \in \mathbb{N}, \quad \text{mat}_{\mathcal{B}}(f^n) = (\text{mat}_{\mathcal{B}}(f))^n$$

De même, si f est un endomorphisme bijectif, on peut montrer qu'il existe une relation forte entre les matrices représentant f et celles représentant f^{-1} . Précisément, on peut montrer que si f est un endomorphisme bijectif, alors dans toute base \mathcal{B} de E , on a

- La matrice $\text{mat}_{\mathcal{B}}(f)$ est inversible.
- $\text{mat}_{\mathcal{B}}(f^{-1}) = (\text{mat}_{\mathcal{B}}(f))^{-1}$

Exemples :

- Dans toute base de E , la matrice de l'application *identité*

$$\text{Id}_E : E \longrightarrow E \\ x \longmapsto x$$

est la matrice I_n .

- Dans le plan muni d'un repère orthonormé $\mathcal{R} = (\mathcal{O}; \vec{i}, \vec{j})$,

- La matrice de la projection orthogonale sur l'axe $(\mathcal{O}x)$ est

$$\pi_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

- La matrice de la projection orthogonale sur l'axe $(\mathcal{O}y)$ est

$$\pi_y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

- La matrice de la symétrie orthogonale par rapport à l'axe $(\mathcal{O}x)$ est

$$\sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

- Dans l'espace muni d'un repère orthonormé $\mathcal{R} = (\mathcal{O}; \vec{i}, \vec{j}, \vec{k})$,

- La matrice de la projection orthogonale sur le plan $(x\mathcal{O}y)$ est

$$\pi_{xy} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

- La matrice de la projection orthogonale sur l'axe $(x\mathcal{O}z)$ est

$$\pi_{xz} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

— La matrice de la symétrie orthogonale par rapport au plan (yOz) est

$$\sigma_{yz} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Changement de bases pour les applications linéaires

Comme cela a été évoqué plusieurs fois, tout espace vectoriel de dimension finie admettant une infinité de bases. On la notion de matrice d'application linéaire est, comme la notion de matrice d'une famille de vecteurs, intimement liée aux bases choisies pour représenter les espaces E et F de départ et d'arrivée. Il est néanmoins possible, comme pour les matrices de vecteurs, de relier entre elles les différentes matrices représentant une même application linéaire dans des bases différentes. Ainsi, dans le cas particulier des endomorphismes,

Théorème (changement de base pour les endomorphismes)

Soient E un espace vectoriel, \mathcal{B}_1 et \mathcal{B}_2 deux bases de E et f un endomorphisme de E . En notant P la matrice de passage de \mathcal{B}_1 à \mathcal{B}_2 , on a

$$\text{mat}_{\mathcal{B}_2}(f) = P^{-1} \times \text{mat}_{\mathcal{B}_1}(f) \times P$$

Notes :

- Le théorème ci dessus se généralise à tout morphisme de E dans F . L'adaptation est laissée au lecteur.
- Cette relation permet de montrer que, bien que deux matrices associées à une même application linéaire n'aient pas les mêmes coefficients, elles ont un certain nombre de propriétés en commun (comme le déterminant ou la trace), appelés *invariants* de matrices et qui représentent en réalité des propriétés du morphisme qui leur est associé.

Théorème du rang

Outre la notion de matrice, d'autres propriétés fondamentales des applications linéaires en dimension fini portent sur les relations existant entre les dimensions des espaces vectoriels de départ et d'arrivée et celles des sous espaces spécifiques $\text{Ker}(f)$ et $\text{Im}(f)$.

Dans ce cadre, on peut dans un premier temps définir *le rang* d'une application linéaire :

Définition (rang d'une application linéaire)

Soient E et F deux espaces vectoriels de dimension finie et f un morphisme de E dans F . On appelle *rang* de f l'entier $\text{rg}(f)$ défini par

$$\text{rg}(f) = \dim(\text{Im}(f))$$

Ainsi, soient E et F deux espaces vectoriels de dimension finie et f un morphisme de E dans F .

1. Puisque $\text{Im}(f)$ est un sous espace vectoriel de F , on a

$$\text{rg}(f) \leq \dim(F)$$

2. Puisque $\text{Ker}(f)$ est un sous espace vectoriel de E , on a

$$\dim(\text{Ker}(f)) \leq \dim(E)$$

D'autre part, d'après la définition de l'image de f , on peut montrer que le sous espace vectoriel est engendré par l'image de n'importe quelle base de E . Précisément, pour toute base $\mathcal{B}_E = \{e_1, \dots, e_p\}$ de E , on a

$$\text{Im}(f) = \text{Vect}(\{f(e_1), \dots, f(e_p)\})$$

(Exercice : à démontrer).

De ce fait, il est clair que

$$\text{rg}(f) \leq \dim(E)$$

Mais le résultat le plus fort concernant ces dimensions est le suivant :

Théorème (théorème du Rang)

Soient E et F deux espaces vectoriels de dimension finie et f un morphisme de E dans F . Alors

$$\text{rg}(f) + \dim(\text{Ker}(f)) = \dim E.$$

Preuve : notons

- $p = \dim E$
- $\ell = \dim(\text{Ker}(f))$
- $\mathcal{B}_K = \{e_1, \dots, e_\ell\}$ une base de $\text{Ker}(f)$.

Puisque $\text{Ker}(f) \subset E$, on a $\ell \leq p$. D'après le théorème de la base incomplète, il est possible d'ajouter $p - \ell$ vecteurs $e_{\ell+1}, \dots, e_p$ à la famille \mathcal{B}_K pour former une base \mathcal{B} de E :

$$\mathcal{B} = \{e_1, \dots, e_\ell, e_{\ell+1}, \dots, e_p\}$$

Mais alors, pour tout vecteur $v \in \text{Im}(f)$, il existe par définition $u \in E$ tel que $v = f(u)$.

En notant alors $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}$ les coordonnées de u dans la base \mathcal{B} , on a

$$u = x_1 \cdot e_1 + \dots + x_\ell \cdot e_\ell + x_{\ell+1} \cdot e_{\ell+1} + \dots + x_p \cdot e_p = \sum_{k=1}^{\ell} x_k \cdot e_k + \sum_{k=\ell+1}^p x_k \cdot e_k$$

et

$$\begin{aligned} v &= f \left(\sum_{k=1}^{\ell} x_k \cdot e_k + \sum_{k=\ell+1}^p x_k \cdot e_k \right) \\ &= \underbrace{\sum_{k=1}^{\ell} x_k \cdot f(e_k)}_{=0_F} + \sum_{k=\ell+1}^p x_k \cdot f(e_k) \\ &= \sum_{k=\ell+1}^p x_k \cdot f(e_k) \in \text{Vect}(\{f(e_{\ell+1}), \dots, f(e_p)\}) \end{aligned}$$

Autrement dit, l'espace vectoriel $\text{Im}(f)$ est engendré par la famille $\mathcal{B}_I = \{f(e_{\ell+1}), \dots, f(e_p)\}$. En termes de dimension, on a donc montré que

$$\dim(\text{Im}(f)) = \text{rg}(f) \leq p - \ell = \dim(E) - \dim(\text{Ker}(f))$$

Pour établir l'égalité cherchée, il reste à démontrer que la famille \mathcal{B}_I est libre. Pour cela, on considère $(p - \ell)$ coefficients $\lambda_{\ell+1}, \dots, \lambda_p \in \mathbb{R}$ tels que

$$\sum_{k=\ell+1}^p \lambda_k \cdot f(e_k) = 0_F$$

Puisque f est linéaire, on a

$$f\left(\sum_{k=\ell+1}^p \lambda_k \cdot e_k\right) = 0_F$$

Autrement dit, le vecteur $\sum_{k=\ell+1}^p \lambda_k \cdot e_k$ appartient au noyau de f . Ce vecteur a donc des coordonnées $\lambda_1, \dots, \lambda_\ell$ dans \mathcal{B}_K :

$$\begin{aligned} \sum_{k=\ell+1}^p \lambda_k \cdot e_k &= \lambda_1 \cdot e_1 + \dots + \lambda_\ell \cdot e_\ell = \sum_{k=1}^{\ell} \lambda_k \cdot e_k \\ \Leftrightarrow \sum_{k=\ell+1}^p \lambda_k \cdot e_k - \sum_{k=1}^{\ell} \lambda_k \cdot e_k &= 0_E \end{aligned}$$

On a donc une combinaison linéaire des e_k égale au vecteur nul. Comme la famille \mathcal{B} est libre (c'est une base de E), on doit avoir $\lambda_k = 0$ pour tout k .

On a en particulier les coefficients $\lambda_{\ell+1} = \dots = \lambda_p = 0$. Autrement dit, la famille $\{f(e_{\ell+1}), \dots, f(e_p)\}$ est libre.

C'est donc une base de $\text{Im}(f)$ et on en déduit que

$$\text{rg}(f) = \dim(\text{Im}(f)) = p - \ell = \dim(E) - \dim(\text{Ker}(f))$$

et le théorème du rang est démontré. \square

L'une des conséquences principales du théorème du rang porte sur la bijectivité des endomorphismes en dimension finie. Précisément,

Proposition

Soient E un espace vectoriel de dimension finie et f un endomorphisme de E . On a

$$f \text{ injective} \Leftrightarrow f \text{ surjective} \Leftrightarrow f \text{ bijective}$$

Preuve : cette propriété est directement issue du théorème du rang. En effet, on a montré plus haut qu'il existait un lien direct entre le caractère injectif et/ou surjectif de f et la dimension des espaces $\text{Ker}(f)$ et $\text{Im}(f)$. Précisément :

- f est injective si et seulement si $\dim(\text{Ker}(f)) = 0$.
- f est surjective si et seulement si $\text{rg}(f) = \dim(E)$.

Mais alors, d'après le théorème du rang,

$$f \text{ est inj.} \Leftrightarrow \dim(\text{Ker}(f)) = 0 \Leftrightarrow \text{rg}(f) = \dim(E) \Leftrightarrow f \text{ est surj.}$$

□

2.5 Diagonalisation des endomorphismes et des matrices carrées

La diagonalisation est une opération essentielle dans de nombreux domaines faisant intervenir les matrices et/ou les applications linéaires. Un cadre naturel permettant de justifier cette opération est porté par les *endomorphismes* d'un espace vectoriel de dimension finie et notamment les endomorphismes du plan ou de l'espace (munis d'un point d'origine). Précisément, un endomorphisme du plan ou de l'espace peut être vu comme une transformation de ce dernier. Or le caractère linéaire d'une telle transformation induit l'existence de directions principales de transformations. La diagonalisation d'un endomorphisme consiste alors à déterminer ces directions principales ainsi que l'intensité des déformations dans ces directions.

2.5.1 Définitions

Diagonalisation des endomorphismes

Définition (diagonalisation d'un endomorphisme)

Soient E un espace vectoriel de dimension finie et f un endomorphisme de E . *Diagonaliser* f consiste à déterminer une base \mathcal{B} de E telle que la matrice $\text{mat}_{\mathcal{B}}(f)$ est *diagonale* :

$$\text{mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Par définition, la matrice de f dans une base $\mathcal{B} = \{e_1, \dots, e_n\}$ de E est de la forme

$$\text{mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

si et seulement si

$$\forall i \in \{1, \dots, n\}, \quad f(e_i) = \lambda_i \cdot e_i$$

Cette propriété conduit à la définition suivante :

Définition (vecteurs propres et valeurs propres I)

Soient E un espace vectoriel de dimension finie et f un endomorphisme de E .

On appelle *vecteur propre de f* tout vecteur non nul de E proportionnel à son image par f .

Autrement dit, $v \in E$ est un vecteur propre de f si et seulement si

$$v \neq 0_E \quad \text{et} \quad \exists \lambda \in \mathbb{R} / f(v) = \lambda \cdot v$$

Le coefficient λ liant v et $f(v)$ est appelé *valeur propre de f associée au vecteur propre v* .

Note : par définition, diagonaliser f consiste à construire une base de E faite de vecteurs propres de f .

Définition (spectre et sous espaces propres I)

Soient E un espace vectoriel de dimension finie et f un endomorphisme de E .

1. On appelle *spectre de f* l'ensemble $\text{Spec}(f)$ des valeurs propres de f .
2. Pour toute valeur propre $\lambda \in \text{Spec}(f)$, on appelle *sous espace propre de f associé à la valeur propre λ* l'ensemble E_λ défini par

$$E_\lambda = \{v \in E / f(v) = \lambda \cdot v\}$$

Notes :

1. Par définition, les valeurs propres de f sont les coefficients que l'on retrouve sur la diagonale de la matrice diagonale issue du processus de diagonalisation.
2. Par définition, l'ensemble E_λ est constitué du vecteur nul 0_E ainsi que l'ensemble des vecteurs propres de f associés à la valeur propre λ .
On peut en outre montrer que pour tout $\lambda \in \text{Spec}(f)$, l'ensemble E_λ est un sous espace vectoriel de E (exercice : à démontrer).

En pratique, diagonaliser un endomorphisme $f \in \mathcal{L}(E)$ consiste donc à construire une base \mathcal{B} de E , faite de vecteurs propres. La théorie développée ci-dessus nous assure

alors que

- La matrice de $D = \text{mat}_{\mathcal{B}}(f)$ dans cette base est **une matrice diagonale**.
- Les coefficients diagonaux de $D = \text{mat}_{\mathcal{B}}(f)$ sont **les valeurs propres de f** .

Diagonalisation des matrices carrées

Étant donné le lien existant entre endomorphismes et matrices carrées, on peut également étendre cette définition aux matrices, indépendamment de la notion d'endomorphisme. Précisément,

Définition (diagonalisation d'une matrice carrée)

Soient $n \in \mathbb{N}^*$ et A une matrice carrée à n lignes. *Diagonaliser* A consiste à **déterminer une matrice inversible $P \in \text{GL}_n(\mathbb{R})$ telle que**

$$P^{-1} \times A \times P = D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Définition (vecteurs propres et valeurs propres II)

Soient $n \in \mathbb{N}^*$ et $A \in \mathcal{M}_n(\mathbb{R})$. On appelle *vecteur propre de A* tout vecteur $X \in \mathbb{R}^n$ tel que

$$X \neq 0_{\mathbb{R}^n} \quad \text{et} \quad \exists \lambda \in \mathbb{R} / A \times X = \lambda.X$$

Le coefficient $\lambda \in \mathbb{R}$ liant les vecteurs X et $A \times X$ est appelé *valeur propre de A associée au vecteur propre X* .

De même,

Définition (spectre et sous espaces propres II)

Soient $n \in \mathbb{N}^*$ et $A \in \mathcal{M}_n(\mathbb{R})$.

1. On appelle *spectre de A* l'ensemble $\text{Spec}(A)$ des valeurs propres de A .
2. Pour toute valeur propre $\lambda \in \text{Spec}(A)$, on appelle *sous espace propre de A associé à la valeur propre λ* l'ensemble E_λ défini par

$$E_\lambda = \{X \in \mathbb{R}^n / A \times X = \lambda.X\}$$

Bien entendu, le spectre et les sous espaces propres d'une matrice vérifient les mêmes propriétés algébriques que pour les endomorphismes.

Par ailleurs, en pratique, diagonaliser une matrice A consiste à déterminer une base \mathcal{B} de \mathbb{R}^n faite de vecteurs propres de A . Le calcul matriciel assure alors que si l'on note la matrice P représentant \mathcal{B} dans la base canonique de \mathbb{R}^n ,

- Le produit $D = P^{-1} \times A \times P$ est diagonal.
- Les coefficients diagonaux de $D = P^{-1}AP$ sont les valeurs propres de A .

2.5.2 Diagonalisation pratique

Dualité endomorphismes/matrice

Il existe bien entendu un lien étroit entre la diagonalisation d'un endomorphisme et la diagonalisation des matrices qui le représentent. Précisément,

Théorème

Soient E un espace vectoriel de dimension finie et f un endomorphisme de E .
Pour toute base de E ,

1. Les valeurs propres de $\text{mat}_{\mathcal{B}}(f)$ sont les valeurs propres de f .
2. Chacun des sous espaces propres de $\text{mat}_{\mathcal{B}}(f)$ peut être associé à un sous espace propre de f .

Ce lien permet en pratique de se concentrer sur la diagonalisation des matrices. D'autre part, on va voir que le calcul matriciel offre des outils efficaces permettant de déterminer l'ensemble des éléments propres (spectre et sous espaces propres) d'une matrice donnée.

Recherche des valeurs propres

Bien que l'objectif final de la diagonalisation d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ soit de déterminer une base de \mathbb{R}^n faite de vecteurs propres de A , la première étape consiste à déterminer l'ensemble du spectre de A . Or on peut montrer que

Proposition/Définition (polynôme caractéristique d'une matrice carrée)

Soit A une matrice carrée à n lignes.

Prop. Les valeurs propres de A sont les solutions λ de l'équation

$$\det(A - \lambda.I_n) = 0$$

Déf. La quantité $\chi_A(\lambda) = \det(A - \lambda.I_n)$ est un polynôme de degré n , appelé *polynôme caractéristique de A* .

Ainsi, les valeurs propres de A sont les racines du polynôme $\chi_A(\lambda)$.

Notes :

1. D'un point de vue pratique, la matrice $A - \lambda.I_n$ est une matrice carrée obtenue en soustrayant λ à chacun des termes diagonaux de A .
2. Si $n = 2$, on obtient alors un polynôme de degré 2. On peut alors déterminer les racines de P . En dimensions supérieures, il n'est pas sûr que l'on puisse déterminer les valeurs propre d'une matrice donnée. Le processus de diagonalisation ne peut alors pas aller à son terme. Cependant, on verra qu'il est parfois possible de profiter des propriétés du déterminant pour déterminer *au cours du calcul* certaines des valeurs propres de A .

Exemple

Soit $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

1. Le polynôme caractéristique $\chi_A(\lambda)$ est

$$\begin{aligned} \chi_A(\lambda) &= \det(A - \lambda.I_2) \\ &= \begin{vmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{vmatrix} \\ &= (1 - \lambda)^2 - 4 \\ &= (1 - \lambda - 2)(1 - \lambda + 2) \\ &= (-1 - \lambda)(3 - \lambda) \end{aligned}$$

2. Les valeurs propres de A sont

$$\text{Spec}(A) = \{-1, 3\}$$

Recherche des vecteurs propres

Une fois déterminé le spectre de A , on doit, pour chaque valeur propre λ_i obtenue, déterminer le sous-espace propre E_{λ_i} associé, i.e. l'ensemble des solutions du système linéaire

$$(S_i) : A \times X = \lambda_i \cdot X \iff (A - \lambda_i \cdot I_n) \times X = 0$$

Exemple : déterminer les sous espaces propres de la matrice $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

Construction de la matrice de passage

Une fois déterminés l'ensemble des sous espaces propres de A , on connaît l'ensemble des vecteurs propres de A . Pour construire la matrice de passage P cherchée, on doit alors construire une famille libre de n vecteurs tirée de ces sous espaces propres.

Or on peut montrer que les sous espaces propres (qui sont des sous espaces vectoriels) sont en somme directe deux à deux. Autrement dit, si l'on extrait une base de chacun de ces sous espaces propres, la réunion de toutes ces bases forme une famille libre de \mathbb{R}^n . Si ce processus produit n vecteurs, la famille ainsi construite est une base de \mathbb{R}^n . En rassemblant ces n vecteurs dans une même matrice, on obtient la matrice P cherchée. La théorie nous assure alors que la matrice $D = P^{-1} \times A \times P$ est une matrice diagonale dont les coefficients diagonaux sont exactement les valeurs propres de A .

Exemple : construire une matrice P telle que

$$P^{-1} \times \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \times P = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$$

Obstruction à la diagonalisabilité

Bien que, en pratique, on travaillera principalement avec des matrices (et des endomorphismes) diagonalisables, il est à noter que le processus exposé ci-dessus n'aboutit pas toujours. En effet, il peut arriver que, une fois déterminé l'ensemble des sous espaces propres de A , il ne soit pas possible de trouver suffisamment de vecteurs propres indépendants pour former une base de \mathbb{R}^n . Cette obstruction est directement reliée à la *dimension* de chacun des sous espaces propres. Précisément, la théorie des polynômes assure que pour

toute matrice $A \in \mathcal{M}_n(\mathbb{R})$, il existe

- un entier $s \in \mathbb{N}^*$,
- s nombres complexes distincts $\lambda_1, \dots, \lambda_s \in \mathbb{C}$,
- s nombres entiers $\alpha_1, \dots, \alpha_s \in \mathbb{N}^*$,

tels que

$$\chi_A(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{\alpha_k} \quad \text{et} \quad \sum_{k=1}^s \alpha_k = n$$

Par ailleurs, on peut alors montrer que pour chacune des valeurs propres $\lambda_k \in \text{Spec}(A)$, on a

$$1 \leq \dim(E_{\lambda_k}) \leq \alpha_k$$

Autrement dit, de chacun des sous espaces propres E_{λ_k} , on peut extraire au plus α_k vecteurs propres indépendants. Pour que la matrice A soit diagonalisable, il faut donc que chacun de ses sous espaces propres soit exactement de dimension α_k (soit le plus gros possible).

Note : si le polynôme caractéristique $\chi_A(\lambda)$ admet n racines distinctes i.e. si

$$s = n \quad \text{et} \quad \forall k \in \{1, \dots, n\}, \alpha_k = 1$$

l'inégalité ci-dessus assure que l'ensemble des sous espaces propres de A forme une famille de n droites indépendantes. En choisissant un vecteur directeur sur chacune de ces droites, on obtient une base de \mathbb{R}^n formée de vecteurs propres de A et A est diagonalisable.

2.5.3 Diagonalisation des matrices symétriques

Produit scalaire et bases orthonormées

À partir de la notion de coordonnées, on peut étendre la notion de produit scalaire et de norme à tout espace vectoriel de dimension finie. Précisément, soit E un espace vectoriel de dimension n et \mathcal{B} une base de E . On appellera *produit scalaires relativement*

à la base \mathcal{B} de deux vecteurs $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}$ et $v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{\mathcal{B}}$ de E la somme

$$\langle u, v \rangle_{\mathcal{B}} = \sum_{k=1}^n x_k \cdot y_k$$

Deux vecteurs u et v de E sont alors dits *orthogonaux* relativement à \mathcal{B} si et seulement si

$$\langle u, v \rangle_{\mathcal{B}} = 0$$

De façon plus générale, deux sous espaces vectoriels \mathcal{S}_1 et \mathcal{S}_2 de E sont dit *orthogonaux* relativement à la base \mathcal{B} si et seulement

$$\forall u_1 \in \mathcal{S}_1, u_2 \in \mathcal{S}_2, \quad \langle u_1, u_2 \rangle = 0$$

On note alors $\mathcal{S}_1 \perp \mathcal{S}_2$.

Par ailleurs, pour tout vecteur $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}} \in E$, on appelle *norme de u* relativement à la base \mathcal{B} le réel

$$\|u\|_{\mathcal{B}} = \sqrt{\langle u, u \rangle_{\mathcal{B}}} = \sqrt{\sum_{k=1}^n x_k^2}$$

Un vecteur $u \in E$ sera alors dit *normé* relativement à la base \mathcal{B} si et seulement si

$$\|u\|_{\mathcal{B}} = 1$$

Enfin, on appelle *famille orthonormée de E* relativement à la base \mathcal{B} toute famille

$$\mathcal{F} = \{e_1, \dots, e_p\}$$

de n vecteurs de E telle que

pour tout couple $(i, j) \in \{1, \dots, p\}^2$,

$$i \neq j \Rightarrow \langle e_i, e_j \rangle_{\mathcal{B}} = 0$$

et pour tout $i \in \{1, \dots, p\}$,

$$\|e_i\|_{\mathcal{B}} = 1$$

On peut alors montrer que toute famille orthonormée est *libre*. Ainsi, toute famille orthonormée \mathcal{F} de n vecteurs de E est *une base orthonormée* de E (une BON de E). On peut alors montrer la matrice (inversible) $P = \text{mat}_{\mathcal{B}}(\mathcal{F})$ vérifie

$$P^{-1} = {}^tP$$

Exercice : montrer que si \mathcal{F} est une BON de E et si l'on note $P = \text{mat}_{\mathcal{B}}(\mathcal{F})$, alors

$${}^tP \times P = I_n$$

Application à la diagonalisation des matrices symétriques réelles

Rappel : une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est dite *symétrique* si et seulement si

$${}^tA = A$$

Or on peut montrer que toute matrice symétrique *réelle* est diagonalisable et que, en outre, les sous espaces propres d'une matrice symétrique sont deux à deux orthogonaux. En pratique, cela signifie que si, lors de la diagonalisation d'une matrice symétrique, on s'attache à construire, pour chaque sous espace propre, une base orthonormée, l'ensemble de ces BON forme alors une BON de E . La matrice P de cette BON vérifie donc ${}^tP = P^{-1}$ et le produit $D = {}^tP \times A \times P$ est une matrice diagonale dont les coefficients diagonaux sont les valeurs propres de A .

Exemple :

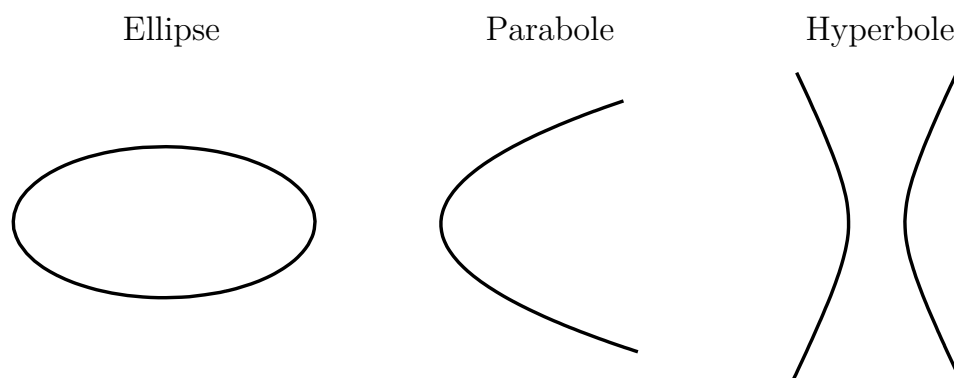
1. Vérifier que les sous espaces propres de la matrice $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ sont deux droites orthogonales entre elles.
2. Construire une BON de \mathbb{R}^2 faite de vecteurs propres de A .
3. Montrer que la matrice P de cette BON dans la base canonique de \mathbb{R}^2 vérifie ${}^tP = P^{-1}$.

Chapitre 3

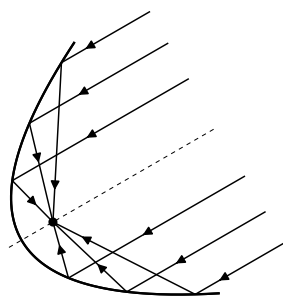
Coniques, quadriques, formes quadratiques

Introduction

On appelle conique toute courbe de type



Les nombreuses propriétés géométriques de ces courbes en ont fait un objet d'étude idéal pour les mathématiciens et ont trouvé de nombreuses applications scientifiques, dont on retrouve des exemples dans le quotidien.



Ainsi, l'une des applications importantes est une propriété des paraboles, dont les caractéristiques géométriques entraînent l'existence d'un point "à l'intérieur" de la parabole auquel se focalise tout rayon entrant dans la parabole parallèlement à son axe.

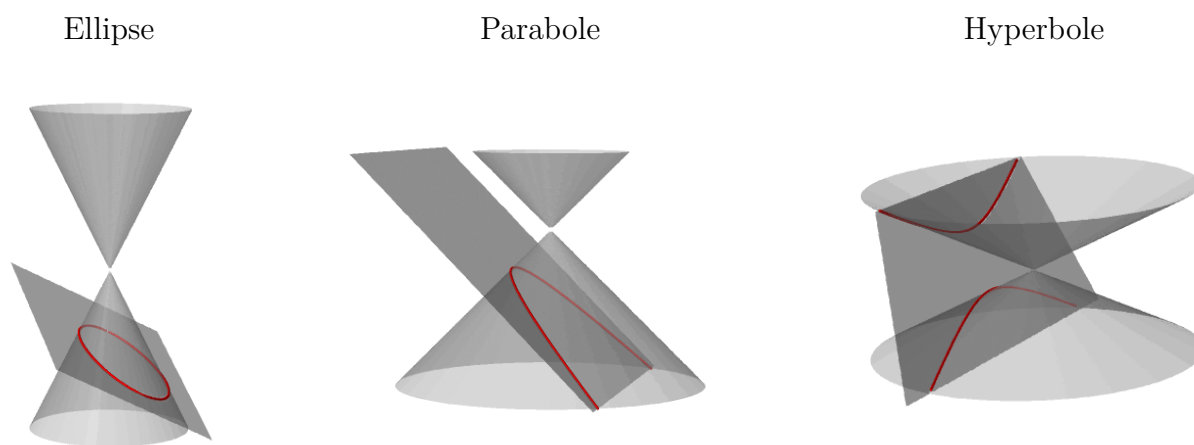
Cette propriété est en particulier exploitée en télécommunication (antennes télescopiques) ou en optique (phares télescopiques). Les fours solaires exploitent également cette propriété pour concentrer les rayons du soleil en un point donné.

Par ailleurs, les quadriques, i.e. les surfaces obtenues en généralisant la notion de conique à la dimension 3, en étendant la notion de conique à la dimension 3, possèdent égale-

ment de nombreuses propriétés géométriques, exploitables en particulier en construction. Ainsi, l'hyperboloïde à une nappe, surface obtenue en rotation d'une hyperbole autour de son axes vertical interviennent dans la conception des cheminées de centrales nucléaires ou de certaines tours (château d'eau).



Ces propriétés, souvent découvertes de façon empiriques, ont très tôt poussé les mathématiciens à étudier les coniques de façon formelle. Les travaux d'Appolonius (IV^{ème} siècle avant J.C.) par exemple, sur l'étude de l'intersection d'un cône de révolution avec un plan de l'espace, font partie des plus anciens travaux mathématiques sur les coniques qui nous soient parvenus. Ils sont notamment à l'origine de la terminologie actuelle, le type de conique dépendant alors de l'inclinaison du plan de coupe par rapport à l'axe du cône :



Archimède (III^{ème} siècle avant J.C.) a également contribué à l'étude des coniques et des quadriques en travaillant sur le calcul d'aires et de volumes délimités par ces courbes et surfaces.

Par ailleurs, c'est encore pour étudier les coniques que les mathématiciens perses et arabes ont posé les bases du calcul algébrique (les notions d'équations, d'inconnues, etc) qui ont été plus tard reprises et développées, notamment par Descartes, pour produire l'outils de la géométrie analytique actuelle.

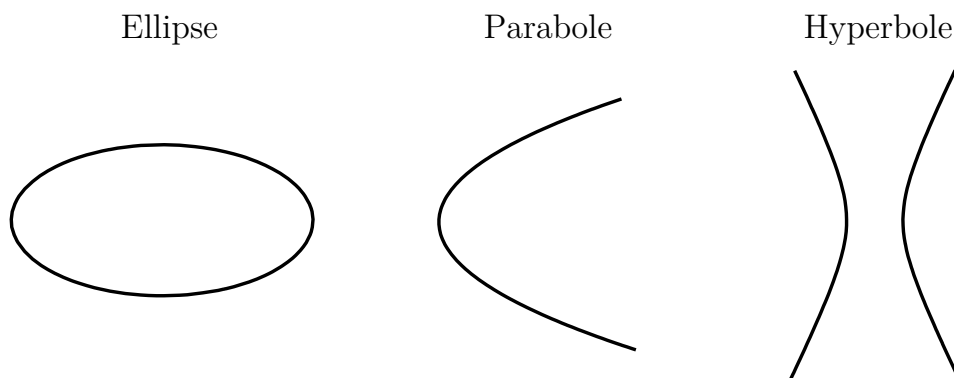
Citons enfin les travaux de l'astronome Kepler dont toute la théorie astronomique est basée sur le fait que les objets se déplaçant dans l'espace (lunes, planètes, comètes, etc) suivent des trajectoires elliptiques, voire paraboliques.

3.1 Définition géométrique 2D

Bien que naturelle, la définition des coniques comme intersection d'un plan et d'un cône a un inconvénient majeur dans son utilisation par le fait qu'elle impose un travail en 3 dimensions alors que les courbes étudiées sont des courbes planes. Le premier travail des mathématiciens a alors été de trouver une autre définition pour ces coniques, basée uniquement sur des éléments du plan et à partir de laquelle ils ont pu démontrer les premières propriétés géométriques, déduites de l'observation :

- Deux axes de symétrie et un centre de symétrie pour les ellipses et les hyperboles
- Un axe de symétrie pour les paraboles
- Quatre points remarquables, appelés sommets pour une ellipse
- Deux sommets pour une hyperbole,
- Un sommet pour une parabole

Exercice : placer ces éléments remarquables sur les coniques ci-dessous :



Définition (conique)

Soient

- F un point du plan,
- \mathcal{D} une droite du plan ne passant pas par F ,
- e un réel strictement positif.

1. On appelle alors conique associée au triplet (F, \mathcal{D}, e) (ou simplement conique (F, \mathcal{D}, e)) la courbe \mathcal{C} formée par l'ensemble des points M du plan tels que

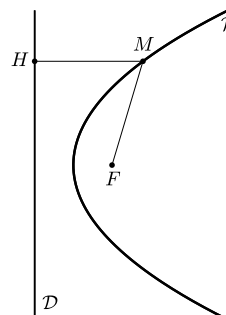
$$d(M, F) = e.d(M, \mathcal{D})$$

où $d(M, F)$ représentant *la distance* de M à F .

2. Le point F est appelé *le foyer* de \mathcal{C} .
3. La droite \mathcal{D} est appelée *la directrice* de \mathcal{C} .
4. Le réel $e > 0$ est appelé *l'excentricité* de \mathcal{C} .

En notant H le projeté orthogonal de M sur \mathcal{D} , la conique \mathcal{C} définie par (F, \mathcal{D}, e) est l'ensemble des points M tels que

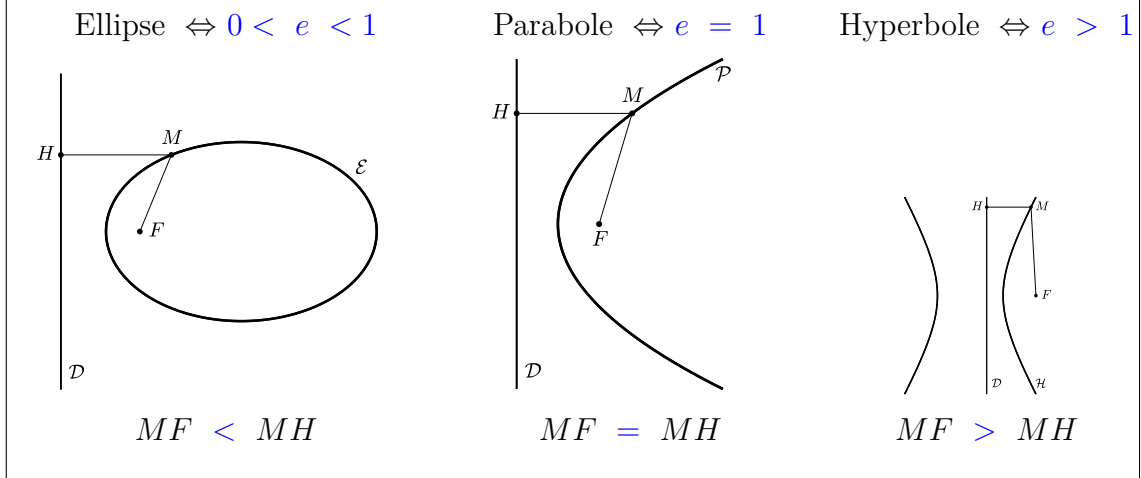
$$MF = e.MH.$$



On peut montrer que cette définition correspond bien aux courbes obtenues par l'intersection de cônes par des plans, bien que l'équivalence ne soit pas immédiate.

De plus, on peut alors caractériser les éléments remarquables des coniques à l'aide de ces trois éléments. Ainsi

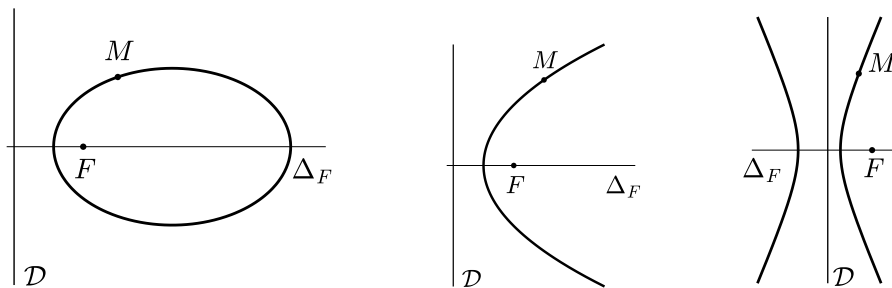
- L'excentricité e varie en fonction de l'angle d'inclinaison du plan de coupe. On peut alors classer les coniques du plan en fonction de leur excentricité :



- L'axe de symétrie commun à tout type de conique (représentant l'horizontale sur les figures ci-dessus) est l'unique droite du plan perpendiculaire à la directrice \mathcal{D} et passant par le foyer F . Cette droite est appelée *l'axe focal* de \mathcal{C} , noté Δ ou Δ_F .

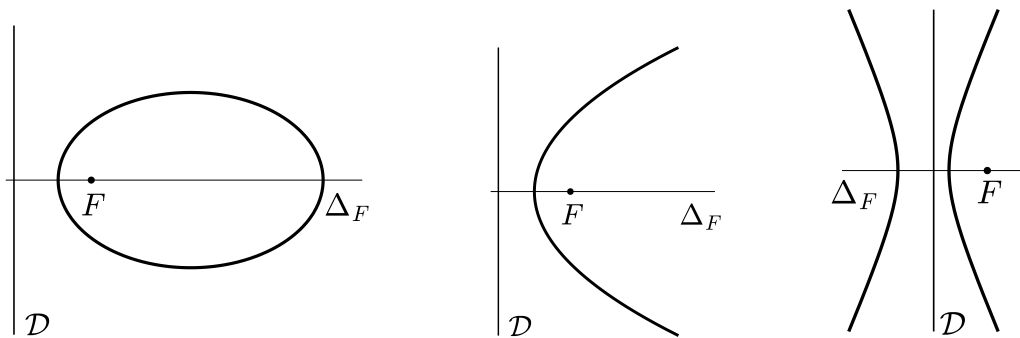
Rappel : un axe de symétrie est une droite du plan telle que le symétrique M' de tout point M de \mathcal{C} par rapport à cette droite est encore un point de \mathcal{C} .

Exercice : placer sur les coniques ci-dessous le symétrique du point M .



- La plupart des sommets d'une conique (tous sauf deux pour l'ellipse) se situent sur cet axe focal.

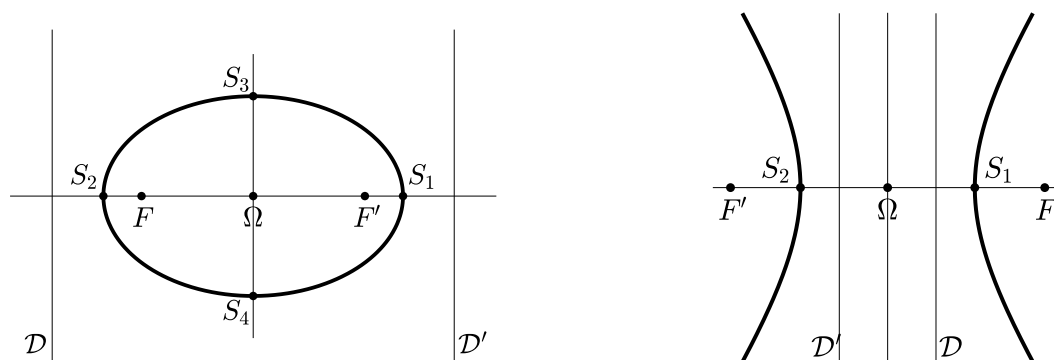
Exercice : placer sur les coniques ci-dessous les sommets appartenant à l'axe focal Δ_F .



- Pour les ellipses et les hyperboles, le second axe de symétrie Δ' est la médiatrice du segment $[SS']$ formé par les deux sommets S et S' situés sur l'axe focal et le centre de symétrie est le milieu Ω de ce segment.

Par ailleurs, ce second axe de symétrie permet de démontrer l'existence d'un second foyer F' et d'une seconde directrice \mathcal{D}' , symétriques respectifs de F et \mathcal{D} par rapport à Δ' .

Enfin, sur leur second axe de symétrie, les ellipses admettent également deux sommets supplémentaires.



Note : de façon générale, on appelle *sommet* d'une conique tout point d'intersection de celle-ci avec l'un de ses axes de symétrie.

Exercice :

1. Montrer à l'aide d'outils purement géométriques que l'axe focale d'une conique est un axe de symétrie.
2. Montrer qu'une conique ne rencontre jamais son foyer ni sa directrice.

3.2 Définition analytique

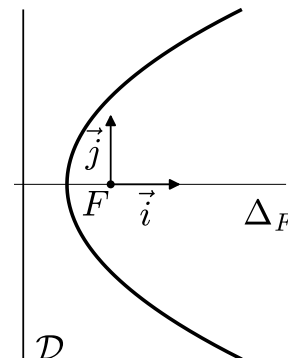
3.2.1 Une première équation

Pour aller plus loin, il devient nécessaire de faire appel à la notion coordonnées. Pour cela, il faut traduire la définition géométrique d'une conique $\mathcal{C} = (F, \mathcal{D}, e)$

$$\mathcal{C} = \{M \in P / d(M, F) = e.d(M, \mathcal{D})\}$$

en une équation portant sur les coordonnées du point M dans un repère fixé du plan.

En choisissant un repère adapté aux données de départ, on simplifie le problème. Précisément, on note $\mathcal{R}_F = (F, \vec{i}, \vec{j})$ le repère orthonormé du plan centré au point F , tel que les vecteurs \vec{i} et \vec{j} soient respectivement directeurs des droites Δ_F et \mathcal{D} .



Dans ce repère \mathcal{R}_F ,

1. Les coordonnées du point F sont

$$F = \begin{pmatrix} 0 \\ 0 \end{pmatrix}_{\mathcal{R}_F}$$

2. En notant $d = d(F, \mathcal{D})$, la droite \mathcal{D} a pour équation

$$\mathcal{D} : x = d$$

3. La droite Δ_F a pour équation

$$\Delta_F : y = 0$$

4. Pour tout point $M = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{R}_F}$, on a

$$d(M, F) = \sqrt{x^2 + y^2} \quad \text{et} \quad d(M, \mathcal{D}) = x + d$$

5. Ainsi, un point $M = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{R}_F}$ appartient à la conique $\mathcal{C} = (F, \mathcal{D}, e)$ si et seulement si

$$\begin{aligned}
 d(M, F) = e.d(M, \mathcal{D}) &\iff \sqrt{x^2 + y^2} = e.(x + d) \\
 &\iff x^2 + y^2 = e^2.(x^2 + 2xd + d^2) \\
 &\iff (1 - e^2)x^2 + y^2 - 2epx - p^2 = 0 \\
 \text{où } p = e.d
 \end{aligned}$$

On obtient ainsi une première équation pour la conique $\mathcal{C} = (F, \mathcal{D}, e)$:

$$\mathcal{C} : (1 - e^2)x^2 + y^2 - 2epx - p^2 = 0$$

À partir de cette équation, il est en particulier possible de retrouver les premières propriétés géométriques mises en évidence plus haut. Ainsi

1. Montrer à l'aide de cette équation que l'axe focal Δ est un axe de symétrie de \mathcal{C} .

Soit $M = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{R}_F}$ un point de \mathcal{C} . Son symétrique M' par rapport à l'axe Δ_F a pour coordonnées

$$M' = \begin{pmatrix} x' \\ y' \end{pmatrix}_{\mathcal{R}_F} = \begin{pmatrix} x \\ -y \end{pmatrix}_{\mathcal{R}_F}$$

Montrer que le point M' appartient également à \mathcal{C} :

$$\begin{aligned}
 (1 - e^2)(x')^2 + (y')^2 - 2ep(x') - p^2 &= (1 - e^2)x^2 + (-y)^2 - 2epx - p^2 \\
 &= (1 - e^2)x^2 + y^2 - 2epx - p^2 \\
 &= 0
 \end{aligned}$$

car le point M appartient à \mathcal{C} . □

2. Montrer que si $e \neq 1$, \mathcal{C} admet deux sommets S et S' sur l'axe Δ et donner leurs coordonnées dans \mathcal{R}_F .

On souhaite déterminer les points $M = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{R}_F}$ du plan qui sont à l'intersection de la conique \mathcal{C} et de l'axe Δ_F dans le cas $e \neq 1$. Or les coordonnées d'un tel point doivent vérifier l'équation de \mathcal{C} et celle de Δ_F .

D'où le système

$$\begin{cases} (1 - e^2)x^2 + y^2 - 2epx - p^2 = 0 \\ y = 0 \end{cases} \iff \begin{cases} y = 0 & (E_1) \\ (1 - e^2)x^2 - 2epx - p^2 = 0 & (E_2) \end{cases}$$

La seconde équation de ce système est une équation polynomiale de degré 2.

On détermine ses racines à l'aide de son discriminant :

$$\begin{aligned} \Delta &= (-2ep)^2 - 4(1 - e^2)(-p^2) \\ &= \cancel{4e^2p^2} + 4p^2 - \cancel{4e^2p^2} \\ &= 4p^2 > 0 \end{aligned}$$

La seconde équation du système admet donc deux racines réelles, ce qui prouve l'existence des deux sommets.

Par ailleurs, ces deux racines réelles sont

$$\begin{aligned} x_1 &= \frac{2ep - \sqrt{4p^2}}{2(1 - e^2)} = \frac{p(e - 1)}{1 - e^2} = -\frac{p}{1 + e} \\ \text{et} \\ x_2 &= \frac{2ep + \sqrt{4p^2}}{2(1 - e^2)} = \frac{p(e + 1)}{1 - e^2} = \frac{p}{1 - e} \end{aligned}$$

Les deux sommets de \mathcal{C} situés sur l'axe focal sont donc les points de coordonnées

$$S_1 = \begin{pmatrix} -\frac{p}{1 + e} \\ 0 \end{pmatrix} \quad \text{et} \quad S_2 = \begin{pmatrix} \frac{p}{1 - e} \\ 0 \end{pmatrix}$$

Note : en étudiant le signe des abscisses x_1 et x_2 trouvées ci-dessus, on peut également déterminer la position de ces sommets par rapport au foyer. On retrouve le résultat observé géométriquement :

— Si $e < 1$ (i.e. si \mathcal{C} est une ellipse), les deux sommets encadrent le foyer F .

— Si $e > 1$ (i.e. si \mathcal{C} est une hyperbole), les deux sommets sont à gauche de F .

3. Montrer que si $e = 1$, \mathcal{C} admet un unique sommet S sur l'axe Δ et donner ses coordonnées dans \mathcal{R}_F .

Si $e = 1$, on a $p = ed = d$ et la seconde équation du système ci-dessus devient

$$(E_2) = -2dx - d^2 = 0 \iff x = -\frac{d}{2}$$

La parabole \mathcal{C} admet donc un unique sommet sur son axe focal, de coordonnées

$$S_0 = \begin{pmatrix} -\frac{d}{2} \\ 0 \end{pmatrix}_{\mathcal{R}_F}$$

4. Donner, dans le cas $e \neq 1$, les coordonnées dans \mathcal{R}_F du centre Ω de \mathcal{C} .

Si $e \neq 1$, les coordonnées du point $\Omega = \begin{pmatrix} x_\Omega \\ y_\Omega \end{pmatrix}_{\mathcal{R}_F}$ sont données par la moyenne des coordonnées des sommets S_1 et S_2 :

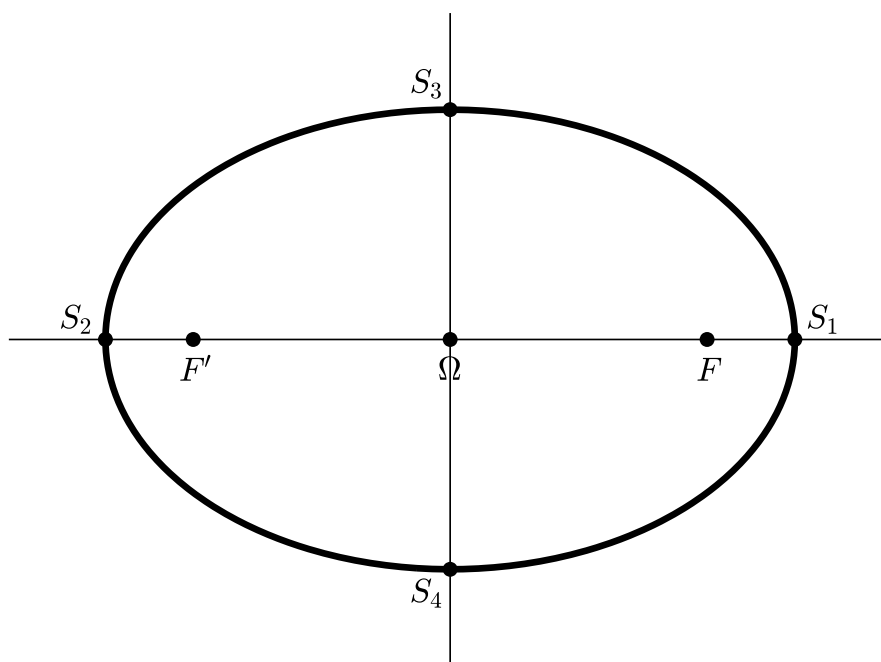
$$\begin{cases} x_\Omega = \frac{x_1 + x_2}{2} = \frac{ep}{1 - e^2} \\ y_\Omega = 0 \end{cases} \implies \Omega = \begin{pmatrix} \frac{ep}{1 - e^2} \\ 0 \end{pmatrix}_{\mathcal{R}_F}$$

Équation réduite d'une conique

Une fois que l'on a, à l'aide d'une première équation, pu établir les principales propriétés géométriques d'une conique \mathcal{C} donnée, il peut construire un second repère, mieux adapté à la conique étudiée, dans lequel l'équation de \mathcal{C} permet d'établir les dernières propriétés. Par ailleurs, dans ce repère, il est également possible d'associer les constantes définissant l'équation en termes encore une fois géométrique. Précisément,

Équation réduite d'une ellipse. Soient \mathcal{E} une ellipse du plan, Ω son centre de symétrie et $\mathcal{R}_\Omega = (x\Omega y)$ un repère orthonormé du plan, centré en Ω et dont l'axe des abscisses correspond à l'axe focal Δ de \mathcal{E} . Il existe deux constantes $a > b > 0$ telles que l'équation de \mathcal{E} dans \mathcal{R}_Ω est

$$(E_{\mathcal{E}}) : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$



Dans ce repère, les coordonnées des sommets de \mathcal{E} sont

$$S_1 = \begin{pmatrix} a \\ 0 \end{pmatrix} \quad S_2 = \begin{pmatrix} -a \\ 0 \end{pmatrix} \quad S_3 = \begin{pmatrix} 0 \\ b \end{pmatrix} \quad S_4 = \begin{pmatrix} 0 \\ -b \end{pmatrix}$$

Exercice : représenter les constantes a et b sur le dessin ci-dessus.

Par ailleurs, les foyers F et F' de \mathcal{E} étant situés sur son axe focal, leurs coordonnées dans \mathcal{R}_Ω sont de la forme

$$F = \begin{pmatrix} c \\ 0 \end{pmatrix} \quad F' = \begin{pmatrix} -c \\ 0 \end{pmatrix}$$

et l'on peut montrer que la constante c ainsi définie vérifie l'égalité

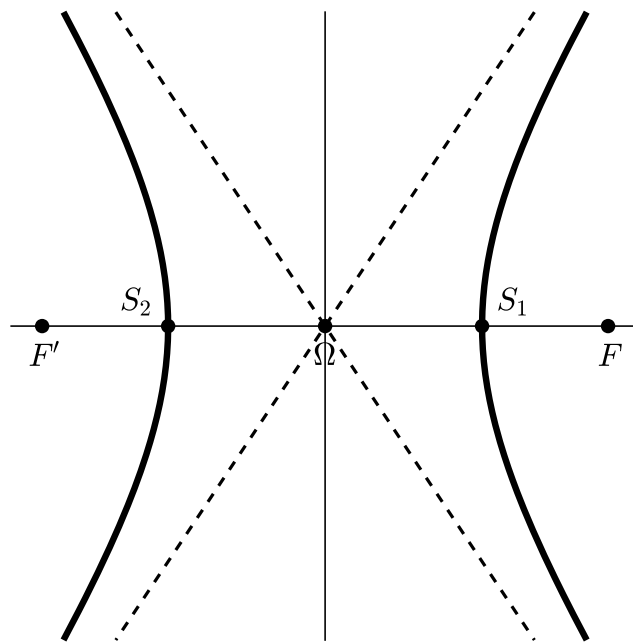
$$a^2 = b^2 + c^2$$

Exercice :

1. Représenter la constante $c > 0$ sur le dessin ci-dessus.
2. Placer sur le dessin ci-dessus un triangle rectangle illustrant l'égalité ci-dessus.

Équation réduite d'une hyperbole. Soient \mathcal{H} une hyperbole du plan, Ω son centre de symétrie et $\mathcal{R}_\Omega = (x\Omega y)$ un repère orthonormé du plan, centré en Ω et dont l'axe des abscisses correspond à l'axe focal Δ de \mathcal{H} . Il existe deux constantes $a, b > 0$ telles que l'équation de \mathcal{H} dans \mathcal{R}_Ω est

$$(E_{\mathcal{H}}) : \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$



Dans ce repère, les coordonnées des sommets de \mathcal{H} sont

$$S_1 = \begin{pmatrix} a \\ 0 \end{pmatrix} \quad S_2 = \begin{pmatrix} -a \\ 0 \end{pmatrix}$$

Par ailleurs, les foyers F et F' de \mathcal{E} étant situés sur son axe focal, leurs coordonnées dans \mathcal{R}_Ω sont de la forme

$$F = \begin{pmatrix} c \\ 0 \end{pmatrix} \quad F' = \begin{pmatrix} -c \\ 0 \end{pmatrix}$$

et l'on peut montrer que la constante c ainsi définie vérifie l'égalité

$$c^2 = a^2 + b^2$$

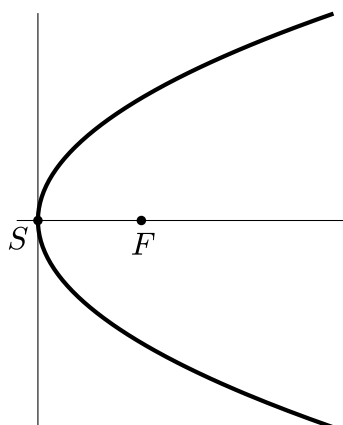
Enfin, on peut montrer que \mathcal{H} admet deux droites asymptotes d'équations

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0 \iff y = \pm \frac{b}{a}x$$

Équation réduite d'une parabole. Soient \mathcal{P} une parabole du plan, S son sommet et $\mathcal{R}_S = (xSy)$ un repère orthonormé du plan, centré en S et dont l'axe des abscisses correspond à l'axe focal Δ de \mathcal{P} . L'équation de \mathcal{P} dans \mathcal{R}_S est

$$(E_{\mathcal{P}}) : y^2 = 2px$$

où $p = d$ est le paramètre de \mathcal{P} .



Dans ce repère, les coordonnées du sommet de \mathcal{P} est

$$S = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

et les coordonnées de son foyer sont

$$F = \begin{pmatrix} \frac{p}{2} \\ 0 \end{pmatrix}$$

Tangentes à une conique

Dans ces repères adaptés, il est également plus simple d'exprimer et d'étudier les tangentes à une conique. On peut en particulier montrer que les équations de ces tangentes se déduisent simplement des équations réduites données ci-dessus. On peut alors démontrer

les propriétés géométriques évoquées en introduction et qui confèrent en particulier les propriétés des coniques en termes de réflexion d'ondes.

Tangentes d'une parabole. Soit \mathcal{P} une parabole du plan d'équation réduite

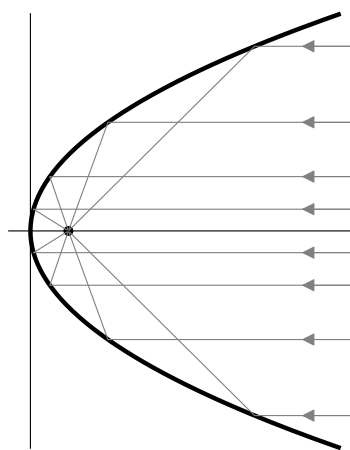
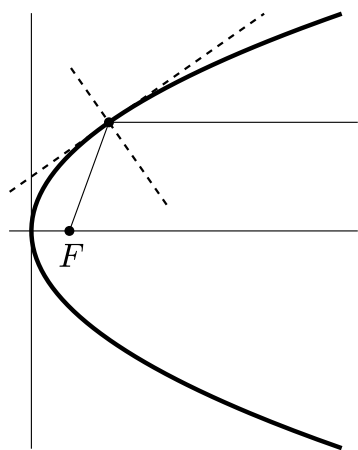
$$y^2 = 2px$$

En tout point $M_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ de \mathcal{P} , la tangente T_0 à \mathcal{P} est la droite d'équation

$$T_0 : yy_0 = p(x + x_0)$$

On peut alors montrer qu'en tout point M_0 , la tangente et la normale à \mathcal{C} en M_0 sont les bissectrices des angles formés par la droite (F, M_0) et la droite horizontale $y = M_0$.

Les lois de l'optique permettent alors de montrer que tout rayon entrant dans la parabole parallèlement à l'axe focal sont réfléchies au foyer et que tout rayon émis au foyer est réfléchi parallèlement à l'axe focal.



Cette propriété est exploitée en pratique pour concentrer un faisceau d'ondes parallèles en un point (antennes, four solaire, etc) ou à construire un faisceau cohérent d'ondes parallèles à partir d'une source ponctuelle (phares, laser, etc).

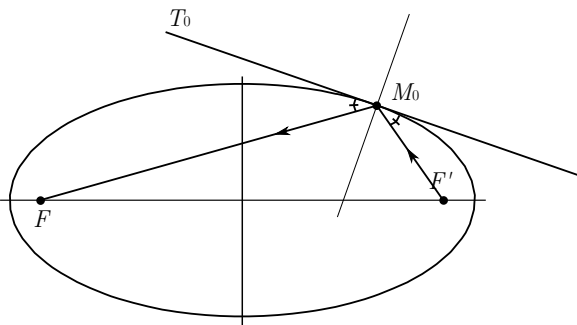
Tangente à un ellipse. Soit \mathcal{E} une ellipse du plan, d'équation réduite

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

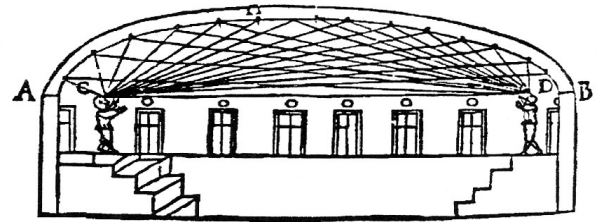
En tout point $M_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ de \mathcal{E} , la tangente T_0 à \mathcal{E} est la droite d'équation

$$T_0 : \frac{x \cdot x_0}{a^2} + \frac{y \cdot y_0}{b^2} = 1$$

On peut alors montrer qu'en tout point M_0 , la normale à \mathcal{C} en M_0 est la bissectrice de l'angles $\widehat{FM_0F'}$, F et F' étant les foyers de \mathcal{C} .



Les lois de l'optique permettent alors de montrer que tout rayon issu de l'un des foyers est renvoyé vers l'autre foyer.



Cette propriété peut être mise à profit pour focaliser des rayons émis d'une source ponctuelle en un autre point de l'espace (comme illustré ci-dessus).

De même, dans certaines stations de métro parisiennes, les quais sont sous une voûte elliptique. Précisément, une coupe transversale d'une station de ce type produit une ellipse dont les foyers étaient des lieux de communications privilégiés entre les poinçonneurs placés de part et d'autre des rails.



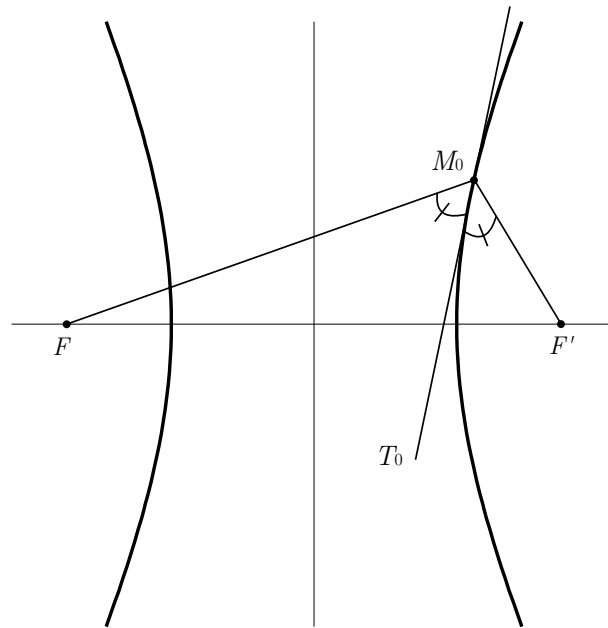
Tangentes à une hyperbole. Soit \mathcal{H} une hyperbole du plan, d'équation réduite

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

En tout point $M_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ de \mathcal{H} , l'équation T_0 à \mathcal{H} est la droite d'équation

$$T_0 : \frac{x x_0}{a^2} - \frac{y y_0}{b^2} = 1$$

On obtient là encore une propriété portant sur les angles formés entre la tangente, la normale et les droites liant M_0 aux foyers : sur une hyperbole de foyers F et F' , la tangente en un point M_0 est la bissectrice intérieure de l'angle $\widehat{FM_0F'}$.



3.2.2 Coniques analytiques

Définition analytique d'une conique

L'approche analytique entamée ci-dessus permet de généraliser la notion de conique via leurs équations.

Définition (conique analytique)

Dans le plan, muni d'un repère orthonormé $\mathcal{R} = (O, \vec{i}, \vec{j})$, on appelle *conique du plan* toute courbe \mathcal{C} du plan dont l'équation dans \mathbb{R} est définie par un polynôme de degré deux :

$$\mathcal{C} : ax^2 + bxy + cy^2 + fx + gy + h = 0, a, b, c, f, g, h \in \mathbb{R}$$

Cette définition, plus générale, permet, pour certaines valeurs des coefficients a à h , de représenter l'ensemble des coniques géométriques, i.e. l'ensemble des coniques définies par un point, une droite et une excentricité. Ainsi, l'équation d'une conique (F, \mathcal{D}, e) dans son repère \mathcal{R}_F correspond à

$$a = 1 - e^2, \quad b = 0, \quad c = 1, \quad f = -2ep, \quad g = 0, \quad h = -p^2$$

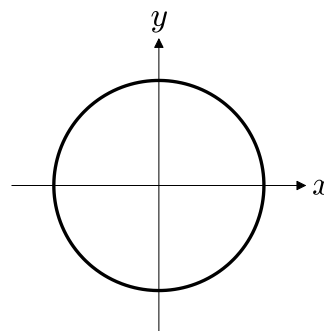
Mais d'autres courbes du plan entrent ainsi dans la famille des coniques.

1. **Les cercles** (qui ne font pas partie des coniques géométriques). Par exemple, le cercle de centre O et de rayon R a pour équation

$$\mathcal{C}_R : x^2 + y^2 = R^2$$

et correspond aux coefficients

$$\begin{aligned} a &= 1, & b &= 0, & c &= 1, \\ f &= 0, & g &= 0, & h &= -R^2 \end{aligned}$$

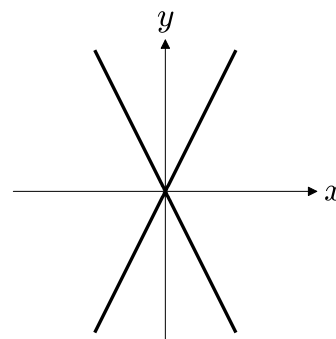


2. **Les droites et couples de droites**, sécantes ou parallèles. Par exemple, la réunion de deux droites linéaires du plan est définie par une équation de la forme

$$\mathcal{D}_1 \cup \mathcal{D}_2 : (a_1x + b_1y)(a_2x + b_2y) = 0$$

et correspond aux coefficients

$$\begin{aligned} a &= a_1a_2, & b &= a_1b_2 + a_2b_1, \\ c &= b_1b_2, \\ f &= g = h = 0 \end{aligned}$$



3. **Les points**. Par exemple le centre O est défini par l'équation

$$\{O\} : x^2 + y^2 = 0$$

et correspond aux coefficients

$$a = 1, \quad b = 0, \quad c = 1, \quad f = g = h = 0$$

4. **L'ensemble vide**, défini par exemple par l'équation

$$\{O\} : -x^2 - y^2 = 0$$

et correspond aux coefficients

$$a = -1, \quad b = 0, \quad c = -1, \quad f = g = h = 0$$

On peut montrer que ce sont là les seuls ensembles du plan que l'on peut décrire à l'aide d'équations de degré 2. Dans ce cadre, les coniques sont également appelées *courbes de degré 2*.

Exercice : pour chacune nouvelles coniques \mathcal{C} ci-dessus,

1. Déterminer si possible une configuration plan de l'espace / cône de l'espace correspondant.
2. Déterminer la classe (ellipse, parabole, hyperbole) peut la classer.

Équations réduite d'une conique analytique

L'étude des propriétés algébriques des polynômes de degré 2 en deux variables permet de retrouver la des différentes coniques du plan (ellipses/paraboles/hyperboles) et de caractériser ces classes en fonction des coefficients du polynôme donnant l'une des ses équations.

Cette classification s'appuie en particulier sur la notion générale d'*équation réduite*.

Définition (équation réduite d'une conique analytique)

On appelle *équation réduite d'une conique* toute équation de la forme

$$(E_1) : \alpha.x^2 + \beta.y^2 = 1, \quad \alpha, \beta \in \mathbb{R}^*$$

$$(E'_1) : \alpha.x^2 + \beta.y^2 = 0 \quad \alpha, \beta \in \mathbb{R}^*$$

$$(E_2) : y^2 = 2\alpha x \quad \alpha \in \mathbb{R}$$

Classification des coniques analytiques

À partir de la notion d'équation réduite, on peut établir une classification précise des différents types de coniques, notamment basée sur le signes des coefficients α et β qui peuvent intervenir dans ces équations réduites. Précisément,

Théorème

Soit \mathcal{C} une conique du plan. Il existe une infinité de bases du plan dans lesquelles \mathcal{C} admet une équation réduite. Par ailleurs,

1. Si \mathcal{C} admet une équation réduite du type (E_1) ou (E'_1) , alors toutes ses équations réduites sont de ce type, et les coefficients α et β associés à ces équations réduites sont de mêmes signes et déterminent la nature de la conique \mathcal{C} :
 - Si α et β sont de même signe, la conique \mathcal{C} est **une ellipse**.
 - Si α et β sont de signes contraires, la conique \mathcal{C} est **une hyperbole**.
2. Si \mathcal{C} admet une équation réduite du type (E_2) , alors toutes ses équations réduites sont de ce type et la conique \mathcal{C} est **une parabole**.

3.2.3 Réduction d'une équation

La théorie des polynômes de degré 2 permet de déterminer, à partir d'une équation de la forme

$$ax^2 + bxy + cy^2 + fx + gy + h = 0$$

valable dans un repère orthonormé $\mathcal{R} = (O; \vec{i}, \vec{j})$ du plan, de déterminer les caractéristiques géométrique de la conique \mathcal{C} associée.

Il est en particulier possible de montrer que la nature (ellipse/parabole/hyperbole) ne dépend que de la partie *quadratique* de l'équation :

$$q(x, y) = ax^2 + bxy + cy^2$$

Précisément, à partir de transformations algébrique de l'expression $q(x, y)$, on peut construire des changements de variables linéaires

$$\begin{cases} \tilde{x} &= rx + sy \\ \tilde{y} &= ux + vy \end{cases}$$

permettant de transformer $q(x, y)$ en une forme quadratique de la forme

$$\tilde{q}(\tilde{x}, \tilde{y}) = \alpha\tilde{x}^2 + \beta\tilde{y}^2$$

dite *forme réduite* de $q(x, y)$.

On peut en outre montrer que quelque soient les transformation permettant d'obtenir cette forme réduite, le signe du produit $\alpha\beta$ des coefficients obtenus, caractéristique de la nature de la conique \mathcal{C} , reste inchangé.

Historiquement, la première méthode de transformation permettant d'obtenir une de ces formes réduites, connue sous le nom de *méthode de Gauss* est basée sur la *mise sous forme canonique* de la forme quadratique $q(x, y)$. Cependant, cette méthode comporte certains inconvénients, notamment du point de vue géométrique. On lui préférera donc la méthode *matricielle*, basée sur la diagonalisation des matrices symétriques.

Définition (matrice d'une forme quadratique)

Soit

$$q(x, y) = ax^2 + bxy + cy^2$$

une forme quadratique. On appelle *matrice de q* la matrice

$$A_q = \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}$$

Dans un premier temps, cette matrice permet de mettre en place la forme matricielle de la forme quadratique q . Précisément, on peut montrer que

$$\forall X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, \quad q(x, y) = {}^tX \times A \times X$$

Exemple : la matrice associée à la forme quadratique

$$q(x, y) = 3x^2 + 4xy + 5y^2$$

est

$$A_q = \begin{pmatrix} 3 & 2 \\ 2 & 5 \end{pmatrix}$$

En notant $X = \begin{pmatrix} x \\ y \end{pmatrix}$, on a alors

$$\begin{aligned}
{}^tX A_q X &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\
&= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 3x + 2y \\ 2x + 5y \end{pmatrix} \\
&= x(3x + 2y) + y(2x + 5y) = 3x^2 + 4xy + 5y^2
\end{aligned}$$

De façon générale, on note alors que les formes quadratiques réduites sont exactement les formes quadratiques associées à des matrices *diagonales*.

En s'appuyant alors sur les éléments propres (valeurs propres et vecteurs propres) de la matrice A_q , on peut alors construire un changement de variable transformant $q(x, y)$ en une forme réduite.

En effet, A_q étant une matrice *symétrique*, il existe une matrice P *orthogonale* (i.e. telle que ${}^tP = P^{-1}$) telle que

$${}^tP \times A \times P = D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

D'un point de vue géométrique, la matrice P peut être associée à un changement de base dans le plan muni du point O . En notant $\tilde{X} = \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix}$ les coordonnées dans ce nouveau repère, on a

$$X = P\tilde{X}$$

et

$$\begin{aligned}
q(x, y) &= {}^tX \times A_q \times X \\
&= {}^t(P\tilde{X}) \times A_q \times (P\tilde{X}) \\
&= {}^t\tilde{X} ({}^tP A_q P) \tilde{X} \\
&= \begin{pmatrix} \tilde{x} & \tilde{y} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} \\
&= \lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2
\end{aligned}$$

On peut ainsi montrer que, finalement, la nature de la conique \mathcal{C} dépend directement du signe des valeurs propres λ_1 et λ_2 de A_q . Précisément,

- Si λ_1 et λ_2 sont non nulles et de même signe, \mathcal{C} est **une ellipse**.
- Si λ_1 et λ_2 sont non nulles et de signes contraires, \mathcal{C} est **une hyperbole**.
- Si λ_1 ou λ_2 est nulle, \mathcal{C} est **une parabole**.

Notons enfin que, puisque

$$\det(A) = ac - \frac{b^2}{4} = \lambda_1 \cdot \lambda_2$$

il est également possible de déduire la nature de \mathcal{C} à partir du discriminant

$$\Delta = b^2 - 4ac$$

de la forme quadratique $q(x, y)$. Précisément,

- \mathcal{C} est une ellipse si et seulement si $\Delta < 0$.
- \mathcal{C} est une hyperbole si et seulement si $\Delta > 0$.
- \mathcal{C} est une parabole si et seulement si $\Delta = 0$.

D'un point de vue géométrique, on peut montrer que le changement de repère associé à la matrice P issue de la diagonalisation de A_q permet en réalité d'aligner les axes du repère sur les directions caractéristiques de la conique \mathcal{C} , i.e. son axe focal et sa directrice.

En injectant le changement de variable induit par la matrice P dans l'équation complète de \mathcal{C} , on obtient, dans ce nouveau repère, une équation de la forme

$$\lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 + \gamma \tilde{x} + \delta \tilde{y} + \mu = 0$$

On peut alors terminer la réduction de cette équation en mettant les polynômes

$$\lambda_1 \tilde{x}^2 + \gamma \tilde{x} \quad \text{et} \quad \lambda_2 \tilde{y}^2 + \delta \tilde{y}$$

sous forme canonique.

3.3 Généralisation aux dimensions supérieures

3.3.1 Formes quadratiques de \mathbb{R}^n

Les notions de polynômes et de matrices permettent de généraliser la notion de conique en dimensions supérieures. Précisément,

Définition (forme quadratique)

Soit $n \in \mathbb{N}^{\geq 2}$. On appelle *forme quadratique de \mathbb{R}^n* tout polynôme *homogène* de degré 2 en les n variables $X = (x_1, \dots, x_n)$:

$$q(X) = q(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=i}^n a_{ij} x_i x_j \quad (1)$$

$$= \sum_{i=1}^n a_{ii} x_i^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^n a_{ij} x_i x_j \quad (2)$$

Exercice : écrire la forme quadratique suivante sous la forme (1) puis sous la forme (2) :

$$q(x_1, x_2, x_3, x_4) = x_1^2 + x_1 x_2 + x_2^2 + x_2 x_3 + x_3^2 + x_3 x_4 + x_4^2 + x_1 x_4 + x_1 x_3 + x_2 x_4.$$

3.3.2 Réduction matricielle des formes quadratiques

Il est possible de généraliser l'ensemble de la théorie permettant la réduction des formes quadratiques en deux variables aux formes quadratiques de \mathbb{R}^n . Précisément, pour une forme quadratique

$$q(x_1, \dots, x_n) = \sum_{i=1}^n a_{ii} x_i^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^n a_{ij} x_i x_j$$

on appelle *matrice de q* la matrice symétrique

$$A_q = \begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \cdots & \frac{a_{1n}}{2} \\ \frac{a_{12}}{2} & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{a_{n-1,n}}{2} \\ \frac{a_{1n}}{2} & \cdots & \frac{a_{n-1,n}}{2} & a_{nn} \end{pmatrix}$$

Là encore, via les éléments propres de la matrice A_q , il est possible d'associer à $q(x_1, \dots, x_n)$ une forme quadratique réduite, constituée d'une somme de carrés

$$\tilde{q}(\tilde{X}) = \sum_{i=1}^n \alpha_i \tilde{x}_i^2$$

3.3.3 Quadriques de l'espace

Classification

Dans le cas $n = 3$, la notion de polynômes de degré 2 en (x, y, z) permet de définir une famille de *surfaces de l'espace* généralisant la notion de conique à la dimension 3 : les quadriques.

Définition (quadrique de l'espace)

On se place dans l'espace muni d'un repère orthonormé \mathcal{R} . On appelle *quadrique* toute surface \mathcal{Q} de l'espace dont l'équation dans \mathcal{R} est de la forme

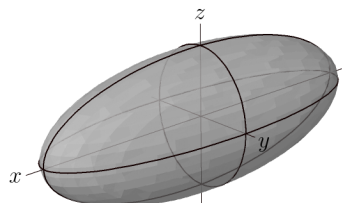
$$\mathcal{Q} : P(x, y, z) = 0$$

où $P(x, y, z)$ est un polynôme de degré 2.

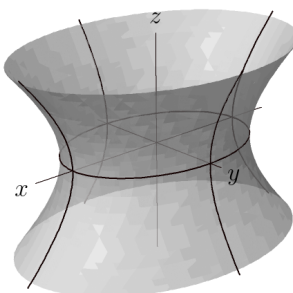
Il est alors également possible d'étendre la notion d'*équation réduite* aux quadriques et l'étude des formes quadratiques en trois variables permet d'établir une classification de ces surfaces, en fonction de la réduction obtenue.

Précisément, il existe essentiellement six types distincts de quadriques de l'espace, reconnaissables à la forme de leur équation dans un repère adapté :

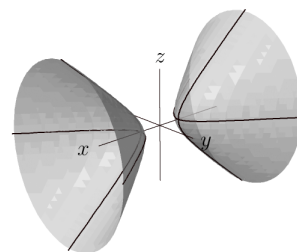
Ellipsoïde



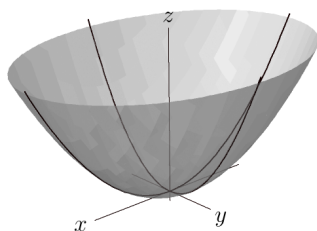
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

Hyperboloïde
à une nappe

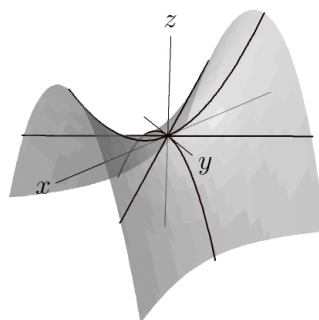
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

Hyperboloïde
à deux nappes

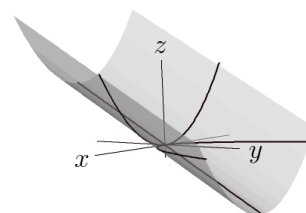
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

Paraboloïde
elliptique

$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$$

Paraboloïde
hyperbolique

$$z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$$

Cylindre
parabolique

$$z = \frac{x^2}{a^2}$$

Chapitre 4

Séries numériques, séries entières

Introduction

D'un point de vue formel, les séries numériques sont des suites numériques particulières, ou plus précisément une façon d'aborder la notion de suite numérique d'un autre point de vue, principalement basé sur la notion de somme.

Formellement, la notion de série numérique permet de donner un sens à la somme d'une infinité de réels (ou complexes).

D'un point de vue pratique, les séries permettent de mettre en évidence de nouveaux objets mathématiques pouvant servir de base à la modélisation mathématique de différents problèmes issus "du monde réel".

Nous verrons ensuite comment les séries permettent de représenter les fonctions réels via une généralisation de la notion de développement limité.

4.1 Définitions

4.1.1 Sommes partielles

Définition (suite des sommes partielles)

Soit (u_n) une suite de nombres réels ou complexes. On appelle *série de terme général* u_n , notée série $\sum u_n$ la suite (S_n) définie par

$$\forall n \in \mathbb{N}, \quad S_n = \sum_{k=0}^n u_k = u_0 + u_1 + \dots + u_n.$$

Par ailleurs,

- Le nombre u_n est alors le *terme général* d'indice n de la série $\sum u_n$.
- La somme S_n est alors la *somme partielle d'ordre n* de la série $\sum u_n$.

Note : si la suite (u_n) n'est définie qu'à partir d'un rang $n_0 \in \mathbb{N}$, la somme partielle d'ordre n de la série $\sum u_n$ est alors $S_n = \sum_{k=n_0}^n u_k$.

4.1.2 Convergence

Une série étant en particulier une suite numérique, on peut naturellement lui attacher la notion de convergence. Ainsi,

Définition (Convergence d'une série numérique)

Soit (u_n) une suite numérique. La série $\sum u_n$ associée est dite *convergente* si la suite (S_n) converge.

On note alors

$$S = \lim_{n \rightarrow +\infty} S_n = \sum_{n=0}^{+\infty} u_n$$

et S est appelée *somme de la série* $\sum u_n$.

Enfin, si la série $\sum u_n$ ne converge pas, elle est dite *diverge*.

Notes :

- La nature (convergente ou divergente) de la série $\sum u_n$ ne dépend pas des premiers termes de la suite (u_n) . Cependant, si la série $\sum u_n$ converge, les premiers termes de la suite (u_n) influencent la valeur de la somme.

— Si la série $\sum u_n$ diverge, l'écriture $\sum_{n=0}^{+\infty} u_n$ n'a pas de sens. Cette quantité n'existe pas.

Définition (reste d'une série convergente)

Soit (u_n) une suite de nombres telle que la série $\sum u_n$ converge. On appelle *reste d'ordre n de la série $\sum u_n$* la somme infinie

$$R_n = \sum_{k=n+1}^{+\infty} u_k$$

En notant S la somme de la série $\sum u_n$, on a alors

$$\forall n \in \mathbb{N}, \quad S = \sum_{n=0}^{+\infty} u_n = S_n + R_n$$

Exercice : Soit (S_n) la série convergente, de terme général u_n et de somme $S = \sum_{n=0}^{+\infty} u_n$, représentée ci-dessous.

$$S = \underbrace{u_0}_{\text{---}} + u_1 + \dots + u_n + u_{n+1} + u_{n+2} + u_{n+3} + \dots$$

⏟

⏟

⏟

⏟

1. Compléter le dessin ci-dessus à l'aide de sommes partielles.
2. Ajouter au dessin une représentation des *restes* de la série $\sum u_n$.

La représentation ci-dessus illustre en particulier le fait que la suite du reste (R_n)

d'une série convergente est également convergente et vérifie

$$\lim_{n \rightarrow +\infty} R_n = 0$$

Ainsi, l'étude du reste d'une série convergente est un outil important dans l'étude d'une série. Elle permet en particulier d'évaluer *la vitesse* à laquelle (S_n) tend vers sa somme S . En effet, les deux suites (S_n) et (R_n) convergent à la même vitesse.

Il est alors possible, en comparant la suite (R_n) à un catalogue de suites tendant vers 0, de placer celle-ci sur une échelle de vitesses.

La première partie de ce cours est dédiée à la présentation de *critères de convergence* pour les séries numériques, i.e. de propriétés portant sur la suite (u_n) assurant la convergence de la suite (S_n) .

Voici un premier critère, donnant une *condition nécessaire* à la convergence d'une série :

Proposition (Condition nécessaire de convergence d'une série numérique)

Pour qu'une série $\sum u_n$ converge, il est *nécessaire* que

$$\lim_{n \rightarrow +\infty} u_n = 0$$

Dans le cas contraire, on dit que la série $\sum u_n$ *diverge grossièrement*.

Attention : cette condition n'est pas suffisante. Il existe en particulier des suites qui tendent vers 0 et dont la série associée diverge.

4.2 Deux classiques

4.2.1 Les séries géométriques

Définition (série géométrique)

On appelle *série géométrique* toute série dont le terme général est de la forme

$$u_n = \alpha \cdot q^n, \quad \alpha, q \in \mathbb{C} \text{ fixés}$$

On a alors

$$\forall n \in \mathbb{N}, \quad S_n = \sum_{k=0}^n \alpha \cdot q^k = \alpha \cdot \frac{1 - q^{n+1}}{1 - q}$$

Exercice : à démontrer.

Les séries géométriques font donc partie des (rares) séries pour lesquelles on sait rapidement déterminer une forme explicite pour la suite (S_n) de ses sommes partielles. Cette forme explicite permet en particulier de déterminer les conditions de convergence et la vitesse de convergence le cas échéant. Précisément, la série $\sum \alpha \cdot q^n$ converge si et seulement si

$$|q| < 1$$

Dans ce cas, on a alors

$$S = \sum_{n=0}^{+\infty} \alpha \cdot q^n = \alpha \cdot \frac{1}{1 - q}$$

et

$$R_n = \sum_{k=n+1}^{+\infty} \alpha \cdot q^k = \alpha \cdot \frac{q^{n+1}}{1 - q}$$

On peut en particulier en conclure que la convergence d'une série géométrique se fait à une vitesse géométrique (R_n tend vers 0 à la vitesse de q^{n+1}).

Exemples :

— Soit $(u_n) : u_n = \left(\frac{1}{2}\right)^2 = \frac{1}{2^n}$.

$$\forall n \in \mathbb{N}, \quad S_n = \sum_{k=0}^n \frac{1}{2^k} = \frac{1 - \frac{1}{2^{n+1}}}{1 - \frac{1}{2}} = 2 - \frac{1}{2^n}$$

Ainsi, la série $\sum \frac{1}{2^n}$ converge et

$$S = 2 \quad \text{et} \quad R_n = \frac{1}{2^n}$$

— Soit $(v_n) : v_n = 2^n$.

$$\forall n \in \mathbb{N}, \quad S_n = \sum_{k=0}^n 2^k = \frac{1 - 2^{n+1}}{1 - 2} = 2^{n+1} - 1 \xrightarrow{n \rightarrow +\infty} +\infty$$

et la série $\sum 2^n$ diverge

4.2.2 Les séries télescopiques

Définition (série télescopique)

On appelle *série télescopique* toute série dont le terme général u_n est de la forme

$$u_n = v_{n+1} - v_n$$

On a alors

$$\forall n \in \mathbb{N}, \quad S_n = \sum_{k=0}^n v_{k+1} - v_k = v_{n+1} - v_0 \quad (*)$$

Exercice : à démontrer.

On peut alors montrer que la série $\sum u_n$ converge si et seulement si

la suite (v_n) converge.

En notant alors ℓ la limite de la suite (v_n) , on a

$$S = \sum_{n=0}^{+\infty} v_{n+1} - v_n = \ell - v_0$$

et le reste d'ordre n d'une série télescopique est alors

$$R_n = \sum_{k=n+1}^{+\infty} v_{k+1} - v_k = \ell - v_{n+1}$$

On peut en particulier en conclure que la convergence d'une série télescopique se fait à la même vitesse que la convergence de la suite (v_n) associée.

Exemple : soit $(u_n) : u_n = \frac{1}{n(n+1)}$. Pour tout $n \geq 1$, on a

$$u_n = \frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$$

Ainsi,

$$\forall n \in \mathbb{N}^*, \quad S_n = \sum_{k=1}^n \frac{1}{k} - \frac{1}{k+1} = 1 - \frac{1}{n+1}$$

On en conclue que la série $\sum \frac{1}{n(n+1)}$ converge et

$$S = \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = 1 \quad \text{et} \quad R_n = \frac{1}{n+1}$$

Note : la notion de série télescopique permet en outre de montrer que la notion de série n'est qu'une façon différente d'aborder la notion de suite. En effet, la formule (*) s'écrit également

$$\forall n \in \mathbb{N}, \quad v_n = v_0 + \sum_{k=0}^{n-1} v_{k+1} - v_k$$

Autrement dit, toute suite (v_n) peut s'écrire sous la forme d'une série.

4.3 Opérations sur l'ensemble des séries

Proposition (somme de séries convergentes)

Soient (u_n) et (v_n) deux suites telles que les séries associées convergent. Alors la série $\sum(u_n + v_n)$ est également convergente et

$$\sum_{n=0}^{+\infty} (u_n + v_n) = \sum_{n=0}^{+\infty} u_n + \sum_{n=0}^{+\infty} v_n$$

Proposition (multiplication par un scalaire)

Soient (u_n) une suite numérique et $\alpha \in \mathbb{C}$ un scalaire fixé.

- Les séries $\sum u_n$ et $\sum \alpha.u_n$ sont de même nature.

— Si les deux séries ci-dessus convergent, alors

$$\sum_{n=0}^{+\infty} \alpha.u_n = \alpha. \sum_{n=0}^{+\infty} u_n$$

Note : formellement, on peut envisager la notion de série pour toute suite *complexe*. Cependant, si (z_n) est une suite complexe, en notant

$$z_n = u_n + iv_n, \quad u_n, v_n \in \mathbb{R}$$

on peut montrer que la série $\sum z_n$ converge si et seulement si les séries $\sum u_n$ et $\sum v_n$ convergent. D'autre part, on a dans ce cas

$$\sum_{n=0}^{+\infty} z_n = \sum_{n=0}^{+\infty} u_n + i \sum_{n=0}^{+\infty} v_n$$

On peut ainsi réduire l'étude des série numériques à l'étude des séries *réelles*.

4.4 Les séries alternées

Le cœur de la théorie des séries numériques porte sur la mise en place de *critères de convergence*. Précisément, il s'agit de déterminer les caractéristiques que doit vérifier une suite (u_n) pour que la série $\sum u_n$ converge. On verra dans les sections suivantes que la plupart de ces critères portent sur les séries à *termes positifs*. Avant cela, on peut énoncer un premier résultat :

Théorème (critère des séries alternées)

Soit (u_n) une suite numérique telle que

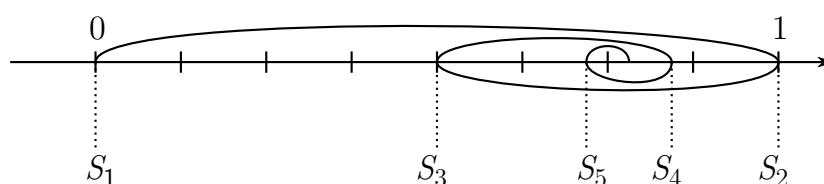
- Pour tout $n \in \mathbb{N}$, u_n et u_{n+1} sont de signes contraires.
- La suite $(|u_n|)$ décroît et tend vers 0.

Alors

- La série $\sum u_n$ converge.
- Le reste R_n de la série $\sum u_n$ vérifie

$$|R_n| \leq |u_{n+1}|$$

Exemple : pour tout $n \in \mathbb{N}^*$, on note $u_n = \frac{(-1)^n}{n}$ et (S_n) la suite des sommes partielles associée. En plaçant les différentes valeurs de (S_n) sur une droite, on constate que la série $\sum u_n$ converge :



Outre cet argument géométrique, on peut montrer d'un point de vue analytique que la série ci-dessus converge. Cependant, ces démonstrations ne permettent pas toujours de déterminer la valeur exacte de la somme.

4.5 Séries à termes positifs

4.5.1 Définition

Une série $\sum u_n$ est dite à *termes positifs* si

$$\forall n \in \mathbb{N}, \quad u_n \geq 0$$

Par abus de langage, on dira encore que la série est à termes positifs si $u_n \geq 0$ à partir d'un certain rang.

La propriété principale d'une série à termes positifs est

Proposition

Une série $\sum u_n$ est à termes positifs si et seulement si la suite (S_n) de ses sommes partielles est **croissante**.

Pour l'étude de ces séries, il est alors possible de d'appliquer les résultats spécifiques aux suites *monotones* :

Proposition

Soit (u_n) une suite à termes positifs et (S_n) la suite de ses sommes partielles.

- Soit (S_n) est majorée et alors elle converge.
- Soit (S_n) diverge vers $+\infty$.

Par ailleurs, les résultats que l'on pourra établir sur les séries à termes positifs permettent, dans certains cas, d'obtenir de l'information sur certaines séries à termes de signes quelconques grâce au résultat ci-dessous :

Définition/Théorème (séries absolument convergentes)

Soit (u_n) une suite numérique quelconque. La série $\sum u_n$ est dite *absolument convergente* si la série $\sum |u_n|$ converge.

Toute série absolument convergente **converge**.

Note : la notion de convergence absolue donne alors un critère *suffisant* de convergence.

4.5.2 Critères de comparaison

On peut alors mettre en place des critères de convergence pour les séries à termes positifs, basés sur des comparaisons entre séries :

Proposition (critère de majoration)

Soient (u_n) et (v_n) deux suites à termes positifs telles que

$$\forall n \in \mathbb{N}, \quad u_n \leq v_n \quad (*)$$

Alors

— Si la série $\sum v_n$ converge, la série $\sum u_n$ converge également et

$$\sum_{n=0}^{+\infty} u_n \leq \sum_{n=0}^{+\infty} v_n$$

— Si la série $\sum u_n$ diverge, alors la série $\sum v_n$ diverge également.

Preuve : il suffit de montrer le premier point, le second étant sa contraposée. D'après la relation (*), on a

$$\forall n \in \mathbb{N}, \quad \sum_{k=0}^n u_k \leq \sum_{k=0}^n v_k$$

Si la série $\sum v_n$ converge, puisqu'elle est à termes positifs, en notant S' sa somme, on a

$$\sum_{k=0}^n v_k \leq \sum_{n=0}^{+\infty} v_n = S'$$

La suite (S_n) des sommes partielles de la série $\sum u_n$ est donc également majorée par S' . Étant croissante, elle est convergente.

Notes :

- Le résultat reste valable si la relation (*) est vérifiée à partir d'un certain rang n_0 .
- Ce critère permet, le cas échéant, d'établir la nature convergente d'une série donnée, mais ne permet pas, en général, de déterminer la valeur de la somme $S = \sum_{n=0}^{+\infty} u_n$.

Proposition (critère de domination)

Soient (u_n) et (v_n) deux suites à termes positifs telles que $u_n = o(v_n)$. Alors

- Si la série $\sum v_n$ converge, la série $\sum u_n$ converge également.
- Si la série $\sum u_n$ diverge, alors la série $\sum v_n$ diverge également.

Preuve : par définition, on a

$$u_n = o(v_n) \iff \frac{u_n}{v_n} \xrightarrow{n \rightarrow +\infty} 0$$

On en déduit que

$$\exists n_0 \in \mathbb{N} / \forall n \geq n_0, \quad \frac{u_n}{v_n} \leq 1 \iff u_n \leq v_n$$

On peut alors appliquer le critère précédent.

Note : là encore, ce critère permet d'établir la convergence d'une série donnée, mais ne permet pas, en général, de calculer sa somme.

Proposition (critères d'équivalence)

Soient (u_n) et (v_n) deux séries à termes positifs telles que $u_n \sim v_n$.

Alors les séries $\sum u_n$ et $\sum v_n$ sont de même nature.

Par ailleurs,

- Si les séries $\sum u_n$ et $\sum v_n$ convergente, alors

$$R_n = \sum_{k=n+1}^{+\infty} u_k \sim R'_n = \sum_{k=n+1}^{+\infty} v_k$$

- Si les séries $\sum u_n$ et $\sum v_n$ divergente, alors

$$S_n = \sum_{k=0}^n u_k \sim S'_n = \sum_{k=0}^n v_k$$

Preuve : à partir de la définition de l'équivalence, on peut établir des encadrement,

valables à partir d'un certain rang, de la forme

$$a.u_n \leq v_n \leq b.u_n$$

On conclue alors en appliquant le critère de majoration. Notons que là encore, l'encadrement n'étant valable qu'à partir d'un certain rang, ce critère ne donne aucune information sur la valeur des sommes éventuelles.

Note : tout ce qui précède s'applique également aux séries à termes négatifs et plus généralement aux séries dont le terme général garde un signe constant à partir d'un certain rang.

4.5.3 Séries de Riemann

Pour exploiter les critères de comparaison ci-dessus, il est alors nécessaire d'avoir une bibliothèque de séries de référence, auxquelles on pourra alors comparer les séries qui nous intéressent. Les premières séries de références sont alors les séries géométriques (de raison $q \geq 0$) pour lesquelles on sait exactement ce qu'il se passe.

Il existe une seconde famille de séries de référence : les séries de Riemann.

Définition (séries de Riemann)

On appelle *Série de Riemann* toute série numérique dont le terme général est de la forme

$$u_n = \frac{1}{n^\alpha}, \quad \alpha \in \mathbb{R} \text{ fixé}$$

Théorème (convergence des séries de Riemann)

Une série de Riemann $\sum \frac{1}{n^\alpha}$ converge si et seulement si $\alpha > 1$.

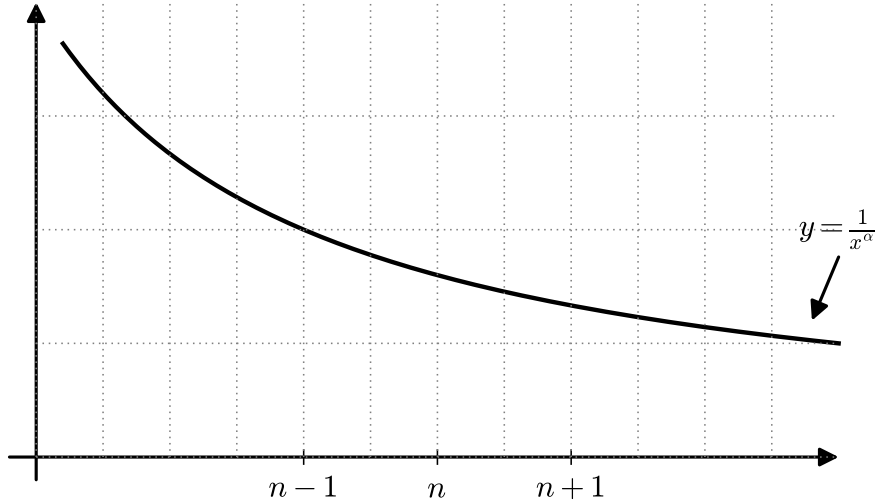
Preuve :

— Si $\alpha \leq 0$, on a

$$\lim_{n \rightarrow +\infty} \frac{1}{n^\alpha} = +\infty$$

et la série $\sum \frac{1}{n^\alpha}$ **diverge grossièrement**.

— Si $\alpha > 0$, l'idée générale consiste à comparer les sommes partielles de la série $\sum u_n$ avec l'intégrale sur $[1, n]$ de la fonction $f_\alpha : t \mapsto \frac{1}{t^\alpha}$:



D'après le graphe ci-dessus, on montre en effet que pour tout $n \geq 2$, on a

$$\int_n^{n+1} \frac{1}{t^\alpha} dt \leq \frac{1}{n^\alpha} \leq \int_{n-1}^n \frac{1}{t^\alpha} dt$$

Or si $\alpha \neq 1$, une primitive de f_α est

$$F_\alpha : t \mapsto \frac{1}{(1-\alpha)t^{\alpha-1}}$$

Ainsi

$$\frac{1}{1-\alpha} \left(\frac{1}{(n+1)^{\alpha-1}} - \frac{1}{n^{\alpha-1}} \right) \leq \frac{1}{n^\alpha} \leq \frac{1}{1-\alpha} \left(\frac{1}{n^{\alpha-1}} - \frac{1}{(n-1)^{\alpha-1}} \right)$$

et, d'après les résultats relatifs aux séries télescopiques, on a

$$\begin{aligned} \frac{1}{1-\alpha} \sum_{k=2}^n \frac{1}{(k+1)^{\alpha-1}} - \frac{1}{k^{\alpha-1}} &\leq \sum_{k=2}^n \frac{1}{k^\alpha} \leq \frac{1}{1-\alpha} \sum_{k=2}^n \frac{1}{k^{\alpha-1}} - \frac{1}{(k-1)^{\alpha-1}} \\ \Leftrightarrow \frac{1}{1-\alpha} \left(\frac{1}{(n+1)^{\alpha-1}} - \frac{1}{2^{\alpha-1}} \right) + 1 &\leq \sum_{k=1}^n \frac{1}{k^\alpha} \leq \frac{1}{1-\alpha} \left(\frac{1}{n^{\alpha-1}} - 1 \right) + 1 \end{aligned}$$

Par encadrement, la suite des sommes partielles de la série $\sum \frac{1}{n^\alpha}$ converge si et seulement si

$$\alpha - 1 > 0 \iff \alpha > 1$$

Notes :

— Dans le cas $\alpha = 1$, une primitive de f_1 étant

$$F_1 : t \mapsto \ln(t) \xrightarrow[t \rightarrow +\infty]{} +\infty$$

on montre que la série $\sum \frac{1}{n}$ diverge.

— Pour $\alpha \in]0, 1]$, les séries de Riemann fournissent des exemples de séries divergentes dont le terme général tend vers 0.

— À l'aide du critère d'équivalence, l'étude des séries de Riemann permet de déterminer la nature de toute série numérique dont le terme général est une fraction rationnelle : si $u_n = \frac{P(n)}{Q(n)}$ alors

$$\text{la série } \sum u_n \text{ converge } \iff \deg(Q) - \deg(P) \geq 2$$

— La méthode utilisée pour étudier la convergence des séries de Riemann peut être appliquée pour toute série dont le terme général est de la forme $u_n = f(n)$, f étant une fonction positive, décroissante et qui tend vers 0 en $+\infty$. On a ainsi un nouveau critère d'étude : le critère de comparaison à une intégrale.

— La démonstration ci-dessus ne permet pas de calculer la valeur de la somme

$$\zeta(\alpha) = \sum_{n=1}^{+\infty} \frac{1}{n^\alpha}$$

pour $\alpha > 1$. À l'heure actuelle, la valeur de cette somme reste inconnue pour la plupart des valeurs de $\alpha > 1$. Précisément, on peut montrer que

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \qquad \sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$$

et que, de façon générale, on ne connaît actuellement la valeur de $\zeta(\alpha)$ que pour $\alpha \in 2\mathbb{N}^*$.

4.5.4 Critères de D'Alembert et de Cauchy

Pour finir, notons qu'il existe, pour les séries à termes positifs, des critères de convergence intrinsèques, basés uniquement sur l'étude du terme général u_n . Ainsi :

Proposition (critère de d'Alembert)

Soit (u_n) une suite à termes strictement positifs telle que le quotient

$$q_n = \frac{u_{n+1}}{u_n}$$

admette une limite ℓ lorsque n tend vers $+\infty$.

- Si $\ell < 1$, alors la série $\sum u_n$ converge.
- Si $\ell > 1$, alors la série $\sum u_n$ diverge.

Exemples :

- Pour tout $\alpha \in \mathbb{R}$, la série $\sum \frac{\alpha^n}{n!}$ converge.
- La série $\sum \binom{n}{3} \alpha^n$ converge si et seulement si $|\alpha| < 1$.

Proposition (critère de Cauchy)

Soit (u_n) une suite à termes strictement positifs telle que la racine

$$r_n = \sqrt[n]{u_n}$$

admette une limite ℓ lorsque n tend vers $+\infty$.

- Si $\ell < 1$, alors la série $\sum u_n$ converge.
- Si $\ell > 1$, alors la série $\sum u_n$ diverge.

Exemples :

- La série $\sum \left(\frac{2n+1}{3n+1}\right)^n$ converge.
- La série $\sum \frac{2^n}{n^\alpha}$ diverge quelque soit $\alpha \in \mathbb{R}$.

Note : aucun de ces critères ne permet de déterminer la nature des séries de Riemann. Pourquoi ?

4.6 Séries entières

La notion de somme infinie évoquée ci-dessus trouve de nombreuses applications dans différentes branches des mathématiques. L'une des plus utilisées est la notion de *série entière*, qui généralise la notion de polynôme en permettant d'envisager les polynômes de degré *infini*.

4.6.1 Définitions

Définition (série entière)

On appelle *série entière* toute fonction de la forme

$$S : x \mapsto \sum_{k=0}^{+\infty} a_k x^k$$

où (a_n) est une suite numérique donnée.

Exemples :

- si $a_n = 1$ pour tout $n \in \mathbb{N}$, on a $S(x) = \sum_{k=0}^{+\infty} x^k$.
- si $a_n = \frac{1}{n}$ pour tout $n \in \mathbb{N}^*$, on a $S(x) = \sum_{k=1}^{+\infty} \frac{x^k}{k}$.
- si $a_n = \frac{1}{n!}$, on a $S(x) = \sum_{k=0}^{+\infty} \frac{x^k}{k!}$.

4.6.2 Rayon de convergence

Une fois fixée une suite (a_n) et la série entière $S : x \mapsto \sum a_n x^n$ associée, il faut, dans un premier temps déterminer le domaine de définition de S .

En terme de séries, la question à laquelle il faut répondre est :

Pour quelles valeurs de $x \in \mathbb{R}$ la série $\sum a_n x^n$ est-elle convergente ?

En guise de réponse générale à cette question, on peut énoncer le théorème suivant :

Théorème/Définition (rayon de convergence)

Soit $S : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ une série entière. Il existe un nombre réel $R \in [0, +\infty]$ tel que le domaine de définition D_S de S est de la forme

$$D_S = (-R, R)$$

R est alors appelé **le rayon de convergence de S** .

Exemples :

— Pour $a_n = 1$ pour tout $n \in \mathbb{N}$, on a

$$S : x \mapsto \sum_{n=0}^{+\infty} x^n$$

D'après les propriétés de la série géométrique, on sait que

$$\text{La série } \sum x^n \text{ converge } \iff |x| < 1$$

D'où

$$R = 1 \quad \text{et} \quad D_S =]-1, 1[$$

Note : à partir de ce premier résultat, on peut déterminer le rayon de convergence et le domaine de définition de toute série entière de la forme

$$S : x \mapsto \sum_{n=0}^{+\infty} \alpha^n \cdot x^n, \quad \alpha \in \mathbb{R} \text{ fixé}$$

Exercice : à faire.

— Dans certains cas, il est également possible de s'appuyer sur le critère de d'Alembert. Ainsi, Soit S la série entière associée à la suite $(a_n) : a_n = \frac{1}{n}$:

$$S(x) = \sum_{n=1}^{\infty} \frac{x^n}{n}$$

Le domaine de définition de S est l'ensemble des x pour lesquels la série $\sum \frac{x^n}{n}$ converge.

En posant

$$u_n = \frac{|x|^n}{n}$$

le critère de d'Alembert s'appuie sur le quotient $\frac{u_{n+1}}{u_n}$:

$$\begin{aligned} \frac{u_{n+1}}{u_n} &= \frac{|x|^{n+1}}{n+1} \cdot \frac{n}{|x|^n} \\ &= \frac{n}{n+1} |x| \\ \frac{u_{n+1}}{u_n} &\xrightarrow{n \rightarrow +\infty} |x| \end{aligned}$$

Ainsi,

— Si $|x| < 1$, la série $\sum \frac{x^n}{n}$ converge. Donc

$$R \geq 1$$

— Si $|x| > 1$, la série $\sum \frac{x^n}{n}$ diverge. Donc

$$R \leq 1$$

Ainsi,

$$R = 1 \quad \text{et} \quad D_S = (-1, 1)$$

Notons que si $|x| = 1$, le critère de d'Alembert ne permet pas de conclure. Il faut donc étudier séparément les cas $x = 1$ et $x = -1$. Or d'après les différents exemples vus en première partie, on sait que

- $x = 1$: la série $\sum \frac{1}{n}$ diverge, donc $1 \notin D_S$.
- $x = -1$: la série $\sum \frac{(-1)^n}{n}$ converge, donc $-1 \in D_S$.

Conclusion : le domaine de définition de la fonction $S : x \mapsto \sum_{n=1}^{+\infty} \frac{x^n}{n}$ est

$$D_S = [-1, 1[$$

- En appliquant la même méthode à la série entière associée à la suite $(a_n) : a_n = \frac{1}{n!}$, on peut montrer que la série $E : x \mapsto \sum \frac{x^n}{n!}$ est définie sur \mathbb{R} tout entier (à vérifier). On a alors

$$R = +\infty$$

Note : formellement, on peut définir une série entière pour n'importe quelle suite numérique (a_n) . Cependant, l'étude du domaine de définition peut alors montrer que le rayon de convergence est nul. Dans ce cas, la fonction $x \mapsto \sum a_n x^n$ n'est définie qu'en $x = 0$.

Exemple : déterminer le rayon de convergence de la série associée à la suite

$$(a_n) : a_n = n!$$

4.6.3 Opérations sur les séries entières

D'un point de vue algébrique, les séries entières (au même titre que les polynômes) sont des objets que l'on peut additionner ou multiplier entre eux. Autrement dit, la somme ou le produit de deux séries entières est encore une série entière. Il existe alors des formules permettant de déterminer les coefficients de cette nouvelle série entière, ainsi que son rayon de convergence.

Proposition

Soient S et T deux séries entières associées respectivement aux suites (a_n) et (b_n) :

$$S(x) = \sum_{n=0}^{\infty} a_n x^n, \quad T(x) = \sum_{n=0}^{\infty} b_n x^n$$

On note respectivement R_a et R_b les rayons de convergence de S et T . Alors

— La somme $S + T$ est une série entière dont les coefficients (s_n) sont donnés par

$$\forall n \in \mathbb{N}, \quad s_n = a_n + b_n$$

D'autre part le rayon de convergence R_s de la somme $S + T$ vérifie

$$R_s \geq \min\{R_a, R_b\}$$

— Le produit $S \times T$ est une série entière dont les coefficients (p_n) sont donnés par

$$\forall n \in \mathbb{N}, \quad p_n = \sum_{k=0}^n a_k b_{n-k}.$$

D'autre part, le rayon de convergence R_p du produit $S \times T$ vérifie

$$R_p \geq \min\{R_a, R_b\}$$

4.6.4 Régularité d'une série entière

Une fois que l'on a déterminé le rayon de convergence d'une série entière, on souhaite savoir pour quelles valeurs de $x \in (-R, R)$ elle est continue, dérivable. Dans le cas des sommes *finies*, on sait que si chaque terme d'une somme est dérivable, alors la somme l'est également et la dérivée de la somme est la somme des dérivées de chacun des termes.

Dans le cas des sommes infinies, le problème est beaucoup plus complexe. Il relève des problèmes dits de passage à la limite sous le signe somme. Précisément, dans le cas des sommes infinies, il est possible de construire des fonctions $(f_n)_{n \in \mathbb{N}}$, toutes définies au voisinage d'un point $a \in \mathbb{R}$ et telles que

$$\lim_{x \rightarrow a} \sum_{n=0}^{+\infty} f_n(x) \neq \sum_{n=0}^{+\infty} \lim_{x \rightarrow a} f_n(x)$$

Les notions de continuité et de dérivabilité étant basées sur la notion de limite, il devient alors légitime d'étudier la régularité des séries entières. Cependant, on peut montrer que toute série entière de rayon de convergence R est continue sur l'ensemble $(-R, R)$ de son domaine de définition et dérivable sur $] -R, R[$. De plus, la dérivée s'obtient en dérivant terme à terme :

$$S(x) = \sum_{k=0}^{+\infty} a_k x^k \implies S'(x) = \sum_{k=0}^{+\infty} k a_k x^{k-1}$$

D'autre part, on peut montrer que la série entière $S'(x)$ a alors le même rayon de convergence que S .

De même, on peut montrer que pour une série entière $S : x \mapsto \sum a_n x^n$, on peut sans restriction intégrer terme à terme sur l'intervalle de convergence. Les primitives de S sont donc les fonctions de la forme

$$\Phi : x \mapsto \sum_{k=0}^{+\infty} \frac{a_k}{k+1} x^{k+1} + \lambda, \quad \lambda \in \mathbb{R}$$

De plus, on peut là encore montrer que la série entière ainsi obtenue a le même rayon de convergence que la série f de départ.

On pourrait alors pousser plus avant l'étude théorique des séries entières (dérivabilités successives, limites aux bord de l'intervalle de convergence,...). Cependant, on a ici les principales notions permettant de comprendre et utiliser les (nombreuses) applications des séries entières.

4.6.5 Développement en séries entières

La première application des séries entières porte sur les fonctions usuelles. Précisément, en généralisant les idées présentes dans la notion de développement limité, on peut écrire sous forme de séries entières toutes les fonctions réelles usuelles. On obtient alors ces fonctions sous la forme d'une somme infinie de monômes de la forme $a_n x^n$. Les D.L. ne sont alors qu'une troncature de cette somme infinie.

Du point de vue des séries, il s'agit de déterminer une série donnée sous forme explicite.

Exemple : pour $a_n = 1$ pour tout $n \in \mathbb{N}$, on a vu plus haut que la série entière

$$S : x \mapsto \sum_{n=0}^{+\infty} x^n$$

est convergente pour tout $x \in]-1, 1[$. Dans ce cas particulier, on sait exprimer la somme $S(x)$ sous forme explicite :

$$\forall x \in]-1, 1[, \quad \sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$$

On dit alors que la somme $S(x)$ ci-dessus est *le développement en série entière* (DSE) de la fonction $[x \mapsto \frac{1}{1-x}]$, valable sur $] - 1, 1[$.

De façon générale, on sait ainsi déterminer le développement en série entière des principales fonctions usuelles (voir table 4.1).

Note :

- Tous les développements en séries entières donnés dans la table 4.1 sont obtenus à partir des formules

$$\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n \quad \text{et} \quad e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$$

- À partir de ces DSE usuels, on peut développer de nombreuses fonctions en s'appuyant sur les opérations (+ et \times) vues plus haut.

Enfin, notons qu'il existe en général des conditions *d'existence* pour le développement en série entière d'une fonction donnée. Nous admettrons dans les exercices que ces conditions sont vérifiées. Notons également la propriété ci-dessous, établissant *l'unicité d'un développement en série entière* pour une fonction donnée :

TABLE 4.1 – Développements en séries entières

Fonction	DSE	Domaine
$x \mapsto \frac{1}{1-x}$	$\sum_{n=0}^{+\infty} x^n = 1 + x + x^2 + x^3 + \dots$	
$x \mapsto \frac{1}{1-x^2}$		
$x \mapsto \frac{1}{1+x}$		
$x \mapsto \frac{1}{1+x^2}$		
$x \mapsto (1+x)^\alpha, \quad \alpha \in \mathbb{R}$	$1 + \sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1)(\alpha-2)\dots(\alpha-(n-1))}{n!} x^n$	
$x \mapsto \ln(1+x)$		
$x \mapsto \arctan(x)$		
$x \mapsto e^x$	$\sum_{n=0}^{+\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \dots$	
$x \mapsto \sin(x)$		
$x \mapsto \cos(x)$		

Proposition (unicité du DSE)

Soient (a_n) et (b_n) deux suites numériques telles que les séries entières

$$S_a(x) = \sum_{n=0}^{+\infty} a_n x^n \quad \text{et} \quad S_b(x) = \sum_{n=0}^{+\infty} b_n x^n$$

convergent sur un intervalle $(-R, R)$ non trivial. Alors

$$(\forall x \in]-R, R[, S_a(x) = S_b(x)) \iff (\forall n \in \mathbb{N}, a_n = b_n)$$

Nous verrons dans la suite comment cette proposition permet de faire appelle au principe *d'identification* bien connu pour les polynômes.

4.6.6 Application aux équations différentielles

Outre la nouvelle représentation qu'offrent les séries entières pour les fonctions usuelles, elles permettent également d'étendre l'ensemble des fonctions à notre disposition pour répondre à différents types de problèmes.

C'est en particulier pour les équations différentielles linéaires.

Le principe consiste à chercher les solutions d'une équation différentielle linéaire donnée sous la forme d'une série entière. La manipulation des différentes sommes permet alors d'extraire de l'équation des conditions sur les coefficients de toute série solution.

Exemple : soit

$$(E) : xy' + y = e^x$$

On suppose qu'il existe une solution $S(x)$ de (E) qui s'écrit sous la forme

$$S(x) = \sum_{n=0}^{\infty} a_n x^n$$

et on note R son rayon de convergence.

D'après ce que l'on a vu plus haut, S est dérivable sur $] - R, R[$ et

$$\forall x \in]-R, R[, S'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

Par ailleurs,

$$\forall x \in]-R, R[, e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$$

La fonction S vérifie alors

$$\begin{aligned} \forall x \in]-R, R[, \quad xS'(x) + S(x) &= e^x \\ \Leftrightarrow x \sum_{n=1}^{\infty} n a_n x^{n-1} + \sum_{n=0}^{\infty} a_n x^n &= \sum_{n=0}^{+\infty} \frac{x^n}{n!} \\ \Leftrightarrow \sum_{n=1}^{\infty} n \cdot a_n x^n + \sum_{n=0}^{\infty} a_n x^n &= \sum_{n=0}^{+\infty} \frac{x^n}{n!} \\ \Leftrightarrow \sum_{n=0}^{\infty} (n+1) a_n x^n &= \sum_{n=0}^{+\infty} \frac{x^n}{n!} \end{aligned}$$

Par identification, on en déduit que

$$\forall n \in \mathbb{N}, \quad (n+1)a_n = \frac{1}{n!} \Leftrightarrow a_n = \frac{1}{(n+1)!}$$

La solution cherchée est alors

$$S : x \mapsto \sum_{n=0}^{+\infty} \frac{x^n}{(n+1)!}$$

Pour terminer l'étude, il faut déterminer le rayon de convergence de la série obtenue. On utilise pour cela le critère de d'Alembert sur le terme général

$$u_n = |a_n \cdot x^n| = \frac{|x|^n}{(n+1)!}$$

$$\begin{aligned} \frac{u_{n+1}}{u_n} &= \frac{|x|^{n+1}}{(n+2)!} \cdot \frac{(n+1)!}{|x|^n} \\ &= \frac{|x|}{n+2} \xrightarrow{n \rightarrow +\infty} 0 \end{aligned}$$

On en déduit

$$R = 0$$

et la série entière S est définie sur \mathbb{R} .

Note : il est possible ici, pour finir, d'obtenir une forme explicite pour la solution S cherchée. En général, il n'est pas toujours possible d'obtenir une forme explicite pour la ou les solution(s) cherchée(s). Cependant, il est alors possible d'obtenir une approximation des solutions cherchées à l'aide des sommes partielles

$$S_n : x \mapsto \sum_{k=0}^n a_k x^k$$

L'approximation est d'autant plus précise que le nombre n de termes considérés est grand (et on peut être aussi précis que l'on veut, quitte à considérer suffisamment de termes).

Exercice : montrer que la série entière

$$S : x \mapsto \sum_{n=0}^{+\infty} \frac{x^n}{(n+1)!}$$

est le Développement en Série Entière de la fonction

$$x \mapsto \frac{e^x - 1}{x}$$

Note : on pourra vérifier que cette fonction admet un prolongement continu sur \mathbb{R} .

Chapitre 5

Fonctions de plusieurs variables

Introduction

Une fonction de plusieurs variables est une fonction f liant une quantité $f(x_1, \dots, x_n)$ à un ensemble de variables x_1, \dots, x_n .

Comme les fonctions d'une variable, elles permettent de représenter des quantités, des positions, etc. Elles permettent donc de modéliser des situations complexes dépendant de plusieurs paramètres. Ainsi,

- la densité (ou la résistance ponctuelle, ou la température) d'une poutre carrée peut, en première approximation, être représentée par une fonction d'une variable x ,
- les caractéristiques d'un plancher peuvent être représentées par des fonctions de deux variables ou trois variables selon la précision souhaitée,
- si l'on étudie des phénomènes qui évoluent au cours du temps, il faut ajouter le temps.

L'étude de ces fonctions permet donc de modéliser et d'étudier en détails l'évolution, la géométrie, la composition de nombreux problèmes physiques.

Les outils développés pour l'analyse de ces fonctions généralisent ceux que l'on connaît en dimension 1 : limites, dérivées, D.L., tangentes, ...

Comme en dimension 1, ces nouveaux outils associés aux lois de la physique appliquées à des phénomènes à plusieurs dimensions produisent des équations liant les fonctions associées à leurs dérivées.

Les équations ainsi obtenues sont dites *Équations aux Dérivées Partielles* (EDP). L'analyse en plusieurs dimensions permet de développer des méthodes de résolution pour certaines de ces équations, notamment celles issues des modèles les plus simples. On en verra quelques exemples, mais au même titre que pour les Équations Différentielles Ordinaires (EDO), pour les modèles plus complexes, les scientifiques sont souvent amenés à se tourner vers des méthodes de résolution approchée, basées sur le calcul numérique. Une partie du travail des chercheurs en sciences "dure" actuellement consiste à développer des outils informatiques (des programmes) donnant une solution approchée fiable à ce type d'EDP.

Nous verrons dans un premier temps comment, en dimension 2, on peut généraliser la notion de courbe connue pour les fonctions d'une variable.

On aura alors un outil géométriques pour comprendre la généralisation des outils plus analytiques aux dimensions supérieures que nous verrons ensuite.

5.1 Définitions

Formellement, une fonction f dépendent de n variables une application qui à un jeu de n coordonnées (x_1, x_2, \dots, x_n) associe une quantité $f(x_1, \dots, x_n)$. L'ensemble des n -uplets pour lesquelles cette quantité existe est le *domaine de définition* de f . C'est une partie D de l'ensemble \mathbb{R}^n . On note

$$f : \begin{array}{l} D \\ X = (x_1, x_2, \dots, x_n) \end{array} \begin{array}{l} \longrightarrow \\ \longmapsto \end{array} \begin{array}{l} \mathbb{R} \\ f(X) = f(x_1, \dots, x_n) \end{array}$$

Exemples :

- Dans le cas $n = 2$, ce domaine de définition correspond à une partie du plan (muni d'un repère orthonormé) :

$f(x, y) = \sin(x + y),$	$D = \mathbb{R}^2$	
$f(x, y) = \frac{1}{x^2 + y^2},$	$D = \mathbb{R}^2 \setminus \{(0, 0)\}$	le plan privé du centre
$f(x, y) = \frac{1}{x - y},$	$D = \mathbb{R}^2 \setminus \{x = y\}$	le plan privé de la première bissectrice
$f(x, y) = \sqrt{1 - x^2 + y^2},$	$D = \{x^2 + y^2 \leq 1\}$	le disque unité

- Dans le cas $n = 3$, ce domaine de définition correspond à une partie de l'espace :

$f(x, y, z) = 2xy + z^2$	$D = \mathbb{R}^3$
$f(x, y, z) = \frac{1}{x - y - z}$	$D = \mathbb{R}^3 \setminus \{x - y - z = 0\}$
$f(x, y, z) = \frac{z}{x^2 + y^2}$	$D = \mathbb{R}^3 \setminus \{x = y = 0\}$
$f(x, y, z) = \ln\left(x^2 + y^2 + \frac{z^2}{4} - 1\right)$	$D = \left\{x^2 + y^2 + \frac{z^2}{4} > 1\right\}$

5.1.1 Graphe d'une fonction

À l'image de ce que l'on connaît pour les fonctions d'une variable, on peut également définir *le graphe* d'une fonction de plusieurs variables.

Définition (graphe d'une fonction)

Soit f une fonction de n variables, définie sur une partie $D \subset \mathbb{R}^n$. On appelle *graphe de f* l'ensemble

$$\begin{aligned} \mathcal{G}_f &= \{(x_1, \dots, x_n, f(x_1, \dots, x_n)), (x_1, \dots, x_n) \in D\} \\ &= \{(X, f(X)), X \in D\} \subset \mathbb{R}^{n+1} \end{aligned}$$

ATTENTION : le graphe de f est un sous ensemble de \mathbb{R}^{n+1} .

Dans le cas $n = 1$, on retrouve la courbe de f :

$$\mathcal{C}_f = \{(x, f(x)), x \in D\} = \{(x, y) \in \mathbb{R}^2 / y = f(x)\}$$

Il s'agit de la courbe d'équation $y = f(x)$, inscrite dans le plan muni d'un repère.

Dans le cas $n = 2$, le graphe de f est ainsi le sous ensemble de \mathbb{R}^3 défini par

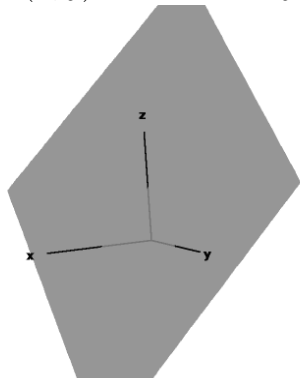
$$\mathcal{S}_f = \{(x, y, f(x, y)), (x, y) \in D\} = \{(x, y, z) \in \mathbb{R}^3 / z = f(x, y)\}$$

Ce graphe peut ici être représenté par *une surface* de l'espace muni d'un repère : la surface d'équation $z = f(x, y)$.

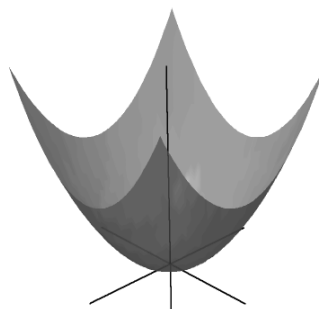
De façon générale, le graphe d'une fonction de n variables est une *l'hypersurface* de \mathbb{R}^{n+1} .

Exemples :

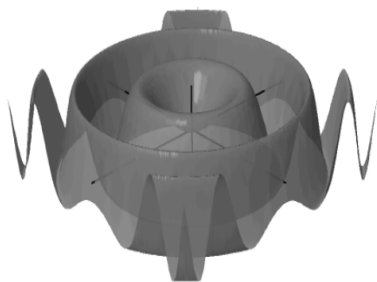
$$\begin{aligned} \ell : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto -x - y + 1 \end{aligned}$$



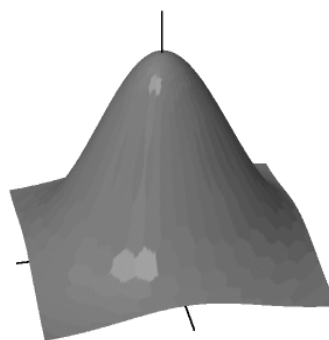
$$\begin{aligned} p : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^2 + y^2 \end{aligned}$$



$$\begin{aligned} s : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \sin(x^2 + y^2) \end{aligned}$$



$$\begin{aligned} g : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto e^{-x^2 - y^2} \end{aligned}$$



5.1.2 Isoclines

Définition (Isoclines)

Soit f une fonction de n variables, définie sur un domaine $D \subset \mathbb{R}^n$. Pour tout $h \in \mathbb{R}$, on appelle *isocline de niveau h* l'ensemble

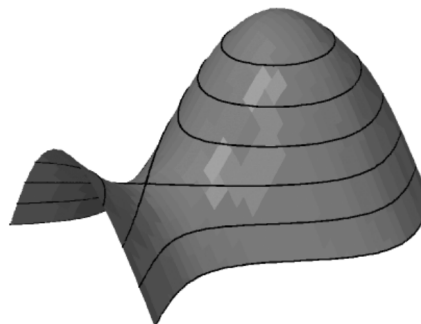
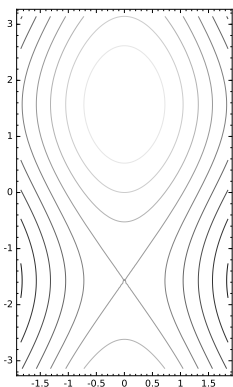
$$I_h = \{(x_1, \dots, x_n) \in D \mid f(x_1, \dots, x_n) = h\}$$

Toute isocline est ainsi un sous ensemble de \mathbb{R}^n .

Dans le cas $n = 2$, les isoclines d'une fonction donnée sont (en général) *des courbes* du plan, appelées aussi *courbes de niveaux*, définies par l'équation *implicite*

$$f(x, y) = h$$

On retrouve également cette isocline sur la surface \mathcal{S}_f en reliant entre eux tous les points d'altitude h , i.e. en coupant \mathcal{S}_f par le plan horizontal d'équation $z = h$.



Dans le cas $n = 3$, les isoclines d'une fonction donnée sont (en général) *des surfaces* de l'espace, définies par l'équation *implicite*

$$f(x, y, z) = h$$

5.2 Fonctions et dérivées partielles

Soit f une fonction de n variables (x_1, \dots, x_n) définie sur un domaine $D \subset \mathbb{R}^n$. Comme pour les fonctions d'une variable, l'étude de f a pour but d'étudier l'évolution de f en fonction des variations des n variables. Cependant, dans le cas général, l'étude de ces variations est rendu plus complexe que dans le cas $n = 1$ du fait de l'indépendance des variables.

L'essentiels du travail d'analyse consiste alors à déterminer dans quelle mesure il est possible d'étudier les variables séparément.

5.2.1 Fonctions partielles

Définition (fonctions partielles)

Soit f une fonction de n variables, définie sur un domaine $D \subset \mathbb{R}^n$ et soit $P = (a_1, \dots, a_n)$ un point de D .

Pour tout $i \in \{1, \dots, n\}$, la i -ème fonction partielle de f est la fonction f_i définie par

$$f_i : x \mapsto f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_n)$$

définie pour tout $x \in \mathbb{R}$ tel que

$$(a_1, \dots, x, \dots, a_n) \in D$$

Il s'agit de fonctions d'**une** variable. Il est en particulier possible de représenter ces différentes fonctions partielles par des courbes, et de les étudier à l'aide des outils d'analyse en dimension 1. L'étude de ces courbes permet alors d'obtenir de l'information concernant la fonction f *au voisinage* du point $P \in D$.

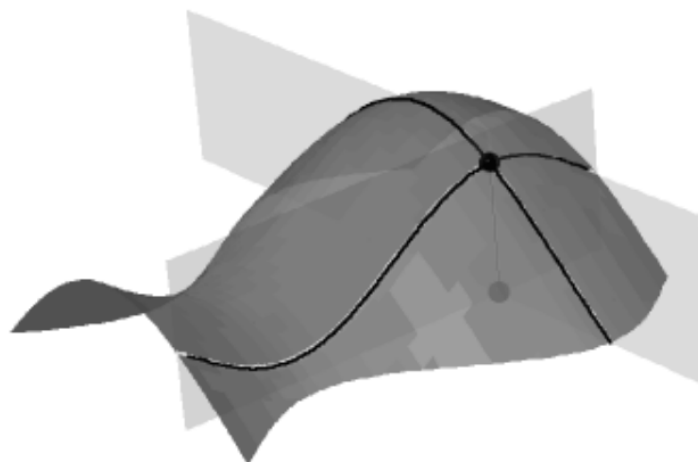
Dans le cas $n = 2$, il est également possible de retrouver les courbes de ces fonctions partielles sur la surface.

Précisément, soit f une fonction de deux variables (x, y) et $P = (a, b)$ un point de son domaine de définition. Les fonctions partielles de f en P sont

$$f_1 : x \mapsto f(x, b) \quad \text{et} \quad f_2 : y \mapsto f(a, y)$$

Les courbes des fonctions f_1 et f_2 sont alors les courbes de la surface \mathcal{S}_f obtenues la coupant par les plans d'équations

$$x = a \quad \text{et} \quad y = b$$



la courbe de la fonction partielle $f_1 : x \mapsto f(x, b)$ est exactement la courbe obtenue en coupant \mathcal{S}_f par le plan vertical d'équation $y = b$.

5.2.2 Dérivées partielles

Les fonctions partielles définies ci-dessus étant en particulier des fonctions d'une variable, il est possible de les étudier avec les outils de l'analyse. On peut en particulier envisager de les dériver. On définit ainsi les *dérivées partielles de f* . Précisément,

Définition (dérivées partielles)

Soit f une fonction de n variables, définie au voisinage d'un point $P = (a_1, \dots, a_n) \in \mathbb{R}^n$ et soit $i \in \{1, \dots, n\}$.

La fonction f est dite dérivable par rapport à sa i -ème variable en P si la limite

$$\lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_i + h, \dots, a_n) - f(a_1, \dots, a_n)}{h}$$

existe dans \mathbb{R} . On note alors

$$\frac{\partial f}{\partial x_i}(a_1, \dots, a_n) = \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_i + h, \dots, a_n) - f(a_1, \dots, a_n)}{h}$$

Note : en note $E_i = (0, \dots, 1, \dots, 0)$, on a également la notation

$$\frac{\partial f}{\partial x_i}(P) = \lim_{h \rightarrow 0} \frac{f(P + h.E_i) - f(P)}{h}$$

Ces dérivées partielles sont alors intimement liées aux variations de f autour de P . Précisément, la i -ème dérivée partielle en P mesure la variation de f en P dans la direction de l'axe (Ox_i) .

On peut alors montrer que, sous certaines conditions de régularité, ces variations “le long des axes du repère” permettent de connaître les variations de f autour de P dans n'importe quelle direction.

Enfin, si ces dérivées partielles existent en tout point de D , elles forment les *fonctions dérivées partielles*

$$\frac{\partial f}{\partial x_i} : (x_1, \dots, x_n) \longmapsto \frac{\partial f}{\partial x_i}(x_1, \dots, x_n)$$

Celles-ci sont à leur tour des fonctions de n variables, qui peuvent admettre des dérivées partielles par rapport à chacune de ces variables. On construit ainsi les dérivées partielles d'ordre 2 (et plus en itérant encore le processus).

Notes :

1. En pratique, le calcul d'une dérivée partielle $\frac{\partial f}{\partial x_i}$ s'effectue comme en dimension 1, les variables x_j , $j \neq i$ étant considérées constantes.
2. Une dérivée partielle d'ordre k s'obtient en dérivant k fois la fonction f de départ, chacune de ces dérivées pouvant se faire par rapport à l'une ou l'autre des variables.

Une fonction f de n variables admet donc

- n dérivées partielles d'ordre 1,
- n^2 dérivées partielles d'ordre 2
- n^k dérivées partielles d'ordre k .

Il est alors naturel de se demander si lors d'un calcul de dérivée partielle d'ordre au moins 2, l'ordre des dérivations importe ou non. Un résultat important de l'analyse multivariée, dû à Schwarz, assure que l'ordre de dérivation n'importe pas. Autrement dit, si les dérivées partielles secondes existent et sont continues, alors

$$\forall i, j \in \{1, \dots, n\}, \quad \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

Dans le cas $i = j$, on note $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_i^2}$.

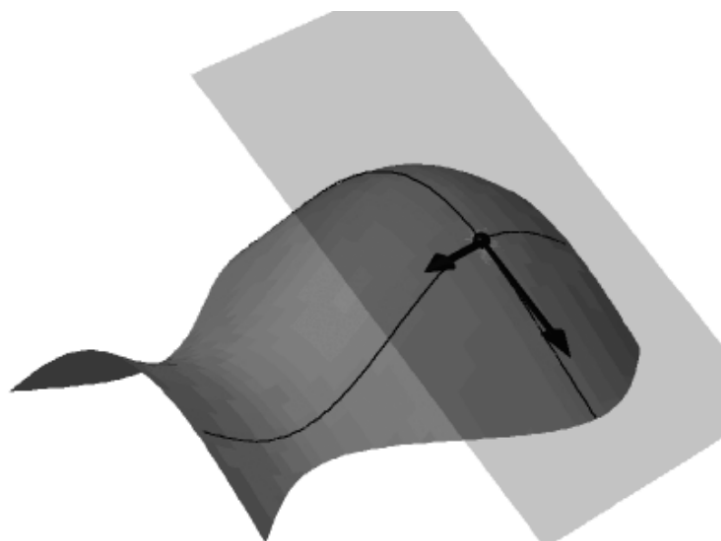
5.3 Plan tangent et différentielle

5.3.1 Définition

Comme en dimension 1, il est possible d'interpréter les dérivées partielles vue plus haut en terme de pente et de vecteur tangent. Ainsi, dans le cas $n = 2$, les fonctions partielles d'une fonction f en un point $P = (a, b)$ correspondant à des courbes tracées sur la surface \mathcal{S}_f , les dérivées partielles $\frac{\partial f}{\partial x}(a, b)$ et $\frac{\partial f}{\partial y}(a, b)$ représentent de coefficients directeurs des tangentes à ces courbes au dessus du point $P = (a, b)$. Elles donnent également les coordonnées de deux vecteurs tangents à la surface \mathcal{S}_f au dessus du point P :

$$\vec{v}_x(P) = \left(1, 0, \frac{\partial f}{\partial x}(a, b)\right), \quad \vec{v}_y(P) = \left(0, 1, \frac{\partial f}{\partial y}(a, b)\right)$$

Ces deux vecteurs étant non colinéaires, ils engendrent un plan de l'espace, appelé *plan tangent* à la surface \mathcal{S}_f en P .



Comme la tangente en dimension 1, ce plan tangent (et en particulier son inclinaison dans l'espace) permet une première description des variations de f autour du point P .

Note : en dimension n quelconque, en tout point $P \in \mathbb{R}^n$ la fonction f admet des dérivées partielles, il est encore possible d'associer à chacune de ces n dérivée partielle $\frac{\partial f}{\partial x_i}(P)$ un vecteur de \mathbb{R}^{n+1} . L'ensemble de ces vecteurs formant une famille libre de \mathbb{R}^{n+1} , elles engendrent un sous espace affine de \mathbb{R}^{n+1} de dimension n , appelé *hyperplan tangent* à l'hypersurface \mathcal{S}_f en P .

5.3.2 Équation du plan tangent

Comme en dimension 1, le lien qu'il existe entre les dérivées partielles d'une fonction f en un point P et son plan tangent se retrouvent dans l'équation de ce plan tangent. On

peut ainsi montrer qu'en dimension 2, l'équation du plan tangent à la surface \mathcal{S}_f d'une fonction f en un point $P = (a, b)$ est

$$T_P : z = \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b) + f(a, b)$$

Exemple : soit \mathcal{S}_f la surface associée à la fonction

$$f : (x, y) \mapsto x^2 + y^2 + 1.$$

1. Déterminer les équations des plans tangents à \mathcal{S}_f aux points d'abscisse $(1, 1)$, $(1, 2)$.
2. Déterminer le(s) point(s) de \mathcal{S}_f admettant un plan tangent horizontal.

Note : à partir de cette équation de plan, on peut exprimer un *vecteur normal* à ce plan tangent :

$$\vec{n}_P = \left(\frac{\partial f}{\partial x}(P), \frac{\partial f}{\partial y}(P), -1 \right)$$

En projetant ce vecteur sur le plan "de base" (xOy) , on obtient le vecteur

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

appelé *vecteur gradient* de f en P .

On peut alors montrer que, dans le plan (xOy) , ce vecteur indique "la ligne de plus grande pente" en P , i.e. la direction à prendre en quittant P pour faire augmenter le plus rapidement la valeur de f .

Par ailleurs, cette propriété reste valable en dimensions supérieures. Nous reviendrons plus en détails sur cette notion de vecteur gradient.

5.3.3 Différentielle

Définition (différentielle)

Soient une fonction f de 2 variables définie sur un domaine $D \subset \mathbb{R}^2$ et $P = (a, b)$ un point de D .

On appelle *différentielle de f en P* l'application df_P définie sur \mathbb{R}^2 par

$$df_P : (x, y) \mapsto \frac{\partial f}{\partial x}(a, b).x + \frac{\partial f}{\partial y}(a, b).y$$

Il s'agit d'une *application linéaire* en (x, y) , i.e. de la forme

$$(x, y) \longmapsto \alpha.x + \beta.y$$

Sa surface associée est donc un plan vectoriel. On reconnaît notamment dans l'expression de $df_P(x, y)$ la partie linéaire de l'équation du plan tangent T_P à \mathcal{S}_f en P .

Précisément, on a

$$T_P : z = df_P(x - a, x - b) + f(a, b)$$

On peut alors montrer que le plan associé à df_P est le plan de l'espace parallèle au plan tangent T_P , passant par le centre du repère. Ce plan contient en particulier toute l'information concernant l'évolution de la fonction f autour du point P et la différentielle df_P permet d'accéder à une estimation des variations de f autour de P .

Dans le cas général, la différentielle d'une fonction f de n variables en un point $P = (a_1, \dots, a_n)$ de son domaine est

$$df_P : X = (x_1, \dots, x_n) \longmapsto \sum_{k=1}^n \frac{\partial f}{\partial x_k}(P).x_k$$

et l'équation de l'hyperplan tangent à \mathcal{S}_f en P dans \mathbb{R}^{n+1} est

$$T_P : x_{n+1} = f(P) + df_P(X - P).$$

Là encore, la différentielle est une fonction linéaire, dont les variations en $O_{\mathbb{R}^n}$ permettent d'estimer les variations de f au voisinage de P .

Exercice : quelle est la matrice de cette application linéaire dans les bases canoniques de \mathbb{R}^n et \mathbb{R} ?

5.3.4 Développements limités

Rappels : si φ est une fonction d'une variable, deux fois dérivable en un point x_0 de son domaine, la formule de Taylor permet d'approcher, au voisinage de x_0 , la quantité $f(x)$ par un polynôme dont les coefficients dépendent des dérivées successives de f au point x_0 . Ainsi, aux ordres 1 et 2, on a

$$\text{Ordre 1 : } f(x) = f(x_0) + f'(x_0).(x - x_0) + o(x - x_0)$$

$$\iff_{x=x_0+h} f(x_0 + h) = f(x_0) + f'(x_0).h + o(h)$$

$$\text{Ordre 2 : } f(x_0 + h) = f(x_0) + f'(x_0).h + \frac{f''(x_0)}{2}.h^2 + o(h^2)$$

D'un point de vue géométrique, le développement limité (DL) à l'ordre 1 correspond à une approximation de la courbe de f par sa tangente au point d'abscisse x_0 . L'erreur commise dans ce cas est alors négligeable devant $h = x - x_0$ au voisinage de x_0 .

Le développement limité à l'ordre 2 correspond lui à une approximation de la courbe de f par une parabole tangente à la courbe de f au point d'abscisse x_0 . L'erreur commise dans ce cas est alors négligeable devant $h^2 = (x - x_0)^2$, toujours au voisinage de x_0 .

À l'aide de la notion de dérivée partielle, il est alors possible de généraliser la notion de développement limité aux dimensions supérieures. Bien qu'il soit formellement possible d'envisager, comme en une variable, un développement limité de tout ordre, on se contente en général des ordres 1 et 2. Ainsi,

Définition (développement limité en n variables)

Soit f une fonction de n variables définie sur un domaine $D \subset \mathbb{R}^n$ et $P = (a_1, \dots, a_n)$ un point de D . Si f admet des dérivées partielles d'ordre 1 et 2 en P , on a, pour tout $H = (h_1, \dots, h_n) \in \mathbb{R}^n$,

$$\text{Ordre 1 : } f(P + H) = f(P) + \overbrace{\sum_{i=1}^n \frac{\partial f}{\partial x_i}(P).h_i}^{=df_P(H)} + o(\|H\|)$$

$$\text{Ordre 2 : } f(P + H) = f(P) + df_P(H) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(P).h_i h_j + o(\|H\|^2)$$

Ainsi, en deux variables, le développement limité à l'ordre 1 en un point $P = (a, b)$ est

$$f(a + h, b + k) = f(a, b) + \frac{\partial f}{\partial x}(a, b).h + \frac{\partial f}{\partial y}(a, b).k + o(\sqrt{h^2 + k^2})$$

On reconnaît ici l'équation du plan tangent à la surface \mathcal{S}_f au point d'abscisse P et la formule assure alors que si l'on approche la surface \mathcal{S}_f par ce plan tangent pour calculer la quantité $f(X)$, l'erreur commise est alors négligeable devant la distance $\|X - P\|$.

De même, à l'ordre 2, on a

$$f(a+h, b+k) = f(a, b) + df_P(h, k) + \frac{1}{2} \left(\frac{\partial^2 f}{\partial x^2}(a, b) \cdot h^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(a, b) \cdot hk + \frac{\partial^2 f}{\partial y^2}(a, b) \cdot k^2 \right) + o(h^2 + k^2)$$

Ici, si l'on néglige le terme d'erreur, on retrouve une équation polynomiale de degré 2. D'un point de vue géométrique, l'approximation ainsi effectuée correspond à une approximation de la surface \mathcal{S}_f par une quadrique de l'espace.

Dans la formule ci-dessus, on voit en particulier apparaître la forme quadratique

$$q(h, k) = \frac{\partial^2 f}{\partial x^2}(a, b) \cdot h^2 + 2 \cdot \frac{\partial^2 f}{\partial x \partial y}(a, b) \cdot hk + \frac{\partial^2 f}{\partial y^2}(a, b) \cdot k^2$$

La matrice de cette forme quadratique est

$$\mathcal{H}_f(P) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(a, b) & \frac{\partial^2 f}{\partial x \partial y}(a, b) \\ \frac{\partial^2 f}{\partial x \partial y}(a, b) & \frac{\partial^2 f}{\partial y^2}(a, b) \end{pmatrix}$$

appelée *matrice hessienne de f en P*.

Cette matrice permet en particulier d'obtenir une écriture matricielle du DL à l'ordre 2 :

$$f(P+H) = f(P) + df_P(H) + \frac{1}{2} {}^t H \times \mathcal{H}_f(P) \times H + o(\|H\|^2)$$

On retrouve ici une notation proche de celle que l'on connaît en dimension 1, qui en outre reste valable en dimensions supérieures.

Note : en pratique, les fonctions étudiées étant construites à partir des fonctions usuelles, il est possible de calculer ces développements limités à l'aide des formules connues en dimension 1.

Exemple : soit $f(x, y) = \ln(1 + x + y)$.

1. Déterminer, à l'aide de la définition, le D.L. de f en $(0, 0)$ à l'ordre 2.
2. Retrouver ce D.L. à l'aide du D.L. de $\ln(1 + u)$ en 0 à l'ordre 2.
3. À l'aide de ce même D.L., donner le D.L. de f à l'ordre n .

5.3.5 Recherche d'extrema

Comme en dimension 1, la notion de dérivées partielles est en particulier un outil important dans la recherche et l'étude des valeurs extrêmes d'une fonction de plusieurs variables.

Points critiques

Définition (points critiques)

Soit f une fonction de n variables, admettant des dérivées partielles en tout point de son domaine $D \subset \mathbb{R}^n$. On appelle *point critique de f* tout point $X \in D$ tel que

$$\forall i \in \{1, \dots, n\}, \quad \frac{\partial f}{\partial x_i}(X) = 0$$

Ainsi, en dimension 1, les points critiques de f sont les solutions de l'équation

$$f'(x) = 0$$

Ils correspondent aux points auxquels la courbe \mathcal{C}_f admet une tangente horizontale, condition nécessaire à l'existence d'extremum (pour les points situés à l'intérieur du domaine de définition).

En dimension 2, les points critiques de f sont les solutions du *système d'équations*

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

D'un point de vue géométrique, il s'agit des points auxquels la surface \mathcal{S}_f admet *un plan tangent horizontal*, condition là encore nécessaire à l'existence d'un extremum local.

De façon général, les points critiques d'une fonction f de n variables correspondent aux points de \mathbb{R}^n auxquels la fonction f "marque un palier". C'est donc parmi ces points là qu'il faut chercher d'éventuels extrema locaux, situés à l'intérieur du domaine de définition de f .

ATTENTION : comme en dimension 1, la recherche des points critique ne constitue qu'un premier tri. Une fois ces points critiques déterminés, il faut les étudier un par un pour savoir s'il s'agit de maximum, de minimum ou ni l'un ni l'autre.

On verra que la suite de l'étude peut, toujours comme en dimension 1, s'appuyer sur les dérivées partielles secondes de la fonction f étudiée et plus exactement sur les matrices hessiennes de f en ces points critiques.

Il est cependant possible, dans certains cas, de déterminer la nature des différents points critiques obtenus en se basant uniquement sur le bon sens.

Exemple : soit $f : (x, y) \mapsto e^{-x^2-y^2}$.

1. Montrer que f admet un unique point critique à déterminer.
2. Montrer que cet unique point critique est un maximum.

Nature des points critiques

Le cas $n = 2$. Soit f , une fonction de deux variables $X = (x, y)$, définie sur une partie $D \subset \mathbb{R}^2$ et admettant un point $P = (a, b)$ comme point critique. Les dérivées partielles de f en P étant nulles, le D.L. de f à l'ordre 2 en P devient :

$$f(x, y) = f(a, b) + \frac{1}{2}(r(x-a)^2 + 2s(x-a)(y-b) + (y-b)^2) + o(\|X - P\|^2)$$

où

$$r = \frac{\partial^2 f}{\partial x^2}(a, b), \quad s = \frac{\partial^2 f}{\partial x \partial y}(a, b), \quad t = \frac{\partial^2 f}{\partial y^2}(a, b)$$

Autrement dit, en notant $q(x, y) = r.x^2 + 2s.xy + t.y^2$, l'allure de la surface \mathcal{S}_f au voisinage de P coïncide avec celle de la quadrique d'équation

$$z = q(x, y)$$

Or la forme de cette quadrique, et donc la nature du point critique $P = (0, 0)$, dépendent alors uniquement du signe des valeurs propres de la matrice de q , i.e. la matrice hessienne

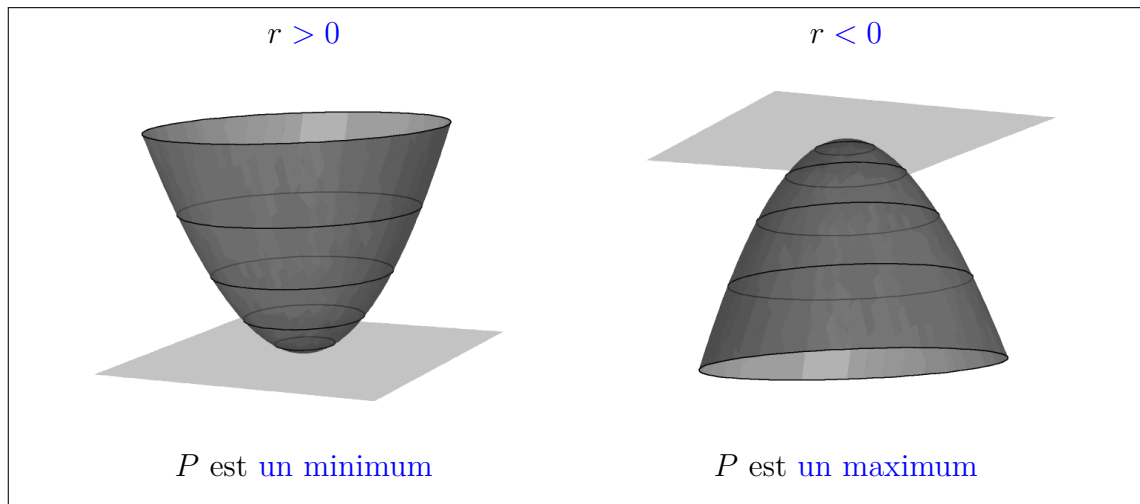
$$\mathcal{H}_f(P) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

En deux variables, ces informations sont également accessibles via *le déterminant*

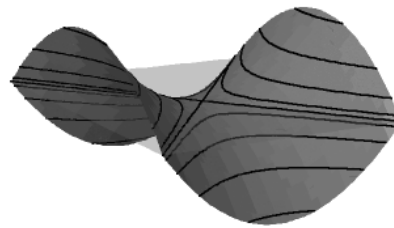
$$\det(\mathcal{H}_f(P)) = rt - s^2$$

On peut alors établir la classification suivante :

- Si $\det(\mathcal{H}_f(P)) > 0$, les deux valeurs propres de $\mathcal{H}_f(P)$ sont de même signe et P est un extremum. Le signe du coefficient r de x^2 donne alors l'orientation de la quadrique associée à q (et donc de \mathcal{S}_f) :



- si $\det(\mathcal{H}_f(P)) < 0$, les deux valeurs propres de $\mathcal{H}_f(P)$ sont de signes contraires et P est alors *un point selle* :



Dans ce cas, tout voisinage de P contient des points X tels que $f(X) > f(P)$ et d'autres tels que $f(X) < f(P)$. Le point P n'est alors pas un extremum local de f .

Note : si $\det(\mathcal{H}_f(P)) = 0$, c'est que l'une des deux valeurs propres de $\mathcal{H}_f(P)$ est nulle. Si l'autre est non nulle, le point P est situé sur une droite de points critiques. L'ensemble de ces points sont alors des extrema et leur nature (maximum ou minimum) dépend là encore du signe de r (ou de t si $r = 0$).

Le cas général

Pour une fonction f de n variables ($n \geq 3$), on n'a plus de représentation géométrique de la situation. Cependant, si P est un point critique de f , le développement limité de f en P à l'ordre 2 s'écrit

$$f(P + H) = f(P) + \frac{1}{2} \cdot {}^t H \times \mathcal{H}_f(P) \times H + o(H^2)$$

où

$$\mathcal{H}_f(P) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(P) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(P) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(P) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(P) \end{pmatrix}$$

et au voisinage de P (i.e. lorsque H est “petit”), le signe de la différence $f(P+H) - f(P)$ dépend alors uniquement du signe de la forme quadratique ${}^t H \times \mathcal{H}_f(P) \times H$.

Or la théorie de réduction des formes quadratiques nous assure qu’il existe un repère de \mathbb{R}^n dans lequel cette forme quadratique s’écrit

$$\tilde{q}(\tilde{H}) = \lambda_1 \tilde{h}_1^2 + \lambda_2 \tilde{h}_2^2 + \dots + \lambda_n \tilde{h}_n^2$$

où les λ_i sont les valeurs propres de \mathcal{H}_f .

Puisqu’il s’agit ici d’une somme (ou différence) de carrés, le signe de cette expression dépend essentiellement du signes des valeurs propres λ_i . Ainsi,

- Si toutes ces valeurs propres sont **positives** on a, pour H petit

$$f(P+H) - f(P) \geq 0 \iff f(P+H) \geq f(P)$$

et P est un minimum.

- Si toutes ces valeurs propres sont **négatives** on a, pour H petit

$$f(P+H) - f(P) \leq 0 \iff f(P+H) \leq f(P)$$

et P est un maximum.

- Si ces valeurs propres ne sont **pas toutes de même signe**, il existe, à partir du point P , des directions dans lesquelles la quantité $f(P+H)$ augmente et d’autre dans lesquelles elle diminue, ces différentes directions étant données par les vecteurs propres de la matrice $\mathcal{H}_f(P)$.

Dans ce cas, le point P n’est pas un extremum. C’est un point selle.

5.4 Champs de vecteurs

5.4.1 Définitions

À partir de la notion de fonction de plusieurs variables définie ci-dessus, il est possible de généraliser encore la notion de fonction via la notion de *fonction vectorielle*.

Définition (Fonction vectorielle et champ de vecteurs)

Soient m et n deux entiers naturels. On appelle *fonction vectorielle* de \mathbb{R}^m dans \mathbb{R}^n toute fonction de la forme

$$F : \begin{array}{ccc} \mathbb{R}^m & \longrightarrow & \mathbb{R}^n \\ X = (x_1, \dots, x_m) & \longmapsto & f(X) = (f_1(X), \dots, f_n(X)) \end{array}$$

On dira en outre qu'une telle fonction F est un *champ de vecteurs* si $m = n$.

Par ailleurs, les fonctions $f_i : \mathbb{R}^m \rightarrow \mathbb{R}$ définies ci-dessus sont appelées *les fonctions coordonnées de F* .

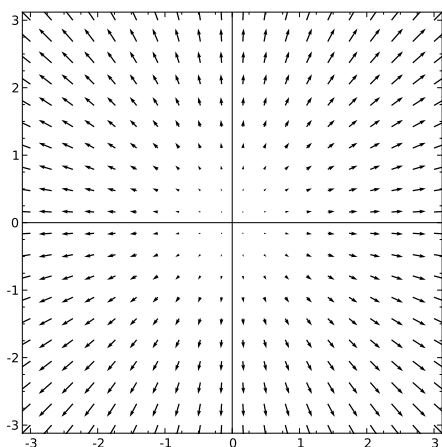
La terminologie *champ de vecteur* définie ci-dessus s'explique par le fait que si $m = n$, la fonction F permet d'associer à chaque point $X = (x_1, \dots, x_n)$ de son domaine de définition un vecteur

$$F(X) = (f_1(X), \dots, f_n(X)) \in \mathbb{R}^n$$

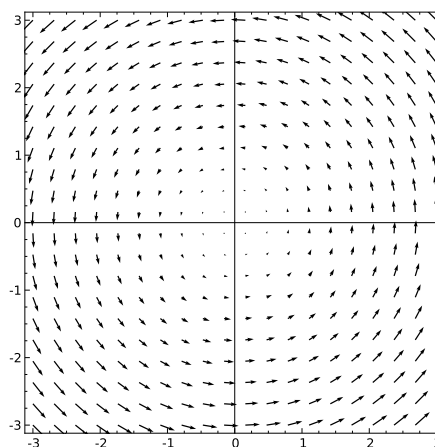
que l'on peut alors "planter" au point X .

Exemples ($n = 2$) :

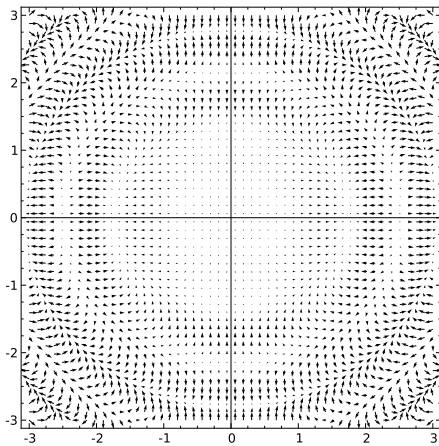
$$F(x, y) = (2x, 2y)$$



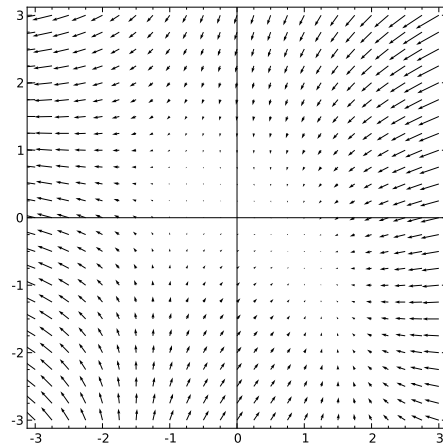
$$G(x, y) = (-y, x)$$



$$H(x, y) = (x \sin(x^2 + y^2), y \cos(x^2 + y^2))$$



$$J(x, y) = \left(-\frac{3}{2}x^2 - y + 1, -x - 2y\right)$$

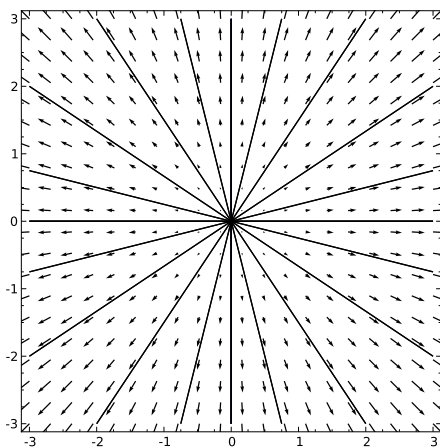


Définition (lignes de champ)

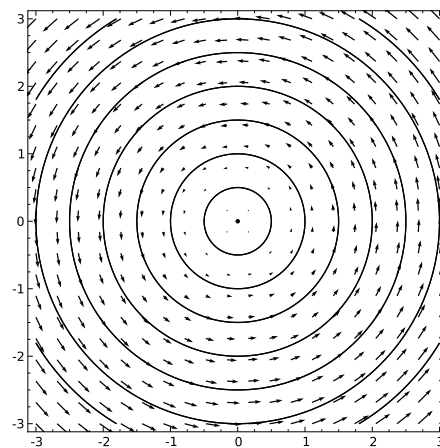
Étant donné un champ de vecteurs F de \mathbb{R}^n , on appelle *ligne de champ de F* toute courbe de \mathbb{R}^n dont le vecteur tangent en chaque point X est colinéaire au vecteur $F(X)$.

Exemples :

$$F(x, y) = (2x, 2y)$$



$$G(x, y) = (-y, x)$$



En pratique, les champs de vecteurs peuvent représenter des champs de forces, de vitesses, de déformations, etc, les lignes de champs pouvant alors représenter des trajectoires, des formes, etc.

5.4.2 Champs de gradient

Définition

Définition (champ de gradient)

Soit

$$f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R} \\ (x_1, \dots, x_n) \longmapsto f(x_1, \dots, x_n)$$

une fonction de n variables (x_1, \dots, x_n) définie sur un sous ensemble $D \subset \mathbb{R}^n$.

En tout point $X = (x_1, \dots, x_n)$ auquel f admet des dérivées partielles, on appelle *gradient de f en X* le vecteur

$$\nabla f(X) = \left(\frac{\partial f}{\partial x_1}(X), \dots, \frac{\partial f}{\partial x_n}(X) \right)$$

Par ailleurs, si le vecteur $\nabla f(X)$ existe pour tout point $X \in D$, on appelle *champ de gradient de f* le champ de vecteurs

$$\nabla f : D \longrightarrow \mathbb{R}^n \\ X = (x_1, \dots, x_n) \longmapsto \nabla f(X) = \left(\frac{\partial f}{\partial x_1}(X), \dots, \frac{\partial f}{\partial x_n}(X) \right)$$

D'un point de vue pratique, le gradient offre ainsi une autre façon de stocker les informations relatives aux variations d'une fonction de plusieurs variables, à l'aide notamment du produit scalaire. Ainsi, soit $P \in D$.

— La différentielle de f en P s'écrit

$$df_P : X = (x_1, \dots, x_n) \longmapsto \nabla f(P) \cdot X \quad (\text{produit scalaire})$$

— L'équation de l'hyperplan tangent au graphe de f en P_0 s'écrit

$$T_P : x_{n+1} = f(P) + \nabla f(P) \cdot X$$

— Le développement limité à l'ordre 1 de f au voisinage de P s'écrit

$$f(P + H) = f(P) + \nabla f(P) \cdot H + o(\|H\|)$$

— Les points critiques de f sont par définition les points $X \in D$ tels que

$$\nabla f(X) = 0_{\mathbb{R}^n}$$

Gradient et isoclines

Outre ces nouvelles représentations, le gradient offre également en dimensions 2 et 3 une nouvelle représentation géométrique des variations de f .

Ainsi, à toute fonction f de deux variables (x, y) , on peut associer :

- une carte topographique, constituée d'un ensemble d'isoclines du plan,
- son champ de gradient ∇f , qui est alors un champ de vecteurs de \mathbb{R}^2 .

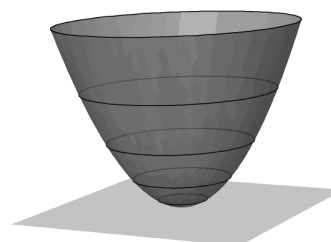
Or en superposant ces deux représentations graphiques 2D de la fonction f , on constate qu'en chaque point $P = (a, b) \in D$, le vecteur $\nabla f(P)$ est orthogonal à l'unique isocline de f passant par P : la courbe d'équation

$$f(x, y) = f(a, b)$$

Exemple : soit

$$f : (x, y) \mapsto x^2 + y^2$$

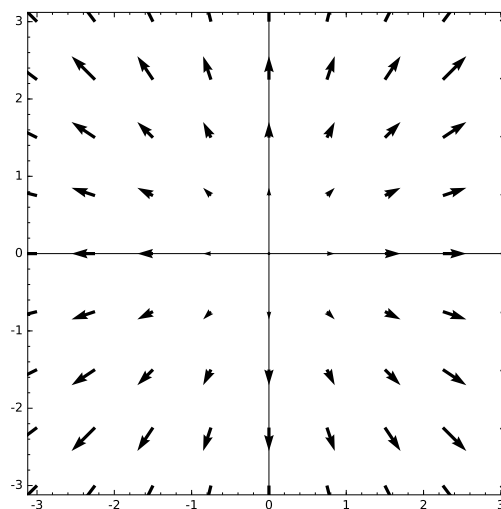
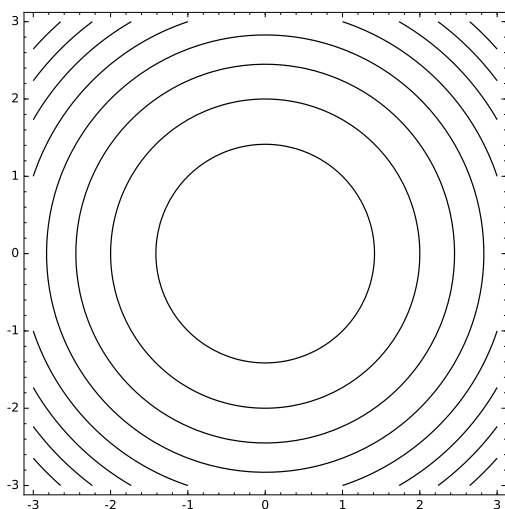
représentée ci-contre.



Les courbes de niveaux de f sont les cercles de centre O

Le champ de gradient de f est

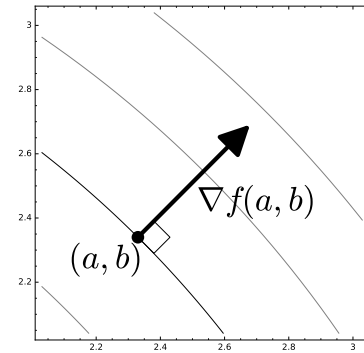
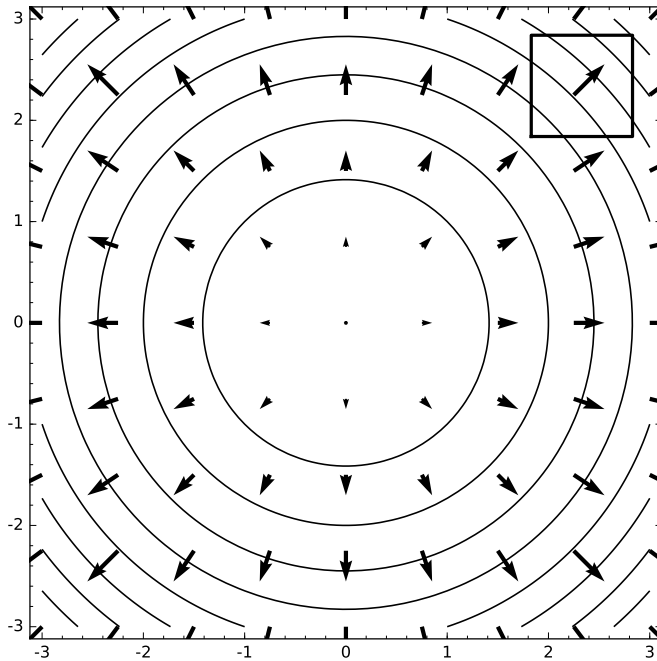
$$\nabla f : (x, y) \mapsto (2x, 2y)$$



Ainsi, si, en dimension 2, on superpose le champs de gradient d'une fonction f à sa

carte topographique, on constate qu'en chaque point $P = (a, b)$ du plan, le vecteur $\nabla f(P)$ est *orthogonal* à l'unique isocline de f passant par P , i.e. l'isocline d'équation

$$f(x, y) = f(P)$$



D'un point de vue qualitatif, cela signifie que contrairement aux isoclines qui indiquent les directions dans lesquelles la fonction f ne varie pas, le gradient indique, lui, la direction dans laquelle la variation de f est la plus importante. Le sens du vecteur $\nabla f(P)$ indique alors le sens "montante" de la fonction et sa norme donne l'intensité de cette variation maximale.

De ce fait, les lignes de champs d'un champ de gradient sont appelées *lignes de plus grande pente*.

Définition (potentiel)

Un champ de vecteur F de \mathbb{R}^n est dit *dérivant d'un potentiel* \mathcal{P} si il existe une fonction $\mathcal{P} : \mathbb{R}^n \rightarrow \mathbb{R}$ telle que

$$F = -\nabla \mathcal{P}$$

Autrement dit, un champ de vecteurs dérive d'un potentiel si c'est le champ de gradient d'une fonction à valeurs réelles. Le signe "-" s'explique par le fait qu'en général, un système dynamique "descend les potentiels".

ATTENTION : tous les champs de vecteurs ne sont pas des champs de gradients. Pour un champs de vecteur $F = (f_1, \dots, f_n)$ donné, il existe différentes méthodes permettant de déterminer s'il dérive d'un potentiel ou non. L'une des méthodes est basée

sur l'intégration successive de chaque fonction coordonnée f_i en fonction de la variable associée x_i . Une autre méthode, notamment appliquée en dimension 2 et 3, est basée sur le théorème de Schwarz. En effet, si $F = (f_1, f_2)$ dérive d'un potentiel, il existe une fonction \mathcal{P} de deux variables telle que

$$f_1 = -\frac{\partial \mathcal{P}}{\partial x} \quad \text{et} \quad f_2 = -\frac{\partial \mathcal{P}}{\partial y}$$

Mais alors, d'après le théorème de Schwarz, on doit avoir

$$\frac{\partial f_2}{\partial x} = \frac{\partial f_1}{\partial y} = -\frac{\partial^2 \mathcal{P}}{\partial x \partial y}$$

La méthode consiste donc à calculer les deux dérivées partielles ci-dessus et à les comparer : si elles sont égales, F dérive d'un potentiel, sinon, F ne dérive pas d'un potentiel.

5.4.3 Variations d'un champs de vecteurs

Il est possible d'appliquer les outils d'analyse vus plus haut à l'étude des variations d'un champs de vecteur ; l'idée étant d'appliquer les outils en question aux fonctions coordonnées du champs de vecteurs F étudié. Il est ainsi possible de définir différents objets mathématiques permettant de décrire le comportement local d'un champs de vecteurs données.

Gradient d'un champs de vecteurs

Définition (gradient d'un champ de vecteurs)

Soit

$$\begin{aligned} F : D \subset \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ X = (x_1, \dots, x_n) &\longmapsto (f_1(X), \dots, f_n(X)) \end{aligned}$$

un champ de vecteurs de \mathbb{R}^n . On appelle *gradient de F en un point $X \in D$* la matrice carrée à n lignes définie par

$$\nabla F(X) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X) & \frac{\partial f_1}{\partial x_2}(X) & \cdots & \frac{\partial f_1}{\partial x_n}(X) \\ \frac{\partial f_2}{\partial x_1}(X) & \frac{\partial f_2}{\partial x_2}(X) & \cdots & \frac{\partial f_2}{\partial x_n}(X) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(X) & \frac{\partial f_n}{\partial x_2}(X) & \cdots & \frac{\partial f_n}{\partial x_n}(X) \end{pmatrix}$$

Cette matrice, également appelée *matrice jacobienne* du champs de vecteurs F en X contient, en lignes, les vecteurs $\nabla f_i(X)$, gradients des fonctions coordonnées de F . Elle contient donc l'ensemble de l'information concernant les variations de ces fonctions coordonnées au point X .

Note : si F dérive d'un potentiel \mathcal{P} , on a, pour tout $X \in D$,

$$\nabla F(X) = -\mathcal{H}_{\mathcal{P}}(X)$$

Divergence

Définition (divergence)

Soit

$$\begin{aligned} F : D \subset \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ X = (x_1, \dots, x_n) &\longmapsto (f_1(X), \dots, f_n(X)) \end{aligned}$$

un champ de vecteurs de \mathbb{R}^n . On appelle *divergence de F en un point $X \in D$* la somme

$$\operatorname{div}(F)(X) = \frac{\partial f_1}{\partial x_1}(X) + \frac{\partial f_2}{\partial x_2}(X) + \dots + \frac{\partial f_n}{\partial x_n}(X) = \sum_{k=1}^n \frac{\partial f_k}{\partial x_k}(X)$$

Intuitivement, la divergence peut s'interpréter comme un terme de mesure locale de l'expansion (ou de la dilatation) du champ. Si on considère en effet un petit cube centré autour d'un point, le terme $\frac{\partial f_i}{\partial x_i}$ mesure le déplacement selon l'axe (Ox_i) de la i -ème composante du champs. La divergence est une somme sur toutes les directions de ces dilatations locales.

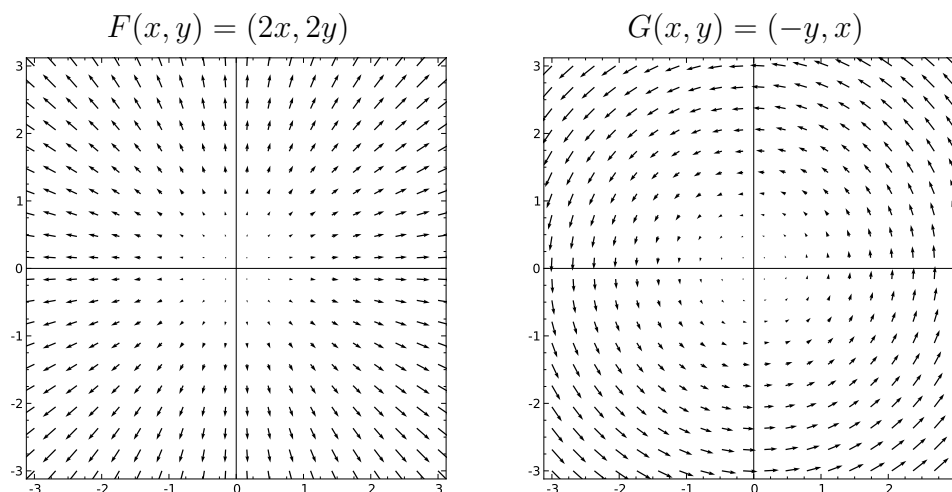
L'une des principales hypothèses (simplificatrices) de la mécanique des fluides est l'hypothèse de non compressibilité. Elle se traduit par le fait que la divergence du champs des vitesses est nulle.

Cette hypothèse permet souvent d'avoir une équation (aux dérivées partielles) supplémentaire pour étudier les mouvements d'une masse de fluide.

Note : si les fonctions coordonnées de F admettent des dérivées partielles en tout point $X \in D$, on peut alors définir *la fonction $\operatorname{div}(F)$* :

$$\begin{aligned} \operatorname{div}(F) : D \subset \mathbb{R}^n &\longrightarrow \mathbb{R} \\ X &\longmapsto \operatorname{div}(F)(X) \end{aligned}$$

Exemples : calculer la divergence en un point $(x, y) \in \mathbb{R}^2$ quelconque des champs suivants :



Rotationnel ($n = 3$)

Définition (rotationnel d'un champ de vecteurs de \mathbb{R}^3)

Soit

$$F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

$$(x, y, z) \longmapsto (f_1(x, y, z), f_2(x, y, z), f_3(x, y, z))$$

un champ de vecteurs de \mathbb{R}^3 . On appelle *rotationnel de F en X* vecteurs

$$\text{Rot}(F)(X) = \left(\frac{\partial f_3}{\partial y}(X) - \frac{\partial f_2}{\partial z}(X), \frac{\partial f_1}{\partial z}(X) - \frac{\partial f_3}{\partial x}(X), \frac{\partial f_2}{\partial x}(X) - \frac{\partial f_1}{\partial y}(X) \right)$$

Intuitivement, le rotationnel d'un champs de vecteur de l'espace donne une approximation de l'axe de rotation des déformations locales de ce champs. On l'interprète facilement comme l'axe d'un tourbillon en mécanique des fluides, le champs de vecteurs représentant la vitesse du liquide.

Note : si les fonctions coordonnées de F admettent des dérivées partielles en tout point $X \in D$, on peut définir *le champ de vecteurs* $\text{Rot}(F)$:

$$\text{Rot}(F) : D \subset \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

$$X \longmapsto \text{Rot}(F)(X)$$

Exemple : soit $F : (x, y, z) \longmapsto (-y, x, 0)$.

1. Calculer $\text{div}(F)(X)$ et $\text{Rot}(F)(X)$.
2. Interpréter les résultats obtenus.

Laplacien

Définition (laplacien)

Soit

$$f : \mathbb{R}^n \longrightarrow \mathbb{R} \\ (x_1, \dots, x_n) \longmapsto f(x_1, \dots, x_n)$$

une fonction à valeurs réelles. On appelle *laplacien de f en $X = (x_1, \dots, x_n)$* la quantité

$$\Delta(f)(X) = \frac{\partial^2 f}{\partial x_1^2}(X) + \dots + \frac{\partial^2 f}{\partial x_n^2}(X) = \sum_{k=1}^n \frac{\partial^2 f}{\partial x_k^2}(X)$$

Cette quantité (scalaire) apparaît dans de très nombreux domaines de la physique et est au cœur de toutes les équations permettant de décrire des phénomènes de diffusion (masse, énergie, chaleur, etc).

Si la fonction f admet des dérivées partielles secondes en tout point de D , on peut alors construire la fonction $\Delta(f)$:

$$\Delta(f) : D \subset \mathbb{R}^n \longrightarrow \mathbb{R} \\ X \longmapsto \Delta(f)(X)$$

Note : le laplacien de f n'est autre que la divergence du gradient de f .

L'opérateur ∇

Il est possible d'unifier l'ensemble des notions vues plus haut à l'aide de l'opérateur ∇ ("nabla"). Il s'agit d'un outil mathématique théorique défini par

$$\nabla = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right)$$

que l'on peut manipuler à la fois comme un vecteur et comme une fonction.

Ainsi,

— Le gradient d'une fonction f est la fonction ∇ appliquée à une fonction f :

$$\nabla(f) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$$

— La divergence d'un champs de vecteurs F peut être vu comme le produit scalaire

$$\operatorname{div} F = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = \sum_{k=1}^n \frac{\partial f_k}{\partial x_k}$$

— Le rotationnel d'un champs de vecteurs F peut être vu comme le produit vectoriel

$$\operatorname{Rot}(F) = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \wedge \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} \frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z} \\ \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x} \\ \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \end{pmatrix}$$

— Le laplacien d'une fonction f correspondant à la divergence de son gradient, on a

$$\Delta f = \nabla \cdot \nabla f = \nabla^2 f$$

5.4.4 Changements de variables de \mathbb{R}^n

Parmi les applications de la notion de champ de vecteur de \mathbb{R}^n , notons également la formalisation de la notion de *changement de variables* en dimensions supérieures.

D'un point de vue pratique, un changement de variable consiste à choisir un nouveau système de coordonnées, en principe mieux adapté au problème modélisé et étudié.

Formellement, un changement de variable $(x_1, \dots, x_n) \rightsquigarrow (u_1, \dots, u_n)$ de \mathbb{R}^n est donné par une fonction

$$\Phi : \begin{array}{ccc} D_u & \longrightarrow & D_x \\ U = (u_1, \dots, u_n) & \longmapsto & \Phi(U) = (x_1(U), \dots, x_n(U)) \end{array}$$

où D_u et D_x sont deux parties de \mathbb{R}^n , respectivement associés aux variables u_i et x_i .

Pour qu'un tel champ de vecteur soit un changement de variable, il faut que l'application Φ soit *bijective*. Autrement dit, il doit exister un champ de vecteurs

$$\Phi^{-1} : \begin{array}{ccc} D_x & \longrightarrow & D_u \\ X = (x_1, \dots, x_n) & \longmapsto & \Phi^{-1}(X) = (u_1(X), \dots, u_n(X)) \end{array}$$

En pratique, on peut montrer qu'un champ de vecteurs Φ est *bijectif* si et seulement pour tout $U \in D_u$, son gradient

$$\nabla\Phi(U) = \left(\begin{array}{c} \frac{\partial u_i}{\partial u_j}(U) \\ \end{array} \right)_{1 \leq i, j \leq n}$$

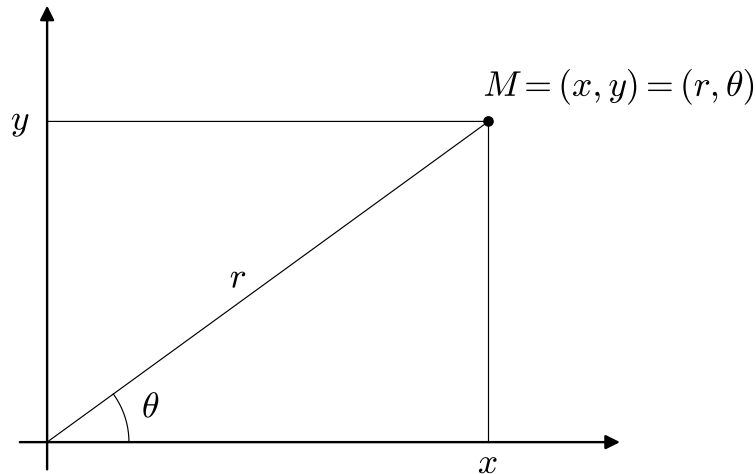
est inversible. Dans ce cas, on peut également montrer que

$$[\nabla(\Phi)(U)]^{-1} = \nabla\Phi^{-1}(\Phi(U))$$

Exemples : en dimensions 2 et 3, on peut ainsi formaliser les changements de variables usuels du plan ou de l'espace.

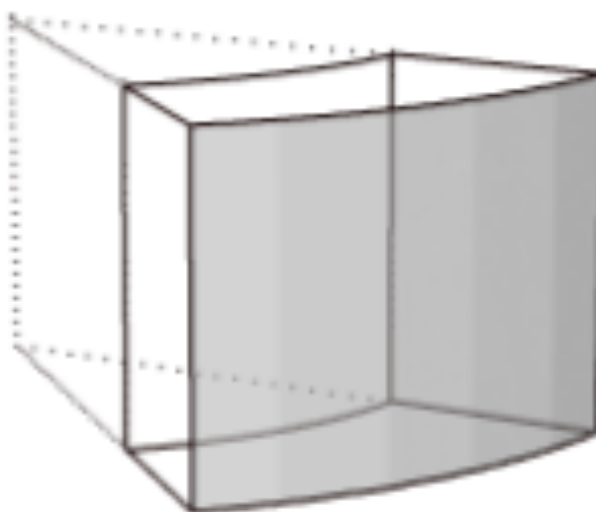
- Dans le plan, le passage au coordonnées *polaires* $(x, y) \rightsquigarrow (r, \theta)$, représentées ci-dessous correspond au champ de vecteurs

$$\begin{array}{lcl} \Phi : \mathbb{R}^+ \times [0, 2\pi[& \longrightarrow & \mathbb{R}^2 \\ (r, \theta) & \longmapsto & (r \cos \theta, r \sin \theta) \end{array}$$



- Dans l'espace,
- le passage en coordonnées *cylindriques* $(x, y, z) \rightsquigarrow (r, \theta, z)$ représentées ci-dessous correspond au champ de vecteurs

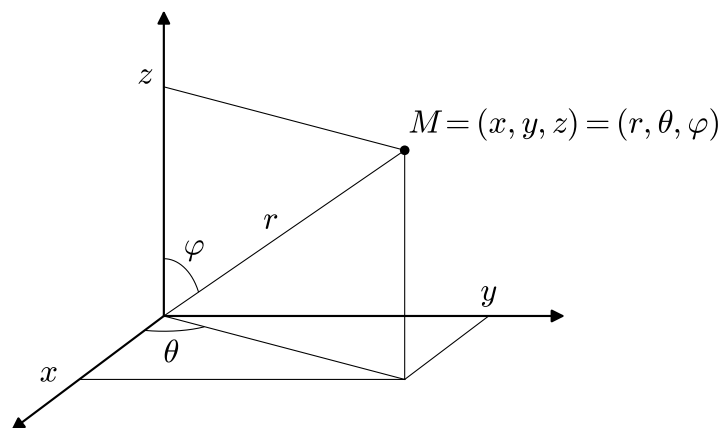
$$\begin{array}{lcl} \Phi : \mathbb{R}^+ \times [0, 2\pi[\times \mathbb{R} & \longrightarrow & \mathbb{R}^3 \\ (r, \theta, z) & \longmapsto & (r \cos \theta, r \sin \theta, z) \end{array}$$



- le passage en coordonnées *sphériques* $(x, y, z) \rightsquigarrow (r, \theta, \varphi)$ représentées ci-dessous correspond au champ de vecteurs

$$\Phi : \mathbb{R}^+ \times [0, 2\pi[\times [0, \pi] \longrightarrow \mathbb{R}^3$$

$$(r, \theta, \varphi) \longmapsto (r \cos \theta \sin \varphi, r \sin \theta \sin \varphi, r \cos \varphi)$$



5.5 Équations aux dérivées partielles

Une équation aux dérivée partielle (EDP) est une équation liant une fonction u de plusieurs variables (x_1, \dots, x_n) à ses dérivées partielles successives. Les EDP servent, au même titre que les équations différentielles, à modéliser l'évolution de systèmes physiques (notamment via les opérateurs vus plus haut). La plupart de ces systèmes dépendant en effet de plusieurs paramètres, les fonctions de plusieurs variables et les EDP permettent de construire des modèles plus fidèles, celles-ci étant basées la plupart du temps sur les

opérateurs différentiels vus plus haut.

Comme pour les équations différentielles ordinaires, il est possible de classer les EDP selon différentes caractéristiques (équations d'ordre 1, d'ordre 2, d'ordre n , équations linéaires, homogènes, etc) et il existe des méthodes de résolutions exactes pour certaines de ces équations.

À l'image des EDO, on retrouve également dans le traitement des EDP la notion de conditions initiales (ou conditions au bord) et d'unicité des solutions.

Cependant, comme pour les EDO, il est rare que ces méthodes de résolutions s'appliquent aux EDP issues des lois physiques. On doit alors recourir à des hypothèses simplificatrices ou se tourner vers des méthodes numériques de résolution approchées.

Parmi les EDP célèbres, citons

1. *L'équation de transport* : $\frac{\partial c}{\partial t} + \operatorname{div}(c.V) = 0$

Elle permet de modéliser le déplacement d'un flot de particules entraînées par un fluide au sein d'une canalisation.

Ici, $c(x, y, z, t)$ représente la concentration du milieu étudié en nombre de particules par unité de volume. Elle dépend à la fois de la position dans la canalisation et du temps. Le champs de vecteur $V = (u(t), v(t), w(t))$ correspond lui au champs de vitesses du fluide porteur.

Il s'agit ici d'une EDP homogène, linéaire, d'ordre 1.

2. *L'équation de la chaleur* : $\Delta u = \frac{1}{\alpha} \frac{\partial u}{\partial t}$

Elle permet de modéliser l'évolution de la température $u(x, y, z, t)$ dans un milieu donné.

Il s'agit ici d'une EDP linéaire d'ordre 2, homogène, à coefficients constants.

3. *L'équation de propagation des ondes* : $\Delta u = \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2}$

Elle permet de modéliser le déplacement d'une onde $u(x, y, z, t)$ (son, lumière, etc) dans un milieu donné.

4. *L'équation de Stokes* : $\eta \Delta V = \nabla p - \rho F$

Elle permet de modéliser l'écoulement d'un fluide dans un milieu confiné. Cette

équation, fondamentale en mécanique des fluides, fait intervenir

- la vitesse V du fluide (sous la forme d'un champs de vecteurs),
- la pression p au sein du fluide,
- une force extérieure agissant sur le fluide (par exemple la gravité),
- des constantes η et ρ caractéristiques du fluide et du milieu dans lequel évolue le fluide.

Chapitre 6

Compléments d'intégration

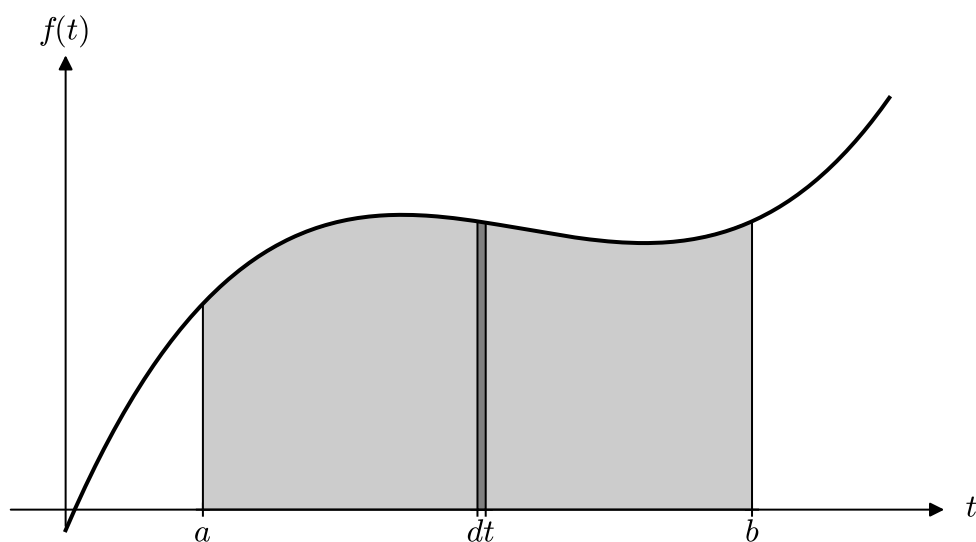
6.1 Intégrales multiples

6.1.1 Motivation

Le calcul intégral a pour objectif de donner un cadre formel à la sommation sur des ensembles infinis ayant la puissance du continu. Ainsi, en une variable, le calcul intégral permet de définir des sommes le long d'un segment $[a, b]$ de la droite réelle. Précisément, si l'on note dt une partie infinitésimale de l'intervalle $[a, b]$, pour toute fonction f d'une variable réelle définie sur $[a, b]$, l'intégrale

$$I = \int_{[a,b]} f(t)dt = \int_a^b f(t)dt$$

correspond à la somme de toutes les valeurs de f le long de l'intervalle $[a, b]$.



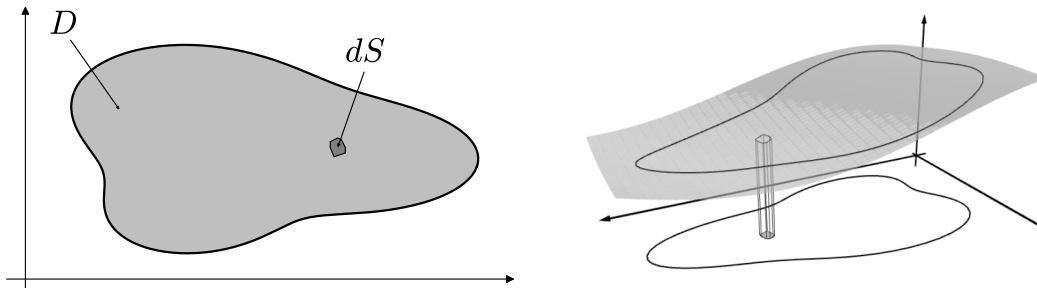
Si f représente une hauteur sur l'axe vertical, la somme de toutes ces hauteurs donne l'aire totale sous la courbe de f . Si f représente une répartition de masse le long d'un segment, l'intégrale de f le long du segment donne la masse totale associée à la répartition, etc.

Or il est possible d'étendre la notion de calcul intégral aux dimensions supérieures.

Ainsi, si D représente une partie du plan, muni d'un repère, en notant dS une partie infinitésimale (2D) de D , pour toute fonction f définie sur D , l'intégrale double

$$I = \iint_D f(M) dS$$

représente la somme de toutes les valeurs de f sur le domaine D .



Là encore, selon la nature de la fonction f , on peut interpréter l'intégrale double de f sur D en termes de surface, de volume, de masse, etc.

De même, si D représente une partie de l'espace muni d'un repère, en notant dV une partie infinitésimale (3D) de D , pour toute fonction f définie sur D , l'intégrale triple

$$I = \iiint_D f(M) dV$$

représente la somme de toutes les valeurs de f sur le domaine D .

Ces extensions aux dimensions supérieures permettent alors de représentées différentes grandeurs caractéristiques des objets 2D ou 3D, notamment étudiés en mécanique :

— Aires et volumes :

— L'aire d'un domaine D du plan est

$$\mathcal{A}(D) = \iint_D dS$$

— Le volume d'un domaine D de l'espace est

$$\mathcal{V}(D) = \iiint_D dV$$

- *Masse* : en notant $\mu : D \rightarrow \mathbb{R}$ une distribution de masse sur le domaine D , la masse totale de D est

$$\mathcal{M}(D_{2D}) = \iint_D \mu(M) dS \quad \text{ou} \quad \mathcal{M}(D_{3D}) = \iiint_D \mu(M) dV$$

- *Centre d'inertie* : par définition, le centre d'inertie (ou centre de gravité) d'un domaine du plan ou de l'espace est *le point moyen* de ce domaine.

Si l'on note O le centre du repère dans lequel on travaille, le centre d'inertie de D est le point G vérifiant

$$\overrightarrow{OG} = \frac{1}{\mathcal{A}(D)} \iint \overrightarrow{OM} dS$$

Ainsi, dans le plan, les coordonnées cartésiennes (x_G, y_G) du centre de gravité G sont

$$x_G = \frac{1}{\mathcal{A}(D)} \iint_D x dS \quad y_G = \frac{1}{\mathcal{A}(D)} \iint_D y dS$$

Ces coordonnées correspondent aux moyennes des fonctions coordonnées des points M de D .

Notes : on peut bien entendu étendre cela aux domaines de l'espace. De plus, en ajoutant une distribution de masse $\mu(x, y)$ ou $\mu(x, y, z)$ définie sur le domaine D , on peut ainsi définir et déterminer le centre d'inertie d'un domaine non homogène.

- *Moment d'inertie* : si l'on se fixe un axe Δ du plan ou de l'espace, le moment d'inertie d'un domaine D par rapport à cet axe est, par définition,

$$I_\Delta = \iint_D d(M, \Delta)^2 dS \quad \text{ou} \quad I_\Delta = \iiint_D d(M, \Delta)^2 dV$$

L'objectif de ce chapitre est de présenter les méthodes pratiques de calcul pour les intégrales multiples. On verra que ce calcul dépend principalement du système de coordonnées choisies pour décrire le domaine D sur lequel on souhaite intégrer. Ce choix influence en particulier l'expression de l'élément infinitésimal dS ou dV . Cependant, dans tous les cas, le domaine D sur lequel on souhaite intégrer est toujours décrit à l'aide de deux variables (u, v) (en 2D) ou trois variables (u, v, w) (en 3D) et toute intégrale sur un même domaine se réduit à une intégrale double de la forme

$$I = \iint_D \varphi(u, v) du dv \quad \text{ou} \quad I = \iiint_D \varphi(u, v, w) du dv dw$$

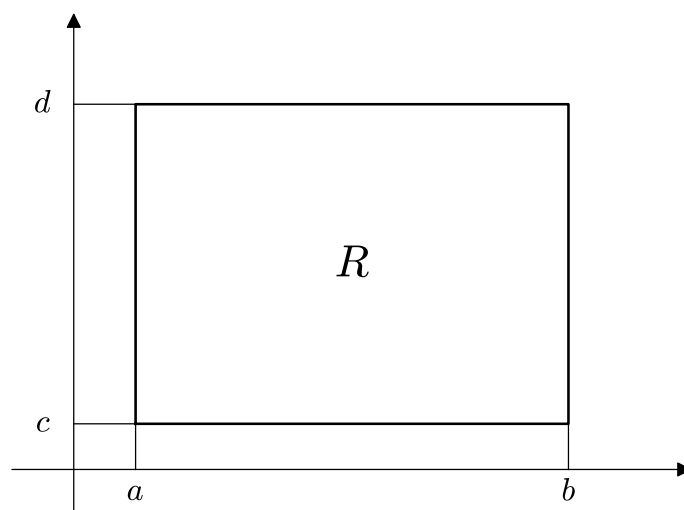
Dans un premier temps, on considérera donc les coordonnées cartésiennes (x, y) ou (x, y, z) . Dans ce cas, l'élément infinitésimal, généré par des variations infinitésimales dx , dy (et dz) des coordonnées donne

$$dS = dx dy$$

$$dV = dx dy dz$$

6.1.2 Intégration sur un rectangle

Les rectangles du plans sont les domaines les plus simples à représenter et manipuler à l'aide des coordonnées cartésiennes associées à un repère orthonormé. Ainsi, étant donné un rectangle du plan



Formellement, l'ensemble des points intérieurs au rectangle R ci-dessus peut être représenté par le sous ensemble de \mathbb{R}^2 défini par

$$R = \{(x, y) \in \mathbb{R}^2 / a \leq x \leq b, c \leq y \leq d\}$$

a, b, c et d étant des réels fixés vérifiant $a < b$ et $c < d$.

Ainsi, si f est une fonction des variables (x, y) définie et continue sur R , le calcul de l'intégrale

$$I = \iint_D f(x, y) dx dy$$

peut se réduire au calcul de deux intégrales simples successives. Précisément,

Théorème (Fubini)

Soit f une fonction continue définie sur le rectangle

$$R = \{(x, y) \in \mathbb{R}^2 / a \leq x \leq b, c \leq y \leq d\}$$

Alors

$$\iint_R f(x, y) dx dy = \int_{x=a}^b \left(\int_{y=c}^d f(x, y) dy \right) dx = \int_{y=c}^d \left(\int_{x=a}^b f(x, y) dx \right) dy$$

En pratique, on peut donc intégrer toute fonction définie sur un rectangle à l'aide de deux intégrations successives en appliquant les méthodes de calcul propres aux intégrales simples. Ainsi, pour le calcul de I , on commence par calculer l'intégrale

$$\tilde{I}(x) = \int_{y=c}^d f(x, y) dy$$

en considérant x comme une constante. On obtient alors une fonction ne dépendant plus que de x , que l'on intègre ensuite sur l'intervalle $[a, b]$. La formule ci-dessus indique que l'on peut également procéder en commençant par traiter la variable x , puis y .

Note : on appelle *fonction à variables séparées* toute fonction de deux variables de la forme

$$f(x, y) = \varphi(x)\psi(y)$$

Si f est une fonction à variables séparées, on peut montrer que

$$I = \iint_R f(x, y) dS = \left(\int_{x=a}^b \varphi(x) dx \right) \left(\int_{y=c}^d \psi(y) dy \right)$$

Exemple : retrouver l'aire du rectangle R ainsi que la position de son centre d'inertie à l'aide des formules précédentes.

Enfin, on peut généraliser cette méthode aux parallélépipèdes de l'espace. Précisément, un parallélépipède de \mathbb{R}^3 est un sous ensemble de \mathbb{R}^3 de la forme

$$P = \{(x, y, z) \in \mathbb{R}^3 / a_1 \leq x \leq b_1, a_2 \leq y \leq b_2, a_3 \leq z \leq b_3\}$$

a_i, b_i étant des réels fixés vérifiant $a_i < b_i$.

Pour toute fonction f dépendant des variables (x, y, z) , on a alors

$$I = \iiint_P f(x, y, z) dV = \int_{x=a_1}^{b_1} \left(\int_{y=a_2}^{b_2} \left(\int_{z=a_3}^{b_3} f(x, y, z) dz \right) dy \right) dx$$

Autrement dit, toute intégrale sur un parallélépipède de l'espace se résume au calcul de trois intégrales simples successives. De plus, les six bornes étant constantes, il est encore possible de traiter les variables dans n'importe quel ordre.

On peut en outre généraliser aux fonctions de n variables, définies et continues sur toute partie de \mathbb{R}^n de la forme

$$D = \{(x_1, \dots, x_n) \in \mathbb{R}^n / \forall i \in [1, n], a_i \leq x_i \leq b_i\}$$

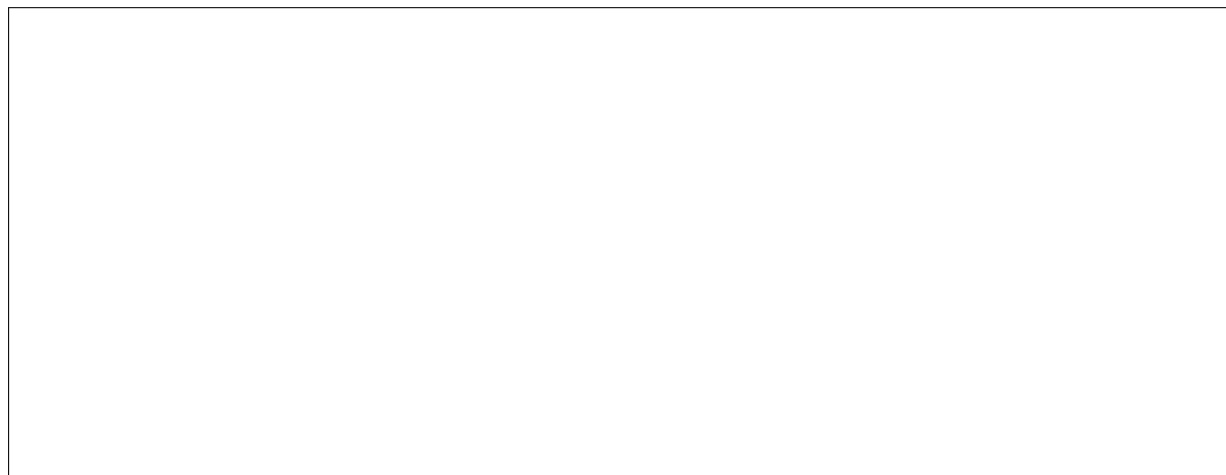
$a_1, b_1, \dots, a_n, b_n$ étant des réels fixés.

6.1.3 Domaines variés

À l'aide des coordonnées cartésiennes, il est possible d'étendre le calcul explicite d'une intégrale multiple à des domaines plus complexes du plan ou de l'espace.

En dimension 2, considérons la partie D du plan délimitée par les courbes de deux fonctions φ_1 et φ_2 définies sur un intervalle $[a, b]$:

$$D = \{(x, y) \in \mathbb{R}^2 / a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\}.$$

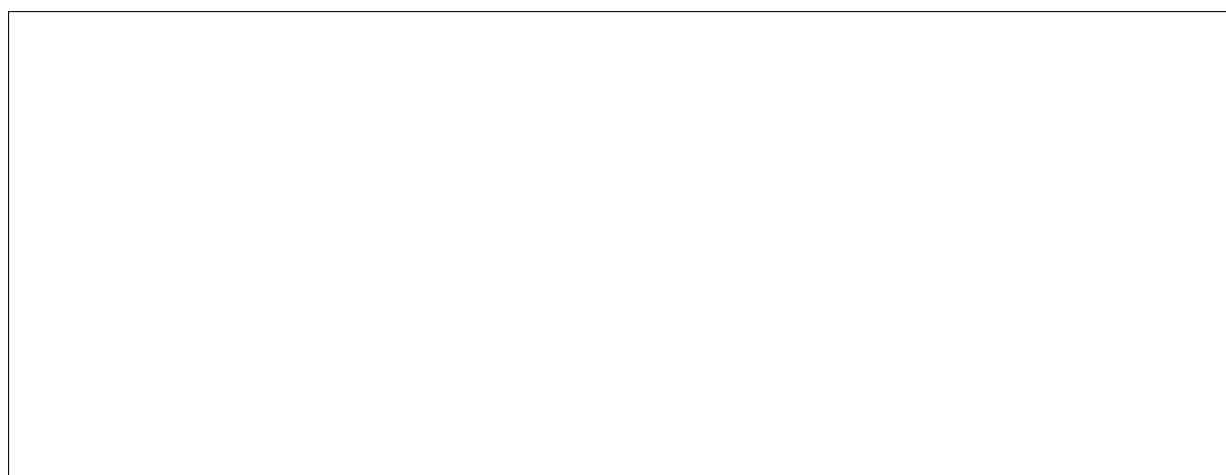


Cette description du domaine D produit les bornes nécessaires à l'interprétation de toute intégrale double sur D en termes d'intégrales simples imbriquées :

$$\iint_D f(x, y) dx dy = \int_{x=a}^b \left(\int_{y=\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy \right) dx$$

Ainsi, comme sur un rectangle, il est possible ici de ramener le calcul de toute intégrale sur D au calcul de deux intégrales simples successives. Cependant, contrairement au cas des rectangles, il est nécessaire, ici, de commencer par traiter la variable y ; commencer par la variable x n'ayant ici aucun sens.

En échangeant les rôles des variables x et y , il est cependant possible d'étendre ces notions aux domaines du plans de la forme



associés aux parties de \mathbb{R}^2 définies par

$$D = \{(x, y) \in \mathbb{R}^2 / c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\}$$

c et d étant des réels fixés et ψ_1 et ψ_2 étant des fonctions connues.

Si f est alors une fonction des variables (x, y) définie sur D , on a

$$\iint_D f(x, y) dx dy = \int_{y=c}^d \left(\int_{x=\psi_1(y)}^{\psi_2(y)} f(x, y) dx \right) dy$$

Exemple : déterminer l'aire et la position du centre d'inertie du triangles (ABC) où

$$A = (0, 0), \quad B = (0, 1), \quad C = (1, 0)$$

Enfin, il est là encore possible de généraliser cela aux dimensions supérieures. Précisément, soit D une partie de \mathbb{R}^3 définie par

$$D = \{(x, y, z) \in \mathbb{R}^3 / a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x), \psi_1(x, y) \leq z \leq \psi_2(x, y)\}$$

où

- a et b sont des réels fixés,
- φ_1 et φ_2 sont des fonctions connues d'une variable,
- ψ_1 et ψ_2 sont des fonctions connues de deux variables

Pour toute fonction f continue des variables (x, y, z) définies sur D , on a

$$\iiint_D f(x, y, z) dx dy dz = \int_{x=a}^b \left(\int_{y=\varphi_1(x)}^{\varphi_2(x)} \left(\int_{z=\psi_1(x,y)}^{\psi_2(x,y)} f(x, y, z) dz \right) dy \right) dx$$

6.1.4 Interprétation géométrique (2D)

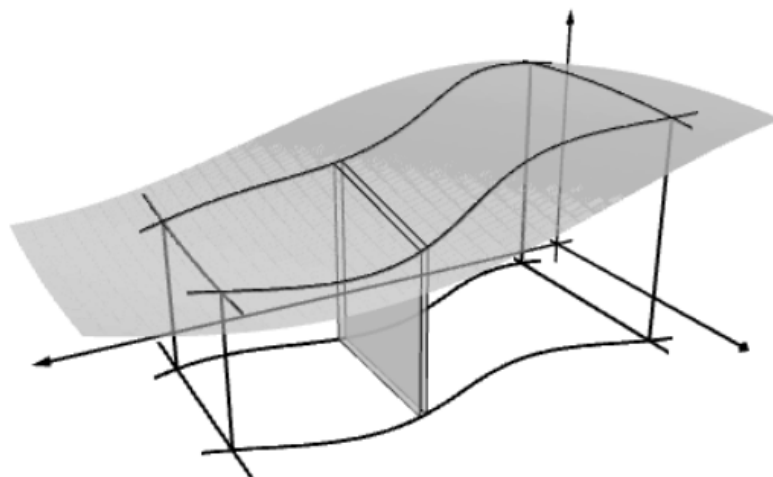
La transformation d'une intégrale double en deux intégrales simples peut se traduire d'un point de vue géométrique. En effet, dans l'espace muni d'un repère orthonormé, l'intégrale

$$I = \iint_D f(x, y) dx dy$$

peut être représentée par la mesure du volume délimité par la surface \mathcal{S}_f représentant f au dessus du domaine D . Chacune des intégrales

$$\tilde{I}(x) = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy$$

peut être représentée par le volume d'une section de ce volume d'épaisseur dx .



L'intégrale $I = \int_a^b \tilde{I}(x) dx$ correspond alors à la somme de tous ces volumes.

On verra en exercices que cette approche, associée un système de coordonnées adapté, permet de décrire et calculer de nombreuses intégrales multiples.

6.1.5 Propriétés du calcul intégral

Le lien existant entre intégrales multiples et intégrales simples permet en pratique d'exploiter les techniques classiques du calcul intégral en une variable (primitives usuelles, intégration par parties, etc) pour évaluer des intégrales multiples. D'un point de vue théorique, ce lien permet également d'établir certaines propriétés (là encore classiques) vérifiées par les intégrales multiples. Précisément,

- *Linéarité de l'intégrale* : soient f et g deux fonctions continues définies sur une partie D du plan et soient λ et μ deux réels quelconques. On a

$$\iint_D (\lambda \cdot f(M) + \mu \cdot g(M)) dS = \lambda \cdot \iint_D f(M) dS + \mu \cdot \iint_D g(M) dS$$

- *Relation de Chasles* : soient D un domaine du plan et f une fonction continue définie sur D . Pour toute partition $(D_i)_{1 \leq i \leq n}$ de D , on a

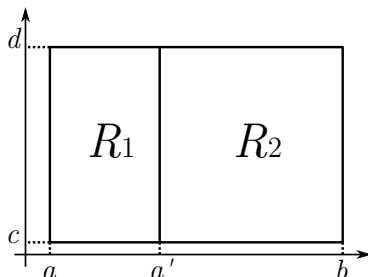
$$\iint_D f(M) dS = \sum_{i=1}^n \iint_{D_i} f(M) dS$$

Notes :

- Les propriétés ci-dessus ont été énoncées pour les intégrales doubles mais restent valables quelque soit le nombre de variables.

- La relation de Chasles permet d'étendre le calcul intégrale en dimensions supérieures aux domaines constitués d'une juxtaposition de domaines intégrables (i.e. que l'on peut décrire de façon à expliciter les bornes de chacune des intégrales simples conduisant au calcul de l'intégrale complète).

Exercice : Considérons un rectangle R du plan, découpé en deux rectangles R_1 et R_2 comme représenté ci-dessous :



Montrer que pour toute fonction f des variables (x, y) définie sur R , on a

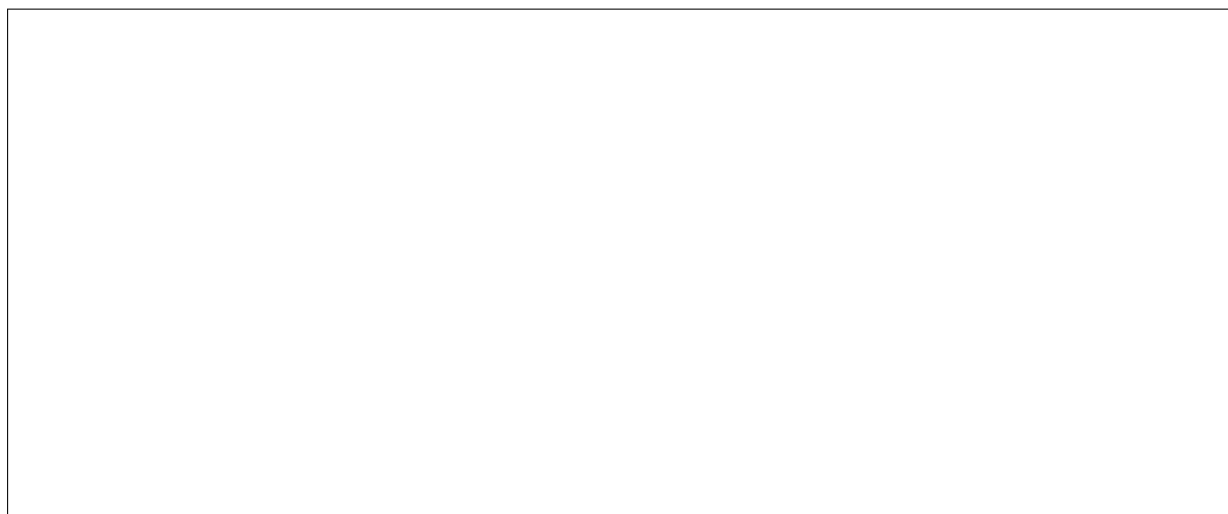
$$\iint_R f(x, y) dx dy = \iint_{R_1} f(x, y) dx dy + \iint_{R_2} f(x, y) dx dy$$

6.1.6 Changement de variables en dimensions supérieures

En pratique, selon la géométrie du problème étudié, les coordonnées cartésiennes ne sont par toujours les plus adaptées au calcul intégral.

Or l'utilisation d'un système alternatif de coordonnées influence directement le calcul d'une intégrale multiple en influençant notamment l'expression de l'élément infinitésimal dV (ou dS en dimension 2).

Ainsi, en coordonnées polaires, un point est repéré par ses coordonnées (r, θ) . Un petit élément de surface dS est alors obtenu en perturbant chacune de ces coordonnées par un couple $(dr, d\theta)$:



On peut alors montrer que $dS = r dr d\theta$. Pour toute fonction f des variables (x, y) définie sur un domaine D du plan, on a alors

$$\iint_D f(M) dS = \iint_{D_{\text{cart.}}} f(x, y) dx dy = \iint_{D_{\text{pol.}}} f(r \cos \theta, r \sin \theta) \cdot r dr d\theta$$

les notations $D_{\text{cart.}}$ et $D_{\text{pol.}}$ désignant respectivement le domaine D décrit à l'aide des coordonnées cartésiennes et polaires.

Avant de présenter les outils théoriques permettant de déterminer l'élément de volume dV associé à un système de coordonnées choisi, rappelons les différents systèmes de coordonnées classiques en dimension 3

Coordonnées alternatives en dimension 3

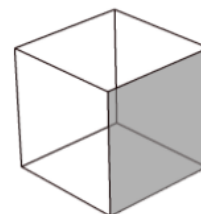
Soit $\mathcal{R} = (\mathcal{O}; \vec{i}, \vec{j}, \vec{k})$ un repère orthonormé de l'espace. Il existe trois systèmes principaux de coordonnées permettant de localiser les points de l'espace à partir du centre \mathcal{O} du repère \mathcal{R} :

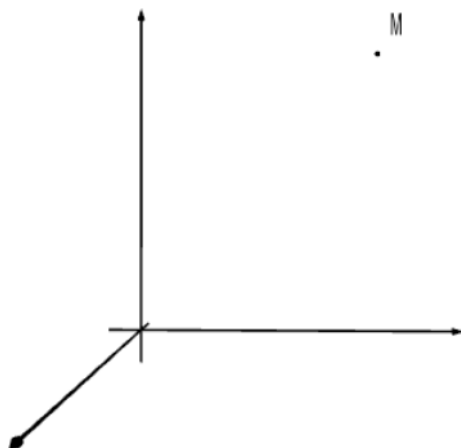
— *Coordonnées cartésiennes*

$$\mathcal{E} = \{(x, y, z), x \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{R}\}$$

Des variations infinitésimales dx , dy et dz des coordonnées cartésiennes génèrent un parallélépipède infinitésimal de volume

$$dV = dx dy dz$$

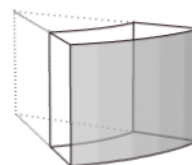
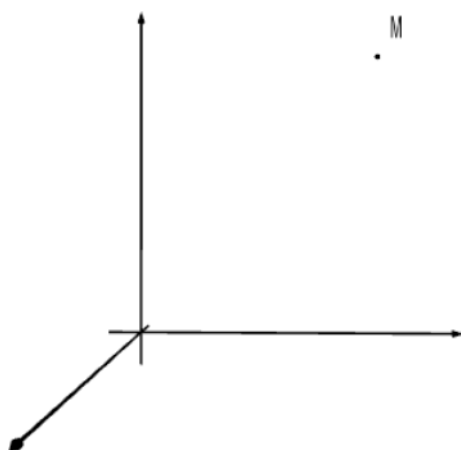


— *Coordonnées cylindriques*

$$\mathcal{E} = \{(r, \theta, z), r \geq 0 \in \mathbb{R}, 0 \leq \theta < 2\pi, z \in \mathbb{R}\}$$

Des variations infinitésimales dr , $d\theta$ et dz des coordonnées cylindriques génèrent une partie infinitésimale de cylindre de volume

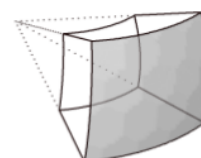
$$dV = r dr d\theta dz$$

— *Coordonnées sphériques*

$$\mathcal{E} = \{(r, \theta, \varphi), r \geq 0, 0 \leq \theta < 2\pi, 0 \leq \varphi \leq \pi\}$$

Des variations infinitésimales dr , $d\theta$ et $d\varphi$ des coordonnées sphériques génèrent une infinitésimale de sphère de volume

$$dV = r^2 \sin \varphi dr d\theta d\varphi$$



Formule générale du changement de variables

Rappel : en une variable, un changement de variable $x \rightsquigarrow u$ sur un intervalle $[a, b]$ est formellement représenté par une bijection

$$\begin{aligned} \varphi : [a, b] &\longrightarrow [c, d] \\ x &\longmapsto u = \varphi(x) \end{aligned}$$

En notant

$$\psi = \varphi^{-1} : [c, d] \longrightarrow [a, b] \\ u \longmapsto x = \psi(u)$$

on a alors

$$\int_a^b f(x) dx = \int_c^d f(\psi(u)) \cdot \psi'(u) du$$

Sur \mathbb{R}^n , un changement de variables $X = (x_1, \dots, x_n) \rightsquigarrow U = (u_1, \dots, u_n)$ est de même formellement représenté par un champ de vecteurs bijectif

$$\begin{aligned} \Phi : D_1 \subset \mathbb{R}^n &\longrightarrow D_2 = \Phi(D_1) \\ X = (x_1, \dots, x_n) &\longmapsto U = (u_1(X), \dots, u_n(X)) \end{aligned}$$

En notant

$$\begin{aligned} \Psi = \Phi^{-1} : D_2 &\longrightarrow D_1 \\ U = (u_1, \dots, u_n) &\longmapsto X = (x_1(U), \dots, x_n(U)) \end{aligned}$$

on a alors

$$\iint \cdots \int_{D_1} f(X) dx_1 \cdots dx_n = \iint \cdots \int_{D_2} f(\Psi(U)) \cdot |\text{Jac}(\Psi)(U)| du_1 \cdots du_n$$

où $\text{Jac}(\Psi)(U)$ est le déterminant de la matrice jacobienne de Ψ en U :

$$\text{Jac}(\Psi)(U) = \begin{vmatrix} \frac{\partial x_1}{\partial u_1}(U) & \cdots & \frac{\partial x_1}{\partial u_n}(U) \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial u_1}(U) & \cdots & \frac{\partial x_n}{\partial u_n}(U) \end{vmatrix}$$

En pratique, un changement de variables $(x_1, \dots, x_n) \rightsquigarrow (u_1, \dots, u_n)$ dans une intégrale de la forme

$$I = \iint \cdots \int_D f(X) dx_1 \cdots dx_n$$

se fait donc en trois étapes :

- Décrire l'ensemble D à l'aide des nouvelles variables (u_1, \dots, u_n) ,
- Identifier la fonctions $g : (u_1, \dots, u_n) = f(\Psi(u_1, \dots, u_n))$,
- Calculer le jacobien $\text{Jac}(\Psi)(U)$ et en déduire l'élément de volume dV associé.

Exemples :

— *Passage en coordonnées polaires*

1. Déterminer explicitement les fonctions coordonnées $x(r, \theta)$ et $y(r, \theta)$ du passage en coordonnées polaires $(x, y) \rightsquigarrow (r, \theta)$.
2. Calculer le jacobien $\text{Jac}(\Psi)(r, \theta)$ et retrouver l'expression de l'élément de surface dS en coordonnées polaires.
3. Retrouver l'aire du disque de rayon R .

— *Passage en coordonnées cylindriques*

1. Déterminer explicitement les fonctions coordonnées $x(r, \theta, z)$, $y(r, \theta, z)$ et $z(r, \theta, z)$ du passage en coordonnées cylindriques $(x, y, z) \rightsquigarrow (r, \theta, z)$.
2. Calculer le jacobien $\text{Jac}(\Psi)(r, \theta, z)$ et retrouver l'expression de l'élément de surface dV en coordonnées cylindriques.

— Retrouver de même l'expression de l'élément de volume dV en coordonnées sphériques.

6.2 Intégrales curvilignes, intégrales de surfaces

Il est également possible d'étendre le calcul intégral à des ensembles "courbes", i.e. des courbes du plan ou de l'espace et des surfaces de l'espace. Pour cela, il est d'abord nécessaire de définir les éléments infinitésimaux des ensembles sur lesquels on souhaite faire du calcul intégral. Cela passe en général par *une paramétrisation* de l'ensemble en question.

6.2.1 Paramétrisation

Paramétrisation d'une courbe

Définition (arc paramétré)

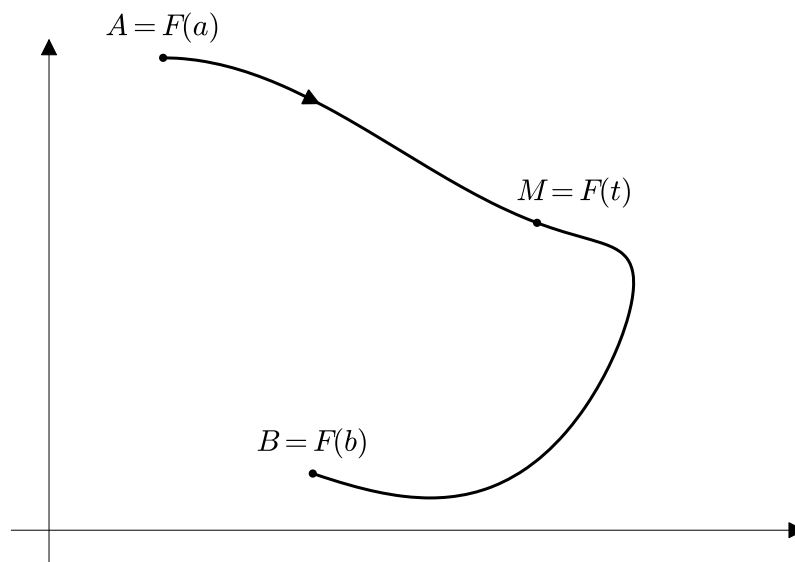
Soient A et B deux points du plan ou de l'espace.

1. On appelle *arc* \widehat{AB} tout chemin continu du plan ou de l'espace reliant les points A et B .
2. On appelle *paramétrisation* de l'arc \widehat{AB} toute fonction vectorielle de la forme

$$F : [a, b] \longrightarrow \mathbb{R}^2 \text{ ou } \mathbb{R}^3 \\ t \longmapsto F(t) = (x(t), y(t)) \text{ ou } F(t) = (x(t), y(t), z(t))$$

telle que

- $F(a) = A$
- $F(b) = B$
- $\forall t \in [a, b], F(t) \in \widehat{AB}$
- $\forall M \in \widehat{AB}, \exists t \in [a, b]$ tel que $M = F(t)$.



Exemples :

- Une paramétrisation du cercle de centre $(0, 0)$ et de rayon $R > 0$ parcouru dans le sens trigonométrique est

$$F : [0, 2\pi[\longrightarrow \mathbb{R}^2 \\ t \longmapsto (R \cdot \cos(t), R \cdot \sin(t))$$

- Une paramétrisation du demi-cercle supérieur de centre $(0, 0)$ et de rayon $R > 0$ parcouru dans le sens trigonométrique est

$$F : [0, \pi] \longrightarrow \mathbb{R}^2 \\ t \longmapsto (R \cdot \cos(t), R \cdot \sin(t))$$

- Si \widehat{AB} est la courbe d'une fonction f au dessus d'un intervalle $[a, b]$, une paramétrisation naturelle de \widehat{AB} est

$$F : [a, b] \longrightarrow \mathbb{R}^2 \\ t \longmapsto (t, f(t))$$

Ainsi, une autre paramétrisation du demi-cercle supérieur de rayon $R > 0$ est

$$F : [-R, R] \longrightarrow \mathbb{R}^2 \\ t \longmapsto (t, \sqrt{1 - t^2})$$

Paramétrisation d'une surface de l'espace

Définition (surface paramétrée)

Soit \mathcal{S} une surface de l'espace, muni d'un repère orthonormé. On appelle paramétrisation de \mathcal{S} toute application

$$G : D \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^3 \\ (u, v) \longmapsto G(u, v) = (x(u, v), y(u, v), z(u, v))$$

telle que pour tout $M \in \mathbb{R}^3$, on a

$$M \in \mathcal{S} \iff \exists (u, v) \in D / G(u, v) = M$$

Exemples :

- Si \mathcal{S} est la surface associée à une fonction f de deux variables définie sur D , une paramétrisation de \mathcal{S} est

$$G : D \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^3 \\ (u, v) \longmapsto (u, v, f(u, v))$$

- En s'appuyant sur les coordonnées cylindriques, construire une paramétrisation du

cylindre vertical de rayon R et de hauteur h .

- En s'appuyant sur les coordonnées sphériques, construire une paramétrisation de la sphère de rayon R .

6.2.2 Intégrales curvilignes

La notion de paramétrisation permet de construire formellement un élément infinitésimal pour une courbe donnée. Précisément, soit \widehat{AB} un arc du plan (ou de l'espace) et $F : [a, b] \rightarrow \mathbb{R}^2$ une paramétrisation de \widehat{AB} . On peut générer un élément infinitésimal de \widehat{AB} en un point $M = F(t) \in \widehat{AB}$ à l'aide d'une variation infinitésimale dt du paramètre t :

$$M = F(t) \rightsquigarrow M' = F(t + dt)$$

Si les fonctions coordonnées $x(t)$ et $y(t)$ de la paramétrisation F sont dérivables en t , la formule de Taylor à l'ordre 1 appliquées à x et y donne

$$\begin{cases} x(t + dt) = x(t) + x'(t).dt + o(dt^2) \\ y(t + dt) = y(t) + y'(t).dt + o(dt^2) \end{cases}$$

En négligeant les termes de degré supérieur ou égal à 2 (en dt), on peut alors assimiler l'arc $\widehat{MM'}$ au vecteur $\vec{d\ell} = \overrightarrow{MM'}$ dont les coordonnées sont

$$\vec{d\ell} = \begin{pmatrix} x'(t).dt \\ y'(t).dt \end{pmatrix} = \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} dt$$

Ce vecteur élémentaire $\vec{d\ell}$, tangent à l'arc \widehat{AB} en M , permet alors de définir et calculer des intégrales le long d'une courbe du plan (ou de l'espace).

Note : indépendamment de la paramétrisation que l'on choisit pour décrire l'arc \widehat{AB} sur lequel on souhaite travailler, on notera

$$\vec{d\ell} = \begin{pmatrix} dx \\ dy \end{pmatrix} \quad \text{et} \quad d\ell = \|\vec{d\ell}\|$$

Longueur d'arc

Définition (longueur d'arc)

Soit \widehat{AB} une courbe du plan. La longueur de \widehat{AB} est

$$L(\widehat{AB}) = \int_{\widehat{AB}} \|\vec{d\ell}\| = \int_{\widehat{AB}} d\ell$$

Par ailleurs, si

$$F : [a, b] \longrightarrow \mathbb{R}^2 \\ t \longmapsto (x(t), y(t))$$

est une paramétrisation de l'arc \widehat{AB} , on a

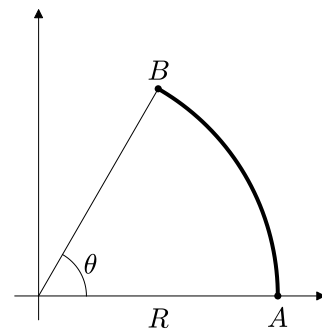
$$L(\widehat{AB}) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2} dt$$

Exemples :

- La notion d'intégrale curviligne permet de calculer la longueur d'un arc de cercle de rayon R et d'angle θ représenté ci-contre.

En effet, une paramétrisation de l'arc \widehat{AB} basée sur les coordonnées polaires est

$$F : [0, \theta_0] \longrightarrow \mathbb{R}^2 \\ t \longmapsto (R \cos(t), R \sin(t))$$



Les fonctions coordonnées de cette paramétrisation sont

$$\begin{cases} x(t) = R \cos(t) \\ y(t) = R \sin(t) \end{cases} \Rightarrow \begin{cases} x'(t) = -R \sin(t) \\ y'(t) = R \cos(t) \end{cases}$$

et

$$d\ell = \sqrt{(-R \cos(t))^2 + (R \sin(t))^2} = R$$

Donc

$$L(\widehat{AB}) = \int_0^{\theta_0} R dt = R\theta_0$$

- La longueur de la courbe d'une fonction f d'une variable réelle au dessus de l'intervalle $[a, b]$ est donnée par

$$L(C_f) = \int_a^b \sqrt{1 + f'(t)^2} dt$$

Intégration d'un champ scalaire le long d'une courbe

De façon générale, le vecteur infinitésimal $\vec{d\ell}$ d'une courbe permet de définir et calculer l'intégrale de n'importe quelle fonction f définie sur un arc \widehat{AB} et à valeurs réelles : si $F : t \mapsto (x(t), y(t))$ est une paramétrisation de l'arc \widehat{AB} définie sur un intervalle $[a, b]$, on a

$$I = \int_{\widehat{AB}} f(M) d\ell = \int_a^b f(x(t), y(t)) \sqrt{x'(t)^2 + y'(t)^2} dt$$

Ainsi,

- Si $f(x, y) = \rho(x, y)$ correspond à une distribution de masse le long de l'arc \widehat{AB} , l'intégrale I correspond à la masse totale M de l'arc.
- Si $f(x, y) = \rho(x, y) \cdot \frac{x}{M}$ ou $f(x, y) = \rho(x, y) \cdot \frac{y}{M}$, l'intégrale I donne les coordonnées cartésiennes du centre d'inertie de l'arc \widehat{AB} .
- Si $f(x, y)$ représente le carré de la distance d'un point $P = (x, y) \in \widehat{AB}$ par rapport à un axe Δ fixé du plan, l'intégrale I donne le moment d'inertie de l'arc par rapport à Δ .

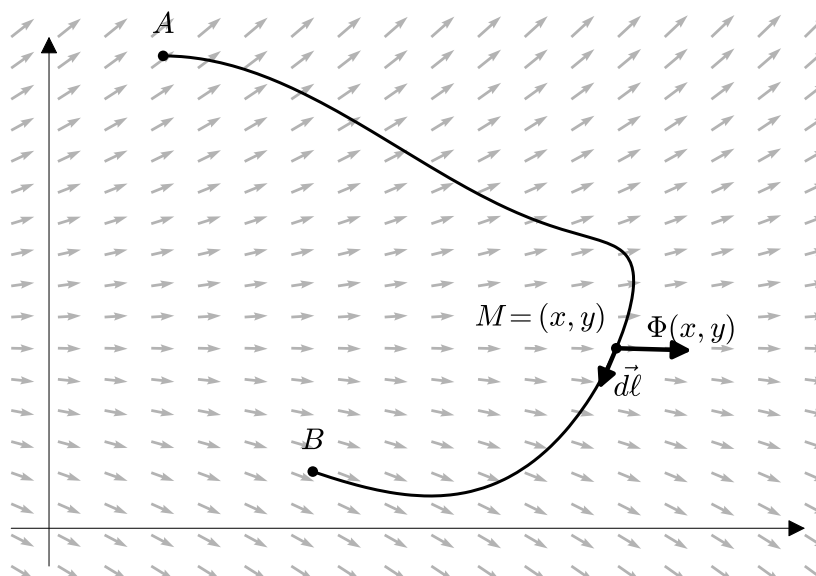
Flux et travail d'un champ de vecteurs

L'expression de l'élément infinitésimal $\vec{d\ell}$ d'une courbe du plan sous la forme d'un vecteur permet également de d'étudier l'interaction d'une courbe du plan donnée avec un champ de vecteurs.

Précisément, soit

$$\begin{aligned} \Phi : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto \Phi(x, y) = (P(x, y), Q(x, y)) \end{aligned}$$

un champ de vecteurs de \mathbb{R}^2 et \widehat{AB} un arc du plan dont on connaît une paramétrisation $F : [a, b] \rightarrow \widehat{AB}$. En chaque point $M = (x(t), y(t)) \in \widehat{AB}$ coexistent un vecteur $\Phi(x(t), y(t))$ et un vecteur élémentaire $\vec{d\ell} = \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} dt$.



Définition (Circulation d'un champ de vecteurs)

— Avec les notations ci-dessus, on appelle *circulation élémentaire* de Φ en un point $M = F(t)$ de l'arc \widehat{AB} le produit scalaire

$$\begin{aligned} dW &= \Phi(M) \cdot \vec{d\ell} = \begin{pmatrix} P(x, y) \\ Q(x, y) \end{pmatrix} \cdot \begin{pmatrix} dx \\ dy \end{pmatrix} \\ &= \begin{pmatrix} P(x(t), y(t)) \\ Q(x(t), y(t)) \end{pmatrix} \cdot \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} dt \\ &= (P(x(t), y(t)) \cdot x'(t) + Q(x(t), y(t)) \cdot y'(t)) dt \end{aligned}$$

— On appelle *circulation* de Φ le long de l'arc \widehat{AB} l'intégrale

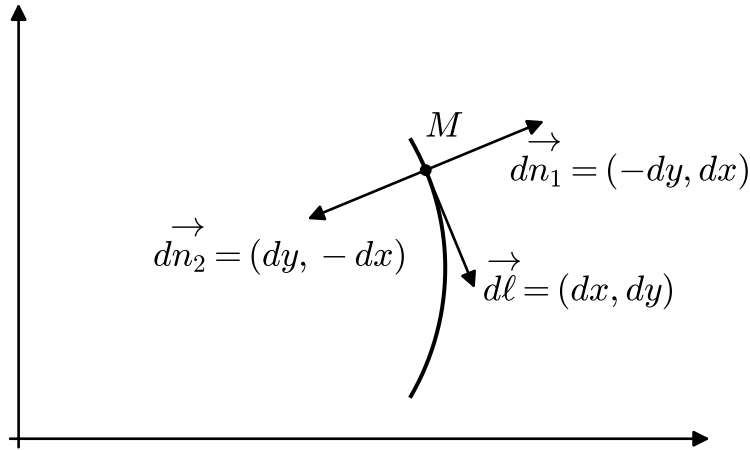
$$W = \int_{\widehat{AB}} dW = \int_a^b (P(x(t), y(t)) \cdot x'(t) + Q(x(t), y(t)) \cdot y'(t)) dt$$

Note : si Φ représente un champ de force agissant sur une particule se déplaçant le long de l'arc \widehat{AB} , la circulation ci-dessus correspond au travail fourni par la force Φ au cours du déplacement, i.e. la différence d'énergie cinétique de la particule entre son point de départ et son point d'arrivée.

D'autre part, à partir du vecteur tangent élémentaire $\vec{d\ell} = \begin{pmatrix} dx \\ dy \end{pmatrix}$, on peut également

définir deux vecteur normaux élémentaires

$$\vec{dn}_1 = \begin{pmatrix} -dy \\ dx \end{pmatrix} \quad \text{ou} \quad \vec{dn}_2 = \begin{pmatrix} dy \\ -dx \end{pmatrix}$$



À partir de ces vecteurs normaux, on peut alors définir la notion de *flux* d'un champ de vecteur à travers une courbe du plan :

Définition (flux d'un champ de vecteurs)

- Avec les notations ci-dessus, on appelle *flux élémentaire* de Φ en un point $M = F(t)$ de l'arc \widehat{AB} le produit scalaire

$$\begin{aligned} d\varphi &= \Phi(M) \cdot \vec{dn} \\ &= \begin{pmatrix} P(x(t), y(t)) \\ Q(x(t), y(t)) \end{pmatrix} \cdot \begin{pmatrix} -y'(t) \\ x'(t) \end{pmatrix} dt \\ &= (Q(x(t), y(t)) \cdot x'(t) - P(x(t), y(t)) \cdot y'(t)) dt \end{aligned}$$

- On appelle *flux* de Φ à travers l'arc \widehat{AB} l'intégrale

$$\varphi = \int_{\widehat{AB}} d\varphi = \int_a^b (Q(x(t), y(t)) \cdot x'(t) - P(x(t), y(t)) \cdot y'(t)) dt$$

Note : dans la définition ci-dessus, le vecteur normal élémentaire \vec{dn} représente le vecteur \vec{dn}_1 ou \vec{dn}_2 selon l'orientation que l'on souhaite étudier. Il est à noter que, puisque $\vec{dn}_2 = -\vec{dn}_1$, les flux associés à chacun de ces vecteurs sont égaux au signe près.

Propriété des intégrales curvilignes

Linéarité et relation de Chasles. Une intégrale curviligne étant évaluée par une intégrale simple, elle hérite directement des propriétés classiques du calcul intégral. Ainsi,

- *Linéarité* : Si Φ et Ψ sont deux champs de vecteurs du plan, pour tout $(\lambda, \mu) \in \mathbb{R}^2$, on a

$$\int_{\widehat{AB}} (\lambda \cdot \Phi(M) + \mu \cdot \Psi(M)) \cdot \vec{d\ell} = \lambda \cdot \int_{\widehat{AB}} \Phi(M) \cdot \vec{d\ell} + \mu \cdot \int_{\widehat{AB}} \Psi(M) \cdot \vec{d\ell}$$

- *Relation de Chasles* : Si C est un point de l'arc \widehat{AB} , plongé dans un champ de vecteurs Φ , on a

$$\int_{\widehat{AB}} \Phi(M) \cdot \vec{d\ell} = \int_{\widehat{AC}} \Phi(M) \cdot \vec{d\ell} + \int_{\widehat{CB}} \Phi(M) \cdot \vec{d\ell}$$

Indépendance à la paramétrisation. Dans toutes les définitions données plus haut, le calcul effectif d'une intégrale curviligne passe par la définition d'une paramétrisation de l'arc sur lequel on souhaite intégrer. C'est cette paramétrisation qui permet en particulier de ramener l'intégrale le long de l'arc à une intégrale "classique" en une variable (i.e. sur un intervalle $[a, b]$ de la droite réelle). Or on peut montrer (notamment à l'aide de la formule du changement de variable dans une intégrale simple) que la valeur de l'intégrale ainsi calculée *ne dépend pas* de la paramétrisation F choisie.

Réversibilité. On peut également montrer que le sens de parcours d'un arc \widehat{AB} influence le signe de toute intégrale curviligne le long de \widehat{AB} :

$$\int_{\widehat{AB}} \Phi(M) \cdot \vec{d\ell} = - \int_{\widehat{BA}} \Phi(M) \cdot \vec{d\ell}$$

Circulation d'un champ de gradient

Nous avons évoqué plus haut le lien existant entre intégrale curviligne et travail (mécanique). Or l'un des résultats fondamentaux de la mécanique du point dit que le travail exercé par un champ gravitationnel sur une masse ponctuelle se déplaçant dans le plan ne dépend pas du chemin emprunté par la-dite masse ponctuelle, mais uniquement du point de départ et du point d'arrivée.

On va voir ici qu'il est possible de démontrer mathématiquement cette propriété et de déterminer les cas (mathématiques) dans lesquels cette propriété est vraie. Précisément, :

Proposition

Si le champ de vecteurs Φ est un champ de gradient du plan (i.e. s'il existe une fonction f de deux variables telle que $\Phi = \nabla f$), alors pour tout arc \widehat{AB} du plan, on a

$$\int_{\widehat{AB}} \Phi(M) \cdot \vec{d\ell} = f(B) - f(A)$$

Preuve : si $\Phi = \nabla f$, on a

$$P(x, y) = \frac{\partial f}{\partial x}(x, y) \quad \text{et} \quad Q(x, y) = \frac{\partial f}{\partial y}(x, y)$$

Mais alors, en introduisant une paramétrisation F de l'arc \widehat{AB} dans l'intégrale curviligne à évaluer, on a

$$\begin{aligned} I &= \int_{\widehat{AB}} \Phi(M) \cdot \vec{d\ell} \\ &= \int_a^b (P(x(t), y(t)) \cdot x'(t) + Q(x(t), y(t)) \cdot y'(t)) dt \\ &= \int_a^b \left(\frac{\partial f}{\partial x}(x(t), y(t)) \cdot x'(t) + \frac{\partial f}{\partial y}(x(t), y(t)) \cdot y'(t) \right) dt \end{aligned}$$

Or on reconnaît, dans la dernière intégrale, la dérivée de la fonction d'une variable

$$t \longmapsto f(x(t), y(t))$$

d'où

$$I = [f(x(t), y(t))]_a^b = f(x(b), y(b)) - f(x(a), y(a)) = f(B) - f(A)$$

On voit en particulier ici que l'intégrale curviligne $\int_{\widehat{AB}} \Phi(M) \cdot \vec{d\ell}$ dépend bien uniquement des points de départ et d'arrivée de l'arc \widehat{AB} .

Formule de Green

La formule de Green-Riemann est une formule permettant de relier une intégrale double sur un domaine D du plan à une intégrale curviligne le long de la frontière ∂D de D :

Théorème (formule de Green)

Soient D un domaine borné du plan et Φ un champ de vecteurs de \mathbb{R}^2 . On a

$$\iint_D \operatorname{div}(\Phi)(M) dS = \int_{\partial D} \Phi(M) \cdot \vec{dn}$$

où \vec{dn} est le vecteur normal élémentaire de la frontière, orienté vers l'extérieur.

Note : cette propriété est à la base de la modélisation de très nombreux phénomènes physiques (diffusion de la chaleur, flux d'un fluide à travers une courbe, etc...). Par ailleurs, cette formule permet également de transformer une intégrale double en une intégrale curviligne et réciproquement. Cela permet alors de calculer l'une ou l'autre des intégrales en se plaçant dans le cas le plus favorable.

Exemple :

1. Déterminer un champ de vecteurs de \mathbb{R}^2 dont la divergence est égale à 1.
2. Retrouver l'aire du disque de rayon R à l'aide d'une paramétrisation de celui-ci et d'une intégrale curviligne.

6.2.3 Intégrales de surfaces

Vecteur normal et élément de surface

À l'image de la notion d'intégrale curviligne, il est possible d'étendre la notion d'intégrale aux *surfaces de l'espace*. Pour cela, il est nécessaire de définir correctement la notion d'élément infinitésimal de surface. Cela passe en particulier par la notion de *vecteur normal à une surface* et par la notion de paramétrisation de surface.

Ainsi, soit \mathcal{S} une surface de l'espace et $G : (u, v) \mapsto G(u, v) = (x(u, v), y(u, v), z(u, v))$ une paramétrisation de \mathcal{S} définie sur une partie $D \subset \mathbb{R}^2$. Si les fonctions x , y et z admettent des dérivées partielles, on peut, en un point $M_0 = G(u_0, v_0)$ de \mathcal{S} , définir les vecteurs

$$\frac{\partial G}{\partial u}(u_0, v_0) = \begin{pmatrix} \frac{\partial x}{\partial u}(u_0, v_0) \\ \frac{\partial y}{\partial u}(u_0, v_0) \\ \frac{\partial z}{\partial u}(u_0, v_0) \end{pmatrix} \quad \text{et} \quad \frac{\partial G}{\partial v}(u_0, v_0) = \begin{pmatrix} \frac{\partial x}{\partial v}(u_0, v_0) \\ \frac{\partial y}{\partial v}(u_0, v_0) \\ \frac{\partial z}{\partial v}(u_0, v_0) \end{pmatrix}$$

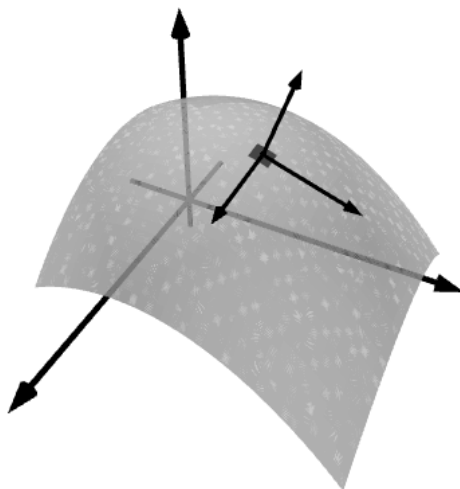
Si ces vecteurs sont linéairement indépendants, ils engendrent un plan de l'espace : le *plan tangent* à \mathcal{S} au point M_0 .

On peut alors, à partir de ces deux vecteurs $\frac{\partial G}{\partial u}(u_0, v_0)$ et $\frac{\partial G}{\partial v}(u_0, v_0)$ construire un *vecteur normal* à \mathcal{S} en M_0 :

$$\vec{N}(u_0, v_0) = \frac{\partial G}{\partial u}(u_0, v_0) \wedge \frac{\partial G}{\partial v}(u_0, v_0)$$

On peut alors définir la mesure d'un élément infinitésimal de surface $d\sigma$ au point M_0 par :

$$d\sigma(M_0) = \|\vec{N}(u_0, v_0)\| dudv = \left\| \begin{pmatrix} \frac{\partial x}{\partial u}(u_0, v_0) \\ \frac{\partial y}{\partial u}(u_0, v_0) \\ \frac{\partial z}{\partial u}(u_0, v_0) \end{pmatrix} \wedge \begin{pmatrix} \frac{\partial x}{\partial v}(u_0, v_0) \\ \frac{\partial y}{\partial v}(u_0, v_0) \\ \frac{\partial z}{\partial v}(u_0, v_0) \end{pmatrix} \right\| dudv$$



L'aire de la surface \mathcal{S} est alors donnée par

$$\mathcal{A}(\mathcal{S}) = \iint_{\mathcal{S}} d\sigma(M) = \iint_D \|\vec{N}(u, v)\| dudv$$

Note : comme pour les intégrales curviligne, il est possible d'ajouter ici une fonction f des variables (x, y, z) (représentant par exemple une distribution de masses sur la surface \mathcal{S}), que l'on intègre sur l'ensemble de la surface, en évaluant f aux points $(x(u, v), y(u, v), z(u, v))$.

Exercice : exprimer, à l'aide d'une intégrale de surface les quantités suivantes :

1. Les coordonnées cartésiennes (x_G, y_G, z_G) du centre d'inertie G de \mathcal{S} .
2. Le moment d'inertie I_{Δ} de \mathcal{S} par rapport à un axe Δ de l'espace.

Flux d'un champ de vecteurs de l'espace

À l'aide de ces notions, il est également possible de définir le flux d'un champ de vecteur de \mathbb{R}^3 à travers une surface de l'espace. On doit pour cela construire un vecteur normal *unitaire* en chaque point de la surface \mathcal{S} :

$$\vec{n}(u_0, v_0) = \frac{\vec{N}(u_0, v_0)}{\|\vec{N}(u_0, v_0)\|}$$

Alors, si $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ est un champ de vecteurs de \mathbb{R}^3 , on appelle *flux élémentaire de* Φ en un point M de \mathcal{S} le produit scalaire

$$d\varphi = \Phi(M) \cdot \vec{n} d\sigma(M)$$

Le flux total du champ de vecteurs Φ à travers \mathcal{S} est alors

$$\varphi = \iint_{\mathcal{S}} d\varphi = \iint_{\mathcal{S}} \Phi(M) \cdot \vec{n} d\sigma$$

Si

$$\begin{aligned} G : D \subset \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (u, v) &\longmapsto G(u, v) \end{aligned}$$

est alors une paramétrisation de \mathcal{S} , on peut alors exprimer cette intégrale de surface sous la forme d'une intégrale double sur le domaine D :

$$\varphi = \iint_D \Phi(G(u, v)) \cdot \vec{n}(u, v) \|\vec{N}(u, v)\| dudv = \iint_D \Phi(G(u, v)) \cdot \vec{N}(u, v) dudv$$

Théorème de Green-Ostrogradski

À l'image du théorème de Green dans le plan, on peut énoncer un théorème analogue pour les surfaces fermées, i.e. les surfaces délimitant une partie fermée de l'espace. Précisément, soit D un domaine de l'espace et ∂D la surface constituant la frontière de D . Pour tout champ de vecteur Φ de l'espace, on a

$$\iiint_D \operatorname{div}(\Phi)(M) dV = \iint_{\partial D} \Phi(M) \cdot \vec{n} d\sigma$$

Théorème de Stokes

Il est également possible d'énoncer un théorème portant sur le rotationnel d'un champ de vecteurs dans l'espace. Précisément, soit \mathcal{S} une surface de l'espace à deux faces, délimitée par une courbe fermée \mathcal{C} . En notant $\vec{d\ell}$ un vecteur élémentaire de \mathcal{C} orientant \mathcal{C} de sorte que \mathcal{S} soit à gauche. Alors

$$\iint_{\mathcal{S}} \operatorname{Rot}(\Phi)(M) \cdot \vec{n} d\sigma = \int_{\mathcal{C}} \Phi(M) \cdot \vec{d\ell}$$