Chapitre A-VIII

Outils mathématiques courants.

Joël SORNETTE met ce cours à votre disposition selon les termes de la licence Creative Commons :

- Pas d'utilisation commerciale.
- Pas de modification, pas de coupure, pas d'intégration à un autre travail.
- Pas de communication à autrui sans citer son nom, ni en suggérant son autorisation.

Retrouvez l'intégralité du cours sur le site joelsornette.fr

RÉSUMÉ :

Comment résumer en quelques dizaines de pages deux à trois années de cours de mathématiques, soit quelques centaines d'heures et quelques milliers de pages de notes ? En ne choisissant que ce qui sert régulièrement en physique (ce qui sert ponctuellement sera traité dans le chapitre concerné) et en ne donnant aucune démonstration, à la rigueur un fil conducteur montrant l'enchaînement des idées. Cela suppose que le lecteur ait déjà vu toutes ces choses et que ce chapitre n'est là que pour lui rafraîchir la mémoire sur un point qui s'estompe.

Table des matières

1 Analyse.

1.a Ce qui ne sera pas traité.

Nous supposerons connue la définition d'une suite numérique de terme général noté u_n qui à tout entier n positif (ou nul éventuellement) associe un nombre réel ou complexe et supposerons maîtrisée la notion de limite d'une suite quand l'indice n tend vers l'infini.

Nous supposerons connue la définition d'une série numérique, réelle ou complexe, de terme général noté u_n qui à tout entier n positif (ou nul éventuellement) associe la somme partielle $S_n = \sum_1^\infty u_n$ (ou $\sum_0^\infty u_n$) et supposerons maîtrisée la notion de limite d'une série quand l'indice n tend vers l'infini.

Nous supposerons connue la notion de fonction à valeurs réelles ou complexes d'une variable réelle ou complexe (la fonction f associe pour certaines valeurs de x réel ou complexe qui constituent le domaine de définition, un réel ou complexe $f(x)$) et supposerons maîtrisées les notions de continuité et de dérivabilité (uniquement dans le cas de variable réelles [1]) ainsi que les règles classiques de dérivation. Nous croyons inutile de nous attarder aux fonctions polynomiales ou fractions rationnelles.

1.b Séries entières.

A une suite numérique de terme général u_n (n positif ou nul), on peut associer la série de terme général $u_n\, x^n$ (ou $u_n\, z^n$ avec z complexe) et si elle existe sa limite quand n tend vers l'infini notée $\sum_0^\infty u_n\, x^n$ qui dépend bien évidemment de x. C'est une façon très courante de définir une fonction d'une variable réelle (ou complexe).

Les mathématiciens démontrent que toute série entière a un *rayon de convergence* noté R parfois infini, tel que pour tout x tel que $|x| < R$ (ou $|z| < R$ pour un complexe), la série converge [2] et diverge pour $r > R$ ou $|z| > R$, que, dans le cas d'une variable réelle la fonction ainsi obtenue est dérivable, que sa dérivée est la somme de la série entière de terme général $n\, u_n\, x^{n-1}$ (dérivation *terme à terme*) qui a le même rayon de convergence et l'on en déduit aisément que la fonction est indéfiniment dérivable.

1.c Fonction exponentielle.

La plus connue des fonctions ainsi définies est la fonction exponentielle, de rayon de convergence infini (définie donc pour tout x) dont la définition est :

$$\exp(x) = \sum_0^\infty \frac{x^n}{n!}$$

1. La dérivabilité de fonctions d'une variable complexe dépasse le cadre de ce cours.
2. la convergence pour $|x| = R$ ou $|z| = R$ se traite au cas par cas.

et aussi pour z complexe :

$$\exp(z) = \sum_{0}^{\infty} \frac{z^n}{n!}$$

Par dérivation terme à terme suivie d'un subtil décalage d'indice, on prouve que l'exponentielle est sa propre dérivée soit :

$$\exp'(x) = \exp(x)$$

Les mathématiciens démontre enfin cette *propriété fonctionnelle* : pour tout couple de réels ou de complexes a et b, on a :

$$\exp(a+b) = \exp(a)\,\exp(b)$$

avec les corollaires suivants :

$\exp(n\,a) = \exp(a)^n$ pour n entier positif ou nul,

$\exp(-a) = \frac{1}{\exp(a)}$ d'où $\exp(n\,a) = \exp(a)^n$ pour n entier positif, négatif ou nul,

$\exp\left(\sqrt[p]{x}\right) = \frac{\exp(x)}{p}$ (avec x positif) d'où $\exp(n\,a) = \exp(a)^n$ pour n fractionnaire (et même réel par continuité).

1.d Fonctions trigonométriques.

• Les fonctions cosinus et sinus.

A partir de la fonction exponentielle complexe, on définit les fonctions cosinus et sinus d'un réel x comme partie réelle et imaginaire de $\exp(i\,x)$ soit en pratique :

$$\begin{cases} \cos(x) = \displaystyle\sum_{0}^{\infty} (-1)^p \frac{x^{2\,p}}{(2\,p)!} \\ \sin(x) = \displaystyle\sum_{0}^{\infty} (-1)^p \frac{x^{2\,p+1}}{(2\,p+1)!} \end{cases}$$

On vérifie aisément que la fonction cosinus est paire et la fonction sinus impaire et que leurs dérivées sont $\cos'(x) = -\sin(x)$ et $\sin'(x) = \cos(x)$ d'où l'on tire l'identité suivante : $\cos^2(x) + \sin^2(x) = 1$ qui prouve à l'évidence que $\cos(x)$ et $\sin(x)$ ont des valeurs comprises entre -1 et $+1$

Les mathématiciens déduisent des propriétés de l'exponentielle les séries de formules suivantes qu'il est très rentable d'apprendre par cœur car elles sont d'un usage courant

dans tous les phénomènes vibratoires et ondulatoires :

$$\begin{cases} \cos(a+b) = \cos(a)\ \cos(b) - \sin(a)\ \sin(b) \\ \cos(a-b) = \cos(a)\ \cos(b) + \sin(a)\ \sin(b) \\ \sin(a+b) = \sin(a)\ \cos(b) + \cos(a)\ \sin(b) \\ \sin(a-b) = \sin(a)\ \cos(b) - \cos(a)\ \sin(b) \end{cases}$$

On en déduit dans le cas particulier $b = a$

$$\begin{cases} \cos(2\,a) = \cos^2(a) - \sin^2(a) = 2\ \cos^2(a) - 1 = 1 - 2\ \sin^2(a) \\ \sin(2\,a) = 2\ \sin(a)\ \cos(a) \end{cases}$$

En revenant au cas général, on arrive par combinaison linéaire puis permutation des deux membres à ces formules qui transforment les produits en somme (linéarisation) :

$$\begin{cases} \cos(a)\ \cos(b) = \dfrac{\cos(a-b) + \cos(a+b)}{2} \\ \\ \sin(a)\ \sin(b) = \dfrac{\cos(a-b) - \cos(a+b)}{2} \quad \text{(formule piégeante)} \\ \\ \sin(a)\ \cos(b) = \dfrac{\sin(a-b) + \sin(a+b)}{2} \end{cases}$$

On en déduit dans le cas particulier $b = a$

$$\begin{cases} \cos^2(a) = \dfrac{1 + \cos(2\,a)}{2} \\ \\ \sin^2(a) = \dfrac{1 - \cos(2\,a)}{2} \\ \\ \sin(a)\ \cos(a) = \dfrac{\sin(2\,a)}{2} \end{cases}$$

En revenant au cas général et en posant $p = a + b$ et $q = a - b$ et permutation des deux membres, on arrive à ces formules qui transforment les sommes en produits (factorisation) :

$$\begin{cases} \cos(p) + \cos(q) = 2\ \cos\left(\dfrac{p+q}{2}\right)\ \cos\left(\dfrac{p-q}{2}\right) \\ \cos(p) - \cos(q) = -2\ \sin\left(\dfrac{p+q}{2}\right)\ \sin\left(\dfrac{p-q}{2}\right) \quad \text{(formule piégeante)} \\ \sin(p) + \sin(q) = 2\ \sin\left(\dfrac{p+q}{2}\right)\ \cos\left(\dfrac{p-q}{2}\right) \\ \sin(p) - \sin(q) = 2\ \sin\left(\dfrac{p-q}{2}\right)\ \cos\left(\dfrac{p+q}{2}\right) \end{cases}$$

Nous n'aborderons pas ici le développement de $\cos(3\,a)$, $\sin(3\,a)$, etc. ni la linéarisation d'expression du type $\cos^4(x)$ car c'est peu fréquent en physique et que l'on peut toujours trouver aisément un formulaire ou, si l'on s'est isolé sur une île déserte sans wi-fi, travailler par étapes successives.

Si l'on définit le nombre π comme la plus petite valeur positive non nulle dont le sinus soit nul (le plus délicat est de montrer qu'il en existe une) donc le cosinus égal à -1 (le plus délicat est de montrer que ce n'est pas $+1$), le développement de $\sin(a+b)$ avec $b=\pi$ donne $\sin(a+\pi) = -\sin(a)$ et donc $\sin(a+2\,\pi) = \sin(a)$ ce qui prouve que la fonction est $2\,\pi$-périodique. A partir de là, on démontre assez aisément :

$$\begin{cases} \cos(x+2\,\pi) = \cos(x) & \text{et} \quad \sin(x+2\,\pi) = \sin(x) \\ \cos(x+\pi) = -\cos(x) & \text{et} \quad \sin(x+\pi) = -\sin(x) \\ \cos\left(x+\dfrac{\pi}{2}\right) = -\sin(x) & \text{et} \quad \sin\left(x+\dfrac{\pi}{2}\right) = \cos(x) \end{cases}$$

• Les fonctions tangente et cotangente.

On définit la fonction tangente par $\tan(x) = \frac{\sin(x)}{\cos(x)}$ (non définie pour $x = (2\,k+1)\frac{\pi}{2}$ avec k entier) et la fonction cotangente par $\mathrm{cotan}(x) = \frac{1}{\tan(x)} = \frac{\cos(x)}{\sin(x)}$ (non définie pour $x = k\,\pi$ avec k entier). On déduit aisément des propriétés des fonctions cosinus et sinus que :

$$\frac{1}{\cos^2(x)} = 1 + \tan^2(x) \qquad \text{et} \qquad \frac{1}{\sin^2(x)} = 1 + \mathrm{cotan}^2(x)$$

On montre aisément que :

$$\tan'(x) = \frac{1}{\cos^2(x)} = 1 + \tan^2(x)$$

$$\mathrm{cotan}'(x) = -\frac{1}{\sin^2(x)} = -[1 + \mathrm{cotan}^2(x)]$$

Ces fonctions sont π-périodique et que $\mathrm{cotan}(x) = -\tan\left(x+\frac{\pi}{2}\right)$ (en pratique on utilise plutôt $\mathrm{cotan}(x) = \frac{1}{\tan(x)}$ et on n'utilise donc guère cette fonction).

Quelques formules utiles :

$$\begin{cases} \tan(a+b) = \dfrac{\tan(a)+\tan(b)}{1-\tan(a)\,\tan(b)} \\ \tan(a-b) = \dfrac{\tan(a)-\tan(b)}{1+\tan(a)\,\tan(b)} \\ \tan(2\,a) = \dfrac{2\,\tan(a)}{1-\tan^2(a)} \end{cases}$$

En posant $t = \tan\left(\frac{x}{2}\right)$, on a :

$$\tan(x) = \frac{2\,t}{1-t^2} \qquad \text{et} \qquad \sin(x) = \frac{2\,t}{1+t^2} \qquad \text{et} \qquad \cos(x) = \frac{1-t^2}{1+t^2}$$

ce qui sera fort utile pour les calculs d'intégrales.

1.e Fonctions hyperboliques.

A partir de la fonction exponentielle, on définit les fonctions cosinus et sinus d'un réel x comme partie paire et impaire de $\exp(x)$ soit en pratique :

$$\begin{cases} \mathrm{ch}(x) = \displaystyle\sum_0^\infty \frac{x^{2\,p}}{(2\,p)!} \\ \mathrm{sh}(x) = \displaystyle\sum_0^\infty \frac{x^{2\,p+1}}{(2\,p+1)!} \end{cases}$$

Leurs propriétés sont proches de celles des fonctions trigonométriques mais avec des signes parfois différents et apprendre par cœur les deux séries de formules est le plus sûr moyen de « s'emmêler les pinceaux ». Je suggère de savoir par cœur que $\mathrm{ch}'(x) = \mathrm{sh}(x)$, $\mathrm{sh}'(x) = \mathrm{ch}(x)$, $\mathrm{ch}^2(x) - \mathrm{sh}^2(x) = 1$, que ces fonctions sont non périodiques, que $\mathrm{ch}(x) \geqslant 1$. Pour toutes les autres formules, on les déduira formellement des formules de trigonométrie en remplaçant cos par ch et sin par i sh et en simplifiant le cas échéant par i.

De là on déduit les formules relatives à la tangente hyperbolique $\mathrm{th}(x) = \frac{\mathrm{sh}(x)}{\mathrm{ch}(x)}$ et de son inverse la cotangente hyperbolique. Retenons que $\frac{1}{\mathrm{ch}^2(x)} = 1 - \mathrm{th}^2(x)$ et que la dérivée de la tangente hyperbolique est $\mathrm{th}'(x) = \frac{1}{\mathrm{ch}^2(x)} = 1 - \mathrm{th}^2(x)$

1.f Fonctions réciproques.

• La théorie et l'exemple de la fonction logarithme népérien.

Si une fonction f continue et monotone sur un intervalle I (éventuellement $\mathbb{R}$) et si l'on note J (éventuellement $\mathbb{R}$) l'ensemble des valeurs qu'elle prend, on définit, sur J, la fonction réciproque f^{-1} par :

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

De plus, si f est dérivable, f^{-1} l'est aussi (sauf l'exception mentionnée ci-dessous) et l'on a :

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

sauf pour les valeurs de x telles que le dénominateur est nul, bien entendu.

Pour l'exploitation de cette formule, il est commode de poser provisoirement $y = f^{-1}(x)$ puis de chercher à exprimer $f'(y)$ en fonction de x (c'est là que réside parfois la difficulté).

Dans la pratique, si f est une fonction classique, on donne à f^{-1} un nom particulier.

Prenons pour exemple la fonction exponentielle, monotone croissante de $\exp(-\infty) = 0$ à $\exp(+\infty) = +\infty$, sa fonction inverse est la fonction logarithme népérien notée ln définie de 0 à $+\infty$, croissante de $\ln(0) = -\infty$ à $\ln(+\infty) = +\infty$. La formule $f^{-1\prime}(x) = \frac{1}{f'(y)}$, avec $y = f^{-1}(x)$ équivalent à $x = f(y)$, devient ici puisque $f(y) = \exp(y)$ et donc d'une part $f'(y) = \exp'(y) = \exp(y)$ (cf supra) et d'autre part $y = f^{-1}(x) = \ln(x)$ équivalent à $x = \exp(y)$:

$$\ln'(x) = \frac{1}{\exp(y)} = \frac{1}{x}$$

Remarque : pour x négatif, on vérifie aisément (dérivée de fonctions composées) qu'une primitive de $\frac{1}{x}$ est $\ln(-x)$; donc si l'on ne connaît pas le signe de x, une primitive est $\ln(|x|)$.

• Fonction trigonométriques inverses.

La fonction cosinus décroît de 1 à -1 quand sa variable croît de 0 à π ; on définit la fonction *arc-cosinus* comme fonction inverse sur ces intervalles. La fonction $\arccos(x)$ décroît donc de π à 0 quand x croît de -1 à 1. En accélérant un peu par rapport à l'exemple précédent, sa dérivée est :

$$\arccos'(x) = -\frac{1}{\sin(y)} = -\frac{1}{\sqrt{1-\cos^2(y)}} = -\frac{1}{\sqrt{1-x^2}}$$

La fonction sinus croît de -1 à 1 quand sa variable croît de $-\frac{\pi}{2}$ à $\frac{\pi}{2}$; on définit la fonction *arc-sinus* comme fonction inverse sur ces intervalles. La fonction $\arcsin(x)$ croît donc de $-\frac{\pi}{2}$ à $\frac{\pi}{2}$ quand x croît de -1 à 1. Sa dérivée est :

$$\arcsin'(x) = \frac{1}{\cos(y)} = \frac{1}{\sqrt{1-\sin^2(y)}} = \frac{1}{\sqrt{1-x^2}}$$

La fonction tangente croît de $-\infty$ à ∞ quand sa variable croît de $-\frac{\pi}{2}$ à $\frac{\pi}{2}$; on définit la fonction *arc-tangente* comme fonction inverse sur ces intervalles. La fonction $\arctan(x)$ croît donc de $-\frac{\pi}{2}$ à $\frac{\pi}{2}$ quand x croît de $-\infty$ à ∞. Sa dérivée est :

$$\arctan'(x) = \frac{1}{1+\tan^2(y)} = \frac{1}{1+x^2}$$

La fonction *arc-cotangente* est peu employée.

• **Fonction hyperboliques inverses.**

La fonction cosinus hyperbolique croît de 1 à ∞ quand sa variable croît de 0 à ∞ ; on définit la fonction *argument-cosinus hyperbolique* comme fonction inverse sur ces intervalles. La fonction $\mathrm{argch}(x)$ croît donc de 0 à ∞ quand x croît de 1 à ∞. Sa dérivée est :

$$\mathrm{argch}'(x) = \frac{1}{\mathrm{sh}(y)} = \frac{1}{\sqrt{\mathrm{ch}^2(y) - 1}} = \frac{1}{\sqrt{x^2 - 1}}$$

La fonction sinus hyperbolique croît de $-\infty$ à ∞ quand sa variable croît de $-\infty$ à ∞ ; on définit la fonction *argument-sinus hyperbolique* comme fonction inverse sur ces intervalles. La fonction $\mathrm{argsh}(x)$ croît donc de $-\infty$ à ∞ quand x croît de $-\infty$ à ∞. Sa dérivée est :

$$\mathrm{argsh}'(x) = \frac{1}{\mathrm{ch}(y)} = \frac{1}{\sqrt{1 + \mathrm{sh}^2(y)}} = \frac{1}{\sqrt{x^2 + 1}}$$

La fonction tangente hyperbolique croît de -1 à 1 quand sa variable croît de $-\infty$ à ∞ ; on définit la fonction *argument-tangente hyperbolique* comme fonction inverse sur ces intervalles. La fonction $\mathrm{argth}(x)$ croît donc de $-\infty$ à ∞ quand x croît de -1 à 1. Sa dérivée est :

$$\mathrm{argth}'(x) = \frac{1}{1 - \mathrm{th}^2(y)} = \frac{1}{1 - x^2}$$

Notons cette approche variante : par construction on a $\exp(y) = \mathrm{ch}(y) + \mathrm{sh}(y)$ d'où l'on tire successivement en introduisant les fonctions inverses

$$\exp(y) = \mathrm{ch}(y) + \sqrt{\mathrm{ch}^2(y) - 1}$$

$$y = \ln[\mathrm{ch}(y) + \sqrt{\mathrm{ch}^2(y) - 1}]$$

$$\mathrm{argch}(x) = \ln[x + \sqrt{x^2 - 1}]$$

et aussi

$$\exp(y) = \mathrm{sh}(y) + \sqrt{\mathrm{sh}^2(y) + 1}$$

$$y = \ln[\mathrm{sh}(y) + \sqrt{\mathrm{sh}^2(y) + 1}]$$

$$\mathrm{argsh}(x) = \ln[x + \sqrt{x^2 + 1}]$$

On montre aussi aisément que :

$$\mathrm{th}(y) = \frac{\mathrm{sh}(y)}{\mathrm{ch}(y)} = \frac{\exp(y) - \exp(-y)}{\exp(y) - \exp(-y)} = \frac{\exp(2\,y) - 1}{\exp(2\,y) - 1}$$

d'où successivement et en utilisant la fonction inverse :

$$\exp(2\,y) = \frac{1+\mathrm{th}(y)}{1-\mathrm{th}(y)}$$

$$y = \frac{1}{2}\,\ln\left[\frac{1+\mathrm{th}(y)}{1-\mathrm{th}(y)}\right] = \ln\left[\sqrt{\frac{1+\mathrm{th}(y)}{1-\mathrm{th}(y)}}\right]$$

$$\mathrm{argth}(x) = \ln\left[\sqrt{\frac{1+x}{1-x}}\right]$$

1.g Dérivées successives. Développement de Taylor.

Soit une fonction de la variable x notée $f(x)$ ou $f^{(0)}(x)$, supposée dérivable et de dérivée notée $f'(x)$ ou $f^{(1)}(x)$. Si cette dérivée est elle-même dérivable, sa dérivée est la *dérivée seconde* de f et on la note $f''(x)$ ou $f^{(2)}(x)$. Si l'on poursuit la démarche tant que l'on peut dériver, on introduit des dérivées troisième , quatrième,..., énième, la seule notation est alors $f^{(3)}(x)$, $f^{(4)}(x)$, ...,$f^{(n)}(x)$

Pour une série entière $f(x) = f^{(0)}(x) = \sum_0^\infty u_k\,x^k$, on a $f^{(1)}(x) = \sum_1^\infty k\,u_k\,x^{k-1}$,..., $f^{(n)}(x) = \sum_n^\infty k\,(k-1)\cdots(k-n+1)\,u_k\,x^{k-n}$ et en particulier $f^{(0)}(0) = u_0$, $f^{(1)}(0) = u_1$,...,$f^{(n)}(0) = n!\,u_n$, ce qui permet d'écrire :

$$f(x) = \sum_0^\infty \frac{f^{(k)}(0)}{k!}\,x^k$$

que l'on appelle développement de Taylor.

On montre qu'il en est de même pour toute fonction indéfiniment dérivable et un changement de variable en $x = x_0 + \xi$ conduit aisément à la généralisation, appelée développement en x_0 :

$$f(x) = \sum_0^\infty \frac{f^{(k)}(x_0)}{k!}\,(x-x_0)^k$$

En physique on utilise fréquemment des formules tronquées comme approximations d'autant meilleures qu'il y a plus de termes, on les appelle *développements limités* et l'on note par exemple pour un développement à l'ordre 2 :

$$f(x) \approx f(0) + f'(0)\,x + \frac{f''(0)}{2}\,x^2 + \cdots$$

Les mathématiciens cherchent à donner une forme au terme négligé ; nous pouvons nous en dispenser au niveau de ce cours.

Remarque : Le premier terme à coefficient non nul d'un développement de Taylor en x_0 d'une fonction est appelée (on simplifie la théorie) équivalent de cette fonction en x_0.

1.h Fonctions composées.

Soient f et g deux fonctions, on appelle fonction composée de g par f la fonction notée (en mathématiques) $f \circ g$ définie par $[f \circ g](x) = f[g(x)]$ et l'on montre que :

$$[f \circ g]'(x) = f'[g(x)]\, g'(x)$$

1.i Primitives et intégrales définies.

Imaginons le problème physique suivant : on dépose une charge Q uniformément répartie sur un segment de droite de longueur L et l'on veut calculer le potentiel électrique créé en un point particulier M en dehors de ce segment à partir de la définition première du potentiel : une charge q en un point P crée en un point M le potentiel $\frac{q}{4\,\pi\,\varepsilon_0\,\|\overrightarrow{PM}\|}$

Comment passer d'une formule valable pour un point à une formule pour une répartition continue ? On commence par une approximation : on découpe le segment en N segments élémentaires égaux de longueur $\ell = \frac{L}{N}$ et portant la charge $q = \frac{Q}{N}$ ou encore $q = \lambda\,\ell$ avec $\lambda = \frac{Q}{L}$. Ces segments ne sont certes toujours pas des points mais on les assimile quand même à une charge ponctuelle q placée au point milieu P_k de chaque segment élémentaire et l'on propose l'approximation :

$$V(M) \approx \sum_1^N \frac{q}{4\,\pi\,\varepsilon_0\,\|\overrightarrow{P_kM}\|}$$

En notant a et $b = a + L$ les abscisses des extrémités du segment, le segment de rang i est entre les abscisses $a_{k-1} = a + (k-1)\,\frac{L}{N}$ et $a_k = a + k\,\frac{L}{N}$ (avec $a_0 = 0$ et $a_N = b$) et les points P_k ont pour abscisses $x_k = \frac{a_{k-1}+a_k}{2}$ avec $a_{k-1} < x_k < a_k$ et l'on peut alors noter :

$$V(M) \approx \sum_1^N (a_k - a_{k-1})\, f(x_k)$$

avec $f(x) = \frac{\lambda}{4\,\pi\,\varepsilon_0\,\|\overrightarrow{P_xM}\|}$ où P_x est le point d'abscisse x. En fait l'approximation reste valable même avec des segments élémentaires inégaux pourvu que le plus grand d'entre eux soit petit et même si dans chaque segment on choisit un autre point que le point milieu.

Une approximation est donc $V(M) \approx \sum_1^N (a_k - a_{k-1})\, f(x_k)$ avec des a_k choisis tels que $a_0 = a < a_1 < a_2 < \cdots < a_{N-2} < a_{N-1} < a_N = b$, des x_k tels que $a_{k-1} \leqslant x_k \leqslant a_k$ pour tous les k entre 1 et N et avec $\sup_1^N a_k - a_{k-1}$ assez petit.

L'approximation est d'autant meilleure que $\sup_1^N a_k - a_{k-1}$ est plus petit et s'il tend vers 0 (ce qui suppose que N tende vers l'infini), la limite est le potentiel recherché... à condition que la limite soit indépendante du choix du découpage. Une condition suffisante (mais non nécessaire) pour que ce soit le cas est que la fonction f soit définie et continue, ce qui est souvent le cas en physique classique. Il est d'usage de noter $\mathrm{d}x$ les $a_k - a_{k-1}$ et

de remplacer le sigma majuscule de sommation par un S minuscule allongé comme il l'était à la Renaissance. On note donc cette limite :

$$V(M) = \int_a^b f(x)\,\mathrm{d}x$$

où le nom de la variable n'a aucune incidence sur le résultat $\int_a^b f(x)\,\mathrm{d}x$ et $\int_a^b f(y)\,\mathrm{d}y$, c'est la même chose. On l'appelle *intégrale définie* de f sur le segment $[a, b]$

Reste à calculer cette limite. On se convainc assez aisément que pour des découpages identiques sur la plage commune aux segments $[a, x]$ et $[a, x+\varepsilon]$ et un seul segment élémentaire ailleurs avec le choix du point au début du segment, $\int_a^{x+\varepsilon} f(\xi)\,\mathrm{d}\xi$ et $\int_a^x f(\xi)\,\mathrm{d}\xi$ diffèrent de $\varepsilon\, f(x)$. De là on démontre que la dérivée de $F(x) = \int_a^x f(\xi)\,\mathrm{d}\xi$ est $f(x)$. La fonction F de dérivée f est appelée *primitive* de f ; elle est définie à une constante additive près. Si F est l'une des primitives de f, on a alors (c'est assez évident) :

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

On note souvent $F(b) - F(a) = [F(x)]_a^b$

Remarque 1 : au vu de la façon dont on a introduit les choses, on a :

$$\int_b^a f(x)\,\mathrm{d}x = -\int_a^b f(x)\,\mathrm{d}x$$

Remarque 2 : il est facile de sentir, plus délicat de démontrer proprement que :

$$\int_a^c f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x$$

mais il faut bien laisser du travail aux mathématiciens.

Remarque 3 : Si a ou b est infini, il suffit de connaître $F(\infty)$ ou $F(-\infty)$ selon le cas. De même si F est discontinue en a ou en b, il suffit de connaître la limite à droite de F en a ou à gauche en b et si la discontinuité est en c entre a et b, on utilise au préalable la remarque 2.

Reste à apprendre à calculer les primitives.

1.j Calcul des primitives et des intégrales définies.

Il est toujours délicat de trouver la primitive d'une fonction qui sort un peu de l'ordinaire. Les logiciels de calcul formel sont d'une aide appréciable mais passent parfois à côté d'une primitive qu'ils ne « voient » pas. Un ami mathématicien est une valeur sûre. Nous nous contenterons ici de quelques techniques classiques.

• **Reconnaître la dérivée d'une fonction connue.**

Si dans f, on reconnaît, à une constante multiplicative près, la dérivée classique d'une autre fonction c'est gagné. La primitive de x^n est $\frac{x^{n+1}}{n+1}$ pour n positif, négatif ou nul, entier, fractionnaire et même réel, sauf toutefois pour $n = -1$ mais alors on sait que $\frac{1}{x}$ a pour primitive $\ln x$.

On aura en tête les dérivées des fonctions trigonométriques et hyperboliques inverses qui sont d'un précieux secours.

• **Changement de variable.**

Si l'on arrive à mettre $f(x)$ sous la forme $f(x) = g'[h(x)]\, h'(x)$ alors une primitive est $F(x) = g[h(x)]$ (cf supra). Dans la pratique, on matérialise la démarche par un changement de variable. Prenons un exemple, soit à calculer :

$$I = \int_0^{\frac{\pi}{2}} \frac{\cos\varphi}{\sqrt{1+\sin^2\varphi}}\,\mathrm{d}\varphi$$

On a bien une forme $f(\varphi) = g'[h(\varphi)]\, h'(\varphi)$ avec $h'(\varphi) = \cos\varphi$ d'où $h(\varphi) = \sin\varphi$ et g défini par $g'(t) = \frac{1}{\sqrt{1+t^2}}$ dont il faudra trouver la primitive qui n'est autre (cf supra) que l'argument-sinus hyperbolique. Dans la pratique, on présente ainsi : on pose $t = \sin\varphi$ d'où [3] $\mathrm{d}t = \cos\varphi\,\mathrm{d}\varphi$ et on change les bornes en remplaçant les bornes relatives à φ par les valeurs correspondantes de t, soit dans notre exemple :

$$I = \int_0^1 \frac{1}{\sqrt{1+t^2}}\,\mathrm{d}t = [\mathrm{argsh}(t)]_0^1 = \mathrm{argsh}(1)$$

Attention, il y a une contrainte essentielle : la nouvelle variable doit être une fonction monotone (croissante ou décroissante) de la première dans l'intervalle d'intégration.

Un usage courant est de se ramener à des fonctions dont on connaît la primitive en escamotant les constantes multiplicatives. Pour calculer $I = \int_a^b \frac{6}{\sqrt{1+4\,x^2}}\,\mathrm{d}x$, on pose $y = 2\,x$ d'où $\mathrm{d}y = 2\,\mathrm{d}x$ et :

$$I = 3\int_{2\,a}^{2\,b} \frac{1}{\sqrt{1+y^2}}\,\mathrm{d}y = 3\,[\mathrm{argsh}(2\,b) - \mathrm{argsh}(2\,a)]$$

Un autre usage classique permet d'éviter la linéarisation de fonctions trigonométriques. Soit à calculer $I = \int_0^{\frac{\pi}{2}} \cos^3(x)\,\mathrm{d}x$ que l'on réécrit ainsi :

$$I = \int_0^{\frac{\pi}{2}} \cos^3(x)\,\mathrm{d}x = \int_0^{\frac{\pi}{2}} \cos^2(x)\,\cos(x)\,\mathrm{d}x = \int_0^{\frac{\pi}{2}} [1-\sin^2(x)]\,\cos(x)\,\mathrm{d}x$$

3. A partir de la dérivée $\frac{\mathrm{d}t}{\mathrm{d}\varphi}$, les mathématiciens introduisent la différentielle $\mathrm{d}t = \dfrac{\mathrm{d}t}{\mathrm{d}\varphi}\,\mathrm{d}\varphi$ et montrent la validité de ce qui suit.

On pose $t = \sin(x)$ d'où $\mathrm{d}t = \cos(x)\,\mathrm{d}x$, d'où :

$$I = \int_0^1 (1 - t^2)\,\mathrm{d}t = \left[t - \frac{t^3}{3}\right]_0^1 = \frac{2}{3}$$

Cette méthode convient pour tous les monômes trigonométriques en $\cos^p(x)\,\sin^q(x)$ où au moins un des deux entiers p et q est impair (sinon, voir plus loin).

• Intégration par parties.

Si la fonction f à intégrer se factorise sous une forme $f(x) = g(x)\,h'(x)$, on utilise la formule de dérivation du produit $g\,h$ pour affirmer que :

$$\int_a^b g(x)\,h'(x)\,\mathrm{d}x = \int_a^b [g(x)\,h(x)]'\,\mathrm{d}x - \int_a^b g'(x)\,h(x)\,\mathrm{d}x = [g(x)\,h(x)]_a^b - \int_a^b h(x)\,g'(x)\,\mathrm{d}x$$

Avec un peu de chance, on tombe sur une intégration plus aisée ; l'exemple classique est $\int_1^2 \ln(x)\,\mathrm{d}x$ où $g(x) = \ln(x)$ et $h'(x) = 1$ d'où $g'(x) = \frac{1}{x}$ et $h(x) = x$ (on choisit la constante d'intégration la plus simple possible bien sûr) et l'on a :

$$\int_1^2 \ln(x)\,\mathrm{d}x = [x\,\ln(x)]_1^2 - \int_1^2 \frac{1}{x}\,x\,\mathrm{d}x = \cdots = 2\,\ln(2) - 1$$

• Cas des fractions rationnelles.

C'est assez technique et, dans le cas le plus général, il vaut mieux se reporter à un cours de mathématiques. Nous ne traiterons ici que le cas le plus fréquemment rencontré en physique, des fonctions de la forme $\frac{Cte}{(x-\alpha)\,(x-\beta)}$ (on prendra par exemple $Cte = 1$). Il suffit de chercher les constantes A et B telles que :

$$\frac{1}{(x-\alpha)\,(x-\beta)} = \frac{A}{(x-\alpha)} + \frac{B}{(x-\beta)}$$

Si l'on multiplie membre à membre par $(x - \alpha)$ puis que l'on donne à x la valeur α on tire successivement :

$$\frac{1}{(x-\beta)} = A + \frac{B\,(x-\alpha)}{(x-\beta)}$$

$$\frac{1}{(\alpha-\beta)} = A + 0$$

On procède symétriquement pour calculer B et l'on a :

$$\frac{1}{(x-\alpha)\,(x-\beta)} = \frac{1}{(\alpha-\beta)}\left[\frac{1}{(x-\alpha)} - \frac{1}{(x-\beta)}\right]$$

dont une primitive est $\dfrac{1}{(\alpha-\beta)}\,[\ln(|x-\alpha|) - \ln(|x-\beta|)]$

• **Linéarisation de fonctions trigonométriques.**

Soit à intégrer une fonction $\cos^{2\,p}(x)\,\sin^{2\,q}(x)$; la méthode précedemment exposée ne convient pas car les exposants sont tous deux pairs. On utilise les formules de trigonométrie pour abaisser les exposants :

$$\cos^{2\,p}(x)\,\sin^{2\,q}(x) = \left(\frac{1+\cos(2\,x)}{2}\right)^p \left(\frac{1-\cos(2\,x)}{2}\right)^q$$

et l'on développe cette expression en somme de monônes. Ceux dont au moins un des exposants est impair sont traités aisément par la méthode précédente et pour les autres, on réitère le processus. On peut aussi utiliser un logiciel de calcul formel qui sera plus rapide.

1.k Equations différentielles.

Une équation différentielle d'ordre 1 est une expression de la forme $F(y', y) = 0$ et l'on appelle solution de cette équation différentielle toute fonction y telle que (on appelle la variable x de façon arbitraire) $F[y'(x), y(x)]$ est la fonction nulle (au moins sur un intervalle de validité).

Une équation différentielle d'ordre 2 est une expression de la forme $F(y'', y', y) = 0$ et l'on appelle solution de l'équation différentielle toute fonction y telle que $F[y''(x), y'(x), y(x)]$ est la fonction nulle (au moins sur un intervalle de validité).

Sur le même principe, on définit les équations différentielles de tout ordre, mais en physique, on rencontre surtout les ordres 1 et 2.

Les seules équations différentielles qu'un physicien doive savoir absolument résoudre sont les équations *à variables séparées* et les équations linéiaires (voir ci-dessous). Pour tout autre type d'équation, il ne faut pas rougir de demander l'aide d'un [4] ami mathématicien car, sauf rares exceptions, il sera ravi vous rendre ce service.

• **Equations différentielles d'ordre un à variables séparés.**

Les équations différentielles d'ordre 1 *à variables séparées* sont celles que l'on peut mettre sous la forme $f(y)\,\frac{\mathrm{d}y}{\mathrm{d}x} = g(x)$ ou, si l'on préfère $f(y)\,\mathrm{d}y = g(x)\,\mathrm{d}x$. Si F et G sont des primitives de f et g, on montre que les solutions vérifient $F(y(x)] = G(x)$ soit encore en introduisant F^{-1}, fonction réciproque de F :

$$y(x) = F^{-1}[G(x)]$$

4. Le français possède un genre neutre qui se confond dans ses formes avec le masculin. Un ami mathématicien peut parfaitement être une amie mathématicienne. Ce n'est pas moi qui suis sexiste, c'est la langue.

Remarque 1 : F et G sont définies à deux constantes additives près, dont seule la différence importe. Si l'on connaît, à un instant initial les valeurs de x et y, on peut déterminer cette constante.

• Equations différentielles linéaires.

Les équations différentielles *linéaires* d'ordre 1 sont celles que l'on peut mettre sous la forme (on appelle la variable x de façon arbitraire) :

$$a(x)\,\frac{\mathrm{d}y}{\mathrm{d}x} + b(x)\,y(x) = c(x)$$

Soit $y_P(x)$ une solution particulière de cette équation ; toutes les autres solutions sont de la forme $y(x) = y_P(x) + Y(x)$ où $Y(x)$ est solution de l'équation dite *homogène* ou *sans second membre* :

$$a(x)\,\frac{\mathrm{d}Y}{\mathrm{d}x} + b(x)\,Y(x) = 0$$

Celle-ci est à variables séparées ; en effet, on peut l'écrire $\frac{1}{Y}\,\frac{\mathrm{d}Y}{\mathrm{d}x} = -\frac{b(x)}{a(x)}$ et l'on sait donc la résoudre (cf supra). Reste donc à savoir comment trouver une solution particulière à la solution *avec second membre*. Il n'y a pas de méthode universelle pour cela et le seul cas qu'il faille savoir traiter est celui où $a(x)$ et $b(x)$ sont des constantes ; on réécrit alors l'équation ainsi :

$$a\,\frac{\mathrm{d}y}{\mathrm{d}x} + b\,y(x) = c(x)$$

Une méthode qui convient souvent est de chercher une solution de la même forme que la fonction $c(x)$. Si c'est un polynôme de degré n, on cherche une solution polynomiale de même degré, si c'est une fonction exponentielle, on on cherche une solution exponentielle de même constante, si c'est une fonction trigonométrique, on on cherche une solution trigonométrique de même pulsation (sous forme de combinaison de sinus et de cosinus), etc.

Dans le dernier cas, la méthode des amplitudes complexes est très confortable ; elle est abondamment détaillée dans le chapitre sur les oscillateurs et nous n'y reviendrons pas ici.

Pour les équations différentielles linéaires d'ordre 2 que l'on peut mettre sous la forme :

$$a(x)\,\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + b(x)\,\frac{\mathrm{d}y}{\mathrm{d}x} + c(x)\,y(x) = d(x)$$

la recherche d'une solution particulière se mène de la même façon quand $a(x)$, $b(x)$ et $c(x)$ sont des constantes et il n'y a pas plus de méthode générale si elles ne le sont pas. Par contre la résolution de l'équation homogène n'est plus simple. On peut montrer que l'ensemble des solutions est l'ensemble des combinaisons linéaires de deux solutions non proportionnelles (c'est donc un espace vectoriel de dimension 2, voir plus loin) mais il

n'y a aucune méthode générale pour trouver ces deux solutions-là. Il faut toutefois savoir que si une astuce a permis d'en trouver une, qu'on note Y_1, on est sauvé : on pose alors $Y(x) = f(x)\,Y_1(x)$ que l'on reporte dans $a(x)\,\frac{\mathrm{d}^2Y}{\mathrm{d}x^2} + b(x)\,\frac{\mathrm{d}Y}{\mathrm{d}x} + c(x)\,Y(x) = 0$ pour obtenir succesivement :

$$a(x)\,[f''\,Y_1 + 2\,f'\,Y_1' + f\,Y_1''] + b(x)\,[f'\,Y_1 + f\,Y_1'] + c(x)\,f\,Y_1 = 0$$

$$a(x)\,Y_1\,f'' + [2\,a(x)\,Y_1' + b(x)\,Y_1]\,f' + [a(x)\,Y_1'' + b(x)\,Y_1' + c(x)\,Y_1]\,f = 0$$

d'où puisque Y_1 est solution ce qui annule le coefficient de f et en posant $F = f'$:

$$a(x)\,Y_1(x)\,\frac{\mathrm{d}F}{\mathrm{d}x} + [2\,a(x)\,Y_1'(x) + b(x)\,Y_1(x)]\,F = 0$$

qui est une équation homogène d'ordre 1 que l'on sait résoudre par séparation des variables.

1.l Fonctions de plusieurs variables.

Nous prendrons l'exemple de fonctions à deux variables notées ici x et y.

• Dérivées partielles.

La dérivée partielle de $f(x,y)$ par rapport à x, notée $\frac{\partial f}{\partial x}$ est la dérivée de $f(x,y)$ où l'on considère y comme un paramètre constant. Par exemple, on n'est pas troublé par l'affirmation que la dérivée de la fonction $f(x) = a\,x^2$ est $\frac{\mathrm{d}f}{\mathrm{d}x} = 2\,a\,x$; on ne devrait pas plus être troublé par l'affirmation que la dérivée partielle de $f(x,y) = y\,x^2$ est $\frac{\partial f}{\partial x} = 2\,y\,x$. Considérer y comme une constante c'est l'assimiler à un paramètre comme l'est le a de $a\,x^2$. On définit de même $\frac{\partial f}{\partial y}$.

Si f est fonction de x et y, il en est de même pour ses dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ dont on peut donc calculer, pour chacune, les dérivées partielles par rapport aux deux variables. On note $\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$, $\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$, $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$ et $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$.

Le théorème de Schwarz, affirme que sous certaines conditions, les dérivées partielles *croisées*, c'est-à-dire $\frac{\partial^2 f}{\partial x \partial y}$ et $\frac{\partial^2 f}{\partial y \partial x}$ sont égales. C'est vrai si ces dérivées secondes sont dérivables, mais c'est parfois vrai quand elles ne le sont pas.

• Equations aux dérivées partielles.

Il s'agit de trouver ici les solutions d'une équation différentielle qui relie les dérivées partielles d'une fonction de plusieurs variables. C'est un sujet mathématiquement épineux

et il n'y a pas grand chose de systématique pour nous aider. Les seules qu'un physicien doive savoir gérer sont d'une part l'équation de propagation :

$$\frac{\partial^2 f}{\partial t^2} = c^2 \frac{\partial^2 f}{\partial x^2}$$

qui est abondamment traitée dans le chapitre D-II sur les ondes stationnaires (solutions factorisées) et progressives (solutions progressives) et d'autre part l'équation de diffusion :

$$\frac{\partial f}{\partial t} = D \frac{\partial^2 f}{\partial x^2}$$

qui est abondamment traitée dans le chapitre E-X sur la diffusion. Nous renvoyons le lecteur à ces chapitres.

1.m Fonctions à valeurs vectorielles.

On étudie ici des fonctions dont les valeurs sont des vecteurs dans un espace vectoriel de dimension finie, réel ou complexe.

Une fois que l'on a compris que les dérivations, intégrations, résolution d'équations différentielles se mènent composante [5] à composante, tout est dit.

2 Algèbre linéaire.

2.a Espaces vectoriels et sous-espaces vectoriels.

• Espaces vectoriels.

On appelle espace vectoriel sur l'ensemble des réels ou des complexes un ensemble d'éléments appelés vecteurs, qu'on note ici $\overrightarrow{u}$, $\overrightarrow{v}$ et analogues, muni

- d'une loi d'addition entre vecteurs
  - commutative : $\overrightarrow{u} + \overrightarrow{v} = \overrightarrow{v} + \overrightarrow{u}$ pour tout $\overrightarrow{u}$ et tout $\overrightarrow{v}$
  - associative : $\overrightarrow{u} + (\overrightarrow{v} + \overrightarrow{w}) = (\overrightarrow{u} + \overrightarrow{v}) + \overrightarrow{w}$ pour tout $\overrightarrow{u}$, tout $\overrightarrow{v}$ et tout $\overrightarrow{w}$, ce qui permet d'écrire $\overrightarrow{u} + \overrightarrow{v} + \overrightarrow{w}$ puisque l'ordre des opérations est indifférent
  - avec un élément neutre noté $\overrightarrow{0}$ tel que pour tout $\overrightarrow{v}$, $\overrightarrow{0} + \overrightarrow{v} = \overrightarrow{v}$
  - et telle que tout $\overrightarrow{v}$ ait un opposé noté $-\overrightarrow{v}$ tel que $\overrightarrow{v} + (-\overrightarrow{v}) = \overrightarrow{0}$
- une loi de multiplication mixte entre un scalaire réel ou complexe (noté ici λ, μ et analogues) et un vecteur
  - distributive vis-à-vis du vecteur : $\lambda \cdot (\overrightarrow{u} + \overrightarrow{v}) = \lambda \cdot \overrightarrow{u} + \lambda \cdot \overrightarrow{v}$ pour tout λ, tout $\overrightarrow{u}$ et tout $\overrightarrow{v}$

5. composantes sur la base vectorielle choisie.

— distributive vis-à-vis du scalaire : $(\lambda + \mu) \cdot \overrightarrow{u} = \lambda \cdot \overrightarrow{u} + \mu \cdot \overrightarrow{u}$ pour tout λ, tout μ et tout $\overrightarrow{u}$
— associative vis-à-vis du scalaire : $(\lambda\,\mu) \cdot \overrightarrow{u} = \lambda \cdot (\mu \cdot \overrightarrow{u})$ pour tout λ, tout μ et tout $\overrightarrow{u}$
— et telle que l'unité scalaire (notée 1 bien sûr) en soit l'élément neutre à gauche : $1 \cdot \overrightarrow{u} = \overrightarrow{u}$ pour tout $\overrightarrow{u}$

On en déduit assez aisément les propriétés suivantes :
— on a $0 \cdot \overrightarrow{u} = \overrightarrow{0}$ pour tout $\overrightarrow{u}$
— on a $\lambda \cdot \overrightarrow{0} = \overrightarrow{0}$ pour tout λ
— on a $(-\lambda) \cdot \overrightarrow{u} = -(\lambda \cdot \overrightarrow{u})$ pour tout scalaire λ et tout $\overrightarrow{u}$
— on a $\lambda \cdot (-\overrightarrow{u}) = -(\lambda \cdot \overrightarrow{u})$ pour tout scalaire λ et tout $\overrightarrow{u}$

Cette présentation axiomatique ne doit pas cacher que les calculs se mènent exactement comme on voudrait qu'ils le fussent.

• Exemples d'espaces vectoriels.

L'ensemble des déplacements dans un plan ou dans l'espace (un déplacement pouvant être décrit par « 1,52 m vers l'avant, 0,73 m vers la gauche et 2,01 m vers le bas ») est un espace vectoriel.

L'ensemble de fonctions d'une variable en est un autre (la fonction $f + g$ est telle que pour tout x, $(f + g)(x) = f(x) + g(x)$ et la fonction $\lambda \cdot f$ telle que pour tout x, $(\lambda \cdot f)(x) = \lambda\, f(x)$)

L'ensemble des suites (ou des séries) numériques en est un troisième (même type de définition de la somme de deux suites et de son produit par un scalaire).

La vérification de ces affirmations est aisée à effectuer.

• Sous-espaces vectoriels.

Si E est un espace vectoriel et si F, sous-ensemble de E, est lui aussi un espace vectoriel, on dit que E est un sous-espace vectoriel de E.

On montre (c'est assez simple) que F est sous-espace vectoriel de E si et seulement si, pour tout couple de scalaires λ et μ et tout couple de vecteurs $\overrightarrow{u}$ et $\overrightarrow{v}$ de F, la combinaison linéaire $\lambda \cdot \overrightarrow{u} + \mu \cdot \overrightarrow{v}$ appartient elle aussi à F.

Si F_1 et F_2 sont deux sous-espaces vectoriels de E, il en est de même pour leur intersection $F_1 \cap F_2$(au sens ensembliste). Cette intersection n'est jamais vide mais peut se réduire à l'ensemble $\{\overrightarrow{0}\}$ formé par le seul vecteur nul.

Si F_1 et F_2 sont deux sous-espaces vectoriels de E, il en est de même pour l'ensemble $F = F_1 + F_2$ des sommes d'un vecteur de F_1 et d'un vecteur de F_2. Si de plus $F_1 \cap F_2 = \{\overrightarrow{0}\}$, on dit que F est *somme directe* de F_1 et F_2 et l'on note $F = F_1 \oplus F_2$. Dans ce cas tout

vecteur de F se décompose de façon unique en somme d'un vecteur de F_1 et d'un vecteur de F_2.

• Exemples de sous-espaces vectoriels.

Dans l'espace vectoriel des suites, sont des sous-espaces vectoriels, entre autres, les ensembles des suites bornées, des suites convergentes, des suites vérifiant une relation de récurrence linéaire.

Dans l'espace vectoriel des fonctions d'un variable réelle, sont des sous-espaces vectoriels, entre autres, les ensembles des fonctions continues (sur $\mathbb{R}$ ou sur un intervalle donné), des fonctions dérivables (idem), des solutions d'une équation linéaire homogène (cf supra).

Soit une famille finie ou infinie (dénombrable ou non) de vecteurs $\overrightarrow{v}_i$ d'un espace vectoriel, on peut montrer (pas difficile mais très subtil) que le plus petit sous-espace vectoriel qui les contienne tous est l'ensemble des combinaisons linéaires d'un nombre fini de ces vecteurs.

2.b Indépendance linéaire. Bases vectorielles. Espaces vectoriels de dimension finie.

• Vecteurs linéairement indépendants.

Un ensemble (ou *famille*) fini de vecteurs $\{\overrightarrow{u}_1, \overrightarrow{u}_2, \cdots, \overrightarrow{u}_n\}$ est dit libre ou linéairement indépendant si la relation $\sum_1^n \lambda_i \overrightarrow{u}_i = \overrightarrow{0}$ n'est possible que si les λ_i sont tous nuls.

Si ce n'est pas le cas, au moins un des vecteurs de la famille peut s'exprimer comme une combinaison linéaire de tous les autres.

• Famille génératrice et base d'un espace vectoriel de dimension finie.

Un ensemble (ou *famille*) fini de vecteurs $\{\overrightarrow{u}_1, \overrightarrow{u}_2, \cdots, \overrightarrow{u}_n\}$ d'un espace vectoriel E est dit générateur si tout vecteur de E peut s'écrire sous forme d'une combinaison linéaire des vecteurs de cet ensemble.

Un ensemble à la fois libre et générateur est qualifié de *base de l'espace vectoriel.* Toutes les bases d'un espace vectoriel, s'il en existe, contiennent le même nombre qu'on appelle alors *dimension* de l'espace vectoriel que l'on dit *de dimension finie.* Si ce n'est pas le cas, on dit qu'il est de dimension infinie et l'on est alors obligé de mélanger algèbre linéaire et analyse (convergence de sommes infinies ou d'intégrales) ; on n'aborde en physique ce délicat problème que dans les séries de FOURIER et transformations de FOURIER (voir dans les chapitres où c'est utile).

Dans toute la suite de cet exposé, on ne parlera plus que d'espaces vectoriels de dimension finie.

Dans un espace vectoriel de dimension finie et pour une base donnée, tout vecteur se décompose de façon unique en une combinaison linéaire des vecteurs de la base.

2.c Formes linéaires. Applications linéaires. Matrices.

• Formes linéaires.

Soit E un espace vectoriel sur l'ensemble des réels $\mathbb{R}$ (ou des complexes $\mathbb{C}$) ; on appelle forme linéaire une fonction f de E sur $\mathbb{R}$ (ou sur $\mathbb{Z}$) telle que pour tout couple de scalaires λ et μ et pour tout couple de vecteurs $\overrightarrow{u}$ et $\overrightarrow{v}$, on ait :

$$f(\lambda \cdot \overrightarrow{u} + \mu \cdot \overrightarrow{v}) = \lambda \cdot f(\overrightarrow{u}) + \mu \cdot f(\overrightarrow{v})$$

Pour tout vecteur $\overrightarrow{u}$ se décomposant sur une base donnée $\{\overrightarrow{v}_k\}$ en $\overrightarrow{u} = \sum_1^n u_k \overrightarrow{v}_k$, on a alors :

$$f(\overrightarrow{u}) = f\left(\sum_1^n u_k \overrightarrow{v}_k\right) = \sum_1^n u_k f(\overrightarrow{v}_k)$$

En notant $a_k = f(\overrightarrow{v}_k)$, on a donc $f(\overrightarrow{u}) = \sum_1^n a_k u_k$.

Si l'on appelle f_i la forme linéaire telle que $f_i(\overrightarrow{v}_k)$ est égal à l'unité si $k = i$ et égal à zéro sinon, alors $f_i(\overrightarrow{u}) = u_i$ d'où, pour tout $\overrightarrow{u}$:

$$f(\overrightarrow{u}) = \sum_1^n a_k f_k(\overrightarrow{u})$$

.

d'où l'on peut déduire que f s'écrit comme fonction linéaire de f_k sous la forme :

$$f = \sum_1^n a_k f_k$$

ce qui prouve que l'ensemble des formes linéaire d'un espace vectoriel E est lui aussi un espace vectoriel, appelé *espace dual*, noté E^*, de même dimension que E et dont une base, appelée *base duale* est l'ensemble des f_k définis à partir de la base de E.

On peut aller plus loin, en étudiant les formes linéaires sur cet ensemble dual, pour construire un *ensemble bidual*, notée E^{**} et dont nous noterons Φ les éléments. Mais on peut alors créer une bijection entre E et son bidual qui à tout vecteur $\overrightarrow{u}$ associe $\Phi_{\overrightarrow{u}}$ défini par $\Phi_{\overrightarrow{u}}(f) = f(\overrightarrow{u})$ pour toute forme linéaire f, ce qui permet d'assimiler, à une bijection près, E et son bidual. Nous admettrons que, pour un espace vectoriel de dimension infinie, cette construction ne définit plus une bijection mais une simple injection et l'on peut trouver dans le bidual des choses étranges (voir les indications sur les distributions dans le chapitre A-X).

• **Applications linéaires.**

Soit E et F (éventuellement égal à E) deux espaces vectoriels ; on appelle application linéaire[6] une fonction $\overrightarrow{f}$ de E sur F telle que pour tout couple de scalaires λ et μ et pour tout couple de vecteurs $\overrightarrow{u}$ et $\overrightarrow{v}$, on ait :

$$\overrightarrow{f}(\lambda \cdot \overrightarrow{u} + \mu \cdot \overrightarrow{v}) = \lambda \cdot \overrightarrow{f}(\overrightarrow{u}) + \mu \cdot \overrightarrow{f}(\overrightarrow{v})$$

Pour tout vecteur $\overrightarrow{u}$ de E de dimension n se décomposant sur une base donnée $\{\overrightarrow{v}_j\}$ de E en $\overrightarrow{u} = \sum_{j=1}^{j=n} u_j \, \overrightarrow{v}_j$, on a alors :

$$\overrightarrow{f}(\overrightarrow{u}) = \sum_{1}^{n} u_j \, \overrightarrow{f}(\overrightarrow{v}_j)$$

Si l'on décompose chacun des $\overrightarrow{f}(\overrightarrow{v}_j)$ sur une base donnée $\{\overrightarrow{w}_i\}$ de F en $\overrightarrow{f}(\overrightarrow{v}_j) = \sum_{i=1}^{i=m} a_{ij} \, \overrightarrow{w}_i$, on a :

$$\overrightarrow{f}(\overrightarrow{u}) = \sum_{1}^{n} \sum_{i=1}^{i=m} u_j \, a_{ij} \, \overrightarrow{w}_i$$

Remarque : si l'on note $\sum_{i=1}^{i=m} U_i \, \overrightarrow{w}_i$ la décomposition de $\overrightarrow{f}(\overrightarrow{u})$ sur la base de F, on a donc :

$$U_i = \sum_{1}^{n} a_{ij} \, u_j$$

En raisonnant comme pour les forme linéaires, on en déduira que l'ensemble des applications linéaires de E dans F est un espace vectoriel de dimension égale au produit des dimensions de E et F, soit avec nos notations $n\,m$.

L'application linéaire pourra être décrite par le tableau à m lignes et n colonnes des a_{ij} que l'on appellera *matrice* associée à l'application linéaire. Nous y reviendrons sous peu.

• **Noyau et image d'une application linéaire.**

Si $\overrightarrow{f}$ est une application linéaire de E dans F, on appelle noyau de $\overrightarrow{f}$ le sous-espace vectoriel (éventuellement réduit au seul $\overrightarrow{0}$) dans E des vecteurs tels que $\overrightarrow{f}(\overrightarrow{u}) = \overrightarrow{0}$ et on appelle image de $\overrightarrow{f}$ le sous-espace vectoriel dans F de tous les vecteurs $\overrightarrow{f}(\overrightarrow{u})$.

On démontre que la somme des dimensions du noyau et de l'image est égale à celle de l'espace vectoriel de départ (soit E).

6. selon que E et F sont égaux ou non, que la fonction est bijective ou non, les mathématiciens la baptise homo-, endo-, iso- ou auto-morphisme.

• **Produit d'applications linéaires et de matrices.**

Soit une application linéaire $\overrightarrow{f}$ de E de base $\{\overrightarrow{e}_k\}$ de dimension n dans F de base $\{\overrightarrow{f}_j\}$ de dimension m qui à $\overrightarrow{u} = \sum_{k=1}^{k=n} u_k \overrightarrow{e}_k$ fait correspondre $\overrightarrow{v} = \overrightarrow{f}(\overrightarrow{u}) = \sum_{j=1}^{j=m} v_j \overrightarrow{f}_j$ avec $v_j = \sum_{k=1}^{k=n} a_{jk} u_k$ (cf supra).

Soit une application linéaire $\overrightarrow{g}$ de F de base $\{\overrightarrow{f}_j\}$ de dimension m dans G de base $\{\overrightarrow{g}_i\}$ de dimension p qui à $\overrightarrow{v} = \sum_{j=1}^{j=m} v_j \overrightarrow{f}_j$ fait correspondre $\overrightarrow{w} = \overrightarrow{g}(\overrightarrow{v}) = \sum_{i=1}^{i=p} w_i \overrightarrow{g}_i$ avec $w_i = \sum_{j=1}^{j=m} b_{ij} v_j$ (cf supra).

On définit le produit de ces applications, noté $g \circ f$ par

$$g \circ f(\overrightarrow{u}) = g[f(\overrightarrow{u})]$$

qui à $\overrightarrow{u} = \sum_{k=1}^{k=n} u_k \overrightarrow{e}_k$ fait correspondre $\overrightarrow{w} = \overrightarrow{g}(\overrightarrow{v}) = \sum_{i=1}^{i=p} w_i \overrightarrow{g}_i$ avec :

$$w_i = \sum_{j=1}^{j=m} b_{ij} \left(\sum_{k=1}^{k=n} a_{jk} u_k \right)$$

Si on la considère comme une application linéaire de E dans G avec $w_i = \sum_{k=1}^{k=n} c_{ik} u_k$, on aura alors :

$$c_{ik} = \sum_{j=1}^{j=m} b_{ij} a_{jk}$$

Cette relation définit le produit des matrices associées aux applications linéaires.

Remarque importante : Le produit d'applications linéaires et donc de matrices n'est pas commutatif [7].

• **Matrice carrées.**

Il s'agit de matrices telles que le nombre n de lignes et m de colonnes sont égales.

La transposée d'une matrice (M) de coefficients a_{ij}, est une matrice notée ${}^t(M)$ de coefficients b_{ij} tels que $b_{ij} = a_{ji}$. Si M est le produit de deux matrices, que l'on note ici $(M) = (A)\,(B)$, on a (notez l'inversion de l'ordre) :

$${}^t(M) = {}^t(B)\,{}^t(A)$$

Une matrice (M) est dite symétrique si ${}^t(M) = (M)$ et antisymétrique si ${}^t(M) = -(M)$ (ses termes *diagonaux* sont alors nuls). Toute matrice est somme, d'une façon unique, d'une matrice symétrique $(\frac{1}{2}\,[(M) + {}^t(M)])$ et d'une matrice anti-symétrique $(\frac{1}{2}\,[(M) - {}^t(M)])$.

7. La question ne se pose en fait que si $E = F = G$ et donc $n = m = p$.

L'inverse d'une matrice (M) est, si elle existe, une matrice notée M^{-1} telle que :

$$(M)^{-1}\,(M) = (M)\,(M)^{-1} = (Id)$$

Où (Id) est la matrice unité de coefficients I_{ij} nuls si $i \neq j$ et égaux à l'unité si $i = j$ (c'est la matrice associée à l'application identité). Si M est le produit de deux matrices, que l'on note encore $(M) = (A)\,(B)$, on a (notez encore l'inversion de l'ordre) :

$$(M)^{-1} = (B)^{-1}\,(A)^{-1}$$

Soit une application d'un espace vectoriel E dans lui-même de matrice (M), carrée donc. Effectuons un changement de base passant des n vecteurs $\{\overrightarrow{e}_j\}$ aux n vecteurs $\{\overrightarrow{f}_i\}$ définis par $\overrightarrow{f}_i = \sum_{j=1}^{j=n} p_{ij}\,\overrightarrow{e}_j$. Assimilons le tableau des coefficients p_{ij} à une matrice carrée, appelée *matrice de passage*. Avec cette nouvelle base, la même application a une nouvelle matrice noté (M') telle que :

$$(M') = (P)^{-1}\,(M)\,(P)$$

2.d Formes bilinéaires et quadratiques, sesquilinéaires et hermitiennes.

Dans un espace vectoriel réel, on appelle forme bilinéaire (notée f) toute application qui à un doublet de vecteurs associe un réel et telle que :

$$\begin{cases} f(\lambda\,\overrightarrow{u} + \mu\,\overrightarrow{v}, \overrightarrow{w}) = \lambda\,f(\overrightarrow{u}, \overrightarrow{w}) + \mu\,f(\overrightarrow{v}, \overrightarrow{w}) \\ f(\overrightarrow{u}, \lambda\,\overrightarrow{v} + \mu\,\overrightarrow{w}) = \lambda\,f(\overrightarrow{u}, \overrightarrow{v}) + \mu\,f(\overrightarrow{u}, \overrightarrow{w}) \end{cases}$$

Si $\{\overrightarrow{e}_i\}$ est une base de l'espace vectoriel et en notant $a_{ij} = f(\overrightarrow{e}_i, \overrightarrow{e}_j)$, on a pour $\overrightarrow{u} = \sum u_i\,\overrightarrow{e}_i$ et $\overrightarrow{v} = \sum v_j\,\overrightarrow{e}_j$:

$$f(\overrightarrow{u}, \overrightarrow{v}) = \sum_i \sum_j a_{ij}\,u_i\,v_j$$

Une forme bilinéaire est dite symétrique si $f(\overrightarrow{u}, \overrightarrow{v}) = f(\overrightarrow{v}, \overrightarrow{u})$ donc telle que $a_{ji} = a_{ij}$ et anti-symétrique si $f(\overrightarrow{u}, \overrightarrow{v}) = -f(\overrightarrow{v}, \overrightarrow{u})$ donc telle que $a_{ji} = -a_{ij}$ (en particulier $a_{ii} = 0$).

On appelle forme quadratique F associée à une forme bilinéaire symétrique f l'application (non linéaire) qui à tout vecteur associe $F(\overrightarrow{u}) = f(\overrightarrow{u}, \overrightarrow{u})$.

Remarque : on montre aisément que $f(\overrightarrow{u}, \overrightarrow{v}) = \frac{1}{2}\,[F(\overrightarrow{u} + \overrightarrow{v}) - F(\overrightarrow{u}) - F(\overrightarrow{v})]$

Dans un espace vectoriel complexe, on appelle forme sesquilinéaire [8] à droite [9] (notée f) toute application qui à un doublet de vecteurs associe un complexe et qui soit linéaire pour

8. du suffixe latin -sesqui (une fois et demie), contraction de semisque (et un demi).

9. Bien sûr, il existe des formes sesquilinéaires à gauche. Pendant mes études les mathématiciens privilégiaient les sesquilinéaires à droite ; désormais ce sont celles à gauche. Quand a eu lieu le changement ? Il me plaît de croire que c'est en 1981 mais cela relève du fantasme personnel.

le premier vecteur et *semi-linéaire* pour le second, c'est-à-dire telle que :

$$\begin{cases} f(\lambda\,\overrightarrow{u}+\mu\,\overrightarrow{v},\overrightarrow{w}) = \lambda\,f(\overrightarrow{u},\overrightarrow{w})+\mu\,f(\overrightarrow{v},\overrightarrow{w}) \\ f(\overrightarrow{u},\lambda\,\overrightarrow{v}+\mu\,\overrightarrow{w}) = \bar{\lambda}\,f(\overrightarrow{u},\overrightarrow{v})+\bar{\mu}\,f(\overrightarrow{u},\overrightarrow{w}) \end{cases}$$

où $\bar{\lambda}$ et $\bar{\mu}$ sont les complexes conjugués de λ et μ

Si $\{\overrightarrow{e}_i\}$ est une base de l'espace vectoriel et en notant $a_{ij} = f(\overrightarrow{e}_i, \overrightarrow{e}_j)$, on a pour $\overrightarrow{u} = \sum u_i\,\overrightarrow{e}_i$ et $\overrightarrow{v} = \sum v_j\,\overrightarrow{e}_j$:

$$f(\overrightarrow{u},\overrightarrow{v}) = \sum_i \sum_j a_{ij}\,u_i\,\bar{v}_j$$

Une forme sesquilinéaire est dite à symétrie hermitienne si $f(\overrightarrow{u},\overrightarrow{v}) = \overline{f(\overrightarrow{v},\overrightarrow{u})}$ donc telle que $a_{ji} = \bar{a}_{ij}$.

On appelle forme hermitienne F associée à une forme sesquilinéaire à symétrie hermitienne f l'application (non linéaire) qui à tout vecteur associe $F(\overrightarrow{u}) = f(\overrightarrow{u},\overrightarrow{u})$.

Remarque : on montre que :

$$f(\overrightarrow{u},\overrightarrow{v}) = \frac{1}{2}\,[F(\overrightarrow{u}+\overrightarrow{v}) + i\,F(\overrightarrow{u}+i\,\overrightarrow{v}) - (1+i)\,F(\overrightarrow{u}) - (1+i)\,F(\overrightarrow{v})]$$

2.e Déterminants. Systèmes d'équations linéaires. Inversion de matrices.

• Déterminants.

Si l'on cherche à construire dans un espace vectoriel E de dimension n et de base $\{\overrightarrow{e}_i\}$ une forme n-linéaire totalement anti-symétrique, c'est-à-dire une fonction de n vecteurs à valeurs scalaires, notée $f(\overrightarrow{u}_1, \overrightarrow{u}_2, \cdots, \overrightarrow{u}_n)$, qui soit linéaire vis-à-vis de chacun des vecteurs (cf forme bilinéaire) et anti-symétrique vis-vis de chaque couple de vecteurs (autrement dit qui change de signe par toute permutation de deux vecteurs), on montre qu'il n'existe qu'une solution à une constante multiplicative près. On appelle déterminant celle pour laquelle on obtient l'unité en l'appliquant aux vecteurs de la base pris dans l'ordre, soit $f(\overrightarrow{e}_1, \overrightarrow{e}_2, \cdots, \overrightarrow{e}_n) = 1$.

Il est d'usage de présenter les n vecteurs $\overrightarrow{u}_i$ sous forme d'un tableau assimilé à une matrice dont les colonnes sont les coordonnées des vecteurs dans la base de E et le déterminant est écrit en remplaçant les parenthèses de la notation matricielle par des barres verticales, ce qui permet de parler du déterminant d'une matrice. Le calcul pratique de déterminants est assez technique et il serait trop long de le développer ici, d'autant plus que les logiciels de calcul formel nous l'épargnent. Le seul qu'il soit utile de connaître est celui de la dimension 2 :

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\,a_{22} - a_{12}\,a_{21}$$

Une propriété à connaître : le déterminant d'un produit de deux matrices est égal au produit des déterminants.

• Systèmes d'équations linéaires. Inversion de matrices.

Là aussi c'est un sujet très technique et très long à développer. Nous nous contenterons d'un cas particulier dont la connaissance est un puissant outil en physique.

Un système de n équations linéaires à n inconnues notées x_i s'écrit sous la forme :

$$\begin{cases} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1n}\,x_n = b_1 \\ a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2n}\,x_n = b_2 \\ \cdots \\ a_{n1}\,x_1 + a_{n2}\,x_2 + \cdots + a_{nn}\,x_n = b_n \end{cases}$$

Si l'on introduit la matrice (A) à n lignes et n colonnes des coefficients a_{ij} et les matrices (X) et (B) à n lignes et une colonne, respectivement des inconnues x_i et des paramètres b_i, on peut écrire de façon condensée $(A)\,(X) = (B)$.

On montre que si le déterminant de la matrice (A) (appelé alors déterminant du système) est non nul, elle est inversible et le système d'équations a une solution unique qui, matriciellement, s'écrit $(X) = (A)^{-1}\,(B)$. La résolution pratique ne pose pas de problème majeur et est vraisemblablement déjà connue du lecteur.

En particulier si les b_i sont tous nuls (on dit que le système est homogène), on ne déduit que les inconnues sont elles aussi toutes nulles.

Un corollaire important est que si l'on veut que la résolution d'un système homogène ne conduise pas à la solution unique où les inconnues sont toutes nulles, il faut que le déterminant du système soit nul. C'est ainsi que l'on trouve la plupart du temps les équations de dispersion (voir la physique vibratoire et ondulatoire).

Si le déterminant d'un système homogène est nul, sauf cas particulier [10], il existe une solution définie à une constante multiplicative près. Les x_i n'ont donc pas une valeur unique mais le rapport de deux des x_i est bien déterminé et il importe de le chercher (voir toujours en mécanique ondulatoire les notions d'impédance).

2.f Diagonalisation. Valeurs propres et vecteurs propres.

• Valeurs propres et vecteurs propres.

Soit f une application linéaire d'un espace vectoriel E de dimension n dans lui-même, de matrice (A). On appelle vecteur propre et valeur propre associée un vecteur $\overrightarrow{u}$ et un scalaire λ (réel ou complexe selon la nature de E) tels que $f(\overrightarrow{v}) = \lambda\,\overrightarrow{v}$

10. Pour les spécialistes, ce n'est plus vrai si le système est de rang inférieur à $n-1$

Si $\overrightarrow{v}$ est vecteur propre pour une valeur propres λ, il est aisé de vérifier que tout vecteur $\mu\,\overrightarrow{v}$ est lui aussi vecteur propre pour la même valeur propre.

Matriciellement, en notant (Id) la matrice identité (cf supra), on veut :

$$(A)\,(V) = \lambda\,(V) = \lambda\,(Id)\,(V)$$

soit :

$$[(A) - \lambda\,(Id)]\,(V) = (0)$$

ce qui impose (cf paragraphe précédent) que le déterminant de la matrice $(A) - \lambda\,(Id)$, appelé *déterminant caractéristique*, soit nul.

L'équation en λ qui exprime que ce déterminant est nul est de degré n (la dimension de E) appelée *équation caractéristique*. Si E est un espace vectoriel complexe, elle admet n solutions, en général deux à deux distinctes. Dans ce cas, s'il l'on effectue un changement de base tel que la nouvelle base soit formé de n vecteurs propres correspondant aux n valeurs propres, dans cette nouvelle base, l'application linéaire a une matrice (A') diagonale (soit $a'_{ij} = 0$ si $i \neq j$ et les a'_{ii} sont les valeurs propres). On dit qu'on a diagonalisé la matrice, ce qui simplifie son utilisation ultérieure. S'il existe au moins une racine double, la diagonalisation n'est pas toujours possible ; on est alors bien obligé de « faire avec ».

Si l'espace vectoriel est réel et que deux racines sont imaginaires conjuguées, on essaie de leur donner un sens physique et si l'on y arrive, on considérera la matrice comme un cas particulier de matrice à coefficients complexes.

Remarque : une fois les valeurs propres trouvées, la recherche des vecteurs propres se ramène à la résolution d'un système homogène (cf supra).

• Diagonalisation d'une forme quadratique.

Soit, dans un espace vectoriel de dimension n, une forme quadratique F associée à une forme bilinéaire f à matrice (A) symétrique (cf supra). On peut monter que l'on peut construire n vecteurs $\overrightarrow{v}_i$ tels que $f(\overrightarrow{v}_i, \overrightarrow{v}_j)$ soit nul si i et j sont différents. Si l'on prend les $\overrightarrow{v}_i$ ainsi construits comme vecteurs d'une nouvelle base, on obtient pour la forme bilinéaire une nouvelle matrice diagonale dont les coefficients diagonaux sont la valeurs des $f(\overrightarrow{v}_i, \overrightarrow{v}_i)$.

La construction n'est pas unique ; par contre, on montre qu'on tombe toujours sur les mêmes nombres de coefficients diagonaux nuls, positifs et négatifs sur la diagonale ; c'est ce que l'on appelle la *signature* de la forme quadratique ; par exemple la métrique de la relativité restreinte définie par le carré de l'intervalle $\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}^2 - \mathrm{d}z^2$ (on trouve la convention opposée) est le résultat d'une telle diagonalisation avec une signature « $+ - - -$ ».

Remarque : il en va de même pour une forme hermitienne.

• **Diagonalisation d'une matrice symétrique dans un espace euclidien.**

On verra un peu plus loin la notion d'espace euclidien et de base orthonormée. Comme le lecteur sait vraisemblablement de quoi il s'agit, nous pouvons traiter dès maintenant ce troisième problème de diagonalisation.

Soit une matrice symétrique réelle (A) considérée comme plongée dans l'ensemble des matrices complexes (cf supra). Si $\overrightarrow{v}$ de composantes v_i possiblement complexes, associée à la valeur propre λ possiblement complexe. On a $\sum a_{ij}\, v_j = \lambda\, v_i$; en multipliant par $\bar{v}_i$ et en sommant, on arrive à :

$$\lambda \sum_i v_i\, \bar{v}_i = \sum_i a_{ii}\, v_i\, \bar{v}_i + \sum_i \sum_{j \neq i} a_{ij}\, (v_i\, \bar{v}_j + \bar{v}_i\, v_j)$$

d'où l'on déduit aisément que λ est réel.

En notant $\overrightarrow{f}$ l'application linéaire associée à (A) égale à ${}^t(A)$ par symétrie, le produit scalaire par un point, si $\overrightarrow{v}$ et $\overrightarrow{w}$ sont deux vecteurs propres associés à des valeurs propres distinctes λ et μ et (V) et (W) les vecteurs colonnes associés et $t(V)$ et $t(W)$ les vecteurs lignes correspondants, on peut écrire puisqu'un scalaire assimilé à une matrice à une ligne et une colonne est forcément une matrice symétrique :

$$\overrightarrow{v} \cdot \overrightarrow{f}\,(\overrightarrow{w}) = {}^t(V)\,(A)\,(W) = {}^t\left[{}^t(V)\,(A)\,(W)\right] = {}^t(W)\,{}^t(A)\,(V) = {}^t(W)\,(A)\,(V) = \overrightarrow{w} \cdot \overrightarrow{f}\,(\overrightarrow{v})$$

d'où successivement :

$$\begin{aligned} \overrightarrow{v} \cdot \overrightarrow{f}\,(\overrightarrow{w}) &= \overrightarrow{w} \cdot \overrightarrow{f}\,(\overrightarrow{v}) \\ \overrightarrow{v} \cdot \mu\, \overrightarrow{w} &= \overrightarrow{w} \cdot \lambda\, \overrightarrow{v} \\ (\mu - \lambda)\, \overrightarrow{v} \cdot \overrightarrow{w} &= 0 \end{aligned}$$

d'où, puisque λ et μ sont distincts, $\overrightarrow{v}$ et $\overrightarrow{w}$ sont orthogonaux.

A partir de là, on démontre assez facilement que l'on peut trouver une nouvelle base orthonormée dans l'espace vectoriel et dans laquelle l'application linéaire a une matrice diagonale. En physique dès que l'on croise une matrice diagonale, on se place dans une base orthonormée où elle devient diagonale, ce qui simplifie la suite de l'étude.

Remarque : il en est de même pour une matrice à symétrie hermitienne.

3 Géométrie.

3.a Espaces affines.

De façon parlante, un espace affine est un ensemble d'éléments appelés points tel que l'ensemble des déplacements d'un point à un autre forme un espace vectoriel. D'un point

de vue axiomatique, on dit que $\mathcal{A}$ est un espace affine associé à l'espace vectoriel E si à tout couple ordonné de deux points (notons-les ici M et N) de $\mathcal{A}$ on associe un vecteur de E, noté $\overrightarrow{MN}$ de façon que l'on ait les propriétés suivantes :

— pour tout triplet de points $\overrightarrow{MN} + \overrightarrow{NP} = \overrightarrow{MP}$ (relation de CHASLES)

— pour tout point M de $\mathcal{A}$ et tout vecteur $\overrightarrow{V}$ de E, il y a un point N de $\mathcal{A}$ et un seul tel que $\overrightarrow{MN} = \overrightarrow{V}$

— pour tout point M, on a $\overrightarrow{MM} = \overrightarrow{0}$ (c'est le cas où M et N sont confondus)

— pour tout couple de points M et N, on a $\overrightarrow{NM} = -\overrightarrow{MN}$

où en fait les deux dernières propriétés peuvent se déduire de la première.

Dans la pratique pour repérer un point de l'espace affine $\mathcal{A}$, on choisit un point particulier appelé origine, en général nommé O et l'on donne la valeur du vecteur $\overrightarrow{OM}$; en particulier en dimension finie, on donne de ce vecteur ses composantes dans une base de l'espace vectoriel de E qui forme avec O un repère de $\mathcal{A}$.

L'espace de la physique est affine de dimension 3. On note habituellement $\overrightarrow{e_x}$, $\overrightarrow{e_y}$ et $\overrightarrow{e_z}$ la base vectorielle.

3.b Plans et droites d'un espace affine tridimensionnel

Dans un espace affine tridimensionnel, on définit une droite à partir d'un point particulier (noté ici A) et d'un vecteur de E (noté ici $\overrightarrow{v}$) comme l'ensemble des points M de $\mathcal{A}$ pour lesquels il existe un réel λ tel que $\overrightarrow{AM} = \lambda\,\overrightarrow{v}$ (il s'agit d'une représentation paramétrique de la droite, le paramètre est λ).

Dans un espace affine tridimensionnel, on définit un plan à partir d'un point particulier (noté ici A) et de deux vecteurs de E (notés ici $\overrightarrow{v}$ et $\overrightarrow{w}$) comme l'ensemble des points M de $\mathcal{A}$ pour lesquels il existe deux réels λ et μ tels que $\overrightarrow{AM} = \lambda\,\overrightarrow{v} + \mu\,\overrightarrow{w}$ (il s'agit d'une représentation paramétrique du plan, les paramètres sont λ et μ).

On montre assez facilement que si tout point M de $\mathcal{A}$ est repéré par ses coordonnées (on note par exemple $\overrightarrow{OM} = x\,\overrightarrow{e_x} + y\,\overrightarrow{e_y} + z\,\overrightarrow{e_z}$), l'ensemble des points d'un plan vérifie une relation de la forme $a\,x + b\,y + c\,z = d$ où a, b, c et d sont des constantes ; cette relation est l'équation cartésienne du plan, ce qui est une autre façon de représenter un plan).

En général, deux plans se coupent selon une droite et les équations de ces deux plans sont une représentation cartésiennes de la droite.

Un plan est un espace affine de dimension deux, on peut y définir une droite de façon paramétrique comme ci-dessus ; si l'on choisit un repère pour le plan, un droite y aura une équation cartésienne de la forme $a\,x + b\,y = c$

3.c Barycentre.

Soient un ensemble de points M_i et un ensemble de coefficients réels m_i de somme non nulle, on définit le barycentre G des points M_i affectés des coefficients m_i par la relation :

$$\sum m_i \overrightarrow{GA_i} = \overrightarrow{0}$$

Comme la relation de CHASLES permet, avec l'origine O, d'écrire $\overrightarrow{GM_i} = \overrightarrow{OM_i} - \overrightarrow{OG}$, on a successivement :

$$\sum m_i \left(\overrightarrow{OM_i} - \overrightarrow{OG}\right) = \overrightarrow{0}$$

$$\sum m_i \overrightarrow{OM_i} = \sum \left(m_i \overrightarrow{OG}\right)$$

$$\sum m_i \overrightarrow{OM_i} = \left(\sum m_i\right) \overrightarrow{OG}$$

$$\overrightarrow{OG} = \frac{\sum m_i \overrightarrow{OM_i}}{\sum m_i}$$

Le barycentre existe donc et est unique. La relation $\sum m_i \overrightarrow{OM_i} = (\sum m_i) \overrightarrow{OG}$ sert parfois de définition au barycentre.

Remarque : on montre avec le même type de calcul que si la somme de m_i est nulle l'application qui à tout point P associe $\sum m_i \overrightarrow{M_iP}$ est une fonction constante.

3.d Espaces affines euclidiens.

Un espace affine est dit euclidien si on munit l'espace vectoriel associé (lui aussi qualifié d'euclidien) d'une forme quadratique F à signature « + + + »(cf supra). On a vu plus haut que l'on peut choir une base vectorielle qui rende la matrice de la forme quadratique diagonale ; en remplaçant les vecteurs de la base par des vecteurs bien choisis qui leur soient proportionnels, on arrive aisément à rendre la matrice égale à la matrice unité. On note désormais $\overrightarrow{v} \cdot \overrightarrow{w}$ pour $F(\overrightarrow{v}, \overrightarrow{w})$ et l'on dit qu'il s'agit du produit scalaire des deux vecteurs ; si ce produit est nul, on dit que les deux vecteurs sont orthogonaux ou perpendiculaires. On note aussi $\|\overrightarrow{v}\| = \sqrt{\overrightarrow{v} \cdot \overrightarrow{v}}$, que l'on appelle norme du vecteur $\overrightarrow{v}$; si celle-ci est égale à l'unité, on dit que le vecteur est normé. Les trois vecteurs de la base sont, par construction, normés et deux à deux orthogonaux ; la base est dite orthonormée.

Un espace euclidien n'a pas qu'une seule base orthonormée. Dans un changement de base orthonormée, la matrice de passage (P) (cf supra) a ces vecteurs colonnes identiques aux vecteurs de la nouvelle base (on note $\overrightarrow{e'_1}$, $\overrightarrow{e'_2}$ et $\overrightarrow{e'_3}$ ses vecteurs). Si on la multiplie par sa transposée ${}^t(P)$, les règles de calcul du produit de matrices (cf supra) montre que le coefficient sur la ligne d'indice i et la colonne d'indice j du produit ${}^t(P)\,(P)$ n'est autre que le produit scalaire de $\overrightarrow{e'_i}$ par $\overrightarrow{e'_j}$ donc 1 si $i = j$ et 0 sinon. Le produit est donc la matrice

unité ce qui prouve que la matrice inverse de (P) est sa transposée, aisée à calculer. On dit que (P) est une *matrice unitaire.*

On peut montrer que le déterminant de la matrice de passage est égal à ± 1 ; s'il est égal à 1, on dit que la nouvelle base est directe, s'il est égal à -1, qu'elle est indirecte.

3.e La trigonométrie plane par la géométrie.

Soit une droite d'un plan définie par un point et un vecteur $\overrightarrow{u}$ choisi unitaire. Si le repère orthonormée du plan est notée $(O, \overrightarrow{e_x}, \overrightarrow{e_y})$, on dit que l'angle que fait la droite avec l'axe Ox ou que fait $\overrightarrow{u}$ avec $\overrightarrow{e_x}$ est l'angle φ défini à $2\,k\,\pi$ près tel que $\overrightarrow{u} \cdot \overrightarrow{e_x} = \cos\varphi$ et $\overrightarrow{u} \cdot \overrightarrow{e_y} = \sin\varphi$. On a donc :

$$\overrightarrow{u} = \cos\varphi\,\overrightarrow{e_x} + \sin\varphi\,\overrightarrow{e_y}$$

Prenons comme nouvelle base du plan $\overrightarrow{e'_x} = \overrightarrow{u}$ et $\overrightarrow{e'_y} = -\sin\varphi\,\overrightarrow{e_x} + \cos\varphi\,\overrightarrow{e_y}$; on vérifie aisément qu'elle est orthonormée directe. Soit un vecteur unitaire $\overrightarrow{v}$ qui fait avec $\overrightarrow{e'_x}$ l'angle ψ ; on a successivement :

$$\overrightarrow{v} = \cos\psi\,\overrightarrow{e'_x} + \sin\psi\,\overrightarrow{e'_y}$$

$$\overrightarrow{v} = \cos\psi\,\left(\cos\varphi\,\overrightarrow{e_x} + \sin\psi\,\overrightarrow{e_y}\right) + \sin\varphi\,\left(-\sin\varphi\,\overrightarrow{e_x} + \cos\varphi\,\overrightarrow{e_y}\right)$$

$$\overrightarrow{v} = (\cos\psi\,\cos\varphi - \sin\psi\,\sin\varphi)\,\overrightarrow{e_x} + (\cos\psi\,\sin\varphi + \sin\psi\,\cos\varphi)\,\overrightarrow{e_y}$$

$$\overrightarrow{v} = \cos(\psi+\varphi)\,\overrightarrow{e_x} + \sin(\psi+\varphi)\,\overrightarrow{e_y}$$

ce qui prouve que $\overrightarrow{v}$ fait avec $\overrightarrow{e_x}$ l'angle $\psi + \varphi$, somme des angles que sont $\overrightarrow{u}$ avec $\overrightarrow{e_x}$ et $\overrightarrow{v}$ avec $\overrightarrow{u}$; il y a additivité des angles géométriques, avec une analogie avec la relation de CHASLES.

3.f Le triangle

La figure **??** p. **??** permet de visualiser les propriétés classiques d'un triangle de sommets A_1, A_2 et A_3 :

- Les trois *médianes* (en rouge) qui joignent les trois sommets aux milieux I_1, I_2 et I_3 des côtés opposés se coupent en un point G, appelé centre de gravité du triangle. Le point G se trouve au tiers de chaque médiane, celui le plus éloigné du sommet.
- Les trois *hauteurs* (en bleu), c'est à dire les perpendiculaires aux trois côtés joignant les sommets opposés aux *pieds des hauteurs* H_1, H_2 et H_3 se coupent en un point H, appelé orthocentre du triangle.
- Les trois *médiatrices* (en vert), c'est à dire les perpendiculaires aux trois côtés passant par leurs milieux I_1, I_2 et I_3 se coupent en un point O, qui est le centre du cercle circonscrit (en vert) au triangle passant par ses sommets.

— Les trois *bissectrices* (en orange), c'est à dire les droites issues des sommets et qui divisent les angles en deux parties égales se coupent en un point C, qui est le centre du cercle inscrit (en orange) au triangle tangent à ses côtés.

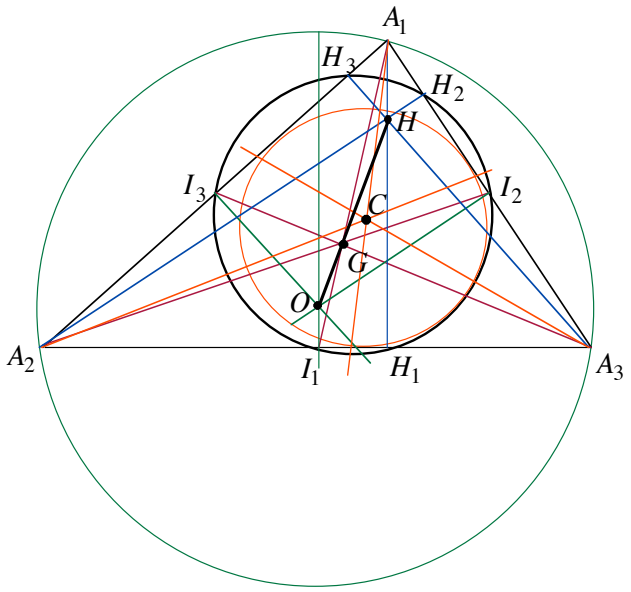

FIGURE 1 – Cercle d'Euler.

On démontre en outre les propriétés suivantes [11] :

— le centre O du cercle circonscrit, le centre de gravité G et l'orthocentre H sont alignés et l'on appelle *droite d'EULER* la droite qui les contient. Le point G est au tiers de OH, celui le plus proche de O.

— le milieu du segment OH est centre d'un cercle, appelé *cercle d'EULER* ou *cercle de FEUERBACH*, passant par les trois pieds des hauteurs, les trois milieux des côtés et les trois milieux des segments joignant les sommets au centre de gravité, c'est pourquoi on l'appelle plus souvent le *cercle des neuf points*.

— Le cercle des neuf points est tangent au cercle inscrit (théorème de FEUERBACH).

Appelons, comme le montre la figure **??** p. **??**, ℓ_1 la longueur du côté opposé au sommet A_1, c'est-à-dire le côté A_2A_3, (et analogues) et φ_1, l'angle au sommet A_1, c'est-à-dire l'angle entre les côtés A_1A_2 et A_1A_3 (et analogues).

Les deux propriétés suivantes sont intéressantes et d'un usage courant :

— Connaissant un angle et la longueur des deux côtés qui le forme, on calcule la longueur de troisième par la formule dite d'AL-KASHI [12] ou de PYTHAGORE généralisée,

11. Je doute fort qu'elles soient utiles au physicien mais je mène un combat d'arrière garde pour la survie de la géométrie.

12. nom apparu en France dans les années 1990, celui d'un mathématicien perse (1380 - 1429) mais le résultat était connu d'Euclide (vers 300 avant J. C. ?), sous une forme voisine, faute de trigonométrie.

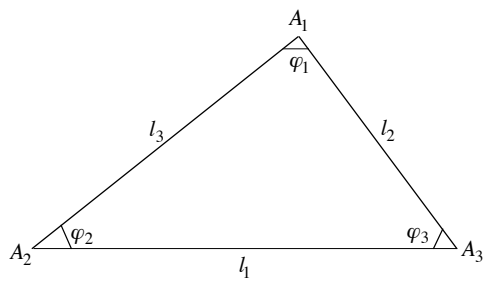

FIGURE 2 – Formule d'Al-Kashi.

par exemple :

$$\ell_3^2 = \ell_1^2 + \ell_2^2 - 2\,\ell_1\,\ell_2\,\cos\varphi_3$$

— les trois angles et les trois longueurs vérifient la *loi des sinus* :

$$\frac{\ell_1}{\sin\varphi_1} = \frac{\ell_2}{\sin\varphi_2} = \frac{\ell_3}{\sin\varphi_3}$$

et même, en introduisant le rayon R du cercle circonscrit, bien que ce soit d'un usage moins courant en physique :

$$\frac{\ell_1}{\sin\varphi_1} = \frac{\ell_2}{\sin\varphi_2} = \frac{\ell_3}{\sin\varphi_3} = 2\,R$$

3.g Courbes planes et gauches.

On peut décrire une courbe de multiples façons. Dans chaque cas, nous irons à l'essentiel.

• L'ordonnée comme fonction de l'abscisse.

On décrit une courbe comme l'ensemble des points d'abscisse x et d'ordonnée y vérifiant une relation de la forme $y = f(x)$.

L'équation d'une droite est de la forme $y = a\,x + b$ où a s'appelle le *coefficient directeur* de la droite.

Le coefficient directeur de la tangente à la courbe d'équation $y = f(x)$ au point d'abscisse x_0 et d'ordonnée $y_0 = f(x_0)$ est $f'(x_0)$ (où bien sûr f' désigne la fonction dérivée de f). L'équation de la tangente est donc, de façon brute :

$$y - y_0 = f'(x_0)\,(x - x_0)$$

Si au voisinage de x_0 la dérivée est une fonction croissante (décroissante), donc si $f''(0)$ est positif (négatif), la courbe tourne sa concavité vers les y croissants (décroissants).

L'aire comprise entre l'axe Ox et la courbe d'équation $y = f(x)$ entre les abscisses a et b est $\int_a^b f(x)\,\mathrm{d}x$ (le résultat est algébrique en fait).

La longueur de la courbe entre ses points d'abscisses a et b est $\int_a^b \sqrt{1 + f'(x)^2}\,\mathrm{d}x$; sa longueur entre une abscisse origine et un point quelconque s'appelle l'*abscisse curviligne de ce point* (il suffit donc d'adapter les bornes de l'intégrale précédente).

Parmi tous les cercles tangents à la courbe d'équation $y = f(x)$ au point de coordonnées x_0 et $f(x_0)$, celui qui « colle le plus à la courbe » [13], appelé *cercle osculateur* [14], a un rayon qu'on appelle *rayon de courbure* (son inverse s'appelle la *courbure*) qui est donné par la formule :

$$\frac{1}{R} = \frac{f''(x_0)}{[1 + f'(x_0)^2]^{\frac{3}{2}}}$$

qui donne un résultat algébrique lié à la concavité de la courbe (cf supra).

• Equation paramétrique.

On décrit une courbe comme l'ensemble des points repérés par un paramètre t (en physique, c'est souvent le temps) selon deux relations de la forme $x = f(t)$ et $y = g(t)$ (pour rendre la suite plus lisible, on notera ces fonctions $x(t)$ et $y(t)$).

Un vecteur tangent à la courbe au point de paramètre t_0 a pour composantes $x'(t_0)$ et $y'(t_0)$ de sorte que l'équation de la tangente (un point de cette tangente aura des coordonnées notées X et Y) soit :

$$\frac{X - x(t_0)}{x'(t_0)} = \frac{Y - y(t_0)}{y'(t_0)}$$

où l'on doit comprendre que si $x'(t_0)$ est nul, alors $[X - x(t_0)]$ aussi (l'équation est alors $X = x(t_0)$) et de même si $y'(t_0)$ est nul, alors $[Y - y(t_0)]$ aussi (l'équation est alors $Y = y(t_0)$). Si $x'(t_0)$ et $y'(t_0)$ sont tous deux nuls, on dit qu'on a affaire à un *point singulier* et on pose un joker (appel à un ami... mathématicien).

La longueur de la courbe entre ses points de paramètres a et b est $\int_a^b \sqrt{x'(t) + y'(t)^2}\,\mathrm{d}t$; sa longueur entre un paramètre origine et un point quelconque s'appelle là encore l'*abscisse curviligne de ce point*.

Le rayon de courbure R et la courbure $\frac{1}{R}$ au point de paramètre t_0 sont donnés par la formule :

$$\frac{1}{R} = \frac{x'(t_0)\,y''(t_0) - y'(t_0)\,x''(t_0)}{[x'(t_0) + y'(t_0)^2]^{\frac{3}{2}}}$$

Pour une courbe fermée, c'est-à-dire telle que pour deux valeurs a et b du paramètre $x(b) = x(a)$ et $y(b) = y(a)$, l'aire $\mathcal{A}$ à l'intérieur de la courbe est donnée par l'une de ces

13. A l'abscisse x voisine de x_0 l'écart entre la courbe et un des cercles tangents est en général du second ordre en $(x - x_0)$; pour le cercle osculateur, il est d'ordre 3.

14. du latin *osculare*, donner un baiser ; romantique, non ?

deux formules :

$$\mathcal{A} = \int_a^b y(t)\, x'(t)\, \mathrm{d}t = -\int_a^b x(t)\, y'(t)\, \mathrm{d}t$$

avec une valeur algébrique, positive si la courbe est décrite de a vers b dans le sens horaire et négative dans le sens trigonométrique.

Remarque : la représentation cartésienne n'est qu'un cas particulier de la représentation paramétrique, l'abscisse x sert de paramètre et la fonction $x(t)$ est remplacée par la fonction qui a x associe x, de dérivée égale à 1 pour tout x et de dérivée seconde nulle.

• Equation cartésienne.

La courbe est représentée comme un ensemble de points dont les coordonnées x et y sont liées par une relation de la forme $f(x, y) = 0$

Pour passer d'une équation paramétrique à une équation cartésienne, il suffit d'éliminer le paramètre entre $x(t)$ et $y(t)$ mais ce n'est pas toujours possible.

Pour passer d'une équation cartésienne à une équation paramétrique, on peut essayer de résoudre $f(x, y) = 0$ comme une équation en y où x est un paramètre (ou l'inverse) pour arriver à quelque chose de la forme $y = \varphi(x)$. Sinon, il faut essayer de se ramener à une identité remarquable ; par exemple à l'équation cartésienne $x^2 + y^2 = 1$, on peut proposer l'équation paramétrique $x(t) = \cos(t)$ et $y(t) = \sin(t)$ puisque $\cos^2(t) + \sin^2(t) = 1$

On verra l'exemple le plus important (les coniques) un peu plus loin.

• Equation polaire.

Si le plan est repéré par une origine O et une base orthonormée de vecteurs $\overrightarrow{e_x}$ et $\overrightarrow{e_y}$, on peut repérer un point par sa distance r au point O et l'angle θ que forment $\overrightarrow{e_x}$ et $\overrightarrow{OM}$; il est d'usage d'associer à tout point une base locale formé du vecteur unitaire de $\overrightarrow{OM}$ noté $\overrightarrow{e_r}$ et du vecteur unitaire qui lui est directement perpendiculaire $\overrightarrow{e_\theta}$

L'équation polaire d'une courbe consiste à considérer r comme fonction de θ ; qu'on notera $r(\theta)$

Un vecteur tangent à la courbe en un de ses points est $r'(\theta)\,\overrightarrow{e_r} + r(\theta)\,\overrightarrow{e_\theta}$ donc l'angle φ que fait la tangente avec $\overrightarrow{OM}$ est tel que $\operatorname{cotan}\varphi = \frac{r'(\theta)}{r(\theta)}$

Un résultat qu'on oublie trop souvent est celui-ci : si la courbe passe par le point O, la tangente fait avec $\overrightarrow{e_x}$ l'angle θ qui est celui pour lequel $r(\theta) = 0$ (correspondant donc au point O).

Nous ne donnons pas ici la formule du rayon de courbure car je n'en ai jamais eu besoin en physique ou alors si peu que je l'ai oublié.

• **Remarque sur les courbes gauches.**

Une courbe dans l'espace est dite *courbe gauche*. Sa représentation la plus aisée est la représentation paramétrique : on donne $x(t)$, $y(t)$ et $z(t)$ composantes de $\overrightarrow{OM}$ en fonction du paramètre t ; pour alléger l'écriture on notera $\overrightarrow{M}(t)$ la fonction qui à t associe $\overrightarrow{OM}$

Un vecteur tangent à la courbe au point de paramètre t est $\overrightarrow{M'}(t)$.

Le cercle osculateur en un point de la courbe de paramètre t est contenu dans le plan qui contient ce point et les vecteurs $\overrightarrow{M'}(t)$ et $\overrightarrow{M''}(t)$, appelé *plan osculateur*. La courbure est donnée par la formule :

$$\frac{1}{R} = \frac{\|\overrightarrow{M''} \wedge \overrightarrow{M'}\|}{\|\overrightarrow{M'}\|^3}$$

où l'on considère que le lecteur sait ce qu'est un produit vectoriel.

3.h Surfaces.

Une surface dans l'espace tridimensionnel peut être décrite soit par une équation cartésienne $f(x, y, z) = 0$ liant les coordonnées des points de la surface, en particulier des équations de la forme $z = f(x, y)$, soit par une équation paramétrique à deux paramètres (disons u et v) de la forme $\overrightarrow{OM} = f(u, v)$.

Une propriété importante en physique est celle-ci : si par un point de la surface et sa normale en ce point, on fait passer un premier plan quelconque et un second qui lui soit orthogonal, ces deux plans coupent la surface selon deux courbes dont les courbures en ce point sont notées $\frac{1}{R_1}$ et $\frac{1}{R_2}$, alors on montre que la somme $\frac{1}{R_1} + \frac{1}{R_2}$ est indépendante du choix du premier plan et on l'appelle courbure de la surface.

3.i Coniques.

Une conique est une courbe plane dont l'équation est de la forme :

$$A\,x^2 + 2\,B\,x\,y + C\,y^2 + 2\,D\,x + 2\,E\,y + F = 0$$

Un changement d'origine puis une rotation des axes permettent, sauf cas particulier, de supprimer les termes en x, y puis en $x\,y$; en divisant enfin par le terme constant, éventuellement changé de signe, on arrive à deux types de coniques classiques, l'ellipse et l'hyperbole. Nous passerons sous silence le cas particulier de la parabole car c'est un cas-charnière et les cas-charnières, qui nécessitent une valeur précise d'un paramètre, n'ont aucun sens physique, à cause des incertitudes.

• **Ellipse.**

Une ellipse est une courbe plane d'équation :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

où l'on choisit Ox de sorte que $a > b$

Une équation paramétrique en est $x = a \cos\varphi$ et $y = b \sin\varphi$; c'est donc une courbe fermée, son intersection avec Ox est formée des points d'abscisse $\pm a$ (le *grand axe*) et son intersection avec Oy des points d'ordonnée $\pm b$ (le *petit axe*).

Il existe sur le grand axe deux points F et F', appelés *foyers*, d'abscisses $\pm c$ avec $c^2 = a^2 - b^2$ tel que la somme des distances FM et $F'M$ de tout point M de l'ellipse aux foyers soit constante et égale à $2\,a$.

Remarque : le cercle est une ellipse particulière avec $a = b$ donc $c = 0$

• **Hyperbole.**

Une hyperbole est une courbe plane d'équation :

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

Une équation paramétrique en est $x = \pm a \operatorname{ch}\varphi$ et $y = b \operatorname{sh}\varphi$; c'est une courbe non-fermée à deux branches selon le signe de x, son intersection avec Ox est formée des points d'abscisse $\pm a$ et elle ne coupe pas Oy ; pour $\varphi \to \pm\infty$, x et y tendent $\pm\infty$ avec un rapport $\frac{y}{x} = \frac{b}{a} \operatorname{th}\varphi \to \pm\frac{b}{a}$ qui montre l'existence de deux asymptotes.

Il existe sur Ox deux points F et F', appelés *foyers*, d'abscisses $\pm c$ avec $c^2 = a^2 + b^2$ tel que la différence des distances FM et $F'M$ de tout point M de l'hyperbole aux foyers soit constante et égale à $2\,a$.

• **Equation polaire.**

Si l'on prend une représentation polaire de l'ellipse ou de l'hyperbole avec l'origine en l'un des foyers et avec $\theta = 0$ dans la direction de Ox, l'équation est dans les deux cas (selon que l'on a choisi F' ou F comme origine) :

$$r = \frac{p}{1 \pm e \cos\theta}$$

avec dans les deux cas $e = \frac{c}{a}$ (c'est l'*excentricité* inférieure à l'unité pour une ellipse et supérieure à l'unité pour une hyperbole, égale à l'unité pour le cas-charnière de la parabole et nulle pour le cas particulier du cercle) et $p = \frac{b^2}{a}$ (c'est le *paramètre* de la conique).

3.j Angles solides.

Définition.

Soit un cône de sommet O s'appuyant sur une courbe fermée quelconque. La surface interceptée sur une sphère de centre O a une taille proportionnelle par homothétie au rayon R de la sphère et donc une aire S proportionnelle au carré de R ; on appelle donc angle solide du cône le rapport $\Omega = \frac{S}{R^2}$, indépendant de R.

L'espace tout entier a donc, puisque l'aire d'une sphère est $4\,\pi\,R^2$, un angle solide de $4\,\pi$.

Angle solide élémentaire.

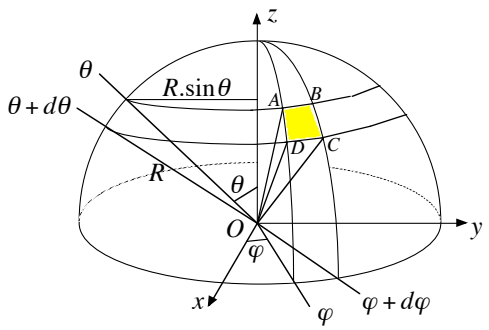

FIGURE 3 – Angle solide élémentaire.

Pour un cône découpant sur la sphère un rectangle entre les longitudes φ et $\varphi + \mathrm{d}\varphi$ et les colatitudes θ et $\theta + \mathrm{d}\theta$ de longueur $AB = R\,\sin\theta\,\mathrm{d}\varphi$ et de largeur $BC = R\,\mathrm{d}\theta$, l'angle solide élémentaire est donc (voir la figure **??** p. **??** qui rappelle les figures que j'ai jadis tracées au tableau noir) :

$$\mathrm{d}\Omega = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi$$

Angle solide sous lequel est vue une surface élémentaire.

Calculons maintenant sous quel angle solide est vue d'un point O une surface élémentaire $\mathrm{d}S$, autour d'un point M, de normale Mz. Il n'y a pas de problème majeur si Mz est parallèle à OM, sinon il faut projeter orthogonalement la surface $\mathrm{d}S$ en une surface $\mathrm{d}\Sigma$ sur la sphère de rayon OM, comme sur la figure **??** p. **??** à gauche (où, pour améliorer la lisibilité, M a été placé sur le bord de $\mathrm{d}S$), et l'on aura alors $\mathrm{d}\Omega = \frac{\mathrm{d}\Sigma}{r^2}$ où r désigne la distance OM.

Or, comme l'indique la partie droite de la même figure, si l'on projette sur un plan un rectangle avec un côté (la longueur) parallèle au plan et l'autre (la largeur) faisant avec lui un angle θ, les longueurs du rectangle et de sa projection sont égales mais la

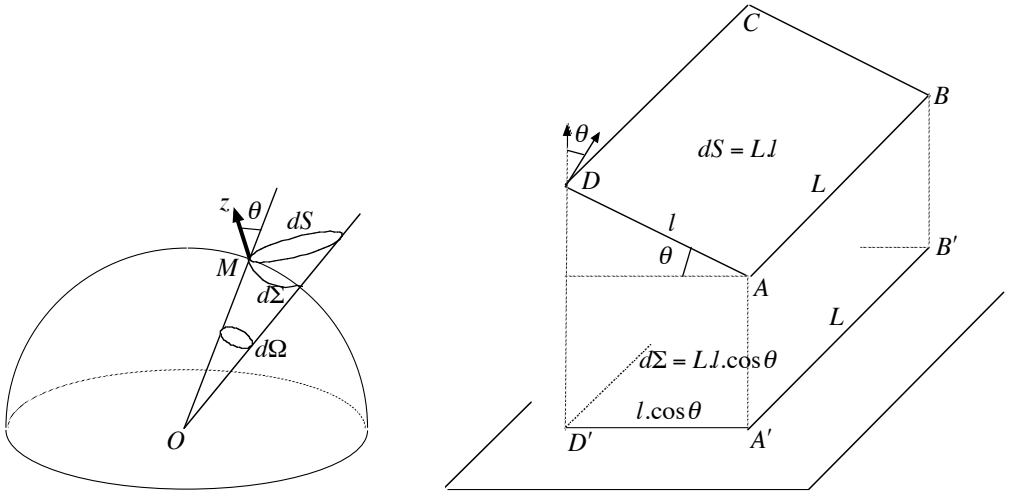

FIGURE 4 – Angle solide sous lequel est vue une surface élémentaire.

largeur de la projection, donc l'aire, a été multipliée par $\cos\theta$; il en est de même pour une surface quelconque, car elle peut être découpée en tas de petits rectangles. Nous avons donc $\mathrm{d}\Sigma = \mathrm{d}S\ \cos\theta$ et $\mathrm{d}\Omega = \frac{\mathrm{d}S\ \cos\theta}{r^2}$ où θ désigne l'angle entre OM et Mz.