

T. Horsin

Equations différentielles. Version du 25-07-2016, 19:36.

*On peut me signaler les erreurs ou les incorrections
qui émaillent sûrement ce document en m'écrivant à
<mailto:thierry.horsin@cnam.fr>*

1 Exemples de problème de physique conduisant à une équation différentielle

1.1 Approximation de l'optique géométrique

A reprendre, problème d'homogénéité.

Les équations de propagation d'une onde électromagnétique (dont la lumière est une forme) sont données par deux équations aux dérivées partielles vectorielles, a priori, qui sont

$$\begin{aligned} \operatorname{curl} H - \frac{1}{c} \partial_t D &= \frac{4\pi}{c} j \\ \operatorname{curl} E + \partial_t B &= 0, \\ \operatorname{div} D &= 4\pi \rho, \\ \operatorname{div} B &= 0 \end{aligned} \tag{1.1}$$

où H est le vecteur champ magnétique, j la densité de courant, D le déplacement électrique et c est homogène à une vitesse, ρ est la densité de charge, B est l'induction magnétique. Écrit comme ceci la vitesse de la lumière vaut 1.

Ces équations sont incomplètes telles que, et on doit rajouter des relations constitutives (matériaux dépendantes) telles que par exemple

$$\begin{aligned} j &= \sigma E, \\ D &= \varepsilon E, \\ B &= \mu H. \end{aligned} \tag{1.2}$$

Ici σ est la conductivité spécifique, ε la permittivité diélectrique et μ la perméabilité magnétique. Ces relations peuvent être largement complexifiées en fonction des matériaux de propagations. En milieu de propagation borné (par exemple dans une cavité) ce système d'équation ((1.1)) doit être complété par des conditions aux limites, et dans le cas d'une propagation dans l'espace entier de conditions de radiations à l'infini.

On cherche des solutions de ces équations (dans le cas où $\rho = 0$ et $j = 0$ sous la forme (sachant que seules les parties réelles nous intéressent)

$$E = E_0(r)e^{-i\omega t}, H = H_0(r)e^{-i\omega t}.$$

Ceci donne

$$\text{curl}H_0 + ik_0\varepsilon E_0 = 0,$$

$$\text{curl}E_0 - ik_0\mu H_0 = 0,$$

$$\text{div}\varepsilon E_0 = 0,$$

$$\text{div}\mu H_0 = 0.$$

Ici on a

$$k_0 = \frac{\omega}{c} = \frac{2\pi}{\lambda_0}$$

où λ_0 est la longueur d'onde dans le vide (de charge et de courant).

Des exemples simples montrent qu'il est intéressant de chercher E_0 et H_0 sous la forme

$$E_0 = \mathcal{H}(r)e^{ik_0S(r)}, H_0 = \mathcal{H}e^{ik_0S(r)},$$

ce qui conduit aux équations (en supposant ε et μ réels) à

$$\nabla S \wedge \mathcal{H} + \varepsilon \mathcal{E} = -\frac{1}{ik_0} \text{curl}\mathcal{H},$$

$$\nabla S \wedge \mathcal{E} - \mu \mathcal{H} = -\frac{1}{ik_0} \text{curl}\mathcal{E},$$

$$\mathcal{E} \cdot \nabla S = -\frac{1}{ik_0} (\mathcal{E} \cdot \nabla \log \varepsilon + \text{div}\mathcal{E}),$$

$$\mathcal{H} \cdot \nabla S = -\frac{1}{ik_0} (\mathcal{H} \cdot \nabla \log \mu + \text{div}\mathcal{H}).$$

L'approximation de l'optique géométrique suppose que k_0 est très grand (la longueur d'onde étant donc très petite) devant les autres quantités ce qui implique que le membre de droite des équations est 0.

$$\nabla \mathcal{S} \wedge \mathcal{H} + \varepsilon \mathcal{E} = -0,$$

$$\nabla \mathcal{S} \wedge \mathcal{E} - \mu \mathcal{H} = -0,$$

$$\mathcal{E} \cdot \nabla \mathcal{S} = 0,$$

$$\mathcal{H} \cdot \nabla \mathcal{S} = 0.$$

Ceci implique pour qu'il y ait une solution non triviale que

$$\frac{1}{\mu} ((\mathcal{E} \cdot \nabla \mathcal{S}) \nabla \mathcal{S} - \mathcal{E} (\nabla \mathcal{S})^2) + \varepsilon \mathcal{E} = 0,$$

et compte tenu de l'hypothèse faite conduit à

$$(\nabla \mathcal{S})^2 = \varepsilon \mu = n^2. \quad (1.3)$$

L'équation (1.3) est appelée équation eikonale ("icone" image en grec).

Les surfaces $\mathcal{S} = cte$ sont appelées front d'onde.

Les rayons lumineux sont les courbes orthogonales aux fronts d'onde.

Si $r(s)$ est la position d'un point sur un rayon paramétrée par longueur d'arc, l'équation d'un rayon devient:

$$nr'(s) = \nabla \mathcal{S} \quad (1.4)$$

Si on différencie cette relation, on obtient

$$\begin{aligned} \frac{d}{ds}(nr'(s)) &= \frac{d}{ds}(\nabla \mathcal{S}) \\ &= r'(s) \cdot \nabla(\nabla \mathcal{S}) \\ &= \frac{1}{n} \nabla \mathcal{S} \cdot \nabla(\nabla \mathcal{S}) \\ &= \frac{1}{2n} \nabla(\nabla \mathcal{S}^2) \\ &= \frac{1}{2n} \nabla(n^2), \end{aligned}$$

ce qui donne

$$\frac{d}{ds} \left(n \frac{dr}{ds} \right) = \nabla n, \quad (1.5)$$

qui est donc l'équation d'un rayon lumineux.

En particulier si n ne dépend pas du point, (le milieu est homogène) alors le rayon se propage en ligne droite.

Qu'advient-il si l'on change de milieu constant.

Le principe de Huygens dit que l'on doit minimiser le temps pour aller d'un point à un autre.

Considérons deux milieux de vitesse de propagation v_1 et v_2 .

Dans chacun des milieux, le déplacement est en ligne droite.

Considérons deux points M_1 et M_2 dans chacun des milieux de coordonnées respectives (x_1, y_1)

(x_2, y_2) avec $y_1 y_2 < 0$ le franchissement se faisant en un point $M_3 = (x_3, 0)$ avec $x_1 \leq x_3 \leq x_2$

Le temps mis pour aller de M_1 à M_3 est $\sqrt{(x_1 - x_3)^2 + y_1^2}/v_1$ et celui pour aller de M_3 à M_2 est $\sqrt{(x_2 - x_3)^2 + y_2^2}/v_2$.

Le temps total est donc

$$\sqrt{(x_1 - x_3)^2 + y_1^2}/v_1 + \sqrt{(x_2 - x_3)^2 + y_2^2}/v_2$$

qui pour être minimal doit vérifier

$$\frac{-(x_1 - x_3)}{v_1 \sqrt{(x_1 - x_3)^2 + y_1^2}} + \frac{-(x_2 - x_3)}{v_2 \sqrt{(x_2 - x_3)^2 + y_2^2}} = 0,$$

ce qui revient exactement à dire que (pour l'instant l'hypothèse des signes n'a pas joué) avec un réarrangement des signes que

$$\frac{\sin(i_1)}{v_1} = \frac{\sin(i_2)}{v_2},$$

ce qui permet de retrouver la loi de Descartes

$$n_1 \sin(i_1) = n_2 \sin(i_2). \quad (1.6)$$

1.2 La fibre optique à gradient d'indice

Une fibre optique est un guide d'onde cylindrique dans lequel se propage une onde qui subit une réflexion totale (en principe) sur les parois.

L'indice de réfraction n dépend de la distance à l'axe de la fibre. Le principe de Fermat donne que si un rayon se propage à partir de l'entrée de la fibre optique en faisant un angle α_1 avec la normale à la fibre alors

$$n(r) \sin(\alpha(r)) = n_1 \sin(\alpha_1).$$

Si l'on cherche la trajectoire de l'onde sous la forme d'une fonction $r(x)$ où x est l'abscisse le long de la fibre, on a

$$n(r(x)) \sin(\alpha(r(x))) = C$$

qui est donc constante.

Or $\alpha(r(x))$ est l'angle entre la tangente à la courbe et l'axe verticale. On a donc

$$r'(x) = \frac{1}{\tan(\alpha(r(x)))}$$

$$r'(x)^2 = \frac{1}{\frac{1}{\cos^2(\alpha(r(x)))} - 1} = \frac{1 - \sin^2(\alpha(r(x)))}{\sin^2(\alpha(r(x)))} = \frac{n(r)^2}{n_1^2 \sin^2(\alpha_1)} - 1.$$

Finalement on obtient

$$r'(x) = \pm \sqrt{\frac{n(r)^2}{n_1^2 \sin^2(\alpha_1)} - 1}. \quad (1.7)$$

Le choix du signe se fait en fonction de l'instant initial $r'(0) = \frac{1}{\tan(\alpha_1)}$.

On propose en général alors une loi $n(r)$.

Par exemple si la fibre optique est de rayon R , on peut prendre

$$n(r) = n_1 \left(1 - \frac{n_1 - n_2}{n_1} \frac{r^2}{R^2} \right).$$

Il y a donc une variation (un gradient d'indice) transversalement à l'axe de la fibre.

Pour information, l'indice n de refraction fait (ce n'est pas le seul) un lien entre les équations de Maxwell de propagation d'une onde électro-magnétique et l'optique géométrique. En effet, il est d'usage de définir $n = \sqrt{\epsilon_r \mu_r}$ où ϵ_r la permittivité diélectrique relative au vide et μ_r la perméabilité magnétique relative au vide. Plus précisément il est commun de définir $n = \frac{c}{v_\phi}$ où c est la vitesse de la lumière dans le vide et v_ϕ la vitesse de phase (la vitesse du point du sommet de la "vague") de la lumière dans le milieu.

Noter que la loi de réfraction $n_1 \sin(i_1) = n_2 \sin(i_2)$ impose en particulier que

$$\left| \frac{n_1}{n_2} \sin(i_1) \right| \leq 1.$$

Si cette quantité vaut 1 alors i_2 mesure un angle droit. Il y a alors réflexion totale. Donc si n_2 est trop petit, par rapport à n_1 , c'est à dire si la vitesse de propagation dans le milieu 2 est plus importante que dans le milieu 1 en fonction de l'angle incident i_1 donné. Autrement dit, la diffraction ne se fait qu'en deçà d'un angle d'incidence limite.

L'équation différentielle (1.7) doit permettre de retrouver cette angle d'incidence limite. L'angle initiale par rapport à l'axe de la fibre optique doit sûrement être assez faible pour que la fibre conduise le rayon lumineux.

Si l'on observe (1.7), chaque fois que la racine s'annule, on va faire face à une difficulté théorique de prolongement de la solution, puisque la solution peut devenir constante ou non.

Pour propager l'onde tout au long de la fibre, on aura donc besoin d'une autre information, qu'on peut mettre en évidence en dérivant (1.7) une nouvelle fois, ce qui permet de faire apparaître une équation différentielle du second ordre.

Notons également que la loi de Descartes (ou Descartes-Snell) se déduit d'hypothèses de conditions aux limites sur les équations de Maxwell et de formes particulières de certaines solutions.

Si on reprend

$$r'(x)^2 = \frac{1}{\frac{1}{\cos^2(\alpha(r(x)))} - 1} = \frac{1 - \sin^2(\alpha(r(x)))}{\sin^2(\alpha(r(x)))} = \frac{n(r)^2}{n_1^2 \sin^2(\alpha_1)} - 1,$$

tant que r' ne s'annule pas, on peut dériver.

Ce qui donne

$$r'(x)r''(x) = \frac{n'(r)n(r)}{n_1^2 \sin^2(\alpha_1)} r'(x).$$

Si $r'(x) \neq 0$ alors on peut simplifier et on trouve

$$r''(x) = \frac{n'(r)n(r)}{n_1^2 \sin^2(\alpha_1)}.$$

On montre que sous des hypothèses raisonnables sur n ce type d'équation a une unique solution locale en temps si on se donne $r(0)$ et $r'(0)$.

Cette solution est de classe C^2 (deux fois dérivable). Les points où $r'(x) = 0$ ne peuvent pas joindre une solution à r constante sauf à perdre la régularité.

Dans notre exemple

$$n'(r) = -2(n_2 - n_1) \frac{r}{R^2}.$$

On a

$$r'(x)^2 = \frac{\left(1 - \frac{n_1 - n_2}{n_1} \frac{r^2}{R^2}\right)^2}{\sin^2(\alpha_1)} - 1.$$

Pour $r = R$, on a

$$r'(x)^2 = \frac{n_2^2}{n_1^2 \sin^2(\alpha_1)} - 1,$$

il est donc nécessaire que

$$\sin^2(\alpha_1) \geq \frac{n_2^2}{n_1^2},$$

pour qu'on n'atteigne pas le bord.

Si r' n'est pas nul, r^2 augmente, donc $\left(1 - \frac{n_1 - n_2}{n_1} \frac{r^2}{R^2}\right)^2$ diminue (si $n_1 > n_2$ ce qui est le cas dans une fibre optique). Il s'ensuit que la condition

$$\sin^2(\alpha_1) \geq \frac{n_2^2}{n_1^2},$$

est aussi suffisante pour que la propagation se fasse.

Une simulation sera effectuée plus loin.

La question sous-jacente à la propagation est de savoir pourquoi "la physique" préfère la solution de classe C^2 . On ne peut pas vraiment répondre à cette question dans le cadre ci-dessus. En effet un rayon lumineux n'est qu'une représentation mathématique d'une onde électro-magnétique qui se propage "majoritairement" dans la direction du rayon. Il faut des considérations énergétiques supplémentaires qui permettent de comprendre cette hypothèse.

En général, l'approximation par rayon lumineux est considérée comme valable lorsque la longueur d'onde est très petite par rapport aux dimensions caractéristiques des milieux de propagations. L'énergie d'un rayon lumineux de fréquence ν est transportée (en l'absence de charge extérieure) par l'onde électromagnétique le définissant et selon le vecteur de Poynting \vec{P} par une équation aux dérivées partielles.

Le vecteur de Poynting est donné par

$$\vec{P} = \frac{1}{\mu_0} \vec{E} \wedge \vec{B}$$

où \vec{E} et \vec{B} désignent le couple champ électromagnétique. Associé à \vec{E} et \vec{B} , la densité d'énergie est

$$u = \frac{\epsilon_0}{2} E^2 + \frac{1}{2\mu_0} B^2$$

qui est transportée par l'équation aux dérivées partielles

$$\frac{\partial u}{\partial t} = -\text{div}(\vec{P}).$$

1.3 Statique des poutres

Dans ce paragraphe, nous considérons un modèle simple de poutre. Les hypothèses sont les suivantes:

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
 Bureau 17.0.14

Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
 Département d'Ingénierie Mathématique

Une poutre est caractérisée par une courbe moyenne paramétrée par son abscisse curviligne. L'abscisse curviligne d'une courbe $t \in [0, 1] \mapsto \gamma(t) \in \mathbb{R}^3$ est une quantité définie lorsque $\gamma'(t) \neq \vec{0}$ pour tout t . On définit alors cette abscisse par (à une constante près)

$$s(t) = \int_0^t \|\gamma'(u)\| du.$$

La représentation de la chaîne d'arpenteur permet de comprendre pourquoi cette quantité est appelée abscisse curviligne.

L'autre hypothèse faite sur une poutre déformée est que lors des déformations considérées les sections normales le restent après déformations.

Les efforts sur une poutre vont être considérées comme ayant une distribution le long de la ligne moyenne donnée par une densité \vec{p} de forces et une densité \vec{m} de moments. Les équations de la statique donnent alors en notant $\vec{R}(s)$ la résultante des forces et $\vec{M}(s)$ la résultante des moments à l'abscisse curviligne s (au point $P(s)$ de la ligne moyenne) et en considérant un petit tronçon de longueur ds de la poutre:

$$\begin{aligned} \vec{R}(s+ds) - \vec{R}(s) + \vec{p} ds &= \vec{0} \\ \vec{M}(s+ds) - \vec{M}(s) + \vec{m} ds + \overrightarrow{P(s)P(s+ds)} \wedge \vec{R}(s) &= \vec{0} \end{aligned}$$

ce qui donne

$$\begin{aligned} \vec{R}'(s) + \vec{p} &= \vec{0} \\ \vec{M}'(s) + \vec{m} + \vec{\tau} \wedge \vec{R}(s) &= \vec{0} \end{aligned}$$

Bien sûr tout ceci ne ferme pas le modèle et l'on doit compléter ces équations.

2 Cas scalaire

2.1 Existence et unicité

On considère $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ une fonction de deux variables. On suppose que f est continue. La plupart des ouvrages traitant des équations différentielles considèrent des fonctions f qui ne sont définies que sur des sous-ensembles de $\mathbb{R} \times \mathbb{R}$, cependant on peut tout à fait traiter la théorie des équations différentielles dans le cadre que j'ai choisi sans perte de généralités en remplaçant f par $f \times \psi$ où ψ est une fonction ad-hoc nulle loin de (t_0, x_0) (cf plus loin) et valant 1 près de (t_0, x_0) . Cette remarque étant faite, on ne se préoccupera plus de ce type de problème au sein de ce document (on peut tout à fait faire de même dans le cas vectoriel).

Etant donné (t_0, x_0) on cherche une fonction

$\phi : I \rightarrow \mathbb{R}$ telle que I est un intervalle ouvert contenant t_0 dérivable telle que

$$\forall t \in I, \phi'(t) = f(t, \phi(t)), \phi(t_0) = x_0. \quad (2.1)$$

Si une telle fonction existe, on dit que le couple (I, ϕ) est une solution du problème (2.1). Ce problème (2.1) s'appelle problème de Cauchy pour f de donnée initiale x_0 en t_0 .

Exemple: Un circuit électrique constituée de deux dipôles respectivement inductant et résistant et soumis à une intensité forcée $I_f(t)$ (au travers d'une impédance équivalente de la fem Z) voit circuler un courant électrique d'intensité $I(t)$ qui satisfait la relation

$$LI'(t) + RI(t) = ZI_f(t), \forall t \geq 0. \quad (2.2)$$

La relation (2.2) est appelée équation différentielle. On vérifiera que dans ce cas là $f(t, x) = I_f(t) - \frac{R}{L}x(t)$.

Associée à une condition $I(0) = I_0$ on obtient un problème de Cauchy.

L'existence d'un couple (I, ϕ) satisfaisant le problème de Cauchy (2.1) est grandement simplifiée dans la situation de l'équation différentielle (2.2). Pour ce type d'équation différentielle, on parle d'équation différentielle du premier ordre linéaire non homogène.

En général, l'existence d'une solution (I, ϕ) du problème (2.1) est un problème ardu.

Dés lors que l'on dispose d'une solution (I, ϕ) de ce problème, alors si I' est un intervalle ouvert contenant t_0 , on a une autre solution (I', ϕ) . La notion d'unicité est donc délicate également. On dispose d'hypothèses suffisantes qui permettent de garantir l'existence et l'unicité d'un couple solution (I, ϕ) où I est un intervalle le plus grand possible:

□ **Théorème 2.1.1** [Cauchy-Lipschitz]. On suppose que f est continue sur et qu'il existe une constante M telle que

$$\forall (t, x, x') \in \mathbb{R}^3, |f(t, x) - f(t, x')| \leq M|x - x'|, \quad (2.3)$$

alors le problème de Cauchy (2.1) possède une unique solution (\mathbb{R}, ϕ) .

Pour une preuve, voir les appendices.

Notre exemple où $f(t, x) = ZI_f(t) - \frac{R}{L}x(t)$ satisfait (2.3) avec $M = \frac{R}{L}$.

2.2 Approximation

Supposons maintenant que l'on dispose d'une solution (I, ϕ) du problème de Cauchy.

Les méthodes d'approximation courantes sont fondées sur le fait que si

$$\phi'(t) = f(t, \phi(t)), \quad \forall t \in I$$

alors on a

$$\forall (t, t') \in I^2, \phi(t') = \phi(t) + \int_t^{t'} f(s, \phi(s)) ds.$$

L'approximation de l'intégrale dans le membre de droite conduit à une méthode de calcul de $\phi(t')$ si l'on connaît $\phi(t)$.

Plusieurs possibilités s'offrent alors en prenant $t = t_0$ et $t' = t_0 + h$ où $h > 0$ est petit:

i. *Méthode explicite*: h est petit et on assimile l'intégrale à l'aire d'un rectangle de base h et de hauteur $f(t_0, \phi(t_0))$ et on obtient

$$\phi(t_0 + h) \sim \phi(t_0) + hf(t_0, \phi(t_0)).$$

Puis connaissant $\phi(t_0 + h)$ (de manière approchée) on recommence.

ii. Méthode implicite: h est petit et on assimile l'intégrale à l'aire d'un rectangle de base h et de hauteur $f(t_0 + h, \phi(t_0 + h))$ et on obtient

$$\phi(t_0 + h) \sim \phi(t_0) + hf(t_0 + h, \phi(t_0 + h)). \quad (2.4)$$

Le situation est plus complexe: connaissant $\phi(t_0 + h)$ on en déduit $\phi(t_0 + h)$, on parle alors de méthode implicite.

iii. Toute méthode d'approximation de l'intégrale $\int_t^{t'} f(s, \phi(s)) ds$ conduit à une méthode de calcul de ϕ qui en général est une combinaison de méthodes implicites et explicites. C'est le cas des méthodes de Runge-Kutta.

2.2.1 Exemple

Reprenons le cas de notre circuit inductant-résistant. Ce qui est connu est l'intensité initiale $I_0 = I(0)$ et l'intensité de forçage $I_f(t)$. On a, en remplaçant les approximations par des égalités:

- en schéma explicite $L(I(h) - I(0))/h + RI(0) = ZI_f(0)$,

- en schéma implicite $L(I(h) - I(0))/h + RI(h) = ZI_f(h)$.

En schéma explicite on obtient

$$I(h) = \left(1 - \frac{R}{L}h\right)I(0) + \frac{h}{L}ZI_f(0).$$

Supposons, pour simplifier que I_f est nulle, on obtient alors

$$I(h) = \left(1 - \frac{R}{L}h\right)I_0.$$

Si on réitère N fois cette expression on obtient

$$I(Nh) = \left(1 - \frac{R}{L}h\right)^N I_0,$$

et si l'on regarde à un temps T fixé, on lie T , N et h par la relation $T = Nh$, ce qui donne $h = \frac{T}{N}$ et finalement

$$I(T) = \left(1 - \frac{R}{L} \frac{T}{N}\right)^N I_0.$$

En prenant la limite quand $N \rightarrow +\infty$ on obtient

$$I(T) = e^{-T \frac{R}{L}} I_0. \quad (2.5)$$

Rappelons que $\left(1 + \frac{x}{n}\right)^n = \exp\left(n \ln\left(1 + \frac{x}{n}\right)\right)$ et donc $\left(1 + \frac{x}{n}\right)^n \rightarrow e^x$ quand $n \rightarrow +\infty$.

En schéma implicite on obtient (toujours en supposant un forçage nulle)

$$\left(\frac{L}{h} + R\right)I(h) = \frac{L}{h}I(0),$$

soit

$$I(h) = \frac{1}{1 + \frac{R}{L}h} I(0).$$

En fait dans ce cas, plus rien n'est implicite, mais c'est une simplification due à la forme de l'équation différentielle.

Toujours en liant T , N et h , on obtient

$$I(T) = \frac{1}{\left(1 + \frac{R}{L} \frac{T}{N}\right)^N} I(0),$$

et en prenant la limite $N \rightarrow +\infty$

on obtient

$$I(T) = \frac{1}{e^{T \frac{R}{L}}} I(0),$$

ce qui est exactement la même formule que (2.5) (ce qui est heureux).

L'avantage de la méthode implicite est que sans condition de taille sur R par rapport à L et h on a toujours

$$|I(h)| < |I(0)|$$

dans le cas d'un schéma implicite, si R et L sont des quantités > 0 (ce qui est naturel).

Cette observation fait du schéma implicite un meilleur outil (sous les conditions de signe ci-dessus) d'approximation dans ce cas là. Cependant cet exemple cache le fait que la relation (2.4) est en général compliquée à traiter. Le paragraphe suivant explique les difficultés liées à ce trait du schéma implicite.

2.2.2 Comment traiter le cas implicite ?

On voit que si l'on remplace \sim dans (2.4) alors on obtient

$$\phi(t_0 + h) = \phi(t_0) + hf(t_0 + h, \phi(t_0 + h)). \quad (2.6)$$

Si on appelle $X_1 = \phi(t_0 + h)$, on voit que

$$X_1 = g(X_1)$$

où g est la fonction donnée par $g(x) = \phi(t_0) + hf(t_0 + h, x)$.

On doit donc savoir résoudre l'équation

$$G(x) = 0, \quad \text{où } G(x) = x - g(x).$$

Plusieurs méthodes sont envisageables: propre à la dimension 1 c'est la méthode de dichotomie.

2.2.2.1 Méthode de dichotomie

On suppose que l'on a deux valeurs x_1 et x_2 tels que $G(x_1)G(x_2) < 0$, on regarde $G(x_1)G(\frac{x_1 + x_2}{2})$.

Si cette quantité est < 0 , on remplace x_2 par $\frac{x_1 + x_2}{2}$ et si cette quantité est > 0 , on remplace x_1 par $\frac{x_1 + x_2}{2}$.

On démontre que si G est continue cette méthode définit une suite qui converge vers un X tel que $G(X) = 0$.

Pour une méthode généralisable à une dimension plus grande:

2.2.2.2 Méthode de Newton

Il s'agit partant d'un x_1 de définir x_2 (s'il existe) par x_2 est l'intersection de la droite de pente $G'(x_1)$ passant par $(x_1, G(x_1))$ avec l'axe des x .

La condition pour que x_2 existe est que $G'(x_1) \neq 0$.

On a alors

$$x_2 = x_1 - \frac{G(x_1)}{G'(x_1)}.$$

Et on remplace x_1 par x_2 .

On définit ainsi une suite dont on espère qu'elle converge vers une solution de $G(X) = 0$.

La difficulté est qu'il faut calculer $G'(x_1)$. En général on remplace $G'(x_1)$ par une approximation de $G'(x_1)$ et on ne met pas à jour ce calcul.

Une autre difficulté est de généraliser cette suite à la dimension supérieure. Nous verrons ultérieurement.

2.3 Cas vectoriel

Cette fois-ci on prend $f : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$, $t_0 \in \mathbb{R}$ et $x_0 \in \mathbb{R}^N$, et le même résultat d'existence et d'unicité est vrai:

□ **Théorème 2.3.1** [Cauchy-Lischitz, cas vectoriel]. On suppose que f est continue et qu'il existe M tel que

$$\forall t \in \mathbb{R}, (x, x') \in \mathbb{R}^N, \|f(t, x) - f(t, x')\| \leq M\|x - x'\|,$$

il existe alors une unique solution (I, ϕ) où I est le plus grand possible du problème de Cauchy

$$\begin{cases} \forall t \in I, \phi'(t) = f(t, \phi(t)), \\ \phi(t_0) = x_0. \end{cases}$$

On a remplacé la valeur absolue par une norme $\|\cdot\|$, c'est à dire un objet permettant de mesurer la longueur des vecteurs.

Une norme doit vérifier pour tout vecteur X et X' .

$$\begin{cases} \|X + X'\| \leq \|X\| + \|X'\| \\ \|X\| \geq 0 \\ \|\alpha X\| = |\alpha| \|X\| \\ \|X\| = 0 \Rightarrow X = 0. \end{cases} \quad (2.7)$$

Par exemple si $X = (X_1, \dots, X_N)$ alors $\|X\| = \sqrt{\sum_{i=1}^N X_i^2}$ est une norme.

Comme dans le cas scalaire on peut envisager plusieurs méthodes d'approximation: explicite, implicite, ou toutes combinaisons de celles-ci.

2.3.1 Cas des systèmes linéaires constants

Cette situation va nous occuper pratiquement tout le semestre.

Le problème de Cauchy est le suivant:

On se donne une matrice A à coefficients réels de taille $N \times N$ et $X_0 \in \mathbb{R}^N$ et on cherche $X : [0, \infty[\rightarrow \mathbb{R}^N$ vérifiant

$$\begin{cases} \forall t \geq 0, X'(t) = AX(t), \\ X(0) = X_0. \end{cases} \quad (2.8)$$

Dans cette situation $f(t, X) = AX$.

Il s'agit de vérifier qu'il existe M tel que

$$\|AX - AX'\| \leq M\|X - X'\|.$$

Une telle inégalité est vraie quelle que soit la matrice et quelle que soit la norme utilisée sur les vecteurs.

Il y a donc existence et unicité d'une solution sur un intervalle le plus grand possible.

On démontre que $X(t) = \exp(tA)X_0$ où

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

2.3.1.1 Schéma d'Euler explicite

On a alors $X(h) \sim (Id + hA)X(0)$ et on réitère.

On obtient avec les mêmes notations (et en supposant que l'approximation devient une égalité) que pour le cas scalaire

$$X(T) = \left(I + \frac{TA}{N}\right)^N X(0).$$

2.3.1.2 Schéma d'Euler implicite

On a alors toujours en supposant que l'approximation devient une égalité

$$(I - hA)X(h) = X(0).$$

On voit donc qu'il est "nécessaire" d'avoir une méthode pour inverser la matrice

$$I - hA,$$

après avoir bien sûr vérifié que $I - hA$ est inversible.

3 Equations différentielles d'ordre plus grand

Cette fois-ci, on s'intéresse à une relation du type

$$\forall t \in I, \phi^{(N)}(t) = f(t, \phi(t), \phi'(t), \dots, \phi^{(N-1)}(t)). \quad (3.1)$$

L'inconnu est toujours le couple (I, ϕ) , mais ϕ est dérivable N fois sur I .

Considérons le vecteur $X(t) = \begin{pmatrix} \phi(t) \\ \phi'(t) \\ \dots \\ \phi^{(N-1)}(t) \end{pmatrix}$.

On a

$$X'(t) = \begin{pmatrix} \phi'(t) \\ \phi''(t) \\ \dots \\ f(t, \phi(t), \dots, \phi^{(N-1)}(t)) \end{pmatrix}$$

que l'on peut donc écrire sous la forme $X'(t) = F(t, X(t))$. On est donc ramené à une équation différentielle du premier ordre vectoriel.

Remarquons qu'une condition initiale en t_0 pour X définit et réciproquement une condition en t_0 pour

$$\phi(t_0), \phi'(t_0), \dots, \phi^{(N-1)}(t_0). \quad (3.2)$$

Il s'ensuit que le problème de Cauchy associé à l'équation (3.1) implique N données initiales fournies par (3.2)

4 Inversion de matrices

La méthode d'Euler implicite dans le cas matriciel nécessite d'inverser une matrice (en fait il faut plutôt résoudre des systèmes).

4.1 Résolution d'un système

Si l'on veut résoudre $AX = B$ (A étant une matrice carrée)

4.1.1 on peut tenter d'utiliser une méthode de Newton:

$X_{n+1} = X_n - A^{-1}(AX_n - B)$ qui nécessite d'avoir calculé A^{-1} , ce que finalement on cherche à faire. Pour palier cet inconvénient, on peut remplacer A^{-1} par une matrice qui l'approche.

Par exemple, si les coefficients diagonaux de A sont dominants on remplacera A^{-1} par la matrice diagonale dont les coefficients sont les inverses des coefficients diagonaux de A . Une telle situation envisageable est celle où en notant les coefficients de la matrice a_{ij} (i pour les lignes et j pour les colonnes), on a

$$\forall i \in \{1, \dots, N\}, |a_{ii}| > \sum_{j \neq i} |a_{ij}|.$$

Il est facile de voir qu'une telle matrice est inversible car si X vérifie $AX = 0$ en prenant i l'indice pour lequel la i -e coordonnées de X est en valeur absolue la plus grande possible, alors

$$a_{ii}x_i + \sum_{j \neq i} a_{ij}x_j = 0 \Rightarrow |a_{ii}| \leq \sum_{j \neq i} \frac{|x_j|}{|x_i|} |a_{ij}| \leq \sum_{j \neq i} |a_{ij}|,$$

prouvant que A est injective et comme c'est une matrice carrée, elle est bijective.

4.1.2 Comment faire sinon

Il y a bien sûr la méthode du pivot de Gauss, qui tout en la mettant en oeuvre permet de découvrir si la matrice A est inversible. Pour des effets d'erreurs d'arrondis, il est préférable de choisir

comme pivot à chaque étape, le plus grand des coefficients en valeur absolue. Cette permutation des lignes (suivant le choix précédent ou non) correspond à changer de base dans l'espace d'arrivée via une matrice P .

Cette méthode se résume en disant que

□ **Théorème 4.1.1** [Décomp. LU]. Si A est une matrice inversible, il existe une matrice P dite de permutation (donc inversible), une matrice telle que

$$A = PLU$$

où L est triangulaire inférieure, U triangulaire supérieure.

Une matrice de permutation est le produit de matrices de transpositions où on permute deux éléments de la base utilisée (correspondant à la permutation des lignes dans le choix d'un éventuel pivot).

Cette méthode est assez rapide, mais n'est pas optimale dans des situations où la matrice A a d'autres propriétés.

Par exemple si A est symétrique, on peut utiliser la

4.1.3 Méthode de Cholesky

Cette méthode fonctionne lorsque A est symétrique et définie positive (c'est à dire que toutes ses valeurs propres sont > 0).

On cherche à écrire $A = {}^tLL$ où L est triangulaire supérieure.

On voit que $L_{11}^2 = a_{11}$ ce qui est possible car $a_{11} = {}^t e_1 A(e_1)$. Choisissons $L_{11} > 0$.

On a ensuite $a_{12} = L_{11}L_{12}$ ce qui détermine L_{12} , puis $a_{1i} = L_{11}L_{1i}$ ce qui détermine L_{1i} . On poursuit coefficient par coefficient.

$a_{22} = L_{22}^2$, puis $a_{23} = L_{12}L_{13} + L_{22}L_{23}$ ce qui détermine L_{23} , etc.. La méthode montre qu'au choix du signe de a_{22} par exemple > 0 il y a unicité de la décomposition.

On obtient

¶**Théorème 4.1.2** Si A est une matrice symétrique définie positive, il existe une matrice triangulaire supérieure L telle que

$$A = {}^tLL,$$

de plus si on impose que les coefficients diagonaux de L sont > 0 (et on le peut) alors L est unique.

4.1.4 Méthode de relaxation

Elle(s) consiste(nt) à écrire $A = M - N$ où M est facilement inversible. Comme on l'a fait remarquer ci-dessus, en général on n'a pas besoin d'inverser A , on veut simplement (souvent) résoudre le système $Ax = b$.

On obtient ici $Mx = b + Nx$. Ce qui ici va se résoudre par

$$x = M^{-1}b + M^{-1}Nx \tag{4.1}$$

impliquant la nécessité de calculer l'inverse de M .

Si M a de bonnes propriétés (par exemple si M est la diagonale de A dans le cas d'une diagonale strictement dominante) on calculera effectivement l'inverse de M et on résoudra l'équation (4.1) par une méthode de point fixe.

5 Systèmes hamiltoniens

Si on se donne $H : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ une fonction assez régulière, on peut considérer le système différentiel

$$\begin{cases} p'(t) = -\frac{\partial H}{\partial q}(p(t), q(t)), \\ q'(t) = \frac{\partial H}{\partial p}(p(t), q(t)), \end{cases} \quad (5.1)$$

avec des conditions initiales $(p(0), q(0))$.

La fonction

$$t \mapsto H(p(t), q(t))$$

est alors constante.

Les méthodes numériques associées doivent donc être en accord avec une telle conservation.

Remarquons que tout système où une quantité est conservée n'est pas nécessairement hamiltonien. Considérons l'exemple du système suivant dit proies-prédateurs.

$$\begin{cases} x'(t) = x(t)(1 - y(t)) \\ y'(t) = y(t)(x(t) - 1) \end{cases}$$

où $x(t)$ désigne la quantité de proies et $y(t)$ désigne la quantité de prédateurs, le tout à l'instant t .

Observons que

$$x'(t) \frac{x(t) - 1}{x(t)} = y'(t) \frac{1 - y(t)}{y(t)}$$

On a donc

$$x(t) - \ln(x(t)) + y(t) - \ln(y(t)) = cte$$

Le hamiltonien devrait ici vérifier

$$\frac{\partial H}{\partial y} = (y - 1)x, \quad \frac{\partial H}{\partial x} = y(x - 1).$$

Ce qui bien sûr n'est pas le cas.

Revenons aux systèmes hamiltoniens.

Supposons qu'on résolve de manière approchée la première équation

$$p(t + \tau) = p(t) - \tau \frac{\partial H}{\partial q}(p(t + \tau), q(t)).$$

On doit pour conserver l'hamiltonien trouver $q(t + \tau)$ pour que $H(p(t + \tau), q(t + \tau)) = H(p(t), q(t))$.

C'est à dire $H(p(t) - \tau \frac{\partial H}{\partial q}(p(t + \tau), q(t)), q(t + \tau)) = H(p(t), q(t))$. Si on écrit $q(t + \tau) = q(t) + \delta_q$ alors

$$H(p(t) - \tau \frac{\partial H}{\partial q}(p(t + \tau), q(t)), q(t) + \delta_q) = H(p(t), q(t))$$

donc si on fait un dl au premier ordre

$$-\tau \frac{\partial H}{\partial p}(p(t), q(t)) \cdot \frac{\partial H}{\partial q}(p(t + \tau), q(t)) + \frac{\partial H}{\partial q}(p(t), q(t)) \delta_q = 0 \text{ à l'ordre 1}$$

On est conduit naturellement à poser $\delta = \tau \frac{\partial H}{\partial p}(p(t + \tau), q(t))$.

On obtient donc une méthode implicite pour p et explicite pour q .

6 Quelques remarques sur le C++

6.1 Les constructeurs de classe

Il est bon d'en fournir un par défaut.

Exemple: dans une classe polynôme,

```
1 Polynome(){
2
3   coeff=null;
4   degre=0;
5 }
```

Un constructeur de copie ``Q est le polynôme P''

```
1 Polynome(const Polynome &P){
2   degre=P.degre;
3   coeff=new float(degre+1);
4   for(int i=0;i<degre+1;i++){
5     coeff(i)=P.coeff(i);
6   }
7 }
```

Un constructeur d'affectation

```
1 Polynome& operator=(const Polynome &P){
2   degre=P.degre;
3   coeff=new float(degre+1);
4   for(int i=0;i<degre+1;i++){
5     coeff(i)=P.coeff(i);
6   }
7 }
```

Un destructeur. Il est d'usage de le déclarer virtuel dans le ".h" lorsque il est amener à être spécialisé dans une classe dérivée.

```
1 ~Polynome(){
2     if(coeff){
3         coeff=NULL;
4     }
5     degre=0;
6 }
```

6.2 Les membres static

Certains membres de classes doivent pouvoir exister sans qu'un élément de la classe ait été instancié.

De tels objets prennent le qualificatif static (dans le ``.h").

```
1 static Polynome polynull;
```

et dans le cpp

```
1 Polynome polynull{
2     degre=-DBL_MAX; // conversion de double vers int
3     coeff=new float(1);
4     coeff(0)=1;
5 }
```

6.3 Les références

Exemple:

```
1 float& f(float x*) {  
2     return x(0);  
3 }
```

on introduit & qui signifie que & P est une référence à l'object P.

Ici si on écrit

```
1 float x(2)={0.1,0.2};  
2 float y=2.0;  
3 f(x)=y;  
4 std::cout<< x(0)<<std::endl;
```

donne 2.

Explication: f(x) retourne une référence vers x(0).

On utilisera les références dans les fonctions pour les types non primitifs pour éviter la copie d'objets volumineux.

```
1 Polynome derive(cons Polynome &P){
2     int degre_;
3     float *coeff_;
4     if(P.degre>=1){
5         degre_=P.degre-1;
6         coeff_=new float(P.degre);
7         for(int i=0;i<=degre_;i++){
8             coeff_(i)=(i+1)*P.coeff(i+1);
9         }
10    }
11    else
12    {
13        degre_=-DBL_MAX;
14        coeff_=NULL;
15    }
16    Polynome res(degre_,coeff_);
17    return res;
18 }
```

Cependant, dans la majorité des situations, une erreur pourrait modifier l'objet référencé par P dans la fonction. Pour éviter ceci, on qualifie la variable de CONST qui permet de signifier que l'objet référencé par P ne doit pas être modifié.

`P.derive()=Q`

dans le main doit donc, ici, donner une erreur à la compilation.

6.4 setter et getter

Si dans la classe Polynome coeff et degre sont déclarés private. Ceci implique que seules les méthodes de la classe Polynome et les méthodes amies (friend) peuvent accéder à ces variables. Si l'on veut pouvoir modifier ces quantités autrement, il faut proposer des méthodes de classes qui vont pouvoir modifier ces variables.

Exemple Une classe Oeil:

Dans Oeil.h

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14

Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

```
1 class Oeil {
2 private:
3     std::string couleur;
4     int correction;
5
6 public:
7     std::string getter1();
8     int getter2();
9
10 void setter1(std::string);
11 void setter2( int);
12 }
```

Dans Oeil.cpp

```
1 std::string getter1(){
2     return couleur;
3 void setter1(string chaine){
4     couleur=chaine;
5     ...
6 .
7 .
8 .
9 }
```

6.5 La surcharge des opérateurs

On souhaite donc surcharger des opérateurs tels que +,-,,/,+=",(),(),... quand cela prend sens.*

Exemple: ordre lexicographique sur \mathbb{R}^2 .

LeBipoint.h

```
1 Class Bipoint {  
2     private:  
3     float(2) coord;  
4     public:  
5     Bipoint(void);/*constructeur par défaut */  
6     Bipoint(float(2));/* On a un constructeur.*/  
7     ~Bipoint();/* Destructeur */  
8     .  
9     .  
10    .  
11    friend operator+(const Bipoint &,const Bipoint &);  
12    .  
13    .  
14    .  
15 }
```

LeBipoint.cpp

```
1 Bipoint operator*(const Bipoint & B1, const Bipoint & B2){  
2     float(2) coeff_;  
3     coeff_(0)=B1.coeff(0)+B2.coeff(0);  
4     coeff_(1)=B1.coeff(1)+B2.coeff(1);  
5     Bipoint res(coeff_);  
6     return res;  
7 }
```

La surcharge des opérateurs fait partie du polymorphisme. Le polymorphisme permet de décliner plusieurs versions d'une même fonction suivant le type des variables d'entrée.

6.6 Classe dérivée

On pourrait encore parler de sous-classe, mais pour des raisons structurelles, le terme dérivée est plus approprié.