Calcul des variations

G. Vinsard

7 février 2010

Sommaire

1 Le problème le plus simple de calcul des variations **2**
1.1 Rappel sur les extrema de fonctions réelles à valeurs réelles 2
1.2 Extrema de fonctions à plusieurs arguments réels et à valeurs réelles 3
1.2.1 Non adéquation des fonctions partielles 4
1.2.2 Géométrisation 5
1.2.3 Continuité dans un espace normé 5
1.2.4 Dérivée au sens de Fréchet dans un espace normé 6
1.2.5 La condition nécessaire de minimum 6
1.3 Traitement du problème le plus simple du calcul des variations 7
1.3.1 Transformation du problème de départ 7
1.3.2 Norme, continuité et différentiabilité 8
1.3.3 Le choix de la norme 9
1.3.4 Inégalité de Poincaré 9
1.3.5 Différentiation de fonctionnelle 10
1.3.6 Condition nécessaire de minimum 10
1.3.7 Équation d'Euler-Lagrange 11
1.3.8 L'objection de du Bois-Reymond 11
1.3.9 Synthèse 12
1.4 Exercices 13

2 Extensions du problème le plus simple de calcul des variations **14**
2.1 Problème dans $C_1^N([0,T])$ 14
2.1.1 Exercices 15
2.2 Problèmes d'équations aux dérivées partielles 15
2.2.1 Transfert de chaleur stationnaire 15
2.2.2 Magnétostatique en domaine borné 15

3 Équation d'Hamilton-Jacobi **16**
3.1 Variation verticale 16
3.2 Variation horizontale 19
3.3 Équation d'Hamilton-Jacobi 20
3.4 Illustration et exercices 21

4 Problème de synthèse **23**

Références **23**

1 Le problème le plus simple de calcul des variations

$\mathcal{V}$ est l'ensemble des fonctions d'un argument réels, à valeurs réelles, définies sur le segment $[0,T]$ et dont les valeurs en 0 et T sont prescrites

$$\mathcal{V} = \{x \in C_1([0,T]) \text{ telle que } x(0) = 0; x(T) = X\} \tag{1}$$

où $C_1([0,T])$ représente les fonctions dérivables une fois dans l'intervalle $[0,T]$ et dont la dérivée est continue dans cet intervalle. L est une fonction à valeur réelle comportant trois argument réels (ou encore un seul argument qui est un vecteur à trois composantes réelles)

$$\begin{array}{rccl} L & : & \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & & (u,v,w) & \longrightarrow & L(u,v,w) \end{array} \tag{2}$$

qui est dérivable une fois par rapport à chacun de ses trois arguments. Il est possible de faire correspondre à chaque fonction x de $\mathcal{V}$ un nombre réel $\mathcal{S}(x)$ défini par

$$\mathcal{S}(x) = \int_0^T L\left(x(t), \dot{x}(t), t\right) \quad dt \tag{3}$$

où

$$\dot{x}(t) = \lim_{\delta t \longrightarrow 0} \frac{x(t+\delta t) - x(t)}{\delta t}$$

suivant la notation de Newton pour la dérivée. La notation de Leibnitz ($x'(T)$ plutôt que $\dot{x}(t)$) n'est pas utilisée et doit être proscrite ici ou le fait d'ajouter un prime à une fonction aura autre signification.

Une fonction à valeur réelle mais et dont l'argument est une fonction (et non pas une valeur, fut-ce la valeur d'un vecteur de plusieurs élements comme L)

$$\begin{array}{rccl} \mathcal{S} & : & \mathcal{V} & \longrightarrow & \mathbb{R} \\ & & x & \longrightarrow & \mathcal{S}(x) = \int_0^T L\left(x(t), \dot{x}(t), t\right) \quad dt \end{array} \tag{4}$$

vient d'être définie. Une telle fonction s'appelle une fonctionnelle et il s'agit maintenant d'introduire quelques éléments utiles pour manipuler les fonctionnelles.

Plus précisément, en accord avec le titre de ce texte, l'objectif final est de résoudre le problème (problème le plus simple du calcul des variations)

$$\text{trouver } x_\infty \in \mathcal{V} \text{ telle que } \forall x \in \mathcal{V} \text{ alors } \mathcal{S}(x_\infty) \leq \mathcal{S}(x) \tag{5}$$

mais autant prévenir un lecteur pressé, on n'épuisera pas le sujet ; loin de là !

1.1 Rappel sur les extrema de fonctions réelles à valeurs réelles

Soit une fonction réelle à valeur réelle f dérivable une fois ; le ou les nombre(s) réel(s) x_∞ dont on puisse dire que

$$\forall x \in \mathbb{R} : f(x_\infty) \leq f(x) \tag{6}$$

s'appellent les arguments minimisant de la fonction f ; la valeur $f(x_\infty)$ s'appelle le minimum de f.

Supposons qu'un ou plusieurs tel(s) nombre(s) x_∞ existe(nt) ; alors (6) peut être réécrite comme

$$\forall x' \in \mathbb{R} : f(x_\infty) \leq f(x_\infty + x') \tag{7}$$

où une simple translation $x = x_\infty + x'$ a été faite, quantificateur portant maintenant sur x'. En notant différemment x' par δx (le δ suivi d'un symbole étant la marque d'un nouveau symbole, dont la nature est la même que le symbole qu'il préfixe, qui est destiné à devenir petit sans toutefois s'annuler complètement) pour des raisons qui vont apparaître tout de suite, (7) peut être transformée en

$$\forall \delta x > 0 : \frac{f(x_\infty + \delta x) - f(x_\infty)}{\delta x} \geq 0 \quad \text{et} \quad \forall \delta x < 0 : \frac{f(x_\infty + \delta x) - f(x_\infty)}{\delta x} \leq 0 \tag{8}$$

Si f est dérivable, i.e. la limite suivante a un sens pour tout x

$$\frac{df}{dx}(x) = \lim_{\delta x \longrightarrow 0} \frac{f(x+\delta x) - f(x)}{\delta x} \tag{9}$$

alors le passage (9) à la limite peut être appliqué à (8) ; ce qui entraîne que

$$\lim_{\substack{\delta x \longrightarrow 0 \\ \delta x > 0}} \frac{f(x_\infty + \delta x) - f(x_\infty)}{\delta x} \geq 0 \quad \text{et} \quad \lim_{\substack{\delta x \longrightarrow 0 \\ \delta x < 0}} \frac{f(x_\infty + \delta x) - f(x_\infty)}{\delta x} \leq 0 \tag{10}$$

Si donc f est dérivable en x_∞ les limites à gauche et à droite de (10) coïncident et il vient

$$0 \leq \frac{df}{dx}(x_\infty) \leq 0 \Longrightarrow \frac{df}{dx}(x_\infty) = 0 \tag{11}$$

Sous réserve que la fonction f soit dérivable en tout point, si il existe un x_∞ satisfaisant à (6) alors ce point satisfait à (11).

Mais bien sûr il peut exister des points satisfaisant à (11) qui ne satisfont pas à (6) : ces points peuvent être

- autre chose que des arguments minimisant de la fonction f ;
  par exemple des arguments maximisant (0 pour $f(x) = -x^2$), ou encore des points d'inflexion (0 pour $f(x) = x^3$) ;
- des arguments minimisant f seulement localement ;
  par exemple si $f(x) = \int_0^x u(u-1)(u-3)du$ alors les x_∞ solution de (11) sont 0, 1 et 3 ; 0 est un argument minimisant de f si on limite l'étendu de recherche des minima à $x < (8-\sqrt{10})/3 - 1/100$ par exemple mais pas pour la droite réelle entière ; c'est 3 qui est l'argument minimisant dans ce dernier cas.

(11) est une condition nécessaire à (6) mais ce n'est pas une condition suffisante.

Condition suffisante Pour trouver une condition suffisante il faut pousser plus loin l'analyse mais ce ne sera pas fait dans le cadre de ce paragraphe introductif dont l'unique fonction est de rappeler le raisonnement qui mène à la condition nécessaire pour la minimisation d'une fonction réelle à valeur réelle dérivable une fois.

Sous-gradient D'autre part on a supposé que la fonction f était différentiable pour trouver la condition nécessaire (11) ; mais si on la suppose différentiable seulement par morceaux (i.e. la fonction est différentiable en général mais il existe un nombre fini de points pour lesquels elle ne l'est pas de façon bénigne : les limites à gauche et à droite de (8) existent mais elle ne coïncident pas) alors on peut encore trouver une condition nécessaire pour qu'un point de discontinuité de la dérivée soit un argument minimisant de la fonction en exprimant que la dérivée à gauche est négative alors que la dérivée à droite est positive.

Une continuation de cette analyse mène à l'idée de sous-gradient qui ne sera pas introduite ici.

1.2 Extrema de fonctions à plusieurs arguments réels et à valeurs réelles

Soit une fonction f à N arguments réels et à valeur réelle, on la note

$$\begin{array}{rrcl} f & : \mathbb{R}^N & \longrightarrow & \mathbb{R} \\ & (x_1, \ldots, x_N) & \longrightarrow & f(x_1, \ldots, x_N) \end{array} \tag{12}$$

l'objectif est de formuler une condition nécessaire analogue à (11) pour le problème de trouver des jeux de nombres $(x_1^\infty, \ldots, x_N^\infty)$ (la marque du minimum ∞ est placée en exposant pour ne pas la confondre avec les indices des variables : rien à voir avec une quelconque puissance) tels que

$$\forall (x_1, \ldots, x_N) \in \mathbb{R}^N : f(x_1^\infty, \ldots, x_N^\infty) \leq f(x_1, \ldots, x_N) \tag{13}$$

Mais au préalable il est nécessaire de discuter du sens à donner à la dérivabilité de f, ce qui va être fait tout de suite.

1.2.1 Non adéquation des fonctions partielles

À partir de la fonction (12) on peut définir les N fonctions partielles

$$\begin{array}{rcccl} f_n & : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & y & \longrightarrow & f(x_1, \ldots, x_{n-1}, y, x_{n+1}, \ldots, x_N) \end{array} \tag{14}$$

et déjà demander que ces fonctions partielles soient dérivables une fois en tous points, ce qui permet de définir la dérivée partielle de la fonction f par rapport à la variable située à la position n dans la liste d'arguments de (12) comme

$$\partial_n f(x_1, \ldots, x_N) = \frac{df_n}{dy}(y) = \lim_{\delta x_n \longrightarrow 0} \frac{f(x_1, \ldots, x_n + \delta x_n, \ldots, x_N) - f(x_1, \ldots, x_n, \ldots, x_N)}{\delta x_n} \tag{15}$$

Autant le reconnaître tout de suite, les fonctions partielles ne sont pas très utiles ; et leur considération peut même être source d'erreur.

Par exemple si la fonction f est telle que $f(x_1, x_2) = x_1 x_2$ les fonctions $f_1(y) = f(y, 0) = y \times 0 = 0$ et $f_2(y) = f(0, y) = 0 \times 0 = 0$ porteraient rapidement à penser que $f(x_1, x_2) = 0$; ce qui est faux.

Même si on évite ce premier piège (d'accord il est grossier), la question de la continuité introduit une nouvelle source d'erreur.

Par exemple si

$$f(x_1, x_2) = \begin{cases} \dfrac{x_1 x_2}{x_1^2 + x_2^2} & \text{si} \quad x_1 \neq 0 \text{ ou (inclusif) } x_2 \neq 0 \\ 0 & \text{si} \quad x_1 = x_2 = 0 \end{cases}$$

on voit que $f_1(y) = f_2(y) = 0$ d'où on supputerait rapidement que les fonctions partielles étant constantes elles sont aussi continue et donc que la fonction f devrait être continue ; ce qui est faux : $f(x, x) = 1/2$ pour $x \neq 0$ et $f(0, 0) = 0$.

Encore un piège qui est lié d'ailleurs au manque d'une définition de la continuité pour les fonctions à plusieurs variables ; celle-ci pourrait être qu'une fonction (12) est continue au point $(0, \ldots, 0)$ si la fonction

$$\phi(t) = f(x_1 t, \ldots, x_N t) \tag{16}$$

est continue en $t = 0$ pour toutes valeurs de $x_1, \ldots, x_N$.

Alors ainsi on pourrait étendre cette définition à la dérivabilité. La fonction (12) serait dérivable au point $(0, \ldots, 0)$ si la fonction (16) est dérivable en $t = 0$ pour toutes valeurs de $x_1, \ldots, x_N$.

Cette définition existe, c'est la dérivée directionnelle ou de Gâteaux et elle contient (jamais deux sans trois) encore un piège. Par exemple pour

$$f(x_1, x_2) = \begin{cases} 1 & \text{si} \quad 0 < x_2 < x_1^2 \\ 0 & \text{si} \quad x_2 \leq 0 \text{ ou } x_1^1 \leq x_2 \end{cases} \tag{17}$$

la fonction (16) est $\phi(t) = 0$ qui est continue et même dérivable aux sens précédents.

Et pourtant la fonction $\psi(u) = f(u, u^2/2)$ n'est pas continue (elle vaut 1 partout sauf en $u = 0$ où elle vaut 0. C'est une drôle de continuité.

Si on enlève la définition de la continuité pour en trouver une autre qui fasse que cette fonction ne soit plus continue, elle reste cependant dérivable au sens de Gâteaux et on trouverait une fonction non continue mais cependant dérivable.

Le moins qu'on puisse penser est que le terrain est miné. On a donc le choix entre

1. rester dans le champ de mine et tenter de définir des chemins de passage en raffinant un peu les définitions ;
2. sortir du champ de mine et le contourner.

C'est la seconde solution qui est préférée dans ce petit cours.

1.2.2 Géométrisation

Les problèmes évoqués dans le paragraphe précédent sont tous liés à l'introduction des fonctions partielles et leurs épigones que sont les dérivées partielles.

C'est cela qu'il s'agit d'éviter et pour l'éviter on peut grouper les N variables $x_1, \ldots, x_N$ réelles en une seule dont la nature sera d'être vectorielle. Pour cela on introduisant le vecteur

$$\vec{x} = x_1\vec{e}_1 + \ldots + x_N\vec{e}_N \tag{18}$$

d'un espace vectoriel E_N à N dimensions dont les x_n sont les coordonnées dans la base $\vec{e}_1, \ldots, \vec{e}_N$.

Plutôt que d'introduire une fonction de N variables comme (12) on introduit une fonction de l'espace E_N comme

$$\begin{array}{cccc} F & : & E_N & \longrightarrow & \mathbb{R} \\ & & \vec{x} & \longrightarrow & F(\vec{x}) \end{array} \tag{19}$$

et la correspondance entre les deux fonctions est que

$$f(x_1, \ldots, x_n, \ldots, x_N) = F(x_1\vec{e}_1 + \ldots + x_N\vec{e}_N) \tag{20}$$

Si une fonction $F(\vec{x})$ est donnée on trouve une fonction $f(x_1, \ldots, x_N)$ par (20) ; ce qui n'est possible que dès lors qu'une base particulière est spécifiée.

Un changement de base change également la correspondance (20). Ce qu'on veut dire par là est qu'un vecteur particulier d'un espace vectoriel E_N ne dépend pas de la base choisie dans cet espace ; aussi si on donne une fonction comme (19), il lui correspond un très grand nombre de fonctions comme (12), autant que de bases possibles dans E_N.

Inversement si on donne une fonction comme (12), par (20) il lui correspond un grand nombre de fonctions comme (15), autant que de bases possibles dans E_N.

Ce dernier inconvénient, que Postnikov avoue franchement ne pas savoir réduire [?, p 39], empêche d'algébriser complètement la géométrie (i.e. faire de la géométrie la géométrie analytique) et inversement aussi de géométriser complètement l'algèbre.

Toutefois si on renonce à jamais changer de base, il est facile d'oublier l'inconvénient et de raisonner dans l'espace E_N plutôt que dans R^N. Et c'est ce qui va être fait.

La question est d'introduire la continuité puis la dérivation de la fonction F par rapport au vecteur $\vec{x}$. Pour cela il faut d'abord munir l'espace vectoriel E_N d'un moyen de mesurer les longueurs des vecteurs, c'est à dire d'une norme.

La norme est une application du type (19) qu'on note cependant un peu différemment

$$\begin{array}{cccc} ||\ || & : & E_N & \longrightarrow & \mathbb{R} \\ & & \vec{x} & \longrightarrow & ||\vec{x}|| \end{array} \tag{21}$$

et qui a pour propriété

$$\begin{array}{ll} ||\vec{x}|| \geq 0 & \text{et} \quad ||\vec{x}|| = 0 \iff \vec{x} = 0 \\ ||\lambda\vec{x}|| = |\lambda| ||\vec{x}|| & \\ ||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}|| & \end{array} \tag{22}$$

L'espace vectoriel E_N muni d'une norme est appelé un espace vectoriel normé ou plus simple un espace normé et sur les espaces normés il est possible d'introduire déjà une définition saine de la continuité puis de la dérivation de fonctions.

1.2.3 Continuité dans un espace normé

On peut reprendre la définition formelle de la continuité, i.e. F est continue au point $\vec{x}$ si

$$\forall \epsilon > 0, \exists \nu > 0 \text{ tel que } ||\vec{x} - \vec{y}|| < \nu \Longrightarrow |F(\vec{x}) - F(\vec{y})| < \epsilon \tag{23}$$

Cela règle la question de la continuité de la fonction (17) : elle n'est pas continue en $\vec{x} = 0\vec{e}_1 + 0\vec{e}_2 = \vec{0}$.

Pour le montrer il suffit de choisir $\epsilon < 1$; alors pour tout ν donné le vecteur $\vec{y} = 2(\sqrt{1 + \nu/2} - 1)\vec{e}_1/||\vec{e}_1|| + (\sqrt{1 + \nu/2} - 1)^2\vec{e}_2/||\vec{e}_2||$ est tel que $F(\vec{y}) = 1$ et on a aussi $||\vec{y} - \vec{x}|| = ||\vec{y}|| \leq 2(\sqrt{1 + \nu/2} - 1) + (\sqrt{1 + \nu/2} - 1)^2 = \nu/2 < \nu$ d'où il s'ensuit que l'implication est fausse.

On n'a pas eu besoin d'expliciter la norme pour obtenir ce résultat ; la définition des propriétés (22) de celle-ci a suffit.

1.2.4 Dérivée au sens de Fréchet dans un espace normé

Une fonction (19) dérivable au sens de Fréchet est une fonction qui peut être écrite

$$F(\vec{x}+\delta\vec{x}) = F(\vec{x}) + \nabla F(\vec{x})(\delta\vec{x}) + o(||\delta\vec{x}||) \tag{24}$$

où

$$\begin{array}{rcccl} \nabla F(\vec{x}) & : & E_N & \longrightarrow & \mathbb{R} \\ & & \vec{y} & \longrightarrow & \nabla F(\vec{x})(\vec{y}) \end{array} \tag{25}$$

est une application linéaire par rapport à son argument $\vec{y}$, i.e. telle que

$$\begin{array}{l} \nabla F(\vec{x})(\lambda\vec{y}) = \lambda\nabla F(\vec{x})(\vec{y}) \\ \nabla F(\vec{x})(\vec{y}+\vec{z}) = \nabla F(\vec{x})(\vec{y}) + \nabla F(\vec{x})(\vec{z}) \end{array} \tag{26}$$

et où

$$\begin{array}{rcccl} o & : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & u & \longrightarrow & o(u) \end{array} \tag{27}$$

est une fonction réelle à valeur réelle qui a la propriété

$$\lim_{u\longrightarrow 0} \frac{o(u)}{u} = 0 \tag{28}$$

Si on revient aux coordonnées par (18) l'application linéaire (25), comme toute application linéaire en dimension finie, est représentée par les N coefficients α_n de manière que

$$\nabla F(\vec{x})(y_1\vec{e}_1 + \ldots + y_N\vec{e}_N) = \alpha_1 y_1 + \ldots + \alpha_N y_N \tag{29}$$

En choisissant dans (24) un vecteur

$$\delta\vec{x} = \delta x_n \vec{e}_n \tag{30}$$

et en utilisant (20) cette formule (24) se transforme en

$$f(x_1,\ldots,x_n+\delta x_n,\ldots,x_N) = f(x_1,\ldots,x_n,\ldots,x_N) + \alpha_n\delta x_n + o(|\delta x_n|||\vec{e}_n||) \tag{31}$$

On peut alors écrire

$$\alpha_n = \frac{f(x_1,\ldots,x_n+\delta x_n,\ldots,x_N) - f(x_1,\ldots,x_n,\ldots,x_N)}{\delta x_n} + ||\vec{e}_n||\frac{o(|\delta x_n|||\vec{e}_n||)}{|\delta x_n|||\vec{e}_n||} \tag{32}$$

et par passage à la limite, prenant en compte que $||\vec{e}_n||$ est une quantité finie et la définition (15), on trouve

$$\alpha_n = \partial_n f(x_1,\ldots,x_N) \tag{33}$$

Le coefficient α_n de l'application linéaire (25) exprimée en coordonnées est la dérivée partielle de la fonction à plusieurs variable définie par (21) à partir de la fonction à argument vectoriel (19).

On rejoint le champ de mine précédent. Mais cette fois ci on dispose d'une poêle à frire ! En effet il est exclu que le coefficient α_n soit mal défini comme il le serait si la dérivée partielle $\partial_n f(x_1,\ldots,x_N)$ n'était pas continue au point $(x_1,\ldots,x_N)$.

Pour résumer, si une fonction (19) dont le domaine de définition est un espace normé est dérivable au sens de Fréchet en un point $\vec{x}$, i.e. on peut écrire (24), alors c'est que la fonction à plusieurs variables définie par (20) a ses dérivées partielles (15) continues.

1.2.5 La condition nécessaire de minimum

Il est maintenant possible de répéter le raisonnement déjà fait pour une fonction réelle d'un argument réel pour le problème qui consiste à chercher $\vec{x}_\infty$ tel que

$$\forall\vec{x}\in E_N : F(\vec{x}_\infty) \leq F(\vec{x}) \tag{34}$$

dans le cas où F est une fonction différentiable au sens de Fréchet (i.e. on peut écrire (24)).

D'abord on récrit (34) en faisant subir à $\vec{x}$ une translation : on cherche $\vec{x}_\infty$ tel que

$$\forall \delta\vec{x} \in E_N : F(\vec{x}_\infty) \leq F(\vec{x}_\infty + \delta\vec{x}) \tag{35}$$

Ensuite on exploite la différentiabilité de F en $\vec{x}$ ce qui permet d'écrire

$$\forall \delta\vec{x} \in E_N : F(\vec{x}_\infty) \leq F(\vec{x}_\infty) + \nabla F(\vec{x}_\infty)(\delta\vec{x}) + o(||\delta\vec{x}||) \tag{36}$$

soit donc

$$\forall \delta\vec{x} \in E_N : -o(||\delta\vec{x}||) \leq \nabla F(\vec{x}_\infty)(\delta\vec{x}) \tag{37}$$

Puis, comme (37) doit être vérifié pour tout $\delta\vec{x}$, elle doit l'être également pour son opposé alors, en divisant par $||\delta\vec{x}||$, on peut écrire encore

$$\forall \delta\vec{x} \in E_N : -\frac{o(||\delta\vec{x}||)}{||\delta\vec{x}||} \leq \frac{1}{||\delta\vec{x}||}\nabla F(\vec{x}_\infty)(\delta\vec{x}) = \nabla F(\vec{x}_\infty)\left(\frac{\delta\vec{x}}{||\delta\vec{x}||}\right) \leq \frac{o(||\delta\vec{x}||)}{||\delta\vec{x}||} \tag{38}$$

Et finalement, en utilisant la définition (28) qui est demandée pour la fonction o, on obtient par passage à la limite

$$\forall \delta\vec{x} \in E_N : 0 \leq \lim_{||\delta\vec{x}|| \longrightarrow 0} \nabla F(\vec{x}_\infty)\left(\frac{\delta\vec{x}}{||\delta\vec{x}||}\right) \leq 0 \tag{39}$$

Or le vecteur $\delta\vec{x}/||\delta\vec{x}||$ est de norme 1 ; cette condition (39) s'écrit encore

$$\forall \vec{x}' \in E_N \text{ tel que } ||\vec{x}'|| = 1 : \nabla F(\vec{x}_\infty)(\vec{x}') = 0 \tag{40}$$

c'est à dire que l'application linéaire (25) doit être à valeur nulle pour tout argument dont la norme est 1.

Si l'application linéaire n'était pas nulle alors il existerait $\vec{y} \neq 0$ tel que $\nabla F(\vec{x}_\infty)(\vec{y}) \neq 0$; comme l'application est linéaire on aurait alors $\nabla F(\vec{x}_\infty)(\vec{y}/||\vec{y}||) \neq 0$, ce qui est contradictoire avec (40).

Il faut donc pour que $\vec{x}_\infty$ soit une solution de (34) que

$$\nabla F(\vec{x}_\infty) = 0 \Longleftrightarrow \forall n \in \{1, \ldots, N\} : \partial_n f(x_1^\infty, \ldots, x_N^\infty) = 0 \tag{41}$$

Condition suffisante La condition (41) n'est que nécessaire. Pour trouver une condition suffisante il faudrait pousser un peu plus loin l'analyse, mais ce ne sera pas fait pour les mêmes raisons que dans le paragraphe de la recherche des conditions nécessaires des fonctions réelles à variable réelle.

1.3 Traitement du problème le plus simple du calcul des variations

Les problèmes (5) et (13) diffèrent de ce qu'une fonction à plusieurs variables ne comporte qu'un nombre fini d'arguments alors qu'une fonctionnelle en comporte une infinité : autant de que de points dans le segment $]0, T[$.

Mais si on oublie cette différence, et qu'on accepte donc de traiter le problème (5) sans trop de précautions, on peut faire immédiatement un traitement analogue au traitement du problème (36). Et c'est ce qu'on va faire dans ce paragraphe.

1.3.1 Transformation du problème de départ

Tout d'abord l'espace $\mathcal{V}$ (1) n'est pas un espace vectoriel. Par contre l'espace

$$\mathcal{V}_0 = \{x \in C_1([0, T]) \text{ telle que } x(0) = 0; x(T) = 0\} \tag{42}$$

en est un. On peut alors considérer tout élément de $\mathcal{V}$ comme la superposition

$$x \in \mathcal{V} \Longleftrightarrow x(t) = \frac{Xt}{T} + y(t) \text{ où } y \in \mathcal{V}_0 \tag{43}$$

À ce changement de variable correspond un changement de fonctionnelle comme

$$\mathcal{J}(y) = \int_0^T L\left(y(t) + \frac{Xt}{T}, \dot{y}(t) + \frac{X}{T}, t\right) \quad dt \tag{44}$$

de manière que

$$\mathcal{S}(x) = \mathcal{J}\left(y - \mathcal{F}\left(\frac{Xt}{T}\right), \dot{y}(t), t\right) \tag{45}$$

où on note $\mathcal{F}(f(t)) = f$ pour pouvoir abstraire une expression de sa variable et en faire ainsi une fonction. Par exemple si x^2 est une expression $\mathcal{F}(x^2)$ est le nom de la fonction $\begin{array}{rcl} : \mathbb{R} & \longrightarrow & \mathbb{R} \\ u & \longrightarrow & u^2 \end{array}$: on a $\mathcal{F}(x^2)(u) = u^2$.

Pour simplifier ce qui suit on peut encore introduire la nouvelle fonction

$$L\left(u + \frac{Xw}{T}, v + \frac{X}{T}, w\right) = \rho(u, v, w) \tag{46}$$

et écrire alors

$$\mathcal{J}(y) = \int_0^T \rho(y(t), \dot{y}(t), t)dt \tag{47}$$

Le problème (5) se transforme alors en

$$\text{trouver } y_\infty \in \mathcal{V}_0 \text{ telle que } \forall y \in \mathcal{V}_0 \text{ on a } \mathcal{J}(y_\infty) \leq \mathcal{J}(y) \tag{48}$$

et y_∞ connue il est possible de trouver x_∞ par (43).

L'intérêt de cette transformation est bien sûr de pouvoir travailler dans l'espace vectoriel $\mathcal{V}_0$ plutôt que dans l'espace (on a envie de penser affine et c'est cela) $\mathcal{V}$.

1.3.2 Norme, continuité et différentiabilité

Une norme sur le modèle de (21) peut être formellement décrite, soit

$$\begin{array}{rrcl} || \; || & : \mathcal{V}_0 & \longrightarrow & \mathbb{R} \\ & y & \longrightarrow & ||y|| \end{array} \tag{49}$$

qui a pour propriété

$$\begin{array}{ll} ||x|| \geq 0 & \text{et} \quad ||x|| = 0 \iff x = 0 \\ ||\lambda x|| = |\lambda| ||x|| & \\ ||x + y|| \leq ||x|| + ||y|| & \end{array} \tag{50}$$

On introduit ensuite une définition de la continuité sur le modèle de (23) : $\mathcal{J}$ est continue en x si

$$\forall \epsilon > 0, \exists \nu > 0 \text{ tel que } ||x - y|| < \nu \Longrightarrow |\mathcal{J}(x) - \mathcal{J}(y)| < \epsilon \tag{51}$$

Et enfin on introduit la dérivée sur le modèle de (24) : une fonctionnelle $\mathcal{J}$ est dérivable au sens de Fréchet si

$$\mathcal{J}(y + \delta y) = \mathcal{J}(y) + \nabla \mathcal{J}(y)(\delta y) + o(||\delta y||) \tag{52}$$

où

$$\begin{array}{rrcl} \nabla \mathcal{J}(y) & : \nu_0 & \longrightarrow & \mathbb{R} \\ & z & \longrightarrow & \nabla \mathcal{J}(y)(z) \end{array} \tag{53}$$

est une fonction linéaire de son argument z, i.e. telle que

$$\begin{array}{l} \nabla \mathcal{J}(y)(\lambda z) = \lambda \nabla \mathcal{J}(y)(z) \\ \nabla \mathcal{J}(y)(z + u) = \nabla \mathcal{J}(y)(z) + \nabla \mathcal{J}(y)(u) \end{array} \tag{54}$$

et où o a la propriété (28).

1.3.3 Le choix de la norme

Une propriété intéressante de la norme en dimension finie était que toutes les normes y sont équivalentes.

Cela signifie que si on donne 2 normes différentes $|| \ ||$ et $|| \ ||'$ alors on peut trouver deux nombre α et β strictement positifs tels que pour y on ait

$$\alpha||y|| \leq ||y||' \leq \beta||y|| \tag{55}$$

Et une utilisation de cette propriété est que la notion de limite ne dépend pas alors de la norme utilisée. Si par exemple on introduit une suite y_n alors

$$\lim_{n \longrightarrow \infty} y_n = 0 \tag{56}$$

se traduit en utilisant la norme $|| \ ||$ par

$$\forall \epsilon > 0, \exists N > 0 \text{ tel que } n > N \Longrightarrow ||y_n|| < \epsilon \tag{57}$$

et donc

$$\forall \epsilon > 0, \exists N > 0 \text{ tel que } n > N \Longrightarrow ||y_n|| < \epsilon \Longrightarrow ||y_n||' < \epsilon/\beta \tag{58}$$

qui s'écrit encore

$$\forall \epsilon' > 0, \exists N' > 0 \text{ tel que } n > N' \Longrightarrow ||y_n||' < \epsilon' \tag{59}$$

c'est à dire que la suite y_n converge aussi en utilisant la norme $|| \ ||'$.

En dimension finie, on peut donc se dispenser de préciser quelle norme est utilisée ; puisque toutes conduisent aux même résultat.

Ce n'est plus vrai en dimension infinie, comme le montre la comparaison de ces deux normes dans $\mathcal{V}_0$:

$$||y|| = \sqrt{\int_0^T (y(t))^2 \quad dt} \quad \text{et} \quad ||y||' = \sqrt{\int_0^T \left(\frac{dy}{dt}(t)\right)^2 \quad dt} \tag{60}$$

($|| \ ||'$ est bien une norme parce que si la dérivée de y est nulle dans $[0, T]$ alors y est une constante, mais dans $\mathcal{V}_0$ cette constante ne peut qu'être nulle puisqu'on a $y(0) = y(T) = 0$.) sur la fonction de $\mathcal{V}_0$

$$y_N(t) = \frac{1}{N} \sin\left(N\pi t/T\right)) \tag{61}$$

On a

$$||y_N|| = \frac{T}{2N^2} \quad \text{et} \quad ||y_N||' = \frac{2\pi^2}{T} \tag{62}$$

d'où on retire que les fonctions $\lim_{N \longrightarrow \infty} y_n$ sont bien différentes en utilisant l'une ou l'autre de ces deux normes : pour $|| \ ||$ c'est la fonction nulle, pour $|| \ ||'$ on n'en sait rien.

En dimension infinie, on ne peut pas se dispenser de préciser quelle norme est utilisée ; elles ne conduisent pas toutes au même résultat.

1.3.4 Inégalité de Poincaré

Le résultat selon lequel l'étude des espaces de fonctions suppose dépend de la norme qu'on utilise est bien ennuyeux ; mais c'est le problème du grand chapitre de l'analyse fonctionnelle et pas vraiment celui de ce petit paragraphe que constitue le calcul des variations.

Pour sortir de ce mauvais lieu, on dispose en effet de l'inégalité de Poincaré qui permet de trier les normes parmi celles qui peuvent être utiles pour le calcul des variations.

Cette inégalité s'écrit

$$\mathcal{V}_0 \text{ donné : } \exists \beta > 0 \text{ tel que } \forall y \in \mathcal{V}_0, \int_0^T (y(t))^2 \quad dt \leq \beta \int_0^T \left(\frac{dy}{dt}(t)\right)^2 \quad dt \tag{63}$$

et elle peut se démontrer en décomposant les éléments de $\mathcal{V}_0$ en série de Fourier (après avoir montré que c'était possible et donc fait de l'analyse fonctionnelle) ce qu'on n'a pas l'intention de faire ici puisque cette inégalité sera démontrée ailleurs et dans un cadre plus général.

On voit ainsi que la norme la plus "dure" entre $|| \ ||$ et $|| \ ||'$ (60) est bien $|| \ ||'$ comme le portait à penser l'exemple (61).

1.3.5 Différentiation de fonctionnelle

On a maintenant tous les éléments utiles pour calculer effectivement $\nabla\mathcal{J}(y)$ quand $\mathcal{J}$ est définie par (47).

D'abord, on utilise la norme $|| \ ||'$ de (60) ce qui permet d'affirmer que

$$\forall \delta y \in \mathcal{V}_0 : ||\delta y||' \longrightarrow 0 \Longrightarrow \forall t \in [0,T], \delta y(t) \longrightarrow 0 \text{ et } \frac{d\delta y}{dt}(t) \longrightarrow 0 \tag{64}$$

ce qui ne serait pas vrai avec la norme $|| \ ||$ comme le montre l'exemple (61).

On peut maintenant chercher à calculer explicitement

$$\mathcal{J}(y+\delta y) = \int_0^T \rho(y(t)+\delta y(t), \dot{y}(t)+\delta\dot{y}(t), t) \quad dt \tag{65}$$

en exploitant le fait que la fonction à trois variables ρ est supposé différentiable (i.e. elle admet trois dérivées partielles continues par rapport à ses trois arguments), car en effet avec (64) on est assuré que $\delta\dot{y}(t)$ est petit.

On trouve alors

$$\begin{aligned} \mathcal{J}(y+\delta y) \quad = \quad \int_0^T \quad & \rho(y(t), \dot{y}(t), t) \\ & + \partial_1\rho(y(t), \dot{y}(t), t)\delta y(t) + \partial_2\rho(y(t), \dot{y}(t), t)\delta\dot{y}(t) \\ & + o(||(\delta y(t), \delta\dot{y}(t))||_{E_2}) \quad dt \end{aligned} \tag{66}$$

où on note

$$||(u,v)||_{E_2} \tag{67}$$

la norme du couple (u,v) figurant un vecteur dans E_2.

Il reste encore à affirmer qu'on peut trouver une fonction o' telle que

$$o'(||\delta y||') = \int_0^T o(||(\delta y(t), \delta\dot{y}(t))||_{E_2}) \quad dt \tag{68}$$

pour que (66) se trouve mis sous la forme (52) avec

$$\nabla\mathcal{J}(y)(\delta y) = \int_0^T \partial_1\rho(y(t), \dot{y}(t), t)\delta y(t) + \partial_2\rho(y(t), \dot{y}(t), t)\delta\dot{y}(t) \quad dt \tag{69}$$

et que la différentielle $\nabla\mathcal{J}(y)$ au point y de $\mathcal{J}$ soit ainsi spécifiée.

1.3.6 Condition nécessaire de minimum

Pour obtenir une condition nécessaire du problème (48), on peut, *mutatis mutendis*, répéter le texte qui s'étend de (35) à (40) d'où on déduit que nécessairement

$$\nabla\mathcal{J}(y_\infty) = 0 \tag{70}$$

Ce qui se traduit par

$$\forall y' \in \mathcal{V}_0 : \int_0^T \partial_1\rho(y_\infty(t), \dot{y}_\infty(t), t)y'(t) + \partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)\dot{y}'(t) \quad dt = 0 \tag{71}$$

Cette expression (71) est extrêmement importante en pratique : elle est la clé de ce qu'on appelle la méthode des éléments finis. On l'appelle la *forme faible* ou la *forme variationnelle* du problème d'extrémisation (48). Et elle permet de faire des calculs numériques dans les cas (fréquents) où on ne connaît pas de solution analytique au problème (48).

Mais il est d'usage de transformer un peu (71) pour obtenir des équation différentielles ordinaires qu'on peut espérer savoir intégrer analytiquement ou encore traiter numériquement autrement qu'avec la forme faible.

1.3.7 Équation d'Euler-Lagrange

Si maintenant on croit que la fonction

$$t \longrightarrow \partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t) \tag{72}$$

est dérivable, alors on peut transformer (71) avec une intégration par partie

$$\begin{aligned} \int_0^T \partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)\dot{y}'(t) \quad dt &= \quad [y'(t)\partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)]_0^T \\ &- \int_0^T y'(t)\frac{d}{dt}\left(\partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)\right) \quad dt \end{aligned} \tag{73}$$

mais comme $y'(0) = y'(T) = 0$ on obtient finalement

$$\int_0^T \partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)\dot{y}'(t) \quad dt = -\int_0^T y'(t)\frac{d}{dt}\left(\partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)\right) \quad dt \tag{74}$$

Et donc au total (71) devient

$$\forall y' \in \mathcal{V}_0 : \int_0^T \left(\partial_1\rho(y_\infty(t), \dot{y}_\infty(t), t) - \frac{d}{dt}\left(\partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)\right)\right) y'(t) \quad dt = 0 \tag{75}$$

qui a le gros avantage de ne plus comporter de dérivées de y'.

Si on suppose qu'il existe un intervalle $[a, b] \subset [0, T]$ et une constante $c > 0$ dans lequel

$$\partial_1\rho(y_\infty(t), \dot{y}_\infty(t), t) - \frac{d}{dt}\left(\partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)\right) \geq c \tag{76}$$

il suffit de choisir

$$y'(t) = \begin{cases} (x-a)^2(x-b)^2 & \text{si} \qquad a < x < b \\ 0 & \text{sinon} \end{cases} \in \mathcal{V}_0 \tag{77}$$

pour trouver qu'alors

$$\int_0^T \left(\partial_1\rho(y_\infty(t), \dot{y}_\infty(t), t) - \frac{d}{dt}\left(\partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)\right)\right) y'(t) \quad dt \geq \frac{c(b-a)^5}{30} \tag{78}$$

ce qui est incompatible avec (75) si $b > a$ et $c > 0$.

On peut faire la même chose pour d'éventuelles valeurs négatives de (76) et on est donc amené à conclure que nécessairement en *presque tous les t*

$$\partial_1\rho(y_\infty(t), \dot{y}_\infty(t), t) - \frac{d}{dt}\left(\partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)\right) = 0 \tag{79}$$

ce qui constitue la *forme forte* ou l'*équation d'Euler-Lagrange* du problème (48).

1.3.8 L'objection de du Bois-Reymond

Cette objection tient à l'hypothèse de départ du paragraphe précédent selon laquelle on suppose que (72) est dérivable.

C'est bien une hypothèse parce que même si on suppose ρ comme différentiable autant de fois qu'on le veut, il n'en reste pas moins qu'on a demandé au départ à y_∞ seulement d'être dérivable (avec dérivée continue) une fois.

Du Bois-Reymond propose alors de ne pas faire cette supposition et de plutôt faire une intégration par partie de l'autre terme de (71), ce qui ne pose pas de problèmes puisque y' est dérivable.

On arrive alors à écrire

$$\begin{aligned} \int_0^T \partial_1\rho(y_\infty(t), \dot{y}_\infty(t), t)y'(t) \quad dt &= \left[y'(t)\int_0^t \partial_1\rho(y_\infty(u), \dot{y}_\infty(u), u) \quad du\right]_0^T \\ &- \int_0^T \dot{y}'(t)\int_0^t \partial_1\rho(y_\infty(u), \dot{y}_\infty(u), u) \quad du \quad dt \end{aligned} \tag{80}$$

soit, compte tenu de ce que $y'(0) = y'(T) = 0$

$$\int_0^T \partial_1\rho(y_\infty(t), \dot{y}_\infty(t), t)y'(t) \quad dt = -\int_0^T \dot{y}'(t) \int_0^t \partial_1\rho(y_\infty(u), \dot{y}_\infty(u), u) \quad du \quad dt \tag{81}$$

et alors (71) peut se mettre sous la forme

$$\forall y' \in \mathcal{V}_0 : \int_0^T \left(-\int_0^t \partial_1\rho(y_\infty(u), \dot{y}_\infty(u), u) \quad du + \partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t) \right) \dot{y}'(t) \quad dt = 0 \tag{82}$$

Il ne reste plus qu'à introduire le lemme *ad hoc* de Paul du Bois-Reymond selon lequel

$$\begin{array}{lll} \text{si} & f \text{ est continue} \quad \text{et} \quad \forall y' \in \mathcal{V}_0 : \displaystyle\int_0^T f(t)\dot{y}'(t) \quad dt = 0 \\ \text{alors} & \qquad\qquad f \text{ est constante } sur[0, T] \end{array} \tag{83}$$

dont la démonstration ne sera pas faite ici, pour voir que (82) conduit à

$$-\int_0^t \partial_1\rho(y_\infty(u), \dot{y}_\infty(u), u) \quad du + \partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t) = \text{Constante} \tag{84}$$

qui est une forme intégrale de l'équation d'Euler-Lagrange (79).

1.3.9 Synthèse

De nombreuses choses ont été évoquées depuis la page **??**. Et il n'est peut-être pas inutile de rappeler ici les résultats importants.

Tout d'abord on a le problème (5)

$$\text{trouver } x_\infty \in \mathcal{V} \text{ telle que } \forall x \in \mathcal{V} \text{ on a } \mathcal{S}(x_\infty) \leq \mathcal{S}(x)$$

dont on propose de chercher une condition nécessaire (toute analyse de condition suffisante est proscrite).

Pour travailler dans un espace vectoriel, on transforme le problème (5) en le problème (48)

$$\text{trouver } y_\infty \in \mathcal{V}_0 \text{ telle que } \forall y \in \mathcal{V}_0 \text{ on a } \mathcal{J}(y_\infty) \leq \mathcal{J}(y)$$

où on a posé (42), (43), (44), (46)

$$\mathcal{V}_0 = \{x \in C_1([0, T]) \text{ telle que } x(0) = 0; x(T) = 0\}$$

$$x \in \mathcal{V} \Longleftrightarrow x(t) = \frac{Xt}{T} + y(t) \text{ avec } y \in \mathcal{V}_0$$

$$\mathcal{J}(y) = \int_0^T \rho\left(y(t), \dot{y}(t), t\right) \quad dt$$

$$\rho(u, v, w) = L\left(u + \frac{Xw}{T}, v + \frac{X}{T}, w\right)$$

L'analyse menée aboutit aux différentes formes de la condition nécessaire qui sont : la forme faible (ou variationnelle) (71)

$$\forall y' \in \mathcal{V}_0 : \int_0^T \partial_1\rho(y_\infty(t), \dot{y}_\infty(t), t)y'(t) + \partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)\dot{y}'(t) \quad dt = 0$$

puis la forme forte (ou d'Euler-Lagrange) (79)

$$\partial_1\rho(y_\infty(t), \dot{y}_\infty(t), t) - \frac{d}{dt}\left(\partial_2\rho(y_\infty(t), \dot{y}_\infty(t), t)\right) = 0$$

ou, si on veut prendre en compte l'objection de du Bois-Reymond, la forme intégrale (84) de cette forme forte

$$-\int_0^t \partial_1\rho(y_\infty(u),\dot{y}_\infty(u),u) \quad du + \partial_2\rho(y_\infty(t),\dot{y}_\infty(t),t) = \text{Constante}$$

Il suffit alors de d'utiliser inversement (42), (43), (44), (46) sur (71) pour retrouver déjà la *forme faible* ou *forme variationnelle* portant sur x_∞ : nécessairement x_∞ est solution de

$$\begin{cases} x_\infty \in \mathcal{V} \text{ i.e. notamment } x_\infty(0) = 0 \text{ et } x_\infty(T) = X \\ \forall y' \in \mathcal{V}_0 : \displaystyle\int_0^T \partial_1 L(x_\infty(t),\dot{x}_\infty(t),t)y'(t) + \partial_2 L(x_\infty(t),\dot{x}_\infty(t),t)\dot{y}'(t) \quad dt = 0 \end{cases} \tag{85}$$

on notera que la variation y' est bien dans $\mathcal{V}_0$ et pas dans $\overline{\mathcal{V}_0}$.

On peut ensuite, toujours en utilisant (42), (43), (44), (46) mais maintenant sur (79) et (84) arriver à : l'équation d'Euler-Lagrange (ou forme forte)

$$\begin{cases} x_\infty \in \mathcal{V} \text{ i.e. notamment } x_\infty(0) = 0 \text{ et } x_\infty(T) = X \\ \partial_1 L(x_\infty(t),\dot{x}_\infty(t),t) - \dfrac{d}{dt}\left(\partial_2 L(x_\infty(t),\dot{x}_\infty(t),t)\right) = 0 \end{cases} \tag{86}$$

et à sa forme intégrale de du Bois-Reymond

$$\begin{cases} x_\infty \in \mathcal{V} \text{ i.e. notamment } x_\infty(0) = 0 \text{ et } x_\infty(T) = X \\ -\displaystyle\int_0^t \partial_1 L(x_\infty(u),\dot{x}_\infty(u),u) \quad du + \partial_2 L(x_\infty(t),\dot{x}_\infty(t),t) = \text{Constante} \end{cases} \tag{87}$$

La forme (85) de la condition nécessaire conduit aux méthodes d'éléments finis pour le calcul effectif de la solution de (5). la forme (86) conduit aux équations différentielles ordinaires permettant soit de trouver des solutions analytiques au problème. La forme (87) permet d'introduire les méthodes numériques de Runge-Kutta.

Mais hélas, il reste un dernier aveux à faire, parfaitement négatif : on n'a pas traité la question de la condition suffisante au problème (5). Derrière cette question se cache pourtant des enjeux importants (la stabilité des solutions), trop important pour être survolés dans ce petit cours.

1.4 Exercices

Le thème de ces exemples est l'apprentissage quasi manuel des calculs qui mènent de la forme fonctionnelle d'un problème à sa forme forte en passant par la forme faible.

On ne cherche pas à recopier sottement les résultats (85) et (86) mais au contraire on refait en abrégé les calculs menat à ces résultats.

Pour chacune des fonctionnelles données de la forme

$$\mathcal{S}(x) = \int_0^T L(t,x(t),\dot{x}(t))dt \quad \textit{où} \quad \dot{x}(t) = \frac{dx}{dt}(t)$$

On développera à l'ordre 1 la fonction $\delta\lambda \longrightarrow \mathcal{S}(x+\delta\lambda x')$ *où* x' *est une fonction* a priori *quelconque, i.e.*

$$\mathcal{S}(x+\delta\lambda x') = \mathcal{S}(x) + \delta\lambda \int_0^T \left(\partial_x L(t,x(t),\dot{x}(t))x'(t) + \partial_{\dot{x}} L(t,x(t),\dot{x}(t))\dot{x'}(t)\right) dt + o(\delta\lambda)$$

puis on écrira les conditions nécessaires pour trouver x_∞ *tel que*

$$\begin{cases} x_\infty(0) = x_0; x_\infty(T) = x_T \\ \forall x \textit{ tel que } x(0) = x_0; x(T) = x_T : \mathcal{S}(x_\infty) \leq \mathcal{S}(x) \end{cases}$$

en prenant soin de noter qu'elles ne sont pas suffisantes. On trouvera donc que

$$\forall x' \ / \ x'(0) = x'(T) = 0 : \int_0^T \left(\partial_x L(t,x_\infty(t),\dot{x}_\infty(t))x'(t) + \partial_{\dot{x}} L(t,x_\infty(t),\dot{x}_\infty(t))\dot{x'}(t)\right) dt = 0$$

en notant bien que $x' \in \mathcal{V}_0$ et pas $\mathcal{V}$.

Ensuite on intégrera par parties le terme contenant $\dot{x'}(t)$ dans l'expression précédente pour trouver

$$\forall x' \ / \ x'(0) = x'(T) = 0 : \int_0^T \left(\partial_x L(t, x_\infty(t), \dot{x}_\infty(t)) - \frac{d}{dt}\left(\partial_{\dot{x}} L(t, x_\infty(t), \dot{x}_\infty(t))\right) \right) x'(t) dt = 0$$

expression à partir de laquelle on conclura qu'il est alors suffisant pour x_∞ d'être tel que

$$\forall t : \partial_x L(t, x_\infty(t), \dot{x}_\infty(t)) - \frac{d}{dt}\left(\partial_{\dot{x}} L(t, x_\infty(t), \dot{x}_\infty(t))\right) = 0$$

ce qui illustre le lien entre les opérations de minimisation de fonctionnelles et résolution d'équations différentielles ordinaires.

On ne pose pas ici de problème sur la nature des espaces de fonctions concernés mais on apprend à opérer avec efficacité.

1. $\mathcal{S}(x) = \displaystyle\int_0^T \frac{m}{2}\left(\frac{dx}{dt}(t)\right)^2 - \frac{k}{2}(x(t))^2 + f(t)x(t) \quad dt$
2. $\mathcal{S}(q) = \displaystyle\int_0^T \left(\frac{l}{2}\left(\frac{dq}{dt}(t)\right)^2 - \frac{1}{2c}(q(t))^2 + e(t)q(t)\right) \exp^{t/\tau} dt \qquad \text{avec } \tau = l/r$
3. $\mathcal{S}(\theta) = \displaystyle\int_0^T \frac{mr^2}{2}\left(\frac{d\theta}{dt}(t)\right)^2 + mgr\cos(\theta(t)) \quad dt$
4. $\mathcal{S}(q) = \displaystyle\int_0^T \left(\int_0^{\frac{dq}{dt}(t)} \phi(u) \quad du\right) - \frac{(q(t)^2)}{2c} + e(t)q(t) \quad dt$

   où ϕ est une fonction continue, impaire et strictement croissate.
5. $\mathcal{S}(x) = -\displaystyle\int_0^T m_0 c^2 \sqrt{1 - \frac{1}{c^2}\left(\frac{dx}{dt}(t)\right)^2} - f(t)x(t) \quad dt$

2 Extensions du problème le plus simple de calcul des variations

Le gros du travail a été fait. Il reste cependant à étendre un peu la portée de ce qui a été établi dans deux directions : les fonctions à valeurs vectorielles et les fonction à plusieurs variables.

La première des directions est facile, la seconde un peu moins et elle n'est signalée qu'à titre indicatif.

2.1 Problème dans $C_1^N([0,T])$

Supposons qu'au lieu d'être une fonction dont la valeur est réelle, x soit une fonction dont la valeur est dans $\mathbb{R}^N$; on suppose que toutes les composantes de x sont dérivables et à dérivée continue, ce qu'on appelle $C_1^N([0,T])$.

Peut-on déduire de l'analyse précédente une condition nécessaire sous forme faible puis forte du problème

$$\text{trouver } x_\infty \in \mathcal{V} \text{ telle que } \forall x \in \mathcal{V} \text{ on a } \mathcal{S}(x_\infty) \leq \mathcal{S}(x) \tag{88}$$

où

$$\mathcal{V} = \{x \in C_1^N([0,T]) \text{ telle que } x(0) = 0; x(T) = X\} \tag{89}$$

et

$$\begin{array}{rcl} \mathcal{S} : \mathcal{V} & \longrightarrow & \mathbb{R} \\ x & \longrightarrow & \mathcal{S}(x) = \displaystyle\int_0^T L(x(t), \dot{x}(t), t) \quad dt \end{array} \tag{90}$$

avec

$$L \; : \; \begin{array}{ccc} \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ (u,v,w) & \longrightarrow & L(u,v,w) \end{array} \tag{91}$$

une fonction à $2N+1$ variables, différentiable au moins deux fois ?

La réponse est oui. Comme il ne s'agit que de trouver une condition nécessaire, on peut toujours faire appel successivement aux problèmes partiels dont les variables sont les composantes x_n $(n = 1 \dots N)$ du vecteur x.

2.1.1 Exercices

Trouver les formes faible et forte des conditions nécessaires d'extrema des problèmes suivants.

1. $\mathcal{S}(q_1,q_2,q_3) = \displaystyle\int_0^T \frac{1}{2}\sum_{n=1}^{3}\sum_{m=1}^{3} M_{mn}\left(\frac{dq_n}{dt}(t)\right)\left(\frac{dq_m}{dt}(t)\right) - \sum_{n=1}^{3}\frac{(q_n(t))^2}{2c_n} \quad dt$
2. $\mathcal{S}(q,Q,x) = \displaystyle\int_0^T \frac{\mu}{2}\left(\frac{dx}{dt}(t)\right)^2 + \frac{l}{2}\left(\frac{dq}{dt}(t)\right)^2 + m(x(t))\left(\frac{dq}{dt}(t)\right)\left(\frac{dQ}{dt}(t)\right) + \frac{L}{2}\left(\frac{dQ}{dt}(t)\right)^2 - Q(t)e(t) \quad dt$
   Attention m est une fonction donnée de $\mathbb{R}$ dans $\mathbb{R}$ dont l'argument est $x(t)$ dans l'expression.
3. $\mathcal{S}(\vec{x}_0,\vec{x}_1,\vec{x}_2) = \displaystyle\int_0^T \frac{1}{2}\sum_{n=0}^{2}\mu_n\left(\frac{d\vec{x}_n}{dt}(t)\right)^2 + \sum_{n=0}^{2}\frac{k\mu_n\mu_{(n+1)\ mod\ 3}}{|\vec{x}_n - \vec{x}_{(n+1)\ mod\ 3}|} \quad dt$ où les $\vec{x}_n$ sont des fonctions à valeurs dans $\mathbb{R}^3$ et où les carrés appliqués à ces fonctions signifient le carré de la norme euclidienne.

2.2 Problèmes d'équations aux dérivées partielles

Ces deux exemples sont destinés à permettre au lecteur de reconnaître s'il a saisi l'esprit du calcul des variations.

2.2.1 Transfert de chaleur stationnaire

On donne la position dans le plan comme $\vec{p} = x\vec{i} + y\vec{j}$, puis un disque dans ce plan comme les positions dont les coordonnées satisfonts à $x^2+y^2 < r^2$.

On range les fonctions à valeur réelle et à deux arguments réels $T \; : \; \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longrightarrow & T(x,y) \end{array}$ définies dans le disque, différentiables une fois, dont les dérivées partielles sont continues et qui s'annulent sur le bord de ce disque dans un ensemble $\mathcal{V}_0$.

On donne de plus une fonction $q \; : \; \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longrightarrow & q(x,y) \end{array}$ définie sur ce disque et on considère la fonctionnelle

$$\mathcal{S}(T) = \int_{x^2+y^2<r^2} \frac{\lambda}{2}\left(\left(\frac{\partial T}{\partial x}(x,y)\right)^2 + \left(\frac{\partial T}{\partial y}(x,y)\right)^2\right) - q(x,y)T(x,y) \quad dxdy$$

Il faut trouver des conditions correspondant à une forme faible puis une forme forte permettant le calcul de $T_\infty \in \mathcal{V}_0$ tel que

$$\forall T \in \mathcal{V}_0 : \mathcal{S}(T_\infty) \leq \mathcal{S}(T)$$

On pourra introduire des notations plus compactes (tensorielles ou utilisant le vecteur $\vec{\nabla}$ de Hamilton ou toutes autres).

2.2.2 Magnétostatique en domaine borné

On donne la position dans l'espace comme $\vec{p} = x\vec{i} + y\vec{j} + z\vec{k}$, puis une sphère dans ce plan comme les positions dont les coordonnées satisfonts à $x^2+y^2+z^2 < r^2$.

On range dans un ensemble $\mathcal{V}_0$ les fonctions à valeur réelle et à trois arguments réels $a \; : \; \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (x,y,z) & \longrightarrow & a(x \end{array}$ définies dans la sphère, différentiables une fois, dont les dérivées partielles sont continues et qui satisfont aux conditions $za_y(x,y,z) = ya_z(x,y,z)$, $xa_z(x,y,z) = za_x(x,y,z)$ et $ya_x(x,y,z) = xa_y(x,y,z)$ quand $x^2+y^2+z^2 = r^2$.

On donne de plus trois fonctions $\begin{array}{cccc} j_n & : & \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & & (x,y,z) & \longrightarrow & j_n(x,y,z) \end{array}$ avec $n = x,y,z$, définie dans cette sphère et on considère la fonctionnelle

$$\mathcal{S}(a_x,a_y,a_z) = \int_{x^2+y^2+z^2<r^2} j_x(x,y,z)a_x(x,y,z) + j_y(x,y,z)a_y(x,y,z) + j_z(x,y,z)a_z(x,y,z) - \frac{1}{2\mu(x,y,z)}\left(\left(\left(\frac{\partial a_z}{\partial y}(x,y,z)\right) - \left(\frac{\partial a_y}{\partial z}(x,y,z)\right)\right)^2 + \left(\left(\frac{\partial a_x}{\partial z}(x,y,z)\right) - \left(\frac{\partial a_z}{\partial x}(x,y,z)\right)\right)^2 + \left(\left(\frac{\partial a_y}{\partial x}(x,y,z)\right) - \left(\frac{\partial a_x}{\partial y}(x,y,z)\right)\right)^2\right) \quad dxdydz$$

On demande les conditions nécessaires (forme faible et forme forte) sur $a_x^\infty, a_y^\infty, a_z^\infty \in \mathcal{V}_0$ tels que

$$\forall a_x,a_y,a_z \in \mathcal{V}_0 : \mathcal{S}(a_x^\infty, a_y^\infty, a_z^\infty) \leq \mathcal{S}(a_x,a_y,a_z)$$

On pourra introduire des notations plus compactes (tensorielles ou utilisant le vecteur $\vec{\nabla}$ de Hamilton ou toutes autres).

3 Équation d'Hamilton-Jacobi

On revient maintenant au problème le plus simple de calcul des variations (5) dont on suppose qu'il existe une solution et on introduit la fonction

$$\begin{array}{cccc} A & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & & (X,T) & \longrightarrow & A(X,T) = \mathcal{S}(x_\infty) \end{array} \tag{92}$$

La question est d'étudier cette fonction à deux arguments réels.

Pour cela il est commode de modifier un peu la notation de (1). Plutôt que $\mathcal{V}$ on appelle

$$\mathcal{V}(X,T) = \{x \in C_1([0,T]) \text{ telle que } x(0) = 0; x(T) = X\} \tag{93}$$

Puis on définit l'application

$$\begin{array}{cccc} \mathcal{A} & : & \mathbb{R}^2 & \longrightarrow & C_1([0,T]) \\ & & (X,T) & \longrightarrow & x_\infty \end{array} \tag{94}$$

de manière que

$$A(X,T) = S(\mathcal{A}(X,T)) \tag{95}$$

3.1 Variation verticale

On étudie tout d'abord la variation de l'application partielle

$$\mathcal{A}_T(X) = \mathcal{A}(X,T) \tag{96}$$

puis cherchera à entrer cette variation dans la fonction $\mathcal{S}$.

Développement de la variation de $\mathbf{x}_\infty$ Si X varie de δX alors x_∞ varie de δx_∞ de manière que de manière que (5) qui s'écrivait

$$\text{trouver } x_\infty \in \mathcal{V}(X,T) \text{ telle que } \forall x \in \mathcal{V}(X,T) \text{ on a } \mathcal{S}(x_\infty) \leq \mathcal{S}(x) \tag{97}$$

devienne

$$\text{trouver } x_\infty + \delta x_\infty \in \mathcal{V}(X+\delta X,T) \text{ telle que } \forall x \in \mathcal{V}(X+\delta X,T) \text{ on a } \mathcal{S}(x_\infty + \delta x_\infty) \leq \mathcal{S}(x) \tag{98}$$

La condition nécessaire de (97) était (85), avec les nouvelles notations

$$\begin{cases} x_\infty \in \mathcal{V}(X,T) \text{ i.e. notamment } x_\infty(0)=0 \text{ et } x_\infty(T)=X \\ \\ \forall y' \in \mathcal{V}_0 = \mathcal{V}(0,T) : \\ \displaystyle\int_0^T \partial_1 L(x_\infty(t),\dot{x}_\infty(t),t)y'(t) + \partial_2 L(x_\infty(t),\dot{x}_\infty(t),t)\dot{y}'(t) \quad dt = 0 \end{cases} \tag{99}$$

celle de (98) est alors

$$\begin{cases} x_\infty + \delta x_\infty \in \mathcal{V}(X+\delta X,T) \text{ i.e. notamment} \\ x_\infty(0)+\delta x_\infty(0)=0 \text{ et } x_\infty(T)+\delta x_\infty(T)=X+\delta X \\ \\ \forall y' \in \mathcal{V}_0 = \mathcal{V}(0,T) : \\ \displaystyle\int_0^T \partial_1 L(x_\infty(t)+\delta x_\infty(t),\dot{x}_\infty(t)+\delta\dot{x}_\infty(t),t)y'(t) \\ +\partial_2 L(x_\infty(t)+\delta x_\infty(t),\dot{x}_\infty(t)+\delta\dot{x}_\infty(t),t)\dot{y}'(t) \quad dt = 0 \end{cases} \tag{100}$$

ou encore, en tenant compte de (99),

$$\begin{cases} \delta x_\infty \in \mathcal{V}(\delta X,T) \text{ i.e. notamment } \delta x_\infty(0)=0 \text{ et } \delta x_\infty(T)=\delta X \\ \\ \forall y' \in \mathcal{V}_0 = \mathcal{V}(0,T) : \\ \displaystyle\int_0^T \partial_1 L(x_\infty(t)+\delta x_\infty(t),\dot{x}_\infty(t)+\delta\dot{x}_\infty(t),t)y'(t) \\ +\partial_2 L(x_\infty(t)+\delta x_\infty(t),\dot{x}_\infty(t)+\delta\dot{x}_\infty(t),t)\dot{y}'(t) \quad dt = 0 \end{cases} \tag{101}$$

Où on voit que l'expression sous l'intégrale ne demande qu'à être développée.

Pour cela on peut considérer que

$$\delta x_\infty = \delta X x_\infty^1 + \frac{\delta X^2}{2} x_\infty^2 + \ldots + \frac{\delta X^n}{n!} x_\infty^n + \ldots \tag{102}$$

où $x_\infty^1 \in \mathcal{V}(1,T)$, $x_\infty^2 \in \mathcal{V}(0,T)$ (c'est bien $\mathcal{V}(0,T)$), $n \geq 2 \Longrightarrow x_\infty^n \in \mathcal{V}(0,T)$ (et où l'exposant de x_∞^n n'est pas une puissance alors que δX^n est la puissance n-ième de δX).

Les fonctions x_∞^n sont évidemment les coefficients différentiels du développement de

$$\mathcal{A}_T(X+\delta X) = \mathcal{A}_T(X) + \delta X \frac{d\mathcal{A}_T}{dX}(X) + \frac{\delta X^2}{2}\frac{d^2\mathcal{A}_T}{dX^2}(X) + \ldots + \frac{\delta X^n}{n!}\frac{d^n\mathcal{A}_T}{dX^n}(X) + \ldots \tag{103}$$

qu'on suppose exister. On a

$$\frac{d^n\mathcal{A}_T}{dX^n}(X) = x_\infty^n \tag{104}$$

Cela dit, si on introduit une fonction $F : C_1([0,T]) \times C_0([0,T]) \longrightarrow \mathbb{R}$ afin d'écrire les choses de façon compacte, on obtient

$$\begin{aligned} F(x_\infty+\delta x_\infty,\dot{x}_\infty+\delta x_\infty) &= \delta X(\partial_1 F_\infty x_\infty^1 + \partial_2 F_\infty \dot{x}_\infty^1) \\ &+ \frac{\delta X^2}{2}\big(\partial_{11}F_\infty (x_\infty^1)^2 + \partial_{12}F_\infty x_\infty^1 \dot{x}_\infty^1 + \partial_{22}F_\infty(\dot{x}_\infty^1)^2 \\ &\qquad +\partial_1 F_\infty x_\infty^2 + \partial_2 F_\infty \dot{x}_\infty^2\big) \\ &+ \frac{\delta X^3}{6}\cdots \end{aligned} \tag{105}$$

où on décide de noter

$$\partial_{nm}F_\infty = \partial_{nm}F(x_\infty,\dot{x}_\infty) \tag{106}$$

On peut maintenant récrire (101) comme

$$\left\{\begin{array}{l}\delta x_\infty \in \mathcal{V}(\delta X, T) \text{ i.e. notamment } \delta x_\infty(0) = 0 \text{ et } \delta x_\infty(T) = \delta X \\ \\ \forall y' \in \mathcal{V}_0 = \mathcal{V}(0, T) : \\ \displaystyle\int_0^T \partial_1 L(x_\infty(t), \dot{x}_\infty(t), t) y'(t) + \partial_2 L(x_\infty(t), \dot{x}_\infty(t), t)\dot{y}'(t) \quad dt \\ \displaystyle +\delta X \int_0^T \begin{array}{l}(\partial_{11} L_\infty(t) x^1_\infty(t) + \partial_{12} L_\infty(t)\dot{x}^1_\infty(t)) y'(t) \\ +(\partial_{21} L_\infty(t) x^1_\infty(t) + \partial_{22} L_\infty(t)\dot{x}^1_\infty(t))\dot{y}'(t) \quad dt\end{array} \\ \displaystyle +\delta X^2/2 \int_0^T \ldots \\ +\ldots \\ = 0\end{array}\right. \tag{107}$$

où, sur le modèle de (106), on a noté

$$\partial_{nm} L_\infty(t) = \partial_{nm} L(x_\infty(t), \dot{x}_\infty(t), t) \tag{108}$$

Avec (99) on sait déjà que le premier terme du développement en série de l'équation variationnelle est nul. En divisant par δX, qui est petit mais pas nul, on obtient alors

$$\left\{\begin{array}{l}\delta x_\infty \in \mathcal{V}(\delta X, T) \text{ i.e. notamment } \delta x_\infty(0) = 0 \text{ et } \delta x_\infty(T) = \delta X \\ \\ \forall y' \in \mathcal{V}_0 = \mathcal{V}(0, T) : \\ \displaystyle\int_0^T \begin{array}{l}(\partial_{11} L_\infty(t) x^1_\infty(t) + \partial_{12} L_\infty(t)\dot{x}^1_\infty(t)) y'(t) \\ +(\partial_{21} L_\infty(t) x^1_\infty(t) + \partial_{22} L_\infty(t)\dot{x}^1_\infty(t))\dot{y}'(t) \quad dt\end{array} \\ \displaystyle +\delta X/2 \int_0^T \ldots \\ +\ldots \\ = 0\end{array}\right. \tag{109}$$

Le développement doit être nul pour toute valeur de δX. Si le terme constant n'était pas nul, alors on pourrait choisir δX suffisamment petit pour que tous les termes qui le suivent soient négligeables devant lui. Il faut donc qu'il soit nul, ce qui fournit le problème permettant le calcul de x^1_∞ comme

$$\left\{\begin{array}{l}x^1_\infty \in \mathcal{V}(1, T) \text{ i.e. notamment } x^1_\infty(0) = 0 \text{ et } x^1_\infty(T) = 1 \\ \\ \forall y' \in \mathcal{V}_0 = \mathcal{V}(0, T) : \\ \displaystyle\int_0^T \begin{array}{l}(\partial_{11} L_\infty(t) x^1_\infty(t) + \partial_{12} L_\infty(t)\dot{x}^1_\infty(t)) y'(t) \\ +(\partial_{21} L_\infty(t) x^1_\infty(t) + \partial_{22} L_\infty(t)\dot{x}^1_\infty(t))\dot{y}'(t) \quad dt\end{array} = 0\end{array}\right. \tag{110}$$

et avec un peu de patience (en fait surtout un bon logiciel de calcul formel) on pourrait obtenir successivement tous les problèmes permettant le calcul des x^n_∞.

Il est remarquable que tous ces problèmes successifs soient linéaires par rapport à leurs solution, avec le corollaire que leur calcul est facile.

En conséquence, on considère que la variation δx_∞ quand X varie de δX est acquise.

Calcul de $\mathbf{A(X + \delta X, t)}$ Il suffit d'exprimer

$$A(X + \delta X, t) = S(x_\infty + \delta x_\infty) \tag{111}$$

avec δx_∞ calculé avec (107) ou, pratiquement, de la forme (102) où les coefficients sont calculés par (110) puis par les problèmes qui la suivent.

En reprenant la forme (3) de $\mathcal{S}$, on obtient qu'à l'ordre 1

$$\mathcal{S}(x_\infty + \delta X x_\infty^1 + \ldots) = \int_0^T L(x_\infty(t) + \delta X x_\infty^1(t) + \ldots, \dot{x}_\infty(t) + \delta X \dot{x}_\infty^1(t) + \ldots, t) \quad dt \tag{112}$$

soit donc

$$\begin{array}{rcl} \mathcal{S}(x_\infty + \delta X x_\infty^1 + \ldots) & = & \mathcal{S}(x_\infty) \\ & + & \delta X \displaystyle\int_0^T \partial_1 L(x_\infty(t), \dot{x}_\infty(t), t) x_\infty^1(t) + \partial_2 L(x_\infty(t), \dot{x}_\infty(t), t) \dot{x}_\infty^1(t) \quad dt \\ & + & \ldots \end{array} \tag{113}$$

d'où on tire que

$$\frac{\partial A}{\partial X}(X,T) = \int_0^T \partial_1 L(x_\infty(t), \dot{x}_\infty(t), t) x_\infty^1(t) + \partial_2 L(x_\infty(t), \dot{x}_\infty(t), t) \dot{x}_\infty^1(t) \quad dt \tag{114}$$

qui n'est pas nulle comme le laisserait croire un examen sans précautions de (99) parce que $x_\infty^1(T) = 1 \neq 0$ (ce n'est pas un élément de $\mathcal{V}(0,T)$).

Cette intégrale (114) peut cependant être calculée par une intégration par parties comme on l'a fait pour passer de (85) à (86). Ainsi

$$\begin{array}{rcl} \dfrac{\partial A}{\partial X}(X,T) & = & [\partial_2 L(x_\infty(t), \dot{x}_\infty(t), t) x_\infty^1(t)]_0^T \\ & + & \displaystyle\int_0^T \left(\partial_1 L(x_\infty(t), \dot{x}_\infty(t), t) - \frac{d}{dt} \partial_2 L(x_\infty(t), \dot{x}_\infty(t), t) \right) x_\infty^1(t) \quad dt \end{array} \tag{115}$$

avec la forme forte (86) on voit que l'intégrale est nulle, et compte tenu que $x_\infty^1(0) = 0$ et $x_\infty^1(T) = 1$, on trouve

$$\frac{\partial A}{\partial X}(X,T) = \partial_2 L(x_\infty(T), \dot{x}_\infty(T), T) \tag{116}$$

ou encore, compte tenu que $x_\infty(T) = X$

$$\frac{\partial A}{\partial X}(X,T) = \partial_2 L(X, \dot{x}_\infty(T), T) \tag{117}$$

ce qui constitue le résultat important de ce paragraphe.

Il est remarquable qu'on n'ait pas besoin de calculer effectivement x_∞^1 pour trouver la valeur de $\dfrac{\partial A}{\partial X}(X,T)$: il suffit de connaître seulement $\dot{x}_\infty(T)$.

3.2 Variation horizontale

Quand on sait tout les efforts qu'il a fallu faire pour obtenir $\dfrac{\partial \mathcal{S}}{\partial X}(X,T)$; quand on voit qu'une variation par rapport à T transforme les bornes de l'intégrale de (3) et que donc pour travailler sur quelque chose de fixe il faut préalablement faire un changement de variable et de fonction sur $x(t)$; on devine que l'approche frontale devient très compliquée d'un point de vue calculatoire.

Cette approche frontale sur le problème le plus simple du calcul des variations est cependant indispensable pour disposer une palette complète d'outils (contondants) utiles à une utilisation effective du calcul des variations dans une physique macroscopique du XXIième siècle dont l'objectif serait de résoudre les problèmes que nous a légués le XIXième siècle (et que le XXième siècle s'est contenté de transmettre les laissant inchangés. Non, là c'est injuste !)

Mais une telle ambition n'est pas l'objectif de ce petit cours qui se limite à quelques considérations de base. Aussi l'approche frontale ne sera pas tentée.

Par contre on peut dire tout de suite quelque chose de simple en examinant (92) écrit sous la forme compacte

$$\begin{array}{rcccl} A & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & & (X,T) & \longrightarrow & A(X,T) = \displaystyle\min_{x \in \mathcal{V}(X,T)} \int_0^T L(x(t), \dot{x}(t), t) \quad dt \end{array} \tag{118}$$

Si X et T varient simultanément de manière que la solution $x_\infty : [0,T] \longrightarrow \mathbb{R}$ trouvée avec (X,T) soit prolongée par

$$x'_\infty : \begin{array}{ccc} [0,T+\delta T] & \longrightarrow & \mathbb{R} \\ t & \longrightarrow & \left\{ \begin{array}{lll} x_\infty(t) & \text{si} & t \in [0,T] \\ \text{autre chose} & \text{sinon} & \end{array} \right. \end{array} \tag{119}$$

alors pour que x'_∞ soit continue et de dérivée continue l'autre chose de (119) ne pourra pas manquer d'être

$$X + \dot{x}_\infty(T)(t-T) \tag{120}$$

pour une variation δT suffisamment petite ; la variation de X correspondante sera donc

$$\delta X = \dot{x}_\infty(T)\delta T \tag{121}$$

et la variation de A sera

$$\delta A = L(X, \dot{x}_\infty(T), T)\delta T \tag{122}$$

En passant à la limite on obtient alors que

$$\frac{\partial A}{\partial X}(X,T)\dot{x}_\infty(T) + \frac{\partial A}{\partial T}(X,T) = L(X, \dot{x}_\infty(T), T) \tag{123}$$

et donc

$$\frac{\partial A}{\partial T}(X,T) = L(X, \dot{x}_\infty(T), T) - \dot{x}_\infty(T)\partial_2 L(X, \dot{x}_\infty(T), T) \tag{124}$$

Avec (124), on obtient la dérivée partielle de A par rapport au temps T comme on avait obtenu la dérivée partielle de A par rapport à X en (117).

3.3 Équation d'Hamilton-Jacobi

Le rapprochement de (117) et (124) met en évidence que les deux dérivées partielles de A ne peuvent être calculées que si on connaît $\dot{x}_\infty(T)$. Mais on peut éliminer cette variable entre les deux équations, ce qui fournit l'équation d'Hamilton Jacobi.

Pour cela on introduit à partir de $L : \mathbb{R}^3 \longrightarrow \mathbb{R}$ une nouvelle fonction

$$H : \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (u,p,w) & \longrightarrow & H(u,p,w) = \max_{v\in\mathbb{R}} (pv - L(u,v,w)) \end{array} \tag{125}$$

en acceptant que L est ainsi faite que ce maximum existe toujours et qu'il est unique.

La condition nécessaire de maximum est qu'il existe v_∞ tel que

$$\partial_2 L(u, v_\infty, w) = p \tag{126}$$

et alors la fonction H s'écrit

$$H(u,p,w) = pv_\infty - L(u, v_\infty, w) \tag{127}$$

Cette fonction a la propriété intéressante que

$$\partial_2 H(u,p,w) = v_\infty \tag{128}$$

En effet, si p varie de δp alors la solution v_∞ de (125) varie de δv_∞ qui contient des termes de tout ordre en δp mais dont on ne retiendra que le premier ordre. On obtient alors que

$$H(u, p+\delta p, w) = (v_\infty + \delta v_\infty)(p + \delta p) - L(u, v_\infty + \delta v_\infty, w) \tag{129}$$

soit donc, en ne retenant que ce qui peut comporter des termes d'ordre plus petits que δp

$$H(u, p+\delta p, w) = H(u,p,w) + v_\infty \delta p + \delta v_\infty(p - \partial_2 L(u, v_\infty, w)) + \ldots \tag{130}$$

et compte tenu de (126)

$$H(u, p + \delta p, w) = H(u, p, w) + v_\infty \delta p + \dots \tag{131}$$

d'où on retire (128).

Ceci étant dit, il ne reste qu'à reconnaître les termes de (127) et (126) dans (124) et (117) pour écrire

$$\frac{\partial A}{\partial T}(X, T) + H\left(X, \frac{\partial A}{\partial X}(X, T), T\right) = 0 \tag{132}$$

qui ne contient plus que les dérivées partielles de A. C'est l'équation d'Hamilton Jacobi.

Si A est trouvé par une résolution directe de (132), compte tenu de (128) on peut retrouver

$$\dot{x}_\infty(T) = \partial_2 H\left(X, \frac{\partial A}{\partial X}(X, T), T\right) \tag{133}$$

Les noms On a évité jusqu'ici de donner trop de noms aux objets manipulés, mais il est maintenant grand temps d'effectuer cette tâche.

D'abord la fonction $L : \mathbb{R}^3 \longrightarrow \mathbb{R}$ s'appelle un lagrangien ; c'est une donnée.

La fonction qu'on déduit de L avec (125) s'appelle un hamiltonien. Le type de transformation qu'on fait subir à L s'appelle une transformation de Legendre.

La fonctionnelle (3) s'appelle l'action associée à la fonction $x : [0, T] \longrightarrow \mathbb{R}$. La fonction A de (118) s'appelle aussi l'action. L'opération de minimisation de (3) s'appelle le principe de moindre action.

Et finalement, comme on l'a déjà dit, l'équation aux dérivées partielles en X et T de A (132) s'appelle l'équation d'Hamilton Jacobi.

À quoi sert tout cela ? – Il appartient aux cours de physique d'y répondre. On se limite ici à expliquer les techniques calculatoires sous-jacentes au développement de ces physiques.

Mais il faut également noter que tout le travail n'a pas été effectué : on a systématiquement écarté la question des conditions suffisantes d'extrema.

3.4 Illustration et exercices

Il n'est peut-être pas inutile d'illustrer sur un exemple simple tous les calculs qui ont été menés.

Prenons

$$L(u, v, w) = \frac{v^2}{2} - \frac{u^2}{2}$$

la forme faible (99) correspondant à ce lagrangien est

$$\begin{cases} x_\infty \in \mathcal{V}(X, T) \text{ i.e. notamment } x_\infty(0) = 0 \text{ et } x_\infty(T) = X \\ \\ \forall y' \in \mathcal{V}_0 = \mathcal{V}(0, T) : \displaystyle\int_0^T -x_\infty(t) y'(t) + \dot{x}_\infty(t) \dot{y}'(t) \quad dt = 0 \end{cases}$$

et la forme forte (86) fournit que

$$\begin{cases} x_\infty(0) = 0 \text{ et } x_\infty(T) = X \\ x_\infty(t) - \dfrac{d^2 x_\infty}{dt^2}(t) = 0 \end{cases}$$

d'où on retire que

$$x_\infty(t) = \frac{X \sin t}{\sin T}$$

ce qui est valable tant que $0 < T < \pi$

Le hamiltonien trouvé par (125) est

$$H(u, p, w) = \frac{p^2}{2} + \frac{u^2}{2}$$

et donc l'équation d'Hamilton-Jacobi (132) est

$$\frac{\partial A}{\partial T}(X,T)+\frac{X^2}{2}+\frac{1}{2}\left(\frac{\partial A}{\partial X}(X,T)\right)^2=0$$

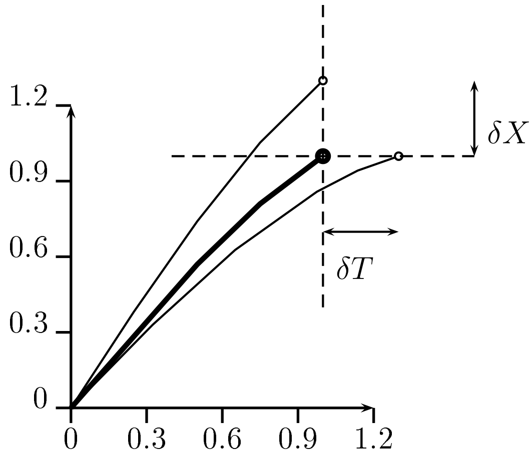
D'autre part le calcul direct de A par (92) donne

$$A(X,T)=\frac{1}{2}\frac{X^2\cos T}{\sin T}$$

dont on vérifie facilement qu'il est bien solution de l'équation d'Hamilton-Jacobi.

Cet exemple peut maintenant rendre un autre service, celui de fournir une dénotation géométrique à ce que sont les variations par rapport à X et T.

Voici d'abord le graphe de $x_\infty(t)$ pour $T=1$ et $X=1$ (en gras) et les graphes pour les deux cas $(T,X)=(1,1.3)$ et $(T,X)=(1.3,1)$



Qu'est δx_∞ quand on fait une variation δX de X en laissant T constant ? n – C'est

$$\delta x_\infty(t)=\frac{\delta X\sin t}{\sin T}\ \text{pour}\ t\in[0,T]$$

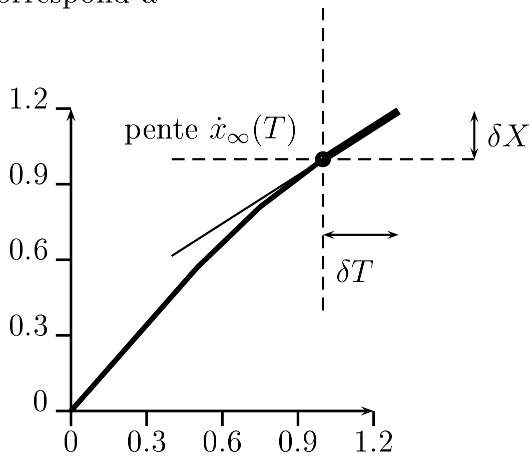
dans ce cas particulier, la série (102) se limite à 1 seul terme qui vaut

$$x^1_\infty(t)=\frac{\sin t}{\sin T}\ \text{pour}\ t\in[0,T]$$

comme le montre le problème (109) qui permet le calcul direct de x^1_∞. Les autres termes sont nuls comme le laisse prévoir le développement de (105) appliqué à ce lagrangien.

Maintenant comment récupère-t-on le terme $\dfrac{\partial A}{\partial T}$ à la page **??** ?

– On suppose connue la solution $x_\infty(t)$ pour T et X fixés et on affirme que cette solution se prolonge au delà de T comme (120), ce qui correspond à



et on trouve la correspondance (121) entre δT et δX adéquate. Il ne reste plus alors qu'à injecter cette nouvelle fonction (119) dans l'intégrale (3) (en modifiant également la borne d'intégration) pour trouver qu'à la limite où δX et δT tendent vers 0 alors on trouve (122) puis finalement (124).

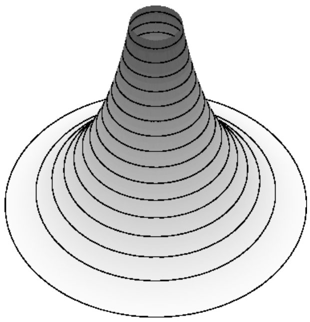
Exercices Reprendre les 5 lagrangiens des exercices de la page **??** et pour chacun d'entre eux trouver :

1. le hamiltonien
2. l'équation d'Hamilton Jacobi

Et, dans les cas où c'est possible, faire tous les calculs comme dans l'illustration.

4 Problème de synthèse

Si on tend une membrane entre deux cerceaux coaxiaux de rayons R_1 et R_2 placés à la hauteur H l'un de l'autre.



on demande quelle est la forme de cette membrane sachant qu'elle doit avoir la surface la plus petite possible.

1. On traitera d'abord la question en décidant que la membrane est représentée par le lieu des points (r, θ, z) (en coordonnées cylindriques) tels que $(z = h(r), \theta, r)$ où h est la fonction qu'il faudra trouver ;
2. On trouvera une limite à cette approche et on la corrigera donc en introduisant plutôt $(z, \theta, r = \rho(z))$, ρ étant la fonction à trouver.