

1 Le calcul variationnel

1.1 Introduction.

Dès que vous avez vu les bases de l'analyse, vous avez appris à répondre à la question suivante : comment trouver le point x pour lequel la fonction $f(x)$ est maximum (ou minimum) ? f est une machine qui prend un nombre en entrée et produit un nombre en sortie. La question ci-dessus en réalité est celle de trouver un extremum *local* : un point qui produit la sortie la plus grande (ou la plus petite) que tous ses voisins immédiats. Nous savons que pour un tel point x^* ,

$$f'(x^*) = 0.$$

Donnons nous maintenant une fonctionnelle \mathcal{S} . Ceci est une machine qui prend *une fonction* en entrée et produit un nombre en sortie. Par exemple

$$\mathcal{S}(f) = \int_a^b [f(x)^2 + f'^2(x)] dx \quad (1.1)$$

est une fonctionnelle qui prend une fonction, ajoute son carré et le carré de sa dérivée et les intègre entre deux bornes pour produire un nombre. Si on entre la fonction $\sin x$ dans cette machine, elle produit le nombre $b - a$. Si on y entre la fonction $\exp x$, elle produit le nombre $2 \exp(2b) - 2 \exp(2a)$.

Le calcul variationnel consiste à répondre à la question suivante : quelle est la fonction f qui produit la plus grande sortie $\mathcal{S}(f)$? La réponse que nous allons voir par la suite est que f doit satisfaire une équation différentielle qui est reliée à la forme de la fonctionnelle \mathcal{S} .

Donnons deux exemples avant d'aller plus loin.

Brachistochrone. L'exemple le plus important historiquement est celui du brachistochrone. Soit un point $A(0, 0)$ situé dans le plan vertical, relié à un point $B(x_1, y_1)$ par un toboggan dont la forme est donnée par la fonction $y = f(x)$. On laisse un objet glisser sans frottement du point A le long du toboggan. Comment choisir la forme du toboggan, (la fonction f), pour que le temps d'arrivée au point B soit minimum ? Vous voyez qu'une fois qu'on se donne un toboggan, c'est à dire une fonction, en utilisant quelques notions de mécanique et de conservation d'énergie, on peut calculer le temps de parcours, c'est à dire un scalaire. Essayons de mettre cela en forme. La vitesse de l'objet à l'ordonnée y vaut $\sqrt{2gy}$. L'élément d'arc $ds = \sqrt{dx^2 + dy^2} = \sqrt{(1 + f'(x)^2)dx}$ est parcouru en un temps $dt = ds/v$. Le temps total du parcours est donc

$$T = \int_0^{x_1} \sqrt{\frac{1 + f'(x)^2}{2gf(x)}} dx$$

Et il faut trouver la fonction f qui minimise cette intégrale¹. Ce problème avait été lancé comme un défi par un des frères Bernoulli vers 1690 (à peine dix ans après l'invention du calcul différentiel) à la communauté scientifiques. Tous les grands (Newton, Leibnitz, l'autre Bernoulli, Hospital, ...) y répondirent. Euler (~1740) a trouvé une solution générale de ce genre de problème et Lagrange (~1780) les a généralisé à la mécanique.

Mécanique analytique. Prenons une particule de masse m qui quitte le point $x = 0$ au temps $t = 0$ et arrive à un autre point $x = x_1$ au temps $t = t_1$. Cette particule est soumise à un potentiel $V(x)$. Le mouvement de la particule est donné par la fonction $x(t)$. La fonction $x(t)$ qui minimise

$$S = \int_0^{t_1} [(m/2)\dot{x}^2(t) - V(x(t))] dt \quad (1.2)$$

est la trajectoire suivie par la particule. Ceci est une nouvelle formulation de la mécanique. Classiquement, nous résolvons l'équation différentielle $F = ma$ où la fonction $F(x) = -dV/dx$ et $a = d^2x/dt^2$ pour remonter à la trajectoire $x(t)$. Ici, la démarche est différente : de toutes les trajectoires possibles qui relie $(0,0)$ à (t_1, x_1) , la particule choisit justement celle qui minimise l'intégrale (1.2). Comme si un dieu calculait le coût (qu'on appelle l'action S) de chaque trajectoire et choisissait la meilleure. Bien sûr, cette formulation de la mécanique et la formulation Newtonienne sont équivalente, bien que la formulation lagrangienne soit beaucoup plus profonde et pratique. Notez que la quantité dans l'intégrale n'est pas l'énergie totale, mais l'énergie cinétique *moins* l'énergie potentielle. On appelle cette quantité le lagrangien.

1.2 Calcul des variations.

Formulons correctement notre problème (nous n'allons pas attaquer le cas le plus général pour l'instant). Soit la fonctionnelle

$$S[f] = \int_a^b \mathcal{L}[f(t), f'(t), t] dt \quad (1.3)$$

Trouver la fonction $f(t)$, avec les conditions $f(a) = y_0$ et $f(b) = y_1$ pour laquelle l'intégrale est un extremum.

Traditionnellement, la fonction \mathcal{L} qui se trouve sous l'intégrale est appelée le lagrangien. C'est une fonction tout ce qu'il y a de plus normal. Par exemple, $\mathcal{L}(x, y) = x^2 + y^2$. Comme à un instant donné, $f(t)$ et $f'(t)$ sont des nombres, il est tout à fait légitime de calculer $\mathcal{L}[f(t), f'(t)]$ qui dans ce cas, vaut $f(t)^2 + f'(t)^2$ comme l'expression que nous avons écrit dans (1.1). En plus, nous avons le droit de prendre les dérivées partielles de \mathcal{L} : par exemple, dans ce cas, $\partial\mathcal{L}/\partial x = 2x$ et $\partial\mathcal{L}/\partial y = 2y$. Il est usuel, si nous avons

1. A première vue, il semble qu'il manque quelque chose à cette formulation : l'intégrale ne contient pas de référence à y_1 et nous n'exigeons apparemment pas que la particule finisse sa trajectoire à l'ordonnée y_1 . Nous y reviendrons plus tard, quand nous aborderons les contraintes.

1 Le calcul variationnel

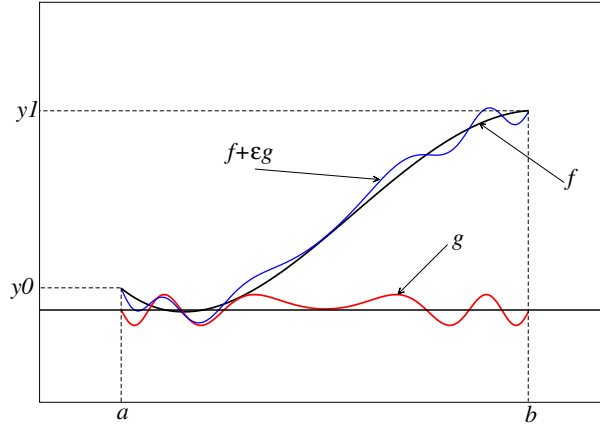


FIGURE 1.1 – Une fonction f et une variation g autour de cette fonction.

noté $\mathcal{L}[f(t), f'(t)]$, de noter ces dérivées partielles par $\partial\mathcal{L}/\partial f$ et $\partial\mathcal{L}/\partial f'$: cela veut juste dire “dérivée partielle par rapport au premier ou deuxième argument”. Dans ce cas, nous aurions eu par exemple $\partial\mathcal{L}/\partial f' = 2f'(t)$. D’ailleurs, $\partial\mathcal{L}/\partial f'$ est ici une fonction de t et on peut par exemple prendre sa dérivée par rapport au temps : $d[\partial\mathcal{L}/\partial f']/dt = 2f''(t)$. Si au début, vous trouvez cette notation abrupte, remplacez f et f' avant les dérivations par x et y , et remettez les à leur place une fois les opérations terminées. Mais on prend vite l’habitude de ces notations. Notez également que l’on ne cherche pas n’importe quelle fonction, mais les fonctions qui prennent des valeurs bien déterminées (y_0 et y_1) aux bords a et b .

Avant d’aller plus loin, revenons un instant au cas d’une fonction $f(x)$ dont on veut trouver l’extremum. Si nous connaissons la valeur de la fonction au point x , alors son accroissement quand on se déplace au point $x + \epsilon$ est donnée par

$$df = f(x + \epsilon) - f(x) = A(x)\epsilon + \text{termes d'ordres } \epsilon^2 + \dots$$

La première partie de l’accroissement (celle qui est importante quand ϵ est petit) est linéaire en ϵ : si nous avons pris un ϵ deux fois plus grand, nous aurions eu un accroissement deux fois plus grand également. La fonction $A(x)$ est le coefficient de proportionnalité entre df et ϵ au point x , et nous avons plus l’habitude de la noter par $f'(x)$. Le point x est un extremum si le coefficient de proportionnalité $f'(x) = 0$, c’est à dire qu’en se déplaçant autour du point x , l’accroissement de f (à l’ordre 1 en ϵ) est nulle.

Nous n’avons qu’à suivre cette méthodologie pour trouver l’extremum de notre fonctionnelle $S[f]$: Nous allons ajouter la fonction $\epsilon g(t)$ à la fonction $f(t)$, et calculer l’accroissement de la fonctionnelle $dS = S[f + \epsilon g] - S[f]$. Avec un peu de chance, cette accroissement comporte un terme linéaire en ϵ :

$$dS = S[f + \epsilon g] - S[f] = A[f, g].\epsilon + \text{termes d'ordre } \epsilon^2 + \dots$$

où $A[f, g]$ est un coefficient de proportionnalité qui dépend des fonctions f et g . Nous disons que f est un extremum si $A[f, g] = 0$ quelque soit g ! Cela est analogue à trouver

1 Le calcul variationnel

l'extremum d'une fonction de plusieurs variables : le point (x^*, y^*, z^*, \dots) est un extremum de la fonction $f(x, y, z, \dots)$ si en ce point, *quelque soit* le déplacement (par exemple $(dx, 0, 0, \dots)$ ou $(0, dy, dz, \dots)$) la partie linéaire de l'accroissement est nulle. Une fonction, comme nous l'avons vu au chapitre 1, n'est finalement qu'un point dans un espace à dimension infini ; une fonctionnelle est comme une fonction d'une infinité de variables. Quelque soit g dans l'expression précédente veut simplement dire quelque soit le déplacement dans cet espace.

Il faut prendre une précaution : nous ne cherchons que des fonctions pour lesquelles $f(a) = y_0$ et $f(b) = y_1$. Comme f satisfait déjà à cette condition, pour que $f + \epsilon g$ la fasse également, nous ne devons considérer que des fonctions g telle que $g(a) = g(b) = 0$. Les fonctions g ne sont donc pas tout à fait quelconque. Nous obtenons :

$$\begin{aligned} S[f + \epsilon g] &= \int_a^b \mathcal{L}[f(t) + \epsilon g(t), f'(t) + \epsilon g'(t)] dt \\ &= S[f] + \epsilon \int_a^b \frac{\partial \mathcal{L}}{\partial f} g(t) dt + \epsilon \int_a^b \frac{\partial \mathcal{L}}{\partial f'} g'(t) dt + \dots \end{aligned} \quad (1.4)$$

où nous avons simplement utilisé le fait que $\mathcal{L}(x + h, y + k) = \mathcal{L}(x, y) + (\partial \mathcal{L} / \partial x)h + (\partial \mathcal{L} / \partial y)k + \dots$ On peut déjà voir la partie linéaire apparaître. Nous pouvons mettre la deuxième intégrale un peu plus en forme en faisant une intégration par partie :

$$\int_a^b \frac{\partial \mathcal{L}}{\partial f'} g'(t) dt = \left[\frac{\partial \mathcal{L}}{\partial f'} g(t) \right]_a^b - \int_a^b \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial f'} \right] g(t) dt$$

La première partie est nulle, puisque la fonction g vaut justement 0 sur les bords. En remettant ce qui reste dans l'expression (1.4), nous avons :

$$dS = \epsilon \int_a^b \left\{ \frac{\partial \mathcal{L}}{\partial f} - \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial f'} \right] \right\} g(t) dt + \dots$$

L'intégrale est notre facteur de proportionnalité entre dS et ϵ . Si la fonction f est un extremum, alors l'intégrale doit être nulle *quelque soit* la fonction g . La seule possibilité est donc que f satisfasse à l'équation différentielle

$$\frac{\partial \mathcal{L}}{\partial f} - \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial f'} \right] = 0 \quad (1.5)$$

qui est appelé l'équation d'Euler-Lagrange. Notez que cette équation est homogène dimensionnellement. Faisons quelques exercices pour nous fixer les idées.

Identité de Beltrami. Évaluons l'expression

$$\begin{aligned} \frac{d}{dt} \left\{ f' \frac{\partial \mathcal{L}}{\partial f'} - \mathcal{L} \right\} &= f'' \frac{\partial \mathcal{L}}{\partial f'} + f' \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial f'} \right) - \frac{\partial \mathcal{L}}{\partial f} f' - \frac{\partial \mathcal{L}}{\partial f'} f'' - \frac{\partial \mathcal{L}}{\partial t} \\ &= f' \left(\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial f'} \right) - \frac{\partial \mathcal{L}}{\partial f} \right) - \frac{\partial \mathcal{L}}{\partial t} \end{aligned}$$

1 Le calcul variationnel

Si \mathcal{L} ne dépend pas explicitement de t , alors $\partial\mathcal{L}/\partial t = 0$; par ailleurs, l'expression entre parenthèse n'est que l'équation d'Euler-Lagrange et vaut zéro par construction. Nous en déduisons que si le lagrangien ne dépend pas explicitement du temps, alors

$$f' \frac{\partial\mathcal{L}}{\partial f'} - \mathcal{L} = \text{Cte}$$

Ceci est appelé l'identité de Beltrami. En mécanique, ceci n'est rien d'autre que la conservation d'énergie (exercice : le démontrer); elle est cependant de portée plus générale et facilite grandement les calculs dans les problèmes où la variable indépendante n'intervient pas explicitement, comme dans le problème du brachistochrone.

Exemple : Mécanique et loi de Newton. En mécanique analytique d'un point matériel, le lagrangien est $\mathcal{L} = T - V$, où T est l'énergie cinétique et V l'énergie potentiel. Si on se restreint au cas unidimensionnel où une particule est soumise à un potentiel $V(x)$, alors pour la particule de trajectoire $x(t)$,

$$\mathcal{L}(x, \dot{x}) = (1/2)m\dot{x}^2 - V(x).$$

Le premier terme de l'équation d'Euler-Lagrange est

$$\frac{\partial\mathcal{L}}{\partial x} = -\frac{dV}{dx}$$

Comme nous l'avons mentionné ci-dessus, le seul terme qui dépend de la première variable x est $V(x)$. Remplacer mentalement la deuxième variable \dot{x} par y dans l'expression du lagrangien avant la dérivation si cela vous dérange. La dérivation par rapport à la deuxième variable (\dot{x}) donne

$$\frac{\partial\mathcal{L}}{\partial\dot{x}} = m\dot{x}$$

et la dérivation par rapport au temps de cette dernière nous donne

$$\frac{d}{dt} \left[\frac{\partial\mathcal{L}}{\partial\dot{x}} \right] = m\ddot{x}$$

et l'équation d'E.L. s'écrit finalement

$$m\ddot{x} + dV/dx = 0$$

Ce qui est bien sûr la même chose que l'équation de Newton $F = ma$.

Allons un peu plus loin. Supposons que le potentiel est constant $V(x) = \text{Cte}$, s'est à dire que la particule se meut dans une région de l'espace où il n'est pas soumis à une force. On peut également dire que cette région de l'espace possède une symétrie d'invariance par translation : deux particules identiques placées à deux endroits différents de l'espace réagiraient exactement de même; dit encore autrement, nous n'avons aucune méthode pour déterminer où l'on se trouve dans l'espace. Dans ce cas, $\partial\mathcal{L}/\partial x = 0$, et l'équation d'E.L. s'écrit

$$\frac{d}{dt} \left[\frac{\partial\mathcal{L}}{\partial\dot{x}} \right] = 0$$

ou encore la quantité $p = \partial\mathcal{L}/\partial\dot{x} = Cte$. Or, p n'est autre chose que la quantité de mouvement $p = m\dot{x}$. Donc, la symétrie d'invariance par translation dans l'espace nous impose la conservation de la quantité du mouvement. C'est un concept extrêmement profond² : à chaque symétrie du lagrangien correspond une loi de conservation. La conservation de l'énergie est liée à la symétrie de translation dans le temps, la conservation du moment cinétique est associée à l'invariance par rotation autour d'un axe et ainsi de suite. En physique, une théorie n'est bien formée que si on peut la formuler sous forme variationnelle et chercher, par la symétrie sous-jacente du lagrangien, ses lois de conservation. Plus tard dans ce chapitre, nous verrons les formulations lagrangienne de l'électromagnétisme et de la mécanique quantique quand nous verrons comment traiter les champs.

1.3 Plusieurs degrés de libertés.

Parfois, souvent, le lagrangien dépend de plusieurs fonctions. Par exemple, une particule dans l'espace habituel possède trois degrés de libertés x_1, x_2, x_3 et $\mathcal{L} = \mathcal{L}(x_1, x_2, x_3, \dot{x}_1, \dot{x}_2, \dot{x}_3)$. Si les x_i sont les coordonnées cartésiennes, nous avons $\mathcal{L} = (m/2) \sum_i \dot{x}_i^2 - V(\mathbf{r})$. La démarche ci-dessus peut-être répétée mot à mot pour démontrer que nous aurons une équation d'E.L. pour *chaque* degrés de liberté :

$$\frac{\partial\mathcal{L}}{\partial x_i} - \frac{d}{dt} \left[\frac{\partial\mathcal{L}}{\partial \dot{x}_i} \right] = 0 \quad (1.6)$$

L'approche lagrangienne ne dépend pas du choix de coordonnées, bien évidemment. Il suffit donc de pouvoir écrire l'énergie cinétique et potentielle pour déduire les équations du mouvement de façon automatique.

Exemple 1 : mouvement dans un champ central. Il est facile de passer de l'expression de l'énergie cinétique en coordonnées cartésiennes (on pose $m = 1$) $T = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$ aux coordonnées sphériques, $T = \dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2$. L'énergie potentiel est $V(r)$, ce qui donne le lagrangien

$$\mathcal{L} = \dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2 - V(r)$$

En écrivant E.L. pour θ :

$$4r\dot{r}\dot{\theta} + 2r^2\ddot{\theta} - 2r^2 \sin \theta \cos \theta \dot{\phi}^2 = 0 \quad (1.7)$$

Nous remarquons que nous avons une symétrie : si nous avons choisi nos axes pour que à $t = 0$, $\theta = \pi/2$ et $\dot{\theta} = 0$ (ce qui revient à choisir le plan xy défini par le rayon vecteur de la particule et sa vitesse), alors $\theta(t) = \pi/2$ vérifie trivialement l'équation (1.7) : la particule reste dans le plan xy . Le lagrangien s'écrit plus simplement donc

$$\mathcal{L} = \dot{r}^2 + r^2\dot{\phi}^2 - V(r)$$

2. Ceci s'appelle le théorème de Noether, du nom de la mathématicienne allemande qui l'a formulé vers 1912.

1 Le calcul variationnel

Nous remarquons que ϕ n'intervient pas dans le lagrangien, le moment associé à cette variable se conserve donc :

$$p_\phi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = 2r^2 \dot{\phi} = \text{cte}$$

Ceci n'est rien d'autre que la conservation du moment cinétique, comme nous l'avions annoncé ci-dessus. Finalement, il faut écrire l'E.L. pour r et résoudre les équations différentielles résultantes. Le lecteur trouvera la solution de ses équations dans les livres de mécanique.

Exemple 2 : Pendule paramétrique. Soit un pendule de masse $m = 1$ restreint au plan xy (on suppose la gravité selon l'axe y) dont la longueur varie de façon périodique dans le temps : $\ell = \ell(t)$. Quelle est son équation de mouvement ?

Le système possède un seul degré de liberté. Nous choisissons l'angle θ que fait le pendule avec l'axe y comme ce degré. L'énergie cinétique est donnée par $T = (\dot{x}^2 + \dot{y}^2)/2$. Comme $x = \ell \sin \theta$ et $y = \ell \cos \theta$, $T = (\dot{\ell}^2 + \ell^2 \dot{\theta}^2)/2$. L'énergie potentielle est $U = gy = -g\ell \cos \theta$. Pour le premier terme de l'équation E.L., nous avons

$$\begin{aligned} \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right] &= \frac{d}{dt} [\ell^2 \dot{\theta}] \\ &= \ell^2 \ddot{\theta} + 2\ell \dot{\ell} \dot{\theta} \end{aligned}$$

Comme par ailleurs, $\partial \mathcal{L} / \partial \theta = -\partial U / \partial \theta = -g\ell \sin \theta$, nous trouvons l'équation du mouvement :

$$\ddot{\theta} + 2 \left(\dot{\ell} / \ell \right) \dot{\theta} + (g/\ell) \sin \theta = 0$$

Si ℓ est constante, nous retrouvons l'équation classique d'un pendule. Si maintenant ℓ oscille faiblement autour d'une position moyenne $\ell = \ell_0 + \ell_1 \sin \omega t$, $\ell_0 \gg \ell_1$, on peut démontrer qu'il y a résonance si la fréquence d'excitation ω est le *double* de la fréquence propre $\Omega = \sqrt{g/\ell_0}$; la phase de l'oscillateur est alors bloquée (à 0 ou π) sur la phase de l'excitation. Ce procédé est largement utilisé dans les circuits électroniques.

Vous avez sans doute remarqué avec quelle facilité l'on obtient les équations du mouvement. En aucun cas on ne doit chercher les forces de réaction des contraintes, comme dans le cas de la formulation Newtonienne. Il existe une interprétation géométrique extrêmement profonde de cette approche qui peut-être trouvée dans les livres de mécanique analytique. Nous ne continuerons pas plus le sujet ici.

Dérivation vectorielle.

Nous avons vu ci-dessus comment dériver les équations d'Euler-Lagrange quand le lagrangien dépend de plusieurs fonctions x_1, x_2, \dots, x_n d'une seule variable t . Très souvent cependant, les fonctions $x_i(t)$ sont les coordonnées d'un *vecteur* et il est vraiment dommage de devoir décomposer le vecteur en ses composantes. Ainsi, au lieu d'écrire $\mathbf{F} = m\mathbf{a}$, nous sommes amenés à écrire n équations du genre $f_x = m\ddot{x}$, $f_y = m\ddot{y}$, ... Nous pouvons *vectoriser* les équations E-L pour nous éviter cette gymnastique inutile et donner un sens géométrique à nos équations.

1 Le calcul variationnel

Pour cela, nous devons généraliser le concept de dérivation. Premièrement, remarquons que le lagrangien est toujours un *scalaire*. Un lagrangien avec un sens géométrique ne doit donc faire intervenir que des opérations sur les vecteurs dont le résultat est un scalaire intrinsèque, c'est à dire un scalaire dont le résultat ne dépend pas du système de coordonnées que nous avons choisi. Le meilleur exemple d'une telle opération est le *produit scalaire*.

Prenons par exemple, dans l'espace à trois dimensions, la fonction $f(\mathbf{r}) = \mathbf{r} \cdot \mathbf{r}$; si nous nous sommes équipées de coordonnées cartésiennes, ceci est un raccourci pour écrire

$$f(x, y, z) = x^2 + y^2 + z^2$$

Nous pouvons donc donner un sens à des expressions tel que $\partial_x f$. Pouvons nous donner un sens à l'expression $\partial f / \partial \mathbf{r}$? La réponse est évidemment oui si nous nous souvenons de la définition de la dérivée. Faisons un *petit* déplacement \mathbf{u} autour du point \mathbf{r} et mesurons la partie linéaire du changement dans la fonction f

$$\begin{aligned} df &= f(\mathbf{r} + \mathbf{u}) - f(\mathbf{r}) \\ &= (\mathbf{r} + \mathbf{u}) \cdot (\mathbf{r} + \mathbf{u}) - \mathbf{r} \cdot \mathbf{r} \\ &= 2\mathbf{r} \cdot \mathbf{u} + O(u^2) \end{aligned}$$

Nous pouvons donc effectivement écrire la partie linéaire comme

$$df = \left(\frac{\partial f}{\partial \mathbf{r}} \right) \cdot \mathbf{u}$$

où $(\partial f / \partial \mathbf{r})$ représente le vecteur $2\mathbf{r}$. Ceci est l'exact équivalent (et la généralisation) de la dérivée de la fonction d'une seule variable scalaire $f(x) = x^2$, où $f'(x) = 2x$. En analyse vectorielle, la quantité $(\partial f / \partial \mathbf{r})$ est souvent notée $\text{grad} f$ ou ∇f ³.

Nous pouvons généraliser le produit scalaire en utilisant les notations matricielle. Dans ce cas, le produit scalaire ci-dessus s'écrit $\mathbf{r}^T \mathbf{r}$ où \mathbf{r}^T est un vecteur colonne associé à \mathbf{r} . De façon encore plus générale, nous pouvons avoir des expressions du genre $f(\mathbf{r}) = \mathbf{r}^T A \mathbf{r}$ où A est une application bilinéaire, qu'on appelle plus communément un tenseur de rang 2. Il n'est pas difficile alors de voir que

$$df = (\mathbf{r}^T (A + A^T)) \cdot \mathbf{u}$$

Si l'on note $A_s = (A + A^T)/2$ la partie symétrique de l'application A , nous avons $\partial f / \partial \mathbf{r} = 2\mathbf{r}^T A_s$. Cette formulation est développée en détail dans le chapitre consacré aux tenseurs. Notons simplement que l'application A peut être utilisée pour gérer simplement les changements de coordonnées. A 2 dimensions, en coordonnées polaires par exemple, nous avons

$$A = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$$

3. Voir les chapitres sur les formes différentielles et le calcul tensoriel.

Exercice. Soit le lagrangien $\mathcal{L} = (1/2)m\dot{\mathbf{x}} \cdot \dot{\mathbf{x}} + V(\mathbf{x})$. Démontrer que les équations d'Euler-Lagrange s'écrivent

$$m\ddot{\mathbf{x}} = -\text{grad}V$$

1.4 Formulation lagrangienne et équation du mouvement d'un champ.

Le domaine où le calcul variationnel montre toute sa puissance est celui de l'obtention d'équations d'évolution d'un champ : la hauteur d'une corde vibrante, le champ électromagnétique, le champ des contraintes élastique dans un solide, le champ des amplitudes en mécanique quantique. Peu importe le champ, il suffit de pouvoir formuler un lagrangien, c'est à dire l'énergie cinétique moins l'énergie potentielle, et le tour est joué. La différence par rapport à ce que nous avons développée ci-dessus est minime. Jusque là, nous avons considéré des lagrangiens fonctions de plusieurs x_i , chacune de ces dernières fonction d'une variable t . Là, nous allons considérer des lagrangiens fonction de plusieurs ϕ_i , chacune de ces dernières fonction de plusieurs variables.

Exemple fondamental. Pour fixer les idées, considérons le cas de la corde vibrante fixée à ces deux extrémités ($x = 0$ et $x = L$). A chaque instant t , la hauteur de la corde à la position x est donnée par $\phi(x, t)$. Étant donnée la forme de la corde à l'instant t_0 ($\phi(x, t_0) = y_0(x)$) et t_1 ($\phi(x, t_1) = y_1(x)$), quelle est l'évolution de la corde qui minimise l'action ? La trajectoire de la corde est alors la surface reliant $y_0(x)$ à $y_1(x)$ dans le temps.

A un instant donnée t , l'énergie cinétique de la corde est donnée par la somme de l'énergie cinétique de tout ses points matériels,

$$T = \int_0^L \rho (\partial\phi/\partial t)^2 dx$$

et son énergie potentielle est l'énergie élastique des déformations

$$V = \int_0^L k (\partial\phi/\partial x)^2 dx \tag{1.8}$$

où ρ est la densité de la ligne et k sa constante élastique. Notant $c^2 = k/\rho$, l'action (le coût d'une trajectoire $\phi(x, t)$) est

$$S[\phi] = \int_{t_0}^{t_1} \int_0^L \left[(\partial\phi/\partial t)^2 - c^2 (\partial\phi/\partial x)^2 \right] dx dt$$

En considérant une variation $\epsilon g(x, t)$ autour de la trajectoire $\phi(x, t)$, nous trouvons qu'à l'ordre 1 en ϵ , la variation de l'action est

$$S[\phi + \epsilon g] - S[\phi] = \epsilon \int_{t_0}^{t_1} \int_0^L \left[\frac{\partial\phi}{\partial t} \frac{\partial g}{\partial t} - c^2 \frac{\partial\phi}{\partial x} \frac{\partial g}{\partial x} \right] dx dt$$

1 Le calcul variationnel

En faisant des intégrations par parties, nous trouvons

$$\delta S = \int_{t_0}^{t_1} \int_0^L \left[\frac{\partial^2 \phi}{\partial t^2} - c^2 \frac{\partial^2 \phi}{\partial x^2} \right] g dx dt$$

$\delta S = 0$ quelque soit la variation g , que si le terme entre [] est nul, c'est à dire

$$\frac{\partial^2 \phi}{\partial t^2} - c^2 \frac{\partial^2 \phi}{\partial x^2} = 0$$

qui est bien sûr l'équation d'onde bien connue.

Équation générale du mouvement de champ. Nous pouvons maintenant généraliser cette approche par des calculs aussi élémentaires que ceux ci-dessus, l'étape la plus technique étant une intégration par partie. Soit un domaine \mathcal{D} dans un espace à n dimensions, où nous cherchons à trouver l'extremum de la fonctionnelle

$$S[u] = \int_D \mathcal{L}(u, u_{,1}, \dots, u_{,n}, u, x_1, \dots, x_n) d^n x$$

où $u_{,i} = \partial u / \partial x_i$. La valeur de la fonction $u(x_1, \dots, x_n)$ sur la frontière du domaine \mathcal{D} étant fixé. L'équation d'Euler-Lagrange pour ce champ est donnée par

$$\sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial \mathcal{L}}{\partial u_{,i}} \right) - \frac{\partial \mathcal{L}}{\partial u} = 0$$

Remarquez que cette expression est une simple généralisation d'EL à une dimension : l'expression de dérivation du moment $(d/dx)(\partial \mathcal{L} / \partial u')$ est remplacée par la somme de toutes les dérivées partielles.

Tenseur énergie-impulsion. Nous pouvons pousser un peu plus loin. Nous avons vu l'égalité de Beltrami pour une fonction d'une variable : quand le lagrangien ne dépend pas de la variable indépendante

$$\frac{d}{dx} \left\{ y' \frac{\partial \mathcal{L}}{\partial y'} - \mathcal{L} \right\} = 0$$

ou si nous voulions l'écrire avec nos notations sophistiquées

$$\frac{d}{dx} \left\{ y_{,x} \frac{\partial \mathcal{L}}{\partial y_{,x}} - \mathcal{L} \right\} = 0$$

ou nous avons noté $y_{,x} = y' = dy/dx$. Si \mathcal{L} représente le lagrangien de la mécanique analytique, le terme entre crochet représente l'Hamiltonien, une quantité scalaire. Nous pouvons généraliser ce concept au lagrangien d'un champ $u(x_1, x_2, \dots, x_n)$ qui ne dépend pas explicitement des variables indépendantes. En notant ⁴

$$T_{ij} = u_{,i} \left(\frac{\partial \mathcal{L}}{\partial u_{,j}} \right) - \delta_{ij} \mathcal{L}$$

4. Rappelons que $\delta_{ij} = 0$ si $i \neq j$ et 1 sinon. Ceci est appelé le symbole de Kronecker

1 Le calcul variationnel

nous pouvons obtenir l'équivalent de l'identité de Beltrami pour un champ :

$$\sum_{j=1}^n \frac{\partial T_{ij}}{\partial x_j} = 0$$

la quantité T , qui joue le rôle de l'Hamiltonien pour le champ, est souvent appelée en physique le tenseur énergie-impulsion. Le lecteur intéressé pourra consulter des livres de géométrie pour voir la signification générale de ce tenseur⁵.

Exercices.

1. Pour une corde vibrante dans un champ de gravitation, nous devons ajouter le terme $V_g = \int_I \rho' \phi dx$ à l'énergie potentielle (1.8) où ρ' est le produit de la densité par l'accélération de la gravité. Dédurre l'équation du mouvement du champ dans ce cas. Même question si la corde se trouve dans un potentiel harmonique $V_h = \int_I \kappa \phi^2 dx$.
2. Calculer le tenseur énergie-impulsion pour la corde vibrante sans potentiel extérieur. Que représente T_{tt} ? Que veut dire l'identité de Beltrami dans ce cas? Que représente alors T_{tx} ?
3. Dédurre les équations du mouvement d'un champ dans un espace à n -dimensions, où l'expression de l'énergie potentielle (interne) est donnée par $V = \int_I k(\nabla\phi)^2 d\mathbf{x}$. Ceci généralise l'équation de la corde vibrante.
4. Dédurre l'équation du mouvement d'un champ dans un espace à n -dimensions anisotrope, où l'expression de l'énergie potentielle (interne) est donnée par

$$V = \int_I \sum_{i,j} a_{ij} \frac{\partial \phi}{\partial x_i} \frac{\partial \phi}{\partial x_j} d\mathbf{x}$$

où pour les coefficients, nous pouvons supposer $a_{ij} = a_{ji}$ (pourquoi?). Ce genre d'expression se rencontre fréquemment dans des problèmes comme l'élasticité des cristaux, où les déformations dans les différentes directions ne sont pas équivalentes.

1.5 Optimisation sous contraintes.

1.5.1 Les déplacements compatibles.

Oublions pour l'instant le problème de l'extremum d'une fonctionnelle, et revenons au problème beaucoup plus simple de l'extremum d'une fonction de plusieurs variables. Soit la fonction $f(x, y)$ dont nous cherchons l'extremum. Nous savons qu'au point extremum, les termes linéaires des variations doivent être nulle

$$df = f(x+h, y+k) - f(x, y) = \left(\frac{\partial f}{\partial x}\right) h + \left(\frac{\partial f}{\partial y}\right) k + O(h^2, k^2)$$

5. Notons quand même que cela ressemble à une expression du genre $\mathbf{div}T = 0$. Classiquement, la divergence est défini pour un vecteur et est fortement associée au flux de ce vecteur à travers une surface. Il n'est pas trop difficile de donner un sens au concept du flux d'un tenseur.

1 Le calcul variationnel

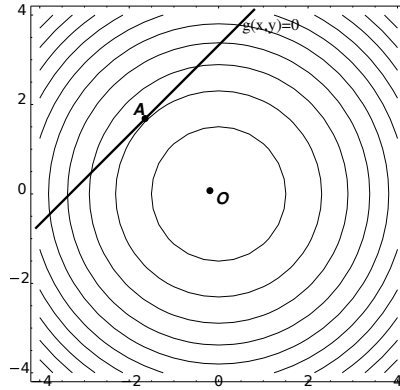


FIGURE 1.2 – La fonction $f(x, y) = x^2 + y^2$ est représentée par ses courbes de niveau. Le minimum absolu de cette fonction est le point $O = (0, 0)$. Le point le “plus bas” de la fonction qui doit également appartenir à la courbe $g(x, y) = 0$ est le point A .

et ceci quelque soit h, k : quelque soit la direction que l’on prend au point extremum, le terme linéaire de la variation doit être nulle.

Supposons maintenant que nous ajoutons une contrainte : nous ne cherchons pas l’extremum absolu de f , mais un point x^* qui satisfasse en plus à la contrainte $g(x, y) = 0$. Par exemple, nous savons que le minimum de la fonction $f(x, y) = x^2 + y^2$ est le point $(0, 0)$. Mais si nous contraignons notre point à se déplacer sur la courbe $y = ax + b$, $b \neq 0$, le point $(0, 0)$ n’est pas atteignable. On peut cependant chercher un point qui minimise f avec la contrainte donnée⁶. Dans le cas simple que nous sommes en train de traiter ici, on peut résoudre une coordonnée par rapport à l’autre et ramener la fonction à une seule variable (degré de liberté) : $f(x) = x^2 + (ax + b)^2$ et chercher maintenant le minimum de cette nouvelle fonction (où les contraintes ont été absorbées). En général cependant, la fonction contrainte $g(x, y) = 0$ est suffisamment compliquée pour qu’on ne puisse pas résoudre une des coordonnées par rapport aux autres.

Revenons à la définition générale : la variation linéaire de la fonction autour du point extremum (x^*, y^*) doit être nulle quelque soit le déplacement (h, k) autour de ce point. En présence des contraintes, nous relâchons cette exigence : il faut que la variation linéaire autour du point (x^*, y^*) soit nulle seulement pour les déplacements (h, k) compatibles avec les contraintes⁷. Les déplacements compatibles avec la contrainte $g(x, y) = 0$ sont données par

$$(\partial g / \partial x)h + (\partial g / \partial y)k = 0$$

Ceci nous donne une équation supplémentaire à considérer avec l’équation

$$df = (\partial f / \partial x)h + (\partial f / \partial y)k = 0$$

6. Imaginer que vous marchez en montagne sur un chemin, et vous vous intéressez au point le plus bas *sur ce chemin* et non pas en général.

7. En mécanique, cela est appelé la méthode des travaux virtuels

1 Le calcul variationnel

Nous sommes arrivés à un système de deux équations à deux inconnus *linéaires*. En résolvant k en fonction de h dans la première équation ; en l'injectant dans la deuxième équation ; et en exigeant maintenant que df soit nulle quelque soit h , nous obtenons⁸

$$(\partial f/\partial x)(\partial g/\partial y) - (\partial f/\partial y)(\partial g/\partial x) = 0 \quad (1.9)$$

Ceci est une condition supplémentaire sur l'extremum de la fonction f .

Exemple. Trouver l'extremum de la fonction $f(x, y) = x^2 + y^2$ avec la contrainte $g(x, y) = -ax + y - b = 0$. En écrivant la condition (1.9), nous obtenons

$$(2x)1 - (2y)(-a) = 0$$

ou autrement dit, $x + ay = 0$. En combinant cette équation avec la contrainte $y = ax + b$, nous trouvons les coordonnées du point extremum : $x^* = -ab/(1 + a^2)$, $y^* = b/(1 + a^2)$.

1.5.2 Les multiplicateurs de Lagrange.

Lagrange a introduit une méthode équivalente à ce que nous venons de voir ci-dessus. Nous voulons optimiser $f(x, y)$ avec la contrainte $g(x, y) = 0$. Lagrange introduit une variable supplémentaire, λ et considère maintenant le problème de l'optimisation (libre) de la fonction $F(x, y, \lambda) = f(x, y) - \lambda g(x, y)$. Cela nous donne

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0 \quad (1.10)$$

$$\frac{\partial F}{\partial y} = \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0 \quad (1.11)$$

$$\frac{\partial F}{\partial \lambda} = g(x, y) = 0 \quad (1.12)$$

L'équation (1.12) n'est rien d'autre que la contrainte et assure que la solution trouvée est bien conforme. Les deux premières équations (1.10,1.11) ne sont rien d'autre que la condition (1.9) si on y élimine λ .

Re-exemple. Minimisons la fonction $f(x, y) = x^2 + y^2$ avec la contrainte $g(x, y) = -ax + y - b = 0$. Nous introduisons la fonction $F(x, y) = x^2 + y^2 - \lambda(-ax + y - b)$. Cela nous donne

$$2x + \lambda a = 0$$

$$2y - \lambda = 0$$

8. On peut également dire que pour que le système $Ah + Bk = 0, Ch + Dk = 0$ ait une solution non trivial $h = k = 0$, il faut que le déterminant de la matrice

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

soit nulle.

Ce qui bien sûr nous donne $x+ay = 0$. Nous trouvons à nouveau le même point optimum ; de plus, si on le souhaite, on peut trouver également $\lambda^* = 2y^* = 2b/(1+a^2)$.

Le coefficient λ est appelé un multiplicateur de Lagrange. La méthode a une signification physique profonde : supposons que nous ayons un point matériel dans le potentiel $f(x, y)$ astreint à rester sur la courbe $g(x, y)$. Au point x^* , le point (qui est stable) est soumis à une force non-nulle de la part du potentiel $\mathbf{F} = (\partial_x f, \partial_y f)$. Pour qu'il puisse rester sur la position x^* , il faut que la courbe $g(x, y)$ exerce une force de réaction sur le point qui vaut justement $\lambda^*(\partial_x g, \partial_y g)$. Si on enlevait la contrainte mais qu'on soumettait le point matériel à cette force supplémentaire, il se mettrait exactement à la même position.⁹

1.5.3 Généralisation des multiplicateurs de Lagrange.

Les multiplicateurs de Lagrange se généralisent de façon naturelle aux fonctions de n variables soumis à m contraintes. Pour optimiser la fonction $f(x_1, \dots, x_n)$ soumise aux contraintes $g_1(x_1, \dots, x_n) = 0, \dots, g_m(x_1, \dots, x_n) = 0$, nous considérons la nouvelle fonction F de $n + m$ variables

$$F(x_1, \dots, x_n; \lambda_1, \dots, \lambda_m) = f(x_1, \dots, x_n) - \lambda_1 g_1(x_1, \dots, x_n) - \dots - \lambda_m g_m(x_1, \dots, x_n)$$

Nous chercherons l'extremum (sans contrainte) de la fonction F .

Exercice. Démontrer le cas général, en suivant le chemin déjà tracé.

La méthode des multiplicateurs de Lagrange se généralise très naturellement aux fonctionnelles. Si nous devons trouver le minimum de la fonctionnelle $S[y] = \int_I \mathcal{L}(y', y, x) dx$ avec la contrainte $\int_I g(y', y) dx = A$, il nous suffit d'introduire la fonctionnelle

$$S'[y; \lambda] = \int_I \{ \mathcal{L}(y', y, x) - \lambda g(y', y) \}$$

et de chercher le minimum de cette fonctionnelle par les techniques habituelles du calcul des variations (expliquer pourquoi nous pouvons omettre la constante A).

Exemple : Le problème isopérimétrique. Trouver la courbe fermée de longueur donnée qui maximise sa surface.

Supposons¹⁰ que nous pouvons représenter la courbe par une fonction $R(\theta)$. Nous cherchons à minimiser la fonctionnelle

$$S[R] = \int_0^{2\pi} (1/2)R^2(\theta) d\theta \tag{1.13}$$

9. Cette dualité force-contrainte a joué un grand rôle dans le développement de la mécanique, de sa formulation Newtonienne dans les années ~1680 à l'apparition du livre de mécanique analytique par Lagrange en ~1780. Lagrange était fier de n'avoir pas inclus une seule figure dans son livre. En mécanique, la méthode de Lagrange nous permet de calculer une position d'équilibre sans calculer explicitement les forces de réactions.

10. Ceci est une très grosse supposition. Le lecteur a intérêt à tracer quelques courbes qui ne sont pas représentable par une fonction $R(\theta)$ pour se rendre compte de la simplification que nous assumons derrière cette supposition.

1 Le calcul variationnel

avec la contrainte

$$\int_0^{2\pi} R(\theta) d\theta = L \quad (1.14)$$

Considérons la nouvelle fonctionnelle

$$S'[R(\theta); \lambda] = \int_0^{2\pi} \left\{ (1/2)R^2(\theta) - \lambda R(\theta) \right\} d\theta$$

Le terme entre $\{\}$ (que nous continuons d'appeler le lagrangien) ne contient même pas de dérivée de R ; l'équation d'Euler-Lagrange nous donne directement

$$R - \lambda = 0$$

c'est à dire $R = \lambda = Cte$. Ceci est bien un cercle. Pour trouver λ , nous utilisons la contrainte (1.14), qui nous donne $\lambda = L/2\pi$. La courbe qui minimise la surface est donc un cercle de rayon $R = L/2\pi$.

Exemple : distribution d'énergie dans un système isolé. Soit un système isolé ayant N particule ($N \rightarrow +\infty$) et soit $c(x)$ le nombre relatif de particules ayant une énergie dans l'intervalle $[x, x + dx]$. Comme le système est isolé, le nombre de particules et l'énergie totale sont conservés :

$$\int_0^{\infty} c(x) dx = 1 \quad (1.15)$$

$$\int_0^{\infty} xc(x) dx = T \quad (1.16)$$

où nous appelons T l'énergie moyenne par particule : $T = E_0/N$ avec E_0 l'énergie totale du système. Le théorème H de Boltzmann nous apprend que la fonction $c(x)$ doit être choisi de façon à ce que la quantité

$$H[c] = \int_0^{\infty} c(x) \cdot \log[c(x)] dx \quad (1.17)$$

soit extremum. La quantité $-H$ est souvent appelé entropie. Avec les deux contraintes (1.15,1.16), nous devons donc chercher l'extremum de la fonctionnelle

$$H'[c; \lambda, \mu] = \int_0^{\infty} c(x) \cdot \log[c(x)] dx - \lambda \int_0^{\infty} c(x) dx - \mu \int_0^{\infty} xc(x) dx$$

Le calcul est assez simple dans ce cas, puisque nous n'avons pas de dérivé dans la fonctionnelle. En cherchant la valeur de la fonctionnelle pour $H'[c + \epsilon g; \lambda, \mu]$ et en nous contentant d'ordre 1 en ϵ , nous trouvons

$$H'[c + \epsilon g; \lambda, \mu] = H'[c; \lambda, \mu] + \epsilon \int_0^{\infty} (1 + \log c - \lambda - \mu x) g(x) dx$$

1 Le calcul variationnel

la fonction $c(x)$ est un extremum de la fonctionnelle H' si les variations linéaires sont nulles $\forall g$, c'est à dire si

$$\log c = (\lambda - 1) + \mu x$$

ou autrement dit $c(x) = Ae^{-\mu x}$; l'utilisation des deux contraintes nous donne $A = -\mu = 1/T$. Autrement dit,

$$c(x) = (1/T)e^{-x/T}$$

1.5.4 Formulation variationnelle des systèmes Sturm-Liouville.

Rappelons qu'un système Sturm-Liouville est une équation aux valeurs propres

$$\alpha(x)y'' + \beta(x)y' + \gamma(x)y = \lambda y$$

Nous avons mentionné (et insisté) que si nous trouvons un poids $w(x)$ qui rend l'opérateur hermitien, alors les solutions des systèmes SL minimisent une certaine fonctionnelle. Nous avons vu que si une telle fonction poids existe, alors elle doit obéir à l'équation $(\alpha w)' = \beta w$ et le système peut alors s'écrire sous la forme alternative

$$\frac{d}{dx}(w\alpha y') + w\gamma y = \lambda w y$$

Remarquer que cette forme ressemble furieusement à une équation d'Euler-Lagrange. On peut faire le chemin inverse : minimiser la fonctionnelle

$$S[y] = \int_I (p(x)y'^2 + q(x)y^2) dx$$

avec la contrainte

$$\int_I w(x)y^2 dx = 1$$

Comme on peut le voir, nous pouvons formellement identifier l'équation d'Euler-Lagrange du système ci-dessus à un système SL en posant $p = w\alpha$, $q = -w\gamma$. A regarder de plus près, la formulation variationnelle d'un système SL, si on assimile la variable x au temps, ressemble beaucoup à un oscillateur harmonique avec une masse et une constante de ressort dépendant du temps. Ceci, comme nous l'avons mentionné, est le point de départ de l'approximation WKB.

1.6 Les conditions aux bords "naturelles" 1.

Revenons à notre formulation initiale de problème d'extremum. Nous voulons trouver l'extremum de la fonctionnelle

$$S[y] = \int_a^b \mathcal{L}(y', y, x) dx$$

avec les conditions aux bords fixée

$$y(a) = y_0 ; y(b) = y_1$$

1 Le calcul variationnel

Pour cela, nous avons écrit la variation de S en fonction d'une petite perturbation $g(x)$ pour obtenir

$$\delta S = \left[\frac{\partial \mathcal{L}}{\partial y'} g \right]_a^b + \int_a^b \left\{ \frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx} \frac{\partial \mathcal{L}}{\partial y'} \right\} g dx \quad (1.18)$$

et nous avons cherché dans quelles conditions, $\delta S = 0$ quelque soit $g(x)$. Nos conditions aux bords nous ont imposées $g(a) = g(b) = 0$, donc le premier terme est nul ; le deuxième terme nous donne les équations d'E-L.

Ceci dit, nous pouvons relâcher nos contraintes, et ne pas exiger que $y(a) = y_0$, $y(b) = y_0$. Le problème serait alors : parmi toutes les courbes entre a et b , trouver celle qui extrême la fonctionnelle. Dans ce cas, l'annulation de δS exige toujours l'annulation de l'intégrale, qui nous donnera comme d'habitude les équation d'Euler Lagrange, et l'annulation du terme de bord. Or, cette fois, comme les bords ne sont plus fixe, nous n'avons plus $g(a) = g(b) = 0$, pour annuler les termes de surface, nous devons exiger

$$\left. \frac{\partial \mathcal{L}}{\partial y'} \right|_{x=a,b} = 0$$

ce qui nous fournit deux nouvelles conditions en remplacement des conditions $y(a) = y_0$, $y(b) = y_0$. On peut bien sur mixer les conditions : fixer y en un bord et laisser y varier sur l'autre bord.

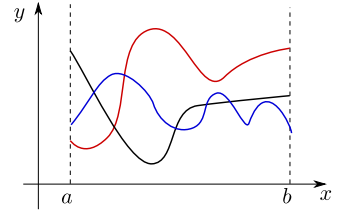


FIGURE 1.3 – Bords libres

Application 1 : Élasticité des poutres. Pour illustrer ce principe, considérons un lagrangien qui contient des dérivées d'ordre 2 : $\mathcal{L} = \mathcal{L}(y'', y', y, x)$. Dans ce cas, la variation s'écrit

$$\delta S = \left[\frac{\partial \mathcal{L}}{\partial y'} g - \frac{d}{dx} \frac{\partial \mathcal{L}}{\partial y''} g + \frac{\partial \mathcal{L}}{\partial y''} g' \right]_a^b + \int_a^b \left\{ \frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx} \frac{\partial \mathcal{L}}{\partial y'} + \frac{d^2}{dx^2} \frac{\partial \mathcal{L}}{\partial y''} \right\} g dx$$

Pour une poutre élastique soumise à une charge $f(x)$ par exemple, l'énergie s'écrit

$$\mathcal{E} = \int_0^L \{ B y''^2 - f(x)y \} dx$$

Si la poutre est encadrée (a), nous avons les conditions $y(0) = y(L) = 0$ et $y'(0) = y'(L) = 0$. Dans ce cas, le terme de bord est automatiquement nulle. Par contre, si la poutre est seulement posé (b), nous avons seulement $y(0) = y(L) = 0$. Pour trouver les deux autres conditions et annuler le terme de bord, nous devons avoir $\partial \mathcal{L} / \partial y'' = 0$, c'est à dire $y''(0) = y''(L) = 0$. Nous laissons au lecteur le soin de trouver les conditions aux limites pour le cas (c).

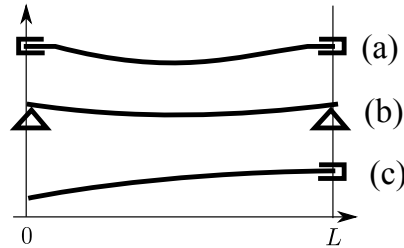


FIGURE 1.4 – poutres fixées.

Application II : angle de contact. Mettons un liquide en contact avec un support solide. Le liquide monte le long de la paroi solide, et par rapport à l'état de référence, l'énergie libre du système s'écrit :

$$F = \gamma_{lg}(\ell - L) + (\gamma_{sl} - \gamma_{sg})h$$

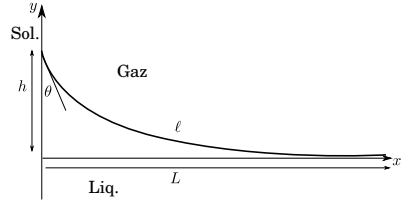
les coefficients γ_{lg}, \dots sont les tensions de surface liquide-gaz, ...; ℓ est la longueur de l'interface liquide-solide après la montée du liquide, L la longueur de cette même interface avant la montée, et nous supposons ces deux valeurs très grandes ($\rightarrow \infty$). h est la longueur de l'interface solide-liquide.

Repérons la courbe du liquide par la fonction $y(x)$. En mettant le problème sous forme variationnelle, nous avons

$$F[y] = \int_0^\infty \gamma_{lg} \left(\sqrt{1 + \dot{y}^2} - 1 \right) dx + (\gamma_{sl} - \gamma_{sg})y(0)$$

avec la contrainte de la conservation de la masse qui est

$$\int_0^\infty y dx = 0$$



Nous sommes en présence d'une optimisation sous contrainte, en présence de bords libre : la hauteur du liquide à $x = 0$ et $x = \infty$ n'est pas fixée. En ajoutant à $y(x)$ une variation $\epsilon g(x)$ où g ne s'annule pas sur les bords de l'intervalle, nous obtenons d'une part une équation d'Euler Lagrange comme d'habitude, et d'autre part une condition non-trivial sur le bord $x = 0$:

$$\left[\frac{\partial \mathcal{L}}{\partial \dot{y}} + (\gamma_{sl} - \gamma_{sg}) \right]_{x=0} = 0$$

Or,

$$\frac{\partial \mathcal{L}}{\partial \dot{y}} = \frac{\gamma_{lg} \dot{y}}{\sqrt{1 + \dot{y}^2}} = \gamma_{lg} \cos \theta$$

où θ est l'angle entre la tangente et l'axe y . On en déduit l'angle de contact solide-liquide

$$\cos \theta = \frac{\gamma_{sg} - \gamma_{sl}}{\gamma_{lg}}$$

La relation ci-dessus est connu sous le nom de la relation d'Young. Comme la mesure de l'angle de contact est facile, on l'utilise en général pour mesurer les tensions de surface.

1.7 Les conditions aux bords naturelles 2.

Dans le paragraphe précédent, l'abscisse x était fixée sur les deux bords, y ayant la liberté de bouger le long d'une droite *verticale*. Nous pouvons généraliser encore plus en relâchant cette contrainte et en permettant aux points sur les bords de se mouvoir le long

1 Le calcul variationnel

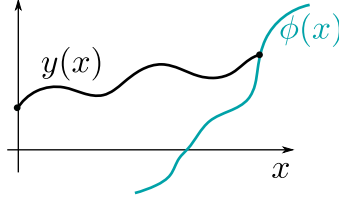


FIGURE 1.5 – Le bord droit de la courbe $y(x)$ est libre de se mouvoir le long d’une courbe $\phi(x)$.

d’une courbe quelconque. Supposons, pour la simplicité, que seulement le bord droit à $x = b$ peut se mouvoir le long d’une courbe $y = \phi(x)$. L’expression (1.18) que nous avons écrit dépend bien sûr du bord, et on peut écrire la variation causé autour de $y(x)$ par

$$S[y, b] - S[y + \epsilon g, b + db] = \left[\frac{\partial \mathcal{L}}{\partial y'} \epsilon g \right]_a^b + \int_a^b \left\{ \frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx} \frac{\partial \mathcal{L}}{\partial y'} \right\} \epsilon g dx + \int_b^{b+db} \mathcal{L}(y', y, x) dx \quad (1.19)$$

$$= \left[\frac{\partial \mathcal{L}}{\partial y'} \epsilon g \right]_a^b + \mathcal{L}(y', y, x) db + \{ \dots \} \quad (1.20)$$

Le terme entre crochet nous donne l’équation d’Euler-Lagrange comme d’habitude sur l’intervalle $[a, b]$. Occupons nous seulement des termes du bord. Nous devons avoir, puisque le point du bord se meut le long de la courbe $y = \phi(x)$,

$$y(b + db) + \epsilon g(b + db) = \phi(x + db)$$

Cela nous donne, à l’ordre 0 en db , $y(b) = \phi(b)$ bien sûr, et de plus, à l’ordre 1 :

$$\epsilon g(b) = (\phi'(b) - y'(b)) db$$

Finalement, la variation de S due au bord s’écrit

$$\delta S = \left\{ \frac{\partial \mathcal{L}}{\partial y'} (\phi'(b) - y'(b)) + \mathcal{L}(y', y, x) \right\} db$$

et la nullité de ce terme nous impose

$$y'(b) \frac{\partial \mathcal{L}}{\partial y'} - \mathcal{L} = \frac{\partial \mathcal{L}}{\partial y'} \phi'(b)$$

Le lecteur peut noter que le terme de gauche est souvent noté H , tandis que $\partial \mathcal{L} / \partial y'$ est souvent noté p . En mécanique analytique, on les appelle l’Hamiltonien et le moment. L’équation s’écrit donc, sur le bord,

$$H = p\phi'(b) \quad (1.21)$$

Résumons : Si $db = 0$, nous n'avons rien à faire de plus dans l'expression (1.20), et nous avons notre condition au bord du paragraphe précédent $p = 0$. Si $db \neq 0$, alors nous devons utiliser la condition plus générale (1.21). En particulier, si $\phi'(b) = 0$ (nous n'admettons que des déplacements le long de l'axe x), alors nous devons avoir

$$H = 0$$

Exemple : angle de contact d'une goutte posé sur un substrat solide non plane.

Exemple 2 : Brachistochrone générale.

1.8 Détour : éléments de géométries non-euclidiennes.

Le cinquième axiome d'Euclide est le suivant : "d'un point en dehors d'une droite, on ne peut tracer qu'une et une seule droite parallèle au premier". Pendant deux millénaires, les mathématiciens ont cru que ceci n'est pas un axiome, mais un théorème que l'on peut démontrer à partir des quatre premiers axiomes. A partir de 1830, il est devenu clair que le *cinquième* mérite le titre d'axiome et qu'on peut tout a fait formuler d'autres géométries qui acceptent d'autres axiomes. Nous allons étudier une version simplifiée de la géométrie riemannienne. Pour cela, nous devons donner un sens précis au mot "droite" : nous le définirons dorénavant comme le chemin le plus court entre deux points ; il est plus habituel alors d'appeler cela une géodésique. Supposons que nous avons muni notre espace bidimensionnel d'un système de coordonnées (x_1, x_2) . La distance entre deux points infiniment voisin est définie par

$$ds^2 = \sum_{i,j} g_{ij} dx_i dx_j$$

où $g_{ij} = g_{ij}(x_1, x_2)$ est appelé le tenseur métrique. Par exemple, si nous avons muni le plan euclidien de coordonnées polaires, $ds^2 = dr^2 + r^2 d\theta^2$, nous avons (en posant $x_1 = r$, $x_2 = \theta$), $g_{11} = 1$, $g_{22} = r^2$, $g_{12} = g_{21} = 0$.

Nous allons considérer dans la suite le cas très particulier où $g_{11} = 1$, $g_{i \neq j} = 0$ et $g_{22} = g^2(x)$, où g est une fonction quelconque¹¹. Le périmètre d'une courbe $y(x)$ reliant deux points est donnée par

$$\ell[y(x)] = \int_1^2 \sqrt{1 + g^2(x)y'^2} dx$$

et il est élémentaire de montrer, d'après les précédents paragraphes, que l'équation d'Euler-Lagrange nous donne immédiatement l'équation de la courbe :

$$y' = a/g\sqrt{g^2 - a^2}$$

11. Nous noterons par habitude les coordonnées (x, y) sans leur associer l'idée de coordonnées cartésiennes

1 Le calcul variationnel

où a est une constante d'intégration. Dans le cas des coordonnées polaire par exemple où $g(x) = x$, nous pouvons intégrer l'équation ci-dessus¹² et obtenir l'équation d'une droite $y = \alpha + \arccos(a/x)$ où a et α sont deux constante d'intégration. Pour vous convaincre que cela est effectivement le cas, Il suffit d'interpréter x comme r et y comme θ , faire un petit schéma et quelques manipulations d'angles .

Prenons le cas plus intéressant pour nous de $g(x) = \sin(x)$. A nouveau, l'intégration s'effectue sans difficulté¹³ et nous obtenons

$$y = \psi + \arccos(\cos \alpha \cot x)$$

où ψ et α sont deux constantes d'intégration. Nous voyons par exemple que pour $\alpha = \pi/2$, nous avons une famille de droites données par $y = \text{Cte}$. Prenons la géodésique $y = 0$ et considérons le point $P = (\pi/2, \pi/2)$ en dehors de cette droite. Toutes les droites traversant ce point doivent avoir le paramètre $\psi = 0$. Il n'est pas alors difficile de voir que toutes ces droites croisent la droite $y = 0$ au point $\cot x = 1/\cos \alpha$.

Nous venons de démontrer dans ce cas que *toutes* les droites traversant P croisent une droite ne contenant pas P ; cela est très différent du cinquième axiome d'Euclide. Le cas que nous venons de traiter correspond à la géométrie sphérique : sur la sphère unité, la distance entre deux points est donnée par $ds^2 = d\theta^2 + \sin^2 \theta d\phi^2$. Mais le point de vue de Riemann est beaucoup plus fondamental que cela : ce qui caractérise l'espace et qui lui donne sa substance est la donnée du tenseur métrique. Les habitants de la surface de la sphère unité ne peuvent pas voir qu'ils sont sur une sphère. Ils peuvent par contre visionner les géodésiques (en suivant les trajets des faisceaux de lumière) et déterminer la nature de leur espace en faisant des mesures par exemple de la somme des angles d'un triangle formé par trois géodésiques. C'est exactement dans ce cadre qu'Einstein a formulé sa théorie de la gravité en 1915, où les masses confèrent de la courbure à l'espace-temps.

A ajouter.

1. Discuter la jauge dans le Lagrangien et éventuellement le théorème de Noether.

Problèmes.

1. Surface minimum. Soit deux cercles concentriques de rayon à priori différents disposés l'un au dessus de l'autre à une hauteur h . Quelle est la forme de la surface d'aire minimum qui relie les deux cercles (fig. 1.6a) ?

2. Energie de courbure. Comment faut-il écrire les équations d'Euler-Lagrange si le lagrangien contient des dérivées secondes ? Plus spécifiquement, supposer que le lagrangien est de la forme $\mathcal{L} = y''(x)^2 + V(y)$. Généraliser ensuite au cas $\mathcal{L} = \mathcal{L}(y'', y', y, x)$.

12. il suffit d'effectuer le changement de variable $u = a/x$

13. Il suffit de poser $u = \cot x$

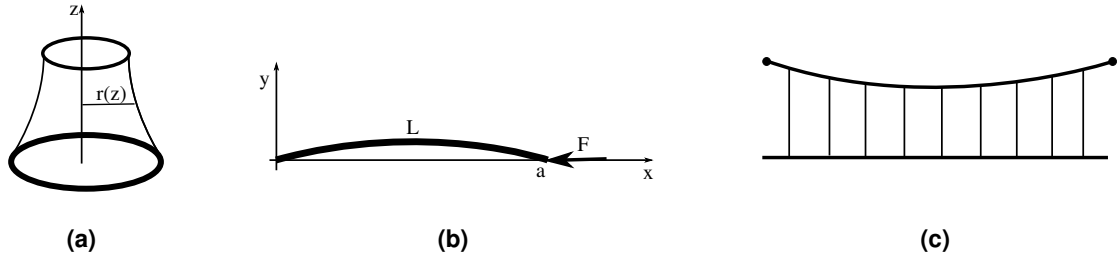


FIGURE 1.6 – (a) : surface minimum entre deux cercle ; (b) flambage d’une barre ; (c) pont suspendu à une chaînette.

3. Brachistochrone. Nous voulons minimiser

$$T = \int_0^{x_1} \sqrt{\frac{1 + y'^2}{y}} dx$$

En utilisant l’identité de Beltrami, démontrer que la courbe $y(x)$ doit obéir à l’équation

$$\frac{-1}{y(1 + y'^2)} = C$$

où C est une constante. Résoudre cette équation du première ordre en démontrant d’abord que

$$dx = dy \sqrt{\frac{y}{2a - y}}$$

où nous avons posé $1/C^2 = 2a$; intégrer cette dernière équation en posant $y = a(1 - \cos \theta)$. [Help : nous devons obtenir l’équation du cycloïde sous forme paramétrique.]

4. Elasticité 1-d. Soit une barre dont on repère les points (avant déformation) par la coordonnées x . On appuie sur la barre parallèlement à son axe ; les points de la barre se déplacent aux coordonnées x' (Figure 1.7). Nous appelons *déplacement* $u(x) = x' - x$ la fonction qui traque cette quantité. L’énergie élastique stockée dans la barre est proportionnelle
au carré du gradient de ce terme :

$$\mathcal{E} = \int_0^L (1/2)ku'(x)^2 dx$$

où k est la constante élastique de la barre (qu’on appelle également le module d’Young). La force a effectué le travail $W = Fu(L)$. L’énergie totale de la barre s’écrit donc

$$\mathcal{E} = \int_0^L \{ (1/2)ku'(x)^2 - Fu'(x) \} dx$$

Démontrer alors que $u(x) = ax$ où a est une constante à déterminer. Déterminer a en utilisant les conditions aux bords

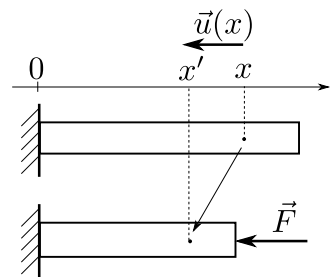


FIGURE 1.7 –

naturelles (section 1.6). En déduite la loi d'élasticité de Hook

$$F = Ku(L)$$

Que vaut K ? quelle est sa dépendance en L ?

5. Flambage d'une poutre. Appuyez sur une règle tenue verticalement sur une table ; au delà d'une certaine force, la règle flambe (fig.1.6b). Ceci est un problème extrêmement important de la résistance des matériaux et conditionne la conception des tours pour qu'elles ne s'écroulent pas (en flambant) sous leurs propres poids. Repérons la barre par son écart à la droite $y(x)$. En supposant faible l'écart de la barre par rapport à la droite, l'énergie de courbure de la barre est donnée par sa courbure locale $B(y''(x))^2$. Nous supposons l'extrémité de la courbe maintenu à $y = 0$, mais pouvant coulisser sur l'axe x et soumise à une force F . Pour trouver la configuration qui minimise l'énergie, nous devons donc trouver l'extremum de la fonctionnelle

$$S[y, a] = \int_0^a B(y''(x))^2 dx - F(L - a)$$

soumis à la contrainte

$$\int_0^a \{1 + (1/2)y'^2\} dx = L$$

Nous avons approximé ici l'élément de ligne $ds = \sqrt{1 + y'^2}$ par son développement de Taylor, en supposant les écarts à la ligne (et leurs dérivées) faible. Démontrez alors que pour $F > F_c$, la poutre droite n'est plus la solution optimum ; calculer F_c .

6. Flambage d'une poutre II. Nous pouvons formuler différemment le problème de flambage, en intégrant directement la contrainte dans le lagrangien. Pour cela il suffit d'utiliser un système de coordonnées plus adaptés que les coordonnées cartésiennes. Repérons un point le long de la poutre par sa longueur d'arc à partir de l'origine s , et par l'angle de la tangente à la courbe en ce point avec l'axe horizontal $\theta(s)$. Ce nouveau système de coordonnées est relié aux coordonnées cartésiennes par la relation

$$\begin{aligned} dx &= \cos \theta ds \\ dy &= \sin \theta ds \end{aligned}$$

La courbure en un point est simplement donnée par $d\theta/ds$. L'énergie s'écrit alors

$$S[\theta] = \int_0^L \{B(\theta')^2 + F \cos \theta\} ds - FL$$

que l'on peut traiter par les équation d'EL sans contraintes : la longueur d'arc s gère automatiquement la constance de la longueur totale de la poutre. Les coordonnées semi-intrinsèque (s, θ) sont très utilisées en géométrie différentielle. Noter de plus la grande similarité de l'action à celle de l'oscillation d'un pendule dans le champs de gravité.

7. Angle de contact. Trouver l'angle de contact d'une goutte de liquide déposée sur une surface solide.

8. isopérimétrique II. Chercher la courbe sous forme paramétrique $x = x(t)$, $y = y(t)$ avec la condition supplémentaire $x(1) = x(0)$ et $y(1) = y(0)$. Écrire au moins les équations d'Euler-Lagrange.

9. Isopérimétrique III. Quelle est la courbe de surface donnée qui minimise son périmètre ?

10. Équation de la chaînette. Une chaîne $y(x)$ est suspendu entre deux points distant de a . La longueur totale de la chaîne est L . Trouver l'équation de la chaîne. Help : A l'évidence, la chaîne doit minimiser l'énergie potentielle, avec une contrainte sur sa longueur. L'énergie potentielle est de la forme $\int \rho y ds$; on doit donc trouver le minimum de

$$H' = \int_{-a/2}^{a/2} \left\{ \rho y \sqrt{1 + y'(x)^2} - \lambda \sqrt{1 + y'(x)^2} \right\} dx$$

où λ est un multiplicateur de Lagrange. Obtenez également l'équation de la chaînette à laquelle on a suspendu un pont (fig.1.6c)

11. Trajectoire complexe. Soit une fonction $y(x)$ complexe ($\mathbb{R} \rightarrow \mathbb{C}$) dont l'action est définie par

$$S[y] = \int_I \mathcal{L}(y, y', y^*, y^{*'}; x) dx$$

Comme par exemple $\mathcal{L} = y' y^{*'} + k y y^*$. Obtenir les équations d'Euler-Lagrange de cette fonction. [Help : il faut démontrer que l'on peut considérer y et y^* comme deux composantes indépendantes, et obtenir une équation d'EL pour chacune].

12. Équation de Schrödinger. Nous cherchons la fonction complexe $\psi(x, t)$ qui optimise l'action associée au lagrangien suivant

$$\mathcal{L} = i\psi^* \partial_t \psi - a(\partial_x \psi^*)(\partial_x \psi) - V(r)\psi^* \psi$$

Trouver l'équation d'Euler Lagrange à laquelle obéit la fonction ψ . Dans la formulation ci-dessus, ψ et ψ^* ne jouent pas le même rôle. Pouvez vous donner une version plus symétrique de ce lagrangien ?

13. Champs électromagnétique dans le vide. Nous définissons un champ $A = (A_0, A_1, A_2, A_3)$ dans un espace à 4 dimensions que l'on repère à l'aide des coordonnées $x = (x_0, x_1, x_2, x_3)$ ¹⁴. Comme nous sommes parfois trop habitué à la séparation en espace (3d) et en temps, nous dissociions parfois les expressions ci-dessus en donnant des noms différents aux différents

14. Nous évitons pour l'instant les exposants et notons les coordonnées x_i au lieu de x^i

1 Le calcul variationnel

composants : nous notons par exemple $A = (\phi, -\mathbf{A})$ où nous appelons ϕ le potentiel et \mathbf{A} le potentiel vecteur ; de même, nous notons $x = (t, \mathbf{x})$ où le premier composant est appelé temps et les trois autres l'espace. Les équations d'électromagnétisme ont été formulé dans le cadre de cette séparation étrange et les différentes dérivées du champ ont reçu des noms différents. Par exemple, on appelle champ électrique le vecteur à 3 dimensions

$$\mathbf{E} = -\partial_t \mathbf{A} - \nabla \phi$$

et champ magnétique

$$\mathbf{H} = \nabla \times \mathbf{A}$$

Revenons à notre formulation générale. Le tenseur électromagnétique est défini par

$$F_{ik} = \frac{\partial A_k}{\partial x_i} - \frac{\partial A_i}{\partial x_k}$$

Ce tenseur est bien sûr anti-symétrique $F_{ik} = -F_{ki}$.

1. Donner l'expression du tenseur F en fonction des champs E_i et H_k .
2. Démontrer que pour trois indices i, j, k , la définition même du tenseur F impose

$$\frac{\partial F_{ij}}{\partial x_k} + \frac{\partial F_{jk}}{\partial x_i} + \frac{\partial F_{ki}}{\partial x_j} = 0$$

démontrer que les seules équations non-triviales sont celles où $i \neq j \neq l$ et cela nous donne 4 équations que nous pouvons regrouper en

$$\begin{aligned} \nabla \times \mathbf{E} &= -\partial_t \mathbf{H} \\ \nabla \cdot \mathbf{H} &= 0 \end{aligned}$$

qui ne sont rien d'autre que les deux premières équations de Maxwell.

Dans l'espace-temps relativiste, la "distance" ds entre deux points voisins est donnée par

$$ds^2 = dx_0^2 - (dx_1^2 + dx_2^2 + dx_3^2) = \sum_{i=0}^3 \epsilon_i dx_i^2$$

où $\epsilon_0 = 1$ et $\epsilon_{i \geq 1} = -1$. Il existe une différence entre une des composantes et les trois autres quant au signe qu'il faut utiliser pour l'élément d'arc. Cette différence apparaît obligatoirement dans toutes les expressions des lois physiques. En particulier, toutes les expressions quadratiques auront une forme similaire à l'expression de l'élément d'arc. Par exemple, l'action du champs électromagnétique est donnée par l'intégrale sur un volume du lagrangien suivant¹⁵ :

$$\mathcal{L} = - \sum_{i,j=0}^3 \epsilon_i \epsilon_j F_{ij}^2$$

¹⁵. Le signe moins n'a pas de conséquence pour nos calculs, mais assure que la solution trouvée est un minimum plutôt qu'un maximum de l'action.

1 Le calcul variationnel

Remarquer la similarité entre cette formulation et la formulation de la théorie de l'élasticité, où le champs A représente le vecteur déplacement et le tenseur F_{ij} le tenseur déformation. Ceci n'est pas un hasard ; Maxwell lui même imaginait l'éther comme un corps élastique et s'inspirait fortement de la théorie de l'élasticité.

1. Dédire que l'on peut mettre le lagrangien sous forme de $\mathcal{L} = \mathbf{E}^2 - \mathbf{H}^2$
2. Dédire les équations du champs

$$\sum_{j=0}^3 \epsilon_i \epsilon_j \frac{\partial F_{ij}}{\partial x_j} = 0$$

3. Mettre ces équations sous la forme plus usuelle des deux autres équations de Maxwell :

$$\nabla \times \mathbf{H} = \partial_t \mathbf{E} \ ; \ \nabla \cdot \mathbf{E} = 0$$