

4. Principe variationnel

Nous avons dérivé les équations de Euler-Lagrange à partir du principe d'Alembert (c'est-à-dire les équations de Newton). Il est également possible de le dériver à partir d'un principe variationnel, le principe de Hamilton, c'est-à-dire sous la forme d'une recherche d'un extremum d'une fonctionnelle. Une fonctionnelle généralise le concept de fonction qui dépend d'une ou plusieurs variables. Une fonctionnelle va dépendre d'une ou plusieurs fonctions. L'avantage d'un principe variationnel réside dans le fait qu'il s'affranchit de tout système de coordonnées.

4.1. Extrema libres et extrema liés

Avant d'aborder la problématique de la recherche d'un extremum d'une fonctionnelle, il peut être utile de rappeler la problématique de la recherche d'un extremum libre d'une fonction à plusieurs variables et d'introduire la problématique de la recherche d'un extremum sous contrainte que l'on résout classiquement par le biais des multiplicateurs de Lagrange. Soit $F(x_1, \dots, x_n)$ une fonction continue et dérivable de n variables $x_i, i = 1, \dots, n$ (Fig. 4.1). Lorsque l'on effectue une variation infinitésimale δx_i des variables cela implique une variation infinitésimale δF de la fonction. La recherche d'un extremum de cette fonction passe par la recherche d'un point stationnaire

$$\delta F = \sum_i \frac{\partial F}{\partial x_i} \delta x_i = 0, \quad \forall \delta x_i. \quad = \text{grad } F \cdot \vec{\delta x} \quad (4.1)$$

C'est-à-dire à partir de ce point et quelque soit la direction dans laquelle on opère une variation infinitésimale des variables, la variation correspondante de la fonction est nulle. On notera que la condition de stationnarité est nécessaire mais n'est en général pas suffisante pour que la condition d'extremum soit respecté. Il est en effet nécessaire d'imposer une condition supplémentaire sur les dérivées d'ordre plus élevé. Par exemple dans le cas d'une fonction d'une variable il est nécessaire d'imposer que la dérivée première change de signe au voisinage du point stationnaire sinon on se trouve dans le cas de figure d'un point d'inflexion. Si l'ensemble des variables x_i sont indépendantes alors l'Eq. (4.1) est équivalente au jeu de n équations

$$\frac{\partial F}{\partial x_i} = 0, \quad i = 1, \dots, n. \quad (4.2)$$

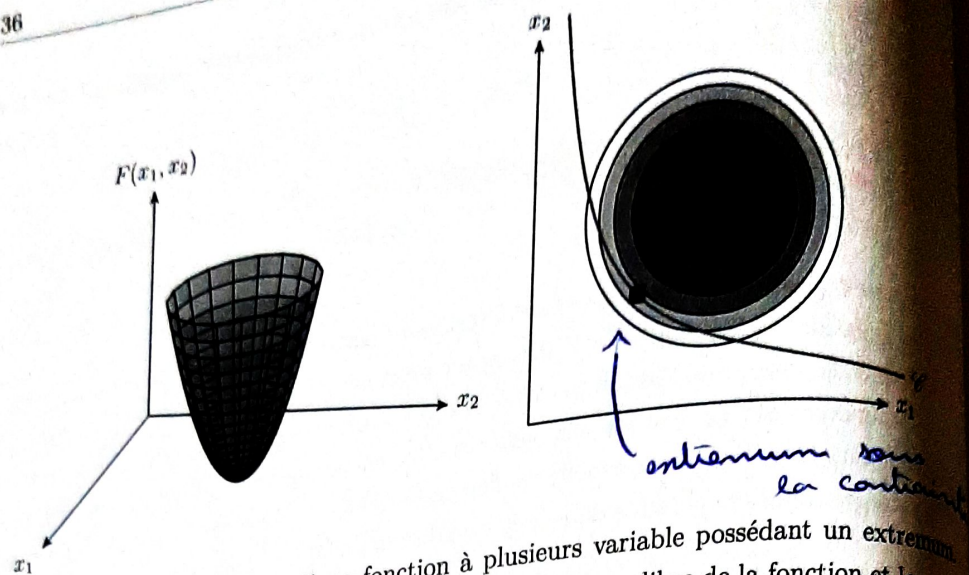


FIGURE 4.1. - Représentations d'une fonction à plusieurs variable possédant un extremum. Le cercle vert indique la position de l'extremum libre de la fonction et le cercle vert indique la position du l'extremum lié de la fonction le long de la courbe $\mathcal{C}$ représentée en rouge.

→ On remarquera que la solution de ce système d'équation correspond bien à un minimum de F si (mais pas seulement) la matrice Hessienne h_{ij} définie par

$$h_{ij} = \frac{\partial^2 F}{\partial x_i \partial x_j}, \quad (4.3)$$

est définie positive. En effet cette condition impose

$$\sum_{ij} h_{ij} \delta x_i \delta x_j > 0, \quad \forall \delta x_i \quad (4.4)$$

⇒ ensemble des valeurs propres positives

On montre facilement que cette condition est remplie si l'ensemble des valeurs propres de la matrice hessienne sont positives.

Nous allons maintenant considérer la recherche d'un extremum dans le cas où les variables sont reliés par une contrainte

$$f(x_1, x_2, \dots, x_n) = 0. \quad (4.5)$$

On parle d'extremum lié (Fig. 4.1). Pour simplifier nous allons supposer qu'il existe une seule relation de contrainte. Il sera aisé de généraliser à un nombre quelconque. Dans ce cas les variations δx_i ne sont plus indépendantes. En effet en effectuant une petite variation de la contrainte on voit qu'il existe une relation

$$\delta f = \sum_i \frac{\partial f}{\partial x_i} \delta x_i = 0, \quad \forall \delta x_i \quad (4.6)$$

on ne peut plus dériver par rapport à $x_i = 0$

Une méthode d'apparence simple pour résoudre ce problème est d'exprimer une des coordonnées en fonction des autres par exemple on peut choisir la coordonnée x_n . A partir de la contrainte on obtient une expression du type

$$x_n = g(x_1, \dots, x_{n-1}). \quad (4.7)$$

On peut alors injecter cette expression dans l'expression de la fonction F . Il ne reste plus qu'à faire la recherche d'un extremum de la fonction $F(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1}))$ dans l'espace de dimension $n-1$. Le problème de cette approche réside dans le fait qu'il est d'une part nécessaire de trouver l'expression de $g(x_1, \dots, x_{n-1})$ et d'autre part cela peut briser certaines symétries. Pourquoi choisir la coordonnée x_n à éliminer? Une approche alternative est l'utilisation des multiplicateurs de Lagrange.

Nous allons sommer les conditions d'extremum et de contrainte Eqs. (4.1) et (4.6) avec un facteur multiplicatif.

$$\underbrace{\delta F}_{=0} + \underbrace{\lambda \delta f}_{=0} = \sum_{i=1}^n \left(\frac{\partial F}{\partial x_i} + \lambda \frac{\partial f}{\partial x_i} \right) \delta x_i = 0, \quad \forall \delta x_i, \forall \lambda \quad (4.8)$$

Cette équation est vraie quelque soit la valeur du paramètre λ . Ce paramètre est appelé multiplicateur de Lagrange. On choisit la valeur de λ tel que

$$\frac{\partial F}{\partial x_n} + \lambda \frac{\partial f}{\partial x_n} = 0 \quad \text{et} \quad \frac{\partial f}{\partial x_n} \neq 0 \quad (4.9)$$

Nous allons voir que le choix de la coordonnée x_n n'a aucune importance et que si nous avions choisi une autre coordonnée par exemple x_1 le résultat aurait été le même. Si λ prend cette valeur alors l'Eq. (4.8) s'écrit

$$\sum_{i=1}^{n-1} \left(\frac{\partial F}{\partial x_i} + \lambda \frac{\partial f}{\partial x_i} \right) \delta x_i = 0, \quad \forall \delta x_i \quad (4.10)$$

Dans cette équation il n'y a que $n-1$ variables δx_i , ils sont donc nécessairement indépendants. On peut donc en conclure

$$\frac{\partial F}{\partial x_i} + \lambda \frac{\partial f}{\partial x_i} = 0, \quad i = 1, \dots, n-1. \quad (4.11)$$

S'il on combine Eq. (4.9) et Eq. (4.10) on en déduit

$$\frac{\partial F}{\partial x_i} + \lambda \frac{\partial f}{\partial x_i} = 0, \quad i = 1, \dots, n \quad (4.12)$$

Ceci démontre que le choix de la coordonnée x_n n'a aucune importance, le choix de tout autre coordonnée mène au même résultat. L'équation 4.12 est équivalent au problème de la recherche d'un extremum libre de la fonction $\bar{F}$ défini par

$$\bar{F} = F + \lambda f \quad (4.13)$$

par la
plus
grande
part
choisir
x_n?

sur l'hypersurface définie par la contrainte

exemple: $F(x,y) = x^2 + y^2 + xy$ *entierement libre: n*
 contrainte: $G(x,y) = x - y + 1 = 0$
 $F(x,y) = x^2 + y^2 + xy + \lambda(x - y + 1)$
 $\frac{\partial F}{\partial x} = 2x + y + \lambda = 0$ (1)
 $\frac{\partial F}{\partial y} = x + 2y - \lambda = 0$ (2)
 $x - y + 1 = 0$ (3)
 (1)+(2): $3x + 3y = 0 \Rightarrow x + y = 0$
 (3): $x - y = -1$
 $\Rightarrow y = \frac{1}{2}$
 $x = -\frac{1}{2}$
 $\lambda = 1/2$

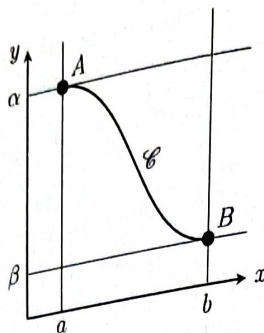


FIGURE 4.2. - Le problème du brachistochrone : Quelle est la courbe $\mathcal{C}$ passant par les points A et B tel que si un mobile glisse sans frottement le long de cette courbe le temps de parcours est minimum ?

L'idée de la méthode des multiplicateurs de Lagrange est d'étendre le problème en ajoutant une inconnue le paramètre λ . Ainsi il y a $n+1$ inconnues $x_1, \dots, x_n$ et λ . Pour déterminer l'ensemble des ces inconnues il y a $n+1$ équations $\partial \bar{F} / \partial x_i = 0, i = 1, \dots, n$ et $f = 0$.

Il est très simple de généraliser à K contraintes $f_\beta(x_1, \dots, x_n), \beta = 1, \dots, K$. Il est facile de démontrer que dans ce cas il suffit de rechercher un extremum pour la fonction

$$\bar{F} = F + \sum_{\beta=1}^K \lambda_\beta f_\beta.$$

Dans ce cas nous avons $n+K$ inconnues $x_1, \dots, x_n$ et $\lambda_1, \dots, \lambda_K$ que l'on peut déterminer en résolvant le système de $n+K$ équations : $\partial \bar{F} / \partial x_i, i = 1, \dots, n$ et $f_\beta = 0, \beta = 1, \dots, K$.

4.2. Calcul variationnel

4.2.1. Brachistochrone

dépend d'une fonction

Pour illustrer le concept de fonctionnelle nous allons introduire le problème classique de la brachistochrone (Fig. 4.2), c'est-à-dire la courbe de descente la plus rapide. Ce problème fut résolu pour la première fois par Jean Bernoulli en 1696. Soit une courbe $y = f(x)$ passant par les points A et B de coordonnées (a, α) et (b, β) si bien que $f(a) = \alpha$ et $f(b) = \beta$. Une masse initialement en A et sans vitesse initiale glisse sans frottement le long de la courbe. Quelle courbe minimise le temps de parcours ? On formalise ce problème en écrivant le théorème de l'énergie cinétique

$$\frac{1}{2}mv^2 = mg(\alpha - y), \quad (4.15)$$

si bien que la vitesse v s'écrit

$$v = \sqrt{\dot{x}^2 + \dot{y}^2} = \left| \frac{dx}{dt} \right| \sqrt{1 + y'(x)^2} = \sqrt{2g} \sqrt{\alpha - y(x)}, \quad (4.16)$$

$$\rightarrow \frac{dt}{dx} = \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2g} \sqrt{\alpha - y(x)}}$$

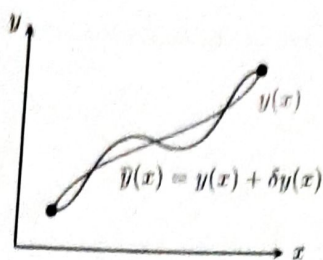


FIGURE 4.3. – La courbe $y(x)$ en noir représente la courbe qui donne un extremum de la fonctionnelle $I[y(x)]$. On effectue une petite variation de cette courbe en définissant la courbe $\bar{y}(x) = y(x) + \delta y(x)$

où $y' = dy/dx$. En supposant que $dx/dt > 0 \forall t$, le temps mis pour parcourir la courbe s'écrit

$$T = \int_a^b \frac{dt}{dx} dx = \frac{1}{\sqrt{2g}} \int_a^b \sqrt{\frac{1 + y'(x)^2}{\alpha - y(x)}} dx \quad (4.17)$$

Le temps de parcours $T = T[y(x)]$ est une fonctionnelle car sa valeur dépend de la courbe $y(x)$. La solution du brachistochrone n'est pas la ligne droite mais la courbe cycloïde. Cette solution est détaillée par l'utilisation de l'équation d'Euler dans les compléments du chapitre (4.A).

4.2.2. Équation d'Euler

Le problème variationnel consiste en la recherche d'une courbe qui produit un extremum pour une fonctionnelle donnée. Soit une fonctionnelle $I[y(x)]$ d'une courbe $y(x)$ définie sur l'intervalle $[a, b]$. Une forme assez générale de fonctionnelle s'écrit

$$I[y(x)] = \int_a^b F(y, y', x) dx, \quad (4.18)$$

où $F(y, y', x)$ peut dépendre de y , de y' et également explicitement de la position x ¹. On cherche un extremum de $I[y(x)]$ en imposant les conditions aux limites $y(a) = \alpha$ et $y(b) = \beta$. Pour trouver cet extremum nous allons effectuer une petite variation de la fonction $y(x)$ (Fig. 4.3). On pourra définir cette variation infinitésimale par un paramètre ϵ qui pourra tendre vers zéro par exemple

$$\delta y(x) = \epsilon \phi(x), \quad (4.19)$$

1. On peut également généraliser pour inclure des dérivées d'ordre supérieurs

tel que $\phi(a) = \phi(b) = 0$. La variation de la fonctionnelle donne

$$\begin{aligned}\delta I &= \delta \int_a^b F(y, y', x) dx = \int_a^b F(y + \delta y, y' + \delta y', x) dx - \int_a^b F(y, y', x) dx \\ &= \int_a^b \delta F(y, y', x) dx = \int_a^b \left(\frac{\partial F}{\partial y} \delta y + \frac{\partial F}{\partial y'} \delta y' \right) dx = \int_a^b \left(\frac{\partial F}{\partial y} \delta y + \frac{\partial F}{\partial y'} \frac{d}{dx} \delta y \right) dx, \quad (4.20)\end{aligned}$$

où nous avons utilisé la commutativité des variations

$$\delta y' = \delta \frac{dy}{dx} = \frac{d}{dx} \delta y. \quad (4.21)$$

En intégrant par partie cette dernière équation on obtient

$$\delta I = \int_a^b \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right) \delta y dx + \left[\frac{\partial F}{\partial y'} \delta y \right]_a^b = 0 \quad \text{car } \delta y(a) = \delta y(b) = 0 \quad (4.22)$$

Le terme de borne de l'intégration par partie est nécessairement nul. En effet on garde en mémoire que nous cherchons une variation de la fonction $y(x)$ tel que les bornes soient fixes. Si bien que

$$\delta y(a) = \delta y(b) = 0. \quad (4.23)$$

Ainsi la condition de stationnarité s'écrit

$$\delta I = \int_a^b \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right) \delta y(x) dx = 0 \quad \forall \delta y(x) \quad (4.24)$$

Dans cette équation on peut choisir librement la variation $\delta y(x)$, on peut par exemple choisir $\delta y(x)$ nul partout sauf en un point². On pourra donc se convaincre que la condition de stationnarité impose

$$\left| \frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = 0. \right. \quad \rightarrow \text{équation différentielle d'ordre 2} \quad (4.25)$$

Cette équation est l'équation d'Euler. On reconnaît la forme que l'on avait trouvé pour les équations d'Euler-Lagrange. Les équations de la mécanique classique peuvent donc s'exprimer sous la forme d'un principe variationnel.

4.3. Principe de moindre action de Hamilton

Le principe de Hamilton s'exprime de la manière suivante. Soit un système décrit par n coordonnées généralisées q_α , $\alpha = 1, \dots, n$ et soit l'action $S[q(t)]$ fonctionnelle de la trajectoire définie par

$$S = \int_{t_1}^{t_2} \mathcal{L}(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t) dt, \quad (4.26)$$

où $\mathcal{L}$ est la fonction de Lagrange. Le chemin $q_\alpha(t)$ pris par la nature entre les instants t_1 et t_2 est tel que l'action soit extremum. $\rightarrow$ telle que $S[q(t)]$ soit un extremum

2. C'est la distribution de Dirac $\delta(x - x_0)$, cf cours de mathématiques

Le principe de Hamilton se démontre assez facilement par la condition de stationnarité. La variation infinitésimale de l'action s'écrit

$$\begin{aligned} \delta S &= \int_{t_1}^{t_2} \delta \mathcal{L} dt = \int_{t_1}^{t_2} \sum_{\alpha} \left(\frac{\partial \mathcal{L}}{\partial q_{\alpha}} \delta q_{\alpha} + \frac{\partial \mathcal{L}}{\partial \dot{q}_{\alpha}} \delta \dot{q}_{\alpha} \right) dt \\ &= \int_{t_1}^{t_2} \sum_{\alpha} \left(\frac{\partial \mathcal{L}}{\partial q_{\alpha}} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_{\alpha}} \right) \right) \delta q_{\alpha} dt + \sum_{\alpha} \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_{\alpha}} \delta q_{\alpha} \right]_{t_1}^{t_2} \end{aligned} \quad (4.27)$$

Lors de la variation du chemin, on impose les conditions aux temps initiaux et finaux, $q_{\alpha}(t_1) = q_{\alpha,1}$ et $q_{\alpha}(t_2) = q_{\alpha,2}$. C'est à dire que parmi l'ensemble des trajectoires possible on se restreint à celles ayant un point initial et un point final bien déterminé. En mécanique pour déterminer le mouvement à partir de la solution générale, il faut $2n$ conditions initiales caractérisé par $2n$ constantes. On peut choisir $q_{\alpha}(t=0)$ et $\dot{q}_{\alpha}(t=0)$ ou bien $q_{\alpha}(t_1)$ et $q_{\alpha}(t_2)$. Ainsi la variation de l'action est donné par

$$\delta S = \int_{t_1}^{t_2} \sum_{\alpha} \left(\frac{\partial \mathcal{L}}{\partial q_{\alpha}} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_{\alpha}} \right) \right) \delta q_{\alpha} dt \quad (4.28)$$

Le principe de stationnarité implique nécessairement que l'intégrant est nul pour tous les temps et quelque soit α . On retrouve donc bien les équations d'Euler-Lagrange à partir du principe de Hamilton.

4.4. Coordonnées cycliques

Lorsque la fonction de Lagrange ne dépend pas explicitement d'une coordonnée q_{α} , on parle alors de coordonnée cyclique. Dans ce cas la dynamique le long de cette coordonnée peut-être résolue simplement. Pour simplifier, nous allons considérer un système à n degrés de liberté et nous allons supposer que la coordonnée q_n est cyclique. Il sera aisé de généraliser à d'autres coordonnées. Par définition on a

$$\frac{\partial \mathcal{L}}{\partial q_n} = 0, \quad \text{car } q_n \text{ cyclique} \quad (4.29)$$

si bien que le moment conjugué p_n à la coordonnée q_n est une quantité conservée que l'on note c_n .

$$p_n = \frac{\partial \mathcal{L}}{\partial \dot{q}_n} = c_n = \text{cste} \quad \text{par rapport au temps} \quad (4.30)$$

Puisque la fonction de Lagrange ne dépend pas de q_n on a

$$p_n = p_n(q_1, \dots, q_{n-1}, \dot{q}_1, \dots, \dot{q}_{n-1}, \dot{q}_n, t) = c_n. \quad (4.31)$$

A partir de l'expression de p_n on peut donc exprimer $\dot{q}_n$ en fonction des autres coordonnées et vitesses.

$$\dot{q}_n = f(q_1, \dots, q_{n-1}, \dot{q}_1, \dots, \dot{q}_{n-1}, c_n, t). \quad (4.32)$$

Cette équation s'intègre formellement

$$q_n = \int dt f(q_1, \dots, q_{n-1}, \dot{q}_1, \dots, \dot{q}_{n-1}, c_n, t).$$

Ainsi q_n peut être éliminé du problème, il suffit de résoudre les équations d'Euler-Lagrange pour les autres degrés de liberté en remplaçant q_n par son expression formelle et on en déduit alors la coordonnée q_n par intégration. On peut toutefois se demander si ce ne serait pas possible d'éliminer la coordonnée q_n avant de dériver les équations d'Euler-Lagrange directement dans un principe variationnel. Cette approche est possible mais il faut toutefois modifier ce principe. En effet, dans le principe variationnel nous avions supposé que les déplacements virtuels étaient indépendants. Clairement s'il on veut prendre en compte le fait que le moment conjugué soit constant impose une relation entre q_n et les autres coordonnées (Eq. 4.33). On définit ainsi la variation δ' tel que l'on se restreint aux trajectoires telles que $p_n = c_n = \text{cste}$, $\delta' = \delta_{p_n=c_n}$. La variation de l'action donne

$$\delta' S = \int_{t_1}^{t_2} \sum_{\alpha=1}^n \left(\frac{\partial \mathcal{L}}{\partial q_\alpha} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) \right) \delta' q_\alpha dt + \sum_{\alpha=1}^n \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \delta' q_\alpha \right]_{t_1}^{t_2}. \quad (4.34)$$

Le premier terme est nul car la solution du problème doit nécessairement respecter les équations d'Euler-Lagrange. Le second terme est le terme de bord. Pour $\alpha \neq n$ on impose que les variations aux temps $t = t_1$ et $t = t_2$ est nul $\delta' q_\alpha(t_1) = \delta' q_\alpha(t_2) = 0$. Ceci n'est cependant pas nécessairement vrai puisqu'il existe une relation entre q_n et les autres coordonnées. Ainsi la variation de l'action s'écrit

$$\delta' S = p_n \delta' [q_n]_{t_1}^{t_2} = \delta' \int_{t_1}^{t_2} p_n \dot{q}_n dt.$$

Nous allons faire passer le terme de droite de l'équation du coté gauche et on obtient

$$\delta' \left[\int_{t_1}^{t_2} \mathcal{L} - p_n \dot{q}_n \right] = 0$$

Dans le cas d'une coordonnée cyclique, on peut donc utiliser un principe variationnel modifié et basé sur la fonction de Lagrange modifiée $\mathcal{L}' = \mathcal{L} - p_n \dot{q}_n$.

$$\delta' S' = 0, \quad S' = \int_{t_1}^{t_2} \mathcal{L}' dt. \quad (4.35)$$

Il suffit alors de résoudre les équations d'Euler-Lagrange avec cette fonction de Lagrange modifiée en prenant soin d'avoir éliminé la variable $\dot{q}_n$ de l'expression. Ce problème est plus simple car la coordonnée q_n a été complètement éliminé avant de dériver les équations du mouvement.

Problème à force centrale Dans le cas du problème à force centrale (problème à deux corps par exemple), on sait que le mouvement est plan on peut donc écrire la fonction de Lagrange directement en fonction des coordonnées polaires.

$$\mathcal{L}(r, \dot{r}, \dot{\theta}) = \frac{1}{2} m \dot{r}^2 + \frac{1}{2} m r^2 \dot{\theta}^2 - V(r).$$

La coordonnée θ est une coordonnée cyclique. Le moment conjugué est donc une quantité conservée.

$$p_\theta = \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mr^2 \dot{\theta} = \text{cste} \quad (4.39)$$

Cette constante correspond à la conservation du moment cinétique. Une première erreur serait d'exprimer $\dot{\theta}$ en fonction de p_θ dans l'expression de la fonction de Lagrange et de dériver les équations d'Euler-Lagrange. Ceci donnerait un résultat faux ! Il est nécessaire de dériver les équations d'Euler-Lagrange puis d'éliminer la coordonnée $\dot{\theta}$. Alternativement s'il on veut éliminer la coordonnée θ avant de dériver les équations d'Euler-Lagrange il est nécessaire d'utiliser la fonction de Lagrange modifiée

$$\begin{aligned} \mathcal{L}'(r, \dot{r}) &= \frac{1}{2} m \dot{r}^2 + \frac{p_\theta^2}{2mr^2} - V(r) - p_\theta \frac{p_\theta}{mr^2} \\ &= \frac{1}{2} m \dot{r}^2 - \frac{p_\theta^2}{2mr^2} - V(r) \\ &= \frac{1}{2} m \dot{r}^2 - V_{\text{eff}}(r) \end{aligned} \quad (4.40)$$

On trouve donc facilement l'expression du potentiel effectif du problème à deux corps

$$V_{\text{eff}}(r) = V(r) + \frac{p_\theta^2}{2mr^2}. \quad (4.41)$$

L'équation d'Euler-Lagrange correspondante donne

$$m\ddot{r} = -\frac{dV_{\text{eff}}}{dr} \quad (4.42)$$

4.5. Principes de Maupertuis, Jacobi et Fermat

Le principe de moindre action de Hamilton permet de déduire les équations d'Euler-Lagrange. Ces équations donnent directement l'évolution temporelle du système. Dans bien des cas, il n'est cependant pas nécessaire (ou trop difficile) de déduire l'évolution temporelle. Il est parfois utile de rechercher directement la trajectoire en éliminant la variable temps. Le principe de moindre action de Maupertuis diffère du principe de Hamilton car il permet d'éliminer le temps. Ce principe historiquement précède le principe de Hamilton et constitue la première formulation d'un principe variationnel. Nous allons cependant ici le déduire du principe de Hamilton. Cette approche via notamment le principe de Jacobi permet une analogie entre la mécanique classique et l'optique avec le principe de Fermat.

Dans le principe de Hamilton, on traite le temps comme un paramètre. On peut aussi le considérer comme une variable indépendante. Ainsi on passe d'un système à n degré de liberté à un système à $n+1$ degrés de liberté, on parle d'espace des configurations étendu et la trajectoire dans cette espace est appelé ligne d'univers.

$$q_\alpha(t) \rightarrow \begin{cases} q_\alpha(\tau), \\ t(\tau), \end{cases} \quad (4.43)$$

où τ est un paramètre qui permet de décrire la trajectoire dans l'espace des configurations étendu. L'action s'écrit en fonction du paramètre τ selon

$$S = \int_{t_1}^{t_2} (q, \dot{q}, t) dt = \int_{\tau_1}^{\tau_2} \mathcal{L}(q, \frac{q'}{t'}, t) t' d\tau,$$

où $t' = dt/d\tau$ et $q'_\alpha = dq_\alpha/d\tau = \dot{q}_\alpha t'$. On note la fonction de Lagrange $\mathcal{L}(q, q', t, t')$ système étendu

$$\mathcal{L}(q, q', t, t') = \mathcal{L}(q, \frac{q'}{t'}, t) t'.$$

Cette fonction dépend en général des $2n+2$ variables q, q', t, t' . Calculons d'abord les moments conjugués $\bar{p}_\alpha$ associés aux variables q_α dans l'espace des configurations étendu.

$$\bar{p}_\alpha = \frac{\partial \mathcal{L}}{\partial q'_\alpha} = \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} = p_\alpha.$$

Les moments conjugués aux variables d'espace sont donc identiques dans l'espace des configurations et l'espace étendu. A la coordonnée temps t on peut associer un moment conjugué $\bar{p}_t$.

$$\bar{p}_t = \frac{\partial \mathcal{L}}{\partial t'} = \mathcal{L} - \sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \frac{q'_\alpha}{t'} = \mathcal{L} - \sum_\alpha p_\alpha \dot{q}_\alpha = -E$$

Le moment conjugué à la coordonnée temporelle est donc l'opposé de l'énergie.³ A partir de la fonction de Lagrange étendu on peut dériver $n+1$ équations d'Euler-Lagrange.

$$\frac{d}{d\tau} \left(\frac{\partial \mathcal{L}}{\partial q'_\alpha} \right) = \frac{\partial \mathcal{L}}{\partial q_\alpha}$$

$$\frac{d}{d\tau} \left(\frac{\partial \mathcal{L}}{\partial t'} \right) = \frac{\partial \mathcal{L}}{\partial t}$$

La première équation abouti aux équations d'Euler-Lagrange dans l'espace des configurations de dimension n et la seconde abouti à la variation temporelle de l'énergie au court du temps.

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) = \frac{\partial \mathcal{L}}{\partial q_\alpha}$$

$$\frac{dE}{dt} = - \frac{\partial \mathcal{L}}{\partial t}$$

On pourra remarquer que l'Eq. (4.51) se dérive des n équations d'Euler si bien qu'il semble que le système d'équation soit sous-déterminé. Nous avons introduit la variable τ , il est cependant nécessaire de la définir en introduisant par exemple la relation $t(\tau)$. Nous allons maintenant considérer le cas d'un système conservatif. Si bien que la fonction de Lagrange ne dépend pas explicitement du temps. Dans ce cas, dans l'espace des configurations

3. En mécanique quantique, le principe de correspondance énonce que le moment conjugué p_α associé à la coordonnée cartésienne x_α doit être remplacé par $p_\alpha \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial x_\alpha}$. Dans l'espace étendu, sachant que le moment conjugué au temps est l'opposé de l'énergie il faut donc remplacer l'énergie E par $E \rightarrow i\hbar \frac{\partial}{\partial t}$. En partant de l'expression de l'énergie $E = T + V$ on obtient alors l'équation de Schrödinger dépendante du temps.

4.5. Principes de Maupertuis, Jacobi et Fermat

45

étendu la coordonnée t est cyclique. Le moment conjugué (l'énergie) est une quantité conservée. On définit un principe variationnel tel que l'on se restreint aux trajectoires vu précédemment. Ainsi il faut utiliser la fonction de Lagrange modifiée.

$$\bar{\mathcal{L}} = \mathcal{L} - p_t t' = \mathcal{L} + Et' = \mathcal{L} t' + \left(\sum_{\alpha} p_{\alpha} q_{\alpha} - \mathcal{L} \right) t' = \sum_{\alpha} p_{\alpha} q'_{\alpha} \quad (4.52)$$

On aboutit donc au principe de Maupertuis

variation $\delta E W = 0$ $W = \int_{s_1}^{s_2} \sum_{\alpha} p_{\alpha} \frac{dq_{\alpha}}{d\tau} d\tau$ $W = \text{action réduite}$ (4.53)

Ce principe fait apparaître l'action W tel que définie par Maupertuis, c'est-à-dire le produit de la masse par la vitesse et par l'espace. On parle aujourd'hui d'action réduite. Dans le principe de Maupertuis le temps a disparu, il s'exprime uniquement en fonction du chemin parcouru dans l'espace des configurations.

On peut considérer le cas d'une particule unique de masse m dans un potentiel $U(\mathbf{r})$. La fonction de Lagrange s'écrit

$$\mathcal{L}(\mathbf{r}, \mathbf{v}) = \frac{1}{2} m \mathbf{v}^2 - U(\mathbf{r}) \quad (4.54)$$

Écrivons l'action réduite de Maupertuis en fonction de l'abscisse curviligne s .

$$W = \int_{s_1}^{s_2} \mathbf{p} \cdot \frac{d\mathbf{r}}{ds} ds = \int_{s_1}^{s_2} m \mathbf{v}^2 \frac{dt}{ds} ds \quad (4.55)$$

$\frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{dt} \frac{dt}{ds} = \frac{d\mathbf{r}}{dt} \frac{1}{v}$
 $\frac{dt}{ds} = \frac{1}{v}$ car vecteur unitaire

On peut exprimer la vitesse le long de l'abscisse curviligne en fonction de l'énergie cinétique $v = \dot{s} = \sqrt{2T/m}$, de plus on utilise la conservation de l'énergie et l'action réduite s'écrit

$$W = \sqrt{2m} \int_{s_1}^{s_2} \sqrt{E - U(\mathbf{r})} ds \quad \text{principe de Jacobi} \quad (4.56)$$

On en déduit donc le principe variationnel de Jacobi

$$\delta_E \int_{s_1}^{s_2} \sqrt{E - U(\mathbf{r})} ds = 0 \quad (4.57)$$

où on voit clairement ici que dans ce principe le temps a disparu. Ce dernier principe est complètement équivalent au principe de Fermat pour l'optique géométrique. Ce principe stipule en effet que la lumière se propage d'un point à l'autre tel que la durée du parcours soit extrémal. Soit T le temps de parcours d'un rayon lumineux

$$T = \int \frac{1}{v} ds = \frac{1}{c} \int n(\mathbf{r}) ds, \quad (4.58)$$

où $n(\mathbf{r})$ est l'indice de réfraction du milieu. Le principe de Fermat s'écrit de manière concise

$$\delta T = 0 \quad (4.59)$$

Donc: $n(\mathbf{r}) \leftrightarrow \sqrt{E - U(\mathbf{r})}$

3.3. Équations d'Euler-Lagrange

Le principe des travaux virtuels se limite aux problèmes de systèmes à l'équilibre. Pour décrire la dynamique, ce principe se généralise sous la forme du principe d'Alembert. Celui-ci s'écrit

$$\sum_i (\mathbf{F}_i - m_i \ddot{\mathbf{r}}_i) \cdot \delta \mathbf{r}_i = 0, \quad \forall \delta \mathbf{r}_i. \quad (3.23)$$

Tout comme dans le cas de l'équilibre les $\delta \mathbf{r}_i$ ne sont pas indépendants et il est nécessaire d'utiliser les coordonnées généralisées. Le principe d'Alembert s'écrit alors

$$\sum_\alpha \left(Q_\alpha - \sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \right) \delta q_\alpha = 0, \quad \forall \delta q_\alpha \quad (3.24)$$

Sachant que l'ensemble des δq_α sont indépendant on a donc

$$\sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} = Q_\alpha, \quad \alpha = 1, \dots, n \quad (3.25)$$

On remarquera que dans la définition des coordonnées cartésiennes en fonction des coordonnées généralisées (Eq. 3.8) la fonction $\mathbf{r}_i$ est une fonction de $\mathbb{Q} \times \mathbb{R}$ à valeur dans E^3 . La dépendance temporelle de cette fonction prend son origine d'une part dans la dépendance explicite par rapport au temps et d'autre part dans la dépendance temporelle de chaque coordonnées $q_\alpha(t)$. La vitesse cartésienne s'écrit donc

$$\dot{\mathbf{r}}_i = \frac{\partial \mathbf{r}_i}{\partial t} + \sum_\alpha \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \dot{q}_\alpha, \quad (3.26)$$

où la dérivée partielle par rapport au temps s'effectue à coordonnées généralisées fixées. On remarque que la vitesse $\mathbf{r}_i$ est une fonction des positions et vitesses généralisées et du temps

$$\mathbf{r}_i = \mathbf{r}_i(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t), \quad T\mathbb{Q} \times \mathbb{R} \rightarrow E^3. \quad (3.27)$$

On remarque également que les vitesses cartésiennes sont linéaires dans les vitesses généralisées. A partir de l'Eq. (3.26) on en déduit l'expression

$$\frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_\alpha} = \frac{\partial \mathbf{r}_i}{\partial q_\alpha}. \quad (3.28)$$

Cette équation caractérise la dépendance des vitesses cartésiennes en fonction des vitesses généralisées. Pour caractériser la dépendance des vitesses cartésiennes en fonction des coordonnées généralisées on pourra dériver l'Eq. (3.26) sachant que les variables q_α et $\dot{q}_\alpha$ sont indépendantes

$$\frac{\partial \dot{\mathbf{r}}_i}{\partial q_\alpha} = \frac{\partial^2 \mathbf{r}_i}{\partial t \partial q_\alpha} + \sum_\beta \frac{\partial^2 \mathbf{r}_i}{\partial q_\alpha \partial q_\beta} \dot{q}_\beta. \quad (3.29)$$

Cette dernière expression se met sous la forme simple

$$\frac{\partial \dot{\mathbf{r}}_i}{\partial q_\alpha} = \frac{d}{dt} \left(\frac{\partial \mathbf{r}_i}{\partial q_\alpha} \right). \quad (3.30)$$

En utilisant les Eqs. (3.28) et (3.30) on peut écrire la partie inertielle du principe d'Alembert (Eq. 3.25) sous la forme

$$\begin{aligned}\sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial \dot{q}_\alpha} &= \sum_i \frac{d}{dt} \left(m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial \dot{q}_\alpha} \right) - \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{r}_i}{\partial \dot{q}_\alpha} \right) \\ &= \sum_i \frac{d}{dt} \left(m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial \dot{q}_\alpha} \right) - \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_\alpha} \\ &= \frac{d}{dt} \left(\frac{\partial}{\partial \dot{q}_\alpha} \left(\sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 \right) \right) - \frac{\partial}{\partial \dot{q}_\alpha} \left(\sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 \right)\end{aligned}\quad (3.31)$$

On reconnaît ici l'expression de l'énergie cinétique

$$T = \sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2, \quad (3.32)$$

si bien que le principe d'Alembert s'écrit sous la forme simple

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) - \frac{\partial T}{\partial q_\alpha} = Q_\alpha, \quad \alpha = 1, \dots, n. \quad (3.33)$$

On remarquera que la dérivée totale par rapport au temps et la dérivée partielle par rapport aux vitesses généralisées ne commutent pas. En effet, on peut voir que

$$\frac{\partial}{\partial \dot{q}_\alpha} \frac{d}{dt} \mathbf{r}_i = \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \neq \frac{d}{dt} \frac{\partial}{\partial \dot{q}_\alpha} \mathbf{r}_i = 0. \quad (3.34)$$

Dans le cas de forces dérivant d'un potentiel $U = U(q_1, \dots, q_n, t)$ le principe d'Alembert s'écrit

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) - \frac{\partial \mathcal{L}}{\partial q_\alpha} = 0, \quad \alpha = 1, \dots, n \quad (3.35)$$

où $\mathcal{L}$ est la fonction de Lagrange ou Lagrangien définie par

$$\mathcal{L}(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t) = T(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t) - U(q_1, \dots, q_n, t). \quad (3.36)$$

On portera une attention particulière au signe négatif qui différencie la fonction de Lagrange de l'énergie mécanique $E_m = T + U$ du système. Les n équations (3.35) constituent les équations de Euler-Lagrange à la base de la mécanique Lagrangienne. Ces équations sont complètement équivalentes aux équations de Newton, leur expression toutefois permettra d'aborder un certain nombre de problème mécanique de manière beaucoup plus simple puisque les contraintes géométriques sont naturellement satisfaites dans ce formalisme. Les équations de Euler-Lagrange (3.35), sont données lorsque l'ensemble des forces dérivent d'un potentiel. Il est facile de généraliser ce principe dans le cas de forces qui ne dérivent pas d'un potentiel. Les équations de Euler-Lagrange s'écrivent alors

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) - \frac{\partial \mathcal{L}}{\partial q_\alpha} = Q_\alpha, \quad \alpha = 1, \dots, n \quad (3.37)$$

où Q_α ne contient que les forces ne dérivant pas d'un potentiel.

Une des propriétés importantes des équations d'Euler-Lagrange (Eq. 3.35) est qu'elles sont indépendantes du jeu de coordonnées généralisés utilisé. Une démonstration de cette propriété est donnée en annexe.

3.4. Exemples

Le pendule simple Pour illustrer l'utilisation des équations d'Euler-Lagrange, nous allons considérer le cas du pendule simple (Fig. 3.4a). Le système est décrit par l'angle θ que fait le pendule avec la verticale si bien que les coordonnées cartésiennes s'écrivent

$$\begin{aligned} x &= \ell \sin \theta, \\ y &= \ell \cos \theta. \end{aligned} \quad (3.38)$$

L'énergie cinétique et l'énergie potentielle du pendule exprimés en fonction de θ et $\dot{\theta}$ s'écrivent

$$T = \frac{1}{2} m \ell^2 \dot{\theta}^2, \quad = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2) \quad (3.39)$$

$$U = -mgl \cos \theta, \quad (3.40)$$

si bien que la fonction de Lagrange s'écrit

$$\mathcal{L}(\theta, \dot{\theta}) = \frac{1}{2} m \ell^2 \dot{\theta}^2 + mgl \cos \theta \quad (3.41)$$

L'équation de Euler-Lagrange donne directement l'équation du pendule

$$\ddot{\theta} + \omega_0^2 \sin(\theta) = 0 \quad (3.42)$$

où $\omega_0^2 = \frac{g}{\ell}$. Dans cette approche, il n'est pas nécessaire d'introduire des forces de contraintes.

3.5. Potentiel généralisé

Les forces dérivant d'un potentiel (Eq. 3.22) sont souvent appelées forces conservatives car un système n'ayant que ce type de forces a une énergie mécanique constante. Cependant il est à noter que ce n'est plus le cas si le potentiel devient explicitement dépendant du temps. Nous préférons donc ne pas utiliser le terme de force conservative mais plutôt le terme de forces dérivant d'un potentiel. La forme des équations de Euler-Lagrange (Eq. 3.35) suggère également de généraliser le concept de potentiel. Nous verrons en effet que certaines forces qui dépendent des vitesses peuvent dériver d'un potentiel généralisé $U(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t)$. Ce potentiel est défini par l'équation

$$Q_\alpha = \frac{d}{dt} \left(\frac{\partial U}{\partial \dot{q}_\alpha} \right) - \frac{\partial U}{\partial q_\alpha}, \quad (3.43)$$

Dans le cas de telles forces, les équations d'Euler-Lagrange (Eq. 3.35) peuvent être utilisés. On désignera ces forces comme mono-géniques pour reprendre le terme introduit par Cornelius Lanczos.

En effet, on a : $\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) - \frac{\partial T}{\partial q_\alpha} = Q_\alpha = \frac{d}{dt} \left(\frac{\partial U}{\partial \dot{q}_\alpha} \right) - \frac{\partial U}{\partial q_\alpha}$