Équation d'Euler–Lagrange $\dfrac{\partial L}{\partial \mathbf{q}} - \dfrac{d}{dt}(\dfrac{\partial L}{\partial \dot{\mathbf{q}}}) = 0$ et hamiltonien

Gilles Leborgne

21 septembre 2018

Table des matières

1 Exemples **2**
1.1 Longueur minimale d'une courbe dans $\mathbb{R}^2$ . . . 2
1.2 Brachistochrone . . . 3
1.3 Géodésique . . . 4
1.4 Dérivée particulaire et notations . . . 5

2 Théorème d'Euler–Lagrange dans $\mathbb{R}$: $\frac{\partial L}{\partial q} - \frac{D}{Dt}\frac{\partial L}{\partial \dot{q}} = 0$ **6**
2.1 Cadre . . . 6
2.2 L'action . . . 6
2.3 Théorème d'Euler–Lagrange dans $\mathbb{R}$ . . . 7
2.4 Cas $\frac{\partial L}{\partial t} = 0$: formule de Beltrami . . . 8
2.5 Equations de Lagrange inchangées par ajout d'une dérivée totale . . . 9
2.6 Exemple de la brachistochrone . . . 9

3 Théorème d'Euler–Lagrange dans $\mathbb{R}^n$: $\frac{\partial L}{\partial \mathbf{q}} - \frac{D}{Dt}\frac{\partial L}{\partial \dot{\mathbf{q}}} = 0$ **10**
3.1 Introduction . . . 10
3.2 Théorème d'Euler–Lagrange dans $\mathbb{R}^n$ . . . 11

4 Rappels de mécanique classique **13**
4.1 Principe fondamental . . . 13
4.2 Énergie cinétique . . . 13
4.3 Énergie potentielle . . . 14
4.4 Système conservatif . . . 14
4.5 Énergie totale . . . 14
4.6 Travail et force dérivant d'un potentiel . . . 15

5 Principe de moindre action d'Hamilton **16**
5.1 Principe fondamental vs équations d'Euler–Lagrange . . . 16
5.2 Principe de moindre action de Hamilton en cartésien . . . 16
5.3 Cadre de la dynamique des systèmes conservatifs usuels . . . 16
5.4 Principe de moindre action de Hamilton . . . 17

6 Équations de Lagrange **17**
6.1 Introduction . . . 17
6.2 Calcul . . . 18

7 Équations de Hamilton **19**
7.1 Introduction . . . 19
7.2 Hamiltonien : approche élémentaire . . . 19
7.2.1 Introduction en cartésien . . . 19
7.2.2 Dans un système de coordonnées . . . 20
7.3 Système de Hamilton . . . 20
7.4 Approche intrinsèque . . . 22
7.4.1 Introduction . . . 22
7.4.2 Hamiltonien . . . 23
7.4.3 Transformée de Legendre . . . 23

8 Forme symplectique et système de Hamilton **24**
8.1 Forme symplectique . . . 24
8.1.1 Espace des phases . . . 24
8.1.2 Forme symplectique usuelle . . . 24
8.2 Hamiltonien et forme symplectique canonique . . . 25
8.2.1 Espace des conditions initiales du système hamiltonien . . . 25
8.2.2 Forme symplectique canonique pour l'hamiltonien . . . 26
8.2.3 Forme différentielle canonique pour l'hamiltonien . . . 27

A Rappels **27**
A.1 Système de coordonnées et bases du système . . . 27
A.1.1 Système de coordonnées . . . 27
A.1.2 Base du système de coordonnées . . . 28
A.1.3 Base duale du système de coordonnées . . . 28
A.1.4 Dérivée partielle : notation . . . 29
A.2 Formes différentielles exactes . . . 29
A.2.1 Forme différentielle . . . 29
A.2.2 Différentielle extérieure . . . 30
A.2.3 Forme différentielle fermée et conservation . . . 30
A.2.4 Forme différentielle exacte . . . 31
A.2.5 Théorème de Poincaré . . . 31
A.3 Produit extérieur et différentielle extérieure . . . 32
A.3.1 Produit extérieur . . . 32
A.3.2 Expression dans un système de coordonnées . . . 32
A.3.3 Différentielle extérieure d'une forme différentielle . . . 33

Références bibliographiques **33**

1 Exemples

1.1 Longueur minimale d'une courbe dans $\mathbb{R}^2$

Dans le plan on veut trouver la courbe la plus courte joignant les points $A = (x_a, y_a)$ et $B = (x_b, y_b)$. (On veut montrer que c'est le segment de droite joignant A à B.)

Si on veut mettre ce problème sous forme mathématique, il faut commencer par comprendre l'énoncé : 1- c'est quoi une courbe ? 2- Sa longueur ? 3- Que veut dire "la plus courte" ?

1- Une courbe (sous forme paramétrée) est une fonction

$$\vec{r} : t \in [a,b] \to \vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \overset{\text{noté}}{=} \vec{x}(t) \in \mathbb{R}^2, \tag{1.1}$$

qui à chaque paramètre t (interprété par exemple comme un temps) associe une position $\vec{x}(t)$. Par exemple, un cercle de rayon R centré à l'origine est donné par $\vec{r}(t) = \begin{pmatrix} R\cos t \\ R\sin t \end{pmatrix}$ pour $t \in [0, 2\pi]$. Ici à chaque angle t (exprimé en radian) on associe la position $\vec{x}$ correspondant à cet angle sur le cercle.

L'image $\Gamma = \mathrm{Im}\vec{r}$ est également appelé une courbe (géométrique).

Et ici $\vec{r}(a) = A$ et $\vec{r}(b) = B$ sont imposés.

2- La longueur de la courbe paramétrée $\vec{r}$ est donnée par :

$$J(\vec{r}) = \int_{t=a}^{b} \sqrt{x'(t)^2 + y'(t)^2}\, dt \overset{\text{noté}}{=} J(\Gamma). \tag{1.2}$$

(Application de Pythagore $= \int ds = \int \sqrt{dx^2 + dy^2}$, voir cours de première année.) On montre immédiatement que J ne dépend pas du paramétrage choisi, i.e. si $\vec{q} : u \in [c,d] \to \vec{q}(u)$ est un autre paramétrage de Γ, on a $J(\vec{r}) = J(\vec{q})$ grâce à la formule de changement de variable dans les intégrales. D'où le nom $J(\Gamma)$ pour le résultat (et $J(\vec{r})$ pour le calcul).

3- Trouver la courbe la plus courte signifie "minimiser J", i.e. : parmi toutes les fonctions $\vec{r} \in C^1$, trouver une fonction $\vec{r}_m$ qui minimise $J(\vec{r})$:

$$J(\vec{r}_m) = \min_{\substack{\vec{r}\in C^1([a,b]);\mathbb{R}^2) \\ \vec{r}(a)=A,\ \vec{r}(b)=B}} J(\vec{r}).$$

Cas particulier : la courbe géométrique Γ est donnée sous forme graphe d'une fonction $f : x \to f(x)$, i.e. $\Gamma = \{(x, f(x)) \text{ pour } x \in [x_a, x_b]\}$. Ici $f(x_a) = y_a$ et $f(x_b) = y_b$ sont imposés. La courbe paramétrée associée est par exemple :

$$\vec{r} : x \in [x_a, y_a] \to \vec{r}(x) = \begin{pmatrix} x \\ f(x) \end{pmatrix} \tag{1.3}$$

(ici on a pris t=x). Et on doit donc minimiser :

$$J(f) = \int_{x=a}^{b} \sqrt{1 + f'(x)^2}\, dx \stackrel{\text{noté}}{=} J(\Gamma), \tag{1.4}$$

sur l'ensemble des fonctions $f \in C^1([x_a, x_b]; \mathbb{R})$ t.q. $f(x_a) = y_a$ et $f(x_b) = y_b$ sont imposés.

1.2 Brachistochrone

Dans $\mathbb{R}^2$ (résolu par Bernoulli en 1696).

Connaissant un point haut $A = (x_a, y_a)$ et un point bas $B = (x_b, y_b)$, on cherche à connaître la courbe joignant A à B telle que : si on lâche un objet de A le long de la courbe, l'objet arrive en un temps minimum en B. On suppose les frottements nuls et la gravité constante.

La solution est utilisée : cela ressemble aux rampes de skate-board ou roller, ou encore les "half-pipes" de snowboard. Mathématiquement ce sont les cycloïdes. Il s'agit de le prouver.

Calcul : par choix du repère $(0, x, y)$, on prend $y_a = 0$ et l'axe des y est dirigé vers le bas.

Soit $E = \{f \in C^1([x_a, x_b]; \mathbb{R}) : f(x_a) = y_a \text{ et } f(x_b) = y_b \text{ imposés}\}$. On cherche une solution dans E.

Calculons le temps mis pour aller de A à B le long du graphe de f. En un point $\vec{x} = (x, y) = (x, f(x))$ du graphe, l'élément de longueur est $ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + f'(x)^2}\, dx$. Si en ce point la vitesse est $||\vec{v}||$, on a $||\vec{v}|| = \dfrac{ds}{dt}$, i.e. $dt = \dfrac{ds}{||\vec{v}||}$. D'où le temps total :

$$T(f) = \int_{x=x_a}^{x_b} \frac{\sqrt{1 + f'(x)^2}\, dx}{||\vec{v}||}.$$

Il reste à exprimer $\vec{v}$ en fonction de x. La conservation de l'énergie (cinétique + potentielle) nous donne, "$\frac{1}{2}m||\vec{v}(x,y)||^2 - mgy = \text{constante}$" en tout point (x, y) de la courbe, et " constante = 0 " puisque vitesse et position initiales sont nulles. On a donc $||\vec{v}(x,y)|| = \sqrt{2gy}$. Donc :

$$T(f) = \int_{x=x_a}^{x_b} \sqrt{\frac{1 + f'(x)^2}{2gf(x)}}\, dx. \tag{1.5}$$

Pour montrer que la fonction f réalisant le minimum de T est une cycloïde, on utilisera l'équation d'Euler–Lagrange (ou plus simplement la formule de Beltrami dans ce cas), cf. paragraphes suivants, exemple 2.10.

Remarque 1.1 Autre présentation : on a (trivialité) :

$$T(f) = \int_{t=0}^{T(f)} dt. \tag{1.6}$$

On veut effectuer un changement de variable dans l'intégrale pour avoir une relation du type $T(f) = \int_{x=x_a}^{x_b} ...\, dx$.

Si $L(f)$ est la longueur de la courbe $\vec{r}(x) = \begin{pmatrix} x \\ f(x) \end{pmatrix}$, "graphe de f", la relation $dt = \dfrac{ds}{||\vec{v}||}$ donne :

$$T(f) = \int_{s=0}^{L(f)} \frac{ds}{||\vec{v}(\vec{r}(x(s))||}.$$

ds est l'élément de longueur (le paramétrage en s est dit intrinsèque), et si on prend le paramétrage cartésien $\vec{r}(x) = \begin{pmatrix} x \\ y = f(x) \end{pmatrix}$ de la courbe, l'élément de longueur est $ds = ||\vec{r}\,'(x)||\,dx = \sqrt{1+f'(x)^2}\,dx$. Et donc :

$$T(f) = \int_{x=x_a}^{x_b} \frac{\sqrt{1+f'(x)^2}\,dx}{||\vec{v}(x, f(x))||}.$$

Puis on se sert du principe de conservation de l'énergie. ▪

1.3 Géodésique

On se place ici dans $\mathbb{R}^3$. On considère une surface paramétrée, pour $\theta \in [\theta_0, \theta_1]$ et $\varphi \in [\varphi_0, \varphi_1]$:

$$\vec{r}(\theta, \varphi) = \begin{pmatrix} x(\theta, \varphi) \\ y(\theta, \varphi) \\ z(\theta, \varphi) \end{pmatrix},$$

et notons $S = \mathrm{Im}(\vec{r})$ l'ensemble des points de la surface géométrique image de $\vec{r}$.

Exemple 1.2 La sphère S centrée en $\vec{0}$ de rayon R est donnée par exemple par :

$$\vec{r}(\theta, \varphi) = R \begin{pmatrix} \cos\theta\cos\varphi \\ \sin\theta\cos\varphi \\ \sin\varphi \end{pmatrix}, \qquad \theta \in [0, 2\pi] \text{ (longitude)}, \quad \varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}] \text{ (latitude)}.$$

▪

Soit une courbe dessinée sur S, pour $u \in [u_0, u_1]$:

$$\vec{z}(u) = \vec{r}(\theta(u), \varphi(u)),$$

dessinée sur cette surface. On s'est implicitement donné des fonctions $\theta : u \in [u_0, u_1] \to \theta(u) \in [\theta_0, \theta_1]$ et $\varphi : u \in [u_0, u_1] \to \varphi(u) \in [\varphi_0, \varphi_1]$.

Définition 1.3 On se donne deux points $A, B \in S$ et une courbe sur S joignant A et B. Si cette courbe est de longueur minimale, elle est appelée une géodésique sur S.

Exemple 1.4 Sur la sphère précédente, si $\theta(u) = \theta_0$ constant et $\varphi(u) = u$ pour $u \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ alors $\vec{z}(u)$ décrit un méridien. ▪

Intéressons-nous ici aux courbes de la forme $\varphi = \varphi(\theta)$ (description 1-D explicite), et donc aux courbes, pour $\theta \in [a, b]$:

$$\vec{z}(\theta) = \vec{r}(\theta, \varphi(\theta)) \tag{1.7}$$

dessinées sur S. Le point de départ est $A = \vec{r}(a, \varphi(a))$ et le point d'arrivée est $B = \vec{r}(b, \varphi(b))$.

Exemple 1.5 Sur la sphère, on suppose donc que la latitude est fonction de la longitude : on se déplace soit vers l'est (exemple A = Angers et B = Berlin), soit vers l'ouest. ▪

On a :

$$\vec{z}\,'(\theta) = \frac{\partial\vec{r}}{\partial\theta}(\theta, \varphi(\theta)) + \frac{\partial\vec{r}}{\partial\varphi}(\theta, \varphi(\theta))\,\varphi'(\theta) = \begin{pmatrix} \frac{\partial x}{\partial\theta} + \frac{\partial x}{\partial\varphi}\varphi'(\theta) \\ \frac{\partial y}{\partial\theta} + \frac{\partial y}{\partial\varphi}\varphi'(\theta) \\ \frac{\partial z}{\partial\theta} + \frac{\partial z}{\partial\varphi}\varphi'(\theta) \end{pmatrix} (\theta, \varphi(\theta)), \tag{1.8}$$

vecteur "vitesse" qui est tangent à la courbe, et la longueur de la courbe est donnée par :

$$J(\vec{z}) = \int_{\theta=\theta_0}^{\theta_1} ||\vec{z}\,'(\theta)||\,d\theta = \int_{\theta=\theta_0}^{\theta_1} L(\theta, \varphi(\theta), \varphi'(\theta))\,d\theta, \tag{1.9}$$

où :

$$L(\theta,\varphi(\theta),\varphi'(\theta)) = (||\frac{\partial\vec{r}}{\partial\theta}(\theta,\varphi(\theta))||^2 + 2(\frac{\partial\vec{r}}{\partial\theta}(\theta,\varphi(\theta)),\frac{\partial\vec{r}}{\partial\varphi}(\theta,\varphi(\theta))\,\varphi'(\theta))_{\mathbb{R}^3} + ||\frac{\partial\vec{r}}{\partial\varphi}(\theta,\varphi(\theta))\,\varphi'(\theta)||^2)^{\frac{1}{2}},$$

notée abusivement :

$$L = (||\frac{\partial\vec{r}}{\partial\theta}||^2 + 2(\frac{\partial\vec{r}}{\partial\theta},\frac{\partial\vec{r}}{\partial\varphi}\,\varphi')_{\mathbb{R}^3} + ||\frac{\partial\vec{r}}{\partial\varphi}\,\varphi'||^2)^{\frac{1}{2}}.$$

La géodésique est la courbe qui minimise J (donnée par (1.9)) sur l'ensemble des fonctions $\vec{z}\in C^2$ t.q. $\vec{z}(a)=A$ et $\vec{z}(b)=B$.

On écrira $J(\vec{z})=\widetilde{J}(\varphi)$, cf. (1.9), et on cherchera $\varphi:\theta\to\varphi(\theta)$ qui minimise la valeur de l'intégrale $\widetilde{J}$.

1.4 Dérivée particulaire et notations

Rappel : si $g : x\in\mathbb{R}\to g(x)\in\mathbb{R}$ est une fonction dérivable en x alors sa dérivée en x est notée :

$$\frac{dg}{dx}(x) = g'(x) \qquad (\overset{\text{déf}}{=} \lim_{h\to 0}\frac{g(x+h)-g(x)}{h}).$$

Dans le cas particulier où la variable est une variable de temps, on note $g : t\in\mathbb{R}\to g(t)\in\mathbb{R}$, et on note également :

$$\frac{dg}{dt}(t) = g'(t) \overset{\text{noté}}{=} \dot{g}(t). \tag{1.10}$$

Rappel : soit $f : (t,x_1,...,x_n)\in[a,b]\times\mathbb{R}^n\to f(t,x_1,...,x_n)\in\mathbb{R}$ une fonction C^1 en $(t,\vec{x}) = (t,x_1,...,x_n)$. Ses dérivées partielles en $(t,\vec{x})$ sont données par :

$$\begin{cases} \dfrac{\partial f}{\partial t}(t,x_1,...,x_n) = \lim\limits_{h\to 0}\dfrac{f(t+h,x_1,...,x_n)-f(t,x_1,...,x_n)}{h}, \\ \dfrac{\partial f}{\partial x_1}(t,x_1,...,x_n) = \lim\limits_{h\to 0}\dfrac{f(t,x_1+h,x_2,...,x_n)-f(t,x_1,x_2,...,x_n)}{h}, \\ \vdots \end{cases}$$

Cas particulier où tous les x_i sont des fonctions de t. On pose, avec $\vec{x}(t)=(x_1(t),...,x_n(t))$:

$$g(t) = f(t,x_1(t),...,x_n(t)) \overset{\text{noté}}{=} f(t,\vec{x}(t)).$$

Définition 1.6 Alors on note :

$$(\dot{g}(t)=) \qquad \frac{dg}{dt}(t) \overset{\text{noté}}{=} \frac{Df}{Dt}(t,\vec{x}(t)) \overset{\text{noté}}{=} \frac{Df(t,\vec{x}(t))}{Dt} \overset{\text{noté}}{=} \frac{D}{Dt}(f(t,\vec{x}(t))) \overset{\text{noté}}{=} \frac{d}{dt}(f(t,\vec{x}(t))),$$

et la fonction $\dfrac{Df}{Dt} : [a,b]\times\mathbb{R}^n\to\mathbb{R}$ ainsi définie est appelée la “dérivée particulaire” de f, ou “la dérivée totale” de f le long de la trajectoire $t\to\vec{x}(t)$.

Donc, par dérivation de fonctions composées :

$$\begin{aligned} \frac{Df}{Dt}(t,\vec{x}(t)) &= \frac{\partial f}{\partial t}(t,x_1(t),...,x_n(t)) \\ &\quad + \frac{\partial f}{\partial x_1}(t,x_1(t),...,x_n(t))\,\frac{dx_1}{dt}(t) + ... + \frac{\partial f}{\partial x_n}(t,x_1(t),...,x_n(t))\,\frac{dx_n}{dt}(t) \\ &= \frac{\partial f}{\partial t}(t,\vec{x}(t)) + \vec{\nabla}f(t,\vec{x}(t))\cdot\vec{v}(t,\vec{x}(t)), \end{aligned}$$

où $\vec{\nabla}f(t,\vec{x}(t)) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(t,\vec{x}(t)) \\ \vdots \\ \frac{\partial f}{\partial x_n}(t,\vec{x}(t)) \end{pmatrix}$ est le “gradient en espace”, $\vec{v}(t,\vec{x}(t)) = \frac{d\vec{x}}{dt}[t) = \begin{pmatrix} \frac{dx_1}{dt}(t) \\ \vdots \\ \frac{dx_n}{dt}(t) \end{pmatrix}$ le vecteur vitesse, et “$\cdot$” le produit scalaire de $\mathbb{R}^n$. Et on note (une trajectoire $t\to\vec{x}(t)$ étant sous-entendue) :

$$\frac{Df}{Dt} = \frac{\partial f}{\partial t} + \vec{v}\cdot\vec{\nabla}f \overset{\text{noté}}{=} \frac{\partial f}{\partial t} + (\vec{v}\cdot\vec{\nabla})f \overset{\text{noté}}{=} \frac{D}{Dt}f \overset{\text{noté}}{=} \frac{d}{dt}f.$$

La notation $(\vec{v}\cdot\vec{\nabla})f$ car $\vec{v}\cdot\vec{\nabla} =^{\text{déf}} v_1\frac{\partial}{\partial x_1}+...+v_n\frac{\partial}{\partial x_n}$ donne $(\vec{v}\cdot\vec{\nabla})(f) = \vec{\nabla}f\cdot\vec{v}$.

Remarque 1.7 Pour disposer du gradient, on a besoin du produit scalaire canonique de $\mathbb{R}^n$. Si on veut se passer de l'utilisation d'un produit scalaire (et ainsi avoir une expression intrinsèque), on se sert de la définition de la différentielle de f en espace, ce qui donne :

$$\frac{Df}{Dt}(t,\vec{x}(t)) = \frac{\partial f}{\partial t}(t,\vec{x}) + df(t,\vec{x}).\vec{v}(t,\vec{x}),$$

où ici $df(t,\vec{x}) =^{\text{déf}} df_t(\vec{x})$ est, à t fixé, la différentielle de f_t en $\vec{x}$ (la différentielle en espace). Dans ce cas le "$\cdot$" désigne la "contraction" $df_t(\vec{x}).\vec{v} =^{\text{déf}} df_t(\vec{x})(\vec{v})$ (notation usuelle pour les applications linéaires). Voir poly "Mécanique : tenseurs 1ère partie, Algèbre linéaire". ▪

2 Théorème d'Euler–Lagrange dans $\mathbb{R}$: $\frac{\partial L}{\partial q} - \frac{D}{Dt}\frac{\partial L}{\partial \dot q} = 0$

2.1 Cadre

On munit $C^1([a,b];\Omega)$ de la norme $||f||_{C^1} = ||f||_\infty + ||f'||_\infty$, où $||g||_\infty = \sup_{t\in[a,b]}|g(t)|$, grâce à laquelle $C^1([a,b];\Omega)$ est un espace complet.

Soit deux réels α_0, α_1, soit $\Omega =]y_1, y_2[$ un intervalle ouvert borné non vide contenant α_0 et α_1, et on note :

$$E = \{q \in C^1([a,b];\Omega) : q(a) = \alpha_0,\ q(b) = \alpha_1\}, \tag{2.1}$$

l'ensemble des fonctions qui sont C^1 à image dans Ω dont les extrémités sont bloquées, faire un dessin. C'est un espace affine d'espace vectoriel associé noté :

$$E_0 = \{\varphi \in C^1([a,b];\Omega) : \varphi(a) = 0,\ \varphi(b) = 0\}. \tag{2.2}$$

On munit E et E_0 de la topologie de C^1 (topologie induite). On note :

$$B_0(0,h) = \{\varphi \in E_0,\ ||\varphi||_{C^1} < h\} = \{\varphi \in E_0,\ ||\varphi||_{C^1} < h\} \tag{2.3}$$

la boule ouverte centre q et de rayon h dans E_0, et pour $q \in E$ et $h > 0$:

$$B(q,h) = \{q\} + B_0(0,h) = \{f \in E : ||f-q||_{C^1} < h\} \tag{2.4}$$

la boule ouverte centre q et de rayon h dans E.

Soit une fonction :

$$L : \begin{cases} [a,b]\times\mathbb{R}\times\mathbb{R} & \to \mathbb{R} \\ \quad (X,Y,Z) & \mapsto L(X,Y,Z), \end{cases} \tag{2.5}$$

qu'on supposera C^2 dans la suite.

Définition 2.1 Quand la fonction L est considérée le long de "trajectoires" $t \in [a,b] \to (X,Y,Z) = (t, q(t), \dot{q}(t))$, où $q \in C^1([a,b];\Omega)$, la fonction L est appelée un lagrangien.

2.2 L'action

Pour L un lagrangien, on note $J : E \to \mathbb{R}$ la fonctionnelle :

$$J(q) = \int_{t=a}^{b} L(t, q(t), \dot{q}(t))\, dt. \tag{2.6}$$

Définition 2.2 $J(q)$ est appelée l'action du lagrangien L le long du chemin q.

Exemple 2.3 La longueur de la courbe explicite (1.3) est donnée par (1.4). Dans ce cas on pose $L(X,Y,Z) = \sqrt{1+Z^2}$. ▪

Exemple 2.4 Pour la brachistochrone (1.5) on pose $L(X,Y,Z) = \sqrt{\frac{1+Z^2}{2gY}}$. ▪

2.3 Théorème d'Euler–Lagrange dans $\mathbb{R}$

But : minimiser J sur E :

$$\text{trouver } q \in E \text{ t.q. } J(q) = \min_{f\in E} J(f),$$

où J est l'action donnée par (2.6) et E est donné par (2.1).

Sous les hypothèses de dérivabilité usuelles, on veut donc que la différentielle $dJ(q)$ (ou dérivée) de J en q soit nulle.

On rappelle que $dJ(q) : E_0 \to \mathbb{R}$ est l'application linéaire (quand elle existe) définie sur l'espace vectoriel E_0 associé à E, cf. (2.2), vérifiant (développement limité au premier ordre), pour tout $\varphi \in E_0$:

$$J(q+h\varphi) = J(q) + h\, dJ(q).\varphi + o(h) \qquad \text{au vois de } h = 0. \tag{2.7}$$

Et, quand $dJ(q)$ existe, pour $\varphi \in E_0$ la quantité suivante existe :

$$dJ(q).\varphi = \lim_{h\to 0} \frac{J(q+h\varphi) - J(q)}{h}, \tag{2.8}$$

et est appelée la dérivée de J en q dans la direction φ.

E étant ouvert, une condition nécessaire pour que J soit minimum en $q \in E$ est la condition d'Euler : les dérivées directionnelles $dJ(q).\varphi$ soient nulles pour toute fonction $\varphi \in E_0$ (variations nulles au voisinage de q).

Théorème 2.5 *On suppose $L \in C^1(\mathbb{R}^3;\mathbb{R})$ telle que ses dérivées $\frac{\partial L}{\partial Y}$ et $\frac{\partial L}{\partial Z}$ sont bornées. Pour que J ait un minimum en $q \in E$, il est nécessaire d'avoir :*

$$\int_a^b \frac{\partial L}{\partial Y}(t, q(t), \dot{q}(t))\, \varphi(t) + \frac{\partial L}{\partial Z}(t, q(t), \dot{q}(t))\, \dot{\varphi}(t)\, dt = 0. \tag{2.9}$$

Si de plus $L \in C^2(\mathbb{R}^3;\mathbb{R})$ et si q réalisant un minimum est C^2, alors, pour tout t (donc en tout point de la courbe q) :

$$\frac{\partial L}{\partial Y} - \frac{D\frac{\partial L}{\partial Z}}{Dt} = 0, \tag{2.10}$$

soit $\frac{\partial L}{\partial Y} = \frac{\partial^2 L}{\partial X \partial Z} + \frac{\partial^2 L}{\partial Y \partial Z}\dot{q} + \frac{\partial^2 L}{\partial^2 Z}\ddot{q}$. Appelée équation d'Euler. Notation abusive usuelle :

$$\frac{\partial L}{\partial q} - \frac{D}{Dt}\Big(\frac{\partial L}{\partial \dot{q}}\Big) = 0, \tag{2.11}$$

soit $\frac{\partial L}{\partial Y} = \frac{\partial^2 L}{\partial t \partial \dot{q}} + \frac{\partial^2 L}{\partial q \partial \dot{q}}\dot{q} + \frac{\partial^2 L}{\partial^2 \dot{q}}\ddot{q}$.

Preuve. Montrons que (2.10) est une condition nécessaire. Soit $q \in E$ et $\varphi \in E_0$ fixés. Soit :

$$F(h) = J(q+h\varphi) = \int_{t=a}^b L(t, (q+h\varphi)(t), (q+h\varphi)'(t))\, dt, \tag{2.12}$$

intégrale qui dépend du paramètre h. On a $dJ(q).\varphi = \lim_{h\to 0} \frac{F(h)-F(0)}{h} = F'(0)$, dérivée de F en 0, cf. (2.8). La question est : peut-on dérivée sous le signe $\int_a^b$? Si oui on a :

$$\begin{aligned} F'(h) &= \int_{t=a}^b \frac{d}{dh}(L(t, (q+h\varphi)(t), (q+h\varphi)'(t)))\, dt \\ &= \int_{t=a}^b \frac{\partial L}{\partial Y}(t, (q+h\varphi)(t), (q+h\varphi)'(t))\, \varphi(t) + \frac{\partial L}{\partial Z}(t, (q+h\varphi)(t), (q+h\varphi)'(t))\, \varphi'(t)\, dt, \end{aligned} \tag{2.13}$$

et la condition d'Euler $F'(0) = dJ(q).\varphi = 0$ donne (2.9).

Pour pouvoir dériver sous le signe $\int_a^b$, vérifions les hypothèses du théorème de convergence dominée de Lebesgue. Soit $g(h,t) = L(t, (q+h\varphi)(t), (q+h\varphi)'(t))$ l'intégrant dans $F(h)$, cf. (2.12). On a :

$$\begin{aligned} |\frac{\partial g}{\partial h}(h,t)| &= |\frac{\partial L}{\partial Y}(t, (q+h\varphi)(t), (q+h\varphi)'(t))\, \varphi(t) + \frac{\partial L}{\partial Z}(t, (q+h\varphi)(t), (q+h\varphi)'(t))\, \varphi'(t)\, dt| \\ &\le ||\frac{\partial L}{\partial Y}||_\infty ||\varphi||_\infty + ||\frac{\partial L}{\partial Z}||_\infty ||\varphi'||_\infty \overset{\text{noté}}{=} C, \end{aligned} \tag{2.14}$$

car L est C^1 de dérivées bornées et φ est C^1 sur $[a,b]$. Et l'inégalité est vraie pour tout $h \in]-h_0, h_0[$, où $h_0 > 0$ est t.q. $h_0\varphi \in E_0$. Comme C est une fonction constante, elle est L^1-intégrable sur $[a,b]$, et le théorème est vérifié : F est dérivable dans $]-h_0, h_0[$ et on peut dériver sous $\int$, d'où (2.13), d'où (2.9) en prenant $h = 0$.

Supposant de plus L et φ C^2, par intégration par parties on a :

$$\int_a^b \frac{\partial L}{\partial Z}(t, q(t), \dot{q}(t))\, \dot{\varphi}(t)\, dt = -\int_a^b \frac{D}{Dt}\Big(\frac{\partial L}{\partial Z}(t, q(t), \dot{q}(t))\Big)\, \varphi(t)\, dt + [\frac{\partial L}{\partial Z}(t, q(t), \dot{q}(t))\, \varphi(t)]_a^b,$$

où le terme de bord est nul puisque $\varphi(a) = \varphi(b) = 0$. Donc pour que q réalise un minimum il faut :

$$\int_a^b \left(\frac{\partial L}{\partial Y}(t, q(t), \dot{q}(t)) - \frac{D}{Dt}\Big(\frac{\partial L}{\partial Z}(t, q(t), \dot{q}(t))\Big) \right) \varphi(t)\, dt = 0, \tag{2.15}$$

ce pour toute fonction $\varphi \in E_0 \bigcap C^2$. Comme C^2 est dense dans C^0, que $\Big(...\Big)$ est $L^1([a, b])$ car C^0, et que « si $\int_a^b \Big(...\Big)(t)\varphi(t)\, dt = 0$ pour tout $\varphi \in C^0$ alors $\Big(...\Big) = 0$ », on déduit que $\Big(...\Big) = 0$, i.e. (2.10). ▪

Remarque 2.6 Preuve de (2.10) "à la physicien" : quand on bouge un peu la courbe q, on a la courbe $q = \delta q$, et l'intégrale $J(q)$ devient $J(q) + \delta J$ où :

$$\delta J = \delta \int L(t, q, \dot{q})\, dt = \int (\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q})\, dt, \tag{2.16}$$

puis $\delta \dot{q} = \delta \frac{dq}{dt} = \frac{d}{dt}\delta q$, d'où (2.9)... Puis intégration par partie comme précédemment. ▪

Exemple 2.7 Longueur d'une courbe en 2-D, cf. (1.4) : on a $L(X, Y, Z) = \sqrt{1 + Z^2} = (1 + Z^2)^{\frac{1}{2}}$, avec donc $\frac{\partial L}{\partial X}(X, Y, Z) = \frac{\partial L}{\partial Y}(X, Y, Z) = 0$, avec $\frac{\partial L}{\partial Z}(X, Y, Z) = Z(1 + Z^2)^{-\frac{1}{2}}$, et donc $\frac{\partial^2 L}{\partial Z^2}(X, Y, Z) = (1 + Z^2)^{-\frac{3}{2}}$. Et l'équation d'Euler est :

$$0 - 0 - 0 - f''(x)\, (1 + f'(x)^2)^{-\frac{3}{2}} = 0,$$

dont la solutiont $f =^{\text{noté}} q$ vérifie $q''(x) = 0$. Donc $q(x) = c_1 x + c_0$ où les c_i sont calculées à l'aide des conditions aux limites $q(a)$ et $q(b)$ données : la courbe de longueur minimale est un segment de droite. ▪

2.4 Cas $\frac{\partial L}{\partial t} = 0$: formule de Beltrami

Cas $L(X, Y, Z) = L(Y, Z)$, i.e. cas $\frac{\partial L}{\partial X} =^{\text{noté}} \frac{\partial L}{\partial t} = 0$. Ici (2.10) devient : trouver q t.q. :

$$\frac{\partial L}{\partial q} - \dot{q}\, \frac{\partial^2 L}{\partial q \partial \dot{q}} - \ddot{q}\, \frac{\partial^2 L}{\partial \dot{q}^2} = 0. \tag{2.17}$$

Puis on remarque que, si $z : t \to z(t)$ est la fonction définie par :

$$z(t) = L(q(t), \dot{q}(t)) - \dot{q}(t)\, \frac{\partial L}{\partial \dot{q}}(q(t), \dot{q}(t)), \tag{2.18}$$

on a par dérivation de fonctions composées (notations abrégées) :

$$\begin{aligned} \dot{z} &= \frac{\partial L}{\partial q}\, \dot{q} + \frac{\partial L}{\partial \dot{q}}\, \ddot{q} - \ddot{q}\, \frac{\partial L}{\partial \dot{q}} - \dot{q}\, \frac{\partial^2 L}{\partial q \partial \dot{q}}\, \dot{q} - \dot{q}\, \frac{\partial^2 L}{\partial \dot{q}^2}\, \ddot{q} \\ &= (\frac{\partial L}{\partial q} - \frac{\partial^2 L}{\partial q \partial \dot{q}}\, \dot{q} - \frac{\partial^2 L}{\partial \dot{q}^2}\, \ddot{q})\dot{q}. \end{aligned} \tag{2.19}$$

Donc (2.17) s'écrit $\dot{z} = 0$. Donc, cf. (2.18), il existe une constante k telle $z(t) = k$ pour tout t, i.e. telle que $L(q(t), \dot{q}(t)) - \dot{q}(t)\, \frac{\partial L}{\partial \dot{q}}(q(t), \dot{q}(t)) = k$, i.e. :

$$L - \dot{q}\, \frac{\partial L}{\partial \dot{q}} = k, \tag{2.20}$$

dite équation de Beltrami.

Exemple 2.8 En 2-D, et longueur de la courbe explicite donnée par (1.4) : on a $L(X,Y,Z) = \sqrt{1+Z^2} = (1+Z^2)^{\frac12}$, avec donc $\frac{\partial L}{\partial Z}(X,Y,Z) = Z(1+Z^2)^{-\frac12}$, et (2.20) donne : il existe $k \in \mathbb{R}$ t.q. :

$$(1+f'(x)^2)^{\frac12} - f'(x)f'(x)(1+f'(x)^2)^{-\frac12} = k,$$

donc

$$1+f'(x)^2 - f'(x)^2 = k(1+f'(x)^2)^{\frac12},$$

donc $f'(x) = (\frac1k - 1)^{\frac12} =^{\text{noté}} a$ constante. Donc nécessairement $f(x) = c_1x + c_0$. Et une telle fonction convient. ▪

2.5 Equations de Lagrange inchangées par ajout d'une dérivée totale

Proposition 2.9 *On suppose L C^2. Soit de plus une fonction C^2 :*

$$f : \begin{cases} \mathbb{R}^2 & \to \mathbb{R} \\ (X,Y) & \mapsto f(X,Y). \end{cases} \tag{2.21}$$

Alors l'équation (2.10) est inchangée si on ajoute $\frac{Df(t,q(t))}{Dt}$ à L.

(Par un choix judicieux de f, la fonctionnelle $L + \frac{Df}{Dt}$ peut être plus simple que L, et on obtiendra la même courbe q réalisant un minimum de J.)

Preuve. Soit $\widetilde L$ définie par :

$$\widetilde L(X,Y,Z) = L(X,Y,Z) + \frac{\partial f}{\partial X}(X,Y) + \frac{\partial f}{\partial Y}(X,Y)\,Z. \tag{2.22}$$

Donc :

$$\frac{\partial \widetilde L}{\partial Y} = \frac{\partial L}{\partial Y} + \frac{\partial^2 f}{\partial Y \partial X} + \frac{\partial^2 f}{\partial Y^2}Z. \quad \text{et} \quad \frac{\partial \widetilde L}{\partial Z} = \frac{\partial L}{\partial Z} + \frac{\partial f}{\partial Y},$$

et :

$$\frac{\partial^2 \widetilde L}{\partial X \partial Z} = \frac{\partial^2 L}{\partial X \partial Z} + \frac{\partial^2 f}{\partial X \partial Y}, \qquad \frac{\partial^2 \widetilde L}{\partial Y \partial Z} = \frac{\partial^2 L}{\partial Y \partial Z} + \frac{\partial^2 f}{\partial Y^2}, \qquad \frac{\partial^2 \widetilde L}{\partial Z^2} = \frac{\partial^2 L}{\partial Z^2}.$$

Notons :

$$\widetilde J(q) = \int_{t=a}^b \widetilde L(t, q(t), \dot q(t))\,dt, \tag{2.23}$$

l'action de $\widetilde L$ le long de q. L'extremum de $\widetilde L$ est donné par les q t.q. pour tout t, cf. (2.10) :

$$0 = \frac{\partial \widetilde L}{\partial Y} - \frac{\partial^2 \widetilde L}{\partial X \partial Z} - \dot q \frac{\partial^2 \widetilde L}{\partial Y \partial Z} - \ddot q \frac{\partial^2 \widetilde L}{\partial^2 Z}, \tag{2.24}$$

soit :

$$\begin{aligned} 0 &= (\frac{\partial L}{\partial Y} + \frac{\partial^2 f}{\partial X \partial Y} + \frac{\partial^2 f}{\partial Y^2}\dot q) - \frac{\partial^2 L}{\partial X \partial Z} - \frac{\partial^2 f}{\partial X \partial Y} - \dot q \frac{\partial^2 L}{\partial Y \partial Z} - \dot q \frac{\partial^2 f}{\partial Y^2} - \ddot q \frac{\partial^2 L}{\partial Z^2} \\ &= \frac{\partial L}{\partial Y} - \frac{\partial^2 L}{\partial X \partial Z} - \dot q \frac{\partial^2 L}{\partial Y \partial Z} - \ddot q \frac{\partial^2 L}{\partial Z^2}, \end{aligned} \tag{2.25}$$

les termes en f s'annulant. Donc un extremum de $\widetilde L$ est un extremum de L et réciproquement. ▪

2.6 Exemple de la brachistochrone

Exemple 2.10 Brachistochrone. Ici $L(X,Y,Z) = \frac{1}{\sqrt{2g}}\frac{(1+Z^2)^{\frac12}}{Y^{\frac12}} = L(Y,Z)$ ne dépend pas de X. Donc une courbe $q =^{\text{noté}} f$ qui réalise le minimum cherché vérifie l'équation (2.20) de Beltrami.

On a :

$$\frac{\partial L}{\partial Z}(Y,Z) = \frac{1}{\sqrt{2g}} \frac{Z}{(1+Z^2)^{\frac12} Y^{\frac12}},$$

et donc il existe une constante $k \in \mathbb{R}$ telle que, cf (2.20), avec $Y = f$ et $Z = f'$:

$$\frac{(1+Z^2)^{\frac12}}{Y^{\frac12}} - Z \frac{Z}{(1+Z^2)^{\frac12} Y^{\frac12}} = \sqrt{2g}\, k.$$

On a donc :

$$1 = \sqrt{2g}\, k\, Y^{\frac12} (1+Z^2)^{\frac12},$$

soit, en posant $c = \dfrac{1}{2gk^2}$:

$$f(1+f'^2) = c, \tag{2.26}$$

dont la solution est une cycloïde pour une roue de rayon $R = \frac{c}{2}$:

Rappel. L'équation paramétrique usuelle de la cycloïde (au signe de y près, ici roue de rayon R qui roule "sous le plafond" et "vers la droite" sans glisser) est, pour $\theta \in [0, 2\pi]$:

$$\vec{r}(\theta) = \begin{pmatrix} x(\theta) \\ y(\theta) \end{pmatrix} = \begin{pmatrix} R(\theta - \sin\theta) \\ R(1 - \cos\theta) \end{pmatrix}.$$

On a $x'(\theta) = R(1 - \cos\theta)$ qui est strictement positif pour $\theta \in]0, 2\pi[$. Donc $x : \theta \to x(\theta)$ est inversible sur $]0, 2\pi[$, d'inverse noté $\theta : x \to \theta(x)$. D'où $f(x) = y(\theta(x)) = (y \circ \theta)(x)$ est dérivable de dérivée $f'(x) = y'(\theta(x))\theta'(x)$, avec $y'(\theta) = R\sin\theta$ et avec $\theta'(x) = \frac{1}{x'(\theta)} = \frac{1}{R(1-\cos\theta)}$ quand $\theta = \theta(x)$. Donc $f'(x) = \frac{\sin\theta}{1-\cos\theta}$ quand $\theta = \theta(x)$. Avec $\cos\theta = \cos(2\frac{\theta}{2}) = 1 - 2\sin^2\frac{\theta}{2}$ et avec $\sin\theta = 2\cos\frac{\theta}{2}\sin\frac{\theta}{2}$. Donc $f'(x) = \frac{\cos\frac{\theta}{2}}{\sin\frac{\theta}{2}}$. D'où $1 + f'(x)^2 = \frac{1}{\sin^2\frac{\theta}{2}}$. D'où $f(x)(1 + f'(x)^2) = \frac{R(1-\cos\theta)}{\sin^2\frac{\theta}{2}} = 2R$. Donc la solution est bien la cycloïde décrite, cf. (2.26). ▪

3 Théorème d'Euler–Lagrange dans $\mathbb{R}^n$: $\frac{\partial L}{\partial \mathbf{q}} - \frac{D}{Dt}\frac{\partial L}{\partial \dot{\mathbf{q}}} = 0$

3.1 Introduction

On s'intéresse aux fonctions à valeurs vectorielles :

$$E = \{\mathbf{q} : t \in [a,b] \to \mathbf{q}(t) = \begin{pmatrix} q_1(t) \\ \vdots \\ q_n(t) \end{pmatrix} \in \mathbb{R}^n,\ \mathbf{q}(a) \text{ et } \mathbf{q}(b) \text{ fixés},\ \mathbf{q} \in C^2\}. \tag{3.1}$$

La dérivée est notée $\frac{d\mathbf{q}}{dt}(t) = \dot{\mathbf{q}}(t)$, et la dérivée seconde $\frac{d^2\mathbf{q}}{dt^2}(t) = \ddot{\mathbf{q}}(t)$.

Soit :

$$L : \begin{cases} \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n & \to \mathbb{R} \\ (t, \vec{y}, \vec{z}) & \mapsto L(t, \vec{y}, \vec{z}) \end{cases} \tag{3.2}$$

une fonction C^2, et soit :

$$J(\mathbf{q}) = \int_{t=a}^{b} L(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))\, dt \tag{3.3}$$

une fonctionnelle à minimiser sur l'ensemble des fonctions $\mathbf{q} \in E$.

Exemple 3.1 Dans $\mathbb{R}^2$. En cartésien on prend $\mathbf{q} = (x, y) = \vec{x} \in \mathbb{R}^2$. Et $\mathbf{q}(t) = (x(t), y(t))$ donne la position à l'instant t. Et $\dot{\mathbf{q}}(t)$ et $\ddot{\mathbf{q}}(t)$ donne les vitesses et accélération au point $\mathbf{q}(t)$ à l'instant t (le long de la trajectoire $t \to \mathbf{q}(t)$). ▪

Exemple 3.2 On prend :

$$\mathbf{q} = \begin{pmatrix} r \\ \theta \end{pmatrix} \overset{\text{noté}}{=} (r,\theta) \in \mathbb{R}_+^* \times \mathbb{R}$$

(paramétrage en polaire), la position géométrique étant donnée par

$$\vec{x} = \vec{\varphi}(\mathbf{q}) = \vec{\varphi}(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}.$$

Et la connaissance de :

$$\mathbf{q}(t) = \begin{pmatrix} r(t) \\ \theta(t) \end{pmatrix}$$

permet de déduire la position $\vec{x}(t) = \vec{\varphi}(\mathbf{q}(t))$ à l'instant t, à savoir :

$$\vec{x}(t) = \begin{pmatrix} r(t)\cos\theta(t) \\ r(t)\sin\theta(t) \end{pmatrix} = \vec{\varphi}(r(t),\theta(t)),$$

soit encore $\vec{x}(t) = (\vec{\varphi}\circ\mathbf{q})(t)$. Donc la vitesse vérifie $\vec{x}\,'(t) = d\vec{\varphi}(\mathbf{q}(t)).\dot{\mathbf{q}}(t)$. Et, par exemple, la longueur de la courbe $\vec{r}$, cf. (1.1) est donnée par (1.2) à savoir :

$$\int_{t=a}^{b} ||\vec{x}\,'(t)||\,dt = \int_{t=a}^{b} ||d\vec{\varphi}(\mathbf{q}(t)).\dot{\mathbf{q}}(t)||\,dt = \int_{t=a}^{b} L(\mathbf{q}(t),\dot{\mathbf{q}}(t))\,dt \overset{\text{noté}}{=} J(\mathbf{q}),$$

où on a posé $L(\mathbf{q},\dot{\mathbf{q}}) =^{\text{déf}} ||d\vec{\varphi}(\mathbf{q}).\dot{\mathbf{q}}||$, ou encore avec les notations (3.2) $L(t,\vec{y},\vec{z}) = ||d\vec{\varphi}(\vec{y}).\vec{z}||$ = norme de la différentielle de $\vec{\varphi}$ en $\vec{y}$ dans la direction $\vec{z}$. Ici L est indépendante de t et on note $L(t,\vec{y},\vec{z}) = L(\vec{y},\vec{z})$. ▪

3.2 Théorème d'Euler–Lagrange dans $\mathbb{R}^n$

Soit E l'espace défini en (3.1).

Théorème 3.3 *Une fonction $\mathbf{q} \in E$ qui réalise le minimum de J vérifie :*

$$\frac{\partial L}{\partial y_i}(t,\mathbf{q}(t),\dot{\mathbf{q}}(t)) - \frac{D}{Dt}\Big(\frac{\partial L}{\partial z_i}(t,\mathbf{q}(t),\dot{\mathbf{q}}(t))\Big) = 0, \qquad \forall i = 1,...,n, \tag{3.4}$$

système de n équations aux n inconnues les n fonctions q_i composantes de $\mathbf{q}$. Ce qui est noté :

$$\frac{\partial L}{\partial y_i} - \frac{D}{Dt}\frac{\partial L}{\partial z_i} = 0, \qquad \forall i = 1,...,n. \tag{3.5}$$

Soit avec les notations de la mécanique : pour tout $i = 1,...,n$ (les n équations) :

$$\frac{\partial L}{\partial q_i} - \frac{D}{Dt}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big) = 0, \qquad \forall i = 1,...,n, \tag{3.6}$$

dites équations d'Euler.

Preuve. Reprendre la démonstration du théorème 2.5. ▪

Remarque 3.4 Le système des n équations (3.5) peut s'écrire :

$$\frac{\partial L}{\partial \vec{y}} - \frac{D}{Dt}\Big(\frac{\partial L}{\partial \vec{z}}\Big) = 0, \qquad \text{i.e.} \qquad \frac{\partial L}{\partial \mathbf{q}} - \frac{D}{Dt}\Big(\frac{\partial L}{\partial \dot{\mathbf{q}}}\Big) = 0. \tag{3.7}$$

▪

Exemple 3.5 On reprend l'exemple 1.1, équation (1.2). Ici $\vec{y} = (x,y)$ et $\vec{z} = (x',y')$.

Ici $L(t,\vec{y},\vec{z}) = ||\vec{z}|| = \sqrt{z_1^2+z_2^2}$, et donc $\frac{\partial L}{\partial t} = \frac{\partial L}{\partial y_i} = 0$, et $\frac{\partial L}{\partial z_i}(t,\vec{y},\vec{z}) = \frac{z_i}{||\vec{z}||}$, pour $i=1,2$. Donc :

$$\frac{\partial L}{\partial z_i}(t,\vec{r}(t),\vec{r}\,'(t)) = \frac{x_i'(t)}{(x_1'(t)^2+x_2'(t)^2)^{\frac12}},$$

pour $i=1,2$, et, avec :

$$\frac{d}{dt}((x_1'(t)^2+x_2'(t)^2)^{\frac12}) = \frac12((x_1'(t)^2+x_2'(t)^2)^{-\frac12}(2x_1'(t)x_1''(t)+2x_2'(t)x_2''(t),$$

on obtient :

$$\begin{aligned}\frac{d}{dt}\Big(\frac{\partial L}{\partial z_i}(t,\vec{r}(t),\vec{r}\,'(t))\Big) &= \frac{x_i''(t)||\vec{r}\,'(t)|| - x_i'(t)\frac{x_1'(t)x_1''(t)+x_2'(t)x_2''(t)}{||\vec{r}\,'(t)||}}{||\vec{r}\,'(t)||^2}\\ &= \frac{x_i''(t)(x_1'(t)^2+x_2'(t)^2) - x_i'(t)(x_1'(t)x_1''(t)+x_2'(t)x_2''(t))}{||\vec{r}\,'(t)||^3}.\end{aligned}$$

Et le système (3.7) donne, pour tout t :

$$\begin{cases} x_1''(t)x_2'(t) - x_1'(t)x_2''(t) = 0,\\ x_2''(t)x_1'(t) - x_2'(t)x_1''(t) = 0,\end{cases}$$

deux fois la même équation, soit, pour tout t :

$$\frac{x_1''}{x_1'}(t) = \frac{x_2''}{x_2'}(t),$$

d'où $\log(\alpha x_1') = \log(\beta x_2')$ pour α,β réels qcq, d'où $\alpha x_1' = \beta x_2'$, d'où $\alpha x_1 = \beta x_2 + b$, équation d'une droite. ▟

Exercice 3.6 Reprendre l'exemple 1.1, équation (1.2), mais cette fois ci en coordonnées polaires.

Réponse. $\vec{x}(r,\theta) = \begin{pmatrix} x(r,\theta)\\ y(r,\theta)\end{pmatrix} = \begin{pmatrix} r\cos\theta\\ r\sin\theta\end{pmatrix}$, et avec pour une courbe, $r=r(t)$ et $\theta=\theta(t)$. La courbe cherchée est donc de type $\vec{g}(t) = \vec{x}(r(t),\theta(t))$, avec (notations abusives pour alléger l'écriture) $\vec{g}\,'(t) = \frac{\partial\vec{x}}{\partial r}(r,\theta)\frac{dr}{dt}(t) + \frac{\partial\vec{x}}{\partial\theta}(r,\theta)\frac{d\theta}{dt}(t) = \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}r' + \begin{pmatrix}-r\sin\theta\\ r\cos\theta\end{pmatrix}\theta' = \begin{pmatrix} r'\cos\theta - r\theta'\sin\theta\\ r'\sin\theta + r\theta'\sin\theta\end{pmatrix}$, et donc $||\vec{g}\,'(t)||^2 = r'(t)^2 + r^2(t)\theta'(t)^2$.

Donc :

$$J_p(\vec{x}) = \int_{t=\alpha}^{\beta} \sqrt{r'(t)^2 + r^2(t)\theta'(t)^2}\,dt. \tag{3.8}$$

Donc ici on pose $\mathbf{q} = (r,\theta) = (y_1,y_2) = \vec{y}$, donc $\dot{\mathbf{q}} = (r',\theta') = (z_1,z_2) = \vec{z}$, et :

$$L(t,\vec{y},\vec{z}) = (z_1^2 + y_1^2 z_2^2)^{\frac12} \qquad (= (r'^2 + r^2\theta'^2)^{\frac12}).$$

On a :

$$\frac{\partial L}{\partial y_1} = \frac{y_1 z_2^2}{L} = \frac{r\theta'^2}{L}, \qquad \frac{\partial L}{\partial y_2} = 0, \qquad \frac{\partial L}{\partial z_1} = \frac{z_1}{L} = \frac{r'}{L}, \qquad \frac{\partial L}{\partial z_2} = \frac{y_1^2 z_2}{L} = \frac{r^2\theta'}{L}. \tag{3.9}$$

D'où :

$$\frac{dL}{dt} = \frac{\partial L}{\partial\vec{y}}\dot{\mathbf{q}} + \frac{\partial L}{\partial\vec{z}}\ddot{\mathbf{q}} = \frac{r\theta'^2 r'}{L} + \frac{r'r''}{L} + \frac{r^2\theta'\theta''}{L} = \frac{rr'\theta'^2 + r'r'' + r^2\theta'\theta''}{L},$$

d'où :

$$\begin{aligned}\frac{d}{dt}(\frac{\partial L}{\partial z_1}) &= \frac{r''L - r'\frac{rr'\theta'^2+r'r''+r^2\theta'\theta''}{L}}{L^2} = \frac{r''(r'^2+r^2\theta'^2) - rr'^2\theta'^2 - r'^2r'' - r^2r'\theta'\theta''}{L^3}\\ &= \frac{r(rr''\theta'^2 - r'^2\theta'^2 - rr'\theta'\theta'')}{L^3}.\end{aligned}$$

et :

$$\begin{aligned}\frac{d}{dt}(\frac{\partial L}{\partial z_2}) &= \frac{(2rr'\theta' + r^2\theta'')L - r^2\theta'\frac{rr'\theta'^2+r'r''+r^2\theta'\theta''}{L}}{L^2}\\ &= \frac{(2rr'\theta'+r^2\theta'')(r'^2+r^2\theta'^2) - r^2\theta'(rr'\theta'^2+r'r''+r^2\theta'\theta'')}{L^3}\\ &= \frac{2rr'^3\theta' + 2r^3r'\theta'^3 + r^2r'^2\theta'' + r^4\theta''\theta'^2 - r^3r'\theta'^3 - r^2r'r''\theta' - r^4\theta'^2\theta''}{L^3}\\ &= \frac{rr'(2r'^2\theta' + r^2\theta'^3 + rr'\theta'' - rr''\theta'}{L^3}\end{aligned}$$

Les équations d'Euler donnent deux fois la même équation :

$$2r'^2\theta' - rr''\theta' + r^2\theta'^3 + rr'\theta'' = 0, \tag{3.10}$$

pas immédiate à résoudre... à moins de revenir en cartésien : on sait que la solution est une droite, et donc (3.10) donne une équation différentielle satisfaite par les droites en coordonnées polaires. ▟

4 Rappels de mécanique classique

En route vers l'hamiltonien.

4.1 Principe fondamental

On considère des particules. Soit $\vec{x}(t)$ la position d'une particule à l'instant t. La fonction $t \to \vec{x}(t)$ désignera la trajectoire d'une particule. Soit $\vec{v}(t,\vec{x})$ la vitesse (eulérienne) de la particule qui se trouve en $\vec{x} = \vec{x}(t)$ à l'instant t, i.e. :

$$\vec{v}(t,\vec{x}(t)) \stackrel{\text{déf}}{=} \frac{d\vec{x}}{dt}(t). \tag{4.1}$$

Si $\vec{v}$ est indépendante de t, i.e. $\frac{\partial \vec{v}}{\partial t}(t,\vec{x}) = 0$ pour tout t, on dit que le mouvement est stationnaire. Dans ce cas, toutes les particules qui passent par un point $\vec{y}$ sont animées de la même vitesse en $\vec{y}$, car si $\vec{y} = \vec{x}(t)$ et $\vec{y} = \widetilde{\vec{x}}(\widetilde{t})$, alors $\frac{d\vec{x}}{dt}(t) = \vec{v}(\vec{x}(t)) = \vec{v}(\vec{y}) = \vec{v}(\widetilde{\vec{x}}(\widetilde{t})) = \frac{d\widetilde{\vec{x}}}{dt}(\widetilde{t})$ (cas stationnaire).

Et l'accélération (eulérienne) de la particule qui se trouve en $\vec{x} = \vec{x}(t)$ à l'instant t est par définition :

$$\vec{\gamma}(t,\vec{x}(t)) \stackrel{\text{déf}}{=} \frac{d^2\vec{x}}{dt^2}(t) \qquad (= \frac{D\vec{v}(t,\vec{x}(t))}{Dt}). \tag{4.2}$$

Principe fondamental de la mécanique (équations de Newton). Soit $\vec{f}$ un "champ de forces" (un champ de vecteurs). Pour une particule de masse m constante soumise à l'instant t au point $\vec{x}$ à la force $\vec{f}(t,\vec{x})$ on a :

$$\vec{f} = m\,\vec{\gamma}, \tag{4.3}$$

au sens, quand à t la particule se trouve en $\vec{x}(t)$, son accélération vérifie :

$$\vec{f}(t,\vec{x}(t)) = m\,\vec{\gamma}(t,\vec{x}(t)) \qquad (= m\frac{d^2\vec{x}}{dt^2}(t)).$$

(Exemple avec $\vec{f}$ le champ de "pesanteur".)

C'est une équation différentielle du second ordre dans $\mathbb{R}^n$ en l'inconnue la fonction $t \to \vec{x}(t)$, "trajectoire de la particule soumise à une force $\vec{f}$". Autrement dit, c'est le système différentiel :

$$m\,\frac{d^2x_i}{dt^2}(t) = f_i(t, x_1(t), ..., x_n(t)), \qquad i = 1, .., n. \tag{4.4}$$

Sa résolution nécessite deux constantes d'intégration dans $\mathbb{R}^n$ données en générale par les conditions initiales $\vec{x}(t_0)$ et $\frac{d\vec{x}}{dt}(t_0) = \vec{v}(t_0,\vec{x}(t_0))$ (position et vitesse initiales : $2n$ constantes scalaires).

4.2 Énergie cinétique

On se place dans un espace vectoriel muni d'un produit scalaire. On prend pour simplifier $\mathbb{R}^n$ muni de son produit scalaire euclidien $(\cdot,\cdot)_{\mathbb{R}^n}$ et de sa norme euclidienne associée $||.||_{\mathbb{R}^n}$.

Définition 4.1 Pour une particule de masse m en $\vec{x}$ à l'instant t, de vitesse (eulérienne) $\vec{v}(t,\vec{x})$, et de vitesse scalaire $v = ||\vec{v}||_{\mathbb{R}^n}$, l'énergie cinétique E_c est définie par :

$$E_c(\vec{v}) = \frac{1}{2}mv^2. \tag{4.5}$$

Notation complète : $E_c(t,\vec{x},\vec{v}) =^{\text{déf}} \frac{1}{2}mv^2 =^{\text{noté}} E_c(\vec{v})$ qui a un sens pour $\vec{v}$ vecteur à t en $\vec{x}$ (i.e. pour $\vec{v}$ un champ de vecteurs instationnaire).

Proposition 4.2 *et définition. On a :*

$$m\vec{v} = \vec{\nabla}E_c(\vec{v}), \tag{4.6}$$

*et $m\vec{v} =^{noté}$ **p** est appelé quantité de mouvement.*

Preuve. On a, pour $\vec{w} \in \mathbb{R}^n$ et $h > 0$:

$$||\vec{v}+h\vec{w}||^2_{\mathbb{R}^n} = ||\vec{v}||^2_{\mathbb{R}^n} + 2h(\vec{v},\vec{w})_{\mathbb{R}^n} + h^2||\vec{w}||^2 = v^2 + 2h(\vec{v},\vec{w})_{\mathbb{R}^n} + o(h),$$

d'où :

$$E_c(\vec{v}+h\vec{w}) = E_c(\vec{v}) + h\,(m\vec{v},\vec{w})_{\mathbb{R}^n} + o(h).$$

Et le développement limité de E au premier ordre donne, au voisinage de $h = 0$:

$$E_c(\vec{v}+h\vec{w}) = E_c(\vec{v}) + h\,(\vec{\nabla}E_c(\vec{v}),\vec{w})_{\mathbb{R}^n} + o(h),$$

vrai pour tout $\vec{w}$, d'où (4.6). ∎

Principe fondamental de la mécanique généralisé. Soit $\vec{f}$ un champ de vecteurs ("un champ de forces"). Une particule de masse m (non nécessairement constante) soumise à l'instant t au point $\vec{x}$ à la force $\vec{f}(t,\vec{x})$ on a :

$$\vec{f} = \frac{d\mathbf{p}}{dt}. \tag{4.7}$$

4.3 Énergie potentielle

Définition 4.3 Soit un ouvert Ω dans $\mathbb{R}^n$ (un ensemble de particules). S'il existe une fonction $E_p \in C^1(\Omega;\mathbb{R})$ telle que la force sur toute particule à la position $\vec{x}$ est indépendante du temps et est donnée par :

$$\vec{f}(\vec{x}) = -\vec{\nabla}E_p(\vec{x}) = -\begin{pmatrix} \frac{\partial E_p}{\partial x_1}(\vec{x}) \\ \vdots \\ \frac{\partial E_p}{\partial x_n}(\vec{x}) \end{pmatrix}, \tag{4.8}$$

alors la fonction E_p est appelée énergie potentielle. (Il est immédiat que si E_p existe alors E_p est définie à une constante près.)

Donc $\vec{f}$ dérive du potentiel E_p ssi $\vec{f}$ est le gradient du potentiel E_p.

Remarque 4.4 Généralement $\vec{f}$ est une fonction $(t,\vec{x}) \to f(t,\vec{x})$. Donc si elle dérive d'un potentiel on a $\frac{\partial \vec{f}}{\partial t} = 0$, et on note $\vec{f} : \vec{x} \to \vec{f}(\vec{x})$. ▪

Exemple 4.5 Oscillateur harmonique 1-D : $E_p(x) = \frac{1}{2}kx^2$ à une constante près (potentiel dû au ressort). ▪

Exemple 4.6 Pesanteur "constante" : on a $E_p(z) = -gz$ à une constante près, z étant l'altitude. ▪

4.4 Système conservatif

Définition 4.7 Un ensemble de particules constituant un ouvert dans $\mathbb{R}^n$ est dit conservatif ssi il existe une fonction potentielle E_p.

Proposition 4.8 *Pour un système conservatif, le principe fondamental (4.3) s'énonce également :*

$$\frac{D}{Dt}\frac{\partial E_c}{\partial \vec{v}} + \frac{\partial E_p}{\partial \vec{x}} = 0, \tag{4.9}$$

autre écriture de $\frac{D}{Dt}(m\vec{v}) = -\vec{\nabla}E_p = \vec{f}$.

Preuve. On a $\frac{\partial E_c}{\partial \vec{v}} = m\vec{v}$ et $\frac{\partial E_p}{\partial \vec{x}} = -f$: on a bien (4.7). ▪

Exemple 4.9 Oscillateur harmonique 1-D : on a $E_c'(v) = mv$ et $E_p'(x) = kx$, et on retrouve l'équation du ressort $mx'' + kx = 0$. ▪

4.5 Énergie totale

Définition 4.10 Dans le cas d'un système conservatif, pour une particule de masse m en $\vec{x}$ à l'instant t animée de la vitesse (eulérienne) $\vec{v}(t,\vec{x})$, l'énergie totale est la fonction E définie par :

$$E(\vec{x},\vec{v}) = E_p(\vec{x}) + E_c(\vec{v}), \tag{4.10}$$

et on note :

$$E = E_p + E_c. \tag{4.11}$$

(L'énergie totale est considérée en les $(\vec{x},\vec{v}) = (\vec{x},\vec{v}(t,\vec{x}))$, sinon on ne peut pas considérer E_c.)

Notation complète de (4.10) :

$$E(\vec{x}(t),\vec{v}(t,\vec{x}(t))) = E_p(\vec{x}(t)) + E_c(\vec{v}(t,\vec{x}(t))). \tag{4.12}$$

Exemple 4.11 Oscillateur harmonique 1-D : on a $E = \frac{1}{2}mv^2 + \frac{1}{2}kx^2$. ▪

Proposition 4.12 *Pour un système conservatif, l'énergie totale se conserve au cours du mouvement, i.e. le long de sa trajectoire, l'énergie totale E d'une particule est constante :*

$$\frac{dE}{dt} = 0. \tag{4.13}$$

Preuve. Soit $g(t) = E(\vec{x}(t), \vec{v}(t, \vec{x}(t))) = E_c(\vec{v}(t, \vec{x}(t))) + E_p(\vec{x}(t))$. On a, avec les notations allégées :

$$\frac{dE}{dt} \stackrel{\text{déf}}{=} \frac{dg}{dt} = \frac{\partial E_c}{\partial \vec{v}} \frac{d\vec{v}}{dt} + \frac{\partial E_p}{\partial \vec{x}} . \frac{d\vec{x}}{dt} = m\vec{v}.\vec{\gamma} - \vec{f}.\vec{v} = (m\vec{\gamma} - \vec{f}).\vec{v} = 0.$$

▪

Remarque 4.13 (4.11) suggère que “les bonnes définitions” de E, E_c et E_p sont données par : ce sont les fonctions définies sur le fibré tangent, i.e. l'ensemble des champs des vecteurs $\widetilde{\vec{v}}$ définis par $\widetilde{\vec{v}}(\vec{x}) = (\vec{x}, \vec{v}(\vec{x}))$ où $\vec{v} : \mathbb{R}^n \to \mathbb{R}^n$ est une fonction C^∞, définies par :
$E_c(\widetilde{\vec{v}}(\vec{x})) = E_c(\vec{v}(\vec{x}))$, où donc $\frac{\partial E_c}{\partial \vec{x}} = 0$,
$E_p(\widetilde{\vec{v}}(\vec{x})) = E_c(\vec{x})$, où donc $\frac{\partial E_p}{\partial \vec{v}} = 0$, et
$E(\widetilde{\vec{v}}(\vec{x})) = E_c(\widetilde{\vec{v}}(\vec{x})) + E_p(\widetilde{\vec{v}}(\vec{x}))$.
C'est la définition donnée dans un cours de géométrie différentielle. ▪

4.6 Travail et force dérivant d'un potentiel

Proposition 4.14 *Soit $\Omega \subset \mathbb{R}^n$ un ouvert connexe.*

La force (supposée continue) dérive d'un potentiel $E_p : \Omega \to \mathbb{R}$ (supposé C^1) ssi, pour tous points $\vec{x}_1, \vec{x}_2 \in \Omega$ et toute courbe Γ dans Ω joignant $\vec{x}_1$ à $\vec{x}_2$, le travail :

$$T(\vec{f}, \Gamma) \stackrel{\text{déf}}{=} \int_\Gamma \vec{f}.d\vec{r} \tag{4.14}$$

ne dépend que des extrémités $\vec{x}_1$ et $\vec{x}_2$ de Γ. Et dans ce cas :

$$T(\vec{f}, \Gamma) = E_p(\vec{x}_1) - E_p(\vec{x}_2). \tag{4.15}$$

Preuve. Soit $\vec{r} : t \in [a,b] \to \vec{r}(t)$ une courbe régulière avec $\vec{r}(a) = \vec{x}_1$ et $\vec{r}(b) = \vec{x}_2$ t.q. $\Gamma = \mathrm{Im}\vec{r}$. Donc :

$$T(\vec{f}, \Gamma) = \int_{t=a}^{b} \vec{f}(\vec{r}(t)).\vec{r}\,'(t)\,dt \tag{4.16}$$

1- Si $\vec{f}$ dérive d'un potentiel E_p alors $t \to \vec{f}(\vec{r}(t)).\vec{r}\,'(t) = -\vec{\nabla}E_p(\vec{r}(t)).\vec{r}\,'(t)$ est la dérivée de $t \to -(E_p \circ \vec{r})(t)$, et donc $T(\vec{f}, \Gamma) = -\int_{t=a}^{b} (E_p \circ \vec{r})'(t)\,dt = -[(E_p \circ \vec{r})(t)]_a^b$, d'où (4.15).

2- Réciproque. Soit $\vec{x}_1, \vec{x} \in \Omega$ fixés, soit $\Gamma_{\vec{x}}$ une courbe dans Ω joignant $\vec{x}_1$ et $\vec{x}$, et soit $\vec{r} : t \in [a,b] \to \vec{r}(t)$ un paramétrage de $\Gamma_{\vec{x}}$.

On définit la fonction $E_p : \Omega \to \mathbb{R}$ par $E_p(\vec{x}) = -T(\vec{f}, \Gamma_{\vec{x}})$, ce qui définit entièrement E_p car T ne dépend pas du chemin.

Comme Ω est ouvert et $\vec{x} \in \Omega$, il existe $h_{\vec{x}} > 0$ t.q. la boule $B(\vec{x}, h_{\vec{x}})$ de centre $\vec{x}$ et de rayon $h_{\vec{x}}$ est entièrement dans Ω. Soit $\vec{y} \in B(\vec{x}, h_{\vec{x}})$, notons $\vec{u} = \vec{y} - \vec{x}$, et notons :

$$[\vec{x}, \vec{y}] = \{\vec{r}(t) = \vec{x} + t\vec{u} \text{ t.q. } t \in [0,1]\} \tag{4.17}$$

le segment de droite joignant $\vec{x}$ à $\vec{y}$, paramétré par $\vec{r} : t \in [0,1] \to \vec{r}(t) = \vec{x} + t\vec{u}$, avec donc $d\vec{r}(t) = \vec{u}\,dt$.

Le travail étant indépendant du chemin, on a $\int_{\Gamma_{\vec{x}} \cup [\vec{x}, \vec{x}+\vec{u}]} = \int_{\Gamma_{\vec{x}}} + \int_{[\vec{x}, \vec{x}+\vec{u}]}$ (relation de Chasles), soit :

$$\int_{[\vec{x}, \vec{x}+\vec{u}]} \vec{f}.d\vec{r} = \int_{\Gamma_{\vec{x}} \cup [\vec{x}, \vec{x}+\vec{u}]} \vec{f}.d\vec{r} - \int_{\Gamma_{\vec{x}}} \vec{f}.d\vec{r} = -E_p(\vec{x} + \vec{u}) + E_p(\vec{x}).$$

Avec :

$$\int_{[\vec{x}, \vec{x}+\vec{u}]} \vec{f}.d\vec{r} = \int_{t=0}^{1} \vec{f}(\vec{x} + t\vec{u}).\vec{u}\,dt = (\vec{f}(\vec{x}) + o(1)).\vec{u} = \vec{f}(\vec{x}).\vec{u} + o(\vec{u}),$$

d'où :

$$\vec{f}(\vec{x}).\vec{u} + o(\vec{u}) = -E_p(\vec{x} + \vec{u}) + E_p(\vec{x}),$$

i.e. $E_p(\vec{y}) = E_p(\vec{x}) - \vec{f}(\vec{x}).(\vec{y} - \vec{x}) + o(\vec{y} - \vec{x})$. D'où E_p est différentiable en $\vec{x}$ avec $\vec{\nabla}E_p(\vec{x}) = -\vec{f}(\vec{x})$. ▪

5 Principe de moindre action d'Hamilton

5.1 Principe fondamental vs équations d'Euler–Lagrange

Plaçons-nous en coordonnées cartésiennes dans $\mathbb{R}^n$.
Pour un système conservatif on a (4.9) :

$$\frac{d}{dt}\frac{\partial E_c}{\partial \vec{v}} + \frac{\partial E_p}{\partial \vec{x}} = 0. \tag{5.1}$$

Définissons donc la fonction L par :

$$L = E_c - E_p, \tag{5.2}$$

la différence entre les énergies cinétiques et potentielles, i.e. $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ définie par :

$$L(\vec{y}, \vec{z}) = E_c(\vec{z}) - E_p(\vec{y}). \tag{5.3}$$

Et les équations de Euler–Lagrange (3.7) se lisent :

$$\frac{d}{dt}\frac{\partial L}{\partial \vec{v}} - \frac{\partial L}{\partial \vec{x}} = 0. \tag{5.4}$$

En comparaison avec (5.1), il y a "juste" le signe qui change...

5.2 Principe de moindre action de Hamilton en cartésien

Théorème 5.1 *On suppose le système conservatif, et on pose $L = E_c - E_p$ comme en (5.3).*
Une solution du principe fondamental (4.3) vérifie les équations d'Euler–Lagrange (5.4).
Et donc une trajectoire $\vec{x} : t \in [a, b] \to \vec{x}(t)$ qui est solution du principe fondamental est également un extrémum de la fonctionnelle :

$$J(\vec{x}) = \int_{t=a}^{b} L(\vec{x}(t), \vec{x}\,'(t))\, dt \qquad (\overset{\text{noté}}{=} \int_{t=a}^{b} L\, dt). \tag{5.5}$$

Preuve. Euler–Lagrange (3.7), à savoir $\frac{\partial L}{\partial \vec{y}} - \frac{d}{dt}(\frac{\partial L}{\partial \vec{z}}) = 0$, avec (5.3) donne $-\vec{\nabla} E_p(\vec{y}) - \frac{d}{dt}(\vec{\nabla} E_c) = 0$. Ce qui est une autre expression du principe fondamental $\vec{f} - \frac{d}{dt}(m\vec{v}) = \vec{0}$ dans le cas des systèmes conservatifs, cf. proposition 4.8. ∎

Définition 5.2 Une coordonnée y_i telle que $\frac{\partial L}{\partial y_i} = 0$ (i.e. L ne dépend pas de y_i), quand $\vec{y} = (y_1, ..., y_n)$, est appelée une coordonnée cyclique.

Corollaire 5.3 *Si y_i est une coordonnée cyclique (ici en cartésien si E_p ne dépend pas de y_i), on a pour $i = 1, ..., n$:*

$$\frac{Dmv_i}{Dt} = 0, \qquad \textit{i.e.} \qquad mv_i = \textit{ constante sur la trajectoire.} \tag{5.6}$$

5.3 Cadre de la dynamique des systèmes conservatifs usuels

Exemple des coordonnées polaires : pour $\vec{r}(r, \theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$, le long d'une trajectoire $t \to \vec{r}(r(t), \theta(t))$ la vitesse est donnée par $||\vec{v}|| = (r'^2 + r^2\theta'^2)^{\frac{1}{2}}$ qui est une fonction non seulement de r' et de θ', mais également de r. Ainsi $\vec{v} = \vec{v}((r, \theta), (r', \theta')) = \vec{v}(\mathbf{q}, \dot{\mathbf{q}})$, où on a posé $\mathbf{q} = (r, \theta)$.

Donc dans le système de coordonnées polaires $\mathbf{q} = (r, \theta)$ l'énergie cinétique s'exprime sous la forme :

$$E_c(\vec{v}(\mathbf{q}, \dot{\mathbf{q}})) = \widetilde{E}_c(\mathbf{q}, \dot{\mathbf{q}}), \tag{5.7}$$

où donc $\widetilde{E}_c = E_c \circ \vec{v}$. Dans la suite, on omet le "tilde" de $\widetilde{E}_c(\mathbf{q}, \dot{\mathbf{q}})$ pour retenir $E_c(\mathbf{q}, \dot{\mathbf{q}})$ (allège l'écriture). Généralisation :

Définition 5.4 Cadre de la dynamique des systèmes conservatifs usuels : on suppose qu'on dispose d'un système de coordonnées $\mathbf{q}$. On suppose que :

1- l'énergie cinétique est une forme quadratique de type :

$$E_c(\mathbf{q},\dot{\mathbf{q}}) = \frac{1}{2}(A(\mathbf{q}).\dot{\mathbf{q}},\dot{\mathbf{q}})_{\mathbb{R}^n} = \frac{1}{2}\dot{\mathbf{q}}^T.A(\mathbf{q}).\dot{\mathbf{q}} \qquad (= \frac{1}{2}\sum_{i,j=1}^n a_{ij}(\mathbf{q})\dot{q}_i\dot{q}_j), \tag{5.8}$$

où $A = [a_{ij}] : \mathbb{R}^n \to \mathbb{R}^{n^2}$ est une fonction matricielle sur $\mathbb{R}^n$ telle que $A(\mathbf{q})$ est symétrique pour tout $\mathbf{q}$,

2- l'énergie potentielle ne dépend que de $\mathbf{q}$ (caractérise la position) :

$$E_p : \mathbf{q} \in \mathbb{R}^n \to E_p(\mathbf{q}) \in \mathbb{R}. \tag{5.9}$$

5.4 Principe de moindre action de Hamilton

Soit $\mathbf{q}$ un système de coordonnées. On définit le lagrangien par :

$$L(\mathbf{q},\dot{\mathbf{q}}) = E_c(\mathbf{q},\dot{\mathbf{q}}) - E_p(\mathbf{q}), \tag{5.10}$$

où E_c et E_p sont les expressions de l'énergie cinétique et de l'énergie potentielle dans le système de coordonnées considéré.

Exemple 5.5 En polaires $E_c = E_c(\mathbf{q},\dot{\mathbf{q}}) = \frac{1}{2}m(r'^2 + r^2\theta'^2)$ et $E_p = E_p(\mathbf{q}) = (E_p \circ \varphi^{-1})(\vec{x})$ si $\vec{\varphi} : \mathbf{q} \to \vec{x} = \vec{\varphi}(\mathbf{q})$ est le changement de variables polaire-cartésien. ▟

Corollaire 5.6 *On se place dans le cadre de la dynamique des systèmes conservatifs usuels.*

Quel que soit le système de coordonnées, on conserve les équations d'Euler–Lagrange :

$$\frac{D}{Dt}\frac{\partial L}{\partial \dot{\mathbf{q}}} - \frac{\partial L}{\partial \mathbf{q}} = 0. \tag{5.11}$$

Une courbe $t \to \mathbf{q}(t)$ satisfaisant cette équation est solution du principe fondamental et réalise un extremum de $J(\mathbf{q}) = \int_a^b L(\mathbf{q}(t),\dot{\mathbf{q}}(t))\,dt$.

Preuve. La courbe solution réalise, sur l'ensemble des courbes Γ de même extrémités, le minimum de $J(\Gamma) = \int_{t=a}^b L(\vec{x}(t),\vec{x}\,'(t))\,dt$, et le minimum de l'intégrale est indépendante du choix du système de coordonnées (indépendant du changement de variables). ▟

6 Équations de Lagrange

6.1 Introduction

En cartésien : prenons l'exemple d'un oscillateur harmonique. L'énergie cinétique est $E_c = \frac{1}{2}mv^2$ et l'énergie potentielle est $E_p = \frac{1}{2}kx^2$, et l'énergie totale $E = E_c + E_p$ se conserve.

L'équation du mouvement est l'équation différentielle du second ordre $m\ddot{x} + kx = 0$ qui s'exprime également comme système différentiel du premier ordre :

$$\begin{cases} m\dot{x} = m\dot{x} \\ m\ddot{x} = -kx \end{cases} \tag{6.1}$$

On pose alors $p = m\dot{x}$, et cela s'écrit :

$$\begin{cases} p = m\dot{x} \\ \dot{p} = kx \end{cases}$$

Ce système s'écrit donc aussi :

$$\begin{cases} p = \dfrac{\partial L}{\partial v} \\ \dot{p} = \dfrac{\partial L}{\partial x} \end{cases} \tag{6.2}$$

On va montrer : pour les systèmes conservatifs, quel que soit le système de coordonnées $\mathbf{q}$, en définissant convenablement $\mathbf{p}$, le système prend la forme :

$$\begin{cases} \mathbf{p} = \dfrac{\partial L}{\partial \dot{\mathbf{q}}} \\ \dot{\mathbf{p}} = \dfrac{\partial L}{\partial \mathbf{q}} \end{cases}$$

Remarque 6.1 Rappel. Usuellement (6.1) est écrit :

$$\frac{d}{dt}\begin{pmatrix} mx \\ m\dot{x} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{pmatrix}\begin{pmatrix} mx \\ m\dot{x} \end{pmatrix}, \qquad \text{i.e.} \qquad X' = AX, \tag{6.3}$$

où $X = \begin{pmatrix} mx_1 \\ mx_2 \end{pmatrix}$ et $A = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{pmatrix}$. La solution du système vérifie $\dot{x}_1 = x_2$ et $\dot{x}_2 = -\frac{k}{m}x_1$, donc $\ddot{x}_1 = -\frac{k}{m}x_1$ et donc la solution x cherchée est donnée par $x = x_1$.

Le système de Lagrange (6.2) ne cherche pas à résoudre ce système différentiel : le système (6.2) est une présentation différente de l'équation du second ordre $m\ddot{x} + kx = 0$ sous la forme $\begin{pmatrix} p \\ \dot{p} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{pmatrix}\begin{pmatrix} mx \\ m\dot{x} \end{pmatrix}$.
Les équations de Hamilton par contre chercherons à résoudre le système (6.3).

6.2 Calcul

Soit $L = E_c - E_p$ donné par (5.10).

Définition 6.2 Dans un système de coordonnées choisi (penser aux polaires), avec (5.10) on définit :

$$\mathbf{p} \stackrel{\text{déf}}{=} \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}}), \tag{6.4}$$

noté :

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}} \qquad (= \frac{\partial E_c}{\partial \dot{\mathbf{q}}}), \tag{6.5}$$

et $\mathbf{p}$ est appelé le moment généralisé (ou la quantité de mouvement généralisée).

(Le terme “moment généralisé” est mal choisi : on pourrait dire “moment paramétrique”.)

Exemple 6.3 En 2-D dans le système cartésien $\mathbf{p} = m\vec{v} = m\begin{pmatrix} x' \\ y' \end{pmatrix}$ est la quantité de mouvement.

Exemple 6.4 En polaires, $E_c = \frac{1}{2}m(r'^2 + r^2\theta'^2)$. Donc $\frac{\partial E_c}{\partial r'} = mr'$ et $\frac{\partial E_c}{\partial \theta'} = mr^2\theta'$. Donc

$$\mathbf{p} = m\begin{pmatrix} r' \\ r^2\theta' \end{pmatrix} \qquad (\neq m\vec{v})$$

est le moment généralisé. À comparer à la quantité de mouvement $m\vec{v} = m\begin{pmatrix} r'\cos\theta - r\theta'\sin\theta \\ r'\sin\theta + r\theta'\cos\theta \end{pmatrix}$.

L'équation d'Euler (5.11) se lit également :

$$\dot{\mathbf{p}} = \frac{\partial L}{\partial \mathbf{q}}, \tag{6.6}$$

d'où :

Définition 6.5 Les équations (6.5) et (6.6) forment le système d'équations de Lagrange :

$$\begin{cases} \mathbf{p} = \dfrac{\partial L}{\partial \dot{\mathbf{q}}}, \\ \dot{\mathbf{p}} = \dfrac{\partial L}{\partial \mathbf{q}}. \end{cases} \tag{6.7}$$

Le sens de (6.7) est : trouver la fonction la fonction $\mathbf{q} : [a, b] \to \mathbb{R}^n$ t.q. :

$$\begin{cases} \mathbf{p}(t) = \dfrac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}(t), \dot{\mathbf{q}}(t)), \\ \dot{\mathbf{p}}(t) = \dfrac{\partial L}{\partial \mathbf{q}}(\mathbf{q}(t), \dot{\mathbf{q}}(t)), \end{cases} \tag{6.8}$$

système différentiel du premier ordre correspdant à l'équation différentielle du second ordre :

$$\frac{D}{Dt}\frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) = \frac{\partial L}{\partial \mathbf{q}}(\mathbf{q}(t), \dot{\mathbf{q}}(t)). \tag{6.9}$$

Exemple 6.6 Oscillateur harmonique 1-D : système de Lagrange $\begin{cases} p = m\dot{x} \\ \dot{p} = -kx \end{cases}$. D'où $\begin{cases} \dot{p} = m\ddot{x} \\ \dot{p} = -kx \end{cases}$, d'où $m\ddot{x} = -kx$ à résoudre.

7 Équations de Hamilton

7.1 Introduction

Ces équations ont été conçues pour les systèmes conservatifs.

Comme les équations de Lagrange, les équations de Hamilton sont une réécriture du principe fondamental "$\vec{f} = m\vec{\gamma}$" pour un système conservatif.

Mais la différence essentielle est : le système de Hamilton est un système différentiel du premier ordre de type :

$$\begin{cases} \dot{\mathbf{q}} = \dfrac{\partial H}{\partial \mathbf{p}} \\ \dot{\mathbf{p}} = -\dfrac{\partial H}{\partial \mathbf{q}} \end{cases} \qquad \text{de type} \qquad \frac{d}{dt}\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = A.\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix},$$

alors que le système de Lagrange est une réécriture de l'équation du second ordre, voir remarque 6.1.

Par exemple pour le ressort en 1-D, le système de Hamilton sera donné par (6.3) dans le cas cartésien sous la forme :

$$\frac{d}{dt}\begin{pmatrix} x \\ m\dot{x} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{m} \\ -k & 0 \end{pmatrix}\begin{pmatrix} x \\ m\dot{x} \end{pmatrix}, \tag{7.1}$$

puisque dans ce cas $q = x$ et $p = m\dot{x}$.

7.2 Hamiltonien : approche élémentaire

7.2.1 Introduction en cartésien

La fonction de Lagrange $L = E_c - E_p$ a un signe "$-$" alors que l'énergie totale $E = E_c + E_p$ a un signe "$+$" :

$$L = \frac{1}{2}mv^2 - E_p \qquad \text{vs} \qquad E = \frac{1}{2}mv^2 + E_p,$$

et comme on s'intéresse aux systèmes conservatifs on a $E =$ constante, relation qui nous intéresse. Or on a :

$$E = mv^2 - L,$$

au sens $E(\vec{x}, \vec{v}) = m||\vec{v}||^2 - L(\vec{x}, \vec{v})$.

Pour se ramener à un système différentiel du premier ordre de type (7.1), on décide de faire un changement de variable :

$$\mathbf{p} \overset{\text{déf}}{=} m\vec{v} \qquad \Leftrightarrow \qquad \vec{v} = \frac{\mathbf{p}}{m},$$

on obtient :

$$\begin{aligned} E(\vec{x}, \vec{v}) = \mathbf{p}.\vec{v} - L(\vec{x}, \vec{v}) &= \frac{1}{m}||\mathbf{p}||^2 - L(\vec{x}, \frac{\mathbf{p}}{m}) \\ &\overset{\text{noté}}{=} H(\vec{x}, \mathbf{p}), \end{aligned} \tag{7.2}$$

appelé hamiltonien. Autrement dit :

$$H(\vec{x}, \mathbf{p}) = \frac{1}{m}||\mathbf{p}||^2 - L(\vec{x}, \frac{\mathbf{p}}{m}), \tag{7.3}$$

et on a récupéré le "bon signe $+$" :

$$H(\vec{x}, \mathbf{p}) = \frac{1}{2}mv^2 + E_p(\vec{x}) = E_c(\vec{v}) + E_p(\vec{x}), \tag{7.4}$$

au nom des variables près...

On décidera de faire systématiquement le changement de variable $\vec{v} \leftrightarrow \mathbf{p}$ où $\mathbf{p}$ est défini par $\mathbf{p} = \dfrac{\partial L}{\partial \dot{q}}$.

7.2.2 Dans un système de coordonnées

Rappel. Soit $\mathbf{q}$ un système de coordonnées, et $\vec{\varphi}(\mathbf{q}) = \vec{x} =^{\text{noté}} \vec{x}(\mathbf{q})$ le changement de système de coordonnées vers les cartésiens (vers la représentation géométrique).

Exemple 7.1 En polaires, $\mathbf{q} = (r, \theta)$ et $\vec{x} = \vec{\varphi}(\mathbf{q}) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$. ▪

Une trajectoire est donnée par $\vec{r}(t) = \vec{x}(t) = \vec{\varphi}(\mathbf{q}(t))$, et la vitesse par

$$\vec{v}(t, \vec{x}(t)) \stackrel{\text{déf}}{=} \vec{r}\,'(t) = d\vec{\varphi}(\mathbf{q}(t)).\dot{\mathbf{q}}(t). \tag{7.5}$$

Définissons la fonction $\widetilde{E}_p$ par :

$$\widetilde{E}_p(\mathbf{q}) = E_p(\vec{x}) \qquad (= E_p(\vec{\varphi}(\mathbf{q}))), \qquad \text{i.e.} \qquad \widetilde{E}_p \stackrel{\text{déf}}{=} E_p \circ \vec{\varphi}, \tag{7.6}$$

et la fonction $\widetilde{E}_c$ par :

$$\widetilde{E}_c(\mathbf{q}, \dot{\mathbf{q}}) = E_c(\vec{v}) \qquad (= E_c(d\vec{\varphi}(\mathbf{q}).\dot{\mathbf{q}})). \tag{7.7}$$

Ainsi le lagrangien s'exprime dans les nouvelles coordonnées à l'aide de la fonction :

$$\widetilde{L}(\mathbf{q}, \dot{\mathbf{q}}) = \widetilde{E}_c(\mathbf{q}, \dot{\mathbf{q}}) - \widetilde{E}_p(\mathbf{q}) \qquad \text{donc} \qquad = L(\vec{x}, \vec{v}) = E_c(\vec{v}) - E_p(\vec{x}).$$

On définit alors :

$$\mathbf{p} \stackrel{\text{déf}}{=} \frac{\partial \widetilde{L}}{\partial \dot{\mathbf{q}}} \qquad (= \frac{\partial \widetilde{E}_c}{\partial \dot{\mathbf{q}}}), \tag{7.8}$$

appelé moment généralisé (on devrait plutôt dire moment paramétrique, voir plus haut).

On va montrer, dans le cadre de la dynamique des systèmes conservatifs usuels, que cette relation va permettre de faire un changement de variable $\mathbf{p} \leftrightarrow \dot{\mathbf{q}}$, en chaque $\mathbf{q}$ donné. On définit alors l'hamiltonien par :

$$\widetilde{H}(\mathbf{q}, \mathbf{p}) = \mathbf{p}.\mathbf{q} - \widetilde{L}(\mathbf{q}, \dot{\mathbf{q}}), \tag{7.9}$$

faire le parallèle avec (7.2).

N.B. : dans la suite on omettra les tildes$^{\sim}$ pour alléger l'écriture.

Exemple 7.2 En polaire on a $\mathbf{q} = (r, \theta)$ et $\dot{\mathbf{q}} = (r', \theta')$, et on a $\widetilde{E}_c(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}m(r'^2 + r^2\theta'^2)$. Avec (7.8) on a $\mathbf{p} = (p_1, p_2) = (\frac{\partial \widetilde{L}}{\partial r'}, \frac{\partial \widetilde{L}}{\partial \theta'})$, le moment généralisé, avec donc $p_1 = \frac{\partial E_c}{\partial r'} = mr'$ et $p_2 = \frac{\partial E_c}{\partial \theta'} = mr^2\theta'$.

Donc également $r' = \frac{p_1}{m}$ et $\theta' = \frac{p_2}{mr^2}$.

On a ainsi le changement de variable $\mathbf{p} \leftrightarrow \dot{\mathbf{q}}$ en chaque $\mathbf{q}$.

D'où l'hamiltonien :

$$H(\mathbf{q}, \mathbf{p}) = \mathbf{p}.\mathbf{q} - (\widetilde{E}_c(\mathbf{q}, \dot{\mathbf{q}}) - \widetilde{E}_p(\mathbf{q})) = p_1 r + p_2\theta - \frac{1}{2m}(p_1^2 + \frac{p_2^2}{r^2}) + \widetilde{E}_p(\mathbf{q}),$$

les variables $\mathbf{q}$ et $\mathbf{p}$ étant ici indépendantes. Autrement dit, posant :

$$H((q_1, q_2), (p_1, p_2)) = q_1 p_1 + q_2 p_2 - \frac{1}{2m}(p_1^2 + \frac{p_2^2}{q_1^2}) + \widetilde{E}_p((q_1, q_2)),$$

la fonction H définit notre hamiltonien en polaires. ▪

7.3 Système de Hamilton

Proposition 7.3 *Dans le cadre de la dynamique des systèmes conservatifs usuels, cf. (5.8), la différentielle de E_c en $\dot{\mathbf{q}}$ est linéaire, avec :*

$$\frac{\partial E_c}{\partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}}) = A(\mathbf{q}).\dot{\mathbf{q}} = \mathbf{p}. \tag{7.10}$$

Et ayant $A(\mathbf{q})$ symétrique définie positive, on a ainsi défini un changement de variables $\mathbf{p} \leftrightarrow \dot{\mathbf{q}}$ donné par :

$$\mathbf{p} = A(\mathbf{q}).\dot{\mathbf{q}} \qquad \Leftrightarrow \qquad \dot{\mathbf{q}} = A(\mathbf{q})^{-1}.\mathbf{p}. \tag{7.11}$$

Preuve. Soit $\vec{y}$ fixé, et soit $g(\vec{z}) = E_c(\vec{y},\vec{z})$. On a $g(\vec{z}) = \frac{1}{2}(A(\vec{y}).\vec{z},\vec{z})_{\mathbb{R}^n}$, d'où $(\vec{\nabla}g(\vec{z}) = A(\vec{y}).\vec{z}$ puisque $A(\vec{y})$ est symétrique.

N.B. : on retrouve la définition (7.8) de $\mathbf{p}$. ▪

Remarque 7.4 On a également :

$$\frac{\partial E_c}{\partial \dot{\mathbf{q}}}.\dot{\mathbf{q}} = 2E_c = \mathbf{p}.\mathbf{q}, \qquad \text{où} \qquad \mathbf{p} = \frac{\partial E_c}{\partial \dot{\mathbf{q}}} = A(\mathbf{q}).\dot{\mathbf{q}}, \tag{7.12}$$

au sens $\frac{\partial E_c}{\partial \dot{\mathbf{q}}}(\mathbf{q},\dot{\mathbf{q}}).\dot{\mathbf{q}} = 2E_c(\mathbf{q},\dot{\mathbf{q}}) = \mathbf{p}.\mathbf{q}$. ▪

Définition 7.5 On se place dans le cadre de la dynamique des systèmes conservatifs usuels. Disposant du système de coordonnées $\mathbf{q}$, de :

$$L(\mathbf{q},\dot{\mathbf{q}}) = E_c(\mathbf{q},\dot{\mathbf{q}}) - E_p(\mathbf{q}), \tag{7.13}$$

et de :

$$\mathbf{p} = A(\mathbf{q}).\dot{\mathbf{q}}, \tag{7.14}$$

cf. (7.12), on pose :

$$H(\mathbf{q},\mathbf{p}) \stackrel{\text{déf}}{=} \mathbf{p}.\dot{\mathbf{q}} - L(\mathbf{q},\dot{\mathbf{q}}). \tag{7.15}$$

La fonction H est appelé hamiltonien.

On a donc :

$$H(\mathbf{q},\mathbf{p}) = \mathbf{p}.(A(\mathbf{q})^{-1}.\mathbf{p}) - L(\mathbf{q}, A(\mathbf{q})^{-1}.\mathbf{p}). \tag{7.16}$$

Exemple 7.6 En coordonnées cartésiennes, on a $\mathbf{p} = m\vec{v}$, on a $H = \frac{\vec{v}}{m}.\vec{v} - \frac{1}{2}m\frac{p^2}{m^2} + E_p(\vec{x})$, soit $H(\vec{x},\mathbf{p}) = \frac{1}{2m}p^2 + E_p(\vec{x})$, où $p^2 = ||\mathbf{p}||^2_{\mathbb{R}^n}$. ▪

Exercice 7.7 Montrer que si on pose :

$$H(\mathbf{q},\mathbf{p},\dot{\mathbf{q}}) \stackrel{\text{déf}}{=} \mathbf{p}.\dot{\mathbf{q}} - L(\mathbf{q},\dot{\mathbf{q}}), \tag{7.17}$$

on a bien, grâce à la définition de $\mathbf{p}$:

$$\frac{\partial H}{\partial \dot{\mathbf{q}}} = 0, \tag{7.18}$$

et donc $H(\mathbf{q},\mathbf{p},\dot{\mathbf{q}}) = H(\mathbf{p},\dot{\mathbf{q}})$.

Réponse. $\frac{\partial H}{\partial \dot{\mathbf{q}}} = \mathbf{p} - \frac{\partial L}{d\dot{\mathbf{q}}} = 0$ par définition de $\mathbf{p}$. ▪

Proposition 7.8 *On se place dans le cadre de la dynamique des systèmes usuels. Ayant* $\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}$, *on a :*

$$H = E_c + E_p, \tag{7.19}$$

i.e. H est l'énergie totale du système exprimé dans les variables $(\mathbf{q},\mathbf{p})$.

N.B. : le sens de (7.19) est $H(\mathbf{q},\mathbf{p}) = E_c(\mathbf{q},\dot{\mathbf{q}}) + E_p(\mathbf{q})$ avec $\dot{\mathbf{q}} = A(\mathbf{q})^{-1}.\mathbf{p}$.

Preuve. On applique (7.12). ▪

Théorème 7.9 *Avec $\mathbf{p}$ défini par (6.5), le système de Lagrange (6.7) est équivalent au système de Hamilton défini par :*

$$\begin{cases} \dot{\mathbf{q}} = \dfrac{\partial H}{\partial \mathbf{p}}, \\[2mm] \dot{\mathbf{p}} = -\dfrac{\partial H}{\partial \mathbf{q}}. \end{cases} \tag{7.20}$$

Le sens de (7.20) est : trouver la fonction $\mathbf{q} : t \to \mathbf{q}(t)$ t.q. pour tout $t \in [a,b]$:

$$\begin{cases} \dot{\mathbf{q}}(t) = \dfrac{\partial H}{\partial \mathbf{p}}(\mathbf{q}(t),\mathbf{p}(t)), \\[2mm] \dot{\mathbf{p}}(t) = -\dfrac{\partial H}{\partial \mathbf{q}}(\mathbf{q}(t),\mathbf{p}(t)), \end{cases}$$

et s'exprime également sous la forme : la solution $\vec{\varphi} : t \to \vec{\varphi}(t) = \begin{pmatrix} \mathbf{q}(t) \\ \mathbf{p}(t) \end{pmatrix}$ est une courbe intégrale du champ de vecteurs $X_H = \begin{pmatrix} \frac{\partial H}{\partial \mathbf{p}} \\ -\frac{\partial H}{\partial \mathbf{q}} \end{pmatrix}$, i.e. $\frac{d\vec{\varphi}}{dt}(t) = X_H(\vec{\varphi}(t))$.

Preuve. Avec (7.15) on a :

$$dH = \frac{\partial H}{\partial \mathbf{p}}\, d\mathbf{p} + \frac{\partial H}{\partial \mathbf{q}}\, d\mathbf{q} = \dot{\mathbf{q}}\, d\mathbf{p} - \frac{\partial L}{\partial \mathbf{q}}\, d\mathbf{q},$$

donc $\frac{\partial H}{\partial \mathbf{p}} = \dot{\mathbf{q}}$ et $\frac{\partial H}{\partial \mathbf{q}} = -\frac{\partial L}{\partial \mathbf{q}}$, et les équations de Lagrange donnent $\frac{\partial L}{\partial \mathbf{q}} = \dot{\mathbf{p}}$. ▪

Exemple 7.10 Oscillateur harmonique 1-D : on a $p = \frac{\partial L}{\partial \dot{x}} = \frac{\partial E_c}{\partial \dot{x}} = mv$. D'où $E_c = \frac{1}{2}\frac{1}{m}p^2$, avec $E_p = \frac{1}{2}kx^2$, donc $H(p,x) = \frac{1}{2}\frac{1}{m}p^2 + \frac{1}{2}kx^2$.

D'où le système de Hamilton $\begin{cases} \dot{p} = -kx \\ \dot{x} = \dfrac{p}{m} \end{cases}$, i.e. (7.1). ▪

7.4 Approche intrinsèque

On rappelle que $\mathbb{R}^{n*} = L(\mathbb{R}^n;\mathbb{R})$ est l'ensemble des formes linéaires sur $\mathbb{R}^n$ (applications linéaires à valeurs dans $\mathbb{R}$) et est appelé le dual de $\mathbb{R}^n$.

7.4.1 Introduction

La définition de $\mathbf{p}$ donnée en (6.4) est basée sur la différentielle du lagrangien $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$: notant $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \in (\mathbb{R}^n)^2$, $V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \in (\mathbb{R}^n)^2$ et $h \in \mathbb{R}$:

$$L(X + hV) = L(X) + h\, dL(X).V + o(h) \tag{7.21}$$

où $dL(X) \in (\mathbb{R}^{2n})^*$ (forme linéaire sur $\mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$) par définition de l'application linéaire tangente, soit :

$$L((X_1, X_2) + h(V_1, V_2)) = L(X_1, X_2) + h\,(dL_1(X), dL_2(X)).\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} + o(h). \tag{7.22}$$

et on a noté $dL(X) = (dL_1(X), dL_2(X))$ où les $dL_i(X) \in \mathbb{R}^{n*}$.

Une autre notation usuelle est $dL_i(X) = \frac{\partial L}{\partial X_i}(X)$. On notera donc :

$$dL(\mathbf{q}, \dot{\mathbf{q}}) = (dL_1(\mathbf{q}, \dot{\mathbf{q}}), dL_2(\mathbf{q}, \dot{\mathbf{q}})) = (\frac{\partial L}{\partial \mathbf{q}}(\mathbf{q}, \dot{\mathbf{q}}), \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}})).$$

Et on posera le long d'une trajectoire solution du principe fondamental (i.e. du système de Lagrange pour les systèmes conservatifs) :

$$\mathbf{p}(t) = \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}(t), \dot{\mathbf{q}}(t)), \tag{7.23}$$

noté :

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}. \tag{7.24}$$

Donc ici $\mathbf{p}(t) \in \mathbb{R}^{n*} = L(\mathbb{R}^n;\mathbb{R})$ est une forme linéaire. Ce n'est pas un vecteur.

Remarque 7.11 On peut représenter les formes linéaires par des vecteurs à condition de disposer d'un produit scalaire, mais alors les expressions de ces vecteurs dépendent du produit scalaire choisi : on perd le caractère intrinsèque (= indépendant de l'observateur).

Exemple. Soit $f : \vec{x} \in \mathbb{R}^n \to \frac{1}{2}||\vec{x}||_M^2 = \frac{1}{2}(M.\vec{x}, \vec{x})_{\mathbb{R}^n} =^{\text{noté}} \frac{1}{2}(\vec{x}, \vec{x})_M$ où M est une matrice symétrique définie positive. On a :

$$f(\vec{x} + h\vec{v}) = f(\vec{x}) + h\, df(\vec{x}).\vec{v} + o(h),$$

où $df(\vec{x})$ est donnée par $df(\vec{x}).\vec{v} = (M.\vec{x}, \vec{v})_{\mathbb{R}^n} = (\vec{x}, \vec{v})_M$. Donc $df(\vec{x})$ est l'application linéaire représentée comme $df(\vec{x}).\vec{v} = (\vec{y}, \vec{v})_{??}$:

1- par le vecteur $\vec{y} = M.\vec{x}$ relativement au produit scalaire canonique $(\cdot,\cdot)_{\mathbb{R}^n}$, ainsi que

2- par le vecteur $\vec{y} = \vec{x}$ relativement au produit scalaire $(\cdot,\cdot)_M$.

Et $\vec{y}$ dépend explicitement du choix du produit scalaire. ▪

7.4.2 Hamiltonien

Soit une fonction C^2 :

$$H : \begin{cases} \mathbb{R}^n \times \mathbb{R}^{n*} & \to \mathbb{R} \\ \quad (\mathbf{q}, \ell) & \mapsto H(\mathbf{q}, \ell). \end{cases} \tag{7.25}$$

Sa différentielle $dH(\mathbf{q},\ell) = (\frac{\partial H}{\partial \mathbf{q}}(\mathbf{q},\ell), \frac{\partial H}{\partial \ell}(\mathbf{q},\ell))$ en un point $(\mathbf{q},\ell)$ est donnée par :

$$\begin{aligned} H(\mathbf{q}+h\vec{w}, \ell+hm) &= H(\mathbf{q},\ell) + h\, dH(\mathbf{q},\ell).\begin{pmatrix}\vec{w}\\ m\end{pmatrix} + o(h) \\ &= H(\mathbf{q},\ell) + h\,\frac{\partial H}{\partial \mathbf{q}}(\mathbf{q},\ell).\vec{w} + h\,\frac{\partial H}{\partial \ell}(\mathbf{q},\ell).m + o(h) \end{aligned}$$

où $\frac{\partial H}{\partial \mathbf{q}}(\mathbf{q},\ell)$ est une forme différentielle, et où $\frac{\partial H}{\partial \ell}(\mathbf{q},\ell) \in \mathbb{R}^{n**} \simeq \mathbb{R}^n$ est un champ de vecteurs.

(On a $\frac{\partial H}{\partial \mathbf{q}} : \mathbb{R}^n \times \mathbb{R}^{n*} \to \mathbb{R}^{n*}$ et $\frac{\partial H}{\partial \ell} : \mathbb{R}^n \times \mathbb{R}^{n*} \to \mathbb{R}^{n**} \simeq \mathbb{R}^n$.)

On se donne le lagrangien L sur $\mathbb{R}^n \times \mathbb{R}^n$ (système supposé conservatif) :

$$L(\mathbf{q}, \dot{\mathbf{q}}) = E_c(\mathbf{q}, \dot{\mathbf{q}}) - E_p(\mathbf{q}).$$

On pose par définition le long d'une solution $t \to \mathbf{q}(t)$ du système de Lagrange (qui donne une solution du principe fondamental dans le cas conservatif) :

$$\mathbf{p}(t) = \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \quad \in \mathbb{R}^{n*}, \tag{7.26}$$

où ici $\frac{\partial L}{\partial \dot{\mathbf{q}}}$ désigne la différentielle (donc pour $L : (\vec{y}, \vec{z}) \to L(\vec{y}, \vec{z})$, on a $\frac{\partial L}{\partial \vec{z}}$ définie par $L((\vec{y}, \vec{z}+\vec{h}) = L(\vec{y}, \vec{z}) + \frac{\partial L}{\partial \vec{z}}(\vec{y}, \vec{z}).(0, \vec{h}) + o(h))$.

Et le long de ces courbes on pose :

$$H(\mathbf{q}, \mathbf{p}) = \mathbf{p}.\mathbf{q} - L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{p})) \overset{\text{noté}}{=} \mathbf{p}.\mathbf{q} - L(\mathbf{q}, \dot{\mathbf{q}}) \qquad (= \mathbf{p}.\mathbf{q} - E_c(\mathbf{q}, \dot{\mathbf{q}}) + E_p(\mathbf{q})). \tag{7.27}$$

N.B. : on a vérifié, dans le cas des systèmes conservatifs, que $\dot{\mathbf{q}} = \dot{\mathbf{q}}(\mathbf{p})$, cf. proposition 7.3 équation (7.11) qu'on réécrit de manière intrinsèque, d'où cette expression.

Les équations de Hamilton s'écrivent :

$$\begin{cases} \dfrac{d\mathbf{q}}{dt} = \dfrac{\partial H}{\partial \mathbf{p}} \quad \in \mathbb{R}^n, \\ \dfrac{d\mathbf{p}}{dt} = -\dfrac{\partial H}{\partial \mathbf{q}} \quad \in \mathbb{R}^{n*}, \end{cases} \tag{7.28}$$

au sens :

$$\begin{cases} \dfrac{d\mathbf{q}}{dt}(t) = \dfrac{\partial H}{\partial \mathbf{p}}(\mathbf{q}(t), \mathbf{p}(t)) \quad \in \mathbb{R}^n, \\ \dfrac{d\mathbf{p}}{dt}(t) = -\dfrac{\partial H}{\partial \mathbf{q}}(\mathbf{q}(t), \mathbf{p}(t)) \quad \in \mathbb{R}^{n*}. \end{cases} \tag{7.29}$$

Cette expression est intrinsèque, et s'exprime également sous la forme : la courbe $t \to (\mathbf{q}(t), \mathbf{p}(t)) \in \mathbb{R}^n \times \mathbb{R}^{n*}$ est une courbe intégrale du champ de vecteur $X_H : \mathbb{R}^n \times \mathbb{R}^{n*} \to \mathbb{R}^{n**} \times \mathbb{R}^{n*}$ donné par :

$$X_H(\mathbf{q}, \mathbf{p}) = \begin{pmatrix} \dfrac{\partial H}{\partial \mathbf{p}}(\mathbf{q}, \mathbf{p}) \\ -\dfrac{\partial H}{\partial \mathbf{q}}(\mathbf{q}, \mathbf{p}) \end{pmatrix}, \qquad \text{soit} \qquad X_H = \begin{pmatrix} \dfrac{\partial H}{\partial \mathbf{p}} \\ -\dfrac{\partial H}{\partial \mathbf{q}} \end{pmatrix}. \tag{7.30}$$

i.e. (7.28) s'écrit $\dfrac{d}{dt}\begin{pmatrix}\mathbf{q}\\ \mathbf{p}\end{pmatrix} = X_H(\mathbf{q}, \mathbf{p})$.

7.4.3 Transformée de Legendre

Voir poly Transformé de Legendre, ou Arnold [2] par exemple.

La transformée de Legendre permet de généraliser cette approche $\mathbf{p} \leftrightarrow \dot{\mathbf{q}}$ au cas où L est convexe en $\dot{\mathbf{q}}$ (le cas parabolique en $\dot{\mathbf{q}}$ étant le cas particulier traité dans le cadre de la dynamique des systèmes conservatifs usuels).

Soit à $\mathbf{q}$ fixé la fonction $L_{\mathbf{q}}(\dot{\mathbf{q}}) = L(\mathbf{q},\dot{\mathbf{q}})$, et soit $Z_{\mathbf{q}}(\mathbf{p},\dot{\mathbf{q}}) = (\mathbf{p},\dot{\mathbf{q}}) - L_{\mathbf{q}}(\dot{\mathbf{q}})$ pour $\mathbf{p} \in \mathbb{R}^{n*}$.

On pose $Y_{\mathbf{q},\mathbf{p}}(\dot{\mathbf{q}}) = Z_{\mathbf{q}}(\mathbf{p},\dot{\mathbf{q}})$, et on cherche un minimum de Y : un minimum est donné en un point $\dot{\mathbf{q}}$ tel que $\frac{\partial Z_{\mathbf{q}}}{\partial \dot{\mathbf{q}}}(\mathbf{p},\dot{\mathbf{q}}) = 0 = \mathbf{p} - \frac{\partial L_{\mathbf{q}}}{\partial \dot{\mathbf{q}}}(\dot{\mathbf{q}})$. Notons $\dot{\mathbf{q}}(\mathbf{p})$ le point trouvé (pour lequel $Y_{\mathbf{q},\mathbf{p}}(\dot{\mathbf{q}}(\mathbf{p}))$ donne un minimum de Y) : ce point vérifie

$$\frac{\partial L_{\mathbf{q}}}{\partial \dot{\mathbf{q}}}(\dot{\mathbf{q}}(\mathbf{p})) = \mathbf{p}.$$

On définit alors la transformée de Legendre de $\dot{\mathbf{q}} \to L_{\mathbf{q}}(\dot{\mathbf{q}})$ comme étant la fonction :

$$\mathbf{p} \to Z_{\mathbf{q}}(\mathbf{p},\dot{\mathbf{q}}(\mathbf{p})) \overset{\text{noté}}{=} H_{\mathbf{q}}(\mathbf{p}) = \overset{\text{noté}}{=} H(\mathbf{q},\mathbf{p}),$$

appelée hamiltonien du système.

La transformée de Legendre existe dès que $L_{\mathbf{q}}$ est convexe, ce qui est bien le cas dans le cadre de la dynamique des systèmes conservatifs usuels.

8 Forme symplectique et système de Hamilton

8.1 Forme symplectique

8.1.1 Espace des phases

On se place dans $\mathbb{R}^n$, i.e. on cherche les trajectoires $t \in [0,T] \to \vec{x}(t) \in \mathbb{R}^n$ suivies par des particules.

Le principe fondamental :

$$\vec{f}(t,\vec{x}(t)) = m\frac{d^2\vec{x}}{dt^2}(t) \tag{8.1}$$

est un système différentiel du second ordre dans $\mathbb{R}^n$ qui se résout à l'aide des condition initiales $\vec{x}(0)$ et $\frac{d\vec{x}}{dt}(0) = \vec{v}(0,\vec{x}(0))$.

L'espace des phases est l'ensemble dans lequel on prend les conditions initiales, soit ici

$$\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n = \{\text{position initiale}\} \times \{\text{vitesse initiale}\}.$$

Et, dès que $\vec{f}$ est Lipschitz en $\vec{x}$ uniformément en t, le théorème de Cauchy–Lipschitz nous donne une bijection entre $\mathcal{P}$ et l'ensemble des solutions :

$$\begin{pmatrix} \vec{x}_0 \\ \vec{v}_0 \end{pmatrix} \in \mathbb{R}^{2n} \quad \Leftrightarrow \quad \exists!\, \vec{x} \text{ solution de (8.1).}$$

Ainsi l'espace des solutions est de dimension $2n$.

Remarque 8.1 Dans le langage de la géométrie différentielle, l'espace des phases est le fibré tangent.

8.1.2 Forme symplectique usuelle

Soit $\vec{\varphi}_1 : \mathbf{q} = (q_1,...,q_n) \in \mathbb{R}^n \to \vec{\varphi}_1(\mathbf{q}) = \vec{x} \in \mathbb{R}^n$ et $\vec{\varphi}_2 : \vec{z} = (z_1,...,z_n) \in \mathbb{R}^n \to \vec{\varphi}_2(\vec{z}) = \vec{y} \in \mathbb{R}^n$ deux systèmes de coordonnées sur $\mathbb{R}^n$ (les domaines de définitions sont des sous-ensembles de $\mathbb{R}^n$).

Ainsi $(\vec{\varphi}_1,\vec{\varphi}_2)$ est un système de coordonnées sur $\mathbb{R}^{2n}$ l'espace des phases, et on dispose d'une base $(\vec{e}_1,...,\vec{e}_n,\vec{e}_{n+1},...,\vec{e}_{2n})$ de $\mathbb{R}^{2n}$, et d'une base duale $(dq^1,...,dq^n,dz^1,...,dz^n)$ de $\mathbb{R}^{2n*}$.

N.B. : ici, il y a ici un problème de notation : un élément $\ell \in (\mathbb{R}^{2n})^*$ agit sur les vecteurs $\vec{\xi} = (\vec{\xi}_1,\vec{\xi}_2) \in \mathbb{R}^n \times \mathbb{R}^n$, i.e. $\ell.\vec{\xi} = \ell.(\vec{\xi}_1,\vec{\xi}_2)$ a un sens.

Or $dq^i \in \mathbb{R}^{n*}$ donc $dq^i \notin (\mathbb{R}^{2n})^*$. Dire que $(dq^1,...)$ est une base de $(\mathbb{R}^{2n})^*$ est donc absurde. On commence par définir les $\widetilde{d}$, pour tout $\vec{\xi} = (\vec{\xi}_1,\vec{\xi}_2) \in \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$:

$$\begin{cases} \widetilde{d}q^i.\vec{\xi} = \widetilde{d}q^i.(\vec{\xi}_1,\vec{\xi}_2) \overset{\text{déf}}{=} dq^i.\vec{\xi}_1, \\ \widetilde{d}z^i.\vec{\xi} = \widetilde{d}z^i.(\vec{\xi}_1,\vec{\xi}_2) \overset{\text{déf}}{=} dz^i.\vec{\xi}_2, \end{cases} \tag{8.2}$$

et il est clair que c'est de la base $(\widetilde{d}q^1,...,\widetilde{d}q^n,\widetilde{d}z^1,...,\widetilde{d}z^n)$ dont on voulait parler.

Pour alléger les notations, on enlève le$\tilde{}$, i.e on pose :

$$\begin{cases} dq^i.\vec{\xi} = dq^i.(\vec{\xi}_1, \vec{\xi}_2) \stackrel{\text{déf}}{=} dq^i.\vec{\xi}_1, \\ dz^i.\vec{\xi} = dz^i.(\vec{\xi}_1, \vec{\xi}_2) \stackrel{\text{déf}}{=} dz^i.\vec{\xi}_2. \end{cases} \tag{8.3}$$

C'est abus de notation est systématique (c'est "évident"...).

Soit $\omega : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \to \mathbb{R}$ la forme bilinéaire alternée définie par :

$$\omega = \sum_{i=1}^n dq^i \wedge dz^i \stackrel{\text{noté}}{=} d\mathbf{q} \wedge d\vec{z} \qquad (= \sum_{i=1}^n dq^i \otimes dz^i - dz^i \otimes dq^i), \tag{8.4}$$

où on utilise les notations (8.3).

Donc si $\vec{\xi} = (\vec{\xi}_1, \vec{\xi}_2) \in \mathbb{R}^n \times \mathbb{R}^n$ et $\vec{\eta} = (\vec{\eta}_1, \vec{\eta}_2) \in \mathbb{R}^n \times \mathbb{R}^n$, on a :

$$\begin{aligned} \omega(\vec{\xi}, \vec{\eta}) = \omega((\vec{\xi}_1, \vec{\xi}_2), (\vec{\eta}_1, \vec{\eta}_2)) &= \sum_{i=1}^n (dq^i.\vec{\xi})(dz^i.\vec{\eta}) - (dz^i.\vec{\xi})(dq^i.\vec{\eta}) \\ &= \sum_{i=1}^n \xi_1^i \eta_2^i - \xi_2^i \eta_1^i \stackrel{\text{noté}}{=} [\vec{\xi}_1]^T.[\vec{\eta}_2] - [\vec{\xi}_2]^T.[\vec{\eta}_1], \end{aligned} \tag{8.5}$$

toujours avec les notations (8.3), les composantes ξ^i et η^i étant clairement prises relativement à la base $(\vec{e}_i)$ du système de coordonnées, i.e. on a posé $\vec{\xi} = \sum_i \xi^i \vec{e}_i$ et $\vec{\eta} = \sum_i \eta^i \vec{e}_i$.

La forme bilinéaire ω est donc représentée dans le système de coordonnées $(\mathbf{q}, \vec{z})$ par la matrice :

$$[\omega] = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \tag{8.6}$$

au sens usuel :

$$\omega((\vec{\xi}_1, \vec{\xi}_2), (\vec{\eta}_1, \vec{\eta}_2)) = \begin{pmatrix} [\vec{\xi}_1]^T & [\vec{\xi}_2]^T \end{pmatrix}.[\omega].\begin{pmatrix} \vec{\eta}_1 \\ \vec{\eta}_2 \end{pmatrix}, \tag{8.7}$$

puisque : $[\omega].\begin{pmatrix} \vec{\eta}_1 \\ \vec{\eta}_2 \end{pmatrix} = \begin{pmatrix} \vec{\eta}_2 \\ -\vec{\eta}_1 \end{pmatrix}$.

En particulier $[\omega].\begin{pmatrix} \vec{0} \\ \vec{e}_j \end{pmatrix} = \begin{pmatrix} \vec{e}_j \\ \vec{0} \end{pmatrix}$, et donc $\omega((\vec{e}_i, \vec{0}), (\vec{0}, \vec{e}_j)) = \delta_{ij}$ est l'élément $(i, n{+}j)$ de la matrice $2n * 2n$ représentant ω.

Remarque 8.2 La matrice de Ω est une "matrice de conservation", qui quand $n = 1$ n'est autre que $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ la matrice de rotation d'un quart de tour.

8.2 Hamiltonien et forme symplectique canonique

8.2.1 Espace des conditions initiales du système hamiltonien

Le système différentiel du second ordre (8.1) de taille n a été transformé en système différentiel hamiltonien du premier ordre (7.28) de taille $2n$. Pour ce système, l'ensemble des conditions initiales est :

$$\text{CIH} = \mathbb{R}^n \times \mathbb{R}^{n*} = \{X_H(0) = \begin{pmatrix} \mathbf{q}(0) \\ \mathbf{p}(0) \end{pmatrix}\}. \tag{8.8}$$

(Condition Initial pour l'Hamiltonien : ce n'est donc pas l'espace des phases.)

L'ensemble des solutions est :

$$\mathcal{X}_H = \{t \in [0,T] \to X_H(t) = \begin{pmatrix} \mathbf{q}(t) \\ \mathbf{p}(t) \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^{n*}\}, \tag{8.9}$$

en bijection avec CIH.

8.2.2 Forme symplectique canonique pour l'hamiltonien

On reprend la démarche du paragraphe 8.1.2, ou la différence essentielle est qu'on travaille sur l'espace $\mathbb{R}^n \times \mathbb{R}^{n*} =$ CIH et non sur l'espace $\mathbb{R}^n \times \mathbb{R}^n$.

On part donc de la base duale $(dq^1, ..., dq^n)$ de $\mathbb{R}^{n*}$, et de la base biduale $(\partial_1, ..., \partial_n)$ de $(\mathbb{R}^{n*})^*$.

On dispose de la base duale $(dq^1, ..., dq^n, \partial_1, ..., \partial_n)$ de $(\mathbb{R}^n \times \mathbb{R}^{n*})^*$.

Comme précédemment on a une absurdité de notation. On commence par définir, pour tout $\vec{x} = (\vec{\xi}, \ell) \in \mathbb{R}^n \times \mathbb{R}^{n*} =$ CIH :

$$\begin{cases} \widetilde{dq}^i.\vec{x} = \widetilde{dq}^i.(\vec{\xi}, \ell) \stackrel{\text{déf}}{=} dq^i.\vec{\xi}, \\ \widetilde{\partial}_i.\vec{x} = \widetilde{\partial}_i.(\vec{\xi}, \ell) \stackrel{\text{déf}}{=} \partial_i.\ell, \end{cases} \tag{8.10}$$

et il est clair que c'est de la base $(\widetilde{dq}^1, ..., \widetilde{dq}^n, \widetilde{\partial}_1, ..., \widetilde{\partial}_n)$ dont on voulait parler. Pour alléger les notations, on enlève le $\widetilde{\ }$, i.e on pose :

$$\begin{cases} dq^i.\vec{x} = dq^i.(\vec{\xi}, \ell) \stackrel{\text{déf}}{=} dq^i.\vec{\xi}, \\ \partial_i.\vec{x} = \partial_i.(\vec{\xi}, \ell) \stackrel{\text{déf}}{=} \partial_i.\ell, \end{cases} \tag{8.11}$$

C'est abus de notation est systématique (c'est le sens "évident"...).

On rappelle que le bidual E^{**} d'un espace E de dimension finie est canoniquement isomorphe à E avec l'isomorphisme :

$$\partial \in E^{**} \leftrightarrow \vec{x} \in E, \qquad \text{où} \qquad \partial(\ell) = \ell.\vec{x} \stackrel{\text{noté}}{=} \vec{x}.\ell, \quad \forall \ell \in E^*. \tag{8.12}$$

D'où la manipulation "aisée" du bidual.

Soit ω : CIH $\to \mathbb{R}$ la forme bilinéaire alternée définie par :

$$\omega = \sum_{i=1}^n dq^i \wedge \partial_i \stackrel{\text{noté}}{=} d\mathbf{q} \wedge d\mathbf{p} \qquad (= \sum_{i=1}^n dq^i \otimes \partial_i - \partial_i \otimes dq^i), \tag{8.13}$$

où on a bien sûr fait appel aux notations (8.11).

Donc si $\vec{x} = (\vec{\xi}, \ell) \in$ CIH $= \mathbb{R}^n \times \mathbb{R}^{n*}$ et $\vec{y} = (\vec{\eta}, m) \in$ CIH $= \mathbb{R}^n \times \mathbb{R}^{n*}$, on a :

$$\begin{cases} \omega(\vec{x}, \vec{y}) = \displaystyle\sum_{i=1}^n (dq^i.\vec{x})(\partial_i.\vec{y}) - (\partial_i.\vec{x})(dq^i.\vec{y}) \\ \qquad = \displaystyle\sum_{i=1}^n \xi^i m_i - \ell_i \eta^i \\ \qquad = [m].[\vec{\xi}] - [\ell].[\vec{\eta}], \end{cases} \tag{8.14}$$

toujours avec les notations (8.3).

La forme bilinéaire ω est donc représenté dans le système de coordonnées $\vec{\varphi} : \mathbf{q} \to \vec{\xi}$ (donc pour la base duale associée à la base biduale) par la matrice :

$$[\omega] = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \tag{8.15}$$

au sens usuel :

$$\omega((\vec{\xi}, \ell), (\vec{\eta}, m)) = \begin{pmatrix} \vec{\xi} & \ell \end{pmatrix}.[\omega].\begin{pmatrix} \vec{\eta} \\ m \end{pmatrix}, \tag{8.16}$$

puisque : $[\omega].\begin{pmatrix} \vec{\eta} \\ m \end{pmatrix} = \begin{pmatrix} m \\ -\vec{\eta} \end{pmatrix}$ et que $\begin{pmatrix} \vec{\xi} & \ell \end{pmatrix}.\begin{pmatrix} m \\ -\vec{\eta} \end{pmatrix} = m.\vec{\xi} - \ell.\vec{\eta}$.

Définition 8.3 Pour une variété Ω caractérisée par un système de coordonnées $\vec{\varphi} : \mathbf{q} \in U \subset \mathbb{R}^n \to \vec{x} \in \Omega \subset \mathbb{R}^p$ (avec $p \geq n$), la forme symplectique ω définie par (8.13) est appelée la forme symplectique canonique sur Ω.

8.2.3 Forme différentielle canonique pour l'hamiltonien

Soit $\theta : U \times \mathbb{R}^{n*} \to (\mathbb{R}^n \times \mathbb{R}^{n*})^*$ défini par, pour un point $\vec{x} = \vec{\varphi}(\mathbf{q}) \in \Omega$ et une forme différentielle α sur Ω :

$$\theta(\mathbf{q}, \alpha) = \sum_i \alpha_i \, dq^i \overset{\text{noté}}{=} \mathbf{p} \, d\mathbf{q}, \tag{8.17}$$

donc avec :

$$\theta(\mathbf{q}, \alpha).(\vec{\xi}, \ell) = \sum_i \alpha_i \, \xi^i, \tag{8.18}$$

cf. notation (8.11).

Définition 8.4 La forme différentielle θ est appelée la forme différentielle canonique pour l'hamiltonien.

Remarque 8.5 Cette définition (8.17) de θ est donnée dans un ouvert de $\mathbb{R}^n$. Si on se place sur une variété Ω, la définition de θ est toujours donnée par (8.17), mais avec :

$$\theta : \Omega \times T_1^0(\Omega) \to T_1^0(T_0^1(\Omega) \times T_1^0(\Omega)),$$

donc dans (8.18) avec α et ℓ des formes différentielles et $\vec{\xi}$ un champ de vecteurs. ▪

On a :

$$d_{\text{ext}}\theta = \sum_i \partial_i \wedge dq^i \overset{\text{noté}}{=} d\mathbf{p} \wedge d\mathbf{q} = -\omega,$$

i.e., au signe près, la différentielle extérieure de θ est la forme symplectique canonique.

En effet :

$$d\theta(\mathbf{q}, \alpha).(\vec{\xi}, \ell) = \lim_{h \to 0} \frac{\theta(\mathbf{q}+h\vec{\xi}, \alpha + h\ell) - \theta(\mathbf{q}, \alpha)}{h} = \sum_i \ell_i \, dq^i,$$

donne :

$$(d\theta(\mathbf{q}, \alpha).(\vec{\xi}, \ell)).(\vec{\eta}, m) = \sum_i \ell_i \eta^i,$$

et donc :

$$d_{\text{ext}}\theta(\mathbf{q}, \alpha)((\vec{\xi}, \ell), (\vec{\eta}, m)) = \sum_i \ell_i \eta^i - m_i \xi^i,$$

qu'on compare avec (8.14).

A Rappels

A.1 Système de coordonnées et bases du système

Cas des variétés de dimension n dans $\mathbb{R}^p$ où $p \ge n$.

A.1.1 Système de coordonnées

Un système de coordonnées sur une variété de dimension n est une application C^∞ :

$$\vec{\varphi} : \begin{cases} U & \to \Omega \\ \mathbf{q} & \mapsto \vec{x} = \vec{\varphi}(\mathbf{q}) \end{cases}$$

où U est un ouvert de $\mathbb{R}^n$ et Ω un ensemble dans $\mathbb{R}^p$ pour un $p \ge n$.

U est appelé espace des paramètres et Ω espace géométrique, et on dit que $\vec{\varphi}$ est le système de coordonnées $\mathbf{q}$.

Exemple 1 : cartésien : $U = \Omega = \mathbb{R}^n$ et $\vec{\varphi} = I$ l'identité (avec donc $\vec{x} = \mathbf{q}$).

Exemple 2 : polaires (donc dans $\mathbb{R}^2$) : $U = \mathbb{R}_+^* \times \mathbb{R}$ et $\Omega = \mathbb{R}^2 - \{\vec{0}\}$, avec $\mathbf{q} = (r, \theta)$ et $\vec{x} = (x, y)$, avec $\vec{\varphi}(r, \theta) = \begin{pmatrix} x = r\cos\theta = x(r,\theta) = \varphi^1(r,\theta) \\ y = r\sin\theta = y(r,\theta) = \varphi^2(r,\theta) \end{pmatrix}$. Ici $n = p = 2$.

Et $\vec{\varphi} : U \to \vec{\varphi}(U)$ est un difféomorphisme si on se restreint à par exemple à $U = R_+^* \times]-\pi, \pi[$ (quand on ne s'intéresse pas aux points dans un voisinage du demi-axe des x négatifs), ou à $U = R_+^* \times]0, 2\pi[$ (quand on ne s'intéresse pas aux points dans un voisinage du demi-axe des x positifs).

Exemple 3 : dans $\mathbb{R}^2$, cercle $C(\vec{0}, R)$ centré en $\vec{0}$ de rayon R : ici $U = [0, 2\pi]$ et $\Omega = C(\vec{0}, R)$, avec $q = \theta$ et $\vec{x} \in C(\vec{0}, R)$, avec $\vec{\varphi}(\theta) = \begin{pmatrix} x = R\cos\theta = x(\theta) = \varphi^1(\theta) \\ y = R\sin\theta = y(\theta) = \varphi^2(\theta) \end{pmatrix}$ (autrement dit $\vec{\varphi}$ donne une paramétrisation du cercle). Ici $n = 1$ et $p = 2$: le cercle est une variété de dimension 1 dans $\mathbb{R}^2$.

A.1.2 Base du système de coordonnées

Étant donné un système de coordonnées $\mathbf{q}$, on dispose d'une base $(\vec{e}_i)_{i=1,\dots,n}$ du système de coordonnées définie par :

$$\vec{e}_i(\vec{x}) = \frac{\partial \vec{\varphi}}{\partial q^i}(\mathbf{q}), \qquad \text{quand} \qquad \vec{x} = \vec{\varphi}(\mathbf{q}). \tag{A.1}$$

Les $\vec{e}_i(\vec{x})$ sont les vecteurs tangents (non unitaires en général) aux lignes de coordonnées $t \to \vec{\varphi}(\mathbf{q}+h\vec{E}_i)$ aux points $\vec{x} = \vec{\varphi}(\mathbf{q})$, où $(\vec{E}_i)$ est la base canonique de $\mathbb{R}^n$.

Donc, si $d\vec{\varphi}(\mathbf{q}) : \mathbb{R}^n \to \mathbb{R}^p$ est l'application linéaire tangente, on a $\vec{e}_i(\vec{x}) = d\vec{\varphi}(\mathbf{q}).\vec{E}_i$, et si $[d\vec{\varphi}(\mathbf{q})] = [\frac{\partial \varphi^i}{\partial q^j}(\mathbf{q})]$ est la matrice jacobienne de $\vec{\varphi}$ (dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$), alors $\vec{e}_j(\vec{x})$ est donné par la j-ème colonne de $[d\vec{\varphi}(\mathbf{q})]$.

Exemple 1 : système cartésien : $[d\vec{\varphi}(\mathbf{q})] = I$ et $(\vec{e}_i)$ est la base canonique indépendante de $\vec{x}$.

Exemple 2 : polaires (donc dans $\mathbb{R}^2$) : $[d\vec{\varphi}(r, \theta)] = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$, et $\vec{e}_1(\vec{x}) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$, et $\vec{e}_2(\vec{x}) = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \end{pmatrix}$.

Exemple 3 : variété = cercle $C(\vec{0}, R)$: $[d\vec{\varphi}(\theta)] = \begin{pmatrix} -R\sin\theta \\ R\cos\theta \end{pmatrix}$, et $\vec{e}_1(\vec{x}) = \begin{pmatrix} -R\sin\theta \\ R\cos\theta \end{pmatrix}$, le seul vecteur dans la base, vecteur qui est tangent au cercle au point $\vec{x} = \vec{\varphi}(\theta)$ du cercle.

Et un vecteur $\vec{v}_{\vec{x}}$ du plan tangent en $\vec{x}$ à la variété s'exprime usuellement à l'aide de la base du système de coordonnées sous la forme $\vec{v}_{\vec{x}} = \sum_{i=1}^n v_{\vec{x}}^i \vec{e}_i(\vec{x})$.

Exemple 3 : le système de coordonnée du cercle est $q = \theta$, et un vecteur $\vec{v}_{\vec{x}}$ tangent au cercle en un point $\vec{x} = \vec{\varphi}(\theta)$ est de la forme $\vec{v}_{\vec{x}} = \lambda\vec{e}_1(\vec{x}) = \lambda \begin{pmatrix} -R\sin\theta \\ R\cos\theta \end{pmatrix}$ où $\lambda \in \mathbb{R}$.

A.1.3 Base duale du système de coordonnées

On dispose également de la base duale $(dq^1(\vec{x}), ..., dq^n(\vec{x}))$ définie en $\vec{x} = \vec{\varphi}(\mathbf{q})$, où, pour tout i, $dq^i(\vec{x})$ est la forme définie par, pour tout j,

$$dq^i(\vec{x})(\vec{e}_j(\vec{x})) = \delta^i_j, \qquad \text{noté} \qquad dq^i.\vec{e}_j = \delta^i_j, \tag{A.2}$$

où δ^i_j est le symbole de Kronecker.

Autrement dit, $dq^i(\vec{x})$ est la projection sur $\mathrm{Vect}\{\vec{e}_i(\vec{x})\}$ parallèlement aux autres directions $\vec{e}_j(\vec{x})$, pour tout $j \neq i$.

Donc, ayant $[d\vec{\varphi}(\mathbf{q})]^{-1}.[d\vec{\varphi}(\mathbf{q})] = [\delta^i_j]$, la forme linéaire $dq^i(\vec{x})$ est immédiatement donné par la i-ème ligne de $[d\vec{\varphi}(\mathbf{q})]^{-1}$.

Exemple 1 : $(dx^1, ..., dx^n)$ où dx^i est donc la projection sur la i-ème composante. Et la forme linéaire dx^1 est représentée par la matrice ligne $[dx^1] = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}$, le calcul matriciel donnant bien $[dx^1].[\vec{e}_1] = 1$, $[dx^1].[\vec{e}_2] = 0$, ...

Exemple 2 : matrice jacobienne $[d\vec{\varphi}(\mathbf{q})] = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$ d'inverse $[d\vec{\varphi}(\mathbf{q})]^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ \frac{-\sin\theta}{r} & \frac{\cos\theta}{r} \end{pmatrix}$.
Donc la base duale est $(dr, d\theta)$ où $dr(\vec{x})$ et $d\theta(\vec{x})$ sont représentées par les matrices lignes

$[dr(\vec{x})] = \cos\theta\, dx + \sin\theta\, dy = \begin{pmatrix} \cos\theta & \sin\theta \end{pmatrix}_{|dx,dy}$ quand $\vec{x} = \vec{\varphi}(\mathbf{q})$, et

$[d\theta(\vec{x})] = \frac{-\sin\theta}{r}\, dx + \frac{\cos\theta}{r}\, dy = \begin{pmatrix} \frac{-\sin\theta}{r} & \frac{\cos\theta}{r} \end{pmatrix}_{|dx,dy}$, quand $\vec{x} = \vec{\varphi}(\mathbf{q})$.

On vérifie bien que $[dr(\vec{x})].[\vec{e}_1(\vec{x})] = 1$, $[dr(\vec{x})].[\vec{e}_2(\vec{x})] = 0$, ainsi que $[d\theta(\vec{x})].[\vec{e}_1(\vec{x})] = 0$ et $[d\theta(\vec{x})].[\vec{e}_2(\vec{x})] = 1$.

Exemple 3 : $d\theta$ où
$[d\theta(\vec{x})] = \frac{-\sin\theta}{r}\,dx + \frac{\cos\theta}{r}\,dy = (\frac{-\sin\theta}{R} \quad \frac{\cos\theta}{R})_{|dx,dy}$
vérifie bien $d\theta(\vec{x}).\vec{e}_1(\vec{x}) = 1$.

A.1.4 Dérivée partielle : notation

Soit $\vec{\varphi} : \mathbf{q} \in U \to \vec{x} \in \Omega$ un système de coordonnées, soit $f : U \to \mathbb{R}$ une fonction, et soit $g : \Omega \to \mathbb{R}$ définie par (push-forward de f par $\vec{\varphi}$) :

$$g \stackrel{\text{déf}}{=} f \circ \vec{\varphi}^{-1}, \qquad \text{i.e.} \qquad g(\vec{x}) = f(\mathbf{q}) \quad \text{quand} \quad \vec{x} = \vec{\varphi}(\mathbf{q}). \tag{A.3}$$

Définition A.1 La notation $\frac{\partial g}{\partial q^i}(\vec{x})$ n'a pas de sens a priori puisque g est défini sur les $\vec{x}$ et non sur les $\mathbf{q}$: on lui donne par définition le sens :

$$\frac{\partial g}{\partial q^i}(\vec{x}) \stackrel{\text{déf}}{=} \frac{\partial f}{\partial q^i}(\mathbf{q}) \quad \text{quand} \quad \vec{x} = \vec{\varphi}(\mathbf{q}). \tag{A.4}$$

Cette notation est systématiquement implicitement utilisée en mécanique.

Exemple 2 Soit $f(r,\theta) = r^2$. Soit $g(x,y) = f(r,\theta) = r^2$ (donc $g(\vec{x}) = x^2+y^2$). Par définition on a $\frac{\partial g}{\partial r}(x,y) = 2r$ et $\frac{\partial g}{\partial \theta}(x,y) = 0$. (Donc $\frac{\partial g}{\partial r}(x,y) = 2\sqrt{x^2+y^2}$.)

A.2 Formes différentielles exactes

Ce sont les formes différentielles qui expriment que "localement les forces dérivent d'un potentiel".

A.2.1 Forme différentielle

Définition A.2 Si Ω est un ouvert de $\mathbb{R}^p$, une forme différentielle sur Ω est une fonction dans $C^\infty(\Omega; \Omega \times (\mathbb{R}^p)^*) =^{\text{noté}} T^0_1(\Omega)$.

Si Ω est une variété dans $\mathbb{R}^p$, voir **Exemple 3** suivant et cours de géométrie différentielle.

Exemple A.3 Soit $f \in C^1(\Omega;\mathbb{R})$, Ω ouvert de $\mathbb{R}^p$. Sa différentielle $\widetilde{d}f : \vec{x} \in \Omega \to (\vec{x}, df(\vec{x})) \in \Omega \times (\mathbb{R}^p)^*$ est une forme différentielle où on rappelle que df est définie en $\vec{x} \in \Omega$ par $df(\vec{x}) \in (\mathbb{R}^p)^*$ où :

$$df(\vec{x}).\vec{u} = \lim_{h\to 0} \frac{f(\vec{x}+h\vec{u}) - f(\vec{x})}{h} \quad \in \mathbb{R},$$

pour tout $\vec{u} \in \mathbb{R}^p$. Et pour alléger les notations on note $\widetilde{d}f = df$. ▪

Définition A.4 Si Ω est une variété de dimension n dans $\mathbb{R}^p$, une forme différentielle sur Ω est un tenseur de type $\binom{0}{1}$, i.e. une application dans $T^0_1(\Omega) =^{\text{déf}} C^\infty(\Omega; T\Omega^*)$ appelé ensemble des tenseurs de type $\binom{0}{1}$.

Exemple 3 : en un point $\vec{x}$ le "plan tangent" $T_{\vec{x}}\Omega$ (de dimension 1) est la droite tangente au cercle en $\vec{x}$, notée $(\vec{x}, \text{Vect}\{\vec{e}_1(\vec{x})\})$.

Soit $T\Omega = \bigcup_{\vec{x}\in\Omega} T_{\vec{x}}\Omega$ est l'union des plans tangents, ici l'union de toutes les droites tangentes au cercle, faire un dessin.

Un champ de vecteurs sur Ω est une application $\widetilde{\vec{v}}$, qui à $\vec{x} \in \Omega$ associe un couple $\widetilde{\vec{v}}(\vec{x}) = (\vec{x}, \vec{v}(\vec{x}))$ où $\vec{v}(\vec{x}) \in T_{\vec{x}}\Omega$. Et $\widetilde{\vec{v}}(\vec{x})$ est représenté par le vecteur $\vec{v}(\vec{x})$ dessiné au point $\vec{x}$. Pour alléger les notations on note $\widetilde{\vec{v}}_{\vec{x}} = \vec{v}_{\vec{x}}$, i.e. on omet le $\tilde{}$.

Par exemple $\widetilde{\vec{e}}_1 : \vec{x} \to \widetilde{\vec{v}}(\vec{x}) = (\vec{x}, \vec{e}_1(\vec{x})) =^{\text{noté}} \vec{e}_1(\vec{x})$: $\vec{e}_1$ est un champ de vecteurs sur le cercle.

(On rappelle qu'un vecteur peut être représenté en n'importe quel point, alors que pour un champ de vecteurs de valeur $\widetilde{\vec{v}}(\vec{x})$, le vecteur $\vec{v}(\vec{x})$ est nécessairement représenté au point $\vec{x}$, i.e. c'est le couple $(\vec{x}, \vec{v}(\vec{x}))$ qui a un sens). appelé ensemble des tenseurs de type $\binom{1}{0}$.

Une forme différentielle est une application $\widetilde{\alpha}$, qui à $\vec{x} \in \Omega$ associe un couple $\widetilde{\alpha}(\vec{x}) = (\vec{x}, \alpha(\vec{x}))$ où $\alpha(\vec{x})$ est une forme linéaire sur $T_{\vec{x}}\Omega$ (donc $\alpha(\vec{x}) \in (T_{\vec{x}}\Omega)^*$). Et la linéarité permet la notation $\alpha_{\vec{x}}(\vec{v}) = \alpha_{\vec{x}}.\vec{v}$.

Par exemple $dq^1 : \vec{x} \to (\vec{x}, dq^1(\vec{x})) =^{\text{noté}} dq^1(\vec{x})$: dq^1 est une forme différentielle sur le cercle.

A.2.2 Différentielle extérieure

La différentielle extérieure d'une forme différentielle $\alpha \in T_1^0(\Omega)$ est la forme bilinéaire alternée notée $d_{\rm ext}\alpha \in T_2^0(\Omega)$ définie par, pour tout $\vec{x} \in \Omega$:

$$d_{\rm ext}\alpha(\vec{x})(\vec{u},\vec{v}) = (d\alpha(\vec{x}).\vec{u}).\vec{v} - (d\alpha(\vec{x}).\vec{v}).\vec{u}, \tag{A.5}$$

où $d\alpha$ est la différentielle usuelle caractérisée par, pour tout $\vec{u} \in T_{\vec{x}}\Omega$:

$$d\alpha(\vec{x}).\vec{u} = \lim_{h\to 0} \frac{\alpha(\vec{x}+h\vec{u}) - \alpha(\vec{x})}{h} \quad \in T_1^0(\Omega), \tag{A.6}$$

avec donc $(d\alpha(\vec{x}).\vec{u}).\vec{v} = \lim_{h\to 0} \frac{\alpha(\vec{x}+h\vec{u}).\vec{v}-\alpha(\vec{x}).\vec{v}}{h} \in \mathbb{R}$, pour tout $\vec{u},\vec{v} \in T_{\vec{x}}\Omega$.

Remarque A.5 N.B. : quand la différentielle usuelle est notée Df (alors qu'ici elle est notée df), la différentielle extérieure est notée df (alors qu'ici elle est notée $d_{\rm ext}f$). ▟

A.2.3 Forme différentielle fermée et conservation

Définition A.6 Une forme différentielle $\alpha \in T_1^0(\Omega)$ est dit fermée dans Ω ssi :

$$d_{\rm ext}\alpha = 0.$$

Dans $\mathbb{R}^n$, avec la base duale $(dq^1, ..., dq^n)$ d'un système de coordonnées, toute forme différentielle s'exprime sous la forme :

$$\alpha = \sum_{i=1}^n \alpha_i dq^i,$$

au sens $\alpha(\vec{x}) = \sum_{i=1}^n \alpha_i(\vec{x})\, dq^i(\vec{x})$ pour $\vec{x} \in \Omega$, où les $\alpha_i \in C^\infty(\Omega;\mathbb{R})$.

Et la différentielle de α en un point $\vec{x}$ est donnée par :

$$d\alpha.\vec{u} = \sum_{i=1}^n d\alpha_i.\vec{u}\, dq^i = \sum_{i,j=1}^n \frac{\partial \alpha_i}{\partial q^j} u^j\, dq^i,$$

au sens $d\alpha(\vec{x}).\vec{u} = \sum_{i=1}^n d\alpha_i(\vec{x}).\vec{u}\, dq^i(\vec{x}) = \sum_{i,j=1}^n \frac{\partial \alpha_i}{\partial q^j}(\vec{x}) u^j\, dq^i(\vec{x})$ quand $\vec{u} = \sum_i u^i \vec{e}_i(\vec{x}) = (u^1, ..., u^n)_{|\vec{e}}$ dans la base $(\vec{e}_i(\vec{x}))_{i=1,...,n}$ du système au point $\vec{x} = \vec{\varphi}(\mathbf{q})$. Et donc, en omettant $(\vec{x})$ pour alléger l'écriture :

$$\begin{aligned} d_{\rm ext}\alpha(\vec{u},\vec{v}) &= \sum_{i=1}^n d\alpha_i.\vec{u}\, dq^i.\vec{v} - \sum_{i=1}^n d\alpha_i.\vec{v}\, dq^i.\vec{u} \\ &= \sum_{i,j=1}^n \frac{\partial \alpha_i}{\partial q^j} u^j v^i - \frac{\partial \alpha_i}{\partial q^j} v^j u^i = \sum_{i,j=1}^n (\frac{\partial \alpha_i}{\partial q^j} - \frac{\partial \alpha_j}{\partial q^i}) u^j v^i \end{aligned}$$

Et donc dire que α est fermée signifie que, pour tout i, j :

$$\frac{\partial \alpha_i}{\partial q^j} = \frac{\partial \alpha_j}{\partial q^i} \tag{A.7}$$

dans Ω.

Remarque A.7 Dans $\mathbb{R}^3$, la relation $\frac{\partial \alpha_i}{\partial q_j} - \frac{\partial \alpha_j}{\partial q_i} = 0$ pour tout i, j s'écrit formellement $\vec{\rm rot}(\alpha) = 0$, et exprime "la conservation" : on "tournera" sans perdre d'information. Voir poly de première année "intégrales sur les courbes et surfaces" pour une interprétation du rotationnel.

Rappel : si $\vec{\rm rot}(\vec{v}) = 0$, alors "$\vec{v} = \vec{\nabla} f$", ce qui s'exprimera pour les formes différentielles : si $\mathrm{rot}(\alpha) = 0$, alors "$\alpha = df$", au moins localement, ce que montre le théorème de Poincaré, paragraphe suivant. Et donc $d_{\rm ext}\alpha = 0$: on ne perd pas d'information (on converve l'information). ▟

Remarque A.8 Une forme différentielle n'est pas fermée en général, avec $\vec{x} = (x, y)$: prendre
$\alpha(\vec{x}) = -y\, dx + x\, dy = \alpha_1(x, y)\, dx + \alpha_2(x, y)\, dy$.
Les équations (A.7) ne sont pas satisfaites. ▟

Remarque A.9 Bis. Une forme fermée permet d'avoir "la conservation" : la proposition 4.14 s'écrit dans ce cas $T(\vec{f},\Gamma) = \int_\Gamma \alpha$; et, dans le cas $\alpha = dg$, le travail vaut "$\int_\Gamma dg = g(B) - g(A)$" : il ne dépend que des extrémités du chemin Γ.

Ce n'est pas le cas avec $\alpha = -y\, dx + x\, dy$ dans $\mathbb{R}^2$. ▟

A.2.4 Forme différentielle exacte

Définition A.10 Une forme différentielle $\alpha \in T_1^0(\Omega)$ est dite exacte dans Ω ssi il existe une fonction f définie dans tout Ω et C^2 telle que $\alpha = df$:

$$\exists f \in C^2(\Omega;\mathbb{R}) \quad \text{t.q.} \quad \alpha = df.$$

Et dans ce cas f est appelée primitive de α (ou encore f est un potentiel).

Proposition A.11 *Si $f \in C^2(\Omega;\mathbb{R})$, alors $d_{\text{ext}}(df) = 0$. D'où, si α est exacte alors α est fermée.*

Preuve. Comme $f \in C^2$ on dispose du théorème de Schwartz $\frac{\partial \frac{\partial f}{\partial x_i}}{\partial x_j} = \frac{\partial \frac{\partial f}{\partial x_j}}{\partial x_i}$. Donc $d_{\text{ext}}(df) = 0$ avec (A.5).

Donc si $\alpha = df$ alors $d_{\text{ext}}\alpha = 0$. ∎

Remarque A.12 Une forme différentielle n'a pas de primitive en général : prendre
$\alpha(\vec{x}) = -y\,dx + x\,dy = \alpha_1\,dx + \alpha_2\,dy.$
qui n'est pas fermée, et qui donc ne peut être exacte. ∎

A.2.5 Théorème de Poincaré

Remarque A.13 Une forme différentielle peut être fermée ($d_{\text{ext}}\alpha = 0$) sans être exacte : dans $\Omega = \mathbb{R}^2 - \{\vec{0}\}$, prendre $\alpha(\vec{x}) = \frac{-y\,dx + x\,dy}{x^2+y^2}$. Ici $\alpha_1(x,y) = -\frac{y}{x^2+y^2}$ et $\frac{\partial \alpha_1}{\partial y}(x,y) = \frac{y^2-x^2}{x^2+y^2}$, et $\alpha_2(x,y) = \frac{x}{x^2+y^2}$ et $\frac{\partial \alpha_2}{\partial x}(x,y) = \frac{y^2-x^2}{x^2+y^2}$. Et localement $\alpha = df$ où $f(x,y) = \arctan(\frac{y}{x})$.

Par contre sur un cercle Γ centré en $\vec{0}$ de rayon R, on a $\int_\Gamma \alpha = \frac{1}{R^2}\int_\Gamma -y\,dx + x\,dy = 2\pi$, alors que si α était exacte, on aurait $\int_\Gamma df = [f]_{\theta=0}^{2\pi} = f(R,0) - f(R,0) = 0$.

Le théorème suivant dit quand même que localement, i.e. si $d_{\text{ext}}\alpha(\vec{x}) = 0$ pour tout $\vec{x}$ dans un voisinage d'un point $\vec{x}_0$ alors $\alpha(\vec{x}) = df(\vec{x})$ dans ce voisinage, i.e. α est exacte localement. ∎

Théorème A.14 *(de H. Poincaré) Si α est fermé dans Ω ouvert de $\mathbb{R}^n$, alors au voisinage de chaque point $\vec{x}$ il existe f telle que $\alpha = df$. I.e. toute forme différentielle fermée est localement exacte : il existe $U \subset \Omega$, il existe $f \in C^2(U;\mathbb{R})$, t.q. $\alpha(\vec{x}) = df(\vec{x})$ pour tout $\vec{x} \in U$.*

N.B. : ce théorème est souvent énonce comme : si Ω est un ouvert étoilé, et si α est fermée dans Ω, alors α est exacte dans Ω. La démonstration donnée le démontre.

Preuve. Soit $\vec{x} \in \Omega$. Soit $\vec{y} \in \Omega$ tel que le segment $[\vec{x},\vec{y}] = \{t\vec{x} + (1-t)\vec{y} : t \in [0,1]\}$ soit entièrement dans Ω. Posons :

$$f(\vec{y}) = \int_{t=0}^1 \alpha(t\vec{y}).\vec{y}\,dt.$$

Vérifions que $df(\vec{x}) = \alpha(\vec{x})$.

Posons :

$$A : \begin{cases} [0,1] \times \Omega & \to \mathbb{R} \\ \quad (t,\vec{y}) & \to A(t,\vec{y}) = \alpha(t\vec{y}).\vec{y}, \end{cases}$$

avec donc $f(\vec{y}) = \int_{t=0}^1 A(t,\vec{y})\,dt$.

À t fixé, posons $A_t(\vec{x}) =^{\text{déf}} A(t,\vec{x})$. Sa différentielle est donnée par :

$$dA_t(\vec{y}).\vec{v} = \alpha(t\vec{y}).\vec{v} + t\,(d\alpha(t\vec{y}).\vec{v}).\vec{y}.$$

Comme $d_{\text{ext}}\alpha = 0$ on a $d\alpha$ symétrique, et donc :

$$dA_t(\vec{y}).\vec{v} = \alpha(t\vec{y}).\vec{v} + t\,(d\alpha(t\vec{y}).\vec{y}).\vec{v},$$

et donc :

$$dA_t(\vec{y}).\vec{v} = (\alpha(t\vec{y}) + t\,d\alpha(t\vec{y}).\vec{y}).\vec{v} = b'(t).\vec{v} \qquad \text{où} \qquad b(t) = t\,\alpha(t\vec{y}).$$

Donc la fonction α étant au moins C^1 :

$$df(\vec{y}).\vec{v} = (\int_{t=0}^1 b'(t)\,dt).\vec{v} = (b(1) - b(0)).\vec{v} = \alpha(\vec{y}).\vec{v}.$$

Ce pour tout $\vec{v} \in \mathbb{R}^n$, donc $\alpha = df$. ∎

A.3 Produit extérieur et différentielle extérieure

On vient de voir que les formes différentielles exactes conservaient localement l'information : $\alpha = df$ implique que le travail $\oint_\Gamma \alpha$ le long d'un chemin fermé est nul.

On peut formaliser cette idée.

A.3.1 Produit extérieur

Définition A.15 Si $\ell, m \in \mathbb{R}^{n*}$ sont deux formes linéaires sur $\mathbb{R}^n$, alors le produit extérieur des deux formes linéaires est la forme bilinéaire alternée définie par :

$$\ell \wedge m = \ell \otimes m - m \otimes \ell,$$

i.e. donnée par, pour tout $\vec{u}, \vec{v} \in \mathbb{R}^n$:

$$(\ell \wedge m)(\vec{u}, \vec{v}) = \ell(\vec{u})m(\vec{v}) - m(\vec{u})\ell(\vec{v}) = \det \begin{pmatrix} \ell(\vec{u}) & m(\vec{u}) \\ \ell(\vec{v}) & m(\vec{v}) \end{pmatrix}.$$

Exemple A.16 Dans $\mathbb{R}^2$, $dx^1 \wedge dx^2$ est appelé le volume (où l'aire en 2-D). En effet, pour $\vec{u}, \vec{v} \in \mathbb{R}^2$, le réel $(dx^1 \wedge dx^2)(\vec{u}, \vec{v}) = u^1v^2 - u^2v^1 = \det(\vec{u}, \vec{v})$ donne l'aire du parallélogramme de côtés $\vec{u}$ et $\vec{v}$. ▟

Exemple A.17 Dans $\mathbb{R}^3$ muni de son produit scalaire canonique, $\omega = dx^2 \wedge dx^3$ est caractérisé par $\omega(\vec{u}, \vec{v}) = u^2v^3 - v^3u^2$ quand $\vec{u} = (u^1, u^2, u^3)$ et $\vec{v} = (v^1, v^2, v^3)$. Si on note $\vec{u}_{23} = (u^2, u^3) \in \mathbb{R}^2$ et $\vec{v}_{23} = (v^2, v^3) \in \mathbb{R}^2$ les projetés de $\vec{u}$ et $\vec{v}$ sur le plan yz, alors $\omega(\vec{u}, \vec{v})$ donne l'aire du parallélogramme limité par $\vec{u}_{23}$ et $\vec{v}_{23}$.

Et de manière générique dans $\mathbb{R}^n$ muni de son produit scalaire canonique, si $\omega = \ell \wedge m$, si pour un vecteur $\vec{u}$ on note $\vec{u}_{\ell m} = (\ell(\vec{u}), m(\vec{u})) \in \mathbb{R}^2$ son projeté sur le plan engendré par les vecteurs $\vec{n}_\ell$ et $\vec{n}_m$ unitaires orientés orthogonaux respectivement à $\mathrm{Ker}\ell$ et à $\mathrm{Ker}m$ (cf. théorème de représentation de Riesz), alors $\omega(\vec{u}, \vec{v})$ donne l'aire orienté du parallélogramme de côté $\vec{u}_{\ell m}$ et $\vec{v}_{\ell m}$. ▟

Remarque A.18 Dans $\mathbb{R}^3$, le produit $\wedge$ permet de définir l'orientation locale d'une surface, à l'aide de $d\omega = dq^1 \wedge dq^2$: ici on se donne une surface paramétrée $\vec{\varphi} : \mathbf{q} = (q^1, q^2) \to \vec{x} = \begin{pmatrix} x = \varphi^1(q^1, q^2) \\ y = \varphi^1(q^1, q^2) \\ z = \varphi^1(q^1, q^2) \end{pmatrix}$, on définit les vecteurs de base de ce système de coordonnées $\vec{e}_i(\vec{x}) = \frac{\partial \varphi^i}{\partial q^i}(\mathbf{q})$ au point $\vec{x} = \vec{\varphi}(\mathbf{q})$, et on définit la base duale $(dq^1(\vec{x}), dq^2(\vec{x}))$ aux points $\vec{x} = \vec{\varphi}(\mathbf{q})$. Et on a :

$$(dq^1 \wedge dq^2)(\vec{e}_1, \vec{e}_2) = 1, \qquad (dq^1 \wedge dq^2)(\vec{e}_2, \vec{e}_1) = -1.$$

▟

A.3.2 Expression dans un système de coordonnées

Si on se donne une base $(\vec{e}_i)_{i=1,\dots,n}$ de base duale $(dq^i)i = 1, ..., n$, une forme linéaire ℓ s'écrit :

$$\ell = \sum_{i=1}^{n} \ell_i \, dq^i \overset{\text{noté}}{=} \begin{pmatrix} \ell_1 & \dots & \ell_n \end{pmatrix}_{|dq},$$

la matrice ligne $[\ell] = \begin{pmatrix} \ell_1 & \dots & \ell_n \end{pmatrix}_{|dq}$ représentant ℓ dans la base (dq^i). Et avec $m = \sum_{i=1}^n m_i \, dq^i = \begin{pmatrix} m_1 & \dots & m_n \end{pmatrix}_{|dq}$, on a :

$$\ell \otimes m = \sum_{i,j=1}^{n} \ell_i m_j dq^i \otimes dq^j \overset{\text{noté}}{=} [\ell_i m_j]_{|dq} = \begin{pmatrix} \ell_1 m_1 & \dots & \ell_1 m_n \\ \vdots & & \vdots \\ \ell_n m_1 & \dots & \ell_n m_n \end{pmatrix}_{|dq}.$$

En effet, on rappelle que :

$$(\ell \otimes m).(\vec{u}, \vec{v}) = \ell(\vec{u})m(\vec{v}) = \sum_{i,j=1}^{n} \ell_i u^i m_j v^j = [\vec{u}]_{|\vec{e}}^T . [\ell_i m_j]_{|dq} . [\vec{v}]_{|\vec{e}}.$$

Donc la matrice de $\ell \wedge m$ dans la base (dq^i) est la matrice antisymétrique :

$$[\ell \wedge m]_{|dq} = \begin{pmatrix} 0 & \ell_1 m_2 - m_1 \ell_2 & ... & \ell_1 m_n - m_1 \ell_n \\ \ell_2 m_1 - m_2 \ell_1 & 0 & & \vdots \\ \vdots & & 0 & \ell_{n-1} m_n - m_{n-1} \ell_n \\ \ell_n m_1 - m_n \ell_1 & ... & \ell_n m_{n-1} - m_n \ell_{n-1} & 0 \end{pmatrix}_{|dq}$$

En particulier $dq^i \wedge dq^i = 0$ pour tout i et :

$$[dq^1 \wedge dq^2]_{|dq} = \begin{pmatrix} 0 & 1 & 0 & ... \\ -1 & 0 & ... & \\ 0 & ... & & \\ \vdots & & & \end{pmatrix}_{|dq},$$

de même pour les autres $dq^i \wedge dq^j$: tous les termes de la matrice sont nuls sauf le terme “ligne i colonne j” qui vaut 1 et le terme “ligne j colonne i” qui vaut -1.

Ainsi $(dq^i \wedge dq^j)_{\substack{i,j=1,\dots,n \\ i<j}}$ est une base de l'espace vectoriel des formes bilinéaires alternées. Et si $n = 2$, l'espace est de dimension 1.

A.3.3 Différentielle extérieure d'une forme différentielle

Proposition A.19 *Pour la forme différentielle donnée par (dans un système de coordonnées) :*

$$\alpha(\vec{x}) = \sum_i \alpha_i(\vec{x}) dq^i(\vec{x}) = (\, \alpha_1(\vec{x}) \quad ... \quad \alpha_n(\vec{x}) \,)_{|dq}, \tag{A.8}$$

on a :

$$d_{\rm ext}\alpha = \sum_i d\alpha_i \wedge dq^i = \sum_{ij} \frac{\partial \alpha_i}{\partial q^j} dq^j \wedge dq^i = \begin{pmatrix} 0 & \frac{\partial \alpha_2}{\partial q^1} - \frac{\partial \alpha_1}{\partial q^2} & ... \\ \frac{\partial \alpha_1}{\partial q^2} - \frac{\partial \alpha_2}{\partial q^1} & 0 & \\ \vdots & & \end{pmatrix}_{|dq}. \tag{A.9}$$

ou encore $d_{\rm ext}\alpha = \sum_{ij} \frac{\partial \alpha_j}{\partial q^i} dq^i \wedge dq^j$.

Preuve. Soit $\alpha = \sum_i \alpha_i dq^i$ où $\alpha_i : \Omega \to \mathbb{R}$. Avec (A.6) :

$$d\alpha.\vec{u} = \sum_i d\alpha_i.\vec{u}\, dq^i,$$

et donc avec (A.5) :

$$\begin{aligned} d_{\rm ext}\alpha(\vec{u},\vec{v}) &= \sum_i d\alpha_i.\vec{u}\,(dq^i.\vec{v}) - \sum_i d\alpha_i.\vec{v}\,(dq^i.\vec{u}) \\ &= \sum_i (d\alpha_i \otimes dq^i)(\vec{u},\vec{v}) - \sum_i (d\alpha_i \otimes dq^i)(\vec{v},\vec{u}) \\ &= \sum_i (d\alpha_i \wedge dq^i)(\vec{u},\vec{v}). \end{aligned}$$

Puis comme $d\alpha_i = \sum_j \frac{\partial \alpha_i}{\partial x^j} dq^j$ (différentielle de la fonction α_i) :

$$d_{\rm ext}\alpha(\vec{u},\vec{v}) = \sum_{ij} (\frac{\partial \alpha_i}{\partial x^j} dq^j \wedge dq^i).(\vec{u},\vec{v})$$

▪

Références

[1] Abraham R., Marsden J.E. : *Foundation of mechanics, 2nd edition*. Addison-Wesley, 1985.

[2] Arnold V.I. : *Mathematical Methods of Classical Mechanics*. Second edition, Springer–Verlag, 1989.

[3] Antaya Hélène : *La cycloïde : du brachistochrone au pendule d'Huyghens*. Bulletin AMQ, vol. XLIV, N°3, octobre 2004, http://newton.mat.ulaval.ca/amq/archives/2004/3/2004-3-part4.pdf

[4] Cartan H. : *Cours de calcul différentiel*. Hermann 1977.

[5] Fitzpatrick Richard : http://farside.ph.utexas.edu/teaching/336k/lectures.pdf.

[6] Marsden J.E., Hughes T.J.R. : *Mathematical Foundations of Elasticity*. Dover publications, 1993.

[7] Matthys Danielle : *Formules de variation première et seconde*. Notes, 2006.

[8] http://fr.wikipedia.org/wiki/G%C3%A9om%C3%A9trie_symplectique

Et beaucoup de livres de mécanique.

Pour les illustrations de la brachistochrone, voir par exemple
http://serge.mehl.free.fr/anx/cv_equ_Euler.html
http://www.mathcurve.com/courbes2d/brachistochrone/brachistochrone.shtml