Notes de cours de l'ISIMA, première année, http://www.isima.fr/~leborgne

Transformée de Legendre : $f(\text{point}) \to f^*(\text{pente})$, et $f^{**} = f$. Introduction, et exemples : thermodynamique (énergie libre), et énergie (hamiltonien)

Gilles Leborgne

3 août 2022

But : changement de variable point x vers pente $p = f'$, donnant le changement de fonction f vers sa transformée de Legendre f^*, ce changement réitéré redonnant $f = f^{**}$.

Table des matières

Notation : « $g := f$ » signifie « g est défini par : g est égal à f ».

1 Transformée de Legendre 1-D

1.1 Introduction formelle

1.1.1 $f^*(p) = px - f(x)$...

Si $f : \left\{\begin{matrix} \mathbb{R} & \to \mathbb{R} \\ x & \to f(x) \end{matrix}\right\}$ alors sa transformée de Legendre $f^* : \left\{\begin{matrix} \mathbb{R} & \to \mathbb{R} \\ p & \to f^*(p) \end{matrix}\right\}$ est définie par

$$f^*(p) := px - f(x) \quad \text{pour } x \text{ t.q. la pente de } f \text{ en } x \text{ est } p. \tag{1.1}$$

1.1.2 ... donc $f^{**} = f$

Donc $f^{**} := (f^*)^* : \left\{\begin{matrix} \mathbb{R} & \to \mathbb{R} \\ x & \to f^{**}(x) \end{matrix}\right\}$ vérifie, pour p t.q. la pente de f^* en p est x,

$$f^{**}(x) = xp - f^*(p) = f(x), \quad \text{donc} \quad f^{**} = f. \tag{1.2}$$

Donc si on connaît f^* alors on reconstruit f.

1.2 Définition de la transformée de Legendre

1.2.1 Développement limité de f et la définition

Soit I un intervalle ouvert de $\mathbb{R}$ (non vide), soit $f \in C^1(I;\mathbb{R})$ et soit $x \in I$. Au voisinage de x, le développement limité de f au premier ordre est

$$\begin{aligned} f(x) &= f(x_0) + p(x-x_0) + o(x-x_0), \quad \text{où} \quad p = f'(x_0) = \text{pente en } x_0, \\ &= \underbrace{px}_{\text{droite linéaire}} + \underbrace{y_p}_{-f^*(p):=px_0-f(x_0)} + o(x-x_0), \end{aligned} \tag{1.3}$$

où donc $y_p = f(x_0) - x_0 p$. Voir figures 1.1-1.2.

Définition 1.1 Si $p \in \mathrm{Im}(f')$ (si p est une pente de f) alors la transformée de Legendre f^* de f est définie avec (1.3) par $f^*(p) = -y_p$, i.e.

$$f^*(p) := px_0 - f(x_0) \quad \text{quand} \quad x_0 = (f')^{-1}(p), \tag{1.4}$$

i.e. quand x_0 est le point où la pente de f vaut p.

Et avec des notations génériques :

$$\boxed{f^*(p) := px - f(x) \quad \text{quand} \quad x = (f')^{-1}(p)} \quad \text{(i.e. quand } x \text{ vérifie } f'(x) = p). \tag{1.5}$$

(Le signe $-$ permet d'avoir $f(x) = xp - f^*(p)$.)

Autrement dit, la transformée de Legendre de f est la fonction f^* définie sur $\mathrm{Im}(f')$ par

$$f^* := Id.(f')^{-1} - f \circ (f')^{-1}, \tag{1.6}$$

i.e., pour tout $p \in \mathrm{Im}(f')$,

$$f^*(p) := p\,(f')^{-1}(p) - f((f')^{-1}(p)), \ \forall p \in \mathrm{Im}(f'), \tag{1.7}$$

i.e., pour tout $x \in I$,

$$f^*(f'(x)) := f'(x)x - f(x). \tag{1.8}$$

Donc :

$$\underline{\text{changement de variable}}\ \boxed{\text{point } x \to \text{pente } p = f'(x)} \xrightarrow{\text{Legendre}} \underline{\text{changement de fonction}}\ \boxed{f(x) \to f^*(p)}.$$

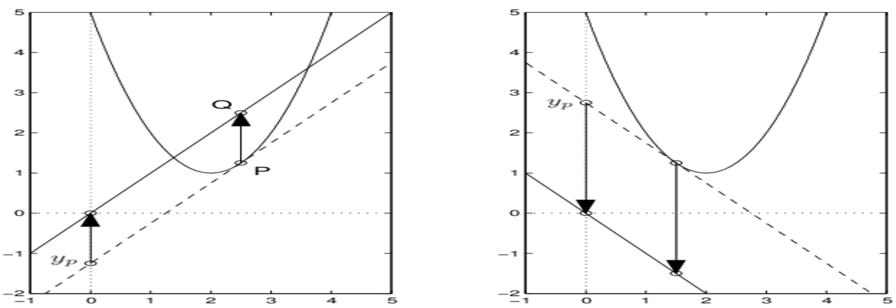

FIGURE 1.1 – Pour $p \in \mathrm{Im}(f')$ (pente donnée), $P = (x_p, f(x_p))$ est le point du graphe de f d'abscisse x_p tel que $f'(x_p) = p$. L'équation de la tangente en P est $y(x) = y_p + px$ (trait pointillé), et$f^*(p) := -y_p = px_p - f(x_p)$. Et $Q = (x_p, px_p)$ est sur la droite linéaire $x \to px$ (trait plein). La flèche $= \vec{PQ} = Q - P = (0, -y_p) = (0, f^*(p))$ donne “la plus grande distance possible entre P et Q”, cf. (1.22).

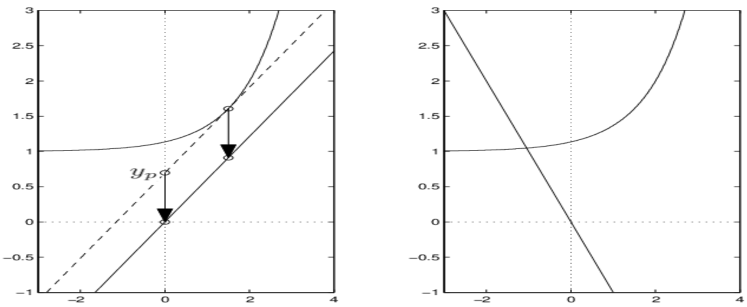

FIGURE 1.2 – Représentation de $f^*(p) = -y_p$, pour une fonction exponentielle. Ici $\mathrm{Im}(f') = \mathbb{R}_+^*$: la figure de droite avec la pente $p < 0$ n'est pas dans le domaine de définition $\mathrm{Im}(f')$ de f^* ; voir exercice 1.7.

Proposition 1.2 *Soit $f \in C^1(I;\mathbb{R})$ strictement convexe (resp. strictement concave). Si $p \in \mathrm{Im}(f')$, i.e. s'il existe $x_p \in I$ t.q. $p = f'(x_p)$ alors x_p est unique, à savoir $x_p = (f')^{-1}(p)$. Donc la transformée de Legendre* $f^* : \left\{ \begin{matrix} \mathrm{Im}(f') & \to \mathbb{R} \\ p & \to f^*(p) \end{matrix} \right\}$ *de f donnée en (1.5) est bien définie.*

Si de plus $f \in C^2(I;\mathbb{R})$ alors f' et $(f')^{-1}$ sont C^1.

Preuve. On a $f' \in C^0(I;\mathbb{R})$ strictement croissante (resp. strictement décroissante), donc bijective $I \to \mathrm{Im}(f')$ avec $\mathrm{Im}(f')$ ouvert, et l'inverse de f' est

$$(f')^{-1} : \begin{cases} \mathrm{Im}(f') & \to I \\ \quad p & \to (f')^{-1}(p) = x_p = \text{le point où la pente vaut } p. \end{cases} \tag{1.9}$$

Et le lemme A.12 donne f' et $(f')^{-1} \in C^{k-1}$ si $f \in C^k$, pour $k = 1, 2$. ∎

1.2.2 $f = f^{**}$

Proposition 1.3 *Quand $f \in C^2(I;\mathbb{R})$ est strictement convexe (resp. concave), $f^* \in C^1(\mathrm{Im}(f');\mathbb{R})$ et*

$$(f^*)' \circ f' = I, \quad \text{i.e.} \quad \boxed{(f^*)' = (f')^{-1}}. \tag{1.10}$$

Donc

$$p = f'(x) \iff x = (f^*)'(p), \quad \text{i.e.} \quad \boxed{p = \frac{df}{dx}(x) \iff x = \frac{df^*}{dp}(p)}. \tag{1.11}$$

Et f^* est strictement convexe (resp. concave), et l'opérateur $^* : f \to f^*$ est involutif :

$$\boxed{f^{**} = f}, \tag{1.12}$$

où $f^{**} := (f^*)^*$, i.e. $f^{**}(x) = xp - f^*(p) = f(x)$ quand $p = ((f^*)')^{-1}(x)$.

Preuve. Soit f une fonction C^2 convexe stricte (si f est concave stricte considérer $-f$). Donc $f^* \in C^1(\mathrm{Im}(f');\mathbb{R})$ est bien définie (prop. 1.2 et (1.6)), et (1.8) donne $(f^*)'(f'(x))f''(x) = f''(x)x + f'(x) - f'(x) = f''(x)x$, donc $(f^*)'(f'(x)) = x$ (car $f''(x) \neq 0$), d'où (1.10).

Donc $p = f'(x)$ ssi $x = (f')^{-1}(p) =^{(1.10)} (f^*)'(p)$, i.e. (1.11).

Et $(f^*)'(f'(x)) = x$ donne $(f^*)''(f'(x))f''(x) = 1$, donc $(f^*)''(p) = \frac{1}{f''(x)} > 0$ quand $p = f'(x)$, donc $(f^*)''$ est bien définie et convexe stricte. Et $f^{**}(q) := qp - f^*(p)$ quand $q = (f^*)'(p) = (f^*)'(f'(x)) =^{(1.10)} (f')^{-1}(f'(x)) = x$, donc $f^{**}(x) = xp - f^*(p) =^{(1.5)} f(x)$, donc $f^{**} = f$. ▪

Exercice 1.4 Vérifier que $(-f)^*(p) = -f^*(-p)$, et redonne $(-f)^{**} = -f$.

Réponse. Soit $p \in \mathrm{Im}(f')$ et soit x t.q. $p = (-f)'(x) = -f'(x)$. Alors $(-f)^*(p) = px - (-f)(x) = px + f(x) = -((-p)x - f(x)) = -f^*(-p)$ comme demandé.

D'où $(-f)^{**}(x) = ((-f)^*)^*(x) = xp - (-f)^*(x) = xp - (-f^*(-p)) = xp + f^*(-p) = xp + (-px - f(x)) = -f(x) = (-f)(x)$: relation vérifiée. ▪

1.2.3 Exemples

Exemple 1.5 $f(x) = \frac{1}{2}mx^2$ donne $f^*(p) = \frac{1}{2}\frac{p^2}{m}$, car $p = f'(x) = mx$, donc $x = \frac{p}{m}$, donc $f^*(p) = \frac{p^2}{m} - \frac{p^2}{2m}$. En particulier $f : x \to f(x) = \frac{x^2}{2}$ est "conservée" par Legendre : $f^*(p) = \frac{p^2}{2}$. ▪

Exercice 1.6 Plus généralement (problèmes élémentaires de recherche de minimum pour une parabole) : soit $\alpha > 0$ et $x_0, y_0 \in \mathbb{R}$, et soit f la fonction parabolique donnée par, pour $x \in \mathbb{R}$:

$$f(x) = y_0 + \frac{\alpha}{2}(x - x_0)^2, \tag{1.13}$$

Montrer que f^* est parabolique (déterminer f^*). Vérifier que $(f^*)^* = f$.

Réponse. Soit $p \in \mathbb{R}$. Ayant $f'(x) = \alpha(x - x_0)$, le point x_p t.q. $f'(x_p) = p$ vaut

$$x_p = x_0 + \frac{p}{\alpha}. \tag{1.14}$$

Donc $f^*(p) = px_p - f(x_p) = p\,(x_0 + \frac{p}{\alpha}) - (y_0 + \frac{\alpha}{2}(\frac{p}{\alpha})^2) = \frac{1}{2\alpha}\,p^2 + x_0\,p - y_0 = \frac{1}{2\alpha}(p + \alpha x_0)^2 - \frac{\alpha x_0^2}{2} - y_0$, donc :

$$f^*(p) = z_0 + \frac{\beta}{2}(p - p_0)^2, \quad \text{où} \quad z_0 = -y_0 - \frac{\alpha x_0^2}{2}, \quad \beta = \frac{1}{\alpha}, \quad p_0 = -\alpha x_0, \tag{1.15}$$

et f^* est bien parabolique. Et

$$(f^*)^*(t) = \tilde{y}_0 + \frac{\gamma}{2}(t - t_0)^2, \quad \text{où} \quad \tilde{y}_0 = (-z_0 - \frac{\beta p_0^2}{2}) = y_0, \quad \gamma = \frac{1}{\beta} = \alpha, \quad t_0 = -\beta p_0 = x_0, \tag{1.16}$$

donc $(f^*)^*(t) = y_0 + \frac{\alpha}{2}(t - x_0)^2$ pour tout t, d'où $f^*(x) = f(x)$ pour tout x. ▪

Exercice 1.7 Soit $\alpha > 0$ et $x_0, y_0 \in \mathbb{R}$ et la fonction de type exponentielle donnée sur $\mathbb{R}$ par, pour $x \in \mathbb{R}$:

$$f(x) = y_0 + \alpha\, e^{x - x_0}. \tag{1.17}$$

Déterminer f^*. Voir figure 1.2. Pour $f(x) = e^x$, montrer $f^*(p) = p\log(p) - p$. Vérifier que $(f^*)^* = f$.

Réponse. Soit $p \in \mathrm{Im}(f')$. Ayant $f'(x) = \alpha\, e^{x - x_0}$, le point x_p t.q. $f'(x_p) = p$ vaut

$$x_p = x_0 + \log(\frac{p}{\alpha}) = x_0 + \log(p) - \log(\alpha). \tag{1.18}$$

Donc

$$\begin{aligned} f^*(p) &= p(x_0 + \log(p) - \log(\alpha)) - f(x_0 + \log(\frac{p}{\alpha})) = p(x_0 - \log(\alpha)) + p\log(p) - (y_0 + \alpha\frac{p}{\alpha}) \\ &= -y_0 + p(x_0 - \log(\alpha) - 1 + \log(p)) = -y_0 + p(x_0 - 1 + \log(\frac{p}{\alpha})). \end{aligned} \tag{1.19}$$

Cas particulier : $y_0 = 0 = x_0$ et $\alpha = 1$. On a $f(x) = e^x$ et $f^*(p) = p\log(p) - p$. Et, pour $p \in \mathbb{R}_+^*$,

$$(f^*)'(p) = x_0 - \log(\alpha) - 1 + \log(p) + p\frac{1}{p} = x_0 - \log(\alpha) + \log(p) = x_0 + \log(\frac{p}{\alpha}) \quad (= x_p). \tag{1.20}$$

Donc, pour $q \in \mathbb{R}$, le point p_q t.q. $(f^*)'(p_q) = q$ vaut

$$p_q = \alpha e^{q - x_0}. \tag{1.21}$$

Donc $(f^*)^*(q) = q\alpha e^{q-x_0} - f^*(\alpha e^{q-x_0}) = \alpha q e^{q-x_0} - (-y_0 + \alpha e^{q-x_0}(x_0 - 1 + q - x_0)) = y_0 + \alpha e^{q-x_0} = f(q)$. ▪

1.2.4 Définition avec le sup

Définition 1.8 La transformée de Legendre de $f : \mathbb{R} \to \mathbb{R}$ convexe est la fonction f^* définie par

$$\boxed{f^*(p) = \sup_{x\in I}\big(px - f(x)\big)} \tag{1.22}$$

Voir figures 1.1 et 1.2. En particulier si le sup est atteint pour $x = x_p$ alors $f^*(p) = px_p - f(x_p)$.

Autrement dit, $\varphi(x,p) = px - f(x)$ alors $f^*(p) = \sup_{x\in I} \varphi(x,p)$, et si le sup est atteint pour $x = x_p$ alors $\frac{\partial\varphi}{\partial x}(x_p,p) = p - f'(x_p) = 0$, i.e. x_p vérifie $f'(x_p) = p$.

Exercice 1.9 Cette définition, cf. (1.22), est un peu plus générale de la définition donnée par (1.5) : Soit $\alpha, \beta \in \mathbb{R}$ et soit $f(x) = \alpha x + \beta$ (affine) pour $x \in \mathbb{R}$. Déterminer f^* avec (4.4). Même question pour $x \in [a, b[$

Réponse. Ici f n'est pas strictement convexe. $f^*(p) = \sup_{x\in\mathbb{R}} px - \alpha x - \beta$, donc $f^*(\alpha) = -\beta$ et $f^*(p) = \infty$ pour $p \neq \alpha$, donc $f^*(p)$ n'est définie que pour $p = \alpha$ (et $\mathrm{Im}(f') = \{\alpha\}$).

Pour $x \in]a, b[$ on a $f^*(p) = \sup_{x\in]a,b[} px - \alpha x - \beta$. Cas $a > 0$: si $p \in [\alpha, \infty[$ alors $f^*(p) = bp - \alpha b - \beta$ (affine) et si $p \in] - \infty, \alpha]$ alors $f^*(p) = ap - \alpha a - \beta$ (affine). Dessin. Autres cas : exercice. ▞

Proposition 1.10 *Soit $f \in C^2(I;\mathbb{R})$ une fonction strictement convexe. (1.22) donne, pour tout $(x,p) \in I \times \mathrm{Im}(f')$,*

$$\boxed{px \le f(x) + f^*(p)} \quad \textit{(Inégalité d'Young).} \tag{1.23}$$

Et on retrouve (1.5) : il existe $x_p \in I$ unique, à savoir donné par $f'(x_p) = p$, t.q. $\sup_{x\in I} \varphi(x,p) = \varphi(x_p,p)$ $(= \max_{x\in I} \varphi(x,p))$ et

$$f^*(p) = px_p - f(x_p) = px - f(x) \quad \textit{quand} \quad f'(x) = p. \tag{1.24}$$

Preuve. Soit $p \in \mathrm{Im}(f')$. La fonction $\varphi_p : x \to \varphi_p(x) := \varphi(x,p) = px - f(x)$ est $C^2(\Omega;\mathbb{R})$ et strictement concave, car ${\varphi_p}'(x) = p - f'(x)$ et ${\varphi_p}''(x) = -f''(x) < 0$. Soit $x_p \in I$ t.q. $p = f'(x_p)$. Avec ${\varphi_p}'(x_p) = 0$ (car $f'(\vec{x}_p) = p$), donc φ_p est maximum au point x_p, donc $\varphi_p(x) \le \varphi_p(x_p)$, pour tout $x \in I$, d'où (1.23), avec $\varphi_p(x_p) = px_p - f(x_p) = f^*(p)$ comme défini en (1.5). ▞

Exercice 1.11 Pour $f(x) = \frac{x^2}{2}$ on a $f^*(p) = \frac{p^2}{2}$, et (4.4) redonne l'inégalité remarquable $xp \le \frac{1}{2}(x^2+p^2)$ (due à $(x-p)^2 \ge 0$). Plus généralement, pour $a > 1$ et $f(x) = \frac{x^a}{a}$, montrer : $f^*(p) = \frac{p^b}{b}$ pour b t.q. $\frac{1}{a} + \frac{1}{b} = 1$ et (4.4) redonne $xp \le \frac{x^a}{a} + \frac{p^b}{b}$ (inégalité d'Young).

Réponse. $f(x) = \frac{x^a}{a}$; f est convexe stricte, et $f^*(p) = px - f(x)$ pour x t.q. $f'(x) = p$, i.e. $x^{a-1} = p$, i.e. $x = p^{\frac{1}{a-1}}$ donc $f^*(p) = p^{\frac{1}{a-1}+1} - \frac{1}{a}(p^{\frac{1}{a-1}})^a = (1-\frac{1}{a})p^{\frac{a}{a-1}}$. On pose $\frac{1}{b} = 1 - \frac{1}{a} = \frac{a-1}{a}$, donc $f^*(p) = \frac{1}{b}p^b$. Puis $\varphi(x,p) = px - f(x) \le \sup_{x'\in\mathbb{R}} \varphi(x',p) = f^*(p)$ donne $px - \frac{x^a}{a} \le \frac{p^b}{b}$. ▞

Exercice 1.12 Soit $f(x) = \frac{\alpha}{2}x^2$ définie sur $[c,d]$ où $c, d \in \mathbb{R}$ et $c < d$ et $\alpha > 0$. Déterminer f^*.

Réponse. La fonction f est strictement convexe sur $]c,d[$ car $\alpha > 0$.

Soit $p \in \mathbb{R}$ et g_p la fonction définie par $g_p(x) = px - f(x)$ sur $[c,d]$. La fonction est continue sur le compact $[c,d]$, donc atteint son maximum sur $[c,d]$: donc $f^*(p) = \sup_{x\in[c,d]} g(x)$ est définie pour tout $p \in \mathbb{R}$. On a $g_p'(x) = p - f'(x) = p - \alpha x$.

1- Si $x = \frac{p}{\alpha} \in]c,d[$, c.à.d. si $p \in \mathrm{Im}(f') =]\alpha c, \alpha d[$, alors g_p atteint son max en $x_p = \frac{p}{\alpha}$.

Donc $f^*(p) = g_p(\frac{p}{\alpha}) = p\frac{p}{\alpha} - f(\frac{p}{\alpha}) = \frac{p^2}{\alpha} - \frac{\alpha}{2}(\frac{p}{\alpha})^2 = \frac{p^2}{2\alpha}$ quand $p \in [\alpha c, \alpha d]$, et f^* est une fonction quadratique.

2- Si $x = \frac{p}{\alpha} > d$: on est dans le cas $p > \alpha d$. Donc $g_p'(x) \ge \alpha d - \alpha x = \alpha(d-x)$. Donc $g_p'(x) \ge 0$ pour $x \in [c,d]$. Donc g_p est croissante sur $[c,d]$ et atteint son max en $x = d$. Donc $f^*(p) = g_p(d) = pd - f(d) = pd - \frac{\alpha}{2}d^2$, et f^* est une fonction affine.

3- Si $x = \frac{p}{\alpha} < c$: on est dans le cas $p < \alpha c$. Donc $g_p'(x) \le \alpha c - \alpha x = \alpha(c-x)$. Donc $g_p'(x) \le 0$ pour $x \in [c,d]$. Donc g_p est décroissante sur $[c,d]$ et atteint son max en $x = c$. Donc $f^*(p) = g_p(c) = pc - f(c) = pc - \frac{\alpha}{2}c^2$, et f^* est une fonction affine.

Résumé : f^* est définie sur $I^* = \mathbb{R}$ par $f^*(p) = \begin{cases} cp - \dfrac{\alpha}{2}c^2 & \text{si} \quad p < \alpha c \quad \text{(affine)}, \\ \dfrac{1}{2\alpha}\,p^2 & \text{si} \quad p \in [\alpha c, \alpha d] \quad \text{(quadratique)}, \\ dp - \dfrac{\alpha}{2}d^2 & \text{si} \quad p > \alpha d \quad \text{(affine)}. \end{cases}$ ▞

2 Application à la thermodynamique

2.1 Energie interne pour une transformation réversible

T = température, P = pression, V = volume, S = entropie. Pour une transformation réversible, la quantité de chaleur élémentaire vaut $\delta Q = T\,dS$, le travail élémentaire vaut $\delta W = -P\,dV$, et l'énergie interne vaut $dU = \delta Q + \delta W$, i.e.

$$dU = T\,dS - P\,dV, \quad \text{au sens} \quad dU(S,V) = T(S,V)\,dS - P(S,V)\,dV, \tag{2.1}$$

les variables indépendantes étant S et V, où donc U est C^1 et

$$T = \frac{\partial U}{\partial S} \text{ et } P = \frac{\partial U}{\partial V}, \quad \text{au sens} \quad T(S,V) = \frac{\partial U}{\partial S}(S,V) \text{ et } P = P(S,V) = \frac{\partial U}{\partial V}(S,V). \tag{2.2}$$

2.2 L'énergie libre F associée (Helmholtz free energy) = cas isochore

Cas d'une transformation réversible à volume constant $V = V_0$ (transformation isochore). On note

$$U_{|V_0}(S) := U(S,V_0), \quad \text{donc} \quad dU_{|V_0}(S) = T_{|V_0}(S)\,dS \quad \text{où} \quad \boxed{T_{|V_0}(S) = \frac{dU_{|V_0}}{dS}(S)} = \text{la pente}. \tag{2.3}$$

(On a supposé $U_{|V_0} \in C^1$.) La température étant facile à mesurer (pas l'entropie) on utilise la transformée de Legendre : supposant $U_{|V_0}$ (localement) strictement convexe, changement de variable $S \to T$, et changement de fonction $U_{|V_0}(S) \to {U_{|V_0}}^*(T)$:

$${U_{|V_0}}^*(T) = TS(T) - U_{|V_0}(S(T)) \quad \text{quand} \quad S(T) = ({U_{|V_0}}')^{-1}(T) = S_{|V_0}(T), \tag{2.4}$$

noté

$${U_{|V_0}}^*(T) = TS - U_{|V_0}(S) \quad \text{quand} \quad S = S_{|V_0} \text{ est t.q. } {U_{|V_0}}'(S) = T. \tag{2.5}$$

Définition 2.1 L'énergie libre à volume constant V_0 est la fonction $F_{|V_0} := -{U_{|V_0}}^*$, donc

$$\boxed{F_{|V_0}(T) := U_{|V_0}(S(T)) - TS(T)} \quad \text{quand} \quad S(T) := ({U_{|V_0}}')^{-1}(T) = S_{|V_0}(T), \tag{2.6}$$

noté

$$F_{|V_0}(T) = U_{|V_0}(S) - TS \quad \text{quand} \quad S \text{ est t.q. } {U_{|V_0}}'(S) = T. \tag{2.7}$$

Et on note

$$\boxed{F = U - TS}, \tag{2.8}$$

au sens $F(T,V) = U(S(T,V),V) - TS(T,V)$ quand $S(T,V) := ({U_{|V}}')^{-1}(T)$, i.e. quand S vérifie $\frac{\partial U}{\partial S}(S,V) = T$.

Récupération de U : on a $F^* = (-U)^*$, donc $F^{**} = -U$: à volume constant.

2.3 Dérivée de F à volume constant

(2.6) donne

$${F_{|V_0}}'(T) = \underbrace{{U_{|V_0}}'(S(T))}_{=T} S'(T) - S(T) - TS'(T) = -S(T), \tag{2.9}$$

donc

$$dF_{|V_0}(T) = -S(T)\,dT \quad \text{noté} \quad \boxed{dF_{|V_0} = -S\,dT}, \tag{2.10}$$

noté $dF = -S\,dT$ (quand $\frac{\partial F}{\partial V} = 0$, i.e. à volume constant).

Remarque 2.2 Approche classique : $F = U - TS$ donne

$$\begin{aligned} dF &= dU - S\,dT - T\,dS \\ &= T\,dS - S\,dT - T\,dS = -S\,dT, \end{aligned} \tag{2.11}$$

où donc $\frac{dF}{dT} = -S$, i.e. $\frac{dF}{dT}(T) = -S(T)$, i.e. $F'(T) = -S(T)$: quand V est constant.

NB : le sens de $(2.11)_1$ est assez obscure car il semble que T et S sont les variables... ce qui n'est pas le cas, cf. (2.6)... Mais c'est une manière d'écrire (2.9)... ∴

3 Application à l'énergie cinétique en 1-D

Approche similaire à celle de l'énergie libre... mais ici on a en plus besoin d'un produit scalaire. En 1-D le produit scalaire entre deux réels x et y est $x \bullet y = cxy$ où $c > 0$ est une constante. On prend $c = 1$ pour simplifier (le 1 représente 1 pied ou 1 mètre par exemple).

3.1 Energie cinétique

Pour $m \in \mathbb{R}_+^*$ ("la masse"), l'énergie cinétique est la fonction parabolique

$$E_c : \begin{cases} \mathbb{R} & \to \mathbb{R} \\ v & \to E_c(v) = \frac{1}{2}mv^2. \end{cases} \tag{3.1}$$

(v est "la vitesse".)

3.2 Quantité de mouvement et transformée de Legendre

On a

$$E_c{}'(v) = mv \overset{\text{noté}}{=} p = \text{appelé la quantité de mouvement}, \tag{3.2}$$

Donc $E_c{}^*(p) \overset{(1.5)}{=} pv - E_c(v) = pv - \dfrac{1}{2}mv^2$ quand $v = \frac{p}{m}$, donc

$$E_c{}^*(p) = \frac{1}{2}\frac{p^2}{m}. \tag{3.3}$$

3.3 Potentiel, lagrangien, hamiltonien

Soit $E_{pot} : x \in \mathbb{R} \to E_{pot}(x) \in \mathbb{R}$ une fonction C^1 ("potentiel" : on ne s'intéresse qu'aux systèmes conservatifs). Le lagrangien est la fonction

$$\boxed{L(x,v) = E_c(v) - E_{pot}(x)} = \frac{1}{2}mv^2 - E_{pot}(x). \tag{3.4}$$

(x = "position", v "vitesse".)

Remarque 3.1 Minimiser le lagrangien pour une particule, soumis à un potentiel, sur l'ensemble des trajectoires reliant deux positions permet d'obtenir la trajectoire de la particule entre ces deux positions ; et l'énergie totale est conservée le long de cette trajectoire. ∴

À x fixé on pose $E_{pot,x} := E_{pot}(x)$ et $L_x(v) := L(x,v)$, donc

$$L_x(v) = \frac{1}{2}mv^2 - E_{pot,x}. \tag{3.5}$$

Autrement dit, on considère l'énergie cinétique à une constante près.

Ayant $L_x'(v) = mv = E_c{}'(v) = p$, la transformée de Legendre $L_x{}^*$ de L_x est donnée par $L_x{}^*(p) = pv - L_x(v) = pv - E_c(v) + E_{pot,x} = E_c{}^*(p) + E_{pot,x}$ quand $v = \frac{p}{m}$, donc

$$L_x{}^*(p) = \frac{1}{2}\frac{p^2}{m} + E_{pot,x} \overset{\text{noté}}{=} H_x(p). \tag{3.6}$$

Ce qui défini l'hamiltonien $H : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$:

$$\boxed{H(x,p) = \frac{1}{2}\frac{p^2}{m} + E_{pot}(x)} = E_c{}^*(p) + E_{pot}(x). \tag{3.7}$$

"Interprétation" : $p = mv$ donne $H(x,p) = \frac{1}{2}mv^2 + E_{pot}(x)$ = énergie totale... mais ici ce n'est pas v (vitesse) la variable mais p (quantité de mouvement = la pente de E_c) :

$$E(x,v) = E_c(v) + E_p(x) \quad \underset{v \to p = mv}{\overset{\text{ch. var.}}{\longrightarrow}} \quad H(x,p) = E_c{}^*(p) + E_{pot}(x). \tag{3.8}$$

4 Transformée de Legendre n-D

4.1 Forme linéaire, différentielle, convexité

Soit $\mathbb{R}_n^* := \mathcal{L}(\mathbb{R}^n;\mathbb{R})$ l'ensemble des formes linéaires (applications linéaires à valeurs scalaires) sur $\mathbb{R}^n$. Pour une forme linéaire $\ell \in \mathbb{R}_n^*$, ses valeurs $\ell(\vec{w}) \in \mathbb{R}$, pour $\vec{w} \in \mathbb{R}^n$, sont notées

$$\ell(\vec{w}) \overset{\text{noté}}{=} \ell.\vec{w}, \tag{4.1}$$

notation usuelle de la "distributivité" $\sim$ linéarité : $\ell(\vec{w}_1+\lambda\vec{w}_2) = \ell(\vec{w}_1)+\lambda\ell(\vec{w}_2)$ est noté $\ell.(\vec{w}_1+\lambda\vec{w}_2) = \ell.\vec{w}_1+\lambda\ell.\vec{w}_2$.

Soit une fonction $f : \vec{x} \in \mathbb{R}^n \to f(\vec{x}) \in \mathbb{R}$. Sa différentielle en un point $\vec{x}$ est la forme linéaire $df(\vec{x}) \in \mathbb{R}_n^* = \mathcal{L}(\mathbb{R}^n;\mathbb{R})$ (application linéaire à valeurs réelles) définie par, quand elle existe : pour tout $\vec{w} \in \mathbb{R}^n$,

$$df(\vec{x}).\vec{w} := \lim_{h\to 0} \frac{f(\vec{x}+h\vec{w})-f(\vec{x})}{h} \tag{4.2}$$

(dérivée dans la direction $\vec{w}$ quand la limite existe).

(Autrement dit, $df(\vec{x}).\vec{w} = g'_{\vec{x},\vec{w}}(0)$ où $g_{\vec{x},\vec{w}}(h) := f(\vec{x}+h\vec{w})$.)

Et alors le graphe de f a pour plan tangent en $\vec{x}$ le plan passant par $(\vec{x}, f(\vec{x}))$ qui est parallèle au plan vectoriel $\mathrm{Ker}(df(\vec{x})) = \{\vec{w} \in \mathbb{R}^n : df(\vec{x}).\vec{w} = 0\}$ = le noyau de la forme linéaire $df(\vec{x})$.

Et f est C^1 ssi $df : \vec{x} \in \mathbb{R} \to df(\vec{x}) \in \mathbb{R}_n^*$ est C^0. Et on poursuit la construction des fonctions C^k pour $k \geq 2$.

On se limite aux f C^2 ; le développement limité de f au second ordre est $f(\vec{x}+h\vec{w}) = f(\vec{x})+h\,df(\vec{x}).\vec{w}+\frac{h^2}{2}\,d^2f(\vec{x})(\vec{w},\vec{w})+o(h^2)$, où $d^2f := d(df)$ (soit $g_{\vec{x},\vec{w}}(h) = g_{\vec{x},\vec{w}}(0)+hg'_{\vec{x},\vec{w}}(0)+\frac{h^2}{2}g''_{\vec{x},\vec{w}}(0)+o(h^2)$).

Une fonction f C^2 est strictement convexe ssi d^2f est définie positive, i.e. ssi $d^2f(\vec{x})(\vec{w},\vec{w}) > 0$ pour tout $\vec{w} \in \mathbb{R}^n - \{\vec{0}\}$, i.e. ssi le graphe de f "ressemble localement" à un paraboloïde strictement convexe.

4.2 Définition de la transformée de Legendre

Définition 4.1 Soit Ω un ouvert dans $\mathbb{R}^n$ et $f \in C^2(\Omega;\mathbb{R})$ strictement convexe. La transformée de Legendre de f est la fonction $f^* : \mathrm{Im}(df) \to \mathbb{R}$ définie par, pour $p \in \mathrm{Im}(df)$,

$$\boxed{f^*(p) := p.\vec{x} - f(\vec{x}) \quad \text{quand} \quad \vec{x} \text{ vérifie } p = df(\vec{x}_p)}, \tag{4.3}$$

i.e. en le point $\vec{x}$ où le plan tangent au graphe de f est parallèle au plan vectoriel $\mathrm{Ker}(p) = \{\vec{v} \in \mathbb{R}^n : p(\vec{v}) = 0\}$ = le noyau de la forme linéaire p.

Autrement dit, $f^*(df(\vec{x})) := df(\vec{x}).\vec{x} - f(\vec{x})$ pour tout $\vec{x} \in \Omega$.

Définition 4.2 Définition plus générale : la transformée de Legendre de $f \in C^2(\Omega;\mathbb{R})$ strictement convexe est la fonction $f^* : \mathrm{Im}(df) \to \mathbb{R}$ définie par

$$\boxed{f^*(p) := \sup_{\vec{x}\in\Omega}(p.\vec{x} - f(\vec{x}))}, \tag{4.4}$$

cette définition pouvant être généralisée aux fonctions convexes non strictement convexes et à un domaine de définition $\supsetneq \mathrm{Im}(df)$.

Donc, si $\varphi : \mathbb{R}^n \times \mathbb{R}_n^* \to \mathbb{R}$ est définie par

$$\varphi(\vec{x}, p) = p.\vec{x} - f(\vec{x}), \tag{4.5}$$

alors (4.4) s'écrit $f^*(p) = \sup_{\vec{x}\in\Omega}\varphi(\vec{x}, p)$.

Proposition 4.3 *Soit $f \in C^2(\Omega;\mathbb{R})$ une fonction strictement convexe. (4.4) donne, pour tout $(\vec{x}, p) \in \Omega \times \mathrm{Im}(df)$,*

$$\boxed{p.\vec{x} \leq f(\vec{x}) + f^*(p)} \quad \text{(Inégalité d'Young)}. \tag{4.6}$$

Et on retrouve (4.3). Réciproquement, si (4.3) alors (4.4).

Preuve. Soit $\vec{x} \in \Omega$, soit $p \in \mathrm{Im}(df)$, soit $\varphi(\vec{x}, p) = p.\vec{x} - f(\vec{x})$. À $p \in \mathrm{Im}(df)$ fixé, soit $\vec{x}_p$ t.q. $p = df(\vec{x}_p)$ et soit $\varphi_p : \vec{x} \to \varphi_p(\vec{x}) := \varphi(\vec{x}, p) = p.\vec{x} - f(\vec{x})$. La fonction φ_p est $C^2(\Omega;\mathbb{R})$, avec $d\varphi_p(\vec{x}) = p - df(\vec{x})$ et $d^2\varphi_p(\vec{x}) = -d^2f(\vec{x}) < 0$, donc φ_p est strictement concave. Avec $d\varphi_p(\vec{x}_p) = 0$ (car $df(\vec{x}_p) = p$), donc φ_p est maximum au point $\vec{x}_p$, donc, pour tout $x \in \Omega$, $\varphi_p(x) \leq \varphi_p(x_p) = p.\vec{x}_p - f(\vec{x}_p) = f^*(p)$ comme défini en (4.3). Donc (4.4) $\Rightarrow$ (4.3) et (4.3) $\Rightarrow$ (4.4). ∎

5 Cas particulier de l'énergie cinétique dans $\mathbb{R}^n$

5.1 Lagrangien associé

Soit $E_c : \vec{v} \in \mathbb{R}^n \to E_c(\vec{v}) \in \mathbb{R}$ une fonction strictement convexe C^2. Et soit $E_{pot} : \vec{x} \in \mathbb{R}^n \to E_{pot}(\vec{x}) \in \mathbb{R}$ une fonction donnée ("potentiel"). Ici $\vec{x}$ est un vecteur position, i.e. un vecteur bipoint relatif à une origine choisie, et $\vec{v}$ est un vecteur vitesse.

Le lagrangien associé est la fonction $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ définie par

$$\boxed{L(\vec{x}, \vec{v}) = E_c(\vec{v}) - E_{pot}(\vec{x})}. \tag{5.1}$$

À $\vec{x}$ fixé on pose $E_{pot,x} := E_{pot}(\vec{x})$ et $L_x(\vec{v}) := L(\vec{x}, \vec{v})$, i.e.

$$L_x(\vec{v}) = E_c(\vec{v}) - E_{pot,x}. \tag{5.2}$$

(L_x est la fonction E_c à une constante près : $\vec{v}$ est la variable et $\vec{x}$ est un paramètre).

Sa transformée de Legendre $L_x{}^* : \mathbb{R}^*_n \to \mathbb{R}$ est donnée par

$$L_x{}^*(p) = p.\vec{v} - L_x(\vec{v}) = E_c{}^*(p) + E_{pot,x} \overset{\text{noté}}{=} H_x(p) \quad \text{quand} \quad \vec{v} \text{ est t.q. } p = dE_c(\vec{v}). \tag{5.3}$$

Interprétation : le changement de variable $\vec{v} \to p =$ "différentielle en $\vec{v}$ donnant la pente en $\vec{v}$" a donné le changement de fonction $L_x \to L_x{}^*$.

5.2 Hamiltonien associé

(5.3) défini l'hamiltonien associé $H : \mathbb{R}^n \times \mathbb{R}^*_n \to \mathbb{R}$:

$$\boxed{H(\vec{x}, p) = E_c{}^*(p) + E_{pot}(\vec{x})}. \tag{5.4}$$

6 Energie cinétique et hamiltonien de vecteurs

Il peut être difficile de comprendre les calculs pour "l'Hamiltonien de l'énergie" : pourquoi une forme linéaire p (et $H(\vec{x}, p)$) plus qu'un vecteur $\vec{p}$ (et $H(\vec{x}, \vec{p})$) ? D'autant plus que $p_i = p^i$ pour les calculs élémentaires ?

6.1 Nécessité d'un produit scalaire pour définir l'énergie cinétique

Dans $\mathbb{R}^n$ muni d'un référentiel on choisit un produit scalaire $(\cdot, \cdot)_g$, on note $(\vec{u}, \vec{w})_g = \vec{u} \bullet \vec{w}$ et $||\vec{u}|| = \sqrt{\vec{u} \bullet \vec{u}} =$ la longueur de $\vec{u}$ relative à ce produit scalaire (unité de longueur en mètre, ou en pied par exemple). L'énergie cinétique est la fonction $E_c : \mathbb{R}^n \to \mathbb{R}$ définie par, avec $m > 0$,

$$E_c(\vec{v}) = \frac{1}{2} m ||\vec{v}||^2. \tag{6.1}$$

Et pour des calculs simples, on utilise $(\vec{e}_i)$ une $(\cdot, \cdot)_g$-b.o.n., i.e. une base $(\cdot, \cdot)_g$-orthonormée, i.e. t.q. $(\vec{e}_i, \vec{e}_j)_g = \delta_{ij}$ pour tout i, j. Alors

$$E_c(\vec{v}) = \frac{1}{2} m \sum_{i=1}^n v_i^2 \quad \text{quand} \quad \vec{v} = \sum_{i=1}^n v_i \vec{e}_i. \tag{6.2}$$

Pour une base $(\vec{e}_i)$ quelconque : $E_c(\vec{v}) = \frac{1}{2} m \sum_{i,j=1}^n g_{ij} v_i v_j$ où $g_{ij} := g(\vec{e}_i, \vec{e}_j)$.

6.2 Utilisation de la quantité de mouvement vectorielle

6.2.1 Quantité de mouvement vectorielle

La quantité de mouvement (classique vectorielle)

$$\vec{p} = m\vec{v} \tag{6.3}$$

permet d'énoncer le principe fondamental de la dynamique (Newton) : $\frac{d\vec{p}}{dt} = \sum$ forces extérieures (si en particulier m est constant alors $m\frac{d\vec{v}}{dt} = \sum$ forces extérieures).

6.2.2 Rappel : représentation de Riesz et gradient associé à une différentielle

Si $\ell \in \mathcal{L}(\mathbb{R}^n;\mathbb{R}) = \mathbb{R}^*_n$ (forme linéaire), et si on choisit un produit scalaire $(\cdot,\cdot)_g$, alors on peut représenter ℓ à l'aide d'un vecteur $\vec{\ell}_g$ (théorème de représentation de Riesz) : $\exists!\vec{\ell}_g \in \mathbb{R}^n$ t.q., pour tout $\vec{w} \in \mathbb{R}^n$,

$$(\vec{\ell}_g, \vec{w})_g = \ell.\vec{w}, \quad \text{noté} \quad \vec{\ell} \bullet \vec{w} = \ell.\vec{w} \tag{6.4}$$

quand un produit scalaire est imposé. De plus $||\vec{\ell}_g|| = ||\ell||$, avec la norme usuelle dans $\mathbb{R}^*_n$ donnée par $||\ell|| = \sup\limits_{||\vec{w}||=1} |\ell.\vec{w}|$.

Donc si $f \in C^1(\mathbb{R}^n;\mathbb{R})$, si $\vec{x} \in \mathbb{R}^n$, alors $df(\vec{x}) \in \mathbb{R}^*_n$ (sa différentielle en $\vec{x}$) peut être représentée par un vecteur, noté $\vec{\text{grad}}_g f(\vec{x})$, appelé le vecteur gradient de f en $\vec{x}$ relatif à $(\cdot,\cdot)_g$: pour tout $\vec{w} \in \mathbb{R}^n$, cf. (6.4),

$$(\vec{\text{grad}}_g f(\vec{x}), \vec{w})_g = df(\vec{x}).\vec{w}, \quad \text{noté} \quad \vec{\text{grad}} f(\vec{x}) \bullet \vec{w} = df(\vec{x}).\vec{w}. \tag{6.5}$$

Interprétation : $\vec{\text{grad}}_g f(\vec{x})$ donne“la direction de plus grande pente” ; en effet, le développement limité au premier ordre s'énonce, pour tout $\vec{w} \in \mathbb{R}^n$,

$$\begin{aligned} f(\vec{x}+h\vec{w}) &= f(\vec{x}) + h\, df(\vec{x}).\vec{w} + o(h) \\ &= f(\vec{x}) + h\, \vec{\text{grad}} f(\vec{x}) \bullet \vec{w} + o(h), \end{aligned} \tag{6.6}$$

et la différence “de hauteur” $f(\vec{x}+h\vec{w}) - f(\vec{x})$ est max pour les $\vec{w} \parallel \vec{\text{grad}} f(\vec{x})$, cf. Cauchy–Schwarz.

6.2.3 Gradient de l'énergie cinétique = quantité de mouvement vectoriel

La différentielle $dE_c(\vec{v}) : \mathbb{R}^n \to \mathbb{R}$ de l'énergie cinétique E_c en $\vec{v}$ est calculé à l'aide de (4.2) : avec $E_c(\vec{v}) = \frac{1}{2} m\, \vec{v} \bullet \vec{v}$ $(= \frac{1}{2} m \sum_{i=1}^n v_i^2$ dans une b.o.n.), on a, pour tout $\vec{w}$,

$$dE_c(\vec{v}).\vec{w} := \lim_{h\to 0} \frac{E_c(\vec{v}+h\vec{w}) - E_c(\vec{v})}{h} = m\vec{v} \bullet \vec{w} = \vec{p} \bullet \vec{w}. \tag{6.7}$$

Donc

$$\vec{\text{grad}} E_c(\vec{v}) = \vec{p} = \vec{p}(\vec{v}) = \text{ la quantité de mouvement }. \tag{6.8}$$

(Dans une $(\cdot,\cdot)_g$-b.o.n. $(\vec{e}_i)$ on a $\vec{\text{grad}} E_c(\vec{v}) = \sum_{i=1}^n p_i \vec{e}_i$.)

6.2.4 Transformée de Legendre de vecteurs de E_c

La transformée de Legendre de E_c est $E_c{}^* : \mathbb{R}^*_n \to \mathbb{R}$ donnée par $E_c{}^*(p) = p.\vec{v} - E_c(\vec{v})$ quand $\vec{v}$ vérifie $dE_c(\vec{v}) = p$ $(\in \mathbb{R}^*_n)$. À la forme linéaire $p \in \mathbb{R}^*_n$ on associe son $(\cdot,\cdot)_g$ vecteur de représentation de Riesz $\vec{p}$: donné par, pour tout $\vec{w} \in \mathbb{R}^n$,

$$\vec{p} \bullet \vec{w} = p.\vec{w}. \tag{6.9}$$

On peut alors définir “la transformée vectorielle” de Legendre $E_c{}^*$ de E_c :

$$\widetilde{E_c}^*(\vec{p}) = E_c{}^*(p). \tag{6.10}$$

Donc

$$\widetilde{E_c}^*(\vec{p}) = \vec{p} \bullet \vec{v} - E_c(\vec{v}) \quad \text{quand} \quad \vec{v} \text{ vérifie } m\vec{v} = \vec{p}. \tag{6.11}$$

Donc

$$\widetilde{E_c}^*(\vec{p}) = \frac{1}{2} \frac{||\vec{p}||^2}{m}. \tag{6.12}$$

6.2.5 Lagrangien et hamiltonien “de vecteurs” associé

On reprend le lagrangien donné en (5.1) :

$$L(\vec{x},\vec{v}) = E_c(\vec{v}) - E_{pot}(\vec{x}), \quad \text{et} \quad L_x(\vec{v}) = \frac{1}{2} m ||\vec{v}||^2 - E_{pot,x} \tag{6.13}$$

à $\vec{x}$ fixé. D'où sa “transformée vectorielle” de Legendre :

$$\widetilde{L_x}^*(\vec{p}) = \frac{1}{2} \frac{||\vec{p}||^2}{m} + E_{pot,x} \overset{\text{noté}}{=} \tilde{H}_x(\vec{p}), \tag{6.14}$$

D'où "l'hamiltonien vectoriel" associé :

$$\boxed{\widetilde{H}(\vec{x},\vec{p}) = E_c{}^*(\vec{p}) + E_{pot}(\vec{x})} = \frac{1}{2}\frac{||\vec{p}||^2}{m} + E_{pot}(\vec{x}). \tag{6.15}$$

"Interprétation" : $\vec{p} = m\vec{v}$ donne $\widetilde{H}(\vec{x},\vec{p}) = \frac{1}{2}m||\vec{v}||^2 + E_{pot}(\vec{x}) =$ énergie totale... mais ici ce n'est pas $\vec{v}$ la variable mais $\vec{p}$ (la quantité de mouvement) :

$$E(\vec{x},\vec{v}) = E_c(\vec{v}) + E_p(\vec{x}) \quad \underset{\vec{v}\to\vec{p}=m\vec{v}}{\overset{\text{ch. var.}}{\longrightarrow}} \quad \widetilde{H}(\vec{x},\vec{p}) = E_c{}^*(\vec{p}) + E_{pot}(\vec{x}). \tag{6.16}$$

6.3 Notations de dualité, et $p_i = p^i$

6.3.1 Convention

(Convention d'Einstein.) Si $(\vec{e}_i)$ est une base dans $\mathbb{R}^n$ (les i sont "en bas"), si $\vec{v} \in \mathbb{R}^n$, alors les composantes de $\vec{v}$ relatives à la base $(\vec{e}_i)$ sont notées v^i (les i sont "en haut") :

$$\vec{v} = \sum_{i=1}^n v^i \vec{e}_i. \tag{6.17}$$

La base duale de la base $(\vec{e}_i)$ est notée (e^i) (les i sont "en haut") : c'est la base de $\mathbb{R}^*_n$ définie par : les $e^i \in \mathbb{R}^*_n = \mathcal{L}(\mathbb{R}^n;\mathbb{R})$ sont (les formes linéaires) définies par

$$(e^i(\vec{e}_j) =) \quad e^i.\vec{e}_j = \delta^i_j \tag{6.18}$$

(vaut 1 si $i = j$ et vaut 0 sinon). Autrement dit, e^i est l'opérateur de projection sur la droite vectorielle $\text{Vect}\{\vec{e}_i\}$ parallèlement aux autres directions $(\vec{e}_j)$. Donc avec (6.17) on a :

$$e^i.\vec{v} = v^i \quad \text{: donne la } i\text{-ème composante d'un vecteur } \vec{v} \text{ dans la base } (\vec{e}_i). \tag{6.19}$$

En effet, e^i étant linéaire, $e^i.(\sum_{j=1}^n v^j\vec{e}_j) = \sum_{i=1}^n v^j e^i.\vec{e}_j = \sum_{i=1}^n v^j\delta^i_j = v^i$.

Si $\ell \in \mathbb{R}^*_n$ (forme linéaire) on note ℓ_i (les i sont "en bas") ses composantes sur la base duale (e^i) :

$$\ell = \sum_{i=1}^n \ell_i e^i, \quad \text{où donc} \quad \ell_i = \ell.\vec{e}_i. \tag{6.20}$$

Donc avec (6.17) on a

$$\ell.\vec{v} = \sum_{i=1}^n \ell_i v^i. \tag{6.21}$$

Et une somme comme $\sum_{i=1}^n \ell_i v^i$ n'est licite pour la convention que si les i sont une fois en haut et une fois en bas. Si cette convention n'est pas respectée alors le résultat "n'est pas intrinsèque" (n'est pas objectif) : il dépend du choix d'un utilisateur (en pratique : dépend d'un produit scalaire ou d'une base).

6.3.2 La différentielle df exprimée dans une base

Pour $\ell = df(\vec{x})$, cf. (4.2), on note $\ell_i = \frac{\partial f}{\partial x^i}(\vec{x})$ ses composantes (les i sont "en bas") :

$$\ell_i = \frac{\partial f}{\partial x^i}(\vec{x}) := df(\vec{x}).\vec{e}_i, \quad \text{et donc} \quad df(\vec{x}) = \sum_{i=1}^n \frac{\partial f}{\partial x^i}(\vec{x})e^i. \tag{6.22}$$

6.3.3 Expression d'un vecteur de représentation de Riesz, et gradient

Soit une base $(\vec{e}_i)$, et (e^i) sa base duale. Si on choisit un produit scalaire $(\cdot,\cdot)_g$, si $g_{ij} = g(\vec{e}_i,\vec{e}_j)$, et si $\ell \in \mathbb{R}^*_n$ (forme linéaire) alors (6.4) donne $(\vec{\ell}_g,\vec{e}_j)_g = \ell.\vec{e}_i$, donc

$$\ell = \sum_{i=1}^n \ell_i e^i \quad \text{et} \quad \vec{\ell}_g = \sum_{i=1}^n \ell^i \vec{e}_i, \quad \text{avec} \quad \sum_{i=1}^n g_{ij}\ell^i = \ell_j. \tag{6.23}$$

(Problème matriciel $[g_{ij}].\begin{pmatrix}\ell^1\\ \vdots\\ \ell^n\end{pmatrix} = \begin{pmatrix}\ell_1\\ \vdots\\ \ell_n\end{pmatrix}$ à résoudre pour trouver les ℓ^i.)

Cas particulier $(\vec{e}_i)$ est une $(\cdot,\cdot)_g$-b.o.n. : on a $g_{ij}=\delta_{ij}$, donc $\ell^i=\ell_i$.

En particulier pour $\ell = df(\vec{x}) \sum_{i=1}^n \frac{\partial f}{\partial x^i}(\vec{x}) e^i$ on a $\vec{\ell}_g = \vec{\mathrm{grad}}_g f(\vec{x}) =^{\text{noté}} \sum_{j=1}^n \ell^j \vec{e}_j$, et

$$(\vec{\mathrm{grad}}_g f(\vec{x}), \vec{e}_i)_g = df(\vec{x}).\vec{e}_i \quad \text{donne} \quad \sum_{i=1}^n g_{ij}\ell^j = \frac{\partial f}{\partial x^i}(\vec{x}), \tag{6.24}$$

d'où les composantes ℓ^i et $\vec{\mathrm{grad}}_g f(\vec{x})$ par résolution d'un problème matriciel.

Cas particulier $(\vec{e}_i)$ est une $(\cdot,\cdot)_g$-b.o.n. : $\ell^i = \frac{\partial f}{\partial x^i}(\vec{x})$ (noter que i est "en haut" à gauche et i est "en bas" à droite : relation non intrinsèque car elle dépend du choix d'un produit scalaire).

6.3.4 Pour l'énergie cinétique

Dans notre cas de l'énergie cinétique, les notations donnent

$$\ell = dE_c(\vec{v}) = p = \sum_{i=1}^n p_i e^i, \quad \text{où} \quad p_i = \frac{\partial E_c}{\partial v^i}(\vec{v}). \tag{6.25}$$

Donc avec un produit scalaire $(\cdot,\cdot)_g$ et une $(\cdot,\cdot)_g$-b.o.n. $(\vec{e}_i)$:

$$\vec{\ell}_g = \vec{\mathrm{grad}}_g E_c(\vec{v}) = \sum_{i=1}^n p^i \vec{e}_i, \quad \text{où} \quad \sum_{i=1}^n g_{ij}p^j = \frac{\partial E_c}{\partial v^i}(\vec{v}). \tag{6.26}$$

Cas particulier $(\vec{e}_i)$ est une $(\cdot,\cdot)_g$-b.o.n. : alors

$$p^i = \frac{\partial E_c}{\partial v^i}(\vec{v}) = p_i ..., \tag{6.27}$$

ce qu'on trouve dans beaucoup de cours "pour simplifier"... mais qui mélange les notations et entraine une confusion entre une fonction et un vecteur, ici confusion entre la covariance (forme linéaire) et la contravariance (vecteur).

A Annexe

A.1 Ensemble convexe

Dans $\mathbb{R}$, un ensemble convexe est un intervalle (borné ou non borné). Soit $n \ge 1$. $\mathbb{R}^n$ est un espace affine, O est une origine, $\vec{\mathbb{R}^n}$ est l'espace vectoriel usuel associé, un point A est repéré à l'aide du vecteur (bi-point) $\overrightarrow{OA}$.

Définition A.1 Le segment de droite reliant deux points $A, B \in \mathbb{R}^n$ est

$$\begin{aligned}[A,B] &:= \{P \in \mathbb{R}^n : \exists t \in [0,1] \text{ t.q. } \overrightarrow{OP} = (1-t)\overrightarrow{OA} + t\overrightarrow{OB}\} \\ &= \{P = A + t\overrightarrow{AB} \overset{\text{noté}}{=} P(t) : t \in [0,1]\}.\end{aligned} \tag{A.1}$$

Un ensemble $\mathcal{C} \subset \mathbb{R}^n$ est convexe ssi pour tout $A, B \in \mathcal{C}$ on a $[A,B] \in \mathcal{C}$. Dessin.

Et le segment de droite vectoriel reliant deux vecteurs $\vec{v}_A, \vec{v}_B \in \vec{\mathbb{R}^n}$ est

$$\begin{aligned}[\vec{v}_A, \vec{v}_B] &:= \{\vec{v} \in \mathbb{R}^n : \exists t \in [0,1] \text{ t.q. } \vec{v} = (1-t)\vec{v}_A + t\vec{v}_B\} \\ &= \{\vec{v} = (1-t)\vec{v}_A + t\vec{v}_B \overset{\text{noté}}{=} \vec{v}(t) : t \in [0,1]\}.\end{aligned} \tag{A.2}$$

A.2 Fonction $\mathbb{R} \to \mathbb{R}$ affine

Définition A.2 Une fonction affine $\varphi : \mathbb{R} \to \mathbb{R}$ est une fonction de la forme : $\exists p \in \mathbb{R}$ t.q. pour tout $x \in \mathbb{R}$:

$$\varphi(x) = px + y_0. \tag{A.3}$$

Le graphe $G(\varphi) = \{(x,y) \in \mathbb{R}^2 : y = \varphi(x)\} \subset \mathbb{R}^2$ de φ est la droite de pente p qui intercepte l'axe "vertical" $x = 0$ en $y_0 = \varphi(0)$.

Une fonction affine φ est trivialement C^∞, avec $p = \varphi'(x)$ = la pente est indépendant de x, et $\varphi^{(n)}(x) = 0$ pour tout $n \geq 2$. Et toute fonction de la forme, pour $x_0 \in \mathbb{R}$:

$$\psi(x) = p(x-x_0) + \tilde{y}_0 \quad (= \varphi(x-x_0) = \text{translatée de } \varphi), \tag{A.4}$$

est également affine (poser $y_0 = \tilde{y}_0 - px_0$ dans (A.3)). Et le graphe $G(\psi) = \{(x,y) \in \mathbb{R}^2 : y = \psi(x)\}$ de ψ est la droite dans $\mathbb{R}^2$ de pente p qui intercepte l'axe "vertical" $x = x_0$ en $\tilde{y}_0$. Dessin.

A.3 Fonction $\mathbb{R} \to \mathbb{R}$ convexe

Définition A.3 Soit $a < b$. Une fonction $f :]a,b[\to \mathbb{R}$ est dite convexe ssi pour tout $x_0, x_1 \in]a,b[$, avec $x_0 \neq x_1$, le segment de droite $[P_0, P_1]$ dans $\mathbb{R}^2$ d'extrémités $P_0 = (x_0, f(x_0))$ et $P_1 = (x_1, f(x_1))$ est au dessus du graphe de f :

$$\forall x_0, x_1 \in]a,b[,\ x_0 \neq x_1,\ \forall t \in [0,1],\ \ (1-t)f(x_0) + tf(x_1) \geq f((1-t)x_0 + tx_1), \tag{A.5}$$

i.e. $(1-t)f(x_0) + tf(x_1) \geq f(x_t)$ où $x_t = (1-t)x_0 + tx_1$ et $t \in [0,1]$. Voir figure A.1.

Et f est strictement convexe ssi l'inégalité ci-dessus est stricte pour tout $t \in]0,1[$.

Et f est concave (resp. strictement concave) ssi $-f$ est convexe (resp. strictement convexe).

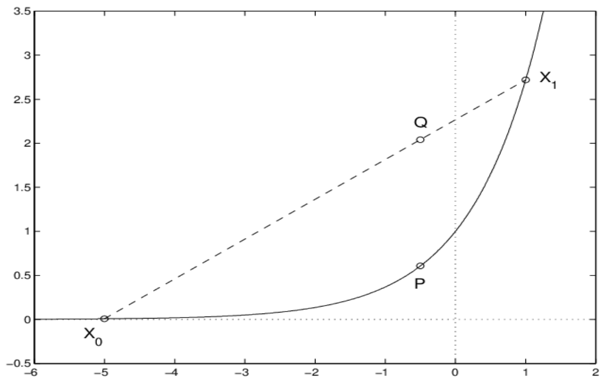

FIGURE A.1 – Fonction f convexe, ici fonction exponentielle (graphe en trait plein). les points $P_0 = (x_0, f(x_0))$ et $P_1 = (x_1, f(x_1))$ sont deux points du graphe. Le segment $[P_0, P_1]$ les joignant (trait pointillé) est au dessus du graphe. Soit, avec $x_t = (1-t)x_0 + tx_1$, le point $Q = (x_t, (1-t)f(x_0) + tf(x_1))$ (sur le segment) est au dessus du point $P = (x_t, f(x_t))$ (sur le graphe).

Exemple A.4 Si f est affine, alors on a égalité dans A.5, car avec $f(x) = px + y_0$, on a $(1-t)f(x_0) + tf(x_1) = (1-t)(px_0 + y_0) + t(px_1 + y_0) = a((1-t)x_0 + tx_1) + y_0$ qui est bien égal à $f(x_t)$. ▟

Exemple A.5 Soit $f : \mathbb{R} \to \mathbb{R}$ donnée par $f(x) = x^2$. Son graphe est une parabole, son graphe est strictement convexe, et f est strictement convexe. ▟

Exercice A.6 Montrer : 1- si $f \in C^1(]a,b[;\mathbb{R}$ et si $f' \neq 0$ alors f est inversible C^1, 2- si $f' > 0$ (resp. $f' < 0$) alors $(f^{-1})' > 0$ (resp. $(f^{-1})' < 0$), 3- Si f C^1 est strictement convexe et $f' > 0$ (resp. $f' < 0$) alors f^{-1} est strictement concave (resp. convexe). (Exemples : $f(x) = e^x$ et $g(x) = e^{-x}$ donnent $x = f^{-1}(y) = \log(y)$ et $x = g^{-1}(y) = -\log(y)$.)

Réponse. 1- f' est C^0 et $f' \neq 0$ alors soit $f' > 0$ soit $f' < 0$ (sinon $\exists x_0$ t.q. $f'(x_0) = 0$ grâce au théorème des valeurs intermédiaires). 2- f est inversible car si $f' > 0$ (resp. < 0) alors $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$ (resp. $f(x_2) < f(x_1)$) donc f est injective, donc f bijective sur son image $f(]a,b[)$ qui vaut $]f(a), f(b)[$ (resp. $]f(b), f(a)[$). Et $\frac{f^{-1}(y_2)-f^{-1}(y_1)}{y_2-y_1} = \frac{x_2-x_1}{f(x_2)-f(x_1)}$ pour $y_1 \neq y_2$ et $y_i = f(x_i)$ donne f^{-1} dérivable. 3- Et si $f' > 0$ alors $((f^{-1})' \circ f)(x) = x$ et $y = f(x)$ donnent $(f^{-1})'(y) = \frac{1}{f'(x)} > 0$ et $(f^{-1})'$ décroissante stricte et f' croissante stricte, et si $f' < 0$ alors $(f^{-1})'(y) = \frac{1}{f'(x)} < 0$ et $(f^{-1})'$ croissante stricte et f' décroissante stricte. ▟

Exercice A.7 Montrer : f est convexe ssi

$$\forall x_0, x_1 \in]a,b[,\ \frac{f(x_0) + f(x_1)}{2} \geq f(\frac{x_0 + x_1}{2}). \tag{A.6}$$

▟

Remarque A.8 Dans (A.5) on ne suppose pas que $x_0 < x_1$, i.e. que $x_0 < x_1$ ou que $x_1 < x_0$, la définition est la même : poser $u = 1-t$, donc $u \in]0,1[$ et $x_t = (1-u)x_1 + ux_0 = \tilde{x}_u$, donc on a inversé le sens de parcourt du segment $[P_0, P_1]$, voir figure (A.1), segment parcouru de P_0 vers P_1 pour quand t croît, et segment parcouru de P_1 vers P_0 pour quand u croît. ∎

Définition A.9 L'épigraphe de f est l'ensemble des points $(x,y) \in \mathbb{R} \times \mathbb{R}$ qui sont au dessus du graphe, c.à.d. :

$$\mathrm{epi}(f) = \{(x,y) \in \mathbb{R} \times \mathbb{R} : y \geq f(x)\}. \tag{A.7}$$

Voir figure A.1 : partie au-dessus du graphe.

Proposition A.10 *f est convexe ssi son épigraphe est convexe. Donc une fonction est convexe ssi, pour tout $x_0, x_1 \in]a,b[$, le segment de droite $[P_0, P_1]$ dans $\mathbb{R}^2$ d'extrémités $P_0 = (x_0, f(x_0))$ et $P_1 = (x_1, f(x_1))$ (points du graphe de f) est tout entier dans l'épigraphe, voir figure A.1.*

(Cette proposition peut servir de définition d'une fonction convexe.)

Preuve. Exercice. ∎

Proposition A.11 *Soit $f :]a,b[\to \mathbb{R}$ convexe, $x_0, x_1 \in]a,b[$ avec $x_0 < x_1$, $\lambda \in]0,1[$, et $x_\lambda = (1-\lambda)x_0 + \lambda x_1$: la pente moyenne entre x_0 et x_λ vérifie :*

$$\frac{f(x_\lambda) - f(x_0)}{x_\lambda - x_0} \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0}, \tag{A.8}$$

c.à.d. la pente moyenne est une fonction croissante de λ (ici $x_0 < x_1$). Faire un dessin. Et

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x_1) - f(x_\lambda)}{x_1 - x_\lambda}. \tag{A.9}$$

En particulier, si f est dérivable dans $]a,b[$, alors pour $x_0, x_1 \in]a,b[$ t.q. $x_0 < x_1$, et pour $\lambda \in]0,1[$:

$$f'(x_0) \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq f'(x_1), \quad \text{i.e.} \quad f'(x_0) \leq f'(x_\lambda) \leq f'(x_1), \tag{A.10}$$

et f' est une fonction croissante sur $]a,b[$, avec inégalités strictes si f est strictement convexe, auquel cas f' est croissante stricte.

Et si $f \in C^2(]a,b[;\mathbb{R})$ alors $f'' \geq 0$ sur $]a,b[$, avec $f'' > 0$ si f est strictement convexe, auquel cas f' est un C^1 difféomorphisme (une application bijective C^1 d'inverse C^1).

Preuve. On a $x_\lambda - x_0 = \lambda(x_1 - x_0)$. Donc (A.8) équivaut à $\frac{f(x_\lambda)-f(x_0)}{\lambda(x_1-x_0)} \leq \frac{f(x_1)-f(x_0)}{x_1-x_0}$, i.e. $f(x_\lambda) - f(x_0) \leq \lambda(f(x_1) - f(x_0))$, i.e. $f(x_\lambda) \leq (1-\lambda)f(x_0) + \lambda f(x_1)$; vrai car f est convexe.

On a $x_1 - x_\lambda = (1-\lambda)(x_1 - x_0)$. Donc (A.9) équivaut à $\frac{f(x_1)-f(x_0)}{x_1-x_0} \leq \frac{f(x_1)-f(x_\lambda)}{(1-\lambda)(x_1-x_0)}$, i.e. $(1-\lambda)(f(x_1) - f(x_0)) \leq f(x_1) - f(x_\lambda)$, i.e. $f(x_\lambda \leq (1-\lambda)f(x_0) + \lambda f(x_1)$; vrai car f est convexe.

Soit $\lambda, \mu \in [0,1]$, $\lambda < \mu$, et $h \in]0, \mu - \lambda[$. On a $\frac{f(x_\mu)-f(x_\lambda)}{(x_\mu - x_\lambda)} \leq \frac{f(x_\mu)-f(x_{\mu-h})}{(x_\mu - x_{\mu-h})} \longrightarrow_{h\to 0} f'(x_\mu)$ pour f dérivable. D'où :

$$\frac{f(x_\mu) - f(x_\lambda)}{x_\mu - x_\lambda} \leq f'(x_\mu), \qquad f'(x_\mu) \leq f'(x_\lambda),$$

Calcul similaire si $\lambda > \mu$. D'où(A.10).

Et les inégalités sont strictes si f est strictement convexe (par définition de la stricte convexité).

D'où si $f \in C^2$, alors $f''(x_\lambda) = \lim_{\mu \to \lambda^+} \frac{f'(x_\mu) - f'(\lambda)}{\mu - \lambda} \geq 0$, avec > 0 si f est strictement convexe (donc f' strictement croissante).

Si f est strictement convexe et C^2, on applique le lemme suivant A.12 avec $h = f'$ qui est C^1 strictement croissante. ∎

Lemme A.12 *Soit I et J deux intervalles ouverts dans $\mathbb{R}$ et $k = 0$ ou 1.*

Si $g \in C^k(I;J)$ est bijective, alors g est monotone stricte et $g^{-1} \in C^k(J;I)$ (donc pour $k = 0$ g est un homéomorphisme C^0, et pour $k = 1$ g est un difféomorphisme C^1). Et g^{-1} est croissante (resp. décroissante) ssi g est croissante (resp. décroissante).

Preuve. Soit g bijective continue. Soit $a, b \in I$, $a < b$. On a $g(a) \neq g(b)$ car g est injective. Supposons $g(a) < g(b)$ (sinon on considère $-g$). Soit $x \in]a, b[$. Montrons que $g(a) < g(x) < g(b)$.

On a $g(x) \neq g(a)$ car g est injective. Si $g(x) < g(a)$, donc $g(x) < g(a) < g(b))$, alors il existe $c \in [x, b]$ tel que $g(c) = g(a)$ (théorème des valeurs intermédiaires car g est continue sur $[x, b]$, voir polycopié “Fonctions de plusieurs variables...”). Absurde pour g injective. Idem si $g(x) \leq g(b)$. Donc $g(a) < g(x) < g(b)$. Donc g est strictement croissante sur $]a, b[$, pour tout $a, b \in I$, donc sur I.

Montrons que g^{-1} est continue. Supposons g strictement croissante (sinon on prend $-g$). Soit $y_0 \in J$ et $x_0 = f^{-1}(y_0)$; il s'agit de montrer : $\forall \beta > 0$, $\exists \eta > 0$, t.q. $\forall y \in]y_0 - \eta, y_0 + \eta[$, on a $g^{-1}(y) \in]g^{-1}(y_0) - \beta, g^{-1}(y_0) + \beta[$. Prenons β suffisamment petit pour que $]x_0 - \beta, x_0 + \beta[\subset I$.

g étant continue, l'image $g(]x_0 - \beta, x_0 + \beta[)$ est un intervalle qui contient y_0 (théorème des valeurs intermédiaires). Et g étant strictement croissante, notant $\eta_1 > 0$ et $\eta_2 > 0$ les réels tels que $g(]x_0 - \beta, x_0 + \beta[) =]y_0 - \eta_1, y_0 + \eta_2[$, alors $\eta = \min(\eta_1, \eta_2)$ convient.

Si de plus $g \in C^1$, alors $g'(x) > 0$ pour tout $x \in I$. Et $g^{-1}(g(x)) = x$. D'où g^{-1} est dérivable de dérivée $(g^{-1})'(y) = \frac{1}{g'(x)}$ où $x = g^{-1}(y)$. Comme $g' \in C^0$ et $g^{-1} \in C^0$, $g' \circ g^{-1} \in C^0$, et est un fonction qui ne s'annule jamais. Donc $(g^{-1})' \in C^0$. ∎

Références

[1] Arnold V. : *Mathematical Methods of Classical Mechanics*. Second Edition, Springer–Verlag, 1989.