Le contrôle optimal sans douleur ... ou presque ; approche économique.

Les puissantes méthodes du contrôle optimal sont d'un accès aride et difficile, ce qui limite leur portée pratique. Les longues et subtiles démonstrations nécessaires pour établir les principaux résultats théoriques semblent réservées à des mathématiciens professionnels. L'utilisateur potentiel risque de se sentir mal à l'aise en utilisant des résultats dont il ne maîtrise pas le sens profond. Ceci est d'autant plus regrettable que, sous réserve d'y mettre beaucoup d'ordre et de méthode, les applications pratiques sont tout à fait accessibles.

Nous basant sur une expérience acquise dans le domaine des études économiques d'entreprise nous essayons dans la présente note de fournir au lecteur un fil conducteur, celui de l'interprétation économique. On pourra ainsi, sans se perdre dans la rigueur mathématique, à la fois démontrer les principales équations caractérisant l'optimum et comprendre leur véritable signification. Une fois cet apprentissage fait, l'écriture de ces équations devient aussi naturelle et automatique que celle d'un programme linéaire.

Nos sources sont essentiellement les références [1] et [2] , qui ne seront malheureusement pas très utiles au lecteur, car ces ouvrages relativement anciens sont devenus introuvables. Nous ne doutons pas que des textes beaucoup plus modernes existent. Sont-ils plus digestes ? Rien n'est moins sûr. Le premier ouvrage cité est un texte mathématique pur, très rigoureux, mais d'une conception et donc d'une lecture difficiles. Son avantage est de fournir des théorèmes complets dans les moindres détails, donc précieux pour les applications. Le second, écrit par un économiste, nous a fourni l'essentiel des idées d'interprétation.

Nous examinons deux versions d'un problème classique de contrôle optimal, puis un problème plus pointu comportant des éléments héréditaires.

(Nous utilisons, sauf exception signalée, des fonctions et des arguments vectoriels. Les seuls éléments scalaires sont la fonction économique et la variable temps.)

1er cas :

Considérons donc le problème suivant :

$$
\mathrm{I}\qquad
\begin{array}{ll}
\left\| \ \text{Max} \int_0^T \mathcal{L}(x,u,t)dt \right. & \textit{multiplicateurs} \\
\| \ \dot{x} = f(x,u,t) & \psi \\
\| \ g(x,u,t) \geqslant 0 & \lambda \\
\| \ g_T[x(T)] \ \geqslant 0 & \nu \\
\| \ x_0 = x(0) &
\end{array}
$$

On forme l'hamiltonien (nous choisissons ici pour la clarté de l'exposé de ne pas y inclure les contraintes) :

$$(1) \qquad \mathcal{H}(x,u,t) = \mathcal{L}(x,u,t) + \psi(t)\, f(x,u,t)$$

Pour une trajectoire optimale $(\bar{x}\ ,\ \bar{u}\ )$ on obtient les équations suivantes, nécessaires mais pas toujours suffisantes[1] :

1. Pour alléger le texte nous renvoyons en annexe le rappel des propriétés détaillées des multiplicateurs.

$$
\begin{array}{ll}
 & |\ -\ \dot{\psi} = \nabla_x \bar{\mathcal{H}} + \lambda \nabla_x \bar{g} \\
 & |\quad 0 = \nabla_u \bar{\mathcal{H}} + \lambda \nabla_u \bar{g} \\
(2) & |\quad \lambda(t)\, g(\bar{x}, \bar{u}, t) = 0 \qquad \text{sans sommation, et} \qquad \lambda(t) \geqslant 0 \\
 & |\quad \nu\, g_T[x(T)] = 0 \qquad \text{et} \qquad \nu \geq 0 \\
 & |\quad \psi(T) = \nu \nabla_x g_T[x(T)]
\end{array}
$$

Bien entendu l'on a également :

$$\mathcal{H}(\bar{x}, u, \psi, t) \leqslant \mathcal{H}(\bar{x}, \bar{u}, \psi, t)$$

Pour plus de clarté vis à vis d'éventuels lecteurs qui ne seraient pas habitués à l'écriture matricielle des équations, précisons l'une des équations en écriture tensorielle détaillée :

$$\dot{\psi}_i = -\frac{\partial \bar{\mathcal{L}}}{\partial x^i} - \psi_k \frac{\partial \bar{f}^k}{\partial x^i} - \lambda_j \frac{\partial \bar{g}^j}{\partial x^i}$$

où tout indice redoublé en position haute et basse est sommé selon la convention d'Einstein.[2]

$\bar{\mathcal{H}}$ veut dire $\mathcal{H}(\bar{x}, \bar{u}, t)$, etc...

Nous allons maintenant essayer de retrouver les principales équations (2) par une approche économique intuitive basée sur un principe général d'optimalité que nous pouvons résumer comme suit : **toute partie d'une trajectoire globalement optimale est elle-même optimale**. Il s'agit là d'une formulation un peu hâtive d'un théorème d'optimalité énoncé par Bellman pour la programmation dynamique.

Définissons la fonction économique $\mathcal{J}(\bar{x}(t), t)$ que nous devons donc rendre maximale non seulement pour $t = 0$ mais pour toute valeur de t entre 0 et T :

$$
\begin{aligned}
(3) \qquad \mathcal{J}[\bar{x}(t), t] &= \mathrm{Max}_{u,\tau} \int_t^T \mathcal{L}(x, u, \tau)\, d\tau = \mathrm{Max}_{u,\tau} \left[\int_t^{t+dt} \mathcal{L}\, d\tau + \int_{t+dt}^T \mathcal{L}\, d\tau \right] \\
&= \mathrm{Max}_{u,\tau} \left[\int_t^{t+dt} \mathcal{L}\, d\tau \right] + \mathcal{J}[\bar{x}(t+dt), t+dt]
\end{aligned}
$$

La maximisation de $\mathcal{L}$ à l'instant t (optimisation statique, à $\bar{x}(t)$ donné) est un problème classique qui conduit aux conditions nécessaires de Kuhn et Tucker :

$$(4) \qquad 0 = \nabla_u \bar{\mathcal{H}} + \lambda \nabla_u \bar{g}$$

D'autre part nous pouvons développer $\bar{\mathcal{J}}$:

$$(5) \qquad \mathcal{J}[\bar{x}(t+dt), t+dt] \simeq \mathcal{J}[\bar{x}(t), t] + (\nabla_x \bar{\mathcal{J}})\, \dot{\bar{x}}\, dt + \frac{\partial \bar{\mathcal{J}}}{\partial t}\, dt$$

$\nabla_x \bar{\mathcal{J}}$ est une fonction de t que nous appellerons $\psi(t)$, variable adjointe, selon la terminologie usuelle[3]. Nous obtenons donc l'équation d'Hamilton-Jacobi :

$$(6) \qquad \mathrm{Max}_{u(t)}[\mathcal{L}(x, u, t) + \psi(t)\, f(x, u, t)] = \mathrm{Max}_{u(t)} \mathcal{H}(x, u, t) = -\frac{\partial \bar{\mathcal{J}}}{\partial t}$$

$$(7) \qquad \mathcal{H}(\bar{x}, \bar{u}, t) = \bar{\mathcal{H}} = -\frac{\partial \bar{\mathcal{J}}}{\partial t}$$

2. Bien entendu la position des indices n'est pas quelconque. Trop souvent les auteurs sont laxistes dans ce domaine, alors qu'une notation strictement tensorielle s'impose dans des espaces non métriques, et se révèle très utile pour l'enchaînement correct des calculs et la détection d'erreurs, même si aucun calcul complexe n'est mis en oeuvre.

3. $\nabla_x \bar{\mathcal{J}}$ fait partie de l'espace dual de E^n, espace de x. Les composantes de ψ seront donc notées ψ_i .

Supposons qu'à l'instant t la valeur de $\bar{x}$ soit augmentée, toutes choses égales par ailleurs, de $\delta\bar{x}$, et que du fait de l'optimisation les valeurs de $\bar{u}(\tau)$ varient de $\delta\bar{u}(\tau)$. A l'instant t on peut écrire symboliquement, sans nous préoccuper du calcul effectif de la matrice de passage (résultat de l'optimisation statique) :

(8) $\qquad \delta\bar{u} = \nabla_x \bar{u}\ \delta\bar{x}$

Compte tenu de la relation $\lambda\bar{g} = 0$, écrite ici avec sommation, on a la relation matricielle :

(9) $\qquad \lambda(\nabla_u \bar{g})(\nabla_x \bar{u}) + \lambda\nabla_x \bar{g} = 0$

Il y a ici un point de rigueur compte tenu du fait que certaines contraintes sont saturées et d'autres non (pour ces dernières on a $\lambda = 0$). Un bon exercice est de refaire tout le calcul en notation tensorielle en distinguant les deux types de contraintes.

En dérivant la relation (7) par rapport à $\bar{x}(t)$ on obtient en tenant compte de (4) et (9) :

$$\nabla_x \bar{\mathcal{H}} + (\nabla_u \bar{\mathcal{H}})(\nabla_x \bar{u}) = \nabla_x \bar{\mathcal{H}} - \lambda(\nabla_u \bar{g})(\nabla_x \bar{u}) = \nabla_x \bar{\mathcal{H}} + \lambda\nabla_x \bar{g} = -\nabla_x\left(\frac{\partial\bar{\mathcal{J}}}{\partial t}\right) = -\frac{\partial(\nabla_x \bar{\mathcal{J}})}{\partial t} = -\dot{\psi}$$

Nous avons donc aisément retrouvé l'équation dite de coordination temporelle :

(10) $\qquad \nabla_x \bar{\mathcal{H}} + \lambda\nabla_x \bar{g} = -\dot{\psi}$

Réécrivons l'équation (7) sous la forme :

(11) $\qquad \mathcal{L}(\bar{x}, \bar{u}, t) + \psi(t)\,\dot{\bar{x}}(t) = -\dfrac{\partial\bar{\mathcal{J}}}{\partial t}$

Si l'on considère $\mathcal{L}$ comme un résultat brut d'activité, au sens flux de recettes moins flux de dépenses pendant la période dt , x comme un stock de produits, u comme un facteur de production ou d'activité, on peut interpréter l'équation (11) ainsi :

La somme du flux $\bar{\mathcal{L}}$ (contribution directe) et de la variation $\psi\,\dot{\bar{x}}$ des stocks à prix constants (contribution indirecte), que nous appellerons donc "résultat économique hors effet de prix sur stocks", est égale (en situation optimale) à la perte de valeur de la fonction économique $\bar{\mathcal{J}}$(t) entre t et $t + dt$, ce qui est le bon sens même.

Du point de vue de l'économiste il est intéressant d'examiner aussi le résultat complet, comprenant la variation de prix sur stocks (la "plus-value sur stocks" qui a fait couler tant d'encre ...) :

(12) $\qquad V(\bar{x}, \bar{u}, t) = \bar{\mathcal{L}} + \psi\,\dot{\bar{x}} + \dot{\psi}\,\bar{x} = \bar{\mathcal{L}} + (\dot{\psi\,\bar{x}})$

$\bar{V}$ est donc le résultat "comptable" de la période considérée, à la différence importante près que les comptables utilisent, en lieu et place des ψ , des prix de revient comptables, ... mais ceci est une autre histoire.

Reprenons alors les équations de l'optimum (4) et (10). En les multipliant respectivement par $\delta\bar{u}$ et $\delta\bar{x}$ et en les additionnant, on obtient (à l'instant t considéré) :

(13) $\qquad \delta\bar{\mathcal{H}} + \lambda\,\delta\bar{g} = -\dot{\psi}\,\delta\bar{x}$

donc :

(14) $\qquad \delta\bar{V} = -\lambda\,\delta\bar{g}$

L'équation (14) permet de donner une interprétation économique simple des multiplicateurs λ . Si l'on rendait la contrainte g^j plus sévère d'une quantité δg^j [c'est à dire $g^j \geqslant \delta g^j$] , le résultat "comptable" de la période $(t, t + dt)$ se trouverait diminué de $\lambda_j\,\delta g^j$. Ce résultat suppose que l'on accepte des variations $(\delta\bar{x},\ \delta\bar{u})$ à l'instant t, reliées par l'équation (8). Si en revanche on impose $\delta\bar{x} = 0$ on a :

$$\delta\bar{V} = -\lambda\,\delta\bar{g} = \delta\bar{\mathcal{H}}$$

ce qui est cohérent avec l'interprétation de λ en optimisation statique.

Nous avons donc retrouvé, sans devoir nous lancer dans des analyses mathématiques hors de portée, les deux principales équations d'optimalité du problème (I). Une autre méthode, plus facile en apparence, mais plus lourde et manquant de rigueur, consisterait à écrire le problème sous forme discrète. Les équations optimales sont obtenues sans difficulté par les classiques conditions K-T. Les équations (1) et (2) apparaissent alors comme des formes limites lorsque l'on fait tendre le pas de temps vers zéro.

1er cas (variante) :

Le problème (I) devient sensiblement plus compliqué si certaines contraintes ne contiennent que des variables d'état. Ce cas est malheureusement fréquent dans les problèmes pratiques.

Considérons par exemple le problème suivant, qui ne comporte que de telles contraintes :

	$\operatorname{Max} \int_0^T \mathcal{L}(x,u,t)dt$	*multiplicateurs*
	$\dot{x} = f(x,u,t)$	ψ
II	$g(x,t) \geqslant 0$	λ
	$g_T[x(T)] \geqslant 0$	ν
	$x_0 = x(0)$	

En dérivant le premier membre des contraintes, nous faisons apparaître les fonctions :

(15) $\quad h(x,u,t) = (\nabla_x g)\, f(x,u,t) + \frac{\partial g}{\partial t}$

c'est à dire en notation tensorielle :

(15bis) $\quad h^j(x,u,t) = \frac{\partial g^j}{\partial x^i}\, f^i(x,u,t) + \frac{\partial g^j}{\partial t}$

L'hamiltonien reste :

(16) $\quad \mathcal{H}(x,u,t) = \mathcal{L}(x,u,t) + \psi(t)\, f(x,u,t)$

Dans une formulation valable pour l'ensemble de la trajectoire les équations de l'optimum sont :

$$
(17) \qquad \left|\begin{array}{ll}
\nabla_x \bar{\mathcal{H}} + \lambda \nabla_x \bar{h} = -\dot{\psi} & \\
\nabla_u \bar{\mathcal{H}} + \lambda \nabla_u \bar{h} = 0 & \\
\dot{\lambda}(t)\, \bar{g} = 0 & \text{sans sommation} \\
\nu\, g_T[x(T)] = 0 \quad \text{et} \quad \nu \geq 0 & \\
\psi(T) = \nu \nabla_x g_T[x(T)] &
\end{array}\right.
$$

Comme dans le cas précédent, avec la même définition pour $\bar{\mathcal{J}}(\mathrm{t})$, on trouve :

(18) $\quad \operatorname{Max}_{u,\tau} [\int_t^{t+dt} \mathcal{L}\, d\tau] + (\nabla_x \bar{\mathcal{J}})\, \dot{x}\, dt = -\frac{\partial \bar{\mathcal{J}}}{\partial t}\, dt$

Mais ici on pose :

(19) $\quad \nabla_x \bar{\mathcal{J}} = \psi + \lambda \nabla_x \bar{g}$

Donc d'après (15) :

(20) $\quad \operatorname{Max}_{u(t)}[(\mathcal{L} + \psi\, \dot{x}) + \lambda h] = -\frac{\partial \bar{\mathcal{J}}}{\partial t} + \lambda \frac{\partial \bar{g}}{\partial t}$

$$\text{(20bis)} \qquad \mathrm{Max}_{u(t)}[\mathcal{H} + \lambda h] = -\frac{\partial \bar{\mathcal{J}}}{\partial t} + \lambda \frac{\partial \bar{g}}{\partial t}$$

L'optimisation instantanée donne évidemment :

$$\text{(21)} \qquad \nabla_u \bar{\mathcal{H}} + \lambda \nabla_u \bar{h} = 0$$

(Pour simplifier l'exposé nous admettons que les λ_j correspondant aux contraintes non saturées peuvent être pris égaux à 0 sur l'arc étudié, en ajustant simultanément les ψ_i ; nous donnons en annexe quelques éléments de démonstration).

La relation (20bis) s'écrit aussi :

$$\text{(22)} \qquad \bar{\mathcal{H}} + \lambda \bar{h} = -\frac{\partial \bar{\mathcal{J}}}{\partial t} + \lambda \frac{\partial \bar{g}}{\partial t}$$

En différentiant par rapport à $\bar{x}$ on obtient :

$$\text{(23)} \qquad \nabla_x \bar{\mathcal{H}} + (\nabla_u \bar{\mathcal{H}})(\nabla_x \bar{u}) + \lambda \nabla_x \bar{h} + \lambda (\nabla_u \bar{h})(\nabla_x \bar{u}) = -\nabla_x \left(\frac{\partial \bar{\mathcal{J}}}{\partial t} - \lambda \frac{\partial \bar{g}}{\partial t}\right)$$

Soit en tenant compte de (21) :

$$\text{(24)} \qquad \nabla_x \bar{\mathcal{H}} + \lambda \nabla_x \bar{h} = -\frac{\partial}{\partial t} (\nabla_x \bar{\mathcal{J}} - \lambda \nabla_x \bar{g}) = -\dot{\psi}$$

ce qui est bien l'équation de coordination de l'optimisation entre les différentes périodes. Le fil conducteur de la démonstration est le même que précédemment.

A titre d'exemple nous examinons en annexe un modèle permettant de faire un intéressant rapprochement avec les analyses lagrangiennes et hamiltoniennes de la mécanique rationnelle.

2ème cas ; prise en compte d'éléments héréditaires :

Nous considérons des problèmes où les équations d'évolution et certaines contraintes à l'instant t sont également fonction d'instants passés, que nous désignerons par θ . Pour avoir un peu de généralité considérons le système suivant :

$$\text{III} \qquad \left\| \begin{array}{l} \mathrm{Max} \int_0^T \mathcal{L}(x,u,t)dt \\ \dot{x} = f(x,u,t) + \int_0^t \varphi[x(\theta), u(\theta), \theta, t]\, d\theta \\ g(x,u,t) + \int_0^t \phi[x(\theta), u(\theta), \theta, t] d\theta \geqslant 0 \\ g_T[x(T)] \;\; \geqslant 0 \\ x_0 = x(0) \end{array} \right.$$

Par analogie avec le problème (I) on est tenté de construire un hamiltonien du type :

$$\text{(25)} \qquad \mathcal{M} = \mathcal{L}(x,u,t) + \psi(t)\left[f(x,u,t) + \int_0^t \varphi[x(\theta), u(\theta), \theta, t]\, d\theta\right] + \lambda(t)\left[g(x,u,t) + \int_0^t \phi[x(\theta), u(\theta), \theta, t] d\theta\right]$$

Mais les multiplicateurs $\psi(t)$ et $\lambda(t)$ devraient aussi intervenir pour des équations écrites à des dates postérieures à t (dans les intégrales).

L'idée de solution est donnée par l'examen d'un problème équivalent discrétisé. On voit immédiatement que l'hamiltonien complet est donné en sommant entre 0 et T les hamiltoniens partiels de type (25).

La transposition au problème continu conduit à considérer l'intégrale $\int_0^T \mathcal{M}[x(t), u(t), t]\, dt$. On peut ensuite procéder à une manipulation classique sur les intégrales doubles en permutant d'abord l'ordre d'intégration, et dans une deuxième étape l'ordre des variables θ, t . On obtient ainsi une nouvelle formulation de l'hamiltonien global :

$$(26) \qquad \int_0^T \mathcal{M}[x(t), u(t), t]\, dt = \int_0^T \mathcal{H}[x(t), u(t), t]\, dt$$

Nous ne donnons que le résultat de ce calcul, qui ne présente pas de difficultés :

$$(27) \qquad \mathcal{H} = \mathcal{L}(x, u, t) + \psi(t) f(x, u, t) + \int_t^T \psi(\theta)\, \varphi[x(t), u(t), t, \theta]\, d\theta + \lambda(t) g(x, u, t) + \int_t^T \lambda(\theta)\, \phi[x(t), u(t), t, \theta]\, d\theta$$

Les équations d'optimisation s'en déduisent immédiatement, par analogie avec le problème (I) :

$$(28) \qquad \nabla_u \mathcal{H}[\bar{x}(t), \bar{u}(t), t] = 0 \qquad \nabla_x \mathcal{H}[\bar{x}(t), \bar{u}(t), t] = -\dot{\psi}(t)$$

Pour nous rassurer nous pouvons vérifier que ces équations sont effectivement la forme limite des équations correspondantes établies en formulation discrète.

G.Ringeisen

août 2004

Bibliographie.

[1] Hestenes   Calculus of variations and optimal control theory   John Wiley&Sons   New York 1966

[2] Albouy   La régulation économique dans l'entreprise   Dunod 1972

Annexe 1.

Pour compléter l'information du lecteur nous rassemblons ici, sans démonstration (mais certaines propriétés sont des conséquences évidentes des conditions K-T classiques), les informations essentielles concernant les variables adjointes ψ et les multiplicateurs λ .

Le cas le plus simple est le problème (I) :

– Les λ sont continus par morceaux entre 0 et T, et continus en chaque point de continuité de $\bar{u}(t)$. On a $\lambda(t) \geqslant 0$ et $\lambda_j = 0$ si $\bar{g}^j > 0$.

– On a $\nu \geqslant 0$ et $\nu_k = 0$ si $g_T^k > 0$.

– Les ψ sont continues et ont des dérivées continues par morceaux.

– La fonction $\bar{\mathcal{H}}$, élargie aux contraintes, est continue, et sur chaque intervalle de continuité de $\bar{u}$ on a $\frac{d}{dt}\mathcal{H} = \frac{\partial}{\partial t}\mathcal{H}$.

Le problème (II) est sensiblement plus compliqué. La différence porte sur les multiplicateurs λ, qui sont toujours continus par morceaux, mais sans indication de signe, et qui ont en outre les propriétés suivantes :

– Ils sont non croissants sur $(0, T)$ et constants (λ_j) sur chaque intervalle où $g^j > 0$.

– Continus quand $\bar{u}$ est continu, et, pour λ_j , en chaque point de discontinuité de $h^j(\bar{x}, \bar{u}, t)$.

Enfin donnons nous un vecteur constant c de même dimension que λ . Il est facile de vérifier que l'utilité marginale du stock $\nabla_x \bar{\mathcal{J}}$ ainsi que les équations de l'optimum ne sont pas modifiées si l'on soumet simultanément les ψ et les λ à la transformation suivante :

$$\psi \longrightarrow \psi + (\nabla_x \bar{g})\, c \qquad\qquad \lambda \longrightarrow \lambda - c$$

Cette propriété nous permet de choisir $\lambda_j = 0$ en un point quelconque de la trajectoire.

Annexe 2.

Toute personne abordant l'étude du contrôle optimal ne peut manquer d'être frappée par le parallélisme avec les notions utilisées dans l'étude de systèmes plus ou moins complexes en mécanique rationnelle.

Il nous paraît intéressant de donner quelques indications sur ce sujet par l'examen du modèle schématique suivant :

Soit $\mathcal{L}(q, \dot{q})$ le lagrangien d'un système mécanique dont on cherche à minimiser l'intégrale entre 0 et T. Soit $g(q) \geqslant 0$ un ensemble de contraintes (liaisons holonomes) et soit λ le vecteur des multiplicateurs. Nous posons $\dot{q} = u$, en prenant u comme variables de commande, auxquelles nous associons les variables adjointes $p(t)$.

L'hamiltonien s'écrit alors

$$\mathcal{H}(q, u) = -\mathcal{L}(q, u) + pu$$

Les contraintes donnent lieu aux fonctions dérivées $(\nabla_q g)\, u$, qui sont nulles pour les contraintes saturées. De plus nous choisissons $\lambda_j = 0$ sur les contraintes non saturées. Nous pouvons donc écrire (optimisation statique)

$$0 = \nabla_u \bar{\mathcal{H}} = -\nabla_u \bar{\mathcal{L}} + p$$

$$\nabla_q \bar{\mathcal{H}} = -\dot{p} \qquad\qquad (\text{car } \lambda \nabla_q \bar{h} = 0)$$

On retrouve donc bien les relations de la mécanique rationnelle :

$$p = \nabla_{\dot{q}} \bar{\mathcal{L}} \qquad\qquad \dot{p} = -\nabla_q \bar{\mathcal{H}} = \nabla_q \bar{\mathcal{L}} \qquad\qquad \dot{q} = \nabla_p \bar{\mathcal{H}}$$

et bien sûr les équations de Lagrange

$$\frac{d}{dt}(\nabla_{\dot{q}}\bar{\mathcal{L}}) - \nabla_q\bar{\mathcal{L}} = 0$$

sans faire de calcul variationnel, puisque l'effet de celui-ci est inclus dans les équations du contrôle optimal.

Lorsque le lagrangien se présente sous la forme

$$\mathcal{L}(q,u) = \mathcal{T}(u) - \mathcal{V}(q) = \text{énergie cinétique} - \text{énergie potentielle}$$

on obtient

$$\mathcal{H}(q,p) = -\mathcal{T}(u) + \mathcal{V}(q) + (\nabla_u\mathcal{T})\,u = \mathcal{T}(u) + \mathcal{V}(q) = \text{énergie totale}$$

et donc la loi de conservation de l'énergie $\dot{\mathcal{H}} = 0$.