

MECANIQUE

PRINCIPES | MÉCANIQUE ANALYTIQUE | MÉCANIQUE CLASSIQUE | MÉCANIQUE
ONDULATOIRE | MÉCANIQUE STATISTIQUE | THERMODYNAMIQUE |
MÉCANIQUE DES MILIEUX CONTINUS

MÉCANIQUE ANALYTIQUE

Nous devons la forme actuelle de la mécanique analytique (appelée aussi parfois "mécanique lagrangienne") aux travaux des frères Bernoulli et particulièrement d'Euler et Lagrange. C'est effectivement en 1696 que commence l'histoire de la vraie physique théorique telle que nous la connaissons aujourd'hui.

Au fait, l'événement de départ de la mécanique analytique provient de l'observation suivante (énoncée au 17ème siècle) : tout système semble évoluer d'un état à un autre toujours en utilisant les "moyens" les plus simples et en conservant une grandeur constante entre les deux états.

Cet énoncé est appelé "principe de moindre action" ou "principe variationnel" ou encore "principe d'économie".

Remarques :

R1. Les "moyens" précités peuvent être : le chemin le plus court, le chemin le plus rapide (les trajectoires spatio-temporelles à plus faibles amplitudes en gros...).

R2. Selon le premier principe fondamental de la physique, la grandeur constante est choisie comme étant l'énergie.

Au fait, bien que cet énoncé puisse paraître comme cohérent, il peut faire douter mais... nous verrons :

1. Qu'en mécanique classique, nous pouvons démontrer la première loi de Newton en admettant ce principe comme vrai et en y superposant le principe de conservation de l'énergie.
2. En électromagnétisme, nous retrouverons toutes les équations de Maxwell (in extenso la loi de Biot-Savart, Faraday, Force de Lorentz, loi de Laplace, etc.) à partir à nouveau des propriétés du principe de moindre action et de conservation de l'énergie.
3. En optique, nous démontrerons que le chemin suivi par la lumière est toujours le plus court et nous permettra donc de démontrer le principe de Fermat à la base de toute l'optique géométrique.
4. En physique atomique, les propriétés du principe de moindre action nous permettront de déterminer certaines propriétés mathématiques des atomes et autres particules (les fermions et les bosons en physique quantique des champs).
5. Le principe de moindre action permet également de démontrer que tout corps, avec ou sans masse, est dévié par un champ d'accélération (déterminé par la géométrie de l'espace... tiens ! y'aurait pas du chemin là dedans... ?) et... permet donc de déterminer l'équation d'Einstein des

champs qui est à la base de toute la relativité générale.

Il va donc sans dire par ces cinq petits exemples les applications phénoménales de ce principe. Au fait, tout peut se résumer en la phrase suivante :

«Toutes nos connaissances de l'Univers se résument au théorème de Noether et au principe de moindre action.»

Historiquement, il est intéressant de savoir que c'est Pierre-Louis Moreau de Maupertuis (un fervant croyant mais personne n'est parfait...) qui a énoncé le premier le principe de moindre action sous la forme (peu scientifique) suivante :

«Tous les phénomènes naturels s'accordent avec le grand principe que la Nature, dans la production de ses effets, agit toujours par les voies les plus simples. [...] On ne peut douter que toutes choses ne soient réglées par un Être suprême qui, pendant qu'il a imprimé à la matière des forces qui dénotent sa puissance, l'a destinée à exécuter des effets qui marquent sa sagesse...»

L'intervention d'Euler et Lagrange dans ce domaine a été de mettre sous forme mathématique ce principe et de démontrer (tenez-vous bien...) qu'il découle d'une simple propriété mathématiques des optima des fonctions continues. Il va sans dire, que sachant que cela a permis de redémontrer toutes les lois de la physique en a dérangé plus d'un...

Ce principe a eu (et a toujours) des répercussions inimaginables et le problème fut d'appliquer l'expression mathématique de ce dernier à tous les phénomènes physiques qui avaient déjà été démontrés de façon expérimentale et empirique à l'époque. Effectuer cette démonstration revenait ainsi à expliquer pourquoi tel phénomène ou tel loi était ainsi plutôt qu'autrement. Imaginez !

Ainsi, le premier à s'attaquer au problème fût donc le Bâlois (Suisse) Léonhard Euler. Mais nous avons également gardé le nom de Lagrange (d'où l'appellation : "formalisme lagrangien") pour définir toute la méthode et le formalisme mathématique construite autour du principe de moindre action.

FORMALISME LAGRANGIEN

Cette technique est assez peu enseignée dans les petites écoles car il faut bien l'avouer le formalisme lagrangien (contenant donc le principe de moindre action sous forme mathématique) fait appel à un niveau d'abstraction un peu plus élevé que les méthodes normales et malgré qu'il soit souvent d'une aide précieuse dans l'élaboration de théories (physique fondamentale, physique quantique, relativité générale, théorie quantique des champs, théorie des supercordes), il en découle rarement de nouvelles solutions (mais plutôt une réduction et une méthode de validation utile et très puissante).

Commençons donc notre travail :

COORDONNÉES GÉNÉRALISÉES

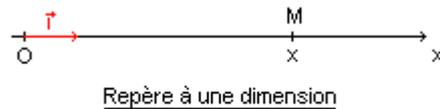
Un réflexe naturel conduit généralement à référer la position d'un point dans l'espace à la seule connaissance de ses trois coordonnées cartésiennes x , y , z (voir ci-dessus les espaces ponctuels). Cette attitude est d'ailleurs le plus souvent justifiée par la simplicité d'un grand nombre de situations rencontrées dans la pratique, où il n'est pas nécessaire de rechercher de méthodes plus élaborées ou de passées dans d'autres systèmes de coordonnées (voir le chapitre de calcul vectoriel).

Pour repérer la position d'un mobile (ou d'un point matériel) en physique il est nécessaire dans un premier d'associer un repère au référentiel. Ainsi, un "repère" est un système de repérage dans l'espace associé au référentiel.

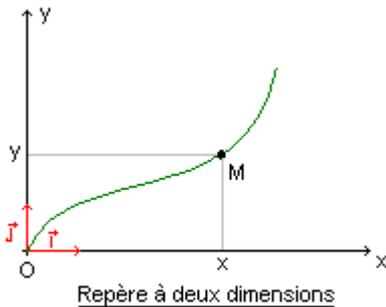
Les repères conventionnels en mécanique classique constituent majoritairement des bases

d'espaces pré-euclidiens canoniques (voir le chapitre d'analyse vectorielle dans la section d'algèbre du site) orientés et où chaque point, ou vecteur de l'espace, peut-être représenté algébriquement par ses valeurs d'affixes (la valeur à l'ordonnée (projection sur l'axe vertical) et la valeur à l'abscisse (projection sur l'axe horizontal)).

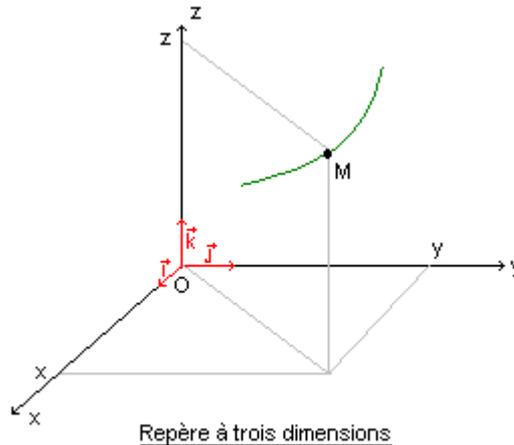
Voici quelques exemples triviaux:



Le mouvement de M s'effectue sur une droite



Le mouvement de M s'effectue dans un plan
(ou plan d'Argand-Cauchy)



Le mouvement de M s'effectue dans l'espace

Remarques:

R1. Comme nous l'avons vu en géométrie différentielle, la distance entre deux points d'une trajectoire courbe en parcourant la courbe est appelée "abscisse curviligne". Sinon, la distance entre deux points d'une trajectoire rectiligne est appelée simplement "abscisse".

R2. Un repère, assimilé à un référentiel, est dit "Galiléen" (c'est rare que nous en fassions explicitement mention en physique par manque de rigueur) si:

- nous pouvons le considérer comme immobile pendant toute l'étude du mouvement du système ou comme étant en translation rectiligne uniforme par rapport à un autre référentiel lui même immobile.

ou:

- comme un système où les lois de Newton sont vérifiées (voir le chapitre de mécanique classique)

Nous renvoyons le lecteur au chapitre de mécanique classique (chapitre suivant) pour une étude plus approfondie sur le "principe de relativité galiléen".

R3. Un repère, assimilé à un référentiel, est dit "barycentrique" (voir le chapitre de géométrie euclidienne) s'il a pour origine le centre de masse (voir le concept de "centre de masse au chapitre de mécanique classique) du corps étudié. Cette remarque est important à prendre en considération lorsque nous verrons le théorème de König en mécanique classique. Ainsi, le "repère de Copernic" est assimilé au centre de gravité (d'inertie) du système solaire et le "repère héliocentrique" ou le "repère de Kepler" au centre d'inertie du Soleil.

R4. Dire qu'un repère orthonormé (O, \vec{i}, \vec{j}) est "direct" signifie que l'angle orienté (\vec{i}, \vec{j}) a pour mesure principale $\pi/2$ (dans le sens horaire). Dire qu'un repère orthonormé (O, \vec{i}, \vec{j}) est "indirect" signifie que l'angle orienté a pour mesure principale $-\pi/2$. Dans tout ce qui suit, si nous spécifions par l'orientation, cela sous-entend que (O, \vec{i}, \vec{j}) est direct.

Il est bien exact que les trois paramètres x, y, z suffisent parfaitement à repérer un point

matériel dans l'espace usuel comme nous en avons déjà fait mention dans notre étude des espaces ponctuels, mais il n'en demeure pas moins qu'il est parfois inévitable, ou même tout simplement plus avantageux, d'utiliser un nombre de paramètres supérieur à trois. Nous pouvons évidemment envisager toutes sortes de paramétrages pour atteindre les coordonnées d'un point dans l'espace, de telle sorte que, d'une façon plus généralisée nous serons amenés à prendre en considération des relations du type (nous ne gardons plus la même écriture que celle que nous avons lors de notre étude des espace ponctuels par cohérence avec les nombreuses références déjà existant sur le sujet):

$$x = x^1 = x^1(q^1, q^2, \dots, q^n)$$

$$y = x^2 = x^2(q^1, q^2, \dots, q^n)$$

$$z = x^3 = x^3(q^1, q^2, \dots, q^n)$$

Les paramètres q^1, q^2, \dots, q^n portent le nom de "coordonnées généralisées", paramètres auxquels un problème sera le plus souvent référé. Connaître leur expression en fonction du temps est le problème fondamental de la dynamique. Cela signifie que nous serons parvenus à une solution quand nous disposerons des relations indépendantes :

$$q^1 = q^1(t)$$

$$q^2 = q^2(t)$$

....

$$q^n = q^n(t)$$

Il est donc important de retenir que le nombre de paramètres q^i définissant le repérage d'un point dans l'espace est au moins égal à trois, sans être nécessairement différent de trois. C'est finalement la nature des situations envisagées qui suggèrent le choix du nombre des paramètres à utiliser (coordonnées cartésiennes, cylindriques, sphériques,...).

Dans une vision plus générale, la configuration instantanée d'un système, quelle qu'en soit la nature, sera déterminée par la connaissance, en fonction du temps, de n paramètres, n définissant le nombre de "degrés de liberté" du système (voir le début du chapitre de la mécanique classique pour voir la définition rigoureuse du concept de degré de liberté)..

Il est tout naturel, mathématiquement, d'associer la manipulation des n paramètres x^α au recours à un hyper-espace à n dimensions, dans lequel les x^α apparaîtraient comme les coordonnées d'un point P représentatif de la configuration d'un système quelconque. Nous donnons à cet espace à n dimensions E_n , le nom "d'espace de configuration".

Mais la rigueur de la mathématique-physique, nous amène à disposer d'une description plus précise des phénomènes en ajoutant cette variable importante qu'est le temps, considérée comme variable indépendante, aux x^α . Nous en arriverons donc fatalement à utiliser un autre hyperespace E_{n+1} auquel nous avons donné le nom "d'espace des événements".

Ce dernier espace de référence revêt un intérêt capital pour un grand nombre de problèmes de la science moderne et se trouve particulièrement bien adapté aux raisonnements de nature relativiste. Les variables indépendantes constituant les coordonnées spatiales et temporelle forment alors ce que nous appelons les "variables d'Euler".

Dans la mesure où les paramètres x^α sont simplement présentés comme des fonctions explicites du temps, le point P décrit une courbe paramétrée, définie par $x^\alpha = x^\alpha(t)$, avec $\alpha = 1, 2, \dots, n$. Cela revient à exploiter simultanément les équations:

$$x^1 = x^1(t)$$

$$x^2 = x^2(t)$$

....

$$x^n = x^n(t)$$

Il arrivera fréquemment que, pour des raisons d'opportunité, nous souhaitons changer de système de coordonnées généralisées, et utiliser un autre ensemble plus compatible avec les spécificités du problème envisagé. Nous substituerons alors au jeu des x^{α} un nouveau jeu de coordonnées q^i . Il est alors évident que nous devons, avant toute chose, nous doter des relations de dépendance existant entre les deux ensembles de coordonnées, à savoir disposer des relations (voir dans la section algèbre du site au chapitre de calcul vectoriel les relations de passages pour les systèmes de coordonnées cartésiennes, polaires, cylindriques et sphériques) :

$$x^1 = x^1(q^1, q^2, \dots, q^n)$$

$$x^2 = x^2(q^1, q^2, \dots, q^n)$$

....

$$x^n = x^n(q^1, q^2, \dots, q^n)$$

Les fonctions x^1, x^2, \dots, x^n seront maintenant supposées définies, continues, de classe C^2 (pour travailler avec l'accélération) par rapport aux q^1, q^2, \dots, q^n et devront conduire à un jacobien différent de zéro (voir le chapitre de calcul intégral et différentiel).

Dans ces conditions, à chaque point E_x de l'espace des configurations des x , noté E_x , correspondra un point E_q de l'espace de configuration des q , noté E_q . Nous avons ainsi effectué une transformation ponctuelle, autrement dit une application de l'espace sur lui-même.

Pour étudier des milieux continus (concept radicalement différent du point matériel), nous aurons cependant deux approches différentes :

1. Soit nous attribuons aux particules du milieu les lois de la mécanique (dynamique des corps en mouvement) ce qui implique un traitement statistique pour en établir une dynamique pour l'ensemble des particules (cette méthode est pratiquée en théorie cinétique des fluides ou encore en mécanique statistique - c'est également une approche dite "réductionniste").
2. Soit nous examinons le mouvement d'ensemble d'un milieu composé par un très grand nombre de particules et nous le considérons comme un milieu continu. Ce concept de "milieu continu" est acceptable pour autant que le volume infinitésimal pris autour d'un point quelconque du fluide soit suffisant pour encore contenir un très grand nombre de particules.

Remarque : par "volume infinitésimal", nous sous-entendons plus précisément un volume suffisamment petit par rapport au volume continu étudié mais suffisamment important par rapport aux distances intermoléculaires. Ce sera donc un volume élémentaire contenant un grand nombre de particules bien qu'il soit mécaniquement considéré comme ponctuel.

Pour décrire le mouvement d'un milieu continu, il y a deux méthodes :

1. Méthode de Lagrange :

Le mouvement du milieu est décrit par une formulation Lagrangienne consistant donc à le caractériser en se donnant un système d'équations au sens newtonien :

$$\begin{aligned}
 x &= f_1(x_0, y_0, z_0, t) \\
 y &= f_2(x_0, y_0, z_0, t) \\
 z &= f_3(x_0, y_0, z_0, t)
 \end{aligned}$$

C'est un système d'équations dans lequel x_0, y_0, z_0 sont les coordonnées cartésiennes d'une particule du milieu continu à l'instant initial $t = 0$. En éliminant le temps, nous obtenons l'équation de la trajectoire de la particule. Par dérivations, nous avons alors la vitesse et l'accélération de la particule.

2. Méthode d' Euler :

Au lieu de suivre le parcours d'un point P , nous portons notre attention sur l'évolution des caractéristiques physiques en un point donné du fluide. Si \vec{v} (une propriété parmi tant d'autres) est la vitesse en un point P au temps t , nous avons:

$$\vec{v} = F(\vec{P}, t)$$

Dans un référentiel cartésien orthonormé, les composantes sont:

$$\begin{aligned}
 v_x &= f_1(x, y, z, t) \\
 v_y &= f_2(x, y, z, t) \\
 v_z &= f_3(x, y, z, t)
 \end{aligned}$$

où nous avons bien évidemment:

$$\vec{v} = v_x \vec{i} + v_y \vec{j} + v_z \vec{k} \quad \text{et} \quad \vec{P} = x\vec{i} + y\vec{j} + z\vec{k}$$

Lorsque nous effectuons un modèle basé sur la méthode d'Euler, nous parlons fréquemment de "système Eulérien". Nous en verrons un premier exemple en mécanique classique lorsque nous formulerons la première loi de Newton pour un point matériel et pour un milieu continu.

PRINCIPE VARIATIONNEL

Le "principe variationnel" n'est donc que la forme mathématique contemporaine du principe de moindre action qui est, comme nous en avons déjà fait mention, à la base du formalisme lagrangien (qui est donc un cadre théorique de la plus haute importance dans les recherches actuelles en physique théorique, et contient la totalité de nos connaissances actuelles en physique).

Rappelons que selon l'énoncé du principe variationnel nous devons trouver dans tout phénomène physique, une certaine quantité qui est naturellement optimisée (minimisée ou maximisée) et qui décrit toutes les variables du système étudié et ainsi son issue.

Remarque : si aucune action extérieure n'intervient dans l'évolution du système, nous parlons alors de "système libre".

Voici la démarche que nous allons suivre, une fois cette démarche présentée, nous nous attaquons à sa formalisation mathématique. Les propositions (les idées) sont les suivantes :

P1. Nous supposons donc le principe variationnel et le principe de conservation de l'énergie comme justes.

P2. L'énergie totale d'un système fermé est constante et constituée de la sommation de l'énergie cinétique et l'énergie potentielle (si nous ne considérons que l'énergie cinétique, alors le système est dit "libre", si les deux énergies sont considérées, nous disons alors que les

système est "généralisé").

P3. Nous définissons une fonction mathématique (dont les variables sont les coordonnées généralisées) appelée "Lagrangien" qui est donnée par la différence entre les deux énergies précitées.

P4. Sur l'évolution d'un système entre deux états, nous cherchons les propriétés de la fonction (du lagrangien) qui donne la minimisation de la variation de la différence des deux énergies (rappelez-vous que le principe variationnel s'appelle aussi principe d'économie et ce n'est pas pour rien) sur l'évolution temporelle ou métrique du système.

Enfin, une fois la propriété de cette fonction déterminée (propriété mise sous la forme que nous appelons "équation d'Euler-Lagrange) nous chercherons toutes les propriétés possibles autres qu'elle a afin d'avoir les outils nécessaires pour la physique théorique et vous allez voir cela marche terriblement bien...

Donc, pour mettre cela sous forme mathématique, nous commençons par poser qu'il existe une fonction réelle de $2n$ variables:

$$(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) \mapsto L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n)$$

que nous appellerons "Lagrangien généralisé" du système, dont l'intégrale satisfait à l'énoncé suivant :

Dans un mouvement naturel partant d'un point $A(q_1^A, \dots, q_n^A)$ à l'instant t_A , arrivant au point $B(q_1^B, \dots, q_n^B)$ à l'instant t_B , l'intégrale suivante, appelée "intégrale d'action" ou simplement "action":

$$S(q, \dot{q}) = \int_{t_A}^{t_B} L[q_1(t), \dots, q_n(t), \dot{q}_1(t), \dots, \dot{q}_n(t)] dt$$

ou dans une écriture abrégée dans le cas d'une paramétrisation temporelle, nous écrivons donc "l'action d'une particule ponctuelle" sous la forme :

$$S(q, \dot{q}) = \int_{t_A}^{t_B} L(q, \dot{q}) dt$$

doit être un extrémum (en fait, "un minimum" ou "un maximum", puisque nous aurions pu tout aussi bien prendre $-L$ au lieu de $+L$ dans le choix de la définition du Lagrangien généralisé).

L'action S est ce que nous appelons communément en physique une "fonctionnelle". Quand une fonction a une variable prend un nombre - l'argument - en entrée et retourne un nombre, une fonctionnelle prend une fonction en entrée et un nombre en retour. Alors qu'une fonction est usuellement définie par ses valeurs en un nombre infini de points, nous pouvons imaginer une fonctionnelle comme une fonction avec une infinité de variables.

ÉQUATION D'EULER-LAGRANGE

Le principe de moindre action énonce que (l'intégrale) S est extrémale si:

$$t \mapsto [q_1(t), \dots, q_n(t), \dot{q}_1(t), \dots, \dot{q}_n(t)]$$

est la trajectoire naturelle effectivement suivie par le système physique.

Considérons alors une trajectoire très voisine à la précédente, que nous noterons:

$$(q_1 + \delta q_1, \dots, q_n + \delta q_n, \dot{q}_1 + \delta \dot{q}_1, \dots, \dot{q}_n + \delta \dot{q}_n)$$

Remarque : nous avons omis maintenant l'écriture des arguments t des fonctions du temps afin d'alléger les écritures.

Si $S(q, \dot{q})$ est bien l'évolution d'un système évoluant selon le principe de moindre action, alors l'action donné par la variation :

$$\delta S \equiv S(q + \delta q, \dot{q} + \delta \dot{q}) - S(q, \dot{q})$$

est nulle pour δq et $\delta \dot{q}$ tendant vers zéro (sous-entendu que tout système physique revient à son état initial).

Ce qui nous amène à écrire :

$$\delta S = \int_{t_A}^{t_B} L(q + \delta q, \dot{q} + \delta \dot{q}) dt - \int_{t_A}^{t_B} L(q, \dot{q}) dt = \int_{t_A}^{t_B} [L(q + \delta q, \dot{q} + \delta \dot{q}) - L(q, \dot{q})] dt = 0$$

Ce qui nous permet de justifier la dénomination de "principe variationnel" :

$$\boxed{\delta S = \delta \int_{t_A}^{t_B} L dt = \int_{t_A}^{t_B} \delta L dt = 0}$$

Ce principe stipule donc que la trajectoire d'une particule (ou d'un système plus général) s'obtient en demandant qu'une certaine fonctionnelle S appelée "action" soit stationnaire par rapport à une variation de la trajectoire. En d'autres termes, si nous effectuons une variation infiniment petite de la trajectoire, la variation doit être nulle. Pour un système mécanique simple, l'action est alors évidemment par le principe de conservation de l'énergie égale à l'intégrale sur la trajectoire de (par définition du lagrangien) de la différence entre l'énergie cinétique et l'énergie potentielle.

Dès lors, dans une théorie pour laquelle les forces dérivent d'un potentiel V , nous sommes naturellement amenés à définir le "Lagrangien" par la relation (il faudra s'en souvenir !) :

$$\boxed{L(q, \dot{q}) = T(q, \dot{q}) - V(q, \dot{q})}$$

où T et V sont la notation traditionnelle dans le formalisme Lagrangien de l'énergie cinétique et de l'énergie potentiel données par :

$$T = \frac{1}{2} \dot{q}^2 \text{ et } V = V(q)$$

Remarque : pour l'étude de la relativité générale, nous ne chercherons pas à ce que la variation de la différence des énergies soit minimale tel que c'est le cas pour les systèmes mécaniques, mais la variation de la longueur d'un arc ds (non dépendant du temps contrairement à l'exemple précédent) dans un espace quelconque lors d'une trajectoire d'un système libre. Ce qui nous amènera à écrire simplement (rappelez-vous en aussi car ce sera très important) :

$$\delta S = \int \delta(ds) = 0$$

Pour revenir à notre application du principe variationnel dans le cas du lagrangien généralisé, nous pouvons alors écrire la différentielle totale exacte (voir chapitre de calcul différentiel et intégral) de dL et nous obtenons alors la relation :

$$\delta S = \int_{t_A}^{t_B} \sum_{i=1}^n \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) dt = 0$$

Intégrons par parties (voir chapitre de calcul différentiel et intégral) le deuxième terme de la somme de l'intégrale précédente :

$$\int_{t_A}^{t_B} \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i dt = \int_{t_A}^{t_B} \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \delta q_i dt = \frac{\partial L}{\partial \dot{q}_i} \delta q_i \Big|_{t_A}^{t_B} - \int_{t_A}^{t_B} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \delta q_i dt$$

Le premier terme de la dernière égalité est nul puisque

$$\begin{aligned} \frac{\partial L}{\partial \dot{q}_i} \delta q_i \Big|_{t_A}^{t_B} &= \frac{\partial L(q_i(t_B), \dot{q}_i(t_B))}{\partial \dot{q}_i} \delta q_i(t_B) - \frac{\partial L(q_i(t_A), \dot{q}_i(t_A))}{\partial \dot{q}_i} \delta q_i(t_A) \\ &= \frac{\partial L(q_i(t_B), \dot{q}_i(t_B))}{\partial \dot{q}_i} (\delta q_i(t_B) - \delta q_i(t_A)) = \frac{\partial L(q_i(t_B), \dot{q}_i(t_B))}{\partial \dot{q}_i} \cdot 0 = 0 \end{aligned}$$

Effectivement, par définition, les δq ont exactement le même ordre de grandeur.

L'expression de l'intégrale de moindre action peut finalement s'écrire :

$$\delta S = \int_{t_A}^{t_B} \sum_{i=1}^n \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i dt = 0$$

Mais les δq_i et $\delta \dot{q}_i$ tendent vers 0 d'une infinité de manières différentes et nous devons cependant avoir néanmoins $\delta S = 0$. Cela veut dire alors que chaque terme sommé de l'intégrale peut être pris indépendamment et doit satisfaire :

$$\forall i \in \{1, \dots, n\}, \delta S = \int_{t_A}^{t_B} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i dt = 0$$

Mais comme les fonctions δq_i et $\delta \dot{q}_i$ peuvent toujours tendre vers zéro de multiple façon, et que cette intégrale doit être quand même nulle, nous en déduisons que ce sont les intégrandes qui sont nuls :

$$\forall i \in \{1, \dots, n\}, \delta S = \frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0$$

Ces n équations, satisfaites par le lagrangien généralisé du système pour le mouvement effectivement suivi, sont appelées "équations d'Euler-Lagrange", ou plus brièvement (mais plus rarement) "équations de Lagrange". Ce sont, comme nous allons le voir, les équations du mouvement du système: résolues, elles donnent l'évolution effective du système dans le temps.

$$\boxed{\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0}$$

Remarque : c'est en étudiant la physique (les chapitres suivant du site) que l'on comprend mieux les applications de cette équation (obtenue quasiment que par des développement purement mathématique !!!) et qu'il devient alors possible de comprendre sa signification. A notre niveau du discours, il est inutile de dire quoi que ce soit. Il faut faire de la physique, et encore de la physique pour la comprendre et la voir apparaître.

Donc dans l'approche lagrangienne, nous apprenons à raisonner à partir de concepts d'énergie potentielle et cinétique, au lieu de concepts de force. Les deux approches sont évidemment équivalentes physiquement, mais les énergies n'étant pas des quantités vectorielles, elles sont conceptuellement plus faciles à utiliser dans une vaste gamme de problèmes. En physique quantique par exemple, la notion de force n'a aucune signification mais les notions d'énergie demeurent valables. C'est une raison de plus pour se familiariser avec leur utilisation. De plus, la force au sens de Newton est une action instantanée à distance. En relativité, une telle chose est impossible. La notion de force est donc une création purement classique et macroscopique contrairement à notre intuition, son intérêt est limité.

Exemple d'application (les autres exemples seront vus pendant notre étude des lois de Newton, de l'électrodynamique, de la relativité restreinte, de la relativité générale, de la physique quantique des champs, etc.):

Dans un premier temps, posons sous une forme mathématique conventionnelle l'équation d'Euler-Lagrange (la notation des coordonnées généralisées n'est pas identique en mathématiques à celle de la physique...):

$$\frac{\partial}{\partial x} \left(\frac{dZ(x, y, y')}{dy'} \right) - \frac{\partial Z(x, y, y')}{\partial y} = 0$$

Prenons un exemple mathématique pratique simple mondialement (voir universellement) connu et très important (nous réutiliserons les développements effectués ici pour l'étude du pendule de Huygens).

L'énoncé du problème est le suivant : déterminer quel est le plus court chemin entre deux points d'un plan (nous devinons que c'est la droite mais il faut le démontrer!).

Dans cet exemple quelque soit le chemin, nous avons un élément de celui-ci qui est donné par la relation de Pythagore dans \mathbb{R}^2 :

$$Z = \sqrt{dx^2 + dy^2} = \sqrt{1 + \frac{dy^2}{dx^2}} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} = \sqrt{1 + y'^2}$$

d'où la différentielle :

$$dZ = \frac{y'}{\sqrt{1 + y'^2}} dy'$$

Puisque Z n'est pas dépendant de y , nous avons :

$$\frac{\partial Z(x, y, y')}{\partial y} = 0$$

Dès lors, il vient que dans ce cas particulier que :

$$\frac{\partial}{\partial x} \left(\frac{dZ(x, y, y')}{dy'} \right) - \frac{\partial Z(x, y, y')}{\partial y} = 0$$

se réduit à :

$$\frac{\partial}{\partial x} \left(\frac{dZ(x, y, y')}{dy'} \right) = c^{te}$$

mais comme Z n'est pas pas non plus dépendant de x , cela signifie que nous avons :

$$\frac{dZ(x, y, y')}{dy'} = c^x$$

Et puisque nous cherchons un optimum :

$$\frac{dZ(x, y, y')}{dy'} = 0 \Rightarrow dZ(x, y, y') = 0$$

Si nous en calculons la primitive :

$$\int dZ(x, y, y') = \int 0 dZ \Rightarrow Z = c^x$$

Or cela semble en contradiction avec un résultat précédent :

$$dZ = \frac{y'}{\sqrt{1+y'^2}} dy'$$

Ceci nous amène donc à écrire :

$$\frac{dZ(x, y, y')}{dy'} = \frac{y'}{\sqrt{1+y'^2}} = 0 \Rightarrow dZ(x, y, y') = \frac{y'}{\sqrt{1+y'^2}} dy' = 0$$

Donc après intégration :

$$\sqrt{1+y'^2} = c^x \Rightarrow y' = \sqrt{(c^{te})^2 - 1} = c_1^{te}$$

Finalement la primitive devient:

$$\int y' dx = y(x) = c_1^{te} \cdot x + c_2^{te}$$

noté autrement en posant $c_1^{te} = a$ et $c_2^{te} = b$ nous obtenons :

$$y(x) = a \cdot x + b$$

Donc le plus court chemin entre deux points dans un plan est l'équation d'une droite... (simple mais puissant)

Remarque :

Si Z est une trajectoire optimale (extrémale), nous devons avoir dans l'espace des fonction la fonction Z telle que sa dérivée est nulle en tout point. Ce qui implique que $c^x = 0$. Dès lors, nous pouvons écrire :

$$\frac{dZ(x, y, y')}{dy'} = 0 \Rightarrow dZ(x, y, y') = 0$$

Nous avons aussi :

$$\frac{y'}{\sqrt{1+y'^2}} = 0 \Rightarrow y' = 0 \Rightarrow y = c^x$$

qui est un cas particulier de $y(x) = a \cdot x + b$ avec $a = 0$.

FORMALISME CANONIQUE

Le formalisme canonique n'introduit pas une nouvelle physique mais propose une nouvelle gamme d'outils pour étudier les phénomènes physiques. Son élément central, "le Hamiltonien", joue un grand rôle en physique quantique.

Comme dans le formalisme de Lagrange nous travaillerons avec des quantités comme l'énergie, T et V plutôt qu'avec des quantités vectorielles comme la force de Newton.

Dans le formalisme de Lagrange, la description d'un système mécanique à n degrés de liberté décrits par les coordonnées générales q_i indépendantes (non contraintes) nous mène à n équations d'Euler-Lagrange:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0 \Leftrightarrow \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0$$

qui sont des équations différentielles du 2ème ordre.

Dans le formalisme canonique (ou de Hamilton), un système mécanique à n degrés de liberté toujours décrits par des q_i indépendants nous mènera à $2n$ équations du premier ordre (plus simple à résoudre).

Chez Lagrange nous comparons principalement des trajectoires et par conséquent les q_i et les \dot{q}_i sont tous indépendants. Chez Hamilton nous devons d'abord apprendre à définir les "moments généralisés", notés p_i , pour remplacer les coordonnées généralisées \dot{q}_i et q_i qui sont aussi tous indépendants.

Remarque : l'origine des moments conjugués sera trivial dès que nous aurons vu un premier exemple concret.

TRANSFORMATION DE LEGENDRE

Cette transformation est souvent utilisée en thermodynamique où elle permet de relier entre eux les différents potentiels thermodynamiques. En mécanique elle permet de définir le hamiltonien à partir du lagrangien. Nous en donnons une description simplifiée et suffisante.

Soit une fonction $f(u, v)$ où u, v sont les deux variables indépendantes dont dépend f .

Définissons:

$$w = \frac{\partial f(u, v)}{\partial v} = w(u, v)$$

La transformation de Legendre permet de définir une fonction $g(u, w)$ qui peut remplacer $f(u, v)$:

$$f(u, v) \mapsto g(u, w) = v \cdot w - f$$

Soit maintenant la différentielle totale de f (voir chapitre de calcul différentiel et intégral) :

$$df = \frac{\partial f}{\partial u} du + \frac{\partial f}{\partial v} dv = \frac{\partial f}{\partial u} du + w dv$$

De la définition de g nous calculons:

$$\begin{aligned}
dg &= wdv + vdw - df \\
&= wdv + vdw - \frac{\partial f}{\partial u} du - wdv \\
&= vdw - \frac{\partial f}{\partial u} du = \frac{\partial g}{\partial w} dw + \frac{\partial g}{\partial u} du
\end{aligned}$$

et nous avons donc:

$$v = \frac{\partial g}{\partial w}, \quad \frac{\partial g}{\partial u} = -\frac{\partial f}{\partial u}$$

HAMILTONIEN

Soit un lagrangien $L(q_i, \dot{q}_i)$ que nous traiterons comme la fonction f ci-dessus avec les q_i jouant le rôle de u et les \dot{q}_i le rôle de v . A la place de w , nous définissons les moments généralisés également appelés "moments canoniques":

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i} = p_i(q_j, \dot{q}_j)$$

avec $i, j = 1, 2, \dots, n$.

Avant de continuer voyons ce que nous permet de faire cette définition :

Nous définissons donc, en analogie avec g , une fonction des q_i et des p_i que nous noterons $H(q_i, p_i)$:

$$\begin{array}{rcc}
g(u, w) & = & vw - f \\
\Downarrow & & \Downarrow \quad \Downarrow \\
H(q_i, p_i) & = & \sum_{i=1}^n \dot{q}_i p_i - L(q_i, \dot{q}_i)
\end{array}$$

Attention! La relation obtenue :

$$H(q_i, p_i) = \sum_{i=1}^n \dot{q}_i p_i - L(q_i, \dot{q}_i)$$

est plus que importante (comme tout le reste d'ailleurs). Nous la retrouverons, entre autres, en physique quantique des champs. Par ailleurs, un très joli exemple de tout ce que nous avons vu maintenant est donné dans le chapitre de mécanique relativiste ou nous calculons le lagrangien et hamiltonien d'une particule libre. Les résultats sont assez pertinents et leur utilité et justesse en électrodynamique plus que étonnante.

Si L dépend du temps (ce qui est quand même assez souvent le cas..) nous avons comme différentielle totale :

$$\begin{aligned}
dL &= \sum_{i=1}^n \frac{\partial L}{\partial q_i} dq_i + \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial L}{\partial t} dt \\
&= \sum_{i=1}^n \frac{\partial L}{\partial q_i} dq_i + \sum_{i=1}^n p_i d\dot{q}_i + \frac{\partial L}{\partial t} dt
\end{aligned}$$

nous calculons aussi la différentielle totale de $H(q_i, p_i)$ et y substituons le résultat obtenu précédemment :

$$dH = \sum_{i=1}^n \dot{q}_i dp_i + \sum_{i=1}^n p_i d\dot{q}_i - \left[\sum_{i=1}^n \frac{\partial L}{\partial q_i} dq_i + \sum_{i=1}^n p_i d\dot{q}_i + \frac{\partial L}{\partial t} dt \right]$$

$$= \sum_{i=1}^n \dot{q}_i dp_i - \sum_{i=1}^n \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial t} dt$$

ce qui montre bien que $H(q_i, p_i)$ est fonction des q_i (et du temps).

Nous pouvons donc aussi écrire pour sa différentielle totale :

$$dH = \sum_{i=1}^n \frac{\partial H}{\partial p_i} dp_i + \sum_{i=1}^n \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial t} dt$$

et comme les q_i et p_i sont indépendants nous identifions, en comparant nos deux expressions que:

$$\frac{\partial H}{\partial q_i} = - \frac{\partial L}{\partial q_i}$$

$$\frac{\partial H}{\partial p_i} = \dot{q}_i$$

$$\frac{\partial H}{\partial t} = - \frac{\partial L}{\partial t}$$

Ces relations sont extrêmement importantes car nous les retrouverons en physique quantique des champs sous une forme un peu plus barbare (mais magnifique aussi...).

Considérons maintenant le deuxième terme du premier membre de l'équation d'Euler-Lagrange. Nous avons :

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = \frac{d}{dt} (p_i) = \frac{\partial L}{\partial q_i} = \dot{p}_i$$

et ainsi, nous obtenons les $2n$ équations ci-dessous :

$$\left. \begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i}; i = 1, 2, \dots, n \\ \dot{p}_i &= - \frac{\partial H}{\partial q_i}; i = 1, 2, \dots, n \end{aligned} \right\} 2n$$

Ces $2n$ équations sont appelées "équations canoniques du mouvement" et sont des équations différentielles du premier ordre.

Remarque : l'apparition du signe moins " - " entre les équations pour les q_i et celles pour leurs moments conjugués, s'appelle une "symétrie symplectique".

De :

$$dH = \sum_{i=1}^n \frac{\partial H}{\partial p_i} dp_i + \sum_{i=1}^n \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial t} dt$$

nous pouvons, sur une trajectoire qui obéit aux équations canoniques, calculer:

$$\begin{aligned} \frac{dH}{dt} &= \sum_{i=1}^n \frac{\partial H}{\partial p_i} \dot{p}_i + \sum_{i=1}^n \frac{\partial H}{\partial q_i} \dot{q}_i + \frac{\partial H}{\partial t} \\ &= - \sum_{i=1}^n \dot{p}_i \dot{q}_i + \sum_{i=1}^n \dot{q}_i \dot{p}_i + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} = - \frac{\partial L}{\partial t} \end{aligned}$$

Remarque : si H ne dépend pas du temps nous avons alors $\partial H / \partial t = 0$, alors H (ainsi que L), sont une "constante du mouvement".

Un exemple s'avère indispensable à ce niveau d'avancement de l'étude formalisme Lagrangien. Nous allons nous restreindre à un cas particulier d'une particule soumis à une force en une dimension. Mais bien que cet exemple et les développements qui y sont liés soient simples nous retrouverons les résultats obtenus ici dans bien d'autres parties du site. Il est donc important de bien l'étudier et de bien le comprendre (ce qui nécessite malheureusement aussi que le contenu du chapitre de mécanique classique soit connu par le lecteur).

Exemple :

Soit une particule de masse m se déplaçant en une dimension (disons x) et soumise à une force dérivant d'un potentiel tel que:

$$F_x = - \frac{\partial V}{\partial x} \Leftrightarrow \ddot{x} = - \frac{1}{m} \frac{\partial V}{\partial x} = - (\nabla V)_x$$

Nous savons que son lagrangien est:

$$L = \frac{1}{2} m \dot{x}^2 - V(x)$$

Nous n'aurons qu'un seul moment (la quantité de mouvement), noté p , conjugué à x et défini par:

$$p = \frac{\partial L}{\partial \dot{x}}$$

équation que nous pouvons (que nous devons !) inverser (de la définition de la quantité de mouvement) :

$$\dot{x} = \frac{p}{m}$$

Nous pouvons noter en ce point le moment p correspond (ô hasard !!) à la composante x de la définition élémentaire $\vec{p} = m\vec{v}$ (ce qui ne sera pas toujours aussi trivialement le cas).

Selon la définition de l'Hamiltonien il vient alors :

$$H = \dot{x}(p)p - L(x, \dot{x}(p)) = \frac{p}{m} p - \frac{m}{2} \left(\frac{p}{m} \right)^2 + V(x) = \frac{p^2}{2m} + V(x)$$

que nous écrivons souvent sous la forme :

$$\boxed{H = T + V}$$

où T est donc l'énergie cinétique exprimée en fonction des moments.

[CROCHETS DE POISSON](#)

Le crochet de Poisson $\{A, B\}_{q,p}$ est la façon standard de noter une certaine opération qui implique les quantités $A(q_i, p_i)$ et $B(q_i, p_i)$ ainsi que l'ensemble des variables canoniques (q_i, p_i) définie par :

$$\boxed{\{A, B\}_{q,p} \equiv \sum_{i=1}^n \left[\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right]}$$

de cette définition nous pouvons déduire certaines propriétés triviales (si ce n'est pas le cas, n'hésitez pas à nous contacter, nous ajouterons les démonstrations) :

$$\begin{aligned} \{A, B\} &= \{B, A\} \\ \{A, B + C\} &= \{A, B\} + \{A, C\} \\ \{A, BC\} &= B\{A, C\} + \{A, B\}C \\ \{A, \{B, C\}\} + \{C, \{A, B\}\} + \{B, \{C, A\}\} &= 0 \end{aligned}$$

où la dernière expression est appelée "identité de Jacobi".

Au delà d'une simple notation, le calcul des crochets de Poisson est assez facile et permet d'obtenir nombre de résultats intéressants. D'autre part, ils sont intimement reliés aux "commutateurs" de la physique quantique que nous étudierons dans le détail dans le chapitre concerné.

Considérons maintenant une fonction quelconque $F(q_i, p_i, t)$ dont la dérivée totale par rapport au temps le long d'une trajectoire s'écrit (vous y reconnaîtrez quelque chose que vous connaissez déjà...) :

$$\frac{dF}{dt} = \sum_{i=1}^n \left[\frac{\partial F}{\partial q_i} \dot{q}_i - \frac{\partial F}{\partial p_i} \dot{p}_i \right] + \frac{\partial F}{\partial t}$$

Si cette trajectoire est une trajectoire physique, elle obéit aux équations canoniques de l'hamiltonien H du système:

$$\boxed{\dot{q}_i = \frac{\partial H}{\partial p_i}, \dot{p}_i = -\frac{\partial H}{\partial q_i}}$$

et alors:

$$\frac{dF}{dt} = \sum_{i=1}^n \left[\frac{\partial F}{\partial q_i} \frac{\partial H}{\partial p_i} + \frac{\partial F}{\partial p_i} \frac{\partial H}{\partial q_i} \right] + \frac{\partial F}{\partial t} = \{F, H\} + \frac{\partial F}{\partial t}$$

En particulier, cette équation permet un calcul facile des constantes du mouvement, $\partial F / \partial t = 0$. En effet, le calcul de $\partial f / \partial t$ est immédiat et le calcul de $\{F, H\}$ un assez simple exercice.

Il existe une famille de résultats intéressants des crochets de Poisson. Parmi les plus importants, calculons certains de ces crochets entre des variables canoniques, coordonnées et moments :

$$\{q_k, q_j\} = \sum_{i=1}^n \left[\frac{\partial q_k}{\partial q_i} \frac{\partial q_j}{\partial p_i} - \frac{\partial q_k}{\partial p_i} \frac{\partial q_j}{\partial q_i} \right]$$

puisque par définition, les coordonnées et moments ne sont pas directement dépendants :

$$\frac{\partial q_j}{\partial p_i} = 0, \frac{\partial q_k}{\partial p_i} = 0$$

d'où :

$$\{q_k, q_j\} = \{q_j, q_k\} = 0$$

et de manière identique :

$$\{p_k, p_j\} = \{p_j, p_k\} = 0$$

Mais :

$$\{q_k, p_j\} = \sum_{i=1}^n \left[\frac{\partial q_k}{\partial q_i} \frac{\partial p_j}{\partial p_i} - \frac{\partial q_k}{\partial p_i} \frac{\partial p_j}{\partial q_i} \right] = \sum_{i=1}^n \frac{\partial q_k}{\partial q_i} \frac{\partial p_j}{\partial p_i} = \sum_{i=1}^n \delta_{ki} \delta_{ji} = \delta_{kj}$$

où rappelons-le, δ_{ij} est le symbole de Kronecker défini par :

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Attention ! $\{q_k, p_j\}$ n'est pas commutatif. Effectivement, le lecteur contrôlera facilement que :

$$\{q_k, p_j\} \neq \{p_k, q_j\}$$

Ce qui implique un résultat assez général que nous retrouverons en physique quantique :

$$\{q_k, p_j\} - \{p_k, q_j\} \neq 0$$

TRANSFORMATIONS CANONIQUES

Nous disons des q_i, p_i que ce sont des "variables canoniques généralisées". Ce n'est pas un euphémisme puisqu'il n'y a pratiquement aucune limite à ce qu'elle peuvent représenter physiquement.

Puisque tel est le cas, il doit exister des transformations entre ces différents choix. Nous noterons Q_i, P_i les nouvelles variables canoniques obtenues suite à une telle transformation.

Nous ne sommes pas surpris par contre de constater que ces transformations sont soumises à des conditions assez sévères. En effet, les q_i, p_i sont généralisés et obéissent à :

$$\{q_k, q_j\} = 0 \quad \{p_k, p_j\} = 0 \quad \{q_k, p_j\} = \delta_{kj}$$

et les équations canoniques :

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

sont invariantes de forme. Ainsi, à la suite d'une transformation des q_i, p_i vers les Q_i, P_i et définissant un nouvel hamiltonien que nous noterons $K(Q_i, P_i)$ nous devons avoir :

$$\{Q_k, Q_j\} = 0 \quad \{P_k, P_j\} = 0 \quad \{Q_k, P_j\} = \delta_{kj}$$

et les équations canoniques :

$$\dot{Q}_i = \frac{\partial K}{\partial p_i} \quad \dot{P}_i = -\frac{\partial K}{\partial q_i}$$

Strictement, les équations de transformation peuvent s'écrire :

$$Q_i = Q_i(q_j, p_j, t)$$

$$P_i = P_i(q_j, p_j, t)$$

avec $i, j = 1, 2, \dots, n$ et doivent pouvoir s'inverser puisque la physique reste indépendante des variables que nous employons pour la décrire, donc nous pouvons écrire les transformations inverses :

$$q_i = q_i(Q_j, P_j, t)$$

$$p_i = p_i(Q_j, P_j, t)$$

avec $i, j = 1, 2, \dots, n$. Les q_i, p_i, Q_i, P_i forment $4n$ variables mais il est évident que seules $2n$ d'entre elles sont indépendantes.