

Partie I

Mécanique analytique

Introduction

La mécanique analytique n'apporte rien de conceptuellement nouveau par rapport aux formulations standard de la dynamique newtonienne (principe fondamental, théorème de l'énergie cinétique et autres points marquants de l'enseignement élémentaire de la mécanique), mais en constitue une formulation très élégante. Parfaitement adaptée à la description de systèmes où les mouvements sont sujets à des contraintes (un cauchemar avec les formulations "standard"), à l'utilisation de techniques de perturbations, ce qui explique son succès toujours certain auprès des astronomes, elle est souvent d'un usage infiniment plus pratique que les formulations plus élémentaires.

Il s'agit aussi d'un cas particulier d'une approche très fructueuse dans des domaines variés de la physique: une méthode variationnelle. En mécanique analytique, nous ne précisons pas les équations locales que doit vérifier à chaque instant le mouvement de la particule. Nous donnerons en fait une condition prescrivant à une intégrale portant sur l'ensemble du mouvement d'être extrémale. Parmi toutes les trajectoires permises par la cinématique, mais parfois absurdes pour la dynamique, il nous faudra choisir la bonne en respectant cette règle. En fait, la description du mouvement en mécanique analytique est très semblable à la description des rayons lumineux avec le principe de Fermat. Là aussi, on doit choisir parmi tous les trajets possibles celui qui rend extrémale une intégrale qui n'est autre que la durée du trajet.

Surtout, et bien qu'il s'agisse d'un formalisme datant, avec Lagrange et Hamilton, de la fin du XVIII^{ème} ou du XIX^{ème} siècle, elle est parfaitement adaptée aux approches modernes de la physique. Elle joue ainsi un rôle essentiel en mécanique statistique, elle est à l'origine de la quantification des dynamiques classiques, elle est fortement apparentée aux formulations modernes de la mécanique quantique en termes d'intégrales de chemin. Elle nous sera enfin d'une grande utilité pour reconstruire l'électromagnétisme à partir de la relativité.

Cette partie se compose de deux chapitres principaux. Dans le premier, qui sera le plus étoffé, nous donnons la formulation lagrangienne de la mécanique analytique, qui est celle que nous utiliserons dans la partie de relativité. Nous insisterons sur la notion de coordonnée généralisée, qui permet de traiter de façon naturelle les contraintes et nous examinerons comment on peut incorporer dans le formalisme un certain nombre d'interactions. Un point important dans ce domaine sera l'établissement de la fonction de Lagrange pour des particules chargées en interaction avec un champ, dont nous montrerons qu'elle redonne bien la force de Lorentz. Enfin, nous déduisons d'un certain nombre de symétries fondamentales de la nature (invariance dans le temps, dans l'espace, invariance par rotation) les lois de conservation essentielles (énergie, impulsion, moment cinétique). Cette approche qui lie les lois de conservation aux propriétés de symétrie est en fait très générale et très puissante.

Le deuxième chapitre sera consacré à une brève revue du formalisme hamiltonien. Brève parce que le sujet est extrêmement vaste, en particulier en ce qui concerne les transformations canoniques et les liens avec la mécanique quantique, brève aussi parce que le formalisme hamiltonien ne sera guère utilisé en grand détail dans les cours de première année.

Chapitre 1

Formulation lagrangienne

1.1 Description du système: coordonnées généralisées

Nous considérerons donc un système composé de N particules matérielles repérées par un indice grec α , variant de 1 à N . Une telle description peut convenir à tout système discret de particules ponctuelles mais aussi à la description du mouvement d'un solide, après une discrétisation convenable en éléments infinitésimaux. Les masses, charges électriques, positions, vitesses et accélérations des particules seront dénotées respectivement $m_\alpha, q_\alpha, \mathbf{r}_\alpha, \mathbf{v}_\alpha = \dot{\mathbf{r}}_\alpha, \mathbf{a}_\alpha = \dot{\mathbf{v}}_\alpha = \ddot{\mathbf{r}}_\alpha$ (nous désignerons souvent dans la suite les dérivées temporelles par des symboles pointés. Les caractères gras représentent des quantités vectorielles).

L'approche standard de la mécanique newtonienne est alors d'écrire le principe fondamental de la dynamique, reliant les accélérations des diverses particules constituant le système aux forces s'exerçant sur elles. L'expression de ces forces est donnée, en fonction de la configuration du système, soit par des lois fondamentales (force de Lorentz, par exemple), soit par des lois phénoménologiques (forces de frottement...). Par exemple, dans le cas de particules en interaction électromagnétique, on écrirait:

$$m_\alpha \mathbf{a}_\alpha = \mathbf{f}_\alpha = q_\alpha (\mathbf{E}(\mathbf{r}_\alpha) + \mathbf{v}_\alpha \times \mathbf{B}(\mathbf{r}_\alpha)), \quad (1.1)$$

où \mathbf{E} et \mathbf{B} sont les champs électrique et magnétique déterminés, en fonction des positions de particules, par la solution des équations de Maxwell.

Si l'écriture de toutes les équations dynamiques du système permet en principe, en y ajoutant les conditions initiales convenables, de déterminer complètement le mouvement, cette résolution peut être très délicate. C'est en particulier le cas quand il existe des contraintes: les positions (ou les vitesses) des particules doivent constamment obéir à un certain nombre de relations. Imaginons, par exemple, le cas de deux pendules accrochés l'un à l'extrémité de l'autre et contraints à se déplacer dans un plan (voir figure 1.1). Dans les formulations classiques, on doit associer à ces différentes liaisons des forces (force de tension des fils constituant les pendules, force de réaction du support commun...). Ces forces sont de nouvelles inconnues dans le problème qui doivent être déterminées en même temps que les variables dynamiques intéressantes. Bien entendu, elles compliquent beaucoup la résolution du problème.

L'idée de la mécanique analytique est de se débarrasser de ces forces inconnues en n'employant que des coordonnées indépendantes qui ne seront soumises à aucune contrainte. Nous les appellerons "coordonnées généralisées". Ces coordonnées sont de nature arbitraire (des positions, des angles...) mais doivent déterminer de façon univoque l'état mécanique du système si on prend en compte les contraintes. On pourra déterminer le mouvement en écrivant une équation différentielle pour chacune de ces coordonnées. Considérons, pour fixer les idées, le cas du double pendule. Il y a a priori six paramètres pour décrire le système (les positions des deux masses). En fait, les contraintes diminuent considérablement la dimensionnalité du problème. D'abord, les fils sont de longueur constante, soit

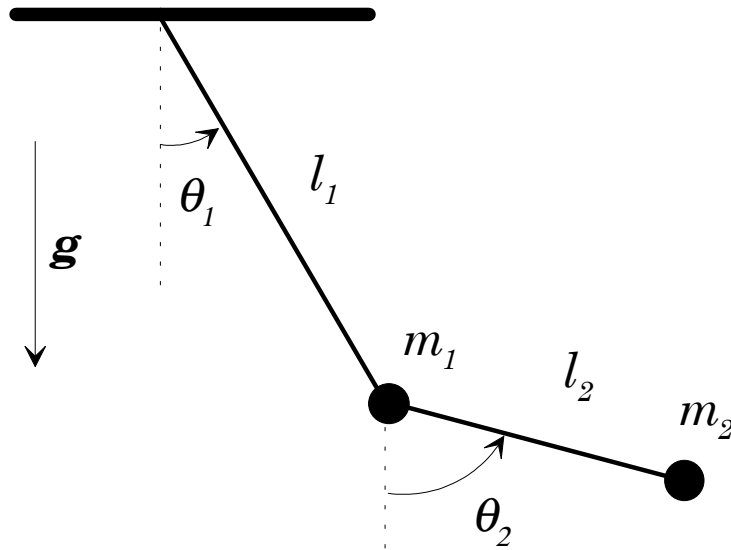


Figure 1.1: Un problème classique de mécanique: deux pendules liés, asservis à se déplacer dans un plan. Les angles θ_1 et θ_2 suffisent à décrire complètement l'état mécanique du système

deux relations. Ensuite, le mouvement s'effectue dans un plan, ce qui fournit encore deux relations (par exemple en écrivant que le produit scalaire de la position avec la normale au plan est nul). Il n'y a donc en fait que deux variables indépendantes qui décrivent le mouvement. Un choix tout à fait naturel de coordonnées généralisées dans ce cas est de prendre les deux angles θ_1 et θ_2 des pendules avec la verticale.

Plus généralement, nous supposons que les liaisons entre les simples coordonnées cartésiennes des particules sont holonomes: il existe $3N - n$ relations du type $f_j(\mathbf{r}_\alpha) = 0$. De telles relations décrivent convenablement toutes les contraintes directes entre coordonnées, à condition qu'elles soient indépendantes du temps (comme celles que nous venons de voir, si la longueur des fils des pendules est invariable)¹. Elles ne décrivent pas les contraintes entre vitesses (par exemple le roulement sans glissement) mais nous verrons plus loin comment on peut en tenir compte. Il ne reste alors que n coordonnées généralisées indépendantes que nous noterons q_i , $i = 1 \dots n$. Soulignons une fois de plus que ces coordonnées ne sont pas nécessairement cartésiennes et n'ont même pas forcément la dimension d'une longueur. Avec des relations holonomes, les positions \mathbf{r}_α ne dépendent que des n coordonnées généralisées et ne dépendent ni des vitesses ni du temps explicitement. Il nous faut maintenant donner les lois permettant d'établir les n équations différentielles déterminant la dynamique des q_i .

¹En fait, la plupart des résultats que nous établirons dans ce chapitre seraient également valables si les relations faisaient intervenir une dépendance explicite en temps, sous la forme $f_j(\mathbf{r}_\alpha, t) = 0$. Les positions dépendraient alors des coordonnées généralisées, mais présenteraient aussi une dépendance explicite en temps. L'énergie cinétique, par exemple, qui est une forme quadratique des dérivées des coordonnées généralisées dans le cas habituel, ferait intervenir des termes linéaires dans ces dérivées, ou même des termes n'en dépendant pas. Nous préciserons, si nécessaire, quels sont les résultats qui dépendent de façon critique de cette hypothèse.

1.2 Principe de moindre action

1.2.1 Énoncé

On postule qu'il existe une fonction $L(q_i, \dot{q}_i, t)$, dite fonction de Lagrange ou lagrangien, homogène à une énergie², qui est telle que l'action

$$S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t) dt, \quad (1.2)$$

soit *extrémale* pour la trajectoire effectivement suivie par le système de t_1 à t_2 entre $q_i(1)$ et $q_i(2)$, valeurs initiales et finales des coordonnées généralisées.

Ce que nous postulons ainsi n'est pas directement un ensemble d'équations différentielles pour les variables dynamiques (que nous pourrions déduire et dont nous montrerons qu'elles sont équivalentes aux formulations standard de la mécanique). Nous nous donnons plutôt un principe variationnel qui postule le caractère extrémal d'une certaine intégrale calculée sur la trajectoire en fonction de celle-ci. Il y a de nombreux autres exemples de principes variationnels en physique. Les lois de l'optique géométrique, par exemple, peuvent se déduire du principe de Fermat qui postule que le rayon lumineux effectivement suivi réalise un extremum (en général un minimum) du temps de parcours.

Le fait que nous prenions une fonction de Lagrange ne dépendant que des positions et des vitesses (mais pas de dérivées d'ordre supérieur) exprime, comme nous le verrons, que les équations fondamentales de la dynamique sont d'ordre deux par rapport au temps. D'autre part, nous spécifions les deux conditions "initiales" nécessaires pour chaque coordonnée en donnant les positions initiales et finales et non les positions et vitesses initiales. Si ces deux formulations sont bien sûr équivalentes, la première est plus avantageuse pour varier l'action sur toutes les trajectoires possibles entre deux points.

Bien sûr, un principe variationnel est d'emploi moins commode en pratique qu'un ensemble d'équations différentielles. Il faut, en principe, imaginer toutes les trajectoires possibles (continues et dérivables) entre les conditions initiales et finales, déterminer l'action sur chacune et déterminer celles qui rendent l'action extrémale. Nous verrons, dans le prochain paragraphe, comment en déduire un système d'équations différentielles beaucoup plus commodes.

1.2.2 Equations de Lagrange

Nous considérons donc deux trajectoires possibles entre $q(1)$ et $q(2)$. L'une, que nous noterons simplement $q_i(t)$, est la trajectoire effectivement suivie. L'autre que nous appellerons "trajectoire variée", infiniment proche, correspond à chaque instant aux positions $q_i(t) + \delta q_i(t)$, où $\delta q_i(t)$ est un accroissement infinitésimal de la position (voir figure 1.2). Ces deux trajectoires doivent obéir aux mêmes conditions initiales et finales. On a donc $\delta q(1) = \delta q(2) = 0$. Nous supposons que les q_i et δq_i sont deux fois différentiables. Le fait que les q_i donnent la trajectoire effectivement suivie a pour conséquence que l'action S est extrémale sur cette trajectoire et ne varie donc pas au premier ordre dans les δq_i quand on passe à la trajectoire variée. Or la variation de l'action s'écrit simplement:

$$\delta S = \int_{t_1}^{t_2} (L(q_i + \delta q_i, \dot{q}_i + \dot{\delta q}_i, t) - L(q_i, \dot{q}_i, t)) dt. \quad (1.3)$$

En développant L au premier ordre dans les δq_i , on a:

$$\delta S = \sum_i \int_{t_1}^{t_2} \frac{\partial L}{\partial q_i} \delta q_i dt + \sum_i \int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_i} \frac{d\delta q_i}{dt} dt. \quad (1.4)$$

²et que l'on saura écrire si on connaît la nature des forces qui s'exercent sur le système

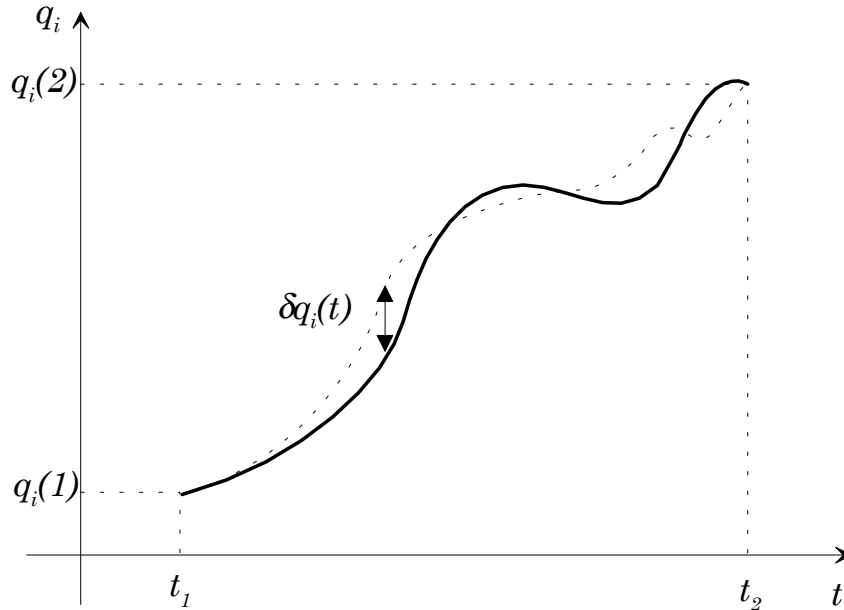


Figure 1.2: Trajectoire effectivement suivie (ligne continue) et trajectoire variée (pointillée). La trajectoire variée s'écarte infinitésimalement de la trajectoire effectivement suivie et coïncide avec celle-ci aux extrémités.

Pour fixer la trajectoire, nous recherchons une condition sur les δq_i . Pour éliminer leurs dérivées temporelles dans l'expression précédente nous intégrons les termes de la seconde somme par parties. On obtient alors:

$$\delta S = \sum_i \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right] \delta q_i dt + \left[\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right]_{t_1}^{t_2}. \quad (1.5)$$

Les accroissements infinitésimaux δq_i s'annulant aux extrémités de la trajectoire, le terme tout intégré est identiquement nul. La somme, elle, ne peut s'annuler pour des δq_i arbitraires (et indépendants) que si les n équations différentielles:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0 \quad (1 \leq i \leq n) \quad (1.6)$$

sont simultanément vérifiées. Ce système différentiel avec les conditions aux limites (fournissant $2n$ conditions indépendantes) détermine complètement les n coordonnées généralisées indépendantes. Ces équations différentielles sont du second ordre par rapport au temps puisque L dépend a priori des \dot{q}_i . Quelques remarques s'imposent à ce point.

- Les équations du mouvement ne changent pas quand L est multipliée par une constante. Cette liberté correspond seulement à un choix d'unités (L a, rappelons-le, la dimension d'une énergie).
- La structure de L doit obéir aux symétries du système physique (invariance par translation dans le temps, dans l'espace...). Nous verrons, dans les prochains paragraphes, que cela conduit à des conséquences importantes en termes de lois de conservation.
- Les équations du mouvement sont inchangées si on ajoute à L la dérivée totale par rapport au temps d'une fonction des coordonnées et du temps. Posons en effet $L' = L + df(q_i, t)/dt$ (la fonction f que nous dérivons ne doit dépendre que des q_i pour que le lagrangien modifié ne dépende que des q_i et \dot{q}_i). L'action S' , calculée avec la nouvelle fonction de Lagrange, ne diffère de S que par un terme de la forme $[f]_{t_1}^{t_2}$, qui ne dépend manifestement pas de la trajectoire suivie entre 1 et 2. On pourra vérifier par simple substitution que les équations de Lagrange ne

changent pas quand on effectue cette modification du lagrangien. Cette liberté (qui n'est pas sans évoquer la liberté de jauge en électromagnétisme) est parfois fort utile pour simplifier la forme du lagrangien.

- La fonction de Lagrange est additive. Considérons en effet deux systèmes physiques indépendants et n'interagissant pas, décrits par les coordonnées généralisées q_1 et q_2 et par les fonctions de Lagrange L_1 et L_2 . Le système global est défini par la réunion des deux ensembles de coordonnées généralisées et il est évident que la fonction de Lagrange associée au système complet est $L = L_1 + L_2$.
- Notons enfin que, dans le cas très simple où les coordonnées généralisées coïncident avec les coordonnées cartésiennes des particules individuelles (ce qui est par exemple le cas quand il n'y a aucune contrainte sur le mouvement), les équations de Lagrange peuvent s'écrire:

$$\nabla_{\mathbf{r}_\alpha} L = \frac{d\nabla_{\mathbf{v}_\alpha} L}{dt}, \quad (1.7)$$

en faisant intervenir les opérateurs gradient par rapport aux positions et vitesses de chaque particule.

Il nous reste maintenant, pour que ce formalisme ait un sens, à donner la forme de la fonction de Lagrange en fonction des interactions que subissent les particules.

1.3 Expressions de la fonction de Lagrange

Cette section est essentielle dans ce chapitre, puisqu'elle nous permettra de traiter effectivement des problèmes de mécanique par le formalisme lagrangien. Nous nous pencherons d'abord sur le cas de la particule libre, puis sur le cas de particules interagissant par des forces conservatives (dérivant d'un potentiel), sur le cas de particules soumises à des forces extérieures au système (avec quelques applications au cas important du mouvement dans le champ de pesanteur) et enfin sur le cas de particules en interaction avec un champ électromagnétique.

1.3.1 Particule unique libre

Ce cas élémentaire n'a de mérite que pédagogique. Il est en effet évident dès l'abord que toute fonction de Lagrange conduisant à des équations s'écrivant $\dot{\mathbf{v}} = 0$ conviendra. Une simple fonction proportionnelle à v^2 vérifie (entre autres) cette propriété. La fonction de Lagrange d'une particule libre unique est donc, à un choix d'unités près, identique à l'énergie cinétique. Nous allons toutefois montrer comment on peut arriver à ce résultat en utilisant les propriétés d'invariance et de symétrie.

Les coordonnées généralisées coïncident dans ce cas avec les coordonnées cartésiennes standard. La particule étant libre, L ne peut explicitement dépendre du temps (invariance par translation dans le temps), ni de la position \mathbf{r} de la particule (invariance par translation spatiale), ni enfin de la direction de sa vitesse \mathbf{v} (invariance par rotation). L doit donc être une fonction du carré du module de la vitesse: $L = f(v^2)$.

Les équations de Lagrange se résument alors à:

$$\nabla_{\mathbf{r}} L = 0 = \frac{d}{dt} \left[\frac{df}{dv^2} \nabla_{\mathbf{v}} v^2 \right], \quad (1.8)$$

qui conduisent bien évidemment, à moins que f ne soit une constante, à $\dot{\mathbf{v}} = 0$ et donc à un mouvement rectiligne uniforme.

Pour préciser davantage la forme de L et de f , il nous faut ajouter une condition supplémentaire: le résultat précédent doit être invariant dans un changement de référentiel galiléen. Considérons pour

cela un référentiel \mathcal{R} , dans lequel la fonction de Lagrange est $L = f(v^2)$ et un référentiel \mathcal{R}' en translation uniforme par rapport à \mathcal{R} avec une vitesse infinitésimale $\boldsymbol{\varepsilon}$. La fonction de Lagrange L' dans \mathcal{R}' doit s'écrire $L' = f(v'^2) = f((\mathbf{v} + \boldsymbol{\varepsilon})^2)$ (la fonction f devant manifestement être la même pour tous les référentiels). En développant au premier ordre en $\boldsymbol{\varepsilon}$, on a: $L' = L + (df/dv^2)2\mathbf{v} \cdot \boldsymbol{\varepsilon}$. Les équations du mouvement dans \mathcal{R} et \mathcal{R}' coïncideront si L et L' ne diffèrent que par une dérivée totale par rapport au temps. Le cas le plus simple où cela se vérifie est quand df/dv^2 est indépendant de v (la dérivée par rapport au temps étant alors simplement $d(2(df/dv^2)\mathbf{r} \cdot \boldsymbol{\varepsilon})/dt$). Le choix le plus simple est donc que f soit simplement proportionnelle à v^2 . Nous poserons donc:

$$L = \frac{1}{2}mv^2 \quad (1.9)$$

et appellerons évidemment “masse de la particule” le coefficient m . Ce coefficient doit être positif. En effet, l'extremum de l'action correspondant à la propagation en ligne droite à vitesse constante est alors un minimum.

Bien sûr, ce raisonnement n'est pas d'une grande rigueur et repose largement sur le critère de simplicité pour identifier complètement la forme du lagrangien. Il illustre en revanche le genre de démarche qu'on doit effectuer pour déterminer la forme du lagrangien correspondant à une nouvelle interaction: respecter d'abord les grandes propriétés de symétrie, respecter les règles de la relativité galiléenne et enfin chercher la forme la plus simple en cas d'ambiguïté. C'est, avec quelques adaptations, la démarche que nous utiliserons plus tard pour déterminer, en relativité, les lagrangiens correspondant à l'interaction électromagnétique.

1.3.2 Système de particules interagissant par des forces dérivant d'un potentiel

Nous supposons d'abord, pour fixer les idées, que les coordonnées généralisées coïncident avec les coordonnées cartésiennes. Nous supposons que la force s'exerçant sur la particule α peut s'écrire $\mathbf{F}_\alpha = -\nabla_{\mathbf{r}_\alpha} U(\mathbf{r}_1, \dots, \mathbf{r}_n, t)$, où la fonction U est une énergie potentielle dépendant a priori de la position de toutes les particules dans le système. Pour ne pas restreindre la généralité, nous permettrons au potentiel de dépendre explicitement du temps. Nous pourrions ainsi traiter, par exemple, le mouvement dans des champs extérieurs variables. Nous chercherons simplement la forme de la fonction de Lagrange qui redonne les équations dynamiques habituelles.

Si U est identiquement nulle, la fonction L se résume à l'énergie cinétique totale: $L = T$ avec $T = \sum_\alpha (1/2)m_\alpha v_\alpha^2$ (ce qui se déduit évidemment du paragraphe précédent et de l'additivité de la fonction de Lagrange pour des systèmes sans interaction mutuelle). Nous vérifierons maintenant que le choix $L = T - U$ donne, pour les équations de Lagrange, les équations standard. En effet, les équations de Lagrange s'écrivent:

$$\nabla_{\mathbf{r}_\alpha} L = \frac{d\nabla_{\mathbf{v}_\alpha} L}{dt}, \quad (1.10)$$

et on a

$$\nabla_{\mathbf{r}_\alpha} L = -\nabla_{\mathbf{r}_\alpha} U = \mathbf{F}_\alpha, \quad (1.11)$$

et

$$\frac{d\nabla_{\mathbf{v}_\alpha} L}{dt} = \frac{dm_\alpha \mathbf{v}_\alpha}{dt} = m_\alpha \mathbf{a}_\alpha = \mathbf{F}_\alpha. \quad (1.12)$$

Cette forme du lagrangien redonne donc bien le principe fondamental de la dynamique tel que nous le connaissions.

Considérons maintenant le cas où le système doit être décrit par des coordonnées généralisées qui ne coïncident pas avec les coordonnées cartésiennes. La fonction de Lagrange elle-même ne doit pas dépendre du choix particulier du système de coordonnées généralisées. Elle doit toujours coïncider avec la différence $T - U$ des énergies cinétiques et potentielles. Pour pouvoir écrire les équations de Lagrange, il faut exprimer ces quantités en fonction des coordonnées généralisées q_i et de leurs dérivées. C'est

toujours possible, puisque les q_i doivent déterminer de façon univoque l'état mécanique du système – quand on prend en compte les contraintes. En inversant ces relations et en les reportant dans les expressions de T et U en fonction des coordonnées cartésiennes on obtient facilement le résultat cherché.

Nous avons supposé que les positions \mathbf{r}_α ne dépendent que des q_i , mais pas de leurs dérivées ni du temps (ce ne serait pas le cas si les liaisons faisaient intervenir une dépendance explicite en temps). On peut donc écrire: $\mathbf{v}_\alpha = \sum_i \dot{q}_i \partial \mathbf{r}_\alpha / \partial q_i$. En substituant dans $T = \sum_\alpha m_\alpha \mathbf{v}_\alpha^2 / 2$, on trouve en général T comme une forme quadratique définie positive des \dot{q}_i , dont les coefficients peuvent dépendre des q_i : $T = \sum_{i,j} A_{i,j}(q_k) \dot{q}_i \dot{q}_j$. Par exemple, on a une expression de ce genre quand on utilise les coordonnées cylindriques pour décrire le mouvement d'une particule unique $T = m(\dot{r}^2 + r^2 \dot{\theta}^2 + \dot{z}^2) / 2$. Notons enfin que, dans les mêmes conditions, U s'exprime simplement comme une fonction des q_i .

1.3.3 Système de particules soumises à des forces extérieures

C'est un cas particulièrement important en mécanique, puisqu'il permet, entre autres, de traiter de mouvements dans le champ de gravitation terrestre. Nous considérons donc un système S dont la dynamique est décrite par n coordonnées généralisées q_i et dont nous cherchons le mouvement. Les particules de S interagissent entre elles par des forces dérivant d'une énergie potentielle et interagissent avec les particules d'un système extérieur \mathcal{S} , décrit par \mathcal{N} coordonnées généralisées Q_j . Nous supposons que les coordonnées généralisées sont bien séparées: l'état de S est complètement déterminé par les q_i seuls. Nous supposons le système \mathcal{S} suffisamment "gros" pour que l'interaction avec S ait une influence négligeable sur sa dynamique (c'est bien sûr le cas pour tout mouvement réaliste dans le champ de pesanteur terrestre). Nous pouvons alors considérer les Q_j comme des données du problème.

La fonction de Lagrange du système $S + \mathcal{S}$ s'écrit $L = T_S + T_{\mathcal{S}} - U(q_i; Q_j)$, où U est l'énergie potentielle dont dérivent toutes les forces du problème (y compris l'interaction entre S et \mathcal{S}), T_S et $T_{\mathcal{S}}$ les énergies cinétiques des deux systèmes. Comme nous avons supposé que les deux ensembles de coordonnées généralisées sont bien séparés, T_S n'est fonction que des q_i et de leurs dérivées et $T_{\mathcal{S}}$ n'est fonction que des seuls Q_j et de leurs dérivées. $T_{\mathcal{S}}$ peut donc être considérée comme une fonction donnée du temps. Une telle fonction étant aussi une dérivée totale par rapport au temps, elle ne joue aucun rôle dans les équations de Lagrange et peut être supprimée. De la même manière, U peut être écrite comme une fonction des q_i seuls et du temps (la dépendance en temps reflétant la dynamique des Q_j). On a finalement: $L = T_S(q_i, \dot{q}_i) - U(q_i, t)$, lagrangien décrivant la dynamique du système S seul.

Ce type de problème étant très fréquemment rencontré en mécanique, nous allons l'illustrer par deux exemples.

Premier exemple

Cet exemple, particulièrement trivial, ne présente guère que l'intérêt d'appliquer les notions introduites dans les paragraphes précédents dans une situation où la mise en équations et les calculs ne présentent aucune difficulté technique. Nous considérerons donc la "machine d'Atwood", pont aux ânes des classes élémentaires (voir figure 1.3). Deux masses m_1 et m_2 , astreintes à se déplacer verticalement dans le champ de pesanteur (accélération \mathbf{g}), pendent aux deux extrémités d'une ficelle passant sur une poulie (tout cela étant inextensible, sans frottements...). Une seule coordonnée généralisée suffit à décrire ce problème en tenant compte des liaisons (mouvement selon la verticale, longueur de ficelle constante). On prendra la position x de la masse m_1 le long d'un axe vertical descendant. La vitesse de m_1 est \dot{x} , celle de m_2 , $-\dot{x}$. On a alors $T = (m_1 + m_2)\dot{x}^2 / 2$ et $U = -m_1 g x - m_2 g(\ell - x)$, où ℓ est une constante. A une constante additive près on a donc $U = (m_2 - m_1)g x$ et

$$L = \frac{1}{2}(m_1 + m_2)\dot{x}^2 - (m_2 - m_1)g x . \quad (1.13)$$

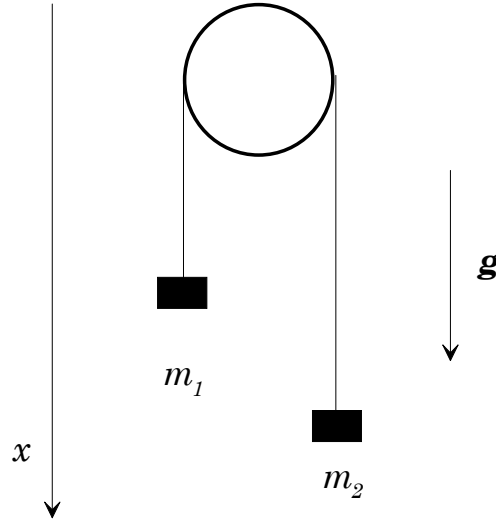


Figure 1.3: La “machine d’Atwood”, exemple élémentaire de mouvements avec liaisons dans un champ extérieur. Deux masses différentes sont reliées par une corde inextensible passant sur une poulie. Le mouvement des masses est uniformément accéléré.

L’unique équation de Lagrange s’écrit alors trivialement:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = (m_1 + m_2)\ddot{x} = \frac{\partial L}{\partial x} = -(m_2 - m_1)g, \quad (1.14)$$

d’où on déduit bien évidemment un mouvement uniformément accéléré avec l’accélération $(m_1 - m_2)g/(m_1 + m_2)$

Pour un exemple aussi trivial, l’écriture du formalisme lagrangien n’apporte pas de simplification décisive. Il eut été aussi simple d’écrire, pour ce problème à un seul degré de liberté, le théorème de l’énergie cinétique. En revanche, l’écriture du principe fondamental de la dynamique nous aurait contraint à faire intervenir des inconnues supplémentaires: les tensions des fils. La puissance de l’écriture lagrangienne ne peut s’exprimer que sur des problèmes un peu plus complexes.

Deuxième exemple

Nous considérerons dans ce paragraphe l’exemple du double pendule (voir figure 1.1). Notons tout de suite que la mise en équations de ce problème est très pénible par les techniques standard de mécanique. Il faut en effet faire intervenir la tension des fils, dont on ne se débarrasse qu’au prix de manipulations fastidieuses. Notons aussi que, pour ce problème à deux degrés de liberté, le théorème de l’énergie cinétique ne nous est d’aucune utilité.

Nous allons écrire les équations de Lagrange. Les deux coordonnées généralisées q_1 et q_2 coïncident avec les angle θ_1 et θ_2 que les pendules font avec la verticale. L’écriture des positions et vitesses des deux masses en fonction des coordonnées généralisées et des longueurs l_1 et l_2 des deux pendules, ne présente aucune difficulté. On en déduit les expressions des énergies cinétiques et potentielles:

$$\begin{aligned} T &= T_1 + T_2 \\ T_1 &= \frac{1}{2}m_1 l_1^2 \dot{\theta}_1^2 \end{aligned} \quad (1.15)$$

$$T_2 = \frac{1}{2}m_2 \left[l_1^2 \dot{\theta}_1^2 + l_2^2 \dot{\theta}_2^2 + 2l_1 l_2 \cos(\theta_1 - \theta_2) \dot{\theta}_1 \dot{\theta}_2 \right] \quad (1.16)$$

$$\begin{aligned} U &= U_1 + U_2 \\ U_1 &= -m_1 g l_1 \cos(\theta_1) \end{aligned} \quad (1.17)$$

$$U_2 = -m_2 g l_2 \cos(\theta_2) - m_2 g l_1 \cos(\theta_1), \quad (1.18)$$

d'où on tire l'expression de la fonction de Lagrange L . L'écriture des deux équations de Lagrange ne pose alors d'autre difficulté qu'algébrique. La mise en équations de ce problème, pratiquement impossible en appliquant simplement le principe fondamental de la dynamique, ne présente aucune difficulté avec l'approche lagrangienne.

La résolution explicite du système différentiel obtenu est toutefois impossible. Il s'agit en effet d'équations non linéaires couplées (à cause des termes en $\cos\theta$). On ne peut obtenir de solutions explicites, comme dans tous les problèmes de pendules, que dans le cadre d'une approximation linéaire, valide dans le domaine des petites oscillations. Nous donnons ici seulement les étapes essentielles du calcul, dont les détails seront facilement retrouvés par le lecteur. Les équations de Lagrange pour θ_1 et θ_2 s'écrivent respectivement après linéarisation:

$$(m_1 + m_2)l_1\ddot{\theta}_1 + m_2l_2\ddot{\theta}_2 + (m_1 + m_2)g\theta_1 = 0 \quad (1.19)$$

$$l_2\ddot{\theta}_2 + l_1\ddot{\theta}_1 + g\theta_2 = 0. \quad (1.20)$$

Un tel système de deux équations linéaires à coefficients constants se résout par les méthodes standard. Les solutions sont des combinaisons linéaires de solutions en $\exp(-i\omega t)$. Les valeurs possibles de ω sont celles qui annulent le déterminant caractéristique:

$$\begin{vmatrix} (m_1 + m_2)(g - l_1\omega^2) & -m_2l_2\omega^2 \\ -l_1\omega^2 & g - l_2\omega^2 \end{vmatrix}. \quad (1.21)$$

On obtient alors une équation bicarrée en ω . Les solutions s'écrivent $\pm\omega_1$ et $\pm\omega_2$. La solution générale est alors une superposition de deux mouvements oscillatoires aux fréquences ω_1 et ω_2 , de phase et d'amplitudes arbitraires. L'écriture explicite de ces phases et amplitudes en fonction des conditions initiales est pénible, mais sans difficultés.

La recherche des deux fréquences ω_1 et ω_2 est un exemple particulièrement simple de recherches de "modes propres" dans des systèmes d'oscillateurs couplés. Très généralement, un système d'oscillateurs linéaires couplés possédant p degrés de liberté, admet p fréquences propres. Chacune de ces fréquences correspond à une configuration particulière des mouvements des degrés de liberté, conduisant à un mouvement purement harmonique (on pensera par exemple aux modes d'oscillations symétriques et antisymétriques de deux pendules couplés par un ressort). Dans le cas présent, il est instructif d'examiner les comportements asymptotiques des deux modes quand $m_1 \rightarrow \infty$ ou $m_2 \rightarrow 0$, ce que nous laissons au lecteur.

1.3.4 Lagrangien de particules chargées dans un champ

Nous abordons ici un exemple très important d'écriture de fonction de Lagrange, d'abord en raison de son utilité pratique, ensuite en raison de l'usage que nous en ferons dans la suite du cours d'électromagnétisme. Nous utiliserons une démarche très pragmatique, en cherchant la forme la plus simple de la fonction de Lagrange qui redonne la force de Lorentz.

Nous considérerons donc un ensemble de N particules, que nous supposerons décrites par leurs coordonnées cartésiennes standard (la généralisation à d'autres systèmes de coordonnées ne posant que des problèmes algébriques) \mathbf{r}_α . Ces particules, chargées, sont placées dans un champ électrique \mathbf{E} et un champ magnétique \mathbf{B} , imposés de l'extérieur. Pour alléger les notations, nous poserons $\mathbf{E}_\alpha = \mathbf{E}(\mathbf{r}_\alpha, t)$ et $\mathbf{B}_\alpha = \mathbf{B}(\mathbf{r}_\alpha, t)$. Nous introduirons enfin les potentiels vecteur $\mathbf{A}_\alpha = \mathbf{A}(\mathbf{r}_\alpha, t)$ et scalaire $V_\alpha = V(\mathbf{r}_\alpha, t)$ permettant de calculer les champs par les relations: $\mathbf{E}_\alpha = -\partial\mathbf{A}_\alpha/\partial t - \nabla V_\alpha$ et $\mathbf{B}_\alpha = \nabla \times \mathbf{A}_\alpha$, où les dérivées spatiales sont à prendre par rapport à la position de la particule α . Notons que la dérivée temporelle du potentiel vecteur est une dérivée partielle. Le champ électromoteur ne fait en effet intervenir que la dépendance explicite en temps du champ magnétique.

Le principe fondamental de la dynamique pour la particule α s'écrit alors:

$$m\ddot{\mathbf{r}}_\alpha = \mathbf{F}_\alpha = q_\alpha(\mathbf{E}_\alpha + \mathbf{v}_\alpha \times \mathbf{B}_\alpha). \quad (1.22)$$

Nous allons transformer cette équation pour la mettre sous une forme qui rappelle celle d'une équation de Lagrange, ce qui nous permettra d'intuiter une forme simple pour la fonction de Lagrange de ce problème.

En faisant intervenir l'expression des champs en fonction des potentiels, on peut écrire:

$$\mathbf{F}_\alpha = q_\alpha \left(-\frac{\partial \mathbf{A}_\alpha}{\partial t} + \mathbf{v}_\alpha \times (\nabla \times \mathbf{A}_\alpha) - \nabla V_\alpha \right). \quad (1.23)$$

En place de la dérivée partielle par rapport au temps de \mathbf{A}_α , on souhaite faire apparaître une dérivée totale par rapport au temps (rappelons que, dans les équations de Lagrange, les dérivées par rapport au temps sont totales). La variation de \mathbf{A}_α par rapport au temps provient de deux causes. D'abord, il peut exister une variation explicite du potentiel vecteur par rapport au temps (exprimée ci-dessus par la dérivée partielle par rapport au temps). Même en l'absence de cette dépendance explicite en temps, le potentiel vecteur "vu" par une particule α varie en raison du mouvement de la particule dans le champ inhomogène. La variation correspondante peut s'écrire simplement $(\mathbf{v}_\alpha \cdot \nabla) \mathbf{A}_\alpha$, terme parfois nommé "dérivée hydrodynamique". On a alors:

$$\frac{d\mathbf{A}_\alpha}{dt} = \frac{\partial \mathbf{A}_\alpha}{\partial t} + (\mathbf{v}_\alpha \cdot \nabla) \mathbf{A}_\alpha, \quad (1.24)$$

et donc

$$\mathbf{F}_\alpha = q_\alpha \left(-\frac{d\mathbf{A}_\alpha}{dt} + (\mathbf{v}_\alpha \cdot \nabla) \mathbf{A}_\alpha + \mathbf{v}_\alpha \times (\nabla \times \mathbf{A}_\alpha) - \nabla V_\alpha \right). \quad (1.25)$$

On reconnaît dans les deux termes centraux deux des termes du développement de $\nabla_{\mathbf{r}_\alpha} (\mathbf{v}_\alpha \cdot \mathbf{A}_\alpha)$. Les deux autres termes de ce développement qui font intervenir des dérivées partielles de \mathbf{v}_α par rapport à la position de la particule α sont manifestement nuls. On obtient donc finalement:

$$\mathbf{F}_\alpha = q_\alpha \left(-\frac{d\mathbf{A}_\alpha}{dt} - \nabla [V_\alpha - \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha] \right). \quad (1.26)$$

En réécrivant alors le premier terme de (1.22) sous la forme:

$$\frac{d}{dt} \nabla_{\mathbf{v}_\alpha} \frac{1}{2} m_\alpha \mathbf{v}_\alpha^2, \quad (1.27)$$

et en remarquant que

$$\frac{d\mathbf{A}_\alpha}{dt} = \frac{d}{dt} \nabla_{\mathbf{v}_\alpha} \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha \quad (1.28)$$

on peut mettre le principe fondamental sous la forme:

$$\begin{aligned} \frac{d}{dt} \nabla_{\mathbf{v}_\alpha} \left[\frac{1}{2} m_\alpha \mathbf{v}_\alpha^2 + q_\alpha \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha \right] &= -q_\alpha \nabla_{\mathbf{r}_\alpha} [V_\alpha - \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha] \\ &= \nabla_{\mathbf{r}_\alpha} \left[\frac{1}{2} m_\alpha \mathbf{v}_\alpha^2 - q_\alpha (V_\alpha - \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha) \right], \end{aligned} \quad (1.29)$$

le gradient par rapport à la position de l'énergie cinétique étant évidemment nul. De la même manière, on peut ajouter dans le gradient du membre de gauche un terme proportionnel au potentiel, $q_\alpha V_\alpha$, qui ne dépend manifestement pas des vitesses. On obtient alors la forme de l'équation de Lagrange en coordonnées cartésiennes, à condition de poser:

$$L = \sum_\alpha [T_\alpha - q_\alpha V_\alpha + q_\alpha \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha]. \quad (1.30)$$

Nous admettrons que cette forme du lagrangien est effectivement convenable. Quelques remarques s'imposent à ce point:

- Le terme “d’énergie potentielle” que nous obtenons ici dépend explicitement de la vitesse des particules. Nous retrouvons ici, sous une autre forme, la non conservativité des forces électromagnétiques: l’énergie mécanique totale n’est pas constante dans le cas le plus général. Ce n’est que si le potentiel vecteur est indépendant du temps (au sens d’une dérivée partielle par rapport au temps) et de plus uniforme que le dernier terme du lagrangien se ramène à une dérivée totale (celle de $\mathbf{A}_\alpha \cdot \mathbf{r}_\alpha$), qui peut être oubliée. Le potentiel électrostatique joue alors le rôle d’une vraie énergie potentielle, ne dépendant que des positions des particules. Ce n’est donc que dans le cadre de l’électrostatique que les forces électromagnétiques sont conservatives, un résultat bien connu sur lequel nous jetons ici un éclairage nouveau. Notons que la non conservation de l’énergie mécanique ne signifie pas une violation de la conservation de l’énergie totale. Nous verrons en effet en relativité que l’énergie et l’impulsion totales sont toujours conservées, à condition de faire intervenir le champ dans le bilan.
- Encore une fois, les champs sont ici imposés de l’extérieur, et ne font pas partie des variables dynamiques du problème. Pour traiter le problème complet de particules en interaction avec leur propre champ, il faudrait adjoindre aux équations de Lagrange que nous venons d’écrire les équations de Maxwell permettant de calculer les champs en fonction des mouvements des particules. Ce n’est que dans la partie suivante que nous apprendrons à le faire de manière consistante dans un formalisme lagrangien englobant le champ.
- L’expression de notre lagrangien fait explicitement intervenir les potentiels. Il nous faut donc vérifier que les équations de Lagrange ne sont pas modifiées quand on effectue une transformation de jauge sur les potentiels. Rappelons en effet que les champs “physiques” ne sont pas modifiés si on effectue la transformation de “jauge”:

$$\mathbf{A} \longrightarrow \mathbf{A}' = \mathbf{A} + \nabla\phi \quad (1.31)$$

$$V \longrightarrow V' = V - \frac{\partial\phi}{\partial t}, \quad (1.32)$$

où ϕ est une fonction arbitraire de l’espace et du temps. Le nouveau lagrangien L' s’exprime alors facilement en fonction de l’ancien, en notant $\phi_\alpha = \phi(\mathbf{r}_\alpha, t)$:

$$\begin{aligned} L' &= L + \sum_\alpha q_\alpha \left(\frac{\partial\phi_\alpha}{\partial t} + (\nabla\phi_\alpha) \cdot \mathbf{v}_\alpha \right) \\ &= L + \sum_\alpha q_\alpha \left(\frac{\partial\phi_\alpha}{\partial t} + (\mathbf{v}_\alpha \cdot \nabla)\phi_\alpha \right) \\ &= L + \sum_\alpha q_\alpha \frac{d\phi_\alpha}{dt}, \end{aligned} \quad (1.33)$$

en reconnaissant dans les deux derniers termes de l’équation centrale la dérivée totale par rapport au temps de ϕ_α , somme de la dérivée partielle et de la “dérivée hydrodynamique”. L' et L ne diffèrent donc que par une dérivée totale par rapport au temps qui ne modifie pas le contenu des équations de Lagrange, comme nous nous y attendions.

1.4 Généralisations

Nous traiterons dans ce paragraphe de deux extensions du formalisme lagrangien³. Nous traiterons d’abord le cas de forces (non électromagnétiques) ne dérivant pas d’un potentiel. Nous verrons par exemple comment traiter le cas des forces de frottement dans le cadre d’un principe variationnel. Dans

³Les résultats de ce chapitre ne seront pratiquement pas utilisés dans la suite du cours. On pourra donc le sauter en première lecture

la seconde partie de ce paragraphe, nous nous pencherons sur le cas des systèmes où les mouvements sont sujets à des liaisons non holonomes. Nous verrons comment on peut facilement incorporer de telles liaisons dans le formalisme lagrangien.

1.4.1 Forces ne dérivant pas d'une énergie potentielle

On suppose connues les forces \mathbf{F}_α s'exerçant sur la particule α . Notre problème est de trouver une expression variationnelle de la dynamique et d'en déduire les équations de Lagrange correspondantes, même si les forces ne dérivent pas d'une énergie potentielle.

Revenons d'abord au cas des forces conservatives, pour mettre le principe de moindre action sous une forme un peu différente. Tel que nous l'avons écrit, le principe de moindre action stipule que $\delta S = \delta \int_{t_1}^{t_2} L dt = 0$, si on estime la variation de l'action entre la trajectoire effectivement suivie et une autre trajectoire infiniment voisine compatible avec les conditions aux limites. On peut aussi écrire $\int_{t_1}^{t_2} (\delta T - \delta U) dt = 0$. Si on note $\delta \mathbf{r}_\alpha$ l'écart pour la particule α entre la trajectoire normale et la trajectoire variée, on a $\delta U = -\sum_\alpha \mathbf{F}_\alpha \cdot \delta \mathbf{r}_\alpha$, le principe de moindre action s'écrivant alors

$$\int_{t_1}^{t_2} (\delta T + \sum_\alpha \mathbf{F}_\alpha \cdot \delta \mathbf{r}_\alpha) dt = 0 . \quad (1.34)$$

On peut exprimer les $\delta \mathbf{r}_\alpha$ en fonction des accroissements δq_i des coordonnées généralisées:

$$\sum_\alpha \mathbf{F}_\alpha \cdot \delta \mathbf{r}_\alpha = \sum_\alpha \sum_i \mathbf{F}_\alpha \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} \delta q_i . \quad (1.35)$$

Posons alors

$$Q_i = \sum_\alpha \mathbf{F}_\alpha \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} . \quad (1.36)$$

Nous désignerons Q_i comme la force généralisée correspondant à la coordonnée généralisée q_i ⁴. Avec cette notation, le principe de moindre action s'écrit:

$$\int_{t_1}^{t_2} (\delta T + \sum_i Q_i \delta q_i) dt = 0 . \quad (1.37)$$

Nous avons établi cette formulation en supposant que les forces dérivent d'une énergie potentielle. Nous admettrons qu'elle reste valable même si ce n'est pas le cas. Nous allons maintenant dériver de cette expression du principe de moindre action les équations de Lagrange correspondantes. L'approche est très similaire à celle que nous employâmes dans le paragraphe 1.2. T étant une fonction des q_i de leurs dérivées et du temps, on peut écrire

$$\delta T = \frac{\partial T}{\partial q_i} \delta q_i + \frac{\partial T}{\partial \dot{q}_i} \delta \dot{q}_i . \quad (1.38)$$

Une intégration par parties élémentaire donne alors:

$$\int_{t_1}^{t_2} \frac{\partial T}{\partial \dot{q}_i} \delta \dot{q}_i dt = \left[\frac{\partial T}{\partial \dot{q}_i} \delta q_i \right]_1^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} \delta q_i dt . \quad (1.39)$$

Le terme tout intégré est identiquement nul, puisque les deux trajectoires coïncident aux limites. En reportant cette expression dans le principe de moindre action et en écrivant que l'intégrale doit s'annuler quelles que soient les variations des coordonnées indépendantes q_i , on obtient les équations de Lagrange:

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = Q_i . \quad (1.40)$$

⁴“Force” qui n'a pas nécessairement la dimension d'une force et qui ne peut être en général attribuée à une particule particulière.

L'écriture de ce système d'équations et sa résolution ne présente aucune difficulté si on connaît les forces s'exerçant sur les particules.

Dans le cas particulier où la force généralisée Q_i peut se mettre sous la forme:

$$Q_i = -\frac{\partial U}{\partial q_i} + \frac{d}{dt} \frac{\partial U}{\partial \dot{q}_i} \quad \text{où } U = U(q_i, \dot{q}_i, t), \quad (1.41)$$

l'équation de Lagrange (1.40) se met sous la forme standard avec $L = T - U$. On retrouve ainsi le cas des forces conservatives mais aussi le cas des forces électromagnétiques. Notons qu'il n'est pas toujours possible de mettre la force généralisée sous cette forme (nous verrons que c'est le cas dans l'exemple traité à la fin de ce paragraphe). Soulignons aussi que nous permettons à U de dépendre des vitesses. Il est donc possible (comme nous l'avons vu pour les forces électromagnétiques) d'englober certaines forces non conservatives dans des équations de Lagrange sous la forme standard.

Très souvent, une partie des forces dérive d'un potentiel (au moins au sens généralisé exprimé par l'équation 1.41). Dans ce cas, en appelant U "l'énergie potentielle" associée aux forces en dérivant et Q_i les forces généralisées correspondant aux forces ne dérivant pas de U , on écrira les équations de Lagrange sous la forme:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = Q_i \quad \text{avec } L = T - U. \quad (1.42)$$

Nous appliquerons, à titre d'exemple, cette démarche à un système où s'exercent des forces de frottement proportionnelles à la vitesse. De telles forces ne sont manifestement pas conservatives. Nous poserons donc (en oubliant pour l'instant toutes les autres forces) $\mathbf{F}_\alpha = -k_\alpha \mathbf{v}_\alpha$. Cette force peut s'écrire $\mathbf{F}_\alpha = -\nabla_{\mathbf{v}_\alpha} \mathcal{F}$ à condition de poser

$$\mathcal{F} = \frac{1}{2} \sum_{\alpha} k_{\alpha} \mathbf{v}_{\alpha}^2. \quad (1.43)$$

Les forces généralisées Q_i s'écrivent:

$$Q_i = - \sum_{\alpha} \nabla_{\mathbf{v}_{\alpha}} \mathcal{F} \cdot \frac{\partial \mathbf{r}_{\alpha}}{\partial q_i}. \quad (1.44)$$

En remarquant alors que:

$$\frac{\partial \dot{\mathbf{r}}_{\alpha}}{\partial \dot{q}_i} = \frac{\partial \mathbf{r}_{\alpha}}{\partial q_i}, \quad (1.45)$$

on a

$$Q_i = - \sum_{\alpha} \nabla_{\mathbf{v}_{\alpha}} \mathcal{F} \cdot \frac{\partial \mathbf{v}_{\alpha}}{\partial \dot{q}_i} = - \frac{\partial \mathcal{F}}{\partial \dot{q}_i}. \quad (1.46)$$

On voit bien qu'une telle expression n'est pas compatible en général avec (1.41).

Les équations de Lagrange en présence de frottements visqueux s'écrivent finalement:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} + \frac{\partial \mathcal{F}}{\partial \dot{q}_i} = 0, \quad (1.47)$$

où L tient compte des éventuelles forces "conservatives".

Pour un mouvement unidimensionnel d'une particule soumise uniquement à une force de frottement, on a $\mathcal{F} = k\dot{x}^2/2$, et $L = m\dot{x}^2/2$. L'équation de Lagrange s'écrit donc $m\ddot{x} + k\dot{x} = 0$, ce qui coïncide bien avec l'expression du principe fondamental (encore une fois, le formalisme lagrangien n'apporte rien pour un problème aussi élémentaire).

1.4.2 Cas des liaisons non holonomes

Nous traiterons ici le cas de systèmes où les variables dynamiques peuvent être reliées par des liaisons non holonomes, c'est à dire ne se mettant pas simplement sous la forme de relations entre les coordonnées cartésiennes (éventuellement dépendant du temps). Nous nous contenterons de traiter le cas où les contraintes s'expriment par des relations entre les vitesses des différentes particules composant le système. Ce cas permet en effet de traiter de la plupart des liaisons du type "roulement sans glissement" dont l'importance est considérable en mécanique du solide. Les techniques que nous introduirons ici peuvent être étendues à d'autres types de contraintes.

Nous procéderons en deux étapes. Nous ne tiendrons d'abord compte que des liaisons holonomes, en introduisant des coordonnées généralisées qui seraient indépendantes si ces liaisons étaient les seules. Nous reviendrons sur le principe de moindre action pour montrer que l'écriture des équations de Lagrange en termes de ces variables dépendantes est impossible. Nous verrons ensuite qu'en introduisant des variables supplémentaires, les multiplicateurs de Lagrange, on peut obtenir un système d'équations différentielles indépendantes permettant, au moins en principe, de résoudre le problème. Nous interpréterons physiquement, enfin, ces variables supplémentaires et montrerons qu'elles décrivent les forces généralisées associées aux liaisons non holonomes.

En tenant compte uniquement des liaisons holonomes, nous définissons un ensemble de n coordonnées généralisées q_i . Pour fixer les idées, nous supposerons que toutes les forces (à part celles associées aux liaisons) dérivent d'une énergie potentielle, éventuellement généralisée comme au paragraphe précédent. Les liaisons non holonomes font que ces n coordonnées généralisées ne sont pas indépendantes. Nous supposerons qu'il existe m relations entre elles (m liaisons) et nous supposerons que ces relations relient les dérivées des q_i par rapport au temps et peuvent donc s'écrire:

$$\sum_{k=1}^n a_{lk} dq_k + a_{lt} dt = 0, \quad (l = 1 \dots m), \quad (1.48)$$

où les coefficients $a_{lk,t}$ sont indépendants du temps.

Reprenons maintenant la dérivation des équations de Lagrange à partir du principe de moindre action, comme dans le paragraphe 1.2. La comparaison de l'action sur la trajectoire effectivement suivie et sur la trajectoire variée n'est pas modifiée, et on écrit:

$$\delta S = \int_{t_1}^{t_2} \sum_k \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} \right] \delta q_k dt. \quad (1.49)$$

De cette équation, on ne peut déduire des équations de Lagrange pour les q_k . Ce ne sont pas en effet des variables indépendantes, dont les accroissements puissent être choisis arbitrairement. Les δq_k sont en fait reliés par les m relations:

$$\sum_k a_{lk} \delta \dot{q}_k = 0. \quad (1.50)$$

Elles résultent simplement de (1.48) en constatant que les a_{lt} s'éliminent quant on écrit les liaisons entre les accroissements des dérivées entre trajectoire de référence et trajectoire variée. Par intégration, les sommes $\sum_k a_{lk} \delta q_k$ sont constantes. Les accroissements de trajectoire s'annulant aux instants initiaux et finaux, cette constante est nulle et on a finalement:

$$\sum_k a_{lk} \delta q_k = 0. \quad (1.51)$$

En d'autres termes, nous avons a priori trop d'équations pour les $n-m$ variables vraiment indépendantes. Plutôt que d'éliminer des variables superfétatoires, nous allons introduire m inconnues supplémentaires (dont nous verrons qu'elles ont un sens physique). Nous considérons donc m fonctions supplémentaires du temps, pour l'instant arbitraires, $\lambda_l(t)$. Elles vérifient bien sûr:

$$\lambda_l \sum_k a_{lk} \delta q_k = 0, \quad (1.52)$$

et donc

$$\int_{t_1}^{t_2} \sum_{k,l} \lambda_l a_{lk} \delta q_k = 0 . \quad (1.53)$$

En introduisant cette relation dans l'expression du principe variationnel, on le met sous la forme:

$$\delta S = \int_{t_1}^{t_2} \sum_{k=1}^n \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} + \sum_{l=1}^m \lambda_l a_{lk} \right] \delta q_k dt . \quad (1.54)$$

Prenons comme variables indépendantes les $n - m$ premières coordonnées généralisées. Au moins formellement, les m dernières peuvent être calculées en fonction de celles ci. Nous choisissons alors les m fonctions λ_l de telle manière que:

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} + \sum_{l=1}^m \lambda_l a_{lk} = 0 , \quad (1.55)$$

pour toutes les valeurs de k entre $n - m + 1$ et n . Ce choix est toujours possible. Les λ_l sont en effet définis comme solutions d'un système linéaire dont le déterminant, formé des coefficients a_{lk} est non nul si les m relations exprimant les contraintes non holonomes sont linéairement indépendantes.

Le principe variationnel s'écrit alors:

$$\delta S = \int_{t_1}^{t_2} \sum_{k=1}^m \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} + \sum_{l=1}^m \lambda_l a_{lk} \right] \delta q_k dt \quad (k = 1 \dots m) . \quad (1.56)$$

Les m premières coordonnées et leurs accroissements étant indépendants, cette intégrale n'est nulle que si les m quantités entre crochets sont identiquement nulles sur la trajectoire effectivement suivie. En ajoutant à ces m équations de Lagrange les $n - m$ relations (1.55) et les m contraintes (1.48), on obtient enfin un système de $n + m$ équations à $n + m$ inconnues (q_i et λ_l):

$$\begin{cases} \frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} + \sum_{l=1}^m \lambda_l a_{lk} = 0 , & (k = 1 \dots n) \\ \sum_{k=1}^m a_{lk} \dot{q}_k + a_{lt} = 0 , & (l = 1 \dots m) . \end{cases} , \quad (1.57)$$

dont la résolution donne la trajectoire effectivement suivie.

Nous avons pu tenir compte des liaisons non holonomes supplémentaires en introduisant des variables additionnelles, ce qui n'est guère économique, surtout quand il s'agit de résoudre effectivement le système. Cependant, les λ_l possèdent une interprétation physique qui rend leur obtention importante. Nous pourrions en effet "simuler" l'effet des liaisons non holonomes en appliquant des forces supplémentaires dans le système (tout mouvement peut toujours être vu comme résultant de forces et pas de contraintes). En fait, ces forces ont une réalité physique. Elles correspondent, par exemple, aux frottements responsables d'un "roulement sans glissement". Ces forces, ne dérivant pas en général d'un potentiel, doivent intervenir dans les équations de Lagrange sous la forme de "forces généralisées" Q_k . Les n équations de Lagrange de ce système sans contraintes supplémentaires s'écriraient, avec la même fonction de Lagrange:

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} + Q_k = 0 , \quad (k = 1 \dots n) . \quad (1.58)$$

En comparant ces équations avec celles de (1.57), on voit que les Q_k doivent être définis par:

$$Q_k = \sum_{l=1}^m \lambda_l a_{lk} . \quad (1.59)$$

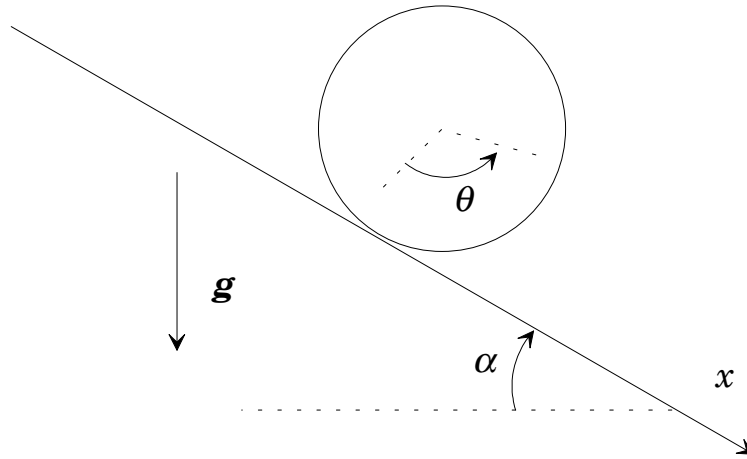


Figure 1.4: Roulement sans glissement d'un cylindre sur un plan incliné. Le mouvement est décrit a priori par deux coordonnées généralisées: la position du point de contact le long du plan et l'angle de rotation du cylindre par rapport à une référence arbitraire. Dans le cas d'un roulement sans glissement, ces deux coordonnées sont reliées par une relation non holonome.

Les fonctions supplémentaires λ_l ne sont donc que des combinaisons linéaires des forces généralisées correspondant aux contraintes. Leur obtention à partir de (1.57) permet donc de calculer les forces de liaison, ce qui justifie amplement leur intérêt.

Pour illustrer ce paragraphe, nous considérerons un problème trivial de dynamique des solides: le roulement sans glissement d'un cylindre creux (la masse est entièrement distribuée sur la surface extérieure) sur un plan incliné⁵ (voir figure 1.4). Si on ne tient compte que des liaisons holonomes (le cylindre est posé sur le plan, il ne se déplace que dans la direction x ...), nous pouvons décrire le mouvement par deux coordonnées généralisées. $q_1 = x$ est simplement l'abscisse du point de contact sur le plan, $q_2 = \theta$ l'angle de rotation du cylindre autour de son axe. Le roulement sans glissement impose l'unique contrainte non holonome supplémentaire $R\dot{\theta} + \dot{x} = 0$, où R est le rayon du cylindre. Nous n'avons donc que deux coefficients a_{lk} : $a_{11} = 1$ et $a_{12} = R$.

La fonction de Lagrange est $L = T - U$. L'énergie cinétique s'écrit:

$$T = \frac{1}{2}M\dot{q}_1^2 + \frac{1}{2}MR^2\dot{q}_2^2, \quad (1.60)$$

où M est la masse du cylindre, et MR^2 son moment d'inertie par rapport à son axe. On a aussi, dans le champ de pesanteur, $U = -Mgq_1 \sin \alpha$. Avec une seule liaison, il existe un seul multiplicateur de Lagrange λ_1 . Les deux équations de Lagrange et la contrainte nous fournissent alors un système de 3 équations à 3 inconnues:

$$\begin{cases} \frac{\partial L}{\partial q_1} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_1} + \lambda_1 = 0 \\ \frac{\partial L}{\partial q_2} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_2} - R\lambda_1 = 0 \\ \dot{q}_1 - R\dot{q}_2 = 0 \end{cases} \quad (1.61)$$

⁵Ce problème est en fait si élémentaire qu'il peut être traité en quelques lignes par le théorème de l'énergie cinétique.

On en déduit immédiatement:

$$\begin{cases} M\ddot{q}_1 - Mg \sin \alpha - \lambda_1 = 0 \\ MR^2\ddot{q}_2 - R\lambda_1 = 0 \\ \dot{q}_1 + R\dot{q}_2 = 0 \end{cases} . \quad (1.62)$$

Après quelques manipulations algébriques sans intérêt, on obtient un mouvement uniformément accéléré pour x avec l'accélération $g \sin(\alpha)/2$, un résultat bien connu de dynamique standard, et $\lambda_1 = -Mg \sin(\alpha)/2$. La force généralisée correspondant à la contrainte de roulement, Q_1 lui est égale. Enfin, on a simplement $Q_1 = F$, où F est la valeur algébrique de la force de frottement s'exerçant, tangentiellement au cylindre, au point de contact⁶.

1.5 Lagrangien et lois de conservation

Dans tout phénomène physique il existe des quantités conservées au cours de l'évolution. Ces quantités, aussi appelées en mécanique "intégrales premières du mouvement", jouent un rôle important en fournissant des renseignements sur la dynamique, même si la trajectoire n'est pas explicitement connue. Par exemple, on peut donner beaucoup de caractéristiques générales des collisions en écrivant la conservation de l'impulsion ou de l'énergie, sans même connaître de façon détaillée la loi régissant l'interaction entre les particules⁷.

Parmi toutes les quantités conservées dans le mouvement, certaines sont triviales et d'autres sans interprétation physique directe. Certaines en revanche, comme l'énergie ou l'impulsion, sont directement reliées à des propriétés physiques très fondamentales. Comme nous le verrons, il existe en général une quantité conservée associée à chaque invariance du système dans une transformation (à chaque propriété de symétrie, au sens le plus large). Nous verrons par exemple que la conservation de l'impulsion découle directement de l'invariance par translation dans l'espace.

Avant d'entrer dans le détail de ces lois de conservation, prouvons d'abord l'existence de quantités conservées dans une évolution lagrangienne. Nous considérerons dans tout ce paragraphe un système exempt des complications et généralisations introduites au paragraphe précédent. Nous ne considérerons donc que des liaisons holonomes. Comme les équations de Lagrange du système pour les n coordonnées généralisées q_i sont du second ordre en temps, la solution explicite du problème fait intervenir $2n$ constantes représentant les conditions aux limites (valeurs des coordonnées aux deux extrémités de la trajectoire). L'une de ces constantes peut toujours être mise sous la forme d'une origine arbitraire t_0 sur le temps. La solution la plus générale peut donc s'écrire:

$$q_i(t + t_0, C_1, \dots, C_{2n-1}), \dot{q}_i(t + t_0, C_1, \dots, C_{2n-1}), \quad (1.63)$$

où t_0 et les C_k sont les $2n$ constantes. Le mouvement étant déterminé, on peut inverser ces relations entre les $2n$ constantes et les coordonnées, et écrire:

$$t_0(q_i, \dot{q}_i, t), C_k(q_i, \dot{q}_i, t). \quad (1.64)$$

L'origine du temps joue un rôle particulier, et n'est pas à proprement parler une intégrale première du mouvement. En revanche, les $2n - 1$ C_k sont bien des fonctions de l'état dynamique du système qui restent constantes au cours du mouvement. Il y a donc, de façon très générale, au moins $2n - 1$ intégrales premières indépendantes dans un mouvement à n degrés de liberté.

Dans la suite de ce paragraphe, nous allons nous concentrer sur trois d'entre elles: l'impulsion, l'énergie et le moment cinétique et montrer comment elles se déduisent de propriétés de symétrie du lagrangien.

⁶Ce résultat peut être établi immédiatement à partir de l'accélération du mouvement

⁷On consultera avec profit le Landau de Mécanique sur ce problème que nous n'aborderons pas du tout dans ce cours.

1.5.1 Invariance par translation dans le temps: énergie

Nous considérons ici un système isolé. La première conséquence est que la fonction de Lagrange ne peut dépendre explicitement du temps. Il n'y a rien en effet dans l'environnement du système pour fixer une origine de temps. Nous supposons de plus que les équations dynamiques de ce système sont des équations de Lagrange standard (sans forces généralisées). On exclut donc explicitement le cas où le système serait soumis à des forces de frottement (difficilement compatibles avec l'hypothèse d'un système isolé).

On a donc $\partial L/\partial t = 0$. La dérivée totale de L par rapport au temps, somme de sa dérivée partielle et des variations temporelles provenant de la variation des coordonnées généralisées, peut donc s'écrire:

$$\frac{dL}{dt} = \sum_i \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i + \sum_i \frac{\partial L}{\partial q_i} \dot{q}_i . \quad (1.65)$$

En utilisant les équations de Lagrange vérifiées par L et les q_i , on met cette dérivée sous la forme:

$$\begin{aligned} \frac{dL}{dt} &= \sum_i \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i + \sum_i \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \\ &= \frac{d}{dt} \left[\sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right] . \end{aligned} \quad (1.66)$$

On trouve ainsi que la quantité

$$\mathcal{E} = \sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \quad (1.67)$$

est une constante ou intégrale première du mouvement. Notons tout de suite qu'elle a la même dimension que L , c'est à dire celle d'une énergie. Pour préciser son interprétation physique, considérons le cas où les forces internes au système dérivent d'une énergie potentielle U ne dépendant que des q_i . On a alors $L = T(q_i, \dot{q}_i) - U$. T est très généralement une forme quadratique des dérivées temporelles des coordonnées généralisées⁸. T vérifie donc le théorème d'Euler:

$$2T = \sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} . \quad (1.68)$$

Comme U ne dépend que des q_i , $\partial L/\partial \dot{q}_i = \partial T/\partial \dot{q}_i$, et donc

$$\mathcal{E} = 2T - L = 2T - (T - U) = T + U . \quad (1.69)$$

La quantité \mathcal{E} coïncide donc dans ce cas simple avec l'énergie mécanique totale du système. Dans un cas plus complexe, nous admettrons donc que l'énergie mécanique du système est définie comme \mathcal{E} et est donc une intégrale première du mouvement. Cette loi de conservation apparaît ici très clairement comme une conséquence directe de l'invariance par translation dans le temps.

L'hypothèse du système isolé exclut bien sûr de traiter le cas d'un ensemble de particules en interaction avec un champ électromagnétique extérieur. Il existe cependant un cas où on peut définir une énergie mécanique conservée pour un tel système: celui d'un champ statique. Les équations de Lagrange s'écrivent en effet normalement avec le lagrangien

$$L = \sum_{\alpha} T_{\alpha} - q_{\alpha} V_{\alpha} + q_{\alpha} \mathbf{A}_{\alpha} \cdot \mathbf{v}_{\alpha} , \quad (1.70)$$

et les notations du paragraphe 1.3 (pour simplifier les écritures nous supposons ici que les coordonnées généralisées sont les coordonnées cartésiennes). \mathbf{A} ne dépendant que de la position des particules et L

⁸ Seulement dans le cas où les contraintes ne font pas explicitement intervenir de dépendance temporelle.

étant explicitement indépendant du temps, le raisonnement précédent s'applique en entier. L'énergie mécanique s'écrit simplement dans ce cas:

$$\mathcal{E} = \sum_{\alpha} \nabla_{\mathbf{v}_{\alpha}} L \cdot \mathbf{v}_{\alpha} - L . \quad (1.71)$$

Elle est conservée et on a:

$$\sum_{\alpha} \nabla_{\mathbf{v}_{\alpha}} L \cdot \mathbf{v}_{\alpha} = 2T + \sum_{\alpha} q_{\alpha} \nabla_{\mathbf{v}_{\alpha}} (\mathbf{A}_{\alpha} \cdot \mathbf{v}_{\alpha}) \cdot \mathbf{v}_{\alpha} = 2T + \sum_{\alpha} q_{\alpha} \mathbf{A}_{\alpha} \cdot \mathbf{v}_{\alpha} . \quad (1.72)$$

Et donc:

$$\mathcal{E} = T + \sum_{\alpha} q_{\alpha} V_{\alpha} , \quad (1.73)$$

résultat évident de statique.

1.5.2 Translation spatiale: conservation de l'impulsion

Nous considérons, comme dans le paragraphe précédent, un système isolé régi par des équations de Lagrange sous la forme standard. La dynamique est visiblement invariante si on déplace globalement le système. Il n'existe rien dans l'environnement pour fixer une origine des coordonnées. Comme ce déplacement est une quantité vectorielle, on a en fait trois conditions d'invariance d'où on peut déduire la conservation de trois quantités.

Nous allons d'abord supposer que le système est décrit par ses coordonnées cartésiennes. Les coordonnées généralisées posent en effet problème, puisqu'une translation spatiale n'a pas forcément une expression simple en termes de ces coordonnées. Nous examinerons ce cas à la fin du paragraphe. Nous considérons une translation d'ensemble du système, qui s'écrit:

$$\mathbf{r}_{\alpha} \longrightarrow \mathbf{r}_{\alpha} + \boldsymbol{\varepsilon} . \quad (1.74)$$

La translation $\boldsymbol{\varepsilon}$ est supposée petite à l'échelle des longueurs caractéristiques du système. On peut aisément exprimer la variation de la fonction de Lagrange dans cette translation. Puisque les vitesses ne changent pas,

$$\delta L = \sum_{\alpha} \nabla_{\mathbf{r}_{\alpha}} L \cdot \boldsymbol{\varepsilon} . \quad (1.75)$$

Cette variation ne peut être nulle quel que soit $\boldsymbol{\varepsilon}$ que si

$$\sum_{\alpha} \nabla_{\mathbf{r}_{\alpha}} L = 0 . \quad (1.76)$$

En utilisant les équations de Lagrange en coordonnées cartésiennes, on met cette condition sous la forme:

$$\frac{d\mathbf{P}}{dt} = 0 , \quad (1.77)$$

avec

$$\mathbf{P} = \sum_{\alpha} \nabla_{\mathbf{v}_{\alpha}} L . \quad (1.78)$$

Nous trouvons donc bien une intégrale première vectorielle pour le mouvement. Dans le cas où $L = T - U$, U ne dépendant que des \mathbf{r}_{α} , $\nabla_{\mathbf{v}_{\alpha}} L = \nabla_{\mathbf{v}_{\alpha}} T$ et

$$\mathbf{P} = \sum_{\alpha} m_{\alpha} \mathbf{v}_{\alpha} , \quad (1.79)$$

indiquant que cette intégrale première vectorielle n'est autre que l'impulsion ou quantité de mouvement totale.

Notons à ce point que l'impulsion totale est définie en terme du gradient par rapport aux vitesses et ne coïncide pas forcément avec l'expression habituelle $\sum_{\alpha} m_{\alpha} \mathbf{v}_{\alpha}$. Ce sera en particulier le cas en électromagnétisme où le potentiel vecteur entre dans la définition de l'impulsion. Comme seuls des potentiels vecteur et scalaire uniformes (c'est à dire des champs nuls) seraient compatibles avec les hypothèses actuelles, nous n'examinerons pas ce cas pour l'instant.

Étudions maintenant le cas où l'existence de contraintes impose le recours à des coordonnées généralisées. L'invariance de l'impulsion totale ne devrait pas dépendre de l'existence de contraintes internes au système. Toutefois, la conservation de \mathbf{P} ne peut plus être établie aussi simplement, puisqu'une translation n'a pas nécessairement une expression simple en termes de coordonnées généralisées. Définissons quand même, par analogie avec le cas précédent, une "impulsion généralisée" p_i associée à la coordonnée généralisée q_i par

$$p_i = \frac{\partial L}{\partial \dot{q}_i} . \quad (1.80)$$

Nous verrons que des impulsions généralisées jouent un rôle central dans le formalisme hamiltonien. En attendant, notons que l'énergie mécanique totale peut se réécrire simplement:

$$\mathcal{E} = \sum_i p_i \dot{q}_i - L . \quad (1.81)$$

Supposons maintenant que la fonction de Lagrange L ne fasse pas explicitement intervenir la coordonnée q_i :

$$\frac{\partial L}{\partial q_i} = 0 . \quad (1.82)$$

Nous dirons alors que q_i est une coordonnée "cyclique". La simple écriture de l'équation de Lagrange pour q_i donne alors:

$$\frac{dp_i}{dt} = 0 . \quad (1.83)$$

L'impulsion associée à une coordonnée cyclique est conservée.

Revenons maintenant à l'invariance par translation. Trois des coordonnées généralisées d'un système isolé peuvent être prises égales aux trois coordonnées cartésiennes du centre d'inertie. De manière évidente, par invariance par translation, ces trois coordonnées sont cycliques. Il en résulte que les trois impulsions associées sont des intégrales premières du mouvement, formant un vecteur qui n'est autre que l'impulsion totale.

1.5.3 Invariance par rotation: moment cinétique

Nous appliquerons ici une méthode très semblable à celle du paragraphe précédent au cas où le système et donc sa fonction de Lagrange sont invariants dans une rotation quelconque autour d'un axe défini par un vecteur unitaire \mathbf{u} . C'est en particulier le cas pour un système isolé. Nous examinerons ici uniquement le cas où le système est défini par les coordonnées cartésiennes des particules.

Nous considérons une rotation infinitésimale d'un angle $\delta\phi$ autour de l'axe. En posant $\delta\phi = \delta\phi \mathbf{u}$, nous pouvons écrire les variations des positions et vitesses dans cette transformation $\delta\mathbf{r} = \delta\phi \times \mathbf{r}$ et $\delta\mathbf{v} = \delta\phi \times \mathbf{v}$. L'invariance par rotation impose que la variation δL de la fonction de Lagrange soit nulle dans cette transformation. Or

$$\delta L = \sum_{\alpha} \nabla_{\mathbf{r}_{\alpha}} L \cdot \delta\mathbf{r}_{\alpha} + \nabla_{\mathbf{v}_{\alpha}} L \cdot \delta\mathbf{v}_{\alpha} . \quad (1.84)$$

En utilisant les équations de Lagrange, qui s'écrivent $\nabla_{\mathbf{r}_{\alpha}} L = \dot{\mathbf{p}}_{\alpha}$, où \mathbf{p}_{α} est l'impulsion de la particule α , on peut écrire:

$$\delta L = \sum_{\alpha} [\dot{\mathbf{p}}_{\alpha} \cdot (\delta\phi \times \mathbf{r}_{\alpha}) + \mathbf{p}_{\alpha} \cdot (\delta\phi \times \mathbf{v}_{\alpha})]$$

$$\begin{aligned}
&= \sum_{\alpha} [\delta\phi \cdot (\mathbf{r}_{\alpha} \times \dot{\mathbf{p}}_{\alpha}) + \delta\phi \cdot (\mathbf{v}_{\alpha} \times \mathbf{p}_{\alpha})] \\
&= \delta\phi \cdot \frac{d\mathbf{L}}{dt},
\end{aligned} \tag{1.85}$$

où nous avons posé

$$\mathbf{L} = \sum_{\alpha} \mathbf{r}_{\alpha} \times \mathbf{p}_{\alpha}. \tag{1.86}$$

La fonction de Lagrange ne peut être conservée pour une rotation arbitraire que si la composante du vecteur \mathbf{L} sur l'axe de rotation est une intégrale première du mouvement. Nous retrouvons ainsi la conservation du moment cinétique par rapport à un axe.

On peut voir aussi, plus simplement, dans ce cas, que le lagrangien est indépendant d'une coordonnée généralisée décrivant le mouvement de rotation autour de l'axe. Il est donc cyclique dans cette coordonnée. L'impulsion généralisée associée à cette coordonnée est constante. Il est facile de vérifier qu'il s'agit bien de la composante le long de cet axe du moment cinétique.

Dans le cas où le système est invariant dans une rotation arbitraire autour d'un axe quelconque, ce qui est le cas d'un système isolé, il en résulte que le moment cinétique \mathbf{L} est une constante du mouvement.

1.6 Action en fonction de la trajectoire

Nous allons utiliser les définitions du paragraphe précédent pour tenter d'exprimer simplement la dépendance de l'action S calculée sur la trajectoire effectivement suivie en fonction des coordonnées spatiales et temporelles du point de départ et du point d'arrivée. En d'autres termes, l'objet de ce paragraphe est de donner les dérivées partielles de $S(q(1), t_1, q(2), t_2)$, action sur la trajectoire suivie considérée comme une fonction des conditions aux limites. Ces résultats nous seront fort utiles dans la suite du cours. Ils nous permettront aussi de jeter un regard nouveau sur les lois de conservation associées aux invariances.

1.6.1 Dépendance en position

Nous considérons ici la variation de S en fonction de la position du point d'arrivée. Nous en déduisons immédiatement la dépendance vis à vis du point de départ. Nous considérons donc ici deux trajectoires *effectivement suivies* entre les instants t_1 et t_2 . La première (trajectoire de référence) s'effectue entre les valeurs $q_i(1)$ et $q_i(2)$ des coordonnées généralisées. La seconde entre les valeurs $q_i(1)$ et $q_i(2) + \delta q_i(2)$, où $\delta q_i(2)$ est un accroissement infinitésimal de la position du point d'arrivée. Les trajectoires, par continuité, restent infiniment voisines pour tous les instants⁹ et on écrira $\delta q_i(t)$ l'écart entre elles.

La variation de l'action en passant d'une trajectoire à l'autre, δS , s'écrit simplement:

$$\delta S = \int_{t_1}^{t_2} \delta L dt, \tag{1.87}$$

où δL est la variation de la fonction de Lagrange entre les deux points à l'instant t . On suit alors un raisonnement identique à celui utilisé pour établir les équations de Lagrange. On écrit d'abord:

$$\delta L = \sum_i \frac{\partial L}{\partial q_i} \delta q_i + \sum_i \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i, \tag{1.88}$$

⁹Si la dynamique du système était chaotique, deux trajectoires très voisines aux points de départ et d'arrivée, peuvent s'écarter notablement l'une de l'autre. Comme nous manipulons des accroissements infinitésimaux, ce problème ne se pose pas.

et on reporte cette expression dans celle de δS . Le terme faisant intervenir $\delta \dot{q}_i$ est alors intégré par parties. On obtient:

$$\delta S = \int_{t_1}^{t_2} \sum_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right] \delta q_i dt + \sum_i \left[\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right]_1 . \quad (1.89)$$

La trajectoire de référence que nous considérons ici est une solution des équations du mouvement. Les équations de Lagrange étant constamment vérifiées, l'intégrale dans l'expression ci-dessus s'annule identiquement. Il ne reste donc que le terme tout intégré, qui se réduit à:

$$\delta S = \sum_i \frac{\partial L}{\partial \dot{q}_i} \delta q_i(2) . \quad (1.90)$$

De manière évidente, si nous avons considéré deux trajectoires différant d'une quantité infinitésimale au point de départ, nous aurions obtenu:

$$\delta S = - \sum_i \frac{\partial L}{\partial \dot{q}_i} \delta q_i(1) . \quad (1.91)$$

Ces deux expressions nous donnent les dérivées partielles de l'action, considérée comme une fonction des coordonnées des points de départ et d'arrivée. En remarquant que $\partial L / \partial \dot{q}_i = p_i$ (impulsion généralisée), on écrira

$$\frac{\partial S}{\partial q_i(2)} = p_i(2) \quad (1.92)$$

$$\frac{\partial S}{\partial q_i(1)} = -p_i(1) \quad (1.93)$$

On peut facilement retrouver, à partir de ce résultat, l'invariance de l'impulsion associée à une coordonnée cyclique. Si q_i est cyclique, la fonction de Lagrange ne dépend pas de q_i et le mouvement doit être invariant dans une translation de la coordonnée q_i . Considérons donc la translation infinitésimale $q_i \rightarrow q_i + \epsilon$. La variation de l'action dans cette transformation est:

$$\delta S = \frac{\partial S}{\partial q_i(2)} \epsilon + \frac{\partial S}{\partial q_i(1)} \epsilon = (p_i(2) - p_i(1)) \epsilon , \quad (1.94)$$

d'où on tire immédiatement la conservation de p_i . Dans le cas où les coordonnées cycliques sont celles du centre d'inertie, on retrouve la conservation de l'impulsion au sens habituel.

1.6.2 Dépendance en temps

Nous considérons maintenant la dépendance de l'action dans le temps d'arrivée t_2 . Nous considérons donc deux trajectoires effectivement suivies par le système. L'une, entre $q(1), t_1$ et $q(2), t_2$ est la trajectoire de référence. L'autre coïncide avec la première jusqu'à l'instant t_2 et continue ensuite pendant un intervalle de temps infinitésimal jusqu'à $t_2 + \delta t_2$, les coordonnées étant alors $q_i(2) + \delta q_i(2)$. On peut écrire de manière évidente la variation de l'action entre ces deux trajectoires comme:

$$\delta S = L(t_2) \delta t_2 , \quad (1.95)$$

mais aussi comme:

$$\delta S = \frac{\partial S}{\partial t_2} \delta t_2 + \sum_i \frac{\partial S}{\partial q_i(2)} \delta q_i(2) . \quad (1.96)$$

En effet l'instant du point d'arrivée et ses coordonnées varient. En utilisant les résultats du paragraphe précédent et en remarquant que $\delta q_i(2) = \dot{q}_i(2) \delta t_2$, on peut écrire:

$$L(t_2) = \frac{\partial S}{\partial t_2} + \sum_i p_i(2) \dot{q}_i(2) . \quad (1.97)$$

En nous souvenant de la définition de l'énergie mécanique totale \mathcal{E}^{10} , nous avons enfin:

$$\frac{\partial S}{\partial t_2} = -\mathcal{E}(2) . \quad (1.98)$$

En reproduisant le même raisonnement pour une variation de l'instant de départ, on obtiendrait de même:

$$\frac{\partial S}{\partial t_1} = \mathcal{E}(1) . \quad (1.99)$$

Avec ces deux expressions et les résultats du paragraphe précédent, nous connaissons toutes les dérivées partielles de l'action en fonction des conditions aux limites imposées à la trajectoire.

Notons également qu'on peut retrouver facilement à partir de ces expressions la conservation de l'énergie mécanique. Si la fonction de Lagrange ne dépend pas explicitement du temps, l'action doit être invariante dans une translation temporelle globale infinitésimale. La variation de l'action dans cette translation s'écrivant évidemment $\delta S = (\mathcal{E}(1) - \mathcal{E}(2))\delta t$, on retrouve l'invariance de l'énergie mécanique.

1.7 Deux théorèmes utiles

Nous considérerons dans ce paragraphe le cas particulier important des énergies potentielles homogènes, c'est à dire telles que:

$$U(\alpha q_1, \dots, \alpha q_n) = \alpha^k U(q_1, \dots, q_n) . \quad (1.100)$$

Ce cas, très spécifique, se rencontre en fait dans une grande variété de problèmes de mécanique:

- Pour l'oscillateur harmonique unidimensionnel $U = Kx^2/2$ est évidemment une fonction homogène avec $k = 2$. Ce résultat est bien sûr valide aussi pour les oscillateurs multidimensionnels.
- Le potentiel gravitationnel ou Coulombien en $1/r$ correspond à $k = -1$.
- Le mouvement dans un champ uniforme (par exemple champ de pesanteur local) correspond enfin à $k = 1$.

Le fait que U soit une fonction homogène a deux conséquences importantes que nous allons examiner dans les prochains paragraphes. D'abord, on peut en déduire des lois d'échelle utiles fixant la dépendance relative de certains paramètres des trajectoires (par exemple la troisième loi de Kepler). La deuxième conséquence est le théorème du viriel, d'une grande importance en mécanique et physique statistique.

1.7.1 Lois d'échelle

Considérons une transformation d'échelle sur les coordonnées généralisées et le temps. Elle peut s'écrire:

$$q_i \longrightarrow \alpha q_i \quad (1.101)$$

$$t \longrightarrow \beta t \quad (1.102)$$

¹⁰Nous n'avons fait dans ce paragraphe aucune hypothèse sur la dépendance en temps de la fonction de Lagrange. \mathcal{E} n'est donc pas nécessairement une quantité conservée. En fait, plutôt que d'énergie totale, nous devrions parler de fonction de Hamilton (voir chapitre suivant).

On en déduit les transformations de quelques quantités:

$$\dot{q}_i \longrightarrow \frac{\alpha}{\beta} \dot{q}_i \quad (1.103)$$

$$U \longrightarrow \alpha^k U \quad (1.104)$$

$$T \longrightarrow \left(\frac{\alpha}{\beta}\right)^2 T. \quad (1.105)$$

Pour établir la dernière ligne, il faut admettre que l'énergie cinétique ne dépend que des \dot{q}_i et pas de q_i . C'est en particulier le cas en coordonnées cartésiennes standard. Nous supposons cette condition remplie dans toute la suite du paragraphe.

En général, le lagrangien L ne se transforme pas de manière simple. En revanche, dans le cas particulier où les facteurs d'échelle sur T et U sont les mêmes, c'est à dire si on choisit les facteurs d'échelle α et β de telle manière que:

$$\left(\frac{\alpha}{\beta}\right)^2 = \alpha^k, \quad (1.106)$$

où encore

$$\beta = \alpha^{1-k/2}, \quad (1.107)$$

la fonction de Lagrange est simplement multipliée par un facteur d'échelle. Un tel facteur est sans effet sur les équations du mouvement. A une trajectoire possible correspond donc après cette transformation d'espace et de temps une autre trajectoire possible. Cette correspondance permet par exemple d'obtenir les lois reliant la période d'un mouvement à l'extension spatiale de la trajectoire. Appliquons ces arguments aux cas particulièrement importants de l'oscillateur harmonique et du mouvement gravitationnel.

- Pour l'oscillateur harmonique, $k = 2$. L'équation (1.107) est donc satisfaite pour n'importe quelle valeur de α à condition que $\beta = 1$. On obtient donc une trajectoire possible en multipliant la coordonnée spatiale par un facteur arbitraire et en laissant inchangée l'échelle de temps. On retrouve ainsi très simplement l'isochronisme des oscillations de l'oscillateur harmonique.
- Pour le mouvement dans un champ gravitationnel uniforme, on a $k = 1$. L'équation (1.107) est donc satisfaite si $\beta = \sqrt{\alpha}$. Deux trajectoires se correspondent donc si leurs extensions spatiales ℓ et ℓ' et temporelles (\mathcal{T} et \mathcal{T}') sont reliées par $\mathcal{T}'/\mathcal{T} = \sqrt{\ell'/\ell}$. On retrouve ainsi, par exemple, la dépendance en $\sqrt{\ell}$ de la période des oscillations d'un pendule simple de longueur ℓ .
- Enfin, pour le cas du mouvement de Kepler dans un potentiel gravitationnel central, on a $k = -1$. La relation (1.107) est donc satisfaite si $\beta = \alpha^{3/2}$. On en déduit alors, par exemple, la troisième loi de Kepler qui relie la période d'une orbite \mathcal{T} à son demi grand axe a : \mathcal{T}^2/a^3 est une constante.

Ces trois exemples illustrent bien la puissance de ces simples lois d'échelle. En fait, ce genre d'arguments se transpose à des domaines très variés et permet souvent d'obtenir des lois très générales par de simples considérations d'échelle ou de dimensionnalité.

1.7.2 Théorème du Viriel

Nous établirons ici, dans le cas des énergies potentielles homogènes, un lien utile entre les énergies potentielles et cinétiques moyennes (moyennées sur un temps long devant les temps caractéristiques du mouvement, la période par exemple). Ce théorème possède de nombreuses applications en mécanique céleste.

L'énergie cinétique T est une forme quadratique des vitesses. On peut donc lui appliquer le théorème d'Euler:

$$2T = \sum_i \frac{\partial T}{\partial \dot{q}_i} \dot{q}_i. \quad (1.108)$$

L'énergie potentielle ne dépendant pas des vitesses généralisées, $\partial T/\partial \dot{q}_i = \partial L/\partial \dot{q}_i$. En utilisant la définition des impulsions généralisées, on peut mettre cette dernière relation sous la forme:

$$2T = \sum_i p_i \dot{q}_i = \frac{d}{dt} \sum_i p_i q_i - \sum_i \dot{p}_i q_i . \quad (1.109)$$

Prenons maintenant la valeur moyenne temporelle de T définie comme:

$$\bar{T} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau T dt . \quad (1.110)$$

Notons que, dans le cas des mouvements périodiques, on peut définir la moyenne temporelle comme l'intégrale sur une période. En incluant dans cette définition de la valeur moyenne l'expression précédente de T , on trouve:

$$2\bar{T} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \left[\sum_i p_i q_i \right]_0^\tau - \frac{1}{\tau} \int_0^\tau \sum_i \dot{p}_i q_i . \quad (1.111)$$

Si le mouvement est *borné*, le terme tout intégré reste fini quand τ tend vers l'infini et ne contribue donc pas à la valeur moyenne (il est strictement nul si le mouvement est périodique et si les intégrales s'étendent sur exactement une période). On a donc:

$$2\bar{T} = - \sum_i \overline{\dot{p}_i q_i} , \quad (1.112)$$

ou encore, en utilisant les équations de Lagrange $\dot{p}_i = \partial L/\partial q_i$,

$$2\bar{T} = - \sum_i \overline{\frac{\partial L}{\partial q_i} q_i} . \quad (1.113)$$

Si, comme dans le paragraphe précédent, T ne dépend que des \dot{q}_i , on a $\partial L/\partial q_i = -\partial U/\partial q_i$. Si enfin U est une fonction homogène des coordonnées de degré k , on peut écrire:

$$2\bar{T} = k\bar{U} . \quad (1.114)$$

Rappelons pour finir les hypothèses nécessaires à l'établissement de cette propriété:

- Les forces dérivent d'une énergie potentielle et les équations de Lagrange s'appliquent sous la forme ordinaire.
- L'énergie potentielle est une fonction des coordonnées homogène et de degré k .
- L'énergie cinétique ne dépend que des vitesses généralisées.
- Le mouvement est borné.

Ces hypothèses contraignantes sont en fait assez souvent remplies. C'est par exemple le cas pour les mouvements périodiques dans un potentiel gravitationnel central, où la relation s'écrit, avec $k = -1$:

$$\bar{T} = -\frac{1}{2}\bar{U} \quad (1.115)$$

ou pour les mouvements dans un potentiel harmonique (nécessairement bornés) pour lesquels $k = 2$ et

$$\bar{T} = \bar{U} . \quad (1.116)$$

Chapitre 2

Formulation hamiltonienne

Nous présenterons brièvement dans ce chapitre la formulation hamiltonienne de la mécanique. Bien qu'elle soit strictement équivalente à la formulation lagrangienne, elle s'avère souvent plus puissante et d'un usage plus commode.

D'abord, d'un point de vue très utilitaire, la formulation hamiltonienne facilite la résolution, analytique ou numérique, des équations du mouvement. Les équations de Lagrange sont des équations différentielles du second ordre. Leur résolution analytique n'est pas toujours facile, si elle est possible. Très souvent (par exemple dès le problème à trois corps en mécanique céleste), on doit recourir à une intégration numérique. Il se trouve que les équations différentielles du second ordre ne se prêtent pas très bien à une telle intégration. Il se pose en particulier des problèmes sévères de stabilité numérique. En revanche, les équations différentielles du premier ordre s'intègrent très aisément et de façon numériquement stable. Comme nous le verrons, les équations de Hamilton sont du premier ordre.

En fait, le formalisme hamiltonien consiste à traiter sur un même plan les positions et les impulsions généralisées associées. Outre la transformation d'une équation de Lagrange du second ordre, portant sur la position seule, en deux équations de Hamilton du premier ordre reliant position et impulsion, cette approche ouvre la voie à des changements de variables (nous dirons des transformations canoniques) très puissants. Plutôt que de changer simplement de coordonnées généralisées, on peut en effet mélanger positions et impulsions dans un changement de variables. Il est même possible ainsi d'échanger le rôle des impulsions et des positions ou de prendre les conditions aux limites comme nouvelles variables. Cette immense liberté dans le choix de la description du problème est au centre de nombreuses méthodes de résolution des problèmes de mécanique.

Un autre avantage décisif de la formulation hamiltonienne est qu'elle se prête à merveille à des méthodes de perturbations (développement de la solution en puissance des perturbations, supposées petites, par rapport à un problème dont la solution est connue). Ces méthodes sont d'une importance primordiale en mécanique céleste (prise en compte, par exemple, des perturbations au mouvement de Kepler dues aux autres planètes). Enfin, et surtout, c'est la formulation hamiltonienne de la mécanique classique qui se prête à la quantification (dite, elle aussi, canonique). Nous n'aborderons pas ce problème, mais un certain nombre des notions que nous introduirons dans ce chapitre parfaitement classique ont une contrepartie dans le formalisme quantique. Si, de manière évidente, la fonction de Hamilton est remplacée par l'opérateur hamiltonien, les crochets de Poisson, par exemple, correspondent aux commutateurs.

2.1 Equations de Hamilton

Nous partons de la formulation lagrangienne établie au chapitre précédent, dans sa forme la plus simple. Nous n'intégrerons donc pas les généralisations aux forces ne dérivant pas d'un potentiel (mises à part les forces électromagnétiques) ni les généralisations à des liaisons non holonomes. On

définit les impulsions généralisées à partir de la fonction de Lagrange des positions, de vitesses et du temps¹ par

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad (2.1)$$

et on peut écrire les équations de Lagrange sous la forme:

$$\dot{p}_i = \frac{\partial L}{\partial q_i} . \quad (2.2)$$

Ces écritures font apparaître L comme une fonction naturelle des q_i et des \dot{q}_i , dont les dérivées partielles respectives sont les \dot{p}_i et les p_i . La “différentielle totale” du lagrangien s’écrit alors:

$$dL = \sum_i p_i d\dot{q}_i + \sum_i \dot{p}_i dq_i . \quad (2.3)$$

La technique des transformations de Legendre, largement illustrée en thermodynamique classique², permet de passer de L à une fonction dont la différentielle s’exprime de façon naturelle en fonction des q_i et des p_i . Posons en effet

$$H = \sum_i p_i \dot{q}_i - L . \quad (2.4)$$

Nous appellerons H la fonction de Hamilton. Notons que, pour les systèmes conservatifs, la définition de la fonction de Hamilton coïncide avec celle de l’énergie totale. En revanche, pour des systèmes où la fonction de Lagrange dépend explicitement du temps, la notion d’énergie mécanique totale perd tout son intérêt, alors que la fonction de Hamilton garde son sens et permet toujours d’écrire les équations du mouvement. C’est ce caractère plus général de la fonction de Hamilton qui justifie de ne pas la confondre avec l’énergie totale. La différentielle de la fonction de Hamilton s’écrit sans difficulté comme:

$$dH = \sum_i \dot{q}_i dp_i - \sum_i \dot{p}_i dq_i . \quad (2.5)$$

H apparaît donc comme une fonction naturelle des p_i et q_i dont les dérivées partielles respectives sont les \dot{q}_i et les \dot{p}_i . Les $2n$ équations de Hamilton, équivalentes aux n équations de Lagrange, qui permettent la résolution du problème avec les conditions aux limites, s’écrivent donc:

$$\frac{\partial H}{\partial p_i} = \dot{q}_i \quad (2.6)$$

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i , \quad (2.7)$$

auxquelles on pourrait ajouter la relation $\partial H / \partial t = -\partial L / \partial t$, qui ne présente d’intérêt que pour un lagrangien et donc un hamiltonien dépendant explicitement du temps. Les positions et les impulsions sont dites variables conjuguées.

Comme attendu, nous avons donc remplacé les n équations différentielles du second ordre que nous donne le formalisme lagrangien, par un nombre double d’équations différentielles du premier ordre qui se prêtent plus facilement à une résolution analytique ou numérique. Les conditions initiales naturelles pour ces équations sont les positions et les impulsions initiales (alors que c’étaient les positions et les vitesses initiales pour les équations de Lagrange). Rappelons que les impulsions ne coïncident pas nécessairement avec les vitesses généralisées, en particulier en présence de champs électromagnétiques).

¹Nous supposerons très souvent dans ce chapitre que la fonction de Lagrange ne dépend pas explicitement du temps. L’énergie mécanique totale est alors une intégrale première du mouvement.

²C’est en effet la transformation qui permet, par exemple, de passer de l’énergie interne U , dont la différentielle s’exprime simplement en fonction de celles du volume et de l’entropie, à l’enthalpie H , faisant intervenir pression et entropie.

Notons enfin la symétrie remarquable de ces équations, qui traitent de façon tout à fait analogue les positions et les impulsions (à un signe près).

Notons aussi que si q_i est une coordonnée cyclique, la fonction de Lagrange et donc la fonction de Hamilton ne dépendent pas explicitement de q_i . On déduit alors immédiatement des équations de Hamilton que p_i est une constante du mouvement.

Examinons maintenant le cas particulier où les forces dérivent d'une énergie potentielle ne dépendant que des positions et où les coordonnées généralisées coïncident avec les coordonnées cartésiennes. La fonction de Lagrange s'écrit $L = T - U$ et les impulsions généralisées coïncident avec la définition habituelle $\mathbf{p}_\alpha = m_\alpha \mathbf{v}_\alpha$. La fonction de Hamilton, qui est identique à l'énergie totale, s'écrit alors $H = T + U$ et les équations de Hamilton s'écrivent:

$$\nabla_{\mathbf{p}_\alpha} H = \mathbf{v}_\alpha \quad (2.8)$$

$$\nabla_{\mathbf{r}_\alpha} H = -\dot{\mathbf{p}}_\alpha . \quad (2.9)$$

En écrivant $H = \sum_\alpha \mathbf{p}_\alpha^2 / 2m_\alpha + U$, on en déduit:

$$\mathbf{p}_\alpha / m_\alpha = \mathbf{v}_\alpha \quad (2.10)$$

$$\mathbf{F}_\alpha = \dot{\mathbf{p}}_\alpha . \quad (2.11)$$

La première de ces deux équations de Hamilton n'est autre que la définition de l'impulsion en termes de la vitesse. La deuxième n'est autre que le principe fondamental de la dynamique, où on a remplacé l'accélération par la dérivée par rapport au temps de l'impulsion. Dans ce cas très simple, l'écriture des équations de Hamilton à partir des équations de Newton revient à prendre comme variables indépendantes position et vitesse, un changement de variable bien connu pour transformer les équations différentielles du second ordre en équations du premier ordre.

Examinons maintenant un autre cas très important: celui de particules chargées dans un champ électromagnétique. Nous supposons encore que les coordonnées sont les coordonnées cartésiennes habituelles. Nous avons montré au chapitre précédent que la fonction de Lagrange s'écrit:

$$L = \sum_\alpha (T_\alpha - q_\alpha V_\alpha + q_\alpha \mathbf{A}_\alpha \cdot \mathbf{v}_\alpha) , \quad (2.12)$$

où V_α et \mathbf{A}_α sont les potentiels scalaires et vecteur vus par la particule α .

Pour passer aux équations de Hamilton, la première étape est de déterminer l'impulsion généralisée \mathbf{p}_α :

$$\mathbf{p}_\alpha = \nabla_{\mathbf{v}_\alpha} L = m_\alpha \mathbf{v}_\alpha + q_\alpha \mathbf{A}_\alpha . \quad (2.13)$$

On voit bien que, sauf dans le cas de l'électrostatique, l'impulsion généralisée ne coïncide pas avec la quantité de mouvement ordinaire.

On peut alors écrire sans difficultés la fonction de Hamilton:

$$H = \sum_\alpha \mathbf{p}_\alpha \cdot \mathbf{v}_\alpha - L = \sum_\alpha \left(\frac{1}{2} m_\alpha \mathbf{v}_\alpha^2 + q_\alpha V_\alpha \right) . \quad (2.14)$$

De façon très remarquable, la fonction de Hamilton écrite sous cette forme ne fait pas intervenir le potentiel vecteur magnétique. En effet, au moins dans le cas où les potentiels sont statiques, nous avons vu au chapitre précédent que l'énergie totale conservée des particules est la somme de leur énergie cinétique et de l'énergie potentielle électrostatique. On pourrait craindre qu'une telle fonction de Hamilton ne conduise à des équations du mouvement qui ne font plus intervenir le champ magnétique. Bien sûr il n'en est rien. On doit en effet, pour écrire les équations de Hamilton, exprimer H comme une fonction des impulsions généralisées, non pas des vitesses. En utilisant le lien entre impulsion et vitesse, on obtient:

$$H = \sum_\alpha \left[\frac{(\mathbf{p}_\alpha - q_\alpha \mathbf{A}_\alpha)^2}{2m_\alpha} + q_\alpha V_\alpha \right] . \quad (2.15)$$

Cette expression du hamiltonien de particules dans un champ nous sera utile dans la quatrième partie. Nous aurons en effet besoin du hamiltonien quantique d'un atome dans un champ de rayonnement. Nous utiliserons simplement l'expression précédente, en remplaçant toutes les quantités relatives à la particule par les opérateurs quantiques correspondants.

A titre d'exercice, écrivons les équations du mouvement à partir de cette expression du hamiltonien, et vérifions que nous retrouvons bien la force de Lorentz habituelle. Pour alléger les notations, nous ne considérerons que le cas d'une particule unique et nous omettrons donc l'indice α . Le lecteur pourra aisément rétablir le cas général.

Notons d'abord qu'une des équations de Hamilton redonne trivialement, comme dans le cas d'un potentiel standard, la définition de l'impulsion généralisée en termes de la vitesse. Nous nous focaliserons donc sur l'autre équation, qui s'écrit:

$$\dot{\mathbf{p}} = -\nabla H = -q\nabla V - \nabla \frac{(\mathbf{p} - q\mathbf{A})^2}{2m}, \quad (2.16)$$

où les gradients s'entendent par rapport à la position \mathbf{r} de la particule. Le gradient du carré scalaire peut se développer comme gradient d'un produit scalaire:

$$\nabla(\mathbf{p} - q\mathbf{A})^2 = 2\{[(\mathbf{p} - q\mathbf{A}) \cdot \nabla](\mathbf{p} - q\mathbf{A}) + (\mathbf{p} - q\mathbf{A}) \times [\nabla \times (\mathbf{p} - q\mathbf{A})]\}. \quad (2.17)$$

Dans cette équation, seul le potentiel vecteur est à considérer comme une fonction de \mathbf{r} . On peut alors simplifier cette expression et la dérivée de l'impulsion s'écrit, en faisant intervenir la vitesse de la particule:

$$\dot{\mathbf{p}} = -\nabla H = -q\nabla V + q(\mathbf{v} \cdot \nabla)\mathbf{A} + q\mathbf{v} \times \mathbf{B}. \quad (2.18)$$

En nous souvenant que:

$$-\nabla V = \mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \quad (2.19)$$

et que

$$(\mathbf{v} \cdot \nabla)\mathbf{A} = \frac{d\mathbf{A}}{dt} - \frac{\partial \mathbf{A}}{\partial t} \quad (2.20)$$

(dérivée hydrodynamique), on a

$$\dot{\mathbf{p}} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B} + q\frac{d\mathbf{A}}{dt}. \quad (2.21)$$

Si enfin on exprime la variation de l'impulsion en fonction de l'accélération, on retrouve bien le principe fondamental avec la force de Lorentz sous sa forme standard. Si nous ne doutions guère du résultat, il était important de l'établir explicitement. Il nous a fait bien sentir la différence essentielle entre quantité de mouvement (au sens de la dynamique newtonienne) et impulsion généralisée.

2.2 Crochets de Poisson

Nous allons introduire dans ce paragraphe une notation que nous n'utiliserons guère dans ce cours, très élémentaire, de mécanique analytique. En revanche, l'analogie et le lien formel très important entre ces crochets de Poisson et les commutateurs de la mécanique quantique rendent importante leur introduction à ce point. Nous préciserons d'ailleurs rapidement ces liens.

Considérons une fonction f quelconque des impulsions, des positions et du temps. On peut en écrire la dérivée totale par rapport au temps, le long de la trajectoire suivie par le système, sous la forme:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_i \frac{\partial f}{\partial p_i} \dot{p}_i + \sum_i \frac{\partial f}{\partial q_i} \dot{q}_i. \quad (2.22)$$

En utilisant les équations de Hamilton, nous mettrons cette dérivée sous la forme:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{H, f\} , \quad (2.23)$$

où $\{H, f\}$ est le crochet de Poisson de f avec H , défini par:

$$\{H, f\} = \sum_i \frac{\partial H}{\partial p_i} \frac{\partial f}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial f}{\partial p_i} . \quad (2.24)$$

Le crochet de Poisson est une forme bilinéaire antisymétrique de ses arguments. C'est aussi le cas des commutateurs entre opérateurs en mécanique quantique. Au delà d'une simple remarque mathématique et d'une évidente analogie de notations, nous allons, tout au long de ce paragraphe, rencontrer de nombreuses similitudes entre crochets de Poisson et commutateurs. En fait les uns sont la version classique des autres.

Notons également que les crochets de Poisson peuvent être utilisés pour décrire l'évolution temporelle d'une densité de probabilité dans l'espace des phases. En effet, si $W(p_i, q_i)$ représente la probabilité pour que les coordonnées généralisées et leurs impulsions conjuguées prennent les valeurs p_i, q_i , W évolue dans le temps selon l'équation:

$$\frac{\partial W}{\partial t} = \{H, W\} . \quad (2.25)$$

Cette équation est en fait le premier terme d'un développement en puissances de \hbar de l'évolution d'une fonction de distribution de quasi-probabilité dans l'espace des phase, très utilisée en mécanique quantique: la distribution de Wigner (développement de Moyal). Au premier ordre en \hbar , comme on pouvait s'y attendre, l'évolution quantique de la densité de probabilité dans l'espace des phases est la même que l'équation d'évolution classique.

Supposons que la fonction f ne fasse pas intervenir explicitement le temps. La condition nécessaire et suffisante pour que f soit une intégrale première du mouvement est alors que $\{H, f\} = 0$. La nullité du crochet de Poisson avec la fonction de Hamilton est équivalente avec la constance de la fonction. Cette propriété est à rapprocher de son équivalent quantique. La condition pour qu'un opérateur F soit une constante du mouvement est en effet que son commutateur $[H, F]$ avec le hamiltonien s'annule.

On peut bien sûr définir le crochet de Poisson de deux fonctions quelconques des positions des impulsions et du temps, estimées sur la trajectoire du système:

$$\{f, g\} = \sum_i \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} . \quad (2.26)$$

On déduit de cette expression les propriétés essentielles de l'algèbre des crochets de Poisson:

$$\{f, g\} = -\{g, f\} \quad (2.27)$$

$$\{f, C\} = 0 \quad (2.28)$$

$$\{f + f', g\} = \{f, g\} + \{f', g\} \quad (2.29)$$

$$\{Cf, g\} = C\{f, g\} \quad (2.30)$$

$$\{f, q_i\} = \frac{\partial f}{\partial p_i} \quad (2.31)$$

$$\{f, p_i\} = -\frac{\partial f}{\partial q_i} , \quad (2.32)$$

où C est une constante arbitraire. On en déduit facilement les crochets de Poisson des impulsions et positions:

$$\{q_i, q_k\} = 0 \quad (2.33)$$

$$\{p_i, p_k\} = 0 \quad (2.34)$$

$$\{p_i, q_k\} = \delta_{i,k} \quad (2.35)$$

Il est intéressant de noter la similitude entre ces “relations de commutation” classiques et leur analogue quantique ($[X_i, P_k] = i\hbar\delta_{ik}$). De façon très générale, nous verrons que les crochets de Poisson des quantités classiques coïncident, à un facteur $-i\hbar$ près, avec les commutateurs quantiques.

Notons une dernière propriété utile des crochets de Poisson, la relation de Jacobi:

$$\{f, \{g, k\}\} + \{g, \{k, f\}\} + \{k, \{f, g\}\} = 0 \quad (2.36)$$

(nullité de la somme de tous les crochets de Poisson obtenus par permutation circulaire des trois fonctions).

Comme nous l'avons déjà mentionné, les crochets de Poisson sont utiles pour la recherche des intégrales premières du mouvement. Précisons encore ce point en montrant que si f et g sont deux intégrales premières du mouvement, leur crochet de Poisson est aussi une intégrale première. En d'autres termes, nous allons montrer que, si $df/dt = 0$ et $dg/dt = 0$, alors $d\{f, g\}/dt = 0$. Nous avons:

$$\begin{aligned} \frac{d}{dt}\{f, g\} &= \frac{\partial}{\partial t}\{f, g\} + \{H, \{f, g\}\} \\ &= \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\} - \{f, \{g, H\}\} - \{g, \{H, f\}\} \\ &= \left\{ \frac{\partial f}{\partial t} + \{H, f\}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} + \{H, g\} \right\} \\ &= \left\{ \frac{df}{dt}, g \right\} + \left\{ f, \frac{dg}{dt} \right\} \\ &= 0 . \end{aligned} \quad (2.37)$$

Si donc on connaît deux intégrales premières du mouvement, on peut en trouver en principe une troisième en prenant leur crochet de Poisson. Rien ne garantit néanmoins que cette troisième intégrale ne soit triviale (nulle en particulier) ou déjà connue. Notons là encore qu'il existe un analogue quantique évident à cette propriété: si deux opérateurs commutent avec le hamiltonien (et sont donc des constantes du mouvement), alors leur commutateur commute lui aussi avec le hamiltonien et donne une troisième constante du mouvement.

Pour clore ce paragraphe, considérons brièvement le cas des trois composantes L_x, L_y et L_z du moment cinétique total \mathbf{L} . Nous nous placerons, pour fixer les idées, dans le cas d'une particule unique, et nous laisserons les généralisations au lecteur. Nous avons montré au chapitre précédent que $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. On en déduit $L_x = yp_z - zp_y$ (et les deux autres composantes par permutation circulaire des indices). Le crochet de Poisson $\{L_x, L_y\}$ fait intervenir les dérivées partielles des composantes concernées du moment cinétique par rapport aux composantes de la position et de l'impulsion. L'expression des composantes du moment cinétique permet de calculer facilement ces dérivées. Après un calcul sans grand intérêt, on obtient:

$$\{L_x, L_y\} = -L_z \quad (2.38)$$

et les trois relations se déduisant de celle ci par permutation circulaire des indices. Notons, une fois de plus, l'analogie entre ces équations et leur contrepartie quantique:

$$[L_x, L_y] = i\hbar L_z , \quad (2.39)$$

relations d'un très grande importance puisqu'elles définissent ce qu'est un moment cinétique quantique.

Nous avons vu au chapitre précédent que l'invariance par rotation arbitraire autour d'un axe impliquait la conservation de la composante du moment cinétique sur cet axe. Imaginons qu'on ait pu montrer la conservation de L_x et L_y . Il en résulte, comme nous venons de le montrer, la conservation de leur crochet de Poisson, c'est à dire de L_z . Il suffit en fait que le moment par rapport à deux axes perpendiculaires soit conservé pour que le moment cinétique total le soit.

Nous pouvons aussi appliquer ces relations à l'étude d'un problème dynamique important. Nous considérerons un système dynamique dont le hamiltonien peut s'écrire $H = \boldsymbol{\Omega} \cdot \mathbf{L}$, où $\boldsymbol{\Omega}$ est un vecteur constant. C'est par exemple le hamiltonien d'une particule dont le moment cinétique est proportionnel au dipôle magnétique, quand elle est plongée dans un champ magnétique uniforme (problème de la "précession de Larmor", dont la version quantique est d'une grande importance dans la compréhension de l'effet Zeeman). Les seules variables dynamiques importantes sont les composantes du moment cinétique. On peut écrire:

$$\frac{d\mathbf{L}}{dt} = \{H, \mathbf{L}\} = \{\boldsymbol{\Omega} \cdot \mathbf{L}, \mathbf{L}\} \quad (2.40)$$

En utilisant les relations de crochets de poisson entre les composantes de \mathbf{L} , on a, par exemple,

$$\frac{dL_x}{dt} = \Omega_y L_z - \Omega_z L_y \quad (2.41)$$

soit encore

$$\frac{d\mathbf{L}}{dt} = \boldsymbol{\Omega} \times \mathbf{L} . \quad (2.42)$$

On retrouve aisément que le moment cinétique précesse autour du vecteur fixe $\boldsymbol{\Omega}$ avec une pulsation Ω .

2.3 Action et hamiltonien

Nous allons revenir sur le principe variationnel. Nous allons écrire l'action en fonction du hamiltonien, et montrer qu'on peut retrouver les équations de Hamilton en écrivant la stationnarité de cette expression de l'action par rapport à certaines variations infinitésimales de la trajectoire. Ce paragraphe ne nous apprendra rien que nous ne sachions déjà, mais il est nécessaire à la cohérence de tout notre édifice.

Le lien entre fonction de Lagrange et fonction de Hamilton nous permet d'écrire l'action sous la forme:

$$S = \int_{t_1}^{t_2} \left(\sum_i p_i \dot{q}_i - H \right) dt , \quad (2.43)$$

avec les notations du premier chapitre.

Ecrivons maintenant, pour retrouver les équations de Hamilton, que l'action est stationnaire pour la trajectoire effectivement suivie. L'accroissement δS de l'action dans un accroissement infinitésimal de la trajectoire doit donc s'annuler. Nous considérerons donc deux trajectoires. La trajectoire de référence est la trajectoire effectivement suivie, définie par $q_i(t)$ et $p_i(t)$ (les variables naturelles dans une approche hamiltonienne sont les positions et les impulsions). L'autre trajectoire (infiniment proche) est définie à chaque instant par $q_i(t) + \delta q_i(t)$ et $p_i(t) + \delta p_i(t)$. Nous imposerons aux deux trajectoires de coïncider à l'instant initial et à l'instant final: $\delta q_i(1) = \delta q_i(2) = 0$. Il est naturel de considérer, dans une approche hamiltonienne, les p_i et les q_i comme des variables indépendantes. Nous n'imposerons donc aucune condition aux δp_i , ni aux extrémités de la trajectoire, ni à aucun instant. Il faut bien voir que nous pouvons ainsi considérer des trajectoires variées qui n'auraient pas de sens du point de vue de la simple cinématique: si les impulsions coïncident avec les quantités de mouvement ($m\mathbf{v}$), varier les vitesses indépendamment des positions implique que, sur la trajectoire variée, les vitesses puissent ne plus être égales aux dérivées des positions (avec une différence au premier ordre dans les petits accroissements). Ce n'est qu'au prix de cette liberté toute mathématique que nous pourrons retrouver l'équation de Hamilton qui contient en fait la définition des impulsions généralisées (de la vitesse dans le cas très simple que nous venons de mentionner).

Avec ces notations, l'accroissement de l'action s'écrit simplement:

$$\delta S = \int_{t_1}^{t_2} \left(\sum_i (\delta p_i \dot{q}_i + p_i \delta \dot{q}_i) - \delta H \right) dt . \quad (2.44)$$

L'accroissement de H entre la trajectoire de référence et la trajectoire variée s'écrit simplement:

$$\delta H = \sum_i \frac{\partial H}{\partial q_i} \delta q_i + \frac{\partial H}{\partial p_i} \delta p_i \quad (2.45)$$

On peut donc mettre δS sous la forme:

$$\delta S = \int_{t_1}^{t_2} \sum_i \delta p_i \left[\frac{dq_i}{dt} - \frac{\partial H}{\partial p_i} \right] dt + \int_{t_1}^{t_2} \sum_i \left[p_i \delta \dot{q}_i - \frac{\partial H}{\partial q_i} \delta q_i \right] dt . \quad (2.46)$$

Considérons plus particulièrement le premier terme de la deuxième intégrale. On peut l'intégrer par parties pour faire apparaître δq_i au lieu de $\delta \dot{q}_i$. Le terme tout intégré dans cette intégration par parties fait intervenir les accroissements δq_i aux instants t_1 et t_2 . Il est donc identiquement nul. En regroupant alors les termes proportionnels aux accroissements des positions et des impulsions, on peut écrire:

$$\delta S = \int_{t_1}^{t_2} \sum_i \left[\frac{dq_i}{dt} - \frac{\partial H}{\partial p_i} \right] \delta p_i dt - \int_{t_1}^{t_2} \sum_i \left[\frac{dp_i}{dt} + \frac{\partial H}{\partial q_i} \right] \delta q_i dt . \quad (2.47)$$

δS ne s'annulera quels que soient les accroissements des positions et impulsions que si les termes entre crochets dans chaque intégrale s'annulent identiquement sur la trajectoire effectivement suivie. On montre bien ainsi que cette trajectoire obéit effectivement aux équations de Hamilton, que nous aurions parfaitement pu établir par ce raisonnement.

Il existe, dans le cadre de la formulation hamiltonienne, un autre principe variationnel, le principe de Maupertuis, qui permet de trouver la forme de la trajectoire, mais pas la loi horaire. Ce principe s'apparente de très près à celui de Fermat, qui permet de déterminer en optique la trajectoire des rayons lumineux. Nous n'aborderons pas ici le principe de Maupertuis, largement discuté dans les manuels³.

2.4 Transformations canoniques

Nous abordons ici ce qui constitue sans doute l'intérêt essentiel de la formulation hamiltonienne. La complète symétrie entre les positions et impulsions va en effet nous permettre d'envisager des changements de variables mêlant ces deux types de quantités. Nous verrons qu'on peut ainsi rendre complètement triviale la dynamique d'un problème, en prenant par exemple comme nouvelles variables les conditions initiales. Bien sûr, la difficulté est de trouver le bon changement de variables. Nous n'aborderons pas, faute de place, les méthodes permettant de les déterminer. Nous profiterons également de ce paragraphe pour établir le théorème de Liouville, qui joue un rôle central en physique statistique et dans l'étude de la dynamique des systèmes complexes.

2.4.1 Principe

Revenons un instant au formalisme lagrangien. Les q_i et leurs vitesses \dot{q}_i ne sont pas des variables indépendantes. Le seul type de changement de variable que l'on puisse envisager est donc de définir de nouvelles coordonnées généralisées Q_i , calculables à partir des q_i et du temps (la relation étant inversible). On peut alors écrire la fonction de Lagrange en fonction des Q_i , les équations de Lagrange correspondantes et résoudre ainsi le problème – en espérant qu'il soit plus simple en terme des nouvelles variables. Nous appellerons une telle transformation des coordonnées généralisées une "transformation ponctuelle". Nous allons maintenant voir qu'elle appartient à un cadre beaucoup plus général.

Dans le formalisme hamiltonien, les variables indépendantes sont les q_i et les p_i . Nous pouvons imaginer un changement de variables très général sous la forme $q_i \rightarrow Q_i(q_i, p_i, t)$ et $p_i \rightarrow P_i(q_i, p_i, t)$.

³On consultera en particulier le Goldstein.

Les définitions des nouvelles positions et des nouvelles impulsions font intervenir toutes les anciennes positions et impulsions. Ce changement de variables nous laisse donc toute liberté. Nous exigerons cependant qu'il soit inversible, pour que l'état du système soit déterminé de façon univoque par les nouvelles coordonnées, ce qui impose que les Q_i et les P_i soient indépendantes.

Pour que ce changement de variable soit utilisable dans un problème de mécanique, il nous faut cependant lui imposer une contrainte supplémentaire. Il faut en effet qu'il existe un nouveau hamiltonien $H'(Q_i, P_i, t)$ qui donne les équations de Hamilton pour les nouvelles variables. Nous dirons alors que la transformation est "canonique".

Supposons que H' existe bien. Le principe variationnel abordé au paragraphe précédent s'écrit :

$$\delta \int_{t_1}^{t_2} \left(\sum_i p_i \dot{q}_i - H \right) dt = 0 \quad (2.48)$$

en terme des anciennes variables et

$$\delta \int_{t_1}^{t_2} \left(\sum_i P_i \dot{Q}_i - H' \right) dt = 0 \quad (2.49)$$

en termes des nouvelles. Pour que ces deux principes variationnels donnent les mêmes équations du mouvement, il suffit que les deux quantités intégrées ne diffèrent que de la dérivée totale par rapport au temps d'une fonction F . En effet, cette différence ne contribue aux intégrales que par un terme de la forme $F(2) - F(1)$, qui ne varie pas quand on varie la trajectoire. On peut mettre cette condition sous la forme :

$$\sum_i p_i dq_i - H dt = \sum_i P_i dQ_i - H' dt + dF, \quad (2.50)$$

dF étant l'accroissement de F entre deux instants voisins. Il suffit donc, pour que la transformation soit canonique, qu'il existe une fonction F telle que :

$$dF = \sum_i p_i dq_i - \sum_i P_i dQ_i + (H' - H) dt \quad (2.51)$$

ou encore :

$$\frac{\partial F}{\partial q_i} = p_i, \quad -\frac{\partial F}{\partial Q_i} = P_i, \quad H' = H + \frac{\partial F}{\partial t} \quad (2.52)$$

Si effectivement la donnée d'une fonction $F(q_i, Q_i, t)$ permet, en écrivant les relations aux dérivées partielles précédentes, de déterminer complètement la transformation, celle-ci sera une transformation canonique et le nouveau hamiltonien sera connu.

Pour montrer que la donnée de F détermine complètement la transformation, considérons les n équations $\partial F(q_i, Q_i, t)/\partial q_i = p_i$. On peut, au moins formellement et sauf cas pathologique, les résoudre en termes des n "inconnues" Q_i et trouver ainsi les n fonctions $Q_i(q_i, p_i, t)$. On peut alors calculer facilement, en fonction des q_i, Q_i et de t les n dérivées partielles $-\partial F/\partial Q_i$ qui donnent les nouvelles impulsions $P_i(q_i, Q_i, t)$. En reportant les expressions des Q_i , on trouve alors les $P_i(q_i, p_i, t)$, ce qui achève de déterminer complètement la transformation. Le nouveau hamiltonien H' peut alors être calculé et on peut écrire les équations de Hamilton en termes des nouvelles variables.

En résumé, la donnée d'une fonction $F(q_i, Q_i, t)$ détermine en général de façon univoque une transformation généralisée et assure que cette transformation soit canonique. Enfin, la donnée de F permet d'exprimer le nouveau hamiltonien. Pour toutes ces raisons F s'appelle la fonction génératrice de la transformation⁴.

⁴Nous avons montré qu'à toute fonction génératrice correspond une transformation (sauf cas pathologique, certaines fonctions ne définissant pas une transformation univoque – en particulier les constantes). Nous n'avons pas établi la réciproque, que nous admettrons.

Nous avons choisi ici, de façon naturelle, de définir F par les variables q_i et Q_i . On peut imaginer de définir une transformation canonique par un autre couple de variables. F étant donnée, considérons en effet la fonction Φ définie par:

$$\Phi = F + \sum_i P_i Q_i , \quad (2.53)$$

et résultant donc d'une transformation de Legendre sur la fonction F . ϕ est donc a priori une fonction naturelle des q_i, P_i . On peut s'en convaincre en écrivant sa différentielle:

$$d\phi = \sum_i p_i dq_i + \sum_i Q_i dP_i + (H' - H)dt . \quad (2.54)$$

Les dérivées partielles de ϕ sont donc:

$$\frac{\partial \phi}{\partial q_i} = p_i, \quad \frac{\partial \phi}{\partial P_i} = Q_i, \quad H' = H + \frac{\partial \phi}{\partial t} . \quad (2.55)$$

Comme dans le cas précédent, on peut montrer aisément que la donnée de ϕ et les relations aux dérivées partielles ci-dessus déterminent complètement la transformation canonique. Pour cela, on résout les n équations $\partial \phi / \partial q_i = p_i$ en termes des $P_i(q_i, p_i, t)$. En reportant ces expressions dans les n autres relations, on achève de déterminer la transformation en obtenant les Q_i .

Nous laissons au lecteur le soin de montrer qu'il existe encore deux expressions possibles pour une transformation canonique. L'une fait intervenir $\Psi(p_i, Q_i, t) = F - \sum_i q_i p_i$ et les relations:

$$\frac{\partial \Psi}{\partial p_i} = -q_i, \quad \frac{\partial \Psi}{\partial Q_i} = -P_i \quad H' = H + \frac{\partial \Psi}{\partial t} . \quad (2.56)$$

L'autre utilise la fonction $\Xi(p_i, P_i, t) = \phi - \sum_i Q_i P_i$ et les relations:

$$\frac{\partial \Xi}{\partial p_i} = -q_i, \quad \frac{\partial \Xi}{\partial P_i} = Q_i \quad H' = H + \frac{\partial \Xi}{\partial t} . \quad (2.57)$$

La donnée au choix de l'une de ces quatre fonctions détermine donc complètement une transformation canonique. Si la fonction génératrice ne fait pas intervenir explicitement le temps, les fonctions de Hamilton coïncident dans les anciennes et les nouvelles représentations (il existe en fait un lien très profond entre les transformations canoniques et les changements de représentation ou les transformations unitaires en mécanique quantique). Le choix immense des fonctions génératrices possibles donne une idée de la puissance de la méthode.

2.4.2 Transformations canoniques et crochets de Poisson

En fait la propriété essentielle des transformations canoniques est qu'elles conservent les crochets de Poisson:

$$\{f, g\}_{p,q} = \{f, g\}_{P,Q} , \quad (2.58)$$

où

$$\{f, g\}_{p,q} = \sum_i \frac{\partial f(p_i, q_i, t)}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} , \quad (2.59)$$

et

$$\{f, g\}_{P,Q} = \sum_i \frac{\partial f(P_i, Q_i, t)}{\partial P_i} \frac{\partial g}{\partial Q_i} - \frac{\partial f}{\partial Q_i} \frac{\partial g}{\partial P_i} . \quad (2.60)$$

On peut bien sûr démontrer cette relation en utilisant la fonction génératrice de la transformation. Il existe cependant une démonstration beaucoup plus intuitive. Limitons nous pour cela au cas où f et g ne dépendent pas explicitement du temps et au cas où la fonction génératrice de la transformation ne dépend pas non plus du temps. Il doit exister un problème de mécanique décrit par les q_i, p_i dont la

fonction $g(p_i, q_i)$ serait le hamiltonien. Le crochet de Poisson $\{f, g\}_{p,q}$ apparaît alors comme la dérivée temporelle de la fonction f le long d'un trajectoire solution de ce problème fictif. La fonction $g(P_i, Q_i)$ est le hamiltonien H' du même problème dans les coordonnées transformées. Le crochet de Poisson $\{f, g\}_{P,Q}$ est la dérivée temporelle de f dans cette représentation. Mais df/dt doit être indépendante de la représentation, ce qui implique l'égalité des crochets de Poisson. Cette relation, établie ici dans un cas un peu particulier, est en fait générale. Elle permet, par exemple, de calculer facilement les crochets de poisson des nouvelles variables dans la représentation des anciennes, ce que nous laisserons au lecteur à titre d'exercice.

2.4.3 Exemples de transformations canoniques

Nous donnons ici, à titre d'illustration, quelques fonctions génératrices définissant des transformations canoniques particulièrement simples. Nous montrerons aussi comment les transformations canoniques permettent de rendre triviale la solution d'un problème de mécanique.

- $\Phi(q_i, P_i, t) = \sum_i q_i P_i$. Le temps n'intervenant pas dans la fonction génératrice, nous avons bien sûr $H' = H$. En appliquant les relations aux dérivées partielles, nous trouvons sans difficultés $p_i = P_i$ et $Q_i = q_i$. Cette fonction génératrice définit donc la transformation unité, ce qui ne présente guère d'intérêt.
- $\Phi(q_i, P_i, t) = \sum_i \phi_i(q_j, t) P_i$. Là encore, on trouve sans difficultés $Q_i = \phi_i(q_j, t)$. Cette fonction génère donc l'ensemble des transformations ponctuelles, éventuellement dépendantes du temps, qui définissent les nouvelles coordonnées en fonction seulement des anciennes. L'avantage de l'approche en termes de transformations canoniques est que l'application des autres relations aux dérivées partielles nous donne $p_i = \sum_k (\partial \phi_k / \partial q_i) P_k$, système qui permet de déterminer les nouvelles impulsions en fonction des anciennes et des coordonnées. Notons que, si les ϕ_i dépendent du temps, les deux hamiltoniens H et H' peuvent différer.
- $F = \sum_i (q_i Q_i)$. On trouve alors immédiatement $p_i = Q_i$ et $P_i = -q_i$. A un signe près, cette fonction réalise l'échange des coordonnées et des impulsions, illustrant le rôle très symétrique que jouent ces notions dans l'approche hamiltonienne.

Nous allons montrer maintenant comment une transformation canonique bien choisie permet de rendre complètement triviale la dynamique d'un système. L'idée est de rendre le nouveau hamiltonien cyclique dans les nouvelles coordonnées. Les nouvelles impulsions sont alors constantes et la dynamique des nouvelles coordonnées se résume à une évolution linéaire dans le temps. La difficulté dans ce genre d'approche, à la base de nombreuses méthodes de résolution de problèmes de mécanique, est bien sûr d'exhiber la transformation canonique convenable, ce qui n'est pas toujours possible. Nous considérerons dans ce paragraphe le problème trivial d'un oscillateur harmonique à une dimension.

Le hamiltonien H s'exprime simplement en fonction de la coordonnée q et de l'impulsion conjuguée p par:

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2, \quad (2.61)$$

somme des énergies cinétiques et potentielles. On considère alors la transformation canonique générée par la fonction

$$F(q, Q, t) = \frac{1}{2} m\omega q^2 \cot Q \quad (2.62)$$

(le choix d'une telle transformation n'est guère possible si on ne connaît déjà la solution du problème). Cette fonction ne dépendant pas explicitement du temps, le nouveau et l'ancien hamiltonien coïncident. A partir de cette fonction génératrice, on trouve

$$p = \frac{\partial F}{\partial q} = m\omega q \cot Q \quad (2.63)$$

et

$$P = -\frac{\partial F}{\partial Q} = \frac{m\omega q^2}{2\sin^2 Q} \quad (2.64)$$

On peut extraire de ces deux relations:

$$q = \sqrt{\frac{2P}{m\omega}} \sin Q, \quad p = \sqrt{2m\omega P} \cos Q. \quad (2.65)$$

Le nouveau hamiltonien peut alors être exprimé, comme il se doit, en termes des nouvelles variables:

$$H' = \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2 = \omega P. \quad (2.66)$$

Comme nous l'espérons, ce nouveau hamiltonien est cyclique dans la nouvelle coordonnée Q . L'impulsion conjuguée, P , est donc une constante. Comme l'énergie mécanique totale est conservée dans ce problème et coïncide avec la fonction de Hamilton, cette constante vaut simplement:

$$P = \frac{\mathcal{E}}{\omega}. \quad (2.67)$$

Elle est donc homogène à une action (produit d'une énergie par un temps). P s'appelle donc variable d'action. On montre que, dans tout problème unidimensionnel, on peut trouver une variable d'action conservée dans l'évolution. L'équation de Hamilton pour Q , $\dot{Q} = \partial H' / \partial P = \omega$, donne simplement:

$$Q = \omega t + \phi. \quad (2.68)$$

Q , évoluant linéairement avec le temps, se nomme variable d'angle. Là encore, dans tout problème unidimensionnel, il existe une variable d'angle conjuguée de la variable d'action. La solution explicite du problème est donc donnée en fonction de deux constantes arbitraires, comme il se doit, l'énergie mécanique totale \mathcal{E} et la phase ϕ , valeur initiale de la nouvelle coordonnée. En utilisant alors la transformation, on peut exprimer la solution en termes des variables initiales, et on trouve:

$$q = \sqrt{\frac{2\mathcal{E}}{m\omega^2}} \sin(\omega t + \phi), \quad (2.69)$$

ce qui n'est pas vraiment un résultat inattendu!

L'apparente simplicité de cette approche, dans ce cas trivial, ne doit pas faire oublier que la grande difficulté est d'exhiber la fonction génératrice adaptée. Nous ne pourrions aborder ici les méthodes variées de résolution fondées sur les transformations canoniques. Le lecteur pourra en trouver une description détaillée dans le Goldstein.

2.4.4 Transformations canoniques et espace des phases

L'état mécanique du système est complètement décrit par la donnée des n q_i et des n p_i . Autrement dit, le système est décrit comme un point dans un espace à $2n$ dimensions que l'on appelle *espace des phases*. Cet espace joue un très grand rôle en physique statistique. L'entropie, par exemple, peut être définie comme le logarithme du nombre de configurations accessibles au système. Compter ce nombre de configurations, c'est compter la surface de l'espace des phases correspondant à un petit intervalle d'énergie. L'espace des phases joue également un rôle très important dans l'étude de la dynamique complexe des systèmes (chaos classique, par exemple). Un point d'équilibre stable correspond à un point dans l'espace des phases, un mouvement périodique à une trajectoire fermée simple, un mouvement chaotique à une trajectoire complexe parcourant rapidement tout le domaine accessible.

Les transformations canoniques transforment un espace des phases en un autre. La géométrie de cette transformation n'est pas complètement arbitraire, en raison des contraintes imposées aux

transformations canoniques. La propriété essentielle est qu’une transformation canonique conserve le volume dans l’espace des phases. Si on considère un domaine V de l’espace des phases des “anciennes coordonnées”, il lui correspond un domaine V' dans le nouvel espace. Pour des raisons évidentes de continuité et dérivabilité des transformations canoniques, V' est un fermé connexe si V l’est. On peut calculer le volume \mathcal{V} du domaine V comme:

$$\mathcal{V} = \int_V dq_1 \cdots dq_n dp_1 \cdots dp_n, \quad (2.70)$$

et calculer de même le volume \mathcal{V}' du domaine V' .

Pour une transformation canonique, $\mathcal{V} = \mathcal{V}'$.

Pour les mathématiciens cette propriété découle de façon immédiate de la structure symplectique de la transformation canonique. Nous allons établir cette propriété de façon moins directe, mais peut être plus accessible. Rappelons d’abord que:

$$\int_{V'} dQ_1 \cdots dQ_n dP_1 \cdots dP_n = \int_V |J| dq_1 \cdots dq_n dp_1 \cdots dp_n, \quad (2.71)$$

où J est le Jacobien du changement de variable, déterminant formé avec toutes les dérivées partielles des nouvelles variables par rapport aux anciennes:

$$J = \begin{vmatrix} \frac{\partial Q_1}{\partial q_1} & \cdots & \frac{\partial P_n}{\partial q_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial Q_1}{\partial p_n} & \cdots & \frac{\partial P_n}{\partial p_n} \end{vmatrix}, \quad (2.72)$$

que nous noterons également

$$J = \frac{\partial(Q_1, \dots, P_n)}{\partial(q_1, \dots, p_n)}. \quad (2.73)$$

Ces notations deviennent tout à fait triviales en dimension 1 et coïncident alors avec les changements de variables standard dans les intégrales. Si nous prouvons que le Jacobien de toute transformation canonique est 1, nous aurons établi la proposition cherchée.

Pour cela, nous aurons besoin de deux propriétés des Jacobiens. D’abord:

$$\frac{\partial(Q_1, \dots, P_n)}{\partial(q_1, \dots, p_n)} = \frac{\frac{\partial(Q_1, \dots, Q_n, P_1, \dots, P_n)}{\partial(q_1, \dots, q_n, P_1, \dots, P_n)}}{\frac{\partial(q_1, \dots, q_n, p_1, \dots, p_n)}{\partial(q_1, \dots, q_n, P_1, \dots, P_n)}}. \quad (2.74)$$

En fait, comme les dérivées partielles ordinaires, les produits et rapports de Jacobiens peuvent se simplifier comme des fractions⁵. De plus:

$$\frac{\partial(Q_1, \dots, Q_n, P_1, \dots, P_n)}{\partial(q_1, \dots, q_n, P_1, \dots, P_n)} = \frac{\partial(Q_1, \dots, Q_n)}{\partial(q_1, \dots, q_n)}. \quad (2.75)$$

On peut donc retirer d’un Jacobien les variables qui apparaissent au “numérateur” et au “dénominateur”. Le changement de variables considéré laisse en effet invariante ces quantités.

En utilisant successivement ces deux propriétés, on met le Jacobien de la transformation canonique sous la forme:

$$J = \frac{\frac{\partial(Q_1, \dots, Q_n)}{\partial(q_1, \dots, q_n)}}{\frac{\partial(p_1, \dots, p_n)}{\partial(P_1, \dots, P_n)}}. \quad (2.76)$$

⁵Nous supposons bien sûr que tous les Jacobiens écrits dans ces équations ont un sens, et en particulier que toutes les transformations “tronquées” sont inversibles, ce qui n’est pas vrai en toute généralité.

La transformation étant canonique, elle est engendrée par une fonction $\Phi(q_i, P_i)$ telle que $\partial\Phi/\partial q_i = p_i$ et $\partial\Phi/\partial P_i = Q_i$. On a donc immédiatement:

$$\frac{\partial Q_i}{\partial q_j} = \frac{\partial^2 \Phi}{\partial q_j \partial P_i} \quad (2.77)$$

et

$$\frac{\partial p_i}{\partial P_j} = \frac{\partial^2 \phi}{\partial P_j \partial q_i} . \quad (2.78)$$

Les dérivées secondes croisées étant égales, toutes les dérivées partielles apparaissant dans le développement du déterminant au numérateur de (2.76) sont égales, terme à terme, à celles apparaissant dans le développement du dénominateur. Cela établit que le Jacobien d'une transformation canonique est de module égal à un et l'invariance du volume dans l'espace des phases.

2.4.5 Transformation générée par l'action et théorème de Liouville

Nous considérerons dans ce paragraphe une transformation canonique très particulière qui fait se correspondre deux états du système à deux instants différents. Considérons en effet un système ayant un lagrangien indépendant du temps (l'énergie totale est donc conservée) dont la dynamique est décrite par les $p_i(t)$ et les $q_i(t)$ et considérons le changement de variables défini par:

$$p_i(t) \longrightarrow P_i(t) = p_i(t + T) \quad (2.79)$$

$$q_i(t) \longrightarrow Q_i(t) = q_i(t + T) , \quad (2.80)$$

où T est une durée fixe. Cette correspondance entre états du système à des instants différents est évidemment une transformation canonique, puisque les nouvelles variables obéissent aux mêmes équations du mouvement que les anciennes. Nous allons voir que la fonction génératrice de cette transformation n'est autre que l'action. Considérons en effet la trajectoire du système entre les instants t et $t + T$ et une trajectoire infiniment voisine obtenue en modifiant les coordonnées aux points de départ et d'arrivée de quantités infinitésimales. Avec les résultats du chapitre précédent, nous pouvons écrire la variation de l'action entre ces deux trajectoires comme:

$$dS = \sum_i p_i(t + T) dq_i(t + T) - \sum_i p_i(t) dq_i(t) \quad (2.81)$$

ou encore

$$dS = \sum_i P_i(t) dQ_i(t) - \sum_i p_i(t) dq_i(t) . \quad (2.82)$$

Cette expression de la différentielle de l'action, que nous pouvons considérer comme une fonction des coordonnées des points de départ et d'arrivée: $S(q_i, Q_i)$, prouve que:

$$\frac{\partial S}{\partial Q_i} = P_i \quad \frac{\partial S}{\partial q_i} = -p_i . \quad (2.83)$$

S est donc bien fonction génératrice de la transformation des q_i en Q_i . Cette transformation joue un rôle central dans la méthode de Hamilton–Jacobi, essentielle pour la résolution de problèmes complexes, et conduisant à la notion importante de séparabilité des variables. Nous ne disposons pas d'un espace suffisant pour traiter convenablement cette méthode. Nous donnerons donc une seule application de la transformation engendrée par l'action.

Considérons un domaine V de l'espace des phases du système. On peut considérer l'ensemble des trajectoires originaires d'un point situé à l'instant t à l'intérieur de ce domaine. Par continuité, ces trajectoires correspondent à l'instant $t + T$ à des points situés dans un nouveau domaine V' de l'espace des phases. Comme la transformation faisant se correspondre les instants t et $t + T$ est canonique,

l'étendue du domaine V' est égale à celle du domaine V . Cette propriété constitue le théorème de Liouville:

Le volume du domaine occupé dans l'espace des phases par un ensemble de trajectoires se conserve au cours du temps.

Ce théorème joue évidemment un rôle important en mécanique statistique. Il prédit, par exemple, la conservation de l'entropie dans une évolution hamiltonienne. Notons que ce théorème ne tient pas en présence de dissipation. N'importe quelle condition initiale conduit en effet à un état de repos où les coordonnées n'évoluent plus. Soulignons aussi le lien entre ce théorème et le théorème de conservation de l'étendue en optique. L'étendue joue le rôle du volume dans un espace des phases {position des rayons/angle }.

Ce paragraphe clôt notre exposé de mécanique analytique. Nous n'avons pas, de loin, donné un exposé exhaustif de ce sujet. Les grands domaines que nous n'aborderons pas sont les méthodes de résolution, telles que la méthode de Hamilton Jacobi. Nous ne dirons rien, non plus, des méthodes de perturbations classiques, si utiles en astronomie. Nous évoquerons brièvement dans la prochaine partie les extensions du formalisme lagrangien à des coordonnées continues (en un mot à des champs), mais sans épuiser non plus ce très vaste sujet. Finalement, nous ne saurions trop recommander au lecteur de se rapporter aux manuels de mécanique quantique pour explorer les liens très profonds entre mécanique quantique et mécanique analytique. Si une présentation de la mécanique quantique à partir de la dynamique classique et de sa quantification canonique n'est pas à recommander pour une première approche de la mécanique quantique, elle est extrêmement enrichissante à un niveau plus avancé.

Appendice 1

Modèle de Bohr

Nous traiterons dans cet appendice du premier modèle réaliste de structure atomique, celui de Bohr en 1913. Il nous sera en effet nécessaire à plusieurs endroits du cours pour des discussions qualitatives du rayonnement atomique. Comme ce modèle fait explicitement référence aux concepts d'action et a donné lieu, au cours de son évolution, à des développements élégants de mécanique classique, il a tout naturellement sa place dans ce cours de mécanique analytique. Nous commencerons par un bref rappel de la situation historique au moment de la formulation du modèle de Bohr. Nous l'exposerons ensuite dans un deuxième paragraphe. Dans un dernier paragraphe, nous préciserons ses limitations et les tentatives, menées entre autres par Bohr et Sommerfeld, pour raffiner le modèle et l'appliquer à d'autres atomes que l'hydrogène. Enfin, nous rappellerons brièvement, pour mémoire, les résultats quantiques rigoureux.

1.1 Un peu d'histoire

Les difficultés de la mécanique classique et les débuts de la mécanique quantique, au tournant du siècle, sont dues à essentiellement deux problèmes: le rayonnement du corps noir et la structure des spectres d'émission ou d'absorption des vapeurs.

Quand on traite en thermodynamique classique le rayonnement d'un corps complètement absorbant en équilibre thermique, on trouve la célèbre loi de Rayleigh Jeans. La densité de puissance spectrale du rayonnement (quantité d'énergie par unité de volume et de fréquence) est proportionnelle au carré de la fréquence. La quantité totale d'énergie électromagnétique contenue dans un corps en équilibre devrait donc être gravement infinie. Ce n'est bien sûr pas le cas et les données expérimentales, relativement précises à la fin du siècle dernier, donnaient un spectre décroissant rapidement à haute fréquence. Pour expliquer ce spectre, Planck introduisit, en 1900, une hypothèse de quantification. Les échanges d'énergie entre matière et rayonnement ne peuvent se faire, à une fréquence donnée, que par multiples entiers d'une quantité fondamentale, proportionnelle à la fréquence, selon la fameuse relation $E = h\nu$. La constante de Planck, h , une fois ajustée aux données expérimentales, l'accord entre les spectres calculés et les spectres expérimentaux se révélait excellent. En fait, Planck considérait cette hypothèse comme heuristique et doutait de sa signification physique. Ce n'est qu'avec Einstein, 5 ans plus tard, que l'idée de quantification de l'énergie électromagnétique fit une avancée notable avec l'introduction de quanta lumineux. Ceux-ci, qu'on devait appeler plus tard photons, sont nécessaires pour analyser, au-delà des valeurs moyennes, les fluctuations du rayonnement. On peut alors interpréter convenablement les propriétés de l'effet photoélectrique, ce qui valut son Nobel à Einstein (la relativité générale paraissait peut-être trop audacieuse pour être couronnée).

L'histoire de la spectroscopie est pour sa part très riche. Dès 1802, Wollaston, avec un spectrographe à prisme, observait des bandes sombres bien résolues dans le spectre solaire, bandes qui ne

sont pas présentes dans le spectre des corps chauffés¹. C'est sans doute là la première observation d'un spectre d'absorption atomique. Wollaston décrit en particulier une bande très intense dans le jaune qui devait être la raie principale du sodium. En 1817, Fraunhofer, encore très jeune, raffine ces mesures et observe les raies "de Balmer" du spectre de l'hydrogène (l'appellation Balmer est beaucoup plus récente comme on le verra). Il invente peu après le réseau de diffraction et l'histoire de la spectroscopie de précision commence. Il mesure avec grand soin (et avec une précision de l'ordre de 10^{-4}) la position des raies de Balmer. La nature de ces raies reste toutefois très controversée. En parallèle avec ces études du spectre solaire, l'étude du spectre d'émission de décharges dans les gaz se poursuit. Masson, en 1851, franchit un pas important en montrant que certaines raies peuvent être attribuées sans ambiguïté à la présence d'hydrogène. C'est la première identification fiable d'un élément avec un spectre de raies. Un nouveau pas important est franchi avec Ångström, qui remarque et explique la coïncidence des raies d'émission de l'hydrogène avec les raies d'absorption dans le spectre solaire. Il conclut que tout élément peut aussi bien absorber ou émettre de la lumière sur une de ses fréquences propres. Des atlas détaillés du spectre solaire sont publiés ensuite par Ångström, Rowland et Huggins (c'est sur les travaux de ce dernier que Balmer s'est, semble-t-il, appuyé).

On a cherché dans le même temps, en manipulant les données spectrales, à dégager des lois auxquelles pourraient obéir les fréquences des raies spectrales, en particulier pour l'hydrogène dont le spectre est simple. On pensa ainsi, pendant un temps, que les différentes fréquences émises par l'hydrogène pourraient être des harmoniques d'une fréquence fondamentale (travaux de Stoney en 1871). Les coïncidences numériques supportant cette approche s'évanouirent rapidement avec les progrès de l'instrumentation et en particulier ceux des spectrographes à réseau.

Un progrès important fut accompli par Balmer en 1885. Déjà âgé, simple instituteur, il manipule les données sur le spectre de l'hydrogène. Il observe que les longueurs d'onde des raies visibles du spectre de l'hydrogène sont proportionnelles à des fractions rationnelles simples faisant intervenir les carrés des nombres entiers, sous la forme $m^2/(m^2 - n^2)$, avec $n = 2$. Ce travail remarquable fut complété par celui de Rydberg, en 1889, qui remarque qu'il vaut mieux considérer les nombres d'ondes (inverses de longueurs d'onde). Les nombres d'ondes des raies de nombreux éléments s'obtiennent en effet simplement comme différences de termes en R/n^2 où n est entier et où R est la maintenant célèbre constante de Rydberg. Le principe de combinaison de Ritz, formulé en 1908, généralise ce travail. La découverte de nouvelles séries de raies de l'hydrogène s'accordant avec les formules de Balmer ($n = 3$ par Paschen en 1908, $n = 1$ (ultraviolet lointain) par Lyman en 1916, $n = 4$ par Brackett en 1922, $n = 5$ par Pfund en 1924, $n = 6$ par Humphrey en 1953) apporta au cours du temps des confirmations remarquables de la formule de Balmer. Bien sûr, la découverte par Michelson et Morley (encore eux) de sous-structures dans les raies de l'hydrogène (nous dirions aujourd'hui de la structure fine) complique un peu le tableau, mais le succès des formules de Balmer ou de Rydberg demeure.

S'il existe une relation aussi simple entre les fréquences, on doit chercher un modèle physique qui les prédise correctement. Le premier modèle "réaliste" de la structure de l'atome d'hydrogène est dû à Thomson, découvreur de l'électron en 1897 au Cavendish Laboratory, fondé par Maxwell environ 20 ans avant (en fait la découverte de l'électron pourrait aussi bien être attribuée à Wiechert, qui travaillait indépendamment). Comme on savait que la matière contient des électrons, Thomson imagine un atome constitué d'une gelée, uniformément chargée positivement, de forme sphérique, dans laquelle se déplacent des électrons indépendants. Le champ électrique produit par la gelée étant proportionnel à la distance, les électrons sont élastiquement liés et effectuent un mouvement harmonique de fréquence donnée. Nous utiliserons assez largement ce modèle très simple dit du "plum-pudding" dans le chapitre sur le rayonnement des sources atomiques. Nous verrons qu'il prédit correctement de nombreux ordres de grandeur.

Ce modèle, en dépit de certains succès, dut être abandonné après les expériences de déviation de

¹Cette brève histoire de la spectroscopie s'inspire d'un article de G.W. Series, dans *The Hydrogen Atom*, Bassani et al. éditeurs, Springer, 1989.

particules α dans des feuilles d'or. En interprétant ces résultats, Rutherford, lui aussi directeur du Cavendish, fut conduit en 1910 à admettre la présence dans la matière de charges positives extrêmement localisées. Il fallait donc renoncer au modèle du plum-pudding et venir à un modèle planétaire de la structure atomique, avec des électrons orbitant sous l'influence de la force de Coulomb autour d'un noyau pratiquement ponctuel. Bien évidemment, un tel système rayonnerait ou absorberait à la fréquence de révolution de l'électron.

Ce modèle présente des difficultés sérieuses. La première est que rien a priori ne fixe les paramètres de l'orbite et donc la fréquence d'émission. On pourrait donc s'attendre à voir les atomes rayonner ou absorber des longueurs d'onde arbitraires. De façon plus grave, ce modèle est manifestement instable. En rayonnant, l'électron en mouvement perd de l'énergie et le rayon de son orbite diminue (nous ferons le calcul explicitement dans la quatrième partie). La fréquence du mouvement augmentant, la perte d'énergie et la chute vers le noyau deviennent de plus en plus rapide. En quelques dizaines de picosecondes, tous les électrons de l'univers auraient dû tomber sur leur noyau en émettant un bref flash de radiation ultraviolette. Cette catastrophe ultraviolette ne s'étant pas produite encore, il faut y voir un grave défaut du modèle.

1.2 Modèle de Bohr

L'attitude de Bohr est tout à la fois pragmatique et extrêmement audacieuse. Puisque rien ne peut expliquer la stabilité et le caractère discret de la structure atomique, c'est qu'il faut introduire dans le modèle une condition supplémentaire de "quantification". Dans cette démarche, Bohr était sans doute guidé par les résultats de Planck. Puisque la constante de Planck décrit la quantification pour le rayonnement électromagnétique, il est assez naturel de tenter de l'utiliser pour la structure atomique. Cette constante ayant la dimension d'une action, il est naturel aussi de quantifier l'action de l'électron sur sa trajectoire².

Nous poserons donc que l'action, calculée sur une orbite, est un multiple entier (évidemment non nul) de la constante de Planck:

$$S = nh . \quad (1.1)$$

Nous allons utiliser cette condition pour déterminer l'énergie de l'orbite, c'est à dire le terme spectral de Ritz qui lui est associé. Les fréquences des différentes transitions s'obtiendront comme des différences de ces termes spectraux.

Le gradient de l'action par rapport à l'extrémité de la trajectoire étant l'impulsion, cette quantité s'écrit évidemment:

$$S = \int \mathbf{p} \cdot d\mathbf{r} , \quad (1.2)$$

\mathbf{r} étant la position de l'électron sur son orbite elliptique. L'action dépend donc de l'énergie de l'orbite, fixée par le demi grand axe de l'ellipse, mais aussi de l'excentricité de celle-ci. Pour fixer les paramètres de l'orbite, il faut imposer une condition supplémentaire. Bohr considère donc seulement des orbites circulaires. Une telle limitation, très arbitraire, n'est justifiée que par son succès. Sur une telle orbite les modules de la vitesse \mathbf{v} et de \mathbf{r} sont constants et on a évidemment:

$$S = 2\pi mrv = 2\pi L = nh \quad (1.3)$$

où m est la masse de l'électron et L la norme du moment angulaire. Notons que nous devrions utiliser ici en toute rigueur la masse réduite de l'électron qui tient compte de l'effet d'entraînement du noyau. Cet effet et même sa variation d'un isotope de l'hydrogène à l'autre sont parfaitement mesurables. Pour simplifier, nous considérerons dans la suite la masse du proton comme infinie. Les constantes que nous définirons sont alors exprimées en fonction de la vraie masse de l'électron et il faudrait appliquer des facteurs correctifs aux formules pour tenir compte de l'effet d'entraînement.

²Nous ne reproduisons pas ici les arguments originels de Bohr, un peu moins accessibles.

La condition de quantification de Bohr s'écrit donc aussi:

$$L = n\hbar , \quad (1.4)$$

où $\hbar = h/2\pi$. C'est sous cette forme qu'elle est le plus souvent écrite.

En 1923, De Broglie introduit une onde associée à toute particule quantique dont la longueur d'onde est donnée par la célèbre relation:

$$\lambda = h/p . \quad (1.5)$$

On constatera sans peine que la relation de quantification de Bohr est équivalente à postuler que le périmètre de l'orbite est égal à un nombre entier de longueurs d'onde de de Broglie. Cette condition "d'onde stationnaire" est très suggestive mais ne pourra être employée de façon quantitative avant le développement de l'équation de Schrödinger en 1926.

En notant e le module de la charge de l'électron et en écrivant l'équilibre de l'électron sur sa trajectoire circulaire, on montre immédiatement que

$$v = \sqrt{\frac{e^2}{4\pi\epsilon_0 r m}} . \quad (1.6)$$

En reportant cette expression dans la condition de quantification, on trouve le rayon de l'orbite:

$$r = a_0 n^2 , \quad (1.7)$$

où a_0 , rayon de Bohr, est défini par:

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} . \quad (1.8)$$

Numériquement, le rayon de Bohr, qui est le rayon de l'état fondamental de l'hydrogène, vaut 0.053 nm.

Il est physiquement intéressant de comparer le rayon de Bohr à une longueur caractéristique formée avec les paramètres de l'électron et la constante de Planck. Il s'agit de la longueur d'onde de Compton de l'électron:

$$\lambda_c = \frac{h}{mc} \quad (1.9)$$

(on se référera au chapitre sur la relativité restreinte pour une description détaillée de l'effet Compton, collision d'un photon énergétique et d'un électron). On peut écrire:

$$a_0 = \frac{\lambda_c}{2\pi\alpha} , \quad (1.10)$$

où:

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \quad (1.11)$$

(c est la vitesse de la lumière dans le vide). Cette constante, sans dimensions, numériquement égale à $1/137$, joue un rôle essentiel dans le modèle de Bohr et au delà dans toute l'électrodynamique quantique. Si, pour des raisons purement historiques, elle est appelée "constante de structure fine", elle mesure en fait la "force" de l'interaction électromagnétique. C'est en effet la seule constante sans dimension formée avec les paramètres de l'électromagnétisme (charge de l'électron et vitesse de la lumière) et la constante de Planck. Pratiquement tous les résultats de l'électrodynamique quantique peuvent se mettre sous la forme d'une fonction simple de cette constante ou d'un développement en ses puissances. Le modèle de Bohr ne fera pas exception à la règle.

Il est maintenant trivial de calculer l'énergie de l'électron sur son orbite et donc le terme spectral de Ritz. On a

$$E = -\frac{R}{n^2} \quad (1.12)$$

(cette énergie, correspondant à un état lié, est évidemment négative). On retrouve bien les termes spectraux en $1/n^2$ de la formule de Balmer. La constante R , qui n'est autre que la constante de Rydberg, peut s'écrire:

$$R = \frac{e^2}{8\pi\epsilon_0 a_0} . \quad (1.13)$$

On vérifiera sans peine qu'elle s'écrit aussi, en termes de la constante de structure fine et de l'énergie relativiste de masse de l'électron, mc^2 , comme:

$$R = mc^2 \frac{\alpha^2}{2} . \quad (1.14)$$

On vérifiera sans peine que la valeur numérique est de 13.6 eV.

Nous avons ici calculé les fréquences de transition comme si la masse du proton était infinie. Pour une masse finie, on trouve encore bien sûr une loi en $1/n^2$, avec une constante de Rydberg légèrement modifiée $R_M = R/(1 + m/M)$ où M est la masse de noyau (1836 m pour l'hydrogène). Les fréquences des raies émises s'interprètent alors simplement. L'atome peut effectuer une transition entre deux niveaux quantiques en émettant ou en absorbant un photon ayant une énergie égale à la différence des énergies du niveau initial et du niveau final. La fréquence de la transition entre les niveaux n et m est donc de la forme $(R/h)(1/n^2 - 1/m^2)$, coïncidant avec la formule de Balmer et les extensions par Rydberg.

Notons enfin qu'on peut, à partir de ces différentes expressions, réécrire la vitesse de l'électron sous la forme:

$$v = c \frac{\alpha}{n} . \quad (1.15)$$

1.3 Au delà du modèle de Bohr

Le modèle de Bohr explique donc parfaitement les fréquences des raies de l'hydrogène en dépit du caractère un peu artificiel des hypothèses de départ (orbite circulaire et condition de quantification ad hoc). Il est cependant insuffisant pour expliquer les structures fines observées très tôt dans le spectre. Un pas important est franchi indépendamment par Sommerfeld et Wilson en 1915. Ils considèrent un mouvement elliptique plus général et imposent des conditions de quantification à tous les couples de variables conjuguées ayant une influence sur la dynamique. En termes modernes, ils quantifient le mouvement radial et la norme du moment angulaire. La formule obtenue est en parfait accord avec les structures fines mesurées (et avec la théorie quantique moderne au même ordre d'approximation).

Bien sûr, Bohr, Sommerfeld et bien d'autres cherchent ensuite à adapter cette "première théorie des quanta" à des systèmes atomiques plus complexes. Ils s'intéressent à l'hélium, le plus simple des systèmes complexes avec ses deux électrons (le spectre expérimental est alors bien connu). Et c'est là où le bât blesse! Le système à trois corps ne peut en effet être traité explicitement en mécanique classique. Les règles de quantification utilisées pour l'hydrogène ne sont pas directement applicables. Commence alors un superbe travail de mécanique céleste visant à mettre la version classique du problème sous une forme propre à la quantification. En dépit d'efforts énormes, tirant parti des techniques les plus sophistiquées de la mécanique analytique, toutes ces tentatives échoueront et il ne sera pas possible de donner une interprétation convaincante du spectre de l'hélium. Une grave crise de la mécanique quantique s'ensuivit. Elle ne sera réglée qu'en 1925-1926 par l'invention simultanée de la mécanique des matrices par Heisenberg et de la mécanique ondulatoire par Schrödinger (on reconnaîtra très rapidement l'équivalence des deux approches, en dépit d'un débat plutôt vif initialement entre Heisenberg et Schrödinger).

Nous ne rappellerons pas ici les développements suivants qui ont permis, en particulier avec Dirac, qui introduit la relativité dans le problème, de donner une théorie complète et rigoureuse de l'atome d'hydrogène et, au prix de techniques de calculs complexes, des éléments plus lourds. Notons toutefois

que l'atome d'hydrogène a continué longtemps à constituer une pierre de touche de la mécanique quantique. La mesure et l'interprétation du déplacement du niveau 2S par rapport au $2P_{1/2}$, nul dans le modèle relativiste de Dirac, le fameux "Lamb shift", a joué un rôle essentiel dans le développement de la théorie quantique des champs et des techniques de traitement des infinis.

Pour mémoire, nous rappellerons brièvement ici les résultats de la mécanique quantique standard non relativiste. On en trouvera une dérivation détaillée dans le Cohen. Les fonctions d'onde propres du Hamiltonien quantique (constitué du seul potentiel Coulombien en $1/r$) sont repérées par trois "nombres quantiques" n, ℓ et m . Elles se mettent, en coordonnées sphériques, sous la forme du produit d'une fonction de r par une fonction de la direction angulaire, une "harmonique sphérique": $\Psi_{n\ell m} = R_{n\ell}(r)Y_{\ell}^m(\theta, \phi)$. ℓ décrit la longueur du moment cinétique: le carré du moment cinétique vaut $\ell(\ell+1)\hbar^2$. ℓ est un nombre entier positif ou nul. m décrit la projection du moment cinétique sur un axe "de quantification", Oz en l'occurrence. Cette projection vaut simplement $m\hbar$ et est comprise entre $-\ell$ et ℓ . Enfin, n , nombre quantique principal, décrit le nombre des "extrema" (nombre de zéros plus un) de la fonction d'onde radiale $R_{n\ell}$. On montre que ℓ est au plus égal à $n-1$. L'énergie du niveau n, ℓ, m est, à ce degré d'approximation, R/n^2 , la valeur prédite par le modèle de Bohr.

On peut donc classer les niveaux par valeurs croissantes du nombre quantique principal. Le niveau $n=1, \ell=m=0$, fondamental, est appelé 1S. Généralement, les niveaux de $\ell=0$ sont notés S, $\ell=1$ P, $\ell=2$ D et $\ell=3$ F. Ces notations, anciennes mais universelles, sont descriptives de l'aspect des raies spectrales: S pour "sharp", P pour "principal", D pour "diffuse", F pour "fundamental". Les raies connectant un niveau S au fondamental, S lui aussi, sont interdites par transition dipolaire électrique. Elles ont donc une durée de vie importante et une très faible largeur spectrale. Elles sont aussi relativement insensibles aux champs électriques ou magnétiques parasites. Les raies d'un niveau P vers le fondamental S sont très autorisées et donc très intenses. Les raies des niveaux D sont très sensibles aux champs électriques parasites dans les décharges et apparaissent larges, diffuses, dans les spectres.

Avec ces notations, les premiers niveaux excités sont 2S et 2P ($m=0, \pm 1$). Les suivants 3S, 3P, 3D... Les fonctions d'onde correspondantes, pour mémoire, sont

$$\Psi_{1S} = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} \quad (1.16)$$

$$\Psi_{2S} = \frac{1}{\sqrt{8\pi a_0^3}} \left(1 - \frac{r}{2a_0}\right) e^{-r/2a_0} \quad (1.17)$$

$$\Psi_{2P, m=\pm 1} = \mp \frac{1}{8\sqrt{\pi a_0^3}} \left(\frac{r}{a_0}\right) e^{-r/2a_0} \sin \theta e^{\pm im\phi} \quad (1.18)$$

$$\Psi_{2P, m=0} = \frac{1}{4\sqrt{2\pi a_0^3}} \left(\frac{r}{a_0}\right) e^{-r/2a_0} \cos \theta . \quad (1.19)$$

Ce modèle n'est qu'approché. Il faut lui ajouter les effets relativistes (qui impliquent en particulier l'existence du spin de l'électron) et les effets d'entraînement du noyau décrivant, ensemble, la structure fine. Il faut aussi tenir compte de l'interaction entre l'électron et le spin nucléaire, des effets de volume du noyau (un proton n'est pas ponctuel) et enfin des corrections radiatives qui décrivent le Lamb shift. Tout cela peut être calculé avec une précision remarquable, puisque l'accord théorie/expérience sur le spectre de l'hydrogène atteint maintenant presque les 10^{-12} en valeur relative, précision limitée seulement par la connaissance de la structure du proton, qui joue un rôle essentiel à ce degré d'exactitude, et par la stabilité des horloges étalon utilisées pour la détermination des fréquences. Avec de tels accords, la constante de Rydberg est sans aucun doute la constante la mieux connue de la physique fondamentale.

Ces fonctions d'onde sont bien éloignées du modèle d'orbite circulaire de Bohr. En fait, les niveaux S qui correspondent à un moment cinétique orbital nul seraient représentés en mécanique classique

par une orbite complètement dégénérée en une droite où l'électron rebondit à chaque période sur le noyau. L'incertitude quantique rend, qualitativement, "incertaine" la direction de cette orbite et les fonctions d'onde quantiques des états S sont à symétrie sphérique. Si on peut retrouver des orbites de Bohr, ou quelque chose qui y ressemble un peu, c'est en allant vers les grands nombres quantiques. Le principe de correspondance nous enseigne en effet que les résultats de la mécanique quantique doivent rejoindre ceux de la mécanique classique quand *tous* les nombres quantiques sont grands. Un grand nombre quantique principal (n de l'ordre de quelques dizaines) correspond à ce qu'on appelle maintenant un état de Rydberg. En effet, pour ces niveaux très excités, l'électron orbite très loin du noyau et des autres électrons et tout se passe comme si l'atome était hydrogénoïde, avec des niveaux en $1/n^2$. Les autres nombres quantiques doivent être grands eux aussi: $\ell = |m| = n - 1$. L'orbitale de ces niveaux, dits "circulaires" dans la littérature moderne, est un tore très mince centré sur un cercle de rayon $a_0 n^2$. C'est évidemment l'orbitale la plus proche du modèle de Bohr. Bien sûr, l'électron ne peut être localisé précisément sur cette orbite circulaire et la densité de présence est uniforme le long du périmètre. Cependant, ces niveaux ressemblent beaucoup au modèle de Bohr et de nombreuses prédictions classiques de ce modèle donnent des ordres de grandeur corrects pour les atomes circulaires.

