

L'ELABORATION DU CALCUL DES VARIATIONS ET SES APPLICATIONS A LA DYNAMIQUE

Florence Viot

1. Introduction

Le calcul des variations a pour principal but de résoudre le problème suivant : déterminer la courbe (ou la surface) qui rend minimale une intégrale donnée. Nous étudierons ici plus particulièrement le cas simple de la détermination de l'arc K défini par l'équation $y = \varphi(x)$, tel que $I = \int_K F(x, y, y')dx$ soit minimale. Se rattache également à cette question l'important problème dit des "isopérimètres", pour lequel on recherche l'arc K qui minimise l'intégrale précédente et qui vérifie de plus une contrainte du type $\int_K G(x, y, y')dx = \text{constante}$.

Dans cet exposé, je me propose de mettre en évidence les étapes essentielles de la genèse du concept de variation, jusqu'à l'élaboration formelle de la méthode analytique que Euler baptisa lui-même "calcul des variations" et à l'établissement de l'une des conditions nécessaires de ce problème, donnée par l'équation d'Euler-Lagrange :

$$\frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) - \frac{\partial F}{\partial y} = 0$$

Nous montrerons enfin combien ce problème procède pour une large part de questions d'origine physique (optique et mécanique), et de quelle façon le calcul des variations a permis la mise en oeuvre d'une nouvelle formulation des principes fondamentaux de la mécanique, ouvrant ainsi le champ à ce qu'il est coutume de nommer maintenant les principes variationnels de la physique.

Notons enfin que le développement mathématique du calcul des variations n'est évidemment pas achevé avec l'établissement de la condition d'Euler-Lagrange, malgré le caractère fondateur de cette équation pour les travaux ultérieurs réalisés en physique. On se souviendra à cet effet des travaux de Legendre (1786) établissant un critère de discrimination entre maxima et minima à l'aide de l'étude de la seconde variation de l'intégrale, critère complété par Jacobi en 1837. On mentionnera également les travaux de Weierstrass (1815-1897) qui reformula l'ensemble des conditions nécessaires et établit pour la première fois une condition suffisante, travaux développés par la suite par Bolza et Hilbert.

2. Les premières mentions de concepts variationnels

Les premières questions de type variationnel étudiées sont sans conteste les problèmes d'isopérimètres. Chez les Grecs déjà, on trouve mention de la recherche de la ligne fermée la

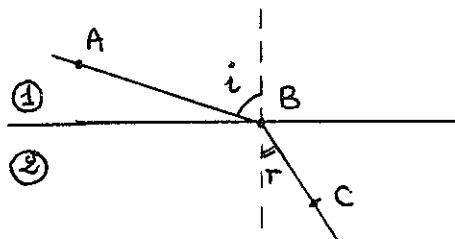
plus courte limitant une surface donnée et de celle des lignes de plus court chemin tracées sur des surfaces données (sphère, cylindre). Ces études perdurent au long des siècles, et sont essentiellement l'occasion de traitements géométriques. Cependant, les problèmes variationnels vont connaître un regain d'intérêt à partir du XVII^e S., lors du développement du calcul infinitésimal.

Citons par exemple, sans les développer, les travaux de Newton (1686) sur la détermination de la forme des surfaces de révolution des solides offrant une moindre résistance à l'avancement dans un fluide : problème dont Newton donnera la solution sans indiquer clairement la méthode utilisée.

a. Le principe de Fermat

Il est frappant de constater que le premier problème variationnel "moderne" posé l'a été en physique, de plus en Optique. Nous montrerons que, loin de constituer un avatar sans lendemain, cette question sera la pierre d'achoppement du développement des principes variationnels de la physique et que des mathématiciens comme Jean Bernoulli feront explicitement usage de sa résolution pour construire leur propre démarche.

Pierre de Fermat (1601-1665) se pose le problème suivant : démontrer à l'aide d'un principe physique clairement énoncé les lois de la réfraction d'un rayon lumineux (à la traversée d'une surface de discontinuité de l'indice). Descartes avait publié ces lois en 1637 dans sa *Dioptrique*, en les justifiant essentiellement à partir de l'analogie entre le trajet suivi par la lumière et la trajectoire d'une balle de jeu de paume et en postulant de plus que la vitesse de la lumière est d'autant plus grande que le milieu est plus dense. Ainsi en déduisait-il la relation $n_1 \sin i = n_2 \sin r$, où le rapport n_1/n_2 apparaît comme étant égal au rapport des vitesses dans les milieux considérés v_1/v_2 .



Fermat, dans une lettre datée du 1^{er} Janvier 1662, propose une autre résolution qui se dégage tant de la méthode des analogies que de l'hypothèse sur les vitesses à laquelle il est farouchement opposé. Il pose alors comme principe que la lumière suit toujours le chemin qui minimise le temps mis pour aller d'un point à un autre, se référant en cela à une grande tradition philosophique des principes dit d'économie qu'il énonce explicitement : "*je reconnais...la vérité de ce principe, que la nature agit toujours suivant les voies les plus courtes*" (1657). Fermat est dès lors ramené à un simple problème de géométrie : étant donné deux points A, C et les vitesses dans les milieux considérés, trouver le point B tel que le trajet ABC soit minimal ; problème qu'il résoud facilement à l'aide de la méthode de maximis et minimis qu'il a développée quelques années auparavant. Il

en déduit, à son grand étonnement, une loi similaire à celle de Descartes sans se rendre clairement compte que dans sa démarche le rapport que j'ai noté (adoptant les conventions modernes) n_1/n_2 doit cette fois-ci être interprété comme identique au rapport v_2/v_1 .

Ainsi apparaît pour la première fois un principe physique spécifiquement variationnel, qui n'est pas d'ailleurs sans entraîner des difficultés philosophiques et épistémologiques certaines. Les cartésiens seront les premiers à dénoncer les aspects téléologiques de ce principe ; de plus le concept de variation en physique est assez délicat, puisqu'il suppose de comparer entre elles différentes trajectoires a priori envisageables (qualifiées généralement de virtuelles) alors qu'il n'existe qu'une seule trajectoire effectivement réalisable. Malgré toutes ces réticences, le principe de Fermat restera cependant un élément moteur dans le développement tant mathématique que physique du calcul des variations.

b. La brachystochrone

Le problème de Jean Bernoulli

En Juin 1696 paraît dans les *Acta eruditorum* un défi lancé par le mathématicien J. Bernoulli (essentiellement à l'adresse de son frère Jacques) et énoncé de la façon suivante :

"Etant donnés deux points A et B dans un plan vertical, déterminer la courbe AMB le long de laquelle un mobile M abandonné en A descend sous l'action de sa propre pesanteur et parvient à l'autre point B dans le moins de temps possible."

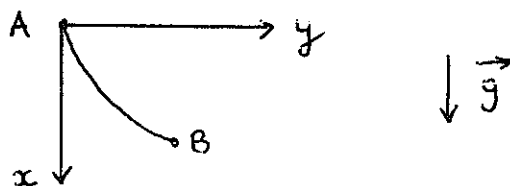
Plus loin, Bernoulli ajoute :

"bien que la droite AB soit le plus court chemin entre A et B, elle n'est cependant pas parcourue dans le temps le plus court ; et pourtant, la courbe cherchée est bien connue des géomètres ; je la nommerai si, avant la fin de l'année, personne d'autre ne l'a nommée."

Ce problème avait été énoncé pour la première fois par Galilée en 1638 dans les *Discorsi* où il affirmait de façon inexacte que la solution est un arc de cercle joignant les deux points considérés. En fait, au moment où paraît ce problème, Jean Bernoulli l'a déjà proposé à Leibniz par lettre séparée datée du 9 Juin 1696, lequel lui retourne aussitôt sa solution le 16 Juin sans qu'il soit d'ailleurs possible de savoir si Leibniz possédait ou non déjà la solution de cette question dans ses propres papiers avant de recevoir la missive de Bernoulli. Leibniz propose également de repousser la date d'expiration de ce défi en raison des délais de livraison des *Acta eruditorum* hors d'Allemagne. En mai 1697, sont finalement publiés dans cette même revue les solutions de Jean Bernoulli et celle de son frère Jacques, accompagnée d'une note de Leibniz indiquant que sa propre solution ne diffère pas notablement de celles-ci. Figurent également une solution anonyme de Newton, déjà publiée en Janvier 1697 dans les *Philosophical Transactions* ainsi qu'un mémoire du marquis de L'Hôpital basé sur une analogie avec le problème de la courbe funiculaire et un autre de Tschirnhaus.

Analyse

Décrivons rapidement, avec les notations actuelles du calcul des variations, la solution de ce problème.



La vitesse du point matériel, abandonné en A sans vitesse initiale, est $v = ds/dt = \sqrt{2gx}$, où s est l'abscisse curviligne, x l'ordonnée du point et g l'intensité de la pesanteur. On peut donc écrire :

$$dt = \frac{ds}{\sqrt{2gx}} = \frac{\sqrt{1+y'^2}dx}{\sqrt{2gx}}$$

où $y' = dy/dx$. Dès lors, la durée de trajet T de A à B s'écrit

$$T = \int_0^{x_B} \frac{\sqrt{1+y'^2}}{\sqrt{2gx}} dx = \int_0^{x_B} F(x, y') dx$$

La courbe $y = f(x)$ telle que T soit minimale obéit à l'équation d'Euler-Lagrange. On en déduit donc

$$\frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = \frac{\partial F}{\partial y} = 0$$

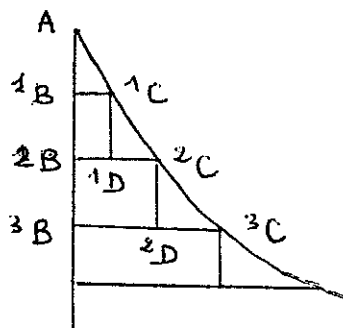
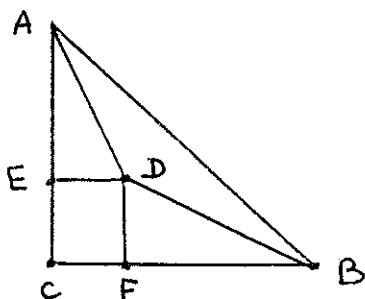
d'où

$$\frac{\partial F}{\partial y'} = \sqrt{\frac{y'^2}{2gx(1+y'^2)}} = \text{cste}$$

relation que l'on peut encore écrire $dy = kds\sqrt{x}$, où k est une constante quelconque. La solution de cette équation est un arc de cycloïde, courbe déjà bien connue depuis les travaux d'Huygens en 1659 sur l'isochronisme des oscillations des pendules.

La solution de Leibniz

Leibniz procède en deux temps : il considère d'abord un segment de droite AB , et se donne un point fixe E situé sur le côté vertical AC .

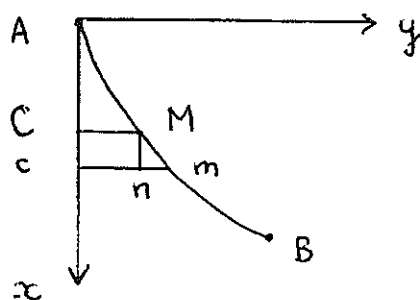


Il recherche alors la position du point D situé sur la parallèle à CB passant par E qui rend minimale le temps de chute $t_{AB} = t_{AD} + t_{DB}$, et montre qu'elle satisfait à une relation géométrique précise. Dès lors il peut s'intéresser au problème que Bernoulli nomme "brachystochrone" (signifiant "temps le plus court") et que lui nomme encore à cette époque "tachystoptote" (signifiant "descente la plus rapide"). Soit AB la courbe cherchée, et $^1B, ^2B, ^3B...$ des points régulièrement espacés sur l'axe vertical. Les points $^1C, ^2C, ^3C...$ de la courbe AB qui s'en déduisent sont suffisamment proches pour que l'on puisse confondre les arcs correspondants avec des portions rectilignes, équivalents aux segments AD, DB utilisés auparavant. L'étude précédente permet alors à Leibniz de montrer que la courbe est solution d'une équation différentielle du type $dy = k\sqrt{x}ds$ où s est l'abscisse curviligne, équation dont nous avons rappelé qu'elle est caractéristique d'une cycloïde.

La solution "indirecte" de Jean Bernoulli

L'originalité de la solution de Jean B. tient à sa volonté de rattacher le problème de la brachystochrone au premier grand principe variationnel énoncé, c'est-à-dire celui de Fermat dont nous avons parlé plus haut. Elle fut de plus riche d'enseignement pour les physiciens mathématiciens qui le suivirent, et constitue sans aucun doute une étape importante dans l'élaboration des principes variationnels de la mécanique, comme nous le verrons. Jean B. lui-même est conscient des implications étonnantes de sa découverte. Ainsi écrit-il : *"j'ai résolu en une seule fois deux problèmes importants, l'un optique, l'autre mécanique, et ai réalisé plus que ce que je demandais aux autres : j'ai montré que les deux problèmes, pris dans deux parties entièrement séparées des Mathématiques, sont de même nature"*.

Dans la leçon 46 des *Leçons de Calcul intégral* (1691-92), Bernoulli a déjà publié ses travaux concernant la courbure des rayons lumineux dans un milieu d'indice continûment variable, étudiée à partir de la loi de Descartes. Il procède dès lors par analogie, en faisant remarquer que la brachystochrone n'est autre que la courbe que suivrait un rayon lumineux dans un milieu de densité (donc d'indice) inversement proportionnel à la vitesse de chute libre d'une particule.

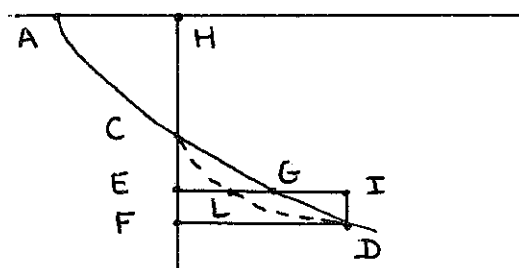


Soit M la position du point à un instant donné, repérée par son ordonnée $x = AC$, et soit c une position infiniment voisine. Bernoulli pose $Cc = dx = Mn$, et note dz l'arc Mm . La loi de Descartes lui permet d'écrire que le sinus de l'angle d'incidence, soit $mn/Mm = dy/dz$, est

proportionnel à la vitesse en M (notée t) ; d'où la relation $dy/dz = t/a$, où a est une constante quelconque. D'après l'hypothèse de Galilée sur la chute des corps $t = \sqrt{ax}$; Bernoulli obtient donc l'équation $dy = \sqrt{x/adz}$, caractéristique d'un arc de cycloïde. (voir le texte présenté en annexe)

La solution de Jacques Bernoulli

Assez proche conceptuellement de celle de Leibniz, cette présentation a l'avantage de mettre clairement en évidence quelques points fondamentaux de la méthode des variations, points que l'on retrouvera très nettement dans la technique élaborée par Euler.



Jacques montre d'abord que si AB est la courbe minimale recherchée, alors tout sous-arc CD de AB minimise également le temps de parcours correspondant t_{CD} , ce qui lui permet de ne s'occuper que de variations locales. Considérant désormais l'arc CD où " C est proche de D ", il introduit un autre arc CLD infiniment proche du premier qu'il caractérise par la position du point L situé sur l'horizontale passant par le point caractéristique G (qui apparaît comme le point de cote verticale médiane entre celles de C et D). Dès lors, le temps de trajet t_{CD} sur l'arc CGD sera minimal si $t_{CG} + t_{GD} = t_{CL} + t_{LD}$ (condition de stationnarité de la quantité étudiée). De là, après des manipulations géométriques que nous ne détaillerons pas ici, Jacques Bernoulli aboutit à la relation : $\frac{EG/\sqrt{HC}}{GI/\sqrt{HE}} = \frac{CG}{GD}$; équation qui en notation différentielle peut s'écrire (en remarquant que $CG \sim ds, HC \sim x, EG \sim dy$) sous la forme désormais classique $dy = kds\sqrt{x}$.

Nous venons donc ici de décrire succinctement les principales résolutions du problème de la brachystochrone ; elles nous ont permis de voir comment certains concepts fondamentaux du calcul des variations se mettent en place : la notion de courbe variée, l'étude locale des problèmes de minimisation et l'utilisation de l'outil différentiel. A travers l'exemple de Jean Bernoulli, nous pouvons de plus pressentir les liens étroits qui vont s'établir, au sein du calcul des variations, entre les concepts mathématiques et les principes fondamentaux de la physique. Notons que le problème spécifique de la brachystochrone va par la suite connaître de nombreuses extensions fécondes : ainsi les Bernoulli étudieront-ils la courbe de descente de durée minimale d'un point fixe à une droite verticale donnée, puis à une droite quelconque ; problème qui débouche sur l'importante

condition de transversalité qui sera formalisée par la suite par Lagrange.

c. Les isopérimètres

Nous ne ferons ici que mentionner l'existence de ce problème et des solutions qui lui furent apportées à cette époque essentiellement par les Bernoulli, à travers par exemple la recherche de la courbe de centre de gravité le plus bas.

L'une des avancées essentielles que leurs travaux permirent de réaliser est, sans conteste, le fait d'avoir compris la nécessité de disposer dans ce type de problème d'un degré de liberté supplémentaire. Ceci se traduit par l'obligation de faire varier la position de (au moins) deux points de la courbe. Cette constatation, que Jean Bernoulli n'a pas perçu dans un premier temps, est clairement utilisée par Jacques puis adoptée par son frère ; elle les conduira à la résolution de nombreux problèmes intéressants d'isopérimètres.

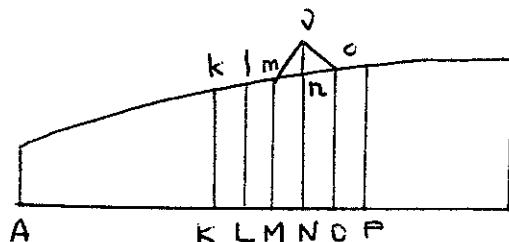
3. L'élaboration du calcul des variations

a. Les travaux d'Euler

L. Euler (1707-1783) est un élève de Jean Bernoulli, et ses premiers travaux sont écrits dans la lignée des problèmes (en particulier d'isopérimètres) posés par les frères Bernoulli. En 1744, il publie un ouvrage très important intitulé "Méthode pour trouver les lignes courbes qui possèdent une propriété de maximum ou de minimum", le *Methodus inveniendi*... Dans ce traité, Euler résout un très grand nombre de problèmes représentant un vaste éventail d'exemples de ce qui va devenir, selon les termes de Lagrange repris plus tard par Euler lui-même, le calcul des variations. Nous allons décrire sa méthode de résolution à travers un exemple simple.

La technique variationnelle d'Euler

Euler cherche à déterminer les propriétés de la courbe anz sur laquelle l'intégrale définie $\int Z dx$, prise entre les abscisses A et Z des points limites supposés fixes, est extrémale. Z est ici fonction de l'abscisse x , de l'ordonnée y du point considéré, et des dérivées successives $p = \frac{dy}{dx}$, $q = \frac{dp}{dx}$... L'intervalle AZ est ensuite divisé en sous-intervalles infiniment petits, repérés par les points équidistants $K, L, M, N, O, \dots$ d'ordonnées $Kk = y_{\dots}$, $Ll = y_{\dots}$, $Mm = y$, $Nn = y'$, $Oo = y''$,... et d'abscisses correspondantes $x_{\dots}$, $x_{\dots}$, $x = AM$, x' , x'' ...



Dès lors, Euler exprime les valeurs des dérivées successives aux points considérés par des relations du type :

$$p = \frac{y' - y}{dx} \quad p' = \frac{y'' - y'}{dx}$$

Par décomposition, Euler écrit l'intégrale étudiée, prise sur l'intervalle AZ , sous la forme

$$\int Z dx = \int_0^{x'} Z dx + Z dx + Z' dx + Z'' dx + \dots$$

expression dans laquelle j'ai rétabli les bornes d'intégration par souci de clarté, et où Z représente la valeur de la fonction au point $(x, y, p...)$, Z' sa valeur au point $(x', y', p'...)$

Afin d'examiner le comportement de l'intégrale, Euler fait ensuite varier l'ordonnée y' de N d'une quantité très petite $n\nu$, variation dont il calcule les conséquences sur l'ensemble des grandeurs $(y, p, q...)$ Ainsi, y' est-il modifié en $y' + n\nu$; p en $p + \frac{n\nu}{dx}$; p' en $-\frac{n\nu}{dx}...$ L'intégrale étudiée, extrémale sur l'arc anz original, n'est quant à elle pas modifiée par cette opération. Or, la seule partie de cette intégrale susceptible de variation est l'expression $Z dx + Z' dx$. Dans le cas particulier où dZ s'exprime sous la forme $dZ = M dx + N dy + P dp$, Euler écrit alors que la variation de $Z dx$ est $P dp dx$ soit $P n\nu$, et que celle de $Z' dx$ est $(N' dy' + P' dp') dx$ soit $N' n\nu dx - P' n\nu$.

La variation totale de l'intégrale a donc pour expression $(P + N' dx - P') n\nu$. Remplaçant alors N' par N , et $P' - P$ par dP , Euler obtient finalement l'équation

$$N - \frac{dP}{dx} = 0$$

qui n'est autre que l'équation connue maintenant sous le nom d'Euler-Lagrange.

Les apports de la méthode d'Euler

Pour la première fois, on dispose d'une méthode générale géométrico-analytique pour appréhender les problèmes variationnels. Malgré le caractère souvent complexe des calculs développés, Euler est en mesure d'appliquer sa méthode à un ensemble très varié de cas, démontrant ainsi son efficacité. Ainsi s'intéresse-t-il au cas d'intégrales définies $\int Z dx$, où Z dépend de $(x, y, p...)$ et d'une quantité elle-même définie de façon intégrale $\Pi = \int_0^x [Z] dx$; problème par exemple lié aux intégrales de fonction dépendant de l'abscisse curviligne : $\Pi = s = \int_0^x \sqrt{1 + p^2} dx$. Cette catégorie de problème constitue une étape importante vers le célèbre problème de Lagrange de minimisation avec contrainte : il s'agit en effet ici de minimiser $\int Z(x, y, y', \Pi) dx$ sous la condition $\Pi' - [Z] = 0$.

Ceci conduira Euler à l'utilisation de fait d'une méthode similaire à la règle des multiplicateurs de Lagrange, sans qu'elle soit pour autant formellement énoncée comme telle.

L'ambiguïté du symbole différentiel chez Euler

Dans un premier temps, Euler utilise la notation différentielle " d " suivant l'usage en vigueur consacré par Leibniz. Ainsi, la différentielle dx est maintenue constante, et la différentielle de

toute grandeur dépendante de x (telle que $y, p, \dots$) est définie comme la différence entre sa valeur au point x et sa valeur au point suivant, distant de dx du précédent ; Euler écrit par exemple $dy = p dx = y' - y$.

Cependant, Euler est amené à donner un autre sens au symbole de différentiation "d". Ainsi, lorsqu'il calcule le changement apporté à l'expression $Z dx + Z' dx$ du fait de la modification de l'ordonnée $y' = N n$ devenue $N \nu$, la grandeur $dZ = P n \nu$ représente cette fois la variation de Z au sens lagrangien du terme. Notons encore qu'Euler écrit en substance dans ce calcul que la modification des diverses variables est donnée par : $dx = 0, dy = 0, dp = \frac{n\nu}{dx} \dots$

La notation "d" est alors clairement utilisée ici dans deux contextes totalement différents, qui se côtoient et se confondent même chez Euler : ainsi passe-t-il insensiblement de la notion leibnizienne de la différentielle à celle de variation, pour revenir ensuite à la notion première dans son équation finale $N - \frac{dP}{dx} = 0$; dP représentant ici explicitement la différentielle usuelle $P' - P$.

b. Les travaux de Lagrange

C'est à Lagrange que revient le mérite d'avoir clairement mis en évidence cette difficulté et de l'avoir résolu élégamment, en introduisant le concept général de variation dont la traduction la plus manifeste est la naissance d'un nouveau symbole : " δ ".

Les premiers pas

C'est en 1754 que le tout jeune Lagrange, alors âgé de dix-huit ans, entreprend de faire part de ses recherches personnelles au savant universellement reconnu qu'est Euler à cette époque, et mentionne en particulier l'existence de ses propres travaux concernant les problèmes de maxima et minima. Mais la première indication claire de ses réflexions se trouve dans la deuxième lettre qu'il écrit à Euler le 12 Août 1755.

S'appuyant sur les considérations développées par celui-ci, Lagrange dit avoir recherché une démonstration plus rapide des résultats fondamentaux mis en évidence par Euler, "sans aucune construction géométrique". A cette occasion, il est intéressant de noter qu'Euler avait mis en évidence une règle mnémotechnique permettant de retrouver rapidement cette équation. Il suffisait, pour une fonction $Z(x, y, p)$ dont la différentielle s'écrit $dZ = M dx + N dy + P dp$, d'annuler le terme $M dx$ puis de remplacer le terme $P dp$ par $-pdP$, et finalement d'annuler l'expression obtenue ; ceci conduisait donc à l'équation $N dy - pdP = 0$, qui n'est autre que l'équation d'Euler (car $dy = p dx$).

Dans la deuxième lettre qu'il adresse à Euler, Lagrange justifie la nécessité d'élaborer une méthode entièrement nouvelle pour traiter ces problèmes par les propos qu'Euler lui-même tenait dans son mémoire original. Celui-ci y avait en effet noté combien il paraissait souhaitable de développer un autre type de méthode, libérée de toute considération géométrique et qui conduirait naturellement à la transformation du terme $P dp$ en $-pdP$ que nous avons évoqué plus haut. Sans

pour autant en déduire que la démarche de Lagrange a été entièrement induite par cette remarque, il est cependant certain qu'elle a dû frapper son esprit inventif, suffisamment du moins pour qu'il la considère a posteriori comme une étape importante de sa réflexion.

Néanmoins, l'élément fondamental de la nouvelle approche développée par Lagrange est sans conteste sa découverte des deux rôles conceptuellement très différents que Euler fait jouer, dans sa méthode, au symbole de différentiation classique d . Pour la première fois, Lagrange est donc amené à introduire un nouveau symbole δ qui permet de caractériser la variation d'une grandeur, sans passer obligatoirement par une représentation géométrique. Ainsi écrit-il :

"je dénoterai par δ la différentielle de y , dans la mesure où c'est y qui doit être différentié, x demeurant constant, pour obtenir la valeur maximum ou minimum de la formule donnée, cela afin de distinguer cette différentielle des autres différences de y qui entrent déjà dans cette formule."

Ce faisant, il développe assez clairement l'algorithme du calcul des variations, en prenant bien soin de distinguer dans l'esprit et dans la notation les deux différentiations qu'Euler confondait sous le symbole d . De plus, il imagine une variation arbitraire et simultanée des ordonnées de tous les points de la courbe, ce qui le libère des calculs lourds et fastidieux qu'Euler avait été obligé d'employer. Il lui suffit de poser de fait les quelques principes analytiques qui vont régir ce nouveau procédé de différentiation : règle de permutation des symboles d et δ , qui entraîne $\delta d^m F(y) = d^m \delta F(y)$ pour toute fonction de la variable y , ainsi que $\delta \int Z = \int \delta Z$ où Z est une fonction quelconque de $x, y, dx, dy, \dots$

Dès lors, Lagrange utilise les procédés usuels de dérivation des fonctions, tout en les appliquant au cas du symbole δ nouvellement introduit. Remarquons de plus que Lagrange impose ici à la variable indépendante x de rester constante dans le processus de variation. Ainsi donc, pour la fonction Z précédemment décrite telle que $dZ = Mdx + Ndy + Pd^2y + Qd^3y + \dots$, il écrit $\delta Z = N\delta y + P\delta dy + Q\delta d^2y + \dots$; ce qui, après intégration et commutation des différentielles, conduit à :

$$\delta \int Z = \int \delta Z = \int N\delta y + \int P\delta dy + \int Q\delta d^2y + \dots$$

Lagrange introduit alors une des étapes fondamentales du nouveau calcul qu'il est en train d'élaborer, en effectuant une intégration par parties qui le conduit à l'expression :

$$\delta \int Z = \int N\delta y + P\delta y - \int dP\delta y + Q\delta dy - dQ\delta y + \int d^2Q\delta y + \dots$$

Lagrange est désormais en mesure de résoudre le problème de la recherche des extrema d'une formule intégrale $\int Z$ calculée entre deux extrémités fixes, telles que $\delta y = 0, d\delta y = 0, \dots$. En effet, posant $\delta \int Z = 0$, il aboutit à

$$\int (N - dP + d^2Q - \dots)\delta y = 0$$

d'où il déduit, sans aucune esquisse de justification, l'équation

$$N - dP + d^2Q - \dots = 0$$

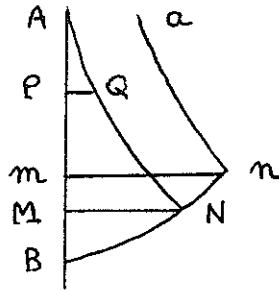
qui n'est autre que l'équation dite d'Euler. Afin de démontrer la grande fécondité de sa nouvelle méthode, Lagrange s'attache ensuite à retrouver la solution d'un problème posé par Euler dans son mémoire original, consistant à déterminer l'extremum de la formule intégrale $\int Z$, où Z dépend non seulement de $x, y, dy, d^2y, \dots$ mais aussi d'une quantité $\Pi = \int(Z)$ où (Z) est fonction de $y, dy, \dots$

La réponse d'Euler fut immédiate et enthousiaste, reconnaissant d'emblée l'immense portée des idées développées par Lagrange. Ainsi lui écrit-il :

"Vous me semblez avoir porté la théorie des maxima et des minima presque à son plus haut point de perfection, et je ne puis assez admirer l'extraordinaire sagacité de votre esprit. Car non seulement j'avais réclamé dans mon traité sur ce sujet une méthode purement analytique qui permette de déduire les règles qui y sont énoncées, mais j'avais encore consacré par la suite bien des efforts à la recherche d'une telle méthode."

L'exemple de la brachystochrone

Le problème étudié par Lagrange a pour but de déterminer la forme de la courbe qui minimise la durée de chute d'un point matériel soumis à la pesanteur (considérée comme uniforme) abandonné en un point origine fixe A , l'autre extrémité de la courbe étant supposée quelconque, mais située sur une ligne donnée BNn .



Il s'agit donc de minimiser pour cette famille d'arcs particulière l'intégrale $\int \frac{ds}{\sqrt{u}}$, où s représente l'abscisse curviligne sur la courbe brachistochrone, et u la "hauteur" du point (analogue à la grandeur $x = AP$ qui caractérise le point courant Q). L'application des règles du nouveau calcul δ conduit à écrire que

$$\delta \int \frac{ds}{\sqrt{u}} = - \int \frac{ds \delta u}{2u^{\frac{3}{2}}} + \int \frac{\delta ds}{\sqrt{u}}$$

Lagrange pose $\delta u = \nu \delta y$, et utilise de plus implicitement la relation $\delta ds ds = \delta dy dy$, issue de l'identité $ds^2 = dy^2 + dx^2$ (avec $\delta x = 0$). Il en déduit

$$\delta \int \frac{ds}{\sqrt{u}} = - \int \frac{ds \nu \delta y}{2u^{\frac{3}{2}}} + \int \frac{dy \delta dy}{ds \sqrt{u}}$$

Après permutation des différentiations d et δ , puis intégration par parties, il obtient

$$\delta \int \frac{ds}{\sqrt{u}} = - \int \left(\frac{ds \nu}{2u^{\frac{3}{2}}} + d \frac{dy}{ds \sqrt{u}} \right) \delta y + \frac{dy}{ds \sqrt{u}} \delta y$$

Cependant, fait remarquer Lagrange, ce calcul ne tient pas compte entièrement du fait que le point d'arrivée est lui aussi variable. La formule précédemment obtenue permet en effet de calculer la variation de temps de parcours lorsque l'on passe de la courbe AQN à toute autre courbe respectant la hauteur de chute AM . Or la courbe variée an correspond elle à une hauteur de chute Am ; il est donc nécessaire de retrancher à l'expression précédente le temps de chute correspondant à mM , soit $\frac{ds}{\sqrt{u}}$. D'où l'expression finale

$$\delta \int \frac{ds}{\sqrt{u}} = - \int \left(\frac{ds\nu}{2u^{\frac{3}{2}}} + d \frac{dy}{ds\sqrt{u}} \right) \delta y + \frac{dy}{ds\sqrt{u}} \delta y - \frac{ds}{\sqrt{u}}$$

La condition de temps minimal conduit donc aux deux expressions cherchées; la première, soit

$$\frac{ds\nu}{2u^{\frac{3}{2}}} + d \frac{dy}{ds\sqrt{u}} = 0$$

déterminant l'équation de la courbe (il s'agit, rappelons-le d'une cycloïde); la seconde, soit

$$\frac{dy}{ds\sqrt{u}} \delta y = \frac{ds}{\sqrt{u}}$$

écrite au point d'arrivée, déterminant les coordonnées de ce point. Lagrange montre d'ailleurs que cette équation traduit la très fondamentale relation d'orthogonalité, qui indique que la courbe solution coupe à angle droit la ligne sur laquelle est astreint à se déplacer le point extrémal.

Mais moins d'un an après cette première démonstration, le 2 septembre 1756 exactement, Lagrange entreprend de réécrire ce calcul en assignant cette fois une variation à la grandeur x . Les règles usuelles du calcul des variations entraînent alors (là encore, j'ai développé quelques étapes que Lagrange a omis de mentionner) :

$$\delta \int \frac{ds}{\sqrt{x}} = \int \frac{\delta ds}{\sqrt{x}} + \int ds \delta \frac{1}{\sqrt{x}}$$

soit, en utilisant cette fois $ds\delta ds = dy\delta dy + dx\delta x$

$$\delta \int \frac{ds}{\sqrt{x}} = \int \left[\frac{dy\delta dy}{ds\sqrt{x}} + \frac{dx\delta dx}{ds\sqrt{x}} - \frac{ds\delta x}{2x^{\frac{3}{2}}} \right]$$

Après permutation des symboles d et δ et intégration par parties, on obtient

$$\delta \int \frac{ds}{\sqrt{x}} = \frac{dy}{ds\sqrt{x}} \delta y + \frac{dx}{ds\sqrt{x}} \delta x - \int \left[\left(d \frac{dy}{ds\sqrt{x}} \right) \delta y + \left(d \frac{dx}{ds\sqrt{x}} + \frac{ds}{2x^{\frac{3}{2}}} \right) \delta x \right]$$

Ici, pour la première fois, Lagrange utilise explicitement l'argument de l'indépendance des variations introduites δx et δy . En conséquence, il annule toutes les expressions qui sont en facteur de ces termes dans l'expression intégrale, et en déduit

$$d \frac{dy}{ds\sqrt{x}} = 0; \quad d \frac{dx}{ds\sqrt{x}} + \frac{ds}{2x^{\frac{3}{2}}} = 0$$

équations qui conduisent toutes deux, comme il l'avait signalé dans une lettre précédente, à la même équation intégrale. De plus, Lagrange obtient l'équation relative au point extrémal, en annulant les termes de l'expression générale situés hors du signe somme, soit : $dy\delta y + dx\delta x = 0$; cette équation traduisant là aussi la condition d'orthogonalité.

Une autre présentation

En 1771, Euler entreprend de définir de façon différente la notion analytique de variation, à l'aide de l'introduction d'un paramètre variable, en l'écrivant en substance comme transition à l'intérieur d'une famille de courbes à un paramètre ; soit :

$$\delta y = \left(\frac{\partial y}{\partial t} \right)_{t=t_0} dt$$

Cette présentation est alors partiellement reprise dans l'ouvrage de Lagrange intitulé *Théorie des fonctions analytiques* (1797), mais c'est essentiellement dans la seconde édition (1806) du traité *Leçons sur le calcul des fonctions* qu'elle sera utilisée.

Lagrange y entreprend de construire une nouvelle présentation du mécanisme variationnel, basé cette fois non plus sur les infiniment petits, mais sur une véritable conception différentielle. Il considère pour cela le problème de la détermination des extrema de la formule intégrale U d'une fonction $y = \varphi(x)$. Cette détermination nécessite, de façon évidente, le calcul de la variation de U lors des petites variations correspondantes de la fonction y . Lagrange imagine alors une modification de la fonction φ en une fonction $\varphi(x, i)$, obéissant à la condition $\varphi(x, i = 0) = \varphi(x) = y$. i joue donc ici le rôle d'un paramètre générant à partir de la courbe originelle $\varphi(x)$ une famille de courbes variées $\varphi(x, i)$. Le paramètre i pouvant être choisi aussi petit que l'on veut, Lagrange écrit la nouvelle fonction $\varphi(x, i)$ sous la forme d'un développement en puissances croissantes :

$$\varphi(x, i) = y + i\dot{y} + \frac{i^2}{2}\ddot{y} + \dots$$

où $y = \varphi(x, 0)$, et $\dot{y}$ représente la dérivée de la fonction $\varphi(x, i)$ par rapport à i au point $i = 0$.

Lagrange assimile alors l'accroissement de la fonction y au terme d'ordre 1 de ce développement, soit $i\dot{y}$, terme totalement arbitraire puisque la dépendance de $\varphi(x, i)$ par rapport à i est quelconque. Cependant, Lagrange définit ici la variation de y non pas par ce terme d'accroissement, mais par le taux d'accroissement correspondant, c'est-à-dire la grandeur $\dot{y}$. Ainsi, la variation de $y = \varphi(x, 0)$ est définie par le rapport $\delta y = \frac{\varphi(x, i) - \varphi(x, 0)}{i}$, calculé en $i = 0$.

De même procède-t-il pour calculer la variation de l'intégrale U d'une fonction $V(x, y, y', \dots)$ donnée. La transformation de y en $\varphi(x, i)$ induit une nouvelle expression de V , donc une nouvelle expression de U , toutes deux développables suivant les puissances croissantes de i . Ainsi, V devient :

$$V + i\dot{V} + \frac{i^2}{2}\ddot{V} + \dots$$

et U s'écrit :

$$U + i\dot{U} + \frac{i^2}{2}\ddot{U} + \dots$$

La notation $\dot{f}$ indiquant de même la dérivée de la nouvelle fonction considérée par rapport à i , au point $i = 0$. $\dot{V}$, $\dot{U}$ représentent donc désormais les variations des fonctions V et U . De

là Lagrange dérive, par des moyens que nous ne développerons pas ici, les équations usuelles traduisant l'existence des extrema de la fonction $U = \int V dx$.

L'approche de Lagrange concernant la présentation du calcul des variations subit ici une modification tout à fait spectaculaire.

4. Le principe de moindre action

a. Le principe de Maupertuis

En 1744 (80 ans après les travaux de Fermat), Maupertuis énonce son principe de moindre action. Il est mû d'une part par la conviction que la nature obéit à un principe d'économie, et d'autre part par le désir de se conformer à la loi newtonienne de proportionnalité des indices de réfraction aux vitesses de la lumière, persuadé que ceci remet en cause les travaux de Fermat. Considérant que quelquefois on a compris le principe d'économie comme un principe de plus court chemin (comme dans la réflexion) et d'autres fois comme un principe de plus court temps (comme dans la réfraction), et qu'il n'y a pas de raison de privilégier le temps par rapport à l'espace (ou vice-versa), Maupertuis emprunte donc une troisième voie : *"le chemin que la nature tient est celui par lequel la quantité d'action est la moindre"*. Dans ce cas, Maupertuis prend pour mesure de l'action la somme des espaces multipliés chacun par la vitesse de propagation, sans aucune justification préalable et en déduit une loi de proportionnalité des sinus pour la réfraction. Il paraît assez clair que Maupertuis a plutôt cherché à dériver des diverses lois connues et admises à l'époque un principe "réconciliateur" en accord avec sa philosophie ; ainsi écrit-il : *"les principes physiques doivent finalement prouver l'existence de Dieu même, et donc être empreints de finalité"*, donc *"notre principe, plus conforme aux idées que nous devons avoir des choses, laisse le monde dans le besoin naturel de la puissance du Créateur et est une suite nécessaire de l'emploi le plus sage de cette puissance"*.

Ce principe vague, qui ne laissera aucun souvenir dans le domaine de l'Optique où seul prévaut le principe de Fermat, va pourtant connaître un développement riche et prometteur dans le domaine de la dynamique. Maupertuis lui-même en réalise les premiers pas, en montrant que ce principe permet de retrouver les lois des chocs mous (ceux pour lesquels la vitesse relative finale est nulle) et des chocs élastiques (qui conservent l'énergie cinétique). Mais c'est à Euler que revient le mérite d'avoir montré la portée de ce principe et à Lagrange celui de l'avoir formalisé clairement et d'avoir énoncé ses conditions d'application.

b. L'énoncé d'Euler

En 1744, Euler publie dans son grand ouvrage *Methodus inveniendi...* un article traitant du mouvement des projectiles dans un milieu non résistant. Il y vérifie que les trajectoires d'un point matériel (dans un certain nombre de cas particuliers connus) correspondent bien à un extremum

de la quantité d'action définie en substance par $\int v ds$, où v est la vitesse du corps et s l'abscisse curviligne décrite.

Ainsi, si aucune force n'agit sur le point, la vitesse est constante et minimiser l'action revient à minimiser la distance parcourue : la trajectoire est donc rectiligne, conformément au principe de l'inertie. De même, Euler applique les résultats de sa méthode variationnelle, démontrés au début de son traité, au cas de la trajectoire parabolique de chute libre des corps. D'après l'hypothèse de Galilée, on a $v^2 = 2gx + \text{cste}$ où x représente l'ordonnée du point ; l'intégrale à minimiser est donc du type $\int ds \sqrt{a + gx} = \int dx \sqrt{(a + gx)(1 + p^2)}$, avec $p = \frac{dy}{dx}$ où y représente l'abscisse du point. Elle obéit bien à la description du problème variationnel le plus simple résolu auparavant par Euler $\int Z(x, y, p) dx$ minimal, avec $N = \frac{\partial Z}{\partial y} = 0$ et $P = \frac{\partial Z}{\partial p} = p \sqrt{(a + gx)/(1 + p^2)}$. L'équation d'Euler démontrée plus haut entraîne donc que l'expression P est constante, Euler montre alors facilement que l'équation différentielle qui s'en déduit est caractéristique d'une parabole. Enfin, et ceci achèvera de convaincre ses lecteurs de la puissance de ce nouveau principe, Euler s'attaque au cas d'un corps soumis à une force centrale et retrouve les principales caractéristiques de la trajectoire.

c. La mise en forme de Lagrange

En 1761, Lagrange publie ses premiers travaux ayant trait à la présentation variationnelle de la mécanique, à l'occasion de son mémoire *Application de la méthode exposée dans le mémoire précédent à la solution de différents problèmes de dynamique* dans les *Miscellanea Taurinensa*. Cette présentation est en effet directement liée aux travaux publiés dans les pages précédentes, exposant les principes généraux du calcul des variations. Elle n'est pas pour autant une simple répétition de ces principes dans le cadre particulier de la dynamique : l'écriture variationnelle de la mécanique a été l'objet de recherches constantes de la part de Lagrange, ainsi que le montre la lecture de sa correspondance avec Euler ; plus qu'une nouvelle exposition, elle constitue pour lui un véritable bouleversement dans la compréhension de cette science.

Ce mémoire *Application de la méthode exposée...* s'appuie explicitement sur les travaux d'Euler publiés dans les appendices du *Methodus inveniendi*, en se proposant d'en élargir le champ d'application à l'ensemble de la dynamique. Lagrange y pose donc :

"PRINCIPE GENERAL. Soient tant de corps qu'on voudra M, M', M'' ...qui agissent les uns sur les autres d'une manière quelconque, et qui soient de plus, si l'on veut, animés par des forces centrales proportionnelles à des fonctions quelconques des distances ; que $s, s', s'', \dots$, dénotent les espaces parcourus par ces corps dans le temps t , et que $u, u', u'', \dots$, soient leurs vitesses à la fin de ce temps ; la formule

$$M \int u ds + M' \int u' ds' + M'' \int u'' ds'' + \dots$$

sera toujours un maximum ou un minimum."

Lagrange s'intéresse d'abord au mouvement d'un "corps M attiré vers tant de centres fixes qu'on voudra par des forces $P, Q, R, \dots$ exprimées par des fonctions quelconques des distances". Le traitement de ce premier problème est assez symptomatique du type de raisonnement de Lagrange, il paraît donc utile de l'examiner en détail. Le principe général s'écrit donc ici

$$\delta \left(\int u ds \right) = 0$$

la masse, supposée constante, n'étant plus prise en compte, et donc virtuellement posée égale à l'unité. Après permutation des signes d'intégration et de variation, et développement de l'intégrand, Lagrange obtient

$$\int (u \delta ds + \delta u ds) = 0.$$

Posant alors la conservation de la force vive, Lagrange écrit

$$\frac{u^2}{2} = \text{const} - \int (P dp + Q dq + R dr + \dots)$$

p, q, r représentant les distances du point matériel aux centres de force. Après différentiation de cette équation à l'aide de δ , on obtient donc,

$$u \delta u = - \int (\delta P dp + P \delta dp + \delta Q dq + Q \delta dq + \delta R dr + R \delta dr + \dots)$$

Lagrange applique alors le procédé caractéristique de sa méthode de variations, consistant à permuter les symboles d et δ , puis à intégrer par parties, et aboutit après quelques manipulations à

$$\int (u \delta ds - P dt \delta p - Q dt \delta q - R dt \delta r - \dots) = 0;$$

utilisant pour cela le fait que $ds = u dt$, et supposant dans un premier temps que P (resp. Q, R) n'est fonction que de p (resp. q, r). Lagrange cherche ensuite à exprimer séparément les deux intégrales introduites, en utilisant un système de coordonnées de référence constitué en premier lieu par les "coordonnées rectangles x, y, z ". Il écrit dès lors ds sous la forme $\sqrt{dx^2 + dy^2 + dz^2}$, et en déduit

$$\int u \delta ds = - \int \left(d \frac{u dx}{ds} \delta x + d \frac{u dy}{ds} \delta y + d \frac{u dz}{ds} \delta z \right) + \frac{u dx}{ds} \delta x + \frac{u dy}{ds} \delta y + \frac{u dz}{ds} \delta z.$$

De même, considérant p, q, r comme fonctions de x, y, z , Lagrange introduit

$$P \delta p + Q \delta q + R \delta r + \dots = \Pi \delta x + \omega \delta y + \Psi \delta z.$$

Reportant ces diverses expressions dans l'équation première, Lagrange aboutit à

$$- \int \left[\left(d \frac{u dx}{ds} + \Pi dt \right) \delta x + \left(d \frac{u dy}{ds} + \omega dt \right) \delta y + \left(d \frac{u dz}{ds} + \Psi dt \right) \delta z \right] + \frac{u dx}{ds} \delta x + \frac{u dy}{ds} \delta y + \frac{u dz}{ds} \delta z = 0.$$

En supposant que les deux extrémités de la trajectoire sont fixes, et arguant du fait que les variations $\delta x, \delta y, \delta z$ sont indépendantes, Lagrange en déduit tout naturellement que chaque intégrand doit être nul et aboutit finalement à un système d'équations "qui serviront à déterminer la courbe décrite par le corps M et sa vitesse à chaque instant" ;

$$d\frac{udx}{ds} + \Pi dt = 0$$

$$d\frac{udy}{ds} + \omega dt = 0$$

$$d\frac{udz}{ds} + \Psi dt = 0$$

ce qui correspond très exactement aux équations, que l'on pourrait qualifier de "newtoniennes", du mouvement à condition de remarquer que $u/ds = 1/dt$ et que (Π, ω, Ψ) apparaissent comme les opposées des composantes des forces suivant les directions repérées par les coordonnées (x, y, z) .

Avec cet exposé, Lagrange est arrivé au terme de son exploitation du principe de moindre action. Et c'est à partir d'un autre principe, que nous ne développerons pas ici, qu'il dérivera les équations fondamentales de la dynamique qui portent désormais son nom ; équations qui s'écrivent en notation moderne :

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = 0$$

où L (le lagrangien) est la différence entre l'énergie cinétique et l'énergie potentielle de la particule, fonction a priori de ses coordonnées q ainsi que de leurs dérivées temporelles $\dot{q}$ et du temps t .

5. Conclusion

De Fermat en passant par les frères Bernoulli jusqu'à Lagrange, nous avons vu s'esquisser l'évolution d'un des grands concepts novateurs de la pensée mathématique, évolution qui paraît indissociable de la notion correspondante qui s'est développée en physique. A travers l'exemple du calcul des variations, je souhaite avoir montré de quelle façon l'historien et le philosophe des sciences se trouvent encore une fois confrontés à la profonde interaction qui lie la formation des objets mathématiques et la mise en oeuvre des principes physiques fondateurs.



Bibliographie

Oeuvres originales

- Pierre de Fermat, *Oeuvres de Fermat*, P.Tannery et C.Henry eds., Paris, 1891-1912.
- *Acta Eruditorum*, Leipzig, juin 1696 et mai 1697.
- L.Euler, *Opera Omnia*, Serie I, vol. XXIV, Carathéodory ed., Bern, 1952.
- P.L. de Maupertuis, "accord de différentes lois de la nature qui avaient jusqu'ici paru incompatibles", *Mémoires de l'Académie des Sciences*, Paris, 1744 ; et *Oeuvres*, Lyon, 1768.
- J.L Lagrange, *Oeuvres*, Serret et Darboux eds., Paris, 1867-1892.

Notons que quelques uns de ces textes sont traduits en anglais dans l'ouvrage :

- D.J.Struik, *A source book in mathematics 1200-1800*, Harvard Univ. Press, Cambridge, 1969.

Articles généraux

On consultera avec profit, pour plus de détails sur certains points particuliers de notre exposé, les ouvrages ou articles suivants :

- H. Goldstine, *A history of the calculus of variations from the 17th through the 19th century*, Springer-Verlag, New-York/ Heidelberg/ Berlin, 1980.
- J. Peiffer, "Le problème de la brachystochrone à travers les relations de Jean I Bernoulli avec L'Hôpital et Varignon", *Studia Leibnitiana*, XVII, 1989.
- J. Peiffer, "Le problème de la brachystochrone, un défi pour les méthodes infinitistes de la fin du XVIIe siècle", *Sciences et techniques en perspective*, XVI, Nantes, 1991.
- M.Blay, *La naissance de la mécanique analytique. La science du mouvement au tournant des XVIIe et XVIIIe siècles*, Presses Univ. de France, 1992.
- C. Fraser, "J.L Lagrange's early contributions to the principles and methods of mechanics", *Archive for the History of Exact Sciences*, 28 (3), 1983.
- C. Fraser, "J.L Lagrange's changing approach to the foundations of the calculus of variations", *Archive for the History of Exact Sciences*, 32 (2), 1985.

ANNEXE 1

JOHANNIS BERNOULLI

Curvatura Radii in Diaphanis non uniformibus, Solutioque Problematis a se in Actis 1696, p. 269, propositi, de invenienda Linea Brachystochrona, id est, in qua grave a dato puncto ad datum punctum brevissimo tempore decurrit; & de Curva Synchrona, seu radiorum unda, construenda.

[...]

(traduction de ce texte p. 56)

Merito quidem miramur, quod HUGENIUS primus invenit, in *Cycloide vulgari* grave facere descensus tautochronos, a quocunque *Cycloidis* puncto incipiat moveri : sed nescio, an non obstupescas plane, cum dixerò, hanc ipsissimam *Cycloidem*, seu *Tautochronam Hugonianam*, esse nostram *Brachystochronam* quaesitam; ad cujus cognitionem duabus viis perveni, indirecta altera, altera directa. Insistendo priori, mirum consensum detexi inter curvitatem radii luminis in medio continue variante, & *Curvam* nostram *brachystochronam*; aliaque observavi, in quibus nescio quid arcani subest, quod proderit in Dioptriciis. Quamobrem verum erit, quod in propositione Problematis asserueram, *non in nuda speculatione, sed in aliis scientiis*, in Dioptriciis puta, *usam habere quam maximum*. Sed ut quæ diximus re ipsa confirmantur, en priorem solvendi modum !

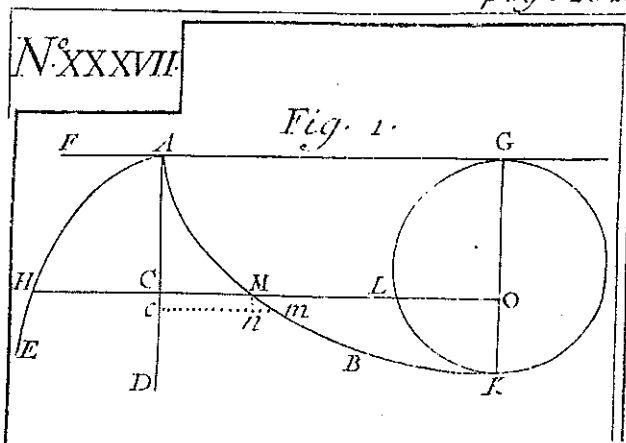
FERMATIUS in Epistola ad DE LA CHAMBRE (Vid. *Epist. CARTESII* Lat. Tom. III, p. 147, & FERMATI Opera Mathem. p. 156 seqq.) stabilivit, radium luminis ex medio rariore in densius transeuntem ita refringi ad perpendicularem, ut, habita ratione temporis, radius (qui a puncto luminante ad punctum illuminatum successive procedere supponitur) viam faciam brevissimam : ex quo principio ostendit, sinum anguli incidentiæ esse ad sinum anguli refractionis in mediorum ratione data directa raritatum, vel reciproca densitatum; id est in ipsa ratione velocitatum, quibus radius media penetrat. Quod postea acutissimus LEIBNITIUS in *Act. Erud.* 1682, p. 185 seqq. & mox Celeb. HUGENIUS in suo Tractatu de Lumine, p. 40, succinctius demonstrarunt, ipsumque principium physicum, vel metaphysicum potius, quod FERMATIUS sua demonstratione geometrica contentus, & facile nimis de jure suo decedens, CLERSELERIO urgente, deferuisse videtur, validissimis argumentis adstruxerunt.

Si nunc concipiamus medium non uniformiter densum, sed

velut per infinitas lamellas horizontaliter interjectas distinctum, quarum interstitia sint repleta materia diaphana raritatis certa ratione accrescentis vel decrescentis; manifestum est, radium, quem ut globulum consideramus, non emanaturum in linea recta, sed in curva quadam (notante id jam & ipso HUGENIO in eodem Tractatu *de Lumine*, sed ipsam curvæ naturam minime determinante) quæ ejus sit naturæ, ut globulus per illam decurrens celeritate continue aucta vel diminuta, pro ratione graduum raritatis, brevissimo tempore perveniat a puncto ad punctum. Constat quoque, cum sinus refractionum in singulis punctis sint respectu ut raritates medii vel celeritates globuli, curvam habere eam proprietatem, ut sinus inclinationum suarum ad lineam verticalem sint ubique in eadem ratione celeritatum. Quibus præmissis, nullo negotio perspicitur, *Curvam Brachystochronam* illam ipsam esse, quam formaret radius transiens per medium, cujus raritates essent in ratione velocitatum, quas grave verticaliter cadendo acquireret. Sive enim velocitatum incrementa dependeant a natura medii magis minusve resistentis, ut in radio; sive abstrahatur a medio, & ab alia causa acceleratio eadem tamen lege generari intelligatur, ut in gravi; cum utroque in casu curva brevissimo tempore percurri supponatur, quid vetat, quo minus altera in alterius locum substitui possit?

TAB. IX.
Nº.
XXXVII.
Fig. 1.

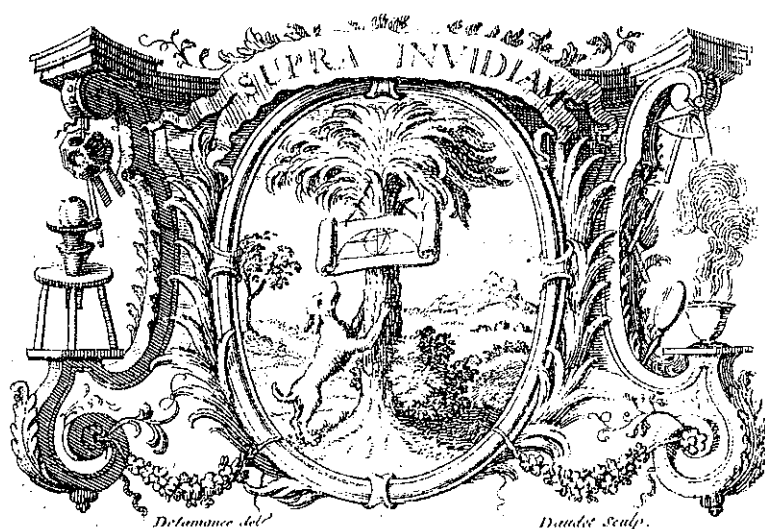
Sic generaliter solvere licet Problema nostrum, quamcunque statuamus accelerationis legem. Eo enim reductum est, ut quæretur curvatura radii in medio secundum raritates, prout libuerit, variante. Esto ergo medium FGD (*Fig. 1*) terminatum ab horizontali FG, in qua punctum radians A, verticalis AD axis curvæ datæ AHE, cujus applicatæ HC determinant raritates medii in altitudinibus AC, vel velocitates radii, seu globuli, in punctis M; radius incurvatus ipse qui quæritur, AMB. Vocentur AC, x ; CH, t ; CM, y ; differentialis Cc, dx ; diff. nm, dy ; diff. Mm, dz ; constans quadam ad arbitrium assumpta, a ; Erit accepta Mm pro sinu toto, mn sinus anguli refractionis seu inclinationis curvæ ad verticalem, & proinde per ea, quæ modo diximus, mn est ad HC in ratione constante, id est, $dy : t = dz : a$; quod hanc suggerit æquationem, $ady = tdz$, seu $aady^2 = ttdx^2 = ttdx^2 + ttdy^2$; quæ reducta generalem dabit æquationem differentialem $dy = tdx : \sqrt{(aa - tt)}$ pro curva AMB quæsita. Atque adeo una opera duo insignia problemata, opticum unum, mechanicum alterum, ultra quam ab aliis petebam, resolvi; ostendique, quamvis ex diversissimis Matheseos partibus sint desumpta, ejusdem tamen esse naturæ.



Sumamus jam specialem casum, & quidem hypothesin communem a GALILÆO primitus introductam & demonstratam, quod velocitates gravium cadentium sint in ratione subduplicata altitudinum emensarum; in hoc enim proprie quæstionis tenor consistit. Quo supposito, curva data AHE erit parabola id est, $tt = ax$ & $t = \sqrt{ax}$, quæ si substituantur in æquatione generali, habebitur hæc $dy = dx \sqrt{\frac{x}{a-x}}$ ex qua concludo Curvam *Brachystochronam* esse *Cycloidem vulgarem*.

Jean Bernoulli

Opera Omnia, I, p189-191



La courbure d'un rayon dans un milieu diaphane uniforme, et la solution du problème proposé par lui-même dans les Actes de 1696, p.269, afin de trouver la Ligne Brachystochrone, c'est-à-dire celle sur laquelle un corps pesant tombe d'un point donné à un autre point donné dans le temps le plus bref; et de construire la Courbe Sychrone, de l'onde des rayons.

[..] Nous admirons à juste titre, le fait que Huygens fut le premier à découvrir qu'un point pesant chute dans le même temps [de façon tautochrone] le long d'une cycloïde quelconque, quelque soit le point de la cycloïde où le mouvement commence : mais je ne sais si on ne sera pas plus admiratif quand j'aurais dit que la *Cycloïde*, ou la *Tautochrone de Huygens*, est bel et bien notre *Brachystochrone* recherchée ; et j'ai acquis cette connaissance de deux manières, l'une indirecte et l'autre directe. Quand j'ai suivi la première, j'ai découvert un admirable accord entre la trajectoire courbe d'un rayon lumineux dans un milieu continûment variable et notre *Courbe Brachystochrone* ; j'ai aussi trouvé d'autres points, dans lesquels il y a je ne sais quoi de mystérieux qui sera utile pour la Dioptrique. Voilà pourquoi il sera vrai, comme je l'avais affirmé dans l'exposé du problème, que ce n'est pas du tout dans la pure spéculation, mais dans d'autres sciences, par exemple la Dioptrique, qu'on tirera le plus grand profit. Mais afin de confirmer mes dires par des faits, voici une première façon de donner une preuve !

Fermat, dans une lettre à de La Chambre [..], a montré qu'un rayon lumineux passant d'un milieu rare à un milieu dense, est dévié vers la perpendiculaire d'une façon telle que (en supposant que le rayon se meuve continûment depuis le point lumineux jusqu'au point illuminé) il suive le chemin qui nécessite le temps le plus court : avec l'aide de ce principe, il montre que le sinus de l'angle d'incidence et le sinus de l'angle de réfraction sont en raison directe de la rareté des milieux, c'est-à-dire dans le rapport même des vitesses avec lesquelles le rayon pénètre ces milieux. C'est ce que par la suite Leibniz, en esprit très fin, dans les *Actes des Savants* 1682, p.185 et peu après le célèbre Huygens dans son *Traité de la Lumière*, p.40, ont démontré de façon plus concise et par de très solides arguments, et ont établi le principe physique, ou mieux métaphysique, lui-même, que Fermat semble avoir abandonné sur l'insistance de Clerselier, se contentant de sa preuve géométrique et renonçant à ses droits trop légèrement.

Maintenant, nous allons considérer un milieu qui n'est pas uniformément dense, mais qui est séparé par des lamelles infinies superposées horizontalement, dont les interstices sont remplis d'une matière diaphane d'une certaine rareté augmentant ou décroissant suivant une loi définie. Il est alors manifeste qu'un rayon, que nous considérerons comme une particule, ne se propagera pas en ligne droite, mais le long d'une trajectoire courbe (ceci a déjà été noté par Huygens dans le *Traité de la Lumière* mentionné ci-dessus, mais il n'a pas déterminé du tout la nature elle-même de la courbe), nature qui est telle que la particule, dont les accroissements ou diminutions de

vitesse dépendent des degrés de rareté du milieu, passera de point en point dans le temps le plus court. Il est établi que, puisque les sinus des angles de réfraction aux points de séparation sont les uns aux autres comme les raretés des milieux, ou comme les vitesses des particules, la courbe a la propriété suivante : les sinus de ses angles d'inclinaison par rapport à la verticale sont partout dans la proportion des vitesses. Mais maintenant, ceci ayant été dit, il apparaît sans peine que la *Courbe Brachystochrone* est la courbe que suivrait un rayon lumineux lors de la traversée d'un milieu dont les raretés seraient dans le rapport des vitesses qu'un corps pesant acquièrerait durant sa chute verticale. En fait, que l'augmentation des vitesses dépende de la nature d'un milieu plus ou moins résistant, comme dans le cas du rayon, ou que nous fassions abstraction du milieu et considérons que, à partir d'une autre cause, la même accélération peut être conçue comme produite par une loi donnée, comme dans le cas du corps pesant ; dans les deux cas, la courbe est parcourue dans le temps le plus court, qui nous interdit de remplacer l'un par l'autre ?

De cette façon, nous pouvons résoudre le problème pour une loi arbitraire de l'accélération. Il est alors réduit à la détermination du chemin suivi par un rayon lumineux à travers un milieu de densité continûment variable. Donc soit FGD (fig.1) le milieu limité par la ligne horizontale FG sur laquelle le point lumineux A est situé ; soit la courbe donnée AHE , avec les axes verticaux AD , ses ordonnées HC déterminant les raretés du milieu à l'altitude AC ou les vitesses des rayons lumineux ou des particules au point M ; soit AMB la courbe du rayon lumineux que nous recherchons. Notons AC, x ; CH, t ; CM, y ; et les différentielles Cc, dx ; les différentielles nm, dy ; les différentielles Mm, dz ; enfin, soit a une constante arbitraire. Alors Mm est le sinus total, mn le sinus de l'angle de réfraction ou de l'angle d'inclinaison de la courbe par rapport à la verticale, et comme nous l'avons vu auparavant, le rapport de mn à HC est constant, c'est-à-dire $dy : t = dz : a$; d'où l'équation $ady = tdz$, ou $aady^2 = ttdz^2 = ttdx^2 + ttdy^2$. Ce qui donne une équation différentielle générale pour la courbe cherchée AMB : $dy = tdx : \sqrt{aa - tt}$. De cette manière, j'ai résolu en une seule fois deux problèmes importants, l'un optique, l'autre mécanique, et j'ai réalisé plus que ce que je demandais aux autres ; j'ai montré que les deux problèmes, pris dans deux parties entièrement séparées des Mathématiques, sont de même nature.

Maintenant prenons un cas spécial, à savoir l'hypothèse commune introduite et démontrée en premier par Galilée, selon laquelle les vitesses des corps en chute libre sont les unes aux autres comme les racines carrées [en raison sous-doublée] des altitudes parcourues ; c'est en cela que consiste réellement la question. Sous cette hypothèse, la courbe donnée AHE sera une parabole, c'est-à-dire que $tt = ax$ et $t = \sqrt{ax}$, si cette valeur est substituée dans l'équation générale, nous obtiendrons : $dy = \sqrt{\frac{x}{a-x}}$, d'où je conclus que la courbe *Brachystochrone* est la *Cycloïde* ordinaire.

Jean Bernoulli, *Opera omnia*, I, p187-193.

trad. F.Viot, avec l'amical soutien de C.Beaujean et J.C.Bos.

ANNEXE 2

J.L. Lagrange

Oeuvres, I, p334-364

ESSAI D'UNE NOUVELLE METHODE

POUR

DÉTERMINER LES MAXIMA ET LES MINIMA

DEN

FORMULES INTÉGRALES INDÉFINIES.

(*Miscellanea Taurinensia*, t. II, 1760-1761.)

Pour peu qu'on soit au fait des principes du Calcul différentiel, on connaît la méthode de déterminer les plus grandes et les moindres ordonnées des courbes; mais il est des questions *de maximis* et *minimis* d'un genre plus élevé et qui, quoique dépendantes de la même méthode, ne s'y appliquent pas si aisément. Ce sont celles où il s'agit de trouver les courbes mêmes, dans lesquelles une expression-intégrale donnée soit un maximum ou un minimum par rapport à toutes les autres courbes.

Le premier Problème de ce genre, que les Géomètres aient résolu, est celui de la *Brachistochrone*, ou ligne de la plus vite descente, que M. Jean Bernoulli proposa vers la fin du siècle passé. On n'y parvint alors que par des voies particulières, et ce ne fut que quelque temps après, et à l'occasion des recherches sur les *Isopérimètres*, que le grand Géomètre dont nous venons de parler et son illustre frère M. Jacques Bernoulli, donnèrent quelques règles générales pour résoudre plusieurs autres questions de même nature. Mais ces règles n'ayant pas assez d'étendue, le célèbre M. Euler a entrepris de réduire toutes les recherches de ce genre à une méthode générale, dans l'ouvrage intitulé : *Methodus inveniendi lineas curvas maximi, minimive proprietate gaudentes : sive solutio*

Problematis isoperimetrici latissimo sensu accepti; ouvrage original et qui brille partout d'une profonde science du calcul. Cependant, quelque ingénieuse et féconde que soit sa méthode, il faut avouer qu'elle n'a pas toute la simplicité qu'on peut désirer dans un sujet de pure analyse. L'Auteur le fait sentir lui-même dans l'Article 39 du Chapitre II de son livre, par ces paroles : « Desideratur itaque methodus a resolutione geometrica et lineari libera, qua pateat in tali investigatione maximi minimique, loco Pdp scribi debere — $p dP$. »

Maintenant voici une méthode qui ne demande qu'un usage fort simple des principes du Calcul différentiel et intégral; mais ayant tout je dois avertir que, comme cette méthode exige que les mêmes quantités varient de deux manières différentes, pour ne pas confondre ces variations, j'ai introduit dans mes calculs une nouvelle caractéristique δ . Ainsi δZ exprimera une différence de Z qui ne sera pas la même que dZ , mais qui sera cependant formée par les mêmes règles; de sorte qu'ayant une équation quelconque $dZ = m dx$, on pourra avoir également $\delta Z = m \delta x$, et ainsi des autres. Cela posé, je viens d'abord au Problème suivant.

1.

PROBLÈME I. — *Étant proposée une formule intégrale indéfinie représentée par $\int Z$, où Z désigne une fonction quelconque déterminée des variables x, y, z , et de leurs différences $dx, dy, dz, d^2x, d^2y, d^2z, \dots$, trouver la relation que ces variables doivent avoir entre elles, pour que la formule $\int Z$ devienne un maximum ou un minimum.*

SOLUTION. — Suivant la méthode connue de *maximis* et *minimis*, il faudra différencier la proposée $\int Z$, en regardant les quantités $x, y, z, dx, dy, dz, d^2x, d^2y, d^2z, \dots$ comme variables, et faire la différentielle, qui en résulte, égale à zéro. Marquant donc ces variations par δ , on aura d'abord, pour l'équation du maximum ou minimum,

$$\delta \int Z = 0,$$

ou, ce qui en est l'équivalent,

$$\int \delta Z = 0.$$

Or, soit Z tel que

$$\begin{aligned} \delta Z = & n \delta x + p \delta dx + q \delta d^2x + r \delta d^3x + \dots \\ & + N \delta y + P \delta dy + Q \delta d^2y + R \delta d^3y + \dots \\ & + v \delta z + w \delta dz + \chi \delta d^2z + \rho \delta d^3z + \dots, \end{aligned}$$

il en viendra l'équation

$$\begin{aligned} & \int n \partial x + \int p \partial dx + \int q \partial d^2 x + \int r \partial d^3 x + \dots \\ & + \int N \partial y + \int P \partial dy + \int Q \partial d^2 y + \int R \partial d^3 y + \dots \\ & + \int v \partial z + \int w \partial dz + \int \chi \partial d^2 z + \int \rho \partial d^3 z + \dots = 0; \end{aligned}$$

mais on comprend aisément que

$$\partial dx = d \partial x, \quad \partial d^2 x = d^2 \partial x,$$

et ainsi des autres; de plus, on trouve, par la méthode des intégrations par parties,

$$\begin{aligned} \int p d \partial x &= p \partial x - \int dp \partial x, \\ - \int q d^2 \partial x &= q d \partial x - dq \partial x + \int d^2 q \partial x, \\ \int r d^3 \partial x &= r d^2 \partial x - dr d \partial x + d^2 r \partial x - \int d^3 r \partial x, \end{aligned}$$

et ainsi du reste; donc l'équation précédente se changera en celle-ci :

$$(A) \quad \left\{ \begin{aligned} & \int (n - dp + d^2 q - d^3 r + \dots) \partial x \\ & + \int (N - dP + d^2 Q - d^3 R + \dots) \partial y \\ & + \int (v - d\omega + d^2 \chi - d^3 \rho + \dots) \partial z \\ & + (p - dq + d^2 r - \dots) \partial x + (q - dr + \dots) d \partial x + (r - \dots) d^2 \partial x + \dots \\ & + (P - dQ + d^2 R - \dots) \partial y + (Q - dR + \dots) d \partial y + (R - \dots) d^2 \partial y + \dots \\ & + (\omega - d\chi + d^2 \rho - \dots) \partial z + (\chi - d\rho + \dots) d \partial z + (\rho - \dots) d^2 \partial z + \dots = 0; \end{aligned} \right.$$

d'où l'on tirera premièrement l'équation indéfinie

$$(B) \quad \left\{ \begin{aligned} & (n - dp + d^2 q - d^3 r + \dots) \partial x \\ & + (N - dP + d^2 Q - d^3 R + \dots) \partial y \\ & + (v - d\omega + d^2 \chi - d^3 \rho + \dots) \partial z = 0, \end{aligned} \right.$$

et ensuite l'équation déterminée

$$(C) \quad \left\{ \begin{aligned} & (p - dq + d^2 r - \dots) \partial x + (q - dr + \dots) d \partial x + (r - \dots) d^2 \partial x + \dots \\ & + (P - dQ + d^2 R - \dots) \partial y + (Q - dR + \dots) d \partial y + (R - \dots) d^2 \partial y + \dots \\ & + (\omega - d\chi + d^2 \rho - \dots) \partial z + (\chi - d\rho + \dots) d \partial z + (\rho - \dots) d^2 \partial z + \dots = 0. \end{aligned} \right.$$

Cette équation se rapporte au dernier point de l'intégrale $\int Z$; mais il faut observer que, comme chacun de ses termes comme $p \partial x$ dépend d'une intégration partielle de la formule $\int p d \partial x$, on peut lui ajouter ou en retrancher une quantité constante. Or, la condition par laquelle cette constante doit se déterminer est qu'elle fasse évanouir le terme $p \partial x$ au point où commence l'intégrale $\int p d \partial x$; il faudra donc retrancher de $p \partial x$

sa valeur en ce point; d'où résulte la règle suivante. Soit le premier membre de l'équation (C), exprimé généralement par M, et soit la valeur de M, au point où commence l'intégrale $\int Z$, désignée par M, et au point où cette intégrale finit, désignée par M', on aura $M' - M = 0$ pour l'expression complète de l'équation (C).

Maintenant, pour se défaire dans les équations trouvées des différences indéterminées $\delta x, \delta y, \delta z, d\delta x, d\delta y, \dots$, on examinera d'abord si, par la nature du Problème, il y a entre elles quelque rapport donné, et les ayant réduites au plus petit nombre possible, on fera ensuite le coefficient de chacune de celles qui resteront égales à zéro. Si elles sont absolument indépendantes les unes des autres, l'équation (B) nous donnera sur-le-champ les trois suivantes :

$$\begin{aligned} n - dp + d^2 q - d^3 r + \dots &= 0, \\ N - dP + d^2 Q - d^3 R + \dots &= 0, \\ v - d\varpi + d^2 \gamma - d^3 \rho + \dots &= 0. \end{aligned}$$

II.

EXEMPLE. — Soit cherchée la courbe brachistochrone dans le vide. Nommant x l'abscisse verticale, et y et z les deux ordonnées horizontales et perpendiculaires l'une à l'autre, la formule qui exprime le temps sera

$$\int \frac{\sqrt{dx^2 + dy^2 + dz^2}}{\sqrt{x}},$$

laquelle étant comparée à $\int Z$, on a

$$Z = \frac{\sqrt{dx^2 + dy^2 + dz^2}}{\sqrt{x}};$$

et différentiant par δ , suivant les règles ordinaires des différentiations,

$$\begin{aligned} \delta Z = & - \frac{\delta x \sqrt{dx^2 + dy^2 + dz^2}}{2x\sqrt{x}} + \frac{dx \delta dx}{\sqrt{x} \sqrt{dx^2 + dy^2 + dz^2}} \\ & + \frac{dy \delta dy}{\sqrt{x} \sqrt{dx^2 + dy^2 + dz^2}} + \frac{dz \delta dz}{\sqrt{x} \sqrt{dx^2 + dy^2 + dz^2}}; \end{aligned}$$

donc, posant, pour abréger,

$$ds = \sqrt{dx^2 + dy^2 + dz^2},$$

on a

$$n = - \frac{ds}{2x\sqrt{x}}, \quad p = \frac{dx}{\sqrt{x} ds}, \quad P = \frac{dy}{\sqrt{x} ds}, \quad \varpi = \frac{dz}{\sqrt{x} ds},$$

et toutes les autres quantités $q, r, N, Q, \dots$, égales à zéro.

III.

Premier cas. — Or, si le Problème est de trouver en général, entre toutes les courbes possibles, celle de la plus vite descente, on aura en ce cas les équations

$$n - dp = 0, \quad -dP = 0, \quad -d\pi = 0,$$

savoir

$$-\frac{ds}{2x dx} - d \frac{dx}{\sqrt{x} ds} = 0, \quad -d \frac{dy}{\sqrt{x} ds} = 0, \quad -d \frac{dz}{\sqrt{x} ds} = 0;$$

ces trois équations devant représenter une courbe unique, il faut qu'elles se réduisent à deux seulement : c'est de quoi il est facile de s'assurer par le calcul, car la seconde étant multipliée par $2 \frac{dy}{\sqrt{x} ds}$ et ajoutée à la troisième multipliée par $2 \frac{dz}{\sqrt{x} ds}$, il vient, à cause de $ds^2 = dx^2 + dy^2 + dz^2$,

$$d \left(\frac{1}{x} - \frac{dx^2}{x ds^2} \right) = 0,$$

savoir, en différentiant et divisant le tout par $\frac{2 dx}{\sqrt{x} ds}$,

$$-\frac{ds}{2x\sqrt{x}} - d \frac{dx}{\sqrt{x} ds} = 0,$$

qui est la première équation.

Présentement, si l'on intègre les deux équations

$$d \frac{dy}{\sqrt{x} ds} = 0, \quad d \frac{dz}{\sqrt{x} ds} = 0,$$

on a

$$\frac{dy}{\sqrt{x} ds} = \frac{1}{\sqrt{a}}, \quad \frac{dz}{\sqrt{x} ds} = \frac{1}{\sqrt{b}},$$

d'où l'on tire d'abord

$$\frac{dy}{dz} = \frac{\sqrt{b}}{\sqrt{a}},$$

ce qui fait voir que la courbe cherchée est toute dans un même plan vertical, et que par conséquent elle est à simple courbure. Pour la mieux connaître, rapportons-la à deux coordonnées prises dans son même plan. Que x soit l'une et t l'autre, on aura

$$\sqrt{y^2 + z^2} = t,$$

et puisque $\frac{dy}{dz} = \frac{\sqrt{b}}{\sqrt{a}}$, on aura en intégrant, sans ajouter de constante, parce que je suppose que l'axe des x passe par la courbe même,

$$z = y \frac{\sqrt{a}}{\sqrt{b}};$$

d'où l'on tire

$$z = t \frac{\sqrt{a}}{\sqrt{a+b}}, \quad y = t \frac{\sqrt{b}}{\sqrt{a+b}}, \quad dy = dt \frac{\sqrt{b}}{\sqrt{a+b}}, \quad ds = \sqrt{dx^2 + dt^2},$$

et enfin

$$\frac{dy}{\sqrt{x} ds} = \frac{\sqrt{b}}{\sqrt{a+b} \sqrt{x}} \frac{dt}{\sqrt{dx^2 + dt^2}} = \frac{1}{\sqrt{a}},$$

ce qui se réduit, en posant $\frac{ab}{a+b} = c$, à

$$dt = \frac{\sqrt{x} dx}{\sqrt{c-x}},$$

équation d'une cycloïde décrite sur une base horizontale par un cercle, dont le diamètre est égal à c .



Portrait de Lagrange