From Lagrangian Mechanics to Optimal Control and PDE Constraints

Martin J. Gander
martin.gander@unige.ch

University of Geneva

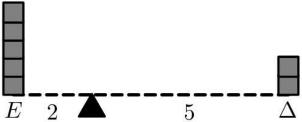
Joint work with Felix Kwok and Gerhard Wanner

MORE workshop, Prague, September 2014

Archimedes' Law of the Lever

Archimedes (287-212 BC): "Two bodies are in equilibrium if their weights are inversely proportional to their arm length"

(Opera, printed 1615 in Paris, BGE Ka459)



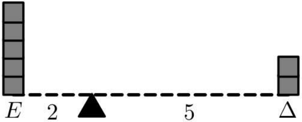
Prop. 6 (rat.) and 7 (irrat.)

Archimedes' Law of the Lever

Archimedes (287-212 BC): "Two bodies are in equilibrium if their weights are inversely proportional to their arm length"


ΑΡΧΙΜΗΔΟΥΣ
ΕΠΙΠΕΔΩΝ ΙΣΟΡΡΟ-
ΠΙΩΝ, Η ΚΕΝΤΡΑΒΑΡΩΝ

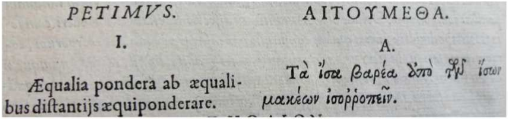
(Opera, printed 1615 in Paris, BGE Ka459)



Prop. 6 (rat.) and 7 (irrat.)

Beautiful proof of Archimedes:

1) Axiom: equal weights at equal distances are in equilibrium


PETIMVS.

I.

Æqualia pondera ab æqualibus diſtantijs æquiponderare.

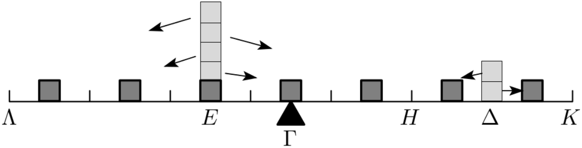
ΑΙΤΟΥΜΕΘΑ.

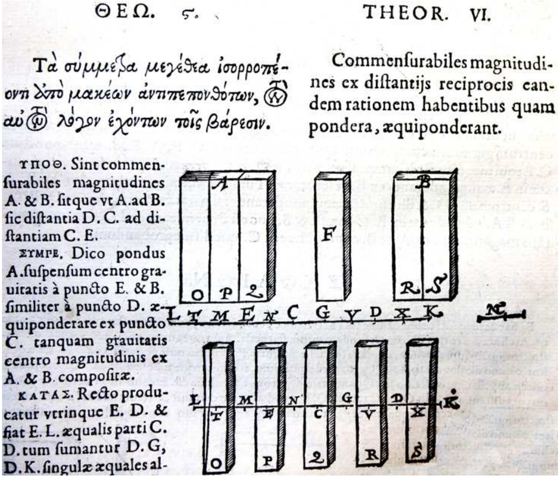
Α.

Τὰ ἴσα βαρέα ἀπὸ τῶν ἴσων μακέων ἰσορροπεῖν.

(Opera 1615 (Paris BGE Ka459)

2) Redistribute weights to reach symmetric configuration





ΘΕΩ. ς.   THEOR. VI.

Τὰ σύμμετρα μεγέθεα ἰσορροπέοντι ἀπὸ μακέων ἀντιπεπονθότων, τὸν αὐτὸν λόγον ἐχόντων τοῖς βάρεσιν.

Commenſurabiles magnitudines ex diſtantijs reciprocis eandem rationem habentibus quam pondera, æquiponderant.

ΥΠΟΘ. Sint commenſurabiles magnitudines A. & B. ſitque vt A. ad B. ſic diſtantia D. C. ad diſtantiam C. E.

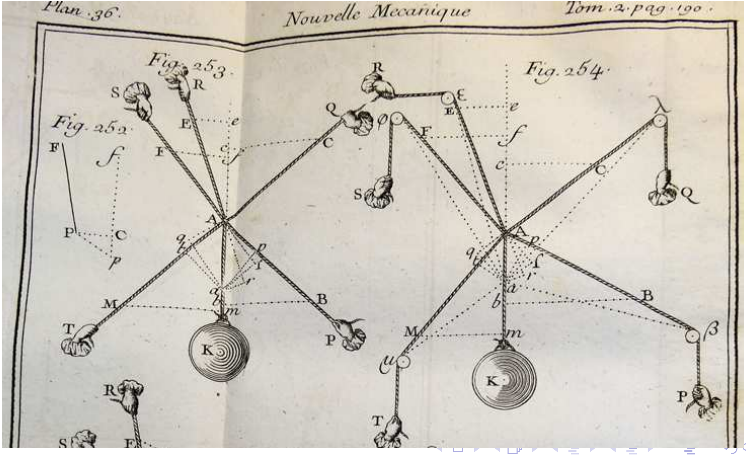
ΣΥΜΠΕ. Dico pondus A. ſuſpenſum centro grauitatis à puncto E. & B. ſimiliter à puncto D. æquiponderare ex puncto C. tanquam grauitatis centro magnitudinis ex A. & B. compoſitæ.

ΚΑΤΑΣ. Recto producatur vtrinque E. D. & fiat E. L. æqualis parti C. D. tum ſumantur D. G, D. K. ſingulæ æquales al-

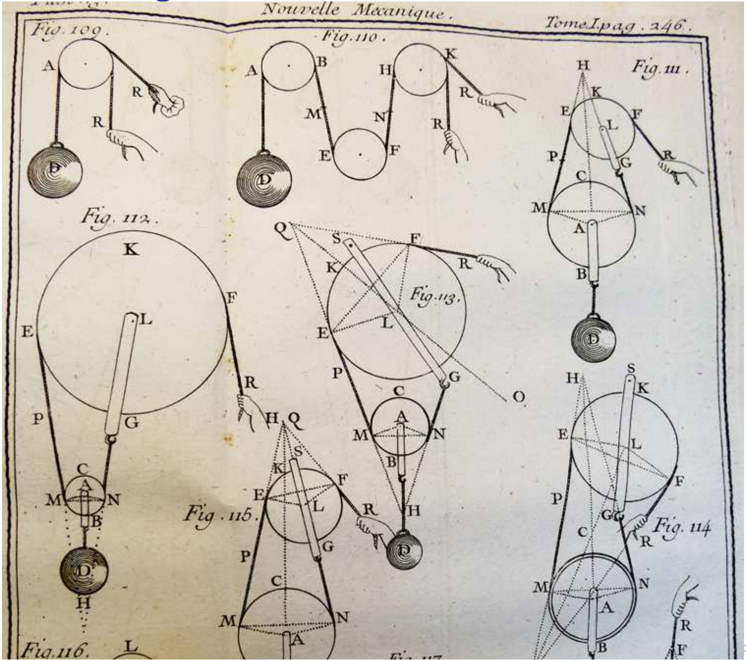
Pierre Varignon

Varignon (1654-1722): Nouvelle Mécanique (frontispiece "dont le projet fut donné en M.DC.LXXXVII") published posthumously in 1725

Hundreds of results illustrated in 64 plates of figures:



Pierre Varignon



Johann Bernoulli

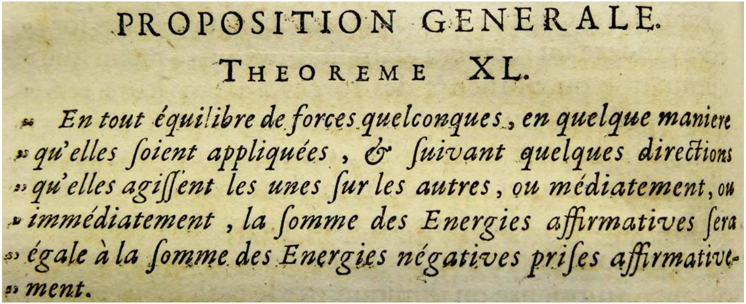
Johann Bernoulli (1667-1748): Invents a general rule ("règle") and announces it to Mr. le Chev. Renau (Bernard Renau d'Eliçagaray, "le petit Renau") in a letter, with copy to Varignon:

> *"Votre projet d'une nouvelle mechanique fourmille d'un grand nombre d'exemples, dont quelques uns à en juger par les figures paroissent assez compliqués; mais je vous deffie de m'en proposer un à votre choix, que je ne resolve sur le champ et comme en jouant par ma dite règle."*

> *"Your project of a new theory of mechanics is swarming of examples, some of which, judging from the figures, appear to be quite complicated; but I challenge you to propose any one of them to me, and I solve it on the spot with my so-called rule"*

Varignon had difficulties admitting that all his work over decades was declared to be so easy and contested the general truth of this rule.

Bernoulli's Rule in Varignon's Book



PROPOSITION GENERALE.

THEOREME XL.

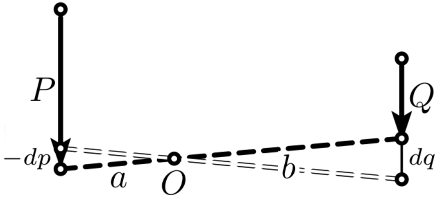
» En tout équilibre de forces quelconques, en quelque maniere
» qu'elles ſoient appliquées, & ſuivant quelques directions
» qu'elles agiſſent les unes ſur les autres, ou médiatement, ou
» immédiatement, la ſomme des Energies affirmatives ſera
» égale à la ſomme des Energies négatives priſes affirmative-
» ment.

"In every equilibrium of arbitrary forces, no matter how they are applied, and in which directions they act the ones on the others, either indirectly or directly, the sum of the positive energies will be equal to the sum of the negative energies taken positively."

Varignon gives however the wrong date 1717 for the letter of Bernoulli, which was later copied by Lagrange.

Lagrange Explains Bernoulli's Rule

Joseph-Louis Lagrange (1736-1813): "Mécanique analytique" from 1788



Assume the lever is in equilibrium: **Archimedes** implies

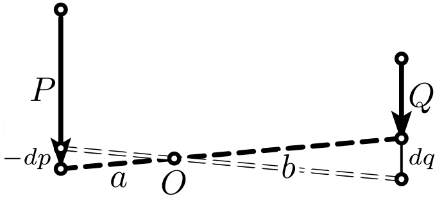
$$Pa = Qb \quad \Longrightarrow \quad \frac{P}{Q} = \frac{b}{a}$$

Bernoulli's idea: apply a "virtual velocity" during an infinitely small time interval to displace the lever. Then

$$a \propto dp \text{ and } b \propto dq \implies \frac{P}{Q} = -\frac{dq}{dp} \implies Pdp + Qdq = 0$$

So Forget Archimedes !

Applying the virtual velocity principle,



we get

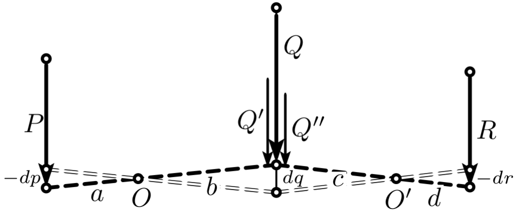
$$Pdp + Qdq = 0.$$

Now if $dp = dx$ for an arbitrary dx, then $dq = -\frac{b}{a}dx$, and hence

$$Pdp + Qdq = 0 \quad \Longrightarrow \quad \left(P - Q\frac{b}{a}\right)dx = 0 \quad \forall dx$$

$$\Longrightarrow \quad \frac{P}{Q} = \frac{b}{a} \quad \text{(Archimedes)}$$

A More Complicated System



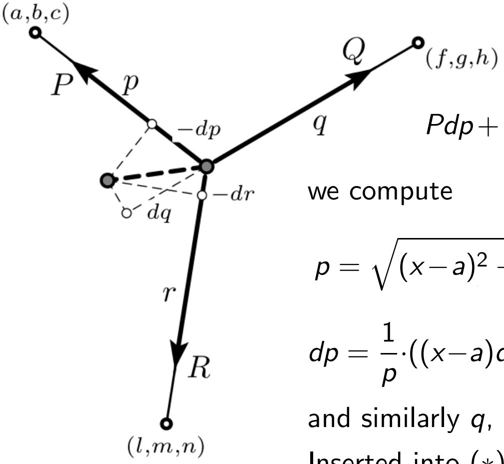
Applying the rule of Bernoulli, we get

$$Pdp + Qdq + Rdr = 0 \qquad (*)$$

If $dp = dx$, then $dq = -\frac{b}{a}dx$, and $dr = -\frac{d}{c}dq = \frac{bd}{ac}dx$, and

$$(*) \implies \left(P - Q\frac{b}{a} + \frac{bd}{ac}R\right) dx = 0 \quad \forall dx$$

A System of Varignon Style



To use

$$Pdp + Qdq + Rdr = 0 \quad (*)$$

we compute

$$p = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$$

$$dp = \frac{1}{p}\cdot((x-a)dx + (y-b)dy + (z-c)dz)$$

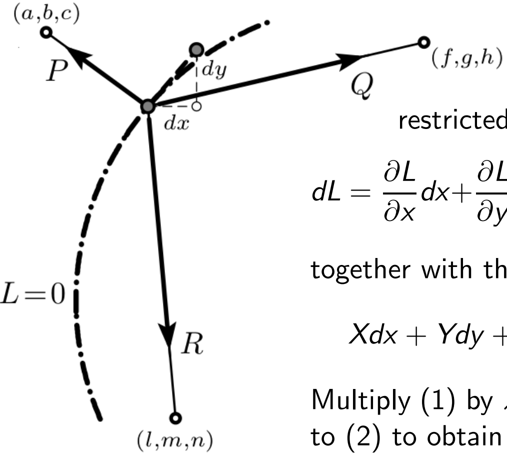
and similarly q, dq, and r, dr.

Inserted into $(*)$ we get for equilibrium

$$Xdx + Ydy + Zdz = 0.$$

with $X = P\frac{x-a}{p} + Q\frac{x-f}{q} + R\frac{x-l}{r}$, $Y = P\frac{y-b}{p} + Q\frac{y-g}{q} + R\frac{y-m}{r}$ and $Z = P\frac{z-c}{p} + Q\frac{z-h}{q} + R\frac{z-n}{r}$.

Same System with Constraint



Motion (dx, dy, dz) restricted to a surface $L = 0$:

$$dL = \frac{\partial L}{\partial x}dx + \frac{\partial L}{\partial y}dy + \frac{\partial L}{\partial z}dz = 0 \quad (1)$$

together with the equation

$$Xdx + Ydy + Zdz = 0 \qquad (2)$$

Multiply (1) by $\lambda = -Z/\frac{\partial L}{\partial z}$ and add to (2) to obtain

$$\left(X + \lambda\frac{\partial L}{\partial x}\right) \cdot dx + \left(Y + \lambda\frac{\partial L}{\partial y}\right) \cdot dy = 0\,, \quad Z + \lambda\frac{\partial L}{\partial z} = 0\,.$$

This means applying the virtual velocity argument, *without constraints*, to $Xdx + Ydy + Zdz + \lambda dL = 0$

"Il n'est pas difficile de prouver par la théorie de l'élimination des équations linéaires..."

The Multiplier Method

Physical interpretation of λ by Lagrange:

$\lambda(\frac{\partial L}{\partial x}, \frac{\partial L}{\partial y}, \frac{\partial L}{\partial z})$ represents the force that holds the particle on the surface $L = 0$

"Équation générale" for ALL problems of equilibria:

$$P\,dp + Q\,dq + R\,dr + \&c + \lambda\,dL + \mu\,dM + \nu\,dN + \&c = 0,$$

- "Méthode très-simple" in Section IV of the first edition from 1788.
- In the second edition from 1811, Lagrange baptizes the method

§ I.

Méthode des Multiplicateurs.

Lagrange Multiplier in Optimization

Lagrange, Théorie de Fonctions Analytiques (1797): "Théorie des fonctions analytiques, contenant les principes du calcul différentiel, dégagés de toute considération d'infiniment petits, d'évanouissans, de limites ou de fluxions, et réduits à l'analyse algébrique des quantités finies"

Constrained optimization problem:

$$f(\mathbf{x}) \longrightarrow \max, \quad \mathbf{g}(\mathbf{x}) = 0,$$

with $f : \mathbb{R}^n \to \mathbb{R}$ and $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$, $m < n$.

Elimination of the constraints:

$$\mathbf{x} = (\mathbf{y}, \mathbf{u}), \quad \mathbf{y} \in \mathbb{R}^m, \ \mathbf{u} \in \mathbb{R}^{n-m}$$

$\mathbf{g}(\mathbf{x}) = 0$ and implicit function theorem $\Longrightarrow \mathbf{y} = \mathbf{y}(\mathbf{u})$

Unconstrained optimization problem:

$$f(\mathbf{y}(\mathbf{u}), \mathbf{u}) \longrightarrow \max.$$

Lagrange Multiplier in Optimization

Necessary condition for a local maximum of $f(\mathbf{y}(\mathbf{u}), \mathbf{u})$:

$$\frac{df}{d\mathbf{u}} = \frac{\partial f}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{u}} + \frac{\partial f}{\partial \mathbf{u}} = 0$$

For $\frac{\partial \mathbf{y}}{\partial \mathbf{u}}$, implicitly differentiate constraint $\mathbf{g}(\mathbf{y}(\mathbf{u}), \mathbf{u}) = 0$

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{u}} + \frac{\partial f}{\partial \mathbf{u}} &= 0, \quad n-m \text{ equations} \\ \frac{\partial \mathbf{g}}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{u}} + \frac{\partial \mathbf{g}}{\partial \mathbf{u}} &= 0, \quad m(n-m) \text{ equations} \\ \mathbf{g} &= 0, \quad m \text{ equations} \end{aligned}$$

$n + m(n-m)$ equations for the n unknowns in $\mathbf{y}$ and $\mathbf{u}$ combined, and the $m(n-m)$ unknowns in the Jacobian $\frac{\partial \mathbf{y}}{\partial \mathbf{u}}$

$\Longrightarrow$ **A very big system because of $\frac{\partial \mathbf{y}}{\partial \mathbf{u}}$.**

Gaussian Elimination of Lagrange

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{y}}\frac{\partial \mathbf{y}}{\partial \mathbf{u}} + \frac{\partial f}{\partial \mathbf{u}} &= 0, \quad n-m \text{ equations} \\ \frac{\partial \mathbf{g}}{\partial \mathbf{y}}\frac{\partial \mathbf{y}}{\partial \mathbf{u}} + \frac{\partial \mathbf{g}}{\partial \mathbf{u}} &= 0, \quad m(n-m) \text{ equations} \\ \mathbf{g} &= 0, \quad m \text{ equations} \end{aligned}$$

Multiply 2nd equation by $\lambda := -\frac{\partial f}{\partial \mathbf{y}}\left(\frac{\partial \mathbf{g}}{\partial \mathbf{y}}\right)^{-1}$ and add to 1st:

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{u}} + \lambda\frac{\partial \mathbf{g}}{\partial \mathbf{u}} &= 0, \quad n-m \text{ equations} \\ \frac{\partial f}{\partial \mathbf{y}} + \lambda\frac{\partial \mathbf{g}}{\partial \mathbf{y}} &= 0, \quad m \text{ equations} \\ \mathbf{g} &= 0, \quad m \text{ equations} \end{aligned}$$

for n unknowns in $\mathbf{y}$ and $\mathbf{u}$ combined, plus m in λ.
Lagrange: get this directly by differentiating

$$\mathcal{L}(\mathbf{u},\mathbf{y},\lambda) := f(\mathbf{y},\mathbf{u}) + \lambda g(\mathbf{y},\mathbf{u})$$

Hestenes' Optimal Control Problem 1950

$$H(t,q,p,A) \leq H(t,q,p,a)$$

must hold for every admissible element (t,q,A).

Thus, H has a maximum value with respect to a_h along a minimizing curve C_0.

$$\begin{aligned} \int_0^T f(\mathbf{y},\mathbf{u})dt &\longrightarrow \text{max}, \\ \dot{\mathbf{y}} &= \mathbf{g}(\mathbf{y},\mathbf{u}), \\ \mathbf{y}(0) &= \mathbf{y}^0, \\ \mathbf{y}(T) &= \mathbf{y}_T. \end{aligned}$$

Introduce the Lagrangian

$$\mathcal{L}(\mathbf{y},\mathbf{u},\lambda) := \int_0^T f(\mathbf{y},\mathbf{u})dt + \int_0^T \lambda\left(\mathbf{g}(\mathbf{y},\mathbf{u}) - \dot{\mathbf{y}}\right)dt,$$

where all the variables depend on time.

Compute the derivatives with respect to the variables $\mathbf{y}$, $\mathbf{u}$, and λ using variational calculus (as Euler did in E420)!

Variational Derivatives

Computing a derivative with respect to $\mathbf{y}$ of

$$\mathcal{L}(\mathbf{y},\mathbf{u},\lambda) := \int_0^T f(\mathbf{y},\mathbf{u})dt + \int_0^T \lambda\left(\mathbf{g}(\mathbf{y},\mathbf{u}) - \dot{\mathbf{y}}\right)dt,$$

we obtain

$$\begin{aligned}\frac{d}{d\varepsilon}\mathcal{L}(\mathbf{y}+\varepsilon\mathbf{z},\mathbf{u},\lambda)|_{\varepsilon=0} &= \int_0^T \frac{\partial f}{\partial \mathbf{y}}\mathbf{z}dt + \int_0^T \lambda\left(\frac{\partial \mathbf{g}}{\partial \mathbf{y}}\mathbf{z} - \dot{\mathbf{z}}\right)dt \\ &= \int_0^T \left(\frac{\partial f}{\partial \mathbf{y}} + \lambda\frac{\partial \mathbf{g}}{\partial \mathbf{y}} + \dot{\lambda}\right)\mathbf{z} - \lambda\cdot\mathbf{z}|_0^T = 0\end{aligned}$$

Since $\mathbf{z}(0) = \mathbf{z}(T) = 0$, but $\mathbf{z}(t)$ arbitrary, we get (and with similar derivatives w.r.t. $\mathbf{u}$ and λ)

$$\begin{aligned} 0 &= \frac{\partial f}{\partial \mathbf{u}} + \lambda\frac{\partial \mathbf{g}}{\partial \mathbf{u}} \\ \dot{\mathbf{y}} &= \mathbf{g}(\mathbf{y},\mathbf{u}) \qquad \mathbf{y}(0) = \mathbf{y}^0,\ \mathbf{y}(T) = \mathbf{y}_T \\ -\dot{\lambda} &= \frac{\partial f}{\partial \mathbf{y}} + \lambda\frac{\partial \mathbf{g}}{\partial \mathbf{y}} \end{aligned}$$

Hamiltonian Structure

Defining the Hamiltonian function

$$H(\mathbf{y},\mathbf{u},\lambda) := f(\mathbf{y},\mathbf{u}) + \lambda \mathbf{g}(\mathbf{y},\mathbf{u}),$$

we see that the first order optimality system is

$$\begin{aligned} \dot{\mathbf{y}} &= \frac{\partial H}{\partial \lambda} &= \mathbf{g}(\mathbf{y},\mathbf{u}), \\ \dot{\lambda} &= -\frac{\partial H}{\partial \mathbf{y}} &= -\frac{\partial f}{\partial \mathbf{y}} - \lambda \frac{\partial \mathbf{g}}{\partial \mathbf{y}}, \end{aligned}$$

a Hamiltonian system, H is preserved along optimal trajectories!

This was already discovered by **Carathéodory 1926:**

$$H(t, x_i, y_i) = -M(t, x_j, \varphi_j, \chi_{k'}) + \sum_j y_j \varphi_j,$$

$$\dot{x}_i = H_{y_i}, \quad \dot{y}_i = -H_{x_i}$$

Necessary Condition from the Hamiltonian

Instead of maximizing the Lagrangian

$$\mathcal{L}(\mathbf{y},\mathbf{u},\lambda) = \int_0^T f(\mathbf{y},\mathbf{u})dt + \int_0^T \lambda\left(\mathbf{g}(\mathbf{y},\mathbf{u}) - \dot{\mathbf{y}}\right) dt$$

which gives along an optimal trajectory $\dot{\mathbf{y}} = \mathbf{g}(\mathbf{y},\mathbf{u})$

$$\int_0^T f(\mathbf{y},\mathbf{u})dt \longrightarrow \max \quad \text{in } \mathbf{u}(t) \text{ as before,}$$

one could also maximize the Hamiltonian

$$H(\mathbf{y},\mathbf{u},\lambda) := f(\mathbf{y},\mathbf{u}) + \lambda\mathbf{g}(\mathbf{y},\mathbf{u}),$$

pointwise for each $t \in [0, T]$ (Hestenes' RAND report 1950)

$$\begin{aligned} H(\mathbf{y},\mathbf{u},\lambda) &\longrightarrow \max \quad \text{with respect to } \mathbf{u}(t), \\ \dot{\mathbf{y}} &= \frac{\partial H}{\partial \lambda} \\ \dot{\lambda} &= -\frac{\partial H}{\partial \mathbf{y}}. \end{aligned}$$

Discovery of Pontryagin

Boltyanski, Gamkrelidze, and Pontryagin (1956): On the theory of optimal processes (in Russian)

Discovery of Pontryagin

Boltyanski, Gamkrelidze, and Pontryagin (1956): On the theory of optimal processes (in Russian)

- Controls are often constraint, e.g. $-1 \leq u_i \leq 1$

Discovery of Pontryagin

Boltyanski, Gamkrelidze, and Pontryagin (1956): On the theory of optimal processes (in Russian)

- Controls are often constraint, e.g. $-1 \leq u_i \leq 1$
- Solution often on the boundary ("bang-bang")
  - $\ddot{y} = \pm M$ (Feldbaum 1949)
  - $\ddot{y} = \pm u$, $|u| \leq M$ (Feldbaum 1953)

Discovery of Pontryagin

Boltyanski, Gamkrelidze, and Pontryagin (1956): On the theory of optimal processes (in Russian)

- Controls are often constraint, e.g. $-1 \leq u_i \leq 1$
- Solution often on the boundary ("bang-bang")
  - $\ddot{y} = \pm M$ (Feldbaum 1949)
  - $\ddot{y} = \pm u$, $|u| \leq M$ (Feldbaum 1953)
- Pontryagin discovers the Hamiltonian formulation without Lagrange multipliers

  The first and the most important step toward the final solution was made by L.S. right after the formulation of the problem, during three days, or better to say, during three consecutive sleepless nights.

  (Gamkrelidze (1999) "Discovery of the Maximum Principle")

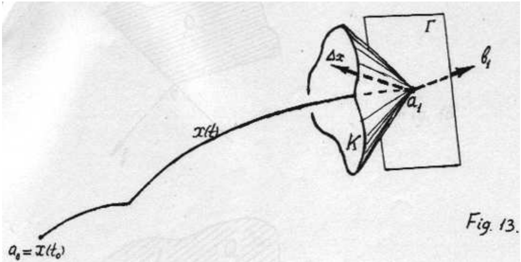
Discovery of Pontryagin

Boltyanski, Gamkrelidze, and Pontryagin (1956): On the theory of optimal processes (in Russian)

- Controls are often constraint, e.g. $-1 \le u_i \le 1$
- Solution often on the boundary ("bang-bang")
  - $\ddot{y} = \pm M$ (Feldbaum 1949)
  - $\ddot{y} = \pm u$, $|u| \le M$ (Feldbaum 1953)
- Pontryagin discovers the Hamiltonian formulation without Lagrange multipliers

> The first and the most important step toward the final solution was made by L.S. right after the formulation of the problem, during three days, or better to say, during three consecutive sleepless nights.

(Gamkrelidze (1999) "Discovery of the Maximum Principle")



Historical Discovery of the Maxmimum Principle

Boltyanski, Gamkrelidze and Pontryagin (1956):
"This fact appears in many cases as a general principle, which we call the *maximum principle*"

Этот факт является частным случаем следующего общего принципа, который мы называем **принципом максимума** (принцип этот доказан нами пока лишь в ряде частных случаев):

Пусть функция $H(x, \psi, u) = \psi_\alpha f^\alpha(x, u)$ *при любых фиксированных* x, ψ *имеет максимум по* u, *когда вектор* u *меняется в замкнутой области* $\overline{\Omega}$; *обозначим этот максимум через* $M(x, \psi)$. *Если* $2n$*-мерный вектор* (x, ψ) *является решением гамильтоновой системы*

$$\left.\begin{aligned} \dot{x}^i &= f^i(x, u) = \frac{\partial H}{\partial \psi_i}, \\ \dot{\psi}_i &= -\frac{\partial f^\alpha}{\partial x^i}\psi_\alpha = -\frac{\partial H}{\partial x^i}, \end{aligned}\right\} \quad i = 1, \ldots, n, \tag{8}$$

где кусочно-непрерывный вектор $u(t)$ *в каждый момент времени удовлетворяет условию* $H(x(t), \psi(t), u(t)) = M(x(t), \psi(t)) > 0$, *то* $u(t)$ *является оптимальным управлением, а* $x(t)$ *— соответствующей оптимальной (в малом) траекторией системы* (1).

Pontryagin (1959): " In the case that Ω is an open set [...], the variational problem formulated here turns out to be a special case of the problem of Lagrange."

Early Work with PDE Constraints

Egorov (1962,1963): Some problems in the theory of optimal control, Optimal control in Banach spaces

Minimum time control problem for the parabolic equation

$$\begin{aligned} \frac{\partial y}{\partial t} + Ay + b(u)y &= f + u && \text{on } \Omega \times (0, T) \\ y &= 0 && \text{on } \partial\Omega \times (0, T) \end{aligned}$$

with initial condition $y(t_0; u) = y_0$ and target $y(t_1; u) = y_T$.

Early Work with PDE Constraints

Egorov (1962,1963): Some problems in the theory of optimal control, Optimal control in Banach spaces

Minimum time control problem for the parabolic equation

$$\begin{aligned} \frac{\partial y}{\partial t} + Ay + b(u)y &= f + u && \text{on } \Omega \times (0, T) \\ y &= 0 && \text{on } \partial\Omega \times (0, T) \end{aligned}$$

with initial condition $y(t_0; u) = y_0$ and target $y(t_1; u) = y_T$.

J.-L. Lions: "Le travail de Yu. V. Egorov contient une étude détaillée de ce problème, mais nous n'avons pas pu comprendre tous les points des démonstrations de cet auteur, les résultats étant très probablement tous corrects."

Early Work with PDE Constraints

Egorov (1962,1963): Some problems in the theory of optimal control, Optimal control in Banach spaces

Minimum time control problem for the parabolic equation

$$\begin{aligned} \frac{\partial y}{\partial t} + Ay + b(u)y &= f + u \quad &\text{on } \Omega \times (0,T) \\ y &= 0 \quad &\text{on } \partial\Omega \times (0,T) \end{aligned}$$

with initial condition $y(t_0; u) = y_0$ and target $y(t_1; u) = y_T$.

J.-L. Lions: "Le travail de Yu. V. Egorov contient une étude détaillée de ce problème, mais nous n'avons pas pu comprendre tous les points des démonstrations de cet auteur, les résultats étant très probablement tous corrects."

Fattorini (1964): Time-optimal control of solutions of operational differential equations (proof of the "bang-bang" property, no maximum principle)

Friedman (1967): Optimal control for parabolic equations

Jacques-Louis Lions

R. M. Temam (SIAM News, July 10, 2001):

> *"A new adventure began for Lions in the early 1960s, when he met (in spirit) another of his intellectual mentors, John von Neumann. By then, using computers built from his early designs, von Neumann was developing numerical methods for the solution of PDEs from fluid mechanics and meteorology. At a time when the French mathematical school was almost exclusively engaged in the development of the Bourbaki program, Lions — virtually alone in France — dreamed of an important future for mathematics in these new directions; he threw himself into this new work, while still continuing to produce high-level theoretical work on PDEs."*

Adjoint Without Lagrange Multipliers

Lions (1968): Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles

$$J(u) = \|Cy(u) - z_d\|_H^2 + (Nu, u)_U, \quad N \text{ self-adjoint}, \geq 0$$

Target $z_d \in H$, state variable $y = y(u) \in V$, $u \in U_{ad}$, a closed convex subset U, and PDE constraint

$$Ay = f + Bu, \qquad A : V \to V'$$

$J(v) - J(u) \geq 0$ for all $v \in U_{ad}$ implies after a short calculation

$$(Cy(u) - z_d, C(y(v) - y(u)))_H + (Nu, v - u)_U \geq 0$$

which is equivalent to

$$(C^*\Lambda(Cy(u) - z_d), y(v) - y(u))_V + (Nu, v - u)_U \geq 0$$

$\Lambda : H \to H'$ canonical isomorphism from H to its dual H'

A Clever Guess

Defining $p(v) \in V$ by $A^*p(v) = C^*\Lambda(Cy(v) - z_d)$, we get

$$
\begin{aligned}
(C^*\Lambda(Cy(u) - z_d), y(v) - y(u))_V + (Nu, v - u)_U \\
&= (A^*p(u), y(v) - y(u))_V + (Nu, v - u)_U \\
&= (p(u), A(y(v) - y(u))_V + (Nu, v - u)_U \\
&= (p(u), B(v - u))_V + (Nu, v - u)_U \\
&= (\Lambda_U^{-1}B^*p(u) + Nu, v - u)_U \geq 0
\end{aligned}
$$

where $B^* : V \to U'$ is the adjoint of B, $\Lambda_U : U \to U'$ is the canonical isomorphism from U to U'.

Hence the adjoint $p(v)$ permits elimination of $y(u)$, and

$$
(\Lambda_U^{-1}B^*p(u) + Nu, u)_U = \inf_{v \in U_{ad}} (\Lambda_U^{-1}B^*p(u) + Nu, v)_U
$$

Lions (1968): "La formulation peut être considérée comme un analogue du principe du maximum de Pontryagin"

Conclusions

We have seen four main ideas:

1. Mechanical systems in equilibrium can easily be analyzed using "virtual velocities"
2. The Lagrange multiplier is just a multiplier from Gaussian elimination
3. Using Lagrange multipliers, one can find the adjoint equation in optimal control
4. One can also find the maximum principle of Pontryagin, noticing that the optimality system is Hamiltonian

"Constrained Optimization: from Lagrangian Mechanics to Optimal Control and PDE Constraints", Gander, Kwok, Wanner, 2014.