Chemin le plus court entre deux points

Philippe BAUCOUR

16 mai 2017

1 Introduction

"Dans un plan, quelle est la distance la plus courte entre deux points ?". La question, posée telle quelle, peut paraître un peu "bizarre" dans la mesure où "tout le mode sait bien" que "la distance la plus courte entre deux points est la ligne droite". OK, mais est-ce que tout le monde est capable de le prouver ? Ça, c'est une question vraiment plus intéressante qui va nous permettre d'aborder des sujets profonds et intéressants.

Je suppose que le niveau du lecteur se situe aux alentours de la terminale ou qu'il se rappelle, au moins, à quoi correspondent une intégrale et une dérivée. En effet, la méthode que l'on va utiliser se rapproche beaucoup de celle que l'on met en œuvre lorsque l'on cherche un extremum sur une courbe $y = f(x)$.

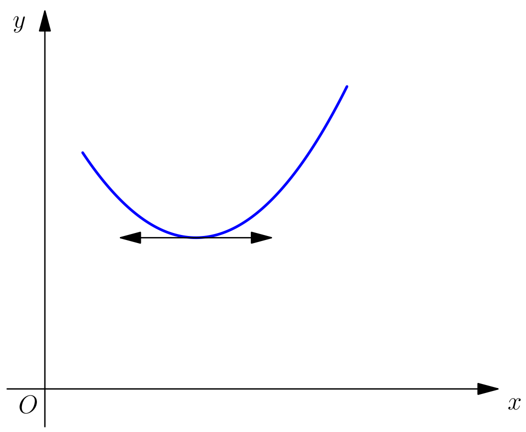

Figure 1 – Dérivée nulle à un extremum (ici un minimum)

Autrement dit on va chercher à annuler une dérivée. On parlera aussi de dérivée partielle mais, pas de panique si vous ne savez pas de quoi il s'agit, je prendrai le temps d'expliquer les choses le moment venu.

2 La méthode

Elle va consister à exprimer le chemin d'un point A à un point B comme une fonction. Quand ce sera fait, nous déterminerons la dérivée de la fonction puis nous annulerons cette dernière. Il n'y a rien de complexe ou de novateur dans cette approche. C'est typiquement celle utilisée pour trouver les minimas et maximas dans les études de fonctions classiques (voir la figure 1).

Commençons par dessiner la situation. Dans la figure 2, entre les points A et B il y a trois exemples de chemins possibles.

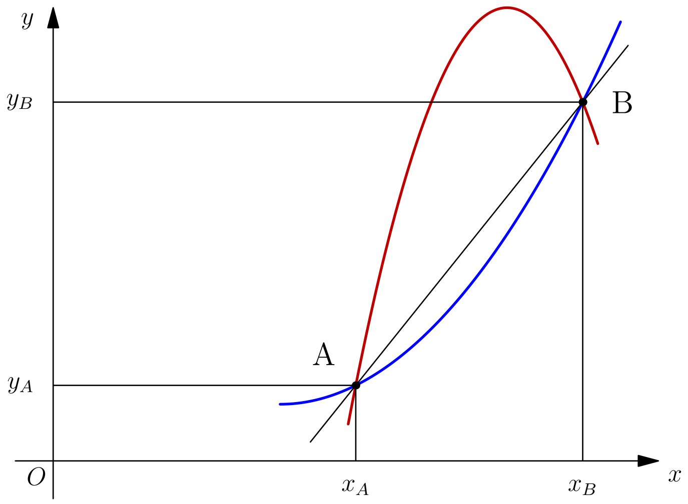

Figure 2 – Différents chemins entre deux points

On peut exprimer la longueur d'un chemin infinitésimal ds par

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2$$

Il n'y a là rien de très compliqué. En effet, si on zoome sur le chemin rouge de la figure 2, la formule précédente n'est rien d'autre que le théorème de Pythagore au pays des Lilliputiens (voir la figure 3).

Ensuite on peut écrire :

$$\begin{aligned}
\mathrm{d}s &= \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} \\
&= \sqrt{\mathrm{d}x^2 (1 + (\frac{\mathrm{d}y}{\mathrm{d}x})^2)} \\
&= \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\, \mathrm{d}x
\end{aligned}$$

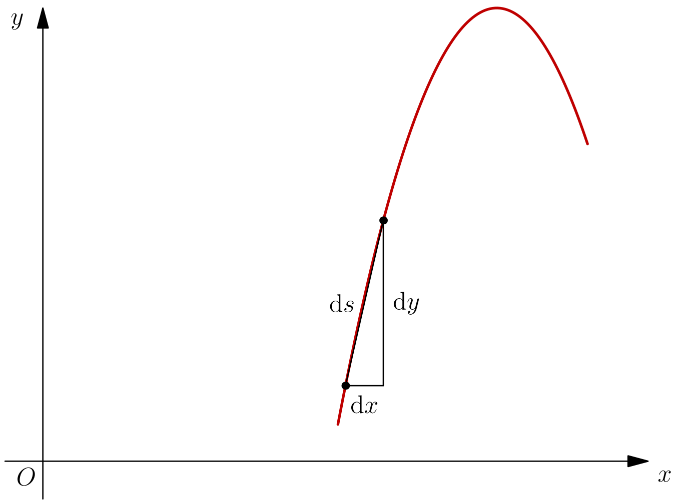

FIGURE 3 – $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2$

Où $\frac{\mathrm{d}y}{\mathrm{d}x}$ est bien la dérivée de y par rapport à x. On peut alléger l'écriture précédente en utilisant la notation "point"

$$\dot{y} = \frac{\mathrm{d}y}{\mathrm{d}x}$$

$$\mathrm{d}s = \sqrt{1 + \dot{y}^2}\,\mathrm{d}x$$

Enfin, on exprimera un chemin quelconque entre les points A et B comme une somme intégrale de chemins élémentaires. On écrit alors :

$$\begin{aligned} S &= \int_{x_A}^{x_B} \mathrm{d}s \\ &= \int_{x_A}^{x_B} \sqrt{1 + \dot{y}^2}\,\mathrm{d}x \end{aligned}$$

3 LA ruse de sioux

En fait c'est pas mal comme début mais on se trouve un peu coincé. En effet, ce que l'on a c'est l'expression d'un chemin. Ce dont on a besoin, c'est l'expression, la formule de tous les chemins possibles. Pour ce faire, on écrit :

$$Chemin_{err}(x) = Chemin_{id}(x) + Erreur(x)$$

En français dans le texte, cela signifie qu'un chemin erroné quelconque entre A et B (membre de gauche de l'égalité) se décompose en la somme du chemin idéal qui va de A à B et d'une erreur de parcours.

Dans la figure 4, S_{err} en rouge c'est le chemin erroné à l'étude. S_{id} en noir, c'est le chemin idéal et E en bleu, c'est l'erreur tout au long du parcours. Il n'est pas trop difficile de voir que "rouge égal noir plus bleu".

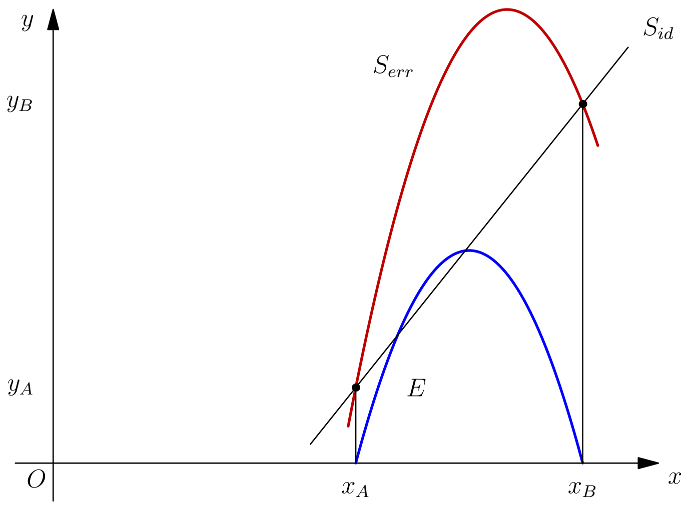

FIGURE 4 – Exemples de chemins entre deux points

Plus succinctement on écrit :

$$S_{err}(x) = S_{id}(x) + E(x)$$

Petite remarque qui a son importance : il faut imposer que $E(x_A) = 0$ et $E(x_B) = 0$ car sinon le chemin "idéal" n'a aucune chance de croiser le chemin erroné S_{err} aux deux extrémités. C'est bien ce que l'on voit sur l'erreur de parcours bleue de la figure 4.

Du point de vue des ordonnées on écrit :

$$y_{err}(x) = y_{id}(x) + E(x)$$

Où $y_{err}(x)$ est l'ordonnée sur le chemin erroné et $y_{id}(x)$ est l'ordonnée sur le chemin idéal.

Quoiqu'il en soit, écrit comme ça, on voit qu'on est toujours dans l'impasse. En effet, à x fixé, tout est fixé : $y_{id}(x)$ et $E(x)$ ont des valeurs bien déterminées. On n'a pas beaucoup de latitude pour discerner un chemin erroné d'un autre. Maintenant, si on se demande ce que l'on pourrait bien faire varier pour faire la différence entre deux chemins, il ne reste que la partie de l'égalité qui tourne autour de $E(x)$. En effet, par définition $y_{id}(x)$ est un chemin idéal, sanctuarisé, auquel on ne veut surtout pas toucher. Essayons une autre écriture de l'erreur et posons $E = \alpha \cdot \varepsilon(x)$. L'ordonnée devient :

$$y_{err}(x) = y_{id}(x) + \alpha \cdot \varepsilon(x)$$

Avant d'aller plus loin, remarquons que pour les mêmes raisons que précédemment il faut que l'on impose que $\varepsilon(x_A) = 0$ et $\varepsilon(x_B) = 0$.

Pour le reste il faut se persuader que ce petit changement d'écriture va tout changer. En effet, à x fixé, on a toujours la même erreur de parcours mais on construit dorénavant cette dernière en multipliant une fonction d'erreur ε par un coefficient α. Ce qui est vraiment bien joué c'est que ce dernier va décoincer la situation, nous donner de l'air et nous servir de variable.

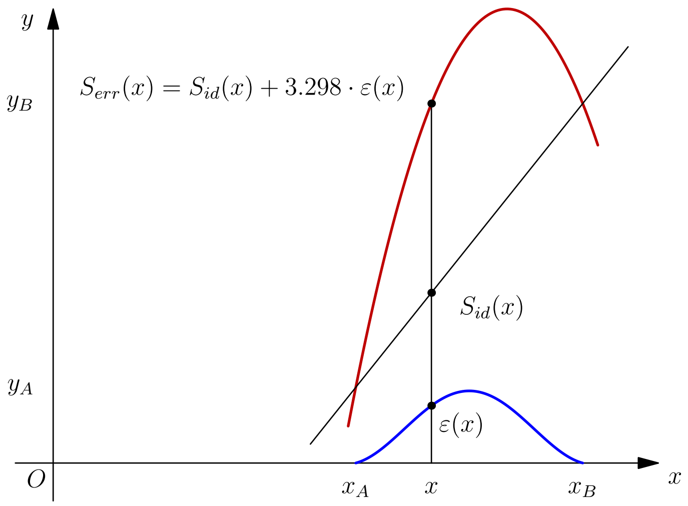

FIGURE 5 – À x fixé $\varepsilon(x)$ et $S_{id}(x)$ sont fixés. C'est α (3.298 ici) qui permet de différencier un chemin d'un autre

Quelle est maintenant la situation ? En fait, à x fixé, la valeur $\varepsilon(x)$ est, elle aussi, fixée et afin de pouvoir atteindre toutes les ordonnées possibles et imaginables on utilise une variable indépendante α afin de multiplier l'erreur $\varepsilon(x)$ à l'abscisse x.

En plus de la figure 5, prenons un exemple numérique pour bien comprendre que α est dorénavant une variable. Je me place au milieu de x_A et x_B. Supposons qu'à cette abscisse l'ordonnée du chemin idéal soit égale à 3. J'étudie le chemin qui, toujours à cette abscisse, passe par l'ordonnée $y = 10$. Si je connais la valeur de $\varepsilon(x)$ à l'abscisse considérée je suis capable de déterminer α en écrivant :

$$10 = \varepsilon(\frac{x_B - x_A}{2}) \cdot \alpha_1 + 3$$

Si maintenant j'étudie un autre chemin qui, pour la même abscisse, passe par l'ordonnée $y = 5$, comme l'erreur $\varepsilon(x)$ reste identique il faudra bien que α change de valeur. En effet, il faudra que :

$$5 = \varepsilon(\frac{x_B - x_A}{2}) \cdot \alpha_2 + 3$$

Dorénavant il faut donc "oublier" la variable x et se faire à l'idée qu'à x fixé et qu'il en va de même pour $\varepsilon(x)$ et $y(x)$. Autrement dit, à partir de maintenant, à x fixé, c'est la variable α qui permettra de différencier un chemin erroné d'un autre.

Autrement dit :

$$y_{err}(\alpha) = y_{id}(x) + \alpha \cdot \varepsilon(x)$$

Avant d'aller plus loin il est très très important de bien comprendre qu'à ε fixée, ce qui fait la différence entre deux chemins, c'est la valeur de α. Dans le raisonnement, ε est une constante et α est la variable que l'on va utiliser dans la dérivation afin de trouver le chemin le plus court. Si vous avez le moindre doute, lisez et relisez cette section avant de poursuivre.

4 En route vers la solution

On a écrit :

$$y_{err}(\alpha) = y_{id}(x) + \alpha \cdot \varepsilon(x)$$

En terme de chemin cela revient à dire que :

$$S_{err}(\alpha) = S_{id}(x) + \alpha \cdot \varepsilon(x)$$

On a donc maintenant une expression valable pour tous les chemins et qui dépend d'une variable α. Pour le coup, ça doit nous rappeler des souvenirs... Exemple : on a une fonction qui donne la hauteur du pendule en fonction du temps t et il faut déterminer à quel moment cette hauteur sera à son minimum (bien sûr c'est à ajustez selon vos souvenirs...). Il faut donc calculer la dérivée du chemin puis l'annuler. Précédemment nous avions écrit que dans le cas général nous avions :

$$S_{err} = \int_{x_A}^{x_B} \sqrt{1 + \dot{y}^2}\, \mathrm{d}x$$

Compte tenu de ce que nous avons dit à propos des ordonnées des points du chemin erroné nous devons écrire :

$$S_{err}(\alpha) = \int_{x_A}^{x_B} \sqrt{1 + \dot{y_{err}}(\alpha)^2}\, \mathrm{d}x$$

Sous la racine on a bien la dérivée par rapport à x de l'ordonnée sur le chemin erroné (y_{err}) qui est élevée au carré.

Pour trouver le chemin $S_{err}(\alpha)$ le plus court, il faut dériver ce dernier par rapport à α

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} S_{err}(\alpha) = \frac{\mathrm{d}}{\mathrm{d}\alpha} \int_{x_A}^{x_B} \sqrt{1 + \dot{y_{err}}(\alpha)^2}\, \mathrm{d}x$$

Comme les bornes d'intégration sont finies il n'y a pas de problème, l'intégrale est dérivable et on peut, dans la joie et la bonne humeur, dire que dans notre cas la dérivée de l'intégrale c'est l'intégrale de la dérivée.

Remarque : La raison pour laquelle il faut faire attention dans ces histoires de dérivée d'intégrale est la suivante. Si on a une fonction f définie par une intégrale :

$$f(x) = \int_{x_A}^{x_B} g(t)\,\mathrm{d}t$$

Alors, pour peu que la fonction G soit une primitive de g on a

$$f(x) = G(x_B) - G(x_A)$$

Maintenant, afin d'évaluer la dérivée de f il faut évaluer l'expression suivante

$$f'(x) = (G(x_B) - G(x_A))'$$

Si x_B et x_A dépendent de x on va avoir une expression qui ressemble à la chose suivante

$$f'(x) = G'(x_B)x'_B - G'(x_A)x'_A$$

Si la dernière égalité vous pose problème prenez le temps de lire les sections 4.1 et 4.2.

On imagine bien que dans certains cas un peu litigieux, les dérivées x'_A et x'_B peuvent avoir tendance à partir en sucette et faire que la fonction f ne soit pas dérivable.

Bref, dans notre cas il est licite de faire passer le dérivation sous l'intégrale. On écrit

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}S_{err}(\alpha) = \int_{x_A}^{x_B} \frac{\mathrm{d}}{\mathrm{d}\alpha}\sqrt{1 + \dot{y_{err}}(\alpha)^2}\,\mathrm{d}x$$

Ce que l'on peut encore écrire sous la forme

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}S_{err}(\alpha) = \int_{x_A}^{x_B} \frac{\mathrm{d}}{\mathrm{d}\alpha}f(y, \dot{y})\,\mathrm{d}x$$

Où $f(y, \dot{y})$ est manifestement une fonction à deux variables. Pas la peine de se laisser impressionner par une fonction qui met en œuvre une variable et sa dérivée. Si on avait $f(x, y)$ ce serait exactement la même chose.

Il faut donc être capable de dériver par rapport à α la fonction f à deux variables. Pour rappel

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}f(x,y) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \alpha} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha}$$

4.1 Rappel à propos des dérivées des fonctions à deux variables

Si la dernière équation ne vous pose pas de problème passez directement à la section 4.3.

Je suis en haut d'une montagne, sur la ligne de crête qui est parallèle à l'axe des x et mes spatules pointent vers l'axe z (voir figure 6). À ma gauche une piste noire pentue à souhait et à ma droite c'est une piste verte qui me tend les bras. Enfin, devant et derrière moi, la ligne de crête est quasiment horizontale.

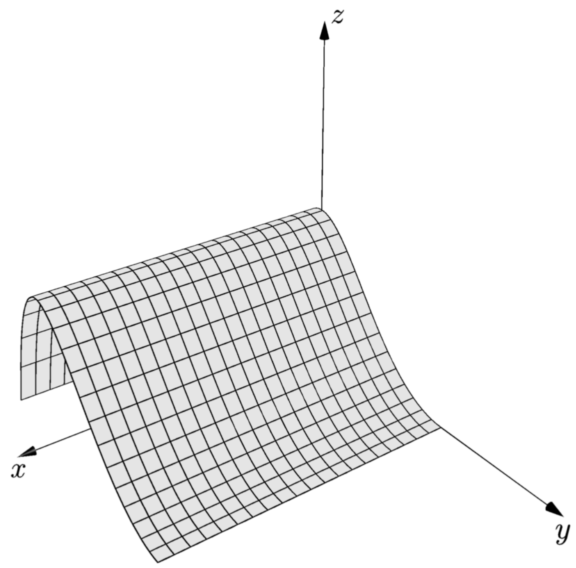

FIGURE 6 – Ligne de crête plate et deux pentes différentes de part et d'autre

Dans ce cadre, la montagne est une courbe en trois dimensions. À chaque position (x, y) correspond une altitude bien précise.

Question : Voir figure 6. Je me déplace sur la ligne de crête. Compte tenu de ce que j'ai écris précédemment qu'elle est alors la pente ? J'ai dit que la ligne de crête était quasiment horizontale. Donc la pente, selon l'axe des x est nulle. La dérivée par rapport à x est donc nulle. La dérivée par rapport à x de la fonction $HauteurMontagne()$ est donc nulle.

Question : Voir figure 7. Je suis passé sur la piste noire. Mes skis sont parallèles à la ligne de crête. Si je pousse sur mes bâtons je me déplace parallèlement à cette dernière. Selon cet axe, quelle est la pente ? En fait c'est exactement la même chose. Je suis à flan de montagne mais comme mes skis sont parallèles à la ligne de crête si je pousse j'observe que je ne monte ni ne descend. En résumé, selon l'axe des x, même quand j'ai quitté la ligne de crête, la pente est nulle. La dérivée par rapport à x de la fonction $HauteurMontagne()$ est donc encore nulle.

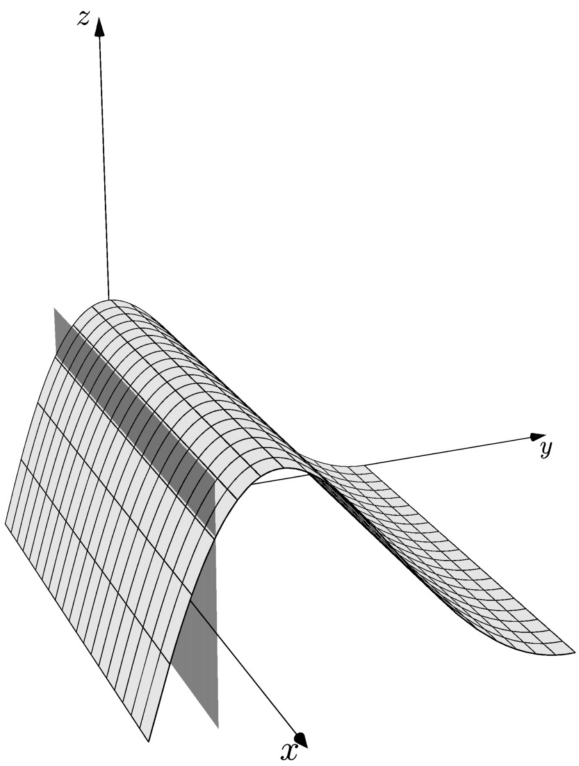

FIGURE 7 – Skis parallèles à la ligne de crête dans la noire

Question : Voir figure 8. Je suis du côté de la piste verte. Mes skis sont dans le sens de la descente (autrement dit, seul y varie). Quelle est alors la pente ? Si je regarde la situation de côté, je suis sur une pente négative car lorsque j'avance ma hauteur diminue. Autrement dit, quand je suis dans le sens de la pente, c'est à dire à x fixé, la dérivée par rapport à y de la fonction $HauteurMontagne()$ est donc négative.

Question : Je suis passé du côté sombre et je me retrouve face à la piste noire. Quelle est alors la pente ? Même constat. La seule chose qui change c'est que si sur la piste verte, pour un mètre d'avancement selon l'axe des y je perdais 10 cm de hauteur, là, sur la piste noire, quand j'avance de 1 m, je perds 50 cm d'altitude (la piste est donc 5 fois plus pentue). Autrement dit, sur la piste noire, quand je suis dans le sens de la pente, quand x est

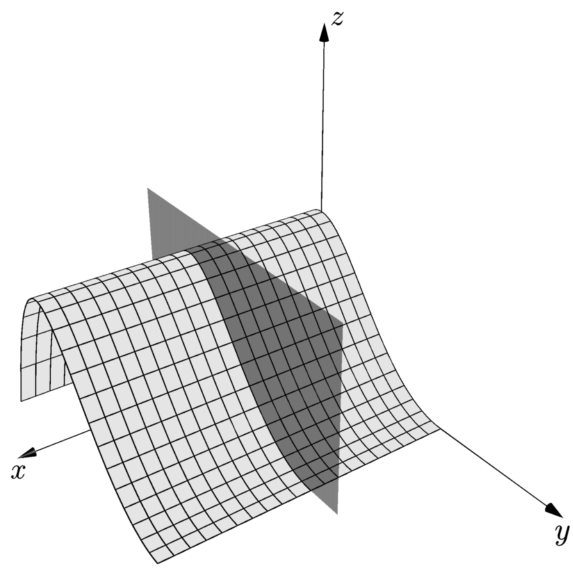

FIGURE 8 – Skis dans le sens de la pente sur la verte

fixé, la dérivée par rapport à y de la fonction $HauteurMontagne()$ est donc négative.

On a vu deux cas extrêmes où on se déplace selon uniquement l'axe des x ou selon uniquement l'axe des y. Bien sûr, entre les deux il existe toute une palette de possibilités. Il suffit de penser au cas où sur la piste noire, je vais à 45° par rapport à l'axe de la piste. Voir figure 9. On sent bien que dans ce cas, la pente est entre le 0° de dénivelé de la ligne de crête et le degré maximum de pente qui est atteint quand on est dans l'axe de la piste.

Une façon d'exprimer cet état de fait est d'écrire :

$$\mathrm{d}f = \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y$$

Dans notre contexte et en bon français cela signifie : la différence totale d'altitude que peut mesurer le skieur est égale à la somme des variations partielles d'altitude. La première variation partielle est celle qui se fait à y constant (et c'est x qui varie). La seconde variation partielle se fait à x constant (et c'est y qui varie).

Dans la formule $\mathrm{d}x$ et $\mathrm{d}y$ sont des petits accroissements selon les axes x et y.

Enfin $\frac{\partial f}{\partial x}$ est la dérivée partielle de la fonction f par rapport à x (on fixe y et on calcule la dérivée de la fonction f classiquement car cette dernière n'a, du coup, plus qu'une seule variable). Prenons un exemple numérique pour fixer les idées. Soit $f(x,y) = x^2 + 3xy + y^5$. Alors on a :

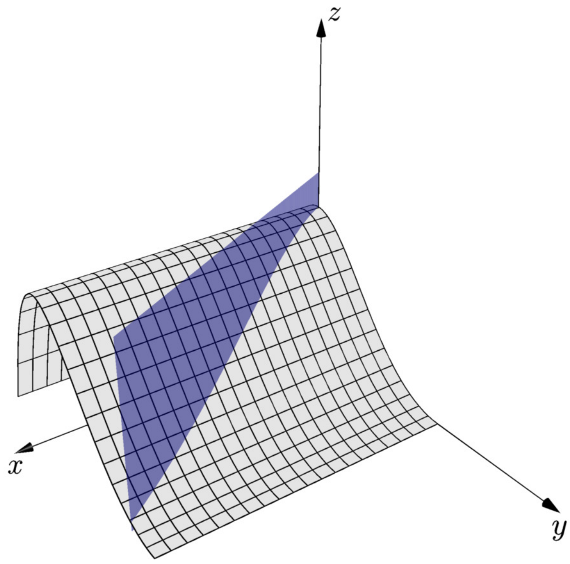

FIGURE 9 – Traversée de la verte en diagonale

$$\frac{\partial f}{\partial x} = 2x + 3y$$

$$\frac{\partial f}{\partial y} = 3x + 5y^4$$

En cas de soucis, pour calculer $\frac{\partial f}{\partial x}$ on peut remplacer y par la lettre k et considérer qu'il s'agit d'une constante. Idem pour l'autre dérivée partielle. Par contre, là c'est x qu'il faut considérer comme constante.

4.2 Dérivée d'une fonction composée

En anglais dans le texte "chain rule". On va commencer par un exemple très simple. Soit la fonction $g(x) = sin(5x)$ dont on cherche la dérivée. Le meilleur moyen de ne pas se planter est de procéder par étape.

J'ai normalement appris que

$$sin(f(x))' = cos(f(x)) \cdot f'(x)$$

Souvent on voit cette règle avec des u. Un truc du style

$$(sin(u))' = cos(u) \cdot u'$$

Franchement c'est peut être plus lourd mais je préfère faire apparaitre la fonction $f(x)$ et sa dérivée. C'est beaucoup plus explicite. On voit vraiment de quoi il s'agit.

A partir de là, le calcul de la dérivée de $g(x)$ ne pose plus trop de problème.

$$\begin{aligned} sin'(5x) &= cos(f(x)) \cdot f'(x) \\ &= cos(5x) \cdot 5 \end{aligned}$$

Ceci étant acquis tout se passe comme si nous avions effectué le calcul suivant

$$\frac{\mathrm{d}}{\mathrm{d}x} g(x) = \frac{\partial g}{\partial f} \frac{\partial f}{\partial x}$$

Attention, si vous êtes un "ayatollah" des notations mathématiques vous risquez de faire une syncope en lisant les lignes qui suivent. Quoiqu'il en soit, avec $f(x) = 5x$ on peut écrire :

$$\begin{aligned} \frac{\partial g}{\partial f} &= \frac{\partial (sin(f))}{\partial f} \\ &= d(sin(f)) \\ &= cos(f) \\ &= cos(5x) \end{aligned}$$

L'idée c'est de dire que la dérivée partielle de g par rapport à f, c'est, dans notre contexte, la dérivée partielle de $sin(f)$ par rapport à f. Tout se passe comme si f était une variable et plus une fonction de x.

De même on peut écrire

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial (5x)}{\partial x} \\ &= d(5x) \\ &= 5 \end{aligned}$$

Là c'est beaucoup plus classique et il n'y a pas de surprise. Au final on retrouve bien

$$sin'(5x) = cos(5x) \cdot 5$$

Bref, on notera que la dérivée d'une fonction composée c'est la multiplication des dérivées partielles et que dans le cas où il n'y a qu'une variable on peut écrire

$$\frac{\mathrm{d}}{\mathrm{d}x} g(x) = \frac{\partial g}{\partial f} \frac{\partial f}{\partial x}$$

Et que dans le cas où il y a plusieurs variables on écrira

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} f(x, y) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \alpha} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha}$$

4.3 Reprenons le cours de notre réflexion

Nous devions calculer la dérivée suivante

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} S_{err}(\alpha) = \int_{x_A}^{x_B} \frac{\mathrm{d}}{\mathrm{d}\alpha} \sqrt{1 + \dot{y_{err}}(\alpha)^2}\,\mathrm{d}x$$

Qui peut s'écrire sous la forme

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} S_{err}(\alpha) = \int_{x_A}^{x_B} (\frac{\mathrm{d}}{\mathrm{d}\alpha} f(y_{err}, \dot{y_{err}}))\,\mathrm{d}x$$

En appliquant les règles de dérivation il vient

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} S_{err}(\alpha) = \int_{x_A}^{x_B} (\frac{\partial f}{\partial y_{err}}\frac{\partial y_{err}}{\partial \alpha} + \frac{\partial f}{\partial \dot{y_{err}}}\frac{\partial \dot{y_{err}}}{\partial \alpha})\,\mathrm{d}x$$

Comme on a

$$y_{err}(\alpha) = y_{id}(x) + \alpha \cdot \varepsilon(x)$$

On en déduit que

$$y'_{err}(\alpha) = y'_{id}(x) + \alpha \cdot \varepsilon'(x)$$

À partir de là on trouve que

$$\frac{\partial y_{err}}{\partial \alpha} = \varepsilon(x)$$

Et que de la même façon

$$\frac{\partial \dot{y_{err}}}{\partial \alpha} = \varepsilon'(x)$$

La dérivation par rapport à α peut donc s'écrire

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} S_{err}(\alpha) = \int_{x_A}^{x_B} (\frac{\partial f}{\partial y_{err}}\varepsilon(x) + \frac{\partial f}{\partial \dot{y_{err}}}\varepsilon'(x))\,\mathrm{d}x$$

Ensuite on sépare les combattants dans deux intégrales distinctes

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} S_{err}(\alpha) = \int_{x_A}^{x_B} \frac{\partial f}{\partial y_{err}}\varepsilon(x)\,\mathrm{d}x + \int_{x_A}^{x_B} \frac{\partial f}{\partial \dot{y_{err}}}\varepsilon'(x)\,\mathrm{d}x$$

Bon, là on va peut être prendre le temps de réfléchir deux minutes. En fait dans la seconde intégrale, ce qui est embêtant c'est ce $\varepsilon'(x)$. Faudrait arriver à le supprimer. Pour cela on peut se dire que $\frac{\partial f}{\partial \dot{y_{err}}}\varepsilon'(x)$ ça ressemble à fg' et essayer de simplifier la seconde intégrale en faisant une intégration par partie qui aurait le mérite de ramener $\varepsilon(x)$ au bercail.

Remarque : En cas de doute à propos de l'intégration par partie, on peut se rappeler que

$$(uv)' = u'v + uv'$$

Alors on peut écrire

$$(uv)' - u'v = uv'$$

D'où l'on tire

$$\begin{aligned}\int uv' &= \int ((uv)' - u'v) \\ &= \int (uv)' - \int (u'v) \\ &= uv - \int u'v\end{aligned}$$

Si on pose $v' = \varepsilon'(x)$ alors $v = \varepsilon(x)$. Ensuite, si on pose $u = \frac{\partial f}{\partial \dot{y_{err}}}$ alors $u' = \frac{d}{dx}\frac{\partial f}{\partial \dot{y_{err}}}$ (l'intégration par partie se fait par rapport à x) et la seconde intégrale s'écrit de la façon suivante :

$$\int_{x_A}^{x_B} \frac{\partial f}{\partial \dot{y_{err}}} \varepsilon'(x)\,\mathrm{d}x = [\frac{\partial f}{\partial \dot{y_{err}}}\varepsilon(x)]_{x1}^{x2} - \int_{x_A}^{x_B} \frac{d}{dx}\frac{\partial f}{\partial \dot{y_{err}}}\varepsilon(x)\,\mathrm{d}x$$

Et là, c'est que du bonheur. En effet, si on se rappelle la petite remarque qui disait que $\varepsilon(x_A) = 0$ et $\varepsilon(x_B) = 0$ alors la première partie du membre de gauche est nulle et il ne reste plus que :

$$\int_{x_A}^{x_B} \frac{\partial f}{\partial \dot{y_{err}}} \varepsilon'(x)\,\mathrm{d}x = -\int_{x_A}^{x_B} \frac{d}{dx}\frac{\partial f}{\partial \dot{y_{err}}}\varepsilon(x)\,\mathrm{d}x$$

On peut donc ré-écrire la dérivée des chemins erronés sous la forme :

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} S_{err}(\alpha) = \int_{x_A}^{x_B} \frac{\partial f}{\partial y_{err}} \varepsilon(x)\,\mathrm{d}x - \int_{x_A}^{x_B} \frac{d}{dx}\frac{\partial f}{\partial \dot{y_{err}}}\varepsilon(x)\,\mathrm{d}x$$

Si on regroupe les deux intégrales en une seule et si on met $\varepsilon(x)$ en facteur, il vient

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}S_{err}(\alpha) = \int_{x_A}^{x_B} (\frac{\partial f}{\partial y_{err}} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y_{err}}})\varepsilon(x)\,\mathrm{d}x$$

Réfléchissons deux minutes... Si on veut que la dérivée de gauche soit nulle pour toutes les erreurs de parcours $\varepsilon(x)$ il faut que la parenthèse sous l'intégrale soit nulle.

Nous devons donc résoudre :

$$\frac{\partial f}{\partial y_{err}} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y_{err}}} = 0$$

Là aussi, prenons le temps de réfléchir deux minutes. On s'est placé dans le cas où la dérivée du chemin est nulle ($\frac{\mathrm{d}}{\mathrm{d}\alpha}S_{err}(\alpha) = 0$). Cela veut donc dire que l'on s'est placé sur le bon chemin, le chemin idéal. Bref, c'est sûr on est plus sur un des chemins erronés et donc on ne joue plus avec y_{err} mais plutôt avec y_{id}. Tout de suite après avoir fait cette remarque, comme on est certains du y dont on parle, on remplace y_{id} par y.

Autrement dit, l'équation précédente s'écrit donc sous la forme :

$$\boxed{\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}} = 0}$$

5 La dernière ligne droite

Nous devons résoudre

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}} = 0$$

Maintenant, on peut redonner à f sa forme explicite pour faire les calculs

$$f = \sqrt{1 + \dot{y}(\alpha)^2}$$

Comme f n'est pas une fonction de y sa dérivée partielle par rapport à y est nulle. On a donc :

$$\frac{\partial f}{\partial y} = 0$$

La dérivée de f par rapport à $\dot{y}$ n'est pas très difficile à trouver

$$\frac{\partial f}{\partial \dot{y}} = \frac{1}{2}(1 + \dot{y}^2)^{-\frac{1}{2}}2\dot{y}$$

Soit encore

$$\frac{\partial f}{\partial \dot{y}} = \frac{\dot{y}}{\sqrt{1 + \dot{y}^2}}$$

L'équation

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}} = 0$$

Devient donc

$$-\frac{d}{dx}\frac{\partial f}{\partial \dot{y}} = 0$$

Enfin si on remplace avec l'expression de la dérivée partielle que l'on vient de calculer, il vient :

$$-\frac{d}{dx}(\frac{\dot{y}}{\sqrt{1+\dot{y}^2}}) = 0$$

On peut se permettre d'oublier le signe négatif (il suffit de tout multiplier par -1).

Maintenant il faut prendre le temps de remarquer que si on cherche un "truc" dont la dérivée par rapport à x est nulle cela veut tout simplement dire que ce "truc" est constant. On a donc :

$$\frac{\dot{y}}{\sqrt{1+\dot{y}^2}} = a$$

Ensuite on fait passer la racine carré à droite et il vient :

$$\dot{y} = a\sqrt{1+\dot{y}^2}$$

On élève les 2 côtés au carré. Comme a est une constante alors a^2 est aussi une constante. Je ne perds pas de temps et je laisse a en lieu et place de a^2. Il vient :

$$\dot{y}^2 = a(1+\dot{y}^2)$$

$$\dot{y}^2(1-a) = a$$

On fait passer $(1-a)$ de l'autre côté. Comme le membre de droite $(\frac{a}{1-a})$ est une constante je laisse la constante a de l'autre côté de l'égalité.

$$\dot{y}^2 = a$$

On prend la racine. La racine d'une constante est une constante donc je laisse toujours a à droite.

$$\dot{y} = a$$

Et là, il n'y a pas 36 solutions, il n'y en a qu'une. Il faut que :

$$\boxed{y = ax + b}$$

Ouf, l'honneur est sauf et on peut dormir tranquille. On vient de démontrer que dans un plan, le chemin le plus court est bien une droite.

6 Conclusion

Ce qu'il y a de remarquable dans la méthode variationnelle et le principe de moindre action (car c'est bien d'eux dont il s'est agit ici) c'est qu'on part sans à priori. On est zen, on est prêt à accepter que pour aller de A à B un point puisse partir autour de la Lune puis revenir alors qu'un autre puisse tirer tout droit. On a pas d'a priori et on fait confiance aux maths pour nous aider à trouver LE chemin le plus court.

En plus, si au départ on parle de distance infinitésimale, bien vite on oublie le problème particulier de la distance entre deux points et on se retrouve à résoudre un problème bien plus général qui se termine à la fin de la quatrième partie de ce billet lorsqu'on retrouve l'équation de Euler-Lagrange.

En d'autres mots, le principe de moindre action est universel. Par exemple on retrouve assez facilement la trajectoire d'un boulet de canon soumis à la gravité terrestre. Là aussi, on part sans à priori, on ne parle pas des forces en présence etc. Le plus beau dans l'histoire c'est qu'on part de la question "mais pourquoi donc une balle a-t-elle cette trajectoire quand je la lance". À la fin on retrouve bien la trajectoire et on en profite pour retrouver aussi la seconde loi de Newton ($\vec{F} = m \cdot \vec{a}$).

Donc avec le principe de moindre action on n'utilise pas des formules apprises (seconde loi de Newton par exemple) pour retrouver la trajectoire de tel ou tel mobile. On fait le contraire. On généralise le problème et ensuite on filtre.

Oh, by the way, en règle générale, lorsqu'on utilise le principe de moindre action pour résoudre un problème de dynamique un peu complexe (pensez à un double pendule par exemple) ce dernier nous isole des problèmes de coordonnées (polaires, sphériques etc.)car on utilise des coordonnées généralisées ce qui simplifie généralement la démarche et les calculs. Là je sens bien qu'il faudrait que j'écrive un second billet pour illustrer ce point particulier...

Pour information, notons que le principe de moindre action permet de retrouver la formule $n_1 sin(i_1) = n_2 sin(i_2)$.

Last but not least, c'est toujours le principe de moindre action que Richard Feynman utilisa dans sa thèse de 1942 et qui lui permit de retrouver, entre autres, la formule de Schrödinger.

Enfin, on peut aussi remarquer qu'au final, on a pas essayé de minimiser tel ou tel paramètre (ici une distance) au contraire, on généralise le problème en essayant de trouver une formule qui décrit tous les paramètres possibles (ici tous les chemins possibles) et ensuite on cherche le plus petit d'entre eux en annulant la dérivée de la formule en question.

Bref, je trouve ça tout bonnement génial.

7 À retenir

— On écrit $S = \int_{x_A}^{x_B} \sqrt{1+\dot{y}^2}\,\mathrm{d}x$

— Un chemin erroné c'est la somme du chemin idéal et d'une erreur de parcours

— $y_{err}(x) = y(x) + \alpha \cdot \varepsilon(x)$ où α est LA variable qui différencie un chemin d'un autre

— On ne fait pas l'étude d'un chemin qu'on tente de rendre minimal mais on trouve la fonction qui décrit tous les chemins et on cherche alors le plus petit d'entre eux.

— L'expression du chemin en fonction de α est $S_{err}(\alpha) = \int_{x_A}^{x_B} \sqrt{1+\dot{y_{err}}(\alpha)^2}\,\mathrm{d}x$

— On dérive par rapport à α $\frac{\mathrm{d}}{\mathrm{d}\alpha} S_{err}(\alpha) = \frac{\mathrm{d}}{\mathrm{d}\alpha} \int_{x_A}^{x_B} \sqrt{1+\dot{y_{err}}(\alpha)^2}\,\mathrm{d}x$

— Pour que la dérivée s'annule il faut que l'équation de Euler-Lagrange $\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}} = 0$ soit respectée

— Ici $f = \sqrt{1+\dot{y}(\alpha)^2}$

— On arrive à $y = ax + b$

8 Aller plus loin

Sur Youtube et Google il faut chercher des termes tels que :

— principe de moindre action

— méthode variationnelle

— principle of least action

— Lagrangien

— Hamiltonien