

# DU CALCUL DIFFERENTIEL AU CALCUL VARIATIONNEL : UN APERCU DE L'EVOLUTION DE P. FERMAT A NOS JOURS<sup>1</sup>

**J.-B. HIRIART-URRUTY**  
**Département de Mathématiques**  
**Université Paul Sabatier de Toulouse**

La Science peut être considérée comme un problème de minimum ; selon Ernst MACH :  
*“Elle consiste à exposer des faits, avec la moindre dépense intellectuelle”.*

Quand il m'arrive de dispenser un enseignement sur le Calcul différentiel, je suggère parfois à mes étudiants de se rendre en pèlerinage à la Salle des Illustres au premier étage du Capitole (= mairie) de Toulouse, ils y trouveront une statue de Pierre FERMAT accompagnée de la légende suivante : “inventeur du calcul différentiel” ; s'ils ont des difficultés avec cette matière ils savent donc à qui se plaindre....Par ailleurs, le Guide du routard Midi-Pyrénées mentionne à propos de Beaumont-de-Lomagne (Tarn-et-Garonne) : “patrie de P. FERMAT, célèbre mathématicien à qui l'on doit le calcul différentiel (encore merci!)”...Calcul différentiel / Calcul intégral, voilà des termes bien étranges pour le profane, des mots qui font peur tant ils sont censés représenter la difficulté en mathématiques. Je me souviens qu'ANTOINE, un chanteur qui eut un certain succès du temps de mon adolescence (avant de se reconvertir dans la publicité pour des lunettes), évoquait dans une de ses chansons le Calcul intégral comme une bête noire que devaient craindre tous les jeunes s'engageant dans des études scientifiques...Lui-même avait une formation d'ingénieur centralien avant de commencer à chanter.

Quand j'ouvre un livre de mathématiques de Première S ou Terminale S (des lycées) et que je me reporte au chapitre sur le Calcul différentiel, je trouve un peu injuste que les commentaires historiques ne mentionnent que NEWTON et LEIBNIZ, alors que FERMAT fut assurément un précurseur...D'ailleurs la “règle de FERMAT” que nous discuterons dans cet exposé est mondialement connue sous cette appellation en Calcul différentiel et variationnel.

Notre propos ici est partagé entre deux objectifs :

- nous concentrer sur *l'un* des aspects du personnage FERMAT et de sa contribution scientifique à la naissance du Calcul différentiel ;
- parler, très succinctement, du *développement historique* (du 17ème au 20ème siècle) du Calcul différentiel, mais évoquer aussi les mathématiques actuelles dont des domaines qui, en quelque sorte, font suite à ce Calcul différentiel ; nous serons ainsi amenés à préciser les raisons pour lesquelles le calcul différentiel a dû évoluer.

---

<sup>1</sup>Version écrite d'un exposé à la Journée FERMAT organisée à l'Hôtel d'Assézat de Toulouse en octobre 2004, sous l'égide de l'Académie des Sciences, Inscriptions et Belles-Lettres de Toulouse.

## 1. Au début fut le 17ème siècle.

Le 17ème siècle est un siècle tendu : guerres (de religion), même si ceci n'est pas spécifique à ce siècle, épidémies de toute sorte, déclin du pastel à Toulouse dont la crise a débuté en 1561, etc., mais du point de vue scientifique c'est le siècle de **la mathématisation du mouvement** : vitesse, accélération, droite tangente à une trajectoire, autant de notions qui vont émerger progressivement...souvent à des fins belliqueuses d'ailleurs ; les calculs étaient motivés, entre autres, par les besoins en astronomie et surtout en balistique : en 1675 est publié par exemple "L'art de jeter des bombes". Dans ce siècle, FERMAT va se retrouver en bonne compagnie, au sein du trio DFP (DESCARTES, FERMAT, PASCAL), mais aussi avec HUYGENS, NEWTON, LEIBNIZ plus tard...Ce 17ème siècle se terminera tout de même, sous Louis XIV, par une grande oeuvre : le Canal du midi ; cela me fait penser à une affirmation de son concepteur P.-P. RIQUET, qui n'est pas sans lien avec l'utilisation pratique du calcul différentiel : "*Ce qui est le plus simple est souvent le meilleur*" ; à méditer...

Le point d'ancrage de l'apport de FERMAT est sa méthode présentée dans "De maximis et minimis" (texte communiqué par FERMAT à ses confrères au Parlement de Bordeaux en 1629, soit treize ans avant la naissance de NEWTON), méthode qui s'applique à la détermination des valeurs qui rendent maximale ou minimale une fonction ainsi que des tangentes aux courbes, ce qui revenait à poser les fondements du Calcul différentiel. La motivation de FERMAT comme de ses successeurs était les **problèmes d'extremum** : le Calcul différentiel est vraiment né de la nécessité de résoudre des problèmes d'extremum, d'**optimisation** dirions-nous de nos jours. C'est le moment de faire un peu de publicité pour un produit multimédia à la conception duquel nous avons participé il y a quelques années : "Fermat : le premier homme du monde" [15]. Daté comme un film (1995), la partie vidéo est un mélange d'histoire des sciences, d'histoire régionale, relaté dans une succession de saynètes...Nous avons utilisé ce document à plusieurs reprises en classes de Première S, Terminale S, et aussi pour illustrer des conférences grand public de popularisation scientifique. Les techniques de diffusion évoluant rapidement, la vidéo est maintenant remplacée par un DVD, le tout à un prix modique.

## 2. La dite règle de FERMAT.

**2.1** Nous sommes appelés à évoquer des *dérivées*...mot bizarre, rien à voir avec les produits dérivés (des financiers)...*Différentiation* de fonctions, encore un terme étrange, rien à voir avec la différenciation cellulaire....Des fonctions qui seraient *dérivables* ou pas...le qualificatif de dérivable n'aide pas d'ailleurs à retrouver ou retenir ce dont il s'agit ; si on pouvait refaire les appellations, je préférerais parler de fonctions *lisses* (d'ailleurs l'Anglais parle encore de fonctions *smooth*, l'Espagnol de fonctions *suave*)...On peut dire que FERMAT travaillait avec des fonctions dérivables (dans les classes de fonctions connues à l'époque, les fonctions polynomiales, trigonométriques) sans en connaître la définition exacte (les objets mathématiques sur lesquels il travaillait n'étaient pas définis en termes de *fonctions* (notion postérieure))...Cela me rappelle ce que répondait le physicien P. DIRAC quand on lui reprochait d'utiliser des concepts et outils mathématiques non définis

formellement : “*Je digère..., je n’ai pas besoin de savoir comment s’opère exactement la digestion*”. L’évolution du concept de “changement” (qui, *in fine*, allait donner naissance à la notion de dérivée) est excellemment analysée dans [3] : “*First the derivative was used, then discovered, explored and developed, and only then, defined*”; l’auteur y souligne que c’est exactement dans l’ordre inverse que nous l’enseignons aujourd’hui, mais ceci est une autre histoire...

Si l’on considère la méthode proposée par FERMAT, avec un tant soit peu de présentation moderne, voici ce que cela donne : *Pour trouver la valeur  $a$  où une fonction est maximum ou minimum : exprimer  $f(a)$ , substituer  $a+h$  à  $a$ , c’est-à-dire exprimer  $f(a+h)$ ; “adégaler”  $f(a)$  et  $f(a+h)$  (adégaler est un terme emprunté à DIOPHANTE pour rappeler que  $f(a+h)$  n’est que “presque égalé” à  $f(a)$ ); dans cette adégalité, ne garder que les termes du premier ordre en  $h$ ; diviser par  $h$ ; en égalant, on obtient maintenant une équation en  $a$  que l’on doit résoudre. Autrement dit, une manière d’utiliser la dérivée  $f'$  pour déterminer les extrema de  $f$  est de chercher les solutions de l’équation  $f'(a) = 0$ . Malgré les attaques dont il fut l’objet, FERMAT garda une pleine confiance dans sa méthode qui, disait-il, “ne trompe jamais et peut s’étendre à des questions très belles”. Il semble par ailleurs qu’il ait su discerner très tôt les divers types d’extrema en négligeant les termes du 3ème ordre dans son développement de  $f(a+h)$ ; “la raison principale” est, disait-il, que les termes du 2ème ordre sont “en plus grande raison” que ceux mesurés par les plus grandes puissances au delà du 2ème ordre. Dans la manière de faire et dans les explications de FERMAT apparaît une notion de “dérivée”, sous l’idée d’un accroissement d’une quantité qui est une fonction de la variation d’une autre quantité que l’on laisse “s’évanouir”. Le délicieux terme de “quantité évanescence”, cher aux mathématiciens analystes des 17ème et 18ème siècles, n’a plus usage en mathématiques depuis fort longtemps; l’Anglais a toutefois conservé une expression parente, “vanishing term”.*

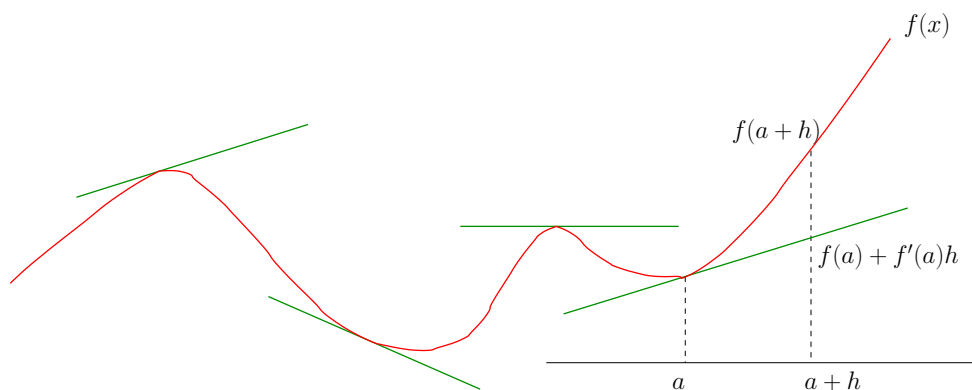


Figure 1

$$f(a+h) = f(a) + \boxed{?} h + h.\epsilon(h) \quad (\epsilon(h) \text{ étant une expression qui tend vers zéro avec } h)$$

$$\boxed{?} = f'(a).$$

Les diverses notations utilisées pour la dérivée de  $f$  en  $a$  sont :  $f'(a)$ ,  $\frac{df}{da}$ ,  $f'(a)$ , etc.

Plus tard, NEWTON dans son idée de *fluxion* dira : “*J’ai trouvé l’indication de cette méthode grâce au tracé par Fermat des tangentes*”.

Ce sont les contributions successives de P. FERMAT, I. NEWTON, G. W. LEIBNIZ (17ème siècle), L. EULER, J. L. LAGRANGE (18ème siècle), A. L. CAUCHY et K. WEIERSTRASS (19ème siècle) qui vont conduire à la définition formelle de la notion de dérivée et de fonction dérivée (appellations dues à LAGRANGE). Si on reprend de manière mathématiquement équivalente la définition donnée plus haut, on a :

**Définition.** *La fonction de la variable réelle  $f$  est dite dérivable en  $a$  lorsque le quotient (dit différentiel)  $[f(a+h) - f(a)] / h$  a une limite (finie) quand l’argument  $h$  tend vers 0.*

Pour les lecteurs étudiants ou professeurs de mathématiques, je signale la caractérisation suivante (qui peut faire l’objet d’un excellent exercice) :

**Théorème** ([2] par exemple). *Soit  $f$  une fonction (définie dans un voisinage de 0) continue en 0, et soit  $c$  un paramètre réel différent de 1 et -1. Alors la dérivabilité de  $f$  en 0 équivaut à l’existence d’une limite (finie) quand  $h$  tend vers 0 dans le quotient suivant*

$$\frac{f(h) - f(c.h)}{(1 - c).h}.$$

*La limite obtenue est la dérivée  $f'(a)$  de  $f$  en  $a$ .*

Le quotient différentiel ci-dessus n’est évidemment pas défini pour la valeur 1 du paramètre  $c$ . Ce qui est plus curieux est que seule la valeur  $-1$  de  $c$  est à écarter dans le théorème ; en effet le résultat annoncé ne subsiste pas avec cette valeur du paramètre : des quotients différentiels symétriques par rapport à l’origine 0 peuvent avoir des limites sans que la fonction  $f$  ait une dérivée en 0 ; l’exemple le plus simple en est la fonction  $f(x) = |x|$ .

**2.2 La règle de Fermat**, telle qu’elle est universellement appelée de nos jours, est donc la suivante :

**Théorème.** *Si une fonction  $f$  est minimisée ou maximisée en un point  $a$ , et si  $f$  est dérivable en  $a$ , alors nécessairement la dérivée  $f'(a)$  de  $f$  en  $a$  est nulle.*

Le Calcul différentiel est le corpus de règles qui permettent de savoir si une fonction est dérivable (on dit aussi différentiable) et de calculer sa dérivée sans revenir chaque fois à la définition : par exemple, la somme de deux fonctions dérivables est dérivable et la dérivée de la fonction somme est la somme des fonctions dérivées. Beaucoup d’opérations usuelles en Mathématiques préservent la dérivabilité, la “lissité” dirions-nous (somme, produit, quotient, exponentiation de fonctions), il en est une tout de même (pourtant importante dans les applications) qui la détruit : prendre le maximum d’un certain nombre (même fini) de fonctions (on en reparle au paragraphe 3.2).

### 2.3 Premières limitations.

Avec la règle de FERMAT on est confronté assez vite à deux limitations (au moins) : Quid lorsque la fonction considérée  $f$  n’est pas dérivable au point candidat  $a$  ? Qu’en est-il

des situations *contraintes*, c'est-à-dire lorsque le problème d'extremum considéré comporte des contraintes ?

Examinons le premier écueil (cf. Figure 2, à gauche). Si l'on s'en tient à la règle de FERMAT, tous les points de non-dérivabilité de  $f$  doivent être retenus *a priori* comme candidats à être extrema locaux (puisqu'ils vérifient par défaut la condition de la règle de FERMAT). Ceci n'est pas satisfaisant, ce qui explique que les mathématiciens se soient évertués au cours des années (y compris les trente dernières) à trouver pour les fonctions non-dérivables un substitut à la notion de dérivée, qui joue en Optimisation un rôle semblable ; nous en dirons quelques mots plus loin.

Deuxième écueil : Que doit-il se passer concernant la dérivée si on doit minimiser une fonction  $f$  sur un ensemble-contrainte tel que le segment  $[u, v]$  (cf. Figure 2, à droite) ? Il est manifestement faux de dire que si une fonction  $f$  est minimisée sur le segment  $[u, v]$  en  $u$ , alors la dérivée de  $f$  en  $u$  doit être nulle. Ceci est une faute grave (que j'ai pourtant vue écrite dans un livre de mathématiques de Terminale S) : "Si une fonction dérivable est minimisée sur un intervalle en un point de cet intervalle, alors la dérivée y est nulle"<sup>2</sup>, qui met en évidence le rôle primordial de la notion de **contrainte** dans un problème d'optimisation. Et d'ailleurs la résolution de problèmes d'extremum en classes de Première et Terminale S se réduit quasi-exclusivement à la détection des points candidats via la règle de FERMAT, c'est-à-dire à la recherche des points annulant la dérivée de la fonction à minimiser. Sans aller très loin dans ce que serait la règle de FERMAT dans un problème d'extremum avec contraintes, voici ce qu'on peut dire facilement :

*Soit  $f$  une fonction dérivable sur un intervalle réel  $I$ , minimisée en un point  $a$  de cet intervalle ; alors*

- si  $a$  est l'extrémité gauche de l'intervalle, la dérivée de  $f$  y est nécessairement positive ;
- si  $a$  est l'extrémité droite de l'intervalle, la dérivée y est nécessairement négative ;
- si  $a$  est à l'intérieur de l'intervalle (c'est-à-dire en bougeant un peu à gauche et à droite de  $a$  on reste encore dans  $I$ ), la dérivée y est nécessairement nulle.

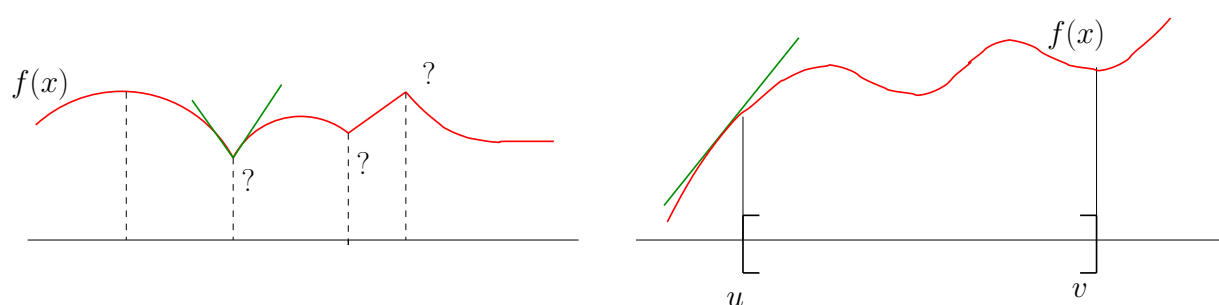


Figure 2

## 2.4 Développements.

Tout ce Calcul différentiel, équissé au-dessus pour des fonctions d'une seule variable, a été développé par les mathématiciens au cours des 18<sup>ème</sup> et 19<sup>ème</sup> siècles (ou plutôt par les

<sup>2</sup>J'accorde à l'auteur le bénéfice du doute : le résultat est vrai si l'on s'en tient à des intervalles *ouverts*.

mathématiciens-physiciens-mécaniciens car on ne faisait pas de distinction à l'époque...) dans plusieurs directions, notamment celles-ci :

- Considération de fonctions de plusieurs variables ( $n$  dimensions), et même “une infinité” (les variables ou inconnues du problème sont elles-mêmes des fonctions, des formes géométriques). La Physique et la Mécanique sont là pour suggérer des **formulations variationnelles** (en configuration “naturelle” ou d'équilibre, la Nature minimise quelque chose : une énergie, une puissance dissipée...). Rappelons ce que disait L. EULER à ce propos : *“Rien ne se passe dans le monde qui ne soit la signification d'un certain maximum ou d'un certain minimum”*.

La règle de FERMAT, généralisant ce qui a été vu au-dessus, est simple : si une fonction  $f$  de  $n$  variables est minimisée en  $a = (a_1, a_2, \dots, a_n)$ , alors le vecteur gradient (le vecteur des dérivées partielles) de  $f$  y est nul

$$\nabla f(\mathbf{a}) = \mathbf{0}.$$

- Prise en compte des contraintes : les candidats minimiseurs doivent satisfaire les contraintes du problème posé (par exemple, si on a à déterminer un débit d'eau minimisant tel ou tel critère, il n'en reste pas moins que la variable-débit doit être de signe positif). La règle de FERMAT admet des adaptations, différentes suivant la forme des contraintes ; elles portent le nom de EULER-LAGRANGE (18ème siècle, pour des contraintes exprimées sous forme d'égalités), de KARUSH-KUHN-TUCKER (deuxième moitié du 20ème siècle, pour des contraintes exprimées sous forme d'inégalités). Voir l'opuscule [8] pour des développements basiques à ce sujet.

Comme dans le cas unidimensionnel, l'optimisation avec contraintes a du mal à entrer dans l'esprit et le savoir-faire de nos étudiants. Je cite volontiers l'anecdote suivante. Le concours d'agrégation externe de mathématiques de 1992 comportait une épreuve (dite d'Analyse numérique), dont j'avais créé le sujet, qui traitait en bonne partie de problèmes d'optimisation avec quelques illustrations [6]. La troisième partie étudiait la minimisation d'une fonction-objectif simple (quadratique) de deux variables sur un ensemble-contrainte simple également, une boule fermée de rayon donné dans le plan. Lorsque la question de l'existence et de la caractérisation des solutions fut posée, une proportion significative de candidats proposa de résoudre l'équation  $\nabla f(x, y) = 0$ , ce qui conduisait à un point de coordonnées  $(x, y)$  situé en dehors de l'ensemble-contrainte ; d'où des conclusions erronées du style : le problème d'optimisation initial n'a pas de solution (alors qu'il s'agissait de minimiser une fonction continue sur une partie compacte du plan!) ; ou bien, la solution se trouve à l'extérieur de l'ensemble-contrainte ! Cette constatation motiva notre note [7] à destination des collègues formateurs.

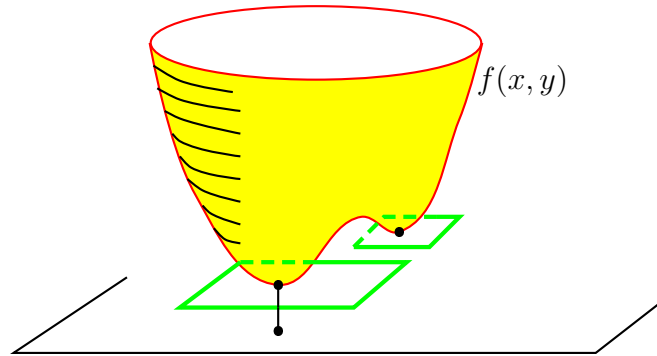


Figure 3

Avant d'aller plus loin mentionnons un autre problème d'extremum auquel est lié le nom de FERMAT.

### 3. Le problème de FERMAT-WEBER.

**3.1** Considérons trois points  $A, B, C$  (non alignés) du plan. Si l'on cherche à minimiser la somme des carrés des distances (euclidiennes usuelles) d'un point courant  $P$  (du plan) aux points  $A, B, C$ , on trouve de manière classique un seul point optimal : le point  $G$  centre de gravité (ou isobarycentre) du triangle  $ABC$ . Le problème est tout autre si on considère, dans le critère à minimiser, les distances (et non les carrés des distances) aux points  $A, B, C$  ; c'est ce qu'on appelle parfois **le problème géométrique de FERMAT**. Dans un contexte d'applications, imaginons que l'on doive placer un hélicoptère  $H$  appelé à desservir trois villes  $A, B, C$  ; le coût du déplacement est proportionnel à la distance aux villes (ce qui est plus logique que la proportionnalité aux carrés des distances) ; question : où faut-il placer l'hélicoptère  $H$  de manière à minimiser la somme des coûts de déplacement ? Nous laissons de côté les contraintes éventuelles du problème (et on comprend qu'il peut y en avoir : l'hélicoptère ne peut être construit dans telle ou telle zone). D'un point de vue mathématique, le problème est le suivant : il s'agit de minimiser la fonction de  $P$

$$f(P) := \text{distance de } P \text{ à } A + \text{distance de } P \text{ à } B + \text{distance de } P \text{ à } C.$$

Le problème peut être résolu par de simples considérations géométriques spécifiques au cas particulier du plan. Du point de vue Optimisation, la fonction-objectif  $f$  ci-dessus est convexe, "coercive à l'infini" (c'est-à-dire  $f(P)$  tend vers l'infini quand (la longueur de)  $P$  tend vers l'infini) ; dans notre problème d'optimisation, l'existence d'une solution et même son unicité ne pose pas de problème majeur. La difficulté vient de la *caractérisation* de

la solution existante; en effet la fonction  $f$  n'est pas différentiable partout où les points candidats à être solutions peuvent se nicher, aux points  $A, B, C$  pour être plus précis. Si, par quelque raisonnement relatif à la position mutuelle des points  $A, B, C$ , on est capable de dire que le point optimal  $H$  se trouve à l'intérieur du triangle  $ABC$ , alors la règle de FERMAT s'applique (puisque  $f$  est différentiable en un tel point  $H$ ): l'équation d'optimalité  $\nabla f(H) = 0$  caractérise  $H$  (cf. Figure 4.a). Si tel n'est pas le cas, il faut faire appel à une condition d'optimalité plus générale, relative aux problèmes convexes, de manière à ne pas écarter les points  $A, B$  ou  $C$  parmi les candidats solutions possibles. Quoi qu'il en soit, le résultat est le suivant : *si les 3 angles du triangle  $ABC$  sont strictement inférieurs à  $120^\circ$ , le point solution  $H$  réalisant le minimum dans le problème géométrique de FERMAT est celui à partir duquel on voit les 3 côtés sous un même angle de  $120^\circ$  (point appelé parfois de FERMAT-TORRICELLI-STEINER); si l'un des angles est supérieur ou égal à  $120^\circ$ , le point  $H$  est le sommet correspondant du triangle.*

Comme on s'en doute, le problème géométrique de FERMAT admet de multiples généralisations, avec des applications pratiques indéniables. Imaginons qu'il faille placer un centre social, dépôt,...destiné à servir  $n$  usagers, clients,... $A_1, A_2, \dots, A_n$ , et que le coût de déplacement soit proportionnel à la distance de déplacement; on a alors à résoudre le problème de localisation dit de FERMAT-WEBER<sup>3</sup> : déterminer le point (du plan) minimisant

$$f(P) := \text{distance de } P \text{ à } A_1 + \text{distance de } P \text{ à } A_2 + \dots \text{distance de } P \text{ à } A_n.$$

Résoudre numériquement ce problème n'est pas très difficile, il relève de techniques à présent standard de minimisation convexe. Dans certains cas, comme celui de quatre points ( $n = 4$ ), la solution du problème de FERMAT-WEBER est explicite (pour quatre points, c'est le point intersection des diagonales du quadrilatère construit à partir des quatre points donnés).

### 3.2 Prise en compte de la non-différentiabilité.

Comme on l'a vu dans l'exemple du paragraphe précédent (problème de FERMAT-WEBER), on peut être amené à minimiser des fonctions qui ne sont pas différentiables. Une vaste classe de problèmes d'optimisation à données non-différentiables, d'importance pratique considérable, est celle où le critère à minimiser est de la forme  $f = \max(f_1, f_2, \dots, f_m)$ ; même si les fonctions  $f_k$  sont très "régulières" (différentiables autant de fois que l'on veut, par exemple), la fonction  $f$ , elle, n'est pas différentiable en tout point. Alors comment fait-on? Quelle est la généralisation de la règle de FERMAT dans de tels cas? On va en donner quelques idées pour certaines classes de fonctions.

Tout d'abord les fonctions *convexes*  $f : x = (x_1, x_2, \dots, x_n) \mapsto f(x)$ . En un point  $a$  on collecte toutes les pentes  $\mathbf{p}$  des droites affines qui minorent  $f$  et qui coïncident avec elle en  $a$  (voir Figure 4) :

$$\{\text{pentes } \mathbf{p} \text{ telles que } f(x) \geq f(a) + \mathbf{p} \cdot (x - a) \text{ pour tout } x\}.$$

---

<sup>3</sup>A. WEBER est un économiste allemand qui a travaillé sur *les problèmes de localisation*. Son ouvrage de 1909 traite de la localisation des industries.



( $\mathbf{p}.u$  désigne ici le produit scalaire usuel entre les vecteurs  $n$ -dimensionnels  $\mathbf{p}$  et  $u$ ).

Cet ensemble, noté  $\partial f(a)$  (graphisme semblable à celui d'une dérivée partielle, ce qui peut prêter à confusion), est ce qu'on appelle *le (ou la) sous-différentiel(le) de  $f$  en  $a$* . Introduit il y a à peine une quarantaine d'années, il joue pour les fonctions convexes  $f$  le rôle de vecteur gradient  $\mathbf{y}$  compris là où le gradient de  $f$  n'existe pas, c'est-à-dire :  $\partial f(a) = \{\nabla f(a)\}$  dès lors que  $f$  est différentiable en  $a$ ; sinon c'est un paquet (convexe) de vecteurs (voir Figure 4.b). Par exemple, si  $f$  est la fonction (de la variable réelle) "valeur absolue de",  $\partial f(a) = \{\pm 1\}$  en un point  $a \neq 0$ , tandis que  $\partial f(0) = [-1, +1]$ . La règle de FERMAT devient ici :

La fonction convexe  $f$  est minimisée en  $a$  si, et seulement si,  $\mathbf{0} \in \partial \mathbf{f}(a)$ .

Pour la classe de *fonctions du type max*, c'est-à-dire  $f = \max(f_1, f_2, \dots, f_m)$ , construite à partir de fonctions  $f_k$  continûment différentiables (= qui ont un gradient continu partout), le bon objet mathématique substitut de la notion de gradient en  $a$  se trouve être (ça se démontre)

$$\begin{aligned} \partial f(a) &: = \text{plus petit polytope convexe contenant les } \nabla f_k(a), k \text{ tel que } f_k(a) = f(a) \\ &= \{\text{barycentres des vecteurs } \nabla f_k(a), k \text{ tel que } f_k(a) = f(a)\}. \end{aligned}$$

Ici encore ce polytope ne contient qu'un seul vecteur, le gradient de  $f$  en  $a$ , lorsque  $f$  est différentiable en  $a$ . Dans le contexte de telles fonctions, la règle de FERMAT devient :

Si  $f = \max(f_1, f_2, \dots, f_m)$  est minimisée en  $a$ , alors  $\mathbf{0} \in \partial \mathbf{f}(a)$ .

Enfin, il y a une trentaine d'années, a été considérée une classe encore plus vaste de fonctions (éventuellement non-différentiables), c'est celle des fonctions dites localement LIPSCHITZ  $f$  (à laquelle appartiennent les deux classes de fonctions évoquées ci-dessus); le franco-canadien F. CLARKE a proposé en 1973 un objet généralisant la notion de gradient, appelé d'ailleurs par lui *gradient généralisé de  $f$*  et noté par le même graphisme  $\partial f$  que précédemment. Pour une présentation simple du gradient généralisé et de ses propriétés, le lecteur peut consulter [4].

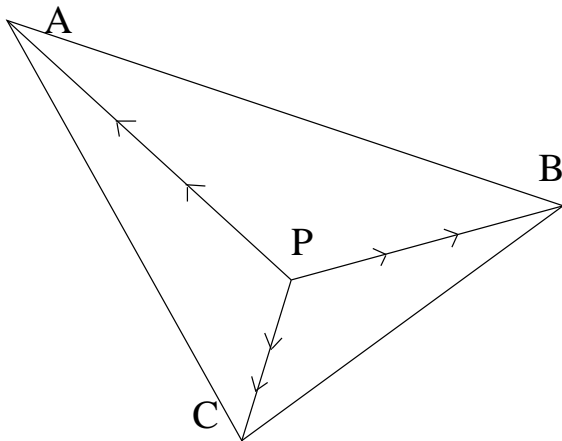


Figure 4.a

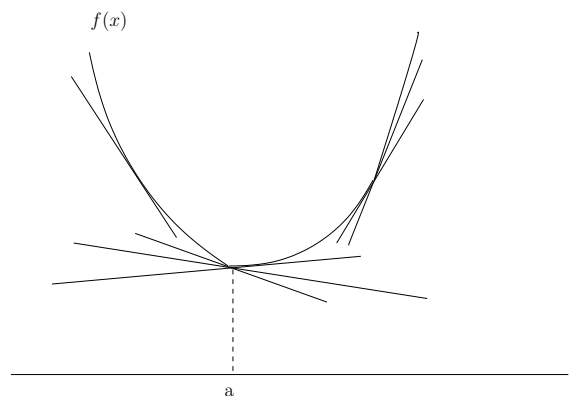


Figure 4.b

#### 4. Le Calcul variationnel ; la Commande optimale.

4.1 Avec la mise sur la place publique, sous forme de défi, en 1696, du problème de la courbe *brachystochrone* par Johann BERNOULLI, peut être datée la naissance publique du *Calcul des variations* ou, dans une appellation plus moderne, l'*Analyse et le Calcul variationnels* (encore que ce qu'on appelle le problème aérodynamique de NEWTON fut posé avant, en 1686, et que G. GALILEE ait abordé le problème de la brachystochrone en 1638). Voici brièvement en quoi consiste la recherche d'une courbe brachystochrone : dans le plan vertical, on souhaite aller de la position  $A$  à la position  $B$  sous le seul effet de la pesanteur, et ce *le plus vite possible* ; quelle est pour cela la trajectoire qu'il faut suivre ? On peut imaginer qu'il s'agisse de construire un toboggan reliant  $A$  à  $B$ , sur lequel, négligeant les forces de frottement, un enfant descende le plus rapidement possible. L'inconnue n'est donc plus une ou plusieurs variables réelles, mais une fonction (= "une infinité de variables"), ce qui complique considérablement le problème. La réponse est connue - elle l'était d'ailleurs au moment où le problème fut posé -, il s'agit d'une *arche de cycloïde* (ou *roulette*).

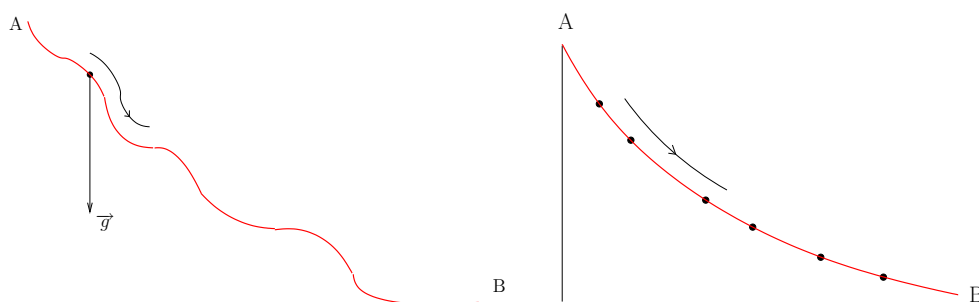


Figure 5

Le Calcul des variations fut développé au 18ème siècle par EULER et LAGRANGE, au 19ème siècle par LEGENDRE, JACOBI, HAMILTON et WEIERSTRASS entre autres. La première moitié du 20ème siècle vit les contributions importantes de BOLZA, BLISS, etc. La condition nécessaire d'optimalité du premier ordre, descendante de la règle de FERMAT pour le calcul des variations, est la condition dite d'EULER (1744)-LAGRANGE(1762) ; nous éludons son explicitation car elle nous amènerait trop loin. La deuxième moitié du 20ème siècle verra naître une forme moderne de l'Analyse et Calcul variationnels, à savoir la *Commande optimale* (ou le *Contrôle optimal*). Précédé par des travaux d'ingénieurs, ce domaine d'études fut impulsé par les besoins de résoudre des problèmes pratiques issus du monde militaire ou spatial ; les deux noms qui ressortent sont ceux de BELLMAN aux Etats-Unis et de PONTRYAGIN et al. en URSS. On en dira quelques mots dans le paragraphe 4.2.

Illustrons, à l'aide d'un exemple, le Calcul variationnel et les questions qui s'y posent.

*Un bateau, perdu en mer, sait qu'il est à une distance de 1 km d'un rivage rectiligne (ce que lui indiquent ses instruments de mesure), mais le brouillard est si épais qu'il ignore*

la direction du rivage. Le bateau, avançant à vitesse constante, voudrait toucher le rivage le plus vite possible. La question est donc : quel est le chemin de longueur minimale que le bateau doit suivre afin d'être sûr de toucher terre ?

Toutes les questions du Calcul variationnel peuvent être posées à propos de cet exemple : l'objet optimal recherché est une trajectoire (= une courbe du plan)...Y a-t-il des solutions (une trajectoire vraiment plus courte que toutes les autres) ? Y en a-t-il plusieurs ? Comment caractériser (mathématiquement) les trajectoires optimales ? Comment en approcher une par un procédé numérique ?

En termes mathématiques : soit le plan avec une origine 0 (la position de départ du bateau) ; il s'agit de trouver *une courbe du plan démarrant en 0, de longueur minimale et qui touche (ou coupe) toute droite du plan à distance 1 km de l'origine.*

La solution, si solution il y a, ne saurait être unique ; en effet, une trajectoire optimale, tournée d'un angle quelconque autour de l'origine est encore optimale.

Une première tentative, montrant au moins la faisabilité de ce qui est demandé sur la trajectoire recherchée est de considérer ceci : le bateau part de l'origine 0 en suivant un rayon (pris dans une direction au hasard) du cercle de rayon 1 km ; au bout du rayon il fait le tour complet du cercle (voir Figure 6.a). Il est ainsi sûr de toucher toute droite du plan à distance de 1 km de l'origine ; il aura parcouru au total  $(2\pi + 1) \approx 7,2832$  km. Mais il y a sans doute mieux à faire...Comment ? J'ai posé la question sous forme de défi à un groupe de jeunes ingénieurs travaillant dans la sous-traitance aéronautique à Toulouse. Leur réponse fut, comme cela est décrit dans la Figure 6.b : le bateau se déplace d'abord au-delà de l'extrémité du rayon de départ, revient vers le cercle le long d'un segment tangent au cercle, suit une bonne partie du tour du cercle, et complète par un nouveau segment de droite tangent au cercle ; l'ouverture angulaire du secteur  $S$  (délimité par le rayon de départ et le rayon dirigé vers le point d'arrivée) étant de (deux fois)  $45^\circ$ . Ce n'est pas si mal ! En effet, un bateau suivant cette trajectoire est sûr de couper (ou de toucher) le rivage rectiligne situé à 1 km de l'origine (point de départ), où qu'il soit ! Et la longueur parcourue est de (après quelques petits calculs)  $(\pi + 2 + \sqrt{2}) \approx 6,5556$  km.

Mais il y a encore mieux à faire dans le même registre. Considérons une ouverture angulaire du secteur  $S_\theta$  d'angle  $\theta$  compris entre 0 et  $45^\circ$  (cf. Figure 6.c) et cherchons la valeur de  $\theta$  qui minimiserait la longueur de la trajectoire correspondante  $C_\theta$ . La longueur  $L(\theta)$  de la trajectoire  $C_\theta$  est  $(2\pi - 4\theta + 2 \cdot \tan(\theta) + \frac{1}{\cos(\theta)})$  km. Elle est minimale pour une valeur de  $\theta_{opt}$  intérieure à l'intervalle  $[0, 45^\circ]$ , en le seul point d'annulation de la dérivée de  $L(\theta)$  (application de la règle de FERMAT) ; cela donne  $\theta_{opt} \approx 36,37^\circ$  et  $L_{opt} := L(\theta_{opt}) \approx 6,4589$  km. La première tentative correspondait à l'ouverture angulaire de  $\theta = 0$ , tandis que la courbe proposée par les jeunes ingénieurs correspondait à l'ouverture angulaire de  $\theta = 45^\circ$ .

Il n'empêche que le mathématicien reste insatisfait : qui nous dit qu'on ne peut pas faire plus court, bref que la trajectoire trouvée au-dessus est vraiment la plus courte possible ? Ceci est l'affaire des conditions suffisantes d'optimalité en Calcul variationnel, que nous n'abordons pas ici. Que le lecteur sache néanmoins qu'on peut faire un peu mieux que la courbe  $C_\theta$  avec  $\theta_{opt}$  explicitée au-dessus...nous livrons le problème à sa sagacité.

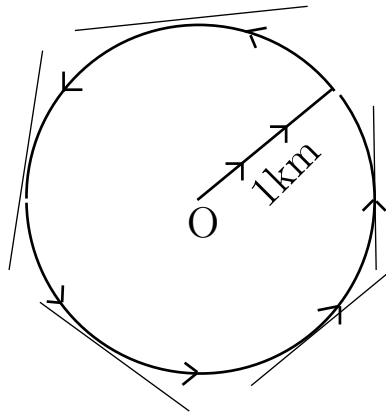


Figure 6.a

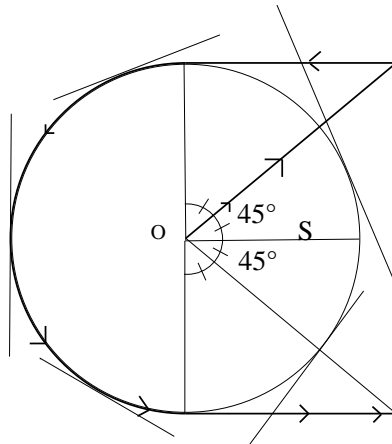


Figure 6.b

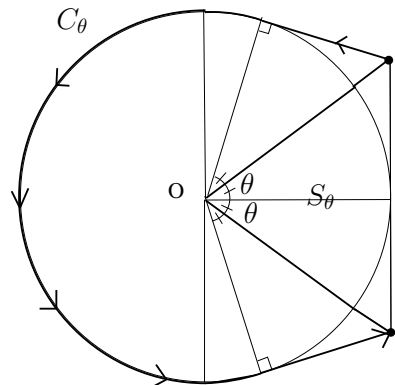


Figure 6.c

#### 4.2. La Commande optimale

A la différence de l'Analyse et du Calcul variationnels où "on laisse faire la Nature", en Commande optimale (on dit aussi Contrôle optimal) on agit sur le système ou on le manoeuvre par l'intermédiaire d'une fonction auxiliaire  $t \mapsto u(t)$  (appelée *fonction de commande* ou *contrôle*) de manière à minimiser un certain critère, tout en satisfaisant un certain nombre de contraintes sur les inconnues. C'est une situation fréquente dans les domaines des sciences de l'ingénieur. D'une manière schématique :

- l'état du système, décrit par  $x(\cdot)$  (vecteur de  $n$  composantes), évolue suivant une équation différentielle (vectorielle)

$$\frac{dx}{dt} = f(t, x(t), \mathbf{u}(t)),$$

où  $u(\cdot)$  est soumis à des contraintes  $u(t) \in U$  (ensemble fixé, par des limitations techniques par exemple) ;

- le critère à minimiser prend la forme générale

$$J := g(t_0, x(t_0), t_f, x(t_f)) + \int_{t_0}^{t_f} l(t, x(t), u(t)) dt ;$$

-  $t_f$  (date terminale),  $x(t_f)$  (position terminale) (et, plus rarement la date de départ  $t_0$  et la position de départ  $x(t_0)$ ) peuvent faire partie des inconnues du problème ;

- il peut y avoir des contraintes sur  $x(\cdot)$  et  $u(\cdot)$ , ponctuelles ou réparties sur toute la durée  $[t_0, t_f]$ .

Un exemple - pas seulement d'école - est le suivant : une rame de métro que l'on prend à l'instant  $t_0 = 0$  dans des conditions initiales ( $x(0) = x_0, \frac{dx}{dt}(0) = v_0$ ) doit être amenée à la position terminale  $x_f$  au repos ( $x(t_f) = x_f, \frac{dx}{dt}(t_f) = 0$ ) ; la commande sur laquelle on peut agir est l'accélération  $\frac{d^2x}{dt^2} = u$ , mais celle-ci est contrainte :  $-u_{\max} \leq u(t) \leq u_{\max}$  ; le critère choisi peut être  $J = t_f$  (la durée de parcours pour aller de  $x_0$  à  $x_f$ ), une combinaison de la durée de parcours et de la consommation  $\int_0^{t_f} |u(t)| dt$ . Il n'est pas du tout évident

de deviner ou détecter la commande optimale  $u(\cdot)$ ...Par exemple, pour le premier exemple de critère (durée de parcours), avec un départ au repos ( $v_0 = 0$ ), la stratégie optimale se trouve être la suivante :  $u(t) = +1$  (accélération maximale possible) jusqu'à atteindre le milieu du parcours ;  $u(t) = -1$  (rétropoussée maximale possible) sur la deuxième moitié du parcours.

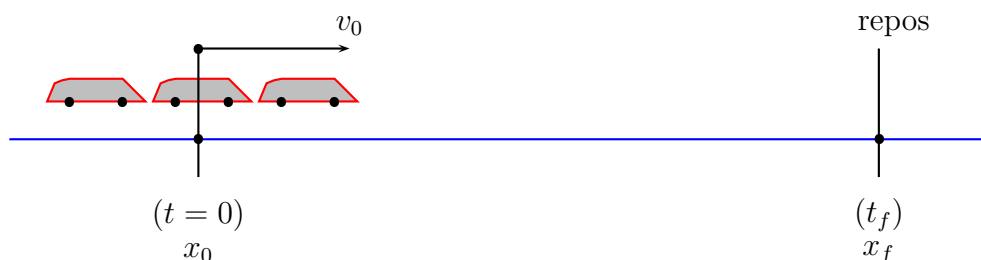


Figure 7

Historiquement, ce sont des problèmes en temps minimal (c'est la durée  $t_f - t_0$  qu'il faut minimiser) posés par les ingénieurs dans le secteur aéronautique et spatial qui ont stimulé les recherches qui conduiront (vers les années 1955-1960) au point d'orgue que représente *le principe du maximum de PONTRYAGIN* (PMP en abrégé). Qu'est-ce-que le PMP ? C'est une condition *nécessaire d'optimalité*, c'est-à-dire un jeu de relations que doit nécessairement vérifier la commande optimale et, le cas échéant la date optimale, dans notre problème. C'est donc encore un petit descendant de la règle à FERMAT, auquel ont rendu hommage régulièrement les savants russes du domaine. Avoir établi le PMP est sans doute l'une des réalisations mathématiques les plus brillantes de la deuxième moitié du 20ème siècle. Le PMP et ses applications restent essentielles dans les Sciences appliquées ; on peut citer comme exemples : le secteur aéronautique et spatial, la Robotique, l'Economie mathématique, La Biomathématique...Au CNES (Centre National d'Etudes Spatiales) qui nous est cher à Toulouse, il ne se passe pas un jour sans que des ingénieurs de recherche utilisent le PMP (pour le transfert d'un satellite d'une orbite à une autre par exemple) ; ils travaillent d'ailleurs dans un bâtiment qui porte le nom de...FERMAT. Pour un livre très récent en français sur la Commande optimale (dans le "monde non-linéaire"), nous suggérons [13].

Pour terminer, nous voudrions signaler un ouvrage récent qui est une ode au Calcul différentiel ou variationnel ainsi qu'à l'Optimisation, il s'agit de la référence [12] : l'apport de FERMAT y est grandement évoqué.

### Références.

[1] P. DUGAC, *FERMAT et la recherche des extremums*, in Histoire de l'Analyse : Autour de la notion de limite et de ses voisinages. Editions Vuibert (2003), p. 39-40.

- [2] P. P. B. EGGERMONT, *Noncentral difference quotients and the derivative*, American Mathematical Monthly, Vol. 95, N°6 (1988), p. 551-553.
- [3] J. V. GRABINER, *The changing concept of change : the derivative from FERMAT to WEIERSTRASS*. Mathematics Magazine, Vol. 56, N°4 (1983), p.195-206.
- [4] J.-B. HIRIART-URRUTY, *Un concept récent pour l'analyse et l'optimisation de fonctions non-différentiables : le gradient généralisé*. Revue de Mathématiques Spéciales (1986), p. 311-321.
- [5] J.-B. HIRIART-URRUTY, *Flâneries mathématiques*. Version écrite de l'exposé au symposium "Convexity and Duality in Optimization", à l'occasion des festivités marquant le 370ème anniversaire de la Rijksuniversiteit de Groningen (Pays-Bas) en juin 1984. Traduction partielle en Anglais publiée dans Lecture Notes in Economics and mathematical Systems 256 (1985), p. 1-9.
- [6] J.-B. HIRIART-URRUTY, *Corrigé détaillé et commenté de l'épreuve "Analyse numérique" du concours externe d'Agrégation de mathématiques 1992*. Revue de Mathématiques Spéciales (1993), p. 252-276.
- [7] J.-B. HIRIART-URRUTY, *Une "bonne" condition nécessaire de minimalité locale*. Revue MATAPLI de la Société des Mathématiques Appliquées et Industrielles, N°40 (1994), p. 35-39.
- [8] J.-B. HIRIART-URRUTY, L'Optimisation. Collection "Que sais-je ? ", Presses Universitaires de France (1996).
- [9] J.-B. HIRIART-URRUTY, *Sur le caractère spectaculaire du théorème de FERMAT-WILES*. Bulletin de l'Association des Professeurs de Mathématiques, N°427 (2000), p. 209-210.
- [10] J.-B. HIRIART-URRUTY, *P. FERMAT : un homme, une époque, de grands résultats*. Exposé au colloque "Un jour avec P. Fermat" au Centre Français de Recherche en Sciences Sociales CEFRES) de Prague le 8 octobre 2001. Traduction en Tchèque publiée dans les cahiers du CEFRES n° 28, Prague (2002), p. 123-129.
- [11] J.-B. HIRIART-URRUTY, *L'Optimisation : deux ou trois choses que je sais d'elle*. Publications du CNES, Centre de compétence technique "Mécanique orbitale" N°148 (2002), et Revue MATAPLI de la Société des Mathématiques Appliquées et Industrielles, N°71 (2003), p. 31-53.
- [12] P. J. NAHIN, When least is best. Princeton University Press (2004).
- [13] E. TRELAT, Contrôle optimal : Théorie et Applications. Collection "Mathématiques concrètes", Editions Vuibert (2005).
- [14] Des Problèmes d'extrema chez FERMAT à la notion de dérivée : aspects historiques et expériences pédagogiques en classe de Première (année scolaire 1989-1990). Sous la direction de M. CLAPIE et M. SPIESSER, MAFPEN-IREM de Toulouse (1991).
- [15] FERMAT, le premier homme du monde. Produit multimédia (livret et cassette vidéo de 25 mn). CRDP Midi-Pyrénées-IREM de Toulouse (1995).

**Addendum** (année universitaire 2005-2006).

Une première diffusion de ce texte auprès de collègues, ainsi que le compte rendu de lecture du livre [12] dans le bulletin de la Société Mathématique de France où je posais également le problème de la courbe optimale énoncé à la fin du paragraphe 4.1, m'ont valu plusieurs réactions de lecteurs : ils répondaient au défi que j'y posais en proposant une courbe meilleure que celle proposée au 4.1, optimale selon eux, sans démonstration de l'optimalité toutefois. Nous présentons ci-dessus cette courbe qui, en effet, est optimale pour le problème posé (Fig. 8a).

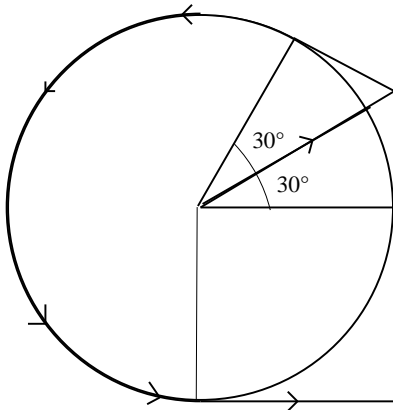


Figure 8.a

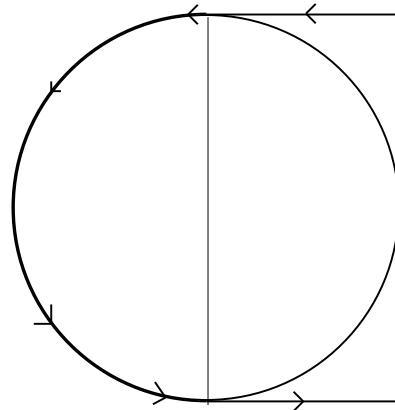


Figure 8.b

Le caractère optimal de cette courbe est démontré dans l'article dont la référence est ci-dessous (que j'ai découvert après la rédaction de mon article), une douzaine de pages de considérations analytiques et géométriques spécifiques au plan....un peu casse-pieds, je dois dire ; je reste demandeur d'une démonstration plus directe. La longueur de cette courbe est  $\frac{7\pi}{6} + \sqrt{3} + 1 \approx 6,3972$  km. Quand on dit "la" courbe optimale, on veut dire que toute autre solution s'obtient à partir de celle explicitée au-dessus par des rotations et réflexions du plan laissant le cercle en place.

Réf. H. JORIS, *Le chasseur perdu dans la forêt : un problème de géométrie plane.* Elemente der Mathematik 35, N°1 (1980), p. 1-14.

Un problème un peu plus simple, mais dont la solution aide à comprendre l'optimalité de la courbe au-dessus, est comme suit. Une compagnie téléphonique a mis des repères ponctuels au sol (des plots par exemple) ; quand on est sur un repère, on sait qu'on est à moins de 1 m du câble téléphonique rectiligne. Un problème est signalé, il faut accéder au câble : comment l'agent dépêché par la compagnie doit-il creuser pour être sûr de tomber sur le câble, en creusant le moins possible ? Plus court que le cercle centré au point repère et de rayon 1 m, bien sûr. C'est donc le problème précédent sans la contrainte sur le point de départ. La "tranchée optimale" se trouve être : un demi-cercle prolongé par deux segments de droite, comme à la fin du parcours de la courbe précédente ; voir Fig. 8b.

Le problème évoqué au 4.1 a une généralisation en dimension trois : *trouver la courbe partant de l'origine qui coupe ou touche tout plan tangent à la sphère de rayon r.* Là, je

n'en suis qu'à collecter des propositions (certaines biscornues)...Le record actuel (qui date de mai 2006) est une courbe de longueur  $13,6699 r$ . Voici une présentation du problème, habillée par T. DRUOT (Service Avant-projets Airbus-France, Toulouse) qui avait réagi à ce défi, et dans laquelle il fait allusion à quelques uns de ces dadas intellectuels.

*Le capitaine Keldeveïn était vraiment dans l'embarras. Après l'explosion du générateur principal de son vaisseau qui avait projeté ce qu'il en restait à 1 parsec au plus du Portail Stellaire, il savait à présent qu'il ne lui restait que soixante heures à vivre, à moins qu'il ne trouve le moyen de rejoindre la porte en moins de 20 parsecs, ce qui représentait la quasi-totalité du trajet possible dans le temps imparti. Du moins c'est ce que prétendait la fraction encore opérationnelle de l'intelligence cybernétique du vaisseau. Par chance, à une telle distance, la porte pouvait être considérée comme une surface infinie. Cependant, n'ayant plus aucun instrument de navigation à sa disposition, il fallait se résoudre à chercher à l'aveuglette. Le capitaine était bien prêt de sombrer dans l'abattement caractéristique des hommes privés d'un coup de toutes les ressources de la technologie qui épaulaient l'humanité depuis si longtemps qu'elles s'intégraient au plus profond des processus de pensée. Son subconscient semblait avoir baissé les bras car il lui renvoyait des images de sa jeunesse, alors qu'il était encore à l'université et qu'il se passionnait pour les jeux mathématiques. L'un de ces professeurs, d'origine basque, avait tout particulièrement contribué à affûter son intelligence. Il se souvenait presque de son nom, mais aussi qu'un jour, il l'avait fait plancher sur un problème qui n'était autre que la version dans le plan du problème qui se posait maintenant si cruellement à sa survie et à celle de son équipage. Alors l'idée se forma qu'ils n'étaient peut-être pas perdus. Mathfolks, dit-il soudain en se tournant vers le pilote du vaisseau, vous souvenez-vous du "problème du voyageur de commerce" ? L'interpellé écarquilla des yeux dans lesquels il eut grand peine à masquer la crainte que le capitaine ait pu tout à coup sombrer dans la folie. Mais le capitaine reprit : mettons-nous au travail voulez-vous...*

Terminons par encore un autre problème, de nature voisine, posé par A. GRIGIS (université de Paris XIII) en réagissant au problème au-dessus : *Quelle est la longueur minimale d'une courbe de l'espace qui peut être vue de tout point de la sphère ?* On imagine aisément que ce problème peut avoir des applications pratiques. Je ne connais pas la réponse.