

# Cours numéro 1 :

## modélisation par suites et fonctions

### 1 Paramètres, variables, relations, fonctions,

#### 1.1 Introduction

Il y a essentiellement deux outils mathématiques qui permettent de modéliser les situations “réelles” : les suites et les fonctions, qui correspondent à des modélisations discrètes et continues. Nous allons d’abord étudier le cas des fonctions, d’une ou plusieurs variables réelles, nous reviendrons plus tard, et brièvement, sur le cas des suites.

Lorsqu’on essaie de modéliser une situation réelle, ce n’est qu’exceptionnellement que se dégagent des variables<sup>1</sup> et des fonctions de manière naturelle et c’est encore plus rare que la situation soit modélisée par **une** fonction d’**une** variable, ce qui est la seule situation mathématique que connaissent vraiment les lycéens. C’est une difficulté pour les professeurs de mathématiques qui souhaitent illustrer leur enseignement par des applications. En général, la situation est caractérisée par un ensemble de paramètres, qui permettent de la décrire, et il y a des relations entre ces paramètres<sup>2</sup>. Dans un premier temps, nous analysons l’origine des paramètres et des relations.

#### 1.2 Exemple 1 : la boîte de conserves

La question est la suivante :

*Vous êtes fabricant de boîtes de conserves. Vous devez produire, au moindre coût, des boîtes cylindriques de volume donné. Que proposez-vous ?*

Les deux quantités qui nous intéressent sont le volume  $V$  de la boîte et son aire  $S$  (qui mesure la quantité de métal). Ces quantités dépendent des dimensions de la boîte, c’est-à-dire du rayon de la base  $R$  et de la hauteur

---

1. Tout de même, en physique, il y a souvent une variable naturelle lorsqu’on étudie l’évolution d’un système, c’est le temps.

2. La plupart des lois de la physique sont ainsi données comme des relations entre des paramètres.

$h : V = \pi R^2 h$  et  $S = 2\pi R h + 2\pi R^2$ . On voit qu'on a des relations entre les divers paramètres, mais sous forme implicite. C'est le cas dans la plupart des modélisations. Dans le cas présent, si l'on fixe  $V$ , on peut aisément tirer  $h$  en fonction de  $R : h = \frac{V}{\pi R^2}$ . On en déduit  $S = 2\pi(R^2 + \frac{V}{\pi R})$  qui est minimum pour  $R = \sqrt[3]{\frac{V}{2\pi}}$ . Il reste à formuler le résultat de manière compréhensible. La valeur de  $R$  vérifie  $2R = \frac{V}{\pi R^2} = h$  : la hauteur doit être égale au diamètre<sup>3</sup>.

On peut aussi tirer  $R$  en fonction de  $h : R = \sqrt{\frac{V}{\pi h}}$  ce qui donne  $S = 2\frac{V}{h} + 2\sqrt{\pi V h}$ . Avec cette forme le calcul est un peu plus compliqué : on voit ici l'importance du choix des paramètres.

On peut poser aussi le problème dual : pour une quantité de métal donnée, quelle forme maximise le volume ? Dans ce cas, si l'on fixe  $S$ , on peut aisément tirer  $h$  en fonction de  $R$ . On obtient  $V = \frac{RS}{2} - \pi R^3$  qui est maximum pour  $R^2 = S/6\pi$ . Si on calcule  $h$ , on retrouve  $h = 2R$ . On peut aussi tirer  $R$  en résolvant l'équation du second degré  $2\pi R^2 + 2\pi R h - S = 0$ . On trouve  $R = \frac{-\pi h + \sqrt{\pi^2 h^2 + 2\pi S}}{2\pi}$  (le signe + est obligatoire). Cette fois, le calcul est nettement plus compliqué.

Dans ce cas, les paramètres (variables ou fonctions) sont définis de manière naturelle et on a facilement des relations entre eux. On obtient des fonctions en résolvant ces relations, c'est-à-dire en exprimant un paramètre en fonction des autres. Le problème peut facilement être traité au lycée.

### 1.3 Exemple 2 : la loi d'Ohm

Il s'agit de la loi qui relie la tension à l'intensité :  $U = RI$  où  $R$  désigne la résistance du conducteur. Que dit cette loi ? Simplement qu'il y a **proportionnalité** entre tension et intensité, le coefficient baptisé résistance étant une caractéristique du conducteur. Dans ce cas, aux quantités physiques, tension et intensité, s'ajoute une quantité supplémentaire définie par la relation de proportionnalité. Ensuite, d'ailleurs, ce coefficient peut être calculé à son tour par la formule  $R = \rho l/s$ . Là encore, ce qu'on dit c'est que  $R$  est proportionnelle à la longueur du conducteur (supposé cylindrique) et inversement proportionnelle à sa surface et on donne un nom au coefficient (résistivité). On introduit ici, à côté des caractéristiques géométriques du conducteur une

---

3. C'est souvent le cas – mais pas toujours – pour les boîtes de conserves du commerce.

nouvelle constante  $\rho$ .

Bref, un premier type de fonctions apparaît ici et dans beaucoup de situations (mais pas toutes) : les fonctions linéaires, qui traduisent la proportionnalité.

L'intérêt de ce type de formule est de permettre de déterminer les caractéristiques d'un conducteur à partir de certaines d'entre elles. Par exemple : on a une ampoule de 100 watts et un courant de 220 volts, on en déduit  $R$  et  $I$  grâce aux formules<sup>4</sup>  $P = UI$  et  $U = RI$ .

## 1.4 Exemple 3 : la dilatation

On donne souvent une formule pour la longueur d'un fil en fonction de la température :  $l = l_0(1 + \alpha(t - t_0))$  où  $l_0$  est la longueur à la température  $t_0$ . On donne aussi alors la variation de volume :  $V = V_0(1 + 3\alpha(t - t_0))$ . D'où sortent de telles formules ? Essentiellement d'une approximation linéaire au premier ordre pour de petites variations du paramètre  $t - t_0$ . On écrit, en fait, la formule de Taylor à l'ordre 1 :  $l = f(t) = f(t_0) + (t - t_0)f'(t_0) + o(t - t_0)$  et le  $\alpha$  n'est autre que  $f'(t_0)/l_0$ . D'ailleurs, lorsqu'on juge cette approximation linéaire trop grossière, on utilise une approximation quadratique  $l = l_0(1 + \alpha t + \beta t^2)$ . Pour passer au volume, on élève au cube  $V = l^3 = l_0^3(1 + \alpha(t - t_0))^3 \sim l_0^3(1 + 3\alpha(t - t_0))$ .

Ici, donc, faute de connaître vraiment la loi de la dilatation, on en prend une approximation, linéaire, voire polynomiale.

## 1.5 Exemple 4 : l'optique géométrique

### 1.5.1 La situation

On s'intéresse à des systèmes optiques, c'est-à-dire une succession de milieux transparents et homogènes séparés par des dioptries<sup>5</sup> ou des miroirs. On étudiera notamment le cas du miroir sphérique. Dans tout ce qui suit on supposera qu'on a affaire à un système centré, c'est-à-dire à un système qui admet un axe de révolution, que l'on tracera toujours horizontal et qui sera orienté de gauche à droite (sens de parcours de la lumière). Le problème est le suivant : étant donné un objet  $A$  (une source lumineuse située sur l'axe), déterminer son image  $A'$  après passage par le miroir ou la lentille.

---

4. Sauf erreur, ces formules sont encore vraies en courant alternatif, au moins s'il n'y a pas de déphasage entre les quantités, ce qui est le cas pour une lampe ou un appareil de chauffage mais ne serait plus vrai pour un moteur. Dans ce cas, il y a un facteur  $\cos \varphi$  qui s'introduit.

5. Un dioptrie est une surface séparant deux milieux dans lesquels la vitesse de la lumière est différente.

Pour traiter ce problème, on supposera toujours qu'on est dans les **conditions de Gauss** : les rayons lumineux sont peu inclinés sur l'axe et proches de l'axe.

Traisons le cas du miroir sphérique (ici concave). Le miroir est modélisé par un arc de cercle de centre  $C$ , situé sur l'axe. On note  $S$  le sommet du miroir (le point d'intersection avec l'axe) et on cherche l'image d'un rayon  $AI$  issu d'un point  $A$  de l'axe, se réfléchissant en  $I$  sur le miroir et recoupant l'axe en  $A'$ . La normale au miroir en  $I$  passe par le centre du miroir. C'est donc la droite  $(IC)$ . La loi de la réflexion indique que l'angle d'incidence  $\widehat{AIC}$  est égal à l'angle de réflexion  $\widehat{CIA'}$  et on le note  $i$ . Par ailleurs, on note respectivement  $\alpha, \beta, \alpha'$  les angles  $\widehat{SAI}$ ,  $\widehat{SCI}$  et  $\widehat{SA'I}$ . La somme des angles des triangles  $ACI$  et  $A'CI$  donne alors  $i = \beta - \alpha = \alpha' - \beta$  et donc  $2\beta = \alpha + \alpha'$ . On a alors le résultat suivant :

**1.1 Proposition.** *Dans les conditions de Gauss, on a la formule (dite de conjugaison) :*

$$\frac{1}{\overline{SA}} + \frac{1}{\overline{SA'}} = \frac{2}{\overline{SC}}.$$

*Démonstration.* Appelons  $R$  le rayon du cercle. Comme on suppose les angles  $\alpha, \beta, \alpha'$  petits, on peut les assimiler à leurs sinus ou leurs tangentes. De plus, comme on a  $\overline{SH} = R(1 - \cos \beta)$  il est voisin de 0 (équivalent à  $\beta^2/2$ ) et on peut donc en définitive écrire :

$$\alpha \simeq \frac{\overline{HI}}{\overline{SA}}, \quad \beta \simeq \frac{\overline{HI}}{\overline{SC}}, \quad \alpha' \simeq \frac{\overline{HI}}{\overline{SA'}}$$

d'où le résultat avec  $\alpha + \alpha' = 2\beta$ .

### 1.5.2 Discussion

Si on calcule en faisant un développement limité, on se ramène à montrer que les quantités suivantes sont "égales" :

$$4\beta^2 - 2\alpha^2\beta^2 + 4\alpha\beta^3 \simeq 4\beta^2 + 4\beta^4 + 2\alpha\beta^5 - \alpha^2\beta^4.$$

On voit que l'erreur est de l'ordre de  $\beta^4$  donc petite si  $\beta$  l'est. D'ailleurs, avec  $\beta \simeq 20^\circ$  les deux membres de l'égalité sont 0,14893 et 0,14851 et avec  $\beta = 10^\circ$ , le premier devient 0,14861.

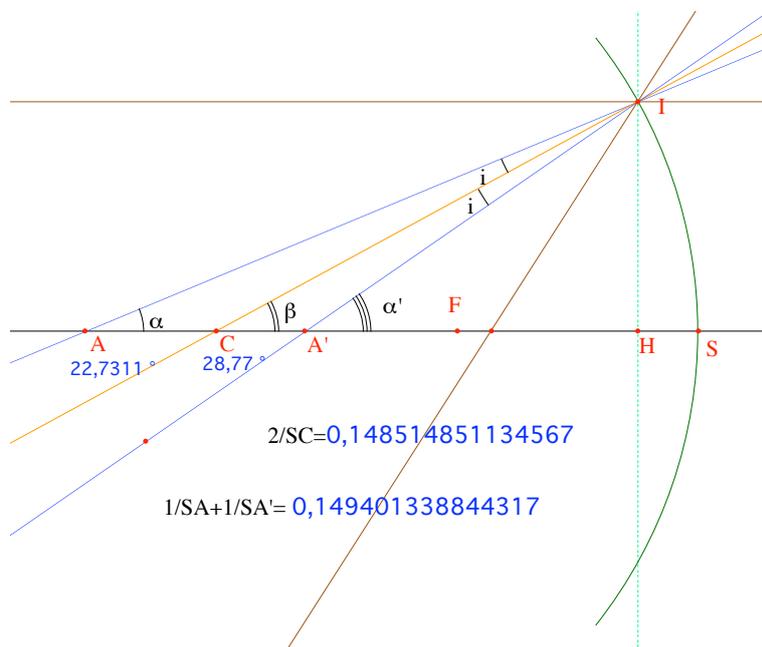


FIGURE 1 – Le miroir sphérique

### 1.5.3 Les fonctions

Pour un système optique général, on appelle **foyer image**  $F'$  l'image d'un point  $A$  situé à l'infini et **foyer objet**  $F$  le point dont l'image est à l'infini. Dans le cas d'un miroir sphérique, la formule de conjugaison montre que ces deux foyers sont égaux et qu'on a  $\overline{SC} = 2\overline{SF}$  autrement dit,  $F$  est le milieu de  $[SC]$ . Avec  $F$ , la formule de conjugaison s'écrit :

$$\frac{1}{\overline{OA'}} + \frac{1}{\overline{OA}} = \frac{1}{\overline{SF}}$$

Dans un système comportant une lentille au lieu d'un miroir, les deux foyers sont distincts et la formule de conjugaison devient :

$$\frac{1}{\overline{OA'}} - \frac{1}{\overline{OA}} = \frac{1}{\overline{OF'}}$$

où  $O$  est le centre de la lentille.

Les physiciens introduisent ici les notations suivantes  $p = \overline{SA}$  (ou  $\overline{OA}$ ),  $p' = \overline{SA'}$  (ou  $\overline{OA'}$ ),  $f = \overline{OF}$  et  $f' = \overline{OF'}$ . Les formules ci-dessus deviennent alors :

$$\frac{1}{p'} - \frac{1}{p} = \frac{1}{f'} \text{ (lentille),} \quad \frac{1}{p'} + \frac{1}{p} = \frac{1}{f'} \text{ (miroir).}$$

On est ici exactement dans la situation évoquée au début : plusieurs paramètres et une relation. D'ailleurs, bon nombre d'exercices du domaine consistent seulement à calculer un paramètre à partir de deux autres supposés connus.

On peut cependant facilement transformer la relation en une fonction :  $p'$  est fonction de  $p$ , avec le paramètre  $f'$  et c'est une fonction homographique que je noterai plutôt avec les variables  $x, y$  (en prévision de leur dérivation) :

$$y = \frac{f'x}{x - f'} \text{ (lentille),} \quad y = \frac{f'x}{x + f'} \text{ (miroir).}$$

Le gain, à la fois mathématique et physique, c'est qu'on a aussitôt le sens de variation de ces fonctions (croissante pour les lentilles, qu'elles soient convergentes ( $f' > 0$ ) ou divergentes ( $f' < 0$ ) et décroissantes pour les miroirs, qu'ils soient concaves ( $f' > 0$ ) ou convexes ( $f' < 0$ ), les courbes représentatives étant, dans tous les cas, des hyperboles équilatères.

On en déduit par exemple, dans le cas d'une lentille, convergente ou divergente, que *si l'on déplace l'objet de gauche à droite, l'image constituée par une lentille en fait autant*, ce qui ne semble pas évident, même pour les experts, surtout dans le cas divergent (voir le livre de Laurence Viennot *En physique : pour comprendre*).

**1.2 Remarque.** Il y a une écriture encore plus simple qui consiste à mettre l'origine en  $F$  pour les objets et  $F'$  pour les images et on a alors les formules  $\overline{FA} \overline{F'A'} = -f'^2$  pour les lentilles et  $\overline{FA} \overline{F'A'} = f'^2$  pour les miroirs. On voit encore plus clairement que les courbes de dépendance de  $A'$  en fonction de  $A$  sont des hyperboles.

## 1.6 Suite de la discussion sur l'origine des fonctions

Les fonctions proviennent parfois de manière naturelle de relations, de principes, de formules diverses, comme on l'a vu ci-dessus. On a vu aussi le cas de la proportionnalité, ou de l'approximation linéaire du premier ordre. On verra dans les chapitres suivants comment l'analyse infinitésimale du phénomène conduit en général à une équation différentielle.

Cependant, dans beaucoup de situations, on a des paramètres, on pressent qu'il y a une relation entre eux, mais on n'a pas d'éléments clairs pour établir une telle relation. On va donc utiliser des méthodes d'interpolation, soit algébriques, soit statistiques. C'est l'objet du paragraphe suivant.

## 2 Interpolation et régression

Lorsqu'il n'y a pas, de manière naturelle, une fonction qui se dégage, on peut essayer de construire une telle fonction qui approche<sup>6</sup> les données. Il y a pour cela deux grandes voies.

### 2.1 Interpolation polynomiale

#### 2.1.1 Le polynôme de Lagrange

La question est la suivante. On a deux paramètres, disons  $a, b$ , dont on connaît  $n+1$  valeurs  $(a_0, b_0), \dots, (a_n, b_n)$  et on cherche une fonction qui vérifie  $f(a_i) = b_i$  pour  $i = 0, \dots, n$ . On cherche *a priori* une fonction  $f$  la plus simple possible, par exemple polynomiale. Le résultat est le suivant :

**2.1 Théorème.** *Soient  $a_0, \dots, a_n$  des réels distincts et  $b_0, \dots, b_n$  des réels quelconques. Il existe un unique polynôme  $P$  de degré  $n$  qui vérifie  $P(a_k) = b_k$  pour tout  $k$ . Ce polynôme est donné par la formule (polynôme d'interpolation de Lagrange<sup>7</sup>) :*

$$P(x) = \sum_{i=0}^n \frac{(x - a_0) \dots \widehat{(x - a_i)} \dots (x - a_n)}{(a_i - a_0) \dots \widehat{(a_i - a_i)} \dots (a_i - a_n)} b_i,$$

où le chapeau signifie que le terme correspondant est omis.

*Démonstration.* L'existence et l'unicité du polynôme sont évidentes pour qui connaît un peu d'algèbre linéaire. En effet, l'application qui à un polynôme  $P$  de degré  $\leq n$  associe  $(P(a_0), \dots, P(a_n))$  est définie sur l'espace vectoriel  $\mathcal{P}_n$  des polynômes de degré  $\leq n$ , qui est de dimension  $n+1$ , elle est à valeurs dans  $\mathbf{R}^{n+1}$  et injective (car un polynôme de degré  $\leq n$  qui admet  $n+1$  racines est nul). Elle est donc surjective, d'où l'existence et l'unicité du polynôme. Une variante plus élémentaire de ce raisonnement consiste à chercher  $P$  sous la forme  $P(x) = \alpha_n x^n + \dots + \alpha_0$  et à résoudre le système correspondant en les  $\alpha_i$ . Il est de Cramer car son déterminant est un Van der Monde. On obtient ainsi l'existence et l'unicité, mais pas la formule explicite.

Pour trouver cette dernière, l'idée est de fabriquer d'abord un polynôme nul en tous les  $a_i$  sauf en  $a_k$  (c'est facile!), de lui imposer ensuite la valeur en  $a_k$  en multipliant par un scalaire et d'ajouter le tout.

---

6. Cela pose tout de même ensuite le problème de la fiabilité des prévisions, notamment en économie. Lorsqu'il n'y a pas de raison objective montrant que tel paramètre dépend de tel autre avec telle loi, les prévisions sont sujettes à caution. En un mot : fonction n'implique pas causalité.

7. Ces questions ont été abordées de manière indépendante par Lagrange et Van der Monde en 1770 à propos de la résolution par radicaux des équations algébriques.

## 2.2 Remarques.

- 1) On peut aussi imposer des conditions sur les dérivées, par exemple les valeurs en  $a_i$  des  $p_i$  premières dérivées.
- 2) Un défaut de cette méthode est qu'on doit recalculer le polynôme si l'on ajoute des points.

### 2.1.2 Un programme sur *xcas*

Par cette méthode, on peut, en particulier, approcher une fonction  $f$  par un polynôme. Voici un programme, sur le logiciel *xcas*, qui calcule le polynôme d'interpolation. On suppose qu'on a défini une fonction  $f$ , on l'approche sur l'intervalle  $[a, b]$ , avec  $n$  pas réguliers. On va donc utiliser les<sup>8</sup>  $a_i = a + i \frac{b-a}{n}$  et les  $b_i = f(a_i)$ . Le programme consiste à calculer d'abord chaque produit  $y_k := \frac{(x - a_0) \dots (x - a_{k-1}) \dots (x - a_n)}{(a_k - a_0) \dots (a_k - a_{k-1}) \dots (a_k - a_n)}$  en partant de 1 et en multipliant récursivement par chaque terme (boucle pour), avec un test pour éviter le cas  $j = k$  (en *xcas*,  $j! = k$  signifie  $j \neq k$ ). On calcule ensuite la somme (qu'on appelle  $z$ ) de ces termes multipliés par  $f(a_k)$  et on obtient  $P(x)$ .

```
interpol(a,b,n,x):={
local j,k,y,z;
z:=0;
pour k de 0 jusque n faire
y:=1;
pour j de 0 jusque n faire
si j!=k alors
y:=y*(x-(a+j*(b-a)/n))/((k-j)*(b-a)/n);
fsi
fpour
z:=z+y*f(a+k*(b-a)/n);
fpour
return z;
}
```

**2.3 Remarque.** Même si l'on multiplie les points il n'est pas sûr que la suite des polynômes de Lagrange converge vers la fonction. Cela peut dépendre du choix des points. Avec des points équidistants sur l'intervalle  $[a, b]$  :  $a_i =$

---

8. Attention, il faut les nommer  $a_j$  ou  $a_k$  car le symbole  $i$  dans *xcas* signifie  $\sqrt{-1}$ .

$a + i\frac{b-a}{n}$ , le polynôme d'interpolation approche très bien<sup>9</sup> la fonction  $\sin x$ , mais il y a divergence dans le cas de  $f(x) = \frac{1}{1+x^2}$  (phénomène de Runge). Avec *xcas* on voit déjà le phénomène en prenant  $f(x) = \frac{1}{x^2 + 0,01}$ ,  $a = -1$ ,  $b = 1$  et  $n = 5$ . Avec  $n = 8$  c'est encore mieux.

## 2.2 Régression linéaire

On part de la même donnée (aux notations près),  $x_1, \dots, x_n$  et  $y_1, \dots, y_n$  et on suppose toujours les  $x_i$  distincts. On a donc un nuage de points. Dans de nombreux cas, au moins lorsque les points du nuage semblent se répartir le long d'une droite, on cherche à l'approcher par une droite. Cette fois, la droite ne passe pas nécessairement par les points, mais elle en est proche au sens des **moindres carrés** :

**2.4 Définition.** Avec les notations précédentes, la droite d'équation  $y = \alpha x + \beta$  est dite approcher le mieux le nuage de points au sens des moindres carrés si la quantité :

$$\varphi(\alpha, \beta) = \sum_{i=1}^n (y_i - (\alpha x_i + \beta))^2$$

est minimum. On l'appelle alors **droite de régression**<sup>10</sup>.

**2.5 Remarque. Attention,** l'approximation ne minimise pas la somme des distances des points du nuage à la droite de régression (on ne minimise que les différences des ordonnées).

### 2.2.1 Étude de $\varphi(\alpha, \beta)$

La fonction  $\varphi(\alpha, \beta)$  est une fonction polynomiale de degré 2 en  $\alpha, \beta$ . Étudions d'abord une telle fonction de manière générale :

### 2.6 Lemme.

1) (Le cas homogène) Soit  $\varphi(\alpha, \beta) = a\alpha^2 + b\alpha\beta + c\beta^2$  une forme quadratique, avec  $a, b, c \in \mathbf{R}$ . Alors  $\varphi$  est nulle au point  $(0, 0)$  et ce point est un minimum<sup>11</sup> de  $\varphi$  si et seulement si on a  $\Delta = b^2 - 4ac < 0$  et  $a, c > 0$ .

2) (Le cas général) Soit  $\varphi(\alpha, \beta) = a\alpha^2 + b\alpha\beta + c\beta^2 + d\alpha + e\beta + f$  avec

9. Au moins sur l'intervalle, car on ne peut empêcher un polynôme de tendre vers l'infini à l'infini.

10. Le mot remonte à Francis Galton (1822-1911).

11. Un minimum strict :  $\varphi$  est  $> 0$  en dehors de  $(0, 0)$ .

$a, \dots, f \in \mathbf{R}$  un polynôme de degré 2. Alors  $\varphi$  admet un minimum strict sur  $\mathbf{R}^2$  si et seulement si on a  $\Delta = b^2 - 4ac < 0$  et  $a, c > 0$ . Le point où le minimum est atteint est donné par

$$\alpha = \frac{2cd - be}{\Delta} \quad \beta = \frac{2ae - bd}{\Delta}.$$

*Démonstration.* 1) Il suffit de considérer les fonctions du second degré en  $\alpha/\beta$  et  $\beta/\alpha$  associées à  $\varphi$ .

2) On sait qu'une condition nécessaire pour que  $(\alpha_0, \beta_0)$  soit un minimum relatif est que les dérivées partielles s'annulent en ce point. Ici, on a  $\frac{\partial \varphi}{\partial \alpha}(\alpha, \beta) = 2a\alpha + b\beta + d$  et  $\frac{\partial \varphi}{\partial \beta}(\alpha, \beta) = b\alpha + 2c\beta + e$ . En résolvant les équations correspondantes on trouve le point annoncé.

Pour voir si le point est effectivement un minimum, on applique la formule de Taylor (qui est exacte pour un polynôme de degré 2 :

$$\begin{aligned} \varphi(\alpha, \beta) &= \varphi(\alpha_0, \beta_0) + (\alpha - \alpha_0) \frac{\partial \varphi}{\partial \alpha}(\alpha_0, \beta_0) + (\beta - \beta_0) \frac{\partial \varphi}{\partial \beta}(\alpha_0, \beta_0) \\ &+ \frac{1}{2} \left( (\alpha - \alpha_0)^2 \frac{\partial^2 \varphi}{\partial \alpha^2}(\alpha_0, \beta_0) + 2(\alpha - \alpha_0)(\beta - \beta_0) \frac{\partial^2 \varphi}{\partial \alpha \partial \beta} + (\beta - \beta_0)^2 \frac{\partial^2 \varphi}{\partial \beta^2} \right). \end{aligned}$$

Le terme de degré 1 disparaît et le terme de degré 2 est exactement la forme quadratique (en  $\alpha - \alpha_0$  et  $\beta - \beta_0$ ) rencontrée en 1) d'où le résultat.

**2.7 Remarque.** On notera que le terme  $f$  n'intervient pas (c'est normal, il n'introduit qu'une translation du résultat).

### 2.2.2 Application

Avec les notations précédentes, on a  $a = \sum_{i=1}^n x_i^2$ ,  $b = 2 \sum_{i=1}^n x_i$ ,  $c = n$ ,  $d = -2 \sum_{i=1}^n x_i y_i$ ,  $e = -2 \sum_{i=1}^n y_i$  et  $f = \sum_{i=1}^n y_i^2$ .

Il est commode d'introduire de nouvelles notions : les moyennes, variances, covariances :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad v(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2, \quad \text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}).$$

On a alors  $a = nv(x) + n\bar{x}^2$ ,  $b = 2n\bar{x}$ ,  $c = n$ ,  $d = -2n(\text{cov}(x, y) + \bar{x}\bar{y})$ ,  $e = -2n\bar{y}$ . Les nombres  $a, c$  sont  $> 0$  et on a  $\Delta = -4n^2v(x)$  et ce nombre est  $< 0$  car les  $x_i$  ne sont pas tous égaux.

On en déduit que la droite de régression existe et que ses coefficients sont donnés par :

$$\alpha = \frac{\text{cov}(x, y)}{v(x)} \quad \text{et} \quad \beta = \bar{y} - \frac{\text{cov}(x, y)}{v(x)} \bar{x}.$$

**2.8 Remarque.** On note que la droite de régression passe par le point moyen  $G = (\bar{x}, \bar{y})$ . Elle est donc déterminée par sa pente  $\alpha = \frac{\text{cov}(x, y)}{v(x)}$  et par le fait qu'elle passe par  $G$  : son équation s'écrivant  $Y - \bar{y} = \alpha(X - \bar{x})$ .

On peut d'ailleurs faire le calcul en prenant  $G$  comme origine. On a  $\varphi(\alpha, \beta) = nv(x)\alpha^2 + n\beta^2 - 2ncov(x, y) + f$  et on étudie d'abord cette fonction en  $\alpha$ , on voit qu'elle admet un minimum par rapport à  $\alpha$  si  $\alpha$  vaut  $\frac{\text{cov}(x, y)}{v(x)}$ . Ce minimum étant  $n\beta^2$ , il est minimum pour  $\beta = 0$ , c'est-à-dire quand la droite passe par l'origine  $G$ .

## 2.3 Les autres régressions

Lorsque les points n'ont pas l'air de s'agglutiner le long d'une droite, mais au contraire d'une parabole, d'une fonction de degré 3, d'une exponentielle, d'un logarithme, on utilise une méthode analogue de minimisation des carrés des distances pour trouver la meilleure courbe possible. Toutes ces régressions sont implantées sur les ordinateurs et les calculatrices. Dans les cas simples, on se ramène à une régression linéaire. Par exemple, dans le cas d'une régression logarithmique, on approche le nuage de points par une courbe  $y = a + b \ln x$  et il suffit pour cela de faire la régression linéaire avec les points  $(\ln x_i, y_i)$ .

## 3 Comment utiliser les fonctions

Où en sommes nous ?

Nous avons décrit le phénomène étudié par une fonction, voire plusieurs, à une, voire plusieurs variables. Quel est maintenant l'objectif ? Il y en a plusieurs, qui vont largement recouper les outils mathématiques dont nous disposons.

### 3.1 Variation des fonctions

Il s'agit de dire quel est le comportement de la fonction qui décrit le phénomène à partir des données actuelles. C'est la problématique de la **prévision**. Bien entendu, c'est un point absolument fondamental. Il suffit pour le comprendre de penser à la finance, à l'écologie, l'économie, la météorologie, etc. On a vu un exemple de cette utilisation avec l'exemple de l'optique.

Pour cela, on dispose, depuis le XVII-ième siècle, d'un outil mathématique très simple et très efficace : la dérivation, les tableaux de variation, etc.

Un très bon exemple, que nous verrons plus tard, concerne l'équation logistique.

## 3.2 Optimisation

Cela va de pair avec le thème précédent : il s'agit de savoir quelle est la "meilleure" valeur des paramètres (voir le cas de la boîte de conserve). Là encore, la dérivée est essentielle. C'est en général l'un des objectifs principaux de la modélisation.

## 3.3 Résolution d'équations

Souvent, on a besoin, face à une fonction  $f(x)$  de déterminer pour quelle valeur de  $x$  on a  $f(x) = a$ . On va pouvoir déployer l'arsenal des méthodes de résolution d'équations numériques (Newton, etc.). On verra ce type de problèmes quand on parlera de radioactivité, d'élimination des médicaments, etc.

## 3.4 Description d'un mouvement

Là, il s'agit en général d'étudier des représentations paramétriques, données par deux fonctions  $x(t), y(t)$ .

# 4 Exemples

Nous verrons de nombreux autres exemples (radioactivité, chimiothérapie) dans les chapitres sur les équations différentielles.

## 4.1 La fonction de coût

Voici un texte d'exercice extrait d'un manuel de Terminale ES :

*Une entreprise qui fabrique des objets estime que le coût total, en milliers d'euros, de production de  $x$  tonnes d'objets s'exprime en fonction de  $x$  par  $C_T(x) = x^3 - 12x^2 + 60x$ .*

*1) Étudier les variations de la fonction  $C_T$  sur  $[0, +\infty[$  et tracer sa représentation graphique ( $F$ ) (1cm pour une tonne, 1cm pour 20000 euros).*

*2) Le coût moyen de fabrication d'un objet est donné par  $C_M(x) = \frac{C_T(x)}{x}$ .*

a) Soit  $O$  l'origine du repère et  $A$  le point de  $(F)$  d'abscisse  $x$ . Expliquer pourquoi  $C_M(x)$  est égal au coefficient directeur de la droite  $(OA)$ .

b) Exprimer  $C_M(x)$  en fonction de  $x$ , étudier ses variations et tracer sa représentation graphique  $(G)$  (1cm pour une tonne, 1cm pour 10000 euros).

3) On appelle coût marginal de  $x$  le coût de fabrication de la  $x + 1$ -ème tonne. On le note  $C_m(x)$  et on admet qu'on a  $C_m(x) = C'_T(x)$ .

a) Étudier les variations de la fonction  $C_m$  et tracer le graphe  $(H)$  de cette fonction dans le même repère que  $(G)$ .

b) Déterminer l'abscisse  $\alpha$  du point d'intersection des courbes  $(G)$  et  $(H)$ . Que peut-on dire de  $C_M(\alpha)$  ?

4) L'entreprise vend sa production 60000 euros la tonne. On note  $B(x)$  le bénéfice réalisé pour la vente de  $x$  tonnes.

a) Calculer  $B(x)$  et étudier les variations de cette fonction.

b) Pour quelle valeur de  $x$  le bénéfice est-il maximum ? Vérifier que, pour cette valeur, le coût marginal est égal au prix de vente unitaire.

Qu'en pensez vous ?

## 5 Optimisation

Ce paragraphe est le plus important du chapitre. Il s'agit d'utiliser les fonctions pour optimiser les réponses à une question. Nous donnons plusieurs exemples de cette situation, en commençant par un vieux problème.

### 5.1 Le problème d'Euclide

Il s'agit de déterminer les dimensions d'un champ rectangulaire délimité par une clôture de longueur  $2l$  donnée afin qu'il soit d'aire maximum. Il s'agit donc, pour  $b, c \geq 0$ , de maximiser  $bc$  avec la contrainte  $b + c = l$ . Il y a plusieurs méthodes, en voici trois.

#### 5.1.1 Par l'étude des fonctions

On tire  $c$  en fonction de  $b$  :  $c = l - b$  et on étudie  $f(b) = b(l - b) = bl - b^2$ . On a  $f'(b) = l - 2b$  et le maximum de  $f$  est atteint en  $l/2$  (on fait le tableau de variation), ce qui correspond au cas du carré.

Variante : si l'on a l'idée du  $l/2$  on pose  $b = \frac{l}{2} + x$  et  $c = \frac{l}{2} - x$  et on a  $\mathcal{A} = \frac{l^2}{4} - x^2$ , évidemment maximum pour  $x = 0$ .

### 5.1.2 Par l'identité remarquable

Le cas du carré correspond à la longueur  $\frac{b+c}{2}$ . On montre l'inégalité  $\left(\frac{b+c}{2}\right)^2 \geq bc$ . C'est simplement  $(b-c)^2 \geq 0$ .

### 5.1.3 Par la géométrie

On contemple la figure ci-dessous :

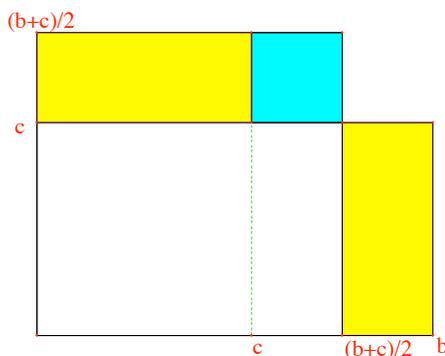


FIGURE 2 – La preuve

### 5.1.4 Par les extrema liés

Si l'on est vraiment pédant, on peut utiliser le théorème 5.3!

## 5.2 Un bel exemple : le problème du maître-nageur

Le problème est le suivant. Le maître-nageur est au bord de la mer. Il aperçoit un baigneur en difficulté dans l'eau et veut lui porter secours. Que doit-il faire pour y parvenir au plus vite (il peut courir sur la plage ou nager) ?

### 5.2.1 Mise en équations

On suppose la côte rectiligne. Le maître-nageur est en un point  $B$ . On suppose que le baigneur est en un point  $A$  et ne bouge pas<sup>12</sup>. On appelle  $H$

---

12. Il peut couler, mais cela n'influe pas sur notre problème ...

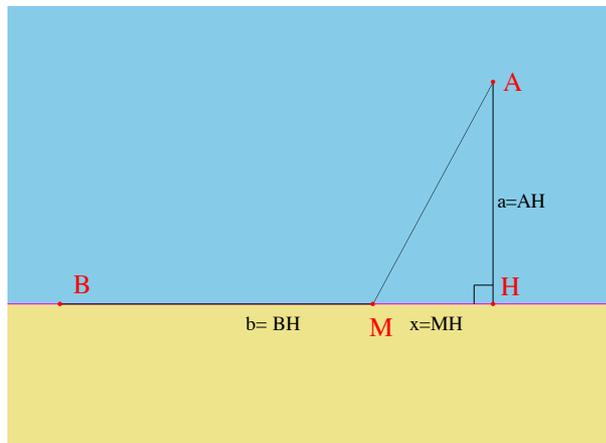


FIGURE 3 – Le maître-nageur et le baigneur

le projeté de  $A$  sur la côte,  $a$  la distance  $AH$ ,  $b$  la distance  $BH$ ,  $w$  la vitesse à pied,  $v$  la vitesse à la nage. Bien sûr, le maître-nageur court plus vite qu'il ne nage, de sorte qu'on a  $w > v$ . Le maître-nageur court jusqu'en  $M$  puis nage de  $M$  à  $A$ .

Il y a plusieurs choix de variables possibles :  $x = BM$ , ou  $x = MH$ , ou  $x = AM$ , ou encore l'angle  $\theta = \widehat{MAH}$ .

On s'aperçoit assez vite que le choix le plus simple est de prendre  $x = MH$ . La fonction à regarder, qui est le temps de parcours, est alors  $f(x) = \frac{b-x}{w} + \frac{\sqrt{a^2+x^2}}{v}$ . Le calcul est simple. On a  $f'(x) = -\frac{1}{w} + \frac{x}{v\sqrt{a^2+x^2}}$  qui admet un zéro pour  $x_0 = \frac{av}{\sqrt{w^2-v^2}}$ . C'est bien un minimum de  $f$  car la dérivée est d'abord négative puis positive. On a donc la solution<sup>13</sup>.

### 5.2.2 Discussion

La bonne question c'est : que doit-on apprendre à l'école des maîtres-nageurs, sachant que, sur place, la personne n'aura pas le temps de faire le calcul<sup>14</sup>? Une première chose à noter est que la distance  $b$  n'intervient pas! Si on admet des vitesses<sup>15</sup>  $v = 5\text{km/h}$  et  $w = 25\text{km/h}$  donc  $w = 5v$ , on

13. Si les vitesses  $v$  et  $w$  sont voisines, il se peut que  $x_0$  soit plus grand que  $b$ . Dans ce cas, le minimum est atteint pour  $x = b$ .

14. À supposer qu'elle sache le faire.

15. Cela fait nager le 100 m en 1 mn 12 et le courir en 14 s 4, ce qui semble raisonnable pour un maître-nageur, qui est un sportif.

trouve  $x_0 = \frac{a}{\sqrt{24}}$  soit environ  $x_0 = a/5$ . Le résultat théorique est donc de courir jusqu'à être environ à  $1/5$  de la distance du nageur au bord.

En fait, on peut facilement calculer  $f(0)$  (le temps qui correspond au cas où l'on court jusqu'en  $H$  puis où l'on nage de  $H$  à  $A$ ). Posons  $w = kv$  avec  $k > 1$ . On trouve  $f(0) = \frac{1}{v} \left( \frac{b}{k} + a \right)$ . On peut aussi calculer  $f(x_0)$  :

$$f(x_0) = \frac{1}{v} \left( \frac{b}{k} + \frac{a\sqrt{k^2 - 1}}{k} \right).$$

Cela permet d'évaluer l'erreur commise en allant en  $H$  plutôt qu'au minimum :

$$f(0) - f(x_0) = \frac{a}{v} \left( 1 - \frac{\sqrt{k^2 - 1}}{k} \right).$$

Avec  $k = 5$ , l'erreur est donc égale à  $0,02 \frac{a}{v}$  et avec  $v = 5km/h$ , donc<sup>16</sup>  $1/v \sim 0,72s/m$ , on voit que l'erreur commise est environ  $0,0145a$ . Pour une distance  $a$  de l'ordre de  $100m$  elle est donc de  $1,45$  seconde : négligeable sur un temps de plus d'une minute.

Une tabulation s'impose pour vérifier les résultats. Supposons  $a = 20m$  et  $b = 100m$ . On a  $v = 50/36 \sim 1,38$  et  $w = 250/36 \sim 6,94$  (en  $m/s$ ). Le minimum de temps est obtenu en courant environ  $96m$ , avec un temps de  $28,5$  secondes. Si on court jusqu'à  $100m$ , le temps est de  $28,8$  secondes : l'erreur est évidemment négligeable.

### 5.2.3 Conclusion

Bref, la règle que l'on doit apprendre aux futurs maîtres-nageurs c'est tout simplement de courir jusqu'à l'aplomb du nageur !

### 5.2.4 Réfraction

Une variante du problème est d'envisager le cas où le maître-nageur n'est pas juste au bord de l'eau. La situation mathématique est alors exactement identique à celle de la réfraction de la lumière. En effet, au lieu d'un maître-nageur, on a alors un rayon lumineux qui franchit une surface limitant deux milieux dans lesquels la vitesse de la lumière est différente. On peut faire le calcul comme ci-dessus, mais il est plus adroit d'utiliser les multiplicateurs de Lagrange. Nous retrouverons ce problème plus loin.

---

16. Attention aux unités!

## 5.3 Cosette et les extrema liés

### 5.3.1 Le problème

Je propose le problème comme une récréation, mais il est tout à fait sérieux.

*Cosette est en vacances à Saint-Tricotin-sur-Pelote (Marne-et-Garonne), mais ce n'est pas pour autant que les Thénardier lui font grâce des corvées. Elle doit partir de la maison avec ses seaux, aller puiser de l'eau dans la Pelote et la porter dans l'abreuvoir des vaches qui est situé plus loin, mais du même côté de la rivière. Aidez-la à trouver le chemin le plus court pour faire son travail.*

*(On distinguera selon que la Pelote est rectiligne, un arc de cercle, ou toute autre courbe.)*

Une variante de ce problème remonte à Héron d'Alexandrie qui souhaite rallier  $A$  à  $B$  au plus vite en allant faire boire son cheval à la rivière.

### 5.3.2 La ruse dans le cas d'une droite

On introduit le point  $A'$  symétrique de  $A$  par rapport à la Pelote, voir figure ci-contre. Le plus court chemin passe par  $M$  intersection de la Pelote avec  $(A'B)$ . En effet, si  $P$  est un autre point on a  $AM + MB = A'M + MB = A'B < AP + PB = AP + PB$  (la ligne droite est le plus court chemin d'un point à un autre).

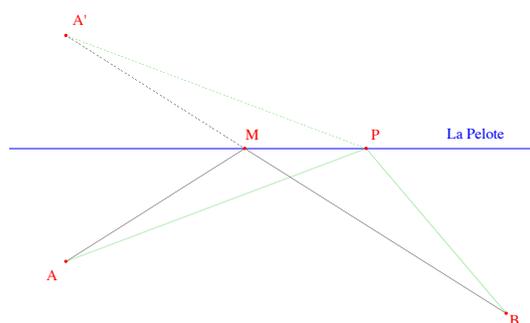


FIGURE 4 – Cosette et la Pelote

Si l'on ne voit pas la ruse on calcule, mais ce n'est pas si facile. On prend la rivière pour axe des  $x$ ,  $A = (0, a)$ , disons  $a > 0$ ,  $B = (c, b)$ , disons  $c, b > 0$ ,  $M = (x, 0)$ . La fonction à minimiser est  $f(x) = \sqrt{x^2 + a^2} + \sqrt{(x - c)^2 + b^2}$ . La dérivée est égale à :

$$f'(x) = \frac{x}{\sqrt{x^2 + a^2}} + \frac{x - c}{\sqrt{(x - c)^2 + b^2}}.$$

Elle s'annule en  $x$  vérifiant  $(a^2 - b^2)x^2 - 2a^2cx + a^2c^2 = 0$ . Le discriminant est  $(abc)^2$ . On obtient deux solutions :  $x = \frac{ac}{a \pm b}$ . La bonne (celle qui correspond

au point symétrique de  $A$ ) correspond au signe  $+$ . Et l'autre ? Elle n'annule pas la dérivée car  $x$  et  $x - c$  doivent être de signes contraires or on a  $x = \frac{ac}{a - b}$  et  $x - c = \frac{bc}{a - b}$ . (Attention à l'élévation au carré).

**5.1 Remarque.** On reconnaît aussi dans cette situation le trajet d'une boule de billard qui rebondit contre une bande ou d'un rayon lumineux qui se réfléchit sur un miroir.

### 5.3.3 Cas général : modélisation

Supposons que la Pelote soit une courbe plane  $P$  d'équation cartésienne  $F(x, y) = 0$  et notons encore  $A$  et  $B$  les points correspondant à la maison et à l'abreuvoir. Soit  $G(x, y)$  la fonction qui associe à un point  $M = (x, y)$  la somme des distances  $MA + MB$ . Si on pose  $A = (a, b)$  et  $B = (c, d)$ , il s'agit de la fonction :

$$G(x, y) = \sqrt{(x - a)^2 + (y - b)^2} + \sqrt{(x - c)^2 + (y - d)^2}.$$

On étudie la restriction à  $P$  de la fonction  $G$  et il s'agit de trouver le minimum de  $G$  sur  $P$ . C'est un problème d'optimisation avec une contrainte qu'on appelle encore problème d'extrema liés.

### 5.3.4 La théorie

Rapelons d'abord que pour une fonction dérivable de deux variables  $F(x, y)$  le gradient au point  $(x, y)$  est le vecteur  $gradF = \left( \frac{\partial F}{\partial x}(x, y), \frac{\partial F}{\partial y}(x, y) \right)$ .

Rappelons un lemme de dérivation :

**5.2 Lemme.** Soit  $F(x, y)$  une fonction de classe  $C^1$  (c'est-à-dire admettant des dérivées partielles continues) et soit  $f(x)$  une fonction dérivable. Alors, la fonction  $u(x) = G(x, f(x))$  est dérivable et on a :

$$u'(x) = \frac{\partial F}{\partial x}(x, f(x)) + f'(x) \frac{\partial F}{\partial y}(x, f(x)).$$

*Démonstration.* Faisons le comme un physicien. On considère un accroissement infinitésimal  $h$  de la variable  $x$ . Par définition de la dérivée, on a  $f(x + h) = f(x) + hf'(x)$  (on néglige ici les infiniment petits d'ordre 2, c'est-à-dire les termes en  $h^2$  et plus). De même, si l'on a des infiniment petits  $h$  et  $k$ , la définition des dérivées partielles donne :

$$F(x + h, y + k) = F(x, y) + h \frac{\partial F}{\partial x}(x, y) + k \frac{\partial F}{\partial y}(x, y).$$

On calcule alors :

$$\begin{aligned} u(x+h) &= F(x+h, f(x+h)) = F(x+h, f(x) + hf'(x)) \\ &= F(x, f(x)) + h \left[ \frac{\partial F}{\partial x}(x, f(x)) + f'(x) \frac{\partial F}{\partial y}(x, f(x)) \right] \end{aligned}$$

et on a le résultat.

**5.3 Théorème.** *Soit  $G$  une fonction de classe  $C^1$  définie sur un ouvert de  $\mathbf{R}^2$  et soit  $P$  une courbe de classe  $C^1$  lisse d'équation  $F(x, y) = 0$ . Si la restriction de  $G$  à  $P$  admet un extremum local, c'est en un point où  $\text{grad } F$  et  $\text{grad } G$  sont colinéaires.*

*Démonstration.* Soit  $M_0 = (x_0, y_0)$  l'extremum local de  $G$  sur  $P$ . Comme la courbe est lisse, le théorème des fonctions implicites montre qu'au voisinage de  $M_0$ ,  $P$  peut s'écrire, disons<sup>17</sup>, sous la forme  $y = f(x)$ . De plus, comme on a identiquement  $F(x, f(x)) = 0$ , on en déduit la formule (c'est le lemme) :

$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))}.$$

La fonction à étudier est  $g(x) = G(x, f(x))$  et on a (encore le lemme) :

$$g'(x) = \frac{\partial G}{\partial x}(x, f(x)) + f'(x) \frac{\partial G}{\partial y}(x, f(x)).$$

Si  $g$  admet un extremum local en  $x_0$ , c'est que  $g'(x_0)$  est nul et on en déduit la relation (en  $x_0$ ) :

$$\frac{\partial G}{\partial x}(x, f(x)) \frac{\partial F}{\partial y}(x, f(x)) - \frac{\partial G}{\partial y}(x, f(x)) \frac{\partial F}{\partial x}(x, f(x)) = 0,$$

qui montre que les gradients sont colinéaires.

### 5.3.5 Explication géométrique : le lien gradient-normale

Le résultat précédent est très naturel si l'on pense géométriquement en termes de gradients et de courbes de niveau. On a une fonction  $F(x, y)$  définie sur le plan. Un exemple emblématique de cette situation est le cas où  $F(x, y)$  est l'altitude du point de la terre de coordonnées  $(x, y)$ . En un tel point il y a deux objets géométriques :

---

17. C'est cela ou  $x = f(y)$  et cela dépend si c'est  $\frac{\partial F}{\partial y}(x_0, y_0)$  qui est non nul, comme ici, ou  $\frac{\partial F}{\partial x}(x_0, y_0)$ .

- La courbe de niveau  $F(x, y) = k$  qui passe par ce point.
- Le gradient de  $F$  en ce point qui définit la direction de la ligne de plus grande pente passant par ce point.

Le lien entre les deux est simple : le gradient est la direction normale à la ligne de niveau (la ligne de niveau c'est là où  $F$  ne varie pas, le gradient c'est la direction où elle varie le plus). C'est ce que dit la proposition suivante :

**5.4 Proposition.** *Soit  $F(x, y) = k$  une courbe de niveau de la fonction  $F$ . La normale en un point  $M$  de cette courbe a pour direction le vecteur  $\text{grad}F$ .*

*Démonstration.* On écrit la courbe sous la forme  $y = f(x)$  au voisinage de  $M$  et le vecteur directeur de la tangente est alors  $(1, f'(x))$ , colinéaire à  $(-\frac{\partial F}{\partial y}(x, f(x)), \frac{\partial F}{\partial x}(x, f(x)))$ . On voit que ce vecteur est orthogonal à  $\text{grad}F$ .

Si maintenant on pense à la fonction  $G$  que l'on étudie le long d'une courbe de niveau de  $F$ , dire qu'elle a un extremum local en un point, c'est dire que sur la courbe  $F = k$  elle ne varie pas, ou très peu, au voisinage de ce point, autrement dit que son gradient n'a pas de composante le long de la ligne de niveau de  $F$ , donc qu'il lui est orthogonal, donc colinéaire au gradient de  $F$ .

### 5.3.6 Rappels sur les ellipses

Revenons à la fonction  $G(M) = MA + MB$ . On connaît bien<sup>18</sup> ses courbes de niveau, ce sont les ellipses de foyers  $A, B$ . Concernant ces ellipses, on a le résultat suivant (sans doute dû à Poncelet) :

**5.5 Proposition.** *Si  $E$  est une ellipse de foyers  $A, B$  et si  $M$  est un point de  $E$ , la normale à  $E$  en  $M$  est la bissectrice de l'angle  $\widehat{AMB}$ .*

*Démonstration.* Il y en a plusieurs. On peut par exemple calculer à partir d'une représentation paramétrique. On écrit l'ellipse sous la forme  $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ , avec  $a > b$ . Ses foyers sont les points  $(\pm c, 0)$  avec  $c^2 = a^2 - b^2$ . On a la représentation paramétrique  $x = a \cos \varphi$ ,  $y = b \sin \varphi$ . On en déduit le vecteur tangent en  $M = (a \cos \varphi, b \sin \varphi)$  en calculant  $x' = -a \sin \varphi$  et  $y' = b \cos \varphi$ . Le vecteur normal est alors  $(b \cos \varphi, a \sin \varphi)$ . Pour voir qu'il est porté par la bissectrice de  $\widehat{AMB}$  on montre qu'il est colinéaire à  $\frac{\overrightarrow{AM}}{AM} + \frac{\overrightarrow{BM}}{BM}$ . On calcule facilement  $\overrightarrow{AM} = (a \cos \varphi + c, b \sin \varphi)$  et  $\overrightarrow{BM} = (a \cos \varphi - c, b \sin \varphi)$ .

18. Enfin, si l'on suit le jury de CAPES, on ne les connaîtra bientôt plus.

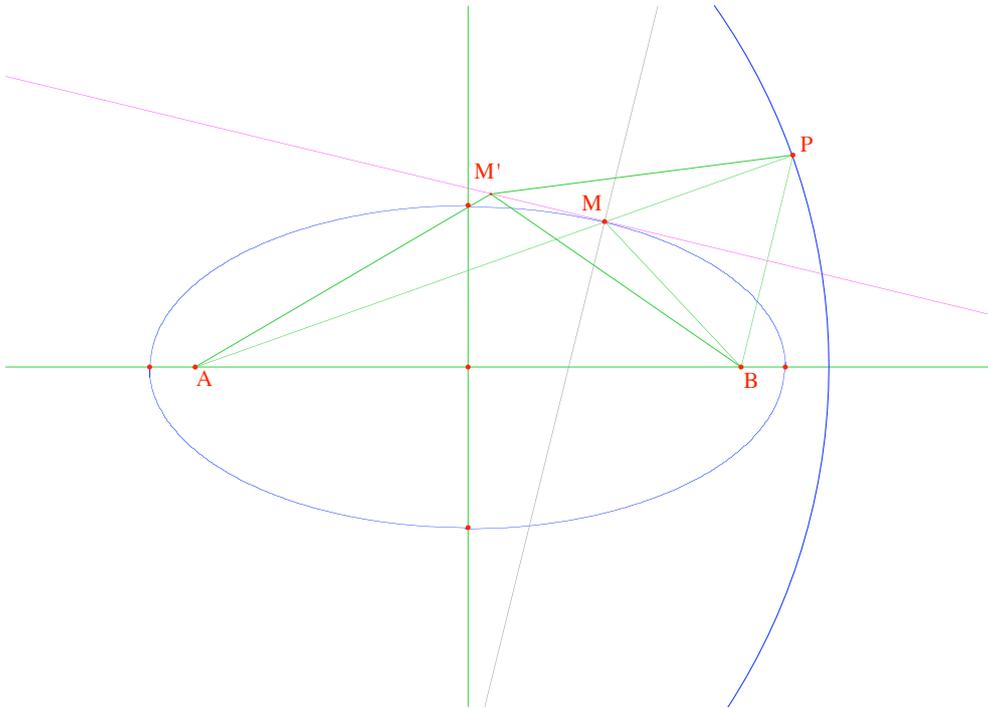


FIGURE 5 – Normale et bissectrice

Le reste du calcul n'est pas difficile et le lecteur le fera à titre d'exercice. On conseille de garder  $AM$  et  $BM$  sans les calculer mais en tenant compte de  $AM + BM = 2a$  et de  $AM^2 - BM^2 = (AM + BM)(AM - BM) = 4ac \cos \varphi$ .

Le calcul n'est pas très difficile, mais la géométrie c'est bien mieux, comme on va le voir maintenant !

Soit  $M$  un point de  $E$ . On trace le cercle directeur associé au foyer  $A$ , c'est-à-dire le cercle de centre  $A$  et de rayon  $2a$ . La demi-droite  $[AM)$  coupe ce cercle en  $P$  et  $M$  est sur la médiatrice  $D$  de  $[PB]$ . En effet, on a  $MA + MB = MA + MP = 2a$ . De plus,  $D$  est la tangente à l'ellipse en  $M$ . En effet, si  $M'$  est sur  $D$  et différent de  $M$ , on a  $M'A + M'B = M'A + M'P > AP = 2a$ , de sorte que  $D$  ne coupe l'ellipse qu'en un point donc est tangente.

Mais  $D$  est bissectrice de  $\widehat{PMB}$  donc bissectrice extérieure de  $\widehat{AMB}$  et la normale est la bissectrice intérieure.

### 5.3.7 Conclusion

Si un point  $M$  de la courbe  $F(x, y) = 0$  réalise un minimum local de  $G(M) = MA + MB$ , les gradients de  $F$  et  $G$  en  $M$  sont colinéaires, donc la normale à la courbe  $F(x, y) = 0$  en  $M$  est la bissectrice de  $\widehat{AMB}$ .

Cela permet de traiter le cas de la droite, on retrouve le fait que le plus court chemin passe par le point d'intersection de la Pelote avec  $A'B$  (il faut que la Pelote soit bissectrice extérieure de  $\widehat{AMB}$ ).

Cela permet aussi de traiter graphiquement le cas où la Pelote est circulaire<sup>19</sup> voir figures ci-dessous. Les directions des gradients des deux fonctions ( $MA + MB$  et l'équation du cercle) apparaissent respectivement en marron et rose et les extrema correspondent au cas où ces droites sont confondues, c'est-à-dire au cas où la bissectrice de l'angle  $\widehat{AMB}$  passe par le centre du cercle. On notera qu'il y a deux solutions correspondant au minimum et au maximum de  $MA + MB$ . Une question naturelle est de construire les points réalisant les extrema à la règle et au compas. On peut montrer que c'est impossible, voir annexe.

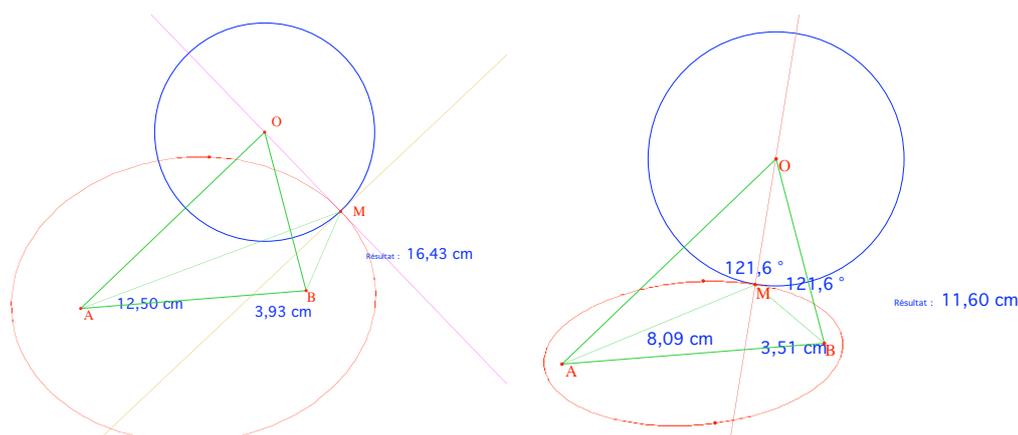


FIGURE 6 – Cosette et la Pelote circulaire, le cas générique et le cas minimal

### 5.3.8 Discussion

Si on réfléchit bien, la solution au problème, tel que je l'ai posé, n'est pas pertinente. En effet, c'est clair que Cosette a tout intérêt à aller au point de la Pelote le plus proche de  $B$  (ce qui est un autre problème). En effet, elle part avec les seaux vides et ce qu'il est important de minimiser c'est le trajet avec les seaux pleins!

Ou alors il faut imaginer une variante (c'est l'avantage des récréations). Par exemple, Cosette doit aller jeter les eaux sales de la maison  $A$  dans la

<sup>19</sup>. Disons un arc de cercle, parce qu'une rivière circulaire, c'est comme une fourmi parlant français, latin et javanais, ça n'existe pas.



de nouveau face à un problème d'extrema liés. Si l'on écrit que les gradients de  $G$  et  $F$  sont colinéaires on trouve  $\frac{\cos \alpha}{v} = \frac{\cos \beta}{w}$ .

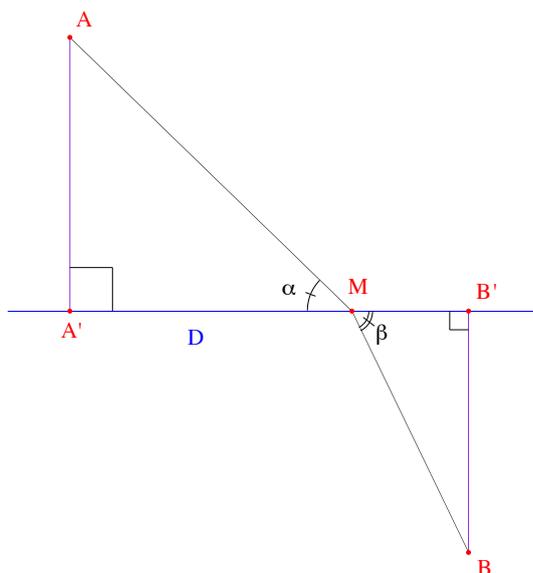


FIGURE 8 – Réfraction

C'est la fameuse relation de Snell-Descartes que l'on formule en général avec les angles que fait le rayon avec les normales à  $D$  c'est à dire  $i = \pi/2 - \alpha$  et  $j = \pi/2 - \beta$ . La relation devient  $\frac{\sin i}{v} = \frac{\sin j}{w}$ .

## 5.5 Annexe : retour sur la Pelote circulaire

Je donne ici une indication très rapide<sup>22</sup> de la solution du problème posé ci-dessus : dans le cas où la Pelote est circulaire, peut-on construire à la règle et au compas les extrema de la fonction  $MA + MB$ . Nous allons voir que la réponse est négative.

On note d'abord que la condition d'égalité des angles  $\widehat{AMO}$  et  $\widehat{BMO}$  se traduit par la relation (\*) :  $MB(\overrightarrow{MO}|\overrightarrow{MA}) = MA(\overrightarrow{MO}|\overrightarrow{MB})$  (en effet, on a  $(\overrightarrow{MO}|\overrightarrow{MA}) = MO.MA.\cos \widehat{AMO}$ ).

On fait le calcul en prenant  $A = (-a, 0)$ ,  $B = (a, 0)$ ,  $O = (b, c)$ . Le cercle  $C$  a pour équation :

$$x^2 + y^2 - 2bx - 2cy + b^2 + c^2 - r^2 = 0.$$

22. Le lecteur est averti que la solution fait appel à des connaissances plus avancées que le reste du cours.

On trouve l'équation suivante pour la relation (\*) (en enlevant le facteur  $4ay$  qui correspond à  $(AB)$ , c'est normal de trouver cette droite dans le lieu) :

$$(cx - by)(x^2 + y^2) - bcx^2 + (a^2 + b^2 - c^2)xy + bcy^2 - a^2cx - a^2by + a^2bc = 0.$$

Il s'agit d'une cubique, qui a six points d'intersection avec le cercle, dont deux sont les points cycliques  $(1, \pm i, 0)$  (en projectif). On élimine ces points en remplaçant dans l'équation de la cubique  $x^2 + y^2$  par sa valeur extraite de l'équation du cercle. Il reste à calculer l'intersection de deux coniques, le cercle et la conique  $\Gamma$  :

$$bcx^2 - bcy^2 + (a^2 + c^2 - b^2)xy + cx(r^2 - b^2 - c^2 - a^2) + by(b^2 + bc - r^2 - a^2) + a^2bc = 0.$$

La question est de savoir si les points d'intersection sont constructibles. Pour voir cela, il suffit de regarder l'équation du troisième degré qui détermine les coniques dégénérées du pinceau, ce qu'on fait en écrivant que le déterminant de la matrice de  $\lambda C + \Gamma$  est nul. Si elle est irréductible, les intersections ne sont pas constructibles. Il suffit même de montrer qu'elle l'est pour une valeur particulière de  $a, b, c, r$ . Si je les prends tous égaux à 1, je trouve :  $4x^3 + 16x^2 + 21x - 1 = 0$ , qui est irréductible sur  $\mathbf{Q}$  (on vérifie qu'elle n'a pas de racine rationnelle). Les extrema ne sont donc pas constructibles car leurs coordonnées contiennent les racines d'un polynôme de degré 3.

## 6 Exercices

### 6.1 Discret-continu : les suites

Pour terminer ce chapitre et aborder le côté discret des modélisations, je propose l'exercice suivant, inspiré de T. Lambre, *L'épreuve sur dossier à l'oral du CAPES*, p. 156. Il s'agit de banque et d'intérêts.

Dans les discussions on prendra en compte deux situations :

- Le placement : un épargnant dépose une certaine somme d'argent à la banque dans le but de la faire fructifier.
- L'emprunt : un particulier emprunte de l'argent à la banque pour effectuer, par exemple, un achat immobilier, qu'il devra rembourser ensuite avec des intérêts.

Dans les deux cas, on se préoccupera de l'intérêt du particulier, pas des banquiers !

### 6.1.1 Intérêts composés

Dire qu'un capital  $C$  est placé avec un intérêt<sup>23</sup>  $i$  l'an selon le principe des intérêts composés signifie qu'il est multiplié chaque année par  $1 + i$ .

1) Calculer le capital  $C_n$  au bout de  $n$  années.

2) On peut aussi placer le capital à un taux mensuel  $j$ , ce qui signifie qu'il est multiplié chaque mois par  $1 + j$ . Calculer le taux mensuel mathématique  $i_m$  qui donne le même capital au bout d'un an que le taux annuel  $i$ .

3) Cette fois, au lieu de couper en 12, on coupe en  $d$ , où  $d$  est un entier  $> 0$ , et on multiplie le capital par  $1 + j$  toutes les  $d$  parties de l'année<sup>24</sup>. Déterminer le taux mathématique  $i_d$  qui donne le même capital au bout d'un an que le taux annuel  $i$ . Calculer le capital obtenu avec cette méthode au bout de  $p$  fractions d'année.

**Réponses :** 1)  $C_n = C(1 + i)^n$

2)  $i_m = (1 + i)^{\frac{1}{12}} - 1$ .

3)  $i_d = (1 + i)^{\frac{1}{d}} - 1$ . Au bout de  $p$  fractions d'année :  $C(1 + i)^{\frac{p}{d}}$ .

### 6.1.2 Taux proportionnel

4) Toujours dans le cas d'un fractionnement de l'année en  $d$  parties, les banquiers proposent parfois un taux dit proportionnel de  $j_d = i/d$ . Cela signifie que, chaque fraction d'année, le capital est multiplié par  $1 + j_d = 1 + \frac{i}{d}$ . Comparer les deux systèmes.

5) Pour un placement d'un nombre non entier d'années, on combine calcul avec le taux  $i$  pour les années entières et taux proportionnel pour l'année non achevée. Calculer le capital obtenu au bout de  $n$  années et  $p$  fractions ( $p < d$ ). Comparer avec le calcul utilisant le taux mathématique.

6) Pour un capital placé, au taux annuel  $i$ , pendant un an trois mois et dix-sept jours, calculer le capital acquis selon le fractionnement en année, mois, quinzaine ou jours<sup>25</sup>. On utilisera la méthode de la question 5) d'usage des taux proportionnels seulement pour la fraction d'année.

**Réponse :** 4) Il s'agit de comparer le taux mathématique  $1 + i_d = (1 + i)^{\frac{1}{d}}$  et le taux proportionnel  $1 + j_d = 1 + \frac{i}{d}$ . Comme  $d$  est entier, on peut élever les deux termes à la puissance  $d$  et il reste à comparer  $1 + i$  et  $(1 + \frac{i}{d})^d$ . Le

---

23. On dit aussi placé à  $100i\%$ , par exemple  $i = 0,05$  correspond à 5%.

24. Dans la pratique, pour les particuliers, le seul autre cas est 24, par quinzaine. En revanche, pour les opérations entre banques on pratique un taux journalier, avec  $d = 360$ .

25. Bien entendu, si le fractionnement est, par exemple, par année, seules les années entières sont prises en compte.

développement du binôme montre que le second terme est plus grand. On a donc  $i_d \leq j_d$  avec égalité seulement pour  $d = 1$  ou  $i = 0$ .

La réponse à donner au banquier qui propose le taux proportionnel est donc oui s'il s'agit d'un placement et non s'il s'agit d'un emprunt.

5)  $C(1+i)^n(1+\frac{i}{d})^p$  avec le taux proportionnel pour l'année écoulée,  $C(1+i)^n(1+i)^{\frac{p}{d}}$  avec le taux mathématique.

Cette méthode minimise les pertes du banquier qui ne propose un taux désavantageux pour lui que sur la dernière année.

6) Je trouve respectivement  $1+i$ ,  $(1+i)(1+\frac{i}{12})^3$ ,  $(1+i)(1+\frac{i}{24})^7$  et enfin  $(1+i)(1+\frac{i}{360})^{107}$ .

### 6.1.3 Intérêts simples

*Pour un placement à court terme, les banquiers proposent parfois un placement avec intérêts simples, toujours au taux proportionnel de  $i/12$ . Cela signifie que chaque mois, le capital  $C$  est augmenté<sup>26</sup> de  $Ci/12$ .*

*7) Calculer le capital obtenu au bout de  $p$  mois d'abord avec les intérêts composés (dans les deux cas vus ci-dessus des taux mensuels  $i_{12}$  et  $j_{12}$ ), puis avec les intérêts simples. Comparer les résultats selon le temps de placement. Envisager ensuite la même situation avec un fractionnement en  $d$  parties.*

#### Réponses :

7) On a  $C_p = C(1+i)^{p/12}$  (composés avec  $i_{12}$ ),  $D_p = C(1+\frac{i}{12})^p$  (composés avec  $j_{12}$ ) et  $S_p = C(1+\frac{pi}{12})$  (simples). On a vu l'inégalité  $C_p \leq D_p$ . La formule du binôme montre aussi l'inégalité  $S_p \leq D_p$ . Il reste à comparer  $C_p$  et  $S_p$ .

On passe au logarithme et on compare  $\ln C_p = \frac{p}{12} \ln(1+i)$  et  $\ln S_p = \ln(1+\frac{pi}{12})$ . On fixe  $i$  et on étudie la fonction différence comme fonction de  $p$ . On voit qu'elle vaut 0 pour  $p = 0$ , qu'elle décroît, puis croît et tend vers  $+\infty$ . Elle s'annule pour la valeur  $p = 12$ . Autrement dit, pour moins de 12 mois, le système à intérêts simples est plus avantageux.

On peut aussi vérifier cela en faisant un développement limité en  $i$  de  $(1+i)^{p/12} - (1+\frac{pi}{12})$ . On trouve  $\frac{p(p-12)}{288}i^2 + o(i^2)$ . Avec un taux de 0,05 et une durée de 6 mois, cela donne un écart de l'ordre de  $3 \times 10^{-4}$ . Ce n'est pas considérable, sauf si l'on place de très grosses sommes ...

En résumé, pour un placement, le système à intérêts composés avec le taux proportionnel est toujours plus avantageux que les deux autres. Le système à intérêts simples est plus avantageux que le système à intérêts composés avec le taux mathématique lorsque la durée du placement est de moins d'un an.

<sup>26</sup>. Ici, contrairement au cas des intérêts composés, le capital  $C$  n'est pas actualisé à chaque opération.

Les réponses sont les mêmes pour un fractionnement en  $d$  parties.

#### 6.1.4 Taux différentiel

8) Lorsque l'année est fractionnée en  $d$  parties et qu'on utilise le taux proportionnel, le capital est multiplié au bout d'une année par  $(1 + \frac{i}{d})^d$ , voir question 4).

Pour les opérations de bourse qui ont lieu en temps réel et parfois à la seconde près, on définit le taux différentiel associé à  $i$  qui doit permettre de calculer le capital au temps  $t$  (exprimé en années, mais avec  $t$  réel).

Justifier la définition du taux différentiel par  $i_\infty = \exp(i) - 1$ . Avec cette définition, par quelle quantité est multiplié le capital au bout de  $t$  années ( $t \in \mathbf{R}$ ) ?

On suppose que le taux annuel est de 5% et que le boursicoteur place un capital d'un million d'euros. Quel est l'écart entre le capital obtenu avec le taux annuel normal et le taux différentiel ?

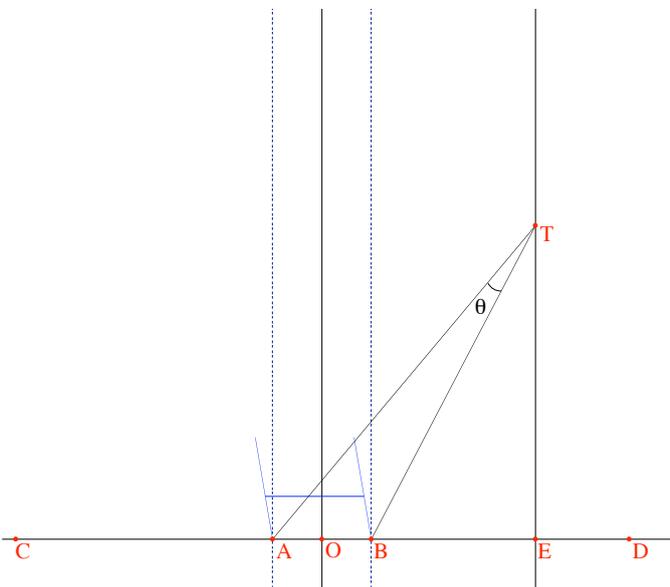
**Réponses :** On a  $1 + i_\infty = \lim_{d \rightarrow +\infty} (1 + \frac{i}{d})^d$  : le taux différentiel est la limite du taux proportionnel relatif à  $d$  quand on subdivise l'année de manière de plus en plus fine. Le capital est alors multiplié par  $e^{it}$  au temps  $t$ .

Avec l'exemple numérique, l'écart est de 1271 euros en faveur du taux différentiel.

## 6.2 Le joueur de rugby

Au rugby, quand une équipe marque un essai en un point  $E$  de la ligne de but  $[CD]$ , la règle est qu'elle doit tenter la transformation en un point  $T$  quelconque de la perpendiculaire à la ligne de but, passant par le point  $E$ . Le but de l'exercice est de déterminer, pour un point  $E$  de la ligne de but, le point  $T$  le plus favorable, c'est-à-dire celui d'où l'on voit les poteaux  $A$  et  $B$  sous un angle maximal.

Dans un repère  $(O, \vec{i}, \vec{j})$ , on note  $(-a, 0)$  et  $(a, 0)$  les coordonnées des poteaux  $A$  et  $B$  où  $a$  est un réel strictement positif.



On suppose qu'un essai a été marqué au point  $E$  du segment  $[BD]$ . Les coordonnées du point  $E$  s'écrivent  $(x, 0)$  avec  $x \geq a$  et  $(x, d)$  sont celles du point  $T$  avec  $d \geq 0$ . On note  $\theta = \widehat{ATB}$  l'angle sous lequel le rugbyman voit les deux poteaux depuis le point  $T$ .

- 1) Calculer  $\tan \theta$  en fonction de  $a$ ,  $x$  et  $d$ .
- 2) On suppose que  $x$  est fixé. Déterminer la valeur de  $d$  pour que l'angle  $\theta$  soit maximal. On note  $d(x)$  cette valeur de  $d$ .
- 3) Étudier les variations de la fonction  $x \mapsto d(x)$  pour  $x \geq a$ . Montrer que la droite d'équation  $y = x$  est asymptote à la courbe représentative de cette fonction.
- 4) Un terrain de rugby a pour côtés  $100\text{ m}$  et  $CD = 68\text{ m}$ . La distance entre deux poteaux est  $5,65\text{ m}$ . Vous êtes entraîneur d'une équipe de rugby. Quel conseil donnez-vous au joueur qui tente les transformations ?

## 6.3 Une boîte

Une boîte parallélépipédique en carton, de dimensions  $x, y, z$  ( $z$  est la hauteur de la boîte), a un volume de  $V$  litres. Trouver les dimensions qui minimisent la surface de carton.

## 6.4 Une autre boîte

Une boîte parallélépipédique en carton, de dimensions  $x, y, z$  ( $z$  est la hauteur de la boîte), a un volume de 32 litres et elle est ouverte en haut (elle n'a donc que cinq faces). Trouver les dimensions (exprimées en décimètres) qui minimisent la surface de carton. Justifier.

## 6.5 Alunissage

On considère le point  $m = (a, b, c)$  de l'espace et la sphère  $S$  de centre  $O = (0, 0, 0)$  et de rayon  $R$ . Trouver le point  $(x, y, z)$  de  $S$  le plus proche de  $m$ . (On pourra utiliser les extrema liés ou la géométrie.)

## 6.6 L'entropie

1) a) Soient  $\lambda$  et  $\mu$  deux nombres  $\geq 0$  tels que  $\lambda + \mu = 1$ . Montrer qu'on a, pour tous  $x, y > 0$ ,  $\lambda \ln x + \mu \ln y \leq \ln(\lambda x + \mu y)$ .

b) En déduire, en raisonnant par récurrence, que si  $\lambda_1, \dots, \lambda_n$  sont des réels positifs vérifiant  $\lambda_1 + \dots + \lambda_n = 1$  et  $x_1, \dots, x_n$  des réels positifs, on a :

$$\lambda_1 \ln x_1 + \dots + \lambda_n \ln x_n \leq \ln(\lambda_1 x_1 + \dots + \lambda_n x_n).$$

2) On considère des nombres réels positifs  $\lambda_i$  vérifiant  $\lambda_1 + \dots + \lambda_n = 1$ . Montrer que la fonction  $H(\lambda_1, \dots, \lambda_n) = -\sum_{i=1}^n \lambda_i \ln \lambda_i$  est maximum lorsque les  $\lambda_i$  sont tous égaux à  $1/n$ .

Cette fonction s'interprète comme l'entropie, donc l'incertitude, d'une expérience aléatoire dont les  $n$  issues ont les probabilités  $\lambda_i$ . On voit que l'incertitude est la plus grande quand l'expérience est équilibrée (on pensera au cas d'un dé : il y a moins d'incertitude si le dé est pipé).