

Université de Montréal

APPROXIMATION DES GÉOMÉTRIES

par

Sébastien Roy

Département de mathématiques et de statistique

Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en mathématiques

avril 1999

© Sébastien Roy, 1999



QA
3
U54
1999
V.017

APPROXIMATION DES GÉOMÉTRIE

APPROXIMATION DES GÉOMÉTRIE

APPROXIMATION DES GÉOMÉTRIE

APPROXIMATION DES GÉOMÉTRIE



APPROXIMATION DES GÉOMÉTRIE

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé

APPROXIMATION DES GÉOMÉTRIES

présenté par

Sébastien Roy

a été évalué par un jury composé des personnes suivantes :

Jacques Bélair

(président-rapporteur)

Michel Delfour

(directeur de recherche)

André Giroux

(membre du jury)

Mémoire accepté le :

26 août 1999

Sommaire

Le sujet de ce mémoire est la recherche de représentations des objets géométriques dans l'espace euclidien.

On rappelle les résultats sur les courbes à un paramètre et la caractérisation des entités géométriques comme la normale, le plan osculateur, les courbures et la torsion. On prend comme exemple les courbes de Bézier.

On explore ensuite les représentations des sous-variétés de codimension un par les courbes ou ensembles de niveau d'une fonction. On spécialise ces résultats aux courbes de niveau quadratiques et l'on évoque rapidement le cas cubique. Enfin, on essaie d'appliquer ces résultats au cas où la fonction est une fonction distance.

Remerciements

Je tiens à remercier les différents membres du personnel du département de mathématiques et statistique de l'Université de Montréal à qui j'ai eu affaire. Pour diverses raisons, ils ont dû répondre dans les deux dernières années à mes questions, que je qualifierais de *non traditionnelles*. Ils l'ont toujours fait avec le sourire. Merci beaucoup!

J'adresse un merci particulier à Monsieur M.Delfour, mon directeur de recherche. Je crois sincèrement que sa patience, ses encouragements et sa disponibilité m'ont permis de finalement mener à terme ce projet.

Pour leur soutien financier, je remercie l'Institut des sciences mathématiques et le Centre de recherches mathématiques.

À la mémoire de Diane.

Table des matières

Sommaire	iii
Remerciements	iv
Table des matières	v
Liste des figures	vi
1 Introduction	1
1 Préambule	1
2 Notations	2
2 Les courbes à un paramètre dans \mathbb{R}^3	4
3 Les courbes de Bézier	11
4 Les courbes de Bézier rationnelles	20
5 Variétés de codimension un par courbes de niveau	30
1 Un peu de théorie	30
2 Quelques applications	36
6 Approximation par des courbes de niveau quadratiques et cubiques	41
7 Les fonctions distance	48
1 La fonction distance	48
2 La fonction distance orientée	58
Références	62

Table des figures

7.1	Distance minimum du point x à la courbe $v = u^2$	58
-----	---	----

Chapitre 1

Introduction

1 Préambule

Le sujet de ce mémoire est la recherche de représentations des objets géométriques dans l'espace euclidien. Il existe mille et une théories mathématiques lorsque vient le temps de parler de courbes et de surfaces d'un espace euclidien. Certaines sont classiques et enseignées dans les premières années d'université. D'autres sont plus techniques et davantage utilisées par des spécialistes. En fait, on a l'impression que les surfaces et les courbes ont été scrutées par diverses méthodes, par divers individus, et ce, depuis toujours. On n'a qu'à penser aux recherches des géomètres, des analystes, des informaticiens... Y a-t-il encore des avenues possibles ? La réponse à cette question est évidemment positive, et il en existe une vers laquelle j'ai orienté la rédaction de ce mémoire.

Plutôt que de choisir des paramétrisations mathématiquement commodes (et connues), il semble intéressant de choisir celles qui font intervenir directement les caractéristiques naturelles des géométries inhérentes comme la normale, la courbure, la torsion, etc. Certains nouveaux résultats sur le sujet laissent en effet entrevoir un potentiel intéressant pour d'autres domaines mathématiques connexes.

Mes lectures ont donc porté sur de la théorie déjà existante afin que je puisse me familiariser avec différents concepts de géométrie différentielle, d'interpolation de fonctions, d'analyse de forme, etc. À travers ça, j'y suis allé de calculs, essayant de faire des liens entre cette théorie existante et mon objectif de faire ressortir des caractéristiques géométriques. J'en suis arrivé à quelques « résultats » et à

quelques murs...C'est donc sur tout ceci que les prochaines pages vont porter.

Concrètement, dans le chapitre 2, on rappelle les résultats sur les courbes à un paramètre et la caractérisation des entités géométriques comme la normale, le plan osculateur, les courbures et la torsion. Afin de fixer les idées, on a choisi comme exemple les courbes de Bézier qui sont très utilisées en conception assistée par ordinateur. On donne les principaux résultats sur les Bézier classiques au chapitre 3 et leur lien avec les polynômes de Bernstein. On complète avec les Bézier rationnelles au chapitre 4.

La deuxième partie du mémoire traite des représentations par courbes ou ensembles de niveau d'une fonction. On montre comment extraire les entités géométriques associées à la ligne de niveau zéro au chapitre 5. Il est alors naturel de faire l'approximation du domaine sous-jacent en approchant la fonction par des fonctions polynomiales par morceaux. Dans cet esprit, on spécialise, au chapitre 6, les résultats du chapitre précédent aux courbes de niveaux quadratiques et l'on évoque rapidement le cas cubique. On explore aussi au chapitre 7 la possibilité d'utiliser les fonctions distance pour lesquelles les expressions se simplifient. L'utilisation de ces dernières est en fait un cas particulier de la description par courbes ou ensembles de niveau.

2 Notations

Voici une liste des principales notations utilisées dans ce mémoire.

\mathbf{x}	vecteur (ou point) \mathbf{x}
$ x $	valeur absolue de x ($x \in \mathbb{R}$)
$\ \mathbf{x}\ $	norme euclidienne de \mathbf{x}
$d(\mathbf{x}, \mathbf{y})$	distance euclidienne entre \mathbf{x} et \mathbf{y}
$\dot{\mathbf{x}}(t)$	dérivée du vecteur position $\mathbf{x}(t)$ par rapport t
$\mathbf{x} \cdot \mathbf{y}$	produit scalaire de deux vecteurs
$\Omega_1 \setminus \Omega_2$	différence d'ensembles
$\partial\Omega$	frontière de Ω
χ_Ω	fonction caractéristique associée à Ω
A^{-}	matrice inverse de la matrice A
A^*	matrice transposée de la matrice A
$\det A$	déterminant de la matrice A

Df	matrice jacobienne d'une fonction f
D^2f	matrice hessienne d'une fonction f
$a:b$	indique <i>la mesure du rapport entre a et b</i>

Dans ce travail, les figures et les équations sont numérotées à l'aide d'un seul nombre. La référence est de la forme $(p.q)$ où p indique le chapitre et q le numéro de l'élément dans ce chapitre.

Chapitre 2

Les courbes à un paramètre dans \mathbb{R}^3

Une des raisons ayant mené au développement de la géométrie différentielle telle qu'on la connaît aujourd'hui, c'est le désir, le besoin de décrire les propriétés d'une courbe (ou d'une surface) de façon locale. C'est ainsi qu'en analysant une courbe de l'espace sous sa forme paramétrique, on en est arrivé à parler de vecteurs tangents, de courbure, etc.

Concrètement, cela signifie qu'on part avec une courbe de l'espace, qu'on considère comme un ensemble de points dont la position est donnée par le vecteur position

$$\mathbf{x} = \mathbf{x}(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix}, \quad t \in [a, b] \subset \mathbb{R}$$

Pour éviter de rencontrer des courbes d'allure « bizarre », on suppose que les composantes de \mathbf{x} admettent toutes des dérivées premières par rapport à t (notées ci-après à l'aide d'un point) et que ces dernières sont continues. De plus, on prend même pour acquis que

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{z}(t) \end{bmatrix} \neq \mathbf{0}, \quad t \in [a, b]$$

Une courbe \mathcal{C} étant un ensemble de points de l'espace, sa représentation paramétrique n'est pas unique ; en général, on peut associer à une courbe une infinité de

représentations de ce type. Par contre, on dit d'une paramétrisation qui vérifie les critères mentionnés ci-haut qu'elle est « régulière » et on qualifie de « lisse » une courbe représentant une telle paramétrisation. À partir d'une telle courbe, on définit la notion de longueur d'arc :

$$L = \int_a^b \|\dot{\mathbf{x}}\| dt$$

Bien que la formule ci-dessus soit exprimée en fonction du paramètre t , la longueur L d'un arc d'une courbe \mathcal{C} est strictement une propriété géométrique de cette courbe. En particulier, L est indépendante de la paramétrisation de \mathcal{C} . En effet, un changement « régulier » de paramètre n'influence pas la forme de la courbe, la longueur de cette dernière doit donc rester la même. Si τ désigne le nouveau paramètre et que $\tau = \tau(t)$, où τ est une fonction différentiable de t , alors

$$\frac{d\mathbf{x}}{dt} dt = \frac{d\mathbf{x}}{d\tau} \frac{d\tau}{dt} dt = \frac{d\mathbf{x}}{d\tau} d\tau$$

montre bien l'indépendance de L par rapport au paramétrage de \mathcal{C} . Ce ne sont que des caractéristiques comme la vitesse à laquelle cette courbe est parcourue ou le sens dans lequel elle est parcourue qui changent.

Il n'existe pas de représentation paramétrique idéale d'une courbe. Le choix du paramètre dépend habituellement du contexte. Cependant, il existe un paramètre « naturel », en ce sens qu'il dépend des propriétés géométriques (la forme et la longueur) de la courbe elle-même et non du repère dans lequel l'équation de la courbe est exprimée. Ce paramètre est une longueur d'arc, mesurée en partant d'un point donné P_0 de la courbe (d'habitude $\mathbf{x}(a)$). Le vecteur position \mathbf{x} d'un point $P = \mathbf{x}(t)$ quelconque de la courbe se définit en fonction de la longueur d'arc s mesurée le long de la courbe, du point initial P_0 à P :

$$\mathbf{x} = \mathbf{x}(s)$$

Dans les pages suivantes, on fera référence à cette représentation sous l'appellation **représentation paramétrique intrinsèque**. Toute courbe lisse admet une représentation de ce type, bien qu'il soit rarement facile (et la plupart du temps impossible) d'exprimer $\mathbf{x}(s)$ explicitement en partant d'une représentation paramétrique quelconque de la courbe. En fait, la représentation paramétrique intrinsèque est davantage utile à des fins théoriques que pratiques. Par contre, si

à partir d'une courbe définie en fonction d'un paramètre quelconque t , on peut évaluer explicitement la longueur

$$s = s(t) = \int_a^t \|\dot{\mathbf{x}}\| dt \quad t \in [a, b] \quad (2.1)$$

et si on peut résoudre l'équation $s = s(t)$ par rapport à t considéré comme une fonction de s [$t = t(s)$], alors on obtient la représentation paramétrique intrinsèque de la courbe :

$$\mathbf{x} = \mathbf{x}(t(s))$$

Maintenant qu'on a une expression mathématique représentant notre courbe \mathcal{C} , on peut en tirer des renseignements concernant son comportement.

Premièrement, on sait que le vecteur $\dot{\mathbf{x}}(t)$ est tangent à la courbe paramétrée d'équation $\mathbf{x} = \mathbf{x}(t)$ au point $\mathbf{x}(t)$. On a précédemment supposé que la norme de ce vecteur était différente de zéro pour tout t appartenant à l'intervalle $[a, b]$, alors en divisant $\dot{\mathbf{x}}(t)$ par sa longueur, on obtient un vecteur unitaire $\mathbf{T}(t)$ tangent à la courbe en $\mathbf{x}(t)$:

$$\mathbf{T}(t) = \frac{\dot{\mathbf{x}}(t)}{\|\dot{\mathbf{x}}(t)\|}$$

On sait par (2.1) que $ds = \|\dot{\mathbf{x}}(t)\| dt$ pour une courbe paramétrée d'équation $\mathbf{x} = \mathbf{x}(t)$. Si on choisit comme paramètre une longueur d'arc (i.e. $\mathbf{x} = \mathbf{x}(s)$) et qu'on note la dérivation par rapport à s , à l'aide du signe $'$, alors $ds = \|\mathbf{x}'(s)\| ds$. Cela implique que $\|\mathbf{x}'(s)\| = 1$. Donc, le vecteur unitaire tangent à la courbe est

$$\mathbf{T}(s) = \mathbf{x}'(s)$$

Toujours à partir de la représentation paramétrique intrinsèque, en supposant que \mathbf{x} est trois fois dérivable sur son domaine de définition et que ses dérivées successives sont continues, on peut définir d'autres fonctions. Étant donné que la longueur de $\mathbf{T}(s)$ est de une unité, ce vecteur tangent vérifie $\mathbf{T}(s) \cdot \mathbf{T}(s) = 1$. En dérivant chaque membre de cette équation par rapport à s , on obtient

$$2\mathbf{T}(s) \cdot \frac{d\mathbf{T}}{ds} = 0$$

ce qui implique que $d\mathbf{T}/ds = \mathbf{x}''(s)$ est perpendiculaire à $\mathbf{T}(s) = \mathbf{x}'(s)$. Soit

$$\kappa = \kappa(s) = \left\| \frac{d\mathbf{T}}{ds} \right\| = \|\mathbf{x}''(s)\| \quad (2.2)$$

Le réel non négatif $\kappa = \kappa(s)$ est appelé **courbure** de \mathcal{C} au point $\mathbf{x}(s)$. C'est une mesure du taux de variation de la direction de la tangente à la courbe en ce point. Si $\kappa(s) \neq 0$, en divisant $d\mathbf{T}/ds$ par sa longueur, soit $\kappa(s)$, on obtient un vecteur unitaire $\mathbf{N}(s)$ de même orientation que $d\mathbf{T}/ds$:

$$\mathbf{N}(s) = \frac{\mathbf{x}''(s)}{\|\mathbf{x}''(s)\|}$$

Ce vecteur $\mathbf{N}(s)$, qui est perpendiculaire à $\mathbf{T}(s)$ et donc à \mathcal{C} en $\mathbf{x}(s)$, est appelé **vecteur normal unitaire** de \mathcal{C} en $\mathbf{x}(s)$. Son sens est donné par le sens de rotation de $\mathbf{T}(s)$ lorsque s croît. On appelle **plan osculateur** de \mathcal{C} en $\mathbf{x}(s)$ le plan \mathbf{O} qui passe par ce point et qui contient les vecteurs \mathbf{T} et \mathbf{N} . Si \mathbf{y} est un point quelconque appartenant à ce plan, l'équation de \mathbf{O} est

$$\det [(\mathbf{y} - \mathbf{x}(s)) \quad \mathbf{x}'(s) \quad \mathbf{x}''(s)] = 0$$

Donc, le plan osculateur d'une courbe tridimensionnelle quelconque varie d'un point à l'autre de cette courbe : en chaque point, c'est le plan qui « contient le mieux » la portion de courbe située au voisinage de ce point. Soit $\kappa(s) \neq 0$, $\rho(s) = \frac{1}{\kappa(s)}$ (appelé rayon de courbure) et

$$\mathbf{x}_c(s) = \mathbf{x}(s) + \rho(s)\mathbf{N}(s)$$

Pour tout s , le point $\mathbf{x}_c(s)$ appartient au plan osculateur de \mathcal{C} en $\mathbf{x}(s)$ et il est situé du côté concave de \mathcal{C} , à une distance $\rho(s)$ de $\mathbf{x}(s)$. On appelle $\mathbf{x}_c(s)$ le centre de courbure de \mathcal{C} en $\mathbf{x}(s)$, et le cercle de rayon $\rho(s)$, contenu dans le plan osculateur et dont le centre est le centre de courbure est appelé **cercle osculateur** de \mathcal{C} en $\mathbf{x}(s)$. Il s'agit du cercle qui « décrit le mieux » le comportement de \mathcal{C} au voisinage de $\mathbf{x}(s)$.

En tout point d'une courbe \mathcal{C} où \mathbf{T} et \mathbf{N} sont définis, on définit un troisième vecteur unitaire \mathbf{B} :

$$\mathbf{B} = \mathbf{T} \times \mathbf{N} \quad \Leftrightarrow \quad \mathbf{N} = \mathbf{B} \times \mathbf{T} \quad \Leftrightarrow \quad \mathbf{T} = \mathbf{N} \times \mathbf{B}$$

appelé **vecteur binormal unitaire** de \mathcal{C} en $\mathbf{x}(s)$. Comme $\mathbf{B}(s) \cdot \mathbf{B}(s) = 1$, alors $\mathbf{B}(s) \cdot (d\mathbf{B}/ds) = 0$ et $d\mathbf{B}/ds$ est perpendiculaire à $\mathbf{B}(s)$. De plus, en dérivant chaque membre de $\mathbf{B} = \mathbf{T} \times \mathbf{N}$ par rapport à s , on obtient

$$\frac{d\mathbf{B}}{ds} = \frac{d\mathbf{T}}{ds} \times \mathbf{N} + \mathbf{T} \times \frac{d\mathbf{N}}{ds} = \kappa\mathbf{N} \times \mathbf{N} + \mathbf{T} \times \frac{d\mathbf{N}}{ds} = \mathbf{T} \times \frac{d\mathbf{N}}{ds}$$

Donc $d\mathbf{B}/ds$ est aussi perpendiculaire à \mathbf{T} . Il s'ensuit que $d\mathbf{B}/ds$ est parallèle à \mathbf{N} et qu'il existe une fonction $\tau(s)$ telle que

$$\frac{d\mathbf{B}}{ds} = -\tau(s)\mathbf{N}(s) \quad (2.3)$$

On appelle le réel $\tau(s)$ **torsion** de \mathcal{C} en $\mathbf{x}(s)$. Ce nombre est une mesure du degré « d'enroulement » de la courbe au voisinage de $\mathbf{x}(s)$, c'est-à-dire qu'il indique dans quelle mesure la courbe « sort » d'un plan. En appliquant les formules $d\mathbf{T}/ds = \kappa\mathbf{N}$ et $d\mathbf{B}/ds = -\tau\mathbf{N}$, on peut calculer $d\mathbf{N}/ds$:

$$\frac{d\mathbf{N}}{ds} = \frac{d}{ds}(\mathbf{B} \times \mathbf{T}) = -\tau\mathbf{N} \times \mathbf{T} + \kappa\mathbf{B} \times \mathbf{N} = \tau\mathbf{B} - \kappa\mathbf{T}$$

Les trois équations

$$\begin{aligned} \frac{d\mathbf{T}}{ds} &= \kappa\mathbf{N} \\ \frac{d\mathbf{N}}{ds} &= -\kappa\mathbf{T} + \tau\mathbf{B} \\ \frac{d\mathbf{B}}{ds} &= -\tau\mathbf{N} \end{aligned}$$

sont appelées **formules de Frenet-Serret**. Ces formules sont très utiles. Par exemple, on peut les utiliser pour déterminer une expression formelle pour $\tau(s)$ (que ne donne pas 2.3). En effet,

$$\begin{aligned} \mathbf{x}''' &= \frac{d}{ds} \left(\frac{d\mathbf{T}}{ds} \right) = \frac{d\kappa}{ds}\mathbf{N} + \kappa \frac{d\mathbf{N}}{ds} \\ \Rightarrow \quad \mathbf{x}'' \times \mathbf{x}''' &= \kappa^2 \mathbf{N} \times \frac{d\mathbf{N}}{ds} \\ &= \kappa^2 \mathbf{N} \times (-\kappa\mathbf{T} + \tau\mathbf{B}) \\ &= \kappa^2 (\kappa\mathbf{B} + \tau\mathbf{T}) \\ \Rightarrow \quad \mathbf{x}' \cdot (\mathbf{x}'' \times \mathbf{x}''') &= \kappa^2 \tau \\ \Rightarrow \quad \tau(s) &= \frac{\mathbf{x}' \cdot (\mathbf{x}'' \times \mathbf{x}''')}{\kappa^2} \end{aligned}$$

La plupart des formules énoncées jusqu'à maintenant ne sont pas très utiles pour analyser une courbe définie par une équation paramétrique non intrinsèque. On peut toutefois les généraliser au cas où le paramètre est quelconque, et ce à l'aide des vecteurs $\dot{\mathbf{x}}$, $\ddot{\mathbf{x}}$ et $\ddot{\mathbf{x}}$ associés à une représentation paramétrique $\mathbf{x} = \mathbf{x}(t)$.

On sait déjà que $\mathbf{T} = \mathbf{T}(t) = \dot{\mathbf{x}}/\|\dot{\mathbf{x}}\|$, c'est-à-dire

$$\begin{aligned} \dot{\mathbf{x}} &= \|\dot{\mathbf{x}}\|\mathbf{T} \\ \Rightarrow \quad \ddot{\mathbf{x}} &= \frac{d\|\dot{\mathbf{x}}\|}{dt}\mathbf{T} + \|\dot{\mathbf{x}}\|\frac{d\mathbf{T}}{dt} \\ &= \frac{d\|\dot{\mathbf{x}}\|}{dt}\mathbf{T} + \|\dot{\mathbf{x}}\|\frac{d\mathbf{T}}{ds}\frac{ds}{dt} = \frac{d\|\dot{\mathbf{x}}\|}{dt}\mathbf{T} + \|\dot{\mathbf{x}}\|^2\kappa\mathbf{N} \\ \Rightarrow \quad \dot{\mathbf{x}} \times \ddot{\mathbf{x}} &= \|\dot{\mathbf{x}}\|^3\kappa\mathbf{T} \times \mathbf{N} = \|\dot{\mathbf{x}}\|^3\kappa\mathbf{B} \end{aligned}$$

Donc, on tire des équations ci-dessus, les expressions suivantes de κ et \mathbf{B} :

$$\kappa = \kappa(t) = \frac{\|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}\|}{\|\dot{\mathbf{x}}\|^3} \quad \mathbf{B} = \mathbf{B}(t) = \frac{\dot{\mathbf{x}} \times \ddot{\mathbf{x}}}{\|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}\|}$$

À l'aide de l'expression de $\ddot{\mathbf{x}}$ ci-haut, on pourrait également calculer \mathbf{N} , mais il est plus simple de se servir directement de $\mathbf{N} = \mathbf{B} \times \mathbf{T}$, avec \mathbf{B} et \mathbf{T} calculés au préalable avec les formules ci-haut. Il reste à calculer la torsion. Pour ce faire, on note que

$$\ddot{\mathbf{x}} = \frac{d}{dt} \left(\frac{d\|\dot{\mathbf{x}}\|}{dt}\mathbf{T} + \|\dot{\mathbf{x}}\|^2\kappa\mathbf{N} \right)$$

Le développement du membre de droite comprend plusieurs termes, mais un seul contient \mathbf{B} : celui obtenu en calculant $\|\dot{\mathbf{x}}\|^2\kappa(d\mathbf{N}/dt) = \|\dot{\mathbf{x}}\|^3\kappa(d\mathbf{N}/ds) = \|\dot{\mathbf{x}}\|^3\kappa(\tau\mathbf{B} - \kappa\mathbf{T})$. Donc,

$$\ddot{\mathbf{x}} = \lambda\mathbf{T} + \mu\mathbf{N} + \|\dot{\mathbf{x}}\|^3\kappa\tau\mathbf{B}$$

où λ et μ sont des fonctions de t à valeurs réelles. On en déduit que

$$(\dot{\mathbf{x}} \times \ddot{\mathbf{x}}) \cdot \ddot{\mathbf{x}} = (\|\dot{\mathbf{x}}\|^3\kappa)^2\tau = \|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}\|^2\tau$$

c'est-à-dire que

$$\tau = \tau(t) = \frac{(\dot{\mathbf{x}} \times \ddot{\mathbf{x}}) \cdot \ddot{\mathbf{x}}}{\|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}\|^2} = \frac{\dot{\mathbf{x}} \cdot (\ddot{\mathbf{x}} \times \ddot{\mathbf{x}})}{\|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}\|^2} = \frac{\det[\dot{\mathbf{x}}, \ddot{\mathbf{x}}, \ddot{\mathbf{x}}]}{\|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}\|^2}$$

À ce point-ci, on dispose d'assez d'éléments pour décrire une courbe de façon locale, ce qui était l'objectif initial de ce chapitre. On peut finalement relier toute cette information par l'entremise de la notion de **repère de Frenet**. Soit $\mathbf{x}(t)$, un point sur notre courbe \mathcal{C} . En supposant toujours que toutes les dérivées dont nous avons besoin existent, alors le développement de Taylor de $\mathbf{x}(t + \Delta t)$ à t est

$$\mathbf{x}(t + \Delta t) = \mathbf{x} + \dot{\mathbf{x}}\Delta t + \frac{1}{2}\ddot{\mathbf{x}}\Delta t^2 + \frac{1}{6}\ddot{\mathbf{x}}\Delta t^3 + \dots$$

Maintenant, si on suppose en plus que $\dot{\mathbf{x}}$, $\ddot{\mathbf{x}}$ et $\ddot{\ddot{\mathbf{x}}}$ sont linéairement indépendants, alors ces trois vecteurs forment un système local de coordonnées dont l'origine est en \mathbf{x} . À partir de ce système, on obtient (toujours localement) un système orthonormé de coordonnées en employant la procédure d'orthogonalisation de Gram-Schmidt. On a déjà rencontré les vecteurs formant cette nouvelle base : \mathbf{T} , \mathbf{N} et \mathbf{B} . Ce sont eux, ayant pour origine $\mathbf{x}(t)$, qui forment ce qu'on appelle le repère de Frenet. Ce dernier varie le long de \mathcal{C} au fur et à mesure que le paramètre t parcourt l'intervalle $[a, b]$.

Chapitre 3

Les courbes de Bézier

Jusqu'à maintenant, il a seulement été question de courbes dont on connaissait la définition formelle, à partir de laquelle on déduisait leurs propriétés géométriques. Dans le présent chapitre, les choses seront abordées d'une toute autre façon : l'approximation de courbe fera son apparition. Même si l'avènement des ordinateurs a beaucoup influencé le développement de nouvelles théories, différents mathématiciens avaient préalablement abordé le sujet. Donc, aujourd'hui, il existe plusieurs façons d'approcher les valeurs réelles d'une courbe. Toutefois, chacune dépend de l'information dont on dispose au départ et des résultats désirés au terme des calculs à effectuer.

En fait, l'approximation de fonctions se résume à deux sortes d'approche. Une première où on a la définition formelle de la fonction sur laquelle on travaille, mais qu'on désire « transformer », le plus fidèlement possible, à l'aide de fonctions plus simples (e.g. polynômes de Taylor). La deuxième, celle sur laquelle on s'attardera, consiste plutôt en la recherche de la « meilleure » fonction possible concordant avec une certaine quantité de données connues à priori (valeurs en certains points de la fonction elle-même ou de l'une de ses dérivées...) sans nécessairement connaître sa définition formelle. À titre d'exemples, mentionnons les différentes techniques d'interpolation polynomiale de Newton, de Lagrange et d'Hermite. On peut aussi parler d'approximation trigonométrique (« Fast Fourier Transforms »), d'approximation à l'aide de fonctions rationnelles. Finalement, provenant du milieu des modélistes graphiques, on a les **courbes de Bézier**. Voyons voir en quoi consistent ces dernières.

Ce qui suit est un cas particulier qui permet de réaliser à quel point cette

méthode est relativement simple car basée sur une idée géométrique. Soit $\mathbf{b}_0, \mathbf{b}_1$ et \mathbf{b}_2 , trois points de \mathbb{R}^3 , et soit $t \in \mathbb{R}$. Considérons

$$\begin{aligned}\mathbf{b}_0^1(t) &= (1-t)\mathbf{b}_0 + t\mathbf{b}_1 \\ \mathbf{b}_1^1(t) &= (1-t)\mathbf{b}_1 + t\mathbf{b}_2 \\ \mathbf{b}_0^2(t) &= (1-t)\mathbf{b}_0^1(t) + t\mathbf{b}_1^1(t).\end{aligned}$$

En insérant les deux premières équations dans la troisième, on obtient

$$\mathbf{b}_0^2(t) = (1-t)^2\mathbf{b}_0 + 2t(1-t)\mathbf{b}_1 + t^2\mathbf{b}_2 \quad (3.1)$$

Lorsqu'on fait varier t sur un intervalle continu, cette expression quadratique décrit un morceau de parabole dans l'espace. On désigne par $\mathbf{b}^2(t)$ le morceau ainsi obtenu (alors que $\mathbf{b}_0^2(t)$ est un point sur ce morceau). En particulier, si t se promène entre 0 et 1, $\mathbf{b}^2(t)$ est à l'intérieur du triangle formé par $\mathbf{b}_0, \mathbf{b}_1$ et \mathbf{b}_2 (i.e l'enveloppe convexe définie par ces trois points). En effet, si on s'y attarde quelques secondes, on remarque que la résultante (3.1) provient de trois interpolations linéaires répétées et est, pour $t \in [0, 1]$, une suite de points qui appartiennent tous à différentes droites joignant les segments $\overline{\mathbf{b}_0\mathbf{b}_1}$ et $\overline{\mathbf{b}_1\mathbf{b}_2}$, qui sont les arêtes dudit triangle. À ce point-ci, il est important de noter que $\mathbf{b}^2(t)$ ne sort jamais du plan contenant le triangle $\Delta\mathbf{b}_0\mathbf{b}_1\mathbf{b}_2$. Toutefois, lorsqu'on généralise le processus décrit ci-haut à une situation où on connaît $n+1$ points ($n > 2$), en plus d'obtenir une courbe de degré n , on a aussi la possibilité de travailler avec une véritable courbe de l'espace. De manière formelle, soit $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_n \in \mathbb{R}^3$ et $t \in \mathbb{R}$, et posons

$$\mathbf{b}_i^r(t) = (1-t)\mathbf{b}_i^{r-1}(t) + t\mathbf{b}_{i+1}^{r-1}(t) \quad \begin{cases} r=1, \dots, n \\ i=0, \dots, n-r \end{cases}$$

avec $\mathbf{b}_i^0(t) = \mathbf{b}_i$. Alors $\mathbf{b}_0^n(t)$ est le point correspondant au paramètre t sur la courbe de Bézier $\mathbf{b}^n(t)$. On appelle **polygone de Bézier** ou **polygone de contrôle** de la courbe $\mathbf{b}^n(t)$, le polygone \mathbf{P} formé par $\mathbf{b}_0, \dots, \mathbf{b}_n$. De même, les sommets \mathbf{b}_i du polygone sont appelés **points de Bézier** ou **points de contrôle**.

Nous avons maintenant une définition décrivant ce que sont les courbes de Bézier. Celle-ci étant récursive, elle nous permet de relativement bien comprendre géométriquement ce que sont ces courbes. Par contre, cet avantage rend plus compliquée l'obtention de renseignements mathématiques comme la notion de

dérivée à ces courbes. En fait, pour faciliter tout développement théorique sur les courbes de Bézier, on a besoin de les définir explicitement. C'est ici qu'entrent en scène les polynômes de Bernstein

$$B_i^n(t) = \binom{n}{i} t^i (1-t)^{n-i}$$

où $\binom{n}{i}$ est la notation employée pour les coefficients binomiaux :

$$\binom{n}{i} = \begin{cases} \frac{n!}{i!(n-i)!} & \text{si } 0 \leq i \leq n \\ 0 & \text{sinon.} \end{cases}$$

À ce jour, ces polynômes ont été étudiés par plusieurs mathématiciens. Une de leurs importantes propriétés en est une de récursivité :

$$B_i^n(t) = (1-t)B_i^{n-1}(t) + tB_{i-1}^{n-1}(t).$$

Les polynômes de Bernstein possèdent une autre propriété importante, celle de former une partition de l'unité, c'est-à-dire :

$$\sum_{j=0}^n B_j^n(t) = 1$$

Finalement, les polynômes de Bernstein sont tels que

$$\begin{aligned} B_j^n(0) &= \delta_{j,0} \quad (= 1 \text{ si } j = 0, \text{ } 0 \text{ sinon}) \\ B_j^n(1) &= \delta_{j,n} \quad (= 1 \text{ si } j = n, \text{ } 0 \text{ sinon}). \end{aligned}$$

Avec ces trois propriétés à l'esprit, on peut maintenant faire le lien avec les courbes de Bézier. En effet, chaque point intermédiaire \mathbf{b}_i^r provenant de la première définition présentée des courbes de Bézier s'exprime en termes des polynômes de Bernstein de degré r et de certains points de Bézier :

$$\mathbf{b}_i^r(t) = \sum_{j=0}^r \mathbf{b}_{i+j} B_j^r(t) \quad \begin{array}{l} r \in \{0, \dots, n\} \\ i \in \{0, \dots, n-r\}. \end{array} \quad (3.2)$$

Si $r = n$, (3.2) nous donne alors cette seconde définition, plus explicite, d'une courbe de Bézier :

$$\mathbf{b}^n(t) = \mathbf{b}_0^n(t) = \sum_{j=0}^n \mathbf{b}_j B_j^n(t). \quad (3.3)$$

À partir de cette définition, il est facile de vérifier que $\mathbf{b}^n(0) = \mathbf{b}_0$ et $\mathbf{b}^n(1) = \mathbf{b}_n$. On peut aussi prouver algébriquement certaines propriétés des courbes de Bézier :

- une transformation affine appliquée à une courbe de Bézier n'en modifie pas la forme
- une transformation affine appliquée au paramètre t d'une courbe de Bézier ($t \mapsto u = t(b - a) + a$) ne donne pas une nouvelle courbe, i.e. :

$$\mathbf{b}^n(t) = \sum_{j=0}^n \mathbf{b}_j B_j^n(t) = \sum_{j=0}^n \mathbf{b}_j B_j^n\left(\frac{u-a}{b-a}\right) = \mathbf{b}^n(u)$$

- pour $t \in [0, 1]$, une courbe de Bézier $\mathbf{b}^n(t)$ reste à l'intérieur de l'enveloppe convexe définie par son polygone de contrôle.
- puisque $B_j^n(t) = B_{n-j}^n(1-t)$, alors

$$\sum_{j=0}^n \mathbf{b}_j B_j^n(t) = \sum_{j=0}^n \mathbf{b}_{n-j} B_j^n(1-t)$$

et donc, peu importe que l'on ait étiqueté les points de Bézier $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_n$ ou $\mathbf{b}_n, \mathbf{b}_{n-1}, \dots, \mathbf{b}_0$, la forme de la courbe finale obtenue est la même dans les deux cas (elles sont simplement parcourues en sens opposé). Parce qu'elle possède cette propriété, on dit donc d'une courbe de Bézier qu'elle est symétrique par rapport à l'ordre avec lequel on parcourt les points de contrôle.

Il est à noter que, jusqu'à présent, le paramètre t ne varie qu'entre 0 et 1. Dans les faits, lorsqu'il est question d'une courbe de Bézier \mathbf{b}^n , on a coutume de supposer que son domaine de définition est cet intervalle $[0, 1]$ en raison de la propriété d'invariance d'une Bézier par rapport à une transformation affine de son paramètre. Il est toutefois bon de mentionner qu'il peut arriver qu'on rencontre des courbes de Bézier définies sur des intervalles différents de $[0, 1]$. Dans le même ordre d'idées, on peut aussi définir une nouvelle courbe de Bézier \mathbf{c}^n (et donc un nouveau polygone de contrôle) correspondant à la partie d'une courbe \mathbf{b}^n engendrée lorsque t varie de 0 à d ($0 < d < 1$). Les nouveaux points de Bézier

\mathbf{c}_j sont donnés par $\mathbf{c}_j = \mathbf{b}_0^j(d)$. En effet, pour $t \in [0, 1]$

$$\begin{aligned}
 \mathbf{c}^n(t) &= \sum_{j=0}^n \mathbf{c}_j B_j^n(t) = \sum_{j=0}^n \mathbf{b}_0^j(d) B_j^n(t) \\
 &= \sum_{j=0}^n \left(\sum_{k=0}^j \mathbf{b}_k B_k^j(d) \right) B_j^n(t) \\
 &= \sum_{k=0}^n \mathbf{b}_k \sum_{j=0}^n B_k^j(d) B_j^n(t) \\
 &= \sum_{k=0}^n \mathbf{b}_k B_k^n(td) \\
 &= \mathbf{b}^n(dt)
 \end{aligned}$$

On dit alors qu'on *subdivise* la courbe de Bézier initiale. En raison de la propriété de symétrie des Bézier, on peut aussi faire cette subdivision au niveau de la « queue » de \mathbf{b}^n , c'est-à-dire définir une nouvelle courbe \mathbf{c}^n correspondant à $\mathbf{b}^n([d, 1])$: les points de Bézier sont $\mathbf{c}_j = \mathbf{b}_j^{n-j}(d)$. Que se passe-t-il à l'extérieur de $[0, 1]$? Et bien, par le même raisonnement que pour la subdivision, on peut également *extrapoler*. On obtient alors des points de Bézier \mathbf{c}_j correspondant, par exemple, à l'intervalle $[1, d]$: $\mathbf{c}_j = \mathbf{b}_{n-j}^j(d)$.

Finalement, grâce aux polynômes de Bernstein, il est plus simple de donner la définition des différentes dérivées d'une courbe de Bézier. Pour ce faire, on commence par calculer la dérivée première des polynômes de Bernstein :

$$\frac{d}{dt} B_i^n(t) = n \left[B_{i-1}^{n-1}(t) - B_i^{n-1}(t) \right].$$

En utilisant (3.3), on obtient donc

$$\frac{d}{dt} \mathbf{b}^n(t) = n \sum_{j=0}^n \left[B_{j-1}^{n-1}(t) - B_j^{n-1}(t) \right] \mathbf{b}_j,$$

qu'on peut simplifier en utilisant $B_j^n(t) \equiv 0$ pour $j \notin \{0, \dots, n\}$ et en effectuant un changement d'indice, pour ainsi obtenir

$$\begin{aligned}
 \frac{d}{dt} \mathbf{b}^n(t) &= \frac{d}{dt} \left(\sum_{j=0}^n \mathbf{b}_j B_j^n(t) \right) = n \sum_{j=0}^{n-1} (\mathbf{b}_{j+1} - \mathbf{b}_j) B_j^{n-1}(t) \\
 &= n \sum_{j=0}^{n-1} \Delta \mathbf{b}_j B_j^{n-1}(t)
 \end{aligned} \tag{3.4}$$

où Δ est le symbole adopté pour dénoter l'opérateur de différence avant (i.e. $\Delta \mathbf{b}_j \stackrel{\text{déf}}{=} \mathbf{b}_{j+1} - \mathbf{b}_j$). La formule (3.4) nous indique donc que la dérivée d'une courbe de Bézier est elle-même une autre courbe de Bézier, sauf qu'au lieu d'être des points de l'espace, les coefficients de cette nouvelle courbe sont des vecteurs ($\Delta \mathbf{b}_j$).

Pour ce qui est des dérivées d'ordres supérieurs, on n'a qu'à appliquer (3.4) à répétition. À la fin du processus, on obtient la formule suivante pour la dérivée $r^{\text{ième}}$

$$\frac{d^r}{dt^r} \mathbf{b}^n(t) = \frac{n!}{(n-r)!} \sum_{j=0}^{n-r} \Delta^r \mathbf{b}_j B_j^{n-r}(t) \quad (3.5)$$

où $\Delta^r \mathbf{b}_j$ est l'équivalent de l'opérateur de différence avant dans le cas où on a des itérations répétées :

$$\begin{aligned} \Delta^r \mathbf{b}_j &\stackrel{\text{déf}}{=} \Delta^{r-1} \mathbf{b}_{j+1} - \Delta^{r-1} \mathbf{b}_j \\ &= \sum_{i=0}^r \binom{r}{i} (-1)^{r-i} \mathbf{b}_{i+j}. \end{aligned}$$

Maintenant que les polynômes de Bernstein nous ont permis d'obtenir (3.5), on peut utiliser le fait que

$$\sum_{j=0}^{n-1} \Delta \mathbf{b}_j = \sum_{j=1}^n \mathbf{b}_j - \sum_{j=0}^{n-1} \mathbf{b}_j = \Delta \sum_{j=0}^{n-1} \mathbf{b}_j$$

pour réexprimer cette expression de la dérivée $r^{\text{ième}}$ d'une courbe de Bézier en fonction des points intermédiaires $\mathbf{b}_i^k(t)$ générés par la définition « géométrique » de cette courbe :

$$\frac{d^r}{dt^r} \mathbf{b}^n(t) = \frac{n!}{(n-r)!} \Delta^r \mathbf{b}_0^{n-r}(t). \quad (3.6)$$

Avec ces deux expressions de la dérivée $\frac{d^r}{dt^r} \mathbf{b}^n(t)$ entre nos mains, on note deux choses. Premièrement, en $t = 0$ et $t = 1$, on voit que (3.5) nous donne

$$\frac{d^r}{dt^r} \mathbf{b}^n(0) = \frac{n!}{(n-r)!} \Delta^r \mathbf{b}_0$$

et

$$\frac{d^r}{dt^r} \mathbf{b}^n(1) = \frac{n!}{(n-r)!} \Delta^r \mathbf{b}_{n-r}.$$

Donc, la dérivée $r^{\text{ième}}$ d'une courbe de Bézier à chacune de ses extrémités dépend de $r + 1$ points de Bézier dans l'entourage de l'extrémité concernée. Deuxièmement, la dérivée première de $\mathbf{b}^n(t)$ obtenue à partir de (3.6) nous indique que

$$\frac{d}{dt}\mathbf{b}^n(t) = n[\mathbf{b}_1^{n-1}(t) - \mathbf{b}_0^{n-1}(t)]$$

c'est-à-dire que, pour t fixé, les points intermédiaires $\mathbf{b}_1^{n-1}(t)$ et $\mathbf{b}_0^{n-1}(t)$ engendrent le vecteur tangent à la courbe de Bézier.

Ici se termine le survol des principales propriétés et caractéristiques d'une courbe de Bézier. Comme pour tout processus d'interpolation, celui de Bézier doit être soumis à une évaluation de base (résultats obtenus, coûts afférents en fonction des résultats recherchés, etc) pour que, par rapport aux autres méthodes, ceux qui l'utilisent sachent quels sont ses points forts et ses points faibles. Une observation faite à la suite d'une telle analyse concerne la capacité de la modélisation de Bézier de fournir une résultante simple si le polygone de contrôle (donc, la forme de la courbe à obtenir...) est complexe. En fait, on a un peu le même problème qu'avec la simple interpolation polynomiale : dès qu'on en veut un peu plus, qu'on voit un peu plus gros, et bien ça devient compliqué ! Toutefois, au même genre de problème, on a le même genre de solution : l'interpolation par morceaux (ou interpolation à l'aide de splines). Pour se faire, on commence par faire un travail d'interpolation sur de petits intervalles, pour ensuite joindre le tout. Dans le cas qui nous intéresse (celui où on réunit une collection de courbes de Bézier), il est certain qu'à priori, on peut « coller » les morceaux sans trop se poser de questions. Toutefois, on se doute bien que si l'on veut que notre approximation globale soit simple et pratique à utiliser, on devra imposer quelques conditions de raccord. Voyons voir ce qu'il en est.

Premièrement, soit \mathbf{C} , la courbe par morceaux, obtenue à la suite de ce processus d'unification. On dénotera par $\mathbf{C}(u)$, le point de cette courbe correspondant au u fixé. Evidemment, u est choisi à l'intérieur de l'un des intervalles $[u_i, u_{i+1}]$, à partir desquels sont construites les différentes courbes de Bézier qu'on appellera \mathbf{C}_i . Sur chacun de ces intervalles, on déterminera également un paramètre local t de la façon suivante :

$$t = \frac{u - u_i}{\Delta_i} \quad (\text{où } \Delta_i = u_{i+1} - u_i).$$

Ainsi chaque courbe de Bézier sera définie sur $[0, 1]$, et on parlera donc de $\mathbf{C}(u)$

et de $\mathbf{C}_i(t)$ lorsqu'il sera question respectivement d'un point sur la courbe globale \mathbf{C} et d'un point sur la courbe de Bézier \mathbf{C}_i . Les points de jonction de la spline \mathbf{C} seront donc tels que $\mathbf{C}(u_i) = \mathbf{C}_i(0) = \mathbf{C}_{i-1}(1)$. Le fait qu'on définisse un paramètre local sur chaque $[u_i, u_{i+1}]$ entraîne que

$$\frac{d\mathbf{C}(u)}{du} = \frac{d\mathbf{C}_i(t)}{dt} \frac{dt}{du} = \frac{d\mathbf{C}_i(t)}{dt} \frac{1}{\Delta_i}. \quad (3.7)$$

Maintenant que les principales notations de la théorie des courbes splines sous forme de Bézier sont connues, terminons ce chapitre en regardant comment on peut « lisser » ces fameuses courbes. Pour les besoins de la cause, soit deux courbes de Bézier \mathbf{C}_0 et \mathbf{C}_1 , avec $\mathbf{b}_0, \dots, \mathbf{b}_n$ et $\mathbf{b}_n, \dots, \mathbf{b}_{2n}$, leur polygone de contrôle respectif. On supposera que lorsqu'elles sont jointes, ces deux courbes forment la courbe polynomiale $\mathbf{C}(u)$, $\mathbf{C}([u_0, u_1])$ correspondant à $\mathbf{C}_0([0, 1])$ et $\mathbf{C}([u_1, u_2])$ à $\mathbf{C}_1([0, 1])$. Comme chaque morceau est indéfiniment différentiable, le seul endroit qui demande vérification si l'on désire une courbe lisse, c'est le point de jonction \mathbf{b}_n . En effet, pour avoir une dérivée première continue en ce point, (3.7) implique que

$$\frac{d\mathbf{C}(u_1)}{du} = \frac{d\mathbf{C}_0(1)}{dt} \frac{1}{\Delta_0} = \frac{d\mathbf{C}_1(0)}{dt} \frac{1}{\Delta_1}.$$

La dernière de ces deux égalités peut être réécrite comme suit

$$\Delta_1 \Delta \mathbf{b}_{n-1} = \Delta_0 \Delta \mathbf{b}_n. \quad (3.8)$$

On remarque que cette équation met en présence des vecteurs (formés à partir de \mathbf{b}_{n-1} , \mathbf{b}_n et \mathbf{b}_{n+1}). En fait, cette équation nous indique que ces trois points doivent être colinéaires pour que l'on puisse espérer avoir une courbe globale C^1 . Mais on remarque également que la longueur des intervalles $[u_0, u_1]$ et $[u_1, u_2]$ a un rôle à jouer dans l'obtention d'une spline lisse. En fait, si on revient à la condition (3.8), et qu'on la transforme afin d'isoler \mathbf{b}_{n+1} , on obtient une combinaison barycentrique de \mathbf{b}_{n-1} et \mathbf{b}_n :

$$\mathbf{b}_{n+1} = (1 - t_2)\mathbf{b}_{n-1} + t_2\mathbf{b}_n$$

avec

$$t_2 = \frac{u_2 - u_0}{u_1 - u_0},$$

c'est-à-dire que t_2 est le paramètre local de u_2 associé à l'intervalle $[u_0, u_1]$. Donc, pour que la spline résultante soit C^1 , il faut que b_{n+1} corresponde au point de Bézier obtenu suite à l'*extrapolation* sur $[1, t_2]$ de \mathbf{C}_0 . En fait, cette condition est généralisable au cas où l'on désire une spline \mathbf{C}^r créée à partir de \mathbf{C}_0 et \mathbf{C}_1 (elle est nécessaire et suffisante) :

$$\mathbf{b}_{n+i} = \mathbf{b}_{n-i}^i(t_2) \quad i = 0, \dots, r.$$

Chapitre 4

Les courbes de Bézier rationnelles

Dans le chapitre précédent, la seule courbe de degré deux à laquelle nous avons accès grâce à la théorie de Bézier était une parabole. Le but de ce chapitre est de présenter une autre catégorie de représentations paramétriques de courbes, les **Bézier rationnelles**, plus générales que les simples Bézier, nous ouvrant donc de nouvelles avenues en ce qui concerne l'approximation. Comme dans le chapitre précédent, avant de donner la définition générale d'une telle courbe (degré n), voyons voir ce qui se passe dans le cas quadratique.

Soit

$$\mathbf{Q}(t) = \frac{\sum_{j=0}^2 w_j \mathbf{P}_j B_j^2(t)}{\sum_{j=0}^2 w_j B_j^2(t)} \quad (4.1)$$

où les w_j sont des scalaires appartenant à \mathbb{R} , les \mathbf{P}_j dénotent des points de \mathbb{R}^3 et les $B_j^2(t)$ sont toujours les polynômes de Bernstein. On donne le nom de **courbe de Bézier rationnelle** de degré deux à $\mathbf{Q}(t)$, le numérateur et le dénominateur pouvant chacun être considéré comme des courbes de Bézier conventionnelles. Les w_j sont appelés **pooids** et les \mathbf{P}_j , **points de contrôle**. Allons-y maintenant de quelques observations.

Aux extrémités (en $t = 0$ et $t = 1$), comme pour les Bézier non rationnelles, on voit que $\mathbf{Q}(t)$ joue un rôle d'interpolant : $\mathbf{Q}(0) = \mathbf{P}_0$ et que $\mathbf{Q}(1) = \mathbf{P}_2$. Étant donné la présence des polynômes de Bernstein dans (4.1), les Bézier rationnelles jouissent, comme le faisait les Bézier traditionnelles, du fait d'être invariantes par

rapport à un changement affine de paramètre. On constate également que si tous les poids w_j sont égaux dans (4.1), on retrouve l'expression d'une courbe de Bézier non rationnelle telle que donnée précédemment. Toujours en lien avec les poids, on remarque que si on multiplie chaque w_j par un facteur commun non nul, la courbe résultante associée à ces nouveaux poids est exactement la même qu'avec les w_j initiaux. En fait, si on ne fait pas de changement de paramétrisation, la représentation (4.1) est unique, c'est-à-dire qu'il existe une seule combinaison de poids $w_0 : w_1 : w_2$ qui permette d'associer un t spécifique à un point spécifique. Par contre, on verra un peu plus loin que si on modifie le paramètre t (par conséquent la manière dont on parcourt la courbe), on peut obtenir différentes représentations de la même courbe initiale, et donc de nouveaux poids \tilde{w}_j tels que $\tilde{w}_0 : \tilde{w}_1 : \tilde{w}_2 \neq w_0 : w_1 : w_2$, c'est-à-dire que le rapport entre les différents poids n'est plus le même.

Afin d'avoir une écriture un peu plus concise pour la suite de ce texte, on emploiera $\mathbf{P}(t)$ et $w(t)$ pour parler respectivement du numérateur et du dénominateur de $\mathbf{Q}(t)$. Aussi, pour le moment, on considérera que $w_j > 0$ pour tout j . (Un cas où un des poids est négatif sera abordé un peu plus loin ; celui où un des poids est égal à 0 nous donne une droite ; finalement, celui où deux des poids sont égaux à 0 ne donne même pas une courbe.)

Ceci étant pris pour acquis, on remarque alors que le coefficient associé à chaque \mathbf{P}_j

$$\frac{w_j B_j^2(t)}{w(t)} \stackrel{\text{déf}}{=} \phi_j(t)$$

est plus grand que 0 si t appartient à l'intervalle $[0, 1]$. De plus, comme la somme de ces quantités ($\sum_{j=0}^2 \phi_j(t)$) est identiquement égale à 1, $\mathbf{Q}(t)$ est donc une combinaison barycentrique des points \mathbf{P}_j . Tout ça mis ensemble implique donc que la partie de courbe $\mathbf{Q}([0, 1])$ est incluse dans l'enveloppe convexe formée par les trois points de contrôle (autrement dit, $\mathbf{Q}(t)$ est bornée sur $[0, 1]$). On fera référence à cette partie de courbe sous l'appellation *partie standard*, alors que la *partie complémentaire* correspondra au reste de la courbe.

Pour ce qui est de la question de la dérivée des courbes de Bézier rationnelles, si on récrit (4.1) de la façon suivante

$$\mathbf{P}(t) = w(t)\mathbf{Q}(t)$$

et qu'on dérive le tout par rapport à t , on obtient

$$\dot{\mathbf{P}}(t) = \dot{w}(t)\mathbf{Q}(t) + w(t)\dot{\mathbf{Q}}(t),$$

et donc

$$\dot{\mathbf{Q}}(t) = \frac{1}{w(t)}[\dot{\mathbf{P}}(t) - \dot{w}(t)\mathbf{Q}(t)]. \quad (4.2)$$

On obtient alors facilement le résultat suivant :

$$\dot{\mathbf{Q}}(0) = \frac{2w_1}{w_0}(\mathbf{P}_1 - \mathbf{P}_0) \quad \dot{\mathbf{Q}}(1) = \frac{2w_1}{w_2}(\mathbf{P}_2 - \mathbf{P}_1).$$

Ainsi, les points de contrôle \mathbf{P}_j jouent le même rôle que le faisaient les \mathbf{b}_j précédemment : ils indiquent la direction prise par les tangentes aux extrémités de la partie standard de la courbe.

Maintenant, si on suppose que $\mathbf{U} = \mathbf{P}_0 - \mathbf{P}_1$ et $\mathbf{V} = \mathbf{P}_2 - \mathbf{P}_1$ sont des vecteurs linéairement indépendants, alors n'importe quel point \mathbf{X} choisi dans le plan formé par le système $(\mathbf{P}_1; \mathbf{U}, \mathbf{V})$ peut être écrit comme

$$\begin{aligned} \mathbf{X} &= \mathbf{P}_1 + \alpha\mathbf{U} + \beta\mathbf{V} \\ &= \alpha\mathbf{P}_0 + (1 - \alpha - \beta)\mathbf{P}_1 + \beta\mathbf{P}_2. \end{aligned} \quad (4.3)$$

En particulier, pour chaque \mathbf{X} appartenant à la partie standard de la Bézier rationnelle, il existe t , α et β tels que

$$\mathbf{Q}(t) = \sum_{j=0}^2 \phi_j(t)\mathbf{P}_j = \mathbf{X} = \alpha\mathbf{P}_0 + (1 - \alpha - \beta)\mathbf{P}_1 + \beta\mathbf{P}_2,$$

c'est-à-dire qu'on a la relation suivante entre les coordonnées de \mathbf{X} dans le système $(\mathbf{P}_1; \mathbf{U}, \mathbf{V})$ et les coordonnées barycentriques du même point par rapport à $\mathbf{P}_0, \mathbf{P}_1$ et \mathbf{P}_2 :

$$\alpha(t) = \phi_0(t) \quad \beta(t) = \phi_2(t) \quad 1 - \alpha(t) - \beta(t) = \phi_1(t).$$

Comme $B_0^2(t)B_2^2(t) = (B_1^2(t))^2/4$, on a donc

$$\begin{aligned} \alpha(t)\beta(t) &= \frac{w_0w_2}{(w(t))^2}B_0^2(t)B_2^2(t) \\ &= \frac{w_0w_2}{4(w(t))^2}(B_1^2(t))^2 \\ &= \frac{w_0w_2}{4(w_1)^2}(1 - \alpha(t) - \beta(t))^2 \quad \forall t. \end{aligned}$$

Ainsi, si l'on note la quantité $w_0w_2/4(w_1)^2$ par k , la courbe de Bézier rationnelle donnée par (4.1) peut être définie par l'équation implicite

$$\alpha\beta = k(1 - \alpha - \beta)^2. \quad (4.4)$$

On sait qu'il existe plusieurs façons, toutes équivalentes, de présenter des coniques. De façon géométrique, elles sont décrites comme des sections d'un cône de révolution par un plan, ou bien comme les projections de paraboles de \mathbb{R}^3 dans un plan. Parfois, on en parle comme étant des courbes algébriques de \mathbb{R}^2 d'ordre deux. Cette dernière « définition » et (4.4) nous permettent ainsi de dire que la courbe paramétrique $\mathbf{Q}(t)$ est une section conique.

Il est ici important de remarquer que pour une courbe donnée, le k défini ci-haut est indépendant de la paramétrisation utilisée. En effet, pour une courbe passant par \mathbf{P}_0 et \mathbf{P}_2 , et dont les tangentes en ces points se rencontrent en \mathbf{P}_1 , le système $(\mathbf{P}_1; \mathbf{U}, \mathbf{V})$ est alors totalement défini et indépendant du paramètre de la courbe décrite sous forme de Bézier rationnelle. Donc, l'équation implicite de cette courbe demeure la même, d'où on en vient à la conclusion que même si on obtient de nouveaux poids \tilde{w}_j , la valeur de $\tilde{w}_0\tilde{w}_2/4(\tilde{w}_j)^2$ demeure semblable (c'est-à-dire que k est invariant par rapport à n'importe quel changement de paramètre).

Le fait qu'on puisse obtenir (4.4) dépend un peu de la présence des coordonnées barycentriques dans le développement exposé jusqu'à maintenant. Ces mêmes coordonnées barycentriques sont également à l'origine de la prochaine remarque.

La théorie associée à ces coordonnées nous indique qu'elles ont un lien important avec la géométrie qu'elles engendrent, plus précisément avec l'aire des surfaces créées par les \mathbf{P}_j et \mathbf{X} . En effet, soit

$$\begin{aligned} a &= (\mathbf{P}_0 - \mathbf{P}_1) \times (\mathbf{P}_2 - \mathbf{P}_1) \cdot \mathbf{N} \\ a_0 &= (\mathbf{P}_2 - \mathbf{X}) \times (\mathbf{P}_1 - \mathbf{X}) \cdot \mathbf{N}/a \\ a_1 &= (\mathbf{P}_0 - \mathbf{X}) \times (\mathbf{P}_2 - \mathbf{X}) \cdot \mathbf{N}/a \\ a_2 &= (\mathbf{P}_1 - \mathbf{X}) \times (\mathbf{P}_0 - \mathbf{X}) \cdot \mathbf{N}/a \end{aligned}$$

où \mathbf{N} est une normale unitaire au plan formé par les \mathbf{P}_j . La valeur absolue de chacune de ces quantités n'est autre qu'un multiple de l'aire des triangles engendrés par les vecteurs dont on fait le produit vectoriel. Cette pondération des aires entraîne que $\sum_{j=0}^2 a_j = 1$. Maintenant, en réécrivant (4.3) de la façon suivante

$$\vec{0} = \alpha(\mathbf{P}_0 - \mathbf{X}) + (1 - \alpha - \beta)(\mathbf{P}_1 - \mathbf{X}) + \beta(\mathbf{P}_2 - \mathbf{X})$$

et en effectuant le produit vectoriel cette équation par $(\mathbf{P}_2 - \mathbf{X})$ et par $(\mathbf{P}_0 - \mathbf{X})$, on obtient alors

$$\alpha a_1 = (1 - \alpha - \beta)a_0 \quad \text{et} \quad \beta a_1 = (1 - \alpha - \beta)a_2,$$

ce qui implique que

$$\alpha = a_0, \quad 1 - \alpha - \beta = a_1 \quad \text{et} \quad \beta = a_2. \quad (4.5)$$

Ainsi, si on insère ces quantités dans (4.4), ça nous donne l'équation de la conique en coordonnées cartésiennes.

Maintenant qu'on a les instruments de base pour décrire les coniques sous forme de Bézier, on est intéressé de savoir à quel type de courbe on a affaire : une parabole, une ellipse ou une hyperbole ? Afin d'en arriver là, on doit toutefois jeter un coup d'oeil à une partie de la courbe, un peu laissée de côté jusqu'à maintenant : la partie complémentaire. Pour se faire, on pourrait étudier le comportement de $\mathbf{Q}(t)$ lorsque t est à l'extérieur de $[0, 1]$. Il sera par contre plus simple de regarder le comportement de $\tilde{\mathbf{Q}}(t)$, la courbe de Bézier rationnelle suivante :

$$\tilde{\mathbf{Q}}(t) = \frac{\tilde{w}_0 \mathbf{P}_0 B_0^2(t) + \tilde{w}_1 \mathbf{P}_1 B_1^2(t) + \tilde{w}_2 \mathbf{P}_2 B_2^2(t)}{\tilde{w}(t)} \quad (4.6)$$

où $\tilde{w}_0 = w_0$, $\tilde{w}_1 = -w_1$ et $\tilde{w}_2 = w_2$ (les mêmes $w_j > 0$ que précédemment), et où donc, $\tilde{w}(t) = \tilde{w}_0 B_0^2(t) + \tilde{w}_1 B_1^2(t) + \tilde{w}_2 B_2^2(t)$.

On remarque que les extrémités du morceau de courbe $\tilde{\mathbf{Q}}([0, 1])$ sont à nouveau \mathbf{P}_0 et \mathbf{P}_2 . Également, pour un t fixé, le point $\tilde{\mathbf{Q}}(t)$ a les quantités suivantes pour coordonnées barycentriques par rapport aux \mathbf{P}_j :

$$\begin{aligned} \tilde{\alpha}(t) &= w_0 B_0^2(t) / \tilde{w}(t) & \tilde{\beta}(t) &= w_2 B_2^2(t) / \tilde{w}(t) \\ 1 - \tilde{\alpha}(t) - \tilde{\beta}(t) &= -w_1 B_1^2(t) / \tilde{w}(t) \end{aligned}$$

Ces trois quantités permettent de satisfaire l'équation implicite (4.4), et donc, $\tilde{\mathbf{Q}}(t)$ est à l'origine de la même section conique que $\mathbf{Q}(t)$. De plus, lorsque que $\tilde{w}(t) \neq 0$, l'équation (4.6) combinée avec la définition de $\mathbf{Q}(t)$ nous indique que les trois points \mathbf{P}_1 , $\mathbf{Q}(t)$ et $\tilde{\mathbf{Q}}(t)$ sont colinéaires :

$$\tilde{\mathbf{Q}}(t) - \mathbf{P}_1 = \frac{w(t)}{\tilde{w}(t)} (\mathbf{Q}(t) - \mathbf{P}_1).$$

Selon nos hypothèses, $w(t)/\tilde{w}(t) \neq 1$ pour $t \in (0, 1)$, et donc on conclut que la fameuse partie complémentaire recherchée de notre courbe de Bézier rationnelle

initiale nous est donnée par $\tilde{\mathbf{Q}}(t)$, $0 \leq t \leq 1$. On a la relation suivante entre cette partie complémentaire et la partie standard :

$$\tilde{\mathbf{Q}}(t) = \mathbf{Q}\left(\frac{t}{2t-1}\right).$$

Evidemment, t doit être différent de $1/2$ pour que cette relation ait lieu. (Il n'existe aucun t réel tel que $\mathbf{Q}(t) = \tilde{\mathbf{Q}}(1/2)$.)

Nous avons obtenu ces quelques derniers résultats en supposant que $\tilde{w}(t) \neq 0$. Par contre, lorsque l'égalité se produit pour un t particulier, l'expression (4.6) n'est pas définie. En fait, ce t qui provoque une singularité dans la définition de $\tilde{\mathbf{Q}}(t)$, correspond à un point à l'infini de cette courbe. Or, $\tilde{w}(t)$ est un polynôme en t de degré deux, et donc, possède au plus deux racines t_1 , t_2 :

$$t_k = \frac{w_0 + w_1 \pm w_1 \sqrt{1 - 4k}}{w_0 + 2w_1 + w_2}. \quad (4.7)$$

(À noter que ces t_k appartiennent à $(0, 1)$, car pour t à l'extérieur de cet intervalle, $\tilde{w}(t) > 0$)

Suite à cette constatation, on peut maintenant effectuer le classement désiré parmi toutes les Bézier rationnelles. En effet, (4.7) nous indique que si $4k > 1$, alors la fonction $\tilde{w}(t)$ n'a pas de racines réelles et est donc bornée entre deux valeurs positives sur $[0, 1]$. Par conséquent, $\tilde{\mathbf{Q}}(t)$ est aussi bornée sur le même intervalle. Comme $\mathbf{Q}(t)$ l'était également, toute la conique est bornée et nous sommes donc en présence d'une ellipse.

Deuxièmement, si $4k = 1$, alors on a une seule racine ($t_1 = t_2$). Mais $\tilde{w}(t)$ est continue sur $[0, 1]$ et est positive aux extrémités ($t = 0, 1$) ce qui implique que cette fonction a le même signe sur tout l'intervalle $[0, 1]$. Si on se réfère aux définitions de $\tilde{\alpha}$ et $\tilde{\beta}$ ci-dessus, cela signifie donc que ces deux quantités sont positives pour tout $t \in [0, 1]$ où elles sont définies, et donc que chaque point associé $\tilde{\mathbf{Q}}(t)$ est situé dans le premier quadrant de $(\mathbf{P}_1; \mathbf{U}, \mathbf{V})$. La courbe résultante est donc une parabole. L'axe focal est donné par la droite engendrée par le segment $\overline{\mathbf{P}_1 \mathbf{Q}(t_1)}$. Finalement, lorsque $4k < 1$, $\tilde{w}(t)$ a deux racines distinctes. On supposera que $t_1 < t_2$. Ayant encore à l'esprit les caractéristiques du polynôme $\tilde{w}(t)$ mentionnées dans le paragraphe précédent, on déduit facilement qu'il est positif sur $[0, t_1)$, négatif sur (t_1, t_2) et à nouveau positif sur $(t_2, 1]$. Donc, pour ces mêmes intervalles, $\tilde{\mathbf{Q}}(t)$ passe du premier quadrant de $(\mathbf{P}_1; \mathbf{U}, \mathbf{V})$ au troisième, puis à nouveau au

premier. C'est le comportement d'une hyperbole. Les asymptotes de cette dernière sont parallèles aux droites engendrées par les segments $\overline{\mathbf{P}_1\mathbf{Q}(t_k)}$. À noter que ces droites ne croisent la conique qu'une seule fois, et ce au point $\mathbf{Q}(t_k)$.

Pour terminer ce tour d'horizon des courbes de Bézier rationnelles de degré deux, on ne peut passer sous silence le lien qui existe entre les poids et la paramétrisation utilisée. À la lumière de ce qui vient d'être dit sur la classification des coniques (sous forme de Bézier rationnelles) à partir du paramètre k , on remarque ce qui suit. Le fait de spécifier un autre point \mathbf{P}_* appartenant à la courbe (en plus de \mathbf{P}_0 et \mathbf{P}_2) détermine totalement le type de conique auquel ces trois points appartiennent. En effet, on n'a qu'à trouver les coordonnées α et β de ce point \mathbf{P}_* dans le système $(\mathbf{P}_1; \mathbf{U}, \mathbf{V})$, puis on les insère dans $\alpha\beta = k(1-\alpha-\beta)^2$, et on obtient le k qui nous permet de conclure. Par exemple, soit

$$\mathbf{P}_0 = (-1, 1) \quad \mathbf{P}_1 = (-4, 0) \quad \mathbf{P}_2 = (-1, -1)$$

ce qui implique que

$$\mathbf{U} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{et} \quad \mathbf{V} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

Soit aussi $\mathbf{P}_* = (-2, 0)$, le troisième point appartenant à la courbe désirée. Les coordonnées de ce dernier dans $(\mathbf{P}_1; \mathbf{U}, \mathbf{V})$ sont $\alpha = 1/3$ et $\beta = 1/3$ et donc

$$k = \frac{(1/3)^2}{(1 - 1/3 - 1/3)^2} = 1 \Rightarrow 4k > 1 \Rightarrow \text{cas elliptique.}$$

Par contre, pour fixer une paramétrisation (et donc, les poids w_j), il faut spécifier une caractéristique de plus. En effet, pour $0 \leq t \leq 1$, les deux Bézier rationnelles suivantes

$$\begin{aligned} \mathbf{Q}_1(t) &= (\mathbf{x}_1(t), \mathbf{y}_1(t)) \\ &= \frac{(-1, 1)(1-t)^2 + (-4, 0)t(1-t) + (-1, -1)t^2}{(1-t)^2 + t(1-t) + t^2} \end{aligned}$$

et

$$\begin{aligned} \mathbf{Q}_2(t) &= (\mathbf{x}_2(t), \mathbf{y}_2(t)) \\ &= \frac{(-1, 1)(1-t)^2 + 2(-4, 0)t(1-t) + 4(-1, -1)t^2}{(1-t)^2 + 2t(1-t) + 4t^2} \end{aligned}$$

peuvent être obtenues grâce à l'information dont on dispose jusqu'à maintenant et sont à l'origine de la même ellipse ($k = 1$), bien qu'elles n'aient pas la même paramétrisation. Pour fixer cette dernière, on « forcera » le comportement de la courbe en spécifiant le t_* tel que $\mathbf{Q}(t_*) = \mathbf{P}_*$ (dans l'exemple précédent, $t_{1*} = 1/2$ et $t_{2*} = 1/3$).

La généralisation de tout ce qui vient d'être dit sur les Bézier rationnelles de degré deux aux courbes de degré n est assez directe. Soit

$$\mathbf{R}(t) \stackrel{\text{déf}}{=} \frac{w_0 \mathbf{P}_0 B_0^n(t) + \cdots + w_n \mathbf{P}_n B_n^n(t)}{w_0 B_0^n(t) + \cdots + w_n B_n^n(t)} = \frac{\mathbf{P}(t)}{w(t)}$$

où les poids w_j appartiennent toujours à \mathbb{R} et les points \mathbf{P}_j à \mathbb{R}^3 . Lorsqu'on fait varier t , $\mathbf{R}(t)$ trace une **courbe de Bézier rationnelle de degré n** . $\mathbf{R}(t)$ est une combinaison barycentrique des $n+1$ points \mathbf{P}_j . Les poids w_j jouent également toujours un rôle de contrôleur de la forme de la courbe résultante : plus un poids est grand par rapport aux autres, plus la courbe est « attirée » par le point correspondant.

Bien qu'il n'ait pas été abordé dans le cas quadratique, il existe un processus de subdivision des courbes de Bézier rationnelles (comme pour les Bézier non rationnelles du chapitre précédent). Pour subdiviser une courbe rationnelle $\mathbf{R}(t) = \sum w_j \mathbf{P}_j B_j^n(t) / \sum w_j B_j^n(t)$ au point $c \in]0, 1[$, il suffit de considérer le numérateur et le dénominateur séparément, chacun comme une Bézier non rationnelle, et d'appliquer à chacun une des méthodes du chapitre précédent (selon le résultat désiré) afin d'obtenir de nouveaux points de contrôle. Par exemple, pour la partie de courbe $\mathbf{R}(ct)$ ($0 \leq t \leq 1$), on a :

$$\left. \begin{array}{l} \text{-numérateur : } \mathbf{c}_j = \sum_{k=0}^j w_k \mathbf{P}_k B_k^j(c) \\ \text{-dénominateur : } d_j = \sum_{k=0}^j w_k B_k^j(c). \end{array} \right\} j = 0, 1, \dots, n$$

Avec ces nouvelles quantités comme points de contrôle, on obtient une courbe $\mathbf{C}(t)$ définie sur $[0, 1]$ correspondant à une partie de $\mathbf{R}(t)$:

$$\mathbf{C}(t) = \frac{\sum \mathbf{c}_j B_j^n(t)}{\sum d_j B_j^n(t)}.$$

Pour ce qui est des dérivées, on a toujours la dérivée première

$$\dot{\mathbf{R}}(t) = \frac{1}{w(t)} [\dot{\mathbf{P}}(t) - \dot{w}(t)\mathbf{R}(t)],$$

obtenue de la même façon que pour $n = 2$, c'est-à-dire en dérivant une fois $\mathbf{P}(t) = w(t)\mathbf{R}(t)$. Si on dérive plus d'une fois cette équation, on peut même obtenir la formule récursive des dérivées d'ordre supérieur d'une courbe de Bézier rationnelle :

$$\frac{d^r}{dt^r}\mathbf{R}(t) = \frac{1}{w(t)} \left[\frac{d^r}{dt^r}\mathbf{P}(t) - \sum_{j=1}^r \binom{r}{j} \frac{d^j w(t)}{dt^j} \frac{d^{r-j}}{dt^{r-j}}\mathbf{R}(t) \right].$$

Ces dernières équations nous donnent

$$\dot{\mathbf{R}}(0) = \frac{nw_1}{w_0}\Delta\mathbf{P}_0 \quad \text{et} \quad \dot{\mathbf{R}}(1) = \frac{nw_{n-1}}{w_n}\Delta\mathbf{P}_{n-1}.$$

ainsi que

$$\ddot{\mathbf{R}}(0) = \frac{n}{w_0} \left[2w_1\left(1 - \frac{nw_1}{w_0}\right)\Delta\mathbf{P}_0 + (n-1)w_2(\mathbf{P}_2 - \mathbf{P}_0) \right]$$

et

$$\ddot{\mathbf{R}}(1) = \frac{n}{w_n} \left[2w_{n-1}\left(\frac{nw_{n-1}}{w_0} - 1\right)\Delta\mathbf{P}_{n-1} + (n-1)w_{n-2}(\mathbf{P}_{n-2} - \mathbf{P}_n) \right].$$

À partir de ces expressions de dérivées aux extrémités, et c'est avec ça qu'on va terminer cette section, on est en mesure d'exprimer la courbure des Bézier rationnelles. En effet, on se rappelle que la courbure d'une courbe paramétrée $\mathbf{x}(t)$ est $\kappa(t) = \|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}\|/\|\dot{\mathbf{x}}\|^3$, et donc, on a pour notre Bézier rationnelle de degré n

$$\kappa(0) = 2 \frac{n-1}{n} \frac{w_0 w_2}{w_1^2} \frac{\text{aire}[\Delta\mathbf{P}_0\mathbf{P}_1\mathbf{P}_2]}{\|\mathbf{P}_0\mathbf{P}_1\|^3}$$

et

$$\kappa(1) = 2 \frac{n-1}{n} \frac{w_{n-2} w_n}{w_{n-1}^2} \frac{\text{aire}[\Delta\mathbf{P}_{n-2}\mathbf{P}_{n-1}\mathbf{P}_n]}{\|\mathbf{P}_{n-1}\mathbf{P}_n\|^3}.$$

De la même façon, on obtient la torsion d'une courbe grâce à

$$\tau(t) = \frac{\det[\dot{\mathbf{x}} \ \ddot{\mathbf{x}} \ \ddot{\mathbf{x}}]}{\|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}\|^2},$$

et donc, pour la courbe paramétrée qui nous intéresse,

$$\tau(0) = \frac{n-2}{4n} \frac{w_0 w_3}{w_1 w_2} \frac{\det[(\mathbf{P}_1 - \mathbf{P}_0) (\mathbf{P}_2 - \mathbf{P}_0) (\mathbf{P}_3 - \mathbf{P}_0)]}{\text{aire}^2[\Delta\mathbf{P}_0\mathbf{P}_1\mathbf{P}_2]}$$

et

$$\tau(1) = \frac{n-2}{4n} \frac{w_{n-3}w_n}{w_{n-2}w_{n-1}} \frac{\det[(\mathbf{P}_{n-1} - \mathbf{P}_n) (\mathbf{P}_{n-2} - \mathbf{P}_n) (\mathbf{P}_{n-3} - \mathbf{P}_n)]}{\text{aire}^2[\Delta\mathbf{P}_{n-2}\mathbf{P}_{n-1}\mathbf{P}_n]}$$

Enfin, que faire pour un t_0 quelconque ? À priori, il est vrai que les différentes quantités inhérentes au calcul de la courbure et de la torsion sont plus simples aux deux extrémités de la courbe que partout ailleurs. Par contre, par subdivision de la Bézier rationnelle $\mathbf{R}(t)$ au point t_0 (en considérant $\mathbf{R}([0, t_0])$ ou $\mathbf{R}([t_0, 1])$), on obtient une nouvelle définition $\mathbf{C}(t)$ d'un morceau de la courbe initiale. Celui-ci est défini sur $[0,1]$ et en 0 ou 1, a la même valeur que $\mathbf{R}(t_0)$. On peut alors appliquer la théorie précédente à ce nouveau morceau pour trouver la courbure et la torsion de $\mathbf{R}(t)$ en t_0 .

Chapitre 5

Variétés de codimension un par courbes de niveau

1 Un peu de théorie

Lorsqu'on travaille sur des cas concrets, les formules du chapitre 2 sont relativement faciles à utiliser. Évidemment, les courbes sur lesquelles on travaille doivent être représentées paramétriquement. Le hic c'est justement que ce n'est pas toujours le cas. Il arrive en effet que la représentation paramétrique soit très difficile, voire impossible à obtenir. On doit alors se rabattre sur d'autres formes de représentation. Par exemple, on peut utiliser la notion de description locale d'un ensemble à l'aide d'un graphe. Une autre approche est celle de la notion d'ensemble C^k . Il y a aussi la notion de représentation par lignes de niveau d'un ensemble. C'est d'ailleurs le but de ce chapitre de voir comment on peut se débrouiller à partir de cette dernière façon pour obtenir certaines entités géométriques vues dans le chapitre 2. Voyons tout d'abord quelques définitions.

Soit Ω un domaine ouvert de \mathbb{R}^N . Soit aussi $C^0(\Omega)$ (ou $C(\Omega)$), l'espace des fonctions continues, définies dans Ω , à valeurs réelles. Alors, si $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{N}^N$ désigne un multi-indice, $|\alpha| = \alpha_1 + \dots + \alpha_N$ l'ordre de dérivation, et

$$\partial^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_N^{\alpha_N}},$$

on définit par récurrence pour $k \geq 1$

$$C^k(\Omega) = \{f \in C^{k-1}(\Omega) \mid \partial^\alpha f \in C^0(\Omega), \forall \alpha \ni |\alpha| = k\}.$$

D'autre part, soit λ tel que $0 < \lambda \leq 1$, alors on dit que $f \in C^{0,\lambda}(\Omega)$ (l'ensemble des fonctions $(0, \lambda)$ -Hölder continues dans Ω) si

$$\exists c > 0 \ni \forall x, y \in \Omega, |f(y) - f(x)| \leq c\|x - y\|^\lambda.$$

Si $\lambda = 1$, f porte également le nom de fonction *lipschitzienne*. Par récurrence pour $k \geq 1$, on a l'ensemble des fonctions (k, λ) -Hölder continues dans Ω :

$$C^{k,\lambda}(\Omega) = \{f \in C^{k-1,\lambda}(\Omega) \mid \partial^\alpha f \in C^{0,\lambda}(\Omega), \forall \alpha \ni |\alpha| = k\}.$$

On posera, pour $\lambda = 0$,

$$C^{k,0}(\Omega) = C^k(\Omega).$$

Finalement, quand f est une fonction à valeurs dans un espace vectoriel ($f : \Omega \rightarrow \mathbb{R}^m$), les espaces correspondant seront notés à l'aide de la dimension de l'espace d'arrivée placée en exposant. Par exemple, on aura $C^{k,\lambda}(\Omega)^m$ (c'est-à-dire que chaque composante appartient à $C^{k,\lambda}(\Omega)$).

Maintenant, soit une fonction continue $f : \mathbb{R}^N \rightarrow \mathbb{R}$. À partir de f , on considère les ensembles suivants :

$$\begin{aligned} \Omega &= \{y \in \mathbb{R}^N \mid f(y) > 0\} \\ f^{-1}(0) &= \{y \in \mathbb{R}^N \mid f(y) = 0\} \end{aligned}$$

On suppose que

$$\left\{ \begin{array}{l} 1- f^{-1}(0) \neq \emptyset \\ 2- \nabla f \neq 0 \text{ dans } f^{-1}(0) \\ 3- \text{il existe un voisinage } V \text{ de } f^{-1}(0) \text{ tel que} \\ \quad f \in C^{k,l}(V) \text{ pour } k \geq 1 \text{ et } 0 \leq l \leq 1. \end{array} \right. \quad (5.1)$$

Dans ces conditions, on a

Théorème 5.1. $\partial\Omega = f^{-1}(0)$.

Démonstration. Par continuité de f , Ω est ouvert et $\mathbb{C}\Omega$ est donc fermé. Toujours par continuité de f ,

$$\overline{\Omega} \subset \{y \in \mathbb{R}^N \mid f(y) \geq 0\}, \quad \overline{\mathbb{C}\Omega} = \mathbb{C}\Omega = \{y \in \mathbb{R}^N \mid f(y) \leq 0\}$$

ce qui implique que $\partial\Omega \subset f^{-1}(0)$. Réciproquement, pour tout x appartenant à la frontière de Ω , soit

$$g(t) = f(x + t\nabla f(x)).$$

On sait qu'il existe un $\delta > 0$ tel que pour tout t , $|t| < \delta$, $x + t\nabla f(x)$ appartient à V et la fonction g est C^1 dans $(-\delta, \delta)$. Ainsi, comme $g(0) = 0$ et $g'(t) = \nabla f(x + t\nabla f(x)) \cdot \nabla f(x)$, on a

$$f(x + t\nabla f(x)) = \int_0^t \nabla f(x + s\nabla f(x)) \cdot \nabla f(x) ds.$$

Grâce à la continuité de ∇f dans V et au fait que $\nabla f(x) \neq 0$, il existe un $\delta' \in]0, \delta[$ tel que

$$\forall s \ni 0 \leq |s| \leq \delta', \quad \nabla f(x + s\nabla f(x)) \cdot \nabla f(x) \geq \frac{1}{2} |\nabla f(x)|^2 > 0.$$

Par conséquent, on a $f(x + t\nabla f(x)) > 0$ pour tout t appartenant à $]0, \delta']$, ce qui implique que n'importe quel point de $f^{-1}(0)$ peut être approximé par une suite

$$\{x_n = x + t_n \nabla f(x) \mid n \geq 1, 0 < t_n \leq \delta', t_n \rightarrow 0\} \subset \Omega$$

On a donc que $f^{-1}(0) \subset \overline{\Omega}$. Utilisant une suite de t_n négatifs, on montre que $f^{-1}(0) \subset \overline{\mathbb{C}\Omega}$, ce qui nous permet de conclure que $f^{-1}(0) \subset \partial\Omega$. \square

Maintenant, à un x arbitraire de $f^{-1}(0)$, on peut associer le vecteur unitaire $e_N(x) = \nabla f(x) / \|\nabla f(x)\|$ et les vecteurs unitaires $e_1(x), \dots, e_{N-1}(x)$ ¹, lesquels forment avec $e_N(x)$, une base orthonormée de \mathbb{R}^N dont l'origine est en x . Toujours à partir du même x , on définit la fonction $g_x : V \rightarrow \mathbb{R}^N$ suivante

$$g_x(y) = \left(\{(y - x) \cdot e_\alpha(x)\}_{\alpha=1}^{N-1}, \frac{f(y)}{\|\nabla f(x)\|} \right) \in C^{k,l}(V)^N$$

Note : à partir de maintenant, l'utilisation d'indices grecs signifiera, si non spécifié autrement, que l'expression concernée varie de 1 à N-1, alors que l'utilisation d'indices romains impliquera une variation de 1 à N.

La transposée de la matrice jacobienne de g_x est donnée par

$$\begin{aligned} Dg_x(y)^* &= [e_1(x), \dots, e_{N-1}(x), \nabla f(y) / \|\nabla f(x)\|] \\ \Rightarrow Dg_x(x)^* &= [e_1(x), \dots, e_{N-1}(x), e_N(x)] = I \\ \Rightarrow Jg_x(x) &= \det(Dg_x(x)) = \det(Dg_x(x)^*) = 1 \end{aligned}$$

Alors par le théorème des fonctions inverses, il existe un voisinage $U(x)$ de $x \subseteq V$ (aussi noté simplement U quand il est évident qu'on fait référence à ce voisinage

¹Ces vecteurs sont les mêmes que e_1, \dots, e_{N-1} de la base canonique de \mathbb{R}^N sauf qu'ils ont subi une translation et/ou une rotation, de sorte que leur origine soit en x et qu'ils soient perpendiculaires à $e_N(x)$

de x), et aussi, un voisinage W de $g_x(x) = \mathbf{0}$ tels que $1^0 g_x(U) = W$ et $2^0 g_x|_U$ a un inverse $h_x : W \rightarrow U$ appartenant à $C^{k,l}(W)^N$. De plus, pour $w \in W$ et $y = g_x^{-1}(w) = h_x(w)$, le théorème nous indique que $Dh_x(w) = Dg_x(y)^{-1}$. En développant cette dernière expression, on obtient

$$Dh_x(w) = Dh_x(g_x(y)) = [\sigma_1(y) \ \sigma_2(y) \ \dots \ \sigma_N(y)]$$

où les composantes $\sigma_i(y)$ sont

$$\begin{cases} \sigma_\alpha(y) = e_\alpha(x) - \frac{e_\alpha(x) \cdot \nabla f(y)}{\nabla f(x) \cdot \nabla f(y)} \nabla f(x) & \alpha = 1, \dots, N-1 \\ \sigma_N(y) = \frac{\|\nabla f(x)\|}{\nabla f(x) \cdot \nabla f(y)} \nabla f(x) \end{cases}$$

À partir de cette description explicite de $Dh_x(w)$, on peut définir la **base covariante** $(a_j(y))$ à un point $y \in \partial\Omega_x \stackrel{\text{déf}}{=} U(x) \cap \partial\Omega$, c'est-à-dire $y = h_x(\xi) = h_x(\xi', 0)$ avec $\xi' \in \mathbb{R}^{N-1}$ et $\xi = (\xi', 0) \in W$:

$$\begin{aligned} a_\alpha(y) &= \frac{\partial h_x}{\partial \xi_\alpha}(\xi', 0) & \alpha = 1, \dots, N-1 \\ &= e_\alpha(x) - \frac{e_\alpha(x) \cdot \nabla f(y)}{\nabla f(x) \cdot \nabla f(y)} \nabla f(x) \\ a_N(y) &= \frac{Dh_x(\xi', 0)^{-*} e_N}{\|Dh_x(\xi', 0)^{-*} e_N\|} & (-* : \text{transposée de l'inverse}) \\ &= \frac{Dg_x(y)^* e_N}{\|Dg_x(y)^* e_N\|} \\ &= \frac{\nabla f(y)}{\|\nabla f(y)\|}. & (\text{normale unitaire vers l'intérieur de } \Omega) \end{aligned}$$

Maintenant, étant donné que cette base ne possède pas nécessairement la propriété d'être orthogonale, on introduit la notion de **base contravariante** $\{a^i\}$ définie à partir de la base covariante $\{a_i\}$ de la façon suivante :

$$a^i \cdot a_j = \delta_{ij}$$

Par définition, les première, deuxième et troisième formes fondamentales **a**, **b** et **c** de l'hypersurface $\partial\Omega$ sont données par les expressions suivantes (toutes évaluées en y) :

$$a_{\alpha\beta} = a_\alpha \cdot a_\beta \quad b_{\alpha\beta} = -a_\alpha \cdot a_{N,\beta} \quad c_{\alpha\beta} = b_\alpha^\lambda b_{\lambda\beta}$$

où

$$a_{N,\beta} = \partial_\beta a_N \stackrel{\text{déf}}{=} \frac{\partial a_N}{\partial \xi_\beta}, \quad b_\alpha^\lambda \stackrel{\text{déf}}{=} a^\lambda \cdot a^\mu b_{\mu\alpha}$$

On utilisera la convention d'Einstein qui dit que l'on doit faire la somme par rapport aux indices répétés, c'est-à-dire que, par exemple, si c et d sont deux vecteurs de \mathbb{R}^N , alors $c_i d_i = c \cdot d = \sum_{i=1}^N c_i d_i$.

Regardons maintenant plus en détail ce qu'est cette deuxième forme fondamentale \mathbf{b} lorsque le domaine Ω est décrit par les courbes de niveau d'une fonction f vérifiant les conditions (5.1).

On sait, grâce à la théorie existante, que pour un y fixé, les $N-1$ valeurs propres λ_γ de $b_{\alpha\beta}$ sont les $N-1$ courbures principales κ_γ de $\partial\Omega$ ($1 \leq \gamma \leq N-1$) en ce y . On se souvient qu'étant donné que ces κ_γ sont une mesure du comportement du vecteur normal à l'hypersurface $\partial\Omega$, leur signe (+/-) dépend *du choix de l'orientation* de ce vecteur :

- $$\left\{ \begin{array}{l} \bullet \text{ si } \kappa > 0, \text{ alors dans la direction principale, celle du} \\ \text{vecteur propre associé à } \kappa, \text{ la surface } \textit{courbe} \text{ dans} \\ \text{le même sens que le vecteur normal;} \\ \bullet \text{ si } \kappa < 0, \text{ alors dans la direction principale associée} \\ \text{à } \kappa, \text{ la surface } \textit{courbe} \text{ dans le sens opposé} \\ \text{au vecteur normal;} \\ \bullet \text{ si } \kappa = 0, \text{ il faut pousser un peu plus notre analyse} \\ \text{du comportement de la surface au point considéré.} \end{array} \right. \quad (5.2)$$

En ce qui nous concerne, comme mentionné ci-dessus, le vecteur normal considéré (a_N) sera toujours celui se dirigeant vers l'intérieur de Ω :

$$a_N(y) = a_N(h(\xi)) = \frac{\nabla f(h(\xi))}{\|\nabla f(h(\xi))\|}.$$

Donc

$$a_{N,\beta} = \left(\frac{\partial(a_N)_1}{\partial\xi_\beta} \quad \frac{\partial(a_N)_2}{\partial\xi_\beta} \quad \dots \quad \frac{\partial(a_N)_N}{\partial\xi_\beta} \right)$$

et si l'on adopte les notations suivantes

$$\begin{aligned} \partial_i f &= \frac{\partial f}{\partial y_i} & \partial_{ij} f &= \frac{\partial^2 f}{\partial y_i \partial y_j} \\ F &= \frac{\nabla f}{\|\nabla f\|} : \mathbb{R}^N \rightarrow \mathbb{R}^N \end{aligned}$$

alors on a

$$\frac{\partial(a_N)_i}{\partial\xi_\beta} = \frac{\partial(F \circ h(\xi))_i}{\partial\xi_\beta} = \partial_j \left(\frac{\partial_i f}{\|\nabla f\|} \right) \circ h(\xi) \partial_\beta h_j(\xi) \quad i = 1, \dots, N.$$

Si on regarde le tout globalement, c'est-à-dire les vecteurs $a_{N,1}, a_{N,2}, \dots, a_{N,N-1}$ comme étant les vecteurs colonnes de la matrice Da_N de dimension $N \times (N-1)$, alors le $i\beta$ -élément de cette dernière est

$$(Da_N)_{i\beta} = (DF \circ h(\xi))_{ij} (Dh)_{j\beta}$$

On remarque la présence de deux matrices jacobiniennes dans l'expression de droite, et en particulier, on se souvient que

$$Dh(\xi) = \left[(Dh)' \quad \frac{\|\nabla f(x)\|}{\nabla f(x) \cdot \nabla f(y)} \nabla f(x) \right]$$

où $(Dh)'$ est la matrice $N \times N-1$ dont les colonnes sont les vecteurs $a_\alpha(y)$.

Maintenant, si on revient aux éléments de DF , aux $\partial_j \left(\frac{\partial_i f}{\|\nabla f\|} \right)$, et qu'on développe ces expressions, on obtient

$$\begin{aligned} \partial_j \left(\frac{\partial_i f}{\|\nabla f\|} \right) &= \frac{\partial_{ij} f \|\nabla f\| - \partial_i f \partial_j \|\nabla f\|}{\|\nabla f\|^2} \\ &= \frac{\partial_{ij} f \|\nabla f\|^2 - \partial_i f \sum_{l=1}^N \partial_l f \partial_{lj} f}{\|\nabla f\|^3} \\ &= \frac{\sum_{l=1}^N [(\partial_l f)^2 (\partial_{ij} f) - (\partial_l f) (\partial_l f) (\partial_{ij} f)]}{\|\nabla f\|^3}. \end{aligned}$$

On peut aussi récrire le tout sous notation matricielle :

$$\begin{aligned} (DF)_{ij} &= \frac{1}{\|\nabla f\|^3} [\|\nabla f\|^2 (D^2 f)_{ij} - (\nabla f)_i (D^2 f \nabla f)_j] \\ &= \frac{1}{\|\nabla f\|^3} [\nabla f \cdot \{(D^2 f)_{ij} \nabla f - (\nabla f)_i (D^2 f) e_j\}] \\ \Rightarrow DF &= \frac{1}{\|\nabla f\|^3} [\|\nabla f\|^2 D^2 f - \nabla f (D^2 f \nabla f)^*] \\ \Rightarrow Da_N &= \left(\frac{1}{\|\nabla f\|} [D^2 f - a_N (D^2 f a_N)^*] \right) \Big|_{h(\xi)} (Dh)' \end{aligned}$$

Donc, on peut obtenir la deuxième forme fondamentale **b** de la façon suivante :

$$\begin{aligned} b_{\alpha\beta}(y) &\stackrel{\text{déf}}{=} (\mathbf{b})_{\alpha\beta}(y) = -a_\alpha(y) \cdot a_{N,\beta}(h(\xi)) \\ &= -a_\alpha(y) \cdot DF \circ h(\xi) a_\beta(y) \end{aligned}$$

car $a_{N,\beta}$ est la $\beta^{\text{ième}}$ colonne de Da_N .

Maintenant, au lieu de regarder ce qui se passe en ce y quelconque appartenant à la fois à $\partial\Omega$ et à un voisinage U de x , on va effectuer notre analyse directement en x . En particulier, la base covariante devient

$$\begin{aligned} a_\alpha(x) &= e_\alpha(x) - \frac{e_\alpha(x) \cdot \nabla f(x)}{\nabla f(x) \cdot \nabla f(x)} \nabla f(x) \\ &= e_\alpha(x) - (e_\alpha(x) \cdot e_N(x)) e_N(x) \\ &= e_\alpha(x) \end{aligned}$$

$$\text{et } a_N(x) = e_N(x).$$

Et donc, si $e_j(x) = \{e_j^k(x)\}_{k=1}^N$, alors la deuxième forme fondamentale \mathbf{b} devient

$$\begin{aligned} b_{\alpha\beta}(x) &= -e_\alpha(x) \cdot DF(x)e_\beta(x) \\ &= -(e_\alpha(x))^* DF(x)e_\beta(x), \end{aligned}$$

c'est-à-dire que

$$\mathbf{b}(x) = -E^* DF(x)E$$

où E désigne la matrice $N \times (N-1)$ dont les colonnes sont les vecteurs $e_\alpha(x)$. Or,

$$DF(x) = \frac{D^2 f(x)}{\|\nabla f(x)\|} - e_N(x)(D^2 f(x)e_N(x))^*,$$

et donc,

$$(e_\alpha(x))^* DF(x) = \frac{1}{\|\nabla f(x)\|} (e_\alpha(x))^* D^2 f(x) - \underbrace{(e_\alpha(x))^* e_N(x)}_0 (D^2 f(x)e_N(x))^*$$

ce qui implique que

$$\mathbf{b}(x) = -\frac{1}{\|\nabla f(x)\|} E^* D^2 f(x)E. \quad (5.3)$$

2 Quelques applications

Avec toute la théorie de la section précédente à notre disposition, attardons nous un peu au cas $N=2$. L'ensemble $f^{-1}(0)$ décrit alors une courbe \mathcal{C} dans le plan et $\mathbf{b}(x)$ ($x \in \mathbb{R}^2$) est une matrice 1×1 . La courbure $\kappa(x)$ de \mathcal{C} est donc donnée par l'élément de $\mathbf{b}(x)$. Par exemple, tous connaissent le résultat classique que pour un cercle de rayon r , la courbure en tout (x,y) est $\kappa \equiv \frac{1}{r}$. Voyons voir

tout de même ce qui se passe lorsqu'on emploie les notions amenées ci-dessus.

Première observation : dans la section précédente, le travail effectué était basé sur une fonction f choisie **à priori**. On se souvient que c'est à partir de cette fonction qu'on a défini Ω , $f^{-1}(0)$, etc. Par contre, ce qu'on s'apprête à faire est un brin différent. En effet, pour arriver au résultat désiré concernant notre cercle, on doit faire attention au *choix* de la fonction f . On sait évidemment que l'équation algébrique du cercle est $x^2 + y^2 = r^2$. Mais que doit-on choisir comme fonction f , $x^2 + y^2 - r^2$ ou $r^2 - x^2 - y^2$? Pour obtenir notre courbure $\frac{1}{r}$, on doit choisir la fonction dont le gradient pointera vers le centre du cercle (Ω se doit donc d'être l'intérieur du disque). Voilà pourquoi le processus de la section précédente est un peu biaisé, bien que si l'on aborde ce problème en ayant en tête (5.2), on a plus vraiment besoin de s'arrêter au choix de f .

Cette remarque terminée, considérons finalement

$$f(x, y) = r^2 - x^2 - y^2.$$

Comme

$$\nabla f(x, y) = \begin{pmatrix} -2x \\ -2y \end{pmatrix}$$

est dirigé vers l'intérieur du cercle pour un $(x, y) \in f^{-1}(0)$, on a

$$e_2(x, y) = \frac{1}{r} \begin{pmatrix} -x \\ -y \end{pmatrix},$$

ce qui implique qu'on peut prendre

$$e_1(x, y) = \frac{1}{r} \begin{pmatrix} y \\ -x \end{pmatrix}.$$

Finalement, on a

$$D^2 f = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

et donc, comme prévu,

$$\begin{aligned} \mathbf{b}(x, y) &= -\frac{1}{2r^3} \begin{pmatrix} y & -x \end{pmatrix} \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} y \\ -x \end{pmatrix} \\ &= \frac{1}{r}. \end{aligned}$$

De façon un peu plus générale, si on part avec

$$f(x, y) = ax^2 + 2bxy + cy^2 + dx + ey + f,$$

alors $f(x, y) = 0$ désigne une conique. Pour cette fonction f , on a

$$\nabla f(x, y) = \begin{pmatrix} 2ax + 2by + d \\ 2bx + 2cy + e \end{pmatrix},$$

(considéré en un point où il est différent de 0)

$$\Rightarrow \begin{cases} e_2(x, y) = \frac{1}{\|\nabla f(x, y)\|} \begin{pmatrix} 2ax + 2by + d \\ 2bx + 2cy + e \end{pmatrix} \\ e_1(x, y) = \frac{1}{\|\nabla f(x, y)\|} \begin{pmatrix} -2bx - 2cy - e \\ 2ax + 2by + d \end{pmatrix} \end{cases}$$

et

$$D^2 f(x, y) = \begin{pmatrix} 2a & 2b \\ 2b & 2c \end{pmatrix}.$$

En bout de ligne, après simplifications, on obtient donc

$$\kappa(x, y) = 2 \frac{-4f(b^2 - ac) + 2bde - ae^2 - cd^2}{\left((2(ax + by) + d)^2 + (2(bx + cy) + e)^2 \right)^{3/2}}. \quad (5.4)$$

On notera que la dernière formule, bien que générale, demeure d'utilisation relativement fastidieuse. Peut-on simplifier les choses ? On peut obtenir une réponse positive à cette question si l'on regarde un peu plus attentivement les courbes avec lesquelles on travaille. En effet, on sait grâce à la théorie des coniques que le discriminant $b^2 - ac$ nous permet de classier les différentes courbes auxquelles on a affaire :

$$b^2 - ac \begin{cases} > 0 & \Rightarrow \{(x, y) \mid f(x, y) = 0\} \text{ décrit une hyperbole} \\ = 0 & \Rightarrow \{(x, y) \mid f(x, y) = 0\} \text{ décrit une parabole} \\ < 0 & \Rightarrow \{(x, y) \mid f(x, y) = 0\} \text{ décrit une ellipse.} \end{cases}$$

Par ailleurs, la géométrie analytique plane nous dit que

$$\begin{cases} x = x' \cos \theta - y' \sin \theta + x_0 \\ y = x' \sin \theta + y' \cos \theta + y_0 \end{cases}$$

ou

$$\begin{cases} x' = (x - x_0) \cos \theta + (y - y_0) \sin \theta \\ y' = (y - y_0) \cos \theta - (x - x_0) \sin \theta \end{cases}$$

correspond à un changement de coordonnées faisant subir à la fois une rotation et une translation à un objet quelconque du plan : la nouvelle origine O' du système $x' \circ y'$ a pour coordonnées (x_0, y_0) par rapport à l'ancien système de coordonnées et l'axe des x' fait un angle θ avec l'axe des x positifs. Tout cela mis ensemble (les trois catégories de coniques et les transformations dans le plan) fait en sorte que si l'on veut discuter de la courbure des coniques de façon générale, on n'a qu'à étudier les trois cas suivants :

(i) $f_e(x, y) = 1 - p^2x^2 - q^2y^2 \quad (p, q > 0)$

(ii) $f_p(x, y) = y - px^2 \quad (p > 0)$

(iii) $f_h(x, y) = p^2x^2 - q^2y^2 - 1 \quad (p, q > 0)$

Dans le premier cas, on a une ellipse centrée en O dont l'axe principal est parallèle à l'axe Ox ou à l'axe Oy , tout dépendant de p et q . Soit $(x, y) \in f_e^{-1}(0)$. Le gradient de f_e est

$$\nabla f_e(x, y) = \begin{pmatrix} -2p^2x \\ -2q^2y \end{pmatrix}. \quad (5.5)$$

Après le calcul de toutes les autres quantités nécessaires, on obtient

$$\kappa_e(x, y) = \frac{(pq)^2}{(p^4x^2 + q^4y^2)^{3/2}}. \quad (5.6)$$

On refait les mêmes calculs pour le cas (ii), où on a une parabole d'axe parallèle à l'axe des y qui passe par O , et on trouve les formules suivantes pour un $(x, y) \in f_p^{-1}(0)$:

$$\nabla f_p(x, y) = \begin{pmatrix} -2px \\ 1 \end{pmatrix} \quad (5.7)$$

$$\kappa_p(x, y) = \frac{2p}{(1 + 4p^2x^2)^{3/2}}. \quad (5.8)$$

Finalement, pour l'hyperbole centrée en O dont l'axe principal est parallèle à l'axe des x (celle décrite par (iii)), on a

$$\nabla f_h(x, y) = \begin{pmatrix} 2p^2x \\ -2q^2y \end{pmatrix} \quad (5.9)$$

$$\kappa_h(x, y) = \frac{(pq)^2}{(p^4x^2 + q^4y^2)^{3/2}}. \quad (5.10)$$

Évidemment, on peut penser à refaire le même genre de travail dans \mathbb{R}^2 , avec des équations cubiques cette fois-là, ou dans \mathbb{R}^3 , avec des surfaces quadratiques. L'approche *générale* est alors relativement compliquée. En effet, les calculs de formules dans les cas

$$f(x, y) = ax^3 + 3bx^2y + 3cxy^2 + dy^3 + ex^2 + 2fxy + gy^2 + hx + iy + j = 0$$

et

$$f(x, y, z) = ax^2 + by^2 + cz^2 + 2(dxy + exz + fyz) + gx + hy + iz + j = 0.$$

sont assez monstrueux dû à la présence des nombreuses inconnues. Pour donner une idée, voici seulement les résultats de base nécessaires à la poursuite du calcul de κ pour la fonction cubique...

$$\begin{aligned} \nabla f(x, y) &= \begin{pmatrix} 3ax^2 + 6bxy + 3cy^2 + 2ex + 2fy + h \\ 3bx^2 + 6cxy + 3dy^2 + 2fx + 2gy + i \end{pmatrix} \\ \Rightarrow D^2 f(x, y) &= \begin{pmatrix} 6ax + 6by + 2e & 6bx + 6cy + 2f \\ 6bx + 6cy + 2f & 6cx + 6dy + 2g \end{pmatrix} \end{aligned}$$

Autrement, bien que la théorie sur ces courbes soit moins développée que celle des coniques, on peut toujours regarder ce qui se passe dans des cas particuliers (paraboloïde, ellipsoïde, etc), sans pour autant malheureusement couvrir toutes les possibilités.

Chapitre 6

Approximation par des courbes de niveau quadratiques et cubiques

Dans les deux derniers chapitres, on a fait un tour d'horizon général d'une théorie connue, celle de Bézier. Celle-ci fonctionne très bien grâce particulièrement au fait qu'on travaille avec des représentations paramétriques ; ce faisant, on se libère en quelque sorte du « carcan » des coordonnées cartésiennes. Mais carcan y a-t-il vraiment ? Peut-on aplanir les difficultés liées à leur utilisation, toujours évidemment dans le cadre de la représentation de domaines ? L'idée derrière ce chapitre est donc de voir quel genre de réponse on peut donner à cette question. Par exemple, comme avec les courbes de Bézier, on aimerait bien, à partir de certaines informations de départ, pouvoir « créer » des courbes algébriques d'une forme préétablie ... voyons voir.

Pour commencer, on va concentrer ses efforts dans \mathbb{R}^2 , avec une catégorie de courbes avec laquelle on a déjà travaillé : les coniques. On aimerait obtenir une fonction

$$\begin{aligned} f(x, y) &= ax^2 + 2bxy + cy^2 + dx + ey + f \\ \Rightarrow \quad \nabla f(x, y) &= \begin{pmatrix} 2ax + 2by + d \\ 2bx + 2cy + e \end{pmatrix} \end{aligned}$$

qui posséderait les caractéristiques suivantes :

$$f(x_1, y_1) = 0 \text{ et } f(x_2, y_2) = 0 \quad (6.1)$$

$$\nabla f(x_1, y_1) = \begin{pmatrix} s_1 \\ t_1 \end{pmatrix} \text{ et } \nabla f(x_2, y_2) = \begin{pmatrix} s_2 \\ t_2 \end{pmatrix}. \quad (6.2)$$

À partir de ces hypothèses, on obtient le système suivant

$$\begin{pmatrix} x_1^2 & 2x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & 2x_2y_2 & y_2^2 & x_2 & y_2 & 1 \\ 2x_1 & 2y_1 & 0 & 1 & 0 & 0 \\ 0 & 2x_1 & 2y_1 & 0 & 1 & 0 \\ 2x_2 & 2y_2 & 0 & 1 & 0 & 0 \\ 0 & 2x_2 & 2y_2 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ s_1 \\ t_1 \\ s_2 \\ t_2 \end{pmatrix} \quad (6.3)$$

Une fois simplifié (en supposant $x_1 \neq x_2$ et $y_1 \neq y_2$), le précédent système est équivalent à

$$\begin{pmatrix} x_1^2 & 2x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ \frac{x_2^2 - x_1^2}{y_2 - y_1} & \frac{2(x_2y_2 - x_1y_1)}{y_2 - y_1} & y_2 + y_1 & \frac{x_2 - x_1}{y_2 - y_1} & 1 & 0 \\ x_1 + x_2 & 2y_2 & \frac{-(y_2 - y_1)^2}{x_1 - x_2} & 1 & 0 & 0 \\ 0 & \frac{(x_2 - x_1)}{y_2 - y_1} & 1 & 0 & 0 & 0 \\ \frac{x_1 - x_2}{y_1 - y_2} & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \frac{t_1(y_2 - y_1)}{x_1 - x_2} \\ \frac{2(y_2 - y_1)}{(y_1 - y_2)} - \frac{(t_1 + t_2)}{2}(y_2 - y_1) \\ \frac{1}{2}[(s_1 + s_2)(x_2 - x_1) + (t_1 + t_2)(y_2 - y_1)] \end{pmatrix} \quad (6.4)$$

Pour qu'il existe une solution réelle à notre problème, il faut donc que

$$(s_1 + s_2)(x_2 - x_1) + (t_1 + t_2)(y_2 - y_1) = 0 \quad (6.5)$$

c'est-à-dire que la résultante vectorielle de l'addition des deux gradients soit perpendiculaire au vecteur dont les extrémités sont les deux points d'où sont issus

ces gradients.

Tout compte fait, ce premier résultat est assez acceptable. Par contre, comme c'est toujours le cas en science, on aimerait simplifier si c'est possible. Par où commencer ? Dû au fait que le résultat (6.4) corresponde à l'approche la plus générale possible du problème, on se doute bien que notre fameuse condition (6.5) demeurera pour n'importe quelle approche simplifiée. Par contre, en se rappelant à ce point-ci l'idée d'analyse *locale* contenue dans le chapitre 5, pourquoi ne pas maintenant essayer de faire quelque chose de similaire ? Par exemple, on pourrait ramener le tout à l'origine O , pour ensuite grâce aux translations et rotations, obtenir le résultat désiré. On cherche donc la fonction

$$f(x, y) = ax^2 + 2bxy + cy^2 + dx + ey$$

telle qu'évidemment $f(0, 0) = 0$ et

$$f(x_1, y_1) = 0 \quad \nabla f(0, 0) = \begin{pmatrix} s_0 \\ t_0 \end{pmatrix} \quad \nabla f(x_1, y_1) = \begin{pmatrix} s_1 \\ t_1 \end{pmatrix}.$$

La deuxième de ces conditions nous indique que $d = s_0$ et $e = t_0$. Il nous reste donc

$$\begin{pmatrix} x_1^2 & 2x_1y_1 & y_1^2 \\ 2x_1 & 2y_1 & 0 \\ 0 & 2x_1 & 2y_1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -s_0x_1 - t_0y_1 \\ s_1 - s_0 \\ t_1 - t_0 \end{pmatrix}$$

qui, une fois échelonné, ressemble à ceci

$$\begin{pmatrix} x_1^2 & 2x_1y_1 & y_1^2 \\ 0 & -2x_1y_1 & -2y_1^2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -(s_0x_1 + t_0y_1) \\ (s_0 + s_1)x_1 + 2t_0y_1 \\ (s_0 + s_1)x_1 + (t_0 + t_1)y_1 \end{pmatrix}$$

(À remarquer la condition nécessaire pour obtenir une solution à notre problème :

$$(s_0 + s_1)x_1 + (t_0 + t_1)y_1 = 0.)$$

Pour simplifier encore davantage, on peut même supposer que notre deuxième point est lui aussi sur l'axe des x , c'est-à-dire que $y_1 = 0$, et les coefficients de f sont alors les suivants :

$$a = -\frac{s_0}{x_1} \quad b = \frac{t_1 - t_0}{2x_1} \quad d = s_0 \quad e = t_0,$$

si on a la condition d'orthogonalité

$$s_0 + s_1 = 0.$$

Mentionnons premièrement que c n'est pas présent dans la liste. La raison est simple. Dans tous les cas mentionnés jusqu'à présent, si la condition d'orthogonalité est respectée, le système possède une variable libre nous donnant ainsi le loisir d'ajouter une hypothèse de départ régissant l'aspect géométrique de la courbe désirée. En fait, dans le dernier cas, on a plus : on peut donner n'importe quelle valeur au c sans modifier la valeur des autres coefficients. Dans les deux premiers cas, au moins un autre coefficient dépend de cette variable libre.

On note également que, malgré qu'on obtienne des formules relativement faciles, ces conditions inhérentes restreignent un peu notre champ d'hypothèses. C'est pour cette raison qu'on regarde maintenant du côté d'une approximation en fonction de la courbure.

On cherche donc toujours une fonction

$$f(x, y) = ax^2 + 2bxy + cy^2 + dx + ey \quad (\dots f(0, 0) = 0)$$

qui cette fois-ci devra respecter

$$f(x_1, 0) = 0 \quad \text{et} \quad \kappa(0, 0) = \kappa_0.$$

Par (5.4) du chapitre 5, on a

$$\kappa(0, 0) = 2 \frac{2bde - ae^2 - cd^2}{(d^2 + e^2)^{3/2}} = \kappa_0$$

La présence d'exposants dans la dernière expression nous enlève malheureusement la possibilité d'établir un système linéaire comme dans le cas des gradients. (Note : si on ajoute la condition $\kappa(0, 0) = \kappa_0$ à celles du dernier cas traité avec les gradients, on peut finalement déterminer notre c (si $d \neq 0$.) Puisqu'il en est ainsi, on va donc une fois de plus se ramener aux trois cas simples que l'on connaît et qui généralisent tout :

- (i) $f_e(x, y) = 1 - p^2(x - \frac{1}{p})^2 - q^2y^2 \quad (p, q > 0)$
- (ii) $f_p(x, y) = y - px^2 \quad (p > 0)$
- (iii) $f_h(x, y) = p^2(x - \frac{1}{p})^2 - q^2y^2 - 1 \quad (p, q > 0).$

Le choix de déplacer l'ellipse et l'hyperbole le long de l'axe des x est dû au fait qu'on veut pouvoir continuer à travailler avec les hypothèses $f_e(0,0) = 0$ et $f_h(0,0) = 0$. Ceci étant dit, si l'on suppose maintenant qu'on connaît la courbure en $(0,0)$ (κ_0), alors on a pour chaque cas :

- (i) $\kappa_e(0,0) \stackrel{\text{déf}}{=} \frac{q^2}{p} = \kappa_0$
- (ii) $\kappa_p(0,0) \stackrel{\text{déf}}{=} 2p = \kappa_0$
- (iii) $\kappa_h(0,0) \stackrel{\text{déf}}{=} \frac{q^2}{p} = \kappa_0$

On détermine donc totalement notre parabole de *référence* en spécifiant sa courbure à l'origine, tandis qu'avec l'ellipse (i) et l'hyperbole (iii), il nous reste une variable libre. Ici s'arrête mon travail avec les coniques. En effet la *contrainte temps* a fait en sorte que j'ai dû arrêter mon travail sur la question ici. Une combinaison d'hypothèses *gradient-courbure* serait certainement à étudier.

Avant de terminer ce chapitre, je veux mentionner que dans l'optique visée dans ce dernier, les courbes cubiques de \mathbb{R}^3 ont brièvement été étudiées avec la même approche que pour les coniques. Par contre, compte tenu de premiers résultats plus ou moins encourageants, je n'ai pas poussé davantage. Voici tout de même les quelques résultats auxquels je suis arrivé.

En premier, l'objectif était d'obtenir une courbe ayant la forme suivante

$$f(x, y) = ax^3 + 3bx^2y + 3cxy^2 + dy^3 + ex^2 + 2fxy + gy^2 + hx + iy + j \quad (6.6)$$

en posant comme hypothèses de départ que

$$\begin{aligned} f(x_1, y_1) = 0 \quad \text{et} \quad f(x_2, y_2) = 0, \\ \nabla f(x_1, y_1) = \begin{pmatrix} s_1 \\ t_1 \end{pmatrix} \quad \text{et} \quad \nabla f(x_2, y_2) = \begin{pmatrix} s_2 \\ t_2 \end{pmatrix}. \end{aligned}$$

Le système linéaire engendré par ces six conditions nous donne une solution comportant quatre variables libres (sur les dix du départ) en supposant $x_1 \neq x_2$ et $y_1 \neq y_2$:

$$\begin{aligned} j &= -y_1i - x_1h - y_1^2g - 2x_1y_1g - x_1^2e - y_1^3d - 3x_1y_1^2d - 3x_1^2y_1b - x_1^3a \\ i &= t_1 - 2y_1g - 2x_1f - 3y_1^2d - 6x_1y_1c - 3x_1^2b \\ &\dots \end{aligned}$$

$$\begin{aligned}
h &= t_1 \frac{y_1 - y_2}{x_2 - x_1} - \frac{(y_2 - y_1)^2}{x_2 - x_1} g - 2y_2 f - (x_1 + x_2) e \\
&\quad - \frac{y_2^3 - 3y_1^2 y_2 + 2y_1^3}{x_2 - x_1} d - \frac{3(x_2 y_2^2 + x_1 y_1^2 - 2x_1 y_1 y_2)}{x_2 - x_1} c \\
&\quad - 3y_2(x_1 + x_2)b - (x_1^2 + x_1 x_2 + x_2^2)a \\
g &= \frac{s_1(x_1 - x_2) + t_1(y_1 - y_2)}{(y_2 - y_1)^2} - \frac{2(x_1 - x_2)}{y_1 - y_2} f - \frac{(x_2 - x_1)^2}{(y_2 - y_1)^2} e \\
&\quad - \frac{-3y_1^2 y_2 + y_2^3 + 2y_1^3}{(y_2 - y_1)^2} d - \frac{3\{2x_1 y_1 - x_2(y_2 + y_1)\}}{(y_2 - y_1)^2} c \\
&\quad - \frac{3(x_1 - x_2)\{2x_1 y_1 - y_2(x_2 + x_1)\}}{(y_2 - y_1)^2} b - \frac{-3x_1^2 x_2 + 2x_1^3 + x_2^3}{(y_2 - y_1)^2} a \\
f &= \frac{s_2 - s_1}{2(y_2 - y_1)} - \frac{x_2 - x_1}{y_2 - y_1} e - \frac{3}{2}(y_1 + y_2)c \\
&\quad - \frac{3(x_2 y_2 - x_1 y_1)}{y_2 - y_1} b - \frac{3(x^2 - x_1^2)}{2(y_2 - y_1)} a \\
d &= \frac{(t_1 + t_2)(y_2 - y_1) + (s_1 + s_2)(x_2 - x_1)}{(y_2 - y_1)^3} - \frac{3(x_2 - x_1)}{y_2 - y_1} c \\
&\quad - \frac{3(x_1 - x_2)^2}{(y_2 - y_1)^2} b - \frac{(x_2 - x_1)^3}{(y_2 - y_1)^3} a.
\end{aligned}$$

Sous cette forme, les quatre variables libres mentionnées ci-avant sont a , b , c et e . Compte tenu du résultat obtenu avec deux points et deux gradients, mon idée d'aller voir ce qui se passait avec trois points et trois gradients est morte dans l'oeuf! Par contre, j'étais toutefois intéressé à savoir s'il y avait une simplification relative du premier essai dans le cas cubique : on désire une fonction comme (6.6) telle que

$$\begin{aligned}
f(0, 0) &= 0 \\
f(x_1, 0) &= 0 \quad (x_1 \neq 0) \\
\nabla f(0, 0) &= \begin{pmatrix} s_0 \\ t_0 \end{pmatrix} \\
\nabla f(x_1, 0) &= \begin{pmatrix} s_1 \\ t_1 \end{pmatrix}.
\end{aligned}$$

Une solution correspondant à ces hypothèses est la suivante :

$$a = \frac{s_0 + s_1}{x_1^2}$$

...

$$\begin{aligned}
 e &= -\frac{2s_0 + s_1}{x_1} \\
 h &= s_0 \\
 i &= t_0 \\
 f &= \frac{1}{2} \left\{ \frac{t_1 - t_0}{x_1} - 3x_1 b \right\}
 \end{aligned}$$

On a donc encore quatre variables libres (b , c , d et g). Les expressions sont toutes plus simples que dans le cas précédent, et ici, il serait intéressant d'envisager l'ajout d'un troisième point d'approximation. La question est de savoir comment choisir ce point afin qu'il soit utile. Étant donné que, contrairement au cas quadratique, on n'a pas un ensemble de candidats généralisant tout l'ensemble des cubiques, ce choix particulier n'est pas évident à priori.

Pour ce qui est de l'approche avec la courbure, les résultats du chapitre 5 en ce qui concerne les cubiques ont diminué mon intérêt à aller plus loin dans cette direction, quoique si l'on fixe la courbure en $(0, 0)$, on obtient

$$\kappa(0, 0) = \frac{2(2fhi - gh^2 - ei^2)}{(h^2 + i^2)^{3/2}}.$$

Chapitre 7

Les fonctions distance

L'objet de ce chapitre est de faire une brève présentation d'un autre outil utile pour l'analyse de forme : la fonction distance. Dans notre vie de tous les jours, les exemples sont nombreux de l'utilisation des distances : mentionner qu'il y a X kilomètres d'une ville A à une ville B, qu'une épreuve sportive se déroule sur Y verges, etc. Le simple fait de préciser ces quantités clarifie ce dont il est question dans nos esprits. C'est un peu la même chose pour la fonction distance : quelqu'un qui sait l'utiliser peut en tirer de l'information géométrique pertinente sur un domaine Ω , lorsqu'elle est associée à ce dernier. Le but visé par l'écriture des prochaines lignes est d'énoncer les principales caractéristiques de cette fonction, bien qu'il existe une théorie assez développée sur le sujet.

1 La fonction distance

Soit donc un sous-ensemble Ω de \mathbb{R}^N . La **fonction distance** d'un point x à l'ensemble Ω est définie de la façon suivante :

$$d_{\Omega}(x) \stackrel{\text{déf}}{=} \begin{cases} \inf_{y \in \Omega} \|y - x\|, & \Omega \neq \emptyset \\ +\infty, & \Omega = \emptyset \end{cases}$$

(À remarquer : la fonction distance est finie dans \mathbb{R}^N si et seulement si $\Omega \neq \emptyset$.)
Voici quelques propriétés que possède d_{Ω} :

- $d_{\Omega}(x) = 0 \iff x \in \overline{\Omega}$

- $x \mapsto d_\Omega(x)$ est uniformément Lipschitz, c'est-à-dire que

$$\forall x, y \in \mathbb{R}^N, \quad |d_\Omega(y) - d_\Omega(x)| \leq \|y - x\|$$

- d_Ω est dérivable presque partout et

$$\|\nabla d_\Omega(x)\| \leq 1 \quad \text{presque partout dans } \mathbb{R}^N$$

Afin de continuer notre survol des principales notions liées à la fonction distance, voici quelques définitions.

Définition 7.1.

- (i) Soit $\Omega \subset \mathbb{R}^N, \Omega \neq \emptyset$ et $x \in \mathbb{R}^N$. On appelle **ensemble des projections de x sur Ω** , l'ensemble

$$\Pi_\Omega(x) \stackrel{\text{déf}}{=} \{z \in \overline{\Omega} \mid \|z - x\| = d_\Omega(x)\}$$

Les éléments de $\Pi_\Omega(x)$ sont appelés **projections** sur Ω et sont représentés par $p_\Omega(x)$.

- (ii) Soit $\Omega \subset \mathbb{R}^N, \Omega \neq \emptyset$ (resp. $\mathbb{C}\Omega \neq \emptyset$), alors l'ensemble des points où la projection sur Ω (resp. $\mathbb{C}\Omega$) n'est pas unique

$$\text{Sk}_{\text{ext}}(\Omega) \stackrel{\text{déf}}{=} \{x \in \mathbb{R}^N \mid \Pi_\Omega(x) \text{ n'est pas un singleton}\}$$

$$\text{(resp. } \text{Sk}_{\text{int}}(\Omega) \stackrel{\text{déf}}{=} \{x \in \mathbb{R}^N \mid \Pi_{\mathbb{C}\Omega}(x) \text{ n'est pas un singleton}\})$$

est appelé le **squelette extérieur** (resp. **intérieur**). L'union des squelettes intérieur et extérieur est appelé le **squelette de Ω** . De plus, pour $x \in \partial\Omega$, $\Pi_\Omega(x)$ et $\Pi_{\mathbb{C}\Omega}(x)$ sont des singletons. Finalement, $\text{Sk}_{\text{ext}}(\Omega) \subset \mathbb{C}\overline{\Omega}$ et $\text{Sk}_{\text{int}}(\Omega) \subset \text{int}\Omega$.

- (iii) Soit $f : \mathbb{R}^N \rightarrow \mathbb{R}$. Alors on dit que f est différentiable au sens de **Gâteaux** en x si la dérivée directionnelle de f existe, c'est-à-dire

$$\forall v \in \mathbb{R}^N, \quad df(x; v) = \lim_{t \searrow 0} \frac{f(x + tv) - f(x)}{t} \text{ existe}$$

et l'application $v \mapsto df(x; v)$ est linéaire.

- (iv) Soit $f : D \subset \mathbb{R}^N \rightarrow \mathbb{R}$. Alors on dit que f est **localement lipschitzienne** si f est uniformément lipschitzienne sur n'importe quel ensemble borné K de D , c'est-à-dire que

$$\forall y, z \in K, \quad |f(z) - f(y)| \leq c\|z - y\|.$$

Nous allons maintenant voir comment la dérivée directionnelle du carré de la fonction distance peut être calculée.

Théorème 7.1. *Soit $\Omega \subset \mathbb{R}^N$ ($\Omega \neq \emptyset$) et $x \in \mathbb{R}^N$. Définissons*

$$f(x) = \frac{1}{2} \left(\|x\|^2 - d_\Omega^2(x) \right).$$

Alors,

(i) *L'ensemble $\Pi_\Omega(x)$ est non vide, compact et*

$$\forall x \in \overline{\mathbb{C}\Omega}, \Pi_\Omega(x) \subset \partial\Omega \text{ et } \forall x \in \overline{\Omega}, \Pi_\Omega(x) = \{x\}$$

(ii) *Pour tout x et v dans \mathbb{R}^N*

$$\begin{aligned} dd_\Omega^2(x; v) &= \lim_{t \searrow 0} \frac{d_\Omega^2(x + tv) - d_\Omega^2(x)}{t} = \min_{z \in \Pi_\Omega(x)} 2(x - z) \cdot v \\ df(x; v) &= \lim_{t \searrow 0} \frac{f(x + tv) - f(x)}{t} = \sup_{z \in \Pi_\Omega(x)} z \cdot v \end{aligned}$$

(iii) *$d_\Omega^2(x)$ est différentiable (au sens de Fréchet) en x si et seulement si elle est différentiable au sens de Gâteaux en x , ce qui se produit si et seulement si $\Pi_\Omega(x)$ est un singleton. En particulier, partout dans $\overline{\Omega}$, d_Ω^2 est différentiable et $\nabla d_\Omega^2 = 0$. Finalement, quand d_Ω^2 est différentiable en x ,*

$$p_\Omega(x) = x - \frac{1}{2} \nabla d_\Omega^2(x).$$

Démonstration de (iii). De (ii) de ce même théorème, on tire que $d_\Omega^2(x)$ est Gâteaux si et seulement si $\Pi_\Omega(x)$ est un singleton. Pour la différentiabilité au sens de Fréchet, on se sert d'une partie du théorème de Rademaker :

Lemme 7.1. *Soit $f : \mathbb{R}^N \rightarrow \mathbb{R}$ une fonction localement lipschitzienne. Alors f est différentiable en x , c'est-à-dire qu'on peut calculer le gradient de f et que*

$$\lim_{y \rightarrow x} \frac{|f(y) - f(x) - \nabla f(x) \cdot (y - x)|}{\|y - x\|} = 0$$

si et seulement si f est différentiable au sens de Gâteaux en x .

Démonstration du lemme. (\Rightarrow) Soit $v \neq \mathbf{0}$ appartenant à \mathbb{R}^N et $t \in \mathbb{R}$

($t > 0$). On pose $y = x + tv$ et donc $y - x = tv$. Comme f est différentiable en x

$$\lim_{t \rightarrow 0} \frac{|f(x + tv) - f(x) - \nabla f(x) \cdot tv|}{\|tv\|} = 0.$$

Comme v est différent de $\mathbf{0}$, on peut réécrire

$$\lim_{t \rightarrow 0} \left| \frac{f(x + tv) - f(x)}{t} - \nabla f(x) \cdot v \right| = 0$$

et donc

$$\lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t} = \nabla f(x) \cdot v$$

qui est exactement la définition de la dérivée directionnelle de f en x dans la direction v

$$df(x; v) = \nabla f(x) \cdot v.$$

(et pour $v = \mathbf{0}$, on a bien $df(x; v) = \nabla f(x) \cdot \mathbf{0} = 0$)

Comme l'application $v \mapsto \nabla f(x) \cdot v$ est linéaire, f est bien différentiable au sens de Gâteaux.

(\Leftarrow) Supposons que f soit Gâteaux en x . Soit L_x , l'application linéaire suivante

$$L_x : v \mapsto df(x; v).$$

Comme la fonction f est localement lipschitzienne en x , il existe donc $\delta_1 > 0$ et $c_1 > 0$ tels que f est uniformément lipschitzienne sur la boule $B(x, \delta_1)$ avec c_1 comme constante de lipschitz. Par conséquent, $\|L_x\| \leq c_1$. La sphère de rayon un $S(0, 1) = \{v \in \mathbb{R}^N \mid \|v\| = 1\}$ est compacte et peut être recouverte par la famille de boules ouvertes $\{B(v, \epsilon/(4c_1)) \mid v \in S(0, 1)\}$ pour un $\epsilon > 0$ quelconque. Ainsi, il existe un recouvrement fini $B(v_n, \epsilon/(4c_1))\}_{n=1}^N$ de $S(0, 1)$, et donc, pour n'importe quel v appartenant à $S(0, 1)$, il existe un n ($1 \leq n \leq N$) tel que $\|v - v_n\| \leq \epsilon/(4c_1)$. Soit maintenant la fonction

$$g(x, v, t) = \frac{f(x + tv) - f(x)}{t} - L_x(v)$$

définie pour $x \in \mathbb{R}^N$, $v \in S(0, 1)$, et $t > 0$. Comme f est Gâteaux en x

$$\exists \delta, 0 < \delta < \delta_1 \quad \ni \quad \forall n (1 \leq n \leq N) \text{ et } \forall t < \delta, \quad |g(x, v_n, t)| < \epsilon/2.$$

Ainsi, pour $\|y - x\| < \delta$

$$\begin{aligned}
& \left| g\left(x, \frac{y-x}{\|y-x\|}, \|y-x\|\right) \right| \\
& \leq \left| g\left(x, v_n, \|y-x\|\right) \right| + \left| \left\| \frac{y-x}{\|y-x\|}, \|y-x\| \right\| - g\left(x, v_n, \|y-x\|\right) \right| \\
& \leq \left| g\left(x, v_n, \|y-x\|\right) \right| + 2c_1 \left\| \frac{y-x}{\|y-x\|} - v_n \right\| \\
& \leq \frac{\epsilon}{2} + 2c_1 \frac{\epsilon}{4c_1} = \epsilon
\end{aligned}$$

et donc, f est différentiable en x . □

Ceci termine la démonstration. □

Maintenant qu'on en sait davantage sur la fonction d_Ω^2 , voyons voir quelles informations elle nous apporte sur d_Ω . Comme mentionné précédemment, la fonction distance est uniformément lipschitzienne et différentiable presque partout sur \mathbb{R}^N . Donc d_Ω^2 est localement lipschitzienne et si d_Ω est différentiable en x , d_Ω^2 l'est aussi et par le Théorème 7.1, $\Pi_\Omega(x)$ contient un élément unique $p_\Omega(x)$.

Théorème 7.2. (i) Soit $\Omega \subset \mathbb{R}^N$ ($\Omega \neq \emptyset$). Si le gradient de d_Ω existe en un point $x \in \mathbb{R}^N$, alors $\Pi_\Omega(x)$ est un singleton et

$$d_\Omega(x) = \|p_\Omega(x) - x\| \text{ et } \nabla d_\Omega(x) = \begin{cases} 0, & \text{si } x \in \overline{\Omega} \\ \frac{x-p_\Omega(x)}{\|x-p_\Omega(x)\|}, & \text{si } x \notin \overline{\Omega} \end{cases}$$

(ii) Soit $\Omega \subset \mathbb{R}^N$ ($\Omega \neq \emptyset$) et $x \in \mathbb{R}^N \setminus \partial\Omega$, d_Ω est différentiable en x si et seulement si d_Ω^2 est différentiable en x .

(iii) Soit $\Omega \subset \mathbb{R}^N$ ($\Omega \neq \emptyset$),

$$\chi_{\overline{\Omega}}(x) = 1 - \|\nabla d_\Omega(x)\|, \quad \chi_{\text{int } \mathbb{C}\Omega}(x) = \|\nabla d_\Omega(x)\|$$

presque partout dans \mathbb{R}^N .

(iv) Si $\Omega \neq \emptyset$, alors

$$\text{Sk}_{\text{ext}}(\Omega) = \{x \in \mathbb{R}^N \mid \nabla d_\Omega^2(x) \neq \emptyset\} \subset \text{int } \mathbb{C}\Omega = \mathbb{C}\overline{\Omega}.$$

Si $\mathbb{C}\Omega \neq \emptyset$, alors

$$\text{Sk}_{\text{int}}(\Omega) = \{x \in \mathbb{R}^N \mid \nabla d_{\mathbb{C}\Omega}^2(x) \neq \emptyset\} \subset \text{int } \Omega$$

(v) Soit $x \in \mathbb{R}^N$, $\alpha \in [0, 1]$, $p \in \Pi_\Omega(x)$ et $x_\alpha = p + \alpha(x - p)$, alors

$$d_\Omega(x_\alpha) = \|x_\alpha - p\| = \alpha\|x - p\| = \alpha d_\Omega(x)$$

et $\forall \alpha \in [0, 1], \Pi_\Omega(p + \alpha(x - p)) \subset \Pi_\Omega(x)$.

En particulier, si $\Pi_\Omega(x)$ est un singleton, alors pour $\alpha \in [0, 1]$, $\Pi_\Omega(x_\alpha)$ est un singleton, et pour $\alpha \in]0, 1]$, $\nabla d_\Omega(x_\alpha)$ existe.

Afin d'appliquer un peu cette théorie, voyons ce que donne un exemple concret dans \mathbb{R}^2 . Considérons comme ensemble Ω , l'ensemble des points formant la courbe de Bézier d'ordre deux (une parabole), c'est-à-dire que

$$\Omega = \{\mathbf{y} \in \mathbb{R}^2 \mid \exists t \in \mathbb{R} \ni \mathbf{y} = \mathbf{b}^2(t)\}.$$

Soit $x \in \mathbb{R}^2$. Si l'on note

$$\begin{aligned} \mathbf{f}(t) &= (1-t)^2 \mathbf{b}_0 + 2t(1-t) \mathbf{b}_1 + t^2 \mathbf{b}_2 - x \\ &= (1-t)^2 (\mathbf{b}_0 - x) + 2t(1-t) (\mathbf{b}_1 - x) + t^2 (\mathbf{b}_2 - x), \end{aligned}$$

alors on a

$$d_\Omega = d_{\partial\Omega}(x) = \inf_{t \in \mathbb{R}} \|\mathbf{f}(t)\|.$$

Pour un x quelconque, le problème se résume donc maintenant à trouver le (les) t qui minimise(nt) $\|\mathbf{f}(t)\|$ parmi les valeurs qui annulent $\frac{d(\|\mathbf{f}(t)\|^2)}{dt}$. Or,

$$\frac{d(\|\mathbf{f}(t)\|^2)}{dt} = 2\mathbf{f}(t) \cdot \dot{\mathbf{f}}(t)$$

où

$$\dot{\mathbf{f}}(t) = 2[(t-1)(\mathbf{b}_0 - x) + (1-2t)(\mathbf{b}_1 - x) + t(\mathbf{b}_2 - x)].$$

Donc

$$\frac{d(\|\mathbf{f}(t)\|^2)}{dt} = 4 \begin{bmatrix} t^3(A + 4B + C - 4D + 2E - 4F) \\ + t^2(-3A - 6B + 9D - 3E + 3F) \\ + t(3A + 2B - 6D + E) \\ + 1(-A + D) \end{bmatrix}$$

où

$$\begin{aligned}
 A &= \|\mathbf{b}_0 - x\|^2 \\
 B &= \|\mathbf{b}_1 - x\|^2 \\
 C &= \|\mathbf{b}_2 - x\|^2 \\
 D &= (\mathbf{b}_0 - x) \cdot (\mathbf{b}_1 - x) \\
 E &= (\mathbf{b}_0 - x) \cdot (\mathbf{b}_2 - x) \\
 F &= (\mathbf{b}_1 - x) \cdot (\mathbf{b}_2 - x).
 \end{aligned}$$

Après simplifications et si l'on pose

$$\mathbf{v}_0 = (\mathbf{b}_0 - \mathbf{b}_1), \quad \mathbf{v}_1 = (\mathbf{b}_1 - x) \quad \text{et} \quad \mathbf{v}_2 = (\mathbf{b}_2 - \mathbf{b}_1),$$

on obtient

$$\begin{aligned}
 c_1 &= A + 4B + C - 4D + 2E - 4F = \|\mathbf{v}_0 + \mathbf{v}_2\|^2 \\
 c_2 &= -3A - 6B + 9D - 3E + 3F \\
 &= -3[(\mathbf{v}_0) \cdot (\mathbf{v}_0 + \mathbf{v}_2)] \\
 c_3 &= 3A + 2B + 6D + E \\
 &= 2\|\mathbf{v}_0\|^2 + (\mathbf{b}_0 - x) \cdot (\mathbf{v}_0 + \mathbf{v}_2) \\
 &= 3\|\mathbf{v}_0\|^2 + \mathbf{v}_0 \cdot \mathbf{v}_2 + \mathbf{v}_1 \cdot (\mathbf{v}_0 + \mathbf{v}_2) \\
 c_4 &= -A + D \\
 &= (\mathbf{b}_1 - \mathbf{b}_0) \cdot (\mathbf{b}_0 - x) \\
 &= -(\mathbf{v}_0 \cdot \mathbf{v}_1 + \|\mathbf{v}_0\|^2).
 \end{aligned}$$

On remarque que c_1 est toujours différent de 0 pour \mathbf{b}_0 , \mathbf{b}_1 et \mathbf{b}_2 non colinéaires. Il ne reste qu'à résoudre l'équation caractéristique

$$t^3 + (c_2/c_1)t^2 + (c_3/c_1)t + (c_4/c_1) = 0. \quad (7.1)$$

La théorie sur la résolution des équations algébriques de degré trois nous donne les solutions suivantes

$$\begin{cases}
 t_1 &= X + Y - \frac{1}{3} \frac{c_2}{c_1} \\
 t_2 &= -\frac{1}{2}(X + Y) - \frac{1}{3} \frac{c_2}{c_1} + \frac{1}{2}i\sqrt{3}(X - Y) \\
 t_3 &= -\frac{1}{2}(X + Y) - \frac{1}{3} \frac{c_2}{c_1} - \frac{1}{2}i\sqrt{3}(X - Y)
 \end{cases}$$

avec

$$X = \sqrt[3]{R + \sqrt{Q^3 + R^2}}, \quad Y = \sqrt[3]{R - \sqrt{Q^3 + R^2}}$$

$$Q = \frac{1}{9c_1^2} [3c_1c_3 - c_2^2], \quad R = \frac{1}{54c_1^3} [9c_1c_2c_3 - 27c_1^2c_4 - 2c_2^3].$$

La quantité $Q^3 + R^2$, qu'on appellera **discriminant**, contrôle la nature des solutions de ce problème :

- (i) si elle est positive, on a un seul t_i réel
- (ii) si elle est nulle, toutes les solutions sont réelles (au moins deux sont égales)
- (iii) si elle est négative, on a trois t_i réels et distincts.

Si on développe les quantités Q et R , en fonction de \mathbf{v}_0 , \mathbf{v}_1 et \mathbf{v}_2 , alors on obtient

$$Q = \frac{1}{3\|\mathbf{v}_0 + \mathbf{v}_2\|^4} \left[(\|\mathbf{v}_0 + \mathbf{v}_2\|^2)(\mathbf{v}_1 \cdot (\mathbf{v}_0 + \mathbf{v}_2)) + 3\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^2 \right. \\ \left. + (\mathbf{v}_0 \cdot \mathbf{v}_2)(\|\mathbf{v}_0\|^2 + \|\mathbf{v}_2\|^2 - \mathbf{v}_0 \cdot \mathbf{v}_2) \right]$$

$$= \frac{1}{3\|\mathbf{v}_0 + \mathbf{v}_2\|^4} \left[(\|\mathbf{v}_0 + \mathbf{v}_2\|^2)(\mathbf{v}_1 \cdot (\mathbf{v}_0 + \mathbf{v}_2)) + 3\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^2 \right. \\ \left. + (\mathbf{v}_0 \cdot \mathbf{v}_2)(\|\mathbf{v}_0 - \mathbf{v}_2\|^2 + \mathbf{v}_0 \cdot \mathbf{v}_2) \right]$$

et

$$R = \frac{1}{2\|\mathbf{v}_0 + \mathbf{v}_2\|^6} \left[\left((\mathbf{v}_0 \cdot \mathbf{v}_2)^2 - (\mathbf{v}_0)^2(\mathbf{v}_2)^2 \right) \left(\|\mathbf{v}_0\|^2 - \|\mathbf{v}_2\|^2 \right) \right. \\ \left. + \|\mathbf{v}_0 + \mathbf{v}_2\|^2 \left[(\mathbf{v}_0 \cdot \mathbf{v}_1)(\mathbf{v}_2 \cdot (\mathbf{v}_0 + \mathbf{v}_2)) - (\mathbf{v}_1 \cdot \mathbf{v}_2)(\mathbf{v}_0 \cdot (\mathbf{v}_0 + \mathbf{v}_2)) \right] \right].$$

Ce qui nous donne

$$Q^3 + R^2 = \frac{1}{108\|\mathbf{v}_0 + \mathbf{v}_2\|^8} \left[4\|\mathbf{v}_0 + \mathbf{v}_2\|^2(\mathbf{v}_1 \cdot (\mathbf{v}_0 + \mathbf{v}_2))^3 \right. \\ + 6(\mathbf{v}_0 \cdot \mathbf{v}_2) \left((\mathbf{v}_0 \cdot \mathbf{v}_2)^2 + 3\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^2 \right) (\mathbf{v}_1 \cdot (\mathbf{v}_0 + \mathbf{v}_2)) \\ + 18\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^2 \left((\mathbf{v}_0 \cdot \mathbf{v}_1)(\mathbf{v}_2 \cdot \mathbf{v}_1) + 2(\mathbf{v}_0 \cdot \mathbf{v}_1)^2 + 2(\mathbf{v}_2 \cdot \mathbf{v}_1)^2 \right) \\ + 3(\mathbf{v}_0 \cdot \mathbf{v}_2)^2 \left(-26(\mathbf{v}_0 \cdot \mathbf{v}_1)(\mathbf{v}_2 \cdot \mathbf{v}_1) + 5(\mathbf{v}_0 \cdot \mathbf{v}_1)^2 + 5(\mathbf{v}_2 \cdot \mathbf{v}_1)^2 \right) \\ + \dots$$

$$\begin{aligned}
& \dots + 6\|\mathbf{v}_0\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2) \left(-5(\mathbf{v}_0 \cdot \mathbf{v}_1)(\mathbf{v}_2 \cdot \mathbf{v}_1) + 2(\mathbf{v}_0 \cdot \mathbf{v}_1)^2 + 11(\mathbf{v}_2 \cdot \mathbf{v}_1)^2 \right) \\
& + 6\|\mathbf{v}_2\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2) \left(-5(\mathbf{v}_0 \cdot \mathbf{v}_1)(\mathbf{v}_2 \cdot \mathbf{v}_1) + 11(\mathbf{v}_0 \cdot \mathbf{v}_1)^2 + 2(\mathbf{v}_2 \cdot \mathbf{v}_1)^2 \right) \\
& + 27\|\mathbf{v}_0\|^4(\mathbf{v}_2 \cdot (\mathbf{v}_1 + \mathbf{v}_2))^2 \\
& + 27\|\mathbf{v}_2\|^4(\mathbf{v}_0 \cdot (\mathbf{v}_1 + \mathbf{v}_0))^2 \\
& + 12(\mathbf{v}_0 \cdot \mathbf{v}_2)^2(\|\mathbf{v}_0\|^2(\mathbf{v}_0 \cdot \mathbf{v}_1) + \|\mathbf{v}_2\|^2(\mathbf{v}_2 \cdot \mathbf{v}_1)) \\
& - 42(\mathbf{v}_0 \cdot \mathbf{v}_2)^2(\|\mathbf{v}_2\|^2(\mathbf{v}_0 \cdot \mathbf{v}_1) + \|\mathbf{v}_0\|^2(\mathbf{v}_2 \cdot \mathbf{v}_1)) \\
& + 4(\mathbf{v}_0 \cdot \mathbf{v}_2)^3(\|\mathbf{v}_0\|^2 + \|\mathbf{v}_2\|^2) \\
& - \left(27\|\mathbf{v}_0\|^4\|\mathbf{v}_2\|^4 + 18\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2)^2 + (\mathbf{v}_0 \cdot \mathbf{v}_2)^4 \right)
\end{aligned}$$

Si l'on réécrit la dernière expression en mettant les quantités contenant le point x en évidence, cela nous donne

$$\begin{aligned}
Q^3 + R^2 &= \frac{1}{108\|\mathbf{v}_0 + \mathbf{v}_2\|^8} \left[4(\mathbf{v}_1 \cdot (\mathbf{v}_0 + \mathbf{v}_2))^3 \|\mathbf{v}_0 + \mathbf{v}_2\|^2 \right. \\
& + 3(\mathbf{v}_0 \cdot \mathbf{v}_1)^2 \left(9\|\mathbf{v}_2\|^4 + 12\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^2 + 4\|\mathbf{v}_0\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2) \right. \\
& \quad \left. + 22\|\mathbf{v}_2\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2) + 5(\mathbf{v}_0 \cdot \mathbf{v}_2)^2 \right) \\
& + 3(\mathbf{v}_2 \cdot \mathbf{v}_1)^2 \left(9\|\mathbf{v}_0\|^4 + 12\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^2 + 4\|\mathbf{v}_2\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2) \right. \\
& \quad \left. + 22\|\mathbf{v}_0\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2) + 5(\mathbf{v}_0 \cdot \mathbf{v}_2)^2 \right) \\
& + 6(\mathbf{v}_0 \cdot \mathbf{v}_1)(\mathbf{v}_2 \cdot \mathbf{v}_1) \left(3\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^2 - 13(\mathbf{v}_0 \cdot \mathbf{v}_2)^2 \right. \\
& \quad \left. - 5(\mathbf{v}_0 \cdot \mathbf{v}_2)(\|\mathbf{v}_0\|^2 + \|\mathbf{v}_2\|^2) \right) \\
& + 6(\mathbf{v}_0 \cdot \mathbf{v}_1) \left(9\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^4 + 3\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2) \right. \\
& \quad \left. + 2\|\mathbf{v}_0\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2)^2 - 7\|\mathbf{v}_2\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2)^2 + (\mathbf{v}_0 \cdot \mathbf{v}_2)^3 \right) \\
& + 6(\mathbf{v}_2 \cdot \mathbf{v}_1) \left(9\|\mathbf{v}_0\|^4\|\mathbf{v}_2\|^2 + 3\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2) \right. \\
& \quad \left. + 2\|\mathbf{v}_2\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2)^2 - 7\|\mathbf{v}_0\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2)^2 + (\mathbf{v}_0 \cdot \mathbf{v}_2)^3 \right) \\
& + 4(\mathbf{v}_0 \cdot \mathbf{v}_2)^3(\|\mathbf{v}_0\|^2 + \|\mathbf{v}_2\|^2) \\
& \left. + 27\|\mathbf{v}_0\|^4\|\mathbf{v}_2\|^4 - 18\|\mathbf{v}_0\|^2\|\mathbf{v}_2\|^2(\mathbf{v}_0 \cdot \mathbf{v}_2)^2 - (\mathbf{v}_0 \cdot \mathbf{v}_2)^4 \right]
\end{aligned}$$

Malgré la présence d'une certaine symétrie dans la dernière expression, on ne peut malheureusement pas la simplifier davantage pour obtenir une formule compacte :

on est forcé à une analyse *cas par cas*.

Maintenant, si l'on revient à (7.1), il peut sembler un peu surprenant d'avoir à résoudre une équation du degré trois (la géométrie du problème nous laissant croire à l'existence d'au plus deux points minimisants). Il ne faut toutefois pas oublier que les racines de cette équation nous indiquent les points critiques de la fonction distance, et non pas nécessairement des points minimisants. Pour se convaincre de ce qui se passe vraiment, étudions un cas simple : la parabole $v = u^2$ dans le plan. Sous forme de courbe de Bézier, elle correspond au cas où on a les points de Bézier suivants

$$\mathbf{b}_0 = (-1, 1) \quad \mathbf{b}_1 = (0, -1) \quad \mathbf{b}_2 = (1, 1).$$

Si $x = (u, v)$, on a alors

$$Q = \frac{1}{24}(1 - 2v) \quad R = \frac{u}{32}$$

et donc

$$Q^3 + R^2 = \left(\frac{1}{24}(1 - 2v)\right)^3 + \left(\frac{u}{32}\right)^2.$$

La courbe formée par $Q^3 + R^2 = 0$ délimite les différents ensembles de points (u, v) selon lesquels sont caractérisées les solutions de (7.1). Par exemple, si $x = (1, 2)$, alors le discriminant est inférieur à 0. On est donc en présence d'un x engendrant trois points d'inflexion distincts de l'équation caractéristique :

$$\begin{array}{lll} t_0 = 0 & \mathbf{b}_2(t_0) = (-1, 1) & d(x, \mathbf{b}_2(t_0)) = \sqrt{5} \\ t_1 = \frac{3 + \sqrt{3}}{4} & \mathbf{b}_2(t_1) = \frac{1}{2}(1 + \sqrt{3}, 2 + \sqrt{3}) & d(x, \mathbf{b}_2(t_1)) = \frac{1}{2}\sqrt{11 - 6\sqrt{3}} \\ t_2 = \frac{3 - \sqrt{3}}{4} & \mathbf{b}_2(t_2) = \frac{1}{2}(1 - \sqrt{3}, 2 - \sqrt{3}) & d(x, \mathbf{b}_2(t_2)) = \frac{1}{2}\sqrt{11 + 6\sqrt{3}} \end{array}$$

Évidemment, $d_\Omega(x)$ est le minimum des quantités de la colonne de droite. Graphiquement, le tout est représenté par la figure 7.1.

Pour un t croissant :

- sur $] -\infty, t_0]$, la distance $d(x, \mathbf{b}_2(t))$ décroît ;
- sur $[t_0, t_2]$, elle croît ;
- elle décroît à nouveau entre t_2 et t_1 ;
- puis finalement, elle se remet à croître sur l'intervalle $[t_1, \infty[$.

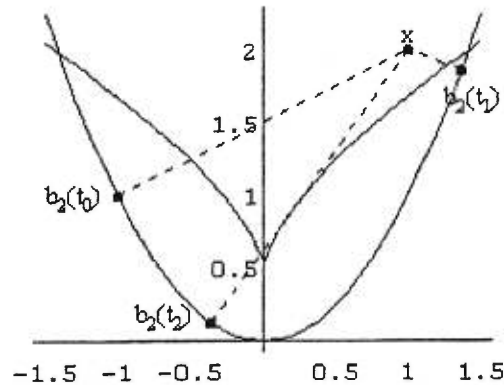


FIG. 7.1: Distance minimum du point x à la courbe $v = u^2$

2 La fonction distance orientée

La fonction distance d_Ω nous donne un assez bon portrait de ce qu'est le domaine Ω . La seule ombre au tableau est le fait que pour tout y appartenant à Ω , $d_\Omega(y) = 0$. On a toutes les raisons de croire que cette simple caractéristique nous prive de précieuses informations sur le domaine en question. C'est pour pallier à cette petite lacune qu'a été introduite la **fonction distance orientée**. On l'appelle ainsi car elle donne une **orientation** à la normale à la frontière.

Soit donc toujours le même ensemble $\Omega \subset \mathbb{R}^N$. On définit cette nouvelle fonction de la manière suivante :

$$\begin{aligned} b_\Omega(x) &\stackrel{\text{déf}}{=} d_\Omega(x) - d_{\mathbb{C}\Omega}(x), \quad \forall x \in \mathbb{R}^N \\ &= \begin{cases} d_{\partial\Omega}(x), & \text{si } x \in \overline{\mathbb{C}\Omega} \\ -d_{\partial\Omega}(x), & \text{si } x \in \overline{\Omega} \end{cases} \end{aligned}$$

De la section précédente, on se souvient que la fonction d_Ω est définie seulement pour $\Omega \neq \emptyset$. Étant donné que pour un sous-ensemble Ω de \mathbb{R}^N , $\partial\Omega \neq \emptyset$ si et seulement si $\Omega \neq \emptyset$ et $\mathbb{C}\Omega \neq \emptyset$, on en conclut donc que pour Ω différent de l'ensemble vide et de \mathbb{R}^N

$$b_\Omega(x) = \begin{cases} d_\Omega(x), & x \in \mathbb{R}^N \setminus \overline{\Omega} = \text{int } \mathbb{C}\Omega \\ 0, & x \in \overline{\Omega} \cap \overline{\mathbb{C}\Omega} = \partial\Omega \\ -d_{\mathbb{C}\Omega}(x), & x \in \mathbb{R}^N \setminus \overline{\mathbb{C}\Omega} = \text{int } \Omega. \end{cases}$$

Voici quelques propriétés de base de la fonction distance orientée. Soit A, B des sous-ensembles de \mathbb{R}^N :

- $b_A \leq b_B \iff \overline{B} \subset \overline{A}$ et $\overline{\mathcal{C}A} \subset \overline{\mathcal{C}B}$
- $b_A = b_B \iff \overline{B} = \overline{A}$ et $\overline{\mathcal{C}A} = \overline{\mathcal{C}B} \iff \overline{B} = \overline{A}$ et $\partial A = \partial B$
- $|b_A| = \max\{d_A, d_{\mathcal{C}A}\} = d_{\partial A}$
- $b_A \geq 0 \iff \overline{\mathcal{C}A} \supset \partial A \supset \overline{A} \iff \partial A = \overline{A}$
- $b_A = 0 \iff \overline{\mathcal{C}A} = \partial A = \overline{A}$.

Aussi, comme la fonction distance orientée est construite à partir de fonctions distance classiques, il est naturel qu'elle se comporte de manière similaire à ces dernières :

Théorème 7.3. *Soit Ω un sous-ensemble de \mathbb{R}^N dont la frontière est non vide.*

(i) *La fonction b_Ω est uniformément lipschitzienne et*

$$\forall x, y \in \mathbb{R}^N, |b_\Omega(y) - b_\Omega(x)| \leq \|y - x\|.$$

(ii) *Pour tout x appartenant à $\partial\Omega$, $\nabla b_\Omega^2(x)$ existe et est égal à 0. Pour tout x n'appartenant pas à $\partial\Omega$, $\nabla b_\Omega^2(x)$ existe si et seulement si $\nabla b_\Omega(x)$ existe.*

(iii) *Si $\nabla b_\Omega(x)$ existe, alors il existe un unique $p_{\partial\Omega}(x) \in \partial\Omega$ tel que*

$$b_\Omega(x) = \begin{cases} \|p_{\partial\Omega}(x) - x\|, & \text{si } x \in \text{int } \mathcal{C}\Omega \\ 0, & \text{si } x \in \partial\Omega \\ -\|p_{\partial\Omega}(x) - x\|, & \text{si } x \in \text{int } \Omega \end{cases}$$

$$\nabla b_\Omega(x) = \begin{cases} \frac{x - p_{\partial\Omega}(x)}{b_\Omega(x)}, & \text{si } x \notin \partial\Omega \\ 0, & \text{presque partout dans } \partial\Omega \end{cases}$$

et $\frac{1}{2}\nabla b_\Omega^2(x) = x - p_{\partial\Omega}(x)$ (7.2)

De plus, $\nabla b_\Omega(x) = \nabla d_{\partial\Omega}(x) = 0$ presque partout dans $\partial\Omega$.

De (ii) et (iii), on note que l'unicité de la projection $p_{\partial\Omega}(x)$ en x est directement liée à l'existence de $\nabla b_\Omega^2(x)$. En effet, (7.2) exprime formellement le lien existant entre ces deux quantités : $p_{\partial\Omega}(x)$ est le gradient de la fonctionnelle

$$f(x) = \frac{1}{2} [\|x\|^2 - b_\Omega^2(x)]. \quad (7.3)$$

Toutes les propriétés de $p_{\partial\Omega}$ peuvent ainsi être obtenues à partir de celles de ∇b_Ω^2 . Finalement, on a l'analogue du Théorème 7.1 :

Théorème 7.4. *Soit Ω , un sous-ensemble de \mathbb{R}^N dont la frontière $\partial\Omega$ est non vide, et x appartenant à \mathbb{R}^N . Alors, pour tout $v \in \mathbb{R}^N$*

$$\begin{aligned} db_{\Omega}^2(x; v) &= \inf_{z \in \Pi_{\partial\Omega}(x)} 2(x - z) \cdot v \\ &= 2 \left[x \cdot v - \sup_{z \in \Pi_{\partial\Omega}(x)} z \cdot v \right] \\ &= 2[x \cdot v - df(x; v)] \end{aligned}$$

où la fonction f mentionnée est celle définie en (7.3) et $\Pi_{\partial\Omega}(x)$ est l'ensemble des projections de x sur $\partial\Omega$ associées à $b_{\Omega}(x)$:

$$\begin{aligned} \Pi_{\partial\Omega}(x) &\stackrel{\text{déf}}{=} \begin{cases} \Pi_{\Omega}(x), & \text{si } x \in \text{int } \mathbb{C}\Omega \\ x, & \text{si } x \in \partial\Omega \\ \Pi_{\mathbb{C}\Omega}(x), & \text{si } x \in \text{int } \Omega \end{cases} \\ &= \{z \in \partial\Omega \mid \|z - x\| = d_{\partial\Omega}(x)\}. \end{aligned}$$

Conclusion

Évidemment, il n'y a pas beaucoup de commentaires nouveaux à apporter concernant la théorie déjà connue qui a été rapportée dans les pages précédentes. Pour ma part, la découverte des courbes de Bézier fut fort intéressante : elles sont simples à construire et les informations géométriques qu'elles *cachent* sont assez faciles à récupérer.

Pour ce qui est de la représentation par ensembles de niveau de sous-variétés géométriques de codimension un, on est en présence d'une théorie efficace. Il est vrai qu'à priori, elle semble donner de moins *jolis* résultats que les représentations paramétriques. Toutefois, on a constaté que certains résultats qui en découlent dans \mathbb{R}^N ($N = 2, 3$) semblent donner lieu à d'intéressantes perspectives. En particulier, l'entrée en scène de la fonction distance dans le contexte de ces ensembles de niveau nous laisse entrevoir un champ d'exploration très vaste. Peu de résultats sont toutefois disponibles à ce jour. Il faudrait en effet remplacer les approximations polynomiales par des fonctions distance. On pourrait alors associer ces dernières à des surfaces de base simples à calculer et à mettre en oeuvre, qui seraient elles, à leur tour, utilisées pour la représentation de surfaces plus complexes. D'autre part, le potentiel réel de cette approche reste encore à démontrer. Y aura-t-il des avantages importants pour l'analyse, le calcul, la mise en oeuvre ? Ceci était évidemment hors d'atteinte dans le cadre d'un mémoire de maîtrise.

Références

- [1] R. Adams, *Calcul différentiel et intégral dans l'espace*, Les Éditions Addison-Wesley Ltée, Montréal, 1989.
- [2] M. Berger et B. Gostiaux, *Differential geometry : manifolds, curves and surfaces*, Springer-Verlag, New York, 1988, (traduit de *Géométrie différentielle : variétés, courbes et surfaces*, seconde édition, Presses Universitaires de France, Paris, 1992).
- [3] W. Boehm, Smooth curves and surfaces, dans G. Farin, ed., *Geometric Modeling : Algorithms and New Trends*, pages 175-184, SIAM, Philadelphia, 1987.
- [4] R. Burden et D. Faires, *Numerical analysis* (fifth edition), PWS Publishing Company, Boston, 1993.
- [5] J. Céa, A.J. Gioan et J. Michel, *Quelques résultats sur l'identification de domaines*, pages 207-232, CALCOLO **10**, 1974.
- [6] D. Chénais, *Un résultat d'existence dans un problème d'identification de domaine*, pages 547-550, C. R. Acad. Sci. Paris Sér. A **276**, 1973.
- [7] C. de Boor, K. Höllig et M. Sabin, High accuracy geometric Hermite interpolation, *Computer Aided Geometric Design*, 4(4) :269-278, 1987.
- [8] M.C. Delfour, Shape analysis via distance functions : local theory, dans M.Delfour, ed., *Boundaries, interfaces and transitions*, pages 1-17, CRM Proceedings and Lecture Notes, AMS Publications, Providence, 1998.
- [9] ———, *Shapes and Geometries : analysis, differential calculus and optimization*, en préparation.
- [10] M.C. Delfour et J.-P. Zolésio, *Shape analysis via distance functions*, pages 129-201, J. Functional Analysis **123** 1994.
- [11] ———, *Oriented distance function in shape analysis and optimization*, dans "Control and Optimal Design of Distributed Parameter Systems", pages 39-72, J. Lagnese, D.L. Russell, L. White, eds, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1994.
- [12] J. Dugunji, *Topology*, Allyn and Bacon Inc., Boston, 1966.
- [13] G. Farin, *Curves and surfaces for computer aided geometric design*, Academic Press Inc., San Diego, 1993.
- [14] A. Gray, *Modern differential geometry of curves and surfaces*, CRC Press, Boca Raton, 1993.
- [15] E. Lee, The rationnal Bézier representation for conics, dans G. Farin, ed., *Geometric Modeling : Algorithms and New Trends*, pages 3-19, SIAM, Philadelphia, 1987.
- [16] J. Sokolowski et J.-P. Zolésio, *Introduction to shape optimization : shape sensitivity analysis*, Springer-Verlag, New York, Berlin, Heidelberg, 1992.
- [17] M.R. Spiegel, *Formules et tables de mathématiques*, McGraw-Hill Inc., Paris, 1974.