APPLICATION DE LA METHODE DE COLLOCATION RBF POUR LA RESOLUTION DE CERTAINES EQUATIONS AUX DERIVEES PARTIELLES

Antoine Filankembo Ouassissou

ACADÉMIE DE BORDEAUX

N° attribué par la Bibliothèque

THÈSE

présentée à

L'UNIVERSITÉ de PAU et des PAYS de l'ADOUR

ÉCOLE DOCTORALE DES SCIENCES EXACTES ET DE LEURS APPLICATIONS

par

Antoine FILANKEMBO OUASSISSOU

pour obtenir le grade de

DOCTEUR

Spécialité : **Mathématiques Appliquées**

Application de la Méthode de Collocation RBF pour la Résolution de Certaines Equations aux Dérivées Partielles

---

Soutenue le 5 juillet 2006

Après avis de :

Mme M.C. LOPEZ de SILANES Prof-Université de Saragosse Rap

M. A. BOUHAMIDI HDR-Université du Littoral "Côte d'Opale" Rap

Devant la commission d'examen formée des rapporteurs et de **:**

M. A. GUESSAB	Prof-Université de Pau	Directeur de Thèse
M. C. GOUT	HDR-Lab. de Mathématiques de l'INSA	
M. J.L. GOUT	Prof-Université de Pau	
M. D. SBIBIH	Prof-Université Mohammed Premier	

-2006-

Remerciements

Je tiens à exprimer toute ma profonde gratitude à Monsieur A. GUESSAB qui m'a proposé le sujet de cette recherche, m'a soutenu dans des moments difficiles et sans l'aide duquel cette thèse n'aurait pas vu le jour. Il m'a donné le goût d'aller toujours de l'avant et a mis à ma disposition un outil de travail adéquat. J'ai su profiter de ses conseils.

Je tiens à remercier Madame M.C. LOPEZ de SILANES et Monsieur A. BOUHAMIDI qui ont bien voulu accepter d'être rapporteur de cette thèse. Qu'ils trouvent ici toute ma reconnaissance.

Je remercie Monsieur J.L. GOUT et Monsieur C. GOUT pour avoir accepté de faire partie de ce jury. Je leur témoigne ma gratitude.

Je voudrais remercier également Monsieur D. SBIBIH qui a accepté de quitter son pays pour venir prendre part à ce jury. Je lui témoigne toute ma gratitude.

Je tiens aussi à remercier Monsieur R. CREFF qui m'a acceuilli au sein du Laboratoire des transferts thermiques et dont j'ai pu profiter de son incitation au travail. Je lui doit reconnaissance.

Je remercie Monsieur S. BLANCHER du LTT qui m'a semblé intéréssé à ce travail par ses questions pertinentes sur le sujet. Je lui en suis reconnaissant.

Toute ma reconnaissance s'adresse à Monsieur J. BATINA du LTT qui m'a beaucoup apporté tant dans le plan social qu'académique et avec qui j'ai travaillé sur la partie numérique ; j'ai pu profiter de son expérience.

Mes remerciements vont aussi à Monsieur M. BATCHI, à Monsieur D. MIZERE et à Monsieur C. WILSON avec qui on a fait chemin ensemble tout au long de ces années de travail dûr.

Je ne saurai oublier de remercier tous mes professeurs de DEA qui m'ont appris leur savoir faire. Je ne les oublierai jamais.

Je remercie particulièrement Mme C. BATINA qui m'a dignement accepté au sein de son foyer pour un séjour qui ne peut passer sous silence. Je lui exprime ici toute ma reconnaissance.

Je remercie toute ma famille qui se trouve au Congo-Brazzaville, et mes amis les plus proches avec qui on a passé de bons moments.

Table de Matières

pages

Introduction

La plupart des problèmes que l'on rencontre dans les phénomènes physiques utilisent des méthodes mathématiques dans leurs résolutions numériques. Ces méthodes numériques de résolution différentes les unes des autres ont depuis fort longtemps fait appel à une subdivision du domaine ou à un maillage. Il s'agit par exemple de la méthode des différences finies, des éléments finis ou des volumes finis. Plusieurs auteurs utilisant ces méthodes ont résolu avec succès de nombreux problèmes liés aux équations aux dérivées partielles. On peut par exemple citer CIARLET [5], RAVIART et THOMAS [43], GOUT [19]. Il va de soi que toutes ces méthodes dans leurs diversités utilisent des approximations numériques et des estimations à priori de la solution au problème afin de rendre compte de l'exactitude de la méthode et de contrôler les erreurs commises dans la résolution. Le but étant évidemment de minimiser ces erreurs et de trouver la meilleure approximation. On peut à ce sujet citer QUARTERONI et VALLI [42] puis GUESSAB [20].

Durant la décennie, plus précisement en 1990, KANSA [28] et [29] a introduit une nouvelle méthode numérique de résolution des équations aux dérivées partielles appelée méthode de collocation par les fonctions radiales de base(RBF). Cette méthode est basée sur la théorie d'interpolation utilisant les fonctions radiales de base et a été utilisée avec succès par les auteurs suivants : POWELL [40], LI, CHEN, PEPPER [31], FASSHAUER [12], HON [21], HON. XZ [23]. Pour des raisons de fiabilité, de simplicité et d'actualité, nous avons donc choisi d'utiliser la méthode de collocation par les RBF dans notre travail.

Dans ce travail, on présente donc une méthode numérique basée sur la quasi-interpolation et sur l'approximation par les fonctions radiales de base pour résoudre certaines équations aux dérivées partielles. Cette méthode n'exige pas une subdivision du domaine ou un maillage comme dans le cas de la méthode des différences finies ou la méthode des éléments finis ou celle des volumes finis. La méthode de la quasi-interpolation s'applique à la résolution des équations aux dirivées partielles par le principe suivant : Une équation aux dérivées partielles étant donnée, on cherche d'abord à quasi-interpoler le terme forcé de l'équation en utilisant les fonctions radiales de base. Une approximation très exacte de la solution peut alors être obtenue par résolution de l'équation fondamentale correspondante et un petit sys-

tème d'équations relatives à la condition initiale ou à la condition limite. Les exemples traités dans ce travail portent sur la dimension 1 dans le cas d'une équation différentielle du second ordre raide et dans le cas du modèle black-scholes, puis sur la dimension 2 dans le cas de la concentration d'un contaminant et dans le cas du champ classique d'un méson de masse donnée. Les estimations d'erreurs données s'appuient sur la fonction puissance de Powell construite à partir d'une classe spéciale de fonctions radiales. Dans le cas du problème raide, l'usage des multiquadriques ou une classe spéciale de fonctions radiales de base montre qu'un choix raisonnable du paramètre de forme optimale est obtenu en prenant la même valeur du paramètre perturbé contenu dans l'équation raide. La solution analytique de l'équation différentielle du second ordre raide est alors très approximativement égale à celle obtenue par quasi-interpolation. Dans le cas du champ classique d'un méson, on établit l'existence et l'unicité de la solution du problème grâce à la théorie des semi- groupes et au théorème du point fixe de Banach.

Dans ce travail nous commencerons d'adord par justifier le choix des fonctions de base utilisées dans la méthode de la quasi-interpolation afin de légitimer le bien-fondé de cette démarche. Voici donc comment est posé le problème en dimension 1 par exemple pour simplifier. On considère $\Omega = [a,b]$; $f \in \mathcal{C}^2[a,b]$ et les points $(x_i)_{0\leq i\leq N}$ tels que $a = x_0 < x_1 < ... < x_N = b$, $f(x_i) = f_i$, $i = 0, ..., N$. Par la méthode d'interpolation on peut trouver une fonction f^* qui peut approcher la fonction f et telle que $f^*(x_i) = f_i$, $i = 0, ..., N$; en résolvant le système d'inconnus $(\lambda_i)_{0\leq i\leq N}$ avec $\sum_{j=0}^{N} \lambda_j \phi_j(x_i) = f_i$, $i = 0, ..., N$ et en prenant $f^* = \sum_{j=0}^{N} \lambda_j \phi_j$ où les ϕ_j sont des fonctions radiales de base. L'idée de la quasi-interpolation est de choisir judicieusement les fonctions ϕ_i telles que $\lambda_i = f(x_i)$, $i = 0, ..., N$. On considère donc l'équation aux dérivées partielles suivante :

$$Lu(x) = f(x), \; x \in \Omega, \tag{1}$$

avec une condition limite

$$Qu(x) = g(x), \; x \in \partial\Omega. \tag{2}$$

On cherche d'abord à quasi-interpoler le terme forcé f de (1) par les fonctions de base $\left(\phi_j\right)_{0\leq j\leq N}$ telles que

$$\phi_j(x) = \phi(|x - x_j|), \forall x \in \Omega = [a, b], \tag{3}$$

où ϕ est une fonction radiale de base. La matrice résultante de la quasi-interpolation aura donc pour coefficients

$$\phi_j(x_i)_{0 \le i,j \le N}. \tag{4}$$

On voit que si N est assez grand, le problème consiste en la résolution d'un système d'équations correspondant à une matrice pleine de dimension élévée pour un choix arbitraire des fonctions radiales $\phi_j = \phi\left(\|. - x_j\|_2\right)$. La nécéssité de bien choisir les fonctions radiales s'impose donc à ce niveau. Ce choix doit donc tenir compte de :

- La structure de la matrice d'interpolation pour maîtriser le problème mal conditionné qui résulterait de l'utilisation des fonctions radiales comme interpolant global.

-Obtenir une bonne estimation d'erreur dans l'analyse numérique du problème.

-Obtenir une bonne valeur du paramètre de forme optimale en vue de la convergence et de l'efficacité de la méthode.

1) **Choix des fonctions de base** $\left(\phi_j\right)_{0 \le j \le N}$ telles que

$$\begin{aligned}
\phi_0(x) &= \frac{1}{2} + \frac{|x - x_1| - (x - x_0)}{2(x_1 - x_0)}, \\
\phi_j(x) &= \frac{|x - x_{j+1}| - |x - x_j|}{2(x_{j+1} - x_j)} - \frac{|x - x_j| - |x - x_{j-1}|}{2(x_j - x_{j-1})}, \ 1 \le j \le N - 1, \\
\phi_N(x) &= \frac{1}{2} + \frac{|x_{N-1} - x| - (x_N - x)}{2(x_N - x_{N-1})}.
\end{aligned}$$

On peut facilement vérifier que les $\left(\phi_j\right)_{0 \le j \le N}$ sont des fonctions affines par morceaux vérifiant

$$\phi_j(x_i) = \delta_{ij}, \ 0 \le i, j \le N.$$

On a donc une matrice nulle partout sauf sur la diagonale principale valant 1 ; ce qui donne une bonne structure de la matrice. La densité h des points est alors donnée par $h = \max_{1 \le j \le N}(x_j - x_{j-1})$. Pour tout $f \in \mathcal{C}^2[x_0, x_N]$ et les données $\{x_j, f_j\}_{j=0}^N$, le quasi-interpolant donné par Wu et Schaback [49] est

$$\mathcal{L}_d f(x) = f_0\alpha_0(x) + f_1\alpha_1(x) + \sum_{j=2}^{N-2} f_j\psi_j(x) + f_{N-1}\alpha_{N-1}(x) + f_N\alpha_N(x), \quad (5)$$

où

$$\begin{aligned}
\alpha_0(x) &= \tfrac{1}{2} + \frac{\phi_1(x) - (x - x_0)}{2(x_1 - x_0)}, \\
\alpha_1(x) &= \frac{\phi_2(x) - \phi_1(x)}{2(x_2 - x_1)} - \frac{\phi_1(x) - (x - x_0)}{2(x_1 - x_0)}, \\
\psi_j(x) &= \frac{\phi_{j+1}(x) - \phi_j(x)}{2(x_{j+1} - x_j)} - \frac{\phi_j(x) - \phi_{j-1}(x)}{2(x_j - x_{j-1})},\ j = 2, ..., N-2, \\
\alpha_{N-1}(x) &= \frac{(x_N - x) - \phi_{N-1}(x)}{2(x_N - x_{N-1})} - \frac{\phi_{N-1}(x) - \phi_{N-2}(x)}{2(x_{N-1} - x_{N-2})}, \\
\alpha_N(x) &= \frac{1}{2} + \frac{\phi_{N-1}(x) - (x_N - x)}{2(x_N - x_{N-1})}.
\end{aligned} \quad (6)$$

La quasi-interpolation de la formule (5) ne demande pas un calcul des dérivées des fonctions et a été prouvée comme ayant la propriété de conservation de la forme avec un ordre de convergence de $O(h^2 \ln h)$. Cette quasi-interpolation est celle utilisée par Y.C Hon [21] pour la résolution de l'équation Black-Scholes par la méthode de quasi-RBFs. Si en outre les points donnés forment une subdivision uniforme, *i.e*,

$$x_j - x_{j-1} = h, 1 \leq j \leq N,$$

alors un ordre d'approximation maximum du quasi-interpolant sur les fonctions radiales de base est obtenu.

$$\mathcal{L}_d f(x) = \sum_{j \in Z} f_j \Psi_j(x), \quad (7)$$

avec

$$\Psi_j = \frac{\psi_{j+1} - \psi_j}{2(x_{j+2} - x_{j-1})} - \frac{\psi_j - \psi_{j-1}}{2(x_{j+1} - x_{j-2})},\ j \in Z, \quad (8)$$

$$\psi_j(x) = \frac{\phi_{j+1}(x) - \phi_j(x)}{2(x_{j+1} - x_j)} - \frac{\phi_j(x) - \phi_{j-1}(x)}{2(x_j - x_{j-1})},\ j \in Z. \quad (9)$$

2) **Interêt de la méthode :** La formule (7) a l'avantage de choisir des points hors du domaine. Exemple :

$$x_{-2} < x_{-1} < x_0 < x_1 < ... < x_{N-1} < x_N < x_{N+1} < x_{N+2}.$$

On appliquera cette méthode pour la résolution du problème raide :

$$\varepsilon u^{''}(x) + u^{'}(x) = f(x), \; x \in [0,1],$$

$$u(0) = a, \; u(1) = b,$$

où ε est un petit paramètre constant. En posant $\mathcal{L}_d f = f^*$, alors (5) s'écrira

$$f^*(x) = \sum_{j=-2}^{N+2} f(x_j) B_j^1(x) \approx f(x),$$

où $\left(B_j^1\right)_{-2 \leq j \leq N+2}$ est la base de B-splines linéaires donnée par

$$\begin{aligned}
B_{-2}^1(x) &= \frac{1}{2} + \frac{|x - x_{-1}| - |x - x_{-2}|}{2(x_{-1} - x_{-2})}, \\
B_j^1(x) &= \frac{|x - x_{j+1}| - |x - x_j|}{2(x_{j+1} - x_j)} - \frac{|x - x_j| - |x - x_{j-1}|}{2(x_j - x_{j-1})}, -1 \leq j \leq N+1, \\
B_{N+2}^1(x) &= \frac{1}{2} + \frac{|x - x_{N+1}| - |x - x_{N+2}|}{2(x_{N+2} - x_{N+1})}.
\end{aligned}$$

L'ensemble de ce travail comprend ainsi quatre chapitres. Dans le premier chapitre on cherche d'abord à formuler le problème d'interpolation par des fonctions radiales dans un espace de Hilbert V, comme on peut le retrouver dans la plupart des articles traitant ce domaine. Comme références, on peut par exemple citer [33] et [34]. La motivation principale de ce chapitre est d'exprimer l'interpolant radial défini en (10) dans la base du sous-espace $W \subset V$ des interpolés, afin d'en trouver une écriture plus commode d'utilisation. Plus précisement, au lieu de

$$s_{f,X}(x) = \sum_{j=1}^{N} \lambda_j \phi\left(\|x - x_j\|_2\right) + p(x), \tag{10}$$

où X et p sont définis en page 12, $s_{f,X}$ peut s'écrire comme dans [33] sous la forme suivante

$$s_{f,X}(x) = \sum_{j=1}^{N} \mu_j K(x, x_j), \tag{11}$$

où les coefficients $\left(\mu_j\right)_{1\leq j\leq N}$ sont déterminés par la résolution du système

$$\sum_{j=1}^{N} \mu_j K(x_i., x_j) = f_i,\ 1 \leq i \leq N. \tag{12}$$

Pour cela, une théorie variationnelle dûe à Golomb Weinberger [18], nous permet d'expliciter les coefficients $(\lambda_j)_{1\leq j\leq N}$ et $(c_r)_{1\leq r\leq l}$ qui interviennent dans la formule

$$\sum_{j=1}^{N} \lambda_j \phi\left(\|x_i - x_j\|_2\right) + \sum_{r=1}^{l} c_r p_r(x_i) = f_i,\ 1 \leq i \leq N, \tag{13}$$

où $l = \dim \pi_{k-1}$ avec π_{k-1} que nous définirons par la suite. L'idée de construire une base $\{K(., x_1), ..., K(., x_N)\}$ de W nous conduit à la construction du noyau reproduisant de l'espace de Hilbert V, comme dans [33], grâce au théorème de représentation de Riesz. Nous avons apporté l'hypothèse (1.37), à savoir : Pour chaque $x \in \mathbb{R}^d$ il existe une constante $M_x > 0$ telle que

$$v(x_i) = 0,\ i = 1, ..., l \Rightarrow |v(x)| \leq M_x \sqrt{\langle v, v\rangle}, \tag{14}$$

comme étant une modification de l'hypothèse analogue considérée jusqu'à présent, c'est à dire : Pour chaque $x \in \mathbb{R}^d$ il existe une constante $M_x > 0$ telle que

$$v(x_i) = 0\ ,\ i = 1, ..., N \Rightarrow |v(x)| \leq M_x \sqrt{\langle v, v\rangle},$$

comme dans [32], [33] et [34]. Par ailleurs, par rapport à la norme $\|.\|$ définie sur V par :

$$\|v\| = \sqrt{(v, v)}\ \forall v \in V,$$

où $(.,.)$ est le produit scalaire sur V défini par

$$(u, v) = \langle u, v\rangle + \sum_{i=1}^{l} u(x_i)\, v(x_i)\ ,\ \forall\ u, v \in V,$$

on veut que $s_{f,X}$ soit la meilleure approximation de f au sens

$$\|f - s_{f,X}\| = \inf_{v\in W} \|f - v\|.$$

Soit alors $d > 0$ un entier donné et $\Omega \subset \mathbb{R}^d$ un ouvert borné de frontière $\partial\Omega$. On notera $\mathcal{C}(\Omega)$ l'ensemble des fonctions continues sur Ω à valeurs réelles et

$$\mathbb{R}_{\geq 0} = \{r \in \mathbb{R} \ : r \geq 0\}.$$

Soit $X = \{x_1, ..., x_N\} \subset \Omega$ un ensemble de points distincts de Ω. On suppose que pour toute fonction $f \in \mathcal{C}(\Omega)$, $f_j = f(x_j)$ est connue $\forall j$, $1 \leq j \leq N$. Le problème est d'interpoler la fonction f sur X par un interpolant $s_{f,X} : \Omega \to \mathbb{R}$ de la forme

$$s_{f,X}(x) = \sum_{j=1}^{N} \lambda_j \phi\left(\|x - x_j\|_2\right) + p(x), \tag{15}$$

i.e.,

$$s_{f,X}(x_i) = f(x_i), \ 1 \leq i \leq N, \tag{16}$$

où p appartient au sous espace π_{k-1} de $\mathcal{C}\left(\mathbb{R}^d\right)$ des polynômes de degré total au plus $k-1$ et

$$\phi : \mathbb{R}_{\geq 0} \to \mathbb{R},$$

est une fonction continue appelée fonction radiale. En notant $\|.\|_2$ la norme euclidienne de $\mathbb{R}^d$, le problème revient donc à trouver des réels λ_j, $1 \leq j \leq N$ tels que

$$\sum_{j=1}^{N} \lambda_j \phi\left(\|x_i - x_j\|_2\right) + p(x_i) = f_i, \ 1 \leq i \leq N. \tag{17}$$

Sous certaines hypothèses, il a été montré dans l'article de Micchelli [38] que le système (17) admet une unique solution. C'est le cas pour des choix particuliers de la fonction radiale ϕ. Les plus communément utilisées sont :

$$\text{Linear splines, } \phi(r) = r, \tag{18}$$

$$\text{Thin-plate-splines, } \phi(r) = r^{\alpha} \ln r \ (\alpha \text{ pair}), \tag{19}$$

$$\text{Fonction radiale de type conique : } \phi(r) = r^{\alpha} (\alpha \text{ impair}), \tag{20}$$

$$\text{Multiquadriques} : \phi(r) = \left(r^2 + c^2\right)^{\frac{\beta}{2}} (\beta \text{ impair}),\ c > 0, \tag{21}$$

$$\text{Gaussians},\ \phi(r) = \exp\left(-cr^2\right),\ c > 0, \tag{22}$$

où le nombre $c > 0$ sera appelé paramètre de forme.

Fonction radiale à support compact :

$$\phi(r) = \begin{cases} \left(1 - \dfrac{r}{\rho}\right)^3 \left(1 + 3\dfrac{r}{\rho} + \left(\dfrac{r}{\rho}\right)^2\right), \text{ si } 0 \leq r \leq \rho \\ 0, \text{ si } r > \rho. \end{cases} \tag{23}$$

Grâce à la théorie variationnelle de Golomb Weinberger [18], on montre que dans le cas de l'interpolation par des fonctions radiales, l'interpolé $s_{f,X}$ possède un ordre d'approximation $\mathcal{O}\left(h^l\right)$ de f, *i.e.*, il existe une constante C telle que

$$\|s_{f,X} - f\|_\infty \leq Ch^l, \tag{24}$$

où h désigne la densité des points donnée par

$$h = \sup_{x \in \Omega} \min_{1 \leq j \leq N} \|x - x_j\|_2, \tag{25}$$

et l est un nombre rationnel positif.

Dans le second chapitre on veut estimer l'erreur d'interpolation du problème formulé en (1.9) et (1.10) dans des cas particuliers de la fonction radiale ϕ. Comme dans [33] et [34], cette estimation d'erreur repose sur la fonction puissance de Powell. Les techniques utilisées dans la déduction des constantes intervenant dans les majorations d'erreurs sont celles qu'on retrouve dans [2], et nous les avons renforcées par de nombreux résultats. La déduction de ces constantes résulte de quelques applications de la méthode de la fonction puissance. Dans le cas des dimensions $d = 1, 2, 3$, et pour $k = 2$, on trouve les meilleures constantes dans les majorations d'erreurs correspondantes respectivement à :

$$\begin{aligned} \pi_1(\mathbb{R})\ ;\ \phi(r) &= r^3, \\ \pi_1\left(\mathbb{R}^2\right)\ ;\ \phi(r) &= r^2 \ln r, \\ \pi_1\left(\mathbb{R}^3\right)\ ;\ \phi(r) &= r. \end{aligned}$$

On travaille ici avec des géométries simples : Dans le premier cas ($d = 1$), x est sur un segment défini par les noeuds d'interpolation $x_1 < x_2 < ... < x_N$. Dans le second cas ($d = 2$), x est sur un segment ou à l'intérieur d'un triangle défini par trois points pris parmi $x_1, x_2..., x_N$. Dans le troisième cas ($d = 3$), x est à l'intérieur ou sur une face d'un cube ou d'un tétraèdre dont les sommets sont pris parmi les N points d'interpolation $x_1, x_2..., x_N$.

Le troisième chapitre est consacré au problème de la quasi-interpolation. Soit donc Ω un ouvert borné de $\mathbb{R}^d$ ayant la propriété de cône et de frontière $\partial\Omega$ lipschitzienne. D'après le premier chapitre, le problème d'interpolation d'une fonction $f \in \mathcal{C}(\Omega)$ sur un ensemble π_{k-1}-unisolvant $X = \{x_1,...,x_N\}$ par un interpolant $s_{f,X}$ de la forme

$$s_{f,X}(x) = \sum_{j=1}^{N} \lambda_j \phi\left(\|x - x_j\|_2\right) + p(x), \text{ où } p \in \pi_{k-1}, \tag{26}$$

peut se ramener à

$$s_{f,X}(x) = \sum_{j=1}^{N} \mu_j K(x, x_j) \text{ où } K(x, x_j) = q_j(x) \ \forall x \in \Omega, \tag{27}$$

les $(q_j)_{1 \le j \le N}$ étant ici les représentants sur V (espace fonctionnel où l'on étudie le problème d'interpolation) des points d'évalution en x_j ; *i.e.*,

$$v(x_j) = (v, q_j), \forall v \in V. \tag{28}$$

Cette méthode demande la résolution d'un système d'équations pour les coefficients inconnnus μ_j. L'idée de la quasi-interpolation que nous utiliserons pour la résolution du problème raide, est de choisir les fonctions de base ϕ_j telles que

$$f_j = s_{f,X}(x_j) = \sum_{k=1}^{N} \lambda_k \phi_k(x_j) = \lambda_j, \ 1 \le j \le N, \tag{29}$$

c'est à dire

$$f(x_j) = \lambda_j, \ 1 \le j \le N. \tag{30}$$

L'idée est d'interpoler le terme forcé de (1) en utilisant une classe spéciale de fonctions radiales de base $\phi\left(\|. - x_j\|_2\right)$. Une approximation très exacte de

la solution peut alors être obtenue par résolution de l'équation fondamentale corrrespondante et un petit système d'équations correspondant à la condition initiale ou à la condition limite. On arrive à interpoler le terme forcé f de (1) en appliquant les résultats précédents pour $N = n + m$ et $s_{f,X} = f^*$. On a donc

$$f^*(x) = \sum_{j=1}^{n+m} \lambda_j \phi\left(\|x - x_j\|_2\right), \tag{31}$$

où

$$(x_j)_{1\leq j\leq n} \subset \Omega \text{ et } (x_j)_{n+1\leq j\leq n+m} \subset \partial\Omega, \tag{32}$$

vérifient

$$f^*(x_k) = \sum_{j=1}^{n+m} \lambda_j \phi\left(\|x_k - x_j\|_2\right) = f(x_k),\ k = 1, ..., n. \tag{33}$$

La solution numérique de (1) et (2) est alors obtenue en résolvant l'équation fondamentale

$$L\Phi(x) = \phi\left(\|x\|_2\right), \tag{34}$$

et en résolvant le système

$$\sum_{j=1}^{n+m} \lambda_j Q\Phi(x_k - x_j) = g(x_k),\ k = n+1, ..., n+m. \tag{35}$$

Nous appliquerons la méthode de collocation par les RBF pour la résolution des autres problèmes traités dans ce travail. Dans l'étude du problème modèle du champ classique d'un méson, on rappelera d'abord quelques notions essentielles sur les opérateurs linéaires non-bornés et sur les semi-groupes ainsi que les résultats des théorèmes qui constituent l'outil éssentiel de l'étude du problème modèle de la mécanique quantique du champ classique d'un méson. Il s'agira donc du problème

$$(\mathcal{P}) \begin{cases} \dfrac{\partial^2 u}{\partial t^2} - \Delta u + m^2 u = -\lambda u^3 & \text{sur } \Omega \times]0, T[\ , \\ u(x,0) = u_0(x) & \text{sur } \Omega, \\ \dfrac{\partial u}{\partial t}(x,0) = v_0(x) & \text{sur } \Omega, \\ u(x,t) = 0 & \text{sur } \partial\Omega \times]0, T[, \end{cases} \tag{36}$$

où Ω est un ouvert borné de $\mathbb{R}^d$ $(d \geq 1)$ de frontière $\partial\Omega$ régulière ; $\lambda \geq 0$ est un paramètre réel. Dans le cas où $\lambda = 0$ on obtient l'équation de Klein-Gordon décrite dans la mécanique quantique relativiste. Cette équation est dans ce cas de la forme $(L + m^2)\,\psi = 0$ où l'opérateur $L = \dfrac{\partial^2}{\partial t^2} - \Delta$ et m est la masse de la particule. Le problème (36) décrit le champ classique u d'un méson de masse m. Le terme non linéaire $-\lambda u^3$ avec $\lambda > 0$, décrit le self interaction du champ. Pour $\lambda > 0$, on montre d'abord l'existence et l'unicité de la solution du problème (36) grâce à la théorie des semis-groupes, des opérateurs linéaires non bornés et au théorème du point fixe de Banach [50] (théorème 1.A). Pour la résolution numérique du problème $(\mathcal{P})$, on utilisera un schéma numérique basé sur la collocation RBF semblable à celui utilisé par Y.C. Hon [21] pour la résolution de l'équation Black-Scholes.

Le quatrième chapitre sera reservé aux tests numériques de quelques modèles et nous terminerons par une conclusion générale et des perspectives.

Chapitre 1

Théorie d'interpolation utilisant les fonctions radiales de base

Dans ce chapitre on cherche d'abord à formuler le problème d'interpolation par des fonctions radiales comme on peut le retrouver dans la plupart des articles traitant ce domaine. Comme références, on peut par exemple citer [33] et [34]. La motivation principale de ce chapitre est d'exprimer l'interpolant radial défini en (1.1) dans la base du sous-espace $W \subset V$ des interpolés, afin d'en trouver une écriture plus commode d'utilisation. Plus précisement, au lieu de

$$s_{f,X}(x) = \sum_{j=1}^{N} \lambda_j \phi\left(\|x - x_j\|_2\right) + p(x), \tag{1.1}$$

$s_{f,X}$ peut s'écrire comme dans [33] sous la forme suivante

$$s_{f,X}(x) = \sum_{j=1}^{N} \mu_j K(x, x_j), \tag{1.2}$$

où les coefficients $\left(\mu_j\right)_{1 \leq j \leq N}$ sont déterminés par la résolution du système

$$\sum_{j=1}^{N} \mu_j K(x_i, x_j) = f_i,\ 1 \leq i \leq N. \tag{1.3}$$

Pour cela, une théorie variationnelle, dûe à Golomb Weinberger [18], nous permettra d'expliciter les coefficients $(\lambda_j)_{1\leq j\leq N}$ et $(c_r)_{1\leq r\leq l}$ qui interviennent dans la formule

$$\sum_{j=1}^{N}\lambda_j\phi\left(\|x_i-x_j\|_2\right)+\sum_{r=1}^{l}c_rp_r\left(x_i\right)=f_i,\ 1\leq i\leq N. \tag{1.4}$$

Ces coefficients seront donnés par les formules

$$\lambda_r=\begin{cases}\mu_r-f\left(x_r\right),\ r=1,...,l\\ \mu_r,\ r=l+1,...,N\end{cases} \tag{1.5}$$

et

$$c_i=f\left(x_i\right)+\sum_{r=1}^{l}f\left(x_r\right)\psi_{x_r}\left(x_i\right)-\sum_{r=1}^{N}\mu_r\psi_{x_r}\left(x_i\right),\ i=1,...,l\,; \tag{1.6}$$

où

$$\psi_z\left(y\right)=\phi\left(\|y-z\|_2\right),\ \forall y,z\in\mathbb{R}^d. \tag{1.7}$$

L'idée de construire une base $\{K\left(.,x_1\right),...,K\left(.,x_N\right)\}$ de W nous conduira à la construction du noyau reproduisant de l'espace de Hilbert V, qui sera en l'occurence la fonction

$$K:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}\ \ (x,y)\mapsto q_x\left(y\right), \tag{1.8}$$

où l'expression de $q_x\left(y\right)$ sera démontré dans la suite au lemme 9, comme dans [33], grâce au théorème de représentation de Riesz. Nous avons apporté l'hypothèse (1.37), à savoir : Pour chaque $x\in\mathbb{R}^d$ il existe une constante $M_x>0$ telle que

$$v\left(x_i\right)=0,\ i=1,...,l\Rightarrow\left|v\left(x\right)\right|\leq M_x\sqrt{\langle v,v\rangle},$$

comme étant une modification de l'hypothèse analogue considérée jusqu'à présent, c'est à dire : Pour chaque $x\in\mathbb{R}^d$ il existe une constante $M_x>0$ telle que

$$v\left(x_i\right)=0,\ i=1,...,N\Rightarrow\left|v\left(x\right)\right|\leq M_x\sqrt{\langle v,v\rangle},$$

comme dans [32], [33] et [34]. Par ailleurs, par rapport à la norme $\|.\|$ définie

sur V par :

$$\|v\| = \sqrt{(v,v)},\ \forall v \in V,$$

où $(.,.)$ est le produit scalaire sur V défini par

$$(u,v) = \langle u,v\rangle + \sum_{i=1}^{l} u(x_i)\, v(x_i),\ \forall u,v \in V,$$

on veut que $s_{f,X}$ soit la meilleure approximation de f au sens

$$\|f - s_{f,X}\| = \inf_{v\in W} \|f - v\|.$$

1.1 Formulation du problème d'interpolation à résoudre

Soit alors $d > 0$ un entier donné et $\Omega \subset \mathbb{R}^d$ un ouvert borné de frontière $\partial\Omega$. On notera $\mathcal{C}(\Omega)$ l'ensemble des fonctions continues sur Ω à valeurs réelles et

$$\mathbb{R}_{\geq 0} = \{r \in \mathbb{R}\ : r \geq 0\}.$$

Soit $X = \{x_1, ..., x_N\} \subset \Omega$ un ensemble de points distincts de Ω. On suppose que pour toute fonction $f \in \mathcal{C}(\Omega)$, $f_j = f(x_j)$ est connue $\forall j, 1 \leq j \leq N$. Le problème est d'interpoler la fonction f sur X par un interpolant $s_{f,X} : \Omega \to \mathbb{R}$ de la forme

$$s_{f,X}(x) = \sum_{j=1}^{N} \lambda_j \phi\left(\|x - x_j\|_2\right) + p(x), \tag{1.9}$$

i.e.,

$$s_{f,X}(x_i) = f(x_i),\ 1 \leq i \leq N, \tag{1.10}$$

où p appartient au sous espace π_{k-1} de $\mathcal{C}\left(\mathbb{R}^d\right)$ des polynômes de degré total au plus $k-1$ et

$$\phi : \mathbb{R}_{\geq 0} \to \mathbb{R},$$

est une fonction continue appelée fonction radiale. $\|.\|_2$ étant la norme euclidienne de $\mathbb{R}^d$, le problème revient donc à trouver des réels λ_j, $1 \leq j \leq N$ tels que

$$\sum_{j=1}^{N} \lambda_j \phi\left(\|x_i - x_j\|_2\right) + p\left(x_i\right) = f_i,\ 1 \leq i \leq N. \tag{1.11}$$

1.1.1 Hypothèses du problème

Unisolvance

On suppose que X est unisolvant par rapport à π_{k-1} *i.e.*,

$$\forall p \in \pi_{k-1}\,;\, p\left(x_j\right) = 0,\, 1 \leq j \leq N \Rightarrow p \equiv 0. \tag{1.12}$$

Condition d'orthogonalité

On suppose en outre que

$$\sum_{j=1}^{N} \lambda_j q\left(x_j\right) = 0,\, \forall q \in \pi_{k-1}. \tag{1.13}$$

1.1.2 Système à résoudre

Si l'on choisit une base $\{p_1, ..., p_l\}$ de π_{k-1}, il est clair que le système (1.11) s'écrit :

$$\sum_{j=1}^{N} \lambda_j \phi\left(\|x_i - x_j\|_2\right) + \sum_{r=1}^{l} c_r p_r\left(x_i\right) = f_i,\ 1 \leq i \leq N, \tag{1.14}$$

et l'égalité (1.13) est équivalente à

$$\sum_{j=1}^{N} \lambda_j p_r\left(x_j\right) = 0,\ 1 \leq r \leq l. \tag{1.15}$$

Les systèmes (1.14) et (1.15) s'écrivent alors sous la forme matricielle

$$\begin{pmatrix} A & P \\ P^T & O \end{pmatrix} \begin{pmatrix} \lambda \\ c \end{pmatrix} = \begin{pmatrix} \beta \\ 0 \end{pmatrix}, \tag{1.16}$$

dans laquelle

$$A_{ij} = \phi\left(\|x_i - x_j\|_2\right),\ i,j = 1,...,N, \tag{1.17}$$

$$P_{ij} = p_j\left(x_i\right),\ i = 1,...,N,\ j = 1,...,l, \tag{1.18}$$

$$\lambda = \left(\lambda_1, ...\lambda_N\right)^T, \tag{1.19}$$

$$c = \left(c_1, ..., c_l\right)^T, \tag{1.20}$$

$$\beta = \left(f_1, ..., f_N\right)^T. \tag{1.21}$$

O est la matrice nulle $l \times l$ et 0 est le vecteur nul de $\mathbb{R}^l$. Avant de discuter la résolution du système (1.16), donnons d'abord les définitions suivantes.

1.1.3 Définitions essentielles à la résolution du système

Définition 1 *Une fonction $\psi \in \mathcal{C}\left(\mathbb{R}^d\right)$ est dite définie positive sur Ω si pour tout ensemble $X = \{x_1, ..., x_N\} \subset \Omega$ de points distincts, et pour tout vecteur $\lambda = (\lambda_i)_{1 \leq i \leq N} \in \mathbb{R}^N$ on a :*

$$\sum_{i,j=1}^{N} \lambda_i \lambda_j \psi\left(x_i - x_j\right) \geq 0,$$

avec égalité si et seulement si $\lambda = 0$.

Définition 2 *Une fonction $\psi \in \mathcal{C}\left(\mathbb{R}^d\right)$ est dite conditionnellement définie positive d'ordre k sur Ω si pour tout ensemble $X = \{x_1, ..., x_N\} \subset \Omega$ de points distincts, et pour tout vecteur $\lambda = (\lambda_i)_{1 \leq i \leq N} \in \mathbb{R}^N$ satisfaisant la condition d'orthogonalité (1.15), on a*

$$\sum_{i,j=1}^{N} \lambda_i \lambda_j \psi\left(x_i - x_j\right) \geq 0, \tag{1.22}$$

avec égalité si et seulement si $\lambda = 0$.

1.1.4 Résolution du système rendu possible

Lorsque les conditions (1.12) et (1.13) sont satisfaites et que la fonction $\psi(x) = \phi(\|x\|_2)$ est conditionnellement définie positive d'ordre k sur Ω, il a été montré dans l'article de Micchelli [38] que le système (1.16) admet une unique solution. C'est le cas pour des choix particuliers de la fonction radiale ϕ. Les plus communément utilisées sont :

$$\text{Linear splines, } \phi(r) = r, \tag{1.23}$$

$$\text{Thin-plate-splines, } \phi(r) = r^{\alpha} \ln r \ (\alpha \text{ pair}), \tag{1.24}$$

$$\text{Fonction radiale de type conique : } \phi(r) = r^{\alpha} (\alpha \text{ impair}), \tag{1.25}$$

$$\text{Multiquadriques : } \phi(r) = \left(r^2 + c^2\right)^{\frac{\beta}{2}} (\beta \text{ impair}),\ c > 0, \tag{1.26}$$

$$\text{Gaussians, } \phi(r) = \exp\left(-cr^2\right),\ c > 0. \tag{1.27}$$

Fonction radiale à support compact :

$$\phi(r) = \begin{cases} \left(1 - \dfrac{r}{\rho}\right)^3 \left(1 + 3\dfrac{r}{\rho} + \left(\dfrac{r}{\rho}\right)^2\right), \text{ si } 0 \leq r \leq \rho \\ 0, \text{ si } r > \rho, \end{cases} \tag{1.28}$$

où le nombre $c > 0$ sera appelé paramètre de forme. Dans les cas (1.23), (1.24) et (1.27) avec $c = 1$, il a été montré dans l'article de Light et Wayne [33], via la transformée de Fourier $\widehat{\psi}$ de la fonction $\psi(x) = \phi(\|x\|_2)$, que la matrice A est définie positive. Dans le cas général l'égalité (1.1) s'écrira

$$s_{f,X}(x) = \sum_{j=1}^{N} \lambda_j \phi\left(\|x - x_j\|_2\right) + \sum_{\substack{\alpha \in Z^d, \alpha \geq 0 \\ |\alpha| \leq k-1}} \mu_{\alpha} x^{\alpha}, \tag{1.29}$$

où

$$\begin{aligned} \alpha &= (\alpha_1, ..., \alpha_d), \\ |\alpha| &= \sum_{i=1}^{d} \alpha_i. \end{aligned}$$

Nous aurons parfois besoin des fonctions radiales à support compact définies positives. Une technique simple de la construction de telles fonctions est basée sur la convolution.

1.2 Fonctions radiales définies positives à support compact

Définition 3 *Soit $\psi : \mathbb{R}_{\geq 0} \to \mathbb{R}$ une fonction continue non nulle à support compact. On définit la fonction ϕ par*

$$\phi\left(\|x\|_2\right) = \int_{\mathbb{R}^d} \psi\left(\|y\|_2\right) \psi\left(\|x-y\|_2\right) dy. \tag{1.30}$$

Par des arguments sur la tranformée de Fourier, le résultat obtenu est une fonction radiale. De plus cette fonction est à support compact par constrution. Elle est en outre définie positive par le fait que

$$\sum_{i,j=1}^{N} \alpha_i \alpha_j \phi\left(\|x_i - x_j\|_2\right) = \int_{\mathbb{R}^d} \left(\sum_{j=1}^{N} \alpha_j \psi\left(\|x_j - y\|_2\right) \right)^2 dy. \tag{1.31}$$

1.3 Ordre d'approximation de l'interpolé

Grâce à la théorie variationnelle de Golomb- Weinberger [18] , on montre que dans le cas de l'interpolation par des fonctions radiales, l'interpolé $s_{f,X}$ possède un ordre d'approximation $\mathcal{O}\left(h^l\right)$ de f, *i.e.*, il existe une constante C telle que

$$\|s_{f,X} - f\|_\infty \leq C h^l, \tag{1.32}$$

où h désigne la densité des points donnée par

$$h = \sup_{x \in \Omega} \min_{1 \leq j \leq N} \|x - x_j\|_2 . \tag{1.33}$$

Dans la suite on notera

$$\psi_x = \phi\left(\|. - x\|_2\right) \text{ pour tout } x \in \mathbb{R}^d. \tag{1.34}$$

De la continuité de ϕ sur $\mathbb{R}_{\geq 0}$ on en déduit que $\psi_x \in \mathcal{C}\left(\mathbb{R}^d\right)$ $\forall x \in \mathbb{R}^d$. On supposera en outre que la fonction $\psi = \phi\left(\|.\|_2\right)$ est conditionnellement définie positive d'ordre k.

1.4 Théorie variationnelle du problème d'approximation par interpolation

Soit V un sous espace vectoriel de $\mathcal{C}\left(\mathbb{R}^d\right)$. On suppose que pour toute fonction $f \in V$, $f_j = f\left(x_j\right)$ est connue $\forall j$, $1 \leq j \leq N$. De cette connaissance on veut alors trouver une autre fonction $u \in V$ telle que $u\left(x\right)$ approche $f\left(x\right)$ et $u\left(x_j\right) = f\left(x_j\right)$ $\forall j$,$1 \leq j \leq N$, où $x \in \mathbb{R}^d$. Comme cela est usuel en théorie d'approximation, l'idée est de trouver u dans un sous espace de V possédant de bonnes propriétés et donc utiliser $u\left(x\right)$ pour approcher $f\left(x\right)$. On suppose que sur V est défini un semi-produit scalaire

$$\langle ., . \rangle : V \times V \to \mathbb{R},$$

jouissant de toutes les propriétés d'un produit scalaire sauf que $\langle v, v \rangle = 0$ n'implique pas nécéssairement $v = 0$. En posant

$$|v|_k = \sqrt{\langle v, v \rangle},$$

pour tout $v \in V$, on définit ainsi une semi-norme sur V, de noyau

$$\mathcal{N} = \left\{v \in V; \langle v, v \rangle = 0\right\}, \text{ avec } \mathcal{N} \neq \{0\}. \tag{1.35}$$

On suppose en outre que $N \geq l$; $\mathcal{N} = \pi_{k-1}$ et $\{x_1, ..., x_l\}$ est unisolvant par rapport à π_{k-1}. L'hypothèse (1.12) est alors satisfaite et on a :

$$\text{si } p \in \pi_{k-1} \text{ et } p\left(x_i\right) = 0,\ 1 \leq i \leq l \text{ alors } p \equiv 0. \tag{1.36}$$

1.4.1 Modification de l'hypothèse originale

On suppose que pour chaque $x \in \mathbb{R}^d$ il existe une constante $M_x > 0$ telle que

$$v(x_i) = 0,\ i = 1, ..., l \Rightarrow |v(x)| \leq M_x \sqrt{\langle v, v \rangle}. \tag{1.37}$$

1.4.2 L'espace de Hilbert V

Nous pouvons maintenant définir sur V un produit scalaire $(.,.)$ et une norme $\|.\|$ par :

$$(u, v) = \langle u, v \rangle + \sum_{i=1}^{l} u(x_i)\, v(x_i),\ \forall\ u,v \in V, \tag{1.38}$$

$$\|v\| = \sqrt{(v, v)},\ \forall v \in V, \tag{1.39}$$

et supposer que V est complet pour cette norme. Notons

$$\begin{aligned} \Gamma_0 &= \{v \in V; v(x_i) = 0,\ 1 \leq i \leq l\}, \\ \Gamma &= \{v \in V; v(x_i) = 0,\ 1 \leq i \leq N\}. \end{aligned} \tag{1.40}$$

D'après (1.38) on a

$$\forall i = 1, ..., l,\ \forall v \in V,\ |v(x_i)| \leq \|v\|, \tag{1.41}$$

ce qui montre que Γ_0 est un sous-espace fermé de V donc

$$\overline{\Gamma} \subset \Gamma_0. \tag{1.42}$$

D'après (1.37) on a

$$\forall i = 1, ..., N,\ \forall v \in \Gamma_0,\ |v(x_i)| \leq M_{x_i} \sqrt{\langle v, v \rangle} \leq M \|v\|, \tag{1.43}$$

où

$$M = \max_{1 \leq i \leq N} \{M_{x_i}\}. \tag{1.44}$$

(1.42) et (1.43) montrent que Γ est un sous-espace fermé de V . On supposera enfin que pour chaque $x \in \mathbb{R}^d$ l'application

$$\delta_x : V \to \mathbb{R},\ v \mapsto v(x) \text{ est bornée.} \tag{1.45}$$

1.4.3 Représentant dans V du point d'évaluation

L'application δ_x étant bornée, par le théorème de représentation de Riesz, il existe donc $q_x \in V$ unique tel que

$$v(x) = (v, q_x) \ \forall v \in V. \tag{1.46}$$

Pour tout $x \in \mathbb{R}^d$, q_x est appelé représentant dans V de δ_x.

1.4.4 Sous-espace des interpolés

On notera

$$W = \left\{ s = \sum_{j=1}^{N} \lambda_j \psi_{x_j} + p \,;\, \lambda_1, ..., \lambda_N \in \mathbb{R} \,;\, p \in \pi_{k-1} \right. \\ \left. \text{et } \sum_{j=1}^{N} \lambda_j q(x_j) = 0 \ \forall q \in \pi_{k-1} \right\}, \tag{1.47}$$

le sous-espace des interpolés $s_{f,X}$ définis en (1.1), des fonctions f de V. Pour tout $f, g \in V$ on peut alors noter

$$\langle f, g \rangle = \sum_{i,j=1}^{N} \lambda_i \mu_j \phi\left(\|x_i - x_j\|_2 \right), \tag{1.48}$$

où λ_i et μ_j sont des réels tels que $s_{f,X} = \sum_{i=1}^{N} \lambda_i \psi_{x_i} + p$ et $s_{g,X} = \sum_{j=1}^{N} \mu_j \psi_{x_j} + q$ appartiennnent à W. On peut alors dégager des propriétés intéressantes de $\langle ., . \rangle$:

$$\forall u, v \in W; u = \sum_{i=1}^{N} \lambda_i \psi_{x_i} + p \,;\, v = \sum_{j=1}^{N} \mu_j \psi_{x_j} + q \text{ où } p, q \in \pi_{k-1},$$

on a

$$\sum_{i,j=1}^{N} \lambda_i \mu_j \phi\left(\|x_i - x_j\|_2 \right) = \sum_{i=1}^{N} \lambda_i v(x_i). \tag{1.49}$$

On a de même

$$\sum_{i,j=1}^{N} \lambda_i \mu_j \phi\left(\|x_i - x_j\|_2\right) = \sum_{i=1}^{N} \mu_i u\left(x_i\right). \tag{1.50}$$

Du fait que la fonction $\psi = \phi\left(\|.\|_2\right)$ est conditionnellement définie positive d'ordre k sur Ω, on en déduit que l'égalité (1.48) définit un produit semi-scalaire sur V de noyau $\mathcal{N} \supset \pi_{k-1}$. On peut également vérifier que l'égalité (1.38) définit un produit scalaire sur V dont la norme induite est définie par l'égalité (1.39). En effet, la seule propriété à vérifier dans (1.38) est :

$$\forall v \in V\,;\ (v, v) = 0 \Rightarrow v = 0.$$

1.4.5 Décomposition Hilbertienne de V en somme directe et conséquence

Soit $(p_j)_{1 \leq j \leq l}$ la base de Lagrange de π_{k-1} associée à $\{x_{1,\,...,}\, x_l\}$ et

$$\mathcal{Q} : V \to \pi_{k-1}\ \ v \mapsto \sum_{j=1}^{l} v\left(x_j\right) p_j. \tag{1.51}$$

On a pour tout $i, 1 \leq i \leq l$,

$$\mathcal{Q}v\left(x_i\right) = \sum_{j=1}^{l} v\left(x_j\right) p_j\left(x_i\right) = v\left(x_i\right). \tag{1.52}$$

On peut alors montrer que $\mathcal{Q}$ est la projection orthogonale de V sur π_{k-1} ; *i.e.*,

$$\pi_{k-1} = \{v \in V; \mathcal{Q}v = v\} \text{ et } \pi_{k-1}^{\perp} = \{v \in V; \mathcal{Q}v = 0\}. \tag{1.53}$$

Par ailleurs on montre facilement que

$$\pi_{k-1}^{\perp} = \Gamma_0 = \{v \in V; v\left(x_i\right) = 0,\ 1 \leq i \leq l\}. \tag{1.54}$$

Soit $f \in V$. Notons

$$C_f = \{v \in V; (v, v) \leq (f, f),\ v\left(x_i\right) = f\left(x_i\right), 1 \leq i \leq N\}. \tag{1.55}$$

Il est clair que $f \in C_f$ et C_f est une partie convexe fermée de V donc C_f admet un élément de norme minimale u noté U_f. Revenons à la notation

$$\Gamma = \{v \in V; v(x_i) = 0,\ 1 \leq i \leq N\}.$$

En écrivant

$$V = \Gamma \oplus \Gamma^{\perp}, \tag{1.56}$$

on a

$$u = u_1 + u_2 \text{ où } u_1 \in \Gamma \text{ et } u_2 \in \Gamma^{\perp}. \tag{1.57}$$

On a donc

$$u_2(x_i) = u(x_i) = f(x_i)\ \ \forall i = 1, ..., N. \tag{1.58}$$

u_2 est donc un interpolant de f sur $X = \{x_1, ..., x_N\}$. Par suite $\|u\| \leq \|u_2\|$. Or d'aprrès (1.57) on a $\|u_2\| \leq \|u\|$. Donc $\|u\| = \|u_2\|$. Comme $\|u\|^2 = \|u_1\|^2 + \|u_2\|^2$, alors $u_1 = 0$. Donc

$$u = u_2 \in \Gamma^{\perp}. \tag{1.59}$$

1.4.6 Localisation de l'interpolant de norme minimale

Soit $x \in \mathbb{R}^d$. Supposons que les applications δ_x, $\delta_{x_1}, ..., \delta_{x_N}$, soient linéairement indépendantes sur V^*, le dual topologique de V. Du fait de (1.45) et par le théorème de représentation de Riesz, il existe $q_x, q_1, ..., q_N \in V$ uniques tels que

$$v(x) = (v, q_x)\ \text{ et } v(x_i) = (v, q_i), 1 \leq i \leq N, \forall v \in V.$$

Il est donc clair que $q_x, q_1, ..., q_N$ sont linéairement indépendantes sur V. En notant

$$S = \text{vect}(q_1, ..., q_N), \tag{1.60}$$

on a

$$S^{\perp} = \{v \in V; (v, q_i) = v(x_i) = 0,\ 1 \leq i \leq N\} = \Gamma. \tag{1.61}$$

Ainsi,

$$\Gamma^{\perp} = S = \text{vect}(q_1, ..., q_N). \tag{1.62}$$

Théorème 4 *Soit $f \in V$ et u l'unique élément de norme minimale de*

$$C_f = \{v \in V; (v,v) \leq (f,f)\,,\ v(x_i) = f(x_i)\,, 1 \leq i \leq N\}\,. \tag{1.63}$$

Pour tout $x \in \mathbb{R}^d$, si q_x est le représentant de δ_x dans V, avec

$$q_{x_i} = q_i,\ 1 \leq i \leq N, \tag{1.64}$$

alors il existe des réels $\lambda_1, ..., \lambda_N$ tels que

$$u = \sum_{i=1}^{N} \lambda_i q_i. \tag{1.65}$$

Les constantes $\lambda_1, ..., \lambda_N$ sont déterminées par les équations

$$f(x_j) = u(x_j) = (u, q_j) = \sum_{i=1}^{N} \lambda_i (q_i, q_j)\,, 1 \leq j \leq N. \tag{1.66}$$

La démonstration du théorème 4 est une conséquence immédiate de (1.59) et (1.62).

1.4.7 Majoration de l'erreur entre f et son interpolant de norme minimale

$q_x,\ q_1, ..., q_N$ étant linéairement indépendantes sur V, alors $q_x \notin \Gamma^{\perp}$. On construit donc un élément

$$\omega = \frac{1}{\|\mathcal{P}q_x\|}\mathcal{P}q_x, \tag{1.67}$$

où

$$\mathcal{P} : V \longrightarrow \Gamma \text{ est la projection orthogonale de } V \text{ dans } \Gamma. \tag{1.68}$$

Cet élément $\omega \in \Gamma$, vérifie

$$\|\omega\| = 1 \text{ et } \omega(x) = (\omega, q_x) = (\omega, \mathcal{P}q_x) = \|\mathcal{P}q_x\|\,. \tag{1.69}$$

Or

$$\begin{aligned} \|\mathcal{P}q_x\| &= \sup\{|(v, \mathcal{P}q_x)|\ ;\ v \in \Gamma,\ \|v\| = 1\} \\ &= \sup\{|(v, q_x)|\ ;\ v \in \Gamma,\ \|v\| = 1\} \\ &= \sup\{|v(x)|\ ;\ v \in \Gamma,\ \|v\| = 1\}\,. \end{aligned} \tag{1.70}$$

On a ainsi construit un unique élément $\omega \in \Gamma$ tel que

$$\|\omega\| = 1 \text{ et } \omega(x) = \sup\{|v(x)| \; ; \; v \in \Gamma, \; \|v\| = 1\}. \tag{1.71}$$

L'élément de norme minimale u de C_f étant noté U_f, on l'appelera l'interpolant de norme minimale de f. Comme $f - U_f \in \Gamma$ et $U_f \in \Gamma^{\perp}$, on a

$$\begin{aligned} |f(x) - U_f(x)|^2 &\leq w(x)^2 \|f - U_f\|^2 \\ &= w(x)^2 \langle f - U_f, f - U_f \rangle \\ &= \{\langle f, f\rangle - \langle f, U_f\rangle - \langle U_f, f\rangle + \langle U_f, U_f\rangle\} w(x)^2 \\ &= \{\langle f, f\rangle - \langle f - U_f, U_f\rangle - 2\langle U_f, U_f\rangle \\ &\quad - \langle U_f, f - U_f\rangle + \langle U_f, U_f\rangle\} w(x)^2 \\ &= \{\langle f, f\rangle - \langle U_f, U_f\rangle\} w(x)^2 \\ &\leq w(x)^2 \langle f, f\rangle, \end{aligned} \tag{1.72}$$

d'où

$$|f(x) - U_f(x)| \leq w(x) \sqrt{\langle f, f\rangle}. \tag{1.73}$$

Théorème 5 *Soit $X = \{x_1, ..., x_N\}$ un ensemble de points de $\mathbb{R}^d$ et u l'interpolant de norme minimale de $f \in V$ en ses points. Soit v l'interpolant de f aux points $x, x_1, ..., x_N$. Dans la notation du théorème 4, on suppose que v est de la forme*

$$v = \lambda q_x + \sum_{i=1}^{N} \lambda_i q_i + p, \text{ où } p \in \pi_{k-1}. \tag{1.74}$$

Alors,

$$|f(x) - u(x)| = |\lambda| \, \|\mathcal{P} q_x\|^2. \tag{1.75}$$

Démonstration. On a $u = \sum_{i=1}^{N} \lambda_i q_i$. Comme $v(x_i) = u(x_i) = f(x_i)$,

$1 \leq i \leq N$, on a $v - u = \lambda q_x + p \in \Gamma$. D'où

$$\begin{aligned} |f(x) - u(x)| &= |v(x) - u(x)| \\ &= |(v-u)(x)| \\ &= |(v-u, q_x)| \\ &= |(v-u, \mathcal{P}q_x)| \\ &= |(\lambda q_x, \mathcal{P}q_x) + (p, \mathcal{P}q_x)| \\ &= |(\lambda q_x, \mathcal{P}q_x) + \langle p, \mathcal{P}q_x \rangle| \\ &= |(\lambda q_x, \mathcal{P}q_x)| \\ &= |\lambda| \, \|\mathcal{P}q_x\|^2. \end{aligned} \tag{1.76}$$

■

1.4.8 Expression de w dans la majoration de l'erreur précédente :

Lemme 6 *Soient $x_1, ..., x_N \in \mathbb{R}^d$ ayant pour représentant $q_1, ..., q_N \in V$ tels que $v(x_i) = (v, q_i)$, $1 \leq i \leq N$ pour tout $v \in V$. Soit $x \in \mathbb{R}^d$ ayant pour représentant $q_x = q_0$. Soit $\omega = \dfrac{1}{\|\mathcal{P}q_x\|}\mathcal{P}q_x$ défini par (1.67). Alors :*

$$\text{il existe } \alpha_0, ..., \alpha_N \in \mathbb{R} \text{ tels que } w = \sum_{i=0}^{N} \alpha_i q_i. \tag{1.77}$$

De plus ces coefficients sont déterminés par les équations

$$\|w\| = 1 \text{ et } w(x_j) = \sum_{i=0}^{N} \alpha_i (q_i, q_j) = 0, \ 1 \leq j \leq N. \tag{1.78}$$

Démonstration. On définit

$$\Gamma_x = \Gamma \cap \{v \in V; v(x) = 0\}. \tag{1.79}$$

En utilisant les répresentants on peut écrire

$$\Gamma_x = \{v \in V; (v, q_i) = 0, \ i = 0, ..., N\}. \tag{1.80}$$

Soit

$$\mathcal{L} : V \longrightarrow \Gamma_x \text{ la projection orthogonale de } V \text{ dans } \Gamma_x.$$

Comme $\omega \in \Gamma$ et $\mathcal{L}\omega \in \Gamma_x \subset \Gamma$, on a $\omega - \mathcal{L}\omega \in \Gamma$. Par ailleurs $\omega(x) - \mathcal{L}\omega(x) = \omega(x)$. Par unicité de ω on a donc $\mathcal{L}\omega = 0$. D'où $\omega \in \Gamma_x^{\perp} = \text{vect}(q_0, ..., q_N)$. Il existe donc $\alpha_0, ..., \alpha_N \in \mathbb{R}$ tels que $w = \sum_{i=0}^{N} \alpha_i q_i$. De plus, les coefficients $\alpha_0, ..., \alpha_N$ sont déterminés par les équations $\|w\| = 1$ et $\sum_{i=0}^{N} \alpha_i (q_i, q_j) = \sum_{i=0}^{N} \alpha_i q_i(x_j) = 0$, $1 \leq j \leq N$. ■

Lemme 7 *Soit $p_1, ..., p_l \in \pi_{k-1}$ tels que $p_i(x_j) = \delta_{ij}, 1 \leq i, j \leq l$. Alors, pour tout $v \in V$,*

$$v(x_i) = (v, p_i), 1 \leq i \leq l. \tag{1.81}$$

Démonstration.

$$\begin{aligned} \text{On a } (v, p_i) &= \langle v, p_i \rangle + \sum_{j=1}^{l} v(x_j) p_i(x_j) \\ &= 0 + \sum_{j=1}^{l} v(x_j) p_i(x_j) = v(x_i). \end{aligned}$$

■

Lemme 8 *Soient $\alpha_0, ..., \alpha_N$ définis au lemme 6 et $p_1, ..., p_l$ définis au lemme 7. Alors,*

$$\sum_{r=0}^{N} \alpha_r p(x_r) = 0, \ \forall p \in \pi_{k-1}. \tag{1.82}$$

Démonstration. On a pour $1 \leq i \leq l$,

$$\sum_{r=0}^{N} \alpha_r p_i(x_r) = \sum_{r=0}^{N} \alpha_r (p_i, q_r) = \sum_{r=0}^{N} \alpha_r q_r(x_i) = w(x_i) = 0.$$

Comme $p_1, ..., p_l$ est une base, le résultat s'en suit. ■

1.4.9 Base du Sous-espace W des interpolés

Considerons maintenant les hypothèses $(\mathcal{H})$ suivantes faites sur V :

(i) $V \subset \mathcal{C}\left(\mathbb{R}^d\right)$, et un produit semi- scalaire $\langle .,. \rangle$ est défini sur V, de noyau π_{k-1}.

(ii) $\{x_1, ..., x_l\}$ est un sous ensemble π_{k-1}- unisolvant et $p_1, ..., p_l \in \pi_{k-1}$ satisfait $p_i(x_j) = \delta_{ij}, 1 \leq i, j \leq l$.

On suppose en outre que la fonction radiale $\phi \in \mathcal{C}(\mathbb{R}_{\geq 0})$ est telle que $\forall x \in \mathbb{R}^d$, il existe $r_x \in V$ tel que

$$r_x(y) = \phi(\|y - x\|_2) - \sum_{i=1}^{l} p_i(x)\, \phi(\|y - x_i\|_2),\ \forall y \in \mathbb{R}^d, \tag{1.83}$$

et

$$(v, r_x) = v(x),\ \forall v \in \Gamma_0, \tag{1.84}$$

avec $(u, v) = \langle u, v \rangle + \sum_{i=1}^{l} u(x_i)\, v(x_i)$, $\forall u, v \in V$.

1.4.10 Expression du représentant q_x du point d'évaluation δ_x en $x \in \mathbb{R}^d$

Lemme 9 *On suppose que les hypothèses $(\mathcal{H})$ faites sur V sont satisfaites. Alors, le représentant q_x du point d'évaluation δ_x en $x \in \mathbb{R}^d$, c'est à dire l'unique élément de V tel que $v(x) = (v, q_x)$ $\forall v \in V$, est de la forme*

$$\begin{aligned} q_x(y) &= \phi(\|y - x\|_2) \\ &- \sum_{i=1}^{l} p_i(x)\, \phi(\|y - x_i\|_2) - \sum_{i=1}^{l} p_i(y)\, \phi(\|x - x_i\|_2) \\ &+ \sum_{i,j=1}^{l} \phi\left(\|x_i - x_j\|_2\right) p_i(x)\, p_j(y) + \sum_{i=1}^{l} p_i(x)\, p_i(y). \end{aligned} \tag{1.85}$$

q_x est le noyau reproduisant.

Démonstration. Soit $\Gamma_0 = \{v \in V;\ v(x_i) = 0,\ 1 \leq i \leq l\}$. On définit $Q : V \to \pi_{k-1}$ par $Qv = \sum_{i=1}^{l} v(x_i)\, p_i = \sum_{i=1}^{l} (v, p_i)\, p_i$. On a vu au (1.51) que Q est la projection orthogonale de V sur $\pi_{k-1} = \Gamma_0^{\perp}$ et donc $I - Q$ est la projection orthogonale sur Γ_0. On a :

$$\begin{array}{ll} (v, (I - \mathcal{Q})\, r_x) & = (v, r_x) = v(x)\ \forall v \in \Gamma_0, \\ (v, (I - \mathcal{Q})\, r_x) & = 0\ \forall v \in \Gamma_0^{\perp}. \end{array} \tag{1.86}$$

$$\begin{aligned} \forall\ v \in V \quad (v, (I - \mathcal{Q})\, r_x) \quad = \quad ((I - \mathcal{Q})\, v, r_x) \quad & = (I - \mathcal{Q})\, v(x) \\ & = v(x) - \mathcal{Q}v(x) \\ & = (v, q_x) - \sum_{i=1}^{l} v(x_i)\, p_i(x) \\ & = (v, q_x) - \sum_{i=1}^{l} (v, p_i)\, p_i(x). \end{aligned} \tag{1.87}$$

D'où

$$(v, (I - \mathcal{Q})\, r_x) = \left(v, q_x - \sum_{i=1}^{l} p_i(x)\, p_i \right) \forall v \in V$$

et

$$(I - \mathcal{Q})\, r_x = q_x - \sum_{i=1}^{l} p_i(x)\, p_i,$$

par suite,

$$q_x = (I - \mathcal{Q})\, r_x + \sum_{i=1}^{l} p_i(x)\, p_i. \tag{1.88}$$

$\forall y \in \mathbb{R}^d$ on a $q_x(y) = r_x(y) - \mathcal{Q}r_x(y) + \sum_{i=1}^{l} p_i(x)\, p_i(y)$. Après calculs et arrangement on aboutit au résultat. Soit maintenant q_x l'élément de V vérifiant

$(I-\mathcal{Q})\, r_x = q_x - \sum_{i=1}^{l} p_i(x)\, p_i$. En utilisant le lemme 7 on a :

$$\begin{array}{rll} \forall v \in V & v(x) & = (v-\mathcal{Q}v)(x) + (\mathcal{Q}v)(x) \\ & & = ((I-\mathcal{Q})\, v, r_x) + (\mathcal{Q}v)(x) \\ & & = (v, (I-\mathcal{Q})\, r_x) + \sum_{i=1}^{l} v(x_i)\, p_i(x) \\ & & = (v, (I-\mathcal{Q})\, r_x) + \sum_{i=1}^{l} (v, p_i)\, p_i(x) \\ & v(x) & = \left(v, (I-\mathcal{Q})\, r_x + \sum_{i=1}^{l} p_i(x)\, p_i \right) \\ & v(x) & = (v, q_x) \ \ \forall v \in V. \end{array} \tag{1.89}$$

■

1.4.11 Autres définitions de $w(x)$ et propriétés

Théorème 10 *On définit $x_0 = x$ et on suppose que $x_0, ..., x_N$ admettent pour représentants $q_0, ..., q_N$ respectivement. Soit $\alpha_0, ..., \alpha_N \in \mathbb{R}$ tels que $w = \sum_{i=0}^{N} \alpha_i q_i$. Alors,*

$$\begin{array}{rcl} (w(x))^2 & = & \sum_{r,s=0}^{N} \beta_r \beta_s \phi(\|x_r - x_s\|_2), \\ \text{où } \beta_r & = & \alpha_r/\alpha_0,\ r = 0, ..., N. \end{array} \tag{1.90}$$

Démonstration. D'après la notation (1.34) on a $\psi_z(y) = \phi(\|y-z\|_2)$,

$\forall y, z \in \mathbb{R}^d$. Alors d'après le lemme 9

$$\begin{aligned}
w &= \sum_{r=0}^{N} \alpha_r q_r \\
&= \sum_{r=0}^{N} \alpha_r \left\{ \begin{array}{c} \psi_{x_r} - \sum_{i=1}^{l} p_i(x_r)\, \psi_{x_i} - \sum_{i=1}^{l} \psi_{x_i}(x_r)\, p_i \\ + \sum_{i,j=1}^{l} \psi_{x_i}(x_j)\, p_j(x_r)\, p_i + \sum_{i=1}^{l} p_i(x_r)\, p_i \end{array} \right\} \\
&= \sum_{r=0}^{N} \alpha_r \psi_{x_r} - \sum_{i=1}^{l} \left(\sum_{r=0}^{N} \alpha_r p_i(x_r) \right) \psi_{x_i} + \rho \text{ où } \rho \in \pi_{k-1} \qquad (1.91) \\
&= \sum_{r=1}^{l} \alpha_r \psi_{x_r} - \sum_{r=1}^{l} \left(\sum_{s=0}^{N} \alpha_s p_r(x_s) \right) \psi_{x_r} + \sum_{r=0,l+1}^{N} \alpha_r \psi_{x_r} + \rho \\
&= \sum_{r=1}^{l} \left(\alpha_r - \sum_{s=0}^{N} \alpha_s p_r(x_s) \right) \psi_{x_r} + \sum_{r=0,l+1}^{N} \alpha_r \psi_{x_r} + \rho \\
&= \sum_{r=0}^{N} \mu_r \psi_{x_r} + \rho,
\end{aligned}$$

où

$$\mu_r = \left\{ \begin{array}{l} \alpha_r \ , \ r = 0,\, l+1, ..., N \\ \alpha_r - \sum_{s=0}^{N} \alpha_s p_r(x_s) \ , \ r = 1, ..., l \end{array} \right. , \qquad (1.92)$$

$$\rho = \sum_{r=0}^{N} \alpha_r \left\{ \sum_{i,j=1}^{l} \psi_{x_i}(x_j)\, p_j(x_r)\, p_i - \sum_{i=1}^{l} \psi_{x_i}(x_r)\, p_i + \sum_{i=1}^{l} p_i(x_r)\, p_i \right\}. \qquad (1.93)$$

Comme $\sum_{s=0}^{N} \alpha_s p_r(x_s) = 0$, $r = 1, ..., l$ d'après le lemme 8, alors

$$\mu_r = \alpha_r, r = 0, ..., N. \qquad (1.94)$$

De même

$$\begin{aligned}
\rho = &- \sum_{i=1}^{l} \left(\sum_{r=0}^{N} \alpha_r \psi_{x_i}(x_r) \right) p_i + \sum_{i,j=1}^{l} \psi_{x_j}(x_i) \left(\sum_{r=0}^{N} \alpha_r p_j(x_r) \right) p_i \\
&+ \sum_{i=1}^{l} \left(\sum_{r=0}^{N} \alpha_r p_i(x_r) \right) p_i,
\end{aligned}$$

$$\rho = - \sum_{i=1}^{l} \left(\sum_{r=0}^{N} \alpha_r \psi_{x_i}(x_r) \right) p_i. \qquad (1.95)$$

Ainsi

$$w = \sum_{r=0}^{N} \alpha_r \psi_{x_r} - \sum_{i=1}^{l} \left(\sum_{r=0}^{N} \alpha_r \psi_{x_i}(x_r) \right) p_i. \tag{1.96}$$

Comme $w(x_i) = 0$, $i = 1, ..., N$ on a

$$\begin{aligned} 1 = \|w\|^2 = (w, w) &= \left(\sum_{s=0}^{N} \alpha_s q_s, w \right) \\ &= (\alpha_0 q_0, w) + \sum_{s=1}^{N} \alpha_s (q_s, w) \\ &= \alpha_0 (q_0, w) + \sum_{s=1}^{N} \alpha_s w(x_s) \\ &= \alpha_0 w(x). \end{aligned}$$

$$\begin{aligned} \text{De plus } (w, w) &= \left(\sum_{s=0}^{N} \alpha_s q_s, w \right) \\ &= \sum_{s=0}^{N} \alpha_s w(x_s) \\ &= \sum_{r,s=0}^{N} \alpha_r \alpha_s \psi_{x_r}(x_s) - \sum_{i=1}^{l} \sum_{r=0}^{N} \alpha_r \psi_{x_i}(x_r) \sum_{s=0}^{N} \alpha_s p_i(x_s) \end{aligned} \tag{1.97}$$

D'où

$$(w, w) = \sum_{r,s=0}^{N} \alpha_r \alpha_s \psi_{x_r}(x_s). \tag{1.98}$$

Finalement, $\sum_{r,s=0}^{N} \beta_r \beta_s \psi_{x_r}(x_s) = \sum_{r,s=0}^{N} (\alpha_r \alpha_s / \alpha_0^2) \psi_{x_r}(x_s) = (1/\alpha_0^2) \alpha_0 w(x) = (w(x))^2$. D'où

$$\sum_{r,s=0}^{N} \beta_r \beta_s \psi_{x_r}(x_s) = (w(x))^2. \tag{1.99}$$

Si l'on définit

$$\gamma_i = -\beta_i, \; i = 1, ..., N, \tag{1.100}$$

alors

$$(w(x))^2 = \sum_{r,s=1}^{N} \gamma_r \gamma_s \phi(\|x_r - x_s\|_2) - 2\sum_{r=1}^{N} \gamma_r \phi(\|x - x_r\|_2) + \phi(0). \quad (1.101)$$

■

Lemme 11 *Soient A et B deux sous-ensemble de $\mathbb{R}^d$ qui sont unisolvants par rapport à π_{k-1}. Soient w_A et w_B associés à A et B respectivement. Si $A \subset B$ alors*

$$w_A \geq w_B.$$

Démonstration. Fixons $x \in \mathbb{R}^d$. Si $\Gamma_A = \left\{v \in V \,;\, v_{|_A} = 0\right\}$ et $\Gamma_B = \left\{v \in V \,;\, v_{|_B} = 0\right\}$, alors $\Gamma_B \subset \Gamma_A$ et donc

$$\begin{array}{rl} w_B(x) & = \sup\left\{|v(x)| \;;\; v \in \Gamma_B,\ \|v\| = 1\right\} \\ & \leq \sup\left\{|v(x)| \;;\; v \in \Gamma_A,\ \|v\| = 1\right\} = w_A(x). \end{array} \quad (1.102)$$

■

Théorème 12 *On suppose que V satisfait les hypothèses $(\mathcal{H})$. Soit U_f l'interpolant de norme minimale de f sur $x_1, ..., x_N \in \mathbb{R}^d$. On suppose que $l \leq N$ et $\{x_1, ..., x_l\}$ unisolvant par rapport à π_{k-1}. Alors*

$$\begin{array}{rl} |f(x) - (U_f)(x)|^2 \leq & \left\{\phi(0) - 2\sum_{r=1}^{l} p_r(x)\,\phi(\|x - x_r\|_2)\right. \\ & \left. + \sum_{r,s=1}^{l} p_r(x)\,p_s(x)\,\phi(\|x_r - x_s\|_2)\right\}\langle f, f\rangle. \end{array} \quad (1.103)$$

Démonstration. De (1.73), on a $(f(x) - (U_f)(x))^2 \leq (w(x))^2 \langle f, f\rangle$; w étant définit par rappart à $\{x_1, ..., x_N\}$, comme $\{x_1, ..., x_l\} \subset \{x_1, ..., x_N\}$,

cette inégalité est encore vraie lorsque w est défini par rapport à $\{x_1, ..., x_l\}$ à cause du lemme 11. D'après le théorème 10,

$$\begin{aligned}(w(x))^2 &= \sum_{r,s=0}^{l} \beta_r \beta_s \phi(\|x_r - x_s\|_2) \text{ où } \beta_r = \alpha_r/\alpha_0 \\ \text{et } w &= \sum_{i=0}^{l} \alpha_i q_i = \alpha_0 q_0 + \sum_{i=1}^{l} \alpha_i q_i,\end{aligned}$$

d'après le lemme 6. Or $\forall v \in V$, $(v, q_i) = v(x_i) = (v, p_i)$, $1 \leq i \leq l$, d'après le lemme 7. Donc,

$$q_i = p_i, 1 \leq i \leq l.$$

Pour $j = 1, ..., l$,

$$\begin{aligned}0 = w(x_j) &= \alpha_0 q_0(x_j) + \sum_{i=1}^{l} \alpha_i p_i(x_j) \\ &= \alpha_0 q_0(x_j) + \alpha_j \\ &= \alpha_0 (q_0, p_j) + \alpha_j \\ &= \alpha_0 p_j(x) + \alpha_j.\end{aligned} \tag{1.104}$$

Ainsi,

$$\alpha_j/\alpha_0 = -p_j(x), j = 1, ..., l, \tag{1.105}$$

et le résultat s'en suit. ■

Corollaire 13 *Soit V satisfaisant les hypothèses $(\mathcal{H})$. On suppose que $v \in V$ vérifie $v(x_1) = ... = v(x_l) = 0$, où $\{x_1, ..., x_l\}$ est un ensemble de points unisolvant par rapport à π_{k-1}. Alors,*

$$\begin{aligned}(v(x))^2 \leq \quad &\left\{ \phi(0) - 2\sum_{r=1}^{l} p_r(x) \phi(\|x - x_r\|_2) \right. \\ &\left. + \sum_{r,s=1}^{l} p_r(x) p_s(x) \phi(\|x_r - x_s\|_2) \right\} \langle v, v \rangle.\end{aligned} \tag{1.106}$$

Démonstration. D'après (1.71), si $v(x_1) = ... = v(x_l) = 0$, alors $(v(x))^2 \leq (w(x))^2 \langle v, v \rangle$, w étant défini ici par rapport à $\{x_1, ..., x_l\}$ dans le théorème 12. ■

1.4.12 Noyau reproduisant de l'espace de Hilbert V

En définissant les fonctions K et Φ par

$$\Phi \ : \ \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}, \ (x,y) \mapsto \phi\left(\|y-x\|_2\right), \tag{1.107}$$

et

$$K : \ \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}, \ (x,y) \mapsto q_x\left(y\right), \tag{1.108}$$

on a

$$\begin{array}{l} K\left(.,x\right) = q_x \ \forall x \in \mathbb{R}^d, \\ \forall x,y \in \mathbb{R}^d, \ K\left(x,y\right) = \Phi\left(x,y\right) - \sum\limits_{i=1}^{l} p_i\left(x\right)\Phi\left(x_i,y\right) - \sum\limits_{j=1}^{l} p_j\left(y\right)\Phi\left(x,x_j\right) \\ + \sum\limits_{i,j=1}^{l} p_i\left(x\right)p_j\left(y\right)\Phi\left(x_i,x_j\right) + \sum\limits_{i=1}^{l} p_i\left(x\right)p_i\left(y\right) \\ K\left(.,x_r\right) = p_r \ \forall r = 1,...,l. \end{array} \tag{1.109}$$

On voit donc que K possède les propriétés suivantes :

$\left.\begin{array}{ll} (a) & K\left(x,y\right) = K\left(y,x\right) \forall x,y \in \mathbb{R}^d, \\ (b) & K\left(.,x\right) \in V \ \forall x \in \mathbb{R}^d, \end{array}\right\}$ d'après 1.85

(c) K est définie positive,

(d) $v\left(x\right) = \left(v, K\left(.,x\right)\right) \ \forall v \in V$ et $x \in \mathbb{R}^d$. Une telle fonction K est appelée noyau générateur de l'espace de Hilbert V.

1.4.13 Transformation de l'interpolé $s_{f,X}$ et écriture dans la base du sous-espace W des interpolés

En utilisant la notation (1.34) et l'écriture (1.4), il est clair que l'interpolant radial de base défini en (1.1) et (1.13) s'écrit,

$$s_{f,X} = \sum_{r=1}^{N} \lambda_r \psi_{x_r} + \sum_{j=1}^{l} c_j p_j. \tag{1.110}$$

Par ailleurs d'après (1.85) et (1.34)

$$\begin{aligned}
\sum_{r=1}^{N}\lambda_r q_r &= \sum_{r=1}^{N}\lambda_r\left\{\psi_{x_r}-\sum_{i=1}^{l}p_i\left(x_r\right)\psi_{x_i}-\sum_{i=1}^{l}\psi_{x_i}\left(x_r\right)p_i\right.\\
&\left.+\sum_{i,j=1}^{l}\psi_{x_i}\left(x_j\right)p_j\left(x_r\right)p_i+\sum_{i=1}^{l}p_i\left(x_r\right)p_i\right\}\\
&=\sum_{r=1}^{N}\lambda_r\psi_{x_r}-\sum_{i=1}^{l}\left(\sum_{r=1}^{N}\lambda_r p_i\left(x_r\right)\right)\psi_{x_i}+\rho \text{ où } \rho\in\pi_{k-1}\\
&=\sum_{r=1}^{l}\lambda_r\psi_{x_r}-\sum_{r=1}^{l}\left(\sum_{s=1}^{N}\lambda_s p_r\left(x_s\right)\right)\psi_{x_r}+\sum_{r=l+1}^{N}\lambda_r\psi_{x_r}+\rho\\
&=\sum_{r=1}^{l}\left(\lambda_r-\sum_{s=1}^{N}\lambda_s p_r\left(x_s\right)\right)\psi_{x_r}+\sum_{r=l+1}^{N}\lambda_r\psi_{x_r}+\rho\\
&=\sum_{r=1}^{N}\lambda_r\psi_{x_r}+\rho,
\end{aligned}\tag{1.111}$$

car $\sum_{s=1}^{N}\lambda_s p_r\left(x_s\right)=0,\ r=1,...,l.$

$$\begin{aligned}
\rho &= \sum_{r=1}^{N}\lambda_r\left\{\sum_{i,j=1}^{l}\psi_{x_i}\left(x_j\right)p_j\left(x_r\right)p_i-\sum_{i=1}^{l}\psi_{x_i}\left(x_r\right)p_i+\sum_{i=1}^{l}p_i\left(x_r\right)p_i\right\}\\
&=-\sum_{i=1}^{l}\left(\sum_{r=1}^{N}\lambda_r\psi_{x_i}\left(x_r\right)\right)p_i+\sum_{i,j=1}^{l}\psi_{x_j}\left(x_i\right)\left(\sum_{r=1}^{N}\lambda_r p_j\left(x_r\right)\right)p_i\\
&+\sum_{i=1}^{l}\left(\sum_{r=1}^{N}\lambda_r p_i\left(x_r\right)\right)p_i.
\end{aligned}\tag{1.112}$$

Comme $\sum_{r=1}^{N}\lambda_r p_k\left(x_r\right)=0,\ k=1,...,l,$

$$\rho=-\sum_{i=1}^{l}\left(\sum_{r=1}^{N}\lambda_r\psi_{x_i}\left(x_r\right)\right)p_i.\tag{1.113}$$

Ainsi,

$$\sum_{r=1}^{N}\lambda_r q_r=\sum_{r=1}^{N}\lambda_r\psi_{x_r}-\sum_{i=1}^{l}\left(\sum_{r=1}^{N}\lambda_r\psi_{x_r}\left(x_i\right)\right)p_i,\tag{1.114}$$

donc,

$$\sum_{r=1}^{N}\lambda_r\psi_{x_r}=\sum_{r=1}^{N}\lambda_r q_r+\sum_{i=1}^{l}\left(\sum_{r=1}^{N}\lambda_r\psi_{x_r}\left(x_i\right)\right)p_i.\tag{1.115}$$

Or d'après (1.110) on a,

$$f(x_i) = s_{f,X}(x_i) = \sum_{r=1}^{N} \lambda_r \psi_{x_r}(x_i) + c_i,\ 1 \leq i \leq l. \tag{1.116}$$

D'après (1.115) et (1.116), (1.110) devient

$$\begin{aligned}
s_{f,X} &= \sum_{r=1}^{N} \lambda_r \psi_{x_r} + \sum_{j=1}^{l} c_j p_j \\
&= \sum_{r=1}^{N} \lambda_r q_r + \sum_{i=1}^{l} \left(\sum_{r=1}^{N} \lambda_r \psi_{x_r}(x_i) \right) p_i + \sum_{j=1}^{l} c_j p_j \\
&= \sum_{r=1}^{N} \lambda_r q_r + \sum_{i=1}^{l} \left(\sum_{r=1}^{N} \lambda_r \psi_{x_r}(x_i) + c_i \right) p_i \\
&= \sum_{r=1}^{N} \lambda_r q_r + \sum_{i=1}^{l} f(x_i)\, p_i \\
&= \sum_{r=1}^{l} \lambda_r q_r + \sum_{i=1}^{l} f(x_i)\, p_i + \sum_{r=l+1}^{N} \lambda_r q_r \\
&= \sum_{r=1}^{l} (\lambda_r + f(x_r))\, q_r + \sum_{r=l+1}^{N} \lambda_r q_r,
\end{aligned} \tag{1.117}$$

car $p_i = q_i$, $1 \leq i \leq l$.par 1.109. Ainsi,

$$s_{f,X} = \sum_{r=1}^{N} \mu_r q_r, \tag{1.118}$$

où

$$\mu_r = \begin{cases} \lambda_r + f(x_r)\ ,\ r = 1, ..., l \\ \lambda_r\ ,\ r = l+1, ..., N. \end{cases} \tag{1.119}$$

Comme

$$q_r = K(., x_r)\ ,\ 1 \leq r \leq N, \tag{1.120}$$

l'interpolant radial de base défini en (1.1) et (1.13) peut alors s'exprimer sous la forme

$$s_{f,X} = \sum_{j=1}^{N} \mu_j K(., x_j)\,, \tag{1.121}$$

où les coefficients μ_j sont donnés par les équations

$$\sum_{j=1}^{N} \mu_j K(x_i., x_j) = f_i,\ 1 \leq i \leq N. \tag{1.122}$$

La matrice définie positive associée à ce système linéaire est donnée par ses coefficients

$$K_{ij} = K(x_i, x_j),\ 1 \leq i, j \leq N. \tag{1.123}$$

L'ensemble $X = \{x_1, ..., x_N\}$ est donc supposé contenir l'ensemble $\{x_1, ..., x_l\}$ qui est π_{k-1}-unisolvant.

$$\{K(., x_1), ..., K(., x_N)\} \text{ peut donc être choisi comme base du sous-espace } W \text{ des interpolés définis en (1.47).} \tag{1.124}$$

1.4.14 coefficients $(\lambda_j)_{1\leq j\leq N}$ et $(c_r)_{1\leq r\leq l}$ qui interviennent dans la formule (1.14)

D'après (1.119) on a

$$\lambda_r = \begin{cases} \mu_r - f(x_r),\ r = 1, ..., l, \\ \mu_r,\ r = l+1, ..., N. \end{cases} \tag{1.125}$$

En posant

$$\theta_r = \mu_r - \lambda_r,\ r = 1, ..., N, \tag{1.126}$$

on a

$$\theta_r = \begin{cases} f(x_r)\ ,\ r = 1, ..., l, \\ 0\ ,\ r = l+1, ..., N. \end{cases} \tag{1.127}$$

D'où pour tout $i = 1, ..., l$ on a

$$\begin{aligned} \sum_{r=1}^{N} \mu_r \psi_{x_r}(x_i) &= \sum_{r=1}^{N} \lambda_r \psi_{x_r}(x_i) + \sum_{r=1}^{N} \theta_r \psi_{x_r}(x_i) \\ &\sum_{r=1}^{N} \lambda_r \psi_{x_r}(x_i) + \sum_{r=1}^{l} \theta_r \psi_{x_r}(x_i) + \sum_{r=l+1}^{N} \theta_r \psi_{x_r}(x_i) \\ &\sum_{r=1}^{N} \lambda_r \psi_{x_r}(x_i) + \sum_{r=1}^{l} f(x_r) \psi_{x_r}(x_i) \end{aligned}$$

Ainsi,

$$\sum_{r=1}^{N}\lambda_r\psi_{x_r}(x_i)=\sum_{r=1}^{N}\mu_r\psi_{x_r}(x_i)-\sum_{r=1}^{l}f(x_r)\psi_{x_r}(x_i),\ i=1,...,l. \tag{1.128}$$

Par ailleurs d'après (1.116) et (1.128) on a

$$c_i=f(x_i)+\sum_{r=1}^{l}f(x_r)\psi_{x_r}(x_i)-\sum_{r=1}^{N}\mu_r\psi_{x_r}(x_i),\ i=1,...,l. \tag{1.129}$$

Ainsi (1.121) est une alternative potentielle de (1.1) pour la résolution des problèmes d'interpolation par les fonctions radiales.

1.4.15 Egalité entre $s_{f,X}$ et l'interpolant de norme minimale u

L'interpolant de norme minimale défini au (1.65) vérifie

$$u(x_i)=\sum_{j=1}^{N}\lambda_j q_j(x_i)=\sum_{j=1}^{N}\lambda_j K(x_i,x_j)=f_i,\ 1\leq i\leq N. \tag{1.130}$$

D'après l'unicité de la solution du système (1.122) on a

$$\lambda_j=\mu_j,\ 1\leq j\leq N. \tag{1.131}$$

D'où

$$u=s_{f,X}=\sum_{r=1}^{N}\mu_r q_r=\sum_{r=1}^{N}\lambda_r\psi_{x_r}+\sum_{j=1}^{l}c_j p_j. \tag{1.132}$$

1.4.16 $s_{f,X}$ meilleure approximation de f

Soit $S=\text{vect}(q_1,...,q_N)$ et

$$\mathcal{Q}:V\longrightarrow S \text{ la projection orthogonale de } V \text{ dans } S. \tag{1.133}$$

$$\forall v \in V,\ \mathcal{Q}v\left(x_i\right) = \left(\mathcal{Q}v, q_i\right) = \left(v, q_i\right) = v\left(x_i\right),\ 1 \leq i \leq N. \tag{1.134}$$

$\mathcal{Q}$ est donc l'opérateur d'interpolation. associé au problème (1.1). Comme

$$\mathcal{Q}V = S, \tag{1.135}$$

S est donc le sous espace des interpolés. Donc

$$S = W. \tag{1.136}$$

Soit $u \in S$ l'interpolé de norme minimale de f étudié précédemment. On a

$$\left(f - u, q_i\right) = \left(f - u\right)\left(x_i\right) = 0,\ 1 \leq i \leq N, \tag{1.137}$$

donc $f - u \in S^{\perp}$. Ainsi,

$$\begin{array}{rl} f & = u + f - u \\ & = \mathcal{Q}f + f - \mathcal{Q}f, \end{array} \tag{1.138}$$

avec $\mathcal{Q}f \in S$ et $f - \mathcal{Q}f \in S^{\perp}$, la décomposition unique de f dans $V = S \oplus S^{\perp}$ implique

$$u = \mathcal{Q}f. \tag{1.139}$$

On a donc

$$\|f - u\| = \|f - \mathcal{Q}f\| = \inf_{v \in W} \|f - v\|. \tag{1.140}$$

Comme $\mathcal{Q}f = u = s_{f,X}$,

$$s_{f,X} \text{ est la meilleure approximation de } f \text{ au sens (1.140).} \tag{1.141}$$

1.4.17 Un exemple d'espace de Hilbert V répondant aux hypothèses Précédentes

Dans ce paragraphe nous allons voir d'après Duchon [8] et [9] comment l'élément r_x définit précédemment peut être choisi lorsque V est l'espace des distributions dont toutes les dérivées partielles d'ordre k sont de carré intrégrables dans $\mathbb{R}^d$. On suppose que

$$k > d/2, \tag{1.142}$$

alors par l'injection de Sobolev,

$$V \hookrightarrow \mathcal{C}\left(\mathbb{R}^d\right). \tag{1.143}$$

On note $\mathcal{D}$ l'espace des fonctions indéfiniment différentiables et à support compact dans $\mathbb{R}^d$. Une distribution est alors une forme linéaire continue sur $\mathcal{D}$ munie d'une topologie appropriée. $[.,.]$ désigne la paire de dualité entre $\mathcal{D}$ et $\mathcal{D}'$ et δ_x est la distribution de dirac en $x \in \mathbb{R}^d$ définie par

$$[v, \delta_x] = v(x) \ \ \forall v \in \mathcal{D}.$$

Le semi-produit scalaire sur V est défini pour tout $u, v \in V$ par

$$\langle u, v \rangle = \sum_{|\alpha|=k} c_\alpha \int_{\mathbb{R}^d} \left(D^\alpha u\right)(x) \left(D^\alpha v\right)(x)\, dx, \tag{1.144}$$

où $c_\alpha = \dfrac{k!}{\alpha!}$, $\{c_\alpha \, ; |\alpha| = k\}$ est choisi tel que la semi-norme correspondante soit rotationnellement invariante. Explicitement, ces paramètres sont spécifiés par l'expansion formelle $\|\xi\|_2^{2k} = \sum\limits_{|\alpha|=k} c_\alpha \xi^{2k}$,

$$x = (x_1, x_2, ..., x_d), \tag{1.145}$$

$$\alpha = (\alpha_1, ..., \alpha_d) \in \mathbb{Z}_+^d, \ |\alpha| = \sum_{r=1}^{d} \alpha_r, \ c_\alpha = \frac{k!}{\alpha!}, \ \alpha! = \alpha_1! \alpha_2! ... \alpha_d!, \tag{1.146}$$

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} ... \partial x_d^{\alpha_d}}. \tag{1.147}$$

Exemple : Dans le cas où $d = k = 2$, on a

$$\langle v, v \rangle = \int \int_{\mathbb{R}^2} \left[\left(\frac{\partial^2 v}{\partial s^2} \right)^2 + 2 \left(\frac{\partial^2 v}{\partial s \partial t} \right)^2 + \left(\frac{\partial^2 v}{\partial t^2} \right)^2 \right] ds dt. \tag{1.148}$$

IL est clair que si $\langle v, v \rangle = 0$, alors toutes les dérivées partielles d'ordre 2 dans l'intégrale sont nulles, donc les dérivées partielles premières de v sont

constantes, et donc $v \in \pi_1(\mathbb{R}^2)$. En général, le noyau de la semi-norme associée à ce produit semi-scalaire est $\pi_{k-1}(\mathbb{R}^d)$. Le produit scalaire sur V étant défini comme précédemment, on définit $\Gamma_0 = \{v \in V; v(x_i) = 0,\ 1 \leq i \leq l\}$, avec $\dim \pi_{k-1} = l$. Fixons $x \in \mathbb{R}^d$. $(\Gamma_0, \|.\|)$ étant un espace de Hilbert, on cherche le représentant de $x \in \mathbb{R}^d$ dans Γ_0 ; c'est à dire l'unique fonction $g_x \in \Gamma_0$ telle que $v(x) = (v, g_x)$, $\forall v \in \Gamma_0$. Supposons d'abord $v \in \Gamma_0$. Comme le produit scalaire est défini par,

$$(u, v) = \langle u, v \rangle + \sum_{i=1}^{l} u(x_i)\, v(x_i), \forall u, v \in V, \tag{1.149}$$

on a

$$\forall w \in V,\ (w, v) = \langle w, v \rangle. \tag{1.150}$$

Si $v \notin \Gamma_0$, on définit

$$\mathcal{Q} : V \to \pi_{k-1},\ v \mapsto \sum_{i=1}^{l} v(x_i)\, p_i, \tag{1.151}$$

où $\{p_1, ..., p_l\}$ est la base de Lagrange de π_{k-1} associée aux points $x_1, ..., x_l$. On a vu que $\mathcal{Q}$ est la projection orthogonale de V sur $\pi_{k-1} = \Gamma_0^{\perp}$, et donc $I - \mathcal{Q}$ est la projection orthogonale sur Γ_0. On a $\forall v \in V$,

$$\begin{aligned} v(x) - (\mathcal{Q}v)(x) &= (I - \mathcal{Q})\, v(x) \\ &= ((I - \mathcal{Q})\, v, g_x) \\ &= \langle v - \mathcal{Q}v, g_x \rangle \\ &= \langle v, g_x \rangle \\ &= \sum_{|\alpha|=k} c_\alpha \int_{\mathbb{R}^d} (D^\alpha v)(y)\, (D^\alpha g_x)(y)\, dy. \end{aligned} \tag{1.152}$$

$$\begin{aligned} \forall v \in \mathcal{D}, \text{ on a } v(x) - (\mathcal{Q}v)(x) &= \sum_{|\alpha|=k} c_\alpha \left[D^\alpha v, D^\alpha g_x\right] \\ &= \left[v, (-1)^k \sum_{|\alpha|=k} c_\alpha D^{2\alpha} g_x\right], \end{aligned} \tag{1.153}$$

de plus,

$$v(x) - (\mathcal{Q}v)(x) = v(x) - \sum_{i=1}^{l} v(x_i)\, p_i(x) = \left[v, \delta_x - \sum_{i=1}^{l} p_i(x)\, \delta_{x_i} \right]. \quad (1.154)$$

(1.153) et (1.154) montre que g_x est une solution de l'équation différentielle distributionnelle

$$(-1)^k \sum_{|\alpha|=k} c_\alpha D^{2\alpha} g_x = \delta_x - \sum_{i=1}^{l} p_i(x)\, \delta_{x_i}. \quad (1.155)$$

Si l'on définit $\psi_x : \mathbb{R}^d \to \mathbb{R}$ par : $\psi_x(y) = \phi(\|y - x\|_2)$ où

$$\phi(r) = \begin{cases} c_{k,d} r^{2k-d} \ln r & \text{si } d \text{ pair}, \\ c_{k,d} r^{2k-d} & \text{si } d \text{ impair}, \end{cases} \quad (1.156)$$

alors d'après l'article [9] de Duchon, on a

$$(-1)^k \sum_{|\alpha|=k} c_\alpha D^{2\alpha} \psi_x = \delta_x, \quad (1.157)$$

avec $c_{k,d}$ des constantes données d'après Duchon par

$$c_{k,d} = \begin{cases} \dfrac{(-1)^{k-\frac{d-1}{2}} \left(k - \frac{d+1}{2}\right)!}{2^d \pi^{\frac{d-1}{2}} (k-1)!\, (2k-d)!} & \text{si } d \text{ est impair}, \\ \dfrac{(-1)^{k-\frac{d}{2}+1}}{2^{2d-1} \pi^{\frac{d}{2}} (k-1)! \left(k - \frac{d}{2}\right)!} & \text{si } d \text{ est pair}. \end{cases} \quad (1.158)$$

Ainsi, une solution particulière de l'équation (1.155) est

$$r_x = \psi_x - \sum_{i=1}^{l} p_i(x)\, \psi_{x_i}, \text{ d'où (1.83).} \quad (1.159)$$

Cet élément r_x n'appartient pas à Γ_0, et donc on prend sa projection orthogonale $(I - \mathcal{Q}) r_x$ sur Γ_0, égale à g_x. D'où

$$\begin{aligned} g_x &= (I - \mathcal{Q}) r_x \\ &= \psi_x - \sum_{i=1}^{l} p_i(x)\, \psi_{x_i} - \sum_{i=1}^{l} \psi_x(x_i)\, p_i + \sum_{i,j=1}^{l} p_i(x)\, \psi_{x_i}(x_j)\, p_j. \end{aligned} \quad (1.160)$$

g_x étant le représentant de $x \in \mathbb{R}^d$ dans Γ_0, et $\mathcal{Q}$ la projection orthogonale définie au (1.151), alors en utilisant le lemme énnoncé au (1.81) on a

$$\begin{aligned} \forall v \in V,\, (v, g_x) &= (v - \mathcal{Q}v, g_x) + (\mathcal{Q}v, g_x) \\ &= v(x) - (\mathcal{Q}v)(x) + 0 \\ &= v(x) - \sum_{i=1}^{l} (v, p_i)\, p_i(x). \end{aligned} \tag{1.161}$$

D'où

$$\begin{aligned} \forall v \in V,\, v(x) &= (v, g_x) + \sum_{i=1}^{l} (v, p_i)\, p_i(x) \\ &= \left(v, g_x + \sum_{i=1}^{l} p_i(x)\, p_i \right). \end{aligned} \tag{1.162}$$

Ainsi, le représentant q_x dans V du point d'évaluation δ_x en $x \in \mathbb{R}^d$ est

$$q_x = g_x + \sum_{i=1}^{l} p_i(x)\, p_i, \text{ d'où (1.85).} \tag{1.163}$$

Puisque $g_x = (I - \mathcal{Q})\, r_x$, alors, $\forall v \in \Gamma_0$, $(v, r_x) = (v, (I - \mathcal{Q})\, r_x) = (v, g_x) = v(x)$. Ainsi,

$$\forall v \in \Gamma_0,\ (v, r_x) = v(x)\text{, d'où (1.84) .} \tag{1.164}$$

Chapitre 2

Estimations d'erreurs par la méthode de la fonction puissance de Powell : cas particulier des thin-plates splines et des fonctions radiales de type conique

Dans ce chapitre on veut estimer l'erreur d'interpolation du problème formulé en (1.9) et (1.10) dans des cas particuliers de la fonction radiale ϕ. Comme dans [33] et [34], cette estimation d'erreur repose sur la fonction puissance de Powell $\Phi\left(x, \widetilde{X}\right)$ définie en (2.20) ; *i.e.*,

$$|f(x) - s_{f,X}(x)| \leq \left(\Phi\left(x, \widetilde{X}\right)\right)^{\frac{1}{2}} |f|_k ,$$

où

$$X = \{x_1, ..., x_N\} \subset \Omega,$$

Ω étant un ouvert borné de $\mathbb{R}^d$ et $\widetilde{X} = \{x_1, ..., x_l\}$ avec $l \leq N$, $\widetilde{X}$ étant supposé π_{k-1}-unisolvant. Les techniques utilisées dans la déduction des constantes intervenant dans les majorations de l'erreur sont celles qu'on retrouve dans

[2], et nous les avons renforcées par les résultats (2.137), (2.138),...,(2.143) puis (2.168), (2.169),...,(2.181) et (2.183), (2.184),...,(2.202). La déduction de ces constantes résulte de quelques applications de la méthode de la fonction puissance de Powell devéloppée dans [39]. Dans le cas des dimensions $d = 1, 2, 3$,et pour $k = 2$, on trouve les meilleurs constantes dans les majorations de l'erreur correspondantes respectivement à :

$$\begin{aligned} \pi_1(\mathbb{R}) \; ; \phi(r) &= r^3 \\ \pi_1\left(\mathbb{R}^2\right) \; ; \phi(r) &= r^2 \ln r \\ \pi_1\left(\mathbb{R}^3\right) \; ; \phi(r) &= r. \end{aligned}$$

On travaille ici avec des géométries simples : Dans le premier cas ($d = 1$), x est sur un segment défini par les noeuds d'interpolation $x_1 < x_2 < ... < x_N$. Dans le second cas ($d = 2$), x est sur un segment ou à l'intérieur d'un triangle défini par trois points pris parmi les N points d'interpolation $x_1, x_2..., x_N$. Dans le troisième cas ($d = 3$), x est à l'intérieur ou sur une face d'un cube ou d'un tétraèdre dont les sommets sont pris parmi les N points d'interpolation $x_1, x_2, \ ..., \ x_N$.

Dans la suite on note donc

$$X = \{x_1, ..., x_N\} \subset \Omega, \tag{2.1}$$

où Ω est un sous ensemble borné de $\mathbb{R}^d$ et

$$\widetilde{X} = \{x_1, ..., x_l\} \text{ avec } l \leq N, \tag{2.2}$$

$\widetilde{X}$ étant supposé π_{k-1}-unisolvant et $p_1, ..., p_l$ est la base de Lagrange de π_{k-1} telle que

$$p_i(x_j) = \delta_{ij}, \ 1 \leq i, j \leq l. \tag{2.3}$$

Comme précédemment, on suppose que

$$k > d/2, \tag{2.4}$$

donc par l'injection de Sobolev,

$$V \hookrightarrow \mathcal{C}\left(\mathbb{R}^d\right). \tag{2.5}$$

Le semi-produit scalaire sur V étant défini pour tout $u, v \in V$ par

$$\langle u, v \rangle = \sum_{|\alpha|=k} c_\alpha \int_{\mathbb{R}^d} \left(D^\alpha u\right)(x) \left(D^\alpha v\right)(x)\, dx, \tag{2.6}$$

le produit scalaire sur V est défini par

$$(u, v) = \langle u, v \rangle + \sum_{i=1}^{l} u(x_i)\, v(x_i), \forall u, v \in V. \tag{2.7}$$

2.1 Fonction puissance de Powell dans le cas des thin-plates splines et les fonctions radiales de type conique

En considerant la fonction radiale ϕ définie par

$$\phi(r) = \begin{cases} r^{2k-d} \ln r, & \text{si } d \text{ pair}, \\ r^{2k-d}, & \text{si } d \text{ impair}, \end{cases} \tag{2.8}$$

on a d'après (2.4),

$$2k - d - 1 \geq 0 \Rightarrow \lim_{r \to 0^+} r^{2k-d} \ln r = \lim_{r \to 0^+} r^{2k-d-1} r \ln r = 0, \tag{2.9}$$

car $\lim\limits_{r \to 0^+} r \ln r = 0$. Par suite,

$$\phi(0) = \lim_{r \to 0} \phi(r) = 0. \tag{2.10}$$

En notant $w_X(x)$ la valeur de $w(x)$ en (1.101) associée à la fonction radiale $c_{k,d}\phi(r)$, où la constante $c_{k,d}$ est donnée par (1.158), et $\phi(r)$ est définie en (2.8), on a :

$$\left(w_X(x)\right)^2 = c_{k,d} \left\{ \sum_{r,s=1}^{N} \gamma_r \gamma_s \phi\left(\|x_r - x_s\|_2\right) - 2 \sum_{r=1}^{N} \gamma_r \phi\left(\|x - x_r\|_2\right) \right\}, \tag{2.11}$$

où

$$\gamma_r = -\alpha_r / \alpha_0,\ r = 1, ..., N, \tag{2.12}$$

d'après (1.90) et (1.100). Les coefficients $\alpha_0, ..., \alpha_N$ sont déterminés au lemme 6. De même, le théorème 12 nous permet de déduire, d'après l'expression entre parenthèse en (1.103), la valeur $w_{\widetilde{X}}(x)$ associée à la fonction radiale $c_{k,d}\phi(r)$. On a :

$$\left(w_{\widetilde{X}}(x)\right)^2 = c_{k,d}\left\{\sum_{r,s=1}^{l} p_r(x)\, p_s(x)\, \phi\left(\|x_r - x_s\|_2\right) - 2\sum_{r=1}^{l} p_r(x)\, \phi\left(\|x - x_r\|_2\right)\right\}. \tag{2.13}$$

On a vu au (1.105) que

$$p_r(x) = -\alpha_r/\alpha_0,\ r = 1, ..., l, \tag{2.14}$$

donc d'après (2.12) et (2.14) on a

$$\gamma_r = p_r(x),\ r = 1, ..., l. \tag{2.15}$$

On sait que d'après le lemme 8

$$\sum_{r=0}^{N} \alpha_r p(x_r) = 0,\ \forall p \in \pi_{k-1}. \tag{2.16}$$

On a donc $\forall p \in \pi_{k-1},\ \alpha_0 p(x) + \sum_{r=1}^{N} \alpha_r p(x_r) = 0$. D'où

$$p(x) = \sum_{r=1}^{N} \left(-\alpha_r/\alpha_0\right) p(x_r). \tag{2.17}$$

Ainsi les coefficients γ_r définis au (2.12) vérifient

$$p(x) = \sum_{r=1}^{N} \gamma_r p(x_r)\ \ \forall p \in \pi_{k-1}. \tag{2.18}$$

En notant

$$\Phi(x, X) = \left(w_X(x)\right)^2, \tag{2.19}$$

et

$$\Phi\left(x, \widetilde{X}\right) = \left(w_{\widetilde{X}}(x)\right)^2, \tag{2.20}$$

avec pour fonction radiale la fonction $c_{k,d}\phi(r)$ définie précédemment, et $U_f =$

$s_{f,X}$, alors l'inégalité (1.73) devient

$$|f(x) - s_{f,X}(x)| \leq (\Phi(x, X))^{\frac{1}{2}} |f|_k . \tag{2.21}$$

Comme

$$w_X(x) \leq w_{\widetilde{X}}(x), \tag{2.22}$$

d'après le lemme 11, on a

$$|f(x) - s_{f,X}(x)| \leq \left(\Phi\left(x, \widetilde{X}\right)\right)^{\frac{1}{2}} |f|_k . \tag{2.23}$$

La fonction $\Phi(x, X)$ est appelée fonction puissance. L'inégalité (2.23) correspond à l'inégalité (1.103) avec pour fonction radiale $c_{k,d}\phi(r)$.

Lemme 14 *Soit ϕ_σ définie par $\phi_\sigma(r) = r^2 \ln(\sigma r)$, où $r \geq 0$ et $\sigma > 0$. Pour de telles valeurs de σ, posons*

$$\Phi_\sigma(\theta) = \sum_{i,j=1}^{3} \theta_i \theta_j \phi_\sigma \|x_i - x_j\|_2 - 2\sum_{i=1}^{3} \theta_i \phi_\sigma \|x - x_i\|_2 , \tag{2.24}$$

et continuons à noter $\Phi_1(\theta)$ par $\Phi(\theta)$, i.e., $\phi_1 = \phi$.

$$\text{Si } x = \sum_{i=1}^{3} \theta_i x_i \text{ avec } \sum_{i=1}^{3} \theta_i = 1, \text{ alors } \Phi_\sigma(\theta) = \Phi(\theta) . \tag{2.25}$$

Démonstration. Ecrivons ϕ_σ comme

$$\begin{array}{rcl} \phi_\sigma(r) & = & r^2 \ln r + r^2 \ln \sigma \\ & = & \phi(r) + r^2 \ln \sigma, \end{array} \tag{2.26}$$

alors, $\Phi_\sigma(\theta)$ devient

$$\begin{array}{rcl} \Phi_\sigma(\theta) & = & \sum\limits_{i,j=1}^{3} \theta_i \theta_j \phi \|x_i - x_j\|_2 + \ln \sigma \sum\limits_{i,j=1}^{3} \theta_i \theta_j \|x_i - x_j\|_2^2 \\ & & -2\sum\limits_{i=1}^{3} \theta_i \phi \|x - x_i\|_2 - 2 \ln \sigma \sum\limits_{i=1}^{3} \theta_i \|x - x_i\|_2^2 \\ & = & \Phi(\theta) + \ln \sigma \left\{ \sum\limits_{i,j=1}^{3} \theta_i \theta_j \|x_i - x_j\|_2^2 - 2\sum\limits_{i=1}^{3} \theta_i \|x - x_i\|_2^2 \right\}. \end{array} \tag{2.27}$$

En remplaçant la condition $x = \sum_{i=1}^{3} \theta_i x_i$ par $\sum_{i=1}^{4} \theta_i x_i = 0$, avec $\theta_4 = -1$ et $x_4 = x$, on a

$$\sum_{i=1}^{4} \theta_i = \sum_{i=1}^{4} \theta_i x_i = 0. \tag{2.28}$$

Or

$$\begin{aligned} \sum_{i,j=1}^{4} \theta_i \theta_j \left\| x_i - x_j \right\|_2^2 &= \sum_{i,j=1}^{3} \theta_i \theta_j \left\| x_i - x_j \right\|_2^2 + 2 \sum_{i=1}^{3} \theta_i \theta_4 \left\| x_i - x_4 \right\|_2^2 \\ &= \sum_{i,j=1}^{3} \theta_i \theta_j \left\| x_i - x_j \right\|_2^2 - 2 \sum_{i=1}^{3} \theta_i \left\| x - x_i \right\|_2^2. \end{aligned} \tag{2.29}$$

Par ailleurs on sait que $\left\| x_i - x_j \right\|_2^2 = \left\| x_i \right\|_2^2 - 2 x_i x_j + \left\| x_j \right\|_2^2$. On a alors,

$$\begin{aligned} \sum_{i,j=1}^{3} \theta_i \theta_j \left\| x_i - x_j \right\|_2^2 - 2 \sum_{i=1}^{3} \theta_i \left\| x - x_i \right\|_2^2 &= \sum_{i,j=1}^{4} \theta_i \theta_j \left\| x_i - x_j \right\|_2^2 \\ &= \sum_{i=1}^{4} \theta_i \left\| x_i \right\|_2^2 \sum_{j=1}^{4} \theta_j - 2 \sum_{i=1}^{4} \theta_i x_i \sum_{j=1}^{4} \theta_j x_j \\ &+ \sum_{j=1}^{4} \theta_j \left\| x_j \right\|_2^2 \sum_{i=1}^{4} \theta_i \\ &= 0. \end{aligned} \tag{2.30}$$

D'où

$$\Phi_\sigma (\theta) = \Phi (\theta). \tag{2.31}$$

■

Lemme 15 *Pour $0 \leq r < \infty$, la fonction ϕ_σ définie par $\phi_\sigma (r) = r^2 \ln \sigma r$ est minorée par $-\frac{1}{2} \left(\sigma^2 e \right)^{-1}$ pour tout $\sigma > 0$.*

Démonstration. Une étude élémentaire des variations de ϕ_σ montre l'existence des extremums de ϕ_σ en $r_0 = 0$ et $r_1 = \sigma^{-1} e^{\frac{-1}{2}}$; le premier donnant lieu au maximum local $\phi_\sigma (r_0) = 0$, tandis que le second tient de minimum global $\phi_\sigma (r_1) = -\frac{1}{2} \left(\sigma^2 e \right)^{-1}$. ■

Lemme 16 *Lorsque tous les vecteurs de $\mathbb{R}^d$ dans (2.11) sont affectés d'un coefficient positif σ tel qu'aucun des paramètres $\gamma_1, ..., \gamma_N$ de (2.12) ne soit modifié, alors la fonction puissance dans (2.19) est affectée du coefficient σ^{2k-d}.*

Démonstration. Désignons par Φ_a et Φ_n les valeurs de la fonction puissance $\Phi(x, X)$ dans (2.19) respectivement avant et après l'affectaction du coefficient σ. Si d est impair, d'après (2.8) on a $\phi(r) = r^{2k-d}$. Donc $\phi(\sigma r) = \sigma^{2k-d}\phi(r)$. Or d'après (2.11),

$$\begin{aligned}
\Phi_n &= c_{k,d}\left\{\sum_{r,s=1}^{N}\gamma_r\gamma_s\phi\left(\sigma\left\|x_r - x_s\right\|_2\right) - 2\sum_{r=1}^{N}\gamma_r\phi\left(\sigma\left\|x - x_r\right\|_2\right)\right\} \\
&= \sigma^{2k-d}c_{k,d}\left\{\sum_{r,s=1}^{N}\gamma_r\gamma_s\phi\left(\left\|x_r - x_s\right\|_2\right) - 2\sum_{r=1}^{N}\gamma_r\phi\left(\left\|x - x_r\right\|_2\right)\right\} \\
&= \sigma^{2k-d}\Phi_a.
\end{aligned} \tag{2.32}$$

Si d est pair, d'après (2.8), on a $\phi(r) = r^{2k-d}\ln r$, donc

$$\begin{aligned}
\phi(\sigma r) &= \sigma^{2k-d}r^{2k-d}\ln\sigma r \\
&= \sigma^{2k-d}r^{2k-d}(\ln\sigma + \ln r) \\
&= \sigma^{2k-d}\left(r^{2k-d}\ln\sigma + \phi(r)\right).
\end{aligned} \tag{2.33}$$

Or d'après (2.11),

$$\begin{aligned}
\Phi_n &= c_{k,d}\left\{\sum_{r,s=1}^{N}\gamma_r\gamma_s\phi\left(\sigma\left\|x_r - x_s\right\|_2\right) - 2\sum_{r=1}^{N}\gamma_r\phi\left(\sigma\left\|x - x_r\right\|_2\right)\right\} \\
&= c_{k,d}\left\{\sigma^{2k-d}\left[\sum_{r,s=1}^{N}\gamma_r\gamma_s\phi\left(\left\|x_r - x_s\right\|_2\right) + \ln\sigma\sum_{r,s=1}^{N}\gamma_r\gamma_s\left\|x_r - x_s\right\|_2^{2k-d}\right.\right. \\
&\quad \left.\left. -2\sum_{r=1}^{N}\gamma_r\phi\left(\left\|x - x_r\right\|_2\right) - 2\ln\sigma\sum_{r=1}^{N}\gamma_r\left\|x - x_r\right\|_2^{2k-d}\right]\right\} \\
&= \sigma^{2k-d}\Phi_a + c_{k,d}\sigma^{2k-d}\left\{\begin{array}{c}\displaystyle\sum_{r,s=1}^{N}\gamma_r\gamma_s\left\|x_r - x_s\right\|_2^{2k-d} \\ \displaystyle -2\sum_{r=1}^{N}\gamma_r\left\|x - x_r\right\|_2^{2k-d}\end{array}\right\}\ln\sigma.
\end{aligned} \tag{2.34}$$

Posons

$$\begin{array}{lll}
\widehat{N} = N+1, & \hat{\gamma}_r = \gamma_r \text{ et} & \hat{x}_r = x_r,\ r = 1, ..., N. \\
& \hat{\gamma}_{N+1} = -1, & \\
& \hat{x}_{N+1} = x, &
\end{array} \tag{2.35}$$

alors,

$$\begin{aligned}\sum_{r,s=1}^{\widehat{N}} \hat{\gamma}_r \hat{\gamma}_s \left\| \hat{x}_r - \hat{x}_s \right\|_2^{2k-d} &= \sum_{r,s=1}^{N} \gamma_r \gamma_s \left\| x_r - x_s \right\|_2^{2k-d} \\ &+2 \sum_{r=1}^{N+1} \hat{\gamma}_r \hat{\gamma}_{N+1} \left\| \hat{x}_r - \hat{x}_{N+1} \right\|_2^{2k-d} \\ &= \sum_{r,s=1}^{N} \gamma_r \gamma_s \left\| x_r - x_s \right\|_2^{2k-d} - 2 \sum_{r=1}^{N+1} \hat{\gamma}_r \left\| \hat{x}_r - x \right\|_2^{2k-d} \\ &= \sum_{r,s=1}^{N} \gamma_r \gamma_s \left\| x_r - x_s \right\|_2^{2k-d} - 2 \sum_{r=1}^{N} \gamma_r \left\| x_r - x \right\|_2^{2k-d}.\end{aligned} \tag{2.36}$$

Or d'après (2.18) on a :

$$\sum_{r=1}^{\widehat{N}} \hat{\gamma}_r p\left(\hat{x}_r\right) = 0 \ \forall p \in \pi_{k-1}. \tag{2.37}$$

On en déduit donc que

$$\sum_{r,s=1}^{N} \gamma_r \gamma_s \left\| x_r - x_s \right\|_2^{2k-d} - 2 \sum_{r=1}^{N} \gamma_r \left\| x_r - x \right\|_2^{2k-d} = 0. \tag{2.38}$$

D'où

$$\Phi_n = \sigma^{2k-d} \Phi_a, \text{ si } d \text{ est pair.} \tag{2.39}$$

■

Théorème 17 *Soit $\Omega \subset \mathbb{R}^d$ un ensemble borné contenant un $\pi_{k-1}\left(\mathbb{R}^d\right)$-unisolvant sous-ensemble. Pour un choix de points d'interpolation, soit $h = h_X$ le nombre*

$$h = \sup_{x \in \Omega} \min_{1 \leq j \leq N} \left\| x - x_j \right\|_2. \tag{2.40}$$

Alors, pour h suffisamment petit, l'erreur de l'interpolant $s_{f,X}$ de $f \in V$ sur X pour la spline naturelle de type (2.8), est bornée par l'inégalité

$$\left| f\left(x\right) - s_{f,X}\left(x\right) \right| \leq c \left| f \right|_k \begin{cases} h^{\frac{1}{2}} \ si \ 2k-d=1, \\ h \left| \ln h \right|^{\frac{1}{2}} \ si \ 2k-d=2 \\ h \ si \ 2k-d \geq 3, \end{cases}, \ x \in \Omega, \tag{2.41}$$

où c est une constante positive qui ne depend que de d, q et Ω et non de X, f ou x.

La preuve du théorème 17 repose sur l'inégalité (2.23) dans laquelle

$$\Phi\left(x,\widetilde{X}\right)=\left(w_{\widetilde{X}}\left(x\right)\right)^{2}=\begin{array}{l}c_{k,d}\left\{\sum\limits_{r,s=1}^{l}p_{r}\left(x\right)p_{s}\left(x\right)\phi\left(\left\|x_{r}-x_{s}\right\|_{2}\right)\right.\\ \left.-2\sum\limits_{r=1}^{l}p_{r}\left(x\right)\phi\left(\left\|x-x_{r}\right\|_{2}\right)\right\}.\end{array}\tag{2.42}$$

Pour tout $x\in\Omega$ fixé, la preuve du théorème 17 demande le choix d'un sous-ensemble $\pi_{k-1}\left(\mathbb{R}^{d}\right)$-unisolvant $\widetilde{X}=\{x_1,...,x_l\}$ de l'ensemble d'interpolation $X=\{x_1,...,x_N\}$. On supposera donc sans perte de généralité que

$$\Omega\cup\widetilde{X}\subset B\left(0,\frac{1}{2}\right),\tag{2.43}$$

où

$$B\left(0,\frac{1}{2}\right)=\left\{z\in\mathbb{R}^{d};\ \ \left\|z\right\|_{2}\leq\frac{1}{2}\right\}.\tag{2.44}$$

Puisque d'après le lemme 16, l'affectation d'un coefficient positif σ aux vecteurs de $\mathbb{R}^d$ n'a pour effet que la multiplication de (2.19) par σ^{2k-d} lorsque les coefficients $\gamma_r=p_r\left(x\right)$, $r=1,...,l$ de (2.15) sont conservés. Par ailleurs, la propriété de l'invariance par multiplication scalaire de notre problème d'interpolation (1.1), *i.e.*,

$$\forall\lambda\in\mathbb{R},\ s_{\lambda f,X}\left(x_i\right)=\lambda f\left(x_i\right),\ 1\leq i\leq N,\tag{2.45}$$

montre que l'interpolant $s_{\overline{f},\sigma X}$ de la nouvelle fonction

$$\overline{f}\left(y\right)=f\left(\sigma^{-1}y\right),\ y\in\mathbb{R}^{d},\tag{2.46}$$

est défini par

$$s_{\overline{f},\sigma X}\left(y\right)=s_{f,X}\left(\sigma^{-1}y\right),\ y\in\mathbb{R}^{d};\tag{2.47}$$

puisqu'en posant

$$y=\sigma x,\ x\in X,\tag{2.48}$$

on a

$$y\in\sigma X\Leftrightarrow\sigma^{-1}y\in X.\tag{2.49}$$

Ainsi

$$f\left(x\right)-s_{f,X}\left(x\right)=\overline{f}\left(\sigma x\right)-s_{\overline{f},\sigma X}\left(\sigma x\right),\ \forall x\in\mathbb{R}^{d}.\tag{2.50}$$

Comme $\Phi\left(\sigma x, \sigma\widetilde{X}\right) = \sigma^{2k-d}\Phi\left(x, \widetilde{X}\right)$ d'après le lemme 16, et

$$|f(x) - s_{f,X}(x)| \leq \left(\Phi\left(x, \widetilde{X}\right)\right)^{\frac{1}{2}} |f|_k$$

d'après (2.23), en posant $y = \sigma x$, $x \in \widetilde{X}$, on a

$$\begin{aligned} \left|\overline{f}(y) - s_{\overline{f},\sigma X}(y)\right| &\leq \left(\sigma^{-(2k-d)}\Phi\left(\sigma x, \sigma\widetilde{X}\right)\right)^{\frac{1}{2}} |f|_k \\ &= \Phi\left(\sigma x, \sigma\widetilde{X}\right)^{\frac{1}{2}} \left(\sigma^{-(2k-d)} |f|_k^2\right)^{\frac{1}{2}} \\ &= \Phi\left(\sigma x, \sigma\widetilde{X}\right)^{\frac{1}{2}} \left|\overline{f}\right|_k, \end{aligned} \tag{2.51}$$

où

$$\left|\overline{f}\right|_k = \left(\sigma^{-(2k-d)} |f|_k^2\right)^{\frac{1}{2}}. \tag{2.52}$$

Ces arguments montrent que nos estimations de l'erreur basées sur (2.23) sont scalairement invariants. Notons que si $\widetilde{X} = \{x_1, ..., x_l\}$ est un sous-ensemble $\pi_{k-1}\left(\mathbb{R}^d\right)$-unisolvant de X fixé, et si on considère une suite $(x_n) \subset \Omega$ de valeurs de x telle que

$$\|x_n - x_1\|_2 \longrightarrow 0, \tag{2.53}$$

alors, d'après (2.15) on peut définir

$$\gamma_i^n = p_i(x_n),\ i = 1, ..., l, \tag{2.54}$$

et par continuité des p_i , $i = 1, ..., l$, on a

$$\gamma_i^n \longrightarrow p_i(x_1),\ i = 1, ..., l. \tag{2.55}$$

D'où

$$\begin{aligned} \gamma_1^n &\longrightarrow 1 \text{ et} \\ \gamma_i^n &\longrightarrow 0,\ i = 2, ..., l. \end{aligned} \tag{2.56}$$

(2.53) et (2.56) implique qu'il existe un entier $n_0 \in \mathcal{N}$ tel que $\forall n \geq n_0$ on a,

$$\begin{cases} \|x_n - x_1\|_2 \leq h \\ |\gamma_1^n - 1| \leq \widetilde{c}h \\ |\gamma_2^n| \leq \widetilde{c}h \\ \vdots \\ |\gamma_l^n| \leq \widetilde{c}h, \end{cases} \tag{2.57}$$

où $\widetilde{c}$ est une constante positive donnée. En posant

$$x_{n_0} = x, \text{ on a } \gamma_i^{n_0} = p_i\left(x_{n_0}\right) = p_i\left(x\right) = \gamma_i,\ i = 1, ..., l. \tag{2.58}$$

D'où

$$\begin{array}{l} \left\|x - x_1\right\|_2 \leq h \\ \left|\gamma_1 - 1\right| \leq \widetilde{c}h \\ \left|\gamma_2\right| \leq \widetilde{c}h \\ \vdots \\ \left|\gamma_l\right| \leq \widetilde{c}h, \end{array} \tag{2.59}$$

par suite

$$\left\|x - x_1\right\|_2 \leq h, \tag{2.60}$$

et

$$\max\left\{\left|\gamma_1 - 1\right|, \left|\gamma_2\right|, ..., \left|\gamma_l\right|\right\} \leq \widetilde{c}h, \tag{2.61}$$

où $\gamma_1, \gamma_2, ..., \gamma_l$ sont les coefficients définis par (2.15).

Lemme 18 *Soit Ω, X et h définis au théorème 17 avec $h \leq e^{-1}$. Fixons un point $x \in \Omega$ et prenons un ensemble $\widetilde{X}$ qui satisfait les conditions du dernier paragraphe. On suppose en outre qu'aucun des points $x_2, ..., x_l$ n'appartient au segment d'extremités x_1 et x. Alors l'expression correspondant (2.20) a la propriété*

$$\Phi\left(x, \widetilde{X}\right) \leq c^2 \left\{ \begin{array}{c} h \text{ si } 2k - d = 1, \\ h^2\left|\ln h\right| \text{ si } 2k - d = 2, \\ h^2 \text{ si } 2k - d \geq 3. \end{array} \right. \tag{2.62}$$

Démonstration. En Transformant l'expression (2.42),

$$\begin{array}{rl} \Phi\left(x, \widetilde{X}\right) = \left(w_{\widetilde{X}}\left(x\right)\right)^2 = & c_{k,d}\left\{ \sum\limits_{r,s=1}^{l} p_r\left(x\right) p_s\left(x\right) \phi\left(\left\|x_r - x_s\right\|_2\right) \right. \\ & \left. -2\sum\limits_{r=1}^{l} p_r\left(x\right) \phi\left(\left\|x - x_r\right\|_2\right) \right\}, \end{array} \tag{2.63}$$

où

$$p_r\left(x\right) = \gamma_r,\ r = 1, ..., l, \tag{2.64}$$

on a

$$\begin{aligned}\Phi\left(x,\widetilde{X}\right) &= 2c_{k,d}\left\{\tfrac{1}{2}\sum_{r,s=1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right)-\sum_{r=1}^{l}\gamma_r\phi\left(\|x-x_r\|_2\right)\right\}\\ &= 2c_{k,d}\left\{\begin{array}{c}\tfrac{1}{2}\sum_{r,s=1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right)-\sum_{r=2}^{l}\gamma_r\phi\left(\|x-x_r\|_2\right)\\ -\gamma_1\phi\left(\|x-x_1\|_2\right)\end{array}\right\}.\end{aligned} \tag{2.65}$$

Or

$$\begin{aligned}\tfrac{1}{2}\sum_{r,s=1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right) &= \tfrac{1}{2}\sum_{s=1}^{l}\gamma_1\gamma_s\phi\left(\|x_1-x_s\|_2\right)\\ &\quad+\tfrac{1}{2}\sum_{r=2}^{l}\sum_{s=1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right)\\ &= \tfrac{1}{2}\sum_{s=2}^{l}\gamma_1\gamma_s\phi\left(\|x_1-x_s\|_2\right)\\ &\quad+\tfrac{1}{2}\sum_{r=2}^{l}\sum_{s=1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right),\end{aligned}$$

et

$$\begin{aligned}\tfrac{1}{2}\sum_{r=2}^{l}\sum_{s=1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right) &= \tfrac{1}{2}\sum_{r=2}^{l}\gamma_1\gamma_r\phi\left(\|x_r-x_1\|_2\right)\\ &\quad+\tfrac{1}{2}\sum_{r=2}^{l}\sum_{s=2}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right).\end{aligned} \tag{2.66}$$

D'où

$$\begin{aligned}\tfrac{1}{2}\sum_{r,s=1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right) &= \sum_{r=2}^{l}\gamma_1\gamma_r\phi\left(\|x_r-x_1\|_2\right)\\ &\quad+\tfrac{1}{2}\sum_{r=2}^{l}\sum_{s=2}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right).\end{aligned} \tag{2.67}$$

On a

$$\begin{aligned}\sum_{r=2}^{l}\sum_{s=2}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right) &= \sum_{r=2}^{l}\sum_{s=2}^{r}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right)\\ &\quad+\sum_{r=2}^{l}\sum_{s=r}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right),\end{aligned} \tag{2.68}$$

car

$$\sum_{r=2}^{l}\gamma_r^2\phi\left(\|x_r-x_r\|_2\right)=0. \tag{2.69}$$

Par ailleurs on vérifie facilement que

$$\sum_{r=2}^{l}\sum_{s=r}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right)=\sum_{r=2}^{l-1}\sum_{s=r+1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right), \tag{2.70}$$

et on montre par réccurence que pour tout entier p tel que

$$3\leq p\leq l, \text{ on a } \sum_{r=2}^{p}\sum_{s=2}^{r}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right)=\sum_{r=2}^{p-1}\sum_{s=r+1}^{p}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right). \tag{2.71}$$

Par suite

$$\sum_{r=2}^{l}\sum_{s=2}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right)=2\sum_{r=2}^{l-1}\sum_{s=r+1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right). \tag{2.72}$$

(2.67) s'écrit donc

$$\begin{aligned}\tfrac{1}{2}\sum_{r,s=1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right) &= \sum_{r=2}^{l-1}\sum_{s=r+1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right)\\ &\quad+\sum_{r=2}^{l}\gamma_1\gamma_r\phi\left(\|x_r-x_1\|_2\right).\end{aligned} \tag{2.73}$$

En reportant ce résultat dans (2.65), on a

$$\begin{aligned}\Phi\left(x,\widetilde{X}\right) &= 2c_{k,d}\left\{\sum_{r=2}^{l-1}\sum_{s=r+1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right)+\sum_{r=2}^{l}\gamma_1\gamma_r\phi\left(\|x_r-x_1\|_2\right)\right.\\ &\quad\left.-\sum_{r=2}^{l}\gamma_r\phi\left(\|x-x_r\|_2\right)-\gamma_1\phi\left(\|x-x_1\|_2\right)\right\},\end{aligned} \tag{2.74}$$

d'où

$$\begin{aligned}\Phi\left(x,\widetilde{X}\right) &= 2c_{k,d}\left\{\sum_{r=2}^{l-1}\sum_{s=r+1}^{l}\gamma_r\gamma_s\phi\left(\|x_r-x_s\|_2\right)\right.\\ &\quad+\sum_{r=2}^{l}\left(\gamma_1-1\right)\gamma_r\phi\left(\|x_r-x_1\|_2\right)\\ &\quad\left.+\sum_{r=2}^{l}\gamma_r\left[\phi\left(\|x_r-x_1\|_2\right)-\phi\left(\|x-x_r\|_2\right)\right]-\gamma_1\phi\left(\|x-x_1\|_2\right)\right\}.\end{aligned} \tag{2.75}$$

Ainsi,

$$\Phi\left(x,\widetilde{X}\right)=2c_{k,d}\left\{T_1+T_2-\gamma_1\phi\left(\left\|x-x_1\right\|_2\right)\right\}, \tag{2.76}$$

où

$$T_1=\sum_{r=2}^{l-1}\sum_{s=r+1}^{l}\gamma_r\gamma_s\phi\left(\left\|x_r-x_s\right\|_2\right)+\sum_{r=2}^{l}\left(\gamma_1-1\right)\gamma_r\phi\left(\left\|x_r-x_1\right\|_2\right), \tag{2.77}$$

et

$$T_2=\sum_{r=2}^{l}\gamma_r\left[\phi\left(\left\|x_r-x_1\right\|_2\right)-\phi\left(\left\|x-x_r\right\|_2\right)\right]. \tag{2.78}$$

On va donc chercher à majorer chacun des trois termes dans la parenthèse de (2.76). On étudie alors les variations de la fonction

$$\varphi\left(t\right)=\left|\phi\left(t\right)\right|\ \ \forall t\in\left[0,1\right]. \tag{2.79}$$

Si d est impair on a

$$\varphi\left(t\right)=t^{2k-d}\text{ et }\varphi'\left(t\right)=\left(2k-d\right)t^{2k-d-1}, \tag{2.80}$$

$$\varphi\text{ est alors monotone strictement croissante sur }\left[0,1\right]. \tag{2.81}$$

Si d est pair on a

$$\varphi\left(t\right)=-t^{2k-d}\ln t\text{ et }\varphi'\left(t\right)=-t^{2k-d-1}\left[1+\left(2k-d\right)\ln t\right], \tag{2.82}$$

$$\varphi\text{ est alors monotone}\begin{cases}\text{strictement croissante sur }\left[0,\exp\left(\frac{-1}{2k-d}\right)\right],\\ \text{strictement décroissante sur }\left[\exp\left(\frac{-1}{2k-d}\right),1\right].\end{cases} \tag{2.83}$$

Afin de majorer le dernier terme de (2.76), notons que (2.61) implique

$$\left|\gamma_1\right|\leq 1+\left|\gamma_1-1\right|\leq 1+\widetilde{c}h\leq 1+\widetilde{c}e^{-1}. \tag{2.84}$$

Par ailleurs puisque $h\leq e^{-1}\leq\exp\left(\frac{-1}{2k-d}\right)$, (2.60) et l'étude des variations de

la fonction φ implique

$$|\phi(\|x-x_1\|_2)| \leq |\phi(h)| = \begin{cases} h^{2k-d}, \text{ si } d \text{ est impair,} \\ h^{2k-d}|\ln h|, \text{ si } d \text{ est pair.} \end{cases} \tag{2.85}$$

Par suite,

$$|-\gamma_1\phi(\|x-x_1\|_2)| \leq \left(1+\widetilde{c}e^{-1}\right) \begin{cases} h^{2k-d}, \text{ si } d \text{ est impair,} \\ h^{2k-d}|\ln h|, \text{ si } d \text{ est pair.} \end{cases} \tag{2.86}$$

Afin de majorer T_1, notons que 2.43 implique

$$\|x_r - x_s\|_2 \leq 1,\ r,s = 1,...,l. \tag{2.87}$$

Posons

$$x_0 = \exp\left(\frac{-1}{2k-d}\right), \tag{2.88}$$

alors $\ln x_0 = \frac{-1}{2k-d}$, d'où

$$x_0^{2k-d} = \exp\left(\ln x_0^{2k-d}\right) = \exp\left((2k-d)\ln x_0\right) = e^{-1}. \tag{2.89}$$

Comme

$$\varphi(x_0) = |\phi(x_0)| = \begin{cases} x_0^{2k-d}, \text{ si } d \text{ est impair,} \\ x_0^{2k-d}|\ln x_0|, \text{ si } d \text{ est pair,} \end{cases} \tag{2.90}$$

on a

$$\varphi(x_0) = \begin{cases} \frac{1}{e}, \text{ si } d \text{ est impair,} \\ \frac{1}{(2k-d)e}, \text{ si } d \text{ est pair.} \end{cases} \tag{2.91}$$

Par ailleurs,

$$\varphi(1) = \begin{cases} 1, \text{ si } d \text{ est impair,} \\ 0, \text{ si } d \text{ est pair.} \end{cases} \tag{2.92}$$

(2.91) et (2.92) implique

$$\max\{\varphi(x_0), \varphi(1)\} = \begin{cases} 1, \text{ si } d \text{ est impair,} \\ \frac{1}{(2k-d)e}, \text{ si } d \text{ est pair.} \end{cases} \tag{2.93}$$

Comme $\max\left\{1, \frac{1}{(2k-d)e}\right\} = 1$, par les propriétés de la monotonie de φ, il est clair que

$$\forall t \in [0,1],\ \varphi(t) = |\phi(t)| \leq \max\{\varphi(x_0), \varphi(1)\} \leq 1. \tag{2.94}$$

(2.61), (2.77), (2.86) et (2.94) implique

$$\begin{array}{rcl} |T_1| & \leq & \sum\limits_{r=2}^{l-1}\sum\limits_{s=r+1}^{l} |\gamma_r|\,|\gamma_s|\,|\phi\left(\|x_r - x_s\|_2\right)| \\ & & +\sum\limits_{r=2}^{l} |(\gamma_1 - 1)|\,|\gamma_r|\,|\phi\left(\|x_r - x_1\|_2\right)| \\ & \leq & \frac{(l-2)(l-1)}{2}\widetilde{c}^2h^2 + (l-1)\,\widetilde{c}^2h^2, \end{array} \tag{2.95}$$

d'où

$$|T_1| \leq \frac{l\,(l-1)}{2}\widetilde{c}^2h^2. \tag{2.96}$$

Afin de majorer T_2, Posons

$$\psi_r\left(\theta\right) = \phi\left(\|\theta x_1 + (1-\theta)\,x - x_r\|_2\right),\ 0 \leq \theta \leq 1,\ r = 2, ..., l. \tag{2.97}$$

Ainsi, (2.78) devient

$$T_2 = \sum_{r=2}^{l}\gamma_r\left[\psi_r\left(1\right) - \psi_r\left(0\right)\right]. \tag{2.98}$$

L'hypothèse qu'aucun des points $x_2, ..., x_l$ n'appartient au segment d'extremités x_1 et x implique que ψ_r a sa dérivée première continue.

$$\psi_r'\left(\theta\right) = F'\left(\theta\right)\phi'\left(F\left(\theta\right)\right), \tag{2.99}$$

où

$$F\left(\theta\right) = \|y\left(\theta\right)\|_2 \text{ avec } y\left(\theta\right) = \theta\left(x_1 - x\right) + x - x_r,\ 0 \leq \theta \leq 1. \tag{2.100}$$

On a

$$\begin{array}{rcl} F\left(\theta\right) & = & \|(1-\theta)\left(x - x_1\right) + x_1 - x_r\|_2 \\ & \leq & \|x - x_1\|_2 + \|x_1 - x_r\|_2 \\ & \leq & h + 1 \\ & \leq & e^{-1} + 1, \end{array} \tag{2.101}$$

$$\begin{array}{rcl} |F\left(\theta + u\right) - F\left(\theta\right)| & = & |\|y\left(\theta + u\right)\|_2 - \|y\left(\theta\right)\|_2| \\ & \leq & \|y\left(\theta + u\right) - y\left(\theta\right)\|_2 \\ & = & u\,\|x_1 - x\|_2, \end{array} \tag{2.102}$$

ainsi

$$\left|\frac{F\left(\theta + u\right) - F\left(\theta\right)}{u}\right| \leq \|x_1 - x\|_2,\ \forall u \in\ ]0, 1[\,. \tag{2.103}$$

Par passage à la limite quand $u \to 0^+$, on a

$$|F'(\theta)| \leq \|x_1 - x\|_2 \ \forall \theta \in [0,1]. \tag{2.104}$$

(2.99), (2.101) et (2.104) implique

$$|\psi_r'(\theta)| \leq \|x_1 - x\|_2 \max_{0 \leq t \leq e^{-1}+1} |\phi'(t)|, \ r = 2, ..., l \ , \forall \theta \in [0,1], \tag{2.105}$$

et comme $|\phi'|$ est une fonction continue sur $[0, +\infty[$ d'après (2.8), la quantité

$$c_0 = \max_{0 \leq t \leq e^{-1}+1} |\phi'(t)|, \tag{2.106}$$

est une constante positive qui ne dépend que de d et k. Par ailleurs l'inégalité généralisée des accroissements finis appliquée à ψ_r sur $[0,1]$ donne

$$|\psi_r(1) - \psi_r(0)| \leq \max_{0 \leq \theta \leq 1} |\psi_r'(\theta)|. \tag{2.107}$$

Par suite (2.60), (2.105), (2.106) et (2.107) implique

$$|\psi_r(1) - \psi_r(0)| \leq c_0 h, \ r = 2, ..., l. \tag{2.108}$$

Ainsi (2.61), (2.98) et (2.108) permet de majorer T_2 par

$$|T_2| \leq \sum_{r=2}^{l} |\gamma_r| \, |\psi_r(1) - \psi_r(0)| \leq (l-1) \, c_0 \widetilde{c} h^2. \tag{2.109}$$

D'après (2.76), on a

$$\Phi\left(x, \widetilde{X}\right) \leq 2 \, |c_{k,d}| \left[|T_1| + |T_2| + |-\gamma_1 \phi(\|x - x_1\|_2)| \right]. \tag{2.110}$$

(2.86), (2.96) et (2.109) implique

$$\begin{array}{rcl} \Phi\left(x, \widetilde{X}\right) & \leq & h \, |c_{k,d}| \left[\begin{array}{c} l(l-1)\, \widetilde{c}^2 h + 2(l-1)\, c_0 \widetilde{c} h \\ +2\left(1 + \widetilde{c} e^{-1}\right) \end{array} \right], \\ \text{si } 2k - d & = & 1, \end{array} \tag{2.111}$$

$$\begin{aligned}\Phi\left(x,\widetilde{X}\right) &\leq h^2\left|\ln h\right|\left|c_{k,d}\right|\left[\begin{array}{c} l\left(l-1\right)\widetilde{c}^2\left|\ln h\right|^{-1} \\ +2\left(l-1\right)c_0\widetilde{c}\left|\ln h\right|^{-1} \\ +2\left(1+\widetilde{c}e^{-1}\right)\end{array}\right], \qquad (2.112)\\ \text{si } 2k-d &= 2,\end{aligned}$$

comme

$$0<h\leq e^{-1}\leq 1, \qquad (2.113)$$

on a

$$\begin{aligned}\Phi\left(x,\widetilde{X}\right) &\leq h\left|c_{k,d}\right|\left[\begin{array}{c} l\left(l-1\right)\widetilde{c}^2+2\left(l-1\right)c_0\widetilde{c} \\ +2\left(1+\widetilde{c}e^{-1}\right)\end{array}\right], \qquad (2.114)\\ \text{si } 2k-d &= 1,\end{aligned}$$

$$\begin{aligned}\Phi\left(x,\widetilde{X}\right) &\leq h^2\left|\ln h\right|\left|c_{k,d}\right|\left[\begin{array}{c} l\left(l-1\right)\widetilde{c}^2+2\left(l-1\right)c_0\widetilde{c} \\ +2\left(1+\widetilde{c}e^{-1}\right)\end{array}\right], \qquad (2.115)\\ \text{si } 2k-d &= 2.\end{aligned}$$

Si $2k-d\geq 3$ en appliquant la formule des accroissements finis à ln sur $[h,1]$,

$$\text{il existe } \alpha\in\left]h,1\right[\text{ tel que } \left|\ln 1-\ln h\right|=\left(1-h\right)\left|\ln'\alpha\right|. \qquad (2.116)$$

Ainsi

$$h^{2k-d-2}\left|\ln h\right|=h^{2k-d-2}\frac{1-h}{\alpha}\leq h^{2k-d-3}\leq 1, \qquad (2.117)$$

car $h\leq 1$ et $2k-d-3\geq 0$. Or d'après (2.86), on a

$$\left|-\gamma_1\phi\left(\left\|x-x_1\right\|_2\right)\right|\leq h^2\left(1+\widetilde{c}e^{-1}\right)\left\{\begin{array}{l} h^{2k-d-2}, \text{ si } d \text{ est impair,} \\ h^{2k-d-2}\left|\ln h\right|, \text{ si } d \text{ est pair,}\end{array}\right. \qquad (2.118)$$

donc,

$$\left|-\gamma_1\phi\left(\left\|x-x_1\right\|_2\right)\right|\leq h^2\left(1+\widetilde{c}e^{-1}\right), \text{ si } 2k-d\geq 3. \qquad (2.119)$$

D'où

$$\Phi\left(x,\widetilde{X}\right)\leq h^2\left|c_{k,d}\right|\left[l\left(l-1\right)\widetilde{c}^2+2\left(l-1\right)c_0\widetilde{c}+2\left(1+\widetilde{c}e^{-1}\right)\right], \text{ si } 2k-d\geq 3. \qquad (2.120)$$

En posant

$$c^2 = |c_{k,d}| \left[l\left(l-1\right)\widetilde{c}^2 + 2\left(l-1\right)c_0\widetilde{c} + 2\left(1+\widetilde{c}e^{-1}\right)\right], \tag{2.121}$$

on a

$$\Phi\left(x,\widetilde{X}\right) \leq c^2 \begin{cases} h \text{ si } 2k-d=1, \\ h^2\left|\ln h\right| \text{ si } 2k-d=2, \\ h^2 \text{ si } 2k-d\geq 3. \end{cases} \tag{2.122}$$

D'où le résultat du lemme 18. ■

Notons que le cas $d = k = 1$ est plus simple car dans ce cas $l = 1$, ce qui implique $T_1 = T_2 = 0$. Avant de donner la preuve du théorème 17 on introduit la notation suivante : Pour un ensemble quelconque $\{v_1, v_2, ..., v_l\}$ de vecteurs de $\mathbb{R}^d$ on note

$$D\left(v_1, v_2, ..., v_l\right) = \det\left(P_j\left(v_i\right)\right)_{1\leq i,j\leq l}, \tag{2.123}$$

où $\{P_1, P_2 ... P_l\}$ est la base des monômes de $\pi_{k-1}\left(\mathbb{R}^d\right)$.

La $\pi_{k-1}\left(\mathbb{R}^d\right)$-unisolvance de $\widetilde{X}$ est donc équivalente à la condition

$$D\left(x_1, x_2, ..., x_l\right) \neq 0. \tag{2.124}$$

On a par ailleurs

$$D\left(x_1, x_2, ..., x_{i-1}, x, x_{i+1}, ..., x_l\right) = \begin{vmatrix} P_1(x_1) & P_2(x_1) & ... & P_{l-1}(x_1) & P_l(x_1) \\ P_1(x_2) & P_2(x_2) & ... & P_{l-1}(x_2) & P_1(x_2) \\ \vdots & \vdots & & \vdots & \vdots \\ P_1(x_{i-1}) & P_2(x_{i-1}) & ... & P_{l-1}(x_{i-1}) & P_l(x_{i-1}) \\ P_1(x) & P_2(x) & ... & P_{l-1}(x) & P_l(x) \\ P_1(x_{i+1}) & P_2(x_{i+1}) & ... & P_{l-1}(x_{i+1}) & P_l(x_{i+1}) \\ \vdots & \vdots & & \vdots & \vdots \\ P_1(x_l) & P_2(x_l) & ... & P_{l-1}(x_l) & P_l(x_l) \end{vmatrix} \tag{2.125}$$

Le développement de ce déterminant suivant la $i-ième$ ligne donne

$$D\left(x_1, x_2, ..., x_{i-1}, x, x_{i+1}, ..., x_l\right) = \sum_{j=1}^{l} (-1)^{i+j} P_j(x) \Delta_{ij}, \tag{2.126}$$

où Δ_{ij} désigne le mineur relatif à $P_j(x_i)$. Or

$$P_j(x) = \sum_{i=1}^{l} p_i(x) P_j(x_i), \; j = 1, ..., l. \tag{2.127}$$

On sait que $\gamma_i = p_i(x)$, $i = 1, ..., l$. En notant

$$\begin{array}{lcl} P & = & (P_j(x_i))_{1\leq i,j\leq l} \\ \gamma & = & (p_i(x))_{1\leq i\leq l} \\ \lambda & = & (P_i(x))_{1\leq i\leq l} \\ co(P) & = & \left((-1)^{i+j}\Delta_{ij}\right)_{1\leq i,j\leq l}, \text{ la comatrice de } P, \end{array} \tag{2.128}$$

l'égalité (2.127) s'écrit sous la forme matricielle

$$\lambda = P^T\gamma, \tag{2.129}$$

d'où

$$\gamma = \left(P^T\right)^{-1}\lambda = \left(P^{-1}\right)^T\lambda, \tag{2.130}$$

avec

$$P^{-1} = \frac{1}{\det(P)}\left(co(P)\right)^T. \tag{2.131}$$

D'où

$$\left(P^{-1}\right)^T = \frac{1}{\det(P)} co(P). \tag{2.132}$$

Ainsi pour tout $i = 1, ..., l$ on a

$$\gamma_i = p_i(x) = \frac{\sum_{j=1}^{l}(-1)^{i+j} P_j(x)\Delta_{ij}}{\det(P)} = \frac{D(x_1, x_2, ..., x_{i-1}, x, x_{i+1}, ..., x_l)}{D(x_1, x_2, ..., x_l)}. \tag{2.133}$$

La preuve du théorème 17 repose sur la propriété de (2.123) qui est donnée dans le lemme suivant. On défini en plus pour chaque $x \in \overline{\Omega}$, le nombre

$$\Delta(x) = \max\left\{|D(x, v_2, ..., v_l)|; \; v_i \in \overline{\Omega}, \; i = 2, ..., l\right\}. \tag{2.134}$$

Lemme 19 *Il existe une constante $M > 0$ qui ne dépend que de d et k telle que pour tout vecteur $y_1, y_2, u_2, u_3, ...u_l$ de la boule $B\left(0, \frac{1}{2}\right)$ nous ayons*

$$|D(y_1, u_2, u_3, ..., u_l) - D(y_2, u_2, u_3, ..., u_l)| \leq M \|y_1 - y_2\|_2 . \tag{2.135}$$

Par conséquent, la fonction $\left\{\Delta(x);\ x \in \overline{\Omega}\right\}$ est lipschitzienne continue.

$$\begin{array}{llll} P_1(y_1) - P_1(y_2) & P_2(y_1) - P_2(y_2) & ... & P_l(y_1) - P_l(y_2) \\ P_1(u_2) & P_2(u_2) & ... & P_l(u_2) \\ \vdots & \vdots & & \vdots \\ P_1(u_l) & P_2(u_l) & ... & P_l(u_l) \end{array}$$

Lemme 20 ***Démonstration. Démonstration.*** *De la définition (2.123) on a*

$$\begin{aligned} &D(y_1, u_2, u_3, ..., u_l) - D(y_2, u_2, u_3, ..., u_l) = \\ &\det \begin{bmatrix} P_1(y_1) - P_1(y_2) & ... & P_l(y_1) - P_l(y_2) \\ P_1(u_2) & ... & P_l(u_2) \\ \vdots & \vdots & \vdots \\ P_1(u_l) & ... & P_l(u_l) \end{bmatrix} \\ &= \sum_{j=1}^{l} (-1)^{1+j} (P_j(y_1) - P_j(y_2)) \Delta_{1j}, \end{aligned} \tag{2.136}$$

où $\{P_1, P_2...P_l\}$ est la base des monômes de $\pi_{k-1}\left(\mathbb{R}^d\right)$. Comme $y_1, y_2 \in B\left(0, \frac{1}{2}\right)$, on a

$$[y_1, y_2] \subset B\left(0, \frac{1}{2}\right) \subset B_0(0, 1), \tag{2.137}$$

où $B_0(0, 1)$ est la boule unité ouverte de $\mathbb{R}^d$ centrée en 0. Par ailleurs

$$\forall j = 1, ..., l,\ P_j : B_0(0, 1) \subset \mathbb{R}^d \longrightarrow \mathbb{R}, \tag{2.138}$$

est continue sur $[y_1, y_2]$, différentiable en tout point de$]y_1, y_2[$. De plus,

$$\text{il existe } C > 0 \text{ telle que } \|dP_j(c)\|_* \leq C \text{ pour tout } c \in]y_1, y_2[. \tag{2.139}$$

En effet,

$$\begin{aligned} \forall h,\ \|h\|_2 \leq 1 \Rightarrow |dP_j(c)\cdot h| &= \left|\sum_{i=1}^{d} \frac{\partial P_j}{\partial x_i}(c)\, h_i\right| \\ &\leq \sum_{i=1}^{d} \left|\frac{\partial P_j}{\partial x_i}(c)\right| |h_i| \\ &\leq \|\nabla P_j(c)\|_2 . \end{aligned} \tag{2.140}$$

Posons

$$C = \max_{1\leq j\leq l} \|\nabla P_j(c)\|_2 . \tag{2.141}$$

Le théorème de l'inégalité des accroissements finis appliqué à P_j sur $[y_1, y_2]$, montre qu'il existe $C > 0$ ne dépendant que de d et k telle que

$$|P_j(y_1) - P_j(y_2)| \leq C \|y_1 - y_2\|_2 ,\ j = 1, ..., l. \tag{2.142}$$

D'après (2.136), on a

$$|D(y_{1,}u_2, u_3, ..., u_l) - D(y_2, u_2, u_3, ..., u_l)| \leq M \|y_1 - y_2\|_2 ,$$

où

$$M = C\sum_{j=1}^{l} |\Delta_{1j}| ,\ \textit{ne dépend que de } d \textit{ et } k. \tag{2.143}$$

La continuité lipschitzienne de $\Delta(x)$ résulte du fait que si $y_1, y_2 \in \overline{\Omega}$ satisfait $\Delta(y_1) \geq \Delta(y_2)$ et si $\Delta(y_1) = |D(y_{1,}u_2, u_3, ..., u_l)|$ avec $u_i \in \overline{\Omega}$, $i = 2, ..., l$, alors

$$\begin{aligned} \Delta(y_1) - \Delta(y_2) &= |D(y_{1,}u_2, u_3, ..., u_l)| - |D(y_{2,}u_2, u_3, ..., u_l)| \\ &\leq |D(y_{1,}u_2, u_3, ..., u_l) - D(y_2, u_2, u_3, ..., u_l)| \\ &\leq M \|y_1 - y_2\|_2, \end{aligned} \tag{2.144}$$

où l'on a utilisé l'hypotèse (2.43) que $\Omega \subset B\left(0, \frac{1}{2}\right)$. ■ ■

Démonstration. Du Théorème 17 **:** Par des arguments donnés au début de cette section, on suppose sans perte de généralité que (2.43) est vérifiée. On pose

$$\Delta = \inf\left\{\Delta(x)\ ;\ x \in \overline{\Omega}\right\}, \tag{2.145}$$

qui est un nombre positif par le fait de la continuité de $\Delta(x)$ sur $\overline{\Omega}$ (lemme 19) et le fait que Ω contient au moins un ensemble de points $\pi_{k-1}\left(\mathbb{R}^d\right)$-unisolvant. Posons

$$h_0 = \frac{\Delta}{(l+2)\,M}, \tag{2.146}$$

où M est donnée au lemme 10. Soit $X = \{x_1, ..., x_N\}$ un ensemble fini de points de $\mathbb{R}^d$ et soit $h = h_X$ donné par (2.40). Nous allons donner la preuve du théorème lorsque

$$h \leq \min\left\{h_0, e^{-1}\right\}. \tag{2.147}$$

Soit alors x un point générique de Ω. L'essentiel de la preuve sera d'établir l'existence des points $x_{1,}x_2, ..., x_l$ de X tels que les conditions du lemme 18 soient satisfaites car dans ce cas le résultat demandé serait une conséquence de (2.23). D'après la définition (2.134), il existe $v_2, ..., v_l$ dans $\overline{\Omega}$ tels que $\Delta(x) = |D(x_{,}v_2, ..., v_l)|$. On en déduit donc de (2.145) que

$$|D(x_{,}v_2, ..., v_l)| \geq \Delta. \tag{2.148}$$

De la définition (2.40) il résulte que

$$\min_{1\leq i\leq N} \|t - x_i\|_2 \leq h,\ \forall t \in \Omega, \tag{2.149}$$

comme $\min\limits_{1\leq i\leq N} \|t - x_i\|_2 = \|t - x_r\|_2$ pour un certain entier $r \in \{1, 2, ..., N\}$, alors $t \in B(x_r, h)$. (2.149) implique donc que

$$\Omega \subset \bigcup_{i=1}^{N} B(x_i, h). \tag{2.150}$$

Or $\bigcup\limits_{i=1}^{N} B(x_i, h)$ est fermée. Donc

$$\overline{\Omega} \subset \bigcup_{i=1}^{N} B(x_i, h). \tag{2.151}$$

Aux points $x, v_2, ..., v_l$ de $\overline{\Omega}$ on associe donc les entiers $r_1, r_2, ..., r_l \in \{1, 2, ..., N\}$ tels que

$$\begin{array}{l} \|x - x_{r_1}\|_2 \leq h, \\ \|v_2 - x_{r_2}\|_2 \leq h, \\ \vdots \\ \|v_l - x_{r_l}\|_2 \leq h. \end{array} \tag{2.152}$$

En notant $x_{1,}x_2, ..., x_l$ les vecteurs $x_{r_1}, x_{r_2}, ..., x_{r_l}$, on a

$$\max\left\{\|x - x_1\|_2, \|v_2 - x_2\|_2, ..., \|v_l - x_l\|_2\right\} \leq h. \tag{2.153}$$

Montrons à présent que

$$\widetilde{X} = \{x_1, ..., x_l\} \text{ est } \pi_{k-1}\left(\mathbb{R}^d\right)\text{-unisolvant, } i.e., D\left(x_{1,}x_2, ..., x_l\right) \neq 0. \tag{2.154}$$

On va utiliser (2.148), (2.153) et appliquer (2.135) l fois. On a

$$\begin{array}{lcl}
\left|D\left(x, v_2, v_3, ..., v_l\right) - D\left(x_{1,}v_2, v_3, ..., v_l\right)\right| & \leq & M\left\|x - x_1\right\|_2 \\
\left|D\left(x_{1,}v_2, v_3, ..., v_l\right) - D\left(x_{1,}x_2, v_3, ..., v_l\right)\right| & \leq & M\left\|v_2 - x_2\right\|_2 \\
\left|D\left(x_{1,}v_2, v_3, ..., v_l\right) - D\left(x_{1,}x_2, x_3, ..., v_l\right)\right| & \leq & M\left\|v_3 - x_3\right\|_2 \\
\vdots & \vdots & \vdots \\
\left|D\left(x_{1,}x_2, x_3, ..., v_{l-1}, v_l\right) - D\left(x_{1,}x_2, x_3, ..., x_{l-1}, v_l\right)\right| & \leq & M\left\|v_{l-1} - x_{l-1}\right\|_2 \\
\left|D\left(x_{1,}x_2, x_3, ..., x_{l-1}, v_l\right) - D\left(x_{1,}x_2, x_3, ..., x_{l-1}, x_l\right)\right| & \leq & M\left\|v_l - x_l\right\|_2.
\end{array} \tag{2.155}$$

Ainsi

$$\left|D\left(x, v_2, v_3, ..., v_l\right) - D\left(x_{1,}x_2, x_3, ..., x_{l-1}, x_l\right)\right| \leq lMh, \tag{2.156}$$

par suite

$$\left|D\left(x_{,}v_2, v_3, ..., v_l\right)\right| - \left|D\left(x_{1,}x_2, x_3, ..., x_l\right)\right| \leq lMh, \tag{2.157}$$

d'où

$$\left|D\left(x_{1,}x_2, x_3, ..., x_l\right)\right| \geq \Delta - lMh \geq \Delta - lMh_0 = 2Mh_0 > 0. \tag{2.158}$$

La dernière étape consiste à prouver l'inégalité

$$\left\|x - v_i\right\|_2 > 2h,\ i = 2, ..., l, \tag{2.159}$$

car dans ce cas (2.153) implique $\|x - x_i\|_2 + h \geq \|x - x_i\|_2 + \|v_i - x_i\|_2 \geq \|x - v_i\|_2 > 2h$. Donc

$$\left\|x - x_i\right\|_2 > h,\ i = 2, ..., l, \tag{2.160}$$

de telle sorte qu'aucun des points $x_2, x_3, ..., x_l$ n'appartient au segment $[x, x_1]$.

Supposons au contraire que

$$\|x - v_k\|_2 \leq 2h, \tag{2.161}$$

pour au moins un entier $k \in \{2, 3, ..., l\}$. Comme $\Omega \subset B\left(0, \frac{1}{2}\right)$, dans (2.135) on pose $y_1 = x$, $y_2 = v_k$, $u_j = v_j$, $j = 2, ..., l$. Par conséquent on a

$$|D\left(x, v_2, v_3, ..., v_l\right)| \leq M \|x - v_k\|_2 . \tag{2.162}$$

D'après (2.146), (2.147), (2.148), (2.161) et (2.162), $\Delta \leq |D\left(x, v_2, ..., v_l\right)| \leq M \|x - v_k\|_2 \leq 2Mh \leq 2Mh_0 < \Delta$, ce qui est contradictoire. Par suite (2.159) est vérifiée. Comme (2.60) est contenu dans (2.153), la seule hypothèse à vérifier dans le lemme 18 est l'hypothèse (2.61) où nous utilisons la définition (2.133) des coefficients $\{\gamma_1, \gamma_2, ..., \gamma_l\}$. Pour $i = 2, ..., l$, notons que $\widetilde{X} \subset B\left(0, \frac{1}{2}\right)$ nous permet d'estimer le numérateur de (2.133). De même dans (2.135) si l'on prend $y_1 = x_1, y_2 = x$ et $u_i = x_i$, $i = 2, ..., l$, alors

$$|D\left(x_1, x_2, ..., x_{i-1}, x, x_{i+1}, ..., x_l\right)| \leq M \|x - x_1\|_2 . \tag{2.163}$$

Ainsi (2.133), (2.60), (2.146), (2.158) et (2.163) implique

$$|\gamma_i| \leq \frac{Mh}{\Delta - lMh_0} = \frac{h}{2h_0}, \; i = 2, ..., l. \tag{2.164}$$

Dans (2.135) si l'on prend $y_1 = x$, $y_2 = x_1$ et $u_i = x_i$, $i = 2, ..., l$, alors

$$\begin{aligned} |\gamma_1 - 1| = \left| \frac{D\left(x, x_2, ..., x_l\right)}{D\left(x_1, x_2, ..., x_l\right)} - 1 \right| = & \frac{|D\left(x, x_2, ..., x_l\right) - D\left(x_1, x_2, ..., x_l\right)|}{|D\left(x_1, x_2, ..., x_l\right)|} \\ & \leq \frac{M \|x - x_1\|_2}{\Delta - lMh_0} \leq \frac{Mh}{2Mh_0} = \frac{h}{2h_0}. \end{aligned} \tag{2.165}$$

Il s'en suit de (2.164) et (2.165), que (2.61) est satisfait en posant

$$\widetilde{c} = \frac{1}{2h_0}, \tag{2.166}$$

qui est une constante qui ne dépend que de d, k et Ω. De plus les modifications de tous les arguments précédents quand $d = k = 1$ sont plus simples. La preuve du théorème 17 est alors complète. ■

2.2 Quelques applications de la méthode de la fonction puissance

Dans le cas où $d = k = 2$ on a $\phi(r) = r^2 \ln r$; $l = \dim \pi_1(\mathbb{R}^2) = 3$. Il s'en suit les théorèmes suivants :

Théorème 21 *(Powell) : Si $x \in \mathbb{R}^2$ est sur un segment joignant deux points donnés x_1, x_2, alors l'erreur entre une fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ et sa thin plate spline interpolant, $s_{f,X}$, est bornée par l'inégalité*

$$|f(x) - s_{f,X}(x)| \leq \left[\frac{\ln 2}{16\pi}\right]^{\frac{1}{2}} |f|_2 \, h, \tag{2.167}$$

où $h = \|x_1 - x_2\|_2$ et $|\cdot|_2$ désigne la semi-norme associée au produit semi-scalaire $\langle ., . \rangle$.

Démonstration. Soit $x \in [x_1, x_2]$ de la forme

$$x = \lambda x_1 + (1 - \lambda) x_2 \text{ pour un certain } \lambda \in [0, 1]. \tag{2.168}$$

Nous pouvons alors poser

$$x = \sum_{i=1}^{3} \theta_i x_i \text{ avec } \theta_1 = \lambda, \ \theta_2 = 1 - \lambda, \ \theta_3 = 0, \tag{2.169}$$

où x_3 est tel que (x_1, x_2, x_3) est un triangle non dégénéré. Prenons $\widetilde{X} = \{x_1, x_2, x_3\}$ et (e_1, e_2) la base canonique de $\mathbb{R}^2$. Soient $q_1, q_2, q \in \pi_1(\mathbb{R}^2)$ tels que

$$\forall y = (y_1, y_2) \in \mathbb{R}^2 \begin{cases} q_1(y) = y^{(1,0)} = y_1 \\ q_2(y) = y^{(0,1)} = y_2 \\ q(y) = y^{(0,0)} = 1. \end{cases} \tag{2.170}$$

On sait d'après (2.18) que $\forall p \in \pi_1(\mathbb{R}^2)$, $p(x) = \sum_{j=1}^{3} \gamma_j p(x_j)$ avec $\gamma_j = p_j(x)$.

En appliquant cette propriété à $q_1(x)$, $q_2(x)$ et $q(x)$, on a

$$\begin{aligned} &(x, e_1) = \left(\sum_{j=1}^{3} \gamma_j x_j, e_1 \right), \\ &(x, e_2) = \left(\sum_{j=1}^{3} \gamma_j x_j, e_2 \right), \\ &\sum_{j=1}^{3} \gamma_j = 1. \end{aligned} \tag{2.171}$$

D'où

$$\begin{aligned} &\left(x - \sum_{j=1}^{3} \gamma_j x_j, e_1 \right) = 0, \\ &\left(x - \sum_{j=1}^{3} \gamma_j x_j, e_2 \right) = 0, \\ &\sum_{j=1}^{3} \gamma_j = 1. \end{aligned} \tag{2.172}$$

(2.172) implique

$$\begin{cases} x = \sum_{j=1}^{3} \gamma_j x_j, \\ \sum_{j=1}^{3} \gamma_j = 1. \end{cases} \tag{2.173}$$

Les $\gamma_j = p_j(x)$, $j = 1, ..., 3$, sont donc les coordonnées barycentriques de x par rapport à x_1, x_2, x_3. Or d'après (2.169)

$$\begin{aligned} &x = \sum_{j=1}^{3} \theta_j x_j, \\ &\sum_{j=1}^{3} \theta_j = 1, \qquad 0 \leq \theta_j \leq 1,\ j = 1, ..., 3. \end{aligned} \tag{2.174}$$

Par unicité de la solution du système (2.174), on a

$$\theta_j = \gamma_j = p_j(x),\ j = 1, ..., 3. \tag{2.175}$$

Ainsi, d'après (2.42), (2.169) et le fait que $c_{2,2} = \frac{1}{8\pi}$ d'après (1.158), on a

$$\Phi\left(x, \widetilde{X}\right) = \frac{1}{8\pi} \left\{ \sum_{r,s=1}^{3} \theta_r \theta_s \phi\left(\|x_r - x_s\|_2 \right) - 2 \sum_{r=1}^{3} \theta_r \phi\left(\|x - x_r\|_2 \right) \right\}, \tag{2.176}$$

d'où

$$\begin{aligned}
\Phi\left(x,\widetilde{X}\right) &= \tfrac{1}{8\pi}\left\{\sum_{r,s=1}^{3}\theta_r\theta_s\phi\left(\|x_r-x_s\|_2\right)-2\sum_{r=1}^{3}\theta_r\phi\left(\|x-x_r\|_2\right)\right\}\\
&= \tfrac{1}{8\pi}\left\{\begin{array}{c}\left[\theta_1\theta_2\phi\left(\|x_2-x_1\|_2\right)+\theta_1\theta_2\phi\left(\|x_1-x_2\|_2\right)\right]\\ -2\left[\theta_1\phi\left(\|x-x_1\|_2\right)+\theta_2\phi\left(\|x-x_2\|_2\right)\right]\end{array}\right\}\\
&= \tfrac{1}{8\pi}\left[2\theta_1\theta_2\phi\left(h\right)-2\theta_1\phi\left(\left(1-\lambda\right)h\right)-2\theta_2\phi\left(\lambda h\right)\right]\\
&= \tfrac{1}{4\pi}\lambda\left(1-\lambda\right)h^2\left[\begin{array}{c}\ln h-(1-\lambda)\ln\left(1-\lambda\right)-(1-\lambda)\ln h\\ -\lambda\ln\lambda-\lambda\ln h\end{array}\right]\\
&= \tfrac{1}{4\pi}\lambda\left(1-\lambda\right)h^2\left[\left(\lambda-1\right)\ln\left(1-\lambda\right)-\lambda\ln\lambda\right].
\end{aligned} \tag{2.177}$$

L'étude des variations sur $[0,1]$ des fonctions

$$\begin{array}{ll}\varphi_1: & \lambda\longmapsto\lambda\left(1-\lambda\right),\\ \varphi_2: & \lambda\longmapsto\left(\lambda-1\right)\ln\left(1-\lambda\right)-\lambda\ln\lambda,\end{array} \tag{2.178}$$

montre que $A_1\left(\frac{1}{2},\frac{1}{4}\right)$ est le maximum de φ_1 et $A_2\left(\frac{1}{2},\ln 2\right)$ est le maximum de φ_2. Ainsi

$$\left.\begin{array}{ll}\lambda\left(1-\lambda\right) & \leq\frac{1}{4}\\ \left(\lambda-1\right)\ln\left(1-\lambda\right) & -\lambda\ln\lambda\quad\leq\ln 2\end{array}\right\},\ 0\leq\lambda\leq 1. \tag{2.179}$$

D'où

$$\Phi\left(x,\widetilde{X}\right)\leq\frac{\ln 2}{16\pi}h^2, \tag{2.180}$$

et

$$\left|f\left(x\right)-s_{f,X}\left(x\right)\right|\leq\left[\frac{\ln 2}{16\pi}\right]^{\frac{1}{2}}\left|f\right|_2 h. \tag{2.181}$$

■

Théorème 22 *(Powell) : Si $x\in\mathbb{R}^2$ est à l'interieur d'un triangle formé par des points donnés x_1,x_2,x_3, alors l'erreur entre une fonction $f:\mathbb{R}^2\longrightarrow\mathbb{R}$ et sa thin plate spline interpolant, $s_{f,X}$, est bornée par l'inégalité*

$$\left|f\left(x\right)-s_{f,X}\left(x\right)\right|\leq\left[\frac{\ln 3}{24\pi}\right]^{\frac{1}{2}}\left|f\right|_2 h, \tag{2.182}$$

où h est la longueur du plus grand côté du triangle et $|\cdot|_2$ désigne la semi-norme associée au produit semi-scalaire $\langle .,.\rangle$.

Démonstration. Puisque x est à l'interieur du triangle, on peut le représenter comme une combinaison linéaire convexe de x_1, x_2, x_3 *i.e*,

$$\begin{array}{l} x = \sum_{i=1}^{3} \theta_i x_i, \\ \sum_{i=1}^{3} \theta_i = 1, \qquad \theta_i \geq 0,\ i = 1, ..., 3. \end{array} \tag{2.183}$$

En utilisant le lemme 14, pour tout $\sigma > 0$, on a

$$\begin{array}{rcl} \Phi(\theta) & = & \Phi_\sigma(\theta) \\ & = & 2\theta_1\theta_2\phi_\sigma \|x_1 - x_2\|_2 + 2\theta_2\theta_3\phi_\sigma \|x_2 - x_3\|_2 \\ & & +2\theta_3\theta_1\phi_\sigma \|x_3 - x_1\|_2 - 2\sum_{i=1}^{3} \theta_i\phi_\sigma \|x - x_i\|_2. \end{array} \tag{2.184}$$

Alors, par le lemme 15, on peut minorer ϕ_σ par $-\frac{1}{2}(\sigma^2 e)^{-1}$ tel que

$$\begin{array}{rcl} \Phi_\sigma(\theta) & \leq & 2\theta_1\theta_2\phi_\sigma \|x_1 - x_2\|_2 + 2\theta_2\theta_3\phi_\sigma \|x_2 - x_3\|_2 \\ & & +2\theta_3\theta_1\phi_\sigma \|x_3 - x_1\|_2 + (\theta_1 + \theta_2 + \theta_3)(\sigma^2 e)^{-1} \\ & = & 2\theta_1\theta_2\phi_\sigma \|x_1 - x_2\|_2 + 2\theta_2\theta_3\phi_\sigma \|x_2 - x_3\|_2 \\ & & +2\theta_3\theta_1\phi_\sigma \|x_3 - x_1\|_2 + (\sigma^2 e)^{-1}, \end{array} \tag{2.185}$$

car $\sum_{i=1}^{3} \theta_i = 1$. En introduisant $\Delta_1 = \|x_2 - x_3\|_2^2$, $\Delta_2 = \|x_1 - x_3\|_2^2$ et $\Delta_3 = \|x_1 - x_2\|_2^2$, on a

$$\Phi_\sigma(\theta) \leq \theta_1\theta_2\Delta_3 \ln \sigma^2\Delta_3 + \theta_2\theta_3\Delta_1 \ln \sigma^2\Delta_1 + \theta_3\theta_1\Delta_2 \ln \sigma^2\Delta_2 + \left(\sigma^2 e\right)^{-1}. \tag{2.186}$$

En tant que fonction de σ^2, le second menbre de l'inégalité (2.186) a pour dérivée la fonction

$$\varphi\left(\sigma^2\right) = \frac{1}{\sigma^2}\left(\theta_1\theta_2\Delta_3 + \theta_2\theta_3\Delta_1 + \theta_3\theta_1\Delta_2 - \frac{1}{e\sigma^2}\right), \tag{2.187}$$

qui s'annulle pour $\sigma^2 = \{e(\theta_1\theta_2\Delta_3 + \theta_2\theta_3\Delta_1 + \theta_3\theta_1\Delta_2)\}^{-1}$. Il est donc clair que le second membre de l'inégalité (2.186) est minimum pour

$$\sigma^2 = \{e(\theta_1\theta_2\Delta_3 + \theta_2\theta_3\Delta_1 + \theta_3\theta_1\Delta_2)\}^{-1}. \tag{2.188}$$

On obtient ainsi

$$\begin{aligned}\Phi(\theta) \quad \leq \quad & \theta_1\theta_2\Delta_3 \ln \Delta_3 + \theta_2\theta_3\Delta_1 \ln \Delta_1 + \theta_3\theta_1\Delta_2 \ln \Delta_2 \\ & - (\theta_1\theta_2\Delta_3 + \theta_2\theta_3\Delta_1 + \theta_3\theta_1\Delta_2) \ln (\theta_1\theta_2\Delta_3 + \theta_2\theta_3\Delta_1 + \theta_3\theta_1\Delta_2).\end{aligned} \tag{2.189}$$

De plus, puisque $\ln \Delta_j \leq \ln h^2$, $j = 1, 2, 3$, on a

$$\Phi(\theta) \leq \Delta \ln h^2 - \Delta \ln \Delta, \tag{2.190}$$

où $\Delta = \theta_1\theta_2\Delta_3 + \theta_2\theta_3\Delta_1 + \theta_3\theta_1\Delta_2$. Or la fonction $\Delta \longmapsto \Delta \ln h^2 - \Delta \ln \Delta$ a pour dérivée la fonction $f(\Delta) = \ln h^2 - \ln \Delta - 1$ qui est strictement décroissante et qui s'annulle pour $\Delta = \frac{h^2}{e}$. Ainsi le second membre de (2.190) est une fonction croissante de Δ pour $0 \leq \Delta \leq \frac{h^2}{e}$. Par ailleurs on a $\Delta \leq (\theta_1\theta_2 + \theta_2\theta_3 + \theta_3\theta_1)\, h^2$. Comme $\theta_1+\theta_2+\theta_3 = 1$, on a $\theta_1\theta_2+\theta_2\theta_3+\theta_3\theta_1 = \theta_1 + \theta_2 - \theta_1\theta_2 - \theta_1^2 - \theta_2^2$. Soit

$$\begin{aligned} f \quad & : K \longrightarrow \mathbb{R} \\ & (x, y) \longmapsto x + y - xy - x^2 - y^2, \end{aligned} \tag{2.191}$$

où K est le compact de $\mathbb{R}^2$ défini par

$$K = \left\{(x, y) \in \mathbb{R}^2;\ 0 \leq x \leq 1,\ 0 \leq y \leq 1,\ x + y \leq 1\right\}, \tag{2.192}$$

et dont sa frontière est

$$\Gamma = Fr(K) = [AB] \cup [AC] \cup [BC], \tag{2.193}$$

avec $A(0, 0)$, $B(1, 0)$, $C(0, 1)$. On a

$$\begin{aligned} \frac{\partial f}{\partial x} &= 1 - y - 2x, \\ \frac{\partial f}{\partial y} &= 1 - x - 2y. \end{aligned} \tag{2.194}$$

(x, y) est un point critique de $f \Longleftrightarrow \dfrac{\partial f}{\partial x} = 0$ et $\dfrac{\partial f}{\partial y} = 0$, *i.e.*,

$$\left\{ \begin{array}{l} 2x + y = 1 \\ x + 2y = 1 \end{array} \right. , \text{ soit } (x, y) = \left(\frac{1}{3}, \frac{1}{3}\right). \tag{2.195}$$

De plus

$$\begin{aligned} r &= \frac{\partial^2 f}{\partial x^2} = -2, \\ s &= \frac{\partial^2 f}{\partial x \partial y} = -1, \\ t &= \frac{\partial^2 f}{\partial y^2} = -2. \end{aligned} \tag{2.196}$$

La matrice Héssienne de f en $\left(\frac{1}{3}, \frac{1}{3}\right)$ a pour déterminant $rt - s^2 = 3 > 0$. Comme $r < 0$, le point $\left(\frac{1}{3}, \frac{1}{3}\right) \in \overset{\circ}{K}$ est l'unique maximum relatif de f dont la valeur est $f\left(\frac{1}{3}, \frac{1}{3}\right) = \frac{1}{3}$. Par ailleurs f étant continue sur le compact K, admet un maximum en un point $(a, b) \in K$. On a

$$f(a, b) = \sup_{(x,y) \in K} f(x, y), \tag{2.197}$$

comme $\left(\frac{1}{3}, \frac{1}{3}\right) \in K$, on a $\frac{1}{3} \leq f(a, b)$. D'autres parts on vérifie aisément que $\forall (x, y) \in \Gamma$, $f(x, y) \leq \frac{1}{4} < \frac{1}{3}$. Il est donc clair que $(a, b) \notin \Gamma$. Comme $\complement_{\mathbb{R}^2}\Gamma = \left(\complement_{\mathbb{R}^2} K\right) \cup \overset{\circ}{K}$ et $(a, b) \in K$, on a $(a, b) \in \overset{\circ}{K}$. L'unicité du maximum relatif de f intérieur à K implique $(a, b) = \left(\frac{1}{3}, \frac{1}{3}\right)$. Ainsi $f\left(\frac{1}{3}, \frac{1}{3}\right) = \frac{1}{3}$ est le maximum absolu de f sur K. D'où

$$f(x, y) \leq \frac{1}{3}, \forall (x, y) \in K, \tag{2.198}$$

par suite on a

$$0 \leq \Delta \leq f(\theta_1, \theta_2) h^2 \leq \frac{1}{3} h^2 \leq \frac{h^2}{e}, \tag{2.199}$$

$$\Phi(\theta) \leq \Delta \ln h^2 - \Delta \ln \Delta \leq \frac{1}{3} h^2 \ln h^2 - \frac{1}{3} h^2 \ln\left(\frac{1}{3} h^2\right) = \frac{1}{3} h^2 \ln 3, \tag{2.200}$$

comme $c_{2,2} = \frac{1}{8\pi}$, on a

$$\Phi\left(x, \widetilde{X}\right) = c_{2,2} \Phi(\theta) \leq \frac{h^2}{24\pi} \ln 3. \tag{2.201}$$

D'où

$$|f(x) - s_{f,X}(x)| \leq \Phi\left(x, \widetilde{X}\right)^{\frac{1}{2}} |f|_2 \leq \left(\frac{\ln 3}{24\pi}\right)^{\frac{1}{2}} |f|_2 h. \tag{2.202}$$

■

Puisque $d = 1$ et $d = 3$ sont les seules autres possibilités vérifiant la condition $k > \frac{d}{2}$ lorsque $k = 2$, nous allons fournir les constantes correspondantes pour ces deux cas.

Dans le cas où $d = 1$ et $k = 2$, la fonction radiale est $\phi(r) = r^3$ et $l = \dim \pi_1(\mathbb{R}) = 2$. Nous sommes donc dans le cas de la cubique spline naturelle d'interpolation aux noeuds $x_1 < x_2 < ... < x_N$.

Proposition 23 *Soit $f \in V = BL^2(\mathbb{R})$ l'espace des distributions dont la dérivée d'ordre 2 est de carré intégrable dans $\mathbb{R}$ et $h = \max_{1 \leq i \leq N-1} (x_{i+1} - x_i)$. Alors l'erreur entre f et sa cubique spline interpolant $s_{f,X}$, est bornée par l'inégalité*

$$|f(x) - s_{f,X}(x)| \leq \frac{1}{4\sqrt{3}} |f|_2 h^{\frac{3}{2}}, \ \forall x \in [x_1, x_N]. \tag{2.203}$$

De plus, le facteur $\frac{1}{4\sqrt{3}}$ ne peut être remplacé par un nombre plus petit indépendant de N et f.

Démonstration. Fixons $x \in [x_i, x_{i+1}]$. x peut alors s'écrire $x = \theta_1 x_i + \theta_2 x_{i+1}$, où $\theta_1 = 1 - \alpha$, $\theta_2 = \alpha$, $\alpha \in [0,1]$ et $i \in \{1, 2, ..., N-1\}$. Prenons $\widetilde{X} = \{x_i, x_{i+1}\}$. Comme $c_{2,1} = \frac{1}{12}$ d'après (1.158), on a

$$\begin{aligned}
\Phi\left(x, \widetilde{X}\right) &= \tfrac{1}{12}\left\{ \sum_{r,s=1}^{3} \theta_r \theta_s \phi\left(\|x_r - x_s\|_2\right) - 2\sum_{r=1}^{3} \theta_r \phi\left(\|x - x_r\|_2\right) \right\} \\
&= \tfrac{1}{12}\left\{ \begin{array}{c} 2\theta_1\theta_2\phi\left(\|x_i - x_{i+1}\|_2\right) - 2\theta_1\phi\left(\|x - x_i\|_2\right) \\ -2\theta_2\phi\left(\|x - x_{i+1}\|_2\right) \end{array} \right\} \\
&= \tfrac{1}{12}\left\{ 2\alpha(1-\alpha)(x_{i+1} - x_i)^3 - 2\left[\begin{array}{c} (1-\alpha)(x - x_i)^3 \\ +\alpha(x_{i+1} - x)^3 \end{array} \right] \right\} \\
&= \frac{\alpha^2 (1-\alpha)^2}{3} (x_i - x_{i+1})^3 \leq \tfrac{1}{48}(x_{i+1} - x_i)^3 \leq \tfrac{1}{48} h^3.
\end{aligned} \tag{2.204}$$

D'où

$$|f(x) - s_{f,X}(x)| \leq \Phi\left(x, \widetilde{X}\right)^{\frac{1}{2}} |f|_2 \leq \frac{1}{4\sqrt{3}} |f|_2 h^{\frac{3}{2}}, \forall x \in [x_1, x_N]. \quad (2.205)$$

On travaille avec $N = l = 2$, *i.e.*, $X = \widetilde{X} = \{x_1, x_2\}$. On sait que si $(S_i)_{1 \leq i \leq 2}$ est la base de fonctions cardinales associées au problème d'interpolation (1.1), *i.e.*, $S_i(x_j) = \delta_{ij}$, $1 \leq i, j \leq 2$, alors $s_{f,X}(x) = \sum_{i=1}^{2} f(x_i) S_i(x)$. D'où $s_{f,X}(x) = 0$, $\forall f \in \Gamma = \{v \in V; v(x_i) = 0, 1 \leq i \leq 2\}$. Par ailleurs, notons que pour $\alpha = \frac{1}{2}$ et $x_{i+1} - x_i = x_2 - x_1 = h$, (2.204) est une égalité. Ainsi dans ce cas le second membre de l'inégalité (2.205) vaut

$$\Phi\left(x, \widetilde{X}\right)^{\frac{1}{2}} |f|_2, \quad (2.206)$$

pour tout f convenable. Or d'après la définition de ψ_x dépendant de (1.156), la définition de g_x donnée en (1.160), et le fait que $\phi(r) = r^3$ s'annule en 0, on a

$$\begin{aligned} g_x(x) &= c_{2,1} \left\{ \sum_{r,s=1}^{2} p_r(x) p_s(x) \phi(\|x_r - x_s\|_2) - 2 \sum_{r=1}^{2} p_r(x) \phi(\|x - x_r\|_2) \right\} \\ &= \Phi\left(x, \widetilde{X}\right). \end{aligned} \quad (2.207)$$

D'autre part on a vu que $g_x \in \Gamma_0 = \Gamma$ est telle que $\forall v \in \Gamma_0$, $v(x) = (v, g_x)$. En prenant $f = g_x$, on a $s_{g,X}(x) = 0$ et $f(x) = g_x(x) = (g_x, g_x) = |g_x|_2^2 = |f|_2^2$. D'où

$$|f(x) - s_{f,X}(x)| = |g_x(x) - s_{g_x,X}(x)| = |g_x|_2^2, \quad (2.208)$$

et

$$\Phi\left(x, \widetilde{X}\right)^{\frac{1}{2}} |f|_2 = \left(\Phi\left(x, \widetilde{X}\right) |f|_2^2\right)^{\frac{1}{2}} = |g_x|_2^2. \quad (2.209)$$

Les relations (2.208) et (2.209) montrent donc que (2.203) est une égalité pour $N = l = 2$, $\alpha = \frac{1}{2}$, $h = x_2 - x_1$ et $f = g_x$. Dans (2.203), le facteur $\frac{1}{4\sqrt{3}}$ ne peut donc être remplacé par un nombre $c \in \left]0, \frac{1}{4\sqrt{3}}\right[$ indépendant de N et f. ■

Dans le cas où $d = 3$ et $k = 2$, la fonction radiale est $\phi(r) = r$ et $l = \dim \pi_1(\mathbb{R}^3) = 4$. Nous sommes donc dans le cas de la fonction radiale linéaire en dimension 3.

Proposition 24 *soit $f \in V = BL^2(\mathbb{R}^3)$. On suppose que $x \in \mathbb{R}^3$ est à l'intérieur ou sur une face d'un tétraèdre dont les sommets x_1, x_2, x_3 et x_4 sont 4 des N points d'interpolation de X. Alors l'erreur entre f et sa spline naturelle linéaire trivariable interpolant $s_{f,X}$, en x est bornée par l'inégalité*

$$|f(x) - s_{f,X}(x)| \leq \frac{1}{4\sqrt{\pi}} |f|_2 h^{\frac{1}{2}}. \tag{2.210}$$

où h est la longueur du plus grand côté du tétraèdre. De plus, le facteur $\frac{1}{4\sqrt{\pi}}$ ne peut être remplacé par un nombre plus petit indépendant de N et f.

Démonstration. Prenons $\widetilde{X} = \{x_1, x_2, x_3, x_4\}$. Alors $x = \sum_{i=1}^{4} \theta_i x_i$ avec $0 \leq \theta_i \leq 1$, $i = 1, 2, 3, 4$ et $\sum_{i=1}^{4} \theta_i = 1$. Comme au théorème 21 par (2.175), on montre que $\theta_i = \gamma_i = p_i(x)$, $i = 1, 2, 3, 4$. Par ailleurs on peut d'après le lemme 18 par (2.76), écrire

$$\Phi\left(x, \widetilde{X}\right) = 2c_{k,d}\left\{T_1 + T_2 - \gamma_1 \phi\left(\|x - x_1\|_2\right)\right\},$$

avec $c_{k,d} = c_{2,3} = -\frac{1}{8\pi}$ d'après (1.158), et T_1,T_2 définis en (2.77) et (2.78) respectivement. $\Phi\left(x, \widetilde{X}\right)$ peut donc s'écrire sous la forme

$$\Phi\left(x, \widetilde{X}\right) = \frac{1}{4\pi}\left\{\sum_{i=1}^{4} \theta_i \|x - x_i\|_2 - \sum_{i=1}^{3}\sum_{j=i+1}^{4} \theta_i \theta_j \|x_i - x_j\|_2\right\}. \tag{2.211}$$

Comme dans la preuve du lemme 16 par (2.35), on adopte la notation $\hat{\theta}_i = \theta_i$, $\hat{x}_i = x_i$, $i = 1, 2, 3, 4$, $\hat{\theta}_5 = -1$, $\hat{x}_5 = x$ et on considère l'expression

$$\begin{aligned}\sum_{i=1}^{5}\sum_{j=1}^{5}\hat{\theta}_i\hat{\theta}_j\left\|\hat{x}_i-\hat{x}_j\right\|_2^2 &= \sum_{i=1}^{4}\sum_{j=1}^{4}\theta_i\theta_j\left\|x_i-x_j\right\|_2^2+2\sum_{i=1}^{4}\theta_i\hat{\theta}_5\left\|x_i-\hat{x}_5\right\|_2^2\\ &= \sum_{i=1}^{4}\sum_{j=1}^{4}\theta_i\theta_j\left\|x_i-x_j\right\|_2^2-2\sum_{i=1}^{4}\theta_i\left\|x-x_i\right\|_2^2\\ &= 2\left\{\sum_{i=1}^{3}\sum_{j=i+1}^{4}\theta_i\theta_j\left\|x_i-x_j\right\|_2^2-\sum_{i=1}^{4}\theta_i\left\|x-x_i\right\|_2^2\right\}.\end{aligned} \tag{2.212}$$

Or $\left\|\hat{x}_i-\hat{x}_j\right\|_2^2=\left\|\hat{x}_i\right\|_2^2-2\hat{x}_i\hat{x}_j+\left\|\hat{x}_j\right\|_2^2$ et $\sum_{i=1}^{5}\hat{\theta}_i\hat{x}_i=0$, $\sum_{i=1}^{5}\hat{\theta}_i=0$, donc $\sum_{i=1}^{5}\sum_{j=1}^{5}\hat{\theta}_i\hat{\theta}_j\left\|\hat{x}_i-\hat{x}_j\right\|_2^2=\sum_{i=1}^{5}\hat{\theta}_i\left\|\hat{x}_i\right\|_2^2\left(\sum_{j=1}^{5}\hat{\theta}_j\right)-2\left(\sum_{i=1}^{5}\hat{\theta}_i\hat{x}_i\right)\left(\sum_{i=1}^{5}\hat{\theta}_j\hat{x}_j\right)$
$+\sum_{j=1}^{5}\hat{\theta}_j\left\|\hat{x}_j\right\|_2^2\left(\sum_{j=1}^{5}\hat{\theta}_i\right)=0$. D'où

$$\sum_{i=1}^{3}\sum_{j=i+1}^{4}\theta_i\theta_j\left\|x_i-x_j\right\|_2^2-\sum_{i=1}^{4}\theta_i\left\|x-x_i\right\|_2^2=0. \tag{2.213}$$

En multipliant (2.213) par $(4\pi h)^{-1}$ et en additionnant le produit à (2.211), on en déduit

$$\begin{aligned}\Phi\left(x,\widetilde{X}\right) &= \frac{\sigma}{4\pi}\left\{\sum_{i=1}^{4}\theta_i\frac{\left\|x-x_i\right\|_2}{h}\left(1-\frac{\left\|x-x_i\right\|_2}{h}\right)\right.\\ &\quad\left.-\sum_{i=1}^{3}\sum_{j=i+1}^{4}\theta_i\theta_j\frac{\left\|x_i-x_j\right\|_2}{h}\left(1-\frac{\left\|x_i-x_j\right\|_2}{h}\right)\right\}\\ &\leq \frac{\sigma}{4\pi}\sum_{i=1}^{4}\theta_i\frac{\left\|x-x_i\right\|_2}{h}\left(1-\frac{\left\|x-x_i\right\|_2}{h}\right) \qquad (2.214)\\ &\leq \frac{h}{16\pi}\sum_{i=1}^{4}\theta_i\\ &= \frac{h}{16\pi},\end{aligned}$$

car $\frac{\left\|x-x_i\right\|_2}{h}$, $i=1,2,3,4$ et $\frac{\left\|x_i-x_j\right\|_2}{h}$, $i,j=1,2,3,4$, $i<j$, appartiennent tous à l'intervalle $[0,1]$, et $\forall t\in[0,1]$, on a $0\leq t\left(1-t\right)\leq\frac{1}{4}$. D'où

$$\left|f\left(x\right)-s_{f,X}\left(x\right)\right|\leq\Phi\left(x,\widetilde{X}\right)^{\frac{1}{2}}\left|f\right|_2\leq\frac{1}{4\sqrt{\pi}}\left|f\right|_2h^{\frac{1}{2}}. \tag{2.215}$$

Nous utilisons les mêmes arguments que dans la preuve de la dernière partie de la proposition 23 pour prouver la dernière partie de la proposition 24 en ramarquant d'abord que (2.214) est une égalité lorsque x est le milieu du plus grand côté du tétraèdre. Pour ce choix donc de x, on pose $f = g_x$ avec g_x défini au (1.160). L'interpolant correspondant à f sur $X = \widetilde{X}$ est donc $s_{g,X} \equiv 0$, comme précédemment. L'inégalité (2.210) est donc satisfaite comme une égalité dans ce cas. Ainsi, dans (2.210), le facteur $\frac{1}{4\sqrt{\pi}}$ ne peut être remplacé par un nombre $c \in \left]0, \frac{1}{4\sqrt{\pi}}\right[$ indépendant de N et f. ■

Se reférant à l'article [2] pour la preuve de la proposition, nous énnonçons sans preuve la proposition suivante :

Proposition 25 *Soit $f \in V = BL^2(\mathbb{R}^3)$. On suppose que $x \in \mathbb{R}^3$ est à l'intérieur ou sur une face d'un cube $h \times h \times h$ dont les sommets sont 8 des N points d'interpolation de X. Alors l'erreur entre f et sa spline naturelle linéaire trivariable interpolant $s_{f,X}$, en x est bornée par l'inégalité*

$$|f(x) - s_{f,X}(x)| \leq \left(\frac{7\sqrt{3} - 3\sqrt{2} - 3}{64\pi}\right)^{\frac{1}{2}} |f|_2 \, h^{\frac{1}{2}}. \tag{2.216}$$

De plus, le facteur $\left(\frac{7\sqrt{3}-3\sqrt{2}-3}{64\pi}\right)^{\frac{1}{2}}$ ne peut être remplacé par un nombre plus petit indépendant de N et f.

Chapitre 3

Le problème de la quasi-interpolation

3.1 Application pour la résolution d'une équation différentielle du second ordre raide

Soit Ω un ouvert borné de $\mathbb{R}^d$ ayant la propriété de cône et de frontière $\partial\Omega$ lipschitzienne. Dans le chapitre I nous avons montré que le problème d'interpolation d'une fonction $f \in \mathcal{C}(\Omega)$ sur un ensemble π_{k-1}-unisolvant $X = \{x_1,...,x_N\}$ par un interpolant $s_{f,X}$ de la forme

$$s_{f,X}(x) = \sum_{j=1}^{N} \lambda_j \phi\left(\|x - x_j\|_2\right) + p(x), \text{ où } p \in \pi_{k-1}, \tag{3.1}$$

peut se ramener à

$$s_{f,X}(x) = \sum_{j=1}^{N} \mu_j K(x, x_j) \text{ où } K(x, x_j) = q_j(x) \ \forall x \in \Omega, \tag{3.2}$$

les $(q_j)_{1\leq j\leq N}$ étant ici les représentants sur V (espace fonctionnel où l'on étudie le problème d'interpolation) des points d'évalution en x_j ; *i.e.*,

$$v(x_j) = (v, q_j), \forall v \in V. \tag{3.3}$$

Cette méthode demande la résolution d'un système d'équations pour les coefficients inconnnus μ_j. L'idée de la quasi-interpolation est de choisir les fonctions de base ϕ_j telles que

$$f_j = s_{f,X}(x_j) = \sum_{k=1}^{N} \lambda_k \phi_k(x_j) = \lambda_j,\ 1 \leq j \leq N, \tag{3.4}$$

c'est à dire

$$f(x_j) = \lambda_j,\ 1 \leq j \leq N. \tag{3.5}$$

Dans cette sous-section l'ensemble $X = \{x_1,...,x_N\}$ comprendra n points intérieurs de Ω et m points du bord $\partial\Omega$; *i.e.*, $N = n + m$, et on posera $s_{f,X} = f^*$. On considère donc l'équation aux dérivées partielles suivante

$$\begin{array}{l} Lu(x) = f(x),\ x \in \Omega \subset \mathbb{R}^d, \\ Qu(x) = g(x),\ x \in \partial\Omega. \end{array} \tag{3.6}$$

Soit

$$(x_j)_{1 \leq j \leq n} \subset \Omega \text{ et } (x_j)_{n+1 \leq j \leq n+m} \subset \partial\Omega. \tag{3.7}$$

3.2 Résolution utilisant la collocation RBF

Par la méthode de collocation RBF on cherche une solution approchée u^*, du problème (3.6), sous la forme

$$u^*(x) = \sum_{j=1}^{n+m} \lambda_j \phi_j(x), \tag{3.8}$$

où $\phi_j(x) = \phi\left(\|x - x_j\|_2\right)$, ϕ étant une fonction radiale. L et Q étant linéaires, en remplaçant u^* dans (3.6), on a

$$\begin{aligned} Lu^*(x_i) &= \sum_{j=1}^{n+m} \lambda_j L\phi_j(x_i) = f(x_i),\ 1 \leq i \leq n, \\ Qu^*(x_i) &= \sum_{j=1}^{n+m} \lambda_j Q\phi_j(x_i) = g(x_i),\ n+1 \leq i \leq n+m. \end{aligned}$$

On résout donc le système d'inconnus $(\lambda_j)_{1 \leq j \leq n+m}$,

$$\begin{aligned} \sum_{j=1}^{n+m} \lambda_j L\phi_j(x_i) &= f(x_i),\ 1 \leq i \leq n, \\ \sum_{j=1}^{n+m} \lambda_j Q\phi_j(x_i) &= g(x_i),\ n+1 \leq i \leq n+m. \end{aligned} \tag{3.9}$$

3.3 Résolution utilisant la quasi-interpolation RBF

Par la méthode utilisant la quasi-interpolation, l'idée est d'interpoler le terme forcé f de (3.6) en utilisant les splines des fonctions radiales de base $\phi\left(\|. - x_j\|_2\right)$. Une bonne approximation u^* de la solution u peut alors être obtenue par résolution de l'équation fondamentale correspondante et un petit système d'équations correspondant à la condition initiale ou à la condition limite. On arrive à interpoler le terme forcé f de (3.6) en appliquant les résultats précédents. On doit donc trouver f^* telle que

$$f^*(x) = \sum_{j=1}^{n+m} \lambda_j \phi\left(\|x - x_j\|_2\right), \tag{3.10}$$

et

$$f^*(x_i) = \sum_{j=1}^{n+m} \lambda_j \phi\left(\|x_i - x_j\|_2\right) = f(x_i),\ i = 1, ..., n. \tag{3.11}$$

Le système (3.11) s'écrit,

$$\sum_{j=1}^{n+m} \lambda_j L\Phi(x_i - x_j) = f(x_i),\ 1 \leq i \leq n, \tag{3.12}$$

si l'on a résolu l'équation fondamentale

$$L\Phi(x) = \phi(\|x\|_2). \tag{3.13}$$

A la relation (3.12) on ajoute la condition de bord

$$\sum_{j=1}^{n+m} \lambda_j Q\Phi(x_i - x_j) = g(x_i),\ i = n+1, ..., n+m, \tag{3.14}$$

correspondant à la deuxième équation de (4.5). La solution approchée u^*, du problème (3.6) est alors obtenue via une solution fondamentale Φ de (3.13), et en résolvant le système

$$\begin{aligned} \sum_{j=1}^{n+m} \lambda_j \phi\left(\|x_i - x_j\|_2\right) &= f(x_i),\ i = 1, ..., n, \\ \sum_{j=1}^{n+m} \lambda_j Q\Phi(x_i - x_j) &= g(x_i),\ i = n+1, ..., n+m. \end{aligned} \tag{3.15}$$

On prend alors

$$u^*(x) = \sum_{j=1}^{n+m} \lambda_j w_j(x), \tag{3.16}$$

où $(w_j)_{1\leq j\leq n+m}$ est telle que

$$\begin{aligned} Lw_j(x) &= L\Phi(x-x_j) = \phi\left(\|x-x_j\|_2\right) = \phi_j(x), \forall x \in \Omega, \\ Qw_j(x) &= Q\Phi(x-x_j), \forall x \in \partial\Omega. \end{aligned} \tag{3.17}$$

Pour la résolution du problème raide on se propose donc d'appliquer la technique de quasi-interpolation suivante. On suppose que f^* s'écrit

$$f^*(x) = \sum_{j=1}^{n+m} f(x_j)\,\phi\left(\|x-x_j\|_2\right), \tag{3.18}$$

i.e., dans (3.10)

$$\lambda_j = f(x_j),\ 1 \leq j \leq n+m. \tag{3.19}$$

Notons qu'un tel choix est possible lorsqu'on prend par exemple

$$\phi_i(x_j) = \delta_{ij},\ 1 \leq i,j \leq n+m. \tag{3.20}$$

Nous avons vu aux chapitres I et II que f^* possède un ordre d'approximation $\mathcal{O}(h^\alpha)$ de f, *i.e.*, il existe une constante C telle que

$$\|f^* - f\|_\infty \leq Ch^\alpha, \tag{3.21}$$

où h désigne la densité des points donnés par

$$h = \max_{x\in\Omega} \min_{x_j} \|x - x_j\|. \tag{3.22}$$

(3.17), (3.18) et (3.21) impliquent

$$\left\| L\sum_{j=1}^{n+m} f(x_j)\, w_j - f \right\|_\infty \leq Ch^\alpha. \tag{3.23}$$

Pour trouver une approximation u^* de la solution u de (3.6), on reécrit f^* sous la forme générale (3.1) ; *i.e.*,

$$f^* = \sum_{j=1}^{n+m} \beta_j\phi_j + p, \text{où } p \in \pi_{k-1}, \tag{3.24}$$

les $\left(\phi_j\right)_{1\leq j\leq n+m}$ étant choisies de manière à vérifier (3.20). Sachant que (3.17) est vérifiée, (3.24) s'écrit

$$f^* = L\left[\sum_{j=1}^{n+m}\beta_j w_j\right] + p.$$

On prend alors

$$u^* = \sum_{j=1}^{n+m}\beta_j w_j + v, \tag{3.25}$$

où $v \in \pi_{k-1}$, est telle que

$$Lv = p. \tag{3.26}$$

Il est donc clair que

$$Lu^* = f^*. \tag{3.27}$$

Par ailleurs (3.20), (3.18), (3.24) et (3.26) impliquent

$$f\left(x_i\right) = f^*\left(x_i\right) = \beta_i + Lv\left(x_i\right),\ 1 \leq i \leq n+m. \tag{3.28}$$

On a donc

$$\beta_i = f\left(x_i\right) - Lv\left(x_i\right),\ 1 \leq i \leq n+m. \tag{3.29}$$

D'après (3.25) u^* s'écrit donc

$$u^* = \sum_{j=1}^{n+m}\left[f\left(x_j\right) - Lv\left(x_j\right)\right]w_j + v. \tag{3.30}$$

L'élément v sera déterminé de telle sorte que u^* vérifie les conditions aux limites de (3.6) ; *i.e.*,

$$Qu^*\left(x_i\right) = g\left(x_i\right),\ n+1 \leq i \leq n+m. \tag{3.31}$$

Pour simplifier, on suppose que $\Omega =]a, b[$ et les points donnés sont

$$a = x_0 < x_1 < ... < x_n = b. \tag{3.32}$$

La densité h des points donnés est alors

$$h = \max_{1\leq j\leq n}\left(x_j - x_{j-1}\right). \tag{3.33}$$

Pour $f \in \mathcal{C}^2[x_0, x_n]$ et les données $\{x_j, f_j\}_{j=0}^n$, le quasi-interpolant donné par Wu et Schaback ([49]) est

$$\mathcal{L}_d f(x) = f_0\alpha_0(x) + f_1\alpha_1(x) + \sum_{j=2}^{n-2} f_j\psi_j(x) + f_{n-1}\alpha_{n-1}(x) + f_n\alpha_n(x), \quad (3.34)$$

où

$$\begin{aligned}
\alpha_0(x) &= \frac{1}{2} + \frac{\phi_1(x) - (x - x_0)}{2(x_1 - x_0)}, \quad (3.35)\\
\alpha_1(x) &= \frac{\phi_2(x) - \phi_1(x)}{2(x_2 - x_1)} - \frac{\phi_1(x) - (x - x_0)}{2(x_1 - x_0)},\\
\psi_j(x) &= \frac{\phi_{j+1}(x) - \phi_j(x)}{2(x_{j+1} - x_j)} - \frac{\phi_j(x) - \phi_{j-1}(x)}{2(x_j - x_{j-1})},\ j = 2, ..., n-2,\\
\alpha_{n-1}(x) &= \frac{(x_n - x) - \phi_{n-1}(x)}{2(x_n - x_{n-1})} - \frac{\phi_{n-1}(x) - \phi_{n-2}(x)}{2(x_{n-1} - x_{n-2})},\\
\alpha_n(x) &= \frac{1}{2} + \frac{\phi_{n-1}(x) - (x_n - x)}{2(x_n - x_{n-1})}.
\end{aligned}$$

Si l'on choisi par exemple pour fonction radiale les splines linéaires, *i.e.* $\phi(r) = r$, on aura $\phi_j(x) = \phi\left(\|x - x_j\|_2\right) = |x - x_j|$, et on vérifie facilement que les fonctions α_0, α_1, $\left(\psi_j\right)_{2\le j\le n-2}$, α_{n-1}, α_n sont telles que pour $i = 0, 1, 2, ..., n-2, n-1, n$, on a :

$$\begin{aligned}
\alpha_0(x_i) &= \delta_{i0}, \quad (3.36)\\
\alpha_1(x_i) &= \delta_{i1},\\
\psi_j(x_i) &= \delta_{ij},\\
\alpha_{n-1}(x_i) &= \delta_{in-1},\\
\alpha_n(x_i) &= \delta_{in}.
\end{aligned}$$

Pour le choix $\phi(r) = r$, si l'on définit les fonctions $(\Psi_j)_{0\le j\le n}$ par

$$\begin{aligned}
\Psi_0 &= \alpha_0, \quad (3.37)\\
\Psi_1 &= \alpha_1,\\
\Psi_j &= \psi_j,\ 2 \le j \le n-2,\\
\Psi_{n-1} &= \alpha_{n-1},\\
\Psi_n &= \alpha_n,
\end{aligned}$$

on peut vérifier que

$$\Psi_j(x) = \begin{cases} 0, & x \leq x_{j-1} \\ \dfrac{x - x_{j-1}}{x_j - x_{j-1}}, & x_{j-1} \leq x \leq x_j \\ \dfrac{x_{j+1} - x}{x_{j+1} - x_j}, & x_j \leq x \leq x_{j+1} \\ 0 & x_{j+1} \leq x, \end{cases} \tag{3.38}$$

et

$$\Psi_j(x_i) = \delta_{ij},\ i, j = 0, ..., n. \tag{3.39}$$

$$\mathcal{L}_d f(x) = \sum_{j=0}^{n} f_j \Psi_j(x), \tag{3.40}$$

est donc un exemple simple de quasi-interpolation. Comme dans ([22]), nous allons appliquer la méthode au problème raide suivant :

$$\varepsilon u''(x) + u'(x) = f(x),\ x \in (0,1), \tag{3.41}$$

$$u(0) = a,\ u(1) = b, \tag{3.42}$$

où $\varepsilon > 0$ est un petit paramètre constant. Il est bien connu que quand le paramètre ε est très petit et $f(x) \equiv 0$, la solution forme une couche limite au voisinage de 0, et pour cela, on choisi proportionnellement plus de points auprès de la couche limite en prenant un total de $(n+5)$ points donnés

$$y_j = sign(j) \left(\frac{j}{n}\right)^p,\ j = -2, -1, ..., (n+2),\ p \in \mathcal{N}^*. \tag{3.43}$$

$$sign(j) = \begin{cases} -1 & \text{si } j < 0, \\ 0 & \text{si } j = 0, \\ 1 & \text{si } j > 0. \end{cases} \tag{3.44}$$

Notons que la méthode permet une sensibilité de choisir des points hors du domaine pourvu que les fonctions de base soient bien définies. Nous supposons en premier lieu une quasi- interpolation du terme forcé de (3.41) en utilisant les fonctions linéaires de base (elles seront changées plus tard en multiqua-

driques si bien que les dérivées des fonctions peuvent aussi être approchées.)

$$f^{*}(x)=\sum_{j=-2}^{n+2} f(y_j)\, B_j^1(x) \cong f(x), \tag{3.45}$$

où B_j^1 est une fonction linéaire de base B-spline donnée par

$$B_j^1(x)=\frac{|x-y_{j+1}|-|x-y_j|}{2(y_{j+1}-y_j)}-\frac{|x-y_j|-|x-y_{j-1}|}{2(y_j-y_{j-1})}. \tag{3.46}$$

On a

$$\begin{aligned} y_{-2} &= \frac{(-1)^{p+1}2^p}{n^p}, \\ y_{-1} &= \frac{(-1)^{p+1}}{n^p}, \\ y_0 &= 0, \\ y_1 &= \frac{1}{n^p}, \\ y_2 &= \frac{2^p}{n^p}, \\ & \\ y_n &= 1, \\ y_{n+1} &= \frac{(n+1)^p}{n^p}, \\ y_{n+2} &= \frac{(n+2)^p}{n^p}. \end{aligned} \tag{3.47}$$

Si p est pair on a :

$$\begin{aligned} y_{-2} &< y_{-1}<y_0<y_1<y_2<...<y_n<y_{n+1}<y_{n+2}, \\ B_{-2}^1(x) &= \frac{1}{2}+\frac{|x-y_{-1}|-|x-y_{-2}|}{2(y_{-1}-y_{-2})}, \\ B_{n+2}^1(x) &= \frac{1}{2}+\frac{|x-y_{n+1}|-|x-y_{n+2}|}{2(y_{n+2}-y_{n+1})}. \end{aligned} \tag{3.48}$$

Si p est impair on a :

$$\begin{aligned} y_0 &< y_{-1} = y_1 < y_{-2} = y_2 < ... < y_n < y_{n+1} < y_{n+2}, \quad (3.49) \\ B^1_{-1}(x) &= B^1_1(x), \\ B^1_{-2}(x) &= B^1_2(x), \\ B^1_{n+2}(x) &= \frac{1}{2} - \frac{|x-y_{n+2}| - |x-y_{n+1}|}{2(y_{n+2}-y_{n+1})}. \end{aligned}$$

$B^1_j(x)$ étant une combinaison linéaire de $\phi_j(x) = \phi\left(\|x-x_j\|_2\right) = |x-x_j|$ ($\phi(r) = r$), avant de résoudre les équations $LW_j(x) = B^1_j(x)$, on résout d'abord l'équation $L\Phi(x) = \phi(\|x\|_2) = |x|$. Notons que si $\Phi(x) = x^2/2 - \varepsilon x$, alors $\varepsilon\Phi''(x) + \Phi'(x) = x$. Ainsi, si $\Phi(x) = (x/2-\varepsilon)|x|$, alors $L\Phi(x) = |x|$, sauf au point $x = 0$. Puisque la fonction $\Phi(x)$ n'est pas $\mathcal{C}^2(\Omega)$, on remplace l'approximation en utilisant les multiquadriques comme suit : Soit

$$\Phi(x) = (x/2-\varepsilon)\sqrt{c^2+x^2}, \quad (3.50)$$

alors $\varepsilon\Phi''(x) + \Phi'(x) \cong |x|$ (l'erreur exacte sera donnée plutart). En outre, si on définit

$$W_j(x) = \frac{\Phi(x-y_{j+1}) - \Phi(x-y_j)}{2(y_{j+1}-y_j)} - \frac{\Phi(x-y_j) - \Phi(x-y_{j-1})}{2(y_j-y_{j-1})}, \quad (3.51)$$

alors $\varepsilon W''_j(x) + W'_j(x) \cong B^1_j(x)$. Le schéma donné de quasi-interpolation pour résoudre l'équation (3.41) est

$$u_q(x) = \sum_{j=-2}^{n+2} f(y_j) W_j(x), \quad (3.52)$$

qui satisfait clairement $Lu_q(x) = \sum_{j=-2}^{n+2} f(y_j) LW_j(x) \cong \sum_{j=-2}^{n+2} f(y_j) B^1_j(x) \cong f(x)$. En outre, pour une certaine fonction $v(x) \in \mathcal{C}^2(\Omega)$, si $v_q(x) = v(x) - \sum_{j=-2}^{n+2} (\varepsilon v''(y_j) + v'(y_j)) W_j(x)$, alors $\varepsilon v''_q(x) + v'_q(x) \cong 0$. De (3.41), la quasi-interpolation pour la solution devient

$$u^*(x) = v(x) + \sum_{j=-2}^{n+2} \left[f(y_j) - (\varepsilon v''(y_j) + v'(y_j))\right] W_j(x). \quad (3.53)$$

Pour satisfaire la condition limite (3.42), on pose $v(x) = v_0 + v_1 x$. Les deux paramètres libres v_0 et v_1 peuvent alors être trouvés en résolvant les équations simples suivantes :

$$u(0) = a = u^*(0) = v_0 + \sum_{j=-2}^{n+2} [f(y_j) - v_1] W_j(0), \tag{3.54}$$

$$u(1) = b = u^*(1) = v_0 + v_1 + \sum_{j=-2}^{n+2} [f(y_j) - v_1] W_j(1). \tag{3.55}$$

Dans le cas où $a = 0$ et $b = 1$, v_0 et v_1 peuvent être obtenus facilement par :

$$v_1 = \frac{1 + \sum_{j=-2}^{N+2} f(y_j) (W_j(0) - W_j(1))}{1 + \sum_{j=-2}^{N+2} (W_j(0) - W_j(1))}, \tag{3.56}$$

$$v_0 = \sum_{j=-2}^{N+2} [v_1 - f(y_j)] W_j(0). \tag{3.57}$$

La solution numérique de l'équation peut alors être obtenue en utilisant (3.53). Le schéma numérique correspondant à l'équation homogène ($f(x) \equiv 0$)

$$\begin{aligned} \varepsilon u''(x) + u'(x) &= 0,\ x \in]0,1[, \\ u(0) &= 0,\ u(1) = 1, \end{aligned} \tag{3.58}$$

dont la solution exacte est donnée par

$$u(x) = \frac{1 - e^{-x/\varepsilon}}{1 - e^{-1/\varepsilon}}, \tag{3.59}$$

donne une très bonne approximation de la solution comme cela sera montré dans les applications numériques.

3.4 Les majorations de l'erreur

En dérivant deux fois la fonction $\Phi(x) = (x/2 - \varepsilon)\sqrt{c^2 + x^2,}$ on trouve

$$\begin{aligned}
\Phi'(x) &= \frac{x^4 - \varepsilon x^3 + \frac{3}{2}c^2x^2 - \varepsilon c^2 x + \frac{1}{2}c^4}{\left(c^2 + x^2\right)^{3/2}}, \qquad (3.60)\\
\Phi''(x) &= \frac{x^3 + \frac{3}{2}c^2x^2 - \varepsilon c^2}{\left(c^2 + x^2\right)^{3/2}},\\
\varepsilon\Phi''(x) + \Phi'(x) &= \frac{x^4 + \frac{3}{2}c^2x^2 + \frac{1}{2}\varepsilon c^2 x - \varepsilon^2 c^2 + \frac{1}{2}c^4}{\left(c^2 + x^2\right)^{3/2}}.
\end{aligned}$$

On a donc $||x| - \varepsilon\Phi''(x) - \Phi'(x)| \leq \left||x| - \dfrac{x^4}{\left(c^2 + x^2\right)^{3/2}}\right|$

$+ \left|\dfrac{\frac{3}{2}c^2x^2 + \frac{1}{2}\varepsilon c^2 x - \varepsilon^2 c^2 + \frac{1}{2}c^4}{\left(c^2 + x^2\right)^{3/2}}\right| = T_1 + T_2$, où

$$T_1 = |x|\left|1 - \frac{|x|^3}{\left(c^2 + x^2\right)^{3/2}}\right| = |x|\left(1 - \frac{|x|^3}{\left(c^2 + x^2\right)^{3/2}}\right), \qquad (3.61)$$

et

$$T_2 = \left|\frac{\frac{3}{2}c^2x^2 + \frac{1}{2}\varepsilon c^2 x - \varepsilon^2 c^2 + \frac{1}{2}c^4}{\left(c^2 + x^2\right)^{3/2}}\right|. \qquad (3.62)$$

Pour tout $x \neq 0$, on a $T_1 = |x|\left(1 - \dfrac{|x|^3}{\left(c^2 + x^2\right)^{3/2}}\right) = |x|\left(1 - \dfrac{1}{\left(1 + \dfrac{c^2}{x^2}\right)^{3/2}}\right)$.

On pose $\alpha = \arctan\dfrac{c}{|x|}$, donc $0 < \alpha < \dfrac{\pi}{2}$ et $\tan\alpha = \dfrac{c}{|x|}$, par suite,

$$T_1 = c\frac{1 - \cos^3\alpha}{\tan\alpha}. \qquad (3.63)$$

On applique donc le théorème des accroissement finis aux fonctions $g(\alpha) = 1 - \cos^3\alpha$ et $l(\alpha) = \tan\alpha$ sur $[0, \alpha]$. Il existe $\theta_1, \theta_2 \in]0, \alpha[$ tels que $g(\alpha) - g(0) = \alpha g'(\theta_1)$ et $l(\alpha) - l(0) = \alpha l'(\theta_2)$. On trouve après calculs,

$$T_1 = c\left(3\sin\theta_1\cos^2\theta_1\cos^2\theta_2\right) = \frac{3c}{2}\left(\sin 2\theta_1\cos\theta_1\cos^2\theta_2\right) \leq \frac{3}{2}c. \qquad (3.64)$$

On a

$$T_2 \leq \frac{\frac{3}{2}c^2x^2 + \frac{1}{2}\varepsilon c^2|x| + \varepsilon^2c^2 + \frac{1}{2}c^4}{(c^2+x^2)^{3/2}} \leq \frac{\frac{3}{2}c^2(x^2+c^2) + \frac{1}{2}\varepsilon c(c^2+x^2) + \varepsilon^2(c^2+x^2)}{(c^2+x^2)^{3/2}} + \frac{\frac{1}{2}c^4}{(c^2+x^2)^{3/2}}.$$

Après simplification on obtient donc

$$T_2 \leq \frac{\frac{3}{2}c^2 + \frac{1}{2}\varepsilon c + \varepsilon^2}{(c^2+x^2)^{1/2}} + \frac{1}{2}c \leq \frac{7}{2}c, \text{ si } \varepsilon \leq c. \tag{3.65}$$

Ainsi

$$||x| - \varepsilon\Phi''(x) - \Phi'(x)| \leq T_1 + T_2 \leq 5c. \tag{3.66}$$

Notons $\overline{h} = \max\limits_{-2\leq j\leq(n+1)}(y_{j+1} - y_j)$, et $\underline{h} = \min\limits_{-2\leq j\leq(n+1)}(y_{j+1} - y_j)$. Puisque $T_1 + T_2 \leq 5c$, nous avons

$$\left|\varepsilon W_j''(x) + W_j'(x) - B_j^1(x)\right| \leq \frac{10c}{\underline{h}}. \tag{3.67}$$

Si $u(x)$ est la solution exacte de l'équation et $u^*(x)$ est la solution numérique utilisant la quasi-interpolation, nous avons alors $\left|\varepsilon(u^*-u)'' + (u^*-u)'\right| \leq \|f\|\frac{9c}{\underline{h}^2} + \|f''\|\overline{h}^2$, où $\|f''\|\overline{h}^2$ est l'erreur sur les fonctions de base linéaires d'interpolation et $\|f\|\frac{9c}{\underline{h}^2}$ est l'erreur causée par la quasi -interpolation aux fonctions de base linéaires d'interpolation. Dans les précédents calculs, si nous utilisons la solution exacte de l'équation fondamentale, l'erreur sera en outre améliorée et sera indépendante du paramètre de raideur ε. De plus, si nous définissons

$$F(x) = \int_0^x \left(1 - e^{\frac{t-x}{\varepsilon}}\right) f(t)\, dt, \tag{3.68}$$

alors la solution exacte de l'équation différentielle avec condition limite homogène peut être écrite comme

$$u_0(x) = F(x) - F(1)\frac{1 - e^{-x/\varepsilon}}{1 - e^{-1/\varepsilon}}. \tag{3.69}$$

Il est alors aisé d'estimer cela par

$$\|F(x)\|_\infty \leq \int_0^x |f(t)|\, dt \leq \int_0^1 |f(t)|\, dt \leq \|f\|_\infty . \tag{3.70}$$

Ainsi, nous avons $\|u_0(x)\|_\infty \leq 2\|F\|_\infty \leq 2\|f\|_\infty$. Puisque les deux solutions exactes et la solution numérique satisfont les mêmes conditions aux limites, la fonction $u_0 = u^* - u$ satisfait la condition limite nulle, et donc nous obtenons l'estimation d'erreur suivante

$$\|u^* - u\|_\infty \leq 2\left(\|f\| \frac{9c}{\underline{h}^2} + \|f''\|_\infty \overline{h}^2\right). \tag{3.71}$$

si $\varepsilon \leq c$. Puisque l'amplitude de ε dans le précédent problème raide est normalement très petit ($\varepsilon << \underline{h}$), l'estimation d'erreur précédente indique que nous pouvons même obtenir une solution plus exacte en dimunuant la valeur de c. C'est très différent selon plusieurs méthodes numériques d'existence pour résoudre les problèmes raides qui d'habitude donnent une solution moins exacte quand ε est plus petit. par exemple , on peut choisir c tel que

$$\varepsilon \leq c \leq \left(\underline{h}\,\overline{h}\right)^2, \tag{3.72}$$

afin que l'estimation d'erreur devienne

$$\|u^* - u\|_\infty \leq k\overline{h}^2. \tag{3.73}$$

Dans les calculs précédents, nous prenons simplement $\varepsilon = c$ pour une vérification des erreurs.

3.5 Résolution du problème modèle du champ classique d'un méson.

Avant d'aborder cette section nous allons d'abord rappeler quelques notions éssentielles sur les opérateurs linéaires non-bornés et sur les semi-groupes ainsi que les résultas des théorèmes qui constituent l'outil éssentiel de l'étude du problème modèle du champ classique d'un méson.

3.5.1 Opérateurs linéaires non-bornés

Dans les paragraphes de 3.5.1 à 3.5.5 nous utilisons les définitions et théorèmes qu'on retrouve dans [4] et [50].

Définition 26 *Soient E et F deux espaces de Banach. On appelle opérateur linéaire non borné de E dans F toute application linéaire $A : D(A) \subset E \longrightarrow F$ définie sur un sous-epace vectoriel $D(A) \subset E$, à valeur dans F. $D(A)$ est le domaine de A.*

On dit que A est borné s'il existe une constante $c \geq 0$ telle que

$$\|Au\| \leq c\|u\| \ \ \forall u \in D(A). \tag{3.74}$$

On appelle ***Graphe de*** *A, la partie $G(A) = \underset{u \in D(A)}{\cup} [u, Au] \subset E \times F$.*

On appelle ***Image de*** *A, la partie $R(A) = \underset{u \in D(A)}{\cup} Au \subset F$.*

On appelle ***Noyau de*** *A, la partie $N(A) = \{u \in D(A) \ ; \ Au = 0\} \subset E$.*

Définition 27 *On dit qu'un opérateur A est fermé si $G(A)$ est fermé dans $E \times F$.*

Remarque 28 *Pour prouver qu'un opérateur A est fermé on procède généralement de la manière suivante. On prend une suite (u_n) dans $D(A)$ telle que $u_n \to u$ dans E et $Au_n \to f$ dans F. Il s'agit ensuite de vérifier que*

(a) $u \in D(A)$

(b) $f = Au$

Si A est fermé, alors $N(A)$ est fermé.

En pratique, la plupart des opérateurs non-bornés que l'on rencontrera sont fermés et à domaine dense dans E.

3.5.2 Adjoint d'un opérateur non-borné

Définition 29 *de l'adjoint A^*. Soit $A : D(A) \subset E \longrightarrow F$ un opérateur non borné à domaine dense. On va définir un opérateur non-borné $A^* : D(A^*) \subset F' \longrightarrow E'$ comme suite. On pose*

$$D(A^*) = \{v \in F' \,;\, \exists\, c \geq 0 \text{ tel que } |\langle v, Au\rangle| \leq c\,\|u\| \;\; \forall u \in D(A)\}\,. \quad (3.75)$$

Il est clair que $D(A^*)$ est un sous espace vectoriel de F'. On va maintenant définir A^*v pour tout $v \in D(A^*)$

Etant donné $v \in D(A^*)$ on considère l'application $g : D(A) \subset E \longrightarrow \mathbb{R}$ définie par $g(u) = \langle v, Au\rangle \;\forall u \in D(A)$. On a $|g(u)| \leq c\,\|u\| \;\forall u \in D(A)$. Grâce au théorème de Hahn-Banach on sait que g peut être prolongée en une application linéaire $f : E \longrightarrow \mathbb{R}$ telle que $|f(u)| \leq c\,\|u\| \;\forall u \in E$; par suite $f \in E'$. On remarquera que le prolongement de g est unique puisque f est linéaire continue sur E et que $D(A)$ est dense. On pose $A^*v = f$. Il est clair que A^* est linéaire. L'opérateur $A^* : D(A^*) \subset F' \longrightarrow E'$ est appélé l'adjoint de A. On a par conséquent la relation fondamentale suivante qui lie A et A^* :

$$\langle v, Au\rangle_{F',F} = \langle A^*v, u\rangle_{E',E} \;\forall u \in D(A)\,,\; \forall v \in D(A^*)\,. \quad (3.76)$$

Remarque 30 *Dans la définition précédente si E est un espace de Hilbert, en identifiant E à son dual E', on définit l'adjoint A^* de l'opérateur linéaire $A : D(A) \subset E \longrightarrow E$, avec $\overline{D(A)} = E$, par $A^* : D(A^*) \subset E \longrightarrow E$ vérifiant $(u, Av) = (A^*u, v)$, $\forall u \in D(A^*)$, $\forall v \in D(A)$ où $D(A^*) = \{u \in E \;: \exists f \in E \,;\; (u, Av) = (f, v)\,, \forall v \in D(A)\}$. Pour $u \in D(A)$, on pose alors $A^*u = f$. L'opérateur A, est dit auto-adjoint si $A = A^*$ i.e.,*

$$(u, Av) = (Au, v)\,,\; \forall u, v \in D(A)\,. \quad (3.77)$$

L'opérateur A est dit skew-adjoint si $A = -A^*$ i.e.,

$$(u, Av) = -(Au, v)\,,\; \forall u, v \in D(A)\,. \quad (3.78)$$

3.5.3 Opérateur symétrique

Définition 31 *Un opérateur linéaire $B : D(B) \subset E \longrightarrow E$ sur l'espace de Hilbert E est dit symétrique si et seulement si $\overline{D(B)} = E$ et*

$$(Bu, v) = (u, Bv), \forall u, v \in D(B). \tag{3.79}$$

L'opérateur B est dit skew-symétrique si et seulement si

$$(Bu, v) = -(u, Bv), \forall u, v \in D(B). \tag{3.80}$$

Définition 32 *Un opérateur $A : D(A) \subset E \longrightarrow E$ sur un espace de Hilbert réel E est dit monotone ou dissipatif si*

$$(Au - Av, u - v) \geq 0 \ \forall u, v \in D(A). \tag{3.81}$$

De plus, l'opérateur A est dit strictement monotone si

$$(Au - Av, u - v) > 0 \ \forall u, v \in D(A), \text{ avec } u \neq v. \tag{3.82}$$

Finalement l'opérateur A est dit fortement monotone s'il existe une constante $c > 0$ telle que

$$(Au - Av, u - v) \geq c \|u - v\|^2 \ \forall u, v \in D(A). \tag{3.83}$$

Si l'opérateur A est linéaire, alors la condition de monotonicité précédente est équivalente à la simple condition de positivité

$$(Au, u) \geq 0 \ \forall u \in D(A), \tag{3.84}$$

et la condition fortement monotone est équivalente à

$$(Au, u) \geq c \|u\|^2 \ \forall u \in D(A). \tag{3.85}$$

Définition 33 *Un opérateur linéaire non-borné $A : D(A) \subset E \longrightarrow E$ sur un espace de Hilbert réel E est dit maximal monotone s'il est monotone et si*

$$R(I + A) = E \text{ i.e., } \forall f \in E, \ \exists u \in D(A) \text{ tel que } u + Au = f. \tag{3.86}$$

Proposition 34 *Soit $A : D(A) \subset E \longrightarrow E$ un opérateur linéaire non-borné sur un espace de Hilbert réel E. Si A est maximal monotone, alors*

a) $D(A)$ est dense dans E

b) A est fermé

c) pour tout $\lambda > 0, (I + \lambda A)$ est bijectif de $D(A)$ sur E, $(I + \lambda A)^{-1}$ est un opérateur borné et $\left\|(I + \lambda A)^{-1}\right\|_{\mathcal{L}(E)} \leq 1$.

3.5.4 Extension de Friedrichs des opérateurs symétriques

On écrit $A \subseteq B$ si l'opérateur B est une extension de l'opérateur A *i.e.*,

$$D(A) \subseteq D(B) \text{ et } Au = Bu \ \forall u \in D(A). \tag{3.87}$$

Soit $A : D(A) \subset E \longrightarrow E$ un opérateur linéaire symétrique. On considère l'équation

$$Au = f,\ u \in D(A). \tag{3.88}$$

Cette équation correspond à un problème aux limites classique pour une équation linéaire élliptique, donc en principe le problème (3.88) n'est pas solvable pour tout $f \in E$. Alors à travers (3.88), on étudie les deux problèmes généralisés

$$A_F u = f,\ u \in D(A_F), \tag{3.89}$$

$$(u, Av) = (f, v) \text{ pour } u \text{ fixé dans } X_E \text{ et pour tout } v \in D(A), \tag{3.90}$$

ici $A \subseteq A_F$. En résumé, on montrera que (3.89) et (3.90) sont équivalents et pour tout $f \in E$, les deux équations ont une unique solution. Cette solution peut aussi être obtenue par le problème variationnel suivant :

$$\frac{1}{2}(u, u)_E - (f, u) = \min!,\ u \in X_E. \tag{3.91}$$

On fait l'hypothèse suivante :

($\mathcal{H}$) L'opérateur linéaire $A : D(A) \subset E \longrightarrow E$ est symérique sur l'espace de Hilbert E avec $\overline{D(A)} = E$ et A est fortement monotone, *i.e* $(Au, u) \geq$

$c\|u\|^2$ pour tout $u \in D(A)$ et $c > 0$ fixé. Le but est de construire les extensions $A \subseteq A_F \subseteq A_E$, où A_F est appelé extension de Friedrichs et A_E est appelé extension énergétique de A. Notons que l'image X_E^* de l'opérateur $A_E : X_E \longrightarrow X_E^*$ demeure dans l'espace d'origine E. L'espace énergétique X_E est le complété de $D(A)$ par rapport au produit scalaire énergétique

$$(u,v)_E = (Au,v) \ \ \forall u,v \in D(A). \tag{3.92}$$

Théorème 35 *(Friedrichs) On suppose que l'hypothèse $(\mathcal{H})$ est vérifiée. Alors*

(a) il existe une extension auto-adjointe $A_F : D(A_F) \subseteq E \longrightarrow E$ de l'opérateur A avec $D(A_F) \subseteq X_E \subseteq E$ et

$$(A_F u,u) \geq c\|u\|^2 \ \ \forall u \in D(A_F).$$

(b) l'opérateur inverse $A_F^{-1} : E \longrightarrow E$ existe et est linéaire, continu, et auto-adjoint. Par conséquent, pour chaque $f \in E$, l'équation (3.88) a une unique solution.

(c) l'opérateur $A_F^{-1} : E \longrightarrow X_E$ est linéaire continu.

(d) les injections $X_E \subseteq E \subseteq X_E^*$ sont continues.

(e) l'opérateur A_F a l'extension $A_E : X_E \longrightarrow X_E^*$, où A_E est l'application duale de X_E^*, i.e., A_E est un homéomorphisme avec

$$\langle A_E u,v\rangle = \|u\|_E^2 \ \text{ pour tout } u \in X_E.$$

De plus,

$$A_F^{-1} f = A_E^{-1} f \ \text{ pour tout } f \in E.$$

(f) si l'injection $X_E \subseteq E$ est compacte, alors l'opérateur $A_F^{-1} : E \longrightarrow E$ est compact.

Corollaire 36 *Pour $f \in E$ donné, les problèmes (3.89), (3.90) , et (3.91) sont mutuellement équivalents. Chaque solution du problème original (3.88) est aussi une solution de (3.89), (3.90), et (3.91).*

Démonstration. Montrons que (3.88)⇒(3.89). En effet, comme $A \subseteq A_F$, toute solution $u \in D(A)$ de (3.88) vérifie $u \in D(A_F)$ et $Au = A_F u$. d'où $A_F u = f$ avec $u \in D(A_F)$. Ainsi (3.88)⇒(3.89). De plus, comme A_F est auto-adjoint, on a $(A_F u, v) = (u, A_F v) = (u, Av)$ $v \in D(A)$. Puisque $D(A_F) \subseteq X_E$, on en déduit que (3.89)⇒(3.90). On montre que les implications réciproques sont aussi vraies. ■

La preuve du théorème de Friedrichs repose sur le corollaire précédent, elle est donnée dans [50].

3.5.5 Notion de semi-groupe

Définition 37 *Un semi-groupe $\{S(t)\}$ sur un espace de Banach E est une famille d'opérateurs $S(t) : E \longrightarrow E$ pour tout $t \geq 0$, vérifiant*

$$\begin{array}{ll} (a) & S(t+s) = S(t)\,S(s) \ \ \forall t, s \geq 0, \\ (b) & S(0) = I \ \text{où } I \text{ est l'identité de } E. \end{array} \tag{3.93}$$

Le générateur $B : D(B) \subseteq E \longrightarrow E$ du semi-groupe $\{S(t)\}$ est définit par

$$Bw = \lim_{t \to 0^+} \frac{S(t)\,w - w}{t} \ \text{où } w \in D(B), \text{ et si cette limite existe.} \tag{3.94}$$

Un groupe à un paramètre $\{S(t)\}$ sur l'espace de Banach E est une famille d'opérateurs $S(t) : E \longrightarrow E$ $\forall t \in \mathbb{R}$ avec (a) et (b) $\forall t, s \in \mathbb{R}$. Toutes fois les générateurs sont aussi appelés générateurs infinitésimaux.

Exemple 38 *On considère l'équation différentielle*

$$\begin{cases} u'(t) = Bu(t) \ \ \text{sur } \mathbb{R}^+ \\ u(0) = w \end{cases} \tag{3.95}$$

Soit $B : E \longrightarrow E$ un opérateur linéaire continu sur l'espace de Banach E. On pose $S(t) = e^{tB}$ où

$$e^{tB} = \sum_{n=0}^{\infty} \frac{t^n B^n}{n!}. \tag{3.96}$$

Cette série converge sur l'espace de Banach $L(E, E)$ pour tout $t \in \mathbb{R}$, *i.e.*, cette série converge par rapport à la norme d'opérateur. $\{S(t)\}$ est un groupe

à un paramètre et la solution de (3.95) est donnée par

$$u(t) = S(t)\, w. \tag{3.97}$$

En effet d'aprrès (3.95)

$$\begin{cases} u'(0) = Bu(0) \\ u(0) = w, \end{cases} \tag{3.98}$$

or

$$\begin{aligned} u'(0) &= \lim_{t\to 0^+} \frac{u(t) - u(0)}{t} = \lim_{t\to 0^+} \frac{u(t) - w}{t} \\ \text{et } Bu(0) &= Bw = \lim_{t\to 0^+} \frac{S(t)\, w - w}{t}. \end{aligned}$$

Considerons maintenant le problème

$$(\mathcal{P}) \begin{cases} \dfrac{\partial^2 u}{\partial t^2} - \Delta u + m^2 u = -\lambda u^3 & \text{sur } \Omega \times]0,T[\ , \\ u(x,0) = u_0(x) & \text{sur } \Omega\ , \\ \dfrac{\partial u}{\partial t}(x,0) = v_0(x) & \text{sur } \Omega\ , \\ u(x,t) = 0 & \text{sur } \partial\Omega \times]0,T[\end{cases} \tag{3.99}$$

où Ω est un ouvert borné de $\mathbb{R}^d$ $(d \geq 1)$ de frontière $\partial\Omega$ régulière. Le nombre $\lambda \geq 0$ est un paramètre réel. Dans le cas où $\lambda = 0$ on obtient l'équation de Klein-Gordon décrite dans la mécanique quantique relativiste. Cette équation est de la forme $(L + m^2)\,\psi = 0$ où l'opérateur $L = \dfrac{\partial^2}{\partial t^2} - \Delta$ et m est la masse de la particule. Le problème (3.99) décrit le champ classique u d'un méson de masse m. Le terme non linéaire $-\lambda u^3$ avec $\lambda > 0$, décrit le self interaction du champ. On peut reformuler (3.99) sous forme d'une équation d'opérateurs

$$\begin{array}{ll} u'' + Au = f(u) \text{ sur }]0,T[& \\ u(0) = u_0,\ u'(0) = v_0 & \text{avec } f(u) = -\lambda u^3. \end{array} \tag{3.100}$$

On pose $X = L^2(\Omega)$, $D(A) = \mathcal{C}_0^\infty(\Omega)$ où $\mathcal{C}_0^\infty(\Omega)$ est l'ensemble des fonctions $u \in \mathcal{C}^\infty(\Omega)$ à support compact K qui dépend de u. $\forall u \in D(A)$, on a

$$\begin{aligned} (Au, u) &= (-\Delta u, u) + m^2 (u, u) \\ (Au, u) &= (\nabla u, \nabla u) + m^2 (u, u) \geq m^2 (u, u). \end{aligned} \tag{3.101}$$

Soit $\widetilde{A} : D\left(\widetilde{A}\right) \subseteq X \longrightarrow X$ l'extension de Friedrichs de A. Alors les espaces d'énergie de A et $-\Delta$ sont les mêmes en norme équivalente, *i.e.*, l'espace d'énergie de A est égal à $X_E = \overset{\circ}{W}{}_2^1(\Omega)$.

Lemme 39 *Soit $\Omega \subset \mathbb{R}^3$ un ouvert borné. Alors l'opérateur*

$$g : \begin{array}{c} \overset{\circ}{W}{}_2^1(\Omega) \longrightarrow L^2(\Omega) \\ u \longmapsto u^3 \end{array} \quad \textit{est localement lipschitzien continu.} \tag{3.102}$$

Démonstration. Soit $u, v \in \mathcal{C}_0^\infty(\Omega)$ tels que $\|u\| \leq r$; $\|v\| \leq r$ ($\|.\|$ norme sur $\overset{\circ}{W}{}_2^1(\Omega)$) par l'inégalité de Hölder basée sur $\frac{1}{3} + \frac{2}{3} = 1$, on a

$$\begin{aligned} \|f(u) - f(v)\|_2 &= \|(u-v)(u^2+uv+v^2)\|_2 \\ &\leq \|u-v\|_6 \|u^2+uv+v^2\|_3^2 \\ &\leq \|u-v\|_6 \left(\|u\|_6^2 + \|u\|_6 \|v\|_6 + \|v\|_6^2\right)^2. \end{aligned} \tag{3.103}$$

On va montrer que $\left(\|u\|_6^2 + \|u\|_6 \|v\|_6 + \|v\|_6^2\right)^2 \leq C$ où C est une constante. ■

Lemme 40 *Soit $\Omega \subset \mathbb{R}^3$ un ouvert borné. Alors l'injection $\overset{\circ}{W}{}_2^1(\Omega) \subseteq L^6(\Omega)$ est continue.*

Démonstration. Soit $x = (\xi, \eta, \zeta)$ et $u \in \mathcal{C}_0^\infty(\Omega)$. L'intégration de $(u^4)_\xi = 4u^3 u_\xi$ donne $|u(x)|^4 = \displaystyle\int_{-\infty}^{x} 4u^3 u_\xi d\xi \leq 4\int_{-\infty}^{+\infty} |u^3 u_\xi| \, d\xi$. En remplaçant ξ par η, on obtient

$$\begin{aligned} |u(x)|^6 \leq\ & 2^3 \left(\int_{-\infty}^{+\infty} \left|u^3 u_\xi\right| d\xi\right)^{1/2} \times \\ & \left(\int_{-\infty}^{+\infty} \left|u^3 u_\eta\right| d\eta\right)^{1/2} \left(\int_{-\infty}^{+\infty} \left|u^3 u_\zeta\right| d\zeta\right)^{1/2}. \end{aligned}$$

On intègre sur Ω et on applique l'inégalité de Hölder basée sur $\frac{1}{2}+\frac{1}{2}=1$. Alors $\|u\|_6^6 \leq K\|u\|_6^{\frac{18}{4}}\|u\|_6^{\frac{6}{4}}$. Ainsi

$$\|u\|_6 \leq K\|u\| \ \ \forall u \in \mathcal{C}_0^\infty(\Omega). \tag{3.104}$$

(3.103) et (3.104)$\Rightarrow$

$$\begin{array}{rl} \|f(u)-f(v)\|_2 & \leq \|u-v\|_6 \left(\begin{array}{c} K^2\|u\|^2 + K^2\|u\|\|v\| \\ +K^2\|v\|^2 \end{array} \right)^2 \\ & \leq 9K^4r^4\|u-v\|_6 \\ & = C\|u-v\|_6, \text{ avec } C = 9K^4r^4 \end{array}$$

■

On considère le problème initial suivant :

$$\left\{ \begin{array}{lcl} u''(t)+Au(t) & = & f(u(t)),\ 0<t<\infty, \\ u(0) & = & u_0,\ u'(0)=v_0. \end{array} \right. \tag{3.105}$$

On fait les hypothèses suivantes :

$(\mathcal{H}_1)$ l'opérateur linéaire $A : D(A) \subseteq X \longrightarrow X$ est auto-adjoint et fortement monotone sur le H espace X sur $K=\mathbb{R}$, C.

Soit X_E l'espace d'énergie de A avec la norme $\|.\|_E$, *i.e.*, X_E est le complété de $D(A)$ par rapport au produit scalaire $(u,v)_E=(Au,v)$.

$(\mathcal{H}_2)$ l'opérateur $f : X_E \longrightarrow X$ est lipschitzienne continu *i.e.*, $\forall R>0$ il existe L tel que $\|f(u)-f(v)\| \leq L\|u-v\|_E$ pour tout $u,v \in X_E$
avec $\|u\|_E \leq R\,;\ \|v\|_E \leq R$.

On pose $v=u'$. Au lieu de (3.105) on considère le système du premier ordre

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} 0 & I \\ -A & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ f(u) \end{pmatrix}. \tag{3.106}$$

On pose de plus $z=(u,v)$ et on écrit (3.105) sous la forme

$$\left\{ \begin{array}{lcl} z'(t) & = & Cz(t)+F(z(t)),\ 0<t<\infty, \\ z(0) & = & z_0. \end{array} \right. \tag{3.107}$$

Soit $Z=X_E \times X$ et $D(C)=D(A)\times X_E$.

Hypothèse : On suppose que l'opérateur $C : D(C) \subseteq Z \longrightarrow Z$ est skew-adjoint et génère un groupe unitaire à un paramètre$\{S(t)\}$.

Corollaire 41 *On suppose* $(\mathcal{H}_1)$ *et* $f \equiv 0$ *i.e.,* $F \equiv 0$. *Alors pour chaque* $z_0 \in D(C)$, *l'équation (3.107) a une solution classique unique dans le sens de la définition ci-dessous donnée par* $z(t) = S(t) z_0$ *pour chaque* $z_0 \in Z$. *La fonction* z *est appelée solution milieu de l'équation (3.107).*

Définition 42 *La fonction* $z :]0, \infty[\longrightarrow Z$ *étant donnée,*

(i) z *est appelée solution classique de (3.107) si et seulement si* $z \in \mathcal{C}^1(]0, +\infty[)$ *; est continue sur* $[0, +\infty[$ *et vérifie (3.107).*

(ii) *Toute solution* z *vérifiant*

$$z(t) = S(t) z_0 + \int_0^t S(t-s) F(z(s)) \, ds, \tag{3.108}$$

est appelée solution milieu du problème initial (3.107).

Démonstration. Du corollaire 42

On pose $B = A^{1/2}$. On a $(u, \overline{u})_E = (Bu, B\overline{u})$ $\forall u, \overline{u} \in X_E$ donc $\forall z, \overline{z} \in Z$, $(z, \overline{z})_Z = (u, \overline{u})_E + (v, \overline{v}) = (Bu, B\overline{u}) + (v, \overline{v})$

où $(.,.)$ est le produit scalaire sur X. Notons que $Cz = (v, -Au)$.

$\forall z, \overline{z} \in D(C)$, $(z, C\overline{z})_Z = (Bu, B\overline{v}) - (v, A\overline{u}) = (Au, \overline{v}) - (v, A\overline{u})$.

$(Cz, \overline{z})_Z = (Bv, B\overline{u}) - (Au, \overline{v}) = (v, A\overline{u}) - (Au, \overline{v})$.

Ainsi $(z, C\overline{z})_Z = -(Cz, \overline{z})_Z$ *i.e.,* C est skew-symétrique.

De plus $R(I \pm C) = Z$. En effet, l'équation $\overline{z} = z \pm Cz$ est équivalente à $\overline{u} = u \pm v$, $\overline{v} = v \pm Au$ et la dernière équation a la solution

$u = (I + A)^{-1}(\overline{u} \pm \overline{v})$, $v = \pm(\overline{u} \pm u)$, part conséquent, les deux opérateurs C et $-C$ sont maximum dissipative. Le théorème suivant conduit à l'assertion. ■

Théorème 43 *Soit* $B : D(B) \subseteq X \longrightarrow X$ *un opérateur linéaire sur le* H*-espace* X *sur* $K = \mathbb{R}$, C. *Alors les assertions suivantes sont équivalentes :*

(i) B *est le générateur d'un groupe unitaire à un paramètre*

(ii) B *et* $-B$ *sont maximaux dissipatifs et* $\overline{D(B)} = X$

(iii) B *est skew-adjoint.*

Notons l'important fait suivant : Puisque $(Au,u) \geq 0 \ \forall u \in D(A)$ la résolution de $(I+A)^{-1} : X \longrightarrow X$ existe

Théorème 44 *On suppose que les hypothèses $(\mathcal{H}_1)$ et $(\mathcal{H}_2)$ précédentes sont vérifiées. Alors pour chaque $z_0 = (u_0, v_0)$ dans Z, il existe des nombres $T > 0$ et $r > 0$ tels que le problème (3.107) ait exactement une solution milieu $z \in Y$ avec $\|z - z_0\|_Y \leq r$, où $Y = \mathcal{C}^1([-T,T], Z)$ est muni de la norme usuelle*

$\|z\|_Y = \max\limits_{-T\leq t\leq T} \|z(t)\|_Z$.

Pour résoudre le problème initial (3.107) avec F non identiquement nulle, on considère l'équation intégrale :

$$z(t) = S(t)\, z_0 + \int_0^t S(t-s)\, F(z(s))\, ds. \tag{3.109}$$

Les solutions de (3.109) sont appelées solutions milieu de (3.107) par le théorème 19.D (voir référence).

Démonstration. Du Théorème 45

Soit $\|.\|$ la norme sur Y. Notons $M = \{z \in Y \ ; \ \|z - z_0\| \leq r\}$, on écrit (3.109) dans la forme suivante du point fixe de l'équation

$$z = Kz, \ z \in M. \tag{3.110}$$

Notons $(\mathcal{H}_2)$ et $\|S(t)\|_Z \leq 1$. On obtient donc

$$\begin{array}{rcl} \|Kz - K\overline{z}\| & \leq & \max\limits_{-T\leq t\leq T} \left| \int_0^t \|F(z(s)) - F(\overline{z}(s))\|_Z \, ds \right|, \\ & \leq & TL(r)\, \|z - \overline{z}\| \text{ pour tout } z, \overline{z} \in M, \end{array} \tag{3.111}$$

ce qui implique

$$\begin{array}{rcl} \|Kz - z_0\| & \leq & \|Kz - Kz_0\| + \|Kz_0 - z_0\| \\ & \leq & TL(r)\, \|z - z_0\| + \|Kz_0 - z_0\| \, \forall z \in M. \end{array}$$

Ainsi on peut choisir r et T dans une voie telle que

$$K(M) \subseteq M \text{ et } K : M \longrightarrow M \text{ est } k - \text{contractante}. \tag{3.112}$$

Comme conséquence, (3.110) a une unique solution par le théorème du point fixe de Banach. ■

Pour $\lambda > 0$, l'existence et l'unicité de la solution du problème sera alors acquise par le théorème du point fixe de Banach [50] (théorème 1.A.). Pour la résolution numérique du problème $(\mathcal{P})$, on utilisera un schéma numérique de collocation par les RBF semblable à celui utilisé par Y.C. Hon [21] pour la résolution de l'équation Black-Scholes.

Chapitre 4

La méthode de collocation RBF et ses applications

4.1 Présentation des fonctions radiales de bases utiles dans la Méthode de collocation RBF

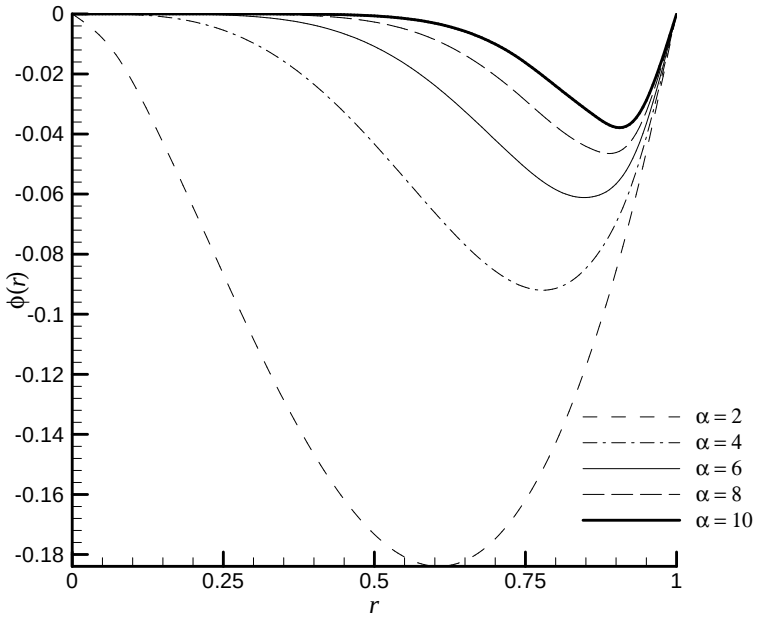

FIG. 4.1 – Thin-plate spline $\phi(r) = r^{\alpha} \ln r$ (α pair)

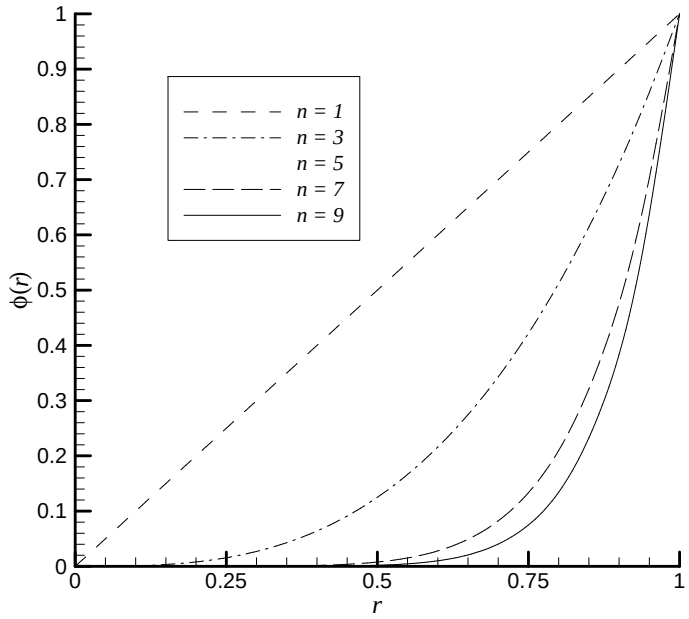

FIG. 4.2 – Fonction radiale de type conique : $\phi(r) = r^n (n \text{ impair})$

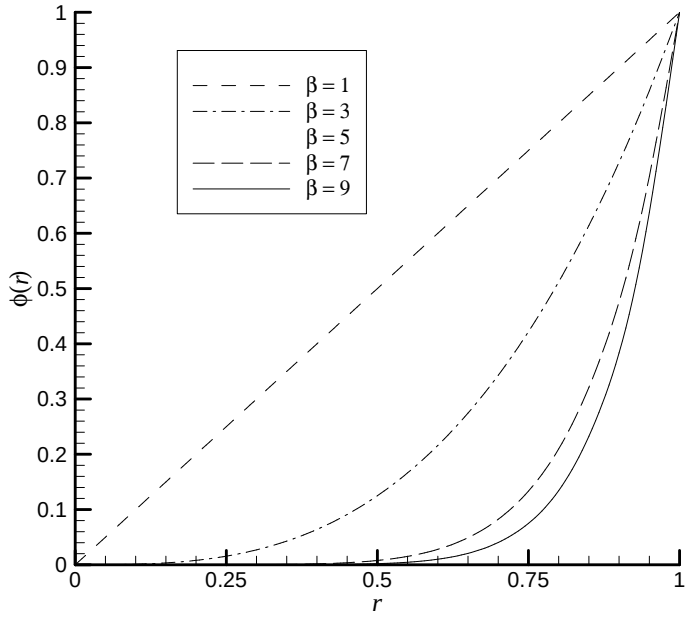

FIG. 4.3 – Multiquadriques : $\phi(r) = (r^2 + c^2)^{\frac{\beta}{2}}$ (β impair), $c > 0$.

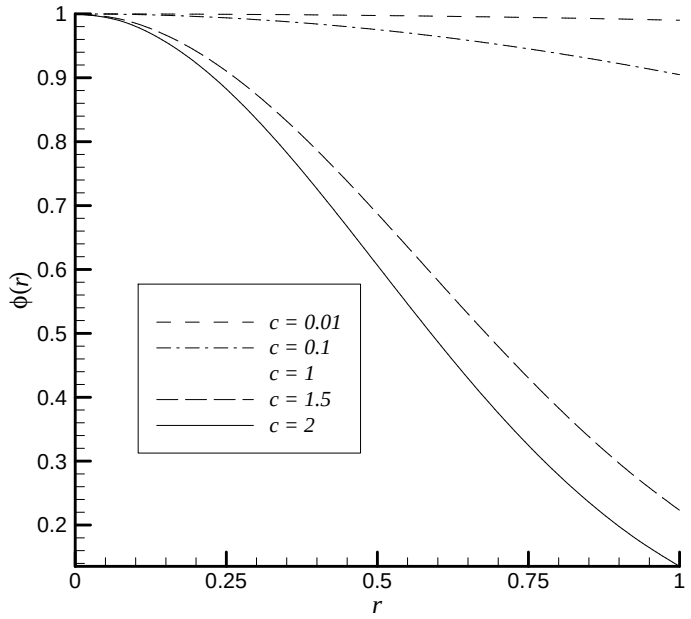

FIG. 4.4 – Gaussians : $\phi(r) = \exp(-cr^2)$, $c > 0$.

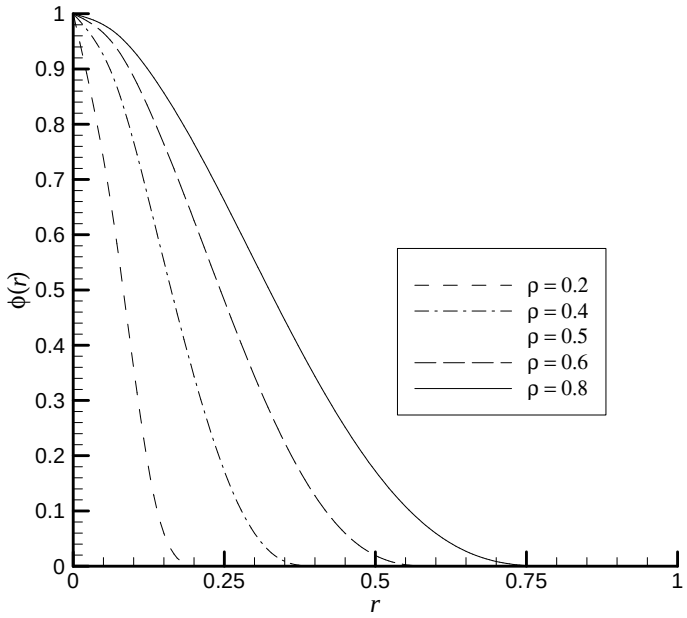

FIG. 4.5 – Fonction radiale à support compact : $\phi(r) = \left(1-\frac{r}{\rho}\right)^3 \left(1+3\frac{r}{\rho}+\left(\frac{r}{\rho}\right)^2\right)$ si $0 \leq r \leq \rho$, $\phi(r) = 0$ si $r > \rho$

Nous allons maintenant passer à l'application de la méthode de collocation RBF.

On considère l'équation aux dérivées partielles suivante

$$\begin{array}{l} Lu\left(x\right)=f\left(x\right),\, x\in\Omega\subset\mathbb{R}^{d},\\ Qu\left(x\right)=g\left(x\right),x\in\partial\Omega. \end{array} \tag{4.1}$$

Soit

$$\left(x_{j}\right)_{1\leq j\leq n}\subset\Omega \text{ et } \left(x_{j}\right)_{n+1\leq j\leq n+m}\subset\partial\Omega. \tag{4.2}$$

La méthode de collocation par les RBF consiste à chercher une solution approchée u^{*}, du problème (4.1), sous la forme

$$u^{*}\left(x\right)=\sum_{j=1}^{n+m}\lambda_{j}\phi_{j}\left(x\right), \tag{4.3}$$

où $\phi_{j}\left(x\right)=\phi\left(\left\|x-x_{j}\right\|_{2}\right)$, ϕ étant la fonction radiale. L et Q étant linéaires, en remplaçant u^{*} dans (4.1), on a

$$\begin{aligned} Lu^{*}\left(x_{i}\right) &= \sum_{j=1}^{n+m}\lambda_{j}L\phi_{j}\left(x_{i}\right)=f\left(x_{i}\right),\, 1\leq i\leq n, \\ Qu^{*}\left(x_{i}\right) &= \sum_{j=1}^{n+m}\lambda_{j}Q\phi_{j}\left(x_{i}\right)=g\left(x_{i}\right),\, n+1\leq i\leq n+m. \end{aligned} \tag{4.4}$$

On résout donc le système d'inconnus $\left(\lambda_{j}\right)_{1\leq j\leq n+m}$,

$$\begin{aligned} \sum_{j=1}^{n+m}\lambda_{j}L\phi_{j}\left(x_{i}\right) &= f\left(x_{i}\right),\, 1\leq i\leq n, \\ \sum_{j=1}^{n+m}\lambda_{j}Q\phi_{j}\left(x_{i}\right) &= g\left(x_{i}\right),\, n+1\leq i\leq n+m, \end{aligned} \tag{4.5}$$

et la solution approchée est donnée par (4.3). Pour un problème d'évolution en dimension d, par exemple de la forme

$$\begin{array}{l} \dfrac{\partial u}{\partial t}+Lu=f \text{ sur } \Omega\times[0,T],\\ Bu=g \text{ sur } \partial\Omega, \end{array} \tag{4.6}$$

on utilise un schéma de discrétisation en temps ; par exemple Cranck-Nicolson

$$\begin{cases} \dfrac{u^{n+1}-u^n}{\Delta t} = \theta\left(Lu^{n+1}+f\left(x,t_{n+1}\right)\right)+\left(1-\theta\right)\left(Lu^n+f\left(x,t_n\right)\right) \\ \text{et des } \left(C.L\right), \end{cases} \tag{4.7}$$

où $\Delta t = t_{n+1}-t_n$ est le pas de temps, $u^n\left(n=0,1,2...\right)$ est la solution au pas de temps $t_n = n\Delta t$ et $\theta = 1/2$. Dans ce cas la solution approchée à l'instant t_n par la méthode de collocation RBF s'écrit

$$u^n\left(x,t_n\right) = \sum_{j=1}^{n+m} \lambda_j\left(t_n\right)\phi_j\left(x\right). \tag{4.8}$$

4.2 Résultats numériques relatifs au problème raide : $\varepsilon u''\left(x\right)+u'\left(x\right)=f\left(x\right)$

4.2.1 Second membre nul

$$\begin{aligned} \varepsilon u''\left(x\right)+u'\left(x\right) &= 0,\ x\in\left(0,1\right), \\ u\left(0\right) &= 0,\ u\left(1\right)=1. \end{aligned} \tag{4.9}$$

solution analytique

$$u\left(x\right) = \frac{1-\exp\left(\dfrac{x}{\varepsilon}\right)}{1-\exp\left(\dfrac{-1}{\varepsilon}\right)}. \tag{4.10}$$

4.2.2 Second membre non nul

$$\begin{aligned} \varepsilon u''\left(x\right)+u'\left(x\right) &= -1,\ x\in\left(0,1\right), \\ u\left(0\right) &= 0,\ u\left(1\right)=1, \end{aligned}$$

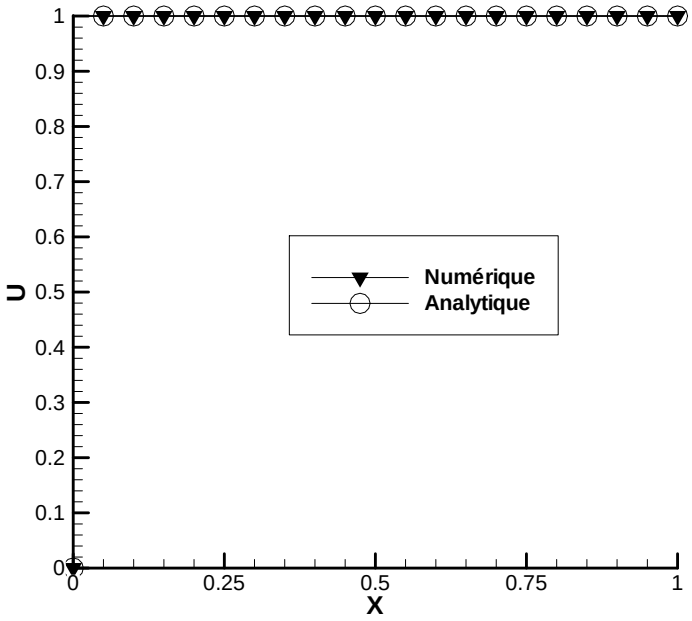

FIG. 4.6 – Problème raide avec second membre nul. $\varepsilon = c = 10^{-5}$, $n = 60$, $p = 6$ Multiquadriques. p définissant (3.43)

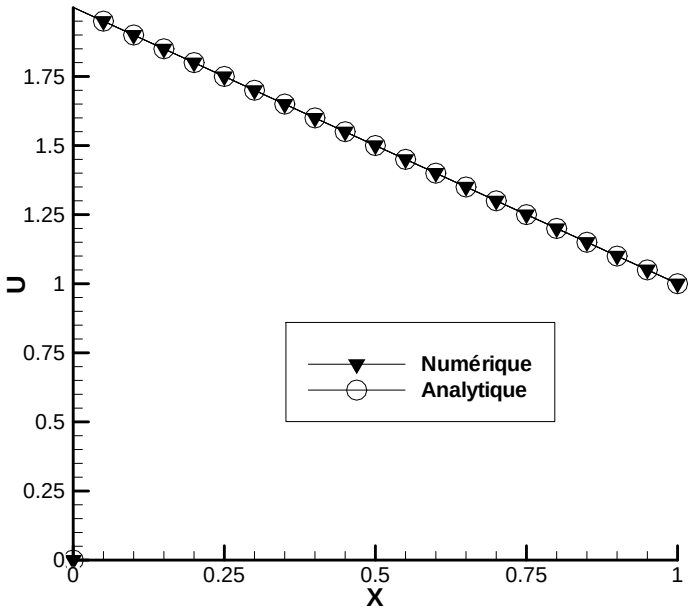

FIG. 4.7 – Problème raide avec second membre non nul : $f(x) = -1$, $\varepsilon = c = 10^{-5}$, $n = 60$, $p = 6$. Multiquadriques. p définissant (3.43)

solution analytique

$$u(x) = \frac{2\left[1 - \exp\left(\frac{-x}{\varepsilon}\right)\right]}{\left[1 - \exp\left(\frac{-1}{\varepsilon}\right)\right]} - x. \tag{4.11}$$

Pour les solutions numériques dans les deux cas selon la méthode de la quasi-interpolation, les calculs ont été éffectués en double précision à l'aide de FORTRAN 77. On a pris 1001 points réguliers $0 : 0.001 : 1$ de l'intervalle $[0, 1]$. L'écart type entre la solution analytique u et la solution numérique u^* est défini par

$$\text{RMSE} = \frac{1}{1001}\sqrt{\sum_{i=1}^{1001}\left(\frac{u^*(x_i) - u(x_i)}{u(x_i)}\right)^2}. \tag{4.12}$$

où u^* est donnée par (3.53) qui utilise (3.50), (3.51), (3.56) et (3.57). La spline linéaire B_j^1 qui permet d'approcher f par f^* dans (3.45), est donnée par (3.46).

Table du RMSE pour la résolution de : $\varepsilon u''(x) + u'(x) = f(x)$ avec $\varepsilon = c$.

$\varepsilon = c$	n	p	RMSE $(f(x) = 0)$	RMSE $(f(x) = -1)$
10^{-8}	200	2	$4,3.10^{-2}$	$6,7.10^{-2}$
10^{-6}	200	2	2.10^{-3}	3.10^{-3}
10^{-4}	200	2	2.10^{-5}	2.10^{-5}
10^{-8}	100	3	$3,2.^{-9}$	$4,7.10^{-9}$
10^{-6}	100	3	$2,6.10^{-9}$	$3,8.10^{-9}$
10^{-4}	100	3	$2,9.10^{-6}$	$2,9.10^{-6}$
10^{-8}	80	4	$3,4.10^{-2}$	5.10^{-2}
10^{-6}	80	4	$1,9.10^{-6}$	3.10^{-6}
10^{-4}	80	4	$3,5.10^{-6}$	$3,5.10^{-6}$
10^{-6}	80	3	$5,6.10^{-10}$	$5,6.10^{-10}$
10^{-3}	70	5	$2,2.10^{-4}$	$2,2.10^{-4}$
10^{-4}	70	5	$3,5.10^{-6}$	$3,5.10^{-6}$
10^{-5}	70	5	$7,5.10^{-7}$	$9,5.10^{-7}$
10^{-3}	60	6	$2,2.10^{-4}$	$2,2.10^{-4}$
10^{-4}	60	6	$3,6.10^{-6}$	$3,9.10^{-6}$
10^{-5}	60	6	$7,1.10^{-6}$	$1,2.10^{-5}$
10^{-6}	60	3	5.10^{-11}	$5,1.10^{-11}$
10^{-6}	40	3	$5,3.10^{-12}$	$5,3.10^{-12}$

Table du RMSE pour la résolution de : $\varepsilon u''(x) + u'(x) = f(x)$ avec $\varepsilon \neq c$.

ε	c	n	p	RMSE $(f(x) = 0)$	RMSE $(f(x) = -1)$
10^{-6}	10^{-8}	100	3	$2,6.10^{-9}$	$1,9.10^{-9}$
10^{-6}	10^{-8}	100	4	$1,6.10^{-2}$	$2,5.10^{-2}$
10^{-5}	10^{-7}	200	5	10^{-4}	$1,6.10^{-4}$
10^{-5}	10^{-7}	200	6	$1,7.10^{-2}$	$2,6.10^{-2}$
10^{-4}	10^{-6}	80	2	$2,6.10^{-4}$	$2,6.10^{-4}$
10^{-4}	10^{-6}	80	6	$5,7.10^{-6}$	$8,4.10^{-6}$
10^{-3}	10^{-5}	70	5	6.10^{-4}	$6,4.10^{-4}$
10^{-3}	10^{-5}	70	4	6.10^{-4}	6.10^{-4}
10^{-2}	10^{-4}	60	3	$1,1.10^{-2}$	10^{-2}
10^{-4}	10^{-5}	60	2	$4,6.10^{-4}$	$4,6.10^{-4}$

Il est aussi interessant d'observer d'après les deux tableaux que les erreurs pour les cas $f(x) = 0$ et $f(x) = -1$ ont le même ordre de grandeur.

4.3 Résultats numériques relatifs à la concentration d'un contaminant :

On considère l'équation générale de transport

$$\frac{\partial c}{\partial t} = D_x \frac{\partial^2 c}{\partial x^2} + D_y \frac{\partial^2 c}{\partial y^2} + D_z \frac{\partial^2 c}{\partial z^2} - V_x \frac{\partial c}{\partial x} - V_y \frac{\partial c}{\partial y} - V_z \frac{\partial c}{\partial z} - \lambda c, \qquad (4.13)$$

avec les conditions aux limites

$$\begin{aligned}
c(x,y,z,t) &= e^{-\lambda t}(c_1+c_2)\left(c_3+c_4 e^{\frac{V_y}{D_y}y}\right)\left(c_5+c_6 e^{\frac{V_z}{D_z}z}\right) \text{ en } x=0,\\
\frac{\partial c}{\partial x}(x,y,z,t) &= e^{-\lambda t}\left(c_3+c_4 e^{\frac{V_y}{D_y}y}\right)\left(c_5+c_6 e^{\frac{V_z}{D_z}z}\right) c_2\frac{V_x}{D_x}e^{\frac{V_x}{D_x}} \text{ en } x=1,\\
\frac{\partial c}{\partial y}(x,y,z,t) &= e^{-\lambda t}\left(c_1+c_2 e^{\frac{V_x}{D_x}x}\right)\left(c_5+c_6 e^{\frac{V_z}{D_z}z}\right) c_4\frac{V_y}{D_y} \text{ en } y=0,\\
\frac{\partial c}{\partial y}(x,y,z,t) &= e^{-\lambda t}\left(c_1+c_2 e^{\frac{V_x}{D_x}x}\right)\left(c_5+c_6 e^{\frac{V_z}{D_z}z}\right) c_4\frac{V_y}{D_y}e^{\frac{V_y}{D_y}} \text{ en } y=1,\\
\frac{\partial c}{\partial z}(x,y,z,t) &= e^{-\lambda t}\left(c_1+c_2 e^{\frac{V_x}{D_x}x}\right)\left(c_3+c_4 e^{\frac{V_y}{D_y}y}\right) c_6\frac{V_z}{D_z} \text{ en } z=0,\\
\frac{\partial c}{\partial z}(x,y,z,t) &= e^{-\lambda t}\left(c_1+c_2 e^{\frac{V_x}{D_x}x}\right)\left(c_3+c_4 e^{\frac{V_y}{D_y}y}\right) c_6\frac{V_z}{D_z}e^{\frac{V_z}{D_z}} \text{ en } z=1,
\end{aligned} \qquad (4.14)$$

et la condition initiale

$$c(x,y,z,0) = \left(c_1+c_2 e^{\frac{V_x}{D_x}x}\right)\left(c_3+c_4 e^{\frac{V_y}{D_y}y}\right)\left(c_5+c_6 e^{\frac{V_z}{D_z}z}\right), \qquad (4.15)$$

où c est la concentration d'un contaminant, $V = (V_x, V_y, V_z)$ est sa vitesse d'infiltration, D_x, D_y, D_z sont ses coefficients de dispersion dans les directions x, y, z respectivement, λ est sa vitesse de chute. La solution exacte du problème s'écrit

$$c(x,y,z,t) = e^{-\lambda t}\left(c_1 + c_2 e^{\frac{V_x}{D_x}x}\right)\left(c_3 + c_4 e^{\frac{V_y}{D_y}y}\right)\left(c_5 + c_6 e^{\frac{V_z}{D_z}z}\right). \tag{4.16}$$

Nous allons tester un modèle $2D$, *i.e.*,

$$D_z = V_z = 0 \text{ dans l'équation (4.13).} \tag{4.17}$$

La solution exacte sera donc

$$c(x,y,t) = e^{-\lambda t}\left(c_1 + c_2 e^{\frac{V_x}{D_x}x}\right)\left(c_3 + c_4 e^{\frac{V_y}{D_y}y}\right). \tag{4.18}$$

4.3.1 Problème stationnaire avec conditions de raccordement

$$c(x,y) = c(x,y,0) = \left(c_1 + c_2 e^{\frac{V_x}{D_x}x}\right)\left(c_3 + c_4 e^{\frac{V_y}{D_y}y}\right). \tag{4.19}$$

Pour simplifier on suppose que

$$D_x = D_y = V_x = V_y = \lambda = 1. \tag{4.20}$$

En tenant compte des conditions de raccordement, on trouve la solution exacte du problème stationnaire

$$c(x,y) = c_2 c_4 e^{x+y}. \tag{4.21}$$

Pour les applications on prendra

$$c_2 = c_4 = 1, \tag{4.22}$$

donc

$$c(x,y) = e^{x+y}. \tag{4.23}$$

La solution numérique du problème stationnaire par la méthode de collocation utilisant les RBF (fonctions radiales de base) a été testée sur le domaine $\Omega = [0,1]^2$. Dans Toute la suite, le RMSE entre la solution analytique u et la solution numérique u^* sera défini par

$$\text{RMSE} = \sqrt{\left(\frac{\sum_{i=0}^{n} \left[u^*(x_i) - u(x_i)\right]^2}{\sum_{i=0}^{n} \left[u(x_i)\right]^2} \right)} \tag{4.24}$$

Table du RMSE pour la résolution de :

$$\begin{array}{l} \dfrac{\partial^2 c}{\partial x^2} + \dfrac{\partial^2 c}{\partial y^2} - \dfrac{\partial c}{\partial x} - \dfrac{\partial c}{\partial y} - c = 0 \text{ sur } \Omega \\ c(x,y) = e^y \text{ en } x = 0, \\ \dfrac{\partial c}{\partial x} = e^{1+y} \text{ en } x = 1, \\ \dfrac{\partial c}{\partial y} = e^x \text{ en } y = 0, \\ \dfrac{\partial c}{\partial y} = e^{1+x} \text{ en } y = 1. \end{array} \tag{4.25}$$

La solution exacte du problème étant donnée en (4.23), on trouve :

$N = n \times n$ (mailles)	$\phi(r) = r^4 \ln r$	$\phi(r) = r^6 \ln r.$
n	RMSE	RMSE
10	$5,2.10^{-3}$	6.10^{-2}
15	$1,6.10^{-2}$	$1,3.10^{-2}$
20	$6,8.10^{-4}$	$5,2.10^{-3}$
25	$3,5.10^{-3}$	$2,5.10^{-3}$
30	2.10^{-3}	$1,3.10^{-3}$
35	$1,2.10^{-3}$	$8,2.10^{-4}$
40	$8,2.10^{-4}$	$5,3.10^{-4}$

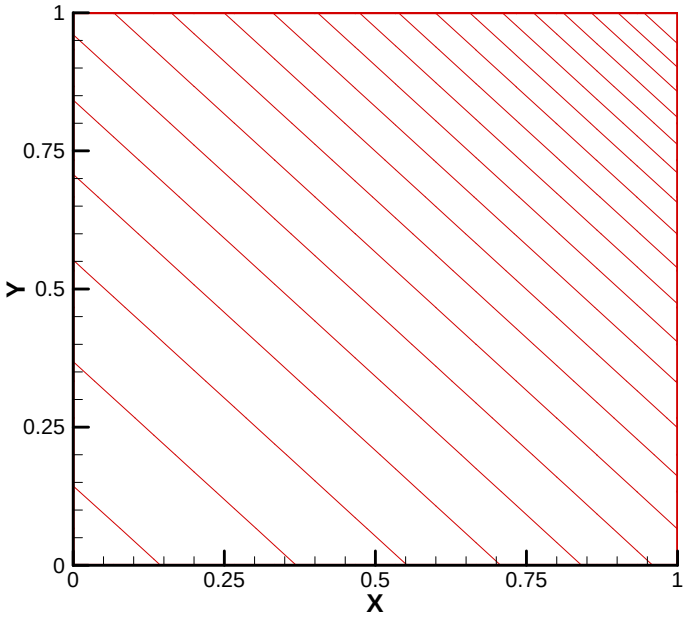

FIG. 4.8 – Problème de la concentration stationnaire avec conditions de raccordement. Isovaleurs de la solution numérique. $\phi(r) = r^5$.

$N = n \times n$ (mailles)	$\phi(r) = r^5$	$\phi(r) = (r^2 + c^2)^{\frac{3}{2}}$, $c = 10^{-5}$	$\phi(r) = (r^2 + c^2)^{\frac{5}{2}}$, $c = 10^{-5}$
n	RMSE	RMSE	RMSE
10	$2,1.10^{-3}$	$1,1.10^{-2}$	2.10^{-3}
15	$2,3.10^{-4}$	$5,6.10^{-3}$	$2,3.10^{-4}$
20	$1,3.10^{-4}$	$3,7.10^{-3}$	$1,3.10^{-4}$
25	$1,2.10^{-4}$	$2,7.10^{-3}$	$1,2.10^{-4}$
30	10^{-4}	$2,1.10^{-3}$	10^{-4}
35	$8,2.10^{-5}$	$1,7.10^{-3}$	$8,2.10^{-5}$
40	$6,6.10^{-5}$	$1,4.10^{-3}$	$6,6.10^{-5}$

On peut alors représenter les isovaleurs de la solution numérique puis le champ de vitesses.

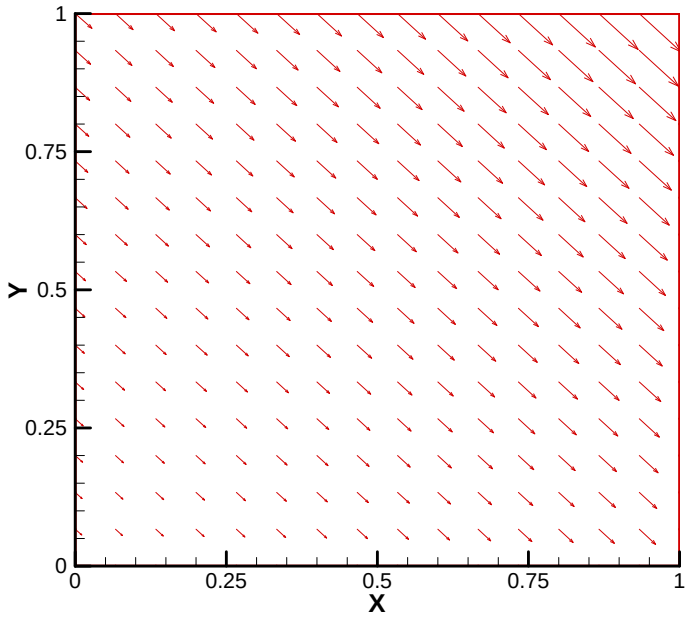

FIG. 4.9 – Problème de la concentration stationnaire avec conditions de raccordement. Champ de vitesses. $\phi(r) = r^5$.

4.3.2 Problème instationnaire avec conditions de raccordement

Avec le choix des constantes données en (4.20), la solution exacte du problème d'évolution s'écrit alors

$$c\,(x,y,t) = e^{x+y-t}. \tag{4.26}$$

On a

$$\begin{aligned} &c\,(0,y,t) = e^{y-t}, \\ &\frac{\partial c}{\partial x}\,(1,y,t) = e^{1+y-t}, \\ &\frac{\partial c}{\partial y}\,(x,0,t) = e^{x-t}, \\ &\frac{\partial c}{\partial y}\,(x,1,t) = e^{x+1-t}. \end{aligned} \tag{4.27}$$

D'où

$$\begin{aligned} &c\,(0,y,t)_{|y=0} = \frac{\partial c}{\partial y}\,(x,0,t)_{|x=0} = e^{-t}, \\ &c\,(0,y,t)_{|y=1} = \frac{\partial c}{\partial y}\,(x,1,t)_{|x=0} = e^{1-t}, \\ &\frac{\partial c}{\partial x}\,(1,y,t)_{|y=0} = \frac{\partial c}{\partial y}\,(x,0,t)_{|x=1} = e^{1-t}, \\ &\frac{\partial c}{\partial x}\,(1,y,t)_{|y=1} = \frac{\partial c}{\partial y}\,(x,1,t)_{|x=1} = e^{2-t}. \end{aligned} \tag{4.28}$$

Les conditions de raccordement sont donc bien vérifiées.

Table du RMSE pour la résolution de :

$$\begin{aligned} &\frac{\partial c}{\partial t} = \frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2} - \frac{\partial c}{\partial x} - \frac{\partial c}{\partial y} - c \text{ sur } \Omega \\ &c\,(x,y,t) = e^{y-t} \text{ en } x = 0, \\ &\frac{\partial c}{\partial x} = e^{1+y-t} \text{ en } x = 1, \\ &\frac{\partial c}{\partial y} = e^{x-t} \text{ en } y = 0, \\ &\frac{\partial c}{\partial y} = e^{1+x-t} \text{ en } y = 1. \end{aligned}$$

$N = n \times n$ (mailles), $\phi(r) = \exp(-cr^2)$						
n	c	RMSE	c	RMSE	c	RMSE
10	10^{-1}	$5,7.10^{-4}$	1	$1,5.10^{-5}$	1,1	$4,9.10^{-6}$
15	10^{-1}	$3,8.10^{-3}$	1	$1,9.10^{-6}$	1,1	$8,2.10^{-6}$
20	10^{-1}	$4,2.10^{-3}$	1	2.10^{-6}	1,1	$5,1.10^{-6}$
25	10^{-1}	$1,7.10^{-3}$	1	6.10^{-6}	1,1	$2,4.10^{-5}$
30	10^{-1}	1.10^{-3}	1	$1,1.10^{-5}$	1,1	4.10^{-6}
35	10^{-1}	$3,3.10^{-2}$	1	$1,2.10^{-4}$	1,1	$1,1.10^{-5}$
40	10^{-1}	$1,5.10^{-2}$	1	$4,3.10^{-5}$	1,1	$8,2.10^{-5}$

$N = n \times n$ (mailles), $\phi(r) = r^{\alpha}$, $200^{ème}$ pas de temps		
n	α	RMSE
10	5	$4,1.10^{-3}$
15	5	$8,5.10^{-4}$
20	5	$2,5.10^{-4}$
25	5	$8,7.10^{-5}$
30	5	$3,2.10^{-5}$
35	5	$1,6.10^{-5}$
40	5	$1,5.10^{-5}$

On peut aisément remarquer que dans ce cas, on obtient les meilleurs résultats pour $c = 1,1$. Nous pouvons alors tracer les isovaleurs de la fonction analytique aux différents pas de temps.

4.3.3 Problème instationnaire sans conditions de raccordement

On a le même problème sans tenir compte des conditions de raccordement.

4.4 Résultats numériques relatifs à un problème économique

On Considère l'équation black-Scholes de l'évaluation d'un prix

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V}{\partial x^2} + Rx\frac{\partial V}{\partial x} - RV = 0, \tag{4.29}$$

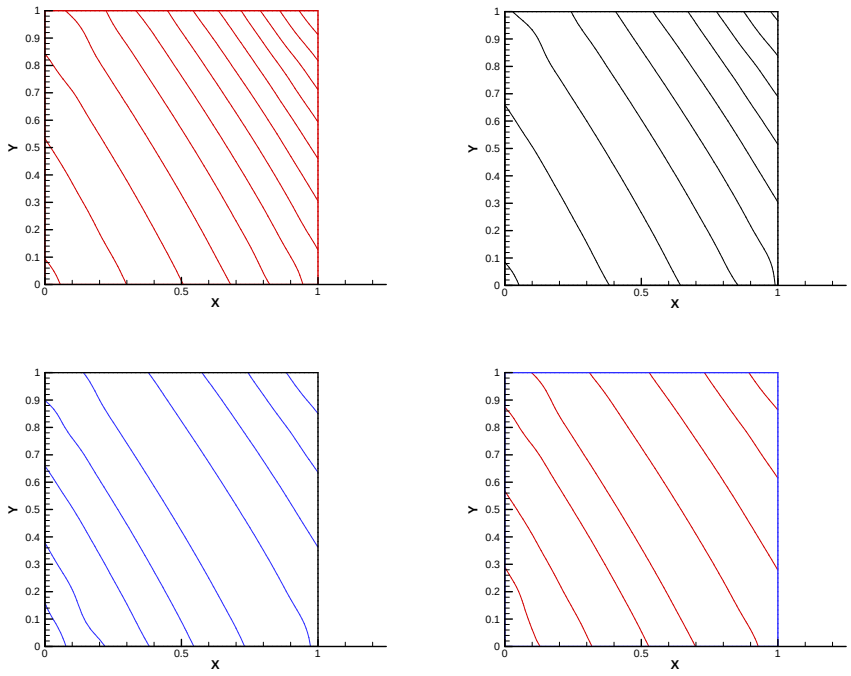

FIG. 4.10 – Problème de la concentration d'un contaminant tenant compte des conditions de raccordement.Isovaleurs de la fonction analytique $n = 0\Delta t$, $n = 50\Delta t$, $n = 70\Delta t$, $n = 202\Delta t$

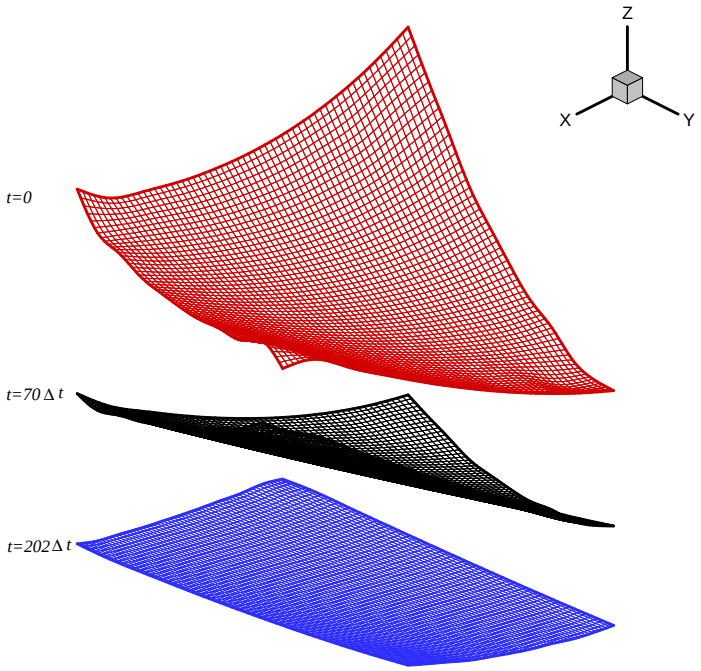

FIG. 4.11 – Problème de la concentration d'un contaminant sans tenir compte des conditions de raccordement ; élévations de la fonction analytique aux pas de temps $n = 0\Delta t$, $n = 70\Delta t$, $n = 202\Delta t$.

où R et le taux d'intérêt sans risques, σ est la volatilité et $V = V(x,t)$ est choix du prix au temps t en fonction du stock x avec $x \in [0, +\infty[$ et $t \in [0,T]$, T étant le temps final de l'expiration du choix. V est en outre soumis aux conditions

$$V(x,T) = \begin{cases} E - x \text{ si } x \leq E \\ 0 \text{ si } x \geq E \end{cases} \tag{4.30}$$

Pour les calculs des prix Européens, on impose les conditions aux limites suivantes :

$$\begin{cases} V(0,t) = Ee^{-R(T-t)} \\ V(S,t) \longrightarrow 0 \ , \ S \longrightarrow +\infty \end{cases} \tag{4.31}$$

La solution exacte de l'équation (4.29) soumis à la condition terminale (4.30) est donnée par Wilmott [46] sous la forme

$$V(x,t) = Ee^{-R(T-t)} Z(-d_2) - xZ(-d_1), \tag{4.32}$$

où $Z(\cdot)$ est la distribution standard cumulative normale, *i.e.*

$$Z(a) = P(0 \leq x \leq |a|) = \int_0^{|a|} f(t)\, dt, \tag{4.33}$$

avec

$$f(t) = \frac{\exp(-t^2/2)}{\sqrt{2\pi}}, \tag{4.34}$$

$$\begin{aligned} d_1 &= \frac{\ln(x/E) + \left(R + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \\ d_2 &= \frac{\ln(x/E) + \left(R - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}. \end{aligned} \tag{4.35}$$

Pour les apllications numériques on prendra :

$$E = 10,\ R = 0,05,\ \sigma = 0,2, \text{ et } T = 0,5 \text{ (année)}. \tag{4.36}$$

Table du RMSE pour la résolution de :

$\dfrac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 x^2 \dfrac{\partial^2 V}{\partial x^2} + Rx\dfrac{\partial V}{\partial x} - RV = 0$
$V(x,T) = \begin{cases} E - x \text{ si } x \leq E \\ 0 \text{ si } x \geq E \end{cases}$
$\begin{cases} V(0,t) = Ee^{-R(T-t)} \\ V(S,t) \longrightarrow 0 \ , \ S \longrightarrow +\infty \end{cases}$

	$\phi(r) = r^5$	$\phi(r) = (r^2 + c^2)^{\frac{3}{2}}$	$\phi(r) = r^4 \ln r$	$\phi(r) = r^3$	$\phi(r)$ =FRSC $\rho = 10^{-3}$
n(noeux)	RMSE	RMSE	RMSE	RMSE	RMSE
16	$2,4.10^{-3}$	$3,7.10^{-3}$	$2,7.10^{-3}$	$3,7.10^{-3}$	2.10^{-2}
31	$6,1.10^{-4}$	$1,1.10^{-3}$	$6,7.10^{-4}$	$1,1.10^{-3}$	$2,6.10^{-2}$
61	$1,6.10^{-4}$	$3,2.10^{-4}$	$1,7.10^{-4}$	$3,2.10^{-4}$	$2,7.10^{-2}$
91	$7,3.10^{-5}$	$1,5.10^{-4}$	$7,5.10^{-5}$	$1,5.10^{-4}$	$2,8.10^{-2}$
121	$4,2.10^{-5}$	$8,4.10^{-5}$	$4,3.10^{-5}$	$8,4.10^{-5}$	$2,8.10^{-2}$
151	$2,7.10^{-5}$	$5,4.10^{-5}$	$2,8.10^{-5}$	$5,4.10^{-5}$	$2,8.10^{-2}$
181	$1,9.10^{-5}$	$3,8.10^{-5}$	2.10^{-5}	$3,8.10^{-5}$	$2,8.10^{-2}$
211	$1,4.10^{-5}$	$2,8.10^{-5}$	$1,4.10^{-5}$	$2,8.10^{-5}$	$2,8.10^{-2}$
241	$1,1.10^{-5}$	$2,2.10^{-5}$	$1,1.10^{-5}$	$2,2.10^{-5}$	$2,8.10^{-2}$
271	9.10^{-6}	$1,7.10^{-5}$	9.10^{-6}	$1,7.10^{-5}$	$2,8.10^{-2}$
301	$7,4.10^{-6}$	$1,4.10^{-5}$	$7,5.10^{-6}$	$1,4.10^{-5}$	$2,8.10^{-2}$

La superposition entre la solution analytique V et la solution numérique V^* est donnée par le graphique ci-dessous.

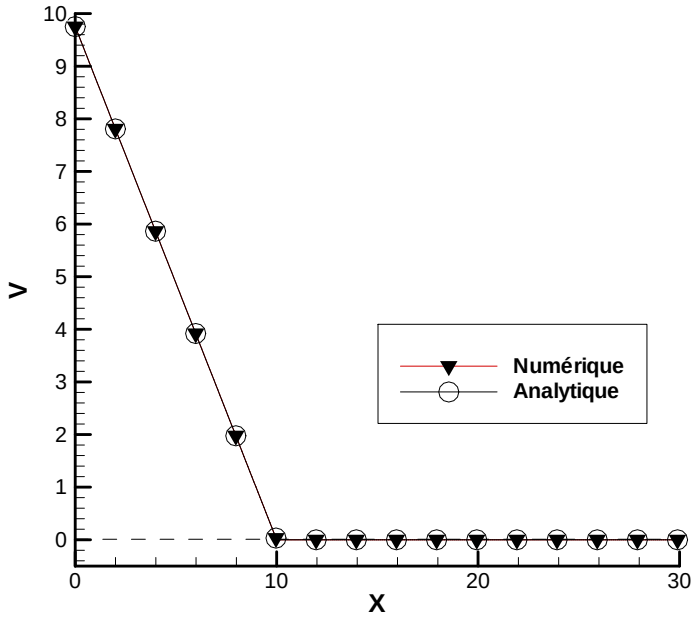

FIG. 4.12 – Problème Economique au pas de temps 200 ($t = 0$). Comparaison solution analytique-solution numérique $\phi(r) = r^5$, $n = 61$

4.5 Résultats numériques relatifs au champ classique d'un méson

Considerons maintenant le problème

$$(\mathcal{P})\begin{cases} \dfrac{\partial^2 u}{\partial t^2} - \Delta u + m^2 u = -\lambda u^3 & \text{sur } \Omega \times]0,T[\ , \\ u(x,0) = u_0(x) & \text{sur } \Omega\ , \\ \dfrac{\partial u}{\partial t}(x,0) = v_0(x) & \text{sur } \Omega\ , \\ u(x,t) = 0 & \text{sur } \partial\Omega \times]0,T[\end{cases} \tag{4.37}$$

où : Ω est un ouvert borné de $\mathbb{R}^N$ $(N \geq 1)$ de frontière $\partial\Omega$ régulière. Pour les applications numériques on prendra $\Omega =]0,1[\times]0,1[$, $\lambda \geq 0$ est un paramètre réel. Dans le cas où $\lambda = 0$ on obtient l'équation de Klein-Gordon décrite dans la mécanique quantique relativiste, et qui se met sous la forme $(L + m^2)\psi = 0$ où l'opérateur $L = \dfrac{\partial^2}{\partial t^2} - \Delta$ et m est la masse de la particule. Le problème (4.37) décrit le champ classique u d'un méson de masse m. Le terme non linéaire $-\lambda u^3$ avec $\lambda > 0$, décrit le self interaction du champ. On peut reformuler (4.37) sous forme d'une équation d'opérateurs

$$\begin{array}{l} u'' + Au = f(u) \text{ sur }]0,T[, \\ u(0) = u_0,\ u'(0) = v_0, \\ \text{avec } f(u) = -\lambda u^3, \\ Au = -\Delta u + m^2 u. \end{array} \tag{4.38}$$

$$\begin{array}{l} u''(t) + Au(t) = f(u(t))\ ,\ 0 < t < \infty \\ u(0) = u_0,\ u'(0) = v_0. \end{array} \tag{4.39}$$

On pose $v = u'$. Au lieu de (4.39), on considère le système du premier ordre

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} 0 & I \\ -A & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ f(u) \end{pmatrix}. \tag{4.40}$$

De plus, on pose $z = (u, v)$ et on écrit (4.39) sous la forme

$$\begin{cases} z'(t) &= Cz(t) + F(z(t))\ ,\ 0 < t < \infty \\ z(0) &= z_0. \end{cases} \tag{4.41}$$

avec

$$C = \begin{pmatrix} 0 & I \\ -A & 0 \end{pmatrix}, F(z) = \begin{pmatrix} 0 \\ f(u) \end{pmatrix}. \tag{4.42}$$

Avant de résoudre le problème $(\mathcal{P})$, rappelons d'abord la méthode de résolution des systèmes non linéaires par l'algorithme de Newton. Nous allons traiter le cas qui nous concerne ; c'est à dire le cas de deux fonctions inconnues. Soit à résoudre un système de deux équations non linéaires, du type :

$$\begin{aligned} F(u,v) &= 0, \\ G(u,v) &= 0, \end{aligned} \tag{4.43}$$

où le couple (u, v) est l'inconnue du système. Ce système peut se mettre sous la forme :

$$\overrightarrow{L}(u,v) = \overrightarrow{0}, \tag{4.44}$$

avec :

$$\overrightarrow{L} = \begin{pmatrix} F \\ G \end{pmatrix}. \tag{4.45}$$

4.5.1 Application à l'équation de Klein-Gordon non linéaire

Soit à résoudre le problème évolutif de l'équation des ondes avec condition initiale $(C.I)$:

$$\begin{cases} \dfrac{\partial \overrightarrow{U}}{\partial t} = C.\overrightarrow{U} + \overrightarrow{L}\left(\overrightarrow{U}\right) \\ \overrightarrow{U}(x,0) = \overrightarrow{U}_0(x) \qquad (C.I) \end{cases} \tag{4.46}$$

où L est un opérateur non linéaire et

$$\overrightarrow{U} = \begin{pmatrix} u \\ v \end{pmatrix}, \text{ avec } v = \frac{\partial u}{\partial t}, \tag{4.47}$$

est le vecteur inconnu. On impose des conditions aux limites $(C.L)$ à chaque instant t. Etant donné un pas de temps Δt, on résout à chaque itérration en temps n par la méthode de Cranck-Nicolson, le problème non linéaire

$$\left\{\begin{array}{c} \dfrac{\overrightarrow{U}^{n+1}-\overrightarrow{U}^{n}}{\Delta t}=\theta\left(C.\overrightarrow{U}^{n+1}+\overrightarrow{L}\left(\overrightarrow{U}^{n+1}\right)\right) \\ +(1-\theta)\left(C.\overrightarrow{U}^{n}+\overrightarrow{L}\left(\overrightarrow{U}^{n}\right)\right) \\ \text{et des } (C.L), \end{array}\right. \tag{4.48}$$

c'est à dire :

$$\overrightarrow{U}^{n+1}-\theta\Delta t\left[C.\overrightarrow{U}^{n+1}+\overrightarrow{L}\left(\overrightarrow{U}^{n+1}\right)\right]=\overrightarrow{U}^{n}+(1-\theta)\,\Delta t\left[\begin{array}{c} C.\overrightarrow{U}^{n} \\ +\overrightarrow{L}\left(\overrightarrow{U}^{n}\right) \end{array}\right], \tag{4.49}$$

où C est la matrice d'opérateurs

$$C=\left(\begin{array}{cc} c_1 & c_2 \\ c_3 & c_4 \end{array}\right). \tag{4.50}$$

On suppose que chaque fonction scalaire inconnue $u, v : \mathbb{R}^N \longrightarrow \mathbb{R}$ peut s'écrire :

$$\left\{\begin{array}{l} u(x)=\displaystyle\sum_{i=0}^{N} u_i\phi_i(x), \\ v(x)=\displaystyle\sum_{i=0}^{N} v_i\phi_i(x). \end{array}\right. \tag{4.51}$$

Les $(\phi_i)_{0\leq i\leq N}$ sont des fonctions radiales données de $\mathbb{R}^n \longrightarrow \mathbb{R}$ (avec $n = N+1$), les inconnues sont les scalaires $(u_i)_{0\leq i\leq N}$ et $(v_i)_{0\leq i\leq N}$. On a donc

$$\overrightarrow{U}(x,t)=\left\{\begin{array}{l} \displaystyle\sum_{i=0}^{N} u_i(t)\,\phi_i(x) \\ \displaystyle\sum_{i=0}^{N} v_i(t)\,\phi_i(x). \end{array}\right. \tag{4.52}$$

On obtient alors :

$$\left[\begin{array}{l} \sum_{i=0}^{N} u_i^{n+1}(t)\,\phi_i(x) \\ -\theta\Delta t\left(c_1\left(\sum_{i=0}^{N} u_i^{n+1}(t)\,\phi_i\right)(x)+c_2\left(\sum_{i=0}^{N} v_i^{n+1}(t)\,\phi_i\right)(x)\right)+L_1\left(\overrightarrow{U}^{n+1}\right) \\ \sum_{i=0}^{N} v_i^{n+1}(t)\,\phi_i(x) \\ -\theta\Delta t\left(c_3\left(\sum_{i=0}^{N} u_i^{n+1}(t)\,\phi_i\right)(x)+c_4\left(\sum_{i=0}^{N} v_i^{n+1}(t)\,\phi_i\right)(x)\right)+L_2\left(\overrightarrow{U}^{n+1}\right) \end{array}\right]-\overrightarrow{S}^{n}(x)=0 \tag{4.53}$$

avec

$$\overrightarrow{S}^{n}(x)=\begin{pmatrix}\overrightarrow{S}_1^{n}(x)\\ \overrightarrow{S}_2^{n}(x)\end{pmatrix}=\overrightarrow{U}^{n}+(1-\theta)\,\Delta t\left[C.\overrightarrow{U}^{n}+\overrightarrow{L}\left(\overrightarrow{U}^{n}\right)\right]. \tag{4.54}$$

D'après la notation $\overrightarrow{L}=\begin{pmatrix}F\\ G\end{pmatrix}$ donnée en (4.45), on a

$$\left\{\begin{array}{l} F(x)=\sum_{j=0}^{N} u_j^{n+1}(t)\,\phi_j(x) \\ -\theta\Delta t\left(c_1\left(\sum_{j=0}^{N} u_j^{n+1}(t)\,\phi_j\right)(x)+c_2\left(\sum_{j=0}^{N} v_j^{n+1}(t)\,\phi_j\right)(x)\right) \\ +L_1\left(\overrightarrow{U}^{n+1}\right)-\overrightarrow{S}_1^{n}(x), \\ G(x)=\sum_{j=0}^{N} v_j^{n+1}(t)\,\phi_j(x) \\ -\theta\Delta t\left(c_3\left(\sum_{j=0}^{N} u_j^{n+1}(t)\,\phi_j\right)(x)+c_4\left(\sum_{j=0}^{N} v_j^{n+1}(t)\,\phi_j\right)(x)\right) \\ +L_2\left(\overrightarrow{U}^{n+1}\right)-\overrightarrow{S}_2^{n}(x), \end{array}\right. \tag{4.55}$$

donc :

$$\left\{\begin{array}{l} F_i = F\left(x_i\right) = \sum_{j=0}^{N} u_j^{n+1}\left(t\right)\phi_j\left(x_i\right) \\ -\theta\Delta t\left(c_1\left(\sum_{j=0}^{N} u_j^{n+1}\left(t\right)\phi_j\right)\left(x_i\right) + c_2\left(\sum_{j=0}^{N} v_j^{n+1}\left(t\right)\phi_j\right)\left(x_i\right)\right) \\ +L_1\left(\overrightarrow{U}^{n+1}\right)\left(x_i\right) - \overrightarrow{S}_1^{\,n}\left(x_i\right), \\ G_i = G\left(x_i\right) = \sum_{j=0}^{N} v_j^{n+1}\left(t\right)\phi_j\left(x_i\right) \\ -\theta\Delta t\left(c_3\left(\sum_{j=0}^{N} u_j^{n+1}\left(t\right)\phi_j\right)\left(x_i\right) + c_4\left(\sum_{j=0}^{N} v_j^{n+1}\left(t\right)\phi_j\right)\left(x_i\right)\right) \\ +L_2\left(\overrightarrow{U}^{n+1}\right)\left(x_i\right) - \overrightarrow{S}_2^{\,n}\left(x_i\right), \end{array}\right. \tag{4.56}$$

d'où

$$\left\{\begin{array}{l} \dfrac{\partial F_i}{\partial u_j} = \phi_j\left(x_i\right) - \theta\Delta t\cdot c_1\left[\phi_j\right]\left(x_i\right) + \dfrac{\partial L_1}{\partial u_j}\left(x_i\right) \\ \dfrac{\partial G_i}{\partial u_j} = -\theta\Delta t\cdot c_3\left[\phi_j\right]\left(x_i\right) + \dfrac{\partial L_2}{\partial u_j}\left(x_i\right) \end{array}\right. \tag{4.57}$$

$$\left\{\begin{array}{l} \dfrac{\partial F_i}{\partial v_j} = -\theta\Delta t\cdot c_2\left[\phi_j\right]\left(x_i\right) + \dfrac{\partial L_1}{\partial v_j}\left(x_i\right) \\ \dfrac{\partial G_i}{\partial v_j} = \phi_j\left(x_i\right) - \theta\Delta t\cdot c_4\left[\phi_j\right]\left(x_i\right) + \dfrac{\partial L_2}{\partial v_j}\left(x_i\right). \end{array}\right. \tag{4.58}$$

Nous avons

$$\left\{\begin{array}{c} L_1 = 0 \\ L_2\left(\overrightarrow{U}\right) = f\left(u\right) = -\lambda u^3 \end{array}\right. \tag{4.59}$$

$$\left\{\begin{array}{ll} c_1 = 0 & c_2 = Id \\ c_3 = -A & c_4 = 0 \end{array}\right. \tag{4.60}$$

avec

$$A\left(u\right) = -\Delta u + m^2 u. \tag{4.61}$$

$\overrightarrow{\alpha}_0$ étant un vecteur "arbitrairement donné", on fabrique la suite :

$$\overrightarrow{\alpha}_{s+1} = \overrightarrow{\alpha}_s - J^{-1}.\overrightarrow{F}\left(\overrightarrow{\alpha}_s\right), \tag{4.62}$$

ce qui équivaut à

$$J.\left(\overrightarrow{\alpha}_{s+1}-\overrightarrow{\alpha}_{s}\right)=-\overrightarrow{F}\left(\overrightarrow{\alpha}_{s}\right), \tag{4.63}$$

c'est à dire à résoudre le système

$$\begin{cases} J.\left(\overrightarrow{Z}_{s}\right)=-\overrightarrow{F}\left(\overrightarrow{\alpha}_{s}\right) \\ \overrightarrow{\alpha}_{s+1}=\overrightarrow{\alpha}_{s}+\overrightarrow{Z}_{s}, \end{cases} \tag{4.64}$$

où J est la matrice jacobienne :

$$J=\begin{bmatrix} \frac{\partial F_0}{\partial u_0} & \frac{\partial F_0}{\partial u_1} & \cdots & \frac{\partial F_0}{\partial u_j} & \cdots & \frac{\partial F_0}{\partial u_N} & \frac{\partial F_0}{\partial v_0} & \frac{\partial F_0}{\partial v_1} & \cdots & \frac{\partial F_0}{\partial v_j} & \cdots & \frac{\partial F_0}{\partial v_N} \\ \frac{\partial F_1}{\partial u_0} & \frac{\partial F_1}{\partial u_1} & \cdots & \frac{\partial F_1}{\partial u_j} & \cdots & \frac{\partial F_1}{\partial v_0} & \frac{\partial F_1}{\partial v_1} & \frac{\partial F_1}{\partial v_1} & \cdots & \frac{\partial F_1}{\partial v_j} & \cdots & \frac{\partial F_1}{\partial v_N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial F_i}{\partial u_0} & \frac{\partial F_i}{\partial u_1} & \cdots & \frac{\partial F_i}{\partial u_j} & \cdots & \frac{\partial F_i}{\partial u_N} & \frac{\partial F_i}{\partial v_0} & \frac{\partial F_i}{\partial v_1} & \cdots & \frac{\partial F_i}{\partial v_j} & \cdots & \frac{\partial F_i}{\partial v_N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial F_N}{\partial u_0} & \frac{\partial F_N}{\partial u_1} & \cdots & \frac{\partial F_N}{\partial u_j} & \cdots & \frac{\partial F_N}{\partial u_N} & \frac{\partial F_N}{\partial v_0} & \frac{\partial F_N}{\partial v_1} & \cdots & \frac{\partial F_N}{\partial v_j} & \cdots & \frac{\partial F_N}{\partial v_N} \\ \frac{\partial G_0}{\partial u_0} & \frac{\partial G_0}{\partial u_1} & \cdots & \frac{\partial G_0}{\partial u_j} & \cdots & \frac{\partial G_0}{\partial u_N} & \frac{\partial G_0}{\partial v_0} & \frac{\partial G_0}{\partial v_1} & \cdots & \frac{\partial G_0}{\partial v_j} & \cdots & \frac{\partial G_0}{\partial v_N} \\ \frac{\partial G_1}{\partial u_0} & \frac{\partial G_1}{\partial u_1} & \cdots & \frac{\partial G_1}{\partial u_j} & \cdots & \frac{\partial G_1}{\partial u_N} & \frac{\partial G_1}{\partial v_0} & \frac{\partial G_1}{\partial v_1} & \cdots & \frac{\partial G_1}{\partial v_j} & \cdots & \frac{\partial G_1}{\partial v_N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial G_i}{\partial u_0} & \frac{\partial G_i}{\partial u_1} & \cdots & \frac{\partial G_i}{\partial u_j} & \cdots & \frac{\partial G_i}{\partial u_N} & \frac{\partial G_i}{\partial v_0} & \frac{\partial G_i}{\partial v_1} & \cdots & \frac{\partial G_i}{\partial v_j} & \cdots & \frac{\partial G_i}{\partial v_N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial G_N}{\partial u_0} & \frac{\partial G_N}{\partial u_1} & \cdots & \frac{\partial G_N}{\partial u_j} & \cdots & \frac{\partial G_N}{\partial u_N} & \frac{\partial G_N}{\partial v_0} & \frac{\partial G_N}{\partial v_1} & \cdots & \frac{\partial G_N}{\partial v_j} & \cdots & \frac{\partial G_N}{\partial v_N} \end{bmatrix}$$

4.5.2 Problème Stationnaire avec $\lambda = 0$

$$\begin{array}{|l|} \hline -\Delta u+m^2u=0 \text{ , sur }]0,1[\times]0,1[, \\ \hline u\left(0,y\right)=u\left(1,y\right)=f\left(y\right), \\ \hline u\left(x,0\right)=u\left(x,1\right)=f\left(x\right), \\ \hline \end{array} \tag{4.65}$$

où

$$f(0) = f(1) = 1. \tag{4.66}$$

On peut déterminer f telle que

$$u(x,y) = f(x) f(y), \tag{4.67}$$

soit la solution exacte de (4.65). On trouve,

$$f(x) = \frac{\left(1 - \exp\left(\frac{-m}{\sqrt{2}}\right)\right) \exp\left(\frac{m}{\sqrt{2}}x\right) + \left(\exp\left(\frac{m}{\sqrt{2}}\right) - 1\right) \exp\left(\frac{-m}{\sqrt{2}}x\right)}{\exp\left(\frac{m}{\sqrt{2}}\right) - \exp\left(\frac{-m}{\sqrt{2}}\right)}. \tag{4.68}$$

4.5.3 Problème stationnaire avec $\lambda > 0$

Le problème stationnaire s'écrit donc

$$\begin{aligned} -\Delta u + m^2 u &= -\lambda u^3 \text{ sur } \Omega, \\ u_{|\Gamma} &= u_0. \end{aligned} \tag{4.69}$$

On cherche une solution approchée u^* de u telle que

$$u^*(x,y) = \sum_{j=1}^{N} \lambda_j \phi_j(x,y). \tag{4.70}$$

On a

$$\Delta u^*(x,y) = \sum_{j=1}^{N} \lambda_j \Delta\phi_j(x,y). \tag{4.71}$$

Soit L l'opérateur défini par

$$L(u) = -\Delta u + m^2 u + \lambda u^3. \tag{4.72}$$

On résout

$$L(u^*)(X_i) = 0, \text{ où } X_i = (x_i, y_i),\ 1 \le i \le N, \tag{4.73}$$

soit

$$\begin{aligned} F_i\left(\overrightarrow{\alpha}\right) &= -\sum_{j=1}^{N}\lambda_j\Delta\phi_j\left(X_i\right)+m^2\sum_{j=1}^{N}\lambda_j\phi_j\left(X_i\right) \quad (4.74)\\ +\lambda\left[\sum_{j=1}^{N}\lambda_j\phi_j\left(X_i\right)\right]^3 &= 0, \quad (4.75)\\ i.e.,\ \overrightarrow{F}\left(\overrightarrow{\alpha}\right) &= 0, \text{ avec } \overrightarrow{F}=\left(F_1,F_2,...,F_N\right), \quad (4.76)\\ \overrightarrow{\alpha} &= \begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_N\end{pmatrix} \quad (4.77)\\ \text{et } F_i &= L\left(u^*\right)\left(X_i\right),\ 1\leq i\leq N. \quad (4.78)\end{aligned}$$

Soit $\overrightarrow{\alpha^0}$ donné et $\left(\overrightarrow{\alpha^s}\right)_{s\in\mathcal{N}}$ une suite telle que

$$\begin{aligned} &\overrightarrow{\alpha^s}\longrightarrow\overrightarrow{\alpha},\ s\longrightarrow+\infty,\\ &\overrightarrow{\alpha^{s+1}}=\overrightarrow{\alpha^s}-\left[J_s^{-1}\right].\overrightarrow{F}\left(\overrightarrow{\alpha^s}\right),\\ &J_s.\overrightarrow{\alpha^{s+1}}=J_s.\overrightarrow{\alpha^s}-\overrightarrow{F}\left(\overrightarrow{\alpha^s}\right).\end{aligned} \quad (4.79)$$

En posant

$$\overrightarrow{Z^s}=\overrightarrow{\alpha^{s+1}}-\overrightarrow{\alpha^s}, \quad (4.80)$$

on résoud

$$J_s.\overrightarrow{Z^s}=-\overrightarrow{F}\left(\overrightarrow{\alpha^s}\right), \text{ et donc, } \overrightarrow{\alpha^{s+1}}=\overrightarrow{\alpha^s}+\overrightarrow{Z^s}. \quad (4.81)$$

$$\begin{aligned} &J=J_s=\left(\tfrac{\partial F_i}{\partial\lambda_j}\right)_{1\leq i,j\leq N}\\ &F_i\left(\overrightarrow{\alpha^s}\right)=\sum_{k=1}^{N}\lambda_k^{(s)}\Delta\phi_k\left(X_i\right)-m^2\sum_{j=1}^{N}\lambda_k^{(s)}\phi_k\left(X_i\right)+\lambda\left(\sum_{j=1}^{N}\lambda_k^{(s)}\phi_k\left(X_i\right)\right)^3\\ &J_{ij}=\left(\tfrac{\partial F_i}{\partial\lambda_j}\right)=\Delta\phi_j\left(X_i\right)-m^2\sum_{j=1}^{N}\phi_j\left(X_i\right)-3\lambda\left(\sum_{j=1}^{N}\lambda_k^{(s)}\phi_k\left(X_i\right)\right)^2.\phi_j\left(X_i\right)\\ &=\Delta\phi_j\left(X_i\right)+m^2\left[\sum_{j=1}^{N}\phi_j\left(X_i\right)+3\lambda\left(\sum_{j=1}^{N}\lambda_k^{(s)}\phi_k\left(X_i\right)\right)^2\right].\phi_j\left(X_i\right)\\ &=-F_i\left(\overrightarrow{\alpha}\right)\end{aligned} \quad (4.82)$$

4.5.4 Problème instationnaire avec $\lambda = 0$

$$(\mathcal{P}) \begin{cases} \dfrac{\partial^2 u}{\partial t^2} - \Delta u + m^2 u = 0 \text{ , sur }]0,1[\times]0,1[\times]0,T[\\ u(X,0) = u_0(X) \text{ , sur}]0,1[\times]0,1[\text{ ,} \\ \dfrac{\partial u}{\partial t}(X,0) = v_0(X) \text{ , sur}]0,1[\times]0,1[, \\ u(X,t) = 0, \text{ sur } \partial\Omega \times]0,T[. \end{cases} \tag{4.83}$$

Solution exacte

$$u(x,y,t) = \cos \omega t \sin \pi x \sin \pi y, \tag{4.84}$$

avec

$$\begin{array}{l} \omega^2 = 2\pi^2 + m^2, \\ u_0(x,y) = \sin \pi x \sin \pi y, \\ v_0(x,y) = 0. \end{array} \tag{4.85}$$

Table du RMSE pour la résolution de :

$$-\Delta u + m^2 u = 0,$$

avec les conditions aux limites de (4.65)

$\phi(r) = r^5.$

n	m	RMSE
10	10^{-2}	$2,8.10^{-5}$
10	1	$1,8.10^{-2}$
10	10^{-1}	2.10^{-4}
15	10^{-2}	9.10^{-6}
20	1	$1,9.10^{-2}$
20	10^{-2}	$4,9.10^{-6}$
20	10^{-1}	2.10^{-4}

Table du RMSE pour la résolution de : $\dfrac{\partial^2 u}{\partial t^2} - \Delta u + m^2 u = 0$, avec les conditions aux limites de (4.83)

$\phi(r) = r^5$, $\omega = \sqrt{2\pi^2 + m^2}$ $138^{\text{ème}}$ pas de temps $m = 1.$

n	Δt	RMSE
10	10^{-2}	$1,27.10^{-3}$
15	10^{-2}	$3,7.10^{-4}$
20	10^{-2}	$1,4.10^{-4}$

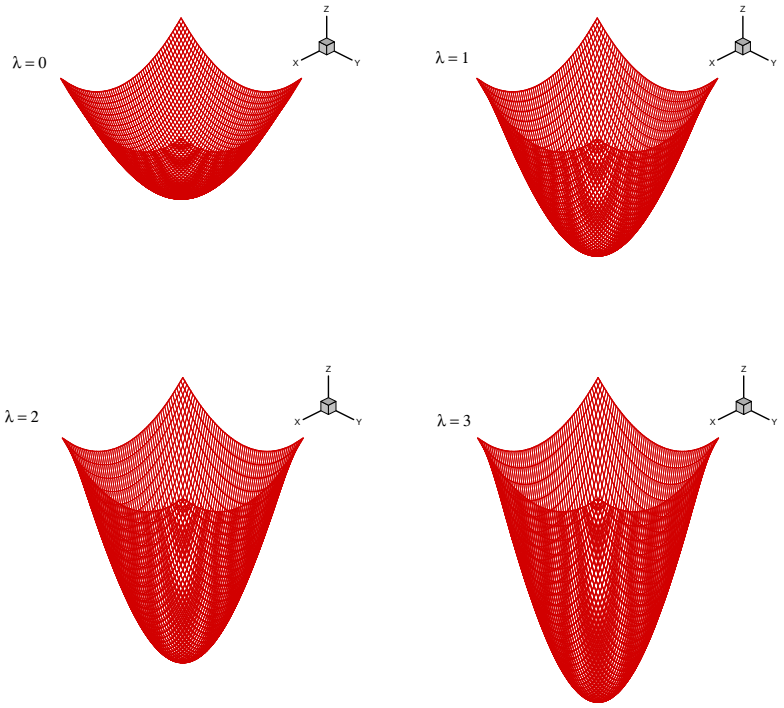

FIG. 4.13 – Solution numérique de l'équation de Klein-Gordon stationnaire.

on a fait varier Δt de 10^{-2} à 10^{-3} et aucun changement significatif sur le RMSE ne s'est produit avec le reste des paramètres inchangés ; ça ne sert donc à rien de faire varier le temps. Le $102^{\text{ème}}$ pas de temps correspond à $t = 102 \times 10^{-2}$.

Nous allons maintenant représenter les différentes fonctions correspondantes à la solution de l'équation de Klein-Gordon dans les différents cas étudiés.

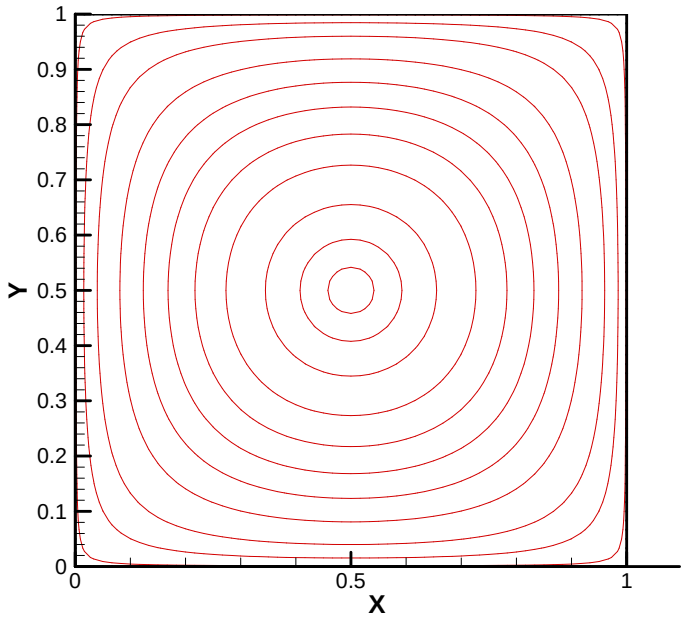

FIG. 4.14 – Solution numérique de l'équation de Klein -Gordon instationnaire linéaire. $3^{ème}$ pas de temps $m = 1$, $\theta = \frac{1}{2}$, $\Delta t = 10^{-3}$.

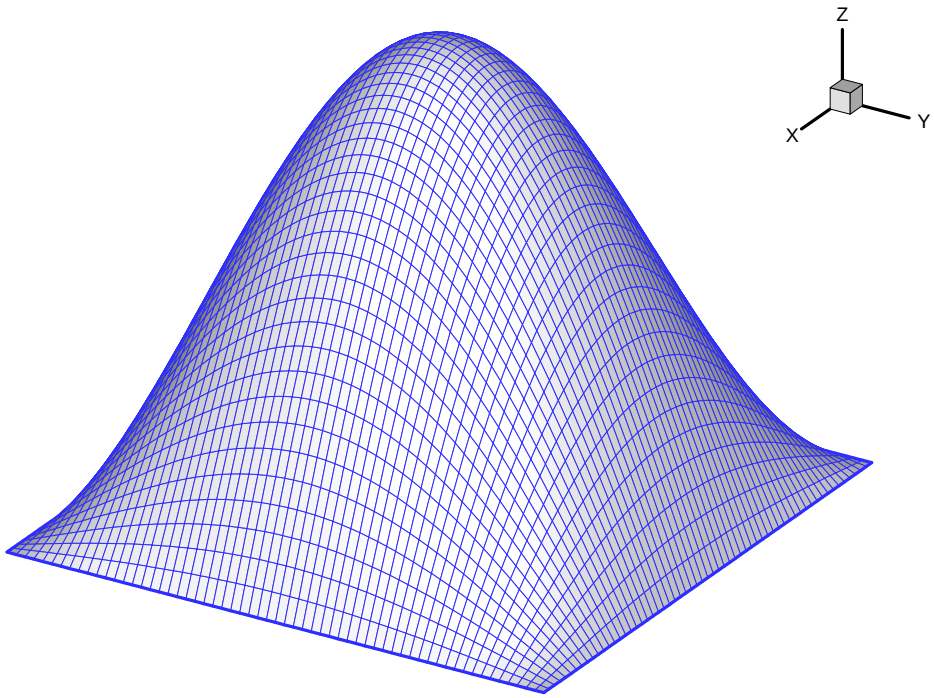

FIG. 4.15 – Solution numérique de l'équation de Klein -Gordon au $3^{ème}$ pas de temps. $\lambda = 0$, $\Delta t = 10^{-3}$, $m = 1$.

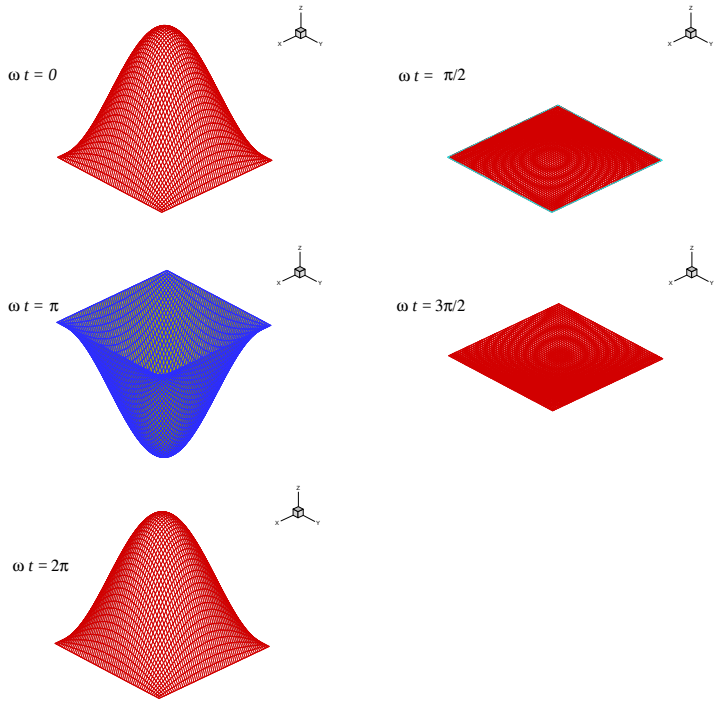

FIG. 4.16 – Solution numérique de l'équation de Klein -Gordon instationnaire et linéaire. $m = 1$, $\theta = \frac{1}{2}$, $\Delta t = 10^{-2}$.

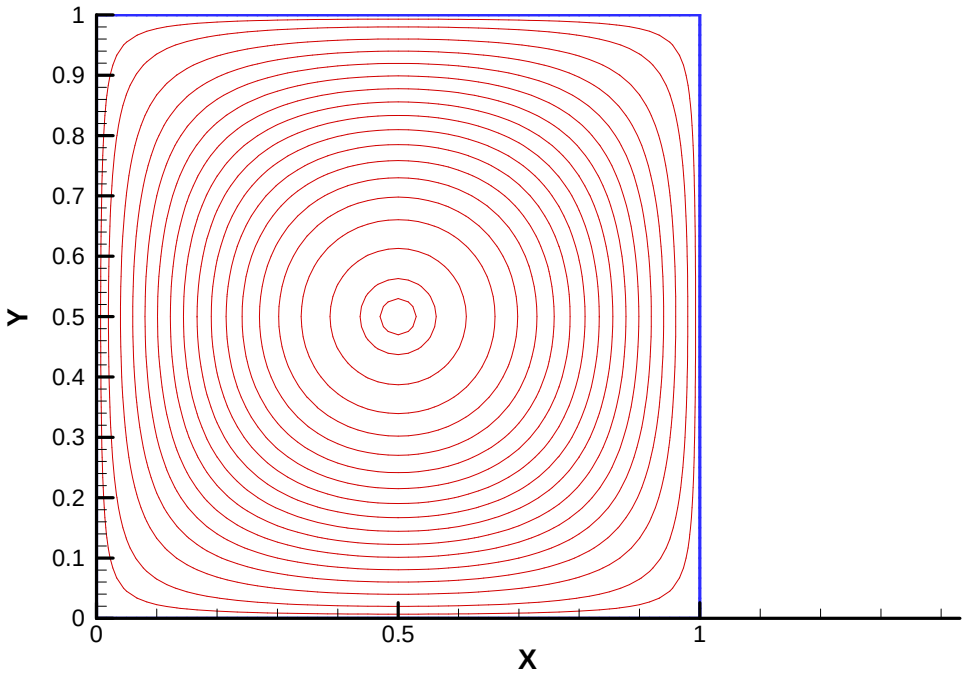

FIG. 4.17 – Solution numérique de l'équation de Klein -Gordon au $3^{ème}$ pas de temps. $\lambda = 1$, $\Delta t = 10^{-3}$, $m = 1$.

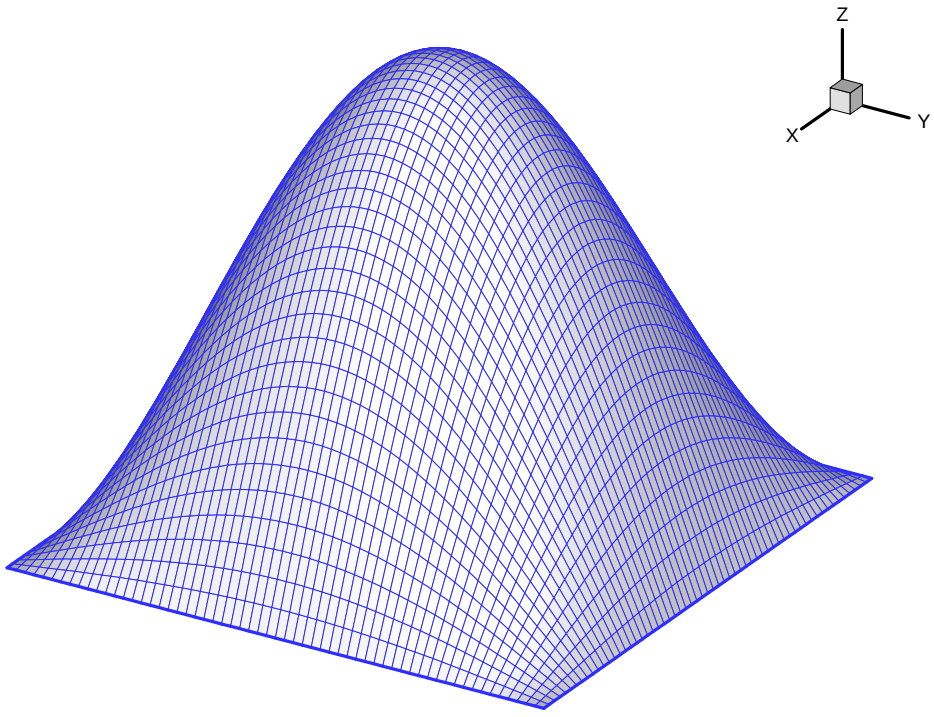

FIG. 4.18 – Solution numérique de l'équation de Klein -Gordon au $3^{ème}$ pas de temps. $\lambda = 1$, $\Delta t = 10^{-3}$, $m = 1$.

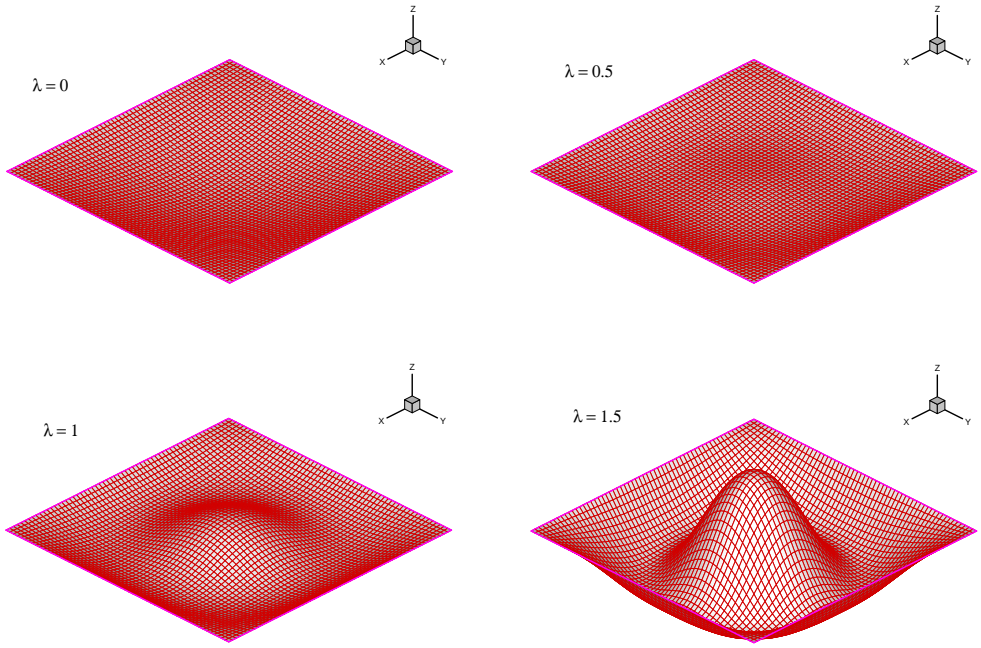

FIG. 4.19 – Solution numérique de l'équation de Klein-Gordon au $102^{ème}$ pas de temps. $m = 1$

Conclusion et perspectives

Le but principal de ce travail a été de tester l'éfficacité de la Méthode de collocation RBF utilisant l'interpolation par les fonctions radiales de base pour résoudre certaines équations aux dérivées partielles. Une telle étude étant motivée par la comparaison au moyen du RMSE entre la solution analytique et la solution numérique selon la méthode, l'aspect traditionnel des EDP beaucoup plus pointilleux sur le cadre fonctionnel de l'étude puis sur des nombreux résultats théoriques abordant la question d'un problème bien posé, a été pour la plus grande partie du travail moins abordé sauf dans le cas du problème non linéaire du champ classique d'un méson. Dès lors, le second membre de nos EDP a été une fonction suffisament régulière. Les problèmes abordés dans ce travail ayant donc à priori une solution analytique (non évidente du point de vue constructive) pour la plupart des cas, nous nous sommes préoccupés de trouver la solution analytique qui vérifie les conditions aux limites et plus particulièrement les conditions de raccordement car la précision de la méthode a été d'autant plus meilleure pour une telle solution que pour une solution analytique qui ne satisfaisait pas les conditions de raccordement. Cette hypothèse s'est confirmée lors du tracé des élévations aux différents pas de temps de la solution analytique instationnaire du problème de la concentration d'un contaminant. Une contribution considérable a été apportée dans ce travail pour la résolution de l'équation non linéaire de Klein-Gordon utilisant la méthode de collocation RBF, le schéma numérique temporel de Crank-Nicolson puis l'algorithme itératif de Newton. Au regard des différents tableaux du RMSE donnés et les différentes représentations graphiques éffectuées de la solution, nous pouvons dans l'ensemble affirmer l'éfficacité et la convergence de la méthode jusqu'à un seuil tolérable du nombre de noeuds d'interpolation. Cependant, soulignons tout de même que la méthode commence à perdre de l'éfficacité lorsque le nombre de noeud d'interpolation devient très élévé, pour preuve en 2D par exemple un nombre de noeuds $n = 600$ disposés sur $[0, 1]$ engendre $N = n^2 = 360000$ noeuds du domaine $\Omega = [0, 1] \times [0, 1]$. Par ailleurs, étant limité par la capacité de la machine, on ne peut se permettre un libre choix d'un nombre de noeuds trop élévé. Enfin, une investigation plus générale de l'étude que nous venons de faire serait par exemple d'approfondir la méthode dans le cas plus général où

le second membre de l'équation aux dérivées partielles serait une distribution des espaces de Sobolev $W^{m,p}(\Omega)$ classiques, car cela permettrait par exemple une comparaison des résultats de la méthode de collocation RBF avec des résultats qu'on obtiendrait via la méthode des éléments finis dans $W^{m,p}(\Omega)$. Le problème reste donc ouvert à ce niveau bien que de nombreux auteurs ont comparé les deux méthodes pour $f \in C^2(\Omega)$ par exemple.

Bibliographie

[1] ABRAMOWITH M and STEGUN I. A : Handbook of Mathematical Functions, Dover Publications, Inc, New York, 1965.

[2] BEJANCU A. : The uniform convergence of multivariate natural splines. Department of Applied Mathematics and Theoretical Physics. Silver Street, Cambridge, England CB3 9EW, 1997.

[3] BEJANCU A. : Local accuracy for radial basis function interpolation on finite uniform grids. J. Approx. Theory 99 (2) 242–257 (1999).

[4] BREZIS H : Analyse Fonctionelle-Theorie et Applications, Masson, Paris, 1983.

[5] CIARLET P.G. : The finite element method for elliptic problems. North Holland Publishing Compagny Amsterdam-New York-Oxford, 1978.

[6] DAUTRAY R and LIONS J-L. : Mathematical Analysis and Numerical Methods for Science and Technology, Vol.1 Physical Origins and Classical Methods. Springer-Verlag, 1990.

[7] DE BOOR C. : A practical guide to splines, Springer-Verlag, New York Heidelberg-Berlin,1978.

[8] DUCHON J. : Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In W. Schempp and K. Zeller, editors, Constructive Theory of Functions of Several Variables, pages 85-100. Berlin : Springer-Verlag, 1977.

[9] DUCHON J : Sur l'erreur d'interpolation des fonctions de plusieurs variables par les D -splines, R.A.I.R.O. An. Num. 12, no. 4, 325–334, 1978.

[10] FARIN G.E. : Curves and surfaces for computer-aided geometric design : a practical guide, Academic Press, San Diego, USA, 1997.

[11] FRANKE R, SCHABACK. R. : Convergence order estimates of meshless collocation methods using radial basis functions, Adv. Comput.Math., **8,** 381-399, 1998.

[12] FASSHAUER G.E. : On the numerical solution of differential equations with radial basis functions. Department of applied Mathematics, Illinois institute of Technology, Chicago, IL 60616, U.S.A. Preprint, to appear in Boundary Element Technology XIII.

[13] FEDOSEYEV, FRIEDMAN A.I.and KANSA, E.J. : Improved multiquadratic method for elliptic partial differential equations via PDE collocation on the boundary. Comput. Math. Appl., (in press), 2000.

[14] FOLEY T.A. and HAGEN H. : Advances in scattered data interpolation, Surv. Math. Ind. 4 , 71-84, 1994.

[15] FRANKE C and SCHABACK R : Solving Partial Differential Equations by Collocation using Radial Basis Functions, Applied Mathematics and Computation 93 , no. 1, 73-82, 1998.

[16] FRANKE C and SCHABACK R : Convergence Orders of Meshless Collocation Methods using Radial Basis Functions, Advances in Computational Mathematics 8 , no. 4, 381-399, 1998.

[17] FRANKE R, Scattered data interpolation : tests of some methods, Math. Cornput. 48 181-200, 1982.

[18] GOLOMB M., WEINBERGER H. F. : Optimal approximation and error bounds, On numerical approximation, R. E. Langer ed., The University of Wisconsin Press, Maddison, 117-190, 1959.

[19] GOUT J.L. : Elements finis polygonaux de Wachspress. Thèse, Université de Pau et des pays de l'Adour, 1980.

[20] GUESSAB A. : Sur les formules de quadrature numérique à nombre minimal de noeux dans un domaine de $\mathbb{R}^d$. Thèse, Université de Pau et des pays de l'Adour, 1987

[21] HON Y.C. : A quasi-radial basis functions method for American options pricing, Comput. Math. Applic. 43 513–524, 2002.

[22] HON Y.C.,WU.Z. : A quasi- interpolation method for solving stiff ordinary differential equations, Int. J. Numer. Meth. Engng., 48 1187-1197, 2000.

[23] HON Y.C., X.Z. : A comparison on using various radial basis function for options pricing. Department of Mathematics, City University of Hong Kong, Kowloon Tong, Hong Kong.

[24] HON Y.C, X.Z. MAO : A multiquadric interpolation method for solving initial value problems, Sci. comput., **12**(1), 51-55, 1997.

[25] HON Y.C., M.W. LU, W.M. XUE, Y.M. ZHU, Multiquadric method for the numerical solution of a biphasic model, Appl. Math. Comput., **88,** 153-175, 1997.

[26] HON Y.C, X.Z. MAO : A radial basis function method for solving options pricing model, Financial Engeneering, **8**(1), 31-49, 1999.

[27] JOHNSON M.J. : Overcoming the boundary effects in surface spline interpolation, IMA J. Numer. Anal. (to appear).

[28] KANSA E.J. : Multiquadrics- A scattered data approximation scheme with applications to computational fluid dynamics I. Surface approximations and partial derivative estimates, comput. Math. Appl., **19**(6-8) 127-145, 1990.

[29] KANSA E.J. : Multiquadrics- A scattered data approximation scheme with applications to computational fluid dynamics II. Solutions to parabolic, hyperbolic, and elliptic differential equations, comput. Math. Appl., **19**(6-8) 147-161, 1990.

[30] KANSA E.J, PH.D. : Motivation for using radial basis functions to solve PDEs. Lawrence Livermore National Laboratory and Embry-Riddle Aeronatical University, 1999.

[31] LI J, CHEN.C.S, PEPPER.D, CHEN.Y, : Mesh-free method for grounwater modeling.

[32] LIGHT W. : Variational Methods for interpolation, particularly by radial basis functions, Technical Report, Dept. of Math. and Comp.Sc., University of Leicester, 1996.

[33] LIGHT W, WAYNE.H. : Error estimates for approximation by radial basis functions, Technical Report, Dept. of Math. and Comp.Sc., University of Leicester, 1996.

[34] LIGHT W, WAYNE.H. : Some remarks on power functions and error estimates for radial basis function interpolation, Technical Report, Dept. of Math. and Comp.Sc., University of Leicester,1996.

[35] LIONS J.L., MAGENES E. : Non-Homogeneous Boundary Value Problems and Applications I,II, Springer-Verlag Berlin Heidelberg New York 1973.

[36] MADYCH W.R and NELSON S.A. : Multivariate interpolation and conditionally positivedefinite functions. II. Mathematics of Computation, 54(189) :211–230, January 1990.

[37] MEINGUET J. : Multivariate interpolation at arbitrary points made simple, Z. Angew. Math. Phys., **30**, 292-304, 1979.

[38] MICCHELLI C. A : Interpolation of scattered data : distance matrices and conditionally positive definite functions. constr. Approx. **2** (1)11-22, 1986.

[39] POWELL M.J.D. : The uniform convergence of thin plate spline interpolation in two dimensions.Numer. Math. 107-128, 1994

[40] POWELL M.J.D. : The theory of radial basis function approximation in 1990. Advances in numerical Analysis, vol.II, ed. W.Light, Oxford science publications Oxford, 105-210, 1992.

[41] POWELL M.J.D. : Radial basis functions for multivariable interpolation : a review, Numerical Analysis, D.F. Griffiths and G. A. Watson (eds), Longman Scientific & Technical (Harlow), 223-241, 1987.

[42] QUARTERONI A, VALLI.A. : Numerical Approximation of Partial Differential Equations, Berlin Heidelberg : Springer-Verlag, 1994.

[43] RAVIART P.A, THOMAS J.M. : Introduction à l'analyse numérique des équations aux dérivées partielles. Masson, 1983.

[44] SCHABACK R, WENDLAND H. : Characterization and construction of radial basis functions, in N. Dyn, D. Leviatan, D. Levin, and A. Pinkus, eds., Multivariate approximation and applications, Cambridge Univ. Press, Cambridge , 1–24, 2001.

[45] WENDLAND H : Sobolev-type error estimates for interpolation by radial basis functions. In : A. Le Mehaute, C. Rabut, L.L. Schumaker, eds., Surface Fitting and Multiresolution Methods, Vanderbilt Univ. Press, pp. 337–344 1977

[46] WILMOT T P, HOWISON S. and DEWYNE J : The Mathematics of Financial Derivatives. New York, Cambridge University Press, 1995.

[47] WONG S.M., HON Y.C., LI T.S., CHUNG S.L., KANSA E.J. : Multizone decomposition for simulation of time-dependent problems using multiquadric scheme , Comput. Math. Applic., **37** (8), 23-43, 1999.

[48] WU Z., SCHABACK R. : Local error estimates for radial basis function interpolationof scattered data, IMA J. Numer.Anal., **13**, 13-27, 1993

[49] WU Z., SCHABACK R. : Shape preserving properties and convergence of univariate multiquadric quasi-interpolation, Acta Math. Appl. Sin.**10(**4**),** 441-446, 1994.

[50] ZEIDLER E. : Nonlinear Functional Analysis and its Applications II/A : Linear Monotone Operators, Springer, New York, 1990.