Je remercie vivement Madame Kulkarni de m'avoir fait l'honneur de juger mon travail avec attention et d'avoir fait un si long voyage pour être présente au jury de cette thèse.

Je suis très heureux de compter Madame Dias D'Almeida parmi mes rapporteurs et je la remercie de s'être déplacée pour assister au jury de cette thèse. Je la remercie de son accueil et de l'intérêt qu'elle a porté à mon travail.

Messieurs Panasenko et Amosov m'ont permis d'élargir le champ de mes connaissances et de recherche. Je remercie particulièrement Monsieur Amosov pour le temps et le travail qu'il a accordé à l'évaluation de ce manuscrit et Monsieur Panasenko de présider le jury.

Je tiens à exprimer mes sincères remerciements à Monsieur Rutily pour avoir été à l'origine de ces recherches. Je le remercie particulièrement de son aide précieuse et de son accueil chaleureux.

Il m'est difficile de devoir résumer en quelques lignes la gratitude et les sentiments que j'éprouve envers Messieurs Ahues et Largillier. Je leur suis infiniment reconnaissant du sérieux dont ils ont fait preuve dans l'encadrement de mon travail, et du soutient sans faille qu'ils m'ont apporté au long de ce parcours.

Je n'aurais peut-être jamais eu le désir de faire des études si mes parents ne m'y avaient pas montré l'intérêt. Je les remercie à la fois d'avoir éveillé en moi l'envie du savoir, et de mettre tout en œuvre pour que je le satisfasse.

Durant ces années, j'ai eu la chance et le plaisir de pouvoir compter sur mes grands parents pour m'assurer, de manière hebdomadaire, des repas pantagruéliques, contenant les acides gras et autres protéines essentiels à la bonne marche de l'esprit.

Mes plus douces pensées vont tout naturellement à Edith qui à su faire preuve d'une patience inouïe pendant cette période difficile.

Je tiens enfin à honorer la mémoire de Monsieur Carasso, dont j'ai eu la chance d'être un de ses étudiants. Je lui suis entièrement reconnaissant de m'avoir fait découvrir l'Analyse Numérique.

Table des matières

Introduction

Ce mémoire traite de la résolution numérique des équations de Fredholm de seconde espèce faiblement singulières, posées dans un espace de Banach. Les méthodes décrites ici sont appliquées plus particulièrement dans le cas de l'espace des fonctions continues sur un intervalle compact et dans le cas de l'espace des fonctions intégrables au sens de Lebesgue sur un intervalle compact. Ce type d'équation intervient dans la résolution du problème de transfert en physique des particules et plus particulièrement dans la modélisation des atmosphères stellaires, motivation première de notre recherche.

Le premier chapitre fixe brièvement le cadre théorique de notre étude. Nous y rappelons notamment différents types de convergence d'une suite d'opérateurs dans un espace de Banach complexe, ainsi que leur propriétés.

Le deuxième chapitre est consacré à la description et à l'analyse de deux méthodes d'approximation de rang fini sur lesquelles sont appliqués trois schémas de raffinement itératif. Nous donnons des majorations pour l'erreur relative associée à chaque méthode et dans chacun des espaces considérés, ainsi que les taux de convergence des schémas de raffinement. On trouvera en outre une description détaillée de la mise en œuvre de ces derniers.

Le troisième chapitre traite de l'application de ces méthodes à la résolution numérique de l'équation de transfert. Cette équation intervient au sein d'un problème beaucoup plus vaste (émanant de la théorie du transfert) dont nous donnons une brève description dans le cadre particulier des atmosphères stellaires. Des expériences numériques, portant sur la validation des méthodes proposées et sur des cas ayant un sens astrophysique, sont présentées. La difficulté de la résolution de cette équation est accrue par la présence d'un très grand paramètre d'intégration. Nous décrivons à la fin de ce dernier chapitre des méthodes asymptotiques permettant de surmonter ce problème.

Enfin, l'annexe B de ce document est un recueil d'articles parus ou à paraître, possédant leurs propres références bibliographiques.

Chapitre 1

Équations intégrales linéaires

1.1 Approximation d'opérateurs linéaires bornés sur un espace de Banach

Dans ce chapitre, X désigne un espace de Banach et $\mathcal{L}(X)$ l'ensemble des opérateurs linéaires bornés sur X. Lorsque qu'il n'y aura aucune ambigüité, et afin d'alléger les notations, $\|\cdot\|$ désignera à la fois la norme dans X et dans $\mathcal{L}(X)$. La norme d'un opérateur $T \in \mathcal{L}(X)$ est définie par

$$\|T\| := \sup\{\|Tx\| \ : \ \|x\| \leq 1\} = \sup\{\|Tx\| \ : \ \|x\| = 1\} = \sup\left\{\frac{\|Tx\|}{\|x\|} \ : \ x \neq 0\right\},$$

et on l'appelle norme subordonnée. Les méthodes numériques présentées dans la suite sont basées sur des approximations d'opérateurs linéaires bornés. Nous allons donc brièvement présenter différents types d'approximations dans l'espace $\mathcal{L}(X)$.

Définition 1 *On dit que $(T_n)_{n\geq 1}$ dans $\mathcal{L}(X)$ est une suite d'approximations de $T \in \mathcal{L}(X)$ si, et seulement si, pour tout $x \in X$,*

$$\lim_{n\to\infty} \|(T_n - T)x\| = 0.$$

On dit aussi que la suite $(T_n)_{n\geq 1}$ converge ponctuellement (ou simplement) vers T.

Définition 2 *On dit que $(T_n)_{n\geq 1}$ dans $\mathcal{L}(X)$ est une suite d'approximations uniforme de $T \in \mathcal{L}(X)$ si, et seulement si,*

$$\lim_{n\to\infty} \|T_n - T\| = 0.$$

Définition 3 *On dit que* $(T_n)_{n\geq 1}$ *dans* $\mathcal{L}(X)$ *est une suite d'approximations collectivement compacte de* $T \in \mathcal{L}(X)$ *si, et seulement si,* $(T_n)_{n\geq 1}$ *converge ponctuellement vers* T *et qu'il existe* $n_0 \geq 0$ *tel que*

$$\bigcup_{n\geq n_0} \{(T_n - T)x \ : \ x \in X, \ \|x\| \leq 1\}$$

est un sous-ensemble relativement compact de X*.*

Définition 4 *On dit que* $(T_n)_{n\geq 1}$ *dans* $\mathcal{L}(X)$ *est* ν*-convergente vers* $T \in \mathcal{L}(X)$ *si, et seulement si, la suite* $(\|T_n\|)_{n\geq 1}$ *est bornée,*

$$\lim_{n\to\infty} \|(T_n - T)T\| = 0 \quad et \quad \lim_{n\to\infty} \|(T_n - T)T_n\| = 0.$$

Les résultats généraux ci-dessous nous seront utiles dans la suite.

Théorème 1 *Soient* $(L_n)_{n\geq 1}$ *dans* $\mathcal{L}(X)$ *une suite d'opérateurs ponctuellement convergente vers* $L \in \mathcal{L}(X)$ *et* T *un opérateur compact. Alors*

$$\lim_{n\to\infty} \|(L_n - L)T\| = 0.$$

Preuve : comme T est compact, $U := \{Tx \ : \ x \in X, \ \|x\| \leq 1\}$ est relativement compact (donc totalement borné) dans l'espace de Banach X. Alors d'après le Théorème de Banach-Steinhaus (voir [54] ou [59]), la convergence ponctuelle de $(L_n)_{n\geq 1}$ est uniforme sur U et donc

$$\lim_{n\to\infty} \|(L_n - L)T\| = \lim_{n\to\infty} \sup\{\|(L_n - L)y\| \ : \ y \in U\} = 0,$$

ce que nous voulions démontrer. ∎

Le corollaire suivant est une conséquence directe du théorème précédent et de la définition de la convergence collectivement compacte.

Corollaire 1 *Soit* $(T_n)_{n\geq 1}$ *dans* $\mathcal{L}(X)$ *une suite d'approximations collectivement compacte d'un opérateur compact* $T \in \mathcal{L}(X)$*. Alors cette suite est* ν*-convergente vers* T*.*

Théorème 2 *Soit* $(T_n)_{n\geq 1}$ *dans* $\mathcal{L}(X)$ *une suite d'approximation uniforme de* $T \in \mathcal{L}(X)$*. Alors cette suite est* ν*-convergente vers* T*.*

Preuve : évidente. ∎

On appelle ensemble résolvant de $T \in \mathcal{L}(X)$ l'ensemble défini par

$$\mathrm{re}(T) := \{z \in \mathbb{C} \ : \ T - zI \text{ est bijectif}\},$$

et, pour tout $z \in \mathrm{re}(T)$, on définit l'opérateur résolvant de T en z par

$$R(T,z) := (T - zI)^{-1}.$$

Ainsi, pour $T \in \mathcal{L}(X)$, $z \in \mathrm{re}(T)$ et $f \in X$, il existe une solution unique φ du problème

$$(T - zI)\varphi = f, \tag{1.1}$$

déterminée par

$$\varphi = R(T,z)f,$$

ce qui justifie le nom donné à l'opérateur $R(T,z)$.

Théorème 3 *Soient $T \in \mathcal{L}(X)$ et $z \in \mathrm{re}(T)$. Alors $R(T,z) \in \mathcal{L}(X)$.*

Preuve : ce résultat est une conséquence du Théorème de l'Application Ouverte (voir [54] ou [59]). ∎

Suivant la complexité de l'opérateur T et du second membre f, l'équation (1.1) devient difficile, voire impossible à résoudre. On se contente alors d'une solution approchée qui est généralement un élément d'une suite convergente dans X vers la solution exacte. Chaque terme φ_n d'une telle suite peut être par exemple la solution du problème approché

$$(T_n - zI)\varphi_n = f, \tag{1.2}$$

où $T_n \in \mathcal{L}(X)$ et $(T_n)_{n\geq 1}$ est une suite d'approximations de T. Il est clair que l'équation (1.2) n'admettra une solution unique que si $z \in \mathrm{re}(T_n)$. Il convient donc de savoir, suivant le type d'approximation de la suite $(T_n)_{n\geq 1}$, si $z \in \mathrm{re}(T_n)$ à partir d'un certain rang. Cette condition est assurée dans le cas de la ν-convergence :

Théorème 4 *Soient $T \in \mathcal{L}(X)$, E un compact de $\mathbb{C}$ contenu dans $\mathrm{re}(T)$ et $(T_n)_{n\geq 1}$ une suite dans $\mathcal{L}(X)$ ν-convergente vers T. Alors il existe $n_0 \geq 1$ tel que*

$$\sup\{\|R(T_n,z)\| \ : \ z \in E, \ n \geq n_0\} < \infty.$$

Preuve : voir le Théorème 2.6 dans [7]. ∎

Une application directe de ce théorème est mise en évidence dans le corollaire suivant qui nous donne une estimation a priori de l'erreur de la solution de l'équation approchée (1.2) par rapport à la solution exacte du problème. Cette estimation sera développée plus en détail dans la suite.

Corollaire 2 *Soient $(T_n)_{n\geq 1}$ une suite dans $\mathcal{L}(X)$ ν-convergente vers T et $(\varphi_n)_{n\geq 1}$ la suite de solutions approchées définies par (1.2). Alors il existe $c > 0$ et $n_0 \geq 0$ tels que pour $n \geq n_0$,*

$$\|\varphi_n - \varphi\| \leq c\|(T - T_n)\varphi\|. \tag{1.3}$$

L'estimation de l'erreur n'est pas le seul moyen de mesurer la qualité de l'approximation φ. Nous pouvons en effet nous intéresser aussi à la précision à laquelle φ_n satisfait l'équation exacte. Cette précision est caractérisée par le résidu de φ_n associé à l'équation (1.1) :

$$r(\varphi_n) := (T - zI)\varphi_n - f. \tag{1.4}$$

Lorsque la suite $(T_n)_{n\geq 1}$ d'approximations de T est uniforme, on obtient l'estimation suivante pour le résidu relatif lorsque n est assez grand :

$$\frac{\|r(\varphi_n)\|}{\|f\|} \leq c\|T - T_n\|, \tag{1.5}$$

où la constante c est celle du Corollaire 2. On remarque qu'on normalise le résidu par la norme du second membre de l'équation (1.1).

Enfin, lorsqu'il n'y aura aucune ambiguïté et afin d'alléger l' écriture, on préférera utiliser les notations $R(z)$ à $R(T, z)$ et $R_n(z)$ à $R(T_n, z)$.

1.2 Opérateurs intégraux à noyau faiblement singulier

Des motivations provenant de la théorie du transfert en Physique que nous présentons dans le chapitre 3 nous ont conduit à considérer deux espaces de Banach X : l'espace $C^0([0, \tau_\star])$ des fonctions continues sur $[0, \tau_\star]$ et l'espace $L^1([0, \tau_\star])$ des classes d'équivalence de fonctions intégrables sur $[0, \tau_\star]$ au sens de Lebesgue, où $\tau_\star > 0$ est un réel donné.

On considère une fonction $\kappa : \{(\tau, \sigma) \in [0, \tau_\star]^2,\ \tau \neq \sigma\} \to \mathbb{C}$ de la forme suivante :

$$\kappa(\tau, \sigma) := \eta(\tau, \sigma) g(|\tau - \sigma|), \quad (\tau, \sigma) \in [0, \tau_\star]^2,\ \tau \neq \sigma, \tag{1.6}$$

où $\eta : [0, \tau_\star]^2 \to \mathbb{C}$ est une fonction continue. On notera

$$\|\eta\|_\infty := \sup_{(\tau,\sigma)\in[0,\tau_\star]^2} |\eta(\tau, \sigma)|. \tag{1.7}$$

On suppose que la fonction $g\ :]0, \tau_\star] \to \mathbb{R}$ est faiblement singulière en zéro au sens suivant :

(1.8) $\displaystyle\lim_{\tau\to 0^+} g(\tau) = +\infty$,

(1.9) $g \in C^0(]0, \tau_\star]) \cap L^1([0, \tau_\star])$,

(1.10) $g(\tau) \geq 0$ pour tout $\tau \in]0, \tau_\star]$,

(1.11) g est une fonction strictement décroissante sur $]0, \tau_\star]$.

Pour plus de détails sur cette caractérisation, voir les commentaires bibliographiques à la fin de cette section.

L'opérateur linéaire

$$x \in X \mapsto \int_0^{\tau_\star} \kappa(\cdot,\sigma)x(\sigma)\,d\sigma, \tag{1.12}$$

est appelé opérateur intégral de noyau faiblement singulier κ. Pour des raisons à la fois théoriques et en rapport avec l'application développée dans le chapitre 3, on considérera les opérateurs intégraux T et Λ définis par

$$(Tx)(\tau) := \int_0^{\tau_\star} \eta(\tau,\sigma)g(|\tau-\sigma|)x(\sigma)\,d\sigma, \tag{1.13}$$

$$(\Lambda x)(\tau) := \int_0^{\tau_\star} g(|\tau-\sigma|)x(\sigma)\,d\sigma. \tag{1.14}$$

Λ correspond à l'opérateur intégral T lorsque la fonction η est constamment égale à 1 sur $[0,\tau_\star]^2$. L'opérateur T étant plus général que Λ, les propriétés vérifiées par T sont aussi vérifiées par Λ. Lorsque T est défini par (1.13), l'équation (1.1) s'écrit

$$\int_0^{\tau_\star} \eta(\tau,\sigma)g(|\tau-\sigma|)\varphi(\sigma)\,d\sigma - z\varphi(\tau) = f(\tau), \quad \tau \in [0,\tau_\star]. \tag{1.15}$$

Avant de démontrer la compacité de T sur $C^0([0,\tau_\star])$ et sur $L^1([0,\tau_\star])$, nous allons démontrer quelques résultats qui seront utiles dans la suite. Définissons tout d'abord la fonction $G : [0,\tau_\star] \to \mathbb{R}$ ci-dessous qui jouera un rôle technique important tout au long de cette section :

$$G(\tau) := \int_0^{\tau} g(\sigma)\,d\sigma. \tag{1.16}$$

On définit de plus, pour tout $\delta > 0$,

$$\varepsilon_1(\kappa,\delta) := \sup\left\{\int_0^{\tau_\star} |\kappa(\tau,\sigma)-\kappa(t,\sigma)|\,d\sigma,\ (\tau,t)\in[0,\tau_\star]^2,\ |\tau-t|\le\delta\right\}, \tag{1.17}$$

$$\varepsilon_2(\kappa,\delta) := \sup\left\{\int_0^{\tau_\star} |\kappa(\tau,\sigma)-\kappa(\tau,s)|\,d\tau.\ (\sigma,s)\in[0,\tau_\star]^2,\ |\sigma-s|\le\delta\right\}. \tag{1.18}$$

L'oscillation d'une fonction x continue sur $[0,\tau_\star]$ par rapport à $\delta > 0$ est définie par

$$\omega_\infty(x,\delta) := \sup\left\{|x(\tau)-x(t)|,\ (\tau,t)\in[0,\tau_\star]^2,\ |\tau-t|\le\delta\right\}. \tag{1.19}$$

Cette définition peut être généralisée, de manière partielle, à toute fonction continue y sur le carré $[0,\tau_\star]^2$: pour tout $\sigma \in [0,\tau_\star]$,

$$\omega_{\infty,1}(y,\delta)(\sigma) := \sup\left\{|y(\tau,\sigma)-y(t,\sigma)|,\ (\tau,t)\in[0,\tau_\star]^2,\ |\tau-t|\le\delta\right\}, \tag{1.20}$$

et pour tout $\tau \in [0, \tau_\star]$,

$$\omega_{\infty,2}(y, \delta)(\tau) \;:=\; \sup\Big\{|y(\tau, \sigma) - y(\tau, s)|,\ (\sigma, s) \in [0, \tau_\star]^2,\ |\sigma - s| \leq \delta\Big\}. \tag{1.21}$$

Remarque 1 *Si x est une fonction continue sur $[0, \tau_\star]$, admettant une dérivée à droite x'_d bornée en chaque point de $]0, \tau_\star[$, et si $\alpha > 0$ est tel que $\|x'_d\| \leq \alpha$, alors*

$$\omega_\infty(x, \delta) \leq \alpha\delta. \tag{1.22}$$

Enfin, la L^1-oscillation d'une fonction $x \in L^1([0, \tau_\star])$ relativement à $\delta > 0$ est définie par

$$\omega_1(x, \delta) := \sup_{\tau \in [0,\delta]} \int_0^{\tau_\star} |x(\tau + t) - x(t)|\, dt, \tag{1.23}$$

où x est prolongée par 0 en dehors de $[0, \tau_\star]$.

Lemme 1 *Soient $\delta > 0$ et y une fonction continue sur le carrée $[0, \tau_\star]^2$. Alors les fonctions $\sigma \in [0, \tau_\star] \mapsto \omega_{\infty,1}(y, \delta)(\sigma) \in \mathbb{R}$ et $\tau \in [0, \tau_\star] \mapsto \omega_{\infty,2}(y, \delta)(\tau) \in \mathbb{R}$, sont continues sur $[0, \tau_\star]$, et*

$$\lim_{\delta \to 0^+} \max\{\|\omega_{\infty,1}(y, \delta)\|_\infty, \|\omega_{\infty,2}(y, \delta)\|_\infty\} = 0. \tag{1.24}$$

Preuve : ceci découle simplement des propriétés de continuité uniforme de y sur le compact $[0, \tau_\star]^2$ de $\mathbb{R}^2$. ∎

Lemme 2 *Soit g vérifiant* (1.9) *et* (1.10). *Alors*

$$\sup_{\tau \in [0,\tau_\star]} \int_0^{\tau_\star} g(|\tau - \sigma|)\, d\sigma = 2\int_0^{\tau_\star/2} g(\sigma)\, d\sigma < +\infty.$$

Preuve : soit $y : [0, \tau_\star] \to \mathbb{R}$ définie par

$$y(\tau) := \int_0^{\tau_\star} g(|\tau - \sigma|)\, d\sigma = G(\tau) + G(\tau_\star - \tau),$$

où G est la fonction définie par (1.16). La fonction y possède une symétrique axiale par rapport à $\tau_\star/2$ et

$$y'(\tau) = g(\tau) - g(\tau_\star - \tau) \left\{ \begin{array}{ll} > 0 & \text{si} \quad 0 < \tau < \tau_\star/2, \\ < 0 & \text{si} \quad \tau_\star/2 < \tau < \tau_\star, \end{array} \right.$$

pour tout τ dans $[0,\tau_\star]$. Ainsi

$$\sup_{\tau\in[0,\tau_\star]}\int_0^{\tau_\star} g(|\tau-\sigma|)\,d\sigma = \max_{\tau\in[0,\tau_\star]} y(\tau) = y(\tau_\star/2) = 2\int_0^{\tau_\star/2} g(\sigma)\,d\sigma,$$

et d'après (1.9) cette dernière intégrale est finie. ∎

On définit alors

$$\gamma := 2\int_0^{\tau_\star/2} g(\sigma)\,d\sigma. \tag{1.25}$$

Lemme 3 *Pour tout $\delta>0$ suffisamment petit,*

$$\varepsilon_1(g,\delta)=\varepsilon_2(g,\delta)\le 4\int_0^{\delta} g(\sigma)\,d\sigma, \tag{1.26}$$

et donc

$$\lim_{\delta\to 0^+}\max\{\varepsilon_1(g,\delta),\varepsilon_2(g,\delta)\}=0. \tag{1.27}$$

Preuve : g jouant le rôle d'un noyau de convolution, la première égalité est évidente. Pour $\delta>0$, t et τ dans $[0,\tau_\star]$ tels que $0<|t-\tau|\le\delta$, définissons $\psi:[0,\tau_\star]\to\mathbb{R}$ par

$$\psi(\sigma):=|g(|t-\sigma|)-g(|\tau-\sigma|)|.$$

Posons

$$r:=\frac{t+\tau}{2}.$$

On peut aisément démontrer que la restriction de ψ à $[t,\tau]$ possède une symétrique axiale par rapport à r, et ainsi

$$\int_0^{\tau_\star}\psi(\sigma)\,d\sigma=\int_0^{t}\psi(\sigma)\,d\sigma+2\int_t^{r}\psi(\sigma)\,d\sigma+\int_\tau^{\tau_\star}\psi(\sigma)\,d\sigma.$$

Or,

$$\begin{aligned}
\int_0^{t}\psi(\sigma)\,d\sigma &= \int_0^{t}(g(t-\sigma)-g(\tau-\sigma))\,d\sigma\\
&= G(t)+G(\tau-t)-G(\tau),\\
\int_t^{r}\psi(\sigma)\,d\sigma &= \int_t^{r}(g(\sigma-t)-g(\tau-\sigma))\,d\sigma\\
&= G((\tau-t)/2)+G((\tau-t)/2)-G(\tau-t),\\
\int_\tau^{\tau_\star}\psi(\sigma)\,d\sigma &= \int_t^{\tau_\star}(g(\sigma-\tau)-g(\sigma-t))\,d\sigma\\
&= G(\tau_\star-\tau)-G(\tau_\star-t)+G(\tau-t),
\end{aligned}$$

où G est la fonction définie par (1.16). Ainsi

$$\begin{aligned}\varepsilon_1(g,\delta) &= 4\int_0^{(\tau-t)/2} g(\sigma)\,d\sigma - \int_t^{\tau} g(\sigma)\,d\sigma - \int_{\tau_\star-\tau}^{\tau_\star-t} g(\sigma)\,d\sigma \le 4\int_0^{(\tau-t)/2} g(\sigma)\,d\sigma \\ &\le 4\int_0^{\delta/2} g(\sigma)\,d\sigma \le 4\int_0^{\delta} g(\sigma)\,d\sigma,\end{aligned}$$

ce que nous voulions démontrer. L'assertion (1.27) découle des propriétés d'intégrabilité de g. ∎

Théorème 5 *Soit g vérifiant* (1.9) *et* (1.10). *Alors*

$$\sup_{\tau\in[0,\tau_\star]} \int_0^{\tau_\star} |\kappa(\tau,\sigma)|\,d\sigma < +\infty \tag{1.28}$$

et

$$\sup_{\sigma\in[0,\tau_\star]} \int_0^{\tau_\star} |\kappa(\tau,\sigma)|\,d\tau < +\infty. \tag{1.29}$$

De plus

$$\lim_{\delta\to0^+} \max\{\varepsilon_1(\kappa,\delta),\varepsilon_2(\kappa,\delta)\} = 0. \tag{1.30}$$

Preuve : (1.28) et (1.29) sont des conséquences directes de la continuité de η et du Lemme 2.

Pour $\delta>0$ assez petit et $(\tau,t)\in[0,\tau_\star]^2$ tels que $|\tau-t|\le\delta$,

$$\begin{aligned}\int_0^{\tau_\star}|\kappa(\tau,\sigma)-\kappa(t,\sigma)|\,d\sigma &= \int_0^{\tau_\star}|\eta(\tau,\sigma)g(|\tau-\sigma|)-\eta(t,\sigma)g(|t-\sigma|)|\,d\sigma \\ &\le \int_0^{\tau_\star}|\eta(\tau,\sigma)||g(|\tau-\sigma|)-g(|t-\sigma|)|\,d\sigma \\ &\quad +\int_0^{\tau_\star}|\eta(\tau,\sigma)-\eta(t,\sigma)|g(|t-\sigma|)\,d\sigma \\ &\le \sup_{\sigma\in[0,\tau_\star]}|\eta(\tau,\sigma)|\int_0^{\tau_\star}|g(|\tau-\sigma|)-g(|t-\sigma|)|\,d\sigma \\ &\quad + \sup_{\sigma\in[0,\tau_\star]}|\eta(\tau,\sigma)-\eta(t,\sigma)|\int_0^{\tau_\star}g(|t-\sigma|)\,d\sigma.\end{aligned}$$

Ainsi, d'après le Lemme 2

$$\varepsilon_1(\kappa,\delta) \le \|\eta\|_\infty\varepsilon_1(g,\delta) + \gamma\|\omega_{\infty,1}(\eta,\delta)\|_\infty, \tag{1.31}$$

où γ est définie par (1.25). Enfin, les Lemmes 1 et 3 impliquent que $\lim\limits_{\delta\to 0^+} \varepsilon_1(\kappa, \delta) = 0$. En utilisant des arguments similaires, on démontre que

$$\varepsilon_2(\kappa, \delta) \leq \|\eta\|_\infty \varepsilon_2(g, \delta) + \gamma \|\omega_{\infty,2}(\eta, \delta)\|_\infty, \tag{1.32}$$

et donc $\lim\limits_{\delta\to 0^+} \varepsilon_2(\kappa, \delta) = 0$, ce qui termine la preuve. ∎

Théorème 6 *Soit T défini par* (1.6)- (1.11) *et* (1.13). *Alors T est un opérateur linéaire compact à la fois de $C^0([0, \tau_\star])$ dans lui-même et de $L^1([0, \tau_\star])$ dans lui-même.*

Preuve : soit $x \in C^0([0, \tau_\star])$ et $\epsilon > 0$. Alors d'après le Théorème 5, il existe $\delta > 0$ tel que $\varepsilon_1(\kappa, \delta) < \dfrac{\epsilon}{\|x\|_\infty}$. Soit $\tau \in [0, \tau_\star]$. Alors, pour tout $t \in [0, \tau_\star]$ tel que $|\tau - t| < \delta$,

$$\begin{aligned} |(Tx)(\tau) - (Tx)(t)| &\leq \int_0^{\tau_\star} |\kappa(\tau, \sigma) - \kappa(t, \sigma)||x(\sigma)|\, d\sigma \\ &\leq \|x\|_\infty \varepsilon_1(\kappa, \delta) < \epsilon. \end{aligned}$$

Ainsi pour tout $x \in C^0([0, \tau_\star])$, Tx est continue en tout point de $[0, \tau_\star]$, ce qui prouve que $C^0([0, \tau_\star])$ est invariant par T. La linéarité de T est évidente.

De plus, soit $x \in C^0([0, \tau_\star])$ tel que $\|x\|_\infty \leq 1$. Alors pour tout $\tau \in [0, \tau_\star]$,

$$\begin{aligned} |(Tx)(\tau)| &\leq \int_0^{\tau_\star} |\eta(\tau, \sigma)| g(|\tau - \sigma|)|x(\sigma)|\, d\sigma \\ &\leq \|x\|_\infty \sup_{\sigma \in [0, \tau_\star]} |\eta(\tau, \sigma)| \int_0^{\tau_\star} g(|\tau - \sigma|)\, dt. \end{aligned}$$

Donc, d'après le Lemme 2,

$$\|T\|_\infty \leq 2\|\eta\|_\infty \int_0^{\tau_\star/2} g(\sigma)\, d\sigma = \gamma \|\eta\|_\infty. \tag{1.33}$$

Ainsi T est borné de $C^0([0, \tau_\star])$ dans lui-même.

Soit $x \in L^1([0, \tau_\star])$. Alors

$$\begin{aligned} \int_0^{\tau_\star} |(Tx)(\tau)|\, d\tau &\leq \int_0^{\tau_\star} \int_0^{\tau_\star} |\kappa(\tau, \sigma)||x(\sigma)|\, d\sigma\, d\tau \\ &\leq \|x\|_1 \sup_{\sigma \in [0, \tau_\star]} \int_0^{\tau_\star} |\kappa(\tau, \sigma)|\, d\tau < +\infty, \end{aligned}$$

d'après le Théorème 5. Donc $L^1([0, \tau_\star])$ est lui aussi invariant par T.

De plus, pour toute fonction x de $L^1([0,\tau_\star])$ telle que $\|x\|_1 \leq 1$,

$$\begin{aligned}\|Tx\|_1 &= \int_0^{\tau_\star}\left|\int_0^{\tau_\star}\eta(\tau,\sigma)g(|\tau-\sigma)x(\sigma)\,d\sigma\right|d\tau \\ &\leq \gamma\|\eta\|_\infty < +\infty. \qquad (1.34)\end{aligned}$$

Donc T est borné de $L^1([0,\tau_\star])$ dans lui-même.

La compacité se déduit en considérant la fonction tronquée g_n définie par

$$g_n(\sigma) := \begin{cases} g(1/n) & \text{si} \quad 0 \leq \sigma < 1/n, \\ g(\sigma) & \text{sinon.}\end{cases}$$

L'opérateur intégral T_n correspondant défini par

$$(T_n x)(\tau) := \int_0^{\tau_\star}\eta(\tau,\sigma)g_n(|\tau-\sigma|)x(\sigma)\,d\sigma$$

est compact à la fois de $C^0([0,\tau_\star])$ dans lui-même et de $L^1([0,\tau_\star])$ dans lui même car son noyau est une fonction continue sur le carré $[0,\tau_\star]^2$. Si $x \in L^1([0,\tau_\star])$, alors

$$\begin{aligned}(T-T_n)x(\tau) &= \int_{\tau-1/n}^{\tau}[g(\tau-\sigma)-g(1/n)]\eta(\tau,\sigma)x(\sigma)\,d\sigma \\ &\quad+\int_\tau^{\tau+1/n}[g(\tau-\sigma)-g(1/n)]\eta(\tau,\sigma)x(\sigma)\,d\sigma \\ &= \int_0^{1/n}[g(s)-g(1/n)]\eta(\tau,\tau-s)x(\tau-s)\,ds \\ &\quad+\int_0^{1/n}[g(s)-g(1/n)]\eta(\tau,\tau+s)x(\tau+s)\,ds.\end{aligned}$$

Donc pour $x \in C^0([0,\tau_\star])$, qui est un sous-ensemble de $L^1([0,\tau_\star])$,

$$\|T-T_n\|_\infty \leq 4\|\eta\|_\infty\int_0^{1/n}g(s)\,ds \to 0, \text{ lorsque } n \to +\infty,$$

et pour $x \in L^1([0,\tau_\star])$, x étant prolongée par 0 en dehors de $[0,\tau_\star]$,

$$\begin{aligned}\int_0^{\tau_\star}|(T-T_n)x(\tau)|\,d\tau &\leq \int_0^{\tau_\star}\int_0^{1/n}|g(s)-g(1/n)|[|\eta(\tau,\tau-s)x(\tau-s)| \\ &\quad+|\eta(\tau,\tau+s)x(\tau+s)|]ds\,d\tau \\ &= \int_0^{1/n}|g(s)-g(1/n)|\int_0^{\tau_\star}[|\eta(\tau,\tau-s)x(\tau-s)| \\ &\quad+|\eta(\tau,\tau+s)x(\tau+s)|]d\tau\,ds \\ &= \|\eta\|_\infty\int_0^{1/n}|g(s)-g(1/n)|\int_0^{\tau_\star}[|x(\tau-s)|+|x(\tau+s)|]\,d\tau\,ds \\ &\leq 4\|x\|_1\|\eta\|_\infty\int_0^{1/n}g(s)\,ds,\end{aligned}$$

donc

$$\|T - T_n\|_1 \le 4\|\eta\|_\infty \int_0^{1/n} g(s)\,ds \to 0 \text{ lorsque } n \to +\infty.$$

Ainsi, dans les deux espaces, T est la limite uniforme d'une suite d'opérateurs compacts.

∎

Remarque 2 *La compacité de T de $C^0([0,\tau_\star])$ dans lui-même peut aussi se déduire du Théorème 5 en appliquant le Théorème d'Arzéla-Ascoli (voir [22]). En effet, Soit U une partie bornée de $C^0([0,\tau_\star])$. La première assertion du Théorème 5 implique que $V := T(U)$ est aussi une partie bornée de cet espace. Soit $y \in V$. Alors il existe $x \in U$ telle que $y = Tx$. Soit $\epsilon > 0$. Alors d'après la deuxième assertion du Théorème 5, il existe $\delta > 0$ tel que $\varepsilon_1(\kappa,\delta) < \dfrac{\epsilon}{\|x\|_\infty}$. Soit $(r,s) \in [0,\tau_\star]^2$ tel que $|r-s| < \delta$. Alors*

$$|y(r) - y(s)| \le \left| \int_0^{\tau_\star} (\kappa(s,t) - \kappa(r,t))x(t)\,dt \right| \le \|x\|_\infty \varepsilon_1(\kappa,\delta) < \epsilon,$$

et donc V est équicontinue. D'après le Théorème d'Arzéla-Ascoli, V est relativement compacte dans $C^0([0,\tau_\star])$ et donc T est compact de $C^0([0,\tau_\star])$ dans lui-même.

Remarque 3 *On trouvera dans [38] des conditions nécessaires et suffisantes pour la compacité d'opérateurs intégraux dans les espaces L^p, où $1 \le p \le \infty$.*

Théorème 7 *Soit Λ défini par* (1.6)- (1.11) *et* (1.14). *Alors*

$$\|\Lambda\|_1 = \|\Lambda\|_\infty = 2\int_0^{\tau_\star/2} g(\sigma)\,d\sigma = \gamma. \tag{1.35}$$

Preuve : lorsque $\eta \equiv 1$ et $x \equiv 1$,

$$(Tx)(\tau) = (\Lambda x)(\tau) = \int_0^{\tau_\star} g(|\tau - \sigma|)\,d\sigma, \quad \tau \in [0,\tau_\star],$$

et ainsi

$$\|\Lambda x\|_\infty = \sup_{\tau \in [0,\tau_\star]} \int_0^{\tau_\star} g(|\tau - \sigma|)\,d\sigma,$$

et d'après l'inégalité (1.33) et le Lemme 2

$$\|\Lambda\|_\infty = 2\int_0^{\tau_\star/2} g(\sigma)\,d\sigma.$$

On définit pour chaque entier $n \geq 2$, la fonction $x_n \in L^1([0,\tau_\star])$ par

$$x_n(\sigma) := \begin{cases} 0 & \text{si} \qquad |\sigma - \tau_\star/2| > 1/n, \\ \dfrac{n}{2} & \text{sinon.} \end{cases}$$

Alors, pour chaque $n \geq 2$, $\|x_n\|_1 = 1$ et il existe σ_n tel que $|\tau_\star/2 - \sigma_n| < 1/n$ et

$$\|\Lambda x_n\|_1 = \frac{n}{2}\int_0^{\tau_\star}\int_{\tau_\star/2-1/n}^{\tau_\star/2+1/n} g(|\tau-\sigma|)\,d\sigma\,d\tau = \int_0^{\tau_\star} g(|\tau-\sigma_n|)\,d\tau, \tag{1.36}$$

donc

$$\lim_{n\to\infty}\|\Lambda x_n\|_1 = \int_0^{\tau_\star} g(|\tau-\tau_\star/2|)\,d\tau = 2\int_0^{\tau_\star/2} g(\tau)\,d\tau = \gamma. \tag{1.37}$$

Enfin, d'après (1.34), pour tout $x \in L^1([0,\tau_\star])$ tel que $\|x\|_1 \leq 1$, $\|\Lambda x\|_1 \leq \gamma$, et ainsi

$$\|\Lambda\|_1 = \gamma,$$

ce que nous voulions démontrer. ∎

Nous terminons ce chapitre en énonçant deux lemmes très utiles pour la construction de bornes d'erreur relative.

Lemme 4 *Soient $x \in C^0([0,\tau_\star])$ et $\delta > 0$. Alors*

$$\omega_\infty(Tx,\delta) \leq \varepsilon_1(\kappa,\delta)\|x\|_\infty. \tag{1.38}$$

Preuve : évidente. ∎

Lemme 5 *Soient $x \in L^1([0,\tau_\star])$ et $\delta > 0$. Alors*

$$\omega_1(Tx,\delta) \leq (\|\eta\|_\infty \varepsilon_1(g,\delta) + \gamma\|\omega_{\infty,1}(\eta,\delta)\|_\infty)\|x\|_1. \tag{1.39}$$

Preuve : remarquons tout d'abord que

$$\sup_{t\in[0,\tau_\star]}\ \sup_{\tau\in[0,\delta]}\ \sup_{\sigma\in[0,\tau_\star]} |\eta(\tau+t,\sigma) - \eta(t,\sigma)| \leq \|\omega_{\infty,1}(\eta,\delta)\|_\infty, \tag{1.40}$$

et que

$$\sup_{\tau\in[0,\delta]}\int_0^{\tau_\star}\int_0^{\tau_\star} |g(|\tau+t-\sigma|) - g(|t-\sigma|)||x(\sigma)|\,d\sigma\,dt \leq \varepsilon_2(g,\delta)\|x\|_1. \tag{1.41}$$

En effet, si $\tau \in [0,\delta]$, alors

$$(1.42)\quad \int_0^{\tau_\star}\int_0^{\tau_\star}|g(|\tau+t-\sigma|)-g(|t-\sigma|)||x(\sigma)|\,d\sigma\,dt = \\ \int_0^{\tau_\star}|x(\sigma)|\int_0^{\tau_\star}|g(|\sigma-\tau-t|)-g(|\sigma-t|)|\,d\sigma\,dt \\ \leq \|x\|_1 \sup_{\sigma\in[0,\tau_\star]}\int_0^{\tau_\star}|g(|(\sigma-\tau)-t|)-g(|\sigma-t|)|\,dt,$$

ce qui, par passage à la borne supérieure par rapport à τ sur $[0,\delta]$, donne (1.41). Soient $x \in L^1([0,\tau_\star])$ et $\tau \in [0,\delta]$. Alors

$$(1.43)\quad \int_0^{\tau_\star}|(Tx)(\tau+t)-(Tx)(t)|\,dt = \\ \int_0^{\tau_\star}\left|\int_0^{\tau_\star}[\eta(\tau+t,\sigma)g(|\tau+t-\sigma|)-\eta(t,\sigma)g(t-\sigma|)]x(\sigma)\,d\sigma\right|dt \\ \leq \int_0^{\tau_\star}\int_0^{\tau_\star}|\eta(\tau+t,\sigma)||g(|\tau+t-\sigma|)-g(|t-\sigma|)||x(\sigma)|\,d\sigma\,dt \\ +\int_0^{\tau_\star}\left|\int_0^{\tau_\star}g(|t-\sigma|)(\eta(\tau+t,\sigma)-\eta(t,\sigma))x(\sigma)\,d\sigma\right|dt \\ \leq \|\eta\|_\infty\int_0^{\tau_\star}\int_0^{\tau_\star}|g(|\tau+t-\sigma|)-g(|t-\sigma|)||x(\sigma)|\,d\sigma\,dt \\ +\sup_{t\in[0,\tau_\star]}\sup_{\sigma\in[0,\tau_\star]}|\eta(\tau+t,\sigma)-\eta(t,\sigma)|\int_0^{\tau_\star}\left|\int_0^{\tau_\star}g(|t-\sigma|)x(\sigma)\,d\sigma\right|dt.$$

Enfin, d'après (1.40) et (1.41) on obtient le résultat souhaité en passant à la borne supérieure en τ sur $[0,\delta]$ dans les deux membres de l'inégalité précédente et en remarquant que $\int_0^{\tau_\star}\left|\int_0^{\tau_\star}g(|t-\sigma|)x(\sigma)\,d\sigma\right|dt = \|\Lambda x\|_1 \leq \gamma\|x\|_1$. ∎

Commentaires bibliographiques

L'introduction de la notion de ν-convergence par Ahues, Largillier et Limaye dans [7] est l'aboutissement de réflexions sur la recherche de conditions suffisantes sur une suite d'opérateur $(T_n)_{n\geq 0}$ pour qu'elle ait toutes les qualités pour se substituer à l'opérateur limite T dans la résolution de l'équation $(T-zI)\varphi = f$ ou dans le problème spectral $T\varphi = \lambda\varphi$. Ces réflexions ont étés initiées par Anselone dans [13] pour la recherche de solutions approchées puis étendues aux problèmes spectraux par Ahues dans [2] et Ahues et Hocine dans [3].

Dans la plupart des ouvrages sur les équations intégrales linéaires, le sens du caractère faiblement singulier d'un noyau est souvent plus restrictif que celui choisi ici.

Par exemple, dans [48] et dans [30], une fonction $\kappa\ : [0, \tau_\star]^2 \to \mathbb{R}$ est dite faiblement singulière si elle est continue pour tout τ et σ dans $[0, \tau_\star]$ tels que $\tau \neq \sigma$ et s'il existe une constante positive M et $\alpha \in]0, 1]$ tels que $|\kappa(\tau, \sigma)| \leq M|\tau - \sigma|^{\alpha - 1}$, $\tau, \sigma \in [0, \tau_\star]$, $\tau \neq \sigma$. On remarque qu'une fonction satisfaisant les hypothèses précédentes satisfait aussi les hypothèses (1.8) - (1.11), ce qui rend notre définition plus générale.

Chapitre 2

Solutions approchées

Cette section est consacrée à construire des approximations de l'opérateur T défini par (1.13), et à donner des estimations a priori de l'erreur relative sur la solution approchée correspondante. Pour tout $\xi \in X^*$, l'adjoint topologique de X, et tout $x \in X$, on définit

$$\langle x\,,\,\xi\rangle := \overline{\xi(x)}. \tag{2.1}$$

Comme ξ est linéaire-conjuguée, alors $\langle x\,,\,\xi\rangle$ est linéaire en x et linéaire-conjugué en ξ, tel un produit scalaire d'un espace préhilbertien complexe.

2.1 Approximations de rang fini

Nous allons étudier ici une certaine classe d'approximations d'opérateurs dont l'espace image est un sous-espace de X de dimension finie. On rappelle que si un opérateur linéaire borné est de rang fini, alors il est compact (résultat classique d'analyse fonctionnelle donné par exemple dans [59] et dans [54]). On rappelle de plus qu'un opérateur linéaire borné de rang fini dans un espace vectoriel normé s'écrit de la manière suivante :

$$T_n := \sum_{j=1}^{n} \langle\cdot\,,\,\ell_{n,j}\rangle e_{n,j}, \tag{2.2}$$

où $n > 0$, et, pour $j \in [\![1, n]\!]$, $\ell_{n,j} \in X^*$ et $e_{n,j} \in X$.

Si z appartient à l'ensemble résolvant de T_n, alors, pour chaque $f \in X$, l'équation approchée

$$T_n\varphi_n = z\varphi_n + f \tag{2.3}$$

admet une solution unique qui dépend continûment de f, d'après le Théorème 3. Sa solution $\varphi_n \in X$ sera alors considérée comme une solution approchée de l'équation (1.1). Si nous appliquons chaque fonctionnelle $\ell_{n,i}$ aux deux membres de l'équation (2.3), alors, pour tout $i \in [\![1, n]\!]$,

$$\sum_{j=1}^{n} \langle \varphi_n \, , \, \ell_{n,j} \rangle \langle e_{n,j} \, , \, \ell_{n,i} \rangle - z \langle \varphi_n \, , \, \ell_{n,i} \rangle = \langle f \, , \, \ell_{n,i} \rangle.$$

Si on pose

$$\mathsf{A}_n(i,j) := \langle e_{n,j} \, , \, \ell_{n,i} \rangle, \;\; \mathsf{b}_n(i) := \langle f \, , \, \ell_{n,i} \rangle \;\; \text{et} \;\; \mathsf{x}_n(j) := \langle \varphi_n \, , \, \ell_{n,j} \rangle, \tag{2.4}$$

alors l'équation (2.3) conduit au système linéaire, de dimension n et d'inconnue x_n,

$$(\mathsf{A}_n - z\mathsf{I}_n)\mathsf{x}_n = \mathsf{b}_n, \tag{2.5}$$

où I_n désigne la matrice identité d'ordre n. On appellera A_n «matrice d'approximation d'ordre n associée à l'opérateur T_n». Une fois ce système résolu, la relation (2.3) donne φ_n sous la forme

$$\varphi_n = \frac{1}{z} \left(\sum_{j=1}^{n} \mathsf{x}_n(j) e_{n,j} - f \right). \tag{2.6}$$

Remarque 4 *Il est important de constater que l'expression de la solution approchée ainsi déduite est donnée en termes des fonctions $e_{n,j}$, $1 \leq j \leq n$. Ainsi, lorsque ces dernières, aussi bien que f, peuvent être évaluées en tout point, il en est de même de la solution approchée φ_n.*

2.1.1 Approximation par projection

On s'intéresse à des approximations obtenues à l'aide d'une famille de projections bornées de rang fini sur X. Une projection bornée π_n de rang fini $n \geq 1$ est définie par

$$\pi_n x := \sum_{j=1}^{n} \langle x \, , \, \xi_{n,j} \rangle e_{n,j}, \quad x \in X, \tag{2.7}$$

où $(e_{n,j})_{j=1}^{n}$ est une base ordonnée de l'espace image de π_n, et $(\xi_{n,j})_{j=1}^{n}$ est une base adjointe de la précédente. L'intersection des noyaux de ces fonctionnelles linéaires-conjuguées bornées donne le noyau de π_n. On construit l'opérateur approché T_n à partir de T et de π_n de la manière suivante :

$$T_n := \pi_n T. \tag{2.8}$$

Ainsi, pour tout $x \in X$

$$T_n x = (\pi_n T)x = \sum_{j=1}^{n} \langle Tx\,,\,\xi_{n,j}\rangle e_{n,j}. \tag{2.9}$$

En identifiant

$$\ell_{n,j} := T^* \xi_{n,j}, \tag{2.10}$$

on retrouve bien la représentation (2.2).

On suppose que $(\pi_n)_{n\geq 1}$ est ponctuellement convergente vers l'opérateur identité dans l'espace de Banach X sur lequel T est défini. Comme T est compact, d'après le Théorème 1, la suite $(T_n)_{n\geq 1}$ converge uniformément vers T. Alors, d'après les Théorèmes 2 et 4, l'opérateur résolvant de T_n en z existe pour n assez grand et est uniformément borné. Ainsi il existe $n_0 \geq 1$, tel que

$$c_0(z) := \sup_{n\geq n_0} \|R_n(z)\| < +\infty, \tag{2.11}$$

et d'après le Corollaire 2, pour $n \geq n_0$,

$$\|\varphi - \varphi_n\| \leq c_0(z)\|(T_n - T)\varphi\| = c_0(z)\|(I - \pi_n)T\varphi\|. \tag{2.12}$$

Nous allons maintenant développer des exemples où des bornes pour l'erreur relative seront proposées, dans le cas de l'espace $C^0([0,\tau_\star])$ et dans le cas de l'espace $L^1([0,\tau_\star])$.

2.1.1.1 Application dans l'espace $C^0([0,\tau_\star])$

On considère une suite d'interpolants affines par morceaux dans $C^0([0,\tau_\star])$: soit $(\tau_{n,j})_{j=1}^n$ une grille sur $[0,\tau_\star]$ telle que

$$0 =: \tau_{n,1} < \tau_{n,2} < \ldots < \tau_{n,n-1} < \tau_{n,n} := \tau_\star. \tag{2.13}$$

On pose

$$h_{n,j} := \tau_{n,j} - \tau_{n,j-1} \quad \text{pour } j \in [\![2, n]\!]. \tag{2.14}$$

On définit, pour $x \in C^0([0,\tau_\star])$ et $j \in [\![1, n]\!]$,

$$\langle x\,,\,\xi_{n,j}\rangle := x(\tau_{n,j}), \tag{2.15}$$

et, pour $\tau \in [0, \tau_\star]$ et $j \in [\![2, n-1]\!]$,

$$e_{n,j}(\tau) \;:=\; \begin{cases} 1 + (\tau - \tau_{n,j})/h_{n,j} & \text{si } \tau \in [\tau_{n,j-1}, \tau_{n,j}], \\ 1 + (\tau_{n,j} - \tau)/h_{n,j+1} & \text{si } \tau \in [\tau_{n,j}, \tau_{n,j+1}], \\ 0 & \text{sinon}, \end{cases}$$

$$e_{n,1}(\tau) \;:=\; \begin{cases} (\tau_{n,2} - \tau)/h_{n,2} & \text{si } \tau \in [0, \tau_{n,2}], \\ 0 & \text{sinon}, \end{cases} \tag{2.16}$$

$$e_{n,n}(\tau) \;:=\; \begin{cases} (\tau - \tau_{n,n-1})/h_{n,n} & \text{si } \tau \in [\tau_{n,n-1}, \tau_\star], \\ 0 & \text{sinon}. \end{cases}$$

On pose

$$h_n := \max\{h_{n,j} \;:\; j \in [\![2, n]\!]\}.$$

Lemme 6 *La suite $(\pi_n)_{n\geq 2}$ définie par (2.7) et (2.13) - (2.16) est bornée.*

Preuve : $\|\pi_n\|_\infty = \sup\{\|\pi_n x\|_\infty : \; x \in C^0([0, \tau_\star]), \; \|x\|_\infty = 1\} = 1.$ ∎

Lemme 7 *Soient $x \in C^0([0, \tau_\star])$ et $(\pi_n)_{n\geq 2}$ la suite définie par (2.7) et (2.13) - (2.16). Alors*

$$\|(I - \pi_n)x\|_\infty \leq \omega_\infty(x, h_n), \tag{2.17}$$

et ainsi $\lim_{n\to\infty} \|(I-\pi_n)x\|_\infty = 0$, lorsque $\lim_{n\to\infty} h_n = 0$.

Preuve : on rappelle que pour tout $\tau \in [0, \tau_\star]$, $\sum_{j=1}^{n} |e_{n,j}(\tau)| = 1$. Alors

$$\|(I - \pi_n)x\|_\infty = \max_{0\leq\tau\leq\tau_\star} \left| \sum_{j=1}^{n} (x(\tau) - x(\tau_{n,j}))e_{n,j}(\tau) \right| \leq \omega_\infty(x, h_n).$$

Enfin, comme $x \in C^0([0, \tau_\star])$, si $\lim_{n\to\infty} h_n = 0$, alors $\lim_{n\to\infty} \omega_\infty(x, h_n) = 0$. ∎

Lemme 8 *Soient T défini par (1.13), π_n définie par (2.7), (2.13) - (2.14) - (2.15) - (2.16) et T_n défini par (2.8). Alors*

$$\|T - T_n\|_\infty \leq \varepsilon_1(\kappa, h_n). \tag{2.18}$$

Preuve : conséquence directe des Lemmes 4 et 7. ∎

Lorsque T est défini par (1.13) les coefficients de la matrice A_n et du vecteur b_n du système (2.5) sont donnés, pour tout i et j dans $[\![1, n]\!]$, par

$$\mathsf{A}_n(i,j) = (Te_{n,j})(\tau_{n,i}) = \int_0^{\tau_\star} \eta(\tau_{n,i},\sigma) g(|\tau_{n,i}-\sigma|) e_{n,j}(\sigma)\, d\sigma, \tag{2.19}$$

$$\mathsf{b}_n(i) = (Tf)(\tau_{n,i}) = \int_0^{\tau_\star} \eta(\tau_{n,i},\sigma) g(|\tau_{n,i}-\sigma|) f(\sigma)\, d\sigma. \tag{2.20}$$

Théorème 8 *Soient φ la solution de (1.15) avec T défini par (1.13) et $f \neq 0$. Soit φ_n la solution de (2.3) avec T_n défini par (2.8) et (2.13) - (2.16). Alors $\varphi \neq 0$ et pour n assez grand,*

$$\frac{\|\varphi_n - \varphi\|_\infty}{\|\varphi\|_\infty} \leq c_0(z) \left(4\|\eta\|_\infty \int_0^{h_n} g(\sigma)\, d\sigma + \gamma \|\omega_{\infty,1}(\eta, h_n)\|_\infty \right), \tag{2.21}$$

où $c_0(z)$ est donnée par (2.11) et calculée avec la norme de $C^0([0, \tau_\star])$.

Preuve : en appliquant le Lemme 7 à l'inégalité (2.12) il vient

$$\|\varphi - \varphi_n\|_\infty \leq c_0(z) \omega_\infty(T\varphi, h_n).$$

Le Lemme 4 implique que

$$\omega_\infty(T\varphi, h_n) \leq \varepsilon_1(\kappa, h_n) \|\varphi\|_\infty.$$

La borne est obtenue en appliquant l'inégalité (1.31) de la preuve du Théorème 5 et le Lemme 3. ∎

Remarque 5 *La borne de l'erreur relative (2.21) montre que*

$$\frac{\|\varphi_n - \varphi\|_\infty}{\|\varphi\|_\infty} = O\left(\int_0^{h_n} g(\sigma)\, d\sigma \right) + O\left(\|\omega_{\infty,1}(\eta, h_n)\|_\infty \right).$$

Le terme $\int_0^{h_n} g(\sigma)\, d\sigma$ mesure l'explosion de la fonction g au voisinage de zéro, et $\|\omega_{\infty,1}(\eta, h_n)\|_\infty$ la variabilité de la fonction η. Notons que lorsque cette dernière est constante, son oscillation est nulle, et par conséquent la borne ne dépend plus que de l'explosion de g au voisinage de 0.

D'après (1.5), le Lemme 8 et l'estimation (1.31), la borne (2.21) est aussi une majoration du résidu relatif de φ_n associé à l'équation (1.1).

2.1.1.2 Application dans l'espace $L^1([0,\tau_\star])$

Soit $(\tau_{n,j})_{j=0}^{n}$ une grille sur $[0,\tau_\star]$ telle que

$$0 =: \tau_{n,0} < \tau_{n,1} < \ldots < \tau_{n,n-1} < \tau_{n,n} := \tau_\star. \tag{2.22}$$

Posons

$$h_{n,j} \quad := \quad \tau_{n,j} - \tau_{n,j-1} \qquad \text{pour } j \in [\![1,n]\!], \tag{2.23}$$

et définissons

$$\begin{aligned} \mu_n &:= \min\{h_{n,i} \,:\, i \in [\![1,n]\!]\}, \\ h_n &:= \max\{h_{n,j} \,:\, j \in [\![1,n]\!]\}, \\ q_n &:= \frac{\mu_n}{h_n}. \end{aligned}$$

Remarque 6 *Dans le cas d'une famille de grilles «quasi-uniformes», il existe une constante q indépendante de n telle que, pour tout n, $q \le q_n$. Dans le cas d'une famille de grilles uniformes, $q_n = 1$ pour tout n.*

On définit, pour $x \in L^1([0,\tau_\star])$,

$$\langle x \,,\, \xi_{n,j}\rangle := \frac{1}{h_{n,j}} \int_{\tau_{n,j-1}}^{\tau_{n,j}} x(\sigma)\, d\sigma, \tag{2.24}$$

et, pour $\tau \in [0,\tau_\star]$,

$$e_{n,j}(\tau) := \begin{cases} 1 & \text{si } \tau \in \,]\tau_{n,j-1}, \tau_{n,j}[, \\ 0 & \text{sinon.} \end{cases} \tag{2.25}$$

Il est clair que $\langle e_{n,j} \,,\, \xi_{n,i}\rangle = \delta_{i,j}$ pour tout i et j dans $[\![1,n]\!]$. On remarque de plus que lorsque $x \in C^0([0,\tau_\star])$, $\lim\limits_{h_{n,i}\to 0^+} \langle x \,,\, \xi_{n,i}\rangle = x(\tau_{n,i})$ pour chaque i dans $[\![1,n]\!]$.

Lemme 9 *Soit $x \in L^1[(0,\tau_\star])$. Si $\lim\limits_{n\to\infty} h_n = 0$, alors $\lim\limits_{n\to\infty} \|(I-\pi_n)x\|_1 = 0$.*

Preuve : la convergence ponctuelle de π_n vers l'identité dans $L^1([0,\tau_\star])$ peut être établie en la démontrant tout d'abord dans $C^0([0,\tau_\star])$ (ceci sera fait en utilisant la continuité uniforme sur $[0,\tau_\star]$ et en utilisant le Théorème de la Valeur Intermédiaire pour chaque intégration sur les sous-intervalles $]\tau_{n,j-1},\tau_{n,j}[$, $j \in [\![1,n]\!]$). Ensuite, on utilise la densité de $C^0([0,\tau_\star])$ dans $L^1([0,\tau_\star])$ et on applique le Théorème de Banach-Steinhaus à la suite bornée $(\pi_n)_{n\ge 1}$:

$$\|\pi_n\|_1 = \sup\{\|\pi_n x\|_1 \;:\; x \in L^1([0,\tau_\star]),\ \|x\|_1 = 1\} = 1, \tag{2.26}$$

car

$$\begin{aligned}
\|\pi_n x\|_1 &\le \int_0^{\tau_\star} \sum_{j=1}^{n} \frac{1}{h_{n,j}} \left(\int_{\tau_{n,j-1}}^{\tau_{n,j}} |x(\sigma)|\, d\sigma \right) e_{n,j}(\tau)\, d\tau \\
&= \sum_{j=1}^{n} \left(\int_{\tau_{n,j-1}}^{\tau_{n,j}} |x(\sigma)|\, d\sigma \right) \left(\frac{1}{h_{n,j}} \int_{\tau_{n,j-1}}^{\tau_{n,j}} d\tau \right) = \|x\|_1,
\end{aligned}$$

et la borne est atteinte pour $x \equiv 1/\tau_\star$. ∎

Lemme 10 *Soit* $x \in L^1([0,\tau_\star])$. *Alors*

$$\|(I-\pi_n)x\|_1 \le \frac{2}{q_n}\omega_1(x,h_n). \tag{2.27}$$

Preuve : pour tout $i \in [\![1, n]\!]$,

$$\begin{aligned}
\int_{\tau_{n,i-1}}^{\tau_{n,i}} |(I-\pi_n)x(\tau)|\, d\tau &\le \frac{1}{h_{n,i}} \int_{\tau_{n,i-1}}^{\tau_{n,i}} \int_{\tau_{n,i-1}}^{\tau_{n,i}} |x(\tau)-x(t)|\, dt\, d\tau \\
&= \frac{2}{h_{n,i}} \int_{\tau_{n,i-1}}^{\tau_{n,i}} \int_{t}^{\tau_{n,i}} |x(\tau)-x(t)|\, d\tau\, dt \\
&= \frac{2}{h_{n,i}} \int_{\tau_{n,i-1}}^{\tau_{n,i}} \int_{0}^{\tau_{n,i}-t} |x(\tau+t)-x(t)|\, d\tau\, dt \\
&\le \frac{2}{h_{n,i}} \int_{\tau_{n,i-1}}^{\tau_{n,i}} \int_{0}^{h_{n,i}} |x(\tau+t)-x(t)|\, d\tau\, dt,
\end{aligned}$$

et ainsi

$$\|(I-\pi_n)x\|_1 \le \frac{2}{\mu_n} \int_0^{h_n} \int_0^{\tau_\star} |x(\tau+t)-x(t)|\, d\tau\, dt \le \frac{2}{q_n}\omega_1(x,h_n),$$

ce que nous voulions démontrer. ∎

Lemme 11 *Soient* T *défini par* (1.13) *et* T_n *défini par* (2.8) *et* (2.22) - (2.25). *Alors*

$$\|T-T_n\|_1 \le \frac{2}{q_n}(\|\eta\|_\infty \varepsilon_1(g,h_n) + \gamma \|\omega_{\infty,1}(\eta,h_n)\|_\infty). \tag{2.28}$$

Preuve : conséquence directe des Lemmes 5 et 10. ∎

Lorsque T est défini par (1.13), les coefficients de la matrice A_n et du second membre b_n du système (2.5) sont donnés, pour tout i et j dans $[\![1, n]\!]$, par

$$\mathsf{A}_n(i,j) = \frac{1}{h_{n,i}} \int_{\tau_{n,i-1}}^{\tau_{n,i}} \int_{\tau_{n,j-1}}^{\tau_{n,j}} \eta(\tau,\sigma) g(|\tau-\sigma|)\, d\sigma\, d\tau, \tag{2.29}$$

$$\mathsf{b}_n(i) = \frac{1}{h_{n,i}} \int_{\tau_{n,i-1}}^{\tau_{n,i}} \int_{0}^{\tau_\star} \eta(\tau,\sigma) g(|\tau-\sigma|) f(\sigma)\, d\sigma\, d\tau. \tag{2.30}$$

Théorème 9 *Soient φ la solution de (1.15) avec T défini par (1.13) et $f \neq 0$. Soit φ_n la solution de (2.3) avec T_n défini par (2.8) et (2.22) - (2.25). Alors $\varphi \neq 0$ et pour n assez grand,*

$$\frac{\|\varphi_n - \varphi\|_1}{\|\varphi\|_1} \leq \frac{2c_0(z)}{q_n} \left(4\|\eta\|_\infty \int_0^{h_n} g(\sigma)\, d\sigma + \gamma \|\omega_{\infty,1}(\eta, h_n)\|_\infty \right), \tag{2.31}$$

où $c_0(z)$ est donnée par (2.11) et calculée avec la norme de $L^1([0,\tau_\star])$.

Preuve : en appliquant le Lemme 10 à l'inégalité (2.12) il vient

$$\|\varphi - \varphi_n\|_1 \leq \frac{2c_0(z)}{q_n} \omega_1(T\varphi, h_n).$$

Le Lemme 5 implique que $\omega_1(T\varphi, h_n) \leq (\|\eta\|_\infty \varepsilon_1(g, h_n) + \gamma \|\omega_{\infty,1}(\eta, h_n)\|_\infty) \|\varphi\|_1$. La borne est obtenue en appliquant le Lemme 3. ∎

Remarque 7 *L'estimation (2.31) montre que*

$$\frac{\|\varphi_n - \varphi\|_1}{\|\varphi\|_1} = O\left(\frac{1}{q_n} \int_0^{h_n} g(\sigma)\, d\sigma \right) + O\left(\frac{1}{q_n} \|\omega_{\infty,1}(\eta, h_n)\|_\infty \right).$$

À l'instar de la borne (2.21), cette borne dépend de l'explosion de la fonction g au voisinage de zéro et de la variabilité de la fonction η. Mais contrairement au cas de l'espace $C^0([0,\tau_\star])$, elle est aussi sensible à la régularité du maillage, mesurée par le rapport q_n. D'après la Remarque 6, si les maillages choisis sont «quasi-uniformes», la borne précédente est alors du même ordre que la borne (2.21). Si $\int_0^{h_n} g(\sigma)\, d\sigma$ tend vers zéro, mais que $\frac{1}{q_n}$ tend vers l'infini, alors pour s'assurer que φ_n tende vers φ, une certaine prudence s'impose dans le choix des grilles du maillage afin que $\lim_{n\to\infty} \frac{1}{q_n} \int_0^{h_n} g(\sigma)\, d\sigma = 0$, comme le montre l'exemple suivant : soit $g :]0,1] \to \mathbb{R}$ définie par

$$g(\sigma) := \frac{1}{\sqrt{\sigma}}.$$

Elle satisfait bien (1.8) (1.11). *Soit* $(\tau_{n,j})_{j=0}^{n}$ *la famille de grilles définie par*

$$\begin{aligned}\tau_{n,0} &:= 0,\\ \tau_{n,j} &:= \frac{1}{2(n-j)}, \quad 1 \leq j \leq n-1,\\ \tau_{n,n} &:= 1.\end{aligned}$$

Alors $h_n = \frac{1}{2}$, $\mu_n = \frac{1}{2(n-1)}$, *et*

$$\frac{1}{q_n}\int_0^{h_n} g(\sigma)\,d\sigma = \frac{2h_n\sqrt{h_n}}{\mu_n} = \sqrt{2}(n-1) \to \infty$$

lorsque n tend vers l'infini, ce qui ne nous permet pas d'affirmer que la suite d'approximations converge. Ce choix de grilles peut intervenir lorsqu'un raffinement du maillage est nécessaire au voisinage d'un certain point (l'origine dans notre exemple). Il est donc important de faire varier le pas maximum en même temps que le pas minimum.

Enfin, d'après (1.5), *le Lemme 11 et l'estimation* (1.31), *la borne* (2.31) *est aussi une borne pour le résidu relatif de* φ_n *associé à l'équation* (1.1).

2.1.2 Approximations par intégration produit

L'équation intégrale (1.15) est approchée par moyennes sur une grille $(\tau_{n,j})_{j=0}^{n}$ sur $[0, \tau_\star]$ telle que

$$0 =: \tau_{n,0} < \tau_{n,1} < \ldots < \tau_{n,n-1} < \tau_{n,n} := \tau_\star, \tag{2.32}$$

où $n \geq 1$. À partir de cette grille, on construit une grille intermédiaire $(\widehat{\tau}_{n,j})_{j=1}^{n}$ satisfaisant

$$\widehat{\tau}_{n,j} \in\,]\tau_{n,j-1}, \tau_{n,j}[\text{ pour tout } j \in [\![1, n]\!]. \tag{2.33}$$

On présente deux approximations différentes suivant l'espace considéré. On définit comme précédemment

$$h_{n,j} := \tau_{n,j} - \tau_{n,j-1} \quad \text{pour } j \in [\![1, n]\!], \tag{2.34}$$

$$h_n := \max\{h_{n,j}, \quad j \in [\![1, n]\!]\}. \tag{2.35}$$

2.1.2.1 Application dans l'espace $C^0([0, \tau_\star])$

On approche la solution φ de (1.15) par la solution φ_n de

$$\sum_{j=1}^{n} \varphi_n(\widehat{\tau}_{n,j})\eta(\tau, \widehat{\tau}_{n,j})\int_{\tau_{n,j-1}}^{\tau_{n,j}} g(|\tau - \sigma|)\,d\sigma - z\varphi_n(\tau) = f(\tau), \quad \tau \in [0, \tau_\star]. \tag{2.36}$$

On remarque que φ_n est entièrement déterminée sur $[0, \tau_\star]$ par ses valeurs aux noeuds de la grille intermédiaire. Si τ prend les valeurs $\widehat{\tau}_{n,i}$ pour $i \in [\![1, n]\!]$ dans (2.36), et si on pose

$$\mathsf{x}_n(j) := \varphi_n(\widehat{\tau}_{n,j}), \quad j \in [\![1, n]\!], \tag{2.37}$$

alors x_n résout le système linéaire

$$(\mathsf{A}_n - z\mathsf{I}_n)\mathsf{x}_n = \mathsf{b}_n, \tag{2.38}$$

où

$$\mathsf{A}_n(i, j) := \eta(\widehat{\tau}_{n,i}, \widehat{\tau}_{n,j}) \int_{\tau_{n,j-1}}^{\tau_{n,j}} g(|\widehat{\tau}_{n,i} - \sigma|)\, d\sigma, \tag{2.39}$$

$$\mathsf{b}_n(i) := f(\widehat{\tau}_{n,i}). \tag{2.40}$$

Afin de construire une borne d'erreur pour φ_n, on remarque que φ_n résout

$$(T_n - zI)\varphi_n = f, \tag{2.41}$$

où

$$(T_n x)(\tau) := \sum_{j=1}^{n} x(\widehat{\tau}_{n,j})\eta(\tau, \widehat{\tau}_{n,j}) \int_{\tau_{n,j-1}}^{\tau_{n,j}} g(|\tau - \sigma|)\, d\sigma \tag{2.42}$$

pour $x \in C^0([0, \tau_\star])$ et $\tau \in [0, \tau_\star]$. On démontre facilement que

$$\|T_n\|_\infty \leq \gamma \|\eta\|_\infty. \tag{2.43}$$

On identifie la famille de fonctions $(e_{n,j})_{j=1}^n$ dans $C^0([0, \tau_\star])$ définies par

$$e_{n,j}(\tau) := \eta(\tau, \widehat{\tau}_{n,j}) \int_{\tau_{n,j-1}}^{\tau_{n,j}} g(|\tau - \sigma|)\, d\sigma \quad \text{pour tout } \tau \in [0, \tau_\star], \tag{2.44}$$

et la famille de fonctionnelles linéaires-conjuguées $(\ell_{n,j})_{j=1}^n$ sur $C^0([0, \tau_\star])$ définies par

$$\langle x\, ,\, \ell_{n,j} \rangle := x(\widehat{\tau}_{n,j}) \quad \text{pour tout } x \in C^0([0, \tau_\star]), \tag{2.45}$$

et nous retrouvons la représentation (2.2).

Théorème 10 *Si h_n tend vers 0 lorsque n tend vers l'infini, alors $(T_n)_{n \geq 1}$ est une suite d'approximations de T collectivement compacte.*

Preuve : soit $x \in C^0([0,\tau_\star])$. Alors

$$\begin{aligned}
\|(T-T_n)x\|_\infty &= \sup_{\tau\in[0,\tau_\star]} \left| \int_0^{\tau_\star} \kappa(\tau,\sigma)x(\sigma)\,d\sigma - \sum_{j=1}^n x(\widehat{\tau}_{n,j})\eta(\tau,\widehat{\tau}_{n,j}) \int_{\tau_{n,j-1}}^{\tau_{n,j}} g(|\tau-\sigma|)\,d\sigma \right| \\
&\le \sup_{\tau\in[0,\tau_\star]} \sum_{j=1}^n \int_{\tau_{n,j-1}}^{\tau_{n,j}} |g(|\tau-\sigma|)||\eta(\tau,\sigma)x(\sigma)-\eta(\tau,\widehat{\tau}_{n,j})x(\widehat{\tau}_{n,j})|\,d\sigma \\
&\le \omega_\infty(x,h_n)\|T\|_\infty + \gamma\|x\|_\infty \|\omega_{\infty,2}(\eta,h_n)\|_\infty .
\end{aligned}$$

Ceci prouve la convergence ponctuelle de $(T_n)_{n\ge 1}$ vers T. Grâce au Principe de la Borne Uniforme, la suite $(\|T_n\|_\infty)_{n\ge 1}$ est bornée et ainsi l'ensemble

$$W := \{T_n x \ : \ x \in C^0([0,\tau_\star]),\ \|x\|_\infty \le 1,\ n>0\}$$

est borné. Il nous reste donc à prouver que W est équicontinu. Soient $\delta>0$, $x \in C^0([0,\tau_\star])$ tel que $\|x\|_\infty \le 1$, $(\tau,t)\in[0,\tau_\star]^2$ tel que $|\tau-t|\le\delta$, et $n>0$. Alors

$$\begin{aligned}
|(T_nx)(\tau)-(T_nx)(t)| &\le \sum_{j=1}^n \Big| x(\widehat{\tau}_{n,j})\Big(\eta(\tau,\widehat{\tau}_{n,j})\int_{\tau_{n,j-1}}^{\tau_{n,j}} g(|\tau-\sigma|)\,d\sigma \\
&\qquad -\eta(t,\widehat{\tau}_{n,j})\int_{\tau_{n,j-1}}^{\tau_{n,j}} g(|t-\sigma|)\,d\sigma\Big)\Big| \\
&\le \|x\|_\infty \sum_{j=1}^n \Big[|\eta(\tau,\widehat{\tau}_{n,j})-\eta(t,\widehat{\tau}_{n,j})| \int_{\tau_{n,j-1}}^{\tau_{n,j}} g(|\tau-\sigma|)\,d\sigma \\
&\qquad +|\eta(t,\widehat{\tau}_{n,j})| \int_{\tau_{n,j-1}}^{\tau_{n,j}} |g(|\tau-\sigma|)-g(|t-\sigma|)|\,d\sigma\Big] \\
&\le \gamma\|\omega_{\infty,1}(\eta,\delta)\|_\infty + \|\eta\|_\infty \varepsilon_1(g,\delta),
\end{aligned}$$

qui tend uniformément vers 0 par rapport à x et n, si $\delta\to 0^+$, d'après les Lemmes 1 et 3. ∎

D'après la preuve précédente, on remarque que

$$\omega_\infty(T_nx,\delta) \le (\gamma\|\omega_{\infty,1}(\eta,\delta)\|_\infty + \|\eta\|_\infty\varepsilon_1(g,\delta))\|x\|_\infty. \tag{2.46}$$

Corollaire 3 *Soient T défini par (1.6) et $(T_n)_{n\ge 1}$ définie par (2.42). Alors il existe $n_0>0$ tel que*

$$c_0(z) := \sup_{n\ge n_0} \|R_n(z)\|_\infty < +\infty. \tag{2.47}$$

Preuve : conséquence directe du Théorème 10, du Corollaire 1 et du Théorème 4. ∎

De plus, d'après le Corollaire 2, pour tout n assez grand,

$$\|\varphi_n - \varphi\|_\infty \leq c_0(z)\|(T - T_n)\varphi\|_\infty. \tag{2.48}$$

Théorème 11 *Soient φ la solution de (1.15) avec T défini par (1.13) et $f \neq 0$. Soit φ_n la solution de (2.3) avec T_n défini par (2.42). Alors $\varphi \neq 0$ et pour n assez grand,*

$$\begin{aligned} \frac{\|\varphi_n - \varphi\|_\infty}{\|\varphi\|_\infty} \leq\ & c_0(z)\left[\frac{\|T\|_\infty}{|z|}\left(4\|\eta\|_\infty \int_0^{h_n} g(\sigma)\,d\sigma + \gamma\|\omega_{\infty,1}(\eta, h_n)\|_\infty\right.\right. \\ & \left.\left. +\|T - zI\|_\infty \frac{\omega_\infty(f, h_n)}{\|f\|_\infty}\right) + \gamma\|\omega_{\infty,2}(\eta, h_n)\|_\infty\right], \end{aligned} \tag{2.49}$$

où $c_0(z)$ est donnée par (2.47).

Preuve : on a vu dans la démonstration du Théorème 10 que

$$\|(T - T_n)\varphi\|_\infty \leq \omega_\infty(\varphi, h_n)\|T\|_\infty + \gamma\|\varphi\|_\infty\|\omega_{\infty,2}(\eta, h_n)\|_\infty.$$

Or, $\varphi = \dfrac{1}{z}(T\varphi - f)$ donc

$$\omega_\infty(\varphi, h_n) \leq \frac{1}{|z|}\left(\omega_\infty(T\varphi, h_n) + \omega_\infty(f, h_n)\right).$$

D'après le Lemme 4,

$$\omega_\infty(T\varphi, h_n) \leq \varepsilon_1(\kappa, h_n)\|\varphi\|_\infty$$

et

$$\omega_\infty(f, h_n) \leq \frac{\omega_\infty(f, h_n)}{\|f\|_\infty}\|(T - zI)\varphi\|_\infty \leq \frac{\omega_\infty(f, h_n)}{\|f\|_\infty}\|T - zI\|_\infty\|\varphi\|_\infty.$$

On obtient la borne (2.49) en appliquant l'inégalité (1.31) de la démonstration du Théorème 5. ∎

Remarque 8 *Contrairement aux estimations obtenues dans le cas des approximations par projection, la borne (2.49) fait intervenir la variabilité de f, le second membre de l'équation. Notons que cette borne ne dépend pas de la régularité du maillage.*

2.1.2.2 Application dans l'espace $L^1([0,\tau_\star])$

On approche la solution φ de (1.15) par la solution φ_n de

$$\sum_{j=1}^{n}\kappa(\,\cdot\,,\widehat{\tau}_{n,j})\int_{\tau_{n,j-1}}^{\tau_{n,j}}\varphi_n(\sigma)\,d\sigma - z\varphi_n = f \quad \text{p.p. dans } [0,\tau_\star]. \tag{2.50}$$

On remarque que φ_n est définie presque partout dans $[0,\tau_\star]$ en termes de ses moyennes locales associées à la grille. Si on intègre, pour $i\in [\![1,n]\!]$, chaque membre de (2.50) sur $[\tau_{n,i-1},\tau_{n,i}]$, si on divise par $h_{n,i}$, et si on pose

$$\mathsf{x}_n(j) := \frac{1}{h_{n,j}}\int_{\tau_{n,j-1}}^{\tau_{n,j}}\varphi_n(\sigma)\,d\sigma, \quad j\in [\![1,n]\!], \tag{2.51}$$

alors x_n résout le système linéaire

$$(\mathsf{A}_n - z\mathsf{I}_n)\mathsf{x}_n = \mathsf{b}_n, \tag{2.52}$$

où

$$\mathsf{A}_n(i,j) := \frac{h_{n,j}}{h_{n,i}}\int_{\tau_{n,i-1}}^{\tau_{n,i}}\kappa(\tau,\widehat{\tau}_{n,j})\,d\tau, \tag{2.53}$$

$$\mathsf{b}_n(i) := \frac{1}{h_{n,i}}\int_{\tau_{n,i-1}}^{\tau_{n,i}} f(\sigma)\,d\sigma. \tag{2.54}$$

Afin de construire une estimation de l'erreur relative, on remarque que φ_n résout

$$(T_n - zI)\varphi_n = f, \tag{2.55}$$

où

$$T_n x := \sum_{j=1}^{n}\kappa(\,\cdot\,,\widehat{\tau}_{n,j})\int_{\tau_{n,j-1}}^{\tau_{n,j}} x(\sigma)\,d\sigma \quad \text{p.p. dans } [0,\tau_\star]. \tag{2.56}$$

On identifie la famille de fonctions $(e_{n,j})_{j=1}^n$ dans $L^1([0,\tau_\star])$ à

$$e_{n,j} := h_{n,j}\kappa(\,\cdot\,,\widehat{\tau}_{n,j}), \text{ p.p. sur } [0,\tau_\star], \tag{2.57}$$

et la famille de fonctionnelles linéaires-conjuguées $(\ell_{n,j})_{j=1}^n$ sur $L^1([0,\tau_\star])$ à

$$\langle x\,,\,\ell_{n,j}\rangle := \frac{1}{h_{n,j}}\int_{\tau_{n,j-1}}^{\tau_{n,j}} x(\sigma)\,d\sigma \quad \text{pour tout } x\in L^1([0,\tau_\star]). \tag{2.58}$$

On retrouve encore la représentation (2.2).

Théorème 12 *Si h_n tend vers 0 lorsque n tend vers l'infini, alors $(T_n)_{n\geq 1}$ est uniformément convergente vers T dans $L^1([0,\tau_\star])$.*

Preuve : soient $x \in L^1([0,\tau_\star])$ et $n > 0$. Alors

$$\begin{aligned}
\|(T_n - T)x\|_1 &= \int_0^{\tau_\star}\left|\int_0^{\tau_\star}\kappa(\tau,\sigma)x(\sigma)\,d\sigma - \sum_{j=1}^{n}\kappa(\tau,\widehat{\tau}_{n,j})\int_{\tau_{n,j-1}}^{\tau_{n,j}} x(\sigma)\,d\sigma\right| d\tau \\
&\leq \int_0^{\tau_\star}\Big(\sum_{j=1}^{n}\int_{\tau_{n,j-1}}^{\tau_{n,j}} |\kappa(\tau,\sigma)-\kappa(\tau,\widehat{\tau}_{n,j})||x(\sigma)|\,d\sigma\Big)\,d\tau \\
&\leq \sum_{j=1}^{n}\int_{\tau_{n,j-1}}^{\tau_{n,j}} |x(\sigma)|\Big[\int_0^{\tau_\star}|\kappa(\tau,\sigma)-\kappa(\tau,\widehat{\tau}_{n,j})|\,d\tau\Big]\,d\sigma.
\end{aligned}$$

Ainsi, pour tout $x \in L^1([0,\tau_\star])$,

$$\|(T_n - T)x\|_1 \leq \varepsilon_2(\kappa, h_n)\|x\|_1, \tag{2.59}$$

et donc si x est tel que $\|x\|_1 \leq 1$,

$$\|(T_n - T)x\|_1 \leq \varepsilon_2(\kappa, h_n).$$

Supposons que $\lim_{n\to\infty} h_n = 0$. Alors, d'après le Théorème 5, $\|(T_n - T)x\|_1$ tend uniformément vers 0 par rapport à x lorsque n tend vers l'infini, ce qui termine la preuve. ∎

Corollaire 4 *Soient T défini par (1.6) et $(T_n)_{n\geq 1}$ définie par (2.56). Alors il existe $n_1 > 0$ tel que*

$$c_1(z) := \sup_{n\geq n_1}\|R_n(z)\|_1 < +\infty. \tag{2.60}$$

Preuve : conséquence directe des Théorèmes 12, 2 et 4. ∎

De plus, d'après le Corollaire 2, on obtient, pour $n \geq n_1$,

$$\|\varphi_n - \varphi\|_1 \leq c_1(z)\|(T - T_n)\varphi\|_1. \tag{2.61}$$

Théorème 13 *Soient φ la solution de (1.15) avec T défini par (1.13) et $f \neq 0$. Soit φ_n la solution de (2.3) avec T_n défini par (2.56). Alors $\varphi \neq 0$ et pour n assez grand,*

$$\frac{\|\varphi_n - \varphi\|_1}{\|\varphi\|_1} \leq c_1(z)(4\|\eta\|_\infty \int_0^{h_n} g(\sigma)\,d\sigma + \gamma\|\omega_{\infty,2}(\eta, h_n)\|_\infty), \tag{2.62}$$

où $c_1(z)$ est donnée par (2.60).

Preuve : on applique l'inégalité (2.59) à (2.61). On obtient la borne grâce à l'inégalité (1.32) de la preuve du Théorème 5. ∎

Remarque 9 *On retrouve une borne relative du type* (2.21) *comme dans le cas des approximations par projection dans l'espace* $C^0([0,\tau_\star])$.

De plus, d'après (1.5), (2.59) *et l'estimation* (1.32), *la borne* (2.62) *est aussi une borne pour le résidu relatif de* φ_n *associé à l'équation* (1.1).

Commentaires bibliographiques

Dans [7] on démontre qu'un opérateur de rang fini peut toujours s'écrire de la façon (2.2).

Deux autres approximations de rang fini peuvent être construites à partir d'une famille de projections bornée. Le choix $T_n := T\pi_n$ est appelé approximation de Sloan ([72]). Le choix $T_n = \pi_n T \pi_n$ est appelé approximation de Galerkin. Les approximations d'un opérateur compact par projection construites plus haut ont l'avantage d'être uniformes, contrairement aux approximations de Sloan et de Galerkin qui, en général, sont seulement ν-convergentes. On trouve dans [48] la description d'autres types d'approximations d'opérateurs intégraux.

2.2 Raffinement itératif de solutions approchées

2.2.1 Trois schémas de raffinement

Plus la précision demandée sur la solution approchée φ_n est grande, c'est-à-dire plus le pas h_n de la grille est petit, plus la dimension du système correspondant est importante. Sa taille peut d'ailleurs devenir vite prohibitive d'un point de vue numérique et, dans ce cas, non seulement la stabilité des algorithmes de résolution de systèmes linéaires peut être compromise, mais le conditionnement de la matrice peut se détériorer.

Les schémas de raffinement donnent une réponse à ce problème : ils nous permettent d'obtenir itérativement la solution exacte d'un grand système linéaire par le biais de la résolution d'une suite de systèmes linéaires de taille modérée et fixe.

Dans le cadre général des problèmes non linéaires, les méthodes de raffinement itératif ont été formalisées dans [74] qui a servi de référence principale aux propositions développées dans [5] pour être appliquées à des problèmes spectraux. Une synthèse de ces méthodes est retenue dans [26].

Les schémas de raffinement itératif que nous présentons ici sont de la forme

$$\left\{ \begin{array}{rcl} y^{(0)} & := & G_n(z)f, \\ y^{(k+1)} & := & y^{(k)} - G_n(z)(Ty^{(k)} - zy^{(k)} - f), \quad k \geq 0, \end{array} \right. \tag{2.63}$$

où $G_n(z)$ désigne une approximation de $R(z)$ qui peut dépendre éventuellement d'un paramètre entier n.

La façon la plus élémentaire de raffiner une solution approchée $\varphi_n = R_n(z)f$ est de choisir $G_n = R_n$ dans (2.63) :

$$\left\{ \begin{array}{rcl} x^{(0)} & := & \varphi_n = R_n(z)f, \\ & & \\ x^{(k+1)} & := & x^{(k)} - R_n(z)(Tx^{(k)} - zx^{(k)} - f), \quad k \geq 0, \end{array} \right. \tag{2.64}$$

comme le propose Atkinson dans [14] pour des équations intégrales de second espèce, à noyau continu, discrétisées par la méthode de Nyström.

Comme $R(z)$ satisfait les identités

$$R(z) = \frac{1}{z}(R(z)T - I) = \frac{1}{z}(T\,R(z) - I),$$

deux nouvelles approximations de $R(z)$ sont ainsi motivées :

$$\widetilde{R}_n(z) := \frac{1}{z}(R_n(z)T - I) \quad \text{et} \quad \widehat{R}_n(z) := \frac{1}{z}(TR_n(z) - I).$$

Ces approximations conduisent aux schémas de raffinement itératif suivants :

$$\left\{ \begin{array}{rcl} \widetilde{x}^{(0)} & := & \widetilde{R}_n(z)f, \\ & & \\ \widetilde{x}^{(k+1)} & := & \widetilde{x}^{(k)} - \widetilde{R}_n(z)(T\widetilde{x}^{(k)} - z\widetilde{x}^{(k)} - f), \quad k \geq 0, \end{array} \right. \tag{2.65}$$

qui constitue une formulation plus générale et plus abstraite de la méthode proposée par Brakhage dans [21] telle qu'elle est reprise par Atkinson dans [14], et

$$\left\{ \begin{array}{rcl} \widehat{x}^{(0)} & := & \widehat{R}_n(z)f, \\ & & \\ \widehat{x}^{(k+1)} & := & \widehat{x}^{(k)} - \widehat{R}_n(z)(T\widehat{x}^{(k)} - z\widehat{x}^{(k)} - f), \quad k \geq 0. \end{array} \right. \tag{2.66}$$

que l'on appellera «schéma (C)», comme l'indique le tableau suivant :

Schémas	$G_n(z)$
Atkinson (A) (2.64)	$R_n(z) := (T_n - zI)^{-1}$
Brakhage (B) (2.65)	$\widetilde{R}_n(z) := \dfrac{1}{z}(R_n(z)T - I)$
Schéma (C) (2.66)	$\widehat{R}_n(z) := \dfrac{1}{z}(TR_n(z) - I)$

Remarque 10 (Lien avec l'accélération) *Nous allons démontrer que la suite d'itérés* $(x^{(k)})_{k\geq 0}$ *produite par le schéma* (2.64) *correspond exactement à la suite d'itérés induite par un schéma d'accélération classique. En effet, si* $R_n(z)$ *existe, alors l'équation originale* $(T - zI)\varphi = f$ *peut s'écrire de la manière équivalente suivante :*

$$(I - R_n(z)(T_n - T))\varphi = R_n(z)f. \tag{2.67}$$

Aussi, si $R_n(z)$ *est telle que*

$$\|R_n(z)(T_n - T)\| < 1, \tag{2.68}$$

l'opérateur $I - R_n(z)(T_n - T)$ *admet une inverse bornée qui, grâce à la série de Neumann, satisfait*

$$\Big(I - R_n(z)(T_n - T)\Big)^{-1} = I + \sum_{j=1}^{\infty}[R_n(z)(T_n - T)]^j. \tag{2.69}$$

(On trouvera une preuve de ce résultat dans [48], un résultat plus général étant démontré dans [54].) Ainsi, sous l'hypothèse (2.68), *la solution* φ *est donnée par*

$$\varphi = R_n(z)f + \sum_{j=1}^{\infty}[R_n(z)(T_n - T)]^j R_n(z)f. \tag{2.70}$$

Cette dernière relation motive de manière naturelle le schéma d'accélération suivant :

$$\varphi^{(k)} = R_n(z)f + \sum_{j=1}^{k}[R_n(z)(T_n - T)]^j R_n(z)f, \quad k \geq 0. \tag{2.71}$$

Nous allons démontrer que les schémas (2.64) et (2.71) génèrent la même suite d'itérés. Soit $(x^{(k)})_{k\geq 0}$ la suite produite par le schéma (2.64). Il est clair que

$$x^{(0)} := R_n(z)f = \varphi^{(0)}.$$

Supposons que, pour $k \geq 0$ fixé, $x^{(k)} = \varphi^{(k)}$ selon (2.71), alors d'après (2.64)

$$\begin{aligned}
x^{(k+1)} &= (I - R_n(z)(T - zI))x^{(k)} + R_n(z)f \\
&= (I - R_n(z)(T - T_n + T_n - zI))\sum_{j=0}^{k}[R_n(z)(T_n - T)]^j R_n(z)f + R_n(z)f \\
&= \sum_{j=0}^{k}[R_n(z)(T_n - T)]^{j+1} R_n(z)f + R_n(z)f \\
&= \sum_{j=0}^{k+1}[R_n(z)(T_n - T)]^j R_n(z)f \\
&= \varphi^{(k+1)},
\end{aligned}$$

ce qui conclut la récurrence.

De même que le calcul d'un résidu qui tend vers 0, la résolution d'un système linéaire homogène peut être instable. Ainsi les théorèmes suivants sont intéressants pour la mise en œuvre effective de ces schémas.

Théorème 14 *Soit $\left(x^{(k)}\right)_{k\geq 0}$, la suite définie par (2.64). Alors, pour tout $k \geq 0$,*

$$x^{(k+1)} := x^{(0)} + R_n(z)f^{(k+1)},$$

où

$$f^{(k+1)} := (T_n - T)x^{(k)}, \quad k \geq 0.$$

Preuve : pour tout $k \geq 0$,

$$\begin{aligned}
x^{(k+1)} &:= x^{(k)} - R_n(z)(Tx^{(k)} - zx^{(k)} - f) \\
&= x^{(k)} - R_n(z)(T - T_n + T_n - zI)x^{(k)} + x^{(0)} \\
&= x^{(0)} + R_n(z)(T_n - T)x^{(k)} \\
&= x^{(0)} + R_n(z)f^{(k+1)},
\end{aligned}$$

ce que nous voulions démontrer. ∎

Théorème 15 *Soit* $\left(\widetilde{x}^{(k)}\right)_{k\geq 0}$, *la suite définie par* (2.65). *Alors, pour tout* $k \geq 0$,

$$\widetilde{x}^{(k+1)} = \widetilde{x}^{(0)} + \frac{1}{z} R_n(z) \widetilde{f}^{(k+1)},$$

où

$$\widetilde{f}^{(k+1)} := (T_n - T) T \widetilde{x}^{(k)}, \quad k \geq 0.$$

Preuve : comme $z \neq 0$, pour tout $k \geq 0$,

$$\begin{aligned}
\widetilde{x}^{(k+1)} &:= \widetilde{x}^{(k)} - \frac{1}{z}(R_n(z)T - I)(T\widetilde{x}^{(k)} - z\widetilde{x}^{(k)} - f) \\
&= \widetilde{x}^{(k)} + \widetilde{x}^{(0)} + \frac{1}{z}(I - R_n(z)T)(T - T_n + T_n - zI)\widetilde{x}^{(k)} \\
&= \widetilde{x}^{(k)} + \widetilde{x}^{(0)} + \frac{1}{z}(I - R_n(z)T)(T - T_n)\widetilde{x}^{(k)} + \frac{1}{z}(I - R_n(z)T)(T_n - zI)\widetilde{x}^{(k)} \\
&= \widetilde{x}^{(0)} + \frac{1}{z}(I - R_n(z)T)(T - T_n)\widetilde{x}^{(k)} + \frac{1}{z}T_n\widetilde{x}^{(k)} - \frac{1}{z}R_n(z)T(T_n - zI))\widetilde{x}^{(k)} \\
&= \widetilde{x}^{(0)} + \frac{1}{z}(T - T_n)\widetilde{x}^{(k)} - \frac{1}{z}R_n(z)\Big[T(T - T_n) + T(T_n - zI)\Big]\widetilde{x}^{(k)} + \frac{1}{z}T_n\widetilde{x}^{(k)} \\
&= \widetilde{x}^{(0)} + \frac{1}{z}T\widetilde{x}^{(k)} - \frac{1}{z}R_n(z)(T - zI)T\widetilde{x}^{(k)} \\
&= \widetilde{x}^{(0)} + \frac{1}{z}\Big[I - R_n(z)(T - zI)\Big]T\widetilde{x}^{(k)} \\
&= \widetilde{x}^{(0)} + \frac{1}{z}\Big[R_n(z)(T_n - zI) - R_n(z)(T - zI)\Big]T\widetilde{x}^{(k)} \\
&= \widetilde{x}^{(0)} + \frac{1}{z}R_n(z)(T_n - T)T\widetilde{x}^{(k)} \\
&= \widetilde{x}^{(0)} + \frac{1}{z}R_n(z)\widetilde{f}^{(k+1)},
\end{aligned}$$

ce que nous voulions démontrer. ∎

Théorème 16 *Soit* $\left(\widehat{x}^{(k)}\right)_{k\geq 0}$, *la suite définie par* (2.66). *Alors pour tout* $k \geq 0$,

$$\widehat{x}^{(k+1)} = \widehat{x}^{(0)} + \frac{1}{z} T R_n(z) \widehat{f}^{(k+1)},$$

où

$$\widehat{f}^{(k+1)} := (T_n - T) \widehat{x}^{(k)}, \quad k \geq 0.$$

Preuve : comme $z \neq 0$, pour tout $k \geq 0$,

$$\begin{aligned}
\widehat{x}^{(k+1)} &:= \widehat{x}^{(k)} - \frac{1}{z}(TR_n(z) - I)(T\widehat{x}^{(k)} - z\widehat{x}^{(k)} - f) \\
&= \widehat{x}^{(k)} + \frac{1}{z}(I - TR_n(z))(T - T_n + T_n - zI)\widehat{x}^{(k)} + \widehat{x}^{(0)} \\
&= \widehat{x}^{(k)} + \frac{1}{z}(I - TR_n(z))(T - T_n)\widehat{x}^{(k)} + \widehat{x}^{(0)} + \frac{1}{z}(I - TR_n(z))(T_n - zI)\widehat{x}^{(k)} \\
&= \widehat{x}^{(0)} + \widehat{x}^{(k)} + \frac{1}{z}(I - TR_n(z))(T - T_n)\widehat{x}^{(k)} + \frac{1}{z}(T_n - zI)\widehat{x}^{(k)} - \frac{1}{z}T\widehat{x}^{(k)} \\
&= \widehat{x}^{(0)} + \frac{1}{z}(I - TR_n(z))(T - T_n)\widehat{x}^{(k)} + \frac{1}{z}(T_n - T)\widehat{x}^{(k)} \\
&= \widehat{x}^{(0)} - \frac{1}{z}TR_n(z)(T - T_n)\widehat{x}^{(k)} \\
&= \widehat{x}^{(0)} + \frac{1}{z}TR_n(z)\widehat{f}^{(k+1)},
\end{aligned}$$

ce que nous voulions démontrer. ∎

2.2.2 Convergence et bornes d'erreurs

Donnons tout d'abord quelques estimations générales pour les erreurs relatives entre les itérés donnés par chaque schéma et la solution exacte du problème.

Théorème 17 *Soient* $(x^{(k)})_{k\geq 0}$ *la suite définie par* (2.64) *et* $f \neq 0$*. Alors* $\varphi \neq 0$*, et pour tout* $k \geq 1$*,*

$$\frac{\|x^{(k)} - \varphi\|}{\|\varphi\|} \leq \|[R_n(z)(T_n - T)]^{k+1}\|. \tag{2.72}$$

Preuve : on remarque tout d'abord que

$$\begin{aligned}
x^{(0)} - \varphi &= R_n(z)f - \varphi \\
&= (R_n(z)(T - zI) - I)\varphi \\
&= (R_n(z)T - zR_n(z) - I)\varphi \\
&= (R_n(z)T - R_n(z)T_n)\varphi \\
&= R_n(z)(T - T_n)\varphi.
\end{aligned}$$

Ainsi, d'après le Théorème 14 et par récurrence sur $k \geq 0$,

$$x^{(k)} - \varphi = \left(R_n(z)(T_n - T)\right)^k (x^{(0)} - \varphi).$$

∎

Théorème 18 *Soient* $(\widetilde{x}^{(k)})_{k\geq 0}$ *la suite définie par* (2.65) *et* $f \neq 0$. *Alors* $\varphi \neq 0$, *et pour tout* $k \geq 1$,

$$\frac{\|\widetilde{x}^{(k)} - \varphi\|}{\|\varphi\|} \leq \left\| \left[\frac{1}{z} R_n(z)(T_n - T)T \right]^{k+1} \right\|. \tag{2.73}$$

Preuve : on remarque tout d'abord que

$$\begin{aligned}
\widetilde{x}^{(0)} - \varphi &= \frac{1}{z}(R_n(z)T - I)(T - zI)\varphi - \varphi \\
&= \frac{1}{z}(R_n(z)TT - zR_n(z)T - T + zI)\varphi - \varphi \\
&= \frac{1}{z}(R_n(z)TT - (zR_n(z) + I)T + zI)\varphi - \varphi \\
&= \frac{1}{z}(R_n(z)TT - R_n(z)T_nT + zI)\varphi - \varphi \\
&= \frac{1}{z}R_n(z)(T - T_n)T\varphi.
\end{aligned}$$

Ainsi, d'après le Théorème 15 et par récurrence sur $k \geq 0$,

$$\widetilde{x}^{(k)} - \varphi = \left(\frac{1}{z} R_n(z)(T_n - T)T \right)^k (\widetilde{x}^{(0)} - \varphi),$$

d'où la majoration recherchée. ∎

Théorème 19 *Soient* $(\widehat{x}^{(k)})_{k\geq 0}$ *la suite définie par* (2.66) *et* $f \neq 0$. *Alors* $\varphi \neq 0$, *et pour tout* $k \geq 1$,

$$\frac{\|\widehat{x}^{(k)} - \varphi\|}{\|\varphi\|} \leq \left\| \left[\frac{1}{z} T R_n(z)(T_n - T) \right]^{k+1} \right\|. \tag{2.74}$$

Preuve : on remarque tout d'abord que

$$\begin{aligned}
\widehat{x}^{(0)} - \varphi &= \frac{1}{z}(TR_n(z) - I)(T - zI)\varphi - \varphi \\
&= \frac{1}{z}(TR_n(z)T - zTR_n(z) - T + zI)\varphi - \varphi \\
&= \frac{1}{z}(TR_n(z)T - T(zR_n(z) + I))\varphi - \varphi \\
&= \frac{1}{z}(TR_n(z)T - TR_n(z)T_n)\varphi \\
&= \frac{1}{z}TR_n(z)(T - T_n)\varphi.
\end{aligned}$$

Ainsi, d'après le Théorème 16 et par récurrence sur $k \geq 0$,

$$\widehat{x}^{(k)} - \varphi = \left(\frac{1}{z} T R_n(z)(T_n - T) \right)^k (\widehat{x}^{(0)} - \varphi),$$

d'où la majoration recherchée. ∎

Ces derniers résultats conduisent aux théorèmes suivants qui donnent des estimations de l'erreur relative pour chaque schéma de raffinement appliqués à des approximations par projection et par intégration produit, et ceci dans le cas des espaces C^0 et L^1. On rappelle que

$$\max\{\|T\|_\infty, \|T\|_1\} \leq \gamma \|\eta\|_\infty. \tag{2.75}$$

On définit de plus

$$\alpha_1(h_n) := \|\eta\|_\infty \varepsilon_1(g, h_n) + \gamma \|\omega_{\infty,1}(\eta, h_n)\|_\infty. \tag{2.76}$$

Lorsque la fonction η est constante par rapport à sa première variable (ce qui est la cas dans certaines de nos applications astrophysiques),

$$\alpha_1(h_n) \leq \|\eta\|_\infty \varepsilon_1(g, h_n) = O\left(\int_0^{h_n} g(\sigma)\, d\sigma \right).$$

2.2.2.1 Approximation par projection

Dans le cas des approximations par projection, $(T_n)_{n\geq 1}$ converge en norme vers T à la fois dans C^0 et L^1 . Ainsi, on peut utiliser le fait que

$$\|R_n(z)(T_n - T)^{k+1}\| \leq [c_0(z)\|T_n - T\|]^{k+1}, \quad k \geq 0. \tag{2.77}$$

On rappelle (voir les Lemmes 8 et 11) que dans le cas des approximations par projection,

$$\|T_n - T\|_\infty \leq \varepsilon_1(\kappa, h_n) \leq \alpha_1(h_n), \tag{2.78}$$

$$\|T_n - T\|_1 \leq \frac{2}{q_n} \alpha_1(h_n). \tag{2.79}$$

Théorème 20 *Soient $(x^{(k)})_{k\geq 0}$, $(\widetilde{x}^{(k)})_{k\geq 0}$ et $(\widehat{x}^{(k)})_{k\geq 0}$ les trois suites définies par (2.64), (2.65) et (2.66) respectivement, et $f \in C^0([0, \tau_\star])$. Si $f \neq 0$, alors $\varphi \neq 0$ et*

$$\max\left\{ \frac{\|x^{(k)} - \varphi\|_\infty}{\|\varphi\|_\infty}, \frac{\|\widetilde{x}^{(k)} - \varphi\|_\infty}{\|\varphi\|_\infty}, \frac{\|\widehat{x}^{(k)} - \varphi\|_\infty}{\|\varphi\|_\infty} \right\} = O\left(\alpha_1(h_n)^{k+1} \right). \tag{2.80}$$

Preuve : c'est une conséquence des Théorèmes 17, 18 et 19, et des inégalités (2.77), (2.78) et (2.75). ∎

Théorème 21 *Soient* $(x^{(k)})_{k\geq 0}$, $(\widetilde{x}^{(k)})_{k\geq 0}$ *et* $(\widehat{x}^{(k)})_{k\geq 0}$ *les suites définies par* (2.64), (2.65) *et* (2.66) *respectivement, et* $f \in L^1([0,\tau_\star])$. *Si* $f \neq 0$, *alors* $\varphi \neq 0$ *et*

(2.81)
$$\max\left\{\frac{\|x^{(k)}-\varphi\|_1}{\|\varphi\|_1}, \frac{\|\widetilde{x}^{(k)}-\varphi\|_1}{\|\varphi\|_1}, \frac{\|\widehat{x}^{(k)}-\varphi\|_1}{\|\varphi\|_1}\right\} = O\left(\left(\frac{1}{q_n}\alpha_1(h_n)\right)^{k+1}\right).$$

Preuve : c'est une conséquence des Théorèmes 17, 18 et 19, et des inégalités (2.77), (2.79) et (2.75). ∎

2.2.2.2 Approximation par intégration produit

Cas de l'espace C^0

Dans le cas des approximations par intégration produit dans $C^0([0,\tau_0])$, $(T_n)_{n\geq 1}$ converge de façon collectivement compacte vers T lorsqu'il est défini par (2.42). Alors $((T_n - T)T)_{n\geq 1}$ et $((T_n - T)T_n)_{n\geq 1}$ sont convergentes en norme vers 0, car T est compact. Puisque

$$\begin{aligned}\Big(R_n(z)(T_n - T)\Big)^2 &= R_n(z)(T_n - T)R_n(z)(T_n - T)\\ &= R_n(z)(T_n - T)R_n(z)T_n + R_n(z)(T_n - T)R_n(z)T\\ &= R_n(z)[(T_n - T)T_nR_n(z) + (T - T_n)R_n(z)T],\end{aligned}$$

et que

$$\begin{aligned}(T - T_n)R_n(z)T &= (T - T_n)(R_n(z) - R(z))T + (T - T_n)R(z)T\\ &= (T - T_n)R_n(z)(T - T_n)R(z)T + (T - T_n)TR(z)\\ &= (T - T_n)R_n(z)(T - T_n)TR(z) + (T - T_n)TR(z),\end{aligned}$$

alors

$$\begin{aligned}&\left\|\Big(R_n(z)(T_n - T)\Big)^2\right\| \leq \\ &c_0(z)\Big[c_0(z)\|(T - T_n)T_n\|\Big(c_0(z)\|R(z)(\|T\| + \|T_n\|) + \|R(z)\|\Big)\|(T_n - T)T\|\Big].\end{aligned} \tag{2.82}$$

Le principe de la Borne Uniforme implique que la suite $(T_n)_{n\geq 1}$ est bornée car elle converge ponctuellement. De plus, dans le cas des approximations par intégration produit dans C^0 on rappelle (voir preuve du Théorème 10) que

$$\|(T_n - T)x\|_\infty \leq \omega_\infty(x, h_n)\|T\|_\infty + \gamma\|x\|_\infty\|\omega_{\infty,2}(\eta, h_n)\|_\infty, \tag{2.83}$$

et les inégalités (1.38) du Lemme 4 et (2.46):

$$\begin{aligned} \omega_\infty(Tx, h_n) &\leq \varepsilon_1(\kappa, h_n)\|x\|_\infty \leq \alpha_1(h_n)\|x\|_\infty, \\ \omega_\infty(T_n x, h_n) &\leq (\gamma\|\omega_{\infty,1}(\eta, h_n)\|_\infty + \|\eta\|_\infty\varepsilon_1(g, h_n))\|x\|_\infty = \alpha_1(h_n)\|x\|_\infty. \end{aligned}$$

Donc

$$\max\{\|(T - T_n)T\|_\infty, \|(T - T_n)T_n\|_\infty\} \leq \alpha_1(h_n)\|T\|_\infty + \gamma\|\omega_{\infty,2}(\eta, h_n)\|_\infty. \tag{2.84}$$

On définit

$$\beta_1(h_n) := \alpha_1(h_n)\|T\|_\infty + \gamma\|\omega_{\infty,2}(\eta, h_n)\|_\infty. \tag{2.85}$$

Finalement,

$$\begin{aligned} &\left\|\Big(R_n(z)(T_n - T)\Big)^2\right\|_\infty \leq \\ &\qquad c_0(z)\beta_1(h_n)\left[c_0(z) + \|R(z)\|_\infty(1 + 2c_0(z)\gamma\|\eta\|_\infty)\right] = O\left(\beta_1(h_n)\right). \end{aligned} \tag{2.86}$$

Nous obtenons le théorème suivant pour l'estimation de l'erreur pour chaque schéma de raffinement, où [.] désigne la fonction partie entière.

Théorème 22 *Soient $(x^{(k)})_{k\geq 0}$ la suite définie par (2.64) et $f \in C^0([0, \tau_\star])$. Si $f \neq 0$, alors $\varphi \neq 0$,*

$$\frac{\|x^{(0)} - \varphi\|_\infty}{\|\varphi\|_\infty} = O(\alpha(h_n)) + O\left(\frac{\omega(f, h_n)}{\|f\|_\infty}\right) + O(\|\omega_{\infty,1}(\eta, h_n)\|_\infty), \tag{2.87}$$

et, pour tout $k \geq 1$,

$$\frac{\|x^{(k)} - \varphi\|_\infty}{\|\varphi\|_\infty} \leq O\left(\beta_1(h_n)^{\left[\frac{k+1}{2}\right]}\right). \tag{2.88}$$

Preuve : si $k=0$, alors $x^{(0)}-\varphi=\varphi_n-\varphi$ et l'estimation de l'erreur relative est donnée par la borne (2.49) du Théorème 11.
Si $k=2l-1$ pour $l>0$, alors d'après le Théorème 17,

$$\begin{aligned}\frac{\|x^{(2l-1)}-\varphi\|_\infty}{\|\varphi\|_\infty} &\leq \|\Big(R_n(z)(T_n-T)\Big)^{2l}\|_\infty \\ &\leq \|\Big(R_n(z)(T_n-T)\Big)^2\|_\infty^l.\end{aligned}$$

Si $k=2l$ pour $l>0$,

$$\begin{aligned}\frac{\|x^{(2l)}-\varphi\|_\infty}{\|\varphi\|_\infty} &\leq \|\Big(R_n(z)(T_n-T)\Big)^{2l+1}\|_\infty \\ &\leq \|R_n(z)\|_\infty(\|T\|_\infty+\|T_n\|_\infty)\|\Big(R_n(z)(T_n-T)\Big)^2\|_\infty^l \\ &\leq 2c_0(z)\|\eta\|_\infty\|\Big(R_n(z)(T_n-T)\Big)^2\|_\infty^l,\end{aligned}$$

d'après (1.33) et (2.43). Alors d'après (2.86), pour tout $k\geq 1$,

$$\frac{\|x^{(k)}-\varphi\|_\infty}{\|\varphi\|_\infty}=O\left(\beta_1(h_n)^{[\frac{k+1}{2}]}\right),$$

ce que nous voulions prouver. ∎

Théorème 23 *Soient $(\widetilde{x}^{(k)})_{k\geq 0}$ la suite définie par (2.65) et $f\in C^0([0,\tau_\star])$. Si $f\neq 0$, alors $\varphi\neq 0$ et, pour $k\geq 0$,*

$$\frac{\|\widetilde{x}^{(k)}-\varphi\|_\infty}{\|\varphi\|_\infty}=O\left(\beta_1(h_n)^{k+1}\right). \tag{2.89}$$

Preuve : si $k\geq 0$, alors d'après le Théorème 18

$$\frac{\|\widetilde{x}^{(k)}-\varphi\|_\infty}{\|\varphi\|_\infty}\leq\left(\frac{c_0(z)}{|z|}\right)^{k+1}\|(T_n-T)T\|_\infty^{k+1}.$$

La borne provient de l'inégalité (2.84). ∎

Théorème 24 *Soient $(\widehat{x}^{(k)})_{k\geq 0}$ la suite définie par (2.66) et $f\in C^0([0,\tau_\star])$. Si $f\neq 0$, alors $\varphi\neq 0$,*

(2.90)
$$\frac{\|\widehat{x}^{(0)}-\varphi\|_\infty}{\|\varphi\|_\infty}=O(\alpha_1(h_n))+O\left(\frac{\omega(f,h_n)}{\|f\|_\infty}\right)+O(\|\omega_{\infty,2}(\eta,h_n)\|_\infty),$$

et, pour $k \geq 1$,

$$\frac{\|\widehat{x}^{(k)} - \varphi\|_\infty}{\|\varphi\|_\infty} = O\left(\beta_1(h_n)^{[\frac{k+1}{2}]}\right). \tag{2.91}$$

Preuve : $\widehat{x}^{(0)} - \varphi = \frac{1}{z} T R_n(z)(T - T_n)\varphi = \frac{1}{z} R_n(z) T(\varphi_n - \varphi)$, et l'estimation de l'erreur relative est donnée par la borne (2.49) du Théorème 11.
Si $k = 2l - 1$ pour $l > 0$, alors

$$\frac{\|\widehat{x}^{(2l-1)} - \varphi\|_\infty}{\|\varphi\|_\infty} \leq \frac{1}{|z|^{2l}} \|T\|_\infty^{2l} \|\Big(R_n(z)(T_n - T)\Big)^2\|_\infty^l.$$

Si $k = 2l$ pour $l > 0$, alors

$$\begin{aligned}
\frac{\|\widehat{x}^{(2l)} - \varphi\|_\infty}{\|\varphi\|_\infty} &\leq \frac{1}{|z|^{2l+1}} \|T R_n(z)(T_n - T)\|_\infty \|\Big(T R_n(z)(T_n - T)\Big)^{2l}\|_\infty \\
&\leq \frac{2}{|z|^{2l+1}} c_0(z)(\gamma \|\eta\|_\infty^{2l+2} \|\Big(R_n(z)(T_n - T)\Big)^2\|_\infty^l,
\end{aligned}$$

d'après (1.33) et (2.43). Alors d'après (2.86), pour tout $k \geq 1$,

$$\frac{\|\widehat{x}^{(k)} - \varphi\|_\infty}{\|\varphi\|_\infty} = O\left(\beta_1(h_n)^{[\frac{k+1}{2}]}\right),$$

ce que ne voulions prouver. ∎

Cas de l'espace L^1

Dans le cas des approximations par intégration produit dans $L^1([0, \tau_0])$, $(T_n)_{n \geq 1}$ est convergente en norme vers T lorsqu'il est défini par (2.56). Ainsi, nous pouvons encore utiliser l'inégalité (2.77). On rappelle de plus (voir preuve du Théorème 12) que dans ce cas,

$$\|T - T_n\|_1 \leq \varepsilon_2(\kappa, h_n). \tag{2.92}$$

Lorsque la fonction η est constante par rapport à sa deuxième variable (ce qui est la cas dans certaines de nos applications astrophysiques),

$$\varepsilon_2(\kappa, h_n) \leq \|\eta\|_\infty \varepsilon_1(g, h_n) = O\left(\int_0^{h_n} g(\sigma)\, d\sigma\right).$$

Théorème 25 *Soient* $(x^{(k)})_{k\geq 0}$, $(\widetilde{x}^{(k)})_{k\geq 0}$ *et* $(\widehat{x}^{(k)})_{k\geq 0}$ *les suites définies par* (2.64), (2.65) *et* (2.66) *respectivement, et* $f \in L^1([0,\tau_*])$. *Si* $f \neq 0$, *alors* $\varphi \neq 0$ *et*

(2.93)
$$\max\left\{\frac{\|x^{(k)}-\varphi\|_1}{\|\varphi\|_1}, \frac{\|\widetilde{x}^{(k)}-\varphi\|_1}{\|\varphi\|_1}, \frac{\|\widehat{x}^{(k)}-\varphi\|_1}{\|\varphi\|_1}\right\} = O\left(\varepsilon_2(\kappa,h_n)^{k+1}\right).$$

Preuve : c'est une conséquence des Théorèmes 17, 18 et 19, et des inégalités (2.77), (2.92) et (2.75). ∎

2.2.3 Tests d'arrêt

Les schémas de raffinement itératifs présentés ici construisent des suites qui convergent vers la solution exacte du problème; il est donc légitime de se poser la question du test d'arrêt des algorithmes correspondants avant leur mise en œuvre, ces tests pouvant porter sur l'erreur ou sur le résidu. La solution exacte n'étant pas connue en général, la première possibilité impliquerait donc le calcul, à chaque itération, des estimations de l'erreur décrites plus haut. Or, dans [8] , où la méthode d'approximation de rang fini a été appliquée à des problèmes admettant des solutions exactes connues, il apparaît que les bornes d'erreurs sont, dans certains cas, pessimistes. Si nous choisissons donc de porter le test d'arrêt sur l'erreur en utilisant les estimations précédentes, nous risquons de demander à l'algorithme plus d'itérations qu'il ne lui en faut en réalité pour obtenir une précision désirée. Nous avons donc choisi de porter le test d'arrêt sur le résidu et nous allons justifier ce choix plus en détail.

On rappelle que les schémas de raffinement itératif s'écrivent de manière générale

$$\begin{aligned} y^{(0)} &= G_n(z)f, \\ y^{(k+1)} &= y^{(0)} + (I - G_n(z)(T-zI))y^{(k)}, \end{aligned}$$

où $G_n(z)$ joue le rôle d'une approximation de l'inverse de l'opérateur $T - zI$, dans le sens où l'on suppose que $\|I - G_n(z)(T-zI)\| < 1$. La solution exacte φ du problème $(T-zI)\varphi = f$, vérifie de façon évidente l'égalité suivante :

$$\varphi = y^{(0)} + (I - G_n(z)(T-zI))\varphi.$$

Désignons par $e^{(k)}$, respectivement par $r^{(k)}$, l'erreur, respectivement le résidu, associés à l'approximation $y^{(k)}$ produite par le schéma :

$$\begin{aligned} e^{(k)} &:= y^{(k)} - \varphi, \\ r^{(k)} &:= (T-zI)y^{(k)} - f. \end{aligned}$$

Ainsi

$$e^{(0)} = -(I - G_n(z)(T - zI))\varphi,$$

et, pour tout $k \geq 0$,

$$e^{(k+1)} := y^{(k+1)} - \varphi = (I - G_n(z)(T - zI))e^{(k)} = -(I - G_n(z)(T - zI))^{k+1}e^{(0)}.$$

Le taux de convergence τ_e de la suite $\left(e^{(k)}\right)_{k\geq 0}$ est donc

$$\tau_e = \|I - G_n(z)(T - zI)\|. \tag{2.94}$$

D'autre part

$$r^{(0)} = (T - zI)y^{(0)} - f = [(T - zI)G_n(z) - I]f,$$

et, pour tout $k \geq 0$,

$$\begin{aligned} r^{(k+1)} &:= (T - zI)y^{(k+1)} - f \\ &= (T - zI)(G_n(z)f + (I - Gn(z)(T - zI))y^k - f \\ &= (T - zI)G_n(z)(f - (T - zI)y^k) + (T - zI)y^{(k)} - f \\ &= -(T - zI)G_n(z)r^{(k)} + r^{(k)} \\ &= [I - (T - zI)G_n(z)]r^{(k)} \\ &= -[I - (T - zI)G_n(z)]^{k+1}r^{(0)}. \end{aligned}$$

Le taux de convergence τ_r de la suite $\left(r^{(k)}\right)_{k\geq 0}$ est donc

$$\tau_r = \|I - (T - zI)G_n(z)\|.$$

On définit les erreurs et résidus relatifs correspondants par

$$\epsilon^{(k)} := \frac{\|e^{(k)}\|}{\|\varphi\|}, \tag{2.95}$$

$$\rho^{(k)} := \frac{\|r^{(k)}\|}{\|f\|}. \tag{2.96}$$

Ainsi,

$$\epsilon^{(k)} \leq \|(I - G_n(z)(T - zI))^{k+1}\|$$

et

$$\rho^{(k)} \leq \|(I - (T - zI)G_n(z))^{k+1}\|.$$

Les suites $\left(\epsilon^{(k)}\right)_{k\geq 0}$ et $\left(\rho^{(k)}\right)_{k\geq 0}$ sont donc dominées par une même suite.

Remarque 11 *Par abus de langage, on appellera «résidu relatif» à la fois la fonction* $\tau \mapsto \dfrac{Ty^{(k)}(\tau) - zy^{(k)}(\tau) - f(\tau)}{\|f\|}$ *et sa norme* $\rho^{(k)}$.

2.2.4 Mise en œuvre

Chaque schéma demande l'évaluation de T en certains points de X. En pratique T n'est pas utilisé dans ce but et un opérateur approché T_m de la famille $(T_\nu)_{\nu\geq 1}$ est préféré, où $m > n$. Le schéma correspondant convergera donc vers la solution exacte φ_m du problème approché

$$(T_m - zI)\varphi_m = f. \tag{2.97}$$

Dans la suite, aucune comparaison n'est faite entre les solutions données par ces schémas. Afin d'alléger le texte, nous utiliserons les mêmes notations, introduites dans la section suivante, pour écrire les différents fonctions, vecteurs et matrices intervenant dans la mise en œuvre de chacun des trois schémas.

2.2.4.1 Notations et définitions

Chaque schéma de raffinement admet une formulation équivalente en terme de deux familles de fonctions auxiliaires $(f^{(k)})_{k\geq 0}$ et $(v^{(k)})_{k\geq 0}$. Nous désignons par $x^{(k)}$ le kième itéré construit par un schéma donné, et pour $k \geq 0$ on définit les vecteurs

$$\begin{array}{llll}
\mathsf{x}_n^{(k)} \in \mathbb{C}^{n\times 1} & \text{tel que, } \forall i \in [\![1, n]\!], & \mathsf{x}_n^{(k)}(i) & := \langle x^{(k)}\,,\, \ell_{n,i}\rangle, \\
\mathsf{x}_m^{(k)} \in \mathbb{C}^{m\times 1} & \text{tel que, } \forall i \in [\![1, m]\!], & \mathsf{x}_m^{(k)}(i) & := \langle x^{(k)}\,,\, \ell_{m,i}\rangle, \\
\mathsf{v}_n^{(k)} \in \mathbb{C}^{n\times 1} & \text{tel que, } \forall i \in [\![1, n]\!], & \mathsf{v}_n^{(k)}(i) & := \langle v^{(k+1)}\,,\, \ell_{n,i}\rangle, \\
\mathsf{b}_n^{(k)} \in \mathbb{C}^{n\times 1} & \text{tel que, } \forall i \in [\![1, n]\!], & \mathsf{b}_n^{(k)}(i) & := \langle f^{(k)}\,,\, \ell_{n,i}\rangle, \\
\mathsf{b}_m \in \mathbb{C}^{m\times 1} & \text{tel que, } \forall i \in [\![1, m]\!], & \mathsf{b}_m(i) & := \langle f\,,\, \ell_{m,i}\rangle,
\end{array}$$

où $(\ell_{n,i})_{i=1}^{n}$ et $(\ell_{m,i})_{i=1}^{m}$ sont les familles de fonctionnelles linéaires-conjuguées intervenant dans la définition des opérateurs approchés de rang fini T_n et T_m respectivement (voir section 2.1, définition (2.2)). Soient A la matrice d'approximation d'ordre n associée à T_n et B la matrice d'approximation d'ordre m associée à T_m. Alors

$$\begin{array}{lll} \mathsf{A} \in \mathbb{C}^{n\times n} & \text{est telle que, } \forall (i,j) \in [\![1,n]\!]^2, & \mathsf{A}(i,j) := \langle e_{n,j}\,,\,\ell_{n,i}\rangle, \\ \mathsf{B} \in \mathbb{C}^{m\times m} & \text{est telle que, } \forall (i,j) \in [\![1,m]\!]^2, & \mathsf{B}(i,j) := \langle e_{m,j}\,,\,\ell_{m,i}\rangle. \end{array}$$

2.2.4.2 Matrices de prolongement et de restriction

La mise en œuvre des schémas de raffinement (2.64), (2.65) et (2.66) fait intervenir les matrices suivantes :

$$\begin{array}{lll} \mathsf{C} \in \mathbb{C}^{n\times m} & \text{telle que, } \forall (i,j) \in [\![1,n]\!]\times[\![1,m]\!], & \mathsf{C}(i,j) := \langle e_{m,j}\,,\,\ell_{n,i}\rangle, \\ \mathsf{D} \in \mathbb{C}^{m\times n} & \text{telle que, } \forall (i,j) \in [\![1,m]\!]\times[\![1,n]\!], & \mathsf{D}(i,j) := \langle e_{n,j}\,,\,\ell_{m,i}\rangle. \end{array}$$

La détermination de leurs coefficients est simplifiée sous certaines hypothèses sur les familles de fonctions $(e_{n,j})_{j=1}^{n}$ et $(e_{m,j})_{j=1}^{m}$ et les familles de fonctionnelles linéaires-conjuguées $(\ell_{n,j})_{j=1}^{n}$ et $(\ell_{m,j})_{j=1}^{m}$. Supposons qu'il existe deux matrices $\mathsf{P} \in \mathbb{C}^{m\times n}$ et $\mathsf{R} \in \mathbb{C}^{n\times m}$ telles que

$$e_{n,j} = \sum_{k=1}^{m} \mathsf{P}(k,j)e_{m,k}, \quad \text{pour tout } j \in [\![1,n]\!], \tag{2.98}$$

$$\ell_{n,i} = \sum_{k=1}^{m} \mathsf{R}(i,k)\ell_{m,k}, \quad \text{pour tout } i \in [\![1,n]\!]. \tag{2.99}$$

Les matrices P et R sont appelées respectivement «matrice de prolongement» et «matrice de restriction». Alors, pour tout $(i,j) \in [\![1,m]\!]\times[\![1,n]\!]$,

$$\mathsf{D}(i,j) = \sum_{k=1}^{m} \mathsf{B}(i,k)\mathsf{P}(k,j),$$

et ainsi

$$\mathsf{D} = \mathsf{BP}.$$

De même, pour tout $(i,j) \in [\![1,n]\!]\times[\![1,m]\!]$, $\mathsf{C}(i,j) = \sum_{k=1}^{m} \mathsf{R}(i,k)\mathsf{B}(k,j)$, et ainsi

$$\mathsf{C} = \mathsf{RB}.$$

Nous allons maintenant détailler la construction des matrices de prolongement et de restriction dans le cas des approximations par projection et dans le cas des approximations par intégration produit appliquées dans chacun des espaces $C^0([0,\tau_\star])$

et $L^1([0,\tau_\star])$. Nous verrons que les hypothèses (2.98) et (2.99) sont satisfaites si on considère des grilles emboîtées.

Cas des approximations par projection

Soient π_n et π_m deux projections de rang fini définies par

$$\pi_n x := \sum_{j=1}^{n} \langle x\,,\,\xi_{n,j}\rangle e_{n,j}, \quad x \in X, \tag{2.100}$$

$$\pi_m x := \sum_{j=1}^{m} \langle x\,,\,e^*_{m,j}\rangle e_{m,j}, \quad x \in X, \tag{2.101}$$

où $(e_{n,j})_{j=1}^n$ est une base ordonnée de l'image de π_n et $(\xi_{n,j})_{j=1}^n$ une base adjointe. De même, $(e_{m,j})_{j=1}^m$ est une base ordonnée de l'image de π_m et $(e^*_{m,j})_{j=1}^m$ une base adjointe. On désigne par X_n et Y_n (resp. X_m et Y_m) l'espace image et le noyau de π_n (resp. π_m). Supposons que T_n et T_m soient construits à l'aide de ces projections de la manière suivante :

$$T_n := \pi_n T \qquad \text{et} \qquad T_m := \pi_m T. \tag{2.102}$$

Alors on identifie

$$\forall j \in [\![1, n]\!],\ \ell_{n,j} := T^*\xi_{n,j} \quad \text{et} \quad \forall k \in [\![1, m]\!],\ \ell_{m,k} := T^* e^*_{m,k}. \tag{2.103}$$

Supposons que $X_n \subseteq X_m$ et que $Y_m \subseteq Y_n$. Alors, pour tout $i \in [\![1, n]\!]$,

$$e_{n,i} = \pi_m e_{n,i} = \sum_{k=1}^{m} \langle e_{n,i}\,,\,e^*_{m,k}\rangle e_{m,k}, \tag{2.104}$$

$$\xi_{n,i} = \sum_{k=1}^{m} \langle e_{m,k}\,,\,\xi_{n,i}\rangle e^*_{m,k}. \tag{2.105}$$

En effet , soit $u \in X$, alors il existe un unique couple $(x_m, y_m) \in X_m \times Y_m$ tel que $u = x_m + y_m$. Alors, pour tout $i \in [\![1, n]\!]$,

$$\begin{aligned}
\langle u\,,\,\xi_{n,i}\rangle &= \langle x_m\,,\,\xi_{n,i}\rangle + \underbrace{\langle y_m\,,\,\xi_{n,i}\rangle}_{=0} \\
&= \sum_{k=1}^{m} \langle u\,,\,e^*_{m,k}\rangle \langle e_{m,k}\,,\,\xi_{n,i}\rangle.
\end{aligned}$$

Aussi, pour tout $i \in [\![1, n]\!]$, $\ell_{n,i} = \sum_{k=1}^{m} \langle e_{m,k}\,,\, \xi_{n,i}\rangle \ell_{m,k}$, et (2.104) est évidente. Enfin,

$$\forall (k,j) \in [\![1, m]\!] \times [\![1, n]\!], \qquad \mathsf{P}(k,j) = \langle e_{n,j}\,,\, e^*_{m,k}\rangle, \tag{2.106}$$

$$\forall (i,k) \in [\![1, n]\!] \times [\![1, m]\!], \qquad \mathsf{R}(i,k) = \langle e_{m,k}\,,\, \xi_{n,i}\rangle. \tag{2.107}$$

Cas où $X = C^0([0, \tau_\star])$

On considère deux grilles $(\tau_{n,j})_{j=1}^n$ et $(\tau_{m,j})_{j=1}^m$ du type (2.13), les fonctions $(e_{n,j})_{j=1}^n$ et $(e_{m,j})_{j=1}^m$ définies par (2.16) ainsi que les fonctionnelles $(\xi_{n,j})_{j=1}^n$ et $(\xi_{m,j})_{j=1}^m$ définies par (2.15) correspondantes. On note π_n et π_m les projections correspondantes. Supposons que ces grilles soient uniformes et telles que

$$r_0 := \frac{m-1}{n-1} \tag{2.108}$$

soit un entier naturel non nul. Alors $X_n \subseteq X_m$ et $Y_m \subseteq Y_n$. En fait, pour tout $j \in [\![1, n]\!]$,

$$e_{n,j} = \pi_m e_{n,j} = \sum_{k=1}^{m} e_{n,j}(\tau_{m,k}) e_{m,k}.$$

Matrice de prolongement

D'après (2.106), pour tout k dans $[\![1, m]\!]$ et tout j dans $[\![1, n]\!]$,

$$\mathsf{P}(k,j) = \langle e_{n,j}\,,\, e^*_{m,k}\rangle = e_{n,j}(\tau_{m,k}). \tag{2.109}$$

Matrice de restriction

D'après (2.107), pour tout i dans $[\![1, n]\!]$ et tout k dans $[\![1, m]\!]$,

$$\mathsf{R}(i,k) = \langle e_{m,k}\,,\, \xi_{n,i}\rangle = e_{m,k}(\tau_{n,i}).$$

C'est-à-dire, d'après (2.108), pour tout i dans $[\![1, n]\!]$ et tout k dans $[\![1, m]\!]$,

$$\mathsf{R}(i,k) = \begin{cases} 1 & \text{si } k = (i-1)r_0 + 1, \\ 0 & \text{sinon.} \end{cases} \tag{2.110}$$

Cas où $X = L^1([0,\tau_\star])$

On considère deux grilles $(\tau_{n,j})_{j=0}^{n}$ et $(\tau_{m,j})_{j=0}^{m}$ du type (2.22), les fonctions $(e_{n,j})_{j=1}^{n}$ et $(e_{m,j})_{j=1}^{m}$ définies par (2.25) ainsi que les fonctionnelles $(\xi_{n,j})_{j=1}^{n}$ et $(\xi_{m,j})_{j=1}^{m}$ définies par (2.24) correspondantes. On note π_n et π_m les projections correspondantes. Supposons que ces grilles soient uniformes et telles que

$$r_1 := \frac{m}{n} \tag{2.111}$$

soit un entier naturel non nul. Alors $X_n \subseteq X_m$ et $Y_m \subseteq Y_n$. En fait, pour tout $j \in [\![1,n]\!]$,

$$e_{n,j} = \pi_m e_{n,j} = \sum_{k=(j-1)r_1+1}^{jr_1} e_{m,k}. \tag{2.112}$$

Matrice de prolongement

D'après (2.106), pour tout k dans $[\![1,m]\!]$ et tout j dans $[\![1,n]\!]$,

$$\mathsf{P}(k,j) = \langle e_{n,j}\,,\, e^*_{m,k}\rangle = \frac{1}{h_{m,k}}\int_{\tau_{m,k-1}}^{\tau_{m,k}} e_{n,j}(\sigma)\,d\sigma.$$

C'est-à-dire, d'après (2.112), pour tout k dans $[\![1,m]\!]$ et tout j dans $[\![1,n]\!]$,

$$\mathsf{P}(k,j) = \begin{cases} 1 & \text{si } k \in [\![(j-1)r_1+1, jr_1]\!], \\ 0 & \text{sinon.} \end{cases} \tag{2.113}$$

Matrice de restriction

D'après (2.107), pour tout i dans $[\![1,n]\!]$ et tout k dans $[\![1,m]\!]$,

$$\mathsf{R}(i,k) = \langle e_{m,k}\,,\, \xi_{n,i}\rangle = \frac{1}{h_{n,i}}\int_{\tau_{n,i-1}}^{\tau_{n,i}} e_{m,k}(\sigma)\,d\sigma.$$

Ainsi, d'après (2.112), pour tout i dans $[\![1,n]\!]$ et tout k dans $[\![1,m]\!]$,

$$\mathsf{R}(i,k) = \begin{cases} h_{m,k}/h_{n,i} & \text{si } k \in [\![(i-1)r_1+1, ir_1]\!], \\ 0 & \text{sinon.} \end{cases} \tag{2.114}$$

La procédure décrite dans ce paragraphe peut être généralisée à des grilles emboîtées quelconques, comme nous allons le voir dans le cas des approximations par intégration produit.

Cas des approximations par intégration produit

On considère deux grilles $(\tau_{n,j})_{j=0}^{n}$ et $(\tau_{m,j})_{j=0}^{m}$ du type (2.32) ainsi que deux grilles intermédiaires $(\widehat{\tau}_{n,j})_{j=0}^{n}$ et $(\widehat{\tau}_{m,i})_{i=0}^{m}$ satisfaisant (2.33). On suppose que

$$\{\tau_{n,j} \,:\, j \in [\![0, n]\!]\} \subseteq \{\tau_{m,j} \,:\, j \in [\![0, m]\!]\} \tag{2.115}$$

et que

$$\{\widehat{\tau}_{n,j} \,:\, j \in [\![1, n]\!]\} \subseteq \{\widehat{\tau}_{m,j} \,:\, j \in [\![1, m]\!]\}. \tag{2.116}$$

Ainsi, il existe deux injections

$$\gamma : [\![0, n]\!] \to [\![0, m]\!] \tag{2.117}$$

et

$$\eta : [\![1, n]\!] \to [\![1, m]\!], \tag{2.118}$$

telles que

$$\forall j \in [\![0, n]\!] \qquad \tau_{n,j} = \tau_{m,\gamma(j)}, \tag{2.119}$$
$$\forall j \in [\![1, n]\!] \qquad \widehat{\tau}_{n,j} = \widehat{\tau}_{m,\eta(j)}. \tag{2.120}$$

Cas où $X = C^0([0, \tau_\star])$

Soient $(e_{n,j})_{j=1}^{n}$ et $(e_{m,j})_{j=1}^{m}$ les fonctions définies par (2.44), $(\xi_{n,j})_{j=1}^{n}$ et $(\xi_{m,j})_{j=1}^{m}$ les fonctionnelles définies par (2.45) correspondantes aux grilles précédentes. On désigne par T_n et T_m les opérateur approchés définis par (2.42) correspondants.

Matrice de prolongement

Si (2.119) est vérifiée, alors, pour tout $j \in [\![1, n]\!]$ et tout $\tau \in [0, \tau_\star]$,

$$e_{n,j}(\tau) = \int_{\tau_{n,j-1}}^{\tau_{n,j}} g(|\tau - \sigma|)\, d\sigma = \sum_{k=\gamma(j-1)+1}^{\gamma(j)} \int_{\tau_{m,k-1}}^{\tau_{m,k}} g(|\tau - \sigma|)\, d\sigma = \sum_{k=\gamma(j-1)+1}^{\gamma(j)} e_{m,k}(\tau).$$

Ainsi, pour tout k, dans $[\![1, m]\!]$, et i, dans $[\![1, n]\!]$,

$$\mathsf{P}(k, j) = \begin{cases} 1 & \text{si } \; k \in [\![\gamma(j-1)+1, \gamma(j)]\!], \\ 0 & \text{sinon.} \end{cases} \tag{2.121}$$

Matrice de restriction

Si (2.120) est vérifiée, alors, pour tout $i \in [\![1, n]\!]$ et tout $x \in C^0([0, \tau_\star])$,

$$\langle x\,,\,\ell_{n,i}\rangle = x(\widehat{\tau}_{n,i}) = x(\widehat{\tau}_{m,\eta(i)}) = \langle x\,,\,\ell_{m,\eta(i)}\rangle.$$

Ainsi, pour tout i, dans $[\![1, n]\!]$, et k, dans $[\![1, m]\!]$,

$$\mathsf{R}(i, k) = \delta_{\eta(i),k}. \tag{2.122}$$

Cas où $X = L^1([0, \tau_\star])$

Soient $(e_{n,j})_{j=1}^n$ et $(e_{m,j})_{j=1}^m$ les fonctions définies par (2.57), $(\xi_{n,j})_{j=1}^n$ et $(\xi_{m,j})_{j=1}^m$ les fonctionnelles définies par (2.58) correspondantes aux grilles précédentes. On désigne par T_n et T_m les opérateur approchés définis par (2.56) correspondants.

Matrice de prolongement

Si (2.120) est vérifiée, alors, pour tout $j \in [\![1, n]\!]$,

$$e_{n,j} = h_{n,j}\, g(|\cdot - \widehat{\tau}_{n,j}|) = \frac{h_{n,j}}{h_{m,\eta(j)}}\, e_{m,\eta(j)} \quad \text{p.p. sur } [0, \tau_\star].$$

Ainsi, pour tout k, dans $[\![1, m]\!]$ et tout j, dans $[\![1, n]\!]$,

$$\mathsf{P}(k, j) = \frac{h_{n,j}}{h_{m,\eta(j)}}\, \delta_{k,\eta(j)}. \tag{2.123}$$

Matrice de restriction

Si (2.119) est vérifiée, alors, pour tout $i \in [\![1, n]\!]$ et tout $x \in L^1([0, \tau_\star])$,

$$\langle x\,,\,\ell_{n,i}\rangle = \frac{1}{h_{n,i}}\int_{\tau_{n,i-1}}^{\tau_{n,i}} x(\sigma)\, d\sigma = \sum_{k=\gamma(i-1)+1}^{\gamma(i)} \frac{1}{h_{n,i}}\int_{\tau_{m,k-1}}^{\tau_{m,k}} x(\sigma)\, d\sigma = \sum_{k=\gamma(j-1)+1}^{\gamma(j)} \frac{h_{m,k}}{h_{n,i}}\langle x\,,\,\ell_{m,k}\rangle.$$

Ainsi, pour tout i, dans $[\![1, n]\!]$ et tout k, dans $[\![1, m]\!]$,

$$\mathsf{R}(i, k) = \begin{cases} h_{m,k}/h_{n,i} & \text{si } k \in [\![\gamma(i-1)+1, \gamma(i)]\!], \\ 0 & \text{sinon.} \end{cases} \tag{2.124}$$

2.2.4.3 Schéma d'Atkinson

Ce schéma correspond à (2.63) avec $G_n(z) := R_n(z)$. D'après le Théorème 14, lorsque T est remplacé par T_m, il s'écrit de la manière suivante :

$$
\left\{ \begin{array}{lcl} x^{(0)} & := & R_n(z)f, \\ f^{(k+1)} & := & (T_n - T_m)x^{(k)}, \quad k \geq 0, \\ x^{(k+1)} & := & x^{(0)} + R_n(z)f^{(k+1)}, \quad k \geq 0. \end{array} \right. \tag{2.125}
$$

Posons $f^{(0)} := f$. On définit, pour tout $k \geq 0$,

$$
v^{(k)} := R_n(z)f^{(k)}. \tag{2.126}
$$

Construction de $x^{(k)}$

On construit $x^{(k)}$ en terme de $\mathsf{x}_n^{(k)}$ et $\mathsf{x}_m^{(k)}$.

Construction de $x^{(0)}$

Soit $\mathsf{v}_n^{(0)}$ la solution du système linéaire $(\mathsf{A} - z\mathsf{I})\mathsf{v}_n^{(0)} = \mathsf{b}_n^{(0)}$. Alors

$$
x^{(0)} = v^{(0)} = \frac{1}{z}\left[\sum_{p=1}^{n} \mathsf{v}_n^{(0)}(p)e_{n,p} - f\right]. \tag{2.127}
$$

Construction de $x^{(k+1)}$

Pour tout $k \geq 0$,

$$
f^{(k+1)} = \sum_{p=1}^{n} \mathsf{x}_n^{(k)}(p)e_{n,p} - \sum_{q=1}^{m} \mathsf{x}_m^{(k)}(q)e_{m,q}. \tag{2.128}
$$

Donc, pour tout $k \geq 0$ et tout $i \in [\![1, n]\!]$,

$$
\mathsf{b}_n^{(k+1)}(i) = \sum_{p=1}^{n} \mathsf{x}_n^{(k)}(p)\langle e_{n,p}\,,\,\ell_{n,i}\rangle - \sum_{q=1}^{m} \mathsf{x}_m^{(k)}(q)\langle e_{m,q}\,,\,\ell_{n,i}\rangle,
$$

et

$$
\mathsf{b}_n^{(k+1)} = \mathsf{A}\mathsf{x}_n^{(k)} - \mathsf{C}\mathsf{x}_m^{(k)}. \tag{2.129}
$$

Pour tout $k \geq 0$,

$$\begin{aligned} v^{(k+1)} &= \frac{1}{z}\left[\sum_{p=1}^{n} \mathsf{v}_n^{(k+1)}(p) e_{n,p} - f^{(k+1)}\right] \\ &= \frac{1}{z}\left[\sum_{p=1}^{n} (\mathsf{v}_n^{(k+1)} - \mathsf{x}_n^{(k)})(p) e_{n,p} + \sum_{q=1}^{m} \mathsf{x}_m^{(k)}(q) e_{m,q}\right], \end{aligned} \tag{2.130}$$

où $\mathsf{v}_n^{(k+1)}$ est la solution du système $(\mathsf{A} - z\mathsf{I})\mathsf{v}_n^{(k+1)} = \mathsf{b}_n^{(k+1)}$.

D'après la définition de $v^{(k+1)}$, on obtient, pour tout $k \geq 0$,

$$x^{(k+1)} = x^{(0)} + v^{(k+1)}. \tag{2.131}$$

Construction de $\mathsf{x}_n^{(k)}$ et $\mathsf{x}_m^{(k)}$

- Pour tout $i \in [\![1, n]\!]$,

$$\mathsf{x}_n^{(0)}(i) = \langle x^{(0)}\,,\, \ell_{n,i}\rangle = \langle v^{(0)}\,,\, \ell_{n,i}\rangle = \mathsf{v}_n^{(0)}.$$

D'où

$$\mathsf{x}_n^{(0)} = \mathsf{v}_n^{(0)}. \tag{2.132}$$

- Pour tout $i \in [\![1, m]\!]$,

$$\begin{aligned} \mathsf{x}_m^{(0)}(i) &= \langle x^{(0)}\,,\, \ell_{m,i}\rangle \\ &= \frac{1}{z}\left[\sum_{p=1}^{n} \mathsf{x}_n^{(0)}(p)\langle e_{n,p}\,,\, \ell_{m,i}\rangle - \langle f\,,\, \ell_{m,i}\rangle\right]. \end{aligned}$$

Donc

$$\mathsf{x}_m^{(0)} = \mathsf{D}\mathsf{x}_n^{(0)} - \mathsf{b}_m. \tag{2.133}$$

- Pour tout $k \geq 1$ et tout $i \in [\![1, n]\!]$,

$$\begin{aligned} \mathsf{x}_n^{(k)}(i) &= \langle x^{(k)}\,,\, \ell_{n,i}\rangle \\ &= \langle x^{(0)}\,,\, \ell_{n,i}\rangle + \langle v^{(k)}\,,\, \ell_{n,i}\rangle \\ &= \mathsf{x}_n^{(0)}(i) + \mathsf{v}_n^{(k)}(i). \end{aligned}$$

D'où

$$\mathsf{x}_n^{(k)} = \mathsf{x}_n^{(0)} + \mathsf{v}_n^{(k)}. \tag{2.134}$$

Pour tout $k \geq 1$, et tout $i \in [\![1, m]\!]$,

$$\begin{aligned}
\mathsf{x}_m^{(k)}(i) &= \langle x^{(k)}\,,\,\ell_{m,i}\rangle \\
&= \langle x^{(0)}\,,\,\ell_{m,i}\rangle + \langle v^{(k)}\,,\,\ell_{m,i}\rangle \\
&= \mathsf{x}_m^{(0)}(i) + \frac{1}{z}\left[\sum_{p=1}^{n}\langle e_{n,p}\,,\,\ell_{m,i}\rangle(\mathsf{v}_n^{(k)} - \mathsf{x}_n^{(k-1)})(p) + \sum_{q=1}^{m}\langle e_{m,q}\,,\,\ell_{m,i}\rangle \mathsf{x}_m^{(k-1)}(q)\right].
\end{aligned}$$

Donc

$$\mathsf{x}_m^{(k)} = \mathsf{x}_m^{(0)} + \frac{1}{z}\left[\mathsf{D}(\mathsf{v}_n^{(k)} - \mathsf{x}_n^{(k-1)}) + \mathsf{B}\mathsf{x}_m^{(k-1)}\right]. \tag{2.135}$$

Algorithme

▷ Construire A, B, C, D, $\mathsf{b}_n^{(0)}$, b_m

▷ $\mathsf{x}_n^{(0)} \leftarrow$ solution de $(\mathsf{A} - z\mathsf{I})\mathsf{x}_n^{(0)} = \mathsf{b}_n^{(0)}$

• $x^{(0)} \leftarrow \underbrace{\frac{1}{z}\left[\sum_{p=1}^{n}\mathsf{x}_n^{(0)}(p)e_{n,p} - f\right]}_{v^{(0)}}$

▷ $\mathsf{x}_m^{(0)} \leftarrow \mathsf{D}\mathsf{x}_n^{(0)} - \mathsf{b}_m$

▷ résidu relatif $\leftarrow \dfrac{\left\| \sum_{q=1}^{m} \mathsf{x}_m^{(0)}(q)e_{m,q} - zx^{(0)} - f\right\|}{\|f\|}$

▷ $\mathsf{x}_m \leftarrow \mathsf{x}_m^{(0)}$

▷ $\mathsf{x}_n \leftarrow \mathsf{x}_n^{(0)}$

Tant que (résidu relatif > tolérance)

▷ $\mathsf{v}_n \leftarrow$ solution de $(\mathsf{A} - z\mathsf{I})\mathsf{v}_n = \mathsf{A}\mathsf{x}_n - \mathsf{C}\mathsf{x}_m$

• $x \leftarrow x^{(0)} + \underbrace{\frac{1}{z}\left[\sum_{p=1}^{n}(\mathsf{v}_n - \mathsf{x}_n)(p)e_{n,p} + \sum_{q=1}^{m}\mathsf{x}_m(q)e_{m,q}\right]}_{v}$

▷ $\mathsf{x}_m \leftarrow \mathsf{x}_m^{(0)} + \dfrac{1}{z}\Big[\mathsf{D}(\mathsf{v}_n - \mathsf{x}_n) + \mathsf{B}\mathsf{x}_m\Big]$

▷ résidu relatif $\leftarrow \dfrac{\left\| \sum_{q=1}^{m} \mathsf{x}_m(q) e_{m,q} - zx - f \right\|}{\|f\|}$

▷ $\mathsf{x}_n \leftarrow \mathsf{x}_n^{(0)} + \mathsf{v}_n$

Fin (tant que).

2.2.4.4 Schéma de Brakhage

Ce schéma correspond au schéma (2.63) avec $G_n(z) := \widetilde{R}_n(z)$, où

$$\widetilde{R}_n(z) := \frac{1}{z}(R_n(z)T - I).$$

D'après le Théorème 15, lorsque T est remplacé par T_m, il s'écrit de la manière suivante :

$$\left\{ \begin{array}{lll} x^{(0)} & := & \widetilde{R}_n(z)f, \\ f^{(k+1)} & := & (T_n - T_m)T_m x^{(k)}, \quad k \geq 0, \\ x^{(k+1)} & := & x^{(0)} + \dfrac{1}{z} R_n(z) f^{(k+1)}, \quad k \geq 0. \end{array} \right. \tag{2.136}$$

On pose $f^{(0)} := T_m f$, et pour tout $k \geq 0$, on définit

$$v^{(k)} := R_n(z) f^{(k)}. \tag{2.137}$$

Construction de $x^{(k)}$

On construit $x^{(k)}$ en fonction de $\mathsf{x}_m^{(k)}$.

Construction de $x^{(0)}$

$$f^{(0)} = T_m f = \sum_{q=1}^{m} \langle f \, , \, \ell_{m,q} \rangle e_{m,q} = \sum_{q=1}^{m} \mathsf{b}_m(q) e_{m,q}. \tag{2.138}$$

Pour tout $i \in [\![1, n]\!]$,

$$\mathsf{b}_n^{(0)}(i) = \langle T_m f \, , \, \ell_{n,i} \rangle = \sum_{q=1}^{m} \langle f \, , \, \ell_{m,q} \rangle \langle e_{m,q} \, , \, \ell_{n,i} \rangle,$$

et ainsi

$$\mathsf{b}_n^{(0)} = \mathsf{C}\mathsf{b}_m. \tag{2.139}$$

Donc

$$v^{(0)} = \frac{1}{z}\left[\sum_{p=1}^{n} \mathsf{v}_n^{(0)}(p)e_{n,p} - \sum_{q=1}^{m} \mathsf{b}_m(q)e_{m,q}\right], \tag{2.140}$$

où $\mathsf{v}_n^{(0)}$ est la solution du système linéaire $(\mathsf{A}-z\mathsf{I})\mathsf{v}_n^{(0)} = \mathsf{b}_n^{(0)}$. Enfin,

$$x^{(0)} = \frac{1}{z}(v^{(0)} - f). \tag{2.141}$$

Construction de $x^{(k+1)}$

Pour tout $p \in [\![1, n]\!]$ et $q \in [\![1, m]\!]$,

$$\begin{aligned}
\langle T_m x^{(k)}\,,\,\ell_{n,p}\rangle &= \sum_{j=1}^{m}\langle x^{(k)}\,,\,\ell_{m,j}\rangle\langle e_{m,j}\,,\,\ell_{n,p}\rangle = (\mathsf{C}\mathsf{x}_m^{(k)})(p),\\
\langle T_m x^{(k)}\,,\,\ell_{m,q}\rangle &= \sum_{j=1}^{m}\langle x^{(k)}\,,\,\ell_{m,j}\rangle\langle e_{m,j}\,,\,\ell_{m,q}\rangle = (\mathsf{B}\mathsf{x}_m^{(k)})(q),
\end{aligned}$$

et ainsi

$$f^{(k+1)} = \sum_{p=1}^{n}(\mathsf{C}\mathsf{x}_m^{(k)})(p)e_{n,p} - \sum_{q=1}^{m}(\mathsf{B}\mathsf{x}_m^{(k)})(q)e_{m,q}. \tag{2.142}$$

Pour tout $i \in [\![1, n]\!]$,

$$\mathsf{b}_n^{(k+1)}(i) = \langle f^{(k+1)}\,,\,\ell_{n,i}\rangle = \sum_{p=1}^{n}(\mathsf{C}\mathsf{x}_m^{(k)})(p)\langle e_{n,p}\,,\,\ell_{n,i}\rangle - \sum_{q=1}^{m}(\mathsf{B}\mathsf{x}_m^{(k)})(q)\langle e_{m,q}\,,\,\ell_{n,i}\rangle,$$

donc

$$\mathsf{b}_n^{(k+1)} = (\mathsf{AC} - \mathsf{CB})\mathsf{x}_m^{(k)}. \tag{2.143}$$

Soit $\mathsf{v}_n^{(k+1)}$ la solution du système linéaire $(\mathsf{A}-z\mathsf{I})\mathsf{v}_n^{(k+1)} = \mathsf{b}_n^{(k+1)}$. Alors

$$v^{(k+1)} = \frac{1}{z}\left[\sum_{p=1}^{n}(\mathsf{v}_n^{(k+1)} - \mathsf{C}\mathsf{x}_m^{(k)})(p)e_{n,p} + \sum_{q=1}^{m}\mathsf{B}\mathsf{x}_m^{(k)}(q)e_{m,q}\right], \tag{2.144}$$

et

$$x^{(k+1)} = x^{(0)} + \frac{1}{z} v^{(k+1)}. \tag{2.145}$$

Construction de $\mathsf{x}_m^{(k)}$

- Pour tout $i \in [\![1, m]\!]$,

$$\begin{aligned}
\mathsf{x}_m^{(0)}(i) &= \langle x^{(0)} \,,\, \ell_{m,i} \rangle \\
&= \frac{1}{z}\Big(\langle v^{(0)} \,,\, \ell_{m,i}\rangle - \langle f \,,\, \ell_{m,i}\rangle \Big) \\
&= \frac{1}{z}\left[\frac{1}{z}\sum_{p=1}^{n} \mathsf{v}_n^{(0)}(p)\langle e_{n,p} \,,\, \ell_{m,i}\rangle - \frac{1}{z}\sum_{q=1}^{m} \mathsf{b}_m(q)\langle e_{m,q} \,,\, \ell_{m,i}\rangle - \langle f \,,\, \ell_{m,i}\rangle \right],
\end{aligned}$$

donc

$$\mathsf{x}_m^{(0)} = \frac{1}{z^2}\Big[\mathsf{D}\mathsf{v}_n^{(0)} - \mathsf{B}\mathsf{b}_m \Big] - \frac{1}{z}\mathsf{b}_m. \tag{2.146}$$

- Pour tout $k \geq 1$, et pour tout $i \in [\![1, m]\!]$,

$$\begin{aligned}
\mathsf{x}_m^{(k)}(i) &= \langle x^{(k)} \,,\, \ell_{m,i} \rangle \\
&= \langle x^{(0)} \,,\, \ell_{m,i} \rangle + \frac{1}{z}\langle v^{(k)} \,,\, \ell_{m,i} \rangle \\
&= \mathsf{x}_m^{(0)}(i) + \frac{1}{z^2}\left[\sum_{p=1}^{n} (\mathsf{v}_n^{(k)} - \mathsf{C}\mathsf{x}_m^{(k-1)})(p)\langle e_{n,p} \,,\, \ell_{m,i}\rangle + \sum_{q=1}^{m} \mathsf{B}\mathsf{x}_m^{(k-1)}(q)\langle e_{m,q} \,,\, \ell_{m,i}\rangle \right].
\end{aligned}$$

Ainsi

$$\mathsf{x}_m^{(k)} = \mathsf{x}_m^{(0)} + \frac{1}{z^2}\Big[\mathsf{D}\mathsf{v}_n^{(k)} + (\mathsf{B}^2 - \mathsf{D}\mathsf{C})\mathsf{x}_m^{(k-1)} \Big]. \tag{2.147}$$

Algorithme

▷ Construire A, B, C, D, $\mathsf{b}_n^{(0)}$, b_m

▷ $\mathsf{v}_n \leftarrow$ solution de $(\mathsf{A} - z\mathsf{I})\mathsf{v}_n = \mathsf{C}\mathsf{b}_m$

- $x^{(0)} \leftarrow \frac{1}{z}\Big[\underbrace{\frac{1}{z}\Big(\sum_{p=1}^{n} \mathsf{v}_n(p) e_{n,p} - \sum_{q=1}^{m} \mathsf{b}_m(q) e_{m,q} \Big)}_{v^{(0)}} - f \Big]$

$\triangleright$ $\mathsf{x}_m^{(0)} \leftarrow \dfrac{1}{z^2}\left[\mathsf{D}\mathsf{v}_n - \mathsf{B}\mathsf{b}_m\right] - \dfrac{1}{z}\mathsf{b}_m$

$\triangleright$ $\mathsf{x}_m \leftarrow \mathsf{x}_m^{(0)}$

$\triangleright$ résidu relatif $\leftarrow \dfrac{\left\| \sum\limits_{q=1}^{m} \mathsf{x}_m(q) e_{m,q} - z x^{(0)} - f \right\|}{\|f\|}$

Tant que (résidu relatif $>$ tolérance)

$\triangleright$ $\mathsf{v}_n \leftarrow$ solution de $(\mathsf{A} - z\mathsf{I})\mathsf{v}_n = (\mathsf{AC} - \mathsf{CB})\mathsf{x}_m$

- $x \leftarrow x^{(0)} + \underbrace{\dfrac{1}{z^2}\left[\sum\limits_{p=1}^{n}(\mathsf{v}_n - \mathsf{C}\mathsf{x}_m)(p) e_{n,p} + \sum\limits_{q=1}^{m} \mathsf{B}\mathsf{x}_m(q) e_{m,q}\right]}_{\frac{1}{z}v}$

$\triangleright$ $\mathsf{x}_m = \mathsf{x}_m^{(0)} + \dfrac{1}{z^2}\left[\mathsf{D}\mathsf{v}_n + (\mathsf{B}^2 - \mathsf{DC})\mathsf{x}_m\right]$

$\triangleright$ résidu relatif $\leftarrow \dfrac{\left\| \sum\limits_{q=1}^{m} \mathsf{x}_m(q) e_{m,q} - z x - f \right\|}{\|f\|}$

Fin (tant que).

2.2.4.5 Schéma (C)

Ce schéma correspond à (2.63) avec $G_n(z) := \widehat{R}_n(z)$, $\widehat{R}_n(z) := \dfrac{1}{z}(TR_n(z) - I)$. D'après le Théorème 16, lorsque T est remplacé par T_m, il s'écrit de la manière suivante :

$$\left\{ \begin{array}{lll} x^{(0)} & := & \widehat{R}_n(z) f, \\ f^{(k+1)} & := & (T_n - T_m) x^{(k)}, \quad k \geq 0, \\ x^{(k+1)} & := & x^{(0)} + \dfrac{1}{z} T_m R_n(z) f^{(k+1)}, \quad k \geq 0. \end{array} \right. \tag{2.148}$$

On pose $f^{(0)} := f$ et, pour tout $k \geq 0$, on définit

$$v^{(k)} := R_n(z) f^{(k)}. \tag{2.149}$$

Donc

$$\left\{\begin{array}{lcl} x^{(0)} & := & \dfrac{1}{z}\left(T_m v^{(0)} - f\right), \\ x^{(k+1)} & = & x^{(0)} + \dfrac{1}{z} T_m v^{(k+1)}, \qquad k \geq 0. \end{array}\right. \tag{2.150}$$

Construction de $x^{(k)}$

On construit $x^{(k)}$ en fonction de $\mathsf{x}_n^{(k)}$ et $\mathsf{x}_m^{(k)}$.

Construction de $x^{(0)}$

$$v^{(0)} = \frac{1}{z}\left(\sum_{p=1}^{n} \mathsf{v}_n^{(0)}(p) e_{n,p} - f\right), \tag{2.151}$$

où $\mathsf{v}_n^{(0)}$ est la solution du système linéaire $(\mathsf{A} - z\mathsf{I})\mathsf{v}_n^{(0)} = \mathsf{b}_n^{(0)}$. Ainsi

$$\begin{aligned} x^{(0)} &= \frac{1}{z}\left(T_m v^{(0)} - f\right) \\ &= \frac{1}{z}\left(\frac{1}{z}\sum_{q=1}^{m}\left[\sum_{p=1}^{n} \mathsf{v}_n^{(0)}(p)\langle e_{n,p}\,,\, \ell_{m,q}\rangle - \langle f\,,\,\ell_{m,q}\rangle\right] e_{m,q} - f\right), \end{aligned}$$

$$x^{(0)} = \frac{1}{z}\left(\frac{1}{z}\sum_{q=1}^{m}(\mathsf{D}\mathsf{v}_n^{(0)} - \mathsf{b}_m)(q) e_{m,q} - f\right). \tag{2.152}$$

Construction de $x^{(k+1)}$

Pour tout $k \geq 0$,

$$f^{(k+1)} = \sum_{p=1}^{n} \mathsf{x}_n^{(k)}(p) e_{n,p} - \sum_{q=1}^{m} \mathsf{x}_m^{(k)}(q) e_{m,q}. \tag{2.153}$$

Donc, pour tout $k \geq 0$ et tout $i \in [\![1, n]\!]$,

$$\mathsf{b}_n^{(k+1)}(i) = \sum_{p=1}^{n} \mathsf{x}_n^{(k)}(p)\langle e_{n,p}\,,\,\ell_{n,i}\rangle - \sum_{q=1}^{m} \mathsf{x}_m^{(k)}(q)\langle e_{m,q}\,,\,\ell_{n,i}\rangle,$$

et ainsi

$$\mathsf{b}_n^{(k+1)} = \mathsf{A}\mathsf{x}_n^{(k)} - \mathsf{C}\mathsf{x}_m^{(k)}. \tag{2.154}$$

Pour tout $k \geq 0$,

$$\begin{aligned} v^{(k+1)} &= \frac{1}{z}\left[\sum_{p=1}^{n} \mathsf{v}_n^{(k+1)}(p)e_{n,p} - f^{(k+1)}\right] \\ &= \frac{1}{z}\left[\sum_{p=1}^{n}(\mathsf{v}_n^{(k+1)} - \mathsf{x}_n^{(k)})(p)e_{n,p} + \sum_{q=1}^{m} \mathsf{x}_m^{(k)}(q)e_{m,q}\right], \end{aligned} \tag{2.155}$$

où $\mathsf{v}_n^{(k+1)}$ est la solution système linéaire $(\mathsf{A} - z\mathsf{I})\mathsf{v}_n^{(k+1)} = \mathsf{b}_n^{(k+1)}$. Ainsi

$$\begin{aligned} T_m v^{(k+1)} &= \frac{1}{z}\sum_{q=1}^{m}\left(\sum_{j=1}^{n}(\mathsf{v}_n^{(k+1)} - \mathsf{x}_n^{(k)})(j)\langle e_{n,j}\,,\,\ell_{m,q}\rangle + \sum_{j=1}^{m}\mathsf{x}_m^{(k)}(j)\langle e_{m,j}\,,\,\ell_{m,q}\rangle\right) e_{m,q} \\ &= \frac{1}{z}\sum_{q=1}^{m}\left(\mathsf{D}(\mathsf{v}_n^{(k+1)} - \mathsf{x}_n^{(k)}) + \mathsf{B}\mathsf{x}_m^{(k)}\right)(q)e_{m,q}. \end{aligned}$$

Donc

$$x^{(k+1)} = x^{(0)} + \frac{1}{z^2}\sum_{q=1}^{m}(\mathsf{D}(\mathsf{v}_n^{(k+1)} - \mathsf{x}_n^{(k)}) + \mathsf{B}\mathsf{x}_m^{(k)})(q)e_{m,q}. \tag{2.156}$$

Construction de $\mathsf{x}_n^{(k)}$ et $\mathsf{x}_m^{(k)}$

• Pour tout $i \in [\![1, n\,]\!]$,

$$\begin{aligned} \mathsf{x}_n^{(0)}(i) &= \langle x^{(0)}\,,\,\ell_{n,i}\rangle \\ &= \frac{1}{z}\left(\frac{1}{z}\sum_{q=1}^{m}(\mathsf{D}\mathsf{v}_n^{(0)} - \mathsf{b}_m)(q)\langle e_{m,q}\,,\,\ell_{n,i}\rangle - \langle f\,,\,\ell_{n,i}\rangle\right), \end{aligned}$$

et ainsi

$$\mathsf{x}_n^{(0)} = \frac{1}{z}\left(\frac{1}{z}\mathsf{C}(\mathsf{D}\mathsf{v}_n^{(0)} - \mathsf{b}_m) - \mathsf{b}_n^{(0)}\right). \tag{2.157}$$

• Pour tout $k \geq 0$ et tout $i \in [\![1, m\,]\!]$,

$$\begin{aligned} \mathsf{x}_n^{(k+1)}(i) &= \langle x^{(0)}\,,\,\ell_{n,i}\rangle \\ &= \langle x^{(0)}\,,\,\ell_{n,i}\rangle + \frac{1}{z^2}\sum_{q=1}^{m}\left(\mathsf{D}(\mathsf{v}_n^{(k+1)} - \mathsf{x}_n^{(k)}) + \mathsf{B}\mathsf{x}_m^{(k)}\right)(q)\langle e_{m,q}\,,\,\ell_{n,i}\rangle, \end{aligned}$$

ainsi

$$\mathsf{x}_n^{(k+1)} = \mathsf{x}_n^{(0)} + \frac{1}{z^2}\left(\mathsf{C}\mathsf{D}(\mathsf{v}_n^{(k+1)} - \mathsf{x}_n^{(k)}) + \mathsf{C}\mathsf{B}\mathsf{x}_m^{(k)}\right). \tag{2.158}$$

• Pour tout $i \in [\![1, n]\!]$,

$$\begin{aligned}
\mathsf{x}_m^{(0)}(i) &= \langle x^{(0)}, \ell_{m,i}\rangle \\
&= \frac{1}{z}\left(\frac{1}{z}\sum_{q=1}^{m}(\mathsf{D}\mathsf{v}_n^{(0)} - \mathsf{b}_m)(q)\langle e_{m,q}, \ell_{m,i}\rangle - \langle f, \ell_{m,i}\rangle\right),
\end{aligned}$$

et ainsi

$$\mathsf{x}_m^{(0)} = \frac{1}{z}\left(\frac{1}{z}\mathsf{B}(\mathsf{D}\mathsf{v}_n^{(0)} - \mathsf{b}_m) - \mathsf{b}_m\right). \tag{2.159}$$

• Pour tout $k \geq 0$ et tout $i \in [\![1, m]\!]$,

$$\begin{aligned}
\mathsf{x}_m^{(k+1)}(i) &= \langle x^{(0)}, \ell_{m,i}\rangle \\
&= \langle x^{(0)}, \ell_{m,i}\rangle + \frac{1}{z^2}\sum_{q=1}^{m}\left(\mathsf{D}(\mathsf{v}_n^{(k+1)} - \mathsf{x}_n^{(k)}) + \mathsf{B}\mathsf{x}_m^{(k)}\right)(q)\langle e_{m,q}, \ell_{m,i}\rangle,
\end{aligned}$$

ainsi

$$\mathsf{x}_m^{(k+1)} = \mathsf{x}_m^{(0)} + \frac{1}{z^2}\left(\mathsf{B}\mathsf{D}(\mathsf{v}_n^{(k+1)} - \mathsf{x}_n^{(k)}) + \mathsf{B}^2\mathsf{x}_m^{(k)}\right). \tag{2.160}$$

Algorithme

▷ Construire A, B, C, D, $\mathsf{b}_n^{(0)}$, b_m

▷ $\mathsf{v}_n \leftarrow$ solution de $(\mathsf{A} - z\mathsf{I})\mathsf{v}_n = \mathsf{b}_n^{(0)}$

• $x^{(0)} \leftarrow \dfrac{1}{z}\left(\dfrac{1}{z}\displaystyle\sum_{q=1}^{m}(\mathsf{D}\mathsf{v}_n - \mathsf{b}_m)(q)e_{m,q} - f\right)$

▷ $\mathsf{x}_n^{(0)} \leftarrow \dfrac{1}{z}\left(\dfrac{1}{z}\mathsf{C}(\mathsf{D}\mathsf{v}_n - \mathsf{b}_m) - \mathsf{b}_n^{(0)}\right)$

▷ $\mathsf{x}_m^{(0)} \leftarrow \dfrac{1}{z}\left(\dfrac{1}{z}\mathsf{B}(\mathsf{D}\mathsf{v}_n - \mathsf{b}_m) - \mathsf{b}_m\right)$

▷ $\mathsf{x}_n \leftarrow \mathsf{x}_n^{(0)}$; $\mathsf{x}_m \leftarrow \mathsf{x}_m^{(0)}$

$\triangleright$ résidu relatif $\leftarrow \dfrac{\left\| \sum_{q=1}^{m} \mathsf{x}_m(q) e_{m,q} - z x^{(0)} - f \right\|}{\|f\|}$

Tant que (résidu relatif $>$ tolérance)

$\triangleright$ $\mathsf{v}_n \leftarrow$ solution de $(\mathsf{A} - z\mathsf{I})\mathsf{v}_n = \mathsf{A}\mathsf{x}_n - \mathsf{C}\mathsf{x}_m$

• $x \leftarrow x^{(0)} + \dfrac{1}{z^2} \sum_{q=1}^{m} \left(\mathsf{D}(\mathsf{v}_n - \mathsf{x}_n) + \mathsf{B}\mathsf{x}_m\right)(q) e_{m,q}$

$\triangleright$ $\mathsf{x}_m^{\text{last}} \leftarrow \mathsf{x}_m$

$\triangleright$ $\mathsf{x}_m \leftarrow \mathsf{x}_m^{(0)} + \dfrac{1}{z^2} \left(\mathsf{B}\mathsf{D}(\mathsf{v}_n - \mathsf{x}_n) + \mathsf{B}^2\mathsf{x}_m\right)$

$\triangleright$ $\mathsf{x}_n \leftarrow \mathsf{x}_n^{(0)} + \dfrac{1}{z^2} \left(\mathsf{C}\mathsf{D}(\mathsf{v}_n - \mathsf{x}_n) + \mathsf{C}\mathsf{B}\mathsf{x}_m^{\text{last}}\right)$

$\triangleright$ résidu relatif $\leftarrow \dfrac{\left\| \sum_{q=1}^{m} \mathsf{x}_m(q) e_{m,q} - z x - f \right\|}{\|f\|}$

Fin (tant que).

2.2.4.6 Exposition et calcul du résidu relatif

Exposition des solutions

On remarque que les trois schémas donnent des solutions approchées de la forme

$$x^{(k)} = \frac{1}{z} \left[\sum_{p=1}^{n} \mathsf{u}_n^{(k)}(p) e_{n,p} + \sum_{q=1}^{m} \mathsf{w}_m^{(k)}(q) e_{m,q} - f \right], \tag{2.161}$$

où $\mathsf{u}_n^{(k)} \in \mathbb{C}^{n \times 1}$ et $\mathsf{w}_m^{(k)} \in \mathbb{C}^{m \times 1}$. D'après les hypothèses (2.98) et (2.99) faites sur les familles de fonctions $(e_{n,j})_{j=1}^{n}$ et $(e_{m,j})_{j=1}^{m}$ et sur les familles de fonctionnelles linéaires-conjuguées $(\ell_{n,j})_{j=1}^{n}$ et $(\ell_{m,j})_{j=1}^{m}$, l'expression précédente devient

$$x^{(k)} = \frac{1}{z} \left[\sum_{q=1}^{m} (\mathsf{P}\mathsf{u}_n^{(k)} + \mathsf{w}_m^{(k)})(q) e_{m,q} - f \right]. \tag{2.162}$$

Dans le cas de l'espace $C^0([0,\tau_\star])$, l'égalité précédente désigne simplement l'égalité entre deux fonctions de cet espace. Dans le cas de l'espace $L^1([0,\tau_\star])$, elle représente l'appartenance de deux fonctions à la même classe d'équivalence : cette égalité est donc, dans ce cas, une égalité presque partout sur $[0,\tau_\star]$. L'exposition d'un itéré $x^{(k)}$ donné peut être effectuée sur une grille $(\tau_{\nu,i})_{i=1}^{\nu}$, qui n'est pas nécessairement l'une des grilles $(\tau_{n,i})_{i=1}^{n}$ ou $(\tau_{m,i})_{i=1}^{m}$ ayant servi à définir les opérateurs approchés T_n et T_m. Dans le cas de l'espace $C^0([0,\tau_\star])$, nous pouvons exhiber la valeur de $x^{(k)}$ en chaque nœud de la grille d'exposition $(\tau_{\nu,i})_{i=1}^{\nu}$. Dans le cas de l'espace $L^1([0,\tau_\star])$, l'itéré que nous voulons exposer peut ne pas être défini en certains noeuds de cette grille. On préfère, dans ce cas, exhiber les moyennes locales $\dfrac{1}{h_{\nu,i}}\displaystyle\int_{\tau_{\nu,i-1}}^{\tau_{\nu,i}} x^{(k)}(\sigma)\,d\sigma$.

Calcul du résidu relatif

À chaque itération on doit calculer le résidu relatif

$$\rho^{(k)} := \frac{\left\| \sum\limits_{q=1}^{m} \mathsf{x}_m(q) e_{m,q} - z x^{(k)} - f \right\|}{\|f\|} \tag{2.163}$$

$$= \frac{\left\| \sum\limits_{q=1}^{m} (\mathsf{x}_m^{(k)} - \mathsf{w}_m^{(k)} - \mathsf{Pu}_n^{(k)})(q) e_{m,q} \right\|}{\|f\|}. \tag{2.164}$$

Cas de l'espace $C^0([0,\tau_\star])$

Dans ce cas on calcule en pratique

$$\rho_{\nu,\infty}^{(k)} := \frac{1}{\|f\|_\infty} \max \left\{ \left| \sum_{q=1}^{m} (\mathsf{x}_m^{(k)} - \mathsf{w}_m^{(k)} - \mathsf{Pu}_n^{(k)})(q) e_{m,q}(\tau_{\nu,j}) \right|,\ 1 \le j \le \nu \right\}, \tag{2.165}$$

qui donne une approximation du résidu relatif calculé avec la norme infinie.

Cas de l'espace $L^1([0,\tau_\star])$

Dans le cas de l'espace L^1, nous devons calculer le résidu en utilisant la norme L^1 dans (2.164) :

$$\rho_1^{(k)} = \frac{1}{\|f\|} \int_0^{\tau_\star} \left| \sum_{q=1}^{m} (\mathsf{x}_m^{(k)} - \mathsf{w}_m^{(k)} - \mathsf{Pu}_n^{(k)})(q) e_{m,q}(\sigma) \right| d\sigma. \tag{2.166}$$

On note $S_{m,q}$ le support de la fonction $e_{m,q}$. Si $(\mathring{S}_{m,q})_{q=1}^{m}$ est une famille d'ouverts deux-à-deux disjoints et si $\bigcup\limits_{q=1}^{m} S_{m,q} = [0, \tau_\star]$, ce qui est le cas dans notre application des approximations par projection présentées précédemment, alors

$$\|T_m x^{(k)} - z x^{(k)} - f\|_1 = \sum_{q=1}^{m} |(\mathsf{x}_m^{(k)} - \mathsf{w}_m^{(k)} - \mathsf{P}\mathsf{u}_n^{(k)})(q)| \int_{S_{m,q}} |e_{m,q}(\sigma)|\, d\sigma.$$

Remarque 12 *Lorsque l'expression* (2.166) *est difficile à calculer, on peut être tenté d'évaluer le résidu suivant*

$$\|(\mathsf{B} - z\mathsf{I})\mathsf{x}_m^{(k)} - \mathsf{b}_m\|,$$

qui représente le résidu de $\mathsf{x}_m^{(k)}$ *associé au système linéaire*

$$(\mathsf{B} - z\mathsf{I})\mathsf{x} = \mathsf{b}_m.$$

On remarque alors que pour tout $i \in [\![1, m]\!]$,

$$\begin{aligned} ((\mathsf{B} - z\mathsf{I})\mathsf{x}_m^{(k)}))(i) - \mathsf{b}_m(i) &= \sum_{q=1}^{m} \langle x^{(k)}\,,\, \ell_{m,q}\rangle\langle e_{m,q}\,,\, \ell_{m,i}\rangle - z\langle x^{(k)}\,,\, \ell_{m,i}\rangle - \langle f\,,\, \ell_{m,i}\rangle \\ &= \langle T_m x^{(k)} - z x^{(k)} - f\,,\, \ell_{m,i}\rangle, \end{aligned}$$

et donc que si la norme matricielle utilisée est la norme 1,

$$\|(\mathsf{B} - z\mathsf{I})\mathsf{x}_m^{(k)} - \mathsf{b}_m\|_1 = \sum_{i=1}^{m} |\langle T_m x^{(k)} - z x^{(k)} - f\,,\, \ell_{m,i}\rangle|.$$

Or, la convergence de la suite $\left(\sum\limits_{i=1}^{m} |\langle T_m x^{(k)} - z x^{(k)} - f\,,\, \ell_{m,i}\rangle|\right)_{k\geq 1}$ *n'implique pas celle de la suite* $\left(\|T_m x^{(k)} - z x^{(k)} - f\|_1\right)_{k\geq 1}$.

2.2.5 Réduction du coût de calcul

L'implémentation des algorithmes de raffinement itératif nécessite, à chaque itération, la résolution du système linéaire

$$(\mathsf{A}_n - z\mathsf{I}_n)\mathsf{x}_n = \mathsf{b}_n,$$

où A_n est une matrice pleine, fixée au début de l'algorithme. Sous certaines hypothèses sur le noyau κ, un grand nombre de coefficients de la matrice A_n sont de l'ordre de zéro lorsque $\tau_\star$ est grand. Nous allons étudier l'effet de l'annulation de ces coefficients dans le cas des approximations par projection dans l'espace $L^1([0,\tau_\star])$, ceci pouvant se généraliser aux autres méthodes d'approximation de rang fini présentées ici.

Soit $(\epsilon_n)_{n\geq 0}$ une suite de réels positifs. On définit

$$\begin{aligned} I_{n,i,j} &:= [\tau_{n,i-1},\tau_{n,i}]\times[\tau_{n,j-1},\tau_{n,j}], \quad (i,j)\in[\![1,n]\!]^2, \\ \mathcal{E}_n &:= \left\{(i,j)\in[\![1,n]\!]^2 \;:\; \sup_{(t,s)\in I_{n,i,j}} \kappa(t,s)\leq \frac{\epsilon_n}{h_{n,j}}\right\}. \end{aligned}$$

Soit κ_n la fonction définie pour tout $(\tau,\sigma)\in[0,\tau_\star]^2$ tel que $\tau\neq\sigma$, par

(2.167)
$$\kappa_n(\tau,\sigma) := \begin{cases} 0 & \text{si } \exists (i,j) \text{ tel que } (\tau,\sigma)\in I_{n,i,j} \text{ et } \displaystyle\sup_{(t,s)\in I_{n,i,j}} \kappa(t,s)\leq \frac{\epsilon_n}{h_{n,j}}, \\ \kappa(\tau,\sigma) & \text{sinon.} \end{cases}$$

Soit K_n l'opérateur intégral induit par le noyau σ_n :

$$(K_n x)(\tau) = \int_0^{\tau_\star} \kappa_n(\tau,\sigma)x(\sigma)\,d\sigma, \quad x\in L^1([0,\tau_\star]),$$

et considérons l'approximation de rang fini

$$\widetilde{T}_n := \pi_n K_n.$$

On désigne par $\widetilde{\mathsf{A}}_n$ la matrice du système linéaire (2.5) correspondant à l'approximation $\widetilde{T}_n$. Alors

$$\widetilde{\mathsf{A}}_n(i,j) = \frac{1}{h_{n,i}}\iint_{I_{n,i,j}} \sigma_n(\tau,\sigma)d\tau\,d\sigma, \quad (i,j)\in[\![1,n]\!]^2.$$

Théorème 26 *Pour tout* $(i,j)\in[\![1,n]\!]^2$,

$$\widetilde{\mathsf{A}}_n(i,j) = \begin{cases} 0 & si\ (i,j)\in\mathcal{E}_n, \\ \mathsf{A}_n(i,j) & sinon, \end{cases}$$

où A_n *est la matrice définie par* (2.29).

Preuve : Évidente. ∎

Théorème 27 *Supposons que la suite* $(\epsilon_n)_{n\geq 0}$ *est telle que* $\lim_{n\to+\infty}\frac{\epsilon_n}{\mu_n}=0$. *Alors* $\left(\widetilde{T}_n\right)_{n\geq 1}$ *est convergente en norme vers* T.

Preuve : Soit $x \in L^1([0, \tau_\star])$. Alors,

$$\begin{aligned}
\|(T - K_n)x\|_1 &\leq \sum_{i=1}^{n}\sum_{j=1}^{n} \iint_{I_{n,i,j}} |\kappa(\tau,\sigma) - \kappa_n(\tau,\sigma)||x(\sigma)|d\tau\, d\sigma \\
&\leq \sum_{(i,j)\in\mathcal{E}_n} \frac{\epsilon_n}{h_{n,j}} \iint_{I_{n,i,j}} |x(\sigma)|d\tau\, d\sigma \\
&\leq \sum_{(i,j)\in\mathcal{E}_n} \frac{\epsilon_n h_{n,i}}{h_{n,j}} \int_{\tau_{n,j-1}}^{\tau_{n,j}} |x(\sigma)|\, d\sigma \leq \frac{\tau_\star \epsilon_n}{\mu_n}\|x\|_1.
\end{aligned}$$

Ainsi, si $\lim\limits_{n\to+\infty} \dfrac{\epsilon_n}{\mu_n} = 0$, $\lim\limits_{n\to+\infty} \|T - K_n\|_1 = 0$. De plus,

$$\begin{aligned}
\|T - \widetilde{T}_n\|_1 &\leq \|T - K_n\|_1 + \|(I - \pi_n)K_n\|_1 \\
&\leq \|T - K_n\|_1 + \|(I - \pi_n)T\|_1 + \|(I - \pi_n)(K_n - T)\|_1,
\end{aligned}$$

et on en déduit la convergence en norme de $\widetilde{T}_n$ vers T. ∎

Commentaires bibliographiques

L'implémentation des algorithmes de raffinement itératif (A), (B) et (C) nécessite, à chaque itération, l'évaluation de plusieurs produits matrice-vecteur. Afin de limiter le coût de calcul dû à ces produits, on pourra s'inspirer des travaux de Beylkin, Coifman et Rokhlin dans [20], où ils décrivent une technique utilisant les propriétés algébriques de la matrice mise en jeu. D'autres techniques ont été développées dans le même sens dans [43], [41] et [68].

Chapitre 3

Application à la théorie du transfert

3.1 La théorie du transfert

La théorie du transfert traite de la propagation d'un champ de rayonnement dans un milieu présentant des hétérogénéités capables d'absorber, d'émettre ou de diffuser de la lumière. Elle s'applique à des milieux discrets dilués, c'est-à-dire constitués de particules de taille beaucoup plus petite que leur distance moyenne. Ces particules peuvent être les constituants d'un gaz ou d'un plasma (électrons, ions, atomes, molécules) ou des particules macroscopiques (macromolécules, gouttelettes, grains, etc ...).

On trouve des applications de cette théorie dans divers domaines, notamment en Géophysique Externe (météorologie, optique des océans ...), en Physique Nucléaire, en Chimie Physique, en Biologie et Médecine, et enfin, en Astrophysique. Notre étude a été motivée par une application à la théorie des atmosphères stellaires, c'est pourquoi nous présenterons dans ce chapitre des exemples numériques consacrés à ce problème. Notons que les atmosphères stellaires ne sont pas le seul domaine d'application de cette théorie en Astrophysique puisque les problèmes des atmosphères planétaires et cométaires, des milieux circumstellaires (étoiles en formation, étoiles à forte perte de masse) des milieux interstellaires (rayonnement galactique diffus, nuages interstellaires) et enfin des noyaux et disques galactiques, font appel à la théorie du transfert, c'est-à-dire, *in fine*, à la résolution de l'équation de transfert que nous allons présenter dans la section suivante.

3.2 L'équation de transfert

La théorie du transfert décrit un champ de rayonnement par ses propriétés en direction et en fréquence en chaque point r et à chaque instant, c'est-à-dire par les propriétés de radiations de direction $\vec{n}$ et de fréquence ν qui atteignent ou quittent ce point à cet instant. Par exemple, l'intensité spécifique, notée I dans la suite, mesure la quantité d'énergie transportée par le champ à travers un élément de surface, dans un cône (élémentaire) de directions et dans un intervalle élémentaire de fréquence. Cette grandeur fondamentale de la théorie du transfert est donc une fonction de la position et de la direction, ainsi que de la fréquence et du temps. Elle vérifie l'équation de transfert, qui consiste en un problème pseudo-intégro-différentiel que nous allons formuler.

3.2.1 Formulation

Nous supposons que le milieu irradié satisfait aux hypothèses «géométriques» suivantes :

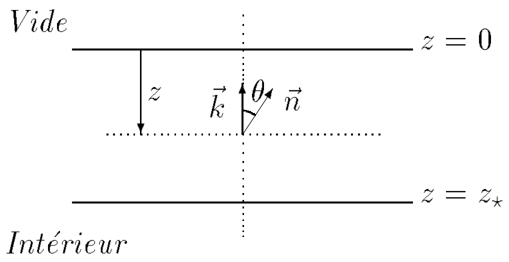
(G_1) le milieu est à symétrie plan-parallèle;

(G_2) il est statique et stationnaire en moyenne;

(G_3) les effets d'anisotropie de la redistribution des photons sont négligés;

(G_4) les conditions aux limites sont à symétrie axiale.

La première hypothèse implique que l'on peut considérer que le milieu est stratifié en couches planes, parallèles, sur lesquelles il est homogène. Les propriétés d'un milieu présentant la symétrie plan-parallèle ne dépendent que de la coordonnée z le long de la normale au plan de stratification, de vecteur unité $\vec{k}$ dirigé vers l'extérieur. Cette coordonnée varie dans un intervalle borné $[0, z_\star]$ si le milieu est limité par deux plans frontières distants de $z_\star \in]0, +\infty[$, ce que nous supposerons. Le champ radiatif a également une dépendance spatiale en $z \in [0, z_\star]$ car nous supposons que les conditions aux limites sont homogènes sur chaque plan frontière. La deuxième hypothèse exclue, entre autres, tout mouvement macroscopique de matière à grande échelle dans le milieu. La troisième hypothèse signifie que la diffusion des photons sur chaque centre diffuseur est isotrope, ce qui veut dire que toutes les directions diffusées sont équiprobables. La quatrième hypothèse implique que les conditions aux limites ne font intervenir que l'angle d'inclinaison des rayons par rapport à la normale $\vec{k}$ aux plans frontière, l'angle d'azimut n'intervenant pas. En tenant compte des autres symétries, le champ radiatif possède la même propriété à tous les niveaux z.



Ces quatre premières hypothèses entraînent que l'intensité spécifique I du champ de rayonnement ne dépend que de trois variables : la hauteur $z \in [0, z_\star]$, le cosinus $\mu \in [-1, 1]$ de l'angle d'inclinaison θ des rayons par rapport à la normale $\vec{k}$ dirigée vers l'extérieur du milieu, et la fréquence ν. Elle vérifie l'équation de transfert qui décrit la propagation de l'énergie portée par une radiation dans une direction et une à fréquence donnée. Sous les hypothèses précédentes, cette équation prend la forme simplifiée suivante : pour tout $z \in]0, z_\star[$, $\nu \in]0, +\infty[$ et $\mu \in [-1, 1]$,

$$\mu \frac{\partial I}{\partial z}(z, \mu, \nu) = \chi(z, \nu) I(z, \mu, \nu) - E(z, \nu) - \int_0^{+\infty} \sigma(z, \nu', \nu) J(z, \nu')\, d\nu'. \tag{3.1}$$

L'équation de transfert exprime que la variation locale d'énergie transportée par une radiation provient du rayonnement créé par les sources situées au voisinage du point considéré : E est le coefficient d'émission (ou émissivité), σ le coefficient différentiel de diffusion et χ le coefficient volumique d'extinction (grandeur strictement positive encore appelée opacité). Ces trois coefficients décrivent la physique des interactions rayonnement-matière. En particulier, σ décrit le phénomène de diffusion des photons sur la couche z; cette diffusion est censée changer la fréquence du photon incident, qui passe de ν à ν'. J est l'intensité moyenne définie pour tout $z \in [0, z_\star]$ et $\nu \in]0, +\infty[$ par

$$J(z, \nu) := \frac{1}{2} \int_{-1}^{1} I(z, \mu, \nu)\, d\mu. \tag{3.2}$$

Les conditions aux limites sont les suivantes :

$$I(0, \mu, \nu) \;=\; I^-(\mu, \nu), \text{ donnée pour } -1 \leq \mu < 0 \text{ et } \nu \in]0, +\infty[, \tag{3.3}$$

$$I(z_\star, \mu, \nu) \;=\; I^+(\mu, \nu), \text{ donnée pour } 0 < \mu \leq 1 \text{ et } \nu \in]0, +\infty[, \tag{3.4}$$

où I^+ et I^- désignent l'intensité spécifique du rayonnement incident sur les deux plans frontière.

Pour tout $z \in [0, z_\star]$, $\nu > 0$ et $\nu' > 0$, on définit le coefficient de diffusion intégré sur la fréquence du photon diffusé par

$$\sigma(z,\nu) := \int_0^{+\infty} \sigma(z,\nu,\nu')\,d\nu', \tag{3.5}$$

et l'indicatrice de diffusion normalisée p par

$$p(z,\nu',\nu) := \frac{\sigma(z,\nu',\nu)}{\sigma(z,\nu')}. \tag{3.6}$$

On remarque que pour tout $z \in [0, z_\star]$ et $\nu > 0$,

$$\int_0^{+\infty} p(z,\nu',\nu)\,d\nu = 1. \tag{3.7}$$

L'opacité χ étant une grandeur strictement positive, nous pouvons la factoriser dans (3.1); ainsi pour tout $z \in]0, z_\star[$, $\nu \in]0, +\infty[$ et $\mu \in [-1, 1]$,

$$\mu \frac{\partial I}{\partial z}(z,\mu,\nu) = \chi(z,\nu)\left[I(z,\mu,\nu) - \frac{E(z,\nu)}{\chi(z,\nu)} - \frac{\sigma(z,\nu)}{\chi(z,\nu)}\int_0^{+\infty} p(z,\nu',\nu)J(z,\nu')\,d\nu'\right]. \tag{3.8}$$

Pour tout $z \in [0, z_\star]$ et $\nu > 0$, on définit de plus la fonction source primaire

$$f_\star(z,\nu) := \frac{E(z,\nu)}{\chi(z,\nu)}, \tag{3.9}$$

ainsi que l'albédo

$$\varpi(z,\nu) := \frac{\sigma(z,\nu)}{\chi(z,\nu)}. \tag{3.10}$$

Lorsque la diffusion est supposée monochromatique (pas de redistribution en fréquence des photons), le coefficient p se réduit à la distribution de Dirac

$$p(z,\nu',\nu) := \delta(\nu' - \nu), \quad z \in [0, z_\star],\ \nu > 0,\ \nu' > 0, \tag{3.11}$$

et l'équation (3.8) devient

$$\mu \frac{\partial I}{\partial z}(z,\mu) = \chi(z)\left(I(z,\mu) - f_\star(z) - \varpi(z)J(z)\right), \quad z \in]0, z_\star[,\ \mu \in [-1, 1]. \tag{3.12}$$

En effet, la fréquence n'étant plus transformée, elle n'a plus besoin d'apparaître explicitement dans les grandeurs intervenant dans l'équation de transfert. Dans la suite nous adoptons ce modèle très simple de diffusion car il contient en substance toute les difficultés posées par la résolution de l'équation de transfert. Nous nous intéressons donc exclusivement à la variante (3.12) de l'équation de transfert initiale (3.1).

Nous allons maintenant effectuer un changement de variable dans cette équation. On associe à l'opacité χ l'échelle de profondeur optique

$$\Theta(z) := \int_0^z \chi(\zeta)\, d\zeta, \quad z \in [0, z_\star]. \tag{3.13}$$

On suppose que l'opacité est une fonction continue de la variable z ce qui provient de l'hypothèse (G_2). Alors Θ est dérivable sur $[0, z_\star]$ et pour tout $z \in [0, z_\star]$

$$\frac{\mathrm{d}\Theta}{\mathrm{d}z}(z) = \chi(z) > 0. \tag{3.14}$$

Ainsi, Θ est une bijection continue strictement croissante de l'intervalle $[0, z_\star]$ sur son image $[0, \tau_\star]$, où

$$\tau_\star := \int_0^{z_\star} \chi(\zeta)\, d\zeta \tag{3.15}$$

désigne l'épaisseur optique de l'atmosphère. Ainsi $\tau_\star$ est finie car $z_\star$ l'est et que l'opacité χ est une fonction continue sur $[0, \tau_\star]$. Comme l'intensité spécifique est une fonction dérivable en tout point z de $]0, z_\star[$ et pour tout $\mu \neq 0$, la fonction

$$\begin{array}{rlcc} \widetilde{I}: & [0, \tau_\star] \times [-1, 1], & \to & \mathbb{R} \\ & (\tau, \mu) & \mapsto & I(\Theta^{-1}(\tau), \mu) \end{array} \tag{3.16}$$

est une fonction dérivable par rapport à sa première variable pour tout $\tau \in]0, \tau_\star[$ et μ non nul. De plus, pour tout $\tau \in]0, \tau_\star[$ et $\mu \in [-1, 1]$ non nul,

$$\frac{\partial \widetilde{I}}{\partial \tau}(\tau, \mu) = \frac{1}{\chi \circ \Theta^{-1}(\tau)} \frac{\partial I}{\partial z}(\Theta^{-1}(\tau), \mu). \tag{3.17}$$

Enfin, d'après l'équation (3.12),

$$\mu \frac{\partial \widetilde{I}}{\partial \tau}(\tau, \mu) = \widetilde{I}(\tau, \mu) - (f_\star - \varpi J) \circ \Theta^{-1}(\tau). \tag{3.18}$$

Afin d'alléger les écritures, toute fonction h de la variable z désignera $h \circ \Theta^{-1}$. De même, nous continuerons à désigner $\widetilde{I}$ par I. Ces fonctions deviennent alors des fonctions de la variable τ. En prenant en compte cet abus d'écriture, l'équation de transfert s'écrit, en terme de la variable τ, de la manière suivante :

$$\mu \frac{\partial I}{\partial \tau}(\tau, \mu) = I(\tau, \mu) - f_\star(\tau) - \varpi(\tau) J(\tau), \quad \tau \in]0, \tau_\star[, \ \mu \in [-1, 1], \tag{3.19}$$

dont les conditions aux limites sont

$$I(0, \mu) = I^-(\mu), \text{ donnée pour } -1 \leq \mu < 0, \tag{3.20}$$

$$I(\tau_\star, \mu) = I^+(\mu), \text{ donnée pour } 0 < \mu \leq 1. \tag{3.21}$$

3.2.2 Formulation intégrale

On définit la fonction source φ du champ radiatif pour tout $\tau \in [0, \tau_\star]$ par

$$\varphi(\tau) := f_\star(\tau) + \varpi(\tau) J(\tau). \tag{3.22}$$

D'après (3.19) - (3.21), I est donnée en fonction de φ par la formule suivante :

$$I(\tau, \mu) = \begin{cases} I^-(\mu) \exp\left(\dfrac{\tau}{\mu}\right) - \dfrac{1}{\mu} \displaystyle\int_0^\tau \varphi(\sigma) \exp\left(\dfrac{\tau - \sigma}{\mu}\right) d\sigma & \text{si } \mu < 0, \\ \varphi(\tau) & \text{si } \mu = 0, \\ I^+(\mu) \exp\left(\dfrac{\tau - \tau_\star}{\mu}\right) + \dfrac{1}{\mu} \displaystyle\int_\tau^{\tau_\star} \varphi(\sigma) \exp\left(\dfrac{\tau - \sigma}{\mu}\right) d\sigma & \text{si } \mu > 0, \end{cases} \tag{3.23}$$

valable pour tout $\tau \in [0, \tau_\star]$. On obtient, par intégration par rapport à μ sur $[-1, +1]$ une expression de J en fonction de φ :

$$J(\tau) = J_\pm(\tau) + \frac{1}{2} \int_0^{\tau_\star} E_1(|\tau - \sigma|) \varphi(\sigma) \, d\sigma, \quad \tau \in [0, \tau_\star], \tag{3.24}$$

où, pour tout $\tau \in [0, \tau_\star]$,

$$J_\pm(\tau) := \frac{1}{2} \int_{-1}^0 I^-(\mu) \exp(\tau/\mu) \, d\mu + \frac{1}{2} \int_0^{+1} I^+(\mu) \exp[-(\tau_\star - \tau)/\mu] \, d\mu. \tag{3.25}$$

E_1 est la première fonction exponentielle intégrale; elle est définie pour tout τ strictement positif par

$$E_1(\tau) := \int_0^1 \frac{\exp(-\tau/\mu)}{\mu}\, d\mu. \tag{3.26}$$

Nous présentons dans la section 3.4.1 des propriétés élémentaires de cette fonction faiblement singulière en zéro. Enfin, après substitution de (3.24) dans (3.22), on obtient l'équation intégrale suivante :

$$\frac{\varpi(\tau)}{2}\int_0^{\tau_\star} E_1(|\tau-\sigma|)\varphi(\sigma)\, d\sigma - \varphi(\tau) = f(\tau), \quad \tau \in [0, \tau_\star], \tag{3.27}$$

où le terme libre f est défini pour tout $\tau \in [0, \tau_\star]$, par

$$f(\tau) := -f_\star(\tau) - \varpi(\tau) J_\pm(\tau). \tag{3.28}$$

Puisque $\tau_\star$ est fini, l'équation (3.27) est une équation de Fredholm de second espèce à noyau de convolution faiblement singulier.

3.3 Application à la théorie des atmosphères stellaires

Le rayonnement qui quitte les étoiles provient des couches superficielles qui constituent par définition leur atmosphère. Ces dernières assurent le lien entre les intérieurs (non directement observés) et le milieu interstellaire. Ainsi les modèles d'atmosphères stellaires sont capitaux pour la compréhension de la structure interne des étoiles et de leur évolution. Ces modèles reposent sur une bonne description du transfert radiatif de l'énergie, car celui-ci joue un rôle crucial dans l'établissement des divers équilibres présents. Si la description de tels phénomènes est relativement facile dans les couches profondes de l'étoile (dans ces milieux le libre parcours moyen des photons est petit devant les distances sur lesquelles varient les principales grandeurs thermodynamiques) ceci n'est plus du tout le cas lorsqu'on s'approche des couches superficielles de l'étoile. Dans ces régions très diluées, l'équation de transfert doit être résolue conjointement aux diverses équations d'équilibre du milieu, le problème global peut alors comprendre jusqu'à 100 millions d'équations !

Un modèle d'atmosphère stellaire est défini par la donnée d'un intervalle de fréquence contenant les fréquences dans lesquelles l'étoile rayonne, et d'un intervalle de profondeur optique $[0, \tau_\star]$ correspondant aux couches superficielles d'où sort le rayonnement.

La théorie classique des atmosphères stellaires repose sur les quatre hypothèses « géométriques » (G_1), (G_2), (G_3), (G_4) du paragraphe 3.2.1. La première hypothèse est valable quand on néglige les effets de courbures (ce qui n'est pas permis dans le cas des étoiles géantes et supergéantes par exemple). On peut alors considérer qu'une atmosphère est stratifiée en couches planes, parallèles entre elles, d'étendue infinie, sur lesquelles elles sont homogènes. La deuxième hypothèse est justifiée pour les étoiles non variables, dont les couches externes sont stables et la luminosité sensiblement constante dans le temps (en moyenne seulement). Elle exclue notamment tout mouvement macroscopique de matière à grande échelle dans l'atmosphère. La troisième hypothèse est couramment faite en théorie des atmosphères stellaires car les centres diffuseurs sont des électrons libres, des ions ou des atomes, qui ont des tailles bien plus faibles que la longueur d'onde des photons dans l'intervalle de fréquence où l'étoile rayonne. La quatrième hypothèse est également vérifiée dans les atmosphères stellaires.

Afin d'appréhender l'ampleur et la difficulté du problème des atmosphères stellaires, nous allons énumérer les différentes équations (couplées ...) intervenant dans cette modélisation. Regroupons tout d'abord les fonctions inconnues en trois catégories :

(α) l'intensité spécifique du champ radiatif I;

(β) les densités numériques n_i des particules lourdes (ions, atomes, molécules dans les atmosphères les plus froides) et celle des électrons libres n_e, c'est-à-dire leur nombre moyen par unité de volume;

(γ) la température électronique T_e.

Les équations des atmosphères stellaires forment un système de $N+3$ équations à $N+6$ inconnues, N désignant le nombre de type de particules lourdes présentes. Ces équations sont :

L'équation de transfert

Sous les hypothèses précédentes, elle s'écrit sous la forme simplifiée (3.12) de la section 3.2.1.

Équations de l'équilibre statistique

La densité de chaque type i de particule lourde intervient dans les N équations suivantes :

$$n_i(z)\sum_{i\neq e} P_{ij}(z) = \sum_{j\neq e} n_j(z)P_{ji}(z), \quad (i \neq e), \quad z \in [0, z_\star], \tag{3.29}$$

où les coefficients P_{ij} sont les taux de transition moyens du type i au type j. Ces équations forment un système de rang $N-1$ et traduisent le fait que ces particules ne sont créées nulle part en moyenne dans un milieu stationnaire et statique (également en moyenne).

L'équation de la neutralité électrique

Les atmosphères sont électriquement neutres en moyenne, car très conductrices : toute accumulation de charge en un point engendre des courants électriques qui les dispersent. D'où la condition de neutralité électrique :

$$\sum_i n_i q_i = 0, \tag{3.30}$$

où q_i désigne la charge de la particule de type i.

L'équation de l'équilibre hydrostatique

Les forces de pression sont équilibrées par les forces gravitationnelles (connues), de sorte que l'atmosphère est en équilibre hydrostatique :

$$\frac{\mathrm{d}P}{\mathrm{d}z} = -\rho(z)g(z), \quad z \in [0, z_\star]. \tag{3.31}$$

P désigne la pression totale du gaz et g l'accélération gravitationnelle sur une couche z fixée. La masse volumique ρ est donnée par la relation suivante

$$\rho(z) = \sum_i m_i n_i(z), \quad z \in [0, z_\star], \tag{3.32}$$

où m_i désigne la masse de la particule de type i. La pression totale comprend la pression cinétique P_c et la pression de radiation P_r :

$$P(z) := P_c(z) + P_r(z), \quad z \in [0, z_\star], \tag{3.33}$$

la pression de radiation se déduisant de l'intensité spécifique par la relation

$$P_r(z) = \frac{2\pi}{c} \int_0^{+\infty} \int_{-1}^{+1} I(z, \mu, \nu) \mu^2 \, d\mu \, d\nu, \tag{3.34}$$

où c est la vitesse de la lumière.

L'équation d'état

La pression cinétique P_c peut être calculée à l'aide de la loi des gaz parfaits, car les atmosphères sont des milieux très dilués assimilables à des plasmas cinétiques classiques :

$$P_c(z) = n(z)kT_e(z), \quad z \in [0, z_\star], \tag{3.35}$$

où k est la constante de Bolzmann et

$$n(z) := \sum_i n_i(z), \quad z \in [0, z_\star], \tag{3.36}$$

est la densité totale des particules lourdes.

L'équation d'énergie

L'équation suivante traduit la conservation de l'énergie dans une atmosphère stellaire :

$$\int_0^{+\infty} E(z,\nu)\, d\nu = \int_0^{+\infty} (\chi(z,\nu) - \sigma(z,\nu)) J(z,\nu)\, d\nu + \epsilon_c(z), \quad z \in [0, z_\star]. \tag{3.37}$$

$\epsilon_c(z)$ désigne la quantité d'énergie non radiative échangée par unité de volume et de temps au point z, énergie transportée essentiellement par convection.

Ces équations appartiennent à trois grands groupes d'équations, chacun d'eux fournissant une des quantités (α), (β) et (γ) introduites à la section 3.2.1 :

premier groupe : il comprend l'équation de transfert (3.1) et la relation (3.2). Elles permettent de calculer (α) en fonction de (β) et (γ);

deuxième groupe : il comprend les équations (3.29) à (3.36), dites «structurelles». Elles permettent de déterminer (β) connaissant (α) et (γ);

troisième groupe : l'équation d'énergie (3.37). Elle permet de déterminer (γ) connaissant (α) et (β).

Ces équations fortement couplées forment un système globalement non linéaire :

- le champ radiatif dépend linéairement des coefficients volumique d'extinction et d'émission du milieu via l'équation de transfert;
- ces coefficients dépendent linéairement des densités particulaires et de la température;
- les densités et la température dépendent de façon non linéaire du champ radiatif.

Pour résoudre ce système de manière itérative, on se donne une distribution simple de température et de population pour résoudre une première fois l'équation de transfert. Les itérations suivantes sont facilitées par un regroupement de ces équations dans les trois catégories décrites plus haut. On commence d'abord par résoudre les équations contenues dans chaque groupe, ce qui permet de déterminer l'une des quantités (α), (β) ou (γ) connaissant les deux autres. On s'intéresse ensuite à la résolution simultanée des équations de deux des trois groupes, ce qui revient à calculer deux des quantités (α), (β) et (γ) connaissant la troisième :

- Calcul de I et T_e pour des densités n_e et n_i données, en résolvant l'équation de transfert et l'équation d'énergie couplées;
- Calcul de I et des des densités n_e et n_i pour une température T_e donnée, en résolvant l'équation de transfert conjointement aux équations de l'équilibre statistique, de l'équilibre hydrostatique et de neutralité électrique. Ce deuxième problème se résout en deux étapes :
  - On se donne la densité particulaire totale n, et on résout les équations de transfert, de l'équilibre statistique et de neutralité électrique couplées. Ce problème est appelé «problème restreint» en théorie des atmosphères stellaires;
  - On calcule n à l'aide de l'équation de l'équilibre hydrostatique et on itère.
- Calcul des densités n_e et n_i ainsi que de la température T_e pour un champ radiatif connu.

L'équation de transfert doit donc être résolue à chaque fréquence pour la détermination des grandeurs thermodynamiques de l'atmosphère (température, pression). Cela implique un nombre très élevé de résolutions, jusqu'à plusieurs centaines de millions ! Heureusement, des méthodes statistiques (voir [25]) permettent de réduire le nombre d'équations de transfert à résoudre, ce dernier restant néanmoins de l'ordre de 10000 ... Cela motive la construction de méthodes rapides pour la résolution de l'équation de transfert.

Commentaires bibliographiques

La théorie du transfert a été développée au début du XX^{e} siècle par des astrophysiciens soucieux de comprendre la physique des atmosphères stellaires. On peut considérer qu'elle est née de deux articles de K. Schwarzschild datant de 1906 et 1914, immédiatement suivis des travaux d'Eddington, Jeans, Milne, ... portant sur les atmosphères et les intérieurs stellaires en équilibre radiatif ([55]). Milne et Hopf publient, dans les années trente, les premiers travaux de synthèse du point de vue physique ([58]) et mathématique ([44]) respectivement. À la même époque, on découvre le neutron et la possibilité que d'autres particules que les photons subissent des diffusions multiples. La théorie du transport des neutrons progresse rapidement pendant la seconde guerre mondiale, du fait du projet Manhattan, aux tragiques retombées militaires. Un deuxième élan lui est donné à la fin des années cinquante, avec le développement puis la commercialisation des réacteurs nucléaires. En fait, les physiciens du transfert et ceux du transport ont élaboré une même théorie, dite du transport des particules, pour l'étude de la diffusion multiple d'un ensemble dilué de particules sur un autre ensemble de particules. Cette théorie s'intéresse aussi bien à la propagation de la lumière dans une atmosphère qu'à la pénétration des neutrons dans la matière, ou à la diffusion d'un gaz sous l'influence d'un champ de force extérieur. Ces deux points font la spécificité de la théorie du transport au sein de la beaucoup plus vaste théorie cinétique des gaz. À la fin des années cinquante, les ouvrages de Chandrasekhar, Kourganoff ([46]), Sobolev, Busbridge (sur les aspects mathématiques dans [23]) pour l'astrophysique, Case, Hoffmann et Placzek, Davidson, pour la neutronique, constituent les références de base d'une bibliographie qui s'enrichie rapidement des contributions de disciplines diverses et de nombreux ouvrages généraux ([66], [62], [80], [31], [29] pour les aspects mathématiques, [76] et [56] sur la théorie des atmosphères stellaires).

3.4 Résolution Numérique

3.4.1 Propriétés des fonctions exponentielles intégrales

La fonction E_1 est la première fonction exponentielle intégrale de la suite de fonctions $(E_\nu)_{\nu\geq 1}$ définie par

$$E_\nu(\tau) := \int_0^1 \frac{\exp(-\tau/\mu)}{\mu^{2-\nu}}\, d\mu = \int_1^{+\infty} \frac{\exp(-\tau\mu)}{\mu^\nu}\, d\mu, \quad \nu \geq 1, \tag{3.38}$$

pour tout $\tau \in]0, \tau_\star]$ si $\nu = 1$, et pour tout $\tau \in [0, \tau_\star]$ si $\nu \geq 2$. Elle est la seule présentant une singularité en $\tau = 0$. En fait, dans un voisinage de 0, E_1 a un comportement logarithmique :

$$\lim_{\tau\to 0^+} \frac{E_1(\tau)}{-\ln\tau} = \lim_{\tau\to 0^+} \frac{\displaystyle\int_1^{+\infty} \exp(-\tau\mu)\, d\mu}{\frac{1}{\tau}} = \lim_{\tau\to 0^+} e^{-\tau} = 1. \tag{3.39}$$

Soit Ei la fonction définie pour tout $x > 0$ par

$$\mathrm{Ei}(\tau) := \int_{-\infty}^{\tau} \frac{\exp(\mu)}{\mu}\, d\mu. \tag{3.40}$$

La fonction Ei est prolongée sur $]-\infty, 0[$ en utilisant la fonction E_1 grâce à la relation suivante :

$$\mathrm{Ei}(-\tau) = -E_1(\tau), \quad \tau > 0. \tag{3.41}$$

Les fonctions suivantes nous serons utiles par la suite :

$$F_\nu(\tau,\alpha) := \int_0^\tau \exp(\alpha\sigma) E_\nu(\sigma)\, d\sigma, \quad \tau \geq 0, \quad \alpha \in \mathbb{R}. \tag{3.42}$$

Pour tout $\tau > 0$,

(3.43)
$$F_1(\tau,\alpha) = \begin{cases} \dfrac{1}{\alpha}[\exp(\alpha\tau)E_1(\tau) - E_1(\tau-\alpha\tau) - \ln(1-\alpha)], & \text{si } \alpha < 1 \text{ et } \alpha \neq 0, \\ 1 - E_2(\tau), & \text{si } \alpha = 0, \\ \exp(\tau)E_1(\tau) + \ln(x) + \gamma, & \text{si } \alpha = 1, \\ \dfrac{1}{\alpha}[\exp(\alpha\tau)E_1(\tau) + \mathrm{Ei}(\alpha\tau-\tau) - \ln(\alpha-1)], & \text{si } \alpha > 1, \end{cases}$$

où $\gamma := 0.5772156649\ldots$ est la constante d'Euler (on trouvera dans [81] une valeur de la constante d'Euler avec 328 chiffres significatifs). On définit les fonctions F_ν pour $\nu \geq 2$ par la relation de récurrence suivante :

(3.44)
$$\alpha F_{\nu+1}(\tau,\alpha) = \exp(\alpha\tau)E_{\nu+1}(\tau) - \frac{1}{\nu} + F_\nu(\tau,\alpha), \quad \tau \geq 0, \quad \alpha \in \mathbb{R}.$$

Nous rappelons quelques propriétés des fonctions exponentielles intégrales (voir [1] et [78]) :

$$\begin{aligned}
&(3.45) & E'_{\nu+1} &= -E_\nu, \quad \nu > 1, \\
&(3.46) & E_\nu(0) &= \frac{1}{\nu-1}, \quad \nu > 1, \\
&(3.47) & \tau E_\nu(\tau) &= e^{-\tau} - \nu E_{\nu+1}(\tau), \quad \tau > 0, \quad \nu \geq 1, \\
&(3.48) & E_{\nu+1}(\tau) &= \frac{e^{-\tau}}{\nu!}\sum_{k=0}^{\nu-1}[\nu-(k+1)]!(-\tau)^k + \frac{(-\tau)^\nu}{\nu!}E_1(\tau), \quad \tau \geq 0, \quad \nu \geq 1.
\end{aligned}$$

Soient a et b deux réels tels que $a < b$ et $\tau \in \mathbb{R}$. Alors,

(3.49)
$$\int_a^b E_\nu(|\tau-\sigma|)\, d\sigma = \begin{cases} \dfrac{2}{\nu} - E_{\nu+1}(\tau-a) - E_{\nu+1}(b-\tau) & \text{si } \tau \in [a,b], \\ |E_{\nu+1}(|\tau-a|) - E_{\nu+1}(|\tau-b|)| & \text{sinon.} \end{cases}$$

Soient a et b deux réels tels que $a < b$. Alors, quel que soit $\tau \in \mathbb{R}$,

$$\int_a^b (\tau-\sigma)E_\nu(|\tau-\sigma|)\,d\sigma = e^{-|\tau-b|} - e^{-|\tau-a|} + \nu[E_{\nu+2}(|\tau-a|) - E_{\nu+2}(|\tau-b|)].$$

On définit pour tout $j > 0$ et tout $\tau \in [0, \tau_\star]$,

$$E_{1,j}(\tau) := \frac{2}{j!}\int_\tau^{+\infty} E_1(\sigma)\sigma^j\,d\sigma. \tag{3.50}$$

On peut montrer ([1]) que

$$E_{1,j}(\tau) = \frac{1}{2}\sum_{k=0}^{j} \frac{\tau^k}{k!} E_{j+2-k}(\tau). \tag{3.51}$$

On remarque que pour tout $j > 0$,

$$E_{1,j}(0) = \frac{1}{2}\frac{1}{j+1}, \tag{3.52}$$

d'après (3.46). Soit $(H_\nu)_{\nu\geq 1}$ la suite de fonctions définies pour tout $b \geq 0$, $\alpha \in \mathbb{R}$ et $\tau \in \mathbb{R}$ par :

$$H_\nu(b,\alpha,\tau) := \int_0^b \exp(\alpha\sigma)E_\nu(|\tau-\sigma|)\,d\sigma. \tag{3.53}$$

Alors pour tout $b \geq 0$, $\alpha \in \mathbb{R}$ et $\tau \geq 0$,

(3.54)
$$H_\nu(b,\alpha,\tau) = \begin{cases} \exp(\alpha\tau)[F_\nu(\tau,-\alpha) + F_\nu(b-\tau,\alpha)] & \text{si } \tau \leq b, \\ \\ \exp(\alpha\tau)[F_\nu(\tau,-\alpha) - F_\nu(\tau-b,-\alpha)] & \text{si } \tau \geq b. \end{cases}$$

Le calcul des coefficients des différentes matrices et vecteurs impliqués dans la résolution numérique de ce problème nécessite l'évaluation de ces fonctions. Nous nous servirons des relations suivantes pour obtenir une approximation de ces évaluations. La fonction E_ν est un cas particulier de la fonction gamma incomplète :

$$E_\nu(\tau) = e^{-\tau}\left(\frac{1}{\tau+\nu}\ \frac{\nu}{\tau+\nu+2-}\ \frac{2(\nu+1)}{\tau+\nu+4-}\cdots\right), \quad \tau > 0.$$

Cette fraction continue ne converge assez rapidement, pour des besoins pratiques, que si $\tau \geq 1$. Pour des valeurs de τ plus petites que 1, on utilise la représentation en série suivante :

(3.55)
$$E_\nu(\tau) = \frac{(-\tau)^{\nu-1}}{(\nu-1)!}(-\ln\tau + \Psi(\nu)) - \sum_{m\notin\{0,\nu-1\}}^{\infty} \frac{(-\tau)^m}{(m-\nu+1)m!}, \quad \tau < 1,$$

où Ψ est la fonction digamma, définie pour tout entier $\nu \geq 1$ par

$$\begin{aligned} \Psi(1) &= -\gamma, \\ \Psi(\nu) &= -\gamma + \sum_{m=1}^{\nu-1} \frac{1}{m}, \quad \nu \geq 1. \end{aligned} \tag{3.56}$$

En pratique, on évalue (3.55) dans l'ordre des puissances ascendantes de τ :

$$\begin{aligned} E_\nu(\tau) &= -\left[\frac{1}{\nu-1} - \frac{\tau}{2-\nu} + \frac{\tau^2}{2(3-\nu)} - \cdots + \frac{(-\tau)^{\nu-2}}{-(\nu-2)!}\right] \\ &\quad + \frac{(-\tau)^{\nu-1}}{(\nu-1)!}(-\ln\tau + \Psi(\nu)) - \left[\frac{(-\tau)^\nu}{\nu!} + \frac{(-\tau)^{\nu+1}}{2(\nu+1)!} + \ldots\right]. \end{aligned}$$

L'expression contenue dans le premier crochet n'est pas prise en compte lorsque $\nu = 1$. On obtient une approximation de $E_\nu(\tau)$, pour $\tau < 1$ en tronquant la série précédente lorsque le dernier terme calculé est plus petit que la précision désirée. Un autre critère d'arrêt est décrit dans [10] et [9]. On utilise les expressions suivantes pour calculer une bonne approximation de Ei :

Pour $\tau > 0$ tel que $\tau < |\ln\epsilon|$, où ϵ est la précision désirée,

$$\mathrm{Ei}(\tau) = \gamma + \ln(\tau) + \sum_{k=1}^{\infty} \frac{\tau^k}{k \times k!}. \tag{3.57}$$

Pour $\tau > |\ln\epsilon|$, on utilise l'expansion asymptotique suivante :

$$\mathrm{Ei}(\tau) \sim \frac{\exp(\tau)}{\tau} \sum_{k=0}^{n} \frac{k!}{\tau^k}, \tag{3.58}$$

où nous utilisons les conventions de notation $0! := 1$ et $\tau^0 := 1$, pour $\tau > 0$. On trouvera dans [1] et [78] de plus amples informations sur ce type de fonctions.

3.4.2 Problèmes choisis

Nous allons appliquer les méthodes numériques décrites dans le chapitre précédent à la résolution de l'équation de transfert sous sa forme intégrale (3.27) que nous rappelons :

$$\frac{\varpi(\tau)}{2} \int_0^{\tau_\star} E_1(|\tau - \sigma|)\varphi(\sigma)\, d\sigma - \varphi(\tau) = f(\tau), \quad \tau \in [0, \tau_\star].$$

Le noyau κ intervenant dans cette équation intégrale, est défini pour tout τ et σ dans $[0, \tau_\star]$, tels que $\tau \neq \sigma$ par

$$\kappa(\tau, \sigma) := \frac{\varpi(\tau)}{2} E_1(|\tau - \sigma|). \tag{3.59}$$

La fonction E_1 satisfait les hypothèses (1.8)- (1.11) (voir la section 3.4.1 précédente) et, d'après l'hypothèse (G_2) de la section 3.2.1, l'albédo est une fonction continue. On peut alors identifier les fonctions η et g des chapitres précédents à

$$\eta(\tau,\sigma) := \frac{\varpi(\tau)}{2}, \tag{3.60}$$

$$g(\tau) := E_1(\tau),\ \tau > 0, \tag{3.61}$$

et appliquer les méthodes décrites dans le chapitre 2 pour résoudre numériquement ce problème dès que le second membre f est une fonction continue sur $[0,\tau_\star]$ ou une représentante de la classe des fonctions intégrables, au sens de Lebesgue, sur $[0,\tau_\star]$.

Dans un milieu réaliste où l'albédo varie, il n'y a pas de loi simple pour exprimer la variation spatiale de ce dernier. À chaque fréquence correspond une loi de variation, des lois associées à deux fréquences différentes pouvant être très différentes. Néanmoins, le comportement général de l'albédo est tel qu'il tend vers 1 au voisinage de la surface et vers une valeur finie, parfois très proche de zéro, loin des couches superficielles (pour des valeurs de τ élevées). Une loi simple décrivant ce type de comportement est une exponentielle négative, systématiquement prise en compte par les travaux de théorie du transfert qui considèrent un albédo variable ([36]). Nous définissons donc ϖ, pour tout $\tau \in [0,\tau_\star]$, par :

$$\varpi(\tau) := \varpi_\star \exp(\alpha\tau), \tag{3.62}$$

où $\alpha \leq 0$ et $\varpi_\star \in\ ]0,1[$. Ainsi, pour tout $\tau \in [0,\tau_\star]$,

$$0 < \varpi(\tau) \leq \varpi_\star < 1. \tag{3.63}$$

Nous considérerons les cas particuliers où $\alpha = 0$ (albédo constant) et $\alpha = -1$.

Nous allons étudier cinq exemples de terme libres : si les trois derniers ont une signification physique particulière, les deux premiers ne présentent qu'un intérêt numérique (validation des méthodes). Soit f_1 le terme libre défini, pour tout τ dans $[0,\tau_\star]$, par

$$f_1(\tau) := \varpi(\tau) - 1 - \frac{\varpi(\tau)}{2}(E_2(\tau) + E_2(\tau_\star - \tau)). \tag{3.64}$$

Alors la solution φ_1 de l'équation (3.27) correspondante vérifie, pour tout $\tau \in [0,\tau_\star]$,

$$\varphi_1(\tau) = 1. \tag{3.65}$$

Soit f_2 le terme libre défini, pour tout τ dans $[0,\tau_\star]$, par

$$f_2(\tau) := \exp(\beta\tau)\Big[\frac{\varpi(\tau)}{2}(F_1(\tau,-\beta) + F_1(\tau_\star - \tau,-\beta)) - 1\Big], \tag{3.66}$$

où F_1 est définie par (3.42). Alors la solution φ_2 de l'équation (3.27) correspondante vérifie pour tout $\tau \in [0, \tau_\star]$,

$$\varphi_2(\tau) := \exp(\beta\tau). \tag{3.67}$$

On s'intéresse particulièrement au cas où φ_2 varie rapidement avec τ afin de simuler le comportement de la fonction source φ dans la région de transition atmosphère-couronne : on prendra $\beta = 5$. Dans le cas d'une atmosphère isotherme (température constante), le second membre de l'équation (3.27) est défini par

$$f_3(\tau) := \varpi(\tau) - 1. \tag{3.68}$$

La solution exacte correspondante sera appelée φ_3. Le terme libre suivant correspond à un saut brusque de la température sur une couche donnée $\tau = \theta$, où $\theta \in]0, \tau_\star]$ (qui peut correspondre à la zone de transition chromosphère-couronne dans une étoile) :

$$f_4(\tau) := \begin{cases} \varpi_\star - 1 & \text{si} \quad \tau \in [0, \theta], \\ 0 & \text{si} \quad \tau \in]\theta, \tau_\star]. \end{cases} \tag{3.69}$$

Nous étudions dans la suite le comportement de la solution φ_4 correspondante lorsque le paramètre θ tend vers zéro. Lorsque sur une couche $\tau = \theta$, où $\theta \in]0, \tau_\star]$, il existe une distribution uniforme de sources isotropes (sources internes), le terme libre est défini, pour tout $\tau \in [0, \tau_\star]$ tel que $\tau \neq \theta$, par

$$f_5(\tau) := -\frac{\varpi_\star}{2} E_1(|\tau - \theta|). \tag{3.70}$$

Ce terme libre intervient lors de la modélisation de nébuleuses plan-parallèles éclairées par une étoile ponctuelle placée sur la couche $\tau = \theta$. Nous étudierons le cas où $\theta = \tau_\star/2$. La solution exacte correspondante sera notée φ_5.

Le tableau suivant donne les espaces fonctionnels correspondants aux termes libres considérés :

Terme libre	Espace
f_1	$C^0([0, \tau_\star])$
f_2	$C^0([0, \tau_\star])$
f_3	$C^0([0, \tau_\star])$
f_4	$L^1([0, \tau_\star])$
f_5	$L^1([0, \tau_\star])$

Enfin, nous considérerons les valeurs suivantes pour les couples $(\varpi_\star, \tau_\star)$:

$\varpi_\star$	$\tau_\star$
0.7	10^2
0.999	10^3

Le premier couple est typique d'un problème posé à une fréquence appartenant à une partie continue du spectre de l'étoile (c'est-à-dire dans une région où les différentes grandeurs physiques varient peu avec la fréquence donc en dehors des raies spectrales). Le dernier est typique d'un problème de transfert posé dans une raie spectrale (correspondant à un profil de raie rectangulaire). On peut se référer à [36] et [28] pour plus de détails sur les domaines de variation de $\varpi_\star$ et $\tau_\star$.

Les coefficients des matrices et seconds membres des systèmes linéaires correspondants à ces termes libres et pour ce choix de ϖ sont donnés, pour chaque méthode, en annexe.

3.4.3 Approximations de rang fini

Les sept premières figures montrent l'erreur commise sur $\varphi_1(\tau) = 1$, $\tau \in [0, \tau_\star]$, par la méthode de projection décrite en 2.1.1.1 et par la méthode d'intégration produit décrite en 2.1.2.1 pour différentes valeurs de $\tau_\star$, ϖ et α.

La figure 3.1 montre l'erreur relative associée à l'approximation construite par la méthode de projection lorsque $\tau_\star = 100$ et que l'albédo est constamment égal à 0.7.

La figure 3.2 montre l'erreur relative associée à l'approximation construite par la méthode de projection lorsque $\tau_\star = 100$ et que l'albédo est variable : $\varpi(\tau) = 0.7\exp(-\tau)$. On remarque que dans ce cas l'erreur reste à son niveau le plus bas au voisinage de $\tau_\star$, et décroît plus vite au voisinage de zéro que dans le cas où l'albédo est constant, comme le montre la figure 3.3, tout en conservant le même ordre de grandeur. Ces approximations ont été construites à partir d'une grille quasi-uniforme définie de la manière suivante :

sous intervalles	nombre de points	pas
$[0, 10]$	251	0.04
$[10, 90]$	551	0.16
$[90, 100]$	251	0.04
taille de la matrice	1000	

Grille 1.

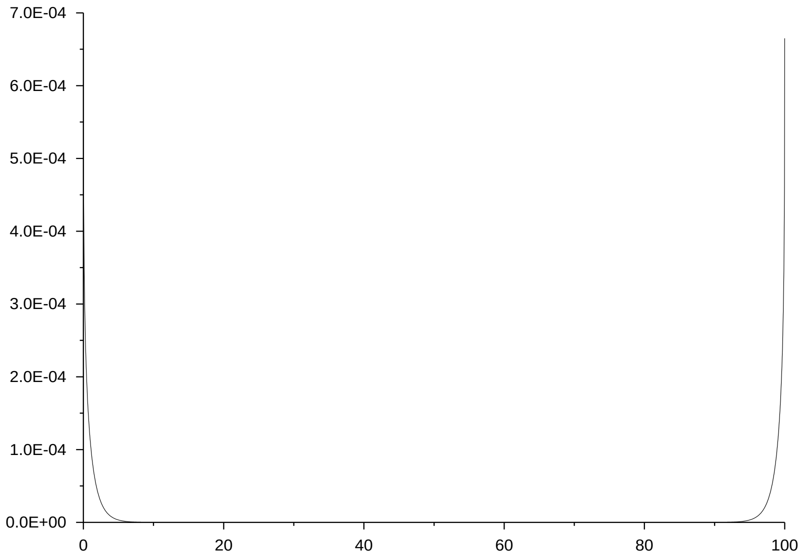


FIG. 3.1 – *Erreur relative associée à l'approximation de φ_1. Projection. Grille 1. $\tau_\star = 100$. $\varpi(\tau) = 0.7$.*

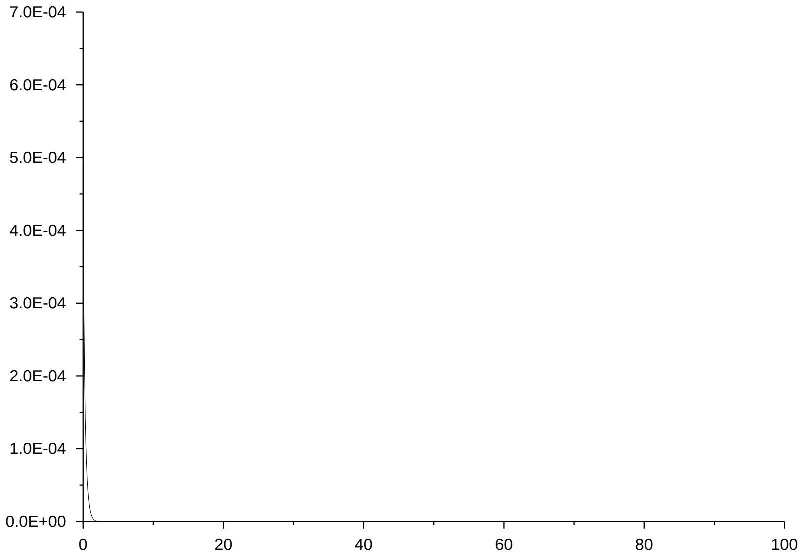


FIG. 3.2 – *Erreur relative associée à l'approximation de φ_1. Projection. Grille 1. $\tau_\star = 100$. $\varpi(\tau) = 0.7\exp(-\tau)$.*

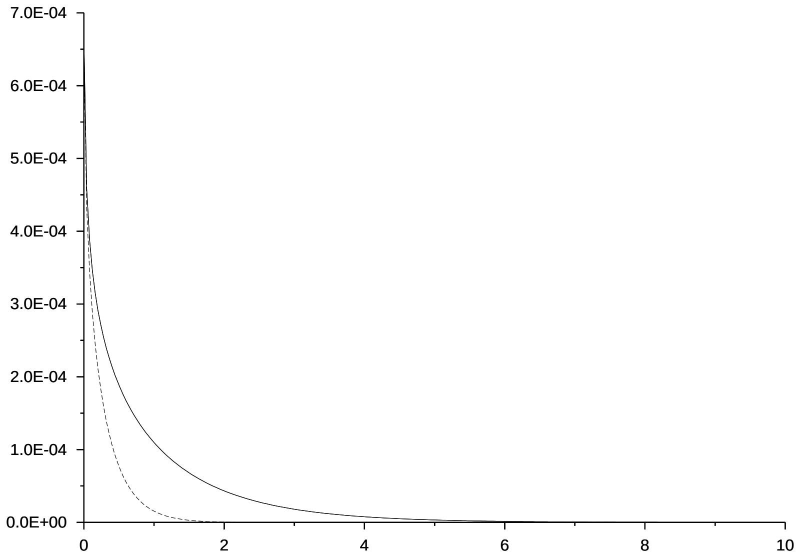


FIG. 3.3 – *Zoom des erreurs relatives associées aux approximations de* φ_1. $\varpi(\tau) = 0.7$ *(continue),* $\varpi(\tau) = 0.7\exp(-\tau)$ *(pointillés). Projection. Grille 1.* $\tau_\star = 100$.

La figure 3.4 montre l'erreur relative associée à l'approximation construite par la méthode d'intégration produit lorsque $\tau_\star = 100$ et que l'albédo est constamment égal à 0.7. La figure 3.5 montre l'erreur relative associée à l'approximation construite par la méthode d'intégration produit lorsque $\tau_\star = 100$ et que l'albédo est variable : $\varpi(\tau) = 0.7\exp(-\tau)$. Ces approximations ont été construites à partir d'une grille uniforme de 501 points.

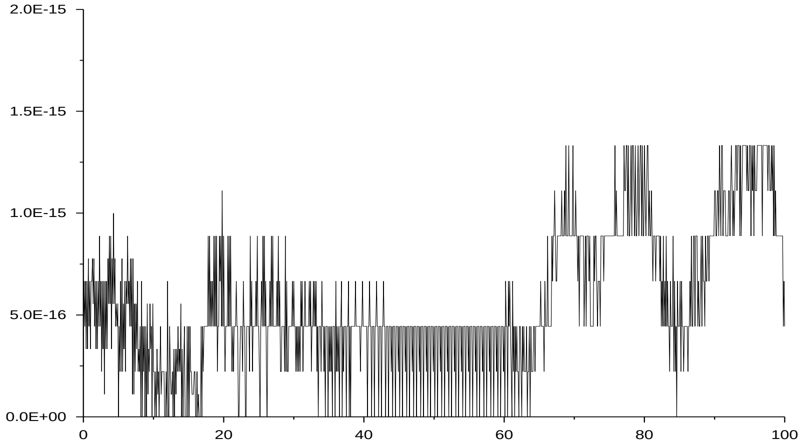


FIG. 3.4 – *Erreur relative associée à l'approximation de φ_1. Intégration produit. Grille uniforme ($n = 500$). $\tau_\star = 100$. $\varpi(\tau) = 0.7$.*

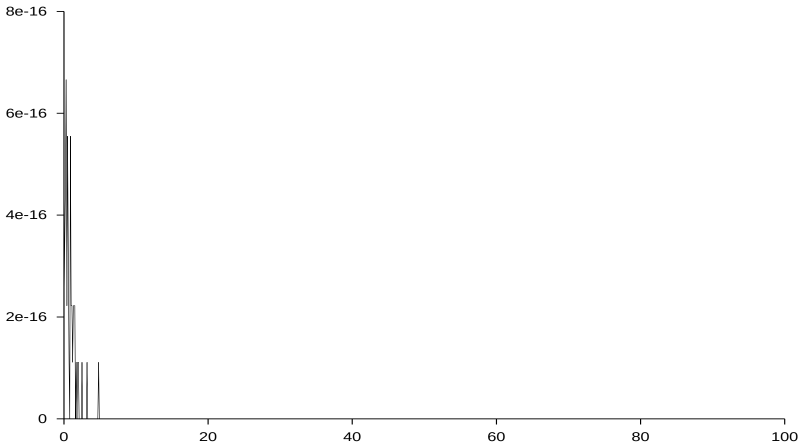


FIG. 3.5 – *Erreur relative associée à l'approximation de φ_1. Intégration produit. Grille uniforme ($n = 500$). $\tau_\star = 100$. $\varpi(\tau) = 0.7\exp(-\tau)$.*

La figure 3.5 montre un zoom de l'erreur relative associée à l'approximation construite par la méthode de projection lorsque $\tau_\star = 1000$ et que l'albédo est constamment égal à 0.999. Cette approximation a été construite à partir de la grille quasi-uniforme définie par

sous intervalles	nombre de points	pas
$[0, 100]$	501	0.2
$[10, 90]$	501	1.6
$[90, 100]$	501	0.2
taille de la matrice	1500	

Grille 2.

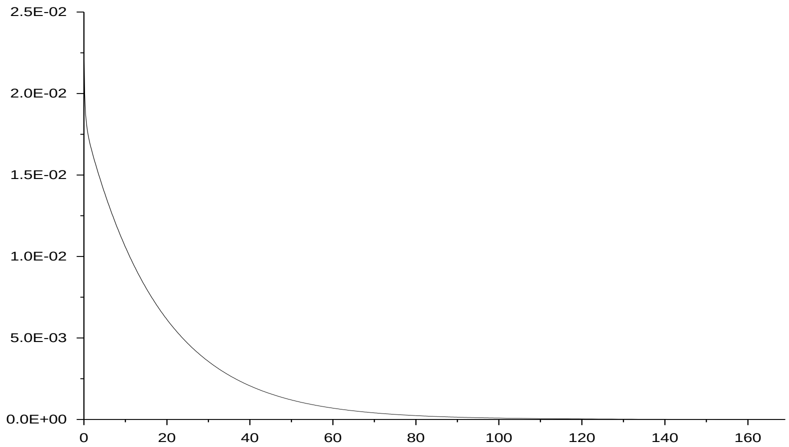


FIG. 3.6 – *Zoom de l'erreur relative associée à l'approximation de φ_1. Projection. Grille 2. $\tau_\star = 1000$. $\varpi(\tau) = 0.999$.*

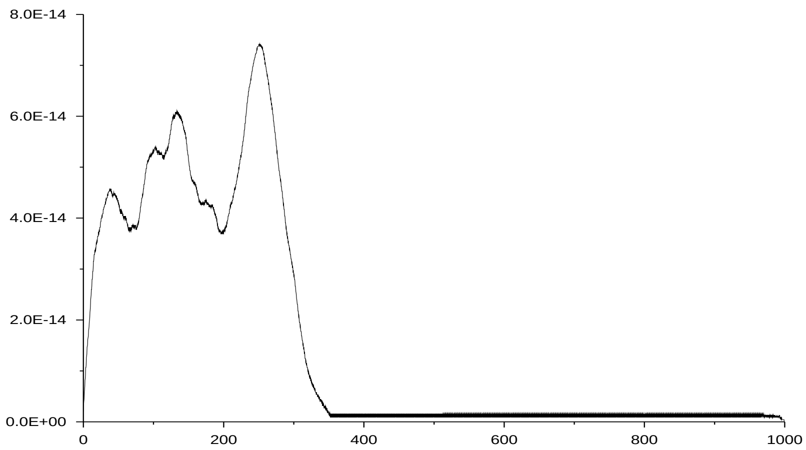


FIG. 3.7 – *Zoom de l'erreur relative associée à l'approximation de φ_1. Intégration produit. Grille uniforme ($n = 1000$). $\tau_\star = 1000$. $\varpi(\tau) = 0.999$.*

Les deux figures suivantes montrent les erreurs relatives associées aux approximations de $\varphi_2 : \tau \in [0 : 100] \mapsto \exp(5\tau)$ constuites à partir de la grille suivante:

sous intervalles	nombre de points	pas
$[0, 90]$	501	0.18
$[90, 100]$	501	0.02
taille de la matrice	1000	

Grille 3.

La figure 3.8 montre une comparaison des erreurs commises par les méthodes de projection et d'intégration produit lorsque l'albédo est constamment égal à $\varpi_\star = 0.7$. Le choix de la grille 3 a été motivé pour privilégier le voisinage de $\tau_\star = 100$.

La figure 3.9 montre un zoom de l'erreur relative associée à l'approximation par intégration produit lorsque l'albédo est variable: $\varpi(\tau) = 0.7 \exp(-\tau)$.

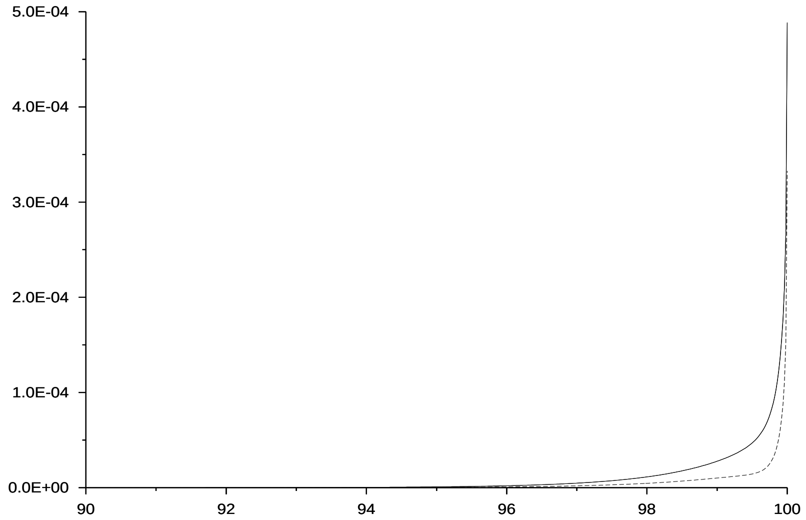


FIG. 3.8 – *Zoom des erreurs relatives associées aux approximations de φ_2. Projection (continue) et intégration produit (pointillés). Grille 3.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$. $\beta = 5$.

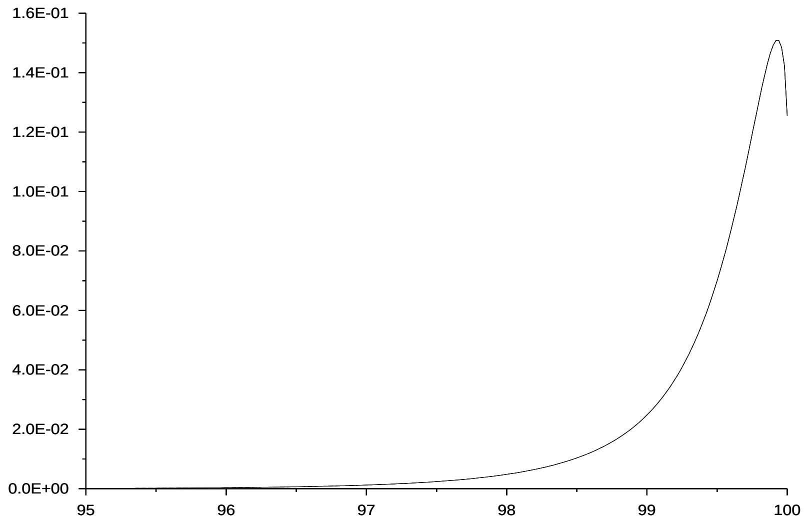


FIG. 3.9 – *Zoom des erreurs relatives associées aux approximations de φ_2. Projection (continue) et intégration produit (pointillés). Grille 3.* $\tau_\star = 100$. $\varpi(\tau) = 0.7\exp(-\tau)$. $\beta = 5$.

Les six figures suivantes sont consacrées à l'étude des approximations de la solution exacte φ_3 correspondant au terme libre f_3. C'est l'exemple que l'on retrouve dans la plupart des ouvrages portant sur la résolution numérique de l'équation de transfert. Nous avons voulu exhiber ces approximations au voisinage de zéro. C'est pourquoi nous avons utilisé la grille d'approximation suivante pour les construire :

sous intervalles	nombre de points	pas
$[0, 10^{-2}]$	101	10^{-4}
$[10^{-2}, 10^{-1}]$	91	10^{-3}
$[10^{-1}, 1]$	91	10^{-2}
$[1, 10]$	91	0.1
$[10, 90]$	501	0.16
$[90, 100]$	301	1/30
taille de la matrice	1170	

Grille 4.

La figure 3.10 montre les solutions approchées construites avec la méthode de projection et d'intégration produit. Les paramètres utilisés sont les suivants : $\tau_\star = 100$, $\varpi(\tau) = 0.7$. On remarquera que les deux courbes sont très proches l'une de l'autre.

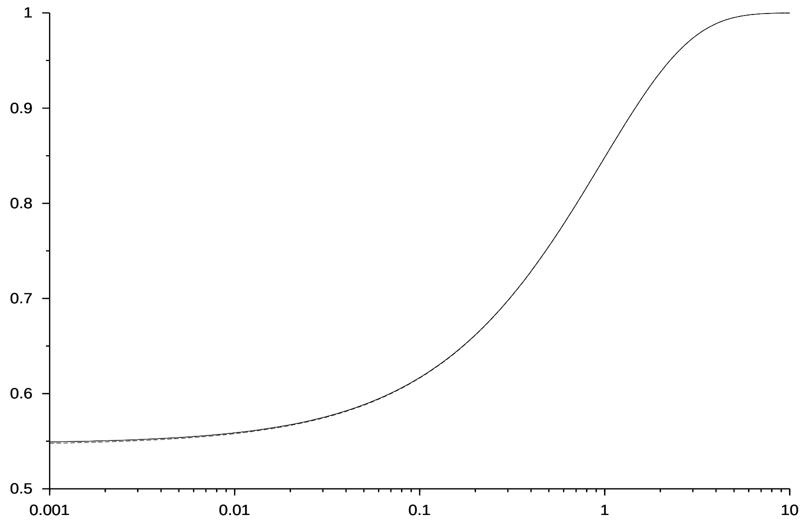


FIG. 3.10 – *Zoom des approximations de φ_3. Projection (continue) et intégration produit (pointillés). Grille 4. $\tau_\star = 100$. $\varpi(\tau) = 0.7$. Abscisses en échelle logarithmique.*

La figure 3.11 montre le résidu relatif associé à l'approximation de φ_3 construite à partir de la méthode de projection exhibée sur la figure précédente.

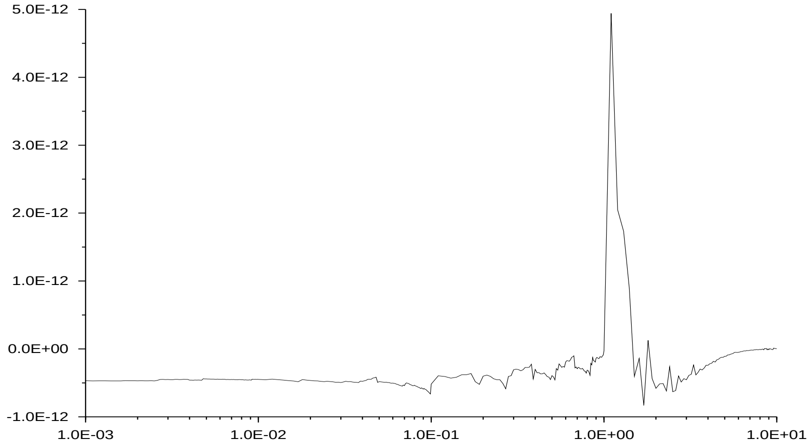


FIG. 3.11 – *Résidu relatif associé à l'approximation de la solution approchée de φ_3. Projection. Grille 4. $\tau_\star = 100$. $\varpi(\tau) = 0.7$. Abscisses en échelle logarithmique.*

La figure 3.12 montre l'approximation de φ_3 construite avec la méthode de projection et la méthode d'intégration produit lorsque l'albédo est variable : $\varpi(\tau) = 0.7\exp(-\tau)$. La figure 3.13 montre le résidu relatif associé.

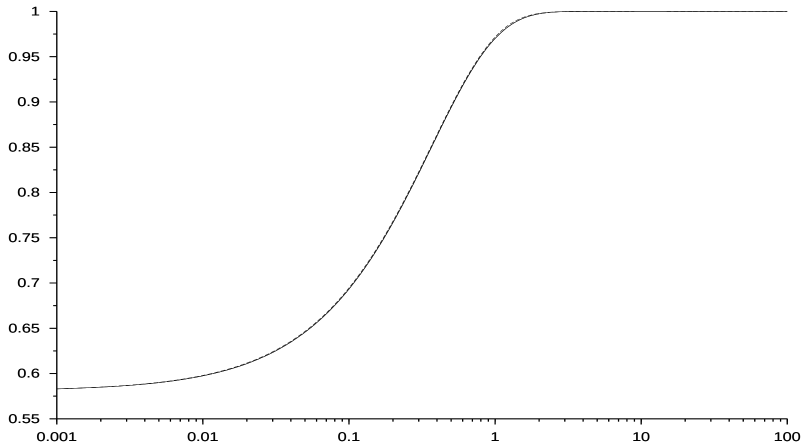


FIG. 3.12 – *Approximations φ_3. Projection (continue) et intégration produit (pointillés). Grille 4. $\tau_\star = 100$. $\varpi(\tau) = 0.7\exp(-\tau)$. Abscisses en échelle logarithmique.*

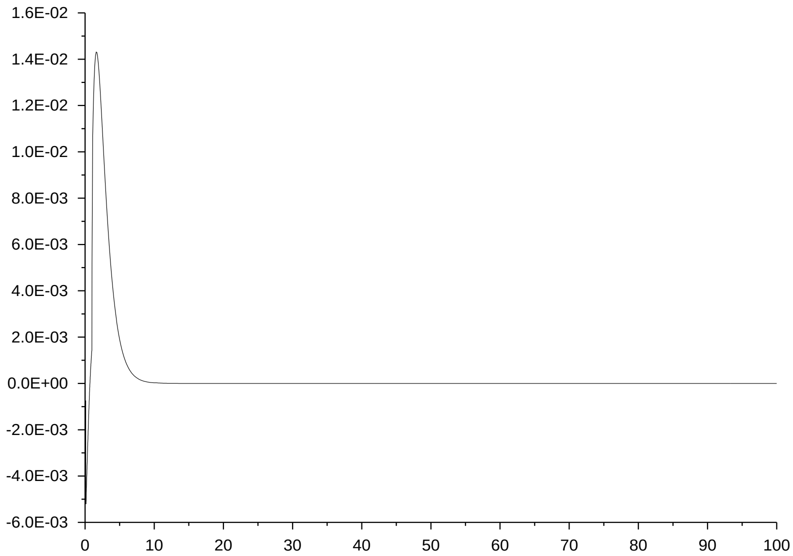


FIG. 3.13 – *Résidu relatif associé à l'approximation de la solution approchée de φ_3. Projection. Grille 4. $\tau_\star = 100$. $\varpi(\tau) = 0.7\exp(-\tau)$. Abscisses en échelle logarithmique.*

Les deux dernières figures montrent les approximations de φ_3 et un résidu relatif correspondant construits par la méthode de projection et d'intégration produit pour des valeur plus élevées de $\tau_\star$ et de $\varpi_\star$. Ces approximations ont étés construites à l'aide de la grille suivante :

sous intervalles	nombre de points	pas
$[0, 1]$	101	10^{-2}
$[1, 100]$	901	0.11
$[100, 900]$	201	4
$[900, 1000]$	1001	10^{-2}
taille de la matrice	2200	

Grille 5.

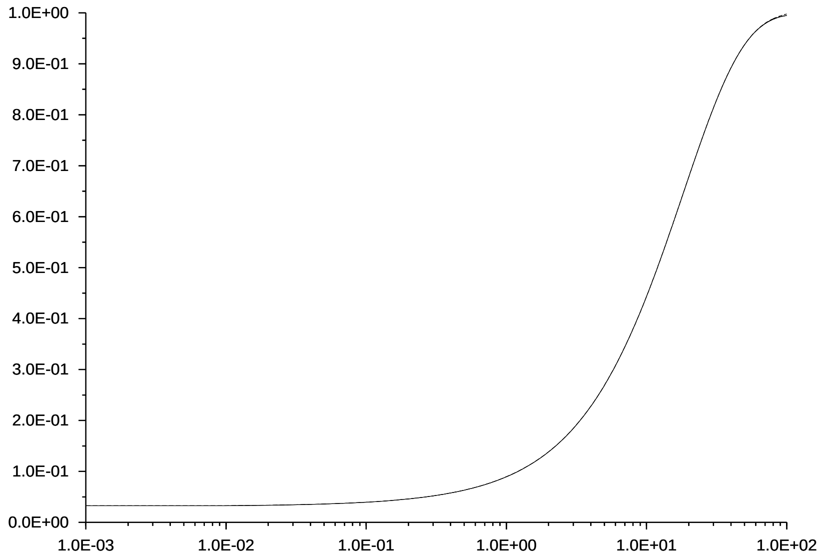


FIG. 3.14 – *Approximation de φ_3. Projection (continue) et intégration produit (pointillés). Grille 5. $\tau_\star = 1000$. $\varpi(\tau) = 0.999$. Abscisses en échelle logarithmique.*

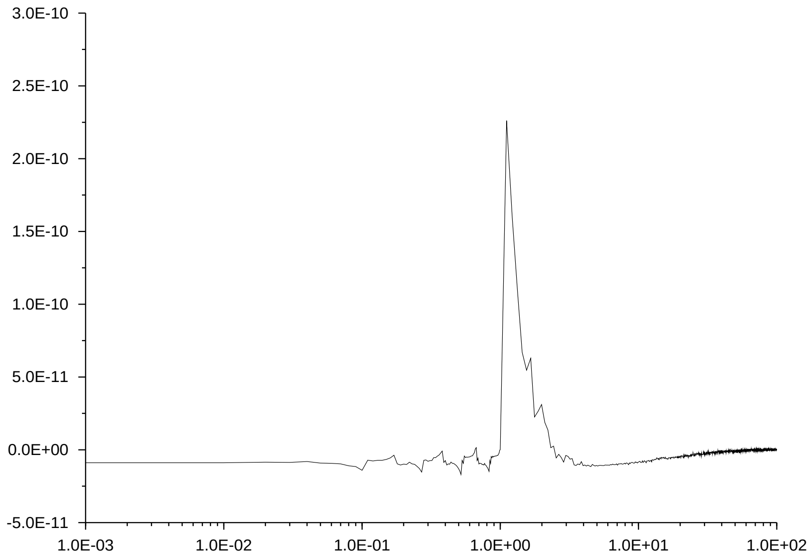


FIG. 3.15 – *Résidu relatif associé à l'approximation de la solution approchée de φ_3. Projection. Grille 5. $\tau_\star = 1000$. $\varpi(\tau) = 0.999$. Abscisses en échelle logarithmique.*

Toutes les approximations précédentes sont des approximations de fonctions continues sur $[0, \tau_\star]$. Les figures suivantes sont consacrées à des approximations de solutions appartenant à l'espace $L^1([0, \tau_\star])$. Ces figures montrent donc les moyennes locales des solutions approchées obtenues. Les sept premières figures traitent de l'approximation de la solution φ_4, dont le terme libre f_4 correspond à un saut de température dans une couche interne de l'atmosphère. Les deux premières figures 3.16 et 3.17 montrent les moyennes locales de la solution approchée obtenue par la méthode de projection décrite en 2.1.1.2 lorsque le saut de température se produit sur la couche $\theta = 50$; l'albédo est constamment égal à 0.7 et $\tau_\star = 100$. La grille d'approximation utilisée est définie de la manière suivante :

sous intervalles	nombre de points	pas
$[0, 10^{-2}]$	101	10^{-4}
$[10^{-2}, 10^{-1}]$	91	10^{-3}
$[10^{-1}, 1]$	91	10^{-2}
$[1, 15]$	101	0.05
$[15, 40]$	101	0.25
$[40, 60]$	270	2/27
$[60, 100]$	51	0.8
taille de la matrice	800	

Grille 6.

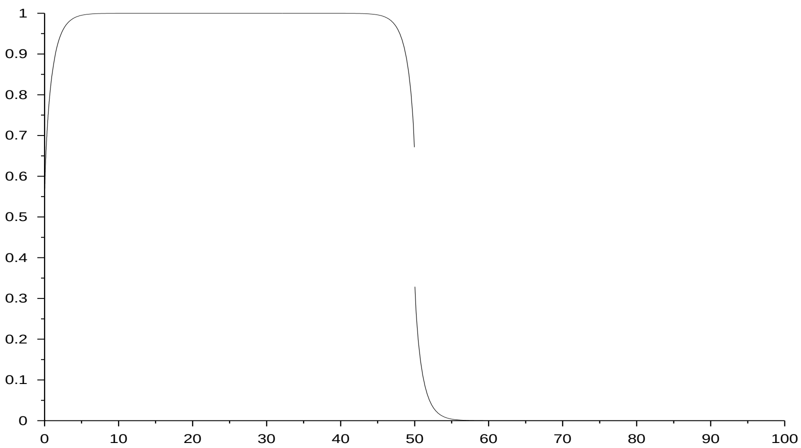


FIG. 3.16 – *Moyennes locales de l'approximation de φ_4. Projection. Grille 6.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$. $\theta = 50$.

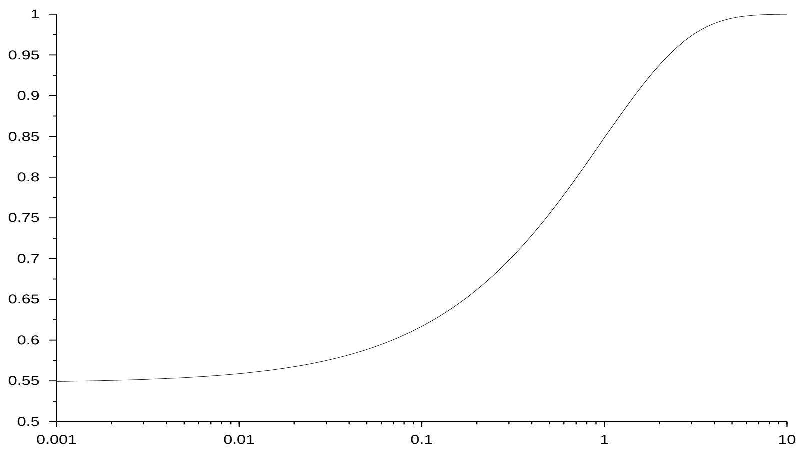


FIG. 3.17 – *Zoom des moyennes locales de l'approximation de* φ_4. *Projection. Grille 6.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$. $\theta = 50$. *Abscisses en échelle logarithmique.*

Les figures 3.18 et 3.19 montrent les moyennes locales de la solution approchée obtenue par la méthode de projection décrite en 2.1.1.2 lorsque le saut de température se produit sur la couche $\theta = 0.1$, c'est à dire proche de la surface ; l'albédo est constamment égal à 0.7 et $\tau_\star = 100$. La grille d'approximation utilisée est définie de la manière suivante :

sous intervalles	nombre de points	pas
$[0, 10^{-2}]$	101	10^{-4}
$[10^{-2}, 10^{-1}]$	91	10^{-3}
$[10^{-1}, 1]$	91	10^{-2}
$[1, 100]$	501	0.198
taille de la matrice	690	

Grille 7.

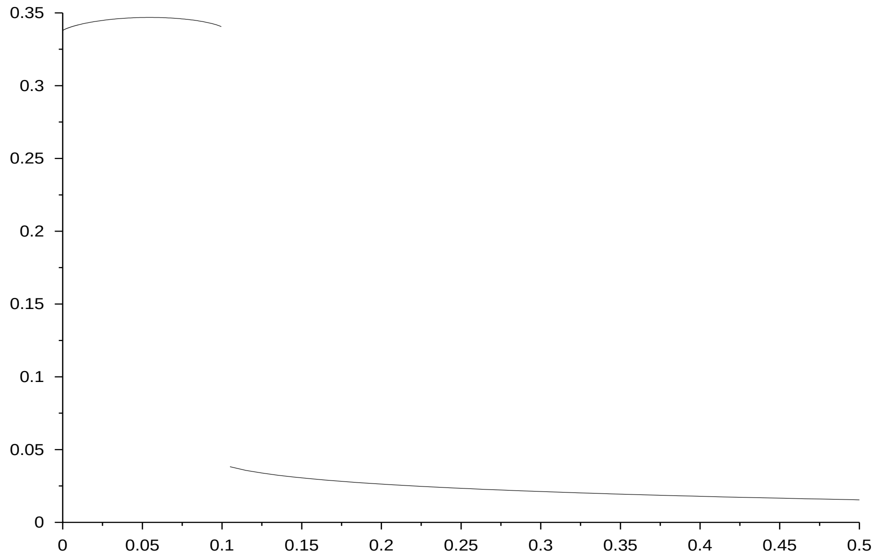


FIG. 3.18 – *Zoom des moyennes locales de l'approximation de* φ_4*. Projection. Grille 7.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$. $\theta = 0.1$.

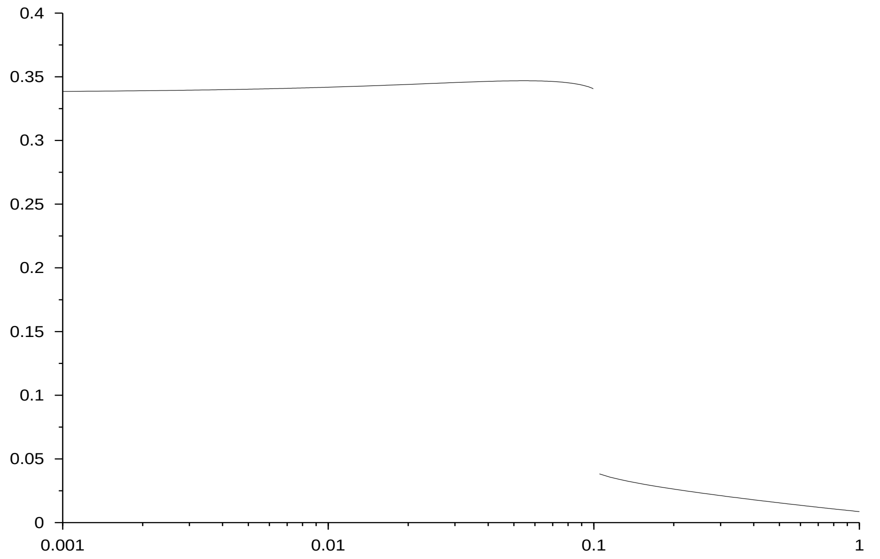


FIG. 3.19 – *Zoom des moyennes locales de l'approximation de* φ_4*. Projection. Grille 7.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$. $\theta = 0.1$. *Abscisses en échelle logarithmique.*

Les figures 3.20 et 3.21 montrent les moyennes locales de la solution approchée obtenue par la méthode d'intégration produit décrite en 2.1.2.2 lorsque le saut de température se produit sur la couche $\theta = 50$, c'est à dire au centre de l'atmosphère, et lorsque l'albédo est variable: $\varpi(\tau) = 0.7\exp(-\tau)$, et pour $\tau_\star = 100$. La grille d'approximation utilisée est la grille 5. La figure 3.22 montre le même type de solution approchée lorsque $\theta = 0.1$. La grille d'approximation est la grille 6.

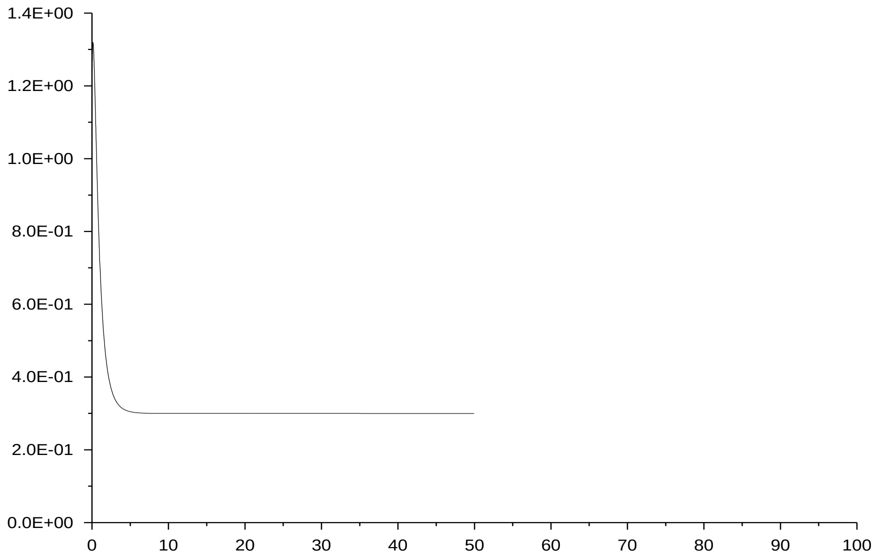


FIG. 3.20 – *Moyennes locales de l'approximation de φ_4. Intégration Produit. Grille 6.* $\tau_\star = 100$. $\varpi(\tau) = 0.7\exp(-\tau)$. $\theta = 50$.

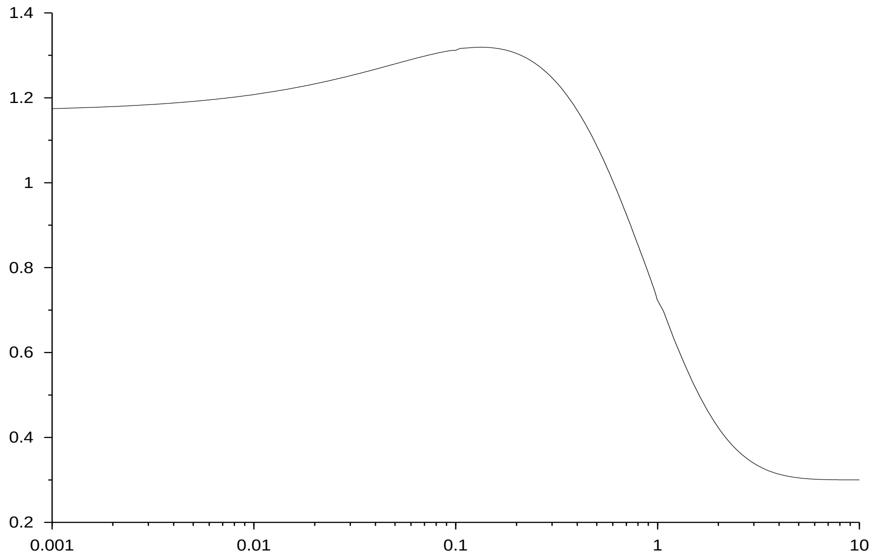


FIG. 3.21 – *Zoom des moyennes locales de l'approximation de φ_4. Intégration Produit. Grille 6. $\tau_\star = 100$. $\varpi(\tau) = 0.7\exp(-\tau)$. $\theta = 50$. Abscisses en échelle logarithmiques.*

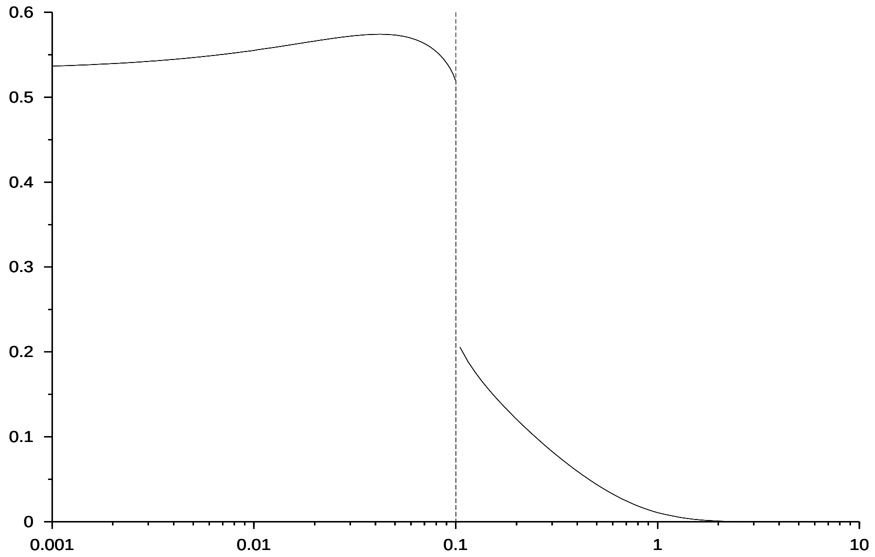


FIG. 3.22 – *Zoom des moyennes locales de l'approximation de φ_4. Intégration Produit. Grille 7. $\tau_\star = 100$. $\varpi(\tau) = 0.7\exp(-\tau)$. $\theta = 0.1$. Abscisses en échelle logarithmiques.*

La figure 3.23 montre les moyennes locales des approximations de φ_5 construites par la méthode d'intégration produit dans L^1. Une comparaison est faite entre un cas où l'albédo est constant et un cas où celui-ci est variable. Nous avons choisi $\tau_\star = 100$, la grille d'approximation utilisée étant définie de la manière suivante :

sous intervalles	nombre de points	pas
$[0, 40]$	201	0.2
$[40, 60]$	401	0.05
$[60, 100]$	201	0.2
taille de la matrice	800	

Grille 8.

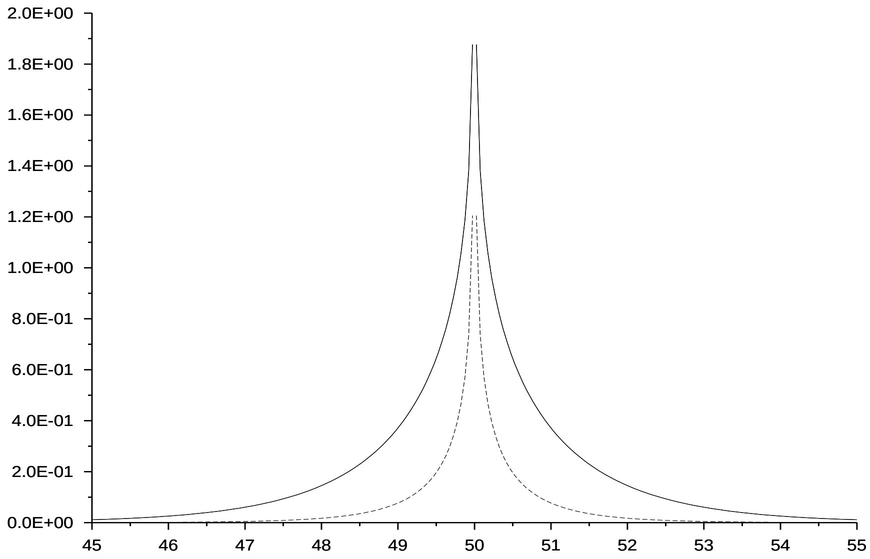


FIG. 3.23 – *Zoom des moyennes locales des approximations de* φ_5*.* $\varpi(\tau) = 0.7$ *(continue) et* $\varpi(\tau) = 0.7\exp(-\tau)$ *(pointillés). Intégration Produit. Grille 8.* $\tau_\star = 100$.

La figure 3.24 montre les moyennes locales de l' approximation de φ_5 construite par la méthode d'intégration produit dans L^1. Nous avons choisi $\tau_\star = 1000$ et ϖ constamment égal à 0.999. La grille d'approximation utilisée est définie de la manière suivante :

sous intervalles	nombre de points	pas
$[0, 400]$	401	1
$[400, 600]$	801	0.25
$[600, 1000]$	401	1
taille de la matrice	1600	

Grille 9.

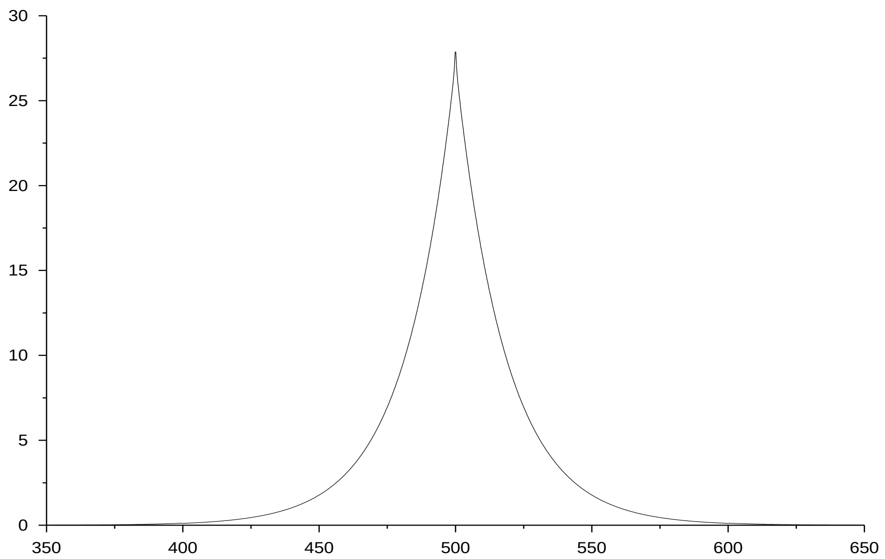


FIG. 3.24 – *Zoom des moyennes locales des approximations de φ_5. Intégration Produit. Grille 9. $\tau_\star = 1000$. $\varpi(\tau) = 0.999$.*

3.4.4 Raffinement itératif

Nous présentons dans cette section des figures montrant le comportement des itérés successifs construits par les trois schémas de raffinement itératif (A), (B) et (C), décrits en 2.2.1, ainsi que le comportement des résidus relatifs associés.

Nous rappelons que ces schémas convergent non pas vers la solution exacte du problème considéré mais vers une solution approchée correspondante à une discrétisation fine de celui-ci. Afin d'alléger les écritures, ces solutions approchées seront notées $\widetilde{\varphi}_i$, $1 \leq i \leq 5$, suivant le second membre considéré.

Nous définissons le dernier itéré d'un schéma comme étant le premier itéré à admettre un résidu relatif inférieur ou égal à 10^{-10}.

Pour des raison de «sécurité», les algorithmes correspondants sont mis en œuvre de façon à respecter un nombre maximum d'itérations.

3.4.4.1 Un exemple dans $C^0([0,\tau_\star])$

La figure 3.25 montre les itérés successifs construits par le schéma d'Atkinson (schéma (A)), basé sur la méthode de projection et approchant la solution $\widetilde{\varphi}_3$. Vingt-six itérations ont étés nécessaires pour obtenir un premier résidu relatif inférieur ou égal à 10^{-10}. Dans cette exemple, l'albédo est constamment égal à 0.7 et $\tau_\star = 100$. Les grilles d'approximation grossière et d'approximation fine sont uniformes, comportant respectivement 51 et 1001 points. Les tailles des matrices d'approximation grossière et fine sont donc respectivement 50 et 1000, soit d'un rapport de 20. La grille d'exposition est définie de la manière suivante, afin d'exhiber ces itérés au voisinage de l'origine :

sous intervalles	nombre de points	pas
$[0, 10]$	501	0.02
$[10, 90]$	501	0.16
$[90, 100]$	501	0.02
taille de la matrice	1500	

Grille 10.

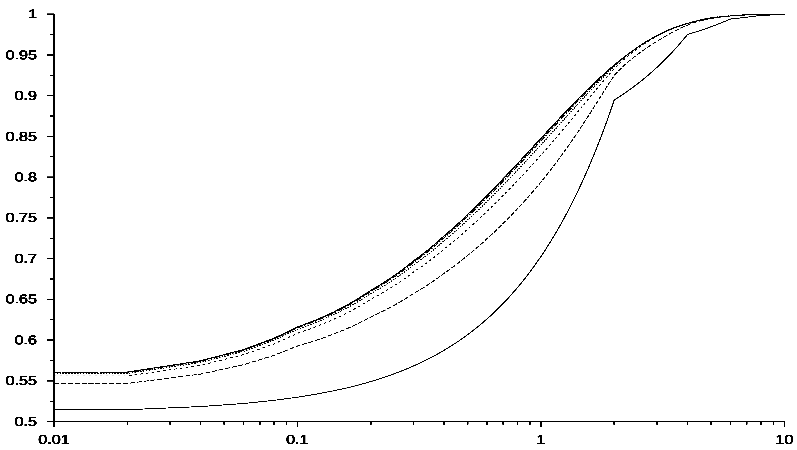
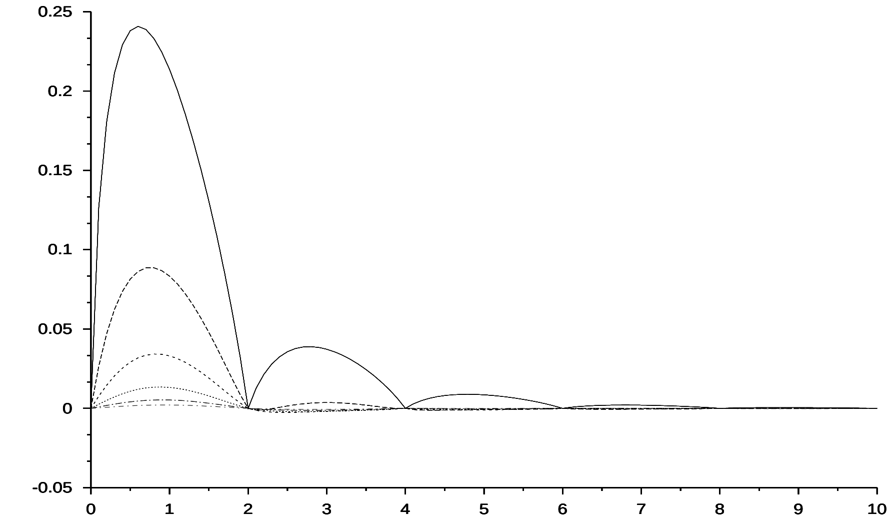


FIG. 3.25 – *Zoom des 26 premières approximations de $\widetilde{\varphi}_3$ du schéma d'Atkinson. Projection. Grille d'approximation grossière et fine uniformes ($n = 51$, $m = 1001$). Grille d'exposition 10. $\tau_\star = 100$. $\varpi(\tau) = 0.7$. Abscisses en échelles logarithmiques.*

Les cinq figures suivantes montrent la fonction résidu relatif correspondant à chaque itération. On remarque un phénomène de super-convergence : le résidu relatif est à son niveau le plus faible aux noeuds de la grille grossière. C'est pourquoi il est utile de choisir une grille d'exposition différente des grilles d'approximation pour pouvoir évaluer le résidu relatif entre leurs noeuds, afin de ne pas stopper le schéma trop tôt. (Voir [53] pour plus de détails sur ce phénomène).



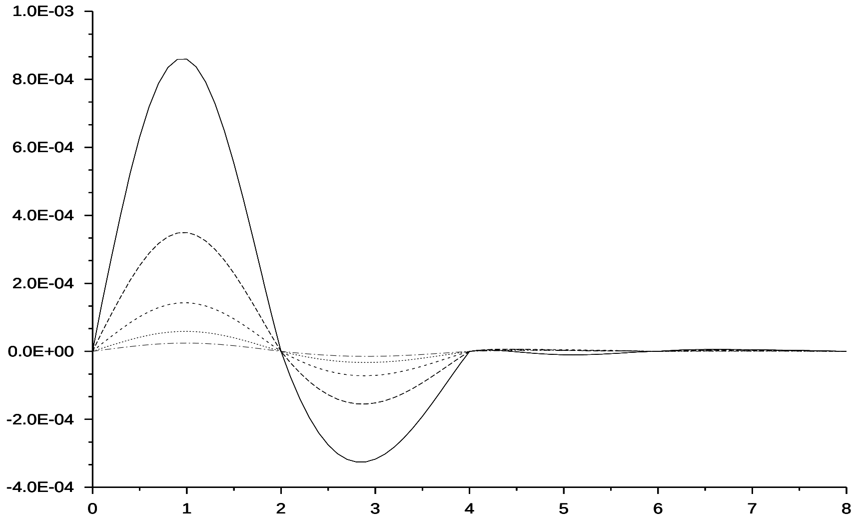


FIG. 3.26 – *Zoom des résidus relatifs associés aux onze premières approximations de* $\widetilde{\varphi}_3$ *du schéma d'Atkinson. Projection. Grille d'approximation grossière et fine uniformes* ($n = 51$, $m = 1001$)*. Grille d'exposition 10.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$.

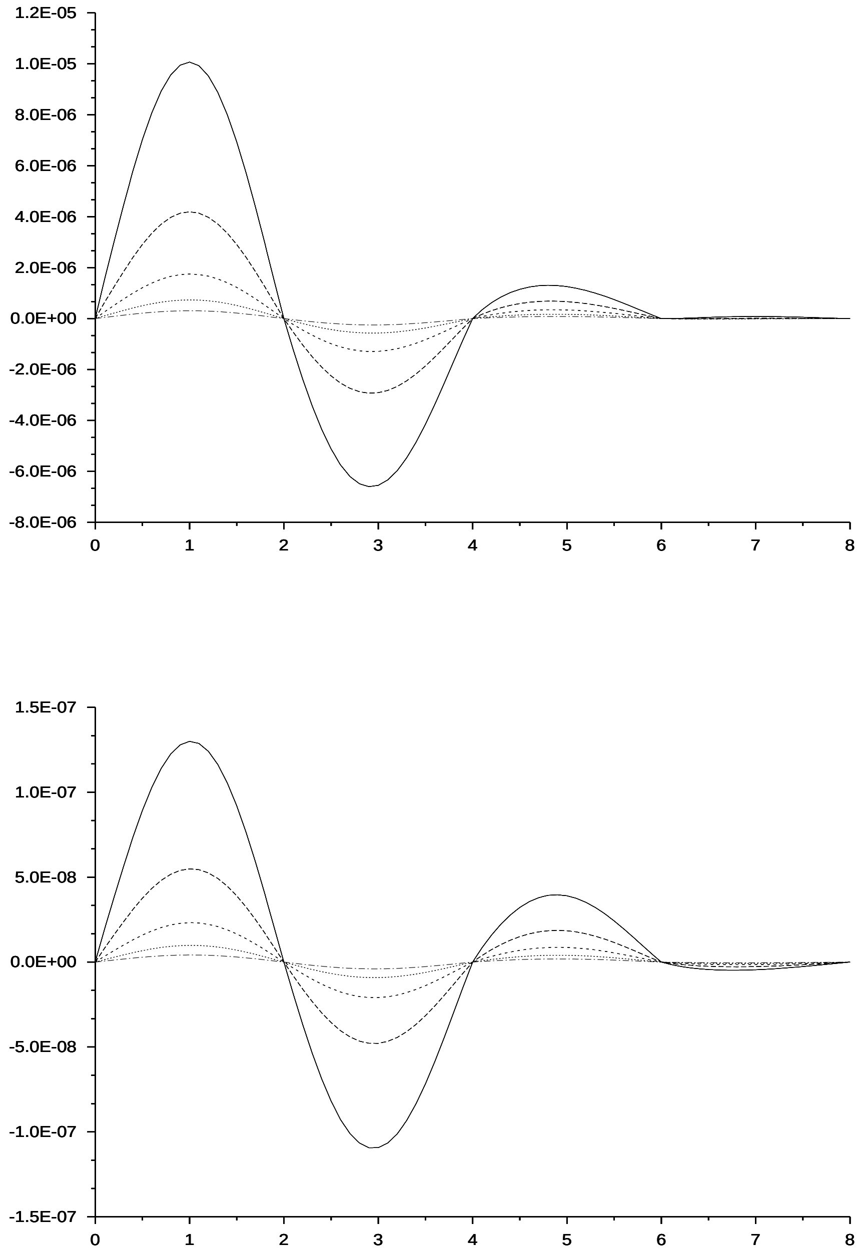


FIG. 3.27 – *Zoom des résidus relatifs des approximations* 11 *à* 20 *de* $\tilde{\varphi}_3$ *du schéma d'Atkinson. Projection. Grille d'approximation grossière et fine uniformes* ($n = 51$, $m = 1001$)*. Grille d'exposition 10.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$.

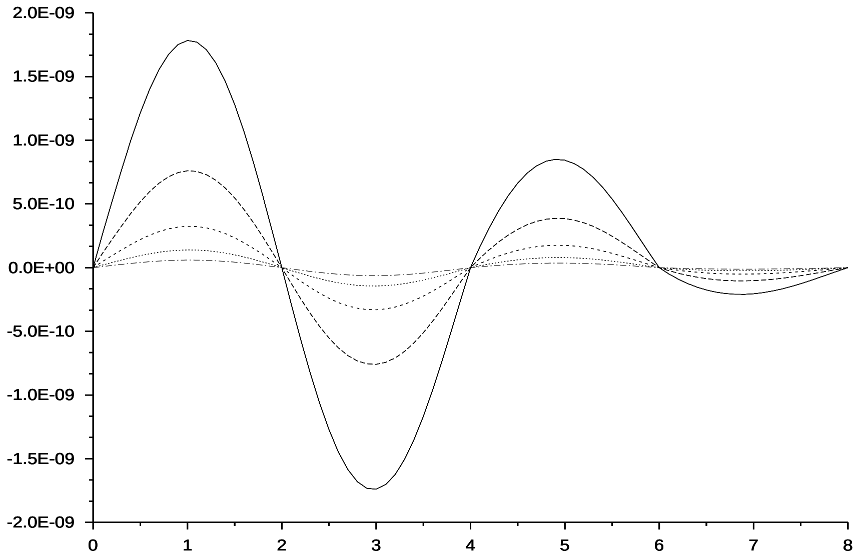


FIG. 3.28 – *Zoom des résidus relatifs associés aux cinq dernières approximations de $\widetilde{\varphi}_3$ du schéma d'Atkinson. Projection. Grille d'approximation grossière et fine uniformes ($n = 51$, $m = 1001$). Grille d'exposition 10. $\tau_\star = 100$. $\varpi(\tau) = 0.7$.*

La figure 3.29 montre les itérés successifs construits par le schéma de Brakhage (schéma (B)) appliqué à la méthode de projection, approchant $\widetilde{\varphi}_3$. Quatorze itérations ont étés nécessaires pour obtenir un premier résidu relatif inférieur ou égal à 10^{-10}. Les figures 3.30 et 3.31 montrent le premier et le dernier résidu relatif calculés par ce schéma. Les grilles d'approximation grossière et d'approximation fine sont uniformes, comportant respectivement 51 et 1001 points. Les tailles des matrices d'approximation ont donc un rapport de 20.

La figure 3.32 montrent les itérés successifs construits par le troisième schéma (schéma (C)) appliqué à la méthode d'intégration produit, approchant la solution $\widetilde{\varphi}_3$. Douze itérations ont étés nécessaires pour obtenir un premier résidu relatif inférieur ou égal à 10^{-10}. Les figures 3.33 et 3.34 montrent les fonctions résidus relatifs associées aux premiers et derniers itérés construits par ce schéma. Les grilles d'approximation grossière et d'approximation fine sont uniformes, comportant respectivement 60 et 900 points afin d'assurer l'emboîtement des grilles d'approximations intermédaires. Les tailles des matrices d'approximation grossière et fine ont donc un rapport de 15.

On remarque que les schémas (B) et (C) convergent plus rapidement que le schéma (A), les deuxièmes itérés correspondants étant beaucoup plus proches de l'itéré final. Les schémas (A) et (B) convergent par valeurs supérieures contrairement au schéma (C). Enfin le phénomène de super-convergence sur les résidus relatifs est encore observé chez ces deux derniers schémas. Dans ces exemples, l'albédo est constamment égal à 0.7 et $\tau_\star = 100$. La grille d'exposition est la grille 9.

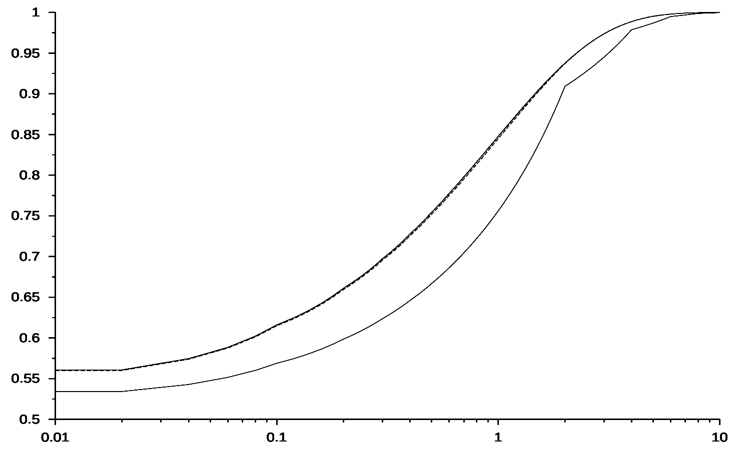


FIG. 3.29 – *Zoom des quatorze premières approximations de* $\widetilde{\varphi}_3$ *du schéma de Brakhage. Projection. Grille d'approximation grossière et fine uniformes (*$n = 51$*,* $m = 1001$*). Grille d'exposition 10.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$. *Abscisses en échelles logarithmiques.*

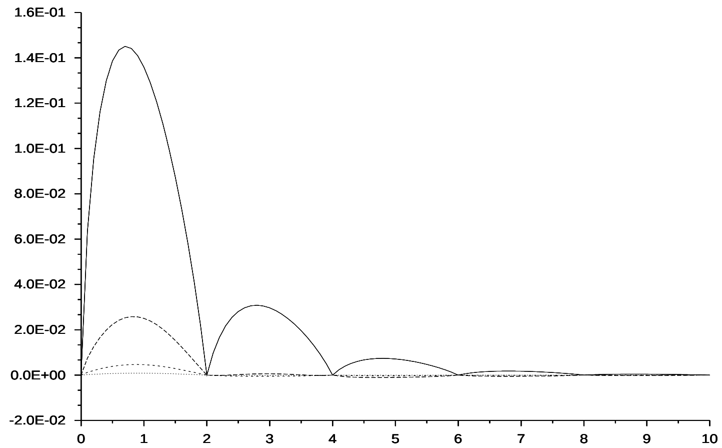


FIG. 3.30 – *Zoom des résidus relatifs associés aux quatre premières approximations de* $\widetilde{\varphi}_3$ *du schéma de Brakhage. Projection. Grille d'approximation grossière et fine uniformes (*$n = 51$*,* $m = 1001$*). Grille d'exposition 10.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$.

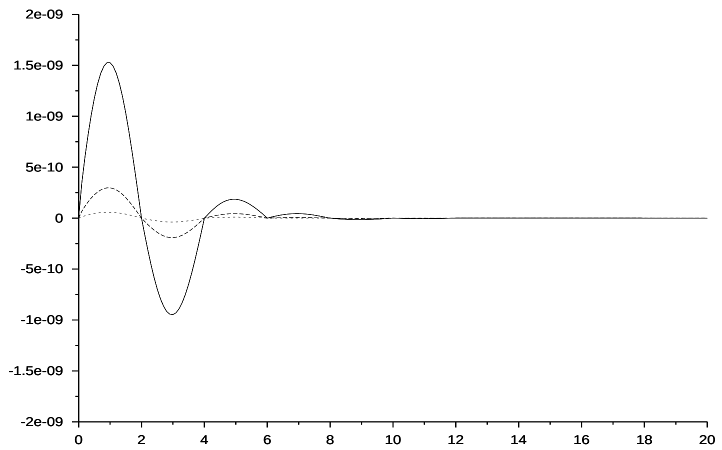


FIG. 3.31 – *Zoom des résidus relatifs associés aux trois dernières approximations de* $\widetilde{\varphi}_3$ *du schéma de Brakhage. Projection. Grille d'approximation grossière et fine uniformes* ($n = 51$, $m = 1001$)*. Grille d'exposition 10.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$.

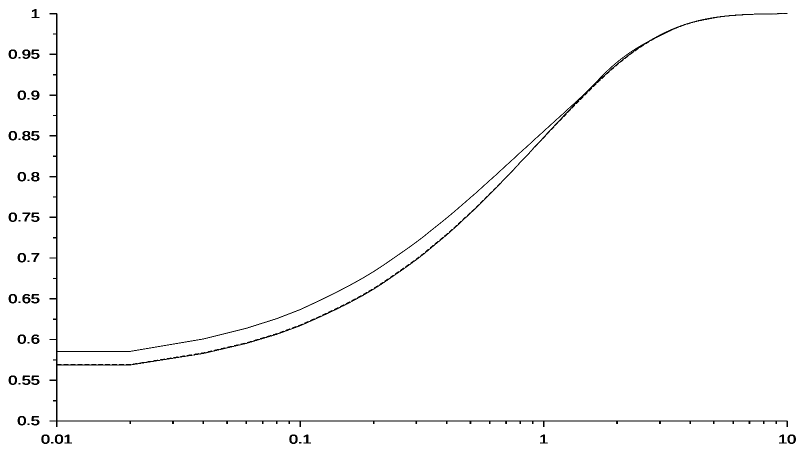


FIG. 3.32 – *Zoom des douze premières approximations de* $\widetilde{\varphi}_3$ *du schéma (C). Intégration produit. Grille d'approximation grossière et fine uniformes* ($n = 60$, $m = 900$)*. Grille d'exposition 10.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$. *Abscisses en échelles logarithmiques.*

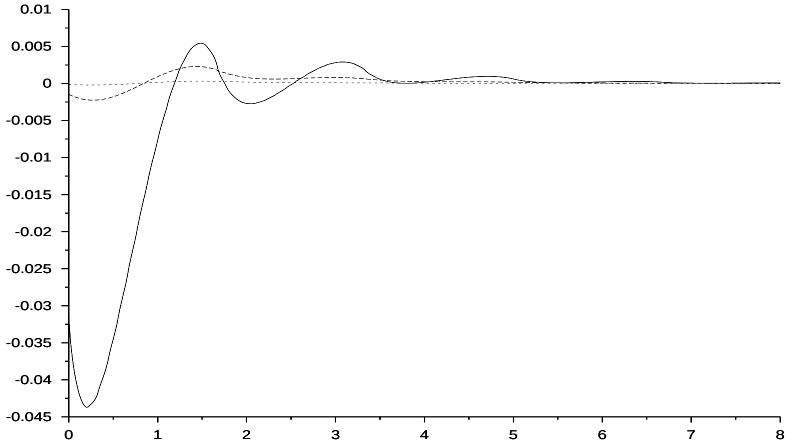


FIG. 3.33 – *Zoom des résidus relatifs associés aux trois premières approximations de* $\widetilde{\varphi}_3$ *du schéma (C). Intégration produit. Grille d'approximation grossière et fine uniformes* ($n = 60$, $m = 900$)*. Grille d'exposition 10.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$.

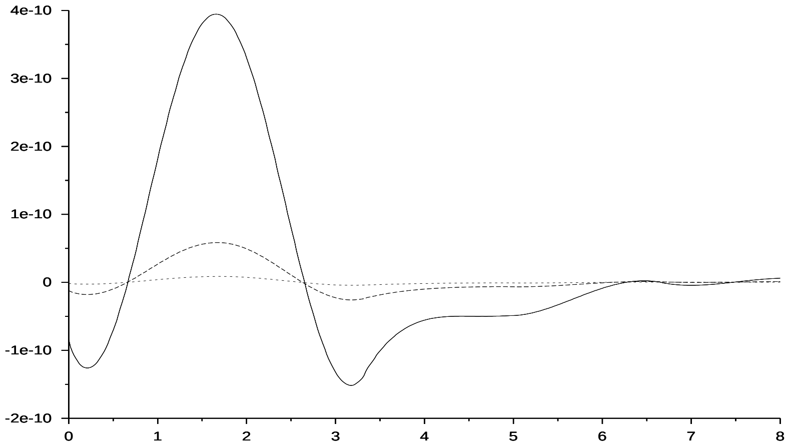


FIG. 3.34 – *Zoom des résidus relatifs associés aux trois dernières approximations de* $\widetilde{\varphi}_3$ *du schéma (C). Intégration produit. Grille d'approximation grossière et fine uniformes* ($n = 60$, $m = 900$)*. Grille d'exposition 10.* $\tau_\star = 100$. $\varpi(\tau) = 0.7$.

Le tableau suivant montre que la convergence des schémas de raffinement dépend très peu du type d'approximation de rang fini utilisée pour leur mise en œuvre. Les expériences correspondantes ont été réalisées avec des grilles uniformes : nous avons fixé la taille de la matrice d'approximation grossière à 200, celle de la matrice d'approximation fine à 1000, et les valeurs $\tau_\star = 100$ et $\varpi = 0.7$. La grille d'exposition est une grille uniforme de 2001 points. La solution visée est $\tilde{\varphi}_3$.

méthode	(A)	(B)	(C)
projection	11	6	6
intégration produit	10	5	5

TAB. 3.1 – *Nombre d'itérations*

Le tableau suivant montre la corrélation entre la vitesse de convergence des différents schémas et le rapport de taille des matrices d'approximations fine et grossière. Les expériences correspondantes ont été réalisées avec des grilles uniformes : nous avons fixé la taille de la matrice d'approximation fine à 1000 et nous notons n la taille de la matrice d'approximation grossière. Les paramètres sont fixés ainsi : $\tau_\star = 100$ et $\varpi = 0.7$. La grille d'exposition est une grille uniforme de 2001 points. Les itérées sont construits par la méthode de projection pour approcher la solution exacte du problème discrétisé avec la grille fine $\tilde{\varphi}_3$.

n	(A)	(B)	(C)
20	43	22	22
50	25	13	13
100	16	13	13
200	11	6	6

TAB. 3.2 – *Nombre d'itérations*

Enfin, le tableau suivant montre le comportement de chaque schéma par rapport à la valeur de $\varpi_\star$ lorsque l'albédo est constant. Les expériences correspondantes ont été réalisées avec des grilles uniformes : nous avons fixé la taille de la matrice d'approximation fine à 1000 et celle de la matrice d'approximation grossière à 200 (soit un rapport de 5) avec $\tau_\star = 100$. La grille d'exposition est une grille uniforme de 2001 points. Les itérées sont construits par la méthode de projection pour approcher la solution exacte du problème discrétisé avec la grille fine $\tilde{\varphi}_3$.

$\varpi_\star$	(A)	(B)	(C)
0.7	11	6	6
0.9	14	8	8
0.99	16	9	9
0.999	16	11	11
0.999999	17	12	11

TAB. 3.3 – *Nombre d'itérations*

3.4.4.2 Exemples dans $L^1([0, \tau_\star])$

La figure 3.35 montre un zoom des moyennes locales du premier et du dernier itérés construits par le schéma d'Atkinson basée sur la méthode de projection pour approcher la solution exacte du problème discrétisé avec la grille fine $\tilde{\varphi}_4$ lorsque l'albédo est variable. Les grilles d'approximation grossière et fine sont uniformes et comportent respectivement 201 et 1001 points sur $[0, 100]$. Six itérations ont été nécessaires pour obtenir un premier résidu relatif inférieur ou égal à 10^{-10} sur une grille d'exposition uniforme de 2001 points.

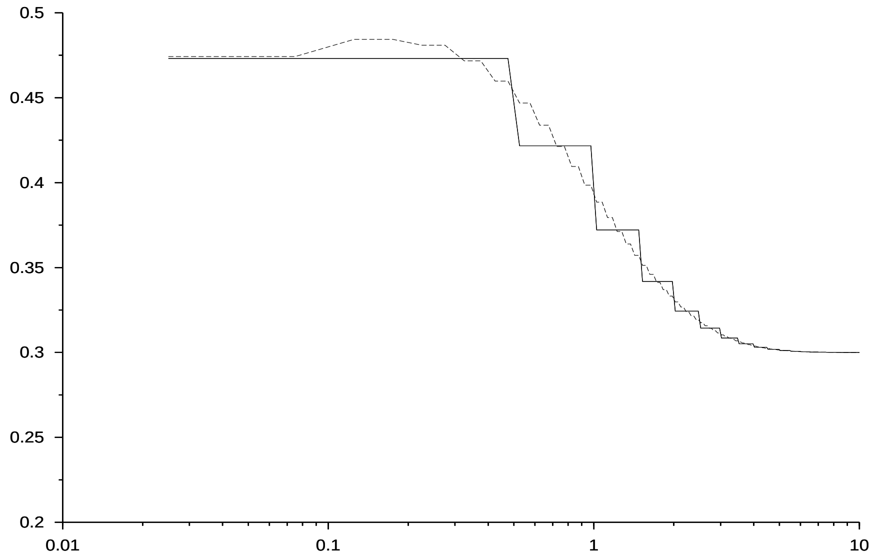


FIG. 3.35 – *Zoom de la première et de la* 6^e *(dernière) approximations de* $\tilde{\varphi}_4$ *du schéma d'Atkinson. Projection. Grilles d'approximations et d'exposition uniformes* ($n = 200$, $m = 1000$, $\nu = 2000$). $\tau_\star = 100$, $\varpi(\tau) = 0.7$. $\theta = 50$.

La figure 3.36 montre un zoom des moyennes locales des deux premiers et du dernier itérés construits par le schéma (C) basé sur la méthode d'intégration produit pour approcher la solution exacte du problème discrétisé avec la grille fine $\tilde{\varphi}_5$ lorsque l'albédo est constamment égal à 0.7. Les grilles d'approximation grossière et fine sont uniformes et comportent respectivement 31 et 1051 points sur $[0, 100]$. Dix-neuf itérations ont été nécéssaires pour obtenir un premier résidu relatif inférieur ou égal à 10^{-10} sur une grille d'exposition uniforme de 2001 points.

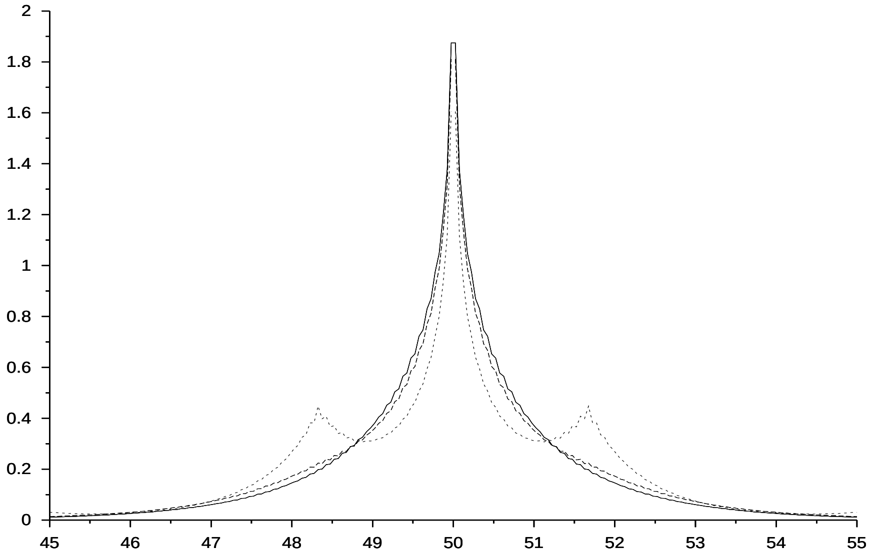


FIG. 3.36 – *Zoom de la première, deuxième et* 19^{e} *(dernière) approximations de* $\tilde{\varphi}_5$ *du schéma (C). Intégration produit. Grilles d'approximations et d'exposition uniformes* ($n = 30$, $m = 1050$, $\nu = 2000$). $\tau_\star = 100$, $\varpi(\tau) = 0.7$. $\theta = 50$.

Le tableau suivant montre la corrélation entre la vitesse de convergence des différents schémas et le rapport de taille des matrices d'approximations fine et grossière. Les expériences correspondantes ont été réalisées avec des grilles uniformes : nous avons fixé la taille de la matrice d'approximation fine à 1050 et nous notons par n la taille de la matrice d'approximation grossière. Les paramètres sont fixés ainsi : $\tau_\star = 100$ et $\varpi = 0.7$. La grille d'exposition est une grille quasi-uniforme raffinée au voisinage de la singularité. Les itérés sont construits par la méthode d'intégration produit pour approcher la solution exacte du problème discrétisé avec la grille fine $\widetilde{\varphi}_5$.

n	(A)	(B)	(C)
10	58	29	29
30	39	19	19
150	14	7	7
210	11	6	6
350	9	5	5

TAB. 3.4 – *Nombre d'itérations*

Le tableau suivant montre le comportement de chaque schéma par rapport à la valeur de $\varpi_\star$, lorsque l'albédo est constant, pour approcher la solution exacte du problème discrétisé avec la grille fine $\widetilde{\varphi}_5$. Ces schémas sont basés sur la méthode d'intégration produit. Les expériences correspondantes ont été réalisées avec des grilles uniformes : nous avons fixé la taille de la matrice d'approximation fine à 1050 et celle de la matrice d'approximation grossière à 210 (soit un rapport de 5) avec $\tau_\star = 100$. La grille d'exposition est une grille quasi-uniforme raffinée au voisinage de la singularité.

$\varpi_\star$	(A)	(B)	(C)
0.7	11	6	6
0.99	13	8	8
0.999	15	9	9
0.999999	(*)	(*)	(*)

(*) ne converge pas.

TAB. 3.5 – *Nombre d'itérations*

3.4.5 Cas d'une grande épaisseur optique

Si les méthodes de raffinement itératif permettent, pour une épaisseur optique fixée, d'obtenir une précision adéquate sur la solution en réduisant considérablement la taille du système linéaire à résoudre, elles ne permettent pas toujours la prise en compte de valeurs importantes de $\tau_\star$: elles sont limitées par le stockage de la matrice d'approximation fine, dont la taille augmente avec la précision. Dans [6], nous avons eu recours à la parallélisation des méthodes de raffinement lorsque peu d'hypothèses sont faites sur la régularité des données, c'est-à-dire sur le terme libre f et sur l'albédo ϖ. Il s'est d'ailleurs avéré que la parallélisation ne devient rentable que pour des valeurs très grandes de $\tau_\star$: en effet, ce type d'implémentation est réellement efficace que lorsque le temps de communication entre les différents processeurs mis en jeu dans la résolution est bien inférieur au temps impartit au calcul lui-même.

La prise en compte d'hypothèses plus fortes sur la régularité de f et ϖ permet, en outre, d'appliquer des méthodes asymptotiques que nous allons brièvement décrire dans la suite. Leur philosophie est de distinguer, dans la solution approchée, une partie régulière qui sera construite analytiquement. Lorsque ces méthodes sont couplées avec une méthode de décomposisiton asymptotique du domaine ([63]), les parties non régulières sont approchées par les solutions d'équations intégrales posées sur un intervalle beaucoup plus petit que l'intervalle initial $[0,\tau_\star]$. Nous décrivons dans la suite les résultats de Amosov et Panasenko développés dans [11].

3.4.5.1 Développement asymptotique

Soient a et b deux réels tels que $a<b$ et x dans $C^0([a,b])$. On définit la norme de x dans cet espace par

$$\|x\|_{[a,b]} := \sup_{\sigma\in[a,b]} |x(\sigma)|, \tag{3.71}$$

et la dérivée de x d'ordre j au point τ de $[0,\tau_\star]$, lorsqu'elle existe, par

$$x^{(j)}(\tau) := \frac{\mathrm{d}^j x}{\mathrm{d}\sigma^j}(\tau). \tag{3.72}$$

Afin d'appliquer une méthode de développement asymptotique, nous allons faire quelques hypothèses sur f et ϖ. Rappelons brièvement l'équation de transfert (3.27) :

$$\frac{\varpi(\tau)}{2}\int_0^{\tau_\star} E_1(|\tau-\sigma|)\varphi(\sigma)\,d\sigma - \varphi(\tau) = f(\tau), \quad \tau\in[0,\tau_\star].$$

On suppose que le terme libre f se décompose en trois fonctions de la manière suivante sur $[0,\tau_\star]$:

$$f := -\varpi f_\ell + (\varpi-1)f_0 - \varpi f_r, \tag{3.73}$$

et qu'il existe un entier $m \geq 0$ tel que :

$$\varpi \in C^{2m}([0,\tau_\star]),\ f_\ell, f_r \in C^0([0,\tau_\star]) \text{ et } f_0 \in C^{2m+2}([0,\tau_\star]). \tag{3.74}$$

On suppose de plus que f_ℓ et f_r sont deux fonctions à décroissance exponentielle, c'est-à-dire qu'il existe une constante positive M telle que pour tout $\tau \in [0,\tau_\star]$,

$$|f_\ell(\tau)| \leq M\exp(-\tau), \tag{3.75}$$
$$|f_r(\tau)| \leq M\exp(-(\tau_\star - \tau)). \tag{3.76}$$

Soit γ la fonction définie par

$$\gamma(\tau) := \frac{\varpi(\tau)}{1-\varpi(\tau)}, \quad \tau \in [0,\tau_\star]. \tag{3.77}$$

On suppose de plus que les fonctions f_0 et γ sont à variation lente, c'est-à-dire qu'il existe une constante positive N et $\delta \in]0,1[$ tels que

$$\|f_0^{(j)}\|_{[0,\tau_\star]} \leq M\delta^j, \quad 0 \leq j \leq 2m+2, \tag{3.78}$$
$$\|\gamma^{(j)}\|_{[0,\tau_\star]} \leq \frac{N}{1-\varpi_\star}\delta^j, \quad 0 \leq j \leq 2m+1. \tag{3.79}$$

On suppose enfin que

$$\varepsilon := \frac{N\delta^2}{1-\varpi_\star} < 1. \tag{3.80}$$

Soient deux réels a et b. On définit l'opérateur $\Lambda_{a,b}$ par

$$(\Lambda_{a,b}\varphi)(\tau) := \frac{1}{2}\int_a^b E_1(|\tau-\sigma|)\varphi(\sigma)\,d\sigma. \tag{3.81}$$

L'équation (3.27) s'écrit alors

$$\varpi(\tau)(\Lambda_{0,\tau_\star}\varphi)(\tau) - \varphi(\tau) = f(\tau), \quad \tau \in [0,\tau_\star]. \tag{3.82}$$

Le but de la méthode de développement asymptotique est de construire une solution approchée $\tilde{\varphi}^{[m]}$ de l'équation (3.27) ayant la forme suivante :

$$\tilde{\varphi}^{[m]} = \varphi_{\text{reg}}^{[m]} + \tilde{\varphi}_\ell^{[m]} + \tilde{\varphi}_r^{[m]},$$

où $\varphi_{\text{reg}}^{[m]}$ est une partie régulière de la solution approchée, $\tilde{\varphi}_\ell^{[m]}$ et $\tilde{\varphi}_r^{[m]}$ étant les solutions respectives de deux équations intégrales auxiliaires posées sur l'intervalle $[0,\tau_\star]$. L'idée est de chercher $\varphi_{\text{reg}}^{[m]}$ sous la forme suivante :

$$\varphi_{\text{reg}}^{[m]} = \sum_{k=0}^{m} \varphi_{k,\text{reg}}. \tag{3.83}$$

Soient σ et τ dans $[0,\tau_\star]$. Grâce à un développement de Taylor, nous pouvons écrire

$$\varphi_{k,\mathrm{reg}}(\sigma) = \sum_{j=0}^{2m-2k+1} \frac{1}{j!}\varphi_{k,\mathrm{reg}}^{(j)}(\tau)(\sigma-\tau)^j + r_k(\sigma,\tau), \tag{3.84}$$

où le reste r_k est tel que

$$|r_k(\sigma,\tau)| \leq \frac{\|\varphi_{k,\mathrm{reg}}^{(2m-2k+2)}\|_{[0,\tau_\star]}}{(2m-2k+2)!}|\sigma-\tau|^{2m-2k+2}. \tag{3.85}$$

Ainsi,

$$\varphi_{\mathrm{reg}}^{[m]}(\sigma) = \varphi_{T}^{[m]}(\sigma,\tau) + r_{T}^{[m]}(\sigma,\tau), \tag{3.86}$$

où

$$\varphi_{T}^{[m]}(\sigma,\tau) := \sum_{k=0}^{m}\sum_{j=0}^{2m-2k+1} \frac{1}{j!}\varphi_{k,\mathrm{reg}}^{(j)}(\tau)(\sigma-\tau)^j, \tag{3.87}$$

$$r_{T}^{[m]}(\sigma,\tau) := \sum_{k=0}^{m} r_k(\sigma,\tau). \tag{3.88}$$

D'après (3.50) et (3.52),

$$\begin{aligned}
\frac{\varpi(\tau)}{2}\int_0^{\tau_\star} E_1(|\tau-\sigma|)\varphi_{T}^{[m]}(\sigma,\tau)\,d\sigma &= \varpi(\tau)\sum_{k=0}^{m}\sum_{j=0}^{2m-2k+1}\frac{1}{2j!}\varphi_{k,\mathrm{reg}}^{(j)}(\tau)\int_0^{\tau_\star}E_1(|\sigma-\tau|)(\sigma-\tau)^j\,d\sigma \\
&= \varpi(\tau)\sum_{k=0}^{m}\sum_{j=0}^{2m-2k+1}\varphi_{k,\mathrm{reg}}^{(j)}(\tau)\frac{1}{2j!}\int_{-\infty}^{+\infty}E_1(|s|)s^j\,ds \\
&\quad -\varpi(\tau)\sum_{k=0}^{m}\sum_{j=0}^{2m-2k+1}(-1)^jE_{1,j}(\tau)\varphi_{k,\mathrm{reg}}^{(j)}(\tau) \\
&\quad -\varpi(\tau)\sum_{k=0}^{m}\sum_{j=0}^{2m-2k+1}E_{1,j}(\tau_\star-\tau)\varphi_{k,\mathrm{reg}}^{(j)}(\tau) \\
&= \varpi(\tau)\sum_{k=0}^{m}\sum_{\alpha=0}^{m-k}\frac{1}{2\alpha+1}\varphi_{k,\mathrm{reg}}^{(2\alpha)}(\tau) - R_\ell^{[m]}(\tau) - R_r^{[m]}(\tau),
\end{aligned}$$

où

$$R_\ell^{[m]}(\tau) := \varpi(\tau)\sum_{k=0}^{m}\sum_{j=0}^{2m-2k+1}(-1)^jE_{1,j}(\tau)\varphi_{k,\mathrm{reg}}^{(j)}(\tau), \tag{3.89}$$

$$R_r^{[m]}(\tau) := \varpi(\tau)\sum_{k=0}^{m}\sum_{j=0}^{2m-2k+1}E_{1,j}(\tau_\star-\tau)\varphi_{k,\mathrm{reg}}^{(j)}(\tau). \tag{3.90}$$

On remarque de plus que

$$\varpi(\tau)\sum_{k=0}^{m}\sum_{\alpha=0}^{m-k}\frac{1}{2\alpha+1}\varphi_{k,\mathrm{reg}}^{(2\alpha)}(\tau)=\varpi(\tau)\varphi_{\mathrm{reg}}^{[m]}(\tau)+\varpi(\tau)\sum_{k=1}^{m}\sum_{\alpha=1}^{k}\frac{1}{2\alpha+1}\varphi_{k-\alpha,\mathrm{reg}}^{(2\alpha)}(\tau).$$

Pour tout $\tau\in[0,\tau_\star]$, on définit

$$\Delta^{[m]}:=\frac{\varpi(\tau)}{2}\int_0^{\tau_\star}E_1(|\tau-\sigma|)r_{_T}^{[m]}(\sigma,\tau)\,d\sigma. \tag{3.91}$$

Soient $\tilde{\varphi}_\ell^{[m]}$ et $\tilde{\varphi}_r^{[m]}$ les solutions respectives des équations intégrales suivantes, posées sur $[0,\tau_\star]$:

$$\varpi(\Lambda_{0,\tau_\star}\tilde{\varphi}_\ell^{[m]})-\tilde{\varphi}_\ell^{[m]} = -\varpi f_\ell+R_\ell^{[m]}, \tag{3.92}$$

$$\varpi(\Lambda_{0,\tau_\star}\tilde{\varphi}_r^{[m]})-\tilde{\varphi}_r^{[m]} = -\varpi f_r+R_r^{[m]}. \tag{3.93}$$

Soit $\tilde{\varphi}^{[m]}$ la fonction définie par

$$\tilde{\varphi}^{[m]}:=\varphi_{\mathrm{reg}}^{[m]}+\tilde{\varphi}_\ell^{[m]}+\tilde{\varphi}_r^{[m]}. \tag{3.94}$$

Alors, pour tout τ dans $[0,\tau_\star]$,

$$\begin{aligned}\varpi(\tau)\Big(\Lambda_{0,\tau_\star}\tilde{\varphi}^{[m]}\Big)(\tau) &= \varpi(\tau)\varphi_{\mathrm{reg}}^{[m]}(\tau)+\varpi(\tau)\sum_{k=1}^{m}\sum_{\alpha=1}^{m-k}\frac{1}{2\alpha+1}\varphi_{k-\alpha,\mathrm{reg}}^{(2\alpha)}(\tau)\\ &\quad+\Delta^{[m]}(\tau)+\tilde{\varphi}_\ell^{[m]}(\tau)+\tilde{\varphi}_r^{[m]}(\tau)-\varpi(\tau)f_\ell(\tau)-\varpi(\tau)f_r(\tau).\end{aligned}$$

Si on définit récursivement $\varphi_{k,\mathrm{reg}}$ par

$$\begin{cases}\varphi_{0,\mathrm{reg}}:=f_0,\\ \varphi_{k,\mathrm{reg}}:=\gamma\sum\limits_{\alpha=1}^{k}\dfrac{1}{2\alpha+1}\varphi_{k-\alpha,\mathrm{reg}}^{(2\alpha)}, \quad 1\le k\le m,\end{cases} \tag{3.95}$$

alors

$$\begin{aligned}\varpi\sum_{k=1}^{m}\sum_{\alpha=1}^{k}\frac{1}{2\alpha+1}\varphi_{k-\alpha,\mathrm{reg}}^{(2\alpha)} &= (1-\varpi)\sum_{k=1}^{m}\gamma\sum_{\alpha=1}^{k}\frac{1}{2\alpha+1}\varphi_{k-\alpha,\mathrm{reg}}^{(2\alpha)}\\ &= (1-\varpi)\varphi_{\mathrm{reg}}^{[m]}-(1-\varpi)f_0.\end{aligned}$$

Or, si φ est la solution de l'équation (3.27), alors, pour tout $\tau\in[0,\tau_\star]$,

$$\varpi(\tau)\Big(\Lambda_{0,\tau_\star}(\varphi-\tilde{\varphi}^{[m]})\Big)(\tau)-(\varphi-\tilde{\varphi}^{[m]})(\tau)=-\Delta^{[m]}(\tau). \tag{3.96}$$

Donc

$$\|\varphi - \tilde{\varphi}^{[m]}\|_{[0,\tau_\star]} \le \frac{1}{1-\varpi_\star}\|\Delta^{[m]}\|_{[0,\tau_\star]}. \tag{3.97}$$

Le lemme suivant va nous permettre d'obtenir une majoration de $\|\Delta^{[m]}\|_{[0,\tau_\star]}$.

Lemme 12 *Soit $(\varphi_{k,\mathrm{reg}})_{k=0}^m$ la famille de fonctions définie par (3.95). Alors, pour tout $k \in [\![0, m]\!]$, et $j \in [\![1, 2m-2k+2]\!]$, il existe une constante $C_{m,k}$ telle que*

$$\|\varphi_{k,\mathrm{reg}}^{(j)}\|_{[0,\tau_\star]} \le C_{m,k}\frac{N^k M}{(1-\varpi_\star)^k}\delta^{2k+j}.$$

Preuve : ce lemme se démontre par induction sur k, en prenant en compte les hypothèses (3.78) et (3.79). Pour les détails voir [12]. ∎

Une première conséquence de ce lemme est qu'il existe une constante C_m telle que :

$$\frac{1}{1-\varpi_\star}\|\Delta^{[m]}\|_{[0,\tau_\star]} \le C_m M \varepsilon^{m+1}. \tag{3.98}$$

Nous avons enfin la majoration

$$\|\varphi - \tilde{\varphi}^{[m]}\|_{[0,\tau_\star]} \le C_m M \varepsilon^{m+1}. \tag{3.99}$$

À ce stade, la difficulté induite par la valeur importante de $\tau_\star$ n'est pas surmontée. En effet, $\tilde{\varphi}_\ell^m$ et $\tilde{\varphi}_r^m$ satisfont respectivement deux équations intégrales posées sur l'intervalle $[0, \tau_\star]$. Les hypothèses précédentes vont néanmoins réduire de manière conséquente l'intervalle d'intégration.

3.4.5.2 Méthode de décomposition du domaine

La méthode de décomposition asymptotique du domaine a pour but de réduire l'intervalle d'intégration dans les deux équations auxiliaires (3.92) et (3.93). Elle repose sur une estimation de la solution de l'équation de Milne

$$\frac{\varpi_\star}{2}\int_0^{+\infty} E_1(|\tau-\sigma|)\psi(\sigma,\varpi_\star)\,d\sigma - \psi(\tau,\varpi_\star) = -\varpi_\star\exp(-\alpha\tau), \quad \tau\in[0,+\infty[. \tag{3.100}$$

Dans [23], il est démontré que cette équation admet une unique solution Ψ de classe $C^0([0,+\infty[) \cap L^\infty(]0,+\infty[)$ et dans [12] il est démontré qu'il existe une constante c, indépendante de $\varpi_\star$, telle que

$$\Psi(\tau,\varpi_\star) \le c\exp(-k(\varpi_\star)\tau), \quad \tau\in[0,+\infty[, \tag{3.101}$$

où $k(\varpi_\star)$ est l'unique solution dans $]0,1[$ de l'équation, dite caractéristique,

$$1-\frac{\varpi_\star}{2}\frac{1}{k}\ln\frac{1+k}{1-k}=0. \tag{3.102}$$

Lemme 13 *Soit φ la solution de l'équation*

$$\frac{\varpi(\tau)}{2}(\varpi\Lambda_{a,b}\varphi)(\tau)-\varphi(\tau)=h(\tau),\quad \tau\in[a,b],$$

où $h\in L^\infty(]0,\infty[)\cap C^0([0,\infty[)$. Soit $\overline{\varphi}$ la solution de l'équation

$$\frac{\varpi(\tau)}{2}(\varpi_\star\Lambda_{0,\infty}\overline{\varphi})(\tau)-\overline{\varphi}(\tau)=\overline{h}(\tau),\quad \tau\in[0,+\infty[,$$

où $|h|\leq\overline{h}$. Alors $|\varphi|\leq\overline{\varphi}$.

Preuve : on écrit φ et $\overline{\varphi}$ sous la forme de leur représentation en serie de Neumann :

$$\varphi=-\sum_{k=0}^{\infty}(\varpi_\star\Lambda_{a,b})^k h,\quad \overline{\varphi}=-\sum_{k=0}^{\infty}(\varpi_\star\Lambda_{0,\infty})^k\overline{h},$$

et on remarque que, pour tout $\tau\in[a,b]$,

$$|\varphi|\leq\sum_{k=0}^{\infty}(\varpi_\star\Lambda_{a,b})^k|h|\leq\sum_{k=0}^{\infty}(\varpi_\star\Lambda_{0,\infty})^k\overline{h}=\overline{\varphi},$$

ce qui termine la preuve. ∎

Lemme 14 *Soient $d\in\,]0,\tau_\star[$ et φ la solution de l'équation*

$$\varpi(\tau)\Big(\Lambda_{0,\tau_\star}\varphi\Big)(\tau)-\varphi(\tau)=h(\tau),\quad \tau\in[0.\tau_\star],$$

où h est continue sur $[0,\tau_\star]$. Soient φ_ℓ^d la solution de l'équation

$$\varpi(\tau)\Big(\Lambda_{0,d}\varphi_\ell^d\Big)(\tau)-\varphi_\ell^d(\tau)=h(\tau),\quad \tau\in[0,d],$$

prolongée par zéro sur $]d,\tau_\star]$, et φ_r^d la solution de l'équation

$$\varpi(\tau)\Big(\Lambda_{\tau_\star-d,\tau_\star}\varphi_r^d\Big)(\tau)-\varphi_r^d(\tau)=h(\tau),\quad \tau\in[\tau_\star-d,\tau_\star],$$

prolongée par zéro sur $[0,\tau_\star-d[$. Alors

$$\|\varphi-\varphi_\ell^d\|_{[0,\tau_\star]} \leq c\|\varphi\|_{[d,\tau_\star]}, \tag{3.103}$$

$$\|\varphi-\varphi_r^d\|_{[0,\tau_\star]} \leq c\|\varphi\|_{[0,\tau_\star-d]}. \tag{3.104}$$

Preuve : soit $\tau \in [0,d]$. Alors

$$\varphi(\tau)-\varphi_\ell^d(\tau)=\frac{\varpi(\tau)}{2}\int_0^d E_1(|\tau-\sigma|)(\varphi-\varphi_\ell^d)(\sigma)\,d\sigma+\frac{\varpi(\tau)}{2}\int_d^{\tau_\star}E_1(|\tau-\sigma|)\varphi(\sigma)\,d\sigma.$$

Or,

$$\left|\frac{\varpi(\tau)}{2}\int_d^{\tau_\star}E_1(|\tau-\sigma|)\varphi(\sigma)\,d\sigma\right|\le\frac{\varpi_\star}{2}\|\varphi\|_{[d,\tau_\star]}E_2(\tau-d)\le\frac{\varpi_\star}{2}\|\varphi\|_{[d,\tau_\star]}\exp(-(\tau-d)),$$

et donc, d'après le Lemme 13,

$$|\varphi(\tau)-\varphi_\ell^d(\tau)|\le\frac{1}{2}\|\varphi\|_{d,\tau_\star}\Psi(\tau_\star-\tau,\varpi_\star),\quad \tau\in[0,d].$$

De plus, pour tout $\tau\in]d,\tau_\star]$,

$$|\varphi(\tau)-\varphi_\ell^d(\tau)|=|\varphi(\tau)|\le\|\varphi\|_{[d,\tau_\star]}.$$

On trouve l'inégalité (3.103) en appliquant (3.101). L'inégalité (3.104) est obtenue d'une manière similaire. ∎

Supposons que

$$d:=\frac{(m+1)}{k(\varpi_\star)}\ln(1/\varepsilon)\in]0,\frac{\tau_\star}{2}]. \tag{3.105}$$

Soit $\varphi_\ell^{[m]}$ la solution de l'équation intégrale

$$\varpi(\tau)(\Lambda_{0,d}\varphi_\ell^{[m]})(\tau)-\varphi_\ell^{[m]}(\tau)=-\varpi(\tau)f_\ell(\tau)+R_\ell^{[m]}(\tau),\quad \tau\in[0,d], \tag{3.106}$$

prolongée par zéro sur $]d,\tau_\star]$. Soit $\varphi_\ell^{[m]}$ la solution de l'équation intégrale

$$\varpi(\tau)(\Lambda_{\tau_\star-d,\tau_\star}\varphi_r^{[m]})(\tau)-\varphi_r^{[m]}(\tau)=-\varpi(\tau)f_r(\tau)+R_r^{[m]}(\tau),\quad \tau\in[\tau_\star-d,\tau_\star]. \tag{3.107}$$

prolongée par zéro sur $[0,\tau_\star-d[$. Soit $\varphi^{[m]}$ la fonction définie sur $[0,\tau_\star]$ par

$$\varphi^{[m]}:=\varphi_{\rm reg}^{[m]}+\varphi_\ell^{[m]}+\varphi_r^{[m]}. \tag{3.108}$$

Alors, d'après le Lemme 14,

$$\|\varphi-\varphi^{[m]}\|_{[0,\tau_\star]}\le\frac{1}{1-\varpi_\star}\|\Delta^{[m]}\|_{[0,\tau_\star]}+c\|\tilde\varphi_\ell^{[m]}\|_{[d,\tau_\star]}+c\|\tilde\varphi_r^{[m]}\|_{[\tau_\star-d,\tau_\star]}. \tag{3.109}$$

Lemme 15 *Soient* $\tilde{\varphi}_\ell^{[m]}$ *et* $\tilde{\varphi}_r^{[m]}$ *les solutions respectives de* (3.93) *et* (3.92)*, et* d *défini par* (3.105)*. Alors*

$$\|\tilde{\varphi}_\ell^{[m]}\|_{[d,\tau_\star]} + \|\tilde{\varphi}_r^{[m]}\|_{[0,\tau_\star - d]} \leq C_m M \varepsilon^{m+1}.$$

Preuve : on montre que le terme libre de l'équation (3.93) est majoré de la manière suivante :

$$|\varpi(\tau) f_\ell^{[m]}(\tau) + R_\ell^{[m]}(\tau)| \leq C_m M(\varpi_\star \exp(-\tau) + \delta^{2m+2}),$$

grâce, notamment, à l'estimation du Lemme 12. On trouve une estimation similaire pour le second membre de l'équation (3.92). L'inégalité est obtenue en appliquant le Lemme 13 et en utilisant le fait que $\dfrac{\delta^{2m+2}}{1-\varpi_\star} \leq \varepsilon^{m+1}$. Pour les détails voir [12]. ∎

Nous avons enfin le théorème suivant qui justifie la méthode de décomposition asymptotique du domaine.

Théorème 28 *Soient* φ *la solution de l'équation* (3.82) *et* $\varphi^{[m]}$ *définie par* (3.108)*. Alors, sous les hypothèses* (3.74) - (3.80)*,*

$$\|\varphi - \varphi^{[m]}\|_{[0,\tau_\star]} \leq C_m M \left(\frac{N\delta^2}{1-\varpi_\star}\right)^{m+1}. \tag{3.110}$$

Preuve : ce lemme est une conséquence directe des Lemmes 12, 13, 14 et 15. ∎

Remarque 13 *Les solutions* $\varphi_\ell^{[m]}$ *et* $\varphi_r^{[m]}$ *des «petites» équations* (3.106) *et* (3.107) *sont arbitrairement annulées sur* $[d, \tau_\star]$ *et* $[0, \tau_\star - d]$ *respectivement, ce qui provoque une discontinuité de saut en* d *et* $\tau_\star - d$ *dans la solution globale, dont l'amplitude ne dépasse pas* $\varphi_\ell^{[m]}(d)$ *et* $\varphi_r^{[m]}(\tau_\star - d)$ *respectivement.*

3.4.5.3 Méthodes hybrides

La mise en œuvre de la méthode de développement asymptotique couplée avec celle de la décomposition asymptotique du domaine nécessite la résolution des équations intégrales (3.106) et (3.107) qui sont posées sur les couches limites de longueur d. Généralement, nous ne pouvons pas trouver de solutions exactes à ces équations, et nous nous contentons de construire des solutions approchées à l'aide des méthodes d'approximation de rang fini décrites dans le chapitre précédent. Nous donnons simplement quelques pistes de réflexion sur l'analyse théorique de l'erreur globale associée à la solution numérique «hybride» obtenue.

Soient $\varphi_{\ell,n}^{[m]}$ et $\varphi_{r,n}^{[m]}$ les solutions approchées de (3.106) et (3.107) construites à l'aide d'une méthode d'approximation de rang fini (projection ou intégration produit dans $C^0([0,d])$ et $C^0([\tau_\star - d, \tau_\star])$ respectivement). La solution approchée «globale» obtenue est donc

$$\varphi_n^{[m]} := \varphi_{\text{reg}}^{[m]} + \varphi_{l,n}^{[m]} + \varphi_{r,n}^{[m]}. \tag{3.111}$$

On obtient alors l'estimation suivante pour l'erreur «globale» :

$$\|\varphi - \varphi_n^{[m]}\|_{[0,\tau_\star]} \leq C_m \varepsilon^{m+1} + \|\varphi_\ell^{[m]} - \varphi_{\ell,n}^{[m]}\|_{[0,d]} + \|\varphi_r^{[m]} - \varphi_{r,n}^{[m]}\|_{[\tau_\star - d, \tau_\star]}. \tag{3.112}$$

Ainsi, si nous voulons conserver l'ordre d'erreur ε^{m+1} acquise sur l'approximation (3.108), il est nécessaire de choisir les grilles d'approximation utilisées pour construire la solutions hybride de telle sorte que les erreurs absolues $\|\varphi_\ell^{[m]} - \varphi_{\ell,n}^{[m]}\|_{[0,d]}$ et $\|\varphi_r^{[m]} - \varphi_{r,n}^{[m]}\|_{[\tau_\star - d, \tau_\star]}$ soient du même ordre.

3.4.5.4 Exemple numérique

Nous allons étudier ici l'approximation hybride $\varphi_n^{[0]}$ de φ_3 construite à partir d'une approximation de rang fini par projection lorsque l'albédo est constamment égal à $\varpi_\star$. Le terme libre f_3 est défini dans ce cas par

$$f_3(\tau) = \varpi_\star - 1.$$

Ainsi, f_ℓ, f_r et f_0 sont définies par

$$f_r(\tau), \quad f_0(\tau) = 1 \quad f_r(\tau) = 0, \quad \tau \in [0, \tau_\star].$$

Les hypothèses de régularité (3.74), (3.75) (3.76) sont donc satisfaites pour tout $m \geq 0$ et $M = 1$. f_0 et ϖ étant constants, les hypothèses (3.78) et (3.79) sont satisfaites pour $N = 1$ et tout $\delta > 0$. C'est pourquoi nous choisirons ce dernier paramètre en fonction de la précision désirée ε :

$$\delta := \sqrt{(1 - \varpi_\star)\varepsilon}.$$

Lorsque $m = 0$, $\varphi_\ell^{[0]}$ et $\varphi_r^{[0]}$ vérifient les équations intégrales

$$\varpi(\tau)(\Lambda_{0,d}\varphi_\ell^{[0]})(\tau) - \varphi_\ell^{[0]}(\tau) = \frac{\varpi_\star}{2} E_2(\tau), \quad \tau \in [0, d], \tag{3.113}$$

$$\varpi(\tau)(\Lambda_{\tau_\star - d, \tau_\star}\varphi_r^{[0]})(\tau) - \varphi_r^{[0]}(\tau) = \frac{\varpi_\star}{2} E_2(\tau_\star - \tau), \quad \tau \in [\tau_\star - d, \tau_\star]. \tag{3.114}$$

Le tableau suivant donne des valeurs approchées de la couche limite d en fonction de $\varpi_\star$ pour $\varepsilon = 0.1$ et $m = 0$. Lorsque $\tau_\star$ est très grand, on remarque que la longueur des intervalles sur lesquels sont posées les équations précédentes peut effectivement être très réduite par rapport à l'intervalle global $[0, \tau_\star]$.

$\varpi_\star$	d
0.7	3
0.9	5
0.99	14
0.999	43
0.999999	1330

TAB. 3.6 – *Couche limite théorique*

La figure 3.37 montre un zoom des solutions approchées hybrides obtenues lorsque différentes valeurs de ε sont choisies. Les solutions des petites équations intégrales sont approchées par la méthode de projection, les grilles utilisées étant du type de la grille 4 (raffinées au voisinage de zéro). Nous avons choisis $\tau_\star = 100$ et $\varpi_\star = 0.7$ afin de limiter la source d'erreur provenant de l'approximation des couches limites. Ce zoom met clairement en évidence la discontinuité de saut que présente la solution hybride obtenue. On remarque de plus que l'amplitude du saut ne dépasse jamais l'ordre de la précision ε désirée. Les figures 3.38, 3.39 et 3.40 montrent les résidus sur l'équation exacte associés à trois approximations hybrides dont les solutions sur les couches limites sont approchées par une méthode de projection avec $\varepsilon = 10^{-1}$, $\varepsilon = 10^{-2}$ et $\varepsilon = 10^{-3}$ respectivement. On remarque que les sauts observés dans les approximations correspondantes se traduisent par des sauts des résidus d'un ordre inférieur à la précision désirée.

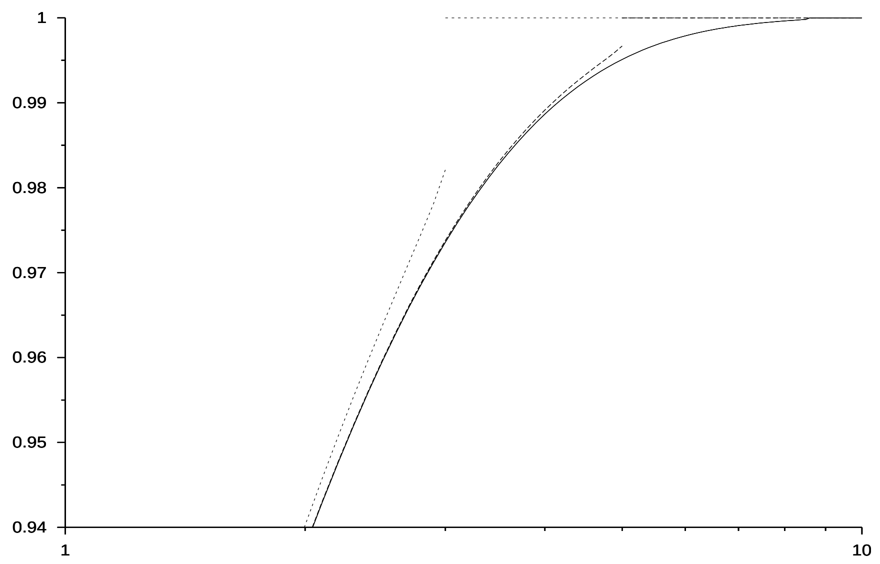


FIG. 3.37 – *Zoom des approximations hybrides construites par les méthodes asymptotiques basées sur une méthode de projection pour* $\varepsilon = 10^{-1}$ *(pointillés) ,* $\varepsilon = 10^{-2}$ *(tirets) et* $\varepsilon = 10^{-3}$ *(continue).* $\tau_\star = 100$*,* $\varpi(\tau) = 0.7$*.*

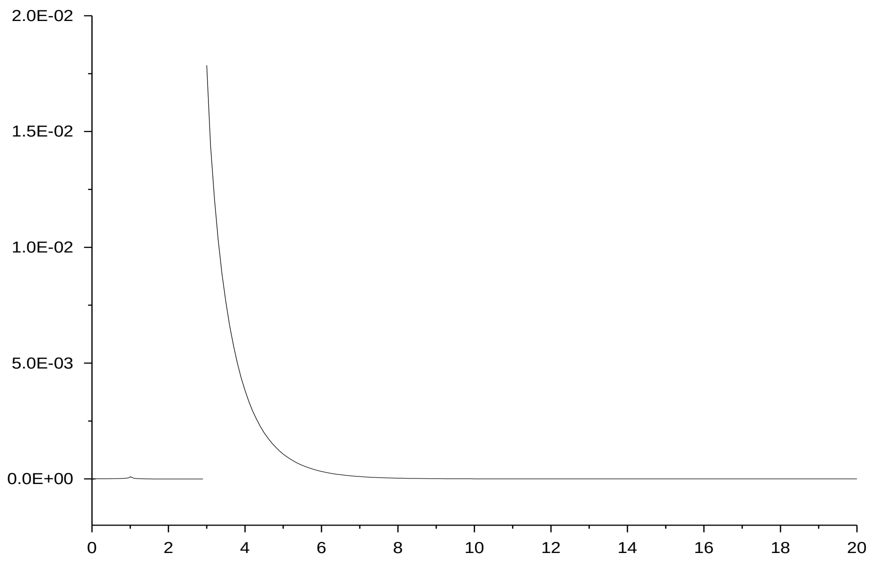


FIG. 3.38 – *Zoom du résidu sur l'équation exacte associé à l'approximation hybride lorsque* $\varepsilon = 10^{-1}$ *(*$d \approx 3$*).* $\tau_\star = 100$*,* $\varpi_\star = 0.7$*. Projection. Grille d'approximation 4.*

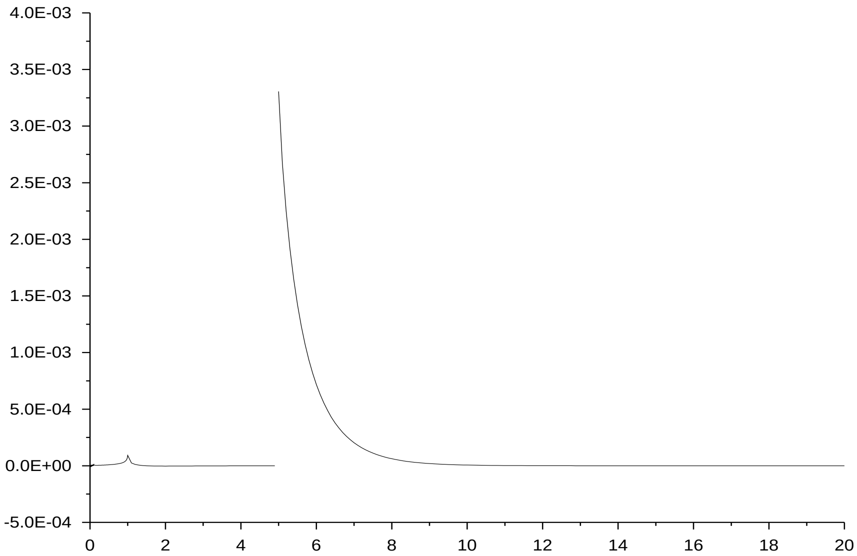


FIG. 3.39 – *Zoom du résidu sur associé à l'approximation hybride lorsque* $\varepsilon = 10^{-2}$ *(*$d \approx 5$*) .* $\tau_\star = 100$*,* $\varpi_\star = 0.7$*. Projection. Grille d'approximation 4.*

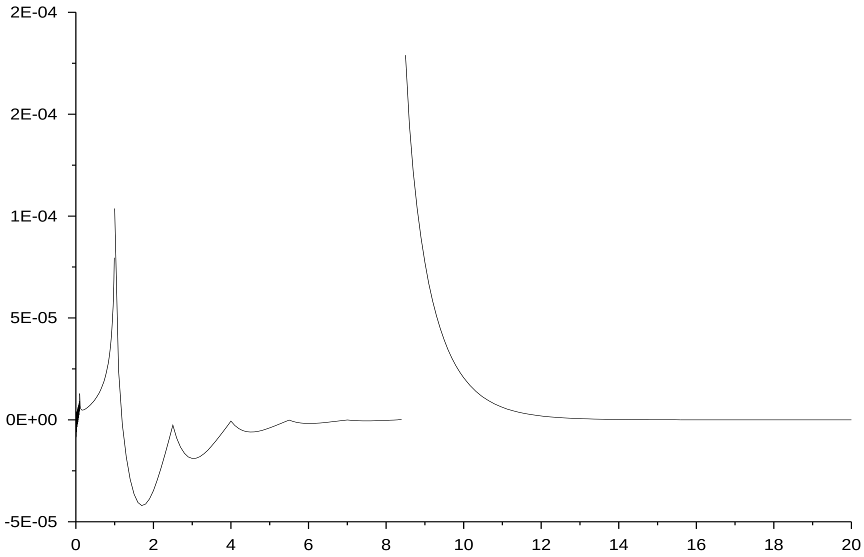


FIG. 3.40 – *Zoom du résidu sur associé à l'approximation hybride lorsque* $\varepsilon = 10^{-3}$ *(*$d \approx 8.5$*).* $\tau_\star = 100$*,* $\varpi_\star = 0.7$*. Projection. Grille d'approximation 4.*

Conclusion et perspectives

Cette étude, motivée par des applications physiques, nous a amené à analyser des méthodes d'approximations de rang fini pour la résolution des équations de Fredholm de second espèce faiblement singulières, posées dans $C^0([0, \tau_\star])$ ou $L^1([0, \tau_\star])$. Nous avons d'ailleurs formulé de manière précise et plus générale la notion de noyau faiblement singulier.

Si l'application de ces méthodes à l'espace $C^0([0, \tau_\star])$ est largement évoquée dans la plupart des ouvrages traitant de ce sujet, il n'en est pas de même de l'espace $L^1([0, \tau_\star])$. Nous avons aussi mis en évidence le rôle de la régularité du maillage et de la singularité du noyau dans l'estimation des bornes d'erreurs relatives des solutions approchées.

La particularité du problème physique considéré nous a conduit à revisiter les schémas de raffinement itératif de solutions approchées comme celui d'Atkinson et de Brakhage et d'en proposer un autre : le schéma (C). Après avoir appliqué nos estimations d'erreur à ces schémas, un effort particulier a été porté sur la mise en œuvre effective de ces derniers, notamment sur le choix du critère d'arrêt et sur l'exposition des solutions cohérente avec l'espace fonctionnel pris en compte. Nous avons proposé une variante permettant de modifier la structure du système linéaire à résoudre à chaque itération (passage d'une matrice pleine à une matrice «creuse» voire «bande»), ce qui peut être utile lors de la parallélisation de ces méthodes.

Enfin, nous n'avons pas voulu négliger la description du problème physique global sous-jacent dont la résolution de l'équation de transfert fait partie. Nous avons ainsi pris en compte dans les applications numériques des exemples se rapprochant au mieux de problèmes posés dans les atmosphères stellaires. Ce point de vue nous a d'ailleurs poussé à étudier et mettre en œuvre des méthodes asymptotiques efficaces pour essayer de surmonter la difficulté majeure de ce problème qui réside en l'amplitude très importante de l'intervalle d'intégration.

Dans l'avenir des efforts particuliers devrons être portés sur la comparaison des méthodes décrites dans ce travail avec des méthodes déjà existantes pour la résolution du problème de transfert (Lambda-itération, méthode ALI) et sur la description mathématique précise de ces dernières. Cela passe notamment par une étude plus approfondie des espaces fonctionnels de référence éventuellement plus vastes que $L^1([0, \tau_\star])$. Nous ne devrons d'ailleurs pas négliger l'adaptation des méthodes décrites ici à des noyaux plus généraux que l'exponentielle intégrale, ce qui nous conduira sans doute à en proposer de nouvelles. Ce travail nous permettra de plus de revisiter la formulation de l'équation de transfert sous sa forme «pseudo-intégro-différentielle» initiale, ce qui semblerait impliquer l'utilisation de méthodes numériques pour la résolution d'équations de Sylvester.

Les méthodes de développement asymptotique de décomposition du domaine, dont l'efficacité est séduisante, devront être étudiées plus en détail afin de pouvoir affaiblir les hypothèses faites sur la régularité des données.

Enfin, indépendamment des applications astrophysiques, une étude complémentaire sur les problèmes spectraux naturellement associés aux équations de seconde espèce reste à faire.

Table des figures

Bibliographie

[1] Abramowitz M. and Stegun I.A. *Handbook of Mathematical Function*. National Bureau Of Standards, Applied Math.Series, Washington D.C. - US Government Printing Office, 1967.

[2] Ahues M. A class of strongly stable operator approximations. *J. Austral. Math. Soc. (Series B)*, 28:435–442, 1987.

[3] Ahues M. and Hocine F. A note on spectral approximation of linear operations. *Appl. Math. Lett.*, 7(2):63–66, 1994.

[4] Ahues M. and Largillier A. Two numerical approximations for a class of weakly singular integral operators. *Applied Numerical Mathematics*, 17:347–362, 1995.

[5] Ahues M., d'Almeida F. and Telias M. On the defect correction method with applications to iterative refinement techniques. Technical Report 324, IMAG, Univ. Grenoble, 1982.

[6] Ahues M., D'Almeida F., Largillier A., Titaud O. and Vasconcelos P. An L^1 refined projection approximate solution of the radiation transfert equation in stellar atmospheres. *Journal of Comp. and Appl. Math.*, 2001.

[7] Ahues M., Largillier A. and Limaye B.V. *Spectral Computations for Bounded Operators*. Chapman and Hall/CRC, 2001.

[8] Ahues M., Largillier A. and Titaud O. The roles of weak singularity and the grid uniformity in relative error bounds. *Accepted for publication in Numerical Functional Analysis and Optimization*, 2001.

[9] Amos D.E. Algorithm 556: Exponential integrals [S13]. *ACM Transactions on Mathematical Software*, 6(3):420–428, 1980.

[10] Amos D.E. Computation of exponential integrals. *ACM Transactions on Mathematical Software*, 6(3):365–377, 1980.

[11] Amosov A. and Panasenko G. Asymptotic analysis of integral equations stated on great interval and its application to stellar radiative transfer modelling. *To appear in C.R. Acad. Sci., Paris*, 2001.

[12] Amosov A., Panasenko G. and Rutily B. One approach to approximate solving to integral equation of radiation transfer for optically thick slab. *To appear in C.R. Acad. Sci., Paris*, 2001.

[13] Anselone P.M. singularity substraction in the numerical solution of intergal equations. *J. Austral. Math. Soc (Series B)*, 22:408–418, 1981.

[14] Atkinson K.E. Iterative variants of the Nyström methods for the numerical solution of integral equations. *Numer. Math.*, 22:17–31, 1973.

[15] Atkinson K.E. *Elementary Numerical Analysis.* John Wiley & Sons, Singapore, second edition, 1993.

[16] Auer L. An Hermitian method for the solution of radiative transfer problems. *J. Quant. Spectrosc. Radiat. Transfer*, 16:931–937, 1976.

[17] Auer L. Acceleration of convergence. In *Stellar Atmospheres: Beyond Classical Models*, NATO ASI, pages 9–17. Kluwer Academic publishers, 1991.

[18] Auer L., Fabiani P., Bendicho P. and Trujillo-Bueno J. Multidimensional radiative transfer with multilevel atoms: ALI method with preconditionning of the rate equations. *AA*, 292:599–615, 1994.

[19] Auer L.L. and Paletou F. Two-dimentional radiative transfer with partial frequency redistribution. i. general method. *Astron. Astrophys.*, 285:675–686, 1994.

[20] Beylkin G., Coifman R. and Rokhlin V. Fast wavelet transforms and numerical algorithms. I. *Communication on Pure and Applied Mathematics*, 44(2):141–183, 1991.

[21] Brakhage H. Über die numerische Behandlung von Integralgleichungen nach der Quadraturformelmethode. *Numer. Math.*, 2:183–196, 1960.

[22] Brezis H. *Analyse Fonctionnelle.* Masson, Paris, 1983.

[23] Busbridge I.W. *The Mathematics of Radiative Transfert.* Cambridge University Press, 1960.

[24] Cannon C.J. *The Transfert of Spectral Line Radiation.* Cambridge University Press, 1985.

[25] Carbon D.F. Line blanketing. In Kalkofen W., editor, *Methods in radiative transfer.* 1984.

[26] Chatelin F. *Spectral Approximation of Linear Operators.* Academic Press, New-York, 1983.

[27] Chatelin F. *Valeurs Propres de Matrices.* Masson, 1988.

[28] Chevallier L. *La thermalisation des électrons dans une atmosphère stellaire.* Thèse de doctorat, Université Claude Bernard - Lyon1, 2000.

[29] Dautray R. et Lions J.L. *Analyse Mathématique et Calcul Numérique pour les Sciences et Techniques*, volume 9. INSTN CEA Masson, 1988.

[30] Delves L.M. and Mohamed J.L. *Computational Methods for Integral Equations.* Cambridge University Press, 1985.

[31] Duderstadt J.J. and Martin W.R. *Transport Theory.* John Wiley and Sons, 1979.

[32] Feautrier. Sur la résolution numérique de l'équation de transfert. Comptes Rendus 258(12):3189-3191, Academie des Sciences de Paris, 1964.

[33] Fontenla J.M. and Rovira M.G. Numerical Evaluation in radiative transfer. *J. Quant. Spectrosc. Radiat. Transfer*, 34(5):389–403, 1985.

[34] Friedman M.B. and Dellwo D.R. Accelerated quadrature methods for linear integral equations. *Journal of Integral Equations*, 8:113–136, 1985.

[35] Gantmacher F.R. *Théorie des Matrices*, volume 1. Dunod, Paris, 1966.

[36] Garcia R.D.M and Siewert, C.E. Radiative transfert in finite inhomogeneous plan-parallel atmospheres. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 27:141–148, 1982.

[37] Golub G.H. and Van Loan C. *Matrix Computations.* The John Hopkins University Press, 1989.

[38] Graham I.G. and Sloan I.H. On the compactness of certain integral operators. *Journal of Mathematical Analysis and Applications*, 68:580–594, 1979.

[39] Graham I.G., Joe S. and Sloan I.H. Iterated Galerkin versus iterated collocation for integral equations of the second kind. *J. Numer. Anal.*, 5:355–369, 1985.

[40] Grammont L. *Analyse numérique des équations de Sylvester généralisées.* Thèse de doctorat, Universite de Saint Etienne, 1994.

[41] Greengard L. and Rokhlin V. A fast algorithm for particle simulations. *Journal of Computational Physics*, 73:187–207, 1987.

[42] Hackbusch W. *Integral Equations. Theory and Numerical Treatment.* Birkäuser-Verlag, Berlin, 1995.

[43] Hackbusch W. and Nowak Z.P. On the fast matrix multiplication in the boundary element method by panel clustering. *Numerical Mathematics*, 56:229–245, 1994.

[44] Hopf E. *Mathematical problems of radiative equilibrium.* Cambridge University Press, Cambridge, 1934.

[45] Kalkofen W. A comparison of differential and integral equation of radiative transfer. *J. Quant. Spectrosc. Radiat. Transfer*, 14:309–316, 1974.

[46] Kourganoff V. *Basic Methods in Transfer Problems.* Oxford University Press, 1952. Dover, New York (1963).

[47] Kress R. *Numerical Analysis.* Number 181 in Graduate Texts in Mathematics. Springer-Verlag, New York, 1998.

[48] Kress R. *Linear Integral Equations*, volume 82 of *Applied Mathematical Sciences.* Springer-Verlag, New York, second edition, 1999.

[49] Largillier A. A numerical quadrature for some weakly singular integral operators. *Appl. Math. Lett.*, 8(1):11–14, 1995.

[50] Largillier A. and Levet M. A note on a collectively compact approximation for a weakly singular integral operators. *Appl. Math. Lett.*, 6(4):87–90, 1993.

[51] Largillier A. and Titaud O. Product integration quadratures for the radiative transfert problem with Hopf's kernel. *Accepted for publication in Integral Methods in Science and Engineering, C. Constanda Ed. Birkhauser-Verlag, Berlin, New-York*, 2001.

[52] LARGILLIER A. and TITAUD O. Spectral approximation of weakly singular integrable kernels using projections. *Accepted for publication in Integral Methods in Science and Engineering, C. Constanda Ed. Birkhauser-Verlag, Berlin, New-York*, 2001.

[53] LEBBAR R. *Sur les propriétés de superconvergence des solutions approchées de certaines équations intégrales et différentielles*. Thèse de doctorat, Université Scientifique et Médicale de Grenoble et Institut National Polytechnique de Grenoble, 1981.

[54] LIMAYE B.V. *Functionnal Analysis*. New Age International Limited, Publishers, Bombay, second edition, 1996.

[55] MENZEL D.H. (Ed). Selected papers on the transfer of radiation. Dover, New York, 1966.

[56] MIHALAS D. *Stellar Atmospheres*. Freeman and Company, San Fransisco, 1978.

[57] MIKHLIN S.G. *Integral Equations*. Pergamon Press, Oxford, 1964.

[58] MILNE E.A. *Thermodynamics of the stars*, volume 3, chapter 2. Springer-Verlag, Berlin, 1930.

[59] NAYLOR A.W. and SELL G.R. *Linear Operator Theory in Engineering and Science*. Number 40 in Applied Mathematical Sciences. Springer-Verlag, Berlin, 1982.

[60] OLSON G.L. and KUNASZ P.B. Short Characteristic solution of the non-dimensional planar slab. *J. Quant. Spectrosc. Radiat. Transfer*, 38(5):325–336, 1987.

[61] OLSON G.L., AUER L. and BUCHLER J.R. A rapidly convergent iterative solution of the non-LTE radiation transfert problem. *J. Quant. Spectrosc. Radiat. Transfer*, 35:431–442, 1986.

[62] OZISIK M.N. *Radiative Transfert*. John Wiley and Sons, 1973.

[63] PANASENKO G.P. Method of asymptotic partial decomposition of domain. *Mathematical Models and Methods in Applied Sciences*, 8(1):139–156, 1998.

[64] PHILLIPS J.L. The use of collocation as a projection method for solving linear operator equation. SIAM *J. Numer. Anal.*, 9(1):14–28, 1972.

[65] POGORZELSKI W. *Integral equations and their applications.* Pergamon Press, Oxford, 1966.

[66] POMRANING P.C. *The Equations of Radiation Hydrodynamics.* Pergamon Press, 1973.

[67] PRENTER P.M. A collocation method for the numerical solution of integral equations. *J. Numer. Anal.*, 10(4):571–581, 1973.

[68] ROKHLIN V. Rapid solution of integral equations of classical potential theory. *Journal of Computational Physics*, 73:325–348, 1987.

[69] RUTILY B. *Solutions exactes de l'équation de transfert et applicationss astrophysiques.* Thèse de doctorat d'État, Universite Claude Bernard - Lyon1, 1992.

[70] RUTILY B. and BERGEAT J. The Neumann solution of the multiple scattering problem in a plane parallel medium -I. the infinite medium. *J. Quant. Spectrosc. Radiat. Transfer*, 33:373–380, 1985.

[71] RYBICKI G.B. and HUMMER D.G. An accelerated lambda iteration method for multilevel radiative transfer. *Astron. Astrophys.*, 245:171–181, 1991.

[72] SLOAN I.H. Iterated Galerkin method for eigenvalue problems. *SIAM, J. Numer. Anal.*, 13:753–760, 1976.

[73] SOBOLEV V.V. *Radiative transfer in the atmospheres of stars and planets.* Gostekhizdat, Moscow, 1956.

[74] STETTER H.J. The defect correction principle and discretization methods. *Numer. Math.*, 29:425–443, 1978.

[75] STOER J. and BURLISCH J. *Introduction to Numerical Analysis.* Springer-Verlag, 1983.

[76] SWIHART Th.L. *Radiation Transfert and Stellar Atmospheres.* Pachart Publishing House, Tucson (Arizona), 1981.

[77] TYRTYSHNIKOV E.E. Fast algorithms for toeplitz and quasi-toeplitz systems. *J. Numer. Anal. Math. Modelling*, 4(5):419–430, 1989.

[78] VAN DE HULST H.C. *Multiple light scattering.* Academic Press, 1980.

[79] VARGA R.S. *Matrix Iterative Analysis.* Prentice-Hall, Englewood Cliffs New York, 1962.

[80] Williams M.M.R. *Mathematical Methods in Particle Transport Theory*. Butterworths, 1971.

[81] Wrench J.W. A new calculation of euler's constant. *Mathematical Tables and Other Aids to Computation*, 6:255, 1952.

Annexe A

Coefficients des matrices et vecteurs

On trouvera dans cette annexe les expressions des coefficients des matrices d'approximation et des seconds membres intervenant dans le système (2.5) dans le cas des approximations par projection et intégration produit. La dépendance des formules au choix du noyau κ et de l'albédo ϖ sera précisée le cas échéant. Les termes libres f_1, f_2, f_3, f_4 et f_5 sont définis respectivement par (3.64), (3.66), (3.68), (3.69) et (3.70).

Coefficients des matrices

Approximations par projection

Application à $C^0([0,\tau_\star])$

Lorsque κ est définie par (3.59), les coefficients de la matrice A_n du système linéaire (2.5) définis par (2.19) sont donnés pour tout i et j dans $[\![1,n]\!]$ par

$$\mathsf{A}_n(i,j) := (Te_{n,j})(\tau_{n,i}) = \frac{\varpi(\tau_{n,i})}{2}\int_0^{\tau_\star} E_1(|\tau_{n,i}-\sigma|)e_{n,j}(\sigma)\,d\sigma.$$

Définissons pour tout $(i,j) \in [\![1,n]\!]^2$

$$d_{n,i,j} := |\tau_{n,i} - \tau_{n,j}|.$$

Les coefficients de A_n peuvent alors être écrits en terme de ces distances et des fonctions F_2 et E_3 grâce à la relation (3.49), de la manière suivante :

$$\begin{aligned}
\mathsf{A}_n(1,1) &= \frac{\varpi(0)}{2}\left(1 + \frac{1}{h_{n,2}}\left[E_3(h_{n,2}) - \frac{1}{2}\right]\right), \\
\mathsf{A}_n(n,n) &= \frac{\varpi(\tau_\star)}{2}\left(1 + \frac{1}{h_{n,n}}\left[E_3(h_{n,n}) - \frac{1}{2}\right]\right).
\end{aligned}$$

Pour tout i dans $[\![2, n]\!]$,

$$\mathsf{A}_n(i,1) = \frac{\varpi(\tau_{n,i})}{2}\left(-E_2(d_{n,i,1}) + \frac{1}{h_{n,2}}\left[E_3(d_{n,i,2}) - E_3(d_{n,i,1})\right]\right).$$

Pour tout i dans $[\![1, n-1]\!]$,

$$\mathsf{A}_n(i,n) = \frac{\varpi(\tau_{n,i})}{2}\left(-E_2(d_{n,i,n}) + \frac{1}{h_{n,n}}\left[E_3(d_{n,i,n-1}) - E_3(d_{n,i,n})\right]\right).$$

Pour tout i dans $[\![1, n]\!]$ et tout j dans $[\![2, n-1]\!]$ tels que $i \neq j$,

$$\mathsf{A}_n(i,j) = \frac{\varpi(\tau_{n,i})}{2}\left(\frac{1}{h_{n,j}}\left[E_3(d_{n,i,j-1}) - E_3(d_{n,i,j})\right] + \frac{1}{h_{n,j+1}}\left[E_3(d_{n,i,j+1}) - E_3(d_{n,i,j})\right]\right).$$

Enfin, pour tout j dans $[\![2, n-1]\!]$,

$$\mathsf{A}_n(j,j) = \frac{\varpi(\tau_{n,j})}{2}\left(2 + \frac{1}{h_{n,j}}\left[E_3(h_{n,j}) - \frac{1}{2}\right] + \frac{1}{h_{n,j+1}}\left[E_3(h_{n,j+1}) - \frac{1}{2}\right]\right).$$

Application à $L^1([0,\tau_\star])$

Lorsque κ est définie par (3.59) et ϖ par (3.62), les coefficients de la matrice A_n du système linéaire (2.5) définis par (2.29) sont donnés pour tout $(i,j) \in [\![1, n]\!]^2$ par

$$\mathsf{A}_n(i,j) = \frac{1}{2h_{n,i}}\int_{\tau_{n,i-1}}^{\tau_{n,i}} \varpi(\tau)\int_{\tau_{n,j-1}}^{\tau_{n,j}} E_1(|\tau-\sigma|)\, d\sigma\, d\tau.$$

On définit, pour tout i et j dans $[\![0, n]\!]$,

$$d_{n,i,j} := |\tau_{n,i} - \tau_{n,j}|.$$

Les coefficients de A_n peuvent alors être écrits en terme de ces distances et de la fonction F_2 de la manière suivante :

Pour tout i et j dans $[\![1, n]\!]$, tels que $j > i$,

$$\begin{aligned}\mathsf{A}_n(i,j) &= \frac{\varpi_\star}{2h_{n,i}}\left[\exp(\alpha\tau_{n,j})(F_2(d_{n,i,j},-\alpha) - F_2(d_{n,i-1,j},-\alpha))\right.\\ &\quad \left.+ \exp(\alpha\tau_{n,j-1})(F_2(d_{n,i-1,j-1},-\alpha) - F_2(d_{n,i,j-1},-\alpha))\right].\end{aligned}$$

Pour tout j dans $[\![1, n]\!]$,

$$\mathsf{A}_n(j,j) = \frac{\varpi_\star}{2h_{n,j}}\left[\exp(\alpha\tau_{n,j})\Big(\frac{2}{\alpha} - F_2(h_{n,j},-\alpha)\Big) - \exp(\alpha\tau_{n,j-1})\Big(\frac{2}{\alpha} + F_2(h_{n,j},\alpha)\Big)\right].$$

Pour tout i et j dans $[\![1,n]\!]$, tels que $j<i$,

$$\begin{aligned}\mathsf{A}_n(i,j) &= \frac{\varpi_\star}{2h_{n,i}}\left[\exp(\alpha\tau_{n,j-1})(F_2(d_{n,j-1,i-1},-\alpha)-F_2(d_{n,j-1,i},-\alpha))\right. \\ &\quad \left.+\exp(\alpha\tau_{n,j})(F_2(d_{n,j,i},-\alpha)-F_2(d_{n,j,i-1},-\alpha))\right].\end{aligned}$$

Lorsque l'albédo est une fonction constamment égale à $\varpi_\star$, ces coefficients sont donnés pour tout i et j dans $[\![1,n]\!]$ par

$$\mathsf{A}_n(i,j) = \begin{cases} \dfrac{\varpi_\star}{2h_{n,i}}\Big[E_3(d_{n,i-1,j})-E_3(d_{n,i-1,j-1})+E_3(d_{n,i,j-1})-E_3(d_{n,i,j})\Big] & \text{si } i\neq j, \\ \varpi_\star\Big(1+\dfrac{1}{h_{n,j}}\Big[E_3(h_{n,j})-\dfrac{1}{2}\Big]\Big) & \text{si } i=j. \end{cases}$$

De plus, si le maillage est uniforme, cette matrice a une structure Toeplitz, ce qui permet d'utiliser des algorithmes rapides pour la résolution du système linéaire correspondant (voir [37] et [77]).

Approximations par intégration produit

Application à l'espace $C^0([0,\tau_\star])$

Lorsque κ est définie par (3.59), les coefficients de la matrice A_n du système linéaire (2.5) définis par (2.39) sont donnés pour tout i et j dans $[\![1,n]\!]$ par

$$\mathsf{A}_n(i,j) = \begin{cases} \dfrac{\varpi(\widehat{\tau}_{n,i})}{2}\Big|E_2(|\tau_{n,j}-\widehat{\tau}_{n,i}|)-E_2(|\tau_{n,j-1}-\widehat{\tau}_{n,i}|)\Big| & \text{si } i\neq j, \\ \dfrac{\varpi(\widehat{\tau}_{n,j})}{2}(2-E_2(\widehat{\tau}_{n,i}-\tau_{n,j-1})-E_2(\tau_{n,j}-\widehat{\tau}_{n,i})) & \text{si } i=j. \end{cases}$$

Application à l'espace $L^1([0,\tau_\star])$

Lorsque κ est définie par (3.59) et ϖ par (3.62), les coefficients de la matrice A_n du système linéaire (2.5) définis par (2.53) sont donnés pour tout i et j dans $[\![1,n]\!]$ par

$$\mathsf{A}_n(i,j) = \frac{h_{n,j}}{h_{n,i}}\int_{\tau_{n,i-1}}^{\tau_{n,i}} \varpi(\tau)E_1(|\tau-\widehat{\tau}_{n,j}|)\,d\tau.$$

Pour tout i et j dans $[\![1,n]\!]$, on pose

$$\widehat{d}_{n,i,j} := |\tau_{n,i}-\widehat{\tau}_{n,j}|.$$

Alors pour tout i et j dans $[\![1, n]\!]$,

$$\mathsf{A}_n(i,j) = \begin{cases} \dfrac{\varpi_\star h_{n,j}}{2h_{n,i}} \exp(\alpha \widehat{\tau}_{n,j})(F_1(\widehat{d}_{i,j}, \alpha) - F_1(\widehat{d}_{n,i-1,j}, \alpha)), & \text{si } j < i, \\ \dfrac{\varpi_\star}{2} \exp(\alpha \widehat{\tau}_{n,j})(F_1(\widehat{d}_{n,j,j}, \alpha) + F_1(\widehat{d}_{n,j-1,j}, -\alpha)), & \text{si } j = i, \\ \exp(\alpha \widehat{\tau}_{n,j})(F_1(\widehat{d}_{i-1,j}, -\alpha) - F_1(\widehat{d}_{n,i,j}, -\alpha)), & \text{si } j > i. \end{cases}$$

Coefficients des seconds membres

Second membre correspondant au terme libre f_1

Approximations par intégration produit

Les coefficients du vecteur second membre du système linéaire (2.5) définis par (2.40) sont donnés dans ce cas pour tout i dans $[\![1, n]\!]$ par

$$\mathsf{b}_{n,i} = \varpi(\widehat{\tau}_{n,i}) - 1 - \frac{\varpi(\widehat{\tau}_{n,i})}{2}(E_2(\widehat{\tau}_{n,i}) + E_2(\tau_\star - \widehat{\tau}_{n,i})).$$

Second membre correspondant au terme libre f_2

Approximations par intégration produit

Les coefficients du vecteur second membre du système linéaire (2.5) définis par (2.40) sont donnés dans ce cas pour tout i dans $[\![1, n]\!]$ par

$$\mathsf{b}_n(i) = \exp(\beta \widehat{\tau}_{n,i}))\Big[\frac{\varpi(\widehat{\tau}_{n,i})x}{2}(F_1(\widehat{\tau}_{n,i}, -\beta) + F_1(\tau_\star - \widehat{\tau}_{n,i}, \beta)) - 1\Big].$$

Second membre correspondant au terme libre f_3

Approximations par projection

Soit κ définie par (3.59) et ϖ par (3.62). Alors le terme libre f_3 défini par (3.68) est continu car $\varpi \in C^0([0, \tau_\star])$. Les coefficients du vecteur second membre du système linéaire (2.5) définis par (2.20) sont donnés dans ce cas pour tout i dans $[\![1, n]\!]$ par

$$\begin{aligned} \mathsf{b}_n(i) \quad := \quad & \frac{\varpi_\star \exp(\alpha \tau_{n,i})}{2}[E_2(\tau_\star - \tau_{n,i}) + E_2(\tau_{n,i}) - 2 \\ & + \varpi_\star \exp(\alpha \tau_{n,i})(F_1(\tau_{n,i}, -\alpha) + F_1(\tau_\star - \tau_{n,i}, \alpha))]. \end{aligned}$$

Approximations par intégration produit

Les coefficients du vecteur second membre du système linéaire (2.5) définis par (2.40) sont donnés dans ce cas pour tout i dans $[\![1,n]\!]$ par

$$\mathsf{b}_n(i) := \varpi(\hat{\tau}_{n,i}) - 1.$$

Second membre correspondant au terme libre f_4

Approximations par projection

Soit κ définie par (3.59) et ϖ par (3.62). Alors les coefficients du vecteur second membre du système linéaire (2.5) définis par (2.30) sont définis dans ce cas pour tout i dans $[\![1,n]\!]$ de la manière suivante :

Si $\theta \leq \tau_{n,i-1}$,

$$\begin{aligned}\mathsf{b}_n(i) &= \frac{\varpi_\star(\varpi_\star - 1)}{2h_{n,i}}\Big[\exp(\alpha\theta)(F_2(\tau_{n,i}-\theta,\alpha) - F_2(\tau_{n,i-1}-\theta,\alpha)) \\ &\quad + F_2(\tau_{n,i-1},\alpha) - F_2(\tau_{n,i},\alpha)\Big].\end{aligned}$$

Si $\theta \in \,]\tau_{n,i-1}, \tau_{n,i}[$,

$$\begin{aligned}\mathsf{b}_n(i) &= \frac{\varpi_\star(\varpi_\star - 1)}{2h_{n,i}}\Big[F_2(\tau_{n,i-1},\alpha) - F_2(\tau_{n,i},\alpha) + \frac{2}{\alpha}(\exp(\alpha\theta) - \exp(\alpha\tau_{n,i-1})) \\ &\quad + \exp(\alpha\theta)(F_2(\tau_{n,i}-\theta,\alpha) - F_2(\theta-\tau_{n,i-1},-\alpha))\Big].\end{aligned}$$

Si $\tau_{n,i} \geq \theta$,

$$\begin{aligned}\mathsf{b}_n(i) &= \frac{\varpi_\star(\varpi_\star - 1)}{2h_{n,i}}\Big[F_2(\tau_{n,i-1},\alpha) - F_2(\tau_{n,i},\alpha) + \frac{2}{\alpha}(\exp(\alpha\tau_{n,i}) - \exp(\alpha\tau_{n,i-1})) \\ &\quad + \exp(\alpha\theta)(F_2(\theta-\tau_{n,i},-\alpha) - F_2(\theta-\tau_{n,i-1},-\alpha))\Big].\end{aligned}$$

Si ϖ est une fonction constamment égale à $\varpi_\star$, ces coefficients sont définis pour tout i dans $[\![1,n]\!]$ par

$$\mathsf{b}_n(i) := \frac{\varpi_\star(\varpi_\star - 1)}{2}\Big[2 - E_2(\tau_{n,i}) - E_2(\tau_\star - \tau_{n,i})\Big].$$

Approximations par intégration produit

Les coefficients du vecteur second membre du système linéaire (2.5) définis par (2.54) sont définis dans ce cas pour tout i dans $[\![1, n]\!]$ par

$$\mathsf{b}_n(i) = \begin{cases} 0 & \text{si } \theta \leq \tau_{n,i-1}, \\ \dfrac{1}{h_{n,i}}(\varpi_\star - 1)(\theta - \tau_{n,i-1}) & \text{si } \tau_{n,i-1} < \theta < \tau_{n,i}, \\ \varpi_\star - 1 & \text{si } \theta \geq \tau_{n,i}. \end{cases}$$

Second membre correspondant au terme libre f_5

Approximations par intégration produit

Soit κ définie par (3.59). Alors les coefficients du vecteur second membre du système linéaire (2.5) définis par (2.54) sont donnés dans ce cas pour tout i dans $[\![1, n]\!]$ par

$$\mathsf{b}_n(i) = \begin{cases} \dfrac{\varpi_\star}{2h_{n,i}}(2 - E_2(\tau_{n,i} - \theta) - E_2(\theta - \tau_{n,i-1})) & \text{si } \theta \in [\tau_{n,i-1}, \tau_{n,i}], \\ \\ \dfrac{\varpi_\star}{2h_{n,i}}|E_2(|\tau_{n,i} - \theta|) - E_2(|\theta - \tau_{n,i-1}|)| & \text{sinon.} \end{cases}$$

Annexe B

Articles