

Titre: Interpolation itérative et compression de données
Title:

Auteurs: Jean-Pierre Mongeau, Serge Dubuc, & Gilles Deslauriers
Authors:

Date: 1990

Type: Rapport / Report

Référence: Mongeau, J.-P., Dubuc, S., & Deslauriers, G. (1990). Interpolation itérative et compression de données (Technical Report n° EPM-RT-90-06).
Citation: <https://publications.polymtl.ca/9877/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/9877/>
PolyPublie URL:

Version: Version officielle de l'éditeur / Published version

Conditions d'utilisation: Tous droits réservés / All rights reserved
Terms of Use:

 **Document publié chez l'éditeur officiel**
Document issued by the official publisher

Institution: École Polytechnique de Montréal

Numéro de rapport: EPM-RT-90-06
Report number:

URL officiel:
Official URL:

Mention légale:
Legal notice:

27 AOUT 1990

(INTERPOLATION ITÉRATIVE ET COMPRESSION
DE DONNÉES)

Jean-Pierre ^{ET AL} Mongeau, étudiant Ph.D.
Serge Dubuc, professeur
Gilles Deslauriers, professeur

avril (1990)

gabriel

INTERPOLATION ITÉRATIVE ET COMPRESSION DE DONNÉES(*)

Jean-Pierre Mongeau⁽¹⁾, étudiant (Ph.D.)

Serge Dubuc⁽²⁾, professeur

Gilles Deslauriers⁽¹⁾, professeur

DÉPARTEMENT DE MATHÉMATIQUES APPLIQUÉES

ÉCOLE POLYTECHNIQUE
Avril 1990

(*) Cette étude a été réalisée grâce aux bourses et subventions octroyées par le F.C.A.R. et le C.R.S.N.G.

(1) Département de mathématiques appliquées,
École Polytechnique, C.P. 6079, Succ. A,
Montréal, Québec, H3C 3A7 Canada.

(2) Département de mathématiques et de statistique,
Université de Montréal, C.P. 6128, Succ. A,
Montréal, Québec, H3C 3J7 Canada.

Ce document a pu être publié grâce à la subvention FCAR ER0287 F.

Tous droits réservés. On ne peut reproduire ni diffuser aucune partie du présent ouvrage, sous quelque forme que ce soit, sans avoir obtenu au préalable l'autorisation écrite de l'auteur.

Dépôt légal, 1^{er} trimestre 1990
Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

Pour ce procurer une copie de ce document, s'adresser aux:

Éditions de l'École Polytechnique de Montréal
École Polytechnique de Montréal
Case postale 6079, Succursale "A"
Montréal (Québec) H3C 3A7
(514) 340-4000

Compter 0,10 \$ par page (arrondir au dollar le plus près) et ajouter 3,00 \$ (Canada) pour la couverture, les frais de poste et la manutention. Régler en dollars canadiens par chèque ou mandat-poste au nom de l'École Polytechnique de Montréal. Nous n'honorons que les commandes accompagnées d'un paiement, sauf s'il y a eu entente préalable, dans le cas d'établissements d'enseignement, de société ou d'organismes canadiens.

Table des matières

Résumé	i
Abstract	ii
Introduction	1
Chapitre 1 - Interpolation itérative	3
Chapitre 2 - Interpolation sur la droite	12
Chapitre 3 - Compression de données	39
Conclusion	55
Bibliographie	56

RÉSUMÉ

Ce rapport est consacré à l'étude de propriétés du processus d'interpolation itérative tel que défini par G. Deslauriers et S. Dubuc. A la base, ce processus permet l'interpolation de fonctions connues sur un sous-groupe de la droite. La cardinalité de ce sous-groupe est infinie. Nous définirons l'interpolation mixte, un processus d'interpolation itérative adapté aux ensembles de cardinalité finie. L'ordre et le comportement (la croissance) de l'erreur d'interpolation sont étudiés. A partir d'une étude du développement en série de Taylor de la fonction à interpoler, nous démontrerons que la croissance de l'erreur au cours de la réalisation du processus itératif est linéaire. En fait, nous démontrerons que l'erreur est multipliée par une constante à chaque étape du processus. Cette constante est indépendante de l'étape en soi. A titre d'application, nous présenterons un processus de compression de données. Une borne de l'erreur globale, définie comme étant le cumul des erreurs inhérentes à chaque étape du processus itératif, et une estimation de cette borne sont présentées. L'estimation, très utile, ne fait intervenir que les entrées de la suite de données initiale.

ABSTRACT

We will study a few properties of the iterative interpolation process. This process was proposed by G. Deslauriers and S. Dubuc. To perform this process, we need to know the interpolated function at an infinite data set. We will propose a new iterative process adapted to the finite data set. Moreover, we will study the interpolation error. We will demonstrate that there exist a positive real number such that the error is multiplied by it at each step of the iterative process. The linear growing of the error allows us to develop an efficient data compression process. A boundary of the rebuilt error is given. This boundary depends only on the values found in the initial sequence of data. Those results and concepts could be used to define an interpolation process and a data compression process in the n dimensional euclidean space.

INTRODUCTION

Une des premières définitions de l'interpolation itérative a été proposé par S. Dubuc [4]. Il a défini ce processus d'interpolation dans le cas particulier d'une fonction $f: \mathbb{R} \rightarrow \mathbb{R}$ connue sur l'ensemble Z . Ce processus permettait de prolonger la fonction f à une fonction F connue sur l'ensemble des nombres dyadiques. Un ensemble que nous savons dense sur la droite \mathbb{R} .

Il construisait la fonction d'interpolation F (le prolongement) à l'aide du procédé itératif suivant: commençons par identifier les valeurs de F sur Z aux valeurs de f ; supposons maintenant qu'on connaisse F sur l'ensemble $\{n/2^m: n \in \mathbb{Z}\}$ ($m \in \mathbb{N} \cup \{0\}$); le prolongement F en un point de l'ensemble $\{(2n-1)/2^{m+1}: n \in \mathbb{Z}\}$ se calcule à l'aide de la somme pondérée suivante: $F((2n-1)/2^{m+1}) =$

$$(-F((n-2)/2^m) + 9F((n-1)/2^m) + 9F(n/2^m) - F((n+1)/2^m))/16 .$$

Remarquez que les poids de cette somme ne dépendent ni du point $n/2^m$, ni de la valeur de m (étape du processus itératif). De plus, bien que ce processus n'utilise que quatre valeurs de F définies en des points situés autour du point d'interpolation, on peut démontrer que la fonction F est continue et différentiable.

Comme nous venons de le voir, le calcul de l'interpolation utilise une suite de valeurs connues de taille infinie. En effet, nous devons connaître f partout sur Z . Le problème que nous aborderons est de définir et d'étudier un processus d'interpolation itérative capable de traiter le cas des ensembles de taille finie. Après une brève description, au chapitre un, de l'interpolation itérative et de certaines de ses propriétés, nous allons identifier, au chapitre deux, un certain nombre de processus d'interpolation itérative sur la droite tel que définis par G. Deslauriers et S. Dubuc [1,2,3,4]. Par la suite, nous ferons une étude détaillée de l'erreur d'interpolation de ce nouveau processus. Finalement, comme exemple d'utilisation du processus d'interpolation, nous présenterons un processus de compression de données. Nous ferons également une étude détaillée de l'erreur de reconstruction des suites comprimées.

CHAPITRE 1, INTERPOLATION ITÉRATIVE

1.1 Description du processus

Un processus d'interpolation itérative sur la droite, de même que dans l'espace \mathbb{R}^n , est défini à partir du triplet (G, T, w) . Sur la droite, ce triplet contient les trois paramètres suivants: G est le sous-groupe \mathbb{Z} de \mathbb{R} ; T est une transformation linéaire sur \mathbb{R} , soit $T(x) = x/b$ où b est un entier plus grand que 1; w est une fonction de \mathbb{R} à valeurs dans \mathbb{R} . Il faut supposer, comme première hypothèse, que la fonction à interpoler est connue partout sur \mathbb{Z} . Définissons une suite de sous-groupes $\{G_n\}_{n=0,1,\dots,\infty}$ de l'espace \mathbb{R} de la façon suivante:

$$G_0 = \mathbb{Z}$$

$$G_1 = T(G_0) = \{ T(x) : x \in G_0 \} = \{ x/b : x \in \mathbb{Z} \}$$

...

$$G_n = T(G_{n-1}) = T^n(G_0) = \{ x/b^n : x \in \mathbb{Z} \}$$

...

Pour toute valeur de b , le sous-groupe limite G_∞ , la réunion des G_n , est dense dans \mathbb{R} . Appelons w : fonction de poids. La nature itérative du processus d'interpolation apparaît ici. A partir des valeurs de la fonction à interpoler, par hypothèse celles définies sur \mathbb{Z} , nous allons calculer les valeurs définies sur G_1 . Connaissant

maintenant les valeurs sur G_1 , nous pouvons répéter le même processus pour connaître les valeurs sur G_2 et, à la limite, les valeurs sur le sous-groupe G_∞ . Le calcul de l'interpolation, lors du passage du sous-groupe G_n au sous-groupe G_{n+1} se fera comme suit: si x est un point de G_{n+1} alors il existe un point y de Z tel que $x = y/b^{n+1}$, la valeur de l'interpolation f en x sera:

$$f(x) = \sum_{z \in Z} w(y/b-z) f(z/b^n) \quad . \quad (1)$$

La valeur de $f(x)$ ne sera donc qu'une somme pondérée des valeurs définies par l'interpolation F sur G_n . Afin d'avoir une fonction f de collocation, nous supposerons que la fonction de poids w est nulle en tout point de G sauf à l'origine. En ce point nous supposerons qu'elle vaut 1. Afin de simplifier un peu plus le processus, nous supposerons que w est non-nulle en un nombre fini de points de G_1 . Cette hypothèse règle définitivement le problème lié à la convergence de la somme définissant l'interpolation. Nous connaissons maintenant la définition des paramètres qui caractérisent un processus d'interpolation itérative.

1.2 Exemple, interpolation itérative de Lagrange

Soit N un entier plus grand que 0 et L_i le polynôme de Lagrange du $2N-1$ ^{ième} degré:

$$L_i(x) = \prod_{\substack{j=-N \\ j \neq i}}^{N+1} \frac{(x-j)}{(i-j)} .$$

Définissons w , la fonction de poids, telle que si $x \in G_1$, $x = n + r/b$, $n \in \mathbb{Z}$ et $r \in \{0, 1, \dots, b-1\}$ alors:

$$w(x) = \begin{cases} L_{-n}(r/b) & \text{si } n \in \{-N+1, \dots, N\} \\ 0 & \text{si non} \end{cases}$$

Cette fonction de poids caractérise le processus d'interpolation itérative de Lagrange de type (b, N) . La définition du processus d'interpolation itérative présenté dans l'introduction correspond à l'interpolation itérative de Lagrange de type $(2, 2)$.

1.3 Théorème de prolongement

A la section 1.1, nous avons défini un processus qui nous permet d'interpoler une fonction en tout point d'un ensemble G_∞ dense dans \mathbb{R} . Voyons sous quelles conditions cette fonction d'interpolation peut être prolongée à tout l'espace euclidien \mathbb{R} . Le théorème suivant se retrouve, sous une forme plus générale, dans le livre de topologie de J. Dugundji [5].

Théorème 1 Soit $F:G_\infty \rightarrow \mathbb{R}$ une fonction, si F est uniformément continue et si G_∞ est dense dans \mathbb{R} , alors il existe une et une seule fonction $\bar{F}: \mathbb{R} \rightarrow \mathbb{R}$ uniformément continue telle que $\bar{F}(x) = F(x)$ pour tout x dans G_∞ .

Démonstration: Soit x un point de \mathbb{R} et $\{x_n\}_{n=1, \dots, \infty}$ une suite dans G_∞ qui converge vers x .

Définissons: $\bar{F}(x) = \lim_{n \rightarrow \infty} F(x_n)$.

Montrons que cette limite existe et ne dépend pas du choix de la suite $\{x_n\}$. Soit $\varepsilon > 0$, par hypothèse F est uniformément continue, il existe donc un nombre $\delta > 0$ tel que si $y, z \in G_\infty$ et $|y-z| < \delta$ alors $|F(y) - F(z)| < \varepsilon$.

Puisque la suite $\{x_n\}$ converge vers x alors il existe un entier N tel que pour tout $n \geq N$: $|x - x_n| < \delta/2$. Ainsi pour tout $n, m \geq N$ nous avons:

$$|x_n - x_m| < \delta \text{ et } |F(x_n) - F(x_m)| < \varepsilon.$$

La suite $\{F(x_n)\}$ est donc une suite de Cauchy. Considérons maintenant une deuxième suite $\{y_n\}$ qui converge vers x , la suite $\{z_n\} = \{x_1, y_1, x_2, y_2, \dots\}$ convergera également vers x . L'argument précédent nous permet de dire que $\{F(z_n)\}$ forme une suite convergente. Les deux sous-suites $\{F(x_n)\}$ et $\{F(y_n)\}$ doivent donc converger vers la même limite. Montrons maintenant que \bar{F} est uniformément continue. Soit x, y deux points de \mathbb{R} et $\{x_n\}, \{y_n\}$ deux suites de G_∞ convergentes vers x et y respectivement.

Soit $\varepsilon > 0$ et $\delta > 0$ issus de la convergence uniforme de F .

Soit N un entier tel que si $n \geq N$ alors:

$$|x - x_n| \leq \delta/3 \text{ et } |\bar{F}(x) - F(x_n)| \leq \varepsilon.$$

Soit M un entier tel que si $n \geq M$ alors:

$$|y - y_n| \leq \delta/3 \text{ et } |\bar{F}(y) - F(y_n)| \leq \varepsilon.$$

Nous avons que : $|\bar{F}(x) - \bar{F}(y)| \leq$

$$|\bar{F}(x) - F(x_N)| + |F(x_N) - F(y_M)| + |\bar{F}(y) - F(y_M)|$$

Si $|x - y| < \delta/3$ alors $|x_N - y_M| < \delta$ et $|F(x_N) - F(y_M)| \leq \varepsilon$. Ainsi:

$$|\bar{F}(x) - \bar{F}(y)| \leq 3\varepsilon \quad . \quad \square$$

1.4 Propriétés élémentaires

Nous présentons, dans cette section, quelques propriétés importantes de l'interpolation itérative. Supposons que f_1, f_2 et f sont des fonctions connues sur G et, supposons que g_1, g_2 et g sont leurs interpolations respectives. L'interpolation itérative respecte les propriétés suivantes:

-Linéarité: si $c_1, c_2 \in \mathbb{R}$ et $f = c_1 f_1 + c_2 f_2$ alors $g = c_1 g_1 + c_2 g_2$

-Invariance sous la translation: si a appartient à G et si pour tout x appartenant à G : $f(x) = f_1(x+a)$,

$$\text{alors } g(x) = g_1(x+a)$$

-Homogénéité: soit k un entier positif, si pour tout point x de G on a: $f(x) = g_1(T^k(x))$ alors pour tout point x de G_∞ : $g(x) = g_1(T^k(x))$.

1.5 Fonction fondamentale

Soit (Z, T, w) le triplet qui définit un processus d'interpolation itérative. L'interpolante fondamentale F du processus sera l'interpolation de la fonction qui vaut 1 à l'origine et 0 partout ailleurs dans Z . Cette fonction d'interpolation joue un grand rôle dans l'étude de l'interpolation itérative. Nous pouvons démontrer, premièrement, que pour tout x dans G_1 :

$$F(x) = w(x).$$

De plus, l'équation (1) qui définit le calcul de l'interpolation (section 1.1) peut se récrire comme suit:

$$f(x) = \sum_{y \in G} f(y) F(x-y).$$

Puisque w est non-nulle en un nombre fini de points de G_1 , F a un support borné. Ainsi, la somme précédente est une somme finie. Les propriétés de l'interpolation seront donc celles de l'interpolante fondamentale F . Nous pouvons aussi montrer que F respecte l'équation fonctionnelle suivante, pour tout point x de G_∞ :

$$F(x/b^n) = \sum_{y \in Z} F(y/b^n) F(x-y)$$

Tous les exemples d'interpolation de ce rapport représentent les interpolantes fondamentales des divers processus étudiés.

1.6 Articles en référence

Nous allons décrire brièvement deux résultats, tirés des références [1,3], qui nous seront utiles. Premièrement, un critère sur la différentiabilité des fonctions d'interpolation a été présenté à la référence [3]. Voyons de quoi il s'agit. Le théorème suivant ne s'appliquera que pour l'interpolation itérative de Lagrange. Sans entrer dans les détails, nous trouvons les définitions et le théorème suivants: Soit P le polynôme trigonométrique:

$$P(\theta) = \sum_{k \in \mathbb{Z}} F(k/b) e^{ik\theta}.$$

Définissons: $S(\theta) = P(\theta) (\sin(b\theta/2)/\sin(\theta/2))^{-2N}$.

$S(\theta)$ peut s'écrire sous la forme: $S(\theta) = \sum_{k \in \mathbb{Z}} s_k e^{ik\theta}$.

Ces deux polynômes trigonométriques servent à construire une matrice dont le module du rayon spectral permet de dire si F est différentiable. En effet, soit B la matrice d'ordre infini:

$$B = (b_{ij})_{-\infty \leq i \leq \infty, -\infty \leq j \leq \infty} \text{ où } b_{i,j} = s_{i-j}.$$

Soit M un entier tel que pour tout entier i appartenant à $[-M, M]$ et pour tout entier j n'appartenant pas à $[-M, M]$ nous avons $b_{i,j} = 0$. Définissons B_M , la matrice "tronquée" d'ordre $2M+1$:

$$B_M = (b_{ij})_{-M \leq i \leq M, -M \leq j \leq M}.$$

Enfin, soit r le rayon spectral de la matrice B_M et E l'exposant tel que $b^E = 1/r$. E est appelé l'exposant critique.

Théorème 1 Si m est un entier positif plus petit que E , alors la dérivée d'ordre m de F existe et est continue. Si α est un nombre réel choisi entre 0 et 1 tel que $m+\alpha < E$, alors la dérivée d'ordre m de F est Lipschitzienne d'ordre α .

La référence [1] traite de l'interpolation dans l'espace euclidien \mathbb{R}^n . Il y est présenté un critère pour la continuité du processus d'interpolation. Ce critère est beaucoup moins restrictif que celui du théorème précédent, on n'impose pas de forme particulière au paramètre w . Nous trouvons:

soit $C_n(h) = \max_{z \in \mathbb{Z}} \{ \sum |F(T^n x - z) - F(T^n y - z)| : x, y \in G \text{ et } |x - y| \leq h \}$,
 soit $S_n = \{ x \in G_n : F(x) \neq 0 \}$ et R_n le rayon du plus petit intervalle centré à l'origine qui contient S_n .

Théorème 2 Si pour tout point x de G_1 : $\sum_{y \in G} F(x-y) = 1$,
s'il existe un entier n et un nombre réel
positif h tel que $h \geq 2R_n/(1-1/b^n)$, $C_n(h) < 2$
et si le sous-groupe engendré par $\{x \in G: |x| \leq h\}$
est G , alors l'interpolante fondamentale est
uniformément continue sur G_∞ .

CHAPITRE 2, INTERPOLATION SUR LA DROITE

Dans ce chapitre, nous étudierons plus particulièrement l'interpolation itérative de Lagrange de type $(2, N)$. En un premier temps, nous allons présenter les fonctions de poids et les graphes des interpolantes des cas où $N=2$ et $N=3$. Dans plusieurs applications, contrairement aux exemples que nous venons d'annoncer, la cardinalité de l'ensemble de points sur lequel nous devons bâtir l'interpolation est finie. Nous identifierons certaines fonctions de poids qui nous permettront de développer un processus itératif pour les ensembles finis. Nous devons démontrer les propriétés de continuité et de différentiabilité de ce nouveau processus. Une section portera sur l'étude de l'erreur d'interpolation. Nous y identifierons l'ordre et le comportement de cette erreur.

2.1.0 Définition et exemples d'interpolation sur la droite

Supposons que: $T(x) = x/2$. Le support de w est supposé être un sous-ensemble fini de G_1 . Il existe donc un entier N tel que l'équation (1) de la section 1.1 s'écrit (lorsque $x \in G_k - G_{k-1}$):

$$F(x) = \sum_{y=-N+1}^N w((2y-1)/2) F(x - (2y-1)/2^k).$$

La valeur de l'interpolation en x ne sera qu'une somme

pondérée de $2N$ valeurs prises par F dans G_{k-1} . Ces $2N$ valeurs se trouvent toutes dans l'intervalle centré en x :

$$[x - (2N-1)/2^k, x + (2N-1)/2^k].$$

Cette structure se représente sous la forme du schéma de la figure 1. L'interpolation sera la somme pondérée par les p_i de $2N$ valeurs prises par F aux points entourant x .

$$\begin{array}{ccccccc}
 & p_N & & p_{N+1} & & \dots & & \dots & & p_{-N+2} & & p_{-N+1} \\
 & | & & | & & \dots & & \dots & & | & & | \\
 x - (2N-1)/2^k & & x - (2N-3)/2^k & & \dots & & x & & \dots & & x + (2N-3)/2^k & & x + (2N-1)/2^k
 \end{array}$$

$$p_i = w((2i-1)/2)$$

Figure 1: schéma d'interpolation

Ce schéma permet de visualiser la position des points, relativement au point x , que l'on retrouve dans la somme de l'interpolation et les poids qui leurs sont associés.

Plutôt que de décrire la fonction de poids à l'aide d'un polynôme de Lagrange, il peut être plus facile de choisir N et w (c'est-à-dire: choisir les poids p_i) de façon à ce que $F(x)$ soit la valeur définie par un polynôme du $(2N-1)^{\text{ième}}$ degré. Ce polynôme passera par les points d'abscisses:

$$\{(x - (2y-1)/2^k, F(x - (2y-1)/2^k) : y = -N+1, -N+2, \dots, N\}.$$

Il est clair que la fonction de poids trouvée par ce procédé est la fonction de poids de l'interpolation itérative de Lagrange de type $(2,N)$.

Les sections suivantes contiennent des exemples de fonctions de poids, de schémas et de graphes d'interpolantes fondamentales, construits à partir des polynômes du troisième $(N=2)$ et du cinquième degré $(N=3)$.

2.1.1 Exemple, polynôme du troisième degré

$$\text{Supposons que: } w(x) = \begin{cases} -1/16 & \text{si } x = \pm 3/2 \\ 9/16 & \text{si } x = \pm 1/2 \\ 1 & \text{si } x = 0 \\ 0 & \text{si non} \end{cases}$$

La figure 1 représente le graphe de l'interpolante fondamentale de ce processus d'interpolation.

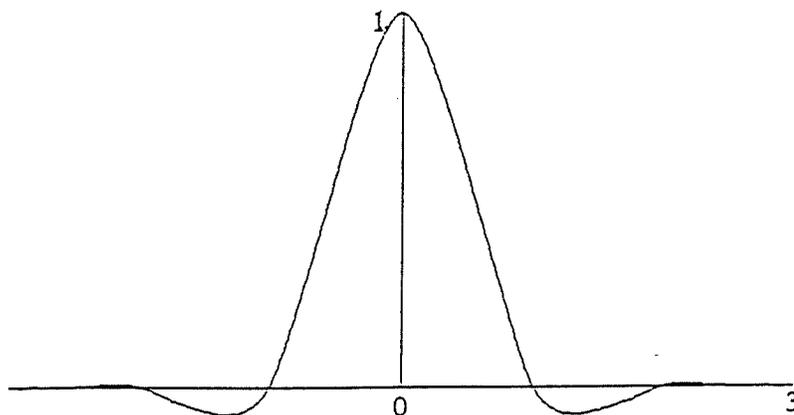


Figure 1: interpolante, polynôme du troisième degré

Nous pouvons démontrer (voir [3]) que ce processus produit une fonction d'interpolation différentiable. Par contre, la dérivée seconde n'existera pas en tout point dyadique. La figure 2 présente le schéma d'interpolation, la valeur de l'interpolation $F(x)$ s'écrit comme une somme pondérée de quatre valeurs de F (si $x \in G_k - G_{k-1}$):

$$F(x) = (-F(x-3/2^k) + 9F(x-1/2^k) + 9F(x+1/2^k) - F(x+3/2^k)) / 16 .$$

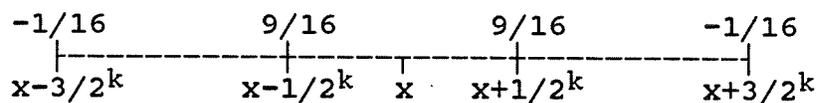


Figure 2: schéma, polynôme du troisième degré

2.1.2 Exemple, polynôme du cinquième degré

$$\text{Supposons que: } w(x) = \begin{cases} 3/256 & \text{si } x = \pm 5/2 \\ -25/256 & \text{si } x = \pm 3/2 \\ 150/256 & \text{si } x = \pm 1/2 \\ 1 & \text{si } x = 0 \\ 0 & \text{si non} \end{cases}$$

La figure 1 représente le graphe de l'interpolante fondamentale de ce processus.

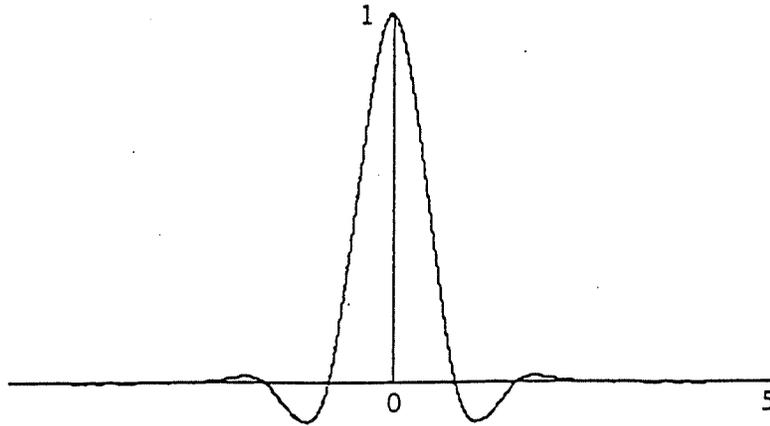


Figure 1: interpolante, polynôme du cinquième degré

Dans cet exemple, nous pouvons démontrer que (voir [3]) la fonction d'interpolation ne sera différentiable que deux fois. La figure 2 présente le schéma d'interpolation. Si x est un point de G_k nous avons:

$$F(x) = (3 F(x-5/2^k) - 25 F(x-3/2^k) + 150 F(x-1/2^k) + 3 F(x+5/2^k) - 25 F(x+3/2^k) + 150 F(x+1/2^k)) / 256$$

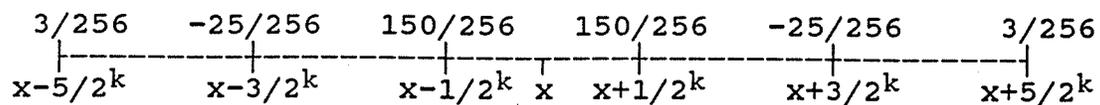


Figure 2: schéma, polynôme du cinquième degré

2.2.0 Schéma asymétrique

Un schéma symétrique est un schéma dont les poids associés aux points $x+(2y-1)/2^k$ et $x-(2y-1)/2^k$ sont égaux ($y \in \{0, 1, \dots, N\}$). Les deux sections précédentes illustrent des exemples de schémas symétriques. Par opposition, un schéma asymétrique est un schéma qui n'est pas symétrique. Les schémas asymétriques sont construits en utilisant, comme nous le faisons avant, un polynôme. Soient N et M deux entiers, la fonction de poids w est choisie de façon à ce que $F(x)$ soit la valeur prise par le polynôme du $M+N-1$ ^{ième} degré passant par les points $(x \in G_k - G_{k-1})$:

$$\{(x-(2y-1)/2^k, F(x-(2y-1)/2^k)) : y = -M+1, -M+2, \dots, N\}.$$

Considérons, comme premiers exemples, les cas où $N=1, M=3$ et où $N=3, M=1$.

2.2.1 Exemple, polynôme du troisième degré

$$\text{Si } N=1, M=3: \quad w(x) = \begin{cases} 1/16 & \text{si } x = -5/2 \\ -5/16 & \text{si } x = -3/2 \\ 15/16 & \text{si } x = -1/2 \\ 5/16 & \text{si } x = 1/2 \\ 1 & \text{si } x = 0 \\ 0 & \text{si non} \end{cases}$$

$$\text{Si } N=3, M=1: \quad w(x) = \begin{cases} 5/16 & \text{si } x = -1/2 \\ 15/16 & \text{si } x = 1/2 \\ -5/16 & \text{si } x = 3/2 \\ 1/16 & \text{si } x = 5/2 \\ 1 & \text{si } x = 0 \\ 0 & \text{si non} \end{cases}$$

Puisque dans chacun des deux cas $C_5(10) \approx 1.554$, le théorème 2 de la section 1.6 nous assure que l'interpolante fondamentale de chacun de ces deux exemples est uniformément continue. La figure 1 présente les schémas d'interpolation de ces deux exemples. Notez la position du point x relativement aux points de la forme $x \pm j/2^k$.

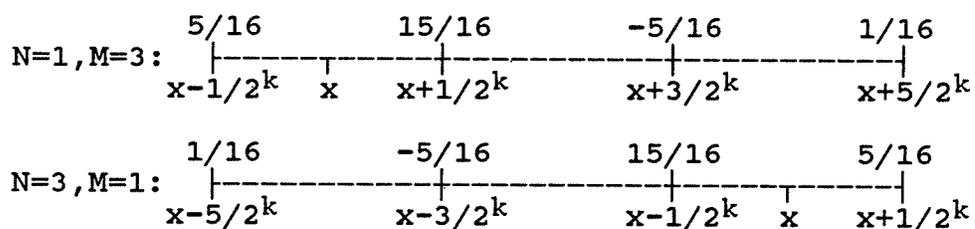


Figure 1: schémas asymétriques, polynôme du troisième degré

2.2.2 Exemple, polynôme du cinquième degré

L'utilisation d'un polynôme du cinquième degré permet d'identifier quatre schémas asymétriques:

$$\text{Si } N=5, M=1: \quad w(x) = \begin{cases} 63/256 & \text{si } x = -1/2 \\ 315/256 & \text{si } x = 1/2 \\ -210/256 & \text{si } x = 3/2 \\ 126/256 & \text{si } x = 5/2 \\ -45/256 & \text{si } x = 7/2 \\ 7/256 & \text{si } x = 9/2 \\ 1 & \text{si } x = 0 \\ 0 & \text{si non} \end{cases}$$

$$\text{Si } N=4, M=2: \quad w(x) = \begin{cases} -7/256 & \text{si } x = -3/2 \\ 105/256 & \text{si } x = -1/2 \\ 210/256 & \text{si } x = 1/2 \\ -70/256 & \text{si } x = 3/2 \\ 21/256 & \text{si } x = 5/2 \\ -3/256 & \text{si } x = 7/2 \\ 1 & \text{si } x = 0 \\ 0 & \text{si non} \end{cases}$$

Tel que dans l'exemple de la section précédente, les fonctions de poids des cas où $N=1, M=5$ et où $N=2, M=4$ sont semblables (elles ne sont pas égales) aux deux fonctions précédentes. Lorsque $N=1, M=5$ ou $N=5, M=1$ le module $C_7(18)$ est à peu près égal à 1.624. De plus, lorsque $N=2, M=4$ ou $N=4, M=2$ le module $C_5(14)$ est à peu près égal à 1.615. Les fonctions d'interpolation associées à ces quatre fonctions de poids sont donc uniformément continues.

2.2.3 Méthode des coefficients indéterminés

La méthode des coefficients indéterminés permet, elle aussi, de trouver les fonctions de poids des sections 2.1.1, 2.1.2, 2.2.1 et 2.2.2. Pour les exemples des sections 2.1.1 et 2.2.1, nous devons résoudre un des sous-systèmes suivants:

$$\begin{bmatrix} . & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & . \\ . & -7 & -5 & -3 & -1 & 1 & 3 & 5 & 7 & . \\ . & 49 & 25 & 9 & 1 & 1 & 9 & 25 & 49 & . \\ . & -343 & -125 & -81 & -1 & 1 & 81 & 125 & 343 & . \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Nous entendons par sous-système, la matrice formée de quatre colonnes consécutives choisies dans la matrice précédente. Si les deux colonnes du centre du sous-système sont les colonnes $(1, -1, 1, -1)^t$ et $(1, 1, 1, 1)^t$ alors la solution sera les poids de la formule symétrique $(-1/16, 9/16, 9/16, -1/16)$. Par contre, si nous décalons le choix des quatre colonnes consécutives nous obtiendrons une formule asymétrique utilisant les quatre points décalés de la même façon par rapport au point d'interpolation.

Lemme 1 Si l'interpolante est continue alors les paramètres a, b, c, d respectent la condition fixée par la première ligne de la matrice du sous-système choisi.

Démonstration: Nous démontrerons ce lemme pour le cas symétrique. Dans les autres cas, il suffit d'utiliser la même technique avec les bonnes équations. La première ligne de tous les sous-systèmes est: $a+b+c+d = 1$. Soit $F(x)$ l'interpolante fondamentale du processus d'interpolation. Soit $\{x_n\}_{n \in \mathbb{N}}$ une suite de points telle que $x_n = 1/2^n$ et soit $h_n = 1/2^n$. Nous avons que pour tout n : $F(x_n) = aF(x_n - 3h_n) + bF(x_n - h_n) + cF(x_n + h_n) + dF(x_n + 3h_n)$, donc: $F(1/2^n) = aF(-1/2^{n-1}) + bF(0) + cF(1/2^{n-1}) + dF(1/2^{n-2})$.

Etant donné la continuité de F , lorsque n tend vers l'infini: $F(0) = a F(0) + b F(0) + c F(0) + d F(0)$.

Puisque $F(0) = 1$ alors: $a+b+c+d = 1$. \square

Lemme 2 Si l'interpolante est différentiable alors les paramètres a, b, c, d respectent la condition fixée par la deuxième ligne de la matrice du sous-système choisi.

Démonstration: Nous démontrerons ce lemme pour le cas symétrique. La condition fixée par la deuxième ligne est: $-3a-b+c+3d = 0$. Puisque F est différentiable alors F est continue et $a+b+c+d = 1$.

Soient $t \in G_m - G_{m-1}$ et $k \in \mathbb{N}$ alors:

$$F(t-1/2^{k+m}) = aF(t-1/2^{k+m-2}) + bF(t-1/2^{k+m-1}) + cF(t) + dF(t+1/2^{k+m-1})$$

ainsi nous avons:

$$F(t) - F(t-1/2^{k+m}) = a (F(t) - F(t-1/2^{k+m-2})) + b (F(t) - F(t-1/2^{k+m-1})) + d (F(t) - F(t+1/2^{k+m-1})) .$$

En divisant cette différence par $1/2^{k+m}$ et en prenant la limite lorsque k tend vers l'infini, nous avons:

$$F'(t) = 4aF'(t) + 2bF'(t) - 2dF'(t) .$$

Choisissons t dans $G_m - G_{m-1}$ tel que $F'(t) \neq 0$, alors:

$4a + 2b - 2d = 1$. Lorsque nous soustrayons cette équation à l'équation: $a+b+c+d = 1$ nous trouvons:

$$-3a - b + c + 3d = 0 . \quad \square$$

Dans le cas des exemples de fonction de poids issues de polynômes du cinquième degré, nous devons résoudre un des sous-systèmes suivants (6 colonnes consécutives):

$$\begin{bmatrix} . & 1 & 1 & 1 & 1 & 1 & 1 & 1 & . \\ . & -7 & -5 & -3 & -1 & 1 & 3 & 5 & 7 & . \\ . & 49 & 25 & 9 & 1 & 1 & 9 & 25 & 49 & . \\ . & -343 & -125 & -81 & -1 & 1 & 81 & 125 & 343 & . \\ . & 2401 & 625 & 243 & 1 & 1 & 243 & 625 & 2401 & . \\ . & -16807 & -3125 & -729 & -1 & 1 & 729 & 3125 & 16807 & . \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Nous pouvons également démontrer que les lignes 1 et 2 représentent des conditions nécessaires pour la continuité et la différentiabilité de l'interpolante fondamentale. Il semble, de plus, que les autres lignes puissent représenter des conditions nécessaires pour l'existence des dérivées d'ordre supérieur.

2.3.0 Ensemble fini et interpolation mixte

Jusqu'à maintenant, nous prenions pour acquis que nous connaissions la fonction à interpoler partout sur Z . Dans bien des applications, cette restriction n'est pas souhaitable. Supposons maintenant que nous ne connaissions une fonction qu'en un nombre fini $M+1$ d'abscisses de l'ensemble $\{0, 1, 2, \dots, M\}$. Le principal intérêt des formules asymétriques est de permettre l'élaboration d'un processus d'interpolation itérative

pour cette suite finie. Considérons les formules issues du polynôme du troisième degré. Le processus se définit comme suit: supposons que nous connaissons la fonction à interpoler pour tous les points de $G_k \cap [0, M]$; nous utilisons les schémas de la section 2.2.1 pour calculer les valeurs de la fonction en $1/2^{k+1}$ et $M-1/2^{k+1}$; de même, nous utilisons le schéma de la section 2.1.1 partout ailleurs dans $(G_{k+1} - G_k) \cap (1/2^{k+1}, M-1/2^{k+1})$. Ce processus se nomme interpolation mixte. Bien que chacun de ces schémas (symétrique et asymétriques) produisent une interpolante fondamentale uniformément continue, rien a priori ne garantit que l'utilisation simultanée de ces schémas produira une interpolation continue. Cette propriété et une condition de différentiabilité sont établies à la section suivante.

2.3.1 Continuité et différentiabilité

L'étude de la continuité et de la différentiabilité est réalisée pour le processus mixte associé au polynôme du troisième degré. Les autres processus mixtes se traitent de la même façon. Établissons tout d'abord quelques lemmes secondaires.

Lemme 1 Soit $\{Y_n\}_{n=0, \dots, M}$ une suite finie telle que:

$$Y_0 = Y_1 = Y_2 = Y_3 = 0.$$

Soit y l'interpolation mixte de cette suite. Si:

$$t \in G_\infty \cap [0, 1] \text{ alors } y(t) = 0.$$

Démonstration: Nous n'avons qu'à montrer que, par induction sur $k \in \mathbb{N}$, si $t \in (G_k \cap [0, 1]) \cup \{1+1/2^k, 1+1/2^{k-1}\}$ alors $y(t) = 0$. \square

Lemme 2 Soit $\{Y_n\}_{n=0, \dots, M}$ une suite finie telle que:

$$Y_M = Y_{M-1} = Y_{M-2} = Y_{M-3} = 0.$$

Soit y l'interpolation mixte de cette suite. Si

$$t \in G_\infty \cap [M-1, M] \text{ alors } y(t) = 0.$$

Démonstration: cette démonstration est similaire à la précédente. \square

Lemme 3 Soit P un polynôme du troisième degré. Soit

$\{Y_n\}_{n=0, \dots, M}$ une suite finie telle que $y_n = P(n)$.

Si y est l'interpolation mixte de cette suite

alors $y(t) = P(t)$ pour tout t dans $[0, M] \cap G_\infty$.

Démonstration: Puisque $y_n = P(n)$ et étant donné le choix des fonctions de poids, nous avons: $y(n/2) = P(n/2)$ pour $n \in \{0, 1, \dots, 2M\}$. Par induction nous avons: $y(n/2^k) = P(n/2^k)$, $n \in \{0, 1, \dots, 2^k M\}$ pour tout k dans \mathbb{N} . \square

Nous pouvons maintenant montrer que le processus mixte est continue.

Théorème 4 Si y est l'interpolation mixte de la suite finie $\{Y_n\}_{n=0, \dots, M}$ alors y est uniformément continue dans $[0, M] \cap G_\infty$.

Démonstration: Découpons l'intervalle $[0, M]$ en trois parts: $[0, 1]$, $[1, M-1]$ et $[M-1, M]$. Nous démontrerons la continuité uniforme de y pour chacun de ces intervalles. Puisqu'au point 1 et $M-1$ les valeurs de y font partie des données initiales alors nous pourrons affirmer que y est uniformément continue sur $[0, M]$. Soit P le polynôme du troisième degré tel que:

$$P(0) = Y_0, P(1) = Y_1, P(2) = Y_2, P(3) = Y_3.$$

Soit $\{g_n\}_{n=0, \dots, M}$ la suite finie telle que:

$$g_n = Y_n - P(n).$$

Nous avons: $g_0 = g_1 = g_2 = g_3 = 0$. Soit g l'interpolation mixte de la suite $\{g_n\}$, si $t \in [0, 1] \cap G_\infty$ alors $g(t) = 0$ et g est uniformément continue dans $[0, 1] \cap G_\infty$. Soit y l'interpolation mixte de la suite $\{Y_n\}$, alors

$$y(t) = g(t) + P(t).$$

Il est clair que la propriété de linéarité s'applique aussi à l'interpolation mixte. En conséquence, $y(t)$ est uniformément continue dans $[0, 1] \cap G_\infty$. De façon similaire y est uniformément continue dans $[M-1, M] \cap G_\infty$. Considérons maintenant l'intervalle $[1, M-1]$. Définissons la suite

infinie $\{g_n\}$ telle que:

$$g_n = 0 \quad \text{si } n < 0 \text{ ou } n > M,$$

$$g_{2n} = Y_n \quad \text{si } 0 \leq n \leq M,$$

$$g_{2n+1} = (-Y_{n-1} + 9Y_n + 9Y_{n+1} - Y_{n+2})/16 \quad \text{si } 0 < n < M-1 .$$

$$g_1 = (5Y_1 + 15Y_2 - 5Y_3 + Y_4)/16$$

$$g_{2M-1} = (Y_{M-3} - 5Y_{M-2} + 15Y_{M-1} + 5Y_M)/16 .$$

Cette suite est la suite des valeurs de y dans $G_1 \cap [0, M]$ et de la valeur 0 partout ailleurs dans G_1 . Si g est l'interpolation mixte de cette suite (telle que définie à la section 2.3.0) alors g est uniformément continue dans G_∞ . De plus nous pouvons démontrer, par induction sur k que si $x \in G_k \cap [1, M-1]$ alors:

$$y(x) = g(x). \quad \square$$

Puisque la fonction d'interpolation est uniformément continue dans $G_\infty \cap [0, M]$, nous parlerons du prolongement continu de y à $\mathbb{R} \cap [0, M]$ plutôt que de la définition de y dans $G_\infty \cap [0, M]$. Ce prolongement permet d'étudier la fonction y et ses dérivées. Nous pouvons trouver dans l'article [3] et à la section 1.6, les conditions qui font qu'un schéma symétrique est différentiable. Le théorème suivant donne une condition qui fait qu'un schéma mixte est également différentiable.

Théorème 5 Si l'interpolante fondamentale d'un schéma symétrique est de classe $C^k(\mathbb{R})$ alors le schéma mixte associé est de classe $C^k(0,M)$.

Démonstration: L'esprit de cette démonstration est identique à l'esprit de la démonstration précédente. En effet, pour démontrer que le schéma mixte est différentiable en un point x , il suffit de définir une suite infinie $\{g_n\}$ avec les valeurs de l'interpolation mixte dans $G_k \cap [0,M]$ (k choisi convenablement) et des 0 partout ailleurs dans $G_k - [0,M]$. Ainsi, sur un intervalle autour de x , l'interpolation itérative de la suite $\{g_n\}$ est égale à l'interpolation mixte. Sur cet intervalle, les propriétés de l'interpolation mixte seront donc celles de l'interpolation itérative. \square

2.4.0 Erreur d'interpolation

Dans cette section nous nous intéresserons à l'étude de l'erreur d'interpolation. Nous verrons que l'ordre de cette erreur est comparable à l'ordre de l'erreur des polynômes de collocation. De plus, nous verrons que dans ce processus itératif, l'erreur commise à une étape est un multiple de l'erreur commise à l'étape précédente. Ce multiple est indépendant de l'étape considérée. Mais avant d'aborder ces problèmes, nous devons établir certaines

propriétés des développements en série de Taylor. Il se peut que ces résultats soient déjà connus. Comme nous n'avons pas trouvé de référence à ce sujet, nous produirons tout de même les démonstrations. Notons que ces résultats restent vrais si nous remplaçons l'intervalle $[0, M]$ par toute la droite \mathbb{R} . Nous nous bornerons cependant à l'étude sur l'intervalle $[0, M]$.

2.4.1 Préliminaire, série de Taylor

Nous étudierons le développement en série de Taylor d'une fonction y de classe C^4 sur un intervalle fermé $[0, M]$. Soit V le sous-ensemble de \mathbb{R}^2 :

$$V = \{ (t, h) : t \in [0, M], t+h \in [0, M] \}.$$

Théorème 1 Si $y \in C^4[0, M]$ alors il existe une fonction $\xi: V \rightarrow [0, M]$ telle que la fonction composée:

$$H(t, h) = y^{iv}(\xi(t, h))h^4/24$$

est continue et:

$$H(t, h) = y(t+h) - y(t) - y'(t)h - y''(t)h^2/2 - y'''(t)h^3/6 .$$

Démonstration: Il existe un nombre ξ_{th} choisi entre t et $t+h$ (pas toujours unique) tel que:

$$y(t+h) = y(t) + y'(t)h + y''(t)h^2/2 + y'''(t)h^3/6 + y^{iv}(\xi_{th})h^4/24$$

$$\text{Soit la fonction: } \xi(t, h) = \begin{cases} \inf\{ \xi_{th} \} & \text{si } h \neq 0 \\ t & \text{si } h = 0 \end{cases} .$$

Supposons que $h \neq 0$. Puisque y^{iv} est continue alors $\{\xi_{th}\}$ est un ensemble fermé et: $\xi(t,h) \in \{\xi_{th}\}$. En conséquence, pour toute valeur de h (y compris 0), l'égalité suivante est vérifiée:

$$y^{iv}(\xi(t,h))h^4/24 =$$

$$y(t+h) - y(t) - y'(t)h - y''(t)h^2/2 - y'''(t)h^3/6.$$

La fonction située à la droite de cette égalité est continue. \square

Théorème 2 $y^{iv}(\xi(t,h))$ est continue.

Démonstration: Si $h \neq 0$ alors $y^{iv}(\xi(t,h)) = 24 H(t,h)/h^4$. Cette fonction est clairement continue. Si $h=0$ alors $y^{iv}(\xi(t,h)) = y^{iv}(t)$. Soit $\{(t_k, h_k)\}_{k=1, \dots, \infty}$ une suite et $t \in [0, M]$ un point tels que:

$$\lim_{k \rightarrow \infty} (t_k, h_k) = (t, 0) .$$

Supposons que: $\lim_{k \rightarrow \infty} y^{iv}(\xi(t_k, h_k)) \neq y^{iv}(\xi(t, 0))$

alors il existe $\epsilon > 0$ et une sous-suite:

$\{(t_{k(m)}, h_{k(m)})\}_{m=1, \dots, \infty}$ de $\{(t_k, h_k)\}$ tels que:

$$|y^{iv}(\xi(t_{k(m)}, h_{k(m)})) - y^{iv}(t)| > \epsilon .$$

puisque $\xi(t_{k(m)}, h_{k(m)}) \in [t_{k(m)} - |h_{k(m)}|, t_{k(m)} + |h_{k(m)}|]$

alors $\lim_{m \rightarrow \infty} \xi(t_{k(m)}, h_{k(m)}) = t$.

La dernière inégalité contredit donc la continuité de y^{iv} . \square

2.4.2 Schéma symétrique

Soit V le sous-ensemble de \mathbb{R}^2 défini comme suit:

$$V = \{ (t, h) : t \in [0, M], t+3h \in [0, M], t-3h \in [0, M] \}$$

Soit $y: [0, M] \rightarrow \mathbb{R}$ une fonction de classe C^4 , définissons l'interpolation $I(t)$, l'erreur d'interpolation $e_k(t)$ et l'erreur d'interpolation locale e_k comme étant:

$$I(t) = (-y(t-3/2^k) + 9y(t-1/2^k) + 9y(t+1/2^k) - y(t+3/2^k)) / 16$$

là où cette équation est bien définie, c'est-à-dire pour les valeurs de t telles que $(t, 1/2^k) \in V$,

$$e_k(t) = \begin{cases} |y(t) - I(t)| & \text{si } t \in G_k - G_{k-1}, (t, 1/2^k) \in V \\ 0 & \text{si non} \end{cases}$$

$$e_k = \max\{ e_k(t), t \in [0, M] \}$$

Lemme 1 Si $y \in C^4([0, M])$, alors il existe quatre fonctions

$\xi_1, \xi_2, \xi_3, \xi_4 : V \rightarrow [0, M]$ telles que:

$$e_k(t) = 3h^4/128 | -9y^{iv}(\xi_1) + y^{iv}(\xi_2) + y^{iv}(\xi_3) - 9y^{iv}(\xi_4) |$$

où $h=1/2^k$, $t \in G_k - G_{k-1}$ et $(t, h) \in V$.

Démonstration: Il suffit de faire la somme des quatre développements en série de Taylor (autour du point t) de: $y(t-3/2^k)$, $y(t-1/2^k)$, $y(t+1/2^k)$ et $y(t+3/2^k)$ dans $I(t)$. \square

Théorème 2 Si $y \in C^4([0, M])$ alors:

$$e_k \leq 15/32 (1/16)^k \sup\{ |y^{iv}(t)| : t \in [0, M] \} .$$

Démonstration: Si $(t, 1/2^k) \in V$, $t \in G_k - G_{k-1}$ alors:

$$e_k(t) =$$

$$3/128 (1/2^k)^4 \left| -9y^{iv}(\xi_1) + y^{iv}(\xi_2) + y^{iv}(\xi_3) - 9y^{iv}(\xi_4) \right|.$$

Puisque $\xi_i \in [t-3/2^k, t+3/2^k]$ et que $[t-3/2^k, t+3/2^k]$ est inclus dans $[0, M]$ alors:

$$e_k \leq 15/32 (1/16)^k \sup\{ |y^{iv}(t)| : t \in [0, M] \} \cdot \square$$

Nous venons de montrer au lemme 1 que l'ordre de l'erreur d'interpolation est égal à $O(h^4)$. Cet ordre est celui de l'erreur d'interpolation des polynômes de collocation du troisième degré. Pour sa part, le théorème 2 fait ressortir le comportement linéaire de l'erreur locale e_k . Le facteur mentionné dans l'introduction est donc égal à 16. Il faut bien comprendre qu'il ne s'agit que de l'erreur locale, c'est-à-dire que l'interpolation I dans G_k est calculée à partir de valeurs exactes de la fonction y dans G_{k-1} . A la section 3.2, nous étudierons l'effet combiné des erreurs locales.

Dans un cadre général, les divers résultats précédents peuvent être réécrits pour les schémas issus de polynômes de degré $2N-1$. Il suffit de choisir une fonction y de classe C^{2N} et de redéfinir convenablement I , V et e_k pour trouver:

Théorème 3 Si $y \in C^{2N}[0,M]$ alors il existe une constante

$C \in \mathbb{R}$, $C > 0$ telle que:

$$e_k \leq C (1/2^{2N})^k \sup\{ |y^{(2N)}(t)| : t \in [0,M] \} .$$

2.4.3 Schéma asymétrique

Les résultats suivants ne sont que la retranscription du lemme 1 et théorème 2 de la section 2.4.2 pour le schéma asymétrique. Une comparaison des deux erreurs locales (symétrique et asymétrique) sera faite dans la prochaine section.

Soit V' le sous-ensemble de \mathbb{R}^2 :

$$V' = \{ (t,h) : t \in [0,M], t+h \in [0,M], t-5h \in [0,M] \}$$

Soit $y: [0,M] \rightarrow \mathbb{R}$ une fonction de classe C^4 , définissons l'interpolation $I'(t)$, l'erreur d'interpolation $e'_k(t)$ et l'erreur d'interpolation locale e'_k comme étant:

$$I'(t) = (5y(t+1/2^k) + 15y(t-1/2^k) - 5y(t-3/2^k) + y(t-5/2^k)) / 16$$

là où cette équation est bien définie, c'est-à-dire pour les valeurs de t telles que $(t, 1/2^k) \in V'$,

$$e'_k(t) = \begin{cases} |y(t) - I'(t)| & \text{si } t \in G_k - G_{k-1}, (t, 1/2^k) \in V' \\ 0 & \text{si non} \end{cases}$$

$$e'_k = \max\{ e'_k(t) : t \in [0,M] \}.$$

Lemme 1 Si $y \in C^4[0, M]$ alors il existe quatre fonctions

$\xi'_1, \xi'_2, \xi'_3, \xi'_4 : V' \rightarrow [0, M]$ telles que:

$$e'_k(t) = h^4/384 \cdot$$

$$|625y^{iv}(\xi'_1) - 405y^{iv}(\xi'_2) + 15y^{iv}(\xi'_3) + 5y^{iv}(\xi'_4)|$$

où $h=1/2^k$, $t \in G_k - G_{k-1}$ et $(t, h) \in V'$.

Théorème 2 Si $y \in C^4([0, M])$ alors:

$$e'_k \leq 175/64 (1/16)^k \sup\{|y^{iv}(t)| : t \in [0, M]\} \cdot$$

Ce lemme et ce théorème traitent le cas du deuxième schéma présenté à la section 2.2.1 ($N=3, M=1$). Des résultats équivalents existent aussi pour le premier schéma de cette section.

2.4.4 Comparaisons des schémas

Choisissons l'un ou l'autre des schémas asymétriques de la section 2.2.1 et le schéma symétrique de la section 2.1.1. Si h est assez petit, c'est-à-dire, puisque $h=1/2^k$, si k est assez grand, alors:

$$y^{iv}(\xi_i) \approx y^{iv}(t)$$

$$y^{iv}(\xi'_i) \approx y^{iv}(t) \cdot$$

De plus: $e_k(t) \approx 3/8 h^4 |y^{iv}(t)|$

$$e'_k(t) \approx 5/8 h^4 |y^{iv}(t)|$$

et: $e_k(t) \approx 3/5 e'_k(t) \cdot$

L'erreur locale d'interpolation d'un des schémas asymétriques est donc à peu près $5/3$ fois supérieure à l'erreur commise par le schéma symétrique. Ce facteur jouera un grand rôle dans l'étude des exemples de compression de données.

2.4.5 Fonction polynômiale

Revenons au schéma symétrique issu du polynôme du troisième degré et considérons le cas où y est un polynôme de degré n plus grand que trois. Sans perdre de généralité, les termes de degré inférieur ou égal à 3 peuvent être omis, l'interpolation étant exacte pour ces degrés.

Soit y un polynôme de la forme:

$$y(t) = a_4 t^4 + a_5 t^5 + \dots + a_n t^n$$

La fonction d'erreur $e_k(t)$ est donnée par ($t \in G_k - G_{k-1}$, $(t, 1/2^k) \in V$):

$$e_k(t) = \left| a_4 (16t^4 - (-(t-3h)^4 + 9(t-h)^4 + 9(t+h)^4 - (t+3h)^4)) \right. \\ \left. + a_5 (16t^5 - (-(t-3h)^5 + 9(t-h)^5 + 9(t+h)^5 - (t+3h)^5)) + \dots \right. \\ \left. + a_n (16t^n - (-(t-3h)^n + 9(t-h)^n + 9(t+h)^n - (t+3h)^n)) \right| / 16$$

où $h = 1/2^k$.

Considérons le terme (pour $i = 4, 5, \dots, n$):

$$16t^i - (-(t-3h)^{i+9}(t-h)^{i+9}(t+h)^i - (t+3h)^i) =$$

$$16t^i + \left(\sum_j C_j^i t^{i-j} (-3h)^j \right) - 9 \left(\sum_j C_j^i t^{i-j} (-h)^j \right)$$

$$- 9 \left(\sum_j C_j^i t^{i-j} (h)^j \right) + \left(\sum_j C_j^i t^{i-j} (3h)^j \right)$$

où $C_j^i = i! / (j!(i-j)!)$.

Ainsi: $16t^i - (-(t-3h)^{i+9}(t-h)^{i+9}(t+h)^i - (t+3h)^i)$

$$= 2h^4 \sum_{j=2}^{[i/2]} C_{2j}^i t^{i-2j} h^{2j-4} (9j-9)$$

où $[x] =$ partie entière de x .

Supposons que: $P(t,h) = | a_4 (144) + a_5 5t (144)$
 $+ a_6 (15 t^2 (144) + h^2 (1440)) +$
 $\dots + a_n 2 \left(\sum_{j=2}^{[n/2]} C_{2j}^n t^{n-2j} h^{2j-4} (9j-9) \right) |$

Ainsi: $e_k(t) = h^4 P(t,h)$.

Lemme 1 Si y est un polynôme du quatrième degré alors:

$$e_k/e_{k+1} = 16.$$

Démonstration: Il suffit de voir que: $P(t,h) = 144 |a_4|$ et $h=1/2^k$. \square

2.4.6 Quotient des erreurs locales

Le quotient e_k/e_{k+1} est proche de la valeur 16 pour une classe de fonctions beaucoup plus large que l'ensemble des polynômes du quatrième degré. Dans les faits sa limite est égale à 16 pour toute fonction de classe C^4 .

Soit $E:V \rightarrow \mathbb{R}$ la fonction définie comme suit:

$E(t, h) = 3/128 \left| -9y^{iv}(\xi_1) + y^{iv}(\xi_2) + y^{iv}(\xi_3) - 9y^{iv}(\xi_4) \right|$
 où les fonctions ξ_i sont choisies conformément au lemme 1 de la section 2.4.2 . Définissons une suite $\{t_k\}_{k=1, \dots, \infty}$ dans \mathbb{R} telle que $(t_k, 1/2^k)$ soit un point de V , $e_k(t_k) = e_k$ et $t_k \in G_k - G_{k-1}$. Définissons aussi une constante:

$$K = \sup \{ E(t, 0) : (t, 0) \in V \} .$$

Lemme 1 $E(t, h)$ est continue dans V .

Démonstration: Théorème 2 section 2.4.1 et lemme 1 section 2.4.2 . \square

Lemme 2 $E(t_k, 1/2^k) =$
 $\max \{ E(t, 1/2^k) : (t, 1/2^k) \in V, t \in G_k - G_{k-1} \}$

Démonstration: $E(t, 1/2^k) = 2^{4k} e_k(t)$ lorsque $t \in G_k - G_{k-1}$ et $(t, 1/2^k) \in V$. \square

Lemme 3 $\lim_{k \rightarrow \infty} E(t_k, 1/2^k) = K$.

Démonstration: Supposons que $\lim E(t_k, 1/2^k) \neq K$ alors il existe une sous-suite $\{(t_{k(n)}, 1/2^{k(n)})\}_{n=1, \dots, \infty}$ choisie dans $\{(t_k, 1/2^k)\}_{k=1, \dots, \infty}$ et un nombre réel $\varepsilon > 0$ tels que: $| E(t_{k(n)}, 1/2^{k(n)}) - K | > \varepsilon$.

Soit $t' \in [0, M]$ un point où $E(t', 0) = K$. Choisissons δ_ε tel que si $t \in B(t', \delta_\varepsilon)$ alors: $| E(t, 0) - E(t', 0) | < \varepsilon/3$.

De plus, choisissons N assez grand pour que: si $n \geq N$ alors $1/2^{k(n)} < \delta_\varepsilon$ et

$$\max\{|E(t, 1/2^{k(n)}) - E(t, 0)| : (t, 1/2^{k(n)}) \in V, t \in G_{k(n)} - G_{k(n)-1}\} < \varepsilon/3$$

Pour un n choisi plus grand que N nous distinguons deux cas:

- 1- $E(t_{k(n)}, 1/2^{k(n)}) > M + \varepsilon$
- 2- $E(t_{k(n)}, 1/2^{k(n)}) < M - \varepsilon$.

Dans le premier cas: $|E(t_{k(n)}, 1/2^{k(n)}) - E(t_{k(n)}, 0)| \leq \max\{|E(t, 1/2^{k(n)}) - E(t, 0)| : t \in U_{k(n)}\} \leq \varepsilon/3$.

En conséquence: $E(t_{k(n)}, 0) > K$, Ceci contredit la définition de K .

Dans le second cas, puisque $1/2^{k(n)} < \delta_\varepsilon$ alors il existe un point $t'_{k(n)}$ tel que:

$$t'_{k(n)} \in (G_{k(n)} - G_{k(n)-1}) \cap B(t', \delta_\varepsilon) .$$

Notons que si $(t'_{k(n)}, 1/2^{k(n)}) \notin V$ nous devons choisir une valeur plus grande pour N .

$$\begin{aligned} \text{Nous avons: } |E(t', 0) - E(t'_{k(n)}, 1/2^{k(n)})| &\leq \\ |E(t', 0) - E(t'_{k(n)}, 0)| + |E(t'_{k(n)}, 0) - E(t'_{k(n)}, 1/2^{k(n)})| &\leq \\ &2\varepsilon/3 \end{aligned}$$

En conséquence: $E(t'_{k(n)}, 1/2^{k(n)}) > E(t_{k(n)}, 1/2^{k(n)})$.

Ceci constitue une contradiction. \square

Théorème 4 Si $y \in C^4[0, M]$ et $y^{iv} \neq 0$ dans $[0, M]$ alors:

$$\lim_{k \rightarrow \infty} e_k / e_{k+1} = 16 \quad .$$

Démonstration: Puisque $y^{iv} \neq 0$ alors il existe un point t tel que $E(t, 0) \neq 0$ et ainsi $K \neq 0$. De plus, puisque E est continue alors il existe un entier N tel que pour tout $k \geq N$ nous avons $e_k \neq 0$. Pour tout $k \geq N$ nous pouvons écrire:

$$e_k / e_{k+1} = 16 E(t_k, 1/2^k) / E(t_{k+1}, 1/2^{k+1}) \quad .$$

Puisque: $\lim_{k \rightarrow \infty} E(t_k, 1/2^k) = K \neq 0$ nous avons:

$$\lim_{k \rightarrow \infty} e_k / e_{k+1} = 16. \quad \square$$

Dans un contexte plus large, supposons que nous étudions un schéma issu d'un polynôme du $2N-1$ ème degré, nous pouvons démontrer que:

Théorème 5 Si $y \in C^{2N}([0, M])$ et $y^{(2N)} \neq 0$ dans $[0, M]$ alors:

$$\lim_{k \rightarrow \infty} e_k / e_{k+1} = 2^{2N} \quad .$$

CHAPITRE 3, COMPRESSION DE DONNÉES

Nous venons de montrer que le comportement de l'erreur d'interpolation était facilement prévisible. Ce comportement et la nature itérative du processus nous laissent penser que l'interpolation mixte est toute indiquée pour développer un processus de compression de données.

3.1 Définition du processus

Nous allons définir le processus de compression et de reconstruction en utilisant les schémas de l'interpolation mixte issus du polynôme du troisième degré. Soit $S = \{Y_i\}_{i=0, \dots, M}$ une suite finie. Supposons que M est une puissance de deux, c'est-à-dire qu'il existe un entier positif n tel que $M = 2^n$. Définissons une suite d'ensembles $\{S_k\}$ de la façon suivante:

$$S_0 = S$$

$$S_1 = \{Y_{2i}\}_{i=0, 1, \dots, M/2}$$

$$S_2 = \{Y_{4i}\}_{i=0, 1, \dots, M/4}$$

...

$\{S_k\}$ est la suite des ensembles comprimés issus de S . Choisissons un entier i tel que $M/2^i$ est plus grand que 3. Ce nombre i sera la profondeur maximale de la compression. Supposons maintenant que S_0 est l'ensemble

des valeurs définies par une fonction y sur $G_i \cap [0, M/2^i]$. Par conséquent, S_i est l'ensemble des valeurs définies par y sur $G_0 \cap [0, M/2^i]$. Avec une étape du processus d'interpolation mixte nous pouvons reconstruire la suite S_{k-1} à partir de la suite S_k . Nous pouvons, avec i étapes du processus mixte, reconstruire la suite S_0 à partir de la suite S_i . Ces deux suites reconstruites seront notées S'_{k-1} et S''_0 . Nous sommes intéressés à connaître la différence entre les suites S_{k-1} , S'_{k-1} et S_0 , S''_0 . La différence entre S_{k-1} et S'_{k-1} a déjà été définie comme étant l'erreur locale d'interpolation. La différence entre les suites S_0 et S''_0 est appelée l'erreur globale d'interpolation. Cette erreur globale servira à évaluer l'effet combiné des erreurs locales. Définissons l'interpolation mixte I comme étant ($t \in G_k - G_{k-1}$, $t \in [0, M/2^i]$):

$$I(t) = \begin{cases} (5y(t-1/2^k) + 15y(t+1/2^k) - 5y(t+3/2^k) + y(t+5/2^k))/16 & \text{si } t=1/2^k \\ (5y(t+1/2^k) + 15y(t-1/2^k) - 5y(t-3/2^k) + y(t-5/2^k))/16 & \text{si } t=M/2^i-1/2^k \\ (-y(t-3/2^k) + 9y(t-1/2^k) + 9y(t+1/2^k) - y(t+1/2^k))/16 & \text{si non} \end{cases}$$

L'erreur d'interpolation locale:

$$e_k(t) = \begin{cases} |y(t) - I(t)| & \text{si } t \in (G_k - G_{k-1}) \cap [0, M/2^i] \\ 0 & \text{si non} \end{cases}$$

$$e_k = \max \{ e_k(t) : t \in [0, M/2^i] \}$$

Supposons que $Y(t)$ est l'interpolation mixte bâtie à partir de la suite S_i . Définissons l'erreur globale d'interpolation comme étant:

$$E(t) = \begin{cases} | Y(t) - y(t) | & \text{si } t \in G_i \cap [0, M/2^i] \\ 0 & \text{si non} \end{cases}$$

$$E_k = \max \{ E(t) : t \in G_k \} .$$

3.2 Etude de l'erreur globale

L'étude de l'erreur locale a déjà été faite à la section 2.4. Dans le cas présent, il faut en plus faire attention aux points de la forme $1/2^k$ et $M/2^i - 1/2^k$. En ces points, le calcul de l'interpolation se fait avec un schéma asymétrique. Il faut donc tenir compte du facteur $5/3$ dont nous discutons à la section 2.4.4. Nous allons maintenant étudier la relation qui existe entre l'erreur globale et l'erreur locale. Le théorème 1 et son corollaire 2 établissent cette relation. Le corollaire 3 et la remarque qui le suit permettent de borner l'erreur globale.

Théorème 1 $E_k \leq e_k + (13/8) E_{k-1}$

Démonstration: Soit:

$$i(Y_1, Y_2, Y_3, Y_4) = (-Y_1 + 9Y_2 + 9Y_3 - Y_4)/16$$

$$\text{et: } e_k(t, Y_1, Y_2, Y_3, Y_4) = | Y(t) - i(Y_1, Y_2, Y_3, Y_4) |$$

$$\text{Nous avons: } e_k(t, Y_1 + \Delta Y_1, Y_2 + \Delta Y_2, Y_3 + \Delta Y_3, Y_4 + \Delta Y_4) \leq$$

$$e_k(t, Y_1, Y_2, Y_3, Y_4) + 5/4 \max\{|\Delta Y_i|\}$$

Si $(t, 1/2^k) \in V$, supposons que $Y_n = Y(t + (2n-5)/2^k)$ et

$\Delta Y_n = Y(t + (2n-5)/2^k) - Y_n$ alors:

$$E(t) \leq e_k(t) + (5/4) E_{k-1}.$$

Pour les points t de la forme $1/2^k$ et $M/2^i - 1/2^k$, il suffit de changer la définition de: $i(t)$, Y_i , ΔY_i et montrer que: $E(t) \leq e_k(t) + (13/8) E_{k-1}$.

Il ne reste qu'à trouver le suprémum de chaque coté de cette inégalité. \square

Corollaire 2 $E_i \leq e_i + (13/8) e_{i-1} + \dots + (13/8)^{i-1} e_1$

Corollaire 3

$$E_i \leq 7/64 \cdot 1/16^i (26^i - 1) \sup\{|Y^{iv}(t)| : t \in [0, M/2^i]\}.$$

Démonstration: Théorème 1 et lemme 2 section 2.4.3. \square

Notons que si nous ne considérons que le schéma symétrique, alors le facteur $7/64$ du corollaire 3 est remplacé par le facteur $3/160$. Dans certaines applications il se peut que la fonction y ne soit pas connue. Il sera

donc impossible, à première vue, de borner l'erreur globale comme le fait le corollaire 3. En combinant le théorème 4 de la section 2.4.6 au corollaire 2, nous trouvons une approximation B_i de la borne E_i : $B_i = (26^i - 1)/16^{i-1} e_1/25$. Cependant, pour que ce nombre soit une bonne approximation il faut que les erreurs locales soient toutes issues de points calculés avec un seul type de schéma (symétrique ou asymétrique). Si tel n'est pas le cas, il faut tenir compte du facteur $5/3$ comme nous le verrons à la section 3.4.

3.3 Remarque sur le nombre de points

A la section 3.1, nous supposons que l'entier M était égal à une puissance de 2. Supposons maintenant qu'il existe un entier k tel que $M/2^k$ ne soit pas divisible par 2 ($M/2^k$ est impair). Puisque cette valeur est impaire, à la $(k+1)^{\text{ième}}$ étape, nous ne pouvons poursuivre le processus de compression. En effet, si nous poursuivons ce processus, nous allons éliminer le point y_M lors du passage de la suite S_k à S_{k+1} . Ce point se situant en bout de suite, nous ne pourrons le retrouver à l'aide d'un des schémas déjà connus.

Une première réponse à ce problème est de développer un schéma asymétrique n'utilisant que les valeurs de y en: $t-7/2^k$, $t-5/2^k$, $t-3/2^k$, $t-1/2^k$ pour trouver la valeur en t . Les poids de ce schéma sont respectivement: $-5/16$, $21/16$, $-35/16$, $35/16$. Si nous calculons le module $C_k(14)$ du théorème 2 de la section 1.6 nous trouvons:

$$\begin{aligned} C_1(14) &= 12.00 \\ C_3(14) &= 43.56 \\ C_5(14) &= 92.37 \\ C_7(14) &= 261.0 \end{aligned}$$

Cette suite semble diverger et ce schéma produit probablement une fonction d'interpolation discontinue. Même si tel n'est pas le cas, la somme (en valeur absolue) des poids de ce schéma est grande. L'erreur, ou plutôt la borne de l'erreur locale, va croître rapidement.

Une autre possibilité serait d'ajouter un point à la fin de la suite S_k . Ce point se calcule à l'aide du polynôme du troisième degré passant par les quatre derniers points de la suite S_k . L'expérience nous montre que cette solution fonctionne bien.

3.4 Exemples

Nous allons maintenant présenter les exemples de compression-reconstruction que nous avons effectués. Remarquons premièrement que tous les calculs ont été réalisés sur ordinateur en utilisant une arithmétique d'à peu près vingt décimales, plus exactement des variables réelles de dix octets. Les valeurs prises pour former l'ensemble de départ S varient toutes entre 0 et 32. Nous estimons donc que tous les résultats présentés ici, entre autres les erreurs locales, sont fiables. En effet, nous estimons que compte tenu des nombres qui forment S , toute erreur locale supérieure à 10^{-16} ne sera pas (trop) entachée de l'erreur due à l'arithmétique de l'ordinateur.

Le premier exemple n'est présenté que pour constater ce que prédit la théorie sur l'erreur d'interpolation et pour décrire le format de présentation des résultats. Nous allons nous attarder davantage sur d'autres exemples plus significatifs.

Définissons l'ensemble de départ S comme suit:

$$S = \{ (2n/4096)^4, n \in \{0, 1, \dots, 4096\} \} .$$

Le tableau 1 présente les résultats obtenus lors de la compression-reconstruction de la suite S . La première colonne contient l'indice du sous-groupe G_k . La seconde

colonne contient l'indice de profondeur de compression ou l'indice de la suite S_k . La troisième colonne nous donne le nombre de points que contient la suite S_k . La quatrième colonne présente les différentes erreurs locales. La dernière colonne contient les quotients des erreurs locales. Finalement, nous trouvons au bas du tableau, l'indice du niveau maximal atteint. Ce niveau maximal est choisi en fonction de l'erreur locale. Dans cet exemple, nous avons arrêté le processus dès que l'erreur locale dépassait la valeur 10^{-6} . Nous trouvons également, au bas du tableau, l'erreur globale telle que calculée par l'ordinateur (voir la définition à la section 3.1). Remarquons, pour éviter une source de confusion, que les indices de G_k et S_k agissent en sens inverse; l'indice de e_k pour sa part suit l'indice de G_k conformément à la définition de e_k .

G_k	S_k	Points	e_k	e_k/e_{k+1}
6	0	4097	8.52×10^{-13}	
5	1	2049	1.36×10^{-11}	16
4	2	1025	2.18×10^{-10}	16
3	3	513	3.49×10^{-09}	16
2	4	257	5.58×10^{-08}	16
1	5	129	8.93×10^{-07}	16
0	6	65		

Arrêt de la compression au niveau: 6

Erreur globale: 9.53×10^{-07}

Tableau 1: compression, polynôme du quatrième degré

Nous constatons premièrement que les quotients sont tous égaux à la valeur 16. Le lemme 1 de la section 2.4.5 nous avait prédit ce résultat. La borne de l'erreur globale donnée par le corollaire 2 de la section 3.2 et l'estimation de cette borne sont toutes deux égales à 1.05×10^{-05} . Cette égalité entre l'estimation de la borne et la borne elle-même vient du fait que les quotients, dans cet exemple, sont tous égaux à 16. Nous remarquons que cette borne est à peu près dix fois supérieure à la vraie valeur de l'erreur globale $E_6 = 9.53 \times 10^{-07}$. Finalement nous constatons que le facteur de compression, défini comme étant le quotient du nombre de points des suites S_0 et S_6 , est égal à 63.03. Nous avons donc réussi à comprimer cette suite d'un facteur 63 tout en conservant une erreur, lors de la reconstruction, égale théoriquement à 10^{-05} et de façon pratique à 10^{-06} .

Dans l'exemple suivant, nous avons à peine augmenté la difficulté en choisissant pour fonction de départ un polynôme du cinquième degré. Soit S l'ensemble de départ défini comme suit:

$$S_0 = \{Y_n\} = \{ (2n/4096)^5, n \in \{0, 1, \dots, 4096\} \} .$$

Le tableau 2 présente les résultats obtenus à partir de cette suite.

G_k	S_k	Points	e_k	e_k/e_{k+1}
5	0	4097	8.51×10^{-12}	
4	1	2049	1.36×10^{-10}	16
3	2	1025	2.17×10^{-09}	16
2	3	513	3.47×10^{-08}	16
1	4	257	5.52×10^{-07}	15.9
0	5	129		

Arrêt de la compression au niveau: 5

Facteur de compression: 31.76

Erreur globale: 5.89×10^{-07}

Tableau 2: compression, polynôme du cinquième degré

Les quotients des erreurs locales cessent d'être constants tout en ne s'éloignant pas de la valeur 16. Cela s'explique par le fait que la fonction $P(t,h)$, définie à la section 2.4.5, n'est elle-même plus une constante. Il s'agit en fait d'un polynôme du premier degré. La figure 3 représente le graphe de la fonction $e_k(t)$ pour diverses valeurs de k . Nous savons que si $t \in G_k - G_{k-1}$ alors $e_k(t) = P(t, 1/2^k)/16^k$. Remarquons que les échelles sont différentes pour chacune des droites représentées sur ce graphique. Avec une échelle fixe, nous n'aurions pas pu tracer toutes ces droites sur une seule feuille. L'alignement des points où e_k est non-nulle confirme le fait que $P(t,h)$ est bien un polynôme de degré 1.

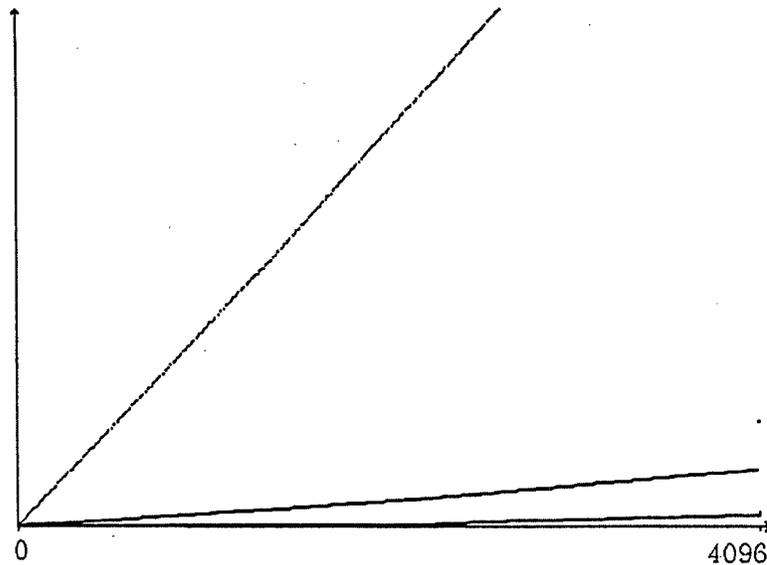


Figure 3: Fonction $e_k(t)$

Ces fonctions $e_k(t)$ ont été calculées directement lors de l'exécution du processus de compression-reconstruction et non avec l'équation de $P(t,h)$. C'est pourquoi le dernier point de chacune des droites est légèrement plus haut (sur le graphe) que le tracé défini par les autres. Souvenons-nous qu'en ces points nous avons utilisé un schéma asymétrique. Il faut donc tenir compte du facteur $5/3$ existant entre les erreurs locales $e_k(t)$ et $e'_k(t)$ tel qu'identifié à la section 2.4.4.

La borne de l'erreur globale est égale à 4.004×10^{-06} . Pour sa part, l'estimation de cette borne est égale à 4.003×10^{-06} . L'estimation est légèrement inférieure à la valeur exacte de la borne car les quotients des erreurs locales sont eux-mêmes légèrement inférieurs à la valeur 16. Dans cet exemple aussi, la valeur exacte de l'erreur globale est 10 fois inférieure à la borne précédente.

Etudions un nouvel exemple. Une fonction sinusoïdale a été choisie pour former la suite initiale S .

$$S = \{ \sin(2\pi n/4096) : n=0,1,\dots,4096 \} .$$

Le tableau 4 présente les résultats obtenus avec cette suite.

G_k	S_k	Points	e_k	e_k/e_{k+1}
5	0	4097	2.074×10^{-12}	
4	1	2049	3.319×10^{-11}	16
3	2	1025	5.310×10^{-10}	16
2	3	513	8.496×10^{-09}	16
1	4	257	1.359×10^{-07}	15.99
0	5	129		

Arrêt de la compression au niveau: 5

Facteur de compression: 31.76

Erreur globale: 1.359×10^{-07}

Tableau 4: compression, fonction sinus

Nous arrivons encore aux mêmes conclusions sur le quotient des erreurs et les valeurs de la borne de l'erreur globale et de son approximation. En effet cette borne et son approximation sont égales à: 9.86×10^{-07} . La borne est 10 fois supérieure à la valeur exacte de l'erreur globale. De plus, l'erreur globale est égale à l'erreur locale e_1 . En regardant de plus près les valeurs de la suite S''_0 nous constatons que le maximum de l'erreur d'interpolation est atteint dès la première étape du processus de reconstruction. Aux étapes suivantes, même si les valeurs utilisées sont entachées d'erreurs, le processus semble "s'auto-corriger". La figure 5 présente les graphes des fonctions $e_k(t)$ pour quelques valeurs de k .

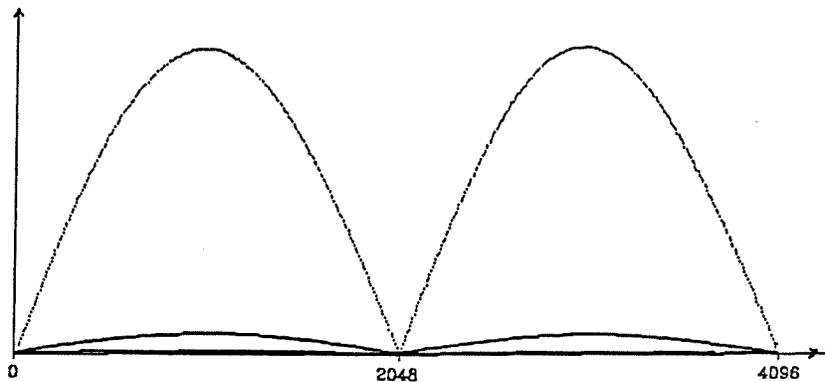


Figure 5: $e_k(t)$, fonction sinus

Les échelles sont encore différentes pour les diverses valeurs de k . Grâce aux lemmes 1 des sections 2.4.2 et 2.4.3, nous savons que la fonction $e_k(t)$, là où elle est non-nulle, doit être la valeur absolue d'une somme de sinus.

Revenons au polynôme du cinquième degré, mais cette fois avec un nombre de points dans S qui n'est pas une puissance de 2 plus 1 (2^n+1). Définissons S de la façon suivante:

$$S = \{ (i/5999)^5, i=0,1,\dots,5999 \} .$$

Le tableau 6 présente les résultats obtenus avec cet ensemble.

G_k	S_k	Points	e_k	e_k/e_{k+1}
10	0	6001	3.47×10^{-14}	
9	1	3001	8.96×10^{-13}	25.8
8	2	1501	1.48×10^{-11}	16.5
7	3	751	2.36×10^{-10}	16.0
6	4	377	2.25×10^{-09}	95.3
5	5	189	5.81×10^{-08}	25.8
4	6	95	9.45×10^{-07}	16.3
3	7	49	8.55×10^{-06}	90.5
2	8	25	2.19×10^{-04}	25.7
1	9	13	3.18×10^{-03}	14.5
0	10	7		

Erreur globale: 3.99×10^{-02}

Tableau 6: compression, polynôme du cinquième degré

Le quotient e_k/e_{k+1} est parfois très loin de la valeur 16. Pour les valeurs de k (indice de G_k) choisies dans l'ensemble: $\{3,6,10\}$ un point a été ajouté à la suite S_{10-k} , le nombre de points de S_{10-k} étant pair. Cet ajout se fait sentir, dans la colonne 5 du tableau, par un quotient e_k/e_{k+1} près de la valeur 9.6. En fait, tel que l'illustre le graphe de la fonction $e_k(t)$ de la figure 7, pour les valeurs de k (indice de G_k) choisies dans l'ensemble $\{0,1,2,4,5,7,8,9\}$, le maximum de l'erreur locale sur G_k est atteint en un point calculé avec un schéma asymétrique. Pour les valeurs de k choisies dans l'ensemble $\{3,6,10\}$, ce maximum est atteint en un point calculé avec le schéma symétrique.

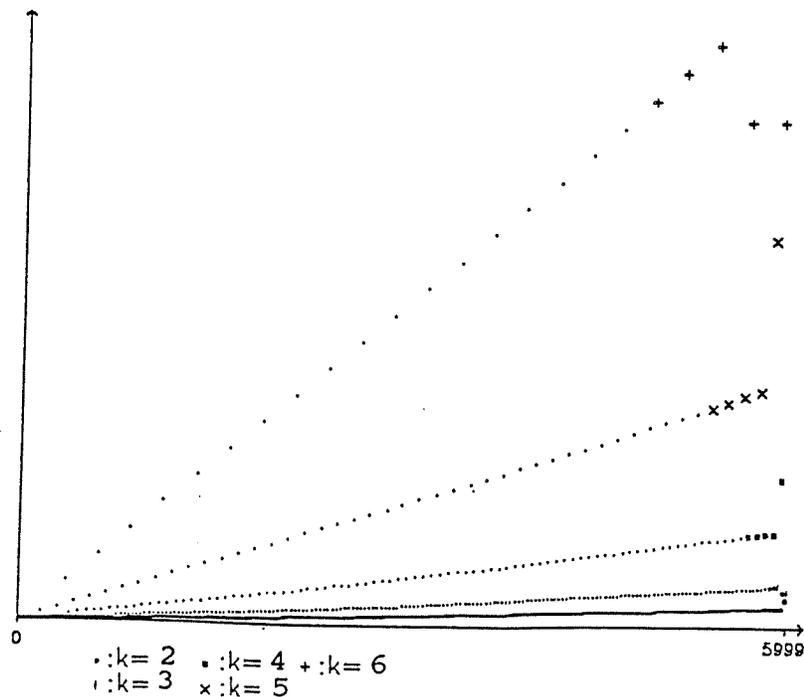


Figure 7: $e_k(t)$, polynôme du cinquième degré

Un facteur $5/3$ agit entre les erreurs des deux types de schémas. Si k appartient à l'ensemble $\{3,6,10\}$ alors:

$$e_k/e_{k+1} \approx 3/5 * 16 = 9.6 \quad .$$

Par contre si k appartient à l'ensemble $\{2,5,9\}$ alors:

$$e_k/e_{k+1} \approx 5/3 * 16 = 26.7 \quad .$$

Ceci explique toutes les différentes valeurs des quotients e_k/e_{k+1} .

Puisque les quotients ne sont plus tous près de la valeur 16, nous ne pouvons plus utiliser l'approximation de la borne de la section 3.2. Mais, chacun des quotients peut être approché ou majoré par un des nombres suivants: 16, $3/5 \times 16$, $5/3 \times 16$. De plus, chacun des quotients e_k/e_1 peut être approché ou majoré par le facteur $5/3 \times 16^{k-1}$ et ce, indépendamment de la suite initiale S . En conséquence, nous pouvons utiliser le nombre $5/3 \times B_i$ comme borne approximative de E_i . Nous avons donc: $5/3 \times B_{10} = 4.35 \times 10^{-01}$, l'erreur globale réelle est encore 10 fois inférieure à cette borne.

CONCLUSION

A titre de conclusion, nous allons faire deux commentaires. Le premier concerne les résultats présentés sur l'interpolation itérative. Connaissant la définition de l'interpolation itérative dans l'espace (voir [1]), nous pouvons généraliser ces résultats et trouver ainsi une famille de schémas d'interpolation qui permet de définir l'interpolation mixte dans l'espace. Les résultats traitant de l'erreur d'interpolation se généralisent aussi directement.

Le second commentaire concerne le processus de compression de données. Nous venons d'étudier un processus efficace et sécuritaire. Nous disons efficace car il ne faut que quelques opérations arithmétiques pour accomplir chaque étape de reconstruction. Le processus d'interpolation mixte est simple et facile à implémenter. Nous disons sécuritaire car nous possédons une borne de l'erreur globale et une approximation de cette borne. Cette approximation peut toujours être évaluée puisqu'elle ne fait intervenir, à la base, que les entrées de la suite initiale S. Nous pouvons donc assez facilement choisir la profondeur maximale de compression compte tenu de l'erreur globale que nous sommes prêts à tolérer.

BIBLIOGRAPHIE

- 1 Deslauriers, G., Dubois, J., Dubuc, S. 1989
Multidimensional iterative interpolation. Rapport
technique, Département de math. et de stat., Université
de Montréal, Montréal.
- 2 Deslauriers, G., Dubuc, S., 1987 Interpolation
dyadique. Fractals, Dimensions non entières et
applications. Paris: Masson , (44-55).
- 3 Deslauriers, G., Dubuc, S., 1989 Symetric iterative
interpolation. Constr. Approx. Vol 5, (49-68).
- 4 Dubuc, S., 1986 Interpolation through an iterative
scheme, J. of Math., Anal. and Appl. 114, (185-204).
- 5 Dugundji, J., 1973, Topology, Allyn and Bacon, inc.,
Boston (302)

ÉCOLE POLYTECHNIQUE DE MONTRÉAL



3 9334 00289654 4