

# Chapitre II

## Interpolation et Approximation

Le problème de l'*interpolation* consiste à chercher des fonctions “simples” (polynômes, polynômes par morceaux, polynômes trigonométriques) passant par des points donnés

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n), \quad (0.1)$$

c.-à-d., on cherche  $p(x)$  avec  $p(x_i) = y_i$  pour  $i = 0, 1, \dots, n$ . Si les valeurs de  $y_i$  satisfont  $y_i = f(x_i)$  où  $f(x)$  est une fonction donnée, il est intéressant d'étudier l'erreur de l'*approximation*

$$f(x) - p(x) = ? \quad (0.2)$$

### Bibliographie sur ce chapitre

J.H. Ahlberg, E.N. Nilson & J.L. Walsh (1967): *The Theory of Splines and Their Applications*. Academic Press, New York. [MA 65/4]

C. de Boor (1978): *A Practical Guide to Splines*. Springer-Verlag. [MA 65/141]

G.D. Knott (2000): *Interpolating Cubic Splines*. Birkhäuser. [MA 65/431]

H.J. Nussbaumer (1981): *Fast Fourier Transform and Convolution Algorithms*. Springer-Verlag.

H. Späth (1995): *One Dimensional Spline Interpolation*. AK Peters. [MA 65/362]

## II.1 Différences divisées et formule de Newton

Étant donnés les  $n + 1$  points (0.1) où les  $x_i$  sont distincts mais pas nécessairement ordonnés, on cherche un polynôme  $p(x)$  de degré  $n$  qui satisfasse

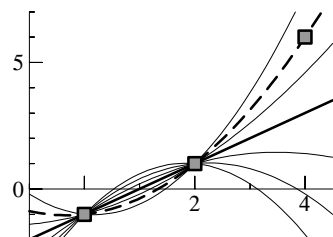
$$p(x_i) = y_i \quad \text{pour} \quad i = 0, 1, \dots, n. \quad (1.1)$$

Cas  $n = 1$ . Le polynôme de degré 1 (une droite) qui passe par  $(x_0, y_0)$ ,  $(x_1, y_1)$  est donné par

$$p(x) = y_0 + (x - x_0) \frac{y_1 - y_0}{x_1 - x_0}. \quad (1.2)$$

Cas  $n = 2$ . Pour obtenir un polynôme de degré 2 (une parabole) qui passe par les trois points  $(x_0, y_0)$ ,  $(x_1, y_1)$ ,  $(x_2, y_2)$ , on ajoute un terme de correction (de degré 2) à la formule (1.2). Comme ce terme doit être zéro aux points  $x_0$  et  $x_1$ , on a nécessairement

$$p(x) = y_0 + (x - x_0) \frac{y_1 - y_0}{x_1 - x_0} + a \cdot (x - x_0)(x - x_1). \quad (1.3)$$



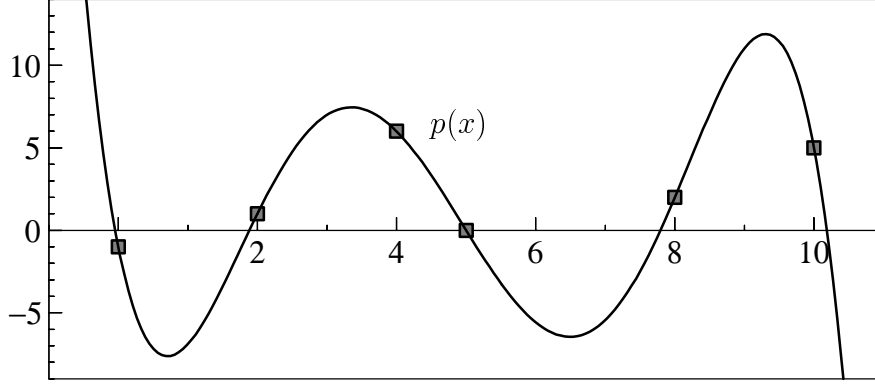


FIG. II.1: Polynôme d'interpolation de degré 5

Le coefficient  $a$  est déterminé par  $p(x_2) = y_2$ . Un calcul simple (soustraire  $p(x_1)$  de  $p(x_2)$  et diviser par  $(x_2 - x_0)(x_2 - x_1)$ ) nous donne

$$a = \frac{1}{x_2 - x_0} \left( \frac{y_2 - y_1}{x_2 - x_1} - \frac{y_1 - y_0}{x_1 - x_0} \right). \quad (1.4)$$

Un exemple pour le cas  $n = 5$  avec des  $x_i$  non équidistants est donné dans la fig. II.1. Pour les cas  $n > 2$  les formules deviennent plus complexes et il est recommandé d'introduire une notation convenable pour simplifier les expressions comme dans (1.4).

**Définition 1.1 (différences divisées)** Pour  $(x_i, y_i)$  donnés ( $x_i$  distincts) on définit

$$\begin{aligned} y[x_i] &:= y_i \\ \delta y[x_i, x_j] &:= \frac{y[x_j] - y[x_i]}{x_j - x_i} \\ \delta^2 y[x_i, x_j, x_k] &:= \frac{\delta y[x_j, x_k] - \delta y[x_i, x_j]}{x_k - x_i} \\ \delta^n y[x_{i_0}, x_{i_1}, \dots, x_{i_n}] &:= \frac{\delta^{n-1} y[x_{i_1}, \dots, x_{i_n}] - \delta^{n-1} y[x_{i_0}, \dots, x_{i_{n-1}}]}{x_{i_n} - x_{i_0}}. \end{aligned}$$

**Théorème 1.2 (formule de Newton)** Le polynôme d'interpolation de degré  $n$  qui passe par les  $n + 1$  points  $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ , où les  $x_i$  sont distincts, est unique et donné par

$$\begin{aligned} p(x) &= y[x_0] + (x - x_0) \delta y[x_0, x_1] + (x - x_0)(x - x_1) \delta^2 y[x_0, x_1, x_2] \\ &\quad + \dots + (x - x_0)(x - x_1) \cdot \dots \cdot (x - x_{n-1}) \delta^n y[x_0, x_1, \dots, x_n]. \end{aligned} \quad (1.5)$$

*Démonstration.* Pour  $n = 1$  et  $n = 2$  la formule (1.5) est équivalente à (1.2) et à (1.3)–(1.4). Pour démontrer le cas général, nous procédons par récurrence. Supposons que

$$p_1(x) = y[x_0] + (x - x_0) \delta y[x_0, x_1] + \dots + (x - x_0) \cdot \dots \cdot (x - x_{n-2}) \delta^{n-1} y[x_0, x_1, \dots, x_{n-1}]$$

soit le polynôme unique de degré  $n - 1$  qui passe par  $(x_i, y_i)$  pour  $i = 0, 1, \dots, n - 1$ . Alors, comme dans (1.3), le polynôme  $p(x)$  est nécessairement de la forme

$$p(x) = p_1(x) + a \cdot (x - x_0)(x - x_1) \cdot \dots \cdot (x - x_{n-1}), \quad (1.6)$$

où  $a$  est déterminé par  $p(x_n) = y_n$ . Il en résulte l'unicité du polynôme d'interpolation.

Pour montrer que  $a = \delta^n y[x_0, x_1, \dots, x_n]$ , ce qui achève la démonstration, nous considérons également le polynôme de degré  $n - 1$

$$p_2(x) = y[x_1] + (x - x_1) \delta y[x_1, x_2] + \dots + (x - x_1) \cdot \dots \cdot (x - x_{n-1}) \delta^{n-1} y[x_1, x_2, \dots, x_n]$$

qui passe par  $(x_i, y_i)$  pour  $i = 1, \dots, n$ . Ensuite, on pose (idée d'Aitken–Neville, 1929, 1932)

$$q(x) := \frac{1}{x_n - x_0} \left( (x_n - x) p_1(x) + (x - x_0) p_2(x) \right). \quad (1.7)$$

Il s'agit d'un polynôme de degré  $n$  qui satisfait la condition (1.1) pour le point  $x_0$  (ici, le facteur  $(x - x_0)$  est nul), pour le point  $x_n$  (ici, le facteur  $(x_n - x)$  est nul), et pour les points  $x_i$ ,  $i = 1, \dots, n - 1$  (ici, les deux polynômes  $p_1(x)$  et  $p_2(x)$  sont égaux à  $y_i$ ). Par l'unicité du polynôme d'interpolation on a alors  $q(x) = p(x)$ . En comparant le coefficient de  $x^n$  dans (1.7) avec celui de (1.6), nous obtenons

$$a = \frac{1}{x_n - x_0} \left( \delta^{n-1} y[x_1, \dots, x_n] - \delta^{n-1} y[x_0, \dots, x_{n-1}] \right) = \delta^n y[x_0, \dots, x_n],$$

ce qui démontre la formule (1.5).  $\square$

TAB. II.1: Différences divisées pour les données de la fig. II.1

$x_i$	$y_i$	$\delta y$	$\delta^2 y$	$\delta^3 y$	$\delta^4 y$	$\delta^5 y$
0	-1					
2	1	1				
4	6	5/2	3/8			
5	0	-6	-17/6	-77/120		
8	2	2/3	5/3	3/4	167/960	
10	5	3/2	1/6	-1/4	-1/8	-287/9600

**Exemple 1.3** Pour les données de la fig. II.1, les différences divisées sont présentées dans le tableau II.1. Le polynôme d'interpolation est alors donné par

$$p(x) = -1 + x + x(x-2)\frac{3}{8} - x(x-2)(x-4)\frac{77}{120} + x(x-2)(x-4)(x-5)\frac{167}{960} - x(x-2)(x-4)(x-5)(x-8)\frac{287}{9600}.$$

ou mieux encore pour la programmation (ou le calcul à la main)

$$p(x) = -1 + x \left( 1 + (x-2) \left( \frac{3}{8} + (x-4) \left( -\frac{77}{120} + (x-5) \left( \frac{167}{960} - (x-8) \frac{287}{9600} \right) \right) \right) \right).$$

*Remarque.* L'ordre des  $\{x_i\}$  n'a aucune importance pour la formule de Newton (1.5). Si l'on permute les données  $(x_i, y_i)$ , on obtient évidemment le même polynôme. Pour l'exemple ci-dessus et pour les  $\{x_i\}$  choisis dans l'ordre  $\{4, 5, 2, 8, 0, 10\}$ , on obtient ainsi

$$p(x) = 6 + (x-4) \left( -6 + (x-5) \left( -\frac{17}{6} + (x-2) \left( \frac{3}{4} + (x-8) \left( \frac{167}{960} - x \frac{287}{9600} \right) \right) \right) \right).$$

En observant que  $\delta^n y[x_{i_0}, \dots, x_{i_n}]$  est une fonction symétrique de ses arguments (par exemple,  $\delta^2 y[x_2, x_3, x_1] = \delta^2 y[x_1, x_2, x_3]$ , voir l'exercice 2), on peut utiliser les valeurs calculées dans le tableau II.1.

Pour diminuer l'influence des erreurs d'arrondi, il est recommandé d'ordonner les  $\{x_i\}$  de manière à ce que les valeurs situées au milieu soient prises d'abord et les valeurs aux extrémités à la fin. Pour ce choix, les expressions  $(x - x_0)$ ,  $(x - x_0)(x - x_1)$ ,  $(x - x_0)(x - x_1)(x - x_2)$ , etc., sont en général plus petites que pour un autre choix et l'amplification des erreurs dans les différences divisées est moins importante.

## II.2 Erreur de l'interpolation

Supposons que les points  $(x_i, y_i)$  soient sur le graphe d'une fonction  $f : [a, b] \rightarrow \mathbb{R}$ , c.-à-d.,

$$y_i = f(x_i), \quad i = 0, 1, \dots, n, \quad (2.1)$$

étudions alors l'erreur  $f(x) - p(x)$  du polynôme d'interpolation  $p(x)$ . Deux exemples sont donnés dans la fig. II.2. A gauche, on voit un polynôme d'interpolation pour la fonction  $f(x) = \sin x$ , et à droite pour la fonction  $1/(1 + x^2)$ . Pour mieux voir l'erreur, on a dessiné la fonction  $f(x)$  en une courbe pointillée.

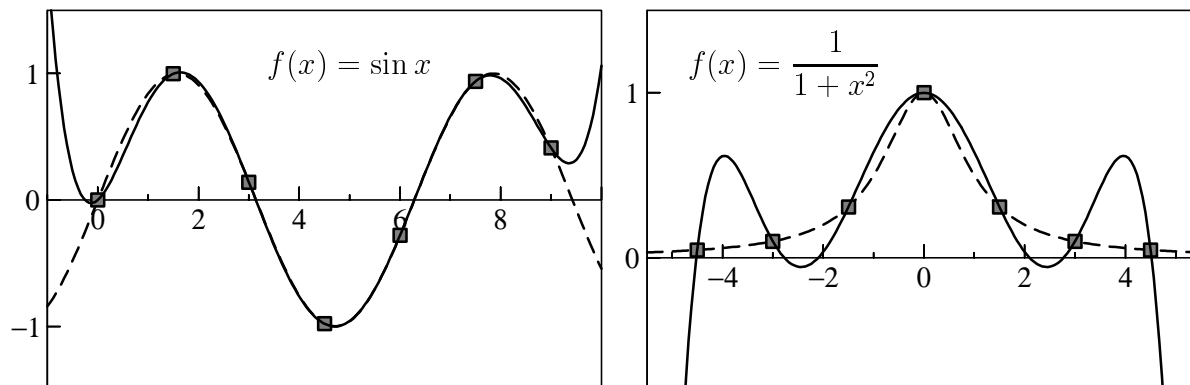


FIG. II.2: Polynôme d'interpolation pour  $\sin x$  (gauche) et pour  $1/(1 + x^2)$  (droite)

Les résultats suivants sont dus à Cauchy (1840, *Sur les fonctions interpolaires*, C.R. XI, p. 775-789, *Oeuvres* ser. 1, vol. V, p. 409-424). Commençons par une relation intéressante entre les différences divisées pour (2.1) et les dérivées de la fonction  $f(x)$ .

**Lemme 2.1** Soit  $f(x)$   $n$ -fois différentiable et  $y_i = f(x_i)$  pour  $i = 0, 1, \dots, n$  ( $x_i$  distincts). Alors, il existe un  $\xi \in (\min x_i, \max x_i)$  tel que

$$\delta^n y[x_0, x_1, \dots, x_n] = \frac{f^{(n)}(\xi)}{n!}. \quad (2.2)$$

*Démonstration.* Soit  $p(x)$  le polynôme d'interpolation de degré  $n$  passant par  $(x_i, y_i)$  et notons  $d(x) = f(x) - p(x)$ . Par définition de  $p(x)$ , la différence  $d(x)$  s'annule en  $n + 1$  points distincts :

$$d(x_i) = 0 \quad \text{pour} \quad i = 0, 1, \dots, n.$$

Comme  $d(x)$  est différentiable, on peut appliquer  $n$  fois le théorème de Rolle (voir le cours d'Analyse I) et on en déduit que

$$d'(x) \quad \text{a } n \text{ zéros distincts dans } (\min x_i, \max x_i).$$

Le même argument appliqué à  $d'(x)$  donne

$$d''(x) \quad \text{a } n - 1 \text{ zéros distincts dans } (\min x_i, \max x_i),$$

et finalement encore

$$d^{(n)}(x) \quad \text{a 1 zéro dans } (\min x_i, \max x_i).$$

Notons ce zéro de  $d^{(n)}(x)$  par  $\xi$ . Alors, on a

$$f^{(n)}(\xi) = p^{(n)}(\xi) = n! \cdot \delta^n y[x_0, x_1, \dots, x_n]. \quad (2.3)$$

La deuxième identité dans (2.3) résulte du fait que  $\delta^n y[x_0, x_1, \dots, x_n]$  est le coefficient de  $x^n$  dans  $p(x)$ .  $\square$

**Théorème 2.2** Soit  $f : [a, b] \rightarrow \mathbb{R}$   $(n + 1)$ -fois différentiable et soit  $p(x)$  le polynôme d'interpolation de degré  $n$  qui passe par  $(x_i, f(x_i))$  pour  $i = 0, 1, \dots, n$ . Alors, pour  $x \in [a, b]$ , il existe un  $\xi \in (\min(x_i, x), \max(x_i, x))$  tel que

$$f(x) - p(x) = (x - x_0) \cdot \dots \cdot (x - x_n) \cdot \frac{f^{(n+1)}(\xi)}{(n + 1)!}. \quad (2.4)$$

*Démonstration.* Si  $x = x_i$  pour un indice  $i \in \{0, 1, \dots, n\}$ , la formule (2.4) est vérifiée car  $p(x_i) = f(x_i)$ . Fixons alors un  $\bar{x}$  dans  $[a, b]$  qui soit différent de  $x_i$  et montrons la formule (2.4) pour  $x = \bar{x}$ .

L'idée est de considérer le polynôme  $\bar{p}(x)$  de degré  $n + 1$  qui passe par  $(x_i, f(x_i))$  pour  $i = 0, 1, \dots, n$  et par  $(\bar{x}, f(\bar{x}))$ . La formule de Newton donne

$$\bar{p}(x) = p(x) + (x - x_0) \cdot \dots \cdot (x - x_n) \cdot \delta^{n+1} y[x_0, \dots, x_n, \bar{x}]. \quad (2.5)$$

Si l'on remplace la différence divisée dans (2.5) par  $f^{(n+1)}(\xi)/(n + 1)!$  (voir le lemme précédent) et si l'on pose  $x = \bar{x}$ , on obtient le résultat (2.4) pour  $x = \bar{x}$  car  $\bar{p}(\bar{x}) = f(\bar{x})$ . Comme  $\bar{x}$  est arbitraire, la formule (2.4) est vérifiée pour tout  $x$ .  $\square$

**Exemple 2.3** Dans la situation de la fig. II.2, on a  $n + 1 = 7$ . Comme la 7<sup>ème</sup> dérivée de  $\sin x$  est majorée par 1, on a que

$$|p(x) - \sin x| \leq |x(x - 1.5)(x - 3)(x - 4.5)(x - 6)(x - 7.5)(x - 9)| \cdot \frac{1}{7!},$$

par exemple

$$|p(4) - \sin 4| \leq 0.035 \quad \text{ou} \quad |p(1) - \sin 1| \leq 0.181.$$

Pour le deuxième exemple,  $f(x) = 1/(1 + x^2)$ , la 7<sup>ème</sup> dérivée est donnée par

$$f^{(7)}(x) = -8! \cdot \frac{(x + 1)(x - 1)x(x^2 - 2x - 1)(x^2 + 2x - 1)}{(1 + x^2)^8},$$

qui est maximale pour  $x \approx \pm 0.17632698$ . On obtient ainsi

$$\left| p(x) - \frac{1}{1 + x^2} \right| \leq |(x^2 - 20.25)(x^2 - 9)(x^2 - 2.25)x| \cdot \frac{4392}{7!}.$$

Alors, l'erreur peut être 4392 fois plus grande que pour l'interpolation de  $\sin x$ .

## II.3 Polynômes de Chebyshev

La formule (2.4) montre que l'erreur de l'interpolation est un produit de la  $(n+1)$ <sup>ème</sup> dérivée de  $f(x)$ , évaluée à un point inconnu, avec l'expression  $(x-x_0) \cdot \dots \cdot (x-x_n)$  qui ne dépend que de la division  $\{x_0, \dots, x_n\}$ .

**Problème intéressante :** chercher, pour un  $n$  donné, la division de  $[a, b]$  pour laquelle

$$L = \max_{x \in [a,b]} |(x-x_0) \cdot \dots \cdot (x-x_n)| \quad \text{est minimal.} \quad (3.1)$$

**Exemple 3.1** Considérons l'intervalle  $[-1, 1]$ , le cas  $n = 2$ , et une distribution symétrique des  $x_i$ . On cherche alors  $p(x) = (x+\alpha)x(x-\alpha) = x^3 - ax$  satisfaisant (3.1). Nous voyons en figure II.3 que, pour  $a = 0$ , nous avons  $L = 1$ ; puis cette valeur diminue quand  $a$  croît, jusqu'au moment où la courbe touche les bornes  $+L$  et  $-L$  à l'intérieur de  $[-1, 1]$  (pour  $a = 3/4$ ); après,  $L$  recommence de grandir. La solution optimale est donc  $p(x) = x^3 - 3x/4$ .

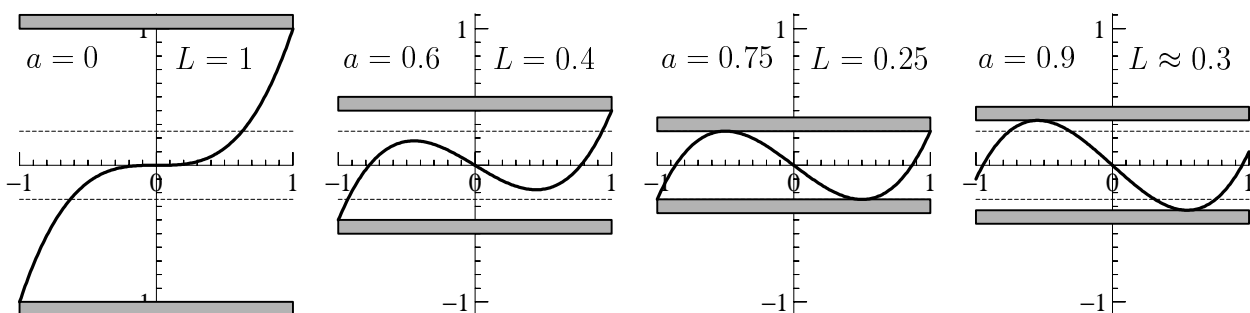


FIG. II.3: Valeurs maximales de  $\tau(x) = x^3 - ax$

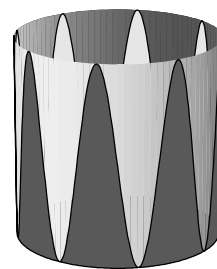
Pour  $n$  arbitraire, la réponse au problème posé se trouve dans un travail de P.L. Chebyshev (transcription française : “Tchebychef”, 1854, *Œuvres* I, p. 109). Elle peut être donnée à l'aide de polynômes de Chebyshev; voir la figure II.4.<sup>1</sup>

### Définition 3.2 (Polynômes de Chebyshev)

Pour  $n = 0, 1, 2, \dots$  et pour  $x \in [-1, 1]$ , on définit

$$T_n(x) = \cos(n \arccos x) \quad (3.2)$$

(projection du cosinus sur un tambour).



### Propriétés des polynômes de Chebyshev.

a) Les fonctions  $T_n(x)$  satisfont la récurrence

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \quad (3.3)$$

Par conséquent,  $T_n(x)$  est un polynôme de degré  $n$  dont le coefficient de  $x^n$  est  $2^{n-1}$ , c.-à-d.,  $T_n(x) = 2^{n-1}x^n + \dots$

b)  $|T_n(x)| \leq 1$  pour  $x \in [-1, 1]$ .

c)  $T_n\left(\cos\left(\frac{k\pi}{n}\right)\right) = (-1)^k$  pour  $k = 0, 1, \dots, n$ .

d)  $T_n\left(\cos\left(\frac{(2k+1)\pi}{2n}\right)\right) = 0$  pour  $k = 0, 1, \dots, n-1$ .

<sup>1</sup>Pour une étude des “courbes blanches” dans la fig. II.4 (à droite) voir la page 209 du livre: Th.J. Rivlin, *Chebyshev Polynomials*. 2nd ed., John Wiley & Sons, 1990 [MA 41/36]

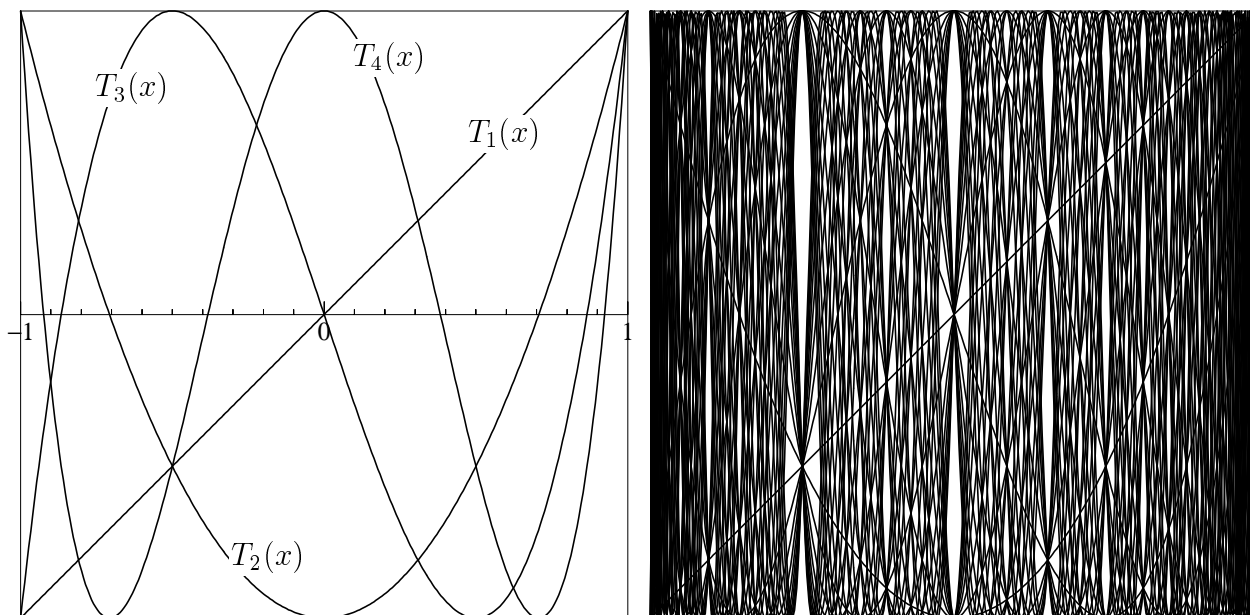


FIG. II.4: Les premiers 4 (à gauche) respectivement 30 (à droite) polynômes de Chebyshev

e) Les polynômes  $T_n(x)$  sont orthogonaux par rapport à la fonction de poids  $1/\sqrt{1-x^2}$

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_n(x) T_m(x) dx = \begin{cases} \pi & \text{si } n = m = 0 \\ \pi/2 & \text{si } n = m \neq 0 \\ 0 & \text{si } n \neq m. \end{cases}$$

*Démonstration.* La formule (3.3) est une conséquence de

$$\cos((n+1)\varphi) + \cos((n-1)\varphi) = 2 \cos \varphi \cdot \cos(n\varphi)$$

si l'on pose  $\cos \varphi = x$  et  $\varphi = \arccos x$ . La même transformation donne

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_n(x) T_m(x) dx = \int_0^\pi \cos(n\varphi) \cos(m\varphi) d\varphi$$

et la propriété (e) résulte de l'orthogonalité de  $\cos(n\varphi)$ .  $\square$

Revenons maintenant à la question de trouver une division satisfaisant à (3.1).

**Lemme 3.3** Soit  $q(x)$  un polynôme de degré  $n$  dont le coefficient de  $x^n$  est  $2^{n-1}$  (comme pour le polynôme de Chebyshev) et soit  $q(x) \not\equiv T_n(x)$ . Alors,

$$\max_{x \in [-1, 1]} |q(x)| > \max_{x \in [-1, 1]} |T_n(x)| = 1. \quad (3.4)$$

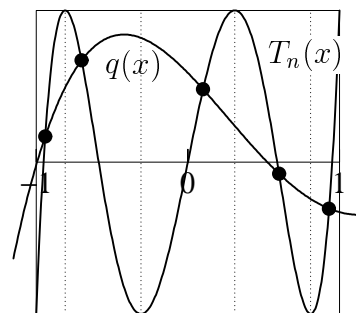
*Démonstration.* Supposons, par l'absurde, que

$$\max_{x \in [-1, 1]} |q(x)| \leq \max_{x \in [-1, 1]} |T_n(x)|$$

(voir la figure pour  $n = 5$ ) et considérons la différence  $d(x) = q(x) - T_n(x)$ . Cette fonction s'annule au moins une fois dans chacun des intervalles fermés

$$\left[ \cos\left(\frac{(k+1)\pi}{n}\right), \cos\left(\frac{k\pi}{n}\right) \right], \quad k = 0, 1, \dots, n-1. \quad (3.5)$$

Alors,  $d(x)$  possède  $n$  zéros dans  $[-1, 1]$  (si une racine  $\alpha \in (-1, 1)$  est à l'extrémité de l'intervalle (3.5), elle doit être comptée deux fois car à un tel point  $T'_n(\alpha) = 0$  et  $q'(\alpha) = 0$ ). Comme  $d(x)$  est un polynôme de degré  $n-1$  (le coefficient de  $x^n$  est le même pour  $q(x)$  et  $T_n(x)$ ), ceci est une contradiction à  $d(x) \not\equiv 0$ .  $\square$



Le lemme précédent montre que

$$\max_{x \in [-1, 1]} |(x - x_0) \cdot \dots \cdot (x - x_n)| \quad \text{est minimal}$$

si et seulement si  $(x - x_0) \cdot \dots \cdot (x - x_n) = 2^{-n} T_{n+1}(x)$ , c.-à-d., si

$$x_k = \cos\left(\frac{(2k+1)\pi}{2n+2}\right), \quad k = 0, 1, \dots, n \quad (3.6)$$

(points de Chebyshev). Pour répondre à la question (3.1), il faut encore utiliser la translation  $x \mapsto \frac{a+b}{2} + \frac{b-a}{2}x$ , qui envoie l'intervalle  $[-1, 1]$  sur  $[a, b]$ . On obtient alors

**Théorème 3.4** *L'expression (3.1) est minimale parmi toutes les divisions  $\{x_0, x_1, \dots, x_n\}$  si et seulement si*

$$x_k = \frac{a+b}{2} + \frac{b-a}{2} \cdot \cos\left(\frac{(2k+1)\pi}{2n+2}\right), \quad k = 0, 1, \dots, n. \quad (3.7)$$

**Exemple 3.5** Considérons la fonction  $f(x) = 1/(1+25x^2)$  sur l'intervalle  $[-1, 1]$  (cf. la fig. II.2). Dans la fig. II.5, on compare le polynôme d'interpolation basé sur des points équidistants ( $n = 8$ ) avec celui basé sur les points de Chebyshev (aussi  $n = 8$ ). On observe une nette amélioration.

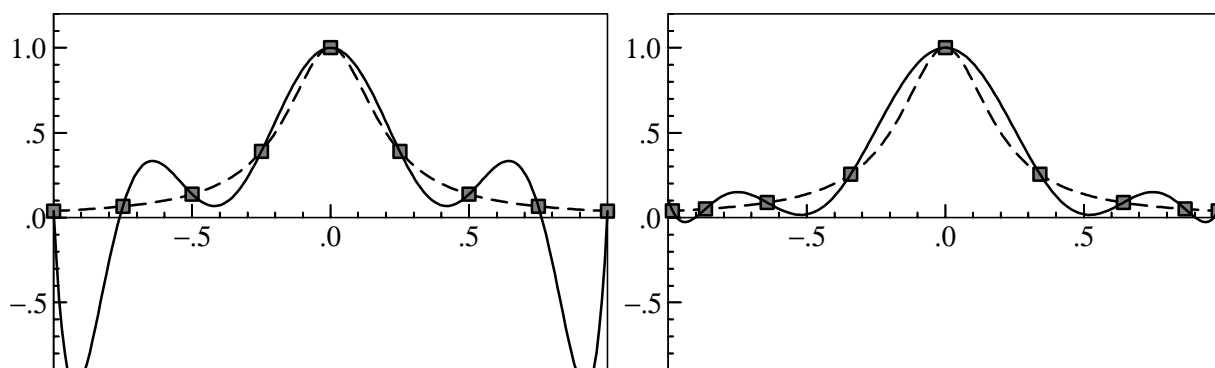


FIG. II.5: Interpolation avec des points équidistants (à gauche) et les points de Chebyshev (à droite)

## II.4 Convergence de l'interpolation

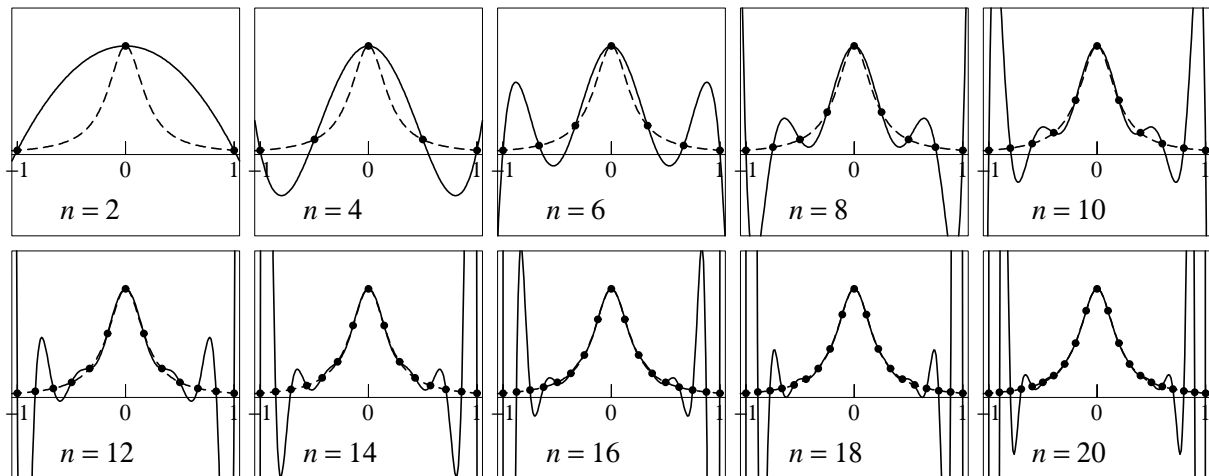
La précision du polynôme d'interpolation (figure II.5 à gauche) est certainement insuffisante, quand  $x$  se rapproche des bords. On pourrait naïvement croire que le degré  $n$  n'est pas encore assez élevé. Mais le contraire se produit : quand on augmente  $n$ , la catastrophe s'intensifie davantage. C'est le *phénomène de Runge* (voir le dessin de la figure II.6).

**Le phénomène de Runge (1901).**<sup>2</sup> Nous considérons des fonctions rationnelles (ou méromorphes)  $f(x)$ , définies sur l'intervalle  $[-1, 1]$ , et la suite des divisions *équidistantes*

$$x_k^{(n)} = -1 + \frac{2k}{n}, \quad k = 0, 1, \dots, n. \quad (4.1)$$

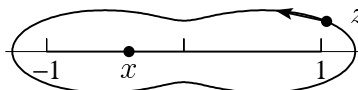
<sup>2</sup>Carl David Tolmé Runge (1856-1927) est le père des mathématiques appliquées en Allemagne. L'article sur ce phénomène a été publié en 1901 dans le journal *Zeitschr. Math. u. Physik* vol. 46.



FIG. II.6: Le phénomène de Runge pour la fonction  $f(x) = 1/(1 + 25x^2)$ 

Nous notons par  $p_n(x)$  le polynôme d'interpolation de degré  $n$  satisfaisant à  $p(x_k^{(n)}) = f(x_k^{(n)})$  pour  $k = 0, 1, \dots, n$ . Le but est d'étudier la convergence de  $p_n(x)$  vers  $f(x)$ . On peut observer dans la fig. II.6 que, dans un certain intervalle autour de l'origine,  $p_n(x)$  converge vers  $f(x)$ , mais au bord de l'intervalle  $[-1, 1]$ , on a divergence.

**Excursion en analyse complexe.** Le meilleur moyen de comprendre ce phénomène est la formule de Cauchy (1831), voir le cours "Analyse II",

$$f(x) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - x} dz. \quad (4.2)$$


Dans cette formule,  $\gamma$  est une courbe fermée autour de  $x$ , telle que  $f(z)$  n'a pas de singularité à l'intérieur de la courbe. Si la courbe est donnée par la paramétrisation  $\gamma : [a, b] \rightarrow \mathbb{C}$  avec  $\gamma(a) = \gamma(b)$ , l'intégrale dans (4.2) est définie par

$$\frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - x} dz = \frac{1}{2\pi i} \int_a^b \frac{f(\gamma(t))}{\gamma(t) - x} \cdot \gamma'(t) dt.$$

**Une formule pour le polynôme d'interpolation.** En utilisant la notation

$$\pi_n(x) := (x - x_0^{(n)}) \cdot (x - x_1^{(n)}) \cdot \dots \cdot (x - x_n^{(n)}) \quad (4.3)$$

pour la division équidistante (4.1), nous obtenons pour le polynôme d'interpolation

$$p_n(x) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - x} \cdot \frac{\pi_n(z) - \pi_n(x)}{\pi_n(z)} dz. \quad (4.4)$$

Ici,  $\gamma$  est une courbe fermée autour du segment  $[-1, 1]$  telle que  $f(z)$  n'ait pas de singularité (pôle) à l'intérieur de la courbe (voir la petite figure dans (4.2)).

En effet, la partie droite de (4.4) est un polynôme de degré  $n$  en  $x$  car  $(\pi_n(z) - \pi_n(x))/(z - x)$  en est un. En posant  $x = x_k^{(n)}$  pour un  $k \in \{0, \dots, n\}$ , on a

$$\frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - x_k^{(n)}} \cdot \frac{\pi_n(z) - \pi_n(x_k^{(n)})}{\pi_n(z)} dz = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - x_k^{(n)}} dz = f(x_k^{(n)}),$$

ce qui montre la formule (4.4). La différence de (4.4) et de (4.2) nous donne la formule suivante pour l'erreur de l'interpolation :

$$f(x) - p_n(x) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - x} \cdot \frac{\pi_n(x)}{\pi_n(z)} dz. \quad (4.5)$$

Il est important de se rappeler que  $\gamma$  est une courbe fermée autour du segment  $[-1, 1]$  telle que  $f(z)$  n'ait pas de singularité à l'intérieur de  $\gamma$ .

**Étude de la convergence.** La formule (4.5) nous permet de procéder de la manière suivante : si, pour un  $x \in [-1, 1]$ , on peut choisir la courbe  $\gamma$  telle que

$$|\pi_n(x)| \leq \alpha^n |\pi_n(z)| \quad \text{pour tout } z \in \gamma \quad (4.6)$$

avec un  $\alpha < 1$ , alors,

$$|f(x) - p_n(x)| \leq \frac{\alpha^n}{2\pi} \int_{\gamma} \frac{|f(z)|}{|z - x|} |dz| \quad (4.7)$$

et on a convergence pour  $n \rightarrow \infty$ . Il faut alors étudier la fonction  $\sqrt[n]{|\pi_n(z)|}$  pour  $n \rightarrow \infty$ .

**Lemme 4.1** Pour la division équidistante (4.1) et pour le polynôme  $\pi_n(z)$  de (4.3) on a

$$\lim_{n \rightarrow \infty} \sqrt[n]{|\pi_n(z)|} = G(z)$$

où

$$G(z) = \exp\left(\frac{1}{2}\Re((z+1)\log(z+1) - (z-1)\log(z-1)) - 1\right). \quad (4.8)$$

*Démonstration.* En prenant le logarithme de  $\sqrt[n]{|\pi_n(z)|}$ , on obtient

$$\log \sqrt[n]{|\pi_n(z)|} = \frac{1}{n} \log |\pi_n(z)| = \frac{1}{n} (\log |z - x_0^{(n)}| + \dots + \log |z - x_n^{(n)}|),$$

ce qui représente une somme de Riemann pour la fonction  $t \mapsto \frac{1}{2} \log |z - t|$ ,  $x_{k+1}^{(n)} - x_k^{(n)} = 2/n$ . On obtient ainsi (observer que  $\log w = \log |w| + i \arg w$ )

$$\lim_{n \rightarrow \infty} \log \sqrt[n]{|\pi_n(z)|} = \frac{1}{2} \int_{-1}^1 \log |z - t| dt = \frac{1}{2} \Re \int_{-1}^1 \log(z - t) dt.$$

En considérant  $z$  comme paramètre, une primitive de la fonction  $g(t) = \log(z - t)$  est donnée par  $(z - t) - (z - t) \log(z - t)$ . Ceci implique

$$\lim_{n \rightarrow \infty} \log \sqrt[n]{|\pi_n(z)|} = \frac{1}{2} \Re((z+1)\log(z+1) - (z-1)\log(z-1) - (z+1) + (z-1)),$$

et démontre la formule du lemme.  $\square$

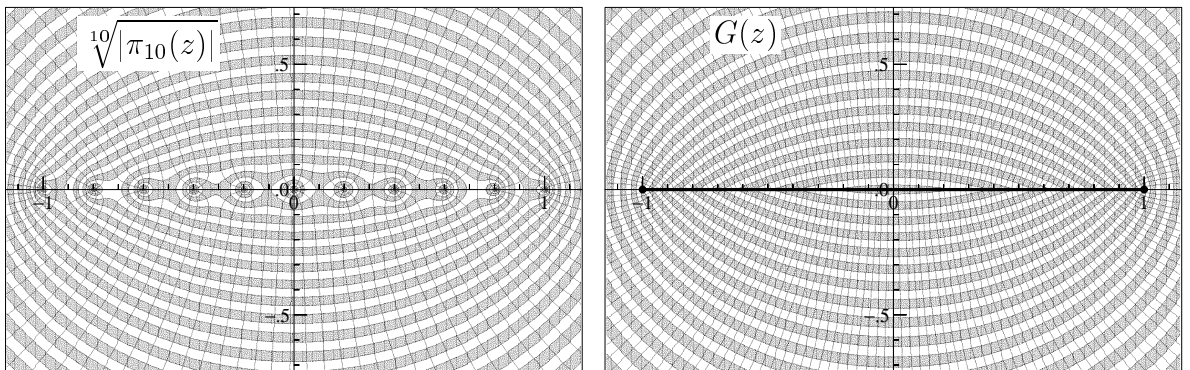
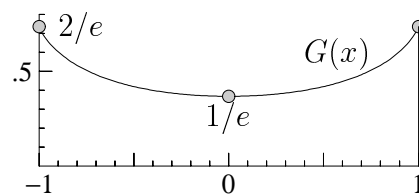


FIG. II.7: Lignes de niveau pour les fonctions  $10\sqrt[10]{|\pi_{10}(x)|}$  et  $G(z)$ ; division équidistante.

Pour un  $x$  réel et compris dans  $(-1, 1)$ , on a

$$G(x) = e^{-1} \sqrt{(1+x)^{1+x} (1-x)^{1-x}},$$

en particulier  $G(0) = 1/e$ ,  $G(\pm 1) = 2/e$ . Les lignes de niveau  $\{z \mid G(z) = \text{Const}\}$  sont dessinées dans la fig. II.7 (dessin à droite).



**Théorème 4.2 (Runge 1901)** Si  $f(z)$  n'a pas de singularité dans  $\{z \mid G(z) \leq G(\beta)\}$ , alors

$$\lim_{n \rightarrow \infty} p_n(x) = f(x) \quad \text{pour} \quad x \in (-\beta, \beta).$$

*Démonstration.* Nous considérons une courbe  $\gamma$  telle que l'intervalle  $[-1, 1]$  ainsi que l'ensemble  $\{z \mid G(z) \leq G(\beta)\}$  sont à l'intérieur, mais tous les pôles de  $f(z)$  à l'extérieur. Pour un  $x \in (-\beta, \beta)$  fixé, ceci entraîne que

$$G(x) < G(\beta) \leq \min_{z \in \gamma} G(z).$$

Comme  $\sqrt[n]{|\pi_n(x)|} \rightarrow G(x)$  et  $\min_{z \in \gamma} \sqrt[n]{|\pi_n(z)|} \rightarrow \min_{z \in \gamma} G(z)$  (convergence uniforme sur  $\gamma$ ; pour le voir il faut élaborer la démonstration du lemme 4.1), on peut trouver un  $\alpha$  ( $0 < \alpha < 1$ ) tel que, pour  $n$  suffisamment grand, on a

$$\sqrt[n]{|\pi_n(x)|} \leq \alpha \cdot \sqrt[n]{|\pi_n(z)|} \quad \text{pour} \quad z \in \gamma.$$

Ceci vérifie (4.6) pour  $x \in (-\beta, \beta)$  et la convergence est une conséquence de (4.7).  $\square$

*Exemple.* La fonction  $f(x) = 1/(1 + 25x^2)$  possède une singularité en  $\pm i/5$ . La ligne de niveau de  $G(z)$  qui passe par  $\pm i/5$  coupe l'axe réel au point  $\beta \approx 0.726$ . Alors, la convergence du polynôme d'interpolation (avec des points équidistants) est établie pour  $|x| < 0.726$  (voir fig. II.6).

**Convergence en utilisant des points de Chebyshev.** Faisons la même étude en remplaçant la division équidistante (4.1) par les points de Chebyshev

$$x_k^{(n)} = \cos\left(\frac{(2k+1)\pi}{2n+2}\right), \quad k = 0, 1, \dots, n. \quad (4.9)$$

L'expérience de la fig. II.6 avec la fonction  $f(x) = 1/(1 + 25x^2)$  est répétée dans la fig. II.8. On observe une convergence uniforme sur toute l'intervalle  $[-1, 1]$ .

Pour comprendre l'absence d'un "phénomène de Runge", considérons le polynôme  $\pi_n(x)$  de (4.3) mais pour la division (4.9). Dans ce cas (voir (3.2) et (3.6))

$$\pi_n(z) = 2^{-n} T_{n+1}(z) = 2^{-n} \cos((n+1)\varphi) = \frac{u^{n+1} + u^{-(n+1)}}{2^{n+1}} \quad (4.10)$$

où  $\varphi = \arccos z$  et  $u = e^{i\varphi}$ .

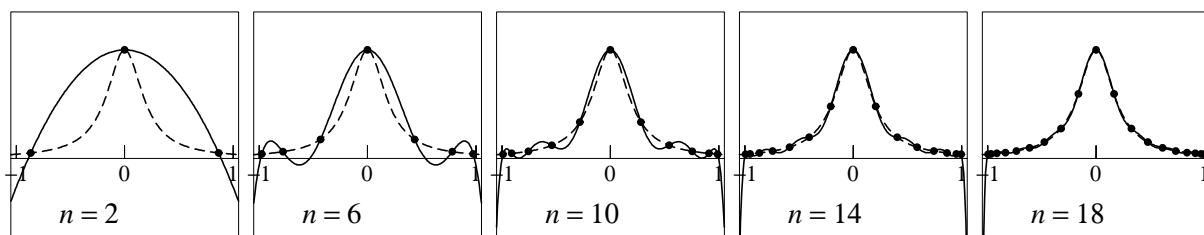


FIG. II.8: Interpolation de la fonction  $f(x) = 1/(1 + 25x^2)$  en utilisant des points de Chebyshev.

La relation  $z = \cos \varphi = \frac{1}{2}(e^{i\varphi} + e^{-i\varphi}) = \frac{1}{2}(u + u^{-1})$  donne l'équation  $u^2 - 2uz + 1 = 0$  pour  $u$ , dont les solutions sont  $u = z + \sqrt{z^2 - 1}$ ,  $u^{-1} = z - \sqrt{z^2 - 1}$ . Élevées à la  $(n+1)$ -ème puissance, la plus grande de ces grandeurs va finir par dominer l'autre. Ainsi,

$$G(z) = \lim_{n \rightarrow \infty} \sqrt[n]{|\pi_n(z)|} = \frac{1}{2} \max\{|z + \sqrt{z^2 - 1}|, |z - \sqrt{z^2 - 1}|\}. \quad (4.11)$$

Les lignes de niveau de la fonction limite deviennent parallèles à l'intervalle  $[-1, 1]$  (voir fig. II.9). Ainsi, un pôle se trouvant en dehors de l'intervalle  $[-1, 1]$  ne peut empêcher le polynôme  $p_n(x)$  de converger sur tout l'intervalle (voir fig. II.8).

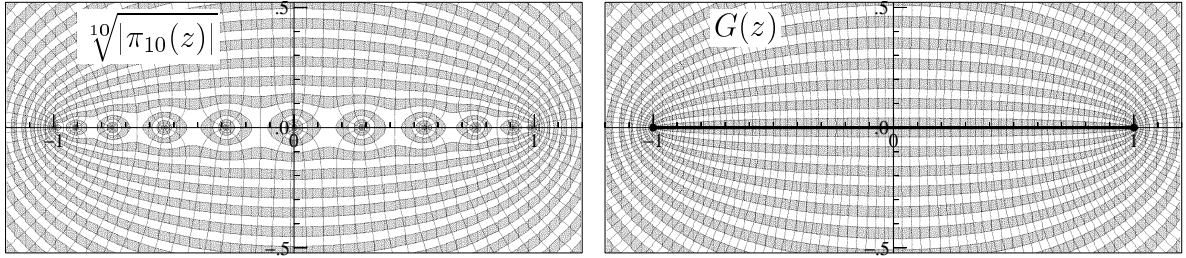


FIG. II.9: Lignes de niveau pour les fonctions  $\sqrt[10]{|\pi_{10}(x)|}$  et  $G(z)$ ; points de Chebyshev.

*Remarque.* La convergence (uniforme) de la suite de polynômes d'interpolation avec les points de Chebyshev peut être démontré pour toute fonction  $f(x)$  qui est continûment différentiable sur  $[-1, 1]$ . La *démonstration* utilise le théorème de Stone-Weierstraß (voir par exemple le chapitre II du livre de Werner & Schaback<sup>3</sup> pour plus de détails).

## II.5 Influence des erreurs d'arrondi sur l'interpolation

**Représentation en virgule flottante.** Chaque nombre réel  $x \neq 0$  peut s'écrire sous la forme

$$x = \pm a \cdot 10^b \quad (5.1)$$

où  $a$ , la *mantisse*, satisfait  $0.1 \leq a < 1$  et l'*exposant*  $b$  est un nombre entier. Cette représentation de  $x$  est unique. Supposons maintenant qu'on ait seulement un nombre fini de chiffres (disons  $\ell$ ) à disposition pour la mantisse, et pas de restriction pour l'exposant. Si  $\bar{a}$  dénote l'arrondi de  $a$ , on va calculer en fait avec le nombre

$$\text{arr}(x) = \pm \bar{a} \cdot 10^b$$

au lieu de  $x$ . Par exemple, le nombre  $\pi = 3.141592653 \dots$  est représenté par

$$\text{arr}(\pi) = 0.31415927 \cdot 10^1 \quad (5.2)$$

si l'on calcule avec  $\ell = 8$  chiffres en base 10.

**Précision de l'ordinateur.** On dénote par  $\text{eps}$  le plus petit nombre positif tel que

$$\text{arr}(1 + \text{eps}) > 1.$$

Pour un calcul en base 10 avec  $\ell$  chiffres dans la mantisse on a

$$\text{arr}(0.\underbrace{10 \dots 0}_{\ell} 49 \dots \cdot 10^1) = 1$$

$$\text{arr}(0.\underbrace{10 \dots 0}_{\ell} 50 \dots \cdot 10^1) = 0.\underbrace{10 \dots 1}_{\ell} \cdot 10^1 > 1.$$

<sup>3</sup>H. Werner & R. Schaback (1979), *Praktische Mathematik II*. Springer-Verlag, 2. Auflage. [MA 65/23]

Dans cette situation, on a alors

$$eps = 5 \cdot 10^{-\ell}. \quad (5.3)$$

Si l'on fait le même calcul en base 2 (comme tous les ordinateurs le font) on obtient

$$eps = 2^{-\ell}. \quad (5.4)$$

Par exemple, sur une SUN workstation on a

$$\begin{aligned} \text{REAL} * 4, & \quad eps = 2^{-24} \approx 5.96 \cdot 10^{-8} \\ \text{REAL} * 8, & \quad eps = 2^{-53} \approx 1.11 \cdot 10^{-16} \\ \text{REAL} * 16, & \quad eps = 2^{-113} \approx 9.63 \cdot 10^{-35}. \end{aligned}$$

**Théorème 5.1** Pour un  $x \neq 0$  on a

$$\frac{|\text{arr}(x) - x|}{|x|} \leq eps, \quad (5.5)$$

c.-à-d., l'erreur relative due à l'arrondissement est bornée par  $eps$ .

*Démonstration.* Soit  $x = a \cdot 10^b$  et  $\text{arr}(x) = \bar{a} \cdot 10^b$ . Si l'on arrondit à  $\ell$  chiffres significatifs, on a  $|\bar{a} - a| \leq 5 \cdot 10^{-\ell-1}$ . Il en résulte

$$\frac{|\text{arr}(x) - x|}{|x|} = \frac{|\bar{a} - a| \cdot 10^b}{|a| \cdot 10^b} \leq \frac{5 \cdot 10^{-\ell-1}}{10^{-1}} = 5 \cdot 10^{-\ell} = eps$$

car  $|a| \geq 1/10$ . □

L'estimation (5.5) peut aussi être écrite sous la forme

$$\text{arr}(x) = x(1 + \epsilon) \quad \text{où} \quad |\epsilon| \leq eps. \quad (5.6)$$

Cette formule est la base pour toute étude d'erreurs d'arrondi.

**Influence des erreurs dans  $y_i$  sur le polynôme d'interpolation.** Supposons que les données  $y_i$  soient erronées et qu'on calcule en fait avec

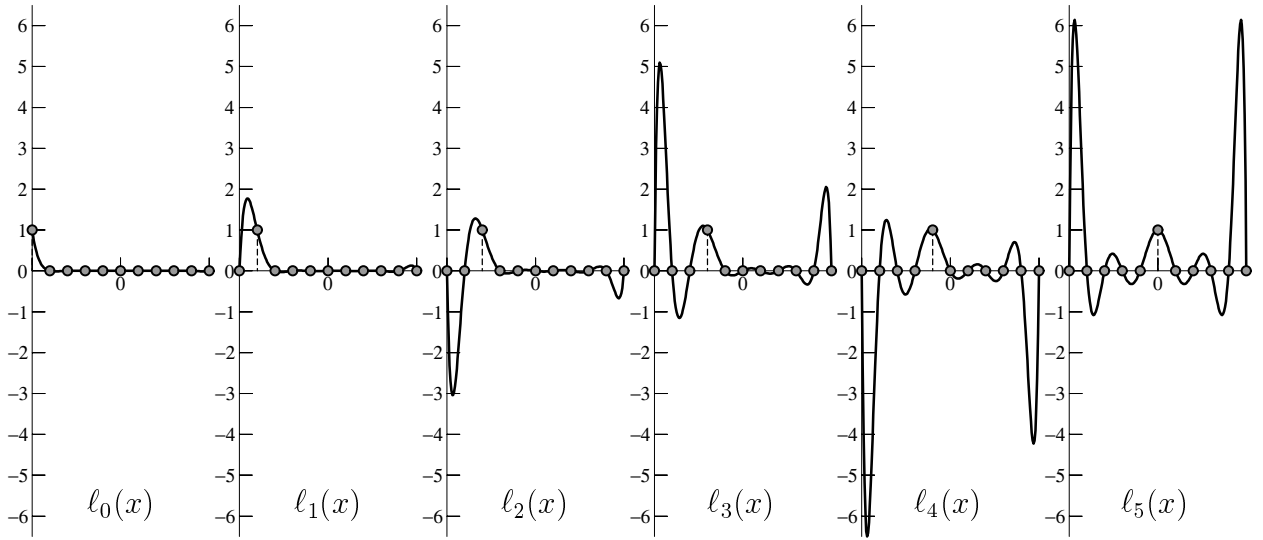
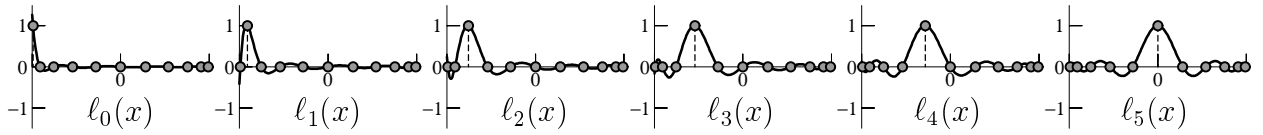
$$\hat{y}_i = y_i(1 + \epsilon_i) \quad \text{où} \quad |\epsilon_i| \leq eps. \quad (5.7)$$

Pour étudier la différence entre le polynôme qui passe par  $(x_i, y_i)$  et celui qui passe par  $(x_i, \hat{y}_i)$ , on utilise la formule du lemme suivant.

**Lemme 5.2 (formule de Lagrange)** Le polynôme d'interpolation  $p(x)$  qui passe par  $(x_i, y_i)$  pour  $i = 0, 1, \dots, n$  est donné par

$$p(x) = \sum_{i=0}^n y_i \ell_i(x) \quad \text{où} \quad \ell_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}. \quad (5.8)$$

*Démonstration.* Le polynôme  $\ell_i(x)$  satisfait  $\ell_i(x_i) = 1$  et  $\ell_i(x_k) = 0$  si  $k \neq i$ . Ainsi, les deux côtés de la formule (5.8) valent  $y_k$  pour  $x = x_k$  ( $k = 0, 1, \dots, n$ ). Comme  $p(x)$  et  $\sum_{i=0}^n y_i \ell_i(x)$  sont des polynômes de degré  $n$ , cette formule est une conséquence de l'unicité du polynôme d'interpolation (voir le théorème 1.2). □

FIG. II.10: Polynômes de Lagrange à points équidistants pour  $n = 10$ FIG. II.11: Polynômes de Lagrange à points de Chebyshev pour  $n = 10$ 

Les polynômes passant par  $(x_i, y_i)$  et  $(x_i, \hat{y}_i)$  sont respectivement

$$p(x) = \sum_{i=0}^n y_i \ell_i(x) \quad \text{et} \quad \hat{p}(x) = \sum_{i=0}^n \hat{y}_i \ell_i(x).$$

Si  $\hat{y}_i$  est donné par (5.7), la différence satisfait  $\hat{p}(x) - p(x) = \sum_{i=0}^n \epsilon_i y_i \ell_i(x)$  et on obtient

$$|\hat{p}(x) - p(x)| \leq \epsilon \cdot \max_{k=0, \dots, n} |y_k| \cdot \sum_{i=0}^n |\ell_i(x)|. \quad (5.9)$$

La fonction  $\sum_{i=0}^n |\ell_i(x)|$  décrit l'amplification de l'erreur dans les données. Sa valeur maximale

$$\Lambda_n := \max_{x \in [a, b]} \sum_{i=0}^n |\ell_i(x)| \quad (5.10)$$

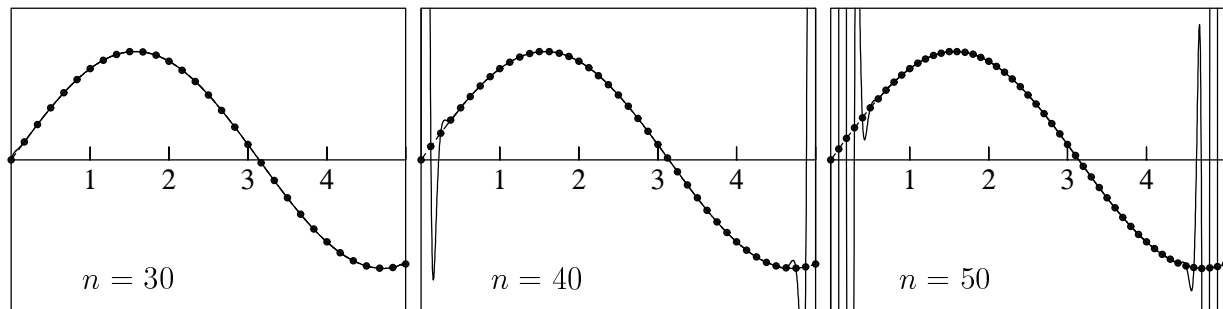
s'appelle la *constante de Lebesgue* associée aux points  $x_0, x_1, \dots, x_n$  et à l'intervalle  $[a, b]$ . Cette constante peut être calculée numériquement (voir les exercices 9 et 10).

**Points équidistants.** Pour la division  $x_k = a + \frac{k}{n}(b - a)$  de l'intervalle  $[a, b]$ , on a

$$\Lambda_{20} \approx 3 \cdot 10^4, \quad \Lambda_{40} \approx 10^{10}, \quad \Lambda_n \approx \frac{2^{n+1}}{e \cdot n \cdot \log n}.$$

**Points de Chebyshev.** Si l'on choisit  $x_k = \frac{a+b}{2} + \frac{b-a}{2} \cos\left(\frac{(2k+1)\pi}{2n+2}\right)$ , on a

$$\begin{aligned} \Lambda_n &\leq 3 && \text{pour } n \leq 20 \\ \Lambda_n &\leq 4 && \text{pour } n \leq 100 \\ \Lambda_n &\approx \frac{2}{\pi} \log n && \text{si } n \text{ est grand.} \end{aligned}$$

FIG. II.12: Influence des erreurs dans les  $y_i$  sur l'interpolation

*Expérience numérique.* Considérons la fonction  $f(x) = \sin x$  sur l'intervalle  $[0, 5]$ . Pour  $n \geq 25$ , l'erreur de l'interpolation est bornée par  $5^{26}/26! \approx 3.69 \cdot 10^{-9}$  quelle que soit la division choisie (voir le théorème 2.2). Prenons les  $y_i = \sin x_i$  en simple précision (erreur relative  $\approx \text{eps} \approx 5.96 \cdot 10^{-8}$ ) et faisons le calcul de  $p(x)$  en double précision. Dans la fig. II.12 nous voyons le polynôme d'interpolation obtenu de cette manière (avec des points équidistants). Son erreur devient bien visible à partir de  $n = 32$ , la valeur où  $\text{eps} \cdot \Lambda_n$  dépasse 1. Evidemment, l'interpolation avec les points de Chebyshev ne montrera pas ce phénomène.

## II.6 Transformée de Fourier discrète (DFT)

Dans le traitement de signaux, où on est confronté à une immense quantité (plusieurs milliers ou millions) de valeurs numériques, la transformée de Fourier est un outil inévitable. De plus, les données ont souvent une certaine périodicité ce qui rend la transformée plus efficace. Elle est, par exemple, très utilisée dans le traitement des sons (voir la fig. II.13) ainsi que pour la compression d'images (voir le paragraphe II.7).

**Transformée de Fourier (continue) ou série de Fourier.** Une série trigonométrique (ou série de Fourier) est une série de la forme

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx). \quad (6.1)$$

En cas de convergence, elle représente une fonction  $2\pi$ -périodique, c.-à-d.  $f(x + 2\pi) = f(x)$  pour tout  $x \in \mathbb{R}$ . Les formules deviennent plus simples en passant aux complexes. Grâce aux identités  $e^{ix} = \cos x + i \sin x$ ,  $\cos x = (e^{ix} + e^{-ix})/2$  et  $\sin x = (e^{ix} - e^{-ix})/2i$ , la série précédente devient simplement

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}. \quad (6.2)$$

La clé fondamentale permettant le calcul des séries trigonométriques a été découverte par Euler (1777, *Opera* vol 16, Pars I, p. 333) et réside dans la *relation d'orthogonalité*

$$\int_0^{2\pi} e^{-i\ell x} \cdot e^{ikx} dx = \int_0^{2\pi} e^{i(k-\ell)x} dx = \begin{cases} 0 & \text{si } k \neq \ell \\ 2\pi & \text{si } k = \ell. \end{cases} \quad (6.3)$$

Cette propriété nous permet de calculer les coefficients  $c_k$ . Il suffit de multiplier (6.2) par  $e^{-i\ell x}$  et intégrer terme par terme<sup>4</sup> de 0 à  $2\pi$ . Tous les termes, sauf un, disparaissent et on obtient

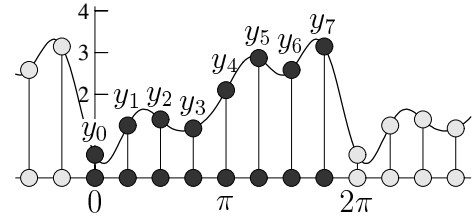
$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx. \quad (6.4)$$

<sup>4</sup>Pour une justification voir les cours *Analyse II* ou *Mathématiques pour Informaticiens*.

Pour marquer la dépendance de  $f(x)$ , nous écrivons souvent  $\hat{f}(k) := c_k$  pour les coefficients de Fourier. Rappelons que les coefficients  $\hat{f}(k)$  convergent vers zéro pour  $k \rightarrow \infty$ , et on a  $\hat{f}(k) = \mathcal{O}(|k|^{-p})$  si le prolongement  $2\pi$ -périodique de  $f(x)$  est  $p$ -fois continûment différentiable.

**Transformée de Fourier discrète (DFT).** Supposons maintenant que la fonction  $2\pi$ -périodique  $f(x)$  soit seulement connue pour les  $x$  de la division équidistante

$$x_j = \frac{2\pi j}{N}, \quad j = 0, 1, \dots, N-1$$



et posons  $y_j = f(x_j)$ . Si nécessaire, on peut prolonger  $\{y_j\}$  à une suite  $N$ -périodique en posant  $y_{j+N} = y_j$  pour tout entier  $j$  ( $N = 8$  dans la petite figure).

Par analogie avec (6.2) on cherche à exprimer cette suite par

$$y_j = \sum_{k=0}^{N-1} z_k e^{ikx_j} = \sum_{k=0}^{N-1} z_k \omega^{kj} \quad \text{avec} \quad \omega = e^{2\pi i/N}. \quad (6.5)$$

Cette fois, la relation d'orthogonalité discrète (observer que  $\omega^N = 1$ )

$$\sum_{j=0}^{N-1} \omega^{-\ell j} \cdot \omega^{kj} = \sum_{j=0}^{N-1} \omega^{(k-\ell)j} = \begin{cases} \frac{1-\omega^{(k-\ell)N}}{1-\omega^{k-\ell}} = 0 & \text{si } k \neq \ell \pmod{N} \\ N & \text{si } k = \ell \pmod{N} \end{cases} \quad (6.6)$$

nous aide à trouver les  $z_k$  à partir de (6.5). Par parfaite analogie avec la preuve ci-dessus pour (6.4), on multiplie l'équation (6.5) par  $\omega^{-\ell j}$  et on additionne de  $j = 0$  à  $j = N-1$ . Tous les termes, sauf un, disparaissent et on obtient la *transformée de Fourier discrète* (DFT)

$$z_k = \frac{1}{N} \sum_{j=0}^{N-1} y_j \omega^{-kj}. \quad (6.7)$$

Si  $y_j = f(x_j)$ , nous écrivons aussi  $\hat{f}_N(k) := z_k$  pour les coefficients de la transformée de Fourier discrète. Comme  $y_N = y_0$ , la valeur  $z_k$  de (6.7) peut être interprétée comme le résultat de la règle du trapèze appliquée à l'intégrale dans (6.4). Toutefois,  $z_k = \hat{f}_N(k)$  n'est pas une approximation de  $c_k = \hat{f}(k)$  pour tout  $k$ , car  $\{z_k\}$  est une suite  $N$ -périodique (ceci est une conséquence immédiate de  $\omega^N = 1$ ) alors que  $\hat{f}(k)$  converge vers zéro si  $k \rightarrow \infty$  (voir le paragraphe II.9 pour une étude détaillée de l'erreur  $\hat{f}(k) - \hat{f}_N(k)$ ).

**Compression des données.** Étant donnée une suite  $N$ -périodique  $\{y_j\}$  et sa transformée de Fourier discrète  $\{z_k\}$ . L'idée est de supprimer dans la représentation (6.5) pour  $y_j$  tous les termes dont la valeur absolue de  $z_k$  est en-dessous d'un certain seuil (par exemple, 3% du coefficient maximal).

*Exemple.* Le premier dessin de la fig. II.13 montre la digitalisation d'un son. On a enregistré 22000 impulsions par seconde, dont 1024 sont dessinées (ceci correspond à  $1024/22 \approx 46.5$  millisecondes). On observe bien une certaine périodicité des données.

Pour les  $N = 1024$  nombres  $\{y_j\}$  on a calculé la transformée de Fourier discrète  $\{z_k\}$ . La suite de leurs valeurs absolues  $\{|z_k|\}$  est dessinée dans la deuxième image de la fig. II.13 pour  $k = 1, \dots, 170$  (comme les  $y_i$  sont réels, on a  $z_{N-k} = z_{-k} = \bar{z}_k$ , et il n'est pas nécessaire de dessiner les valeurs pour  $k \geq N/2$ ; pour  $170 < k < N/2$  les  $|z_k|$  sont inférieurs à 0.072).

La théorie de ce paragraphe est basée sur le fait que  $f(x)$  est une fonction périodique. Cependant, la période du signal n'est visiblement pas égale à  $N = 1024$ , mais elle est plutôt proche de



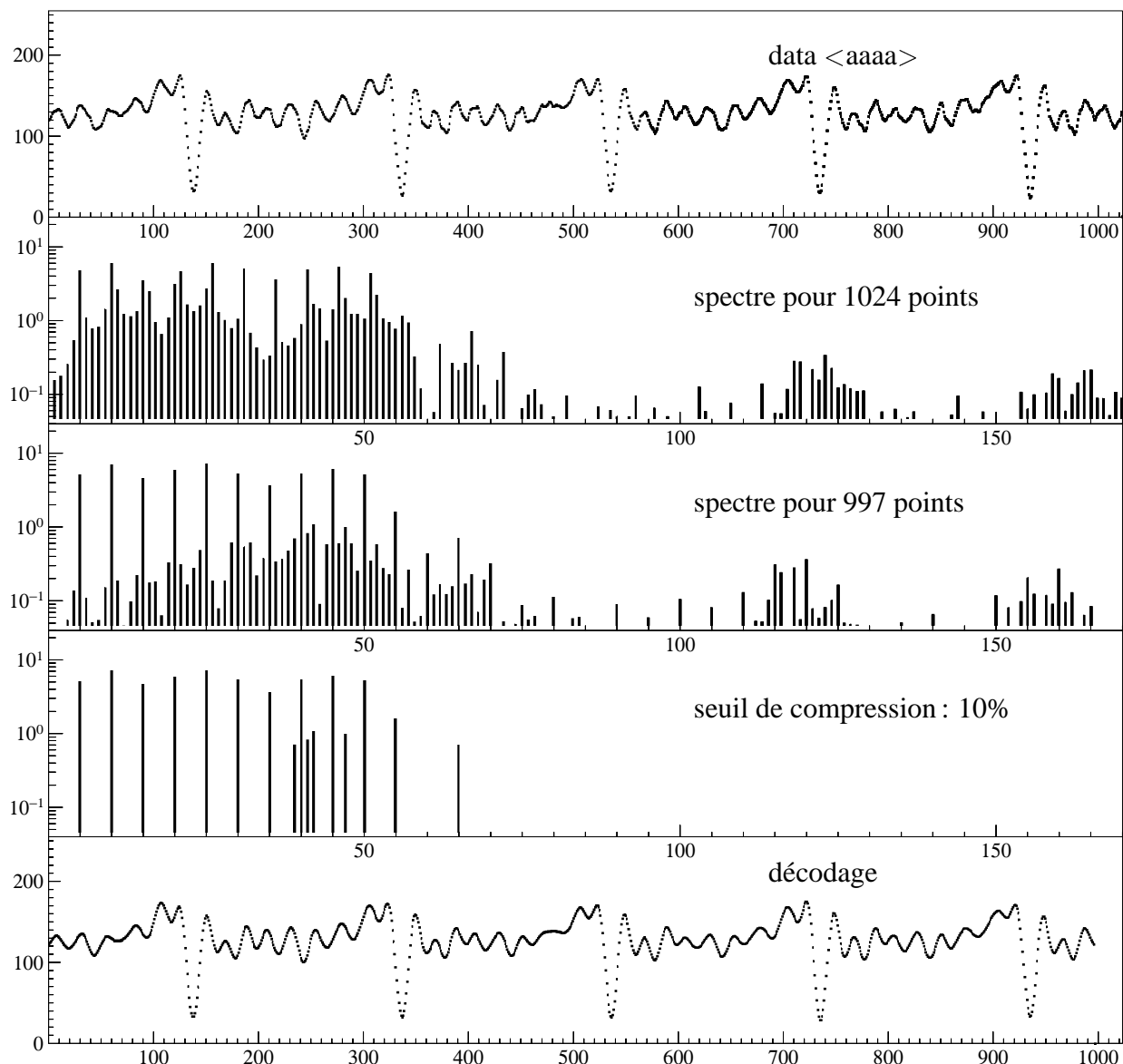


FIG. II.13: Spectrogramme, compression et décodage du son “a” prononcé par Martin

$N = 997$ . Si l’on calcule 1024 valeurs de  $f(x)$  sur une période exacte (par interpolation linéaire) ainsi que leur transformée de Fourier discrète, on obtient la troisième image de la fig. II.13. Cette fois, on peut beaucoup mieux observer la fréquence principale ( $5 * 22000/997 \approx 110$  Hz) ainsi que les harmoniques (multiples de la fréquence principale).

Maintenant nous supprimons tous les coefficients  $z_k$  dont la valeur absolue est inférieure à 10% de la valeur maximale. Les 16 coefficients restant sont dessinés dans le quatrième dessin de la fig. II.13. Pour un seuil de 3% ils resteraient 46 coefficients, et pour un seuil de 1% pas plus de 87. Ainsi, la vraie information contenue dans le signal ne contient que 16 (respectivement 46 ou 87) nombres complexes au lieu des 997 valeurs réelles  $y_j$ .

Pour décoder le signal, nous utilisons la formule (6.5) avec les  $z_k$  restant après la compression. Le résultat (décodage) est dessiné en bas de la fig. II.13. On peut constater que le signal original est très bien reproduit.

La conclusion de cette expérience est la suivante : au lieu de stocker les 997 valeurs du signal original, il suffit de stocker quelques coefficients de la transformée de Fourier discrète sans perdre de l’information visible.

## II.7 Transformée cosinus discrète (DCT) et JPEG

L'algorithme du paragraphe précédent marche très bien si les données ont une certaine périodicité. Si elles ne sont pas périodiques (par exemple pour la compression d'images), on utilise souvent une variante de la transformée de Fourier discrète. Cette variante a en plus l'avantage d'éviter le calcul avec des nombres complexes.

**Transformée de Fourier en cosinus.** Soit  $f(x)$  une fonction continue, définie sur l'intervalle  $[0, \pi]$ . On la prolonge en une fonction paire par  $f(-x) = f(x)$  et en une fonction  $2\pi$ -périodique par  $f(x + 2\pi) = f(x)$ . La série de Fourier (6.1) d'une telle fonction contient seulement les termes en cosinus et s'écrit

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx. \quad (7.1)$$

Sur l'intervalle  $[0, \pi]$ , les fonctions  $\cos kx$  (pour  $k \geq 0$ ) vérifient la *relation d'orthogonalité*

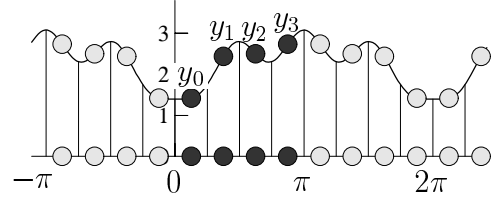
$$\int_0^{\pi} \cos \ell x \cdot \cos kx \, dx = \frac{1}{2} \int_0^{\pi} (\cos(\ell + k)x + \cos(\ell - k)x) \, dx = \begin{cases} 0 & \text{si } k \neq \ell \\ \pi/2 & \text{si } k = \ell \neq 0 \\ \pi & \text{si } k = \ell = 0. \end{cases} \quad (7.2)$$

La démarche habituelle (multiplier (7.1) par  $\cos \ell x$  et intégrer terme par terme de 0 à  $\pi$ ) nous donne

$$a_k = \frac{2}{\pi} \int_0^{\pi} f(x) \cos kx \, dx. \quad (7.3)$$

**Transformée cosinus discrète (DCT).** Comme  $f(0) \neq f(\pi)$  en général, nous considérons les  $N$  points au milieu des sous-intervalles, c.-à-d. les points

$$x_j = \frac{(2j+1)\pi}{2N}, \quad j = 0, 1, \dots, N-1$$



et nous posons  $y_j = f(x_j)$  ( $N = 4$  dans la petite figure). Par analogie avec (7.1), nous exprimons cette suite par (voir la fig. II.14 pour les fonctions de base  $\cos kx_j$ )

$$y_j = \frac{z_0}{2} + \sum_{k=1}^{N-1} z_k \cos kx_j. \quad (7.4)$$

Avec la *relation d'orthogonalité discrète* (pour  $0 \leq k, \ell \leq N-1$ )

$$\sum_{j=0}^{N-1} \cos \ell x_j \cdot \cos kx_j = \frac{1}{2} \sum_{j=0}^{N-1} (\cos(\ell + k)x_j + \cos(\ell - k)x_j) = \begin{cases} 0 & \text{si } k \neq \ell \\ N/2 & \text{si } k = \ell \neq 0 \\ N & \text{si } k = \ell = 0. \end{cases} \quad (7.5)$$

nous trouvons (multiplier l'équation (7.4) par  $\cos \ell x_j$  et additionner de  $j = 0$  à  $j = N-1$ ) la *transformée cosinus discrète (DCT)*

$$z_k = \frac{2}{N} \sum_{j=0}^{N-1} y_j \cos kx_j. \quad (7.6)$$

La valeur  $z_k$  de (7.6) est le résultat de la règle du point milieu appliquée à l'intégrale dans (7.3).

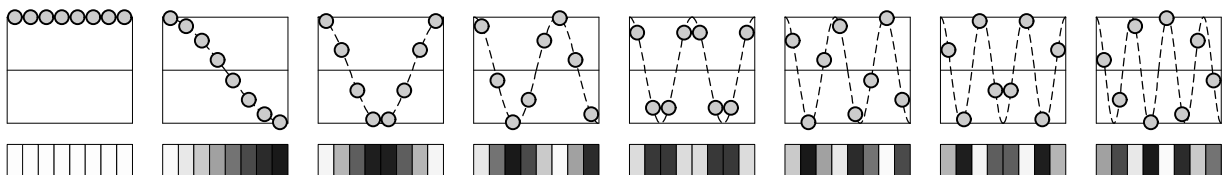
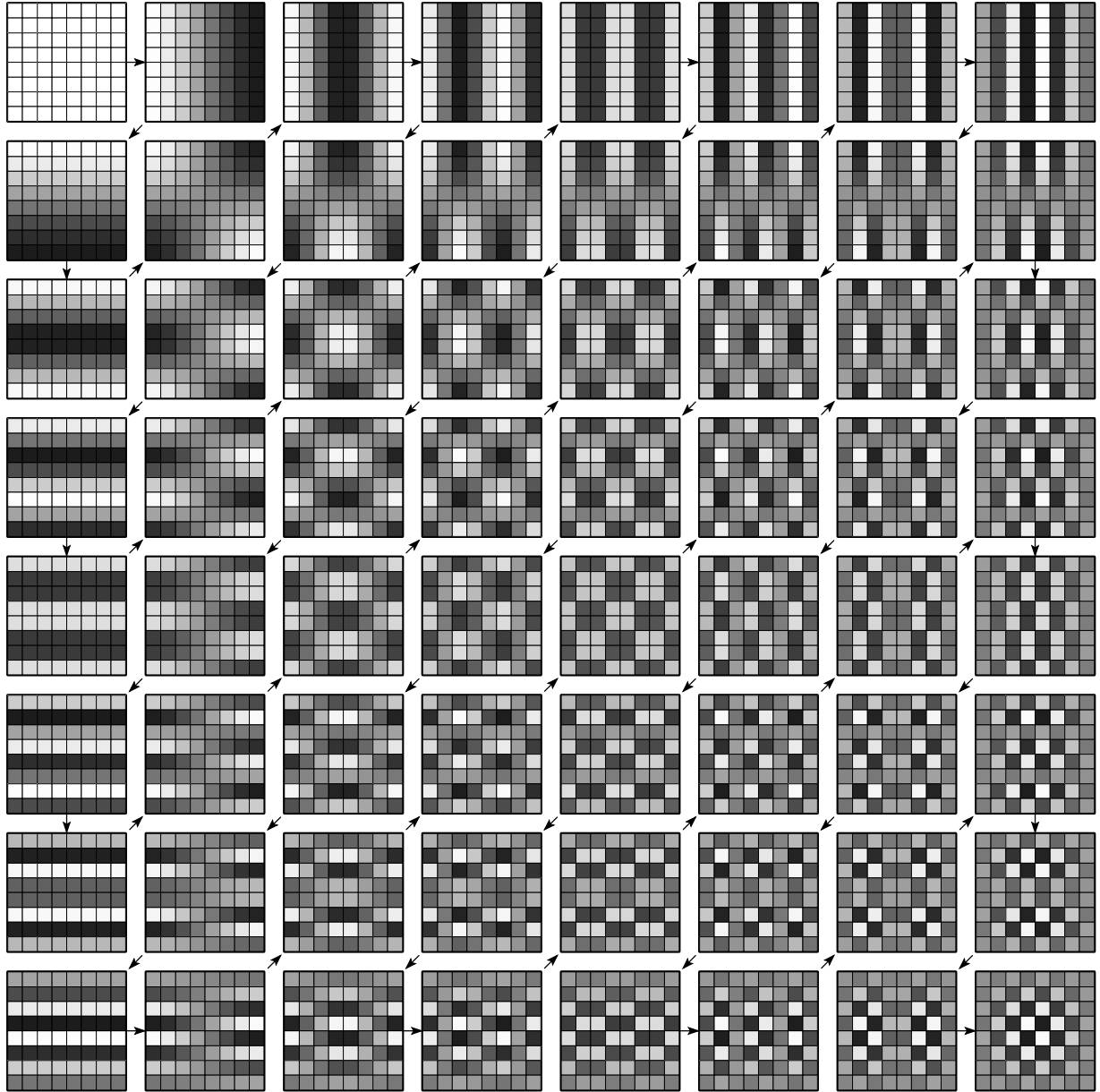


FIG. II.14: Fonctions de base pour la transformée cosinus discrète ( $N = 8$ )

FIG. II.15: Fonctions de base pour la transformée cosinus discrète en 2 dimensions ( $8 \times 8$  pixels)

**Transformée cosinus discrète en dimension 2.** Une image digitale est donnée par un tableau  $\{Y_{i,j}\}$ , où  $i$  parcourt les pixels verticalement et  $j$  horizontalement. La valeur de  $Y_{i,j}$  représente le niveau de gris (ou la couleur) du pixel  $(i, j)$ . Motivé par l'algorithme précédent, nous essayons d'exprimer ces données par

$$Y_{i,j} = \sum_{k=0}^{N-1} \sum_{\ell=0}^{N-1} \tilde{Z}_{k,\ell} \cos kx_i \cos \ell x_j \quad (7.7)$$

où  $x_j = (2j+1)\pi/2N$  comme pour la transformée cosinus discrète. Pour compenser le facteur  $1/2$  dans (7.4), nous utilisons la notation  $\tilde{Z}_{0,0} = Z_{0,0}/4$ ,  $\tilde{Z}_{k,0} = Z_{k,0}/2$ ,  $\tilde{Z}_{0,\ell} = Z_{0,\ell}/2$  et  $\tilde{Z}_{k,\ell} = Z_{k,\ell}$ . La relation d'orthogonalité (7.5) nous permet de calculer les  $Z_{k,\ell}$  par la formule

$$Z_{k,\ell} = \frac{4}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} Y_{i,j} \cos kx_i \cos \ell x_j. \quad (7.8)$$

Les fonctions de base  $\cos kx_i \cos \ell x_j$  ( $k, \ell = 0, \dots, N-1$ ) sont illustrées dans la fig. II.15.

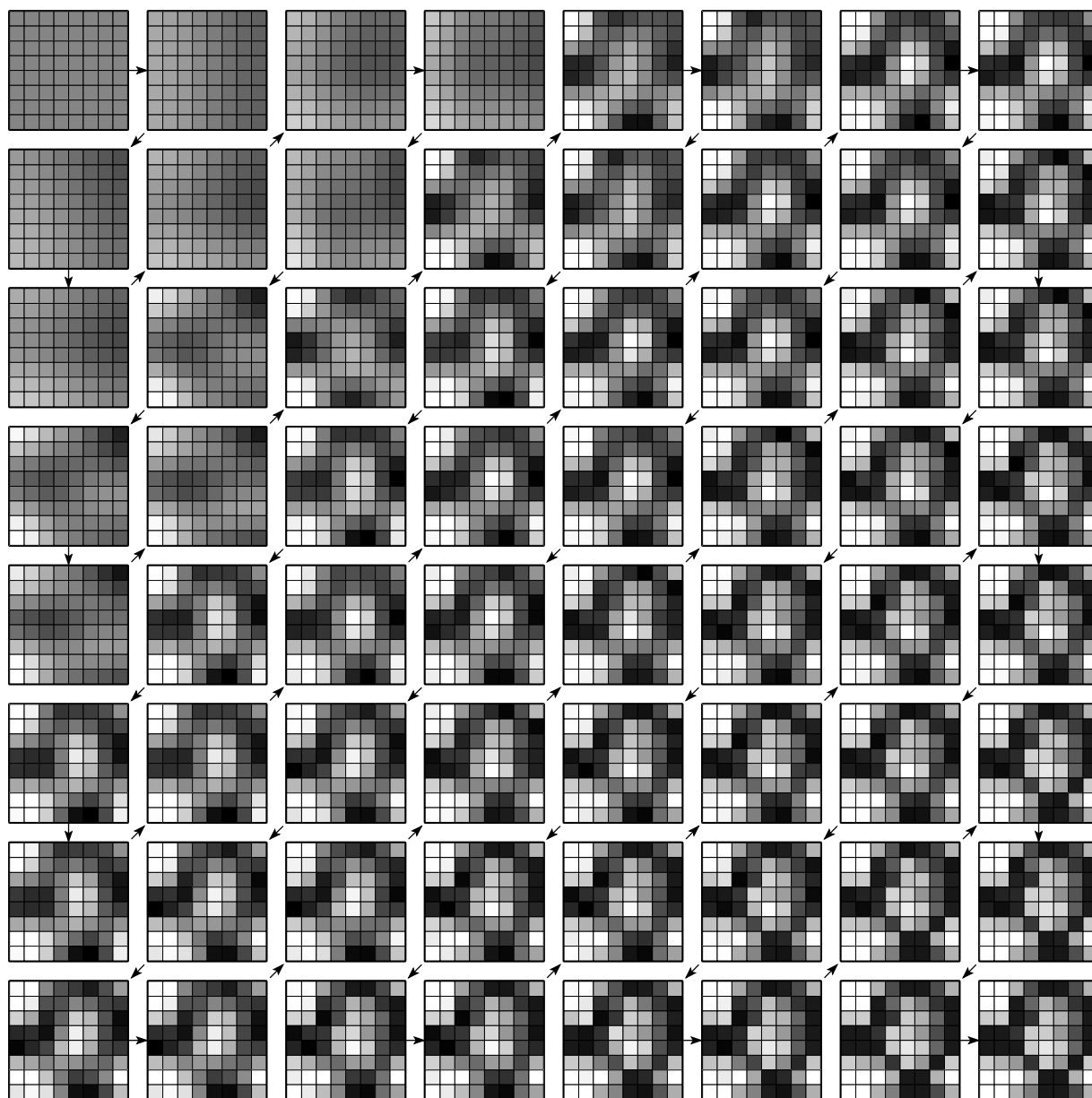
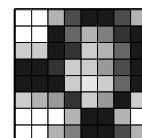


FIG. II.16: Une image toujours mieux reconstituée en zig-zag par JPEG

**JPEG.** L'utilisation de cette base dans la compression d'images est due à Ahmed, Natarajan et Rao (IEEE 1974) <sup>5</sup>. On décompose l'image entière en blocs de  $8 \times 8$  pixels et on calcule pour chaque bloc la transformée cosinus discrète (7.8). Les coefficients  $Z_{k,\ell}$  sont alors quantifiés et ceux qui sont en-dessous d'un seuil sont remplacés par zéro. Pour des images contenant des parties uniformes (par exemple : ciel bleu), seulement deux ou trois coefficients vont rester ce qui donne un important facteur de compression.

La reconstruction de l'image à droite (un conglomérat entre Astérix, Mickey Mouse et Donald Duck, dessiné par Gerhard) est illustrée dans la fig. II.16. Le premier dessin (en haut à gauche) correspond au premier terme de la somme (7.7). En rajoutant un terme après l'autre, et en suivant un chemin astucieux en zig-zag, la reconstruction est démontrée.



<sup>5</sup>Pour des explications plus détaillées et références voir *JPEG still image data compression standard* par W. B. Pennebaker et J. L. Mitchell, Van Nostrand Reinhold, New York, 1992.

## II.8 Transformée de Fourier rapide (FFT)

Pour  $\{y_j\}_{j=0}^{N-1}$  donnés, un calcul direct de la transformée de Fourier discrète  $\{z_k\}_{k=0}^{N-1}$  (voir (6.7)) nécessite  $\approx N^2$  multiplications et additions. Dans ce paragraphe, nous présentons un algorithme (un des “Top 10 Algorithms” du 20ème siècle) qui fait le même travail en  $N \log_2 N$  opérations. Cet algorithme est dû à Cooley & Tukey (1965); il est basé sur des idées de Runge (1925).

Nous supprimons le facteur  $1/N$  dans (6.7) et nous utilisons la notation

$$\mathcal{F}_N y = z \quad \text{où} \quad z_k = \sum_{j=0}^{N-1} y_j \omega_N^{-kj}, \quad \omega_N = e^{2\pi i/N} \quad (8.1)$$

**Lemme 8.1** Soient  $u = (u_0, u_1, \dots, u_{N-1})$ ,  $v = (v_0, v_1, \dots, v_{N-1})$  et définissons

$$y = (u_0, v_0, u_1, v_1, \dots, u_{N-1}, v_{N-1}). \quad (8.2)$$

Alors, pour  $k = 0, 1, \dots, N-1$ , on a ( $\omega_{2N} = e^{2\pi i/2N} = e^{\pi i/N}$ )

$$\begin{aligned} (\mathcal{F}_{2N} y)_k &= (\mathcal{F}_N u)_k + \omega_{2N}^{-k} (\mathcal{F}_N v)_k \\ (\mathcal{F}_{2N} y)_{k+N} &= (\mathcal{F}_N u)_k - \omega_{2N}^{-k} (\mathcal{F}_N v)_k. \end{aligned} \quad (8.3)$$

*Démonstration.* En utilisant  $\omega_{2N}^2 = \omega_N$ , une séparation des indices paires et impaires nous donne pour un  $k$  arbitraire

$$(\mathcal{F}_{2N} y)_k = \sum_{j=0}^{2N-1} y_j \omega_{2N}^{-jk} = \sum_{\ell=0}^{N-1} \underbrace{y_{2\ell}}_{u_\ell} \underbrace{\omega_{2N}^{-2\ell k}}_{\omega_N^{-\ell k}} + \sum_{\ell=0}^{N-1} \underbrace{y_{2\ell+1}}_{v_\ell} \underbrace{\omega_{2N}^{-(2\ell+1)k}}_{\omega_{2N}^{-k} \cdot \omega_N^{-\ell k}} = (\mathcal{F}_N u)_k + \omega_{2N}^{-k} (\mathcal{F}_N v)_k.$$

La deuxième formule de (8.3) résulte de  $\omega_N^N = 1$  et de  $\omega_{2N}^N = -1$ .  $\square$

La formule (8.3) nous permet de calculer (avec  $N$  multiplications et  $2N$  additions) le vecteur  $\mathcal{F}_{2N} y$  à partir de  $\mathcal{F}_N u$  et  $\mathcal{F}_N v$ . La même procédure peut être appliquée récursivement aux suites  $u$  et  $v$  si elles ont une longueur paire. Si l’on suppose que  $N = 2^m$ , on obtient l’algorithme présenté dans le schéma suivant (pour  $N = 8 = 2^3$ )

$$\begin{array}{c} \mathcal{F}_N \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{pmatrix} \begin{array}{l} \swarrow \\ \searrow \end{array} \begin{array}{c} \mathcal{F}_{N/2} \begin{pmatrix} y_0 \\ y_2 \\ y_4 \\ y_6 \end{pmatrix} \\ \mathcal{F}_{N/2} \begin{pmatrix} y_1 \\ y_3 \\ y_5 \\ y_7 \end{pmatrix} \end{array} \begin{array}{l} \swarrow \quad \searrow \\ \swarrow \quad \searrow \end{array} \begin{array}{c} \mathcal{F}_{N/4} \begin{pmatrix} y_0 \\ y_4 \end{pmatrix} \\ \mathcal{F}_{N/4} \begin{pmatrix} y_2 \\ y_6 \end{pmatrix} \\ \mathcal{F}_{N/4} \begin{pmatrix} y_1 \\ y_5 \end{pmatrix} \\ \mathcal{F}_{N/4} \begin{pmatrix} y_3 \\ y_7 \end{pmatrix} \end{array} \begin{array}{l} \swarrow \quad \searrow \\ \swarrow \quad \searrow \\ \swarrow \quad \searrow \\ \swarrow \quad \searrow \end{array} \begin{array}{l} \mathcal{F}_{N/8} y_0 = y_0 \\ \mathcal{F}_{N/8} y_4 = y_4 \\ \mathcal{F}_{N/8} y_2 = y_2 \\ \mathcal{F}_{N/8} y_6 = y_6 \\ \mathcal{F}_{N/8} y_1 = y_1 \\ \mathcal{F}_{N/8} y_5 = y_5 \\ \mathcal{F}_{N/8} y_3 = y_3 \\ \mathcal{F}_{N/8} y_7 = y_7 \end{array} \end{array} \quad (8.4)$$

La programmation de cet algorithme se fait de droit à gauche. D’abord, on met les  $\{y_j\}$  dans l’ordre exigé par l’algorithme (8.4). Après, on effectue les opérations de (8.3) comme indiqué dans le schéma (8.4). Pour une explication détaillée de la programmation voir le livre “Numerical Recipies”.<sup>6</sup>

<sup>6</sup>W.H. Press, B.R. Flannery, S.A. Teukolsky & W.T. Vetterling (1989): *Numerical Recipies. The Art of Scientific Computing* (FORTRAN Version). Cambridge University Press.

TAB. II.2: Comparaison de nombres d'opérations

$N$	$N^2$	$N \log_2 N$	quotient
$2^5 = 32$	$\approx 10^3$	160	$\approx 6.4$
$2^{10} \approx 10^3$	$\approx 10^6$	$\approx 10^4$	$\approx 100$
$2^{20} \approx 10^6$	$\approx 10^{12}$	$\approx 2 \cdot 10^7$	$\approx 5 \cdot 10^4$

Pour passer d'une colonne à une autre (dans le schéma (8.4)) on a besoin de  $N/2$  multiplications complexes et de  $N$  additions (ou soustractions). Comme  $m = \log_2 N$  passages sont nécessaires, on a le résultat suivant.

**Théorème 8.2** *Pour  $N = 2^m$ , le calcul de  $\mathcal{F}_N y$  peut être effectué en  $\frac{N}{2} \log_2 N$  multiplications complexes et  $N \log_2 N$  additions complexes.*

Pour mieux illustrer l'importance de cet algorithme, nous comparons dans le tableau II.2 le nombre d'opérations nécessaires pour le calcul de  $\mathcal{F}_N y$  – avec ou sans FFT.

**L'inverse de la transformée de Fourier discrète.** Pour le décodage il faut calculer les  $y_j$  à partir des  $z_k$  à l'aide de la formule (6.5). Pour en obtenir un algorithme rapide, il suffit de remplacer  $\omega_{2N}^{-k}$  dans (8.3) par  $\omega_{2N}^k$ .

**Transformée cosinus rapide.** Pour les données réelles  $(y_0, y_1, \dots, y_{N-1})$ , nous considérons le vecteur  $\tilde{y} = (y_0, y_1, \dots, y_{N-1}, y_{N-1}, \dots, y_1, y_0)$  illustré dans le petit dessin avant la formule (7.4). La relation  $z_k = \frac{1}{2} \omega_{2N}^{-k/2} (\mathcal{F}_{2N} \tilde{y})_k$  pour la transformée cosinus discrète (7.6) nous permet de trouver un algorithme rapide (même deux fois plus rapide que la FFT).<sup>7</sup>

**Transformée cosinus rapide en dimension 2.** Pour le calcul de (7.8) il faut évaluer la somme sur  $j$  pour chaque  $i \in \{0, 1, \dots, N-1\}$ , et ensuite encore la somme sur  $i$ . Ceci correspond à  $N+1$  transformées cosinus discrètes. Le travail est alors proportionnel à  $(N+1)N \log_2 N$  à la place de  $N^4$  pour un calcul direct. En plus, les  $N$  premières DCT peuvent être calculées en parallèle.

## II.9 Interpolation trigonométrique

Pour la division équidistante

$$x_j = \frac{2\pi j}{N}, \quad j = 0, 1, \dots, N$$

de l'intervalle  $[0, 2\pi]$  et pour  $y_0, y_1, \dots, y_{N-1}$  donnés, on cherche un polynôme trigonométrique (une combinaison linéaire finie de fonctions  $e^{ikx}$ ) passant par  $(x_j, y_j)$  pour  $j = 0, 1, \dots, N-1$ .

**Théorème 9.1 (polynôme d'interpolation trigonométrique)** *Pour  $y_0, y_1, \dots, y_{N-1}$  donnés, soit  $\{z_k\}$  sa transformée de Fourier discrète (6.7). Si  $N$  est pair, le polynôme trigonométrique*

$$p_N(x) = \sum_{k=-N/2}^{N/2} z_k e^{ikx} := \frac{1}{2} (z_{-N/2} e^{-iNx/2} + z_{N/2} e^{iNx/2}) + \sum_{|k| < N/2} z_k e^{ikx} \quad (9.1)$$

satisfait  $p_N(x_j) = y_j$  pour  $j = 0, 1, \dots, N-1$ .

**Démonstration.** La suite  $z_k e^{ikx_j}$  est  $N$ -périodique en  $k$ . Ainsi,  $p_N(x_j) = \sum_{k=-N/2}^{N/2} z_k e^{ikx_j} = \sum_{k=0}^{N-1} z_k e^{ikx_j}$ . Comme  $e^{ikx_j} = \omega_N^{kj}$ , une comparaison avec (6.5) montre que  $p_N(x_j) = y_j$ .  $\square$

<sup>7</sup>Van Loan, *Computational Frameworks for the Fast Fourier Transform*. SIAM, Philadelphia, 1992. [MA 65/326]

*Remarque.* Si les  $y_k$  sont réels, les  $\{z_k\}$  satisfont à  $z_{-k} = \bar{z}_k$  et le polynôme  $p_N(x)$  est une fonction réelle, c.-à-d. une combinaison réelle de  $\sin kx$  et  $\cos kx$ .

**Erreur de l'interpolation trigonométrique.** Supposons maintenant que  $y_j = f(x_j)$  où

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \quad c_k = \hat{f}(k) \quad (9.2)$$

est une série de Fourier qui converge absolument pour tout  $x$ . Pour étudier l'erreur  $f(x) - p_N(x)$ , où  $p_N(x)$  est le polynôme d'interpolation trigonométrique du théorème 9.1, il faut savoir comment les  $z_k = \hat{f}_N(k)$  de la transformée de Fourier discrète approximent les coefficients de Fourier  $c_k = \hat{f}(k)$ .

**Lemme 9.2** Si la série (9.2) est absolument convergente, alors

$$\hat{f}_N(k) - \hat{f}(k) = \sum_{\ell \neq 0} \hat{f}(k + \ell N). \quad (9.3)$$

*Démonstration.* La transformée de Fourier discrète avec  $y_j = f(x_j)$  de (9.2) est donnée par

$$\hat{f}_N(k) = \frac{1}{N} \sum_{j=0}^{N-1} \left( \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{inx_j} \right) \omega_N^{-kj} = \sum_{n=-\infty}^{\infty} \hat{f}(n) \left( \frac{1}{N} \sum_{j=0}^{N-1} \omega_N^{(n-k)j} \right) = \sum_{\ell=-\infty}^{\infty} \hat{f}(k + \ell N).$$

La dernière égalité est une conséquence de la relation d'orthogonalité (6.6).  $\square$

On sait que, si la fonction  $f(x)$  est globalement lisse, les coefficients de Fourier convergent rapidement vers zéro. Pour une telle fonction,  $\hat{f}_N(k)$  est une bonne approximation de  $\hat{f}(k)$  pour  $|k| \leq N/2$ , mais elle est mauvaise pour  $|k| > N/2$ .

**Théorème 9.3** Soit  $f(x)$  donnée par (9.2) avec une série absolument convergente. Alors, le polynôme trigonométrique (9.1) pour  $y_j = f(x_j)$  satisfait pour tout  $x \in \mathbb{R}$

$$|f(x) - p_N(x)| \leq 2 \sum'_{|k| \geq N/2} |\hat{f}(k)|. \quad (9.4)$$

*Démonstration.* On soustrait (9.1) de (9.2) :

$$f(x) - p_N(x) = \sum'_{k=-N/2}^{N/2} (\hat{f}(k) - \hat{f}_N(k)) e^{ikx} + \sum'_{|k| \geq N/2} \hat{f}(k) e^{ikx}.$$

L'assertion est donc une conséquence de (9.3) et de l'inégalité de triangle.  $\square$

Ce théorème permet une interprétation intéressante. Considérons une fonction  $2\pi$ -périodique de fréquence maximale  $M$  (c.-à-d.,  $\hat{f}(k) = 0$  pour  $|k| > M$ ). Alors, le polynôme trigonométrique  $p_N(x)$  donne le résultat exact ( $p_N(x) = f(x)$  pour tout  $x$ ) si

$$N > 2M. \quad (9.5)$$

Ce résultat – le *théorème d'échantillonnage* – nous donne une formule pour le nombre d'échantillons nécessaires pour une représentation exacte d'une telle fonction.

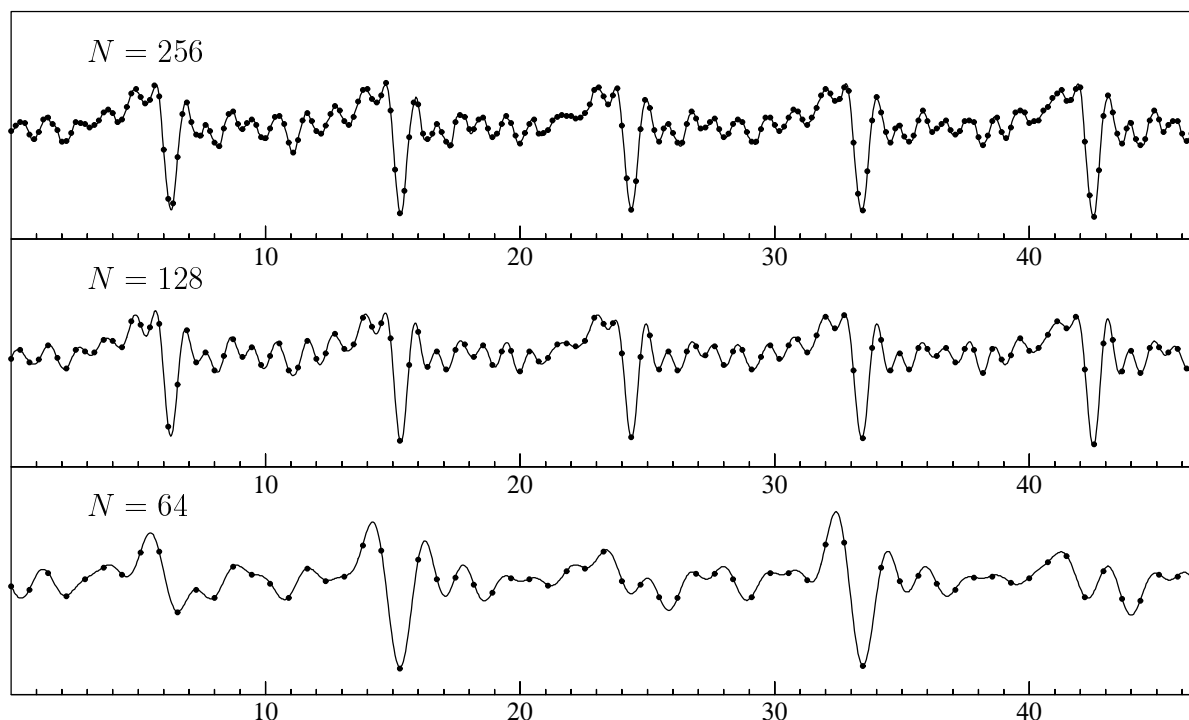


FIG. II.17: Interpolation trigonométrique

Reprenons l'exemple de la fig. II.13. Dans le deuxième dessin de cette figure on voit que les fréquences dominantes du son sont situées dans l'intervalle  $|k| \leq 60$ . Comme la longueur de l'intervalle dans la fig. II.13 est de  $1024/22000$  secondes, la valeur maximale  $M = 60$  correspond à  $60 * 22000/1024 \approx 1390$  Hz. Alors,  $N = 128$  échantillons sont suffisants pour représenter correctement le signal. Dans la fig. II.17 sont dessinés les polynômes trigonométriques  $p_N(x)$  (pour  $N = 64, 128$  et  $256$ ) passant par

$$y_j \quad \text{pour} \quad j = 0 \pmod{1024/N}$$

où  $y_j$  sont les données de la fig. II.13. On voit que la représentation est bonne à partir de  $N = 128$ . Il suffit alors d'utiliser chaque 8<sup>ème</sup> échantillon (une autre possibilité de compression des données).

## II.10 Interpolation par fonctions spline

Le mot “spline” (anglais) signifie “languette élastique”. On s'intéresse à la courbe décrite par une languette forcée de passer par un nombre fini de points donnés (disons par  $(x_i, y_i)$  pour  $i =$

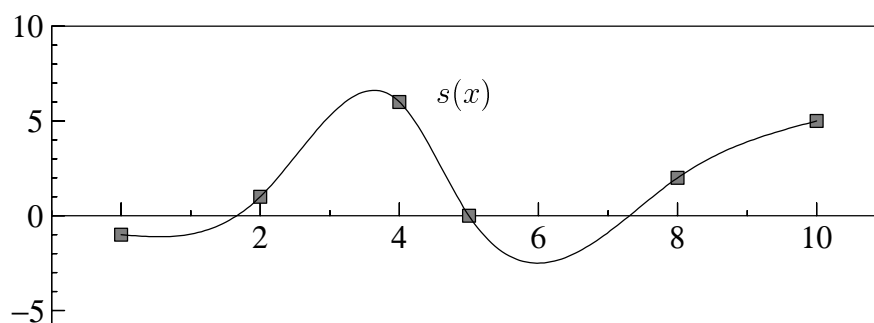


FIG. II.18: Spline cubique (à comparer avec fig. II.1)



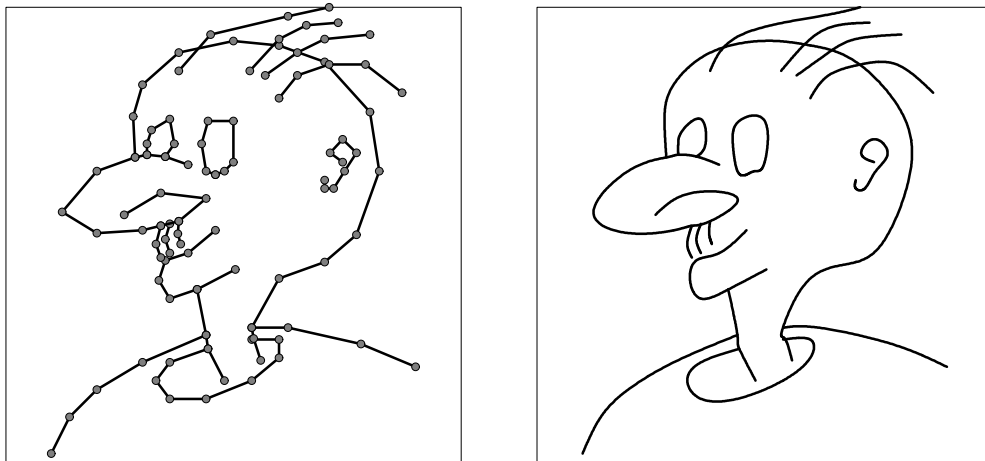


FIG. II.19: Un dessin en zig-zag (à gauche) et en splines (à droite)

$0, 1, \dots, n$ ). La fig. II.18 montre le spline passant par les mêmes données que pour la fig. II.1 (pour pouvoir le comparer avec le polynôme d'interpolation).

La théorie des splines a été développée dans les années 1950 par I.J. Schoenberg pour servir au calcul scientifique (approximations, ...); de nos jours elle est constamment appliquée pour la représentation de courbes et surfaces en Computer Graphics (voir la fig. II.19).

**Formulation mathématique du problème :** pour des points  $(x_i, y_i)$  avec des  $x_i$  ordonnés, on cherche une fonction  $s : [a, b] \rightarrow \mathbb{R}$  (où  $a = x_0, b = x_n$ ) satisfaisant à

- (S1)  $s(x_i) = y_i$  pour  $i = 0, 1, \dots, n$ ;
- (S2)  $s(x)$  est 2 fois continûment différentiable (c.-à-d. de classe  $\mathcal{C}^2$ ) ;
- (S3)  $\int_a^b (s''(x))^2 dx \rightarrow \min.$

L'intégrale dans (S3) représente l'énergie de la languette déformée qui, par le principe de Maupertius, est supposée minimale.

**Théorème 10.1** Soit  $a = x_0 < x_1 < \dots < x_n = b$  une division donnée et soit  $s : [a, b] \rightarrow \mathbb{R}$  une fonction vérifiant (S1), (S2) et qui est un polynôme de degré 3 sur chaque sous-intervalle  $[x_{i-1}, x_i]$ .

Pour toute fonction  $f : [a, b] \rightarrow \mathbb{R}$  vérifiant (S1), (S2) et

$$s''(b)(f'(b) - s'(b)) = s''(a)(f'(a) - s'(a)) \quad (10.1)$$

on a alors que

$$\int_a^b (s''(x))^2 dx \leq \int_a^b (f''(x))^2 dx. \quad (10.2)$$

**Démonstration.** Chaque fonction vérifiant (S1) et (S2) peut être écrite sous la forme  $f(x) = s(x) + \epsilon h(x)$  où  $\epsilon \in \mathbb{R}$  et  $h(x)$  est de classe  $\mathcal{C}^2$  satisfaisant à

$$h(x_i) = 0 \quad \text{pour } i = 0, 1, \dots, n. \quad (10.3)$$

La condition (10.2) devient alors

$$\begin{aligned} \int_a^b (s''(x))^2 dx &\leq \int_a^b (s''(x) + \epsilon h''(x))^2 dx \\ &= \int_a^b (s''(x))^2 dx + 2\epsilon \int_a^b s''(x)h''(x) dx + \epsilon^2 \int_a^b (h''(x))^2 dx. \end{aligned}$$

Si l'on fixe  $h(x)$ , cette condition est satisfaite pour tout  $\epsilon$  (positive et négative) si et seulement si

$$\int_a^b s''(x)h''(x) dx = 0 \quad (10.4)$$

ce qui est équivalent à (après une intégration par parties)

$$s''(x)h'(x) \Big|_a^b - \int_a^b s'''(x)h'(x) dx = 0 \quad (10.5)$$

L'hypothèse (10.1) implique que la première expression de (10.5) est nulle. Comme  $s'''(x)$  est constant sur  $(x_{i-1}, x_i)$ , disons égal à  $\alpha_i$ , la deuxième expression de (10.5) devient

$$\int_a^b s'''(x)h'(x) dx = \sum_{i=1}^n \alpha_i \int_{x_{i-1}}^{x_i} h'(x) dx = \sum_{i=0}^n \alpha_i (h(x_i) - h(x_{i-1})) = 0$$

par (10.3). Ainsi, (10.4) et par conséquent (10.2) aussi sont vérifiés.  $\square$

Le théorème précédent montre que les candidats à la solution de (S1-S3) sont des fonctions de classe  $C^2$  qui sont des polynômes de degré 3 par morceaux.

**Définition 10.2 (spline cubique)** Soit  $a = x_0 < x_1 < \dots < x_n = b$  une division de  $[a, b]$ . Une fonction  $s : [a, b] \rightarrow \mathbb{R}$  s'appelle *spline (cubique)* si elle est 2 fois continûment différentiable et si, sur chaque intervalle  $[x_{i-1}, x_i]$ , elle est un polynôme de degré 3.

Pour satisfaire la condition (10.1), on a plusieurs possibilités:

- *spline naturel*: on suppose que

$$s''(a) = 0 \quad \text{et} \quad s''(b) = 0. \quad (10.6)$$

- *spline scellé*: on suppose données les pentes aux extrémités

$$s'(a) = p_0 \quad \text{et} \quad s'(b) = p_n. \quad (10.7)$$

- *spline périodique*: on suppose que

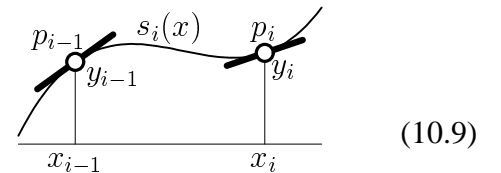
$$s'(a) = s'(b) \quad \text{et} \quad s''(a) = s''(b). \quad (10.8)$$

Evidemment, pour le spline scellé, la condition (10.1) est seulement satisfaite si la fonction  $f(x)$  vérifie aussi la condition (10.7). Alors,  $s(x)$  minimise l'intégrale de (S3) seulement dans la classe de fonctions dont les pentes sont fixées aux extrémités. Pour le spline périodique, la situation est analogue.

Le but suivant est de dériver une construction du spline vérifiant  $s(x_i) = y_i$  pour  $i = 0, 1, \dots, n$  et une des conditions (10.6)-(10.8).

**Interpolation d'Hermite.** Considérons un seul sous-intervalle  $[x_{i-1}, x_i]$  et cherchons un polynôme  $s_i(x)$  de degré 3 vérifiant

$$\begin{aligned} s_i(x_{i-1}) &= y_{i-1}, & s_i(x_i) &= y_i, \\ s'_i(x_{i-1}) &= p_{i-1}, & s'_i(x_i) &= p_i. \end{aligned}$$



(10.9)

La solution peut être obtenue par la formule de Newton en remplaçant les deux dernières conditions de (10.9) par  $s_i(x_{i-1} + \epsilon) = y_{i-1} + \epsilon p_{i-1}$  et  $s_i(x_i - \epsilon) = y_i - \epsilon p_i$ , et en considérant la limite  $\epsilon \rightarrow 0$ . Les différences divisées correspondantes aux données (10.9) sont présentées dans le tableau II.3 ( $h_{i-1} := x_i - x_{i-1}$ ). En utilisant les valeurs encadrées pour la formule de Newton, on obtient

$$\begin{aligned} s_i(x) &= y_{i-1} + (x - x_{i-1}) \delta y[x_i, x_{i-1}] \\ &\quad + \frac{(x - x_{i-1})(x - x_i)}{h_{i-1}^2} \left( (p_i - \delta y[x_i, x_{i-1}])(x - x_{i-1}) + (p_{i-1} - \delta y[x_i, x_{i-1}])(x - x_i) \right). \end{aligned} \quad (10.10)$$

TAB. II.3: Différences divisées pour l'interpolation d'Hermite

$x_{i-1}$	$y_{i-1}$	$p_{i-1}$		
$x_{i-1}$	$y_{i-1}$		$h_{i-1}^{-1}(\delta y[x_i, x_{i-1}] - p_{i-1})$	
		$\delta y[x_i, x_{i-1}]$		$h_{i-1}^{-2}(p_i + p_{i-1} - 2\delta y[x_i, x_{i-1}])$
$x_i$	$y_i$		$h_{i-1}^{-1}(p_i - \delta y[x_i, x_{i-1}])$	
$x_i$	$y_i$	$p_i$		

**Construction du spline interpolant.** Pour chaque choix de pentes  $p_0, p_1, \dots, p_n$ , la fonction  $s : [a, b] \rightarrow \mathbb{R}$ , définie par  $s(x) = s_i(x)$  pour  $x \in [x_{i-1}, x_i]$ , satisfait à

- a)  $s(x_i) = y_i$  pour  $i = 0, 1, \dots, n$ ;
- b)  $s(x)$  est de classe  $C^1$  et  $s'(x_i) = p_i$  pour  $i = 0, 1, \dots, n$ ;
- c) sur chaque intervalle  $[x_{i-1}, x_i]$ ,  $s(x)$  est un polynôme de degré 3.

Pour construire le spline interpolant, il reste à déterminer les pentes  $p_0, p_1, \dots, p_n$  de manière à ce que  $s''(x)$  soit continue, c.-à-d.,

$$s''_i(x_i) = s''_{i+1}(x_i), \quad i = 1, \dots, n-1, \quad (10.11)$$

et qu'une des conditions (10.6)-(10.8) soit satisfaite. En dérivant (10.10) deux fois, on obtient

$$\begin{aligned} s''_i(x_i) &= \frac{2}{h_{i-1}}(2p_i + p_{i-1} - 3\delta y[x_i, x_{i-1}]) \\ s''_i(x_{i-1}) &= -\frac{2}{h_{i-1}}(p_i + 2p_{i-1} - 3\delta y[x_i, x_{i-1}]). \end{aligned}$$

La condition (10.11) devient alors (pour  $i = 1, \dots, n-1$ )

$$\frac{p_{i-1}}{h_{i-1}} + p_i \cdot 2 \cdot \left( \frac{1}{h_{i-1}} + \frac{1}{h_i} \right) + \frac{p_{i+1}}{h_i} = 3 \left( \frac{\delta y[x_i, x_{i-1}]}{h_{i-1}} + \frac{\delta y[x_{i+1}, x_i]}{h_i} \right). \quad (10.12)$$

Ceci donne  $n-1$  équations linéaires pour les  $n+1$  inconnues  $p_0, p_1, \dots, p_n$ . Les deux dernières conditions sont données par le type du spline. Par exemple, pour le *spline scellé*, les valeurs de  $p_0$  et  $p_n$  sont explicitement données et le système linéaire ainsi obtenu s'écrit matriciellement

$$\underbrace{\begin{pmatrix} 2\left(\frac{1}{h_0} + \frac{1}{h_1}\right) & \frac{1}{h_1} & & & \\ \frac{1}{h_1} & 2\left(\frac{1}{h_1} + \frac{1}{h_2}\right) & \frac{1}{h_2} & & \\ & \frac{1}{h_2} & 2\left(\frac{1}{h_2} + \frac{1}{h_3}\right) & \ddots & \\ & & \ddots & \ddots & \frac{1}{h_{n-2}} \\ & & & \frac{1}{h_{n-2}} & 2\left(\frac{1}{h_{n-2}} + \frac{1}{h_{n-1}}\right) \end{pmatrix}}_A \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_{n-1} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_{n-1} \end{pmatrix} \quad (10.13)$$

avec des  $c_i$  donnés par (10.12). La matrice  $A$  est symétrique et tridiagonale.

**Théorème 10.3** Soit  $Ap = c$  avec des  $c_i$  satisfaisant  $|c_i| \leq \gamma$  pour tout  $i$ . Alors,

$$|p_i| \leq \frac{h}{2} \gamma \quad \text{où} \quad h = \max_{i=0, \dots, n-1} h_i. \quad (10.14)$$

En particulier,  $A$  est inversible (poser  $\gamma = 0$ ).

*Démonstration.* Soit  $p$  une solution de  $Ap = c$  et choisissons l'indice  $\ell$  de manière à ce que  $|p_\ell| \geq |p_j|$  pour tout  $j$ . La ligne  $\ell$  du système  $Ap = c$  donne

$$p_\ell \cdot 2 \cdot \left( \frac{1}{h_{\ell-1}} + \frac{1}{h_\ell} \right) = -\frac{p_{\ell-1}}{h_{\ell-1}} - \frac{p_{\ell+1}}{h_\ell} + c_\ell$$

et, en prenant sa valeur absolue,  $|p_\ell| \cdot 2 \cdot \left( \frac{1}{h_{\ell-1}} + \frac{1}{h_\ell} \right) \leq |p_\ell| \cdot \left( \frac{1}{h_{\ell-1}} + \frac{1}{h_\ell} \right) + \gamma$ , on en déduit

$$|p_\ell| \leq \frac{h_\ell \cdot h_{\ell-1}}{h_\ell + h_{\ell-1}} \gamma \leq \frac{\max(h_\ell, h_{\ell-1})}{2} \gamma \leq \frac{h}{2} \gamma$$

car  $h_\ell + h_{\ell-1} \geq 2 \cdot \min(h_\ell, h_{\ell-1})$ . □

*Conclusion.* Le spline scellé *existe* toujours et il est *unique*. La même démonstration s'applique aussi au spline naturel et au spline périodique.

La résolution du système linéaire (10.13) se fait par élimination. On élimine la variable  $p_1$  dans la ligne 2 à l'aide de la ligne 1, puis la variable  $p_2$  dans la ligne 3 à l'aide de la ligne 2, etc. On obtient alors un système bidiagonal qui est facile à résoudre.

## II.11 L'erreur du spline

Soient une fonction différentiable  $f : [a, b] \rightarrow \mathbb{R}$  et une division  $a = x_0 < x_1 < \dots < x_n = b$ . Considérons le spline  $s(x)$  satisfaisant (spline scellé)

$$s(x_i) = f(x_i), \quad i = 0, 1, \dots, n \quad (8.1a)$$

$$s'(x_0) = f'(x_0), \quad s'(x_n) = f'(x_n). \quad (8.1b)$$

Ce spline est donné par (10.10) où les coefficients  $p_i$  satisfont (10.12),  $p_0 = f'(x_0)$  et  $p_n = f'(x_n)$ . Le but est d'étudier l'erreur  $f(x) - s(x)$  pour  $x \in [a, b]$ .

L'idée importante est de considérer (sur  $[x_{i-1}, x_i]$ ) également le polynôme d'interpolation d'Hermite  $q_i(x)$ , défini par

$$q_i(x_{i-1}) = f(x_{i-1}), \quad q_i(x_i) = f(x_i), \quad q'_i(x_{i-1}) = f'(x_{i-1}), \quad q'_i(x_i) = f'(x_i). \quad (11.1)$$

Il est aussi donné par (10.10) si l'on remplace  $p_j$  par  $f'(x_j)$ . On va estimer séparément les deux termes dans la formule

$$f(x) - s_i(x) = (f(x) - q_i(x)) + (q_i(x) - s_i(x)). \quad (11.2)$$

**Théorème 11.1** Soit  $f(x)$  de classe  $\mathcal{C}^4$  et  $q_i(x)$  le polynôme de degré 3 satisfaisant (11.1) (interpolation d'Hermite). Alors, pour  $x \in [x_{i-1}, x_i]$ , on a

$$|f(x) - q_i(x)| \leq \frac{h_{i-1}^4}{384} \cdot \max_{\xi \in [x_{i-1}, x_i]} |f^{(4)}(\xi)|. \quad (11.3)$$

*Démonstration.* Si l'on calcule le polynôme d'interpolation de degré 3 passant par

$$(x_{i-1}, f(x_{i-1})), \quad (x_{i-1} + \epsilon, f(x_{i-1} + \epsilon)), \quad (x_i - \epsilon, f(x_i - \epsilon)), \quad (x_i, f(x_i))$$

avec la formule de Newton, on obtient pour  $\epsilon \rightarrow 0$  exactement le polynôme  $q_i(x)$  (voir le tableau II.3 des différences divisées). L'erreur de ce polynôme peut être majorée par (voir le para-

graphe II.2)

$$|(x - x_{i-1})(x - x_{i-1} - \epsilon)(x - x_i + \epsilon)(x - x_i)| \cdot \frac{|f^{(4)}(\xi)|}{4!}.$$

Donc, pour  $\epsilon \rightarrow 0$ ,

$$|f(x) - q_i(x)| \leq (x - x_{i-1})^2(x - x_i)^2 \cdot \frac{1}{4!} \cdot \max_{\xi \in [x_{i-1}, x_i]} |f^{(4)}(\xi)|.$$

Comme la fonction  $(x - x_{i-1})^2(x - x_i)^2$  possède son maximum au milieu de l'intervalle  $[x_{i-1}, x_i]$ , on obtient l'estimation (11.3).  $\square$

Pour estimer la deuxième expression de (11.2), nous soustrayons la formule (10.10) pour  $s_i(x)$  de la formule analogue pour  $q_i(x)$  et nous obtenons

$$q_i(x) - s_i(x) = \frac{(x - x_{i-1})(x - x_i)}{h_{i-1}^2} \left( (f'(x_i) - p_i)(x - x_{i-1}) + (f'(x_{i-1}) - p_{i-1})(x - x_i) \right). \quad (11.4)$$

Il nous faut encore une estimation de  $f'(x_i) - p_i$ .

**Lemme 11.2** Soit  $f(x)$  de classe  $\mathcal{C}^4$  sur  $[a, b]$  et notons  $h = \max_i h_i$  où  $h_i = x_{i+1} - x_i$ . Si les  $p_i$  satisfont (10.12),  $p_0 = f'(x_0)$  et  $p_n = f'(x_n)$ ,

$$|f'(x_i) - p_i| \leq \frac{h^3}{24} \cdot \max_{x \in [a, b]} |f^{(4)}(x)|. \quad (11.5)$$

Si la division est équidistante et si  $f \in \mathcal{C}^5[a, b]$ , alors

$$|f'(x_i) - p_i| \leq \frac{h^4}{60} \cdot \max_{x \in [a, b]} |f^{(5)}(x)|. \quad (11.6)$$

*Démonstration.* Voici les idées pour le cas équidistant. Les  $p_i$  sont définis par (voir (10.12))

$$\frac{1}{h} (p_{i-1} + 4p_i + p_{i+1}) - \frac{3}{h^2} (f(x_{i+1}) - f(x_{i-1})) = 0. \quad (11.7)$$

Pour estimer la différence  $f'(x_i) - p_i$ , nous calculons le défaut qu'on obtient en remplaçant  $p_i$  par  $f'(x_i)$  dans (11.7) :

$$\frac{1}{h} (f'(x_{i-1}) + 4f'(x_i) + f'(x_{i+1})) - \frac{3}{h^2} (f(x_{i+1}) - f(x_{i-1})) =: d_i. \quad (11.8)$$

En utilisant la formule de Taylor (voir le cours d'Analyse I) pour  $f(x_i \pm h)$  et  $f'(x_i \pm h)$ , on obtient

$$d_i = h^3 \int_0^1 \left( \frac{(1-t)^3}{3!} - 3 \frac{(1-t)^4}{4!} \right) (f^{(5)}(x_i + th) + f^{(5)}(x_i - th)) dt.$$

Comme la fonction  $(1-t)^3/3! - 3(1-t)^4/4!$  ne change pas de signe sur  $[0, 1]$ , ceci nous donne l'estimation

$$|d_i| \leq h^3 \underbrace{\int_0^1 \left( \frac{(1-t)^3}{3!} - 3 \frac{(1-t)^4}{4!} \right) dt}_{= 1/30} \cdot \max_{x \in [x_{i-1}, x_{i+1}]} |f^{(5)}(x)|.$$

Notons  $d = (d_1, \dots, d_{n-1})^T$  et  $e = (e_1, \dots, e_{n-1})^T$  où  $e_i = f'(x_i) - p_i$ . En prenant la différence entre (11.8) et (11.7), on obtient  $Ae = d$  où  $A$  est la matrice de (10.13). Le théorème 10.3 implique

$$|e_i| \leq \frac{h}{2} \max_j |d_j| \leq \frac{h^4}{60} \cdot \max_{x \in [a, b]} |f^{(5)}(x)|,$$

ce qu'il fallait démontrer.  $\square$

**Théorème 11.3 (erreur du spline scellé)** Soit  $f : [a, b] \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^4$ ,  $a = x_0 < x_1 < \dots < x_n = b$  une division arbitraire et  $s(x)$  le spline qui passe par  $(x_i, f(x_i))$  pour  $i = 0, 1, \dots, n$  et qui satisfait  $s'(x_0) = f'(x_0)$  et  $s'(x_n) = f'(x_n)$ . Alors, avec  $h_i = x_{i+1} - x_i$  et  $h = \max_i h_i$  on a

$$|f(x) - s(x)| \leq \frac{5}{384} \cdot h^4 \cdot \max_{\xi \in [a, b]} |f^{(4)}(\xi)|. \quad (11.9)$$

Si de plus la division est équidistante et si  $f \in \mathcal{C}^5[a, b]$ ,

$$|f(x) - s(x)| \leq \frac{h^4}{384} \cdot \max_{\xi \in [a, b]} |f^{(4)}(\xi)| + \frac{h^5}{240} \cdot \max_{\xi \in [a, b]} |f^{(5)}(\xi)|. \quad (11.10)$$

*Démonstration.* Considérons de nouveau le cas équidistant. Pour  $x \in [x_{i-1}, x_i]$ , nous estimons séparément les deux termes dans (11.2). L'estimation du premier terme résulte de (11.3). Pour le deuxième terme, nous utilisons (11.4) et (11.6). Ceci donne

$$|q_i(x) - s_i(x)| \leq \frac{(x - x_{i-1})(x_i - x)}{h_{i-1}^2} \cdot h_{i-1} \cdot \frac{h^4}{60} \cdot \max_{\xi \in [a, b]} |f^{(5)}(\xi)| \leq \frac{h^5}{4 \cdot 60} \cdot \max_{\xi \in [a, b]} |f^{(5)}(\xi)|. \quad \square$$

**Exemple.** Considérons encore une fois (voir les fig. II.5 et II.6) la fonction

$$f(x) = \frac{1}{1 + 25x^2} \quad \text{sur} \quad [-1, 1] \quad (11.11)$$

et calculons le spline interpolant (scellé) pour la division équidistante  $x_i = -1 + 2i/n$ ,  $i = 0, 1, \dots, n$ . Dans la fig. II.20, les 4 dessins du haut montrent le spline  $s(x)$  pour  $n = 3$ ,  $n = 9$ ,  $n = 27$  et  $n = 81$ . La fonction  $f(x)$  est dessinée en pointillés. Les 4 dessins du bas montrent les erreurs. Cette fois, la fonction  $h^4 \cdot f^{(4)}(x)/384$  ( $h = 2/n$ ) est incluse en pointillés (on a choisi l'échelle sur l'axe  $y$  de manière à ce que  $h^4 \cdot f^{(4)}(0)/384$  soit toujours au même endroit). Pour des petits  $h$ , quand le deuxième terme de (11.10) est négligeable, on peut très bien observer la validité de l'estimation (11.10).

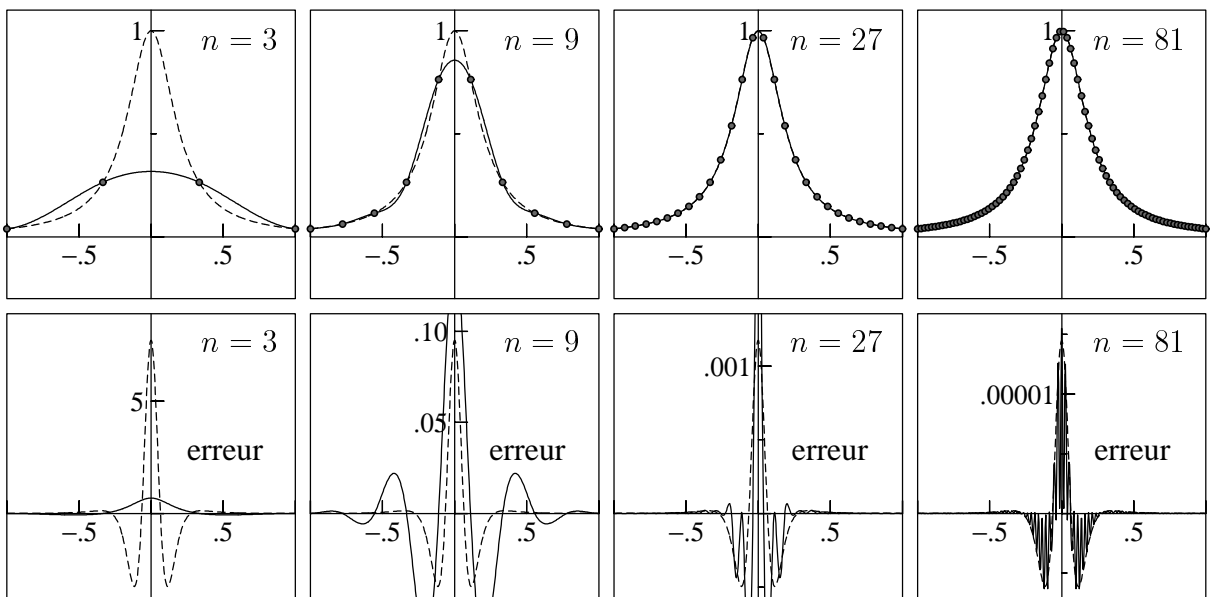


FIG. II.20: Spline interpolant (scellé) pour la fonction (11.11)

## II.12 Exercices

1. Calculer le polynôme d'interpolation passant par les points

$$(0, 0), \quad (1, 3), \quad (3, 1), \quad (5, 2), \quad (8, 2),$$

en utilisant la formule de Newton.

2. Démontrer que la différence divisée  $\delta^n y[x_0, x_1, \dots, x_n]$  est une fonction symétrique; c.-à-d.

$$\delta^n y[x_{\sigma(0)}, x_{\sigma(1)}, \dots, x_{\sigma(n)}] = \delta^n y[x_0, x_1, \dots, x_n]$$

pour toute permutation  $\sigma$  de  $\{0, 1, \dots, n\}$ .

3. Nous savons que l'erreur pour l'interpolation linéaire de  $f$  aux points  $x_0, x_1$  est

$$f(x) - p(x) = (x - x_0)(x - x_1) \frac{f''(\zeta(x))}{2}, \quad x_0 < x < x_1,$$

si  $f \in \mathcal{C}^2[x_0, x_1]$ . Déterminer la fonction  $\zeta(x)$  explicitement dans le cas où  $f(x) = \frac{1}{x}$ ,  $x_0 = 1$ ,  $x_1 = 2$ , et trouver  $\max_{1 \leq x \leq 2} \zeta(x)$  et  $\min_{1 \leq x \leq 2} \zeta(x)$ .

4. On veut tabuler la fonction  $y = \sin x$  aux points équidistants  $x_j = jh$ ,  $j \geq 0$ .

(a) Majorer l'erreur d'interpolation dans l'intervalle  $[x_i, x_{i+1}]$  lorsqu'on fait passer un polynôme de degré 3 par  $x_{i-1}, x_i, x_{i+1}, x_{i+2}$ .

(b) Quel  $h$  peut-on prendre pour que cette erreur soit  $\leq 10^{-8}$  pour tout  $i > 0$ ?

5. (a) Montrer que la formule de quadrature

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx \approx \frac{\pi}{s} \sum_{k=1}^s f(c_k), \quad c_k = \cos\left(\frac{(2k-1)\pi}{2s}\right) \quad (12.1)$$

est exacte pour tout polynôme  $f(x)$  de degré  $\leq 2s - 1$ .

*Indication.* Vérifier (12.1) pour les polynômes de Chebyshev  $T_0(x), \dots, T_{2s-1}(x)$ .

(b) Effectuez le changement de variables  $x = \cos(t)$  dans (12.1), à quelle formule de quadrature se ramène t'on ?

6. Considérer la fonction  $f(x) = x^3$  sur l'intervalle  $[-1, 1]$ .

(a) Déterminer  $p$  tel que la droite  $d(x) = px$  satisfait

$$\max_{x \in [-1, 1]} |f(x) - d(x)| \rightarrow \min.$$

(b) Pour quels  $x_i$ , la droite trouvée peut être interprétée comme un polynôme d'interpolation pour  $f(x)$  ?

(c) Y-a-t'il une relation avec les points de Chebyshev? Si oui, expliquer laquelle.

7. Pour un polynôme  $p(x) = a_0 + a_1x + \dots + a_nx^n$ , l'algorithme de Horner permet de calculer  $b_0 = p(x_0)$  par :

$$\begin{aligned} b_n &= a_n \\ b_i &= a_i + x_0 b_{i+1}, \quad i = n-1, \dots, 1, 0. \end{aligned}$$

Notons  $q(x) = b_1 + b_2x + \dots + b_nx^{n-1}$ .

a) Montrer que  $p(x) = b_0 + (x - x_0)q(x)$  et que  $p'(x_0) = q(x_0)$ .

b) Généraliser l'algorithme de Horner pour calculer  $p(x_0)$  et  $p'(x_0)$  en même temps.

8. Sur l'intervalle  $[-1, 1]$ , on considère la fonction

$$f(x) = \frac{3x - 4}{(x^2 - 9)(5x^2 - 8x + 4)}.$$

Faire une représentation graphique de  $f(x)$  et calculer ses pôles. Utiliser la figure II.7 pour déterminer approximativement l'intervalle maximal  $[\alpha, \beta] \subset [-1, 1]$  où la suite des polynômes d'interpolations (pour les divisions équidistantes  $x_i = -1 + 2i/n$ ) converge (quand  $n \rightarrow \infty$ ).

9. Pour une division  $x_0 < x_1 < \dots < x_n$  ( $n \geq 2$ ), étudier la fonction

$$\Lambda(x) = \sum_{i=0}^n |\ell_i(x)|, \quad \ell_i(x) = \prod_{j \neq i} \frac{(x - x_j)}{(x_i - x_j)}. \quad (12.2)$$

a) Dessiner  $\Lambda(x)$  pour  $n = 2$  en considérant la division  $\{-1, 0, 1\}$ .

b) Montrer que, sur l'intervalle  $[x_{j-1}, x_j]$ , on a l'identité

$$\Lambda|_{[x_{j-1}, x_j]}(x) = \sum_{i=0}^n \epsilon_i \ell_i(x), \quad \text{avec} \quad \epsilon_i = \begin{cases} (-1)^{j-i+1} & \text{si } i \leq j-1 \\ (-1)^{i-j} & \text{si } i \geq j. \end{cases} \quad (12.3)$$

c) Montrer que la fonction  $\Lambda(x)$  de (12.2) ne possède qu'un seul maximum local sur chaque intervalle  $[x_{j-1}, x_j]$ .

*Indication.* Etudier les extréma du polynôme (12.3).

10. Calculer les constantes de Lebesgue

$$\Lambda_n = \max_{x \in [-1, 1]} \sum_{i=0}^n |\ell_i(x)|, \quad n = 1, 2, 3, \dots$$

a) pour la division équidistante  $x_k = -1 + 2k/n$ ,

b) pour les points de Chebyshev  $x_k = -\cos((2k+1)\pi/(2n+2))$ .

Pour calculer le maximum de la fonction  $f(x) = \sum_{i=0}^n |\ell_i(x)|$  sur  $[x_{j-1}, x_j]$  utiliser la recherche de Fibonacci.

11. Calculer à la main la transformée de Fourier discrète de la suite  $\{0, 1, 2, 3, 0, -3, -2, -1\}$ .

12. Pour une fonction continue  $f : [a, b] \rightarrow \mathbb{R}$  et une formule de quadrature  $(b_i, c_i)_{i=1}^s$  à poids  $b_i$  positifs, estimer la différence

$$I_\Delta - \hat{I}_\Delta,$$

où  $I_\Delta$  est le résultat pour la division  $\Delta = \{a = x_0 < x_1 < \dots < x_n = b\}$  (sans tenir compte des erreurs d'arrondi), et  $\hat{I}_\Delta$  est le résultat obtenu en tenant compte des erreurs d'arrondi dans l'évaluation de  $f(x)$ , c.-à-d. en remplaçant  $f(x_j + c_i h_j)$  par

$$\hat{y}_{ij} = f(x_j + c_i h_j)(1 + \epsilon_{ij}), \quad |\epsilon_{ij}| \leq \epsilon.$$

Justifier la condition (6.1) du paragraphe I.6.

13. (a) Calculer les coefficients  $\hat{f}(k)$  de la série de Fourier pour la fonction  $2\pi$ -périodique

$$f(x) = \begin{cases} 4x/\pi & \text{pour } |x| < \pi, \\ 0 & \text{pour } x = \pi. \end{cases}$$

(b) Calculer la transformée de Fourier discrète pour  $\{y_k\}_{k=0}^N$  où  $y_l = f(2\pi l/N)$ ,  $l = 0, \dots, N$  avec  $f$ , la fonction de l'exercice précédent.

(c) Vérifier le résultat obtenu dans l'exercice 11 pour  $N = 8$ .

(d) Estimer la différence  $|z_k - \hat{f}_N(k)|$ .



14. Pour deux suites  $N$ -périodiques  $y$  et  $z$ , on définit la *convolution*  $y * z$  par

$$(y * z)_k = \sum_{j=0}^{N-1} y_{k-j} z_j.$$

Montrer que  $y * z$  est également  $N$ -périodique et que

$$\mathcal{F}_N(y * z) = \mathcal{F}_N y \cdot \mathcal{F}_N z \quad (12.4)$$

où la multiplication dans (12.4) est effectuée élément par élément.

15. Démontrer l'égalité de Parseval (1806) pour la transformée de Fourier discrète.

$$N \sum_{k=0}^{N-1} |z_k|^2 = \sum_{k=0}^{N-1} |y_k|^2.$$

16. Soit  $N$  un nombre pair et  $p_N(x)$  le polynôme trigonométrique qui passe par  $(x_\ell, y_\ell)$  pour  $\ell = 0, 1, \dots, N-1$  (voir la formule (9.1)). Montrer que

$$p_N(x) = \sum_{\ell=0}^{N-1} y_\ell S_N(x - x_\ell) \quad (12.5)$$

où

$$S_N(x) = \frac{\sin(xN/2)}{N} \cdot \frac{\cos(x/2)}{\sin(x/2)}. \quad (12.6)$$

17. La fonction  $S_N(x)$  de (12.6) permet d'étudier l'influence des erreurs dans les  $y_\ell$  sur le polynôme trigonométrique interpolant. Montrer que  $S_N(x)$  est  $2\pi$ -périodique et que pour tout  $x \in [-\pi, \pi]$

$$\left| S_N(x) - \frac{\sin(xN/2)}{xN/2} \right| \leq \frac{2}{\pi \cdot N}.$$

18. Calculer la main le spline naturel ( $s''(x_0) = s''(x_n) = 0$ ) qui passe par les points

$$(-3, 0), \quad (-2, 0), \quad (-1, 0), \quad (0, 1), \quad (1, 0), \quad (2, 0), \quad (3, 0).$$

Dessiner les graphes (si possible avant les calculs!) de  $s(x)$ ,  $s'(x)$ ,  $s''(x)$ ,  $s'''(x)$ .

*Indication.* Pour des données symétriques par rapport  $x = 0$ , le spline naturel est une fonction paire. Avec cette observation, on peut réduire la dimension du système des  $p_i$ .

19. Pour la fonction

$$f(x) = \begin{cases} x^3 & \text{si } x \geq 0 \\ 0 & \text{si } x \leq 0, \end{cases}$$

calculer les différences

$$\Delta f(x), \quad \Delta^2 f(x), \quad \Delta^3 f(x), \quad \text{et} \quad B(x) = \Delta^4 f(x)$$

où  $\Delta g(x) := g(x+1) - g(x)$  et  $\Delta^2 g := \Delta(\Delta g)$ , etc.

Montrer que  $B(x)$  est une fonction "spline" (appelée "B-spline") à support compact, c.-à-d. qu'elle est nulle en dehors d'un intervalle borné.

Calculer les valeurs de  $B(x)$  pour  $x$  entier et dessiner  $B(x)$ .

20. (*Spline périodique*). Soient donnés  $(x_i, y_i)$  pour  $i = 0, 1, \dots, n$  avec  $y_n = y_0$ . Montrer l'existence et l'unicité d'un spline qui passe par tous les  $(x_i, y_i)$  et qui satisfait

$$s'(x_n) = s'(x_0), \quad s''(x_n) = s''(x_0).$$

21. (Formule de *Newton-Gregory*). Soit  $f : [a, b] \rightarrow \mathbb{R}$  une fonction différentiable et  $x_i = a + ih$ , (avec  $h = (b - a)/n$ ) une division équidistante de  $[a, b]$ . Montrer que l'intégration du spline, défini par  $s(x_i) = f(x_i)$  pour  $i = 0, 1, \dots, n$ ,  $s'(a) = f'(a)$  et  $s'(b) = f'(b)$ , donne la formule de quadrature suivante:

$$\int_a^b f(x)dx \approx h \left( \frac{1}{2}f(x_0) + f(x_1) + \dots + f(x_{n-1}) + \frac{1}{2}f(x_n) \right) - \frac{h^2}{12} (f'(x_n) - f'(x_0)) .$$

Pourquoi cette formule est-elle exacte pour tout polynôme de degré 3?