Université de Rennes 2024-2025
Agrégation externe

Compléments sur l'interpolation

Miguel Rodrigues

La partie la moins avancée discutée ici peut être trouvée dans les livres classiques d'analyse numérique en général [Dem16, Fil13, Sch04]. Les estimations plus précises se trouvent essentiellement dans [CM92, CM97, Rom05].

Ces notes de cours sont consacrées principalement à l'interpolation polynomiale comme outil d'approximation de fonctions d'une variable réelle à valeur réelle. Cependant, la partie existence et construction de l'interpolation se transfère immédiatement à un corps quelconque. Presque tout se transfère, avec des modifications, aux fonctions holomorphes d'une variable complexe. L'analyse pour des fonctions définies sur un espace vectoriel réel de dimension finie demanderait des considérations géométriques plus poussées. À l'inverse, aucun changement ne serait nécessaire pour considérer des fonctions à valeurs dans un tel espace vectoriel, puisqu'il suffirait d'appliquer coordonnée par coordonnée.

1 Interpolation de Lagrange

1.1 Généralités

Dans tout ce texte, on se donne $(a,b) \in \mathbf{R}^2$, $a < b$ et $f : [a,b] \to \mathbf{R}$ et pour $n \in \mathbf{N}$ on note $\mathbf{R}_n[X]$ l'espace des polynômes réels de degré au plus n.
Pour tout choix de points deux à distincts, appelés *nœuds d'interpolation*, $\mathbf{x} \in [a,b]^{n+1}$, explicitement $\mathbf{x} = (x_0, \cdots, x_n)$, avec $x_j \neq x_\ell$ si $j \neq \ell$, il existe une unique polynôme réel P de degré au plus n tel que

$$\forall j \in [\![0,n]\!],\ P(x_j) = f(x_j)\,.$$

Ce polynôme est appelé polynôme d'interpolation de Lagrange de f aux nœuds $\mathbf{x}$. Dans la suite on le notera $P_f^{\mathbf{x}}$.
Pour démontrer l'affirmation ci-dessus, il suffit de montrer que l'application

$$E_{\mathbf{x}} : \mathbf{R}_n[X] \to \mathbf{R}^{n+1}\,,\ P \mapsto (P(x_j))_{0\leq j\leq n}$$

est surjective, ce que l'on fera ci-dessous en exhibant une famille de polynômes envoyée sur la base canonique de $\mathbf{R}^{n+1}$.
La terminologie *interpolation* vient du fait qu'à partir des valeurs aux nœuds on a a construit une fonction $P_f^{\mathbf{x}}$ que l'on peut aussi évaluer *entre* les nœuds. Évidemment rien n'empêche de l'utiliser aussi à l'extérieur des nœuds pour faire de l'extrapolation. L'interpolation de Lagrange n'utilise que des valeurs de f, pas des valeurs de ses dérivées, ce qui est compatible avec le fait que celles-ci puissent venir d'un calcul approché discret ou de mesures expérimentales.
L'intérêt de remplacer f par $P_f^{\mathbf{x}}$ est multiple. Le plus évident correspond au cas où l'on ne connaît pas f mais où l'on est capable de l'évaluer en quelques points (par des simulations, des mesures,...). Un autre

intérêt est que l'on replace une fonction arbitraire, vivant dans un espace fonctionnel, de dimension infinie, par une fonction polynomiale de degré au plus n, donc vivant dans un espace de dimension finie. Cette étape de discrétisation est cruciale quand on veut faire des calculs numériques sur f et c'est sous cette forme qu'intervient l'interpolation dans des problèmes *a priori* déconnectés comme celui de quadrature, c'est-à-dire d'intégration numérique.

Remarque 1 *Il est très instructif de comparer la théorie développée ici pour les polynômes de Lagrange avec celles, mieux connues, pour les polynômes de Taylor pour lesquels on choisit un point x_* et impose*

$$\forall j \in [\![0, n]\!],\ P^{(j)}(x_*) = f^{(j)}(x_*)\,.$$

1.2 Choix de bases

Pour effectuer un calcul concret de $P_f^{\mathbf{x}}$, il faut choisir une base de $\mathbf{R}_n[X]$ et résoudre l'équation affine associée à la matrice, dans cette base, de l'application d'évaluation $E_{\mathbf{x}}$.
Si l'on choisit la base canonique de $\mathbf{R}_n[X]$, $(X^j)_{0\le j\le n}$, on obtient comme matrice de $E_{\mathbf{x}}$ une matrice de Vandermonde, $[x_\ell^j]_{j,\ell}$. Le conditionnement de telles matrices croît très vite avec le degré n, ce qui les rend inutiles en pratique.
À l'inverse, la *base de Lagrange* $(L_j^{\mathbf{x}})_{0\le j\le n}$ définie par

$$L_j^{\mathbf{x}}(X) := \prod_{\substack{\ell=0\\ \ell\neq j}}^{n} \frac{X - x_j}{x_\ell - x_j}$$

est envoyée sur la base canonique de $\mathbf{R}^{n+1}$ de sorte que la matrice de $E_{\mathbf{x}}$ dans cette base est la matrice identité. Ainsi

$$P_f^{\mathbf{x}} \;=\; \sum_{j=0}^{n} f(x_j)\, L_j^{\mathbf{x}}\,.$$

L'un des principaux défauts de la forme de Lagrange est qu'elle est trop globale en les nœuds (parce que symétrique) et qu'elle se prête en particulier mal à l'ajout d'un nouveau point *a posteriori*. À ce propos, on notera que même pour $n = 1$ et que l'on calcule à la main, la forme utilisée n'est pas celle de Lagrange mais

$$P_f^{(x_0,x_1)}(X) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}\,(X - x_0)\,.$$

La méthode la plus utilisée en pratique, la *forme de Newton*, généralise cette dernière formule. Elle correspond à choisir comme base $(\pi_j^{\mathbf{x}})_{0\le j\le n}$ où, pour $0 \le j \le n+1$,

$$\pi_j^{\mathbf{x}}(X) \;=\; \prod_{\ell=0}^{j-1}(X - x_\ell)\,.$$

Puisque les $(\pi_j^{\mathbf{x}})_{0\le j\le n}$ sont ordonnés pour la division, une fois que l'on dispose des coefficients de $P_f^{\mathbf{x}}$ dans cette base, son évaluation, rapide et stable, s'obtient facilement par la méthode de Ruffini-Horner. Un calcul récursif des coefficients est permis par le fait que si $n \ge 1$ et $\mathbf{x} \in [a, b]^{n+1}$ et $\mathbf{y} \in [a, b]^n$ sont tels que $y_j = x_j$, $0 \le j \le n-1$, alors

$$P_f^{\mathbf{x}} \;=\; P_f^{\mathbf{y}} \;+\; f[x_0, \cdots, x_n]\,\pi_n^{\mathbf{x}}$$

pour un certain $f[x_0, \cdots, x_n]$ à déterminer. Puisque $\pi_j^{\mathbf{x}}$ ne dépend que des $(j+1)$ premières coordonnées de $\mathbf{x}$, on déduit alors

$$P_f^{\mathbf{x}} \;=\; \sum_{\ell=0}^{n} f[x_0, \cdots, x_\ell]\, \pi_\ell^{\mathbf{x}}\,.$$

On montre par ailleurs que $f[x_0] = f(x_0)$ et si $n \geq 1$

$$f[x_0, \cdots, x_n] \;=\; \frac{f[x_1, \cdots, x_n] - f[x_0, \cdots, x_{n-1}]}{x_n - x_0}\,.$$

Cette dernière égalité se déduit du fait que

$$P_f^{(x_0,\cdots,x_n}(X) \;=\; \frac{X - x_0}{x_1 - x_0}\, P_f^{(x_1,\cdots,x_n)}(X) + \left(1 - \frac{X - x_0}{x_1 - x_0}\right) P_f^{(x_0,\cdots,x_{n-1})}(X)\,.$$

Les coefficients se calculent donc récursivement et à cause de la forme de la relation de récurrence sont appelés *différences divisées.*

Remarque 2 *La forme de Newton se compare plus facilement aux développements de Taylor. En effet, lorsque* $\mathbf{x}$ *converge vers* $(x_*, \cdots, x_*)$*, chaque* $\pi_j^{\mathbf{x}}$ *converge vers* $(X - x_*)^j$ *et, si* f *est de classe* $\mathcal{C}^n$*, chaque* $f[x_0, \cdots, x_j]$ *converge vers* $f^{(j)}(x_*)/j!$*. Ainsi les différences divisées fournissent une approximation des dérivées par différences finies.*

2 Asymptotique en le degré

On s'intéresse maintenant aux estimations de comparaison entre f et $P_f^{\mathbf{x}}$.

2.1 Formules de reste

Théorème 3 *Supposons* f *de classe* $\mathcal{C}^{n+1}$ *sur* $[a, b]$*. Alors pour tout choix de nœuds deux à deux distincts* $\mathbf{x} \in \mathbf{R}^{n+1}$*, pour tout* $x \in [a, b]$*,*

1. *on a*

$$f(x) \;=\; P_f^{\mathbf{x}}(x) \;+\; \pi_{n+1}^{\mathbf{x}}(x) \int_{\Delta_n} f^{(n+1)}(x_0 + t_0(x_1 - x_0) + \cdots + t_{n-1}(x_n - x_{n-1}) + t_n(x - x_n))\, \mathrm{d}\mathbf{t}$$

où Δ_n *désigne le simplexe*

$$\Delta_n := \{\, \mathbf{t} = (t_0, \cdots, t_n) \,|\, 0 \leq t_n \leq t_{n-1} \leq \cdots \leq t_1 \leq t_0 \leq 1 \,\}\;;$$

2. *il existe* $\xi_x \in [a, b]$ *tel que*

$$f(x) \;=\; P_f^{\mathbf{x}}(x) \;+\; \pi_{n+1}^{\mathbf{x}}(x)\, \frac{f^{(n+1)}(\xi_x)}{(n+1)!}\,.$$

La première formule se montre par récurrence, la seconde s'en déduit par le théorème des valeurs intermédiaires (puisque $1/(n+1)!$ est le volume de Δ_n).
Dans la limite $\mathbf{x} \to (x_*, \cdots, x_*)$, on retrouve les formules de reste de Taylor usuelles.
Pour utiliser les estimations ci-dessus, la proposition élémentaire suivante est cruciale.

Proposition 4 *Il existe une constante universelle C (indépendante de a, b, n, f et $\mathbf{x}$) telle que si $a < b$, $n \in \mathbf{N}$, et $\mathbf{x} \in [a,b]^{n+1}$ à coordonnées deux à deux distinctes, alors*

$$\max_{[a,b]} |\pi^{\mathbf{x}}_{n+1}| \;\leq\; C \left(\frac{b-a}{R}\right)^{n+1}$$

où R peut être choisi comme

1. $R = 1$ *dans tous les cas ;*
2. $R = 2$ *si les $x_0, \cdots, x_n$ sont symétriques par rapport à $(a+b)/2$;*
3. $R = e$ *si les $x_0, \cdots, x_n$ sont équidistants de a à b, c'est-à-dire*

$$x_j = a + (b-a)\,\frac{j}{n}\,;$$

4. $R = 4$ *si les $x_0, \cdots, x_n$ sont les points de Tchebychev, c'est-à-dire les zéros du $(n+1)$-ième polynôme de Tchebychev*[1]*,*

$$x_j \;=\; \frac{a+b}{2} + \frac{b-a}{2}\,\cos\left(\frac{2j+1}{n+1}\,\frac{\pi}{2}\right)\,.$$

Le cas $R = 2$ est lié à $x(1-x) \leq 1/2$ quand $0 \leq x \leq 1$; celui $R = e$ est lié à la formule de Stirling.
Une manière simple et efficace de faire converger l'erreur $f - P^{\mathbf{x}}_f$ vers 0 est donc de fixer n et de faire tendre $(b-a)$ vers 0. Évidemment pour couvrir un segment fixé, il faut alors découper ce segment en morceaux assez petits et appliquer l'interpolation polynomiale sur les petits morceaux. C'est ce que l'on appelle l'*interpolation polynomiale par morceaux* et c'est cet usage de l'interpolation polynomiale qui est utile en pratique et relié aux méthodes de quadrature,...
On va maintenant essayer de se convaincre que l'approche inverse consistant à fixer $[a,b]$ et à envoyer n vers l'infini est vouée à l'échec. Le constat est analogue à celui des développements de Taylor : sauf pour les fonctions analytiques, les développements de Taylor en x_* ne sont pertinents qu'à n fixé dans la limite $x \to x_*$, pas à x fixé dans la limite $n \to \infty$.
Ce problème est suffisamment important pour qu'il soit souvent avantageux de fixer $k \in \mathbf{N}$, k beaucoup plus petit que n et de résoudre un problème des moindres carrés : déterminer $P \in \mathbf{R}_k[X]$ qui minimise la norme euclidienne de $(P(x_j) - f(x_j))_{0 \leq j \leq n}$ parmi les polynômes de degré au plus k. On évite ainsi les phénomènes de *surapprentissage* : la surperformance sur les nœuds se fait au détriment des performances ailleurs. Voir les figures ci-dessous.
Afin de mener l'analyse, commençons par remarquer que pour que la combinaison des estimations ci-dessus n'explose pas il faut contrôler uniformément en n

$$\frac{(b-a)^n}{n!\,R^n}\,\max_{[a,b]} |f^{(n)}|\,,$$

ce qui arrive lorsque f est développable en séries entières avec un rayon de convergence en $(a+b)/2$ au moins égal à

$$(b-a)\,\left(\frac{1}{2} + \frac{1}{R}\right)\,.$$

1. Polynômes orthogonaux pour la mesure sur $[a,b]$, absolument continue par rapport à la mesure de Lebesgue, de densité

$$x \mapsto \frac{1}{\sqrt{(b-x)(x-a)}}.$$

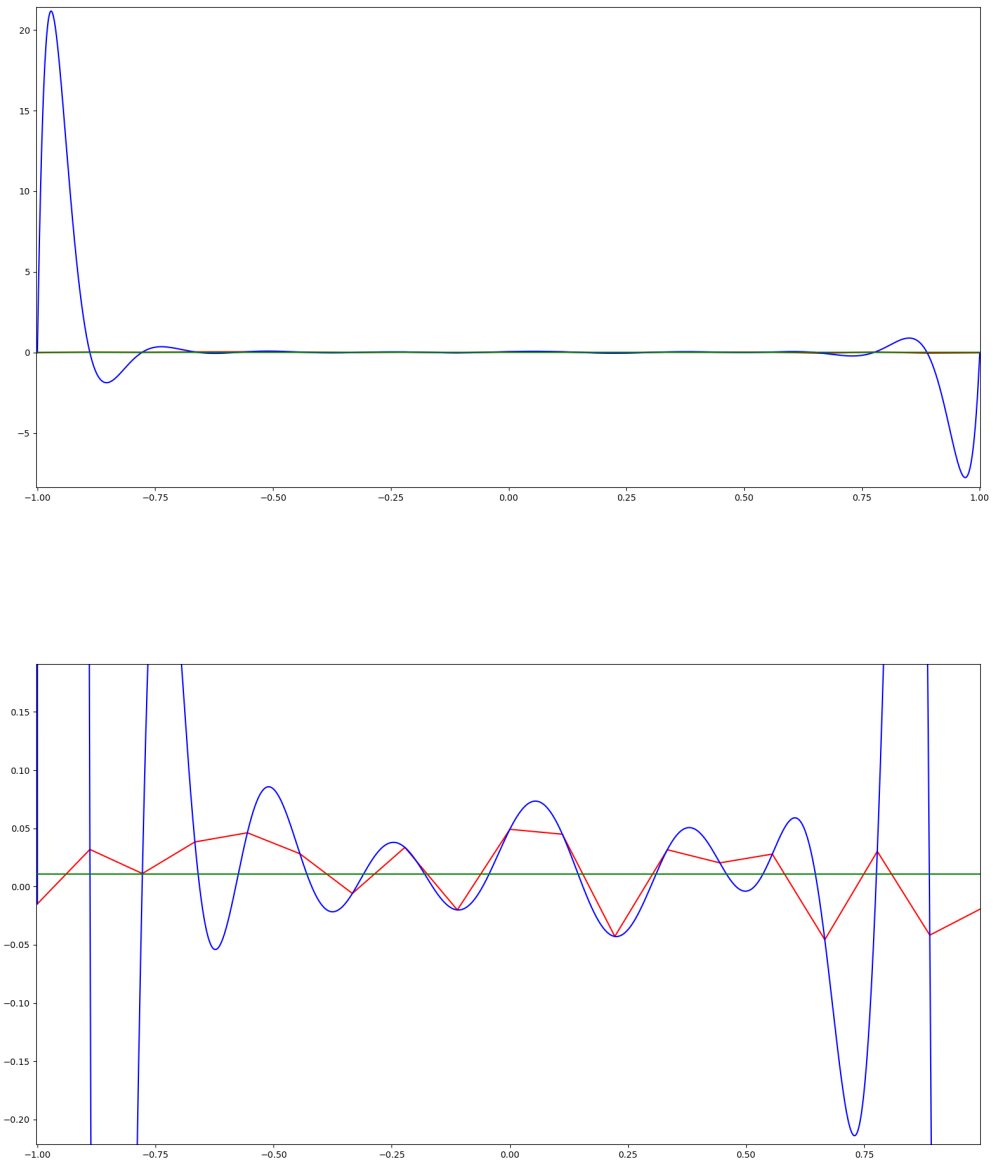

FIGURE 1 – On a tiré ici uniformément dans $[-0,05, 0,05]$ des valeurs en 19 points équidistants de $[-1, 1]$. On compare en rouge l'interpolation affine par morceaux, en vert la meilleure approximation constante au sens des moindres carrés et en bleu l'interpolation polynomiale.

Si le rayon est strictement plus grand, on déduit la convergence de $\max_{[a,b]} |f - P_f^{\mathbf{x}}|$ vers zéro quand $n \to \infty$. Les considérations ci-dessus sont basées sur le fait que si f est développable en séries entières sur $[\frac{a+b}{2} - r, \frac{a+b}{2} + r]$ avec $r > \frac{b-a}{2}$, alors

$$\frac{1}{n!} \max_{[a,b]} |f^{(n)}| \overset{n\to\infty}{=} \mathcal{O}((r - \tfrac{b-a}{2})^{-n})$$

ce que l'on montre par exemple à partir de la formule de Cauchy

$$\frac{1}{n!} f^{(n)}(x) = \frac{1}{2\mathrm{i}\pi} \int_{\partial D\left(\frac{a+b}{2}, r\right)} \frac{f(z)}{(z-x)^{n+1}} \, \mathrm{d}z$$

où l'on a également noté f l'extension analytique de f au plan complexe.
Les estimations de cette sous-section ne sont pas très précises, mais il est tout de même vrai que pour pas mal de choix de points, y compris les points équidistants (les seuls vraiment pratiques), même si f est développable en série entière l'erreur peut explosée. C'est le *phénomène de Runge*. Explicitement avec $(a, b) = (-1, 1)$,

$$f : [-1, 1] \to \mathbf{R}, \qquad x \mapsto \frac{1}{1 + 25x^2} \tag{2.1}$$

et la suite de choix de points

$$\mathbf{x}^{[n]} := \left(-1 + 2\frac{j}{n}\right)_{0 \le j \le n}$$

on a

$$\max_{[-1,1]} |f - P_f^{\mathbf{x}^{[n]}}| \overset{n\to\infty}{\longrightarrow} \infty .$$

La divergence se fait par l'accumulation d'oscillations proches des bords de $[-1, 1]$.

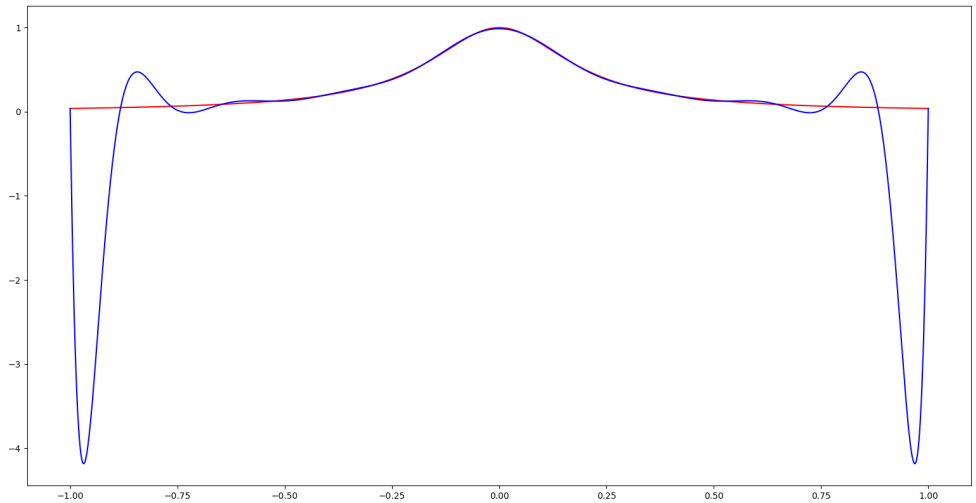

FIGURE 2 – En rouge le graphe de la fonction définie par (2.1). En bleu celui de son polynôme interpolé à partir de 18 points équidistants. Le calcul du polynôme a été obtenu avec une commande de régression polynomiale, `numpy.polyfit` de `python` en l'occurrence.

2.2 Approximation polynomiale

Des estimations plus précises sont obtenues en comparant l'erreur d'interpolation à la meilleure approximation polynomiale sur $[a,b]$. Le théorème de Weierstrass assure que si f est continue alors

$$d_n(f) := \inf_{P\in\mathbf{R}_n[X]} \max_{[a,b]} |f-P| \overset{n\to\infty}{\longrightarrow} 0\,.$$

Il est naturel de se demander à quel point l'interpolation polynomiale est loin de réaliser cet infimum. Une réponse fait intervenir la norme de l'application d'interpolation

$$I_{\mathbf{x}} \,:\, \mathcal{C}^0([a,b]) \to \mathcal{C}^0([a,b])\,, \quad f \mapsto P_f^{\mathbf{x}}\,,$$

c'est-à-dire

$$\Lambda_{\mathbf{x}} := \|I_{\mathbf{x}}\| = \sup_{\substack{f\in\mathcal{C}^0([a,b])\\ f\not\equiv 0}} \frac{\max_{[a,b]}|I_{\mathbf{x}}(f)|}{\max_{[a,b]}|f|}\,.$$

Ce nombre est appelé *constante de Lebesgue*. On montre facilement que

$$\Lambda_{\mathbf{x}} \,=\, \max_{[a,b]} \left(\sum_{j=0}^{n} |L_j^{\mathbf{x}}| \right)\,.$$

Le lemme suivant montre que l'erreur d'approximation est contrôlée par le produit d'une constante de stabilité $\Lambda_{\mathbf{x}}$ et d'une erreur d'approximabilité $d_n(f)$. Le principe est de portée très générale, et la forme générale du lemme est appelée lemme de Lebesgue.

Lemme 5 *Pour tout choix de nœuds deux à deux distincts $\mathbf{x}\in[a,b]^{n+1}$, on a*

$$\max_{[a,b]} |f-P_f^{\mathbf{x}}| \,\le\, (1+\Lambda_{\mathbf{x}})\, d_n(f)\,.$$

Démonstration. Pour tout $Q\in\mathbf{R}_n[X]$, on a $f-P_f^{\mathbf{x}} = f-Q-P_{f-Q}^{\mathbf{x}}$ donc

$$\max_{[a,b]} |f-P_f^{\mathbf{x}}| \,\le\, (1+\Lambda_{\mathbf{x}})\, |f-Q|\,.$$

Le résultat s'en déduit en passant en la borne inférieure sur Q. ∎

Une analyse fine montre le théorème suivant.

Théorème 6 *1. Il existe une constante universelle c (indépendante de a, b, n et $\mathbf{x}$) telle que si $a<b$, $n\in\mathbf{N}$, et $\mathbf{x}\in[a,b]^{n+1}$ à coordonnées deux à deux distinctes, alors*

$$\Lambda_{\mathbf{x}} \ge \frac{2}{\pi}\,\ln(n+1) + c\,.$$

2. Pour tout $a<b$, si $\mathbf{x}^{[n]}$ note la suite des points de Tchebychev sur $[a,b]$ à l'ordre n, alors

$$\Lambda_{\mathbf{x}^{[n]}} \overset{n\to\infty}{\sim} \frac{2}{\pi}\,\ln(n+1)\,.$$

3. Pour tout $a<b$, si $\mathbf{x}^{[n]}$ note la suite des points équidistants sur $[a,b]$ à l'ordre n, alors

$$\Lambda_{\mathbf{x}^{[n]}} \overset{n\to\infty}{\sim} \frac{2^{n+1}}{e\,n\,\ln(n)}\,.$$

De ce point de vue les points équidistants sont catastrophiques alors que les points de Tchebychev sont asymptotiquement optimaux et, comme on va le voir ci-dessous, ils fonctionnement théoriquement de manière tout à fait satisfaisante pour les fonctions régulières. Néanmoins ceux-ci sont peu utilisés en pratique (avec n grand) parce qu'ils ne sont pas réellement explicites.
Grâce au théorème de la borne uniforme, un corollaire du théorème précédent est que pour tout choix de suites $\mathbf{x}^{[n]}$ de nœuds d'interpolation il existe un choix de fonction continue f telle que $\max_{[a,b]} |P_f^{\mathbf{x}^{[n]}}|$ tende vers $+\infty$ lorsque $n \to \infty$. Cependant avec de la régularité sur f on peut compenser la croissance de $\Lambda_{\mathbf{x}^{[n]}}$. Avec les points équidistants il faut de l'analyticité avec un rayon d'analyticité assez grand. Avec les points de Tchebychev on peut montrer qu'il suffit que f soit absolument continue, c'est-à-dire que sa dérivée (au sens des distributions) soit intégrable ou dit encore autrement qu'elle soit la primitive d'une fonction intégrable.
Contentons-nous de quelques résultats pour les fonctions $\mathcal{C}^\ell$.

Théorème 7 *Soit $a < b$ et $\ell \in \mathbf{N}$. Il existe une constante C (indépendante de n et f) telle que pour tout $n \in \mathbf{N}$ et $f \in \mathcal{C}^\ell([a,b])$,*

$$d_n(f) \leq C \, \frac{\max_{[a,b]}(|f| + \cdots + |f^{(\ell)}|)}{(n+1)^\ell} \, .$$

Démonstration. On va d'abord montrer comment se ramener à de l'approximation trigonométrique.
En multipliant les développements de Taylor adéquats par des fonctions de troncature puis en rendant paire puis périodique, on obtient $g : \mathbf{R} \to \mathbf{R}$ de classe $\mathcal{C}^\ell$, 2π-périodique et paire telle que pour $\theta \in [\frac{\pi}{4}, \frac{3\pi}{4}]$

$$g(\theta) \, = \, f\left(\frac{a+b}{2} + \frac{b-a}{2} \sqrt{2} \, \cos(\theta) \right) \, ,$$

et dont la norme $\mathcal{C}^\ell$ est contrôlée par un multiple de celle de f.
Il suffit alors de montrer la borne L^∞ sur la différence entre g et un polynôme en $\cos(\cdot)$ de degré au plus n. Or il découle des formules trigonométriques classiques que pour tout $k \in \mathbf{N}$, $\cos(k \cdot)$ est un polynôme réel en $\cos(\cdot)$ de degré k (les polynômes associés sont les polynômes de Tchebychev).
Pour conclure il suffit d'appliquer le théorème d'approximation trigonométrique de Jackson, qui se démontre par récurrence sur ℓ. Plutôt que de fournir sa démonstration, nous fournissons une version dégradée par un facteur $\sqrt{n}$ mais plus courte.
Introduisons les coefficients de Fourier de g

$$\forall k \in \mathbf{Z} \, , \quad c_k(g) \, := \, \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}\, k\, x} g(x) \mathrm{d}x \, .$$

Comme g est paire et réelle, ses coefficients de Fourier sont réels et pairs en k donc sa série de Fourier à un ordre n

$$S_n(g) := \sum_{k=-n}^{n} c_k(g) \mathrm{e}^{\mathrm{i}\, k \cdot} \, = \, c_0(g) + \sum_{k=1}^{n} c_k(g) \, \cos(k \cdot)$$

est un polynôme réel en $\cos(\cdot)$ de degré au plus n. De plus on a pour tout n

$$\begin{aligned}
\|g - S_n(g)\|_{L^\infty(\mathbf{R})} &\leq \sum_{|k|\geq n+1} |c_k(g)| \\
&\leq \sqrt{\sum_{|k|\geq n+1} \frac{1}{|k|^{2\ell}}} \times \sqrt{\sum_{k\in\mathbf{Z}} |k|^{2\ell} |c_k(g)|^2} \\
&\leq \sqrt{2} \sqrt{\int_n^\infty \frac{\mathrm{d}x}{x^{2\ell}}} \times \sqrt{\sum_{k\in\mathbf{Z}} |c_k(g^{(\ell)})|^2} \\
&\leq \frac{\|g^{(\ell)}\|_{L^2([-\pi,\pi])}}{\sqrt{\pi(2\ell-1)}} \frac{1}{n^{\ell-\frac{1}{2}}} .
\end{aligned}$$

■

Le théorème de Jackson possède une réciproque partielle : si g peut être approché trigonométriquement à l'ordre n avec une erreur en $\mathcal{O}(n^{-(k+\alpha)})$, $0 < \alpha < 1$, alors g est $\mathcal{C}^k$ et $g^{(k)}$ est α-höldérienne.
La borne que nous montrons se rate d'un facteur $\sqrt{n}$ parce qu'elle utilise $\|g^{(\ell)}\|_{L^2}$ plutôt que $\|g^{(\ell)}\|_{L^\infty}$. La borne optimale pour l'approximation par les séries de Fourier tronquées est en $\mathcal{O}(\ln(n+1)\, n^{-\ell})$, ce qui se montre avec le lemme de Lebesgue et le théorème de Jackson.
La démonstration ci-dessus illustre le fait qu'une bonne approximation trigonométrique s'obtient de manière plus naturelle qu'une approximation polynomiale. Par ailleurs, l'interpolation trigonométrique avec des points équidistants est particulièrement efficace ; *cf.* l'analyse pratique de la transformée de Fourier discrète. Cela explique le rôle des points de Tchebychev pour l'interpolation polynomiale, qui correspondent aux points équidistants pour l'interpolation trigonométrique.
Certaines démonstrations du théorème de Weierstrass fournissent explicitement des polynômes qui approchent f, par exemple *via* les polynômes de Bernstein. Néanmoins ces approximations explicites sont en général sous optimale. Ainsi pour une fonction $\mathcal{C}^1$ l'approximation de Bernstein donne une erreur en $\mathcal{O}(n^{-1/2})$ plutôt qu'en $\mathcal{O}(n^{-1})$. Cet ordre en $n^{-1/2}$ a la même origine que celui dans la loi des grands nombres/le théorème central limite et la méthode de Monte Carlo. Il s'agit d'une vraie limitation de la méthode : contrairement aux approximations trigonométriques, la vitesse de convergence n'est pas meilleure pour des fonctions plus régulières.

Références

[CM92] M. Crouzeix and A.-L. Mignot. *Analyse numérique des équations différentielles.* Masson, 1992.

[CM97] M. Crouzeix and A.-L. Mignot. *Exercices d'analyse numérique des équations différentielles.* Masson, 1997.

[Dem16] J.-P. Demailly. *Analyse numérique et équations différentielles.* ÉDP Sciences, 2016.

[Fil13] F. Filbet. *Analyse numérique - Algorithme et étude mathématique.* Dunod, 2013.

[Rom05] J.-E. Rombaldi. *Interpolation & approximation - Analyse pour l'agrégation.* Vuibert, 2005.

[Sch04] M. Schatzman. *Analyse numérique.* Dunod, 2004.