Prépa. Agrég. Analyse, avril 2007.

Weierstrass et Stone-Weierstrass

Le théorème classique de Weierstrass

Pour démontrer le théorème d'approximation polynomiale de Weierstrass sur la droite réelle, considérons la fonction réelle γ définie sur $\mathbb{R}$ par

$$\forall x \in \mathbb{R}, \quad \gamma(x) = \frac{\mathrm{e}^{-x^2/2}}{\sqrt{2\pi}}.$$

Il s'agit bien entendu de la densité gaussienne standard ; cette fonction γ est réelle ≥ 0 et son intégrale sur $\mathbb{R}$ est égale à 1. Par ailleurs, elle est développable en série entière de rayon de convergence infini. On peut dès lors imaginer qu'une convolution $f * \gamma$, pour f continue à support compact, sera encore développable en série entière de rayon infini, donc approchable sur tout compact par des fonctions polynomiales (tout simplement les sommes partielles de la série entière) ; d'après les résultats classiques d'approximation par convolution, on sait qu'on peut approcher f uniformément par des convolées qui sont « en gros » du type $f * \gamma$ (et qui par conséquent sont aussi des fonctions entières) : c'est la stratégie conceptuellement très simple que nous emploierons pour prouver le théorème de Weierstrass sur $\mathbb{R}$; c'est en fait la méthode originelle[a] employée par Weierstrass en 1885. Commençons par un lemme très gentil.

Lemme 1. *Si f est mesurable bornée sur $\mathbb{R}$, la fonction convolée $f * \gamma$ est égale sur $\mathbb{R}$ à la somme d'une série entière de rayon de convergence infini (autrement dit, la convolée $f * \gamma$ est la restriction* à $\mathbb{R}$ *d'une* fonction entière[b] *définie sur $\mathbb{C}$).*

Preuve. — On écrit

$$\sqrt{2\pi}\,(f * \gamma)(x) = \int_{\mathbb{R}} f(t)\,\mathrm{e}^{-(x-t)^2/2}\;\mathrm{d}t = \mathrm{e}^{-x^2/2} \int_{\mathbb{R}} \mathrm{e}^{xt}\, f(t)\,\mathrm{e}^{-t^2/2}\;\mathrm{d}t.$$

On va développer e^{xt} en série entière, justifier l'interversion série-intégrale pour montrer que pour tout x réel, on a

$$h(x) := \int_{\mathbb{R}} \mathrm{e}^{xt}\, f(t)\,\mathrm{e}^{-t^2/2}\;\mathrm{d}t = \sum_{n=0}^{+\infty} \Bigl(\int_{\mathbb{R}} \frac{t^n}{n!}\, f(t)\,\mathrm{e}^{-t^2/2}\;\mathrm{d}t \Bigr)\, x^n,$$

ce qui prouvera que h est développable en série entière de rayon infini. Pour finir, le produit $\mathrm{e}^{-x^2/2}\, h(x)$ est développable lui aussi, avec rayon de convergence infini, par les résultats classiques sur les produits de séries entières.

Pour justifier l'interversion dans l'expression

$$\int_{\mathbb{R}} \Bigl(\sum_{n=0}^{+\infty} \frac{x^n t^n}{n!}\, f(t)\,\mathrm{e}^{-t^2/2} \Bigr)\,\mathrm{d}t$$

il suffit de montrer que la nouvelle série de fonctions où tous les termes de la série ci-dessus ont été remplacés par leur valeur absolue définit une fonction intégrable ≥ 0 de la variable t ; or f est bornée, disons par M, donc

$$\sum_{n=0}^{+\infty} \frac{|x|^n |t|^n}{n!}\, |f(t)|\,\mathrm{e}^{-t^2/2} \leq \mathrm{M}\;\mathrm{e}^{|xt| - t^2/2} \leq \mathrm{M}\;\mathrm{e}^{x^2}\,\mathrm{e}^{-t^2/4}$$

qui est bien intégrable en t sur $\mathbb{R}$; ainsi, l'interversion est justifiée et la preuve du lemme terminée.

À partir du lemme 1, il est assez facile d'obtenir le théorème d'approximation de Weierstrass. Considérons une fonction f réelle ou complexe sur $\mathbb{R}$, continue à support compact. La fonction f est donc bornée et uniformément continue sur $\mathbb{R}$; par les résultats habituels d'approximation par convolution, on sait que la suite des fonctions f_n définies par

$$\forall x \in \mathbb{R}, \quad f_n(x) = \int_{\mathbb{R}} f(x - t/n)\gamma(t)\,\mathrm{d}t$$

converge uniformément([c]) vers f sur $\mathbb{R}$. Si on définit F_n par $\mathrm{F}_n(s) = f(s/n)$, on voit que F_n est mesurable bornée sur $\mathbb{R}$ et on voit que

$$\forall x \in \mathbb{R}, \quad f_n(x) = \int_{\mathbb{R}} \mathrm{F}_n(nx - t)\gamma(t)\,\mathrm{d}t = (\mathrm{F}_n * \gamma)(nx).$$

La fonction $\mathrm{F}_n * \gamma$ est entière([d]) d'après le lemme 1 ; il en résulte que f_n est aussi une fonction entière (si φ est entière et n fixé, la fonction $x \to \varphi(nx)$ est évidemment entière aussi).

Faisons un dernier pas pour obtenir l'approximation par des fonctions polynomiales. Si $\varphi(x) = \sum_{k=0}^{+\infty} a_k x^k$ est une fonction entière, on sait d'après les résultats classiques sur les séries entières que les fonctions polynomiales $\mathrm{P}_{\mathrm{N}}(x) = \sum_{k=0}^{\mathrm{N}} a_k x^k$, définies par les sommes partielles de la série entière, convergent vers φ uniformément sur tout compact de $\mathbb{R}$. Si K est un compact de $\mathbb{R}$, on peut donc remplacer chaque fonction entière f_n par une fonction polynomiale P_n qui approche f_n à 2^{-n} près sur le compact K. On a ainsi prouvé le théorème d'approximation polynomiale.

Théorème (Weierstrass). *Pour toute fonction f continue à support compact sur $\mathbb{R}$, il existe*([e]) *une suite de fonctions entières qui tend vers f uniformément sur $\mathbb{R}$; en particulier, il existe une suite de fonctions polynomiales qui tend vers f uniformément sur tout compact.*

Si la fonction f est réelle, les séries entières (et par conséquent les polynômes) construits ci-dessus sont à coefficients réels ; si f est complexe, on doit évidemment prendre des coefficients complexes. Pour passer au théorème de Weierstrass sur $\mathbb{R}^d$, on peut considérer la fonction réelle γ_d définie sur $\mathbb{R}^d$ par

$$\forall x = (x_1, \ldots, x_d) \in \mathbb{R}^d, \quad \gamma_d(x) = \gamma(x_1)\ldots\gamma(x_d) = (2\pi)^{-d/2} \exp\bigl(-\sum_{i=1}^{d} x_i^2/2\bigr).$$

Dans la preuve du lemme 1, il suffirait de remplacer le développement en série de e^{xt} en une seule variable par celui de

$$\mathrm{e}^{x \cdot t} = \mathrm{e}^{x_1 t_1 + \cdots + x_d t_d}$$

en d variables, et de généraliser les propriétés des séries entières en une variable au cas de séries entières en plusieurs variables. Si on préfère éviter ces histoires de séries entières en plusieurs variables, on laissera tomber l'affirmation sur les fonctions entières et on passera tout de suite à la convergence uniforme sur tout compact pour les sommes partielles de ces séries, sommes partielles qui sont clairement des fonctions polynomiales de plusieurs variables.

On va développer une autre approche, utilisant les *partitions de l'unité*, qui permet de ramener facilement le cas de plusieurs dimensions au cas de $\mathbb{R}$.

Lemme 2. *Soient* F *une partie fermée d'un espace métrique* (E, d) *et* $\mathrm{U}_1, \ldots, \mathrm{U}_\mathrm{N}$ *des ouverts de* E *qui recouvrent* F *; il existe des fonctions* $\varphi_1, \ldots, \varphi_\mathrm{N} \geq 0$ *continues sur* E *telles que chaque* φ_j *soit nulle en dehors de* U_j, $j = 1, \ldots, \mathrm{N}$ *et*

$$\forall x \in \mathrm{F}, \quad \sum_{j=1}^{\mathrm{N}} \varphi_j(x) = 1\,; \quad \forall y \in \mathrm{E}, \quad \sum_{j=1}^{\mathrm{N}} \varphi_j(y) \leq 1.$$

Preuve. — Si l'un des ouverts, par exemple U_1, est égal à E tout entier, il suffit de poser $\varphi_1 = \mathbf{1}$ et $\varphi_j = 0$ pour $j \neq 1$; dans le cas contraire, posons pour chaque $j = 1, \ldots, \mathrm{N}$ et pour tout $x \in \mathrm{E}$

$$\psi_j(x) = \operatorname{dist}(x, \mathrm{U}_j^c) = \inf\{d(x, y) : y \notin \mathrm{U}_j\}.$$

Il est clair que ψ_j est continue sur E, $\psi_j > 0$ sur U_j et $\psi_j = 0$ hors de U_j. Considérons aussi

$$\chi(x) = \operatorname{dist}(x, \mathrm{F})\,;$$

cette fonction est continue ≥ 0 sur E. La somme $\psi = \chi + \sum_{j=1}^{\mathrm{N}} \psi_j$ est continue > 0 en tout point de E puisque les U_j recouvrent F et que $\chi > 0$ en dehors de F. On pose alors pour chaque indice j entre 1 et N

$$\forall x \in \mathrm{E}, \quad \varphi_j(x) = \psi_j(x)/\psi(x).$$

La fonction φ_j est continue ≥ 0 sur E. Pour tout $y \in \mathrm{E}$, on a

$$\sum_{j=1}^{\mathrm{N}} \varphi_j(y) = \Bigl(\sum_{j=1}^{\mathrm{N}} \psi_j(y)\Bigr)/\psi(y) = \Bigl(\sum_{j=1}^{\mathrm{N}} \psi_j(y)\Bigr)/\Bigl(\chi(y) + \sum_{j=1}^{\mathrm{N}} \psi_j(y)\Bigr) \leq 1.$$

Si $x \in \mathrm{F}$, on a $\chi(x) = 0$ et par conséquent $\sum_{j=1}^{\mathrm{N}} \varphi_j(x) = 1$.

Soit a un réel > 0 ; appliquons le lemme 2 à $\mathrm{E} = \mathbb{R}$ et $\mathrm{F} = [-a, a]$; d'après le lemme 2, pour tout $\delta > 0$ donné on peut trouver un entier N et des fonctions $\varphi_1, \ldots, \varphi_\mathrm{N}$ continues sur $\mathbb{R}$, chacune ayant son support contenu dans un intervalle de longueur $< \delta$, telles que $\varphi_j \geq 0$, $\sum_{j=1}^{\mathrm{N}} \varphi_j \leq 1$ sur $\mathbb{R}$ et $\sum_{j=1}^{\mathrm{N}} \varphi_j(x) = 1$ pour tout $x \in [-a, a]$ (dans le cas de $\mathbb{R}$, on peut trouver des solutions plus particulières[f] et plus explicites au problème de partition de l'unité). On utilisera cette information dans la preuve qui suit.

Désignons par $\operatorname{Dec}(\mathbb{R}^d)$ l'ensemble des fonctions *décomposées*, c'est-à-dire les fonctions φ continues à support compact sur $\mathbb{R}^d$ qui sont de la forme

$$\forall x = (x_1, \ldots, x_d) \in \mathbb{R}^d, \quad \varphi(x) = \varphi_1(x_1)\varphi_2(x_2)\ldots\varphi_d(x_d),$$

où les fonctions $\varphi_1, \ldots, \varphi_d$ sont continues à support compact sur $\mathbb{R}$.

Proposition. *Pour toute fonction continue f sur un compact* K *de* $\mathbb{R}^d$ *et pour tout* $\varepsilon > 0$, *il existe une combinaison linéaire g de fonctions de* $\mathrm{Dec}(\mathbb{R}^d)$ *telle que*

$$\forall x \in \mathrm{K}, \quad |f(x) - g(x)| < \varepsilon.$$

On peut supposer que

$$\max_{y \in \mathbb{R}^d} |g(y)| \le \max_{x \in \mathrm{K}} |f(x)|.$$

Preuve. — On écrira la preuve dans le cas de $\mathbb{R}^2$, pour simplifier. On se donne un compact K de $\mathbb{R}^2$, on choisit $a > 0$ tel que K soit contenu dans le carré $\mathrm{C}_a = [-a, a]^2$, et on se donne une fonction continue f sur K. Soit $\varepsilon > 0$ donné ; par la continuité uniforme de la fonction f, il existe $\delta > 0$ tel que $|f(y) - f(x)| < \varepsilon$ dès que la distance entre $x, y \in \mathrm{K}$ est $< \delta$. Considérons une famille de fonctions continues $\varphi_1, \ldots, \varphi_\mathrm{N} \ge 0$ sur $\mathbb{R}$, φ_j étant à support dans un intervalle U_j de longueur $< \delta/2$, telle que la fonction somme S définie sur $\mathbb{R}$ par

$$\forall t \in \mathbb{R}, \quad \mathrm{S}(t) = \sum_{j=1}^{\mathrm{N}} \varphi_j(t)$$

vérifie $\mathrm{S} \le 1$ sur $\mathbb{R}$ et $\mathrm{S} = 1$ sur $[-a, a]$. Posons pour tout $x = (x_1, x_2) \in \mathbb{R}^2$ et tous les indices $i, j \in \{1, \ldots, \mathrm{N}\}$

$$\varphi_{i,j}(x) = \varphi_i(x_1)\varphi_j(x_2).$$

Cette fonction $\varphi_{i,j}$ est ≥ 0, nulle en dehors du rectangle ouvert $\mathrm{U}_{i,j} = \mathrm{U}_i \times \mathrm{U}_j$, qui est de diamètre $< \delta$, et $\varphi_{i,j}$ est dans $\mathrm{Dec}(\mathbb{R}^2)$. Il est clair que pour tout $x = (x_1, x_2) \in \mathbb{R}^2$ on a

$$\sum_{i,j=1}^{\mathrm{N}} \varphi_{i,j}(x) = \Bigl(\sum_{i=1}^{\mathrm{N}} \varphi_i(x_1)\Bigr)\Bigl(\sum_{j=1}^{\mathrm{N}} \varphi_j(x_2)\Bigr) = \mathrm{S}(x_1)\mathrm{S}(x_2).$$

Il en résulte que pour tout $x = (x_1, x_2) \in \mathrm{C}_a$ on a $\sum_{i,j=1}^{\mathrm{N}} \varphi_{i,j}(x) = 1$, et que pour tout $y \in \mathbb{R}^2$, on a $\sum_{i,j=1}^{\mathrm{N}} \varphi_{i,j}(y) \le 1$. Désignons par A l'ensemble des couples (i, j) tels que l'ouvert $\mathrm{U}_{i,j}$ rencontre K. Si $x \in \mathrm{K}$ et $(i, j) \notin \mathrm{A}$, le point x n'est pas dans l'ouvert $\mathrm{U}_{i,j}$, donc $\varphi_{i,j}(x) = 0$; on a ainsi

$$\forall x \in \mathrm{K}, \quad \sum_{(i,j) \in \mathrm{A}} \varphi_{i,j}(x) = \sum_{i,j=1}^{\mathrm{N}} \varphi_{i,j}(x) = 1. \tag{1}$$

Pour chaque $(i, j) \in \mathrm{A}$, choisissons un point $x_{i,j} \in \mathrm{K} \cap \mathrm{U}_{i,j}$ et posons

$$\forall y \in \mathbb{R}^2, \quad g(y) = \sum_{(i,j) \in \mathrm{A}} \varphi_{i,j}(y) f(x_{i,j}).$$

La fonction g est combinaison linéaire de fonctions de $\mathrm{Dec}(\mathbb{R}^2)$. Il est clair que si on pose $\mathrm{M} = \max_{x \in \mathrm{K}} |f(x)|$, on a

$$\forall y \in \mathbb{R}^2, \quad |g(y)| \le \sum_{(i,j) \in \mathrm{A}} \varphi_{i,j}(y) |f(x_{i,j})| \le \mathrm{M} \sum_{(i,j) \in \mathrm{A}} \varphi_{i,j}(y) \le \mathrm{M}.$$

Soit $x \in \mathrm{K}$; d'après la relation (1) on peut écrire

$$f(x) - g(x) = \sum_{(i,j)\in \mathrm{A}} \varphi_{i,j}(x)\big(f(x) - f(x_{i,j})\big).$$

Pour chaque terme $\varphi_{i,j}(x)(f(x) - f(x_{i,j}))$ non nul, on a $\varphi_{i,j}(x) \neq 0$, donc $x \in \mathrm{U}_{i,j}$ et $\|x - x_{i,j}\| < \delta$, puisque le diamètre de $\mathrm{U}_{i,j}$ est $< \delta$; il en résulte que $|f(x) - f(x_{i,j})| < \varepsilon$ et

$$|f(x) - g(x)| \le \sum_{(i,j)\in \mathrm{A}} \varphi_{i,j}(x)\,|f(x) - f(x_{i,j})| \le \varepsilon \sum_{(i,j)\in \mathrm{A}} \varphi_{i,j}(x) = \varepsilon,$$

ce qui termine la preuve.

Remarque. La proposition permet de retrouver par itération le fait que la fonction f définie sur K peut se prolonger [g] en fonction continue sur $\mathbb{R}^d$ (une version facile du *lemme d'Urysohn*).

Théorème (Weierstrass en plusieurs dimensions). *Pour toute fonction continue f sur un compact* K *de $\mathbb{R}^d$, il existe une suite de fonctions polynomiales qui tend vers f uniformément sur* K.

Preuve. — Continuons avec $d = 2$ et avec un carré C_a contenant le compact K. D'après la proposition précédente, on peut trouver g de la forme

$$g = \sum_{\alpha=1}^{\mathrm{N}} c_\alpha \varphi_\alpha$$

telle que $|f(x) - g(x)| < \varepsilon/2$ pour tout $x \in \mathrm{K}$ et $\varphi_\alpha \in \mathrm{Dec}(\mathbb{R}^2)$ pour tout $\alpha = 1, \ldots, \mathrm{N}$. Il est clair qu'il suffit de pouvoir approcher sur C_a chaque fonction φ_α par une fonction polynomiale P_α, avec par exemple

$$\max_{x\in \mathrm{C}_a} |\varphi_\alpha(x) - \mathrm{P}_\alpha(x)| < \frac{\varepsilon}{2(1 + \sum_\beta |c_\beta|)}\,;$$

en effet, la fonction $\mathrm{P} = \sum_\alpha c_\alpha \mathrm{P}_\alpha$ sera polynomiale et on aura pour tout $x \in \mathrm{C}_a$

$$|g(x) - \mathrm{P}(x)| \le \sum_{\alpha=1}^{\mathrm{N}} |c_\alpha|\,|\varphi_\alpha(x) - \mathrm{P}_\alpha(x)| < \varepsilon/2,$$

ce qui donne finalement

$$\max_{x\in \mathrm{K}} |f(x) - \mathrm{P}(x)| < \varepsilon.$$

On voit donc qu'il suffit de savoir approcher arbitrairement bien sur le carré C_a une fonction $\varphi \in \mathrm{Dec}(\mathbb{R}^2)$,

$$\varphi(x_1, x_2) = \varphi_1(x_1)\varphi_2(x_2).$$

Soit δ donné tel que $0 < \delta < 1$; d'après Weierstrass sur la droite, chaque fonction φ_j, $j = 1, 2$, peut être approchée sur $[-a, a]$ par une fonction polynomiale P_j ; on en déduit une approximation de φ par

$$p(x) = \mathrm{P}_1(x_1)\mathrm{P}_2(x_2)$$

qui est polynomiale des variables x_1, x_2 : si $|\varphi_i(t) - \mathrm{P}_i(t)| \le \eta < \delta/3$ pour tout i et tout $|t| \le a$, on voit [h] que $|\varphi(x) - p(x)| \le \eta + (1 + \eta)\eta < 3\eta < \delta$.

Weierstrass complexe
Sur $\mathbb{C}^d$, on considérera les d coordonnées complexes $z_j = x_j + \mathrm{i} y_j$, $j = 1, \ldots, d$; identifiant $\mathbb{C}^d$ à $\mathbb{R}^{2d}$ on pourra d'après ce qui précède, pour toute fonction complexe f continue sur un compact $\mathrm{K} \subset \mathbb{C}^d$, approcher uniformément sur K la fonction f par des fonctions polynomiales à coefficients complexes des variables réelles x_j, y_j ; de plus, comme on a $x_j = \frac{1}{2}(z_j + \overline{z}_j)$ et $y_j = \frac{1}{2\mathrm{i}}(z_j - \overline{z}_j)$, on peut remplacer les fonctions polynomiales à coefficients complexes en x_j, y_j par des fonctions polynomiales p en $z_j, \overline{z}_j$, $j = 1, \ldots, d$ à coefficients complexes, de la forme

$$\forall z = (z_1, \ldots, z_d) \in \mathbb{C}^d, \quad p(z) = \mathrm{P}(z, \overline{z}) = \sum_{m,n} c_{m,n}\, z_1^{m_1}\, \overline{z}_1^{\,n_1} \ldots z_d^{m_d}\, \overline{z}_d^{\,n_d}$$

où la somme ci-dessus est étendue à un nombre fini de couples (m, n) de multi-indices $m = (m_1, \ldots, m_d)$, $n = (n_1, \ldots, n_d)$ formés de d entiers ≥ 0.

Le théorème de Stone-Weierstrass

On peut trouver une démonstration standard de Stone-Weierstrass dans tous les bons ouvrages (par exemple Hirsch-Lacombe). La démonstration suivante est un peu moins habituelle : elle consiste à utiliser des idées cependant très habituelles (comme celle de partition de l'unité), pour ramener le théorème de Stone-Weierstrass au théorème classique de Weierstrass.

On dit qu'une famille $\mathcal{G}$ de fonctions continues complexes (ou réelles) sur un espace topologique X *sépare les points* si pour tous $x \neq y$ dans X, il existe une fonction $g \in \mathcal{G}$ telle que $g(x) \neq g(y)$.

Théorème 1 (préliminaire à Stone-Weierstrass). *On suppose que* X *est un espace topologique compact non vide et* $\mathcal{G}$ *une famille de fonctions continues complexes (ou réelles) sur* X *qui sépare les points de* X. *Pour toute fonction* f *complexe (ou réelle) continue sur* X *et tout* $\varepsilon > 0$, *il existe un entier* N, *des fonctions* $g_1, \ldots, g_{\mathrm{N}} \in \mathcal{G}$ *et une fonction* h *complexe (ou réelle) continue à support compact sur* $\mathbb{C}^{\mathrm{N}}$ *(ou sur* $\mathbb{R}^{\mathrm{N}}$*) tels que*

$$\forall x \in \mathrm{X}, \quad \bigl| f(x) - h\bigl(g_1(x), \ldots, g_{\mathrm{N}}(x)\bigr) \bigr| < \varepsilon.$$

Preuve. — Démontrons le cas complexe ; le cas réel en découle, en considérant $\mathbb{R}$ comme un sous-ensemble de $\mathbb{C}$, et en remplaçant à la fin la fonction complexe h sur $\mathbb{C}^{\mathrm{N}}$ par la partie réelle de sa restriction à $\mathbb{R}^{\mathrm{N}}$, considéré comme sous-ensemble de $\mathbb{C}^{\mathrm{N}}$.

Soient f une fonction complexe continue sur X et $\varepsilon > 0$; on considère le sous-ensemble

$$\mathrm{U} = \{(x, y) \in \mathrm{X} \times \mathrm{X} : |f(x) - f(y)| < \varepsilon\}$$

qui est un ouvert du compact $\mathrm{X} \times \mathrm{X}$, puisque f est continue. Son complémentaire

$$\mathrm{K} = \{(x, y) \in \mathrm{X} \times \mathrm{X} : |f(x) - f(y)| \geq \varepsilon\} = (\mathrm{X} \times \mathrm{X}) \setminus \mathrm{U}$$

est donc compact. Pour chaque $(x, y) \in \mathrm{K}$, on a $x \neq y$, donc il existe par hypothèse une fonction $g \in \mathcal{G}$ telle que $g(x) \neq g(y)$; l'ensemble $\mathrm{V}_g = \{(u, v) : |g(u) - g(v)| > 0\}$ est un ouvert de K contenant (x, y), donc la famille des $(\mathrm{V}_g)_{g \in \mathcal{G}}$ est un recouvrement ouvert du compact K. Par Borel-Lebesgue, on peut sélectionner une famille finie $g_1, \ldots, g_{\mathrm{N}}$ dans $\mathcal{G}$ telle que les ouverts V_{g_j} correspondants, $j = 1, \ldots, \mathrm{N}$ recouvrent K. La fonction

$$(x, y) \to \Bigl(\sum_{j=1}^{\mathrm{N}} |g_j(x) - g_j(y)|^2\Bigr)^{1/2}$$

est continue > 0 sur le compact K, donc son minimum est > 0 ; désignons ce minimum par $\delta > 0$. On définit pour tout $x \in \mathrm{X}$

$$\mathrm{G}(x) = (g_1(x), \ldots, g_{\mathrm{N}}(x)) \in \mathbb{C}^{\mathrm{N}} ;$$

l'application G est continue de X dans $\mathbb{C}^{\mathrm{N}}$, et on a si $(x, y) \in \mathrm{K}$

$$\|\mathrm{G}(x) - \mathrm{G}(y)\| = \Bigl(\sum_{j=1}^{\mathrm{N}} |g_j(x) - g_j(y)|^2\Bigr)^{1/2} \geq \delta.$$

On a donc pour tout $(x, y) \in \mathrm{X} \times \mathrm{X}$, selon que $(x, y) \in \mathrm{K}$ ou $(x, y) \in \mathrm{U}$:

$$\|\mathrm{G}(x) - \mathrm{G}(y)\| \geq \delta \ \ \textbf{ou} \ \ |f(x) - f(y)| < \varepsilon. \tag{2}$$

Désignons par $\mathrm{K}_1 = \mathrm{G}(\mathrm{X}) \subset \mathbb{C}^{\mathrm{N}}$ l'image du compact X ; l'ensemble K_1 est un sous-ensemble compact de $\mathbb{C}^{\mathrm{N}}$. Considérons un recouvrement de K_1 par une famille finie $(\mathrm{U}_\alpha)_{\alpha \in \mathrm{I}}$ d'ouverts de $\mathbb{C}^{\mathrm{N}}$ de diamètre $< \delta$ et qui rencontrent K_1 ; pour chaque $\alpha \in \mathrm{I}$ soit z_α un point de $\mathrm{K}_1 \cap \mathrm{U}_\alpha$; puisque z_α est dans l'image de X par G, on peut trouver $x_\alpha \in \mathrm{X}$ tel que $\mathrm{G}(x_\alpha) = z_\alpha$. D'après le lemme 2 on peut trouver une famille $(\varphi_\alpha)_{\alpha \in \mathrm{I}}$ de fonctions continues ≥ 0 sur $\mathbb{C}^{\mathrm{N}}$, telle que φ_α soit nulle en dehors de U_α pour chaque $\alpha \in \mathrm{I}$, que $\sum_{\alpha \in \mathrm{I}} \varphi_\alpha \leq 1$ sur $\mathbb{C}^{\mathrm{N}}$ et $\sum_{\alpha \in \mathrm{I}} \varphi_\alpha(z) = 1$ pour tout $z \in \mathrm{K}_1$. Posons

$$\forall z \in \mathbb{C}^{\mathrm{N}}, \quad h(z) = \sum_{\alpha \in \mathrm{I}} \varphi_\alpha(z) f(x_\alpha).$$

Cette fonction h est continue sur $\mathbb{C}^{\mathrm{N}}$, et à support compact car chaque φ_α est nulle en dehors du borné U_α. On va montrer que la fonction $x \in \mathrm{X} \to h(\mathrm{G}(x))$ est proche de f ; on a

$$f(x) - h(\mathrm{G}(x)) = \sum_{\alpha \in \mathrm{I}} \varphi_\alpha(\mathrm{G}(x))\,(f(x) - f(x_\alpha)) ;$$

si un terme particulier $\varphi_\alpha(\mathrm{G}(x))(f(x) - f(x_\alpha))$ de la somme précédente est non nul, on a d'abord $\varphi_\alpha(\mathrm{G}(x)) \neq 0$, donc $\mathrm{G}(x)$ est dans U_α, ainsi que $z_\alpha = \mathrm{G}(x_\alpha)$; comme le diamètre de U_α est $< \delta$, on voit que $\|\mathrm{G}(x) - \mathrm{G}(x_\alpha)\| < \delta$, ce qui implique d'après la condition (2) que $|f(x) - f(x_\alpha)| < \varepsilon$; on a donc

$$\bigl|\varphi_\alpha(\mathrm{G}(x))\,(f(x) - f(x_\alpha))\bigr| \leq \varepsilon\, \varphi_\alpha(\mathrm{G}(x))$$

pour tout $\alpha \in \mathrm{I}$, donc

$$|f(x) - h(\mathrm{G}(x))| = \bigl|\sum_{\alpha \in \mathrm{I}} \varphi_\alpha(\mathrm{G}(x))\,(f(x) - f(x_\alpha))\bigr| \leq \varepsilon \sum_{\alpha \in \mathrm{I}} \varphi_\alpha(\mathrm{G}(x)) = \varepsilon,$$

ce qui achève la preuve.

Théorème 2 (une forme de Stone-Weierstrass). *On suppose que* X *est un espace topologique compact non vide et* $\mathcal{G}$ *une famille de fonctions complexes continues sur* X *qui sépare les points de* X*. Pour toute fonction* f *complexe continue sur* X *et tout* $\varepsilon > 0$*, il existe un entier* N*, des fonctions* $g_1, \ldots, g_\mathrm{N} \in \mathcal{G}$ *et une fonction* p *sur* $\mathbb{C}^\mathrm{N}$*, polynomiale des variables* $z_j, \overline{z}_j$*,* $j = 1, \ldots, \mathrm{N}$ *tels que*

$$\forall x \in \mathrm{X}, \quad \bigl| f(x) - p\bigl(g_1(x), \ldots, g_\mathrm{N}(x)\bigr) \bigr| < \varepsilon.$$

Supposons maintenant que $\mathcal{G}$ *soit une famille de fonctions réelles continues sur* X *qui sépare les points de* X *; pour toute fonction* f *réelle continue sur* X *et tout* $\varepsilon > 0$*, il existe un entier* N*, des fonctions* $g_1, \ldots, g_\mathrm{N} \in \mathcal{G}$ *et une fonction* P *sur* $\mathbb{R}^\mathrm{N}$*, polynomiale à coefficients réels de* N *variables réelles, tels que*

$$\forall x \in \mathrm{X}, \quad \bigl| f(x) - \mathrm{P}\bigl(g_1(x), \ldots, g_\mathrm{N}(x)\bigr) \bigr| < \varepsilon.$$

Preuve. — Les preuves des deux cas sont similaires, en utilisant l'énoncé de Weierstrass approprié. Donnons le cas complexe. D'après le théorème 1, on peut trouver une fonction continue h sur $\mathbb{C}^\mathrm{N}$ et $g_1, \ldots, g_\mathrm{N} \in \mathcal{G}$ tels que $|f(x) - h(g_1(x), \ldots, g_\mathrm{N}(x))| < \varepsilon/2$ pour tout $x \in \mathrm{X}$. D'après Weierstrass complexe, on peut trouver une fonction p polynomiale des variables $z_1, \ldots, z_\mathrm{N}, \overline{z_1}, \ldots, \overline{z_\mathrm{N}}$ telle que $|h(z) - p(z)| < \varepsilon/2$ pour tout z du compact $\mathrm{K}_1 \subset \mathbb{C}^\mathrm{N}$ formé des points $\mathrm{G}(x) = (g_1(x), \ldots, g_\mathrm{N}(x))$, $x \in \mathrm{X}$. Finalement,

$$\forall x \in \mathrm{X}, \quad \bigl| f(x) - p(\mathrm{G}(x)) \bigr| < \varepsilon,$$

le résultat annoncé.

Remarque 1. Pour faire tourner la démonstration du théorème 1 il suffit que $\mathcal{G}$ sépare les couples de points (x, y) de X tels que $|f(x) - f(y)| \geq \varepsilon$; en particulier il suffit que $\mathcal{G}$ sépare les couples (x, y) tels que $f(x) \neq f(y)$. Cette variante permet par exemple de traiter l'ensemble singleton $\mathcal{G} = \{x \to \mathrm{e}^{\mathrm{i}\,x}\}$ sur le compact $\mathrm{X} = [0, 2\pi]$ sans devoir parler de quotient : si f est continue 2π-périodique, l'exponentielle $x \to \mathrm{e}^{\mathrm{i}\,x}$ sépare les points x, y tels que $f(x) \neq f(y)$; il en résulte que les polynômes en $z, \overline{z}$ de l'exponentielle permettent d'approcher la fonction continue périodique f. Ces polynômes de l'exponentielle sont les fonctions de la forme

$$x \to \sum_{k=-\mathrm{N}}^{\mathrm{N}} c_k \, \mathrm{e}^{\mathrm{i}\,kx},$$

qu'on appelle *polynômes trigonométriques*.

Remarque 2. On pourrait approcher la fonction h du théorème 1 par une fonction p d'une famille donnée $\mathcal{P}$ de fonctions sur $\mathbb{C}^\mathrm{N}$ (ou $\mathbb{R}^\mathrm{N}$), famille telle que pour tout compact K de $\mathbb{C}^\mathrm{N}$ (ou $\mathbb{R}^\mathrm{N}$), l'ensemble des restrictions à K des fonctions de $\mathcal{P}$ soit dense dans $\mathrm{C}(\mathrm{K})$. C'est le cas si $\mathcal{P}$ est la classe des fonctions polynomiales (d'un nombre quelconque de variables ; c'est le théorème de Weierstrass classique), mais bien d'autres cas sont envisageables.

Remarque 3. Faisons une application spécifique de la remarque 2 : il n'est pas extrêmement difficile de montrer (i) que toute fonction réelle continue sur $\mathbb{R}^d$, à support dans un compact K de $\mathbb{R}^d$, peut être approchée uniformément sur K par une fonction de la forme

$$y \in \mathbb{R}^d \to \max\bigl(a_1(y), \ldots, a_{\mathrm{M}}(y)\bigr) - \max\bigl(b_1(y), \ldots, b_{\mathrm{N}}(y)\bigr)$$

où les fonctions a_i, b_j sont des fonctions affines sur $\mathbb{R}^d$. Si dans la preuve du théorème 2, on remplace l'emploi du théorème de Weierstrass classique par cette nouvelle information, on obtient la version réticulée de Stone-Weierstrass : si E est un espace vectoriel de fonctions réelles continues sur le compact X, qui est stable par sup fini, contient les constantes et sépare les points de X, alors E est dense dans $\mathrm{C}_{\mathbb{R}}(\mathrm{X})$.

Stone-Weierstrass exprimé avec une algèbre de fonctions

Commençons par le cas réel. On a dit dans le théorème 2 que si $\mathcal{G}$ est une famille de fonctions réelles continues sur le compact X, qui sépare les points de X, toute fonction réelle continue f sur X pourra être approchée uniformément sur X par une fonction de la forme $\psi = \mathrm{P}(g_1, \ldots, g_{\mathrm{N}})$, avec $g_j \in \mathcal{G}$ et P fonction polynomiale de N variables réelles. Plus précisément, cela signifie que ψ est combinaison linéaire à coefficients réels de fonctions de la forme

$$x \in \mathrm{X} \to g_1(x)^{k_1} \ldots g_{\mathrm{N}}(x)^{k_{\mathrm{N}}},$$

avec $k_1, \ldots, k_{\mathrm{N}}$ entiers ≥ 0. En particulier, lorsque $k_1 = \cdots = k_{\mathrm{N}} = 0$, on introduit la fonction constante $\mathbf{1}$ dans la combinaison linéaire ; l'ensemble de toutes les fonctions ψ de la forme précédente est la $\mathbb{R}$-algèbre engendrée par la famille $\mathcal{G}$ et par les fonctions constantes.

Supposons réciproquement que l'ensemble $\mathcal{A}$ soit une algèbre de fonctions réelles, contenant les fonctions constantes et séparant les points de X : dans ce cas toutes les fonctions ψ sont encore dans $\mathcal{A}$, ce qui implique par le théorème 2 que toute fonction continue f peut être approchée uniformément par un élément de $\mathcal{A}$.

Théorème (Stone-Weierstrass, cas réel). *On suppose que* X *est un espace topologique compact non vide et* $\mathcal{A}$ *une* $\mathbb{R}$*-algèbre de fonctions continues réelles sur* X, *qui sépare les points de* X *et contient les fonctions constantes. L'algèbre* $\mathcal{A}$ *est dense dans* $\mathrm{C}_{\mathbb{R}}(\mathrm{X})$, *pour la norme uniforme.*

Remarque. Revenons à l'énoncé du théorème 2, en supposant de plus qu'il existe un point $x_0 \in \mathrm{X}$ tel que $f(x_0) = 0$ et que $g(x_0) = 0$ pour toute $g \in \mathcal{G}$; le théorème 2 fournit un polynôme P de N variables réelles tel que $|f(x_0) - \mathrm{P}(g_1(x_0), \ldots, g_{\mathrm{N}}(x_0))| < \varepsilon$, c'est-à-dire que $|\mathrm{P}(0, \ldots, 0)| < \varepsilon$. Mais $\mathrm{P}(0, \ldots, 0)$ est le coefficient constant du polynôme P : quitte à perdre un ε supplémentaire, on peut supposer que P n'a pas de terme constant. Il en résulte que si $\mathcal{A}$ est une $\mathbb{R}$-algèbre de fonctions continues réelles sur X, nulles au point x_0 et qui sépare les points de X, l'algèbre $\mathcal{A}$ est dense dans $\mathrm{C}_{\mathbb{R}}(\mathrm{X})$.

Considérons l'espace $\mathrm{C}_0(\mathbb{R})$ des fonctions réelles continues sur $\mathbb{R}$ qui tendent vers 0 à l'infini. Cet espace s'identifie naturellement au sous-espace des fonctions continues sur le compactifié d'Alexandrov de $\mathbb{R}$, qui sont nulles au point à l'infini. D'après les lignes précédentes, si une algèbre $\mathcal{A}$ de fonctions réelles contenue dans $\mathrm{C}_0(\mathbb{R})$ sépare les points de $\mathbb{R}$, et si de plus pour tout $x \in \mathbb{R}$ il existe une fonction $g \in \mathcal{A}$ telle que $g(x) \neq 0$ (séparation du point à l'infini), alors $\mathcal{A}$ est dense dans $\mathrm{C}_0(\mathbb{R})$.

Dans le cas complexe, l'approximation du théorème 2 était obtenue par une fonction $\psi = p(g_1, \ldots, g_N)$, avec $g_j \in \mathcal{G}$ et p une fonction polynomiale en z_j et $\overline{z}_j$, $j = 1, \ldots, N$. Cela signifie que ψ est combinaison linéaire à coefficients complexes de fonctions de la forme

$$x \in X \to g_1(x)^{k_1} \overline{g_1(x)}^{\ell_1} \ldots g_N(x)^{k_N} \overline{g_N(x)}^{\ell_N},$$

ce qui montre qu'on doit aussi utiliser ici les fonctions complexes conjuguées telles que $x \to \overline{g_1(x)}$. Pour obtenir la traduction en termes d'algèbres de fonctions, il nous faudra supposer que la fonction $\overline{g}$ est encore dans $\mathcal{A}$, pour toute $g \in \mathcal{A}$.

Théorème (Stone-Weierstrass, cas complexe). *On suppose que* X *est un espace topologique compact non vide et* $\mathcal{A}$ *une* $\mathbb{C}$*-algèbre de fonctions continues sur* X, *qui sépare les points de* X, *contient les fonctions constantes,* et qui de plus est stable par conjugaison complexe. *L'algèbre* $\mathcal{A}$ *est dense dans* $C_{\mathbb{C}}(X)$, *pour la norme uniforme.*

Notes.

(a) L'article de Weierstrass est intitulé *Über die analytische Darstellbarkeit sogenannter willkürlicher Functionen einer reellen Veränderlichen*, il est paru dans les *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin (1885), pp. 633-639.* En 1885, Weierstrass a 70 ans, et il enseigne son fameux cours à Berlin depuis presque 30 ans ; j'imagine (ça n'engage que moi) qu'il devait connaître le résultat depuis un moment, et qu'il s'est décidé à le rédiger cette année-là.

On peut trouver une version française de l'argument de Weierstrass dans le livre d'Émile Borel *Leçons sur les fonctions de variables réelles*, dont la première édition est parue chez Gauthier-Villars en 1905. On peut consulter l'article de Weierstrass sur le web : passer par *Wikipedia*, version anglaise ; on trouve un lien à la fin de l'article intitulé *Stone-Weierstrass theorem*.

(b) Une *fonction entière* F est une fonction définie et holomorphe sur $\mathbb{C}$; d'après la théorie des fonctions holomorphes, on sait que dans ce cas la série de Taylor de F à l'origine (ou bien n'importe où ailleurs, d'ailleurs) a un rayon de convergence infini, et que la somme de cette série représente F en tout point de $\mathbb{C}$,

$$\forall z \in \mathbb{C}, \quad F(z) = \sum_{n=0}^{+\infty} \frac{F^{(n)}(0)}{n!} z^n.$$

(c) Rappelons l'argument : on se donne une fonction f bornée et uniformément continue sur $\mathbb{R}$. Étant donné $\varepsilon > 0$, il existe $\delta > 0$ tel que pour tout x on ait $|f(x) - f(x-t)| < \varepsilon/2$ dès que $|t| < \delta$. Par ailleurs il existe M tel que $|f(x)| \leq M$ pour tout x. Comme γ est intégrable sur $\mathbb{R}$ on peut trouver n assez grand pour que

$$\int_{|s| \geq \delta n} \gamma(s) \, ds < \frac{\varepsilon}{4M}.$$

Comme γ est d'intégrale 1, on a pour tout x

$$f(x) - f_n(x) = \int_{\mathbb{R}} \big(f(x) - f(x - s/n)\big) \gamma(s) \, ds$$

$$= \int_{|s|<\delta n} \bigl(f(x) - f(x - s/n)\bigr)\gamma(s)\,\mathrm{d}s + \int_{|s|\ge\delta n} \bigl(f(x) - f(x - s/n)\bigr)\gamma(s)\,\mathrm{d}s.$$

Dans la première intégrale, on a $|f(x) - f(x - s/n)| < \varepsilon/2$ puisque $|s/n| < \delta$, donc

$$|f(x) - f_n(x)| \le \frac{\varepsilon}{2} \int_{\mathbb{R}} \gamma(s)\,\mathrm{d}s + 2\mathrm{M} \int_{|s|\ge\delta n} \gamma(s)\,\mathrm{d}s < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(d) La preuve du lemme 1 utilise des propriétés des intégrales dépendant d'un paramètre, ce qui semble bien normal dans le cadre de la préparation à l'agrégation. Mais on pourrait, dans le cas d'une fonction f continue à support compact, par exemple à support contenu dans $[-1, 1]$, remplacer l'intégrale (de Riemann)

$$f_n(x) = \int_{\mathbb{R}} f(s/n)\gamma(nx - s)\,\mathrm{d}s = n \int_{-1}^{1} f(t)\gamma(nx - nt)\,\mathrm{d}t$$

par une somme de Riemann, pour k assez grand, évidemment entière comme combinaison linéaire de fonctions entières

$$\varphi_n(x) = \frac{n}{k} \sum_{j=-k}^{k-1} f(j/k)\gamma(nx - nj/k).$$

En prenant $k = n^2$ grand, on trouve une approximation entière assez simple et explicite pour une fonction f continue sur $\mathbb{R}$, à support dans $[-1, 1]$, à savoir

$$x \to \frac{1}{n\sqrt{2\pi}} \sum_{j=-n^2}^{n^2-1} f(j/n^2) \exp\bigl(-(nx - j/n)^2/2\bigr).$$

Le lecteur pourra chercher une preuve directe de cette affirmation. Il pourra commencer par résoudre l'exercice suivant, avant de l'appliquer à des fonctions $\varphi(t)$ fabriquées à partir de «morceaux» de $t \to \gamma(nx - t)$:

si φ est une fonction ≥ 0 intégrable sur $\mathbb{R}$, croissante sur $]-\infty, b]$ et nulle sur $]b, +\infty[$, on a pour tout entier $n \ge 1$

$$\Bigl|\frac{1}{n} \sum_{j\in\mathbb{Z}} \varphi(j/n) - \int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x\Bigr| \le \frac{\varphi(b)}{n}.$$

Si ψ est une fonction ≥ 0 intégrable sur $\mathbb{R}$, croissante sur $]-\infty, b]$ et décroissante sur $[b, +\infty[$, on a pour tout $a \in \mathbb{R}$ et tout $n \ge 1$

$$\Bigl|\frac{1}{n} \sum_{j\in\mathbb{Z}, j<na} \psi(j/n) - \int_{-\infty}^{a} \psi(x)\,\mathrm{d}x\Bigr| \le \frac{2\psi(b)}{n}.$$

(e) Si on examine la preuve du lemme 1, on voit qu'on a en fait : *pour toute fonction f bornée et uniformément continue sur $\mathbb{R}$, il existe une suite de fonctions entières qui tend vers f uniformément sur $\mathbb{R}$*. De son côté, Weierstrass énonce que *pour toute fonction f bornée et continue sur $\mathbb{R}$, il existe une suite de fonctions entières qui tend vers f uniformément sur toute partie bornée de $\mathbb{R}$.*

($^{\mathbf{f}}$) Considérons la fonction triangle T égale à $1-|x|$ pour $|x|\le 1$ et 0 sinon. On voit facilement que

$$(*)\qquad \forall x\in\mathbb{R},\quad \sum_{n\in\mathbb{Z}} \mathrm{T}(x-n)=1.$$

Si $\delta>0$ et si on pose $\varphi_\delta(x)=\mathrm{T}(x/\delta)$, on obtiendra $\sum_{n\in\mathbb{Z}}\varphi_\delta(x-n\delta)=1$, avec des fonctions de support arbitrairement petit (support de longueur 2δ). Si on considère un intervalle fermé borné $[-a,a]$, il est clair qu'on pourra trouver N assez grand pour que

$$\forall x\in[-a,a],\quad \sum_{n=-\mathrm{N}}^{\mathrm{N}}\varphi_\delta(x-n\delta)=1.$$

Une méthode générale pour obtenir la relation $(*)$ est la suivante : on part d'une fonction ψ nulle hors de $[-\delta,\delta]$, de carré sommable, d'intégrale 1 et on pose pour un $b>0$

$$\chi(x)=\psi(x+b)-\psi(x-b).$$

Il est clair que χ est nulle hors de $[-\delta-b,\delta+b]$, et χ est d'intégrale nulle sur $\mathbb{R}$. On pose ensuite

$$\varphi(x)=\int_{-\infty}^{x}\chi(t)\,\mathrm{d}t.$$

On voit que φ est continue, à support dans $[-\delta-b,\delta+b]$. Maintenant, on va prouver que

$$\sum_{n\in\mathbb{Z}}\varphi(x-2nb)=1.$$

Cette série de fonctions est *localement finie*, c'est-à-dire que tout point possède un voisinage dans lequel un nombre fini seulement de ces fonctions est non nul. Il est clair que la somme est constante, par le calcul de la dérivée,

$$\sum_{n\in\mathbb{Z}}\varphi'(x-2nb)=\sum_{n\in\mathbb{Z}}\chi(x-2nb)=\sum_{n\in\mathbb{Z}}\psi(x+b-2nb)-\sum_{n\in\mathbb{Z}}\psi(x-b-2nb)=0,$$

mais la valeur de la constante est moins évidente (sauf si $\delta<b$: dans ce cas $\varphi(0)=1$ et $\varphi(2nb)=0$ pour tout $n\ne 0$) ; calculons-la avec une normalisation qui nous est plus familière, en posant $b=\pi$. La formule de Poisson nous dit que pour toute fonction raisonnable f sur $\mathbb{R}$, on a

$$\sum_{n\in\mathbb{Z}}f(2n\pi)=\frac{1}{2\pi}\sum_{n\in\mathbb{Z}}\widehat{f}(n),$$

où on a posé

$$\forall y\in\mathbb{R},\quad \widehat{f}(y)=\int_{\mathbb{R}}f(x)\,\mathrm{e}^{-\mathrm{i}\,xy}\,\mathrm{d}x.$$

Comme $\chi(x)=\psi(x+\pi)-\psi(x-\pi)$, on voit que

$$\widehat{\chi}(y)=(\mathrm{e}^{\mathrm{i}\,\pi y}-\mathrm{e}^{-\mathrm{i}\,\pi y})\,\widehat{\psi}(y)=2\,\mathrm{i}\sin(\pi y)\,\widehat{\psi}(y).$$

Puisque χ est la dérivée de φ, on a $\widehat{\chi}(y) = \mathrm{i}y\widehat{\varphi}(y)$, donc

$$\frac{1}{2\pi}\,\widehat{\varphi}(y) = \frac{2\mathrm{i}\sin(\pi y)}{2\mathrm{i}\pi y}\,\widehat{\psi}(y) = \frac{\sin(\pi y)}{\pi y}\,\widehat{\psi}(y)$$

où $\sin(\pi y)/(\pi y)$ est interprété comme valant 1 quand $y = 0$; si x est fixé et si on pose $f(u) = \varphi(u + x)$, la formule de Poisson donne

$$\sum_{n\in\mathbb{Z}} \varphi(x + 2n\pi) = \sum_{n\in\mathbb{Z}} f(2n\pi) = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \widehat{f}(n) = \sum_{n\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}xn}\,\frac{\sin(\pi n)}{\pi n}\,\widehat{\psi}(n)$$

où seul le terme $n = 0$ est non nul,

$$\sum_{n\in\mathbb{Z}} f(2n\pi) = \widehat{\psi}(0) = \int_{\mathbb{R}} \psi(x)\,\mathrm{d}x = 1.$$

($^{\mathbf{g}}$) Posons pour h continue sur K

$$\|h\|_{\infty,\mathrm{K}} = \max_{x\in\mathrm{K}} |h(x)|.$$

Supposons $\|f\|_{\infty,\mathrm{K}} = 1$ pour simplifier l'écriture. On commence avec $\varepsilon = \frac{1}{2}$ à trouver par la proposition une fonction g_1 continue sur $\mathbb{R}^d$ telle que $\|f - g_1\|_{\infty,\mathrm{K}} < \frac{1}{2}$ et $\|g_1\|_\infty \le 1$. Ensuite on considère $f_1 = f - g_1$ sur K, on pose $\varepsilon = \frac{1}{4}$ et on trouve g_2 sur $\mathbb{R}^d$ telle que

$$\|f_1 - g_2\|_{\infty,\mathrm{K}} = \|f - g_1 - g_2\|_{\infty,\mathrm{K}} < \frac{1}{4}$$

et $\|g_2\|_\infty \le \|f_1\|_{\infty,\mathrm{K}} < \frac{1}{2}$. Au pas suivant, on pose $f_2 = f_1 - g_2$ et on trouve g_3 telle que $\|f - g_1 - g_2 - g_3\|_{\infty,\mathrm{K}} < \frac{1}{8}$ et $\|g_3\|_\infty < \frac{1}{4}$. En continuant ainsi on définit une série $\sum_{k=1}^{+\infty} g_k$ normalement convergente sur $\mathbb{R}^d$, dont la somme G est continue sur $\mathbb{R}^d$ et égale à f sur K : c'est le prolongement cherché.

($^{\mathbf{h}}$) On sait que $0 \le \varphi_i(x_1) \le 1$; si $|\mathrm{P}_i(x_1) - \varphi_i(x_1)| < \eta$ pour tout $x_1 \in [-a, a]$, il en résulte que $|\mathrm{P}_i(x_1)| < 1 + \eta$; ensuite,

$$|\varphi_i(x_1)\varphi_j(x_2) - \mathrm{P}_i(x_1)\mathrm{P}_j(x_2)| = |(\varphi_i(x_1) - \mathrm{P}_i(x_1))\varphi_j(x_2) + \mathrm{P}_i(x_1)(\varphi_j(x_2) - \mathrm{P}_j(x_2))|$$

qui est $\le \eta + (1 + \eta)\eta$.

($^{\mathbf{i}}$) Le résultat a plusieurs énoncés et plusieurs preuves. Pour l'énoncé que nous avons donné, la preuve suivante est assez naturelle.

On commence par rappeler le fait très connu que si φ est convexe sur un ouvert convexe U, on peut approcher φ uniformément sur tout compact $\mathrm{K} \subset \mathrm{U}$ par des fonctions ψ de la forme $\psi(x) = \max\bigl(a_1(x), \ldots, a_\mathrm{M}(x)\bigr)$ où les a_j sont affines : pour chaque $y \in \mathrm{K}$, on peut trouver une fonction affine d'appui a_y telle que $a_y(y) = \varphi(y)$ et $a_y \le \varphi$ sur U. Soit $\varepsilon > 0$; par continuité l'ensemble $\mathrm{V}_y = \{\varphi - a_y < \varepsilon\}$ est un ouvert de U qui contient y ; par compacité on pourra couvrir K par un nombre fini d'ouverts $\mathrm{U}_{y_1}, \ldots \mathrm{U}_{y_\mathrm{M}}$. Cela implique que $0 \le \varphi - \max(a_{y_1}, \ldots, a_{y_\mathrm{M}}) < \varepsilon$ sur K.

Pour continuer, étant donnée f continue sur $\mathbb{R}^d$ et un compact K, on introduit C convexe et compact tel que K soit contenu dans l'intérieur de C. On peut approcher f uniformément sur C par une fonction g de classe C^2 sur $\mathbb{R}^d$. Il est intuitivement clair que les valeurs propres de la matrice hessienne $\mathrm{H}(y) = \mathrm{D}^2_y g$ des dérivées partielles secondes de g en un point y doivent rester bornées quand y varie dans le compact C, mais pour le justifier, il est difficile de parler de la continuité en y d'une valeur propre individuelle de la matrice hessienne $\mathrm{H}(y)$; on peut employer le stratagème suivant : la quantité

$$\sum_{i,j=1}^{d} |\mathrm{H}_{i,j}(y)|^2$$

(le carré de la norme de Hilbert-Schmidt de la matrice symétrique $\mathrm{H}(y)$) est certainement continue en y, et elle est égale à la somme des carrés des valeurs propres de $\mathrm{H}(y)$. Soit M^2 le maximum sur C de cette quantité ; on en déduit que toutes les valeurs propres de $\mathrm{H}(y)$ sont $\le \mathrm{M}$ quand y varie dans C. Par conséquent

$$\varphi(y) = g(y) + \frac{\mathrm{M}}{2} \sum_{j=1}^{d} y_j^2$$

est convexe sur l'intérieur U de C, et g apparaît comme différence de deux fonctions convexes. Par la remarque préliminaire, on peut approcher g sur le compact $\mathrm{K} \subset \mathrm{U}$ par une fonction de la forme

$$y \in \mathbb{R}^d \to \max\bigl(a_1(y), \ldots, a_{\mathrm{M}}(y)\bigr) - \max\bigl(b_1(y), \ldots, b_{\mathrm{N}}(y)\bigr).$$

Pour survoler une autre preuve, on note que l'ensemble E des fonctions de la forme précédente est un espace vectoriel stable par translation et dilatation, et qu'il contient au moins une fonction ≥ 0 à support compact, non nulle, par exemple

$$\psi(x_1, \ldots, x_d) =$$

$$= \max(x_1 - 1, -x_1 - 1, \ldots, x_d - 1, -x_d - 1, 0) - \max(x_1 - 1, -x_1 - 1, \ldots, x_d - 1, -x_d - 1)$$
$$= \max(0, \min(1 - |x_1|, \ldots, 1 - |x_d|)).$$

Par convolution, on peut approcher f par $f * \psi_n$ avec $\psi_n \in \mathrm{E}$ et par sommes de Riemann, on remplace la convolée $f * \psi_n$ par une combinaison linéaire de translatées de ψ_n, qui est dans E.

Une dernière preuve utilise les triangulations de $\mathbb{R}^d$. Décrivons le cas plus simple où $d = 2$: on peut découper $\mathbb{R}^2$ en triangles, de façon que chaque sommet appartienne à 6 triangles. Pour chaque sommet s de la triangulation, on peut trouver une fonction $\varphi_s \in \mathrm{E}$ égale à 1 en s, nulle en tous les autres sommets et affine sur les six triangles contenant s. La famille des (φ_s) forme une partition de l'unité. Par homothétie on peut supposer que les triangles sont arbitrairement petits ; la formule usuelle $\sum_s f(s)\varphi_s$ donnera alors une fonction de E qui approchera f sur un compact K donné.