Bull. de l'APMEP, septembre-octobre 2006.

Le théorème fondamental de l'algèbre : Une démonstration par le calcul différentiel et l'optimisation

Jean-Baptiste Hiriart-Urruty(*)

« Ce qui est simple est souvent le meilleur »,
P.-P. Riquet (concepteur du canal du Midi).

Ce qu'on appelle le théorème fondamental de l'Algèbre affirme que tout polynôme admet au moins une racine complexe ; on a eu récemment, dans le bulletin de l'APMEP [1], une évocation historique relatant la contribution de J. R. ARGAND. Quand j'étais étudiant en premier cycle universitaire, j'ai entendu de mes professeurs (et pas seulement moi) « qu'on allait tout nous démontrer dans les cours de mathématiques ... sauf le théorème fondamental de l'Algèbre » ; bref sa démonstration nous était mentionnée comme un Graal inaccessible à notre niveau. Or il n'en est rien... Il existe une multitude de démonstrations possibles, dont certaines sont vraiment accessibles en deuxième année d'études en mathématiques post Bac (même après la réforme LMD...). Des articles de revue comme [2, 3], le chapitre 4 de la partie A de [6], un site web :

http ://math.fullerton.edu/mathews/c2002/funtheorem/funtheorem.html,

y sont consacrés. Y a t-il des démonstrations du théorème fondamental de l'Algèbre qui soient plus simples que les autres ? La réponse est une question de goût et, comme souvent en mathématiques, tout dépend de ce que l'on suppose connu avant d'en aborder la démonstration. Reconnaissons toutefois que dans la plupart des démonstrations bien connues et diffusées, les techniques et résultats concernant les fonctions de la variable complexe y jouent un rôle prépondérant [2]. Nous proposons ci-dessous une démonstration très différente (au moins dans la dernière étape), basée sur le Calcul différentiel (tel qu'on l'apprend en première année d'université) et l'Optimisation (également des programmes de première et deuxième années d'université). L'idée de cette manière de conclure n'est pas de moi, je l'ai apprise dans divers écrits pédagogiques sous la direction d'un collègue russe, A. V. TIKHOMIROV, professeur à l'université de Moscou [4, 5]. J'ai pensé qu'il était utile de la faire mieux connaître.

1. Appuis sur un socle commun de connaissances...

– Soit $P(z) = a_0 + a_1 z + \ldots + a_n z^n$ le polynôme à coefficients complexes a_k dont on va s'occuper, avec $n \geq 1$ et $a_n \neq 0$ (et également $a_0 \neq 0$, sinon il n'y a rien à démontrer, 0 est une racine). Notons tout de suite qu'on ne perd pas en généralité en supposant les a_k réels ; en effet, si $Q(z) := P(z)\overline{P(\overline{z})}$, Q est un nouveau polynôme, à coefficients

(*) Département de Mathématiques. Université Paul Sabatier de Toulouse.

réels cette fois, et si z^* est une racine de Q, z^* ou $\overline{z^*}$ est une racine de P ; ceci nous permettrait d'alléger un petit peu les notations mais sans plus.

– My TAYLOR is everywhere…

Nous sommes en première année d'études post Bac. Quand on nous demande de minimiser une fonction $\varphi : \mathbf{R} \to \mathbf{R}$ dérivable (suffisamment de fois), on passe par les dites conditions nécessaires de minimalité dont voici la version la plus simple : en 0 supposé minimiser (localement) φ, on a nécessairement $\varphi'(0) = 0$ et $\varphi''(0) \geq 0$. En général, on s'en tient là… Le problème est que ces conditions ne sont pas informatives si $\varphi''(0) = 0$, les fonctions polynomiales t^k, avec $k \geq 3$, sont là pour illustration ; il faut donc aller regarder plus loin. Considérons pour cela ce qu'on appelle *les conditions nécessaires d'optimalité d'ordre supérieur*. Elles sont directement déduites de la formule de développement de TAYLOR-YOUNG au point candidat 0. Voici ce que cela donne. Soit $\varphi : \mathbf{R} \to \mathbf{R}$ une fonction indéfiniment dérivable en 0, telle que :

- 0 est un point de minimisation locale pour φ ;
- il existe p tel que $\varphi^{(p)}(0) \neq 0$.

Alors, si s désigne le premier entier pour lequel $\varphi^{(s)}(0) \neq 0$ ($s \geq 2$ forcément), s est nécessairement pair et $\varphi^{(s)}(0) > 0$.

À vrai dire, nous n'allons utiliser ce résultat que dans un contexte on ne peut plus simple, celui des fonctions φ polynomiales.

– Quand la compacité fait défaut…

Nous sommes en deuxième année d'études et on nous demande de minimiser une fonction continue (de deux variables) $f : \mathbf{R}^2 \to \mathbf{R}$ sur une partie compacte de $\mathbf{R}^2$. Le théorème de WEIERSTRASS est là pour assurer l'existence d'un point de minimisation (et un de maximisation d'ailleurs). Quid quand on nous demande de minimiser sur tout le plan $\mathbf{R}^2$? Eh bien, la compacité peut être compensée par une « coercivité à l'infini » de f, c'est-à-dire en forçant $f(x, y)$ à tendre vers l'infini quand la norme de (x, y) tend vers l'infini. De façon plus précise, étant donné $f : \mathbf{R}^2 \to \mathbf{R}$, si $f(x, y) \to +\infty$ quand $\|(x, y)\| \to +\infty$, on est sûr que f est bornée inférieurement sur $\mathbf{R}^2$ et que sa borne inférieure est atteinte. La démonstration en est très simple : soit M une valeur prise par f ; puisque f est coercive à l'infini, il existe r tel que les valeurs de $f(x, y)$ dépassent M quand $\|(x, y)\|$ est supérieur à r ; inutile donc d'aller minimiser f en dehors de la boule fermée de centre 0 et de rayon r : minimiser f sur $\mathbf{R}^2$ équivaut à minimiser f sur cette partie compacte qu'est la boule fermée de centre 0 et de rayon r ; on est ainsi ramené au contexte habituel d'application du théorème de WEIERSTRASS.

2. La démonstration elle-même.

Considérons la fonction suivante $f(z) := |P(z)|^2$; mettre au carré est important pour « lisser » les choses (penser à la technologie des « moindres carrés »), la fonction module au carré ayant le bon goût d'être dérivable. La fonction f peut être vue comme une fonction du nombre complexe $z = x + iy$, ou bien des deux variables réelles x et y, comme on voudra.

Dans la première approche correcte de la démonstration du théorème fondamental de l'Algèbre, ARGAND (au début du 19ème siècle) avait considéré la minimisation de $z \to |P(z)|$ sur **C**, mais il n'avait pas justifié que l'infimum de $|P(z)|$ sur **C** était atteint ; la même difficulté apparaît chez GAUSS et CAUCHY.

Revenons à notre fonction f. Premières observations : f est continue et $f(z) \to +\infty$ quand $|z| \to +\infty$. Voyons cela.

La continuité ne pose pas de difficulté. Quant à la deuxième propriété annoncée, observons la « dominance du terme de plus haut degré » dans l'expression polynomiale en $|z|$ apparaissant ci-dessous :

$$f(z) = \left|\sum_{k=0}^{n} a_k z^k\right| \geq \left(|a_n| \cdot |z|^n - \sum_{k=0}^{n-1} |a_k| \cdot |z|^k\right)^2$$

$$\geq |a_n|^2 \cdot |z|^{2n}\left(1 + \mathrm{O}\left(\frac{1}{|z|}\right)\right) \to +\infty \text{ quand } |z| \to +\infty.$$

Cette condition de coercivité de f sur **C**, alliée à sa continuité, nous assure donc que f est bornée inférieurement sur **C** et que sa borne inférieure est atteinte : il existe z^* tel que $f(z^*) = \inf_{\mathbf{C}} f$. Il s'agit à présent de montrer que ce z^* fait notre affaire, c'est-à-dire que $f(z^*) = 0$.

Note : sachant que f est positive, et que l'on souhaite mettre en évidence un point où f s'annule, il est naturel de penser à *minimiser f*.

Maintenant qu'on tient ce point de minimisation z^*, il est tout aussi naturel *d'exploiter les conditions nécessaires d'optimalité en ce point.* Pour cela, et dans un souci de simplification d'écriture, on commence par se décaler vers l'origine en considérant $\tilde{P}(z) := P(z^* + z)$. De fait, $\tilde{P}(z)$ est également polynomial de degré n :

$$\tilde{P}(z) = b_0 + b_1 z + \ldots + b_n z^n,$$

et notre objectif est de démontrer que 0 est racine de $\tilde{P}$, **c'est-à-dire que $b_0 = 0$.**

Dans l'expression ci-dessus de $\tilde{P}(z)$, désignons par s le premier entier ≥ 1 pour lequel $b_1 = b_2 = \ldots = b_{s-1} = 0$ et $b_s \neq 0$; ainsi $\tilde{P}(z) = b_0 + b_s z^s + \ldots + b_n z^n$. À présent, pour $\theta \in [0, 2\pi[$, considérons la fonction numérique φ_θ *de la variable réelle t* définie par $\varphi_\theta(t) := \left|\tilde{P}(t\,\mathrm{e}^{\mathrm{i}\theta})\right|^2$ (ça n'est rien d'autre que la fonction f évaluée le long d'une droite passant par z^* et dirigée par $\mathrm{e}^{\mathrm{i}\theta}$). Bien évidemment, φ_θ est minimisée pour $t = 0$, et ce, **quel que soit** $\theta \in [0, 2\pi[$. C'est cette propriété que nous allons exploiter. Détaillons un peu plus l'expression polynomiale en la variable t de $\varphi_\theta(t)$:

$$\varphi_\theta(t) = \left(b_0 + b_s t^s \mathrm{e}^{\mathrm{i}s\theta} + \ldots + b_n t^n \mathrm{e}^{\mathrm{i}n\theta}\right)\left(\overline{b_0} + \overline{b_s} t^s \mathrm{e}^{-\mathrm{i}s\theta} + \ldots + \overline{b_n} t^n \mathrm{e}^{-\mathrm{i}n\theta}\right)$$

$$= |b_0|^2 + \left(b_s \overline{b_0} t^s \mathrm{e}^{\mathrm{i}s\theta} + b_0 \overline{b_s} t^s \mathrm{e}^{-\mathrm{i}s\theta}\right) + \text{termes en } t^k \text{ avec } k \geq s+1$$

$$= |b_0|^2 + 2t^s \,\mathrm{Re}(b_s \overline{b_0}\, \mathrm{e}^{\mathrm{i}s\theta}) + \mathrm{o}(t^s).$$

Par suite, s ≥ 2 (puisque $\varphi'_\theta(0) = 0$) et :

$$\varphi_\theta^{(k)}(0) = 0 \text{ pour } k = 1, \ldots, s-1$$

$$\varphi_\theta^{(s)}(0) = 2s!\,\mathrm{Re}(b_s \overline{b_0}\, \mathrm{e}^{\mathrm{i}s\theta})$$

(bien garder à l'esprit que s est indépendant de θ et que $s \geq 2$).

Raisonnons par l'absurde en supposant $b_0 \neq 0$. Sachant que $b_s \neq 0$ (par définition), la quantité $\mathrm{Re}(b_s \overline{b_0}\, \mathrm{e}^{\mathrm{i}s\theta})$ peut encore s'écrire sous forme développée :

$$\mathrm{Re}[(u+iv)\cdot(\cos s\theta + \mathrm{i}\sin s\theta)]$$

avec $u + iv := b_s \overline{b_0} \neq 0$, soit encore

$$u\cos s\theta - v\sin s\theta = \sqrt{u^2+v^2}\,\cos(s\theta + \rho).$$

Cette expression – et donc $\varphi_\theta^{(s)}(0) = 2s!\,\mathrm{Re}(b_s \overline{b_0}\, \mathrm{e}^{\mathrm{i}s\theta})$ – balaye tout l'intervalle $\left]-\sqrt{u^2+v^2}, \sqrt{u^2+v^2}\right[$ quand θ parcourt l'intervalle $[0, 2\pi[$. Or les conditions d'optimalité d'ordre supérieur nous indiquent que le premier k pour lequel $\varphi_\theta^{(k)}(0) \neq 0$ est pair et $\varphi_\theta^{(k)}(0) \geq 0$. On devrait donc avoir s pair (jusque là pas de problème) et $\varphi_\theta^{(s)}(0) \geq 0$, ce qui n'est pas le cas pour tous les $\theta \in [0, 2\pi[$. L'hypothèse de départ est donc absurde, c'est son contraire qui est vrai : $b_0 = 0$. La démonstration est terminée.

Références.

[1] O. Kouteynoff, *La démonstration par Argand du théorème fondamental de l'algèbre*, Bulletin de l'APMEP n° 462 (2006), p. 122-137.

[2] E. Fieux, P. Lassère, F. Rodriguez, *Approches analytiques du théorème de D'Alembert-Gauss : un bestiaire.* À paraître dans la Revue des mathématiques de l'enseignement supérieur (ex-Revue de mathématiques spéciales).

[3] E. Fieux, P. Lassère, F. Rodriguez, *Le théorème de D'Alembert-Gauss, retour du bestiaire : option topologie et analyse réelle.* En préparation.

[4] V. Alexeev, V. Tikhomirov, S. Fomine, *Commande optimale*, Éditions Mir (1982).

[5] V. Alexeev, E. Galeev, V. Tikhomirov, *Recueil de problèmes d'optimisation*, Éditions Mir (1987).

[6] F. Guénard (traducteur), *Les nombres. Leur histoire, leur place et leur rôle de l'Antiquité aux recherches actuelles*, Éditions Vuibert (1998).

[7] J.-B. Hiriart-Urruty, *L'Optimisation.* Collection « Que sais-je ? », Presses Universitaires de France (1996).