Balade newtonienne entre analyse et arithmétique

Journées X-UPS 2023

Antoine Chambert-Loir

Université Paris Cité, Institut de Mathématiques de Jussieu-Paris Rive Gauche, F-75013, Paris, France
Courriel: *antoine.chambert-loir@u-paris.fr*

Résumé. — Inventés par Kurt Hensel à la toute fin du 19e siècle sur le modèle des séries en une indéterminée, les nombres p-adiques sont devenus non seulement un outil indispensable de l'arithmétique contemporaine, mais un sujet d'étude en soi. Dans ce texte, issu de deux exposés aux journées X-UPS 2023, j'expliquerai leur construction, comment ils s'insèrent dans un cadre plus vaste, entre analyse et arithmétique, où deux constructions portant le nom d'Isaac Newton jouent un rôle central: la méthode de Newton et la notion de polygone de Newton.

Abstract. — Invented by Kurt Hensel at the very end of 19th century on the model of power series in one indeterminate, the p-adic numbers have not only become an indispensable tool of contemporary arithmetic, but a research topic per se. In this text, stemming out two talks at the 2023 X-UPS lectures, I shall explain their construction, how they fit in a vaster framework, between analysis and arithmetic, where two constructions bearing the name of Isaac Newton play a central role : Newton's method and Newton's polygon.

Table des matières

1. Valeurs absolues, topologies, exemples 2
2. Complétude, locale compacité. Normes 10
3. Méthode de Newton 19
4. Polytopes et polygones de Newton 29
5. Fonctions et équations analytiques 38

Références 49

1. Valeurs absolues, topologies, exemples

***Définition 1.1* (Kürschák, 1913).** — *Soit* A *un anneau commutatif. On appelle* valeur absolue*, ou* semi-norme multiplicative*, sur* A *une application* $m\colon \mathrm{A} \to \mathbf{R}_+$ *vérifiant les propriétés suivantes :*

(1) *On a* $m(0) = 0$ *et* $m(1) = 1$ (normalisation)*;*
(2) *Pour* $a, b \in \mathrm{A}$*, on a* $m(ab) = m(a)m(b)$ (multiplicativité)*;*
(3) *Pour* $a, b \in \mathrm{A}$*, on a* $m(a + b) \leqslant m(a) + m(b)$ (inégalité triangulaire).

Exemple 1.2. — Les lecteurices de ce texte connaissent probablement deux exemples de valeurs absolues : la valeur absolue usuelle sur le corps $\mathbf{R}$ des nombres réels, et le module sur le corps $\mathbf{C}$ des nombres complexes.

1.3. — Soit A un anneau commutatif et soit m une valeur absolue sur A.

Supposons que A est un corps. Alors, 0 est le seul élément de A de valeur absolue nulle. En effet, si $a \in \mathrm{A}$ est non nul, on a $1 = m(1) = m(a \cdot a^{-1}) = m(a) \cdot m(a^{-1})$, donc $m(a) \neq 0$.

C'est le cas le plus important, et l'on peut s'y ramener comme suit. Appelons *noyau* de la valeur absolue l'ensemble P des $a \in \mathrm{A}$ tels que $m(a) = 0$; c'est un idéal premier de A. Il existe sur l'anneau quotient A/P une unique application telle que $m(\overline{a}) = m(a)$ pour tout $a \in \mathrm{A}$. Cette application est encore une valeur absolue. Par construction, elle ne s'annule qu'en 0. Il existe alors sur le corps des fractions F de A/P une unique application m telle que $m(\overline{a}/\overline{b}) = m(\overline{a})/m(\overline{b})$ pour tous $\overline{a}, \overline{b} \in \mathrm{A}/\mathrm{P}$ tels que $m(\overline{b}) \neq 0$. Cette application est encore une valeur absolue.

1.4. — La motivation initiale de Kürschák (1913) pour l'introduction de la notion de valeur absolue sur un corps commutatif F est qu'elle permet de définir une notion de limite sur F. Plus d'un siècle plus tard, expliquons comment faire dans le cas plus général d'un anneau.

Soit donc m une valeur absolue sur un anneau A. De la définition, on déduit que $(x, y) \mapsto m(x - y)$ est une semi-distance sur A : la symétrie découle de ce que $1 = m(1) = m(-1)^2$ (propriété (1.1/2)), donc $m(-1) = 1$, et l'inégalité triangulaire est la propriété (1.1/2). Cette semi-distance est une distance si et seulement si m ne s'annule qu'en 0.

Pour $a \in \mathrm{A}$ et $r \in \mathbf{R}_+$, on peut alors définir les boules $\mathrm{B}(a, r)$ et $\mathrm{B}^\circ(a, r)$, respectivement circonférenciées et non circonférenciées, dans A : ce sont

les ensembles des $x \in \mathrm{A}$ tels que $m(x-a) \leqslant r$, resp. $m(x-a) < r$. La semi-distance m définit une topologie sur A pour laquelle les voisinages d'un point $a \in \mathrm{A}$ sont les ensembles qui contiennent une boule non circonférenciée $\mathrm{B}^\circ(a,r)$, avec $r > 0$.

Proposition 1.5. — *Soit* A *un anneau commutatif et soit m une valeur absolue sur* A*; munissons* A *de la topologie définie par m.*

(1) *L'addition et la multiplication de* A *sont continues,*
(2) *Si* A *est un corps, l'inversion est continue sur* $\mathrm{A}^\times$.
(3) *L'application m est uniformément continue.*
(4) *La topologie de* A *est séparée si et seulement si m ne s'annule qu'en* 0.

Démonstration. — La continuité de l'addition résulte de l'inégalité triangulaire : pour $a, b, x, y \in \mathrm{A}$, on écrit

$$m((x+y)-(a+b)) = m((x-a)+(y-b)) \leqslant m(x-a)+m(y-b),$$

si bien que $x+y$ tend vers $a+b$ lorsque $x \to a$ et $y \to b$. Pour la multiplication, on écrit de même

$$\begin{aligned} m(xy-ab) &= m((x-a)(y-b)+a(y-b)+(x-a)b) \\ &\leqslant m(x-a)m(y-b)+m(a)m(y-b)+m(x-a)m(b), \end{aligned}$$

si bien que xy tend vers ab lorsque $x \to a$ et $y \to b$.

Supposons maintenant que A soit un corps et démontrons que l'inversion est continue. Remarquons que si $m(a) \leqslant 1/2$, alors $m(1/(1-a)) \leqslant 2$. En effet, on écrit $\frac{1}{1-a} = 1 + \frac{a}{1-a}$, donc

$$m(1/(1-a)) \leqslant 1 + m(a)m(1/(1-a)) \leqslant 1 + \frac{1}{2}m(1/(1-a)),$$

d'où on déduit l'inégalité voulue. Alors, pour $a, x \in \mathrm{A}^*$, on écrit

$$\frac{1}{x} - \frac{1}{a} = \frac{a-x}{ax} = \frac{a-x}{a^2}\frac{1}{1-(a-x)/a}.$$

Si $m(a-x) \leqslant \frac{1}{2}m(a)$, alors $m((a-x)/a)) \leqslant 1/2$, puis

$$m(1/x - 1/a) \leqslant 2m(a-x)/m(a^2),$$

d'où la continuité requise.

L'uniforme continuité de m résulte de la variante

$$|m(a) - m(b)| \leqslant m(a-b)$$

de l'inégalité triangulaire : comme $a = (a - b) + b$, on a $m(a) \leqslant m(a - b) + m(b)$, d'où $m(a) - m(b) \leqslant m(a - b)$; par symétrie, $m(b) - m(a) \leqslant m(b - a) = m(a - b)$.

Si m ne s'annule qu'en 0, l'application $(x, y) \mapsto m(x - y)$ est une distance, donc la topologie de A est séparée. Inversement, un élément $a \in \mathrm{A}$ tel que $m(a) = 0$ appartient à tout voisinage de 0, donc est nul si la topologie de A est séparée. □

***Exemple 1.6* (Valeur absolue triviale).** — Sur n'importe quel corps commutatif F, on peut également observer que l'on définit également une valeur absolue m en posant $m(0) = 0$ et $m(a) = 1$ pour tout $a \in \mathrm{F}$ tel que $a \neq 0$. Cette valeur absolue est dite *triviale* ; d'ailleurs, la définition originelle d'une valeur absolue par Kürschák (1913) l'excluait.

Pour la valeur absolue triviale, la boule non circonférenciée de centre 0 et de rayon 1 ne contient que 0. La topologie de F est donc *discrète*.

Inversement, supposons que la topologie de F définie par une valeur absolue m soit discrète, et démontrons que m est la valeur absolue triviale. Par hypothèse, il existe un nombre réel $r > 0$ tel que la boule ouverte $\mathrm{B}(0, r)$ ne contienne que le point 0. Soit $x \in \mathrm{F}^\times$ et démontrons que $m(x) = 1$. Quitte à changer x en $1/x$, on peut supposer que $m(x) < 1$. Il existe alors un entier $n \geqslant 1$ tel que $m(x)^n < r$; puisque $m(x^n) = m(x^n)$, on a donc $m(x^n) < r$, d'où $x^n = 0$ d'après le choix de r. Ainsi, $x = 0$.

***Exemple 1.7* (Valeurs absolues p-adiques).** — Considérons le cas du corps $\mathbf{Q}$ des nombres rationnels. Soit p un nombre premier. La valeur absolue p-adique m_p que nous allons définir est essentiellement équivalente à la *valuation p-adique* v_p qui est peut-être plus classique.

Si a est un nombre rationnel non nul, il existe un unique entier relatif n telle que a puisse s'écrire $a = p^n b/c$, où b et c sont des entiers non multiples de p. En effet, si $a = p^n b/c = p^{n'} b'/c'$, alors $p^n bc' = p^{n'} b'c$, et l'égalité $n = n'$ découle de l'unicité de la décomposition en facteurs premiers. Cet entier n s'appelle la valuation p-adique de a ; on le note $v_p(a)$. Posons alors $m_p(a) = p^{-v_p(a)}$.

On pose enfin $m_p(0) = 0$ et $v_p(0) = +\infty$.

Vérifions que m_p est une valeur absolue sur $\mathbf{Z}$; on l'appelle la *valeur absolue p-adique*. La propriété (1.1/1) est évidente et la propriété (1.1/2) découle de la définition. La propriété (1.1/3) est en fait vraie sous une forme plus forte :

(4) Pour $a, b \in \mathbf{Q}$, on a $m_p(a+b) \leqslant \sup(m_p(a), m_p(b))$ (*inégalité ultramétrique*).

Il suffit de traiter le cas où a et b sont non nuls ; on écrit $a = p^n u/v$, $b = p^m r/s$, où u, v, r, s sont des entiers non multiples de p. Si $m \leqslant n$, alors $a + b = p^m(p^{n-m}u/v + r/s) = p^m(p^{n-m}us + rv)/sv$; si m' est l'exposant de p dans la décomposition en facteurs premiers de $p^{n-m}us + rv$, on a donc $v_p(a+b) = m + m'$, et en particulier $v_p(a+b) \geqslant m = v_p(b) = \inf(v_p(a), v_p(b))$, d'où l'inégalité voulue. Le cas où $n \leqslant m$ est symétrique.

Remarquons tout de suite la *formule du produit :*

$$m_\infty(a) \prod_{p \text{ premier}} m_p(a) = m_0(a)$$

qui relie valeur absolue usuelle m_∞, les valeurs absolues p-adiques m_p, où p parcourt l'ensemble des nombres premiers, et la valeur absolue triviale m_0 sur $\mathbf{Q}$. L'égalité est évidente si $a = 0$. Si $a \neq 0$, il s'agit de démontrer

$$m_\infty(a) \prod_{p \text{ premier}} m_p(a) = 1\,;$$

on écrit pour cela la décomposition en facteurs premiers, $a = \pm \prod_p p^{v_p(a)}$, si bien que

$$m_\infty(a) \prod_p m_p(a) = \left(\prod_p p^{v_p(a)}\right) \prod_p p^{-v_p(a)} = 1 = m_0(a).$$

Exemple 1.8. — Le cas où F est le corps des fractions rationnelles en une indéterminée T à coefficients dans un corps commutatif k est particulièrement important. Pour tout polynôme irréductible $\pi \in k[\mathrm{T}]$, on dispose d'une *valuation* π-adique v_π sur $k(\mathrm{T})$, telle que $v_\pi(f) = n$ si $f \in k(\mathrm{T})$ s'écrit sous la forme $\pi^n g/h$, où g et h sont des polynômes non divisibles par π. On en déduit une *valeur absolue* π-adique sur $k(\mathrm{T})$, définie par $m_\pi(f) = e^{-v_\pi(f)\deg(\pi)}$, où la normalisation choisie ($c = e^{-\deg(\pi)}$) anticipe la formule du produit.

On dispose également d'une valuation « à l'infini » v_∞, définie par $v_\infty(g/h) = -\deg(g) + \deg(h)$ si g et h sont des polynômes non nuls, et $v_\infty(0) = +\infty$. Enfin, on pose $m_\infty(f) = e^{-v_\infty(f)}$ pour tout $f \in k(\mathrm{T})$.

Ces normalisations donnent lieu à la formule du produit : pour toute fraction rationnelle $f \in k(\mathrm{T})$, on a l'égalité

$$m_\infty(f) \cdot \prod_\pi m_\pi(f) = m_0(f)$$

qui traduit simplement l'égalité

$$\sum_\pi v_\pi(f) \deg(\pi) = \deg(f)$$

si f est un polynôme non nul.

Exemple 1.9. — L'exemple des valeurs absolues p-adiques sur $\mathbf{Q}$ et celui des valeurs absolues sur un corps de fractions rationnelles en une indéterminée associées à un polynôme irréductible se généralisent *mutatis mutandis* au cas du corps des fractions F d'un anneau factoriel A, en remplaçant le nombre premier p ou le polynôme irréductible π par un élément irréductible π de A.

On commence en effet par définir la valuation π-adique d'un élément $f \in \mathrm{F}$ par $v_\pi(f) = +\infty$ si $f = 0$, et par $v_\pi(f) = n$ si n est un entier tel qu'il existe u et v dans A, non multiples de π, tels que $f = \pi^n u/v$. Pour tout nombre réel $c > 1$, on en déduit une *valeur absolue* π-adique sur F, définie par $m_\pi(f) = c^{-v_\pi(f)}$. Elle est ultramétrique.

Il n'y a pas, en général, de façon privilégiée de choisir c. De toutes façons, lorsque c varie, ces valeurs absolues définissent toutes la même topologie.

Lemme 1.10. — *Soit m une valeur absolue sur un anneau commutatif* A. *Les propriétés suivantes sont équivalentes :*

(1) *La valeur absolue m est ultramétrique;*
(2) *Pour tout $a \in \mathrm{A}$ tel que $m(a) \leqslant 1$, on a $m(1+a) \leqslant 1$;*
(3) *Pour tout $n \in \mathbf{N}$, on a $m(n) \leqslant 1$.*
(4) *La suite $(m(n))$ ne tend pas vers l'infini.*

Démonstration. — L'implication (1)⇒(2) est évidente : si m est ultramétrique et si $a \in \mathrm{A}$ vérifie $m(a) \leqslant 1$, alors $m(1+a) \leqslant \sup(m(1), m(a)) = 1$.

L'implication (2)⇒(3) se démontre par une récurrence immédiate.

Prouvons l'implication (3)⇒(1). Soit donc $a, b \in \mathrm{A}$. Soit n un entier naturel non nul; la formule du binôme pour $(a+b)^n$ et l'inégalité triangulaire

entraînent

$$m(a+b)^n = m((a+b)^n) \leqslant \sum_{k=0}^{n} m\Big(\tbinom{n}{k} a^{n-k} b^k\Big) \leqslant \sum_{k=0}^{n} m(a)^{n-k} m(b)^k$$

puisque $m\big(\tbinom{n}{k}\big) \leqslant 1$ pour tout k. On en déduit l'inégalité

$$m(a+b)^n \leqslant (n+1)\sup(m(a), m(b))^n.$$

En prenant les racines n-ièmes et en faisant tendre n vers l'infini, on en déduit l'inégalité ultramétrique. □

Remarque 1.11. — Soit F un corps muni d'une valeur absolue m qui n'est pas ultramétrique. Alors, F est de caractéristique zéro. En effet, il existerait sinon un entier n tel que $m(n)$ soit maximal ; en particulier, $m(n^2) \leqslant m(n)$, d'où $m(n) \leqslant 1$ puisque $m(n) \neq 0$.

Proposition 1.12. — *Soit m et m' des valeurs absolues sur un corps commutatif* F. *Si la topologie définie par m' est plus fine que la topologie définie par m, il existe un nombre réel $\rho \geqslant 0$ tel que $m'(x) = m(x)^\rho$ pour tout $x \in \mathrm{F}^\times$, et inversement.*

Démonstration. — Soit B l'ensemble des $x \in \mathrm{F}$ tels que $m(x) \leqslant 1$; c'est un voisinage de 0 pour la topologie définie par m, donc c'est aussi un voisinage de 0 pour la topologie définie par m'. Il existe donc un nombre réel r tel que $0 < r < 1$ et tel que pour tout $x \in \mathrm{F}$ tel que $m'(x) \leqslant r$, on ait $m(x) \leqslant 1$.

Supposons que 0 soit le seul élément de F tel que $m'(x) \leqslant r$; alors la topologie de F définie par m' est la topologie discrète, donc m' est la valeur absolue triviale de F (exemple 1.6) et l'on a $m'(x) = m(x)^0$ pour tout $x \in \mathrm{F}$.

Il existe sinon un élément $a \in \mathrm{F}$ tel que $0 < m'(a) \leqslant r$; posons $\rho = \log(m(a))/\log(m'(a))$. Soit x un élément de F tel que $m'(x) < 1$. Pour tout entier $n \geqslant 1$, il existe un unique entier $p \geqslant 1$ tel que $m'(a)^{p+1} < m'(x)^n \leqslant m'(a)^p$; en fait, p est la partie entière de $n\log(m'(x))/\log(m'(a))$. Comme $m'(x^n/a^p) = m'(x)^n/m'(a)^p \leqslant 1$, on a $m'(a \cdot x^n/a^p) \leqslant m'(a) \leqslant r$, donc $m(a \cdot x^n/a^p) \leqslant 1$ puis $m(x) \leqslant m(a)^{(p-1)/n}$. En faisant tendre n vers l'infini, on obtient $m(x) \geqslant m(a)^{\log(m'(x))/\log(m'(a))} = m'(x)^\rho$. De même, comme $m'(a^{p+1}/x^n) = m'(a)^{p+1}/m'(x)^n < 1$, on a $m'(a \cdot a^{p+1}/x^n) \leqslant m'(a) \leqslant r$, donc $m(a \cdot a^{p+1}/x^n) \leqslant 1$ puis $m(x) \geqslant m(a)^{(p+2)/n}$. En faisant tendre n vers l'infini, on obtient $m(x) \geqslant m'(x)^\rho$. □

Exemples 1.13. — (1) La topologie discrète est plus fine que toute topologie, et est donnée par la valeur absolue triviale. Cela correspond au cas $\rho = 0$ dans la proposition.

(2) D'après cette proposition, les valeurs absolues m sur $\mathbf{R}$ qui définissent une topologie plus fine que la topologie usuelle sont de la forme $m(x) = |x|^\rho$, où ρ est un nombre réel $\geqslant 0$. Le cas $\rho = 0$ correspond à la topologie discrète. On laisse aux lecteurices le soin de vérifier que l'on a $\rho \leqslant 1$ et qu'inversement, toute telle valeur correspond bien à une valeur absolue sur $\mathbf{R}$.

Le cas du corps des nombres complexes est analogue.

On peut cependant considérer sur $\mathbf{R}$ ou $\mathbf{C}$ des valeurs absolues très exotiques au sens où les topologies définies ne sont pas comparables à la topologie classique. Un exemple fameux est donné par la démonstration de Monsky (1970) du théorème selon lequel dans toute décomposition d'un carré en un nombre fini triangles de même aire, le nombre de triangles de la décomposition est pair. Dans cette preuve, Monsky introduit une valeur absolue 2-adique sur $\mathbf{R}$, pour laquelle $m(2) = 1/2$.

(3) Les topologies p-adiques sur $\mathbf{Q}$ sont deux à deux incomparables, et elles sont incomparables à la topologie définie par la valeur absolue usuelle. Elles sont toutes moins fines que la topologie discrète définie par la valeur absolue triviale.

***Théorème 1.14* (Ostrowski, 1916).** — (1) *Soit m une valeur absolue sur $\mathbf{Q}$. Alors m est la valeur absolue triviale, une puissance de la valeur absolue usuelle, ou une puissance d'une valeur absolue p-adique.*

(2) *Soit k un corps et soit m une valeur absolue sur $k(\mathrm{T})$ dont la restriction à k est la valeur absolue triviale. Alors m est la valeur absolue triviale, une puissance de la valeur absolue associée au degré, ou une puissance d'une valeur absolue associée à un polynôme irréductible.*

Démonstration. — (1) Soit a un entier naturel tel que $a > 1$. Soit x un entier $\geqslant 1$; soit n un entier $\geqslant 1$ et soit p le plus petit entier tel que $x^n < a^p$ et soit $x^n = x_0 + ax_1 + \cdots + a^{p-1}x_{p-1}$ la décomposition en base a de x^n, où $0 \leqslant x_0, \dots, x_{p-1} \leqslant a-1$. On en déduit $m(x^n) \leqslant \sup(m(1), \dots, m(a-1))pm(a)^p$, d'où

$$m(x) \leqslant \sup(m(1), \dots, m(a-1))^{1/n} p^{1/n} \sup(1, m(a))^{p/n}.$$

En faisant tendre n vers l'infini, p/n tend vers $\log(x)/\log(a)$ et on obtient l'inégalité $m(x) \leqslant \sup(1, m(a))^{\log(x)/\log(a)}$, d'où $m(x)^{1/\log(x)} \leqslant \sup(1, m(a))^{1/\log(a)}$.

S'il existe $x \geqslant 1$ tel que $m(x) > 1$, on en déduit que $m(a) > 1$ pour tout entier $a > 1$, puis que $m(x)^{1/\log(x)} \leqslant m(a)^{1/\log(a)}$ pour tous $a, x > 1$. En notant ρ cette valeur constante, on a donc $m(x) = x^\rho$ pour tout entier $x > 1$, puis $m(x) = |x|^\rho$ pour tout $x \in \mathbf{Q}$.

Sinon, on a $m(x) \leqslant 1$ pour tout entier $x \geqslant 2$: la valeur absolue m est ultramétrique.

Si $m(n) = 1$ pour tout entier non nul n, alors $m(a) = 1$ pour tout nombre rationnel non nul a, et m est la valeur absolue triviale.

Sinon, soit p le plus petit entier $\geqslant 1$ tel que $m(p) < 1$. On a $p > 1$; si u et v sont des entiers > 1 tels que $p = uv$, alors $m(u) = m(v) = 1$ par minimalité de p, donc $m(p) = 1$. Ainsi, p est un nombre premier. Soit u un entier tel que $m(u) < 1$ et soit $u = kp + r$ sa division euclidienne par p; si $r \neq 0$, on a $m(r) = 1$ par minimalité de p, donc $m(p) \geqslant m(k)m(p) = m(kp) = m(r - u) = 1$, et cela contredit la définition de p. Ainsi, $r = 0$ et u est multiple de p. On en déduit aisément que $m(a) = m(p)^{v_p(a)} = m_p(a)^\rho$, où $\rho = -\log(m(p))/\log(p)$.

(2) Dans ce cas, m est ultramétrique.

Si $m(\mathrm{T}) \leqslant 1$, alors $m(f) \leqslant 1$ pour tout polynôme $f \in k[\mathrm{T}]$. Soit I l'ensemble des polynômes $f \in k[\mathrm{T}]$ tels que $m(f) < 1$; c'est un idéal de $k[\mathrm{T}]$, et c'est même un idéal *premier*. Si $\mathrm{I} = (0)$, alors $m(f) = 1$ pour tout polynôme non nul f, d'où l'on déduit que m est la valeur absolue triviale sur $k(\mathrm{T})$. Sinon, il existe un polynôme irréductible π tel que $\mathrm{I} = (\pi)$, et on en déduit que m est une puissance de la valeur absolue m_π.

Il reste à traiter le cas où $m(\mathrm{T}) > 1$. Dans ce cas, on a $m(a\mathrm{T}^n) = m(\mathrm{T})^n$ pour tout $a \in k^\times$, et l'inégalité ultramétrique entraîne que $m(f) = m(\mathrm{T})^{\deg(f)}$. En posant $\rho = \log(m(\mathrm{T}))$, on obtient $m(f) = m_\infty(\mathrm{T})^\rho$. □

Exemple 1.15. — Soit F un corps muni d'une valeur absolue ultramétrique m. On étend m à l'anneau $\mathrm{F}[\mathrm{T}]$ en posant

$$m(f) = \sup(m(a_n)), \qquad f = \sum_n a_n \mathrm{T}^n.$$

On vérifie directement que l'application ainsi définie vérifie les axiomes (1.1/1) et (1.1/3) d'une valeur absolue ; elle vérifie même l'inégalité ultramétrique (1.7/4).

Moins évidente est la propriété (1.1/2), c'est-à-dire que l'on a $m(fg) = m(f)m(g)$ pour tous $f, g \in \mathrm{F}[\mathrm{T}]$. La propriété est évidente si $f = 0$ ou $g = 0$; supposons donc que $f \neq 0$ et $g \neq 0$. Soit $\mathrm{A} = \mathrm{F}^\circ$ l'anneau de valuation de F, soit P son idéal maximal $\mathrm{F}^{\circ\circ}$ et soit $k = \mathrm{A}/\mathrm{P}$ son corps résiduel. Posons $f = \sum a_n \mathrm{T}^n$, $g = \sum b_n \mathrm{T}^n$, $h = \sum c_n \mathrm{T}^n$; soit r un entier tel que $m(f) = a_r$ et soit s un entier tel que $m(g) = b_s$. on a $f = a_r(f/a_r)$, $g = b_s(g/b_s)$, et la définition entraîne

$$m(fg) = m(a_r)m(b_s)m((f/a_r)(g/b_s)),$$
$$m(f)m(g) = m(a_r)m(b_s)m(f/a_r)m(g/b_s).$$

Pour démontrer l'égalité voulue, on peut ainsi supposer que $m(f) = m(g) = 1$, et il s'agit de démontrer que $m(h) = 1$; l'hypothèse signifie que f et g appartiennent à $\mathrm{A}[\mathrm{T}]$ et que leurs images $\overline{f}$ et $\overline{g}$ dans $k[\mathrm{T}]$ sont non nulles. Comme l'anneau $k[\mathrm{T}]$ est intègre, le produit $\overline{f}\overline{g}$ n'est pas nul ; puisque ce produit est l'image du polynôme $fg \in \mathrm{A}[\mathrm{T}]$, on a $m(c_n) \leqslant 1$ pour tout n, et il existe un entier t tel que $m(c_t) = 1$. Ainsi, $m(h) = 1$.

Les lecteurices auront peut-être reconnu quelques cas particuliers de cet énoncé de multiplicativité : lorsque F est le corps des fractions d'un anneau factoriel R et que la valeur absolue m est associée à un élément irréductible, il est souvent connu comme le « lemme de Gauss » : il apparaît comme une étape cruciale. dans la démonstration du théorème de Gauss que l'anneau $\mathrm{R}[\mathrm{T}]$ est factoriel.

Par construction, $m(f) \neq 0$ pour $f \neq 0$; l'application m se prolonge alors de manière unique en une valeur absolue sur le corps $\mathrm{F}(\mathrm{T})$ des fractions rationnelles — on l'appelle la *valeur absolue de Gauss*.

2. Complétude, locale compacité. Normes

2.1. — Soit A un anneau commutatif et soit m une valeur absolue sur A. Le séparé–complété $\widehat{\mathrm{A}}$ de A pour la topologie définie par m est le quotient de l'ensemble des suites de Cauchy dans A par la relation d'équivalence pour laquelle $(a_n) \sim (b_n)$ si et seulement si $m(a_n - b_n) \to 0$. Les suites de Cauchy

forment un sous-anneau de l'ensemble des suites, et la relation d'équivalence en question est celle associée à l'idéal formé des suites qui tendent vers 0. Ainsi, $\widehat{\mathrm{A}}$ possède une structure d'anneau pour laquelle l'application canonique $\mathrm{A} \to \widehat{\mathrm{A}}$ est un morphisme d'anneau. Le noyau de ce morphisme est le noyau P de la valeur absolue m ; il est en particulier injectif lorsque A est un corps.

Comme l'application $m\colon \mathrm{A} \to \mathbf{R}_+$ est uniformément continue (et passe au quotient par l'idéal P), l'application canonique s'étend par densité en une unique application continue de $\widehat{\mathrm{A}}$ dans $\mathbf{R}_+$; c'est encore une valeur absolue. Concrètement si $\alpha \in \widehat{\mathrm{A}}$ est la classe d'une suite de Cauchy (a_n), alors la suite $m(a_n)$ converge et $m(\alpha)$ est sa limite. Les axiomes d'une valeur absolue se vérifient immédiatement.

Lemme 2.2. — *Soit* F *un corps et soit* m *une valeur absolue sur* F. *Alors* $\widehat{\mathrm{F}}$ *est un corps.*

Démonstration. — Soit α un élément non nul de $\widehat{\mathrm{F}}$ et soit (a_n) une suite de Cauchy qui représente α. Comme (a_n) ne tend pas vers 0 dans F, la suite $(m(a_n))$ ne tend pas vers 0 dans $\mathbf{R}_+$, si bien que $m(\alpha) > 0$. En particulier, il existe un nombre réel $r > 0$ tel que $m(a_n) \geqslant r$ pour tout entier n assez grand. Pour tout entier n, posons $b_n = 1/a_n$ si $a_n \neq 0$, et $b_n = 0$ sinon. Pour n, p assez grands, on a $b_n - b_p = (a_p - a_n)/a_p a_n$, ce qui entraîne que la suite (b_n) est de Cauchy ; notons β sa classe dans $\widehat{\mathrm{F}}$. La suite $(a_n b_n)$ est stationnaire de limite 1, on a donc $\alpha\beta = 1$. □

2.3. — Soit F un corps muni d'une valeur absolue m qui n'est pas ultramétrique ; on sait que F est de caractéristique zéro (remarque 1.11) et la restriction de m à son sous-corps premier $\mathbf{Q}$ est une valeur absolue sur $\mathbf{Q}$ qui n'est pas ultramétrique. D'après le théorème d'Ostrowski, il existe un nombre réel $\rho > 0$ tel que $m(a) = |a|^\rho$ pour tout $a \in \mathbf{Q}$. En particulier, la suite $(m(n))$ tend vers l'infini lorsque l'entier n tend vers l'infini. Ou encore, pour tout $x \in \mathrm{F}$ tel que $m(x) \neq 0$ et tout nombre réel M, il existe un entier n tel que $m(nx) \geqslant \mathrm{M}$. Cette propriété, reminiscente de la propriété d'Archimède pour le corps $\mathbf{R}$ des nombres réels, conduit à dire que m est une *valeur absolue archimédienne.*

De plus, la structure de $\mathbf{Q}$-algèbre sur le complété $\widehat{\mathrm{F}}$ de F se prolonge en une structure de $\mathbf{R}$-algèbre.

Si $\widehat{\mathrm{F}}$ contient un élément i tel que $i^2 = -1$, un théorème d'Ostrowski affirme que $\widehat{\mathrm{F}}$ est isomorphe à $\mathbf{C}$, et que sa valeur absolue est une puissance de la valeur absolue usuelle. Dans le cas contraire, on démontre que la valeur absolue m de $\widehat{\mathrm{F}}$ s'étend en une valeur absolue sur le corps $\widehat{\mathrm{F}}(i)$, lequel est donc isomorphe à $\mathbf{C}$ d'après le premier cas.

Cela montre que la théorie des corps munis d'une valeur absolue n'apporte de nouveauté par rapport à l'analyse classique que dans le cas ultramétrique.

2.4. — Soit A un anneau commutatif et soit m une valeur absolue m ultramétrique sur A. L'inégalité ultramétrique donne lieu à des des comportements topologiques/géométriques assez différents de ceux de l'analyse réelle.

(1) *Pour tous $a, b \in \mathrm{A}$ tel que $m(b) < m(a)$, on a $m(a+b) = m(a)$.*

Par l'inégalité ultramétrique, on a déjà $m(a+b) \leqslant \sup(m(a), m(b)) = m(a)$. En écrivant $a = (a+b)-b$, l'inégalité ultramétrique fournit également $m(a) \leqslant \sup(m(a+b), m(-b))$. Si $m(a+b) \geqslant m(-b)$, on obtient l'inégalité $m(a) \leqslant m(a+b)$; l'autre cas est absurde puisqu'il entraîne $m(a) \leqslant m(-b) = m(b)$.

(2) *Si deux boules se rencontrent, celle de plus petit rayon est contenue dans l'autre. En particulier, tout point d'une boule en est un centre.* Soit en effet $a, b \in \mathrm{A}$ et $r, s \in \mathbf{R}_+$, et soit c un point commun aux boules $\mathrm{B}(a, r)$ et $\mathrm{B}(b, s)$. Alors $m(c-a) \leqslant r$ et $m(c-b) \leqslant s$; si $r \leqslant s$, on a donc $m(a-b) = m((c-b)-(c-a)) \leqslant \sup(r, s) = s$, donc $a \in \mathrm{B}(b, s)$. Par suite, si $x \in \mathrm{B}(a, r)$, on a $m(x-b) = m((x-a)-(a-b)) \leqslant \sup(m(x-a), m(a-b)) \leqslant s$, si bien que $\mathrm{B}(a, r) \subseteq \mathrm{B}(b, s)$.

(3) *Les boules circonférenciées sont à la fois ouvertes et fermées.* La boule $\mathrm{B}(a, r)$ est ouverte, car si $x \in \mathrm{B}(a, r)$, on a $\mathrm{B}(x, r) \subseteq \mathrm{B}(a, r)$ (en fait égalité). Elle est fermée car si $x \notin \mathrm{B}(a, r)$, on a $\mathrm{B}(x, r) \cap \mathrm{B}(a, r) = \varnothing$ (sinon, on aurait $\mathrm{B}(x, r) \subseteq \mathrm{B}(a, r)$, contredisant l'hypothèse $x \notin \mathrm{B}(a, r)$).

Tout point de A possède une base de voisinages ouverts et fermés, à savoir les boules circonférenciées. Par suite, *la topologie de* A *est totalement discontinue.*

(4) *Une suite (a_n) est de Cauchy si et seulement si la suite des différences $(a_{n+1} - a_n)$ tend vers* 0. En effet, si cette dernière propriété est satisfaite,

l'égalité

$$a_{n+k} - a_n = \sum_{i=0}^{k-1} (a_{n+i+1} - a_{n+i})$$

entraîne

$$m(a_{n+k} - a_n) \leqslant \sup_{0 \leqslant i < k} m(a_{n+i+1} - a_{n+i}),$$

d'où l'on déduit que la suite (a_n) est de Cauchy.

Plus généralement, un monde ultramétrique est un monde dans lequel la politique des petits pas ne permet pas d'aller bien loin.

2.5. — Soit F un corps muni d'une valeur absolue ultramétrique m. L'ensemble A des éléments $a \in \mathrm{F}$ tels que $m(a) \leqslant 1$ est un sous-anneau de F ; par construction, pour tout élément non nul a de F, on a $a \in \mathrm{A}$ ou $1/a \in \mathrm{A}$: l'anneau A est ce qu'on appelle un *anneau de valuation* du corps F.

Un élément $a \in \mathrm{A}$ est inversible dans A si et seulement si $m(a) = 1$: en effet, supposant $a \neq 0$, il faut et il suffit que $1/a$ appartienne à A, c'est-à-dire $m(1/a) \leqslant 1$, d'où $m(a) \geqslant 1$.

Par suite, l'ensemble M des éléments $a \in \mathrm{A}$ tels que $m(a) < 1$ est un idéal de A et c'est le seul idéal maximal de l'anneau A.

Ainsi, et malgré la vocation analytique de l'introduction de la notion de valeur absolue, les valeurs absolues ultramétriques explorent un monde mathématique très proche de l'algèbre.

2.6. — La restriction à $\mathrm{F}^\times$ de l'application $x \mapsto -\log(m(x))$ est un morphisme de groupes de $\mathrm{F}^\times$ dans $\mathbf{R}$; son image Γ est donc un sous-groupe de $\mathbf{R}$ et il y a trois possibilités pour l'image de ce morphisme :

– Elle est triviale : c'est le cas où la valeur absolue m est la valeur absolue triviale, et l'anneau A coïncide avec le corps F ;

– C'est un sous-groupe isomorphe à $\mathbf{Z}$; il existe donc un élément $a \in \mathrm{A}$ tel que $-\log(m(a))$ est l'unique générateur strictement positif de Γ. Cet élément a est un générateur de l'idéal M, qui est donc principal. On peut en déduire que A est un anneau principal (si I est un idéal non nul de A, il existe un plus grand entier n tel que $-\log(m(x)) \geqslant -n\log(m(a))$ pour tout $x \in \mathrm{A}$, et l'on a $\mathrm{I} = \langle a^n \rangle$) qui n'est pas un corps (car a n'est pas inversible dans A) ;

– C'est un sous-groupe dense de $\mathbf{R}$. Dans ce cas, l'idéal M n'est pas principal, il n'est même pas de type fini[1], de sorte que l'anneau A n'est pas noethérien.

On dit la *valuation de* F *est discrète* si le groupe Γ est isomorphe à $\mathbf{Z}$. La terminologie est traditionnelle, mais malheureuse, car elle n'a rien à voir avec la notion topologique (qui correspond au cas de la valeur absolue triviale).

Proposition 2.7. — *Soit* F *un corps muni d'une valeur absolue ultramétrique non triviale. Les propriétés suivantes sont équivalentes :*

(1) *La topologie de* F *est localement compacte;*

(2) *La topologie de l'anneau de valuation* A *est compacte;*

(3) *La topologie de* F *est complète, sa valuation est discrète et son corps résiduel* A/M *est fini.*

Si elles sont satisfaites, A *est homéomorphe à un ensemble de Cantor.*

Démonstration. — (1)⇒(2). Supposons F localement compact. Il existe alors un nombre réel $r > 0$ tel que $\mathrm{B}(0, r)$ est compact. La fonction $a \mapsto m(a)$ sur $\mathrm{B}(0, r)$ est continue, elle est donc majorée et atteint sa borne supérieure; cela ramène au cas où il existe $a \in \mathrm{A}$ tel que $r = m(a)$. L'application $x \mapsto x/a$ de F dans lui-même est un homéomorphisme et l'image de $\mathrm{B}(0, r)$ est la boule $\mathrm{B}(0, 1)$. Ainsi, A est compact.

(2)⇒(3). Supposons maintenant A compact. Tout d'abord, l'espace métrique F est localement compact, donc complet.

Soit $p\colon \mathrm{A} \to \mathrm{A}/\mathrm{M}$ l'application de réduction; pour $a, b \in \mathrm{A}$, les conditions $p(a) = p(b)$ et $m(a - b) < 1$ sont donc équivalentes. Pour tout $u \in \mathrm{A}/\mathrm{M}$, choisissons un élément $s(u) \in \mathrm{A}$ tel que $p(s(u)) = u$; alors A est la réunion des boules non circonférenciées $\mathrm{B}^\circ(s(u), 1)$, pour $u \in \mathrm{A}/\mathrm{M}$; elles sont ouvertes et deux à deux disjointes; comme A est compact, elles sont en nombre fini, ce qui prouve que A/M est fini.

La réunion des $\mathrm{B}^\circ(s(u), 1)$, pour $u \in \mathrm{A}/\mathrm{M}$ et $u \neq 0$ est une partie ouverte de A; c'est le complémentaire de $\mathrm{B}^\circ(s(0), 1) = \mathrm{B}^\circ(0, 1) = \mathrm{M}$ qui est donc une partie fermée; par suite, la fonction continue $a \mapsto m(a)$ sur $\mathrm{B}^\circ(0, 1)$ atteint sa borne supérieure, ce qui prouve qu'il existe $a \in \mathrm{M}$ tel que $m(x) \leqslant m(a)$ pour tout $x \in \mathrm{M}$. Observons que $a \neq 0$ (sinon, $\mathrm{M} = 0$, donc A est un corps, donc

[1] Dans un anneau de valuation comme l'anneau A, tout idéal de type fini est principal…

$A = F$ et la valeur absolue de F est triviale) ; pour $x \in M$, on a $m(x/a) \leqslant 1$, donc $x/a \in M$ et $x \in \langle a \rangle$; ainsi, l'idéal M est principal et la valuation de F est discrète.

(3)⇒(1). Fixons un générateur a de l'idéal M et une application $u \mapsto s(u)$ de A/M dans A tel que $p(s(u)) = u$ pour tout $u \in A/M$. Pour toute suite (u_n) de $(A/M)^{\mathbf{N}}$, la série $\sum s(u_n)a^n$ dans A converge, car son terme général tend vers 0, puisque F est complet. L'application $\varphi\colon (A/M)^{\mathbf{N}} \to A$ telle que $\varphi(u) = \sum_n s(u_n)a^n$ est continue, lorsque $(A/M)^{\mathbf{N}}$ est muni de la topologie produit des topologies discrètes. En effet, si les n premiers termes de u et v coïncident, alors $\varphi(u) - \varphi(v)$ est multiple de a^n, donc $m(\varphi(u) - \varphi(v)) \leqslant m(a)^n$. Si, de plus, si $u_n \neq v_n$, alors $m(\varphi(u) - \varphi(v)) = m(a)^n$, ce qui prouve que φ est injective.

Démontrons que φ est surjective. Soit $x \in A$; posons $x_0 = x$ et $u_0 = p(x)$; comme $x - s(u_0) \in M$, il existe un unique élément $x_1 \in A$ tel que $x - s(u_0) = ax_1$. Par récurrence, on construit des suites (u_n) dans A/M et (x_n) dans A telles que $x = s(u_0) + as(u_1) + \cdots + a^n s(u_n) + a^{n+1}x_{n+1}$ pour tout entier n. On constate que $x = \varphi(u)$, donc φ est surjective.

L'application $\varphi\colon (A/M)^{\mathbf{N}} \to A$ est bijective, continue ; c'est donc un homéomorphisme car $(A/M)^{\mathbf{N}}$ est compact et A est séparé. Lorsque A/M est de cardinal 2 (ce qui signifie que $F = \mathbf{Q}_2$ ou $F = \mathbf{F}_2((T))$), la caractérisation des points de l'ensemble triadique de Cantor K_3 par leur développement triadique exhibe un homéomorphisme de cet ensemble avec $\{0, 1\}^{\mathbf{N}}$. En général, l'espace topologique produit $(A/M)^{\mathbf{N}}$, produit infini d'un espace topologique discret fini (avec au moins deux éléments), est infini, métrisable, compact, séparable et totalement discontinu : un résultat classique de topologie générale entraîne qu'il est homéomorphe à K_3. □

Remarques 2.8. — (1) De la démonstration de (3)⇒(1), on peut extraire l'énoncé : si F est complet, sa valuation non triviale et discrète, alors l'anneau de valuation A est homéomorphe à l'espace « prodiscret » $(A/M)^{\mathbf{N}}$.

(2) Lorsque la valeur absolue de F est triviale, la topologie de F est discrète, donc localement compacte puisque pour tout $a \in F$, le singleton $\{a\}$ est un voisinage ouvert compacts de a. En revanche, comme l'anneau A coïncide avec F, il n'est compact que si F est fini.

Exemple 2.9. — Soit k un corps. Munissons l'anneau $\mathrm{A} = k[\mathrm{T}]$ de la valeur absolue m_{T} pour laquelle $m(\mathrm{T}) = 1/e$ et $m(a) = 1$ pour tout $a \in k^\times$. Le complété $\widehat{\mathrm{A}}$ de A s'identifie naturellement à l'anneau $k[[\mathrm{T}]]$ des séries formelles en T.

Exemple 2.10. — On appelle *corps des nombres p-adiques* le complété $\mathbf{Q}_p$ du corps des nombres rationnels pour la valeur absolue p-adique m_p (normalisée par $m_p(p) = 1/p$). Son anneau de valuation est noté $\mathbf{Z}_p$; c'est l'ensemble des entiers p-adiques.

La construction du séparé–complété entraîne que $\mathbf{Z}_p$ est l'adhérence dans $\mathbf{Q}_p$ de l'anneau $\mathbf{Z}$, que son idéal maximal est l'adhérence de l'idéal $\langle p \rangle$, et que le corps résiduel de $\mathbf{Z}_p$ s'identifie au corps fini $\mathbf{Z}/p\mathbf{Z}$. La démonstration de la proposition précédente fournit ainsi un homéomorphisme de $(\mathbf{Z}/p\mathbf{Z})^{\mathbf{N}}$ sur $\mathbf{Z}_p$. Cet homéomorphisme dépend du choix d'éléments $s(u)$ de $\mathbf{Z}_p$ pour $u \in \mathbf{Z}/p\mathbf{Z}$. S'il existe un choix meilleur que d'autres, multiplicatif (voir l'exemple 3.4), il est naturel de choisir $s(u)$ dans $\{0, \dots, p-1\}$. C'est d'ailleurs ce que fait Hensel (1904) lorsqu'il introduit les nombres p-adiques, comme des sortes de nombres entiers donnés par un développement illimité en base p. En revanche, si nous serions tentés d'écrire $\dots s_n s_{n-1} \dots s_0$ un tel nombre, Hensel l'écrit $s_0 s_1 s_2 \dots$! Dans cet article, Hensel ne donne que peu d'indication sur l'origine de ses idées, mais on peut en deviner une trace dans son article (Hensel, 1901) où il élabore explicitement une analogie entre le développement en base p des nombres entiers (et même des nombres algébriques) et le développement en série de puissances d'une fonction algébrique.

On peut préférer une présentation plus algébrique des nombres p-adiques. Pour tout entier n, l'application de $\mathbf{Z}$ dans $\mathbf{Z}_p$ induit un isomorphisme d'anneaux de $\mathbf{Z}/p^n\mathbf{Z}$ sur $\mathbf{Z}_p/p^n\mathbf{Z}_p$. Les inverses de ces isomorphismes fournissent un homomorphisme d'anneaux

$$\mathbf{Z}_p \to \prod_n \mathbf{Z}/p^n\mathbf{Z}$$

qui est injectif, et dont l'image (la limite projective de ce système d'anneaux, notée $\varprojlim_n \mathbf{Z}/p^n\mathbf{Z}$), est le sous-anneau de $\prod_n \mathbf{Z}/p^n\mathbf{Z}$ formé des suites (a_n) telles que $a_{n+1} \equiv a_n \pmod{p^n}$. De plus, cet homomorphisme induit un

homéomorphisme de $\mathbf{Z}_p$ sur son image lorsque $\prod_n \mathbf{Z}/p^n\mathbf{Z}$ est muni de la topologie produit des topologies discrètes.

Définition 2.11. — *Soit* F *un corps muni d'une valeur absolue* $|\cdot|$ *et soit* V *un* F*-espace vectoriel. Une fonction* $\|\cdot\| : \mathrm{V} \to \mathbf{R}_+$ *est appelée* norme *si*

(1) *Pour* $v \in \mathrm{V}$, *on a* $\|v\| = 0$ *si et seulement si* $v = 0$*;*
(2) *On a* $\|av\| = |a|\,\|v\|$ *pour tout* $a \in \mathrm{F}$ *et tout* $v \in \mathrm{F}$*;*
(3) *On a* $\|v + w\| \leqslant \|v\| + \|w\|$ *pour tous* $v, w \in \mathrm{V}$.

2.12. — Une norme sur V définit une distance sur V, donc une topologie pour laquelle l'addition de V et la multiplication sont continues.

D'après la proposition suivante, deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sur V définissent la même topologie sur V si et seulement si elles sont équivalentes, c'est-à-dire s'il existe des nombres réels $c, c' > 0$ tels que $c' \|v\|_2 \leqslant \|v\|_1 \leqslant c \|v\|_2$ pour tout $v \in \mathrm{V}$.

Proposition 2.13. — *Soit* F *un corps muni d'une valeur absolue non triviale. soit* V *et* W *des* F*-espaces vectoriels normés et soit* $f : \mathrm{V} \to \mathrm{W}$ *une application linéaire. Les conditions suivantes sont équivalentes :*

(1) *L'application* f *est continue;*
(2) *L'application* f *est continue en* 0*;*
(3) *Il existe un nombre réel* c *tel que* $\|f(v)\| \leqslant c\,\|v\|$ *pour tout* $v \in \mathrm{V}$.

Démonstration. — L'implication (1)⇒(2) est évidente, et l'hypothèse (3) entraîne que f est c-lipschitzienne, donc continue.

Supposons (2) et démontrons (3); c'est là que nous devrons utiliser l'hypothèse que la valeur absolue de F n'est pas triviale.

La continuité de f entraîne qu'il existe un nombre réel $r > 0$ tel que si $\|v\| < r$, alors $\|f(v)\| \leqslant 1$. Puisque la valeur absolue de F n'est pas triviale, il existe un élément $a \in \mathrm{F}$ tel que $0 < |a| < 1$; nous allons démontrer que $\|f(v)\| \leqslant (1/|a|r)\,\|v\|$. En effet, soit $v \in \mathrm{V}$; l'inégalité est évidente si $v = 0$; supposons donc $v \neq 0$ et soit n le plus petit entier relatif tel que $\|a^n v\| < r$. On a donc $|a|^n\,\|f(v)\| = \|f(a^n v)\| \leqslant 1$; d'autre part, la définition de n entraîne que $\|a^{n-1}v\| \geqslant r$, donc $1 \leqslant (|a|^{n-1}/r)\,\|v\|$. Par suite, $\|f(v)\| \leqslant (1/|a|r)\,\|v\|$, comme il fallait démontrer. □

Théorème 2.14. — *Soit* F *un corps muni d'une valeur absolue pour laquelle il est complet.*

(1) *Soit* V, V′ *des* F*-espaces vectoriels normés et soit* $f\colon \mathrm{V} \to \mathrm{V}'$ *une application linéaire. Si* V *est de dimension finie, alors* f *est lipschitzienne, et en particulier continue.*

(2) *Soit* V *un* F*-espace vectoriel de dimension finie. Toutes les normes sur* V *sont équivalentes; la topologie qu'elles induisent sur* V *est complète.*

(3) *Soit* V *un* F*-espace vectoriel normé et soit* W *un sous-espace vectoriel de* V. *Si* W *est de dimension finie, alors* W *est fermé.*

Démonstration. — On raisonne le théorème par récurrence sur la dimension de V. Soit $(e_1, \dots, e_n)$ une base de V.

(1) Si f n'est pas continue, il existe pour tout $k \geqslant 1$ un vecteur non nul $v^{(k)} \in \mathrm{V}$ tel que $\|v^{(k)}\| < k^{-1}\|f(v^{(k)})\|$. Écrivons $v^{(k)} = a_1e_1 + \cdots + a_ne_n$, où $a_1, \dots, a_n \in \mathrm{F}$; par homogénéité, on peut supposer que l'une des coordonnées a_m est égale à 1 et toutes les autres sont de valeur absolue $\leqslant 1$. Puis, quitte à considérer une sous-suite et à réordonner la base $(e_1, \dots, e_n)$, on peut supposer que cette coordonnée est celle d'indice n; posons $\mathrm{W} = \langle e_1, \dots, e_{n-1}\rangle$. Par hypothèse de récurrence, l'espace W est fermé dans V.

Observons que $\|f(v^{(k)})\| \leqslant \sum_{j=1}^n \|f(e_j)\|$ pour tout k, si bien que $\|v^{(k)}\| \to 0$. Cela entraîne que la suite $(v^{(k)} - e_n)$ tend vers $-e_n$; comme elle est contenue dans W, cela prouve que $-e_n \in \mathrm{W}$, ce qui est absurde.

(2) En appliquant (1) à l'application identique de V pour deux normes arbitraires sur V, on obtient une des deux inégalités requises par l'équivalence des normes, et l'autre s'en déduit par symétrie.

Par ailleurs, le choix de la base $(e_1, \dots, e_n)$ induit une norme $\|\cdot\|$ sur V telle que $\|\sum a_ie_i\| = \sup(|a_1|, \dots, |a_n|)$ pour tout $(a_1, \dots, a_n) \in \mathrm{F}^n$. Il est facile de voir que cette norme rend V complet.

(3) La norme de V définit une norme sur W qui le rend complet, d'après (2). Comme un sous-espace complet d'un espace métrique est fermé, cela entraîne que W est fermé. □

3. Méthode de Newton

Soit F un corps muni d'une valeur absolue ultramétrique et non triviale pour laquelle il est complet. On s'intéresse dans ce chapitre aux équations polynomiales à une ou plusieurs variables à coefficients dans F.

3.1. — Isaac Newton, dans sa *Méthode des fluxions* (1671, mais seulement publiée en 1736) puis dans une lettre du 13 juin 1676 à son ami Henry Oldenburg, diplomate et homme de sciences, a introduit la méthode pour résoudre les équations polynomiales qui porte aujourd'hui son nom. Le contexte est la résolution d'équations dépendant d'un paramètre — ce qu'il appelle « équations affectées ». Citons cette lettre (Turnbull, 1960) :

> Les extractions de racines affectées, d'équations à plusieurs termes littéraux, ressemblent par la forme à leurs extractions en nombres, mais la méthode de Viète et de notre compatriote Oughtred est moins adaptée à cet usage. C'est pourquoi j'ai été amené à en concevoir une autre…

Partant d'une solution « approchée » a d'une équation $f(x) = 0$, Newton récrit l'équation sous la forme $f(a + y) = 0$ en l'inconnue y et la développe sous la forme $g(y) = 0$; si $g(y) = f(a) + f_1(a)y + f_2(a)y^2 + \dots$, il considère que $b = -f(a)/f_1(a)$ est une solution approchée de cette dernière équation, puis il itère le procédé.

La présentation moderne de cette méthode, telle qu'elle a pris forme aux 18e et 19e siècle (avec notamment Lagrange et Fourier), consiste à itérer l'application

$$a \mapsto a - \frac{f(a)}{f'(a)}.$$

Par rapport au texte de Newton, elle combine deux modifications importantes, tirées des travaux de Joseph Raphson et Thomas Simpson. Tout d'abord, la suite (a_n) ainsi construite est vue comme une suite d'approximations d'une solution de l'équation initiale, alors que Newton considérait plutôt la suite $(a_1, a_2 - a_1, \dots)$, chaque terme étant vu comme solution approchée d'une équation différente. D'autre part, le dénominateur $f'(a)$ est interprété comme une différentielle (une « fluxion » !) et non comme le terme $f_1(a)$ dans le développement $f(a+x) = f(a) + f_1(a)x + \dots$.

Je renvoie aux deux articles de Kollerstrom (1992) et Christensen (1996) pour plus de détails sur cette histoire.

Il ne s'agira pas ici de développer le succès de cette méthode en analyse numérique, en théorie des systèmes dynamiques (de la méthode KAM aux plus médiatiques ensembles de Julia et Mandelbrot), mais de montrer son impact dans le contexte des corps munis d'une valeur absolue ultramétrique.

On commence par expliquer la méthode de Newton en dimension 1.

Proposition 3.2. — *Soit $\varphi \in \mathrm{F}[\mathrm{T}]$ un polynôme dont les coefficients appartiennent à l'anneau de valuation* A *de* F. *Soit $a \in \mathrm{A}$ tel que $|\varphi(a)| < |\varphi'(a)|^2$; la méthode de Newton pour φ issue du point a converge vers un élément $\alpha \in \mathrm{A}$ qui est l'unique élément de* A *tel que $\varphi(\alpha) = 0$ et $|\alpha - a| < |\varphi'(a)|$; plus précisément, on a $|\alpha - a| = |\varphi(a)/\varphi'(a)|$. De plus, la convergence est* quadratique : *on a $|a_n - \alpha| \leqslant |\varphi'(a)| \cdot (|\varphi(a)/\varphi'(a)^2|)^{2^n}$ pour tout entier $n \geqslant 0$.*

Démonstration. — Posons $c = |\varphi(a)/\varphi'(a)|$; les hypothèses entraînent donc $c < |\varphi'(a)| \leqslant 1$. On écrit alors la formule de Taylor à l'ordre 1,

$$\varphi(a + \mathrm{T}) = \varphi(a) + \varphi'(a)\mathrm{T} + \mathrm{T}^2\psi(\mathrm{T}),$$

où $\psi \in \mathrm{A}[\mathrm{T}]$. Soit $b = -\varphi(a)/\varphi'(a)$; on a $|b| = c < 1$ par hypothèse, d'où $|\varphi(a+b)| = c^2|\psi(b)|$. Puisque ψ est à coefficients dans A, on a donc $|\psi(b)| \leqslant 1$, d'où $|\varphi(a + b)| \leqslant c^2$. Par ailleurs, en écrivant

$$\varphi'(a + \mathrm{T}) = \varphi'(a) + \mathrm{T}\lambda(\mathrm{T}),$$

où $\lambda \in \mathrm{A}[\mathrm{T}]$, on constate que

$$|\varphi'(a + b) - \varphi'(a)| \leqslant |b| = c < |\varphi'(a)|,$$

si bien que $|\varphi'(a + b)| = |\varphi'(a)|$. Posons $a_1 = a + b$. On a donc $a_1 \in \mathrm{A}$, $|a_1 - a| = c$, $|\varphi'(a_1)| = |\varphi'(a)|$, et $|\varphi(a_1)| \leqslant c^2$. En particulier, on a

$$|\varphi(a_1)/\varphi'(a_1)| \leqslant c^2/|\varphi'(a_1)| = c^2/|\varphi'(a)| < c < |\varphi'(a)| = |\varphi'(a_1)|$$

et

$$|\varphi(a_1)/\varphi'(a_1)^2| \leqslant c^2/|\varphi'(a)|^2 = |\varphi(a)/\varphi'(a)^2|^2.$$

Cela entraîne qu'il existe une unique suite (a_n) dans A telle que $a_0 = a$, $|\varphi(a_n)| < |\varphi'(a_n)|^2$ et $a_{n+1} = a_n - \varphi(a_n)/\varphi'(a_n)$ pour tout n. Plus précisément, on a $|\varphi'(a_n)| = |\varphi'(a)|$,

$$|a_{n+1} - a_n| \leqslant |\varphi(a_n)|/|\varphi'(a_n)| \leqslant c,$$

et

$$|\varphi(a_{n+1})| \leqslant |\varphi(a_n)/\varphi'(a)|^2$$

pour tout $n \in \mathbf{N}$.

Les inégalités

$$|\varphi(a_{n+1})/\varphi'(a_{n+1})^2| \leqslant (|\varphi(a_n)/\varphi'(a_n)^2|)^2$$

entraînent la majoration

$$|\varphi(a_n)/\varphi'(a_n)^2| \leqslant |\varphi(a)/\varphi'(a)^2|^{2^n},$$

pour tout $n \in \mathbf{N}$. Posons $h = |\varphi(a)/\varphi'(a)^2|$, de sorte que $|a_{n+1} - a_n| \leqslant |\varphi'(a)|h^{2^n}$ pour tout entier n. En particulier, la suite $(a_{n+1} - a_n)$ tend vers 0 ; comme F est complet, la suite (a_n) converge ; sa limite est un élément α de A tel que $|\varphi(\alpha)| = 0$ et $|\alpha - a| \leqslant c$. Comme $|\varphi'(a_n)| = |\varphi(a)|$ pour tout n, on a aussi $|\varphi'(\alpha)| = |\varphi(a)|$. Enfin,

$$|\alpha - a_n| \leqslant \sup_{m \geqslant n} |a_{m+1} - a_m| \leqslant |\varphi'(a)|h^{2^n}.$$

Soit b un élément de A tel que $|b - a| < |\varphi'(a)|$ et $\varphi(b) = 0$. En écrivant $\varphi(b) = \varphi(a) + (b - a)\varphi'(a) + (b - a)^2\psi(b - a)$, il vient $(b - a)(\varphi'(a) + (b - a)\psi(b - a)) = -\varphi(a)$. On a $|\psi(b - a)| \leqslant 1$, car $\psi \in \mathrm{A}[\mathrm{T}]$ et $b - a \in \mathrm{A}$; comme $|b - a| < |\varphi'(a)|$, on a donc $|\varphi'(a) + (b - a)\psi(b - a)| = \varphi'(a)$, d'où l'égalité $|b - a| = |\varphi(a)/\varphi'(a)|$.

En particulier, $|\alpha - a| = |\varphi'(a)|$.

En développant la relation $\varphi(\alpha + \mathrm{T}) = \varphi'(\alpha)\mathrm{T} + \mathrm{T}^2\theta(\mathrm{T})$, où $\theta \in \mathrm{A}[\mathrm{T}]$, on obtient l'égalité $(b - \alpha)(\varphi'(\alpha) + (b - \alpha)\theta(b - \alpha)) = 0$. Puisque $|b - \alpha| = |(b-a)-(b-a)| < |\varphi'(a)|$ et $|\theta(b-\alpha)| \leqslant 1$, on a $\varphi'(\alpha)+(b-\alpha)\theta(b-\alpha) \neq 0$. Par suite, $b - \alpha = 0$ et $b = \alpha$. □

Corollaire 3.3. — *Soit $\varphi \in \mathrm{A}[\mathrm{T}]$ et soit $a \in \mathrm{A}$ tel que $|\varphi(a)| < 1$ et $|\varphi'(a)| = 1$. Si* F *est complet, il existe un unique élément $\alpha \in \mathrm{A}$ tel que $|a-\alpha| < 1$ et $\varphi(\alpha) = 0$. On a de plus $|a - \alpha| = |\varphi(\alpha)|$.*

Observons que l'hypothèse signifie que lorsqu'on considère le polynôme $f \in k[\mathrm{T}]$ déduit de φ par réduction modulo l'idéal maximal de A, l'image $\overline{a}$ de a est racine de f mais pas de f'.

Exemple 3.4. — Prenons par exemple un polynôme de la forme $\varphi = \mathrm{T}^d - a$, de sorte que $\varphi' = d\mathrm{T}^{d-1}$. Supposons $a \notin \mathrm{M}$, de sorte que sa réduction $\overline{a}$ n'est pas nulle. Soit $c \in k$ tel que $c^d = \overline{a}$. D'après la proposition, une condition suffisante pour qu'il existe un élément $\alpha \in \mathrm{A}$ tel que tel que $\alpha^d = a$ et $\overline{\alpha} = c$ est qu'il existe $\alpha \in \mathrm{A}$ tel que $\overline{\alpha} = c$ et $|\alpha^d - a| < |d|^2$.

a) Si $d \cdot 1_{\mathrm{A}} \notin \mathrm{M}$, c'est-à-dire si d n'est pas multiple de la caractéristique de k, on a $|d| = 1$, donc tout élément $\alpha \in \mathrm{A}$ tel que $\overline{\alpha} = c$ vérifie le critère.

Pour un exemple concret, on peut prendre $\varphi = \mathrm{T}^2 + 1$ dans le corps $\mathbf{Q}_p$, où p est un nombre premier tel que $p \equiv 1 \pmod 4$: le polynôme φ est alors scindé dans $\mathbf{Q}_p$.

Observons que le polynôme $\mathrm{T}^{p-1} - 1$ est scindé sur $\mathbf{F}_p$. Il existe ainsi pour tout élément $m \in \mathbf{F}_p^\times$ un unique élément $\theta_m \in \mathbf{Z}_p$ tel que $\overline{\theta_m} = m$ et $\theta_m^{p-1} = 1$. Ces éléments s'appellent parfois les représentants de Teichmüller de $\mathbf{F}_p^\times$. Ils forment un sous-groupe cyclique d'ordre $p - 1$.

b) Si $d \cdot 1_{\mathrm{A}} \in \mathrm{M}$, il peut ne pas exister d'élément $\alpha \in \mathrm{A}$ tel que $\alpha^d = a$ et $\overline{\alpha} = c$. C'est par exemple le cas pour l'équation $\mathrm{T}^2 + 1$ sur $\mathbf{Q}_2$, avec $c = 1$. En effet, si $\alpha^2 + 1 = 0$, on a $(\alpha - 1)^2 = \alpha^2 - 2\alpha + 1 = -2\alpha$; alors $v_2(\alpha) = 0$ et $2v_2(\alpha - 1) = 1$, ce qui est absurde.

En revanche, si $|a - 1| < |d|^2$, on peut prendre $\alpha = 1$ dans le critère, de sorte que a est une puissance d-ième.

3.5. — La méthode de Newton fonctionne également à plusieurs variables. Munissons F^n de la norme donnée par $\|(a_1, \dots, a_n)\| = \sup(|a_1|, \dots, |a_n|)$. Soit $\varphi_1, \dots, \varphi_n \in \mathrm{F}[\mathrm{T}]$ des polynômes à coefficients dans F en n indéterminées $\mathrm{T} = (\mathrm{T}_1, \dots, \mathrm{T}_n)$ et soit $\Phi = (\varphi_1, \dots, \varphi_n) \colon \mathrm{F}^n \to \mathrm{F}^n$ l'application polynomiale qu'on en déduit. Sa différentielle $\mathrm{D}\Phi$ est une application de F^n dans l'espace des matrices de type $n \times n$ dont les coefficients sont donnés par les dérivées partielles formelles $\partial\varphi_i / \partial\mathrm{T}_j$ des polynômes φ_i. On définit également le jacobien de Φ comme l'application polynomiale $\mathrm{J}\Phi \colon a \mapsto \det(\mathrm{D}\Phi(a))$.

La méthode de Newton pour Φ est l'itération de l'application rationnelle de F^n dans lui-même donnée par

$$a \mapsto \Phi(a) - (\mathrm{D}\Phi(a))^{-1}\Phi(a),$$

définie pourvu que $\mathrm{J}\Phi(a) \neq 0$.

Théorème 3.6. — *Soit* F *un corps muni d'une valeur absolue ultramétrique pour laquelle il est complet.*

Soit $\Phi\colon \mathrm{F}^n \to \mathrm{F}^n$ *une application polynomiale définie par des polynômes à coefficients dans l'anneau de valuation* A *de* F*;*

Soit $a \in \mathrm{A}^n$ *tel que* $\|\Phi(a)\| < |\mathrm{J}\Phi(a)|^2$*, il existe un unique point* $\alpha \in \mathrm{F}^n$ *tel que* $\|\alpha - a\| \leqslant \|\Phi(a)\| / |\mathrm{J}\Phi(a)|$ *et* $\Phi(\alpha) = 0$*. Il est obtenu par la méthode de Newton issue de* a *et vérifie* $\|a - \alpha\| \leqslant \|\Phi(a)\| / |\mathrm{J}\Phi(a)|$.

Démonstration. — La preuve est essentiellement la même qu'en dimension 1 ; la formule de Taylor est le seul point qui demande quelques arguments supplémentaires. A priori, on peut écrire

$$\Phi(a + \mathrm{T}) = \Phi(a) + \mathrm{D}\Phi(a) \cdot \mathrm{T} + \Psi(\mathrm{T}),$$

où Ψ est donné par des polynômes à coefficients dans A et dont chaque monôme est de degré au moins 2. Par suite, on a $\|\Psi(b)\| \leqslant \|b\|^2$ pour tout $b \in \mathrm{F}^n$. De même, la formule de la comatrice entraîne que $\mathrm{J}\Phi(\mathrm{T}) \cdot \mathrm{D}\Phi(\mathrm{T})^{-1}$ est une matrice dont les entrées sont des polynômes à coefficients dans A ; en particulier,

$$\|\mathrm{D}\Phi(a)^{-1} \cdot b\| \leqslant |\mathrm{J}\Phi(a)|^{-1} \|b\|$$

pour tout $b \in \mathrm{F}^n$. Le reste de la démonstration est laissé aux lecteurs et lectrices. □

3.7. — Avant d'énoncer le corollaire suivant, faisons quelques rappels sur le résultant de deux polynômes. Soit $\mathrm{P}, \mathrm{Q} \in \mathrm{A}[\mathrm{T}]$ deux polynômes à coefficients dans un anneau commutatif A et soit p, q deux entiers naturels tels que $\deg(\mathrm{P}) \leqslant p$ et $\deg(\mathrm{Q}) \leqslant q$. Le résultant $\mathrm{Res}_{p,q}(\mathrm{P}, \mathrm{Q})$ est alors défini comme le déterminant d'une matrice (« de Sylvester ») de taille $p + q$ consistant à écrire, sur q lignes, les coefficients de P, décalés d'une colonne d'une ligne sur la suivante, puis, sur p lignes, les coefficients de Q, décalés d'une colonne d'une ligne sur la suivante. De manière plus conceptuelle, et surtout plus éclairante, c'est le déterminant de l'application linéaire

$$\mathrm{A}[\mathrm{T}]_{<q} \times \mathrm{A}[\mathrm{T}]_{<p} \to \mathrm{A}[\mathrm{T}]_{<p+q}, \quad (\mathrm{U}, \mathrm{V}) \mapsto \mathrm{UP} + \mathrm{VQ},$$

où l'espace vectoriel source est muni de la base $(1, \mathrm{T}, \ldots, \mathrm{T}^{q-1}, 1, \ldots, \mathrm{T}^{p-1})$ et l'espace vectoriel but de la base $(1, \mathrm{T}, \ldots, \mathrm{T}^{p+q-1})$.

Supposons que l'anneau A soit un corps F. Alors, le résultant s'annule si et seulement si cette application linéaire n'est pas injective, et l'on vérifie

que cela se produit si et seulement si l'une des deux conditions suivantes est vérifiée :

(1) Les polynômes P et Q ne sont pas premiers entre eux ;
(2) On a $\deg(P) < p$ et $\deg(Q) < q$.

Posons en effet $D = \gcd(P, Q)$ et écrivons $P = DP_1$ et $Q = DQ_1$; on a la relation $Q_1P - P_1Q = 0$ qui prouve que l'application linéaire précédente n'est pas injective lorsque $\deg(Q_1) < q$ et $\deg(P_1) < q$. Cela se produit si $\deg(D) > 0$ ou si l'on a simultanément $\deg(Q) < q$ et $\deg(Q) < p$. Inversement, supposons que les polynômes P, Q sont premiers entre eux et que l'application linéaire considérée n'est pas injective ; considérons un élément non nul (U, V) de son noyau ; la relation $UP + VQ = 0$ et la coprimalité de P et Q entraîne que Q divise U et que P divise V. On en déduit $\deg(Q) < q$ et $\deg(P) < p$.

Par construction, le résultant $\mathrm{Res}_{p,q}(P, Q)$ est une application polynomiale en les coefficients de P et Q ; en particulier, si $\varphi\colon A \to B$ est un morphisme d'anneaux commutatifs, et si P^φ et Q^φ désignent les polynômes de $B[T]$ obtenus en appliquant φ aux coefficients de P et Q, on a

$$\mathrm{Res}_{p,q}(P^\varphi, Q^\varphi) = \varphi(\mathrm{Res}_{p,q}(P, Q)).$$

Observons que l'on peut fort bien avoir $\deg(P^\varphi) < \deg(P)$ ou $\deg(Q^\varphi) < \deg(Q)$, si bien que la généralité consistant à ne pas fixer $p = \deg(P)$ et $q = \deg(Q)$ est ici nécessaire.

Corollaire 3.8. — *Soit $\varphi \in A[T]$ un polynôme de degré n et soit $\psi, \eta \in A[T]$ deux polynômes tels que $\deg(\varphi) = \deg(\psi) + \deg(\eta)$ et $\|\varphi - \psi\eta\| < |\mathrm{Res}(\psi, \eta)|^2$, où $\mathrm{Res}(\psi, \eta)$ désigne le résultant des polynômes φ et ψ (relativement à leurs degrés). Il existe alors un unique couple (γ^*, η^*) dans $A[T]$ tels que $\varphi = \gamma^*\eta^*$, $\deg(\psi^* - \psi) < \deg(\psi)$, $\|\psi^* - \psi\| < |\mathrm{Res}(\psi, \eta)|$, $\deg(\eta^* - \eta) < \deg(\eta)$ et $\|\eta^* - \eta\| < |\mathrm{Res}(\psi, \eta)|$.*

Démonstration. — Posons $p = \deg(\psi)$, $q = \deg(\eta)$ et $n = p + q = \deg(\varphi)$. On identifie l'espace affine U des polynômes g tels que $\deg(g - \psi) < \deg(\psi)$ à l'espace vectoriel à l'espace vectoriel F^p, celui V des polynômes h tels que $\deg(h - \eta) < \deg(\eta)$ à l'espace vectoriel F^q, et celui W des polynômes f tels que $\deg(f - \varphi) < \deg(\varphi)$ à l'espace vectoriel F^n. Considérons l'application Φ

de $\mathrm{U} \times \mathrm{V}$ dans W donnée par $(g, h) \mapsto gh$. Elle est polynomiale. Sa différentielle en un couple (g, h) est l'application linéaire $(u, v) \mapsto uh + v\eta$ de $\mathrm{F}[\mathrm{T}]_{<q} \times \mathrm{F}[\mathrm{T}]_{<p}$ dans $\mathrm{F}[\mathrm{T}]_{<n}$; par suite, son jacobien $\mathrm{J}\Phi(g, h)$ en un couple (g, h) est égal au *résultant* $\mathrm{Res}(g, h)$. L'assertion résulte donc du théorème. □

Corollaire 3.9. — *Soit* $\varphi \in \mathrm{F}[\mathrm{T}]$ *un polynôme irréductible et unitaire tel que* $\varphi(0) \in \mathrm{A}$. *Alors* $\varphi \in \mathrm{A}[\mathrm{T}]$.

Démonstration. — Notons $\varphi = \sum_{m=0}^{n} a_m \mathrm{T}^m$. Il s'agit de prouver que $|a_m| \leqslant 1$ pour tout entier m ; notons p le plus grand entier tel que $|a_p| = \sup(|a_0|, \ldots, |a_n|)$. Si $p = 0$ ou $p = n$, on a $|a_p| \leqslant 1$ par hypothèse, d'où $\sup(|a_0|, \ldots, |a_n|) \leqslant 1$. Raisonnons par l'absurde ; on a donc $0 < p < n$. On écrit

$$a_p^{-1}\varphi = b_n \mathrm{T}^n + \cdots + b_{p+1}\mathrm{T}^{p+1} + \mathrm{T}^p + b_{p-1}\mathrm{T}^{p-1} + \cdots + b_0,$$

de sorte que $|b_0|, \ldots, |b_{p-1}| \leqslant 1$ et $|b_{p+1}|, \ldots, |b_n| < 1$. Posons alors $\psi = b_n \mathrm{T}^{n-p} + 1$ et $\eta = \mathrm{T}^p + b_{p-1}\mathrm{T}^{p-1} + \cdots + b_0$. Ce sont des éléments de $\mathrm{A}[\mathrm{T}]$; le polynôme $a_p^{-1}\varphi - \psi\eta$ est de degré $< n$ et ses coefficients sont tous de valeur absolue < 1. De plus, le résultant $\mathrm{Res}(\psi, \eta)$ (relativement à leurs degrés $n - p$ et p) est un élément de A dont l'image dans le corps résiduel k de F est le résultant des polynômes $\overline{\psi} = 1 + \overline{b_n}\mathrm{T}^{n-p}$ et $\overline{\eta} = \mathrm{T}^p$ relativement aux degrés $n - p$ et p. Comme ces polynômes n'ont pas de racine commune et comme $p = \deg(\overline{\eta})$, ce résultant n'est pas nul. Par suite, $|\mathrm{Res}(\psi, \eta)| = 1$. Le théorème entraîne donc qu'il existe un unique couple (ψ^*, η^*) de polynômes tels que $\deg(\psi^* - \psi) < n - p$, $\deg(\eta^* - \eta) < p$, $\|\psi^* - \psi\| < 1$, $\|\eta^* - \eta\| < 1$ et $a_p^{-1}\varphi = \psi^*\eta^*$. Alors, $\varphi = a_p\psi^*\eta^*$, ce qui contredit l'irréductibilité de φ. □

Donnons tout de suite une application importante.

Théorème 3.10. — *Soit* F *un corps muni d'une valeur absolue ultramétrique pour laquelle il est complet et soit* E *une extension algébrique de* F. *Il existe une unique valeur absolue sur* E *qui prolonge celle de* F*; elle applique un élément* $a \in \mathrm{E}$ *de polynôme minimal* P *sur* $|\mathrm{P}(0)|^{1/\deg(\mathrm{P})}$.

Démonstration. — Traitons d'abord le cas où la valeur absolue de F est triviale ; démontrons que la valeur absolue triviale est la seule valeur absolue sur F qui prolonge celle de E Soit $b \in \mathrm{E} - \mathrm{F}$ et soit $a_0, \ldots, a_n$ des éléments

de F, non tous nuls, tels que $a_n b^n + \cdots + a_0 = 0$; il s'agit de prouver que $|b| = 1$. Comme la valeur absolue de E est nécessairement ultramétrique, il existe deux entiers $p < q$ tels que $|a_p b^p| = |a_q b^q|$ et $a_p, a_q \neq 0$. Alors $|a_p| = |a_q| = 1$, puis $|b|^p = |b|^q$, d'où $|b| = 1$.

Supposons maintenant que la valeur absolue de F n'est pas triviale. Comme E est réunion d'extensions finies, il suffit de traiter le cas où E est une extension finie de F. Alors, E est un F-espace vectoriel de dimension finie.

Soit m_1 et m_2 des valeurs absolues sur E qui prolongent celle de F ; alors m_1 et m_2 sont des normes sur E. Comme la dimension de E est finie, comme toutes les normes sur E sont équivalentes et les valeurs absolues m_1 et m_2 définissent la même topologie sur E. Il existe donc un nombre réel $\rho > 0$ tel que $m_1(a) = m_2(b)^\rho$ pour tout $b \in \mathrm{E}$. En considérant un élément $b \in \mathrm{F}$ tel que $|b| \neq 1$, on obtient $\rho = 1$. Cela démontre que la valeur absolue de F se prolonge d'au plus une façon en une valeur absolue de E.

Pour $a \in \mathrm{E}$, considérons a comme un endomorphisme F-linéaire de E et posons $m(a) = |\det(a)|^{1/d}$, où $d = [\mathrm{E} : \mathrm{F}]$. On a $m(0) = 0$; $m(1) = 1$; $m(ab) = m(a)m(b)$ pour $a, b \in \mathrm{E}$; pour $a \in \mathrm{F}$, on a $m(a) = |\det(a)|^{1/d} = |a^d|^{1/d} = |a|$. Il reste à démontrer que $m(a + b) \leqslant \sup(m(a), m(b))$ pour tout $a, b \in \mathrm{E}$; pour cela, il suffit de prouver que $m(1 + a) \leqslant 1$ si $a \in \mathrm{E}$ et $m(a) \leqslant 1$. Soit $\varphi \in \mathrm{F}[\mathrm{T}]$ le polynôme minimal de a ; si son degré est e, le corps F est de façon naturelle un espace vectoriel de dimension d/e sur son sous-corps $\mathrm{F}(a)$. En considérant une base de F sur $\mathrm{F}(a)$, on voit que le polynôme caractéristique de a est égal à $\varphi^{d/e}$. Par suite, $\det(a) = \varphi^{d/e}(0)$ et $m(a) = |\varphi(0)|^{1/e}$. Comme le polynôme minimal de $1 + a$ est égal à $\varphi(\mathrm{T} - 1)$, on a $m(1 + a) = |\varphi(-1)|^{1/e}$. Par hypothèse, φ est irréductible, unitaire, et $\varphi(0) \in \mathrm{A}$ puisque $|\varphi(0)| = m(a)^e \leqslant 1$. D'après le corollaire, $\varphi \in \mathrm{A}[\mathrm{T}]$; alors $\varphi(-1) \in \mathrm{A}$ et $m(1 + a) \leqslant 1$.

La dernière relation a été établie au cours de la démonstration. □

3.11. — Soit F un corps muni d'une valeur absolue $|\cdot|$ ultramétrique pour laquelle il est complet. Supposons que F n'est pas algébriquement clos, considérons-en une clôture algébrique $\overline{\mathrm{F}}$ et munissons-la de l'unique valeur absolue qui prolonge celle de E.

Supposons d'abord que $[\overline{\mathrm{F}} : \mathrm{F}]$ est fini ; dans ce cas, $\overline{\mathrm{F}}$ est complet. Notons que cette situation est plutôt exceptionnelle. En effet, un théorème d'Artin

& Schreier (1927) affirme que cela ne se produit que si F est « réel clos » (F est ordonnable, et les polynômes vérifient le théorème des valeurs intermédiaires) auquel cas $[\overline{\mathrm{F}} : \mathrm{F}] = 2$ et $\overline{\mathrm{F}} = \mathrm{F}(\sqrt{-1})$. En particulier, F est de caractéristique zéro et les éléments négatifs n'ont pas de racine carrée. Vérifions aussi que la valeur absolue de F est triviale. Comme elle est ultramétrique, le théorème d'Ostrowski entraînerait sinon que F contient un corps p-adique. La contradiction vient de ce que d'après l'exemple 3.4, tout entier congru à 1 modulo p (modulo 8 si $p = 2$) possède une racine carrée dans $\mathbf{Q}_p$, donc dans F, même s'il est strictement négatif.

Ainsi, on peut se concentrer sur le cas où $[\overline{\mathrm{F}} : \mathrm{F}]$ est infini. Dans ce cas, $\overline{\mathrm{F}}$ n'est pas complet. (La méthode, un peu technique, consiste à construire une série rapidement convergente dont les termes successifs sont de degrés de plus en plus grand, de sorte que la limite n'est pas algébrique.) On peut alors considérer son complété $\widehat{\overline{\mathrm{F}}}$, et la proposition suivante affirme que c'est un corps algébriquement clos.

Proposition 3.12. — *Si* F *est un corps algébriquement clos muni d'une valeur absolue, alors son complété* $\widehat{\mathrm{F}}$ *est encore algébriquement clos.*

Démonstration. — Compte tenu du §2.3, on se borne à traiter le cas où la valeur absolue de F est ultramétrique. Considérons alors un polynôme unitaire irréductible $\varphi \in \widehat{\mathrm{F}}[\mathrm{T}]$; démontrons que son degré d est égal à 1. Pour cela, il suffit de démontrer qu'il possède une racine, racine que nous allons construire à l'aide de la méthode de Newton en partant d'une racine d'un polynôme à coefficients dans F qui est assez proche de φ.

Nous allons nous contenter de traiter le cas où φ est un polynôme *séparable*, c'est-à-dire que ses racines de φ dans une clôture algébrique de $\widehat{\mathrm{F}}$ sont simples ; de manière équivalente, les polynômes φ et φ' sont premiers entre eux. Cette condition est automatique lorsque F est de caractéristique zéro ; elle ne l'est pas lorsque F est de caractéristique p, le polynôme φ est alors un polynôme en T^p et on élimine ce cas sans trop de peine par des techniques qui n'apportent rien au propos de ce texte. Le fait que φ soit séparable se traduit aussi par le fait que le *discriminant* de φ, défini comme le résultant $\mathrm{Res}_{d,d-1}(\varphi, \varphi')$, n'est pas nul.

Quitte à remplacer φ par un polynôme de la forme $\varphi(c\mathrm{T})/c^d$, on peut aussi supposer que tous ses coefficients sont de valeur absolue au plus 1.

Soit $\psi \in \mathrm{F}[\mathrm{T}]$ un polynôme unitaire de degré d que nous prendrons assez proche de φ. Ses coefficients sont en particulier de valeur absolue au plus 1. Notons $a_1, \dots, a_d$ les racines de ψ dans F ; elles vérifient $|a_1|, \dots, |a_d| \leqslant 1$. (Si $\psi = \sum_{n=0}^{d} c_n \mathrm{T}^n$, on a $|c_n a^n| \leqslant |a|^n$ pour tout n, et $|c_d a^d| = |a|^d$ puisque $c_d = 1$; si $a \in \mathrm{F}$ vérifie $|a| > 1$, le terme d'indice d est de plus grande valeur absolue, à savoir $|a|^d$, et c'est le seul, si bien que $|\psi(a)| = |a|^d$; en particulier, $\psi(a) \neq 0$.)

Soit a l'un quelconque des a_j. Sous réserve que l'on ait $|\varphi(a)| < |\varphi'(a)|^2$, la méthode de Newton fournit un élément $\alpha \in \widehat{\mathrm{F}}$ tel que $|\alpha - a| < 1$ et $|\varphi(\alpha)| = 0$. Il suffit donc de justifier que cette inégalité est vérifiée lorsque que $\|\varphi - \psi\|$ est assez petit.

Comme $|a| \leqslant 1$ et $\psi(a) = 0$, on a $|\varphi(a)| = |\varphi(a) - \psi(a)| \leqslant \|\varphi - \psi\|$. Démontrons comment minorer $|\varphi'(a)|$.

En écrivant

$$|\varphi'(a) - \psi'(a)| \leqslant \|\varphi' - \psi'\| \leqslant \|\varphi - \psi\|,$$

on voit que $|\varphi'(a)| = |\psi'(a)|$ si $\|\varphi - \psi\| < |\psi'(a)|$; il suffit donc de minorer cette quantité. Pour cela, on fait appel aux formules classiques calculant le le discriminant de ψ en fonction de ses racines :

$$\mathrm{D}(\psi) = \prod_{1 \leqslant i < j \leqslant d} (a_j - a_i)^2 = (-1)^{(d-1)(d-2)/2} \prod_{j=1}^{d} \psi'(a_j).$$

Comme $|a_j| \leqslant 1$ pour tout j et $\|\psi\| \leqslant 1$, on a $|\psi'(a_j)| \leqslant 1$ pour tout j, si bien que $|\mathrm{D}(\psi)| = \prod_{j=1}^{d} |\psi'(a_j)| \leqslant |\psi'(a)|$. Enfin, la définition du discriminant comme déterminant d'une matrice formée des coefficients de ψ et ψ' l'exprime comme un polynôme à coefficients entiers en les coefficients de ψ. Grâce à l'inégalité ultramétrique, on constate donc que $|\mathrm{D}(\varphi) - \mathrm{D}(\psi)| \leqslant \|\varphi - \psi\|$. Si de plus $\|\varphi - \psi\| < |\mathrm{D}(\varphi)|$, il en résulte que $|\mathrm{D}(\psi)| = |\mathrm{D}(\varphi)|$.

En conclusion, dès que ψ est choisi de sorte que $\|\varphi - \psi\| < |\mathrm{D}(\varphi)|$, on a l'inégalité $|\varphi'(a)| \geqslant |\mathrm{D}(\varphi)|$.

Comme on a aussi $|\mathrm{D}(\varphi)| \leqslant 1$, il suffit de prendre ψ tel que $\|\varphi - \psi\| < |\mathrm{D}(\varphi)|^2$ pour que la méthode de Newton issue de n'importe laquelle de ses racines fournisse une racine de φ. $\square$

4. Polytopes et polygones de Newton

4.1. — Soit $\varphi = \sum_{m\in\mathbf{N}^n} a_m \mathrm{T}^m \in \mathrm{F}[\mathrm{T}]$ un polynôme en n indéterminées $\mathrm{T}_1, \ldots, \mathrm{T}_n$ et à coefficients dans F. Pour tout $r \in \mathbf{R}_+^n$ et $m \in \mathbf{N}^n$, on pose $r^m = r_1^{m_1} \ldots r_n^{m_n}$ et on définit

$$\|\varphi\|_r = \sup_{m\in\mathbf{N}^n} |a_m| r^m.$$

La notation à l'aide de multi-indices a pour but de laisser le lecteur ou la lectrice que ça arrangerait croire qu'il n'y a qu'une seule indéterminée.

On appelle support de φ l'ensemble (fini) S_φ des $m \in \mathbf{N}^n$ tels que $a_m \neq 0$. On appelle *polytope de Newton* de φ l'enveloppe convexe Π_φ de S_φ. On dit qu'un point $x \in \mathbf{R}^n$ est un sommet de Π_φ s'il existe un hyperplan affine H de $\mathbf{R}^n$ tel que $\mathrm{H} \cap \Pi_\varphi = \{x\}$; si f est une forme affine définissant H, la restriction de f à Π_φ a un signe constant. On démontre que les sommets de φ sont des points de S_φ et que Π_φ est l'enveloppe convexe de ses sommets.

Les égalités $\mathrm{S}_\varphi = \varnothing$, $\Pi_\varphi = \varnothing$ et $\varphi = 0$ sont équivalentes.

Proposition 4.2. — *L'application $\varphi \mapsto \|\varphi\|_r$ est une valeur absolue sur l'anneau* $\mathrm{F}[\mathrm{T}]$ *telle que $\|a\varphi\|_r = |a|\,\|\varphi\|_r$ pour tout $a \in \mathrm{F}$ et tout $\varphi \in \mathrm{F}[\mathrm{T}]$.*

Le cas où $n = 1$ et $r = 1$ est l'exemple 1.15 de la valeur absolue de Gauss ; la démonstration ci-dessous est plus géométrique.

Démonstration. — Les égalités $\|0\|_r = 0$ et $\|1\|_r = 1$ sont évidentes, de même que l'égalité $\|a\varphi\|_r = |a|\,\|\varphi\|_r$ pour $a \in \mathrm{F}$ et $\varphi \in \mathrm{F}[\mathrm{T}]$, et l'inégalité ultramétrique $\|\varphi + \psi\|_r \leqslant \sup(\|\varphi\|_r, \|\psi\|_r)$ pour $\varphi, \psi \in \mathrm{F}[\mathrm{T}]$. Démontrons que $\|\cdot\|_r$ est multiplicative.

Soit $\varphi, \psi \in \mathrm{F}[\mathrm{T}]$; prouvons $\|\varphi\psi\|_r = \|\varphi\|_r\,\|\psi\|_r$. Il suffit de traiter le cas où φ, ψ sont non nuls. Posons $\varphi = \sum a_m \mathrm{T}^m$, $\psi = \sum b_m \mathrm{T}^m$ et $\varphi\psi = \sum c_m \mathrm{T}^m$. Pour tout $m \in \mathbf{N}^n$, on a $c_m = \sum_{p+q=m} a_p b_q$, de sorte que

$$|c_m| r^m \leqslant \sup_{p+q=m} |a_p||b_q| r^m = \sup_{p+q=m} |a_p| r^p\, |b_q| r^q \leqslant \|\varphi\|_r\,\|\psi\|_r\,.$$

Inversement, notons $\mathrm{S}_{\varphi,r}$ l'ensemble des $m \in \mathbf{N}^n$ tels que $|a_m| r^m = \|\varphi\|_r$ et définissons $\mathrm{S}_{\psi,r}$ et $\mathrm{S}_{\varphi\psi,r}$ de façon analogue. Soit $m \in \mathbf{N}^n$. S'il n'y a aucun couple (p, q) tel que $p \in \mathrm{S}_{\varphi,r}$, $q \in \mathrm{S}_{\psi,r}$ et $p + q = m$, le calcul précédent montre que $|c_m| r^m < \|\varphi\|_r\,\|\psi\|_r$. Par suite, $\mathrm{S}_{\varphi\psi,r}$ est contenu dans la somme de Minkowski $\mathrm{S}_{\varphi,r} + \mathrm{S}_{\psi,r}$. Soit m un sommet de l'enveloppe convexe des $p + q$,

pour $p \in \mathrm{S}_{\varphi,r}$ et $q \in \mathrm{S}_{\psi,r}$; alors il n'existe qu'un seul couple (p,q) comme ci-dessus, et le même calcul garantit $|c_m|r^m = |a_p|r^p|b_q|r^q = \|\varphi\|_r \|\psi\|_r$. Par suite, $\|\varphi\psi\|_r = \|\varphi\|_r \|\psi\|_r$, comme il fallait démontrer. □

Remarque 4.3. — Dans le cas particulier où F est muni de la valeur absolue triviale, la démonstration précédente met en évidence le résultat géométrique suivant : *Soit φ, ψ des éléments non nuls de* $\mathrm{F}[\mathrm{T}]$. *Alors le polytope de Newton $\Pi_{\varphi\psi}$ de leur produit $\varphi\psi$ est la somme de Minkowski $\Pi_\varphi + \Pi_\psi$ des polytopes de Newton Π_φ et Π_ψ, c'est-à-dire l'ensemble des $x + y$, pour $x \in \Pi_\varphi$ et $y \in \Pi_\psi$.*

On peut en déduire des théorèmes d'irréductibilité dans l'anneau $\mathrm{F}[\mathrm{T}]$: si, par exemple, le polytope de Newton Π_φ n'est pas réduit à un point et n'est pas la somme de Minkowski $\mathrm{P} + \mathrm{Q}$ de deux polytopes non réduits à un point et à sommets entiers, alors φ est irréductible dans $\mathrm{F}[\mathrm{T}]$. Gao (2001) montre le lien avec les critères classiques d'Eisenstein et de Dumas et donne quelques exemples concrets où le polytope de Newton est contenu dans un triangle du plan.

4.4. — Parce que nous avons un peu de mal à considérer que la loi d'action de $\mathbf{R}$ sur $\mathbf{R}_+^*$ donnée par $(t,x) \mapsto x^t$ fait de $\mathbf{R}_+^*$ un espace affine sur $\mathbf{R}$, on passe au logarithme et on associe à tout polynôme $\varphi \in \mathrm{F}[\mathrm{T}]$ la fonction τ_φ sur $\mathbf{R}^n$, parfois appelée le polynôme tropical[2] associé à φ, qui est définie par $\tau_\varphi(x) = \log(\|\varphi\|_{e^x})$, Explicitement si $\varphi = \sum c_m \mathrm{T}^m$, on a

$$\tau_\varphi(x) = \sup_{m \in \mathrm{S}_\varphi} (\log(|c_m|) + m \cdot x).$$

C'est une fonction convexe continue, affine par morceaux, en tant que borne supérieure d'une famille finie de fonctions affines. Si τ_φ est affine sur un ouvert non vide V de $\mathbf{R}^n$ il existe un unique point $m \in \mathrm{S}_\varphi$ tel que τ_φ coïncide avec $x \mapsto \log(|c_m|) + m \cdot x$ sur V. Sur l'ensemble des $t \in \mathrm{F}^n$ tels que $\log(|t|) \in \mathrm{V}$, on a alors $\log(|\varphi(t)|) = \tau_\varphi(\log(|t|))$; en particulier, φ ne s'annule pas sur cet ensemble.

[2] La légende affirme que l'adjectif *tropical* a été introduit pour rendre hommage au mathématicien brésilien Imre Simon. Il reflète *de facto* une vision du monde eurocentrée. Son utilisation est ainsi critiquée dans ce contexte mathématique, en particulier par des mathématicien·nes d'origine sud-américaine qui, cependant, continuent à l'utiliser, peut-être dans l'attente d'une terminologie moins problématique ; nous faisons de même.

4.5. — La transformée de Legendre de τ_φ est définie par

$$v_\varphi(p) = \sup_x \big(p \cdot x - \tau_\varphi(x)\big),$$

où x parcourt $\mathbf{R}^n$. Elle est convexe et semi-continue inférieurement, en tant que borne supérieure d'une famille de fonctions affines, mais comme cette famille est infinie (ai-je supposé que l'espace est de dimension > 0 ?...), elle peut valoir $+\infty$. On l'appelle le *polygone de Newton* de φ.

Soit $m \in \mathrm{S}_\varphi$; on a donc $\tau_\varphi(x) \geqslant m \cdot x + \log(|c_m|)$ pour tout $x \in \mathbf{R}^n$, de sorte que $m \cdot x - \tau_\varphi(x) \leqslant -\log(|c_m|)$, si bien que $v_\varphi(m) \leqslant -\log(|c_m|)$; en particulier, $v_\varphi(m)$ est finie. Il en résulte que v_φ est finie et continue sur le polytope de Newton Π_φ de φ, et on peut démontrer qu'elle vaut $+\infty$ en dehors.

En fait, v_φ est la plus grande fonction v_φ sur Π_φ qui soit convexe, semi-continue inférieurement, et vérifie $v_\varphi(m) \leqslant -\log(|c_m|)$ pour $m \in \mathrm{S}_\varphi$; elle est affine par morceaux. Son épigraphe dans $\mathbf{R}^n \times \mathbf{R}$ est l'enveloppe convexe supérieure des points de la forme $(m, -\log(|c_m|))$, pour $m \in \mathrm{S}_\varphi$. En particulier, elle vaut $-\log(|c_m|)$ en tout sommet m de Π_φ.

4.6. — La dualité des fonctions convexes a une formulation symétrique en termes de « sous-différentiels » :

$$p \in \partial\tau_\varphi(x) \quad \Leftrightarrow \quad x \in \partial v_\varphi(p).$$

Comme les fonctions τ_φ et v_φ sont affines par morceaux, cela fournit une dualité géométrique entre domaines d'affinité de τ_φ et v_φ pour lesquels la dimension est complémentaire. On a vu le lien entre polytopes maximaux sur lesquels τ_φ est affine et points anguleux de v_φ. Inversement, il y a correspondance entre points anguleux de τ_φ et polytopes maximaux sur lesquels v_φ est affine.

4.7. — C'est dans une lettre d'Isaac Newton à Henry Oldenburg qu'apparaît ce polygone, dans le contexte de la résolution d'une équation en deux variables :

$$y^6 - 5xy^5 + (x^3/a)y^4 - 7a^2x^2y^2 + 6a^3x^3 + b^2x^4 = 0$$

pour laquelle Newton cherche à exprimer y en termes d'une série de puissances de x ; il fait à peu près le dessin ci-dessous, si ce n'est que nous plaçons les étiquettes et les points aux intersections et non dans les cases : les points

marqués sont les couples (m, n) tels que $x^n y^m$ apparaît dans le polynôme ci-dessus, et le trait gras est une « règle » qu'on placerait sous ces points, le plus haut possible. Newton extrait du polynôme les termes pour lesquels le

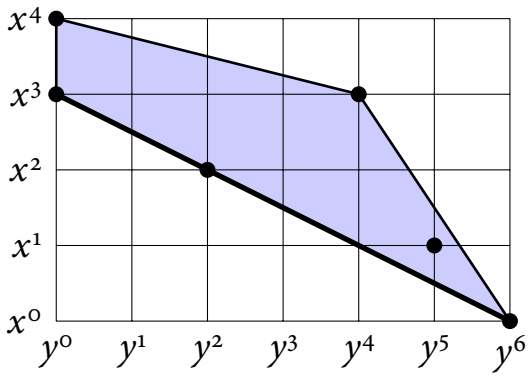

FIGURE 1. L'exemple de Newton : $\varphi = y^6 - 5xy^5 + (x^3/a)y^4 - 7a^2x^2y^2 + 6a^3x^3 + b^2x^4 = 0$

couple (m, n) est sur le trait gras, ici :

$$y^6 - 7a^2x^2 + 6a^3x^3 = 0$$

qu'il réduit en $v^6 - 7v^2 + 6 = 0$, en posant $y = v\sqrt{ax}$. C'est une équation « bicubique » dont les solutions sont les racines carrées de 1, 2 et -3 ; Newton écrit alors

$$y = +\sqrt{ax}, \quad -\sqrt{ax}, \quad +\sqrt{2ax}, \quad -\sqrt{2ax},$$

expressions qu'il voit comme le premier terme de la solution, car il évite (sans même les mentionner) les solutions $y = \pm\sqrt{-3ax}$. Pour chacun de ces termes p, il récrit alors l'équation pour $y - p$, et répète l'argument.

Dans notre langage, il convient de considérer l'anneau des séries formelles $\mathbf{C}[[x]]$ à coefficients complexes, ou plutôt son corps des fractions $\mathrm{F} = \mathbf{C}((x))$, muni d'une valeur absolue ultramétrique pour laquelle $|a| = 1$ pour tout $a \in \mathbf{C}^*$ et $|x| < 1$; pour fixer les idées et simplifier les logarithmes ci-dessous, nous prenons $|x| = 1/e$. Le polynôme de deux variables devient alors le polynôme d'une variable

$$\varphi = \mathrm{T}^6 - 5x\mathrm{T}^5 + (x^3/a)\mathrm{T}^4 - 7a^2x^2\mathrm{T}^2 + (6a^3x^3 + b^2x^4)$$

à coefficients dans F. Si on note $\varphi = \sum c_m \mathrm{T}^m$, on a donc $\log(|c_0|) = -3$, $\log(|c_2|) = -2$, $\log(|c_4|) = -3$ et $\log(|c_5|) = -1$. Ainsi, à une symétrie par

rapport à l'axe horizontal près, le polygone de Newton de φ est précisément donné par le dessin que Newton a réalisé.

On devine dans l'exemple de Newton que les racines y de φ, y compris les deux que Newton n'a pas considérées, auront pour valeur absolue $e^{-1/2}$. C'est effectivement le cas : les pentes du polygone de Newton sont les logarithmes des valeurs absolues des racines.

Théorème 4.8. — *Soit* F *un corps muni d'une valeur absolue ultramétrique et soit* $\varphi = \sum c_j \mathrm{T}^j \in \mathrm{F}[\mathrm{T}]$ *un polynôme en une indéterminée tel que* $c_0 = \varphi(0) \neq 0$*; soit* d *son degré.*

(1) *Pour toute racine* a *de* φ, $\log(|a|)$ *est une pente de* v_φ.

(2) *Si* φ *est scindé, il existe une numérotation* $\{a_1, \dots, a_d\}$ *de ses racines telle que pour tout entier* $j \in \{1, \dots, d\}$, *la pente de* v_φ *sur l'intervalle* $[j-1; j]$ *soit égale à* $\log(|a_j|)$.

(3) *Si* F *est complet et* φ *est irréductible, le polygone de Newton de* φ *est affine, de pente* $\log(|c_d|/|c_0|)/d$.

Démonstration. — Notons $\varphi = \sum_{m=0}^{d} c_m \mathrm{T}^m$. Par hypothèse, $c_0 \neq 0$ et $c_d \neq 0$, de sorte que $\Pi_\varphi = [0; d]$.

(1) Soit a une racine de φ et soit M l'ensemble des entiers $m \in \{0, \dots, d\}$ tels que $|c_m||a|^m$ soit maximal, c'est-à-dire $\tau_\varphi(\log(|a|)) = \log(|c_m|) + m\log(|a|)$. Puisque $\varphi(a) = \sum_{m=0}^{d} c_m a^m = 0$, l'ensemble M possède au moins deux éléments. La fonction $h\colon t \mapsto t\log(|a|) - \tau_\varphi(\log(|a|))$ est affine ; on a $h(m) = -\log(|c_m|)$ pour $m \in \mathrm{M}$, et $h(m) < -\log(|c_m|)$ sinon. Puisque h est affine, elle est convexe et la définition de v_φ entraîne l'inégalité $h \leqslant v_\varphi$. Puisque $v_\varphi - h$ est positive et convexe, l'ensemble de ses zéros est convexe : les fonctions h et v_φ coïncident sur l'enveloppe convexe de M. Cela prouve que $\log(|a|)$ est une pente de v_φ.

(2) La fonction v_φ est convexe, affine par morceaux ; ses pentes sont donc en ordre croissant. Ordonnons les racines de φ dans l'ordre de leurs valeurs absolues, disons $a_1, \dots, a_d$. Soit h la fonction de $[0; d]$ dans $\mathbf{R}$ qui est affine sur chaque intervalle $[m-1; m]$ et telle que

$$h(m) = -\log(|a_{d-m+1}|) - \cdots - \log(|a_d|) - \log(|c_d|)$$

pour tout $m \in \{1, \dots, d\}$. Sa pente sur l'intervalle $[m-1; m]$ est $-\log(|a_{d-m}|)$; ses pentes sont donc croissantes, de sorte que h est convexe.

Les coefficients de φ s'expriment comme les fonctions symétriques élémentaires des a_i : pour tout entier m tel que $0 \leqslant m \leqslant d$, on a

$$c_m = (-1)^{d-m} c_d \sum_{i_1 < \cdots < i_{d-m}} a_{i_1} \dots a_{i_{d-m}}.$$

Dans cette expression, le terme de plus grande valeur absolue est obtenu pour $(i_1, \dots, i_{d-m}) = (d-m+1, \dots, d)$, de sorte que

$$|c_m| \leqslant |c_d| |a_{d-m+1}| \dots |a_d|$$

pour tout entier $m \in \{0, \dots, d\}$, c'est-à-dire $h(m) \leqslant -\log(|c_m|)$. Puisque la fonction h est convexe et que v_φ est la plus grande fonction convexe semi-continue inférieurement qui vaut au plus $-\log(|c_m|)$ en m, on a $h(t) \leqslant v_\varphi(t)$ pour tout $t \in [0; d]$.

Soit Σ l'ensemble des entiers $m \in \{0, \dots, d\}$ tels que $m = d$ ou $|a_{d-m}| < |a_{d-m+1}|$. Si $m \in \Sigma$, il n'y a qu'un seul terme, dans l'expression de c_m donnée ci-dessus, qui soit de valeur absolue maximale, si bien que l'on a en fait $h(m) = -\log(|c_m|)$. On a donc $h(m) = v_\varphi(m)$.

Sur un intervalle $[m, p]$ dont les extrémités sont des éléments consécutifs de Σ, la fonction ψ est affine, donc la fonction $v_\varphi - h$ est convexe, semi-continue inférieurement et positive ; comme elle s'annule à ses extrémités, elle est constante sur cet intervalle. Cela prouve que $v_\varphi = h$ et conclut la démonstration de (2).

(3) Supposons maintenant que F soit complet. Nous allons donner un argument par la théorie de Galois en nous restreignant au cas où φ est séparable (le cas général s'en déduit en écrivant $\varphi = \psi(\mathrm{T}^q)$, où q est une puissance de la caractéristique de F et ψ est un polynôme séparable). Soit E une extension algébrique finie de F, galoisienne, dans laquelle le polynôme φ est scindé ; munissons-la de l'unique valeur absolue qui prolonge celle de F. Soit G le groupe de Galois de l'extension E/F ; c'est le groupe (fini) des automorphismes de corps de E qui sont l'identité sur F. Pour tout $\sigma \in \mathrm{G}$, l'application $x \mapsto |\sigma(x)|$ de E dans $\mathbf{R}_+$ est une valeur absolue sur E dont la restriction à F est la valeur absolue initiale. Comme F est complet, la valeur absolue de E est la seule qui prolonge celle de E, si bien que $|\sigma(x)| = |x|$ pour tout $x \in \mathrm{E}$.

Par ailleurs, le groupe G stabilise l'ensemble $\{a_1, \dots, a_d\}$ des racines de φ ; comme φ est irréductible, l'action qui en résulte est transitive. Par suite,

toutes les racines de φ ont même valeur absolue. En appliquant l'analyse de (2) au corps E, il en résulte que le polygone de Newton de φ n'a qu'une seule pente. Comme il vaut $\log(|c_0|)$ en 0 et $\log(|c_d|)$ en d, la formule annoncée en résulte. □

Corollaire 4.9. — *Soit* F *un corps, soit m une valeur absolue ultramétrique non triviale sur* F*, dont la valuation est discrète. Soit π un générateur de l'idéal maximal de l'anneau de valuation* A *de* F*. Soit $\varphi \in \mathrm{F}[\mathrm{T}]$ un polynôme unitaire ; soit d son degré. Supposons que son polygone de Newton soit affine, de pente $(r/d)\log(|\pi|)$, où r et d sont premiers entre eux. Alors φ est irréductible.*

Démonstration. — On peut supposer que F est complet. Soit ψ un facteur irréductible unitaire de φ. Puisque les pentes de son polygone de Newton sont de la forme $\log(|a|)$, où a est une racine de ψ, donc de φ, le polygone de Newton de ψ est affine, et sa pente est $(r/d)\log(|\pi|)$. Posons $\psi = \mathrm{T}^e + c_{e-1}\mathrm{T}^{e-1} + \cdots + c_0$; on a nécessairement $\log(|c_0|) = e(r/d)\log(|\pi|)$. Nécessairement, d divise er. Comme r et d sont premiers entre eux, d divise e, d'où $d = e$. Ainsi, $\psi = \varphi$. □

Exemple 4.10. — Un cas particulier du corollaire est le *critère d'Eisenstein* qui correspond au cas d'un polynôme $\varphi = \mathrm{T}^d + c_{d-1}\mathrm{T}^{d-1} + \cdots + c_0$ tel que $c_0, \ldots, c_{d-1}$ appartiennent à l'idéal maximal M de l'anneau de valuation de F, et tel que c_0 n'appartienne pas à M^2. Avec les notations du corollaire, le polygone de Newton de φ est affine, à pente $(1/d)\log(|\pi|)$; le polynôme φ est donc irréductible.

Pour un exemple spécifique et très classique, soit p un nombre premier impair et posons

$$\varphi = \frac{(\mathrm{T}+1)^p - 1}{\mathrm{T}} = \mathrm{T}^{p-1} + \binom{p}{p-1}\mathrm{T}^{p-2} + \cdots + \binom{p}{2}\mathrm{T} + \binom{p}{1}.$$

Ses racines complexes sont de la forme $-1 + \omega$, où $\omega^p = 1$ et $\omega \neq 1$; autrement dit, il est égal à $\Phi_p(\mathrm{T}+1)$, où Φ_p est le polynôme cyclotomique d'indice p. Le coefficient dominant de φ vaut 1, et comme p est premier, tous les autres coefficients sont multiples de p ; le coefficient constant, égal à p, n'est pas multiple de p^2. En appliquant le critère d'Eisenstein au corps $\mathbf{Q}$, muni de la

valeur absolue p-adique, on en déduit que φ est irréductible. En fait, il est même irréductible comme élément de $\mathbf{Q}_p[\mathrm{T}]$.

Exemple 4.11. — Le critère d'Eisenstein montre comment déduire du polygone de Newton un critère d'irréductibilité « local ». Coleman (1987) a expliqué comment combiner plusieurs valeurs absolues pour en déduire des résultats d'irréductibilité « globaux ». Il donne en particulier l'exemple du polynôme

$$\varphi = 1 + \mathrm{T} + \frac{1}{2}\mathrm{T}^2 + \cdots + \frac{1}{n!}\mathrm{T}^n$$

à coefficients dans $\mathbf{Q}$.

Fixons un nombre premier p, munissons $\mathbf{Q}$ de la valeur absolue p-adique et calculons son polygone de Newton. C'est l'enveloppe convexe supérieure de l'ensemble des points de la forme $(m, -\log(|1/m!|_p)) = (m, -v_p(m!)\log(p))$, pour $m \in \{0, \dots, n\}$.

Écrivons le développement en base p de n :

$$n = a_m p^m + a_{m-1}p^{m-1} + \cdots + a_0,$$

où $a_0, \dots, a_m \in \{0, \dots, p-1\}$. Soit $(m_1, \dots, m_s)$ la suite strictement décroissante des entiers $k \in \{0, \dots, m\}$ tel que $a_k \neq 0$; pour tout i, posons $n_i = a_{m_1}p^{m_1} + \cdots + a_{m_i}p^{m_i}$. Démontrons que les sommets de τ_φ sont les points de la forme $(n_i, -v_p(n_i!)\log(p))$, pour $i \in \{1, \dots, s\}$.

Rappelons d'abord comment évaluer la valuation p-adique de $n!$: parmi tous les entiers entre 1 et n, $\lfloor n/p \rfloor$ sont multiples de p, $\lfloor n/p^2 \rfloor$ sont multiples de p^2, etc., de sorte que

$$v_p(n!) = \sum_{k=1}^{\infty} \lfloor n/p^k \rfloor.$$

Alors, $\lfloor n/p^k \rfloor = a_m p^{m-k} + \cdots + a_k$ pour $k \leqslant m$, et est nul pour $k > m$, si bien que

$$v_p(n!) = \sum_{k=1}^{m}\sum_{j=k}^{m} a_j p^{j-k} = \sum_{j=1}^{m} a_j \sum_{k=1}^{j} p^{j-k} = \sum_{j=1}^{m} a_j \frac{p^j - 1}{p-1} = \frac{n - (a_0 + \cdots + a_m)}{p-1}.$$

En notant $\sigma(n) = a_0 + \cdots + a_m$; on a donc $v_p(n!) = (n - \sigma(n))/(p-1)$.

Par définition, v_φ est la plus grande fonction convexe semi-continue inférieurement telle que $v_\varphi(m) \leqslant -v_p(m!)\log(p)$ pour tout $m \in \{0, \dots, n\}$. Par

suite, $h\colon t \mapsto v_\varphi(t)(p-1)/\log(p)+t$ est la plus grande fonction convexe semi-continue inférieurement telle que $h(m) \leqslant \sigma(m)$ pour tout $m \in \{0,\dots,n\}$.

Comme la fonction nulle vérifie ces inégalités, on a $h(t) \geqslant 0$ pour tout t. Par suite, $h(0) = 0$. Si $(a-1)p^{m_1} < m \leqslant ap^{m_1}$, on a $\sigma(m) \geqslant a$, et $\sigma(m) \geqslant a+1$ si $a_1p^{m_1} < m \leqslant n$. Par suite, la fonction donnée par $t \mapsto t/a_1p^{m_1}$ vérifie les inégalités voulues ; cela prouve l'inégalité $t/p^{m_1} \leqslant h(t)$ pour tout $t \in [0;n]$. Comme $h(a_1p^{m_1}) \leqslant \sigma(a_1p^{m_1}) = a_1$, on a donc $h(a_1p^{m_1}) = a_1$ et $h(t) = t/p^{m_1}$ sur l'intervalle $[0;a_1p^{m_1}]$.

On raisonne maintenant de même sur l'intervalle $[a_1p^{m_1},n]$. Pour tout entier m dans cet intervalle, on a $\sigma(m) = 1+\sigma(m-a_1p^{m_1})$. Posant $h_1(t) = h(t+a_1p^{m_1})-a_1$, on constate que h_1 est à l'entier $n-a_1p^{m_1}$ ce que h est à l'entier n. Cela prouve que h est la fonction affine par morceaux de pente $1/p^{m_1}$ sur l'intervalle $[0;n_1]$, de pente $1/p^{m_2}$ sur l'intervalle $[n_1;n_2]$, …, de pente $1/p^{m_s}$ sur l'intervalle $[n_{s-1};n]$. Les sommets de h sont donc les entiers n_i, et ceux de v_φ aussi.

Les pentes de v_φ sont donc les nombres réels

$$\left(\frac{1}{p^{m_i}(p-1)}-\frac{1}{p-1}\right)\log(p) = -\frac{p^{m_i}-1}{p^{m_i}(p-1)}\log(p).$$

Démontrons que φ est irréductible dans $\mathbf{Q}[\mathrm{T}]$. Considérons un facteur irréductible ψ de φ ; il s'agit de prouver que $\deg(\psi) = \deg(\varphi)$.

Soit p un nombre premier qui divise n et soit $\psi_1 \in \mathbf{Q}_p[\mathrm{T}]$ un facteur irréductible de ψ. Comme $\mathbf{Q}_p$ est complet, le théorème 4.8, (3), entraîne que le polygone de Newton (pour la valeur absolue p-adique) de ψ_1 n'a qu'une seule pente, et cette pente v est de la forme $k\log(p)/\deg(\psi_1)$, pour un certain entier relatif k. D'autre part, v apparaît dans les pentes du polygone de Newton de φ, de sorte qu'il existe $i \in \{1,\dots,s\}$ tel que $kp^{m_i} = -\deg(\psi_1)(p^{m_i}-1)/(p-1)$. En particulier, p^{m_i} divise $\deg(\psi_1)$. Or, avec les notations précédentes, on $m_s = v_p(n)$, si bien que $p^{v_p(n)}$ divise $\deg(\psi_1)$. Comme c'est le cas de chaque facteur irréductible ψ_1 de ψ, il s'ensuit que $p^{v_p(n)}$ divise $\deg(\psi)$.

Comme ceci vaut pour tout nombre premier p qui divise n, on en conclut que $\deg(\varphi) = n = \prod_p p^{v_p(n)}$ divise $\deg(\psi)$. Par conséquent, $\deg(\psi) = \deg(\varphi)$, et cela conclut la preuve que φ est irréductible.

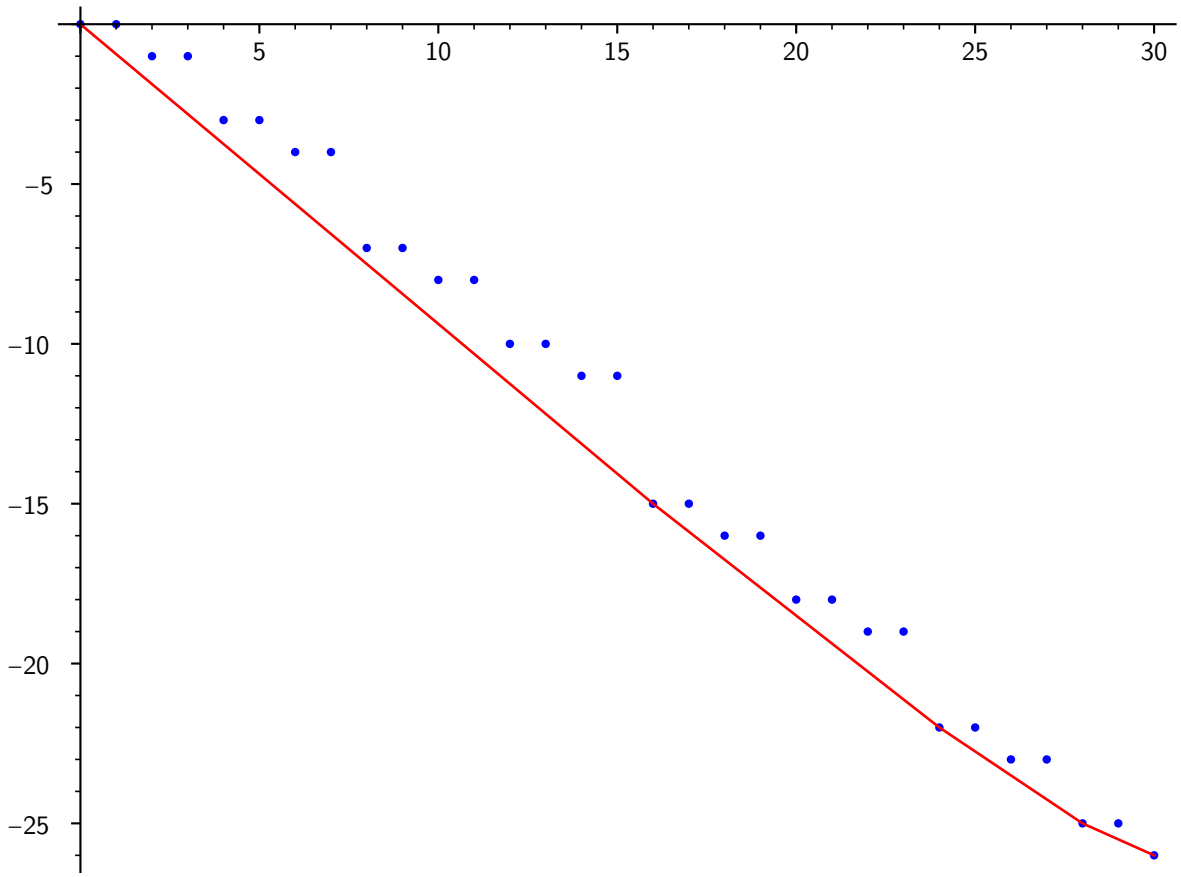

FIGURE 2. Polygone de Newton du polynôme $\sum_{n=0}^{30} T^n/n!$ pour la valeur absolue 2-adique. Les 4 pentes correspondent à la décomposition $30 = 2^4 + 2^3 + 2^2 + 2$

En précisant cet argument, COLEMAN (1987) redémontre un théorème de SCHUR (1930) selon lequel le groupe de Galois du polynôme φ est le groupe alterné $\mathfrak{A}_n$ si n est multiple de 4, et le groupe symétrique sinon.

5. Fonctions et équations analytiques

Soit F un corps muni d'une valeur absolue pour laquelle il est complet. On peut développer une théorie des séries entières dans F dont les débuts sont tout à fait parallèles à celle des séries entières à coefficients complexes. Pour simplifier, on n'expose que le cas d'une seule indéterminée.

5.1. — Soit $\varphi = \sum_{n=0}^{\infty} c_n \mathrm{T}^n$ une série formelle en une indéterminée T à coefficients dans F. On définit son *rayon de convergence* $\mathrm{R}_\varphi \in [0; +\infty]$ par la formule usuelle :

$$\mathrm{R}_\varphi = \left(\limsup_n |c_n|^{1/n}\right)^{-1}.$$

Par définition, si $|a| < \mathrm{R}_\varphi$, alors $c_n a^n$ tend vers 0, tandis que si $|a| > \mathrm{R}_\varphi$, alors $c_n a^n$ n'est pas bornée.

Le cas de la valeur absolue triviale est peu intéressant : on a $\mathrm{R}_\varphi = 1$ si $\varphi \neq 0$, et pour $a \neq 0$, la série $\sum c_n a^n$ ne converge que si φ est un polynôme. Le cas d'une valeur absolue archimédienne est probablement bien connu ; on supposera donc dans la suite que la valeur absolue est ultramétrique et non triviale.

5.2. — Pour toute série formelle $\varphi \in \mathrm{F}[[\mathrm{T}]]$, disons $\varphi = \sum c_n \mathrm{T}^n$, et tout nombre réel $\mathrm{R} \geqslant 0$, posons $\|\varphi\|_\mathrm{R} = \sup_n |c_n| \mathrm{R}^n$; c'est un élément de $[0; +\infty]$. C'est une fonction croissante de R, continue à gauche : pour $\mathrm{R} > 0$, on a

$$\|\varphi\|_\mathrm{R} = \sup_{r<\mathrm{R}} \|\varphi\|_r .$$

Soit en effet A tel que $\mathrm{A} < \|\varphi\|_\mathrm{R}$; il existe donc $n \geqslant 0$ tel que $\mathrm{A} < |c_n|\mathrm{R}^n \leqslant \|\varphi\|_\mathrm{R}$; il existe alors r tel que $r < \mathrm{R}$ et $\mathrm{A} < |c_n| r^n$, d'où $\mathrm{A} < \|\varphi\|_r \leqslant \|\varphi\|_\mathrm{R}$.

5.3. — Pour $a \in \mathrm{F}$ et $\varphi, \psi \in \mathrm{F}[[\mathrm{T}]]$ tels que $\|\varphi\|_\mathrm{R}, \|\psi\|_\mathrm{R} < \infty$, on a les relations :

(1) Si $\mathrm{R} > 0$, alors $\|\varphi\|_\mathrm{R} = 0$ si et seulement si $\varphi = 0$ (pour $\mathrm{R} = 0$, on a $\|\varphi\|_0 = |c_0|$) ;

(2) On a $\|\varphi + \psi\|_\mathrm{R} \leqslant \sup(\|\varphi\|_\mathrm{R}, \|\psi\|_\mathrm{R})$;

(3) Avec la convention $0 \cdot \infty = 0$, on a $\|a\varphi\|_\mathrm{R} = |a|\, \|\varphi\|_\mathrm{R}$ et $\|\varphi\psi\| = \|\varphi\|\, \|\psi\|$.

Cela démontre que l'ensemble des séries φ telles que $\|\varphi\|_\mathrm{R} < +\infty$ est une sous-algèbre de $\mathrm{F}[[\mathrm{T}]]$ sur laquelle $\|\cdot\|_\mathrm{R}$ est une norme et une valeur absolue. Seule la propriété de multiplicativité n'est pas immédiate : posons $\varphi = \sum a_n \mathrm{T}^n$ et $\psi = \sum b_n \mathrm{T}^n$; alors, $\varphi\psi = \sum_n c_n \mathrm{T}^n$, où $c_n = \sum_{k=0}^n a_{n-k} b_k$ pour tout k. Par hypothèse, on a $|a_{n-k}|\mathrm{R}^{n-k} \leqslant \|\varphi\|_\mathrm{R}$ et $|b_k|\mathrm{R}^k \leqslant \|\psi\|_\mathrm{R}$ pour tous $k, n \in \mathbf{N}$ tels que $k \leqslant n$; par l'inégalité ultramétrique, on en déduit

$$|c_n|\mathrm{R}^n \leqslant \sup_k (|a_{n-k}||b_k|\mathrm{R}^n) \leqslant \|\varphi\|_\mathrm{R}\, \|\psi\|_\mathrm{R} ,$$

de sorte que $\|\varphi\psi\|_{\mathrm{R}} \leqslant \|\varphi\|_{\mathrm{R}} \|\psi\|_{\mathrm{R}}$. Pour démontrer l'autre inégalité, considérons des nombres réels α et β tels que $0 \leqslant \alpha < \|\varphi\|_{\mathrm{R}}$ et $0 \leqslant \beta < \|\psi\|_{\mathrm{R}}$. Soit alors $r \in \mathbf{R}$ tel que $0 \leqslant r < \mathrm{R}$ et $\alpha < \|\varphi\|_r$ et $\beta < \|\psi\|_r$. Comme $0 \leqslant r < \mathrm{R}$, les suites $|a_n|r^n$ et $|b_n|r^n$ tendent vers 0, de sorte qu'il existe un plus grand entier $m \in \mathbf{N}$ tel que $|a_m|r^m = \|\varphi\|_r$ et un plus grand entier $p \in \mathbf{N}$ tel que $|b_p|r^p = \|\psi\|_r$. Posons $n = m + p$; dans l'expression $c_n = \sum_{k=0}^{n} a_{n-k}b_k$, le terme $a_m b_p$ est le seul qui soit de valeur absolue maximale, de sorte que $|c_n|r^n = |a_m|r^m|b_p|r^p = \|\varphi\|_r \|\psi\|_r$. Alors, $\|\varphi\psi\|_r = \|\varphi\|_r \|\psi\|_r > \alpha\beta$. A fortiori, $\|\varphi\psi\|_{\mathrm{R}} \geqslant \alpha\beta$, et l'inégalité $\|\varphi\psi\|_{\mathrm{R}} \geqslant \|\varphi\|_{\mathrm{R}} \|\psi\|_{\mathrm{R}}$ s'en déduit en faisant tendre α et β vers $\|\varphi\|_{\mathrm{R}}$ et $\|\psi\|_{\mathrm{R}}$ respectivement.

Lemme 5.4. — *Soit* R *un nombre réel strictement positif. Les séries* $\sum c_n \mathrm{T}^n$ *telles que* $(|c_n|\mathrm{R}^n)$ *tende vers* 0 *forment une sous-algèbre de* $\mathrm{F}[[\mathrm{T}]]$ *dont les éléments ont tous rayon de convergence au moins* R.

La restriction de $\|\cdot\|_{\mathrm{R}}$ *à cette algèbre est une norme et une valeur absolue pour laquelle elle est complète.*

Lorsque $\mathrm{R} = 1$, cette algèbre est notée $\mathrm{F}\langle \mathrm{T}\rangle$ et est appelée *algèbre de Tate* en l'indéterminée T ; ses éléments sont aussi appelés *séries formelles restreintes*. La définition générale de cette algèbre semble due à Berkovich (1990) ; nous l'appellerons l'*algèbre de Tate* de rayon R et la noterons $\mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$.

Démonstration. — Les démonstrations se déduisent pour l'essentiel de ce qui précède. Démontrons la stabilité par produit. Si $\varphi = \sum a_n \mathrm{T}^n$, $\psi = \sum b_n \mathrm{T}^n$ et $\varphi\psi = \sum c_n \mathrm{T}^n$, on a $c_n = \sum_{p+q=n} a_p b_q$ pour tout n ; démontrons que $|c_n|\mathrm{R}^n$ tend vers 0. Soit $\varepsilon > 0$; soit $\mathrm{N} \geqslant 0$ tel que $|a_n|\mathrm{R}^n \leqslant \varepsilon$ et $|b_n|\mathrm{R}^n \leqslant \varepsilon$ pour $n \geqslant \mathrm{N}$. Alors, si $n \geqslant 2\mathrm{N}$, toute décomposition $n = p+q$ entraîne $p \geqslant \mathrm{N}$ ou $q \geqslant \mathrm{N}$; dans le premier cas, $|a_p|\mathrm{R}^p|b_q|\mathrm{R}^q \leqslant \varepsilon \|\psi\|$, et dans le second, $|a_p|\mathrm{R}^p|b_q|\mathrm{R}^q \leqslant \varepsilon \|\varphi\|$. Ainsi, $|c_n|\mathrm{R}^n \leqslant \varepsilon \sup(\|\varphi\|, \|\psi\|)$. Cela prouve que $|c_n|\mathrm{R}^n$ tend vers 0.

Cette algèbre est complète : soit (φ_m) une suite de Cauchy dans $\mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$; pour tout m, posons $\varphi_m = \sum c_{m,n} \mathrm{T}^n$.

Soit n un entier. On a $|c_{m,n} - c_{p,n}|\mathrm{R}^n \leqslant \|\varphi_m - \varphi_p\|$, pour tous m, p, ce qui prouve que la suite $(c_{m,n})_m$ vérifie le critère de Cauchy. Puisque F est complet, elle converge dans F ; notons c_n sa limite. Soit φ la série formelle $\sum c_n \mathrm{T}^n$; démontrons qu'elle appartient à $\mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$ et que (φ_m) converge vers φ dans $\mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$.

Soit $\varepsilon > 0$ et soit M un entier tel que $\|\varphi_m - \varphi_p\| \leqslant \varepsilon$ pour $m, p \geqslant M$. Fixons $m \geqslant M$; en faisant tendre p vers l'infini, on obtient $|c_{m,n} - c_n|R^n \leqslant \varepsilon$ pour tout n. Puisque $(|c_{m,n}|R^n)_m$ tend vers 0, il existe un entier N tel que $|c_{m,n}|R^n \leqslant \varepsilon$ pour tout $n \geqslant N$. Alors $|c_n|R^n \leqslant \sup(|c_{m,n} - c_n|, |c_{m,n}|)R^n \leqslant \varepsilon$ pour $n \geqslant N$. Cela prouve que $(|c_n|R^n)$ tend vers 0, donc $\varphi \in F\langle T/R\rangle$. Les mêmes inégalités prouvent que $\|\varphi_m - \varphi\| \leqslant \varepsilon$ pour $m \geqslant M$. Cela démontre que (φ_m) tend vers φ. □

5.5. — Soit $\varphi \in F\langle T/R\rangle$; notons $\varphi = \sum c_n T^n$. Soit $a \in F$ tel que $|a| \leqslant R$. Par hypothèse, la suite de terme général $c_n a^n$ tend vers 0 ; Comme F est complet, la série $\sum c_n a^n$ converge ; on note $\varphi(a)$ sa somme. Il y a même convergence normale, de sorte que l'application $a \mapsto \varphi(a)$ ainsi définie est continue sur la boule circonférenciée $B(0, R)$.

De plus, l'application $\varphi \mapsto \varphi(a)$ est un morphisme d'algèbres de $F\langle T/R\rangle$ dans F.

Il y a également un résultat pour la composition des séries formelles : soit $\varphi \in F[[T]]$ et soit $\psi \in F\langle T/S\rangle$. Notons $\varphi = \sum_{m=0}^{\infty} c_m T^m$. Si la suite $(|c_m| \, \|\psi\|_S^m)$ tend vers 0, la série $\sum c_m \psi^m$ converge dans l'algèbre complète $F\langle T/S\rangle$. On note $\varphi \circ \psi$ sa limite. C'est en particulier le cas si $\|\psi\|_S < R_\varphi$, ou si $\varphi \in F\langle T/R\rangle$ et $\|\psi\|_S \leqslant R$.

Notons deux cas particuliers importants :

(1) Soit $a \in F$; supposons $\varphi \in F\langle T/R\rangle$ et $|a| \leqslant R$. Alors, on peut former la série formelle $\varphi_a = \varphi(a + T)$; elle appartient à $F\langle T/R\rangle$; pour tout $b \in B(0; R)$, on a $\varphi_a(b) = \varphi(a + b)$. Les coefficients de φ_a sont donnés par la « formule de Taylor ». Il existe une unique suite $(\varphi^{[n]})$ d'éléments de $F\langle T/R\rangle$ telle que

$$\varphi_a(T) = \sum_{n=0}^{\infty} \varphi^{[n]}(a) T^n.$$

Explicitement, si $\varphi = \sum c_m T^m$, on a

$$\varphi^{[n]} = \sum_{m=0}^{\infty} c_{m+n}(m + n) \dots (m + 1) T^n.$$

Lorsque F est de caractéristique zéro, $\varphi^{[n]}$ est reliée à la n-ième dérivée formelle de φ par la formule $\varphi^{[n]} = \varphi^{(n)}/n!$. En particulier, φ est indéfiniment dérivable sur $B(0; R)$, de dérivée $\varphi^{[1]}$.

(2) Le rayon de convergence de la série formelle $1/(1-\mathrm{T})$ est égal à 1. Ainsi, toute série $\psi \in \mathrm{F}\langle \mathrm{T}/\mathrm{S}\rangle$ telle que $\|\psi - 1\|_{\mathrm{S}} < 1$ est inversible dans $\mathrm{F}\langle \mathrm{T}/\mathrm{S}\rangle$. Nous déterminerons ci-dessous les éléments inversibles de cette algèbre.

Proposition 5.6. — *Soit $\varphi \in \mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$; notons $\varphi = \sum_{n=0}^{\infty} c_n \mathrm{T}^n$. Les conditions suivantes sont équivalentes :*

(1) *La série φ est inversible dans $\mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$;*
(2) *Pour tout entier $n > 0$, on a $|c_n|\mathrm{R}^n < |c_0|$;*
(3) *On a $\|\varphi - \varphi(0)\|_{\mathrm{R}} < |\varphi(0)|$.*

Démonstration. — Supposons la condition (3) satisfaite et démontrons que φ est inversible. Quitte à remplacer φ par $\varphi/\varphi(0)$, on se ramène au cas où $\varphi(0) = 1$ et où $\|1 - \varphi\|_{\mathrm{R}} < 1$. On peut alors considérer la série convergente $\sum (1-\varphi)^n$ dans l'algèbre complète $\mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$, et sa limite ψ vérifie $(1-\varphi)\psi = \psi - 1$, c'est donc l'inverse de φ.

Supposons maintenant que φ soit inversible dans $\mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$ et soit ψ son inverse. Posons $\varphi = \sum_{n=0}^{\infty} a_n \mathrm{T}^n$, $\psi = \sum_{n=0}^{\infty} b_n \mathrm{T}^n$ et $\varphi\psi = \sum_{n=0}^{\infty} c_n \mathrm{T}^n$ (on a donc $c_0 = 1$ et $c_n = 0$ si $n > 0$). Soit m le plus grand entier tel que $|a_m|\mathrm{R}^m = \|\varphi\|_{\mathrm{R}}$ et soit n le plus grand entier tel que $|b_n|\mathrm{R}^n = \|\psi\|_{\mathrm{R}}$. Dans l'expression $c_{n+m} = \sum_{k=0}^{m+n} a_{m+n-k} b_k$, le terme $a_m b_n$ est le seul de valeur absolue maximale, si bien que $|c_{n+m}| = |a_m||b_n|$. Puisque $c_k = 0$ pour $k > 0$, cela entraîne $m = n = 0$. On a donc $|a_k|\mathrm{R}^k < |a_0|$ pour tout entier $k > 0$, ce qui prouve (2).

Supposons enfin que l'assertion (2) est vérifiée. Comme la suite $(|a_k|\mathrm{R}^k)$ tend vers 0, il existe un plus entier $n > 0$ tel que $|a_n|\mathrm{R}^n = \sup_{k>0} |a_k|\mathrm{R}^k = \|\varphi - a_0\|_{\mathrm{R}}$. Puisque $|a_n|\mathrm{R}^n < |a_0|$ par hypothèse et $a_0 = \varphi(0)$, il vient donc $\|\varphi - \varphi(0)\|_{\mathrm{R}} < |\varphi(0)|$. □

5.7. — La théorie des polygones de Newton admet une variante pour les fonctions analytiques d'une variable. Soit $\varphi \in \mathrm{F}[[\mathrm{T}]]$; notons-la $\varphi = \sum c_n \mathrm{T}^n$.

Pour tout $x \in \mathbf{R}$, on pose

$$\tau_\varphi(x) = \log(\|\varphi\|_{e^x}) = \sup_n \log(|c_n|) + nx,$$

formule dans laquelle on convient d'exclure les entiers n tels que $c_n = 0$, ou bien l'on pose $\log(|0|) = -\infty$… Si le rayon de convergence de φ est nul, alors τ_φ est constante, égale à $+\infty$. Si $\varphi = 0$, alors τ_φ est constante, égale à $-\infty$. Dans la suite de la discussion, on exclut implicitement ces deux cas triviaux.

Soit $d = \mathrm{ord}(\varphi)$ le plus petit entier naturel tel que $c_d \neq 0$. Lorsque x tend vers $-\infty$, on a $\tau_\varphi(x) = \log(|c_d|) + dx$.

Lorsque $|c_n|e^{rn}$ tend vers 0, en particulier lorsque $r < \log(\mathrm{R}_\varphi)$, ou bien lorsque $r \leqslant \log(\mathrm{R})$ et $\varphi \in \mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$, il existe un plus grand entier N tel que $\log(|c_n|) + nr = \tau_\varphi(r)$. Alors, il suffit de considérer les entiers $n \leqslant \mathrm{N}$ dans la borne supérieure qui définit $\tau_\varphi(r)$, si bien que ν_φ est une fonction convexe, affine par morceaux, sur l'intervalle $]-\infty; r]$. La fonction τ_φ est croissante, semi-continue inférieurement, continue à gauche, convexe sur $\mathbf{R}$, à valeurs dans $\mathbf{R} \cup \{+\infty\}$.

Soit $x \in \mathbf{R}$ et soit $\mathrm{R} = e^x$; on suppose que $\varphi \in \mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$. Pour tout $a \in \mathrm{F}$ tel que $|a| = \mathrm{R}$, on peut donc évaluer $\varphi(a)$. Par l'inégalité ultramétrique, on a $\log(|\varphi(a)|) \leqslant \tau_\varphi(a)$.

Comme $|c_n|\mathrm{R}^n$ tend vers 0 et $x \leqslant \log(\mathrm{R})$, l'ensemble N des entiers n tels que $\log(|c_n|) + nx = \tau_\varphi(x)$ est fini et non vide. S'il est réduit à un seul élément n, on a alors $|\varphi(a)| = |c_n|\mathrm{R}^n$, c'est-à-dire $\log(|\varphi(a)|) = \tau_\varphi(a)$; en particulier, $\varphi(a) \neq 0$. Comme dans le cas des polynômes, les valeurs absolues des zéros de φ sont donc liées aux points anguleux de τ_φ.

5.8. — On définit le polygone de Newton ν_φ de φ comme la transformée de Legendre de τ_φ : pour $t \in \mathbf{R}$, on pose

$$\nu_\varphi(t) = \sup_x \big(xt - \tau_\varphi(x)\big).$$

Bien entendu, dans cette borne supérieure, il suffit de se borner aux nombres réels x tels que $x < \log(\mathrm{R}_\varphi)$.

Soit $d = \mathrm{ord}(\varphi)$. Pour x tendant vers $-\infty$, on a $\tau_\varphi(x) = xd + \log(|c_d|)$, donc $\nu_\varphi(t) \geqslant -(d-t)x - \log(|c_d|)$. Ainsi, on a $\nu_\varphi(t) = +\infty$ si $t < d$. D'autre part, si on écrit $\varphi = \mathrm{T}^d\varphi_1$, on a $\tau_\varphi = dx + \tau_{\varphi_1}$, si bien que $\nu_\varphi(t) = \nu_{\varphi_1}(t-d)$. On pourra ainsi se ramener au cas où $\mathrm{ord}(\varphi) = 0$, c'est-à-dire $\varphi(0) \neq 0$.

Si φ est un polynôme de degré n, on a de même $\tau_\varphi(x) = nx + \log(|c_n|)$ pour x tendant vers $+\infty$, donc $\nu_\varphi(t) = +\infty$ si $t > n$. On supposera souvent implicitement dans la suite que φ n'est pas un polynôme.

Par dualité en théorie des fonctions convexes, ν_φ est la plus grande fonction convexe semi-continue inférieurement sur $\mathbf{R}$ (à valeurs dans $\mathbf{R} \cup \{+\infty\}$) telle que $\nu_\varphi(m) \leqslant -\log(|c_m|)$ pour tout $m \in \mathbf{N}$ tel que $c_m \neq 0$. Son graphe est le bord de l'enveloppe convexe supérieure des points $(m, -\log(|c_m|))$.

Cela donne un moyen pratique de la construire dans des cas concrets. On part en effet de l'égalité $v_\varphi(0) = -\log(|c_0|)$ qui sera justifiée ci-dessous. Posons $p = \inf_{n>0}(-\log(|c_n|) + \log(|c_0|))/n$; la fonction $h\colon t \mapsto -\log(|c_0|) + pt$ est la plus grande fonction affine qui coïncide avec v_φ en 0 et minore v_φ. Soit N l'ensemble des entiers $n > 0$ tels que la borne supérieure précédente soit atteinte en n. S'il est vide ou s'il est infini, alors $v_\varphi = h$. Supposons que N soit non vide et fini et soit n son plus grand élément. La fonction $v_\varphi - h$ sur $[0; n]$ est convexe, semi-continue inférieurement, positive, et vaut 0 en 0 et n ; elle est donc nulle, si bien que v_φ coïncide avec h sur l'intervalle $[0; n]$. On reprend alors l'argument précédent considérant la plus grande fonction affine de pente $\geqslant p$ qui minore v_φ et coïncide avec v_φ en n, etc.

Exemple 5.9. — Supposons que F soit le corps des nombres rationnels, muni de la valeur absolue p-adique. Pour simplifier les notations, on suppose que le logarithme est pris en base p, c'est-à-dire que $\log(p) = 1$.

(1) Considérons le cas où $\varphi = e^{\mathrm{T}} = \sum_{n=0}^{\infty} \mathrm{T}^n/n!$.

On a vu la relation $-\log(1/|n!|) = (\sigma(n) - n)/(p-1)$, où $\sigma(n)$ est la somme des chiffres du développement en base p de n. Par suite, la fonction affine $t \mapsto -t/(p-1)$ est inférieure ou égale à $-\log(|1/n!|)$ en n ; on a donc $v_\varphi(t) \geqslant -t/(p-1)$ pour tout $t \in \mathbf{R}_+$. Comme $v_\varphi(0) \leqslant -\log(|1|) = 0$, on a aussi $v_\varphi(0) = 0$. Prenons $n = p^m$; on a $\sigma(n) = 1$, donc $v_\varphi(n) \leqslant (1-n)/(p-1)$. Par convexité, il vient $v_\varphi(t) \leqslant t/n(p-1) - t/(p-1)$ si $t \in [0; n]$. Faisons tendre m (donc n) vers $+\infty$; il en résulte l'inégalité $v_\varphi(t) \leqslant -t/(p-1)$. Ainsi, le polygone de Newton de φ est donné par $t \mapsto -t/(p-1)$.

(2) Prenons maintenant $\varphi = \log(1 + \mathrm{T}) = \sum_{n=1}^{\infty}(-1)^{n-1}\mathrm{T}^n/n$. Puisque $\operatorname{ord}(\varphi) = 1$, son polygone de Newton est infini entre 0 et 1. Démontrons que pour tout entier $m \geqslant 0$, on a $v_\varphi(p^m) = -m$ et que la restriction de v_φ à l'intervalle $[p^m; p^{m+1}]$ est affine.

Notons h la fonction affine par morceaux ainsi définie ; il est immédiat qu'elle est décroissante et convexe. pour tout $m \in \mathbf{N}$, on a $h(p^m) = -m = -\log(|1/p^m|)$, donc $h(p^m) \leqslant v_\varphi(p^m)$. Soit $n \in \mathbf{N}$ tel que $n \geqslant 1$; soit m l'unique entier naturel tel que $p^m \leqslant n < p^{m+1}$; puisque $n < p^{m+1}$, on a $v_p(n) \leqslant m$, donc $-\log(|1/n|) \geqslant -m = h(p^m)$; a fortiori, $-\log(|1/n|) \geqslant -(m+1) = h(p^{m+1})$. Sur l'intervalle $[p^m; p^{m+1}]$, la fonction h est affine, inférieure ou égale à v_φ, et coïncide avec v_φ, aux extrémités. Puisque v_φ est convexe, ces deux fonctions coïncident sur cet intervalle.

Théorème 5.10. — *Soit* R *un nombre réel strictement positif et soit* $\varphi \in \mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$ *une série formelle restreinte. On suppose que* φ *n'est pas un polynôme et que* $\varphi(0) \neq 0$.

(1) *La fonction* v_φ *est définie sur* $\mathbf{R}_+$ *; elle est continue, convexe et affine par morceaux.*

(2) *Les nombres réels* t *tels que* v_φ *soit finie et non dérivable en* t *sont des entiers tels que* $v_\varphi(t) = -\log(|c_t|)$. *Soit* $(n_0, n_1, \dots)$ *la suite strictement croissante de ces entiers. On a* $n_0 = 0$.

(3) *Soit* $a \in \mathrm{F}$ *tel que* $|a| \leqslant \mathrm{R}$ *et* $\varphi(a) = 0$*; alors* $\log(|a|)$ *est une pente de* v_φ.

(4) *Soit* m *un entier tel que* n_m *est défini et soit* μ *la pente de* v_φ *sur l'intervalle* $[n_{m-1}; n_m]$. *On a* $\mu < \log(\mathrm{R}_\varphi)$. *Si, de plus,* $\mu \leqslant \log(\mathrm{R})$, *il existe un unique couple* (P, ψ) *où* $\mathrm{P} \in \mathrm{F}[\mathrm{T}]$ *est un polynôme de degré* $n_m - n_{m-1}$ *tel que* $\mathrm{P}(0) = 1$ *et dont le polygone de Newton n'a qu'une pente, égale à* μ, *et où* $\psi \in \mathrm{F}\langle \mathrm{T}/\mathrm{R}\rangle$ *est telle que* $\varphi = \mathrm{P}\psi$. *De plus,* μ *n'est pas une pente de* v_ψ.

Démonstration. — (1) Soit n un entier naturel tel que $c_n \neq 0$; rappelons aussi que $c_0 \neq 0$, par hypothèse. Toute fonction convexe qui est majorée par $-\log(|c_0|)$ en 0 et par $-\log(|c_n|)$ en n est majorée sur l'intervalle $[0; n]$, de sorte que v_φ est finie sur cet intervalle. Puisque φ n'est pas un polynôme, il existe des entiers n arbitrairement grands tels que $c_n \neq 0$. Cela prouve que v_φ est finie sur $\mathbf{R}_+$. Étant convexe et semi-continue inférieurement, elle est donc continue sur cet intervalle.

(2) Soit h l'unique fonction affine par morceaux qui est affine sur chaque intervalle de la forme $[n; n+1]$ et coïncide avec v_φ sur cet intervalle. Elle est convexe et continue, et vérifie $v_\varphi \leqslant h$. Par construction, on a $h(n) \leqslant -\log(|c_n|)$ pour tout n ; par suite, $h = v_\varphi$. Cela prouve que v_φ est affine sur chaque intervalle de la forme $[n; n+1]$.

La fonction h qui coïncide avec v_φ sur $[1; +\infty]$, est affine sur $[0; 1]$ et vaut $-\log(|c_0|)$ en 0 est convexe, semi-continue inférieurement, et majore v_φ ; elle coïncide donc avec v_φ. Par suite, $v_\varphi(0) = -\log(|c_0|)$.

Soit n un entier tel que $n \geqslant 1$ et $v_\varphi(n) < -\log(|c_n|)$; supposons que les dérivées à droite et à gauche de v_φ en n soient distinctes. Pour $\varepsilon > 0$ assez petit, l'unique fonction h sur $\mathbf{R}_+$ qui est affine sur chaque intervalle de la forme $[m; m+1]$, coïncide avec v_φ hors de $[n-1; n+1]$ et qui vaut $v_\varphi(n) + \varepsilon$ en n est encore convexe, continue, et vérifie $h(m) \leqslant -\log(|c_m|)$ pour tout $m \in \mathbf{N}$. Cela contredit la définition de v_φ.

(3) Soit $x \in \mathbf{R}$ tel que $x \leqslant \log(\mathrm{R})$. Pour tout $a \in \mathrm{F}$ tel que $|a| = e^x$, on peut donc évaluer $\varphi(a)$. Par l'inégalité ultramétrique, on a $\log(|\varphi(a)|) \leqslant \tau_\varphi(a)$. Supposons $\varphi(a) = 0$. Comme $|c_n|\mathrm{R}^n$ tend vers 0 et $x \leqslant \log(\mathrm{R})$, l'ensemble N des entiers n tels que $\log(|c_n|) + nx = \tau_\varphi(x)$ est fini et non vide. Pour $n \in \mathrm{N}$, on a

$$v_\varphi(n) \geqslant nx - \tau_\varphi(x) = -\log(|c_n|),$$

de sorte que $v_\varphi(n) = -\log(|c_n|)$.

Si l'ensemble N est réduit à un seul élément n, on a alors $|\varphi(a)| = |c_n|e^{nx}$, c'est-à-dire $\log(|\varphi(a)|) = \tau_\varphi(a)$, ce qui contredit l'hypothèse $\varphi(a) = 0$.

Notons alors n et n' le plus petit et le plus grand élément de N, de sorte que $n < n'$, et soit h l'unique fonction affine qui coïncide avec v_φ en n et en n'. Démontrons que v_φ et h coïncident sur l'intervalle $[n; n']$. Par convexité, on a $v_\varphi \leqslant h$ sur cet intervalle et $v_\varphi \geqslant h$ en dehors. Il suffit donc de prouver que $v_\varphi = \sup(h, v_\varphi)$ et donc que l'on a $h(m) \leqslant -\log(|c_m|)$ pour tout $m \in \mathbf{N}$; on peut pour cela supposer que $m \in [n; n']$. Écrivons $m = (1-u)n + un'$, où $u \in [0; 1]$. L'inégalité $\log(|c_m|) + mx \leqslant \tau_\varphi(x)$ se récrit

$$\begin{aligned}\log(|c_m|) &\leqslant -mx + \tau_\varphi(x)\\ &= (1-u)(-nx + \tau_\varphi(n)) + u(-n'x + \tau_\varphi(n'))\\ &= (1-u)(-h(n)) + u(-h(n')) = -h(m),\end{aligned}$$

de sorte que $h(m) \leqslant -\log(|c_m|)$, ainsi qu'il fallait démontrer.

Ainsi, v_φ est affine sur l'intervalle $[n; n']$, de pente

$$\frac{v_\varphi(n') - v_\varphi(n)}{n' - n} = \frac{(n'x - \tau_\varphi(x)) - (nx - \tau_\varphi(x))}{n' - n} = x.$$

Cela prouve que $x = \log(|a|)$ est une pente de v_φ.

(4) Par hypothèse, les pentes de v_φ sur l'intervalle $[n_m; +\infty[$ sont strictement supérieures à μ ; il existe donc un nombre réel $\mu' > \mu$ tel que

$$v_\varphi(t) \geqslant v_\varphi(n_m) + \mu'(t - n_m)$$

pour tout $t \geqslant n_m$. En particulier, pour tout entier $n \geqslant n_m$, on a

$$-\log(|c_n|) \geqslant v_\varphi(n) \geqslant v_\varphi(n_m) + \mu'(n - n_m),$$

si bien que la suite $(|c_n|e^{n\mu'})$ est bornée, si bien que $\mu < \mu' \leqslant \log(\mathrm{R}_\varphi)$.

Supposons $\mu \leqslant \log(\mathrm{R})$. Pour tout entier N, posons $\varphi_\mathrm{N} = \sum_{n=0}^{\mathrm{N}} c_n \mathrm{T}^n$. Supposons $\mathrm{N} > n_j$. Alors, la restriction du polygone de Newton

de φ_N à l'intervalle $[0; n_m]$ coïncide avec celui de φ. Sa restriction à l'intervalle $[n_m; N]$ ne coïncide pas forcément avec celui de φ, c'est d'ailleurs ce qu'indique la confrontation des exemples 4.11 et 5.9, mais le polygone de Newton de φ_N domine celui de φ sur cet intervalle.

On déduit du théorème 4.8 qu'il existe un unique polynôme $P_N \in F[T]$ de degré $n_m - n_{m-1}$ tel que $P_N(0) = 1$, dont le polygone de Newton est affine de pente μ, et qui divise φ_N dans $F[T]$. Posons $\psi_N = \varphi_N/P_N$. Lorsque $N \to +\infty$, les polynômes P_N convergent dans $F[T]$ vers un polynôme P dont le polygone de Newton est affine, de pente μ, les polynômes ψ_N convergent dans $F\langle T/R\rangle$ vers une série formelle ψ et l'on a $\varphi = P\psi$. Par ailleurs, les pentes de ψ_N qui sont $< \mu$ sont celles de φ, et les autres sont $\geqslant \mu'$. En passant à la limite, on voit que μ n'est pas une pente de ψ. □

Proposition 5.11. — *Soit $\varphi \in F\langle T/R\rangle$, soit $P \in F[T]$ un polynôme non nul; en notant $P = \sum_{k=0}^{d} c_k T^k$, on suppose que $\|P\|_R = |c_d| R^d$. Il existe un unique couple (ρ, ψ), où $\rho \in F[T]$ est un polynôme de degré $< d$ et $\psi \in F\langle T/R\rangle$, tel que $\varphi = P\psi + \rho$. De plus, on a $\|\rho\|_R \leqslant \|\varphi\|_R$ et $\|P\|_R \|\psi\|_R \leqslant \|\varphi\|_R$.*

Démonstration. — Avant d'établir l'existence et l'unicité, démontrons les inégalités qui concluent cet énoncé. Considérons une relation $\varphi = \psi P + \rho$, où $\rho \in F[T]$ est un polynôme de degré $< d$ et $\psi \in F\langle T/R\rangle$. Les inégalités à démontrer sont évidentes lorsque $\psi = 0$; supposons donc $\psi \neq 0$, et considérons le plus grand entier, n, tel que $\|\psi\|_R$ soit donné par le coefficient de ψ d'exposant n; alors $\|\varphi - \rho\|_R = \|\psi P\|_R$ est donné par le coefficient d'exposant $n + d$. Puisque $\deg(\rho) < d$, on a donc $\|\varphi - \rho\|_R \leqslant \|\varphi\|_R$, d'où $\|\psi\|_R \|P\|_R \leqslant \|\varphi\|_R$. Puis $\|\rho\|_R = \|\varphi - \psi P\| \leqslant \sup(\|\varphi\|_R, \|\psi P\|_R) \leqslant \|\varphi\|_R$.

Établissons maintenant l'unicité d'une décomposition comme dans l'énoncé. Si (ρ, ψ) et (ρ', ψ') sont deux tels couples, on a $0 = (\psi' - \psi)P + (\rho' - \rho)$. Les inégalités précédentes entraînent $\psi' - \psi = 0$ puis $\rho' - \rho = 0$.

Démontrons maintenant l'existence d'un tel couple (ρ, ψ). Pour tout entier N, notons φ_N l'unique polynôme de $F[T]$ de degré $\leqslant N$ tel que $\operatorname{ord}(\varphi - \varphi_N) > N$. Considérons alors la division euclidienne de φ_N par P, disons $\varphi_N = Q_N P + R_N$. Pour $N \geqslant M > d$, on a donc

$$\varphi_N - \varphi_M = (Q_N - Q_M)P + (R_N - R_M).$$

Par conséquent, on a

$$\|Q_N - Q_M\|_R \|P\|_R, \|R_N - R_M\| \leqslant \|\varphi_N - \varphi_M\|.$$

Lorsque M, N tendent vers l'infini, $\|\varphi_M - \varphi_N\|_R$ tend vers 0, donc $\|Q_N - Q_M\|_R$ tend vers 0, de même que $\|R_M - R_N\|_R$. Par suite, la suite (R_N) converge vers un polynôme ρ de degré $< d$, et la suite (Q_N) converge vers un élément ψ de $F\langle T/R\rangle$. En passant à la limite, on a $\varphi = \psi Q + \rho$. □

***Proposition 5.12* (Théorème de « préparation » de Weierstrass)**
Soit $\varphi \in F\langle T/R\rangle$, notons $\varphi = \sum_{n=0}^{\infty} c_n T^n$, et soit N le plus grand entier n tel que $|c_N| R^N = \|\varphi\|_R$. Il existe un unique couple (P, ψ) tel que $\varphi = \psi P$, où P est polynôme unitaire de degré N dans $F[T]$ et ψ est un élément inversible $\psi \in F\langle T/R\rangle$.

Démonstration. — C'est une variante de l'argument fait pour isoler la partie de pente donnée dans le polygone de Newton. Pour simplifier les notations, on suppose $c_N = 1$. Pour tout entier M, soit $\varphi_M = \sum_{n=0}^{M} c_n T^n$. Si $M > N$, le point $(N, -\log(|c_N|))$ est un sommet du polygone de Newton du polynôme φ_M et il existe un unique polynôme unitaire $P_M \in F[T]$ de degré N divisant φ_M et dont le polygone de Newton coïncide avec celui de φ sur l'intervalle $[0; N]$. Posons $\psi_M = \varphi_M / P_M$. Lorsque l'entier M tend vers l'infini, les suites (P_M) et (ψ_M) convergent vers un polynôme unitaire $P \in F[T]$ et vers une série formelle $\psi \in F\langle T/R\rangle$. Le polygone de Newton de P coïncide avec celui de φ sur $[0; N]$. En particulier, si l'on note $P = \sum_{n=0}^{N} a_n T^n$, on a $|a_n| R^n \leqslant |a_N| R^n$ pour tout $n \in \{0, \dots, N\}$. Notons $\psi = \sum_{n=0}^{\infty} b_n T^n$. Si n est le plus grand entier tel que $|b_n| R^n = \|\psi\|_R$, on a $|c_{N+n}| R^{N+n} = \|\psi\|_R \|P\|_R = \|\varphi\|_R$, de sorte que $n = 0$. Compte tenu de la proposition 5.6, cela prouve que ψ est inversible dans $F\langle T/R\rangle$. □

***Corollaire 5.13* (Strassmann, 1928).** — *Un élément non nul de $F\langle T/R\rangle$ n'a qu'un nombre fini de zéros dans $B(0; R)$.*

Démonstration. — Soit φ un élément non nul de $F\langle T/R\rangle$. Avec les notations de la proposition, les zéros de φ sont les zéros de P, car un élément inversible de $F\langle T/R\rangle$ ne s'annule pas sur $B(0; R)$. □

Remarque 5.14. — Soit φ un élément non nul de $F\langle T/R\rangle$. On démontre comme en analyse complexe que les zéros de φ dans la boule circonférenciée $B(0;R)$ sont isolés. Si F est localement compact (et sa valeur absolue non triviale), cette boule $B(0;R)$ est compacte, si bien que l'ensemble des zéros de φ dans $B(0;R)$ est fini. Cette conclusion n'est cependant pas évidente dans le cas général et on peut voir ce théorème de Strassmann comme une trace de la compacité du « disque analytique » de rayon R tel que défini, par exemple, dans la théorie de Berkovich (1990).

Références

E. Artin & O. Schreier (1927), « Eine Kennzeichnung der reell abgeschlossenen Körper ». *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, **5** (1), p. 225–231.

V. G. Berkovich (1990), *Spectral Theory and Analytic Geometry over Non-Archimedean Fields*, Mathematical Surveys and Monographs **33**, American Mathematical Society, Providence, RI.

C. Christensen (1996), « Newton's method for resolving affected equations ». *Coll. Math. J.*, **27** (5), p. 330–340.

R. F. Coleman (1987), « On the Galois groups of the exponential Taylor polynomials ». *Enseign. Math. (2)*, **33** (3-4), p. 183–189.

S. Gao (2001), « Absolute Irreducibility of Polynomials via Newton Polytopes ». *Journal of Algebra*, **237** (2), p. 501–520.

K. Hensel (1901), « Ueber die Entwickelung der algebraischen Zahlen in Potenzreihen ». *Math. Ann.*, **55** (2), p. 301–336.

K. Hensel (1904), « Neue Grundlagen der Arithmetik ». *Journal für die reine und angewandte Mathematik*, **1904** (127), p. 51–84.

N. Kollerstrom (1992), « Thomas Simpson and 'Newton's method of approximation' : An enduring myth ». *The British Journal for the History of Science*, **25** (3), p. 347–354.

J. Kürschák (1913), « Über Limesbildung und allgemeine Körpertheorie. » *Journal für die reine und angewandte Mathematik*, **1913** (142), p. 211–253.

P. Monsky (1970), « On Dividing a Square Into Triangles ». *The American Mathematical Monthly*, **77** (2), p. 161. *2317329*.

A. Ostrowski (1916), « Über einige Lösungen der Funktionalgleichung $\varphi(x)\cdot\varphi(y)=\varphi(xy)$ ». *Acta Math.*, **41** (1), p. 271–284.

I. Schur (1930), « Gleichungen ohne Affekt ». *Sitzungsberichte Preuss. Akad. Wiss. Phys.-Math. Klasse*, p. 443–449.

R. Strassmann (1928), « Über den Wertevorrat von Potenzreihen im Gebiet der p-adischen Zahlen. » *Journal für die reine und angewandte Mathematik*, **1928** (159), p. 13–28.

H. W. Turnbull (éditeur) (1960), *The correspondence of Issac Newton, 1676–1687*, **2**, Published for the Royal Society at the University Press, Cambridge [Eng.].