

# Astérisque

ALEXIS MARIN

## **Géométrie des polynômes. Coût global moyen de la méthode de Newton**

*Astérisque*, tome 152-153 (1987), Séminaire Bourbaki, exp. n° 670, p. 19-36

[http://www.numdam.org/item?id=SB\\_1986-1987\\_\\_29\\_\\_19\\_0](http://www.numdam.org/item?id=SB_1986-1987__29__19_0)

© Société mathématique de France, 1987, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (<http://smf4.emath.fr/Publications/Asterisque/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

## GÉOMÉTRIE DES POLYNOMES COUT GLOBAL MOYEN DE LA MÉTHODE DE NEWTON

(d'après M. Shub et S. Smale)

par Alexis MARIN

La suite  $1, 3/2, 17/12, \dots$  définie par  $x_0 = 1$  et la relation de récurrence  $x_{n+1} = \frac{1}{2}(x_n + 2/x_n)$  converge très rapidement vers  $\sqrt{2}$ . Le troisième terme  $x_3 = 577/408 \sim 1,414215\dots$  possède déjà les six premiers chiffres du développement décimal de  $\sqrt{2} \sim 1,414213\dots$  alors que  $x_2 = 17/12 \sim 1,416\dots$  n'en avait que trois. Cette suite, connue des Babyloniens, est produite par la *méthode de Newton*. Pour approcher une solution d'une équation  $f(x) = 0$ , où  $f$  est de classe  $C^2$ , Newton prend une *valeur initiale*  $x_0$  et, itérant l'application  $x \mapsto N_f(x) = x - f(x)/f'(x)$ , il engendre la suite  $x_n = N_f^n(x_0)$ . (Notez que  $N_f(x) = \frac{1}{2}(x + 2/x)$  si  $f(x) = x^2 - 2$ .) Les points fixes de  $N = N_f$  sont les zéros de  $f$ , et puisque  $N' = f f'' / (f')^2$ , un zéro simple  $w$  de  $f$  (i.e.  $f(w) = 0 \neq f'(w)$ ) est un *point fixe superattractif* de  $N$  (i.e.  $N(w) = w$  et  $N'(w) = 0$ ), en particulier il y a une constante  $K$  et un voisinage  $V$  de  $w$  dont tout point  $x$  vérifie  $|N(x) - w| \leq K |x - w|^2$ . Cette *convergence quadratique* explique le "doublement de précision" à chaque itération observé ci-dessus.

Cependant, si  $x_0$  n'est pas suffisamment proche d'une racine de  $f$ , la suite de Newton peut avoir un comportement chaotique ou s'accumuler sur un cycle, ce dernier phénomène se produisant dans un ouvert de l'espace  $P_d$  des polynômes de degré  $d$ , dès le degré 3, et ce pour un ouvert de valeurs initiales. Pour trouver une bonne valeur initiale, on peut explorer systématiquement le domaine de  $f$ , mais le coût d'un tel balayage est prohibitif; on préfère *tirer au sort* une valeur initiale, et itérer quatre ou cinq fois l'application de Newton; avec un peu de chance, la valeur de  $f$  sur le dernier itéré est pratiquement nulle, sinon on relance les dés... L'expérience confirme que, sauf pour les polynômes pathologiques et *rare*s, les joueurs de dés sont gagnants.

Pour justifier cette pratique, il faut munir les espaces de polynômes et de valeurs initiales d'une mesure de probabilité. Smale considère

l'espace

$$P_d(1) = \{p(v) = v^d + a_{d-1}v^{d-1} + \dots + a_0 \mid a_i \in \mathbf{C}, |a_i| \leq 1\}$$

muni de la mesure de Lebesgue  $\mu$  sur les coefficients, normalisée par  $\mu(P_d(1)) = 1$ . On ne perd rien à se limiter à  $P_d(1)$  car, pour tout polynôme  $q$  de degré  $d > 1$ , il existe  $a$  et  $b$  positifs avec  $p(v) = aq(bv)$  dans  $P_d(1)$ ; d'autre part (voir 5.4), les racines de  $p \in P_d(1)$  sont de module inférieur à 2. Connaissant  $p(v_0)$  et  $p'(v_0)$ , Smale a un critère pour savoir si, partant de  $v_0$ , la méthode de Newton converge quadratiquement ([Sm5], cf. §6). Il vient de montrer ([Sm6]) que, si pour un polynôme  $p$  de  $P_d(1)$  la moyenne du nombre de tirages dans le disque  $\{|v| \leq 3\}$  pour obtenir une valeur initiale  $v_0$  vérifiant son critère est supérieur à  $n$  alors  $p$  se trouve dans un ensemble exceptionnel  $\mathcal{E}_{n,d}$  de mesure bornée par (constante) $(d^5/n)$ . Dans [ShSm1], il y avait un résultat dans le même sens, mais sans critère permettant de savoir si on a gagné (cf. 2.3).

Dès 1976, lors de ses travaux d'économie mathématique ([Sm1]), puis avec Hirsch ([HiSm]), Smale avait introduit des *méthodes de Newton globales* (cf. §§3 et 5) n'ayant qu'une convergence linéaire mais dont le coût global peut être estimé plus finement. En 1981, il dégagait une jolie démonstration du théorème de d'Alembert-Gauss ([Sm2] et §3) et prouvait le résultat suivant: *Pour tout  $0 < \mu < 1$ , il y a un ensemble exceptionnel  $U_d(\mu)$  de mesure  $\mu$  dans  $P_d(1)$  tel que, pour tout  $p$  hors de  $U_d(\mu)$ , une méthode de Newton globale partant de  $v_0 = 0$  fournit en moins de  $(100(d+2))^9/\mu^7$  itérations une valeur à partir de laquelle le "doublement de précision" à chaque itération de Newton se produit.* En 1983 avec Shub ([ShSm1] et [ShSm2]), il introduit des méthodes d'ordre supérieur (dont chaque itération nécessite le calcul de  $\log d$  dérivées de  $p$  si  $p \in P_d(1)$ ). Partant d'un point  $v_0$  pris au hasard sur un cercle de grand rayon, Shub et Smale obtiennent un bon zéro approché en  $N_d(\mu)$  itérations où  $N_d(\mu)$  croît linéairement en  $d$  :

$$N_d(\mu) = L_1 d(|\log \mu|/\mu)^{1+1/\log d} + L_2 \quad \text{avec} \quad 20 > L_1 > L_2 \quad ,$$

et ils estiment le nombre d'opérations arithmétiques nécessaires. En 1984, dans un survol ([Sm4]) où, *sans prétendre produire des algorithmes rivalisant avec ceux couramment utilisés*, il cherche à expliquer pourquoi *les algorithmes efficaces de l'analyse sont rapides en moyenne* bien qu'ils piétinent sur des cas dégénérés, et il décrit une méthode globale du premier ordre qui permet de retrouver avec de meilleures bornes la plupart des résultats de [ShSm1] et [ShSm2] en n'ayant besoin à chaque itération de n'évaluer que  $p$  et  $p'$  (cf. §4).

Le grand avantage de la méthode de Newton sur les autres procédés de résolution approchée d'une équation d'une variable réelle (dichotomie, sécante, ...) est qu'elle s'applique en toute dimension et

même dans un Banach, ce qui permet de traiter des équations fonctionnelles ([KA], chap. 18). Renegar a étendu aux systèmes de  $n$  équations polynomiales en  $n$  inconnues complexes les résultats de Shub-Smale [R1], et [Sm6] contient des résultats valables dans un Banach.

Dans le cas d'un polynôme d'une variable complexe, l'application de Newton a une extension rationnelle  $(x, p) \mapsto N_p(x)$  à la sphère de Gauss  $S$  : La méthode de Newton est un algorithme *purement itératif rationnel* et l'on peut se demander s'il existe un autre algorithme *purement itératif rationnel*  $G : S \times P_d \rightarrow S$  qui soit de plus *génériquement convergent*, i.e.

$$\{(x, p) \in S \times P_d \mid G_p^n(x) \text{ ne converge pas vers une racine de } p\}$$

est de mesure nulle. Pour  $d > 3$ , la thèse de McMullen ([McM]) répond: "non". Ceci n'arrêta pas Shub et Smale: avec la conjugaison complexe en plus des opérations arithmétiques, ils ont construit un algorithme *purement itératif génériquement convergent* pour les systèmes de  $n$  polynômes en  $n$  variables complexes, et, pour  $n = 1$ , un tel algorithme qui, près des racines, est la méthode de Newton ([ShSm3] et §7).

D'autre part, Dantzig (cf. [Sm3]) avait conjecturé que:  $\rho(m, n)$  le nombre moyen d'itérations nécessaires pour résoudre par la méthode du simplexe un problème linéaire à  $m$  contraintes et  $n$  variables est, pour  $m$  fixé, linéaire en  $n$ . En 1982, Smale résolvait par les mêmes méthodes ce fameux problème; il donne une fonction  $K(\varepsilon, m)$  telle que, pour tout  $\varepsilon > 0$ ,  $\rho(m, n) \leq Kn^\varepsilon$  ([Sm3]). Depuis, des algorithmes polynomiaux en  $m$  et  $n$  ont été trouvés ([R2], [Sm6]).

Tous ces résultats illustrent l'intérêt de l'introduction d'idées élémentaires de topologie et de résultats généraux comme le théorème de la variété stable dans les problèmes concrets de calcul (cf. [Sh] et §7).

Dans cet exposé, je me limiterai aux polynômes d'une variable complexe.

Mes remerciements vont à J.J. Risler qui m'a convaincu d'étudier les travaux de Smale pour son séminaire, ainsi qu'à L. Guillou pour ses commentaires perspicaces sur ce texte.

## 1. Méthode de Newton, chaos, et non convergence générique.

Si un polynôme réel  $f(t) = \prod_{i=1}^d (t - x_i)$  a toutes ses racines réelles et distinctes, il en est de même de sa dérivée  $f'(t) = d \prod_{j=1}^{d-1} (t - c_j)$ , et zéros et points critiques alternent  $x_1 < c_1 < x_2 < \dots < x_d$  dans  $\mathbf{R}$ . Soit  $f$  un tel polynôme de degré  $d \geq 3$  et notons  $N = N_f$ . Pour l'orbite  $\{t_n = N^n(t)\}$  d'un point  $t$ , il y a trois possibilités:

- 1)  $t_n$  converge vers un point fixe de  $N$  (un zéro de  $f$ );

- 2)  $t_n$  n'est pas défini pour  $n > n_0$  car  $t_{n_0}$  est un point critique;
- 3)  $t_n$  est défini pour tout  $n$ , mais ne converge pas.

Soit  $S_1, S_2, S_3$  la partition correspondante de  $\mathbf{R}$ . Comme les points fixes  $x_i$  de  $N$  sont attractifs ( $|N'(x_i)| = 0 < 1$ ),  $S_1$  est ouvert. Soient  $b_1, \dots, b_d$  les bandes  $]c_{i-1}, c_i[$  (ici  $c_0 = -\infty$  et  $c_d = +\infty$ ). Un coup d'oeil à la figure 1 nous assure que  $N$  envoie des points proches des  $c_i$  dans  $b_1$  ou  $b_d$  et que  $S_1$  contient ces deux bandes ; donc  $S_2$  est disjoint de  $\bar{S}_3$ , et  $S_3$  est un fermé invariant par  $N$  inclus dans  $b_2 \cup \dots \cup b_{d-1}$ .

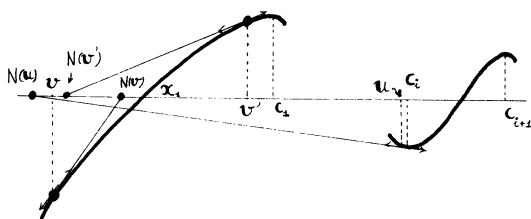


FIGURE 1

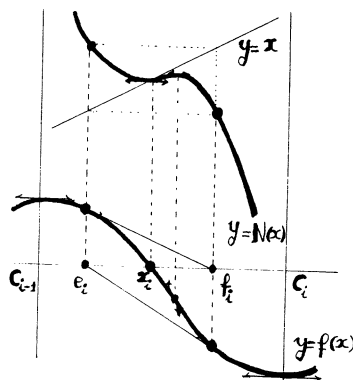


FIGURE 2

Soit  $\Omega$  l'espace des suites infinies en les symboles  $b_2, \dots, b_{d-1}$  muni de la topologie produit; le décalage  $D$  agit sur  $\Omega$ . L'application  $h : S_3 \rightarrow \Omega$  définie par  $N^n(x) \in h(b)_n$  est continue et envoie le système dynamique  $(S_3, N)$  dans  $(\Omega, D)$ .

**THÉORÈME 1.1** (Barna, Saari et Urenho [Ba], [SU]). *Si le polynôme  $f$  de degré  $d \geq 3$  a tous ses zéros réels et distincts, alors  $S_3$  est un Cantor de mesure nulle, l'application  $h$  est surjective et toute orbite périodique non constante de  $D$  se relève en une orbite de  $h$  de même période, les orbites constantes se relevant en des orbites de période 2.*

**DÉMONSTRATION:** Pour  $1 < i < d$ , soit  $\beta_i = ]e_i, f_i[$  la composante connexe de  $x_i$  dans  $S_1$ , alors  $S_3 \subset B = \bigcup \{b_i \setminus \beta_i \mid 1 < i < d\}$  et  $N$  échange  $e_i$  et  $f_i$  (cf. Figure 2).

Comme  $\lim_{t \rightarrow c_i} N'(t) = -\infty$  et  $N'(x_i) = 0$ , l'équation

$$N'(x) = -a^2 \tag{*}$$

a au moins une solution dans chacun des  $2d-2$  intervalles de  $B$  et il n'y a qu'une solution par intervalle car (\*) est de degré  $2d-2$ , donc  $N'$  est strictement monotone sur chaque intervalle sur lequel elle est négative,

en particulier croissante sur  $]c_{i-1}, e_i]$  et décroissante sur  $[f_i, c_i[$ . Par un calcul astucieux ([Ba], p.202-206), Barna obtient  $N'(e_i) < -1$  et  $N'(f_i) < -1$ ; donc  $N$  est dilatante sur  $B$ , et a fortiori sur  $S_3 \subset B$ . Comme  $S_3$  est invariant, sa mesure doit être nulle. Le reste du théorème est facile.  $\square$

Dans [SU] et [HM], on trouvera d'autres informations sur la dynamique de  $N$ . Hurley et Martin ont aussi observé que si  $T_f$  est une application  $C^2$  définie dans le complémentaire d'un ensemble fini  $J$ , tendant vers l'infini près de  $J$  et dont les points fixes sont attractifs, alors deux points fixes consécutifs sont séparés par une asymptote du graphe de  $T_f$ , et  $T_f$  présente le même chaos que  $N_f$ .

Le théorème de Barna reste vrai s'il y a des racines multiples. En ce cas, comme  $S_2$  est dénombrable, la mesure de  $\mathbf{R} \setminus S_1$  est nulle, ce qui semble donner raison aux joueurs. Cependant, pour tout  $d > 2$ , il y a des polynômes  $p$  de degré  $d$  avec  $p''(0) = 0$  et  $p(0) = -p'(0) \neq 0 \neq p(1) = p'(1)$  et donc pour lesquels  $\{0, 1\}$  est un cycle superattractif d'ordre 2 de  $N$  (par exemple (†) le polynôme  $p(z) = \frac{1}{2}z^3 - z + 1$ ). Tout polynôme  $q$  proche de  $p$  possède un cycle attractif d'ordre 2. Il s'en suit que *la méthode de Newton n'est pas un algorithme génériquement convergent*.

La dynamique holomorphe donne d'autres informations sur les orbites de  $N$  (voir [Sm4], [DH] et [F]).

## 2. Les méthodes d'Euler-Newton, et un critère au but de convergence.

Soient  $p : \mathbf{C} \rightarrow \mathbf{C}$  un polynôme et  $v$  un point régulier de  $p$  (i.e.  $p'(v) \neq 0$ ); il y a, défini au voisinage de  $z = p(v)$ , un unique inverse à droite de  $p$  envoyant  $z$  sur  $v$ , notons le  $p_v^{-1}$ , de rayon de convergence  $r_v$  en  $v$ . Notons  $D_v = D(z, r_v)$ . Si  $r_v > |z|$ , Euler calcule une racine  $w$  de  $p$  grâce au développement de Taylor de  $p_v^{-1}$  en  $z$  :

$$w = p_v^{-1}(0) = p_v^{-1}(z - z) = \sum_{n=0}^{\infty} \frac{1}{n!} \left\{ \frac{d^n}{dz^n} p_v^{-1} \right\}_z (-z)^n \quad (2.1)$$

Le calcul des termes de (2.1) coûtant de plus en plus cher, Euler préfère tronquer cette série à l'ordre  $k$  et itérer l'application  $v \mapsto E_k(v)$  ainsi obtenue. Remarquons que  $E_1(v) = v - p(v)/p'(v)$  est l'application de Newton.

---

(†) D'après le théorème de Barna, un tel polynôme a des racines complexes.

**PROPOSITION 2.1** ([ShSm1], [Sm4]). *Pour chaque  $k$ , il y a une constante  $c_k < 1$  telle que si  $w$  est une racine simple de  $p$  et  $v_0 \in p_w^{-1}(c_k D_w)$ , alors pour tout  $n$  l'itéré  $v_n = E_k^n(v_0)$  est dans  $p_w^{-1}(D_w)$ , la suite  $v_n$  converge vers  $w$ , et*

$$|p(v_n)| \leq b^{(k+1)^n - 1} |p(v_0)| \quad , \quad \text{où} \quad b = |p(v_0)| / c_k r_w < 1 \quad .$$

*La suite  $c_k$  est croissante, tend vers  $(2 - \sqrt{2})/4$  quand  $k \rightarrow \infty$ , et*

$$c_1 = 1/9 < c_2 < 1/8 < c_3 < 1/7 < c_4 < (2 - \sqrt{2})/4 < 1/6 \quad .$$

Un *bon zéro approché* (pour la méthode d'Euler d'ordre  $k$ ) est un  $v_0$  auquel on peut appliquer la proposition 2.1. Nous ne démontrerons et n'utiliserons 2.1 qu'avec  $k = 1$ . Pour  $k > 1$ , voir [ShSm1, p.115-121].

**REMARQUE 2.2.** La connaissance de  $r_w$  est inaccessible à un observateur en  $v = v_0$ ; on peut cependant minorer  $r_w$  par  $\rho_p$  le module de la plus petite valeur critique, qui peut s'estimer à l'aide du discriminant. Le critère obtenu a peu d'intérêt du point de vue numérique, car le calcul du discriminant coûte trop cher; il sera cependant utile pour donner des estimées probabilistes.

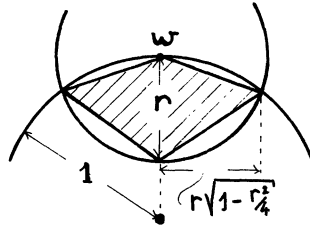
**PROPOSITION 2.3** ([ShSm1]). *Soit  $D$  le disque unité; la probabilité dans  $D \times P_d(1)$  pour que  $v_0$  dans  $D$  soit un bon zéro approché d'un polynôme  $p$  de  $P_d(1)$  est minorée par  $c/d^5$  avec  $c > 4 \cdot 10^{-4}$ .*

**DÉMONSTRATION DE 2.3:** Un point  $z$  est une valeur critique de  $p$  si et seulement si le polynôme  $p - z$  est de discriminant zéro. Comme le discriminant  $\mathcal{D} : P_d \rightarrow \mathbf{C}$  est de degré  $d - 1$  en le terme constant, on a, en posant  $\mathcal{E}_d(\rho) = \{p \in P_d(1) \mid \rho_p < \rho\}$ , le lemme suivant:

**LEMME 2.4.** *La mesure de  $\mathcal{E}_d(\rho)$  est majorée par  $(d - 1)\rho^2$ .  $\square$*

Comme le produit des racines de  $p$  est  $a_0$ , un polynôme  $p$  de  $P_d(1)$  (qui vérifie  $|a_0| \leq 1$ ) a une racine  $w$  dans  $D$ . Si de plus  $p$  est hors de  $\mathcal{E}_d(\rho)$  pour un  $\rho > 0$ ,  $w$  est un zéro simple et  $|p'(w)| \leq 1 + 2 + \dots + d = \frac{1}{2}d(d + 1)$ ; donc, grâce au théorème du quart (cf. [O], p. 3), le disque  $D'$  centré en  $w$  et de rayon  $r = \rho/18d(d + 1)$  est inclus dans  $p_w^{-1}(1/9D_w)$ , et donc constitué de bons zéros approchés. L'aire de  $D \cap D'$  est supérieure à  $\sigma = r^2 \sqrt{1 - r^2/4}$  (cf. Figure 3). Soit  $0 < \mu < 1$ . En prenant  $\rho = \sqrt{\mu/(d - 1)}$  dans 2.4, la probabilité cherchée est minorée par  $\frac{1}{\pi} \int_0^1 \sigma(\mu) d\mu \geq c/d^5$ .  $\square$

FIGURE 3



NOTATIONS 2.5. Le polynôme  $p$  étant fixé et  $v$  régulier pour  $p$ , on pose:  $\beta(v) = |p(v)/p'(v)|$ ,  $\gamma(v) = |p^{(k)}(v)/k!p'(v)|^{1/(k-1)}$  et  $\alpha(v) = \beta(v)\gamma(v)$ . Si l'on a besoin de préciser  $p$ , on écrit  $\beta(v, p)$ ,  $\gamma(v, p)$ ,  $\alpha(v, p)$ .

DÉMONSTRATION DE 2.1 LORSQUE  $k = 1$ : Calculons  $p(v_1)$  par la série de Taylor de  $p$  en  $v = v_0$ ; comme  $k = 1$ ,  $v_1 - v = \beta(v)$  et les deux premiers termes s'annulent, les autres étant majorés par  $|p(v)|(\alpha(v))^{k-1}$  d'où le

LEMME 2.6. *Si  $v$  est régulier pour  $p$  et  $\alpha(v) < \alpha$ , alors  $|p(v_1)| < \frac{\alpha}{1-\alpha} |p(v)|$ .*  $\square$

LEMME 2.7. *Si  $v$  est régulier pour  $p$ , alors  $\alpha(v) < 4|p(v)|/r_v$ .*

DÉMONSTRATION DE 2.7: Comme  $p$  est l'inverse de la fonction univalente  $p_v^{-1} : D_v \rightarrow \mathbf{C}$ , les coefficients  $a_k$  de son développement de Taylor en  $v$  sont bornés par un théorème de Löwner ([H], p. 137):

$$|a_k| \leq B_k |a_1|^k r_v^{k-1} \quad \text{où} \quad B_k = 2^k \frac{1 \cdot 3 \dots (2k-1)}{1 \cdot 2 \dots k \cdot (k+1)} \leq 4^{k-1}$$

d'où le lemme.  $\square$

Sous les hypothèses de 2.1, comme  $r_v > r_w - |p(v)|$ , on déduit de ces deux lemmes  $|p(v)| \leq A|p(v)|^2$  où  $A|p(v)| < b$ . D'autre part, comme sur  $D(0, r_w/9)$  les inverses  $p_v^{-1}$  et  $p_w^{-1}$  coïncident, et

$$|v_1 - v| = |p(v)/p'(v)| \leq r_w/9 |p'(v)| \leq r_v/8 |p'(v)| < r_v/4 |p'(v)| \quad ,$$

le théorème du quart assure que  $v_1$  est dans  $p_w^{-1}(D_w)$ . Donc, par récurrence,  $|p(v_{n+1})|^2 < A|p(v_n)|^2$  avec  $A|p(v_0)| < b$ , d'où la proposition.  $\square$



### 3. Le théorème de d'Alembert-Gauss, et une méthode de Newton globale.

L'ensemble  $\Sigma$  des valeurs critiques d'un polynôme  $p$  de degré  $d > 0$  est fini ( $|\Sigma| \leq d - 1$ ); comme  $p$  est une application ouverte, il y a un  $v$  tel que le rayon  $R_z = \{tz \mid 0 < t \leq 1\}$  ( $z = p(v)$ ) évite  $\Sigma$  (cf. Figure 4); le point  $v$  est donc régulier,  $p_v^{-1}$  se prolonge analytiquement à un voisinage de  $R_z$  et  $w = \lim_{t \rightarrow 0} p_v^{-1}(tz)$  est une racine de  $p$ . Par des majorations faciles, on peut obtenir un ensemble de  $d$  points de grand module contenant un tel  $v$ : cette démonstration est constructive. Elle conduisit Smale ([Sm2]) à étendre  $p_v^{-1}$  à des secteurs angulaires.

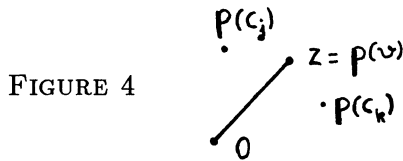


FIGURE 4

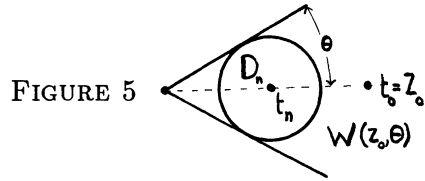


FIGURE 5

L'*étroitesse*  $e(v)$  (relativement à  $p$ ) d'un point  $v$  non racine de  $p$  est le plus grand  $\theta \geq 0$  tel que  $p_v^{-1}$  se prolonge au secteur:

$$W(z, \theta) = \{z' \in \mathbf{C} \mid |\text{Arg}(z'/z)| < \theta\}$$

(rappelons que  $z = p(v)$  et convenons que  $W(z, 0) = \{z\}$  (cf. [G])).

**THÉORÈME 3.1.** *Il y a des constantes  $M$  et  $L$  dépendant de  $\theta$  ( $L > 0 < M < 1$ ) telles que, pour  $1 - M \leq K < 1$  et  $v_0$  un point d'*étroitesse*  $\theta > 0$  relativement à un polynôme non constant  $p$ , et si l'on note  $t_i = K^i p(v_0)$  la suite  $v_i$  définie par:*

$$v_i = N_{p-t_i}(v_{i-1}) = v_{i-1} - \frac{p(v_{i-1}) - t_i}{p'(v_{i-1})} \quad (\#)$$

*converge vers une racine de  $p$ . De plus:*

$$|p(v_n)| \leq K^n (1 + L) |p(v_0)| \quad (\#\#)$$

*$M = \sin \theta / (24 + \sin \theta)$  et  $L = \frac{2}{3} M$  conviennent.*

**COROLLAIRE 3.2.** *Soit  $p$  un polynôme de  $P_d(1)$  et  $\varepsilon > 0$ . Si  $v_0$  est un point d'*étroitesse*  $\theta > 0$  et de module  $R > 1$ , et  $v_n$  est la suite définie par (#) avec  $K = 1 - M$ , alors, si  $n \geq N(d, \theta, R, \varepsilon)$ , on a  $|p(v_n)| \leq \varepsilon$ , où  $N(d, \theta, R, \varepsilon) = K_1 d + K_2 + K_3 \log(1/\varepsilon)$  avec  $K_1 = K_3 \log R$ ,  $K_2 = K_3 \log \frac{R(1+L)}{R-1}$  et  $K_3 = \frac{1}{\lceil \log(1-M) \rceil} < \frac{1}{M}$ .  $\square$*

DÉMONSTRATION DE 3.1: Posons  $z_n = p(v_n)$ ,  $r_n = r_{v_n}$  et  $D_n = D(t_n, t_n \sin \theta)$ . Comme  $t_n = K^n |p(v_0)|$ , l'inégalité (##) suit de:

$$(a_n) \quad |z_n - t_n| \leq L |t_n|.$$

(a<sub>0</sub>)

est vraie car  $t_0 = z_0$ . De (a<sub>n</sub>), il vient:

$$(b_n) \quad |z_n - t_{n+1}| \leq |z_n - t_n| + |t_n - t_{n+1}| \leq (L + M) |t_n|.$$

Comme  $D_n$  est inclus dans  $W(t_0, \theta)$  (cf. Figure 5), on tire de (a<sub>n</sub>) et (b<sub>n</sub>):

$$r_n \geq |t_n| \sin \theta - |z_n - t_n| \geq (\sin \theta - L) |t_n| \geq J |z_n - t_{n+1}| \quad ,$$

où  $J = (\sin \theta - L)/(L + M)$ . Supposons

$$J > 4 \quad . \quad (*)$$

Les lemmes 2.6 et 2.7 appliqués au polynôme  $p - t_{n+1}$  donnent:

$$|z_{n+1} - t_{n+1}| \leq \frac{4 |z_n - t_{n+1}|^2}{r_n - 4 |z_n - t_{n+1}|} \leq I |t_{n+1}| \quad ,$$

où  $I = 4(L + M)/M(J - 4)$ . Donc, si de plus

$$I < \sin \theta \quad , \quad (*')$$

le point  $z_{n+1}$  est dans  $W(z_0, \theta)$ , et si enfin

$$I < L \quad , \quad (*'')$$

la condition (a<sub>n+1</sub>) est vérifiée. Que les valeurs de  $M$  et  $L$  données dans 3.1 vérifient les trois hypothèses (\*) n'est plus qu'un exercice de calcul.  $\square$

#### 4. Les algorithmes probabilistes.

Soit  $S_R$  le cercle  $\{v \mid |v| = R\}$  muni de la mesure de Lebesgue normalisée par  $\mu(S_R) = 1$ . Soit  $p$  un polynôme de  $P_d(1)$  et posons  $E_R(\theta) = \{v \in S_R \mid e(v) \leq \theta\}$ ; on établira au §5 la:

PROPOSITION 4.1. *L'ensemble  $E_R(\theta)$  est inclus dans la réunion de  $2(d-1)$  arcs de  $S_R$  chacun de mesure:  $\frac{1}{\pi d} \left[ \theta + 2 \arcsin \frac{1}{R-1} \right]$ .*

On notera  $\mu_{\theta,R}$  la somme de leur mesure.

ALGORITHME  $\mathcal{A}(\theta, R)$  4.2 ([Sm4]). Soit  $p$  un polynôme de  $P_d(1)$  et  $\varepsilon > 0$ . Les constantes  $M$ ,  $K = 1 - M$ , et  $N$  (dépendant de  $d$ ,  $\theta$ ,  $R$ , et  $\varepsilon$ ) étant celles de 3.2, procéder comme suit:

- 1° Choisir au hasard un  $v_0$  sur  $S_R$ .
- 2° Calculer  $v_N$  en itérant  $N$  fois (3.1).
- 3° Si  $|p(v_N)| \leq \varepsilon$  fin; sinon retourner au 1°.

La moyenne du nombre de cycles dans cet algorithme est majorée par  $t = t_{\theta, R} = 1/(1 - \mu_{\theta, R})$ .  $\square$

#### QUELQUES VALEURS NUMÉRIQUES.

	$M$	$t$	$tK_1$	$tK_2$	$tK_3$
$R = 3, \theta = \pi/12$	: 1/94	6	615	231	560
$R = 6, \theta = \pi/6$	: 1/49	2,44	212	24	119

Lorsque, pour  $R$  donné, on cherche à minorer  $tK_1$ , l'optimum semble être autour de  $R = 8$  et  $\theta = 7\pi/36$  et est assez stable (Pour  $6 \leq R \leq 16$ ,  $tK_1$  varie entre 206 et 223.) Assez paradoxalement, il est plus économique de partir d'une grande valeur initiale (cf. §5).

Si vous ne faites pas confiance au hasard, soit  $k = 2(d-1)\ell t_{\theta, R}$  un entier avec  $\ell > 1$ . Placez alors  $k$  points également répartis sur  $S_R$ . Comme  $E_k(\theta)$  est dans une réunion de  $2(d-1)$  arcs, il ne peut contenir plus de  $T_k = 2(d-1) + k\mu_{\theta, R}$  de ces points. Tirant au sort les valeurs initiales dans cette urne de  $k$  points, on est sûr que l'algorithme 4.2 n'a pas plus de  $T_k$  cycles; la moyenne du nombre de tirages est majorée par  $t_k = t_{\theta, k}\ell/(\ell-1)$ . Le choix de  $\ell$  est affaire de goût: un grand  $\ell$  améliore la moyenne, mais recule la barrière de sécurité  $T_k$ .

THÉORÈME 4.3. Soit  $0 < \mu < 1$ . Il y a une fonction  $N(\mu, d)$  telle que, si  $R = 1 + 4d^2/\pi\mu$  et  $(v_0, p)$  est hors d'un ensemble exceptionnel de mesure  $\mu$  dans  $S_R \times P_d(1)$ , il suffit, partant de  $v_0$  et fixant  $\theta = \frac{1}{2}\pi\mu$  de  $N(\mu, d)$  itérations de la formule 3.1(♯) pour obtenir un bon zéro approché de  $p$ . On a :

$$N(\mu, d) = K_1(\mu)(d + \frac{1}{2}) \frac{\log(1/\mu) + K_2(d)}{\mu} + K_3(d)$$

où  $K_1$  est croissante,

$$15,2 \leq K_1(\mu) \leq 25 \quad , \quad K_1(1/4) < 15,5 < K_1(1/2) < 18 \quad ,$$

$$(d + \frac{1}{2})K_2 = 3 \log d + \log(36/\pi) \quad , \quad \text{et} \quad K_3(d) = 2/3 + K_1\pi/4d^2 \quad .$$

DÉMONSTRATION DE 4.3: La mesure des  $(v_0, p)$  tels que, soit  $\rho_p < \rho$ , soit  $e_p(v_0) < \theta$ , est majorée par

$$\mu = (d-1)\rho^2 + \frac{2(d-1)}{\pi d} \left[ \theta + 2 \arcsin \frac{1}{R-1} \right].$$

Il suffit de poser  $\rho = \sqrt{\mu}/d$  et d'appliquer 3.2 avec  $\varepsilon = \rho/9$ , et, pour  $\theta$  et  $R$ , les valeurs données dans 4.3.  $\square$

## 5. La géométrie des polynômes (cf. [ShSm1]).

Une idée naturelle pour trouver un zéro d'un polynôme  $p$  est de minimiser la fonction réelle  $g(v) = \frac{1}{2} |p(v)|^2$  en descendant le long des lignes de gradient. Or,  $X(v) = \text{grad } g(v) = \overline{p'(v)} p(v)$  est colinéaire au champ de Newton  $n(v) = p(v)/p'(v)$  ( $X(v) = |p'(v)|^2 n(v)$ ) et donc  $X$  a mêmes courbes intégrales que l'équation différentielle de Newton  $v' = -n(v)$ . La méthode d'Euler, avec le pas 1 pour résoudre cette équation différentielle, produit la même itération que la méthode de Newton. Remarquons aussi que, comme  $p'(v)n(v) = p(v) = z$ , le polynôme  $p$  envoie son champ de Newton sur le champ radial du but. Donc les lignes de gradient que nous cherchons à descendre sont aussi les feuilles du feuilletage  $\mathcal{F}_p$ , préimage par  $p$  des rayons de  $\mathbf{C}$  à savoir les courbes que Smale suivait dans sa démonstration du théorème fondamental de l'algèbre. Dans [Sm2] (et aussi [ShSm1]), Smale avait suivi ces lignes de gradient de plus près en relaxant le pas de la méthode d'Euler, i.e. en itérant  $v \mapsto v - hp(v)/p'(v)$  où  $0 < h < 1$ .

De manière imagée, si le graphe de  $g$  est une montagne que l'on cherche à descendre, au lieu à chaque pas de viser le refuge attendu, Smale [Sm2] vise une étape intermédiaire, disons à mi-distance. La méthode du §3 ([Sm4]) est plus sûre: avant de partir, on détermine toutes ses étapes, ce qui permet d'éviter de mauvais embranchements aux cols. Pour expliquer le paradoxe du §4, on peut dire que, de plus haut, la vue permet de distinguer la vallée principale des vallées secondaires et donc on peut faire un meilleur plan de route.

Le feuilletage  $\mathcal{F}_p$ , orienté par  $X$ , a des singularités de deux types (cf. Figure 6): des sources aux racines de  $p$  et des selles de singe à  $m-1$  pattes aux points critiques de multiplicité  $m$  et de valeur critique non nulle ( $0 = p'(c) = \dots = p^{(m-1)}(c)$  et  $p(c)p^{(m)}(c) \neq 0$ ). Soit  $\Gamma(p)$  la réunion des variétés instables des points critiques de  $\mathcal{F}_p$ ; c'est un arbre plongé dans  $\mathbf{C}$ , l'arbre de Shub-Smale du polynôme  $p$ . Posons  $\Gamma(p)^+$  l'union de  $\Gamma(p)$  et des variétés stables des singularités de type selle (éventuellement multiples) de  $\mathcal{F}_p$ . Les composantes de  $\Gamma(p)^+/\Gamma(p)$  sont en nombre inférieur ou égal à  $2(d-1)$  (ce maximum est atteint quand les racines de  $p$  sont simples, les points critiques sont de multiplicité 2 et n'ont pas de liaisons entre eux dans  $\mathcal{F}_p$ ).

FIGURE 6

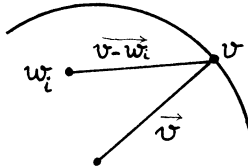
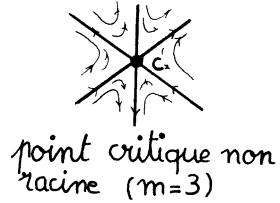


FIGURE 7

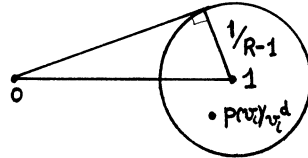


FIGURE 8

REMARQUE 5.1. Les arbres plongés dans  $\mathbf{C}$  qui peuvent se réaliser comme arbres de Shub-Smale d'un polynôme peuvent être facilement déterminés et, à l'aide des arbres de Shub-Smale, on peut paramétrer les polynômes par leurs valeurs critiques (cf. [M] et [STW]).

LEMME 5.2 (Gauss). Soit  $p$  un polynôme dont toutes les racines sont à l'intérieur d'un cercle  $S$ , alors le champ de Newton est transverse à  $S$ .

COROLLAIRE 5.3. Soit  $S$  un cercle contenant dans son intérieur tous les zéros d'un polynôme  $p$ , alors:

- (i) l'arbre  $\Gamma(p)$  est dans l'intérieur de  $S$ ;
- (ii)  $\Gamma(p)^+ \cap S$  contient au plus  $2(d-1)$  points. □

LEMME 5.4. Soit  $p$  un polynôme de  $P_d(1)$ ,  $R \geq 2$ ,  $S_R$  le cercle  $|v| = R$  et  $v_1, v_2$  deux points de  $S_R$ . Alors :

- (i)  $p(v_i)$  est non nul;
- (ii) 
$$d \left| \operatorname{Arg} \frac{v_1}{v_2} \right| - 2 \arcsin \frac{1}{R-1} \leq \left| \operatorname{Arg} \frac{p(v_1)}{p(v_2)} \right| \leq d \left| \operatorname{Arg} \frac{v_1}{v_2} \right| + 2 \arcsin \frac{1}{R-1} .$$

DÉMONSTRATION DE 4.1: Soit  $R \geq 2$  et  $p$  dans  $P_d(1)$ ; remarquons qu'un point  $v_0$  de  $S_R$  est dans  $E_R(\theta)$  si et seulement s'il y a un  $v'$  dans  $S_R \cap \Gamma(p)^+$  avec  $|\operatorname{Arg}(p(v)/p(v'))| < \theta$ . D'après 5.4(i) et 5.2,  $\mathcal{F}_p$  est transverse à  $S_R$  et l'un des arcs limités par  $v$  et  $v'$  sur  $S_R$  a tous ses points  $u$  vérifiant  $|\operatorname{Arg}(p(u)/p(v'))| < \theta$ . La proposition 4.1 suit alors de l'inégalité de gauche dans 5.4 (ii). □

DÉMONSTRATION DE 5.2 (voir Figure 7) : Soient  $w_i$  les racines de  $p$  et soit  $v$  sur  $S$ . On a :

$$|n(v)|^{-2} v \bar{n}(v) = v \frac{p'(v)}{p(v)} = \sum_{i=1}^d \frac{v}{v - w_i} .$$

Or, comme  $w_i$  est intérieur à  $S$ , on a  $\operatorname{Re}(v(\overline{v - w_i})) > 0$ ; donc  $\operatorname{Re}(v/(v - w_i)) > 0$  et  $\operatorname{Re}(v \bar{n}(v)) > 0$ .  $\square$

DÉMONSTRATION DE 5.4 (voir Figure 8) : Soit  $p(v) = v^d + a_1 v^{d-1} + \dots + a_d$  pour  $|v_j| = R$ . On a :

$$\left| 1 - \frac{p(v_j)}{v_j^d} \right| \leq \sum_{i=1}^d \frac{|a_i|}{|v_j|^d} < \frac{1}{|v_j|} \frac{1}{1 - 1/|v_j|} = \frac{1}{R-1} \leq 1$$

car  $R \geq 2$ , d'où (i); et pour (ii):

$$\begin{aligned} \left| \operatorname{Arg} \frac{p(v_1)}{p(v_2)} - \operatorname{Arg} \left( \frac{v_1}{v_2} \right)^d \right| &= \operatorname{Arg} \left| \frac{p(v_1)}{v_1^d} \frac{v_2^d}{p(v_2)} \right| \leq \\ &\leq \left| \operatorname{Arg} \frac{p(v_1)}{v_1^d} \right| + \left| \operatorname{Arg} \frac{p(v_2)}{v_2^d} \right| \leq 2 \arcsin \frac{1}{R-1} . \end{aligned}$$

$\square$

## 6. Un critère de convergence calculable en $v_0$ .

Soit  $v$  un point régulier d'un polynôme  $p$ . En 2.5, on a défini  $\alpha(v, p) = \beta(v, p) \gamma(v, p)$  où  $\beta(v, p)$  est la norme du vecteur de Newton  $p(v)/p'(v)$ , alors que le calcul de  $\gamma(v)$  nécessite la connaissance de toutes les dérivées de  $p$  en  $v$ . La proposition suivante permet de majorer  $\gamma(v)$  (et donc  $\alpha(v)$ ) en n'ayant à évaluer que  $p$  et  $p'$  en  $v$  :

PROPOSITION 6.1 ([Sm5], [Sm6]). Soit  $\varphi(r) = 1 + \dots + r^d$  et  $p(v) = a_0 + a_1 v + \dots + a_d v^d$  un polynôme de degré  $d$ , alors :

$$\gamma(v, p) \leq \frac{\|p\| (\varphi'(|v|))^2}{|p'(v)| \varphi(|v|)} , \quad \text{où} \quad \|p\| = \operatorname{Max} |a_i| .$$

THÉORÈME 6.2 ([Sm5]). Soit  $p$  un polynôme et  $v_0$  un point régulier tel que  $\alpha(v_0, p) < 3 - 2\sqrt{2}$ , alors la suite de Newton  $v_{n+1} = v_n - p(v_n)/p'(v_n)$  converge vers une racine de  $p$ . De plus, il y a une bijection croissante  $K : [0, 3 - 2\sqrt{2}[ \longrightarrow ]0, 1[$  telle que: Si  $\alpha(v_0, p) < \alpha$ , alors  $\|v_{n+1} - v_n\| \leq K(\alpha)^{2^n - 1} \|v_1 - v_0\|$  et  $\alpha_0 = K^{-1}(\frac{1}{2}) > 0, 14$ .

Pour  $p$  dans  $P_d(1)$ , le nombre  $\alpha(0, p)$  est majoré par  $|a_0|/|a_1|^2$  qui est indépendant de  $d$ . Ainsi, l'ensemble des polynômes de  $P_d(1)$ , dont la suite de Newton partant de zéro converge, a une mesure minorée indépendamment de  $d$ .

DÉMONSTRATION DE 6.1: En majorant les coefficients de  $p$  par  $\|p\|$ , on obtient  $|p^{(k)}(v)/k!| \leq \|p\| \varphi^{(k)}(|v|)/k!$ , d'où en faisant  $k = 1$  on obtient  $\|p\| \varphi'(|v|)/|p'(v)| \geq 1$ . Un jeu sur les coefficients du binôme ([Sm6], p.12-14) permet d'établir que  $\alpha(r, \varphi) \leq 1$  pour tout  $r \geq 0$ , d'où

$$\left| p^{(k)}(v)/k! \right| \leq \|p\| \varphi^{(k)}(|v|)/k! \leq \|p\| (\varphi'(|v|))^k / (\varphi(|v|))^{k-1} \quad , \quad \text{et}$$

$$\gamma(v, p) \leq \sup_{k>1} \left[ \frac{\|p\| (\varphi'(|v|))^k}{|p'(v)| (\varphi(|v|))^{k-1}} \right]^{1/k-1} \leq \frac{\varphi'}{\varphi} \sup_{k>1} \left[ \frac{\|p\| \varphi'}{|p'|} \right]^{1/k-1} .$$

Le dernier suprémum est atteint pour  $k = 2$ , d'où la proposition.  $\square$

DÉMONSTRATION DE 6.2: Pour  $0 \leq t < \gamma^{-1}$ , soit  $\psi$  la fonction  $\psi(t) = \frac{1}{1-\gamma t} - 2\gamma t + \alpha - 1$ ; comme en 2.6, on établit que, pour  $\|v - v_0\| \leq t\gamma^{-1}$  on a:

$$\left| \frac{p(v_0)}{p'(v_0)} \right| \leq \frac{\psi(0)}{-\psi'(0)} \quad , \quad \text{et} \quad \left| \frac{p''(v)}{p'(v_0)} \right| \leq \frac{\psi''(t)}{-\psi'(0)} .$$

Kantorovitch ([KA], théorème 3, p.234) nous dit que, si  $\psi(t) = 0$  a une racine dans  $[0, \gamma^{-1}[$ , alors la suite  $v_n$  converge vers une racine  $w$  de  $p$ , que la suite de Newton  $t_n$  de  $\psi$  partant de  $t_0 = 0$  converge vers la plus petite racine  $t_*$  de  $\psi(t) = 0$  et  $|v_{n+1} - v_n| \leq t_{n+1} - t_n$ .

Il ne reste plus qu'à estimer  $t_{n+1} - t_n$  pour obtenir les majorations du théorème 6.2.  $\square$

REMARQUES. Comme Kantorovitch énonce son théorème dans un Banach, le théorème 6.2 est valable pour une fonction analytique définie sur un Banach (il en est bien sûr de même pour 6.1). Smale démontre 6.2 plus directement en majorant à chaque itération de Newton les quantités  $\beta$  et  $\gamma$ . N'utilisant pas le théorème de Kantorovitch, cette approche se généralise aux méthodes d'ordre supérieur ([C]).

### 7. Un algorithme génériquement convergent ([ShSm3]).

Soit  $G_d$  l'ouvert de  $P_d$  formé des polynômes  $p(v) = a_0 + a_1v + \dots + a_dv^d$  ayant leurs racines et leurs points critiques distincts:  $P_d \setminus G_d$  étant algébrique,  $G_d$  est un ouvert dense de mesure pleine. Soient  $p$  dans  $G_d$  et  $v$  dans la sphère de Gauss  $S = \mathbf{C} \cup \infty$ . Définissons alors:

$$k_p(v) = \frac{\varphi(|v|) |p'(v)|^2}{2(\varphi'(|v|))^2 |p(v)| \|p\|} \quad , \quad h_p(v) = \text{Min}(1, k_p(v)) \quad ,$$

et  $T_p(v) = v - h_p(v)p(v)/p'(v)$  (où  $\varphi$  et  $\|p\|$  sont définis au §6).

Les points fixes de  $T_p$  sont les racines et les points critiques de  $p$ . L'application  $\tilde{T} : S \times G_d \rightarrow S$  ainsi définie est continue et  $C^\infty$  près des points critiques de  $p$ . Quand  $p$  est fixé par le contexte, nous écrirons  $T$  au lieu de  $T_p$ .

REMARQUE. Les formules ci-dessus définissent une application continue  $S \times P_d \rightarrow S$ , mais si  $p$  a des racines multiples, cette application n'est pas lisse près des racines multiples. Remarquons d'autre part que l'application de Newton  $\tilde{N} : S \times P_d \rightarrow S$  n'est pas continue (en  $P_d$ ) aux racines multiples, car quand le polynôme varie, un point critique se rapprochant d'une racine voit son image passer brusquement de l'infini à cette racine.

THÉORÈME 7.1. *Pour tout  $p$  de  $G_d$ , il y a un fermé  $V_p$  de mesure nulle dans  $S$  tel que, pour tout  $v$  hors de  $V_p$  la suite des itérés  $T^k(v)$ ,  $k \geq 0$ , converge vers un zéro de  $p$ .*

*De plus,  $T$  est la méthode de Newton dans un voisinage de chaque racine de  $p$ .*

PROPOSITION 7.2. *Soit  $p$  dans  $G_d$  et  $c$  un point critique de  $p$ , alors*

$$V^s(c) = \{v \mid T^k(v) \rightarrow c \text{ quand } k \rightarrow \infty\}$$

*est un fermé de mesure nulle dans  $S$ .*

DÉMONSTRATION DE 7.2: Comme le jacobien de  $T$  est presque partout non nul, il suffit de montrer qu'un voisinage de  $c$  dans  $V^s(c)$  est de mesure nulle. Comme  $T$  est lisse au voisinage de  $c$ , il suffit, selon le théorème de la variété stable ([HPS]), de prouver que  $T'(c)$  est hyperbolique (en effet ce théorème impliquera alors que  $V^s(c)$  est, près de  $c$ , un arc lisse). On a

$$T'(c)(V) = V - \lambda \bar{V} \quad , \quad \text{où} \quad \lambda = \frac{p(c)\bar{p}''(c)\varphi(|c|)}{2(\varphi'(|c|))^2 \|p\| |p(c)|}$$



Les valeurs propres de  $T'(c)$  sont donc  $1 \pm |\lambda|$ ; or, comme on a  $\varphi(r)\varphi''(r) \leq 2(\varphi'(r))^2$ ,

$$|\lambda| \leq \frac{|p''(c)|}{\|p\|} \frac{1}{\varphi''(|c|)} \leq 1 \quad .$$

□

DÉMONSTRATION DE 7.1: Posons  $V_p = \bigcup \{V^s(c) \mid p'(c) = 0\}$ . Soit  $v$  hors de  $V_p$  et  $v_n = T^n(v)$ , alors  $p'(v) \neq 0$  et comme, d'après 6.1,  $k(v) < \frac{1}{2}\alpha(v)$ , on obtient par une majoration analogue à celle de 2.6 :

$$|p(v_{n+1})| \leq h(v_n) |p(v_n)| < |p(v_n)| \quad ,$$

où l'application continue  $h : \mathbf{C} \rightarrow [0, 1]$  n'est nulle qu'aux points critiques de  $p$ .

Comme  $v$  n'est dans aucun des  $V^s(c)$ , cela implique que la suite décroissante  $|p(v_n)|$  tend vers zéro et  $v_n$  tend vers une racine de  $p$  (deux racines de  $p$  ne peuvent être adhérentes à  $v_n$  puisque, comme  $T$  coïncide avec l'application de Newton près des racines de  $p$ , ces dernières sont des points fixes attractifs de  $T$ ). □

## BIBLIOGRAPHIE

- [Ba] B. BARNA, *Über Divergenzpunkte des Newtonsche Verfahrens zur Bestimmung von Wurzeln algebraischer Gleichungen III*, Publ. Math. Debrecen **8** (1961), 193-207.
- [C] J. CURRY, *On zero finding methods of higher order from data at one point*, MSRI Berkeley (1986).
- [DH] A. DOUADY AND J.H. HUBBARD, *On the dynamics of polynomial like mappings*, Ann. Sci. E.N.S., 4e série, t.18 (1985), 287-343.
- [F] J.-P. FRANCOISE, *Estimations uniformes pour les domaines de convergence de la méthode de Newton*, Séminaire de Géométrie Algébrique Réelle (J.J. Risler), Publ. Univ. Paris VII, **24** (1986).
- [G] O. GABBER, *Diverses interventions au Séminaire Bourbaki*, tradition orale.
- [H] W. HAYMAN, "Multivalent functions", Cambridge Univ. Press, Cambridge, 1958.
- [HPS] M. HIRSCH, C. PUGH AND M. SHUB, *Invariant manifolds*, Lectures Notes in Math. **583** (1977), Springer, New-York.

- [HM] M. HURLEY AND C. MARTIN, *Newton's algorithm and chaotic dynamical systems*, SIAM J. Math. Anal. **15** (1984), 238-252.
- [Ka] L. KANTOROVITCH ET G. AKILOV, *Analyse fonctionnelle*, t.2, Editions MIR, Moscou, 1981.
- [McM] C. McMULLEN, "Families of Rational Maps and Iterative Root-Finding Algorithms", Ph. D., Harvard Univ., Mai 1985, à paraître.
- [M] A. MARIN, *Les arbres de Shub-Smale*, Séminaire de Géométrie algébrique réelle (J.J. Risler), Publ. Univ. Paris VII **24** (1986).
- [O] J. OESTERLE, *Démonstration de la conjecture de Bieberbach*, exposé n°649 du Séminaire Bourbaki (juin 1985).
- [R1] J. RENEGAR, *On the efficiency of Newton's method in approximating all zeros of a system of complex polynomials*, à paraître dans Mathematics of Operations Research.
- [R2] J. RENEGAR, *A polynomial-time algorithm based on Newton's method for linear programming*, MSRI Berkeley (1986).
- [SU] D. SAARI AND J. URENKO, *Newton's method, circle maps and chaotic motions*, Amer. Math. Monthly **91** (1984), 3-17.
- [Sh] M. SHUB, *The geometry and topology of dynamical systems, and algorithms for numerical problems*, notes préparées pour des conférences données à D.D.A, Université de Peking, Beijing, Chine, Août-Septembre 1983.
- [ShSm1&2] M. SHUB AND S. SMALE, *Computational complexity on the geometry of polynomials and a theory of cost*, Part I, Ann. Sci. E.N.S. (4) t.18 (1985), 107-161; Part II, SIAM J. Computing **15** (1986), 145-161.
- [ShSm3] M. SHUB AND S. SMALE, *On the existence of generally convergent algorithms*, Jour. of Complexity **2** (1986), 2-11.
- [STW] M. SHUB, D. TISCHLER AND R. WILLIAMS, *The Newtonian graph of a complex polynomial*, soumis à SIAM J. of Math. Analysis.
- [Sm1] S. SMALE, *A Convergent process of price adjustment and global Newton methods*, J. Math. Econom. **3** (1976), 107-120.
- [Sm2] S. SMALE, *The fundamental theorem of algebra and complexity theory*, Bull. Amer. Math. Soc. **4** (1981), 107-120.
- [Sm3] S. SMALE, *The Problem of the average speed of the simplex method*, in "Mathematical Programming: the state of the Art, Bonn 1982" (editors Bachem et al.), Springer, 1983.
- [Sm4] S. SMALE, *On the efficiency of algorithms of analysis*, Bull. Amer. Math. Soc. **13** (1985), 87-121.

- [Sm5] S. SMALE, *Newton's method estimates from data at one point*, à paraître dans: *Proceedings of a Conference at Laramie in Honor of Gail S. Young*, Springer New York (1986).
- [Sm6] S. SMALE, *Algorithms for solving equations*, MSRI Berkeley, 1986 (texte écrit pour le congrès mondial de Berkeley).

Vous trouverez de nombreuses autres références dans les bibliographies de [Sm2], [Sm4], [Sm6]; en plus, [Sm6] contient une discussion de l'état actuel des problèmes posés dans [Sm2] et [Sm4].

Alexis MARIN  
UA 1169 du CNRS  
Université Paris-Sud  
Mathématiques, bât 425  
91405 ORSAY