Ultrametric Smale's α-theory

Jazz G. Suchen, Josué Tonelli-Cueto

Ultrametric Smale's α-theory

Jazz G. Suchen
Institute for Interspecies Studies
Berlin, Germany
jazz.g.suchen@outlook.com

Josué Tonelli-Cueto
Inria Paris & IMJ-PRG
Paris, France
josue.tonelli.cueto@bizkaia.eu

Abstract

We present a version of Smale's α-theory for ultrametric fields, such as the p-adics and their extensions, which gives us a multivariate version of Hensel's lemma.

Hensel's lemma [4, §3.4] gives us sufficient condition for lifting roots mod p^k to roots in $\mathbb{Z}_p$. Alternatevely, Hensel's lemma gives us sufficient conditions for Newton's method convergence towards an approximate root. Unfortunately, in the multivariate setting, versions of Hensel's lemma are scarce [2]. However, in the real/complex world, Smale's α-theory [3] gives us a clean sufficient criterion for deciding if Newton's method will converge quadratically. In the p-adic setting, Breiding [1] proved a version of the γ-theorem, but he didn't provide a full α-theory. In this short communication, we provide an ultrametric version of Smale's α-theory for square systems—initially presented as an appendix in [5]—, together with an easy proof.

In what follows, and for simplicity[1], $\mathbb{F}$ is a non-archimedian complete field of characteristic zero with (ultrametric) absolute value $|\ |$ and $\mathcal{P}_{n,\boldsymbol{d}}[n]$ the set of polynomial maps $f:\mathbb{F}^n\to\mathbb{F}^n$ where f_i is of degree d_i. In this setting, we will consider on $\mathbb{F}^n$ the ultranorm given by $\|x\|:=\max\{x_1,\dots,x_n\}$, its associated distance $\mathrm{dist}(x,y):=\|x-y\|$, and on k-multilinear maps $A:(\mathbb{F}^n)^k\to\mathbb{F}^q$ the induced ultranorm, which is given by

$$\|A\|:=\sup_{v_1,\dots,v_k\neq 0}\frac{\|A(v_1,\dots,v_k)\|}{\|v_1\|\cdots\|v_k\|}. \tag{1}$$

In this context, we can define Smale's parameters as follows. Below D_xf denotes the differential map of f at x and D_x^kf the k-linear map induced by the kth order partial derivatives of f at x.

Definition 1 (Smale's parameters)**.** Let $f\in\mathcal{P}_{n,\boldsymbol{d}}[n]$ and $x\in\mathbb{F}^n$. We define the following:

(a) *Smale's* α: $\alpha(f,x):=\beta(f,x)\gamma(f,x)$, if D_xf is non-singular, and $\alpha(f,x):=\infty$, otherwise.

(b) *Smale's* β: $\beta(f,x):=\|\mathrm{D}_xf^{-1}f(x)\|$, if D_xf is non-singular, and $\alpha(f,x):=\infty$, otherwise.

(c) *Smale's* γ: $\gamma(f,x):=\sup_{k\geq 2}\left\|\mathrm{D}_xf^{-1}\frac{\mathrm{D}_x^kf}{k!}\right\|^{\frac{1}{k-1}}$, if D_xf is invertible, and $\gamma(f,x):=\infty$, otherwise.

[1] We focus on characteristic zero and polynomials to avoid technical details related to Taylor series.

If $\mathrm{D}_x f$ is invertible, then the *Newton operator*,

$$\mathrm{N}_f : x \mapsto x - \mathrm{D}_x f^{-1} f(x),$$

is well-defined at x. For a point x, the *Newton sequence* is the sequence $\{\mathrm{N}_f^k(x)\}$. Note that this sequence is well-defined (i.e., $\mathrm{N}_f^k(x)$ makes sense for all k) if and only if $\mathrm{D}_{\mathrm{N}_x^k(x)} f$ is invertible at every k. Also note that

$$\beta(f, x) = \|x - \mathrm{N}_f(x)\|,$$

so β measures the length of a Newton step.

Theorem 1 (Ultrametric α/γ-theorem). *Let $f \in \mathcal{P}_{n,\boldsymbol{d}}[n]$ and $x \in \mathbb{F}^n$. Then the following are equivalent:*

$$(\alpha)\ \alpha(f, x) < 1 \quad \text{and} \quad (\gamma)\ \operatorname{dist}(x, f^{-1}(0)) < 1/\gamma(f, x)$$

Moreover, if any of the above equivalent conditions holds, then the Newton sequence, $\{\mathrm{N}_f^k(x)\}$, is well-defined and it converges quadratically to a non-singular zero ζ of f. More specifically, for all k, the following holds:

(a) $\alpha(f, \mathrm{N}_f^k(x)) \leq \alpha(f, x)^{2^k}$. (b) $\beta(f, \mathrm{N}_f^k(x)) \leq \beta(f, x)\alpha(f, x)^{2^k-1}$. (c) $\gamma(f, \mathrm{N}_f^k(x)) \leq \gamma(f, x)$.

(Q) $\|\mathrm{N}_f^k(x) - \zeta\| = \beta(f, \mathrm{N}_f^k(x)) \leq \alpha(f, x)^{2^k-1}\beta(f, x) < \alpha(f, x)^{2^k}/\gamma(f, x)$.

In the univariate p-adic setting, we have that for $f \in \mathbb{Z}_p[X]$ and $x \in \mathbb{Z}_p$,

$$\gamma(f, x) \leq 1/|f'(x)|,$$

since $|1/k! f^{(k)}(x)| \leq 1$ and $|f'(x)| \leq 1$. Therefore we can see that the condition $|f(x)| < |f'(x)|^2$ of Hensel's lemma implies $\alpha(f, x) < 1$ for a p-adic integer polynomial. In this way, we can see that Theorem 1 generalizes Hensel's lemma to the multivariate case.

Moreover, in the univariate setting, we can show the following proposition which gives a precise characterization of Smale's γ in the ultrametric setting as the separation between 'complex' roots—not only a bound as it happens in the complex/real setting.

Proposition 2 (Ultrametric separation theorem for γ). *[5, Theorem 3.15] Fix an algebraic closure $\overline{\mathbb{F}}$ of $\mathbb{F}$ with the corresponding extension of the ultranorm. Let $f \in \mathbb{F}[X]$ and $\zeta \in \overline{\mathbb{F}}$ a simple root, then*

$$\frac{1}{\gamma(f, \zeta)} = \operatorname{dist}(\zeta, f^{-1}(0) \setminus \{\zeta\}). \qquad \square$$

Proof of Theorem 1

The proof of the theorem relies in the following three lemmas, stated for $f \in \mathcal{P}_{n,\boldsymbol{d}}[n]$ and $x, y \in \mathbb{F}^n$.

Lemma 3. *If $\gamma(f, x)\|x - y\| < 1$, then $\mathrm{D}_y f$ is invertible and $\|\mathrm{D}_y f^{-1} \mathrm{D}_x f\| = 1$.*

Lemma 4 (Variations of Smale's parameters). *If $\rho := \gamma(f, x)\|x - y\| < 1$, then:*

(a) $\alpha(f, y) \leq \max\{\alpha(f, x), \rho\}$. (b) $\beta(f, y) \leq \max\{\beta(f, x), \|y - x\|\}$. (c) $\gamma(f, y) = \gamma(f, x)$.

Moreover, if $\|y - x\| < \beta(f, x)$, all are equalities.

Lemma 5 (Variations along Newton step)**.** *If $\alpha(f,x)<1$, then:*

(a) $\alpha(f,\mathrm{N}_f(x))\leq\alpha(f,x)^2$. (b) $\beta(f,\mathrm{N}_f(x))\leq\alpha(f,x)\beta(f,x)$. (c) $\gamma(f,\mathrm{N}_f(x))=\gamma(f,x)$.

In particular, $\mathrm{N}_f(\mathrm{N}_f(x))$ *is well-defined.*

Proof of Theorem 1. If $\alpha(f,x)<1$, then, using induction and Lemma 5, we obtain that (a), (b) and (c) hold. But then the sequence $\{\mathrm{N}_f^k(x)\}$ converges since $\lim_{k\to\infty}\|\mathrm{N}_f^{k+1}(x)-\mathrm{N}_f^k(x)\|=0$ and so it is a Cauchy sequence. Finally, (Q) follows from noting that for $l\geq k$

$$\|\mathrm{N}_f^l(x)-\mathrm{N}_f^k(x)\|\leq\alpha(f,x)^{2^{l-k}}\beta(f,\mathrm{N}_f^k(x))$$

and taking infinite sum together with the equality case of the ultrametric inequality. In particular, we have $\operatorname{dist}(x,f^{-1}(0))=\|x-\zeta\|=\beta(f,x)<1/\gamma(f,x)$. This shows that ($\alpha$) implies ($\gamma$).

For the other direction, assume that $\operatorname{dist}(x,f^{-1}(0))<1/\gamma(f,x)$. Then $\gamma(f,x)$ is finite, since otherwise $\operatorname{dist}(x,f^{-1}(0))<0$, which is impossible. Let $\zeta\in\mathbb{F}^n$ be a zero of f such that $\operatorname{dist}(x,\zeta)<1/\gamma(f,x)$. Then $0=f(\zeta)=f(x)+\mathrm{D}_xf(\zeta-x)+\sum_{k=2}^\infty\frac{\mathrm{D}_x^kf}{k!}(\zeta-x,\dots,\zeta-x)$, and so

$$-\mathrm{D}_xf^{-1}f(x)=\zeta-x+\sum_{k=2}^\infty\mathrm{D}_xf^{-1}\frac{\mathrm{D}_x^kf}{k!}(\zeta-x,\dots,\zeta-x).$$

Now, the higher order terms satisfy that $\left\|\mathrm{D}_xf^{-1}\frac{\mathrm{D}_x^kf}{k!}(\zeta-x,\dots,\zeta-x)\right\|\leq(\gamma(f,x)\|\zeta-x\|)^{k-1}\|\zeta-x\|<\|\zeta-z\|$ and so, by the equality case of the ultrametric inequality, $\beta(f,x)=\|\zeta-x\|<1/\gamma(f,x)$, as desired. □

Now, we prove the auxiliary lemmas 3, 4 and 5

Proof of Lemma 3. We have that $\mathrm{D}_xf^{-1}\mathrm{D}_yf=\mathbb{I}+\sum_{k=1}^\infty\mathrm{D}_xf^{-1}\frac{\mathrm{D}_x^{k+1}f(y-x,\dots,y-x)}{k!}$. Now, under the given assumption, $\left\|\mathrm{D}_xf^{-1}\frac{\mathrm{D}_x^{k+1}f(y-x,\dots,y-x)}{k!}\right\|\leq(\gamma(f,x)\|y-x\|)^{k-1}<1$ for $k\geq2$, and so, by the the ultrametric inequality, $\|\mathrm{D}_xf^{-1}\mathrm{D}_yf-\mathbb{I}\|<1$. Therefore $\sum_{k=0}^\infty(\mathbb{I}-\mathrm{D}_xf^{-1}\mathrm{D}_yf)^k$ converges, and it does so to the inverse of $\mathrm{D}_xf^{-1}\mathrm{D}_yf$. Since, by assumption D_xf is invertible, so it is D_yf.

Finally, by the invertibility of D_yf, we have that $\mathrm{D}_yf^{-1}\mathrm{D}_xf=\sum_{k=0}^\infty(\mathbb{I}-\mathrm{D}_xf^{-1}\mathrm{D}_yf)^k$, and so, by the equality case of the ultrametric inequality, $\|\mathrm{D}_yf^{-1}\mathrm{D}_xf\|=1$, as desired. □

Proof of Lemma 4. We first prove (c) and then (b). (a) follows from (b) and (c) immediately.

(c) We note that under the given assumption, for $k\geq2$,

$$\left\|\mathrm{D}_xf^{-1}\frac{\mathrm{D}_y^kf}{k!}\right\|\leq\gamma(f,x)^{k-1}. \tag{2}$$

For this, we expand the Taylor series of $\frac{\mathrm{D}_y^kf}{k!}$ (with respect y) and note that its lth term is dominated by

$$\gamma(f,x)^{k+l-1}\|y-x\|^l,$$

which, by the ultrametric inequality, gives the above inequality. In this way, for $k\geq2$,

$$\left\|\mathrm{D}_yf^{-1}\frac{\mathrm{D}_y^kf}{k!}\right\|\leq\left\|\mathrm{D}_yf^{-1}\mathrm{D}_xf\right\|\left\|\mathrm{D}_xf^{-1}\frac{\mathrm{D}_y^kf}{k!}\right\|\leq\gamma(f,x)^{k-1}$$

by Lemma 3 and (2). Thus $\gamma(f,y)\leq\gamma(f,x)$. Now, due to this, the hypothesis $\gamma(f,y)\|x-y\|<1$ holds, and so, by the same argument, $\gamma(f,x)\leq\gamma(f,y)$, which is the desired equality.

(b) Arguing as in (c), we can show that

$$\left\|\mathrm{D}_xf^{-1}f(y)\right\|\leq\max\{\|\mathrm{D}_xf^{-1}f(x)+y-x\|,\gamma(f,x)\|y-x\|^2\} \tag{3}$$

by noting that the general term (of the Taylor series of $\mathrm{D}_xf^{-1}f(y)$ with respect y) is dominated by $\gamma(f,x)^{k-1}\|y-x\|^k<\gamma(f,x)\|y-x\|^2$. Now, $\beta(f,y)\leq\|\mathrm{D}_yf^{-1}\mathrm{D}_xf\|\,\|\mathrm{D}_xf^{-1}f(y)\|$, and so, by Lemma 3 and (3),

$$\beta(f,y)\leq\max\{\|\mathrm{D}_xf^{-1}f(x)+y-x\|,\gamma(f,x)\|y-x\|^2\}\leq\max\{\beta(f,x),\|y-x\|\}.$$

For the equality case, note that, by the same argument, we have $\beta(f,x)\leq\max\{\beta(f,y),\|y-x\|\}=\beta(f,y)$ where the equality on the right-hand side follows from $\beta(f,x)>\|y-x\|$. □

Proof of Lemma 5. (a) follows from combining (b) and (c), and (c) from Lemma 4 (c). We only need to show (b). We use equation (3) in the proof of Lemma 4 with $y=\mathrm{N}_f(x)$. By (3) and Lemma 3,

$$\beta(f,\mathrm{N}_f(x))\leq\max\{\|D_xf^{-1}f(x)+N_f(x)-x\|,\gamma(f,x)\|\mathrm{N}_f(x)-x\|^2\}.$$

Now, $\mathrm{N}_f(x)-x=-\mathrm{D}_xf^{-1}f(x)$, so the above becomes $\beta(f,\mathrm{N}_f(x))\leq\max\{0,\gamma(f,x)\beta(f,x)^2\}$, which gives the desired claim. □

Acknowledgements

J.G.S. is supported by a Companion Species Research Fellowship funded by the Durlacher Foundation. J.T.-C. is supported by a postdoctoral fellowship of the 2020 "Interaction" program of the Fondation Sciences Mathématiques de Paris, and partially supported by ANR JCJC GALOP (ANR-17-CE40-0009).

J.G.S. and J.T.-C. are thankful to Elias Tsigaridas for useful suggestions, and to Evgenia Lagoda for moral support.

References

[1] P. Breiding. On a p-adic Newton Method. Master's thesis, Universität Göttingen, 2013.

[2] Keith Conrad. A multivariate Hensel's lemma. Manuscript at `https://kconrad.math.uconn.edu/blurbs/gradnumthy/multivarhensel.pdf`.

[3] J.-P. Dedieu. *Points fixes, zéros et la méthode de Newton*, volume 54 of *Mathématiques & Applications*. Springer, 2006.

[4] F. Q. Gouvêa. *p-adic numbers. An Introduction.* Universitext. Springer, 2nd edition, 1997.

[5] J. Tonelli-Cueto. A p-adic Descartes solver: the Strassman solver, 3 2022. arXiv:2203.07016.