Journal of Number Theory 226 (2021) 307–357

Contents lists available at ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt

General Section

The distribution relation and inverse function theorem in arithmetic geometry ☆

Yohsuke Matsuzawa *, Joseph H. Silverman *

Department of Mathematics, Box 1917, Brown University, Providence, RI 02912, USA

ARTICLE INFO

Article history:
Received 31 August 2020
Received in revised form 21 January 2021
Accepted 5 March 2021
Available online 19 April 2021
Communicated by A. Pal

MSC:
primary 11G50
secondary 14G40, 37P30, 47J07, 58C15

Keywords:
Arithmetic distance function
Inverse function theorem
Arithmetic distribution relation

ABSTRACT

We study arithmetic distribution relations and the inverse function theorem in algebraic and arithmetic geometry, with an emphasis on versions that can be applied uniformly across families of varieties and maps. In particular, we prove two explicit versions of the inverse function theorem, the first via general distribution and separation inequalities that may be of independent interest, the second via a careful implementation of classical Newton iteration.

© 2021 Elsevier Inc. All rights reserved.

Contents

1. Introduction 308
2. Notation and terminology 309
2.1. Bounded sets 310

☆ The first author's research is supported by a JSPS Overseas Research Fellowship. The second author's research is supported by Simons Collaboration Grant #712332.

* Corresponding authors.
E-mail addresses: matsuzawa@math.brown.edu (Y. Matsuzawa), jhs@math.brown.edu (J.H. Silverman).

https://doi.org/10.1016/j.jnt.2021.03.016
0022-314X/© 2021 Elsevier Inc. All rights reserved.

2.2. Local heights attached to subschemes and arithmetic distance functions 311
3. A counterexample to the arithmetic distribution relation in [8] 312
4. Applications to quasi-integral points 314
5. Distribution, separation, and the inverse function theorem 315
5.1. The distribution inequality/relation 315
5.2. The separation inequality 321
5.3. The inverse function theorem: Version I 324
6. A version of continuity of roots 331
7. The inverse function theorem: Version II 334
8. Newton's method (non-archimedean case) 339
9. Newton's method (Archimedean case) 352
References 357

1. Introduction

In this article, we study arithmetic distribution relations and the inverse function theorem in algebraic and arithmetic geometry, with an emphasis on versions that can be applied uniformly across families of varieties and maps. Roughly speaking, such results have the following general form:

Theorem Template *(Distribution Relation). Let $\varphi : W \to V$ be a map between metric spaces. Assume that V, W, φ satisfy suitable hypotheses. Then for all $P \in W$ and all $q \in V$,*

$$d_V\big(\varphi(P), q\big) \gg\ll \prod_{Q \in \varphi^{-1}(q)} d_W(P,Q)^{e_\varphi(Q)}$$

for appropriately defined local multiplicities $e_\varphi(Q)$.

Theorem Template *(Inverse Function Theorem). Let $\varphi : W \to V$ be a map between metric spaces. Assume that V, W, φ satisfy suitable hypotheses. There are exponents $n, m > 0$ and a subset $\mathcal{R}_\varphi \subset W$ so that if $P \in W$ and $q \in V$ satisfy*

$$d_V\big(\varphi(P), q\big) \ll d_V(P, \mathcal{R}_\varphi)^n,$$

then there is a (unique) $Q \in W$ satisfying

$$\varphi(Q) = q \quad \text{and} \quad d_W(P,Q) \ll d_V\big(\varphi(P), q\big) \cdot d_V(P, \mathcal{R}_\varphi)^{-m}.$$

In both of these formulations, the implicit constants may depend on V, W, φ, and also possibly on the distance of (P, q) to some sort of "boundary" of $W \times V$.

Remark 1.1. In this paper, we use "arithmetic distance" δ instead of usual distance d since it is compatible with height theory in arithmetic geometry. Roughly speaking, the two "distance functions" δ and d are related each other via the formula $-\log d(\cdot,\cdot) = \delta(\cdot,\cdot)$.

We are interested in the case that V and W are algebraic varieties defined over either a global field or a complete field and $\varphi : W \to V$ is a generically finite morphism. In this setting, versions of the distribution relation and inverse function theorem were given in [8], where the inverse function theorem was proven by combining the distribution relation with a separation result. Unfortunately, the proof of the stated distribution relation in [8] is incorrect due to a miscalculation with multiplicities.

The primary goals of this article may thus be summarized as follows:

- Give a counterexample to the arithmetic distribution relation in [8, Proposition 6.2(b)] and explain where the error in the proof occurs. See Section 3.
- Prove that the distribution relation in [8] is correct if $\dim V = 1$ or if φ is étale. See Section 5.
- Prove that the distribution relation is valid in general as a $\ll$ inequality. Using the distribution inequality and a separation estimate, prove the inverse function theorem as in [8] with the ramification divisor replaced by a certain annihilator subscheme of W. For maps $\varphi : W \to V$ of degree d, this gives the inverse function theorem with exponents $(n, m) = (d, d-1)$. See Section 5.
- Indicate briefly why the $\dim V = 1$ and φ étale cases of the inverse function theorem suffice for various arithmetic applications, such as bounding the number of quasi-integral points on elliptic curves [9] and in orbits of maps on $\mathbb{P}^1$ [10]. See Section 4.
- Give an alternative independent proof of the inverse function theorem over complete fields with optimal exponents $(n, m) = (2, 1)$. We do everything uniformly on quasi-projective varieties, so the proof, via an étale map reduction to $\mathbb{A}^N$ and then classical Newton iteration, requires a lot of finicky computation, as well as requiring handling the non-archimedean and archimedean cases separately. See Sections 7, 8, 9.

2. Notation and terminology

We mostly follow the terminology and notation from [3] and [8], including in particular the following.

- A *variety* over a field k is an irreducible and reduced scheme of finite type over k.
- For two closed subschemes $X, Y \subset V$ of a scheme V, the sum $X + Y$ is the closed subscheme which is defined by $\mathcal{I}_X \mathcal{I}_Y$ where $\mathcal{I}_X$ and $\mathcal{I}_Y$ are ideal sheaves of X and Y.
- In Sections 3–5 we let K be a field with a complete set of inequivalent absolute values M_K that are normalized so that the product formula holds. We fix an algebraic closure $\overline{K}$ for K and write $M(\overline{K})$ for the set of absolute values on $\overline{K}$ extending the absolute values in M_K. In Sections 6–9, we let K be a field that is complete with respect to an absolute value $|\cdot|$.
- An M_K*-constant* is a function $\gamma : M(\overline{K}) \to \mathbb{R}_{\geq 0}$ such that $\gamma(v)$ depends only on the restriction $v|_K$ and $\{v|_K : \gamma(v) \neq 0\}$ is a finite subset of M_K.

- Let V/K be a variety. A function $\mu : V(\overline{K}) \times M(\overline{K}) \to \mathbb{R}$ is M_K-*bounded* if there exists an M_K-constant γ satisfying $\mu(P,v) \le \gamma(v)$ for all $(P,v) \in V(\overline{K}) \times M(\overline{K})$.
- We write $O(M_K)$ to indicate relations that hold up to an M_K-constant. For example, we use the following notation, where the right-hand property is required to hold for some $C > 0$ and some M_K-constant γ.

$$\begin{aligned} f = g + O(M_K) &\iff |f-g| \le \gamma. \\ f \le g + O(h) + O(M_K) &\iff f \le g + C|h| + \gamma. \\ f = g + O(h) + O(M_K) &\iff |f-g| \le C|h| + \gamma. \\ f \ll g + O(M_K) &\iff f \le Cg + \gamma. \\ f \gg\ll g + O(M_K) &\iff f \ll g + O(M_K) \\ &\qquad \text{and } g \ll f + O(M_K). \end{aligned}$$

2.1. Bounded sets

In this section we recall some standard definitions regarding bounded subsets of schemes defined over valued fields. For further material on bounded sets, see for example [3, Ch. 10, Sec. 1].

Definition 2.1. Let $\big(K, |\cdot|\big)$ be a complete field. Let $U = \operatorname{Spec} A$ be an affine scheme of finite type over K. A subset $B \subset U(K)$ is called *bounded* or *affine bounded* if for every $f \in A$, we have

$$\sup_{x \in B} \big|f(x)\big| < \infty.$$

Remark 2.2. Let $\varphi\colon U \longrightarrow U'$ be a morphism between affine schemes of finite type over K.

(1) If $B \subset U(K)$ is a bounded subset of U, then $f(B) \subset U'(K)$ is a bounded subset of U'.
(2) If φ is finite and $B' \subset U'(K)$ is a bounded subset of U', then the set $\varphi^{-1}(B') = \{x \in U(K) : \varphi(x) \in B'\}$ is a bounded subset of U.

Definition 2.3. Let $\big(K, |\cdot|\big)$ be a complete field. Let X be a scheme of finite type over K. A subset $B \subset X(K)$ is called *bounded* if there is a finite open affine cover $\{U_i\}_{i=1}^r$ of X and affine bounded subsets $B_i \subset U_i(K)$ such that $B = \bigcup_{i=1}^r B_i$.

Definition 2.4. Let $\big(K, |\cdot|\big)$ be a complete field. Let $X = \operatorname{Spec} A$ be an affine scheme of finite type over K. A *standard bounded subset* is a subset $B \subset X(K)$ of the form

$$B = \big\{x \in X(K) : \big|f_1(x)\big| \le b_1, \dots, \big|f_r(x)\big| \le b_r\big\},$$

where $A = K[f_1, \dots, f_r]$ and $b_1, \dots, b_r > 0$. Note that a standard bounded subset is an affine bounded subset.

Remark 2.5. More generally, for a field K with a set of absolute values M_K, one says that a subset $X \subset \mathrm{Spec}(A)(K) \times M_K$ is M_K-*bounded* if for every $f \in A$ there is an M_K-constant γ_f such that

$$\sup_{\substack{x \in \mathrm{Spec}(A)(K) \\ \text{s.t. } (x,v) \in X}} \bigl|f(x)\bigr|_v \le e^{\gamma_f(v)}.$$

2.2. Local heights attached to subschemes and arithmetic distance functions

We briefly recall from [8] the notation and construction of local height functions attached to closed subschemes, arithmetic distance functions, and local height functions attached to the boundary of a quasi-projective scheme. We refer the reader to [8] for further details. However, we note that most of the existing literature, including [8], assumes that the base scheme is irreducible and reduced, but in this paper we at times use base schemes that lack these properties. We refer the reader to [7] for an extension of the theory of local heights to schemes that are not necessarily irreducible or reduced.

Let V/K be a projective variety. We can assign to each proper closed subscheme $X \subset V$ a local height function

$$\lambda_X : V(\overline{K}) \times M_K \longrightarrow \mathbb{R} \cup \infty,$$

uniquely determined up to an M_K-bounded function by the properties that if $X = D$ is an effective divisor, then $\lambda_X = \lambda_D$ is the usual Weil local height, and if X and Y are closed subschemes, then $\lambda_{X \cap Y} = \min\{\lambda_X, \lambda_Y\}$. (The intersection $X \cap Y$ is defined to be the scheme whose ideal sheaf is $\mathcal{I}_{X \cap Y} = \mathcal{I}_X + \mathcal{I}_Y$.) These subscheme local heights have a number of natural functorial properties, as described in [8, Theorem 2.1], including functoriality $\lambda_{\varphi^* X} = \lambda_X \circ \varphi$ for morphisms φ.

Local height function λ_X is bounded below up to M_K-bounded function, so we can always assume it is non-negative if the difference by M_K-bounded function does not matter.

Let $\Delta(V) \subset V \times V$ be the diagonal. The *arithmetic distance function* on V is the local height

$$\delta_V = \lambda_{\Delta(V)}.$$

It is well-defined up to an M_K-bounded function, and satisfies a number of standard properties described in [8, Proposition 3.1], including the following two triangle inequalities, where we omit $v \in M_K$ from the notation:

$$\delta_V(P,R) \geq \min\bigl\{\delta_V(P,Q),\, \delta_V(Q,R)\bigr\}.$$
$$\lambda_X(Q) \geq \min\bigl\{\lambda_X(P),\, \delta_V(P,Q)\bigr\}.$$

If V is merely quasi-projective, we embed V in a projective variety $\overline{V}$ and define the boundary ∂V to equal $\overline{V} \smallsetminus V$ with its induced-reduced scheme structure. Then to each closed subscheme $X \subset V$ we can assign a local height function λ_X that is well-defined up to $O(\lambda_{\partial V})$. These quasi-projective local height functions inherit the functorial properties of the projective local heights, except that every relation holds only up to $O(\lambda_{\partial V})$.

Remark 2.6. We also take this opportunity to correct some typographical errors in [8]. In the definition of the union $X \cup Y$ in (iii) on page 196 of [8], the ideal should be denoted $\mathcal{I}_{X\cup Y}$. And in [8, Theorem 2.1(e)], the local height should be for the union, not the intersection, so the displayed formula in the statement of (e) on page 198 and the proof of (e) on page 199 should read

$$\max\{\lambda_X, \lambda_Y\} \leq \lambda_{X\cup Y} \leq \lambda_X + \lambda_Y.$$

Similarly, the formula in [8, Theorem 4.1(e)] should be

$$\max\{h_X, h_Y\} \leq h_{X\cup Y} \leq h_X + h_Y.$$

3. A counterexample to the arithmetic distribution relation in [8]

The arithmetic distribution relation as stated in [8] says the following:

Proposition 3.1. ([8, Proposition 6.2(b)]) *Let $\varphi : W \to V$ be a finite map of smooth quasi-projective varieties. Let $P \in W$ and $q \in V$. Then*

$$\delta_V(\varphi(P), q; v) = \sum_{\substack{Q\in W(\overline{K})\\ \varphi(Q)=q}} e_\varphi(Q)\delta_W(P,Q;v) + O(\lambda_{\partial(W\times V)}(P,q;v)).$$

Here $e_\varphi(Q)$ is the ramification index of φ at Q, so for example, if Q is not in the ramification locus of φ, then $e_\varphi(Q/q) = 1$.

Example 3.2. We give a simple counter-example to Proposition 3.1. We consider the map

$$\begin{aligned} W = \mathbb{P}^1 \times \mathbb{P}^1 &\xrightarrow{\ \varphi\ } \mathbb{P}^1 \times \mathbb{P}^1 = V \\ \bigl([x,y],[z,w]\bigr) &\longrightarrow \bigl([x^2,y^2],[z,w]\bigr). \end{aligned}$$

We fix an absolute value v on K, and for notational convenience, we drop v from the notation. We take

$$q = \bigl([0,1],[0,1]\bigr) \quad \text{and} \quad P = \bigl([a,1],[b,1]\bigr) \quad \text{with } |a| < 1 \text{ and } |b| < 1.$$

Then

$$\varphi^{-1}(q) = \{Q\} = \big\{\big([0,1],[0,1]\big)\big\} \quad \text{with } e_\varphi(Q/q) = 2,$$

and

$$\varphi(P) = \big([a^2,1],[b,1]\big).$$

Under our assumption that $|a| < 1$ and $|b| < 1$, we see that

$$\delta_V(\varphi(P),q) = -\log\max\big\{|a^2|,|b|\big\}, \tag{1}$$
$$\delta_W(P,Q) = -\log\max\big\{|a|,|b|\big\}. \tag{2}$$

Proposition 3.1 says that

$$\delta_V(\varphi(P),q) = 2\delta_W(P,Q) + O(1), \tag{3}$$

but by varying a and b, we see from (1) and (2) that the right-hand side of (3) may be strictly larger than the left-hand side. Indeed, for any fixed $1 \le \kappa \le 2$, we find that

$$\delta_V(\varphi(P),q)/\delta_W(P,Q) \longrightarrow \kappa \quad \text{as } |b| = |a|^\kappa \to 0.$$

Example 3.3. We use Example 3.2 to indicate where the proof of Proposition 3.1 in [8] goes wrong. For this example only, we use notation from [8], which may be slightly different from the notation used in the rest of this paper. (See also Remark 5.11 for another similar example.)

The proof in [8] begins with the case that $\varphi : W \to V$ is a Galois cover, say with Galois group $\mathrm{Aut}(W/V) = \{\tau_1,\dots,\tau_n\}$. It is then asserted that

$$(\varphi\times\varphi)^*\Delta(V) = \sum_{i=1}^{n}(1\times\tau_i)^*\Delta(W). \tag{4}$$

This equality of schemes is not correct. Thus in Example 3.2, writing $\mathfrak{I}(X)$ for the ideal sheaf of a scheme X, we have

$$\begin{aligned}\mathfrak{I}\big((\varphi\times\varphi)^*\Delta(V)\big) &= \big((x_1y_2)^2-(y_2x_1)^2, z_1w_2-z_2w_1\big),\\ \mathfrak{I}\big(\Delta(W)+(1\times\tau)^*\Delta(W)\big) &= (x_1y_2-y_2x_1, z_1w_2-z_2w_1)\\ &\quad\cdot(x_1y_2+y_2x_1, z_1w_2-z_2w_1).\end{aligned}$$

This gives an inclusion of ideals,

$$\mathfrak{I}\big(\Delta(W)+\tau^*\Delta(W)\big) \subset \mathfrak{I}\big((\varphi\times\varphi)^*\Delta(V)\big),$$

but the ideals are not equal, since for example the right-hand ideal contains $z_1w_2 - z_2w_1$ and the left-hand ideal does not. Of course, the underlying reduced schemes are the same, since

$$\sqrt{\mathfrak{I}\big(\Delta(W) + \tau^*\Delta(W)\big)} = \mathfrak{I}\big((\varphi \times \varphi)^*\Delta(V)\big).$$

In general, there is a distribution inequality; see Section 5.1.

Remark 3.4. As shown by Example 3.2, the arithmetic distance relation described in Proposition 3.1 is incorrect in the stated generality. There are, however, two important cases for which the proof of Proposition 3.1 in [8] is correct, and thus for which the distribution relation and its application to the inverse function theorem are valid.

- The varieties V and W are smooth and the map $\varphi : W \to V$ is étale. In this case there is no ramification, so the relevant ideal sheaves are automatically reduced
- The varieties V and W are smooth of dimension 1, in which case the ramification divisor is 0-dimensional.

See Lemma 5.6 for the subscheme formula that is key to proving the distribution relation in these two cases.

4. Applications to quasi-integral points

Various quantitative versions of the inverse function theorem have been used by the second author [2,9,10] to study integral points. In this section we briefly indicate the relevance of the present paper to these earlier results.

The paper [9] proves uniform height estimates for quasi-S-integral points on families of elliptic curves, and more generally on families of abelian varieties. The proof uses a method of Siegel that involves taking the pull-back of the multiplication-by-m map. For the application in [9], one looks at an abelian scheme, i.e., a family of abelian varieties $A \to T$ over a not necessarily complete base variety T, and applies the inverse function theorem to the multiplication-by-m map $[m] : A \to A$. The map $[m]$ is étale, so as noted in Remark 3.4, the proofs of the distribution relation and the inverse function theorem in [8] are correct, so [9] does not require the present paper.

The paper [10] proves an analogue of Siegel's integral point theorem for f-orbits of points in $\mathbb{P}^1$, where $f \in \bar{\mathbb{Q}}(z)$ is a rational map of degree at least 2. A key step in the proof uses Siegel's pull-back idea, but in this case the inverse function theorem is applied to an iterate $f^{\circ n}$ of f. The map $f^{\circ n} : \mathbb{P}^1 \to \mathbb{P}^1$ is highly ramified, but the inverse function theorem on $\mathbb{P}^1$ is much easier than the general case, and a self-contained proof of the required theorem is given in [10]. Uniform versions of the results in [10] were given by Hsia and the second author [2]. The proof includes an application of the inverse function theorem from [8] to the highly ramified iterates of a family of rational maps

$f : \mathbb{P}^1_T \to \mathbb{P}^1_T$ over a base variety T. Thus the inverse function theorem references in [2] should be replaced by references to the present paper.[1]

5. Distribution, separation, and the inverse function theorem

In this section, we prove some estimates on arithmetic distance functions involving inverse images by a finite morphism. The core inequality is the distribution inequality described in Theorem 5.2. We combine it with the separation inequality in Proposition 5.7 to prove a quantitative multivariable inverse function theorem; see Theorem 5.12. This section thus provides a correction to [8, §6].

5.1. The distribution inequality/relation

Definition 5.1. Let $\varphi\colon W \longrightarrow V$ be a finite flat morphism between schemes of finite type over a field k. Let k' be an algebraically closed field containing k. For $x \in W(k')$, define the *multiplicity* of φ at x by

$$e_\varphi(x) = \operatorname{length}_{\mathcal{O}_{W_{k'},x}} \mathcal{O}_{W_{k'},x}/\varphi^*\mathfrak{m}_{\varphi(x)}\mathcal{O}_{W_{k'},x}$$

where $W_{k'} = W \times_{\operatorname{Spec} k} \operatorname{Spec} k'$ and $V_{k'} = V \times_{\operatorname{Spec} k} \operatorname{Spec} k'$, where x and $\varphi(x)$ are closed points of $W_{k'}$ and $V_{k'}$, respectively, and where $\mathfrak{m}_{\varphi(x)}$ is the maximal ideal of $\mathcal{O}_{V_{k'},\varphi(x)}$. Note that

$$e_\varphi(x) = \dim_{k'} \mathcal{O}_{W_{k'},x}/\varphi^*\mathfrak{m}_{\varphi(x)}\mathcal{O}_{W_{k'},x},$$

since k' is algebraically closed. If φ has constant degree d, then for any $y \in V(k')$, we have

$$\sum_{x\in W(k'),\varphi(x)=y} e_\varphi(x) = d.$$

Theorem 5.2 *(Distribution inequality). Let $\varphi\colon W \longrightarrow V$ be a generically étale finite flat morphism between quasi-projective geometrically integral varieties over K.*

(1) *For all $(P, q, v) \in W(\overline{K}) \times V(\overline{K}) \times M(\overline{K})$, we have*

$$\begin{aligned}\delta_V(\varphi(P), q; v) &\\ \le \sum_{\substack{Q\in W(\overline{K})\\ \varphi(Q)=q}} & e_\varphi(Q)\delta_W(P, Q; v) + O\big(\lambda_{\partial(W\times V)}(P, q; v)\big) + O(M_K). \end{aligned} \tag{5}$$

[1] We remark that it was noted explicitly in [2, Section 3] that "it is undoubtedly possible to give a direct, albeit long and messy, proof of the desired [inverse function theorem] result." The proof of the inverse function theorem via Newton iteration in Section 7 may be viewed as such a long, messy, and direct proof.

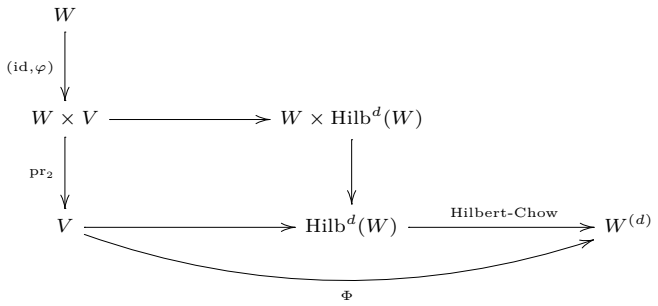


Fig. 1. The map Φ that inverts φ.

(2) *Suppose V and W are smooth. Then (5) is an equality in each of the following situations*:
- $\dim V = \dim W = 1$.
- *φ is étale.*

5.1.1. Inequalities between closed subschemes

We prove some containments between closed subschemes, from which we deduce Theorem 5.2. This can be done over an arbitrary field k, so we let $\varphi\colon W \longrightarrow V$ be a generically étale finite flat morphism between quasi-projective geometrically integral varieties over k.

Let $d = \deg \varphi$. Since φ is flat, we see that

$$W \xrightarrow{1\times\varphi} W \times V \xrightarrow{\mathrm{proj}_2} V$$

is a flat family of zero-dimensional closed subschemes of W of length d. Thus it defines a morphism $V \to \mathrm{Hilb}^d(W)$. Let Φ denote the composite $V \to \mathrm{Hilb}^d(W) \to W^{(d)}$, where $W^{(d)} = W^d/S_d$ is the d-times symmetric power of W and at a geometric point $y \in V$, we have $\Phi(y) = \{x_1, \dots, x_d\}$, where $x_1, \dots, x_d$ are the points in the inverse image of y by φ, listed with multiplicity. The second morphism is the Hilbert-Chow morphism. See Fig. 1.

We form the fiber product

$$\begin{array}{ccc} Z = V \times_{W^{(d)}} W^d & \xrightarrow{\psi} & W^d \\ \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle \pi} \\ V & \xrightarrow[\Phi]{} & W^{(d)} \end{array}$$

where π is the quotient morphism. We remark that Z need not be either reduced or irreducible.

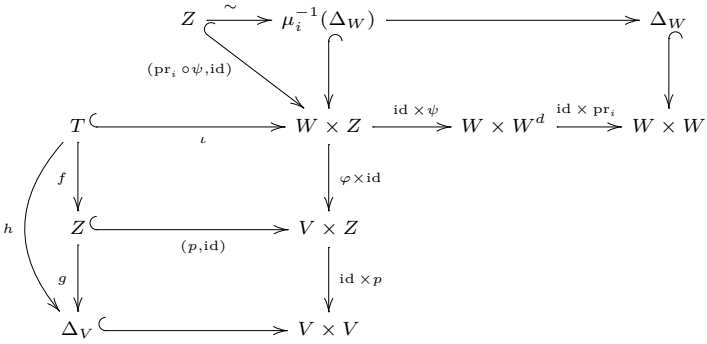


Fig. 2. Fitting together all of the spaces and maps.

Lemma 5.3. *The morphisms π and φ are universally open, i.e., every base change of π and φ is an open map.*

Proof. Since φ is a flat morphism of finite type between Noetherian schemes, it is universally open.

Note that π is finite surjective, and in particular it is universally submersive, i.e., every base change $\pi' \colon X \to Y$ of π is surjective and the topology on Y is the quotient topology of X. Let $U \subset X$ be an open subset. Then ${\pi'}^{-1}(\pi'(U)) = S_d \cdot U$, since S_d acts transitively on every geometric fiber of π'. Since $S_d \cdot U$ is open, submersivity of π' implies $\pi'(U)$ is also open. □

We consider the diagram in Fig. 2, where $\mu_i = (\mathrm{id} \times \mathrm{pr}_i) \circ (\mathrm{id} \times \psi)$, and where all of the squares are cartesian.

Lemma 5.4.

(1) $\mathrm{Support}(T) = \mathrm{Support}\Big(\sum_{i=1}^{d} \mu_i^{-1}(\Delta_W)\Big)$.
(2) *Every irreducible component of T dominates Δ_V via h and T is generically reduced, i.e., satisfies the condition* (R0) (*cf.* [5, (17. I)]).
(3) *For every irreducible component E of $W \times Z$, we have $(\varphi \times p)(E) = V \times V$ and $\mu_i(E) = W \times W$ as sets.*
(4) *If $(x, z) \in (\varphi \times p)^{-1}(\Delta_V)(\bar{k})$ is contained in exactly r of the sets*

$$\mu_1^{-1}(\Delta_W)(\bar{k}), \ \ldots, \ \mu_d^{-1}(\Delta_W)(\bar{k}),$$

then $e_\varphi(x) = r$.

Proof. (1) This follows from the construction.
(2) By Lemma 5.3, h is an open map, and therefore every irreducible component of T

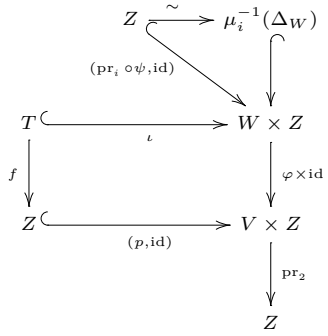


Fig. 3. Another commutative diagram.

dominates Δ_V. Let $U \subset V$ be a dense open subset over which φ is étale. Then $\varphi \times p$ is étale over $U \times U$. This implies h is étale over the dense open subset $\Delta_U \subset \Delta_V$, and hence T is generically reduced.

(3) By Lemma 5.3, the map $\varphi \times p$ is an open map. It is also finite, and hence we get $(\varphi \times p)(E) = V \times V$. The last statement follows from this and the following commutative diagram:

$$\begin{array}{ccc} W \times Z & \xrightarrow{\mu_i} & W \times W \\ {\scriptstyle \varphi \times p}\downarrow & \swarrow {\scriptstyle \varphi \times \varphi} & \\ V \times V & & \end{array} .$$

(4) This follows from the construction. □

We next consider the diagram given in Fig. 3, where we view all of the schemes as schemes over the Z at the bottom of the diagram. Let $j\colon Z_{\mathrm{red}} \to Z$ be the reduced scheme. Base change of the diagram in Fig. 3 along the morphism j gives the diagram in Fig. 4, where the subscripts $(-)_{Z_{\mathrm{red}}}$ stands for base change to Z_{red}.

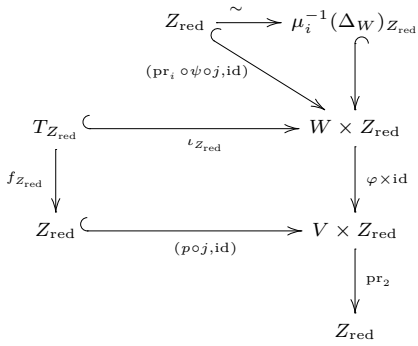
Lemma 5.5. *The scheme $T_{Z_{\mathrm{red}}}$ is reduced, and in particular, we have*

$$T_{Z_{\mathrm{red}}} \subset \sum_{i=1}^{d} \mu_i^{-1}(\Delta_W)_{Z_{\mathrm{red}}}$$

as closed subschemes of $W \times Z_{\mathrm{red}}$.

Proof. Lemma 5.4(1) tells us that these two closed subschemes have the same the support, so it is enough to show that $T_{Z_{\mathrm{red}}}$ is reduced.

First, note that $T_{Z_{\mathrm{red}}} \to T$ is a thickening of schemes. Thus by Lemma 5.4(2), the scheme $T_{Z_{\mathrm{red}}}$ satisfies (R0). By the diagram (6), the map $f_{Z_{\mathrm{red}}}$ is the base change of $\varphi \times \mathrm{id}$



(6)

Fig. 4. Base change to Z_{red}.

and therefore it is flat finite. Since Z_{red} is reduced, it satisfies the condition (S1) (cf. [5, (17. I)] for the definition of the condition (S1)), and therefore $T_{Z_{\mathrm{red}}}$ also satisfies (S1). Thus $T_{Z_{\mathrm{red}}}$ is (R0) and (S1), which is equivalent to being reduced (cf. [5, (17.I)] or [6, p183]). □

In general, Z is not reduced, and the containment $T \subset \sum_{i=1}^{d} \mu_i^{-1}(\Delta_W)$ is not true. However, with some additional assumptions, we can show that these closed subschemes are equal.

Lemma 5.6.

(1) *If $\varphi \times p$ is flat, then $W \times Z$ and T are reduced.*
(2) *If $\dim W = \dim V = 1$ and W, V are smooth, then $\varphi \times p$ is flat and*

$$T = \sum_{i=1}^{d} \mu_i^{-1}(\Delta_W).$$

(3) *If V is smooth and φ is étale, then $\varphi \times p$ is étale and*

$$T = \sum_{i=1}^{d} \mu_i^{-1}(\Delta_W).$$

Proof. (1) By the same argument in the proof of Lemma 5.4 (2), $W \times Z$ and T are (R0). By the assumption, $W \times Z$ and T are flat over $V \times V$ and Δ_V respectively. This implies they are (S1) and we are done.
(2) Since W is a smooth curve, the symmetric power $W^{(d)}$ is smooth and $\pi\colon W^d \to W^{(d)}$ is flat. This implies $p\colon Z \to V$ is flat and therefore $\varphi \times p$ is flat. Since $V \times V$ is a smooth surface, it is in particular Cohen-Macaulay, and the scheme $W \times Z$, which is flat finite over $V \times V$, is also Cohen-Macaulay. By the reducedness of $W \times Z$ and

Lemma 5.4 (3), $\mu_i^{-1}(\Delta_W) \subset W \times Z$ are effective Cartier divisors. Thus $\sum_{i=1}^d \mu_i^{-1}(\Delta_W)$ is also an effective Cartier divisor and in particular it is (S1). Since $\mu_i^{-1}(\Delta_W) \simeq Z$ are reduced and any two of them do not have common irreducible components (cf. Lemma 5.4 (4)), $\sum_{i=1}^d \mu_i^{-1}(\Delta_W)$ is generically reduced. This proves $\sum_{i=1}^d \mu_i^{-1}(\Delta_W)$ is reduced and we are done.
(3) If φ is étale, π is étale at every point over $\Phi(V)$. Thus p is étale, and therefore $\varphi \times p$ is étale. In particular, $\mu_i^{-1}(\Delta_W) \simeq Z$ are reduced. Since φ is étale, $\mu_i^{-1}(\Delta_W)$ are disjoint (cf. Lemma 5.4 (4)) and therefore $\sum_{i=1}^d \mu_i^{-1}(\Delta_W)$ is reduced and we are done. □

5.1.2. Proof of distribution inequality

Proof of Theorem 5.2. We consider the diagram in Fig. 5, where $\mu_i = (\mathrm{id} \times \mathrm{pr}_i) \circ (\mathrm{id} \times \psi)$ and all the squares are cartesian. Note that $W \times Z_{\mathrm{red}}$ is reduced since W is geometrically integral. By Lemma 5.4(3), $(\mathrm{id} \times j) \circ (\varphi \times p)$ maps all associated points of $W \times Z_{\mathrm{red}}$ to the generic point of $V \times V$, and μ_i maps them to the generic point of $W \times W$. Therefore we get

$$\begin{aligned}\delta_V \circ (\varphi \times (p \circ j)) &= (\lambda_{\Delta_V} + O(\lambda_{\partial(V\times V)})) \circ (\varphi \times (p \circ j)) \\ &= \lambda_{T_{Z_{\mathrm{red}}}} + O(\lambda_{\partial(W\times Z_{\mathrm{red}})})\end{aligned}$$

and

$$\begin{aligned}\sum_{i=1}^d \delta_W \circ \mu_i \circ (\mathrm{id} \times j) &= \sum_{i=1}^d (\lambda_{\Delta_W} + O(\lambda_{\partial(W\times W)})) \circ \mu_i \circ (\mathrm{id} \times j) \\ &= \sum_{i=1}^d \lambda_{\mu_i^{-1}(\Delta_W)_{Z_{\mathrm{red}}}} + O(\lambda_{\partial(W\times Z_{\mathrm{red}})}) \\ &= \lambda_{\sum_{i=1}^d \mu_i^{-1}(\Delta_W)_{Z_{\mathrm{red}}}} + O(\lambda_{\partial(W\times Z_{\mathrm{red}})}).\end{aligned}$$

Lemma 5.5 tells us that

$$\lambda_{T_{Z_{\mathrm{red}}}} \le \lambda_{\sum_{i=1}^d \mu_i^{-1}(\Delta_W)_{Z_{\mathrm{red}}}} + O(\lambda_{\partial(W\times Z_{\mathrm{red}})}).$$

This implies

$$\delta_V \circ (\varphi \times (p \circ j)) \le \sum_{i=1}^d \delta_W \circ \mu_i \circ (\mathrm{id} \times j) + O(\lambda_{\partial(W\times Z_{\mathrm{red}})}).$$

Here note that $p \circ j$ is a finite morphism, and in particular it is a proper morphism, so we have

$$\lambda_{\partial(W\times Z_{\mathrm{red}})} \gg\ll \lambda_{\partial(W\times V)} \circ (\mathrm{id} \times (p \circ j)).$$

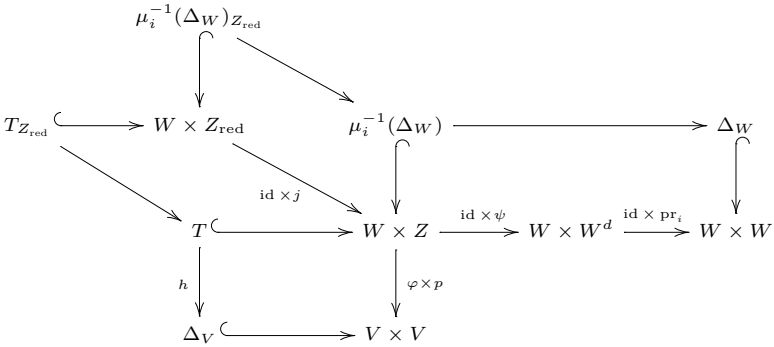


Fig. 5. Yet another commutative diagram.

Thus we get

$$\delta_V \circ (\varphi \times (p \circ j)) \leq \sum_{i=1}^{d} \delta_W \circ \mu_i \circ (\mathrm{id} \times j) + O(\lambda_{\partial(W \times V)} \circ (\mathrm{id} \times (p \circ j))). \tag{7}$$

Now take arbitrary $(P, q, v) \in W(\overline{K}) \times V(\overline{K}) \times M(\overline{K})$. Take a point $z \in Z_{\mathrm{red}}(\overline{K})$ such that $p \circ j(z) = q$. Then the definition of Z tells us that $\psi \circ j(z)$ is a point of the form

$$(Q_1, \ldots, Q_d) \in W(\overline{K})^d$$

where Q_i's are preimages of q by φ counted with multiplicity. Plugging the point (P, z) into (7), we get

$$\begin{aligned} \delta_V(\varphi(P), q; v) &\leq \sum_{i=1}^{d} \delta_W(P, Q_i; v) + O(\lambda_{\partial(W \times V)}(P, q; v)) \\ &= \sum_{\substack{Q \in W(\overline{K}) \\ \varphi(Q) = q}} e_\varphi(Q) \delta_W(P, Q; v) + O(\lambda_{\partial(W \times V)}(P, q; v)). \end{aligned}$$

For the second statement, by Lemma 5.6, if W, V are smooth, and if either they have dimension 1 or if φ is étale, then we have the equality

$$\lambda_{T_{Z_{\mathrm{red}}}} = \lambda_{\sum_{i=1}^{d} \mu_i^{-1}(\Delta_W)_{Z_{\mathrm{red}}}} + O(\lambda_{\partial(W \times Z_{\mathrm{red}})}).$$

Thus by the same argument, we get the desired equality. □

5.2. The separation inequality

In this section we quantify the assertion that distinct inverse image points cannot be too close to one another.

Proposition 5.7 *(Separation). Let W and V be quasi-projective varieties over K, and let $\varphi\colon W \to V$ be a generically étale, generically finite morphism. Let $\mathcal{I}$ be the annihilator ideal sheaf of $\Omega_{W/V}$. Then for $v \in M(\overline{K})$ and for all points $Q, Q' \in W(\overline{K})$ such that $\varphi(Q) = \varphi(Q')$ and $Q \neq Q'$:*

(1) *In general we have*

$$\delta_W(Q, Q'; v) \leq \lambda_{\mathcal{I}}(Q; v) + O(\lambda_{\partial(W\times W)}(Q, Q'; v)) + O(M_K).$$

(2) *If φ is proper, then*

$$\delta_W(Q, Q'; v) \leq \lambda_{\mathcal{I}}(Q; v) + O(\lambda_{\partial V}(\varphi(Q); v)) + O(M_K).$$

(3) *Let $\mathrm{Fit}_0(\Omega_{W/V})$ be the 0-th Fitting ideal of $\Omega_{W/V}$. Then (1) and (2) are true with $\lambda_{\mathcal{I}}$ replaced by $\lambda_{\mathrm{Fit}_0(\Omega_{W/V})}$.*
(4) *Assume that V and W are smooth and that φ is finite, and let $R(\varphi)$ be the ramification divisor of φ. Then (1) and (2) are true with $\lambda_{\mathcal{I}}$ replaced by $\lambda_{R(\varphi)}$.*

Lemma 5.8. *Let k be a field. Let W, V be varieties over k and let $\varphi\colon W \to V$ be a generically étale generically finite morphism. Let $\mathcal{I}$ be the annihilator ideal sheaf of $\Omega_{W/V}$. Then*

$$(\mathrm{pr}_1^*\mathcal{I})\mathcal{I}_{\Delta(W)} + \mathcal{I}_{\Delta(W)}^2 \subset (\varphi\times\varphi)^*\mathcal{I}_{\Delta(V)} + \mathcal{I}_{\Delta(W)}^2$$

on $W \times W$. In particular, since the 0-th Fitting ideal sheaf $\mathrm{Fit}_0(\Omega_{W/V})$ is contained in $\mathcal{I}$, we have

$$(\mathrm{pr}_1^*\,\mathrm{Fit}_0(\Omega_{W/V}))\mathcal{I}_{\Delta(W)} + \mathcal{I}_{\Delta(W)}^2 \subset (\varphi\times\varphi)^*\mathcal{I}_{\Delta(V)} + \mathcal{I}_{\Delta(W)}^2.$$

Moreover, if W, V are smooth and φ is finite, then

$$(\mathrm{pr}_1^*\mathcal{I}_{R(\varphi)})\mathcal{I}_{\Delta(W)} + \mathcal{I}_{\Delta(W)}^2 \subset (\varphi\times\varphi)^*\mathcal{I}_{\Delta(V)} + \mathcal{I}_{\Delta(W)}^2 \tag{8}$$

where $\mathcal{I}_{R(\varphi)}$ is the ideal sheaf of the ramification divisor $R(\varphi)$.

Proof. Consider the exact sequence

$$\varphi^*\Omega_{V/k} \longrightarrow \Omega_{W/k} \longrightarrow \Omega_{W/V} \longrightarrow 0. \tag{9}$$

Since the sheaf of differentials is the conormal sheaf of the diagonal, we can rewrite the sequence (9) as

$$(\varphi\times\varphi)^*\mathcal{I}_{\Delta(V)}/\mathcal{I}_{\Delta(W)}(\varphi\times\varphi)^*\mathcal{I}_{\Delta(V)} \longrightarrow \mathcal{I}_{\Delta(W)}/\mathcal{I}_{\Delta(W)}^2 \longrightarrow \Omega_{W/V} \longrightarrow 0.$$

Then using the fact that $\mathcal{I}\Omega_{W/V} = 0$, we get

$$(\mathrm{pr}_1^*\,\mathcal{I})\mathcal{I}_{\Delta(W)} + \mathcal{I}_{\Delta(W)}^2 \subset (\varphi\times\varphi)^*\mathcal{I}_{\Delta(V)}\mathcal{O}_{W\times W} + \mathcal{I}_{\Delta(W)}^2.$$

The gives everything except for the last assertion of the lemma, which follows from the fact that if W and V are smooth and φ is finite, then $\mathrm{Fit}_0(\Omega_{W/V}) = \mathcal{I}_{R(\varphi)}$. □

Remark 5.9. Let $\mathcal{I}$ is the annihilator ideal sheaf of $\Omega_{W/V}$. Then if W and V are smooth and φ is finite, we have

$$\mathcal{I}_{R(\varphi)} = \mathrm{Fit}_0(\Omega_{W/V}) \subset \mathcal{I}.$$

Further, if we let $V\big(\mathrm{Fit}_0(\Omega_{W/V})\big)$ denote the closed subset defined by the ideal sheaf $\mathrm{Fit}_0(\Omega_{W/V})$, then we also have

$$V\big(\mathrm{Fit}_0(\Omega_{W/V})\big) = \mathrm{Support}\,\Omega_{W/V}.$$

Thus if $R(\varphi)$ is reduced, then $\mathcal{I}_{R(\varphi)} = \mathcal{I}$.

Remark 5.10. When W and V are smooth curves, it follows easily from the proof of Lemma 5.8 that we have an equality

$$(\mathrm{pr}_1^*\,\mathcal{I}_{R(\varphi)})\mathcal{I}_{\Delta(W)} + \mathcal{I}_{\Delta(W)}^2 = (\varphi\times\varphi)^*\mathcal{I}_{\Delta(V)} + \mathcal{I}_{\Delta(W)}^2.$$

Remark 5.11. The containment (8) in Lemma 5.8 may be strict in general in dimension greater than 1. For example, let

$$\varphi\colon \mathbb{A}^2 \longrightarrow \mathbb{A}^2, \quad (x,y)\mapsto (x^2+xy+y^2, xy+1).$$

Then the defining function of the ramification divisor is $2(x^2-y^2)$, and

$$\begin{aligned}(\mathrm{pr}_1^*\,\mathcal{I}_{R(\varphi)})\mathcal{I}_{\Delta(W)} + \mathcal{I}_{\Delta(W)}^2 &\\ = (x^2-y^2)(x-z,y-w) + (x-z,y-w)^2, &\qquad (10)\\ (\varphi\times\varphi)^*\mathcal{I}_{\Delta(V)} + \mathcal{I}_{\Delta(W)}^2 &\\ = \big((x^2+xy+y^2)-(z^2+zw+w^2), (xy+1)-(zw+1)\big) &\\ + (x-z,y-w)^2. &\qquad (11)\end{aligned}$$

It is easy to see that $xy-zw$ is in the ideal (11), but that it is not in (10); cf. Example 3.3.

Proof of Proposition 5.7. The inclusion of ideal sheaves in Lemma 5.8 translates into the inequality of local height functions

$$\min\{\lambda_{\mathcal{I}}\circ\mathrm{pr}_1 + \delta_W, 2\delta_W\} \geq \min\{\delta_V\circ(\varphi\times\varphi), 2\delta_W\} - C_1\lambda_{\partial(W\times W)} + O(M_K).$$

For $Q, Q' \in W(\overline{K})$ such that $\varphi(Q) = \varphi(Q')$ and $Q \neq Q'$, we have

$$\delta_V((\varphi \times \varphi)(Q, Q'); v) = \delta_V(q, q) = \infty,$$

and therefore

$$\begin{aligned}\lambda_{\mathcal{I}}(Q; v) + \delta_W(Q, Q'; v) \\ \geq 2\delta_W(Q, Q'; v) - C_1\lambda_{\partial(W\times W)}(Q, Q'; v) + O(M_K).\end{aligned}$$

This implies

$$\lambda_{\mathcal{I}}(Q; v) \geq \delta_W(Q, Q'; v) - C_1\lambda_{\partial(W\times W)}(Q, Q'; v) + O(M_K),$$

which proves the first statement.

If we suppose that φ is proper, then we have

$$\lambda_{\partial(W\times W)} \gg\ll \lambda_{\partial(V\times V)} \circ (\varphi \times \varphi) + O(M_K),$$

and we also always have

$$\lambda_{\partial(V\times V)} \gg\ll \lambda_{\partial V} \circ \mathrm{pr}_1 + \lambda_{\partial V} \circ \mathrm{pr}_2 + O(M_K).$$

These imply that for all $Q, Q' \in W(\overline{K})$ such that $\varphi(Q) = \varphi(Q')$, we have

$$\lambda_{\partial(W\times W)}(Q, Q'; v) \gg\ll \lambda_{\partial V}(\varphi(Q); v) + O(M_K).$$

Thus we get the second statement □

5.3. The inverse function theorem: Version I

In this section we prove a version of the inverse function theorem. The statement and proof are modeled after the statement and proof in [8], but corrected by the use of the alternative ramification local height coming from the corrected versions of the distribution and separation lemmas proven in Sections 5.1 and 5.2.

Theorem 5.12 *(Inverse function theorem: Version I). Fix/define the following quantities*:

- V *and* W *are quasi-projective geometrically integral varieties defined over* K.
- $\varphi\colon W \to V$ *is a generically étale finite flat surjective morphism of degree* d *defined over* K.
- $\mathrm{Ann}(\Omega_{W/V})$ *is the annihilator ideal sheaf of* $\Omega_{W/V}$.
- $A(\varphi) \subset W$ *is the closed subscheme defined by* $\mathrm{Ann}(\Omega_{W/V})$.
- δ_W *and* δ_V *are arithmetic distance functions on* W *and* V.

- $\lambda_{A(\varphi)}$ *is a local height function associated with* $A(\varphi)$.
- $\lambda_{\partial(W\times V)}$ *is a local height boundary function for* $W \times V$.

(a) *There exist constants* $C_2, C_3 \in \mathbb{R}_{>0}$ *and* M_K*-constants* C_4, C_5 *such that the following holds*:
If the triple $(P, q, v) \in W(\overline{K}) \times V(\overline{K}) \times M(\overline{K})$ *satisfies*

$$\delta_V(\varphi(P), q; v) \geq d\lambda_{A(\varphi)}(P; v) + C_2\lambda_{\partial(W\times V)}(P, q; v) + C_4(v),$$

then there exists a point $Q \in W(\overline{K})$ *satisfying*

$$\begin{aligned} \varphi(Q) = q \quad &\textit{and} \\ \delta_W(P, Q; v) \geq \delta_V\big(\varphi(P), q; v\big) &- (d-1)\lambda_{A(\varphi)}(P; v) \\ &- C_3\lambda_{\partial(W\times V)}(P, q; v) - C_5(v). \end{aligned}$$

(b) *If we take* C_4 *to be an appropriate positive real number, instead of an* M_K*-constant, and if we also assume that* $P \notin A(\varphi)$*, then the point* Q *in* (a) *is unique.*

Example 5.13. The point Q may depend on the absolute value v, as well as on P and q. We illustrate with an example, which for convenience we write using affine coordinates. We start with distinct non-archimedean absolute values v and w that do not divide 2, and we choose elements $\alpha, \beta \in K$ satisfying

$$|\alpha|_v < 1, \quad |\alpha|_w = 1, \quad |\beta|_v = 1, \quad |\beta|_w < 1.$$

We consider the map $\varphi(x) = x^2$ and points

$$P = 1, \quad Q = \frac{\alpha^n - \beta^n}{\alpha^n + \beta^n}, \quad q = Q^2.$$

The ramification divisor of φ is $R(\varphi) = (0)$, so for any absolute value u we have

$$\lambda_{R(\varphi)}(P; u) = \lambda_{(0)}(1; u) = -\log|1|_u = 0.$$

We next determine the distance from $\varphi(P)$ to q. Taking u to be v or w, we compute

$$\begin{aligned} \delta_{\mathbb{P}^1}\big(\varphi(P), q; u\big) = -\log|1 - Q^2|_u &= -\log\left|\frac{4\alpha^n\beta^n}{(\alpha^n+\beta^n)^2}\right|_u \\ &= \begin{cases} -n\log|\alpha|_v & \text{if } u = v, \\ -n\log|\beta|_w & \text{if } u = w. \end{cases} \end{aligned}$$

Hence if n is sufficiently large, then we have satisfied the assumptions of the inverse function theorem for both v and w, so in each case there is a point in $\varphi^{-1}(q) = \{\pm Q\}$ that is appropriately close to P. But for $u \in \{v, w\}$, we find that

$$\begin{aligned}\delta_{\mathbb{P}^1}\big(P, \pm Q; u\big) = -\log|-1 \pm Q|_u &= -\log\left|-1 \pm \frac{\alpha^n - \beta^n}{\alpha^n + \beta^n}\right|_u \\ &= \begin{cases} -\log\left|\dfrac{2\alpha^n}{\alpha^n + \beta^n}\right|_u & \text{if sign is } -, \\ -\log\left|\dfrac{2\beta^n}{\alpha^n + \beta^n}\right|_u & \text{if sign is } +, \end{cases} \\ &= \begin{cases} -n\log|\alpha|_v & \text{if } u = v, \text{ sign is } -, \\ -n\log|\beta|_w & \text{if } u = w, \text{ sign is } +, \\ 0 & \text{otherwise.} \end{cases}\end{aligned}$$

Hence for the v-adic absolute value, the inverse function theorem requires us to take $-Q$, while for the w-adic absolute value we must take Q.

Remark 5.14. The morphism φ is étale outside $A(\varphi)$. In particular, $e_\varphi(Q) = 1$ for all points $Q \in (W \smallsetminus A(\varphi))(\overline{K})$.

Remark 5.15. Suppose that W and V are smooth. Then in general, we have $A(\varphi) \subset R(\varphi)$ as closed subschemes. Indeed,

$$\varphi^*\Omega_V \longrightarrow \Omega_W \longrightarrow \Omega_{W/V} \longrightarrow 0$$

is a locally free resolution of $\Omega_{W/V}$ and the ideal of $R(\varphi)$ is locally generated by the determinant of the first map. Hence

$$\lambda_{A(\varphi)} \leq \lambda_{R(\varphi)} + O(\lambda_{\partial W}) + O(M_K),$$

so if W and V are smooth, then Theorem 5.12 is true with $\lambda_{R(\varphi)}$ in place of $\lambda_{A(\varphi)}$.

Proof of Theorem 5.12. We fix boundary functions $\lambda_{\partial W}$ and $\lambda_{\partial V}$. Since

$$\lambda_{\partial(W\times V)} \gg\ll \lambda_{\partial W} \circ \mathrm{pr}_1 + \lambda_{\partial V} \circ \mathrm{pr}_2 + O(M_K),$$

we may take $\lambda_{\partial(W\times V)} = \lambda_{\partial W} \circ \mathrm{pr}_1 + \lambda_{\partial V} \circ \mathrm{pr}_2$. For simplicity, we also take other boundary functions on products in this form: for instance $\lambda_{\partial W^3} = \lambda_{\partial W} \circ \mathrm{pr}_1 + \lambda_{\partial W} \circ \mathrm{pr}_2 + \lambda_{\partial W} \circ \mathrm{pr}_3$. Note also that we have

$$\lambda_{\partial W} \gg\ll \lambda_{\partial V} \circ \varphi + O(M_K),$$

since φ is a proper morphism.

In the following, we write C_i for positive real constants and $C_i(v)$ for M_K-constants. These constants are allowed to depend on the varieties W and V, on the map φ, and on our choice of local height and distance functions $\delta_W, \delta_V, \lambda_{A(\varphi)}, \lambda_{\partial W}, \lambda_{\partial V}$.

By Theorem 5.2 and Proposition 5.7, we have:

• **Distribution inequality**:

$$\delta_V(\varphi(P), q; v) \le \sum_{Q \in W(\overline{K}), \varphi(Q)=q} e_\varphi(Q)\delta_W(P, Q; v) + C_6\lambda_{\partial(W\times V)}(P, q; v) + C_7(v)$$

for all $(P, q, v) \in W(\overline{K}) \times V(\overline{K}) \times M(\overline{K})$.

• **Separation inequality**:

$$\delta_W(Q, Q'; v) \le \lambda_{A(\varphi)}(Q; v) + C_8\lambda_{\partial V}(\varphi(Q); v) + C_9(v)$$

for all $(Q, Q', v) \in W(\overline{K}) \times W(\overline{K}) \times M(\overline{K})$ such that $\varphi(Q) = \varphi(Q')$ and $Q \neq Q'$.

Let us fix arbitrary $(P, q, v) \in W(\overline{K}) \times V(\overline{K}) \times M(\overline{K})$. For $Q, Q' \in W(\overline{K})$ such that $\varphi(Q) = \varphi(Q') = q$ and $Q \neq Q'$, by the triangle inequality and the separation inequality, we have

$$\begin{aligned}
&\min\{\delta_W(P, Q'; v), \delta_W(P, Q; v)\} \\
&\quad \le \delta_W(Q, Q'; v) + C_{10}\lambda_{\partial(W^3)}(P, Q, Q'; v) + C_{11}(v) \\
&\quad \le \lambda_{A(\varphi)}(Q; v) + C_8\lambda_{\partial V}(\varphi(Q); v) + C_{10}\lambda_{\partial(W^3)}(P, Q, Q'; v) + C_{12}(v) \\
&\quad \le \lambda_{A(\varphi)}(Q; v) + C_{13}\lambda_{\partial V}(q; v) + C_{10}\lambda_{\partial W}(P; v) + C_{14}(v).
\end{aligned}$$

Here we use the relation $\lambda_{\partial(W^3)}(P, Q, Q'; v) \le \lambda_{\partial W}(P; v) + C_{15}\lambda_{\partial V}(q; v) + O(M_K)$.

Let $Q \in W(\overline{K})$ be a point with $\varphi(Q) = q$ such that

$$\delta_W(P, Q; v) = \max\{\delta_W(P, Q'; v) : Q' \in W(\overline{K}),\ \varphi(Q') = q\}.$$

Then for every $Q' \in W(\overline{K})$ such that $\varphi(Q') = q$ and $Q' \neq Q$, we have

$$\begin{aligned}
\delta_W(P, Q'; v) &= \min\{\delta_W(P, Q'; v), \delta_W(P, Q; v)\} \\
&\le \lambda_{A(\varphi)}(Q; v) + C_{13}\lambda_{\partial V}(q; v) + C_{10}\lambda_{\partial W}(P; v) + C_{14}(v). \qquad (12)
\end{aligned}$$

Bounding the right-hand side of the distribution inequality by (12), we get

$$\begin{aligned}
&\delta_V(\varphi(P), q; v) \\
&\quad \le e_\varphi(Q)\delta_W(P, Q; v) \\
&\qquad + \big(d - e_\varphi(Q)\big)\big(\lambda_{A(\varphi)}(Q; v) + C_{13}\lambda_{\partial V}(q; v) + C_{10}\lambda_{\partial W}(P; v) + C_{14}(v)\big) \\
&\qquad + C_6\lambda_{\partial(W\times V)}(P, q; v) + C_7(v) \\
&\quad \le e_\varphi(Q)\delta_W(P, Q; v) + (d - e_\varphi(Q))\lambda_{A(\varphi)}(Q; v) \\
&\qquad + C_{16}\lambda_{\partial(W\times V)}(P, q; v) + C_{17}(v).
\end{aligned}$$

In summary, we have proved that for all

$$(P, q, v) \in W(\overline{K}) \times V(\overline{K}) \times M(\overline{K})$$

there is a point $Q \in W(\overline{K})$ satisfying $\varphi(Q) = q$ such that

$$\begin{aligned}
\delta_V(\varphi(P), q; v) \le e_\varphi(Q)\delta_W(P, Q; v) + (d - e_\varphi(Q))\lambda_{A(\varphi)}(Q; v) \\
+ C_{16}\lambda_{\partial(W\times V)}(P, q; v) + C_{17}(v)
\end{aligned} \tag{13}$$

and

$$\delta_W(P, Q; v) = \max\big\{\delta_W(P, Q'; v) : Q' \in W(\overline{K}),\ \varphi(Q') = q\big\}.$$

Next, we compare $\lambda_{A(\varphi)}(Q; v)$ and $\lambda_{A(\varphi)}(P; v)$. By the triangle inequality, we have

$$\begin{aligned}
&\min\{\lambda_{A(\varphi)}(Q; v), \delta_W(P, Q; v)\} \\
&\qquad \le \lambda_{A(\varphi)}(P; v) + C_{18}\lambda_{\partial(W\times W)}(P, Q; v) + C_{19}(v).
\end{aligned}$$

Case 1. Suppose that

$$\delta_W(P, Q; v) \le \lambda_{A(\varphi)}(P; v) + C_{18}\lambda_{\partial(W\times W)}(P, Q; v) + C_{19}(v).$$

By the distribution inequality and the choice of Q, we have

$$\begin{aligned}
&d\delta_W(P, Q; v) + C_6\lambda_{\partial(W\times V)}(P, q; v) + C_7(v) \\
&\quad \ge \sum_{Q'\in W(\overline{K}), \varphi(Q')=q} e_\varphi(Q')\delta_W(P, Q'; v) + C_6\lambda_{\partial(W\times V)}(P, q; v) + C_7(v) \\
&\quad \ge \delta_V(\varphi(P), q; v).
\end{aligned}$$

Thus we get

$$\begin{aligned}\delta_V(\varphi(P),q;v) &\le d\left(\lambda_{A(\varphi)}(P;v)+C_{18}\lambda_{\partial(W\times W)}(P,Q;v)+C_{19}(v)\right)\\ &\quad +C_6\lambda_{\partial(W\times V)}(P,q;v)+C_7(v)\\ &\le d\lambda_{A(\varphi)}(P;v)+C_{20}\lambda_{\partial(W\times V)}(P,q;v)+C_{21}(v).\end{aligned}$$

Hence if we assume that

$$\delta_V(\varphi(P),q;v)\ge d\lambda_{A(\varphi)}(P;v)+C_2\lambda_{\partial(W\times V)}(P,q;v)+C_4(v), \tag{14}$$

then we get

$$\begin{aligned}&d\lambda_{A(\varphi)}(P;v)+C_2\lambda_{\partial(W\times V)}(P,q;v)+C_4(v)\\ &\qquad\le d\lambda_{A(\varphi)}(P;v)+C_{20}\lambda_{\partial(W\times V)}(P,q;v)+C_{21}(v).\end{aligned}$$

Now if $P\in A(\varphi)$, or equivalently if $\lambda_{A(\varphi)}(P;v)=\infty$, then (14) implies that $q=\varphi(P)$, so we can simply take $Q=P$. Thus we may assume that $P\notin A(\varphi)$, in which case we get

$$C_2\lambda_{\partial(W\times V)}(P,q;v)+C_4(v)\le C_{20}\lambda_{\partial(W\times V)}(P,q;v)+C_{21}(v). \tag{15}$$

Thus if we take $C_2>C_{20}$ and take C_4 large enough as an M_K-constant, then for any $(P,q,v)\in W(\overline{K})\times V(\overline{K})\times M(\overline{K})$, either the inequality (15) does not hold, or else

$$\lambda_{\partial(W\times V)}(P,q;v)=C_4(v)=C_{21}(v)=0.$$

In the latter case, we have

$$\delta_V(\varphi(P),q;v)=d\lambda_{A(\varphi)}(P;v)$$

and

$$\begin{aligned}&\delta_W(P,Q;v)\\ &=\lambda_{A(\varphi)}(P;v)+C_{18}\lambda_{\partial(W\times W)}(P,Q;v)+C_{19}(v)\\ &=\delta_V(\varphi(P),q;v)-(d-1)\lambda_{A(\varphi)}(P;v)+C_{18}\lambda_{\partial(W\times W)}(P,Q;v)+C_{19}(v)\\ &\ge\delta_V(\varphi(P),q;v)-(d-1)\lambda_{A(\varphi)}(P;v)-C_{22}\lambda_{\partial(W\times V)}(P,q;v)-C_{23}(v).\end{aligned}$$

This proves the first statement for Case 1.

We also note that if we take C_4 to be a large positive constant, rather than an M_K-constant, and if we also assume that $P\notin A(\varphi)$, then Case 1 does not happen.

Case 2. We are reduced to the case that for appropriate choices for the constants C_2 and C_4, we may assume that

$$\begin{aligned}\lambda_{A(\varphi)}(Q;v) &\le \lambda_{A(\varphi)}(P;v) + C_{18}\lambda_{\partial(W\times W)}(P,Q;v) + C_{19}(v)\\ &\le \lambda_{A(\varphi)}(P;v) + C_{24}\lambda_{\partial(W\times V)}(P,q;v) + C_{25}(v). \end{aligned} \tag{16}$$

By the same reasoning as earlier, we may assume that $P \notin A(\varphi)$, or equivalently, that $\lambda_{A(\varphi)}(P;v) < \infty$. Then (16) tells us that $Q \notin A(\varphi)$ and $e_\varphi(Q) = 1$.

Applying (13) and (16), we get

$$\begin{aligned}&\delta_V(\varphi(P),q;v)\\ &\le \delta_W(P,Q;v) + (d-1)\left(\lambda_{A(\varphi)}(P;v) + C_{24}\lambda_{\partial(W\times V)}(P,q;v) + C_{25}(v)\right)\\ &\qquad + C_{16}\lambda_{\partial(W\times V)}(P,q;v) + C_{17}(v)\\ &\le \delta_W(P,Q;v) + (d-1)\lambda_{A(\varphi)}(P;v) + C_{26}\lambda_{\partial(W\times V)}(P,q;v) + C_{27}(v),\end{aligned}$$

which is what we want. This completes the proof of (a), i.e., the existence of a point Q having the specified properties.

(b) It remains to show that if we take C_4 to be a sufficiently large absolute constant, rather than an M_K-constant, then Q is unique. Assume that $P \notin A(\varphi)$. We first choose $C_2,\dots,C_5$ so that (a) holds, but we then replace C_4 with a large positive real number. As noted earlier, this means that Case 1 in the proof of (a) does not occur.

Suppose that there is a point $Q' \in W(\overline{K})$ such that $\varphi(Q') = q$ and $Q' \ne Q$ and Q' satisfies

$$\begin{aligned}&\delta_V(\varphi(P),q;v)\\ &\quad\le \delta_W(P,Q';v) + (d-1)\lambda_{A(\varphi)}(P;v) + C_3\lambda_{\partial(W\times V)}(P,q;v) + C_5(v).\end{aligned}$$

This estimate and (12) and (16) yield

$$\begin{aligned}&d\lambda_{A(\varphi)}(P;v) + C_2\lambda_{\partial(W\times V)}(P,q;v) + C_4\\ &\le \delta_V(\varphi(P),q;v)\\ &\le \delta_W(P,Q';v) + (d-1)\lambda_{A(\varphi)}(P;v) + C_3\lambda_{\partial(W\times V)}(P,q;v) + C_5(v)\\ &\le d\lambda_{A(\varphi)}(P;v) + C_{28}\lambda_{\partial(W\times V)}(P,q;v) + C_{29}(v).\end{aligned}$$

If we replace C_2 and C_4 with larger constants that depend on C_{28}, C_{29}, and $\lambda_{\partial(W\times V)}$, we obtain a contradiction. □

Remark 5.16. The proof of Theorem 5.12 shows that there is an M_K-constant γ such that if $C_4\colon M(\overline{K}) \longrightarrow \mathbb{R}_{\ge 0}$ satisfies the strict inequality $C_4(v) > \gamma(v)$ for all $v \in M(\overline{K})$, and if we assume that $P \notin A(\varphi)$, then we still get the uniqueness of Q.

6. A version of continuity of roots

In Sections 6–9, we shift our focus to a field K that is complete with respect to a fixed absolute value. We thus let $\big(K, |\cdot|\big)$ be a complete field, and since there is only one absolute value, we drop v from the our notation for local heights and arithmetic distance functions.

A theorem such as Theorem 5.12 may be viewed as a quantitative higher dimensional variant of the classical theorem that the roots of a univariate polynomial vary continuously with its coefficients. In our next result, we apply Theorem 5.12 to a certain morphism to prove such a result. We use this later to prove a stronger inverse function theorem; see Theorem 7.1.

Proposition 6.1. *Let $\big(K, |\cdot|\big)$ be a complete field with a non-trivial absolute value $|\cdot|$. Let $D \in \mathbb{R}_{>0}$ and $n \in \mathbb{Z}_{>0}$. Then there are constants $C_{30}, C_{31} \in \mathbb{R}_{>0}$ such that the following holds. Suppose that*:

- *$f, g \in K[t]$ are monic polynomials of degree n;*
- *$|f| \le D$ and $|g| \le D$, where $|f|$ and $|g|$ are the Gauss norms[2];*
- *There is an $\alpha \in K$ such that*

$$f(\alpha) = 0 \quad \text{and} \quad |f - g| \le e^{-C_{30}} |f'(\alpha)|^n. \tag{17}$$

Then there is $\beta \in K$ such that

$$g(\beta) = 0 \quad \text{and} \quad |\alpha - \beta| |f'(\alpha)|^{n-1} \le e^{C_{31}} |f - g|.$$

Remark 6.2. See Corollary 7.4, for a stronger version Proposition 6.1. But when $\big(K, |\cdot|\big)$ is non-archimedean, we need Proposition 6.1 to prove Proposition 8.1, to which we ultimately reduce Corollary 7.4.

Proof of Proposition 6.1. We consider the following morphism:

$$\begin{aligned} \varphi\colon \mathbb{A}^{n+1}_K &\longrightarrow \mathbb{A}^{n+1}_K \\ (x_0, \dots, x_{n-1}, t) &\longmapsto (x_0, \dots, x_{n-1}, t^n + x_{n-1}t^{n-1} + \cdots + x_1 t + x_0). \end{aligned}$$

The map φ is a generically étale, finite, surjective morphism of degree n. Its ramification divisor is

$$R(\varphi) = \big(J := nt^{n-1} + (n-1)x_{n-1}t^{n-2} + \cdots + x_1 = 0\big) \subset \mathbb{A}^{n+1}_K.$$

[2] The Gauss norm of a polynomial is the maximum of the absolute values of its coefficients.

We fix an algebraic closure $\overline{K}$ of K and an extension of $|\cdot|$ to $\overline{K}$, which we also denote by $|\cdot|$. We consider the affine bounded subset

$$B := \{\xi \in \mathbb{A}^{n+1}(\overline{K}) \mid \|\xi\| \le D\},$$

where $\|\cdot\|$ denotes the sup norm on the coordinates. Since φ is finite, $\varphi^{-1}(B) \subset \mathbb{A}^{n+1}(\overline{K})$ is also an affine bounded subset. We apply Theorem 5.12 to $\varphi_{\overline{K}}$ on various bounded subsets, where we note that since we are over a local field, the set M_K consists of a single absolute value. We may take

$$\begin{aligned}
\delta_{\mathbb{A}^{n+1}}(\xi,\eta) &= \log \frac{1}{\|\xi-\eta\|} && \text{on } \big(\varphi^{-1}(B)\cup B\big)\times\big(\varphi^{-1}(B)\cup B\big),\\
\lambda_{R(\varphi)}(\xi) &= \log \frac{1}{|J(\xi)|} && \text{on } \varphi^{-1}(B),\\
\lambda_{\partial(\mathbb{A}^{n+1}\times\mathbb{A}^{n+1})}(\xi,\eta) &= 0 && \text{on } \varphi^{-1}(B)\times B.
\end{aligned}$$

Theorem 5.12 and Remark 5.15 tell us that there are positive constants C_{30}, C_{31} such that for $\xi \in \varphi^{-1}(B)$ and $\eta \in B$, if

$$\log \frac{1}{\|\varphi(\xi)-\eta\|} \ge n \log \frac{1}{|J(\xi)|} + C_{30},$$

then there is a $\zeta \in \varphi^{-1}(B)$ satisfying $\varphi(\zeta) = \eta$ and

$$\log \frac{1}{\|\xi-\zeta\|} + (n-1)\log \frac{1}{|J(\xi)|} + C_{31} \ge \log \frac{1}{\|\varphi(\xi)-\eta\|}.$$

Rewriting this, we find that if $\xi \in \varphi^{-1}(B)$ and $\eta \in B$ satisfy

$$\|\varphi(\xi)-\eta\| \le e^{-C_{30}}|J(\xi)|^n,$$

then there is $\zeta \in \varphi^{-1}(B)$ satisfying

$$\varphi(\zeta) = \eta \quad \text{and} \quad \|\xi-\zeta\||J(\xi)|^{n-1} \le e^{C_{31}}\|\varphi(\xi)-\eta\|.$$

We apply this to the points

$$\xi = (a_0,\dots,a_{n-1},\alpha) \quad \text{and} \quad \eta = (b_0,\dots,b_{n-1},0)$$

associated to the polynomials

$$f = t^n + a_{n-1}t^{n-1} + \cdots + a_0 \quad \text{and} \quad g = t^n + b_{n-1}t^{n-1} + \cdots + b_0.$$

It follows that if

$$|f-g| = \|\varphi(\xi)-\eta\| \le e^{-C_{30}}|J(\xi)|^n = e^{-C_{30}}|f'(\alpha)|^n,$$

then there is a $\beta \in \overline{K}$ such that $g(\beta)=0$ and

$$|\alpha-\beta||f'(\alpha)|^{n-1} \le \|\xi-(b_0,\dots,b_{n-1},\beta)\||f'(\alpha)|^{n-1} \le e^{C_{31}}|f-g|.$$

It remains to show that $\beta \in K$, where we may need to increase the value of C_{30}. To this end, we may assume that $f'(\alpha)\neq 0$, since otherwise (17) tells us that $f=g$, so we may take $\beta=\alpha$.

Using the assumption that $f'(\alpha)\neq 0$, we can estimate

$$\begin{aligned}
\big|g'(\beta)\big| &\ge \big|f'(\alpha)\big| - \big|f'(\alpha)-g'(\beta)\big| \\
&\ge \big|f'(\alpha)\big| - C_{32}|\alpha-\beta| \quad \text{for some } C_{32}>0 \text{ depending} \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{on } n, D, \text{ and } |\cdot|. \\
&\ge \big|f'(\alpha)\big| - C_{32}e^{C_{31}}\frac{|f-g|}{\big|f'(\alpha)\big|^{n-1}} \\
&\ge \big|f'(\alpha)\big| - C_{32}e^{C_{31}}e^{-C_{30}}\big|f'(\alpha)\big| \\
&= C_{33}\big|f'(\alpha)\big| \quad \text{where } C_{33}> \text{ if we take an} \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{appropriately large value for } C_{30}.
\end{aligned} \tag{18}$$

In particular, we note that $g'(\beta)\neq 0$.

Suppose $\beta \notin K$. Let $\beta' \in \overline{K}$ be a $\mathrm{Gal}(\overline{K}/K)$-conjugate of β with $\beta'\neq\beta$. Then

$$|\beta'-\beta| \ge C_{34}|g'(\beta)| \ge C_{35}|f'(\alpha)|$$

where the first inequality is elementary and the second is (18). Note that we also have

$$|\alpha-\beta| \le e^{C_{31}-C_{30}}|f'(\alpha)|.$$

Thus we get

$$|\beta'-\beta| \ge C_{35}e^{C_{30}-C_{31}}|\alpha-\beta|.$$

When K is archimedean, we have $K=\mathbb{C}$ or $\mathbb{R}$ with absolute value $|\cdot|^\epsilon$ for some $\epsilon\in(0,1]$ where $|\cdot|$ is the usual absolute value. If $K=\mathbb{C}$, β must have been contained in K. If $K=\mathbb{R}$, the above inequality implies

$$|2\operatorname{Im}\beta|^\epsilon \ge C_{35}e^{C_{30}-C_{31}}|\operatorname{Im}\beta|^\epsilon.$$

This does not hold if we take C_{30} large enough.

When K is non-archimedean, we take C_{30} large enough so that

$$|\beta'-\beta| > |\alpha-\beta|.$$

It follows from Krasner's Lemma [4, Ch. II, Sec. 2, Prop. 3] that $\beta \in K(\alpha) = K$. □

7. The inverse function theorem: Version II

In Theorem 5.12 we proved an inverse function theorem that is uniform over a possibly infinite collection of absolute values. In this section we work over a single complete field and use a higher dimensional version of Newton's iterative method to prove the following stronger statement.

Theorem 7.1 *(Inverse function theorem II). Let $\big(K, |\cdot|\big)$ be a complete field. Let W, V be smooth quasi-projective varieties over K, and let $\varphi\colon W \longrightarrow V$ be a generically finite generically étale morphism. Let $E \subset W$ be the closed subscheme defined by the 0-th fitting ideal sheaf of $\Omega_{W/V}$. Fix arithmetic distance functions δ_W, δ_V, a local height function λ_E, and a boundary function $\lambda_{\partial V}$. Let $B \subset W(K)$ be a bounded subset.*

Then there are constants $C_{36}, C_{37}, C_{38}, C_{39} > 0$ and a bounded subset $\widetilde{B} \subset W(K)$ containing B such that for all $P \in B$ and $q \in V(K)$ satisfying

$$P \notin E \quad \text{and} \quad \delta_V(\varphi(P), q) \geq 2\lambda_E(P) + C_{36}\lambda_{\partial V}(q) + C_{37},$$

there is a unique $Q \in \widetilde{B}$ satisfying

$$\varphi(Q) = q \quad \text{and} \quad \delta_W(P, Q) \geq \delta_V(\varphi(P), q) - \lambda_E(P) - C_{38}\lambda_{\partial V}(q) - C_{39}.$$

Remark 7.2. If W is projective, then we may take $B = \widetilde{B} = W(K)$, since projective varieties are covered by finitely many affine bounded subsets.

Remark 7.3. If φ is a finite morphism, then E is equal to the ramification divisor of φ.

Corollary 7.4 *(A variant of continuity of roots). Let $\big(K, |\cdot|\big)$ be a complete field. Let $D \in \mathbb{R}_{>0}$ and $n \in \mathbb{Z}_{>0}$. Then there are positive constants $C_{40}, C_{41} > 0$ such that the following holds. Suppose that*:

- *$f, g \in K[t]$ are monic polynomials of degree n;*
- *$|f| \leq D$ and $|g| \leq D$;*
- *There is an $\alpha \in K$ such that*

$$f(\alpha) = 0 \quad \text{and} \quad |f - g| \leq C_{40}|f'(\alpha)|^2.$$

Then there is $\beta \in K$ such that

$$g(\beta) = 0 \quad \text{and} \quad |\alpha - \beta||f'(\alpha)| \leq C_{41}|f - g|.$$

Proof. The proof is the same as the proof of Proposition 6.1, except that we use Theorem 7.1, instead of Theorem 5.12. Note that if $\varphi\colon W \longrightarrow V$ is a finite surjective morphism of smooth varieties, then the closed subscheme defined by the 0-th Fitting ideal of $\Omega_{W/V}$ is equal to the ramification divisor of φ. □

We first prove the uniqueness of the Q in Theorem 7.1. Then, in order to prove existence, we consider the archimedean and non-archimedean cases separately. We start with a lemma that says if x is in a bounded subset, then all y that are sufficiently close to x also lie in a bounded subset.

Lemma 7.5. *Let $\big(K, |\cdot|\big)$ be a complete field. Let X be a quasi-projective variety over K. Fix an arithmetic distance function δ_X and a boundary function $\lambda_{\partial(X\times X)}$. Let $B \subset X(K)$ be a bounded subset.*

(a) *There are constants $C_{42}, C_{43} > 0$ so that*

$$\bigcup_{x\in B} \big\{y \in X(K) : \delta_X(x,y) \geq C_{42}\lambda_{\partial(X\times X)}(x,y) + C_{43}\big\}$$

is confined in a bounded subset $B' \subset X(K)$.

(b) *If B is an affine bounded subset of some open affine subset $U \subset X$, then B' may be chosen to be an affine bounded subset of the same open set U.*

(c) *If further the absolute value is non-archimedean and B is a standard bounded subset of U, then we may take $B' = B$.*

Proof. Without loss of generality, we may assume that there is an open affine subset $U \subset X$ such that $B \subset U(K)$ and that B is affine bounded in U. Let $x \in B$ and $y \in X(K)$ be two points that satisfy

$$\delta_X(x,y) \geq C_{42}\lambda_{\partial(X\times X)}(x,y) + C_{43}.$$

Let $Z = X \smallsetminus U$. We equip Z with its reduced structure and fix a local height λ_Z. The triangle inequality [8, Prop. 3.1(c)] gives

$$\min\{\lambda_Z(y), \delta_X(x,y)\} \leq \lambda_Z(x) + C_{44}\lambda_{\partial(X\times X)}(x,y) + C_{45},$$

where C_{44} and C_{45} are independent of x and y.

Suppose that

$$\delta_X(x,y) \leq \lambda_Z(x) + C_{44}\lambda_{\partial(X\times X)}(x,y) + C_{45}. \tag{19}$$

Then we have

$$C_{42}\lambda_{\partial(X\times X)}(x,y) + C_{43} \leq \lambda_Z(x) + C_{44}\lambda_{\partial(X\times X)}(x,y) + C_{45}.$$

Since $x \in B$, we know that

$$\lambda_Z(x) \le \sup_{\xi \in B} \lambda_Z(\xi) =: M < \infty,$$

and hence

$$(C_{42} - C_{44})\lambda_{\partial(X\times X)}(x,y) + C_{43} - C_{45} \le M. \tag{20}$$

Since $\lambda_{\partial(X\times X)}$ is bounded below, we see that (20) is false for sufficiently large C_{42} and C_{43} and hence (19) is also false.

We may thus assume that

$$\lambda_Z(y) \le \lambda_Z(x) + C_{44}\lambda_{\partial(X\times X)}(x,y) + C_{45}.$$

In particular, $\lambda_Z(y) < \infty$ and $y \in U$.

Let $f_1, \dots, f_r$ be K-algebra generators for the ring $\mathcal{O}(U)$. Then there are constants C_{46}, C_{47} such that

$$\begin{aligned}\delta_X(\xi,\eta) \le{}& \log \frac{1}{\max_{1\le i\le r}\{|f_i(\xi) - f_i(\eta)|\}} \\ &+ C_{46}(\lambda_Z(\xi) + \lambda_Z(\eta) + \lambda_{\partial(X\times X)}(\xi,\eta)) + C_{47}\end{aligned}$$

for all $\xi, \eta \in U(K)$. Plugging in $\xi = x$ and $\eta = y$, we get

$$\begin{aligned}C_{42}\lambda_{\partial(X\times X)}(x,y) + C_{43} &\le \delta_X(x,y) \\ &\le \log \min_{1\le i\le r} \big|f_i(x) - f_i(y)\big|^{-1} \\ &\quad + C_{46}\big(\lambda_Z(x) + \lambda_Z(y) + \lambda_{\partial(X\times X)}(x,y)\big) + C_{47} \\ &\le \log \min_{1\le i\le r} \big|f_i(x) - f_i(y)\big|^{-1} \\ &\quad + C_{46}\big(2\lambda_Z(x) + (C_{44}+1)\lambda_{\partial(X\times X)}(x,y) + C_{45}\big) + C_{47} \\ &\le \log \min_{1\le i\le r} \big|f_i(x) - f_i(y)\big|^{-1} \\ &\quad + C_{48}\lambda_Z(x) + C_{49}\lambda_{\partial(X\times X)}(x,y) + C_{50}.\end{aligned}$$

Thus

$$\begin{aligned}(C_{42} - C_{49})\lambda_{\partial(X\times X)}(x,y) + C_{43} - C_{50} - C_{48}M \\ \le \log \min_{1\le i\le r} \big|f_i(x) - f_i(y)\big|^{-1}.\end{aligned}$$

The boundary function $\lambda_{\partial(X\times X)}$ is bounded below, and the values $\big|f_1(x)\big|,\dots,\big|f_r(x)\big|$ are bounded for $x\in B$, so we see that the values $\big|f_1(y)\big|,\dots,\big|f_r(y)\big|$ are bounded. Hence y is contained in an affine bounded subset of $U(K)$.

Finally let $\big(K,|\cdot|\big)$ be non-archimedean, and let B be a standard bounded subset of U. We choose K-algebra generators $f_1,\dots,f_r$ and positive constants $b_1,\dots,b_r$ so that

$$B=\big\{x\in U(K):\big|f_1(x)\big|\le b_1,\dots,\big|f_r(x)\big|\le b_r\big\}.$$

Choosing C_{42} and C_{43} sufficiently large ensures that $\big|f_i(x)-f_i(y)\big|\le b_i$ for all i. This implies $\big|f_i(y)\big|\le b_i$, since $\big(K,|\cdot|\big)$ is non-archimedean, and thus we see that $y\in B$. □

Proof of uniqueness in Theorem 7.1. As usual, we may assume that $\lambda_{\partial V}\ge 0$. Suppose now that both of the points $Q,Q'\in\widetilde{B}$ satisfy the conclusion of Theorem 7.1. Our goal is to show that $Q=Q'$.

Since $\varphi(P)$ varies over the bounded subset $\varphi(B)$ and λ_E is bounded below, Lemma 7.5 tells us that there is a bounded subset $B'\subset V(K)$ such that $q\in B'$ if we take $C_{36},\dots,C_{39}$ to be sufficiently large. Set

$$M=\sup_{\eta\in B'}\lambda_{\partial V}(\eta)<\infty.$$

Note that if we increase the values of $C_{36},\dots,C_{39}$, we can use the same bounded set B' and the same value for M. We estimate

$$\begin{aligned}\min\{\delta_W(P,Q),\delta_W(P,Q')\}&\ge\delta_V(\varphi(P),q)-\lambda_E(P)-C_{38}\lambda_{\partial V}(q)-C_{39}\\&\ge\lambda_E(P)+C_{37}-C_{39}-C_{38}M.\end{aligned}$$

On the other hand, by the triangle inequality [8, Prop. 3.1(c)] and separation (Proposition 5.7), we have

$$\min\{\delta_W(P,Q),\delta_W(P,Q')\}\le\delta_W(Q,Q')+C_{51}\le\lambda_E(Q)+C_{52}$$

for constants C_{51},C_{52} independent of P,q,Q,Q'.

By the triangle inequality again, we have

$$\min\{\lambda_E(Q),\delta_W(P,Q)\}\le\lambda_E(P)+C_{53}$$

for some constant C_{53}. If $\delta_W(P,Q)\le\lambda_E(P)+C_{53}$, then

$$\lambda_E(P)+C_{37}-C_{39}-C_{38}M\le\lambda_E(P)+C_{53},$$

which gives a contradiction if C_{37} is sufficiently large.

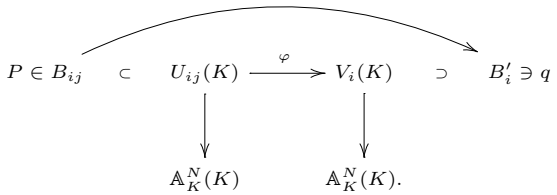


Fig. 6. Maps from affine bounded sets to affine space.

On the other hand, if $\lambda_E(Q) \le \lambda_E(P) + C_{53}$, then

$$\lambda_E(P) + C_{37} - C_{39} - C_{38}M \le \lambda_E(Q) + C_{52} \le \lambda_E(P) + C_{52} + C_{53},$$

which again gives a contradiction if C_{37} is sufficiently large. □

Proof of existence of Theorem 7.1. The bulk of the proof of the existence of Q is based on local calculations that are somewhat different in the archimedean and non-archimedean cases. So in the remainder of this section we give the part of the proof that is common to both cases, and refer the reader to Sections 8 and 9 for the remainder of the proof.

As usual, we may assume that take $\lambda_{\partial V} \ge 0$. We start with an open affine cover $\{V_i\}_{i=1}^r$ of V such that each V_i admits an étale morphism to an affine space $\mathbb{A}_K^N$. For each i we fix an open affine cover $\{U_{ij}\}_{j=1}^{s_i}$ of $\varphi^{-1}(V_i)$ such that each U_{ij} admits an étale morphism to $\mathbb{A}^N$.

We choose generators for the rings $\mathcal{O}_V(U_{ij})$ so that the associated standard bounded subsets $B_{ij} \subset U_{ij}(K)$ satisfy $B \subset \bigcup_{i,j} B_{ij}$. By Lemma 7.5, there are constants C_{54}, C_{55}, and standard bounded subsets $B_i' \subset V_i(K)$ such that:

- $\varphi(B_{ij}) \subset B_i'$
- if $q \in V(K)$ satisfies

$$\delta_V(\varphi(P), q) \ge 2\lambda_E(P) + C_{54}\lambda_{\partial V}(q) + C_{55}$$

for some $P \in B_{ij}$, then $q \in B_i'$.

Thus we can reduce the existence of Q to the situation summarized in Fig. 6, which essentially reduces the problem to affine space.

Case 1: $(K, |\cdot|)$ is non-archimedean. In this case, the proof of existence will be done in Section 8; see Proposition 8.1.

Case 2: $(K, |\cdot|)$ is archimedean. We may take $K = \mathbb{R}$ or $\mathbb{C}$ with the usual absolute value. Then for any $x \in B_{ij}$, there are compact neighborhoods

$$x \in T_1 \subset T_2 \subset U_{ij}(K) \quad \text{and} \quad \varphi(x) \in T' \subset U_i(K)$$

such that:

- $T_1 \subset T_2^\circ$, where T_2° denotes the interior of T_2;
- $\varphi(T_1) \subset (T')^\circ$;
- a small open neighborhood of T_2 is homeomorphic to its image by $U_{ij} \longrightarrow \mathbb{A}_K^N$;
- a small open neighborhood of T' is homeomorphic to its image by $V_i \longrightarrow \mathbb{A}_K^N$.

Since B_{ij} is compact, we can cover B_{ij} by finitely many sets that look like T_1. Further, if C_{54}, C_{55} are sufficiently large, then for all $q \in B_{ij}$ such that there exists some $P \in T_1$ with

$$\delta_V(\varphi(P), q) \geq 2\lambda_E(P) + C_{54}\lambda_{\partial V}(q) + C_{55},$$

we necessarily have $q \in T'$.

In this case, Proposition 9.1, which we prove in Section 9, completes the proof of existence, once we observe that (1) the function $e^{-\delta_W}$ is comparable to the usual metric on T_2; (2) the function $e^{-\delta_V}$ is comparable to the usual metric on T'; (3) if J is a local equation for E, then $e^{-\lambda_E}$ is comparable to $|J|$; (4) the function $\lambda_{\partial V}$ is bounded on T'. □

8. Newton's method (non-archimedean case)

When we work over a non-archimedean complete field such as $\mathbb{C}_p$, we lose the local compactness, and thus we cannot localize problems as is done over the complex numbers. Nevertheless, similar arguments work if we use bounded subsets instead of small neighborhoods. The goal of this section is to prove the following proposition.

Proposition 8.1. *Let $\big(K, |\cdot|\big)$ be a non-archimedean complete field. Let V/K and W/K be smooth affine varieties, let $\varphi\colon W \longrightarrow V$ be a generically finite generically étale morphism defined over K and suppose that there are étale morphisms π and ν to affine space as in the following diagram*:

$$\begin{array}{ccc} W & \xrightarrow{\varphi} & V \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\nu} \\ \mathbb{A}_K^N & & \mathbb{A}_K^N. \end{array}$$

Let $E \subset W$ be the closed subscheme defined by the 0th Fitting ideal sheaf $\mathrm{Fit}_0(\Omega_{W/V})$ of the relative sheaf of differentials $\Omega_{W/V}$. Fix arithmetic distance functions δ_W and δ_V and a local height function λ_E.

Let $B \subset W(K)$ and $B' \subset V(K)$ be standard bounded subsets such that $\varphi(B) \subset B'$. Then there are constants $C_{56}, C_{57} > 0$ such that the following holds. If $P \in B$ and $q \in B'$ are points satisfying

$$P \notin E \quad \text{and} \quad \delta_V(\varphi(P), q) \geq 2\lambda_E(P) + C_{56},$$

then there exists a point $Q \in B$ satisfying

$$\varphi(Q) = q \quad \text{and} \quad \delta_W(P, Q) \geq \delta_V(\varphi(P), q) - \lambda_E(P) - C_{57}.$$

We start with some preliminary results. The following proposition, which we will apply to étale morphism from an affine open subset to an affine space, enables us to go from affine coordinates back to our original variety. It is exactly the inverse function theorem for étale morphisms.

Proposition 8.2. *Let $\big(K, |\cdot|\big)$ be a non-archimedean complete field. Let X/K and Y/K be affine varieties, and let $\pi\colon X \longrightarrow Y$ be an étale morphism defined over K. Let $B \subset X(K)$ be a standard bounded set, and let $B' \subset Y(K)$ be a bounded set. Fix arithmetic distance functions δ_X and δ_Y on X and Y. Then there are constants $C_{58}, C_{59} > 0$ such that if $x \in B$ and $y \in B'$ satisfying*

$$\delta_Y(\pi(x), y) \geq C_{58},$$

then there exists a unique $z \in B$ satisfying

$$\pi(z) = y \quad \text{and} \quad \delta_X(x, z) \geq \delta_Y(\pi(x), y) - C_{59}.$$

Proof. (Proof of Existence): Step 1. We first localize the problem so that we can use Chevalley's structure theorem of étale morphisms. Let

$$X = \operatorname{Spec} S \quad \text{and} \quad Y = \operatorname{Spec} R.$$

Let elements $f_1, \dots, f_r \in S$, $g_1, \dots, g_r, g_{r+1}, \dots, g_s \in R$ be any collections of elements that have the following properties:

- $(f_1, \dots, f_r) = S$.
- $(g_1, \dots, g_s) = R$.
- $g_1, \dots, g_s$ generate R as a K-algebra.
- For each $i = 1, \dots, r$, we have $\pi(\operatorname{Spec} S_{f_i}) \subset \operatorname{Spec} R_{g_i}$, i.e.,

$$\begin{array}{ccc} X = \operatorname{Spec} S & \xrightarrow{\pi} & \operatorname{Spec} R = Y \\ \uparrow & & \uparrow \\ \operatorname{Spec} S_{f_i} & \longrightarrow & \operatorname{Spec} R_{g_i} \end{array}$$

For each $i = 1, \dots, r$ we define a subset $B_i \subset B$ by

$$B_i = \left\{ x \in B : \ \left|\frac{f_1}{f_i}(x)\right| \le 1, \dots, \left|\frac{f_r}{f_i}(x)\right| \le 1 \right\}.$$

Then B_i is a standard bounded subset of $\operatorname{Spec} S_{f_i}$ and $\bigcup_{i=1}^{r} B_i = B$. Note that for each $i = 1, \dots, r$, the set $\pi(B_i)$ is a bounded subset of $\operatorname{Spec} R_{g_i}$.

We claim that there is a number $C_{60} > 0$ such that

$$B_i' := \{ y \in B' \mid \delta_Y(\pi(x), y) \ge C_{60} \text{ for some } x \in B_i \}$$

is a bounded subset of $\operatorname{Spec} R_{g_i}$ for $i = 1, \dots, r$. Indeed, we may assume that δ_Y is

$$\delta_Y(y, y') = \log \frac{1}{\max_{1 \le j \le s} \{ |g_j(y) - g_j(y')| \}}$$

on $(\pi(B) \cup B') \times (\pi(B) \cup B')$. Since $1/g_i$ is regular on $\operatorname{Spec} R_{g_i}$ and $\pi(B_i)$ is a bounded subset of $\operatorname{Spec} R_{g_i}$, there is $C_{61} > 0$ such that

$$|g_i(\pi(x))| \ge C_{61} \quad \text{for all } x \in B_i.$$

Thus, for any $y \in B'$, if $C_{60} > 0$ is large enough and $\delta_Y(\pi(x), y) \ge C_{60}$ for some $x \in B_i$, then we get $|g_i(y)| \ge C_{61}$. This proves the claim.

Hence to prove existence, we may replace $\pi \colon \operatorname{Spec} S \longrightarrow \operatorname{Spec} R$, B, and B' with $\pi \colon \operatorname{Spec} S_{f_i} \longrightarrow \operatorname{Spec} R_{g_i}$, B_i, and B_i'.

Step 2. By Step 1 and Chevalley's structure theorem of étale morphisms (cf. [1, Theorem 2.3.5]), we may assume that $Y = \operatorname{Spec} R$ and $X = \operatorname{Spec} S$, where

$$S = \Big(R[t]/\big(f(t)\big) \Big)_{g(t)}.$$

Here $f(t), g(t) \in R[t]$ are polynomials with $f(t)$ a monic, and

$$f'(t) \in \Big(R[t]/\big(f(t)\big) \Big)_{g(t)}^{\times}.$$

The situation is summarized in the following diagram:

$$\begin{array}{ccc} \operatorname{Spec}\Big(R[t]/\big(f(t)\big) \Big)_{g(t)} & & B \ni x \\ \Big\downarrow \pi & & \Big\downarrow \pi \\ \operatorname{Spec} R & & B' \ni y \xleftrightarrow[\text{close to one another}]{\text{points that are}} \pi(x) \end{array}$$

We write

$$f(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_0,$$
$$g(t) = b_m t^m + b_{m-1}t^{m-1} + \cdots + b_0,$$

where $a_i, b_j \in R$. We write $f\big((a_i);t\big)$ and $g\big((b_j);t\big)$ if we need to specify the coefficients.

We choose $c_0, \ldots, c_l \in R$ so that

$$R = K[a_0, \ldots, a_{n-1}, b_0, \ldots, b_m, c_0, \ldots, c_l].$$

Then we may take

$$\begin{aligned}&\delta_Y(y, y')\\ &= \log \min_{i,j,k}\Big\{\big|a_i(y) - a_i(y')\big|^{-1}, \big|b_j(y) - b_j(y')\big|^{-1}, \big|c_k(y) - c_k(y')\big|^{-1}\Big\}\end{aligned}$$

on $\big(\pi(B) \cup B'\big) \times \big(\pi(B) \cup B'\big)$.

Let $x \in B$ and $y \in B'$. We are going to apply Proposition 6.1 to the polynomials $f\big((a_i(x));t\big)$ and $f\big((a_i(y));t\big)$, and to the quantity $t(x)$, which is a root of $f\big((a_i(x));t\big)$. If $\delta_Y\big(\pi(x), y\big) \geq C_{62}$, then for all $i = 0, \ldots, n-1$ we get

$$\big|a_i(\pi(x)) - a_i(y)\big| \leq e^{-C_{62}}. \tag{21}$$

Since $\pi(x)$ and y move in bounded subsets, there is a number $D > 0$, which is independent of x, y, such that

$$|a_i(\pi(x))| \leq D \quad \text{and} \quad \big|a_i(y)\big| \leq D \quad \text{for all } i = 0, \ldots, n-1.$$

Since $f'(t)$ and $1/f'(t)$ are regular functions on X, there is a $D' > 0$, which is independent of x, such that

$$\frac{1}{D'} \leq \big|f'\big((a_i(x));t(x)\big)\big| \leq D'.$$

Then by Proposition 6.1, if C_{62} is large enough, there is a number $C_{63} > 0$ such that the following holds. For all $x \in B$ and $y \in B'$ satisfying

$$\delta_Y\big(\pi(x), y\big) \geq C_{62},$$

there is $\beta \in K$ such that

$$f\big((a_i(y));\beta\big) = 0 \text{ and } \big|t(x) - \beta\big| \leq C_{63} \max_{\leq i \leq n-1}\big\{\big|a_i(\pi(x)) - a_i(y)\big|\big\}. \tag{22}$$

Note that (y, β) defines a K-valued point z of $\operatorname{Spec} R[t]/\big(f(t)\big)$. By (22), the point z moves in a bounded subset of $\operatorname{Spec} R[t]/\big(f(t)\big)$ as x and y move.

By the choice of δ_Y and (21) and (22), there is a constant $C_{64} > 0$ independent of x and y such that

$$\left|g\big((b_j(\pi(x)));t(x)\big) - g((b_j\big(\pi(z)\big));t(z))\right| \le C_{64}e^{-\delta_Y(\pi(x),y)} \le C_{64}e^{-C_{62}}. \tag{23}$$

Since $\left|g\big((b_j(\pi(x)));t(x)\big)\right|^{-1}$ is bounded above when x runs over B, we see that

$$\left|g\big((b_j(\pi(z)));t(z)\big)\right|^{-1}$$

is also bounded above as x and y move. Hence z is contained in a bounded subset B_1 of $X = \mathrm{Spec}(R[t]/(f(t)))_{g(t)}$. On $(B \cup B_1) \times (B \cup B_1)$, we may take

$$\delta_X(\xi,\eta) = \log\min\Big\{e^{\delta_Y(\pi(\xi),\pi(\eta))}, \big|t(\xi)-t(\eta)\big|^{-1}, \big|(1/g)(\xi)-(1/g)(\eta)\big|^{-1}\Big\}.$$

Then we get

$$\begin{aligned}\delta_X(x,z) &\ge \log\min\{e^{\delta_Y(\pi(x),y)}, C_{63}^{-1}e^{\delta_Y(\pi(x),y)}, C_{65}e^{\delta_Y(\pi(x),y)}\}\\ &\ge \delta_Y\big(\pi(x),y\big) - C_{66}\\ &\ge C_{62} - C_{66}\end{aligned}$$

where $C_{66} = \log\max\{1, C_{63}^{-1}, C_{65}\}$. Thus if C_{62} is large enough, by Lemma 7.5, we get $z \in B$. (Note that we can take C_{63}, C_{65} independent of C_{62}.) This z is the point that we want.

(Proof of Uniqueness): We apply Proposition 5.7 to the morphism $\pi\colon X \longrightarrow Y$. Since $\Omega_{X/Y} = 0$, there is a number $C_{67} > 0$ such that

$$\delta_X(z,z') \le C_{67}$$

for all $z, z' \in B$ such that $\pi(z) = \pi(z')$ and $z \ne z'$. Suppose that there are constants $C_{58}, C_{59} > 0$ and points $x, z, z' \in B$ and $y \in B'$ satisfying

- $\delta_Y(\pi(x), y) \ge C_{58}$;
- $z \ne z'$ and $\pi(z) = \pi(z') = y$;
- $\delta_X(x,z) + C_{59} \ge \delta_Y(\pi(x),y)$ and $\delta_X(x,z') + C_{59} \ge \delta_Y(\pi(x),y)$.

Then find that

$$\begin{aligned}C_{58} - C_{59} \le \delta_Y(\pi(x),y) - C_{59} &\le \min\big\{\delta_X(x,z), \delta_X(x,z')\big\}\\ &\le \delta_X(z,z') + C_{68}\\ &\le C_{67} + C_{68}\end{aligned}$$

where C_{68} comes from the triangle inequality on $B \times B \times B$. Thus if $C_{58} - C_{59}$ is large enough, more precisely if $C_{58} > C_{59} + C_{67} + C_{68}$, then we get a contradiction, which proves that z is unique. □

We next compare the distance between two points to the distance between their images under an étale morphism.

Corollary 8.3. *Let $\big(K, |\cdot|\big)$ be a non-archimedean complete field. Let X/K and Y/K be affine varieties, and let $\pi\colon X \longrightarrow Y$ be an étale morphism defined over K. Fix arithmetic distance functions δ_X and δ_Y on X and Y, respectively. Let $B \subset X(K)$ be a bounded subset. Then there are constants $C_{69}, C_{70} > 0$ such that for all points $x, y \in B$ satisfying*

$$\delta_X(x,y) \geq C_{69},$$

we have

$$\delta_X(x,y) + C_{70} \geq \delta_Y\big(\pi(x), \pi(y)\big).$$

Proof. We may assume B is a standard bounded subset. We apply Proposition 8.2 to the map $\pi\colon X \to Y$ and the bounded sets B and $B' = \pi(B)$, and we let C_{69}, C_{70} be the constants appearing in the conclusion of that proposition.

In general, there is a constant C_{71} such that

$$\delta_X \leq \delta_Y \circ (\pi \times \pi) + C_{71}.$$

Hence if $\delta_X(x,y) \geq C_{72}$ with $x, y \in B$, then we have

$$\delta_Y(\pi(x), \pi(y)) \geq C_{72} - C_{71}.$$

Thus if we take C_{72} so that $C_{72} - C_{71} \geq C_{69}$, then there is a point $z \in B$ such that

$$\pi(z) = \pi(y) \quad \text{and} \quad \delta_X(x,z) \geq \delta_Y(\pi(x), \pi(y)) - C_{70}.$$

Hence

$$C_{72} - C_{71} - C_{70} \leq \min\big\{\delta_X(x,y), \delta_X(x,z)\big\} \leq \delta_X(y,z) + C_{73},$$

where C_{73} comes from the triangle inequality on $B \times B \times B$.

We apply Proposition 5.7 to the morphism $\pi\colon X \to Y$. Since $\Omega_{X/Y} = 0$, there is a number $C_{74} > 0$ such that

$$\delta_X(w,w') \leq C_{74} \quad \text{for all } w, w' \in B \text{ with } \pi(w) = \pi(w') \text{ and } w \neq w'.$$

In our situation, if $y \neq z$, then $\delta_X(y,z) \leq C_{74}$, so we find that

$$C_{72} - C_{71} - C_{70} \leq C_{74} + C_{73}.$$

Thus if we take C_{72} sufficiently large, i.e., so that $C_{72} - C_{71} - C_{70} > C_{74} + C_{73}$, then $y = z$ and

$$\delta_X(x,z) \geq \delta_Y\big(\pi(x),\pi(y)\big) - C_{70},$$

which concludes the proof of Corollary 8.3. □

Our next result is an algebraic version of Taylor's theorem up to second order terms that is uniform on bounded sets.

Lemma 8.4. *Let $\big(K,|\cdot|\big)$ be a complete field, let X/K be an affine variety, fix an arithmetic distance function δ_X on X, and let $B \subset X(K)$ be a bounded subset of X.*

(1) *Let $f \in \mathcal{O}(X)$ be a regular function. Then there is a number $C > 0$ such that for all $a, b \in B$, we have*

$$\big|f(a) - f(b)\big| \leq Ce^{-\delta_X(a,b)}. \tag{24}$$

(2) *Let*

$$\pi\colon X \longrightarrow \mathbb{A}^N_K = \operatorname{Spec} K[x_1,\dots,x_N]$$

be an étale morphism. For any regular function $f \in \mathcal{O}(X)$, define $\partial f/\partial x_i$ by

$$\Omega_X \simeq \pi^*\Omega_{\mathbb{A}^N} \simeq \bigoplus_{i=1}^N \mathcal{O}_X d(\pi^* x_i), \qquad df \longleftrightarrow \sum_{i=1}^N \frac{\partial f}{\partial x_i} d(\pi^* x_i).$$

Then for any $f \in \mathcal{O}(X)$, there is a number $C > 0$ such that for all $a, b \in B$, we have

$$\left| f(a) - f(b) - \sum_{i=1}^N \frac{\partial f}{\partial x_i}(b)(a_i - b_i) \right| \leq Ce^{-2\delta_X(a,b)}, \tag{25}$$

where $\pi(a) = (a_1,\dots,a_N)$ and $\pi(b) = (b_1,\dots,b_N)$. N.B. The crucial quantity in (25) is the 2 appearing in the exponent of $e^{-2\delta_X(a,b)}$.

Proof. (1) Let $f = f_1, f_2, \dots, f_r$ be K-algebra generators of $\mathcal{O}(X)$. Then we may take

$$\delta_X(a,b) = \log \min_{1\leq i\leq r} \big|f_i(a) - f_i(b)\big|^{-1}$$

for $(a,b) \in B \times B$. Then (24) is a tautology.

(2)

Let I be the ideal of diagonal in $X \times X$. By definition we have $\Omega_X = I/I^2$, and the differential of a regular function $g \in \mathcal{O}(X)$ is given by $dg = g\otimes 1 - 1\otimes g$. Thus the image of

$$F = (f \otimes 1 - 1 \otimes f) - \sum_{i=1}^{N} \left(1 \otimes \frac{\partial f}{\partial x_i}\right) (\pi^* x_i \otimes 1 - 1 \otimes \pi^* x_i) \tag{26}$$

in $\Omega_X = I/I^2$ is

$$df - \sum_{i=1}^{N} \frac{\partial f}{\partial x_i} d(\pi^* x_i) = 0.$$

This means that $F \in I^2$. Thus we have

$$\lambda_F + C \geq 2\lambda_I = 2\delta_X \quad \text{on } B \times B$$

for some constant $C > 0$. In other words,

$$\left|F(a,b)\right| \leq e^C e^{-2\delta_X(a,b)} \quad \text{for } (a,b) \in B \times B.$$

Since

$$F(a,b) = f(a) - f(b) - \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(b)(a_i - b_i),$$

this gives the desired inequality. □

Proof of Proposition 8.1. We fix coordinates of the affine space $\mathbb{A}^N$. For $\xi = (\xi_1, \ldots, \xi_N) \in K^N$, we write $\|\xi\| = \max_{1 \leq i \leq N} |\xi_i|$, and similarly $\|M\|$ denotes the maximum of the absolute values of the entries of the matrix M.

We choose numbers $b, b' > 0$ so that the standard bounded subsets

$$B_0 = \{\xi \in K^N \mid \|\xi\| \leq b\} \subset \mathbb{A}^N(K),$$
$$B'_0 = \{\xi \in K^N \mid \|\xi\| \leq b'\} \subset \mathbb{A}^N(K),$$

satisfy $\pi(B) \subset B_0$ and $\nu(B') \subset B'_0$. The following diagram summarizes our setting:

$$\begin{array}{ccccccc}
B & & \subset & W(K) & \xrightarrow{\varphi} & V(K) & \supset & B' \\
\pi \downarrow & & & \pi \downarrow & & \downarrow \nu & & \downarrow \nu \\
B_0 & & \subset & \mathbb{A}^N_K(K) & & \mathbb{A}^N_K(K) & \supset & B'_0.
\end{array}$$

Since π and ν are étale morphisms, we have

$$\begin{array}{ccc} \varphi^*\Omega_V & \xrightarrow{D\varphi} & \Omega_W \\ \uparrow\wr & & \uparrow\wr \\ \varphi^*\nu^*\Omega_{\mathbb{A}^N_K} & & \pi^*\Omega_{\mathbb{A}^N_K} \\ \uparrow\wr & & \uparrow\wr \\ \mathcal{O}_W^N & & \mathcal{O}_W^N \end{array}$$

where the vertical arrows are isomorphisms. Thus $D\varphi$ is represented by an $N\times N$-matrix with entries in $\mathcal{O}(W)$. We identify $D\varphi$ with this matrix, and we let

$$J = \det D\varphi.$$

Since φ is generically étale, we know that J is a non-zero regular function on W. Note that J generates $\mathrm{Fit}_0(\Omega_{W/V})$, and hence we may take

$$\lambda_E = \log|J|^{-1}.$$

We take as our arithmetic distance function $\delta_{\mathbb{A}^N}$ on $\mathbb{A}^N$ the function

$$\delta_{\mathbb{A}^N}(\xi,\eta) = \log\|\xi-\eta\|^{-1} \quad \text{for } \xi,\eta\in K^N,$$

or equivalently,

$$e^{-\delta_{\mathbb{A}^N}(\xi,\eta)} = \|\xi-\eta\|.$$

We henceforth use without comment this identification of the usual norm on K^N and the arithmetic distance function. In particular, we have

$$\big\|\pi(x)-\pi(x')\big\| \ll e^{-\delta_W(x,x')} \quad \text{and} \quad \big\|\nu(y)-\nu(y')\big\| \ll e^{-\delta_V(y,y')}.$$

We are going to work in the following set:

$$S = \left\{ (x,y)\in (W\times V)(K) : \begin{array}{l} x\in B,\ y\in B',\ J(x)\neq 0 \\ e^{-\delta_V(\varphi(x),y)} \le C_{75}\big|J(x)\big|^2 \end{array} \right\},$$

where C_{75} is a positive number that we will take sufficiently small during the proof so as to ensure the various desired properties. In the following, the labeled constants are positive numbers that depend only on

$$W, V, \varphi, B, B', B_0, B_0', \delta_W, \delta_V, \lambda_E, \pi, \nu. \tag{27}$$

In particular, they do not depend on the points chosen on our varieties. Sometimes we omit the phrase such as "there exists a constant $C > 0$ such that. . . ."

Since J is a regular function on W, we have

$$\big|J(x)\big| \le C_{76} \quad \text{for all } x \in B. \tag{28}$$

Claim 1 *(Pullback via π). There are constants $C_{77}, C_{78} > 0$ such that if $x \in B$ and $\zeta \in B_0$ satisfy*

$$\|\pi(x) - \zeta\| \le C_{77},$$

then there exists a unique $z \in B$ satisfying

$$\pi(z) = \zeta \quad \textit{and} \quad e^{-\delta_W(x,z)} \le C_{78}\big\|\pi(x) - \zeta\big\|.$$

We observe that Claim 1 is essentially a restatement of Proposition 8.2.

We now state and prove several useful inequalities.

Claim 2 *(Key Inequalities). There are numbers $C_{79}, C_{80} > 0$ such that for all $(x, y) \in S$, we have*:

$$\big\|(D\varphi)(x)^{-1}\big(\nu(\varphi(x)) - \nu(y)\big)\big\| \le C_{79}\frac{1}{|J(x)|}e^{-\delta_V(\varphi(x),y)}. \tag{29}$$

$$\big\|(D\varphi)(x)^{-1}\big(\nu(\varphi(x)) - \nu(y)\big)\big\| \le C_{79}C_{75}^{1/2}e^{-\delta_V(\varphi(x),y)/2}. \tag{30}$$

$$\big\|(D\varphi)(x)^{-1}\big(\nu(\varphi(x)) - \nu(y)\big)\big\| \le C_{79}C_{75}|J(x)| \le C_{80}C_{75}. \tag{31}$$

Proof of Claim 2. We have

$$\begin{aligned}\big\|(D\varphi)(x)^{-1}\big(\nu(\varphi(x)) - \nu(y)\big)\big\| &\le \big\|(D\varphi)(x)^{-1}\big\| \cdot \big\|\nu(\varphi(x)) - \nu(y)\big\| \\ &\le C_{79}\frac{1}{|J(x)|}e^{-\delta_V(\varphi(x),y)}.\end{aligned}$$

Here the first inequality follows from the triangle inequality. For the second inequality, write

$$(D\varphi)(x)^{-1} = \frac{1}{J(x)}\operatorname{adj}\big((D\varphi)(x)\big)$$

where $\operatorname{adj}((D\varphi)(x))$ is the adjoint matrix. The entries of $\operatorname{adj}((D\varphi)(x))$ are regular functions on W, and thus $\big\|\operatorname{adj}((D\varphi)(x))\big\|$ is bounded on B. This proves (29), and then (30) and (31) follow from (29), the inequality

$$e^{-\delta_V(\varphi(x),y)} \le C_{75}|J(x)|^2,$$

and the fact that $|J|$ is bounded on B. □

The preceding material allows us to reduce the proof of Proposition 8.1 to the following statement: There are constants $C_{75}, C_{81} > 0$ such that for all $(P, q) \in S$, there exists a point $Q \in B$ satisfying

$$\varphi(Q) = q \quad \text{and} \quad e^{-\delta_W(P,Q)} \le C_{81} \frac{1}{\big|J(P)\big|} e^{-\delta_V(\varphi(P),q)}.$$

So we start with an arbitrary $(P, q) \in S$, and we will construct the requisite Q as the limit of a sequence of points $Q_0, Q_1, \cdots \in W(K)$ defined in the following way.

We start by choosing C_{75} sufficiently small so that

$$C_{80}C_{75} \le C_{77} \quad \text{and} \quad C_{80}C_{75} \le b.$$

Algorithm Used to Construct of $Q_0, Q_1, \ldots$.

(1) Set $Q_0 = P$.
(2) Given $Q_0, \ldots, Q_i$ satisfying $(Q_j, q) \in S$ for $j = 0, \ldots, i$, we consider the quantity

$$\eta := \pi(Q_i) - (D\varphi)(Q_i)^{-1}\big(\nu\big(\varphi(Q_i)\big) - \nu(q)\big) \in \mathbb{A}^N(K).$$

Since $(Q_i, q) \in S$, we know from (31) that

$$\Big\|(D\varphi)(Q_i)^{-1}\big(\nu\big(\varphi(Q_i)\big) - \nu(q)\big)\Big\| \le C_{80}C_{75}.$$

It follows that

$$\|\eta\| \le \max\Big\{\big\|\pi(Q_i)\big\|, C_{80}C_{75}\Big\} \le b,$$

so $\eta \in B_0$, and also that

$$\big\|\pi(Q_i) - \eta\big\| \le C_{80}C_{75} \le C_{77}.$$

Hence Claim 1 tells us that there is a unique point $Q_{i+1} \in B$ satisfying

$$\pi(Q_{i+1}) = \eta \quad \text{and} \quad e^{-\delta_W(Q_i,Q_{i+1})} \le C_{78}\|\pi(Q_i) - \eta\|. \tag{32}$$

In order to ensure that we can continue this procedure, and to prove that Q_i converges to a point Q having the desired properties, we verify the following assertions.

Claim 3. *Let α be an arbitrary real number satisfying $0 < \alpha < 1$. If C_{75} is small enough, depending only on (27) and the choice of α, then the following are true*:

(1) $(Q_{i+1}, q) \in S$, *so we can continue the algorithm to create* Q_{i+2}.
(2) *For all* $j = 0, 1, \dots,$ *we have*

$$\big|J(Q_j)\big| = \big|J(P)\big|. \tag{33}$$

(3) *For all* $j = 0, 1, \dots,$ *we have*

$$e^{-\delta_V(\varphi(Q_j),q)} \le \alpha^j e^{-\delta_V(\varphi(P),q)}. \tag{34}$$

We prove Claim 3 by induction. More precisely, we assume that (33) and (34) are true for $j = 0, \dots, i$ and we prove that $(Q_{i+1}, q) \in S$ and that (33) and (34) are true for $j = i+1$. In this induction step, we may replace C_{75} with a smaller value, but the value of C_{75} is always independent of i.

First, we note that

$$\begin{aligned}
&e^{-\delta_V(\varphi(Q_{i+1}),q)} \\
&\le C_{82} \max\{e^{-\delta_V(\varphi(Q_{i+1}),\varphi(Q_i))}, e^{-\delta_V(\varphi(Q_i),q)}\} && \text{triangle inequality} \\
&\le C_{82} \max\{e^{-\delta_V(\varphi(Q_{i+1}),\varphi(Q_i))}, C_{75}C_{76}^2\} && \text{by } (Q_i, q) \in S \text{ and (28)} \\
&\le C_{82} \max\{C_{83} e^{-\delta_W(Q_{i+1},Q_i)}, C_{75}C_{76}^2\} && \delta_W \ll \delta_V \circ (\varphi \times \varphi) + O(1) \\
&\le C_{82} \max\{C_{83}C_{78}\|\pi(Q_i) - \eta\|, C_{75}C_{76}^2\} && \text{by (32)} \\
&\le C_{82} \max\{C_{83}C_{78}C_{80}C_{75}, C_{75}C_{76}^2\} && \text{by the construction of } \eta \\
&= C_{84}C_{75},
\end{aligned}$$

where $C_{84} = C_{82} \max\{C_{83}C_{78}C_{80}, C_{76}^2\}$. Therefore, if C_{75} is sufficiently small, then we may apply Corollary 8.3 to the étale morphism $\nu\colon V \to \mathbb{A}_K^N$ and the points $\varphi(Q_{i+1})$ and q to obtain

$$e^{-\delta_V(\varphi(Q_{i+1}),q)} \le C_{85}\big\|\nu(\varphi\big(Q_{i+1}\big)) - \nu(q)\big\|.$$

Hence

$$\begin{aligned}
&e^{-\delta_V(\varphi(Q_{i+1}),q)} \\
&\le C_{85}\big\|\nu(\varphi(Q_{i+1})) - \nu(q)\big\| \\
&\le C_{85}\big\|\nu(\varphi(Q_{i+1})) - \nu(\varphi(Q_i)) - (D\varphi)(Q_i)(\pi(Q_{i+1}) - \pi(Q_i)) \\
&\qquad\qquad + (D\varphi)(Q_i)(\pi(Q_{i+1}) - \pi(Q_i)) + \nu(\varphi(Q_i)) - \nu(q)\big\| \\
&= C_{85}\big\|\nu(\varphi(Q_{i+1})) - \nu(\varphi(Q_i)) - (D\varphi)(Q_i)(\pi(Q_{i+1}) - \pi(Q_i))\big\| \\
&\le C_{85}C_{86} e^{-2\delta_W(Q_{i+1},Q_i)},
\end{aligned}$$

where the third equality follows from the construction of Q_{i+1} and the last inequality follows from Lemma 8.4(2). By the construction of Q_{i+1}, we get

$$
\begin{aligned}
e^{-\delta_V(\varphi(Q_{i+1}),q)} &\le C_{85}C_{86}e^{-2\delta_W(Q_{i+1},Q_i)} \\
&\le C_{85}C_{86}C_{78}^2\big\|\pi(Q_i)-\eta\big\|^2 \\
&= C_{85}C_{86}C_{78}^2\big\|(D\varphi)(Q_i)^{-1}(\nu(\varphi(Q_i))-\nu(q))\big\|^2 \qquad (35)\\
&\le C_{85}C_{86}C_{78}^2C_{79}^2C_{75}^2|J(Q_i)|^2, \qquad (36)
\end{aligned}
$$

where the last inequality follows from the fact $(Q_i,q)\in S$ and (31).

Now note that

$$
\begin{aligned}
\big|J(Q_{i+1})-J(Q_i)\big| &\le C_{87}e^{-\delta_W(Q_{i+1},Q_i)} && \text{by Lemma 8.4(1)}\\
&\le C_{87}C_{78}\big\|\pi(Q_i)-\eta\big\| && \text{by construction of } Q_{i+1}\\
&= C_{87}C_{78}\big\|(D\varphi)(Q_i)^{-1}(\nu(\varphi(Q_i))-\nu(q))\big\| \\
& && \text{by definition of } \eta\\
&\le C_{87}C_{78}C_{79}C_{75}|J(Q_i)| && \text{by (31)}.
\end{aligned}
$$

Thus if C_{75} is small enough so that $C_{87}C_{78}C_{79}C_{75}<1$, then

$$|J(Q_{i+1})| = |J(Q_i)| = |J(P)|.$$

In particular, this shows that $J(Q_{i+1})\neq 0$ and verifies (33) for $j=i+1$.

Plugging $|J(Q_{i+1})| = |J(Q_i)|$ into (36), we get

$$e^{-\delta_V(\varphi(Q_{i+1}),q)} \le C_{85}C_{86}C_{78}^2C_{79}^2C_{75}^2|J(Q_{i+1})|^2.$$

Hence if C_{75} is sufficiently small to ensure that $C_{85}C_{86}C_{78}^2C_{79}^2C_{75}\le 1$, then we get

$$e^{-\delta_V(\varphi(Q_{i+1}),q)} \le C_{75}\big|J(Q_{i+1})\big|^2.$$

This completes the proof that $(Q_{i+1},q)\in S$.

Finally, we prove (34) for $j=i+1$. By (35) and (30), we have

$$
\begin{aligned}
e^{-\delta_V(\varphi(Q_{i+1}),q)} &\le C_{85}C_{86}C_{78}^2\Big\|(D\varphi)(Q_i)^{-1}\Big(\nu\big(\varphi(Q_i)\big)-\nu(q)\Big)\Big\|^2\\
&\le C_{85}C_{86}C_{78}^2C_{79}^2C_{75}e^{-\delta_V(\varphi(Q_i),q)}.
\end{aligned}
$$

Therefore, it suffices to take C_{75} small enough so that it satisfies

$$C_{85}C_{86}C_{78}^2C_{79}^2C_{75}\le\alpha.$$

This completes the proof of Claim 3.

To finish the proof of Proposition 8.1, we prove that the sequence $\{Q_i\}$ converges and that its limit has the desired properties. In order to talk about the limit, we fix closed immersions of W and V into large affine spaces and identify them with their images.

By construction, (30), and (34), we have

$$\begin{aligned} e^{-\delta_W(Q_i,Q_{i+1})} &\le C_{78}C_{79}C_{75}^{1/2}e^{-\delta_V(\varphi(Q_i),q)/2} \\ &\le C_{78}C_{79}C_{75}^{1/2}e^{-\delta_V(\varphi(P),q)/2}\alpha^{i/2}. \end{aligned}$$

The right-hand side goes to 0 as $i \to \infty$. Hence $\{Q_i\}$ is a Cauchy sequence (since we are in the non-archimedean setting), so it has a limit $Q \in B$. By (34), we have

$$\varphi(Q) = q.$$

Let $\iota\colon W \subset \mathbb{A}^n$ be our chosen embedding, so we may define δ_W by

$$e^{-\delta_W(x,x')} = \bigl\|\iota(x) - \iota(x')\bigr\|.$$

Then, again using the fact that our absolute value $|\cdot|$ is non-archimedean, we have

$$\begin{aligned} e^{-\delta_W(Q_i,P)} &\le \max\{e^{-\delta_W(Q_i,Q_{i-1})}, e^{-\delta_W(Q_{i-1},Q_{i-2})}, \dots, e^{-\delta_W(Q_1,P)}\} \\ &\le \max\left\{C_{79}\frac{e^{-\delta_V(\varphi(Q_{i-1}),q)}}{|J(Q_{i-1})|}, \dots, C_{79}\frac{e^{-\delta_V(\varphi(P),q)}}{|J(P)|}\right\} \\ &\le C_{79}\frac{1}{|J(P)|}e^{-\delta_V(\varphi(P),q)}, \end{aligned}$$

where the second inequality follows from (29), and the last inequality follows from (33) and (34). Taking the limit as $i \to \infty$, we get

$$e^{-\delta_W(Q,P)} \le C_{79}\frac{1}{|J(P)|}e^{-\delta_V(\varphi(P),q)}$$

and we are done. □

9. Newton's method (Archimedean case)

In this section we turn to archimedean case and prove the following result.

Proposition 9.1. *Let $\bigl(K, |\cdot|\bigr)$ be a complete archimedean field, i.e., $K = \mathbb{R}$ or $\mathbb{C}$. Let $N > 0$ be an integer, let $b_2 > b_1 > 0$ be numbers, and define bounded sets*

$$B_1 = \bigl\{x \in K^N : \|x\| \le b_1\bigr\} \quad \text{and} \quad B_2 = \bigl\{x \in K^N : \|x\| \le b_2\bigr\}.$$

Let $U \subset K^N$ be an open neighborhood of B_2, let

$$\varphi = (\varphi_1, \dots, \varphi_N)\colon U \longrightarrow K^N$$

be an analytic map, and let

$$D\varphi = \frac{\partial(\varphi_1, \ldots, \varphi_N)}{\partial(x_1, \ldots, x_N)} \quad \text{and} \quad J = \det D\varphi.$$

Then there are constants $C_{88}, C_{89} > 0$ such that the following holds. For all $P \in B_1$ and $q \in K^N$ satisfying

$$J(P) \neq 0 \quad \text{and} \quad \big\|\varphi(P) - q\big\| \leq C_{88}\big|J(P)\big|^2,$$

there exists a point $Q \in B_2$ satisfying

$$\varphi(Q) = q \quad \text{and} \quad \|P - Q\| \leq C_{89}\frac{\big\|\varphi(P) - q\big\|}{|J(P)|}.$$

Proof. In this proof, the labeled constants, always assumed positive, depend only on N, b_1, b_2, and φ. And since J is a continuous function, there is $C_{90} > 0$ such that

$$\big|J(x)\big| \leq C_{90} \quad \text{for all } x \in B_2. \tag{37}$$

Let $B' \subset K^N$ be a sufficiently large bounded subset so that

$$\bigcup_{P \in B_1} \Big\{ q \in K^N : \big\|\varphi(P) - q\big\| \leq \big|J(P)\big|^2 \Big\} \subset B'.$$

We choose a $C_{91} > 0$ so that for all $x \in B_2$ with $J(x) \neq 0$ and all $q \in B'$, we have

$$\big\|(D\varphi)(x)^{-1}(\varphi(x) - q)\big\| \leq C_{91}\frac{\big\|\varphi(x) - q\big\|}{|J(x)|}. \tag{38}$$

We fix a small positive number $0 < \epsilon < 1$, and we let $\eta > 0$ be a small positive number that we will specify later. For $i = 1, 2, \ldots$, we set

$$c_i = \eta^{(2-\epsilon)^i},$$

and we take η sufficiently small to ensure that

$$b_1 + C_{90}C_{91}\sum_{j \geq 0} c_j \leq b_2.$$

Now let $P \in B_1$ and $q \in B'$ be points satisfying

$$J(P) \neq 0 \quad \text{and} \quad \big\|\varphi(P) - q\big\| \leq \eta\big|J(P)\big|^2.$$

We start with

$$Q_0 = P.$$

Suppose that we have constructed $Q_0, \dots, Q_i \in K^N$ with $J(Q_i) \neq 0$. We then define the next point in the sequence by

$$Q_{i+1} = Q_i - (D\varphi)(Q_i)^{-1}\big(\varphi(Q_i) - q\big).$$

The following claim will be used to show that this sequence converges to a point having the desired properties.

Claim 4. *If η is sufficiently small, then the following are true for all $i = 0, 1, \dots$.*

(a) $\|Q_i\| \le b_1 + C_{90}C_{91}\sum_{j=0}^{i-1} c_j$, *so in particular,* $Q_i \in B_2$.
(b) $J(Q_i) \neq 0$.
(c) $\big\|\varphi(Q_i) - q\big\| \le c_i\big|J(Q_i)\big|^2$.

Proof of Claim 4. The proof is by induction on i. If $i = 0$, everything is true by assumption. Suppose that (a), (b), and (c) are true for i.

We first calculate

$$\begin{aligned}
\|Q_{i+1}\| &\le \|Q_i\| + \big\|(D\varphi)(Q_i)^{-1}(\varphi(Q_i) - q)\big\| \\
&\le \|Q_i\| + C_{91}\frac{\big\|\varphi(Q_i) - q\big\|}{\big|J(Q_i)\big|} && \text{by (38) and } Q_i \in B_2 \\
&\le \|Q_i\| + C_{91}c_i\big|J(Q_i)\big| && \text{by (c) for } i \\
&\le \|Q_i\| + C_{91}C_{90}c_i && \text{since } Q_i \in B_2 \\
&\le b_1 + C_{90}C_{91}\sum_{j=0}^{i} c_j && \text{by (a) for } i.
\end{aligned}$$

This proves (a) for $i+1$.

Since φ and J are analytic, there are constants $C_{92}, C_{93} > 0$ such that for all $x, x' \in B_2$ we have

$$\big\|J(x') - J(x)\big\| \le C_{92}\|x' - x\|, \tag{39}$$

$$\big\|\varphi(x') - \varphi(x) - (D\varphi)(x)(x' - x)\big\| \le C_{93}\|x' - x\|^2. \tag{40}$$

Since $Q_{i+1}, Q_i \in B_2$, by (39) we have

$$\begin{aligned}
\big|J(Q_{i+1}) - J(Q_i)\big| &\le C_{92}|Q_{i+1} - Q_i| \\
&= C_{92}|D\varphi(Q_i)^{-1}(\varphi(Q_i) - q)| \\
&\le C_{92}C_{91}\frac{\big\|\varphi(Q_i) - q\big\|}{\big|J(Q_i)\big|}.
\end{aligned} \tag{41}$$

Suppose that $J(Q_{i+1}) = 0$. Then we get

$$\big|J(Q_i)\big|^2 \le C_{92}C_{91}\big\|\varphi(Q_i)-q\big\| \le C_{92}C_{91}c_i\big|J(Q_i)\big|^2 \le C_{92}C_{91}\eta\big|J(Q_i)\big|^2.$$

Since $J(Q_i)\neq 0$ by induction hypothesis, this does not happen if η is small enough so that $C_{92}C_{91}\eta<1$. This proves (b) for $i+1$.

We estimate

$$\begin{aligned}|J(Q_{i+1})| &\ge \big|J(Q_i)\big|\left(1-C_{92}C_{91}\frac{\big\|\varphi(Q_i)-q\big\|}{\big|J(Q_i)\big|^2}\right) \quad \text{from (41)},\\ &\ge \big|J(Q_i)\big|(1-C_{92}C_{91}c_i) \quad \text{from (c) for } i. \end{aligned} \tag{42}$$

Further, since $Q_{i+1},Q_i\in B_2$, we can use (40) to deduce that

$$\Big\|\underbrace{\varphi(Q_{i+1})-\varphi(Q_i)-(D\varphi)(Q_i)(Q_{i+1}-Q_i)}_{\text{By definition of } Q_{i+1},\text{ this equals } \varphi(Q_{i+1})-q.}\Big\| \le C_{93}\|Q_{i+1}-Q_i\|^2. \tag{43}$$

We can estimate the right-hand side of (43) as follows.

$$\begin{aligned} C_{93}\|Q_{i+1}-Q_i\|^2 &= C_{93}\big\|(D\varphi)(Q_i)^{-1}(\varphi(Q_i)-q)\big\|^2\\ &\le C_{93}C_{91}^2\frac{\big\|\varphi(Q_i)-q\big\|^2}{\big|J(Q_i)\big|^2}\\ &\le C_{93}C_{91}^2\frac{c_i^2\big|J(Q_i)\big|^4}{\big|J(Q_i)\big|^2} \qquad \text{by (c) for } i,\\ &\le C_{93}C_{91}^2\big|J(Q_i)\big|^2c_i^2. \end{aligned}$$

Thus we get

$$\begin{aligned}\big\|\varphi(Q_{i+1})-q\big\| &\le C_{93}C_{91}^2\frac{\big\|\varphi(Q_i)-q\big\|^2}{\big|J(Q_i)\big|^2}\\ &\le C_{93}C_{91}^2\big|J(Q_i)\big|^2c_i^2.\end{aligned} \tag{44}$$

By (42), we have

$$\begin{aligned}\big\|\varphi(Q_{i+1})-q\big\| &\le C_{93}C_{91}^2\frac{\big|J(Q_{i+1})\big|^2}{(1-C_{92}C_{91}c_i)^2}c_i^2\\ &= \frac{C_{93}C_{91}^2}{(1-C_{92}C_{91}c_i)^2}\eta^{\epsilon(2-\epsilon)^i}\big|J(Q_{i+1})\big|^2\eta^{(2-\epsilon)^{i+1}}\\ &= \frac{C_{93}C_{91}^2}{(1-C_{92}C_{91}c_i)^2}\eta^{\epsilon(2-\epsilon)^i}\big|J(Q_{i+1})\big|^2c_{i+1}\\ &\le \frac{C_{93}C_{91}^2}{(1-C_{92}C_{91}\eta)^2}\eta^{\epsilon}\big|J(Q_{i+1})\big|^2c_{i+1}\end{aligned}$$

Thus if we choose η small enough so that it satisfies

$$\frac{C_{93}C_{91}^2}{(1-C_{92}C_{91}\eta)^2}\eta^{\epsilon} \le 1,$$

then we get

$$\big\|\varphi(Q_{i+1}) - q\big\| \le \big|J(Q_{i+1})\big|^2 c_{i+1},$$

and we are done with the proof of Claim 4. □

We can use Claim 4 to prove that the sequence Q_i is a Cauchy sequence via the following calculation:

$$\begin{aligned}
\|Q_n - Q_m\| &\le \sum_{i=m}^{n-1} \|Q_{i+1} - Q_i\| \\
&\le \sum_{i=m}^{n-1} \Big(\|Q_{i+1} - q\| + \|Q_i - q\| \Big) \\
&\le \sum_{i=m}^{n-1} \Big(c_{i+1}\big|J(Q_{i+1})\big|^2 + c_i\big|J(Q_i)\big|^2 \Big) \quad \text{from Claim 4(c)}, \\
&\le \sum_{i=m}^{n-1} C_{90}^2(c_{i+1} + c_i) \quad \text{from (37), since } Q_i, Q_{i+1} \in B_2, \\
&\xrightarrow[n,m\to\infty]{} 0 \quad \text{since } c_i = \eta^{(2-\epsilon)^i} \text{ and } 0 < \eta < 1.
\end{aligned}$$

Hence the limit

$$Q := \lim_{i\to\infty} Q_i \in B_2$$

exists. Further, using the continuity of $D\varphi$ and the definition of Q_{i+1}, we see that

$$\lim_{i\to\infty} \varphi(Q_i) - q = \lim_{i\to\infty} (D\varphi)(Q_i)(Q_{i+1} - Q_i) = (D\varphi)(Q)(Q - Q) = 0,$$

which proves that

$$\varphi(Q) = \lim_{i\to\infty} \varphi(Q_i) = q.$$

It remains to show that Q is close to P. By (38), we have

$$\|Q - P\| = \|Q - Q_0\| \le \sum_{i\ge 0} C_{91} \frac{\big\|\varphi(Q_i) - q\big\|}{\big|J(Q_i)\big|}.$$

Using (42) and (44) gives

$$\begin{aligned}\frac{\|\varphi(Q_{i+1})-q\|}{|J(Q_{i+1})|} &\le C_{93}C_{91}^2\frac{\|\varphi(Q_i)-q\|^2}{|J(Q_i)|^2}\frac{1}{|J(Q_i)|(1-C_{92}C_{91}c_i)}\\ &\le \frac{C_{93}C_{91}^2}{1-C_{92}C_{91}c_i}\cdot c_i\cdot\frac{\|\varphi(Q_i)-q\|}{|J(Q_i)|}\\ &\le \frac{C_{93}C_{91}^2}{1-C_{92}C_{91}\eta}\cdot\eta\cdot\frac{\|\varphi(Q_i)-q\|}{|J(Q_i)|}.\end{aligned}$$

We take η sufficiently small so that we have

$$\tau := \frac{C_{93}C_{91}^2}{1-C_{92}C_{91}\eta}\eta < 1.$$

Then we get

$$\frac{\|\varphi(Q_i)-q\|}{|J(Q_i)|} \le \tau^i\frac{\|\varphi(P)-q\|}{|J(P)|},$$

and hence

$$\|Q-P\| \le C_{91}\frac{1}{1-\tau}\frac{\|\varphi(P)-q\|}{|J(P)|},$$

which completes the proof of Proposition 9.1. □

References

[1] Lei Fu, Etale Cohomology Theory, revised edition, Nankai Tracts in Mathematics, vol. 14, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015, x+611 pp.
[2] Liang-Chung Hsia, Joseph H. Silverman, A quantitative estimate for quasi-integral points in orbits, Pac. J. Math. 249 (2) (2011) 321–342.
[3] Serge Lang, Fundamentals of Diophantine Geometry, Springer-Verlag, New York, 1983.
[4] Serge Lang, Algebraic Number Theory, second edition, Graduate Texts in Mathematics, vol. 110, Springer-Verlag, New York, 1994.
[5] Hideyuki Matsumura, Commutative Algebra, second edition, Mathematics Lecture Note Series, vol. 56, Benjamin/Cummings Publishing Co., Inc., Reading, Mass, 1980, xv+313 pp.
[6] Hideyuki Matsumura, Commutative Ring Theory. Translated from the Japanese by M. Reid, second edition, Cambridge Studies in Advanced Mathematics, vol. 8, Cambridge University Press, Cambridge, 1989, xiv+320 pp.
[7] Yohsuke Matsuzawa, Height functions associated with closed subschemes, preprint, https://arxiv.org/abs/2008.08153, 2020, in preparation.
[8] Joseph H. Silverman, Arithmetic distance functions and height functions in Diophantine geometry, Math. Ann. 279 (2) (1987) 193–216.
[9] Joseph H. Silverman, A quantitative version of Siegel's theorem: integral points on elliptic curves and Catalan curves, J. Reine Angew. Math. 378 (1987) 60–100.
[10] Joseph H. Silverman, Integer points, Diophantine approximation, and iteration of rational maps, Duke Math. J. 71 (3) (1993) 793–829.