Math. Ann. 335, 285–310 (2006) **Mathematische Annalen**

DOI: 10.1007/s00208-006-0750-y

Canonical height functions for affine plane automorphisms

Shu Kawaguchi

Received: 8 November 2005 / Revised version: 4 January 2006
Published online: 7 April 2006 – © Springer-Verlag 2006

Abstract. Let $f : \mathbb{A}^2 \to \mathbb{A}^2$ be a polynomial automorphism of dynamical degree $\delta \geq 2$ over a number field K. We construct height functions defined on $\mathbb{A}^2(\overline{K})$ that transform well relative to f, which we call canonical height functions for f. These functions satisfy the Northcott finiteness property, and a $\overline{K}$-valued point on $\mathbb{A}^2(\overline{K})$ is f-periodic if and only if its height is zero. As an application, we give an estimate on the number of points with bounded height in an infinite f-orbit.

Mathematics Subject Classification (2000): Primary: 11G50, Secondary: 32H50

Key words. canonical height, affine plane automorphism, Hénon map

Introduction and the statement of the main results

One of the basic tools in Diophantine geometry is the theory of height functions. On Abelian varieties defined over a number field, Néron and Tate developed the theory of canonical height functions that behave well relative to the $[n]$-th power map (cf. [10, Chap. 5]). On certain K3 surfaces with two involutions, Silverman [15] developed the theory of canonical height functions that behave well relative to the two involutions. For the theory of canonical height functions on some other projective varieties, see for example [2], [17], [8]. In this paper, we show the existence of canonical height functions on the affine plane for polynomial automorphisms of dynamical degree ≥ 2.

Consider a polynomial automorphism $f : \mathbb{A}^2 \to \mathbb{A}^2$ given by

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix},$$

where $p(x, y)$ and $q(x, y)$ are polynomials in two variables. The degree d of f is defined by $d := \max\{\deg p, \deg q\}$. The dynamical degree δ of f is defined by

S. Kawaguchi
Department of Mathematics, Faculty of Science, Kyoto University, Kyoto, 606-8502, Japan
(e-mail: kawaguch@math.kyoto-u.ac.jp)

$$\delta := \lim_{n\to+\infty} \left(\deg f^n\right)^{\frac{1}{n}},$$

which is an integer with $1 \le \delta \le d$. We generally assume that $d \ge 2$.

For any polynomial (auto)morphism $f : \mathbb{A}^2 \to \mathbb{A}^2$, we write $\overline{f} : \mathbb{P}^2 \to \mathbb{P}^2$ for the extension of f to $\mathbb{P}^2$. Recall that a polynomial automorphism $f : \mathbb{A}^2 \to \mathbb{A}^2$ is said to be *regular* if the unique point of indeterminacy of $\overline{f}$ is different from the unique point of indeterminacy of $\overline{f^{-1}}$. Then regular polynomial automorphisms are precisely the polynomial automorphisms with $\delta = d$. In the space of polynomial automorphisms of degree d, regular polynomial automorphisms, which include the Hénon maps, constitute general members.

The other extreme is polynomial automorphisms of dynamical degree $\delta = 1$. Such automorphisms are precisely *triangularizable* automorphisms, i.e., polynomial automorphisms $f : \mathbb{A}^2 \to \mathbb{A}^2$ that are conjugate, in the group of polynomial automorphisms, to polynomial automorphisms of the form

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + P(y) \\ by + c \end{pmatrix},$$

where $ab \ne 0$ and $P(y)$ is a polynomial in y. For more details, see the survey of Sibony [13] and the references therein. See also §3.

Over a number field, Silverman [16] studied arithmetic properties of quadratic Hénon maps, and then Denis [3] studied arithmetic properties of Hénon maps and some classes of polynomial automorphisms. Marcello [11], [12] studied arithmetic properties of some other classes of polynomial automorphisms of the affine spaces, including regular polynomial automorphisms.

Our first result shows the existence of height functions that behave well relative to polynomial automorphisms of $\mathbb{A}^2$.

Theorem A. *Let $f : \mathbb{A}^2 \to \mathbb{A}^2$ be a polynomial automorphism of dynamical degree $\delta \ge 2$ over a number field K.* (*This is equivalent to saying that f is a polynomial automorphism that is not triangularizable.*) *Let $h : \mathbb{A}^2(\overline{K}) \to \mathbb{R}$ be the naive logarithmic height function. Then there exists a function $\widehat{h} : \mathbb{A}^2(\overline{K}) \to \mathbb{R}$ with the following properties*:[1]

(i) $\widehat{h} \gg\ll h$;
(ii) $\widehat{h} \circ f + \widehat{h} \circ f^{-1} = \left(\delta + \frac{1}{\delta}\right)\widehat{h}.$

We call any function $\widehat{h}$ satisfying (i) *and* (ii) *a* canonical height function *for f. Then $\widehat{h}$ enjoys the following uniqueness property*: *if $\widehat{h}'$ is also a canonical height function for f and satisfies $\widehat{h}' = \widehat{h} + O(1)$, then $\widehat{h}' = \widehat{h}$.*

Recall that a point $x \in \mathbb{A}^2(\overline{K})$ is said to be f-periodic if $f^m(x) = x$ for some positive integer m. Property (i) in Theorem A implies that $\widehat{h}$ satisfies the Northcott

[1] Here $\gg\ll$ in (i) means that there are positive constants a_1, a_2 and constants b_1, b_2 so that $a_1 h + b_1 \le \widehat{h} \le a_2 h + b_2$.

finiteness property. Namely, for any positive number M and positive integer D, the set

$$\{x \in \mathbb{A}^2(\overline{K}) \mid [K(x) : K] \le D,\ \widehat{h}(x) \le M\}$$

is finite. This leads to the following corollary, which shows that the set of $\overline{K}$-valued f-periodic points is not only a set of bounded height but also characterized as the set of height zero with respect to a canonical height function for f.

Corollary B. *With the notation and assumptions as in Theorem A,*
(1) $\widehat{h}(x) \ge 0$ *for all* $x \in \mathbb{A}^2(\overline{K})$.
(2) $\widehat{h}(x) = 0$ *if and only if* x *is* f*-periodic.*

As an application of canonical height functions, we obtain an estimate on the number of points with bounded height in an infinite f-orbit. First we introduce some notation and terminology. For a canonical height function $\widehat{h}$ for f, we set

$$\widehat{h}^+ = \frac{1}{1-\delta^{-2}}\left(\widehat{h} - \frac{1}{\delta}\widehat{h}\circ f^{-1}\right), \quad \widehat{h}^- = \frac{1}{1-\delta^{-2}}\left(\widehat{h} - \frac{1}{\delta}\widehat{h}\circ f\right).$$

Then $\widehat{h}^{\pm}$ are non-negative functions, and $\widehat{h}^+(x) = 0$ if and only if $\widehat{h}^-(x) = 0$ if and only if x is f-periodic (cf. Corollary 4.3). For a point $x \in \mathbb{A}^2(\overline{K})$, let $O_f(x) := \{f^l(x) \mid l \in \mathbb{Z}\}$ denote the f-orbit of x. For a non f-periodic point $x \in \mathbb{A}^2(\overline{K})$, we set

$$\widehat{h}(O_f(x)) = \frac{\log\left(\widehat{h}^+(x)\widehat{h}^-(x)\right)}{\log\delta}. \tag{0.1}$$

Then the right-hand side of (0.1) depends only on the orbit $O_f(x)$ and the choice of the height function $\widehat{h}$, and not on the particular choice of the point x in the orbit. Moreover, as a function of x, we have $\widehat{h}(O_f(x)) \gg\ll \min_{y\in O_f(x)} \log\widehat{h}(y)$ on $\mathbb{A}^2(\overline{K}) \setminus \{f\text{-periodic points}\}$ (cf. Lemma 5.1).

For regular polynomial automorphisms of degree $d \ge 2$, it is known that, for any non f-periodic point $x \in \mathbb{A}^2(\overline{K})$, one has

$$\lim_{T\to+\infty} \frac{\#\{y \in O_f(x) \mid h(y) \le T\}}{\log T} = \frac{2}{\log d}.$$

(See [16, Theorem C], [3, Théorème 2], and [12, Théorème A].) The next theorem gives a refinement and generalization.

Theorem C. *Let* $f : \mathbb{A}^2 \to \mathbb{A}^2$ *be a polynomial automorphism of dynamical degree* $\delta \ge 2$ *over a number field* K. *Then for all infinite orbits* $O_f(x)$,

$$\#\{y \in O_f(x) \mid h(y) \le T\} = \frac{2}{\log\delta}\log T - \widehat{h}(O_f(x)) + O(1) \tag{0.2}$$

as $T \to +\infty$. *Here the* $O(1)$ *bound depends only on* f *and the choice of* $\widehat{h}$, *independent of the orbit* $O_f(x)$.

It is interesting that the dynamical degree of f appears in the right-hand side of (0.2). We remark that, when f is not regular, i.e., $(2 \leq)\ \delta < \deg f$, even the following weaker estimate seems new:

$$\lim_{T \to +\infty} \frac{\#\{y \in O_f(x) \mid h(y) \leq T\}}{\log T} = \frac{2}{\log \delta}.$$

The contents of this paper is as follows. In §1 we briefly review some properties of height functions. In §2 we show that if f is a regular polynomial automorphism of degree $d \geq 2$ then there is a constant c such that

$$h(f(x)) + h(f^{-1}(x)) \geq \left(d + \frac{1}{d}\right) h(x) - c \tag{0.3}$$

for all $x \in \mathbb{A}^2(\overline{K})$. In §3 we recall Hénon maps, Friedland–Milnor's theorem on the conjugacy classes of polynomial automorphisms, and some properties of dynamical degrees of polynomial automorphisms. In §4 we prove Theorem A and Corollary B in a more general setting of polynomial automorphisms of $\mathbb{A}^n$ whose conjugates satisfy an inequality similar to (0.3). In §5 we prove Theorem C in this more general setting. On certain K3 surfaces, Silverman counted the number of points with bounded height in a given infinite chain ([15, §3]). Our method of proof of Theorem C is inspired by his method.

1. Quick review on height theory

In this section, we briefly review the properties of height functions that we will use in this paper.

Let K be a number field and O_K its ring of integers. For $x = (x_0 : \cdots : x_n) \in \mathbb{P}^n(K)$, the naive logarithmic height of x is defined by

$$h(x) = \frac{1}{[K:\mathbb{Q}]} \left[\sum_{P \in \operatorname{Spec}(O_K) \setminus \{0\}} \max_{0 \leq i \leq n} \{-\operatorname{ord}_P(x_i)\} \log \#(O_K/P) + \sum_{\sigma : K \hookrightarrow \mathbb{C}} \max_{0 \leq i \leq n} \{\log |\sigma(x_i)|\} \right].$$

This definition naturally extends to all points $x \in \mathbb{P}^n(\overline{\mathbb{Q}})$ as to give the naive logarithmic height function $h : \mathbb{P}^n(\overline{\mathbb{Q}}) \to \mathbb{R}$.

We begin by recalling the following two basic properties of height functions.

Theorem 1.1 (Northcott's finiteness theorem, [14] Corollary 3.4). *For any positive number M and positive integer D, the set*

$$\left\{ x \in \mathbb{P}^n(\overline{\mathbb{Q}}) \,\middle|\, [\mathbb{Q}(x) : \mathbb{Q}] \leq D,\ h(x) \leq M \right\}$$

is finite.

Theorem 1.2 ([14] Theorem 3.3, [10] Chap. 4, Prop. 5.2).

(1) (Height machine) *There is a unique way to attach, for any projective variety X defined over $\overline{\mathbb{Q}}$, a map*

$$h_X : \mathrm{Pic}(X) \longrightarrow \frac{\{\textit{real-valued functions on } X(\overline{\mathbb{Q}})\}}{\{\textit{real-valued bounded functions on } X(\overline{\mathbb{Q}})\}}, \quad L \mapsto h_{X,L}$$

with the following properties:

(i) $h_{X,L\otimes M} = h_{X,L} + h_{X,M} + O(1)$ *for any* $L, M \in \mathrm{Pic}(X)$;

(ii) *If* $X = \mathbb{P}^n$ *and* $L = \mathcal{O}_{\mathbb{P}^n}(1)$, *then* $h_{\mathbb{P}^n,\mathcal{O}_{\mathbb{P}^n}(1)} = h + O(1)$;

(iii) *If* $f : Y \to X$ *is a morphism of projective varieties and* L *is a line bundle on* X, *then* $h_{Y,f^*L} = h_{X,L} \circ f + O(1)$.

(2) (Positivity of height) *Let* X *be a projective variety defined over* $\overline{\mathbb{Q}}$ *and* L *a line bundle on* X. *We set* $B = \mathrm{Supp}(\mathrm{Coker}(H^0(X, L) \otimes \mathcal{O}_X \to L))$. *Then there exists a constant* c_1 *such that* $h_{X,L}(x) \geq c_1$ *for all* $x \in (X \setminus B)(\overline{\mathbb{Q}})$.

A rational map $f = [F_0 : F_1 : \cdots : F_n] : \mathbb{P}^n \dashrightarrow \mathbb{P}^n$ defined over $\overline{\mathbb{Q}}$ is said to be of degree d if the F_i's are homogeneous polynomials of degree d over $\overline{\mathbb{Q}}$, with no common factors. Let $Z_f \subset \mathbb{P}^n(\overline{\mathbb{Q}})$ denote the locus of indeterminacy.

Theorem 1.3 ([10] Chap. 4, Lemma 1.6). *Let* $f : \mathbb{P}^n \dashrightarrow \mathbb{P}^n$ *be a rational map of degree* d *defined over* $\overline{\mathbb{Q}}$. *Then there exists a constant* c_2 *such that*

$$h(f(x)) \leq d\, h(x) + c_2$$

for all $x \in \mathbb{P}^n(\overline{\mathbb{Q}}) \setminus Z_f$.

2. Geometric properties of regular polynomial automorphisms

In this section, we show (0.3) for regular polynomial automorphisms of $\mathbb{A}^2$. First we recall the definition of regular polynomial automorphisms of $\mathbb{A}^2$. Consider a polynomial automorphism of degree $d \geq 2$ of the form

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p(x, y) \\ q(x, y) \end{pmatrix},$$

where $p(x, y)$ and $q(x, y)$ are polynomials in two variables, and d is the maximum of $\deg p$ and $\deg q$. Let $\overline{f} : \mathbb{P}^2 \dashrightarrow \mathbb{P}^2$ be the extension of f given in homogeneous coordinates as

$$\overline{f}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} Z^d p(X/Z, Y/Z) \\ Z^d q(X/Z, Y/Z) \\ Z^d \end{bmatrix}.$$

Let $f^{-1} : \mathbb{A}^2 \to \mathbb{A}^2$ be the inverse of f, and $\overline{f^{-1}} : \mathbb{P}^2 \dashrightarrow \mathbb{P}^2$ be its extension. We denote by H the line at infinity.

Since $\overline{f}$ is not a morphism, there is a point $\mathbf{p}$ on $\mathbb{P}^2$ at which $\overline{f}$ is not defined. Then $\mathbf{p}$ lies on H, and there is a curve C on $\mathbb{P}^2$ that contracts to $\mathbf{p}$ by $\overline{f^{-1}}$, i.e., $\overline{f^{-1}}(C) = \mathbf{p}$ (cf. [1, Lemme II.10]). Since f^{-1} is an automorphism on $\mathbb{A}^2$, C must be equal to H. Hence $\mathbf{p} = \overline{f^{-1}}(H)$, which implies that $\mathbf{p} \in H$ is the unique point of indeterminacy of $\overline{f}$. Similarly $\overline{f^{-1}}$ has a unique point of indeterminacy on H, denoted by $\mathbf{q}$. A polynomial automorphism of $\mathbb{A}^2$ is said to be *regular* if $\mathbf{p} \neq \mathbf{q}$.

By elimination of indeterminacy, by successively blowing up points starting from $\mathbf{p} \in \mathbb{P}^2$, we obtain a projective surface W and a composition of blow-ups $\pi_W : W \to \mathbb{P}^2$ such that $\overline{f} \circ \pi_W : W \dashrightarrow \mathbb{P}^2$ becomes a morphism. We take W so that the number of blow-ups needed for elimination of indeterminacy is minimal. Noting that π_W induces an isomorphism $\pi_W^{-1}(\mathbb{P}^2 \setminus \{\mathbf{p}\}) \to \mathbb{P}^2 \setminus \{\mathbf{p}\}$, we take $\mathbf{q}' \in W$ with $\pi_W(\mathbf{q}') = \mathbf{q}$. In a parallel way as for $\mathbf{p}$, $\overline{f^{-1}} \circ \pi_W : W \dashrightarrow \mathbb{P}^2$ becomes a morphism after a finite number of blow-ups starting at $\mathbf{q}'$.

To summarize, there is a projective surface V obtained by successive blow-ups of $\mathbb{P}^2$ at $\mathbf{p}$ and then successive blow-ups at $\mathbf{q}$ in a parallel way as for $\mathbf{p}$ such that, if $\pi : V \to \mathbb{P}^2$ denotes the morphism of blow-ups, then $\overline{f} \circ \pi$ extends to a morphism $\varphi : V \to \mathbb{P}^2$ and $\overline{f^{-1}} \circ \pi$ extends to a morphism $\psi : V \to \mathbb{P}^2$. As for W, we take V so that the number of blow-ups needed for elimination of indeterminacy is minimal.

$$\begin{tikzcd} & V \arrow[dl, "\psi"'] \arrow[d, "\pi"] \arrow[dr, "\varphi"] & \\ \mathbb{P}^2 & \mathbb{P}^2 \arrow[l, dashed, "\overline{f^{-1}}"'] \arrow[r, dashed, "\overline{f}"] & \mathbb{P}^2 \end{tikzcd} \tag{2.1}$$

Before stating the next theorem, we fix some notation and terminology. Let $\rho : Y \to X$ be a morphism of smooth projective surfaces. For an irreducible curve C on Y, its push-forward is defined by

$$\rho_*(C) := \begin{cases} \deg(\rho|_C : C \to f(C))\, f(C) & \text{(if } f(C) \text{ is a curve),} \\ 0 & \text{(if } f(C) \text{ is a point).} \end{cases}$$

This extends linearly to the homomorphism ρ_* from divisors on Y to divisors on X. For two divisors Z_1, Z_2 on Y, we write $Z_1 \geq Z_2$ if $Z_1 - Z_2$ is effective. A divisor Z on Y is said to be nef if $Z \cdot C \geq 0$ for any curve C on Y.

Theorem 2.1. *Let $f : \mathbb{A}^2 \to \mathbb{A}^2$ be a regular polynomial automorphism of degree $d \geq 2$. Let H denote the line at infinity. Let V be as in (2.1). Then, as a $\mathbb{Q}$-divisor on V,*

$$D := \varphi^* H + \psi^* H - \left(d + \frac{1}{d}\right)\pi^* H$$

is effective.

Proof. The proof we present here, which simplifies the proof we gave in the initial draft, is due to Noboru Nakayama.

As above, let $\pi_W : W \to \mathbb{P}^2$ be a composition of blow-ups of $\mathbb{P}^2$ starting at $\mathbf{p}$ such that $\varphi_W := \overline{f} \circ \pi_W : W \dashrightarrow \mathbb{P}^2$ becomes a morphism. Let H_W be the proper transform of H by π_W, and E_W the exceptional curve on W given by the last blow-up of π_W. Since φ_W is a morphism and W is taken so that the number of blow-ups is minimal, we see that φ_W sends E_W to H isomorphically.

We consider $\pi_W^* H$ and $\varphi_W^* H$. We write $\pi_W^* H = aH_W + bE_W + M_W$ and $\varphi_W^* H = a' H_W + b' E_W + I_W$, where a, b, a', b' are non-negative integers, and M_W, I_W are effective divisors on W with $\mathrm{Supp}(E_W) \not\subseteq \mathrm{Supp}(M_W)$, $\mathrm{Supp}(E_W) \not\subseteq \mathrm{Supp}(I_W)$ such that M_W, I_W are contracted to $\mathbf{p}$ by π_W.

We determine a, b, a', b'. Since π_W is a birational morphism, $\pi_{W*}\pi_W^* H = H$. It follows that $a = 1$. Similarly, $\varphi_{W*}\varphi_W^* H = H$ yields $b' = 1$. On the other hand, let $[H]$ denote the cohomology class of H in $H^2(\mathbb{P}^2, \mathbb{Z})$. Since the degree of $f : \mathbb{A}^2 \to \mathbb{A}^2$ is d, we get $\varphi_{W*}\pi_W^*[H] = d[H] \in H^2(\mathbb{P}^2, \mathbb{Z})$. It follows that $\varphi_{W*}\pi_W^* H = dH$ and $b = d$. Since the degree of $f^{-1} : \mathbb{A}^2 \to \mathbb{A}^2$ is also d, we get $\pi_{W*}\varphi_W^* H = dH$ and $a' = d$. Putting together, we have

$$\begin{aligned} \pi_W^* H &= H_W + dE_W + M_W, \\ \varphi_W^* H &= dH_W + E_W + I_W. \end{aligned}$$

Since the effective divisor $\pi_W^* H$ is nef, Lemma 2.2 below yields that

$$\varphi_W^*(dH) = \varphi_W^*(\varphi_{W*}\pi_W^* H) = (\varphi_W^*\varphi_{W*})\pi_W^* H \geq \pi_W^* H.$$

We thus get

$$dI_W \geq M_W. \tag{2.2}$$

In a parallel way as for $\mathbf{p}$, let $\pi_U : U \to \mathbb{P}^2$ be a composition of blow-ups of $\mathbb{P}^2$ starting at $\mathbf{q}$ such that $\psi_U := \overline{f^{-1}} \circ \pi_U : U \dashrightarrow \mathbb{P}^2$ becomes a morphism. Let H_U be the proper transform of H by π_U, and F_U the exceptional curve on U given by the last blow-up of π_U. The morphism ψ_U sends F_U to H isomorphically. In a parallel way, we get

$$\begin{aligned} \pi_U^* H &= H_U + dF_U + N_U, \\ \psi_U^* H &= dH_U + F_U + J_U, \\ dJ_U &\geq N_U, \end{aligned} \tag{2.3}$$

where N_U, J_U are effective divisors on U with $\mathrm{Supp}(F_U) \not\subseteq \mathrm{Supp}(N_U)$, $\mathrm{Supp}(F_U) \not\subseteq \mathrm{Supp}(J_U)$ such that N_U, J_U are contracted to $\mathbf{q}$ by π_U.

By the construction of V, there are birational morphisms $\alpha : V \to W$ and $\beta : V \to U$ such that the following diagram is commutative.

$$\begin{tikzcd}
& & V \arrow[dl, "\beta"'] \arrow[dr, "\alpha"] \arrow[dd, "\pi"] \arrow[ddll, bend right, "\psi"'] \arrow[ddrr, bend left, "\varphi"] & & \\
& U \arrow[dl, "\psi_U"'] \arrow[dr, "\pi_U"] & & W \arrow[dl, "\pi_W"'] \arrow[dr, "\varphi_W"] & \\
\mathbb{P}^2 & & \mathbb{P}^2 \arrow[ll, dashed, "\overline{f^{-1}}"'] \arrow[rr, dashed, "\overline{f}"] & & \mathbb{P}^2
\end{tikzcd}$$

Let $H^\#$ on V be the proper transform of H by π. Let E, M, I on V be the proper transforms of E_W, M_W, I_W by α, respectively. Let F, N, J be the proper transforms of F_U, N_U, J_U by β, respectively. Then the following equalities hold:

$$\pi^* H = H^\# + dE + dF + M + N, \tag{2.4}$$

$$\varphi^* H = d(H^\# + dF + N) + E + I, \tag{2.5}$$

$$\psi^* H = d(H^\# + dE + M) + F + J. \tag{2.6}$$

By (2.4)–(2.6), we get

$$\begin{aligned} D &= \varphi^* H + \psi^* H - \left(d + \frac{1}{d}\right) \pi^* H \\ &= \left(d - \frac{1}{d}\right) H^\# - \frac{1}{d} M + I - \frac{1}{d} N + J. \end{aligned}$$

Since $dI \geq M$ and $dJ \geq N$ by (2.2) and (2.3), we see that D is effective. $\square$

Lemma 2.2. *Let $\rho : Y \to X$ be a birational morphism of smooth projective surfaces. Let Z be an effective divisor on Y. If Z is nef, then $\rho^* \rho_* Z \geq Z$.*

Proof. First we treat a case when ρ is the blow-up of X at a point $x \in X$. Let E denote the exceptional curve on Y. We write $Z = a_1 C_1 + \cdots + a_k C_k + bE$, where $C_1, \cdots, C_k, E$ are distinct irreducible and reduced curves, and $a_1, \cdots, a_k, b$ are non-negative integers. Then $\rho_* Z = a_1 \rho(C_1) + \cdots + a_k \rho(C_k)$. Hence $\rho^* \rho_* Z = a_1(C_1 + m_1 E) + \cdots + a_k(C_k + m_k E)$, where m_i is the multiplicity of the curve $\rho(C_i)$ at x. Note that $m_i = C_i \cdot E$.

Since Z is nef, we get

$$\begin{aligned} Z \cdot E &= a_1(C_1 \cdot E) + \cdots + a_1(C_k \cdot E) + b(E \cdot E) \\ &= a_1 m_1 + \cdots + a_k m_k - b \geq 0. \end{aligned}$$

Hence $a_1 m_1 + \cdots a_k m_k \geq b$ and we get $\rho^* \rho_* Z \geq Z$.

In general, we decompose ρ into a composition of blow-ups: $\rho = \rho_l \circ \cdots \circ \rho_2 \circ \rho_1$, where each ρ_i is the blow-up at a point. Put $\rho' := \rho_l \circ \cdots \circ \rho_2$, and

$Z' := \rho_{1*}Z$. Since the projection formula yields $(\rho_{1*}Z) \cdot C = Z \cdot (\rho_1^* C)$ for any curve, we see that Z' is nef. Then, by induction, $\rho'^* \rho'_* Z' \geq Z'$. Pulling back by ρ_1, we get $\rho_1^*(\rho'^* \rho'_* Z') \geq \rho_1^* Z'$. Thus

$$\rho^* \rho_* Z = \rho_1^* \rho'^* \rho'_* (\rho_{1*} Z) \geq \rho_1^* (\rho_{1*} Z) \geq Z.$$

Now we prove (0.3).

Theorem 2.3. *Let $f : \mathbb{A}^2 \to \mathbb{A}^2$ be a regular polynomial automorphism of degree $d \geq 2$ defined over a number field K. Then, there exists a constant c such that*

$$h(f(x)) + h(f^{-1}(x)) \geq \left(d + \frac{1}{d}\right) h(x) - c$$

for all $x \in \mathbb{A}^2(\overline{K})$.

Proof. We can prove Theorem 2.3 as in [16, Theorem 3.1]. We take $x \in \mathbb{A}^2(\overline{K})$. Since $\pi : V \to \mathbb{P}^2$ gives an isomorphism $\pi|_{\pi^{-1}(\mathbb{A}^2)} : \pi^{-1}(\mathbb{A}^2) \to \mathbb{A}^2$, there is a unique point $\widetilde{x} \in V$ with $\pi(\widetilde{x}) = x$. By Theorem 2.1, we have

$$\begin{aligned} h_{V, \mathcal{O}_V(\varphi^* H)}(\widetilde{x}) + h_{V, \mathcal{O}_V(\psi^* H)}(\widetilde{x}) = &\left(d + \frac{1}{d}\right) h_{V, \mathcal{O}_V(\pi^* H)}(\widetilde{x}) \\ &+ h_{V, \mathcal{O}_V(D)}(\widetilde{x}) + O(1). \end{aligned}$$

It follows from Theorem 1.2(1) that

$$h_{V, \mathcal{O}_V(\varphi^* H)}(\widetilde{x}) = h_{\mathbb{P}^2, \mathcal{O}_V(H)}(\varphi(\widetilde{x})) + O(1) = h_{\mathbb{P}^2, \mathcal{O}_V(H)}(f(x)) + O(1).$$

We similarly have

$$\begin{aligned} h_{V, \mathcal{O}_V(\psi^* H)}(\widetilde{x}) &= h_{\mathbb{P}^2, \mathcal{O}_V(H)}(f^{-1}(x)) + O(1), \\ h_{V, \mathcal{O}_V(\pi^* H)}(\widetilde{x}) &= h_{\mathbb{P}^2, \mathcal{O}_V(H)}(x) + O(1). \end{aligned}$$

On the other hand, since $\pi(\mathrm{Supp}(D)) \subseteq \mathrm{Supp}(H)$, we have $\widetilde{x} \notin \mathrm{Supp}(D)$. Since D is effective by Theorem 2.1, it follows from Theorem 1.2(2) that there is a constant c_2 independent of $\widetilde{x}$ such that $h_{V, \mathcal{O}_V(D)}(\widetilde{x}) \geq c_2$. Hence we get the assertion. $\square$

3. Hénon maps, conjugacy classes of polynomial automorphisms, and dynamical degrees

In this section, we review Hénon maps, Friedland–Milnor's theorem on the conjugacy classes of polynomial automorphisms, and some properties of dynamical degrees of polynomial automorphisms, which will be used in §4. We also give explicit forms of $\varphi^* H$, $\psi^* H$ and $\pi^* H$ in Theorem 2.1 for Hénon maps.

A Hénon map is a polynomial automorphism of the form

$$f\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}p(x)-ay\\x\end{pmatrix}, \tag{3.1}$$

where $a\neq 0$ and p is a polynomial of degree $d\geq 2$. Let $\overline{f}:\mathbb{P}^2\dashrightarrow\mathbb{P}^2$ (resp. $\overline{f^{-1}}:\mathbb{P}^2\dashrightarrow\mathbb{P}^2$) be the birational extension of f (resp. f^{-1}). Then $\overline{f}$ has the unique point of indeterminacy $\mathbf{p}={}^t[0,1,0]$, and $\overline{f^{-1}}$ has the unique point of indeterminacy $\mathbf{q}={}^t[1,0,0]$. In particular, Hénon maps are examples of regular polynomial automorphisms.

We recall Friedland–Milnor's theorem [5, §2], which is based on Jung's theorem [7]. Let

$$E=\left\{f:\mathbb{A}^2\to\mathbb{A}^2,\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}ax+P(y)\\by+c\end{pmatrix}\;\middle|\;\begin{matrix}a,b\in\overline{\mathbb{Q}}^{\times},c\in\overline{\mathbb{Q}}\\P(y)\in\overline{\mathbb{Q}}[Y]\end{matrix}\right\} \tag{3.2}$$

be the group of triangular automorphisms (also called elementary automorphisms, or de Jonquères automorphisms).

Theorem 3.1 ([5], §2). *Let $f:\mathbb{A}^2\to\mathbb{A}^2$ be a polynomial automorphism over $\overline{\mathbb{Q}}$. Then there is a polynomial automorphism $\gamma:\mathbb{A}^2\to\mathbb{A}^2$ over $\overline{\mathbb{Q}}$ such that $g:=\gamma^{-1}\circ f\circ\gamma$ is one of the following types*:

(i) *g is a triangular automorphism*;
(ii) *g is a composition of Hénon maps.*

Note that Friedland–Milnor proved the theorem over $\mathbb{C}$, but the theorem holds over $\overline{\mathbb{Q}}$ by the specialization argument in [3, Lemme 2].

A polynomial automorphism f is said to be *triangularizable* if it is conjugate to a triangular automorphism.

Here we recall properties of dynamical degrees of polynomial automorphisms $f:\mathbb{A}^2\to\mathbb{A}^2$. The dynamical degree of f is defined by

$$\delta(f):=\lim_{n\to+\infty}\left(\deg f^n\right)^{\frac{1}{n}}$$

(cf. [13, Définition 1.4.7]). Suppose $g=\gamma^{-1}\circ f\circ\gamma$ is conjugate to f. Then, since $g^n=\gamma^{-1}\circ f^n\circ\gamma$, we have $\deg f^n-2\deg\gamma\leq\deg g^n\leq\deg f^n+2\deg\gamma$. It follows that $\delta(f)=\delta(g)$. Thus dynamical degrees depend only on conjugacy classes of polynomial automorphisms.

For polynomial automorphisms $g_1,g_2:\mathbb{A}^2\to\mathbb{A}^2$ with degree at least 2 and their extensions $\overline{g_1},\overline{g_2}:\mathbb{P}^2\dashrightarrow\mathbb{P}^2$, one has

$$\deg(g_1\circ g_2)\leq(\deg g_1)(\deg g_2), \tag{3.3}$$

with equality if and only if the unique point $\mathbf{q}_{g_1}$ of indeterminacy of $\overline{g_1^{-1}}$ is different from the unique point $\mathbf{p}_{g_2}$ of indeterminacy of $\overline{g_2}$ (cf. [13, Proposition 1.4.3]).

We remark that a composition g of Hénon maps is a regular polynomial automorphism, because the indeterminacy point of $\overline{g}$ is ${}^t[0, 1, 0]$ while the indeterminacy point of $\overline{g^{-1}}$ is ${}^t[1, 0, 0]$. We recall the following results.

Proposition 3.2. *Let $f : \mathbb{A}^2 \to \mathbb{A}^2$ be a polynomial automorphism. Let d be the degree of f and δ the dynamical degree of f.*

(1) *δ is an integer with $1 \le \delta \le d$.*
(2) *$\delta = 1$ if and only if f is triangularizable.*
(3) *Suppose $d \ge 2$. Then $\delta = d$ if and only if f is a regular polynomial automorphism.*

Proof. We rely on the results of Furter [4] to give a quick proof. We put $\tau = \frac{\deg(f^2)}{\deg f}$. Then Furter showed that either (i) $\tau \le 1$ or (ii) τ is an integer greater than or equal to 2. Moreover, (i) occurs if and only if f is triangularizable ([4, Proposition 5]). In the case (ii), one has $\deg f^n = \tau^n \cdot \deg f$ ([4, Proposition 4]).

(1) In the case (i), f is conjugate to a triangular automorphism g. Then the definition (3.2) yields that $\deg g^n \le \deg g$, whence $\delta(f) = \delta(g) = 1$. In the case (ii), the dynamical degree of f is equal to an integer $\tau \ge 2$.
(2) It follows from the above proof of (1).
(3) Since d is assumed to be ≥ 2, (3.3) shows that f is a regular polynomial automorphism if and only if $\tau = \deg f$ (≥ 2). Since $\tau = \delta(f)$ if $\tau \ge 2$, we get the assertion. □

Since Hénon maps are basic objects in the dynamics of polynomial automorphisms of $\mathbb{A}^2$ (cf. Theorem 3.1), it is worth giving explicit forms of $\varphi^* H$, $\psi^* H$ and $\pi^* H$ in Theorem 2.1 for Hénon maps of degree $d \ge 2$, as Silverman [16] did for quadratic Hénon maps. In particular, this gives a different proof of Theorem 2.1 in case of Hénon maps.

For this, we need an explicit description of blow-ups at (infinitely near) points on $\mathbb{P}^2$ that resolve the point of indeterminacy of a Hénon map $\overline{f}$. The case $\deg g = 2$ was carried out by Silverman [16, §2], and the general case by Hubbard–Papadopol–Veselov [6, §2] in their compactification of Hénon maps in $\mathbb{C}^2$ as dynamical systems. Let us put together their results in the following theorem. (Note that, for the next theorem, the field of definition of f can be any field, and $p(x)$ need not be monic.)

Theorem 3.3 ([6], §2).

(1) *Let f be a Hénon map in (3.1), and $\overline{f} : \mathbb{P}^2 \dashrightarrow \mathbb{P}^2$ its birational extension. Then $\overline{f}$ becomes well-defined after a sequence of $2d - 1$ blow-ups. Explicitly, blow-ups are described as follows*:
   (i) *First blow-up at* $\mathbf{p}$;
   (ii) *Next blow up at the unique point of indeterminacy, which is given by the intersection of the exceptional divisor and the proper transform of H*;

(iii) *For the next* $d-2$ *times after* (ii), *blow-up at the unique point of indeterminacy, which is given by the intersection of the last exceptional divisor and the proper transform of the first exceptional divisor*;

(iv) *For the next* $d-1$ *times after* (iii), *blow-up at the unique point of indeterminacy, which lies on the last exceptional divisor but not on the proper transform of the other exceptional divisors.*

(2) *Let* $\overline{f_{2d-1}} : W \to \mathbb{P}^2$ *be the extension of the Hénon map after the sequence of* $2d-1$ *blow-ups. Let* $E_i^{'}$ *denote the proper transform of* i*-th exceptional divisor* ($i = 1, \dots, 2d-1$). *Then* $\overline{f_{2d-1}}$ *maps* $E_i^{'}$ ($i = 1, \dots, 2d-2$) *to* $\mathbf{q}$, *while* $E_{2d-1}^{'}$ *is mapped to* H *by an isomorphism.*

(3) $E_1^{'2} = -d$, $E_i^{'2} = -2$ ($i = 2, \dots, 2d-2$), *and* $E_{2d-1}^{'\,2} = -1$.

In particular, for Hénon maps, V in (2.1) is the projective surface obtained by successive $2d-1$ blow-ups of $\mathbb{P}^2$ at $\mathbf{p}$ as in Theorem 3.3 and then successive $2d-1$ blow-ups at $\mathbf{q}$ in a parallel way as in Theorem 3.3.

Let E_i ($1 \le i \le 2d-1$) be the proper transform of i-th exceptional divisor on V on the side of $\mathbf{p}$, and F_j ($1 \le j \le 2d-1$) be the proper transform of j-th exceptional divisor on V on the side of $\mathbf{q}$. Let $H^\#$ be the proper transform of H. The configuration of $H^\#$, E_i and F_j is illustrated in Figure 1.

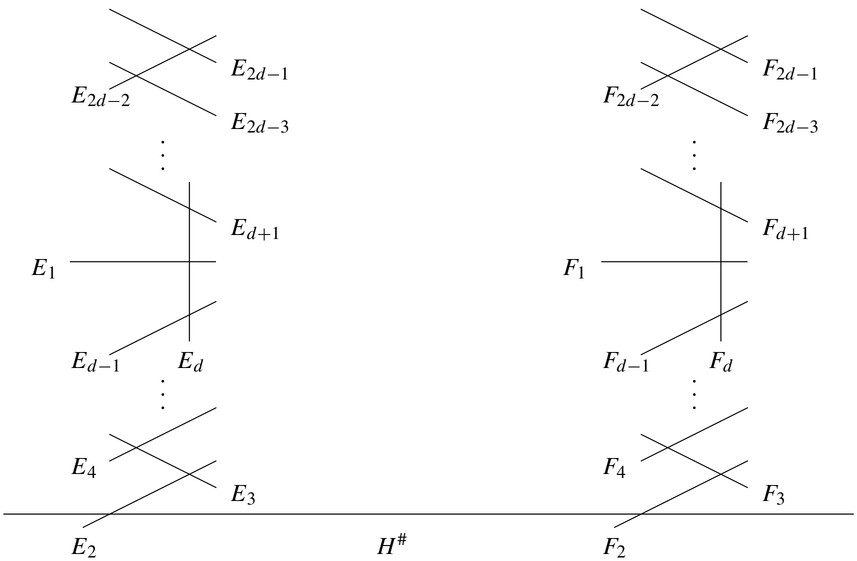

Fig. 1. The configuration after blow-ups. The line $H^\#$ has the self-intersection number -3. The lines E_1 and F_1 have the self-intersection numbers $-d$. The lines $E_2, E_3, \dots, E_{2d-2}$ and $F_2, F_3, \dots, F_{2d-2}$ have the self-intersection numbers -2. The lines E_{2d-1} and F_{2d-1} have the self-intersection numbers -1

Proposition 3.4. *Let $f : \mathbb{A}^2 \to \mathbb{A}^2$ be a Hénon map of degree $d \geq 2$. Let the notation be as above.*

(1) *As divisors on V, we have*

$$\pi^* H = H^\# + \sum_{i=1}^{d} iE_i + \sum_{i=d+1}^{2d-1} dE_i + \sum_{j=1}^{d} jF_j + \sum_{j=d+1}^{2d-1} dF_j,$$

$$\varphi^* H = dH^\# + E_1 + \sum_{i=2}^{d} dE_i + \sum_{i=d+1}^{2d-1} (2d-i)E_i + \sum_{j=1}^{d} jdF_j + \sum_{j=d+1}^{2d-1} d^2 F_j,$$

$$\psi^* H = dH^\# + \sum_{i=1}^{d} idE_i + \sum_{i=d+1}^{2d-1} d^2 E_i + F_1 + \sum_{j=2}^{d} dF_j + \sum_{j=d+1}^{2d-1} (2d-j)F_j.$$

(2) *The effective $\mathbb{Q}$-divisor D in Theorem 2.1 is expressed as*

$$D = \frac{d^2-1}{d} H^\# + \frac{d-1}{d} E_1 + \sum_{i=2}^{d} \frac{d^2-i}{d} E_i + \sum_{i=d+1}^{2d-1} (2d-i-1)E_i$$
$$+ \frac{d-1}{d} F_1 + \sum_{j=2}^{d} \frac{d^2-j}{d} F_j + \sum_{j=d+1}^{2d-1} (2d-j-1)F_j.$$

Proof. We will show the expression for $\varphi^* H$. Since φ maps $H^\#$, E_i $(1 \leq i \leq 2d-2)$ and F_j $(1 \leq j \leq 2d-1)$ to the point $\mathbf{q}$, we have

$$\varphi^* H \cdot H^\# = 0, \qquad \varphi^* H \cdot E_i = 0, \qquad \varphi^* H \cdot F_j = 0$$

for $1 \leq i \leq 2d-2$ and $1 \leq j \leq 2d-1$. Since φ maps E_{2d-1} to H isomorphically, we have

$$\varphi^* H \cdot E_{2d-1} = 1.$$

Noting that the Picard group of V is generated by $H^\#$, E_i, F_j $(1 \leq i, j \leq 2d-1)$, we set $\varphi^* H = aH^\# + \sum_{i=1}^{2d-1} b_i E_i + \sum_{j=1}^{2d-1} c_j F_j$. From the above information and the information of the configuration after blow-ups (cf. Figure 1), we have the system of linear equations

$-3a + b_2 + c_2 = 0;$	$-db_1 + b_d = 0,$ $a - 2b_2 + b_3 = 0,$ $b_{i-1} - 2b_i + b_{i+1} = 0,$ $b_1 + b_{d-1} - 2b_d + b_{d+1} = 0,$ $b_{2d-2} - b_{2d-1} = 1;$	$-dc_1 + c_d = 0,$ $a - 2c_2 + c_3 = 0,$ $c_{j-1} - 2c_j + c_{j+1} = 0,$ $c_1 + c_{d-1} - 2c_d + c_{d+1} = 0,$ $c_{2d-2} - c_{2d-1} = 0;$

where $i = 3, \ldots, d-1, d+1, \ldots, 2d-2$ and $j = 3, \ldots, d-1, d+1, \ldots, 2d-2$. By solving this system, we obtain the expression for $\varphi^* H$. Similarly we obtain the formula for $\psi^* H$. The formula for $\pi^* H$ follows from the construction of V. (We can also show this by using $\pi^* H \cdot H^\# = 1, \pi^* H \cdot E_i = 0$ and $\pi^* H \cdot F_j = 0$ for all i and j.) The assertion (2) follows from (1). $\square$

Remark 3.5. Using classical results of Jung [7] and van der Kulk [9], it is possible to explicitly compute D for any regular polynomial automorphisms f of degree $d \geq 2$, as in Proposition 3.4 for Hénon maps. In this case, coefficients of D are expressed in terms of the polydegree $(d_1, \ldots, d_l)$ of f (cf. [5, §3]). Note that, for Hénon maps f of degree $d \geq 2$, its polydegree is (d), i.e., $l = 1$ and $d_1 = d$.

4. Canonical height functions

In this section, we will prove Theorem A and Corollary B in a more general setting of polynomial automorphisms of $\mathbb{A}^n$. We note that, when $n \geq 3$, the degree of a polynomial automorphism and the degree of its inverse may not be the same. For details about the dynamics of polynomial automorphisms, we refer the reader to the survey [13]. We begin by fixing some notation and terminology.

Definition 4.1. *For two real-valued functions λ, λ' defined on $\mathbb{A}^n(K)$, we write $\lambda \gg\ll \lambda'$ if there exist positive constants a_1, a_2 and constants b_1, b_2 such that $a_1\lambda(x) + b_1 \leq \lambda'(x) \leq a_2\lambda(x) + b_2$ for all $x \in \mathbb{A}^n(K)$.*

Theorem 4.2. *Let $g : \mathbb{A}^n \to \mathbb{A}^n$ be a polynomial automorphism over a number field K. Let δ and δ_- denote the degrees of g and g^{-1}, respectively. Assume that $\delta \geq 2$ and that there exists a constant $c \geq 0$ such that*

$$\frac{1}{\delta} h(g(x)) + \frac{1}{\delta_-} h(g^{-1}(x)) \geq \left(1 + \frac{1}{\delta\delta_-}\right) h(x) - c. \tag{4.1}$$

(1) *The following limits are finite*:

$$\widehat{h}^+(x) = \limsup_{l \to +\infty} \frac{1}{\delta^l} h(g^l(x)), \qquad \widehat{h}^-(x) = \limsup_{l \to +\infty} \frac{1}{\delta_-^l} h(g^{-l}(x)). \tag{4.2}$$

They satisfy

$$\begin{aligned} \widehat{h}^+ \circ g &= \delta\, \widehat{h}^+, & \widehat{h}^- \circ g^{-1} &= \delta_- \widehat{h}^-, \\ \widehat{h}^+ \circ g^{-1} &= \frac{1}{\delta} \widehat{h}^+, & \widehat{h}^- \circ g &= \frac{1}{\delta_-} \widehat{h}^-. \end{aligned} \tag{4.3}$$

(2) *Define $\widehat{h} : \mathbb{A}^n(\overline{K}) \to \mathbb{R}$ by*

$$\widehat{h}(x) = \widehat{h}^+(x) + \widehat{h}^-(x). \tag{4.4}$$

Then $\widehat{h}$ satisfies:

(i) $h(x) \gg\ll \widehat{h}(x)$;

(ii) $\frac{1}{\delta}\widehat{h}(g(x)) + \frac{1}{\delta_-}\widehat{h}\left(g^{-1}(x)\right) = \left(1 + \frac{1}{\delta\delta_-}\right)\widehat{h}(x)$.

(3) *It is clear that* $\widehat{h}^{\pm}$ *and* $\widehat{h}$ *in* (1) *and* (2) *are non-negative functions. We have the following equivalences*:

$$\widehat{h}(x) = 0 \Longleftrightarrow \widehat{h}^{+}(x) = 0 \Longleftrightarrow \widehat{h}^{-}(x) = 0 \Longleftrightarrow x \text{ is } g\text{-periodic}.$$

(4) *Any function* $\widehat{h} : \mathbb{A}^n(\overline{K}) \to \mathbb{R}$ *satisfying* (2-i) *and* (2-ii) *is called a* canonical height function *for g. Suppose that* $\widehat{h}' : \mathbb{A}^n(\overline{K}) \to \mathbb{R}$ *is also a canonical height function for g and satisfies* $\widehat{h}' = \widehat{h} + O(1)$. *Then* $\widehat{h}' = \widehat{h}$.

Corollary 4.3. *Let g be as in the statement of Theorem 4.2, including the assumption that g satisfies (4.1). We denote by* $\widehat{h}_g^{+}, \widehat{h}_g^{-}$ *the functions given in (4.2), and by* $\widehat{h}_g$ *the canonical height function for g given in (4.4). Let* $\gamma : \mathbb{A}^n \to \mathbb{A}^n$ *be a polynomial automorphism over K, and define a polynomial automorphism* $f : \mathbb{A}^n \to \mathbb{A}^n$ *by* $f = \gamma \circ g \circ \gamma^{-1}$.

(1) *Define* $\widehat{h} : \mathbb{A}^n(\overline{K}) \to \mathbb{R}$ *by*

$$\widehat{h}(x) = \widehat{h}_g(\gamma^{-1}(x)).$$

Then $\widehat{h}$ *satisfies*:

(i) $h(x) \gg\ll \widehat{h}(x)$;

(ii) $\frac{1}{\delta}\widehat{h}(f(x)) + \frac{1}{\delta_-}\widehat{h}\left(f^{-1}(x)\right) = \left(1 + \frac{1}{\delta\delta_-}\right)\widehat{h}(x)$.

(*Notice that* δ, δ_- *are respectively the degrees of* g, g^{-1} *and not necessarily equal to the degrees of* f, f^{-1}.)

(2) *Any function* $\widehat{h} : \mathbb{A}^n(\overline{K}) \to \mathbb{R}$ *satisfying the above properties* (i) *and* (ii) *is called a* canonical height function *for f. Then* $\widehat{h}$ *satisfies the uniqueness property as described in Theorem* 4.2(4), *with f in place of g.*

(3) *For any canonical height function* $\widehat{h}$ *for f, we set*

$$\widehat{h}^{+}(x) = \frac{1}{1 - (\delta\delta_-)^{-1}}\left(\widehat{h}(x) - \frac{1}{\delta_-}\widehat{h}(f^{-1}(x))\right),$$
$$\widehat{h}^{-}(x) = \frac{1}{1 - (\delta\delta_-)^{-1}}\left(\widehat{h}(x) - \frac{1}{\delta}\widehat{h}(f(x))\right).$$

Then $\widehat{h}^{\pm}$ *enjoy the transformation formulas (4.3) with f in place of g, and satisfy* $\widehat{h} = \widehat{h}^{+} + \widehat{h}^{-}$.

(4) *Any canonical height function* $\widehat{h}$ *for f and the functions* $\widehat{h}^{\pm}$ *defined in* (3) *are non-negative functions. They satisfy the equivalences as described in Theorem* 4.2(3), *with f in place of g.*

Proof of Theorem A and Corollary B. Admitting Theorem 4.2 and Corollary 4.3, we will prove Theorem A and Corollary B. We may replace K by a finite extension field. Since the dynamical degree δ of $f : \mathbb{A}^2 \to \mathbb{A}^2$ is assumed to be greater than or equal to 2, Theorem 3.1 and Proposition 3.2 yield that there is a polynomial automorphism γ so that $g := \gamma^{-1} \circ f \circ \gamma$ is a composition of Hénon maps. Since a composition of Hénon maps is a regular polynomial automorphism (cf. lines before Proposition 3.2), it follows from Proposition 3.2(3) that the degree of g is equal to the dynamical degree of g. Noting that the dynamical degrees of f and g are the same, this means that $\deg g = \delta$. From Theorem 2.3, g satisfies (4.1), and so we can apply Theorem 4.2 to g. Then Theorem A and Corollary B follow from Corollary 4.3. $\square$

Proof of Theorem 4.2. (1) We remark that the definition (4.1) of $\widehat{h}^{\pm}$ has some similarity to the definition of Green currents on $\mathbb{A}^n(\mathbb{C})$ for g (cf. [13, Définition 2.2.5]), and to Silverman's definition of canonical heights on certain K3 surfaces [15, §3].

Let us see that the values $\widehat{h}^{\pm}(x)$ are finite by showing the following claim.

Claim 4.3.1. There exist constants $c^{\pm}$ such that $\widehat{h}^{\pm}(x) \leq h(x) + c^{\pm}$ for all $x \in \mathbb{A}^n(\overline{K})$.

Proof. By Theorem 1.3, there exists a constant c_2 such that $\frac{1}{\delta}h(g(x)) \leq h(x) + \frac{c_2}{\delta}$ for all $x \in \mathbb{A}^n(\overline{K})$. We show

$$\frac{1}{\delta^l}h(g^l(x)) \leq h(x) + \left(\sum_{i=1}^{l} \frac{1}{\delta^i}\right) c_2$$

by the induction on l. Indeed, since $\frac{1}{\delta}h(g^{l+1}(x)) \leq h(g^l(x)) + \frac{c_2}{\delta}$, we have

$$\frac{1}{\delta^{l+1}}h(g^{l+1}(x)) \leq \frac{1}{\delta^l}h(g^l(x)) + \frac{c_2}{\delta^{l+1}} \leq h(x) + \left(\sum_{i=1}^{l+1} \frac{1}{\delta^i}\right) c_2.$$

By putting $c^+ = c_2 \sum_{i=1}^{+\infty} \frac{1}{\delta^i} = \frac{c_2}{\delta - 1}$, we obtain

$$\widehat{h}^+(x) = \limsup_{l \to +\infty} \frac{1}{\delta^l}h(g^l(x)) \leq h(x) + c^+.$$

The estimate for $\widehat{h}^-$ is shown similarly. (Note that it follows from $\delta \geq 2$ that $\delta_- \geq 2$.) $\square$

To see the transformation formulas (4.3), we observe

$$\begin{aligned}\widehat{h}^+ (g(x)) &= \limsup_{l \to +\infty} \frac{1}{\delta^l} h\left(g(g^l(x))\right)\\ &= \delta \limsup_{l \to +\infty} \frac{1}{\delta^{l+1}} h\left(g^{l+1}(x)\right) = \delta \widehat{h}^+(g(x)).\end{aligned}$$

This gives the first transformation formula, and the other formulas are shown similarly. This completes the proof of (1).

(2) We begin by showing $\widehat{h} \gg h$.

Claim 4.3.2. We have

$$\widehat{h}(x) \geq h(x) - \frac{\delta\delta_-}{(\delta-1)(\delta_- - 1)}c$$

for all $x \in \mathbb{A}^n(\overline{K})$, where c is the constant given in (4.1).

Proof. We set $h' = h - \frac{\delta\delta_-}{(\delta-1)(\delta_- -1)}c$. Then we have for all $x \in \mathbb{A}^n(\overline{K})$

$$\frac{1}{\delta}h'(g(x)) + \frac{1}{\delta_-}h'(g^{-1}(x)) \geq \left(1 + \frac{1}{\delta\delta_-}\right) h'(x). \tag{4.5}$$

Then we have

$$\begin{aligned} \frac{1}{\delta^2}h'(g^2(x)) + \frac{1}{\delta\delta_-}h'(x) \quad &\geq \left(1 + \tfrac{1}{\delta\delta_-}\right) \tfrac{1}{\delta}h'(g(x)), \\ \frac{1}{\delta\delta_-}h'(x) + \frac{1}{\delta_-^2}h'(g^{-2}(x)) &\geq \left(1 + \tfrac{1}{\delta\delta_-}\right) \tfrac{1}{\delta_-}h'(g^{-1}(x)). \end{aligned}$$

Adding these two inequalities and using (4.5) again, we obtain

$$\frac{1}{\delta^2}h'(g^2(x)) + \frac{1}{\delta_-^2}h'(g^{-2}(x)) \geq \left(1 + \frac{1}{(\delta\delta_-)^2}\right) h'(x).$$

Inductively, we have

$$\frac{1}{\delta^{2^l}}h'(g^{2^l}(x)) + \frac{1}{\delta_-^{2^l}}h'(g^{-2^l}(x)) \geq \left(1 + \frac{1}{(\delta\delta_-)^{2^l}}\right) h'(x).$$

(Though not necessary for the proof, one can also show

$$\frac{1}{\delta^m}h'(g^m(x)) + \frac{1}{\delta_-^m}h'(g^{-m}(x)) \geq \left(1 + \frac{1}{(\delta\delta_-)^m}\right) h'(x)$$

for every $m \in \mathbb{Z}_{>0}$.) By letting $l \to +\infty$, we obtain

$$\begin{aligned} &\limsup_{l\to+\infty} \frac{1}{\delta^{2^l}}h'(g^{2^l}(x)) + \limsup_{l\to+\infty} \frac{1}{\delta_-^{2^l}}h'(g^{-2^l}(x)) \\ &\geq \limsup_{l\to+\infty} \left(\frac{1}{\delta^{2^l}}h'(g^{2^l}(x)) + \frac{1}{\delta_-^{2^l}}h'(g^{-2^l}(x))\right) \geq h'(x). \end{aligned} \tag{4.6}$$

Since

$$\begin{aligned} \widehat{h}^+(x) &= \limsup_{m\to+\infty} \frac{1}{\delta^m}h(g^m(x)) \\ &= \limsup_{m\to+\infty} \frac{1}{\delta^m}\left(h'(g^m(x)) + \frac{\delta\delta_-}{(\delta-1)(\delta_- -1)}c\right) \geq \limsup_{l\to+\infty} \frac{1}{\delta^{2^l}}h'(g^{2^l}(x)) \end{aligned}$$

and similarly $\widehat{h}^-(x) \geq \limsup_{l\to+\infty} \frac{1}{\delta_-^{2l}} h'(g^{-2^l}(x))$, it follows that the left-hand side of (4.6) is less than or equal to $\widehat{h}(x)$, while the right-hand-side is $h(x) - \frac{\delta\delta_-}{(\delta-1)(\delta_--1)}c$. Thus we get the desired inequality. □

It follows from Claim 4.3.1 and Claim 4.3.2 that $\widehat{h}$ satisfies property (i). Indeed we have

$$h(x) - \frac{\delta\delta_-}{(\delta-1)(\delta_--1)}c \leq \widehat{h}(x) \leq 2h(x) + c^+ + c^-. \tag{4.7}$$

Property (ii) is checked from the transformation formulas (4.3).

(3) We will show that $x \in \mathbb{A}^n(\overline{K})$ is g-periodic if and only if $\widehat{h}(x) = 0$. Suppose $\widehat{h}(x_1) = 0$. Then by (ii) and the non-negativity of $\widehat{h}$, we have $\widehat{h}(g(x_1)) = 0$ and $\widehat{h}(g^{-1}(x_1)) = 0$. Take an extension field L of K such that x_1 is defined over L. Since $\widehat{h} \gg\ll h$, $\widehat{h}$ satisfies the Northcott finiteness property. Thus the set

$$\{g^l(x_1) \mid l \in \mathbb{Z}\} \quad \left(\subseteq \{x \in \mathbb{A}^n(L) \mid \widehat{h}(x) = 0\}\right)$$

is finite, and so x_1 is g-periodic. On the other hand, suppose $\widehat{h}(x_2) =: a > 0$. Then $\frac{1}{\delta}\widehat{h}(g(x_2)) + \frac{1}{\delta_-}\widehat{h}(g^{-1}(x_2)) = \left(1 + \frac{1}{\delta\delta_-}\right)\widehat{h}(x_2) = \left(1 + \frac{1}{\delta\delta_-}\right)a$. Thus we have

$$\widehat{h}(g(x_2)) \geq \frac{1+\delta\delta_-}{\delta+\delta_-}a \quad \text{or} \quad \widehat{h}(g^{-1}(x_2)) \geq \frac{1+\delta\delta_-}{\delta+\delta_-}a.$$

Since $\frac{1+\delta\delta_-}{\delta+\delta_-} > 1$, this shows that the set $\{g^l(x_2) \mid l \in \mathbb{Z}\}$ is not a set of bounded height. Thus x_2 cannot be g-periodic.

Next we will show that $\widehat{h}(x) = 0 \iff \widehat{h}^+(x) = 0 \iff \widehat{h}^+(x) = 0$. Since $\widehat{h}^\pm$ are non-negative and $\widehat{h} = \widehat{h}^+ + \widehat{h}^-$, it is clear that $\widehat{h}(x) = 0$ implies $\widehat{h}^+(x) = \widehat{h}^-(x) = 0$. We will show that $\widehat{h}^+(x) = 0$ implies $\widehat{h}(x) = 0$. A key observation here is again that $\widehat{h}$ satisfies the Northcott finiteness property. Suppose $\widehat{h}^+(x) = 0$. Then

$$\widehat{h}(g^l(x)) = \widehat{h}^+(g^l(x)) + \widehat{h}^-(g^l(x)) = \delta^l\widehat{h}^+(x) + \frac{1}{\delta_-^l}\widehat{h}^-(x) = \frac{1}{\delta_-^l}\widehat{h}^-(x).$$

Let L be a finite extension of K over which x is defined. Then

$$\{g^l(x) \in \mathbb{A}^n(L) \mid l \geq 0\} \quad \subseteq \quad \{y \in \mathbb{A}^n(L) \mid \widehat{h}(y) \leq \widehat{h}^-(x)\}$$

is finite. Hence x is g-periodic. Similarly we see that $\widehat{h}^-(x) = 0$ implies $\widehat{h}(x) = 0$. This completes the proof of (3).

(4) To show the uniqueness property, suppose $\widehat{h}'$ is also a canonical height function for g such that $\lambda := \widehat{h}' - \widehat{h}$ is bounded on $\mathbb{A}^n(\overline{K})$. Set $M := \sup_{x\in\mathbb{A}^n(\overline{K})} |\lambda(x)|$.

Then

$$\left(1+\frac{1}{\delta\delta_-}\right)M = \left(1+\frac{1}{\delta\delta_-}\right)\sup_{x\in\mathbb{A}^n(\overline{K})}|\lambda(x)|$$
$$= \sup_{x\in\mathbb{A}^n(\overline{K})}\left|\frac{1}{\delta}\lambda(g(x)) + \frac{1}{\delta_-}\lambda(g^{-1}(x))\right| \le \left(\frac{1}{\delta}+\frac{1}{\delta_-}\right)M.$$

Since $1+\frac{1}{\delta\delta_-}-\frac{1}{\delta}-\frac{1}{\delta_-} = \frac{(\delta-1)(\delta_--1)}{\delta\delta_-} > 0$, we have $M = 0$, whence $\widehat{h} = \widehat{h}'$. □

Proof of Corollary 4.3. (1) Let $\widehat{h}_g$ be the canonical height function for g constructed in Theorem 4.2(2). We will show that $\widehat{h} := \widehat{h}_g \circ \gamma^{-1}$ satisfies properties (i) and (ii) of Corollary 4.3. By (4.7), we have $\widehat{h}_g(\gamma^{-1}(x)) \le 2h(\gamma^{-1}(x)) + c^+ + c^-$. Theorem 1.3 yields that there is a constant $c_{\gamma^{-1}}$ such that $h(\gamma^{-1}(x)) \le (\deg\gamma^{-1})\, h(x) + c_{\gamma^{-1}}$ for all $x \in \mathbb{A}^n(\overline{K})$. Thus

$$\widehat{h}(x) \le 2(\deg\gamma^{-1})\, h(x) + (2c_{\gamma^{-1}} + c^+ + c^-). \tag{4.8}$$

On the other hand, Theorem 1.3 yields that there is a constant c_γ such that $h(\gamma(x)) \le (\deg\gamma)\, h(x) + c_\gamma$ for all $x \in \mathbb{A}^n(\overline{K})$. Hence

$$h(\gamma^{-1}(x)) \ge (\deg\gamma)^{-1}h(x) - (\deg\gamma)^{-1}c_\gamma.$$

By (4.7) we have $\widehat{h}_g(\gamma^{-1}(x)) \ge h(\gamma^{-1}(x)) - \frac{\delta\delta_-}{(\delta-1)(\delta_--1)}c$, and so

$$\widehat{h}(x) \ge (\deg\gamma)^{-1}h(x) - (\deg\gamma)^{-1}c_\gamma - \frac{\delta\delta_-}{(\delta-1)(\delta_--1)}c. \tag{4.9}$$

Now property (i) follows from (4.8) and (4.9).

Property (ii) follows from

$$\begin{aligned}\widehat{h}(f(x)) + \widehat{h}(f^{-1}(x)) &= \widehat{h}_g(\gamma^{-1}(f(x))) + \widehat{h}_g(\gamma^{-1}(f^{-1}(x)))\\ &= \widehat{h}_g(g(\gamma^{-1}(x))) + \widehat{h}_g(g^{-1}(\gamma^{-1}(x)))\\ &= \left(1+\frac{1}{\delta\delta_-}\right)\widehat{h}_g(\gamma^{-1}(x)) = \left(1+\frac{1}{\delta\delta_-}\right)\widehat{h}(x),\end{aligned}$$

where we used Theorem 4.2(2-ii) in the third equality.

(2) This follows from the proof of Theorem 4.2(4), with f in place of g.

(3) By property (1-ii), we readily see $\widehat{h} = \widehat{h}^+ + \widehat{h}^-$. Also, property (1-ii) gives

$$\widehat{h}\circ f - \frac{1}{\delta_-}\widehat{h} = \delta\left(\widehat{h} - \frac{1}{\delta_-}\widehat{h}\circ f^{-1}\right).$$

This shows the first transformation formula: $\widehat{h}^+(f(x)) = \delta\ \widehat{h}^+(x)$. The other transformation formulas are checked similarly.

(4) To show that $\widehat{h}$ is non-negative, we assume the contrary, so that there exists $x' \in \mathbb{A}^n(\overline{K})$ with $\widehat{h}(x') =: a' < 0$. Then we have

$$\widehat{h}(f(x')) \leq \frac{1+\delta\delta_-}{\delta+\delta_-}a' \quad \text{or} \quad \widehat{h}(f^{-1}(x')) \leq \frac{1+\delta\delta_-}{\delta+\delta_-}a'.$$

This implies that $\widehat{h}$ is not bounded from below. Since $h \gg\ll \widehat{h}$, this is a contradiction.

We have shown that $\widehat{h} \geq 0$. Then $\widehat{h}^+(f^l(x)) + \widehat{h}^-(f^l(x)) = \widehat{h}(f^l(x)) \geq 0$ for any $l \in \mathbb{Z}$ and $x \in \mathbb{A}^n(\overline{K})$. This is equivalent to

$$\widehat{h}^+(x) \geq -\frac{1}{(\delta\delta_-)^l}\widehat{h}^-(x).$$

By letting $l \to +\infty$, we have $\widehat{h}^+(x) \geq 0$. Similarly we have $\widehat{h}^-(x) \geq 0$.

Thus $\widehat{h}, \widehat{h}^\pm$ are all non-negative functions. The proof of Theorem 4.2(3), with f in place of g, then gives the desired equivalences. □

In the remainder of this section, we would like to discuss the condition (4.1) in Theorem 4.2. The next proposition shows that the constant $(1 + \frac{1}{\delta\delta_-})$ in (4.1) cannot be replaced by any larger number.

Proposition 4.4. *Let $g : \mathbb{A}^n \to \mathbb{A}^n$ a polynomial automorphism of degree $\delta \geq 2$ over a number field K. Let δ_- denote the degree of g^{-1}. Then*

$$\liminf_{\substack{x \in \mathbb{A}^n(\overline{K}) \\ h(x)\to\infty}} \frac{\frac{1}{\delta}h(g(x)) + \frac{1}{\delta_-}h(g^{-1}(x))}{h(x)} \leq 1 + \frac{1}{\delta\delta_-}. \tag{4.10}$$

Proof. To derive a contradiction, suppose (4.10) does not hold. Then there are positive numbers $a > 1+\frac{1}{\delta\delta_-}$ and M such that, for any $x \in \mathbb{A}^n(\overline{K})$ with $h(x) \geq M$, one has

$$\frac{1}{\delta}h(g(x)) + \frac{1}{\delta_-}h(g^{-1}(x)) \geq ah(x). \tag{4.11}$$

We take a non g-periodic point $x_0 \in \mathbb{A}^n(\overline{K})$ and fix it. We set $O_g(x_0) = \{g^m(x_0) \mid m \in \mathbb{Z}\}$. Since the set $\{y \in O_g(x_0) \mid h(y) < M\}$ is finite by the Northcott finiteness property, we see together with (4.11) that there is a constant $c \geq 0$ such that, for all $y \in O_g(x_0)$, one has

$$\frac{1}{\delta}h(g(y)) + \frac{1}{\delta_-}h(g^{-1}(y)) \geq ah(y) - c.$$

Noting $a > 1 + \frac{1}{\delta\delta_-} \geq \frac{1}{\delta} + \frac{1}{\delta_-}$, we set $c' := \left(a - \frac{1}{\delta} - \frac{1}{\delta_-}\right)^{-1} c$ and $h' := h - c'$. Then h' satisfies

$$\frac{1}{\delta}h'(g(y)) + \frac{1}{\delta_-}h'(g^{-1}(y)) \geq ah'(y)$$

for all $y \in O_g(x_0)$. As in the proof of Claim 4.3.2, we get

$$\frac{1}{\delta^2} h'(g^2(y)) + \frac{1}{\delta_-^2} h'(g^{-2}(y)) \geq \left(a^2 - \frac{2}{\delta\delta_-}\right) h'(y).$$

We set $a_1 = a^2 - \frac{2}{\delta\delta_-}$. Since

$$\begin{aligned} a_1 - 1 - \frac{1}{(\delta\delta_-)^2} &= a^2 - \frac{2}{\delta\delta_-} - 1 - \frac{1}{(\delta\delta_-)^2} \\ &> \left(1 + \frac{1}{\delta\delta_-}\right)^2 - \frac{2}{\delta\delta_-} - 1 - \frac{1}{(\delta\delta_-)^2} = 0, \end{aligned}$$

we have $a_1 > 1 + \frac{1}{(\delta\delta_-)^2}$. Thus, if we define a sequence $\{a_l\}_{l=0}^{+\infty}$ by $a_0 = a$ and $a_{l+1} = a_l^2 - \frac{2}{(\delta\delta_-)^{2^l}}$, then we get inductively

$$\frac{1}{\delta^{2^l}} h'(g^{2^l}(y)) + \frac{1}{\delta_-^{2^l}} h'(g^{-2^l}(y)) \geq a_l h'(y).$$

On the other hand, it follows from Theorem 1.3 and the argument in Claim 4.3.1 that there is a constant c'' independent of $l \in \mathbb{Z}_{>0}$ such that, for all $y \in O_g(x_0)$,

$$2h'(y) + c'' \geq \frac{1}{\delta^{2^l}} h'(g^{2^l}(y)) + \frac{1}{\delta_-^{2^l}} h'(g^{-2^l}(y)).$$

Thus $2h'(y) + c'' \geq a_l h'(y)$ for all $l \in \mathbb{Z}$ and all $y \in O_g(x_0)$. Since x_0 is a non g-periodic point, the Northcott finiteness property implies that the set $\{h(y) \mid y \in O_g(x_0)\}$ is unbounded. Thus $h'(y) > 0$ for some $y \in O_g(x_0)$. Since $\lim_{l\to+\infty} a_l = +\infty$ by the following Lemma 4.5(1), this is a contradiction. □

Lemma 4.5. *Let $D \geq 4$. Let $\{a_l\}_{l=0}^{+\infty}$ be a sequence defined by $a_0 = a$ and $a_{l+1} = a_l^2 - 2D^{-2^l}$.*

(1) *If $a > 1 + \frac{1}{D}$, then* $\lim_{l\to+\infty} a_l = +\infty$.
(2) *If $a = 1 + \frac{1}{D}$, then* $\lim_{l\to+\infty} a_l = 1$.
(3) *If $1 \leq a < 1 + \frac{1}{D}$, then* $\lim_{l\to+\infty} a_l = 0$.

Proof. We show (1). Set $\varepsilon_l = a_l - 1 - D^{-2^l}$. In particular $\varepsilon_0 = a - 1 - D^{-1} > 0$. Since $\varepsilon_{l+1} = a_{l+1} - 1 - D^{-2^{l+1}} = 2\varepsilon_l(1 + D^{-2^l}) + \varepsilon_l^2$, we get $\varepsilon_{l+1} > 2\varepsilon_l > \cdots > 2^{l+1}\varepsilon_0$. Hence $\lim_{l\to+\infty} \varepsilon_l = +\infty$ and thus $\lim_{l\to+\infty} a_l = +\infty$

We show (2). In this case, we have $a_l = 1 + D^{-2^l}$. Thus $\lim_{l\to+\infty} a_l = 1$.

Finally we show (3). On one hand, we get by induction $a_l \geq 2D^{-2^{l-1}}$ for $l \geq 1$, and in particular $a_l \geq 0$ for $l \geq 1$. On the other hand, we claim for sufficiently large l that $a_l < 1$. Indeed, we assume the contrary and suppose $a_l \geq 1$ for all l. By induction, we get $a_l < 1 + D^{-2^l}$. We set $\lambda_l = 1 + D^{-2^l} - a_l$, and so $0 < \lambda_l \le D^{-2^l}$. Then $a_{l+1} = a_l^2 - 2D^{-2^l} = (1 + D^{-2^l} - \lambda_l)^2 - 2D^{-2^l} = 1 + D^{-2^{l+1}} - 2\lambda_l(1 + D^{-2^l}) + \lambda_l^2$. Hence we get $\lambda_{l+1} = 2\lambda_l(1 + D^{-2^l}) - \lambda_l^2 \ge 2\lambda_l$, which says that $\lim_{l\to+\infty} \lambda_l = +\infty$. This is a contradiction. Hence there is an l_0 with $a_{l_0} < 1$. Since $(0 \le)\ a_{l_0+k} \le a_{l_0}^{2^k}$, we get $\lim_{l\to+\infty} a_l = 0$. □

We denote by L the left-hand side of (4.10). It follows from Theorem 2.3 that, if g is a regular polynomial automorphism of $\mathbb{A}^2$ of degree $\delta \ge 2$, then $\delta = \delta_-$ and $L = 1 + \frac{1}{\delta^2}$. We remark that Marcello [12, Théorème 3.1] showed that, if g is a regular polynomial automorphism of $\mathbb{A}^n$ (this means the set of indeterminacy $Z_{\overline{g}}$ and $Z_{\overline{g^{-1}}}$ are disjoint, cf. [13, Définition 2.2.1]), then $L \ge 1$. It would be interesting to know what polynomial automorphisms g on $\mathbb{A}^n$ satisfy (4.1).

5. The number of points with bounded height in an f-orbit

In this section, we will prove Theorem C. As in §4 we will show Theorem C in a more general setting. The arguments below are inspired by those of Silverman on certain K3 surfaces [15, §3].

Throughout this section, let $f : \mathbb{A}^n \to \mathbb{A}^n$ be a polynomial automorphism over a number field K as in Corollary 4.3. Namely, we assume that there is a polynomial automorphism $\gamma : \mathbb{A}^n \to \mathbb{A}^n$ such that $g := \gamma^{-1} \circ f \circ \gamma$ is a polynomial automorphism satisfying the condition (4.1) in Theorem 4.2. Let $\widehat{h}$ be a canonical height function for f, and $\widehat{h}^{\pm}$ the functions defined in Corollary 4.3(3).

For $x \in \mathbb{A}^n(\overline{K})$, we define the *$f$-orbit* of x by

$$O_f(x) := \{f^l(x) \mid l \in \mathbb{Z}\}.$$

For an f-orbit $O_f(x)$, we define the *canonical height of $O_f(x)$* (with respect to $\widehat{h}$) by

$$\widehat{h}(O_f(x)) = \frac{\log \widehat{h}^+(x)}{\log \delta} + \frac{\log \widehat{h}^-(x)}{\log \delta_-} \qquad \in \mathbb{R} \cup \{-\infty\}.$$

Lemma 5.1.

(1) *The value $\widehat{h}(O_f(x))$ depends only on the orbit $O_f(x)$ and the choice of the height function $\widehat{h}$, and not on the particular choice of the point x in the orbit. Moreover, $\widehat{h}(O_f(x)) = -\infty$ if and only if $O_f(x)$ is a finite orbit.*

(2) *Assume that $O_f(x)$ is an infinite orbit. Then we have*

$$\widehat{h}(O_f(x)) + \epsilon_1 \le \left(\frac{1}{\log \delta} + \frac{1}{\log \delta_-}\right) \min_{y \in O_f(x)} \log \widehat{h}(y) \le \widehat{h}(O_f(x)) + \epsilon_2,$$

where the positive constants ϵ_1 and ϵ_2 are given by

$$\epsilon_1 = \frac{1}{\log \delta} \log\left(1 + \frac{\log \delta}{\log \delta_-}\right) + \frac{1}{\log \delta_-} \log\left(1 + \frac{\log \delta_-}{\log \delta}\right),$$

$$\epsilon_2 = \epsilon_1 + \left(\frac{1}{\log \delta} + \frac{1}{\log \delta_-}\right) \log \max\{\delta, \delta_-\}.$$

Proof. (1) follows from Corollary 4.3(3)(4). To prove (2), set

$$p = 1 + \frac{\log \delta}{\log \delta_-} \quad \text{and} \quad q = 1 + \frac{\log \delta_-}{\log \delta}.$$

Then $p > 1$, $q > 1$, and $\frac{1}{p} + \frac{1}{q} = 1$. By Hölder's inequality, we have

$$\begin{aligned}\widehat{h}(y) &= \widehat{h}^+(y) + \widehat{h}^-(y)\\ &= \frac{1}{p}\left(p^{\frac{1}{p}}\widehat{h}^+(y)^{\frac{1}{p}}\right)^p + \frac{1}{q}\left(q^{\frac{1}{q}}\widehat{h}^-(y)^{\frac{1}{q}}\right)^q \geq p^{\frac{1}{p}}q^{\frac{1}{q}}\widehat{h}^+(y)^{\frac{1}{p}}\widehat{h}^-(y)^{\frac{1}{q}}.\end{aligned}$$

Hence, $\frac{1}{p}\log p + \frac{1}{q}\log q + \frac{1}{p}\log \widehat{h}^+(y) + \frac{1}{q}\log \widehat{h}^-(y) \leq \log \widehat{h}(y)$. Since

$$\frac{1}{p}\log \widehat{h}^+(y) + \frac{1}{q}\log \widehat{h}^-(y) = \left(\frac{1}{\log \delta} + \frac{1}{\log \delta_-}\right)^{-1}\widehat{h}(O_f(x)),$$

we obtain $\widehat{h}(O_f(x)) + \epsilon_1 \leq \left(\frac{1}{\log \delta} + \frac{1}{\log \delta_-}\right)\min_{y \in O_f(x)} \log \widehat{h}(y)$.

On the other hand, we have $\widehat{h}(f^l(x)) = \delta^l \widehat{h}^+(x) + \delta_-^{-l}\widehat{h}^-(x)$ for $l \in \mathbb{Z}$. We set $g(t) = \delta^t \widehat{h}^+(x) + \delta_-^{-t}\widehat{h}^-(x)$ for $t \in \mathbb{R}$, and

$$t_0 := \frac{\log(\widehat{h}^-(x)\log \delta_-) - \log(\widehat{h}^+(x)\log \delta)}{\log \delta + \log \delta_-}.$$

Then one sees that g takes its minimum at t_0, with $g(t_0) = p^{\frac{1}{p}}q^{\frac{1}{q}}\widehat{h}^+(x)^{\frac{1}{p}}\widehat{h}^-(x)^{\frac{1}{q}}$. Consequently as a function of $l \in \mathbb{Z}$, $\widehat{h}(f^l(x))$ takes its minimum at $l = [t_0]$ or $l = [t_0] + 1$, where $[t_0]$ denotes the largest integer less than or equal to t_0. Then we get

$$\begin{aligned}\widehat{h}(f^{[t_0]}(x)) &= \delta^{[t_0]}\widehat{h}^+(x) + \delta_-^{-[t_0]}\widehat{h}^-(x)\\ &= \delta^{-(t_0-[t_0])}\delta^{t_0}\widehat{h}^+(x) + \delta_-^{t_0-[t_0]}\delta_-^{-t_0}\widehat{h}^-(x)\\ &< \max\{\delta, \delta_-\}\left(\delta^{t_0}\widehat{h}^+(x) + \delta_-^{-t_0}\widehat{h}^-(x)\right)\\ &= \max\{\delta, \delta_-\}p^{\frac{1}{p}}q^{\frac{1}{q}}\widehat{h}^+(x)^{\frac{1}{p}}\widehat{h}^-(x)^{\frac{1}{q}}.\end{aligned}$$

Similarly we get

$$\begin{aligned}\widehat{h}(f^{[t_0]+1}(x)) &= \delta^{1+[t_0]-t_0}\delta^{t_0}\widehat{h}^+(x) + \delta^{-(1+[t_0]-t_0)}\delta_-^{-t_0}\widehat{h}^-(x)\\ &< \max\{\delta, \delta_-\}p^{\frac{1}{p}}q^{\frac{1}{q}}\widehat{h}^+(x)^{\frac{1}{p}}\widehat{h}^-(x)^{\frac{1}{q}}.\end{aligned}$$

This shows $\left(\frac{1}{\log \delta} + \frac{1}{\log \delta_-}\right)\min_{y \in O_f(x)} \log \widehat{h}(y) \leq \widehat{h}(O_f(x)) + \epsilon_2$. $\square$

Theorem 5.2. *let $f : \mathbb{A}^n \to \mathbb{A}^n$ be a polynomial automorphism over a number field K as in Corollary 4.3. Let $\widehat{h}$ be a canonical height function for f. For any non f-periodic point $x \in \mathbb{A}^n(\overline{K})$, we define counting functions*

$$N(x,T) := \#\{y \in O_f(x) \mid h(y) \leq T\},$$
$$\widehat{N}(x,T) := \#\{y \in O_f(x) \mid \widehat{h}(y) \leq T\}.$$

We also define a comparison function

$$\Gamma(x,T) := \left(\frac{1}{\log\delta} + \frac{1}{\log\delta_-}\right)\log T - \widehat{h}\left(O_f(x)\right).$$

(1) *We have*

$$\widehat{N}(x,T) = 0 \qquad \textit{if } \Gamma(x,T) < 0,$$
$$\left|\widehat{N}(x,T) - \Gamma(x,T)\right| \leq \frac{\log 2}{\log\delta} + \frac{\log 2}{\log\delta_-} + 1 \quad \textit{if } \Gamma(x,T) \geq 0.$$

(2) *For all infinite orbits $O_f(x)$, we have*

$$N(x,T) = \Gamma(x,T) + O(1) \qquad \textit{as } T \to +\infty,$$

where the $O(1)$ bound depends only on f and the choice of $\widehat{h}$, independent of the orbit $O_f(x)$.

Proof. (1) Suppose $\Gamma(x,T) < 0$. Then for any $y \in O_f(x)$, Lemma 5.1(2) yields that

$$\left(\frac{1}{\log\delta} + \frac{1}{\log\delta_-}\right)\log\widehat{h}(y) \geq \widehat{h}(O_f(x)) + \epsilon_1 > \left(\frac{1}{\log\delta} + \frac{1}{\log\delta_-}\right)\log T.$$

Thus $\widehat{h}(y) > T$ for any $y \in O_f(x)$, which gives the former part of (1).

Let us show the latter part of (1). Since $O_f(x)$ is an infinite orbit, the map $\mathbb{Z} \ni l \mapsto f^l(x) \in \mathbb{A}^n(\overline{K})$ is one-to-one. Then

$$\begin{aligned}\#\{y \in O_f(x) \mid \widehat{h}(y) \leq T\} &= \#\{l \in \mathbb{Z} \mid \widehat{h}(f^l(x)) \leq T\} \\ &= \#\{l \in \mathbb{Z} \mid \delta^l \widehat{h}^+(x) + \delta_-^{-l}\widehat{h}^-(x) \leq T\}.\end{aligned}$$

It follows from Lemma 5.3 below that

$$\begin{aligned}-1 + \frac{\log\frac{T}{2\widehat{h}^+(x)}}{\log\delta} + \frac{\log\frac{T}{2\widehat{h}^-(x)}}{\log\delta_-} &\leq \#\{y \in O_f(x) \mid \widehat{h}(y) \leq T\} \\ &\leq 1 + \frac{\log\frac{T}{\widehat{h}^+(x)}}{\log\delta} + \frac{\log\frac{T}{\widehat{h}^-(x)}}{\log\delta_-},\end{aligned}$$

for $T \geq \widehat{h}^+(x)^{\frac{\log\delta_-}{\log\delta+\log\delta_-}}\widehat{h}^-(x)^{\frac{\log\delta}{\log\delta+\log\delta_-}}$ or equivalently $\left(\frac{1}{\log\delta} + \frac{1}{\log\delta_-}\right)\log T \geq \widehat{h}(O_f(x))$.

On the other hand, we have

$$
\begin{aligned}
-1 &+ \frac{\log \frac{T}{2\widehat{h}^{+}(x)}}{\log \delta} + \frac{\log \frac{T}{2\widehat{h}^{-}(x)}}{\log \delta_{-}} \\
&= -1 - \frac{\log 2}{\log \delta} - \frac{\log 2}{\log \delta_{-}} + \left(\frac{1}{\log \delta} + \frac{1}{\log \delta_{-}}\right) \log T - \widehat{h}(O_f(x)), \\
1 + \frac{\log \frac{T}{\widehat{h}^{+}(x)}}{\log \delta} &+ \frac{\log \frac{T}{\widehat{h}^{-}(x)}}{\log \delta_{-}} = 1 + \left(\frac{1}{\log \delta} + \frac{1}{\log \delta_{-}}\right) \log T - \widehat{h}(O_f(x)).
\end{aligned}
$$

This shows the latter part of (1), up to verifying Lemma 5.3.

(2) Since $h \gg\ll \widehat{h}$ by property (i) of Corollary 4.3, there exist a positive constant a_2 and a constant b_2 such that $\widehat{h} \leq a_2 h + b_2$. Then we have

$$
\begin{aligned}
N(x, T) &= \#\{y \in O_f(x) \mid h(y) \leq T\} \\
&\leq \#\{y \in O_f(x) \mid \widehat{h}(y) \leq a_2 T + b_2\} \\
&\leq \left(\frac{1}{\log \delta} + \frac{1}{\log \delta_{-}}\right) \log(a_2 T + b_2) - \widehat{h}(O_f(x)) + 1 + \frac{\log 2}{\log \delta} + \frac{\log 2}{\log \delta_{-}} \\
&\leq \Gamma(x, T) + O(1) \qquad \text{as } T \to +\infty.
\end{aligned}
$$

Using $a_1 h + b_1 \leq \widehat{h}$ for some positive constant a_1 and constant b_1, we have $N(x, T) \geq \Gamma(x, T) + O(1)$ as T tends to $+\infty$. □

Lemma 5.3. *Let $A, B, T > 0$ be positive numbers. If $T \geq A^{\frac{\log \delta_{-}}{\log \delta + \log \delta_{-}}} B^{\frac{\log \delta}{\log \delta + \log \delta_{-}}}$, then we have*

$$
-1 + \frac{\log \frac{T}{2A}}{\log \delta} + \frac{\log \frac{T}{2B}}{\log \delta_{-}} \leq \#\{l \in \mathbb{Z} \mid \delta^l A + \delta_{-}^{-l} B \leq T\} \leq 1 + \frac{\log \frac{T}{A}}{\log \delta} + \frac{\log \frac{T}{B}}{\log \delta_{-}}.
$$

Proof. If $l \in \mathbb{Z}$ satisfies $\delta^l A + \delta_{-}^{-l} B \leq T$, then $\delta^l A \leq T$ and $\delta_{-}^{-l} B \leq T$. Note that $\frac{\log \frac{B}{T}}{\log \delta_{-}} \leq \frac{\log \frac{T}{A}}{\log \delta}$ is equivalent to $T \geq A^{\frac{\log \delta_{-}}{\log \delta + \log \delta_{-}}} B^{\frac{\log \delta}{\log \delta + \log \delta_{-}}}$. Then, for $T \geq A^{\frac{\log \delta_{-}}{\log \delta + \log \delta_{-}}} B^{\frac{\log \delta}{\log \delta + \log \delta_{-}}}$, we have

$$
\begin{aligned}
\#\{l \in \mathbb{Z} \mid \delta^l A + \delta_{-}^{-l} B \leq T\} &\leq \# \left\{ l \in \mathbb{Z} \;\middle|\; \frac{\log \frac{B}{T}}{\log \delta_{-}} \leq l \leq \frac{\log \frac{T}{A}}{\log \delta} \right\} \\
&\leq 1 + \frac{\log \frac{T}{A}}{\log \delta} + \frac{\log \frac{T}{B}}{\log \delta_{-}}.
\end{aligned}
$$

On the other hand, if $l \in \mathbb{Z}$ satisfies $\delta^l A \leq \frac{T}{2}$ and $\delta_{-}^{-l} B \leq \frac{T}{2}$, then $\delta^l A + \delta_{-}^{-l} B \leq T$. Thus,

$$\#\{l \in \mathbb{Z} \mid \delta^l A + \delta_-^{-l} B \leq T\} \geq \# \left\{ l \in \mathbb{Z} \;\middle|\; \frac{\log \frac{2B}{T}}{\log \delta_-} \leq l \leq \frac{\log \frac{T}{2A}}{\log \delta} \right\}$$
$$\geq -1 + \frac{\log \frac{T}{2A}}{\log \delta} + \frac{\log \frac{T}{2B}}{\log \delta_-}.$$

□

Proof of Theorem C. As we saw in the proof of Theorem A and Corollary B, for any polynomial automorphism $f : \mathbb{A}^2 \to \mathbb{A}^2$ of dynamical degree ≥ 2, there is a polynomial automorphism $\gamma : \mathbb{A}^2 \to \mathbb{A}^2$ such that $g := \gamma^{-1} \circ f \circ \gamma$ satisfies (4.1) in Theorem 4.2. Then, applying Theorem 5.2(2) to f, we obtain Theorem C. □

Acknowledgements. The author deeply thanks Professor Noboru Nakayama for simplifying the proof of (0.3), and the referee for giving many kind and helpful comments.

References

1. Beauville, A.: Surfaces algébriques complexes. Astérisque, 54, Soc. Math. France, Paris, 1978
2. Call, G., Silverman, H.: Canonical heights on varieties with morphisms. Compositio Math. **89**, 163–205 (1993)
3. Denis, L.: Points périodiques des automorphismes affines. J. Reine Angew. Math. **467**, 157–167 (1995)
4. Furter, J.-P.: On the degree of iterates of automorphisms of the affine plane. Manuscripta Math. **98**, 183–193 (1999)
5. Friedland, S., Milnor, J.: Dynamical properties of plane polynomial automorphisms. Ergodic Theory Dynam. Systems **9**, 67–99 (1989)
6. Hubbard, J., Papadopol, P., Veselov, V.: A compactification of Hénon mappings in $\mathbb{C}^2$ as dynamical systems. Acta Math. **184**, 203–270 (2000)
7. Jung, H.: Über ganze birationale Transformationen der Ebene. J. Reine Angew. Math. **184**, 161–174 (1942)
8. Kawaguchi, S.: Canonical heights, invariant currents, and dynamical eigensystems of morphisms for line bundles. To appear in J. Reine Angew. Math.
9. van der Kulk, W.: On polynomial rings in two variables. Nieuw Arch. Wiskunde **1**, 33–41 (1953)
10. Lang, S.: Fundamentals of Diophantine geometry, Springer, 1983
11. Marcello, S.: Sur les propiétés arithmétiques des itérés d'automorphismes réguliers. C.R. Acad. Sci. Paris Sér. I Math. **331**, 11–16 (2000)
12. Marcello, S.: Sur la dynamique arithmétique des automorphismes de l'espace affine. Bull. Soc. Math. France **131**, 229–257 (2003)
13. Sibony, N.: Dynamique des applications rationnelles de $\mathbb{P}^k$. In: Dynamique et géométrie complexes (Lyon, 1997), 97–185, Soc. Math. France, 1999
14. Silverman, J.: The theory of height functions. in Arithmetic geometry, Springer, 151–166 1986
15. Silverman, J.: Rational points on $K3$ surfaces: a new canonical height. Invent. Math. **105**, 347–373 (1991)
16. Silverman, J.: Geometric and arithmetic properties of the Hénon map. Math. Z. **215**, 237–250 (1994)
17. Zhang, S.: Small points and adelic metrics. J. Algebraic Geom. **4**, 281–300 (1995)