Compositio Mathematica **105:** 43–54, 1997.

Geometric height inequalities and the Kodaira–Spencer map

MINHYONG KIM
Department of Mathematics, Columbia University, New York, NY 10027 and Department of Mathematics, University of Arizona, Tucson, AZ 85721

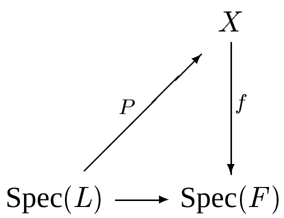
Received 5 April 1995; accepted in final form 30 August 1995

Abstract. We extend a result of Vojta on height inequalities for algebraic points on curves over function fields to include the case of positive characteristic. The main tool used is the Kodaira–Spencer map and destabilizing flags for vector bundles on curves.

Key words: curves, function fields, heights, destabilizing flag.

1.

Let X be a curve defined over a number field F and let K be the canonical class of X. Given an algebraic point



we define the geometric (logarithmic) discriminant of P to be

$$d(P) := \frac{1}{[L\colon F]} \log \mathbf{D}_{L/F},$$

where $\mathbf{D}_{L/F}$ is the ordinary discriminant of L over F. Denote by $h(P)$ the (logarithmic) geometric height with respect to the canonical class. Vojta has made the following conjecture bounding the height of algebraic points on X in terms of the discriminant:

For any $\epsilon > 0$, *there is a constant* $C = C(X, \epsilon)$ *such that for all algebraic points* P, *we have*

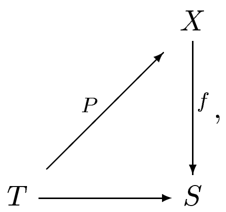
$$h(P) \leqslant (1+\epsilon)d(P) + C.$$

We see that if P ranges only over F-rational points, the right-hand side is bounded so Faltings' Theorem (Mordell conjecture) is implied by the above inequality.

It is however, a much stronger statement with a remarkable set of Diophantine consequences, as surveyed in [2]. From the point of view of giving bounds for the height of points rational over number fields, it makes explicit the dependence of such a bound on the field, by isolating out the purely geometric quantity $C(X, \epsilon)$.

We are interested here in examining the geometric analogue of this inequality, specifically, extending to positive characteristic a result of Vojta, which we go on to describe. So let X be a smooth projective surface over the perfect field k. We are interested in a situation where X admits a map $f: X \to S$ to a smooth projective curve S, also defined over k, with function field F in such a way that the fibers of f are geometrically connected curves and the generic fiber X_F is smooth, of genus $g \geqslant 2$. (So our X is actually the analog of the minimal regular model, suitably compactified at infinity via Arakelov theory, of the curve defined over the number field discussed above, rather than the curve itself.)

Given a diagram



where T is a smooth projective curve mapping to S, we can view P as an algebraic point of X_F. The *canonical height* of P is defined by the formula

$$h(P) := \frac{1}{[T\colon S]} \deg P^*\omega,$$

where $\omega = \omega_X := K_X \otimes f^* K_S^{-1}$ denotes the relative dualizing sheaf for $X \to S$. This is a representative for the class of height functions on $X_F(\bar{F})$ associated to the canonical sheaf K_{X_F}. Define the relative discriminant of T to be

$$d(T) := \frac{2g(T) - 2}{[T\colon S]},$$

where $g(T)$ is the genus of T. The goal of this paper is to extend the following estimate for the height, which is due to Vojta, to positive characteristics:

Assume that k is of characteristic 0. Then for any $\epsilon > 0$ there is a constant $C = C(X, \epsilon)$ such that

$$h(P) \leqslant (2 + \epsilon) d(P) + C. \qquad (*)$$

We list just a few other inequalities of this sort:

$$h(P) \leqslant 2(2g-1)(2g+3)(d(P)+s), \qquad \text{(Szpiro)}$$

$$h(P) \leqslant 2(2g-1)^2(d(p)+s), \qquad \text{(Esnault–Viehweg)}$$

$$h(P) \leqslant (2g-1)^2(d(p)+s), \qquad \text{(Shepherd–Barron)}$$

$$h(P) \leqslant (2g-1)d(p)+O(1) \qquad (g>2), \qquad \text{(Moriwaki)}$$

$$h(P) \leqslant (2g-1)(d(p)+3s)-\omega^2, \qquad \text{(Sheng-Li Tan)}$$

where s denotes the number of singular fibers of f. Among these, the ones of Szpiro, Moriwaki, and Shepherd–Barron work in positive characteristic. However, various complications arise due to the existence of non-separable morphisms, which we now describe in brief. Note that in the complex case, inequality (*) above is trivially true in case $X \to S$ is a constant family, that is, $X = C \times S \xrightarrow{p_2} S$. The relative dualizing sheaf in this case is just $p_1^*\Omega_C$ and a strong version of the inequality above, with a 1 instead of a $2+\epsilon$ and constant zero follows easily from the Hurwitz formula. But we can as easily check that constant families are exactly the troublesome ones in positive characteristic: so let S and C be smooth projective curves over $\mathbf{F}_p$. We can construct E, a smooth projective curve equipped with non-trivial maps $f\colon E \to S$ and $g : E \to C$ in a manner such that f is separable of degree d_1 and g is of degree d_2. Consider the twists $g \circ F^n : E \to C$ with powers of the Frobenius morphism which give rise to algebraic points $P_n := (g \circ F^n) \times f\colon E \to C \times S$. Then it is clearly seen that the discriminant $d(P_n)$ remains constant and equal to $(2g(E)-2)/d_1$, while $h(P_n) = (d_2/d_1)p^n(2g(C)-2)$.

However, the actual situation is even more complicated than this as can be seen by the following construction due to Felipe Voloch:

By the Kodaira–Parshin construction gives us a diagram [4]

$$\begin{array}{ccc} X & \xrightarrow{h} & Y := S \times S \\ {\scriptstyle f}\downarrow & & \downarrow \\ T & \longrightarrow & S \end{array} ,$$

where $T \to S$ is finite unramified, f is smooth and h is ramified only along the diagonal. Now, we have the algebraic points P_n of Y constructed previously (with $C = E = S$), which are now in fact rational points. Note that the image of all the P_n, $n > 0$, which we denote by the same letters, are transverse to the diagonal, so that the inverse image in X, $E_n = q^{-1}(P_n)$ is smooth. Let g' denote the genus of the generic fiber of X and let d denote the degree of the map $q\colon X \to Y$. Then,

$$\begin{aligned}\langle E_n.\omega_X\rangle &= \frac{(2g'-2)}{(d(2g-2))}\langle E_n.q^*\omega_Y\rangle + O(\langle E_n.\omega_X\rangle^{1/2})\\ &= \frac{(2g'-2)}{(2g-2)}\langle (P_n).\omega_Y\rangle + O(\langle E_n.\omega_X\rangle^{1/2})\\ &= (2g'-2)p^n + O(\langle E_n.\omega_X\rangle^{1/2}) = (2g'-2)p^n + O(p^{n/2}).\end{aligned}$$

Here, the estimate of the remainder term uses Neron's theorem on heights with respect to divisors algebraically equivalent to zero [5] (Thm. 2.11). On the other hand, the geometric discriminant can be computed using the Hurwitz formula:

$$2g(E_n) - 2 = d(2g(S) - 2) + \deg(R_n),$$

where R_n is the ramification divisor of E_n over P_n. But by the transversality mentioned above, deg R_n is equal to d times the intersection number $\langle P_n.\Delta\rangle$. This is just the number of $\mathbf{F}_{p^n}$ points of S, and hence, is equal to $p^n + O(p^{n/2})$ by the Riemann hypothesis for S. That is, $d(E_n) = p^n + O(p^{n/2})$. When considering algebraic points, this example shows that just having a non-trivial map to a constant curve, that is, a non-trivial F/k-trace of the Jacobian, can cause sections of large height to 'lift'.

In this paper, we prove two inequalities, which illustrate in a transparent manner the dependence of inequalities of this sort on the Kodaira–Spencer map. Recall that the Kodaira–Spencer map is constructed on any open set $U \subset S$ over which f is smooth from the exact sequence

$$0 \to f^*\Omega^1_U \to \Omega^1_{X_U} \to \Omega^1_{X_U/U} \to 0,$$

by taking the coboundary map $KS\colon f_*(\Omega^1_{X_U/U}) \to \Omega^1_U \otimes R^1 f_*(\mathcal{O}_{X_U})$.

THEOREM 1. *Suppose the Kodaira–Spencer map of the* $f : X \to S$ *(defined on some open subset of* S*) is non-zero. Then, when* $g \geqslant 3$,

$$h(P) \leqslant (2g-2)d(p) + O(h(P)^{1/2}).$$

For $g = 2$,

$$h(P) \leqslant (2+\epsilon)d(P) + C(X,\epsilon).$$

Notice that in view of the example above, the first inequality is the best possible one that holds in general.

THEOREM 2. *Suppose the Kodaira–Spencer map of* X/S *is an isomorphism on some open subset of* S*. Then Vojta's inequality holds*:

$$h(P) \leqslant (2+\epsilon)d(P) + C(X,\epsilon).$$

2.

Note that the inequalities of Theorems 1 and 2 are invariant with respect to a change of domain $P: T \to T' \to X$ so we may as well assume that the map $T \to X$ is generically 1-1. In fact, both the height and the discriminant depend only on the image $P(T)$ inside X. That is,

$$h(P) = \frac{1}{[K(P(T)) : F]} \langle P(T).\omega \rangle,$$

where $K(P(T))$ is the residue field of X at $P(T)$, and

$$d(P) = \frac{1}{[K(P(T)) : F]} (2g(P(T)) - 2),$$

where the genus refers to the geometric genus.

Recall also that we can associate to any line bundle L_F on X_F a height function h_{L_F} on $X_F(\bar{F})$ by choosing a regular model $f: X \to S$ as above and a model L (an extension to X) of L_F and defining

$$h(P_F) := \deg P^* L / [T : S],$$

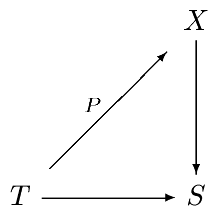
where $P: T \to X$ denotes the obvious extension of the algebraic point P_F: Spec$(F(T)) \to X_F$.

For the convenience of the reader, we recall the proof of the well-known fact that changing X or L changes the function h_{L_F} by a quantity bounded on $X_X(\bar{F})$. It suffices to examine the intersection number between L and an irreducible horizontal divisor H on X. But given models L_1 and L_2, we get $L_1(f^*(-D)) \subset L_2 \subset L_1(f^*(D))$ for some divisor D on S. By intersecting with H and dividing by the degree of H over S, we get the up-to-$O(1)$ independence of L. To examine the dependence on X, in suffices to check what happens when we blow-up a point $r: X' \to X$. Given the horizontal divisor H, we know that $\langle r^* L.r^* H \rangle = \langle L.H \rangle$. On the other hand, the closure inside X' of H_F, is the strict transform H' of H. We have $r^* H = H' + mE$ where E is the exceptional divisor and m is the multiplicity of H at the blown-up point. Thus $\langle r^* L.H' \rangle = \langle r^* L.r^* H \rangle = \langle L.H \rangle$.

In the sequel, we will have occasion to change the model X and several line bundles by finite operations that do not affect the generic fiber. The preceding argument says that this does not affect height inequalities up to a bounded quantity.

On the other hand, for the estimates in the theorems, the following lemma allows us to worry only about those P which are separable over S:

LEMMA 1. *In the diagram*



suppose $T\to X$ is generically unramified but $T\to S$ is inseparable. Then $h(P)\leqslant d(P)+C$, where the constant depends only on X.

Proof. Recall the definition of $\omega := K_X\otimes f^*K_S^{-1}$. We have the exact sequence of sheaves

$$0\to f^*\Omega^1_S\to\Omega^1_X\to\Omega^1_{X/S}\to 0.$$

After a suitable blow-up $r\colon X'\to X$ (which changes all the heights by $O(1)$) we can get the quotient Q of $r^*\Omega^1_X$ by the saturation L of $r^*f^*\Omega^1_S$ to be locally free [3] (Lemma 2). Now, the fact that T is inseparable over S causes the map

$$L\,|\,T\to\Omega^1_T$$

induced by $r^*\Omega^1_X\,|\,T\simeq\Omega^1_X\,|\,T\to\Omega^1_T$ to be zero, since L is generically the same as $f^*\Omega^1_S$. Hence we get a non-zero map

$$Q\,|\,T\to\Omega^1_T.$$

Thus, we get $c_1(Q\,|\,T)\leqslant 2g_T-2$, hence,

$$\langle K_X.T\rangle-\langle L.T\rangle\leqslant 2g_T-2.$$

But $L\subset r^*f^*(\Omega^1_S(D))$ for some divisor D on S, so that $\langle\omega.T\rangle\leqslant 2g_T-2+\deg(D)[T:S]$. □

3.

In this section, we repeat some of the arguments from [6] in our setting for the convenience of the reader and to make the modifications necessary for the positive characteristic case.

We use the convention of indicating with a subscript F an object restricted to the generic fiber: e.g., $(\Omega^1_X)_F := \Omega^1_X|X_F$ and $\omega_F := \omega|X_F$. Then we have an exact sequence

$$0\to\mathcal{O}_{X_F}\to(\Omega^1_X)_F\to\omega_F\to 0,$$

so that $\deg(\Omega^1_X)_F=\deg\omega_F=2g-2$. We wish to apply the Riemann–Roch theorem to the sheaves

$$\mathcal{E}_n := \mathrm{Sym}^{(2+\epsilon)n}(\Omega^1_X)_F\otimes\omega_F^{-n}.$$

Here, as in the following, we assume that ϵ is rational and that n is large enough for all the expressions to make sense. We get

$$\deg\mathcal{E}_n=[(2+\epsilon)n((2+\epsilon)n+1)/2-((2+\epsilon)n+1)n](2g-2)\geqslant\epsilon n^2$$

and rank $\mathcal{E}_n=(2+\epsilon)n+1$ so that

$$H^0(\mathcal{E}_n)\geqslant\epsilon n^2-((2+\epsilon)n+1)(g-1).$$

Thus, $\mathcal{E}_n$ has a non-zero section for large n. But this is just an injection $\omega_F^n \to \mathrm{Sym}^{(2+\epsilon)n}(\Omega_X^1)_F$ which thus extends to an injection

$$\omega^n \to \mathrm{Sym}^{(2+\epsilon)n}(\Omega_X^1) \otimes f^*M,$$

for some invertible sheaf M on S. Now, given any algebraic point

$$\begin{array}{ccc} & & X \\ & \overset{P}{\nearrow} & \downarrow{\scriptstyle f} \\ T & \longrightarrow & S \end{array}$$

we get a sequence of maps

$$P^*\omega^n \to P^*\mathrm{Sym}^{(2+\epsilon)n}(\Omega_X^1) \otimes P^*f^*M \to \Omega_T^{(2+\epsilon)n} \otimes P^*f^*M.$$

If the composed map is non-zero, we see, by taking degrees and dividing by $n[T\colon S]$, that

$$h(P) \leqslant (2+\epsilon)d(P) + \deg(M)/n$$

and the point satisfies the desired inequality.

If the composed map is zero, we say that $P(T) \subset X$ is a *degenerate* curve.

To analyze the degenerate curves, we use the projective bundle

$$\mathbf{P}(\Omega_X^1) := \mathrm{Proj}_X\left(\bigoplus_{i=0}^{\infty} \mathrm{Sym}^i \Omega_X\right).$$

Denoting by $p\colon \mathbf{P}(\Omega_X^1) \to X$ the projection, recall that we have a canonical isomorphism

$$\mathrm{Sym}^n \Omega_X^1 \simeq p_*\mathcal{O}(n),$$

from which we get the surjection $p^*\mathrm{Sym}^n\Omega_X^1 \to \mathcal{O}(n) \to 0$.

Now, since $P: T \to X$ is a generically 1-1 map from a smooth curve T to X, the surjection $P^*\Omega_X^1 \to \Omega_T$ to some invertible subsheaf of the sheaf of differentials of T induces a lifting $t_P : T \to \mathbf{P}(\Omega_X^1)$ of P. This lifting is such that we get a commutative diagram

$$\begin{array}{ccc} t_P^* p^* \Omega_X^1 & \longrightarrow & t_P^* \mathcal{O}(1) \\ {\scriptstyle \simeq}\downarrow & & \downarrow \\ P^*\Omega_X^1 & \longrightarrow & \Omega_T^1 \end{array},$$

so that in the situation above, for the subsheaf $\omega^n \otimes f^*M^{-1}$ of $\mathrm{Sym}^{(2+\epsilon)n}\Omega^1_X$, we get a commutative diagram

$$\begin{array}{ccccc} t_P^*p^*(\omega^n \otimes f^*M^{-1}) & \longrightarrow & t_P^*p^*\,\mathrm{Sym}^{(2+\epsilon)}\Omega^1_X & \longrightarrow & t_P^*\mathcal{O}((2+\epsilon)n) \\ \downarrow\simeq & & \downarrow\simeq & & \downarrow \\ P^*(\omega^n \otimes f^*M^{-1}) & \longrightarrow & P^*\mathrm{Sym}^{(2+\epsilon)}\Omega^1_X & \longrightarrow & \Omega_T^{(2+\epsilon)n} \end{array}.$$

Denoting by s the (non-zero) section of $p^*(\omega^{-n} \otimes f^*M) \otimes \mathcal{O}((2+\epsilon)n)$ given by the composed map

$$p^*(\omega^n \otimes f^*M^{-1}) \to p^*\mathrm{Sym}^{(2+\epsilon)}\Omega^1_X \to \mathcal{O}((2+\epsilon)n),$$

we see that $P(T)$ is degenerate if and only if $t_P(T) \subset \mathrm{div}(s)$.

Thus, we get the following result of Vojta: *Fix a rational number $\epsilon > 0$. Then there exists a constant C and a finite collection of irreducible reduced divisors E_i in $\mathbf{P}(\Omega^1_X)$ such that for all algebraic points $P: T \to X$ such that $t_P(T)$ is not in any of the E_i's*

$$h(P) \leqslant (2+\epsilon)d(P) + C.$$

Now, let $Y' \subset \mathbf{P}(\Omega^1_X)$ be one of the divisors above, $g': Y' \to X$ the map induced by $p: \mathbf{P}(\Omega^1_X) \to X$, and suppose $t_P(T) \subset Y'$. If Y' is not dominant over X, then the $P(T)$ must be equal to the image curve $g'(Y')$ and these form a finite set. So we assume that $g': Y' \to X$ is dominant map of surfaces. Pullback induces a diagram

$$\begin{array}{ccc} g'^*Y' \subset g'^*\mathbf{P}(\Omega^1_X) = \mathbf{P}(g'^*\Omega^1_X) & & \\ \searrow \quad \downarrow & & \\ Y' & & \end{array}$$

except that now there is a section $s: Y' \to g'^*Y'$. Also, $t_P: T \to Y' \subset \mathbf{P}(\Omega^1_X)$ lifts to $h: T \to g'^*\mathbf{P}(\Omega^1_X)$ induced by

$$t_P^* g'^*\Omega^1_X = P^*\Omega^1_X \to \Omega^1_T.$$

One checks readily that $h = s \circ t_P$. On Y', pulling back the canonical exact sequence on $\mathbf{P}(g'^*\Omega^1_X)$ gives an exact sequence

$$0 \to Q \to g'^*\Omega^1_X \to s^*\mathcal{O}(1)$$

in particular, a subsheaf $Q \subset g'^*\Omega^1_X$ such that $t_P^*Q \to t_P^*g'^*\Omega^1_X \to \Omega^1_T$, or equivalently,

$$Q \to g'^*\Omega^1_X \to t_{P*}\Omega^1_T$$

is the zero map. All the above data, including the degenerate curves, lift to a desingularization of Y' which we denote by Y. Denote the map down to X by g.

4. Proof of theorems

For any $Y \xrightarrow{g} X$ as constructed above, we can use Stein factorization to obtain a diagram

$$\begin{array}{ccc} Y & \longrightarrow & X \\ \downarrow & & \downarrow \\ B & \longrightarrow & S \end{array}$$

such that Y/B has connected fibers. Let E be the function field of B so that we have again the natural notation Y_E etc. We get a map $Y \to X_B$, and hence, a map from Y to the regular minimal model X' of X_B. From the argument in the preceding section, we get the canonical exact sequence

$$0 \to Q \to g^*\Omega_X^1 \to M \to 0,$$

such that for each degenerate curve $C \subset X$ which lifts to $\tilde{C} \subset Y$ as described above, we get a non-zero map $M \to \Omega_{\tilde{C}}$. Therefore, $M.\tilde{C} \leqslant 2g_{\tilde{C}} - 2$. Now, suppose

$$\deg(2+\epsilon)M_E \geqslant \deg g^*\omega_E$$

(as **Q**-divisors). Then, by [6] Thm. 2.11, we get

$$h_{g^*\omega_E}(\tilde{C}) \leqslant (2+\epsilon)h_{M_E}(\tilde{C}) + O(h^{1/2}_{g^*\omega_E}) \leqslant (2+\epsilon)d(\tilde{C}) + O(h^{1/2}_{g^*\omega_E}).$$

(Here we abuse notation a bit and denote by $\tilde{C}$ also the point corresponding to it.) But it is readily seen that the original height $h(C)$ on X we are interested in is equal to the height $h_{g^*\omega_E}(\tilde{C})$, so

$$h(C) \leqslant (2+\epsilon)d(\tilde{C}) + O(h^{1/2}_{g^*\omega_E}).$$

Since we could have argued above with $\frac{1}{2}\epsilon > 0$ rather than $\epsilon > 0$, we get therefore a constant C depending only on $\frac{1}{2}\epsilon$ (and thus, on ϵ) and X (as we run over all the Y's) such that

$$h(C) \leqslant (2+\epsilon)d(\tilde{C}) + C.$$

Therefore, we may *assume that* $\deg((2+\epsilon)M_E) < \deg(g^*\omega_E)$.

CLAIM. Put $A := Q_E$ and $B := M_E$. If $\deg(2+\epsilon)M_E < \deg g^*\omega_E$, then the exact sequence

$$0 \to A \to g^*\Omega^1_{X E} \to B \to 0$$

is a destabilizing flag for $g^*\Omega^1_{X E}$, i.e., $A - B$ is torsion or has positive degree.

Proof of Claim. Note that we still have a map $g^*\omega_E^n \to g^* \operatorname{Sym}^{(2+\epsilon)n}(\Omega^1_X)_E$ which is zero when composed with the quotient map to $B^{(2+\epsilon)n}$, by the degree assumption. Therefore, by considering the filtration on $g^* \operatorname{Sym}^{(2+\epsilon)n}(\Omega^1_X)_E$ induced by the exact sequence, we get a non-zero map

$$g^*\omega_E^n \to B^{(2+\epsilon)n-i} \otimes A^i = B^{(2+\epsilon)_n} \otimes (A \otimes B^*)^i$$

for some positive i. By the degree assumption again,

$$B^{(2+\epsilon)n} \otimes \omega_E^{-n} \hookrightarrow \mathcal{O}_{Y_E}$$

for some n (which we may increase by some quantity independent of all the $\tilde{C}$'s again). So $(A \otimes B^*)^i$ has a section for some i, proving the claim.

Recall the fact that the destabilizing flag is unique unless $g^*\Omega^1_{X E} = A \oplus B$ where $A - B$ is torsion: Suppose

$$0 \to A' \to g^*\Omega^1_{X E} \to B' \to 0$$

is a different one. Then

$$A' \hookrightarrow B$$

and

$$A \hookrightarrow B',$$

so $B = A'(D)$ and $B' = A(D)$ for some effective D. Then we get

$$A - B = B' - A' - 2D = -(A' - B') - 2D.$$

For both $A - B$ and $A' - B'$ to be positive or torsion, we need $D = 0$ and hence $B' = A, B = A'$.

But if $(g^*\Omega^1_X)_E = A \oplus B$ where $A - B$ is torsion, then

$$\deg B = \tfrac{1}{2} \deg(g^*\Omega^1_X)_E = \deg(g^*\omega)_E,$$

contradicting the italicized assumption above. Thus we may assume that the destabilizing flag is unique. Since the degree assumption continues to hold after base change to any another curve mapping to Y_E, the destabilizing flag will be unique after any such base-change. So we can use flat descent to conclude that the destabilizing flag is defined over X_E. That is, if denote by h the map from the minimal regular model X' of X_E to X, we get a saturated subsheaf $L \hookrightarrow h^*\Omega^1_X \to \Omega^1_{X'}$

which generically base changes to A. Then one checks immediately that any degenerate curve $C \subset X$ whose tangent map factors through Y' lifts to a *solution* for the Pfaffian equation $L \hookrightarrow \Omega^1_{X'}$ in the sense of [1], [2], that is, the composed map

$$L \hookrightarrow \Omega^1_{X'} \to \Omega_C$$

is zero.

As mentioned previously, since the inequalities to be proved as well as the hypotheses of the theorems are invariant under base-change and choice of models, we can argue directly with X' rather than X, and thus reduce to the following situation:

There exists a finite set of Pfaffian equations $L \subset \Omega^1_X$ such that each degenerate curve is a solution for one of these equations.

Proof of Theorem 1. Fix such an L. There exists a blow-up $r: X' \to X$ such that the quotient of $r^*\Omega_X$ by the saturation of r^*L is locally free. Call that quotient G for $P: T \to X$ degenerate, we see that $P^*G \subset \Omega^1_T$. Denoting by h_G the height with respect to G, (restricted to the generic fiber) this implies that $h_G(P) \leqslant d(P) + O(1)$.

We now study the degree of G_F. Clearly, $\deg G_F = \deg(\Omega^1_X)_F - \deg L_F = 2g-2-\deg L_F$. But $\deg L_F < 2g-2$. For otherwise, that is, if $\deg L_F \geqslant 2g-2 > 0$ then L_F would not be contained in $f^*\Omega^1_F$ (which is trivial), whence $L_F \subset (\Omega^1_X)_F$ would map isomorphically to $\Omega_{X_F/F}$, being a non-zero map of a line bundle on curve to another line bundle of not greater degree. That is, after adjusting by a constant if necessary, we would have a splitting of

$$0 \to f^*\Omega^1_F \to (\Omega^1_X)_F \to \Omega_{X_F/F} \to 0,$$

whence the Kodaira–Spencer map would be zero. So $\deg G_F > 0$ and $G_F^{-(2g-2)} \otimes K_X^{\deg G_F}$ is of degree 0. Thus, by Neron's theorem [6] (Thm 2.11), we get

$$\begin{aligned} \deg(G_F)h_K(P) &\leqslant (2g-2)h_G(P) + O(h_K^{1/2}(p)) \\ &\leqslant (2g-2)d(P) + O(h_K^{1/2}(p)) \end{aligned}$$

and the theorem follows. Note that $2g-2 > 2+\epsilon$ for small $\epsilon > 0$ iff $g > 2$, whence we have the division into the two cases in the statement of the theorem. □

Proof of Theorem 2. As in the preceding proof, consider the degree of G_F. If $\deg L_F \leqslant g-1$ then $\deg G_F \geqslant g-1$ so that $h_K \leqslant 2h_G + O(h_K^{1/2}) \leqslant 2d(P) + O(h_K^{1/2})$ and we are done. Otherwise, $\deg L_F \geqslant g$ so that L_F has a section. Now consider the inclusion $f_*L_F \subset f_*(\Omega^1_X)_F$ and the exact sequence

$$0 \to \Omega^1_F \to f_*(\Omega^1_X)_F \to f_*\Omega^1_{X_F/F} \to \Omega^1_F \otimes R^1 f_*\mathcal{O}_{X_F}$$

where the last arrow is the Kodaira–Spencer map. By Lemma 1 we may assume that $L_F \cap f^*\Omega_F = 0$ (since otherwise, all degenerate T for such L would be inseparable over S) so that $f_*L_F \cap \Omega_F = 0$. The preceding remarks imply that f_*L_F is of rank at least one and gives rise to an element of $f_*(\Omega^1_X)_F$ not contained in Ω_F. Thus, we get a non-zero element contained in the kernel of the Kodaira–Spencer map. □

References

1. Jouanolou, J. P.: Hypersurface solutions d'une equation de Pfaff analytique, *Math. Ann.* 232 (1978) 239–245.
2. Kim, M.: Pfaffian equations and the Cartier operator, *Compositio Mathematica*, to be published.
3. Lang, S.: Old and new conjectures in Diophantine inequalities, *Bull. Amer. Math. Soc.* 23 (1990) 37–75.
4. Moriwaki, A.: Height inequality of nonisotrivial curves over function fields, *J. Alg. Geom.* 3 (1994) 249–263.
5. Martin-Deschamps, M.: La construction de Kodaira–Parshin, *Astérisque* 127 (1985) 261–272.
6. Serre, J.-P.: *Lectures on the Mordell–Weil Theorem*, Vieweg, Braunschweig (1990).
7. Vojta, P.: On algebraic points on curves, *Comp. Math.* 78 (1991) 29–36.