数理解析研究所講究録 1451 巻 2005 年 72-111

AN ATTEMPT TOWARD DIOPHANTINE ANALOGUE OF RAMIFICATION COUNTING IN NEVANLINNA THEORY : TRUNCATED COUNTING FUNCTION IN SCHMIDT'S SUBSPACE THEOREM (PRELIMINARY VERSION)

RYOICHI KOBAYASHI

Graduate School of Mathematics, Nagoya University

ABSTRACT. We establish a new framework of Diophantine geometry which introduces truncated counting function to Schmidt Subspace Theorem. This is a Diophantine analogue of the ramification counting function in Nevanlinna theory. This framework canonically splits the Diophantine inequality in the Parametric Subspace Theorem into Archimedean and non-Archimedean parts. Using this framewark, we will propose some conjectures on the effective version of the Roth theorem with truncated counting function.

0. Introduction.

In [V, Chapt.6], Vojta described systematically the similarities between the proofs of the Cartan-Ahlfors-Weyl Theorem[1] and the Schmidt Subspace Theorem[2].

In this article, we push this direction further. As in Vojta [V], our strategy is to bridge Nevanlinna theory and Diophantine approximation. What is novel in this article is to bridge the Nevanlinna-Cartan theory on Wronskian and the theory of successive minima in geometry of numbers by means of establishing a Diophantine analogue of the truncated counting functions in Nevanlinna-Cartan theory ([N],[C]). We establish this analogue via a new Diophantine analogue of Nevanlinna's lemma on logarithmic derivative.

It was Vojta who first formulated and proved a Diophantine analogue of Nevanlinna's lemma on logarithmic derivative (see [V, Theorem 6.4.3] and [V, Theorem 6.6.1]). In this article, we introduce a new geometric framework in Diophantine approximation and prove the higher jet version of [V, Theorem 6.4.3] in our framework.

Let us fix a number field k. Let a finite set of linear forms in general position be given and D the linear divisor defined by these linear forms. Then, for each point $x \in \mathbb{P}^n(k)\backslash D$, we can canonically associate the finite set S_x^n of non-Archimedean places of k by selecting those places over which the Zariski closures of x and some component of D over the ring of integers $\mathcal{O}_k$ intersect with multiplicity $\geq n$. Our new view point is to combine the association $x \mapsto S_x^n$ with Bombieri-Vaaler's theory

---

[1] This is on approximation to hyperplanes in $\mathbb{P}^n(\mathbb{C})$ by holomorphic curves.

[2] This is on approximation to hyperplanes in $\mathbb{P}^n(k)$ by rational points.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX on geometry of numbers [B-V], by which we are able to establish a new Diophantine analogue of Nevanlinna's lemma on logarithmic derivative. This leads to a strong version of Schmidt's Subspace Theorem with the truncated counting function, i.e., a strict Diophantine analogue of the Nevanlinna-Cartan theory.

The proximity function $m(x, F_i)$ (resp. the counting function $N(x, F_i)$) measures the Archimedean (resp. non-Archimedean) approximation of x to the hyperplane $F_i = 0$. By truncating $N(x, F_i)$ at level n, we get the truncated counting function $N^n(x, F_i)$. More generally, given a finite set S of places including all Archimedean ones, the S-proximity function $m_S(x, F_i)$ measures the approximation relative to the places in S and the S-counting function $N_S(x, F_i)$ does the same relative to the places outside of S. Let's return to the original situation. By truncating $N(x, F_i)$ at level n, we get the truncated counting function $N_n(x, F_i)$ and we define the **residual counting function**

$$N^n(x, F_i) := N(x, F_i) - N_n(x, F_i) \ .$$

This counts only intersections having the multiplicity ($= m$) not smaller than n with weight $m - n$. The absolute logarithmic height function $\mathrm{ht}(x)$ measures the total arithmetic complexity of $x \in \mathbb{P}^n(k)$.

The main result of this article is formulated as follows.

Main Theorem (Theorem 4.3). *Let $F = \{F_i\}_{i=0}^N$ be a set of linear forms in $\mathbb{P}^n(k)$ in general position. Let $\varepsilon > 0$. Then there exists a finite union of linear subspaces $E(F, \varepsilon)$ and a constant $C(F, \varepsilon)$ such that for all $x \in \mathbb{P}^n(k) \backslash E(F, \varepsilon)$ the approximation inequality*

$$\sum_{i=0}^{N} m(x, F_i) + \sum_{i=0}^{N} N^n(x, F_i) \leq (n + 1 + \varepsilon)\, \mathrm{ht}(x) + C(F, \varepsilon)$$

holds.

The presence of the residual counting function in the left hand side strengthens the Schmidt Subspace Theorem. We hope that this will be useful in the attempt toward the effective version Schmidt's Subspace Theorem.

The plan of this article is as follows. Because our method is based on the analogy between Diophantine approximation and Nevanlinna theory, we included a brief introduction to Nevanlinna theory in §1 and §3. We then establish their Diophantine analogue in §2 and in §4.

In the course of the proof of the Main Theorem, we establish a new framework in Diophantine geometry. Let k be any number field and S any fixed finite set of places of k containing all Archimedean ones. The basic Diophantine functions (i.e., the proximity and counting function) are defined with respect to the fixed S and the set S is fixed in the whole story. In our new framework, we introduce the varying S and generalize Vojta's Theorem [V, Theorem 6.4.3][3] in this setting. The varying S means the following. We consider $\mathbb{P}^n(k)$ together with a linear divisor D defined over k. To each $x \in \mathbb{P}^n(k) - D$, we select all non-Archimedean places v of k with

[3] The role of Vojta's Theorem [V, Theorem 6.4.3] is two-fold in the geometry of Diophantus-Nevanlinna analogy. One is that as the Diophantine analogue of the Lemma on logarithmid deribvative. The other is that as a re-formulation of the Parametric Subspace Theorem.

the property that the Zariski closures of x and D (over the ring O_k of integers) intersect with multiplicity $\geq n$ over the place v. We define S_x^n the collection of such non-Archimedean places v and define an association $x \mapsto S_x^n$. The varying S just means $S(x) = S_\infty \cup S_x^n$. Here we encounter the difficulty stemming from a non-uniform number of non-Archimedean places. In §2, we overcome this difficulty by geometry of numbers with an appropriate choice of the weights in the length function (the associated star-body should have a good "shape" with respect to the non-Archimedean places involved).

In §4, we push the analogy further and formulate this in the shape, Theorem 4.2, completely analogous to the Lemma on logarithmic derivative. Theorem 4.2 appears as a new version of the Parametric Subspace Theorem (because this generalizes [V, Theorem 6.4.3]) from which we deduce the Main Theorem. As is clear from the statement of Theorem 4.2, the new version of Parametric Subspace Theorem splits into the Archimedean and non-Archimedean parts.

The use of the Roth lemma in the proof of the Roth and Schmidt Subspace Theorem is the origin of the ineffectiveness of these theorems. In §5, we propose some conjectures toward the effectiveness from the new framework introduced in §2. The Diophantus-Nevanlinna analogue we establish in this article is based on the Nevanlinna-Cartan theory which consists of the Lemma on logarithmic derivative and the Wronskian formalism (see §3). In the most primitive sense, Parametric Subspace Theorem is the Diophantine analogue of the Lemma on logarithmic derivative and Schmidt's proof of "PSST $\Rightarrow$ SST" is the Diophantine analogue of the Wronskian formalism[4]. Vojta refined this analogue by establishing the "Type A analogue" of the Lemma on logarithmic derivative[5] (cf. Theorem 2.1) by showing "PSST $\Rightarrow$ Type A". Then Vojta's proof of "Type A $\Rightarrow$ SST" turns out to be the Diophantine analogue of the Wronskian formalism at this stage. §5 is an attempt toward pushing this direction further. Reformulating Vojta's proof of "PSST $\Rightarrow$ Type A" in our new framework, we get the "Type B" analogue of the Lemma on logarithmic derivative (see Theorem 4.2). The "Type B" analogue (in Theorem 4.2) splits into the inequalities (16) over S_∞ and certain conditions over S_x^n. Here it is remarkable that we can propose a Diophantine inequality which seems to be much simpler compared to the original Roth type inequality which, we conjecture, would effectively bound the exceptions to the inequality (16). However, since logically "Type A statement $\Rightarrow$ Type B", we cannot avoid establishing "Type A statement" ($\Leftarrow$ PSST $\Leftarrow$ Roth Theorem). In short, we need ineffective "Type A statement" to conclude that the effective bound on the exceptions to (16) gives the effective bound to the "Type A statement". Since the "Type B" analogue of the Lemma on logarithmic derivative obeys the same geometric pattern as in the Nevanlinna theory, we can establish the Diophantine analogue of the Wronskian formalism together with the Diophantine analogue of the Wronskian itself. In the Roth case ($n = 1$), the Archimedean part of the above splitting reduces to a simple Diophantine inequality for $\mathbb{Z}^2$ embedded in $\mathbb{R}^2$. We conjecture that there exists an effective bound for the height of solutions to the opposite inequality (see §5, (17) and (18)). Once we were able to prove this conjecture, we would argue inductively on n and finally get the effective Schmidt Subspace Theorem *with residual counting function*,

[4] PSST (resp. SST) is an abbreviation of the Parametric Subspace Theorem (resp. the Subspace Theorem).

[5] This is [V, Theorem 6.4.3] on which this article is based.

which implies the effective version of the abc-conjecture and its generalization.

1. Lemma on logarithmic derivative.

We refer to [Y], [K1,2] for technical details of this section.

We will freely use standard notation in Nevanlinna theory which we briefly review here. Let X be a smooth projective variety and D any effective divisor. Nevanlinna theory provides a natural framework for the study of approximation to hypersurfaces by transcendental holomorphic curves.

Let $f : \mathbb{C} \to X$ be a holomorphic curve whose image is not entirely contained in $\operatorname{Supp}(D)$.

The proximity function

$$m_{f,D}(r) = \int_0^{2\pi} -\log \operatorname{dist}_{\mathrm{Euc}}(f(re^{i\theta}), D)\frac{d\theta}{2\pi}$$

measures the Euclidean approximation of f to D. Here the Euclidean distance is measured by using a smooth Hermitian norm of $\mathcal{O}_X(D)$ and a defining equation of D.

The counting function

$$N_{f,D}(r) = \sum_{0<|a|<r} \operatorname{ord}_a(\sigma(f)) \log\left|\frac{r}{a}\right| + \operatorname{ord}_0(\sigma(f)) \log r$$

measures the approximation of f to D by counting the number of roots of $\sigma(f(z)) = 0$. The Poisson-Jensen formula implies that the height function

$$T_{f,D}(r) = m_{f,D}(r) + N_{f,D}(r)$$

depends (up to bounded functions) only on the linear equivalence class of D (First Main Theorem in Nevanlinna theory).

Let X be a smooth projective variety and D any effective divisor. We will use the superscript $^{(j)}$ to indicate the j-th jet object. For a holomorphic curve $f : \Delta \to X$, we define the j-th canonical jet lift $f^{(j)} : \mathbb{C} \to X^{(j)}$ by $f^{(j)}(z) = (f(z), f'(z), \ldots, f^{(j)})$. Two germs of holomorphic curves $f_i : \Delta_i \to X$ $(i = 1, 2)$ passing through $x \in X$ (i.e., $f_i(0) = x$) are said to be j-equivalent if and only if f_1 and f_2 have the same Taylor series at $z = 0$ up to order j. The j-th jet space $X^{(j)}$ is by definition the set of all j-equivalence classes of germs of holomorphic curves in X. We write $\pi^{(j)} : X^{(j)} \to X$ for the canonical projection. Let s be a holomorphic function defined on an open set $U \subset X$ and let $f : \Delta \to X$ be a representative of an element of $(\pi^{(j)})^{-1}U$. Then the association

$$(s, f) \mapsto \left.\frac{d^j}{dz^j}\right|_{z=0} s(f(z))$$

canonically defines a holomorphic function $d^j s$ on $(\pi^{(j)})^{-1}U$. Let a proper subscheme Z of X be locally given in terms of the generators of the defining ideal by $Z = V(s_1, \ldots, s_t)$. By V_{reg} we mean the regular part of V and we set

$$Z^{(j)} = \text{the Zariski closure of } V_{\mathrm{reg}}(s_1, \ldots, s_t, ds_1, \ldots, ds_t, \ldots, d^j s_1, \ldots, d^j s_t)$$

and call it the j-th jet space of Z. Let ∞ be the divisor at infinity of (any) projective completion $\overline{X^{(j)}}$ of the jet space $X^{(j)}$.

We are now ready to state a modern version of Nevanlinna's lemma on logarithmic derivative ([Y],[K1,2]).

Theorem 1.1. *Let (X, D) be as above. Let $f : \mathbb{C} \to X$ be a holomorphic curve such that $f(\mathbb{C}) \not\subset \mathrm{Supp}\,(D)$. Let j be any positive integer. Then we have*

$$\begin{cases} m_{f,Z}(r) \leq m_{f^{(j)},Z^{(j)}}(r) + S_f(r) /\!/ , \\ m_{f^{(j)},\infty}(r) \leq S_f(r) /\!/ . \end{cases}$$

Here the symbol $S_f(r)$ indicates a small order error function

$$S_f(r) = O(\log^+(rT_{f,E}(r)))$$

in the asymptotic sense as $r \to \infty$. The symbol $/\!/$ means that the said inequality holds outside a Borel set of finite Lebesgue measure.

The Lemma on logarithmic derivative is important in the following two points:
(i) [On approximation.] Suppose that a holomorphic curve $f : \mathbb{C} \to X$ approximates a proper subscheme Z. Then the first inequality of Theorem 1.1 implies that any j-jet of the holomorphic curve f likely approximates the j-jet space of Z in $X^{(j)}$. Moreover the second inequality implies that any j-jet of f does not approximate the divisor at infinity of any projective completion of $X^{(j)}$.
(ii) [Non sensitivity on targets and subschemes.] The inequalities of Theorem 1.1 are of the same form for any holomorphic curve in any target and with respect to approximation to any proper subscheme.

In the next section, we will discuss what Diophantine analogue is possible about the Lemma on logarithmic derivative from the view point of (i) and (ii).

2. A Diophantine analogue of the Lemma on logarithmic derivative.

In this section we closely follow [V, Chapt. 6] to prove a Diophantine analogue of Nevanlinna's lemma on logarithmic derivative for points of $\mathbb{P}^n(k)$ approximating hyperplanes[6].

We first introduce basic definitions in Diophantine approximation on projective varieties.

Let k be a fixed number field and X a smooth projective variety defined over k and D a divisor. Let v be any place (finite or infinite) of k. To define the Diophantine analogue of proximity/counting/height functions, we need to extend X to an arithmetic variety $\mathcal{X}$ over $\mathrm{Spec}\,(\mathcal{O}_k)$ having $X(k)$ as a fiber over a generic point. The Weil function associated to v is defined by

$$\lambda_{D,v} : X - D \longrightarrow \mathbb{R}_{\geq 0} \; ; \; x \mapsto \lambda_{D,v}(x) := -\log \mathrm{dist}_v(x, D) \; .$$

By using Weil functions, we introduce fundamental functions in Diophantine approximation.

Let S be a finite set of places of k containing all infinite ones. For $v \in M_k$ we set $d_v = [k_v : \mathbb{Q}_p]/[k : \mathbb{Q}]$ (where v divides p). Let $x \in X(k)$.
(i) The S-proximity function is defined by

$$m_S(D, x) = \sum_{v \in S} d_v \lambda_{D,v}(x) \; .$$

[6] The Diophantine analogue we prove in this article works only on the approximation to hyperplanes in projective spaces. We will prove in a future paper a more general analogue on the approximation to general subschemes of general projective varieties.

This measure the approximation to D with respect to places in S.
(ii) The S-counting function is defined by

$$N_S(D,x) = \sum_{v \notin S} d_v \lambda_{D,v}(x) \ .$$

This measures the approximation to D with respect to places outside of S.

Let $\sigma = 0$ be a defining equation of D over k and let p be the prime corresponding to the restriction of v to $\mathbb{Q}$. Then

$$N_S(D,x) = \frac{1}{[k:\mathbb{Q}]} \sum_{v \notin S} \deg_v(\sigma(x)) \log N(v) = \sum_{v \notin S} \frac{\deg_v(\sigma(x))}{\deg_v p} \log p \ .$$

The proximity and counting functions may change drastically if we change D in its linear equivalence class. The quantity invariant under linear change of D will define a complete intersection theory of points and divisors in Diophantine approximation.

The absolute logarithmic height function is defined by the sum of all Weil functions[7]:

$$\mathrm{ht}_D(x) = m_S(D,x) + N_S(D,x) \ .$$

The product formula implies that the absolute logarithmic height function $\mathrm{ht}(x)$ (up to bounded functions) depends only on the linear equivalence class of the divisor D (this is the Diophantine analogue of the First Main Theorem in Nevanlinna Theory).

The asymptotic behavior of these functions is defined for infinite set of k-rational points of X. Therefore,

$$\text{the collection of the images of } f|_{\mathbb{C}(r)} : \mathbb{C}(r) \to X \text{ for unbounded set } \{r\} \\ \overset{\text{analogy}}{\longleftrightarrow} \text{ an infinite set } \{x\} \subset X(k)$$

where $\mathbb{C}(r) = \{z \in \mathbb{C}; |z| < r\}$ and $f : \mathbb{C} \to X$ is an entire holomorphic curve. The prime structure is intrinsic in Diophantine geometry, while its Nevanlinna analogue is to *distinguish* the "map" $f : \mathbb{C} \to X$ from its "image" $f(\mathbb{C})$:

$$(\mathbb{C}(r), \text{point measure}) \overset{\text{analogy}}{\longleftrightarrow} \text{finite places of } k \ ,$$
$$(\partial\mathbb{C}(r), \frac{d\theta}{2\pi}) \overset{\text{analogy}}{\longleftrightarrow} \text{infinite places of } k \ .$$

Note that there is no "canonical" Diophantine notion of the "Nevanlinna theoretic non-Archimedean places". The above defined analogue depends on the holomorphic curve f and the divisor $D = (\sigma)$ ($a \in \mathbb{C}(r)$ is a "non-Archimedean place" if and only if $\mathrm{ord}_a(\sigma(f))$ is positive). The main idea of this article is to reverse the orientation of considering analogies and to consider the "Diophantine analogue" of the non-existence of the "canonical definition" of the "Nevanlinna theoretic finite places".

---

[7] The absolute logarithmic hight function does not depend on the choice of the field K over which x is defined.

To formulate a Diophantine analogue of the Lemma on logarithmic derivative, we need a Diophantine analogue of the derivative for points of $\mathbb{P}^n(k)$ (note that the derivative is defined intrinsically for holomorphic curves but not for rational points).

For this purpose, we identify $T_{[x]}\mathbb{P}_n$ ($[x] \in \mathbb{P}_n(k)$ where $x \in \mathcal{O}_{k,S}^{n+1} - \{0\}$ without common non S-unit factor) with $k^{n+1} \wedge x \cong k^{n+1}/\langle x \rangle > \cong k^n$ up to $\mathcal{O}(-1)_x$ so that we can work in linear algebra[8]. This space inherits a canonical lattice structure $\mathcal{O}_{k,S}^{n+1} \wedge x \subset k^{n+1} \wedge x$.

To proceed, we consider two families of lines in $\mathbb{P}_{n+1}(k)$ (= the projective completion of k^{n+1}). One is the family $\mathcal{F}_1$ of lines in $\mathbb{P}(k^{n+1})$ passing through $[x] \in \mathbb{P}_n(k)$ (the hyperplane at infinity). The other family $\mathcal{F}_2$ consists of those lines passing through the origin of k^{n+1} (naturally parameterized by the hyperplane $\mathbb{P}_n(k)$ at infinity). For any point p of $\mathbb{P}_n(k)$ we pick a point $y \in k^{n+1}$ in the corresponding line in $\mathcal{F}_2$ in such a way that $y \in \mathcal{O}_{k,S}^{n+1}$ without common non S-unit factor.

Let $x, v_2, \ldots, v_{n+1}$ be a basis of $\mathcal{O}_{k,S}^{n+1}$ and set $y = y_1 x + y_2 v_2 + \cdots + y_{n+1} v_{n+1}$. The S-unit factor is undetermined for the above choice of y. However, the absolute logarithmic height

$$\overline{h}(y \wedge x) := \frac{1}{[k:\mathbb{Q}]} \sum_{v \in S} [k_v : \mathbb{Q}_v] \log \max_{2 \le i \le n+1} |y_i|_v$$

is independent of the undetermined S-unit factor. The set of all points in $\mathbb{P}_n(k)$ with the same $y \wedge x \in \mathcal{O}_{k,S}^{n+1} \wedge x$ forms an infinite set having the following properties:
(a) This infinite set accumulates at the "center of gravity" $[x] \in \mathbb{P}_n(k)$.
(b) This infinite set lies on the line of $\mathbb{P}_n(k)$ which is the intersection of $\mathbb{P}_n(k)$ and the 2-plane in $\mathbb{P}_{n+1}(k)$ determined by x and y.
We interpret this infinite set as an analogue of a holomorphic curve $c : \mathbb{C} \to \mathbb{P}_n(\mathbb{C})$ with $f(0) = [x]$ and $y \wedge x \in \mathcal{O}_{k,S}^{n+1} \wedge x \subset T_{[x]}\mathbb{P}_n(k)$ an analogue of the derivative $c'(0)$. More explicitly, we associate to $y \wedge x \in k^{n+1} \wedge x$ (y being as above) a holomorphic curve $c(z)$ in the following way. The line in $\mathbb{P}_n(k)$ determined by the set of all points in $\mathbb{P}_n(k)$ having the same $y \wedge x \in \mathcal{O}_{k,S}^{n+1}$ is parameterized as $z \mapsto c(z) := [x + zy]$ (x and y being linearly independent). This is what we want to have. Indeed, we introduce a system of homogeneous coordinates so that $x = (x_0 : \cdots : x_n)$ and $y = (y_0 : \cdots : y_n)$ and assume that $x_0 + y_0 \neq 0$. Then, for z with $|z|$ small, we have

$$\begin{aligned} c(z) &= [x_0 + zy_0 : \cdots : x_n + zy_n] \\ &= [1 : (x_1 + zy_1)x_0^{-1}(1 + z\frac{y_0}{x_0})^{-1} : \cdots : (x_n + zy_n)x_0^{-1}(1 + z\frac{y_n}{x_n})^{-1}] \\ &= [1 : \frac{x_1}{x_0} + z\frac{x_0 y_1 - x_1 y_0}{x_0} + O(z^2) : \cdots : z\frac{x_0 y_n - x_n y_0}{x_0} + O(z^2)] \,. \end{aligned}$$

In a similar way, given $y^{(1)}$ and $y^{(2)}$ in $\mathcal{O}_{k,S}^{n+1}$, we can associate a holomorphic curve $c(z)$ with the property that $c(0) = [x]$ and $y^{(1)} \wedge x$ (resp. $y^{(2)} \wedge x$) as an analogue of $c'(0)$ (resp. $c''(0)$) by setting

$$c(z) := \left[x + zy^{(1)} + z^2 \left(y^{(2)} - \frac{y^{(1)}}{x_0} \right) \right] .$$

[8] The Euler exact sequence implies the isomorphism $k^{n+1}/\langle x \rangle \cong T_{[x]}\mathbb{P}^n \otimes \mathcal{O}(-1)_{[x]}$.

Indeed, we have

$$c(z) = \left[1 : \frac{x_1}{x_0} + z\frac{x_0 y_1^{(1)} - x_1 y_0^{(1)}}{x_0} + z^2 \frac{x_0 y_1^{(2)} - x_i y_0^{(2)}}{x_0} : \right.$$
$$\left. \cdots : \frac{x_n}{x_0} + z\frac{x_0 y_n^{(1)} - x_n y_0^{(1)}}{x_0} + z^2 \frac{x_0 y_n^{(2)} - x_n y_0^{(2)}}{x_0}\right].$$

Thus, given $y^{(1)} \wedge x$ and $y^{(2)} \wedge x$ in $\mathcal{O}_{k,S}^{n+1} \wedge x \in T_{[x]}\mathbb{P}_n(k)$, we can associate a holomorphic curve in $\mathbb{P}_n(\mathbb{C})$ in a canonical way. Under this correspondence, $\overline{h}(y \wedge x)$ being large is an analogue of $|f'(0)|$ being large. Under the above situation we have $\frac{1}{k!}\frac{d^k c}{dt^k}(0)$ corresponds to $y^{(k)}$ $(k = 1, 2)$. Moreover, we can generalize the above argument to any given $y^{(1)}, \ldots, y^{(j)}$ $(j = 1, \ldots, n)$. This suggests that a Diophantine analogue of the sequence of jets of a holomorphic curve is defined by successively taking linearly independent sequence of vectors in the lattice $\mathcal{O}_{k,S}^{n+1} \wedge x$. Minkowski's geometry of numbers (the theory of successive minima) together with the appropriate choice of convex bodies provide us a geometric framework for the Diophantine analogue of the derivatives. It is Vojta's observation in [V, Chapt 6] that the Diophantine analogue of the derivatives should be built modelled after the Lemma on logarithmic derivative in the Nevanlinna theory. Indeed, the Lemma on logarithmic derivative (Theorem 1.1) suggests the best choice of the convex bodies in the theory of successive minima.

The following theorem is a generalization of [V, Theorem 6.4.3]. We fix any finite set S of places of k containing all Archimedean ones. For k-rational point x, we introduce a temporary notion of the relative logarithmic height by setting $\mathrm{ht}'(x) = [k : \mathbb{Q}]\,\mathrm{ht}(x)$ ($\mathrm{ht}(x)$ being the absolute logarithmic height).

Theorem 2.1. *Let $F_0, \ldots, F_N$ be a set of linear forms in k^{n+1} in general position. Let $\varepsilon > 0$. Then there exists a finite set $\mathcal{S}$ of proper linear subspaces of k^{n+1} such that if $x \in k^{n+1}$ is not a vector in the union of the linear subspaces in $\mathcal{S}$, then there exists a sequence $x^{(1)}, \ldots, x^{(n)} \in \mathcal{O}_{k,S}^{k+1}$ of vectors such that $x \wedge x^{(1)} \wedge \cdots \wedge x^{(n)} \neq 0$ and for each $p = 1, \ldots, n$, the following inequality holds: For $p = 1, \ldots, n$, we set $x^{\leq p-1} = x \wedge x^{(1)} \wedge \cdots \wedge x^{(p-1)}$ and $F_{i,p} = F_i \wedge F_{n-p+2} \wedge \cdots \wedge F_n$. Then, after suitably re-ordering the F's, we have*

$$\sum_{v \in S} \log \frac{||(x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \cdot F_{i,p}||_v}{||x^{\leq p-1}||_v ||x^{\leq p-1} \cdot F_{i,p}||_v} < \varepsilon\, \mathrm{ht}'(x)$$

for all $i = 0, \ldots, N$ and for all x such that $x^{\leq p-1} \cdot F_{i,p} \neq 0$. If $x^{\leq p-1} \cdot F_{i,p} = 0$ then $(x^{\leq p-2} \wedge x^{(p)}) \cdot F_{i,p} = 0$.

As $x^{\leq 0} = x$ and $x^{\leq -1} \wedge x^{(1)} = x^{(1)}$, [V, Theorem 6.4.3] corresponds to putting $p = 1$ in Theorem 2.1. The following lemma is a "higher jet analogue" of [V, Lemma 6.4.4].

Lemma 2.2. *Let k be a field with absolute value $|\cdot|$. Let $F_0, \ldots, F_N$ be $N + 1$ linear forms in general position. If for some index i,*

$$|(x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \cdot F_{i,p}| > A\,|x^{\leq p-1}||x^{\leq p-1} \cdot F_{i,p}|\,,$$

then there exists a constant $c > 0$ (the constant c depends only on the F's) such that after re-ordering the F's (the re-ordering depends on the F's and also on $x^{\le p-1}$)

$$|(x^{\le p-1} \wedge (x^{\le p-2} \wedge x^{(p)})) \cdot (F_{j,p} \wedge F_{n-p+1,p})| > cA\,|x^{\le p-1}||x^{\le p-1} \cdot F_{j,p}|$$

holds for some index j (the index j depends on the F's, $x^{\le p-1}$ and also on $x^{\le p-2} \wedge x^{(p)}$).

Proof. We introduce the lexicographical order to the set of all $i = (i_1, \dots, i_p)$ satisfying $0 \le i_1 < \cdots < i_p \le N$ and write $F_i = F_{i_1} \wedge \cdots \wedge F_{i_p}$. After re-ordering the F's, we may assume that $|x^{\le p-1} \cdot F_i|$ is ordered lexicographically:

$$|x^{\le p-1} \cdot F_{0,1,\dots,p-1}| \le \cdots \le |x^{\le p-1} \cdot F_{N-p+1,\dots,N}| \; .$$

In particular we may assume that

$$|x^{\le p-1} \cdot F_{0,p}| \le \cdots \le |x^{\le p-1} \cdot F_{n-p+1,p}| \; .$$

Noting that $k^{n+1} \wedge (x \wedge x^{(1)} \wedge \cdots \wedge x^{(p-2)}) \cong k^{n+1}/\langle x, x^{(1)}, \dots, x^{(p-2)}\rangle$, we see that $F_{0,p}, \dots, F_{n-p+1,p}$ form a basis of $(k^{n+1}/\langle x, x^{(1)}, \dots, x^{(p-2)}\rangle)^*$. Let $F^*_{0,p}, \cdots, F^*_{n-p+1,p}$ be its dual basis. Let $a \ll b$ indicate that there exists a constant c such that $a \le cb$. We use this abbreviation if the constant depends on the arguments in a uniform way. Otherwise we will take special care. For instance, if there are two uniform constants c_1 and c_2 such that $c_1 b \le a \le c_2 b$, we write $a \gg\ll b$. Hereafter the constant implicit in $\ll$, *etc.* depends only on the F's uniformly. We then have $|x^{\le p-1}| \gg\ll |x^{\le p-1} \cdot F_{n-p+1,p}|$. As the lemma dose not change if we add a scalar multiple of $x^{(p-1)}$ to $x^{(p)}$, we may assume $(x^{\le p-2} \wedge x^{(p)}) \cdot F_{n-p+1,p} = 0$.

We now claim that the assumption of Lemma 2.2 implies

$$\max_{0 \le j \le n-p+1} \frac{|(x^{\le p-1} \wedge (x^{\le p-2} \wedge x^{(p)})) \cdot F_{j,p}|}{|x^{\le p-1} \cdot F_{j,p}|} \gg A \; .$$

To prove this, we introduce a unit vector u proportional to $x^{\le p-1}$. Then the vectors $F^*_{0,p}, \cdots, F^*_{0,n-p}$ and u form a new basis of $k^{n+1}/\langle x, x^{(1)}, \dots, x^{(p-2)}\rangle$. It follows from the rule of re-ordering F's that the transition matrix associated to this base change has bounded coefficients. The same is true for its inverse. The coordinates of $(x^{\le p-1} \wedge (x^{\le p-2} \wedge x^{(p)})) \cdot F_{i,p}$ relative to this basis are computed by evaluating its dual basis. The j-th coordinate for $0 \le j \le n-p$ is

$$\begin{vmatrix} x^{\le p-1} \cdot F_{i,p} & (x^{\le p-2} \wedge x^{(p)}) \cdot F_{i,p} \\ 0 & (x^{\le p-2} \wedge x^{(p)})_j \end{vmatrix} = (x^{\le p-1} \cdot F_{i,p})(x^{\le p-2} \wedge x^{(p)})_j$$

as all j-th coordinates (for $j \le n-p$) of x vanish. Here $(x^{\le p-2} \wedge x^{(p)})_j$ is the j-th coordinate of $x^{\le p-2} \wedge x^{(p)}$ relative to this basis. On the other hand, its $(n-p+1)$-st coordinate is

$$\begin{vmatrix} x^{\le p-1} \cdot F_{i,p} & (x^{\le p-2} \wedge x^{(p)}) \cdot F_{i,p} \\ |x^{\le p-1}| & 0 \end{vmatrix} = -((x^{\le p-2} \wedge x^{(p)}) \cdot F_{i,p})|x^{\le p-1}| \; .$$

as $x^{\leq p-2} \wedge x^{(p)}$ is a linear combination of $F^*_{i,p}$'s for $0 \leq i \leq n-p$ (this implies that its $(n-p+1)$-st coordinate vanishes). Therefore we have

$$\begin{aligned}&|(x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \cdot F_{i,p}|\\ \gg\ll &\max\{|x^{\leq p-1} \cdot F_{i,p}||x^{\leq p-2} \wedge x^{(p)}|, |(x^{\leq p-2} \wedge x^{(p)}) \cdot F_{i,p}||x^{\leq p-1}|\}\ .\end{aligned}$$

Dividing this by $|x^{\leq p-1} \cdot F_{i,p}||x^{\leq p-1}|$, we see that the assumption of Lemma 2.2 becomes

$$\max\left\{\frac{|(x^{\leq p-2} \wedge x^{(p)}) \cdot F_{i,p}|}{|x^{\leq p-1} \cdot F_{i,p}|}, \frac{|x^{\leq p-2} \wedge x^{(p)}|}{|x^{\leq p-1}|}\right\} \gg A\ .$$

If the first term is larger, we have the claim. Suppose that the second term is larger. As $(x^{\leq p-2} \wedge x^{(p)}) \cdot F_{n-p+1,p} = 0$, there exists some j such that

$$|x^{\leq p-2} \wedge x^{(p)} \cdot F_{j,p}| \gg |x^{\leq p-2} \wedge x^{(p)}| \gg A|x^{\leq p-1}|\ .$$

On the other hand, $|x^{\leq p-1}| \gg |x^{\leq p-1} \cdot F_{j,p}|$ for this j. Re-ordering F's again if necessary, we have the claim.

We now return to the original basis $\{F^*_{i,p}\}$ $(0 \leq i \leq n-p+1)$. Let j be the index such that

$$\frac{|(x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \cdot F_{j,p}|}{|x^{\leq p-1} \cdot F_{j,p}|} \gg A$$

holds as in the claim. We then have

$$\begin{aligned}&|(x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \cdot (F_{j,p} \wedge F_{n-p+1,p})|\\ =&\left|\det\begin{pmatrix} x^{\leq p-1} \cdot F_{j,p} & (x^{\leq p-2} \wedge x^{(p)}) \cdot F_{j,p} \\ x^{\leq p-1} \cdot F_{n-p+1,p} & x^{\leq p-2} \wedge x^{(p)} \cdot F_{n-p+1,p}\end{pmatrix}\right|\\ \gg&|(x^{\leq p-2} \wedge x^{(p)}) \cdot F_{j,p}||x^{\leq p-1}| \qquad (\text{as } (x^{\leq p-2} \wedge x^{(p)}) \cdot F_{n-p+1,p} = 0)\\ \gg&A|x^{\leq p-1} \cdot F_{j,p}||x^{\leq p-1}|\ .\end{aligned}$$

We have thus proved Lemma 2.2. □

Proof of Theorem 2.1 : We consider the following sequence of statements indexed by $p = 1, \ldots, n$:

Statement (S_p): "Let $F_0, \cdots, F_N$ be a set of linear forms on k^{n+1} in general position. Let $\varepsilon > 0$. Then there exists a finite set $\mathcal{S}_p$ of points in $\Lambda^p k^{n+1}$ such that if a sequence $x, x^{(1)}, \ldots, x^{(p-1)}$ satisfies the condition that $x^{\leq p-1}$ is not a scalar multiple of a vector in $\mathcal{S}_p$, then there exists a $x^{(p)} \in \mathcal{O}^{k+1}_{k,S}$ such that $x \wedge x^{(1)} \wedge \cdots \wedge x^{(p)} \neq 0$ and the inequality

$$\sum_{v \in S} \log \frac{||(x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \cdot F_{i,p}||_v}{||x^{\leq p-1}||_v||x^{\leq p-1} \cdot F_{i,p}||_v} < \varepsilon\, \mathrm{ht}'(x) \tag{1}$$

holds for all $i = 0, \ldots, N$ and for all x such that $x^{\leq p-1} \cdot F_{i,p} \neq 0$. If $x^{\leq p-1} \cdot F_{i,p} = 0$ then $(x^{\leq p-2} \wedge x^{(p)}) \cdot F_{i,p} = 0$."

We proceed by induction on p. The case $p = n$ is Theorem 2.1 which we want to prove.

The statement (S_1) coincides with [V, Theorem 6.4.3]. Suppose that the statement (S_q) is true up to $q = p-1$. What we have to prove is that the statement (S_p) holds. For this we closely follow the arguments in [V, pp.107-111]. Assume that the statement (S_p) is false. Then there exists an infinite sequence of $\{x^{\leq j}\}_{j=1}^{p-1}$ for each of which there exists no $x^{(p)}$ which satisfies the inequality (1). In the case $p = 1$, Vojta [V, p.107] considered an infinite sequence of x with no suitable x'. Our $\{x^{\leq j}\}_{j=1}^{p-1}$ replaces Vojta's x in [V, p.107]. We will show that the non-existence of a suitable $x^{(p)}$ is equivalent to the statement that a certain first successive minima is large. Then we will use Davenport's lemma [V, Lemma 6.2.1] and some multilinear algebra to arrive at the infinite sequence of $\{x^{\leq j}\}_{j=1}^{p-1}$ which contradicts to the Parametric Subspace Theorem [V, Theorem 6.4.2].

For each $v \in S$ and $\{x^{\leq j}\}_{j=1}^{p-1}$ in the sequence, we consider k with absolute value $||\cdot||_v$ and re-order the F's as in the proof of Lemma 2.2[9]. We use the subscript v to indicate that the re-ordering of the F's is with respect to the place v. For instance, we write F_i (resp. $F_{i,p}$) in this situation as $F_{v;i}$ (resp. $F_{v;i,p}$). So we have

$$||x^{\leq p-1} \cdot F_{v;0,p}||_v \leq \cdots \leq ||x^{\leq p-1} \cdot F_{v;n-p+1,p}||_v \, .$$

In particular we have

$$\prod_{v \in S} ||x^{\leq p-1} \cdot F_{v;n-p+1,p}||_v \gg\ll \exp(\mathrm{ht}'(x^{\leq p-1})) \, .$$

As F's are given and x's are integral, this estimate is uniform in S. Moreover, we may assume that

$$\prod_{v \in S} ||x^{\leq p-2} \cdot F_{v;\geq n-p+2}||_v \gg\ll \exp(\mathrm{ht}'(x^{\leq p-2}))$$

uniformly in S. Here $F_{v;\geq n-p+2} = F_{v;n-p+2} \wedge \cdots \wedge F_{v;n}$. By passing to an appropriate infinite subsequence, we may assume that $F_{v;i,p}$ $(i = 0, \ldots, n-p+1)$ do not depend on $\{x^{\leq j}\}_{j=1}^{p-1}$ in the sequence. Since the inequality (1) is false, Lemma 2.2 implies that there exist indices i (depending on v, $x^{\leq p-2}$ and $x^{\leq p-1}$) such that

$$\text{(2)} \qquad \sum_{v \in S} \log \frac{||(x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \cdot (F_{v;i,p} \wedge F_{v;n-p+1,p})||_v}{||x^{\leq p-1}||_v ||x^{\leq p-1} \cdot F_{v;i,p}||_v} > \varepsilon \, \mathrm{ht}'(x^{\leq p-1})$$

holds[10]. We want to interpret this inequality in terms of Bombieri-Vaaler's adelic version of successive minima ([B-V], see also [V, pp.90-96]). On $k^{n+1} \wedge x^{\leq p-1}$, we consider the system of successive minima with respect to the lattice structure induced from that of $\mathcal{O}_{k,S}^{n+1}$ and the star body given by the length function determined

[9] Re-ordering F's separately for different v's in S has its Nevanlinna analogue. That is, given a holomorphic curve $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$, dividing the circle $\partial\mathbb{C}(r)$ into sub-arcs C_i, so that $f|_{C_i}$ approximates different portion of the divisor defined by the linear forms $F_0, \ldots, F_N$.

[10] If S varies in non-uniform way, the passage from Lemma 2.2 to (2) will need special care.

by the following data[11]:

$$
\begin{aligned}
(3) \qquad & L_{v;i,p}(x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \\
& = (x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \cdot ((F_{v;n-p+1} \wedge F_{v;\geq n-p+2}) \wedge F_{v;i,p}) \ , \\
A_{v;i,p} & = \frac{1}{||x^{\leq p-1}||_v ||x^{\leq p-1} \cdot F_{v;i,p}||_v ||x^{\leq p-2} \cdot F_{v;\geq n-p+2}||_v}
\end{aligned}
$$

for $i = 0, \dots, n-p$ (Note that $F_{v;n-p+1,p} = F_{v;n-p+1} \wedge F_{v;\geq n-p+2}$). The inequality (2) implies that the first successive minimum is large:

$$\lambda_1^{[k:\mathbb{Q}]} \gg \exp(\varepsilon \,\mathrm{ht}'(x^{\leq p-1})) \ .$$

Our strategy is to compare this inequality with the estimates of the successive minima derived from the Bombieri-Vaaler theory [B-V] (see [V, Theorem 6.1.11]) and to derive a situation which violates the Parametric Subspace Theorem [V, Theorem 6.4.2]. In order to do this, we need to measure the relative volume of $F_{v;i,p} \wedge F_{v;n-p+1,p}$ $(0 \leq i \leq n-p)$ relative to the lattice

$$
\begin{aligned}
(\mathcal{O}_{k,S}^{n+1} \wedge x^{\leq p-2}) \wedge x^{\leq p-1} & \subset (k^{n+1} \wedge x^{\leq p-2}) \wedge x^{\leq p-1} \\
& \cong (k^{n+1} \wedge x^{\leq p-2}) / \langle x^{\leq p-1} \rangle \\
& \cong k^{n+1} / \langle x, x^{(1)}, \cdots, x^{(p-1)} \rangle \ .
\end{aligned}
$$

To compute the relative volume, we consider the standard basis $\{e_i\}_{i=0}^{n}$ and the associated coordinate functions $\{x_i\}_{i=0}^{n}$. We assume that $(e_0 \wedge \cdots \wedge e_{p-2})(x^{\leq p-2}) \neq 0$ and $(e_0 \wedge \cdots \wedge e_{p-1})(x^{\leq p-1}) \neq 0$. Then

$$(e_p \wedge x^{\leq p-2}) \wedge x^{\leq p-1}, \dots, (e_n \wedge x^{\leq p-2}) \wedge x^{\leq p-1}$$

form a basis for a sublattice of $(\mathcal{O}_{k,S}^{n+1} \wedge x^{\leq p-2}) \wedge x^{\leq p-1}$. We need to know its index. To compute the index, let $v_{p-1} := x^{\leq p-1}, v_p, \dots, v_n$ form a basis of the lattice $\mathcal{O}_{k,S}^{n+1} \wedge x^{\leq p-2}$. Then the index is

$$
\begin{aligned}
& ((v_p^* \wedge v_{p-1}^*) \wedge \cdots \wedge (v_n^* \wedge v_{p-1}^*)) \\
& \quad \cdot (((e_p \wedge x^{\leq p-2}) \wedge v_{p-1}) \wedge \cdots \wedge ((e_n \wedge x^{\leq p-2}) \wedge v_{p-1})) \\
& = \det \begin{pmatrix} v_p^*(e_p \wedge x^{\leq p-2}) & \cdots & v_p^*(e_n \wedge x^{\leq p-2}) \\ \vdots & \cdots & \vdots \\ v_n^*(e_p \wedge x^{\leq p-2}) & \cdots & v_n^*(e_n^* \wedge x^{\leq p-2}) \end{pmatrix} .
\end{aligned}
$$

[11] The set up for the successive minima on k^n relative to S is the following. For each $v \in S$ let $L_{v,1}, \dots, L_{v,n}$ be n linearly independent linear forms with coefficients in k and let $A_{v;1}, \dots, A_{v;n}$ be positive real numbers. Given such data, we define the length function by

$$f(x)^{[k:\mathbb{Q}]} = \prod_{v \in S} \max_{1 \leq i \leq n} A_{v;i} ||L_{v;i}(x)||_v \ .$$

For details, see [V, p.90-96] for this length function, especially the associated star body and its relation to the Bombieri-Vaaler theory.

On the other hand, $e_{p-1}\wedge x^{\le p-2}, e_p\wedge x^{\le p-2},\dots,e_n\wedge x^{\le p-2}$ also form a basis of the lattice $\mathcal{O}_{k,S}^{n+1}\wedge x^{\le p-2}$. Writing v_{p-1} as a linear combination of $e_{p-1}\wedge x^{\le p-2}, e_p\wedge x^{\le p-2},\dots,e_n\wedge x^{\le p-2}$ and applying $v_{p-1}^*,\dots,v_n^*$, we have

$$\begin{aligned}
1&=\det\begin{pmatrix} v_{p-1}^*(e_{p-1}\wedge x^{\le p-2}) & v_{p-1}^*(e_p\wedge x^{\le p-2}) & \cdots & v_{p-1}^*(e_n\wedge x^{\le p-2})\\ v_p^*(e_{p-1}\wedge x^{\le p-2}) & v_p^*(e_p\wedge x^{\le p-2}) & \cdots & v_p^*(e_n\wedge x^{\le p-2})\\ \vdots & \vdots & \cdots & \vdots\\ v_n^*(e_{p-1}\wedge x^{\le p-2}) & v_n^*(e_p\wedge x^{\le p-2}) & \cdots & v_n^*(e_n\wedge x^{\le p-2})\end{pmatrix}\\
&=\det\begin{pmatrix} 1/x_{p-1} & v_{p-1}^*(e_p\wedge x^{\le p-2}) & \cdots & v_{p-1}^*(e_n\wedge x^{\le p-2})\\ 0 & v_p^*(e_p\wedge x^{\le p-2}) & \cdots & v_p^*(e_n\wedge x^{\le p-2})\\ \vdots & \vdots & \vdots & \vdots\\ 0 & v_n^*(e_p\wedge x^{\le p-2}) & \cdots & v_n^*(e_n\wedge x^{\le p-2})\end{pmatrix}
\end{aligned}$$

where $v_{p-1}=x_{p-1}(e_{p-1}\wedge x^{\le p-2})+\cdots$. So, the index is

$$\prod_{v\in S}\left\|\det\begin{pmatrix} v_p^*(e_p\wedge x^{\le p-2}) & \cdots & v_p^*(e_n\wedge x^{\le p-2})\\ \vdots & \cdots & \vdots\\ v_n^*(e_p\wedge x^{\le p-2}) & \cdots & v_n^*(e_n^*\wedge x^{\le p-2})\end{pmatrix}\right\|_v=\prod_{v\in S}||x_{p-1}||_v\,.$$

The volume of $F_{v;i,p}\wedge F_{v;n-p+1,p}$ $(0\le i\le n-p)$ relative to the sublattice of $(\mathcal{O}_{k,S}^{n+1}\wedge x^{\le p-2})\wedge x^{\le p-1}$ formed by $e_p\wedge x^{\le p-2},\cdots,e_n\wedge x^{\le p-2}$ of index $\prod_{v\in S}||x_{p-1}||_v$ is

$$\begin{aligned}
&\prod_{v\in S}||((e_p\wedge x^{\le p-2})\wedge x^{\le p-1})\wedge\cdots\wedge((e_n\wedge x^{\le p-2})\wedge x^{\le p-1}))\\
&\qquad\cdot((F_{v;0,p}\wedge F_{v;n-p+1,p})\wedge\cdots\wedge(F_{v;n-p,p}\wedge F_{v;n-p+1,p}))||_v^{-1}\\
&=\Bigg[\prod_{v\in S}||F_{v;n-p+1,p}\cdot x^{\le p-1}||_v^{n-p}||(x^{\le p-1}\wedge(e_p\wedge x^{\le p-2})\wedge\cdots\wedge(e_n\wedge x^{\le p-2}))\\
&\qquad\cdot(F_{v;0,p}\wedge\cdots\wedge F_{v;n-p,p}\wedge F_{v;n-p+1,p})||_v\Bigg]^{-1}\\
&\gg\ll\Bigg[\exp((n-p)\mathrm{ht}'(x^{\le p-1}))\\
&\qquad\cdot\prod_{v\in S}\Bigg(||x_{p-1}||_v\,||x^{\le p-2}||_v^{n-p+1}\\
&\qquad\quad\cdot||(e_{p-1}\wedge e_p\wedge\cdots\wedge e_n)\cdot(F_{v;0}\wedge\cdots\wedge F_{v;n-p+1})||_v\Bigg)\Bigg]^{-1},
\end{aligned}$$

where the constant implicit in $\gg\ll$ is uniform in S. Note that the index $\prod_{v\in S}||x_{p-1}||_v$ appears as a factor in the expression of the volume. Therefore this factor disappears if we consider the absolute volume.

Let $\lambda_1,\dots,\lambda_{n-p+1}$ be the system of successive minima for the length function determined by the data (3). By the Bombieri-Vaaler theory [B-V] on successive minima (see also [V, Theorem 6.1.11]), we have the estimate

$$
\begin{aligned}
&(\lambda_1\lambda_2\cdots\lambda_{n-p+1})^{[k:\mathbb{Q}]}\\
&\gg\ll \exp((n-p)\mathrm{ht}'(x^{\leq p-1})+(n-p+1)\mathrm{ht}'(x^{\leq p-2}))\cdot \prod_{\substack{v\in S\\ 0\leq i\leq n-p}} A_{v;i,p}\\
&\gg\ll \frac{1}{\prod_{v\in S}\prod_{i=0}^{n-p+1}||x^{\leq p-1}\cdot F_{v;i,p}||_v}
\end{aligned}
$$

uniform in S. Following [V, p.108], we apply Davenport's lemma [V, Lemma 6.2.1][12] with $\rho_i=\rho/\lambda_i$, where we choose ρ so that $\rho_1\rho_2\cdots\rho_{n-p+1}=1$. As a result, we infer that there exist constants $\rho_{v;i}$ $(i=0,\dots,n-p)$ with the following properties:

(i) Consider the star body given by the length function $f'(x^{\leq p-1}\wedge(x^{\leq p-2}\wedge x^{(p)}))$ determined by the same $L_{v;i,p}$ as in (3) but with a new $A_{v;i,p}$ which differs from (3) by the factor $\rho_{v;i}$, i.e.,

$$
f'(x^{\leq p-1}\wedge(x^{\leq p-2}\wedge x^{(p)}))=\prod_{v\in S_\infty}\max_{0\leq i\leq n-p}\rho_{v,i}A_{v;i,p}L_{v;i,p}(x^{\leq p-1}\wedge(x^{\leq p-2}\wedge x^{(p)}))\ .
$$

Then the associated successive minima λ_i' $(i=1,\dots,n-p+1)$ satisfy

$$
\lambda_i' \gg\ll \left[\prod_{v\in S}\prod_{i=0}^{n-p+1}||x^{\leq p-1}\cdot F_{v;i,p}||_v\right]^{-\frac{1}{(n-p+1)[k:\mathbb{Q}]}}
$$

uniform in S.

(ii) Let N_v be defined by $N_v=2$ if v is complex, $=1$ if v is real and $=0$ if v is non-Archimedean.

Using

$$
\lambda_1^{[k:\mathbb{Q}]} \gg \exp(\varepsilon\,\mathrm{ht}'(x^{\leq p-1}))
$$

we infer that $\rho_{v;i}$ $(i=0,\dots,n-p)$ satisfy

(4)
$$
\begin{aligned}
\rho_{v;i} &\ll (\lambda_1'/\lambda_1)^{N_v} \qquad [\text{for all } v\in S,\ i=0,\dots,n-p]\\
&\ll \left[\prod_{v\in S}\prod_{0\leq i\leq n-p+1}||x^{\leq p-1}\cdot F_{v;i,p}||_v\right]^{-\frac{N_v}{(n-p+1)[k:\mathbb{Q}]}}\cdot\exp(-\varepsilon\,\mathrm{ht}'(x^{\leq p-1}))^{\frac{N_v}{[k:\mathbb{Q}]}}
\end{aligned}
$$

uniform in S.

Now we consider $\Lambda^{n-p}(x^{\leq p-1}\wedge(\mathcal{O}_{k,S}^{n+1}\wedge x^{\leq p-2}))$ (this is of rank $n-p+1$) and the successive minima defined by the length function determined by the data

[12] Davenport's lemma involves the procedure of scaling by unit, making certain factors indexed by S_∞ into the same order of magnitude. This necessarily produces error depending only on k and S. So we must take care of the extra error stemming from "scaling by unit". From geometric view point, Davenport's lemma plays the role of "choosing a good gauge" in differential geometry. Indeed, "scaling by unit" is analogous to applying a gauge transformation.

$(L_{v;\widehat{m}}, A_{v;\widehat{m}})$ $(m = 0, \dots, n-p)$ where $L_{v;\widehat{m}}$ is a linear form on $\Lambda^{n-p}(x^{\leq p-1} \wedge (k^{n+1} \wedge x^{\leq p-2})) \cong k^{n+1}/\langle x, x^{(1)}, \dots, x^{(p-1)}\rangle$ defined by

$$(F_{v;0,p} \wedge F_{v;n-p+1,p}) \wedge \cdots \wedge (\widehat{F_{v;m,p} \wedge F_{v;n-p+1,p}}) \wedge \cdots \wedge (F_{v;n-p,p} \wedge F_{v,n-p+1,p})$$

and $A_{v;\widehat{m}}$ is defined by

$$A_{v;\widehat{m}} = A_{v;0} \cdots \widehat{A_{v;m}} \cdots A_{v;n-p} \,.$$

Then [V, Proposition 6.3.10] applied to $\Lambda^{n-p}(x^{\leq p-1} \wedge (\mathcal{O}_{k,S}^{n+1} \wedge x^{\leq p-2}))$ implies that the associated successive minima $\mu_1, \dots, \mu_{n-p+1}$ satisfy

$$\mu_i \gg\ll \prod_{v \in S} \prod_{0 \leq i \leq n-p+1} ||x^{\leq p-1} \cdot F_{v;i,p}||_v^{-\frac{n-p}{(n-p+1)[k:\mathbb{Q}]}}$$

uniform in S. This implies that there exists a full sublattice of $\Lambda^{n-p}(x^{\leq p-1} \wedge (\mathcal{O}_{k,S}^{n+1} \wedge x^{\leq p-2}))$ with a basis $v_1, \dots, v_{n-p+1}$ (depending on $x^{(i)}$, $i \leq p-1$), such that, after scaling each v_j by an appropriate unit[13], we have

$$\begin{aligned}
(*) \qquad & ||v_j \cdot ((F_{v;0,p} \wedge F_{v;n-p+1,p}) \wedge \cdots \\
& \quad \wedge (\widehat{F_{v;m,p} \wedge F_{v;n-p+1,p}}) \wedge \cdots \wedge (F_{v;n-p,p} \wedge F_{0,n-p+1,p}))||_v \\
\ll & \prod_{0 \leq i \leq n-p+1} ||x^{\leq p-1} \cdot F_{v;i,p}||_v^{-\frac{n-p}{n-p+1}} \prod_{\substack{0 \leq i \leq n-p \\ i \neq m}} \frac{1}{A_{v;i}\rho_{v,i}} \\
\ll & \prod_{0 \leq i \leq n-p+1} ||x^{\leq p-1} \cdot F_{v;i,p}||_v^{-\frac{n-p}{n-p+1}} \cdot ||x^{\leq p-1}||_v^{n-p} \\
& \cdot \prod_{\substack{0 \leq i \leq n-p \\ i \neq m}} ||x^{\leq p-1} \cdot F_{v;i,p}||_v \cdot ||x^{\leq p-2} \cdot F_{v;\geq n-p+2}||_v \cdot \rho_{v,i}^{-1} \\
\ll & ||x^{\leq p-1}||_v^{n-p-1} \cdot \frac{\rho_{v,m}}{||x^{\leq p-1} \cdot F_{v;m,p}||_v} \\
& \cdot \prod_{0 \leq i \leq n-p+1} ||x^{\leq p-1} \cdot F_{v;i,p}||_v^{\frac{1}{n-p+1}} \cdot ||x^{\leq p-2} \cdot F_{v;\geq n-p+2}||_v^{n-p} \,,
\end{aligned}$$

for each $v \in S$, as $||x^{\leq p-1} \cdot F_{v;n-p+1,p}||_v \gg\ll ||x^{\leq p-1}||_v$ and $\prod_i \rho_{v,i} = 1$ hold for all $v \in S$. Here, the constants implicit in $\gg\ll$ are uniform in S. However, if we treat

[13] Scaling each v_j by a unit, we make the quantity

$$||v_j \cdot ((F_{v;0,p} \wedge F_{v;n-p+1,p}) \wedge \cdots \wedge (\widehat{F_{v;m,p} \wedge F_{v;n-p+1,p}}) \wedge \cdots \wedge (F_{v;n-p,p} \wedge F_{0,n-p+1,p}))||_v$$

having the magnitude satisfying

$$\ll \prod_{0 \leq i \leq n-p+1} ||x^{\leq p-1} \cdot F_{v;i,p}||_v^{-\frac{n-p}{n-p+1}} \prod_{\substack{0 \leq i \leq n-p \\ i \neq m}} \frac{1}{A_{v;i}\rho_{v,i}} \,.$$

Some error depending on k and S necessarily arises in this procedure. We choose the unit so that the error is minimum.

varying S, we have to take care of the extra error produced in the application of "scaling by unit" technique. On the other hand, $v_j \in \Lambda^{n-p}(x^{\leq p-1} \wedge (\mathcal{O}_{k,S}^{n+1} \wedge x^{\leq p-2}))$ can be written as

$$v_j = \sum_{l \in I_j} (x^{\leq p-1} \wedge (x^{\leq p-2} \wedge u_{l,1})) \wedge \cdots \wedge (x^{\leq p-1} \wedge (x^{\leq p-2} \wedge u_{l,n-p}))$$

with $u_{l,i} \in \mathcal{O}_{k,S}^{n+1}$ $(i = 1, \ldots, n-p)$. So, the left hand side of the above inequality is equivalent to

(**)

$$\begin{aligned}
&||v_j \cdot ((F_{v;0,p} \wedge F_{v;n-p+1,p}) \wedge \cdots \\
&\qquad \wedge (F_{v;m,p} \widehat{\wedge F_{v;n-p+1,p}}) \wedge \cdots \wedge (F_{v;n-p,p} \wedge F_{0,n-p+1,p}))||_v \\
=&||x^{\leq p-1}||_v^{n-p-1} || \sum_l x^{\leq p-1} \wedge (x^{\leq p-2} \wedge u_{l,1}) \wedge \cdots \wedge (x^{\leq p-2} \wedge u_{l,n-p}) \\
&\qquad \cdot (F_{v;0,p} \wedge \cdots \wedge \widehat{F_{v;m,p}} \wedge \cdots \wedge F_{v;n-p+1,p})||_v \\
=&||x^{\leq p-1}||_v^{n-p-1} ||x^{\leq p-2} \cdot F_{v;\geq n-p+2}||_v^{n-p} \\
&\qquad \cdot || \sum_l (x^{(p-1)} \wedge u_{l,1} \wedge \cdots \wedge u_{l,n-p}) \cdot (F_{v,0} \wedge \cdots \wedge \widehat{F_{v,m}} \wedge \cdots \wedge F_{v,n-p+1})||_v \,.
\end{aligned}$$

Note that

$$F_{v,m}^* = F_{v,0} \wedge \cdots \wedge \widehat{F_{v,m}} \wedge \cdots \wedge F_{v,n-p+1} \in \Lambda^{n-p+1}((k^{n+1})^*/\langle F_{v,n-p+2}, \ldots, F_{v,n}\rangle) \,.$$

As $n-p+1 = \dim((k^{n+1})^*/\langle F_{v,n-p+2}, \ldots, F_{v,n}\rangle)-1$, this space is identified with the dual of $(k^{n+1})^*/\langle F_{v,n-p+2}, \ldots, F_{v,n}\rangle$, according to the identification $\Lambda^{\dim V-1}V^* \cong V$ defined by $V \ni v \mapsto v \cdot \omega \in \Lambda^{\dim V-1}V^*$, where ω is a non-zero element of $\Lambda^{\dim V}V^*$. Therefore $F_{v,m}^*$ $(m = 0, \ldots, n-p+1)$ form a basis dual to $F_{v,0}, \ldots, F_{v,n-p+1}$ (a basis of $(k^{n+1})^*/\langle F_{v,n-p+2}, \ldots, F_{v,n}\rangle$). Define vectors $u_j \in \Lambda^{n-p}\mathcal{O}_{k,S}^{n+1}$ by

$$u_j := \sum_{l \in I_j} u_{l,1} \wedge \cdots \wedge u_{l,n-p} \,.$$

Then $x^{(p-1)} \wedge u_j$ $(j = 1, \ldots, n-p+1)$ form a basis of a full sublattice of $\{F \in \Lambda^p(\mathcal{O}_{k,S}^{n+1})^* \,;\, x^{(p-1)} \cdot F = 0\}$. Comparing (*) and (**), we have

$$\text{(5)} \qquad \begin{aligned}
&||(x^{(p-1)} \wedge u_j) \cdot (F_{v,0} \wedge \cdots \wedge \widehat{F_{v,m}} \wedge \cdots \wedge F_{v,n-p+1})||_v \\
&\ll \frac{\rho_{v,m}}{||x^{\leq p-1} \cdot F_{v;m,p}||_v} \prod_{0 \leq i \leq n-p+1} ||x^{\leq p-1} \cdot F_{v;i,p}||_v^{\frac{1}{n-p+1}}
\end{aligned}$$

for each $v \in S_\infty$ and

(6)

$$\begin{aligned}
&\prod_{v \in S} \prod_{m=0}^{n-p} \max_{1 \leq j \leq n-p+1} ||(x^{(p-1)} \wedge u_j) \cdot (F_{v,0} \wedge \cdots \wedge \widehat{F_{v,m}} \wedge \cdots \wedge F_{v,n-p+1})||_v \\
\ll &\prod_{v \in S} ||x^{\leq p-1} \cdot F_{v;n-p+1,p}||_v
\end{aligned}$$

both uniform in S. Here we used the property $\prod_{m=0}^{n-p}\rho_{v,m}=1$ (see [V, pp.97-98] for the choice of $\rho_{v,m}$ in the proof of Davenport's lemma). One might have a strange impression, because we take the product over the whole S in (6) while the estimate $(*)$ holds only for Archimedean places. This will disappear, if one recalls that in Davenport's lemma all contribution from non-Archimedean places transfered to S_∞ by the "gauge transformation" induced by multiplying a suitable S-unit.

On the other hand, under the identification

$$(k^{n+1})^*/\langle F_{v,n-p+2},\ldots,F_{v,n}\rangle \cong (k^{n+1})^*\wedge(F_{v,n-p+2}\wedge\cdots\wedge F_{v,n}),$$

we define

$$F_{v,k}\cdot x^{\leq p-1}=F_{v;k,p}\cdot x^{\leq p-1}.$$

As $x^{\leq p-2}\cdot F_{v;n-p+2}\neq 0$ and $x^{\leq p-1}\cdot F_{v;n-p+1,p}\neq 0$, both $\{F^*_{v,k}\}_{0\leq k\leq n-p}\cup\{x^{\leq p-1}\}$ and $\{F^*_{v,k}\}_{0\leq k\leq n-p+1}$ form bases of $(k^{n+1})^*/\langle F_{v,n-p+2},\ldots,F_{v,n}\rangle$. So, $F^*_{v,n-p+1}$ is written as

$$F^*_{v,n-p+1}=a_0x^{\leq p-1}+\sum_{k=0}^{n-p}F^*_{v,k}$$

with $a_0=\frac{1}{x^{\leq p-1}\cdot F_{v,n-p}}$ and $a_k=-\frac{x^{\leq p-1}\cdot F_{v,k}}{x^{\leq p-1}\cdot F_{v,n-p}}$ for $k=0,\ldots,n-p$. Therefore we have

$$F^*_{v,n-p+1}\equiv-\sum_{k=0}^{n-p}\frac{x^{\leq p-1}\cdot F_{v,k}}{x^{\leq p-1}\cdot F_{v,n-p}}F^*_{v,k}\pmod{x^{\leq p-1}}. \tag{7}$$

Plugging (7) into $||(x^{(p-1)}\wedge u_j)\cdot F^*_{v,n-p+1}||_v$ and using (5), we have

$$\begin{aligned}&||(x^{(p-1)}\wedge u_j)\cdot(F_{v,0}\wedge\cdots\wedge F_{v,n-p})||_v\\ \ll&\sum_{0\leq k\leq n-p}\frac{||x^{\leq p-1}\cdot F_{v,k}||_v}{||x^{\leq p-1}\cdot F_{v,n}||_v}||(x^{(p-1)}\wedge u_j)\cdot F^*_{v,k}||_v\\ \ll&\frac{1}{||x^{\leq p-1}\cdot F_{v,n}||_v}\cdot\max_{0\leq k\leq n-p}\rho_{v,k}\cdot\prod_{0\leq i\leq n-p+1}||x^{\leq p-1}\cdot F_{v;i,p}||_v^{\frac{1}{n-p+1}}\end{aligned}$$

uniform in S. Taking the product over $v\in S$ and using (4), we have

(8)

$$\begin{aligned}&\prod_{v\in S}\max_{1\leq j\leq n-p+1}||(x^{(p-1)}\wedge u_j)\cdot F^*_{v,n-p+1}||_v\\ \ll&\prod_{v\in S}\frac{1}{||x^{\leq p-1}\cdot F_{v,n}||_v}\cdot\max_{1\leq j\leq n-p+1}||(x^{(p-1)}\wedge u_j)\cdot F^*_{v,k}||_v\\ \ll&\prod_{v\in S}\frac{1}{||x^{\leq p-1}\cdot F_{v,n}||_v}\cdot\prod_{v\in S}\Big[\prod_{v\in S}\prod_{0\leq i\leq n-p+1}||x^{\leq p-1}\cdot F_{v;i,p}||_v\Big]^{-\frac{N_v}{(n-p+1)[k:\mathbb{Q}]}}\\ &\quad\cdot\prod_{v\in S}\exp(-\varepsilon\,\mathrm{ht}'(x^{\leq p-1}))^{\frac{N_v}{[k:\mathbb{Q}]}}\cdot\prod_{v\in S}\prod_{0\leq i\leq n-p+1}||x^{\leq p-1}\cdot F_{v;i,p}||_v^{\frac{1}{n-p+1}}\\ \ll&\frac{\exp(-\varepsilon\,\mathrm{ht}'(x^{\leq p-1}))}{\prod_{v\in S}||x^{\leq p-1}\cdot F_{v,n}||_v}\end{aligned}$$

uniform in S. Taking the product of (6) and (8) we have

$$(\dagger) \qquad \prod_{v\in S}\prod_{m=0}^{n-p+1} \max_{1\le j\le n-p+1} ||(x^{(p-1)}\wedge u_j)\cdot F^*_{v,m}||_v \ll \exp(-\varepsilon\,\mathrm{ht}'(x^{\le p-1}))$$

uniform in S. This yields the second condition of [V, Theorem 6.4,2].

Next we check that the first condition of [V, Theorem 6.4.2] also holds. Indeed, (4) implies that there exists a positive constant c' uniform in S such that $\rho_{v;i}$ $(0\le i\le n-p)$ satisfy

$$\rho_{v;i} \ll \begin{cases} 1 & \text{(if } v \text{ is non-Archimedean)} \\ \exp(c'\mathrm{ht}'(x^{\le p-1})) & \text{(if } v \text{ is Archimedean)} \end{cases}$$

uniform in S. Plugging this into (5) we see that there exists a positive constant c uniform in S such that

$$(\dagger\dagger) \qquad ||(x^{(p-1)}\wedge u_j)\cdot F^*_{v,m}|| \ll \exp(c\,\mathrm{ht}'(x^{\le p-1}))$$

holds for all $v\in S$, $j=1,\cdots,n-p+1$ and m $(0\le m\le n-p)$ which is uniform in S. This yields the first condition of [V, Theorem 6.4.2]. This completes the proof of Theorem 2.1. □

We suppose that $v\in S$ is non-Archimedean. Then we can get an upper bound for the intersection multiplicity of (the Zariski closure of) $x^{\le p-1}$ and $F_{v,i,p}=0$ over v in terms of $\exp(\mathrm{ht}'(x^{\le p-1}))$ in the following way. First we note

$$\prod_{w\in S} ||x^{\le p-1}\cdot F_{v,i,p}||_w \gg 1 .$$

Indeed, as $x^{\le p-1}$ is S-integral and $F_{v;i,p}$ is one of finitely many vectors with k-coefficients, the product theorem implies this with a lower bound uniform in S. Also we have

$$||x^{\le p-1}\cdot F_{v;i,p}||_w \ll ||x^{\le p-1}||_w ,$$

if $w\ne v$. It follows from these two inequalities that

$$\begin{aligned} 1 &\ll ||x^{\le p-1}\cdot F_{v;i,p}||_v \prod_{w\ne v} ||x^{\le p-1}\cdot F_{v;i,p}||_w \\ &\ll ||x^{\le p-1}\cdot F_{v;i,p}||_v \prod_{w\ne v} ||x^{\le p-1}||_w^{|S|-1} \\ &= ||x^{\le p-1}\cdot F_{v;i,p}||_v \frac{\exp((|S|-1)\mathrm{ht}'(x^{\le p-1}))}{||x^{\le p-1}||_v^{|S|-1}} . \end{aligned}$$

This implies

$$||x^{\le p-1}\cdot F_{v;i,p}||_v \gg \frac{||x^{\le p-1}||_v^{|S|-1}}{\exp((|S|-1)\mathrm{ht}'(x^{\le p-1}))} .$$

This provides an upper bound for the intersection multiplicity.

As will be shown in §4, Schmidt's Subspace Theorem is proved by the same pattern as Cartan's Second Main Theorem (i.e., plugging the "Lemma on logarithmic derivative" into the "Wronskian formalism"). In particular, Theorem 2.1 (the Diophantine analogue of the Lemma on logarithmic derivative) together with the Diophantine analogue of the Wronskian formalism yield Schmidt's Subspace Theorem. In order to introduce the truncated counting function into Schmidt's Subspace Theorem, houever, Theorem 2.1 turns out to be not enough.

In the rest of §2, we describe necessary modifications. Schmidt's Subspace Theorem is concerned with the Diophantine approximation of points $\{x\}$ to a divisor D consisting of hyperplanes in general position[14] in $\mathbb{P}^n(k)$. In the Nevanlinna theory, we have the concept of the residual counting function $N^n_{f,D}(r)$, which counts the ramification indices of the holomorphic curve $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ at places where f intersects D with multiplicity $\geq n$. This is intrinsically defined, based on the Wronskian $f' \wedge f'' \wedge \cdots \wedge f^{(n)}$. Our strategy is to establish the Diophantine analogue of the geometry behind the Nevanlinna theoretic truncated counting function and to introduce the Diophantine analogue of the truncated counting function into Schmidt's Subspace Theorem. Let D be a linear divisor (fixed once for all) in $\mathbb{P}^n(k)$ defined by linear forms F_i $(i = 0, 1, \ldots, n)$. For a rational point x and fixed D such that $x \notin \operatorname{Supp}(D)$, we consider the intersection of the Zariski closures of x and D in the arithmetic scheme $\mathbb{P}^n(\mathcal{O}_k)$ over $\operatorname{Spec}(\mathcal{O}_k)$ (i.e., consider $F_i(x)$ as a "function" on $\operatorname{Spec}(\mathcal{O}_k)$ and consider its zeros counted with multiplicities). Let S^n_x be the set of non-Archimedean places over which the Zariski closures of x and some component of D intersect with multiplicity $\geq n$. Then we consider, instead of a single rational point x, a pair (x, S^n_x) of x and the canonically associated finite set of non-Archimedean places S^n_x, and develop Diophantine approximation for these pairs (this framework is established once we fix a linear divisor D). This is our new framework in Diophantine approximation, which we imported from the Nevanlinna theory.

We are going to modify Theorem 2.1 so that the result of the same type holds for pairs (x, S^n_x). We here list the difference of Theorem 2.1 and its variant modified into the form useful in the proof of the Schmidt Subspace Theorem with residual counting function:

(i) In Theorem 2.1, the finite set of places S is fixed. However, in its modification, S should be $S(x) = S_\infty \cup S^n_x$, which does depend on x. In particular, the length functions should be defined as a sum over the places in $S(x)$.

(ii) We apply the geometry of numbers (the succssive minima) to the lattice of all algebraic integers $\mathcal{O}^{n+1}_k$. In other words, in Theorem 2.1, all x's were S-integral. However, in its modification, all x's should be S_∞-integral (i.e., the notion of the integrality should not depend on x).

In particular, (i) and (ii) imply that we apply the successive minima to the lattice $\mathcal{O}^{n+1}_k$ of ordinary algebraic integers, but the length function should be defined as a sum over all places in $S(x)$ (not only Archimedean places).

For the above purpose we examine whether the arguments in the proof of Theorem 2.1 remain true if we replace the fixed S by the varying $S(x) = S_\infty \cup S^n_x$ and try to apply the same strategy with respect to $S(x)$ instead of S. It turns out that there are three places in the proof of Theorem 2.1 which requires special care.

[14] We call such a divisor a linear divisor.

(i) *The first place.* The first is the place where Lemma 2.2 is applied to individual members of S to conclude (2). If S is varying, we have to take care of the error. As k is fixed, there is no problem on Archimedean places. The problem may happen only from the varying non-Archimedean part. However, since the F_i's are given linear forms, there exists the finite set S_F (determined by F_i's) of non-Archimedean places with the property that, in the proof of Theorem 2.1, non-trivial constants occur only when S_x^n touches S_F. Therefore, even if S varies, only the part of S which touches the fixed S_F may cause the problem. But, since S_F is finite and fixed, no problem occurs from non-Archimedean places, too.

(ii) *The second place.* The second is the place where the "scaling by unit" technique (based on the Dirichlet Unit Theorem) is used (e.g., the part where we applied Davenport's lemma)[15] and we should be careful about the error emerging in this procedure, too.

Scaling by unit has the effect making all factors in $\prod_{v\in S}\|\cdots\|_v$ having the desired (e.g., the same) order of magnitude. For instance, after making all factors having the same order of magnitude, we take the geometric mean of all factors (indexed by $S_\infty \cup S_x^n$) under consideration. In any case, this procedure necessarily produces error which depends only on the number field k and the finite set of non-Archimedean places under consideration (i.e., S_x^n)[16]. In general, $S(x) = S_\infty \cup S_x^n$ is varying. Let us estimate the magnitude of the error. The arguments in the proof of Theorem 2.1 are in the product form which is transformed into the sum form by taking the logarithm. Interchanging the arithmetic mean and the logarithm transforms the product form into the sum form with the "error" depending on S_x^n. However, in the proof of Theorem 2.1, "scaling by unit" technique was used only on the set S_∞ of Archimedean places of the given number field k. Therefore Dirichlet's Unit Theorem implies that the maximum of the absolute values of the "average" and the result of the "scaling by unit" over each $v \in S_\infty$ is bounded above by a constant depending only on the given k. After scaling by unit, we make all factors a_v having the same order of magnitude $A\times$(something uniform), A being the average. Therefore the error under consideration

$$\left|\frac{1}{[k:\mathbb{Q}]}\sum_{v\in S_\infty}\log a_v - \log\frac{1}{[k:\mathbb{Q}]}\sum_{v\in S_\infty} a_v\right|$$

is at most of order

$$\left|\frac{1}{[k:\mathbb{Q}]}[k:\mathbb{Q}](\log A + (\text{something uniform})) - \log(A\times(\text{something uniform}))\right|,$$

which is clearly uniform.

(iii) *The third place.* The third place is not explicit in the proof of Theorem 2.1. It concerns directly with the non-uniformity of the contribution from the non-Archimedean places. For instance, if we would like to get the same conclusion even when there is no control on $|S_x^n|$, we must control the error stemming from the places

[15] All parts in the proof of Theorem 2.1 where "scaling by unit" technique is used are so indicated in the corresponding footnotes.

[16] If S_x happens to be fixed (as in the situation of Theorem 2.1), then the error is of course uniform.

in S_x^n in the proof of Theorem 2.1. Otherwise we cannot get uniform estimates. We note that if the contribution from non-Archimedean places dominates that from the Archimedean places in the $S(x)$-proximity function, it will have a nature of the usual counting function restricted to places belonging to S_x^n. So we must control the behavior of the counting function. Now, let D be a linear divisor defined by the linear forms $F_0, \ldots, F_N$ in general position just as in Theorem 2.1. Let $x \in \mathbb{P}^n(k) - \mathrm{Supp}\,(D)$. We consider the association $x \mapsto S_x^n$ where S_x^n is the finite set of non-Archimedean places of k over which the Zariski closures (over Spec $(\mathcal{O}_k)$) of x and D intersect with multiplicity m greater than or equal to n. The condition $m \geq n$ imposed on S_x^n makes the arguments of the proof of Theorem 2.1 applicable to the situation of varying $S = S(x)$. If the non-Archimedean places dominates the Archimedean ones in the $S(x)$-proximity function, the inductive application of the first successive minimum indexed by $p\,(= 1, \ldots, n)$ yields the conclusion of Theorem 2.1 if and only if the non-Archimedean part in the inductive first successive minima behaves exactly as lattice theory predicts, and this is possible if and only if $m \geq n$.

Let us formulate this more carefully. To get uniform estimates, we have to control the contribution from the non-Archimedean places in S_x^n and for this purpose, we first need to modify the "weights" $A_{v;i,p}$ in the length function (3) used in the "inductive first successive minima" indexed by $p = 1, 2, \ldots, n$ in the proof of Theorem 2.1. The modification should be done respecting the "geometric effect" of the weights. We recall that the weights $A_{v;i,p}$ have a geometric effect on the "shape" of the symmetric star bodies defined by (3). Namely, if the weight in front of $L_{v;i,p}$ is "very large", this "collapses" to that consisting of the star bodies which are "widely spread" in the direction of the hyperplane defined by $F_{v;i} \cdot x^{(p)} = 0$ (for $p = 1, \ldots, n$). So, it is most probable that the first successive minimum associated to this system of star bodies will pick up such $x^{(p)}$ which is parallel to the hyperplane $F_{v,i} = 0$ to the extent determined by the degree of parallelness of $x^{(p-1)}$ to $F_{v;i} = 0$ in the v-adic sense (we define that degree of parallelness of $x^{(p-1)}$ to $F_{v;i} = 0$ is large, if $\mathrm{ord}_v(F_{v;i} \cdot x^{(p-1)})$ is large). In summary, if v is a non-Archimedean place, $A_{v;i,p}$ being very large implies $F_{v;i} \cdot x^{(p)}$ being divisible by accordingly high power of the prime v.

With this "geometric effect" understood, we modify the weights as follows. The general idea is that, at each step of the inductive first successive minima introduced in the proof of Theorem 2.1, we replace the usual counting functions involved in the definition of the weights by the residual counting functions at level 1 ($p = 1, \ldots, n$)[17]. For the first step (i.e., $p = 1$; the case considered in [V, Theorem 6.4.3]), we modify the weighs by replacing $||x \cdot F_{v;i,1}||_v$ by its "residual version" at level 1 ($v \in S(x) - S_\infty$)[18]. Here, by the residual version at level t of $||x \cdot F_{v;i,1}||_v$, we mean the version which is defined by replacing $\mathrm{ord}_v(x \cdot F_{v;i,1})$ in the definition of $||x \cdot F_{v;i,1}||_v$ by its level t residual version $\max\{\mathrm{ord}_v(x \cdot F_{v;i,1}) - t, 0\}$. In a similar way, we introduce this modification using the residual counting function at level 1 at every p-th step for $p = 1, \ldots, n$. Namely, we replace $||x^{\leq p-1} \cdot F_{v;i,p}||_v$ by its residual version at level 1, i.e., replace $\mathrm{ord}_v(x^{\leq p-1} \cdot F_{v;i,p})$ in the definition of $||x^{\leq p-1} \cdot F_{v;i,p}||_v$ by its level 1 residual version $\max\{\mathrm{ord}_v(x^{\leq p-1} \cdot F_{v;i,p}) - 1, 0\}$.

With this modification on the weights, we repeat the arguments of the proof of Theorem 2.1. The effect of the modification of weights is that we are able to get

[17] This modification changes only non-Archimedean weights.

[18] We call this the modification of type I.

uniform estimates even if we have no control on S_x^n.

We proceed as follows. For each non-Archimedean places v in $S(x)$, we construct the "ladder"

$$H_v^{(0)} \subset H_v^{(1)} \subset \cdots H_v^{(p)} \subset \cdots \subset H_v^{(n)} \quad (p = 0, 1, \ldots, n)$$

of sublattices of $x \wedge \mathcal{O}_k^{n+1}$ (the natural lattice structure of algebraic integers of $x \wedge k^{n+1}$) in the following way:

$$H_v^{(0)} = \{x \wedge x^{(0)} \in x \wedge \mathcal{O}_k^{n+1} \mid \min_{0 \le i < n} \{\mathrm{ord}_v(F_{v;i}(x^{(0)})) - \mathrm{ord}_v(F_{v;i}(x))\} \ge 0\} ,$$

$$H_v^{(1)} = \{x \wedge x^{(1)} \in x \wedge \mathcal{O}_k^{n+1} \mid \min_{0 \le i < n} \{\mathrm{ord}_v(F_{v;i}(x^{(1)})) - \mathrm{ord}_v(F_{v;i}(x))\} \ge -1\} ,$$

$$\cdots$$

$$H_v^{(t)} = \{x \wedge x^{(t)} \in x \wedge \mathcal{O}_k^{n+1} \mid \min_{0 \le i < n} \{\mathrm{ord}_v(F_{v;i}(x^{(t)})) - \mathrm{ord}_v(F_{v;i}(x))\} \ge -t\} ,$$

$$\cdots$$

$$H_v^{(n)} = \{x \wedge x^{(n)} \in x \wedge \mathcal{O}_k^{n+1} \mid \min_{0 \le i < n} \{\mathrm{ord}_v(F_{v;i}(x^{(n)})) - \mathrm{ord}_v(F_{v;i}(x))\} \ge -n\} .$$

Then we have

$$\mathrm{Vol}_v(H_v^{(p)}) = \mathrm{Vol}_v(H_v^{(p)}/H_v^{(p-1)})\mathrm{Vol}_v(H_v^{(p-1)}/H_v^{(p-2)}) \cdots \mathrm{Vol}_v(H_v^{(1)}/H_v^{(0)})\mathrm{Vol}_v(H^{(0)})$$

for each $v \in S(x) - S_\infty$.

On the other hand, we modify the usual length function on $x \wedge k^{n+1}$ so that the associated successive minima is compatible with the inductive first successive minima on the system of the modified length functions (3)[19]. Namely we modify the length function by

(9)

$$L(x \wedge x') = (x \wedge x') \cdot (F_{v;i} \wedge F_{v;n}), \quad (0 \le i < n) ;$$

$$A_{v;i}^{\text{modified}} = \begin{cases} A_{v;i} \quad (\text{if } v \text{ is Archimedean}) , \\ \text{the weight given by the level } n \text{ residual version of } \dfrac{1}{||x \cdot F_{v;i}||_v} \\ \quad (\text{if } v \text{ is non-Archimedean}) , \end{cases}$$

where $x \in \mathcal{O}_k^{n+1}$ is given and x' is an unknown vector also in $\mathcal{O}_k^{n+1}$. Under this modification, we can prove that the inductive first successive minima $\lambda_1^{(1)}, \ldots, \lambda_1^{(n)}$ with respect to the length function (3) with the above modified weights[20] are uniformly equivalent to the usual successive minima $\lambda_1, \ldots, \lambda_n$ with respect to the above modified usual length function (9)[21] on $x \wedge k^{n+1}$, in the sense that

$$\lambda_1^{(1)} \gg\ll \lambda_1, \ldots, \lambda_1^{(n)} \gg\ll \lambda_n$$

[19] we call this the modification of type II.

[20] Here, we use the modification of type I.

[21] Here we use the modification of type II.

holds. To prove this, we recall Bombieri-Vaaler's version [V, Theorem 6.1.11] of the Minkowski second theorem:

$$\begin{aligned}&(\lambda_1\lambda_2\cdots\lambda_n)^{[k:\mathbb{Q}]}\\ \gg\ll\ &\exp((n-1)\mathrm{ht}'(x))\prod_{v\in S(x)}\left(\det(F_{v;0},\ldots,F_{v;n-1})\prod_{0\le i<n}A_{v;i}^{\text{modified}}\right)\end{aligned}$$

where $A_{v;i}^{\text{modified}}$ is as introduced above, i.e., the level n residual version of the old weight $A_{v;i}$ (the modification of type II). The ladder $H_v^{(p)}$ of sublattices is defined only relative to $v\in S(x)-S_\infty$. However, for notational convenience, we introduce an "arbitrary flag" for each $v\in S_\infty$ (this has no substantial meaning). Under this notational convention, the right hand side of the above expression is

$$\begin{aligned}&\gg\ll\prod_{v\in S(x)}(\mathrm{Vol}_v(H_v^{(n)}))^{-1}\\ &=\Bigg(\prod_{v\in S(x)}\mathrm{Vol}_v(H_v^{(n)}/H_v^{(n-1)})\mathrm{Vol}_v(H_v^{(n-1)}/H_v^{(n-2)})\\ &\qquad\cdots\mathrm{Vol}_v(H_v^{(1)}/H_v^{(0)})\mathrm{Vol}_v(H_v^{(0)})\Bigg)^{-1},\end{aligned}$$

which implies that the vectors $\widetilde{x}^{(1)},\ldots,\widetilde{x}^{(n)}$ obtained by the successive minima associated to the length function (9) form a basis of $H_v^{(n)}$ (in particular $\widetilde{x}^{(p)}\in H_v^{(n)}$ for all $p=1,\ldots,n$) for each v, i.e., the part

$$\left(\prod_{v\in S(x)-S_\infty}\mathrm{Vol}_v(H_v^{(n)}/H_v^{(n-1)})\mathrm{Vol}_v(H_v^{(n-1)}/H_v^{(n-2)})\cdots\mathrm{Vol}_v(H_v^{(1)}/H_v^{(0)})\right)^{-1}$$

in the above expression is the v-power which the inverse v-volume of the parallelepiped generated by $\widetilde{x}^{(1)},\ldots,\widetilde{x}^{(n)}$ gains as an effect of the modification of the weights. On the other hand, we have

$$\left(\prod_{v\in S(x)}\mathrm{Vol}_v(\text{parallelepiped generated by }x^{(1)},\ldots,x^{(n)})\right)^{-1}\gg\ll(\lambda_1^{(1)}\cdots\lambda_1^{(n)})^{[k:\mathbb{Q}]}$$

and the choice of the weights in the modified inductive first successive minima implies that $x^{(1)},\ldots,x^{(n)}$ generates a sublattice of $H_v^{(n)}$ and therefore we have

$$\mathrm{Vol}_v(\text{parallelepiped generated by }x^{(1)},\ldots,x^{(n)})^{-1}\le\mathrm{Vol}(H_v^{(n)})^{-1}$$

for each v. This implies

$$\begin{aligned}(\lambda_1^{(1)}\cdots\lambda_1^{(n)})^{[k:\mathbb{Q}]}&\gg\ll\left(\prod_{v\in S(x)}\mathrm{Vol}_v(\text{parallelepiped generated by }x^{(1)},\ldots,x^{(n)})\right)^{-1}\\ &\le(\lambda_1\cdots\lambda_n)^{[k:\mathbb{Q}]}.\end{aligned}$$

The procedure of the successive minima $\{\lambda_p\}_{p=1}^n$ based on the length function (9) (using the level n residual version) chooses lattice points from wider possibilities compared to those $\{\lambda_1^{(p)}\}_{p=1}^n$ based on (3) modified (using the level 1 residual version at each step). We therefore have

$$\lambda_1^{(p)} \gg \lambda_p .$$

This implies that the inverse v-volume of the parallelepiped generated by $x^{(1)}, \ldots, x^{(n)}$ in fact gain at least the same amount of the v-powers as the inverse v-volume of the parallelepiped generated by $\widetilde{x}^{(1)}, \ldots, \widetilde{x}^{(n)}$ does after the modification $A_{v;i} \mapsto A_{v;i}^{\text{modified}}$. These two estimates imply the uniform equivalence of $\lambda_1^{(p)}$ and λ_p $(p = 1, \ldots, n)$. We therefore have

$$(\lambda_1^{(1)} \cdots \lambda_1^{(n)})^{[k:\mathbb{Q}]} \gg\ll (\lambda_1 \cdots \lambda_n)^{[k:\mathbb{Q}]} .$$

As the "ladder" of the v-divisibility of $\det(F_{v;0}, \ldots, F_{v;n-1})A_{v;i}^{\text{modified}}$ corresponds to the "ladder" of the sublattices $H_v^{(0)} \subset H_v^{(1)} \subset \cdots \subset H_v^{(n)}$ if $x \in H^{(0)}$, the above estimate and the volume formula forces (up to error uniform in $S(x)$) the vectors $x^{(1)}, \ldots, x^{(n)}$ of $\mathcal{O}_k^{n+1}$ in the modified inductive first successive minima to lie in the above defined "ladder" in the sense that

$$x^{(1)} \in H_v^{(1)}, \ldots, x^{(n)} \in H_v^{(n)} \tag{10}$$

holds, if $x \in H_v^{(0)}$ holds[22]. In particular, we must assume that m (the multiplicity of the intersection of of x and D over the place v) to be not smaller than n, for the above process to make sense.

We have thus proved that although the association $x \mapsto S_x^n$ is itself not uniform, the behavior of $x^{(1)}, \ldots, x^{(n)}$ are perfectly controlled in the sense of (10). In particular we have

$$N^n(x, F_i) \leq N_{S_\infty}(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0) - N_{S(x)}(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0)$$

where the counting functions in the right hand side measures the v-adic approximation of $x^{(1)} \wedge \cdots \wedge x^{(n)}$ to zero. This provides a kind of uniform estimates over the non-Archimedean places in S_x^n. Here, for finite set of places S including S_∞, the counting dunction N_S counts the approximation relative to the non-Archimedean places outside of S.

Finally, we need to show that (†) and (††) obtained in the modified situation violate the Parametric Subspace Theorem [V, Theorem 6.4.2]. In the Parametric Subspace Theorem, the set of places involved must be fixed. However, in the modified situation, we used the successive minima with respect to the length functions involving all places in $S(x)$. Therefore, all places in $S(x)$ are involved in (†) and (††). This means that, although we want to build the situation violating the Parametric Subspace Theorem, the inequalities in (†) and (††) in the modified situation are "disturbed" by the x-dependent non-Archimedean places in $S(x)$. However, we are done, if we can show that the inequalities in (†) and (††) reduce to those with respect to the places in S_∞ (this expectation is natural, because the intersection of all $S(x)$'s for various x's is just S_∞).

We show that Theorem 2.1 is modified under the assumption that S_x^n consists of non-Archimedean places over which the Zariski closure of x and some component of D intersect with multiplicity $\geq n$:

[22] The conclusion (10) may not always hold. However, the error stemming from "not always hold" is uniform in $S(x)$.

Theorem 2.3. *Let $F_0, \ldots, F_N$ be a set of linear forms in k^{n+1} in general position. Let $\varepsilon > 0$. Then there exists a finite set S of proper linear subspaces of k^{n+1} with the following property. If $x \in k^{n+1}$ is not a vector in the union of the linear subspaces in S, then we can inductively construct a sequence $x^{(1)}, \ldots, x^{(n)} \in \mathcal{O}_k^{n+1}$ of vectors with the following properties:*
(i) $x, x^{(1)}, \ldots, x^{(n)}$ are linearly independent:

$$x \wedge x^{(1)} \wedge \cdots \wedge x^{(n)} \neq 0 \; .$$

(ii) $\mathrm{ord}_v(x^{(t)} \cdot F_i)$ decreases 1 as t increases 1, i.e., if $\mathrm{ord}_v(x \cdot F_i) \geq n$, we have

$$\mathrm{ord}_v(x^{(t)} \cdot F_i) = \mathrm{ord}_v(x \cdot F_i) - t$$

for $t = 1, 2, \ldots, n$.
(iii) If we set $x^{\leq p-1} = x \wedge x^{(1)} \wedge \cdots \wedge x^{(p-1)}$ and $F_{i,p} = F_i \wedge F_{n-p+2} \wedge \cdots \wedge F_n$ for $p = 1, \ldots, n$, we have the following inequality: after suitably re-ordering the F's, we have

$$\sum_{v \in S(x)} \log \frac{||(x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \cdot F_{i,p}||_v}{||x^{\leq p-1}||_v ||x^{\leq p-1} \cdot F_{i,p}||_v^{\mathrm{modified}}} < \varepsilon \, \mathrm{ht}'(x)$$

for all $i = 0, \ldots, N$ and for all x such that $x^{\leq p-1} \cdot F_{i,p} \neq 0$. If $x^{\leq p-1} \cdot F_{i,p} = 0$ then $(x^{\leq p-2} \wedge x^{(p)}) \cdot F_{i,p} = 0$. Here, $||x^{\leq p-1} \cdot F_{i,p}||_v^{\mathrm{modified}}$ means that if $v \in S_x^n$ we replace $\mathrm{ord}_v(x^{\leq p-1} \cdot F_{i,p})$ in the original definition by its level 1 residual version $\max\{\mathrm{ord}_v(x^{\leq p-1} \cdot F_{i,p}) - 1, 0\}$, and if $v \in S_\infty$, we need no modification.

Proof. The conclusions (i) and (ii) follow from the above arguments after the proof of Theorem 2.1. To show (iii), we need to establish its relation to the Parametric Subspace Theorem [V, Theorem 6.4.2]. However, (ii) implies that the non-Archimedean places in $S(x)$ have no contribution to the left hand side of the inequality in (iii). This implies that the sum over the $S(x)$ reduces to the sum over S_∞. Therefore, if (iii) does not hold for infinite number of $x^{\leq p-1}$, we arrive at the situation which violates the Parametric Subspace Theorem [V, Theorem 6.4.2] (on the lattice $\mathcal{O}_k^{n+1}$ with S_∞ as the fixed set places). $\square$

3. Nevanlinna-Cartan theory.

In this section we describe the geometry (Nevanlinna-Cartan theory) which connects the Lemma on logarithmic derivative to the approximation inequality of holomorphic curves to a linear divisor in $\mathbb{P}^n(\mathbb{C})$. Yamanoi [Y] was able to characterize the nature of this geometry in general setting: Let X be a complex smooth projective variety and D any effective divisor. Then the existence of a holomorphic map W from a jet space $X^{(j)}$ to a certain line bundle $L \to X$ (S_0 being its zero-section) satisfying the condition $D^{(j)} \subset W^*(S_0)$ (in the scheme theoretical sense) characterizes the geometry behind Nevanlinna-Cartan theory. The holomorphic map $W : X^{(j)} \to L$ and the condition $D^{(j)} \subset W^*(S_0)$ are respectively the abstract version of the Wronskian and the "linearity condition" in the original Nevanlinna-Cartan theory.

Let's return to our situation. Let D be a linear divisor of $\mathbb{P}^n(\mathbb{C})$ defined by the linear forms $F_0, \dots, F_N$ in general position. Let $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ be a holomorphic curve such that $f(\mathbb{C}) \not\subset \operatorname{Supp}(D)$. Then Nevanlinna's lemma on logarithmic derivative (Theorem 1.1) states that

$$(9) \qquad \begin{cases} m_{f,D}(r) \leq m_{f^{(k)},D^{(k)}}(r) + S_f(r)/\!/ \\ m_{f^{(k)},\infty}(r) \leq S_f(r)/\!/ \end{cases}$$

hold for any nonnegative integer k. Here, $f^{(k)} : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})^{(k)}$ is the k-th jet lift of f and

$$D^{(k)} = \cup_{j=0}^{N} V(F_j, dF_j, \dots, d^k F_j)$$

is the union of the k-th jet space of individual hyperplanes defined by $F_j = 0$. Note that we are insisting on the linearity and use the individual defining equation $F_j = 0$ instead of the product $F_0 \cdots F_N = 0$ to define the jet space $D^{(k)}$. Theorem 1.1 still holds in this situation (see, for instance, [K1,2]).

We choose $(\zeta_0 : \zeta_1 : \dots : \zeta_n)$ a system of homogeneous coordinates of $\mathbb{P}^n(\mathbb{C})$. Then $z_i = \zeta_i/\zeta_0$ form a system of affine coordinates of $\mathbb{C}^n = \mathbb{P}^n(\mathbb{C}) - \{z_0 = 0\}$. The Wronskian $\det(d^j z_i)_{i,j=1}^n$ of affine coordinates $z_1, \dots, z_n$ defines a holomorphic map

$$W : \mathbb{P}^n(\mathbb{C})^{(n)} \longrightarrow K^{-1}_{\mathbb{P}^n(\mathbb{C})}.$$

Set $f_i(z) = z_i \circ f(z)$. Then $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ is expressed as $f(z) = (f_1(z), \dots, f_n(z))$ in terms of the above affine coordinate system $(z_1, \dots, z_n)$. Using jet lifts $f^{(j)}$ $(j = 1, \dots, n)$ of f, we can make up a holomorphic curve

$$W_f := W \circ f : \mathbb{C} \longrightarrow K^{-1}_{\mathbb{P}^n(\mathbb{C})} ; \quad \mathbb{C} \ni z \mapsto W_f(z) = \det(d^j f_i(z))_{i,j=1}^n \in K^{-1}_{\mathbb{P}^n(\mathbb{C})} .$$

The Lemma on logarithmic derivative (9) implies that, up to error of order $S_f(r) = O(\log^+(rT_f(r)))$, a holomorphic curve $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ approximates D if and only if $f^{(j)} : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})^{(j)}$ approximates $D^{(j)}$ for any nonnegative integer j. On the other hand, as the Wronskian is defined in terms of linear coordinates $z_1, \dots, z_n$, the n-th jet space of any hyperplane is sent to the zero-section of $K^{-1}_{\mathbb{P}^n(\mathbb{C})}$ via the Wronskian map $W : \mathbb{P}^n(\mathbb{C})^{(n)} \to K_{\mathbb{P}^n(\mathbb{C})}$. Let D_j be a hyperplane defined by $F_j = 0$ and S_0 (resp. S_∞) the 0-section (resp. ∞-section) of the anticanonical bundle $K^{-1}_{\mathbb{P}^n(\mathbb{C})} \to \mathbb{P}^n(\mathbb{C})$. Then we have the "linearity condition"

$$D^{(n)} \subset W^*(S_0)$$

in the scheme theoretical sense. Combining this with the first inequality in (9) gives

$$m_{f,D}(r) \le m_{W_f,S_0}(r) + S_f(r)/\!/ .$$

On the other hand, there is a linear equivalence

$$S_0 + \pi^* K_{\mathbb{P}^n(\mathbb{C})} = S_\infty \quad (\pi \text{ being the projection } K^{-1}_{\mathbb{P}^n(\mathbb{C})} \to \mathbb{P}^n(\mathbb{C})) .$$

which was implicitly used by Nevanlinna [N] in the case of $n = 1$ and by Cartan [C] in the case of general n. The above inequality together with this linear equivalence yield the approximation inequality (Cartan's Second Main Theorem) in the following way:

$$\begin{aligned}
&m_{f,D}(r) + N_{W_f,S_0}(r) + T_{f,K_{\mathbb{P}^n(\mathbb{C})}}(r)/\!/ \\
\le& m_{W_f,S_0}(r) + N_{W_f,S_0}(r) + T_{f,K_{\mathbb{P}^n(\mathbb{C})}}(r) + S_f(r)/\!/ \\
\le& T_{W_f,S_0+\pi^* K_{\mathbb{P}^n(\mathbb{C})}}(r) + S_f(r)/\!/ \\
\le& T_{W_f,S_\infty}(r) + S_f(r)/\!/ .
\end{aligned}$$

We now use the second inequality in (9) to conclude the approximation inequality

$$m_{f,D}(r) + N_{W_f,S_0}(r) \le (n+1)T_f(r) + S_f(r)/\!/$$

where $T_f(r)$ is the height function relative to the hyperplane bundle $\mathcal{O}_{\mathbb{P}^n(\mathbb{C})}(1)$.

Next we look at only such "non-Archimedean" places in $\mathbb{C}(r) = \{z \in \mathbb{C}\,;\ |z| < r\}$ such that $F_j(f(z)) = 0$ for some j. Let us fix one particular j and suppose that $F_j(f(z)) = 0$. As F_j is linear, there exists an affine coordinate system $w_1, \cdots, w_n$ such that $F_j = 0$ is equivalent to $w_n = 0$. If z is a multiple root of $F_j(f(z)) = 0$ with multiplicity m, then

$$d^j(w_n \circ f) = O(z^{m-j})$$

for $j \le m-1$. If $m > n$, then the Wronskian matrix of $(w_1 \circ f, \ldots, w_n \circ f)$ is of the form

$$\begin{pmatrix}
d(w_1 \circ f) & d(w_2 \circ f) & \cdots & d(w_{n-1} \circ f) & O(z^{m-1}) \\
d^2(w_1 \circ f) & d^2(w_2 \circ f) & \cdots & d^2(w_{n-1} \circ f) & O(z^{m-2}) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
d^{n-1}(w_1 \circ f) & d^{n-1}(w_2 \circ f) & \cdots & d^{n-1}(w_{n-1} \circ f) & O(z^{m-n+1}) \\
d^n(w_1 \circ f) & d^n(w_2 \circ f) & \cdots & d^n(w_{n-1} \circ f) & O(z^{m-n})
\end{pmatrix} .$$

This implies that if $m > n$ then the intersection multiplicity of W_f and S_0 dominates the intersection multiplicity of f and D (at z) minus n. Let $N_{n,f,D_j}(r)$ denote the truncated counting function at level n. We introduce the residual counting function $N^n_{f,D_j}(r)$ at level n by

$$\begin{aligned}
N^n_{f,D_j}(r) &= N_{f,D_j}(r) - N_{n,f,D_j}(r) \\
&= \sum_{0<|a|<r} \max\{\mathrm{ord}_a(F_j \circ f) - n, 0\} \log\left|\frac{r}{a}\right| \\
&\quad + \max\{\mathrm{ord}_0(F_j \circ f) - n, 0\} \log r .
\end{aligned}$$

Then the above observation implies

$$N^n_{f,D}(r) \leq N_{W_f,S_0}(r) .$$

Therefore we have Cartan's Second Main Theorem for the approximation to linear divisors of holomorphic curves into $\mathbb{P}^n(\mathbb{C})$ with residual counting function at level n:

$$m_{f,D}(r) + N^n_{f,D}(r) \leq (n+1)T_f(r) + S_f(r) /\!/ .$$

Of course the above argument makes sense only if the Wronskian W_f is not identically zero. On the other hand, $W_f \equiv 0$ if and only if $f(\mathbb{C})$ is contained in a proper linear subspace. By induction on the dimension of the linear closure of f, we infer that there exists a union Z of finitely many proper subspaces such that if $f(\mathbb{C})$ is not contained in Z the above inequality holds (with n replaced by the dimension of the linear closure of f).

4. A Diophantine analogue of Nevanlinna-Cartan theory.

In §2 (in the proof of Theorem 2.3) we established the Diophantine analogue $S(x) = S_\infty \cup S^n_x$ of the notion of the "Nevanlinna theoretic non-Archimedean places" (depending on holomorphic curve and divisor under consideration). Because of its importance, we here repeat the construction. By taking the Zariski closure over $\operatorname{Spec}(\mathcal{O}_k)$, a rational point x of $\mathbb{P}^n(k)$ is identified with a section of $\mathbb{P}^n(\mathcal{O}_k) \to \operatorname{Spec}(\mathcal{O}_k)$. Therefore, in this setting, a rational point (resp. a divisor) of $\mathbb{P}^n(k)$ becomes a curve (resp. a divisor) of the arithmetic scheme $\mathbb{P}^n(\mathcal{O}_k)$. The intersection of x and a linear divisor D in $\mathbb{P}^n(k)$ is thus defined. Therefore for each rational point x not contained in $\operatorname{Supp}(D)$, we can define the finite set S^n_x of non-Archimedean places over which x and D intersect with multiplicity $\geq n$. We consider this S^n_x as the Diophantine analogue of the "Nevanlinna theoretic non-Archimedean places".

In this section, we combine Theorem 2.3 and the Wronskian formalism in §3. This is to compare infinite set of rational points of $\mathbb{P}^n(k)$ and a holomorphic curve in $\mathbb{P}^n(\mathbb{C})$ from a view point of their intersection with a given linear divisor. We have already established in §2 two versions of the Diophantine analogue of the Lemma on logarithmic derivative, which culminated in Theorem 2.1 and 2.3. Here, we first prove the Schmidt Subspace Theorem by combining Theorem 2.1 and the Wronskian technique (in §3). We regard this proof as the simplest "model case" which should be suitably modified (although the necessary modification will be far from simple) according to the nature of problems we want to solve.

Let $\overline{v}$ denote the image in $T_{[x]}\mathbb{P}^n(k)$ of $v \in k^{n+1}$ under the identification $k^{n+1} \wedge x \cong k^{n+1}/\langle x \rangle \cong T_{[x]}\mathbb{P}^n(k) \otimes \mathcal{O}(-1)_{[x]}$ (see the discussion just before Theorem 2.1).

Theorem 4.1. *Let D be a linear divisor of $\mathbb{P}^n(k)$ in general position. Let $D^{(p)}$ denote the union of the p-th jet space of all irreducible components of D. Let S be a fixed finite set of places of k containing all Archimedean ones. Then there exists a finite union $\mathcal{S}$ of proper linear subspaces of $\mathbb{P}^n(k)$ such that if $x \notin \mathcal{S}$ then there exist $\overline{x}^{(1)}, \ldots, \overline{x}^{(n)} \in T_{[x]}\mathbb{P}^n(\mathcal{O}_{k,S})$ such that*

$$\begin{cases} m_S(x, D) \leq m_S(\overline{x}^{(p)}, D^{(p)}) + \varepsilon \operatorname{ht}(x) \\ m_S(\overline{x}^{(p)}, \infty) \leq \varepsilon \operatorname{ht}(x) \end{cases}$$

for all $p = 1, \ldots, n$, where the S-proximity functions $m_S(\cdot,\cdot)$ are defined as in §2 using the v-adic distances and ∞ represents the divisor at infinity of the projective completion of $T\mathbb{P}^n(k)$.

Proof. Let $x^{(1)}, \ldots, x^{(n)}$ be as in Theorem 2.3 and let $\overline{x}^{(1)}, \ldots, \overline{x}^{(n)}$ be the image in $T_{[x]}\mathbb{P}^n(k)$ under the above identification. We show that these $\overline{x}^{(j)}$'s satisfy the desired inequalities. In the proof of Lemma 2.2 we have shown that the assumption

$$|(x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \cdot F_{i,p}| > A\,|x^{\leq p-1}||x^{\leq p-1} \cdot F_{i,p}| \quad \exists i$$

is equivalent to

$$\max\left\{ \frac{|(x^{\leq p-2} \wedge x^{(p)}) \cdot F_{i,p}|}{|x^{\leq p-1} \cdot F_{i,p}|}, \frac{|x^{\leq p-2} \wedge x^{(p)}|}{|x^{\leq p-1}|} \right\} \gg A\,.$$

On the other hand, $T_{[x]}\mathbb{P}^n(k)$ is identified with $k^{n+1} \wedge x$. Therefore, if $p = 1$ in the above inequality, the quantity $\frac{|x^{\leq p-2} \wedge x^{(p)}|}{|x^{\leq p-1}|}$ taken the product over $v \in S$ is equivalent to the proximity function $m(\overline{x}^{(1)}, \infty)$ to ∞ under the above identification. On the other hand, the quantity $\frac{|(x^{\leq p-2} \wedge x^{(p)}) \cdot F_{i,p}|}{|x^{\leq p-1} \cdot F_{i,p}|}$ taken the similar product over S is equivalent to the proximity function to the first jet space $D^{(1)}$ of the linear divisor D. By induction on p, we get the desired inequalities. At each step indexed by $p = 1, \cdots, n$, there arises a finite union of exceptional points (in wedge products). At the first step there arises a finite exception for x. In the second step, there arises again a finite exception for $x \wedge x^{(1)}$. In terms of x as a vector in k^{n+1}, it will be a finite set of proper linear subspaces of dimension ≤ 1. At the p-th step, there arises a finite exception for $x^{(p-1)}$. In terms of x as a vector in k^{n+1}, it will be a finite set of proper linear subspaces of dimension $\leq p-1$. The statement on the exceptional subspaces $\mathcal{S}$ is a consequence of this observation. □

We can view the totality of $x, x^{(1)}, \ldots, x^{(n)}$ as a point of the n-th jet space $\mathbb{P}^n(k)^{(n)}$ of $\mathbb{P}_n(k)$. The Wronskian

$$W : \mathbb{P}^n(k)^{(n)} \longrightarrow K^{-1}_{\mathbb{P}^n(k)}$$

is a morphism. The image of $(x, x^{(1)}, \ldots, x^{(n)})$ under the Wronskian W coincides with $x^{(1)} \wedge \cdots \wedge x^{(n)}$ and the Wronskian morphism W sends $D^{(n)}$ to S_0 (the zero-section of $K^{-1}_{\mathbb{P}^n(k)}$). Moreover we have a linear equivalence

$$S_0 + \pi^* K_{\mathbb{P}^n(k)} = S_\infty$$

where S_∞ represents the divisor at infinity of the projective completion of $K^{-1}_{\mathbb{P}^n(k)}$. Therefore we have

$$\begin{aligned}
&m(x, D) + N(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0) + h(x, K_{\mathbb{P}^n(k)}) \\
\leq\, & m(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0) + N(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0) + h(x, K_{\mathbb{P}^n(k)}) + \varepsilon\, \mathrm{ht}\,(x) \\
&\qquad \text{[by the first inequality in Theorem 4.1]} \\
\leq\, & h(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0 + \pi^* K_{\mathbb{P}^n(k)}) + \varepsilon\, \mathrm{ht}\,(x) \\
\leq\, & h(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_\infty) + \varepsilon\, \mathrm{ht}\,(x) \\
&\qquad \text{[by the above linear equivalence]} \\
\leq\, & \varepsilon\, \mathrm{ht}\,(x) \\
&\qquad \text{[by the second inequality in Theorem 4.1]}\,.
\end{aligned}$$

This is the Schmidt Subspace Theorem. We remark here that the counting function $N(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0)$ essentially has no contribution (as we see from the definition of the successive minima).

Next, we introduce the residual counting function to the Schmidt Subspace Theorem by the same argument as the model case with only exception that Theorem 2.1 is replaced by Theorem 2.3. The point here is to execute the successive minima with respect to the point (x)-dependent finite set $S(x) = S_\infty \cup S_x^n$ of places instead of the fixed S. By the above procedure, we are able to show that these $\overline{x}^{(j)}$'s behave, over the non-Archimedean places in S_x^n, just like the derivatives $f^{(1)}(z), \ldots, f^{(n)}(z)$ of the holomorphic curve f do when f intersects D with multiplicity $m \geq n$.

Theorem 2.3 together with the corresponding Lemma 2.2 imply the following:

Theorem 4.2. *Let D be a linear divisor of $\mathbb{P}^n(k)$ in general position. Let $D^{(p)}$ denote the union of the p-th jet space of all irreducible components of D. Then there exists a finite union S of proper linear subspaces of $\mathbb{P}^n(k)$ such that, if $x \notin S$, then there exist $\overline{x}^{(1)}, \ldots, \overline{x}^{(n)} \in T_{[x]}\mathbb{P}^n(\mathcal{O}_k)$ which satisfy the inequalities*

$$m_{S_\infty}(x, D) \leq m_{S_\infty}(\overline{x}^{(p)}, D^{(p)}) + \varepsilon \operatorname{ht}(x) \tag{11}$$

$$m_{S_\infty}(\overline{x}^{(p)}, \infty) \leq \varepsilon \operatorname{ht}(x) \tag{12}$$

and the condition

$$x^{(p)} \in H_v^{(p)} \quad \forall v \in S_x \tag{13}$$

for all $p = 1, 2, \ldots, n$. Here, $S(x)$ is the finite set of places of k defined by $S(x) = S_\infty \cup S_x^n$ where S_x^n is the set of non-Archimedean places of k over which the section $x : \operatorname{Spec}(\mathcal{O}_k) \to \mathbb{P}^n(\mathcal{O}_k)$ and the linear divisor D in $\mathbb{P}^n(\mathcal{O}_k)$ intersect with multiplicity $m \geq n$.

The condition (13) implies

$$N^n(x, D) \leq N_{S_\infty}(x, D)(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0) - N_{S(x)}(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0)\ . \tag{14}$$

This inequality exactly plays the role of the Diophantine analogue of the Wronskian formalism in Nevanlinna theory in the proof of the Main Theorem. The point here is that Theorem 4.2 (i.e., the Diophantine analogue of the Lemma on logarithmic derivative for varying $S(x)$) splits into two statements: One is (11) and (12) for the Archimedean places which has a typical form of the Lemma on logarithmic derivative, and the other is (13) (or (14)) for the non-Archimedean places in S_x^n which is the Diophantine analogue of the Wronskian formalism.

To show that (13) implies (14), we suppose that one of the components of D and x intersects over $v \in S_x^n$ with multiplicity m and assume that $m \geq n$. Then, by the theory of successive minima relative to $S(x)$, the order of the v-divisibility of the "Wronskian" $x^{(1)} \wedge \cdots \wedge x^{(n)}$ at v is by definition that of the determinant of the "Wronskian matrix"

$$\begin{pmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_{n-1}^{(1)} & x_n^{(1)} \\ x_1^{(2)} & x_1^{(2)} & \cdots & x_{n-1}^{(2)} & x_n^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1^{(n-1)} & x_2^{(n-1)} & \cdots & x_{n-1}^{(n-1)} & x_n^{(n-1)} \\ x_1^{(n)} & x_2^{(n)} & \cdots & x_{n-1}^{(n)} & x_n^{(n)} \end{pmatrix}.$$

It follows from the proof of Theorem 2.3 (arguments before it) that we may assume that

$$\mathrm{ord}_v(x_n^{(j)}) = m - j.$$

This is the Diophantine analogue of $d^j(w_n \circ f) = O(z^{m-j})$ in the Wronskian matrix for holomorphic curves into $\mathbb{P}^n(\mathbb{C})$ discussed in §3. We thus have the Diophantine analogue (14) of the Wronskian formalism in Nevanlinna theory.

We are now ready to execute the same procedure as in the Schmidt Subspace Theorem with Theorem 4.1 replaced by Theorem 4.2. Applying the Diophantine analogue of the Wronskian formalism (14), we have:

$$\begin{aligned}
& m_{S_\infty}(x, D) + N^n_{S_\infty}(x, D) + h(x, K_{\mathbb{P}^n(k)}) \\
\leq & m_{S_\infty}(x, D) + N_{S_\infty}(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0) - N_{S(x)}(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0) \\
& \quad + h(x, K_{\mathbb{P}^n(k)}) + \varepsilon\, \mathrm{ht}'(x)\ .
\end{aligned}$$

Next, we proceed applying the Diophantine analogue ((11) and (12) in Theorem 4.2) of the Lemma on logarithmic derivative:

$$\begin{aligned}
& m_{S_\infty}(x, D) + N^n(x, D) + h(x, K_{\mathbb{P}^n(k)}) \\
\leq & m_{S_\infty}(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0) + N_{S_\infty}(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0) \\
& \quad - N_{S(x)}(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0) + h(x, K_{\mathbb{P}^n(k)}) + \varepsilon\, \mathrm{ht}\,(x) \\
& \qquad [\text{by the inequality (11) in Theorem 4.2}] \\
\leq & h(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_0 + \pi^* K_{\mathbb{P}^n(k)}) + \varepsilon\, \mathrm{ht}\,(x) \\
\leq & h(x^{(1)} \wedge \cdots \wedge x^{(n)}, S_\infty) + \varepsilon\, \mathrm{ht}\,(x) \\
& \qquad [\text{by the linear equivalence } S_0 + \pi^* K_{\mathbb{P}^n(k)} = S_\infty] \\
\leq & \varepsilon\, \mathrm{ht}\,(x) \\
& \qquad [\text{by the inequality (12) in Theorem 4.2}]\ .
\end{aligned}$$

We have thus get the Main Theorem stated in the introduction:

Theorem 4.3 (Schmidt Subspace Theorem with Residual Counting Function). *Let $F = \{F_i\}_{i=0}^N$ be a set of linear forms in $\mathbb{P}^n(k)$ in general position. Let $\varepsilon > 0$. Then there exists a finite union of linear subspaces $E(F, \varepsilon)$ and a constant $C(F, \varepsilon)$ such that for all $x \in \mathbb{P}^n(k) \backslash E(F, \varepsilon)$ the approximation inequality*

$$\sum_{i=0}^{N} m(x, F_i) + \sum_{i=0}^{N} N^n(x, F_i) \leq (n + 1 + \varepsilon)\, \mathrm{ht}(x) + C(F, \varepsilon)$$

holds.

5. Conjectures toward effectiveness.

In Theorem 2.3, we have shown the following: For $x \in \mathcal{O}_k^{n+1}$, possibly with exceptions consisting of finitely many proper linear subspaces of dimension $\leq p-1$, we can inductively construct a sequence $x^{(1)}, \ldots, x^{(p)}$ in $\mathcal{O}_k^{n+1}$ such that the inequality

$$\sum_{v \in S(x)} \log \frac{||(x^{\leq p-1} \wedge (x^{\leq p-2} \wedge x^{(p)})) \cdot F_{i,p}||_v}{||x^{(p)}||_v ||x^{\leq p-1} \cdot F_{i,p}||_v^{\text{modified}}} < \varepsilon \, \mathrm{ht}'(x) \tag{15}$$

holds. In the proof of Theorem 2.3, we have shown that this existence theorem (for $x^{(p)}$'s) is nearly equivalent to the Parametric Subspace Theorem. On the other hand, the only known proof of the Parametric Subspace Theorem is the proof by contradiction via the application of the Roth lemma. This proof does not give any information on the bound of the height of the exceptional proper linear subspaces.

However, as we have shown in the proof of Theorem 2.3 and Theorem 4.2, the inequality of type (15) plays the role of the Diophantine analogue of the Lemma on logarithmic derivative and the Wronskian formalism in Nevanlinna theory. In fact, this splits into two statements one of which is on the Archimedean places and the other on the non-Archimedean places, respecting the non-uniform non-Archimedean places involved in the argument. This enabled us to have uniform estimates in Theorem 4.2. The former one is

$$\begin{cases} m_{S_\infty}(x, D) \leq m_{S_\infty}(\overline{x}^{(p)}, D^{(p)}) + \varepsilon \, \mathrm{ht}\,(x) \\ m_{S_\infty}(\overline{x}^{(p)}, \infty) \leq \varepsilon \, \mathrm{ht}\,(x) \end{cases} \tag{16}$$

which is strongly analogous to the Lemma on logarithmic derivative (Theorem 1.1) in Nevanlinna theory. We then consider x's which do not obey the system of inequalities (16) (for any choice of $x^{(p)}$). Ths advantage of doing so lies on the expectation that bounding such x's would be significantly simpler. The reason is that the condition

(17)
$$m_{S_\infty}(x, D) > m_{S_\infty}(\overline{x}^{(p)}, D^{(p)}) + \varepsilon \, \mathrm{ht}\,(x) \quad \text{or} \quad m_{S_\infty}(\overline{x}^{(p)}, \infty) > \varepsilon \, \mathrm{ht}\,(x)$$
$$\forall \, \overline{x}^{(p)} \in T_{[x]}\mathbb{P}^n(\mathcal{O}_k)$$

would be significantly easier to handle compared to (15) (which is almost equivalent to the Parametric Subspace Theorem).

The structure of the analogy between Schmidt's Subspace Theorem and the Nevanlinna-Cartan Theory is logically complicated. In particular, we must first prove the ineffective Schmidt's Subspace Theorem and then introduce a geometric idea for the consideration on the effectiveness. To explain this complexity, we introduce two kinds of Diophantine analogues of Lemma on logarithmic derivative in Nevanlinna theory. One is that of Type A and the other is that of Type B. The analogue of Type A is the inequality in Theorem 2.3. That of Type B consists of the system of the inequalities in Theorem 4.2 (together with the condition (13)). We have no direct proof for the Type B analogue. namely, at this stage, we can prove the Type B analogue only "via proving that of Type A". We proved the

Type A analogue by reducing it to the Parametric Subspace Theorem [V, Theorem 6.4.2]. The essential part is to *define* the Diophantine analogue of the derivatives by applying Minkowski's geometry of numbers (successive minima) with certain length functions (the Type A analogue is based on the length function to which the successive minima is applied to define $x', \dots, x^{(n)}$ and this is the reason why we cannot avoid the Type A analogue). On the other hand, the Parametric Subspace Theorem reduces to the Roth lemma[23] and therefore the result is not effective. However, the Type A analogue implies the Type B analogue together with the ineffective finiteness statement. Then the Type B analogue turns out to split into the inequalities over S_∞ and the conditions (13) over the non-Archimedean places in $S(x)$.

Our **conjecture** is that *we are able to effectively bound the solutions of these inequalities over* S_∞.

This discussion is summarized in the following table, where LLD (resp. WF, PSST and SST) is the abbreviation of Lemma on logarithmic derivative (resp. Wronsakian formalism, Parametric Schmidt's Subspace Theorem and Schmidt's Subspace Theorem). Type A stands for the Type A Diophantine analogue of LLD (Theorem 2.1) and Type B does theType B Diophantine analogue of LLD (Theorem 4.2). Note that there are implications PSST $\Rightarrow$ Type A (Theorem 2.1) and Type A $\Rightarrow$ Type B in the situation modified with varying $S(x)$ (Theorem 4.2). Note that Type B is proved only via proving Type A.

Nevanlinna Theory	Diophantine approximation		
LLD	PSST	Type A	Type B
WF	PSST $\Rightarrow$ SST	Type A $\Rightarrow$ SST	Diophantine analogue WF

Proving a Diophantine inequality is equivalent to proving the smallness of the set of solutions of the opposite inequality. The residual counting function in the Schmidt Subspace Theorem implies a stronger Diophantine inequality and therefore there might be more exceptions. This means that bounding the height of the solutions of the opposite inequality will become more nontrivial and harder.

The easiest case is the following. Suppose that all linear forms F_i $(i = 0, \dots, N)$ are defined over a fixed number field k. We consider the approximation to hyperplanes by k-rational points. The approximation inequality

$$\sum_{i=0}^{N} (m_S(x, F_i) + N^n(x, F_i)) \leq (n + 1 + \varepsilon)\,\mathrm{ht}\,(x) + C(\varepsilon)$$

becomes "trivial" for small N, i.e., $N \leq n$. Indeed, the "First Main Theorem"

$$\mathrm{ht}\,(x) = m_S(x, F_i) + N_S(x, F_i) + O(1)$$

[23] The proof of the Roth lemma is the origin of the ineffectiveness of the Roth and Schmidt Theorems. The Roth lemma is the most difficult part of the proof of the Roth theorem and also plays the central role in the Schmidt Subspace Theorem.

implies that the opposite approximation inequality reduces to

$$\begin{aligned}(N+1)\mathrm{ht}\,(x)+O(1) &\geq \sum_{i=0}^{N}(m(x,F_i)+N(x,F_i)) \\ &\geq \sum_{i=0}^{N}(m_S(x,F_i)+N^n(x,F_i)) \\ &> (n+1+\varepsilon)\,\mathrm{ht}\,(x)+C(\varepsilon)\ .\end{aligned}$$

If $N \leq n$ this reduces to

$$(n-N+\varepsilon)\,\mathrm{ht}\,(x) < C'(\varepsilon)$$

which implies an effective bound on the height of all solutions of the opposite inequality. However, even for small N, if F's are defined in an extension K of k, the approximation inequality is as nontrivial as that in the case of general N. Therefore, if we believe the existence of the geometry which unifies Diophantine approximation and Nevanlinna theory, the above argument (using the "First Main Theorem") has no essential meaning and we should find more geometric way explaining the bound of the height of the solutions of the opposite inequality even in the above easiest situation. Later we examine the simplest case and show that the condition (17) (split from (15)) in fact implies the effectiveness.

We suppose next that the linear forms are defined in the extension K of k and we consider the approximation to hyperplanes by k-rational points, just as in the classical Roth theorem (we assume that the valuations of k is extended to those of K appropriately). The behavior of the Weil height under the field extension implies that the above mentioned argument does not work. However, the arguments in the proof of Theorems 2.1 and 2.3 still work with respect to the extended valuations.

The above consideration suggests us to "assume" (as a working hypothesis) that the structure of the proof of Theorems 2.1, 2.3, 4.1 and 4.2 contain something essential for the existence of an effective bound of the height of the solutions of the opposite approximation inequality.

We begin with the simplest case. Set $n=1$, $k=\mathbb{Q}$ and $F=x_0$ where (x_0,x_1) are coordinates of $\mathbb{Q}^2$ and $F=0$ represents the point ∞ of $\mathbb{P}^1(\mathbb{Q})$. This simplest case is most important in the attempt toward the effectiveness. In fact, the following simple argument turns out to be the non-trivial first step. We would like to show that for any given positive number ε, there exists an effective bound for the solutions to the Diophantine problem (17):

$$\begin{aligned}&m_{S_\infty}(\overline{x}^{(1)},S_0) < m_{S_\infty}(x,F)-\varepsilon\,\mathrm{ht}\,(x) \quad \text{or} \quad m_{S_\infty}(\overline{x}^{(1)},\infty) > \varepsilon\,\mathrm{ht}\,(x) \\ &\forall\, \overline{x}^{(1)} \in T\mathbb{P}^1(\mathbb{Z})\ .\end{aligned}$$

In this case, it is equivalent to the following:

(18)

$$\begin{aligned}&\mathrm{dist}_{\,\mathrm{Euc}}(x^{(1)},S_0)^{-1} < \mathrm{dist}_{\,\mathrm{Euc}}(x,F)^{-1}H(x)^{-\varepsilon} \quad \text{or} \quad \mathrm{dist}_{\,\mathrm{Euc}}(x^{(1)},\infty)^{-1} > H(x)^{\varepsilon} \\ &\forall\, x^{(1)} \in \mathbb{Z}^2\ ,\end{aligned}$$

where S_0 represents the zero in the first jet space (= tangent space) of $\mathbb{P}^1(\mathbb{Q})$ and dist_{Euc} means to measure the Euclidean distance. Each quantity in the above inequalities has the following geometric meaning:

$$\begin{aligned}
&\text{dist}_{\text{Euc}}(x, F)^{-1} = |\text{slope}| \ , \\
&\text{dist}_{\text{Euc}}(x^{(1)}, S_0)^{-1} = \min\{|\text{width}|^{-1}, |\text{slope}|\} \ , \\
&\text{dist}_{\text{Euc}}(x^{(1)}, \infty)^{-1} = |\text{width}| \ .
\end{aligned}$$

Here, |slope| represents the maximum of the absolute value of the usual slope in $\mathbb{R}^2$ of the line determined by the point $x \in \mathbb{Z}^2$ and 1. Moreover, $|\text{width}|^{-1}$ represents the maximum of the inverse of the Fubini-Study length of the vector $x^{(1)}$ (measured as a tangent vector in $T_{[x]}\mathbb{P}^1(\mathbb{C})$) and 1. Finally, |width| represents the maximum of the Fubini-Study length of the vector $x^{(1)}$ and 1. As the Fubini-Study metric of $\mathbb{P}^1(\mathbb{C})$ is given by $\frac{|dz|}{1+|z|^2}$, we have

$$\begin{aligned}
|\text{width}| &:= \max\left\{\frac{||x^{(1)}||_{\text{Euc}}}{H(x)^2}, 1\right\} \ , \\
|\text{width}|^{-1} &:= \max\left\{H(x) \cdot \frac{H(x)}{||x^{(1)}||_{\text{Euc}}}, 1\right\} \ ,
\end{aligned}$$

for x with $H(x) \geq 100$ (for instance).

We then **conjecture** that *the x's satisfying the second inequality of (18) have an effective bound.*

As for the first inequality in (18), the necessary condition satisfied by x is $|\text{width}|^{-1} \leq |\text{slope}|$. It is explicitly written as

$$\max\left\{\frac{H(x)^2}{||x^{(1)}||_{\text{Euc}}} \ , \ 1\right\} \ \leq \ \frac{H(x)}{|x_0|} \ , \qquad \forall x^{(1)} \in \mathbb{Z}^2 \ . \tag{19}$$

We **conjecture** that *such x's have an effective bound.*

We thus have the following effective Roth-type conjecture:

Conjecture 5.1. *Let $F_0, \ldots, F_N$ be distinct linear forms on $\mathbb{P}^1(\mathbb{Q})$. Then for any ε, there exists an effectively computable constant $C(\varepsilon, F)$ such that*

$$\sum_{i=0}^{N} m(x, F_i) + N^1(x, F_i) \leq (2+\varepsilon)\,\text{ht}\,(x) + C(\varepsilon, F) \quad \forall x \in \mathbb{P}^1(\mathbb{Q}) \ .$$

Using the definition of the residual counting function $N(x, F_i) - N_1(x, F_i) = N^1(x, F_i)$ and the "First Main Theorem" $m(x, F_i) + N(x, F_i) = \text{ht}\,(x)$, we can rewrite Conjecture 5.1 as

$$\text{ht}\,(x) \leq (1+\varepsilon) \sum_{i=0}^{2} N_1(x, F_i) + C(\varepsilon) \ .$$

Now we take $F_0 = x_0$, $F_1 = x_1$ and $F_2 = -x_1 - x_2$. Then Conjecture 5.1 in the product form becomes the statement of the effective version of the abc-conjecture:

Conjecture 5.2 (abc-conjecture). *For any $\varepsilon < 0$, there exists an effectively computable constant $C(\varepsilon)$ such that for all mutually prime integers a, b and c satisfying $a + b + c = 0$, the inequality*

$$\max\{|a|, |b|, |c|\} \leq C(\varepsilon) \Bigg(\prod_{\substack{p:\ \mathrm{prime} \\ p \mid abc}} p \Bigg)^{1+\varepsilon}$$

holds.

The case $n = 1$ and k being any number field is the same except we must consider $\mathbb{R}^{2r_1} \times \mathbb{C}^{2r_2}$ into which $\mathcal{O}_k^2$ has a co-compact embedding, where r_1 and r_2 are the number of real and (conjugate pair of) complex places of k ($r_1 + 2r_2 = [k : \mathbb{Q}]$). Indeed, let $(x_1, x_2) \in k^2$ and $(x_1^{(\iota)}, x_2^{(\iota)})$ the point of $\mathbb{R}^2$ or $\mathbb{C}^2$ corresponding to the real or complex embedding of k. Let $L(x) = a_1x_1 + a_2x_2$ be a linear form defined over k and $L^{(\iota)}$ the linear form in $\mathbb{R}^2$ or $\mathbb{C}^2$ corresponding to the real or complex embedding of k. Then, considering the linear equation $L(x) = 0$ in k^2 is equivalent to considering the system of linear equations $L^{(\iota)}(x^{(\iota)}) = 0$. For each ι we think of $L^{(\iota)}(x^{(\iota)}) = 0$ as an equation defined in $\mathbb{R}^{r_1} \times \mathbb{C}^{2r_2}$. We can thus argue quantitatively the analogue of (18) over the co-compact lattice of algebraic integers.

Next we consider the case n being general and k being any number field. Let $F_0, \ldots, F_N$ be linear forms with k coefficients in general position. In multiple dimensional case we cannot separate F's because these define hyperplanes having nonempty intersections. Let r_1 and r_2 be the number of real and (conjugate pair of) complex places. The Diophantine problem (17) is described as

$$\sum_{i=0}^{N} m(\overline{x}^{(p)}, F_i^{(1)}) < \sum_{i=0}^{N} m(x, F_i) - \varepsilon\,\mathrm{ht}\,(x) \quad \text{or} \quad m(\overline{x}^{(p)}, \infty) > \varepsilon\,\mathrm{ht}\,(x)$$
$$\forall\ \overline{x}^{(p)} \in T_{[x]}\mathbb{P}^n(\mathcal{O}_k)\ .$$

Here, $\overline{x}^{(p)} \in T_{[x]}\mathbb{P}^n(k)$ is well defined (and therefore the approximation to $F_i^{(1)}$ is also well-defined). Indeed, let $f : \mathbb{C} \to X$ be a holomorphic curve into any complex manifold X and let $(z_1, \ldots, z_n)$, $(w_1, \ldots, w_n)$ be two systems of holomorphic local coordinates around $f(z)$. Write z_i and w_j simply for the compositions $z_i \circ f$ and $w_j \circ f$. Then we have $\frac{dz_i}{dz} = \sum_{j=1}^{n} \frac{\partial z_i}{\partial w_j} \frac{dw_j}{dz}$. Differentiating this $(p-1)$ times gives $\frac{d^pz_i}{dz^p} \equiv \sum_{j=1}^{n} \frac{\partial z_i}{\partial w_j} \frac{d^pw_j}{dz^p}$ modulo differentials up to order $p-1$. So $\frac{d^pz_i}{dz^p}$ for any system of holomorphic local coordinates behaves like a tangent vector.

The above Diophantine problem is equivalent to the problem of bounding the height of x's satisfying the following inequalities on $\mathcal{O}_k^{n+1}$ which is a co-compact lattice in $\mathbb{R}^{(n+1)r_1} \times \mathbb{C}^{(n+1)r_2}$:

$$\text{(20)} \qquad \prod_{\substack{v \in S_\infty \\ 0 \leq i \leq N}} \mathrm{dist}_v(x^{(p)}, F_i^{(1)})^{-1} < \prod_{\substack{v \in S_\infty \\ 0 \leq i \leq N}} \mathrm{dist}_v(x, F_i)^{-1} H(x)^{-\varepsilon}$$
$$\text{or} \quad \prod_{\substack{v \in S_\infty \\ 0 \leq i \leq N}} \mathrm{dist}_v(x^{(p)}, \infty)^{-1} > H(x)^{\varepsilon}\ ,$$
$$\forall\ x^{(p)} \in \mathcal{O}_k^{n+1}\ .$$

The reason why we conjectured that we able to bound the solutions to (17) in the case of $n = 1$ was that in (19) we find $H(x)$ twice in the left hand side while only once in the right hand side.

As the first inequality of (20) suggests, the inequality of type (19) still holds if x locates in "general" position (note that the linear embedding $\mathbb{P}^1(\mathbb{C}) \hookrightarrow \mathbb{P}^n(\mathbb{C})$ is isometric and totally geodesic with respect to the Fubini-Study metric of projective spaces). However, this "happy" situation on the first inequality no longer holds, if x locates at some special position relative to the hyperplanes determined by F's. On the other hand, the second inequality of (17) has nothing to do with F's and so it causes no difficulty in multi-dimensional situation. Therefore, we are exclusively interested in the first inequality of (17).

Since Theorems 2.1/4.1 and 2.3/4.2 hold outside of a collection of finitely many proper linear subspaces and once the effectiveness of the simplest situation of $n = 1$ were proved, the induction process on n would proceed. So, What we must do now is to determine which linear subspaces are "special" in the sense that the "happy" situation of (19) no longer holds.

The criterion is simple. Given a proper linear subspace, we just count how many $H(x)$'s which appears in both sides and compare them. If we find more $H(x)$'s in the left, we are able to have an effective bound. Otherwise, we have no hope to get effective bound only from the Archimedean places and therefore we have to take the non-Archimedean ones in S_x into account. In doing so, we can argue inductively on the dimension of linear spaces, because, in Conjecture 5.1 we have already established the effective version of the Roth theorem.

How to reduce to lower dimensional cases is explained as follows. Given a proper linear subspace we consider the restrictions of F's. If F's are still in general position, we can reduce the dimension. Suppose that the restricted F's are no longer in general position. If this happens, we need case by case consideration.

This happens, if, for instance, we choose F's to be a configuration of four lines in general position in $\mathbb{P}^2$ and take any diagonal line L connecting P and Q which are two of six intersection points. Let $\{x\}$ be an infinite sequence of k-rational points contained in L. Assume that the sequence approximates P in the Archimedean places. In this case, $H(x)$'s appear twice in both sides of (19) and considering only Archimedean places is not enough. In fact, this sequence will approximates Q in the non-Archimedean places and We must argue synthetically taking both approximations into account. Now let us suppose that L is not exceptional. Then we have

$$\sum_{i=0}^{3} m(x, F_i) + N^2(x, F_i) \leq (3 + \varepsilon)\,\mathrm{ht}\,(x) + C(\varepsilon)$$

and as x lies in L, this inequality reduces to the following inequality for a point in L:

$$2(m(x, P) + m(x, Q) + N^1(x, P) + N^1(x, Q)) \leq (3 + \varepsilon)\,\mathrm{ht}\,(x) + C(\varepsilon)\ .$$

However, If we forget the ambient space and consider the sequence as that in L ($\cong \mathbb{P}^1(k)$), we have the Roth theorem

$$m(x, P) + m(x, Q) + N^1(x, P) + N^1(x, Q) \leq (2 + \varepsilon)\,\mathrm{ht}\,(x) + C(\varepsilon)$$

and it is well-known that this inequality is best possible which means that no improvement on $2+\varepsilon$ is possible. So, the conclusion of the same inequality as the Schmidt Subspace Theorem for points of L violates the Roth theorem. Therefore L must be an exceptional subspace.

Next we show that any line L passing through just one intersection point P is not exceptional. Let Q and R be other two intersections of L and the line configuration. Suppose that L is not exceptional. Then

$$2m(x,P)+m(x,Q)+m(x,R)+\cdots \leq (3+\varepsilon)\,\mathrm{ht}\,(x)+C(\varepsilon)\ .$$

On the other hand, the Roth theorem and the "First Main Theorem" imply

$$\begin{aligned}2m(x,P)+m(x,Q)+\cdots &\leq \mathrm{ht}\,(x)+(m(x,P)+m(x,Q)+\cdots)\\ &\leq \mathrm{ht}\,(x)+(2+\varepsilon)\,\mathrm{ht}\,(x)+C(\varepsilon)=(3+\varepsilon)\,\mathrm{ht}\,(x)+C(\varepsilon)\end{aligned}$$

and therefore we have no contradiction. This means that L is not exceptional.

The reason why the above combinatorial argument (based on the Roth theorem) synthesizes Archimedean and non-Archimedean places lies in the fact that under the assumption $m_{S_\infty}(x^{(p)},\infty)<\varepsilon\,\mathrm{ht}\,(x)$, the inequality in Theorem 2.3 (this is a version of Parametric Subspace Theorem) and the Roth-Schmidt approximation inequality is almost equivalent[24].

It is now clear how to determine the exceptional linear subspaces in the general case (n being general, k being any number field and F's linear forms defined over k in general position). First of all we say that a given linear subspace V is exceptional if the conclusion of the Schmidt Subspace Theorem on $\mathbb{P}^n(k)$ translated to an approximation inequality on V contradicts the ($\dim V$)-version of the Schmidt Subspace Theorem. To check this, we have only to argue just we did to check lines when $n=2$ and F's are four lines in general position. Although the argument becomes combinatorially more complicated as n becomes larger, we are able to algorithmically organize it.

The conclusion is that the set of all maximal exceptional subspaces are determined by combinatorial argument based on the inductive use of the (lower dimensional) Schmidt Subspace Theorem.

We say that an exceptional subspace is maximal if this is maximal among all exceptional subspaces with respect to the inclusion.

Conjecture 5.3 (Effective Schmidt Subspace Theorem). *Let $F=\{F_i\}_{i=0}^N$ be a set of linear forms in $\mathbb{P}^n(k)$ in general position. Let $\varepsilon>0$. Then there exists an effectively computable finite union of linear subspaces $E(F,\varepsilon)$ (we call them exceptional subspaces) and an effectively computable constant $C(F,\varepsilon)$ such that for all $x\in\mathbb{P}^n(k)\backslash E(F,\varepsilon)$ the approximation inequality*

$$\sum_{i=0}^N m(x,F_i)+\sum_{i=0}^N N^n(x,F_i)\leq (n+1+\varepsilon)\,\mathrm{ht}(x)+C(F,\varepsilon)$$

[24] This is evident from the proof of the Schmidt Subspace Theorem given in §4 and the proof of Theorem 4.3.

holds. Moreover, for given ε and F's, we have a combinatorial algorithm to determine all maximal exceptional subspaces. In particular, there exists an effective bound for the height of exceptional subspaces.

We execute the above algorithm in the simplest case, i.e., n being general, $k = \mathbb{Q}$ and F's being $F_i = x_i$ for $i = 0, 1, \ldots, n$ and $F_{n+1} = -x_0 - x_1 - \cdots - x_n$. In this case it is easy to determine all maximal exceptional subspaces. These are diagonals, i.e., those hyperplanes which are defined by the "sub-sum" $\sum_{i \in I} x_i = 0$ over any proper subset $I \subset \{0, 1, \ldots, n\}$ such that $|I| \geq 2$ (the case of $|I| = 1$ is excluded from the begining because we are arguing in the setting of the Schmidt Subspace Theorem with respect to the linear forms F's).

We check this in the case of $n = 3$. Set $F_i = x_i$ for $i = 0, \ldots, 3$ and $F_4 = -(x_0 + \cdots + x_3)$. First we consider $\mathbb{P}^1$ over which F's restricts in general position. By the effective Roth, any such $\mathbb{P}^1$ is not exceptional. Next, suppose that F's do not restrict in general position. If complementary two and three of five intersection points coincide, the left hand side of the Schmidt Subspace Theorem becomes $3m(x, P) + 2m(x, Q)$ plus residual counting function terms, which (as a total) is not larger than $5\,\mathrm{ht}\,(x)$ by the "First Main Theorem" and indeed this estimate is best possible for some infinite sequence. This violates the Schmidt Subspace Theorem for $n = 3$. So, such $\mathbb{P}^1$ is exceptional. If just three coincide, the left hand side becomes $3m(x, P) + m(x, Q) + m(x, R)$ plus residual counting function terms, which is not larger than $(4 + \varepsilon)\mathrm{ht}\,(x)$ by the effective Roth and the "First Main Theorem". This does not violate the Schmidt Subspace Theorem for $n = 3$. We can argue similarly in the case that just two pairs coincide. Thus we have shown that the exceptional $\mathbb{P}^1$'s are characterized by the condition that two and three of five intersection points coincide. We cannot bound the height of such $\mathbb{P}^1$'s. However, these $\mathbb{P}^1$'s turn out to be not maximal. These are classified into a finite number of 1-parameter family and each family is contained in some $\mathbb{P}^2$ defined by certain subsum $= 0$, which, as we show below, is exceptional. Next, we consider a $\mathbb{P}^2$ on which the restriction of F's are in general position. This case is reduced to the Schmidt Subspace Theorem on $\mathbb{P}^2$ with four lines in general position. In this case the only exceptions are diagonal lines. However as we saw above, the left hand side of the Schmidt Subspace Theorem is $2(m(x, P) + m(x, Q))$ plus residual counting function terms and is not larger than $4\,\mathrm{ht}\,(x)$. This violates the Schmidt Subspace Theorem for $n = 2$ but does not for $n = 3$. So any $\mathbb{P}^2$ in general position with F's is not exceptional. If $\mathbb{P}^2$ is determined by the line $L = \{F_0 = F_1 = 0\}$ and a point P on $F_2 = F_3 = 0$, the right hand side of the Schmidt Subspace Theorem becomes $2m(x, L) + 2m(x, P)$ plus residual counting function terms, which is not larger than $4\,\mathrm{ht}\,(x)$ by the "First Main Theorem". So, this does not violate the Schmidt Subspace Theorem and any such $\mathbb{P}^2$ is not exceptional. If $\mathbb{P}^2$ is determined by the line $L = \{F_0 = F_1 = 0\}$ and the point $P = \{F_2 = F_3 = F_4 = 0\}$, the left hand side of the Schmidt Subspace Theorem becomes $3m(x, L) + 2m(x, P)$ plus residual counting function terms which is not larger than $5\,\mathrm{ht}\,(x)$ by the "First Main Theorem". This violates the Schmidt Subspace Theorem and this case corresponds to the subsum $x_0 + x_1 = 0$. Similarly $\mathbb{P}^2$ determined by the line $L = \{F_3 = F_4 = 0\}$ and the point $P = \{F_0 = F_1 = F_2 = 0\}$ corresponds to the subsum $x_0 + x_1 + x_2 = 0$. We thus conjecture the following:

Conjecture 5.4 (generalized abc-conjecture). *For any $\varepsilon < 0$ there exists an effectively computable constant $C(\varepsilon)$ such that the following holds: If $a_0, a_1, \ldots, a_n, a_{n+1}$*

are mutually prime integers satisfying the condition $a_0 + a_1 + \cdots + a_n + a_{n+1} = 0$ *and* $\sum_{i \in I} a_i \neq 0$ *for any proper subset* $I \subset \{0, 1, \ldots, n\}$*, then*

$$\max\{|a_0|, |a_1|, \cdots, |a_{n+1}|\} \leq C(\varepsilon)\Bigg(\prod_{\substack{p\,:\,\text{prime} \\ p \,|\, a_0 a_1 \cdots a_{n+1}}} p^{\min\{\mathrm{ord}_p(a_0 a_1 \cdots a_{n+1}), n\}} \Bigg)^{1+\varepsilon}$$

holds.

References

[A] L. Ahlfors, *The theory of meromorphic curves*, Acta Soc. Sci. Finn. N.S. A. **III** (1941), 1-31.

[B-V] E. Bombieri and J. Vaaler, *On Siegel's lemma*, Inv. Math. **73** (1983), 11-32.

[C] H. Cartan, *Sur les zéros des combinaisons linéaires de p fonctions holomorphes donnés*, Mathematica **7** (1933), 5-31.

[K1] R. Kobayashi, *Nevanlinna's Lemma on logarithmic derivative and integral geometry*, preprint (2000).

[K2] R. Kobayashi, *Toward Nevanlinna theory as a geometric model of Diophantine approximation*, to appear in AMS SUGAKU exposition (2002).

[KS] S. Kobayashi, *Hyperbolic Complex Spaces*, Springer-Verlag, 1998.

[L] S. Lang, *Number Theory III*, Encyclop. Math. Sc., Vol. 60, Springer-Verlag, 1991.

[N] R. Nevanlinna, *Analytic Functions*, Springer-Verlag, 1970.

[R] K.F.Roth, *Rational approximations to algebraic numbers*, Mathematica **2** (1955), 1-20.

[S] W. Schmidt, *Diophantine Approximation*, LNM 785, Springer Verlag, 1980.

[V] P. Vojta, *Diophantine Approximation and Value Distribution Theory*, LNM 1239, Springer Verlag, 1987.

[Y] K. Yamanoi, *Algebro-geometric version of Nevanlinna's lemma on logarithmic derivative and applications*, to appear in Nagoya Math. Jour. (2002).