Invent math. 116, 109–138 (1994)

Inventiones mathematicae

Diophantine approximations on projective spaces

Gerd Faltings[1] **and Gisbert Wüstholz**[2]

[1] Princeton University and Institute for Advanced Study, Princeton, NY, USA
[2] Département Mathematik, ETH Zürich, CH-8092 Zürich, Switzerland

Oblatum 10-II-1993

To Armand Borel

1 Introduction

In this paper we study diophantine approximations on projective spaces of arbitrary dimension. Especially we give a new proof for the results of W.M. Schmidt. Historically diophantine approximation has been the method for proving finiteness results in diophantine geometry. In the case of the projective line the fundamental result is Roth's theorem. This has been generalized to higher dimension by Schmidt. His basic result is the subspace theorem (see [S2]). However the proof is very complicated and involved. It relates on deep results in the geometry of numbers.

Recently a new development in the theory was started by Vojta which lead to a breakthrough for diophantine approximations on abelian varieties. He found a new way to use heights and the theorem of Mordell and Weil. His main insight was that on products of abelian varieties one has many more line bundles than just the product bundles. His discovery inspired Faltings to find a new general result in the theory of diophantine approximations, the product theorem. It very much simplifies the theory and its application to abelian varieties lead to some generalization of Mordell's conjecture (see [Fa]). As a second application Faltings also gave a lower bound for the distance of a rational point to a hypersurface on an abelian variety.

In the present paper we study the consequences of the product theorem in the more classical case of projective spaces. Here not so many line bundles are available as for abelian varieties. Therefore the results will be much weaker in general. However in the case Schmidt is considering the result is best possible up to an ε. The advantage in the present situation is that we do not have to cope with complicated line bundles as in the abelian case; this simplifies the proofs considerably. On the other hand the estimates one needs have to be much more precise. In order to obtain the best possible exponents in the approximation theorems we have to be very careful.

Now let us explain our results. We fix a number field K and chose a subvariety $E \subset \mathbb{P}^n$ defined over K and study the distance $d_v(E; x)$ of a K-rational point of $\mathbb{P}^n \setminus E$ to E where v is a place of K. It is natural to ask for an estimate of the type

$$d_v(E; x) \gg H(x)^{-\kappa}$$

such that the exponent κ is as small as possible. Here $H(x)$ denotes the absolute Weil height of x. We shall give such an estimate depending only on the geometry of E. It will turn out that the assertion is only non-trivial for subvarieties E which are not geometrically irreducible. For example Schmidt's theorem deals with a linear subspace defined over a finite extension of K or equivalently with the K-scheme obtained by taking the union of all its conjugates.

Our paper is organized as follows. In section 2 we give a review of the theory given in [Fa]. In particular we explain the product theorem. In the third section we define the geometric invariant. It is based on the theory Hilbert polynomials which leads to the definition of a probability measure associated with the geometry of E. In the next section we compute this invariant in the case of Schmidt's theorem. The basic tool here is a new theory of stability for vector spaces with filtration. After this we count dimensions of certain vector spaces which is crucial for applying Siegel's lemma which we state next. Then we need some results on differential operators on subvarieties of projective varieties. In particular we have to adapt the product theorem to our situation. In our case we have to apply the product theorem not to products of projective spaces but to products of projective varieties since the proof of our main result is by induction so that products of projective varieties appear and we have to deal with it. Then, in Sect. 8, we give our main result and prove it there. After this we give applications and further results. In particular we shall show how the linear case can be applied in the non-linear situation. This is done in Sect. 9. We mention that Theorem 9.3 gives an answer to a question raised by Schmidt in [S1]. Then in Sect. 10 we show how the unit equation can be treated in our theory. In Sect. 11 we discuss the norm form equation and generalize the results of Schmidt. In Sect. 12 we prove a gap principle which is useful in counting the number of solutions of diophantine equations.

2 Preliminaries

For a number field K and a rational point $x \in \mathbb{P}^n(K)$ we denote by $H(x)$ the absolute Weil height of x. It is invariant under extensions of K. Its logarithm $h(x)$ is the degree over the integers $\mathcal{O}_K$ of K of the fiber of $\mathcal{O}(1)$ at x, divided by the degree $[K:\mathbb{Q}]$. Here the line bundle $\mathcal{O}(1)$ has to carry suitable hermitian metrics at the infinite places. Also if v is a place of K we denote by $\|\cdot\|_v$ generally the v-adic norm on K_v, the v-adic completion of K, normalized in such a way that $\|p\|_v = p^{-1}$ if v is a finite place dividing p. Otherwise $\|\cdot\|_v$ is the usual norm on $\mathbb{R}$ or $\mathbb{C}$. In order to compensate for this we count places v with multiplicity if we sum over all of them. More precisely in such sums taken over all places of K we count a place v with multiplicity equal to the degree of K_v over $\mathbb{Q}_p$ if v divides p or equal to the degree of K_v over $\mathbb{R}$ if v is an infinite place. Equivalently we can sum over all embeddings of L into the algebraic closure of $\mathbb{Q}_p$ or into $\mathbb{C}$ respectively.

The theory of heights was extended in [Fa] to varieties of arbitrary dimension. Accordingly for any subvariety $X \subseteq \mathbb{P}^n$ of the projective space which is defined and irreducible over $\mathbb{Q}$ we can define its height $h(X)$ which controls various constants relevant for X. In the special case when the dimension of X is zero it coincides with the height $h(x)$ introduced above.

The next important tool from [Fa] is the product theorem. Let K be a field of characteristic zero and $P = P_1 \times \ldots \times P_m$ be a product of projective spaces $P_j = \mathbb{P}^{n_j}$, $j = 1, \ldots, m$. For positive integers $d_1, \ldots, d_m$ we consider the line bundle $\mathscr{L} = \mathcal{O}(d_1, \ldots, d_m)$. If f is a global section of $\mathscr{L}$ the index $i(x;f)$ of f at a point $x \in P$ is the weighted multiplicity of f at a point x with weights $1/d_j$, $j = 1, \ldots, m$. For any positive real number σ there is a subscheme $Z_\sigma \subseteq P$ consisting of those x for which the index is at least σ. If the field is algebraically closed the product theorem states that for any sufficiently small positive ε there exists an integer $r = r(\varepsilon)$ depending on ε with the following property. Suppose that $d_j/d_{j+1} \geqq r$ for $j = 1, \ldots, m-1$. Then for any $x \in Z_\sigma$ there exist irreducible subvarieties $Z_j \subseteq P_j$, $1 \leqq j \leqq m$, such that $x \in Z_1 \times \ldots \times Z_m \subseteq Z_{\sigma - \varepsilon}$. Furthermore the degrees of the subvarieties Z_j can be bounded by a constant depending only on ε. If the field K is not algebraically closed then the theorem holds over a finite extension of K whose degree can be bounded or over K if we allow that the Z_j are only K-irreducible but not necessarily geometrically irreducible. However in applications the Z_i will tend to have a smooth K-rational point so that they are in fact geometrically irreducible. Also there is an arithmetic extension of the product theorem if K is a number field. Namely if f has integral coefficients then for suitable positive constants c_1, c_2

$$\sum_j d_j h(Z_j) \leqq c_1(d_1 + \ldots + d_m) + c_2 \sum \log \|f\|_v$$

where the sum on the right hand side is over all infinite places and $\|f\|_v$ is any reasonable norm of f like the L^2-norm or the maximum norm.

In order to apply the product theorem later we need some further theory. For details we refer to [Fa]. Let $X \subseteq \mathbb{P}^n$ be a projective variety of dimension d and defined over a number field K. Then there exists a composition $\pi: X \to \mathbb{P}^d$ of good projections. A good projection

$$p: X \subset \mathbb{P}^n \setminus \{x\} \to \mathbb{P}^{n-1}$$

has as center a point $x \in \mathbb{P}^n(K)$ which has a representation $x = (x_0, \ldots, x_n)$ such that the x_j are integers bounded by $[K:\mathbb{Q}] \deg(X)$ and such that the distances $d_v(X; x)$ are bounded below by a positive constant depending only $\deg(X)$ for all infinite places of K. After a modification the projection π can be extended to a projection $\tilde{\pi}: \tilde{X} \to \mathbb{P}^d$ of integral models.

We need also estimates for the heights of preimages of varieties. Suppose that $Y \subset \mathbb{P}^d$ is a proper subvariety defined over K. Then it can be defined by homogeneous polynomials with coefficients in K such that their logarithmic heights are bounded in terms of $h(Y)$. These polynomials define also a subvariety Y' of $\mathbb{P}^n$ whose height is bounded by a multiple of $h(Y)$. It follows that

$$h(X \cap Y') \leqq c(h(X) + h(Y))$$

for some positive constant depending on the degree of X and Y.

Let now $X_1, \ldots, X_m$ be such varieties with $X_j \subseteq \mathbb{P}^{n_j}$. Then we have seen that there exist good projections $\pi_j\colon X_j \to \mathbb{P}^{d_j}$, $d_j = \dim X_j$, $j = 1, \ldots, m$. This can be combined to give a projection

$$\pi\colon X \to P' = \mathbb{P}^{d_1} \times \ldots \times \mathbb{P}^{d_m}$$

where

$$X = X_1 \times \ldots \times X_m \subset P = \mathbb{P}^{n_1} \times \ldots \mathbb{P}^{n_m}.$$

Finally we need some information about how the index behaves with respect to good projections. So let $\pi\colon X \to \mathbb{P}^d$ be again a good projection. Then there exists a homogeneous polynomial G with integer coefficients bounded by $\exp(c_1\, h(X) + c_2)$ and degree at most $(n-d)\,[K:\mathbb{Q}]\,\deg(X)$ such that G does not vanish identically on X but annihilates the module of relative differentials $\Omega^1_{X/\mathbb{P}^d}$. It follows that for any differential operator D on $\mathbb{P}^d$ of order 1 the operator GD extends to a differential operator on X.

Take now again m such varieties $X_1, \ldots, X_m$ with $X_j \subseteq \mathbb{P}^{n_j}$ and put $X = X_1 \times \ldots \times X_m$, $P = \mathbb{P}^{n_1} \times \ldots \times \mathbb{P}^{n_j}$ and $P' = \mathbb{P}^{d_1} \times \ldots \times \mathbb{P}^{d_m}$. Then we obtain polynomials $G_1, \ldots, G_m$ as above. We consider a point $x = (x_1, \ldots, x_m)$ in X with $G_j(x_j) \neq 0$ for $j = 1, \ldots, m$. Then the index of any section f of $\mathcal{O}(t_1, \ldots, t_m)$ on X can be computed with respect to the lifts of the differential operators on P and if g is the norm of f with respect to π we find that

$$i(x;f) \leqq i(\pi(x);g).$$

These considerations show that it suffices to have a product theorem for multiprojective spaces in order to deduce one for products of projective varieties. In fact if f is a section of $\mathcal{O}(t_1, \ldots, t_m)$ on X with index at least σ at $x \in X$ then the norm g of f has index at least σ at $\pi(x)$. We find then a product variety $Z' = Z'_1 \times \ldots \times Z'_m$ on P' with $\pi(x) \in Z'$ and the properties already stated. We put $Z_j = \pi_j^{-1} Z'_j$ and $Z = Z_1 \times \ldots \times Z_m$. This is a proper subvariety of X and contains x. Hence x is contained in a proper subvariety Z of X. The degrees and heights of the subvarieties can be calculated by means of the preceeding remarks. If on the other hand $x = (x_1, \ldots, x_m)$ is a point with the property that $G_j(x_j) = 0$ for some j then we simply replace X_j by an irreducible component of the intersection of X_j with the zero set of G_j which contains x_j. Again this is a proper subvariety of X and the degree and height can be bounded easily in terms of $h(X)$ and the degree of X.

3 Graded algebras and probability

Let $R = \bigoplus_{n \geqq 0} R_n$ be a graded Noetherian algebra over a field $K = R_0$ generated by R_1. Let $\{F^j, j \geqq 0\}$ be a decreasing sequence of graded ideals with $F^0 = R$, $F^1 \subset R_+ = \bigoplus_{n \geqq 1} R_n$ and $F^i F^j \subseteq F^{i+j}$. We assume that $\bigcap_{j \geqq 0} F^j = 0$ and this implies that on each R_n we obtain a separated finite filtration. Finally we assume that the bigraded algebra $S = \operatorname{gr}^F R$ is finitely generated. We write $S = \bigoplus_{n,j \geqq 0} S_n^j$ with $S_n^j = F_n^j / F_n^{j+1}$.

More generally let $M=\bigoplus_{n\geqq 0} M_n$ be a graded R-module with a separated filtration $F^j(M)$ by graded submodules such that $F^i\cdot F^j(M)\subseteq F^{i+j}(M)$ and such that $N=\operatorname{gr}^F M$ is a finitely generated module over S. We write $N=\bigoplus_{n,j\geqq 0} N_n^j$. Our assumptions imply that M is a finitely generated R-module and the induced filtrations on graded submodules or quotient modules satisfy it too.

We define the Hilbert series of M in the usual way as

$$P(M;Y)=\sum_{n\geqq 0}\chi_M(n)\,Y^n,$$

and the Hilbert series of N as

$$P(N;X,Y)=\sum_{j,n\geqq 0}\chi_N(j,n)\,X^j\,Y^n,$$

where $\chi_M(n)=\dim M_n$ and $\chi_N(j,n)=\dim N_n^j$. Let (a_i,b_i) for $1\leqq i\leqq l$ be the bidegrees of generators for S with a_i equal to the degree coming from the filtration and b_i from the graduation and put $d=\dim M$, the Krull-dimension of M. Then $\chi_M(n)$ is a polynomial in n of degree $d-1$ for $n\gg 1$.

Lemma 3.1 *There are polynomials $P_i(X,Y)$, $Q_i(X,Y)$, $1\leqq i\leqq\lambda$, with integer coefficients such that*

$$P(N;X,Y)=\sum_{i=1}^{\lambda}P_i(X,Y)/Q_i(X,Y).$$

The polynomials Q_i are products of at most d factors of the form $1-X^a\,Y^b$ and the pair (a,b) is in the set of (a_i,b_i) for $1\leqq i\leqq l$.

Proof. The proof is by induction on the dimension of M. It is obvious for $d=0$ since then P is a polynomial. Furthermore the Hilbert series P is additive on short exact sequences. Therefore we may assume that N has only one associated prime. In fact there is a filtration $N=N_0\supset N_1\supset\ldots\supset N_r=0$ of N such that $N_i/N_{i+1}\cong S/\mathfrak{p}_i$, $0\leqq i\leqq r-1$, where $\mathfrak{p}_i$ is an associated prime of N (see [Ma, Theorem 6.4]). Therefore

$$P(N;X,Y)=\sum_i P(N_i/N_{i+1};X,Y)$$

and the modules N_i/N_{i+1} have only one associated prime ideal $\mathfrak{p}_i$ each.

If $d>0$ we can find a bihomogeneous generator t of S such that $t\notin\mathfrak{p}$. Hence multiplication by t is injective on N and maps N_n^j into $N_{n+b_i}^{j+a_i}$ where (a_i,b_i) is the degree of t. This leads to

$$\chi_N(j+a_i,n+b_i)-\chi_N(j,n)=\chi_{N/tN}(j+a_i,n+b_i).$$

We multiply this relation by $X^{j+a_i}\,Y^{n+b_i}$. The result follows in the usual way on summing these equations. □

We note that $P(M;Y)=P(N;1,Y)$. The functions $\chi_N(j,n)$ encode much information about N and we shall determine their asymptotic behavior. We start with the simplest case, namely that $N=S$ and S is a polynomial algebra on r generators. In this case $P(S; X, Y)=Q(X, Y)^{-1}$ with $Q=(1-X^{a_1}Y^{b_1})\dots(1-X^{a_r}Y^{b_r})$ and

$$\chi(j,n)=\sum_{\substack{a_1t_1+\dots+a_rt_r=j\\ b_1t_1+\dots+b_rt_r=n}} 1$$

where $t_1, \dots, t_r$ are non-negative integers. We compare the sum with the integral $\int_{V(j,n)} dF$ where $V(j,n)$ is the domain in $\mathbb{R}^r$ given by $a_1x_1+\dots+a_rx_r=j$, $b_1x_1+\dots+b_rx_r=n$, $x_1, \dots, x_r\geqq 0$ and where dF is the volume element. If A denotes the maximum of the quotients a_i/b_i for $i=1, \dots, r$ then $V(j,n)$ is not empty only if $j\leqq An$. Without loss of generality we may assume that $A=a_r/b_r$. The evaluation of the integral above leads up to a constant to the expression

$$(n-j/A)^{r-2}\int_V dF.$$

Here dF is the volume element of the region $V\subseteq\mathbb{R}^{r-1}$ defined by

$$x_1\geqq 0, \dots, x_{r-1}\geqq 0$$

and

$$(A-a_1/b_1)x_1+\dots+(A-a_{r-1}/b_{r-1})x_{r-1}=1.$$

It follows that for j fixed

$$\chi(j,n)=c(n-j/A)^{r-2}+o(n^{r-2})$$

if $j\equiv 0 \bmod a$ and $n\equiv 0 \bmod b$ where a and b are the greatest common divisors of $a_1, \dots, a_r$ and of $b_1, \dots, b_r$ respectively and c is a positive constant. The $o(n^{r-2})$ is uniform in j. In the remaining cases $\chi(j,n)=0$.

We consider next an expression of the form $P(X,Y)/Q(X,Y)$ where P is a polynomial. The rational function again has a power series expansion with coefficients of the type above.

In order to determine the coefficients $\chi_N(j,n)$ we have to sum up such rational functions but with varying denominators. This means that a and b vary. If we replace them by their lowest common multiple a' and b' respectively which is the lowest common multiple of $a_1, \dots, a_l$ and $b_1, \dots, b_l$ respectively we get for $\chi_N(j,n)$ once again an expression as above where j and n run through the residue classes modulo a' and modulo b' respectively. The following lemma follows immediately.

Lemma 3.2 *There exists a real constant c such that for $j\equiv 0$ modulo a' and $n\equiv 0$ modulo b' we have*

$$\chi_N(j,n)=c(n-j/A)^{r-2}+o(n^{r-2})$$

with $o(n^{r-2})$ uniform in j.

We now define a probability measure ρ_n on the real line. For $x\in\mathbb{R}$ we put

$$\rho_n([x,\infty))=\max_{j/n\geqq x}(\dim F^j M_n/\dim M_n).$$

Then, if $n \equiv 0$ modulo b', we find that

$$\rho_n = \sum_{j \equiv 0(a')} (\chi_N(j, n)/\chi_M(n))\, \delta_{j/n}$$

where $\delta_{j/n}$ is the delta function at the point j/n. The support of ρ_n is in the interval $[0, A]$ and by Lemma 3.1

$$\rho_n = c \sum_{j \equiv 0(a')} ((1-j/(nA))^{r-2} + o(1))\, n^{-1}\, \delta_{j/n}.$$

For any continuous function Φ on $[0, A]$ we put

$$\int \Phi\, d\rho_n = \sum \Phi(j/n)\, \rho_n(j/n)$$

where the summation is over all j with $0 \leqq j/n \leqq A$. Let ρ be the measure $(1-x/A)^{r-2}\, dx$ on [0, A].

Then

$$\int \Phi\, d\rho_n = c \sum \Phi(j/n)((1-j/(nA))^{r-2} + o(1))\, n^{-1}$$

and this converges to a multiple of $\int \Phi\, d\rho$.

If $n \neq 0$ modulo b' we make use of the fact that R is generated by homogeneous elements of degree one. As before we may reduce the study of the measures ρ_n to the case that M has only one associated prime ideal so that we find a homogeneous element t of degree one with the property that multiplication by t is injective on M. This multiplication induces an injection from $F^j M_n$ into $F^j M_{n+1}$. Since $\dim M = (1+o(1)) \dim M_{n+1}$ we find that

$$\rho_n([x, \infty)) \leqq (1+\varepsilon_n)\, \rho_{n+1}([x, \infty))$$

for all n and with real numbers ε_n such that $\varepsilon_n \to 0$ when n tends to infinity. For any non-negative integer n we define k by $k = [n/b']$. Then

$$\rho_{kb'}([x, \infty)) \leqq (1+\varepsilon_n')\, \rho_n([x, \infty)) \leqq (1+\varepsilon_n'')\, \rho_{(k+1)b'}([x, \infty))$$

for ε_n' and ε_n'' also tending to zero with n going to infinity. This shows that for n in a fixed residue class of b' the limit exists and coincides with the limit taken over all n with $n \equiv 0$ modulo b'. Therefore the measures ρ_n converge to a multiple of ρ independently of the residue class of b' which proves the following

Lemma 3.3 *The probability measures ρ_n converge in measure to a measure ρ_∞ which is a multiple of ρ.*

Example 1 Let $R = K[T_0, \ldots, T_d]$ be the polynomial ring in $d+1$ variables and f a homogeneous polynomial of degree r. We put $F^j R = I^j$ where I is the ideal generated by f over R. The ring $\mathrm{gr}^F R$ is generated by the images of $T_0, \ldots, T_d$ and by the image of f which is in S_r^1. Hence $A = 1/r$. Since $\rho_\infty = c\rho$ and ρ_∞ is normalized the constant can be determined and one obtains for it the value rd. The interval $[x, 1/r]$ has measure $(1-rx)^d$ and the expectation value of ρ_∞ is

$$E(\rho_\infty) = \int \xi\, d\rho_\infty = \frac{1}{r(d+1)}.$$

We shall see that this expectation value explains the exponents in the generalization of Roth's theorem.

Example 2 Let $X \subseteq \mathbb{P}^d$ be a projective variety and assume that f does not vanish on X. Then we obtain again a graded Ring $R/I(X)$ where $I(X)$ is the ideal of X. Similar calculations show that $\rho_\infty[x, 1/r] = (1 - rx)^{\dim X}$. The expectation value is $E(\rho_\infty) = 1/(r(\dim X + 1))$.

In our applications we have to consider more generally filtrations which are supported by a discrete set $\mathcal{S}$ of non-negative real numbers rather than by non-negative integers. The same considerations as in the proof of Lemma 3.2 show that also in this case the probability measures ρ_n converge to a measure ρ_∞. We can therefore relax the hypothesis of Lemma 3.2 accordingly.

4 Filtrations

Let L be a field of characteristic 0 and V a finite dimensional vector space over L. We consider for real numbers

$$0 \leqq p_0 < p_1 < \ldots < p_m < p_{m+1}$$

a filtration

$$V = F^{p_0} V \supset F^{p_1} V \supset \ldots \supset F^{p_m} V \supset F^{p_{m+1}} V = 0$$

and define an invariant $\mu(V)$ by

$$\mu(V) = \sum p_j \dim((F^{p_j} V / F^{p_{j+1}} V)/\dim V).$$

The vector space V is called semistable if for every subspace $0 \neq V' \subset V$ with induced filtration we have $\mu(V') \leqq \mu(V)$ or equivalently that $\mu(V) \leqq \mu(V'')$ for each quotient V'' of V with $0 < \dim V'' < \dim V$. The notation is very similar to semistability of vector bundles on curves. In fact there is a close relation and we shall see below how this connection arises and how one can make use of it.

First we recall some facts which will be useful in the subsequent discussion. Let

$$O \to V' \to V \to V'' \to O$$

be a short exact sequence of vector spaces with filtration induced by the filtration of V and with $0 < \dim V' < \dim V$. Then $\mu(V) \geqq \mu(V')$ if and only if $\mu(V'') \geqq \mu(V')$. Indeed, under these assumptions the corresponding sequence of vector spaces $F^{p_j} V'$, $F^{p_j} V$ and $F^{p_j} V''$ is also exact so that

$$\dim(F^{p_j} V') = \dim(F^{p_j} V) + \dim(F^{p_j} V'')$$

for all $j = 0, \ldots, m+1$. this implies that

$$\dim(F^{p_j} V / F^{p_{j+1}} V) = \dim(F^{p_j} V' / F^{p_{j+1}} V') + \dim(F^{p_j} V'' / F^{p_{j+1}} V'').$$

On multiplying by p_j and adding the resulting equations we obtain

$$(\dim(V') + \dim(V''))\, \mu(V) = \dim(V')\, \mu(V') + \dim(V'')\, \mu(V'').$$

Now we observe that for any real numbers x', x'', y', y'' with $x', x'' \geqq 0$ and $y', y'' > 0$ we have

$$\frac{x'+x''}{y'+y''} \geqq \frac{x'}{y'}$$

if and only if

$$\frac{x''}{y''} \geqq \frac{x'}{y'}.$$

Therefore we get $\mu(V) \geqq \mu(V')$ if and only if we have $\mu(V'') \geqq \mu(V')$. This is exactly what we wanted to verify.

If V_1 and V_2 are filtered vector spaces a filtered homomorphism is a homomorphism $f\colon V_1 \to V_2$ such that $f(F^{p_j} V_1) \subseteq F^{p_j} V_2$. A filtered homomorphism is called strict if $f(F^{p_j} V_1) = f(V_1) \cap F^{p_j} V_2$. In other words the filtration on $f(V_1)$ induced by the filtration of V_1 coincides with the filtration induced by the filtration of V_2. Let V_1 and V_2 be semistable and $f\colon V_1 \to V_2$ a filtered homomorphism. Then if $\mu(V_1) > \mu(V_2)$ it follows that $f=0$. Otherwise we would obtain a factorization of f by filtered maps

$$V_1 \to V_1' \cong V_2' \to V_2$$

where V_1' is a quotient of V_1 and V_2' a subspace of V_2 both endowed with the induced filtrations. By semistability

$$\mu(V_1) \leqq \mu(V_1') = \mu(V_2') \leqq \mu(V_2)$$

which is a contradiction. Furthermore if $\mu(V_1) = \mu(V_2)$ the homomorphism f is strict.

Using this property we get a canonical Harder-Narasimhan filtration on V. This is a flag

$$V = V_k \supset V_{k-1} \supset \ldots \supset V_1 \supset V_0 = 0$$

of vector spaces with V_{i+1}/V_i semistable and with

$$\mu(V_{i+1}/V_i) < \mu(V_i/V_{i-1})$$

for all i with $1 \leqq i \leqq k-1$.

In order to prove the existence we may assume that V is not semistable and we define m to be the maximum of the numbers $\mu(V')$ for all $V' \subseteq V$. Among the subspaces V'' with $\mu(V'') = m$ we choose V_1 such that the rank of V_1 is maximal. Then obviously V_1 is semistable. Furthermore we have $\mu(V_1) > \mu(W)$ for each nonzero subquotient W of V/V_1. In fact let W_1 be the inverse image of W under the canonical projection. Then rank $W_1 >$ rank V_1 and by the maximality of V_1 we get $\mu(W_1) < \mu(V_1)$. Using the equivalence above the inequality follows as stated. The uniqueness now follows easily: Let V_2 be another subspace with $\mu(V_2) = m$ and rank $V_1 =$ rank V_2 different from V_1. Then $V_1 + V_2 \supset V_1$ and so $(V_1 + V_2)/V_1$ is a non-zero subquotient of V/V_1. Hence $\mu(V_1) > \mu((V_1 + V_2)/V_1)$. Since there exists a sequence to filtered homomorphisms

$$V_2 \to V_2 + V_2 \to (V_1 + V_2)/V_1$$

we obtain a contradiction to what we have demonstrated above.

Clearly we could also work with several filtrations $\{F_w^{p_j,w}\}$ indexed by w for $1 \leqq w \leqq l$. For each w we obtain a μ-invariant $\mu_w(V)$. We use now the invariant

$$\mu(V) = \mu_1(V) + \ldots + \mu_l(V)$$

and obtain a Harder-Narasimhan filtration correspondingly. We could also work with more general functions $\mu(V)$ and all what we need is that they satisfy

$$(\dim(V') + \dim(V''))\,\mu(V) = \dim(V')\,\mu(V') + \dim(V')\,\mu(V'')$$

for exact sequences

$$O \to V' \to V \to V'' \to O$$

of vector spaces with filtration and filtered homomorphisms.

A particular case where several filtrations arise is obtained when V is a filtered vector space and $G \subseteq \mathrm{GL}(V)$ a finite subgroup. Then we obtain filtrations $\{F_\sigma^{p_j} V\}$ for $\sigma \in G$ such that $F_\sigma^{p_j} V$ is the image of the subspace $F^{p_j} V$ under σ. Hence for each $\sigma \in G$ we get a μ-invariant $\mu_\sigma(V)$. The sum of these invariants $\mu_\sigma(V)$ over all $\sigma \in G$ gives a μ-invariant $\mu(V)$. Accordingly we obtain a canonical Harder-Narasimhan filtration. It is invariant under the group G. In fact the elements of G permute the filtration. If $V' \subseteq V$ is a subspace then the image $\tau(V')$ of V' under $\tau \in G$ satisfies

$$\tau(V') \cap F_\sigma^{p_j} V \cong V' \cap F_{\sigma\tau^{-1}}^{p_j} V.$$

Hence $\mu_\sigma(\tau(V')) = \mu_{\sigma\tau^{-1}}(V')$ and summation over σ gives $\mu(\tau(V')) = \mu(V')$. It follows that the image of a Harder-Narasimhan filtration under $\tau \in G$ is again a Harder-Narasimhan filtration. The uniqueness of the Harder-Narasimhan filtration now gives the result.

We apply this remark in the following situation. Let $L \supseteq K$ be a finite extension of the field K of characteristic 0 and let V be a vector space of finite dimension over K. We consider a filtration $\{F^{p_j} V\}$ on the vector space $W = V \otimes_K L$ which gives us a μ-invariant on W and we define a μ-invariant on V by $\mu(V) = \mu(W)$. This is a function with the properties needed in order obtain a Harder-Narasimhan filtration on V. The general considerations above apply as follows. Let $\Sigma = \mathrm{Hom}_K(L, \bar{K})$ and embed W into $V \otimes_K \bar{K}$ by means of the $\sigma \in \Sigma$. The subspaces $F_\sigma^{p_j}$ generated by $\sigma(F^{p_j})$ define a finite set of filtrations $\{F_\sigma^{p_j}\}$ on $V \otimes_K \bar{K}$ with μ-invariants μ_σ. The Galois group G of $\bar{K}$ over K acts on $V \otimes_K \bar{K}$ and has a faithful representation in $\mathrm{GL}(V \otimes_K \bar{K})$ which permutes the filtration. As above we obtain a Harder-Narasimhan filtration on $V \otimes_K \bar{K}$ which is invariant under the Galois group G. It is induced from a filtration on V which coincides with the Harder-Narasimhan filtration on V with respect to μ since μ and $\Sigma \mu_\sigma$ differ only by a non-zero factor. Thus semistability can be checked over $\bar{K}$.

Next we shall study the behavior of the Harder-Narasimhan filtration when the filtration $\{F^{p_j}\}$ is changed. If $\{F^{p'_j}\}$ is another filtration with $\max|p_j - p'_j|$

small enough then the Harder-Narasimhan filtration for $\{F^{p_j}\}$ and $\{F^{p'_j}\}$ coincide. The same holds if one takes the filtration $\{F^{tp_j}\}$ instead of $\{F^{p_j}\}$ for positive real numbers t. So again the Harder-Narasimhan filtration is stable under this replacement. It follows that in order to prove semistability we may always assume that the p_j are integers. By the first remark we may replace the set p_j, $0 \leqq j \leqq m+1$, by a new set p'_j of rational numbers close to p_j and then by the second remark we may multiply by the least common denominator to obtain a filtration whose indexes are rational integers. This does not change the Harder-Narasimhan filtration.

We come now to the main result of this section.

Theorem 4.1 *If V_1 and V_2 are semistable then also the tensor product $V_1 \otimes V_2$ is semistable.*

Before we prove the result we discuss briefly the strategy for the proof. It consists of using the corresponding result for vector bundles over algebraic curves obtained by Narasimhan and Seshadri (see [NS]). We start with a filtered vector space over K and construct an algebraic curve Y over K and a vector bundle $\mathscr{E}$ on Y such that the μ-invariant on V and the μ-invariant for the vector bundle are the same. We recall that the latter is given by

$$\mu(\mathscr{E}) = \deg(\mathscr{E})/\operatorname{rank}(\mathscr{E}) = \int_{Y(\mathbb{C})} c_1(\mathscr{E})/\operatorname{rank}(\mathscr{E})$$

where $c_1(\mathscr{E})$ denotes the first Chern class of $\mathscr{E}$. Then we have to relate the corresponding Harder-Narasimhan filtrations.

Proof of the theorem. We have seen above that we may assume that the field is algebraically closed. But then we have to work with several filtrations $\{F_v^{p_j}\}$ with $v = 1, \ldots, l$ instead of one single. Furthermore we may assume that the real numbers p_j, $0 \leqq j \leqq m+1$, are all integers. We choose l pairwise different points $x_1, \ldots, x_l$ on the projective line $X = \mathbb{P}^1_K$ over K and fix a positive integer N. Let $\pi\colon Y \to X$ be a cyclic cover of X of degree N totally ramified over $x_1, \ldots, x_l$. Let $y_j = \pi^{-1} x_j$, $1 \leqq j \leqq l$. Starting with V we construct a vector bundle $\mathscr{E}(V)$ over Y as follows. The stalk of $\mathscr{E}(V)$ at $y \neq y_1, \ldots, y_l$ is given by $\mathscr{E}(V)_y = V \otimes \mathcal{O}_{Y,y}$. At $y = y_v$ for $v = 1, \ldots, l$ it is

$$\mathscr{E}(V)_y = \sum_{j=0}^{m} \mathfrak{m}_y^{-p_j} F_v^{p_j}(V) \otimes \mathcal{O}_{Y,y}$$

where $\mathfrak{m}_y$ denotes the maximal ideal in $\mathcal{O}_{Y,y}$. By construction the bundle $\mathscr{E}(V)$ is equivariant with respect to $\operatorname{Gal}(Y/X) = \mathbb{Z}/N\mathbb{Z}$. We calculate its degree. It has a filtration such that the associated graded bundle is given by

$$\operatorname{gr}\mathscr{E}(V) = \bigoplus_{v=1}^{l} \bigoplus_{j=0}^{m} \mathcal{O}(p_j \cdot y_v) \otimes (F_v^{p_j}(V)/F_v^{p_{j+1}}(V)).$$

Its Chern class $c_1(\mathscr{E}(V))$ is equal to $c_1(\operatorname{gr}\mathscr{E}(V)$ and given by

$$\sum_{v,j} c_1(\mathcal{O}(p_j \cdot y_v) \otimes \operatorname{gr}^j_{F_v}(V)).$$

Since

$$c_1(\mathcal{O}(p_j \cdot y_v) \otimes \mathrm{gr}^j_{F_v}(V)) = p_j \cdot \dim(\mathrm{gr}^j_{F_v}(V))\, c_1(\mathcal{O}(y_v))$$

we get by integration

$$\mu(\mathcal{E}(V)) = \sum_{v,j} p_j \dim(\mathrm{gr}^j_{F_v}(V)) \deg \mathcal{O}(y_v)/\dim V.$$

The right hand side is equal to the sum $\mu(V)$ of μ-invariants corresponding to the various filtrations $\{F_v^j\}$. We note also that the construction commutes with tensor products and clearly V is semistable if $\mathcal{E}(V)$ is semistable. This follows at once by considering submodules of $\mathcal{E}(V)$ induced form subspaces of V. Also the converse is true: First of all the Harder-Narasimhan filtration of $\mathcal{E}(V)$ is equivariant under $\mathbb{Z}/N\mathbb{Z}$. Hence it is induced from a filtration on the constant vector bundle $V \otimes \mathcal{O}_X$ on X. If a subbundle $\mathcal{E}'$ of $V \otimes \mathcal{O}_X$ is not constant its degree is at most -1. This follows from the fact that the degree decreases on non-constant subbundles. The subbundle $\mathcal{E}^*$ of $\mathcal{E}(V)$ induced by $\mathcal{E}'$ has fibres

$$\mathcal{E}^*_y = \sum_{j=0}^{m} \mathfrak{m}_y^{-p_j} F^{p_j}(\mathcal{E}'_x) \otimes \mathcal{O}_{Y,y}$$

where $x = x_v$, $y = y_v$ for $v = 1, \ldots, l$ and

$$F^{p_j}(\mathcal{E}'_x) = (F^{p_j}(V) \otimes \mathcal{O}_{X,x}) \cap \mathcal{E}'_x.$$

Similar calculations as above show that the degree of $\mathcal{E}^*$ is given by

$$\begin{aligned} \deg \mathcal{E}^* &= \text{constant} + N \deg \mathcal{E}' \\ &\leqq \text{constant} - N. \end{aligned}$$

It follows that for N big the degree becomes negative. We deduce that for N big the smallest non-trivial subbundle in the Harder-Narasimhan filtration of $\mathcal{E}$ is induced by a subspace of V. If V is semistable this subspace must be all if V. This is what we wanted to show.

The discussion above shows that if V_1 and V_2 are semistable the same holds for $\mathcal{E}(V_1)$ and $\mathcal{E}(V_2)$. By the result of Narasimhan and Seshadri (see [NS]) we know that $\mathcal{E}(V_1 \otimes V_2) = \mathcal{E}(V_1) \otimes \mathcal{E}(V_2)$ is semistable and this implies that $V_1 \otimes V_2$ is semistable which proves the theorem. □

As an application let V be a filtered vector space as above. On $V^{\otimes n}$ we obtain an induced filtration so that

$$\mu(V^{\otimes n}) = n\mu(V).$$

We consider the symmetric algebra $R = S[V]$ generated by V and assume that V is semistable. Let $I \subseteq R$ be a homogeneous ideal. Then the quotient R/I is finitely generated and graded. Further the space $(R/I)_n$ of elements of degree n is a quotient of $V^{\otimes n}$. Since $V^{\otimes n}$ is semistable we have

$$\mu((R/I)_n) \geqq \mu(V^{\otimes n}) = n\mu(V).$$

The filtration on $V^{\otimes n}$ induces a filtration $\{F^p\}$ on the quotient R/I of the tensor algebra. Associated with it we obtain a sequence of measures ρ_n as in Sect. 1 converging to ρ_∞. Since

$$\rho_n(p/n) = \dim(F_n^p/F_n^{p'})/\dim((R/I)_n).$$

and

$$\mu((R/I)_n) = \sum p \dim(F_n^p/F_n^{p'})/\dim((R/I)_n),$$

where $F^p \supset F^{p'}$ is any step in the filtration, we see that

$$E(\rho_n) = \frac{1}{n}\,\mu((R/I)_n)$$

for all n. In the limit for $n \to \infty$ we get

$$E(\rho_\infty) = \lim_{n\to\infty} \frac{1}{n}\,\mu((R/I)_n) \geqq \mu(V).$$

In this way we have found an easy way to calculate a lower bound for $E(\rho_\infty)$.

The same considerations apply when one starts with a finite extension $L \supseteq K$, a vector space V over K and a filtration on $W = V \bigotimes_K L$ as before. One obtains again a probability measure ρ_∞ with $E(\rho_\infty) \geqq \mu(V)$.

5 The law of large numbers

Let $d\rho$ be a probability measure on the real line supported on the interval I of finite length. For any measurable set S we put $\rho(S) = \int_S d\rho$. The expectation value and variance are denoted by $E(\rho)$ and $\mathrm{Var}(\rho)$. Further for $\varepsilon > 0$ we let $I(\varepsilon)$ be the set defined by $|X - E(\rho)| \geqq \varepsilon$ and $\chi_{I(\varepsilon)}$ its characteristic function. Then

$$\chi_{I(\varepsilon)} \leqq \frac{(X-E(\rho))^2}{\varepsilon^2}\,\chi_{I(\varepsilon)} \leqq \frac{(X-E(\rho))^2}{\varepsilon^2}$$

so that we obtain Tschebyscheff's inequality

$$\rho(I(\varepsilon)) \leqq \int \frac{(X-E(\rho))^2}{\varepsilon^2}\,d\rho = \frac{1}{\varepsilon^2}\,\mathrm{Var}(\rho).$$

More generally let $d\rho_1, \ldots, d\rho_m$ be probability measures as before. They define a product measure $d\rho_1 \otimes \ldots \otimes d\rho_m$ on the product space. Its convolution $d\rho = d\rho_1 * \ldots * d\rho_m$ is the direct image under the addition map $\mu\colon \mathbb{R}^m \to \mathbb{R}$ and has support on the interval $\mu(I^m)$. Its expectation value is given by

$$\begin{aligned} E(\rho) &= \int X \mu_*\, d\rho_1 \otimes \ldots \otimes d\rho_m \\ &= \int \mu^*(X)\, d\rho_1 \otimes \ldots \otimes d\rho_m \\ &= E(\rho_1) + \ldots + E(\rho_m). \end{aligned}$$

Similarly its variance is given by

$$V(\rho)=V(\rho_1)+\ldots+V(\rho_m)$$

and Tschebyscheff can be written as

$$\rho(I(m\varepsilon))\leqq\frac{V(\rho)}{m\varepsilon^2}\cdot\frac{1}{m}.$$

In applications the first factor is bounded independently of m so that the left hand side tends to zero with $1/m$ when m tends to infinity. We call this observation the law of large numbers.

Our result is applied in the following situation. Let $X\subseteq\mathbb{P}^n$ be a projective variety, $\Gamma_d:=\Gamma(X,\mathcal{O}(d))$ and $\{F^p\}$ a filtration on $R=\bigoplus_{d\geqq 0}\Gamma_d$ as in Sect. 2 with $p\geqq 0$ real. Accordingly we obtain a probability measure $d\rho_d$ which converges to a limit measure $d\rho_\infty$. More generally put $X=X_1\times\ldots\times X_m$ with $X_j\subseteq\mathbb{P}^n$, $R_j=\bigoplus_{d_j\geqq 0}\Gamma(X_j,\mathcal{O}(d_j))$ and let $\{F_j^p\}$ be a filtration on R_j. We consider

$$\Gamma_{d_1,\ldots,d_m}=\Gamma(X,\mathcal{O}(d_1,\ldots,d_m))$$

and put

$$R=\bigoplus\Gamma_{d_1,\ldots,d_m}.$$

Clearly $R=R_1\cdot\ldots\cdot R_m$ is the symmetric product of the rings $R_1,\ldots,R_m$. On R we take the product filtration $\{F^p\}$ such that F^p is generated by elements of the form $r_1\cdot\ldots\cdot r_m$ with $r_j\in F_j^{p_j}$ and $p_1+\ldots+p_m\geqq p$. By the definition of the probability measures we have

$$\dim(F^\xi\Gamma_{d_1,\ldots,d_m})/\dim(\Gamma_{d_1,\ldots,d_m})=\int_\xi^\infty d\rho_{1,d_1}*\ldots*d\rho_{m,d_m}$$

for any ξ in the support of the convolution measure. By Lemma 3.2 we have

$$\rho_{j,d_j}=\rho_{j,\infty}+o(1/d_j)$$

in the weak topology. Hence

$$\int_\xi^\infty d\rho_{1,d_1}*\ldots*d\rho_{m,d_m}=\int_\xi^\infty d\rho_{1,\infty}*\ldots*d\rho_{m,\infty}+o(1/d_1)+\ldots+o(1/d_m).$$

If $\varepsilon>0$ is any real number the law of large numbers implies that for $|E-X|\leqq m\varepsilon$ we have

$$\int_{|E-X|\leqq m\varepsilon} d\rho_{1,\infty}*\ldots*d\rho_{m,\infty}\geqq 1-\frac{\operatorname{Var}\rho}{m\varepsilon^2}\frac{1}{m}.$$

On noting that

$$\operatorname{Var}(\rho)=\operatorname{Var}(\rho_1)+\ldots+\operatorname{Var}(\rho_m)\leqq m\cdot\max(\operatorname{Var}(\rho_j))$$

we find that the right hand side of the inequality above is at least

$$1-\max(\mathrm{Var}(\rho_j))/\varepsilon^2 m.$$

Furthermore if $d_1, \ldots, d_m$ are sufficiently large we see that the following proposition holds.

Proposition 5.1 *For $d_1, \ldots, d_m$ sufficiently large we have*

$$\dim(F^{E(\rho)-m\varepsilon}\Gamma_{d_1,\ldots,d_m}/F^{E(\rho)+m\varepsilon}\Gamma_{d_1,\ldots,d_m}) \geqq (1-c/m)\dim(\Gamma_{d_1,\ldots,d_m})$$

where c is a positive constant independent of m.

6 Siegel's lemma

In the construction of the auxiliary polynimial we shall need a modification of the results obtained by Bombieri and Vaaler in their paper [BV] which is the general reference for this section. Let L be a number field of degree r over the rationals, N a positive integer and $W \subset L^N$ a subspace of dimension $M<N$ which we assume to be generated by elements with height at most H. Then by Theorem 9 together with (2.6) in the paper quoted above we find a basis $w_1, \ldots, w_{N-M}$ of the orthogonal complement $W^\perp$ of W in $(\mathcal{O}_L)^N$ such that

$$\prod_{l=1}^{N-M} H(w_l) \leqq C(L)^{N-M} H^M$$

for some constant $C(L)$ which can be given explicitly. It depends only on L. Here $\mathcal{O}_L$ denotes as usual the ring of integers of L.

More generally let $W_1, \ldots, W_s$ be such subspaces and $M_1, \ldots, M_s$ their dimensions. Then $W_j^\perp$ has a basis w_{jl} in $(\mathcal{O}_L)^N$, $l=1, \ldots, N-M_j$, such that the above inequalities hold with w_l, M and H replaced by w_{jl}, M_j an H_j respectively. Let $y_1, \ldots, y_M$ be a maximal linearly independent subset of the vectors w_{jl}, $1 \leqq j \leqq s$, $1 \leqq l \leqq M_j$. Then, since the height is at least 1,

$$\prod_{k=1}^{M} H(y_k) \leqq \prod_{j=1}^{s} \prod_{l=1}^{N-M} H(w_{jl}) \leqq C(L)^{\Sigma(N-M_j)} \prod_{j=1}^{s} H_j^{M_j}.$$

The orthogonal complement of the space generated by $y_1, \ldots, y_M$ is $W_1 \cap \ldots \cap W_s$. If $rM<N$ then by Theorem 12 in [BV] we get $N-rM$ linearly independent vectors $x_1, \ldots, x_{N-rM}$ in $\mathbb{Z}^N \cap W_1 \cap \ldots \cap W_s$ such that

$$\prod_{l=1}^{N-rM} H(x_l) \leqq C(L)^{N-rM+\Sigma(N-M_j)} \prod_{j=1}^{s} H_j^{M_j}.$$

This leads to the following version of Siegel's lemma.

Proposition 6.1 *Suppose that the Codimension M of the subspace $W_1 \cap \ldots \cap W_s$ satisfies $0 < rM < N$. Then there exists an element $0 \neq x \in \mathbb{Z}^N \cap W_1 \cap \ldots \cap W_s$ such that*

$$\max_j(|x_j|) \leqq C \cdot (H_1^{M_1} \ldots H_s^{M_s})^{1/(N-rM)}.$$

Here C is a positive constant which depends only on the field L.

7 Differential operators

Let $X_K \subseteq \mathbb{P}^n$ be a projective variety of dimension d defined over a number field K with ring of integers $\mathcal{O}_K$. We denote by X the closure of X_K in $\mathbb{P}^n$ over $\mathcal{O}_K$ and we consider a good projection π: $X \to P = \mathbb{P}^d$. This is a rational map which is not defined at the intersections of X with the centers of projections. However there exists a modification $\tilde{X} \to X$ such that π becomes regular on $\tilde{X}$. By Corollary 2.14 of [Fa] there also exists a homogeneous polynomial G of degree at most $(n-d)$ $[K:\mathbb{Q}]$ $\deg(X)$ such that G annihilates the module of relative differentials $\Omega^1_{\tilde{X}/P}$. Moreover the polynomial G can be choosen in such a way that the coefficients of G are rational integers with absolute values at most equal to $\exp(c_1 h(X) + c_2)$ and such that G does not vanish identically on X. Dualizing and taking global sections we see that for every derivation ∂ on P the derivation $G\partial$ extends to a derivation $\widetilde{G\partial}$ on $\tilde{X}$ such that $\pi_* \tilde{\partial} = G\partial$.

We assume now that all associated prime ideals of X have characteristic 0. Let x be a point in $X(\mathcal{O}_K)$ and $\partial_1, \ldots, \partial_l$ any derivations on X. For non-negative integers $\tau_1, \ldots, \tau_l$ with $\tau_1 + \ldots + \tau_l = \tau$ we consider the differential operator

$$D = (\partial_1^{\tau_1}/\tau_1!) \ldots (\partial_l^{\tau_l}/\tau_l!).$$

Let J be the ideal sheaf of x, $d \geqq 0$ an integer and $f \in \Gamma(X, J^\tau \cdot \mathcal{O}(d))$. Then $D(f)(x)$ is a well-defined section of the fibre of $\mathcal{O}(d)$ at x.

Lemma 7.1 *The section $D(f)(x)$ is in $\mathcal{O}_K$.*

Proof. We choose an affine neighborhood U of x in X. Then $U = \operatorname{Spec} R$ where $R = \mathcal{O}_K[T_1, \ldots, T_n]/I$ for some ideal I. Here $T_1, \ldots, T_n$ are independent variables. The derivations ∂_j on U can be lifted to derivations $\tilde{\partial}_j$ of $\mathcal{O}_K[T_1, \ldots, T_n]$ which we can express as

$$\tilde{\partial}_j = \sum a_{ij}\, \partial/\partial T_i$$

with $a_{ij} \in \mathcal{O}_K[T_1, \ldots, T_n]$. They satisfy $\tilde{\partial}_j(I) = 0$. In this way D lifts to a differential operator $\tilde{D}$ on $\mathcal{O}_K[T_1, \ldots, T_n]$. Over U the line bundle $\mathcal{O}(d)$ becomes trivial so that $\mathcal{O}(d) = \mathcal{O}_U \cdot h$ for some section h over U. Hence we may write f in the form $f = \varphi \cdot h$ for some $\varphi \in R$. Again φ can be lifted to an element $\tilde{\varphi}$ in $\mathcal{O}_K[T_1, \ldots, T_n]$. Since f is in $\Gamma(X, J^\tau \cdot \mathcal{O}(d))$ we find that φ is in I_x^τ where I_x is the ideal of x in R. If $\tilde{I}_x$ denotes the ideal of x in $\mathcal{O}_K[T_1, \ldots, T_n]$ then plainly $\tilde{\varphi} \in \tilde{I}_x^\tau$. We may assume also that for $j = 1, \ldots, n$ we have $T_j(x) = 0$. If this is not the case we just replace T_j by $T_j - T_j(x)$, $1 \leqq j \leqq n$. Therefore we can write $\tilde{\varphi}$ as

$$\tilde{\varphi} = \sum b_{i_1, \ldots, i_n} T_1^{i_1} \ldots T_n^{i_n}$$

where $b_{i_1,\ldots,i_n} \in \mathcal{O}_K$ and the sum is over all sets $i_1, \ldots, i_n$ with $i_1 + \ldots + i_n \geqq \tau$. It follows that $\tilde{D}(\tilde{\varphi})(x)$ is a linear combination of the $b_{i_1,\ldots,i_n}$ with $i_1 + \ldots + i_n = \tau$. Hence

$$Df(x) = \tilde{D}(\tilde{\varphi})(x) \cdot h(x) \in \mathcal{O}_K. \qquad \square$$

We also need a sharpening of the previous lemma. For this let again $\pi: X \to P$ be a good projection and F the norm of G. Then F defines a subvariety $Y \subset P$ such that π is étale on $X \setminus \pi^{-1}(Y)$. We assume that $x \in X(K)$ and that $\pi(x) = y \notin Y$. Then for every place v of K the v-adic distance $d_v(y, Y)$ from y to Y satisfies

$$d_v(y, Y) \gg \|F(y)\|_v / \|F\|_v.$$

Hence there exists a ball B_v around y of radius $\rho_v = d_v(y, Y)$ such that on the set $Z = \pi^{-1}(B_v)$ the morphism π is étale. It follows that over B_v there exists a section σ of π with $\sigma(y) = x$. We put $\tilde{B} = \sigma(B_v)$. Then over $\tilde{B}$ the line bundle $\mathcal{O}(1)$ becomes trivial. We may choose a section h of the restriction of $\mathcal{O}(1)$ to $\tilde{B}$ such that on $\tilde{B}$ we have $\mathcal{O}(1) = \mathcal{O}_{\tilde{B}} \cdot h$. This section can be choosen in such a way that its norm satisfies $1 \leqq \|h\| \leqq 2$. As a consequence every section of $\mathcal{O}(1)$ over $\tilde{B}$ can be written as

$$f = \varphi \cdot h$$

with some analytic function φ such that

$$|\varphi|_v \leqq \| \mathrm{f} \|_v.$$

For a set $a_1, \ldots, a_d$ of non-negative integers we put $\mathbf{a} = (a_1, \ldots, a_d)$ and define the differential operator

$$\Delta(\mathbf{a}) = ((\partial/\partial z_1)^{a_1}/a_1!) \ldots ((\partial/\partial z_d)^{a_d}/a_d!).$$

Lemma 7.2 *Let $\varphi_1, \ldots, \varphi_l$ be analytic in the unit multidisc of dimension d with supremum norm $\|\varphi_j\| \leqq 1$, $j = 1, \ldots, l$. For non-negative integers $b_1, \ldots, b_l$ with $b = b_1 + \ldots + b_l$ we consider the product $\varphi = \varphi_1^{b_1} \ldots \varphi_l^{b_l}$. Then, if $a = a_1 + \ldots + a_d$, we have*

$$|\Delta(\mathbf{a})\, \varphi(0)| \leqq 2^{db+a} \sum \prod |\varphi_j(0)|^{b_j - e_j}|$$

where the sum is over all sets of non-negative integers $a_{11}, \ldots, a_{ld}$ with $a_{1j} + \ldots + a_{lj} = a_j$ and the product is over all j with $a_{j1} + \ldots + a_{jd} = e_j \leqq b_j$.

Proof. The proof follows from Lemma 6.2 in [Fa]. First one observes that it follows from the hypothesis that

$$|((\partial/\partial z)^a/a!)\, \varphi^b(0)| \leqq 1$$

if $a > b$. Then one proves the result in the case $l = 1$ by induction on the number of variables. In the case $l > 1$ the result follows using Leibnitz' rule. $\square$

We remark that the inequality of Lemma 7.2 also holds for p-adic analytic functions in even a sharper form. In fact the factor 2^{db+a} can be omitted and the sum can be replaced by the maximum. Also we remark that a similar result holds when φ is replaces by a product $\varphi_1 \cdot \varphi$ with some holomorphic function φ_1 such that the supremum norm is bounded by 1. The only change is that the sum is over all sets of non-negative integers $a_{11}, \ldots, a_{ld}$ with $a_{1j}+\ldots+a_{lj} \leqq a_j$, $1 \leqq j \leqq d$. Again the remarks above apply in the non-archimedean case. Also all the theory holds with $\mathcal{O}(1)$ replaced by $\mathcal{O}(d)$ for some $d \geqq 1$. Further it can be extended easily to schemes of the form $X = X_1 \times \ldots \times X_m$ and line bundles of the shape $\mathcal{O}(d_1, \ldots, d_m)$.

8 The main result

We shall now apply the techniques developed in the preceding sections. Let $K \subseteq L$ be number fields and let $\mathscr{P}$ be a finite set of places of L. For each $w \in \mathscr{P}$ we fix a finite index set I_w and choose for each $\alpha \in I_w$ a non-zero homogeneous polynomial $f_{w,\alpha}$ in $n+1$ variables $T_0, \ldots, T_n$ with coefficients in L. For each pair w, α with $w \in \mathscr{P}$ and $\alpha \in I_w$ we let $c_{w,\alpha} \geqq 0$ be a real number. Finally for each place w of L we fix a w-adic norm on the line bundle $\mathcal{O}(1)$ which extends the standard p-adic or real norm. The norms on $\mathcal{O}(1)$ induce norms on the powers of $\mathcal{O}(1)$ which we denote all by $\|\cdot\|_w$. We intend to show that there are only finitely many K-rational points $x = (x_0 : \ldots : x_n)$ in $\mathbb{P}^n(K)$ which satisfy the system of inequalities

$$(1) \qquad \|f_{w,\alpha}\|_w < H(x)^{-c_{w,\alpha}} \quad w \in \mathscr{P}, \alpha \in I_w.$$

We shall first consider the case that all $f_{w,\alpha}$ are linear forms and we put $V = \Gamma(\mathbb{P}^n_K, \mathcal{O}(1))$ and $V_L = V \underset{K}{\otimes} L$. For any fixed place w in $\mathscr{P}$ and any positive real number p we consider the subspace of V_L generated by the linear forms $f_{w,\alpha}$ for which $c_{w,\alpha} \geqq p$. In this way we get a finite set of subspaces

$$V_L = W^0 \supset W^1 \supset \ldots \supset W^e \supset W^{e+1} = 0$$

of V_L. We define $p_j = p_{w,j}$ as the minimum of the numbers $c_{w,\alpha}$ taken over the indices α given by the generators of W^j for $0 \leqq j \leqq e$. If W^0 is not generated by the $f_{w,\alpha}$, $\alpha \in I_w$, we put $p_0 = p_{w,0} = 0$ and also we put $p_{e+1} = p_e + 1$. In this way we get for each $w \in \mathscr{P}$ a filtration as in Sect. 3. For $w \notin \mathscr{P}$ we take the trivial filtration given by

$$V_L = F_w^0 \supset F_w^1 = 0.$$

The results of Sect. 2 apply and we obtain for each place w of L an invariant $\mu_w = \mu_w(V)$ with $\mu_w = 0$ for $w \notin \mathscr{P}$. We assume now that the filtrations are jointly semistable. This means that for each non-zero proper subspace $W \subset V$ we have

$$\sum_w \mu_w(W) \leqq \sum_w \mu_w(V).$$

Here the sum is over all places of w counted with multiplicities according to the conventions we made in Sect. 1.

Theorem 8.1 *Assume that all $f_{w,\alpha}$ are linear and that they define a jointly semistable filtration on V. Assume furthermore that $\sum_w \mu_w(V) > [L:\mathbb{Q}]$ where the sum is over all places of L counted with multiplicities. Then the number of points $x \in \mathbb{P}^n(K)$ with* (1) *is finite.*

Let us remark that the result is sharp in the sense that (1) always has infinitely many solutions if $\sum_w \mu_w(V) < [L:\mathbb{Q}]$. This follows from an elementary counting argument as in the easy part of Theorem 2A in [S2].

Before we prove the result we discuss briefly how the hypothesis of the theorem behaves with respect to finite field extensions $L' \supseteq L$. Let $\mathscr{P}'$ be the set of places w' of L' which extend the places w of L in $\mathscr{P}$. We define new linear forms and exponents by $f_{w',\alpha} = f_{w,\alpha}$ and $c_{w',\alpha} = c_{w,\alpha}$ if w' divides w. Then

$$\|f_{w',\alpha}(x)\|_{w'} = \|f_{w,\alpha}(x)\|_{w'} = \|f_{w',\alpha}(x)\|_w < H(x)^{-c_{w',\alpha}}$$

for $w' \in \mathscr{P}'$ and $\alpha \in I_w = I_{w'}$. Furthermore we have

$$\sum \mu_{w'}(V) = [L':L] \cdot \sum \mu_w(V)$$

so that the left hand side exceeds $[L':\mathbb{Q}]$ if and only if the second factor on the right hand side exceeds $[L:\mathbb{Q}]$. We may therefore freely make finite field extensions without changing the hypothesis.

We begin now with the proof of the theorem and assume that there are infinitely many solutions. Let $\varepsilon > 0$ be a sufficiently small real number and r sufficiently large as in Faltings' product theorem. Then we choose a set $x_1, \ldots, x_m$ of solutions of (1) with the following properties. If h_j denotes the logarithmic Weil height of x_j we require that h_1 is sufficiently large and that

(2) $$h_{j+1} \geqq r h_j, \quad j = 1, \ldots, m-1.$$

Depending on this choice of solutions we choose positive integers $d_1, \ldots, d_m$. To begin with let d_1 be so large that

(3) $$\varepsilon d_1 h_1 \geqq h_m$$

and define d_j for $j = 2, \ldots, m$ by

(4) $$d_j = \left[\frac{d_1 h_1}{h_j}\right] + 1.$$

Then clearly

$$d_1 h_1 \leqq d_j h_j \leqq (1+\varepsilon) d_1 h_1$$

and, as a consequence,

$$d_{j+1} h_{j+1} \leqq (1+\varepsilon) d_j h_j$$

for $1 \leqq j \leqq m-1$.

Inductive proposition 8.2 *For each j with $0 \leqq j \leqq mn$ there exist positive real numbers γ, c_1, c_2 and a product variety*

$$Z = Z_1 \times \ldots \times Z_m \subseteq (\mathbb{P}^n)^m$$

of dimension at most $mn - j$ defined over K which contains $(x_1, \ldots, x_m)$. Furthermore the degree of Z_i is at most c_1 for $1 \leqq i \leqq m$ and

(5) $$d_1 h(Z_1) + \ldots + d_m h(Z_m) \leqq c_2 (d_1 + \ldots + d_m)$$

where $h(Z_i)$ is the height of Z_i, $1 \leqq i \leqq m$.

The inductive proposition implies the theorem. Namely for $j=mn$ we get $Z=(x_1, \ldots, x_m)$ and (5) implies that

$$h_i = h(x_i) \leqq c_3 \frac{h_i}{h_1}, \qquad 1 \leqq i \leqq m,$$

for some positive constant c_3. It follows that $h_1 \leqq c_3$ which contradicts the choice of x_1.

We turn now to the proof of the inductive proposition. For $j=0$ we may put $\gamma = c_1 = c_2 = 1$ and $Z_i = \mathbb{P}^n$ for $1 \leqq i \leqq m$. Therefore we may assume that the inductive proposition is proved for j and we are going to verify it for $j+1$. Also we may assume that $(x_1, \ldots, x_m)$ is a smooth point.

If the product variety Z constructed for j has dimension strictly less than $nm-j$ then Z also satisfies the inductive proposition for $j+1$ with the same choice of constants. Otherwise $\dim Z = mn-j$ and we put $X_i = Z_i$, $i=1, \ldots, m$, and $X = X_1 \times \ldots \times X_m$. Let $u_0, \ldots, u_n$ be homogeneous coordinates on $\mathbb{P}^n$ and let x_{ij} be the restriction of $\mathrm{pr}_i^* u_j$ to X. Our aim is to construct a global section $0 \neq f \in \Gamma(X \underset{K}{\otimes} L, \mathcal{O}(d_1, \ldots, d_m))$ such that f has sufficiently big positive index at $x = (x_1, \ldots, x_m)$. To achieve this the section f is constructed in such a way that for each $w \in \mathcal{P}$ it can be written as a linear combination with coefficients in $\mathbb{Z}$ of monomials of the form

$$(6) \qquad \prod_{i,j} x_{ij}^{k_{w,i,j}} \cdot \prod_{i,\alpha} \mathrm{pr}_i^* (f_{w,\alpha})^{j_{w,\alpha,i}}$$

such that

$$(7) \qquad \sum_{i,\alpha} j_{w,\alpha,i} \cdot c_{w,\alpha}/d_i \geqq m v_w$$

where the v_w are real numbers with $\sum \mu_w > \sum v_w = r > [L:\mathbb{Q}]$. We also want that the coefficients of the monomials in (6) have absolute values at most $\exp(c_4(d_1 + \ldots + d_m))$. This will be a consequence of Siegel's lemma. In order to apply it we have to count the conditions and unknowns by the techniques from Sect. 3. In our situation the filtration F_w is induced on the homogeneous coordinate ring of X from the filtrations on the homogeneous coordinate rings of X_i, $i=1, \ldots, m$. They are obtained from the linear forms $f_{w,\alpha}$ in the way as described at the beginning of the section.

In Sect. 1 we have associated with these data a probability measure $\rho_{\infty,i,w}$ for each pair i, w with $1 \leqq i \leqq m$, $w \in \mathcal{P}$. The stability assumption implies that the sum of the expectation values of all $\rho_{\infty,i,w}$ is at least $\sum \mu_w > r$. Choose $\delta > 0$ sufficiently small and let E_w denote the sum over all i with $1 \leqq i \leqq m$ of the expectation values of $\rho_{\infty,i,w}$. If we put

$$\Gamma = \Gamma(X \underset{K}{\otimes} L, \mathcal{O}(d_1, \ldots, d_m))$$

we see from Proposition 6.1 that the space $V_w = F_w^{E_w - m\delta} \Gamma$ has dimension at least

$$(1 - O(1/m)) \dim \Gamma.$$

Taking the intersection W of all V_w for $w \in \mathscr{P}$ we see that the dimension of W is at least

$$(1-\delta)\dim \Gamma$$

provided that m is sufficiently large.

By Siegel's lemma we find a non-zero section $f \in W$. It is contained in $F^{E_w - m\delta}\Gamma$ for each $w \in \mathscr{P}$ so that (7) is satisfied since δ was choosen small enough.

We shall next show that the section f has a sufficiently big positive index at $x=(x_1, \ldots, x_m)$. We fix a place $w \in \mathscr{P}$. By construction f is in $F^{E_w - m\delta}\Gamma$.

Now we consider good projections $\pi_j: X_j \to P_j = \mathbb{P}^{d_j}$, $j=1, \ldots, m$, and obtain a projection $\pi: X \to P'$ where $P' = P_1 \times \ldots \times P_m$. The theory in Sect. 7 applies and we get for each j with $1 \leqq j \leqq m$ a homogeneous polynomial F_j wich defines a subscheme $Y_j \subset P_j$ such that π_j is étale outside Y_j. If $\pi(x_i) \in Y_i$ for some i we replace X_i by an irreducible component Z_i of $\pi_i^{-1}(Y_i)$ which contains x_i. For $j \neq i$ we put $Z_j = X_j$. Then $Z = Z_1 \times \ldots \times Z_m$ satisfies the hypothesis of the inductive proposition with constants c_1 and c_2 which can be easily determined. Therefore we may assume that X_j is smooth in x_j and that π_j is étale in x_j for all j. Hence we get balls $B_{j,w}$ for $w \in \mathscr{P}$ around $y_j = \pi_j(x_j)$ and sections σ_j of π_j over $B_{j,w}$. The radius of the ball $B_{j,w}$ can be taken equal to the distance $d_w(y_j, Y_j)$ which is at least $\exp(c_5\, h(x_j) + c_6)$ for some effective constants c_5 and c_6. We put $\tilde{B}_j = \sigma_j(B_{j,\omega})$, $\tilde{B} = \tilde{B}_1 \times \ldots \times \tilde{B}_m$ and $B = B_{1,w} \times \ldots \times B_{m,w}$. We choose $\kappa > 0$ sufficiently small compared with ε and we want to show that the index σ of the section f is at least κ. For this we consider a differential operator on B of the form

$$D = \prod((\partial/\partial z_{ij})^{t_{ij}}/t_{ij}!)$$

where the t_{ij} are non-negative integers such that with $t_j = \sum t_{ij}$ we have

$$\frac{t_1}{d_1} + \ldots + \frac{t_m}{d_m} = \sigma.$$

We multiply D by $G_1^{t_1} \ldots G_m^{t_m}$ and then the resulting differential operator $\tilde{D}$ extends to X. The restriction of $\tilde{D} f$ to x is a well-defined integral section of the fibre of $F_w^{E_w - m\delta - c_w \sigma}$ over x where we have put $c_w = \max_\alpha(c_{w,\alpha})$. Every differential operator on X can be expressed as a linear combination of such differential operators.

We now write $(\tilde{D} f)(x)$ as a linear combination of monomials of the form

$$\prod_{\alpha, i} \mathrm{pr}_i^* f_{w,\alpha}(x)^{j'_{w,\alpha,i}}$$

for non-negative integers $j'_{w,\alpha,i}$ with

$$\sum j'_{w,\alpha,i}\, c_{w,\alpha}/d_i \geqq E_w - m\delta - c_w\, \sigma.$$

For $w \in \mathscr{P}$ the norms of these monomials at w do not exeed

$$\prod_i H(x_i)^{-\sum j'_{w,\alpha,i} c_{w,\alpha}}.$$

If we replace $H(x_i)$ by $H(x_1)^{d_i/d_1}$ the product is at most equal to

$$H(x_1)^{-(E_w - m\delta - c_w \sigma) d_1}.$$

The absolute values of the coefficients of the monomials are bounded by $\exp(c_7(d_1+\ldots+d_m))$. For $w \notin \mathscr{P}$ the absolute values of $(\tilde{D}f)(x)$ are at most $\exp(c_8(d_1+\ldots+d_m))$ for the archimedean places and at most 1 for the non-archimedean places. Here we use Lemma 7.1.

If we take the product over all absolute values of $(\tilde{D}f)(x)$ we find that for σ and δ sufficiently small

$$1 \leqq H(x_1)^{([L:\mathbb{Q}]+\varepsilon-\sum_w E_w)\, d_1}.$$

The right hand side is less than one provided that $\sum E_w > [L:\mathbb{Q}]+\varepsilon$ so that we finally find that σ is bounded below by some positive multiple $c_9\,\varepsilon$ of ε. Since by semistability $\sum E_w \geqq \sum \mu_w$ the hypothesis shows that the above inequality holds.

We apply now the product theorem in the way as explained at the end of Sect. 1 and find a product variety

$$Z = Z_1 \times \ldots \times Z_m \subseteq (\mathbb{P}^n)^m$$

which has the required properties. This finishes the proof of the inductive proposition and also completes the proof of Theorem 8.1.

Let us remark that we proved actually the more general result that for any subscheme $X \subseteq \mathbb{P}^n$ the set of $x \in X(K)$ with (1) is finite. This remark will be useful in the further applications.

9 Further results and comments

We shall now discuss first what happens without the assumption on the semistability. Let $V = KT_0 \oplus \ldots \oplus KT_n$ with independent variables $T_0, \ldots, T_n$ and with the filtrations given by the linear forms $f_{w,\alpha}$, $w \in \mathscr{P}$, $\alpha \in I_w$, on $V_L = V \underset{K}{\otimes} L$. We consider the first non-trivial step in the Harder-Narasimhan filtration of V. It is given by a semistable subspace $W \subseteq V$. Let $\mathbb{P}^n = \mathbb{P}(V)$ be the projective space consisting of one dimensional quotient spaces of V. The subspace W corresponds to a projection

$$\pi: \mathbb{P}(V) \setminus \mathbb{P}(V/W) \to \mathbb{P}(W)$$

from an open set of $\mathbb{P}(V)$. The induced filtration on W is jointly semistable. We consider the set of $f_{w,\alpha}$ which lie in $W \underset{K}{\otimes} L$. We assume that $\mu(W) \geqq [L:\mathbb{Q}]$. Then there are only finitely many points $\pi(x) \in \mathbb{P}(W)(K)$. For x running throught the K-rational points in a fixed fibre $y = \pi(x)$ the height tends to infinity. On the other hand there is some $f_{w,\alpha} \in W \underset{K}{\otimes} L$ with $c_{w,\alpha} > 0$ which is a non-zero constant on the fibre of y because of semistability of W. Therefore there are only finitely many K-rational points in each fibre for which the projection is defined and (1) holds.

For the remaining points one applies induction to the quotient V/W. We obtain therefore

Theorem 9.1 *Assume that all $f_{w,\alpha}$, $w\in\mathscr{P}$, $\alpha\in I_w$, are linear and let $W\subseteq V$ denote the first non-trivial step in the Harder-Narasimhan filtration associated with the filtration given by the $f_{w,\alpha}$. Assume furthermore that*

$$\sum \mu_v(W) > [L:\mathbb{Q}]$$

where the sum is over all places of L counted with multiplicities. Then there are only finitely many K-rational points $x\in\mathbb{P}(V)(K)$ with (1) which are not in $\mathbb{P}(V/W)(K)$.

By the construction of W we always have $\mu(W)\geqq\mu(V)$ so that in particular the assertion of Theorem 9.1 holds for rational points x in the complement of a proper linear subspace. This leads to the following version of Schmidt's subspace theorem.

Corollary 9.2 *There are finitely many proper subspace $W\subset V$ such that the set of K-rational points $x\in\mathbb{P}^n(K)$ which are not in $\mathbb{P}(V/W)$ and which satisfy*

$$\prod_{w\in\mathscr{P}}\prod_{\alpha\in I_w}\|f_{w,\alpha}(x)\|_w < H(x)^{-(n+1)[L:\mathbb{Q}]-\delta}$$

is finite.

Proof. Put $\delta=2\varepsilon$. If the set is infinite we find real numbers $c_{w,\alpha}$, $w\in\mathscr{P}$, $\alpha\in I_w$, such that $\sum c_{w,\alpha}\geqq(n+1)[L:\mathbb{Q}]+\varepsilon$ and infinitely many K-rational points in the set such that

$$\|f_{w,\alpha}(x)\|_w < H(x)^{-c_{w,\alpha}},\qquad w\in\mathscr{P},\alpha\in I_w.$$

We find that

$$\mu(V)=\sum c_{w,\alpha}/(n+1) > [L:\mathbb{Q}]$$

and this contradicts Theorem 9.1. □

From the corollary one easily deduces Schlickewei's extension of Schmidt's subspace theorem which also includes finite places.

Let us explain how Theorem 8.1 implies Roth's theorem. We choose an algebraic number α not in $\mathbb{Q}$ and consider the linear form

$$f=T_0-\alpha T_1$$

in two variables T_0 and T_1. For L we take any finite Galois extension of $K=\mathbb{Q}$ which contains α. For one infinite place w of L we let f_w be equal to f and κ be any real number greater than 2. The other infinite places are all conjugate to w under $\mathrm{Gal}(L/\mathbb{Q})$ and we choose the conjugate to f there. More precisely, if w' is another place of L then there exists some $\gamma\in\mathrm{Gal}(L/\mathbb{Q})$ such that

$$\|x\|_{w'}=\|\gamma^{-1}x\|_w,\qquad x\in L.$$

We choose $f_{w'}=f^\gamma=T_0-\gamma(\alpha)\,T_1$. If $x\in K$ we have

$$\|f_{w'}(x)\|_{w'}=\|f^\gamma(x)\|_{w'}=\|\gamma(f(x))\|_{w'}=\|f(x)\|_w.$$

Since α is irrational the stability assumption is satisfied and then all $\mu_w(V)$ are equal to $\kappa/2$. This implies that

$$\sum \mu_w(V) = \kappa[L:\mathbb{Q}]/2 > [L:\mathbb{Q}].$$

and Theorem 8.1 implies Roth's theorem.

Finally we turn to forms of higher degree. We shall reduce this to the case of linear forms. However here the results are probably not best possible but only the best which can be shown by our methods. The $f_{w,\alpha}$ define a filtration $\{F^p\}$ on the polynomial ring in $n+1$ variables so that we obtain measures ρ_w with expectation values $E(\rho_w)$.

Theorem 9.3 *Assume that* $\sum E(\rho_w) > [L:\mathbb{Q}]$ *where the sum is over all places of* L *with multiplicities. Then the set of points* $x \in \mathbb{P}^n(K)$ *which satisfy* (1) *is not Zariski-dense in* $\mathbb{P}^n$. *The same holds if we replace* $\mathbb{P}^n$ *by any closed subscheme and the* ρ_w *by the measures defined with respect to the coordinate ring of the subscheme.*

Proof. Without loss of generality we may assume that the forms $f_{w,\alpha}$ all have degree r. Let i: $\mathbb{P}^n \to \mathbb{P}^N$ be the r-fold Segre embedding given by monomials of degree r. Then $i^* \mathcal{O}(1) = \mathcal{O}(r)$ and $f_{w,\alpha} = i^* L_{w,\alpha}$ for sections $L_{w,\alpha}$ in $\Gamma(\mathbb{P}^N, \mathcal{O}(1))$. Furthermore the system of inequalities (1) translate into

$$\|L_{w,\alpha}(i(x))\|_w < H(i(x))^{-\frac{c_{w,\alpha}}{r}},$$

$w \in \mathscr{P}$, $\alpha \in I_w$. The linear forms $L_{w,\alpha}$ define a filtration $\{G^{p'}\}$ on

$$R = \bigoplus_{t \geqq 0} \Gamma(\mathbb{P}^n, i^* \mathcal{O}(t))$$

with weights p'. The filtrations are related by

$$G_t^{p/r} = F_{rt}^p$$

so that the measures $\rho_{r,w,t}$ corresponding to the G-filtration and the measures $\rho_{w,rt}$ corresponding to the F-filtration agree and this implies that

$$E(\rho_{r,w,t}) = \sum \frac{p}{r} \cdot \frac{1}{t} \rho_{r,w,t}\left(\frac{p}{r} \cdot \frac{1}{t}\right) = E(\rho_{w,rt}).$$

Taking limits we find that $E(\rho_{r,w,\infty}) = E(\rho_{w,\infty})$. Therefore

$$\sum E(\rho_{r,w}) = \sum E(\rho_w) > [L:\mathbb{Q}]$$

for $\rho_{r,w} = \rho_{r,w,\infty}$ and $\rho_w = \rho_{w,\infty}$. By Theorem 9.1 the result follows. $\square$

10 The unit equation

Let V be a vector space over K of dimension $n+1$ generated by $T_0, \ldots, T_n$ and S a finite set. For each $v \in S$ let w_v be a weight function defined on the set $T_0, \ldots, T_n$ and with non-negative real values such that

$$\sum_{v \in S} w_v(T_j) = 1$$

for $j=0, \ldots, n$ and

$$\min_j w_v(T_j) = 0.$$

We put $V_j = K \cdot T_j$, $0 \leqq j \leqq n$, and define filtrations $\{F_v^p\}$ for $v \in S$ by

$$F_v^p V = \bigoplus_{w_v(T_j) \geqq p} V_j.$$

Here $p \in \{w_v(T_j);\ 0 \leqq j \leqq n\}$. Each V_j is a semistable filtered vector space with $\mu(V_j) = \sum_{v \in S} w_v(T_j) = 1$. Therefore V is a semistable filtered vector space with

$$\mu(V) = \frac{1}{n+1} \sum_v \sum_j w_v(T_j) = 1.$$

We denote by p_j the vector with components $w_v(T_j)$ for $v \in S$. Then two components V_i and V_j of V are isomorphic as filtered vector spaces if and only if $p_i = p_j$. Let Σ be the set of vectors p_j. We put

$$V_\sigma = \bigoplus_{j \in I_\sigma} V_j$$

for $\sigma \in \Sigma$ where I_σ is the set of j with $p_j = \sigma$. This gives a direct decomposition

$$V = \bigoplus_{\sigma \in \Sigma} V_\sigma$$

of V into isotypic components V_σ for which we also have $\mu(V_\sigma) = 1$. For a subspace $W \subseteq V$ and for $\sigma \in \Sigma$ we define $W_\sigma = W \cap V_\sigma$. Then we have $W = \bigoplus_{\sigma \in \Sigma} W_\sigma$ if and only if $\mu(W) = 1$.

We consider the quotient $\tilde{V} = V/K(T_0 + \ldots + T_n)$. Since V is semistable all quotients of $\tilde{V}$ have μ-invariant $\geqq 1$. Therefore the slopes $\mu(\tilde{V}_{i+1}/\tilde{V}_i)$ in the Harder-Narasimhan filtration

$$\tilde{V} = \tilde{V}_k \supset \tilde{V}_{k-1} \supset \ldots \supset \tilde{V}_1 \supset \tilde{V}_0 = 0$$

are at least 1. If $\mu(\tilde{V}_{i+1}/\tilde{V}_i) = 1$ for some i then $i = k-1$ since the slopes are strictly decreasing in i. Hence $\mu(\tilde{V}/\tilde{V}_{k-1}) = 1$. But then $\mu(V/W) = 1$ if W is the inverse image of $\tilde{V}_{k-1}$ in V under the canonical projection. Since $\mu(V) = 1$ and since

$$(\dim V)\,\mu(V) = \dim(V/W)\,\mu(V/W) + \dim W/\mu(W)$$

we get $\mu(W)=1$. We decompose W into its isotypic components

$$W=\bigoplus W_\mu.$$

By definition W contains $T_0+\ldots+T_n$. It follows that W_μ contains $\sum T_i$ where the sum is over I_μ. The minimal such subspace W which can occur is the space generated by the elements $\sum_{j\in I_\sigma} T_j$ for $\sigma\in\Sigma$. Furthermore the condition that $\min_j w_v(T_j)=0$ implies that not all p_j are equal. Hence we obtain strict subspaces.

We apply now the theory to the unit equation. Let $\mathscr{P}$ be a finite set of places of a number field K including all the infinite places and let U be the set of $\mathscr{P}$-units of K. We are looking for the solutions $x=(x_0, \ldots, x_n)$ in U^{n+1} such that

$$x_0+\ldots+x_n=0.$$

To determine them, let $T_0, \ldots, T_n$ be coodinate functions on K^{n+1}. They generate a vector space V over K. We have

$$\|T_j(x)\|_v=\frac{|x_j|_v}{\|x\|_v}, \quad j=1, \ldots, n,$$

where $\|x\|_v=\max|x_j|_v$. Since $x_j\in U$ for $j=0, \ldots, n$ we have $\|x\|_v=1$ for $v\notin\mathscr{P}$ and by the product formula

$$\prod_{w\in\mathscr{P}} \|T_j(x)\|_w=H(x)^{-[K:\mathbb{Q}]}.$$

We define weights $p_w(T_j)$ by

$$\|T_j(x)\|_w=H(x)^{-p_w(T_j)[K:\mathbb{Q}]}.$$

They satisfy

$$\sum_{w\in\mathscr{P}} p_w(T_j)=1, \quad j=0, \ldots, n,$$

and

$$\min_j(p_w(T_j))=0, \quad w\in\mathscr{P}.$$

As before we denote by $\tilde{V}$ by the quotient of V by the subspace generated by $T_0+\ldots+T_n$. This corresponds to an injection

$$i\colon \mathbb{P}(\tilde{V})\to\mathbb{P}(V)=\mathbb{P}^n$$

and x induces a point $\xi\in\mathbb{P}^n(K)$ which is in the image of i. The weights $p_w(T_j)$ define a filtration on V with μ-invariant

$$\mu(V)=\frac{1}{n+1}\sum_{w\in\mathscr{P}}\sum_{j=0}^{n} p_w(T_j)=1.$$

The filtration induces a filtration on $\tilde{V}$ with

$$\mu(\tilde{V}) = \frac{n+1}{n}.$$

In fact any n out of the sections $t_0, \ldots, t_n$ given by $t_j = i^* T_j$ are linearly independent and therefore only the n largest weights matter for any fixed $w \in \mathscr{P}$. Since the smallest weight is zero the result follows. If $\tilde{V}$ is semistable then by Theorem 8.1 there are only finitely many solutions. Otherwise we look at the Harder-Narasimhan filtration of $\tilde{V}$. If all slopes are greater than 1 again we get only finitely many solutions. In the remaining case we find that by the general theory up to finitely many exeptions at least one proper subsum of $x_0 + \ldots + x_n$ must vanish.

11 Norm forms

Let again K be a number field, $L \supseteq K$ a finite field extension and $X \subseteq \mathbb{P}^n$ a projective variety over K with $r = \dim X$. We fix an integer $l > r+1$ and positive integers $d_1, \ldots, d_l$. Let f_j be a non-zero section in $\Gamma(X \underset{K}{\otimes} L, \mathcal{O}(d_j))$ for $j = 1, \ldots, l$ such that for any subset $T \subset \{1, \ldots, l\}$ with $|T| = r+1$ the sections $f_j, j \in T$, have no common zero on X. Also let $\mathscr{P}$ be a finite set of places of L including the infinite places and let $\mathcal{O}$ be the ring of $\mathscr{P}$-integers. We consider the equation

$$\prod_{j=1}^{l} f_j(x) = c \tag{8}$$

for $0 \neq c \in L$ with $x \in X(K)$. We are interested in the solutions $x = (x_0, \ldots, x_n) \in \mathcal{O}^{n+1}$ with $x \in X(K)$. We assume that $D > (l-1)d$ where $D = \sum d_j$ and $d = \max_j (d_j)$.

Theorem 11.1 *The solutions in $X(K)$ of* (8) *which can be represented in the form $x = (x_0, \ldots, x_n) \in \mathcal{O}^{n+1}$ lie in finitely many hypersurfaces. The number of these hypersurfaces can be bounded effectively.*

Proof. Let $\varepsilon > 0$ be a real number such that $(1-3\varepsilon)D > (l-1)\, d$ and x be such a solution. Then for $w \in \mathscr{P}$

$$\prod_{j=1}^{l} \| f_j(x) \|_w = |c|_w \, \|x\|_w^{-\Sigma d_j}.$$

We assume that the height of x is sufficiently large. Then there exist real numbers $p_{j,w}$ with

$$\| f_i(x) \|_w < H(x)^{-p_{j,w}}$$

such that

$$\sum_{j,w} p_{j,w} > 1 - \varepsilon [L:\mathbb{Q}] \sum_j d_j$$

and

$$l \sum_w |\min_j (p_{j,w})| < \varepsilon [L:\mathbb{Q}] \, D/d$$

For each w let $T_w \subset \{1, \ldots, l\}$ be a subset with $|T_w| = r+1$ and such that $\sum_j p_{j,w}/d_j$ is maximal where the sum is over all $j \in T_w$. Then the ring

$$R = \bigoplus_{t \geqq 0} \Gamma(X \otimes_K L, \mathcal{O}(t))$$

is finite over the ring S generated by elements f_j, $j \in T_w$. It follows that

$$R = S e_1 \oplus \ldots \oplus S e_N$$

for some $N \geqq 1$ and elements $e_j \in R$. Let $\{F_w^p\}$ be the filtration on R defined by the f_j, $j = 1, \ldots, l$. The filtration gives a probability measure $\rho_{X,w}$ as usual. It induces a filtration on S and we have

$$E(\rho_{X,w}) \geqq \frac{1}{r+1} \sum_{j \in T_w} p_{j,w}/d_j .$$

We put and obtain $p'_{j,w} = p_{j,w}/d_j$

$$\binom{l-2}{r} \sum_{j \neq j(w)} p'_{j,w} = \sum_T \sum_{j \in T} p'_{j,w} \leqq \binom{l-1}{r+1} \sum_{j \in T_w} p'_{j,w}$$

where in the double sum we sum over all $T \subset \{1, \ldots, l\}$ with $|T| = r+1$ and with $j(w) \in T$. Here $j(w)$ is any fixed index with $p_{j(w),w} = \min_j (p_{j,w})$. It follows that

$$\sum_{j \in T_w} p'_{j,w} \geqq \frac{r+1}{l-1} \sum_{j \neq j(w)} p'_{j,w} .$$

Furthermore the minimum of the function

$$\sum_{j,w} p_{j,w}/d_j$$

under the side conditions above is bounded from below by $(1-2\varepsilon)[L:\mathbb{Q}] D/d$. Therefore we find that

$$\sum_w E(\rho_{X,w}) > [L:\mathbb{Q}] .$$

The theorem now follows from Theorem 9.3. □

As an application we take $0 \neq f \in \Gamma(X \underset{K}{\otimes} L, \mathcal{O}(d))$ and let f_i, $j=1, \ldots, l$ be the conjugates of f. Then we find that the solution of

$$\operatorname{Norm}(f)=c$$

with $0 \neq c \in K$ lie in finitely many hypersurfaces.

12 The gap principle

Let $K \subseteq L$ be number fields and $\mathcal{P}$ a finite set of places of L. For independent variables $T_0, \ldots, T_n$ we put $V=KT_0 \oplus \ldots \oplus KT_n$ and we let $f_{w,\alpha}$, $\alpha=0, \ldots, n$, be a basis for $V \underset{K}{\otimes} L$ for each $w \in \mathcal{P}$.

We fix a positive real number ρ. For any set of non-negative real numbers $p_{w,\alpha}$, $w \in \mathcal{P}$, $\alpha=0, \ldots, n$, we put

$$\mu=\frac{1}{n+1} \sum_{w,\alpha} p_{w,\alpha}.$$

We consider the set of points in $\mathbb{P}(V)(K)$ with

$$\text{(9)} \qquad \|f_{w,\alpha}(x)\|_w \leqq H(x)^{-p_{w,\alpha}} \|x\|_w, \qquad \alpha=0, \ldots, n, \qquad w \in \mathcal{P},$$
$$R \leqq H(x) \leqq R^{1+\rho},$$

where R is a positive real number and the $p_{w,\alpha}$ are any set of non-negative real numbers with

$$\text{(10)} \qquad \mu \geqq (1+2\rho)[L:\mathbb{Q}].$$

Theorem 12.1 *There exist positive constants R_0 and N depending effectively only on ρ and $[L:\mathbb{Q}]$ such that for $R>R_0$ the set of solutions of* (9) *is contained in a union of at most N hyperplanes.*

Proof. We always may assume that the set of vectors p with components $p_{w,\alpha}$, $w \in \mathcal{P}$, $\alpha=0, \ldots, n$, for which there exists a solution of (9) under the condition (10) is contained in the compact region defined by

$$(1+2\rho)[L:\mathbb{Q}] \leqq \mu \leqq (1+3\rho)[L:\mathbb{Q}].$$

By compactness there exists a finite set of vectors p' in the compact region such that any solution of (9) with (10) is already a solution for some p' in the finite set provided that we replace 2ρ in (10) by some ρ' with $\rho<\rho'<2\rho$. Let p' with components $p'_{w,\alpha}$, $w \in \mathcal{P}$, $\alpha=0, \ldots, n$, be in the finite set and $x^{(0)}, \ldots, x^{(n)}$ solutions of (9) for this p'. Then

$$\|\det(x^{(\beta)}\|_w \ll \|\det(f_{w,\alpha}(x^{(\beta)}))\|_w$$

and

$$\|\det(f_{w,\alpha}(x^{(\beta)}))\|_w \ll \prod_{\beta} \|x^{(\beta)}\|_w \cdot R^{-\Sigma_\alpha p'_{w,\alpha}}$$

for $w \in \mathscr{P}$. Therefore

$$\prod \|\det(x^{(\beta)})\| \ll R^{(1+\rho)(n+1)[L:\mathbb{Q}] - \Sigma_{w,\alpha} p'_{w,\alpha}}$$

where the product is taken over all places of L with the usual convention and where the sum is over all $w \in \mathscr{P}$. Since

$$\sum_{w,\alpha} p'_{w,\alpha} = (n+1)\mu$$

we find that the exponent of R is negative. Therefore the right hand side becomes less than one so that $\det(x^{(\beta)}) = 0$. This implies that the solutions with fixed p' are all contained in a fixed hyperplane. This proves the theorem. □

We should remark also that the result easily extends to the case where the $f_{w,\alpha}$ are non-linear.

References

[BV] Bombieri, E., Vaaler, J.: On Siegel's Lemma. Invent. Math. **73**, 11–32 (1983)
[Fa] Faltings, G.: Diophantine Approximation on Abelian Varieties. Ann. Math. **129**, 549–576 (1991)
[Ma] Matsumura, H.: Commutative ring theory. Cambridge: Cambridge University Press (1989)
[NS] Narasimhan, M.S., Seshadri, C.S.: Stable and unitary vector bundles on a compact Riemann surface. Ann. Math. **82**, 540–567 (1965)
[S1] Schmidt, W.M.: Approximation to Algebraic numbers. Enseign. Math. **17**, 187–253 (1971)
[S2] Schmidt, W.M.: Diophantine Approximation. (Lect. Notes Math., vol. 785) Berlin Heidelberg New York: Springer (1980)