10. Norm forms

Objekttyp: **Chapter**

Zeitschrift: **L'Enseignement Mathématique**

Band (Jahr): **17 (1971)**

Heft 1: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **24.09.2024**

10. Norm forms

10.1. Let K be an algebraic number field of degree t. There are t isomorphisms of K into the complex numbers; denote the images of an element α of K under these isomorphisms by $\alpha^{(1)}, \ldots, \alpha^{(t)}$. Let $L(\mathbf{x}) = \alpha_1 x_1 + \ldots + \alpha_n x_n$ be a linear form with coefficients in K. For $1 \leqq i \leqq t$ put $L^{(i)}(\mathbf{x}) = \alpha_1^{(i)} x_1 + \ldots + \alpha_n^{(i)} x_n$. The norm

$$\mathcal{N}(L(\mathbf{x})) = L^{(1)}(\mathbf{x}) \ldots L^{(t)}(\mathbf{x})$$

is a form of degree t with rational coefficients. A form obtained in this way will be called a *norm form*. It is easy to see that every form $F(\mathbf{x})$ which has rational coefficients and is irreducible over the rationals but which is a product of linear forms with algebraic coefficients, is a constant times a norm form. In particular when $n = 2$, every form with rational coefficients which is irreducible over the rationals is essentially a norm form.

10.2. We may as well discuss more general products of linear forms with real or complex algebraic coefficients. For any real or complex number α we denote its complex conjugate by $\bar{\alpha}$. The complex conjugate of a linear form $L(\mathbf{x}) = \alpha_1 x_1 + \ldots + \alpha_n x_n$ is defined by $\bar{L}(\mathbf{x}) = \bar{\alpha}_1 x_1 + \ldots + \bar{\alpha}_n x_n$. We shall call linear forms $L_1, \ldots, L_t$ a *Symmetric System* if, except for the ordering, $\bar{L}_1, \ldots, \bar{L}_t$ are the same as the given forms.

Theorem 10A (Schmidt, 1971b). *Suppose $L_1, \ldots, L_t$ is a Symmetric System of linear forms with algebraic coefficients. Suppose $\eta > 0$. The following two conditions are equivalent:*

(a) *There is a constant $c_1 = c_1(L_1, \ldots, L_t; \eta)$ and there are infinitely many integer points $\mathbf{x}$ with*

$$| L_1(\mathbf{x}) \ldots L_t(\mathbf{x}) | \leqq c_1 |\mathbf{x}|^{t-\eta} .$$

(b) *There is a rational subspace S^d of dimension d with $1 \leqq d \leqq n$ and there is a Symmetric System of linear forms $L_{i_1}, \ldots, L_{i_m}$ with $1 \leqq m \leqq t$ and $i_1 < \ldots < i_m$ whose restrictions to S^d have rank r with*

$$r \leqq dm/\eta \quad \textit{and} \quad r < d . \tag{10.1}$$

This theorem again contains Roth's Theorem. It can be deduced from the Subspace Theorem.

10.3. We shall now discuss diophantine equations

$$\mathcal{N}(L(\mathbf{x})) = a \tag{10.2}$$

where a is a constant. As $\mathbf{x}$ runs through the integer points, $\mathcal{N}(L(\mathbf{x}))$ runs through certain rationals with bounded denominators. Hence there are constants a for which (10.2) has infinitely many integer solutions $\mathbf{x}$ precisely if there are constants b for which the inequality

$$|\mathcal{N}(L(\mathbf{x}))| \leqq b \tag{10.3}$$

has infinitely many solutions. This will in fact be the case if the coefficients of L are linearly dependent over the rationals, so that we shall assume in the sequel that the coefficients are linearly independent.

We shall say that a linear form with coefficients in K is *full* in K if its coefficients form a field basis of K. Suppose the linear form $L(\mathbf{x})$ is full in K where K is neither the rational field nor an imaginary quadratic field. Further assume for a moment that the coefficients of L form in fact an integer basis of K. By Dirichlet's unit theorem K contains infinitely many units, and hence there are infinitely many integer points $\mathbf{x}$ with $\mathcal{N}(L(\mathbf{x})) = 1$. By studying units of certain subrings of K one sees more generally that *if* $L(\mathbf{x})$ *is full in* K *where* K *is not rational or imaginary quadratic, then* (10.3) *has infinitely many solutions if* b *is large enough.* We shall say that a linear form $L(\mathbf{x})$ *represents* a linear form $L'(\mathbf{y})$ (where the number of components of $\mathbf{y}$ need not be n) if there is a constant c such that for every integer point $\mathbf{y}$ there is an integer point $\mathbf{x}$ with $L'(\mathbf{y}) = cL(\mathbf{x})$. Now suppose $L(\mathbf{x})$ is a linear form with coefficients in K. We shall call $L(\mathbf{x})$ *degenerate* if it represents a linear form $L'(\mathbf{y})$ which is full in a subfield K' of K which is neither rational nor imaginary quadratic. For example, $L(\mathbf{x}) = \sqrt{2}\,x_1 + \sqrt{3}\,x_2 + \sqrt{6}\,x_3$ is not full in $K = \mathbf{Q}(\sqrt{2}, \sqrt{3})$, but it represents the form $2y_1 + \sqrt{6}y_2 = \sqrt{2}(\sqrt{2}y_1 + \sqrt{3}y_2)$ which is full in $K' = \mathbf{Q}(\sqrt{6})$ and thus $L(\mathbf{x})$ is degenerate. From what we said above it follows that *for degenerate* $L(\mathbf{x})$ *and for large* b *the inequality* (10.3) *has infinitely many integer solutions.* A detailed proof may be found, e.g., in Borevich and Shafarevich (1966, ch. 2).

10.4. The converse also holds:

THEOREM 10B (Schmidt, 1971b). *Suppose* $L(\mathbf{x})$ *is a non-degenerate linear form with linearly independent coefficients in a number field* K. *Then for* *any fixed b, the inequality* (10.3) *has only finitely many solutions in integer points* **x**.

When the number of variables $n = 2$, this becomes Thue's result on his equation $F(x, y) = m$ where $F(x, y)$ is a binary form. When $n = 3$ the theorem was shown by Skolem (1935) when K has degree $t = 5$ and by Chabauty (1938) for general degree, but these authors made the additional assumption that among the isomorphisms of K into the complex numbers there are at least two pairs of complex conjugates. Skolem and Chabauty used a p-adic method. The general case $n = 3$ was settled by Schmidt (1967b). Before the results of §7 were known, Györy (1968) assumed the hypothetical truth of Corollary 7B and derived results about norm forms in an arbitrary number of variables. Since he did not have Theorem 10A as a tool, his results are relatively weak. See also Györy (1969), where he proves some special cases of Theorem 10B. Ramachandra (1969) dealt with special norm forms and derived for them an asymptotic formula for the number of solutions of (10.3), thus generalizing Mahler's (1933c) result.

Theorem 10B and the theorems of Skolem and Chabauty are non-effective. Effective bounds for the size of the solutions of certain rather special equations with norm forms were given by Skolem (1937) (for further references see Skolem (1938)) and Feldman (1970b). See also the references given at the end of §7.2.

If $L(\mathbf{x})$ is full in K where K is neither rational nor imaginary quadratic, then the solutions of an equation (10.2) may be parametrized by using the group of units of K. More generally one can show that if $L(\mathbf{x})$ is degenerate, then all solutions of (10.2) with finitely many exceptions belong to finitely many parameter families. For example, in the equation

$$(10.4) \qquad \mathcal{N}(\sqrt{2}x_1 + \sqrt{3}x_2 + \sqrt{6}x_3) = a\,,$$

all but finitely many solutions have $x_3 = 0$ or $x_2 = 0$ or $x_1 = 0$ and hence come from solutions of $\mathcal{N}(\sqrt{2}x_1 + \sqrt{3}x_2) = a$ or $\mathcal{N}(\sqrt{2}x_1 + \sqrt{6}x_3) = a$ or $\mathcal{N}(\sqrt{3}x_2 + \sqrt{6}x_3) = a$. Hence all but finitely many solutions come from one of the three equations

$$\mathcal{N}'(2x_1 + \sqrt{6}x_2) = \pm 2\sqrt{a}, \quad \mathcal{N}''(x_1 + \sqrt{3}x_3) = \pm \frac{1}{2}\sqrt{a},$$

$$\mathcal{N}'''(x_2 + \sqrt{2}x_3) = \pm \frac{1}{3}\sqrt{a}\,,$$

where $\mathcal{N}', \mathcal{N}'', \mathcal{N}'''$ are the norms from the fields $K' = \mathbf{Q}(\sqrt{6})$, $K'' = \mathbf{Q}(\sqrt{3})$ and $K''' = \mathbf{Q}(\sqrt{2})$. The solutions of these three equations can be easily described in terms of the units of the fields K', K'' and K'''. In particular (10.4) has only finitely many solutions unless a is a perfect square.

10.5. Roth's Theorem implied not only that for an irreducible binary form $F(x, y)$ of degree $t \geqq 3$ there are only finitely many solutions of $|F(x, y)| < a$, but according to Theorem 2C there are only finitely many solutions of $|F(x, y)| < (|x|+|y|)^{v}$ if $v < t-2$. In the present context it is reasonable to expect that " in general " there are only finitely many integer points $\mathbf{x}$ with

$$|\mathcal{N}(L(\mathbf{x}))| < |\mathbf{x}|^{v} \tag{10.5}$$

if

$$v < t - n\,. \tag{10.6}$$

Using Minkowski's Linear Forms Theorem one can easily show that unless $n = 1$ or $n = 2$ and no conjugate of L has real coefficients, there are infinitely many $\mathbf{x}$ with $|\mathcal{N}(L(\mathbf{x}))| \leqq c\,|\mathbf{x}|^{t-n}$; hence $t - n$ in (10.6) is best possible.

Suppose $K = \mathbf{Q}(\alpha)$ is a number field of degree t and suppose $1 \leqq r \leqq t$. We shall say that K is *r times transitive* if for any r distinct conjugates $\alpha^{(i_1)}, \ldots, \alpha^{(i_r)}$ of α there is an element φ of the Galois group of $\mathbf{Q}(\alpha^{(1)}, \ldots, \alpha^{(t)})$ (i.e. the least normal extension of K) with $\varphi(\alpha^{(1)}) = \alpha^{(i_1)}, \ldots, \varphi(\alpha^{(r)}) = \alpha^{(i_r)}$. This definition is clearly independent of the primitive element α.

THEOREM 10C (Schmidt, in preparation). *Suppose the coefficients of* $L(\mathbf{x}) = \alpha_1 x_1 + \ldots + \alpha_n x_n$ *lie in a number field K and are linearly independent over the rationals. Suppose that K is generated by the quotients* α_i/α_j $(1 \leqq i, j \leqq n)$ *and that K is* $(n-1)$*-times transitive. Finally assume that any n of the conjugates of L* $(\mathbf{x})$ *are linearly independent. Then for every v with* (10.6) *there are only finitely many integer points* $\mathbf{x}$ *satisfying* (10.5).

COROLLARY 10D. *Suppose L* $(\mathbf{x})$ *is as above and suppose G* $(\mathbf{x})$ *is a polynomial of total degree* $v < t - n$. *Then the equation*

$$\mathcal{N}(L(\mathbf{x})) = G(\mathbf{x})$$

has only finitely many integer solutions.

This contains Corollary 2D.

10.6. Both Theorems 10B and 10C are derived from Theorem 10A. We shall briefly discuss the argument for Theorem 10B. We have to show that $L(\mathbf{x})$ is degenerate if the inequality

$$|\mathcal{N}(L(\mathbf{x}))| = |L^{(1)}(\mathbf{x}) \dots L^{(t)}(\mathbf{x})| \leqq c = c\,|\mathbf{x}|^{t-t}$$

has infinitely many solutions. By the case $\eta = t$ of the assertion (a) ⇒ (b) of Theorem 10A there is a subspace S^d and there is a Symmetric System $L^{(i_1)}, \dots, L^{(i_m)}$ of forms whose restrictions to S^d have a rank r with

$$r \leqq dm/t \quad \textit{and} \quad r < d\,. \tag{10.7}$$

One can reduce the situation to the special case where $L(\mathbf{x}) = x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$ and $K = \mathbf{Q}(\alpha_2, \dots, \alpha_n)$, and where $d = n$. The conditions (10.7) now become $r \leqq nm/t$ and $r < n$, and with

$$q = t/n$$

they become

$$rq \leqq m \quad \textit{and} \quad r < n\,. \tag{10.8}$$

Now $rq < m$ is impossible (this would imply infinitely many solutions of $|\mathcal{N}(L(\mathbf{x}))| < |\mathbf{x}|^{-\delta}$ for some $\delta > 0$), and hence the rank r of every Symmetric System $L^{(i_1)}, \dots, L^{(i_m)}$ satisfies

$$m \leqq rq\,. \tag{10.9}$$

But by (10.8) there is a special Symmetric System $L^{(i_1)}, \dots, L^{(i_\mu)}$ of rank ρ with

$$\mu = \rho q \quad \textit{and} \quad \rho < n\,. \tag{10.10}$$

We choose μ and ρ as small as possible with this property. We may assume without loss of generality that the forms $L^{(i_1)}, \dots, L^{(i_\mu)}$ are $L^{(1)}, \dots, L^{(\mu)}$.

In what follows, α will be a primitive element of K, i.e. an element with $K = \mathbf{Q}(\alpha)$. We have to distinguish two cases.

(A) For every element φ of the Galois group of $\mathbf{Q}(\alpha^{(1)}, \dots, \alpha^{(t)})$, the two sets $\{\alpha^{(1)}, \dots, \alpha^{(\mu)}\}$ and $\{\varphi(\alpha^{(1)}), \dots, \varphi(\alpha^{(\mu)})\}$ are identical or disjoint.

(B) We have not (A).

In the case (A) it turns out that μ divides t and that $L(\mathbf{x})$ represents a full linear form $L'(\mathbf{y})$ in a field K' of degree t/μ, where K' is neither rational nor imaginary quadratic, and hence $L(\mathbf{x})$ is degenerate. Let us discuss what happens in the case (B). For simplicity we shall assume that K is totally real.

There is an element φ of the Galois group such that the sets $\{\alpha^{(1)}, ..., \alpha^{(\mu)}\}$ and $\{\varphi(\alpha^{(1)}), ..., \varphi(\alpha^{(\mu)})\}$ are neither identical nor disjoint. We may assume without loss of generality that

$$\{\varphi(\alpha^{(1)}), ..., \varphi(\alpha^{(\mu)})\} = \{\alpha^{(1)}, ..., \alpha^{(l)}, \alpha^{(\mu+1)}, ..., \alpha^{(2\mu-l)}\}.$$

Here $1 \leqq l \leqq \mu - 1$. The forms $L^{(1)}, ..., L^{(\mu)}$ have rank ρ, and hence also $L^{(1)}, ..., L^{(l)}, L^{(\mu+1)}, ..., L^{(2\mu-l)}$ have rank ρ. Denote the rank of $L^{(1)}, ..., L^{(l)}$ by r_1 and the rank of $L^{(1)}, ..., L^{(\mu)}, ..., L^{(2\mu-l)}$ by r_2. It is easily seen that $r_2 \leqq 2\rho - r_1$, i.e. that

$$r_1 + r_2 \leqq 2\rho .$$

Since μ was chosen as small as possible with (10.10), and since $l \leqq \mu - 1$, we have $l < r_1 q$. The number $2\mu - l$ of elements of $L^{(1)}, ..., L^{(\mu)}, ..., L^{(2\mu-l)}$ satisfies $2\mu - l \leqq r_2 q$ by (10.9). Thus

$$2\mu = l + (2\mu - l) < r_1 q + r_2 q \leqq 2\rho q ,$$

which contradicts (10.10). Hence (B) is impossible if K is totally real.

We have in fact used the hypothesis that K is totally real, for in general $L^{(1)}, ..., L^{(l)}$ need not be a Symmetric System, and $l < r_1 q$ need not hold. The situation is therefore somewhat more complicated if K is not totally real.

11. Generalizations and open problems

11.1. The theorems of §7 and §10 can almost certainly be generalized to include p-adic valuations. I understand that work on this question is being done now. (p-adic versions of the results of §2 were discussed in §4.5). Next, suppose that K is an algebraic number field and that $\alpha_1, ..., \alpha_l$ are algebraic numbers such that $1, \alpha_1, ..., \alpha_l$ are linearly independent over K. It is likely that *for every $\delta > 0$ there are only finitely many l-tuples of elements $\beta_1, ..., \beta_l$ of K with*

$$(11.1) \qquad |\alpha_i - \beta_i| < \mathscr{H}(\beta)^{-1-(1/l)-\delta} \quad (i = 1, ..., l),$$

where $\mathscr{H}(\beta)$ is a suitably defined height of $\beta = (\beta_1, ..., \beta_l)$. A possible definition for $\mathscr{H}(\beta)$ is

$$\mathscr{H}(\beta) = \prod_v \max(1, \|\beta_1\|_v, ..., \|\beta_l\|_v),$$