ACTA ARITHMETICA
XVIII (1971)

Forms over $\mathfrak{p}$-adic fields

by

P. A. B. PLEASANTS (Cardiff)

In grateful memory of Professor Davenport

1. Introduction. A well-known conjecture of Artin, which remained unsettled for many years, was that every homogeneous polynomial of degree d in n variables over a $\mathfrak{p}$-adic field represents zero non-trivially if $n > d^2$. Ax and Kochen [1] proved by a very difficult method that this is true when the order of the residue class field is greater than a bound depending only on d, but their method did not give an effective estimate for this bound. An example due to Terjanian [16] shows that the conjecture is not true in general. The conjecture is always true when $d = 2$ or 3, however, the proof of the quadratic case being due to Hasse [9]. Proofs of the cubic case have been given by Demyanov [8] (except for the case when the residue class field has characteristic 3), Lewis [11], Springer [14], and Davenport [5]. Laxton and Lewis [10] have also given a proof of the conjecture for $d = 5$, 7 and 11 (subject to the order of the residue class field being large enough) using the deep theorem of Weil about the number of points on algebraic curves. (The case $d = 5$ had already been treated by Birch and Lewis [2] by a similar method.) The work of Laxton and Lewis has been superseded by that of Ax and Kochen except in so far as more effective estimates for the size of residue class field required can be deduced from the former.

The treatments of $\mathfrak{p}$-adic forms in [11], [14], [2], [5], [3] and [10] have several features in common — in particular they all depend on extending non-singular zeros over the residue class field to non-singular zeros over the $\mathfrak{p}$-adic field and on transforming the original polynomial by multiplying certain sets of variables by a prime element of the field. The treatments in [5], [3] and [10] are simplified by the use of certain invariants: Davenport introduced an invariant $h(C)$ in [5], which was generalized to forms of degree greater than 3 by Birch and Lewis in [3], and in [10] Laxton and Lewis used an invariant $\mathscr{I}(F)$. In the present paper a further simplification is achieved that avoids the use of such invariants or of reduced polynomials of any kind but retains the other features mentioned. The method also gives some information about non-singular zeros of p-adic forms of degrees 7 and 11.

Our main lemma is essentially an alternative account of the part of Laxton's and Lewis's work in which they reduce the problem of finding a non-trivial zero of a p-adic form to that of finding a non-singular zero of a related form over the residue class field. Consequently this paper, like [10], gives no information about forms of degrees other than the ones mentioned, since for forms of other degrees it is possible for the corresponding forms over the residue class field to have no non-singular zeros. In § 5 we deal with forms over the residue class field in the cubic case, when the use of Weil's theorem is not needed.

2. Notation. We call two homogeneous polynomials, F_1 and F_2, over a field $\mathscr{K}$ *equivalent* if there is a non-singular linear transformation T over $\mathscr{K}$ such that $F_1(\boldsymbol{x}) \equiv F_2(T\boldsymbol{x})$. (It is not assummed that the coefficients of T are integers, should $\mathscr{K}$ contain integers of any kind.) The order, $o(F)$, of a homogeneous polynomial F is the smallest integer m such that F is equivalent to a form that contains only m variables explicitly. It is a consequence of a remark at the beginning of § 2 of [7] that $o(F)$ does not change when the field $\mathscr{K}$ is extended in any way. A form F in n variables is called *degenerate* if $o(F) < n$. A vector $\boldsymbol{x}_0$ such that $F(\boldsymbol{x}_0) = 0$ is a *non-trivial* zero of F if its coordinates are not all zero, and is a *non-singular* zero if $\dfrac{\partial F}{\partial x_i}(\boldsymbol{x}_0) \neq 0$ for some i.

We define $h(F)$ (1), for a form F in n variables over a field $\mathscr{K}$, as the smallest integer h for which $F(\boldsymbol{x})$ is equivalent to a form of the shape $x_1G_1(x_1, \ldots, x_n) + \ldots + x_hG_h(x_1, \ldots, x_n)$, where $G_1, \ldots, G_h$ are forms of degree one less than the degree of F. This invariant was introduced for cubic forms by Davenport and Lewis in [6]. An alternative definition of $h(F)$ is that $n-h(F)$ is the greatest dimension of any linear space over $\mathscr{K}$ on which F vanishes identically (2). Clearly $h(F) \leqslant o(F)$, and F has a non-trivial zero over $\mathscr{K}$ if and only if $h(F) < n$. Unlike $o(F)$, $h(F)$ may decrease when the field $\mathscr{K}$ is extended.

Throughout this paper K is a field that is complete with respect to a discrete non-archimedean valuation, $\mathfrak{o}$ is the ring of integers of K, and π is a fixed prime element of K. We shall assume that the residue class field $k = \mathfrak{o}/\pi\mathfrak{o}$ is finite. We denote the order of an element a of K by $\operatorname{ord} a$ — so that $\operatorname{ord} a$ is the largest integer r for which $a \in \pi^r\mathfrak{o}$ (and is related to the order of a form only in name). If $F(\boldsymbol{x})$ is a form with

(1) This is not related to the invariant $h(C)$ mentioned in § 1.

(2) If $\mathscr{K}$ is a finite field, to say that F vanishes identically on the linear space L asserts more than merely that every point of L is a zero of F.

coefficients in $\mathfrak{o}$ we denote by $F^*(\boldsymbol{x})$ the image of F in the residue class field.

3. The main lemma. The following lemma reduces the problem of finding a non-trivial zero of a form F over K to that of finding non-singular zeros over the residue class field k of forms of the same degree as F for which the invariant h is greater than the degree.

LEMMA 1. *Let d be fixed. If every form f over k of degree d with $h(f) > d$ has a non-singular zero over k then every form F over K of degree d with $h(F) > d^2$ has a non-singular zero over K.*

Proof. Let $F(\boldsymbol{x})$ be a form over K of degree d in n variables, where $n > d^2$, that has no non-singular zeros over K. There is no loss of generality in assuming that the coefficients of F are integers of K. Non-singular zeros of $F^*(\boldsymbol{x})$ over the residue class field k give rise to non-singular zeros of $F(\boldsymbol{x})$ over K by the familiar process known as 'Newton approximation' or 'Hensel's Lemma' (see Lemma 1 of [11], for example). It follows that $F^*(\boldsymbol{x})$ has no non-singular zeros, and so $h(F^*) \leqslant d$. Hence there is a non-singular linear transformation $\boldsymbol{x} = T^*\boldsymbol{y}$ over k taking $F^*(\boldsymbol{x})$ into $g(\boldsymbol{y}) = y_1g_1(\boldsymbol{y}) + \ldots + y_dg_d(\boldsymbol{y})$, where $g_1, \ldots, g_d$ are forms of degree $d-1$. T^* can be lifted to a linear transformation $\boldsymbol{x} = T\boldsymbol{y}$ over K, by choosing the coefficients of T from the appropriate residue classes, and then T takes $F(\boldsymbol{x})$ into $G(\boldsymbol{y})$, where G has integer coefficients and $G^* = g$. Also $\operatorname{ord}(\det T) = 0$. If we combine T with the linear transformation defined by

$$y_i = \begin{cases} \pi z_i & (i = 1, \ldots, d), \\ z_i & (i = d+1, \ldots, n), \end{cases}$$

we obtain a linear transformation T_1 with $\operatorname{ord}(\det T_1) = d$ which takes $F(\boldsymbol{x})$ into $\pi F_1(\boldsymbol{z})$, where F_1 is a form with integer coefficients.

Since F has no non-singular zeros neither does F_1, and so the same reasoning applies to F_1. Hence there exists a linear transformation T_2 with $\operatorname{ord}(\det T_2) = 2d$ such that $F(T_2\boldsymbol{z}) = \pi^2F_2(\boldsymbol{z})$, where F_2 is a form with integer coefficients. The argument can be repeated to give, for each positive integer r, a linear transformation T_r with $\operatorname{ord}(\det T_r) = rd$ such that $F(T_r\boldsymbol{z}) = \pi^rF_r(\boldsymbol{z})$, where F_r is a form with integer coefficients.

By applying elementary row and column operations to the matrix of T_r we can express T_r as $P_rD_rQ_r$, where P_r and Q_r are integral and unimodular (i.e. their determinants have order zero) and D_r is a diagonal transformation of the type

$$x_i = \pi^{\sigma(i)}z_i \quad (i = 1, \ldots, n),$$

where

$$\sigma(1) \leqslant \sigma(2) \leqslant \ldots \leqslant \sigma(n) \tag{1}$$

and

$$\sigma(1)+\sigma(2)+\ldots+\sigma(n) = rd. \tag{2}$$

On making the substitution $z = Q_r^{-1}u$, where

$$u_i = \begin{cases} \pi^{\sigma(m)-\sigma(i)}v_i & (i = 1, \ldots, m), \\ 0 & (i = m+1, \ldots, n) \end{cases}$$

for some m with $1 \leqslant m \leqslant n$, we have

$$F(T_r z) = F(v_1\pi^{\sigma(m)}p_1^{(r)}+\ldots+v_m\pi^{\sigma(m)}p_m^{(r)}),$$

where $p_i^{(r)}$ is the ith column of the matrix of P_r. Hence $F(v_1\pi^{\sigma(m)}p_1^{(r)}+\ldots+v_m\pi^{\sigma(m)}p_m^{(r)})$ is a form in $v_1, \ldots, v_m$ all of whose coefficients have order at least r, and so $F(v_1p_1^{(r)}+\ldots+v_mp_m^{(r)})$ is a form in $v_1, \ldots, v_m$ all of whose coefficients have order at least $r-d\sigma(m)$. It follows from (1) and (2) that $\sigma(m) \leqslant rd/(n-m+1)$, and so if we choose $m = n-d^2$ we have $r-d\sigma(m) \geqslant r/(d^2+1)$.

Since $\mathfrak{o}$ is compact with respect to the valuation topology of K, some subsequence of the sequence $\{P_r\}$ of linear transformations tends to a limit P. Then P is non-singular and

$$F(v_1p_1+\ldots+v_{n-d^2}p_{n-d^2}) \equiv 0,$$

where p_i is the ith column of the matrix of P. Hence F vanishes identically on the linear space of dimension $n-d^2$ spanned by $p_1, \ldots, p_{n-d^2}$, and so $h(F) \leqslant d^2$.

The reduction of T_r to diagonal form could have been avoided if we were content with a single non-trivial zero of F instead of a linear space of zeros; for some column $t_i^{(r)}$ of the matrix of T_r must have the maximum order of its elements $\leqslant rd/n$, and then if $w^{(r)}$ is a scalar multiple of $t_i^{(r)}$ the maximum order of whose coordinates is zero we have

$$\operatorname{ord}(F(w^{(r)})) \leqslant \operatorname{ord}(F(t_i^{(r)}))-rd^2/n \leqslant r(1-d^2/n).$$

Then $F(w) = 0$ and $w \neq O$, where w is a limit point of the sequence $\{w^{(r)}\}$.

In Lemma 11 of [10] Laxton and Lewis show by means of Weil's theorem about the number of points on algebraic curves that a certain type of form F^* of degree d over a finite field k has a non-singular zero over k if $d = 2, 3, 5, 7$ or 11 and k has sufficiently many elements. The only information about F^* that is needed for their proof is that $h(F^*) > d$ (the conclusion of Lemma 8 of [10] being an immediate consequence of this assumption). (Birch and Lewis had already shown ([2], Lemma 9) that a form F^* of degree 3 or 5 over a finite field with sufficiently many elements has a non-singular zero provided only that $o(F^*) > d$.) Consequently Laxton's and Lewis's result can be combined with our Lemma 1 to give the following slightly strengthened version of the theorem of [10].

THEOREM 1. *If $d = 2, 3, 5, 7$ or 11 and the residue class field k is sufficiently large, then every form F of degree d over K with $h(F) > d^2$ has a non-singular zero over K.*

This includes the theorem of [10], since if $n > d^2$ then either $h(F) > d^2$ or $h(F) < n$. For forms of degree 2, 3 or 5 the weaker hypothesis $o(F) > d^2$ is sufficient to ensure that F has a non-singular zero (as is proved in [2]), but this is not always true for forms of other degrees — a form of the shape $F(x) = G(x)Q^2(x)$, where G has no non-singular zeros and Q is a quadratic form of rank n, has order n but no non-singular zeros. Clearly Theorem 1 never provides non-singular zeros for forms in d^2+1 variables, as $h \leqslant d^2$ for such forms.

4. Remarks. The reason why the method of Laxton and Lewis works only for forms of the degrees stated is that the hypothesis of Lemma 1 is false for other degrees. Over any field $\mathscr{K}$ there are forms of any degree $\geqslant 2$ with h arbitrarily large. Hence for a degree d that is composite or is a sum of composite numbers there exist forms f that are products of powers of forms for each of which h is large. Such forms f have no non-singular zeros, and yet $h(f)$ is the minimum value of h for any factor of f and so is large. This example also shows that the conclusion of Lemma 1 and Theorem 1 itself are false for forms of other degrees.

The usefulness of Theorem 1 as a criterion for the existence of non-singular zeros of F is limited by the fact that if F is a form over the rational numbers Q, say, and K is a p-adic field then $h(F)$ over K may well be less than $h(F)$ over Q. It is not even true that these two values of h are equal for almost all p. In [13] Selmer gave examples of cubic forms C in three variables that are soluble p-adically for every p and yet insoluble over Q: hence $h(C) = 3$ over Q but $h(C) \leqslant 2$ over every p-adic field. More recently Swinnerton-Dyer [15], Mordell [12], and Cassels and Guy [4] have given examples of cubic forms in four variables that are everywhere locally soluble but are not globally soluble. Forms in more variables or of higher degree (or both) having all the local values of h less than the global value can be constructed from these examples in several ways.

If F is a form over K of degree d in n variables such that[3] $\operatorname{ord}(F(Tz))-n^{-1}d\operatorname{ord}(\det T)$ is bounded above for all linear transformations T, then it is not difficult to show that F is equivalent to a form that is 'weakly reduced' in the sense defined in [3] (where 'equivalent' now has the extended meaning used in [3]). The argument used in the proof of Lemma 1 then shows that a form with no non-trivial zero is equivalent to a weakly reduced form. Although weaker than the result

[3] $\operatorname{ord} F(T(z))$ denotes the maximum of the orders of the coefficients of $F(T(z))$, considered as a polynomial in z.

that every non-degenerate form is equivalent to a weakly reduced form, which is proved in [3], this is sufficient to give the result in the last section of [2] about the shape of forms that have no non-trivial zero.

The method of Lemma 1 does not seem capable of being applied to simultaneous systems of forms.

5. Cubic forms. When $d = 2$ or 3 it is not necessary to use Weil's theorem to complete the proof of Artin's conjecture. In the cubic case the following lemma is useful. It does not appear to have been explicitly stated before but it is implicit in the proof of Lemma 2.4 of [5] and the proof given here is Davenport's.

LEMMA 2. *If a non-degenerate cubic form over a field $\mathscr{K}$ has a non-trivial zero over $\mathscr{K}$ then it has a non-singular zero over $\mathscr{K}$.*

(For a non-degenerate quadratic form it is well known that *every* non-trivial zero is non-singular.)

Proof. The proof depends on the fact that a quadratic form over $\mathscr{K}$ that is zero for all values of its variables is identically zero. This can be seen by first choosing values of the variables with all but one variable zero and then choosing values with all but two variables zero. The corresponding statement for forms of higher degree is not true in general — for example, the cubic form $x_1^2x_2+x_1x_2^2$ over the field with two elements is always zero.

Let $C(x_1, \ldots, x_n)$ be a cubic form over $\mathscr{K}$ in n variables having a non-trivial zero $\boldsymbol{x}$ but no non-singular zero. After a suitable non-singular transformation we can suppose that $\boldsymbol{x} = (1, 0, \ldots, 0)$. Then C has the shape

$$x_1^2L(x_2, \ldots, x_n)+x_1Q(x_2, \ldots, x_n)+C'(x_2, \ldots, x_n),$$

where L is a linear form, Q is quadratic, and C' is cubic. Since $\boldsymbol{x}$ is a singular zero $L \equiv 0$.

Suppose that values of $x_2, \ldots, x_n$ could be found for which $Q(x_2, \ldots, x_n) \neq 0$. Then the point

$$\big(-C'(x_2, \ldots, x_n)/Q(x_2, \ldots, x_n), x_2, \ldots, x_n\big)$$

would be a non-singular zero of C, since $\partial C/\partial x_1 = Q(x_2, \ldots, x_n) \neq 0$. Hence $Q(x_2, \ldots, x_n)$ vanishes for all values of the variables, and so $Q \equiv 0$. Thus C is equivalent to C', and so is degenerate.

COROLLARY. *Suppose that $\mathscr{K}$ and r are such that every cubic form over $\mathscr{K}$ in r or more variables has a non-trivial zero over $\mathscr{K}$. Then every cubic form C over $\mathscr{K}$ with $o(C) \geqslant r$ has a non-singular zero over $\mathscr{K}$.*

Proof. Suppose that $o(C) = s \geqslant r$. Then C is equivalent to a form $C'(x_1, \ldots, x_s)$ in s variables. C' has a non-trivial zero over $\mathscr{K}$ and is non-degenerate when considered as a form in the s variables $x_1, \ldots, x_s$. Therefore C' has a non-singular zero over $\mathscr{K}$, by the lemma, and hence so does C.

Used in conjunction with Lemma 1 this corollary gives Artin's conjecture for cubic forms.

THEOREM 2. *Every cubic form C over K with $o(C) \geqslant 10$ has a non-singular zero over K.*

Proof. By a well known theorem of Chevalley every cubic form over k in 4 or more variables has a non-trivial zero. Hence, by the corollary, every cubic form c over k with $o(c) \geqslant 4$ has a non-singular zero over k, and it follows from Lemma 1 that every cubic form over K in 10 or more variables has a non-trivial zero over K. By a second application of the corollary, therefore, every cubic form C over K with $o(C) \geqslant 10$ has a non-singular zero over K.

Every cubic form C in two or more variables over the field $\boldsymbol{R}$ of real numbers has a non-trivial zero over $\boldsymbol{R}$, as can be seen by considering a path avoiding the origin that joins a point where C is positive to a point where C is negative. Hence the corollary to Lemma 2 has the following immediate consequence (which is Lemma 6.1 of [5]).

Every cubic form over $\boldsymbol{R}$ that is not the cube of a linear form has a non-singular zero over $\boldsymbol{R}$.

References

[1] J. Ax and S. Kochen, *Diophantine problems over local fields I*, Amer. J. Math. 87 (1965), pp. 605–630.

[2] B. J. Birch and D. J. Lewis, *p-adic forms*, J. Indian Math. Soc. 23 (1959), pp. 11–32.

[3] — — *On p-adic forms*, Michigan Math. J. 9 (1962), pp. 53–57.

[4] J. W. S. Cassels and M. J. T. Guy, *On the Hasse principle for cubic surfaces*, Mathematika 13 (1966), pp. 111–120.

[5] H. Davenport, *Cubic forms in thirty-two variables*, Philos. Trans. Roy. Soc. London, Ser. A, 251 (1959), pp. 193–232.

[6] — and D. J. Lewis, *Exponential sums in many variables*, Amer. J. Math. 84 (1962), pp. 649–665.

[7] — — *Non-homogeneous cubic equations*, J. London Math. Soc. 39 (1964), pp. 657–671.

[8] V. B. Demyanov, *On cubic forms in discretely normed fields* (Russian), Dokl. Akad. Nauk SSSR 74 (1950), pp. 889–891.

[9] H. Hasse, *Darstellbarkeit von Zahlen durch quadratische Formen in einem beliebigen algebraischen Zahlkörper*, J. Reine Angew. Math. 153 (1924), pp. 113–130.

[10] R. R. Laxton and D. J. Lewis, *Forms of degrees 7 and 11 over p-adic fields*, Proc. Sympos. Pure Math., Vol VIII (1965), pp. 16–21, Amer. Math. Soc., Providence, Rhode Island.

[11] D. J. Lewis, *Cubic homogeneous polynomials over p-adic number fields*, Ann. of Math. (2) 56 (1952), pp. 473–478.

[12] L. J. Mordell, *On the conjecture for the rational points on a cubic surface*, J. London Math. Soc. 40 (1965), pp. 149–158.
[13] E. S. Selmer, *The Diophantine equation $ax^3+by^3+cz^3=0$*, Acta Math. 85 (1951), pp. 203–362.
[14] T. A. Springer, *Some properties of cubic forms over fields with a discrete valuation*, Nederl. Akad. Wetensch. Proc., Ser. A, 58 (1955), pp. 512–516.
[15] H. P. F. Swinnerton-Dyer, *Two special cubic surfaces*, Mathematika 9 (1962), pp. 54–56.
[16] G. Terjanian, *Un contre-exemple à une conjecture d'Artin*, C. R. Acad. Sci. Paris, Ser. A–B, 262 (1966), p. 612.

Received on 18. 3. 1970

ACTA ARITHMETICA
XVIII (1971)

Frieze patterns

by

H. S. M. Coxeter (Toronto)

In memory of Harold Davenport,
my companion of college days

1. Introduction. The idea of a frieze pattern is most quickly conveyed by means of an example, such as the following pattern of order 7:

0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	...
...	1	2	2	3	1	2	4	1	2	2	3
1	3	5	2	1	7	3	1	3	5	...	
...	2	1	7	3	1	3	5	2	1	7	3
1	2	4	1	2	2	3	1	2	4	...	
...	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	0	0	...	

Apart from the borders of zeros and ones, the essential property is that every four adjacent numbers forming a square

$$\begin{array}{ccc} & b & \\ a & & d \\ & c & \end{array}$$

satisfy the "unimodular" equation $ad-bc=1$. Moreover, we insist that all the numbers (except the borders of zeros) shall be positive. The surprising conclusion is that every such pattern is periodic. More precisely, it is symmetrical by a glide: the product of a horizontal translation and a horizontal reflection.

After giving some historical background, we shall prove this periodicity and deduce some cyclic sequences based on continued fractions. Finally, we shall give a necessary and sufficient condition for a frieze pattern to consist of integers.

2. Frieze patterns of order 5. The story begins in 1602, when Nathaniel Torporley (1564–1632) began to investigate the five "parts"