HOW TO SOLVE A QUADRATIC EQUATION IN RATIONALS

D. W. MASSER

1. *Introduction*

The title alludes to a similar title of the paper [**3**] by Grunewald and Segal, in which it is shown how to solve a quadratic equation in integers. This latter procedure seems to be quite difficult, and the algorithm outlined in [**3**] is rather involved, although it is completely effective in the logical sense.

Kornhauser has given fairly good explicit 'search bounds' for integral solutions if the number of variables is two or essentially at least five. Thus let $P(X_1, \dots, X_n)$ be a polynomial of total degree at most 2, with rational integer coefficients, and write $H = H(P)$ for the maximum of the absolute values of the coefficients.

For the case $n = 2$, it is shown in [**4**] that if the equation

$$P(X_1, X_2) = 0 \tag{1}$$

has a solution in rational integers $X_1 = x_1$, $X_2 = x_2$, then it has one with

$$\max\{|x_1|, |x_2|\} \leqslant (14H)^{5H}. \tag{2}$$

For the case $n \geqslant 5$, assuming that the quadratic homogeneous part of P is non-singular, it is shown in [**5**] that if the equation

$$P(X_1, \dots, X_n) = 0 \tag{$*$}$$

has a solution in rational integers $X_1 = x_1, \dots, X_n = x_n$, then it has one with

$$\max\{|x_1|, \dots, |x_n|\} \leqslant (n^3H)^{50n}(H + \sqrt{H}). \tag{3}$$

Both (2) and (3) are actually close to best possible in their dependence on H: consideration of a generalized Pell-type equation shows that any bound $\exp(f(H))$ in (2) with $\lim_{H\to\infty} f(H)/H = 0$ cannot be valid, and variations of a counterexample of Kneser show that any bound like (3) with exponent less than $n/2$ also cannot be valid.

As far as I know, there are no explicit search bounds at all in the literature for the cases $n = 3$, $n = 4$ of three and four variables.

However, if P happens to be homogeneous, then much more precise search bounds are known, in this context for non-trivial solutions. These are traditionally stated for 'classically integral' forms

$$P(X_1, \dots, X_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij} X_i X_j$$

with integers $p_{ij} = p_{ji}$. Thus the 'off-diagonal' coefficients are even. In [**1**] (see

Received 6 January 1997; revised 24 May 1997.

1991 *Mathematics Subject Classification* 11D09.

Bull. London Math. Soc. 30 (1998) 24–28

also [**2**, Lemma 8.1, p. 87]) Cassels proved that if (∗) has an integral solution $X_1 = x_1, \ldots, X_n = x_n$ with $x_1, \ldots, x_n$ not all zero, then there is one with

$$\max\{|x_1|, \ldots, |x_n|\} \leqslant (3L)^{(n-1)/2}, \tag{4}$$

where

$$L = L(P) = \sum_{i=1}^{n} \sum_{j=1}^{n} |p_{ij}|.$$

Furthermore, the original Kneser counterexample shows that the exponent $(n-1)/2$ is best possible for every n. The result (4) has since been generalized in a number of directions, most recently by Schlickewei and Schmidt (see, for example, [**6**]).

Now for rational solutions. If P remains homogeneous, then these are essentially the same as integral solutions, and so we shall regard P as not necessarily homogeneous. So it must now be written as

$$P(X_1, \ldots, X_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} X_i X_j + 2 \sum_{i=1}^{n} p_{i0} X_i + p_{00}$$

for integers $p_{ij} = p_{ji}$, and we define

$$L(P) = \sum_{i=0}^{n} \sum_{j=0}^{n} |p_{ij}|.$$

The Corollary of [**3**, p. 2] asserts the existence of an algorithm to decide if (∗) has a rational solution, via the Hasse–Minkowski Theorem. The main result of the present paper is a search bound of the same precision as Cassels's estimate (4).

Accordingly, for $\xi = (\xi_1, \ldots, \xi_n)$ in $\mathbf{Q}^n$, define

$$H(\xi) = \prod_{v} \max\{1, |\xi_1|_v, \ldots, |\xi_n|_v\},$$

where the product is taken over all standard valuations v of the rational field $\mathbf{Q}$; if we write $\xi_1 = x_1/x_0, \ldots, \xi_n = x_n/x_0$ for coprime integers $x_0 \neq 0, x_1, \ldots, x_n$, then

$$H(\xi) = \max\{|x_0|, |x_1|, \ldots, |x_n|\}.$$

This makes it clear that for any B there are only finitely many ξ in $\mathbf{Q}^n$ with $H(\xi) \leqslant B$.

THEOREM. *Suppose that* (∗) *has a rational solution* $X_1 = \xi_1, \ldots, X_n = \xi_n$ *for* $\xi = (\xi_1, \ldots, \xi_n)$. *Then it has one with*

$$H(\xi) \leqslant (3L)^{(n+1)/2}.$$

We shall also show that the exponent $(n+1)/2$ is best possible.

Finally, in [**3**] an analogue for polynomials of the Hasse–Minkowski Theorem for forms was proved, namely the following.

THEOREM (Grunewald–Segal). *Suppose that* (∗) *has a solution over the real numbers* $\mathbf{R}$ *and over each p-adic completion* $\mathbf{Q}_p$. *Then it has a solution over* $\mathbf{Q}$.

We shall give another short proof of this fact.

2. *Proofs*

It is more convenient to work with forms, and it is clear that the above theorem on small rational zeros is equivalent to the following estimate for integral zeros.

PROPOSITION. *Suppose that $F(X_0, \dots, X_n)$ is a classically integral quadratic form. If it has an integral zero $(x_0, \dots, x_n)$ with $x_0 \neq 0$, then it has one with*

$$\max\{|x_0|, \dots, |x_n|\} \leqslant (3L)^{(n+1)/2},$$

where $L = L(F)$.

As we said above, the exponent $(n+1)/2$ is best possible. It may be found slightly surprising that the weak condition $x_0 \neq 0$ pushes up the 'Cassels exponent' (in this case $n/2$) by a positive amount.

We prove the Proposition by induction on n. We start with $n = 1$ and

$$F(X_0, X_1) = aX_0^2 + 2bX_0X_1 + cX_1^2$$

for integers a, b, c with

$$L = |a| + 2|b| + |c|.$$

If $c = 0$, then $b \neq 0$ and now $X_0 = 2b, X_1 = -a$ is the required small zero. Otherwise, if $c \neq 0$, then the discriminant $b^2 - ac$ must be the square of an integer d, and

$$X_0 = x_0 = c, \quad X_1 = x_1 = -b + d$$

is a zero. Further,

$$4d^2 \leqslant 4|b|^2 + (|a| + |c|)^2 \leqslant L^2,$$

and therefore

$$|x_0| \leqslant L, \quad |x_1| \leqslant L,$$

as required.

Now assume that $n \geqslant 2$ and that the Proposition has been proved for forms in fewer than $n+1$ variables. To prove it for $F(X_0, \dots, X_n)$, we first apply Cassels's result to find a small integral zero $(x_0, \dots, x_n)$, with $x_0, \dots, x_n$ not all zero, satisfying

$$\max\{|x_0|, \dots, |x_n|\} \leqslant (3L)^{n/2}.$$

If $x_0 \neq 0$, then we are finished.

Otherwise, if $x_0 = 0$, then we can find an even smaller zero. For then the form $F(0, X_1, \dots, X_n)$ has a non-trivial integral zero, and now Cassels's result with one fewer variable supplies integers $x_1', \dots, x_n'$, not all zero, with

$$F(0, x_1', \dots, x_n') = 0, \quad M \leqslant (3L)^{(n-1)/2} \tag{5}$$

for

$$M = \max\{|x_1'|, \dots, |x_n'|\}.$$

Assume now that the point $x' = (x_0', x_1', \dots, x_n')$ (with $x_0' = 0$) is non-singular on the variety V defined by $F(X) = 0$ for $X = (X_0, \dots, X_n)$. This point is 'at infinity', and we attempt to shift away from infinity by drawing chords. Take a generic point t, and let the chord joining x' to t cut V in a new point. This new point is easily calculated (compare [**2**, p. 88]) to be proportional to

$$x'' = F(t, t)\, x' - 2F(t, x')\, t, \tag{6}$$

where $F(X, X')$ is the bilinear form attached to F, with $F(X, X) = F(X)$.

In other words, x'' is a zero of F for every t. We can now choose $t = (t_0, \ldots, t_n)$ so that the coordinate x_0'' of $x'' = (x_0'', \ldots, x_n'')$ is non-zero. Indeed,

$$x_0'' = -2F(t, x')\, t_0$$

has the shape $t_0(f_0 t_0 + \cdots + f_n t_n)$ for $f_0, \ldots, f_n$ independent of t; and x' being non-singular means that $f_0, \ldots, f_n$ are not all zero. Now it is easy to find $t_0, \ldots, t_n$, each either 0 or 1, such that $x_0'' \neq 0$ as desired. (For example, we should take $t_0 = 1$; this suffices if $f_1 = \cdots = f_n = 0$, and otherwise, if say $f_1 \neq 0$, then either $t_1 = 0$ or $t_1 = 1$ will do.)

Next, the definitions of L and M lead to

$$|F(t, t)| \leqslant L, \quad |F(t, x')| \leqslant LM,$$

so (6) gives

$$|x_i''| \leqslant L|x_i'| + 2LM|t_i| \leqslant 3LM \quad (0 \leqslant i \leqslant n).$$

Thus (5) yields

$$\max\{|x_0''|, \ldots, |x_n''|\} \leqslant 3L(3L)^{(n-1)/2} = (3L)^{(n+1)/2},$$

and therefore our desired small zero of F, at least if x' is non-singular.

If x' is singular, then we reduce to fewer variables and use induction. But the coefficients must stay under control. To ensure this, we can assume without loss of generality that $x_n' \neq 0$. Now x' is independent of the first n standard unit vectors $u_0, \ldots, u_{n-1}$, so we can define new variables $Y_0, \ldots, Y_n$ by

$$X = (X_0, \ldots, X_n) = Y_0 u_0 + \cdots + Y_{n-1} u_{n-1} + Y_n x'. \tag{7}$$

In particular, $X_0 = Y_0$. Because x' is singular, it follows that

$$F(X_0, \ldots, X_n) = F(Y_0 u_0 + \cdots + Y_{n-1} u_{n-1}) = G(Y_0, \ldots, Y_{n-1})$$

is a form G in the fewer variables $Y_0, \ldots, Y_{n-1}$. Furthermore, its coefficients constitute a subset of the coefficients of F. The given integral zero of F with $X_0 \neq 0$ supplies an integral zero of G with $Y_0 \neq 0$. So the induction hypothesis provides integers $y_0, \ldots, y_{n-1}$, with $y_0 \neq 0$, such that $G(y_0, \ldots, y_{n-1}) = 0$ and

$$\max\{|y_0|, \ldots, |y_{n-1}|\} \leqslant (3L)^{n/2}.$$

Now the values $Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}, Y_n = 0$ in (7) lead to $X_0 = y_0, \ldots, X_{n-1} = y_{n-1}$, $X_n = 0$, so that $(y_0, \ldots, y_{n-1}, 0)$ is our required small integral zero of F. This completes the proof of the Proposition.

To see that the exponent $(n+1)/2$ is best possible, it is more convenient to return to inhomogeneous polynomials as in the Theorem. Fix an integer $q \geqslant 2$, and consider

$$P(X_1, \ldots, X_n) = 2X_1 - 2q^2 - (X_2 - qX_1)^2 - \cdots - (X_n - qX_{n-1})^2,$$

where the squared linear terms are absent if $n = 1$. This has rational and even integral zeros, for example

$$X_1 = q^2, \quad X_2 = q^3, \quad \ldots, \quad X_n = q^{n+1}.$$

We proceed to show that any rational or even real zero $X_1 = \xi_1, \ldots, X_n = \xi_n$ is almost as big, and in fact

$$\xi_n \geqslant \tfrac{1}{2} q^{n+1}. \tag{8}$$

It is certainly clear that $\xi_1 \geqslant q^2$, so (8) is true if $n = 1$. If $n \geqslant 2$, then with $\eta_i = \xi_{i+1} - q\xi_i$, we have

$$\xi_n - q^{n-1}\xi_1 = \sum_{i=1}^{n-1} q^{n-1-i}\eta_i,$$

so the Cauchy–Schwarz inequality leads to

$$(\xi_n - q^{n-1}\xi_1)^2 \leqslant Q \sum_{i=1}^{n-1} \eta_i^2,$$

with

$$Q = \sum_{i=1}^{n-1} q^{2(n-1-i)} \leqslant 2q^{2(n-2)}.$$

Because $P(\xi_1, \ldots, \xi_n) = 0$, the right-hand side of the above is exactly

$$2Q(\xi_1 - q^2) \leqslant 2Q\xi_1.$$

It follows that

$$\xi_n \geqslant q^{n-1}\xi_1 - 2q^{n-2}\sqrt{\xi_1} = q^{n-2}\sqrt{\xi_1}(q\sqrt{\xi_1} - 2),$$

which is at least $q^{n-1}(q^2-2) \geqslant \frac{1}{2}q^{n+1}$, as asserted in (8).

In particular, $H(\xi) \geqslant \frac{1}{2}q^{n+1}$ for any rational zero ξ. But we see easily that

$$L = 2 + 2q^2 + (n-1)(1 + 2q + q^2) \leqslant 4nq^2.$$

Therefore $H(\xi) \geqslant cL^{(n+1)/2}$ for some $c > 0$ independent of q, and the exponent $(n+1)/2$ is best possible, as claimed.

Finally, the strategy used in the proof of our Theorem leads to a quick proof of the Hasse–Minkowski Theorem for quadratic polynomials; again it is more convenient to consider the equivalent version for forms F with zeros away from infinity $X_0 = 0$, as in the Proposition. If F has everywhere non-trivial local zeros, then the classical Hasse–Minkowski Theorem provides a non-trivial global zero x'. If x' is not at infinity, then we are finished. If x' is at infinity, then we attempt to shift away to x'', as in (6). We succeed if x' is non-singular; otherwise, if x' is singular, we reduce to fewer variables and use induction. The induction start, $n = 1$, is also easy, because if the discriminant is a square in $\mathbf{R}$ and each $\mathbf{Q}_p$, then it is a square in $\mathbf{Q}$ (see, for example, [**2**, Lemma 3.1, p. 78]).

References

1. J. W. S. Cassels, 'Bounds for the least solutions of homogeneous quadratic equations', *Proc. Cambridge Philos. Soc.* 51 (1955) 262–264; 52 (1956) 604.

2. J. W. S. Cassels, *Rational quadratic forms* (Academic Press, London, 1978).

3. F. J. Grunewald and D. Segal, 'How to solve a quadratic equation in integers', *Math. Proc. Cambridge Philos. Soc.* 89 (1981) 1–5.

4. D. Kornhauser, 'On the smallest solution to the general binary quadratic diophantine equation', *Acta Arith.* 55 (1990) 83–94.

5. D. Kornhauser, 'On small solutions of the general non-singular quadratic diophantine equation in five or more unknowns', *Math. Proc. Cambridge Philos. Soc.* 107 (1990) 197–211.

6. H. P. Schlickewei and W. M. Schmidt, 'Isotrope Unterräume rationaler quadratischer Formen', *Math. Z.* 201 (1989) 191–208.

Mathematisches Institut
Universität Basel
Rheinsprung 21
4051 Basel
Switzerland