ACTA ARITHMETICA
XVIII (1971)

On the product of three homogeneous linear forms

by

H. P. F. Swinnerton-Dyer (Cambridge)

1. Introduction. Let L_1, L_2, L_3 be three homogeneous linear forms in u, v, w with real coefficients and determinant Δ; and suppose that the lower bound of $|L_1 L_2 L_3|$ for integer values of u, v, w not all zero is 1. In a series of papers written some thirty years ago, Davenport investigated the values of Δ for which this is possible. In the fourth and last of these [2] he showed that

(1) $$\Delta = 7 \quad \text{or} \quad \Delta = 9 \quad \text{or} \quad \Delta > 9.1;$$

and he showed that in the first two cases $L_1 L_2 L_3$ must be the norm form associated with the totally real cubic field of discriminant 49 or 81 respectively.

The corresponding problem for the product of two linear forms is much easier, and was essentially completely solved by Markov. There are countably many possible values of Δ less than 3, each of which has the form

$$\Delta = (9 - 4n^{-2})^{1/2}$$

for some integer n; the first few values of n are $1, 2, 5, 13, 29, \ldots$ and there is an algorithm for constructing all the permissible values of n. Moreover, after multiplication by suitable constants, L_1 and L_2 are a linear form defined over a real quadratic field and its conjugate over the rationals. On the other hand, there are forms L_1, L_2 with $\Delta = 3$ which do not satisfy this last condition; and in any neighbourhood of 3 there are uncountably many distinct permissible values of Δ.

It is natural to ask whether there are any analogous phenomena for the present problem, and in particular whether the permissible values of Δ have a finite point of accumulation. It follows from standard techniques in the Geometry of Numbers that there is such a finite point of accumulation if and only if there is an admissible form $L_1 L_2 L_3$ such that (even after multiplication by suitable constants) L_1, L_2 and L_3 cannot be written as a linear form defined over a totally real cubic field and its two conjugates over the rationals. These questions seem very difficult to decide. The results of Cassels and Swinnerton-Dyer [1] tend to suggest that the answers are negative; in particular they show that what appears (by analogy with the case of two linear forms) to be the most natural way of constructing such products $L_1L_2L_3$ cannot succeed. If the answers to these questions are indeed negative, then the methods of Davenport [2] can be used to obtain all permissible values of Δ less than any preassigned bound. Conversely, if one finds all the permissible values of Δ less than a large enough bound, the behaviour of the sequence of values of Δ will strongly suggest what the answers to these two questions are. I have recently written programs to implement Davenport's method on a computer, and have thereby determined the eighteen permissible values of Δ less than 17, one of which arises from two essentially different products $L_1L_2L_3$. These values of Δ are

$$(2)\quad \begin{cases} 7,\ 9,\ \sqrt{148} = 12.1655,\ 63/5 = 12.6,\ 13\ \text{(twice)},\ 14,\ 351/25 \\ = 14.04,\ 189/13 = 14.5385,\ 133/9 = 14.7778,\ \sqrt{229} = 15.1327, \\ 259/17 = 15.2353,\ 559/35 = 15.9714,\ \sqrt{257} = 16.0312, \\ 273/17 = 16.0588,\ \sqrt{(539/2)} = 16.4165,\ 117/7 = 16.7143, \\ \sqrt{(22736/81)} = 16.7539,\ \sqrt{(2597/9)} = 16.9869. \end{cases}$$

The associated products are described in detail in § 5. In my opinion these results provide overwhelming evidence that the answers to the two questions posed at the beginning of this paragraph are negative; but they give no clue to the methods that would be needed for a proof.

One consequence of this conclusion is of particular interest. It is well known that if K is any algebraic number field other than the rationals, then each ideal class in K contains a non-zero ideal whose absolute norm is less than $|d|^{1/2}$, where d is the discriminant of K. Indeed, for any fixed r, s (in the standard notation), there is a constant $C_{r,s} < 1$ such that each ideal class contains a non-zero ideal of absolute norm at most $C_{r,s}|d|^{1/2}$; for example Davenport's result (1) shows that the least value of $C_{3,0}$ is $\frac{1}{7}$. Because this theorem is central to the only known way of computing the class number of an arbitrary K, it is important to obtain it in as strong a form as possible. If the conjecture made at the end of the last paragraph is true, it follows that, at least for totally real cubic fields, each ideal class of K contains a non-zero ideal whose absolute norm is $o(|d|^{1/2})$; and presumably this will happen whenever $r+s > 2$. If one is looking for empirical evidence, it is natural to ask for each $n > 0$ what is the totally real cubic field K_n of least discriminant d_n with the property that there is an ideal class of K_n containing an ideal of absolute norm n but no ideal of smaller absolute norm. I am indebted to Professor

H. J. Godwin for the following table, which is derived from his tables of class-numbers of all totally real cubic fields with $d < 20000$.

TABLE I. Fields K_n with $d_n < 20000$

n	d_n	$nd_n^{-1/2}$	h
1	49	0.1429	1
2	1957	0.0452	2
3	2597	0.0589	3
4	4312	0.0609	3
5	3969	0.0794	3
7	8281	0.0769	3
9	17689	0.0679	3

This table is too short to provide any clear indication, but the irregularities which it displays are interesting — especially in view of the missing entry at $n = 6$, which must have $d_6 > 20000$ and so $nd^{-1/2} < 0.0425$. (But it should be noted that K_6 and even more K_8 must have fairly large class number, and it may be because of this that they have large discriminants.)

When a theorem has been proved with the help of a computer, it is impossible to give an exposition of the proof which meets the traditional test — that a sufficiently patient reader should be able to work through the proof and verify that it is correct. Even if one were to print all the programs and all the sets of data used (which in this case would occupy some forty very dull pages) there can be no assurance that a data tape has not been mispunched or misread. Moreover, every modern computer has obscure faul s in its software and its hardware — which so seldom cause errors that they go undetected for years — and every computer is liable to transient faults. Such errors are rare, but a few of them have probably occurred in the course of the calculations reported here. However, the calculation consists in effect of looking for a rather small number of needles in a six-dimensional haystack; almost all the calculation is concerned with parts of the haystack which in fact contain no needles, and an error in those parts of the calculation will have no effect on the final results. Despite the possibilities of error, I therefore think it almost certain that the list of permissible $\Delta \leqslant 17$ in (2) is complete; and it is inconceivable that an infinity of permissible $\Delta \leqslant 17$ have been overlooked.

Nevertheless, the only way to verify these results (if this were thought worth while) is for the problem to be attacked quite independently, by a different program written by someone else for a different machine. This corresponds exactly to the situation in most experimental sciences. It seems right therefore to describe this work in the way one would describe an experiment — that is, to give the results and an account of the methods used which is detailed enough to enable the work to be repeated and (I hope) to convince the reader that the problem has been attacked in a sensible and competent way.

The methods of the present paper are those of Davenport [2], which such modifications as are needed if the calculation is not to involve intelligent choice. I am deeply indebted to Professor Davenport for his advice and encouragement during this work. I am also grateful to Professor M. V. Wilkes, the Director of the Cambridge University Mathematical Laboratory, for making available to me the substantial amount of machine time needed for the calculations.

2. The division into cases. It is desirable to normalize the linear forms L_i as far as possible; and here there is a complication because we cannot assume in advance that $|L_1L_2L_3|$ attains its minimum. It is essential for what follows that $|L_1L_2L_3|$ should take the value 1, and there are two ways to ensure this:

(i) We consider in the first instance only those forms $|L_1L_2L_3|$ which attain their minimum. By a general theorem, to each permissible value of Δ there do correspond such forms; and if for a given value of Δ the only forms of this type which occur are ones which are isolated in the sense of Lemma 2 below (as happens for $\Delta \leqslant 17$), then for this value of Δ there are no forms $|L_1L_2L_3|$ which do not attain their minimum.

(ii) We choose a small constant $\varepsilon > 0$ and replace the hypothesis that the minimum of $|L_1L_2L_3|$ is 1 by the hypothesis that $|L_1L_2L_3|$ does take the value 1 and that

$$|L_1L_2L_3| > 1-\varepsilon$$

for all integers u, v, w not all zero. The machine calculations remain unaltered, since the factor $(1-\varepsilon)$ is swallowed up in the allowance made for round-off error; but the theory becomes more complicated because some of the formulae acquire additional terms $O(\varepsilon)$.

Davenport [2] adopted the second of these methods, because at that time the general theory had not been developed; however the first method is simpler and we adopt it here.

After a unimodular integral transformation on u, v, w we may therefore assume that $|L_1L_2L_3| = 1$ at $(1, 0, 0)$; and by multiplying the L_i by suitable constants whose product is ± 1 we can further assume that

(3) $$L_i = u+\alpha_i v+\beta_i w \quad (i = 1, 2, 3).$$

Consider now the positive definite quadratic form

(4) $$(L_1-L_2)^2+(L_2-L_3)^2+(L_3-L_1)^2 = 2(Av^2+Bvw+Cw^2)$$

say, whose determinant is

$$4AC-B^2 = 3\Delta^2.$$

By applying a suitable integral unimodular transformation to v, w we can further assume that this quadratic form is reduced, so that $|B| \leqslant A \leqslant C$. Write

$$S(\gamma) = \tfrac{1}{2}\{(\gamma_1-\gamma_2)^2+(\gamma_2-\gamma_3)^2+(\gamma_3-\gamma_1)^2\}$$

so that $A = S(\alpha)$ and $C = S(\beta)$; then it follows from what we have just done that $\Delta \leqslant \Delta_0 = 17$ implies

(5) $$S(\alpha) \leqslant \Delta_0, \quad S(\beta) \leqslant \tfrac{1}{4}\{S(\alpha)+3\Delta_0^2/S(\alpha)\}.$$

It is only through this pair of conditions that we shall use the fact that the quadratic form (4) is reduced. The second condition does impose an upper bound on $S(\beta)$, for it will be shown in Lemma 1 below that $S(\alpha) \geqslant 7$.

None of the α_i, β_i can be rational, because none of the L_i can take the value 0. For any integers n_1, n_2, n_3 denote by $\{n_1, n_2, n_3\}$ the set of points $(\gamma_1, \gamma_2, \gamma_3)$ such that

(6) $$|(m+\gamma_1 n)(m+\gamma_2 n)(m+\gamma_3 n)| \geqslant 1$$

for all integers m, n not both zero, and also

$$n_i > \gamma_i > n_i - 1 \quad (i = 1, 2, 3).$$

Clearly each of α and β is in just one such set. Reverting to (3), writing $-v$ for v replaces each α_i by $-\alpha_i$, and writing $u+nv$ for u replaces each α_i by α_i+n; and neither of these action essentially alters $L_1L_2L_3$ nor affects the reduction conditions (5). To each $\{n_1, n_2, n_3\}$ we associate a *region type*, which is the union of $\{n_1, n_2, n_3\}$ and all the sets that can be derived from it by permuting the n_i, adding the same integer to all of them, and possibly changing the sign of all of them. Clearly this induces a partition of the union of all $\{n_1, n_2, n_3\}$ into disjoint region types; and any γ satisfying (6) belongs to just one $\{n_1, n_2, n_3\}$ and to just one region type.

LEMMA 1. *$S(\gamma) \geqslant 7$ for every γ satisfying* (6). *Moreover, the only types containing a point γ with $S(\gamma) \leqslant 33$ are those in Table II below; and for each such type the number in the third column is a lower bound* (not necessarily best possible) *for $S(\gamma)$.*

Proof. For any given type it is enough to look at the points in an arbitrary one of the components $\{n_1, n_2, n_3\}$. Without loss of generality we can assume $n_1 \leqslant n_2 \leqslant n_3$; and (6) certainly implies $n_3 - n_1 \geqslant 2$. Thus

$$S(\gamma) = \tfrac{3}{4}(\gamma_3-\gamma_1)^2+\tfrac{1}{4}(2\gamma_2-\gamma_1-\gamma_3)^2 > \tfrac{3}{4}(n_3-n_1-1)^2,$$

so that $S(\gamma) \leqslant 33$ implies $n_3 - n_1 < 8$ and hence there are only finitely many types that need to be considered.

Suppose first that $n_1 < n_2 < n_3$ and define r, s, t by

$$(m+\gamma_1 n)(m+\gamma_2 n)(m+\gamma_3 n) = m^3+rm^2n+smn^2+tn^3;$$

thus $S(\gamma) = r^2-3s$. The six conditions obtained from (6) by setting $m = n_i$ or n_i-1, $n = -1$ are linear inequalities on r, s, t and define a convex polyhedron in (r, s, t) space; and to find the minimum of r^2-3s in this polyhedron is a matter of standard technique. This method was used by Davenport ([2], Lemma 11) to prove that $S(\gamma) \geqslant 7$ always and to obtain lower bounds for $S(\gamma)$ in region types A, B and C, and a detailed account can be found there. Because of the number of cases to be considered (and the possibility of a more elaborate subdivision which turned out not to be worth while) it was implemented on a computer, giving the results shown in Table II.

TABLE II. Region types compatible with $S(\gamma) \leqslant 33$

Region type	Typical set	Lower bound for $S(\gamma)$
A	$\{-1, 0, 2\}$	7
B	$\{-2, 0, 2\}$	8.75
C	$\{-2, 1, 2\}$	10.11
D	$\{-2, 0, 3\}$	13.56
E	$\{-3, 1, 2\}$	15.75
F	$\{-3, 2, 2\}$	21
G	$\{-3, 0, 3\}$	19.95
H	$\{-3, 1, 3\}$	20.44
I	$\{-3, 2, 3\}$	23.56
J	$\{-3, 3, 3\}$	29.67
K	$\{-3, 0, 4\}$	28.01
L	$\{-3, -1, 4\}$	29.36

If say $n_2 = n_3$, the argument becomes more elaborate but is essentially of the same kind. This concludes the description of the proof of the lemma.

We use henceforth the region type names of Table II. It follows from (5) with $\Delta_0 = 17$ and from $S(\gamma) \geqslant 7$ that $S(\beta) \leqslant 33$; so the only types that have to be considered are those of Table II. We shall say for example that a set of forms (3) is of type BC if α is of type B and β is of type C; then by (5) and Table II the only types that need to be considered are AA to AL, BA to BI, CA to CI, DA to DE and EA to EE. Having reached this conclusion, we make no further use of (5); henceforth therefore there is symmetry between α and β, and we need not consider for example both type BC and type CB.

Now suppose, as a typical case, that we are looking for forms of type AC. By means of the linear transformations mentioned above, and the possibility of permuting the L_i, we can without loss of generality assume that α is in $\{-1, 0, 2\}$ and that β is in $\{-2, 1, 2\}$ or one of the sets derived from it by permutation of coordinates. Each type therefore breaks up into six cases (or less if there is any symmetry). Not all these cases need to be examined in detail. For example, if α is in $\{-1, 0, 2\}$ and β in $\{-2, 1, 2\}$, then $\alpha-\beta$ is in $\{2, -1, 0\}$ or $\{2, -1, 1\}$ since the other possible $\{n_1, n_2, n_3\}$ are empty; and since each of these regions for $\alpha-\beta$ is of type A the resulting set of forms (3) will already have been found (with a change of variables) in the investigation of type AA. Similarly with type AB, if α is in $\{-1, 0, 2\}$ and β in $\{0, 2, -2\}$, consideration of $\alpha+\beta$ shows that we can further assume $\alpha_1+\beta_1 < -2$ and $\alpha_2+\beta_2 > 1$. In this way we have reduced our problem to a finite number of cases, some of which have extra conditions. There turn out to be fifty-three such cases; this enumeration was done by hand and, since it is the part of the work most liable to error, was checked by Professor Davenport.

3. The investigation of individual cases. The next part of the investigation, which accounts for almost all the machine time used, consists in reducing these fifty-three cases to subcases; the aim is to come down to a limited number of subcases, in each of which the α_i and β_i are so tightly bounded that the methods of § 5 can be applied.

At any moment in the calculation, an individual case or subcase is described by means of three convex polygons R_1, R_2, R_3; here R_i is a polygon in the (α_i, β_i) plane and the conditions so far imposed on α_i, β_i are equivalent to requiring the point (α_i, β_i) to lie in R_i. Because of the machinery of § 2, R_i is initially a unit square — or a half-square or quarter-square if there are supplementary conditions. In the course of the calculation the R_i are reduced in size by the two processes of *refinement* and *subdivision* described below. The object of the calculation is for each eventual subcase either to reduce one of the R_i to the empty set or to reduce all the R_i to sets of very small diameter, so that one can guess an admissible associated set of forms L_j and then use Lemma 2 to prove that this is the only admissible set of forms associated with this triplet of R_i. Empirically, each triplet of very small R_i produced by these calculations is associated with such a set of forms L_j; had this not been so, the proof of (2) could not have been completed, but instead one would have found a small region of very great interest. In the programs as originally written, 'small diameter' was taken to mean diameter less than 10^{-7}; some definition has to be built into the program because this is one of the conditions for stopping the iteration in the refinement. However, examination of intermediate results showed that in a few cases the program was refining an already small region in a highly inefficient manner; and in these cases the program was stopped when the R_i had diameter about 10^{-5} and the calculation was finished by hand.

An individual step of the refinement works as follows, where to simplify the notation we assume that we are refining R_1. Choose suitable integers u, v, w in a way that will be described below; here v and w are not both zero and u is the integer nearest to either the upper or the lower bound of $-a_1v-\beta_1w$ in R_1. Let M_2 be an upper bound for $|u+a_2v+\beta_2w|$ in R_2, and define M_3 similarly. (The R_i are described inside the computer by keeping lists both of their sides and of their vertices, each ordered in the obvious way. The M_i are now easy to compute because each of them is attained at one of two well-defined vertices of R_i.) Since $|L_1L_2L_3| \geqslant 1$, we must certainly have

$$(7) \qquad |u+a_1v+\beta_1w| \geqslant (1-10^{-8})/M_2M_3$$

where the factor $(1-10^{-8})$, instead of the natural 1, is included so as to overcome any errors caused by round-off. (The computer used works to 13 significant figures in the decimal scale, and a detailed examination of the programs and the upper bounds for v, w implicit in them shows that this allowance is more than adequate. An earlier version of the program corrected for round-off at each step, but this was very slow.) It follows that (a_1, β_1) cannot lie in the strip S_1 defined by

$$(1-10^{-8})/M_2M_3 > u+a_1v+\beta_1w > -(1-10^{-8})/M_2M_3.$$

There are now four possibilities for the relation between R_1 and S_1:

(i) R_1 does not overlap S_1. In this case (7) is of no value; and it is unlikely that this triplet u, v, w will ever again be useful for refining R_1. In fact it can only be useful if subsequent refinements of R_2 and R_3 sufficiently diminish M_2 and M_3.

(ii) S_1 completely covers R_1. Now the case being considered is impossible.

(iii) S_1 overlaps R_1, and $R_1\setminus S_1$ is connected. Now we can replace R_1 by $R_1\setminus S_1$, which is a strictly smaller convex polygon; thus we have refined R_1.

(iv) S_1 overlaps R_1 but $R_1\setminus S_1$ is not connected. Now (7) is not immediately useful, but this triplet u, v, w may become useful when R_1 has been further refined.

At first sight, it would seem that (iv) should be used to subdivide cases; but this would lead to difficulties in the organization of storage space, and in any case it is essential not to let subcases proliferate too rapidly. Indeed, experience shows that one should only subdivide when there is no further scope for refinement.

A refinement of R_2 or R_3 may diminish M_2 or M_3, and therefore increase the size of the strip S_1 for given u, v, w; this is why refinement must be an iterative process. To obtain a worth-while improvement in R_1

the strip S_1 must not be too narrow — thus either v and w must both be small integers or u, v, w must be such that one of M_2 and M_3 is substantially smaller than the magnitudes of u, v, w would suggest. Even with the most favourable choice of u, v, w, M_2 will not be substantially smaller than

$$\{\mathrm{Max}(|v|, |w|)\}\{\text{Diameter of } R_2\}$$

and a combination of theoretical and empirical arguments suggest that one should confine oneself to triplets u, v, w such that $M_2M_3 < 100$ and

$$|v|, |w| < 2\{\text{largest of the diameters of the } R_i\}^{-2/3}.$$

The two processes of refinement and subdivision are carried out by the same program. On starting to consider a case or subcase, the program sets up for each i a list of triplets u, v, w which are likely to be useful in refining R_i; these lists are based on criteria similar to those of the previous paragraph, but more complicated. These lists remain fixed throughout the calculations on a given case or subcase. After these lists have been set up, the program alternates between a refinement mode and a subdivision mode, starting in the refinement mode. At any given moment the program has four data lists, each item in any list being a triplet of regions R_i. The lists are

(i) R^*, the triplet of regions R_i currently being worked on;

(ii) $\mathscr{L}_1$, a list of triplets which will be further processed as part of the present case or subcase;

(iii) $\mathscr{L}_2$, a list of triplets which cannot usefully be further processed as part of the present case or subcase, but which will have to be processed in later runs of the program with new lists of useful u, v, w;

(iv) $\mathscr{L}_3$, a list of triplets which are so small that they do not need to be further refined by this program.

Of these lists, R^* consists of one triplet and $\mathscr{L}_1$ contains at most six triplets; both these must be held in immediate access store. $\mathscr{L}_2$ is being built up to be used as data in a subsequent run of the program; it may contain several hundred triplets but can be held in backing store. $\mathscr{L}_3$ must be printed out, but is very short and usually empty. Initially, R^* is given by the data which determine the case or subcase being considered, and $\mathscr{L}_1, \mathscr{L}_2$ and $\mathscr{L}_3$ are all empty.

During a refinement stage, the triple R^* is progressively refined in the way described some paragraphs ago, using in turn each of the triplets, u, v, w in each of the three lists. (There are modifications to ensure that triplets u, v, w which are almost certainly not useful are not used.) This refinement is repeated until either the set of R_i involved is shown to be impossible or the improvement in R_i obtained by going through the list of u, v, w once more is negligible — this being defined to mean that none of the upper or lower bounds for α_i or β_i in R_i is improved by as much as 10^{-8}. When either of these two events occurs, the program shifts into the subdivision mode. If at this point there is a triplet R^*, the program subdivides it; if not, it goes straight into the 'tidying up' step.

The process of subdividing R^* works as follows:

(i) If all three R_i in R^* have diameter less than 10^{-7}, put this triplet into the list $\mathscr{L}_3$ and proceed to the 'tidying up' stage.

(ii) Otherwise, for a preassigned integer N_0 which is fixed throughout the consideration of a given case or subcase, let $\mathrm{Max}\,\alpha_1$ denote the least upper bound of α_1 in R_1 and similarly for $\mathrm{Min}\,\alpha_1$, and write

$$n = \text{Integer part of } (N_0 \,\mathrm{Max}\,\alpha_1).$$

If $n > N_0 \,\mathrm{Min}\,\alpha_1$, put the triplet defined by $R_1 \cap (\alpha_1 < n/N_0), R_2, R_3$ into the list $\mathscr{L}_1$, take the triplet defined by $R_1 \cap (\alpha_1 > n/N_0), R_2, R_3$ to be the new triplet R^*, and return to the refinement phase. We shall call this process 'splitting on α_1'.

(iii) If it is not possible to split on α_1, try successively to split on $\beta_1, \alpha_2, \beta_2, \alpha_3, \beta_3$.

(iv) If none of these are possible, put the current triplet R^* into $\mathscr{L}_2$ and proceed to the tidying up stage.

The tidying up stage is reached whenever the current R^* has been disposed of in any way other than by splitting. It works as follows:

(i) If $\mathscr{L}_1$ is not empty, take as the new R^* that triplet in $\mathscr{L}_1$ which was most recently put into $\mathscr{L}_1$; and delete it from $\mathscr{L}_1$. Then return to the refinement stage.

(ii) If $\mathscr{L}_1$ is empty, consideration of this case or subcase is complete. The program now either terminates or takes a new case or subcase from the data tape.

When the program ends, one is left with the two lists $\mathscr{L}_2$ and $\mathscr{L}_3$. In practice, $\mathscr{L}_2$ may contain up to a thousand triplets (each of which corresponds to a subcase of the original case), while $\mathscr{L}_3$ contains two or three triplets at most. The way of dealing with $\mathscr{L}_3$ is described in § 4. Each triplet in $\mathscr{L}_2$ represents a new subcase, which is fed back into the program again with a new and larger value of N_0. The values of N_0 used were 12, 24, 48, 96, 200, 500 and 1000; it is not economical to increase N_0 too rapidly, because this has the effect that a region is rejected only in very small bits.

The organization of the splitting, and the 'last in, first out' treatment of $\mathscr{L}_1$, are designed to ensure that $\mathscr{L}_1$ is kept small. This is essential when working on a multi-access computer on which programs which need a great deal of immediate access store have very low priority.

There is one further complication still to be mentioned. The description above explains how the program works if it is dealing with a case or subcase completely. This was done for the cases of types AA to AH, BB and BC. But for most of the other initial cases, a large part of the six-dimensional region $R_1\times R_2\times R_3$ consists of points for which the corresponding forms $L_1L_2L_3$ have $\Delta > \Delta_0 = 17$, and for the present purpose these points do not have to be considered. The way in which the condition $\Delta \leqslant 17$ was used is as follows. We have

$$\Delta = \alpha_1(\beta_2-\beta_3)+\alpha_2(\beta_3-\beta_1)+\alpha_3(\beta_1-\beta_2)$$

and we can assume the coordinates so ordered that $\Delta > 0$. If for example the triplet of regions R_i is such that $\mathrm{Min}\,\beta_3 > \mathrm{Max}\,\beta_1$, then to replace α_2 by $\mathrm{Min}\,\alpha_2$ can only decrease Δ. Replacing in this way $\alpha_2, \alpha_3, \beta_2$ and β_3 by their upper or lower bounds (whichever is appropriate) we obtain a linear form $\Lambda_1(\alpha_1, \beta_1)$ such that in $R_1\times R_2\times R_3$ we have $\Delta \geqslant \Lambda_1(\alpha_1, \beta_1)$. Hence we may impose on R_1 the additional condition $\Lambda_1(\alpha_1, \beta_1) \leqslant 17$, and similarly for R_2 and R_3. These conditions were used at the beginning of each refinement stage.

Ideally, the result of the calculations described in this section would have been simply a list of very small regions, each of which could have been dealt with by the methods of § 4. In practice, because of the fairly crude methods used to produce the list of u, v, w for the refinement process, there was one other region which resisted the efforts of the computer. For this region

$$-3.461752 > \alpha_1 > -3.461860, \quad -0.803853 > \beta_1 > -0.803980,$$
$$-0.683282 > \alpha_2 > -0.683344, \quad 2.885317 > \beta_2 > 2.885243,$$
$$2.935778 > \alpha_3 > 2.935357, \quad 0.447989 > \beta_3 > 0.447860;$$

and it was verified by hand that R_1 is covered by the strips given by $(u, v, w) = (1, -9, 40)$, $(7, -14, 69)$, $(15, -24, 122)$, $(21, -29, 151)$, $(22, -38, 191)$, $(28, -43, 220)$, $(43, -67, 342)$, $(62, 20, -9)$, $(101, 38, -38)$, $(146, 51, -38)$, $(163, 58, -47)$, $(225, 78, -56)$, $(388, 136, -103)$, $(417, 146, -110)$. These values are given explicitly to show the intricacy of the patchwork that is involved; and the way in which everything fits together is one of the main reasons for believing the conjecture stated in the Introduction.

4. The isolation theorem. We have still to show how to deal with the triplets of small regions which appear in the lists $\mathscr{L}_3$ — or in some cases were removed from the lists $\mathscr{L}_2$ because the computer was making heavy weather of them. For each of these triplets it was possible to find an associated admissible form $L_1L_2L_3$ by intelligent guesswork — had this not been so one would have had to proceed as in the end of § 3 — and it remains to show that this is the only admissible form associated with this triplet. In other words one must show that any admissible form $\hat{L}_1\hat{L}_2\hat{L}_3$ near $L_1L_2L_3$ is essentially the same as $L_1L_2L_3$. This requires an 'isolation theorem'; and an isolation theorem of the kind required is proved as Theorem 2 of [1]. However although the proof is constructive the theorem itself is for simplicity not stated in a constructive form; we therefore reformulate it here in the form in which it is actually used.

Suppose that $L_1L_2L_3$ is admissible, where the L_i are given by (3), L_1 is defined over a totally real cubic field, and L_2 and L_3 are the conjugates of L_1 over the rationals. These conditions hold for the admissible forms which we are trying to isolate. Any form near $L_1L_2L_3$ can be given by linear forms $\hat{L}_i$ such that

$$\hat{L}_1 = (1+\varepsilon_{11})L_1+\varepsilon_{12}L_2+\varepsilon_{13}L_3$$

and so on, where the ε_{ij} are small. For simplicity we consider instead the linear forms L_i^* where

$$L_1^* = (1+\varepsilon_{11})^{-1}\hat{L}_1 = L_1+\theta_{12}L_2+\theta_{13}L_3$$

and so on; and the explicit bounds for the regions R_i provide (very small) explicit upper bounds for the size of the ε_{ij} and θ_{ij} that have to be considered. We have therefore to show how to find explicit (not too small) constants δ, δ_0, which will depend on $L_1L_2L_3$, such that

(8) $\left\{\begin{array}{l}\textit{if } 0<\operatorname{Max}|\theta_{ij}|<\delta_0 \textit{ then there exist integers } u, v, w \textit{ not all zero such} \\ \textit{that } |L_1^*L_2^*L_3^*|<1-\delta.\end{array}\right.$

Here δ has to be large enough to ensure

$$(1-\delta)(1+\varepsilon_{11})(1+\varepsilon_{22})(1+\varepsilon_{33})<1. \tag{9}$$

It is convenient to break the problem up into twelve cases, depending on which is the absolutely largest of the θ_{ij} and what is its sign. Thus for example, having chosen δ to ensure (9) we look for δ_{12}^+ such that (8) holds with $\delta_0=\delta_{12}^+$ provided that the θ_{ij} also satisfy the condition

$$\operatorname{Max}|\theta_{ij}| = |\theta_{12}| \quad \text{and} \quad \theta_{12}>0. \tag{10}$$

If we did not have to worry about the magnitude of δ_{12}^+ we would proceed as follows. Choose integers u, v, w such that $L_1L_2<0$, and denote by ξ_i the value of L_i for these values of u, v, w. Provided $\theta_{12}>0$ is small enough, we can find a unit ζ of the lattice (that is, an integral unimodular transformation on u, v, w which takes each L_i into ζ_iL_i) such that $|\zeta_2|\gg|\zeta_3|\gg|\zeta_1|$ and $\theta_{12}\xi_2\zeta_2/\xi_1\zeta_1$ is as close to -1 as we like. The equations $L_i=\xi_i\zeta_i$ define integers u, v, w; and in view of (10) we have $L_2^*=L_2(1+o(1))$, $L_3^*=L_3(1+o(1))$ and

$$L_1^* = L_1+\theta_{12}L_2+o(L_1) = L_1\{1+\theta_{12}\xi_2\zeta_2/\xi_1\zeta_1+o(1)\} = o(L_1).$$

Thus $L_1^*L_2^*L_3^* = o(L_1L_2L_3) = o(\xi_1\xi_2\xi_3)$, and we can certainly ensure that this is absolutely less than $1-\delta$. However, the argument in this form leads to too small a value of δ_{12}^+ in some cases. This accounts for the extra complications below.

Let η be a unit of the lattice, chosen once for all, such that

$$|\eta_2|>|\eta_3|>|\eta_1|, \quad \eta_1\eta_2>0$$

and of course $\eta_1\eta_2\eta_3=\pm1$; and let T be the corresponding integral unimodular transformation on u, v, w. Choose m integral points $P_j = (u_j, v_j, w_j)$ for $j=1,2,\ldots,m$ and for all $n>0$ write $P_{j+mn}=T^nP_j$. Choose m positive constants

$$c_0>c_1>\ldots>c_{m-1}>c_m = c_0\eta_1/\eta_2$$

and for all $n>0$ write $c_{j+mn}=c_j(\eta_1/\eta_2)^n$. Denote by λ_{ij} the value of L_i at P_j and for convenience write

$$\mu_{2j} = 1+c_{j-1}(|\lambda_{1j}|+|\lambda_{3j}|)/|\lambda_{2j}|,$$
$$\mu_{3j} = 1+c_{j-1}(|\lambda_{1j}|+|\lambda_{2j}|)/|\lambda_{3j}|,$$
$$\mu_j = \frac{1-\delta}{\mu_{2j}\mu_{3j}|\lambda_{1j}\lambda_{2j}\lambda_{3j}|}-\frac{c_{j-1}|\lambda_{3j}|}{|\lambda_{1j}|}.$$

Clearly increasing j by m will decrease μ_{2j} and μ_{3j} and will increase μ_j.

LEMMA 2. *Suppose that for $j=1,2,\ldots,m$ we have*

$$\lambda_{1j}\lambda_{2j}<0, \quad |1+c_j\lambda_{2j}/\lambda_{1j}|<\mu_j, \quad |1+c_{j-1}\lambda_{2j}/\lambda_{1j}|<\mu_j; \tag{11}$$

then (8) *with the additional condition* (10) *holds with* $\delta_0=\delta_{12}^+=c_0$.

Proof. On considering the effect of increasing j by m, it is clear that (11) holds for all $j>0$. Since $c_j\to0$ as $j\to\infty$, we can choose j so that

$$c_{j-1}\geqslant\theta_{12}\geqslant c_j; \tag{12}$$

and so every θ is absolutely bounded by c_{j-1}. Now consider the values of the L_i^* at P_j; we have

$$|L_2^*/\lambda_{2j}| = |1+\lambda_{1j}\theta_{21}/\lambda_{2j}+\lambda_{3j}\theta_{23}/\lambda_{2j}|\leqslant\mu_{2j},$$
$$|L_3^*/\lambda_{3j}| = |1+\lambda_{1j}\theta_{31}/\lambda_{3j}+\lambda_{2j}\theta_{32}/\lambda_{3j}|\leqslant\mu_{3j},$$
$$|L_1^*/\lambda_{1j}| = |1+\lambda_{2j}\theta_{12}/\lambda_{1j}+\lambda_{3j}\theta_{13}/\lambda_{1j}|$$
$$\leqslant|1+\theta_{12}\lambda_{2j}/\lambda_{1j}|+|\lambda_{3j}\theta_{13}/\lambda_{1j}|$$
$$<\mu_j+c_{j-1}|\lambda_{3j}/\lambda_{1j}| = (1-\delta)/\mu_{2j}\mu_{3j}|\lambda_{1j}\lambda_{2j}\lambda_{3j}|,$$

the last inequality coming from (11) and (12). Multiplying these three results together gives $|L_1^*L_2^*L_3^*|<1-\delta$ as required. This proves the lemma.

TABLE III. List of admissible forms

Form number	Field equation	d	α	β	$d^{1/2}\Delta^{-1}$	Δ	Ideal
I	$\theta^3-2\theta^2-\theta+1$	7^2	θ	$\theta^2-2\theta$	1	7	Principal
II	$\theta^3-3\theta^2+1$	9^2	θ	$\theta^2-3\theta+1$	1	9	Principal
III	$\theta^3+2\theta^2-2\theta-2$	148	θ	$\theta^2+2\theta-1$	1	12·1655	Principal
IV	$\theta^3+9\theta^2+6\theta-1$	63^2	$(2\theta^2+15\theta-3)/5$	$(\theta^2+10\theta+11)/5$	5	12·6	Non-principal
V	$\theta^3+\theta^2-4\theta+1$	13^2	θ	$\theta^2+2\theta-2$	1	13	Principal
VI	$\theta^3+4\theta^2-25\theta-1$	91^2	$(2\theta^2+3\theta-5)/21$	$(\theta^2+12\theta-13)/21$	7	13	Non-principal
VII	$\theta^3-2\theta^2-\theta+1$	7^2	θ^2	$\theta^2-2\theta$	$\frac{1}{2}$	14	No
VIII	$\theta^3+\theta^2-4\theta+1$	13^2	$(3\theta^2-7)/5$	$(27\theta^2+45\theta-3)/25$	25/27	14·04	No
IX	$\theta^3-2\theta^2-\theta+1$	7^2	$(27\theta^2-45\theta-3)/13$	$(3\theta^2-18\theta+4)/13$	13/27	14·5385	No
X	$\theta^3+16\theta^2+41\theta-1$	133^2	$(\theta^2+12\theta+2)/9$	$(\theta^2+21\theta+11)/27$	9	14·7778	Non-principal
XI	$\theta^3+\theta^2-5\theta+2$	229	θ	$\theta^2+2\theta-3$	1	15·1327	Principal
XII	$\theta^3+44\theta^2+41\theta-1$	259^2	$(6\theta^2+253\theta+40)/119$	$(\theta^2+45\theta+1)/17$	17	15·2353	Non-principal
XIII	$\theta^3+64\theta^2+61\theta-1$	3913^2	$(2\theta^2+127\theta+11)/35$	$(\theta^2+66\theta+18)/35$	245	15·9714	Non-principal
XIV	$\theta^3-5\theta+3$	257	θ	$\theta^2+\theta$	1	16·0312	Principal
XV	$\theta^3+51\theta^2+48\theta-1$	819^2	$(5\theta^2+254\theta+26)/51$	$(40\theta^2+193\theta+82)/153$	51	16·0588	No
XVI	$\theta^3-5\theta^2-15\theta-1$	4312	$(\theta^2-2\theta-5)/8$	$(\theta^2-6\theta-1)/4$	4	16·4165	Non-principal
XVII	$\theta^3+12\theta^2+9\theta-1$	117^2	$(\theta^2+15\theta+5)/7$	$(2\theta^2+23\theta+3)/7$	7	16·7143	Non-principal
XVIII	$\theta^3+\theta^2-37\theta-1$	5684	$(\theta^2+2\theta+1)/9$	$(\theta+1)/3$	9/2	16·7539	No
XIX	$\theta^3+8\theta^2+12\theta-1$	2597	$(\theta^2+4\theta-1)/3$	$(\theta^2+7\theta+2)/3$	3	16·9869	Non-principal
XX	$\theta^3+309\theta^2+306\theta-1$	94563^2	$(116\theta^2+35809\theta+20952)/5491$	$(21\theta^2+6530\theta+3225)/5491$	5491	17·2215	Non-principal

The only difficulty in applying this lemma is the choice of the points $P_1, P_2, \ldots, P_m$. If there is a point P for which $L_1L_2L_3 = \pm 1$ and the L_i have the desired signs, then one can take $m = 1$ and choose P_1 to be P or a transform of it by some power of T. This happens in all but four of the cases listed in § 5. In the remaining four cases the P_j were chosen from among the triplets (u, v, w) corresponding to the sides of the regions R_i, and their transforms by powers of T. Once the P_j have been chosen, it is easy to choose the c_j by means of (11), since one knows in advance that μ_j will be approximately $|\lambda_{1j}\lambda_{2j}\lambda_{3j}|^{-1}$.

5. The results. In Table III we list the nineteen inequivalent admissible forms $L_1L_2L_3$ with $\Delta \leqslant 17$, together with the one other form which falls within the cases (AA to AH, BB and BC) that were dealt with completely. The first column gives a reference number for the form. The second column gives a cubic polynomial whose zeros θ_i are used to define α_i and β_i, and hence L_i by (3). The third column gives the discriminant of the totally real cubic field $Q(\theta)$, this being written as a square when $Q(\theta)$ is a normal extension of Q. The fourth and fifth columns give α_i and β_i as functions of θ_i. The sixth column gives the rational number $d^{1/2}\Delta^{-1}$, thus enabling Δ to be found explicitly, and the seventh column gives Δ in decimal form. The last column records whether the Z-module generated by $1, \alpha, \beta$ is a (fractional) ideal in $Q(\theta)$, and if so whether it is a principal ideal.

In lattice language, VII corresponds to a sublattice of index 2 in I, and perhaps should not really be counted as an independent form. The forms VIII, X, and XVIII represent ± 1 in two essentially different ways and so have two possible descriptions, only one of which is given here; and XVI represents ± 1 in three essentially different ways. The forms which do not have units of arbitrary signature are XI, XIV, XVIII and XIX; for each of these just half the signatures are possible.

References

[1] J. W. S. Cassels and H. P. F. Swinnerton-Dyer, *On the product of three homogeneous linear forms and indefinite ternary quadratic forms*, Phil. Trans. Roy. Soc., A, 248 (1955), pp. 73–96.

[2] H. Davenport, *On the product of three homogeneous linear forms IV*, Proc. Camb. Phil. Soc. 39 (1943), pp. 1–21.

TRINITY COLLEGE
Cambridge, England

Received on 1. 4. 1970