FURTHER CONTRIBUTION TO THE GEOMETRY OF NUMBERS FOR NON-CONVEX REGIONS

BY

L. J. MORDELL

1. Introduction. The most important contribution to the question of finding the minimum value of $|f(x, y)|$ for integer values of x, y, where f is a real function of its arguments, is given by Minkowski's fundamental theorem in the Geometry of Numbers which can be stated as follows:

THEOREM. *Let $f(x, y)$ be a single-valued function of x and y which satisfies the following conditions:*
  (I) $f(0, 0) = 0$.
  (II) $f(x, y) > 0$ *unless* $x = y = 0$.
  (III) $f(tx, ty) = tf(x, y)$, *where* $t > 0$.
  (IV) $f(x_1 + x_2, y_1 + y_2) \leq f(x_1, y_1) + f(x_2, y_2)$.
  (V) $f(-x, -y) = f(x, y)$.
Then there exists at least one set of integer values of x, y such that

$$(1) \qquad\qquad 0 < f(x, y) \leq 2/J^{1/2},$$

where J is the area of the region $f(x, y) \leq 1$.

The result can be put in a slightly different form by writing

$$x = \alpha\xi + \beta\eta, \qquad y = \gamma\xi + \delta\eta$$

where $\Delta \equiv \alpha\delta - \beta\gamma > 0$, and making ξ, η run through all integer values. Then x, y describes the points of a lattice L say, and a point of L now satisfies the conditions

$$(2) \qquad\qquad 0 < f(x, y) \leq 2J^{-1/2}\Delta^{1/2}.$$

This in general is not the best possible result. When f is given and L varies but with Δ fixed, Minkowski has considered the question of replacing the constant $2J^{-1/2}$ in (2) by the least possible value. He shows that this leads to an extremal problem of the following type:

PROBLEM. *To find the minimum value of $|ad - bc|$ when a, b, c, d satisfy the equations*

$$(3) \qquad f(a, b) = 1, \qquad f(c, d) = 1, \qquad f(a + c, b + d) = 1,$$

that is, to find a parallelogram of minimum area one of whose vertices is at the origin O and whose other three vertices are on the boundary $f(x, y) = 1$.

These problems, though simple enough in theory, often prove very difficult in practice.

The geometrical significance of Minkowski's conditions I–V is that the region in (1) is bounded, convex and symmetrical with respect to the origin. Apparently, he did not attempt to extend his results to non-convex or infinite domains though he deduced a result for the product of linear forms by applying the inequality of the arithmetic and geometric means. Very recently [3]([1]), however, I showed that such regions were amenable to treatment and gave a number of best possible results, for example, for the regions

$$\left| x^3 + y^3 \right| \leqq 1; \quad \left| xy(x + y) \right| \leqq 1; \quad \left| x^p + y^p \right| \leqq 1, \ .33 \cdots < p \leqq 1;$$
$$\left| x^4 - y^4 \right| \leqq 1,$$

by using methods of some generality.

Other proofs for the first two regions were found by Davenport [1] and myself. The last two examples, however, are particular cases of a general theorem for regions $\mathfrak{R}$ resembling those defined in these examples. My method, which was very different from that of Minkowski, was to develop a theory adequate in principle for the type of region dealt with, and then to reduce the question to an ordinary maximum and minimum problem, really an elementary extremal problem. The field of variation of a variable point P, say, defining the problem is an appropriate small region abutting the given region $\mathfrak{R}$. The next step is to show that P can be restricted to an appropriate small part of the boundary of $\mathfrak{R}$. Then the problem takes the form: to find the minimum value of the area of the parallelogram $OPQR$, where Q, R are points on the boundary of $\mathfrak{R}$ or perhaps, and this is a simplification, on lines associated with $\mathfrak{R}$. The geometrical configuration is often such as to suggest the position of P for which the minimum value occurs, but the actual proof may be troublesome and difficult.

Further, my methods are such as to suggest a method of attack for many non-convex or infinite domains of a simple character. I also suggested the generalization of the region typified by $\left| x^3 + y^3 \right| \leqq 1$ as a further problem.

Mahler [2] then took up the subject and showed that Minkowski's ideas could be applied to very general types of non-convex regions, namely star-domains, that is, those in which every radius vector from the origin meets the boundary in just one point. The generality of his method means that concrete applications may require the consideration of many sub-cases and attention to a great many details, especially numerical ones, since the discussion involves a number of extremal problems; for it is precisely the special properties of the region $\mathfrak{R}$ which then become of paramount importance. This is exemplified in the case $\left| x^5 + y^5 \right| \leqq 1$ which he considered later in a manuscript which I have read but which is not to be submitted for publication.

([1]) Numbers in brackets refer to the references cited at the end of the paper.

The more general case

$$(4) \qquad\qquad |x^n + y^n| \leq 1, \quad n > 1,$$

where $n = p/q$ and p, q are both odd positive integers, had been in my thoughts for some time. Here $x^n = \operatorname{sgn} x|x|^n$, and so our region is

$$|\operatorname{sgn} x|x|^n + \operatorname{sgn} y|y|^n| \leq 1.$$

In this form, it has, as remarked by Dr. Mahler, a meaning for all real n; and my discussion of (4) later on in this paper also applies to this.

I take now, however, the region $|f(x, y)| \leq 1$, say $\mathfrak{R}$, where f is such that $\mathfrak{R}$ is very similar to the region (4), and develop a self-contained treatment which requires no knowledge of previous work. In its course, the conditions imposed upon $f(x, y)$ are satisfied when $f(x, y) = x^n + y^n$, or for that matter $\operatorname{sgn} x|x|^n + \operatorname{sgn} y|y|^n$. Shortly after I began, Dr. Mahler also commenced working independently upon this special case when $n \geq 4$ as a generalization of his work on the case $n = 5$.

Before any results can be established, we must have some idea or picture of a geometrical configuration defined by four simultaneous equations in four unknowns given in (1) of §3. I am greatly indebted to Dr. Mahler for the approximate solution of these equations in the case $n = 5$ above. I do not use these numerical results, but they have been of the greatest value in picturing the state of affairs, though of course our ideas and methods are entirely different. I am also much indebted to him for some general inequalities[2] which he obtained in his work on (4) when $n \geq 4$, and which enable me to finish off my work very simply in the special case of (4), when $n \geq 4$. He has also added to my obligations by drawing several diagrams including the one given in this paper. I am also indebted to Dr. Mahler, Mrs. Ollerenshaw, and Prof. Davenport for many valuable comments on my manuscript and for the removal of many obscurities.

2. **The function $f(x, y)$ and the region $|f(x, y)| \leq 1$.** We suppose $f(x, y)$ defined for all real x, y as a single-valued, symmetrical, continuous, real function of x, y, homogeneous (and, without loss of generality, of dimension one), and we suppose also that $f(x, y) = -f(-x, -y)$. We suppose that $f(x, y)$ is a non-decreasing function of x when y is constant. We take the units of x, y so that the curve $f(x, y) = 1$ meets the axes of x, y at the points $(1, 0)$, $(0, 1)$, and suppose that the curve crosses the tangents $x = 1$, $y = 1$ at these points, as occurs, for example, when they are points of inflexion. Let $\mathfrak{R}$ be the region $|f(x, y)| \leq 1$. We suppose that the part of $\mathfrak{R}$ in the first quadrant is a convex region, and that the part of the boundary of $\mathfrak{R}$ lying in the fourth quadrant

above $x+y=0$ is concave, so that except for its end points the straight line joining any two points of this boundary lies outside the region $0 \leqq f(x, y) \leqq 1$. We suppose that the line $x+y=0$ is an asymptote to $f(x, y)=1$, and that every other line meets the curve in 1 or 3 points except that the points $(1, 0)$, $(0, 1)$ may occur any number of times as intersections of the corresponding tangents $x=1$, $y=1$. Clearly any line $y=px$ from the origin except $x+y=0$ meets $f(x, y) = 1$ in exactly one point, since $f(x, px) = xf(1, p) = 1$.

The equation $f(x, y) = 1$ defines $y=g(x)$ as a single-valued, steadily decreasing, continuous function of x which we suppose twice differentiable for $-\infty < x < \infty$ except at $x=1$. Then $g'(0)=0$, $g'(1) = -\infty$. Also if the line $x-y=0$ meets the curve at (x_0, y_0), then $g'(x_0) = -1$. From the convexity and concavity conditions the following results are obvious:

$$\text{for} \quad -\infty \leqq x \leqq 0, \qquad\qquad -1 \leqq g'(x) \leqq 0;$$
$$\text{for} \quad\;\; 0 \leqq x \leqq x_0, \qquad\qquad 0 \geqq g'(x) \geqq -1;$$
$$\text{for} \quad\;\; x_0 \leqq x \leqq 1, \qquad\qquad -1 \geqq g'(x) \geqq -\infty;$$
$$\text{for} \quad\;\; 1 \leqq x \leqq \infty, \qquad\qquad -\infty \leqq g'(x) \leqq -1.$$

Also $g(x)+x \to 0$ as $x \to \pm \infty$; and for $x < 0$ or $x > 1$, $1-x > g(x) > -x$.

Also $x = g(y)$.

A simple case of the functions typified by $f(x, y)$ is

$$f(x, y) = (x^n + y^n)^{1/n},$$

where $n > 1$ is a rational number of the form p/q, where p, q are both odd and positive. It will suffice to prove that all lines other than $x+y=0$, $x=1$, $y=1$ meet the curve $f(x, y) = 1$ in either one or three real points. Then in particular the tangent at any point other than $(1, 0)$, $(0, 1)$ meets the curve in one other real point.

For let the line be $rx+sy=1$, where we have excluded $(r, s)=(1, 0)$, $(0, 1)$, (∞, ∞), that is, $x+y=0$. Then $t=x/y$ is given by the equation([3])

$$E(t) \equiv t^n + 1 - (rt + s)^n = 0.$$

Now

$$E'(t)/n = t^{n-1} - r(rt + s)^{n-1},$$

and so the equation $E'(t)=0$ has two real roots if $r>0$ and none if $r<0$. If $r<0$, $E(\infty) = +\infty$, $E(-\infty) = -\infty$. Hence if $r<0$, and clearly also if $r=0$ (or if $s \leqq 0$), there will be one real intersection since $E(t)$ is a monotone function of t. If $r>0$ (or if $s>0$), the values of $E(t)$ for real t form three monotone sets and so clearly $E(t)$ has either one or three real zeros.

3. A system of equations. We suppose that $f(x, y)$ is as defined in §2 and

---

([3]) We use the notation $E(t)$, E for various equations and expressions occurring throughout the paper, not necessarily the same at different places.

now consider the system of equations

(1) $\quad f(a, b) = 1, \quad f(c, d) = 1, \quad f(a + c, b + d) = 1, \quad f(c - a, d - b) = 1.$

Clearly, if $(a, b, c, d) = (A, B, C, D)$ is one solution, then $(a, b, c, d) = (B, A, D, C)$ is also one, leading to two sets of solutions. We shall prove that for those in one set, say the first set,

$$a > 1, \quad b < 0, \quad c < 0, \quad d > 1.$$

We shall then show that

$$a > d, \quad b < c.$$

Since then $a+c>b+d$, and $f(b+d, b+d) < f(a+c, b+d) = 1 < f(a+c, a+c)$, we have

(2) $\qquad (b + d)f(1, 1) < 1, \qquad (a + c)f(1, 1) > 1.$

Hence if (x_0, x_0) or (y_0, y_0) is the point on $f(x, y) = 1$ with $x = y$,

$$b + d < y_0, \qquad a + c > y_0.$$

We shall also give, in §(6) below, sufficient conditions for the existence of a solution with $b > -1/2$, and we shall suppose from §4 onwards that such a solution exists.

We deduce from (1), since $a = g(b)$, $c = g(d)$, and so on, the two equations

(3) $\qquad F(b, d) \equiv g(b) + g(d) - g(b + d) = 0,$

(4) $\qquad G(b, d) \equiv g(d) - g(b) - g(d - b) = 0.$

We now study the equation (3).

LEMMA 1. *If b, d are any real numbers satisfying $F(b, d) = 0$, then $bd < 0$.*

Now

(5) $\qquad \partial F(b, d)/\partial b = g'(b) - g'(b + d).$

Suppose first that b and d are both negative. Then $\partial F(b, d)/\partial b > 0$, since the concavity condition means that $g''(x) \geq 0$ for $x < 0$, and so $g'(x)$ decreases when $x < 0$ and decreases. Hence $F(b, d)$ decreases when $b \leq 0$ and decreases. Also $F(0, d) = g(0) = 1$, $F(-\infty, d) = g(d) + d > 0$, and so $F(b, d)$ cannot vanish when $b < 0$, $d < 0$.

Suppose next that b and d are both positive or zero. Several cases now arise according to the positions of b, d in the intervals $(0, 1)$ or $(1, \infty)$.

Suppose first that $0 \leq b \leq 1$, $0 \leq d \leq 1$. Then from (3), since $g(b) \geq 0$, $g(d) \geq 0$, $g(b+d) \geq 0$ and so $b+d \leq 1$. Now from (5), $F(b, d)$ steadily increases when b increases and $0 \leq b \leq 1 - d$. Also $F(0, d) = 1$, $F(1-d, d) = g(1-d) + g(d) > 0$. Hence $F(b, d)$ cannot vanish for $0 \leq b \leq 1 - d$.

Suppose next that $b>1$, $d\geq0$. From (5), $F(b, d)$ steadily decreases as b increases since the concavity condition means that $g'(x)$ increases when $x>1$ increases. Also $F(1, d)=g(d)-g(1+d)>0$, $F(\infty, d)=g(d)+d>0$. Hence we cannot have $b>1$, $d\geq0$.

Similarly we cannot have $b\geq0$, $d>1$. The three cases considered mean that we have excluded $b\geq0$, $d\geq0$. Hence, as we have also excluded $b\leq0$, $d\leq0$, we must have $bd<0$ and b, d cannot have the same sign, for example, if $b<0$, then $d>0$.

Lemma 2. *If $b<0$, then $d>1$.*

For suppose $d\leq1$. If $b+d>0$, then $0<g(b+d)<1$, and so from (3), since $g(b)>1$, $F(b, d)>0$. Hence we must have $b+d<0$. Now $F(b, d)$ increases as b decreases from $-d$ to $-\infty$, since from (5), $\partial F(b, d)/\partial b<0$ because $g'(x)$ decreases as x decreases when $x<0$. Also $F(-d, d)=g(-d)+g(d)-1>0$, $F(-\infty, d)=g(d)+d>0$. Hence $F(b, d)\neq0$ for $b<0$, $d\leq1$, and so if $b<0$, then $d>1$.

Then from $a=g(b)$ and so on, $a>1$, $c<0$. By symmetry, if $d<0$, then $b>1$, and from $a=g(b)$ and so on, $a<0$, $c>1$. Hence in (1), all the solutions of the first set have $b<0$[4].

We can now study the equation (4) subject to the restriction $b<0$, $d>1$, and prove the following lemma.

Lemma 3. *$G(b, d)=0$ defines d as a steadily increasing function of b, when $b<0$ and decreases. Also*

$$1 < d < g(b), \qquad g(b) + g(1 - b) < 0,$$

or, since $a=g(b)$, $d<a$. Also $d<1-b$.

Clearly $\partial G(b, d)/\partial b = -g'(b)+g'(d-b)<1-1=0$. Next $\partial G(b, d)/\partial d = g'(d)-g'(d-b)<0$, since $g'(x)$ steadily increases when $x>1$ and increases. Since

$$\frac{\partial G(b, d)}{\partial d}\frac{dd}{db} + \frac{\partial G(b, d)}{\partial b} = 0,$$

$dd/db<0$, and so d increases as b decreases.

Next $G(b, d)$ as a function of d steadily decreases as d increases and $d\geq1$. Also

---

[4] Professor Davenport gives the following simple proof that $a>1$, $b<0$, $c<0$, $d>1$. At least one of a, c must be positive. For if $a\leq0$ and $c\leq0$, then, by the concavity of $f(x, y)=1$ for $x\leq0$, we would have $f((a+c)/2, (b+d)/2)\geq1$, contrary to (1). Similarly, at least one of b, d must be positive. Also at least one of a, b must be negative. For if $a\geq0$, $b\geq0$ then $a+c\geq c$, $b+d\geq d$ whence $f(a+c, b+d)>f(c, d)=1$, since the inequality must be strict in at least one case. Similarly at least one of c, d must be negative. There are only two possibilities: $a<0$, $b>0$, $c>0$, $d<0$ or $a>0$, $b<0$, $d>0$, $c<0$. Without loss of generality we can consider the latter, that is, $a>1$, $b<0$, $c<0$, $d>1$.

$$G(b, g(b)) = b - g(b) - g[g(b) - b]$$
$$< - E - g(E), \text{ say,}$$
$$< 0,$$

since $E > 1$. Hence $G(b, d)$ cannot vanish for $d \geq g(b)$ and therefore $d < g(b) < 1 - b$. Also, since $d > 1$,

$$G(b, 1) = - g(b) - g(1 - b) > 0.$$

Hence

$$\psi(b) = g(b) + g(1 - b) < 0.$$

Also

$$\psi'(b) = g'(b) - g'(1 - b).$$

Since $b < 0$,

$$0 > g'(b) > - 1, \qquad - \infty < g'(1 - b) < - 1.$$

Hence $\psi'(b) > 0$ and so $\psi(b)$ decreases steadily when $b < 0$ and decreases. Since $\psi(0) = 1$, values of b must satisfy $0 > b_0 > b$, where b_0 is the solution of $g(b) + g(1 - b) = 0$. Then also $b_0 > - 1/2$ if $g(-1/2) + g(3/2) < 0$, a condition which is satisfied in our special case $g(x) = (1 - x^n)^{1/n}$ since $3^n - 1 > 2^{n+1}$ for $n \geq 3$.

We return to the equation $F(b, d) = 0$ and study d as a function of b when $b < 0$ and $1 < d < 1 - b$. *We now impose a condition upon $f(x, y)$, namely that $g'(-x) > g'(x)$ for $0 < x \leq 1$.* This condition is obviously satisfied if, for $0 \leq x \leq 1$, $g(x)$ can be expanded in a series of positive ascending powers of x all of whose coefficients, except possibly the constant term, are negative, for example, when $g(x) = (1 - x^n)^{1/n}$, where $n > 1$. We then show that d decreases steadily when b decreases from 0 to $-1/2$.

For from (5),

$$\partial F(b, d)/\partial b = g'(b) - g'(b + d) > g'(- b) - g'(b + d).$$

But $g'(x)$ decreases as x increases from $x = 0$ to $x = 1$, and hence, if $-b < b + d$ or $d + 2b > 0$, a condition certainly satisfied if $b > - 1/2$, then $\partial F(b, d)/\partial b > 0$. Also

$$\partial F(b, d)/\partial d = g'(d) - g'(b + d).$$

Since $d > 1$, $- \infty < g'(d) < -1$, and since the point $(b+d, a+c)$ has a vectorial angle greater than $45°$, $0 > g'(b+d) > -1$, and so $\partial F(b, d)/\partial d < 0$. Hence $dd/db > 0$ and the result.

Now $G(b, d) = 0$ defines d as a steadily increasing function of b when $b < 0$ and decreases; also when $d = 1$, $b = b_0$, where b_0 is defined by $g(b_0) + g(1 - b_0) = 0$. But from $F(b, d) = 0$, $d > 1$ when $b = b_0$. Hence the equations $G(b, d) = 0$, $F(b, d) = 0$ will have exactly one solution with $-1/2 < b < b_0$ if $d_1 > d_2$, where

$$G(- 1/2, d_1) = 0, \qquad F(- 1/2, d_2) = 0,$$

and also by Lemma 3, $1 < d_1, d_2 < g(b)$.

LEMMA 4. *A sufficient condition for this is now shown to be*

$$(6) \qquad -2g(-1/2) - g(3/2) + g[-1/2 + g(-1/2)] > 0.$$

For, since

$$g(b) + g(d_2) - g(b + d_2) = 0,$$

$$-g(d_2) = g(b) - g(b + d_2) < g(b) - g[b + g(b)].$$

Also

$$g(d_1) - g(b) - g(d_1 - b) = 0,$$

$$-g(d_1) = -g(b) - g(d_1 - b) > -g(b) - g(1 - b).$$

But $d_1 > d_2$ if $-g(d_1) > -g(d_2)$, that is, if

$$-g(b) - g(1 - b) > g(b) - g[b + g(b)],$$

or

$$(7) \qquad -2g(b) - g(1 - b) + g[b + g(b)] > 0.$$

The result follows on putting $b = -1/2$.

If we have proved that $b > -1/2$ (and so also $c > -1/2$), the inequality $g(x) < 1 - x$ gives $a = g(b) < 1 - b < 3/2$. Write

$$\Delta = ad - bc.$$

Then also

$$(8) \qquad \Delta > 3/4.$$

We can also show that the equations (3), (4) cannot have more than one solution for which $b < 0$, $b + d < y_0$. For on subtracting the equations (3), (4),

$$E \equiv 2g(b) - g(b + d) + g(d - b) = 0.$$

Now

$$\partial E/\partial d = -g'(b + d) + g'(d - b) < 1 - 1 = 0.$$

$$\partial E/\partial b = 2g'(b) - g'(b + d) - g'(d - b)$$

$$= g'(b) - g'(b + d) + g'(b) - g'(d - b) > 0.$$

Hence $dd/db > 0$ and so d is a steadily decreasing function of b when $b < 0$ and decreases and of course $b + d < y_0$. Since $G(b, d) = 0$ defines d as a steadily increasing function in the same circumstances, the equations $E = 0$, $G(b, d) = 0$ cannot have more than one root.

4. The region $|f(x, y)| \leq 1$. The sets of values (a, b), (c, d), $(a+c, b+d)$, $(c-a, d-b)$ whose existence has just been established define four points Q, R, S, T respectively. These, with their images in the origin, say, Q', R', S', T', form a parallelogram whose vertices S, T, S', T' and middle points

of sides lie on the boundary of $\Re$, that is, on $|f(x, y)| = 1$.

It will be convenient to use X, Y to denote coordinates referred to axes OQ, OR, and so we can write

$$x = aX + cY, \qquad y = bX + dY.$$

Then Q, R, S, T become the points

$$(X, Y) = (1, 0), \ (0, 1), \ (1, 1), \ (-1, 1),$$

respectively; thus the line QS becomes $X = 1$. It meets the curve $f(x, y) = 1$ in one further point U.

We now make an important assumption about our geometrical configuration. We suppose that U is either below OQ or at Q. A sufficient condition for this is now shown to be

$$\lambda = cf_a + df_b \leqq 0,$$

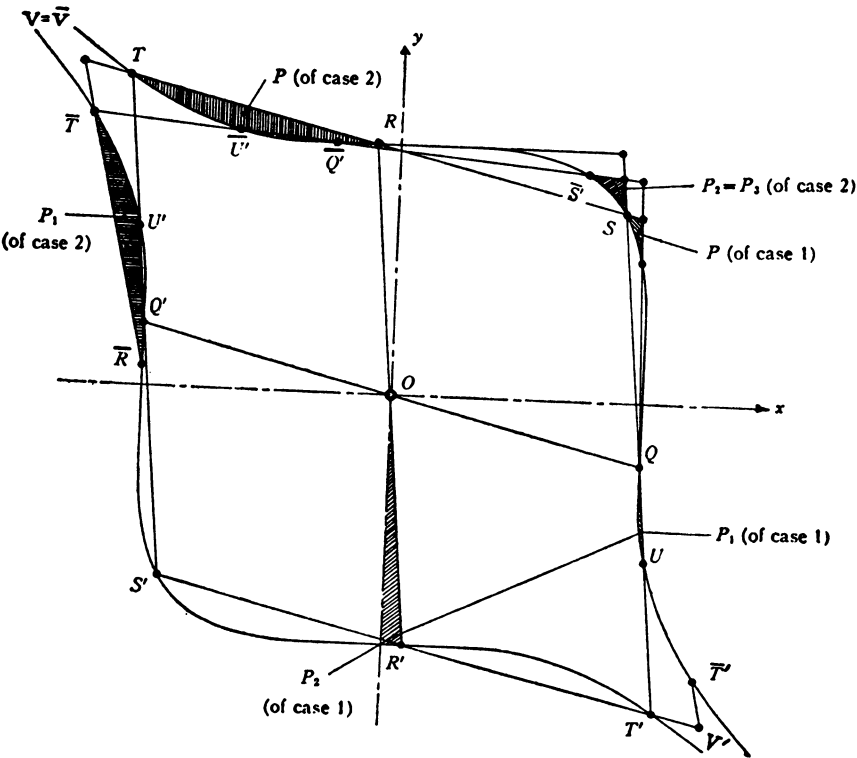where $f_a = \partial f(a, b)/\partial a, f_b = \partial f(a, b)/\partial b$.

FIG. 1.

For QS, that is, $X = 1$, meets the curve $f(x, y) = 1$ where

$$E(Y) = f(a + cY, b + dY) - 1 = 0.$$

Now $E(-1) = -2$, $E(0) = 0$, $E(1) = 0$. Since $E(Y) = 0$ has only three real roots, the third root will be negative, that is, U will be below OQ if $E(Y)$ is decreasing for small $Y \geq 0$. But $E'(0) = cf_a + df_b$, and the result follows, the equality sign holding only when U is at Q.

Regardless, however, of the position of U,

$$\mu = af_c + bf_d < 0$$

where $f_c = \partial f(c, d)/\partial c$, $f_d = \partial f(c, d)/\partial d$.

For TRS meets the curve $f(x, y) = 1$ where

$$E_1(X) \equiv f(aX + c, bX + d) - 1 = 0.$$

Also $E_1(-1) = 0$, $E_1(0) = 0$, $E_1(1) = 0$, $E_1(X) \sim Xf(a, b) \rightarrow + \infty$ as $X \rightarrow \infty$; and so, since $E_1(X) = 0$ has only three real roots, $E_1'(0) < 0$.

LEMMA 5. *A sufficient condition for this geometrical configuration, that is, that U should be below OQ, is that $2g'(-1/2) + 1 > 0$ and*

$$g(1 + y) + g(1 - y) - 2g'(y) < 0$$

for $0 \geq y \geq -1/2$.

The condition is $cf_a + df_b < 0$, which, since $a - g(b) = 0$, that is, $f_b/f_a = -g'(b)$, and as is shown in a moment $f_a > 0$, becomes $c - dg'(b) < 0$. Since

$$c + a = g(d + b), \qquad c - a = g(d - b), \qquad 2c = g(d + b) + g(d - b),$$

the condition is

$$g(d + b) + g(d - b) - 2dg'(b) < 0.$$

Consider now the function of the two independent variables

$$E = g(x + y) + g(x - y) - 2xg'(y),$$

where $x \geq 1$, and $0 \geq y \geq -1/2$.

$$\partial E/\partial x = g'(x + y) + g'(x - y) - 2g'(y).$$

Now

$$
\begin{array}{ll}
0 \geq g'(x + y) \geq - \infty & \text{if} \quad 0 \leq x + y \leq 1, \\
-1 \geq g'(x + y) \geq - \infty & \text{if} \quad x + y \geq 1, \\
-1 \geq g'(x - y) \geq - \infty & \text{since} \quad x - y \geq 1, \\
g'(y) \geq g'(-1/2) & \text{since} \quad 0 \geq y \geq - 1/2.
\end{array}
$$

Hence $\partial E/\partial x \leq -1 - 2g'(-1/2)$. We assume now that $2g'(-1/2) + 1 > 0$, a

condition obviously satisfied when $g(y) = (1 - y^n)^{1/n}$ since $1 - 2(1/2)^{n-1} > 0$ when $n > 2$. Hence E is a steadily decreasing function of x. Hence if $E < 0$ when $x = 1$, then $E < 0$ when $x \geqq 1$. Since we know that $0 > b \geqq -1/2$, the result follows. If we know that b satisfies the inequalities $b_1 \geqq b \geqq b_0$, line 2 of Lemma 5 will have now $2g'(b) + 1 > 0$, and line 4, $b_1 \geqq y \geqq b_0$, instead of $2g'(-1/2) + 1 > 0$ and $0 \geqq y \geqq -1/2$ respectively.

We now prove some inequalities involving the partial derivatives f_a, f_b, f_c, f_d, namely:

LEMMA 6. $f_a > 1, f_d > 1, 0 < f_b, f_c < 2/3, f_c < f_b, f_d < f_a, f_b f_d - f_a f_c > 0$.

From the concavity condition, the tangent at Q lies between the curve $f(x, y) = 1$ and the line QS, that is, $X = 1$, and hence makes intercepts on the x, y axes greater than those made by QS. The line $X = 1$ is given by $dx - cy = ad - bc = \Delta$, and the tangent at Q by $xf_a + yf_b = 1$. Hence

$$\Delta/d < 1/f_a < 1, \qquad \Delta/-c < 1/f_b.$$

Since from (8) of §3, $\Delta > 3/4$, and $d > 0$, $-c > 0$, then $f_a > 1$, $f_b > 0$, and $f_b < -c/\Delta < (1/2)/(3/4) = 2/3$. Similarly from the tangent at R which makes intercepts on the x, y axes greater than those made by RS,

$$\Delta/a < 1/f_d < 1, \qquad \Delta/-b < 1/f_c,$$

and the result follows.

Next, since $d < a$, the point (d, c) lies on the curve $f(x, y) = 1$ between $(1, 0)$ and (a, b). Hence from the concavity condition, the tangent at Q makes smaller intercepts on the x, y axes than the tangent at (d, c), and has a greater slope to the x axis than the tangent at (d, c). Hence first $1/f_a < 1/f_d$, $1/f_b < 1/f_c$, and so $f_d < f_a$, $f_c < f_b$. Next $-f_d/f_c < -f_a/f_b$ or $f_b f_d - f_a f_c > 0$.

We now show that the tangent at Q meets RS produced at a point whose distance from R is less than $2RS$.

For the tangent at Q is $xf_a + yf_b = 1$, that is $\lambda Y + X = 1$, where

$$\lambda = cf_a + df_b > cf_a > -f_a/2.$$

The intercept on $Y = 1$ is given by $X = 1 - \lambda > 1$, since $\lambda < 0$ from the hypothesis that U lies below Q. Also

$$1/f_a > (ad - bc)/d = a - (bc/d) > 1 - (1/4d) > 3/4,$$

and so $f_a < 4/3$, $\lambda > -2/3$, $1 - \lambda < 5/3 < 2$, as stated.

The tangent also meets the line $y = 1$ where $x = (1 - f_b)/f_a$, that is, $0 < x < 1$.

Similarly the tangent at $R(c, d)$ meets the line $x = 1$, where $y = (1 - f_c)/f_d$, that is, $0 < y < 1$.

Denote now by $\overline{Q}(-b, -a)$ and so on, the image of $Q(a, b)$ and so on, in the asymptote $x + y = 0$. Then $\overline{Q}'$ is the image of Q in the line $x - y = 0$. I prove now that $\overline{Q}'$ lies between the points $R(c, d)$ and $T(c - a, d - b)$. The con-

dition for this is that $c-a<b<c$ which is obvious from previous results since $b>-1/2, a>1$.

From the region $\mathfrak{R}$ we cut off a finite region $\mathfrak{S}$ as follows. Let RT and $\overline{R}\overline{T}$ meet on the asymptote at the point V or $\overline{V}$, and so $R'T'$ and $\overline{R}'\overline{T}'$ meet at V'. Then $\mathfrak{S}$ is the finite part of $\mathfrak{R}$ cut off by the lines $VT, \overline{V}\overline{T}, V'T', \overline{V}'\overline{T}'$.

5. The main theorem. We can now state that the main object of this paper is to find sufficient conditions for the truth of the following theorem.

THEOREM. *Let L be any lattice of determinant $\Delta = ad-bc$. Then a point of L other than O lies in $\mathfrak{S}$. This may always be taken as an inner point of $\mathfrak{S}$ except when L is the critical lattice*

$$x = aX + cY, \qquad y = bX + dY,$$

where X, Y run through all integer values, or its image in the line $x+y=0$, namely

$$x = bX + dY, \qquad y = aX + cY.$$

Then $\mathfrak{S}$ contains on its boundary the points (a, b), (c, d), $(a+c, b+d)$, $(c-a, d-b)$ of L and also their images in $x+y=0$.

We find as sufficient a condition really of an extremal character, namely the following:

CONDITION. *The hyperbola through S whose asymptotes are the tangents at Q and R does not intersect the curve $f(x, y)=1$ in real points lying between S and the point where the tangent at Q meets the curve $f(x, y)=1$ again, or for that matter the point $x=1, y=0$.*

It is obvious from homogeneity considerations that the theorem implies that any lattice of determinant less than Δ has one of its points other than the origin O as an inner point of $\mathfrak{S}$.

The question arises whether a critical lattice of $\mathfrak{S}$ has points other than O in the interior of $\mathfrak{R}$, that is, do integers X, Y, not both zero exist such that

$$(1) \qquad\qquad |f(aX + cY, bX + dY)| < 1.$$

In some cases, $f(x, y)$ may reduce to a polynomial in X, Y with rational coefficients, and then the answer may be easy. Thus when $f(x, y)=x^3+y^3$, on adding and subtracting the last two equations in (1) the condition becomes

$$|X^3 - XY^2 + Y^3| < 1,$$

which is obviously impossible. But if $f(x, y)=x^n+y^n$, where n is a positive odd integer greater than 3, (1) has algebraic numbers for its coefficients and it is not an easy matter to answer the question. I have made no serious effort to do so at present, but the question seems to be worth consideration.

6. Lattices. We gather together some elementary properties of lattices

required later and these are stated as lemmas.

LEMMA 7. *Let $p(x_1, y_1)$ and $q(x_2, y_2)$ be two points of L such that the area of the triangle Opq is $\Delta/2$, that is, $|x_1y_2 - x_2y_1| = \Delta$. Then L can be generated by vector addition from p and q, that is, every point of L can be represented in the form $\rho p + \sigma q$, with integers ρ, σ.*

LEMMA 8. *Let $p(x_1, y_1)$ be a point of L other than O such that no point of L is an inner point of Op. Let l be a line parallel to Op and at distance Δ/Op from it, so that l has the equation $xy_1 - yx_1 = \pm\Delta$. Then any segment of length Op contains exactly one point of L except that it will contain two points of L when the end points of the segment are points of L.*

LEMMA 9. *Any parallelogram of area not less than 4Δ and centre at O contains a point of L other than O.*

LEMMA 10. *A parallelogram $OACB$ of area Δ cannot contain two points of L other than O unless either they are collinear with O, or one is at A and the other on BC, or one is at B and the other on AC.*

LEMMA 11. *If $p(x_1, y_1)$ and $q(x_2, y_2)$ are two points of L such that the area of the triangle Opq is Δ, then the midpoint of either Op, or of Oq, or of pq is a point of L.*

7. Reduction of the problem to two cases. The parallelogram $STS'T'$ has area 4Δ and so contains a point P of L other than O. There are now several alternatives possible:

(I) P is an inner point of $\mathfrak{R}$ and then obviously of $\mathfrak{S}$;

(II) P is a boundary point of both the parallelogram and of $\mathfrak{R}$, that is, P is at Q, R, S, T, U;

(III) P lies in the small curvilinear region QU, excluding the end points Q, U allowed for in (II), say the region $\{QU\}$;

(IV) P lies in the small curvilinear region RT, excluding the end points R, T counted in (II), say the region $\{RT\}$.

If (I) holds there is nothing more to be proved. Suppose next that (II) holds and let first P be at Q. Then by Lemma 8, either one point of L lies in RS and is an inner point of $\mathfrak{R}$ and of $\mathfrak{S}$, or R, S are both points of L which then becomes the critical lattice $x = aX + cY$, $y = bX + dY$. A similar argument holds for R, S, T on using the lines $QS, Q'R, QR$ respectively. When P is at U, a line through R parallel to OU obviously meets the boundary $f(x, y) = 1$ in a point to the right of QS and so gives a segment lying in $\mathfrak{R}$ and $\mathfrak{S}$ and of length greater than OU and so contains a point of L which is an inner point of $\mathfrak{S}$.

We now deal with the cases (III) and (IV) and, by Lemma 10, these are mutually incompatible. Then also from the same lemma, there cannot be

more than one point of L in either region $\{QU\}$ or $\{RT\}$ unless they are collinear with O. Then we take P to be that point nearest the origin O.

8. **The first case.** We prove now the following lemma.

LEMMA 12. *A point of L cannot lie in the region $\{QU\}$ if the hyperbola through S whose asymptotes are the tangents at Q and R does not intersect the curve $f(x, y) = 1$ in real points lying between S and the point where the tangent at Q meets again the curve $f(x, y) = 1$.*

For let a point of L, say for convenience of notation P_1 given by (X, Y) $= (\Xi_1, H_1)$ or by $(x, y) = (\xi_1, \eta_1)$, lie in $\{QU\}$. The tangents at Q, Q' form with the lines TS, $T'S'$ a parallelogram of area 4Δ and this contains a point $P(x, y) = (\xi, \eta)$, or $(X, Y) = (\Xi, H)$, of L other than O. We may suppose that P does not lie in $\mathfrak{S}$, and P cannot lie in $\{TR\}$ since cases III, IV are incompatible. Hence there is no loss of generality in supposing that P lies in a region abutting on and below RS produced, outside $\mathfrak{S}$ and to the left of the tangent at Q. Clearly, P lies above the line $y = 0$ from the concavity of the boundary $f(x, y) = 1$ with $y < 0$.

We have proved in §4 that the tangent at Q makes an intercept of length less than $2RS$ measured from R on RS produced. We now prove that P, P_1 form a basis of L, that is, $\Xi H_1 - H \Xi_1 = -1$.

For, since $0 < H \leqq 1$, $1 \leqq \Xi < 2$, $0 \leqq \Xi_1 < 1$, $-1 < H_1 < 0$,

$$\left| \Xi H_1 - H \Xi_1 \right| < 2 + 1;$$

and, since the left-hand side is a nonzero integer,

$$\Xi H_1 - H \Xi_1 = -1 \quad \text{or} \quad -2.$$

We can exclude the -2, for when this holds $P_{11}(\Xi_{11}, H_{11}) = (P - P_1)/2$ is by Lemma 11 a point of L, and is also an inner point of $\mathfrak{S}$; for clearly

$$0 < \Xi_{11} < 1, \qquad 0 < H_{11} < 1,$$

and so the result follows.

Now consider the point $P_2 = P_1 - P$. On using x, y; X, Y coordinates respectively, we see that, since $x_1 > 1$, $x < 1$, $X_1 < 1$, $X > 1$, then

$$\xi_2 > 0, \qquad \Xi_2 < 0.$$

Hence P_2 lies in the angle whose vertex is at O and whose sides pass through the points $(x, y) = (0, -1)$, and R'. Now P_1, P_2 both lie on the line $x\eta - y\xi = \Delta$ since P, P_1 form a basis of L, and P_2 is the point of this line which is also the point of L nearest P_1 and to the left of P_1. This line meets each of the curves $f(x, y) = \pm 1$ in only one real point. For the intersections are given by the equations

$$f(x, (x\eta - \Delta)/\xi) \pm 1 = 0,$$

the left-hand sides of which are steadily increasing functions of x since $\xi > 0$,

$\eta > 0$. Hence any point on the line lying between the intersections is an inner point of $\mathfrak{S}$. I prove presently that the condition enunciated in the theorem means that the line has an intercept between the tangents at Q and R' greater than OP. Hence the intercept must contain at least one point of L. This point cannot be P_1 and so it can be taken as P_2. Now $\xi_2 > 0$, and since the tangent at R' cuts $f(x, y) = -1$ in a point to the left of $x = 0$, that is, the points of the tangent between this point and R' are inner points of $\mathfrak{S}$, then P_2 is an inner point of $\mathfrak{R}$.

Let the line $X\mathrm{H} - Y\Xi = 1$ meet the tangents at Q, R' at the points (X_Q, Y_Q), $(X_{R'}, Y_{R'})$. It is sufficient to show that

$$X_Q - X_{R'} > \Xi.$$

Write, as in §4,

$$\lambda = cf_a + df_b, \qquad \mu = af_c + bf_d,$$

so that $\lambda < 0$, $\mu < 0$. The tangent at Q is $\lambda Y + X - 1 = 0$ and meets the line $-\Xi Y + \mathrm{H}X - 1 = 0$ where

$$(\Xi + \lambda \mathrm{H})X_Q = \Xi + \lambda.$$

The tangent at R' is $Y + 1 + \mu X = 0$ and meets $-\Xi Y - 1 + \mathrm{H}X = 0$ where

$$(\mu \Xi + \mathrm{H})X_{R'} = 1 - \Xi.$$

Hence the condition $X_Q - X_{R'} > \Xi$ becomes

$$\frac{\Xi + \lambda}{\Xi + \lambda \mathrm{H}} - \frac{1 - \Xi}{\mu \Xi + \mathrm{H}} > \Xi.$$

Now $\Xi + \lambda \mathrm{H} > 0$ since P lies to the left of the tangent at Q. Next $\mu \Xi + \mathrm{H} > 0$ since P lies above the line through O parallel to the tangent at R'. For the line meets $x = 1$ below $y = 0$ and P lies above $y = 0$. Hence the condition becomes

$$\lambda \mu \Xi + \lambda \mathrm{H} + \mu \Xi^2 + \Xi \mathrm{H} - (\Xi - \Xi^2 + \lambda \mathrm{H}(1 - \Xi)) > \Xi(\lambda \mathrm{H} + \Xi)(\mu \Xi + \mathrm{H}),$$

or since $\Xi > 0$,

$$\lambda \mu + \mu \Xi + \mathrm{H} - 1 + \Xi + \lambda \mathrm{H} > (\lambda \mathrm{H} + \Xi)(\mu \Xi + \mathrm{H}),$$

or

$$\lambda \mu > (\lambda \mathrm{H} + \Xi)(\mu \Xi + \mathrm{H}) - (\lambda \mathrm{H} + \Xi) - (\mu \Xi + \mathrm{H}) + 1,$$

or

$$(\Xi + \lambda \mathrm{H} - 1)(\mathrm{H} + \mu \Xi - 1) < \lambda \mu.$$

Considered as an inequality in Ξ, H, this defines a region bounded by a

hyperbola whose asymptotes are the tangents at Q, R and which passes through the point S, $(\Xi, \mathrm{H}) = (1, 1)$. Now P lies to the left of the tangent at Q and below the tangent at R. Hence each of the left-hand factors of the inequality is negative and so the inequality means that P must lie between the hyperbola and its asymptotes $\Xi + \lambda \mathrm{H} - 1 = 0$, $\mathrm{H} + \mu \Xi - 1 = 0$, and is to the left of the first asymptote and below the second. But by the hypothesis, the curve $f(x, y) = 1$ and this hyperbola have no intersection between S and the point where the tangent at Q meets the curve $f(x, y) = 1$ again, that is, the arc of the hyperbola between S and the line $y = 0$ lies entirely in $\mathfrak{S}$. Hence the result.

We assume now that the first case has been disposed of, that is, that the condition of §5 is satisfied.

9. **The second case.** We have proved now that a point P of L must lie in the region $\{RT\}$, and we take P so that no point of L other than O or P lies in OP.

If, however, we had argued with the parallelogram $\overline{S}T\overline{S}'T'$, the image of $STS'T'$ in $x + y = 0$, we should again have had four cases to consider. The first three are dealt with exactly as before, and, by symmetry, the hypothesis is such as to rule out the possibility of P lying in the region $\{\overline{Q}'\overline{U}'\}$. Hence P lies in that part of $\{RT\}$ not included in $\{\overline{Q}'\overline{U}'\}$. Also by symmetry a point of L, say P_1, lies in that part of $\{\overline{R}\overline{T}\}$ not included in $\{Q'U'\}$.

We have another parallelogram of area 4Δ whose sides are the tangents at R, R' and the lines ST', $S'T$. Hence there is a point P_2 of L in a region exterior to the part of $\mathfrak{R}$ included between the points R, S, but included between the line QS produced and the tangent at R. By symmetry there is a point P_3 of L in the region which is the image of the preceding one in the line $x - y = 0$. I prove first that $P_2 = P_3$. Now the tangent at R cuts the line $x = 1$ at a point $(1, y)$ where $y = (1 - f_c)/f_d$, that is, from Lemma 6, $0 < y < 1$. Hence the points P_2, P_3 lie in the square $0 \leq x < 1$, $0 \leq y < 1$.

Define the point $P_4(x_4, y_4) = P_2 - P_3$. Then $|x_4| < 1$, $|y_4| < 1$. Hence *either* P_4 is in the second or fourth quadrant parts of the square $|x| < 1$, $|y| < 1$, and is then an inner point of $\mathfrak{R}$ and $\mathfrak{S}$, *or* P_4 lies in the first quadrant or third quadrant of the square.

Suppose first that P_4 lies in $0 \leq x < 1$, $0 \leq y < 1$. Then $x_2 \geq x_3$, $y_2 \geq y_3$. Hence $a + c \geq x_2 \geq x_3 \geq b + d$. Since $b + d > 1/2$ and $x_2 < 1$, then $2x_3 > x_2$ and so P_4 is distinct from P_3. Further $y_3 \geq b + d$, since P_2 is below the tangent at R, above RS and to the left of QS, and P_3 is to the left of the tangent at $\overline{R}'$ and to the right of $\overline{S}'\overline{R}'$. Hence

$$y_4 = y_2 - y_3 \leq 1 - b - d, \qquad x_4 \leq a + c - b - d.$$

Now $b + d > 1/2$ and so $y_4 < b + d$, $x_4 < a + c$, and so P_4 is an inner point of $\mathfrak{R}$ and so must be at O.

Suppose next that P_4 lies in the third quadrant. Then $x_3 \geq x_2$, $y_3 \geq y_2$ and so

$$a + c \geqq y_3 \geqq y_2 \geqq b + d, \qquad x_2 \geqq b + d,$$
$$- x_4 = x_3 - x_2 \leqq 1 - b - d,^{\cdot} \qquad - y_4 \leqq a + c - b - d,$$

and so $-P_4$ is an inner point of $\Re$ and $\mathfrak{S}$. Clearly P_4 is distinct from $-P_2$.

Hence P_2, P_3 coincide and so P_2 must lie in a small curvilinear triangle two of whose sides are QS produced and $\overline{Q}'\overline{S}'$ produced and whose curved side is the boundary of $\Re$ or $\mathfrak{S}$ included between the points $(x, y) = (a+c, b+d)$ and $(b+d, a+c)$.

There is a point P_1 of L in the region $\{\overline{RT}\}$ such that no point of L is an inner point of OP_1. Clearly P_1, P form a basis of L since the triangle OPP_1 can contain no points of L other than O, P, P_1. Next P_1, P_2 form a basis of L since the line P_1P_2 lies beneath the line joining $\overline{T}$ and $\overline{S}'$ and P lies above the line $\overline{TS}'$, and so the triangle OP_2P_1 contains no points of L other than O, P_2, P_1. Hence since P and P_2 are on the same side of OP_1, the line PP_2 is parallel to OP_1 and $P_2P = kOP_1$, where k is an integer greater than 0, and so $x_2 - x = - kx_1$. But, since the x, y coordinates of the lowest boundary point of $\{\overline{RT}\}$ are $-d$, $-c$, clearly $-x_1 > d$, and so

$$kd < x_2 - x < a + c - c + a = 2a.$$

But from §3, $a < 3/2, d > 1$, and so $k < 3$. Next $k \neq 2$, for if $P - P_2 = 2P_1$,

$$2y_1 = y - y_2, \qquad 2x_1 = x - x_2.$$

Hence $2y_1 < d - b - d - b$, that is, $y_1 < -b$. Also $y_1 > 0$, since $y > 1$ and $y_2 < 1$. Next $2x_1 > c - a - c - a$, that is, $x_1 > -a$. Also $x_1 < -1$, since P_1 is not in $\mathfrak{S}$. Hence P_1 is restricted to the small area external to $f(x, y) = -1$, above $y = 0$ and to the right of $x = -a$, that is, also bounded by the line drawn from $(-a, -b)$ perpendicular to the y axis. But P_1 is the midpoint of the line joining P and $-P_2$. But this is impossible since P and $-P_2$ are on one side of the line TS' and P_1 is on the other side. Hence $k = 1$ and $P_1 = P - P_2$. Using X, Y coordinates, $Y_1 = Y - Y_2 < 0$, since P is below and P_2 is above the line TRS. Now consider the point

$$P_5 = P + P_1.$$

Clearly $Y_5 = Y + Y_1 < Y < 1$ and so P_5 lies strictly below the line RT. By symmetry P_5 also lies strictly to the right of the line $\overline{RT}$. Hence P_5 is an inner point of $\mathfrak{S}$ since it is different from P and P_1.

10. The extremal problem arising in the first case. The extremal problem now reduces to finding conditions for the equations

$$f(x, y) = 1, \qquad (xf_a + yf_b - 1)(xf_c + yf_d - 1) - \lambda\mu = 0,$$

where

$$\lambda = (a + c)f_a + (b + d)f_b - 1 = cf_a + df_b \leqq 0,$$
$$\mu = (a + c)f_c + (b + d)f_d - 1 = af_c + bf_d < 0,$$

to have no solutions with $0 < y < b+d$.

On substituting $x = g(y)$ in the second equation, we have

$$0 = E(y) = 1 - (f_b + f_d)y + f_b f_d y^2 + [(f_a f_d + f_b f_c)y - f_a - f_c]g(y)$$
$$+ f_a f_c [g(y)]^2 - \lambda\mu,$$

and wish to find conditions that $E(y) < 0$ for $0 \leq y < b+d$, and we know that $E(b+d) = 0$. For this purpose, we shall use the following lemma.

LEMMA 13. *Suppose that a twice differentiable function $E(y)$ is defined for $0 \leq y \leq q$, that $E(0) < 0$, $E(q) = 0$, $E''(y) > 0$ for $0 \leq y \leq p < q$, $E'(y) > 0$ for $p \leq y \leq q$. Then $E(y) < 0$ for $0 \leq y < q$.*

Since $E'(y) > 0$ for $p \leq y \leq q$, $E(y)$ must increase from $y = p$ to $y = q$ and hence $E(y) < 0$ for $p \leq y < q$. The same holds for $0 \leq y < p$, for if not suppose that $E(y) \geq 0$ somewhere in this interval. Then by the mean value theorem there exists y_1 with $0 < y_1 < y$ for which $E'(y_1) > 0$, and also y_2 with $y < y_2 < p$ for which $E'(y_2) < 0$. This contradicts the hypothesis that $E'(y)$ is increasing for $0 \leq y \leq p$.

We shall apply this lemma later with $p = 1/2$, $q = b+d$.

It will be convenient to deal now with some inequalities involving the function $f(x, y)$. Some of these follow from the definition of $f(x, y)$, while some others will now be postulated as satisfied by $f(x, y)$.

Let (y_0, y_0) be the point on the curve $f(x, y) = 1$ or $x = g(y)$ for which $x = y$ so that $y_0 = g(y_0)$, $g'(y_0) = -1$. Then we prove

$$1 \geq yg'(y) + g(y) \geq 0$$

for $0 \leq y \leq y_0$. For $g(0) = 1$, $g'(0) = 0$, and

$$\frac{d}{dy}[yg'(y) + g(y)] = yg''(y) + 2g'(y) \leq 0$$

for $0 \leq y \leq y_0$, since $g'(y) \leq 0$, $g''(y) \leq 0$.

Next we *postulate* that

$$y + g(y)g'(y) \geq 0$$

for $0 \leq y \leq y_0$. This holds when $g(y) = (1 - y^n)^{1/n}$ and $y_0^n = 1/2$, since $g'(y) = -y^{n-1}(1-y^n)^{1/n-1}$, and so the inequality becomes

$$y\left[1 - \frac{y^{n-2}}{(1-y^n)^{1-2/n}}\right] \geq 0, \quad \text{that is,} \quad 1 - 2y^n \geq 0.$$

Now we *postulate* that, for $0 \leq y \leq 1/2$,

$$E_1(y) \equiv 1 + g(y)g''(y) \geq 0.$$

When $g(y) = (1 - y^n)^{1/n}$, $g''(y) = -(n-1)y^{n-2}(1-y^n)^{1/n-2}$,

$$E_1(y) = 1 - (n - 1)y^{n-2}/(1 - y^n)^{2-2/n},$$

and so $E_1(y) \geqq 0$ if $(1-y^n)^{2-2/n} \geqq (n-1)y^{n-2}$ for $0 \leqq y \leqq 1/2$.

Clearly it suffices if

$$(1 - 2^{-n})^{2-2/n} \geqq (n - 1)/2^{n-2},$$

or

$$(1 - 2^{-n})^2 \geqq (n - 1)/2^{n-2},$$

or

$$1 \geqq 1/2^{n-1} + (n - 1)/2^{n-2},$$

that is,

$$2^{n-1} \geqq 2n - 1.$$

This obviously holds when $n \geqq 4$.

Next we *postulate* that for $0 < y \leqq y_0$,

(1) $$\psi(y) = yg'(y) + g(y) - g'(y) - 1 > 0.$$

Since $\psi(0) = 0$, $\psi(y_0) = 0$, a sufficient condition when $g(y) = (1-y^n)^{1/n}$ is that

$$\psi'(y) = (y - 1)g''(y) + 2g'(y)$$

should have only one zero y_1 in $0 < y < y_0$. For $\psi'(y) > 0$ for small $y > 0$ from the convexity condition, and so $\psi(y)$ increases in $0 < y < y_1$ and decreases in $y_1 < y < y_0$.

We can now deal with the conditions for $E(y) < 0$ for $0 \leqq y \leqq b+d$. First

$$E(0) = (1 - f_a)(1 - f_c) - \lambda\mu < 0,$$

since $f_a > 1$, $0 < f_c < 1$, $\lambda \leqq 0$, $\mu < 0$. Next

$$E'(y) = - f_b - f_d + 2f_bf_dy + (f_af_d + f_bf_c)[yg'(y) + g(y)]$$
$$- (f_a + f_c)g'(y) + 2f_af_cg(y)g'(y).$$

One part of $E'(y)$ is

$$2f_bf_dy + 2f_af_cg(y)g'(y) = (2f_bf_d - 2f_af_c)y + 2f_af_c[y + g(y)g'(y)]$$
$$\geqq 2(f_bf_d - f_af_c)y.$$

Another part of $E'(y)$ is

$$(f_af_d + f_bf_c)(yg'(y) + g(y)) - (f_a + f_c)g'(y) - f_d.$$

Since $f_a > f_d > 1$, $f_b > 0$, $f_c > 0$, and $g'(y) < 0$, this part is greater than

$$f_d(yg'(y) + g(y) - g'(y) - 1) = f_d\psi(y) \geqq 0$$

for $0 \leqq y \leqq y_0$. Hence

$$E'(y) \geqq f_d\psi(y) + 2(f_bf_d - f_af_c)y - f_b > \psi(y) - f_b,$$

since $f_bf_d - f_af_c > 0$ from §4.

By hypothesis, $\psi(y)$ has only one turning point $y=y_1$ in the interval $0 \leqq y \leqq y_0$, and $\psi(0) = \psi(y_0) = 0$. Hence for $1/2 \leqq y \leqq b+d$,

$$\psi(y) \geqq \min \left[\psi(1/2), \psi(b+d) \right],$$

and so $E'(y) \geqq 0$ for $1/2 \leqq y \leqq b+d$ if

(2)
$$g(1/2) - g'(1/2)/2 - 1 - f_b \geqq 0,$$
$$g(b+d) + (b+d-1)g'(b+d) - 1 - f_b \geqq 0.$$

The second condition can be dispensed with if $b+d \leqq y_1$.

Finally

$$E''(y) = 2f_b f_d + (f_a f_d + f_b f_c)(yg''(y) + 2g'(y))$$
$$- (f_a + f_c)g''(y) + 2f_a f_c(g(y)g''(y) + g'(y)^2)$$
$$= M + N,$$

say, where

$$M = 2f_b f_d + 2f_a f_c(g(y)g''(y) + g'(y)^2),$$

and since from §4, $f_b f_d > f_a f_c > 0$,

$$M > 2f_a f_c(1 + g(y)g''(y) + g'(y)^2) > 0,$$

by hypothesis. Next

$$N = (f_a f_d + f_b f_c)(yg''(y) + 2g'(y)) - (f_a + f_c)g''(y).$$

Now $E''(y) > 0$ if $N \geqq 0$. Since $g''(y) < 0$ for $0 < y < 1$, $N \geqq 0$ if

$$(f_a f_d + f_b f_c)(-y - 2g'(y)/g''(y)) + f_a + f_c \geqq 0$$

for $0 < y < 1/2$.

We *postulate* that for $0 < y < 1/2$,

$$y + \frac{2g'(y)}{g''(y)} \leqq \frac{f_a + f_c}{f_a f_d + f_b f_c},$$

a condition easily investigated if the left-hand side steadily increases for $0 \leqq y \leqq 1/2$, for then it becomes

$$K = \frac{1}{2} + \frac{2g'(1/2)}{g''(1/2)} \leqq \frac{f_a + f_c}{f_a f_d + f_b f_c}.$$

This from Lemma 6 is certainly satisfied if

$$K(f_a^2 + f_b^2) \leqq 1.$$

11. **The case** $f(x, y) = x^n + y^n$, $n \geqq 3$. I discuss now the application to this case of the system of equations considered in §3. Now $g(y) = (1 - y^n)^{1/n}$, and on writing $b = -\beta$, the inequality (7) at the end of §3 becomes

$$E(\beta) \equiv [(1 + \beta)^n - 1]^{1/n} - 2(1 + \beta^n)^{1/n} + \{1 + [\beta - (1 + \beta^n)^{1/n}]^n\}^{1/n} > 0.$$

The result proved there was that if $E(\beta) > 0$ for $0 \leqq \beta_1 \leqq \beta < 1$, then the system of equations has exactly one solution $\beta = \beta_0$, with $0 < \beta_0 < \beta_1$.

By differentiation it is easily seen that

$$[(1 + \beta)^n - 1]^{1/n} - (1 + \beta^n)^{1/n}, \qquad \beta - (1 + \beta^n)^{1/n}$$

are increasing functions of β for $\beta > 0$.

We prove first that $E(\beta) > 0$ if $\beta_1 = 3/4$, that is, $3/4 \leqq \beta \leqq 1$. Clearly

$$E(\beta) > [(7/4)^n - 1]^{1/n} - [1 + (3/4)^n]^{1/n} - (1 + 1)^{1/n}$$
$$+ [1 - \{(1 + (3/4)^n)^{1/n} - 3/4\}^n]^{1/n}.$$

Now $[(7/4)^n - 1]^{1/n} = (7/4)[1 - (4/7)^n]^{1/n}$ and $-[1 + (3/4)^n]^{1/n}$ are obviously increasing functions of n for $n \geqq 1$, and $[1 + (3/4)^n]^{1/n} < 1 + (1/n)(3/4)^n$. Hence

$$E(\beta) \geqq [(7/4)^3 - 1]^{1/3} - [1 + 27/64]^{1/3} - 2^{1/3}$$
$$+ \{1 - [1/4 + (1/n)(3/4)^n]^n\}^{1/n}.$$

Since the last term steadily increases as n increases from $n = 1$,

$$E(\beta) \geqq (279/64)^{1/3} - (91/64)^{1/3} - 2^{1/3} + [1 - (25/64)^3]^{1/3}$$
$$> 1.633 - 1.125 - 1.260 + [1 - (.390625)^3]^{1/3}$$
$$> 1.633 - 2.385 + (.94039)^{1/3}$$
$$> - .8 + .9 > 0.$$

Suppose next that $1/2 \leqq \beta \leqq 3/4$. Then clearly

$$E(\beta) \geqq [(3/2)^3 - 1]^{1/3} - (1 + 1/2^3)^{1/3} - (1 + 27/64)^{1/3}$$
$$+ \{1 - [1/2 + (1/3)(1/2^3)]^3\}^{1/3}$$
$$> (19/8)^{1/3} - (9/8)^{1/3} - (91/64)^{1/3} + [1 - (13/24)^3]^{1/3}$$
$$> 1.334 - 1.041 - 1.125 + [1 - (.5542)^3]^{1/3}$$
$$> 1.334 - 2.166 + (.8410)^{1/3}$$
$$> - .832 + .943 > 0.$$

Suppose next that $2/5 \leqq \beta \leqq 1/2$. Then

$$E(\beta) > [(7/5)^3 - 1]^{1/3} - [1 + (2/5)^3]^{1/3} - (1 + 1/8)^{1/3}$$
$$+ \{1 - [1 - 2/5 + (1/3)(2/5)]^3\}^{1/3}$$
$$> (218/125)^{1/3} - (133/125)^{1/3} - (9/8)^{1/3} + [1 - (233/375)^3]^{1/3}$$
$$> 1.203 \cdots - 1.021 \cdots - 1.041 \cdots + .9 \cdots > 0.$$

Hence $b > -2/5$. It is also easily shown that if $n \geq 4$, then $b > -.3$. For if $.3 \leq \beta \leq 2/5$,

$$E(\beta) > [(1.3)^4 - 1]^{1/4} - 1 - (1/4)(.3)^4 - 1 - (1/4)(.4)^4$$
$$+ \{1 - [.7 + (1/4)(.3)^4]^4\}^{1/4}$$
$$> 1.166 - 2 - .0021 - .0064 + .93 > 0.$$

We have to show now that the geometrical configuration is of the type discussed, that is,

$$ca^{n-1} + db^{n-1} \leq 0.$$

When $n = 3$, it is easily seen on adding the last two equations in (1) of §3 that this holds with the equality sign. I suggest the following conjecture.

CONJECTURE. *For $n = p/q > 3$, p and q odd positive integers,*

$$ca^{n-1} + db^{n-1} < 0.$$

We can prove this directly for $n = 5, 7$. When $n = 5$, from

$$(a + c)^5 + (b + d)^5 = 1, \qquad (-a + c)^5 + (-b + d)^5 = 1,$$

by addition, since $c^5 + d^5 = 1$,

$$a^4c + b^4d + 2(a^2c^3 + b^2d^3) = 0.$$

Hence we have to show that

$$E = a^2c^3 + b^2d^3 > 0.$$

Now

$$E > a^2b^3 + b^2,$$
$$E/b^2 > 1 + b(1 - b^5)^{2/5} > 1 - (1/2)(1 + 1/32)^{2/5} > 0.$$

When $n = 7$, from the addition of similar equations,

$$a^6c + b^6d + 5(a^4c^3 + b^4d^3) + 3(a^2c^5 + b^2d^5) = 0.$$

Hence we have to show that

$$E' = 5(a^4c^3 + b^4d^3) + 3(a^2c^5 + b^2d^5) > 0.$$

Now

$$E' = 5b^4d^3 + 3a^2c^5 + 3b^2d^5 + 5a^4c^3 > b^4(5 + 3a^2b) + b^2(3 + 5a^4b).$$

Also

$$5 + 3a^2b = 5 + 3(1 - b^7)^{2/7}b > 5 - (3/2)(1 + 1/128)^{2/7} > 0,$$
$$3 + 5a^4b = 3 + 5(1 - b^7)^{4/7}b > 3 - (5/2)(1 + 1/128)^{4/7} > 0.$$

Hence $E' > 0$.

We now prove the conjecture for $n \geq 4$ by using Lemma 5. Writing $-y$ for y, we have to show that

$$[(1 + y)^n - 1]^{1/n} - [1 - (1 - y)^n]^{1/n} - \frac{2y^{n-1}}{(1 + y^n)^{(n-1)/n}} > 0,$$

where y runs through the range of values now possible from the end of §11 for $-b$, that is, $(y+1)^n - y^n > 2$, $0 < y < 2/5$. It suffices to show that

$$E = [(1 + y)^n - 1]^{1/n} - [1 - (1 - y)^n]^{1/n} - 2y^{n-1} > 0.$$

Suppose first that $ny \leq 1$. Since $(1 - 1/n)^n$ increases with n, and $(y+1)^n > 2$,

$$E \geq 1 - [1 - (1 - 1/n)^n]^{1/n} - 2/n^{n-1} > 1 - [1 - (3/4)^4]^{1/n} - 2/n^{n-1}$$
$$> (1/n)(3/4)^4 - 2/n^{n-1} > 0$$

since $(3/4)^4 > 2/4^2$.

Suppose next that $ny > 1$, $y < 1/3$. Then

$$E > [(1 + 1/n)^n - 1]^{1/n} - [1 - (2/3)^n]^{1/n} - 2/3^{n-1}$$
$$> (e^{1-1/2n} - 1)^{1/n} - 1 - 2/3^{n-1} \quad (\text{since} \quad \log(1 + 1/n) > 1/n - 1/2n^2)$$
$$> (e - 1 - e/2n)^{1/n} - 1 - 2/3^{n-1} \quad (\text{since} \quad e^{-x} > 1 - x \text{ for } x > 0)$$
$$> (1.71 - 1.36/n)^{1/n} - 1 - 2/3^{n-1}$$
$$> (1.37)^{1/n} - 1 - 2/3^{n-1} \quad (\text{since} \quad n \geq 4)$$
$$> (1/n) \log 1.37 - 2/3^{n-1}$$
$$> .31/n - 2/3^{n-1} > 0.$$

Suppose finally that $1/3 \leq y < 2/5$. Then

$$E > [(4/3)^n - 1]^{1/n} - 1 + (1/n)(3/5)^n - 2(2/5)^{n-1}.$$

The first term increases with n, as is clear on writing it as $(4/3)(1 - (3/4)^n)^{1/n}$. Hence

$$E > (175/81)^{1/4} - 1 - 2(2/5)^3 > 1.21 - 1 - .128 > 0.$$

We come now to the conditions assumed to hold for the extremal problem. In §10 we defined

$$\psi(y) = yg'(y) + g(y) - g'(y) - 1,$$

which, since $g(y) = (1 - y^n)^{1/n}$, becomes

$$\psi(y) = \frac{-y^{n-1}(y - 1)}{(1 - y^n)^{(n-1)/n}} + (1 - y^n)^{1/n} - 1 = \frac{1 - 2y^n + y^{n-1}}{(1 - y^n)^{(n-1)/n}} - 1.$$

We have to verify first that $\psi(y)$ has only one turning point in $0 < y < 1/2^{1/n}$. Since

$$\psi'(y) = (y - 1)g''(y) + 2g'(y),$$

$\psi'(y) = 0$ gives

$$(y - 1)g''(y)/g'(y) + 2 = 0.$$

Since

$$g'(y) = - y^{n-1}/(1 - y^n)^{(n-1)/n}, \quad \frac{g''(y)}{g'(y)} = \frac{n-1}{y} + \frac{(n-1)y^{n-1}}{1 - y^n} = \frac{n-1}{y(1 - y^n)}.$$

Hence

$$(n - 1)(1 - y)/y(1 - y^n) = 2,$$

or

$$y\frac{1 - y^n}{1 - y} = \frac{n-1}{2}.$$

The left-hand side is monotone increasing in y for $0 < y < 1$. The equation will have a root with $y < (1/2)^{1/n}$ if

$$(1/2)^{1/n} > (n - 1)(1 - (1/2)^{1/n}),$$

that is, if

$$(2)^{1/n} < 1 + 1/(n - 1).$$

This is obvious since $(2)^{1/n} < 1 + 1/n < 1 + 1/(n-1)$.

We have now to verify the two inequalities (2) of §11

$$\psi(1/2) \geqq f_b, \qquad \psi(b + d) \geqq f_b.$$

The first inequality becomes

$$(1 - 1/2^n)^{-(n-1)/n} - 1 > b^{n-1}.$$

For this it suffices if

$$b^{n-1} < \{(n - 1)/n\}\{1/2^n\},$$

or, since $b > -.3$, if

$$(.6)^{n-1} < 1/2 - 1/2n.$$

This is clearly true for $n > 4$ since $.216 + .125 < .5$.

The second inequality follows from one due to Dr. Mahler, namely

$$(b + d)^n < (3 - (5)^{1/2})/2$$

if $n \geqq 4$. I give later, in Lemma 14, a proof very different from his. The condition $\psi(b+d) \geqq f_b$ becomes

$$E_1 = \frac{1 - 2(b+d)^n + (b+d)^{n-1}}{[1 - (b+d)^n]^{(n-1)/n}} - b^{n-1} - 1 > 0.$$

Write $k = ((5)^{1/2} - 1)/2 = .618 \cdots$, and so $k^2 + k = 1$. Since

$$(b+d)^n < k^2 < 1/2^{1/n},$$

and $\psi(y)$ has only one turning point, it suffices to assume that we have to prove that

$$E = \frac{1 - 3 + (5)^{1/2} + k^{(2n-2)/n}}{k^{(n-1)/n}} - b^{n-1} - 1 > 0.$$

Since $(-2 + (5)^{1/2})/k = (-4 + 2(5)^{1/2})/((5)^{1/2} - 1) = (3 - (5)^{1/2})/2 = k^2$,

$$E = k^{2+1/n} + k^{1-1/n} - b^{n-1} - 1.$$

Now $e^{-t} \geq 1 - t$ for all t. On putting $\log k = -nt$, we have

$$k^{1/n} > 1 + (1/n) \log k,$$

and also

$$k^{-1/n} > 1 - (1/n) \log k.$$

Hence

$$
\begin{aligned}
E &\geq k^2(1 + (1/n) \log k) + k(1 - (1/n) \log k) - 1 - b^{n-1} \\
&> (1/n)(k^2 - k) \log k - b^{n-1} \\
&> (1/n)(.236) \log 1.618 \cdots - b^{n-1} \\
&> .113/n - b^{n-1} \\
&> .113/n - (.3)^{n-1} \qquad\qquad\qquad\qquad\qquad\text{(if } n \geq 4), \\
nE &> .113 - n(.3)^{n-1},
\end{aligned}
$$

and since $n(.3)^{n-1}$ decreases steadily for $n > 1/\log(10/3)$,

$$nE > .113 - 4(.027) = .113 - .108 > 0.$$

We now prove Mahler's inequalities:

LEMMA 14. *If* $n \geq 3$, $(a+c)^n > (1/2)(-1 + (5)^{1/2})$, $(b+d)^n < (1/2)(3 - (5)^{1/2})$.

It suffices to prove the first since $(a+c)^n + (b+d)^n = 1$.
Put $(a+c)^n = P$, $(b+d)^n = Q$. Then

$$P + Q = 1, \qquad \frac{(c^2 - a^2)^n}{P} + \frac{(d^2 - b^2)^n}{Q} = 1.$$

Write $r = (c^2 - a^2)^n$, $s = (d^2 - b^2)^n$. Since $d < a$ and $c^2 < b^2$, $r < 0$, then $r + s < 0$, $s - r > 0$.

We prove first that $s-r \leq 2$, that is, on putting $b=-\beta$, $c=-\gamma$, we have

$$E = [(1 + \beta^n)^{2/n} - \gamma^2]^n + [(1 + \gamma^n)^{2/n} - \beta^2]^n \leq 2 \quad \text{for} \quad 0 \leq \beta, \gamma \leq 1/2.$$

Since E is symmetrical in β, γ, it obviously suffices to show that E as a function of β decreases steadily for $0 \leq \beta \leq 1/2$. Now

$$\frac{1}{2n\beta} \frac{\partial E}{\partial \beta} = \beta^{n-2}(1 + \beta^n)^{2/n-1}[(1 + \beta^n)^{2/n} - \gamma^2]^{n-1} - (d^2 - \beta^2)^{n-1}$$

$$< \beta^{n-2}(1 + \beta^n) - (1 - \beta^2)^{n-1} \qquad \text{(since } d>1\text{)}$$

$$< \frac{1}{2^{n-2}}\left(1 + \frac{1}{2^n}\right) - \left(\frac{3}{4}\right)^{n-1}.$$

Hence

$$\frac{2^{n-1}}{2n\beta} \frac{\partial E}{\partial \beta} < 2 + \frac{1}{2^{n-1}} - \left(\frac{3}{2}\right)^{n-1} \leq 0 \qquad \text{(for } n \geq 3\text{)}$$

since $1/2^{n-1}$ decreases and $(3/2)^{n-1}$ increases with n. Since $b<0$, clearly $s-r<2$. Next from

$$r/P + s/(1 - P) = 1, \quad r(1 - P) + sP = P - P^2,$$

and so

$$P^2 + P(s - r - 1) + r = 0.$$

Hence as $r<0$ and $P>0$, P is the greater root of this quadratic and so

$$2P = 1 - s + r + [(s - r - 1)^2 - 4r]^{1/2}$$

$$= 1 - s + r + [(s - r)^2 - 2r - 2s + 1]^{1/2}$$

$$> 1 - s + r + [(s - r)^2 + 1]^{1/2}$$

since $r+s<0$. The right-hand side is a decreasing function of $s-r$ since $d(1-x+(x^2+1)^{1/2})/dx = -1+x/(x^2+1)^{1/2}<0$, and as $s-r<2$,

$$2P > -1 + 5^{1/2}.$$

We come now to the final condition of §10, namely,

$$[1/2 + 2g'(1/2)/g''(1/2)](f_a^2 + f_b^2) \leq 1.$$

But

$$g(y) = (1 - y^n)^{1/n}, \quad g'(y) = -y^{n-1}(1 - y)^{1/n-1},$$

$$g''(y) = -(n - 1)y^{n-2}(1 - y)^{1/n-2}, \quad g'(y)/g''(y) = y(1 - y^n)/(n - 1).$$

Hence the condition above is certainly satisfied if

$$E = (1/2 + 1/(n - 1))(a^{2n-2} + b^{2n-2}) - 1 \leq 0.$$

Now

$$E < \frac{n+1}{2(n-1)} a^{2n} - 1 + \frac{n+1}{2(n-1)} b^{2n-2}$$

$$< \frac{n+1}{2(n-1)}\left(1+\frac{1}{2^n}\right)^2 - 1 + \frac{n+1}{2(n-1)}\left(\frac{1}{2}\right)^{2n-2} \quad \left(\text{since} \quad |b| < \frac{1}{2}\right)$$

$$< \frac{5}{6}\left(1+\frac{1}{16}\right)^2 - 1 + \frac{5}{6}\frac{1}{64} \qquad \text{(if } n > 4)$$

$$< .95 - 1 + .02 < 0.$$

This finishes the application to $f(x, y) = x^n + y^n$. Hence the main theorem in §5 holds for $n > 4$. I conjecture that it holds for $n \geq 3$. It would not be difficult to reduce 4 to a somewhat smaller number.

REFERENCES

1. H. DAVENPORT
     The minimum of a binary cubic form, J. London Math. Soc. vol. 18 (1943) pp. 168–176.
2. K. MAHLER
     Lattice points in two dimensional star domains. To appear in Proc. London Math. Soc.
     Note on lattice points in star domains, J. London Math. Soc. vol. 17 (1942) pp. 130–133.
     On lattice points in an infinite star domain, J. London Math. Soc. vol. 18 (1943) pp. 233–238.
     On lattice points in the domain $|xy| \leq 1$, $|x+y| \leq 5^{1/2}$, *and applications to asymptotic formulae
          in lattice point theory* (2 parts), Proc. Cambridge Philos. Soc. vol. 40 (1944) pp. 107–120.
3. L. J. MORDELL
     On numbers represented by binary cubic forms, Proc. London Math. Soc. (2) vol. 48 (1943)
          pp. 198–228.
     The minimum of a binary cubic form, I, II, J. London Math. Soc. vol. 18 (1943) pp. 201–210;
          210–218.
     On the geometry of numbers in some non-convex regions, Proc. London Math. Soc. (2) vol. 48
          (1945) pp. 339–390.

MANCHESTER UNIVERSITY,
     MANCHESTER, ENGLAND