J. Austral. Math. Soc. **24** (Series A) (1977), 266–285.

SIMULTANEOUS DIOPHANTINE APPROXIMATION

J. M. MACK

To George Szekeres on his 65th birthday

(Received 9 June 1975; revised 6 September 1976)

Communicated by Jane Pitman

Abstract

Using a method suggested by E. S. Barnes, it is shown that the simultaneous inequalities $r(p-\alpha r)^2 < c$, $r(q-\beta r)^2 < c$ have an infinity of integral solutions p, q, r (with $r > 0$), for arbitrary irrationals α and β, provided that $c > 1/2.6394$. This improves an earlier result of Davenport, who shows that the same conclusion holds if $c > 1/46^{1/4} = 1/2.6043\cdots$.

1. Introduction

Let α, β be irrational numbers. Davenport (1952) has shown that the simultaneous inequalities

$$(1) \qquad r(p-\alpha r)^2 < c, \qquad r(q-\beta r)^2 < c$$

have an infinity of integral solutions p, q, r (with $r > 0$) provided that

$$c > \frac{1}{46^{1/4}} = \frac{1}{2.6043\cdots}.$$

In the opposite direction, Cassels (1955) has shown that if $c < 2/7 = 1/3.5$, there exist α and β for which the inequalities (1) have only a finite number of solutions. Both results are obtained by using the fact that if C is the infimum of constants c such that the inequalities (1) admit an infinity of solutions for all choices of α and β, then $C = 1/\Delta$, where Δ is the lattice constant of the three-dimensional star body defined by the inequality $|z|\max(x^2, y^2) < 1$. This result, which is a particular case of a general theorem of Davenport (1955), is first mentioned in Cassels (1955), although the analogous result for a closely related problem had been obtained much earlier by Davenport and Mahler (1946). The value of C is unknown.

https://doi.org/10.1017/S1446788700020292 Published online by Cambridge University Press

Davenport obtained his estimate by using a technique previously employed by Mullender (1950). Essentially, a method of Mordell is used to find a lower bound for Δ by reducing the problem to a two-dimensional problem in the geometry of numbers, and by using the known lattice constant of the star-body $|z|(x^2+y^2)<1$. In the present paper, a different reduction will be used to show that *the inequalities* (1) *have an infinity of solutions provided that*

$$c>\frac{1}{2.6394}.$$

This method again reduces the problem of finding a lower bound for Δ to a two-dimensional problem; however the regions which arise are bounded, whereas the corresponding regions studied by Davenport and by Mullender are unbounded star-domains.

In §2, we describe Barnes' method, and we analyse the two-dimensional regions obtained from it in §§3–5. The constructions used to obtain the requisite lower bound for Δ are described briefly in §6, suppressing most of the routine calculations. Some final comments are made in §7.

2. Reduction to a two-dimensional problem

Let L, M, N be real linear forms of determinant 1 in the variables u, v, w, and let

$$\mu=\inf(|N|\max(L^2,M^2)),$$

where the infimum is taken over integral u, v, w, not all zero. Suppose $\mu>0$, and assume first that μ is attained. Thus there exist integers u_0, v_0, w_0, not all zero, such that (in an obvious notation)

$$\mu=|N_0|\max(L_0^2,M_0^2).$$

We suppose $L_0^2\leqq M_0^2$, so that $|L_0/M_0|=t$, where $0\leqq t\leqq 1$. Define new linear forms

$$X=L/M_0,\qquad Y=M/M_0,\qquad Z=N/N_0,$$

of determinant $1/\mu$. We may assume that the forms X, Y, Z take the values t, 1, 1 respectively at u_0, v_0, w_0. Now consider the three-dimensional lattice Λ given by

$$x=X-tZ,\qquad y=Y-Z,\qquad z=Z,$$

for integral u, v, w. Λ has determinant $1/\mu$, and the definition of μ implies that

(2) $$|z|\max((x+tz)^2,\ (y+z)^2)\geqq 1$$

https://doi.org/10.1017/S1446788700020292 Published online by Cambridge University Press

for all points of Λ other than the origin. Further, since $(0,0,1)$ is a primitive point of Λ, there is a basis of Λ with respect to which its points are given by $x = x_1 v + x_2 w$, $y = y_1 v + y_2 w$, $z = u + z_1 v + z_2 w$, with integral u, v, w. Let $\mathscr{L}$ be the two-dimensional lattice obtained by considering only the x and y coordinates. $\mathscr{L}$ has determinant $1/\mu$, and (2) implies that $\mathscr{L}$ has the property that for any lattice point (x, y) of $\mathscr{L}$ other than the origin, there is a real number κ such that

$$\inf(|u+\kappa|\max((x+t(u+\kappa))^2, \quad (y+u+\kappa)^2) \geqq 1,$$

where the infimum is taken over all integers u. Therefore, if we define $S(t)$ $(0 \leqq t \leqq 1)$ to consist of those points (x, y) such that for any real κ there is a λ congruent to κ modulo 1 for which

$$|\lambda|\max((x+t\lambda)^2, \quad (y+\lambda)^2) < 1,$$

it follows that $\mathscr{L}$ is admissible for $S(t)$. If as t varies from 0 to 1 the lattice constant, $\Delta(t)$ say, of $S(t)$ is at least Δ_0, then $\Delta_0 \leqq 1/\mu$. In the case where μ is not attained, we may obtain the same result by applying the above argument for a sequence of values μ_n tending to the infimum μ, and observing that the corresponding sequence of sets thus obtained satisfies the conditions of a theorem of Mahler (c.f. Cassels 1959, p. 140) which asserts that the sequence of lattice constants then tends to the lattice constant of $S(t)$. Rewriting the inequality above as $\mu \leqq 1/\Delta_0$, and recalling the definition of μ, we see that this inequality implies that the lattice constant of the star-body $|z|\max(x^2, y^2) < 1/\Delta_0$ is at least 1, whence the lattice constant Δ defined in §1 satisfies $\Delta \geqq \Delta_0$. Hence our result will be established if we can show that $\Delta_0 \geqq 2.6394$.

3. Determination of $S(t)$

The regions $S(t)$ may in theory be determined as follows. First, determine the region

$$\boldsymbol{R}(1) = \{(\lambda, y) : |\lambda|(y+\lambda)^2 < 1\},$$

so that for each y, the set

$$I(y) = \{\lambda : (\lambda, y) \in \boldsymbol{R}(1)\}$$

is known. Then, for each t in $0 \leqq t \leqq 1$, determine

$$R(t) = \{(\lambda, x) : |\lambda|(x+t\lambda)^2 < 1\},$$

and, for fixed y and t, study the set

$$I(y) \cap \{\lambda : (\lambda, x) \in R(t)\} \tag{3}$$

as a function of x. Those x for which this set covers the reals mod 1 yield points $(x, y) \in S(t)$. As y varies, we obtain the whole of $S(t)$ in this way, and as t varies, we obtain all the regions $S(t)$. In order to see how to implement this programme, we need first to examine the shape of $R(t)$. If we define functions $L_t = L_t(\lambda)$, $U_t = U_t(\lambda)$ by

$$(4) \qquad L_t(\lambda) = -\lambda t - |\lambda|^{-1/2}, \qquad U_t(\lambda) = -\lambda t + |\lambda|^{-1/2}$$

then

$$R(t) = \{(\lambda, x): L_t < x < U_t\} \qquad (0 \leqq t \leqq 1),$$

while

$$\boldsymbol{R}(1) = \{(\lambda, y): L_1 < y < U_1\}.$$

From these descriptions, we see that $R(t)$ is symmetric in the origin and contains the lines $\lambda = 0$ and $x + t\lambda = 0$. Except for $t = 0$ (when there is no turning point T), $R(t)$ (shown in Figure 2) has a shape of the form shown in Figure 1, which depicts $\boldsymbol{R}(1)$ for $y \geqq 0$. Figure 1 shows that the set $I(y)$ defined above is either an interval of a disjoint union of two intervals. Further, the concavity of the boundary curves of $R(t)$, together with the fact that for each λ, $(\lambda \times \mathbf{R}) \cap R(t)$ is an interval, suggests that the set (3) at first increases with x and then decreases as x increases. This result would imply that $S(t) \cap (\mathbf{R} \times \{y\})$ is either empty or else of the form $I \times \{y\}$, where I is an interval. That this latter result is true follows most readily from the observation that, for each y and t considered,

$$I(y) \times \{ty\} \subset R(t),$$

since it then follows from the shape of $R(t)$ that (3) is monotone decreasing with respect to $|x - ty|$. We collect together in a Lemma this result and two other immediate results which will be used frequently in the sequel.

LEMMA 1. (i) *Let I be a closed interval of length* 1 *such that $I \subset I(y)$ and $I \times \{x\} \subset R(t)$. Then $(x, y) \in S(t)$.*

(ii) *Let I be a closed interval of length* 2 *and $J \subset I$ an open interval of length at most* 1 *such that $I \backslash J \subset I(y)$ and $(I \backslash J) \times \{x\} \subset R(t)$. Then $(x, y) \in S(t)$.*

(iii) *If $(x', y') \in S(t)$, then $\{x : (x, y') \in S(t)\}$ is an open interval.*

Since $S(t)$ is symmetric in the origin, the practical problem reduces to the following: determine those $y \geqq 0$ for which $I(y)$ covers the reals mod 1, and then for each such y, and each t, determine numbers $m(t, y)$ and $M(t, y)$ such that $\{x : m(t, y) < x < M(t, y)\} \times \{y\} \subset S(t)$ and for which $M(t, y) - m(t, y)$ is

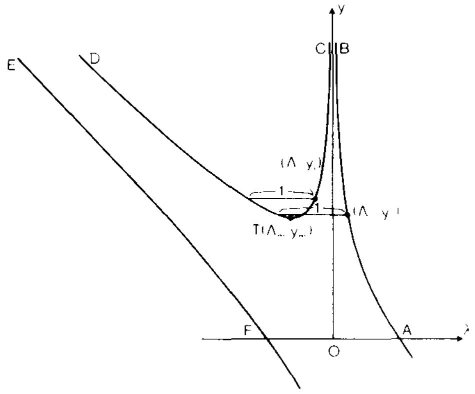

Fig. 1. The region $\boldsymbol{R}(1)$

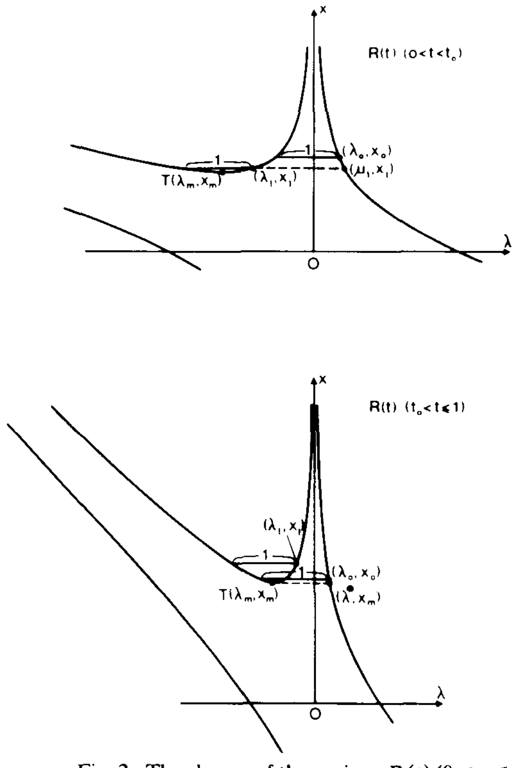

Fig. 2. The shapes of the regions $R(t)$ $(0 < t \leqq 1)$

https://doi.org/10.1017/S1446788700020292 Published online by Cambridge University Press

as large as possible. We begin by identifying certain points on the boundary of $R(t)$, and values of x, y, and t which will be significant later.

The turning points T on the boundary of $\boldsymbol{R}(1)$, and on the boundary of $R(t)$ for $t>0$ (see Figs. 1, 2) strongly influence the shape of $S(t)$. The coordinates (λ_m, x_m) of T on $R(t)$ are given by

$$(5) \qquad \lambda_m = -(2t)^{-2/3}, \qquad x_m = 3(t/4)^{1/3},$$

and we denote the coordinates of T on $\boldsymbol{R}(1)$ by (Λ_m, y_m). Thus

$$(6) \qquad \Lambda_m = -2^{-2/3} = -0.6300, \qquad y_m = 3(1/4)^{1/3} = 1.8899.$$

[Approximate values are rounded off to four decimal places wherever they appear.] The value $\lambda^* > 0$ such that (λ^*, x_m) is on the boundary of $R(t)$ (see Figure 2) may be found by putting $\lambda^* = c^2 t^{-2/3}$, where $c>0$ then satisfies the equation

$$c^3 + 3(1/4)^{1/3}c - 1 = 0.$$

Solving this, we find

$$(7) \qquad \lambda^* = 0.2238t^{-2/3}.$$

The difference $\lambda^* - \lambda_m$ decreases as t increases, and the equation $\lambda^* - \lambda_m = 1$ holds for $t = t^*$ say, where

$$(8) \qquad t^* = 0.7889.$$

For any t, we shall need the values $\lambda_0 > 0$ and $x_0 > 0$ such that (λ_0, x_0) and $(\lambda_0 - 1, x_0)$ both lie on the boundary of $R(t)$ (see Figure 2). Note that the latter point is to the right of T for $t < t^*$, and to the left of T for $t > t^*$. Solving the relevant equations, we find

$$(9) \qquad \lambda_0 = (1-(1-4\beta^2)^{1/2})/2, \quad \text{where} \quad \beta = ((1+t^2)^{1/2}-1)/t^2,$$

and

$$(10) \qquad x_0 = U_t(\lambda_0),$$

where U_t is given by (4). It follows from this that x_0 is an increasing function of t. Again, we shall denote the coordinates of the point corresponding to (λ_0, x_0) on $\boldsymbol{R}(1)$ by (Λ_0, y_0), so that

$$(11) \qquad \Lambda_0 = \lambda_0(1) = 0.2200, \qquad y_0 = x_0(1) = 1.9123.$$

Another pair of interest is (λ_1, x_1), where (λ_1, x_1) and $(\lambda_1 - 1, x_1)$ both lie on the boundary of $R(t)$ and $\lambda_m < \lambda_1 < 0$ (see Figure 2 again). λ_1 is in fact a root of the equation

$$(12) \qquad t^4\lambda^4 + (4t^2 - 2t^4)\lambda^3 + (t^4 - 6t^2)\lambda^2 + 2t^2\lambda + 1 = 0,$$

and

(13) $$x_1 = U_t(\lambda_1).$$

When $t = 1$, equation (12) factorises, with two of its roots given by

(14) $$\lambda = (-1 - 2\sqrt{2} \pm \sqrt{(5 + 4\sqrt{2})})/2 = -0.2820, -3.5465.$$

Hence, for $\boldsymbol{R}(1)$, the coordinates of the corresponding point (Λ_1, y_1) are given by

(15) $$\Lambda_1 = -0.2820, \qquad y_1 = 2.1652.$$

The remaining value for λ in (14) is also of interest, since it corresponds to a value of y such that (λ, y) and $(\lambda - 1, y)$ lie on the arcs TD, EF respectively of $\boldsymbol{R}(1)$ (see Figure 1). Examination of $\boldsymbol{R}(1)$ shows that for values of y just greater than this value, the set $I(y)$ cannot cover the reals mod 1. Hence $S(t)$ has no points (x, y) with y just greater than this particular value of y, which we denote by $y_{\max}$. We have

(16) $$y_{\max} = U_1(-3.5465) = 4.0800.$$

We remark here that there do exist values $y > y_{\max}$ yielding points $(x, y) \in S(t)$, at least for some values of t. For example, $(4.525, 4.525) \in S(1)$. This shows that $S(t)$ is not always connected, and hence is not always a two-dimensional star-body. However, as we can see no way of using that part of $S(t)$ lying outside $|y| \leqq y_{\max}$, we discuss it no further.

We define one more value of λ related to $x_1(t)$. The value $\mu_1 = \mu_1(t)$ is defined by the conditions

(17) $$\mu_1 > 0, \qquad x_1(t) = U_t(\mu_1).$$

The difference $x_0(t) - x_1(t)$, initially positive, decreases steadily as t increases, vanishing for a unique value $t = t_0$. t_0 is thus defined by the equation

(18) $$x_0(t_0) = x_1(t_0),$$

and the value of t_0 is approximately 0.46.

4. Estimation of $m(t, y)$

We now obtain values for $m(t, y)$, thus giving a lower bound for the interval of x-values such that $(x, y) \in S(t)$ for each t and y. We shall discuss later whether or not these values can be improved. We consider separately certain ranges of values for y.

(i) $y = 0$. We have, by Lemma 1(i) with $I = [-\lambda_0, 1-\lambda_0]$,

(19) $$m(t,0) = -x_0(t) \qquad (0 \leqq t \leqq t^*),$$

where $x_0 = x_0(t)$ is given by (10). For $t^* \leqq t \leqq 1$, the turning point, $-T$ say, on the boundary of $R(t)$ in $x < 0$, limits the values of λ that can be used, and consequently

(20) $$m(t,0) = -x_m = -3(t/4)^{1/3} \qquad (t^* \leqq t \leqq 1).$$

(ii) $0 < y \leqq y_m$. Define $y^* = y^*(t)$ by

$$y^* = U_1(1-\lambda_0(t)) \qquad (0 \leqq t \leqq t^*),$$
$$y^* = U_1(1-\lambda^*(t)) \qquad (t^* \leqq t \leqq 1),$$

where λ_0 and λ^* are defined by (9) and (7) respectively. Then we may put

(21) $$m(t,y) = m(t,0) \qquad (0 < y < y^*),$$

since Lemma 1(i) applies exactly as for (i) above, with the same sets I.

For $y^* \leqq y \leqq y_m$, the largest value of $\lambda \in \overline{I(y)}$ is less than $1-\lambda_0(t)$ (or $1-\lambda^*(t)$), and is that $\lambda > 0$ such that $y = U_1(\lambda)$. For this value of λ, we may take $I = [\lambda - 1, \lambda]$ and hence obtain

(22) $$m(t,y) = L_t(\lambda - 1) \qquad (y^* \leqq y = U_1(\lambda) \leqq y_m).$$

We note that the curve $(m(t,y), y)$, for $y^* \leqq y \leqq y_m$, is a smooth concave curve.

(iii) $y = y_m$. As $y \to y_m -$, $m(t,y)$ approaches the value given by (22) with $\lambda = \lambda^*(1)$. As $y \to y_m +$, since any set used in Lemma 1 must cover Λ_m mod 1, we see that by choosing an interval I in Lemma 1(i) with its right-hand endpoint on the arc TD in Figure 1, we may choose $m(t,y)$ for $y > y_m$ so that

$$m(t, y_m +) = L_t(\Lambda_m - 1).$$

We note that the difference $m(t, y_m +) - m(t, y_m -)$ is positive for $t = 0$ and increases with t.

(iv) $y > y_m$. As remarked above, by choosing λ such that $(\lambda, U_1(\lambda))$ lies on TD in Figure 1, we may put

$$m(t,y) = L_t(\lambda - 1) \qquad (y_m < y \leqq y_{\max}).$$

By doing this, we ignore those $\lambda \in I(y)$ with $\lambda > \Lambda_m$, and for some $y > y_m$, we can use these λ to improve our estimate. Recall that y_0 and y_1 are given by (11) and (15) respectively. We now put, as above,

(23) $$m(t,y) = L_t(\lambda - 1) \qquad (y_m < y \leqq y_0 \quad \text{and} \quad y_1 < y \leqq y_{\max}),$$

while for the remaining values of y, we appeal to Lemma 1(ii) with $J = [\lambda' - 2, \lambda']$, and where (λ', y) lies on the arc AB in Figure 1. It follows from this that we may put

$$m(t, y) = L_t(\lambda' - 2) \qquad (y_0 < y < y_1). \tag{24}$$

In each of the intervals where (23) or (24) are used, the curves $(m(t, y), y)$ are again concave. Further, (24) implies that there is a jump in the value of $m(t, y)$ as y passes through y_1.

Finally, we remark that examination of the estimates obtained above shows that no improvement in $m(t, y)$ is possible, except perhaps by defining values of y analogous to y_0 and y_1, but where the difference in λ-values is 2 instead of 1, and studying $m(t, y)$ between these two values of y. Any improvement obtained would be small, and would have no effect on the argument in §6. Thus $(m(t, y), y)$ effectively gives part of the boundary of $S(t)$.

5. Estimation of $M(t, y)$

The estimation of an upper bound for the interval of x-values such that $(x, y) \in S(t)$ is further complicated by the fact that the arguments necessarily involve that part of the boundary of $R(t)$ upon which the turning point T lies. As t increases, the boundary near T has more and more effect on our estimates. For this reason, and also because the estimates we give are valid over intervals of y-values which vary with t, it is better to describe the estimates for ranges of values of t than for ranges of values of y. Briefly, as t increases, the effect of T is to influence the estimation first for y near $y_{\max}$, and then for smaller values of y.

(i) $0 \leqq t \leqq 0.0516$. Since $\lambda_m(0.0516) = -4.5465$, and since from (16) $y_{\max} = L_1(-4.5465)$, we see that T on $R(t)$ has no influence on $S(t)$ for t in the range being considered.

Since $S(t)$ is symmetric in the origin, we should be able to choose

$$M(t, 0) = -m(t, 0) = x_0(t), \tag{25}$$

and this is possible, since we may take $I = [\lambda_0(t) - 1, \lambda_0(t)]$ in Lemma 1(i), where λ_0 and x_0 are given by (9) and (10) respectively. Further, since the same interval I is applicable for each value of y such that

$$y \leqq U_1(\lambda_0(t)), \tag{26}$$

we may put

$$M(t, y) = x_0(t) \qquad (0 \leqq y \leqq U_1(\lambda_0(t))). \tag{27}$$

https://doi.org/10.1017/S1446788700020292 Published online by Cambridge University Press

As y increases, the positive λ such that $y = U_1(\lambda)$ decreases, and this forces us to decrease our estimate. For this λ we may put

$$M(t, y) = U_t(\lambda - 1) \qquad (\lambda > 0,\ U_1(\lambda_0(t)) \leqq y < y_m), \tag{28}$$

by using $I = [\lambda - 1, \lambda]$ in Lemma 1(i). In particular,

$$M(t, y_m -) = U_t(\lambda^*(1) - 1) = 1.135 + 0.7762t. \tag{29}$$

For $y > y_m$, the fact that the relevant part of the boundary of $R(t)$ remains the curve to the right of T means that we should choose I to be as far to the right as possible in $I(y)$. Thus we may repeat the argument leading to (23) and (24) in §4 (iv). With λ and λ' precisely as defined there, we may select

$$M(t, y) = U_t(\lambda - 1) \qquad (y_m < y \leqq y_0 \quad \text{and} \quad y_1 < y \leqq y_{\max}), \tag{30}$$

$$M(t, y) = U_t(\lambda' - 2) \qquad (y_0 < y < y_1). \tag{31}$$

These imply that $M(t, y)$ decreases as y increases. Further, on examining the case $t = 0$ and comparing with the results of §4, we find $M(0, y) = -m(0, y)$ for $0 \leqq y \leqq y_{\max}$, which is desirable since $S(0)$ is symmetric in the y-axis.

(ii) $0.0516 \leqq t \leqq 0.075$. For t in this range,

$$-4.5465 \leqq \lambda_m(t) \leqq -3.5465,$$

and consequently the estimates for $M(t, y)$ obtained in (i) remain valid, except when $\lambda - 1 < \lambda_m(t)$; for this case we use

$$M(t, y) = x_m(t) \qquad (U_1(\lambda_m(t) + 1) \leqq y \leqq y_{\max}). \tag{32}$$

This is so because Lemma 1(i) applies with I as before (i.e., as used in obtaining (23) and (30)) to $x = x_m(t)$. The estimates for $M(t, y)$ are thus (27) and then (28) for $0 \leqq y < y_m$, (30) for $y_m < y < y_0$, (31) for $y_0 < y < y_1$, (30) for $y_1 < y < U_1(\lambda_m(t) + 1)$, and (32) for $U_1(\lambda_m(t) + 1) \leqq y \leqq y_{\max}$.

(iii) $0.075 \leqq t \leqq 0.267$. In this range, T on $R(t)$ has moved sufficiently far to the right for us to be able to use the curve EF in Figure 1 to estimate $M(t, y)$ for some values of y such that $y \geqq y_m$, while still using previous estimates for other y. To be precise, we shall use (32) only for $U_1(\lambda_m(t) + 1) \leqq y \leqq U_1(\lambda_m(t))$, and, for larger y, if (λ, y) lies on EF, we shall put

$$M(t, y) = U_t(\lambda + 1) \qquad (U_1(\lambda_m(t)) < y \leqq y_{\max}). \tag{33}$$

We may do this, since Lemma 1(i) applies with $I = [\lambda, \lambda + 1]$.

We note that for $t = 0.24$ approximately,

$$U_1(\lambda_m(t) + 1) = y_m,$$

while for $t = 0.267$ approximately,

https://doi.org/10.1017/S1446788700020292 Published online by Cambridge University Press

$$U_1(\lambda_m(t)) = y_m.$$

Hence, if $t \geqq 0.24$, (32) may be used only for $y_m < y \leqq U_1(\lambda_m(t))$, while for $t > 0.267$, (32) is no longer applicable. As for (31), since the argument justifying it is inapplicable as soon as $\lambda_m(t)$ lies in the interval $J = [\lambda' - 2, \lambda']$ used to obtain (24), we prefer to dispense with it.

The estimates for $M(t, y)$ are therefore (27) and (28) for $0 \leqq y < y_m$, (30) for $y_m < y < U_1(\lambda_m(t) + 1)$ (and so only for $0.075 \leqq t \leqq 0.24$), (32) for $U_1(\lambda_m(t) + 1) \leqq y \leqq U_1(\lambda_m(t))$ if $0.075 \leqq t \leqq 0.24$, (32) for $y_m < y \leqq U_1(\lambda_m(t))$ if $0.24 \leqq t \leqq 0.267$, and (33) for the remaining values of y.

As a result of these estimates, we see that $M(t, y)$ increases with y for $y > U_1(\lambda_m(t))$, since (33) is an increasing function of y. We may further show that when $M(t, y)$ is defined by (33), the curve $(M(t, y), y)$ is smooth and convex. Further, by (30),

$$M(t, y_m +) = U_t(\Lambda_m - 1) = 0.7833 + 1.63t \qquad (0 \leqq t \leqq 0.24), \tag{34}$$

while

$$M(t, y_m +) = x_m(t) \qquad (0.24 \leqq t \leqq 0.267), \tag{35}$$

by (32), while from (33) and our remark above,

$$M(t, y) \geqq x_m \qquad (y > y_m) \tag{36}$$

for $0.0516 \leqq t \leqq 0.267$, and (36) is trivially true for $0 \leqq t \leqq 0.0516$.

(iv) $0.267 \leqq t \leqq t_0$. Recall that t_0 is defined by (18), so that in the present range of values of t, $x_0(t) \geqq x_1(t)$.

For $y > y_m$, we shall define $M(t, y)$ by (33), so that (36) holds, and in fact

$$M(t, y_m +) = U_t(-1.5198) = 0.811 + 1.5198t, \tag{37}$$

because $y_m = L_1(-4^{2/3}) = L_1(-2.5198)$.

For $0 < y < y_m$, we use (27) and (28) to define $M(t, y)$, until t reaches the value where $M(t, y_m -) = x_1(t)$. From (29) and (13), this occurs when $\lambda_1(t) = -0.7762$, and so, from (12), when $t = 0.385$ approximately. For $0.385 < t \leqq t_0$, the use of (28) is restricted to the range

$$U_1(\lambda_0(t)) \leqq y \leqq U_1(\lambda_1(t) + 1),$$

because for any larger value of y less than y_m, we may put

$$M(t, y) = x_1(t), \tag{38}$$

by using Lemma 1(ii) with $J = [\lambda - 2, \lambda]$, where $(\lambda, U_1(\lambda))$ lies on the arc AB in Figure 1.

Note that $M(t_0, y) = M(t_0, 0)$ for $0 \leqq y < y_m$, by the definition of t_0, and that $M(t, y_m -)$ is a lower bound for $M(t, y)$ for $0.267 \leqq t \leqq t_0$ and $0 \leqq y < y_m$.

(v) $t_0 \leqq t \leqq t^*$. For $t > t_0$, $x_0(t) < x_1(t)$. We therefore use (27) only for $0 \leqq y \leqq L_1(\lambda_0(t) - 2)$, after which we may use the arc EF in Figure 1. If $(\lambda - 2, y(= L_1(\lambda - 2)))$ lies on EF, then we may put

$$M(t, y) = U_t(\lambda), \tag{39}$$

valid for

$$L_1(\lambda_0(t) - 2) \leqq y \leqq L_1(\mu_1(t) - 2), \tag{40}$$

where, by (17), $(\mu_1(t), x_1(t))$ lies on the boundary of $R(t)$. This estimate is justified by using $J = [\lambda - 2, \lambda]$ in Lemma 1(ii).

For the remaining values of y less than y_m, we may use (39) with $\lambda = \mu_1(t)$, giving $M(t, y) = x_1(t)$ and thus agreeing with (38). However, (33) can be used to give a better estimate for $M(t, y)$ for some values of y less than y_m if, with λ as used above in (39), $I = [\lambda - 2, \lambda - 1]$ satisfies $I \times \{x\} \subset R(t)$ for some $x > x_1(t)$. By (37) and (38), we see that

$$M(t, y_m +) = M(t, y_m -)$$

when

$$U_t(-1.5198) = x_1(t) = U_t(\lambda_1(t) - 1),$$

and so when $\lambda_1(t) = -0.5198$, which occurs for $t = 0.575$ approximately. For $t > 0.575$, we may use (33) for all y such that $y \geqq L_1(\lambda_1(t) - 2)$.

Summing up, we have the following results for the present range of values of t. $M(t, y)$ is given by (27) for $0 \leqq y \leqq L_1(\lambda_0(t) - 2)$, and by (39) for those y specified in (40). For $t_0 \leqq t < 0.575$, $M(t, y)$ is given by (38) for $L_1(\mu_1(t) - 2) < y < y_m$, and by (33) for $y > y_m$. For $0.575 \leqq t \leqq t^*$, $M(t, y)$ is given by (38) for

$$L_1(\mu_1(t) - 2) < y < L_1(\lambda_1(t) - 2),$$

and by (33) for $y \geqq L_1(\lambda_1(t) - 2)$.

It follows from the above that $M(t, y)$ is an increasing function of y for $t \geqq 0.575$.

(vi) $t^* < t \leqq 1$. An argument similar to that used in obtaining (20) for $m(t, 0)$ when $t > t^*$ shows that we must put $M(t, 0) = x_m(t)$, and we must then have

$$M(t, y) = x_m(t) \qquad (0 \leqq y \leqq L_1(\lambda_m(t) - 1), \tag{41}$$

for until y reaches the upper value given in (41), we cannot use the arc EF in Figure 1 to improve $M(t, y)$. Since $x_0(t) > x_m(t)$, we put

$$M(t,y)=U_t(\lambda-1) \qquad (L_1(\lambda_m(t)-1<y\leqq L_1(\lambda_0(t)-2),$$

where $(\lambda-2,y)$ is on EF, and so, for the given range of y, $(\lambda-1, M(t,y))$ is on the boundary of $R(t)$ between T and $(\lambda_0(t)-1, x_0(t))$.

For larger values of y, we use (39) for those y given by (40), then (38) for $L_1(\mu_1(t)-2)<y<L_1(\lambda_1(t)-2)$, and finally (33) for $y\geqq L_1(\lambda_1(t)-2)$.

As a result of these estimates, we see that $M(t,y)$ remains an increasing function of y (as remarked at the end of (v) for $t\geqq 0.575$), and further, that a comparison of the estimates for $m(1,y)$ given in §4 with the estimates for $M(1,y)$ given above shows that they are symmetric about $y=x$, i.e., $M(1,m(1,y))=y$. We note that $S(1)$ is also symmetric about $y=x$.

This completes the task of finding estimates for $M(t,y)$ for $0\leqq t\leqq 1$ and $0\leqq y\leqq y_{\max}$. Examination of the arguments used shows that there are two places where (slight) improvements may be possible. First, as described at the end of §4, by investigating x-values analogous to $x_0(t)$ and $x_1(t)$, but where the appropriate values of λ differ by 2 instead of 1, $M(t,y)$ could perhaps be improved for some t and some y. Second, by examining more closely the set $I(y)$ for y near y_m, and its relation to those λ such that $(\lambda, x_1(t))\in R(t)$, one can improve (33) (and so (37)) for y very close to but just greater than y_m, for $0.385<t<0.575$. However neither of these improvements would alter the result obtained in §6.

6. Calculation of a lower bound for $\Delta(t)$

The results of sections 4 and 5 give values for $m(t,y)$ and $M(t,y)$ such that for each t $(0\leqq t\leqq 1)$ and each y $(0\leqq y\leqq y_{\max})$,

$$\{(x,y): m(t,y)<x<M(t,y)\}\subset S(t). \tag{42}$$

By the symmetry of $S(t)$ in the origin, we may put, for $0\geqq y\geqq -y_{\max}$,

$$m(t,y)=-M(t,-y), \qquad M(t,y)=-m(t,-y),$$

and then (42) holds for $|y|\leqq y_{\max}$. We shall use (42) to construct for each t a convex symmetric parallelogram or hexagon inscribed in $S(t)$. A lower bound for the area of these inscribed figures leads immediately by Minkowski's convex body theorem to a lower bound for the lattice constant $\Delta(t)$ of $S(t)$. Before embarking on the construction, we remark that the principal difficulty in obtaining a good estimate for $\Delta(t)$ using the above method occurs for t near 0.9. For other values of t, we have nevertheless tried to obtain reasonably good estimates for $\Delta(t)$, even though these can have no effect on the final result.

The constructions will be given explicitly for $y\geqq 0$, since they can be extended by symmetry to $y < 0$. Figures 3, 4, 5 and 6 show the constructions used for $t = 0$, 0.4, 0.5 and 1, respectively, inside the relevant part of the corresponding regions $S(t)$. The regions drawn are of course based on the left hand side of (42), but we have previously remarked that except for some small intervals of values of y, the values for $m(t, y)$ and $M(t, y)$ obtained in sections 4 and 5 cannot be improved. It is also clear that any component of $S(t)$ lying outside $|y| \leqq y_{max}$ has no effect on the present calculations.

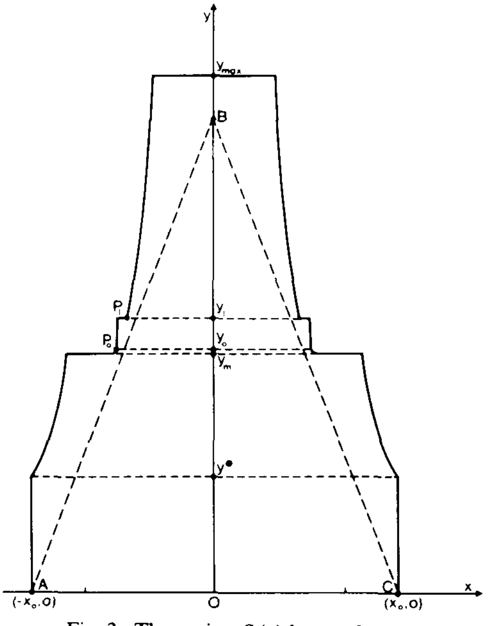

Fig. 3. The region $S(t)$ for $t = 0$.

Three points which lie on or close to the boundary of $S(t)$ and which are of use in the subsequent constructions will now be identified and labelled.

The point $(m(t, y_0), y_0)$, where y_0 is given by (11) and $m(t, y_0)$ by (23) with $\lambda = \Lambda_0 - 1 = -0.7800$, will be approximated by the interior point

$$P_0 = (1.78t - 0.749, 1.912).$$

The point $(m(t, y_1+), y_1)$, where y_1 is given by (15) and $m(t, y_1+)$ by (23) with $\lambda = \Lambda_1 - 1 = -1.2820$, will be approximated by the interior point

$$P_1 = (2.282t - 0.661, 2.165).$$

For $t \geqq 0.267$, the point $(M(t, y_m+), y_m)$, where y_m is given by (6) and $M(t, y_m+)$ by (37), will be approximated by the interior point

$$M = (1.519t + 0.811, 1.889).$$

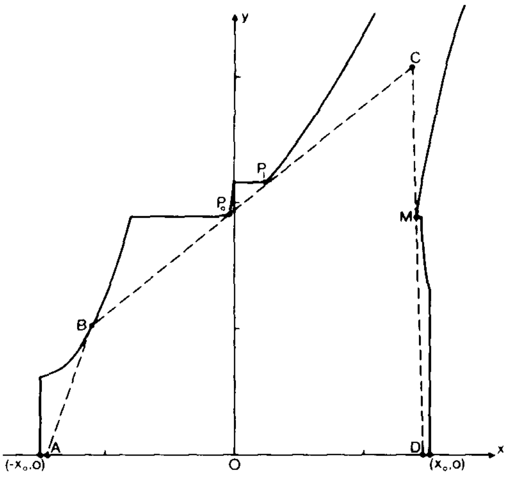

Fig. 4. The region $S(t)$ for $t = 0.4$.

The construction used varies with t, and therefore we consider certain subintervals of $0 \leqq t \leqq 1$ in turn.

(i) $0 \leqq t \leqq 0.29$. (See Figure 3.) We show the triangle with vertices $A(-\sqrt{2}, 0)$, $B(3.75t, 3.75)$ and $C(\sqrt{2}, 0)$ lies inside $S(t)$ for $0 \leqq t \leqq 0.29$. By the results of §3, and in particular of §4 (iv), the edge AB lies in $S(t)$ if it passes to the right of P_0 and P_1, and if its slope is less than the slope of the tangent to the curve $x = m(t, y)$ at y_1, where $m(t, y)$ is given by (23). This is so because if $y > y_1$, the relevant boundary of $S(t)$ lies above this tangent. The necessary simple calculations show these conditions are met for $t \leqq 0.296$. By the results on $M(t, y)$ obtained in §5, it suffices to study BC for $y \geqq y_m$.

For $0 \leqq t \leqq 0.145$, we note that if (30) is used to estimate $M(t, y)$ for $y \geqq y_m$, then the boundary of $S(t)$ lies on or to the right of the concave curve thus obtained. The tangent to this curve at $y = y_1$ is to the right of BC for $y \geqq y_m$.

For $0.145 \leqq t \leqq 0.267$, x_m is, by (36), a lower bound for $M(t, y)$ in $y \geqq y_m$, while for $0.267 \leqq t \leqq 0.29$, when $M(t, y)$ is defined by the increasing function (33), a lower bound for $M(t, y)$ in $y \geqq y_m$ is given by (37). In each case it is easily verified that BC lies inside $S(t)$.

Hence the triangle ABC, and so the convex symmetric parallelogram obtained by adding the reflection of B in the origin, lies entirely in $S(t)$ for $0 \leqq t \leqq 0.29$. The area of this parallelogram is $15/\sqrt{2} > 10.6$.

(ii) $0.29 \leqq t \leqq t'$, where t' is the value of t at which $M(t, y_m+)$ (given by (33)) equals $M(t, 0)$ (given by (27)). (t' lies between 0.487 and 0.488). (See Figure 4.) For $t \geqq 0.29$, the points $A(-1.465, 0)$ and $D(1.465, 0)$ lie in $S(t)$, by

§4(i) and the equation $M(t,0) = -m(t,0)$. For $y = U_1(0.45)$, $m(t,y)$ is given by (22) and we approximate $(m(t,y), y)$ by the interior point $B(0.55t - 1.348, 1.04)$. The point C is defined to be the intersection of the lines BP_0 and DM. The figure obtained from $ABCD$ and its reflection in the origin is easily verified to be a convex symmetric hexagon for the given range of t. It remains to check that it lies inside $S(t)$.

Since the curve $x = m(t,y)$ is concave near B, AB is inside $S(t)$. From the results of §4, BC lies to the right of the curve $x = m(t,y)$ if it passes to the right of the chord P_0P_1, and if its slope is less than that of the tangent to this curve from above at P_1, and greater than that of the tangent from below at P_0. Carrying out the calculations, we find these conditions are all met for $t \leqq 0.59$, hence for t in the range being considered.

For the side CD, we discuss $y \geqq y_m$ and $y < y_m$ separately. For $y \geqq y_m$, where $M(t,y)$, given by (33), increases with y and determines a convex curve, we note that for $t \leqq 0.431$ the slope of CD is negative and hence all is well, while for $0.431 \leqq t \leqq t'$, the slope of CD is greater than the slope at M of this convex curve, and again all is well.

For $0 \leqq y < y_m$, note that by §5 (iv), $M(t, y_m -)$ is a lower bound for $M(t,y)$ if $t \leqq t_0$, and by §5 (v), $M(t,0)$ is a lower bound for $M(t,y)$ if $t \geqq t_0$. Further, $M(t, y_m +) \leqq M(t, y_m -)$ and $M(t, y_m +) \leqq M(t,0)$ for $t \leqq t'$. Hence CM lies inside $S(t)$ whenever $M(t, y_m -) \geqq 1.465$ (the coordinate of D), and so for $0.425 \leqq t \leqq t'$. For $0.29 \leqq t \leqq 0.425$, CM has negative slope and meets the line $x = M(t, y_m -)$ in a point, W say, with ordinate less than 1. From (9) and (27), it follows that $M(t,0)$ is a lower bound for $M(t,y)$ for $0 \leqq y \leqq 1$ and $t \geqq 0.29$, and so CW lies in $S(t)$. WM clearly lies in $S(t)$. Hence CM lies in $S(t)$ and consequently so does CD.

The area of $ABCD$ is the sum of the area of the triangle ABD, which is independent of t, and the area, A say, of the triangle BCD. The coordinates of C, and hence the value of A, can be given explicitly in terms of t. When this is done, the derivative $A'(t)$ is negative for $0.29 \leqq t \leqq 0.5$, hence a lower bound for A is given by $A(0.5) = 3.8716$. Since the area of ABD is 1.5236, we find that the resulting convex symmetric hexagon has area greater than 10.79 for $0.29 \leqq t \leqq t'$.

(iii) $t' \leqq t \leqq 0.69$. (See Figure 5.) In this range of t, we construct an inscribed hexagon as follows. Put $A(-x_0(t), 0)$ and $E(x_0(t), 0)$ on the x-axis. The tangent to the curve $x = m(t,y)$ at the point $N((t/2) - \sqrt{2}, \sqrt{2} - (1/2))$ meets the perpendicular to the x-axis through A at the point B on the boundary of $S(t)$, and it meets the chord P_0P_1 at the interior point C of $S(t)$. The chord P_0P_1 meets the perpendicular to the x-axis through E at the interior point D of $S(t)$. $ABCDE$, together with its reflection in the origin,

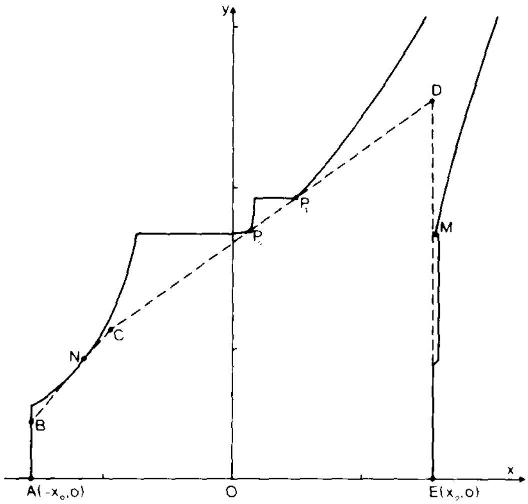

Fig. 5. The region $S(t)$ for $t = 0.5$.

forms a convex symmetric hexagon, which is readily verified to lie within $S(t)$ — recall that $M(t, y_m +) \geqq M(t, 0)$ for $t \geqq t'$. The area $A(t)$ of the polygon $ABCDE$ can be calculated explicitly in terms of t, but it is rather complicated. We can find a lower bound for $A(t)$ by the following method. Fix the vertices A and E at their positions A', E' say corresponding to $t = \tau_i$, and allow N, P_0, P_1 (and hence C) to vary with t for $t \geqq \tau_i$. Since $x_0(t)$ increases with t, the area $A_i(t)$ say of $A'B'CD'E'$ (the resulting polygon) is a lower bound for $A(t)$ for $t \geqq \tau_i$, and $A_i(t)$ can be shown to decrease as t increases. We choose $\tau_1, \cdots, \tau_7$ equal respectively to 0.48, 0.57, 0.62, 0.65, 0.67, 0.68 and 0.69, and then find the smallest value of $A_i(\tau_{i+1})$ to be $A_6(\tau_7) = 5.2839$. (Calculation of $A(t)$ shows that its least value is $A(\tau_7) = 5.3031$). Consequently we conclude that the corresponding inscribed hexagon has area at least 10.5678 for $t' \leqq t \leqq 0.69$.

(iv) $0.69 \leqq t \leqq 1$. (See Figure 6.) We have remarked before that values of t near 0.9 present the greatest difficulty. We shall use one construction for $0.69 \leqq t \leqq 0.9$, and another for $0.91 \leqq t \leqq 1$, and discuss briefly the modifications necessary to cover the range $0.9 \leqq t \leqq 0.91$. It is convenient to begin with $t = 1$. Figure 16 shows $S(1)$. P is the point (x, y_m) where, by (23), $x = L_1(\Lambda_m - 1)$.

The points Q, Q_0 and Q_1 are the reflections of P, P_0, P_1 in the line $y = x$. N is the point $((1/2) - \sqrt{2}, \sqrt{2} - (1/2))$, and the tangent to the boundary of $S(1)$ at N is perpendicular to ON. The perpendicular bisector of ON meets the boundary of $S(1)$ at a point P' lying between P and P_0. The hexagon obtained by taking the tangents at N and at P' and their reflections in the

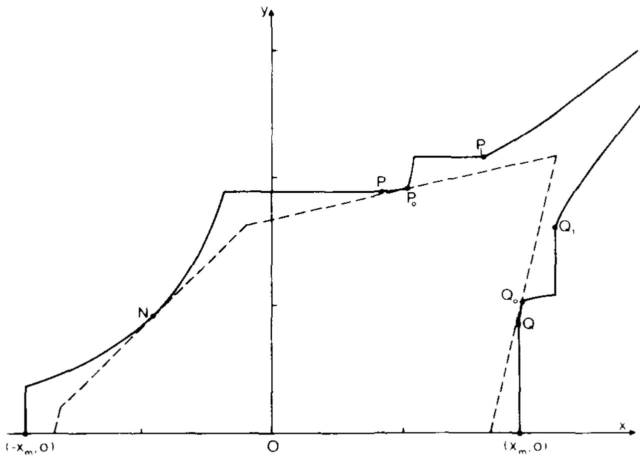

Fig. 6. The region $S(t)$ for $t = 1$.

lines $y = \pm x$ is in fact the optimal inscribed convex hexagon for $S(1)$. So that we may generalise to other t, we approximate it by replacing the tangent at P' by the tangent from below at P_0, and regarding its reflection in $y = x$ as the tangent from below at Q_0. Denote the hexagon obtained from these three tangents (and their reflections in 0) by $H(t)$. (The coordinates of Q_0 for general t are given by (39) and (40) and are $(x_0(t), L_1(\lambda_0(t) - 2))$. The arc QQ_0 appears only for $t > t^* = 0.7889$. The coordinates of Q_1 are, by (33) and §5 (vi), $(x_1(t), L_1(\lambda_1(t) - 2))$.) The area, $A(t)$ say, of $H(t)$ can be given as a function of t, but is too complicated to be of use. If the coordinates of Q_0 are held fixed at their value for $t = \tau$, but N and P_0 allowed to vary, then the resulting hexagon $H_\tau(t)$ has area $A_\tau(t)$ which decreases as t increases and which is a lower bound for $A(t)$ $(t \geqq \tau)$. Explicit computer calculation of $A(t)$, and then of $A_\tau(t)$ for suitably spaced τ, shows that $A(t)$ increases with t and that a lower bound for $A(t)$ in $0.91 \leqq t \leqq 1$ is 10.5573. Again there is no difficulty in showing $H(t)$ to be properly inscribed in $S(t)$.

For $0.69 \leqq t \leqq 0.9$ we use the hexagon $H_1(t)$ formed by the tangent at N, the chords P_0P_1 and Q_0Q_1, and the images of these three lines in the origin. $H_1(t)$ is inscribed in $S(t)$, and its area is again too complicated to discuss explicitly, although it can again be approximated arbitrarily closely by introducing a further parameter τ fixing Q_0 and Q_1. Calculations show that the area of $H_1(t)$ is a decreasing function of t and is not less than 10.5612 (its value at $t = 0.9$) for $0.69 \leqq t \leqq 0.9$.

There remains the interval $0.9 \leqq t \leqq 0.91$. An investigation of the lattice generated by P_0 and Q_0 shows that it has a point T_0 on the boundary of $S(t)$ near N for $t = 0.9073$. For this value of t, it is possible to modify both $H(t)$

and $H_1(t)$ so that T_0, P_0 and Q_0 are the midpoints of their respective sides. This is done by replacing the tangents at P_0 and Q_0 by tac-lines. The resulting construction yields the following procedure for dealing with the outstanding values of t. We replace the point N, which is obtained by choosing $\lambda = 1/2$ in (22), by the adjacent point T, obtained by choosing $\lambda = 0.5047$ in (22). If $H(t)$ is modified by replacing the tangent at N by the tangent at T, the tangent at P_0 by a tac-line at P_0 of slope 0.269, and the tangent at Q_0 by a tac-line at Q_0 of slope 6.66, the resulting hexagon has area greater than 10.5572 for $0.9073 \leqq t \leqq 0.91$. If $H_1(t)$ is modified by replacing the tangent at N by the tangent at T and the chord P_0P_1 by a tac-line at P_0 of slope 0.457, the resulting hexagon has area greater than 10.5572 for $0.9 \leqq t \leqq 0.9073$. In each case the modified hexagons are inscribed in $S(t)$, and consequently we have a lower bound of 10.5572 for the area of a convex symmetric inscribed hexagon, in the range $0.9 \leqq t \leqq 0.91$.

Collecting together the estimates obtained in (i)–(iv), we see that 10.5572 is a lower bound for the areas of the convex symmetric parallelograms or hexagons inscribed in $S(t)$ $(0 \leqq t \leqq 1)$, hence, by an application of Minkowski's convex body theorem,

$$\Delta(t) \geqq 10.5572/4 = 2.6394,$$

which was required to prove the result stated in §1.

7. Conclusion

We have remarked before that $S(t)$ has a component outside the range $|y| \leqq y_{max}$, at least for $t = 1$. Until one has some idea of the critical lattices for the component of $S(t)$ studied above, it is hard to see how to improve upon the present result by using all of $S(t)$. By examining the construction used to obtain good hexagons for t near 0.9073, it is clear that the "spike" in $S(t)$ (the part between y_1 and y_{max}) must be used more effectively if critical lattices are to be found. In fact, by modifying that construction so that P_0 is moved left, a lattice of determinant 2.88 can be obtained which is admissible for $S(1)$. This implies that no method based on the argument of §2 can close the gap which separates the present upper and lower estimates for the simultaneous approximation constant C. There is at present no reason to suppose any particular value for C, so any improvement on the bounds obtained above for $\Delta(t)$ would be of interest, especially for large t. The exhaustive computer investigation carried out in order to find good hexagons for t near 0.9 precludes the possibility of any improvement resulting from different choices of inscribed convex symmetric regions for t near 0.9. There are other general methods for finding lattice constants (e.g., the method of Mordell described in

Cassels (1959), §III.6) or bounds for them, but applying them to $S(t)$ will not be easy. A further difficulty is provided by the fact that for t near 1, the shape of the boundary of $S(t)$ between the points P_0 and P_1 (see Figure 4) implies that the component of $S(t)$ containing the origin is not a star-body.

Acknowledgement

I am indebted to Eric Barnes for suggesting the method used in this paper, and to George Szekeres for much advice and encouragement. This paper is based on work, done some years ago at the University of Sydney, which formed part of the author's doctoral thesis (Mack (1971)). I am grateful to the referee for suggesting ways of rendering intelligible the original version of §§3–5, and for other helpful comments which have been incorporated into the paper, and to June Jeffrey for the carefully drawn diagrams.

References

J. W. S. Cassels (1955), 'Simultaneous diophantine approximation', *J. London Math. Soc.* **30**, 119–121.

J. W. S. Cassels (1959), *An introduction to the geometry of numbers* (Springer, 1959).

H. Davenport (1952), 'Simultaneous diophantine approximation', *Proc. London Math. Soc.* (3) **2**, 406–416.

H. Davenport (1955), 'On a theorem of Furtwängler', *J. London Math. Soc.* **30**, 186–195.

H. Davenport and K. Mahler (1946), 'Simultaneous diophantine approximation', *Duke Math. J.* **13**, 105–111.

J. M. Mack (1971), *Simultaneous diophantine approximation*, Ph.D. thesis, University of Sydney.

P. Mullender (1950), 'Simultaneous approximation', *Annals of Mathematics* **52**, 417–426.

Department of Pure Mathematics,
University of Sydney,
N.S.W. 2006,
Australia.