A PROBLEM IN DIOPHANTINE APPROXIMATION

By H. T. J. NORTON.

[Received and read June 8th, 1916.—Revised October 31st, 1916.]

THE purpose of this note is to discuss a problem of which a particular case is this :—if ϑ is a real number, to find an upper limit to the smallest natural number n which is such that the difference between $n^2\vartheta$ and the nearest integer is less than some assigned positive number $1/\lambda$. Mr. Hardy and Mr. Littlewood* have shown that for all values of ϑ and λ an n with the property supposed exists, and that it is less than a number which is independent of ϑ, but they say† that they have not "succeeded in finding a definite function $\Phi(\lambda)$, the same for all ϑ's, such that

$$|\overline{n^2\vartheta}| \leqslant 1/\lambda \quad \text{for} \quad n \leqslant \Phi."$$

The result which is obtained here is that

$$n < \lambda^{2\lambda(1+\epsilon)},$$

where $\epsilon \to 0$ as $\lambda \to \infty$.

The most general theorems of this kind which are known have been proved by Minkowski.‡ One of them is a generalisation of the following proposition to manifolds of many dimensions. Suppose that rectilinear coordinates (X, Y) are taken in a plane and that H is a closed contour in the plane, which contains the origin in its interior, is symmetrical about the origin—that is to say, if (X, Y) lies on H so does $(-X, -Y)$—and is nowhere concave—that is, if P_1 and P_2 are points on H, then every point between P_1 and P_2 on the line joining them lies upon H or in its interior : then if the area contained by H is not less than 4,

---

* *Acta Mathematica*, Vol. 37, pp. 155–190. Extensions of the theorems proved by Mr. Hardy and Mr. Littlewood have been published lately by Mr. Fowler in this journal and Herr H. Weyl in the *Annalen*; they are, however, in a different direction and are not related to the subject of the present note.

† *l.c.*, p. 174. The notation $|\overline{n^2\vartheta}|$ means the difference, taken positively, between $n^2\vartheta$ and the nearest integer.

‡ *Geometrie der Zahlen* (1910), pp. 76, 77, 218, 219.

there are two points with integral coordinates, besides the origin, which lie within or upon the contour. For instance, if H is the parallelogram bounded by the lines

$$X = \pm q, \quad X\vartheta - Y = \pm \frac{1}{q},$$

the area contained by H is 4, the other conditions are fulfilled and there are, therefore, two integral points besides the origin upon or within the contour. If (ξ, η) is one of them

$$|\xi| \leqslant q, \quad |\xi\vartheta - \eta| \leqslant \frac{1}{q},$$

and hence follows the well-known theorem:—if ϑ and q are any real numbers, there exists a positive integer n which is less than or equal to q, and is such that the difference between $n\vartheta$ and the nearest integer is less than or equal to $1/q$. Minkowski's second theorem has the same character, but is more complicated.

From these two theorems most of the results known can be inferred very simply—as, for instance, Dirichlet's* theorem that if $\vartheta_1 \dots \vartheta_m$ are real numbers we can choose an n less than λ^m so that the difference between $n\vartheta_i$ and the nearest integer is less than $1/\lambda$, or Tchebycheff's theorem that, if ϑ is irrational, we can choose n so that the fractional part of $n\vartheta$ differs from any assigned number between 0 and 1 by as little as we like and that for an infinity of values of n the error is less than $3/n$, or Kronecker's generalisation of this, that if $\vartheta_1 \dots \vartheta_m$ are linearly independent irrationals we can choose n so that the fractional part of each of the numbers $n\vartheta_i$ differs by as little as we like from any arbitrarily assigned numbers between 0 and 1. The theorem of Mr. Hardy and Mr. Littlewood which is indicated above does not, however, appear to follow so easily. In this case, one is led to consider the closed contour H' which is bounded by the two lines

$$X = \pm \Phi$$

and the two parabolas

$$X^2\vartheta - Y = \pm \frac{1}{\lambda}.$$

The theorem is that if Φ exceeds a limit which depends only upon λ, there are two integral points, other than the origin, in the interior of H'. Since this contour is concave, it is not possible to apply Minkowski's theorem directly; and on the other hand it does not seem to be easy to alter the condition of non-concavity so as to bring contours like H' into the reach

* For elementary proofs of these theorems see Hardy and Littlewood, *l.c.*, pp. 158–165.

of the theorem, without destroying the whole argument. The proof which is given here is a modification of the argument used by Dirichlet. It leads to an upper limit which increases with great rapidity as λ tends to infinity, and I am very doubtful whether one increasing much more slowly could not be found. But a more exact determination of the limit seems to present serious difficulties.

For the sake of convenience and compactness of expression, I shall use a special notation. Let "$A \equiv B$" mean "A differs from B by an integer", and let η be a symbol which may denote different numbers on different occasions but only those which are not greater than 1. "$A \equiv \eta B$", then, asserts :—there exists an integer X such that

$$-|B| \leqslant A - X \leqslant |B|.$$

The general theorem* of Mr. Hardy and Mr. Littlewood, which is referred to above, may then be stated as follows :—

"If m, κ, and λ are three whole numbers, there exists a function $\Phi(m, \kappa, \lambda)$ such that, if $f_1(X) \ldots f_m(X)$ are any m polynomials in X with real coefficients and of degree not greater than κ, then the congruences

$$n \,.\, f_i(n) \equiv \eta/\lambda \quad (i = 1, 2, \ldots, m)$$

are satisfied by some positive integer $n \leqslant \Phi(m, \kappa, \lambda)$."

The particular case of this theorem, in which $\kappa = 0$, is the theorem of Dirichlet, which is mentioned above, and it is known† that we may put

$$\Phi(m, 0, \lambda) = \lambda^m - 1.$$

My object is to find an explicit formula for Φ which is valid for all values of κ. I take the case $\kappa = 1$; the generalisation by induction to an arbitrary value of κ does not present any difficulty.

Let λ be any positive integer and

$$a_i X + b_i \quad (i = 1, 2, \ldots, m)$$

any m linear functions of X with real coefficients. If

$$g_i(X) = X(a_i X + b_i),$$

---

* Cf. *l.c.*, Theorems 1.21 and 1.31.

† I do not know whether it has been determined exactly how small $\Phi(m, 0, \lambda)$ may be supposed to be. Mr. Hardy and Mr. Littlewood (*l.c.*, p. 159) give the value λ^m. The argument used by Minkowski (*l.c.*, p. 108) would give $\lambda^m - 1$. If Dirichlet's argument is used in the slightly modified form in which it is stated below, the value which is found is $\lambda^m - 2^m + 1$. But differences of this magnitude are too small to alter the formulæ which I obtain for higher values of κ.

the problem is to find an upper limit to the smallest positive n which satisfies the congruences

(1) $$g_i(n) \equiv \eta/\lambda \quad (i = 1, 2, \ldots, m).$$

Suppose that $a_1 \ldots a_j \ldots$ is any sequence of integers, and that $p_0 \ldots p_j \ldots$ is a sequence defined as follows :—$p_0 = 1$; p_1 is the smallest positive integer which satisfies

$$p_1 \,.\, 2a_i \equiv \eta a_1^{-1} \quad (i = 1, 2, \ldots, m);$$

and generally p_j is the smallest positive integer such that

$$\left.\begin{array}{ll} p_j \,.\, 2a_i & \equiv \eta \,.\, a_j^{-1} \\ p_j \,.\, 2(1+p_1)\,a_i & \equiv \eta \,.\, a_j^{-1} \\ \ldots \quad \ldots \quad \ldots & \ldots \quad \ldots \\ p_j \,.\, 2\,(1+p_1+\ldots\, p_{j-1})\,a_i & \equiv \eta \,.\, a_j^{-1} \end{array}\right\} \quad (i = 1, 2, \ldots, m).$$

By the theorem of Dirichlet which has been quoted above, p_j exists and

(2) $$a_j^{mj} > p_j > 0.$$

Write, now, $q_s = \sum_0^s p_j$, and it then follows from the definitions that

(3) $$2p_t \,.\, q_s \,.\, a_i \equiv \eta a_t^{-1}$$

$$(i = 1, 2, \ldots, m;\ s = 0, 1, \ldots;\ t = s+1, s+2, \ldots),$$

and from (2) that

(4) $$0 < q_{t-1} < q_t < 1 + \sum_1^t a_j^{mj};$$

and, further, in virtue of (3) and the identical equation

$$g_i(X) = g_i(X') + g_i(X-X') + 2(X-X')\,X'a_i,$$

$$\begin{aligned} g_i(q_t) &= g_i(q_s) + g_i(q_t - q_s) + 2\,(q_t - q_s)\,q_s a_i \\ &= g_i(q_s) + g_i(q_t - q_s) + \sum_{r=s+1}^{t} 2p_r q_s a_i \ (t > s) \\ &\equiv g_i(q_s) + g_i(q_t - q_s) + \eta \sum_{s+1}^{t} a_j^{-1}, \end{aligned}$$

that is to say

(5) $$g_i(q_t) - g_i(q_s) \equiv g_i(q_t - q_s) + \eta \sum_{s+1}^{t} a_j^{-1}$$

$$(i = 1, 2, \ldots, m;\ s = 0, 1, \ldots;\ t = s+1, s+2, \ldots).$$

The result can now be obtained from (4) and (5) by following the line of argument used by Dirichlet.

Let R be an m-dimensional manifold in which the coordinates range from 0 to 1, and, taking any integer μ, suppose the range of each coordinate is divided into μ equal parts; R will then be sub-divided into μ^m parts in such a way that the corresponding coordinates of two points, which lie in the same part, do not differ by more than $1/\mu$. Now R contains 2^m integral points, the coordinates of which are permutations of 0 and 1, and each of them lies on the boundary of one and of only one of the sub-divisions of R; these sub-divisions, therefore, can be separated into two classes, firstly the 2^m "corner" ones, which contain each an integral point, and secondly the remainder which number μ^m-2^m. Consider those points of R of which the coordinates are congruent to

$$\{g_1(q_t),\ g_2(q_t),\ \ldots,\ g_m(q_t)\} \quad (t=0,\ 1,\ \ldots,\ \mu^m-2^m).$$

To each value of t corresponds a point of R and to each, in general, only one point, but there are more than one if any of the numbers $g_i(q_t)$ is an integer; in the latter case, it may be supposed that some particular one is selected. Since there are μ^m-2^m+1 values of t, one of two things is true: either there is a value of t, for which the corresponding point of R lies in a "corner" part, and in this case, since each such part contains an integral point,

$$g_i(q_t) \equiv \eta/\mu \quad (i=1,\ 2,\ \ldots,\ m;\ t \leqslant \mu^m-2^m);$$

or else there are two distinct values of t, t_1 and t_2, such that the corresponding points both lie in the same sub-division of R, and in this case

$$g_i(q_{t_1})-g_i(q_{t_2}) \equiv \eta/\mu \quad (i=1,\ 2,\ \ldots,\ m),$$

and hence by (5)

$$g_i(q_{t_1}-q_{t_2}) \equiv \eta\left(\mu^{-1}+\sum_{t_2+1}^{t_1} a_j^{-1}\right) \quad (0 \leqslant t_2 < t_1 \leqslant \mu^m-2^m).$$

In either case, it follows from (4) that

$$g_i(n) \equiv \eta\left(\mu^{-1}+\sum_{1}^{\mu^m-2^m} a_j^{-1}\right) \quad (i=1,\ 2,\ \ldots,\ m), \tag{6}$$

for some integer n which satisfies the inequalities

$$0<n<1+\sum_{1}^{\mu^m-2^m} a_j^{mj}. \tag{7}$$

Arbitrary positive values can be assigned to μ and $a_1 \ldots a_j \ldots$ in (6) and (7), and if values are assigned in any way that makes

$$\mu^{-1}+\sum_{1}^{\mu^m-2^m} a_j^{-1} \leqslant \lambda^{-1}, \tag{8}$$

(7) gives an upper limit to the smallest positive n which satisfies (1).

If, for instance, a_j is taken equal to $\lambda^{(m+1)\beta}$ and β is determined so that $\beta \to 1$ and $\lambda^{\beta-1} \to \infty$ as λ tends to infinity, then (8) is satisfied by taking μ equal to $\lambda(1+\epsilon)$, and (7) gives

$$n < \lambda^{m(m+1)\lambda^m(1+\epsilon)}, \tag{9}$$

where $\epsilon \to 0$ as $\lambda \to \infty$. More refined methods of dealing with (7) and (8) appear to lead to results of the same order of magnitude. It follows from this argument that if in the theorem of Mr. Hardy and Mr. Littlewood referred to above κ is equal to 1, then the expression on the right hand side of (9) can be substituted for Φ. The significance of the result is made more evident if Φ is supposed to be given instead of λ. In this form, the theorem asserts that if q be an integer, then among the numbers $1, 2, \ldots, q$ there is at least one, n, such that

$$n\,.\,f_i(n) \equiv \eta/(\log q)^{\frac{1}{m}-\delta} \quad (i=1, 2, \ldots, m),$$

where δ tends to zero as $q \to \infty$. It is easy to show that the order of δ is at any rate not greater than $\log\log\log q/m \log\log q$.

Corresponding results for an arbitrary value of κ can be obtained by induction. Assuming that Φ is known for $\kappa = \kappa_0 - 1$, the argument for the case $\kappa = \kappa_0$ proceeds step by step as in the case $\kappa = 1$; the only difference consists in the relations which define p_j and the identity from which (5) is deduced.

The identity which is required in the general case is merely that if $g(X)$ is a polynomial in X from which the absolute term is missing and if

$$y_t = \sum_{0}^{t} X_j,$$

then $g(y_t) - g(y_s) - g(y_t - y_s)$

$$= \sum_{s+1}^{t} X_j \left[\{g(y_j) - g(y_{j-1}) - g(y_j - y_s) + g(y_{j-1} - y_s)\}/X_j \right] \quad (t > s).$$

If g is of degree κ_0, the function between square brackets on the right hand side is a polynomial of degree $\kappa_0 - 2$ in X_j, the coefficients of which are polynomials in y_{j-1} and y_s. If it is written equal to $h_i(X_j, y_{j-1}, y_s)$, then the congruences which define p_j are

$$p_j h_i(p_j, q_{j-1}, q_s) \equiv \eta a_j^{-1} \quad (i = 1, 2, \ldots, m; \; s = 0, 1, \ldots, j-1),$$

and we have instead of (2)

$$\Phi(mj, \kappa_0 - 1, a_j) > p_j > 0,$$

and instead of (7) $$0 < n < \sum_{1}^{\mu''' - 2^m} \Phi(mj, \kappa_0 - 1, a_j) + 1.$$

Writing $$l_0 q = q, \quad l_1 q = \log q, \quad l_2 q = \log\log q, \quad \ldots,$$

the general theorem is as follows :—

"If m, q, and κ are integers and if $f_1(X) \ldots f_m(X)$ are m arbitrary polynomials in X with real coefficients and of degree not greater than κ, then among the numbers 1, 2, ..., q there is at least one, n, which satisfies the congruences

(10) $$n \,.\, f_i(n) \equiv \eta/(l_\kappa q)^{\frac{1}{m} - \delta} \quad (i = 1, 2, \ldots, m),$$

where δ tends to zero as q tends to infinity and is of order not greater than $l_{\kappa+2} q / m l_{\kappa+1} q$."

The value of $\Phi(m, \kappa, \lambda)$ which is implied by (10) increases so rapidly, as $\lambda \to \infty$, that it appears improbable that this result could not be substantially improved, but I have not been able to prove that it could. In the case $m = 1$, $\kappa = 0$, we can take $\Phi(1, 0, \lambda) = \lambda - 1$, and we cannot take $\Phi(1, 0, \lambda)$ less than $\lambda - 1$, since if ϑ is equal to $1/\lambda - 1$ then $r\vartheta$ is not congruent to η/λ for any r which is less than $\lambda - 1$. In the case $m = 1$, $\kappa = 1$, it can be shewn that we can *not* take $\Phi(1, 1, \lambda) = C\lambda$, where C is constant; that is to say, that for all values of C there are an infinity of λ's and ϑ's for which the difference between $r^2\vartheta$ and the nearest integer is greater than $1/\lambda$ if r is less than or equal to $C\lambda$. More than this I have not been able to prove.