Now $R \ll J$, so by our assumption on V we may omit the first term on the right. Then

$$JV^4 \ll DJ^{2/3}T^{2/3}(DT)^\varepsilon,$$

which implies the assertion (6.4), in the case $D \ll T^B$. But otherwise (1.27) follows directly from the fourth moment and the estimate $L(1/2+it, \chi) \ll (DT)^{1/4+\varepsilon}$.

References

[B–I] E. Bombieri and H. Iwaniec, *On the order of* $\zeta(1/2+it)$, Ann. Scuola Norm. Sup. Pisa, Cl. Sci. 13 (1986), 449–472.

[G] A. Good, *The square mean of Dirichlet series associated with cusp forms*, Mathematika 29 (1982), 278–295.

[H1] D. R. Heath-Brown, *The twelfth power moment of the Riemann zeta-function*, Quart. J. Math. Oxford 29 (1978), 443–462.

H2] — *The growth rate of the Dedekind Zeta-function on the critical line*, Acta Arith. 49 (1988), 323–339.

[H–W] M. N. Huxley and N. Watt, *Exponential sums and the Riemann zeta function*, Proc. London Math. Soc. (3) 57 (1988), 1–24.

[Iv] A. Ivić, *The Riemann Zeta-Function*, John Wiley & Sons, 1985.

[Iw] H. Iwaniec, *Fourier coefficients of cusp forms and the Riemann zeta-function*, Séminaire de Théorie des Nombres, Univ. Bordeaux 1979/80, exposé no 18, 36 pp.

[J1] M. Jutila, *Lectures on a method in the theory of exponential sums*, Tata Institute of Fundamental Research, Lectures in Mathematics and Physics 80, Bombay 1987.

[J2] — *The fourth power moment of the Riemann zeta-function over a short interval*, Coll. Soc. János Bolyai 52, Number Theory (Budapest 1987), North-Holland, Amsterdam 1990, 221–244.

[J3] — *Mean value estimates for exponential sums*, Number Theory, Ulm 1987, Lecture Notess in Mathematics 1380, Springer, 1989, 120–136.

[J4] — *On exponential sums involving the Ramanujan function*, Proc.Indian Acad. Sci. (Math. Sci.) 97 (1987), 157–166.

[J5] — *Exponential sums connected with quadratic forms*, Proc. of the First Conference of the Canadian Number Theory Association, Banff, Alberta, 1988, Walter de Gruyter, 1990, 271–286.

[Me1] T. Meurman, *The mean twelfth power of Dirichlet L-functions on the critical line*, Ann. Acad. Sci. Fenn. Ser. A I, Dissertationes, 52 (1984), 44 pp.

[Me2] — *On exponential sums involving the Fourier coefficients of Maass wave forms*, J. Reine Angew. Math. 384 (1988), 192–207.

[Mo] H. L. Montgomery, *Topics in Multiplicative Number Theory*, Lecture Notes in Mathematics 227, Springer, 1971.

[S] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Princeton University Press, 1971.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TURKU
SF-20500 Turku 50
Finland

Received on 21.2.1989 (1909)

ACTA ARITHMETICA
LVII (1991)

Khintchine-type theorems on manifolds

by

M. M. Dodson (York), B. P. Rynne and J. A. G. Vickers (Southampton)

To the memory of Professor V. G. Sprindžuk

1. Introduction. Sprindžuk made fundamental contributions to the difficult problem of extending classical results on metric Diophantine approximation to submanifolds, or in his terminology, from the case of "independent variables" to that of "dependent variables" [9]. In this paper some Khintchine-type theorems are obtained for a fairly general class of manifolds.

For any vectors $\boldsymbol{x} = (x_1, \ldots, x_k)$, $\boldsymbol{y} = (y_1, \ldots, y_k)$ in $\boldsymbol{R}^k$ we write

$$\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{k} x_i y_i \quad \text{and} \quad |\boldsymbol{x}| = \max\{|x_i|: i = 1, \ldots, k\}.$$

For any real number t let

$$\|t\| = \inf\{|t-p|: p \in \boldsymbol{Z}\}.$$

Let $\psi(r)$, $r = 1, 2, \ldots$, be a sequence of numbers with $\psi(r) \in [0, 1/2]$. It follows from Groshev's generalisation of Khintchine's theorem ([9], Chap. 1, Theorem 12) that for almost all $\boldsymbol{x} \in \boldsymbol{R}^k$ the inequality

$$\|\boldsymbol{q} \cdot \boldsymbol{x}\| < \psi(|\boldsymbol{q}|) \tag{1.1}$$

has finitely many solutions $\boldsymbol{q} \in \boldsymbol{Z}^k$ if the series

$$\sum_{r=1}^{\infty} \psi(r) r^{k-1} \tag{1.2}$$

converges and infinitely many solutions if the series diverges (providing $\psi(r)$ satisfies certain monotonicity conditions when $k = 1$ or 2). Khintchine's theorem on simultaneous Diophantine approximation ([9], Chap. 1, Theorem 8) asserts that the dual system of inequalities

$$\|qx_i\| < \psi(|q|), \quad i = 1, \ldots, k, \tag{1.3}$$

has finitely many solutions $q \in \boldsymbol{Z}$ for almost all $\boldsymbol{x} \in \boldsymbol{R}^k$ if the series

$$\sum_{r=1}^{\infty} \psi(r)^k \tag{1.4}$$

converges (the corresponding result holds when (1.4) diverges if ψ satisfies appropriate monotonicity conditions; we will not consider this case but details are given in [5] and [9]).

When the points x in the inequalities (1.1) and (1.3) are restricted to lie on a smooth submanifold $M \subset \mathbf{R}^k$ of dimension $m < k$, the above results tell us nothing since M has measure zero. However, the same questions as in the Euclidean case can still be considered but with respect to the induced measure on M. Despite the substantial difficulties posed by the functional relationship between the coordinates of the points on the manifold there has been progress in carrying over the classical theory to manifolds. Surveys of known results and further references are given in [4], [9] and [10]. In one direction, conditions have been sought which ensure that for almost all $x \in M$ the inequality (1.3), with the special sequence $\psi(r) = cr^{-1/k-\varepsilon}$ (or, equivalently, inequality (1.1) with $\psi(r) = cr^{-k-\varepsilon}$), has at most finitely many solutions for arbitrary $c, \varepsilon > 0$ (see [9]). Such manifolds are called extremal. It is an immediate consequence of the results below that the manifolds we consider are extremal; in [6] manifolds satisfying slightly weaker conditions have been shown to be extremal. In the other direction, it has been shown that if M belongs to a very general class of manifolds then for almost all $x \in M$ the inequality (1.3) with $\psi(r) = cr^{-1/k}$, $c > 0$, has infinitely many solutions ([1], Theorem 1). This result, together with Khintchine's transfer principle ([5], Chap. V, corollary to Theorem II), shows that when $\psi(r) = cr^{-k}$, $c > 0$, (1.1) has infinitely many solutions for almost all $x \in M$.

For general sequences ψ the position is less satisfactory, particularly when the series are divergent, although some special cases are known. For example, if M consists of the cartesian product of sufficiently many curves then some partial results have been obtained for the inequality (1.3) in both the divergent and convergent cases (for further details see [2], [3], [4] and Section 12, Chap. 2 of [9]). When M is the (1-dimensional) curve $\{(x, \ldots, x^k) \colon x \in \mathbf{R}\}$ and (1.2) converges, then (1.1) holds infinitely often almost nowhere on M ([4], Theorem 3).

In this paper we extend these results to general C^3 manifolds satisfying certain non-zero curvature conditions (which exclude curves).

Much of the notation and terminology we will use is taken from [6]. In particular, we suppose that M is a C^3 manifold of dimension m and codimension n embedded in $\mathbf{R}^k$ $(k = m+n)$. For any point $x \in M$, $T_x M^\perp$ will denote the normal space of M at x, and for any $\gamma \in T_x M^\perp$, $\varkappa^i(x, \gamma)$, $i = 1, \ldots, m$, will denote the principal curvatures of M at x with respect to γ (further details are given in [8]).

DEFINITION. We say that the manifold M satisfies *condition K1* at $x \in M$ if, for any $\gamma \in T_x M^\perp$, at least two of the principal curvatures $\varkappa^i(x, \gamma)$ are non-zero and have the same sign.

Now suppose that $x \in M$ and N is a codimension 1 hyperplane which intersects M transversely at x. Then the set $N \cap M$ is a C^3 $(m-1)$-dimensional manifold embedded in the plane N in a neighbourhood of x. The above definition of condition K1 for a manifold (with respect to an embedding in $\mathbf{R}^{m+n}$) can readily be extended to the case of an embedding in a hyperplane. This enables us to make the following

DEFINITION. The manifold M satisfies *condition K2* at $x \in M$ if there exists a codimension 1 hyperplane $N(x) \subset \mathbf{R}^{m+n}$ which intersects M transversely at x with the property that the set $N(x) \cap M$ satisfies condition K1 at x with respect to its embedding in $N(x)$.

The conditions K1 and K2 impose some restrictions on the sizes of m and n. In particular, for K1 and K2 to hold it is necessary that $m \geqslant 2$ and $m \geqslant 3$ respectively. Thus, we need $k = m+n \geqslant 3$ for K1 or K2 to hold. Further restrictions on the relative sizes of m and n can be deduced as in Section 2 of [6] where a similar condition is discussed.

To illustrate the geometrical meaning of the above conditions consider the case where M is an m-dimensional hypersurface in $\mathbf{R}^{m+1}$, i.e. when M has codimension 1. In this case, at any point $x \in M$ the unit normal vector γ (with respect to the Euclidean norm) is uniquely determined (up to a factor of ± 1) and the Gaussian curvature $K(x)$ of M at x is given by the product $K(x) = \varkappa^1(x, \gamma) \ldots \varkappa^m(x, \gamma)$. Thus, when $m = 2$ condition K1 at x is equivalent to the condition that $K(x) > 0$, so that M must be "bowl shaped" rather than "saddle shaped". However, when $m \geqslant 3$ the condition $K(x) \neq 0$ is sufficient (but not necessary) for condition K1 to hold. To deal with condition K2 we first note that for any $m \geqslant 3$, $n \geqslant 1$, condition K2 implies K1, but not conversely in general. However, if M is a hypersurface then K1 implies K2 (these results can be deduced from the alternative formulations of the conditions discussed in Section 3 and in the proof of Lemma 4.1 below). Thus, in this case K2 holds if M has non-zero Gaussian curvature.

We will now state our main results.

THEOREM 1.1. *Suppose that condition K1 holds at almost all $x \in M$. If the series* (1.2) *is convergent then, for almost all $x \in M$, the inequality* (1.1) *has at most finitely many solutions $q \in \mathbf{Z}^{m+n}$.*

THEOREM 1.2. *Suppose that condition K2 holds at almost all $x \in M$. If the series* (1.2) *is divergent then, for almost all $x \in M$, the inequality* (1.1) *has infinitely many solutions $q \in \mathbf{Z}^{m+n}$.*

THEOREM 1.3. *Suppose that condition K1 holds at almost all $x \in M$. If the series* (1.4) *is convergent then, for almost all $x \in M$, the system of inequalities* (1.3) *has at most finitely many solutions $q \in \mathbf{Z}$.*

These theorems can be generalised in the same way that Theorem 12 (Groshev's theorem) is generalised by Theorems 13, 14 and 15 in Ch. 1 of [9].

These generalisations will be omitted since they can readily be derived from the results below and in [9]. Note that Theorem 1.1 extends (but does not include) Theorem 3 of [4]. This suggests that Theorem 1.1 should hold for curves which satisfy a suitable curvature condition.

2. Preliminary estimates. In the following we will use the notation $e(t) = \exp(2\pi i t)$ and $|X|$ for the Lebesgue measure of a measurable set X in Euclidean space.

LEMMA 2.1. *Let $\Omega \subset \mathbf{R}^l$ be an open set and let $g\colon \Omega \to \mathbf{R}$ be a C^3 function and let $G(\varrho) = \{\mathbf{u} \in \Omega\colon \|g(\mathbf{u})\| < \varrho\}$ for all $\varrho \in [0, 1/2]$. Suppose that the inequality*

$$(2.1)\qquad \Big|\int_{\Omega} e(jg(\mathbf{u}))\,d\mathbf{u}\Big| \leqslant \alpha j^{-1}$$

holds for some $\alpha > 0$ and all positive integers j. Then

$$(2.2)\qquad \big||G(\varrho)| - 2\varrho|\Omega|\big| \leqslant 8\alpha\varrho(1 + |\log \varrho|).$$

Proof. When $\varrho = 0$ the result is clear, so we suppose that $\varrho > 0$. The Fourier coefficients $a_j(\varrho)$, $j \in \mathbf{Z}$, of the function $\chi_\varrho\colon \mathbf{R} \to \mathbf{R}$ defined by $\chi_\varrho(t) = 1$ if $\|t\| < \varrho$ and $\chi_\varrho(t) = 0$ otherwise, are given by $a_0(\varrho) = 2\varrho$, $a_j(\varrho) = \{e(j\varrho) - e(-j\varrho)\}/(2\pi i j)$, $j \neq 0$, so that $|a_j(\varrho)| \leqslant \min\{2\varrho, 2j^{-1}\}$ for all $j \neq 0$, $\varrho \in [0, 1/2]$. Now

$$|G(\varrho)| = \int_{\Omega} \chi_\varrho(g(\mathbf{u}))\,d\mathbf{u} = 2\varrho|\Omega| + \sum_{j \neq 0} a_j(\varrho) \int_{\Omega} e(jg(\mathbf{u}))\,d\mathbf{u},$$

and the result follows from (2.1) and the estimate for $|a_j(\varrho)|$.

The next two lemmas will be used in applying Lemma 2.1 in the following sections. The first is a variant of Lemma 4.2 of [11].

LEMMA 2.2. *Let $I \subset \mathbf{R}$ be an interval and let $g\colon I \to \mathbf{R}$ be a C^2 function such that*

$$|g'(u)| \geqslant \alpha, \quad |g''(u)| \leqslant c\alpha, \quad u \in I,$$

for some constants c, $\alpha > 0$. Then for any subinterval $A \subset I$,

$$\Big|\int_A e(g(u))\,du\Big| \leqslant (1 + c|A|)\pi^{-1}\alpha^{-1}.$$

Proof. Letting b and a denote the upper and lower end points of A respectively we have, on integrating by parts,

$$\begin{aligned}\Big|\int_a^b e(g(u))\,du\Big| &= \Big|\int_a^b g'(u)^{-1} e(g(u))\,g'(u)\,du\Big| \\ &= \Big|g'(b)^{-1}\big(e(g(b)) - e(g(a))\big) + \int_a^b g''(u)\,g'(u)^{-2}\big(e(g(u)) - e(g(a))\big)\,du\Big|/(2\pi) \\ &\leqslant (\pi\alpha)^{-1} + c(\pi\alpha)^{-1}|A|.\end{aligned}$$

LEMMA 2.3. *Let $g\colon \mathbf{R}^2 \to \mathbf{R}$ be a C^3 function and suppose that*

$$(2.3)\qquad \frac{\partial^2 g}{\partial u_i^2}(\mathbf{u}) \geqslant \alpha, \quad i = 1, 2, \qquad \Big|\frac{\partial^2 g}{\partial u_1 \partial u_2}(\mathbf{u})\Big| \leqslant \alpha/2, \qquad |D^3 g(\mathbf{u})| \leqslant c\alpha,$$

for all $\mathbf{u} \in \mathbf{R}^2$, for some constants c, $\alpha > 0$. Then for any convex set $A \subset \mathbf{R}^2$,

$$(2.4)\qquad \Big|\int_A e(g(\mathbf{u}))\,d\mathbf{u}\Big| \leqslant 16(1 + c|A|)\pi^{-1}\alpha^{-1}.$$

Proof. It follows from the inequalities (2.3) that there exists a unique point $\mathbf{u}_{00} \in \mathbf{R}^2$ such that

$$(2.5)\qquad \frac{\partial g}{\partial u_i}(\mathbf{u}_{00}) = 0, \quad i = 1, 2,$$

(see [6], § 3). We may assume that $\mathbf{u}_{00} = \mathbf{0}$. Introducing polar coordinates $\mathbf{u}(r, \theta) = (r\cos\theta, r\sin\theta)$ and letting $h(r, \theta) = g(\mathbf{u}(r, \theta))$, it follows from (2.3) and (2.5) that

$$(2.6)\qquad \frac{\partial h}{\partial r}(0, \theta) = 0, \quad \frac{\partial^2 h}{\partial r^2}(r, \theta) \geqslant \alpha/2, \quad \Big|\frac{\partial^3 h}{\partial r^3}(r, \theta)\Big| \leqslant 8c\alpha,$$

for all $r \geqslant 0$, $\theta \in [0, 2\pi]$. In the new coordinates we have

$$(2.7)\qquad \int_A e(g(\mathbf{u}))\,d\mathbf{u} = \int_0^{2\pi} \zeta(\theta)\,d\theta,$$

where

$$\zeta(\theta) = \int_{A(\theta)} e(h(r, \theta))\,r\,dr, \quad \theta \in [0, 2\pi]$$

and $A(\theta) = \{r \geqslant 0\colon \mathbf{u}(r, \theta) \in A\}$. Since A is convex, the set $A(\theta)$ is an interval; its upper and lower end points will be denoted by $b(\theta)$ and $a(\theta)$ respectively. Now, the integral $\zeta(\theta)$ can be rewritten in the form

$$(2.8)\qquad \zeta(\theta) = \int_a^b e(h(r))\,h'(r)\frac{r}{h'(r)}\,dr,$$

where the dependence of h, a and b on θ has been suppressed and the dashes denote differentiation with respect to r (by (2.6), $h'(r) > 0$ for $r > 0$). Integrating (2.8) by parts we obtain

$$(2.9)\qquad |\zeta(\theta)| = \Big|\gamma(b)\big(e(h(b)) - e(h(a))\big) - \int_a^b \gamma'(r)\big(e(h(r)) - e(h(a))\big)\,dr\Big|/(2\pi),$$

where $\gamma(r) = r/h'(r)$. Using (2.6), we have

$$|\gamma(r)| \leqslant 2\alpha^{-1}, \quad |\gamma'(r)| \leqslant 16c\alpha^{-1}, \quad r \geqslant 0,$$

and the result follows from (2.7) and (2.9).

3. Proof of Theorem 1.1. We will henceforth assume that M has the form

$$M = \{x \in R^{m+n}: x = \xi(u), u \in \Omega\} = \xi(\Omega),$$

where Ω is a bounded open subset of R^m and the parametrisation function $\xi: \Omega \to R^{m+n}$ is C^3 and has the form

$$\xi(u) = (u_1, \ldots, u_m, \varphi_1(u), \ldots, \varphi_n(u)) = (u, \varphi(u)), \quad u \in \Omega,$$

where $\varphi(u) \in R^n$. In addition, we assume that the function ξ and its first three derivatives are uniformly bounded on Ω. These assumptions are not restrictive since it is always possible to cover M with a countable collection of local coordinate neighbourhoods on which the assumptions hold (after relabelling the coordinate axes if necessary) and it is clearly sufficient to prove the theorems on these local neighbourhoods. Also, the mapping $\xi: \Omega \to M$ maps sets of measure zero in Ω to sets of measure zero in M. Thus, to prove the theorems it suffices to show that the appropriate subsets of Ω, rather than M, have measure zero.

For any C^r function $f: V \to R^s$ defined on an open set $V \subset R^l$ let $D^i f(x)$, where $0 \leqslant i \leqslant r$, denote the ith order Fréchet derivative of f at the point $x \in V$. In particular, when f is real valued we will identify $Df(x)$ with the vector

$$\left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_l}(x)\right)$$

and $D^2 f(x)$ with the Hessian matrix of f at x.

As in [6] we define a function $\Phi_v: \Omega \to R$, for each $v \in R^n$, by

$$\Phi_v(u) = v \cdot \varphi(u), \quad u \in \Omega.$$

By calculating the second fundamental form of M at a point $x = \xi(u) \in M$ using the parametrisation ξ (a similar calculation using a slightly different parametrisation is given in [8], Chapter VII, Example 3.3), it can be shown that condition K1 holds at x if and only if for each non-zero $v \in R^n$, the matrix $D^2\Phi_v(u)$ has at least two non-zero eigenvalues with the same sign.

Since by hypothesis the set of points $x \in M$ at which condition K1 does not hold has measure zero in M, this set cannot affect the conclusions of the theorem and hence can be neglected (similarly in the proofs of Theorems 1.2 and 1.3 we will neglect the set of points at which the relevant condition does not hold). We now choose a point $x^0 = \xi(u^0)$ in M at which condition K1 holds. For notational simplicity and without loss of generality, we will suppose that $x^0 = 0$, $u^0 = 0$. Since $D^2\Phi_v(0)$ is a continuous function of v it follows from the above formulation of condition K1 that there exists a constant $\delta > 0$ such that for each $v \in S^n_\infty(1)$ (where $S^n_\infty(\varrho) = \{w \in R^n: |w| = \varrho\}$) the matrix $D^2\Phi_v(0)$ has at least two eigenvalues of the same sign with magnitude greater than 2δ.

In order to use the results of Section 2 we need a function $\tilde{\Phi}_v: R^m \to R$ for each $v \in R^n$, such that $\tilde{\Phi}_v(u)$ is a C^3 function of $(u, v) \in R^m \times R^n$,

$$|D^2\tilde{\Phi}_v(u) - D^2\Phi_v(0)| \leqslant \delta/(2m^2), \quad (u, v) \in R^m \times S^n_\infty(1), \tag{3.1}$$

and $\tilde{\Phi}_v(u) = \Phi_v(u)$ for all $(u, v) \in B^m_2(\varepsilon) \times R^n$, for some $\varepsilon > 0$, where $B^m_2(\varepsilon) = \{w \in R^m: |w|_2 < \varepsilon\}$ and $|w|_2 = (w \cdot w)^{1/2}$. Such a function can be constructed as in Lemma 3 of [7].

It follows from (3.1) that

$$|U_v(D^2\tilde{\Phi}_v(u) - D^2\Phi_v(0))U_v^{-1}| \leqslant \delta/2, \quad (u, v) \in R^m \times S^n_\infty(1), \tag{3.2}$$

where U_v is an orthogonal matrix which diagonalises the matrix $D^2\Phi_v(0)$.

Let

$$K = \sup\{|D\xi(u)|: u \in \Omega\} = \max\{1, \sup\{|D\varphi(u)|: u \in \Omega\}\}.$$

For any numbers $u, v \geqslant 0$ the notation $u \ll v$ will mean $u \leqslant cv$, where the positive constant c depends at most on m, n, δ and K. We now shrink M further by putting $\Omega = \Omega(\varepsilon) = B^m_2(\varepsilon)$ and $M = M(\varepsilon) = \xi(\Omega(\varepsilon))$. The assumption that $x^0 = 0$ implies that $M(\varepsilon) \subset B^{m+n}_2(\hat{\varepsilon})$, for some $\hat{\varepsilon}$ with $\varepsilon \leqslant \hat{\varepsilon} \ll \varepsilon$.

Lemma 3.1. *For any vector $q \in Z^{m+n}$, $q \neq 0$, we have*

$$\left|\int_\Omega e(q \cdot \xi(u))\,du\right| \ll |q|^{-1}. \tag{3.3}$$

Proof. We choose a non-zero vector $q \in Z^{m+n}$, and consider this fixed. For any $x \in R^{m+n}$, write $x = (x^{(1)}, x^{(2)}) \in R^m \times R^n$. Let $v = q^{(2)}/|q^{(2)}|$ (if $q^{(2)} = 0$ let $v = 0$) and define the functions $g: \Omega \to R$, $\tilde{g}: R^m \to R$ by

$$\begin{aligned} g(u) &= q \cdot \xi(u) = q^{(1)} \cdot u + q^{(2)} \cdot \varphi(u) = q^{(1)} \cdot u + |q^{(2)}|\,\Phi_v(u), \\ \tilde{g}(u) &= q^{(1)} \cdot u + |q^{(2)}|\,\tilde{\Phi}_v(u). \end{aligned} \tag{3.4}$$

Since $\tilde{\Phi}_v(u) = \Phi_v(u)$ on $\Omega \times R^n$, $\tilde{g}(u) = g(u)$ for all $u \in \Omega$. Now suppose that $|q^{(2)}| \leqslant |q|/(2nK)$. Then it follows from (3.4) that there is an integer j, $1 \leqslant j \leqslant m$, say $j = 1$, such that

$$\left|\frac{\partial g}{\partial u_1}(u)\right| \geqslant \tfrac{1}{2}|q|, \quad u \in \Omega. \tag{3.5}$$

For any $u \in R^m$, write $u = (u_1, u')$, where $u' = (u_2, \ldots, u_m) \in R^{m-1}$, and for any $u' \in R^{m-1}$, let $\Omega'(u') = \{u_1 \in R: (u_1, u') \in \Omega\}$. For each $u' \in B^{m-1}_2(\varepsilon)$ the set $\Omega'(u')$ is an interval, so by Lemma 2.2 and (3.5)

$$\left|\int_{\Omega'(u')} e(g(u_1, u'))\,du_1\right| \ll |q|^{-1}.$$

Thus by Fubini's theorem, (3.3) holds in this case.

To deal with the case where $|q^{(2)}| > |q|/(2nK)$ we first make the orthogonal transformation of the u coordinates in R^m determined by the matrix U_v, defined above. For simplicity we will not introduce a new notation for the transformed variables and functions of these variables. By the construction of this transformation, the matrix $D^2\Phi_v(\mathbf{0})$ is now diagonal. Furthermore, by relabelling the coordinates if necessary, we may suppose without loss of generality that the first two elements of the diagonal of $D^2\Phi_v(\mathbf{0})$ are each greater than 2δ. Therefore, it follows from (3.2) and the equation $D^2\tilde{g} = |\mathbf{q}^{(2)}| D^2\tilde{\Phi}_v$ that

$$\begin{aligned} &\frac{\partial^2\tilde{g}}{\partial u_i^2}(\mathbf{u}) \geqslant |\mathbf{q}^{(2)}|\,\delta, \quad i=1,2, \\ &\left|\frac{\partial^2\tilde{g}}{\partial u_1\,\partial u_2}(\mathbf{u})\right| \leqslant |\mathbf{q}^{(2)}|\,\delta/2, \end{aligned} \qquad \mathbf{u}\in\mathbf{R}^m. \tag{3.6}$$

In this case for $\mathbf{u}\in\mathbf{R}^m$, write $\mathbf{u} = (u_1, u_2, \mathbf{u}'')$, $\mathbf{u}'' = (u_3, \ldots, u_m)\in\mathbf{R}^{m-2}$ and for $\mathbf{u}''\in\mathbf{R}^{m-2}$, let $\Omega''(\mathbf{u}'') = \{(u_1, u_2)\in\mathbf{R}^2\colon (u_1, u_2, \mathbf{u}'')\in\Omega\}$. For each $\mathbf{u}''\in B_2^{m-2}(\varepsilon)$ the set $\Omega''(\mathbf{u}'')$ is convex so by Lemma 2.3 and (3.6)

$$\Big|\int_{\Omega''(\mathbf{u}'')} e(\tilde{g}(u_1, u_2, \mathbf{u}''))\,du_1\,du_2\Big| \ll |\mathbf{q}|^{-1}.$$

Hence the result again follows by Fubini's theorem, which completes the proof of the lemma.

We now define the sets

$$B(\mathbf{x}, b; \varrho) = \{\mathbf{u}\in\Omega\colon |\mathbf{x}\cdot\xi(\mathbf{u}) - b| < \varrho\} \subset \Omega = B_2^m(\varepsilon),$$

$$B(\mathbf{x}; \varrho) = \{\mathbf{u}\in\Omega\colon \|\mathbf{x}\cdot\xi(\mathbf{u})\| < \varrho\} = \bigcup_{b\in\mathbf{Z}} B(\mathbf{x}, b; \varrho),$$

for any $\mathbf{x}\in\mathbf{R}^{m+n}$, $b\in\mathbf{Z}$, $\varrho\in[0, 1/2]$. Combining Lemmas 2.1 and 3.1 yields

LEMMA 3.2. *If ε is as above then, for any $\varrho\in[0, 1/2]$ and $\mathbf{q}\in\mathbf{Z}^{m+n}$, $\mathbf{q}\neq\mathbf{0}$,*

$$\big||B(\mathbf{q}; \varrho)| - 2\varrho|\Omega|\big| \leqslant c_1\varrho|\mathbf{q}|^{-1}(1 + |\log\varrho|), \tag{3.7}$$

where $c_1 > 0$ is a constant.

Now let $\psi^*(0) = \psi^*(1) = 1/2$, and

$$\psi^*(r) = \max\{\psi(r), r^{-m-n-1}\}, \quad r = 2, 3, \ldots$$

Clearly the series

$$\sum_{r=1}^{\infty} r^{m+n-1}\psi^*(r) \tag{3.8}$$

is convergent if and only if the series (1.2) is convergent. In addition,

$$r^{-1}(1 + |\log\psi^*(r)|) \ll r^{-1}\log r, \tag{3.9}$$

for all $r \geqslant 2$, and hence by (3.7)

$$|\Omega|\,\psi^*(|\mathbf{q}|) \leqslant |B(\mathbf{q}; \psi^*(|\mathbf{q}|))| \leqslant 3|\Omega|\,\psi^*(|\mathbf{q}|), \tag{3.10}$$

for all sufficiently large $\mathbf{q}$. It now follows from the Borel–Cantelli lemma that the inequality

$$\|\mathbf{q}\cdot\xi(\mathbf{u})\| < \psi^*(|\mathbf{q}|) \tag{3.11}$$

has at most finitely many solutions for almost all $\mathbf{u}\in\Omega$. Since $\psi(r) \leqslant \psi^*(r)$ for all r, this proves Theorem 1.1.

4. Proof of Theorem 1.2. We now suppose that the series (1.2) is divergent and the manifold M satisfies condition K2 almost everywhere. Since condition K2 implies K1 it follows that condition K1 holds almost everywhere on M and so Theorem 1.1 holds on M. Thus, for almost all $\mathbf{x}\in M$ the inequality

$$\|\mathbf{q}\cdot\mathbf{x}\| < |\mathbf{q}|^{-m-n-1}, \tag{4.1}$$

has at most finitely many solutions. Therefore, the inequality (3.11) has at most finitely many more solutions than (1.1) for almost all $\mathbf{x}\in M$, and it is sufficient to prove Theorem 1.2 with the function ψ replaced by ψ^*. For notational convenience we will now relabel ψ^* as ψ.

To prove the theorem it is necessary to estimate the measure of the set $B(\mathbf{p}; \varrho)\cap B(\mathbf{q}; \sigma)$ for linearly independent vectors $\mathbf{p}, \mathbf{q}\in\mathbf{Z}^{m+n}$ and for $\varrho, \sigma\in[0, 1/2]$. To obtain an appropriate estimate for the measure of this set for arbitrary pairs of vectors $\mathbf{p}$ and $\mathbf{q}$ is rather difficult so we restrict our attention to the set of vectors lying in a suitably chosen cone. To define this cone we now suppose, in addition to the assumptions at the beginning of Section 3, that condition K2 holds at the point $\mathbf{x}^0 = \mathbf{0}$ and let $\mathbf{n}^0$ denote a unit normal vector to the hyperplane $N^0 = N(\mathbf{0})$ (see the definition of condition K2). Also, for any vectors $\mathbf{x}, \mathbf{y}\in\mathbf{R}^{m+n}$, let $\theta(\mathbf{x}, \mathbf{y})$ denote the unique number satisfying $0 \leqslant \theta(\mathbf{x}, \mathbf{y}) < 2\pi$ and $\cos\theta(\mathbf{x}, \mathbf{y}) = \mathbf{x}\cdot\mathbf{y}/(|\mathbf{x}|_2|\mathbf{y}|_2)$. For any positive number β, let

$$C_\beta = \{\mathbf{x}\in\mathbf{R}^{m+n}\colon \theta(\mathbf{x}, \mathbf{n}^0) < \beta\}.$$

LEMMA 4.1. *If $\varepsilon, \beta > 0$ are sufficiently small then for all $\varrho, \sigma\in[0, 1/2]$ and all sufficiently large linearly independent integer vectors $\mathbf{p}, \mathbf{q}\in C_\beta$,*

$$|B(\mathbf{p}; \varrho)\cap B(\mathbf{q}; \sigma)| \ll \varrho\sigma|\Omega| + \varrho\sigma|\log\sigma|\,|\mathbf{q}|^{-1}\varepsilon\theta(\mathbf{p}, \mathbf{q})^{-1}. \tag{4.2}$$

Proof. We prove the result for arbitrary sufficiently large linearly independent vectors $\mathbf{x}, \mathbf{y}\in C_\beta$. We first deal with the special case where $\mathbf{n}^0 = (1, 0, \ldots, 0)$ and $N^0 = \{\mathbf{x}\in\mathbf{R}^{m+n}\colon x_1 = 0\}$. Using the coordinates $(x_2, \ldots, x_{m+n})$ in the plane N^0 it is clear that the set $N^0\cap M$, regarded as a subset of N^0, consists of the manifold whose parametrisation is given by the function

$$\mathbf{u}' \to (\mathbf{u}', \varphi(0, \mathbf{u}')), \quad \mathbf{u}'\in B_2^{m-1}(\varepsilon).$$

By the definition of condition K2 this manifold must satisfy condition K1 at the point $\mathbf{x}^0$. Since the above parametrisation is of the form considered in Section 3 it follows from the parametric formulation of condition K1 discussed there that K1 holds if and only if, for each non-zero $\mathbf{v}\in\mathbf{R}^n$, the matrix

$D'^2\Phi_v(\mathbf{0})$ has at least two non-zero eigenvalues with the same sign (D' denotes differentiation with respect to $\mathbf{u}'$). Since $D'^2\Phi_v(\mathbf{u})$ is a continuous function of $(\mathbf{u}, \mathbf{v})$ there exist $\delta > 0$, $\varepsilon^0 > 0$, such that for each $\mathbf{v} \in S^n_\infty(1)$, $\mathbf{u} \in B^m_2(\varepsilon^0)$, the matrix $D'^2\Phi_v(\mathbf{u})$ has at least two eigenvalues of the same sign with magnitude greater than δ.

Now suppose that $\beta < \frac{1}{2}(m+n)^{-1}$, so that for all $\mathbf{x} \in C_{2\beta}$ we have $x_1 > 0$, $|\mathbf{x}| = x_1$ and

(4.3) $$|\mathbf{x}^{(1)'}| \leqslant \beta|\mathbf{x}|, \quad |\mathbf{x}^{(2)}| \leqslant \beta|\mathbf{x}|,$$

where $\mathbf{x}^{(1)'} = (x_2, \ldots, x_m)$.

Before dealing with arbitrary vectors $\mathbf{x}, \mathbf{y} \in C_\beta$ we first suppose that $\mathbf{y} = (y, \mathbf{0}, \mathbf{0})$, $y > 0$. Let $\mathbf{x} \in C_{2\beta}$, $a \in \mathbf{Z}$ and write $\hat{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$, $\hat{a} = a/|\mathbf{x}|$ and $\theta = \theta(\mathbf{x}, \mathbf{y})$. It is clear that $\theta \gg \max\{|\hat{\mathbf{x}}^{(1)'}|, |\hat{\mathbf{x}}^{(2)}|\}$. We now define the function $g\colon B^m_2(\varepsilon^0) \to \mathbf{R}$ by

(4.4) $$g(\mathbf{u}) = \hat{\mathbf{x}}\cdot\boldsymbol{\xi}(\mathbf{u}) - \hat{a} = u_1 + \hat{\mathbf{x}}^{(1)'}\cdot\mathbf{u}' + \hat{\mathbf{x}}^{(2)}\cdot\boldsymbol{\varphi}(\mathbf{u}) - \hat{a}, \quad \mathbf{u} \in B^m_2(\varepsilon^0).$$

For each $\varepsilon \in (0, \varepsilon^0)$ let $L = \varrho|\mathbf{x}|^{-1}$ and

$$B_\varepsilon(\mathbf{x}, a; \varrho) = \{\mathbf{u} \in B^m_2(\varepsilon)\colon |\mathbf{x}\cdot\boldsymbol{\xi}(\mathbf{u}) - a| < \varrho\} = \{\mathbf{u} \in B^m_2(\varepsilon)\colon |g(\mathbf{u})| < L\}.$$

There exists a constant $c_2 > 0$ such that for all $\varepsilon \in (0, \varepsilon^0)$ the set $B_\varepsilon(\mathbf{x}, a; \varrho) = \varnothing$ if $|\hat{a}| \geqslant c_2\varepsilon$. Now let $Z^0_t = \{\mathbf{u} \in B^m_2(\varepsilon^0)\colon g(\mathbf{u}) = t\}$ for each $t \in \mathbf{R}$. It is clear that $B_{\varepsilon^0}(\mathbf{x}, a; \varrho) = \bigcup_{|t| \leqslant L} Z^0_t$.

Since $\mathbf{x} \in C_{2\beta}$ it follows from (4.3) that for sufficiently small β

(4.5) $$1/2 \leqslant \left|\frac{\partial g}{\partial u_1}(\mathbf{u})\right| \leqslant 2, \quad \mathbf{u} \in B^m_2(\varepsilon^0).$$

Thus, for each $t \in [-L, L]$ the set Z^0_t is a C^3 manifold of codimension 1 in $B^m_2(\varepsilon^0)$ and has a parametrisation of the form

$$\mathbf{u}' \to \mathbf{z}(\mathbf{u}', t) = (z_1(\mathbf{u}', t), u_2, \ldots, u_m),$$

where the function $\mathbf{u}' \to \mathbf{z}(\mathbf{u}', t) \in B^m_2(\varepsilon^0)$ is defined on an open set $\Omega_t \subset B^{m-1}_2(\varepsilon^0)$. Using (4.3) and (4.4) it can be seen that for sufficiently large $|\mathbf{x}|$ and sufficiently small positive β and ε ($< \varepsilon^0$) the function $\mathbf{u}' \to \mathbf{z}(\mathbf{u}', t)$ is defined on the ball $B^{m-1}_2(\varepsilon)$ for all $t \in [-L, L]$, $|\hat{a}| \leqslant c_2\varepsilon$. Again we restrict the manifold M by putting $\Omega = \Omega(\varepsilon) = B^m_2(\varepsilon)$ and $M = M(\varepsilon) = \boldsymbol{\xi}(\Omega(\varepsilon))$. Thus, we now have

$$B(\mathbf{x}, a; \varrho) = B_\varepsilon(\mathbf{x}, a; \varrho) \subset \bigcup_{|t| \leqslant L} Z_t,$$

where $Z_t = \mathbf{z}(B^{m-1}_2(\varepsilon), t) \subset Z^0_t$ for each $t \in [-L, L]$. Also, the assumed form of $\mathbf{y}$ implies that $B(\mathbf{y}; \sigma) = \{\mathbf{u} \in B^m_2(\varepsilon)\colon \|yu_1\| < \sigma\}$.

It follows from the implicit function theorem that the function $(\mathbf{u}', t) \to z_1(\mathbf{u}', t)$ is a C^3 function on the open set $\Gamma = B^{m-1}_2(\varepsilon) \times [-L, L]$. By implicit differentiation we find that for $(\mathbf{u}', t) \in \Gamma$

(4.6) $$\frac{\partial z_1}{\partial u_i}(\mathbf{u}', t) = -\frac{\partial g}{\partial u_i}\bigg/\frac{\partial g}{\partial u_1}, \quad i = 2, \ldots, m,$$

(4.7) $$\frac{\partial^2 z_1}{\partial u_i \partial u_j}(\mathbf{u}', t)$$
$$= -\left\{\frac{\partial g}{\partial u_1}\frac{\partial^2 g}{\partial u_i \partial u_j} - \frac{\partial g}{\partial u_j}\frac{\partial^2 g}{\partial u_1 \partial u_i} - \frac{\partial g}{\partial u_i}\frac{\partial^2 g}{\partial u_1 \partial u_j} - \frac{\partial g}{\partial u_i}\frac{\partial g}{\partial u_j}\bigg/\frac{\partial g}{\partial u_1}\right\}\bigg/\left(\frac{\partial g}{\partial u_1}\right)^2,$$
$$i, j = 2, \ldots, m,$$

where the argument of the functions on the right hand sides of (4.6) and (4.7) is $\mathbf{z}(\mathbf{u}', t)$. It follows from (4.6) and the above estimates that

(4.8) $$|D'z_1(\mathbf{u}', t)| \ll \beta, \quad (\mathbf{u}', t) \in \Gamma.$$

To estimate $|Z_t \cap B(\mathbf{y}; \sigma)|$, $|t| \leqslant L$, we introduce the sets

$$R_t = \{\mathbf{u}' \in B^{m-1}_2(\varepsilon)\colon \|yz_1(\mathbf{u}', t)\| < \sigma\}.$$

When $|t| \leqslant L$, $Z_t \cap B(\mathbf{y}; \sigma) \subset \mathbf{z}(R_t, t)$, and so by (4.8), $|Z_t \cap B(\mathbf{y}; \sigma)| \ll |R_t|$. We will use Lemma 2.1 to estimate $|R_t|$.

It follows from (4.4) that there exists a constant $\eta > 0$ such that if

(4.9) $$|\hat{\mathbf{x}}^{(2)}| < \eta|\hat{\mathbf{x}}^{(1)'}|,$$

then for some integer i, $2 \leqslant i \leqslant m$,

$$\left|\frac{\partial g}{\partial u_i}(\mathbf{u})\right| \geqslant \tfrac{1}{2}|\hat{\mathbf{x}}^{(1)'}|, \quad \mathbf{u} \in B^m_2(\varepsilon^0).$$

Thus, if $\hat{\mathbf{x}}$ satisfies (4.9) then by (4.5) and (4.6)

$$\left|\frac{\partial z_1}{\partial u_i}(\mathbf{u}', t)\right| \geqslant \tfrac{1}{4}|\hat{\mathbf{x}}^{(1)'}| \gg \theta, \quad (\mathbf{u}', t) \in \Gamma.$$

This estimate together with a similar argument to that used for the first case in the proof of Lemma 3.1 shows that

(4.10) $$\left|\int_{B^{m-1}_2(\varepsilon)} e(yz_1(\mathbf{u}', t))\,d\mathbf{u}'\right| \ll y^{-1}\theta^{-1}, \quad |t| \leqslant L.$$

Hence, by Lemma 2.1

(4.11) $$|R_t| \ll \sigma(\varepsilon^{m-1} + \theta^{-1}y^{-1}|\log\sigma|), \quad |t| \leqslant L.$$

Now suppose that $\hat{\mathbf{x}}$ satisfies

(4.12) $$|\hat{\mathbf{x}}^{(2)}| \geqslant \eta|\hat{\mathbf{x}}^{(1)'}|.$$

The elements of the matrix $D'^2 z_1(\mathbf{0}, t)$ are given by (4.7), and we see from this formula that

$$D'^2 z_1(\mathbf{0}, t) = -\frac{\partial g}{\partial u_1}(z(\mathbf{0}, t))^{-1} D'^2 \Phi_{\hat{x}^{(2)}}(z(\mathbf{0}, t)) + E_t,$$

where E_t denotes the $(m-1)\times(m-1)$ matrix whose elements $(E_t)_{i-1,j-1}$, $i, j = 2, \ldots, m$, are obtained from the right hand side of (4.7) by deleting the first term inside the brackets. It follows from the discussion at the beginning of the proof that the matrix $D'^2\Phi_{\hat{x}^{(2)}}(z(\mathbf{0}, t))$ has at least two eigenvalues with the same sign, and modulus $\gg |\hat{x}^{(2)}| \gg \theta$. However, by examining the terms in (4.7) it can be seen, using the definition of g in (4.4) together with (4.5) and (4.12), that $|E| \ll |\hat{x}^{(2)}|^2$. Thus, for sufficiently small β the matrix $D'^2 z_1(\mathbf{0}, t)$ has at least two eigenvalues with the same sign and modulus greater than $2c_3\theta$, for some constant $c_3 > 0$. It can easily be seen that if ε and β are sufficiently small, then this estimate holds uniformly for any $\hat{x} \in C_{2\beta}$ satisfying (4.12). In addition, by adapting the technique used in Section 3 above, a C^3 function $\tilde{z}_1$: $\mathbf{R}^{m-1} \times [-L, L] \to \mathbf{R}$ can be constructed such that $\tilde{z}_1(\mathbf{u}', t) = z_1(\mathbf{u}', t)$, for all $(\mathbf{u}', t) \in \Gamma$, and

$$|D'^2\tilde{z}_1(\mathbf{u}', t) - D'^2 z_1(\mathbf{0}, t)| \leqslant c_3\theta/(2m^2), \qquad (\mathbf{u}', t) \in \mathbf{R}^{m-1} \times [-L, L],$$

for each $\hat{x} \in C_{2\beta}$ satisfying (4.12). Now, these estimates together with a similar argument to that used for the second case in the proof of Lemma 3.1 show that (4.10), and hence (4.11), holds. Thus, we have shown that (4.11) holds for any $\hat{x} \in C_{2\beta}$.

By implicit differentiation and (4.5) we see that

$$\left|\frac{\partial z_1}{\partial t}(\mathbf{u}', t)\right| = \left|\frac{\partial g}{\partial u_1}(z(\mathbf{u}', t))\right|^{-1} \leqslant 2, \qquad (\mathbf{u}', t) \in \Gamma,$$

hence the above results show that

$$|B(\mathbf{x}, a; \varrho) \cap B(\mathbf{y}; \sigma)| \subset \Big|\bigcup_{|t| \leqslant L} Z_t \cap B(\mathbf{y}; \sigma)\Big| \ll \sup\{|R_t|: |t| \leqslant L\}\, 2L$$
$$\ll \sigma(\varepsilon^{m-1} + \theta^{-1}y^{-1}|\log\sigma|)\varrho|\mathbf{x}|^{-1}.$$

Since the number of integers a for which the set $B(\mathbf{x}, a; \varrho)$ is non-empty is bounded by $2c_2\varepsilon|\mathbf{x}|$, we have

(4.13) $$|B(\mathbf{x}; \varrho) \cap B(\mathbf{y}; \sigma)| \ll \varrho\sigma\varepsilon^m + \varrho\sigma|\log\sigma|\, y^{-1}\varepsilon\theta^{-1}.$$

Since $\varepsilon^m \ll |B_2^m(\varepsilon)|$, this proves the result for all $\mathbf{x} \in C_{2\beta}$ for the particular choice of $\mathbf{y} = (y, \mathbf{0}, \mathbf{0})$.

Now suppose that $\mathbf{x}, \mathbf{y} \in C_\beta$ are arbitrary. We can choose an orthogonal matrix T_y so that $T_y\mathbf{y} = (|\mathbf{y}|_2, \mathbf{0}, \mathbf{0})$, $T_y C_\beta \subset C_{2\beta}$ and T_y is of the form $I + O(\beta)$ for small β. Using the matrix T_y to change coordinates in $\mathbf{R}^{m+n}$ we see that, in the new coordinate system, the vector $\mathbf{y}$ has the form assumed above, $\mathbf{x} \in C_{2\beta}$ and the coordinates of the set $B(\mathbf{x}; \varrho) \cap B(\mathbf{y}; \sigma)$ are given by $T_y(B(\mathbf{x}; \varrho) \cap B(\mathbf{y}; \sigma))$. Since T_y is orthogonal

$$|B(\mathbf{x}; \varrho) \cap B(\mathbf{y}; \sigma)| = |T_y(B(\mathbf{x}; \varrho) \cap B(\mathbf{y}; \sigma))|$$

and we may estimate this, using the new coordinates, by the above method. Also, in the new coordinates, M has a parametrisation ξ_y: $B_2^m(\varepsilon) \to \mathbf{R}^{m+n}$ of the same form as before with $\xi_y \to \xi$ uniformly on $B_2^m(\varepsilon)$ as $\beta \to 0$. Thus for all sufficiently small β, it can be seen that we obtain an estimate of the form (4.13) which holds uniformly for all $\mathbf{x}, \mathbf{y} \in C_\beta$.

Finally, to complete the proof we must deal with the general case where the hyperplane $N^0 \neq \{\mathbf{x} \in \mathbf{R}^{m+n}: x_1 = 0\}$. To do this observe that since N^0 is transverse to M at $\mathbf{x}^0$, there is an orthogonal basis in $\mathbf{R}^{m+n}$ such that, with respect to the new coordinate system defined by this basis, M has a parametrisation of the above form near the point $\mathbf{x}^0$ and the vector $\mathbf{n}^0$ has components $(1, 0, \ldots, 0)$. Thus, in this coordinate system the hyperplane N^0 has the form considered above and so the general result follows from the result in the special case above. This completes the proof of the lemma.

From now on we will use the notation $B(\mathbf{q}) = B(\mathbf{q}; \psi(|\mathbf{q}|))$. Suppose that ε and β are as in Lemma 4.1. Then, for all sufficiently large linearly independent integer vectors $\mathbf{p}, \mathbf{q} \in C_{2\beta}$, it follows from Lemma 4.1 and (3.9) that

(4.14) $$|B(\mathbf{p}) \cap B(\mathbf{q})| \ll \psi(|\mathbf{p}|)\psi(|\mathbf{q}|)|\Omega|\,\theta(\mathbf{p}, \mathbf{q})^{-1}.$$

Now let

$$I_\beta = \{\mathbf{u} \in \Omega: \mathbf{u} \in B(\mathbf{q}) \text{ for infinitely many } \mathbf{q} \in C_\beta\}.$$

To prove the theorem it is sufficient to show that $|I_\beta| = |\Omega|$. We will use the estimates (3.10) and (4.14). However, (4.14) requires that the integer vectors be linearly independent. To ensure that this holds some further definitions are needed. Let P denote the set of primitive integer vectors in $\mathbf{Z}^{m+n}$ (i.e. vectors whose components have greatest common divisor 1). If $\mathbf{p}, \mathbf{q} \in P$, and $\mathbf{p} \neq \pm\mathbf{q}$, then $\mathbf{p}$ and $\mathbf{q}$ are linearly independent ([9], p. 38). Now let $P_\beta = P \cap C_\beta$, and for $r = 1, 2, \ldots$, let $P_\beta(r) = \{\mathbf{q} \in P_\beta: |\mathbf{q}| = r\}$. If $\mathbf{p}, \mathbf{q} \in P_\beta$ and $\beta < \pi$, then $\mathbf{p} \neq -\mathbf{q}$ and so distinct vectors in P_β are linearly independent. In addition, the discussion following Theorem 14, Ch. 1 of [9] shows that

(4.15) $$\beta r^{m+n-1} \ll |P_\beta(r)| \ll \beta r^{m+n-1},$$

for all sufficiently large r ($|P_\beta(r)|$ denotes the cardinality of the set $P_\beta(r)$).

By Lemma 5, Ch. 1 of [9] we have, for any positive integer N_1,

(4.16) $$|I_\beta| \geqslant \limsup_{N_2 \to \infty} \frac{\left(\sum_{r=N_1}^{N_2} \sum_{\mathbf{q} \in P_\beta(r)} |B(\mathbf{q})|\right)^2}{\sum_{r,s=N_1}^{N_2} \sum_{\mathbf{p} \in P_\beta(r)} \sum_{\mathbf{q} \in P_\beta(s)} |B(\mathbf{p}) \cap B(\mathbf{q})|}.$$

It follows from (3.10), (4.14) and (4.15) that for sufficiently large N_1

(4.17)
$$\sum_{r=N_1}^{N_2} \sum_{q \in P_\beta(r)} |B(q)| \gg \beta|\Omega| \sum_{r=N_1}^{N_2} r^{m+n-1}\psi(r),$$

(4.18)
$$\sum_{r,s=N_1}^{N_2} \sum_{p \in P_\beta(r)} \sum_{q \in P_\beta(s)} |B(p) \cap B(q)|$$
$$\ll \beta|\Omega| \sum_{r=N_1}^{N_2} r^{m+n-1}\psi(r) + |\Omega| \sum_{\substack{r,s=N_1 \\ r \leqslant s}}^{N_2} \psi(r)\psi(s) \sum_{p \in P_\beta(r)} \sum_{\substack{q \in P_\beta(s) \\ q \neq p}} \theta(p, q)^{-1}.$$

We now estimate the quantity

(4.19)
$$\sum_{\substack{q \in P_\beta(s) \\ q \neq p}} \theta(p, q)^{-1}$$

for an arbitrary vector $p \in P_\beta(r)$, $r \leqslant s$.

The line spanned by the vector p intersects the hypercube $S_\infty^{m+n}(s)$ at the point $p(s) = (s/|p|)\,p$, and it can be seen that for any vector $q \in S_\infty^{m+n}(s)$,

(4.20)
$$|q - p(s)|\, s^{-1} \ll \theta(p, q) \ll |q - p(s)|\, s^{-1}.$$

Also, there is at most one integer vector $q \in S_\infty^{m+n}(s)$ with

(4.21)
$$|q - p(s)| < 1/2,$$

and, for any integer $j \geqslant 1$, the number of integer vectors $q \in S_\infty^{m+n}(s)$ satisfying

$$j - 1/2 \leqslant |q - p(s)| < j + 1/2$$

is $\ll j^{m+n-2}$. Thus the contribution of the vectors $q \in S_\infty^{m+n}(s)$ not satisfying (4.21) to the sum (4.19) is bounded by

$$\ll \sum_{1 \leqslant j \ll \beta s} j^{m+n-2} j^{-1} s \ll \beta s^{m+n-1}.$$

To estimate the contribution of the vector q satisfying (4.21) (if it exists) we note that for any linearly independent vectors $p, q \in Z^{m+n}$,

$$\theta(p, q) \gg |p|^{-1}|q|^{-1}$$

(this follows from the inequality $(p \cdot q)^2 \leqslant (|p|_2|q|_2)^2 - 1$, which comes from Schwarz' inequality and the fact that p, q are distinct elements of Z^{m+n}). This inequality shows that the contribution of the vector q satisfying (4.21) to (4.19) is bounded by $\ll rs \leqslant s^2$, and hence (4.19) is bounded by $\ll \beta s^{m+n-1}$, for sufficiently large s, since $m+n \geqslant 4$. This estimate together with (4.15) shows that the second term on the right hand side of (4.18) is bounded by

$$\ll \beta^2 |\Omega| \Big(\sum_{r=N_1}^{N_2} r^{m+n-1}\psi(r)\Big)^2$$

and hence, by (4.16), (4.17) and (4.18)

$$|I_\beta| \gg \limsup_{N_2 \to \infty} |\Omega| / \Big[1 + \Big(\beta \sum_{r=N_1}^{N_2} r^{m+n-1}\psi(r)\Big)^{-1}\Big],$$

for sufficiently large N_1. Since the series (1.2) is divergent it follows that $|I_\beta| \gg |\Omega|$, i.e. there exists a constant $c_4 > 0$ such that

(4.22)
$$|I_\beta| \geqslant c_4|\Omega|.$$

This result shows that the set I_β has positive measure. To prove the theorem we must show that $|I_\beta| = |\Omega|$. First note that the estimate (4.22) is uniform in the sense that if ε and β are sufficiently small and if $\hat{\Omega} \subset \Omega = B_2^m(\varepsilon)$ is an open ball then

(4.23)
$$|I_\beta \cap \hat{\Omega}| \geqslant c_5|\hat{\Omega}|,$$

where $c_5 > 0$ is independent of $\hat{\Omega}$ (this follows from the uniformity of the estimates in the proof of (4.22)). Now suppose that $|I_\beta| < |\Omega|$. Then there exists a point $u_d \in \Omega$ at which the Lebesgue density of the set I_β is zero. Therefore, there exists a sufficiently small open ball $\Omega_d \subset \Omega$ surrounding u_d such that $|I_\beta \cap \Omega_d| \leqslant c_5|\Omega_d|/2$. However, this contradicts (4.23) so we must have $|I_\beta| \geqslant |\Omega|$. Thus, since $I_\beta \subset \Omega$ we must have $|I_\beta| = |\Omega|$, which completes the proof of the theorem.

5. Proof of Theorem 1.3. For each integer $q > 0$ let

$$C(q; \psi) = \{u \in \Omega\colon \|q\xi_i(u)\| < \psi(q),\ i = 1, \ldots, m+n\}.$$

By Lemma 8, Ch. 2 of [9],

$$|C(q; \psi)| \ll \psi(q)^{m+n} \sum_{|p| \leqslant \psi(q)^{-1}} \Big|\int_\Omega e(p \cdot \xi(u))\, du\Big|$$

and so by Lemma 3.1 we have for all positive q,

(5.1)
$$|C(q; \psi)| \ll \psi(q)^{m+n}\Big(1 + \sum_{\substack{|p| \leqslant \psi(q)^{-1} \\ p \neq 0}} q^{-1}|p|^{-1}\Big)$$
$$\ll \psi(q)^{m+n}(1 + q^{-1}\psi(q)^{-m-n+1}).$$

Now let $\psi^*(0) = \psi^*(1) = 1/2$, and

$$\psi^*(r) = \max\{\psi(r), r^{-\mu-1/(m+n)}\}, \quad r = 2, 3, \ldots,$$

for some $\mu > 0$. Clearly the series

(5.2)
$$\sum_{r=1}^{\infty} \psi^*(r)^{m+n}$$

is convergent if and only if the series (1.4) is convergent. In addition,

$$r^{-1}\psi^*(r)^{-m-n+1} \leqslant r^{\mu(m+n-1)-1/(m+n)},$$

for all $r \geqslant 2$, and hence if we choose $\mu < (m+n)^{-1}(m+n-1)^{-1}$, then

$$|C(q; \psi^*)| \ll \psi^*(q)^{m+n},$$

for all sufficiently large q. It now follows from the Borel–Cantelli lemma that the system of inequalities

$$\|q\xi_i(\boldsymbol{u})\| < \psi^*(q), \quad i = 1, \ldots, m+n,$$

has at most finitely many solutions for almost all $\boldsymbol{u} \in \Omega$, which proves Theorem 1.3.

References

[1] R. C. Baker, *Metric Diophantine approximation on manifolds*, J. London Math. Soc. 14 (1976), 43–48.
[2] V. I. Bernik, *The asymptotic behaviour of the number of solutions of some systems of inequalities in the theory of Diophantine approximations of dependent quantities* (Russian), Vestsi Akad. Navuk BSSR Ser. Fiz.-Mat. Navuk 1973 (1), 10–17.
[3] — *An analogue of Khintchine's theorem in the metric theory of Diophantine approximations of dependent quantities* (Russian), ibid. 1977 (6), 44–49.
[4] — *Applications of measure theory and Hausdorff dimension to the theory of Diophantine approximation*, in *New Advances in Transcendence Theory* (ed. A. Baker), Cambridge University Press, 1988.
[5] J. W. S. Cassels, *An Introduction to Diophantine Approximation*, Cambridge University Press, 1957.
[6] M. M. Dodson, B. P. Rynne and J. A. G. Vickers, *Metric Diophantine approximation and Hausdorff dimension on manifolds*, Math. Proc. Cambridge Philos. Soc. 105 (1989), 547–558.
[7] — — — *Averaging in multifrequency systems*, Nonlinearity 2 (1989), 137–148.
[8] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry II*, Interscience, New York 1969.
[9] V. G. Sprindžuk, *Metric Theory of Diophantine Approximations* (Engl. trans. R. A. Silverman), Wiley, New York 1979.
[10] — *Achievements and problems in Diophantine approximation theory*, Russian Math. Surveys 35 (1980), 1–80.
[11] E. C. Titchmarsh, *The Theory of the Riemann Zeta Function*, Oxford University Press, London 1951.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF YORK
York YO1 5DD, England
FACULTY OF MATHEMATICAL STUDIES
UNIVERSITY OF SOUTHAMPTON
Southampton S09 5NH, England

Received on 9.5.1989 (1937)

ACTA ARITHMETICA
LVII (1991)

Representation of primes by the principal form of discriminant $-D$ when the classnumber $h(-D)$ is 3

by

KENNETH S. WILLIAMS* (Ottawa, Ont.) and
RICHARD H. HUDSON (Columbia, S.C.)

0. Notation and preliminary result. Throughout this paper p denotes a prime > 3. We shall be concerned with binary quadratic forms $ax^2+bxy+cy^2$, written (a, b, c), which are integral (that is, a, b, c are integers), positive-definite (that is, $a > 0$, $b^2-4ac < 0$) and primitive (that is, $\mathrm{GCD}(a, b, c) = 1$). The discriminant of the form (a, b, c) is the negative integer b^2-4ac. On the set of all such forms of fixed discriminant $-D$ $(D > 0)$, we define an equivalence relation $\sim$ as follows: we write $(a, b, c) \sim (a', b', c')$ if there exist integers p, q, r, s with $ps-qr = +1$ such that

$$a(px+qy)^2+b(px+qy)(rx+sy)+c(rx+sy)^2 = a'x^2+b'xy+c'y^2.$$

It is well known that there are only finitely many such equivalence classes. The number of classes is called the classnumber of forms of discriminant $-D$ and is denoted by $h(-D)$. The principal form of discriminant $-D$ is the form p_{-D} given by

$$(0.1) \qquad p_{-D} = \begin{cases} (1, 0, D/4), & \text{if } D \equiv 0 \pmod 4, \\ (1, 1, (D+1)/4), & \text{if } D \equiv 3 \pmod 4. \end{cases}$$

A positive integer m is said to be represented by the form (a, b, c) if there exist integers x and y such that $m = ax^2+bxy+cy^2$. If the prime p (not dividing $2D$) is represented by a form of discriminant $-D$, it is well known that the Legendre symbol $\left(\frac{-D}{p}\right) = +1$. In this paper we shall be concerned with the representability of a prime p (> 3) by the principal form p_{-D} of discriminant $-D$ when $h(-D) = 3$.

Recent deep work of Goldfeld, Gross, Mestre, Oesterlé and Zagier (see [6], [7], [12], [13], [14], [20]) has led to the complete determination of all the imaginary quadratic fields with classnumber 3 [12: Théorème 4], namely,

* Research supported by Natural Sciences and Engineering Research Council of Canada Grant A-7233.