Analytic Methods for the Distribution of Rational Points on Algebraic Varieties

D.R. Heath-Brown
Mathematical Institute, Oxford

Lecture 1 — Introduction to the Hardy-Littlewood Circle Method

The most important analytic method for handling equidistribution questions about rational points on algebraic varieties is undoubtedly the Hardy-Littlewood circle method. There are a number of good texts available on the circle method, but the reader may particularly wish to study the books by Davenport [4] and Vaughan [11].

In this lecture we shall consider an irreducible form $F(X_1, \dots, X_n) \in \mathbb{Z}[X_1, \dots, X_n]$ of degree d which defines a hypersurface $F = 0$ in $\mathbb{P}^{n-1}$. The fundamental questions will be: are there any rational points on the hypersurface? If so, are they equidistributed in a suitable sense? We shall address these issues by looking at integer points on the affine cone. Thus to ask about equidistribution with respect to measures on $\mathbb{P}^{n-1}(\mathbb{R})$, for example, we may take a small box

$$\mathcal{R} := \prod_{i=1}^{n} (\kappa_i, \lambda_i] \subset \mathbb{R}^n,$$

in which

$$\kappa_i \lambda_i > 0, \quad (1 \le i \le n),$$

and examine the asymptotic behaviour of

$$N_{\mathcal{R}}(B) := \#\{\mathbf{x} \in \mathbb{Z}^n \cap B\mathcal{R} : \ F(\mathbf{x}) = 0\}$$

as $B \to \infty$. Here $B\mathcal{R} := \{\mathbf{x} : B^{-1}\mathbf{x} \in \mathcal{R}\}$. If we were to take a number of small boxes $\mathcal{R}_i$, of equal volumes, and we discovered that $N_{\mathcal{R}_i}(B)$ were asymptotically equal for each box, then we could say that the solutions of $F(\mathbf{x}) = 0$ were equidistributed amongst the boxes. Of course only boxes containing real solutions of $F(\mathbf{x}) = 0$ can feature, and indeed one finds in

practice that points on the real locus $F = 0$ need to be appropriately weighted in order to achieve equidistribution. In other words what one attempts to establish is the equidistribution of integer solutions to $F(\mathbf{x}) = 0$, with respect to a suitable measure, on the real locus $F(\mathbf{x}) = 0$, and results on the behaviour of $N_{\mathcal{R}}(B)$ should be interpreted in this light.

Notice that the function $N_{\mathcal{R}}(B)$ may count a given projective point several times, with $\mathbf{x}$ being different scalar multiples of the same vector. However it is usually a trivial matter to adjust for this. In suitable circumstances one finds that

$$N_{\mathcal{R}}(B) \sim c(\mathcal{R})B^{n-d}$$

as $B \to \infty$, for a certain constant $c(\mathcal{R})$, and the equidistribution problem then becomes one of the behaviour of $c(\mathcal{R})$ as the box $\mathcal{R}$ varies.

The second aspect of equidistribution which we shall consider concerns the distribution of rational points within the p-adics, or more generally in the adèles. For this it is sufficient to select any vector $\mathbf{a} \in \mathbb{Z}^n$ and any modulus $m \in \mathbb{N}$ such that

$$(m, a_1, \dots, a_n) = 1,$$

and investigate the modified counting function

$$N_{\mathcal{R}}(B; m, \mathbf{a}) := \#\{\mathbf{x} \in \mathbb{Z}^n \cap B\mathcal{R} : \mathbf{x} \equiv \mathbf{a} \pmod{m}, F(\mathbf{x}) = 0\}.$$

If m is composed of high powers of sufficiently many primes, then the behaviour of $N_{\mathcal{R}}(B; m, \mathbf{a})$ as both $\mathcal{R}$ and $\mathbf{a}$ vary is enough to describe completely the adèlic distribution of rational points on $F = 0$. Indeed the relative size of $N_{\mathcal{R}}(B; m, \mathbf{a})$ as $\mathbf{a}$ varies, describes the equidistribution of rational points, or lack of it, within the adéles.

Before proceeding further we should point out that although we have taken the groundfield as $\mathbb{Q}$ above, it is quite possible to consider problems over an arbitrary number field. The reader is referred to the book by Wang [13] for the circle method in this context.

Just what asymptotic behaviour should we expect for $N_{\mathcal{R}}(B)$? Since F is a form of degree d, we will have $F(\mathbf{x}) = O(B^d)$ for $\mathbf{x} \in B\mathcal{R}$. Hence the 'probability' that a particular integer value of $F(\mathbf{x})$ vanishes should be around B^{-d}. The number of integer vectors $\mathbf{x}$ in the region $B\mathcal{R}$ is of order B^n, so one might suppose that $N_{\mathcal{R}}(B)$ is likely to be of order precisely B^{n-d}. Of course this can only be meaningful when $n > d$. Moreover it can happen that the region $B\mathcal{R}$ does not even contain a real solution $\mathbf{x} \in \mathbb{R}^n$ of the equation $F(\mathbf{x}) = 0$. None the less we can state the following guiding principle.

Heuristic Principle *When $n > d$ we have*

$$B^{n-d} \ll N_{\mathcal{R}}(B) \ll B^{n-d}$$

unless there is a reason why not!

We should perhaps explain here how the Vinogradov $\ll$ notation is used. The statement $f \ll g$ means that there exists a positive "constant" c, such that $|f| \leq cg$ for all values under consideration. It follows that the statement $0 \ll g$ is somewhat uninformative, holding for any non-negative g ! When we say that c is "constant" we mean that it does not depend on any variable, but may depend on parameters which are regarded as fixed. Thus, in our heuristic principle the implied constants are independent of B, but might depend on $F, \mathcal{R}, n$ and d.

The starting point for the Hardy-Littlewood circle method is the generating function

$$S(\alpha; B\mathcal{R}) = S(\alpha) := \sum_{\mathbf{x} \in \mathbb{Z}^n \cap B\mathcal{R}} \exp\{2\pi i \alpha F(\mathbf{x})\},$$

in which we take α to be a real variable, usually on $[0, 1]$. The function $\exp\{2\pi i\alpha\}$ occurs so frequently that we introduce the standard notation $e(\alpha) := \exp\{2\pi i\alpha\}$. Thus we may think of $e(.)$ as being a character on $\mathbb{R}^+/\mathbb{Z}^+$. We now have the key fact that if $n \in \mathbb{Z}$ then

$$\int_0^1 e(\alpha n) d\alpha = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases} \tag{1}$$

Applying this to the generating function $S(\alpha)$ produces the formula

$$N_{\mathcal{R}}(B) = \int_0^1 S(\alpha) d\alpha. \tag{2}$$

What should we expect about the size of $S(\alpha)$? Let us suppose that the intervals $[\kappa_i, \lambda_i]$ defining $\mathcal{R}$ satisfy $-1 \leq \kappa_i < \lambda_i \leq 1$. This can always be achieved by re-scaling. Moreover let us set

$$V(\mathcal{R}) := \prod_{i=1}^{n} (\lambda_i - \kappa_i). \tag{3}$$

Then if we write $\mathcal{N}(B)$ for the number of integer vectors in $B\mathcal{R}$ we see that

$$\mathcal{N}(B) = \{V(\mathcal{R}) + o(1)\} B^n.$$

Now if $0 < \alpha < 1$ we might expect the numbers $e(\alpha F(\mathbf{x}))$ to be randomly scattered around the unit circle. Indeed if they are sufficiently random we might expect, from the central limit theorem, that $S(\alpha)$ should have a normal

distribution in some sense. This leads us to expect that $S(\alpha)$ should usually be of order not much more than $B^{n/2}$. In other words, we expect that the terms in $S(\alpha)$ cancel to exactly the extent predicted by crude probabilistic arguments.

In contrast, when α is near to 0 (or to 1) we certainly can not expect such cancellation. To make this precise, let us write $||F||$ for the sum of the moduli of the coefficients of F, so that $|F(\mathbf{x})| \leq ||F||B^d$ for $\mathbf{x} \in B\mathcal{R}$. Then if

$$\alpha \in \mathcal{J} := [-(8||F||B^d)^{-1}\,,\,(8||F||B^d)^{-1}],$$

say, we find that $|\alpha F(\mathbf{x})| \leq 1/8$ for $\mathbf{x} \in B\mathcal{R}$. Thus

$$\frac{1}{\sqrt{2}} = \cos(\frac{\pi}{4}) \leq \Re\{e(\alpha F(\mathbf{x}))\} \leq 1,$$

and hence

$$\frac{1}{\sqrt{2}}\mathcal{N}(B) \leq \Re\{S(\alpha)\} \leq \mathcal{N}(B).$$

Thus if we write

$$N_{\mathcal{R}}(B;\mathcal{J}) = \int_{\mathcal{J}} S(\alpha)d\alpha$$

for that part of the integral (2) in which α is suitably small, we find that

$$\{c_1 + o(1)\}B^{n-d} \leq N_{\mathcal{R}}(B;\mathcal{J}) \leq \{c_2 + o(1)\}B^{n-d}, \tag{4}$$

with $c_1 = V(\mathcal{R})/(4||F||\sqrt{2})$ and $c_2 = V(\mathcal{R})/(4||F||)$.

The precise values of c_1 and c_2 are unimportant, but there are several significant observations to be made. Firstly we see that the integration for α near 0 produces a term which is of exactly the order B^{n-d} we previously predicted heuristically. Conversely, if $N_{\mathcal{R}}(B)$ does not behave as the naive heuristics predict, so that it is not of exact order B^{n-d}, then the circle method is almost certainly doomed to fail. Thirdly, in an ideal world we may hope that the range in which α is not close to the endpoints of the interval $[0,1]$ will only make a contribution of order about $B^{n/2}$. Of course, this will be smaller than the expected main term when $n > 2d$. Thus we might hope to prove that $N_{\mathcal{R}}(B)$ is of order B^{n-d}, provided that $n > 2d$. The final point is more technical, and is a converse to the third. In view of the central limit theorem we expect that $|S(\alpha)|$ is usually exactly of order $B^{n/2}$. Thus we cannot hope to prove results for the case $n \leq 2d$ unless there is demonstrable cancellation in the values of $S(\alpha)$ as we average over α.

In any event, the basic message is that the method can only work, in its simplest form, if the number of variables exceeds $2d$. Moreover, in order to

succeed we must prove that the sum $S(\alpha)$ cancels in the way that the central limit theorem predicts.

It is now time to admit that our initial distinction between values α near 0 or 1, and values in the remainder of the range was far too simple. It is quite possible that the coefficients of F are all even, for example, in which case we will have the same behaviour around $\alpha = 1/2$ as we previously saw near $\alpha = 0$. More generally, it will often be the case that the values of $F(\mathbf{x})$ are not uniformly distributed to some particular modulus q, in which case $S(\alpha)$ may be large when α is close to a fraction of the form a/q. To make this more precise, we define

$$S_q(a) := \sum_{\mathbf{x} (\mathrm{mod}\ q)} e(\frac{a}{q} F(\mathbf{x})),$$

where the summation condition means that $\mathbf{x}$ runs over $\mathbb{Z}^n/q\mathbb{Z}^n$, or more concretely that $\mathbf{x}$ runs over integer vectors for which

$$0 \le x_i \le q-1, \quad (1 \le i \le n).$$

Clearly $e(aF(\mathbf{x})/q) = e(aF(\mathbf{y})/q)$ whenever $\mathbf{x} \equiv \mathbf{y} (\mathrm{mod}\ q)$, so that

$$S(\frac{a}{q}) = \sum_{\mathbf{x} (\mathrm{mod}\ q)} e(\frac{a}{q} F(\mathbf{x})) \#\{\mathbf{y} \in B\mathcal{R} : \mathbf{y} \equiv \mathbf{x}\ (\mathrm{mod}\ q)\}.$$

Moreover

$$\#\{y \in (Ba, Bb] : y \equiv x\ (\mathrm{mod}\ q)\} = (b-a)\frac{B}{q} + O(1),$$

whence

$$\#\{\mathbf{y} \in B\mathcal{R} : \mathbf{y} \equiv \mathbf{x}\ (\mathrm{mod}\ q)\} = V(\mathcal{R})\frac{B^n}{q^n} + O(B^{n-1}),$$

with $V(\mathcal{R})$ given by (3). It follows that

$$\begin{aligned} S(\frac{a}{q}) &= \sum_{\mathbf{x} (\mathrm{mod}\ q)} e(\frac{a}{q} F(\mathbf{x})) V(\mathcal{R})\frac{B^n}{q^n} + \sum_{\mathbf{x} (\mathrm{mod}\ q)} O(B^{n-1}) \\ &= V(\mathcal{R})\frac{B^n}{q^n} S_q(a) + O(q^n B^{n-1}). \end{aligned}$$

Thus, for any fixed fraction a/q, we see that if $S_q(a)$ is non-zero, then $S(a/q)$ will be of exact order B^n. Hence indeed $S(\alpha)$ will also be of exact order B^n when α is close to a/q.

It follows, via the analysis which lead to (4), that values of α which are close to any fraction of small denominator may make contributions of the size predicted by our original heuristic argument. Thus it is appropriate to split the range $[0,1]$ into two subsets, the "major arcs" in which α is close to a rational number with small denominator, in some precisely defined sense, and the remaining set, known as the "minor arcs". This terminology stems from the original formulation of the circle method, in which one integrated around a circle in the complex plane, instead of using the interval $[0,1]$.

The reader will observe that everything we have described carries over immediately to $N_{\mathcal{R}}(B;m,\mathbf{a})$. All that is necessary is to work with the modified generating function

$$S(\alpha;B\mathcal{R};m,\mathbf{a}) := \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap B\mathcal{R}\\ \mathbf{x}\equiv\mathbf{a}(\mathrm{mod}\ m)}} \exp\{2\pi i\alpha F(\mathbf{x})\}.$$

Having summarised the key features of the circle method we complete this lecture with a review of results. Statements in the literature are formulated in a variety of different ways. For example, some authors work only with $\mathcal{R} = [-1,1]^n$; most do not consider the condition $\mathbf{x} \equiv \mathbf{a} \pmod m$; and sometimes one is content with establishing the existence of solutions rather than deriving an asymptotic formula for $N_{\mathcal{R}}(B)$. Thus the references we shall give will usually not contain the precise assertions we present below.

We begin with the case $d = 2$, so that F is a quadratic form. Here one has the following basic result, which we shall prove in the next two lectures.

Theorem 1 *Let F be a diagonal quadratic form. If $n \geq 5$ there is a constant $c = c(\mathcal{R};m,\mathbf{a})$ such that*

$$N_{\mathcal{R}}(B;m,\mathbf{a}) = cB^{n-2} + o(B^{n-2})$$

as $B \to \infty$.

We shall see a full interpretation of the constant c in the next lecture, but we should say at once that $c = 0$ if and only if the local conditions on $\mathbf{x}$ imply that $F(\mathbf{x})$ never vanishes.

Notice that the condition $n \geq 5$ corresponds to the inequality $n > 2d$ which we encountered earlier. In fact one can do a little better for quadratics, using a special variant of the circle method (see Heath-Brown [7]), and show that

$$N_{\mathcal{R}}(B;m,\mathbf{a}) = cB^2 + o(B^2)$$

when $n = 4$ and the determinant of the quadratic form is not a square. However if the determinant is a square one gets a different behaviour

$$N_{\mathcal{R}}(B;m,\mathbf{a}) = cB^2\log B + o(B^2\log B).$$

One can even handle forms with $n = 3$, in which case one has

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = cB \log B + o(B \log B).$$

We now turn to forms of degree 3. If $F(\mathbf{x})$ is an "additive" or "diagonal" form, of the shape

$$F(\mathbf{x}) = c_1 x_1^3 + \ldots + c_n x_n^3,$$

the analysis can be pushed rather further than in the general case, and one has

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = cB^{n-3} + o(B^{n-3})$$

as soon as $n \geq 8$. This follows from the methods of Vaughan [12]. In this case we are not yet able to handle values of n as small as $n = 2d + 1$.

For non-diagonal cubic forms things are harder, but Hooley [10] has shown essentially that

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = cB^{n-3} + o(B^{n-3})$$

for $n \geq 9$, providing that F is non-singular. Finally, for arbitrary cubic forms our basic heuristic may fail completely, as the example

$$F(X_1, \ldots, X_n) = X_1^3 + X_2(X_3^2 + \ldots + X_n^2)$$

shows. Here one has $F(\mathbf{x}) = 0$ whenever $x_1 = x_2 = 0$. Thus one has $N_{\mathcal{R}}(B) \gg B^{n-2}$ for suitable regions $\mathcal{R}$.

For general degrees the diagonal case is again easier. Here one can show

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = cB^{n-d} + o(B^{n-d})$$

for $n \geq n_1(d)$ with a value of $n_1(d)$ such that

$$n_1(d) \sim d^2 \log d$$

by the method of Ford [6]. For general non-singular forms one can establish the same asymptotic relation for $N_{\mathcal{R}}(B; m, \mathbf{a})$ by the method of Birch [1], under the condition $n \geq n_2(d)$, where $n_2(d) = 1 + (d-1)2^d$.

The overall picture is thus that the circle method only works if the number of variables is sufficiently large in terms of the degree, where "sufficiently large" refers to a function which tends to infinity with d. None the less, where the method does work it provides a complete answer to the equidistribution question. For singular varieties the heuristic formula may fail altogether.

We end this lecture by remarking that the techniques described here can be extended to varieties defined by the vanishing of a system of forms $F_1(\mathbf{x}) =$

$\ldots = F_k(\mathbf{x}) = 0$. All that is necessary is to define a multi-variable generating function

$$S(\alpha_1, \ldots, \alpha_k) := \sum_{\mathbf{x} \in \mathbb{Z}^n \cap B\mathcal{R}} \exp\{2\pi i(\alpha_1 F_1(\mathbf{x}) + \ldots + \alpha_k F_k(\mathbf{x}))\},$$

and integrate over $[0,1]^k$. Examples of results which can be proved in this way may be found in the works of Davenport and Lewis [3], for diagonal forms, and Birch [1] for general forms.

Lecture 2 — Major Arcs and Local Factors in the Hardy-Littlewood Circle Method

In this lecture we shall work with the integral

$$N_{\mathcal{R}}(B; m, \mathbf{a}) = \int_0^1 S(\alpha; B\mathcal{R}; m, \mathbf{a}) d\alpha \tag{5}$$

analogous to (2), and consider the contribution coming from α in the "major arcs". For the time being we shall merely assume that $F(X_1, \dots, X_n) \in \mathbb{Z}[X_1, \dots, X_n]$ is a form of degree d in n variables. Later we shall have to impose further restrictions on F and d.

Clearly it is time to be more specific about what we mean by the "major arcs". We shall use a small positive parameter $\delta < 1/3$, independent of B, which we shall specify later, see (11). It will be convenient to write $Q := B^\delta$. For each positive integer $q \le Q$, and each non-negative integer $a < q$ such that $(a, q) = 1$, we proceed to define the interval

$$I(a, q) := \left[\frac{a}{q} - \frac{Q}{B^d} \, , \, \frac{a}{q} + \frac{Q}{B^d} \right].$$

We now observe that these various intervals are disjoint providing that B is large enough, as we henceforth assume. For if $a/q \neq a'/q'$ then

$$\left| \frac{a}{q} - \frac{a'}{q'} \right| \ge \frac{1}{qq'} \ge \frac{1}{Q^2} > 2\frac{Q}{B^d}.$$

With this in mind we define the set of "major arcs" as

$$\mathfrak{M} = \bigcup_{q \le Q} \bigcup_{\substack{0 \le a < q \\ (a,q)=1}} I(a, q).$$

Strictly speaking we do not have $\mathfrak{M} \subseteq [0, 1]$ since the left endpoint of $I(0, 1)$ is negative. Thus we will replace the range of integration in (5) by the interval $[-Q/B^d \, , \, 1 - Q/B^d]$ to avoid this problem. We then define the "minor arcs" to be the set

$$\mathfrak{m} = \left[-\frac{Q}{B^d} \, , \, 1 - \frac{Q}{B^d} \right] \setminus \mathfrak{M}.$$

It should be pointed out at this stage that there is considerable flexibility in choosing the major arc set, and for other applications a rather different choice may be appropriate. Indeed in some cases one chooses the major arcs so as to comprise the entire interval $[0, 1]$. The key points to note about our

choice (indeed any choice) for the set $\mathfrak{M}$, are firstly that the intervals $I(a,q)$ are longer than B^{-d} by a factor which tends to infinity; and secondly that one uses all denominators up to a bound Q which also tends to infinity. The reader will note that the description given in the first lecture fits neither of these conditions!

We now face the crucial task of approximating $S(\alpha) = S(\alpha; B\mathcal{R}; m, \mathbf{a})$ on the various intervals $I(a,q)$. We will certainly need more precise estimates than we obtained in the first lecture. It will be convenient to put $\theta = \alpha - a/q$. Then if $\mathbf{y} \equiv \mathbf{x} (\mathrm{mod}\ q)$ we have

$$e(\alpha F(\mathbf{y})) = e(\frac{a}{q}F(\mathbf{y}))e(\theta F(\mathbf{y})) = e(\frac{a}{q}F(\mathbf{x}))e(\theta F(\mathbf{y})).$$

In fact we also have the congruence condition $\mathbf{x} \equiv \mathbf{a}\,(\mathrm{mod}\ m)$ to contend with. The reader may prefer to take $m = 1$ in what follows, but we shall give the details for the general case. We have

$$\begin{aligned} S(\alpha; B\mathcal{R}; m, \mathbf{a}) &= \sum_{\substack{\mathbf{x} (\mathrm{mod}\ mq) \\ \mathbf{x} \equiv \mathbf{a} (\mathrm{mod}\ m)}} \sum_{\substack{\mathbf{y} \in B\mathcal{R} \\ \mathbf{y} \equiv \mathbf{x} (\mathrm{mod}\ mq)}} e(\alpha F(\mathbf{y})) \\ &= \sum_{\substack{\mathbf{x} (\mathrm{mod}\ mq) \\ \mathbf{x} \equiv \mathbf{a} (\mathrm{mod}\ m)}} e(\frac{a}{q}F(\mathbf{x})) \sum_{\substack{\mathbf{y} \in B\mathcal{R} \\ \mathbf{y} \equiv \mathbf{x} (\mathrm{mod}\ mq)}} e(\theta F(\mathbf{y})). \qquad (6) \end{aligned}$$

(Here, and elsewhere in these notes, we shall follow the convention that summations are over integers, and integer vectors, without explicitly writing $\mathbf{y} \in \mathbb{Z}^n \cap B\mathcal{R}$, for example.) The inner sum above contains an exponential in which $\theta F(\mathbf{y})$ is relatively small, of order $Q = B^{\delta}$, and smoothly varying. We can therefore hope to approximate the sum fairly well by the corresponding integral.

In fact we have

$$\sum_{\substack{\mathbf{y} \in B\mathcal{R} \\ \mathbf{y} \equiv \mathbf{x} (\mathrm{mod}\ mq)}} e(\theta F(\mathbf{y})) = (mq)^{-n} I(B; \theta) + O(B^{n-1+2\delta}), \qquad (7)$$

where

$$I(B; \theta) := \int_{B\mathcal{R}} e(\theta F(\mathbf{z})) dz_1 \dots dz_n. \qquad (8)$$

We should stress at this point that, in writing the above error term, we have taken the form F and the box $\mathcal{R}$, along with $m, \mathbf{a}$ and the parameter δ, to be fixed. Thus the error term is of the shape $\rho B^{n-1+2\delta}$, where ρ is bounded in terms of B, but not necessarily as a function of $F, \mathcal{R}, m, \mathbf{a}$ and δ. This remark will apply to all the error terms we shall meet henceforth.

The proof of (7) is somewhat technical, and might reasonably be skipped on a first reading. All that is important is the form of the main term, and the fact that the error term contains an exponent which is strictly less than n. However we present below the necessary argument.

For each relevant integer point $\mathbf{y}$ we define

$$\mathcal{C}(\mathbf{y}) = \{\mathbf{z} \in \mathbb{R}^n : y_i < z_i \leq y_i + mq, \text{ for } 1 \leq i \leq n\}.$$

Then if $\mathbf{z} \in \mathcal{C}(\mathbf{y})$ we have $F(\mathbf{z}) - F(\mathbf{y}) = O(qB^{d-1})$, since $|z_i - y_i| \leq mq \leq mQ \leq mB$ for each index i, and m is regarded as fixed. It follows that $\theta(F(\mathbf{z}) - F(\mathbf{y})) = O(QB^{-d}.qB^{d-1}) = O(Q^2B^{-1})$, whence $e(\theta F(\mathbf{z}) - \theta F(\mathbf{y})) = 1 + O(Q^2B^{-1})$ and thus

$$e(\theta F(\mathbf{z})) = e(\theta F(\mathbf{y})) + O(B^{-1+2\delta}).$$

If we integrate over $\mathcal{C}(\mathbf{y})$ we may then deduce that

$$\int_{\mathcal{C}(\mathbf{y})} e(\theta F(\mathbf{z}))dz_1 \dots dz_n = (mq)^n e(\theta F(\mathbf{y})) + O((mq)^n B^{-1+2\delta}),$$

so that

$$e(\theta F(\mathbf{y})) = (mq)^{-n} \int_{\mathcal{C}(\mathbf{y})} e(\theta F(\mathbf{z}))dz_1 \dots dz_n + O(B^{-1+2\delta}).$$

We have constructed the cubes $\mathcal{C}(\mathbf{y})$ so as to be disjoint. Moreover the box $B\mathcal{R}$ contains $O(B^n)$ integer points $\mathbf{y}$. Hence, when we sum over the various values of $\mathbf{y}$ we find that

$$\sum_{\substack{\mathbf{y} \in B\mathcal{R} \\ \mathbf{y} \equiv \mathbf{x} (\mathrm{mod}\ mq)}} e(\theta F(\mathbf{y})) = (mq)^{-n} \int_{\mathcal{C}} e(\theta F(\mathbf{z}))dz_1 \dots dz_n + O(B^{n-1+2\delta}),$$

where $\mathcal{C}$ is the union of the cubes $\mathcal{C}(\mathbf{y})$. If we have a point $\mathbf{z}$ which belongs to one of $B\mathcal{R}$ or $\mathcal{C}$ but not to the other, then $\mathbf{z}$ must lie at a distance at most $O(mq) = O(q)$ from the boundary of $B\mathcal{R}$. Since $q \leq B$ it follows that

$$\int_{\mathcal{C}} e(\theta F(\mathbf{z}))dz_1 \dots dz_n = \int_{B\mathcal{R}} e(\theta F(\mathbf{z}))dz_1 \dots dz_n + O(qB^{n-1}).$$

On comparing our various estimates we therefore derive the required approximation (7).

We now return to the main thread of the argument by inserting (7) into (6) to deduce that

$$S(\alpha; B\mathcal{R}; m, \mathbf{a}) = \frac{S_q(a; m, \mathbf{a})}{(mq)^n} I(B; \theta) + O(q^n B^{n-1+2\delta}),$$

where we have defined

$$S_q(a;m,\mathbf{a}) := \sum_{\substack{\mathbf{x} (\mathrm{mod}\ mq)\\ \mathbf{x}\equiv\mathbf{a} (\mathrm{mod}\ m)}} e(\frac{a}{q}F(\mathbf{x})).$$

We proceed to integrate over $I(a,q)$, producing

$$\int_{I(a,q)} S(\alpha; B\mathcal{R}; m, \mathbf{a})d\alpha = \frac{S_q(a;m,\mathbf{a})}{(mq)^n}J(B) + O(Q^{n+1}B^{n-d-1+2\delta})$$

with

$$J(B) := \int_{-QB^{-d}}^{QB^{-d}} I(B,\theta)d\theta. \tag{9}$$

We can then sum for a coprime to q and for $q \le Q$ to yield

$$\begin{aligned}\int_{\mathfrak{M}} S(\alpha; B\mathcal{R}; m, \mathbf{a})d\alpha &= S(Q)J(B) + O(Q^{n+3}B^{n-d-1+2\delta})\\ &= S(Q)J(B) + O(B^{n-d-1+(n+5)\delta}),\end{aligned} \tag{10}$$

where

$$S(Q) := \sum_{q\le Q}\sum_{\substack{0\le a<q\\(a,q)=1}} \frac{S_q(a;m,\mathbf{a})}{(mq)^n}.$$

The error term here is $o(B^{n-d})$, which is satisfactory for Theorem 1, providing that we choose

$$\delta = \frac{1}{n+6}, \tag{11}$$

as we shall indeed do.

Thus to complete the treatment of the major arcs we need to say more about the sum $S(Q)$ and the integral $J(B)$. We begin with the latter. By substituting $\mathbf{z} = B\mathbf{w}$ in (8) and $\theta = B^{-d}\phi$ in (9) we find that

$$J(B) = B^{n-d}\int_{-Q}^{Q}\{\int_{\mathcal{R}} e(\phi F(\mathbf{w}))dw_1\dots dw_n\}d\phi.$$

We now impose an additional technical requirement on the form F, namely that

$$F_1(w_1,\dots,w_n) > 0 \quad \forall \mathbf{w}\in\mathcal{R}, \tag{12}$$

where

$$F_1(w_1,\dots,w_n) := \frac{\partial F(w_1,\dots,w_n)}{\partial w_1}.$$

(With a little more effort one can analyse $J(B)$ satisfactorily under the more natural assumption that F is non-singular in $\mathcal{R}$.) We can then solve the equation $u = F(w_1, \dots, w_n)$ for w_1, using the Implicit Function Theorem. This produces a function $w_1 = f(u, w_2, \dots, w_n)$, say. We proceed to define

$$g(u) := \int \frac{dw_2 \dots dw_n}{F_1(f(u, w_2, \dots, w_n), w_2, \dots, w_n)},$$

where the integration is subject to $(w_1, \dots, w_n) \in \mathcal{R}$. The function g is continuously differentiable, and has compact support. (The first of these two assertions requires more than a moment's thought.) With these definitions we have

$$\int_{\mathcal{R}} e(\phi F(\mathbf{w}))dw_1 \dots dw_n = \int_{-\infty}^{\infty} e(\phi u)g(u)du = \hat{g}(-\phi),$$

where $\hat{g}$ denotes the Fourier transform. It then follows that

$$\lim_{B\to\infty} \int_{-B^\delta}^{B^\delta} \{\int_{\mathcal{R}} e(\phi F(\mathbf{w}))dw_1 \dots dw_n\}d\phi = \lim_{T\to\infty} \int_{-T}^{T} \hat{g}(-\phi)d\phi = g(0),$$

whence

$$J(B) \sim B^{n-d}c_\infty \quad (B \to \infty) \tag{13}$$

where c_∞ is the so called "singular integral", given by

$$c_\infty := g(0) = \int_{\mathbf{w}\in\mathcal{R},\, F(\mathbf{w})=0} \frac{dw_2 \dots dw_n}{F_1(w_1, \dots, w_n)},$$

with a certain amount of abuse of notation. The singular integral is to be regarded as the density of real points of $F = 0$ in the box $\mathcal{R}$. In fact one can show, subject only to the condition that F is non-singular in $\mathcal{R}$, that

$$c_\infty = \lim_{\varepsilon\downarrow 0} \frac{1}{2\varepsilon} \mathrm{Meas}\{\mathbf{w} \in \mathcal{R} : |F(\mathbf{w})| < \varepsilon\}.$$

We end our discussion of $J(B)$ by remarking that (13) shows in particular that $J(B) = O(B^{n-d})$.

We turn now to the sum $S(Q)$. Just as $J(B)$ reflects the behaviour of F with respect to the real valuation $|.|_\infty$, we shall see that $S(Q)$ contains the corresponding p-adic information. Our treatment will require the following lemma, which we will use to separate out the factors corresponding to individual primes.

Lemma 1 *Let $q = rs$ and $m = jk$, with rj coprime to sk. Choose integers $\overline{r}$ and $\overline{r}$ such that $r\overline{r} \equiv 1 (\mathrm{mod}\ sk)$ and $s\overline{s} \equiv 1 (\mathrm{mod}\ rj)$. Then*

$$S_q(a; m, \mathbf{a}) = S_r(a\overline{s}; j, \mathbf{a}) S_s(a\overline{r}; k, \mathbf{a}).$$

For the purposes of the proof we choose integers $\overline{j}$ and $\overline{k}$ with $j\overline{j} \equiv 1 (\mathrm{mod}\ sk)$ and $k\overline{k} \equiv 1 (\mathrm{mod}\ rj)$. One can then verify that if $\mathbf{u}$ runs over vectors modulo rj with $\mathbf{u} \equiv \mathbf{a} \pmod{j}$, and $\mathbf{v}$ runs over vectors modulo sk with $\mathbf{v} \equiv \mathbf{a} \pmod{k}$, then

$$\mathbf{w} := sk\overline{s}\overline{k}\mathbf{u} + rj\overline{r}\overline{j}\mathbf{v}$$

runs over vectors modulo qm with $\mathbf{w} \equiv \mathbf{a} (\mathrm{mod}\ m)$. Moreover we have

$$F(\mathbf{w}) \equiv (sk\overline{s}\overline{k})^d F(\mathbf{u}) + (rj\overline{r}\overline{j})^d F(\mathbf{v}) \equiv s\overline{s}F(\mathbf{u}) + r\overline{r}F(\mathbf{v}) \pmod{q},$$

whence

$$e(\frac{a}{q}F(\mathbf{w})) = e(\frac{a\overline{s}}{r}F(\mathbf{u}))e(\frac{a\overline{r}}{s}F(\mathbf{v})).$$

The lemma then follows on summing over $\mathbf{u}$ and $\mathbf{v}$.

We now examine the sum

$$S(q, m, \mathbf{a}) := \sum_{\substack{0 \le a < q \\ (a,q)=1}} S_q(a; m, \mathbf{a})$$

occurring in the definition of $S(Q)$. With the above notation we note that if b and c run over distinct residue classes coprime to r and s respectively, then $a = sb + rc$ will run over distinct residue classes coprime to rs. Then Lemma 1 implies that

$$\begin{aligned} S(q, m, \mathbf{a}) &= \sum_{\substack{0 \le b < r \\ (b,r)=1}} \sum_{\substack{0 \le c < s \\ (c,s)=1}} S_q(sb + rc; m, \mathbf{a}) \\ &= \sum_{\substack{0 \le b < r \\ (b,r)=1}} \sum_{\substack{0 \le c < s \\ (c,s)=1}} S_r((sb + rc)\overline{s}; j, \mathbf{a}) S_s((sb + rc)\overline{r}; k, \mathbf{a}) \\ &= \sum_{\substack{0 \le b < r \\ (b,r)=1}} \sum_{\substack{0 \le c < s \\ (c,s)=1}} S_r(b; j, \mathbf{a}) S_s(c; k, \mathbf{a}) \\ &= S(r, j, \mathbf{a}) S(s, k, \mathbf{a}). \end{aligned}$$

It follows that if $q = \prod p^e$ and $m = \prod p^f$ then we have

$$S(q, m, \mathbf{a}) = \prod_p S(p^e, p^f, \mathbf{a}). \tag{14}$$

Thus we can think of $S(Q)$ as being a partial sum corresponding to an Euler product. In what follows we shall follow the above notation, writing $f = \nu_p(m)$ for each prime p, for example.

Our next goal will be to replace $S(Q)$ by the corresponding infinite sum. This will require estimates for the size of $S(q, m, \mathbf{a})$ which may be derived by the methods in the next lecture. Thus, for the time being we content ourselves with working on the following hypothesis.

Hypothesis *There is a real number $\eta > 0$, depending only on n and d, such that*

$$S_q(a; m, \mathbf{a}) \ll q^{n-2-\eta}$$

whenever a and q are coprime.

We should remark that if q is prime, and if F is non-singular modulo q, then one has

$$S_q(a; m, \mathbf{a}) \ll q^{n/2},$$

by Deligne's Riemann Hypothesis for varieties over finite fields [5]. When $n \geq 5$ this result yields good information for composite moduli q which have no square factor, by Lemma 1. However there are serious problems in extending such estimates to prime powers.

Under the hypothesis above we see that $S(q, m, \mathbf{a}) \ll q^{n-1-\eta}$, so that

$$\sum_{q=1}^{\infty} q^{-n} S(q, m, \mathbf{a}) \tag{15}$$

is absolutely convergent. It then follows from (14) that

$$\sum_{q=1}^{\infty} q^{-n} S(q, m, \mathbf{a}) = \prod_p c_p,$$

where

$$c_p := \sum_{e=0}^{\infty} p^{-en} S(p^e, p^f, \mathbf{a})$$

with $f = \nu_p(m)$. Indeed we will have

$$\sum_{q>Q} q^{-n} |S(q, m, \mathbf{a})| \ll Q^{-\eta},$$

whence

$$S(Q) = m^{-n} \prod_p c_p + O(Q^{-\eta}).$$

We can now combine this with (10) and (13) to conclude that

$$\int_{\mathfrak{M}} S(\alpha; B\mathcal{R}; m, \mathbf{a}) d\alpha \sim B^{n-d} m^{-n} c_\infty \prod_p c_p \quad (B \to \infty), \tag{16}$$

subject to the hypothesis above on the size of $S_q(a; m, \mathbf{a})$, and on our previous assumption (12).

The sum (15) is known as the "singular series". The original formulation of the circle method used a function of a complex variable integrated around a circle, and there was a singularity at each point $e(a/q)$ with $q \leq Q$. Thus the series arose from the contributions of these various singularities.

We close this lecture by giving an important interpretation of the factors c_p, which depends on the following lemma.

Lemma 2 *Define*

$$N(p^e, p^f, \mathbf{a}) := \#\{\mathbf{x} \pmod{p^{e+f}} : \mathbf{x} \equiv \mathbf{a} \pmod{p^f},\ p^e | F(\mathbf{x})\}.$$

Then

$$S(p^e, p^f, \mathbf{a}) = p^e N(p^e, p^f, \mathbf{a}) - p^{n+e-1} N(p^{e-1}, p^f, \mathbf{a})$$

for $e \geq 1$.

By definition we have

$$S(p^e, p^f, \mathbf{a}) = \sum_{a=0}^{p^e-1} S_{p^e}(a; p^f, \mathbf{a}) - \sum_{a=0}^{p^{e-1}-1} S_{p^e}(ap; p^f, \mathbf{a}).$$

Moreover we have

$$S_{p^e}(ap; p^f, \mathbf{a}) = p^n S_{p^{e-1}}(a; p^f, \mathbf{a}),$$

since

$$e(\frac{ap}{p^e} F(\mathbf{x})) = e(\frac{a}{p^{e-1}} F(\mathbf{y}))$$

for every $\mathbf{x}$ such that $\mathbf{x} \equiv \mathbf{y} \pmod{p^{e+f-1}}$. To complete the proof of the lemma it we observe that

$$N(p^e, p^f, \mathbf{a}) = \sum_{a=0}^{p^e-1} S_{p^e}(a, p^f, \mathbf{a}),$$

since

$$\sum_{a=0}^{q} e(\frac{a}{q} w) = \begin{cases} q, & q | w, \\ 0, & q \nmid w. \end{cases}$$

We proceed to consider the behaviour of $N(p^e, p^f, \mathbf{a})$ as e tends to infinity. An induction argument based on Lemma 2 shows that

$$\sum_{e=0}^{E} p^{-en} S(p^e, p^f, \mathbf{a}) = p^{-E(n-1)} N(p^E, p^f, \mathbf{a})$$

for every $E \geq 0$, whence

$$c_p = \lim_{e \to \infty} p^{-e(n-1)} N(p^e, p^f, \mathbf{a}).$$

We may therefore think of c_p as being the density of points on $F = 0$ in $\mathbb{A}^n(\mathbb{Z}_p)$, subject to the condition $\mathbf{x} \equiv \mathbf{a} (\mathrm{mod}\ m)$. Thus the product $c_\infty \prod_p c_p$ in (16) can be interpreted as a product of local densities, one for each place of $\mathbb{Q}$. It can be shown that $c_p > 0$ if and only if the variety $F = 0$ has a non-singular p-adic point $\mathbf{x}$ with $\mathbf{x} \equiv \mathbf{a} (\mathrm{mod}\ m)$.

Now, suppose that the circle method succeeds in showing that

$$N_{\mathcal{R}}(B; m, \mathbf{a}) \sim B^{n-d} m^{-n} c_\infty \prod_p c_p \quad (B \to \infty)$$

for every $\mathbf{a}$ and m. We can then show that the rational points on $F = 0$ are dense in the adèlic points, providing that the non-singular adèlic points are dense in the set of all adèlic solutions of $F = 0$. To see this, let $\mathcal{N}$ be an adèlic neighbourhood containing a point on $F = 0$. According to our assumption, $\mathcal{N}$ will contain a non-singular adèlic point $a = (\mathbf{a}_\infty, \mathbf{a}_2, \mathbf{a}_3, \ldots)$ on $F = 0$. It follows that there exists $\varepsilon > 0$ and p_0 such that $\mathcal{N}$ contains the set of all adèles $(\mathbf{b}_\infty, \mathbf{b}_2, \mathbf{b}_3, \ldots)$ satisfying $|\mathbf{b}_\infty - \mathbf{a}_\infty| < \varepsilon$ and $|\mathbf{b}_p - \mathbf{a}_p| < \varepsilon$ for $p \leq p_0$. Thus if $\mathbf{x} \in \mathbb{Z}^n$ we will have $B^{-1}\mathbf{x} \in \mathcal{N}$ providing that $\mathbf{x} \in B\mathcal{R}$ and $\mathbf{x} \equiv B\mathbf{a} \pmod{m}$, for a suitable box $\mathcal{R}$ and modulus m. Since $\mathbf{a}$ is non-singular, the circle method provides an asymptotic formula for the number of relevant rational points of the form $B^{-1}\mathbf{x}$, in which the constant $c_\infty \prod_p c_p$ is strictly positive. The claim then follows.

Lecture 3 — The Minor Arcs in the Hardy-Littlewood Circle Method

In this third lecture we shall consider the minor arc integral

$$\int_{\mathfrak{m}} S(\alpha; B\mathcal{R}; m, \mathbf{a}) d\alpha.$$

Our primary goal is to show that this is $o(B^{n-d})$. However we also have a secondary goal, which is to establish the hypothesis used in the previous lecture to establish the convergence of the singular series. Up to now we have needed only mild conditions on the form F and its degree d, but unfortunately we shall now impose rather severe restrictions.

We shall begin by assuming that F is a diagonal form, so that it is of the shape

$$F(\mathbf{x}) = c_1 x_1^d + \ldots + c_n x_n^d$$

for certain non-zero integer coefficients c_j. Under this assumption the generating function factorizes as

$$S(\alpha; B\mathcal{R}; m, \mathbf{a}) = \prod_{j=1}^{n} S_j(\alpha; B; m, a_j), \tag{17}$$

where

$$S_j(\alpha; B; m, a_j) := \sum_{\substack{x \in \mathbb{Z} \cap [B\kappa_j, B\lambda_j] \\ x \equiv a_j \,(\mathrm{mod}\ m)}} e(\alpha c_j x^d).$$

For this factorisation to take place it is important that $\mathcal{R}$ is a box, as well as having F diagonal.

We are now able to consider a one-variable exponential sum, rather than an n-variable sum, and this produces a considerable simplification. We begin by using Hölder's inequality to show that

$$\begin{aligned}
&\left| \int_{\mathfrak{m}} S(\alpha; B\mathcal{R}; m, \mathbf{a}) d\alpha \right| \\
&\quad \leq \int_{\mathfrak{m}} \prod_{j=1}^{n} |S_j(\alpha; B; m, a_j)| d\alpha \\
&\quad \leq \sup_{\alpha \in \mathfrak{m}} \prod_{j \leq n-4} |S_j(\alpha; B; m, a_j)| \int_{\mathfrak{m}} \prod_{j=n-3}^{n} |S_j(\alpha; B; m, a_j)| d\alpha \\
&\quad \leq \sup_{\alpha \in \mathfrak{m}} \max_{j \leq n-4} |S_j(\alpha; B; m, a_j)|^{n-4} \left\{ \prod_{j=n-3}^{n} \int_{\mathfrak{m}} |S_j(\alpha; B; m, a_j)|^4 d\alpha \right\}^{1/4}
\end{aligned}$$

$$\leq \sup_{\alpha\in\mathfrak{m}} \max_{j\leq n-4} |S_j(\alpha;B;m,a_j)|^{n-4} \max_{n-3\leq j\leq n} \int_{\mathfrak{m}} |S_j(\alpha;B;m,a_j)|^4 d\alpha$$
$$\leq \sup_{\alpha\in\mathfrak{m}} \max_{j\leq n-4} |S_j(\alpha;B;m,a_j)|^{n-4} \max_{n-3\leq j\leq n} \int_0^1 |S_j(\alpha;B;m,a_j)|^4 d\alpha.$$

It therefore follows that there are indices j and k for which

$$\left|\int_{\mathfrak{m}} S(\alpha;B\mathcal{R};m,\mathbf{a})d\alpha\right|$$
$$\leq \sup_{\alpha\in\mathfrak{m}} |S_j(\alpha;B;m,a_j)|^{n-4} \int_0^1 |S_k(\alpha;B;m,a_k)|^4 d\alpha. \tag{18}$$

We first examine the integral. We have

$$|S_k(\alpha;B;m,a_k)|^4 = \sum_{x_1,x_2,x_3,x_4} e(\alpha c_j(x_1^d + x_2^d - x_3^d + x_4^d)),$$

where the variables are subject to $x_i \in [B\kappa_k, B\lambda_k]$ and $x_i \equiv a_k (\mathrm{mod}\ m)$. On applying (1) we therefore find that

$$\int_0^1 |S_k(\alpha;B;m,a_k)|^4 d\alpha = \#\{(x_1,x_2,x_3,x_4) : x_1^d + x_2^d = x_3^d + x_4^d\}, \tag{19}$$

with the same constraints on the variables x_i as before. Since $-1 \leq \kappa_k < \lambda_k \leq 1$ we deduce that

$$\int_0^1 |S_k(\alpha;B;m,a_k)|^4 d\alpha$$
$$\leq \#\{(x_1,x_2,x_3,x_4) \in \mathbb{Z}^4 \cap [-B,B]^4 : x_1^d + x_2^d = x_3^d + x_4^d\}.$$

We can estimate this quantity relatively easily, using the following well-known bound.

Lemma 3 *For any $\varepsilon > 0$ there is a constant $c(\varepsilon)$ such that*

$$\#\{m \in \mathbb{N} : m|N\} \leq c(\varepsilon)N^{\varepsilon}$$

for every positive integer N.

When $x_1^d > x_3^d$ we apply this to $N = x_1^d - x_3^d$, observing that $m = |x_2 - x_4|$ will be a positive integer divisor of N. Moreover we may note that the values of $N = x_4^d - x_2^d$ and $x_4 - x_2$ determine at most $d-1$ possible values of x_2 and x_4. It follows that each pair x_1, x_3 with $x_1^d > x_3^d$ leads to $O(B^{d\varepsilon})$ quadruples

(x_1, x_2, x_3, x_4). The case $x_1^d < x_3^d$ is of course similar, so that there are a total of $O(B^{2+d\varepsilon})$ solutions with $x_1^d \neq x_3^d$. Finally we observe that if $x_1^d = x_3^d$ then $x_2^d = x_4^d$. Thus we get $O(B^2)$ solutions in total this way. It therefore follows that

$$\int_0^1 |S_k(\alpha; B; m, a_k)|^4 d\alpha \ll B^{2+d\varepsilon} \tag{20}$$

for any fixed $\varepsilon > 0$. One of the remarkable aspects of the circle method is the way in which the proof of the asymptotic behaviour of $N_{\mathcal{R}}(B)$ is made to depend on the behaviour of the counting function for a completely different variety, namely $x_1^d + x_2^d = x_3^d + x_4^d$. Indeed for this latter variety a crude upper bound is sufficient.

It remains to bound an individual value of $S_j(\alpha; B; m, a_j)$, so as to deal with the supremum on the right hand-side of (18). Here we shall handle only the case $d = 2$. Thus for the rest of this lecture we shall assume that $d = 2$, with the specific goal of proving Theorem 1. The procedure we use dates back to the work of Weyl [14]. The general argument is discussed.... We shall give an exposition here that is specific to our situation. For convenience of notation we put

$$\beta := \alpha c_j, \quad \kappa := \kappa_j, \quad \lambda := \lambda_j, \quad \text{and} \quad a := a_j,$$

so that

$$S_j(\alpha; B; m, a_j) = \sum_{\substack{\kappa B < x \leq \lambda B \\ x \equiv a \pmod m}} e(\beta x^2).$$

The argument begins with the expansion

$$|S_j(\alpha; B; m, a_j)|^2 = \sum_{x,y} e(\beta(x^2 - y^2)).$$

We then set $x = y + mz$, so that z runs over integer values in the range $|z| \leq (\lambda - \kappa)B/m$. Thus in particular we have $|z| \leq 2B$. For each value of z the conditions on y become

$$y + mz \in (\kappa B, \lambda B], \quad y \in (\kappa B, \lambda B], \quad \text{and} \quad y \equiv a \pmod m. \tag{21}$$

We now substitute $y = a + mw$ and set

$$\mu := \{\max(\kappa B - mz\,,\, \kappa B) - a\}/m, \quad \nu := \{\min(\lambda B - mz\,,\, \lambda B) - a\}/m,$$

so that the range (21) for y corresponds to the range $\mu < w \leq \nu$ for w. The substitutions we have made yield

$$x^2 - y^2 = m^2 z^2 + 2amz + 2m^2 zw,$$

whence

$$\begin{aligned}|S_j(\alpha;B;m,a_j)|^2 &= \sum_{|z|\le 2B} e(\beta(m^2z^2+2amz))\sum_{\mu<w\le\nu} e(2\beta m^2zw)\\ &\le \sum_{|z|\le 2B}\Big|\sum_{\mu<w\le\nu} e(2\beta m^2zw)\Big|.\end{aligned}$$

Now in general, for any integers $W_1<W_2$, and for any $\gamma\in\mathbb{R}\setminus\mathbb{Z}$, we have

$$\begin{aligned}\sum_{W_1\le w\le W_2} e(\gamma w) &= \frac{e(\gamma(W_2+1))-e(\gamma W_1)}{e(\gamma)-1}\\ &= \frac{e(\gamma(W_2+1/2))-e(\gamma(W_1-1/2))}{2i\sin(\pi\gamma)},\end{aligned}$$

whence

$$\Big|\sum_{W_1\le w\le W_2} e(\gamma w)\Big| \le \frac{1}{|\sin(\pi\gamma)|}.$$

We also have the trivial bound W_2-W_1+1. It follows, on writing $\xi:=2\beta m^2$, that

$$\sum_{\mu<w\le\nu} e(2\beta m^2zw) \ll \min\{B\,,\,\frac{1}{|\sin(\pi\xi z)|}\},$$

and hence

$$|S_j(\alpha;B;m,a_j)|^2 \ll \sum_{|z|\le 2B}\min\{B\,,\,\frac{1}{|\sin(\pi\xi z)|}\}.$$

It is convenient to introduce at this point the periodic "saw-tooth" function

$$\theta:\mathbb{R}\to[-\frac{1}{2}\,,\,\frac{1}{2}],\quad \theta(t):=t-\max\{n\in\mathbb{Z}:n\le t+\frac{1}{2}\}.$$

This has the property that $|\theta(t)|\ll|\sin\pi t|\ll|\theta(t)|$. We proceed to decompose the available range for z into subsets

$$Z_h:=\{z\in\mathbb{Z}:|z|\le 2B,\ \frac{h}{B}\le\theta(\xi z)<\frac{h+1}{B}\},$$

where h runs over integers with $|h|\le 1+B/2$. Then if $|h|\ge 2$ we have

$$\frac{1}{|\sin(\pi\xi z)|}\ll\frac{1}{|\theta(\xi z)|}\ll\frac{B}{1+|h|}$$

for $z \in Z_h$, while if $|h| \le 1$ we have

$$B \ll \frac{B}{1+|h|}.$$

Thus in either case we find that

$$\min\{B\,,\,\frac{1}{|\sin(\pi\xi z)|}\} \ll \frac{B}{1+|h|},$$

whence

$$|S_j(\alpha;B;m,a_j)|^2 \ll (\max_h \#Z_h) \sum_{|h|\le 1+B/2} \frac{B}{1+|h|} \ll (\max_h \#Z_h) B\log B.$$

Now if $z_1, z_2 \in Z_h$ it follows that either $z_2 - z_1$ or $z_1 - z_2$ must be in Z_0, and hence that $\#Z_h \le 1 + 2\#Z_0$ for every value of h. Since $0 \in Z_0$ we have $1 + 2\#Z_0 \le 3\#Z_0$. We may therefore summarise our findings as follows.

Lemma 4 *With the notation above we have*

$$|S_j(\alpha;B;m,a_j)|^2 \ll B(\log B)\#\{z \in \mathbb{Z} : |z| \le 2B,\ 0 \le \theta(\xi z) < B^{-1}\}.$$

Before proceeding further with our treatment of the minor arcs we pause to deduce an important corollary of this lemma. If we choose

$$\kappa_j = m, \quad \lambda_j = 2m, \quad \alpha = \frac{a}{q} \quad \text{and} \quad B = q$$

then we will have $\xi = 2m^2 c_j a/q$ and

$$S_q(a;m,\mathbf{a}) = S(\alpha;B\mathcal{R};m,\mathbf{a}).$$

Moreover we find that

$$S_q(a;m,\mathbf{a}) = S(\alpha;B\mathcal{R};m,\mathbf{a}) = \prod_{j=1}^{n} S_j(\alpha;B;m,a_j),$$

by (17). Lemma 4 shows that

$$|S_j(\alpha;B;m,a_j)|^2$$
$$\ll q(\log q)\#\{z \in \mathbb{Z} : |z| \le 2q,\ 0 \le \theta(2m^2 c_j a z/q) < q^{-1}\}.$$

Since $2m^2 c_j a z/q$ is a fraction with denominator q, the values of z we have to consider are precisely those for which $q|2m^2 c_j a z$. However a and q are

coprime, so the condition reduces to $q|2m^2c_jz$. Moreover m and the various c_j are fixed, so that there are at $O(1)$ values of z in any range of length q. It follow that

$$|S_j(\alpha; B; m, a_j)|^2 \ll q(\log q)$$

and hence that

$$S_q(a; m, \mathbf{a}) \ll q^{n/2}(\log q)^{n/2}.$$

This therefore verifies the hypothesis made in our second lecture, as soon as $n \geq 5$, with any η in the range $0 < \eta < 1/2$.

We must now complete our estimation of $S_j(\alpha; B; m, a_j)$ in the case in which α lies in the minor arc set $\mathfrak{m}$. We should first recall that the major arcs were the union of intervals

$$I(a, q) = \left[\frac{a}{q} - \frac{Q}{B^2}\,,\, \frac{a}{q} + \frac{Q}{B^2}\right], \quad (q \leq Q),$$

where $Q = B^\delta$. We proceed to use Dirichlet's approximation theorem to find a rational approximation u/v to α satisfying

$$\left|\alpha - \frac{u}{v}\right| \leq \frac{Q}{B^2 v}, \quad (u, v) = 1, \quad v \leq B^2 Q^{-1}. \tag{22}$$

It follows from this that $|\alpha - u/v| \leq QB^{-2}$, and since $\alpha \notin \mathfrak{M}$ we deduce that

$$v > Q. \tag{23}$$

Now for $|z| \leq 2B$ we have

$$|\xi z - \frac{2m^2c_juz}{v}| \leq \frac{4m^2|c_j|Q}{Bv} \leq \frac{4m^2|c_j|}{B},$$

by (23). Thus if $0 \leq \theta(\xi z) \leq B^{-1}$ we must have

$$|\theta(\frac{2m^2c_juz}{v})| \leq \frac{1 + 4m^2|c_j|}{B}.$$

It follows that

$$2m^2c_juz \equiv l \pmod{v} \tag{24}$$

for some integer l satisfying

$$|l| \leq v\,\frac{1 + 4m^2|c_j|}{B} \ll vB^{-1}.$$

The number of possible values for l is therefore $O(1 + vB^{-1})$. Moreover for each l the congruence (24) determines $O(1)$ residue class modulo v, since u

and v are coprime and $2m^2c_j$ is a fixed non-zero integer. Finally the range $|z| \le 2B$ contains at most $1+4Bv^{-1}$ integers from each residue class modulo v. On combining this information we find that

$$\#\{z \in \mathbb{Z} : |z| \le 2B,\, 0 \le \theta(\xi z) < B^{-1}\} \ll (1 + vB^{-1})(1 + v^{-1}B)$$

$$\ll 1 + vB^{-1} + v^{-1}B \ll BQ^{-1},$$

in view of (22) and (23).

Finally we can apply Lemma 4 to deduce that

$$|S_j(\alpha; B; m, a_j)|^2 \ll B(\log B).BQ^{-1},$$

whence

$$|S_j(\alpha; B; m, a_j)| \ll B^{1-\delta/2}(\log B)^{1/2}. \tag{25}$$

On combining this with (18) and (20) we deduce, in the case $d = 2$, that

$$\left| \int_{\mathfrak{m}} S(\alpha; B\mathcal{R}; m, \mathbf{a}) d\alpha \right| \ll \{B^{1-\delta/2}(\log B)^{1/2}\}^{n-4} B^{2+2\varepsilon}$$

for any fixed $\varepsilon > 0$. This gives us a satisfactory bound $o(B^{n-2})$ on choosing ε small enough, providing that $n \ge 5$. This completes the proof of Theorem 1.

Lecture 4 — Combining Analytic and Geometric Methods

We have seen in the previous lectures that when the circle method succeeds it actually proves that the rational points on the variety under consideration are dense in the adèlic points. However there are many varieties which do not have this latter property, because there is a non-empty "Brauer-Manin obstruction". Consequently the circle method is doomed to failure in such cases. While it would be out of place in these lectures to give a full description of the Brauer-Manin obstruction, it is hoped that the example below will convey an idea of the general situation.

In this lecture we shall examine two cases in which analytic methods have been successfully combined with a "descent" process to rescue the above situation. The first of these is due to Heath-Brown and Skorobogatov [8]. Let K be an algebraic number field of degree $d \geq 2$, with integral basis $\omega_1, \ldots, \omega_d$, and define a form of degree d by setting

$$N(\mathbf{x}) := N(x_1, \ldots, x_d) := N_{K/\mathbb{Q}}(\omega_1 x_1 + \ldots + \omega_d x_d),$$

with the obvious abuse of notation. One is then interested in rational solutions to the equation

$$ct(t-1) = N(\mathbf{x}) \neq 0, \tag{26}$$

where c is a given non-zero integer, and where t and $\mathbf{x}$ are required to satisfy certain local conditions. (More generally, one can study $ct^a(t-1)^b = N(\mathbf{x}) \neq 0$ for arbitrary fixed integer exponents a and b.) These varieties provide a number of examples in which the rational points *fail* to be equidistributed.

Let us examine the case in which $K = \mathbb{Q}(\cos(2\pi/7))$. This is the the real subfield of the cyclotomic field generated by the primitive 7-th roots of unity. It is an abelian cubic field. We shall use in particular the fact that $n \in \mathbb{N}$ is a norm from K if and only it is a product of prime powers p^e such that $p^e \equiv \pm 1 (\mathrm{mod}\ 7)$ whenever $p \neq 7$. We shall take $c = 2$. Then there are solutions to (26) with $t = 28$, and also with $t = 29$. Thus there is a 2-adic solution (namely $t = 28$) with $|t|_2 < 1$, and a 7-adic solution (namely $t = 29$) with $|t-1|_7 < 1$. This produces an adèlic solution satisfying both constraints simultaneously.

We proceed to show however that there is no rational solution satisfying both constraints. To do this we write $t = u/v$ with u, v being coprime integers. We then see that $2u(u-v)v$ is a norm. It follows from the coprimality that exactly one of u, $u-v$ and v is even whence, using the coprimality condition again, we may conclude that one of the triples $u/4, u-v, v$, or $u, (u-v)/4, v$ or $u, u-v, v/4$ consists entirely of integral norms. In each case

this gives rise to an equation of the type $\pm 4N_1 \pm N_2 \pm N_3 = 0$, where each N_i is a norm. Since we must have $N_i \equiv 0$ or $\pm 1 (\mathrm{mod}\ 7)$ we deduce that we must have $7|N_1$. It follows that whichever one of $u, u-v$ or v is even must also be a multiple of 7. Hence if $t = u/v$ satisfies $|t|_2 < 1$ we must also have $|t|_7 < 1$. Thus we cannot have $|t|_2 < 1$ and $|t-1|_7 < 1$ simultaneously.

We have therefore shown that, for this example, the rational points are not dense in the adèlic points. Very roughly the argument consisted on using divisibility information to pass to a "descent variety" (which in this case was $\pm 4N_1 \pm N_2 \pm N_3 = 0$), for which there are local constraints not implicit in the original variety.

We proceed to examine the general situation. In homogeneous form the equation (26) becomes

$$cT(T-U)U^{d-2} = N(\mathbf{X}) \neq 0, \tag{27}$$

which involves $d+2$ variables. In our preliminary discussion of the circle method we saw that one generally requires more that $2d$ variables for the method to succeed, so one cannot expect the circle method to work here. Indeed it transpires that there is in general a non-empty Brauer-Manin obstruction for this variety.

One way to think of the descent process for the variety (27) is to introduce a non-zero parameter λ, and to look for solutions in which

$$T = \lambda N(\mathbf{y}), \quad U = w^d$$

and

$$c\lambda^2 N(\mathbf{y}) - N(\mathbf{z}) = c\lambda w^d \tag{28}$$

for suitable nonzero integral $\mathbf{y}, \mathbf{z}$ and w. With the above choices we see that

$$cT(T-U)U^{d-2} = w^{d(d-2)} N(\mathbf{y}) N(\mathbf{z}) \neq 0,$$

and since the set of norms is closed under multiplication this provides us with a solution to (27). It turns out that, if the local conditions on t and $\mathbf{x}$ are acceptable for the Brauer-Manin condition, then one can always find a value for λ for which (28) is locally solvable.

These manoeuvres have had two positive effects. Firstly they have replaced the original problem (27) with a new one (28), in which the number of variables, namely $2d+1$, is now strictly greater than $2d$. Thus there is at least some prospect of the circle method working. The second point is a more subtle one. It turns out that there is no Brauer-Manin obstruction for varieties of the form (28), so again there is no *a priori* reason for the circle method to fail.

We shall not give the full argument needed to handle (28) by the circle method, but there is one point that does deserve to be discussed. It is far from true that the circle method works precisely when the number of variables satisfies $n > 2d$. In rare cases one can handle fewer variables; more frequently success requires distinctly more variables. Thus it is worthwhile seeing how, in this case, we can manage with as few as $2d + 1$ variables.

We saw, in the second lecture, how the major arcs could be handled under quite general conditions. The real difficulty lies with the minor arcs. For (28) it is natural to use a generating function

$$\sum_{\mathbf{y},\mathbf{z},w} e(\alpha\{c\lambda^2 N(\mathbf{y}) - N(\mathbf{z}) - c\lambda w^d\}),$$

with suitable summation conditions for the variables. This generating function will factor as

$$S_1(c\lambda^2\alpha)S_2(-\alpha)S_3(-c\lambda\alpha),$$

where

$$S_1(\beta) = \sum_{\mathbf{y}} e(\beta N(\mathbf{y})),$$

$$S_2(\beta) = \sum_{\mathbf{z}} e(\beta N(\mathbf{z}))$$

and

$$S_3(\beta) = \sum_{w} e(\beta w^d).$$

The minor arc procedure which led to (18) now shows that

$$\left|\int_{\mathfrak{m}} S_1(c\lambda^2\alpha)S_2(-\alpha)S_3(-c\lambda\alpha)d\alpha\right|$$

$$\leq \sup_{\alpha\in\mathfrak{m}} |S_3(-c\lambda\alpha)| \max_{k=1,2} \int_0^1 |S_k(c_k\alpha)|^2 d\alpha,$$

where $c_1 = c\lambda^2$ and $c_2 = -1$.

The sum S_3 is the degree d version of the quadratic sum we investigated in (25), and may be handled by an extension of Weyl's method. It is the mean-value of S_k for $k = 1$ and 2, which interests us here. On expanding the square and integrating termwise we find, in complete analogy to (19), that

$$\int_0^1 |S_k(c_k\alpha)|^2 d\alpha = \#\{\mathbf{y}_1, \mathbf{y}_2 : N(\mathbf{y}_1) = N(\mathbf{y}_2)\}$$

with the vectors $\mathbf{y}_1, \mathbf{y}_2$ in appropriate ranges. If we take, for illustrative purposes, all coordinates of $\mathbf{y}_1, \mathbf{y}_2$ to lie in the range $[1, B]$, we see that the trivial bound for $S_k(c_k\alpha)$ is $O(B^d)$. On the other hand, for each given $\mathbf{y}_1$ we claim that the number of $\mathbf{y}_2$ with $N(\mathbf{y}_1) = N(\mathbf{y}_2)$ is $O(B^\varepsilon)$, for any small fixed $\varepsilon > 0$. It follows that

$$\int_0^1 |S_k(c_k\alpha)|^2 d\alpha \ll B^{d+\varepsilon},$$

so that we essentially have square-root cancellation. It is this remarkably sharp mean-value bound which is the key to success.

To verify the claim above we consider the equation $N_{K/\mathbb{Q}}(\gamma) = a$ with γ an algebraic integer in K, whose various conjugates are all $O(B)$. The number of integral ideals of norm a is at most $\tau(a)^d$, where $\tau(a)$ is the number of divisors of a. According to Lemma 3 this is $O(a^{d\varepsilon})$ for any $\varepsilon > 0$. Allowing for possible associates, each ideal of norm a corresponds to $O((\log B)^d)$ possible integers γ, in view of the constraints on the conjugates of γ. The claimed result them follows, on re-defining ε.

The remainder of this lecture is devoted to a second example in which analytic methods must be combined with a "descent" argument. Full details are in the author's paper [9].

In this case we shall examine a variety defined by a pair of simultaneous equations of the shape

$$V: \quad \begin{cases} L_1(x_1, x_2)L_2(x_1, x_2) = x_3^2 + x_4^2 \\ L_3(x_1, x_2)L_4(x_1, x_2) = x_5^2 + x_6^2 \end{cases} \tag{29}$$

where $L_1, \ldots, L_4$ are linear forms defined over $\mathbb{Z}$, no two of which are proportional. It is known that the variety V can fail to satisfy the Hasse Principle, and that even if there are rational points, they may fail to be dense in the adèlic points. Thus a direct application of the circle method is impossible. Indeed for pairs of quadratic forms the best result in the literature is that of Cook [2], which relates only to diagonal forms, and requires at least 9 variables.

We shall consider points on V (more precisely, on the affine cone over V), for which $\mathbf{x} = (x_1, x_2) \in B\mathcal{R}$. Here $\mathcal{R} \subset \mathbb{R}^2$ is open, bounded and convex, with a piecewise continuously differentiable boundary, and B is a large positive parameter. Note that if $\mathbf{x} \in B\mathcal{R}$ then the sizes of $x_3, \ldots, x_6$ are all restricted to be $O(B)$. We further assume that $L_i(\mathbf{x}) > 0$ for $1 \leq i \leq 4$, for all $\mathbf{x} \in \mathcal{R}$. One further technical condition needs to be imposed, namely that

$$L_1(x_1, x_2) \equiv L_2(x_1, x_2) \equiv \nu x_1 \pmod 4$$

and

$$L_3(x_1, x_2) \equiv L_4(x_1, x_2) \equiv \nu' x_1 \pmod 4,$$

for appropriate $\nu, \nu' = \pm 1$. It is then appropriate to impose a congruence restriction on $\mathbf{x}$ and to work with

$$(B\mathcal{R})^{(2)} := \{\mathbf{x} \in B\mathcal{R} : x_1 \equiv 1 \pmod 2\}.$$

Introducing the arithmetic function

$$r(n) := \#\{(a, b) \in \mathbb{Z}^2 : a^2 + b^2 = n\},$$

we see that the number of solutions of (29) under consideration is given by

$$S(B) := \sum_{\mathbf{x} \in (B\mathcal{R})^{(2)}} r(L_1(\mathbf{x})L_2(\mathbf{x}))r(L_3(\mathbf{x})L_4(\mathbf{x})).$$

The principle result of [9] is then the following.

Theorem 2 *Let σ_∞ and σ_p be the local densities for the variety V with equations (29), for the set $(B\mathcal{R})^{(2)}$. If $\sigma_p = 0$ for any prime p, then V has no rational point with $(x_1, x_2) \in (B\mathcal{R})^{(2)}$. If $\sigma_p \neq 0$ for every prime p, then*

$$S(B) = \{1 + \varepsilon\}\sigma_\infty \prod_p \sigma_p + o(B^2)$$

for a certain rational constant $\varepsilon \in [-1, 1]$.

If $\varepsilon = -1$ then V has no rational point with $(x_1, x_2) \in (B\mathcal{R})^{(2)}$.

Note that σ_∞ will be a positive constant multiple of B^2, so that we have a genuine asymptotic formula if the other local factors are positive, and if $\varepsilon \neq -1$. In fact the constant ε is given explicitly. It is a finite product of certain local factors. We therefore see how the Hardy-Littlewood asymptotic formula holds with the additional factor $1 + \varepsilon$, which reflects Brauer-Manin considerations. Indeed, if all the σ_p are non-zero, one has $1 + \varepsilon = 0$ precisely when there is a Brauer-Manin obstruction.

Just as we saw how (27) led to (28), we may replace the variety (29) by the system

$$\begin{array}{ll} L_1(\mathbf{x}) = d(y_1^2 + y_2^2), & L_2(\mathbf{x}) = d(y_3^2 + y_4^2), \\ L_3(\mathbf{x}) = d'(y_5^2 + y_6^2), & L_4(\mathbf{x}) = d'(y_7^2 + y_8^2), \end{array} \tag{30}$$

for different choices $d, d' \in \mathbb{N}$. Here we use the fact that $d(y_1^2+y_2^2) \times d(y_3^2+y_4^2)$ is a sum of two squares, $x_3^2+x_4^2$, say. The underlying process can be expressed more precisely using the identity

$$r(mn) = \frac{1}{4} \sum_{d|m,n} \mu(d)\chi(d)r(m/d)r(n/d),$$

where χ is the non-principal character modulo 4. This shows that

$$S(B) = \frac{1}{16} \sum_{d,d'} \mu(d)\mu(d')\chi(dd')S(B;d,d'), \tag{31}$$

with

$$S(B;d,d') := \sum_{\mathbf{x}\in(B\mathcal{R})^{(2)}} r(L_1(\mathbf{x})/d)r(L_2(\mathbf{x})/d)r(L_3(\mathbf{x})/d')r(L_4(\mathbf{x})/d').$$

Here we set $r(q) = 0$ if q is not an integer. The inner sum on the right clearly counts solutions to (30).

Just to check on progress, let us suppose one were to eliminate $\mathbf{x} = (x_1, x_2)$ from the equations (30). This would produce a pair of diagonal quadratic equations in the 8 variables $y_1, \dots, y_8$. Now we have noted already that Cook's work [2] would allow us to handle such a system via the circle method, if only there were 9 or more variables. However it transpires that although the circle method is not itself available for the varieties (30) there is an alternative analytic method which one can use, based on the fact that

$$r(m) = 4 \sum_{d|m} \chi(d).$$

The argument is quite delicate, and produces a result of the form

$$S(B;d,d') = C_{d,d'}B^2 + O(B^2(\log B)^{-\delta}) \tag{32}$$

for some small positive constant δ, for any fixed $d, d' \in \mathbb{N}$. It should be observed that our proof of Theorem 1 saves a positive power of the parameter B, while in contrast the above estimate saves only a power of $\log B$. This is a reflection of the fact that we are working very much on the border of what is currently feasible. Although the result does not come from the circle method, the constant $C_{d,d'}$ is none the less the product of the local densities for the equations (30). Thus we have exactly the result one would expect from the circle method, even though the method of proof is different.

Once (32) has been established it is possible to sum up the various terms in (31), using the asymptotic formula (32) for each pair d, d'. (There are important issues concerning uniformity in d, d' which need to be dealt with, but these can in fact be overcome.) This leads to a result of the form

$$\sum_{\mathbf{x}\in(B\mathcal{R})^{(2)}} r(L_1(\mathbf{x})L_2(\mathbf{x}))r(L_3(\mathbf{x})L_4(\mathbf{x})) = CB^2 + o(B^2)$$

with

$$C = \frac{1}{16} \sum_{d,d'} \mu(d)\mu(d')\chi(dd')C_{d,d'}.$$

To complete the proof one then computes this sum, and relates it to the product of local densities for (29).

We conclude with a numerical illustration, relating to the system

$$x_1(x_1 + 12x_2) = x_3^2 + x_4^2, \quad (x_1 + 4x_2)(x_1 + 16x_2) = x_5^2 + x_6^2.$$

In this case one can show that $0 < 1 + \varepsilon < 2$ and that

$$\{1 + \varepsilon\}\sigma_\infty \prod_p \sigma_p = 2$$

for the region

$$\mathcal{R} = \{0 < x_1, x_1 + 16x_2 < 1\}.$$

(It would appear that the occurrence of an integer value for the overall density is no more than a fluke!) One now has the following numerical values.

B	$S(B)$	$S(B)/2B^2$
1000	1993472	0.9967...
2000	8030592	1.0038...
4000	32057728	1.0018...
8000	1276046726	0.9969...
16000	511437824	0.9989...
32000	2043518720	0.9978...

Acknowledgements

Part of this work was undertaken while the author was attending the Diophantine Geometry session at the Centro di Ricerca Matematica Ennio De Giorgi in Pisa. The hospitality and financial support of the centre is gratefully acknowledged.

Thanks are also due to Dr T. Browning, whose careful reading of a first draft of these notes led to many improvements and corrections.

References

[1] B.J. Birch, Forms in many variables, *Proc. Roy. Soc. Ser. A,* 265 (1961/1962) 245-263.

[2] R.J. Cook, Simultaneous quadratic equations, *J. London Math. Soc. (2),* 4 (1971), 319-326.

[3] H. Davenport and D.J. Lewis, Simultaneous equations of additive type, *Philos. Trans. Roy. Soc. London Ser. A,* 264 (1969), 557-595.

[4] H. Davenport, Analytic methods for Diophantine equations and Diophantine inequalities, (CUP, Cambridge, 2005).

[5] P. Deligne, La conjecture de Weil. I, *Inst. Hautes Études Sci. Publ. Math.,* 52 (1980), 137-252.

[6] K. Ford, New estimates for mean values of Weyl sums, *Internat. Math. Res. Notices,* 1995, no. 3, 155-171.

[7] D.R. Heath-Brown, A new form of the circle method, and its application to quadratic forms, *J. Reine Angew. Math.,* 481 (1996), 149-206.

[8] D.R. Heath-Brown, and A.N. Skorobogatov, Rational solutions of certain equations involving norms, *Acta Math.,* 189 (2002), 161-177.

[9] D.R. Heath-Brown, Linear relations amongst sums of two squares, *Number theory and algebraic geometry*, 133-176, London Math. Soc. Lecture Note Ser., 303, (Cambridge Univ. Press, Cambridge, 2003).

[10] C. Hooley, On nonary cubic forms. III, *J. Reine Angew. Math.,* 456 (1994), 53-63.

[11] R.C. Vaughan, *The Hardy-Littlewood method, 2nd Edition* Cambridge Tracts in Mathematics, 125. (Cambridge University Press, Cambridge-New York, 1997).

[12] R.C. Vaughan, On Waring's problem for cubes, *J. Reine Angew. Math.,* 365 (1986), 122-170.

[13] Y. Wang, *Diophantine equations and inequalities in algebraic number fields*, (Springer-Verlag, Berlin, 1991).

[14] H. Weyl, Über die Gleichverteilung von Zahlen mod Eins, *Math, ann.,* 77 (1916), 313-352.

Mathematical Institute,
24-29, St. Giles',
Oxford OX1 3LB

rhb@maths.ox.ac.uk