ON SIMULTANEOUS ADDITIVE EQUATIONS AND WARING'S PROBLEM.

by

TREVOR DION WOOLEY.

A thesis submitted for the Degree of Doctor of Philosophy of the University of London and for the Diploma of Membership of the Imperial College.

Department of Mathematics,
Imperial College,
Queen's Gate,
London SW7 2BZ.

ABSTRACT.

This thesis is divided into three parts.

In Part I we consider the non-trivial solubility in p-adic integers of a pair of additive equations of degrees $k > n > 1$:

$$\left.\begin{aligned} f(\underset{=}{x}) &= a_1x_1^k + \ldots + a_sx_s^k = 0 \\ g(\underset{=}{x}) &= b_1x_1^n + \ldots + b_sx_s^n = 0 \end{aligned}\right\}, \qquad (*)$$

where the coefficients $a_1,\ldots,a_s$, and $b_1,\ldots,b_s$ are rational integers. It appears that the situation with $k \neq n$ has not previously been investigated. We develop methods involving p-normalisation procedures and exponential sums which enable us to show:

(i) when $k = 3$, $n = 2$, and $s \geq 11$, the equations (*) have a non-trivial solution in p-adic integers for every rational prime p. An example shows that this result is best possible.

(ii) if $s > 2(k+n)$ and $p > k^4n^2$, then the equations (*) have a non-trivial solution in p-adic integers. The condition on s is best possible.

In Part II we consider the non-trivial solubility of the equations (*) in rational integers. We show that when $k = 3$, $n = 2$ and $s \geq 14$, then subject to certain natural conditions on $a_1,\ldots,a_s$, $b_1,\ldots,b_s$, the equations (*) have a non-trivial rational solution. To prove this result, we first generalise Vaughan's iterative method to the case of simultaneous equations, and then use a variant of the Hardy-Littlewood method, subject to complications.

In Part III we investigate Waring's Problem. Let $G(k)$ denote the smallest number s such that every sufficiently large natural number n is the sum of at most s kth powers of natural numbers. We begin by showing that certain technical refinements of Vaughan's new

iterative method permit, in certain circumstances, an upper bound $G(k) \le H(k)$ to be replaced by $G(k) \le H(k)-1$, with further savings possible for $k > 12$. We then describe a new method which extends the ideas of Vaughan's iterative method. This permits us to improve substantially all previous upper bounds for $G(k)$ when $k \ge 6$. In particular, Vinogradov's upper bound

$$G(k) < (2+o(1))k\log k \text{ as } k \longrightarrow \infty ,$$

dating from 1959, is replaced by

$$G(k) < (1+o(1))k\log k \text{ as } k \longrightarrow \infty .$$

Παντες μεν Κιλικες κακοι 'ανερες; 'εν δε Κιλιξιν

εἷς 'αγαθος Κινυρης, και Κινυρης δε Κιλιξ.

Demodocus.

CONTENTS

PART III. ON WARING'S PROBLEM.

CHAPTER 5. ON WARING'S PROBLEM: SOME REFINEMENTS.

CHAPTER 6. LARGE IMPROVEMENTS IN WARING'S PROBLEM.

TABLES.

ACKNOWLEDGEMENTS.

It is a great pleasure for me to express my gratitude to the Imperial Mathematics Department for the excellent working environment, and stimulating atmosphere, which has contributed towards any success I may have achieved here. Thanks must go, in particular, to the postgraduates with whom I have "shared" an office during the past two years: Gwyneth Jones, Matthew Richards, Jonathan Weinreich and Richard Ashton.

I have benefited from interesting conversations with Professor Pollington and Dr. Coleman, and from the kind advice of Professor Chalk and Dr. Chen. My supervisor during my time at Imperial, Professor Vaughan, deserves special commendation for his very helpful advice, comments and suggestions. It is difficult to imagine that I would have been so productive had it not been for his pervasive enthusiasm.

Lastly, I must thank my family for enduring a pure mathematician for so long. Their support, and that of my friends, has been invaluable.

This work was made possible by the receipt of a Science and Engineering Research Council research grant.

Declaration on joint work with Professor R.C.Vaughan (supervisor).

The thesis is the candidates own work, except for two areas:
(i) The appendix "On a problem related to one of Littlewood and Offord". Here the idea of using exponential sums, and parts (A) and (B)(i) of the proof are the work of the supervisor. Also, Lemma 2.1

is due to the supervisor, this being a correction of the candidate's previous argument.

(ii) In the first part of the work on Waring's Problem, the work on the development and use of refined major arc estimates for the generating functions (§§7-9 of Chapter 5) is the work of the supervisor. Also, the idea of using Vinogradov's Mean Value Theorem in §10 is due to the supervisor.

NOTATION.

The notation throughout this thesis is consistent with that in current use by analytic number theorists, unless otherwise stated. We list below the notational conveniences most commonly adopted in this thesis:

$e(\alpha)$ denotes $e^{2\pi i\alpha}$,

$|\cdot|_p$ denotes the usual p-adic valuation, normalised with $|p|_p = p^{-1}$,

« and » refer to Vinogradov's well-known notation, so that we write $f \ll g$ to mean $f = O(g)$,

$p^r \| n$ is used to denote that $p^r | n$ but $p^{r+1} \nmid n$,

$[x]$ denotes the integer part of the real variable x,

$\|x\|$ is used to denote $\underset{y \in \mathbb{Z}}{\text{Min}} \{ |x-y| \}$.

We also make extensive use of vector notation for brevity. For example, $(c_1, \ldots, c_t)$ is abbreviated to $\underline{c}$.

For the sake of conciseness we adopt a localised approach to the labelling of results and equations. Thus, for example, the name of the nth lemma of section m of the pth chapter, when referring from the qth chapter, is:

$$\begin{cases} \text{"Lemma } p.m.n\text{" or "Lemma } m.n \text{ of Chapter } p\text{"} & \text{if } q \neq p, \\ \text{"Lemma } m.n\text{"} & \text{if } q = p. \end{cases}$$

This approach is brief, and causes no confusion.

CHAPTER 0.

INTRODUCTION AND HISTORICAL SURVEY.

1. DIOPHANTINE EQUATIONS, AND ARTIN'S CONJECTURE.

When he proposed his list of problems for the twentieth century at the 1900 International Congress, Hilbert asked for a method for deciding whether or not an arbitrary diophantine equation has a solution (Hilbert's tenth problem). That is, does a universal algorithm exist which decides, given a polynomial $f = f(x_1,\ldots,x_n)$ with integer coefficients, whether or not the equation $f = 0$ is soluble in integers $x_1,\ldots,x_n$. Matijasevič [1970, 1971] completed the solution of Hilbert's tenth problem —— in the negative. Despite this negative resolution of the original problem, there remains active interest in developing methods for establishing the solubility of certain restricted classes of equations. Interest has focused on two areas:

(i) Let $F(x,y)$ denote a polynomial with integer coefficients in the two variables x and y, and let m be some integer. Under certain circumstances it is possible to decide whether or not the equation $F(x,y) = m$ has only finitely many solutions in integers. In addition, if the equation has only finitely many solutions, it can frequently be shown that it has "very few" solutions (see A.Baker [1975], Chapter 4).

(ii) Let $G = G(x_1,\ldots,x_n)$ denote a homogeneous polynomial with integer coefficients in the variables $x_1,\ldots,x_n$. Under certain circumstances it can be shown that when n is sufficiently large, the equation $G = 0$ has a non-trivial

solution in integers (that is, a solution with not all the x_i zero). When a solution is known to exist, it can frequently be shown that "many" solutions exist.

We shall consider only the second of these areas (the first more properly belongs to transcendental or algebraic number theory).

Consider a homogeneous polynomial $G = G(x_1, \ldots, x_n)$ of degree k with integer coefficients. It is plain that if the diophantine equation

$$G(x_1, \ldots, x_n) = 0 \qquad (0.1)$$

is to be soluble non-trivially in rational integers, then necessarily,

(i) the equation $G = 0$ must be soluble non-trivially over $\mathbb{R}$,

(ii) for every rational prime p, the equation $G = 0$ must be soluble non-trivially over the p-adic field $\mathbb{Q}_p$. Equivalently, for every rational prime p, the congruence

$$G(x_1, \ldots, x_n) \equiv 0 \pmod{p^n}$$

must be soluble for each natural number n with one at least of the x_i coprime with p.

Conditions (i) and (ii) are known collectively as the *local solubility* conditions. If the equation (0.1) is soluble non-trivially in rational integers, then it is said to satisfy the *global solubility* condition. Under many circumstances we would expect local solubility to be a sufficient condition for global solubility (a generalised Hasse principle), but even the global solubility of cubic equations is not yet fully understood.

In 1936 Artin conjectured (see Artin [1965], p.x) that if $n > k^2$ then every homogeneous polynomial of degree $k > 0$ in $\mathbb{Q}_p[x_1, \ldots, x_n]$ has a non-trivial zero in $\mathbb{Q}_p$. For many years, all work done on this

problem appeared to support this conjecture. Interest was generated in restrictions of the problem, and it was probably Artin's conjecture, at least in part, that prompted Davenport and Lewis to work on additive equations in the sixties. The original conjecture has been shown to be false by Terjanian [1966], who exhibited a counter-example of degree 4 in 18 variables over $\mathbb{Q}_2$. Indeed, the conjecture is not just false, but "very false", for Arkhipov and Karatsuba [1981] have shown that there are arbitrarily large values of r, and homogeneous polynomials of degree $k = k(r)$ in at least k^r+1 variables, which possess no non-trivial solution over $\mathbb{Q}_p$ for some rational prime p. Although now thoroughly discredited for general polynomials, the Artin conjecture, now reincarnated as the "Artin question", gains some credibility from the restricted problem of additive equations.

2. P-ADIC SOLUBILITY OF ADDITIVE EQUATIONS.

Let c_{ij} $(1 \le i \le t,\ 1 \le j \le s)$ be rational integers, and k_i $(1 \le i \le t)$ be natural numbers. Consider the simultaneous diophantine equations

$$F_i(\underset{=}{x}) = c_{i1}x_1^{k_i} + \ldots + c_{is}x_s^{k_i} = 0 \quad (1 \le i \le t)\ . \tag{0.2}$$

The Artin question for this system of additive equations asks whether or not, for each rational prime p, the simultaneous equations (0.2) are non-trivially soluble over $\mathbb{Q}_p$ whenever

$$s > \sum_{i=1}^{t} k_i^2\ .$$

Davenport and Lewis [1963] were able to answer the Artin question in the affirmative for single additive equations (the case $t = 1$),

and went on to consider simultaneous additive equations ($t > 1$). For pairs of equations with $k_1 = k_2 = k$, Davenport and Lewis [1967] were able to answer the Artin question in the affirmative only for odd k, whilst for even k they showed that for each rational prime p, the equations (0.2) are soluble non-trivially whenever $s \geq 7k^3$. [When the powers are all equal, say to k, the general case with t equations gives somewhat weaker results, and the best results currently available, due to Low, Pitman and Wolff [1988], require at least $[2t^2k\log k + 1]$ variables when k is large and odd, and $[48tk^3\log(3tk^2)]$ when $k \geq 3$.]

In an earlier paper, Davenport and Lewis [1966] considered the local and global solubility of pairs of additive cubics. In particular, they were able to show that when $s \geq 16$, $t = 2$ and $k_1 = k_2 = 3$, then for each rational prime p, the equations (0.2) have a non-trivial p-adic solution. Further, they were able to find an example in 15 variables possessing no non-trivial 7-adic solution. So this result of Davenport and Lewis is in some sense best possible.

Cook [1985] refined the methods of Davenport and Lewis, using a computational check of cases to deal with stubborn small primes. He showed that 13 variables are sufficient to guarantee the non-trivial p-adic solubility of a pair of additive cubic equations whenever $p \neq 7$. Further, it is not difficult to show that there is an infinite set of primes for which 12 variables are insufficient. To see this, let $p \equiv 1 \pmod 3$, and take ω to be a cubic non-residue $(\mathrm{mod}\ p)$. Then the equation

$$x_1^3 - \omega y_1^3 + p(x_2^3 - \omega y_2^3) + p^2(x_3^3 - \omega y_3^3) = 0 \qquad (0.3)$$

has no non-trivial p-adic solution, since each of the congruences

$x_1^3 \equiv \omega y_1^3 \pmod{p}$ has no non-trivial solution (mod p). Thus, by considering two disjoint copies of the equation (0.3), we obtain a pair of equations in 12 variables having no non-trivial p-adic solution.

The prime 7 has therefore been shown to exhibit exceptional behaviour with regard to pairs of additive cubics. Later work (see Atkinson and Cook [1989]) has shown that there are exceptional primes for pairs of additive equations of higher degree (for example, 11 is exceptional for fifth powers). However, there are no exceptional primes for pairs of additive quadratic equations —— an argument similar to that applied above shows that for primes $p > 2$, an example in 8 variables exists possessing no non-trivial p-adic solution, and moreover Demjanov [1956] has shown that 9 variables are sufficient to guarantee the non-trivial p-adic solubility of a pair of additive quadratic equations.

The classification of exceptional primes is an area of current interest in the study of additive equations.

It is notable that to date, there appears to have been no consideration of the p-adic solubility of simultaneous additive equations in which the powers k_i are not all equal. In Part I of this thesis, we consider what is probably the simplest "non-trivial" example of a problem of this type. Thus we consider the p-adic solubility of equations of the type

$$\left.\begin{aligned} c_1x_1^3 + \ldots + c_sx_s^3 &= 0 \\ d_1x_1^2 + \ldots + d_sx_s^2 &= 0 \end{aligned}\right\} \tag{0.4}$$

with $c_i, d_i \in \mathbb{Z}$.

We are able to prove:

Theorem 0.1. *The equations* (0.4) *are soluble non-trivially in p-adic integers provided only that* $s \geq 11$.

An example in 10 variables for primes $p \equiv 1 \pmod 3$ shows that the condition on s is essentially best possible. Thus there are no exceptional primes for simultaneous additive equations of the form (0.4). In deriving this result, the method we use is essentially an elaboration of the methods of Davenport and Lewis, combined with the computational approach of Cook. Thus we use a variational principle to show that all equations of the form (0.4) are in some sense equivalent to a pair of equations which have many variables explicit (mod p). After disposing of a few standard cases, it is then possible to use exponential sums, *via* the identity

$$p^{-1}\sum_{r=0}^{p-1} e(rh/p) = \begin{cases} 1 & h \equiv 0 \pmod p \\ 0 & h \not\equiv 0 \pmod p, \end{cases}$$

to show that the derived equations have a solution (mod p) which can be lifted, by means of a standard argument from p-adic analysis (Hensel's Lemma), to a non-trivial p-adic solution. It then follows that the equations (0.4) have a non-trivial p-adic solution.

The outline of the method just given is a gross simplification. In particular, small primes (especially the prime 7) are very troublesome, and we are forced to go into some detail in order to dispose of them. The essential non-linearity introduced by considering such examples makes the analysis in many respects more difficult than that of Davenport and Lewis, and of Cook, and one of the goals of the first part of this thesis has been to initiate

progress towards a systematic method of dealing with such systems of equations.

A step in the proof of Theorem 0.1 entails estimating the number of singular solutions to a certain pair of congruences. In order to do this, we use a lemma counting the number of solutions to a certain "subset sum" congruence problem. This problem is rather similar to an archimedean problem of Littlewood and Offord [1943]. We include an appendix to examine this problem in some detail.

In the second chapter of Part I, we consider pairs of additive kth and nth power equations. We are able to derive a result which generalises one of Atkinson and Cook [1989], who considered the case with $k = n$.

Theorem 0.2. *Suppose that $k \geq n > 1$, and that p is a rational prime with $p > k^4n^2$. Then a pair of additive kth and nth power equations in at least $2(k+n)+1$ variables are simultaneously soluble non-trivially in p-adic integers.*

When $p \equiv 1 \pmod{kn}$, it is easy to construct an example of a pair of additive kth and nth power equations in $2(k+n)$ variables possessing no simultaneous non-trivial p-adic solution. Thus Theorem 0.2 shows that although there may be exceptional primes, there cannot be too many of them. Theorem 0.2 follows from a refinement and generalisation of the arguments used to prove Theorem 0.1.

3. RATIONAL SOLUBILITY OF ADDITIVE EQUATIONS.

Now consider the problem of the non-trivial global solubility of equations of the form (0.2).

Additive quadratic equations are sufficiently simple for a variety of algebraic methods to be brought to bear on the problem of global solubility. Thus it was shown in the nineteenth century that provided not all of the coefficients are of the same sign, then an additive quadratic equation in five variables is non-trivially soluble in rational integers. However, these methods do not easily generalise to equations of higher degree, and it remained until this century for progress to be made on the general problem.

The Hardy-Littlewood method provides a systematic means of passing from local solubility to global solubility of systems of diophantine equations. The modern versions of this method are now rather technical in nature, but we can illustrate many of the salient features by considering the rational solubility of a single additive kth power equation.

Let ζ_i and ξ_i be positive real numbers with $\zeta_i < \xi_i$ $(1 \le i \le s)$, and let P be a large positive real number. For $1 \le i \le s$, let

$$f_i(\alpha) = \sum_{\zeta_i P < x \le \xi_i P} e(\alpha x^k) .$$

Then by using the identity

$$\int_0^1 e(\alpha h)\, d\alpha = \begin{cases} 1 & h = 0 \\ 0 & h \in \mathbb{Z}\setminus\{0\}, \end{cases}$$

we deduce that the number of solutions, $R(P)$, of the additive equation

$$a_1 x_1^k + \ldots + a_s x_s^k = 0 \qquad (0.5)$$

in integers $x_1, \ldots, x_s$ inside the box

$$\mathcal{B} = (\zeta_1 P, \xi_1 P] \times (\zeta_2 P, \xi_2 P] \times \ldots \times (\zeta_s P, \xi_s P]$$

is given by

$$R(P) = \int_0^1 \prod_{i=1}^{s} f_i(a_i \alpha)\, d\alpha \ . \tag{0.6}$$

So we see that the number of solutions of the equation (0.5) inside $\mathcal{B}$ corresponds to the value of the integral (0.6). In such circumstances we frequently speak of the diophantine equation (meaning (0.5)) underlying the integral (0.6), and conversely, of the integral representation (meaning (0.6)) of the diophantine equation (0.5).

Roughly speaking, the idea which underlies the Hardy-Littlewood method is the splitting of the range of integration $[0,1]$ into two sets as follows. In one set, the so-called *minor arcs*, $\mathfrak{m}$, we put the points α with poor rational approximation (α is not close to a rational a/q with q "small"). In the complement of these minor arcs, the *major arcs*, $\mathfrak{M}$, we put the points α with good rational approximation.

By permitting P to become very large, on the major arcs we are able to obtain an asymptotic formula for the generating functions $f_i(\alpha)$ in terms of the distance of α from the nearest good rational approximation to α. By making an appropriate choice for the ζ_i and ξ_i, we can then show that the contribution from the major arcs to the integral (0.6) is asymptotic to

$$C\mathfrak{S}P^{s-k} + o(P^{s-k}) \ ,$$

in which C is a constant depending on the non-trivial real solubility of the equation (0.5), and $\mathfrak{S}$ is a constant depending on the non-trivial solubility of the equation (0.5) over all p-adic fields. Further, provided that the equation (0.5) has a non-trivial,

non-singular solution over both the real field, and every p-adic field, one can show that C and $\mathfrak{S}$ are both positive.

On the minor arcs we proceed in two stages. In the first stage we obtain an estimate for the integral

$$I_t(P) = \int_0^1 |f(\alpha)|^{2t} d\alpha ,$$

where

$$f(\alpha) = \sum_{x \leq P} e(\alpha x^k) ,$$

and with $2t < s$. The value of this integral is the number of solutions of the diophantine equation

$$x_1^k + \ldots + x_t^k = y_1^k + \ldots + y_t^k$$

with $1 \leq x_i, y_i \leq P$. By going back and forth between integral representations and diophantine equations, it is possible to manipulate equations to obtain an estimate of the form

$$I_t(P) \ll P^{2t-k+\varepsilon},$$

when t is sufficiently large in terms of k.

In the second stage, we exploit the "bad rational approximability" property of points in $\mathfrak{m}$ to obtain an estimate of the form

$$\sup_{\alpha \in \mathfrak{m}} |f_1(\alpha)| \ll P^{1-\rho+\varepsilon},$$

in which ρ is a positive number depending on k. Such estimates stem, essentially, from the theory of uniform distribution. By combining this latter estimate with the above mean value estimate, we obtain for $s > 2t$ a bound for the contribution of the minor arcs to the integral (0.6) which is $o(P^{s-k})$. Thus

$$R(P) \gg P^{s-k} \longrightarrow \infty \text{ as } P \longrightarrow \infty ,$$

and so there are many rational solutions to the equation (0.5).

All recent analyses of additive equations have depended on an approach of the form just outlined, although elaborations of the method are required in order to obtain the best results currently available. In the case of simultaneous equations we obtain as an integral representation a multi-dimensional integral. However, when the k_i are all equal, by carefully manipulating the equations it is possible to use linearity to make a change of variables which permits the integral over the minor arcs to be written as the product of several single dimensional integrals. Thus the original simultaneous problem may be dealt with by using methods available from the study of the corresponding single equation problem.

For brevity, we shall define $G^*(\underline{k}) = G^*(k_1,\dots,k_t)$ to be the least integer r such that for all $s \geq r$, and all c_{ij} $(1 \leq i \leq t,$ $1 \leq j \leq s)$ satisfying all conditions imposed by local solubility considerations, the simultaneous equations (0.2) have a non-trivial rational solution. Also, for the purposes of this discussion, we adopt the convention that "$G^*(\underline{k}) \leq r$" is to mean that r satisfies the above conditions subject to some further, not unreasonable, conditions (for example, we may require the equations to possess non-singular solutions over the real field).

We have already mentioned that $G^*(2) \leq 5$. Using methods based on the strategy outlined above, results for single additive equations have reached a reasonably satisfactory state. For cubics, we have seen already that p-adic solubility considerations imply that in general, at least 7 variables will be required to guarantee non-trivial rational solubility. Thus the result $G^*(3) \leq 7$ (R.Baker [1990]) is in some sense best possible. For general exponents, it is of interest to consider a rational "Artin question"

for additive equations, and for single equations this problem has recently been resolved by Vaughan. Davenport and Lewis [1963] were able to show that $G^*(k) \le k^2+1$ for all values of k other than $7 \le k \le 17$. This gap has slowly been filled, culminating with the case $k = 10$ (Vaughan [1989a]).

For systems of equations our knowledge is very much less satisfactory. However, the barriers to more precise results might be said to derive more from the incomplete state of our knowledge of the p-adic problem, than from limitations in the Hardy-Littlewood method itself. For pairs of equations, and small values of k, previous authors have obtained

$$\text{"}G^*(2,2) \le 9\text{"} \quad \text{(Cook [1971])},$$

and

$$G^*(3,3) \le 14 \quad \text{(Brüdern)}.$$

The latter result (not yet in print) is the most recent improvement on the result $G^*(3,3) \le 16$ (Vaughan [1977]), which in view of the example of Davenport and Lewis [1966] is essentially best possible (when "16" is replaced by any smaller integer, 7-adic solubility can no longer be guaranteed).

More general results have been obtained by previous authors. However, once again the case when the k_i are not all equal appears not to have been considered.

In Part II of this thesis, we consider the non-trivial global solubilty of the equations (0.4). We are able to prove the following result.

Theorem 0.3. ("$G^*(3,2) \le 14$"). *The simultaneous additive equations* (0.4) *have a non-trivial solution in rational integers if the following conditions hold.*

(*a*) *the quadratic equation in* (0.4) *is indefinite, and has at least five variables explicit, and*

(*b*) *the cubic equation in* (0.4) *has at least seven variables explicit, and*

(*c*) *the simultaneous equations* (0.4) *have a non-trivial real solution, and*

(*d*) (*i*) $s \ge 14$, *or*

(*ii*) *at least* 6 *of the* d_i *are zero, or*

(*iii*) *at least* 4 *of the* c_i *are zero.*

The non-linearity inherent in the system (0.4) forces us to adopt a fundamentally multi-dimensional approach to the proof of Theorem 0.3, as we are unable to use the linearising ideas available for systems of equations of the same degree. Although results stemming from Vinogradov's mean value theorem (see, for example, Hua [1965], Chapter V) are of some use, to obtain more precise results we are forced to generalise the modern approaches to the Hardy-Littlewood method.

In the application of the Hardy-Littlewood method to additive problems of the type described above, a fundamental rôle is played by estimates for the number of solutions of auxiliary equations of the form

$$x_1^{k_i} + \ldots + x_s^{k_i} = y_1^{k_i} + \ldots + y_s^{k_i} \quad (1 \le i \le t), \tag{0.7}$$

in which $1 \le x_j, y_j \le P$. One idea for improving classical estimates,

in which the x_j and y_j range over the entire interval, is to restrict the variables to lying in intervals of the form

$$P_j < x_j, y_j < 2P_j \text{ for } j = 1, \ldots, s,$$

where $P_1 \geq P_2 \geq \ldots$. The use of diminishing ranges does not, however, seem well suited to cases where the k_i are not all equal. The problem is that the method makes savings by exploiting features of the real character of solutions, which become less pronounced when the k_i are not all equal. Vaughan [1989a,b,c] has shown that when $t = 1$ a more efficient approach is to impose restrictions on the arithmetic character of the solutions.

In the first chapter of part II of this thesis, we shall demonstrate that this approach remains effective when $t > 1$, although there are then a number of algebraic, as well as analytic, questions to be answered if we are to take full advantage of the method. We consider the equations (0.7) with $x_j, y_j \in \mathcal{A}(P,R)$, for a suitable R, where

$$\mathcal{A}(P,R) = \{ n : n \leq P, \ p \text{ prime}, \ p|n \text{ implies } p \leq R \}.$$

We then relate the number of solutions of (0.7) to the number of solutions of the simultaneous equations

$$x_1^{k_i} + \ldots + x_r^{k_i} + m^{k_i}(u_1^{k_i} + \ldots + u_{s-r}^{k_i})$$

$$= y_1^{k_i} + \ldots + y_r^{k_i} + m^{k_i}(v_1^{k_i} + \ldots + v_{s-r}^{k_i}) \quad (1 \leq i \leq t) \qquad (0.8)$$

with $x_j, y_j \leq P$, $M < m \leq MR$, and $u_j, v_j \in \mathcal{A}(P/M, R)$. By making use of homogeneity and Hölder's inequality, we are then able to relate the number of solutions of the equations (0.8) to the number of solutions of (0.7) with s replaced by a range of values not too far from s.

The second chapter of Part II is devoted to the application of the Hardy-Littlewood method to our problem. Here there are a number

of difficulties over a standard application, many being caused by our non-standard generating functions. However, the underlying strategy remains the same, and we are able to bring the plan to a successful conclusion. We use the methods alluded to above on the minor arcs in our application of the Hardy-Littlewood method. The treatment of the major arcs is complicated in two respects. Firstly, we are dealing with an inherently non-linear problem, which causes difficulties even in a classical approach to the problem. Secondly, we have restricted the variables to lie in the set $\mathcal{A}(P,R)$, and this causes complications. It should be emphasised that in order to deal successfully with the major arcs, it is crucial that the set $\mathcal{A}(P,R)$ is relatively dense. Fortunately, when $R \geq P^{\eta}$ with $\eta > 0$, we have card $\mathcal{A}(P,R) \gg P$.

4. WARING'S PROBLEM.

Given a natural number k, consider now the diophantine problem of representing a natural number n as the sum of s kth powers of natural numbers:

$$x_1^k + \ldots + x_s^k = n \ . \tag{0.9}$$

Define $g(k)$ to be the least s such that every natural number n is represented in the form (0.9). In 1770, Waring asserted that $g(k)$ satisfies

$g(2) = 4$, $g(3) = 9$, $g(4) = 19$, and (later) $g(k) < \infty$ for $k \geq 4$.

Lagrange proved that $g(2) = 4$ in the eighteenth century, and during the nineteenth century the problem was solved for many values of k.

It was Hilbert [1909] who finally succeeded in proving that

$$g(k) < \infty \text{ for all } k \in \mathbb{N},$$

although the proof he gave guaranteed only the existence of such a number.

The integer

$$n = 2^k\left[\left(\frac{3}{2}\right)^k\right] - 1 , \tag{0.10}$$

is smaller than 3^k, and so can be represented as the sum of kth powers of 1 and 2 only. Thus we may deduce that

$$g(k) \geq 2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2 ,$$

and indeed it would seem that this holds with equality (this assertion remains unproven). For a discussion of the current state of play on this problem, the interested reader should see the introduction of Vaughan [1981b]. Note that it has been proved since the writing of that book that $g(4) \leq 20$ (Balasubramanian [1985]). Indeed, Balasubramanian, Deshouillers and Dress [1986a,b] claim to have proved the bound $g(4) \leq 19$, and hence $g(4) = 19$. However, it should be emphasised that there are some doubts about the extensive computational checks required in their "proof", and moreover a complete proof has yet to be published.

The number (0.10) clearly has some rather special features, and in general fewer than $g(k)$ kth powers suffice to represent all "large" natural numbers n. Thus we are led to consider $G(k)$, which we define to be the least natural number s such that all sufficiently large natural numbers are the sum of at most s kth powers of natural numbers. The evaluation of $G(k)$ is very much more difficult than that of $g(k)$, and to date, $G(k)$ is known precisely only in the cases $G(2) = 4$ (Lagrange) and $G(4) = 16$ (Davenport [1939b]).

Hardy and Littlewood were the first workers to make substantial progress on the problem of reducing $G(k)$. Using a method based on the ideas outlined in §2, they were able to show (Hardy and Littlewood [1922]) that

$$G(k) \leq (k-2)2^{k-1} + 5 \text{ for all } k \in \mathbb{N}. \qquad (0.11)$$

For smaller k, arguments based on the use of diminishing ranges in the Hardy-Littlewood method had led, by the Forties, to the bounds $G(3) \leq 7$ (Linnik [1943]), $G(4) = 16$ (Davenport [1939b]), $G(5) \leq 23$, $G(6) \leq 36$, $G(7) \leq 53$ (Davenport [1942], and his methods for $k = 7$), $G(8) \leq 73$, $G(9) \leq 99$, $G(10) \leq 122$ (Narasimhamurti [1941]). Moreover Davenport [1939b] was able to show that all natural numbers n satisfying the necessary congruence conditions (a local solubility hypothesis) are the sum of at most 14 biquadrates. We abbreviate this result by writing $G^{\#}(4) \leq 14$.

Meanwhile, Vinogradov was using ideas based on diminishing ranges, and mean value theorems, to obtain results for large k substantially smaller than that given by (0.11). Initially, he was able to show that as $k \longrightarrow \infty$ we have $G(k) = O((k\log k)^2)$, but later improved this result to $G(k) < (C+o(1)).k\log k$, with $C = 6$ (see Vinogradov [1934]). Over the next twenty-five years, Vinogradov [1959] was able to reduce the value of C to 2, obtaining the bound

$$G(k) < k(2\log k + 4\log\log k + 2\log\log\log k + 13) \text{ for } k \geq 170000.$$

Over the thirty years since then, only the $o(1)$ term in the bound $G(k) < (2+o(1))k\log k$ has been reduced, the most recent result being due to Vaughan [1989a].

In Part III of this thesis we improve substantially all previous upper bounds for $G(k)$ when $k \geq 6$. We are also able to improve the bound for $G(5)$, although it is reported that this has already been achieved by Brüdern.

The first chapter of this section is devoted to some rather technical refinements which enable small improvements to be made in bounds for $G(k)$. The interest in these methods probably derives more from the applications of such methods outside Waring's problem as to the improvements themselves. To illustrate the use of these methods directly, we show how to make the modest improvements in Waring's problem alluded to.

In the second chapter of this section, we go on to describe a new method in Waring's problem which leads to rather substantial improvements in the existing upper bounds. Indeed, the method leads to an asymptotic halving of the existing upper bounds, and so at last we may substantially improve Vinogradov's upper bound $G(k) \leq (2+o(1))k\log k$.

The major results of this part of the thesis may be summarised in the following two theorems:

Theorem 0.6. *As* $k \longrightarrow \infty$, *we have*

$$G(k) < k(\log k + \log\log k + O(1)) .$$

Theorem 0.7. *For* $5 \leq k \leq 20$, *we have*

$$G(k) \leq F(k),$$

where $F(k)$ *is as given by Table* 0.1.

Table 0.1

k	$F(k)$	k	$F(k)$	k	$F(k)$	k	$F(k)$
5	18	9	55	13	87	17	120
6	27	10	63	14	95	18	129
7	36	11	70	15	103	19	138
8	47	12	79	16	112	20	146

These bounds may be compared with the most recent results in print, themselves a substantial improvement on those previously known, due to Vaughan [1989a,c]. He has shown that $G^{\#}(4) \le 12$, $G(5) \le 19$, $G(6) \le 29$, $G(7) \le 41$, $G(8) \le 57$, $G(9) \le 75$, $G(10) \le 93$, $G(11) \le 109$, $G(12) \le 125$, $G(13) \le 141$, $G(14) \le 156$, $G(15) \le 171$, $G(16) \le 187$, $G(17) \le 202$, $G(18) \le 217$, $G(19) \le 232$, $G(20) \le 248$.

We remark that the above results are intended as something of a demonstration of the power of the method, and improvements may be obtained by refining the method, especially for smaller k. This is a matter we intend to return to in papers subsequent to this thesis.

As has become standard in applications of the Hardy-Littlewood method, bounding $G(k)$ depends fundamentally on estimates for the number of solutions of auxiliary equations of the form

$$x_1^k + \ldots + x_s^k = y_1^k + \ldots + y_s^k \tag{0.12}$$

with the x_j and y_j lying in certain subsets $\mathcal{A}_j$ of $[1,P]\cap\mathbb{Z}$. Vaughan's substantial innovation was to set each of the $\mathcal{A}_j$ equal to a set $\mathcal{A}$ with suitable arithmetic properties, and to use these properties to relate the number of solutions of the equation (0.12) to the number of solutions of the equation

$$x^k + m^k(u_1^k + \ldots + u_{s-1}^k) = y^k + m^k(v_1^k + \ldots + v_{s-1}^k) \tag{0.13}$$

with $(xy,m) = 1$, $x,y \le P$, $M < m \le M'$, and $u_j, v_j \in \mathcal{B}$, where $\mathcal{B}$ has

similar properties to $\mathcal{A}$, but with $\mathcal{B} \subset [1, P/M] \cap \mathbb{Z}$. The condition $(xy, m) = 1$ enables us, roughly speaking, to assume that $x \equiv y \pmod{m^k}$, and we are effectively able to take first differences in a very efficient manner, by considering $\Psi(x, y) = m^{-k}(x^k - y^k)$ (which is, of course, an integer). We may then exploit the homogeneity of the u_j and v_j, *via* Hölder's inequality, to relate the number of solutions of (0.13) to that of equations of the form (0.12) with $\mathcal{A}$ replaced by $\mathcal{B}$, and s replaced by a range of values not too far from s. In this way we may obtain estimates for the number of solutions of the equation (0.12) iteratively, going (in the simplest version of the method) from $s-1$ to s.

To explain the limitations of Vaughan's method, let us define the modified forward difference operator, Δ_1^*, by

$$\Delta_1^*(f(x); h; m) = m^{-k}(f(x+hm^k) - f(x)),$$

and define Δ_j^* recursively by

$$\Delta_{j+1}^*(f(x); h_1, \dots, h_{j+1}; m_1, \dots, m_{j+1}) = \Delta_1^*(\Delta_j^*(f(x); h_1, \dots, h_j; m_1, \dots, m_j); h_{j+1}; m_{j+1}) .$$

Then classical estimates for generating functions used in the context of Waring's problem depend on taking standard differences repeatedly, and then using estimates stemming from that for the divisor function,

$$d(n) \ll n^{\varepsilon} \text{ for } n \in \mathbb{N}.$$

Thus we consider exponential sums depending on

$$\Delta_j^*(f(z); h_1, \dots, h_j; 1, \dots, 1)$$

with $f(z) = (z - h_1 - \dots - h_j)^k$. In the approach of which Vaughan's method is the most refined version, we consider exponential sums depending on

$$\Delta_j^*(f(z); h_1, \dots, h_j; m, 1, \dots, 1)$$

with $f(z) = (z - h_1 m^k - h_2 - \ldots - h_j)^k$. Thus we have taken one efficient difference, and for the remaining differences we are forced to return to standard differencing.

Our new idea is to further exploit the arithmetic properties of $\mathcal{A}$ so as to continue taking differences, each difference being taken nearly as efficiently as the first. Thus we consider exponential sums depending on

$$\Delta_j^*(f(z); h_1, \ldots, h_j; m_1, \ldots, m_j)$$

with $f(z) = (z - h_1 m_1^k - \ldots - h_j m_j^k)^k$. It transpires that this approach puts no obstacles in the way of previous applications of Vaughan's iterative method, since the improvements effected do not change the character of the auxiliary equations used.

Finally, it should be noted that the ideas involved in our new method are of significance beyond the application in Waring's Problem to which we have put it in this thesis. Thus large improvements may be achieved in estimates stemming from Vinogradov's mean value theorem, in estimates for fractional parts of polynomials, and much else. In particular, the estimate "$G^*(3,2) \leq 14$" may be replaced by "$G^*(3,2) \leq 13$" in Theorem 0.3. These are all matters the author intends to consider in papers subsequent to this thesis.

PART I.

SIMULTANEOUS ADDITIVE EQUATIONS: THE LOCAL PROBLEM.

Modified versions of Chapters 1 and 2 have been accepted for publication by the London Mathematical Society and Mathematika, respectively as the first and third parts in the series "On simultaneous additive equations".

CHAPTER 1.

ON THE P-ADIC SOLUBILITY OF PAIRS OF EQUATIONS, ONE CUBIC AND ONE QUADRATIC.

1. INTRODUCTION.

Let c_{ij} $(1 \le i \le t,\ 1 \le j \le s)$ be rational integers, and k_i $(1 \le i \le t)$ be natural numbers. Consider the simultaneous diophantine equations

$$F_i(\underline{x}) = c_{i1}x_1^{k_i} + \ldots + c_{is}x_s^{k_i} = 0 \quad (1 \le i \le t)\ . \tag{1.1}$$

We say that the simultaneous equations (1.1) satisfy the *p-adic solubility condition* if they have a non-trivial solution in p-adic integers (i.e. a solution with not all the x_i zero). Also, we say that the equations satisfy the *congruence condition* if they satisfy the p-adic solubility condition for every rational prime p.

Define $\Gamma_p^*(\underline{k}) = \Gamma_p^*(k_1,\ldots,k_t)$ to be the least integer r such that for all $s \ge r$, and all c_{ij} $(1 \le i \le t,\ 1 \le j \le s)$, the equations (1.1) satisfy the p-adic solubility condition. Also, define $\Gamma^*(\underline{k})$ to be $\underset{p \text{ prime}}{\text{Sup}}\ \Gamma_p^*(\underline{k})$.

There has been much interest in the problem of evaluating $\Gamma^*(\underline{k})$ over the last quarter of a century, mostly stemming from the pioneering investigations of Davenport and Lewis [1963, 1966, 1967, 1969]. However, it seems that thus far the only cases which have been considered are those in which the k_i are all equal. In this chapter we shall consider the easiest "non-trivial" case with unequal exponents, namely $\Gamma^*(3,2)$, and shall be concerned with the equations

$$\left.\begin{aligned} F(\underline{x}) &= c_1x_1^3 + \ldots + c_sx_s^3 = 0 \\ G(\underline{x}) &= d_1x_1^2 + \ldots + d_sx_s^2 = 0 \end{aligned}\right\} \tag{1.2}$$

where $c_i, d_i \in \mathbb{Z}$ for $i = 1,\ldots,s$. We shall prove:

Theorem 1.1. *We have* $\Gamma^*(3,2) = 11$.

Although this problem is of independent interest, knowledge of the p-adic solubility of the equations (1.1) is also an essential prerequisite to an application of the standard method for dealing with the corresponding rational problem, namely the Hardy-Littlewood method. We shall go on to consider the solubility of the equations (1.2) over the rational integers in Chapters 3 and 4.

Theorem 1.1 may be compared with results obtained by previous workers for small values of t and k_i:

$$\Gamma^*(2) = 5 \text{ (classical)}, \quad \Gamma^*(3) = 7 \text{ (Lewis [1957])},$$

$$\Gamma^*(2,2) = 9 \quad \text{(Demjanov [1956])},$$

$$\Gamma^*(3,3) = 16 \text{ (Davenport and Lewis [1966])}.$$

Cook [1985] has given a more precise analysis of the p-adic solubility of pairs of additive cubic equations, showing that $\Gamma_p^*(3,3) \leq 13$ for $p \neq 7$, with equality holding whenever $p \equiv 1 \pmod 3$. It is worth noting that the exceptional behaviour exhibited by the prime 7 for pairs of cubics is in contrast with the situation for pairs of equations, one cubic and one quadratic, where no prime demonstrates such "exceptional" behaviour.

It is to be hoped that our consideration of this particular example may stimulate interest in the more general problem, and to this end many of our methods are set out in a quite general form. The proof of the theorem is based on generalisations of the methods of Davenport and Lewis [1966, 1967, 1969] and Cook [1985], although owing to the inherent non-linearity of the system (1.2), numerous complications occur.

We begin in §2 by generalising the p-normalisation methods of Davenport and Lewis [1966, 1967, 1969] to our non-linear set-up. The primes 2 and 3 have special difficulties associated with them, and are dealt with separately in §5. The primes $p \equiv 5 \pmod{12}$ can be treated particularly simply, and these we also consider in §5. For primes other than 2 and 3, we are able to use Hensel's Lemma to show in §3 that unless the resultant p-normalised system is of a particular form, then it has a non-trivial p-adic solution. In §4 we use exponential sums to bound the primes for which we are still unable to guarantee a non-trivial solution. Lemma 4.3 enables us to bound the number of solutions of a certain system of congruences rather effectively, and this gives us a bound on the set of primes requiring further investigation of much greater strength than would otherwise be possible. We are then left with a finite set of congruences for which we require solutions non-singular (mod p), and these may be checked using a computer. We mention some of the economies which may be made in this task in §6. Finally, in §7, we conclude the proof of the upper bound implicit in the theorem, and give an example of a system of the form (1.2) for which 10 variables are insufficient to guarantee a non-trivial p-adic solution.

2. THE P-ADIC NORMALISATION OF A SYSTEM OF ADDITIVE FORMS.

Here we attempt to generalise the successful treatment of simultaneous kth power equations given by Davenport and Lewis [1966, 1967, 1969]. Not all of the details will be required for the system (1.2), but the general method is of interest in its own right, and so is included in any case.

Suppose that $k_1, \ldots, k_t$ are positive integers, and $\underline{\underline{F}} = (F_1, F_2, \ldots, F_t)$ is the simultaneous system of $t \le s$ diagonal forms

$$F_i(\underline{\underline{x}}) = F_i(x_1, \ldots, x_s) = a_{i1}x_1^{k_i} + \ldots + a_{is}x_s^{k_i} \quad (1 \le i \le t) \qquad (2.1)$$

with $k = \text{Max}\{ k_i : 1 \le i \le t \}$. We suppose that no $F_i(\underline{\underline{x}})$ is a linear combination of the other $F_j(\underline{\underline{x}})$ (i.e. the system is not linearly dependent).

Let

$$K = \prod_{1 \le j \le t} k_j ,$$

and

$$S = \left\{ (j_1, \ldots, j_t) \in \{1, \ldots, s\}^t : j_i \ne j_{i'} \text{ for } i \ne i' \right\}$$

so that $|S| = s(s-1)\ldots(s-t+1) = M$, say. When $\sigma = (j_1, \ldots, j_t) \in S$ define

$$D_\sigma(\underline{\underline{F}}) = \det \left(a_{i j_m}^{k_i'} \right)_{1 \le i, m \le t}$$

where $k_i' = K/k_i$. Then we define

$$\partial(\underline{\underline{F}}) = \prod_{\sigma \in S} D_\sigma(\underline{\underline{F}}) . \qquad (2.2)$$

As an example, for the system (1.2) we have

$$\partial(F, G) = \prod_{i \ne j} (c_i^2 d_j^3 - c_j^2 d_i^3) .$$

The apparently rather complicated form of ∂, with powers attached to the coefficients, is required to ensure that $\partial(\underline{\underline{F}})$ is in some sense "homogeneous" with respect to the coefficients of $\underline{\underline{F}}$. This homogeneity leads to results of the form obtained by Davenport and Lewis [1966, 1967, 1969].

To simplify notation, we define an equivalence relation $\mathcal{R}$ on $\{1, \ldots, t\}$ by

$$i \mathcal{R} j \text{ if and only if } k_i = k_j .$$

We denote the equivalence class containing i by $[i]$, and the set of equivalence classes by $I = \{1, \ldots, t\}/\mathcal{R}$.

Lemma 2.1. *Given a system* $\underline{\underline{F}}$ *of the form* (2.1), $\partial(\underline{\underline{F}})$ *satisfies the following:*

(i) if $\nu_1,\dots,\nu_s$ *are integers, and*

$$\underline{\underline{F}}' = \underline{\underline{F}}(p^{\nu_1}x_1,\dots,p^{\nu_s}x_s),$$

then

$$\partial(\underline{\underline{F}}') = p^{tKM\nu/s}\partial(\underline{\underline{F}}) ,$$

where $\nu = \nu_1 + \dots + \nu_s$;

(ii) If λ_{ij} $(1 \le i,j \le t)$ *are rational numbers, and the system* $\underline{\underline{F}}''$ *is defined by*

$$F_i''(\underline{\underline{x}}) = \sum_{j \in [i]} \lambda_{ij}F_j(\underline{\underline{x}}) , \quad (i = 1,\dots,t),$$

then we have

$$\partial(\underline{\underline{F}}'') = J^M.\partial(\underline{\underline{F}}) ,$$

where

$$J = \prod_{\alpha \in I} det \left[\lambda_{ij}^{k_i'} \right]_{i,j \in \alpha} .$$

Proof: (i) Write ρ_i for p^{ν_i}, and a'_{ij} for the coefficient of $x_j^{k_i}$ in $F_i'(\underline{\underline{x}})$. Then if $\sigma = (j_1,\dots,j_t) \in S$, we have

$$\begin{aligned} D_\sigma(\underline{\underline{F}}') &= \det \left[a'^{k_i'}_{ij_m} \right]_{1 \le i,m \le t} \\ &= \det \left[\rho_{j_m}^K a_{ij_m}^{k_i'} \right]_{1 \le i,m \le t} \\ &= \rho_{j_1}^K \dots \rho_{j_t}^K .\det \left[a_{ij_m}^{k_i'} \right]_{1 \le i,m \le t} \end{aligned}$$

Then

$$\partial(F') = \prod_{\sigma \in S} \left[D_\sigma(\underline{\underline{F}}).\rho_{j_1}^K \dots \rho_{j_t}^K \right]$$

and $\sum_{\sigma \in S} (\nu_{j_1} + \dots + \nu_{j_t}) = tM\nu/s.$

(ii) Write a''_{ij} for the coefficient of $x_j^{k_i}$ in $F_i''(\underline{\underline{x}})$. Then if $\sigma = (j_1,\dots,j_t) \in S$, we have

$$\det \left[a''^{k_i'}_{ij_m} \right]_{1 \le i,m \le t} = (\det H).\det \left[a_{ij_m}^{k_i'} \right]_{1 \le i,m \le t}$$

where the rows and columns of the matrix H can be rearranged to give the block matrix

$$\begin{bmatrix} C_1 & & & \\ & C_2 & & \\ & & \ddots & \\ & & & C_r \end{bmatrix}$$

where $r = |I|$, and each C_i corresponds to a unique $\alpha \in I$, having determinant

$$\det \left[\lambda_{ij}^{k_i'} \right]_{i,j \in \alpha}.$$

The result follows on noting that the number of factors in the product (2.2) is M, which is even for $t > 1$.

This completes the proof of the lemma.

We shall follow Davenport and Lewis [1966, 1967, 1969], and call two systems of the form (2.1), with integral coefficients, *equivalent* if one system can be obtained from the other by a combination of the operations (i) and (ii) of Lemma 2.1; here we shall always demand that the determinant $J \neq 0$, so that the operation (ii) is non-trivial. Notice that the operations (i) and (ii) are commutative. Also, if the system $\underline{\underline{F}}(\underline{x}) = 0$ has a non-trivial solution over $\mathbb{Q}_p$, then so does any system equivalent to $\underline{\underline{F}}$.

We shall assume throughout §§2-5 that $\partial(\underline{\underline{F}}) \neq 0$ (we shall see in §7 that this is not too demanding a condition for us to fulfill). This property is plainly preserved under equivalence.

A system $\underline{\underline{F}}$ will be called *p-normalised* if the power of the rational prime p dividing $\partial(\underline{\underline{F}})$ is minimal amongst the set of all forms equivalent to $\underline{\underline{F}}$. That this is possible follows because this power is non-negative. A p-normalised system is not unique, for any operation of type (ii) with J not divisible by p transforms a

p-normalised system into another p-normalised system. Such a transformation is known as a *unimodular change of basis*.

To any system $\underline{\underline{F}}$ with integral coefficients there corresponds a system $\underline{\underline{F}}^*$ with coefficients in $\mathbb{Z}/p\mathbb{Z}$, these coefficients being congruent (mod p) to the corresponding coefficients of $\underline{\underline{F}}$. Plainly, there may be variables explicit in $\underline{\underline{F}}$ but not in $\underline{\underline{F}}^*$. The number of variables occurring explicitly in a form is known as its *rank*.

Lemma 2.2. *A p-normalised system $\underline{\underline{F}}$ can be written in the form*

$$F_i = F_{i,0} + pF_{i,1} + \ldots + p^{k-1}F_{i,k-1} \quad (i = 1,\ldots,t)$$

where the $F_{i,j}$ are forms in m_j variables, and these sets are disjoint for $j = 0,\ldots,k-1$. Moreover, each of the m_j variables occurs in one at least of $F_{i,j}$ $(i = 1,\ldots,t)$ with a coefficient not divisible by p.

The following inequalities hold:

(i) we have

$$m_0 + \ldots + m_{j-1} \geq js/k \quad \text{for } j = 1,\ldots,k_{\min},$$

where k is the harmonic mean of the k_i and $k_{\min}$ is the least of the k_i.

(ii) suppose that $\alpha \in I$ with $k_i = h$ for $i \in \alpha$, and we form any L linear combinations

$$f_i^*(\underline{\underline{x}}) = \sum_{j \in \alpha} \lambda_{ij}^* F_j^*(\underline{\underline{x}}) \quad (i = 1,\ldots,L)$$

where $L \leq |\alpha|$, and the $(\lambda_{ij}^*)_{j \in \alpha}$ are linearly independent over $\mathbb{Z}/p\mathbb{Z}$. Then, denoting by $Q(L,h)$ the number of variables occurring explicitly in one at least of these combinations, we have

$$Q(L,h) \geq Ls/th \quad \text{for } L = 1,\ldots,|\alpha|.$$

(iii) denoting by $q_{i,j}$ *the number of variables occurring explicitly in the form* $F^*_{i,j}$, *we have*

$$m_0 + m_1 + \ldots + m_{j-1} + q_{i,j} \geq js/k + s/tk_i$$

$$\text{for } 0 \leq j \leq \text{Min}\{ k_{\min}, k_i - 1 \}.$$

Proof: It is plainly possible to express a p-normalised system in the form

$$F_i = F_{i,0} + pF_{i,1} + \ldots \quad (i = 1, \ldots, t)$$

where we put in $F_{i,j}$ those terms $a_{in} x_n^{k_i}$ for which p^j is the highest power of p dividing every a_{in} $(i = 1, \ldots, t)$. Then the sets of variables occurring in $F_{i,j}$ $(j = 0, 1, \ldots)$ are plainly disjoint.

But if $j \geq k$, the forms $F_{i,j}$ are empty. For if $\underline{\underline{F}}$ is a p-normalised system then the power of p dividing $\partial(\underline{\underline{F}})$ is minimal, and so if $a_{in} x_n^{k_i}$ were terms in $F_i(\underline{\underline{x}})$ $(i = 1, \ldots, t)$ with a_{in} all divisible by p^k, we could reduce the power of p dividing $\partial(\underline{\underline{F}})$ by an operation of type (i), namely that of putting $x_n = p^{-1} x'_n$, while preserving the integrality of the coefficients.

(i) let $x_1, \ldots, x_m$, where $m = m_0 + \ldots + m_{j-1}$, denote the variables in $F_{i,0}, \ldots, F_{i,j-1}$ $(1 \leq j \leq k_{\min})$. Then the system

$$F'_i(\underline{\underline{x}}) = p^{-j} F_i(px_1, \ldots, px_m, x_{m+1}, \ldots, x_s) \quad (i = 1, \ldots, t)$$

has integral coefficients and is equivalent to the system $\underline{\underline{F}}$. By Lemma 2.1, we have

$$\partial(\underline{\underline{F'}}) = p^{-jM.(k'_1 + \ldots + k'_t) + tMKm/s} \partial(\underline{\underline{F}}) .$$

Then by the definition of a p-normalised system,

$$m \geq js.(k'_1 + \ldots + k'_t)/(tK)$$

$$= js/k .$$

(ii) let $\alpha \in I$ with $k_i = h$ for $i \in \alpha$, and rearrange equations so that $\alpha = \{1, \ldots, r\}$. Take any L linear combinations of the $F_j(\underline{\underline{x}})$ $(j \in \alpha)$,

$$f_i(\underline{\underline{x}}) = \sum_{j=1}^{r} \lambda_{ij} F_j(\underline{\underline{x}}) , \quad \text{for } i = 1, \ldots, L,$$

where $1 \le L \le r$, and the $(\lambda_{ij})_{1 \le j \le r}$ are linearly independent over $\mathbb{Z}/p\mathbb{Z}$. Each set can be completed to give a set of r linear combinations independent (mod p). Then $f_1,\dots,f_r$ are derived from the F_j $(1 \le j \le r)$ by operation (ii) of Lemma 2.1 with J not divisible by p. Let $Q = Q(L,h)$ be the number of variables that occur in at least one of $f_1,\dots,f_L$ with coefficient not divisible by p, and take these variables to be $x_1,\dots,x_Q$. Then the system

$$\left.\begin{aligned} f'_i(\underline{x}) &= p^{-1}f_i(px_1,\dots,px_Q,x_{Q+1},\dots,x_s) \qquad i = 1,\dots,L \\ f'_i(\underline{x}) &= f_i(px_1,\dots,px_Q,x_{Q+1},\dots,x_s) \qquad i = L+1,\dots,r \end{aligned}\right\}$$

has integral coefficients, and is derived from $\underline{f}$ by a combination of operation (i) with $\nu = Q$, and operation (ii) with $J = p^{-L}J_0$, where J_0 is not divisible by p. Then on writing $h' = K/h$, we have

$$\begin{aligned} \partial(\underline{f}') &= p^{tKMQ/s - MLh'}J_0^M\,\partial(\underline{f}) \\ &= p^{tKMQ/s - MLh'}J_0^M\left(\det(\lambda_{ij}^{h'})_{i,j\in a}\right)^M\partial(\underline{F}) . \end{aligned}$$

Thus

$$|\partial(\underline{f}')|_p = p^{MLh' - tKMQ/s}|\partial(\underline{F})|_p .$$

But $\underline{f}'$ is equivalent to $\underline{F}$, and so by the definition of a p-normalised system we have

$$tKMQ/s \ge MLh',$$

and hence

$$Q \ge Ls/th .$$

(iii) Suppose that the number of variables occurring explicitly in $F^*_{n,j}$ is q. Let these variables be $x_{m+1},\dots,x_{m+q}$, where $m = m_0 + \dots + m_{j-1}$, and let $x_1,\dots,x_m$ be the variables occurring in $F_{i,0},\dots,F_{i,j-1}$ $(1 \le i \le t)$. Then for $0 \le j \le \mathrm{Min}\{k_{min}, k_n-1\}$, the system

$$\left.\begin{aligned} F''_i(\underline{x}) &= p^{-j}F_i(px_1,\dots,px_{m+q},x_{m+q+1},\dots,x_s) \qquad i \ne n, \\ F''_n(\underline{x}) &= p^{-j-1}F_n(px_1,\dots,px_{m+q},x_{m+q+1},\dots,x_s) \end{aligned}\right\}$$

has integral coefficients, and is equivalent to $\underline{\underline{F}}$. But by Lemma 2.1, we have

$$\partial(\underline{\underline{F''}}) = p^{tKM(m+q)/s - (j(k'_1+\ldots+k'_t) + k'_n)M} \partial(\underline{\underline{F}}) ,$$

and hence, by the definition of a p-normalised system,

$$tKM(m+q)/s \geq M(j(k'_1 + \ldots + k'_t) + k'_n) .$$

Then

$$m + q \geq sj/k + s/(tk_n) .$$

This completes the proof of the lemma.

Notice that as far as the p-normalisation of a system goes, a measure of its "average" degree is simply the harmonic mean of the k_i.

Lemma 2.3. *Each of the $F_{i,j}$ can be written in the form*

$$F_{i,j} = F_{i,j,0} + pF_{i,j,1} + p^2F_{i,j,2} + \ldots$$

for $i = 1,\ldots,t$ and $j = 0,\ldots,k-1$, where the $F_{i,j,h}$ are forms in $r_{i,j,h}$ variables, and these sets of variables are disjoint for $h = 0,1,2,\ldots$. Moreover, each of the $r_{i,j,h}$ variables occuring in $F_{i,j,h}$ have coefficient not divisible by p.

Using the notation of Lemma 2.2, the following hold:

(i) $\sum_{h=0}^{\infty} r_{i,j,h} = m_j$,

(ii) $r_{i,j,0} = q_{i,j}$,

(iii) if we define $R_{i,H}$ by

$$R_{i,H} = \sum_{j=0}^{H} \left[\sum_{h=0}^{H-j} r_{i,j,h} \right] ,$$

then we have

$$R_{i,H} \geq (H+1)s/(tk_i) , \quad \text{for } H = 0,\ldots,k_i-1.$$

Proof: All the results of the lemma, except part (iii), are immediate from Lemma 2.2.

Let $H \le k_n - 1$, $R = R_{n,H}$, and let $x_1, \dots, x_R$ denote the variables occurring in the form

$$\left[F_{n,0,0} + pF_{n,0,1} + \dots + p^H F_{n,0,H}\right] + p\left[F_{n,1,0} + pF_{n,1,1} + \dots + p^{H-1}F_{n,1,H-1}\right] + \dots + p^H F_{n,H,0}$$

i.e. the variables explicit in the form F_n reduced (mod p^{H+1}). Then the system

$$\left.\begin{aligned} F'_n(\underline{x}) &= p^{-H-1}F_n(px_1, \dots, px_R, x_{R+1}, \dots, x_s), \\ F'_i(\underline{x}) &= F_i(px_1, \dots, px_R, x_{R+1}, \dots, x_s), \text{ for } i \ne n, \end{aligned}\right\}$$

has integral coefficients and is equivalent to the system $\underline{F}$. By Lemma 2.1, we have

$$\partial(\underline{F}') = p^{-(H+1)Mk'_n + tKMR/s}\,\partial(\underline{F}).$$

Then, by the definition of a p-normalised system,

$$R \ge (H+1)sk'_n/(tK) = (H+1)s/(tk_n).$$

This completes the proof of the lemma.

Using Lemmata 2.2 and 2.3, we obtain for the system (1.2) with $s \ge 11$ (with an obvious minor change in notation):

$$\left.\begin{aligned} m_0 \ge 5, \quad q_{F,0} = R_{F,0} \ge 2, \quad q_{G,0} = R_{G,0} \ge 3, \\ m_0 + m_1 \ge 10, \quad m_0 + q_{F,1} \ge 7, \quad m_0 + q_{G,1} \ge 8, \\ R_{F,1} \ge 4, \quad R_{G,1} \ge 6, \quad R_{F,2} \ge 6. \end{aligned}\right\} \qquad (2.3)$$

3. TACKLING THE "STANDARD" CASES WHEN P > 5.

On the assumptions $\partial(F,G) \neq 0$ and $s \geq 11$, we may rearrange variables and change notation (for this section only) to write

$$\left.\begin{aligned} F_0^* &= a_1x_1^3+\ldots+a_ux_u^3+\ b_1y_1^3+\ldots+b_vy_v^3 \\ G_0^* &= \qquad\qquad c_1y_1^2+\ldots+c_vy_v^2+\ d_1z_1^2+\ldots+d_wz_w^2 \end{aligned}\right\} \equiv 0 \pmod{p} \qquad (3.1)$$

where none of the a_1,b_1,c_1,d_1 are divisible by p, and by (2.3) we have

$$m_0 = u+v+w \geq 5, \quad q_{G,0} = v+w \geq 3, \quad q_{F,0} = u+v \geq 2 . \qquad (3.2)$$

In this section we shall show that subject to the hypothesis

(H) for all congruences of the form (3.1) satisfying $u+v+w \geq 5$, $u \leq 1$, *and* $w \leq 1$, *we have a solution non-singular (mod p),*

when $p > 5$, all p-normalised systems of the form (1.2) have a non-trivial p-adic solution. We achieve this by showing that we may set certain of the variables to zero in such a way that the resulting pair of equations is equivalent to a system having a solution non-singular (mod p). By an application of Hensel's Lemma, such a solution gives us a non-trivial p-adic solution to the system.

Given F and G of the form (1.2), and $\underline{x} \in \mathbb{Z}_p^s$, we shall define

$$\Delta(i,j) = \Delta(F,G;i,j) = 6x_ix_j(c_id_jx_i - c_jd_ix_j), \text{ and}$$

$$\Delta^*(F,G) = \max_{1\leq i<j\leq s} \{|\Delta(i,j)|_p\} .$$

Thus if $\underline{x}$ is a solution to the congruences $F \equiv G \equiv 0 \pmod{p}$, it is non-singular (mod p) when $\Delta^*(F,G) = 1$.

We shall require some well-known lemmata.

The following version of Hensel's Lemma plainly extends to systems in $n > 2$ variables, by fixing $n-2$ of them.

Lemma 3.1 (*see Greenberg* [1969], *Proposition* (5.20)). *Suppose that* $F(X,Y)$, $G(X,Y) \in \mathbb{Z}_p[X,Y]$, *and that* $a_0, b_0 \in \mathbb{Z}_p$ *satisfy*

$$Max\ \{|F(a_0,b_0)|_p\ , |G(a_0,b_0)|_p\} < |\Delta_0|_p^2\ ,$$

where

$$\Delta_0 = \Delta(F,G)|_{(a_0,b_0)} = \left(\frac{\partial F}{\partial X}\frac{\partial G}{\partial Y} - \frac{\partial G}{\partial X}\frac{\partial F}{\partial Y} \right)_{(a_0,b_0)}$$

is non-zero, $\partial F/\partial X$ *etc. being formal derivatives. Then there is a unique* $(a,b) \in \mathbb{Z}_p \times \mathbb{Z}_p$ *such that* $F(a,b) = G(a,b) = 0$ *and*

$$Max\{|a - a_0|_p\ ,\ |b - b_0|_p\} \le p^{-1}|\Delta_0|_p\ .$$

Lemma 3.2 (*Chowla, Mann and Straus* [1959]). *For any positive integer* k, *and rational prime* p, *write* $\delta = (k, p-1)$. *If* $p > 2\delta+1$ *and* $\alpha_1\alpha_2\ldots\alpha_n \not\equiv 0 \pmod p$, *then*

$$\alpha_1 x_1^k + \ldots + \alpha_n x_n^k$$

represents at least $Min\{\ p\ ,\ 1 + ((2n-1)(p-1)/\delta)\ \}$ *residues* (mod p).

Corollary 3.2.1. *Let* $p > 7$ *and* $\alpha_1\alpha_2 \not\equiv 0 \pmod p$. *Then* $\alpha_1 x_1^3 + \alpha_2 x_2^3$ *represents all residues mod* p.

Corollary 3.2.2. *Let* $p > 5$ *and* $\alpha_1\alpha_2 \not\equiv 0$ (*mod* p). *Then* $\alpha_1 x_1^2 + \alpha_2 x_2^2$ *represents all residues mod* p.

Lemma 3.3 (*Lewis* [1957], *Theorem* 1). *The congruence*

$$ax^3 + by^3 + cz^3 \equiv 0\ (mod\ p)$$

is always soluble with one at least of $x,y,z \not\equiv 0$ (*mod* p).

Lemma 3.4 (*see Davenport and Lewis* [1966] *Lemma* 5). *Suppose that* $abc \not\equiv 0$ (*mod* p). *Then the congruence* $ax^3 + by^3 \equiv c$ (*mod* 7) *is soluble unless* $b \equiv \pm a$ *and* $c \equiv \pm 4a$.

Notice that, in particular, the corollaries to Lemma 3.2 say that all non-zero residues are represented *non-trivially* by the respective forms. Notice also that if $abc \not\equiv 0 \pmod{p}$, then the conclusion of Lemma 3.3 implies that at least two of x,y,z are non-zero (mod p).

We first consider the cases where $q_{F,0} \geq 3$ and $p > 7$. We then go on to deal with the prime 7, and then the cases with $q_{F,0} = 2$.

Lemma 3.5. *Let $p > 7$. Then all p-normalised systems of the form* (1.2) *satisfying $s \geq 11$, $q_{F,0} \geq 3$, and either $u > 1$ or $w > 1$, have a non-trivial solution over $\mathbb{Z}_p$.*

Proof: We divide into cases according to the value of v.

(i) $v = 0$.

Then by (3.2), $w \geq 3$ and $u \geq 3$. By Lemma 3.3 we can solve the congruence

$$a_1x_1^3 + \ldots + a_ux_u^3 \equiv 0 \pmod{p}$$

with one at least of the variables non-zero, say x_i. Also, by Corollary 3.2.2, we can solve the congruence

$$d_1z_1^2 + \ldots + d_wz_w^2 \equiv 0 \pmod{p}$$

independently with one at least of the variables non-zero, say z_j. Then the system is soluble (mod p) with

$$\Delta^*(F,G) \geq |6x_iz_j(a_id_jx_i)|_p = 1.$$

(ii) $v = 1$.

Then by (3.2), $w \geq 2$ and $u \geq 2$. Set $y_1 = 1$. By Corollary 3.2.1 we can solve the congruence

$$a_1x_1^3 + \ldots + a_ux_u^3 + b_1 \equiv 0 \pmod{p}$$

with one at least of the variables non-zero, say x_i. Also, by Corollary 3.2.2, we can solve the congruence

$$c_1 + d_1z_1^2 + \ldots + d_wz_w^2 \equiv 0 \pmod{p}$$

independently with one at least of the variables non-zero, say z_j. Then the system is soluble (mod p) with

$$\Delta^*(F,G) \geq |6x_i z_j (a_i d_j x_i)|_p = 1.$$

(iii) $v \geq 2$.

There are two cases:

(a) $w \geq 2$.

(α) $p \equiv 1 \pmod 3$.

Since $q_{F,0} \geq 3$, by Lemma 3.3 we can solve the congruence

$$a_1 x_1^3 + \ldots + a_u x_u^3 + b_1 y_1^3 + \ldots + b_v y_v^3 \equiv 0 \pmod p$$

non-trivially (here the x_i may not be present if $u = 0$). Further, we can plainly ensure that at least one of the y_i is non-zero (for example, by taking variables other than y_1 and y_2 to be suitable residues). Let $1, c, c'$ be the three distinct cubic roots of unity (mod p). Then fixing the y_i so as to solve the cubic, we can plainly find $g_i \in \{1, c, c'\}$ for $i = 1, \ldots, v$, so that

$$c_1 (g_1 y_1)^2 + \ldots + c_v (g_v y_v)^2 \not\equiv 0 \pmod p. \tag{3.3}$$

Then we have a solution to the cubic with the quadratic (3.3) non-zero.

(β) $p \not\equiv 1 \pmod 3$ and $u > 0$.

Then every non-zero residue (mod p) is a cubic residue. We may put $x_2 = \ldots = x_u = y_2 = \ldots = y_v = 0$ and solve the congruence $a_1 x_1^3 + b_1 y_1^3 \equiv 0 \pmod p$ non-trivially.

(γ) $p \not\equiv 1 \pmod 3$ and $u = 0$.

Then $v \geq 3$. Also, we have $b_i \equiv \beta_i^3 \pmod p$ for some $\beta_i \not\equiv 0$ ($i = 1,2,3$). But if

$$c_1 \beta_2^2 + c_2 \beta_1^2 \equiv c_2 \beta_3^2 + c_3 \beta_2^2 \equiv c_3 \beta_1^2 + c_1 \beta_3^2 \equiv 0 \pmod p,$$

then

$$c_1 (\beta_1 \beta_2 \beta_3)^2 \equiv -c_2 (\beta_1^2 \beta_3)^2 \equiv c_3 (\beta_1^2 \beta_2)^2 \equiv -c_1 (\beta_1 \beta_2 \beta_3)^2 \pmod p .$$

Then $(\beta_1\beta_2\beta_3)^2 \equiv 0 \pmod{p}$, which is a contradiction. So one of $(\beta_2,-\beta_1,0)$, $(0,\beta_3,-\beta_2)$ and $(-\beta_3,0,\beta_1)$ is a solution in (y_1,y_2,y_3) to the congruence

$$b_1y_1^3 + b_2y_2^3 + b_3y_3^3 \equiv 0 \pmod{p},$$

such that $c_1y_1^2 + c_2y_2^2 + c_3y_3^2 \not\equiv 0 \pmod{p}$.

In any of the cases (α), (β) and (γ), we may solve the cubic congruence in (3.1) in such a way that

$$c_1y_1^2 + \ldots + c_vy_v^2 \not\equiv 0 \pmod{p},$$

and so that at least one of the y_i is non-zero, say y_j. But by Corollary 3.2.2, fixing these values of y_i, we may solve the quadratic

$$d_1z_1^2 + \ldots + d_wz_w^2 \equiv -(c_1y_1^2 + \ldots + c_vy_v^2) \not\equiv 0 \pmod{p},$$

with one at least of the z_i non-zero, say z_k. Then the system is soluble with

$$\Delta^*(F,G) = |6y_jz_k(b_jd_ky_j)|_p = 1.$$

(b) $u \geq 2$ and $w \leq 1$.

Then by (3.2), $v \geq 3-w$. By Corollary 3.2.2 we can solve the congruence

$$c_1y_1^2 + \ldots + c_vy_v^2 + d_1z_1^2 \equiv 0 \pmod{p}$$

non-trivially (here the term in z_1 may not be present if $w = 0$). Further, we can ensure that at least one of the y_i is non-zero (for example, by setting the variables other than y_1 and y_2 to be suitable residues), say y_1. Notice that by replacing any y_i by $-y_i$, we still have a solution to the quadratic. Fix these values of y_i, and consider the cubic form

$$a_1x_1^3 + a_2x_2^3 .$$

By Corollary 3.2.1, since $p > 7$ this form fails to represent non-trivially at most one reduced residue (mod p), namely the zero residue. But for suitable choices of + and - signs, there are at

least two distinct reduced residues amongst

$$\pm b_1 y_1^3 \pm \ldots \pm b_v y_v^3$$

(because one at least of the y_i is non-zero). Then one at least of these two distinct residues provides a non-trivial solution in x_1 and x_2 to the congruence

$$a_1 x_1^3 + a_2 x_2^3 \pm b_1 y_1^3 \pm \ldots \pm b_v y_v^3 \equiv 0 \pmod{p}.$$

So we may suppose that x_1 is non-zero, and therefore that the system is soluble with

$$\Delta^*(F,G) \geq |6x_1 y_1 (a_1 c_1 x_1)|_p = 1.$$

Then in cases (i), (ii) and (iii) we have established that the congruences (3.1) have a solution non-singular (mod p). We may therefore apply Hensel's Lemma to deduce that the equations (1.2) have a non-trivial solution over $\mathbb{Z}_p$.

This completes the proof of the lemma.

The prime 7 is problematic because the congruence $x^3 + y^3 + 4z^3 \equiv 0 \pmod 7$ has solutions only when $z \equiv 0 \pmod 7$. We now modify the ideas above to by-pass this eventuality.

Lemma 3.6. *All 7-normalised systems of the form* (1.2) *satisfying* $s \geq 11$, $q_{F,0} \geq 3$, *and either* $u > 1$ *or* $w > 1$ *have a non-trivial solution over* $\mathbb{Z}_7$.

Proof: We divide into cases according to the value of v.

(i) $v = 0$.

The proof follows in precisely the same way as in Lemma 3.5.

(ii) $v = 1$.

Then by (3.2), $w \geq 2$ and $u \geq 2$. By Lemma 3.4, we can solve the congruence

$$a_1 x_1^3 + \ldots + a_u x_u^3 + b_1 y_1^3 \equiv 0 \pmod 7$$

with $y_1 \not\equiv 0 \pmod 7$, and at least one of the x_i non-zero, *unless* $u = 2$, $a_2 \equiv \pm a_1$ and $b_1 \equiv \pm 4a_1$. If the latter is not the case, the proof proceeds as in Lemma 3.5 (ii). However, if the latter is the case, by a change of variables and a temporary change in notation, we may suppose that the system is

$$\left.\begin{aligned} F^* &= x_1^3 + x_2^3 + 4y_1^3 \\ G^{**} &= 7(b_1x_1^2 + b_2x_2^2) + c_1y_1^2 + d_1z_1^2 + \ldots + d_tz_t^2 \end{aligned}\right\}$$

where 7 does not divide c_1, at most 7 divides $d_1, \ldots, d_t$, and by (2.3) we have

$$t = m_0 + q_{G,1} - q_{F,0} \geq 5.$$

Here, F^* denotes the variables explicit in $F \pmod p$, and G^{**} those explicit in $G \pmod{p^2}$.

Put $y_1 = 0$. If three or more of $d_1, \ldots, d_t$ are not divisible by 7, say $d_1, \ldots, d_w$, then we can solve the congruences

$$\left.\begin{aligned} F_0^* &= x_1^3 + x_2^3 \equiv 0 \\ G_0^* &= d_1z_1^2 + \ldots + d_wz_w^2 \equiv 0 \end{aligned}\right\} \pmod 7$$

independently with $x_1 = 1$, $x_2 = -1$, and one at least of the z_i non-zero, say z_j. Then the system is soluble with

$$\Delta^*(F,G) \geq |6x_1z_j(d_jx_1)|_7 = 1.$$

If only two of $d_1, \ldots, d_t$ are not divisible by 7, without loss of generality d_{t-1} and d_t, then put $z_{t-1} = z_t = 0$ and divide G through by 7. We then obtain a system of the form

$$\left.\begin{aligned} F_0^* &= x_1^3 + x_2^3 \equiv 0 \\ G'^*_0 &= C_1x_1^2 + C_2x_2^2 + d_1z_1^2 + \ldots + d_wz_w^2 \equiv 0 \end{aligned}\right\} \pmod 7 \qquad (3.4)$$

where 7 does not divide any of the d_i, the C_i are integers, and $w \geq 3$. We can solve the cubic congruence by putting $x_1 = 1$ and $x_2 = -1$. Then by Corollary 3.2.2, we may solve the quadratic congruence for the z_i, say with z_j non-zero. So the reduced system is soluble with

$$\Delta^*(F,G) \geq |6x_1z_j(d_jx_1)|_p = 1.$$

So in this case the reduced system has a solution non-singular (mod 7), and hence the equations (1.2) have a non-trivial solution over $\mathbb{Z}_7$, by Hensel's Lemma.

(iii) $v \geq 2$.

There are two cases:

(a) $w \geq 2$.

The same argument as in Lemma 3.5(iii)(a)(α) applies.

(b) $u \geq 2$ and $w \leq 1$.

By (3.2), $v \geq 3 - w$. Observe that by a multiplicative change of variables, $c_1y_1^2 + c_2y_2^2 \pmod 7$ can be transformed into one of the forms

$$\pm(y_1^2 + y_2^2) \text{ or } \pm(y_1^2 - y_2^2) .$$

By taking certain of the variables to be zero, we have

$$c_1y_1^2 + \ldots + c_vy_v^2 + d_1z_1^2 + \ldots + d_wz_w^2 = c_1y_1^2 + c_2y_2^2 + dz^2,$$

where we have adopted the convention of writing dz^2 for $d_1z_1^2$ if $w = 1$, and for $c_3y_3^2$ if $w = 0$. Then $c_1y_1^2 + c_2y_2^2 + dz^2 \equiv 0 \pmod 7$ is equivalent to one of the congruences

$$\begin{aligned} y_1^2 + y_2^2 + z^2 &\equiv 0, \text{ with solution } (1,2,4), \\ +(y_1^2 - y_2^2) + z^2 &\equiv 0, \text{ with solution } (2,1,2), \\ -(y_1^2 - y_2^2) + z^2 &\equiv 0, \text{ with solution } (1,2,2), \\ -(y_1^2 + y_2^2) + z^2 &\equiv 0, \text{ with solution } (2,2,1). \end{aligned}$$

Adopt the convention of writing by^3 for $b_3y_3^3$ if $w = 0$, and for 0 if $w = 1$. Then in each of these cases, the quadratic is soluble with y_1, y_2 and z each non-zero, so that for suitable choices of + and -, the expression

$$\pm b_1y_1^3 \pm b_2y_2^3 \pm by^3$$

takes at least three distinct values. Then the congruence

$$c_1y_1^2 + c_2y_2^2 + dz^2 \equiv 0$$

may be solved in such a way that, by Lemma 3.3, the congruence

$$a_1x_1^3 + a_2x_2^3 + b_1y_1^3 + b_2y_2^3 + by^3 \equiv 0$$

has a solution with at least one of the x_i non-zero, say x_1. Then the system is soluble with

$$\Delta^*(F,G) \geq |6x_1y_1(a_1c_1x_1)|_7 = 1.$$

Then in cases (i), (ii) and (iii), we have deduced that by setting certain of the variables to zero, the equations (1.2) are equivalent to a system which either has a non-trivial 7-adic solution, or else has a solution non-singular (mod 7). An application of Hensel's Lemma now gives us a non-trivial 7-adic solution to the equations (1.2).

This completes the proof of the lemma.

So subject to hypothesis *(H)*, and the assumptions $\partial(F,G) \neq 0$, $s \geq 11$ and $q_{F,0} \geq 3$, we have shown that every system of the form (1.2) with $p > 5$ has a non-trivial p-adic solution. We now go on to show that when $q_{F,0} = 2$, we nonetheless have a non-trivial p-adic solution.

Lemma 3.7. *Let $p > 5$, and suppose that hypothesis (H) holds. Then all p-normalised systems of the form* (1.2) *with $s \geq 11$ and $q_{F,0} = 2$ have a non-trivial solution over $\mathbb{Z}_p$.*

Proof: We use the inequalities (2.3) to divide into cases:

(i) Suppose that $R_{F,0} = 2$ and $R_{F,1} \geq 5$.

We may rearrange variables and change notation to write

$$\left.\begin{aligned} F^{**} &= A_1X_1^3 + A_2X_2^3 + p(a_1x_1^3 + \ldots + a_ux_u^3 + b_1y_1^3 + \ldots + b_vy_v^3) \\ G^{*} &= B_1X_1^2 + B_2X_2^2 + c_1y_1^2 + \ldots + c_vy_v^2 + d_1z_1^2 + \ldots + d_wz_w^2 \end{aligned}\right\}$$

where p does not divide any of the a_i, b_i, c_i, d_i, A_i, the B_i are integers, and from (2.3) we have

$$u + v = R_{F,1} - R_{F,0} \geq 3,$$

$$u + v + w = m_0 + q_{F,1} - q_{F,0} \geq 5,$$

$$v + w = m_0 - q_{F,0} \geq 3.$$

Here F^{**} denotes the variables explicit in $F \pmod{p^2}$, and G^* denotes those explicit in $G \pmod{p}$.

Independently of $R_{F,1}$, if $p = 7$ and $A_1 \equiv \pm A_2 \pmod 7$, then since $v+w \geq 3$, we have a solution to the system non-singular (mod 7), as for (3.4). Henceforth we may therefore assume that if $p = 7$, then $A_1 \not\equiv \pm A_2 \pmod 7$.

We now set $X_1 = X_2 = 0$, and consider the congruences obtained by a change of variables:

$$\left.\begin{aligned} F'^* &= (p^{-1}F)^* \\ G'^* &= G^* \end{aligned}\right\} \equiv 0 \pmod p$$

This system of congruences is either of a form in which hypothesis *(H)* applies, or else is of the same form as the systems we considered in Lemmata 3.5 and 3.6, and with $q_{F',0} \geq 3$. The system therefore has a solution non-singular (mod p), and hence the equations (1.2) must have a non-trivial p-adic solution, *unless* $p = 7$, $u = 2$, $v = 1$, $w = 2$, $a_2 \equiv \pm a_1 \pmod 7$ and $b_1 \equiv \pm 4a_1 \pmod 7$.

But in this last case we consider the congruences obtained by setting $X_1 = X_2 = y_1 = z_1 = z_2 = 0$, and making a change of variables:

$$\left.\begin{aligned} F'^* &= (7^{-1}F)^* \\ G'^* &= (7^{-1}G)^* \end{aligned}\right\} \equiv 0 \pmod 7.$$

By a rearrangement of variables and change in notation, by (2.3) the system takes the form (3.4) with

$$w + 2 = (m_0 + m_1) - m_0 \geq 5.$$

Then the congruences have a solution non-singular (mod 7), and hence the equations (1.2) must have a non-trivial 7-adic solution.

(ii) Suppose that $R_{F,0} = 2$, $R_{F,1} = 4$, and $R_{F,2} \geq 7$.

We may rearrange variables and change notation to write

$$\left.\begin{aligned} F &= A_1X_1^3+A_2X_2^3 + p(A_3X_3^3+A_4X_4^3) + p^2(a_1x_1^3+\ldots+a_ux_u^3+b_1y_1^3+\ldots+b_vy_v^3) \\ &\quad + p^3(B_1z_1^3+\ldots+B_wz_w^3) + p^{3m+n}(e_1u_1^3+\ldots+e_ru_r^3) \\ G &= C_1X_1^2+C_2X_2^2 + C_3X_3^2+C_4X_4^2 + p(D_1x_1^2+\ldots+D_ux_u^2) + c_1y_1^2+\ldots+c_vy_v^2 \\ &\quad + d_1z_1^2+\ldots+d_wz_w^2 + p(f_1u_1^2+\ldots+f_ru_r^2) \end{aligned}\right\}$$

where p does not divide any of the a_i, b_i, c_i, d_i, A_i and e_1, the B_i, C_i, D_i, f_i and $e_2, \ldots, e_r$ are integers, we have

$$m \geq 1 \text{ and } 0 \leq n \leq 2,$$

and by (2.3),

$$u + v = R_{F,2} - R_{F,1} \geq 3,\ v + w = m_0 + q_{F,1} - R_{F,1} \geq 3,$$
$$4 + u + v + w + r \geq 11. \tag{3.5}$$

We may suppose that the f_i are not divisible by p, for if $p|f_i$ and $p^3|(p^{3m+n}e_i)$ for some i, then δ is not minimal (we can make a transformation of the type $u_i \longmapsto p^{-1}u_i$ without affecting the integral character of the coefficients).

Independently of $R_{F,2}$, if $p = 7$ and $A_3 \equiv \pm A_4 \pmod 7$, then we set $X_1 = X_2 = 0$ and divide what is left of F by 7. Since $v+w \geq 3$, the system takes the form (3.4), and we therefore have a non-trivial 7-adic solution to the equations (1.2). Henceforth we may therefore assume that if $p = 7$, then $A_3 \not\equiv \pm A_4 \pmod 7$.

(a) Suppose first that $u+v+w \geq 5$.

Then we set $X_1 = \ldots = X_4 = 0$ and consider the congruences obtained by a change of variables:

$$\left.\begin{aligned} F''^* &= (p^{-2}F)^* \\ G''^* &= G^* \end{aligned}\right\} \equiv 0 \pmod p$$

This system of congruences is either of a form in which hypothesis *(H)* applies, or else is of the same form as the systems we

considered in Lemmata 3.5 and 3.6, and with $q_{F'',0} \geq 3$. The system therefore has a solution non-singular (mod p), and hence the equations (1.2) must have a non-trivial p-adic solution, *unless* $p = 7$, $u = 2$, $v = 1$, $w = 2$, $a_2 \equiv \pm a_1 \pmod 7$ and $b_1 \equiv \pm 4a_1 \pmod 7$.

But in this last case we consider the system obtained by setting $X_1 = \ldots = X_4 = y_1 = z_1 = z_2 = 0$, and make a change of variables:

$$\left.\begin{aligned} F''^* &= (7^{-2}F)^* \\ G''^* &= (7^{-1}G)^* \end{aligned}\right\} \equiv 0 \pmod 7 .$$

Since

$$r \geq 7 - u - v - w = 2, \text{ and}$$

$$m_0 + m_1 - R_{F,1} - R_{G,0} \geq 10 - 4 - 3 = 3,$$

possibly by setting one or more variables to zero, we have a system of the form

$$\left.\begin{aligned} F''^* &= x_1^3 + x_2^3 \\ G''^* &= D_1x_1^2 + a_1x_2^2 + b_1y_1^2 + \ldots + b_{r'}y_{r'}^2 \end{aligned}\right\} \equiv 0 \pmod 7 \qquad (3.6)$$

where 7 does not divide $b_1, \ldots, b_{r'}$, $r' \geq 2$, and if $r' = 2$, then 7 does not divide a_1.

But we may solve the cubic congruence by putting $(x_1, x_2) = (\alpha_1, -\alpha_2)$ for any $\alpha_1, \alpha_2 \in \{1,2,4\}$. If $r' > 2$, then we may solve the quadratic congruence non-trivially, by Corollary 3.2.2, say with y_1 non-zero. If $r' = 2$, then since for some choice of α_1 and α_2, the expression $D_1\alpha_1^2 + a_1\alpha_2^2$ must be a non-zero residue, we may also solve the quadratic congruence for y_1 and y_2, by Corollary 3.2.2, say with y_1 non-zero. Then this system is soluble with

$$\Delta^*(F'', G'') \geq |6x_1y_1(b_1x_1)|_p = 1.$$

So the system has a solution non-singular (mod 7), and hence the equations (1.2) have a non-trivial 7-adic solution.

(b) Suppose now that $u+v+w < 5$.

Then by (3.5), we have $r \geq 3$.

If $n = 0$, put

$$X_3 = X_4 = x_1 = \ldots = x_u = y_1 = \ldots = y_v = z_1 = \ldots = z_w = 0,$$

and replace X_i by $p^m X_i$ for $i = 1,2$,

if $n = 1$, put

$$X_1 = X_2 = x_1 = \ldots = x_u = y_1 = \ldots = y_v = z_1 = \ldots = z_w = 0,$$

and replace X_i by $p^m X_i$ for $i = 3,4$,

if $n = 2$, put

$$X_1 = \ldots = X_4 = 0,$$

and replace x_i by $p^m x_i$ for $i = 1,\ldots,u$, y_i by $p^m y_i$ for $i = 1,\ldots,v$, and z_i by $p^m z_i$ for $i = 1,\ldots,w$.

Then we may divide through by p^{3m+n} in what is left of F, and by p in what is left of G. Then we obtain, by a rearrangement of variables and change of notation, the system

$$F^{(m)*} = a_1 x_1^3 + \ldots + a_u x_u^3 + b_1 y_1^3 + \ldots + b_v y_v^3$$

$$G^{(m)*} = + c_1 y_1^2 + \ldots + c_v y_v^2 + d_1 z_1^2 + \ldots + d_w z_w^2$$

where

$$u \geq 2, \quad u+v \geq 3, \quad v+w \geq 3,$$

and

$$u+v+w \geq \begin{cases} 5 \text{ for } n = 0,1 \\ 6 \text{ for } n = 2 . \end{cases}$$

When $p \neq 7$ the system of congruences $F^{(m)*} \equiv G^{(m)*} \equiv 0 \pmod{p}$ is seen to be of the same form as the systems we considered in Lemma 3.5. When $p = 7$ we have $A_{2i-1} \equiv A_{2i} \pmod 7$ for $i = 1,2$, and hence when $n = 0,1$ the system of congruences $F^{(m)*} \equiv G^{(m)*} \equiv 0 \pmod 7$ is seen to be of the same form as the systems we were able to solve in Lemma 3.6, and with $q_{F^{(m)},0} \geq 3$. When $p = 7$ and $n = 2$ we have $u+v+w \geq 6$, so that by setting at most one variable to zero we obtain

a system of the form we were able to solve in Lemma 3.6. The system therefore has a solution non-singular (mod p), and hence the equations (1.2) must have a non-trivial p-adic solution.

(iii) Suppose that $R_{F,0} = 2$, $R_{F,1} = 4$, and $R_{F,2} = 6$.

By a rearrangement of variables and change of notation, the p-normalised system now takes the form,

$$\left.\begin{aligned} F &= A_1X_1^3+A_2X_2^3+ p(A_3X_3^3+A_4X_4^3) + p^2(A_5X_5^3+A_6X_6^3) \\ &\qquad + p^{3m+n}(a_1x_1^3+\ldots+a_rx_r^3) \\ G &= B_1X_1^2+B_2X_2^2+ B_3X_3^2+B_4X_4^2 + B_5X_5^2+B_6X_6^2 \\ &\qquad + (b_1x_1^2+\ldots+b_rx_r^2) \end{aligned}\right\} \qquad (3.7)$$

where

$$m \geq 1,\ 0 \leq n \leq 2,\ \text{and } r \geq 5,$$

and where the B_i and b_i are integers, and p does not divide any of the A_i, or a_1. Further, by the same argument we applied to the f_i in (ii), at most p divides any of the b_i.

Let $S_0 = \{i: p\nmid b_i,\ \text{and } 1 \leq i \leq r\}$, and $S_1 = \{1,\ldots,r\}\backslash S_0$. Then one or other of card S_0 and card S_1 is at least 3, since $r \geq 5$. Suppose that card $S_t \geq 3$, and rename any three of the variables with indices in S_t as x_1, x_2, x_3. Put $x_4 = x_5 = 0$. Now we have the system

$$\left.\begin{aligned} F &= A_1X_1^3+A_2X_2^3+ p(A_3X_3^3+A_4X_4^3) + p^2(A_5X_5^3+A_6X_6^3) \\ &\qquad + p^{3M+N}(a_1x_1^3+a_2x_2^3+a_3x_3^3) \\ G &= B_1X_1^2+B_2X_2^2 + B_3X_3^2+B_4X_4^2 + B_5X_5^2+B_6X_6^2 \\ &\qquad + p^t(b_1x_1^2+b_2x_2^2+b_3x_3^2) \end{aligned}\right\}$$

where p does not divide any of the A_i or b_i, $t = 0$ or 1, $M \geq 1$, $0 \leq N \leq 2$, and at most two of a_1, a_2, a_3 are divisible by p.

If $N = 0$, put

$$X_3 = \ldots = X_6 = 0,$$

and replace X_i by p^MX_i for $i = 1,2$,

if $N = 1$, put

$$X_1 = X_2 = X_5 = X_6 = 0,$$

and replace X_i by $p^M X_i$ for $i = 3,4$,

if $N = 2$, put

$$X_1 = \ldots = X_4 = 0,$$

and replace X_i by $p^M X_i$ for $i = 5,6$.

Then we may divide through by p^{3M+N} in what is left of F, and by p^t in what is left of G, to obtain, by a rearrangement of variables and change of notation, the system

$$\begin{aligned} F'^* &= a_1 x_1^3 + \ldots + a_u x_u^3 + b_1 y_1^3 + \ldots + b_v y_v^3 \\ G'^* &= \qquad\qquad + c_1 y_1^2 + \ldots + c_v y_v^2 + d_1 z_1^2 + \ldots + d_w z_w^2 \end{aligned} \tag{3.8}$$

where

$$u+v \geq 3, \quad v+w \geq 3, \quad \text{and} \quad u+v+w \geq 5.$$

Now we know that $A_{2i-1} \not\equiv A_{2i} \pmod 7$ for $i = 1,2$, and hence *unless* $N = 2$, $p = 7$, $u = 2$, $v = 1$, $w = 2$, $a_1 \equiv \pm a_2 \pmod 7$ and $b_1 \equiv \pm 4a_1 \pmod 7$, then the system of congruences $F'^* \equiv G'^* \equiv 0 \pmod p$ is seen to be of the same form as the systems we considered in Lemmata 3.5 and 3.6, and with $q_{F'^*,0} \geq 3$. In this case the system therefore has a solution non-singular (mod p), and hence the equations (1.2) must have a non-trivial p-adic solution.

In the problematic case, returning to our notation previous to equation (3.8), we must have $A_5 \equiv \pm A_6 \pmod 7$. If B_5 and B_6 are both divisible by 7, we may set $X_1 = \ldots = X_4 = 0$, and divide through what is left of F by 7^2, and what is left of G by 7^t. In this case we have a system of the form (3.4), and hence a non-trivial 7-adic solution to the equations (1.2).

Returning to the notation implicit in equation (3.7), we may assume that B_5 and B_6 are not both divisible by 7. Suppose now in

addition that at least four of the b_i are divisible by 7, so without loss of generality 7 divides $b_1,\ldots,b_4$. Then we set

$$X_1 = \ldots = X_4 = x_5 = \ldots = x_r = 0,$$

and replace X_i by $p^M X_i$ for $i = 5,6$. By dividing through what is left of F by 7^{3M+2}, and what is left of G by 7, we obtain a system of the form (3.8) with $u+v \geq 3$, $v+w \geq 4$, and $u+v+w \geq 6$. Then, possibly by setting one of the variables to zero, we again obtain a system of the form (3.4), and hence a non-trivial 7-adic solution to the equations (1.2).

Then we may assume that at least two of the b_i are not divisible by 7 (since $r \geq 5$), and at least one of B_5 and B_6 is not divisible by 7. Set $X_1 = \ldots = X_4 = 0$. Then by a multiplicative change of variables and a change in notation, we have a system of the form

$$\left.\begin{aligned} (7^{-2}F^{\nu})^* &= x_1^3 + x_2^3 \\ (G^{\nu})^* &= A_1x_1^2 + a_1x_2^2 + b_1y_1^2 + b_2y_2^2 \end{aligned}\right\} \equiv 0 \pmod 7$$

where 7 does not divide a_1, b_1 or b_2. As for the congruences (3.6), this system has a solution non-singular (mod 7), and hence the equations (1.2) have a non-trivial 7-adic solution.

This completes the proof of the lemma.

4. ESTIMATES FOR EXPONENTIAL SUMS.

In this section we use exponential sums to show that hypothesis *(H)* is satisfied for all primes exceeding some prime p_0. Our first step is to obtain an estimate for the number of solutions to a pair of congruences (Lemma 4.3), which enables us to provide a good bound on a fundamental exponential sum.

Consider the congruences

$$\left.\begin{aligned} a_1x_1^3 + \ldots + a_5x_5^3 \\ b_1x_1^2 + \ldots + b_5x_5^2 \end{aligned}\right\} \equiv 0 \pmod{p} \qquad (4.1)$$

where at most one of the a_i, and at most one of the b_i is zero, and where a_i and b_i are not simultaneously zero. The number, N, of solutions (mod p) to the congruences (4.1) is given by the exponential sum

$$N = p^{-2} \sum_{u,v \pmod p} T_1(\underline{u}) \ldots T_5(\underline{u}) \ , \qquad (4.2)$$

where

$$T_j(\underline{u}) = \sum_{x \bmod p} e((a_j u x^3 + b_j v x^2)/p) \ , \text{ for } j = 1,\ldots,5.$$

The term with $u = v = 0$ in (4.2) contributes p^3 to N, which is the "expected" number of solutions to the congruences (4.1). We aim to show that the remaining terms, together with the number of singular solutions, do not exceed p^3-1 in modulus, and hence that the congruences (4.1) have a non-singular solution. For pairs of congruences of the same degree it is essentially only the ratios of the coefficients that determine the singularity condition. Unfortunately, singularity conditions are more complicated in our case, and we must therefore provide an estimate for the number of singular solutions of the congruences (4.1) (see Lemma 4.5).

Before proceeding to the proof of Lemma 4.3, we give some well-known estimates for exponential sums. These estimates are a consequence of the argument of Corollary 6D to Theorem 6C of Chapter IV of Schmidt [1976]. We give a sketch proof for the sake of completeness.

Lemma 4.1. *Let*

$$T(u) = \sum_{x \bmod p} e(ux^3/p), \quad \text{and } S(u) = \sum_{x \bmod p} e(ux^2/p).$$

Then when p is a rational prime congruent to 1 (mod 3), we have

$$\sum_{u=1}^{p-1} |T(u)|^2 = 2p(p-1)$$

$$\sum_{u=1}^{p-1} |T(u)|^4 = 6p^2(p-1)$$

$$\sum_{u=1}^{p-1} |T(u)|^6 \le 22p^3(p-1) .$$

When $p \equiv 2 \pmod 3$, each of these sums is zero. Also, for any positive integer r and prime $p > 2$, we have

$$\sum_{u=1}^{p-1} |S(u)|^{2r} = p^r(p-1).$$

Proof: First consider the cubic exponential sums.

(i) Suppose that $p \equiv 1 \pmod 3$.

By Lemma 4.3 of Vaughan [1981b], we have for $p \nmid u$

$$T(u) = \overline{\chi}(u)\tau(\chi) + \chi(u)\tau(\overline{\chi}),$$

where

$$\tau(\chi) = \sum_{x=1}^{p} \chi(x)e(x/p) ,$$

and χ is a primitive cubic character modulo p. As in Cook [1985] §4,

$$|T(u)|^2 = 2p + 2.\mathrm{Re}(\chi(u)\tau(\chi)\overline{\tau}(\overline{\chi})) , \text{ and}$$

$$|T(u)|^4 = 6p^2 + 8p.\mathrm{Re}(\chi(u)\tau(\chi)\overline{\tau}(\overline{\chi})) + 2.\mathrm{Re}(\chi^2(u)\tau^2(\chi)\overline{\tau}^2(\overline{\chi})) .$$

Summing over u gives us the first two results, and in the same manner,

$$\sum_{u=1}^{p-1} |T(u)|^6 = 20p^3(p-1) + \sum_{u=1}^{p-1} 2.\mathrm{Re}(\tau^3(\chi)\overline{\tau}^3(\overline{\chi}))$$

$$\le 22p^3(p-1),$$

since $|\tau(\chi)| = p^{1/2}$.

(ii) Suppose that $p \equiv 2 \pmod 3$.

For $r \ge 1$, the number of solutions to the congruence

$$x_1^3 + \ldots + x_r^3 \equiv y_1^3 + \ldots + y_r^3 \pmod p \qquad (4.3)$$

is given by

$$N = p^{-1} \sum_{u=0}^{p} |T(u)|^{2r}.$$

But since every residue is a cubic residue (mod p), the number of solutions to the congruence (4.3) is precisely p^{2r-1}. The result now follows.

The results for S follow in a similar, though simpler manner.

This completes the proof of the lemma.

The proof we give for Lemma 4.3 is entirely elementary, using nothing more complicated than the following well-known result on the quadratic residue symbol.

Lemma 4.2. *Let $f(x) = ax^2 + bx + c$, where a,b,c are integers, and let p be an odd prime not dividing a. Further, let $d = b^2 - 4ac$ be the discriminant of f. Then denoting the quadratic residue symbol by $\left[\frac{\cdot}{p} \right]$, we have*

$$\sum_{x=1}^{p} \left[\frac{f(x)}{p} \right] = \begin{cases} -\left[\frac{a}{p} \right] & \text{if } p \text{ does not divide } d \\ (p-1)\left[\frac{a}{p} \right] & \text{if } p \text{ divides } d. \end{cases}$$

Lemma 4.3. *Suppose that $p \neq 2,3$ is a rational prime. Then the number of solutions, S, of the simultaneous equations*

$$\left.\begin{aligned} x_1^3 + x_2^3 &= y_1^3 + y_2^3 \\ x_1^2 + x_2^2 &= y_1^2 + y_2^2 \end{aligned}\right\} \tag{4.4}$$

over $\mathbb{F}_p$ satisfies $S \leq 4p^2 - 5p + 2$.

Proof: Suppose that (x_1, x_2, y_1, y_2) satisfies (4.4), and let

$$\lambda = x_1 + x_2 - y_1 - y_2. \tag{4.5}$$

Write $s_k(\underline{z}) = z_1^k + z_2^k$ for $k \geq 1$. Then noting the identity

$$s_1(\underline{z})^3 - 3s_2(\underline{z})s_1(\underline{z}) + 2s_3(\underline{z}) = 0,$$

we have from (4.4),

$$(\lambda + s_1(\underline{y}))^3 - 3s_2(\underline{y})(\lambda + s_1(\underline{y})) + 2s_3(\underline{y}) = 0 .$$

So if the equations (4.4) are to have a solution, then λ must satisfy the cubic equation

$$\lambda(\lambda^2 + 3\lambda(y_1 + y_2) + 6y_1y_2) = 0 . \qquad (4.6)$$

There are two cases.

(a) $\lambda = 0$.

Then from (4.5) we have $x_1 + x_2 = y_1 + y_2$, and from (4.4) and (4.5) we have $x_1x_2 = y_1y_2$. So for any $s \in \mathbb{F}_p$,

$$(s - x_1)(s - x_2) = (s - y_1)(s - y_2).$$

Then $(x_1 - y_1)(x_1 - y_2) = 0$, and hence either $x_1 = y_1$, or $x_1 = y_2$. Thus $\lambda = 0$ gives precisely two solutions to (4.4) and (4.5) for each pair $(y_1, y_2) \in \mathbb{F}_p^2$ with $y_1 \neq y_2$, namely $(x_1, x_2) = (y_1, y_2)$ and $(x_1, x_2) = (y_2, y_1)$. But when $y_1 = y_2$ we plainly get the single solution $x_1 = x_2 = y_1$, so the total number of distinct solutions corresponding to $\lambda = 0$ is $2p(p-1) + p = 2p^2 - p$.

(b) $\lambda \neq 0$.

Then from (4.6), λ satisfies

$$\lambda^2 + 3\lambda(y_1 + y_2) + 6y_1y_2 = 0 . \qquad (4.7)$$

For a given (y_1, y_2), the number of solutions in λ of this quadratic equation is $\left[\frac{\Delta}{p} \right] + 1$, where $\Delta = 9y_1^2 - 6y_1y_2 + 9y_2^2$ is the discriminant of the quadratic in (4.7). Then the number, T, of triples (y_1, y_2, λ) satisfying (4.7) is

$$\sum_{y_1=1}^{p} \sum_{y_2=1}^{p} \left(\left[\frac{9y_1^2 - 6y_1y_2 + 9y_2^2}{p} \right] + 1 \right)$$

$$= p^2 + \sum_{y_1=1}^{p} \sum_{y_2=1}^{p} \left[\frac{9y_1^2 - 6y_1y_2 + 9y_2^2}{p} \right] .$$

Fix y_1 and consider $9y_1^2 - 6y_1y_2 + 9y_2^2$ as a quadratic in y_2. It has discriminant $-288y_1^2$, which is divisible by p if and only if $p = 2$ or 3, or if p divides y_1. Then using Lemma 4.2 we deduce that

$$\sum_{y_1=1}^{p}\sum_{y_2=1}^{p}\left\lfloor \frac{9y_1^2 - 6y_1y_2 + 9y_2^2}{p} \right\rfloor = \left\lfloor \sum_{y_1=1}^{p-1} - \left\lfloor \frac{9}{p} \right\rfloor \right\rfloor + (p-1)\left\lfloor \frac{9}{p} \right\rfloor = 0 .$$

So $T = p^2$. But $\lambda = 0$ is a solution of (4.7) whenever p divides $6y_1y_2$, that is, when $y_1 = p$ or $y_2 = p$, and we have already counted the solutions with $\lambda = 0$. So the number of triples (y_1, y_2, λ) satisfying (4.7) with $\lambda \neq 0$ is $p^2 - (2p - 1)$.

Given a triple (y_1, y_2, λ), substitute (4.5) into the quadratic in (4.4). We deduce that x_2 satisfies the quadratic

$$2y_1y_2 + 2\lambda(y_1 + y_2) + \lambda^2 + 2x_2^2 = 2x_2(y_1 + y_2 + \lambda). \qquad (4.8)$$

Given a solution x_2 to (4.8), x_1 is uniquely determined by (4.5), so for each triple (y_1, y_2, λ), there are at most 2 pairs (x_1, x_2) satisfying (4.4) and (4.5).

The total number of solutions to the equations (4.4) is therefore bounded by

$$S \le 2(p^2 - 2p + 1) + 2p^2 - p = 4p^2 - 5p + 2.$$

This completes the proof of the lemma.

Corollary 4.3.1. *Writing* $T(\underline{\underline{u}}) = \sum_{x \bmod p} e((ux^3 + vx^2)/p)$, *we have*

$$p^{-2}\sum_{u=1}^{p-1}\sum_{v=1}^{p-1} |T(\underline{\underline{u}})|^4 \le \begin{cases} 3p^2 - 12p + 9 & \text{for } p \equiv 1 \pmod 3 \\ 3p^2 - 6p + 3 & \text{for } p \equiv 2 \pmod 3. \end{cases}$$

Proof: The exponential sum

$$p^{-2}\sum_{u=0}^{p-1}\sum_{v=0}^{p-1} |T(\underline{\underline{u}})|^4$$

is the number of solutions of the simultaneous congruences

$$\left.\begin{aligned} x_1^3 + x_2^3 &\equiv x_3^3 + x_4^3 \\ x_1^2 + x_2^2 &\equiv x_3^2 + x_4^2 \end{aligned}\right\} \pmod p ,$$

which is bounded above by $4p^2-5p+2$, by Lemma 4.3. Also, by Lemma 4.1,

$$\sum_{\substack{u=0 \\ u=0 \text{ or } v=0}}^{p-1}\sum_{v=0}^{p-1} |T(\underline{u})|^4 = |T(\underline{0})|^4 + \sum_{u=1}^{p-1}|T(u,0)|^4 + \sum_{v=1}^{p-1}|T(0,v)|^4$$

$$= \begin{cases} p^4 + 7p^2(p-1) & \text{for } p \equiv 1 \pmod 3 \\ p^4 + p^2(p-1) & \text{for } p \equiv 2 \pmod 3 . \end{cases}$$

Then

$$p^{-2}\sum_{u=1}^{p-1}\sum_{v=1}^{p-1} |T(\underline{u})|^4 \le \begin{cases} 3p^2 - 12p + 9 & \text{for } p \equiv 1 \pmod 3 \\ 3p^2 - 6p + 3 & \text{for } p \equiv 2 \pmod 3 \end{cases}$$

This completes the proof of the corollary.

Lemma 4.4. *We have*

(i) if no a_i *is zero and no* b_i *is zero, then*

$$|N - p^3| \le \begin{cases} 6p^{5/2} + (2(33)^{1/2}-23)p^{3/2} \\ \qquad + (17-2(33)^{1/2})p^{1/2} & \text{for } p \equiv 1 \pmod 3 \\ 6p^{5/2} - 11p^{3/2} + 5p^{1/2} & \text{for } p \equiv 2 \pmod 3 \end{cases}$$

(ii) if there is an i for which a_i *is zero, but no* b_j *is zero, then*

$$|N - p^3| \le \begin{cases} 3p^{5/2} - 11p^{3/2} + 8p^{1/2} + 6p^2 - 6p & \text{for } p \equiv 1 \pmod 3 \\ 3p^{5/2} - 5p^{3/2} + 2p^{1/2} & \text{for } p \equiv 2 \pmod 3 \end{cases}$$

(iii) if there is an i for which b_i *is zero, but no* a_j *is zero, then*

$$|N - p^3| \le \begin{cases} 6p^{5/2} + (2(33)^{1/2}-24)p^{3/2} \\ \qquad + (18-2(33)^{1/2})p^{1/2} + p^2 - p & \text{for } p \equiv 1 \pmod 3 \\ 6p^{5/2} - 12p^{3/2} + 6p^{1/2} + p^2 - p & \text{for } p \equiv 2 \pmod 3 \end{cases}$$

(iv) if there is an i for which a_i *is zero, and a j for which* b_j *is zero, then*

$$|N - p^3| \le \begin{cases} (12p^5 - 96p^4 + 264p^3 - 288p^2 + 108p)^{1/2} \\ \qquad + 7p^2 - 7p & \text{for } p \equiv 1 \pmod 3 \\ (12p^5 - 48p^4 + 72p^3 - 48p^2 + 12p)^{1/2} \\ \qquad + p^2 - p & \text{for } p \equiv 2 \pmod 3 \end{cases}$$

Proof: Separating out the main term, we obtain

$$N - p^3 = p^{-2}\Bigg[\Bigg(\sum_{u=1}^{p-1} \sum_{v=1}^{p-1} T_1(\underline{u}) \ldots T_5(\underline{u}) \Bigg) + \Bigg(\sum_{v=1}^{p-1} T_1(0,v) \ldots T_5(0,v) \Bigg) + \Bigg(\sum_{u=1}^{p-1} T_1(u,0) \ldots T_5(u,0) \Bigg) \Bigg]$$

$$= S_1 + S_2 + S_3 \text{ , say.}$$

There are four cases:

(i) No a_i is zero, and no b_i is zero.

Then by Hölder's inequality,

$$\left| \sum_{u=1}^{p-1} \sum_{v=1}^{p-1} T_1(\underline{u}) \ldots T_5(\underline{u}) \right| \leq \prod_{i=1}^{5} \Bigg(\sum_{u=1}^{p-1} \sum_{v=1}^{p-1} |T_i(\underline{u})|^5 \Bigg)^{1/5}$$

$$\leq \Big(\operatorname*{Sup}_{\underline{u}} |T_I(\underline{u})| \Big) \sum_{u=1}^{p-1} \sum_{v=1}^{p-1} |T_I(\underline{u})|^4$$

for some $I \in \{1,\ldots,5\}$, where the supremum is over $u \neq 0$ and $v \neq 0$. By the Riemann Hypothesis for finite fields (see e.g. Schmidt [1976]), we have for any polynomial $f(x)$ of degree k with integer coefficients, and with p not dividing the content of $f(x) - f(0)$, the estimate

$$\left| \sum_{x \bmod p} e(f(x)/p) \right| \leq (k-1)p^{1/2}.$$

Using this estimate, making a change of variables, and applying Corollary 4.3.1, we obtain

$$|S_1| \leq \begin{cases} 6p^{5/2} - 24p^{3/2} + 18p^{1/2} & \text{for } p \equiv 1 \pmod 3 \\ 6p^{5/2} - 12p^{3/2} + 6p^{1/2} & \text{for } p \equiv 2 \pmod 3 \text{ .} \end{cases}$$

Also, by a similar argument to the above, there is a J such that

$$|S_2| \leq p^{-2} \sum_{v=1}^{p-1} |T_J(0,v)|^5.$$

By Lemma 4.1 and Cauchy's inequality, on making a change of variables,

$$|S_2| \leq p^{1/2}(p-1).$$

Also, as above, there is a K such that

$$|S_3| \le p^{-2} \sum_{u=1}^{p-1} |T_K(u,0)|^5.$$

If $p \equiv 2 \pmod 3$, then the sum is zero. Otherwise, by Cauchy's inequality and Lemma 4.1,

$$|S_3| \le p^{-2} \left[\sum_{u=1}^{p-1} |T_K(u,0)|^6 \right]^{1/2} \left[\sum_{u=1}^{p-1} |T_K(u,0)|^4 \right]^{1/2}$$

$$\le p^{-2}(22p^3(p-1).6p^2(p-1))^{1/2}$$

$$= 2(33)^{1/2}p^{1/2}(p-1) .$$

The result now follows on adding the above estimates for $|S_1|$, $|S_2|$ and $|S_3|$.

(ii) there is an i for which a_i is zero, but no b_j is zero.

By a rearrangement of variables we may suppose that $a_1 = 0$. Then by Hölder's inequality and the Riemann Hypothesis for finite fields,

$$\left| \sum_{u=1}^{p-1} \sum_{v=1}^{p-1} T_1(\underline{u}) \ldots T_5(\underline{u}) \right| \le \left(\sup_{\underline{u}} T_1(\underline{u}) \right) \prod_{i=2}^{5} \left[\sum_{u=1}^{p-1} \sum_{v=1}^{p-1} |T_i(\underline{u})|^4 \right]^{1/4}$$

$$\le p^{1/2} \sum_{u=1}^{p-1} \sum_{v=1}^{p-1} |T_I(\underline{u})|^4$$

for some $I \in \{2,\ldots,5\}$. Making a change of variable and using Corollary 4.3.1, we have

$$|S_1| \le \begin{cases} 3p^{5/2} - 12p^{3/2} + 9p^{1/2} & \text{for } p \equiv 1 \pmod 3 \\ 3p^{5/2} - 6p^{3/2} + 3p^{1/2} & \text{for } p \equiv 2 \pmod 3 . \end{cases}$$

Also, in a similar way to (i), we have

$$|S_2| \le p^{1/2}(p-1),$$

and noting that $T_1(u,0) = p$, we have

$$|S_3| \le p^{-2} \left[p \sum_{u=1}^{p-1} |T_K(u,0)|^4 \right] ,$$

for some integer K, and hence,

$$|S_3| \begin{cases} \le 6p(p-1) & \text{for } p \equiv 1 \pmod 3 \\ = 0 & \text{for } p \equiv 2 \pmod 3 . \end{cases}$$

The result now follows on adding the above estimates for $|S_1|$, $|S_2|$ and $|S_3|$.

(iii) there is an i for which b_i is zero, but no a_j is zero.

In a similar way to (i) and (ii),

$$|N - p^3| \leq \begin{cases} 6p^{5/2} - 24p^{3/2} + 18p^{1/2} + p(p-1) \\ \qquad + 2(33)^{1/2}p^{1/2}(p-1) & \text{for } p \equiv 1 \pmod 3 \\ 6p^{5/2} - 12p^{3/2} + 6p^{1/2} + p(p-1) & \text{for } p \equiv 2 \pmod 3 . \end{cases}$$

(iv) there is an i for which a_i is zero, and a j for which b_j is zero.

We may rearrange variables so that $a_1 = 0$ and $b_2 = 0$. In a similar way to (i) and (ii), we obtain

$$p^{-2} \left| \sum_{u=1}^{p-1} \sum_{v=1}^{p-1} T_1(\underline{u}) \ldots T_5(\underline{u}) \right| \leq (2p^{1/2})(p^{1/2})p^{-2} \sum_{u=1}^{p-1} \sum_{v=1}^{p-1} |T_I(\underline{u})|^3$$

for some I, and hence by Cauchy's inequality,

$$|S_1| \leq 2p\left[p^{-2} \sum_{u=1}^{p-1} \sum_{v=1}^{p-1} |T_I(\underline{u})|^2\right]^{1/2} \left[p^{-2} \sum_{u=1}^{p-1} \sum_{v=1}^{p-1} |T_I(\underline{u})|^4\right]^{1/2}. \qquad (4.9)$$

But

$$p^{-2} \sum_{u=0}^{p-1} \sum_{v=0}^{p-1} |T_I(\underline{u})|^2$$

is the number of solutions to the system of congruences

$$\left.\begin{aligned} a_I x_1^3 &\equiv a_I x_2^3 \\ b_I x_1^2 &\equiv b_I x_2^2 \end{aligned}\right\} \pmod p,$$

which is plainly p, since a_I and b_I are both non-zero. Then using Lemma 4.1 we deduce that

$$\sum_{u=1}^{p-1} \sum_{v=1}^{p-1} |T_I(\underline{u})|^2 = \begin{cases} p^3 - p^2 - 2p(p-1) - p(p-1) & \text{for } p \equiv 1 \pmod 3 \\ p^3 - p^2 - p(p-1) & \text{for } p \equiv 2 \pmod 3 . \end{cases}$$

By Corollary 4.3.1 and (4.9), we then deduce that

$$|S_1| \leq \begin{cases} (12p^5 - 96p^4 + 264p^3 - 288p^2 + 108p)^{1/2} & \text{for } p \equiv 1 \pmod 3 \\ (12p^5 - 48p^4 + 72p^3 - 48p^2 + 12p)^{1/2} & \text{for } p \equiv 2 \pmod 3 . \end{cases}$$

And as in (i) and (ii),

$$|S_2| \leq p(p-1), \text{ and}$$

$$|S_3| \begin{cases} \leq 6p(p-1) & \text{for } p \equiv 1 \pmod 3 \\ = 0 & \text{for } p \equiv 2 \pmod 3 . \end{cases}$$

The result now follows on adding our estimates for $|S_1|$, $|S_2|$ and $|S_3|$.

This completes the proof of the lemma.

Lemma 4.5. *Suppose that* $p \neq 2,3$. *Then the number of distinct singular solutions of the congruences* (4.1) *is at most*

(*a*) $9p-8$ *if no* a_i *is zero, and no* b_i *is zero,*

(*b*) $5p-4$ *if there is an* i *for which either* $a_i \equiv 0$ *or* $b_i \equiv 0$, *but there are no* i *and* j, $i \neq j$, *with* $a_i \equiv b_j \equiv 0$,

(*c*) $2p-1$ *if there is an* i *for which* a_i *is zero, and a* j *for which* b_j *is zero.*

Proof: Suppose that $(\alpha_1,\ldots,\alpha_5)$ is any solution of the congruences (4.1) singular (mod p). Suppose also that precisely v of the α_i are non-zero, and rearrange variables so that $\alpha_i \not\equiv 0$ for $i = 1,\ldots,v$. Plainly, as a_i and b_i are not simultaneously zero, we have $v = 0$ or $v \geq 2$. Since $\underline{\alpha}$ is a singular solution and $p \neq 2,3$, we must have for every $i,j \in \{1,\ldots,5\}$,

$$\alpha_i\alpha_j(a_ib_j\alpha_i - a_jb_i\alpha_j) \equiv 0 \pmod{p},$$

and hence $a_ib_j\alpha_i \equiv a_jb_i\alpha_j \pmod{p}$ for all $i,j \in \{1,\ldots,v\}$.

Suppose that $v \geq 2$. There are two cases:

(i) there is no $i \in \{1,\ldots,v\}$ for which either $a_i \equiv 0$ or $b_i \equiv 0$.

Then

$$\alpha_j \equiv \begin{cases} \dfrac{a_1b_j}{b_1a_j}\,\alpha_1 & \text{for } j = 1,\ldots,v \\ 0 & \text{for } j = v+1,\ldots,5. \end{cases}$$

There are plainly $p-1$ such non-trivial solutions, taking $\alpha_1 \equiv 1,\ldots,p-1$.

(ii) there are $i, j \in \{1, \ldots, v\}$ for which $a_i \equiv 0$ or $b_j \equiv 0$.

Then we can find $i, j \in \{1, \ldots, v\}$ such that a_j and b_i are both non-zero, but $a_j b_i \alpha_j \equiv a_i b_j \alpha_i \equiv 0 \pmod{p}$. Then $\alpha_j \equiv 0$, which is a contradiction.

So we conclude that (i) holds whenever $v \neq 0$, and from

$$\left.\begin{aligned} a_1\alpha_1^3 + \ldots + a_5\alpha_5^3 &\equiv 0 \\ b_1\alpha_1^2 + \ldots + b_5\alpha_5^2 &\equiv 0 \end{aligned}\right\} \pmod{p}$$

we obtain

$$(b_1^3/a_1^2 + \ldots + b_v^3/a_v^2).\alpha_1^3 \equiv 0 \pmod{p}.$$

Rearrange variables so that $a_i \not\equiv 0$ and $b_i \not\equiv 0$ for $i = 1, \ldots, n$, and either $a_i \equiv 0$ or $b_i \equiv 0$ for $n < i \leq 5$. Let $\beta_i \equiv b_i^3/a_i^2$ for $i = 1, \ldots, n$. Using the notation of Appendix A, there are at most $N(p; \underline{\beta})$ distinct congruences of the form

$$\beta_{i_1} + \ldots + \beta_{i_t} \equiv 0 \pmod{p},$$

with $1 \leq i_1 < i_2 < \ldots < i_t \leq n$ and $0 \leq t \leq n$, one of which is the "empty" congruence. For each such non-empty congruence, we have p-1 non-trivial singular solutions (given by the p-1 non-zero values of α_1). Thus, by Theorem 1.1 of Appendix A, the number of singular solutions, including $\underline{0}$, to the congruences (4.1) is at most

(a) $(10 - 1)(p-1) + 1 = 9p-8$ if no a_i is zero, and no b_i is zero,

(b) $(6 - 1)(p-1) + 1 = 5p-4$ if there is an i for which either $a_i \equiv 0$ or $b_i \equiv 0$, but there are no i and j, $i \neq j$, with $a_i \equiv b_j \equiv 0$,

(c) $(3 - 1)(p-1) + 1 = 2p-1$ if there is an i for which a_i is zero, and a j for which b_j is zero.

This completes the proof of the lemma.

Lemmata 4.4 and 4.5 show that the number, N^*, of solutions to the simultaneous congruences (4.1) which are non-singular (mod p), satisfies $|N^* - p^3| \leq E$, where E satisfies:

(i) if no a_i is zero and no b_i is zero, then

$$E \le \begin{cases} 6p^{5/2} + (2(33)^{1/2}-23)p^{3/2} \\ \qquad + (17-2(33)^{1/2})p^{1/2} + 9p - 8 & \text{for } p \equiv 1 \pmod 3 \\ 6p^{5/2} - 11p^{3/2} + 5p^{1/2} + 9p - 8 & \text{for } p \equiv 2 \pmod 3 \end{cases}$$

(ii) if there is an i for which a_i is zero, but no b_j is zero, then

$$E \le \begin{cases} 3p^{5/2} - 11p^{3/2} + 8p^{1/2} + 6p^2 - p - 4 & \text{for } p \equiv 1 \pmod 3 \\ 3p^{5/2} - 5p^{3/2} + 2p^{1/2} + 5p - 4 & \text{for } p \equiv 2 \pmod 3 \end{cases}$$

(iii) if there is an i for which b_i is zero, but no a_j is zero, then

$$E \le \begin{cases} 6p^{5/2} + (2(33)^{1/2}-24)p^{3/2} \\ \qquad + (18-2(33)^{1/2})p^{1/2} + p^2 + 4p - 4 & \text{for } p \equiv 1 \pmod 3 \\ 6p^{5/2} - 12p^{3/2} + 6p^{1/2} + p^2 + 4p - 4 & \text{for } p \equiv 2 \pmod 3 \end{cases}$$

(iv) if there is an i for which a_i is zero, and a j for which b_j is zero, then

$$E \le \begin{cases} (12p^5 - 96p^4 + 264p^3 - 288p^2 + 108p)^{1/2} \\ \qquad + 7p^2 - 5p - 1 & \text{for } p \equiv 1 \pmod 3 \\ (12p^5 - 48p^4 + 72p^3 - 48p^2 + 12p)^{1/2} \\ \qquad + p^2 + p - 1 & \text{for } p \equiv 2 \pmod 3 \end{cases}$$

It is now easily verified that these inequalities (i) - (iv) ensure that $|N^* - p^3| < p^3$ for all primes p satisfying:

$$\text{when no } a_i \text{ is zero} \begin{cases} p > 31 & \text{for } p \equiv 1 \pmod 3 \\ p > 29 & \text{for } p \equiv 2 \pmod 3 \end{cases}$$

$$\text{when an } a_i \text{ is zero} \begin{cases} p > 13 & \text{for } p \equiv 1 \pmod 3 \\ p > 5 & \text{for } p \equiv 2 \pmod 3 \end{cases}$$

Thus, for all rational primes, except possibly

$$\begin{cases} \text{when no } a_i \text{ is zero,} & 2,3,5,7,11,13,17,19,23,29,31, \\ \text{when an } a_i \text{ is zero,} & 2,3,5,7,13 \end{cases}$$

the simultaneous congruences (4.1) have a solution non-singular

(mod p). In §§5 and 6 we shall be concerned with the treatment of these remaining primes, and the particular cases for which they cause difficulties.

5. THE PRIMES P CONGRUENT TO 5 (MOD 12), AND THE PRIMES 2 AND 3.

In this section we start by giving a rather simple proof that $\Gamma_p^*(3,2) \leq 11$ for all primes $p \equiv 5 \pmod{12}$. We then consider the primes 2 and 3. Since both primes divide the Jacobian $\Delta^*(F,G)$ of the system (1.2), we are unable simply to apply Hensel's Lemma to solutions of the p-normalised equations (mod p). We show that if we can find $\underline{x} \in \mathbb{Z}_3^s$ non-singular (mod 9) satisfying the cubic equation (mod 3^2), and the quadratic equation (mod 3), then a version of Hensel's Lemma can be used to infer the existence of a non-trivial 3-adic solution. Thus, by checking a finite number of cases using a computer, we may verify that all 3-normalised equations of the form (1.2) possess a non-trivial 3-adic solution. Although a similar procedure might also be applied to the prime 2, we choose to apply a modification of the argument applied to the primes $p \equiv 5 \pmod{12}$, which involves checking a few cases directly.

Lemma 5.1. *We have* $\Gamma_p^*(3,2) \leq 11$ *for all primes* $p \equiv 5$ (*mod* 12).

Proof: If $p \equiv 5 \pmod{12}$, then -1 is a quadratic residue (mod p), and every non-zero $a \in \mathbb{Z}/p\mathbb{Z}$ is a cubic residue (mod p). Consider a system of the form (1.2) in $s \geq 11$ variables. Suppose that c is a quadratic non-residue (mod p), and let $K = \{0,1,c,p,pc\}$. Then if $b \in \mathbb{Z}_p \setminus \{0\}$, we have $b = ga^2$ for some $a \in \mathbb{Z}_p \setminus \{0\}$ and $g \in K \setminus \{0\}$ (this

is a simple exercise in the use of Hensel's Lemma; see, for example, Cassels [1986], Chapter 4, Lemma 3.2 and its corollary).

We partition $\{1,\dots,s\}$ into the five sets

$$\mathfrak{U}_g = \{i:\ c_i = ga^2 \text{ for some } a \in \mathbb{Z}_p\backslash\{0\}\ \} \qquad (g \in K),$$

so that

$$\sum_{g \in K} |\mathfrak{U}_g| = s.$$

Let M denote the number of $g \in K\backslash\{0\}$ for which $\mathfrak{U}_g \neq \emptyset$. Then $M \leq 4$, and writing $[x]$ for the integer part of x, we have

$$|\mathfrak{U}_0| + \sum_{g \in K\backslash\{0\}} \left[\tfrac{1}{2}|\mathfrak{U}_g|\right] \geq |\mathfrak{U}_0| + \sum_{\substack{g \in K\backslash\{0\} \\ \mathfrak{U}_g \neq \emptyset}} \tfrac{1}{2}\left(|\mathfrak{U}_g| - 1\right)$$

$$\geq \tfrac{1}{2} \sum_{g \in K} |\mathfrak{U}_g| - \tfrac{1}{2} M$$

$$\geq \tfrac{1}{2}(s - 4)\ .$$

Thus we may rearrange the variables so that for $i = 1,2,3,4$, either $2i-1 \in \mathfrak{U}_q$ and $2i \in \mathfrak{U}_q$ for some $q \in K\backslash\{0\}$, or else $2i-1 \in \mathfrak{U}_0$. In the latter case $d_{2i-1}x^2 + d_{2i}y^2 = 0$ has the non-trivial solution $(x,y) = (1,0)$. In the former case $d_{2i-1}x^2 + d_{2i}y^2 = q((ax)^2+(by)^2)$ for some $a,b \in \mathbb{Z}_p\backslash\{0\}$, and we have the non-trivial solution $(x,y) = (\omega b, a)$, where ω satisfies $\omega^2 = -1$ (such an ω exists by Hensel's Lemma, for we can find ω' with $\omega'^2 \equiv -1 \pmod{p}$). Then

$$G = (d_1x_1^2+d_2x_2^2) + \dots + (d_7x_7^2+d_8x_8^2) + d_9x_9^2 + d_{10}x_{10}^2 + \dots + d_sx_s^2$$

where each expression in parentheses may be made zero by setting $(x_{2i-1}, x_{2i}) = (a_{2i-1}, a_{2i})$, say. Put $a_9 = \dots = a_s = 0$, and for $i = 1,2,3,4$ put $C_i = c_{2i-1}a_{2i-1}^3 + c_{2i}a_{2i}^3$. For each i, there is a $g \in L = \{0,1,p,p^2\}$ and an $a \in \mathbb{Z}_p\backslash\{0\}$ such that $C_i = ga^3$, since $p \equiv 2 \pmod 3$ (this is also a simple exercise in the use of Hensel's Lemma; see, for example, Cassels [1986], Chapter 4, Lemma 3.4).

Consider the equation

$$C_1X_1^3 + C_2X_2^3 + C_3X_3^3 + C_4X_4^3 = 0 \text{ over } \mathbb{Z}_p . \tag{5.1}$$

If any C_i is zero, we have the non-trivial solution $X_i = 1$, $X_j = 0$ $(j \neq i)$. So we may assume that no C_i is zero, and hence that there is a rearrangement of variables such that $C_1 = ga^3$ and $C_2 = gb^3$ for some $a,b \in \mathbb{Z}_p$ and $g \in L\backslash\{0\}$. Then (5.1) has the non-trivial solution

$$(X_1,\ldots,X_4) = (b,-a,0,0).$$

In either case the equation (5.1) has some non-trivial solution $(b_1,\ldots,b_4)$ over $\mathbb{Z}_p$, and hence the system (1.2) has the non-trivial solution

$$(x_1,\ldots,x_s) = (a_1b_1, a_2b_1, a_3b_2, a_4b_2, a_5b_3, a_6b_3, a_7b_4, a_8b_4, 0,\ldots,0).$$

This completes the proof of the lemma.

Lemma 5.2. *We have* $\Gamma_2^*(3,2) \leq 11$.

Proof: Consider the equations (1.2) with $s \geq 11$. We may assume that for every i, at most one of c_i and d_i is zero (for otherwise we plainly have a non-trivial solution). Let $K = \{0,1,2,4\}$. Then by Hensel's Lemma we may partition the indices $1,\ldots,s$ into the four sets

$$\mathfrak{U}_g = \{i:\ c_i = ga^3 \text{ for some } a \in \mathbb{Z}_2\backslash\{0\}\},\ (g \in K),$$

so that

$$\sum_{g \in K} |\mathfrak{U}_g| = s.$$

Let M denote the number of $g \in K\backslash\{0\}$ for which $\mathfrak{U}_g \neq \emptyset$, and let $r = |\mathfrak{U}_0|$. Then $M \leq 3$ and

$$S = |\mathfrak{U}_0| + \sum_{g \in K\backslash\{0\}} \left[\frac{1}{2}|\mathfrak{U}_g|\right] \geq r + \sum_{\substack{g \in K\backslash\{0\} \\ \mathfrak{U}_g \neq \emptyset}} \frac{1}{2}\left(|\mathfrak{U}_g| - 1\right)$$

$$\geq \frac{1}{2}(s+r-3) .$$

There are two cases.

(i) Suppose that $r > 0$.

Then $S \geq 5$, so we may rearrange variables so that for $i = 1,\ldots,5$, either $2i-1 \in \mathfrak{U}_q$ and $2i \in \mathfrak{U}_q$ for some $q \in K\backslash\{0\}$, or else $2i-1 \in \mathfrak{U}_0$. In each case $c_{2i-1}x^3 + c_{2i}y^3 = 0$ has a non-trivial solution over $\mathbb{Z}_2$, say (a_{2i-1}, a_{2i}). Define $D_i = d_{2i-1}a_{2i-1}^2 + d_{2i}a_{2i}^2$ for $i = 1,\ldots,5$. Then since $\Gamma^*(2) = 5$, we can find a non-trivial solution to the equation

$$D_1u_1^2 + \ldots + D_5u_5^2 = 0 ,$$

and hence a non-trivial solution to the system (1.2) over $\mathbb{Z}_2$.

(ii) Suppose that $r = 0$.

Then $S \geq 4$. If $S \geq 5$, the same argument as in (i) applies, so we may assume that $S = 4$. By a rearrangement and multiplicative change of variables we may suppose that $c_{2i-1} = c_{2i} \in K\backslash\{0\}$ for $i = 1,2,3,4$. Further, we may assume that one at least of the remaining 3 indices, without loss of generality $i = 9$, is such that $c_7 = c_8 = c_i$. Let $\mathcal{B}_i = \{2i,\ 2i-1\}$ for $i = 1,2,3$, and $\mathcal{B}_4 = \{7,8,9\}$.

Consider the cubic equations corresponding to each set $\mathcal{B}_i$. The equation $x^3 + y^3 = 0$ has the non-trivial solution $(u,-u)$ for any $u \in \mathbb{Z}_2\backslash\{0\}$. Also, the equation $x^3 + y^3 + z^3 = 0$ has the non-trivial solution (au,bu,cu) for any $u \in \mathbb{Z}_2\backslash\{0\}$, with $c = 2^r$ for any $r \geq 1$ and some $(a,b) \equiv (1,-1) \pmod 8$ (by an application of Hensel's lemma).

Write $D_i = d_{2i-1} + d_{2i}$ $(i = 1,2,3)$, $D_4 = d_7a^2 + d_8b^2 + d_9c^2$, and consider the equation

$$Q(D_1,D_2,D_3,D_4) = D_1u_1^2 + \ldots + D_4u_4^2 = 0 . \qquad (5.2)$$

If we can find a non-trivial solution $(u_1,\ldots,u_4)$ to the equation

(5.2) over $\mathbb{Z}_2$, then in view of our change of variables, the equations (1.2) have the non-trivial solution

$$(u_1, -u_1, u_2, -u_2, u_3, -u_3, au_4, bu_4, cu_4, 0, \ldots, 0).$$

We now consider for which quadratics $Q(\underline{\underline{D}})$ a non-trivial solution exists. We plainly have a non-trivial solution to the quadratic $Q(\underline{\underline{D}})$ if any D_i is zero, so we may suppose that no D_i is zero. By making a multiplicative change of variables, we may suppose that $(D_i, 8) \leq 2$ for each i, that $(D_1 D_2, 2) = 1$, and that $D_1 \equiv 1 \pmod 8$. Further, if for any i and j we have $D_i \equiv D_j \pmod 8$ and $(D_i D_j, 2) = 1$, then we may suppose that $D_1 \equiv D_2 \pmod 8$. If we can now find a solution to the quadratic equation (mod 8) for which D_i and u_i are both coprime with 2 for some i (a solution we shall call "non-singular (mod 8)"), then by Hensel's Lemma there is a non-trivial solution to the quadratic over $\mathbb{Z}_2$.

By rearranging the D_i, it is readily verified that Table 5.1 lists all the distinct possible cases, and hence that we may restrict attention to five quadratics for which we can find no solution non-singular (mod 8).

Table 5.1. (Here γ and δ denote arbitrary integers.)

Quadratic, Q	Non-singular solution (mod 8)	Quadratic, Q	Non-singular solution (mod 8)
$D_1\ D_2\ D_3\ D_4$	$u_1\ u_2\ u_3\ u_4$	$D_1\ D_2\ D_3\ D_4$	$u_1\ u_2\ u_3\ u_4$
1 7 γ δ	1 1 0 0	1 1 5 5	none
3 5 γ δ	1 1 0 0	1 1 1 2	1 1 2 1
1 1 6 γ	1 1 1 0	1 1 3 2	1 1 2 1
1 5 2 γ	1 1 1 0	1 1 2 2	none
1 1 1 1	none	1 3 2 2	1 1 1 1
1 1 1 3	1 0 2 1	1 3 2 6	none
1 1 1 5	1 1 1 1	1 3 6 6	1 1 1 1
1 1 3 3	1 1 1 1	1 5 6 6	none

We now aim to show that by using an alternative permissable choice of (a,b,c), we may change the value of D_4 so as to avoid the five problematic quadratics. Consider the coefficient $D_4 = d_7a^2 + d_8b^2 + d_9c^2$. By rearranging variables we may write $D_4 = d\mathcal{R}(a,b,c)$, with $\mathcal{R}(a,b,c) = a^2 + \alpha b^2 + \beta c^2$ for some $d,\alpha,\beta \in \mathbb{Z}_2$. By rearranging variables, we may suppose that if $\alpha \equiv \beta \pmod 8$, and $(\alpha\beta,2) = 1$, then $\alpha \equiv 1 \pmod 8$. It may be verified that Table 5.2 lists the only distinct possibilities for (α,β) when one of α and β

Table 5.2.

α β mod 8	a_1 b_1 c_1	$\mathcal{R}(a_1,b_1,c_1)$ mod 8	a_2 b_2 c_2	$\mathcal{R}(a_2,b_2,c_2)$ mod 8
1 1	1 -1 0	2	1 -1 2	6
1 3	1 -1 0	2	1 0 -1	4
1 5	1 -1 0	2	1 0 -1	6
1 7	1 -1 0	2	1 -1 2	6
3 5	1 -1 0	4	1 0 -1	6
3 7	1 -1 0	4	0 1 -1	2
5 7	1 -1 0	6	0 1 -1	4
1 2	1 -1 0	2	1 0 -1	3
1 4	1 -1 0	2	1 0 -1	5
1 6	1 -1 0	2	1 0 -1	7
1 0	1 -1 0	2	1 0 -1	1
3 2	1 0 -1	3	0 1 -1	5
3 4	1 0 -1	5	0 1 -1	7
3 6	1 0 -1	7	0 1 -1	1
3 0	1 0 -1	1	0 1 -1	3
5 2	1 -1 0	6	1 0 -1	3
5 4	1 -1 0	6	1 0 -1	5
5 6	1 -1 0	6	1 0 -1	7
5 0	1 -1 0	6	1 0 -1	1
7 2	0 1 -1	1	1 0 -1	3
7 4	0 1 -1	3	1 0 -1	5
7 6	0 1 -1	5	1 0 -1	7
7 0	0 1 -1	7	1 0 -1	1

is not divisible by 2. Thus we can find (a_1, b_1, c_1) and (a_2, b_2, c_2) so that, by making the change of variables which allows us to assume that $(D_4, 8) \leq 2$, we have two values distinct (mod 8) for this reduced coefficient.

Also, using γ and δ to denote arbitrary integers coprime with 2, we have the cases with $(\alpha, \beta) = (\gamma 2^r, \delta 2^s)$ for some $s \geq r \geq 1$. Here we obtain

$$\mathcal{R}(1,-1,0) = 1 + \gamma 2^r \text{ and } \mathcal{R}(0,1,-1) = \gamma 2^r + \delta 2^s.$$

Write $\gamma 2^r + \delta 2^s = \eta 2^t$ with η not divisible by 2 (if $\eta = 0$ then we can easily solve the quadratic non-trivially). Then unless t is even, by taking account of changes of variable which allow us to assume that $(D_4, 8) \leq 2$, the coefficients represented by $\mathcal{R}(1,-1,0)$ and $\mathcal{R}(0,1,-1)$ are distinct (mod 8). But if $t = 2v$ is even, then

$$2^{-t}\mathcal{R}(0,1,-1) \equiv \eta \pmod 8, \text{ and } 2^{-t}\mathcal{R}(2^{v+1},1,-1) \equiv \eta + 4 \pmod 8,$$

and these residues are distinct (mod 8). Further η is not divisible by 2, so taking account of changes of variable which allow us to assume that $(D_4, 8) \leq 2$, the coefficients D_4 given by these solutions are distinct.

In all of these cases we obtain a choice of values for D_4. Making the change of variable which allows us to assume that $(D_4, 8) \leq 2$, we see that there are two values distinct (mod 8) for this "reduced" coefficient. But by examining Table 5.1, it is readily seen that one choice, at least, of coefficient available to us gives a quadratic which we can solve, no matter where the coefficient D_4 has been rearranged to.

Then we can find a solution non-singular (mod 8) to the "reduced" quadratic equation, and hence there is a non-trivial solution to the equations (1.2) over $\mathbb{Z}_2$.

This completes the proof of the lemma.

We now attend to the prime 3. We first prove a version of Hensel's Lemma not dissimilar to one used in Davenport and Lewis [1966], §5 for the prime 3.

Lemma 5.3. *Suppose that we have a solution* $\underline{x}$ *to the congruences*

$$\left.\begin{aligned} F(\underline{x}) &= c_1x_1^3 + \ldots + c_sx_s^3 \equiv 0 \pmod 9 \\ G(\underline{x}) &= d_1x_1^2 + \ldots + d_sx_s^2 \equiv 0 \pmod 3 \end{aligned}\right\}$$

for which there are indices i *and* j *with*

$$x_ix_j(c_id_jx_i - c_jd_ix_j) \not\equiv 0 \pmod 3.$$

Then there is a non-trivial 3-adic solution to the equations (1.2).

Proof: By induction. Suppose that for some $\mu \geq 2$, we have a solution to the simultaneous congruences

$$\left.\begin{aligned} F(\underline{x}) &\equiv 0 \pmod{3^{\mu}} \\ G(\underline{x}) &\equiv 0 \pmod{3^{\mu-1}} \end{aligned}\right\}$$

satisfying, for some i and j,

$$x_ix_j(c_id_jx_i - c_jd_ix_j) \not\equiv 0 \pmod 3 . \tag{5.3}$$

Write $y_i = x_i + k_i3^{\mu-1}$, for some k_i to be determined. Then

$$\left.\begin{aligned} F(\underline{y}) &\equiv F(\underline{x}) + 3^{\mu}(c_1k_1x_1^2 + \ldots + c_sk_sx_s^2) \pmod{3^{\mu+1}} \\ G(\underline{y}) &\equiv G(\underline{x}) + 3^{\mu-1}(2d_1k_1x_1 + \ldots + 2d_sk_sx_s) \pmod{3^{\mu}} \end{aligned}\right\}.$$

If we can solve the simultaneous congruences

$$\left.\begin{aligned} 3^{-\mu}F(\underline{x}) + c_1k_1x_1^2 + \ldots + c_sk_sx_s^2 &\equiv 0 \\ 3^{1-\mu}G(\underline{x}) + 2d_1k_1x_1 + \ldots + 2d_sk_sx_s &\equiv 0 \end{aligned}\right\} \pmod 3 \tag{5.4}$$

then we have a solution $\underline{y}$ to the simultaneous congruences

$$\left.\begin{aligned} F(\underline{y}) &\equiv 0 \pmod{3^{\mu+1}} \\ G(\underline{y}) &\equiv 0 \pmod{3^{\mu}} \end{aligned}\right\}$$

satisfying

$$y_iy_j(c_id_jy_i - c_jd_iy_j) \equiv x_ix_j(c_id_jx_i - c_jd_ix_j) \not\equiv 0 \pmod 3 .$$

But the linear forms in $\underline{k}$ in the simultaneous congruences (5.4) are non-proportional (mod 3), by (5.3). Then we may solve for the k_i,

and hence the inductive hypothesis holds with μ replaced by $\mu+1$. But by hypothesis the result holds for $\mu = 2$, so the result holds for all $\mu \geq 2$. So we have a non-trivial solution $\underline{\underline{x}}$ to the congruences $F(\underline{\underline{x}}) \equiv G(\underline{\underline{x}}) \equiv 0 \pmod{3^{\mu}}$ for all μ, and hence a non-trivial 3-adic solution to the system (1.2).

This completes the proof of the lemma.

By §2, we may write the equations (1.2) in 3-normalised form $F = G = 0$, with

$$\left.\begin{aligned} F^{**} &= a_1x_1^3+\ldots+a_ux_u^3+ b_1y_1^3+\ldots+b_vy_v^3 && \equiv 0 \pmod 9\\ G^{*} &= c_1y_1^2+\ldots+c_vy_v^2 +d_1z_1^2+\ldots+d_wz_w^2 && \equiv 0 \pmod 3 \end{aligned}\right\} \quad (5.5)$$

where c_i and d_i are coprime with 3, each a_i and b_i are divisible by at most 3, and where F^{**} denotes the variables explicit in $F \pmod 9$, and G^{*} denotes those explicit in $G \pmod 3$.

Let u_1 denote the number of a_i not divisible by 3, and v_1 the number of b_i not divisible by 3. Then by (2.3), we have

$$m_0+ q_{F,1} = u+v+w \geq 7,\ q_{G,0}= v+w \geq 3,\ R_{F,1}= u+v \geq 4,$$
$$m_0= u_1+v+w \geq 5, \text{ and } q_{F,0}= u_1+v_1 \geq 2.$$

Lemma 5.4. *Suppose that* $w \geq 3$. *Then there is a solution to the congruences* (5.5) *satisfying condition* (5.3) *of Lemma* 5.3.

Proof: First solve the cubic equation

$$a_1x_1^3+ \ldots + a_ux_u^3+ b_1y_1^3+ \ldots + b_vy_v^3 \equiv 0 \pmod 9.$$

This is possible non-trivially, since $u+v \geq 4$, and $u_1+v_1 \geq 2$ together imply that by a change of variable, one of the following congruences subsists in the above congruence:

$$x^3 + y^3 \equiv 0 \pmod 9 \text{ or } x^3 + 2y^3 + 3z^3 \equiv 0 \pmod 9.$$

Both congruences are soluble non-trivially, and further, soluble with a variable non-zero (mod 3) having a cubic coefficient non-zero

(mod 3). Fixing the solution to the cubic, we can then independently solve the congruence

$$d_1 z_1^2 + \ldots + d_w z_w^2 \equiv -(c_1 y_1^2 + \ldots + c_v y_v^2) \pmod 3.$$

Since $w \geq 3$, this may always be solved non-trivially, by Corollary 3.2.2. But now, as the cubic coefficients of the z_i are congruent to zero (mod 3), condition (5.3) of Lemma 5.3 is satisfied.

This completes the proof of the lemma.

Lemma 5.5. *Subject to the truth of the hypothesis (H3) :*

(H3) for all congruences of the form (5.5) *with* $u+v+w \geq 7$, $u+v \geq 4$, $v+w \geq 3$, $u_1+v+w \geq 5$, $u_1+v_1 \geq 2$, *and* $w < 3$, *we have a solution satisfying* (5.3),

every 3-normalised system of the form (1.2) *with* $s \geq 11$ *has a non-trivial 3-adic solution.*

Proof: This follows directly by an application of our version of Hensel's Lemma (Lemma 5.3) combined with *(H3)* and Lemma 5.4.

This completes the proof of the lemma.

6. THE COMPUTATIONAL METHOD FOR THE PRIMES 3,7,11,13,19,23 AND 31.

(a) Primes other than 3.

Consider the simultaneous congruences

$$\left.\begin{aligned} a_1 x_1^3 + \ldots + a_5 x_5^3 &\equiv 0 \\ b_1 x_1^2 + \ldots + b_5 x_5^2 &\equiv 0 \end{aligned}\right\} \pmod p \qquad (6.1)$$

where $0 \leq a_i, b_i \leq p-1$, at most one each of the a_i and b_i is zero, and a_i and b_i are not simultaneously zero. To establish the validity of hypothesis *(H)*, following the conclusions of the previous two

sections, we have only to check that the congruences (6.1) have a solution non-singular (mod p) for every permissible choice of the a_i and b_i. Further, if $p \equiv 1 \pmod 3$ and $p > 13$, or if $p \equiv 2 \pmod 3$, then we may also assume that no a_i is zero.

By the change of variables $x_i \longrightarrow \omega x_i$, for a suitable ω, we may assume that each b_i is 0, 1, or c, where c is any quadratic non-residue (mod p). Also, by independently making the change of variables $x_i \longrightarrow -x_i$, we may assume that $0 \leq a_i \leq (p-1)/2$. Further, by multiplying through by a suitable non-zero factor and rearranging variables, we may assume that $b_1 \leq b_2 \leq \ldots \leq b_5$, $b_1 = 0$ or 1, and $b_2 = b_3 = 1$. Finally, if for any j we have $b_j = b_{j+1}$, then we may assume that $a_j \leq a_{j+1}$.

Suppose that $\mathcal{S}$ is the set of all $(a_1,\ldots,a_5,b_1,\ldots,b_5)$, satisfying the above criteria. If for some $\alpha,\beta,\gamma,\kappa,\lambda,\mu$ we can find a non-singular solution to the congruences

$$\left.\begin{aligned} \alpha x^3 + \beta y^3 + \gamma z^3 &\equiv 0 \\ \kappa x^2 + \lambda y^2 + \mu z^2 &\equiv 0 \end{aligned}\right\} \pmod p,$$

then any simultaneous congruences of the form (6.1) with

$$(a_i, a_j, a_k, b_i, b_j, b_k) \equiv (\alpha,\beta,\gamma,\kappa,\lambda,\mu) \pmod p,$$

for some distinct $i,j,k \in \{1,\ldots,5\}$, will admit a solution non-singular (mod p). We used this idea to economise further on the number of cases to be checked, starting by testing all forms in three variables satisfying the above criteria, and storing those for which we could find no solution non-singular (mod p). We then took each of these forms, and tested all forms in four variables satisfying the above criteria obtained by considering any permissable coefficients for a fourth variable, again storing all forms possessing no non-singular solution. If necessary, we then proceeded in like manner to a fifth variable.

Using this pseudo-sieve method, we used a computer to check that all simultaneous congruences of the form (6.1) possessed solutions non-singular (mod p).

(b) The prime 3.

Consider the simultaneous congruences

$$\left.\begin{aligned} a_1x_1^3 + \ldots + a_7x_7^3 &\equiv 0 \pmod 9 \\ b_1x_1^2 + \ldots + b_7x_7^2 &\equiv 0 \pmod 3 \end{aligned}\right\} \qquad (6.2)$$

where $0 \le a_i \le 8$, $0 \le b_i \le 2$, and the coefficients satisfy the conditions of hypothesis *(H3)*. To establish hypothesis *(H3)*, following the conclusions of the previous section, we have only to check that the congruences (6.2) have a solution satisfying condition (5.3), for every permissable choice of the a_i and b_i. By the change of variables $x_i \longrightarrow -x_i$, we may assume that $0 \le a_i \le 4$. By multiplying the quadratic congruence through by a suitable non-zero factor and rearranging variables, we may then assume that the number of b_i equal to 1 is at least as great as the number equal to 2. Also, by a rearrangement of variables we may assume that $b_1 \le b_2 \le \ldots \le b_7$. Finally, if for any j we have $b_j = b_{j+1}$, then we may assume that $a_j \le a_{j+1}$.

As the system of congruences (6.2) is more complicated than (6.1), the pseudo-sieve approach is less successful. However, by noting that 0, 1^3 and 2^3 represent all the cubic residues (mod 9), we are able to economise by testing for solutions of (6.2) using only $x_i \in \{0,1,2\}$. We were thus able to use a computer to check that all simultaneous congruences of the form (6.2) possessed solutions satisfying condition (5.3).

7. FINDING A P-ADIC SOLUTION.

The computational checks outlined in the previous section complete the proof that hypotheses *(H)* and *(H3)* hold, and hence we have shown, subject to the assumptions that $s \geq 11$ and $\partial(F,G) \neq 0$, that one of the following hold:

(i) the system $F = G = 0$ has a non-trivial solution over $\mathbb{Z}_p$,

(ii) the system $F \equiv G \equiv 0 \pmod{p}$ has a non-singular solution $\pmod{p}$.

In the latter case, a Hensel's Lemma argument will also give us a non-trivial p-adic solution. We now remove the assumption $\partial(F,G) \neq 0$. For this purpose we use the argument of §5 of Davenport and Lewis [1967]. Although our argument will differ only in the details, it is included for the sake of completeness.

Lemma 7.1. *For any system of the form* (1.2) *in at least* 11 *variables (irrespective of the value of* $\partial(F,G)$*),* $F = G = 0$ *has a non-trivial p-adic solution.*

Proof: We have proved that $F = G = 0$ has a non-trivial p-adic solution provided $\partial(F,G) \neq 0$. Suppose now that $\partial(F,G) = 0$. For every integer μ there exist forms

$$\left.\begin{aligned} F^{(\mu)} &= c_1^{(\mu)}x_1^3 + \ldots + c_s^{(\mu)}x_s^3 \\ G^{(\mu)} &= d_1^{(\mu)}x_1^2 + \ldots + d_s^{(\mu)}x_s^2 \end{aligned}\right\}$$

with rational integer coefficients such that $\partial(F^{(\mu)},G^{(\mu)}) \neq 0$ and such that

$$c_i^{(\mu)} - c_i \equiv d_i^{(\mu)} - d_i \equiv 0 \pmod{p^\mu} \quad \text{for } i = 1,\ldots,s.$$

But then the equations $F^{(\mu)} = G^{(\mu)} = 0$ have a non-trivial p-adic solution $\underline{\xi}^{(\mu)} = (\xi_1^{(\mu)},\ldots,\xi_s^{(\mu)})$. By homogeneity, we can suppose that the $\xi_i^{(\mu)}$ are p-adic integers, and that one at least is not divisible by p. By the compactness of the p-adic integers, the set $\{\underline{\xi}^{(\mu)}\}$ has

an accumulation point $\underline{\xi} \neq 0$. Then if μ goes to infinity through a suitable sequence, we have that

$$\lim_{\mu} \underline{\xi}^{(\mu)} = \underline{\xi}$$

exists in the p-adic sense, and

$$\begin{aligned} |F(\underline{\xi}^{(\mu)})|_p &= |F(\underline{\xi}^{(\mu)}) - F^{(\mu)}(\underline{\xi}^{(\mu)})|_p \\ &= |\sum_i (c_i^{(\mu)} - c_i)\xi_i^{(\mu)3}|_p \\ &\leq p^{-\mu}. \end{aligned}$$

By continuity, we therefore have $F(\underline{\xi}) = 0$, and similarly $G(\underline{\xi}) = 0$. Hence the system (1.2) has a non-trivial p-adic solution.

This completes the proof of the lemma.

Thus any system in at least 11 variables has, for every rational prime p, a non-trivial solution in p-adic integers, i.e. $\Gamma^*(3,2) \leq 11$. We now demonstrate that $\Gamma^*(3,2) \geq 11$, which will complete the proof of Theorem 1.1.

LEMMA 7.2. *We have* $\Gamma_p^*(3,2) \geq 11$ *for all primes* $p \equiv 1 \pmod 3$.

Proof: Let $p \equiv 1 \pmod 3$ be a rational prime, c be a cubic non-residue (mod p), and ω be a quadratic non-residue (mod p). Consider the simultaneous equations in 10 variables:

$$\left.\begin{aligned} (x_1^3+cx_2^3) + p(x_3^3+cx_4^3) + p^2(x_5^3+cx_6^3) &= 0 \\ (x_7^2-\omega x_8^2) + p(x_9^2-\omega x_{10}^2) &= 0 \end{aligned}\right\} \quad (7.1)$$

The cubic equation has no non-trivial solution over $\mathbb{Z}_p$, since c is a cubic non-residue (mod p). In addition, the quadratic equation has no non-trivial solution in $\mathbb{Z}_p$, since ω is a quadratic non-residue (mod p). Hence the system (7.1) has no non-trivial solution over $\mathbb{Z}_p$.

This completes the proof of the lemma.

CHAPTER 2.

A RESULT ON THE P-ADIC SOLUBILITY OF PAIRS OF EQUATIONS.

1. INTRODUCTION.

Let c_i and d_i $(1 \le i \le s)$ be rational integers, and k and n be natural numbers. We shall consider the solubility over the p-adic integers $\mathbb{Z}_p$ of the pair of additive equations

$$\left.\begin{aligned} f(\underline{x}) &= c_1 x_1^k + \ldots + c_s x_s^k = 0 \\ g(\underline{x}) &= d_1 x_1^n + \ldots + d_s x_s^n = 0 \end{aligned}\right\} \tag{1.1}$$

Let $k \ge n > 1$, and let $p \equiv 1 \pmod{k}$ and $p \equiv 1 \pmod{n}$. Suppose that q is a kth power non-residue (mod p), and q' is a nth power non-residue (mod p). Then the equations

$$\left.\begin{aligned} \sum_{i=1}^{k} p^{i-1}(x_i^k - q y_i^k) &= 0 \\ \sum_{i=1}^{n} p^{i-1}(x_i'^n - q' y_i'^n) &= 0 \end{aligned}\right\} \tag{1.2}$$

have no simultaneous non-trivial solution over $\mathbb{Z}_p$. Thus for infinitely many primes p, we have $\Gamma_p^*(k,n) > 2(k+n)$. When $k = n$, Atkinson and Cook [1989] have made progress in the opposite direction:

Theorem 1.1. *Suppose that $k \ge 2$ and p is a rational prime with $p > k^6$. Then $\Gamma_p^*(k,k) \le 4k+1$. This result is essentially best possible in that it fails when $4k+1$ is replaced by $4k$.*

This result may be compared with the result for a single equation:

Theorem 1.2 (*see Dodson* [1966], *Lemma* 2.4.1). *Suppose that* $k \geq 2$, *and* p *is a rational prime with* $p > k^4$. *Then* $\Gamma_p^*(k) \leq 2k+1$. *This result is essentially best possible in that it fails when* $2k+1$ *is replaced by* $2k$.

The last lines of both theorems follow by considering examples of the form (1.2).

Thus the situation in which the exponents in equation (1.1) are equal has been resolved rather satisfactorily for all but a small set of primes. However, when $k \neq n$, little is known. In Chapter 1, we gave methods which are of use in considering the p-adic solubility of general systems of homogeneous additive equations, and we demonstrated that $\Gamma^*(3,2) = 11$. Here we shall refine the methods given in that chapter to establish the following result.

Theorem 1.3. *Suppose that* $k \geq n \geq 1$, *and that* p *is a rational prime with* $p > k^4n^2$. *Then*

$$\Gamma_p^*(k,n) \leq \begin{cases} 2(k+n)+1 & k \geq n > 1 \\ 2k+2 & k > n = 1 \\ 3 & k = n = 1 \ . \end{cases}$$

This result is essentially best possible in that it fails to hold when the right hand side is reduced.

Thus the conclusion of Theorem 1.3 rather neatly bridges the gap between Theorems 1.1 and 1.2. It would seem likely that the following generalisation holds:

Conjecture. *Let* $k_i' = k_i$ *when* $k_i > 1$ *and* $k_i' = \frac{1}{2}$ *when* $k_i = 1$. *Then for all rational primes* p *satisfying*

$$p > (Max\{k_1, k_2, \ldots, k_t\})^2 (k_1 \ldots k_t)^2$$

we have

$$\Gamma_p^*(k_1, \ldots, k_t) \leq 2(k_1' + k_2' + \ldots + k_t') + 1 . \tag{1.3}$$

If true, this result would be best possible, in that the bound (1.3) could not be reduced, by virtue of examples analogous to (1.2).

In principle, one could use a computer to establish the p-adic solubility of the equations (1.1) for the small primes excluded by the conditions of Theorem 1.3, although it should be emphasised that much preparation may be necessary (see, for example, Chapter 1, §3, where the prime 7 causes problems). Since knowledge of the p-adic solubility of the equations (1.1) is an essential prerequisite to an application of the Hardy-Littlewood method (the standard method for establishing the non-trivial rational solubility of the equations (1.1)), such questions are not without importance.

Finally, we note that there would appear to be no fundamental difficulty in extending the methods contained herein to systems of more than two simultaneous equations.

2. A REFINED P-ADIC NORMALISATION PROCEDURE.

First note that Theorem 1.3 follows at once when $k = n = 1$, and by Theorem 1.1, it also follows when $k = n$. We may therefore suppose that $k > n \geq 1$. The case with $n = 1$ follows by a slight alteration of the proof for the case $n > 1$, so we shall not give the details here.

Suppose then that $k > n > 1$, and consider the equations (1.1). We now refine the p-normalisation procedure adopted in Chapter 1 so as to more readily extract information about the coefficients of f. For $1 \le r \le s$, write $R = \{1,2,\dots,r\}$, $T = \{r+1,\dots,s\}$, and $t = |T|$, so that $r+t = s$. Now write $M = r(r-1)$, and $N = 2(r-1)n$. Define

$$\partial(f,g) = \prod_{\substack{i \neq j \\ i,j \in R}} (c_i^n d_j^k - c_j^n d_i^k) . \prod_{h \in T} c_h^N . \tag{2.1}$$

Write $\underline{\underline{f}}$ for (f,g).

Lemma 2.1. *Given a system of the form* (1.1), $\partial(\underline{\underline{f}})$ *satisfies the following:*

(i) if $\nu_1,\dots,\nu_s$ *are integers, and*

$$\underline{\underline{f}}' = \underline{\underline{f}}(p^{\nu_1}x_1,\dots,p^{\nu_s}x_s),$$

then

$$\partial(\underline{\underline{f}}') = p^{2knM\nu/r}\partial(\underline{\underline{f}}) ,$$

where $\nu = \nu_1 + \dots + \nu_s$;

(ii) if λ *and* μ *are rational numbers, and the system* $\underline{\underline{f}}''$ *is defined by*

$$f''(\underline{\underline{x}}) = \lambda f(\underline{\underline{x}}),\ g''(\underline{\underline{x}}) = \mu g(\underline{\underline{x}}) ,$$

then we have

$$\partial(\underline{\underline{f}}'') = (\lambda^n\mu^k)^M\lambda^{Nt}\partial(\underline{\underline{f}}) .$$

Proof: (i) Writing c_i' and d_i' respectively for the coefficients of x_i^k in $f'(\underline{\underline{x}})$ and x_i^n in $g'(\underline{\underline{x}})$, we have

$$c_i'^n d_j'^k - c_j'^n d_i'^k = p^{kn(\nu_i+\nu_j)}(c_i^n d_j^k - c_j^n d_i^k) ,$$

and

$$c_h'^N = p^{2knM\nu_h/r} c_h^N .$$

Then by (2.1), we have

$$\partial(\underline{\underline{f}}') = p^{2knM(\nu_1+\dots+\nu_r)/r} . p^{2knM(\nu_{r+1}+\dots+\nu_s)/r} \partial(\underline{\underline{f}}) ,$$

and the result follows.

(ii) follows simply from the definition of $\partial(f,g)$.

This completes the proof of the lemma.

As in Chapter 1, §7 (following from Davenport and Lewis [1967]), we are able to use a p-adic compactness argument to permit us to assume that $\partial(\underline{\underline{f}}) \neq 0$. This property is plainly preserved under equivalence. Before proceeding to the next lemma, recall the notation of §2 of Chapter 1.

Lemma 2.2. *A p-normalised system $\underline{\underline{f}}$ can be written in the form*

$$\left.\begin{aligned} f &= f_0 + pf_1 + \ldots + p^{k-1}f_{k-1} \\ g &= g_0 + pg_1 + \ldots + p^{k-1}g_{k-1} \end{aligned}\right\}$$

where the f_j and g_j are forms in m_j variables, and these sets are disjoint for $j = 0,\ldots,k-1$. Moreover, each of the m_j variables occurs in one at least of f_j and g_j with a coefficient not divisible by p.

The following inequalities hold for all $1 \leq r,t \leq s$ with $r+t = s$:

(i) we have

$$m_0 + \ldots + m_{j-1} \geq jr/k + jt/k \quad \text{for } j = 1,\ldots,n\ ,$$

where $k = 2kn/(k+n)$ is the harmonic mean of k and n.

(ii) with $h = f$ or g, denote by $q_{h,j}$ the number of variables occurring explicitly in the form h_j^*. We have

$$m_0 + m_1 + \ldots + m_{j-1} + q_{f,j} \geq jr/k + jt/k + r/(2k) + t/k \quad \text{for } j = 0,\ldots,n,$$

and

$$m_0 + m_1 + \ldots + m_{j-1} + q_{g,j} \geq jr/k + jt/k + r/(2n) \quad \text{for } j = 0,\ldots,n-1.$$

Proof: All the results except for (i) and (ii) follow as in Lemma 2.2 of Chapter 1.

(i) let $x_1,\ldots,x_m$, where $m = m_0 + \ldots + m_{j-1}$, denote the variables in $\underline{\underline{f}}_0,\ldots,\underline{\underline{f}}_{j-1}$ $(1 \le j \le n)$. Then the system

$$\underline{\underline{f}}'(\underline{\underline{x}}) = p^{-j}\,\underline{\underline{f}}(px_1,\ldots,px_m,x_{m+1},\ldots,x_s)\ ,$$

has integral coefficients and is equivalent to the system $\underline{\underline{f}}$. By Lemma 2.1 we have

$$\partial(\underline{\underline{f}}') = p^{-jM.(k+n)\ -\ jNt\ +\ 2knMm/r}\partial(\underline{\underline{f}})\ .$$

Then by the definition of a p-normalised system,

$$\begin{aligned} m &\ge jr(k+n)/(2kn) + jrNt/(2knM) \\ &= jr/k + jt/k\ . \end{aligned}$$

(ii) Suppose that the number of variables occurring in f_j^* is q. Let these variables be $x_{m+1},\ldots,x_{m+q}$, where $m = m_0 + \ldots + m_{j-1}$, and let $x_1,\ldots,x_m$ be the variables occurring in $\underline{\underline{f}}_0,\ldots,\underline{\underline{f}}_{j-1}$. Then for $0 \le j \le n$, the system

$$\left.\begin{aligned} f''(\underline{\underline{x}}) &= p^{-j-1}f(px_1,\ldots,px_{m+q},x_{m+q+1},\ldots,x_s) \\ g''(\underline{\underline{x}}) &= p^{-j}g(px_1,\ldots,px_{m+q},x_{m+q+1},\ldots,x_s) \end{aligned}\right\}$$

has integral coefficients, and is equivalent to $\underline{\underline{f}}$. But by Lemma 2.1, we have

$$\partial(\underline{\underline{f}}'') = p^{2knM(m+q)/r\ -\ (j(k+n)+n)M\ -\ (j+1)Nt}\,\partial(\underline{\underline{f}})\ ,$$

and hence, by the definition of a p-normalised system,

$$2knM(m+q)/r \ge M(j(k+n) + n + 2tn(j+1)/r)\ .$$

Then

$$m + q \ge rj/k + r/(2k) + (j+1)t/k\ .$$

The result follows similarly for $q_{g,j}$.

This completes the proof of the lemma.

Lemma 2.3. *Suppose that* $s \geq 2(k+n)+1$, *and that the system* (1.1) *is p-normalised. Then we may rearrange variables and change notation to write*

$$\left.\begin{aligned} f^* &= a_1x_1^k+\ldots+ a_ux_u^k+ b_1y_1^k+\ldots+ b_vy_v^k \\ g^* &= c_1y_1^n+\ldots+ c_vy_v^n+ d_1z_1^n+\ldots+ d_wz_w^n \end{aligned}\right\} \qquad (2.2)$$

where none of the a_1,b_1,c_1,d_1 *are divisible by p, and*

$$u+v+w \geq 5, \quad v+w \geq 3, \quad u+v \geq 3 .$$

Proof: We put $r = 4n+1$, $t = s-r \geq 2(k-n)$, and apply Lemma 2.2. We obtain by part (i) of that lemma,

$$\begin{aligned} m_0 &\geq (4n+1).(k+n)/(2kn) + 2 - 2n/k \\ &> 2 + 2n/k + 2 - 2n/k \end{aligned}$$

and so $m_0 \geq 5$. Also, by Lemma 2.2(ii), we have

$$q_{f,0} \geq (4n+1)/(2k) + 2 - 2n/k > 2 ,$$

and

$$q_{g,0} \geq (4n+1)/(2n) > 2 ,$$

and so $q_{f,0} \geq 3$ and $q_{g,0} \geq 3$.

This completes the proof of the lemma.

We now state two lemmata which will permit us to eliminate certain cases of (2.2).

Lemma 2.4. *Suppose that* $p \nmid abc$ *and* $p > k^4$. *Then the congruence*

$$ax^k + by^k + cz^k \equiv d \pmod{p} \qquad (2.3)$$

has a non-trivial solution (x,y,z).

Proof: For $d \equiv 0 \pmod{p}$, the result follows from Dodson [1966], Lemma 2.4.1. Meanwhile, for $d \not\equiv 0 \pmod{p}$ the result follows easily from the proof of that lemma.

This completes the proof of the lemma.

Given f and g of the form (1.1), and $\underline{x} \in \mathbb{Z}_p^s$, we shall define

$$\Delta(i,j) = \Delta(f,g;i,j) = knx_i^{n-1}x_j^{n-1}(c_i d_j x_i^{k-n} - c_j d_i x_j^{k-n}), \text{ and}$$

$$\Delta^*(f,g) = \max_{1 \le i < j \le s} \{|\Delta(i,j)|_p\} .$$

Thus if $\underline{x}$ is a solution to the congruences $f \equiv g \equiv 0 \pmod{p}$, it is non-singular (mod p) when $\Delta^*(f,g) = 1$.

Lemma 2.5. *All p-normalised systems of the form* (1.1) *satisfying* $s \ge 2(k+n)+1$, $p > k^4$, *and either* $u > 2$ *or* $w > 2$, *have a non-trivial solution over* $\mathbb{Z}_p$.

Proof: By setting certain of the variables to zero, we may assume from Lemma 2.3 that $u+v+w = 6$, $u+v \ge 3$, $v+w \ge 3$, and either $u = 3$ or $w = 3$. We shall assume that $u = 3$, the case with $w = 3$ following in a similar manner.

Since $v+w \ge 3$, by Lemma 2.4 we can solve the congruence

$$c_1 y_1^n + \ldots + c_v y_v^n + d_1 z_1^n + \ldots + d_w z_w^n \equiv 0 \pmod{p} \qquad (2.4)$$

with one at least of the variables non-zero, without loss of generality either y_1 or z_1. Also by Lemma 2.4, fixing the y_i and z_i so as to solve the congruence (2.4), we have a non-trivial solution to the congruence

$$a_1 x_1^k + \ldots + a_u x_u^k \equiv -(b_1 y_1^k + \ldots + b_v y_v^k) \pmod{p} ,$$

since $u = 3$. Without loss of generality, x_1 is non-zero. Then the system is soluble (mod p) with

$$\Delta^*(f,g) \ge \text{Max}\{|knx_1^{n-1}z_1^{n-1}(a_1 d_1 x_1^{k-n})|_p , |knx_1^{n-1}y_1^{n-1}(a_1 c_1 y_1^{k-n})|_p\}$$
$$= 1.$$

An application of Hensel's Lemma (Lemma 3.1 of Chapter 1) completes the proof of the lemma.

3. USE OF EXPONENTIAL SUMS.

After Lemmata 2.3 and 2.5, we may assume, by setting certain of the variables to zero, that the equations (1.1) are of the form (2.2), with

$$u+v+w = 5,\ u+v \geq 3,\ v+w \geq 3,\ u \leq 2 \text{ and } w \leq 2. \tag{3.1}$$

Let us write the equations in the form

$$\left.\begin{aligned} A_1x_1^k + \ldots + A_5x_5^k \\ B_1x_1^n + \ldots + B_5x_5^n \end{aligned}\right\} \equiv 0 \pmod p \tag{3.2}$$

where the A_i and B_i are not simultaneously zero, and these coefficients satisfy condition (3.1). The number, N, of solutions (mod p) to the congruences (3.2) is given by the exponential sum

$$N = p^{-2} \sum_{\alpha,\beta(\mathrm{mod}\ p)} T_1(\underline{\alpha})\ldots T_5(\underline{\alpha}) , \tag{3.3}$$

where

$$T_j(\underline{\alpha}) = T_j(\alpha,\beta) = \sum_{x \bmod p} e((A_j\alpha x^k + B_j\beta x^n)/p) \text{ , for } j = 1,\ldots,5.$$

Lemma 3.1. *Suppose that $p > k$, and $A_iB_i \not\equiv 0$ (mod p). Then we have*

$$p^{-2} \sum_{\underline{\alpha} \bmod p} |T_i(\underline{\alpha})|^4 \leq knp^2.$$

Proof: We have that

$$p^{-2} \sum_{\underline{\alpha} \bmod p} |T_i(\underline{\alpha})|^4$$

is the number of solutions over $\mathbb{F}_p$ to the simultaneous equations

$$\left.\begin{aligned} x_1^k + x_2^k = y_1^k + y_2^k \\ x_1^n + x_2^n = y_1^n + y_2^n \end{aligned}\right\} \tag{3.4}$$

Suppose that $(\underline{x},\underline{y})$ is any solution of (3.4). Write $d = (k,n)$. Then by eliminating y_2 from the equations (3.4) we obtain

$$(x_1^k + x_2^k - y_1^k)^{n/d} = (x_1^n + x_2^n - y_1^n)^{k/d}. \tag{3.5}$$

On noting that k/d and n/d cannot both be even, there are two possibilities:

(i) k/d and n/d are not both odd.

Then for each of the p^2 possible choices for x_1 and x_2, there are at most kn/d solutions in y_1 to the equation (3.5). This is because y_1 satisfies some non-trivial polynomial of degree kn/d, and hence there are at most kn/d solutions in y_1 to (3.5). y_2 is then determined from the equations

$$\left.\begin{aligned} y_2^k &= x_1^k + x_2^k - y_1^k \\ y_2^n &= x_1^n + x_2^n - y_1^n \end{aligned}\right\} .$$

Thus, for some z, we have that y_2 satisfies

$$y_2^d = z ,$$

to which there are at most d solutions. So in this case the total number of solutions to (3.4) is

$$\leq (kn/d)p^2.d = knp^2.$$

(ii) k/d and n/d are both odd.

Suppose first that

$$\left.\begin{aligned} x_1^k + x_2^k &= 0 \\ x_1^n + x_2^n &= 0 \end{aligned}\right\} . \qquad (3.6)$$

There are at most dp solutions to the equations (3.6), since for each of the p possible choices for x_2, we can eliminate to obtain an equation of the form $x_1^d = z$ for x_1. Repeating the same argument for y_1 and y_2, we deduce that the total number of solutions of this type is at most $(dp)^2$.

Suppose now that (3.6) does not hold. Then picking any of the remaining $\leq p^2$ possible choices for x_1 and x_2, the non-trivial polynomial in y_1 in (3.5) is of degree at most $n(k/d - 1)$. So in this case there are at most $n(k/d - 1)p^2.d$ solutions to (3.4).

Then the total number of solutions in the second case is

$$\leq knp^2 - ndp^2 + d^2p^2 \leq knp^2.$$

This completes the proof of the lemma.

Lemma 3.2. *Write*

$$f_i(\underline{\underline{\alpha}}) = T_i(\underline{\underline{\alpha}}) \ , \ g_j(\alpha) = T_j(\alpha,0) \ , \ \text{and} \ h_m(\beta) = T_m(0,\beta). \qquad (3.7)$$

Suppose also that A_i, B_i, A_j *and* B_m *are each non-zero. Then we have*

(*i*) $p^{-2} \sum_{\underline{\underline{\alpha}} \bmod p} |f_i(\underline{\underline{\alpha}})|^2 |g_j(\alpha)|^2 \le knp^2,$

(*ii*) $p^{-2} \sum_{\underline{\underline{\alpha}} \bmod p} |f_i(\underline{\underline{\alpha}})|^2 |h_m(\beta)|^2 \le knp^2,$

(*iii*) $p^{-2} \sum_{\underline{\underline{\alpha}} \bmod p} |g_j(\alpha)|^2 |h_m(\beta)|^2 \le knp^2.$

Proof: The method of proof is typified by the proof of (i).

$$p^{-2} \sum_{\underline{\underline{\alpha}} \bmod p} |f_i(\underline{\underline{\alpha}})|^2 |g_j(\alpha)|^2$$

is the number of solutions over $\mathbb{F}_p$ to the simultaneous equations

$$\left.\begin{aligned} A_i x^k + A_j y^k &= A_i x'^k + A_j y'^k \\ B_i x^n &= B_i x'^n \end{aligned}\right\} \qquad (3.8)$$

The number of solutions to the nth power equation in (3.8) is at most $n(p-1)+1$. Given any such solution, the number of solutions to the kth power equation

$$A_j(y^k - y'^k) = -A_i(x^k - x'^k)$$

is at most kp, since for each of the p possible choices for y', there are at most k possible solutions in y.

Then the total number of solutions to the equations (3.8) is

$$\le (n(p-1)+1).kp \le knp^2 \ .$$

This completes the proof of the lemma.

Lemma 3.3. *If* $p > k$, *then we have*

$$|N - p^3| \le (kn-1)(k-1)p^{5/2}.$$

Proof: Write

$$\sum_{\underline{\underline{\alpha}}}^{*} \text{ for } \sum_{\substack{\underline{\underline{\alpha}} \bmod p \\ \underline{\underline{\alpha}} \ne \underline{\underline{0}}}} .$$

Then defining f, g and h as in (3.7), we have by (2.2) that

$$\sum_{\underline{\alpha}}^{*} \prod_{i=1}^{5} T_i(\underline{\alpha}) = \sum_{\underline{\alpha}}^{*} \left(\prod_{i=1}^{u} g_i(\alpha) \prod_{j=u+1}^{u+v} f_j(\alpha,\beta) \prod_{m=u+v+1}^{5} h_m(\beta) \right) .$$

Then by Hölder's inequality,

$$\begin{aligned} &\left| \sum_{\underline{\alpha}}^{*} \prod_{i=1}^{5} T_i(\underline{\alpha}) \right| \\ &\quad \le \prod_{i=1}^{u} \prod_{j=u+1}^{u+v} \prod_{m=u+v+1}^{5} \left(\sum_{\underline{\alpha}}^{*} |g_i(\alpha)|^u |f_j(\alpha,\beta)|^v |h_m(\beta)|^w \right)^{1/(uvw)} \\ &\quad \le \sum_{\underline{\alpha}}^{*} |g_I(\alpha)|^u |f_J(\alpha,\beta)|^v |h_M(\beta)|^w \end{aligned}$$

for some indices I, J and M corresponding to the maximum of the sums in the second expression. For the moment, write f, g, h respectively for $f_J(\alpha,\beta)$, $g_I(\alpha)$, and $h_M(\beta)$. Then by (3.1), we can find positive real numbers λ, μ, ν, ρ, such that $\lambda + \mu + \nu + \rho = 1$, and

$$g^u f^{v-1} h^w = (f^2 g^2)^\lambda . (g^2 h^2)^\mu . (f^2 h^2)^\nu . (f^4)^\rho ,$$

by repeated subdivision. So by Hölder's inequality,

$$\sum_{\underline{\alpha}}^{*} |g_I(\alpha)|^u |f_J(\alpha,\beta)|^{v-1} |h_M(\beta)|^w \le S_\lambda S_\mu S_\nu S_\rho ,$$

where

$$S_\lambda = \left(\sum_{\underline{\alpha}}^{*} |g_I(\alpha)|^2 |f_J(\alpha,\beta)|^2 \right)^\lambda ,$$

and S_μ, S_ν, S_ρ are sums over g^2h^2, f^2h^2 and f^4, in obvious notation (here we adopt the natural convention of taking S_λ to be 1 when $\lambda = 0$, and similarly for μ, ν, ρ). But

$$\begin{aligned} p^{-2} \sum_{\underline{\alpha}}^{*} |g_I(\alpha)|^2 |f_J(\alpha,\beta)|^2 + p^2 &= p^{-2} \sum_{\underline{\alpha}} |g_I(\alpha)|^2 |f_J(\alpha,\beta)|^2 \\ &\le knp^2 \end{aligned}$$

by Lemma 3.2, so that

$$S_\lambda \le ((kn-1)p^4)^\lambda .$$

Similarly for S_μ, S_ν and S_ρ. Hence

$$\Big| \sideset{}{^*}\sum_{\underline{\alpha}} \prod_{i=1}^{5} T_i(\underline{\alpha}) \Big| \le \Big(\sup_{\substack{\underline{\alpha} \bmod p \\ \underline{\alpha} \not\equiv \underline{0}}} |f_J(\underline{\alpha})| \Big) (kn-1)p^4.$$

But by the Riemann Hypothesis for finite fields (see for example, Schmidt [1976]), we have $|f_J(\underline{\alpha})| \le (k-1)p^{1/2}$ for all $\underline{\alpha} \not\equiv \underline{0}$. The result now follows from (3.3).

This completes the proof of the lemma.

In the next lemma we use a result contained in Appendix A to estimate the number of singular solutions to the congruences (3.2). (The reader may wish to compare this lemma with Lemma 4.5 of Chapter 1, where a stronger result is given for a simpler problem).

Lemma 3.4. *Suppose that* $p \nmid kn$. *Then the number of distinct solutions to the congruences* (3.2) *which are singular* (*mod* p) *is at most*

$$(9(k-n)^4 + 2k)p .$$

Proof: Suppose that $\alpha_1, \dots, \alpha_5$ is any solution of the congruences (3.2) singular (mod p). Suppose that precisely q of the α_i are non-zero, and rearrange variables so that $\alpha_i \not\equiv 0 \pmod p$ for $i = 1, \dots, q$. Plainly, as A_i and B_i are not simultaneously zero, we have $q = 0$ or $q \ge 2$. Since $\underline{\alpha}$ is a singular solution and $p \nmid kn$, we must have

$$A_i B_j \alpha_i^{k-n} \equiv A_j B_i \alpha_j^{k-n} \pmod p \tag{3.9}$$

for all i and j in $\{1, \dots, q\}$.

Suppose that $q \ge 2$. There are three cases:

(i) There is an $i \in \{1,\dots,q\}$ with $A_i \equiv 0 \pmod p$.

Then for all $j \neq i$ with $j \in \{1,\dots,q\}$ we have from (3.9) that

$$A_j B_i \alpha_j^{k-n} \equiv 0 \pmod p ,$$

and hence, as $B_i \alpha_j^{k-n} \not\equiv 0 \pmod p$, we obtain $A_j \equiv 0 \pmod p$ for $j = 1,\dots,q$. But then by (3.1) we have $q \leq 2$, and hence $q = 2$. So $\underline{\alpha}$ is a non-trivial solution to the congruence $B_1\alpha_1^n + B_2\alpha_2^n \equiv 0 \pmod p$, to which there are at most $n(p-1)$ non-trivial solutions.

So there are at most $n(p-1)$ solutions in this case.

(ii) There is an $i \in \{1,\dots,q\}$ with $B_i \equiv 0 \pmod p$.

A similar argument reveals that there are at most $k(p-1)$ solutions possible in this case.

(iii) There is no $i \in \{1,\dots,q\}$ for which either $A_i \equiv 0$ or $B_i \equiv 0 \pmod p$.

Then from (3.9),

$$(\alpha_j/\alpha_i)^{k-n} \equiv (A_iB_j)/(B_iA_j) \pmod p$$
$$\text{for all } i \neq j \text{ in } \{1,\dots,q\}. \qquad (3.10)$$

The congruences (3.10) must be soluble (mod p) for all $i \neq j$ with $i,j \in \{1,\dots,q\}$, for otherwise there could be no singular solutions of this type. But the congruence

$$x^{k-n} \equiv (A_1B_j)/(B_1A_j) \pmod p \qquad (3.11)$$

has at most $k-n$ solutions $\xi_1^{(j)},\dots,\xi_{k-n}^{(j)}$, so there are at most $(k-n)^{q-1}$ possible choices for $(\xi^{(2)},\dots,\xi^{(q)})$ satisfying the congruences

$$\left.\begin{aligned} A_1 + A_2(\xi^{(2)})^k + \dots + A_q(\xi^{(q)})^k &\equiv 0 \\ B_1 + B_2(\xi^{(2)})^n + \dots + B_q(\xi^{(q)})^n &\equiv 0 \end{aligned}\right\} \pmod p .$$

Hence

$$\beta_1 + \dots + \beta_q \equiv 0 \pmod p ,$$

where

$$\beta_i \equiv \left(\omega B_i^k/A_i^n\right)^{1/(k-n)} \pmod p \quad \text{for } i = 1,\dots,q , \qquad (3.12)$$

and ω may be taken to be any (fixed) factor required to guarantee the existence of the $(k-n)$th root. Note that our choice of the $\xi^{(j)}$ effectively forces a particular choice of roots in (3.12).

Rearrange variables so that $A_i \not\equiv 0$ and $B_i \not\equiv 0 \pmod{p}$ for $i = 1,\dots,g$, and either $A_i \equiv 0$ or $B_i \equiv 0 \pmod{p}$ for $g < i \le 5$. Using the notation of Theorem 1.1 of Appendix A, for each set of choices of the root in (3.11), there are at most $N(p;\underline{\underline{\beta}})$ distinct congruences of the form

$$\beta_{i_1} + \dots + \beta_{i_t} \equiv 0 \pmod{p}$$

with $1 \le i_1 < i_2 < \dots < i_t \le g$ and $0 \le t \le g$, including the empty congruence. For each such non-empty congruence, we have at most $(k-n)^4(p-1)$ non-trivial solutions to the congruences (3.2) (given by the $p-1$ non-zero values of α_1, and the $(k-n)^4$ possible choices of the roots in (3.11)). Thus, by Theorem 1.1 of Appendix A, we find that the number of singular solutions of type (iii) to the congruences (3.2) is at most

$$(10-1)(k-n)^4(p-1) .$$

Then taking all possible singular solutions into account, the total number of singular solutions to the equations (3.2) is at most

$$(k+n)(p-1) + 9(k-n)^4(p-1) + 1 < (9(k-n)^4 + 2k)p .$$

This completes the proof of the lemma.

Lemma 3.5. *Suppose that $p > k^4n^2$ is a rational prime, and the congruences* (3.2) *satisfy the conditions implicit in* (3.1). *Then the congruences admit a non-singular solution (mod p).*

Proof: By Lemmata 3.3 and 3.4, the number, N^*, of non-singular solutions to the congruences (3.2) satisfies

$$|N^* - p^3| \le (kn-1)(k-1)p^{5/2} + (9(k-n)^4 + 2k)p .$$

Since $p > k^4n^2$, we have

$$(kn-1)(k-1) + p^{-3/2}(9(k-n)^4 + 2k) - k^2n$$
$$< -(n+1)k+1+(k^6n^3)^{-1}(9(k-n)^4+2k) \qquad (3.13)$$

Recall that we are assuming that $k > n > 1$. Then $n \geq 2$, and $k \geq 3$, so that the right hand side of (3.13) is

$$\leq -3k+1+11/k^2 < 0 \ . \qquad (3.14)$$

Then by (3.13) and (3.14), we have $|N^* - p^3| < k^2np^{5/2}$, and hence

$$N^* > p^3 - k^2np^{5/2} > 0.$$

Thus there is at least one non-singular solution (mod p) to the congruences (3.2), as required.

This completes the proof of the lemma.

We may now use Hensel's Lemma (Lemma 3.1 of Chapter 1) to deduce that the equations (1.1) have a non-trivial p-adic solution whenever $p > k^4n^2$.

This completes the proof of Theorem 1.3.

PART II.

SIMULTANEOUS ADDITIVE EQUATIONS: THE RATIONAL PROBLEM.

A modified version of Part II has been submitted for publication to the Journal für die reine und angewandte Mathematik as the second part of the series "On simultaneous additive equations".

CHAPTER 3.

THE NEW ITERATIVE METHOD FOR SIMULTANEOUS EQUATIONS.

1. INTRODUCTION TO CHAPTERS 3 AND 4.

Let c_{ij} $(1 \le i \le t,\ 1 \le j \le s)$ be rational integers, and k_i $(1 \le i \le t)$ be natural numbers. Consider the simultaneous diophantine equations

$$F_i(\underline{x}) = c_{i1}x_1^{k_i} + \ldots + c_{is}x_s^{k_i} = 0 \quad (1 \le i \le t)\ . \tag{1.1}$$

Define $G^*(\underline{k}) = G^*(k_1,\ldots,k_t)$ to be the least integer r such that for all $s \ge r$, and all c_{ij} $(1 \le i \le t,\ 1 \le j \le s)$ satisfying all conditions imposed by local solubility considerations (in particular, satisfying the congruence condition as defined in §1 of Chapter 1), the simultaneous equations (1.1) have a non-trivial rational solution.

There has been much interest in the problem of establishing bounds for $\Gamma^*(\underline{k})$ and $G^*(\underline{k})$ over the last quarter of a century, mostly stemming from the pioneering investigations of Davenport and Lewis [1963, 1966, 1967, 1969]. However, until Chapter 1, it seems that only equations with the k_i all equal had been considered. In that chapter we gave methods of use in considering the p-adic solubility of the equations (1.1). In particular, we considered the p-adic solubility of the equations

$$\left.\begin{aligned} F(\underline{x}) &= c_1x_1^3 + \ldots + c_sx_s^3 = 0 \\ G(\underline{x}) &= d_1x_1^2 + \ldots + d_sx_s^2 = 0 \end{aligned}\right\} \tag{1.2}$$

where $c_i, d_i \in \mathbb{Z}$ for $i = 1,\ldots,s$, and we were able to show that $\Gamma^*(3,2) = 11$. In this section we go on to investigate the rational solubility of the equations (1.2). We shall prove:

Theorem 1.1 ("$G^*(3,2) \le 14$"). *The simultaneous equations* (1.2) *have a non-trivial solution in rational integers if the following conditions hold:*

(a) *the quadratic equation in* (1.2) *is indefinite, and has at least* 5 *variables explicit, and*

(b) *the cubic equation in* (1.2) *has at least* 7 *variables explicit, and*

(c) *the simultaneous equations* (1.2) *have a non-trivial real solution, and*

(d) (*i*) $s \ge 14$, *or*

(*ii*) *at least* 6 *of the* d_i *are zero, or*

(*iii*) *at least* 4 *of the* c_i *are zero.*

Theorem 1.1 may be compared with results obtained by previous workers for small values of t and k_i:

$$G^*(2) \le 5 \text{ (classical)}, \quad G^*(3) \le 7 \text{ (R.Baker [1990])},$$

$$G^*(2,2) \le 9 \text{ (Cook [1971])},$$

$$G^*(3,3) \le 15 \text{ (R.Baker and Brüdern [1988])} .$$

The result of Cook for $G^*(2,2)$ involves the condition that any linear form in the equations must be indefinite with at least five variables explicit. The result for $G^*(3,3)$ is the latest in a long series: Davenport and Lewis [1966] obtained $G^*(3,3) \le 18$, "18" being reduced to "17" by Cook [1972] and to "16" by Vaughan [1977]. In view of the recent paper Vaughan [1989], it seems likely that "15" could be reduced to "14" using current technology (and in fact, in a recent conversation, Dr. J.Brüdern indicated that he had proved $G^*(3,3) \le 14$ by using ideas from Vaughan [1986a]).

It is to be hoped that our consideration of this particular example may stimulate interest in the more general problem, and to

this end some of our methods are set out in a quite general form. The proof of Theorem 1.1 is an application of the Hardy-Littlewood method, based on a generalisation of the methods of Vaughan [1989a] to simultaneous equations of differing degree. These methods permit us to prove:

Theorem 1.2. *Let* $\mathcal{A}(P,R) = \{ n \in [1,P] : p$ *prime and* $p|n$ *implies* $p \le R \}$. *Suppose that* $\varepsilon > 0$, $0 < \eta < \eta_0(\varepsilon)$ *and* $P > 1$. *Then*

$$\iint_{[0,1]^2} \Big| \sum_{x \in \mathcal{A}(P,P^{\eta})} e(\alpha x^3 + \beta x^2) \Big|^{10} d\alpha \, d\beta \ll P^{35/6 + \varepsilon}.$$

This may be compared with a result which may be obtained by classical methods:

$$\iint_{[0,1]^2} \Big| \sum_{0<x\le P} e(\alpha x^3 + \beta x^2) \Big|^{10} d\alpha \, d\beta \ll P^{6+\varepsilon}.$$

In proving Theorem 1.2, we use an estimate which may be of independent interest:

Theorem 1.3. *Suppose that* n *and* m *are integers with* $m > n \ge 1$. *Then the number of solutions,* S, *of the simultaneous diophantine equations*

$$\left.\begin{aligned} x_1^m + x_2^m + x_3^m &= y_1^m + y_2^m + y_3^m \\ x_1^n + x_2^n + x_3^n &= y_1^n + y_2^n + y_3^n \end{aligned}\right\}$$

with $0 < x_i, y_i \le P$ $(i = 1,2,3)$, *satisfies* $S \ll P^{3+\varepsilon}$.

This estimate, which is equivalent to the estimate

$$\iint_{[0,1]^2} \Big| \sum_{0<x\le P} e(\alpha x^m + \beta x^n) \Big|^{6} d\alpha \, d\beta \ll P^{3+\varepsilon},$$

generalises the result

$$\iint_{[0,1]^2} \Big| \sum_{0<x\le P} e(\alpha x^3 + \beta x) \Big|^6 d\alpha\, d\beta \ll P^{3+\varepsilon},$$

derived in Hua [1965], and recently used by Heath-Brown [1988] in work on Weyl sums. Both results, save for the presence of the P^{ε}, are essentially best possible, by considering diagonal solutions to the corresponding diophantine equations.

We note that the methods of this section could be applied to the corresponding generalised Waring problem, of simultaneously representing integers N and M in the form

$$x_1^3 + \ldots + x_s^3 = N, \quad x_1^2 + \ldots + x_s^2 = M.$$

We shall not pursue this application here.

In the application of the Hardy-Littlewood method to additive problems of the type described above, a fundamental rôle is played by estimates for the number of solutions of auxiliary equations of the form

$$x_1^{k_i} + \ldots + x_s^{k_i} = y_1^{k_i} + \ldots + y_s^{k_i} \quad (1 \le i \le t), \tag{1.3}$$

in which $1 \le x_j, y_j \le P$. One idea for improving classical estimates, in which the x_j and y_j range over the entire interval, is to restrict the variables to lying in intervals of the form

$$P_j < x_j, y_j < 2P_j \text{ for } j = 1, \ldots, s,$$

where $P_1 \ge P_2 \ge \ldots$. The use of diminishing ranges does not, however, seem well suited to cases where the k_i are not all equal. The problem is that the method makes savings by exploiting features of the real character of solutions, which become less pronounced when the k_i are not all equal. Vaughan [1989a,b,c] has shown that when $t = 1$ a more efficient approach is to impose restrictions on the arithmetic character of the solutions. As we shall demonstrate in this paper, this approach remains effective when $t > 1$, although

there are then a number of algebraic, as well as analytic, questions to be answered if we are to take full advantage of the method.

We consider the equations (1.3) with $x_j, y_j \in \mathcal{A}(P,R)$, for a suitable R, where

$$\mathcal{A}(P,R) = \{ n : n \le P,\ p \text{ prime}, p|n \text{ implies } p \le R \} . \qquad (1.4)$$

We then relate the number of solutions of (1.3) to the number of solutions of the simultaneous equations

$$x_1^{k_i}+\ldots+ x_r^{k_i}+ m^{k_i}(u_1^{k_i}+\ldots+ u_{s-r}^{k_i})$$

$$= y_1^{k_i}+\ldots+ y_r^{k_i}+ m^{k_i}(v_1^{k_i}+\ldots+ v_{s-r}^{k_i}) \quad (1 \le i \le t) \qquad (1.5)$$

with $x_j, y_j \le P$, $M < m \le MR$, and $u_j, v_j \in \mathcal{A}(P/M,R)$. By making use of homogeneity and Hölder's inequality, we are then able to relate the number of solutions of the equations (1.5) to the number of solutions of (1.3) with s replaced by a range of values not too far from s. Although we shall not describe all possible variations of the method in this thesis, since we are primarily interested in the equations (1.2), we hope in a subsequent paper to describe some of the further consequences of the method.

We use the methods alluded to above on the minor arcs in our application of the Hardy-Littlewood method. The treatment of the major arcs is complicated in two respects. Firstly, we are dealing with an inherently non-linear problem, which causes difficulties even in a classical approach to the problem. Secondly, we have restricted the variables to lie in the set $\mathcal{A}(P,R)$, and this causes complications.

We follow Vaughan [1989a] as far as possible when dealing with the auxiliary equations. However, we aim to give a reasonably self contained exposition, so that although some of our proofs may be obtained with little difficulty from results of Vaughan [1989a] for single equations, we nonetheless include most of the details.

In §2 we establish the reduction formula relating the number of solutions of the systems of equations (1.3) and (1.5). In order to make use of arithmetic properties of the equations, we provide an estimate for the number of solutions of a certain system of congruence equations in §3. In §4 we give estimates for the number of solutions of certain diophantine equations, deriving Theorem 1.3, and then we go on to use the conclusions of §§2 and 3 to deduce Theorem 1.2. In the second chapter of this section we consider the additive equations (1.2). In §1 we show that the conditions of Theorem 1.1 cannot be substantially relaxed. In §2 we make some simplifying observations, and describe the major and minor arcs required for our application of the Hardy-Littlewood method. In §3 we consider the minor arcs, where we must take care in dealing with any zero coefficients that are present in the equations (1.2). As we know rather little about exponential sums over the set $\mathcal{A}(P,R)$, we are forced to hard-prune the major arcs in §5. In §6 we then consider the pruned major arcs, finally establishing Theorem 1.1.

Throughout, ε will denote a sufficiently small positive number.

2. THE FUNDAMENTAL LEMMA.

We first derive a fundamental lemma which relates the number of solutions of the equations (1.3) to that of the equations (1.5). This is merely an extension of Lemma 2.1 of Vaughan [1989a], and although our proof differs only in the details, we nonetheless give it in full, for the sake of completeness. We then go on to show how algebraic features of the equations can be used to simplify these auxiliary equations.

With $\mathcal{A}(P,R)$ defined by (1.4), let $S_s(P,R)$ denote the number of solutions of the simultaneous diophantine equations

$$x_1^{k_i} + \ldots + x_s^{k_i} = y_1^{k_i} + \ldots + y_s^{k_i} \quad (1 \le i \le t), \tag{2.1}$$

in which

$$0 < t < s, \quad 0 < k_t < \ldots < k_2 < k_1 = k, \quad \text{and } x_j, y_j \in \mathcal{A}(P,R). \tag{2.2}$$

(If $t \ge s$, we can solve the system of diophantine equations by elimination to obtain $S_s(P,R) \ll P^s$, so we lose nothing by disregarding this case.)

Suppose that r is an integer satisfying $1 \le r < s$, and θ is a real number with $0 < \theta < 1$. Let $T_s(P,R,\theta;r)$ denote the number of solutions of the simultaneous diophantine equations

$$\left.\begin{aligned} x_1^{k_i} + \ldots + x_r^{k_i} + m^{k_i}(u_1^{k_i} + \ldots + u_{s-r}^{k_i}) \\ = y_1^{k_i} + \ldots + y_r^{k_i} + m^{k_i}(v_1^{k_i} + \ldots + v_{s-r}^{k_i}) \end{aligned}\right\} (1 \le i \le t) \tag{2.3}$$

with

$$1 \le x_j, y_j \le P, \quad (x_j y_j, m) = 1 \quad (1 \le j \le r), \tag{2.4}$$

$$P^\theta < m \le \text{Min}\{ P, P^\theta R \}, \quad u_j, v_j \in \mathcal{A}(P^{1-\theta}, R), \quad (1 \le j \le s-r). \tag{2.5}$$

The following lemma relates S_s to T_s. (The author is indebted to Professor Vaughan for the case $r > s/2$ of the lemma).

Lemma 2.1. *Let* $\theta = \theta(s,k_1,k_2,\ldots,k_t;r)$ *satisfy* $0 < \theta < 1$ *and suppose that* $1 \le D \le P$. *Then, for any* $1 \le r < s$,

$$S_s(P,R) \ll \Big(\sum_{d>D} (S_s(P/d,R))^{1/s} \Big)^s + S_s(D^{1-\theta}P^\theta, R)$$
$$+ P^\varepsilon \Big(\sum_{d \le D} ((P/d)^\theta R)^{c(r,s)} (T_s(P/d,R,\theta;r))^{1/s} \Big)^s,$$

where

$$c(r,s) = \begin{cases} 2 - (2r+1)/s & \text{for } 1 \le r \le s/2 \\ 1 - 1/s & \text{for } s/2 < r < s. \end{cases}$$

Proof: We shall use vector notation to avoid repeated suffices where possible. Thus, for example, we shall write

$$\underline{\alpha} \text{ for } (\alpha_1,\ldots,\alpha_t), \text{ and } \underline{d}_{=j=}^{\underline{k}}\underline{\alpha} \text{ for } (d_j^{k_1}\alpha_1,\ldots,d_j^{k_t}\alpha_t),$$

with obvious modifications where appropriate. We shall also write $\mathcal{U}_t^*$ for the t-dimensional unit cube.

For a given solution of (2.1) satisfying (2.2), let

$$d_j = (x_j, y_j) \text{ for } 1 \le j \le s.$$

Now let S' denote the number of those solutions for which $d_j > D$ for at least one j, let S'' denote the number for which

$$d_j \le D \tag{2.6}$$

for every j, and

$$\text{Max}\{x_j, y_j\} \le d_j^{1-\theta} P^{\theta} \tag{2.7}$$

for at least one j, and let S''' denote the number for which $d_j \le D$ for every j, and (2.7) holds for no j. Then

$$S_s(P,R) \le 3.\text{Max}\{S', S'', S'''\}.$$

We consider three cases:

(i) Suppose that $S' \ge \text{Max}\{S'', S'''\}$.

Then

$$S_s(P,R) \le 3S'.$$

Let

$$f(\underline{\underline{\alpha}}; Q, R) = \sum_{x \in \mathcal{A}(Q,R)} e(\alpha_1 x^{k_1} + \ldots + \alpha_t x^{k_t}) . \tag{2.8}$$

Then

$$S' \ll \sum_{d>D} \int_{\mathcal{U}_t^*} |f(\underline{\underline{d^k \alpha}}; P/d, R)^2 . f(\underline{\underline{\alpha}}; P, R)^{2s-2}| \, d\underline{\underline{\alpha}} .$$

Hence, by Hölder's inequality,

$$S_s(P,R) \ll \sum_{d>D} (S_s(P/d,R))^{1/s} (S_s(P,R))^{1-1/s},$$

and the lemma follows in case (i).

(ii) Suppose that $S'' \ge \text{Max}\{S''', S'\}$.

Then

$$S_s(P,R) \le 3S'' .$$

For a solution counted by S'' we have (2.6) for all j and (2.7) for some j, say $j = i$. Thus

$$d_1 \le D \text{ and } \mathrm{Max}\{x_1, y_1\} \le d_1^{1-\theta} P^{\theta},$$

so that

$$\mathrm{Max}\{x_1, y_1\} \le D^{1-\theta} P^{\theta}.$$

Hence

$$S'' \ll \int_{\mathfrak{U}_t^*} |f(\underline{\underline{\alpha}}; D^{1-\theta}P^{\theta}, R)^2 . f(\underline{\underline{\alpha}}; P, R)^{2s-2}| \, d\underline{\underline{\alpha}} .$$

Then by Hölder's inequality,

$$S_s(P,R) \ll (S_s(D^{1-\theta}P^{\theta}, R))^{1/s} . (S_s(P,R))^{1-1/s},$$

and so the result follows in case (ii).

(iii) Suppose that $S''' \ge \mathrm{Max}\{S', S''\}$.

Then

$$S_s(P,R) \le 3S''' . \tag{2.9}$$

Given a solution of (2.1) counted by S''', we have for every j,

$$d_j \in \mathcal{A}(D,R) \text{ and } \mathrm{Max}\{x_j, y_j\} > d_j^{1-\theta} P^{\theta}.$$

Let

$$u_j = x_j/d_j \ , \ v_j = y_j/d_j \ ,$$

so that

$$(u_j, v_j) = 1 \text{ and } \mathrm{Max}\{u_j, v_j\} > (P/d_j)^{\theta},$$

and let m_j denote the smallest divisor of $\mathrm{Max}\{u_j, v_j\}$ exceeding $(P/d_j)^{\theta}$. Since none of the prime divisors of $\mathrm{Max}\{u_j, v_j\}$ exceed R, we have

$$m_j \in \mathcal{A}(P,R) \text{ and } (P/d_j)^{\theta} < m_j \le \mathrm{Min}\{P/d_j, (P/d_j)^{\theta} R\} . \tag{2.10}$$

Thus

$$S''' \ll \sum_{\eta_1} \dots \sum_{\eta_s} S'''(\eta_1, \dots, \eta_s) \tag{2.11}$$

where the summation is over $\eta_1, \dots, \eta_s$ with $\eta_j = \pm 1$ and where $S'''(\eta_1, \dots, \eta_s)$ is the number of solutions of the system of equations

$$\sum_{j=1}^{s} \eta_j d_j^{k_i}(x_j^{k_i} - m_j^{k_i} y_j^{k_i}) = 0 \quad (i = 1, \ldots, t),$$

with

$$d_j \in \mathcal{A}(D,R),\ x_j \in \mathcal{A}(P/d_j,R),\ (x_j,m_j) = 1,\ y_j \in \mathcal{A}(P/(d_jm_j),R)$$

and m_j satisfying (2.10).

Let

$$f_{\underline{m}}(\underline{\alpha};Q,R) = \sum_{\substack{x \in \mathcal{A}(Q,R) \\ (x,m)=1}} e(\alpha_1 x^{k_1} + \ldots + \alpha_t x^{k_t}), \tag{2.12}$$

$$F_j(\underline{\alpha}) = f_{m_j}(\eta_j \underline{d}^{k}\underline{\alpha}\ ; P/d_j, R)\ f(-\eta_j \underline{d}^{k}\underline{m}^{k}\underline{\alpha}\ ; P/(d_jm_j), R)\ .$$

Then

$$S''(\eta_1,\ldots,\eta_s) \le \int_{\mathcal{U}_t^*} \prod_{j=1}^{s} \Bigg(\sum_{d_j \in \mathcal{A}(D,R)} \sum_{m_j}{}' F_j(\underline{\alpha}) \Bigg)\, d\underline{\alpha}\ ,$$

where $\sum_{m_j}{}'$ denotes summation over m_j satisfying (2.10).

We now divide into cases.

(a) If $1 \le r \le s/2$, then we let

$$X_j(\underline{\alpha}) = |f_{m_j}(\underline{d}^{k}\underline{\alpha}\ ; P/d_j, R)^{2r} f(\underline{d}^{k}\underline{m}^{k}\underline{\alpha}\ ; P/(d_jm_j), R)^{2s-2r}|\ , \tag{2.13}$$

and

$$Y(\underline{\alpha}) = |\prod_{j=1}^{s} f_{m_j}(\underline{d}^{k}\underline{\alpha}\ ; P/d_j, R)|. \tag{2.14}$$

Then, by (2.11),

$$S''' \ll \sum_{\substack{d_1 \\ d_j \in \mathcal{A}(D,R)}} \cdots \sum_{d_s} \sum_{m_1}{}' \cdots \sum_{m_s}{}' \int_{\mathcal{U}_t^*} Y(\underline{\alpha})^{\frac{s-2r}{s-r}} \prod_{j=1}^{s} \left(X_j(\underline{\alpha})^{\frac{1}{2s-2r}} \right) d\underline{\alpha}\ . \tag{2.15}$$

By Hölder's inequality we have

$$\int_{\mathcal{U}_t^*} Y(\underline{\alpha})^{\frac{s-2r}{s-r}} \prod_{j=1}^{s} \left(X_j(\underline{\alpha})^{\frac{1}{2s-2r}} \right) d\underline{\alpha}$$

$$\le \left(\int_{\mathcal{U}_t^*} Y(\underline{\alpha})^2\, d\underline{\alpha} \right)^{\frac{s-2r}{2s-2r}} \prod_{j=1}^{s} \left(\int_{\mathcal{U}_t^*} X_j(\underline{\alpha})\, d\underline{\alpha} \right)^{\frac{1}{2s-2r}}$$

and by (2.8), (2.12) and (2.14), and by considering the underlying diophantine equations, we have

$$\int_{\mathfrak{U}_t^*} Y(\underline{\underline{\alpha}})^2 \, d\underline{\underline{\alpha}} \le \int_{\mathfrak{U}_t^*} Z(\underline{\underline{\alpha}})^2 \, d\underline{\underline{\alpha}} ,$$

where

$$Z(\underline{\underline{\alpha}}) = \Big| \prod_{j=1}^{s} f(\underline{\underline{d}}_j^k \underline{\underline{\alpha}} \, ; P/d_j, R) \Big| . \tag{2.16}$$

Therefore, by Hölder's inequality and (2.13), we have

$$\sum_{m_1}^{\prime} \cdots \sum_{m_s}^{\prime} \int_{\mathfrak{U}_t^*} Y(\underline{\underline{\alpha}})^{\frac{s-2r}{s-r}} \prod_{j=1}^{s} \Big(X_j(\underline{\underline{\alpha}})^{\frac{1}{2s-2r}} \Big) \, d\underline{\underline{\alpha}}$$

$$\le \Big(\int_{\mathfrak{U}_t^*} Z(\underline{\underline{\alpha}})^2 \, d\underline{\underline{\alpha}} \Big)^{\frac{s-2r}{2s-2r}} U(P,R,\theta) ,$$

where

$$U(P,R,\theta) = \Big(\sum_{m_1}^{\prime} \cdots \sum_{m_s}^{\prime} 1 \Big)^{\frac{2s-2r-1}{2s-2r}} \Big(\sum_{m_1}^{\prime} \cdots \sum_{m_s}^{\prime} \prod_{j=1}^{s} \Big(\int_{\mathfrak{U}_t^*} X_j(\underline{\underline{\alpha}}) \, d\underline{\underline{\alpha}} \Big) \Big)^{\frac{1}{2s-2r}} .$$

But by using (2.3), (2.4) and (2.5), and considering the underlying diophantine equations, we have

$$U(P,R,\theta) \ll \Big(\prod_{j=1}^{s} ((P/d_j)^{\theta} R)^{2s-2r-1} \, T_s(P/d_j, R, \theta; r) \Big)^{\frac{1}{2s-2r}} .$$

Therefore, by (2.15) and Hölder's inequality,

$$S''' \ll \Big(\sum_{d_1}^{} \cdots \sum_{d_s}^{} \int_{\mathfrak{U}_t^*} Z(\underline{\underline{\alpha}})^2 \, d\underline{\underline{\alpha}} \Big)^{\frac{s-2r}{2s-2r}} \Big(\sum_{\substack{d_1 \\ d_j \in \mathcal{A}(D,R)}}^{} \cdots \sum_{\substack{d_s \\ d_j \in \mathcal{A}(D,R)}}^{} \prod_{j=1}^{s} V(d_j)^{1/s} \Big)^{\frac{s}{2s-2r}}$$

where

$$V(d) = ((P/d)^{\theta} R)^{2s-2r-1} \, T_s(P/d, R, \theta; r). \tag{2.17}$$

By (2.16),

$$\sum_{\substack{d_1 \\ d_j \in \mathcal{A}(D,R)}}^{} \cdots \sum_{d_s}^{} \int_{\mathfrak{U}_t^*} Z(\underline{\underline{\alpha}})^2 \, d\underline{\underline{\alpha}}$$

is the number of solutions of the system of equations

$$d_1^{k_i}x_1^{k_i} + \ldots + d_s^{k_i}x_s^{k_i} = d_1^{k_i}y_1^{k_i} + \ldots + d_s^{k_i}y_s^{k_i} \quad (1 \le i \le t),$$

with $d_j \in \mathcal{A}(D,R)$, $x_j, y_j \in \mathcal{A}(P/d_j,R)$. Hence, by using the well-known estimate for the divisor function, it is

$$\ll P^{\varepsilon}S_s(P,R).$$

Therefore, by (2.9),

$$S_s(P,R) \ll P^{\varepsilon}\Big[\sum_{d \in \mathcal{A}(D,R)} V(d)^{1/s}\Big]^s. \tag{2.18}$$

This, in view of (2.17), completes the proof of case (iii)(a).

(b) If $s/2 < r < s$, then we let

$$X_j(\underline{\alpha}) = |f_{m_j}(\underline{d}_{=j=}^{k}\underline{\alpha}\,; P/d_j,R)^{2r} f(\underline{d}_{=j=j=}^{k}\underline{m}_{=j=}^{k}\underline{\alpha}\,; P/(d_jm_j),R)^{2s-2r}|\,,$$

and

$$Y(\underline{\alpha}) = |\prod_{j=1}^{s} f(\underline{d}_{=j=j=}^{k}\underline{m}_{=j=}^{k}\underline{\alpha}\,; P/(d_jm_j),R)|.$$

Then, by (2.11),

$$S''' \ll \sum_{\substack{d_1 \\ d_j \in \mathcal{A}(D,R)}} \cdots \sum_{d_s} \sum_{m_1}' \cdots \sum_{m_s}' \int_{\mathcal{U}_t^*} Y(\underline{\alpha})^{\frac{2r-s}{r}} \prod_{j=1}^{s}\Big(X_j(\underline{\alpha})^{\frac{1}{2r}}\Big)\, d\underline{\alpha}\,.$$

The argument now proceeds in a manner similar to case (a). By repeated use of Hölder's inequality, we obtain

$$\sum_{m_1}' \cdots \sum_{m_s}' \int_{\mathcal{U}_t^*} Y(\underline{\alpha})^{\frac{2r-s}{r}} \prod_{j=1}^{s}\Big(X_j(\underline{\alpha})^{\frac{1}{2r}}\Big)\, d\underline{\alpha}$$

$$\ll \Big(\sum_{m_1}' \cdots \sum_{m_s}' \int_{\mathcal{U}_t^*} Y(\underline{\alpha})^2 d\underline{\alpha}\Big)^{\frac{2r-s}{2r}} \Big(\prod_{j=1}^{s}((P/d_j)^{\theta}R)^{s-1}\, T_s(P/d_j,R,\theta;r)\Big)^{\frac{1}{2r}}$$

Therefore, by using Hölder's inequality once again, as in case (a) we obtain

$$S''' \ll \Big(\sum_{\substack{d_1 \\ d_j \in \mathcal{A}(D,R)}} \cdots \sum_{d_s} \sum_{m_1}' \cdots \sum_{m_s}' \int_{\mathcal{U}_t^*} Y(\underline{\alpha})^2 d\underline{\alpha}\Big)^{\frac{2r-s}{2r}} \Big(\sum_{\substack{d_1 \\ d_j \in \mathcal{A}(D,R)}} \cdots \sum_{d_s} \prod_{j=1}^{s} V(d_j)^{1/s}\Big)^{\frac{s}{2r}}$$

where

$$V(d) = ((P/d)^{\theta}R)^{s-1}\, T_s(P/d, R, \theta; r). \tag{2.19}$$

But now

$$\sum_{\substack{d_1 \\ d_j \in \mathcal{A}(D,R)}} \cdots \sum_{d_s} \sum_{m_1}' \cdots \sum_{m_s}' \int_{\mathfrak{U}_t^*} Y(\underline{\underline{\alpha}})^2 \, d\underline{\underline{\alpha}}$$

is the number of solutions of the system of equations

$$(m_1 d_1 x_1)^{k_i} + \ldots + (m_s d_s x_s)^{k_i} = (m_1 d_1 y_1)^{k_i} + \ldots + (m_s d_s y_s)^{k_i}$$

$$(1 \le i \le t),$$

with $d_j \in \mathcal{A}(D,R)$, the m_j satisfying (2.10), and

$$x_j, y_j \in \mathcal{A}(P/(d_j m_j), R).$$

Hence, by using the well-known estimate for the divisor function, this is

$$\ll P^{\varepsilon} S_s(P,R).$$

(In fact a more precise analysis at this point would do rather better). Therefore, by (2.9),

$$S_s(P,R) \ll P^{\varepsilon}\Big[\sum_{d \in \mathcal{A}(D,R)} V(d)^{1/s} \Big]^s.$$

This, in view of (2.19), completes the proof of case (iii)(b).

This completes the proof of the lemma.

Henceforward we shall suppose that θ satisfies

$$0 < \theta \le 1/k\,, \tag{2.20}$$

and put

$$M = P^{\theta},\ H = PM^{-k},\ Q = PM^{-1}. \tag{2.21}$$

We shall suppose that P is large in terms of ε, and that R is at most a fairly small power of P. We shall elaborate on this latter condition in due course.

For $\underline{\underline{w}} \in (\mathbb{Z}/m^{k_1}\mathbb{Z}) \times \ldots \times (\mathbb{Z}/m^{k_t}\mathbb{Z})$, define $\mathcal{B}_r(\underline{\underline{w}}, m; \underline{\underline{k}})$ to be the set of solutions distinct (mod m^k) of the simultaneous congruences

$$z_1^{k_i} + \ldots + z_r^{k_i} \equiv w_i \pmod{m^{k_i}} \quad (1 \le i \le t), \tag{2.22}$$

with the z_i satisfying $(z_i, m) = 1$.

Also, define $T_s^*(P, R, \theta; r)$ to be the number of solutions to the simultaneous equations (2.3) satisfying (2.4), (2.5), and in addition

$$x_j \equiv y_j \pmod{m^k} \quad (1 \le j \le r), \tag{2.23}$$

with each solution $(\underline{\underline{x}}, \underline{\underline{y}}, \underline{\underline{u}}, \underline{\underline{v}})$ weighted by counting it with multiplicity $\mathrm{card}(\mathcal{B}_r(\underline{\underline{w}}, m; \underline{\underline{k}}))$, where $w_i \equiv x_1^{k_i} + \ldots + x_r^{k_i} \pmod{m^{k_i}}$ for $1 \le i \le t$.

Lemma 2.2. *We have*

$$T_s(P, R, \theta; r) \ll T_s^*(P, R, \theta; r).$$

Proof: For a given m, r and $\underline{\underline{k}}$, write $\mathcal{B}(\underline{\underline{w}})$ for $\mathcal{B}_r(\underline{\underline{w}}, m; \underline{\underline{k}})$. From (2.3),

$$x_1^{k_i} + \ldots + x_r^{k_i} \equiv y_1^{k_i} + \ldots + y_r^{k_i} \pmod{m^{k_i}} \text{ for all } 1 \le i \le t,$$

so that each solution of (2.3) can be classified according to the common residue class (mod m^{k_i}) of $x_1^{k_i} + \ldots + x_r^{k_i}$ and $y_1^{k_i} + \ldots + y_r^{k_i}$, for each i. Let

$$g_m(\alpha, \underline{\underline{z}}) = g_m(\alpha, z_1, \ldots, z_r)$$

$$= \sum_{\substack{x_1 \le P \\ x_1 \equiv z_1 \pmod{m^k}}} \cdots \sum_{\substack{x_r \le P \\ x_r \equiv z_r \pmod{m^k}}} e(\alpha_1 s_1(\underline{\underline{x}}) + \ldots + \alpha_t s_t(\underline{\underline{x}}))$$

where we write $s_i(\underline{\underline{x}})$ for $(x_1^{k_i} + \ldots + x_r^{k_i})$.

Then we have

$$T_s(P, R, \theta; r) \le \left[\sum_{P^\theta < m \le \mathrm{Min}\{P, P^\theta R\}} U_m \right] \tag{2.24}$$

where

$$U_m = \int_{\mathcal{U}_t^*} G_m(\underline{\alpha}) . |f(\underline{m}^{\underline{k}}\underline{\alpha} ; P^{1-\theta}, R)|^{2s-2r} \, d\underline{\alpha} \tag{2.25}$$

and

$$G_m(\underline{\alpha}) = \sum_{w_1=1}^{m^{k_1}} \cdots \sum_{w_t=1}^{m^{k_t}} \left| \sum_{\underline{z} \in \mathcal{B}(\underline{w})} g_m(\underline{\alpha}, \underline{z}) \right|^2 .$$

Hence, by Cauchy's inequality,

$$G_m(\underline{\alpha}) \leq \sum_{w_1=1}^{m^{k_1}} \cdots \sum_{w_t=1}^{m^{k_t}} \mathrm{card}(\mathcal{B}(\underline{w})) \sum_{\underline{z} \in \mathcal{B}(\underline{w})} |g_m(\underline{\alpha}, \underline{z})|^2 . \tag{2.26}$$

This, in view of (2.24) and (2.25), completes the proof of the lemma.

Now let

$$\mathcal{B}_r^* = \mathcal{B}_r^*(\underline{k}) = \max_{\underline{w}, m} \mathrm{card}(\mathcal{B}_r(\underline{w}, m; \underline{k})) , \tag{2.27}$$

where the maximum is taken over all $\underline{w} \in (\mathbb{Z}/m^{k_1}\mathbb{Z}) \times \cdots \times (\mathbb{Z}/m^{k_t}\mathbb{Z})$, and all $M < m \leq MR$.

Also, let $\tilde{T}_s(P, R, \theta; r)$ be the number of solutions to the simultaneous equations (2.3) satisfying (2.4), (2.5), and (2.23).

The following corollary to Lemma 2.2 is often a useful simplification.

Corollary 2.2.1. *We have*

$$T_s(P, R, \theta; r) \ll \mathcal{B}_r^* . \tilde{T}_s(P, R, \theta; r) .$$

Proof: Immediate from the lemma.

$\mathcal{B}_r(\underline{w}, m; \underline{k})$ will be unusually large only when $\underline{w}$ admits singular solutions to the congruences (2.22). But for such solutions to

exist, the variables z_i must lie in rather restricted sets, and this can sometimes be exploited when it comes to estimating exponential sums.

We now consider $\tilde{T}_s(P,R,\theta;r)$, the number of solutions of the equations (2.3) satisfying (2.4), (2.5) and (2.23). Put

$$z_j = x_j + y_j,\ h_j = (x_j - y_j)m^{-k}\ (1 \le j \le r).$$

Then for each j we have $2x_j = z_j + h_j m^k$ and $2y_j = z_j - h_j m^k$. But plainly

$$\tilde{T}_s(P,R,\theta;r) \ll \Upsilon_0 + \Upsilon_1 + \ldots + \Upsilon_r\ , \tag{2.28}$$

where $\Upsilon_n = \Upsilon_n(P,R,\theta;r,s)$ is the number of solutions of (2.3) subject to (2.4), (2.5), (2.23), and in addition

$$x_i = y_i \text{ for } n+1 \le i \le r.$$

Then

$$\Upsilon_n \le P^{r-n}.\Upsilon_n^*(P,R,\theta;r,s)\ , \tag{2.29}$$

where

$$\Upsilon_n^*(P,R,\theta;r,s) = \sum_{\eta_1} \ldots \sum_{\eta_n} \Upsilon_n^*(P,R,\theta;r,s;\eta_1,\ldots,\eta_n)\ ,$$

and the summation is over $\eta_1,\ldots,\eta_n$ with $\eta_j = \pm 1$, and where $\Upsilon_n^*(P,R,\theta;r,s;\eta_1,\ldots,\eta_n)$ is the number of solutions of the system of equations

$$\sum_{j=1}^{n} \eta_j \Phi_i(z_j,h_j,m) = 2^{k_i}\left(v_1^{k_i} + \ldots + v_{s-r}^{k_i} - u_1^{k_i} - \ldots - u_{s-r}^{k_i}\right)$$

$$(1 \le i \le t), \tag{2.30}$$

with $u_j, v_j \in \mathcal{A}(Q,R)$ $(1 \le j \le s-r)$, $1 \le z_j \le 2P$, $1 \le h_j \le H$, $M < m \le MR$, and where

$$\Phi_i(z,h,m) = m^{-k_i}((z + hm^k)^{k_i} - (z - hm^k)^{k_i})\ (1 \le i \le t). \tag{2.31}$$

We now introduce the exponential sum

$$F_n^{\sigma}(\underline{\alpha}) = \sum_{M<m\le MR} \prod_{i=1}^{n} \left(\sum_{h\le H} \sum_{0<z\le 2P} e(\eta_i \underline{\alpha\Phi}) \right)\ , \tag{2.32}$$

where we have written $\underline{\alpha\Phi}$ for $(\alpha_1\Phi_1(z,h,m) + \ldots + \alpha_t\Phi_t(z,h,m))$, and

where σ denotes $(\eta_1,\dots,\eta_n)$. Thus, by (2.8), (2.30) and (2.32), we have

$$\Upsilon_n^*(P,R,\theta;r,s) = \int_{\mathcal{U}_t^*} \sum_{\sigma} F_n^{\sigma}(\underline{\underline{\alpha}}).|f(\underline{\underline{2}}^k\underline{\underline{\alpha}};Q,R)|^{2s-2r}\,d\underline{\underline{\alpha}}\ ,$$

where the summation over σ denotes that we are to sum over all possible choices of $\sigma = (\pm1,\dots,\pm1)$. In particular, if we define

$$F_n(\underline{\underline{\alpha}}) = \sum_{M<m\le MR}\Big|\sum_{h\le H}\sum_{0<z\le 2P} e(\underline{\underline{\alpha}}\Phi)\Big|^n, \tag{2.33}$$

then we have

$$\Upsilon_n^*(P,R,\theta;r,s) \ll \int_{\mathcal{U}_t^*} F_n(\underline{\underline{\alpha}}).|f(\underline{\underline{2}}^k\underline{\underline{\alpha}};Q,R)|^{2s-2r}\,d\underline{\underline{\alpha}}\ . \tag{2.34}$$

3. ESTIMATING B_r^*.

In this section we estimate $\mathfrak{B}_2^*$ for the equations (1.2) by using elementary properties of congruences.

Consider the simultaneous congruences

$$\left.\begin{aligned} f = z_1^3 + z_2^3 &\equiv u_1 \\ g = z_1^2 + z_2^2 &\equiv u_2 \end{aligned}\right\} \pmod{p^s}\ . \tag{3.1}$$

If (z_1,z_2) is any solution of the congruences (3.1), we define

$$\Delta(f,g) = \begin{vmatrix} 3z_1^2 & 3z_2^2 \\ 2z_1 & 2z_2 \end{vmatrix} = 6z_1z_2(z_1 - z_2)\ . \tag{3.2}$$

We call a solution $\underline{\underline{z}}$ of (3.1) (mod p^s) singular (mod p^r) if $|\Delta(f,g)|_p = p^{-r}$. If $|\Delta(f,g)|_p = 1$, then we call the solution non-singular (mod p).

Lemma 3.1. *Suppose that p is a rational prime and $(u_1,u_2) \in (\mathbb{Z}/p^s\mathbb{Z})^2$. Let $\mathcal{C}(p,r)$ denote the number of distinct solutions (z_1,z_2) of the simultaneous congruences (3.1) with $(z_i,p) = 1$ for $i = 1,2$, and singular (mod p^r) ($r \le s$). Then we have*

$$\mathcal{C}(p,r) \ll Min\{p^{s-r}, p^r\},$$

where the implicit constant is independent of p.

Proof: Consider primes $p \neq 2,3$. Suppose that the congruences (3.1) have a solution (z_1,z_2) singular (mod p^r) for some $0 \le r \le s$, and with $(z_1,p) = 1$ for $i = 1,2$. Then by (3.2) we have $(z_1 - z_2, p^s) = p^r$.

If $r = 0$, then by eliminating z_2 and considering the single polynomial equation (mod p), we see that there are at most $O(1)$ non-singular solutions to the congruences (3.1) distinct (mod p). We shall consider all solutions (z_1',z_2') of (3.1) non-singular (mod p), with $z_i' \equiv z_i \pmod p$ ($i = 1,2$) for some fixed solution (z_1,z_2) of (3.1) non-singular (mod p), and with $(z_i,p) = 1$ for $i = 1,2$. Plainly, we can write $z_i' = z_i + k_i p^\tau$ with $0 < \tau \le s$, $(k_1,k_2,p) = 1$ and $0 < k_i \le p^{s-\tau}$ ($i = 1,2$).

If $r > 0$, then $z_1 \equiv z_2 \pmod{p^r}$. Considering (3.1) (mod p^r), we obtain

$$\left.\begin{aligned} 2z_1^3 &\equiv u_1 \\ 2z_1^2 &\equiv u_2 \end{aligned}\right\} \pmod{p^r},$$

and hence, as $(z_i,p) = 1$ for $i = 1,2$, we have $z_1 \equiv u_2^{-1}u_1 \pmod{p^r}$, whenever $2u_1^2 \equiv u_2^3 \pmod{p^r}$ and $(u_2,p) = 1$ (if the latter conditions do not hold, then the congruences are insoluble). Then all solutions of (3.1) singular (mod p^r) are mutually congruent (mod p^r). We shall consider all distinct solutions (z_1',z_2') of (3.1), with $z_i' \equiv z_i \pmod{p^r}$ ($i = 1,2$) for some fixed solution (z_1,z_2) of (3.1) singular (mod p^r), and with $(z_i,p) = 1$ for $i = 1,2$. Plainly, we can

write $z_i' = z_i + k_i p^\tau$ with $\tau \le s$, $(k_1, k_2, p) = 1$, $0 < k_i \le p^{s-\tau}$ $(i = 1,2)$. We may also assume that $\tau \ge r$, for the uniqueness of $\underline{z} \pmod{p^r}$ implies that $\underline{z}' \equiv \underline{z} \pmod{p^r}$.

In either case, we have

$$\left.\begin{aligned}(z_1 + k_1 p^\tau)^3 + (z_2 + k_2 p^\tau)^3 &\equiv u_1\\ (z_1 + k_1 p^\tau)^2 + (z_2 + k_2 p^\tau)^2 &\equiv u_2\end{aligned}\right\} \pmod{p^s},$$

so on substituting from (3.1),

$$\left.\begin{aligned}3(z_1^2 k_1 + z_2^2 k_2)p^\tau + 3(z_1 k_1^2 + z_2 k_2^2)p^{2\tau} + (k_1^3 + k_2^3)p^{3\tau} &\equiv 0\\ 2(z_1 k_1 + z_2 k_2)p^\tau + (k_1^2 + k_2^2)p^{2\tau} &\equiv 0\end{aligned}\right\} \pmod{p^s} \qquad (3.3)$$

We now divide into cases:

(i) Suppose that $r + \tau \ge s$.

As $(z_1, p) = 1$, from (3.3) we have $k_1 + k_2 \equiv 0 \pmod{p^{s-\tau}}$. So we have at most $\phi(p^{s-\tau})$ distinct choices for $(k_1, k_2) \pmod{p^{s-\tau}}$.

(ii) Suppose that $r + \tau < s$ and $\tau > r$.

From (3.3) we have

$$\left.\begin{aligned}3z_1^2 k_1 + 3z_2^2 k_2 &\equiv 0\\ 2z_1 k_1 + 2z_2 k_2 &\equiv 0\end{aligned}\right\} \pmod{p^{r+1}}$$

and by elimination,

$$6z_2 k_2 (z_1 - z_2) \equiv 0 \pmod{p^{r+1}}.$$

But $(z_1 - z_2, p^s) = p^r$ and $(z_2, p) = 1$, so $p \mid k_2$, and hence also $p \mid k_1$. This contradicts the assumption $(k_1, k_2, p) = 1$, so that this case can never occur.

(iii) Suppose that $r + \tau < s$ and $\tau = r$.

Then since $\tau > 0$, we have $r > 0$. Also, $z_1 \equiv z_2 \pmod{p^r}$, and from (3.3) we obtain $k_1 + k_2 \equiv 0 \pmod{p^r}$. So we can write $z = z_1$, $z_2 = z + \zeta p^r$, $k = k_1$ and $k_2 = -k + \kappa p^r$ for some integers κ and ζ. So from (3.3) we have

$$\left.\begin{aligned}3(-2\zeta z k + z^2 \kappa)p^{2r} + 6zk^2 p^{2r} &\equiv 0\\ 2(z\kappa - k\zeta)p^{2r} + 2k^2 p^{2r} &\equiv 0\end{aligned}\right\} \pmod{p^{2r+1}}$$

and hence, by elimination, $\kappa \equiv 0 \pmod{p}$, and $k(k - \zeta) \equiv 0 \pmod{p}$.

But we cannot have $k \equiv 0 \pmod{p}$, for this would contradict $(k_1, k_2, p) = 1$. Then $k \equiv \zeta \pmod{p}$, which implies that $z_1' \equiv z_2 \pmod{p^{r+1}}$ and $z_2' \equiv z_1 \pmod{p^{r+1}}$. But in this case we may relabel variables so that $z_i' \equiv z_i \pmod{p^{r+1}}$, and apply the same arguments as above, but with $\tau \geq r+1$.

Thus, for some constant C independent of p, we obtain at most

$$C \sum_{\substack{\tau=r \\ s-\tau \leq r}}^{s} \phi(p^{s-\tau}) = C \operatorname{Min}\{p^{s-r}, p^r\}$$

distinct solutions of the congruences (3.1) $\pmod{p^s}$ which are singular $\pmod{p^r}$.

It is not difficult to check that the result for $p = 2$ and 3 differs only in the constant, and this completes the proof of the lemma.

Lemma 3.2. *We have* $\mathcal{B}_2^*(3,2) \ll (MR)^{2+\varepsilon}$.

Proof: We wish to estimate the maximum number of solutions, distict $\pmod{m^3}$, to the simultaneous congruences

$$\left.\begin{aligned} z_1^3 + z_2^3 &\equiv u_1 \pmod{m^3} \\ z_1^2 + z_2^2 &\equiv u_2 \pmod{m^2} \end{aligned}\right\} \tag{3.4}$$

with

$$(z_i, m) = 1 \text{ for } i = 1,2, \tag{3.5}$$

as we allow u_1 and u_2 to vary over $\mathbb{Z}/m^3\mathbb{Z}$ and $\mathbb{Z}/m^2\mathbb{Z}$ respectively, and as we allow m to vary over $M < m \leq MR$.

Suppose that (z_1, z_2) is any solution of the simultaneous congruences

$$\left.\begin{aligned} z_1^3 + z_2^3 &\equiv u_1 \\ z_1^2 + z_2^2 &\equiv u_2 \end{aligned}\right\} \pmod{m^2} \tag{3.6}$$

satisfying (3.5). There are $O(m^{1+\varepsilon})$ solutions (z_1', z_2'), distinct $\pmod{m^3}$, of the congruences (3.4) satisfying (3.5) corresponding to (z_1, z_2). These are given by taking any of the m possible choices of

$z_1' \equiv z_1 \pmod{m^2}$, and then solving for z_2' from the congruence

$$z_2'^3 \equiv u_1 - z_1'^3 \pmod{m^3} .$$

The latter congruence has $O(m^\varepsilon)$ solutions distinct (mod m^3).

Thus

$$\mathcal{B}_2^*(3,2) \ll (MR)^{1+\varepsilon}.\mathcal{C}^* ,$$

where $\mathcal{C}^*$ is the maximum number of solutions satisfying (3.5), distinct (mod m^2), to the simultaneous congruences (3.6) as we allow $\underline{u}$ to vary over $(\mathbb{Z}/m^2\mathbb{Z})^2$, and as we allow m to vary over $M < m \le MR$.

Suppose that $p^s \| m^2$, and (z_1, z_2) is any solution of (3.6) with

$$(z_1 - z_2, p^s) = p^r \text{ and } (z_i, p) = 1 \text{ for } i = 1,2. \qquad (3.7)$$

Then by Lemma 3.1, the number of solutions to the simultaneous congruences

$$\left.\begin{aligned} z_1^3 + z_2^3 &\equiv u_1 \\ z_1^2 + z_2^2 &\equiv u_2 \end{aligned}\right\} \pmod{p^s}$$

satisfying (3.7) is $\ll \mathrm{Min}\{p^{s-r}, p^r\} \le p^{s/2}$.

Applying the Chinese Remainder Theorem, we deduce that there is an absolute constant C such that for each divisor d of m^2, there are at most

$$S(d) = \prod_{\substack{p \text{ prime} \\ p \mid m^2}} C.p^{s/2}$$

solutions to the congruences (3.6) satisfying the condition that for each prime $p \mid m$, with $p^r \| d$, the solution satisfies the condition (3.7). Denoting the number of prime divisors of n by $\omega(n)$, we have $\omega(m^2) \ll \log m/\log\log m$ (see, for example, Hardy and Wright [1979], §22.10), and hence

$$S(d) \ll C^{\log m/\log\log m}.m \ll m^{1+\varepsilon}.$$

Since the number of divisors of an integer n is $O(n^\varepsilon)$, we obtain $\mathcal{C}^* \ll (MR)^{1+2\varepsilon}$, and hence, in view of the above comments, $\mathcal{B}_2^* \ll (MR)^{2+3\varepsilon}$.

This completes the proof of the lemma.

4. BOUNDING THE NUMBER OF SOLUTIONS OF THE AUXILIARY EQUATIONS.

For the example in which we are interested, there would appear to be no satisfactory method of providing "minor arc" estimates for the exponential sum $F_n(\underline{\underline{\alpha}})$, and so we are forced to estimate the integral (2.34) in terms of the number of solutions of various diophantine equations. We begin with a number of preliminary lemmata.

Henceforward all definitions from §2 will assume the obvious restricted meanings appropriate for the case $t = 2$, $k_1 = 3$ and $k_2 = 2$.

Lemma 4.1. *Suppose that m and n are integers with $m > n \geq 1$. Then the number of solutions, S, of the simultaneous diophantine equations*

$$\left.\begin{aligned} x_1^m + x_2^m + x_3^m &= y_1^m + y_2^m + y_3^m \\ x_1^n + x_2^n + x_3^n &= y_1^n + y_2^n + y_3^n \end{aligned}\right\}$$

with $0 < x_i, y_i \leq P$, $(i = 1,2,3)$, satisfies

$$S \ll P^{3+\varepsilon}.$$

Proof: S is plainly the number of solutions of the simultaneous diophantine equations

$$\left.\begin{aligned} x_1^m + x_2^m - x_3^m &= y_1^m + y_2^m - y_3^m \\ x_1^n + x_2^n - x_3^n &= y_1^n + y_2^n - y_3^n \end{aligned}\right\} \tag{4.1}$$

with $0 < x_i, y_i \leq P$, $(i = 1,2,3)$.

Consider any solution $(\underline{\underline{x}}, \underline{\underline{y}})$ of (4.1). We obtain

$$(x_1^m + x_2^m - x_3^m)^n - (x_1^n + x_2^n - x_3^n)^m = (y_1^m + y_2^m - y_3^m)^n - (y_1^n + y_2^n - y_3^n)^m \tag{4.2}$$

and on factorising each side, for some polynomial $Q(t_1, t_2, t_3)$ homogeneous in t_1, t_2, t_3, of degree $nm-2$, and with integer coefficients, we have

$$(x_1-x_3)(x_2-x_3)Q(x_1, x_2, x_3) = (y_1-y_3)(y_2-y_3)Q(y_1, y_2, y_3) .$$

(i) Suppose first that the right hand side of (4.2) is zero.

The number of solutions of the equation

$$(y_1^m + y_2^m - y_3^m)^n - (y_1^n + y_2^n - y_3^n)^m = 0 \qquad (4.3)$$

with $0 < y_1 \leq P$, is at most $O(P^2)$. This is because $m > n$, so that by picking any one of the possible P^2 choices for y_1 and y_2, the left hand side of (4.3) is a non-trivial polynomial in y_3 (note, for example, that the binomial expansion of $(A-y_3^m)^n$ has fewer terms than that of $(B-y_3^n)^m$). Hence (4.3) determines y_3 up to a multiplicity of at most $O(1)$.

Consider any one of these solutions (y_1, y_2, y_3), and take any of the P possible choices for x_3 in (4.1). Write

$$M = y_1^m + y_2^m - y_3^m + x_3^m \text{ , and } N = y_1^n + y_2^n - y_3^n + x_3^n \text{ .}$$

Then from (4.1), we have

$$\left.\begin{aligned} x_1^m + x_2^m &= M \\ x_1^n + x_2^n &= N \end{aligned}\right\} \qquad (4.4)$$

Then $(M - x_1^m)^n = (N - x_1^n)^m$, and as $m > n$, this again determines a non-trivial polynomial in x_1. Thus we have determined x_1 up to a multiplicity of at most $O(1)$, and hence also x_2, from (4.4), up to a multiplicity of at most $O(1)$.

Then the number of solutions of (4.1) with the right hand side of (4.2) equal to zero is at most $O(P^2.P) \ll P^3$.

(ii) Suppose now that the right hand side of (4.2) is non-zero.

Then by picking any one of the $O(P^3)$ possible choices for (y_1, y_2, y_3), we have for some non-zero integer K,

$$(x_1-x_3)(x_2-x_3)Q(x_1, x_2, x_3) = K \text{ .}$$

Using the divisor function, we have at most $(d(K))^2 \ll P^{\varepsilon}$ possible solutions of this last equation for x_1-x_3 and x_2-x_3, say

$$x_1 = x_3 + d_1 \text{ and } x_2 = x_3 + d_2 \text{ .}$$

Substituting into (4.1), as y_1, y_2, y_3 have been fixed, we obtain for some non-trivial polynomial f in x_3, $f(x_3) = 0$. Hence there are at most $O(1)$ possible solutions for x_3.

Then the total number of possible solutions to (4.1) in this case is $O(P^3.P^{\varepsilon})$.

This completes the proof of the lemma.

Theorem 1.3 is plainly equivalent to Lemma 4.1. We have the following corollary.

Corollary 4.1.1. *For any $R \le P$, we have $S_3(P,R) \ll P^{3+\varepsilon}$.*

Proof: Plainly $S_3(P,R) \le S_3(P,P)$, and the latter is $\ll P^{3+\varepsilon}$ by the lemma.

Lemma 4.2 (*see Hua* [1965], *Theorem* 4). *Let $\psi(x)$ denote a polynomial of degree k with integer coefficients, and let*

$$T(\alpha) = \sum_{x=1}^{P} e(\psi(x)\alpha) \ .$$

Then when $1 \le j \le k$, we have

$$\int_0^1 |T(\alpha)|^{2^j} d\alpha \ll \Delta^{\varepsilon}.P^{2^j-j+\varepsilon},$$

where Δ is the greatest common divisor of the coefficients of $\psi(x)$, and the implicit constant depends only on k and j.

For integers m and y satisfying $M < m \le MR$ and $1 \le y \le 4P$ respectively, define

$$\psi(z;m,y) = 4mz^3 - 3yz^2,$$

and let $N_t(Q;m,y)$ denote the number of solutions of the diophantine equation

$$\psi(u_1;m,y) + \ldots + \psi(u_t;m,y) = \psi(v_1;m,y) + \ldots + \psi(v_t;m,y)$$

with $1 \le u_i, v_i \le Q$. Also, define

$$N_t(Q) = \operatorname*{Max}_{m,y} N_t(Q;m,y) .$$

Lemma 4.3. *We have*

(i) $\Upsilon_2^*(P,R,\theta;2,s) \ll PMRH^2.S_{s-2}(Q,R) + P^{1+\epsilon}MRH.N_{s-2}(Q)$,

(ii) $\Upsilon_1^*(P,R,\theta;2,s) \ll P^{1/2}MRH.S_{s-2}(Q,R)$
$$+ P^{\epsilon}(PH)^{1/2}MR(N_{s-2}(Q).S_{s-2}(Q,R))^{1/2},$$

(iii) $\Upsilon_0^*(P,R,\theta;2,s) \ll MR.S_{s-2}(Q,R)$.

Proof: First note that by (2.8), and considering the underlying diophantine equations, we have

$$\int_{\mathfrak{U}_2^*} |f(\underline{\underline{2}}^k\underline{\underline{\alpha}};Q,R)|^{2t}\, d\underline{\underline{\alpha}} \ll S_t(Q,R) \tag{4.5}$$

for each integer t.

(i) By Cauchy's inequality, we have from (2.33) that

$$F_2(\underline{\underline{\alpha}}) \le H \sum_{M<m\le MR} \sum_{h\le H} \Big| \sum_{0<z\le 2P} e(\underline{\underline{\alpha\Phi}}) \Big|^2,$$

and hence, from (2.34) and by considering the underlying diophantine equations, we deduce that

$$\Upsilon_2^*(P,R,\theta;2,s) \ll H.V , \tag{4.6}$$

where V is the number of solutions to the simultaneous diophantine equations

$$\left.\begin{aligned} 4((u_1^3 - v_1^3) + \ldots + (u_{s-2}^3 - v_{s-2}^3)) &= 3h(z_1^2 - z_2^2) \\ (u_1^2 - v_1^2) + \ldots + (u_{s-2}^2 - v_{s-2}^2) &= hm(z_1 - z_2) \end{aligned}\right\} \tag{4.7}$$

with

$$1 \le z_1, z_2 \le 2P \ , \ 1 \le h \le H \ , \ M < m \le MR \ ,$$
$$\text{and } u_i, v_i \in \mathcal{A}(Q,R) \text{ for } i = 1,\ldots,s-2.$$

The number of solutions of the equations (4.7) with $z_1 = z_2$ is

$$\ll PMRH.S_{s-2}(Q,R) . \tag{4.8}$$

Suppose then that $z_1 \neq z_2$, and write

$$x = z_1 - z_2 \text{ and } y = z_1 + z_2 .$$

By elimination, we deduce from (4.7) that the u_i and v_i satisfy the equation

$$\psi(u_1;m,y) + \ldots + \psi(u_{s-2};m,y) = \psi(v_1;m,y) + \ldots + \psi(v_{s-2};m,y)$$

Then given m and y, there are at most $N_{s-2}(Q;m,y)$ solutions in the u_i and v_i to the equations (4.7). But given $\underline{u}$, $\underline{v}$, y and m, and by using estimates for the divisor function, there are $O(P^{\varepsilon})$ solutions in h and x to the equations (4.7). We may then determine z_1 and z_2 directly from x and y. Thus the total number of solutions to the equations (4.7) with $x \neq 0$ is

$$\ll P^{\varepsilon} \sum_{m,y} N_{s-2}(Q;m,y) \ll P^{1+\varepsilon} MR.N_{s-2}(Q) . \tag{4.9}$$

On collecting together (4.6), (4.7), (4.8) and (4.9), we reach the desired conclusion.

(ii) By the Cauchy-Schwarz inequalities, we deduce from (2.34) that

$$\Upsilon_1^*(P,R,\theta;2,s) \ll \left(S_{s-2}(Q,R)\right)^{1/2}\left(\int_{\mathcal{U}_2^*} |F_1(\underline{\alpha})|^2 \, |f(\underline{2}^k\underline{\alpha};Q,R)|^{2s-4} \, d\underline{\alpha}\right)^{1/2}$$

$$\ll \left(S_{s-2}(Q,R)\right)^{1/2}\left(MR \int_{\mathcal{U}_2^*} F_2(\underline{\alpha}).|f(\underline{2}^k\underline{\alpha};Q,R)|^{2s-4} \, d\underline{\alpha}\right)^{1/2}$$

by (2.33). Thus, by (2.34) and part (i) of this lemma, we have

$$\Upsilon_1^*(P,R,\theta;2,s) \ll (S_{s-2}(Q,R))^{1/2}.(PM^2R^2H^2.S_{s-2}(Q,R) + P^{1+\varepsilon}M^2R^2H.N_{s-2}(Q))^{1/2},$$

and the result follows.

(iii) Here we obtain directly from (2.33) and (2.34) the bound

$$\Upsilon_0^*(P,R,\theta;2,s) \ll MR.S_{s-2}(Q,R) .$$

This completes the proof of the lemma.

We now draw together the conclusions of our analysis thus far to prove Theorem 1.2.

First note that by Lemma 4.2 and Schwarz's inequality, we have

$$N_3(Q) \ll Q^{7/2+\varepsilon} .$$

Thus, by (2.21), Corollary 4.1.1 and Lemma 4.3, we have

$$\Upsilon_2^*(P,R,1/6;2,5) \ll P^{14/3+\varepsilon}R + P^{55/12+\varepsilon}R ,$$

$$\Upsilon_1^*(P,R,1/6;2,5) \ll P^{11/3+\varepsilon}R + P^{29/8+\varepsilon}R ,$$

$$\Upsilon_0^*(P,R,1/6;2,5) \ll P^{8/3+\varepsilon}R .$$

Then by (2.28), (2.29), Corollary 2.2.1 and Lemma 3.2, we obtain

$$T_5(P,R,1/6;2) \ll (MR)^{2+\varepsilon}.P^{14/3+\varepsilon}R \ll P^{5+2\varepsilon}R^3 . \qquad (4.10)$$

Notice that by using trivial estimates we have $P^5 \ll S_5(P,R) \ll P^{10}$. Then there is some real number $\mu = \mu(R)$, with $5 < \mu \le 10$, such that for $2 \le R \le P$ we have

$$S_5(P,R) \ll P^{\mu} .$$

We now apply Lemma 2.1 with $D = P^{\theta}$. We obtain

$$S_5(P,R) \ll \Big(\sum_{d > P^{\theta}} (P/d)^{\mu/5} \Big)^5 + P^{10(2\theta-\theta^2)}$$

$$+ P^{\varepsilon}\Big(\sum_{d \le P^{\theta}} \Big[(P/d)^{\theta}R\Big].T_5(P/d,R,\theta;2)^{1/5}\Big)^5 .$$

With $\theta = 1/6$, the second term on the right hand side of the above expression is $o(P^5)$. The first term on the right hand side is $O(P^{5\mu/6})$. This leaves the last term on the right hand side, which by (4.10) is

$$\ll R^8P^{35/6+3\varepsilon} .$$

Then

$$S_5(P,R) \ll R^8P^{35/6+3\varepsilon}$$

for each $2 \le R \le P$, and for any $\varepsilon > 0$. Then by taking $R = P^{\eta}$ with $\eta \le \varepsilon/8$, we have

$$S_5(P,P^{\eta}) \ll P^{35/6+4\varepsilon}.$$

This completes the proof of Theorem 1.2.

CHAPTER 4.

PAIRS OF ADDITIVE EQUATIONS, ONE QUADRATIC AND ONE CUBIC.

1. A DISCUSSION OF THE CONDITIONS OF THEOREM 1.1 OF CHAPTER 3.

In this section we shall show that conditions (a), (b) and (c) of Theorem 3.1.1 are in some sense best possible, in that if any one of them is removed, then the equations (3.1.2) may fail to have a non-trivial rational solution.

(a)(i) the quadratic equation in (3.1.2) is indefinite.

Consider the equations

$$\left.\begin{aligned} c_1x_1^3+\ldots+c_{s-6}x_{s-6}^3+(y_1^3+cy_2^3)+p(y_3^3+cy_4^3)+p^2(y_5^3+cy_6^3) &= 0\\ d_1x_1^2+\ldots+d_{s-6}x_{s-6}^2 &= 0 \end{aligned}\right\} \qquad (1.1)$$

where the c_i are non-zero, p is a rational prime with $p \equiv 1 \pmod 3$, c is a cubic non-residue (mod p), and the d_i are either all strictly positive, or all strictly negative. Notice that although the quadratic here is definite, it does have a "sign change". On taking $s \geq 14$, it is easily verified that conditions (b), (c) and (d) of Theorem 3.1.1 are satisfied.

Suppose that $(\underline{x}, \underline{y})$ is any rational solution of the equations (1.1). Since the quadratic equation is definite, we must have $x_1 = \ldots = x_{s-6} = 0$. But since c is a cubic non-residue (mod p), on substituting $x_1 = \ldots = x_{s-6} = 0$ into the cubic equation, we find that the cubic equation has no non-trivial solution over $\mathbb{Z}_p$, and hence we must have $y_1 = \ldots = y_6 = 0$. Then the system (1.1) has only the trivial solution.

(ii) the quadratic equation in (3.1.2) must have at least 5 variables explicit.

This is a local solubility consideration. Consider, for example, equations in 10 variables of the form considered in the proof of Lemma 7.2 of Chapter 1. Such equations have no non-trivial solutions.

(b) the cubic equation in (3.1.2) must have at least 7 variables explicit.

Consider the system of equations

$$\left.\begin{aligned} (x_1^3 + 2x_2^3) + 7(x_3^3 + 2x_4^3) + 7^2(x_5^3 + 2x_6^3) &= 0 \\ -14x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 + x_7^2 + \ldots + x_s^2 &= 0 \end{aligned}\right\} \qquad (1.2)$$

This system satisfies condition (c) of Theorem 3.1.1, for we may put $x_1 = \ldots = x_5 = 1$, and then solve the cubic equation for x_6 over $\mathbb{R}$. The cubic equation will give a solution with $0 < |x_6| < 1$, and we may then put $x_7 = \ldots = x_s = \alpha$, and solve the quadratic equation for α. This will give a non-trivial real solution for the system of equations over $\mathbb{R}$. Notice that conditions (a) and (d) of Theorem 3.1.1 are satisfied on taking $s \geq 14$.

We now attempt to solve the system over $\mathbb{Z}$. Since 2 is a cubic non-residue (mod 7), the cubic equation has no non-trivial solution over $\mathbb{Z}$, as in the discussion of condition (a). Then we are forced to put $x_1 = \ldots = x_6 = 0$, and since the remainder of the quadratic is positive definite, the remaining variables must also be zero. Thus the system of equations (1.2) has no non-trivial rational solution.

(c) the simultaneous equations (3.1.2) have a non-trivial real solution.

This is plainly a necessary condition for a non-trivial rational solution to exist. However, it is perhaps worth noting that the quadratic equation may be indefinite, and yet a non-trivial real solution to the system still fails to exist. To see this, consider the system of equations

$$\left.\begin{aligned} -M^3x_1^3 + x_2^3 + x_3^3 + \ldots + x_s^3 &= 0 \\ -x_1^2 + M^2x_2^2 + M^2x_3^2 + \ldots + M^2x_s^2 &= 0 \end{aligned}\right\}$$

where $M > s^{1/4}$ is a positive integer. On eliminating x_1, we obtain the equation

$$M^4(x_2^2 + \ldots + x_s^2) - (x_2^3 + \ldots + x_s^3)^{2/3} = 0 ,$$

which has no real solution, since

$$\begin{aligned} (x_2^3 + \ldots + x_s^3)^{2/3} &\leq s.(\text{Max } |x_i|)^2 \\ &< M^4(\text{Max } |x_i|)^2 \\ &\leq M^4(x_2^2 + \ldots + x_s^2) . \end{aligned}$$

2. PRELIMINARIES TO AN APPLICATION OF THE HARDY-LITTLEWOOD METHOD.

We start by showing that the conditions of Theorem 3.1.1 allow us to assume that the equations (3.1.2) have a non-singular real solution $(\eta_1,\ldots,\eta_s)$ for which no η_i is zero, and also a non-singular p-adic solution for every rational prime p. We then dispose of the cases (d)(ii) and (d)(iii) of Theorem 3.1.1. Finally, having made these simplifications, we set up the apparatus required for our application of the Hardy-Littlewood method.

Lemma 2.1. *Suppose that conditions* (a), (b) *and* (c) *of Theorem* 3.1.1 *hold for the equations* (3.1.2). *Then one of the following holds*:

(*i*) *the equations* (3.1.2) *have a non-trivial rational solution, or*

(*ii*) *the equations* (3.1.2) *have a real solution* $(\eta_1,\ldots,\eta_s)$ *for which there is an* i *such that* c_i, d_i *and* η_i *are all non-zero.*

Proof: By a rearrangement of variables and change in notation, we may write the equations (3.1.2) in the form

$$\left.\begin{aligned} a_1x_1^3+\ \ldots\ +a_ux_u^3+\ b_1y_1^3+\ \ldots\ +b_vy_v^3 &= 0\\ c_1y_1^2+\ \ldots\ +c_vy_v^2+\ d_1z_1^2+\ \ldots\ +d_wz_w^2 &= 0 \end{aligned}\right\} \qquad (2.1)$$

where none of the a_i, b_i, c_i or d_i are zero, and by conditions (a) and (b) of Theorem 3.1.1, we have $u + v \geq 7$ and $v + w \geq 5$. There are four cases:

(a) Suppose that $v = 0$.

Since $\mathrm{Max}\{G^*(2),\Gamma^*(2)\} = 5$ (classical) and $\mathrm{Max}\{G^*(3),\Gamma^*(3)\} = 7$ (respectively R.Baker [1990] and Lewis [1957]), both the quadratic and cubic equations in (2.1) are independently soluble non-trivially in rational integers, and so the equations (2.1) certainly have a non-trivial rational solution.

(b) Suppose that $v > 0$ and $u > 0$.

By condition (a) of Theorem 3.1.1, the quadratic in (2.1) is indefinite in at least five variables, and so we may solve the quadratic equation over $\mathbb{R}$ with at least one of the y_i non-zero, say y_1. Substituting these values of y_i and z_i into the cubic equation, we may then solve the cubic equation for the x_i over $\mathbb{R}$. This gives us a non-trivial real solution to the system (2.1) with b_1, c_1 and y_1 each non-zero.

(c) Suppose that $v > 0$, $u = 0$ and $w > 0$.

By condition (c) of Theorem 3.1.1, there is a non-trivial real solution to the equations (2.1). Either a y_i is non-zero, in which case we are done, or else there is a change of sign amongst the d_i, say $d_1 < 0$ and $d_2 > 0$. But then we may solve the cubic equation over $\mathbb{R}$ with at least one of the y_i non-zero, say y_1. Substitute these y_i into the quadratic, and solve over $\mathbb{R}$ for the z_i, using the fact that this quadratic is indefinite. Thus we have a non-trivial real solution to the equations (2.1) with b_1, c_1 and y_1 each non-zero.

(d) Suppose that $v > 0$ and $u = w = 0$.

By condition (c) of Theorem 3.1.1, there is a non-trivial real solution to the equations (2.1). Then the equations are soluble with at least one of the y_i non-zero, and we are done.

This completes the proof of the lemma.

Lemma 2.2. *Suppose that there is a real solution* $\underline{\underline{x}} = (\eta_1, \ldots, \eta_s)$ *to the equations* (3.1.2) *for which there is an* i *such that* c_i, d_i *and* η_i *are all non-zero. Then one of the following holds:*

(*i*) *the equations* (3.1.2) *have a non-trivial rational solution, or*

(*ii*) *locally, there is an* $(s-2)$*-dimensional subspace* $\mathcal{S}$ *of positive* $(s-2)$*-volume in the neighbourhood of* $\underline{\underline{\eta}}$ *on which* $F = G = 0$. *In particular, there is a real solution* $\underline{\underline{\eta'}} = (\eta_1', \ldots, \eta_s')$ *to the equations* (3.1.2) *for which no* η_i' *is zero.*

Proof: If for some i both c_i and d_i are zero, then the equations (3.1.2) plainly have a non-trivial rational solution. Therefore we may assume that for each i at least one of c_i and d_i is non-zero. By a rearrangement of variables, we may assume that there is a real solution $(\eta_1, \ldots, \eta_s)$ for which c_1, d_1 and η_1 are all non-zero.

Suppose that we have a non-trivial real solution $(\eta_1,\ldots,\eta_s)$ for which precisely v of the η_i are non-zero (with $s \geq v > 0$). Notice that as c_1, d_1 and η_1 are all non-zero, we may assume that $v \geq 2$, and by a rearrangement of the variables $x_2,\ldots,x_s$, that $\eta_i = 0$ for $i = v+1,\ldots,s$.

Consider the expression

$$\Delta = \Delta(i,j) = \begin{vmatrix} 3c_i\eta_i^2 & 3c_j\eta_j^2 \\ 2d_i\eta_i & 2d_j\eta_j \end{vmatrix} = 6\eta_i\eta_j(c_id_j\eta_i - c_jd_i\eta_j) \ . \qquad (2.2)$$

There are two cases:

(a) there exist $i,j \in \{1,\ldots,v\}$ with $\Delta(i,j) \neq 0$.

By rearranging variables, we may suppose that $i = 1$ and $j = 2$. Then by the Implicit Function Theorem (see, for example, Apostol [1957], Theorem 7-6), there is an $(s-2)$-dimensional neighbourhood, T_0, of $(\eta_3,\ldots,\eta_s)$ and a function

$$(\eta_1',\eta_2') : T_0 \longrightarrow \mathbb{R}^2,$$

such that

$$\left.\begin{aligned} F(\eta_1'(\eta_3',\ldots,\eta_s'),\eta_2'(\eta_3',\ldots,\eta_s'),\eta_3',\ldots,\eta_s') &= 0 \\ G(\eta_1'(\eta_3',\ldots,\eta_s'),\eta_2'(\eta_3',\ldots,\eta_s'),\eta_3',\ldots,\eta_s') &= 0 \end{aligned}\right\}$$

for all $(\eta_3',\ldots,\eta_s') \in T_0$. Then locally, there is an $(s-2)$-dimensional subspace $\mathcal{S}$ of positive $(s-2)$-volume in the neighbourhood of $\underline{\underline{\eta}}$ on which $F = G = 0$. Further, by choosing η_i' with $|\eta_i'-\eta_i|$ sufficiently small for $i = 3,\ldots,s$, we can find η_1' and η_2' with $F(\underline{\underline{\eta}}') = G(\underline{\underline{\eta}}') = 0$. Since η_1 and η_2 are non-zero, by using continuity we can also ensure that η_1' and η_2' are non-zero, and hence that η_i' is non-zero for $i = 1,\ldots,s$. So we conclude that case (ii) of the lemma applies.

(b) $\Delta(i,j) = 0$ for all $i,j \in \{1,\ldots,v\}$.

Then by (2.2),

$$c_id_j\eta_i - c_jd_i\eta_j = 0 \text{ for all } i,j \in \{1,\ldots,v\} \ . \qquad (2.3)$$

There are two cases:

(1) both c_i and d_i are non-zero for every $i \in \{1,\ldots,v\}$.

Then by (2.3), we have

$$\eta_j = \begin{cases} \dfrac{c_1 d_j}{d_1 c_j}\,\eta_1 & \text{for } j = 2,\ldots,v, \\ 0 & \text{for } j = v+1,\ldots,s. \end{cases}$$

Hence we may put $\eta_1 = 1$, and we have a non-trivial rational solution $(\eta_1,\ldots,\eta_v,0,..,0)$ to (3.1.2). Then we are in part (i) of the lemma.

(2) there is an $i \in \{1,\ldots,v\}$ for which either $c_i = 0$ or $d_i = 0$.

As $c_1 d_1 \eta_1 \eta_i \neq 0$, and by (2.3), $c_i d_1 \eta_i = c_1 d_i \eta_1$, we obtain $c_i = d_i = 0$, which contradicts the supposition made at the beginning of the proof of the lemma.

This completes the proof of the lemma.

Lemma 2.3. *Suppose that* $s \geq 11$, *and the simultaneous equations* (3.1.2) *satisfy conditions* (a), (b) *and* (c) *of Theorem* 3.1.1. *Then the equations* (3.1.2) *have a non-trivial solution in rational integers if either of the following conditions hold*:

(a) *at least* 6 *of the* d_i *are zero,*

(b) *there is a rearrangement of variables such that* $c_1,\ldots,c_4$ *are zero, and* $d_1,\ldots,d_4$ *are not all of the same sign.*

Proof: (a) If 6 or more of the d_i are zero, by a rearrangement of variables we may suppose that $d_1 = \ldots = d_6 = 0$. By condition (a) of Theorem 3.1.1, $d_7,\ldots,d_s$ cannot all be of the same sign. But $s \geq 11$ and $\mathrm{Max}\{\Gamma^*(2), G^*(2)\} = 5$, so there is a non-trivial integer solution $(a_7,\ldots,a_s)$ to the equation

$$d_7 x_7^2 + \ldots + d_s x_s^2 = 0 .$$

Let $C_i = c_i$ $(i = 1,\ldots 6)$, and

$$C_7 = c_7 a_7^3 + \ldots + c_s a_s^3 .$$

Then since $\mathrm{Max}\{\Gamma^*(3), G^*(3)\} = 7$, there is a non-trivial integer solution $(u_1,\dots,u_7)$ to the equation

$$C_1u_1^3 + \dots + C_7u_7^3 = 0 .$$

This gives a non-trivial integer solution

$$(u_1,\dots,u_6,u_7a_7,u_7a_8,\dots,u_7a_s)$$

to the system (3.1.2).

(b) If $c_1,\dots,c_4$ are zero, and $d_1,\dots,d_4$ are not all of the same sign, a similar argument to that in (a) demonstrates that there is a non-trivial solution to the system (3.1.2).

This completes the proof of the lemma.

Lemma 2.4. *Let $c_1,\dots,c_t$ $(t \ge 7)$ be rational integers, and*

$$f(\underline{x}) = c_1x_1^3 + \dots + c_tx_t^3 . \tag{2.4}$$

Suppose that $(\eta_1,\dots,\eta_t)$ is a real solution of the equation $f(\underline{x}) = 0$ with $0 < \eta_i < 1$ for $i = 1,\dots,t$. Then given any $0 < \alpha < 1$ and P sufficiently large, there is a non-trivial solution in integers $(y_1,\dots,y_t)$ to the equation $f(\underline{x}) = 0$, and satisfying

$$(1-\alpha)\eta_iP < y_i \le (1+\alpha)\eta_iP \quad (i = 1,\dots,t).$$

Proof: We modify the argument of Vaughan [1989b] so as to deal with additive cubic equations.

Suppose that ε is a sufficiently small positive number, that $\eta = \eta(\varepsilon)$ is a small positive number depending at most on ε, and that P is sufficiently large in terms of ε, η, α and $c_1,\dots,c_t$. Let $\theta = 1/8$, $\tau = 10^{-10}$, $R = P^{\eta}$, and let M_s $(s = 1,\dots,5)$ be real numbers satisfying

$$P^{\theta} \le M_s \le P^{\theta+\tau}.$$

We consider the number $R(M_1,\dots,M_5)$ of solutions of the equation

$$c_1p_1^3y_1^3 + \dots + c_5p_5^3y_5^3 + c_6x_6^3 + c_7x_7^3 + \dots + c_tx_t^3 = 0 ,$$

with the p_s prime, and

$$\left.\begin{aligned} &(1-\alpha/3)\eta_r P < x_r \le (1+\alpha/3)\eta_r P \quad (r = 6,\dots,t), \\ &p_s \equiv 2 \pmod 3,\ M_s < p_s \le (1+\alpha/3)M_s,\ y_s \in \mathcal{A}(P/M_s, R), \\ &(1-\alpha/3)\eta_s P/M_s < y_s \le (1+\alpha/3)\eta_s P/M_s \quad (s = 1,\dots,5). \end{aligned}\right\} \tag{2.5}$$

Then, with modifications involving only adjustments of implicit constants, and by making the trivial bound

$$\Big|\sum_{x\le P} e(\beta x^3)\Big| \ll P\ ,$$

the argument of Vaughan [1989b] yields

$$R(M_1,\dots,M_5) \gg P^{t-3}(\log P)^{-5} \longrightarrow \infty \text{ as } P \longrightarrow \infty.$$

Now $p_s > R$, so each solution obtained in this way is unique. Then by (2.5) we have a non-trivial integral solution of (2.4) satisfying

$$(1-\alpha/3)\eta_r P < x_r \le (1+\alpha/3)\eta_r P \text{ for } r = 6,\dots,t, \text{ and}$$

$$(1-\alpha/3)\eta_s P < p_s y_s \le (1+\alpha/3)^2\eta_s P \text{ for } s = 1,\dots,5\ .$$

This completes the proof of the lemma.

Lemma 2.5. *Suppose that the simultaneous equations* (3.1.2) *satisfy conditions* (a), (b) *and* (c) *of Theorem* 3.1.1. *Then the equations* (3.1.2) *have a non-trivial solution in rational integers if at least four of the* c_i *are zero.*

Proof: Suppose that the equations (3.1.2) satisfy conditions (a), (b) and (c) of Theorem 3.1.1. Then by Lemmata 2.1 and 2.2, we may assume that the equations have a non-singular real solution $(\eta_1,\dots,\eta_s)$ for which no η_i is zero. Further, we may assume that $0 < \eta_i < 1$ for $i = 1,\dots,s$, since whenever necessary the c_i can be replaced by $-c_i$ and η_i^3 by $-\eta_i^3$, and by homogeneity η_i can be replaced by $\eta_i/(\text{Max } \eta_j)$.

We may plainly assume that for each i at least one of c_i and d_i is non-zero. Suppose that at least 4 of the c_i are zero. Rearrange

variables in (3.1.2) so that $c_1,\dots,c_t$ are non-zero, and so that $c_{t+1},\dots,c_s$ are zero.

By Lemma 2.3(b), we may assume that $d_{t+1},\dots,d_s$ are all positive. Then

$$(d_1\eta_1^2 + \dots + d_t\eta_t^2) = -(d_{t+1}\eta_{t+1}^2 + \dots + d_s\eta_s^2) < 0 . \qquad (2.6)$$

But by condition (b) of Theorem 3.1.1 and Lemma 2.4, given any $0 < \alpha < 1$, and P sufficiently large, there is an integral solution $(y_1,\dots,y_s)$ to the cubic equation in (3.1.2) satisfying

$$(1-\alpha)\eta_i P < y_i \le (1+\alpha)\eta_i P \quad (i = 1,\dots,t),$$

and with $y_{t+1},\dots,y_s$ free variables, since $c_{t+1},\dots,c_s$ are zero. Then by (2.6), on taking α sufficiently small, we deduce that there is an integral solution $\underline{y}$ to the cubic equation in (3.1.2) satisfying

$$D = d_1y_1^2 + \dots + d_ty_t^2 < 0 .$$

Then by the classical result $\mathrm{Max}\{G^*(2),\Gamma^*(2)\} \le 5$, there exists a non-trivial integral solution $(v_t,\dots,v_s)$ to the equation

$$Du_t^2 + d_{t+1}u_{t+1}^2 + \dots + d_su_s^2 = 0 .$$

But then $(y_1v_t, y_2v_t, \dots, y_tv_t, v_{t+1}, v_{t+2}, \dots, v_s)$ is a non-trivial integral solution to the system (3.1.2).

This completes the proof of the lemma.

We now show that we may assume that there is a non-singular p-adic solution to the equations (3.1.2). Let $M_n(q)$ denote the number of solutions to the simultaneous congruences

$$\left.\begin{aligned} c_1m_1^3 + \dots + c_sm_s^3 &\equiv 0 \\ d_1m_1^2 + \dots + d_sm_s^2 &\equiv 0 \end{aligned}\right\} \pmod{q} . \qquad (2.7)$$

Lemma 2.6. *Suppose that the equations* (3.1.2) *satisfy conditions* (*a*), (*b*) *and* (*c*) *of Theorem* 3.1.1, *and* $s \geq 11$. *Then one of the following holds*:

(*i*) *there is a non-trivial rational solution to the equations* (3.1.2), *or*

(*ii*) *for each rational prime* p *there is a number* $u = u(p) < \infty$ *such that for all* $t \geq u$,

$$M_n(p^t) \geq p^{(t-u)(s-2)}.$$

Proof: As in previous lemmata, we may assume that for each i, at least one of c_i and d_i is non-zero.

By Theorem 1.1 of Chapter 1 we have $\Gamma^*(3,2) = 11$, and hence there is a non-trivial p-adic solution to the system $F = G = 0$. For any p-adic solution, $\underline{a} = (a_1, \ldots, a_s)$, denote the number of non-zero a_i by $v = v(\underline{a})$. Then we may clearly choose $\underline{a}$ so that $v(\underline{a})$ is maximal amongst all the p-adic solutions to the system $F = G = 0$. We rearrange variables so that $a_i = 0$ for $i = v+1, \ldots, s$, and consider the expression

$$\Delta = \Delta(i,j) = 6a_ia_j(c_id_ja_i - c_jd_ia_j).$$

There are 2 cases:

(a) there exist $i, j \in \{1, 2, \ldots, v\}$ with $\Delta(i,j) \neq 0$.

Rearrange variables so that $i = 1$, $j = 2$, and write

$$A(y_1, y_2) = c_1y_1^3 + c_2y_2^3, \text{ and } B(y_1, y_2) = d_1y_1^2 + d_2y_2^2,$$

and for fixed $y_3, \ldots, y_s$, put

$$C(y_3, \ldots, y_s) = c_3y_3^3 + \ldots + c_sy_s^3,$$

and

$$D(y_3, \ldots, y_s) = d_3y_3^2 + \ldots + d_sy_s^2.$$

Suppose that $|\Delta(1,2)|_p^2 = p^{1-u} > 0$. Choose y_i $(i = 3, \ldots, s)$ with

$$y_i \equiv a_i \pmod{p^t} \quad (t \geq u). \tag{2.8}$$

Then we have

$$\left.\begin{aligned} c_1a_1^3 + c_2a_2^3 + c_3y_3^3 + c_4y_4^3 + \ldots + c_sy_s^3 &\equiv 0 \\ d_1a_1^2 + d_2a_2^2 + d_3y_3^2 + d_4y_4^2 + \ldots + d_sy_s^2 &\equiv 0 \end{aligned}\right\} \pmod{p^u}$$

But then

$$\text{Max}\{|A(a_1,a_2)+C(y_3,\ldots,y_s)|_p\ ,\ |B(a_1,a_2)+D(y_3,\ldots,y_s)|_p\} \le p^{-u} < |\Delta(1,2)|_p^2$$

so by Lemma 1.3.1 there is a $(b_1,b_2) \in \mathbb{Z}_p \times \mathbb{Z}_p$ with $A(b_1,b_2) + C(y_3,\ldots,y_s) = 0$ and $B(b_1,b_2) + D(y_3,\ldots,y_s) = 0$. But by (2.8) there are p^{t-u} choices for each y_i $(i = 3,\ldots,s)$, and hence at least $p^{(t-u)(s-2)}$ possible solutions to the simultaneous congruences (2.7) with $q = p^t$. Then $M_n(p^t) \ge p^{(t-u)(s-2)}$ for all $t \ge u$.

(b) for all $i,j \in \{1,2,\ldots,v\}$ we have $\Delta(i,j) = 0$.

Then

$$c_id_ja_i - c_jd_ia_j = 0 \text{ for all } i,j \in \{1,2,\ldots,v\}. \tag{2.9}$$

There are 3 cases:

(1) there is no $i \in \{1,2,\ldots,v\}$ for which either $c_i = 0$ or $d_i = 0$.

Then by (2.9) we have

$$a_j = \begin{cases} \dfrac{c_1d_j}{d_1c_j} \cdot a_1 & \text{for } j = 1,\ldots,v \\ 0 & \text{for } j = v+1,\ldots,s \end{cases}$$

and we plainly have a non-trivial rational solution to the equations (3.1.2).

(2) there is an $i \in \{1,\ldots,v\}$ for which $c_i = 0$.

By a rearrangement of variables we have $i = 1$. Then by (2.9), $c_jd_1a_j = 0$ for every $j \in \{2,3,\ldots,v\}$, and as $d_1a_j \neq 0$, we must have $c_j = 0$ for every $j \in \{1,2,\ldots,v\}$. Then by condition (b) of Theorem 3.1.1, the equation

$$c_{v+1}x_{v+1}^3 + \ldots + c_sx_s^3 = 0 \tag{2.10}$$

must have at least 7 variables explicit, and so in particular we have $v \le s-7$. Also, $\Gamma^*(3) = 7$, and so the equation (2.10) has a

non-trivial p-adic solution $(b_{v+1},\ldots,b_s)$, say. Let $D = d_{v+1}b_{v+1}^2 + \ldots + d_s b_s^2$. Then we shall have found a non-trivial p-adic solution to (3.1.2) with $v+1$ variables non-zero provided that we can find a non-trivial $(u_1,\ldots,u_{v+1})$, with no u_i zero, such that

$$d_1u_1^2 + \ldots + d_vu_v^2 + Du_{v+1}^2 = 0. \tag{2.11}$$

Now, we know that this equation has the non-trivial solution $\underline{u} = (a_1,\ldots,a_v,0)$, with none of the a_i zero. Then as $d_1 \neq 0$, we have for some $w \geq 0$,

$$|2d_1a_1|_p^2 = p^{-w} > 0. \tag{2.12}$$

Put

$$u_{v+1} = p^{w+1}.$$

Then,

$$d_1a_1^2 + d_2a_2^2 + d_3a_3^2 + \ldots + d_va_v^2 + Du_{v+1}^2 \equiv 0 \pmod{p^{w+1}}. \tag{2.13}$$

Fixing u_{v+1}, a standard application of Hensel's Lemma for a single equation in the variable u_1 gives us a p-adic integer u_1 satisfying the equation

$$d_1u_1^2 + d_2a_2^2 + d_3a_3^2 + \ldots + d_va_v^2 + Du_{v+1}^2 = 0, \tag{2.14}$$

and further, u_1 is non-zero, since we have

$$\begin{aligned} d_1u_1^2 &= -(d_2a_2^2 + d_3a_3^2 + \ldots + d_va_v^2 + Du_{v+1}^2) \\ &\equiv d_1a_1^2 \pmod{p^{w+1}} \\ &\not\equiv 0 \pmod{p^{w+1}} \end{aligned}$$

respectively by (2.14), (2.13), and (2.12). Then we have a solution $(u_1,\ldots,u_{v+1})$ to (2.11) with no u_i zero, and hence a p-adic solution to (3.1.2) with $v+1$ variables non-zero. But this contradicts the maximality of v, and so this case can never occur.

(3) there is an i for which $d_i = 0$ for some $i \in \{1,\ldots,v\}$.

A similar argument to that in (2) applies.

This completes the proof of the lemma.

Notice that Lemmata 2.3 and 2.5 provide a proof of cases (d)(ii) and (iii) of Theorem 3.1.1. Henceforth we may therefore assume that $s \geq 14$, and

$$\left.\begin{array}{l} \text{the number of zero } d_i \text{ is } \leq 5 \\ \text{the number of zero } c_i \text{ is } \leq 3 \end{array}\right\} \tag{2.15}$$

and hence that there are at least 6 values of i for which both c_i and d_i are non-zero. Suppose that for $1 \leq i \leq s$ precisely m of the d_i are zero, precisely n of the c_i are zero, and let $h = s-m-n$. Then we may rearrange the variables with indices $1,\dots,s$ so that $d_i = 0$ for $i = 1,\dots,m$, $c_i \neq 0$ and $d_i \neq 0$ for $i = m+1,\dots,m+h$, and $c_i = 0$ for $i = m+h+1,\dots,s$.

By Lemmata 2.1 and 2.2, we may assume that

$$\begin{array}{c} \text{the equations (3.1.2) have a non-singular real solution} \\ (\eta_1,\dots,\eta_s) \text{ such that } 0 < \eta_i < \tfrac{1}{2} \text{ for } i = 1,\dots,s \ , \end{array} \tag{2.16}$$

since whenever necessary the c_i can be replaced by $-c_i$ and η_i^3 by $-\eta_i^3$, and by homogeneity η_i can be replaced by $\eta_i/(2.\text{Max } \eta_j)$. Also, by Lemma 2.6, we may assume that

$$\begin{array}{c} \text{for every rational prime } p \text{ there is a } u = u(p) < \infty \\ \text{such that for all } t \geq u,\ M_n(p^t) \geq p^{(t-u)(s-2)}. \end{array} \tag{2.17}$$

Let P be large (in terms of ε, $c_1,\dots,c_s$, $d_1,\dots,d_s$, $\eta_1,\dots,\eta_s$) and let α_i ($i = 2,3$) be real variables. Also, let

$$t_2 = \underset{1\leq i\leq s}{\text{Max}} |d_i| \ , \quad t_3 = \underset{1\leq i\leq s}{\text{Max}} |c_i| \ , \tag{2.18}$$

and

$$\delta = 10^{-2},\ 0 < \eta < \eta_0(\varepsilon) \text{ and } R = P^{\eta}. \tag{2.19}$$

Here η_0 is sufficiently small so that it satisfies the conditions necessary for both Theorem 3.1.2, and also Theorem 4.4 of Vaughan [1989a] to hold.

Write

$$\xi_i = \tfrac{1}{2}\eta_i, \quad \zeta_i = 2\eta_i, \tag{2.20}$$

$$F_i(\underline{\alpha}) = F_i(\alpha_2, \alpha_3) = \sum_{\xi_i P < x \le \zeta_i P} e(c_i\alpha_3 x^3 + d_i\alpha_2 x^2), \tag{2.21}$$

$$f_i(\underline{\alpha}) = f_i(\alpha_2, \alpha_3) = \sum_{\substack{\xi_i P < x \le \zeta_i P \\ x \in \mathcal{A}(P,R)}} e(c_i\alpha_3 x^3 + d_i\alpha_2 x^2), \tag{2.22}$$

$$F(\underline{\alpha}) = \sum_{0 < x \le P} e(\alpha_3 x^3 + \alpha_2 x^2), \tag{2.23}$$

$$f(\underline{\alpha}) = \sum_{x \in \mathcal{A}(P,R)} e(\alpha_3 x^3 + \alpha_2 x^2). \tag{2.24}$$

When we wish to stress the presence of zero coefficients, we shall write

$$g_i(\alpha) = f_i(0,\alpha), \quad H_i(\beta) = F_i(\beta,0), \tag{2.25}$$

$$g(\alpha) = f(0,\alpha), \quad H(\beta) = F(\beta,0). \tag{2.26}$$

Our objective is to estimate the number $R(P)$ of solutions of (3.1.2) in rational integers x_i which satisfy

$$\xi_i P < x_i \le \zeta_i P \quad \text{and} \quad x_i \in \mathcal{A}(P,R) \quad (i = 1, \ldots, m+h-4), \text{ and} \tag{2.27}$$

$$\xi_i P < x_i \le \zeta_i P \quad (i = m+h-3, \ldots, s). \tag{2.28}$$

We shall show that $R(P) \longrightarrow \infty$ as $P \longrightarrow \infty$, using a variant of the Hardy-Littlewood method, thereby completing the proof of Theorem 3.1.1. Let

$$Q_i = 18 t_i P^{i-1}, \ (i = 2,3), \quad \mathcal{U}_2^* = (Q_2^{-1}, 1+Q_2^{-1}) \times (Q_3^{-1}, 1+Q_3^{-1}). \tag{2.29}$$

Then,

$$R(P) = \iint_{\mathcal{U}_2^*} \mathcal{F}(\underline{\alpha})\mathcal{G}(\underline{\alpha})\mathcal{H}(\underline{\alpha})\, d\underline{\alpha}, \tag{2.30}$$

where

$$\left.\begin{aligned} \mathcal{F}(\underline{\alpha}) &= \prod_{i=1}^{m+h-4} f_i(\underline{\alpha}), \quad \mathcal{H}(\underline{\alpha}) = \prod_{i=m+h-3}^{m+h} F_i(\underline{\alpha}), \\ \text{and } \mathcal{G}(\underline{\alpha}) &= \prod_{i=m+h+1}^{s} F_i(\underline{\alpha}). \end{aligned}\right\} \tag{2.31}$$

The open square $\mathcal{U}_2^*$ is dissected in the following way. When $(a_2, a_3, q) = 1$ and $1 \le a_i \le q \le P$ $(i = 2,3)$, we denote a typical major arc by

$$\mathfrak{M}(q, \underline{a}) = \mathfrak{M}(q, a_2, a_3) = \{ \underline{\alpha} : |q\alpha_i - a_i| < Q_i^{-1} \text{ for } i = 2,3 \}. \tag{2.32}$$

The $\mathfrak{M}(q, \underline{a})$ are disjoint, since whenever $(a_2, a_3, q) \neq (a_2', a_3', q')$ (as ordered triples), $q, q' \le P$ and $(a_2, a_3, q) = (a_2', a_3', q') = 1$ (as highest common factors), we have

$$\left| \frac{a_i}{q} - \frac{a_i'}{q'} \right| \ge \frac{1}{qq'} > \left(\frac{1}{q} + \frac{1}{q'} \right) Q_i^{-1} \text{ , for one of } i = 2,3.$$

Let $\mathfrak{M}$ denote the union of the major arcs, and define the minor arcs, $\mathfrak{m}$, by

$$\mathfrak{m} = \mathcal{U}_2^* \setminus \mathfrak{M} . \tag{2.33}$$

Now let $2 \le W \le P^{1/2}$, and define the pruned major arcs,

$$\mathfrak{N}(q, \underline{a}) = \{ \underline{\alpha} : |q\alpha_i - a_i| < WP^{-1}Q_i^{-1} \text{ for } i = 2,3 \}, \tag{2.34}$$

and let $\mathfrak{N}$ denote the union of the $\mathfrak{N}(q, \underline{a})$ with $1 \le a_i \le q \le W$, $(a_2, a_3, q) = 1$. Plainly, $\mathfrak{N}(q, \underline{a}) \subset \mathfrak{M}(q, \underline{a})$ and $\mathfrak{N} \subset \mathfrak{M}$. Let $\mathfrak{n} = \mathcal{U}_2^* \setminus \mathfrak{N}$, so that

$$\mathfrak{n} = (\mathfrak{M} \setminus \mathfrak{N}) \cup \mathfrak{m} . \tag{2.35}$$

Throughout, ε is a sufficiently small positive number and the implied constants in the O, $\ll$, and $\gg$ notations depend on at most ε, $c_1, \dots, c_s$, $d_1, \dots, d_s$, $\eta_1, \dots, \eta_s$.

3. THE MINOR ARCS.

As usual, our minor arc estimate depends on bounds for the number of solutions of certain systems of diophantine equations. By use of the trivial inequality

$$|z_1 z_2 \dots z_n| \le |z_1|^n + \dots + |z_n|^n \tag{3.1}$$

and Hölder's inequality, we are able to put these equations into standard forms, the number of solutions of which may be easily estimated.

From Theorem 3.1.2, we have

$$\iint_{\mathfrak{U}_2^*} |f(\underline{\underline{\alpha}})|^{10}\, d\underline{\underline{\alpha}} \ll P^{35/6+\varepsilon}, \tag{3.2}$$

Also, by considering the underlying diophantine equations we deduce from Theorem 3.1.3 that

$$\iint_{\mathfrak{U}_2^*} |f(\underline{\underline{\alpha}})|^{6}\, d\underline{\underline{\alpha}} \ll P^{3+\varepsilon}. \tag{3.3}$$

By considering the underlying diophantine equations, from Theorem 4.4 of Vaughan [1989a] we have

$$\int_0^1 |g(\alpha)|^6\, d\alpha \ll P^{13/4+\varepsilon}. \tag{3.4}$$

Also, using classical estimates we have

$$\int_0^1 |H(\alpha)|^4\, d\alpha \ll P^{2+\varepsilon}. \tag{3.5}$$

We shall obtain a suitable minor arc estimate for $f(\underline{\underline{\alpha}})$ using the following lemma:

Lemma 3.1 (*R. Baker* [1986], *Theorem* 5.1). *Write*

$$\psi(x) = \alpha_k x^k + \ldots + \alpha_1 x,$$

$$S_m(\psi) = \sum_{n=1}^{N} e(m\psi(n)),$$

and $K = 2^{k-1}$. *Let* $N > C(k,\varepsilon)$, *and let* M *be a natural number and* T *a positive number such that*

$$(MNT^{-1})^{K+\varepsilon} \le N . \tag{3.6}$$

Suppose that

$$\sum_{m=1}^{M} |S_m(\psi)| \ge T.$$

Then there exists a natural number q *and integers* $v_1,\ldots,v_k$ *such that*

$$q < (MNT^{-1})^k . N^{\varepsilon}, \quad (q,v_1,\ldots,v_k) = 1, \quad (q,v_2,\ldots,v_k) \le MN^{\varepsilon},$$

$$|\alpha_j q - v_j| < M^{-1}(MNT^{-1})^k . N^{\varepsilon-j} \quad \text{for } j = 1,2,\ldots,k.$$

Lemma 3.2. *Suppose that* $m+1 \le i \le m+h$, $1 \le j \le m$, *and* $m+h+1 \le k \le s$. *Then we have*

(*i*) $$\iint_{\mathfrak{U}_2^*} |f_i(\underline{\alpha})|^4 |g_j(\alpha_3)|^6 \, d\underline{\alpha} \ll P^{21/4+\varepsilon},$$

(*ii*) $$\iint_{\mathfrak{U}_2^*} |f_i(\underline{\alpha})|^6 |H_k(\alpha_2)|^4 \, d\underline{\alpha} \ll P^{21/4+\varepsilon},$$

(*iii*) $$\iint_{\mathfrak{U}_2^*} |g_j(\alpha_3)|^6 |H_k(\alpha_2)|^4 \, d\underline{\alpha} \ll P^{21/4+\varepsilon},$$

(*iv*) $$\iint_{\mathfrak{U}_2^*} |f_i(\underline{\alpha})|^{10} \, d\underline{\alpha} \ll P^{35/6+\varepsilon}.$$

Proof: The first three results follow in a manner typified by case (i).

(i) By considering the underlying diophantine equation, we have that

$$\iint_{\mathfrak{U}_2^*} |f_i(\underline{\alpha})|^4 |g_j(\alpha_3)|^6 \, d\underline{\alpha}$$

is bounded above by the number of solutions to the simultaneous diophantine equations

$$\left.\begin{aligned} c_i(x_1^3+x_2^3) + c_j(y_1^3+y_2^3+y_3^3) &= c_i(x_1'^3+x_2'^3) + c_j(y_1'^3+y_2'^3+y_3'^3) \\ d_i(x_1^2+x_2^2) &= d_i(x_1'^2+x_2'^2) \end{aligned}\right\} \qquad (3.7)$$

with

$$x_r, x_r' \in \mathcal{A}(P,R)\ ,\ \xi_i P < x_r, x_r' \le \zeta_i P\ ,\ (r = 1,2),\ \text{and}$$

$$y_r, y_r' \in \mathcal{A}(P,R)\ ,\ \xi_j P < y_r, y_r' \le \zeta_j P\ (r = 1,2,3).$$

By (3.5), and considering the underlying diophantine equation, the number of solutions of the quadratic equation in (3.7) is at most

$$\int_0^1 |H(\alpha)|^4\, d\alpha \ll P^{2+\varepsilon}.$$

Given any such solution, let

$$X = c_i(x_1^3 + x_2^3 - x_1'^3 - x_2'^3)\ .$$

Then the number of solutions of the equation

$$c_j(y_1^3 + y_2^3 + y_3^3 - y_1'^3 - y_2'^3 - y_3'^3) = -X\ ,$$

with $y_r, y_r' \in \mathcal{A}(P,P^\eta)$ and $\xi_j P < y_r, y_r' \le \zeta_j P$ $(r = 1,2,3)$, is at most

$$\int_0^1 |g(c_j\alpha)|^6 e(\alpha X)\, d\alpha \le \int_0^1 |g(c_j\alpha)|^6\, d\alpha \ll P^{13/4+\varepsilon}\ ,$$

by considering the underlying diophantine equation, and using (3.4). Thus the number of solutions of (3.7) is

$$\ll P^{2+\varepsilon}.P^{13/4+\varepsilon} = P^{21/4+2\varepsilon}.$$

(iv) By (2.22), (2.24), and considering the underlying diophantine equations, we have by using (3.2) that

$$\iint_{\mathfrak{U}_2^*} |f_i(\underline{\underline{\alpha}})|^{10}\, d\underline{\underline{\alpha}} \le \iint_{\mathfrak{U}_2^*} |f(\underline{\underline{\alpha}})|^{10}\, d\underline{\underline{\alpha}} \ll P^{35/6+\varepsilon}.$$

This completes the proof of the lemma.

Lemma 3.3. *We have*

$$\iint_{\mathfrak{U}_2^*} |\mathfrak{F}(\underline{\alpha})\mathfrak{G}(\underline{\alpha})|\, d\underline{\alpha} \ll P^{s-8-2\delta}.$$

Proof: By (2.31), we have

$$|\mathfrak{F}(\underline{\alpha})\mathfrak{G}(\underline{\alpha})| = \left(\prod_{i=m+1}^{m+h-4} |f_i(\underline{\alpha})|\right)\left(\prod_{j=1}^{m} |g_j(\alpha_3)|\right)\left(\prod_{k=m+h+1}^{s} |H_k(\alpha_2)|\right).$$

By (3.1), we obtain

$$|\mathfrak{F}(\underline{\alpha})\mathfrak{G}(\underline{\alpha})| \le \left(\sum_{i=m+1}^{m+h-4} |f_i(\underline{\alpha})|^{h-4}\right)\left(\sum_{j=1}^{m} |g_j(\alpha_3)|^m\right)\left(\sum_{k=m+h+1}^{s} |H_k(\alpha_2)|^n\right)$$
$$= \sum_{i=m+1}^{m+h-4}\sum_{j=1}^{m}\sum_{k=m+h+1}^{s} |f_i(\underline{\alpha})|^{h-4}\, |g_j(\alpha_3)|^m\, |H_k(\alpha_2)|^n.$$

Hence,

$$\iint_{\mathfrak{U}_2^*} |\mathfrak{F}(\underline{\alpha})\mathfrak{G}(\underline{\alpha})|\, d\underline{\alpha} \ll \max_{i,j,k} \iint_{\mathfrak{U}_2^*} |f_i(\underline{\alpha})|^{h-4} |g_j(\alpha_3)|^m |H_k(\alpha_2)|^n\, d\underline{\alpha},$$

where the maximum is taken over the ranges of i, j, k in the previous equation. Then there are integers I, J, K such that

$$\iint_{\mathfrak{U}_2^*} |\mathfrak{F}(\underline{\alpha})\mathfrak{G}(\underline{\alpha})|\, d\underline{\alpha} \ll \iint_{\mathfrak{U}_2^*} |f_I(\underline{\alpha})|^{h-4} |g_J(\alpha_3)|^m |H_K(\alpha_2)|^n\, d\underline{\alpha}. \tag{3.8}$$

By (2.15), we may assume that $m \le 5$, and $n \le 3$. For the moment write, for convenience's sake, $f^{h-4}g^mH^n$ for

$$|f_I(\underline{\alpha})|^{h-4} |g_J(\alpha_3)|^m |H_K(\alpha_2)|^n.$$

Then using (3.1) we can simplify the expression in (3.8) to a form in which we can use Lemma 3.2. First note that by bounding f, g and H trivially, we can always write

$$f^{h-4}g^mH^n \ll P^{s-14} f^L g^M H^N,$$

where $L \le h-4$, $M \le m$, $N \le n$, and $L+M+N = 10$. Then by repeated use of (3.1), we obtain

$$f^{h-4}g^{m}H^{n} \ll P^{s-14}(f^{10-N}H^{N} + f^{5-N}g^{5}H^{N})$$
$$\ll P^{s-14}(f^{10} + f^{6}H^{4} + fg^{5}H^{4} + f^{5}g^{5})$$
$$\ll P^{s-14}(f^{10} + f^{6}H^{4} + f^{4}g^{6} + g^{6}H^{4}) .$$

Then from Lemma 3.2, we have

$$\iint_{\mathfrak{U}_2^*} |\mathcal{F}(\underline{\underline{\alpha}})\mathcal{G}(\underline{\underline{\alpha}})|\, d\underline{\underline{\alpha}} \ll P^{s-14}(P^{21/4+\varepsilon} + P^{35/6+\varepsilon})$$
$$\ll P^{s-8-2\delta} .$$

This completes the proof of the lemma.

Lemma 3.4. *We have*

$$\iint_{\mathfrak{m}} |\mathcal{F}(\underline{\underline{\alpha}})\mathcal{G}(\underline{\underline{\alpha}})\mathcal{H}(\underline{\underline{\alpha}})|\, d\underline{\underline{\alpha}} \ll P^{s-5-\delta}.$$

Proof: Using inequality (3.1), we have

$$\sup_{\underline{\underline{\alpha}} \in \mathfrak{m}} |\mathcal{H}(\alpha)| \ll \max_{m+h-3 \le i \le m+h} \sup_{\underline{\underline{\alpha}} \in \mathfrak{m}} |F_i(\underline{\underline{\alpha}})|^4.$$

Suppose that the maximum in the above expression occurs at $i = I$, and let $t_3' = c_I$ and $t_2' = d_I$.

Suppose that $F(t_2'\alpha_2, t_3'\alpha_3) \ge P^{3/4+\varepsilon}$. Then by Lemma 3.1, for P sufficiently large, there exist integers a_2, a_3, and q with

$$1 \le q < P^{1-\varepsilon}, \quad (a_2, a_3, q) = 1,$$

and such that

$$|qt_i'\alpha_i - a_i| < P^{1-i-\varepsilon}, \quad \text{for } i = 2,3.$$

Put

$$q' = |[t_2', t_3']|q \ , \text{ and } a_i' = t_i'^{-1}a_i|[t_2', t_3']| \ , \text{ for } i = 2,3.$$

Then since $\underline{\underline{\alpha}} \in \mathfrak{U}_2^*$, we have from (2.29) that $1 \le a_i' \le q'$ for $i = 2,3$. Thus for P sufficiently large, we have

$$1 \le a_2', a_3' \le q' < P, \quad (a_2', a_3', q') = 1,$$

and

$$|q'\alpha_i - a_i'| < Q_i^{-1} \quad (i = 2,3).$$

Consequently, if $F(t_2'\alpha_2, t_3'\alpha_3) \geq P^{3/4+\varepsilon}$, then by (2.32) we have $\underline{\alpha} \in \mathfrak{M}$, and hence if $\underline{\alpha} \in \mathfrak{m}$, then we must have

$$F(d_I\alpha_2, c_I\alpha_3) \ll P^{3/4+\varepsilon}. \tag{3.9}$$

Then for $\underline{\alpha} \in \mathfrak{m}$, we have

$$\sum_{\zeta_I P < x \leq \zeta_I P} e(c_I\alpha_3 x^3 + d_I\alpha_2 x^2) \ll \Big| \sum_{0 < x \leq \zeta_I P} e(c_I\alpha_3 x^3 + d_I\alpha_2 x^2) \Big| + \Big| \sum_{0 < x \leq \zeta_I P} e(c_I\alpha_3 x^3 + d_I\alpha_2 x^2) \Big|$$

so from (3.9), $F_I(\underline{\alpha}) \ll P^{3/4+\varepsilon}$, and hence by Lemma 3.3,

$$\begin{aligned} \iint_{\mathfrak{m}} |\mathcal{F}(\underline{\alpha})\mathcal{G}(\underline{\alpha})\mathcal{H}(\underline{\alpha})| \, d\underline{\alpha} &\ll \Big(\sup_{\underline{\alpha} \in \mathfrak{m}} |F_I(\underline{\alpha})| \Big)^4 \iint_{\mathcal{U}_2^*} |\mathcal{F}(\underline{\alpha})\mathcal{G}(\underline{\alpha})| \, d\underline{\alpha} \\ &\ll P^{3+4\varepsilon} \cdot P^{s-8-2\delta} \\ &\ll P^{s-5-\delta}. \end{aligned}$$

This completes the proof of the lemma.

4. GENERATING FUNCTIONS.

Here we give estimates for the various generating functions that arise in our treatment of the major arcs. For the F_1, we are able to use classical estimates, similar to those found in Vaughan [1981b], §7. However, the f_1 are more difficult, requiring a treatment based on Vaughan [1989a], Lemmata 5.3 and 5.4.

Let

$$S_1(q, \underline{a}) = S_1(q, a_2, a_3) = \sum_{x=1}^{q} e((d_1a_2x^2 + c_1a_3x^3)/q) \, . \tag{4.1}$$

Lemma 4.1. *Suppose that* $(q, a_2, a_3) = 1$. *Then*

$$S_i(q, \underline{a}) \ll \begin{cases} (a_3, q)^{1/3} . q^{2/3+\varepsilon} & \text{for } i = 1, \dots, m \\ q^{2/3+\varepsilon} & \text{for } i = m+1, \dots, m+h \\ (a_2, q)^{1/2} . q^{1/2+\varepsilon} & \text{for } i = m+h+1, \dots, s \ . \end{cases}$$

Proof: There are three cases.

(i) $i = 1, \dots, m$.

Then $d_i = 0$ and $c_i \neq 0$, so by R.Baker [1986], Lemma 4.2, we have

$$S_i(q, \underline{a}) \ll (c_i a_3, q)^{1/3} q^{2/3+\varepsilon} \ll (a_3, q)^{1/3} q^{2/3+\varepsilon}.$$

(ii) $i = m+1, \dots, m+h$.

Then $c_i \neq 0$ and $d_i \neq 0$, so by R.Baker [1986], Lemma 4.2, we have

$$S_i(q, \underline{a}) \ll (c_i a_3, d_i a_2, q)^{1/3} q^{2/3+\varepsilon} \ll q^{2/3+\varepsilon}.$$

(iii) $i = m+h+1, \dots, s$.

Then $c_i = 0$ but $d_i \neq 0$. The proof proceeds as in (i).

This completes the proof of the lemma.

When $\underline{\alpha} \in \mathfrak{M}(q, \underline{a})$, let

$$\beta_j = \alpha_j - a_j/q \ , \ (j = 2,3), \tag{4.2}$$

and

$$v_i(\underline{\beta}) = \int_{\zeta_i P}^{\zeta_i P} e(c_i \beta_3 \gamma^3 + d_i \beta_2 \gamma^2) \, d\gamma \ . \tag{4.3}$$

Lemma 4.2. *Suppose that* $1 \leq i \leq s$, *and that* q, a_2, *and* a_3 *are integers with* $(q, d_i a_2, c_i a_3) = d$. *Suppose also that* $\underline{\alpha} \in \mathfrak{M}(q, \underline{a})$. *Then with* β_j *defined by* (4.2), *and writing*

$$V_i(\underline{\alpha}, q, \underline{a}) = q^{-1} S_i(q, \underline{a}) v_i(\underline{\beta}) \ , \tag{4.4}$$

we have

$$F_i(\underline{\alpha}) - V_i(\underline{\alpha}, q, \underline{a}) \ll d^{1/3} q^{2/3+\varepsilon}.$$

Proof: This follows from Lemma 4.4 of R.Baker [1986].

Let $\rho(x)$ denote Dickman's function, defined for real x by

$$\left.\begin{aligned} \rho(x) &= 0 \text{ when } x \le 0, \\ \rho(x) &= 1 \text{ when } 0 < x \le 1, \\ \rho &\text{ is continuous for } x > 0, \\ \rho &\text{ is differentiable for } x > 1, \\ x\rho'(x) &= -\rho(x-1) \text{ for } x > 1. \end{aligned}\right\} \tag{4.5}$$

We further define

$$w_1(\underline{\underline{\beta}}) = \int_{\xi_1 P}^{\xi_1 P} \rho\left(\frac{\log\gamma}{\log R} \right) e(c_1\beta_3\gamma^3 + d_1\beta_2\gamma^2)\, d\gamma . \tag{4.6}$$

With β_j defined by (4.2), write

$$W_1(\underline{\underline{\alpha}}, q, \underline{\underline{a}}) = q^{-1} S_1(q, \underline{\underline{a}}) w_1(\underline{\underline{\beta}}) . \tag{4.7}$$

We have the estimate:

Lemma 4.3 (*Vaughan* [1989a], *Lemma* 5.3). *Let τ be a fixed positive number, and suppose that $R \le X \le R^{\tau}$. Then*

$$card(\mathcal{A}(X,R)) = X.\rho\left(\frac{logX}{logR} \right) + O(X/logX) .$$

We also have a result analogous to Lemma 2.6 of Vaughan [1981b].

Lemma 4.4. *Let $c_1, c_2, \ldots$ be any sequence of complex numbers, and suppose that F has a continuous derivative on $[0,X]$. Then*

$$\sum_{m \in \mathcal{A}(X,R)} c_m F(m) = \left(\sum_{m \in \mathcal{A}(X,R)} c_m \right) F(X) - \int_0^X \left(\sum_{m \in \mathcal{A}(\gamma,R)} c_m \right) F'(\gamma)\, d\gamma .$$

Proof: We have

$$F(m) = F(X) - \int_m^X F'(\gamma)\, d\gamma .$$

Therefore

$$\sum_{m \in \mathcal{A}(X,R)} c_m F(m) = \left(\sum_{m \in \mathcal{A}(X,R)} c_m \right) F(X) - \sum_{m \in \mathcal{A}(X,R)} \int_m^X c_m F'(\gamma)\, d\gamma ,$$

and the result follows on interchanging the order of summation and integration.

We prove the next lemma in rather greater generality than is strictly necessary for our purposes.

Lemma 4.5. *Let* $R = P^{\eta}$ *for some* $0 < \eta < \eta_0(\varepsilon, k) < 1$, *and let*

$$h(\underline{\underline{\alpha}}) = \sum_{x \in \mathcal{A}(P,R)} e(\alpha_k x^k + \ldots + \alpha_1 x) , \tag{4.8}$$

$$S(q, \underline{\underline{a}}) = \sum_{x=1}^{q} e((a_k x^k + \ldots + a_1 x)/q) , \tag{4.9}$$

$$w(\underline{\underline{\beta}}) = \int_R^P \rho\left(\frac{\log\gamma}{\log R} \right) e(\beta_k \gamma^k + \ldots + \beta_1 \gamma)\, d\gamma . \tag{4.10}$$

Let $\beta_i = \alpha_i - a_i/q$ *for* $i = 1, \ldots, k$, *and write*

$$W(\underline{\underline{\alpha}}, q, \underline{\underline{a}}) = q^{-1} S(q, \underline{\underline{a}}) w(\underline{\underline{\beta}}) . \tag{4.11}$$

Suppose that $q \le R$, *and* $(a_1, \ldots, a_k, q) \ll 1$. *Then*

$$h(\underline{\underline{\alpha}}) = W(\underline{\underline{\alpha}}, q, \underline{\underline{a}}) + O\left(\frac{qP}{\log P} \left(1 + P|\beta_1| + \ldots + P^k|\beta_k|\right) \right) . \tag{4.12}$$

Proof: As in the proof of Vaughan [1989a] Lemma 5.4, we have

$$\sum_{\substack{x \in \mathcal{A}(m,R) \\ x \equiv r \pmod q}} 1 = q^{-1} \sum_{x \in \mathcal{A}(m,R)} 1 + O(P/\log P) \quad (R < m \le P).$$

Therefore

$$\sum_{x \in \mathcal{A}(m,R)} e((a_k x^k + \ldots + a_1 x)/q) = q^{-1} S(q, \underline{\underline{a}}) \sum_{x \in \mathcal{A}(m,R)} 1 + O(qP/\log P)$$

$$= q^{-1} S(q, \underline{\underline{a}}) . m\rho\left(\frac{\log m}{\log R} \right) + O(qP/\log P) \tag{4.13}$$

by Lemma 4.3 and (4.9). Then by using Lemma 4.4, we have

$$\sum_{x \in \mathcal{A}(P,R)} e((a_k x^k + \ldots + a_1 x)/q)\, e(\beta_k x^k + \ldots + \beta_1 x) = S_0 - S_1 , \tag{4.14}$$

where

$$S_0 = e(\beta_k P^k + \ldots + \beta_1 P) \sum_{x \in \mathcal{A}(P,R)} e((a_k x^k + \ldots + a_1 x)/q)$$

$$= e(\beta_k P^k + \ldots + \beta_1 P).q^{-1}S(q,\underline{\underline{a}}).P\rho\left(\frac{\log P}{\log R}\right) + O\left(\frac{qP}{\log P}\right), \quad (4.15)$$

by (4.13), and

$$S_1 = \int_0^P \left(\sum_{x \in \mathcal{A}(\gamma,R)} e((a_k x^k + \ldots + a_1 x)/q) \right) \frac{d}{d\gamma}\left(e(\beta_k \gamma^k + \ldots + \beta_1 \gamma)\right) d\gamma .$$

But we have

$$\int_0^R \left(\sum_{x \in \mathcal{A}(\gamma,R)} e((a_k x^k + \ldots + a_1 x)/q) \right) \frac{d}{d\gamma}\left(e(\beta_k \gamma^k + \ldots + \beta_1 \gamma)\right) d\gamma$$

$$\leq \int_0^R \gamma(|\beta_1| + \ldots + k|\beta_k|\gamma^{k-1})\, d\gamma$$

$$\ll R\left(|\beta_1|P + \ldots + |\beta_k|P^k \right).$$

Then by (4.13) we have

$$S_1 = q^{-1}S(q,\underline{\underline{a}}) \int_R^P \gamma\rho\left(\frac{\log\gamma}{\log R}\right) \frac{d}{d\gamma}\left(e(\beta_k \gamma^k + \ldots + \beta_1 \gamma)\right) d\gamma - \Delta , \quad (4.16)$$

where

$$\Delta \ll R\left(|\beta_1|P + \ldots + |\beta_k|P^k\right) + \frac{qP}{\log P} \int_R^P \left|\frac{d}{d\gamma}\left(e(\beta_k \gamma^k + \ldots + \beta_1 \gamma)\right)\right| d\gamma$$

$$\ll \frac{qP}{\log P}\left(|\beta_1|P + \ldots + |\beta_k|P^k \right).$$

Then by (4.8), (4.14), (4.15) and (4.16), we have

$$h(\underline{\underline{\alpha}}) = q^{-1}S(q,\underline{\underline{a}}) \left(e(\beta_k P^k + \ldots + \beta_1 P).P\rho\left(\frac{\log P}{\log R}\right) \right.$$

$$\left. - \int_R^P \gamma\rho\left(\frac{\log\gamma}{\log R}\right) \frac{d}{d\gamma}\left(e(\beta_k \gamma^k + \ldots + \beta_1 \gamma)\right) d\gamma \right)$$

$$+ \Delta + O(qP/\log P) ,$$

and integrating by parts, we obtain

$$h(\underline{\underline{\alpha}}) = q^{-1}S(q,\underline{\underline{a}}) \int_R^P e(\beta_k \gamma^k + \ldots + \beta_1 \gamma).\left(\rho\left(\frac{\log\gamma}{\log R}\right) + \frac{1}{\log R}\,\rho'\left(\frac{\log\gamma}{\log R}\right)\right) d\gamma$$

$$+ \Delta + O(qP/\log P) .$$

By using (4.5), and observing that for $\gamma \in [R,P]$, $\rho\left(\frac{\log\gamma}{\log R}\right)$ is a monotonic decreasing function of γ, and further that $1 \ll \rho\left(\frac{\log\gamma}{\log R}\right) \ll 1$, we deduce that

$$\int_R^P e(\beta_k\gamma^k+\ldots+\beta_1\gamma).\frac{1}{\log R}\,\rho'\left(\frac{\log\gamma}{\log R}\right)\,d\gamma \ll \int_R^P \frac{1}{\log\gamma}\,\rho\left(\frac{\log\gamma}{\log R}-1\right)\,d\gamma$$
$$\ll \int_R^P \frac{d\gamma}{\log\gamma}$$
$$\ll P/\log P\ .$$

Then by (4.10) and (4.11), we have

$$h(\underline{\underline{\alpha}}) - W(\underline{\underline{\alpha}},q,\underline{\underline{a}}) \ll \frac{qP}{\log P}\,(1 + |\beta_1|P + \ldots + |\beta_k|P^k)\ .$$

This completes the proof of the lemma.

Lemma 4.6. *We have*

$$v_1(\underline{\underline{\beta}}) \ll P(1 + P^2|\beta_2| + P^3|\beta_3|)^{-1/3},$$

and

$$w_1(\underline{\underline{\beta}}) \ll P(1 + P^2|\beta_2| + P^3|\beta_3|)^{-1/3}.$$

Proof: The first bound follows from Vaughan [1981b] Theorem 7.3 on making a change of variable. The second bound follows on noting that $\rho\left(\frac{\log\gamma}{\log R}\right)$ is a monotonic decreasing function of γ, and futher that $1 \ll \rho\left(\frac{\log\gamma}{\log R}\right) \ll 1$ for $\gamma \in [R,P]$. Hence the argument of the proof of Vaughan [1981b] Theorem 7.3 may be applied with only trivial modifications.

This completes the proof of the lemma.

5. PRUNING THE MAJOR ARCS.

Since our knowledge of the behaviour of the generating functions associated with the set $\mathcal{A}(P,R)$ is rather imprecise, we are forced to hard prune the major arcs. Here then, we show that the contribution from $\mathfrak{M}\setminus\mathfrak{N}$ to the integral $R(P)$ is comparatively small.

Lemma 5.1. *Suppose that* $m+1 \le i \le m+h$, $1 \le j \le m$, *and* $m+h+1 \le k \le s$. *Then we have*

$$(i) \quad \iint_{\mathcal{U}_2^*} |f_i(\alpha_2,\alpha_3)|^8 |H_k(\alpha_2)|^5 \, d\underline{\alpha} \ll P^8,$$

$$(ii) \quad \iint_{\mathcal{U}_2^*} |g_j(\alpha_3)|^8 |H_k(\alpha_2)|^5 \, d\underline{\alpha} \ll P^8,$$

$$(iii) \quad \iint_{\mathcal{U}_2^*} |f_i(\alpha_2,\alpha_3)|^6 |g_j(\alpha_3)|^8 \, d\underline{\alpha} \ll P^9.$$

Proof: (i) First note that as a simple application of the Hardy-Littlewood method, we have

$$\int_0^1 |H_I(\alpha)|^5 \, d\alpha \ll P^3, \tag{5.1}$$

for each I for which d_I is non-zero.

Let S_1 be the number of solutions of the equation

$$x_1^3 + \ldots + x_4^3 = y_1^3 + \ldots + y_4^3 ,$$

with $x_r, y_r \in \mathcal{A}(P,R)$ and $\xi_i P < x_r, y_r \le \zeta_i P$ $(r = 1,\ldots,4)$. Then we can plainly extend the range of the variables, and apply Vaughan [1986a], Theorem 2 to obtain

$$S_1 \ll P^5. \tag{5.2}$$

Also, we have

$$\int_0^1 |f_i(\underline{\alpha})|^8 \, d\alpha_3 = \int_0^1 \sum_{\underline{x},\underline{y}} e(s_3(\underline{x},\underline{y})\alpha_3 + s_2(\underline{x},\underline{y})\alpha_2) \, d\alpha_3 ,$$

where the summation over $\underline{x}$, $\underline{y}$ denotes summation over the variables

$x_r, y_r \in \mathcal{A}(P,R)$ with $\xi_1 P < x_r, y_r \le \zeta_1 P$ $(r = 1,\dots,4)$, and where

$$s_m(\underline{\underline{x}}, \underline{\underline{y}}) = (x_1^m - y_1^m) + \dots + (x_4^m - y_4^m) .$$

Write $c(\underline{\underline{x}}, \underline{\underline{y}})$ for $e(s_2(\underline{\underline{x}}, \underline{\underline{y}})\alpha_2)$. Then

$$\int_0^1 |f_1(\underline{\underline{\alpha}})|^8 \, d\alpha_3 = \sum_{\underline{\underline{x}}, \underline{\underline{y}}} c(\underline{\underline{x}}, \underline{\underline{y}}) \int_0^1 e(s_3(\underline{\underline{x}}, \underline{\underline{y}})\alpha) \, d\alpha ,$$

$$= \sum_{\underline{\underline{x}}, \underline{\underline{y}}} c(\underline{\underline{x}}, \underline{\underline{y}}) \delta(\underline{\underline{x}}, \underline{\underline{y}}) ,$$

where

$$\delta(\underline{\underline{x}}, \underline{\underline{y}}) = \begin{cases} 1 \text{ if } x_1^3 + \dots + x_4^3 = y_1^3 + \dots + y_4^3 \\ 0 \text{ otherwise.} \end{cases}$$

Hence

$$\int_0^1 |f_1(\underline{\underline{\alpha}})|^8 \, d\alpha_3 \le \sum_{\underline{\underline{x}}, \underline{\underline{y}}} \delta(\underline{\underline{x}}, \underline{\underline{y}}) = S_1 . \tag{5.3}$$

Thus by (5.2) and (5.3), we have

$$\iint_{\mathcal{U}_2^*} |f_1(\underline{\underline{\alpha}})|^8 |H_k(\alpha_2)|^5 \, d\underline{\underline{\alpha}} \ll P^5 \int_0^1 |H_k(\alpha)|^5 \, d\alpha$$

$$\ll P^8,$$

by using (5.1).

Case (ii) follows in a like, though simpler manner.

(iii) $\iint_{\mathcal{U}_2^*} |f_1(\alpha_2, \alpha_3)|^6 |g_j(\alpha_3)|^8 \, d\underline{\underline{\alpha}}$ is the number of solutions of the simultaneous equations

$$\left.\begin{aligned} c_1(x_1^3+x_2^3+x_3^3)+c_j(y_1^3+\dots+y_4^3) &= c_1(x_1'^3+x_2'^3+x_3'^3)+c_j(y_1'^3+\dots+y_4'^3) \\ d_1(x_1^2+x_2^2+x_3^2) &= d_1(x_1'^2+x_2'^2+x_3'^2) \end{aligned}\right\} \tag{5.4}$$

with

$$x_r, x_r' \in \mathcal{A}(P,R), \ \xi_1 P < x_r, x_r' \le \zeta_1 P \text{ for } r = 1,2,3, \text{ and}$$

$$y_r, y_r' \in \mathcal{A}(P,R), \ \xi_j P < y_r, y_r' \le \zeta_j P \text{ for } r = 1,\dots,4.$$

The proof now follows the pattern of Lemma 3.2(i), using the estimates

$$\int_0^1 |H_i(\alpha)|^6 \, d\alpha \ll P^4 \ , \text{ and } \int_0^1 |g_j(\alpha)|^8 \, d\alpha \ll P^5,$$

obtained respectively from (5.1) and (5.2).

This completes the proof of the lemma.

Lemma 5.2. *Suppose that both* c_i *and* d_i *are non-zero, and* $t > 9$. *Then we have*

$$\iint_{\mathfrak{M}} |F_i(\underline{\underline{\alpha}})|^t \, d\underline{\underline{\alpha}} \ll P^{t-5},$$

and for some $\sigma > 0$,

$$\iint_{\mathfrak{M}\setminus\mathfrak{N}} |F_i(\underline{\underline{\alpha}})|^t \, d\underline{\underline{\alpha}} \ll W^{-\sigma} P^{t-5}.$$

Proof: Define $V_i(\underline{\underline{\alpha}})$ for $\underline{\underline{\alpha}} \in \mathfrak{M}$ by putting $V_i(\underline{\underline{\alpha}}) = V_i(\underline{\underline{\alpha}}, q, \underline{a})$ when $\underline{\underline{\alpha}} \in \mathfrak{M}(q, \underline{a})$, $1 \le a_j \le q \le P$ $(j = 2,3)$ and $(a_2, a_3, q) = 1$. We shall write $\mathcal{M}$ for either $\mathfrak{M}$ or $\mathfrak{M}\setminus\mathfrak{N}$, and $\mathcal{M}(q, \underline{a})$ for either $\mathfrak{M}(q, \underline{a}) \cap \mathfrak{M}$ or $(\mathfrak{M}(q, \underline{a}) \cap \mathfrak{M})\setminus(\mathfrak{N}(q, \underline{a}) \cap \mathfrak{N})$. In addition, we let $Y = 1$ if $\mathcal{M} = \mathfrak{M}$, and $Y = W$ if $\mathcal{M} = \mathfrak{M}\setminus\mathfrak{N}$. Also, we write β_i for $\alpha_i - a_i/q$ $(i = 2,3)$ whenever $\underline{\underline{\alpha}} \in \mathcal{M}(q, \underline{a})$.

Suppose that $1 \le a_j \le q \le P$ $(j = 2,3)$ and $(a_2, a_3, q) = 1$. Then by Lemma 4.2, for $\underline{\underline{\alpha}} \in \mathfrak{M}(q, \underline{a})$ we have

$$|F_i(\underline{\underline{\alpha}}) - V_i(\underline{\underline{\alpha}}, q, \underline{a})| \ll q^{2/3+\varepsilon}. \tag{5.5}$$

Then for $\underline{\underline{\alpha}} \in \mathcal{M}(q, \underline{a})$, we have

$$|F_i(\underline{\underline{\alpha}})|^t - |V_i(\underline{\underline{\alpha}})|^t \ll \mathcal{T}_1 + \mathcal{T}_2 \ , \tag{5.6}$$

where

$$\mathcal{T}_1 = (q^{2/3+\varepsilon})^t \text{ and } \mathcal{T}_2 = (q^{2/3+\varepsilon})|V_i(\underline{\underline{\alpha}})|^{t-1}.$$

Now

$$\iint_{\mathcal{M}} \mathcal{T}_1 \, d\underline{\underline{\alpha}} \ll \sum_{q \le P} \sum_{\substack{1 \le a \le q \\ (a,b,q)=1}} \sum_{1 \le b \le q} (q^{2/3+\varepsilon})^t \iint_{\mathcal{M}(q,a,b)} d\underline{\underline{\alpha}}$$

$$\ll \sum_{q \le P} (q^{2/3+\varepsilon})^t P^{-3},$$

by (2.29), (2.32) and (2.34). Hence, provided $2t/3 - 2 < t - 5$ (that is $t > 9$), we have

$$\iint_{\mathfrak{M}} \mathcal{T}_1 \, d\underline{\alpha} \ll P^{t-5}Y^{-\sigma}, \tag{5.7}$$

for some $\sigma > 0$.

Also, by (4.4) and Lemma 4.1,

$$\iint_{\mathfrak{M}} \mathcal{T}_2 \, d\underline{\alpha} \ll \sum_{q\le P} \sum_{\substack{1\le a\le q \\ (a,b,q)=1}} \sum_{1\le b\le q} q(q^{-1/3+\varepsilon})^t \iint_{\mathfrak{M}(q,a,b)} |v_1(\underline{\beta})|^{t-1} d\underline{\alpha} \tag{5.8}$$

But by Lemma 4.6, we have

$$\iint_{\mathfrak{M}(q,a,b)} |v_1(\underline{\beta})|^T \, d\underline{\alpha} \ll P^T \iint_{\mathfrak{M}(q,a,b)} (1 + P^2|\beta_2| + P^3|\beta_3|)^{-T/3} \, d\underline{\alpha} .$$

Then if $T > 6$, the right hand side of the last expression is, by (2.29), (2.32) and (2.34),

$$\ll P^{T-5}\mathrm{Min}\{ 1 , (q/Y)^{T/3-2} \} , \tag{5.9}$$

Now, for $t > 9$, recalling that $W \le P^{1/2}$, we have

$$\sum_{q\le Y} q^{3-t/3+t\varepsilon} \mathrm{Min}\{ 1 , (q/Y)^{(t-1)/3-2} \} \ll Y^{2-(t-1)/3} \sum_{q\le Y} q^{2/3+t\varepsilon}$$
$$\ll PY^{-\sigma},$$

for some $\sigma > 0$, and

$$\sum_{Y<q\le P} q^{3-t/3+t\varepsilon} \mathrm{Min}\{ 1 , (q/Y)^{(t-1)/3-2} \} \ll P \sum_{q>Y} q^{2-t/3+t\varepsilon}$$
$$\ll PY^{-\sigma},$$

for some $\sigma > 0$. Then by (5.8) and (5.9) (with $T = t-1$), we have

$$\iint_{\mathfrak{M}} \mathcal{T}_2 \, d\underline{\alpha} \ll P^{t-5}Y^{-\sigma}. \tag{5.10}$$

Finally, we have

$$\iint_{\mathfrak{M}} |V_1(\underline{\alpha})|^t d\underline{\alpha} \ll \sum_{q\le P} \sum_{\substack{1\le a\le q \\ (a,b,q)=1}} \sum_{1\le b\le q} (q^{-1/3+\varepsilon})^t \iint_{\mathfrak{M}(q,a,b)} |v_1(\underline{\beta})|^t d\underline{\alpha} . \tag{5.11}$$

But for $t > 9$, we have

$$\sum_{q\le Y} q^{2-t/3+t\varepsilon} \mathrm{Min}\{ 1 , (q/Y)^{t/3-2} \} \ll Y^{2-t/3} \sum_{q\le Y} q^{t\varepsilon}$$
$$\ll Y^{-\sigma},$$

for some $\sigma > 0$, and

$$\sum_{q>Y} q^{2-t/3+t\varepsilon} \operatorname{Min}\{ 1 , (q/Y)^{t/3-2} \} \ll \sum_{q>Y} q^{2-t/3+t\varepsilon} \ll Y^{-\sigma},$$

for some $\sigma > 0$. Then by (5.9) and (5.11), we have

$$\iint_{\mathcal{M}} |V_i(\underline{\underline{\alpha}})|^t \, d\underline{\underline{\alpha}} \ll P^{t-5} Y^{-\sigma}. \tag{5.12}$$

The results follow on collecting together (5.6), (5.7), (5.10) and (5.12).

This completes the proof of the lemma.

Lemma 5.3. *Suppose that* c_i, d_i *and* d_j *are non-zero, and* $t > 6$. *Then we have*

$$\iint_{\mathfrak{M}} |F_i(\underline{\underline{\alpha}})|^t |H_j(\alpha_2)|^5 \, d\underline{\underline{\alpha}} \ll P^t,$$

and for some $\sigma > 0$,

$$\iint_{\mathfrak{m}\setminus\mathfrak{M}} |F_i(\underline{\underline{\alpha}})|^t |H_j(\alpha_2)|^5 \, d\underline{\underline{\alpha}} \ll W^{-\sigma} P^t.$$

Proof: Adopt the same notation (as regards $V_i(\underline{\underline{\alpha}})$, $\mathcal{M}$, Y and $\underline{\underline{\beta}}$) as in the previous lemma. Now by Lemma 4.2, we have

$$\left(|F_i(\underline{\underline{\alpha}})|^t - |V_i(\underline{\underline{\alpha}})|^t \right) |H_j(\alpha_2)|^5 \ll \mathcal{I}_1 + \mathcal{I}_2 , \tag{5.13}$$

where

$$\mathcal{I}_1 = (q^{2/3+\varepsilon})^t \, |H_j(\alpha_2)|^5$$

and

$$\mathcal{I}_2 = (q^{2/3+\varepsilon}) |V_i(\underline{\underline{\alpha}})|^{t-1} |H_j(\alpha_2)|^5.$$

Then

$$\iint_{\mathcal{M}} \mathcal{I}_1 \, d\underline{\underline{\alpha}} \ll \sum_{q\le P} \sum_{\substack{1\le a\le q \\ (a,b,q)=1}} \sum_{1\le b\le q} (q^{2/3+\varepsilon})^t \iint_{\mathcal{M}(q,a,b)} |H_j(\alpha_2)|^5 d\underline{\underline{\alpha}} \tag{5.14}$$

But by (2.29), (2.32) and (2.34), we have

$$\sum_{1\le a\le q} \int_{a/q-(qQ_2)^{-1}}^{a/q+(qQ_2)^{-1}} |H_j(\alpha_2)|^5 \, d\alpha_2 \le \int_0^1 |H_j(\alpha)|^5 \, d\alpha \ll P^3,$$

by (5.1), and hence from (5.14), (2.29), (2.32) and (2.34), we have

$$\iint_{\mathcal{M}} \mathcal{T}_1 \, d\underline{\alpha} \ll \sum_{q\le P} \sum_{1\le b\le q} (q^{2/3+\varepsilon})^t . P^3 \int_{b/q-(qQ_3)^{-1}}^{b/q+(qQ_3)^{-1}} d\alpha_3$$

$$\ll \sum_{q\le P} P(q^{2/3+\varepsilon})^t$$

So provided $2t/3 + 2 < t$ (that is $t > 6$), we have

$$\iint_{\mathcal{M}} \mathcal{T}_1 \, d\underline{\alpha} \ll P^t Y^{-\sigma}, \tag{5.15}$$

for some $\sigma > 0$.

Also, by (4.4) and Lemma 4.1,

$$\iint_{\mathcal{M}} \mathcal{T}_2 \, d\underline{\alpha} \ll \sum_{q\le P} \sum_{\substack{1\le a\le q \\ (a,b,q)=1}} \sum_{1\le b\le q} q(q^{-1/3+\varepsilon})^t . \mathcal{I}(t-1) \tag{5.16}$$

where

$$\mathcal{I}(T) = \iint_{\mathcal{M}(q,a,b)} |v_1(\underline{\beta})|^T |H_j(\alpha_2)|^5 \, d\underline{\alpha} .$$

But by Lemma 4.6, we have

$$\mathcal{I}(T) \ll P^T \iint_{\mathcal{M}(q,a,b)} (1 + P^3|\beta_3|)^{-T/3} |H_j(\alpha_2)|^5 \, d\underline{\alpha} .$$

Then if $T > 3$, by separating variables and making use of (5.1), we obtain

$$\sum_{1\le a\le q} \iint_{\mathcal{M}(q,a,b)} |v_1(\underline{\beta})|^T |H_j(\alpha_2)|^5 \, d\underline{\alpha}$$

$$\ll P^{T-3}\mathrm{Min}\{1, (q/Y)^{T/3-1}\} \left(\int_0^1 |H_j(\alpha_2)|^5 \, d\alpha_2 \right)$$

$$\ll P^T \mathrm{Min}\{1, (q/Y)^{T/3-1}\} . \tag{5.17}$$

Now, for $t > 6$, recalling that $W \le P^{1/2}$, we have

$$\sum_{q\le Y} q^{2-t/3+t\varepsilon} \mathrm{Min}\{1 , (q/Y)^{(t-1)/3-1}\} \ll Y^{1-(t-1)/3} \sum_{q\le Y} q^{2/3+t\varepsilon}$$

$$\ll PY^{-\sigma},$$

for some $\sigma > 0$, and

$$\sum_{Y<q\le P} q^{2-t/3+t\varepsilon} \mathrm{Min}\{ 1 , (q/Y)^{(t-1)/3-1} \} \ll P \sum_{q>Y} q^{1-t/3+t\varepsilon}$$

$$\ll PY^{-\sigma},$$

for some $\sigma > 0$. Then from (5.16) and (5.17), we have

$$\iint_{\mathcal{M}} \mathcal{T}_2 \, d\underline{\alpha} \ll P^t Y^{-\sigma}, \tag{5.18}$$

for some $\sigma > 0$.

Finally, we have

$$\iint_{\mathcal{M}} |V_1(\underline{\alpha})|^t |H_j(\alpha_2)|^5 \, d\underline{\alpha}$$
$$\ll \sum_{q \le P} \sum_{\substack{1 \le a \le q \\ (a,b,q)=1}} \sum_{1 \le b \le q} (q^{-1/3+\varepsilon})^t \iint_{\mathcal{M}(q,a,b)} |V_1(\underline{\beta})|^t |H_j(\alpha_2)|^5 \, d\underline{\alpha} .$$

As in Lemma 5.2, but using the results above, we again deduce in this case that for $t > 6$,

$$\iint_{\mathcal{M}} |V_1(\underline{\alpha})|^t |H_j(\alpha_2)|^5 \, d\underline{\alpha} \ll P^t Y^{-\sigma}, \tag{5.19}$$

for some $\sigma > 0$.

Collecting together (5.13), (5.15), (5.18) and (5.19), this completes the proof of the lemma.

Lemma 5.4. *Suppose that both c_i and d_i are non-zero. Then we have*

$$\iint_{\mathcal{U}_2^*} |f_1(\underline{\alpha})|^{14} \, d\underline{\alpha} \ll P^9.$$

Proof: The integral is the number of solutions, S, of the simultaneous diophantine equations

$$\left.\begin{aligned} c_1(x_1^3 + \ldots + x_7^3) &= c_1(y_1^3 + \ldots + y_7^3) \\ d_1(x_1^2 + \ldots + x_7^2) &= d_1(y_1^2 + \ldots + y_7^2) \end{aligned}\right\}$$

with $x_r, y_r \in \mathcal{A}(P, P^\eta)$ and $\xi_i P < x_r, y_r \le \zeta_i P$ $(r = 1, \ldots, 7)$. But by allowing x_6, x_7, y_6 and y_7 to range throughout the interval $(\xi_i P, \zeta_i P]$, and considering the underlying diophantine equations, we plainly have

$$S \le \iint_{\mathcal{U}_2^*} |f_1(\underline{\alpha})|^{10} |F_1(\underline{\alpha})|^4 \, d\underline{\alpha} .$$

We can estimate this integral by using the Hardy-Littlewood method.

Thus, by Lemma 3.2(iv) and (3.9), we have

$$\iint_{\mathfrak{m}} |f_1(\underline{\alpha})|^{10}|F_1(\underline{\alpha})|^4\, d\underline{\alpha} \ll \Big(\sup_{\alpha \in \mathfrak{m}} |F_1(\underline{\alpha})| \Big)^4 \iint_{\mathfrak{U}_2^*} |f_1(\underline{\alpha})|^{10}\, d\underline{\alpha}$$

$$\ll (P^{3/4+\varepsilon})^4 . P^{35/6 + \varepsilon}$$

$$\ll P^{9-\sigma},$$

for a $\sigma > 0$, and hence by Hölder's inequality and Lemma 5.2, we have

$$S \ll P^{9-\sigma} + \Big(\iint_{\mathfrak{U}_2^*} |f_1(\underline{\alpha})|^{14}\, d\underline{\alpha} \Big)^{5/7} \Big(\iint_{\mathfrak{M}} |F_1(\underline{\alpha})|^{14}\, d\underline{\alpha} \Big)^{2/7}$$

$$\ll P^{9-\sigma} + S^{5/7}(P^9)^{2/7},$$

giving us $S \ll P^9$, as required.

This completes the proof of the lemma.

Lemma 5.5. *For $\mathcal{B} \subseteq [0,1]^2$, let*

$$I(\mathcal{B}) = \iint_{\mathcal{B}} |\mathcal{F}(\underline{\alpha})\mathcal{G}(\underline{\alpha})\mathcal{H}(\underline{\alpha})|\, d\underline{\alpha}\ .$$

Suppose that $s \geq 14$. Then

$$I([0,1]^2) \ll P^{s-5},$$

and there is a positive number σ such that

$$I(\mathfrak{n}) \ll P^{s-5}W^{-\sigma}.$$

Proof: By Lemma 3.4, we have

$$I(\mathfrak{m}) = \iint_{\mathfrak{m}} |\mathcal{F}(\underline{\alpha})\mathcal{G}(\underline{\alpha})\mathcal{H}(\underline{\alpha})|\, d\underline{\alpha} \ll P^{s-5-\delta}.$$

Adopt the same notation (as regards $\mathcal{M}$ and Y) as in Lemma 5.2. Then by using inequality (3.1) in the same way as in the proof of Lemma 3.3, we have

$$\iint_{\mathcal{M}} |\mathcal{F}(\underline{\alpha})\mathcal{G}(\underline{\alpha})\mathcal{H}(\underline{\alpha})|\, d\underline{\alpha} \ll \iint_{\mathcal{M}} f^{h-4}g^m H^n F^4\, d\underline{\alpha}\ ,$$

for some indices I, J, K, L and integers h, m, n, where we have abbreviated $|f_I|$ to f, $|g_J|$ to g, $|H_K|$ to H, and $|F_L|$ to F, and where $h \geq s-8$, $m \leq 5$, and $n \leq 3$. By using the inequality (3.1) once

again, and by making the trivial bound $f \ll P$, we have

$$f^{h-4}g^{m}H^{n} \ll P^{s-14}(f^{10-n}H^{n} + f^{5-n}g^{5}H^{n}) ,$$

and

$$f^{10-n}H^{n} \ll f^{10} + f^{5}H^{5},$$

$$f^{5-n}g^{5}H^{n} \ll g^{5}H^{5} + f^{5}g^{5}.$$

Then

$$\int\int_{\mathcal{M}} |\mathcal{F}(\underline{\alpha})\mathcal{G}(\underline{\alpha})\mathcal{H}(\underline{\alpha})| \, d\underline{\alpha} \ll P^{s-14}\Big(\int\int_{\mathcal{M}} f^{10}F^{4} \, d\underline{\alpha} + \int\int_{\mathcal{M}} f^{5}H^{5}F^{4} \, d\underline{\alpha}$$

$$+ \int\int_{\mathcal{M}} g^{5}H^{5}F^{4} \, d\underline{\alpha} + \int\int_{\mathcal{M}} f^{5}g^{5}F^{4} \, d\underline{\alpha}\Big) \quad (5.20)$$

But by Hölder's inequality,

$$\int\int_{\mathcal{M}} f^{10}F^{4} \, d\underline{\alpha} \ll \Big(\int\int_{\mathfrak{U}_2^*} f^{14} \, d\underline{\alpha}\Big)^{5/7}\Big(\int\int_{\mathcal{M}} F^{14} \, d\underline{\alpha}\Big)^{2/7}$$

$$\ll (P^{9})^{5/7}(P^{9}Y^{-\sigma})^{2/7}$$

$$\ll P^{9}Y^{-\tau}, \text{ for some } \tau > 0, \quad (5.21)$$

by Lemmata 5.2 and 5.4. Also, by Hölder's inequality,

$$\int\int_{\mathcal{M}} f^{5}H^{5}F^{4} \, d\underline{\alpha} \ll \Big(\int\int_{\mathfrak{U}_2^*} f^{9}H^{5} \, d\underline{\alpha}\Big)^{5/9}\Big(\int\int_{\mathcal{M}} H^{5}F^{9} \, d\underline{\alpha}\Big)^{4/9}$$

$$\ll (P^{9})^{5/9}(P^{9}Y^{-\sigma})^{4/9}$$

$$\ll P^{9}Y^{-\tau}, \text{ for some } \tau > 0, \quad (5.22)$$

by Lemmata 5.1(i) and 5.3, on making the trivial bound for f. Further, by Hölder's inequality

$$\int\int_{\mathcal{M}} g^{5}H^{5}F^{4} d\underline{\alpha} \ll \Big(\int\int_{\mathfrak{U}_2^*} g^{9}H^{5} d\underline{\alpha}\Big)^{5/9}\Big(\int\int_{\mathcal{M}} H^{5}F^{9} d\underline{\alpha}\Big)^{4/9}$$

$$\ll (P^{9})^{5/9}(P^{9}Y^{-\sigma})^{4/9}$$

$$\ll P^{9}Y^{-\tau}, \text{ for some } \tau > 0, \quad (5.23)$$

by Lemmata 5.1 and 5.3, on making the trivial bound for g. Finally,

by Hölder's inequality

$$\iint_{\mathfrak{M}} f^5 g^5 F^4 d\underline{\alpha} \ll \left(\iint_{\mathfrak{U}_2^*} f^{14} d\underline{\alpha} \right)^{5/56} \left(\iint_{\mathfrak{U}_2^*} f^6 g^8 d\underline{\alpha} \right)^{5/8} \left(\iint_{\mathfrak{M}} F^{14} d\underline{\alpha} \right)^{2/7}$$

$$\ll (P^9)^{5/56} (P^9)^{5/8} (P^9 Y^{-\sigma})^{2/7}$$

$$\ll P^9 Y^{-\tau}, \text{ for some } \tau > 0, \tag{5.24}$$

by Lemmata 5.1(iii), 5.2 and 5.4.

Then by (5.20)-(5.24), we have $I(\mathfrak{M}) \ll P^{s-5} Y^{-\tau}$ for some $\tau > 0$, and hence

$$I([0,1]^2) = I(\mathfrak{m}) + I(\mathfrak{M}) \ll P^{s-5},$$

and

$$I(\mathfrak{n}) = I(\mathfrak{M} \setminus \mathfrak{N}) + I(\mathfrak{m}) \ll P^{s-5} W^{-\tau} + P^{s-5-\delta}$$

$$\ll P^{s-5} W^{-\tau}.$$

This completes the proof of the lemma.

6. DEALING WITH THE PRUNED MAJOR ARCS.

When $s \geq 14$, from (2.33), (2.35) and Lemma 5.5, we have for some $\sigma > 0$,

$$\iint_{\mathfrak{U}_2^*} \mathcal{F}(\underline{\alpha}) \mathcal{G}(\underline{\alpha}) \mathcal{H}(\underline{\alpha}) \, d\underline{\alpha} = \iint_{\mathfrak{N}} \mathcal{F}(\underline{\alpha}) \mathcal{G}(\underline{\alpha}) \mathcal{H}(\underline{\alpha}) \, d\underline{\alpha} + O(P^{s-5} W^{-\sigma}) \, . \tag{6.1}$$

We now obtain an asymptotic formula for the term

$$\iint_{\mathfrak{N}} \mathcal{F}(\underline{\alpha}) \mathcal{G}(\underline{\alpha}) \mathcal{H}(\underline{\alpha}) \, d\underline{\alpha}$$

in (6.1), and this will permit us to complete the proof of Theorem 3.1.1. The treatment we give mostly follows familiar lines, although the details are complicated by the nature and treatment of the problem.

Lemma 6.1. *Suppose that* $W \le R$. *Then*

$$\iint_{\mathfrak{N}} \mathcal{F}(\underline{\underline{\alpha}})\mathcal{G}(\underline{\underline{\alpha}})\mathcal{H}(\underline{\underline{\alpha}})\, d\underline{\underline{\alpha}} = \sum_{q \le W} S(q)J^*(q,P,W) + O(P^{s-5}W^4/\log P)\,, \tag{6.2}$$

where

$$S(q) = \sum_{\substack{a=1 \\ (a,b,q)=1}}^{q} \sum_{b=1}^{q} \left(\prod_{i=1}^{s} q^{-1} S_i(q,a,b) \right), \tag{6.3}$$

and

$$J^*(q,P,W) = \int_{-q^{-1}WP^{-3}}^{q^{-1}WP^{-3}} \int_{-q^{-1}WP^{-2}}^{q^{-1}WP^{-2}} \left(\prod_{i=1}^{m+h-4} w_i(\underline{\underline{\beta}}) \right) \left(\prod_{i=m+h-3}^{s} v_i(\underline{\underline{\beta}}) \right) d\underline{\underline{\beta}}\,. \tag{6.4}$$

Proof: Recall the definitions (4.2), (4.6) and (4.7). Suppose that $1 \le a_j \le q \le W$ $(j = 2,3)$, $(a_2, a_3, q) = 1$, and $\underline{\underline{\alpha}} \in \mathfrak{N}(q, \underline{a})$. Then by (2.34) and Lemma 4.5, for $i = 1,\dots,m+h-4$ we have

$$f_i(\underline{\underline{\alpha}}) - W_i(\underline{\underline{\alpha}}, q, \underline{a}) \ll \frac{qP}{\log P}\left(1 + q^{-1}W\right) \ll WP/\log P\,, \tag{6.5}$$

and by Lemma 4.2, for $i = m+h-3,\dots,s$ we have

$$F_i(\underline{\underline{\alpha}}) - V_i(\underline{\underline{\alpha}}, q, \underline{a}) \ll W. \tag{6.6}$$

Therefore, by (2.31), (2.34), (6.5) and (6.6), we have

$$\begin{aligned} \iint_{\mathfrak{N}} \left| \mathcal{F}(\underline{\underline{\alpha}})\mathcal{G}(\underline{\underline{\alpha}})\mathcal{H}(\underline{\underline{\alpha}}) - \left(\prod_{i=1}^{m+h-4} W_i(\underline{\underline{\alpha}}, q, \underline{a}) \right) \left(\prod_{i=m+h-3}^{s} V_i(\underline{\underline{\alpha}}, q, \underline{a}) \right) \right| d\underline{\underline{\alpha}} \\ \ll WP^{s-1} \iint_{\mathfrak{N}} d\underline{\underline{\alpha}} + \frac{WP}{\log P} \cdot P^{s-1} \iint_{\mathfrak{N}} d\underline{\underline{\alpha}} \\ \ll \frac{W^2 P^s}{\log P} (WP^{-2}) \cdot (WP^{-3}) \\ = P^{s-5} \cdot W^4/\log P\,. \end{aligned}$$

Then,

$$\begin{aligned} \iint_{\mathfrak{N}} \mathcal{F}(\underline{\underline{\alpha}})\mathcal{G}(\underline{\underline{\alpha}})\mathcal{H}(\underline{\underline{\alpha}})\, d\underline{\underline{\alpha}} = \iint_{\mathfrak{N}} \left(\prod_{i=1}^{m+h-4} W_i(\underline{\underline{\alpha}}, q, \underline{a}) \right) \left(\prod_{i=m+h-3}^{s} V_i(\underline{\underline{\alpha}}, q, \underline{a}) \right) d\underline{\underline{\alpha}} \\ + O(P^{s-5}W^4/\log P) \end{aligned}$$

The result now follows by factorising out in the usual way, using definitions (4.4) and (4.7).

This completes the proof of the lemma.

Lemma 6.2. *Let*

$$J = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\prod_{i=1}^{m+h-4} w_i(\underline{\beta}) \right) \left(\prod_{i=m+h-3}^{s} v_i(\underline{\beta}) \right) d\underline{\beta} \tag{6.7}$$

Then for $s > 8$, J *converges absolutely,* $J \ll P^{s-5}$, *and*

$$J^*(q,P,W) - J \ll (q/W)^{1/3} P^{s-5}. \tag{6.8}$$

Proof: By Lemma 4.6,

$$J \ll P^s \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1+|\beta_2|P^2)^{-n/3} (1+|\beta_3|P^3)^{-m/3} (1+|\beta_2|P^2+|\beta_3|P^3)^{-h/3} d\underline{\beta}$$

By (2.15), $n \le 3$ and $m \le 5$, so for $s > 8$ we have

$$(1+|\beta_2|P^2)^{-n/3} (1+|\beta_3|P^3)^{-m/3} (1+|\beta_2|P^2+|\beta_3|P^3)^{-h/3}$$
$$\le (1+|\beta_2|P^2)^{-4/3} (1+|\beta_3|P^3)^{-4/3}.$$

We therefore obtain

$$J \ll P^s . P^{-2} . P^{-3} = P^{s-5}, \text{ for any } s > 8.$$

So the integral converges absolutely and $J \ll P^{s-5}$. Also, by (6.4), and using the same argument as above, for $s > 8$ we have

$$J^*(q,P,W) - J \ll P^s (I_1 + I_2) ,$$

where

$$I_1 = \int_{q^{-1}WP^{-3}}^{\infty} \int_0^{\infty} (1+|\beta_2|P^2)^{-4/3} (1+|\beta_3|P^3)^{-4/3} d\underline{\beta} ,$$

and

$$I_2 = \int_0^{\infty} \int_{q^{-1}WP^{-2}}^{\infty} (1+|\beta_2|P^2)^{-4/3} (1+|\beta_3|P^3)^{-4/3} d\underline{\beta} .$$

But

$$I_1 \ll P^{-2}((q/W)^{1/3} P^{-3}) ,$$

and

$$I_2 \ll P^{-3}((q/W)^{1/3} P^{-2}) ,$$

and this completes the proof of the lemma.

Lemma 6.3. *We have*

$$J = \mathfrak{C}P^{s-5} + O(P^{s-5}/\log P) ,$$

where $\mathfrak{C}$ is a positive constant.

Proof: Making a change of variable in (4.3), we have

$$v_1(\underline{\beta}) = P\int_{\xi_1}^{\zeta_1} e(d_1(\beta_2 P^2)\gamma^2 + c_1(\beta_3 P^3)\gamma^3)\, d\gamma ,$$

and by making a change of variable in (4.6), we have

$$w_1(\underline{\beta}) = P\int_{\xi_1}^{\zeta_1} \rho\left(\frac{\log P}{\log R} + \frac{\log\gamma}{\log R} \right) e(d_1(\beta_2 P^2)\gamma^2 + c_1(\beta_3 P^3)\gamma^3)\, d\gamma .$$

By making a change of variable in (6.7), we deduce that $J = \mathfrak{C}(P).P^{s-5}$, where $\mathfrak{C}(P)$ is given by

$$\mathfrak{C}(P) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\mathcal{B}} \mathcal{P}(\underline{\gamma}).e(F(\underline{\gamma})\beta_3 + G(\underline{\gamma})\beta_2)\, d\underline{\gamma}\, d\underline{\beta} ,$$

where

$$\mathcal{P}(\underline{\gamma}) = \prod_{i=1}^{m+h-4} \rho(\ \log(P\gamma_1)/\log R\ ) ,$$

and $\underline{\gamma} = (\gamma_1,\dots,\gamma_s)$, $\mathcal{B}$ is the box defined by $\xi_i<\gamma_i<\zeta_i$ $(i = 1,\dots,s)$, and where F and G are as defined by (3.1.2).

By (2.16) (see Lemma 2.2(ii)) the equations $F(\underline{\gamma}) = G(\underline{\gamma}) = 0$ define an $(s-2)$-dimensional subspace $\mathcal{S}$ of $\mathcal{B}$ which passes through the point $(\eta_1,\dots,\eta_s)$ in the interior of $\mathcal{B}$. Further, the $(s-2)$-volume of $\mathcal{S}$ is positive. Then applying Fourier's integral formula twice to the integral $\mathfrak{C}(P)$, in the form

$$\lim_{\lambda\to\infty} \int_{-\lambda}^{\lambda} V(t).e(t\omega)\, d\omega = V(0),$$

we obtain $\mathfrak{C}(P) = \int_{\mathcal{S}} \mathcal{P}(\underline{\gamma})\, d\mathcal{S} \geq 0$. But since $R = P^{\eta}$, we have

$$\begin{aligned}\rho(\ \log(P\gamma_1)/\log R\ ) &= \rho(\eta^{-1} + \eta^{-1}\log\gamma_1/\log P)\\ &= \rho(\eta^{-1}) + O(1/\log P) .\end{aligned}$$

Thus, on noting that $\int_{\mathcal{S}} d\mathcal{S} = \mu(\mathcal{S}) > 0$, we deduce that

$$\mathfrak{C}(P) = (\rho(\eta^{-1}))^{m+h-4} \int_{\mathcal{S}} d\mathcal{S} + O(1/\log P)$$

$$= \mathfrak{C} + O(1/\log P) ,$$

with $\mathfrak{C}$ a positive constant.

This completes the proof of the lemma.

Lemma 6.4. *Suppose that* $(a_2, a_3, q) = (b_2, b_3, r) = (q, r) = 1$. *Then,*

$$S_1(\, qr,\ a_2 r + b_2 q,\ a_3 r + b_3 q\,) = S_1(q, a_2, a_3)\ S_1(r, b_2, b_3) .$$

Proof: This is essentially Vaughan [1981b] Lemma 2.10, the presence of the c_i and d_i making no difference to the proof.

Lemma 6.5. *The function* $S(q)$ *is multiplicative.*

Proof: This is essentially Vaughan [1981b] Lemma 2.11, combined with Lemma 6.4.

Lemma 6.6. *Suppose that* $s \geq 12$. *Then*

$$\sum_{q \leq W} q^{1/3} |S(q)| \ll W^{\varepsilon}.$$

Proof: By (6.3) and Lemma 4.1, we have

$$S(q) \ll \sum_{\substack{a=1 \\ (a,b,q)=1}}^{q} \sum_{b=1}^{q} (q^{-1/3+\varepsilon})^s (a,q)^{m/3} (b,q)^{n/3},$$

and further, by (2.15) we have $m \leq 5$ and $n \leq 3$. But for $t \geq 1$, by using properties of the divisor function, we have

$$\sum_{c=1}^{q} (c,q)^t \leq \sum_{d|q} (q/d) d^t \ll q^{t+\varepsilon},$$

and for $t < 1$, we have

$$\sum_{c=1}^{q} (c,q)^t \leq \sum_{d|q} (q/d) d^t \ll q^{1+\varepsilon}.$$

Hence

$$S(q) \ll (q^{-1/3+\varepsilon})^s \Big[\sum_{a=1}^{q} (a,q)^{5/3} \Big] . \Big[\sum_{b=1}^{q} (b,q) \Big]$$

$$\ll q^{(8-s)/3+2s\varepsilon} .$$

Then for $s \geq 12$, we obtain for any $\varepsilon > 0$,

$$S(q) \ll q^{-4/3+\varepsilon}, \tag{6.9}$$

and by using Lemma 6.5 to factorise into prime powers,

$$\sum_{q \leq W} q^{1/3} |S(q)| \leq \prod_{p \leq W} \Big(1 + \sum_{l=1}^{\infty} (p^l)^{1/3} |S(p^l)| \Big)$$

$$\leq \prod_{p \leq W} (1 + Cp^{\varepsilon}/p)$$

for some constant $C = C(s)$. But then

$$\sum_{q \leq W} q^{1/3} |S(q)| \ll W^{2\varepsilon} .$$

This completes the proof of the lemma.

Lemma 6.7. *Suppose that $s \geq 12$. Then for some $\sigma > 0$ we have*

$$\sum_{q \leq W} S(q) J^*(q,P,W) = J.\mathfrak{S}(W) + O(W^{-\sigma} P^{s-5}) ,$$

where

$$\mathfrak{S}(W) = \sum_{q \leq W} S(q) .$$

Further, $\mathfrak{S} = \sum_{q=1}^{\infty} S(q)$ *converges absolutely, and* $\mathfrak{S}(W) - \mathfrak{S} \ll W^{-\tau}$ *for some* $\tau > 0$.

Proof: By Lemmata 6.2 and 6.6, we have

$$\sum_{q \leq W} S(q) (J^*(q,P,W) - J) \ll P^{s-5} W^{-1/3} \sum_{q \leq W} q^{1/3} |S(q)|$$

$$\ll P^{s-5} W^{-\sigma},$$

for some $\sigma > 0$. Also, if $s \geq 12$, we have by (6.9),

$$S(q) \ll q^{-4/3+\varepsilon}.$$

Hence $\mathfrak{S} = \sum_{q=1}^{\infty} S(q)$ is absolutely convergent, and $\mathfrak{S} \ll 1$. Further,

$$\mathfrak{S}(W) - \mathfrak{S} \ll \sum_{q > W} q^{-4/3+\varepsilon} \ll W^{-\tau},$$

for some $\tau > 0$, and this completes the proof of the lemma.

Combining the conclusions of Lemmata 6.1, 6.2, 6.3 and 6.7 with (6.1), we deduce that for $s \geq 14$, provided that $W \leq R$, we have

$$R(P) = \iint_{\mathfrak{U}_2^*} \mathfrak{F}(\underline{\alpha})\mathfrak{G}(\underline{\alpha})\mathfrak{H}(\underline{\alpha})\, d\underline{\alpha} = \mathfrak{S}\mathfrak{C}P^{s-5} + O\left(P^{s-5}(W^4/\log P + W^{-\sigma})\right),$$

where σ is a positive constant depending only on $c_1,\dots,c_s$, $d_1,\dots,d_s$ and $\eta_1,\dots,\eta_s$. By taking W to be a suitable power of $\log P$, we obtain

$$R(P) = \mathfrak{S}\mathfrak{C}P^{s-5} + o(P^{s-5}),$$

and hence Theorem 3.1.1 will follow provided we can show that $\mathfrak{S} > 0$. For then $R(P) \gg P^{s-5} \longrightarrow \infty$ as $P \longrightarrow \infty$, and hence there must be a non-trivial rational solution to the equations (3.1.2).

For each prime p, define formally

$$T(p) = \sum_{h=0}^{\infty} S(p^h).$$

Lemma 6.8. *Suppose that $s \geq 12$. Then $T(p)$ converges absolutely, so does $\prod_p T(p)$, and $\mathfrak{S} = \prod_p T(p)$. Further, there is a positive number p_0, depending at most on $c_1,\dots,c_s$, $d_1,\dots,d_s$, such that $\frac{1}{2} < \prod_{p \geq p_0} T(p) < \frac{3}{2}$.*

Proof: For $s \geq 12$, by (6.9) we have $S(p^h) \ll (p^{-4/3+\varepsilon})^h$, and therefore $\sum_{h=0}^{\infty} S(p^h)$ is absolutely convergent. So $T(p)$ converges absolutely. The rest of the lemma is essentially Vaughan [1981b] Theorem 2.4.

Lemma 6.9. *For each natural number q, we have*

$$\sum_{d \mid q} S(d) = q^{2-s} M_n(q).$$

Proof: Recalling (2.7), this is only a slight modification of Vaughan [1981b], Lemma 2.12.

Given a rational prime p, put $q = p^t$ in Lemma 6.9. Then

$$T(p) = \lim_{t \to \infty} p^{t(2-s)} M_n(p^t)$$

provided that either this limit exists, or that in Lemma 6.8 exists. Then by (2.17), $T(p) \geq p^{-u(p).(s-2)} > 0$, and so by Lemma 6.8 we have $\mathfrak{G} > 0$. In view of the above remarks, this completes the proof of Theorem 3.1.1.

PART III.

ON WARING'S PROBLEM.

Modified versions of Chapters 5 and 6 have been submitted for publication, respectively to the London Mathematical Society and Annals of Mathematics. The former is a paper joint with Professor Vaughan.

CHAPTER 5.

ON WARING'S PROBLEM: SOME REFINEMENTS.

1. INTRODUCTION.

Let $G(k)$ denote the least number s such that every sufficiently large natural number is the sum of at most s kth powers of natural numbers. In this chapter we introduce a number of refinements to the methods described in Vaughan [1989a,c]. In certain circumstances they permit an upper bound $G(k) \leq H(k)$ to be replaced by $G(k) \leq H(k) - 1$, with more substantial savings being possible for $k > 12$. Thus we are able to establish:

Theorem 1.1. *We have* $G(5) \leq 18$ *and* $G(6) \leq 28$.

Table 1.1.

k	10	11	12	13	14	15	16	17	18	19	20
$F(k)$	92	108	124	139	153	168	183	198	213	228	243

Theorem 1.2. *When* $10 \leq k \leq 20$ *we have* $G(k) \leq F(k)$, *where* $F(k)$ *is given by Table* 1.1.

These bounds may be compared with the upper bounds $G(5) \leq 19$, $G(6) \leq 29$, $G(10) \leq 93$, $G(11) \leq 109$, $G(12) \leq 125$, $G(13) \leq 141$, $G(14) \leq 156$, $G(15) \leq 171$, $G(16) \leq 187$, $G(17) \leq 202$, $G(18) \leq 217$, $G(19) \leq 232$, $G(20) \leq 248$ of Theorems 1.1 and 1.7 of Vaughan [1989a]. We note that Brüdern [1990] has also established $G(5) \leq 18$ by a refinement of the arguments of Vaughan [1989a,c].

We define

$$\mathcal{A}(P,R) = \{ x \le P : p \text{ prime}, p|x \text{ implies } p \le R \} \tag{1.1}$$

and when k is a positive integer, let $S_s(P,R)$ denote the number of solutions to the equation

$$x_1^k + \ldots + x_s^k = y_1^k + \ldots + y_s^k \tag{1.2}$$

with $x_i, y_i \in \mathcal{A}(P,R)$. The proof of Theorem 1.1 is motivated by the observation that by simply combining the methods of Vaughan [1989a,c], one can show that for each $\varepsilon > 0$, there is an $\eta > 0$ such that $S_8(P,P^\eta) \ll_\varepsilon P^{13+\varepsilon}$ $(k = 5)$ and $S_{14}(P,P^\eta) \ll_\varepsilon P^{22+\varepsilon}$ $(k = 6)$. Moreover the dominant contribution in these estimates arises from the "major arcs". We therefore develop an argument based on just one more iteration of the method, but relative to minor arcs only, similar to that contained in Vaughan [1989b]. This is embodied in the following theorem.

Theorem 1.3. *Let* $k \ge 5$, $\theta = (2^{k-2}-1)/(k.2^{k-2}+1)$, *and* $t \ge 2k-2$. *Suppose that* λ *and* μ *are positive real numbers satisfying*

$$\mu > 2t-2-k, \quad \lambda > 2t-k , \tag{1.3}$$

$$\mu(1-\theta) + 1 + (2t-2)\theta = \lambda , \tag{1.4}$$

and

$$\lambda(1-\theta) + 1 + 2t\theta < 2t+2-k , \tag{1.5}$$

and that for each $\varepsilon > 0$, *there is an* $\eta > 0$ *such that*

$$\left.\begin{aligned} S_{t-1}(P,P^\eta) \ll_\varepsilon P^{\mu+\varepsilon}, \ S_t(P,P^\eta) \ll_\varepsilon P^{\lambda+\varepsilon}, \\ \text{and } S_{t+1}(P,P^\eta) \ll_\varepsilon P^{2t+2-k+\varepsilon}. \end{aligned}\right\} \tag{1.6}$$

Then $G(k) \le 2t+2$, *provided only that when* k *is a power of 2 one has* $2t+2 \ge 4k$.

We note that the additional condition on t when k is a power of 2 is required only to guarantee the positivity of the "singular

series", and could be relaxed under a suitable local solubility hypothesis.

With regard to Theorem 1.2, an inspection of the proof of Theorem 1.7 of Vaughan [1989a] reveals that one of the kth powers in the representation is present solely to facilitate the estimation of the contribution from the "major arcs". To prove Theorem 1.2, therefore, we develop a new treatment of the major arcs which makes the extra kth power superfluous.

To describe the main conclusion of this analysis we shall require some notation. Let

$$1/2 \le \phi \le k/(k+1) ,$$

and

$$\mathfrak{m} = \{ \alpha : \text{if } q \in \mathbb{N} \text{ and } \|q\alpha\| \le P^{\phi-k} \text{ then } q > P^{\phi} \} ,$$

where $\|x\|$ denotes $\min_{y \in \mathbb{Z}} |x-y|$. We suppose that the positive number $\sigma = \sigma(k)$ has the property that for each $\varepsilon > 0$, there is a set $\mathfrak{C} = \mathfrak{C}_{\varepsilon,k} \subset [1,P] \cap \mathbb{Z}$ such that

$$\text{card } \mathfrak{C} \gg_{\varepsilon} P/\log P$$

and the exponential sum

$$h(\alpha) = \sum_{x \in \mathfrak{C}} e(\alpha x^k)$$

satisfies

$$\sup_{\mathfrak{m}} |h(\alpha)| \ll_{\varepsilon} P^{1-\sigma+\varepsilon} .$$

Theorem 1.4. *Suppose that* $k \ge 3$, $t \ge 2k+1$,

$$\lambda + k < 2t + \sigma u ,$$

and that for each $\varepsilon > 0$ *there is an* $\eta > 0$ *such that*

$$S_t(P,P^{\eta}) \ll_{\varepsilon} P^{\lambda+\varepsilon}.$$

Then

$$G(k) \le 2t+u .$$

In establishing Theorem 1.2 through the use of Theorem 1.4, a further improvement over the methods described in Vaughan [1989a,c] is wrought by the use of estimates stemming from Vinogradov's mean value theorem. When $k > 12$, the use of such estimates proves more effective than Weyl differencing in all but the initial phases of the iteration process.

We give the proof of Theorem 1.3 in §§2-5, and in §6 the special cases $k = 5$, 6 are considered, thereby establishing Theorem 1.1. We then provide two "major arc" estimates in §§7 and 8, these enabling us to prove Theorem 1.4 in §9. In §10 we describe the estimates arising from Vinogradov's mean value theorem. Finally, in §11 we combine the approaches of Vaughan [1989a,c] with the conclusions of §9 and §10 to deduce Theorem 1.2.

2. PRELIMINARIES TO THE PROOF OF THEOREM 1.3.

We shall prove Theorem 1.3 by using a variant of the Hardy-Littlewood method. We suppose that ε is a sufficiently small positive number, that $\eta = \eta(\varepsilon,k)$ is a small positive number depending at most on ε and k, that $\tau = \tau(\varepsilon,\eta,k)$ is a small positive number depending at most on ε, η and k, and that n is sufficiently large in terms of ε, η and τ.

Let

$$P = n^{1/k}, \quad \sigma = 2^{1-k}, \quad \theta = (1-2\sigma)/(k+2\sigma), \quad \text{and } R = P^{\eta}. \tag{2.1}$$

Also, let $M_1,\dots,M_{2t}$ be real numbers satisfying

$$P^{\theta} \le M_s \le P^{\theta+\tau}, \tag{2.2}$$

and write

$$Q_s = PM_s^{-1}. \tag{2.3}$$

For t, λ and μ satisfying the hypotheses of Theorem 1.3, we consider the number $r(n;\underline{\underline{M}}) = r(n;M_1,\dots,M_{2t})$ of solutions to the equation

$$x_1^k + x_2^k + p_1^k y_1^k + p_2^k y_2^k + \dots + p_{2t}^k y_{2t}^k = n \ , \tag{2.4}$$

with the primes p_s satisfying

$$p_s \equiv -1 \pmod{k}, \ M_s < p_s \le 2M_s \ , \tag{2.5}$$

and with

$$x_r \le P \ \text{ for } r = 1,2 \ , \quad y_s \in \mathcal{A}(Q_s,R) \ \text{ for } s = 1,\dots,2t \ . \tag{2.6}$$

We shall show that

$$\sum_{M_1} \dots \sum_{M_{2t}} r(n;\underline{\underline{M}}) \gg n^{(2t+2)/k - 1} \ , \tag{2.7}$$

where the multiple sum is over all choices of M_s of the form

$$M_s = 2^u P^\theta \ , \tag{2.8}$$

and satisfying (2.2). Since $p_s > R$, each solution of (2.4) gives rise to a unique representation of n as the sum of $2t+2$ kth powers of positive integers in the sense that the ordered $(2t+2)$-tuple $x_1, x_2, p_1y_1, \dots, p_{2t}y_{2t}$ is unique. Hence the verification of (2.7) is sufficient to establish Theorem 1.3.

We henceforth assume that the M_s are of the form (2.8). Let

$$e(\beta) = e^{2\pi i\beta}, \tag{2.9}$$

$$F(\alpha) = \sum_{x \le P} e(\alpha x^k) \ , \tag{2.10}$$

$$g_s(\alpha) = \sum_{x \in \mathcal{A}(Q_s,R)} e(\alpha x^k) \ , \tag{2.11}$$

$$H(\alpha;L) = \sum_{x \le L} e(\alpha x^k) \ . \tag{2.12}$$

Then

$$r(n;\underline{\underline{M}}) = \int_0^1 \mathcal{F}(\alpha) e(-\alpha n) \, d\alpha \tag{2.13}$$

where

$$\mathcal{F}(\alpha) = F(\alpha)^2 \prod_{s=1}^{2t} \Big(\sum_{p_s} g_s(\alpha p_s^k) \Big) \ , \tag{2.14}$$

and the p_s satisfy (2.5).

We shall require the inequalities contained in the following lemma.

Lemma 2.1. *We have*

(*i*) $\lambda + 2(1-\sigma)(1-\theta) + \theta < 2t + 2 - k$,

(*ii*) $\lambda + 1 - \sigma + k\theta < 2t + 2 - k$,

(*iii*) $\lambda + 1 + (1/2 - 1/k)(1-\sigma+k\theta) < 2t + 2 - k$,

(*iv*) $\lambda + 4(1-\sigma) < 2t + 4 - k$.

Proof: Each inequality is deduced by repeated use of (1.5).

(i) We have

$$\begin{aligned}\lambda + 2(1-\sigma)(1-\theta) + \theta &= \lambda(1-\theta) + 2t\theta + (\lambda-2t)\theta + 2(1-\sigma)(1-\theta) + \theta \\ &< 2t + 1 - k + (1-k)\theta/(1-\theta) + 2(1-\sigma)(1-\theta) + \theta \\ &\le 2t + 1 - k - (1-2\sigma)/(1+\sigma) + 2(1-\sigma)(1-\theta) + \theta \\ &< 2t + 2 - k + \sigma - 2(1-\sigma)\theta + \theta\end{aligned}$$

and $\sigma - (1-2\sigma)\theta < \sigma - 1/(2k) < 0$.

(ii) Similarly, we have

$$\begin{aligned}\lambda + 1 - \sigma + k\theta &< 2t + 2 - k + (1-k\theta)\theta/(1-\theta) - \sigma \\ &< 2t + 2 - k + 4\sigma\theta - \sigma\end{aligned}$$

and $4\theta < 1$ for $k \ge 5$.

(iii) In a manner similar to (i), we have

$$\begin{aligned}\lambda + 1 + (1/2 - 1/k)(1-\sigma+k\theta) &< 2t + 2 - k + (1-k)\theta/(1-\theta) \\ &\quad + (1/2 - 1/k)(1 - \sigma + k\theta) \\ &< 2t + 2 - k - 1 + 3\sigma \\ &\quad + (2-\sigma)(1/2 - 1/k) \\ &< 2t + 2 - k - 2/k + 3\sigma\end{aligned}$$

and $3\sigma < 2/k$.

(iv) In a manner similar to (i), we have

$$\lambda + 4(1-\sigma) < 2t + 5 - k - 4\sigma + (1-k)\theta/(1-\theta)\ ,$$

and

$$1 - 4\sigma - \frac{(k-1)(1-2\sigma)}{k - 1 + 4\sigma} = -2\sigma + \frac{4\sigma(1-2\sigma)}{k - 1 + 4\sigma}$$
$$< -2\sigma + \sigma(1 - 2\sigma) < 0$$

since $k - 1 + 4\sigma > 4$.

This completes the proof of the lemma.

3. AUXILIARY INTEGRALS.

The following lemma is particularly useful, and may be compared with the inequality $S_t(P, P^\eta) \ll P^{\lambda+\varepsilon}$.

Lemma 3.1. *Let*

$$I_1 = \int_0^1 |\sum_p g_s(\alpha p^k)|^{2t} \, d\alpha \ ,$$

where the summation is over all primes p satisfying

$$p \equiv -1 \pmod{k}, \text{ and } M_s < p \le 2M_s \ . \tag{3.1}$$

Then we have

$$I_1 \ll P^{\lambda+k\tau+\varepsilon}.$$

Proof: For the sake of convenience, in the proof of this lemma we write M for M_s, Q for Q_s, and g for g_s. Also, let $H = PM^{-k}$.

Let u be any positive integer with $1 \le u \le t(2P)^k$, and consider the number of solutions, $I_2(u)$, to the equation

$$p_1^k x_1^k + \ldots + p_t^k x_t^k = u \ , \tag{3.2}$$

with the p_r satisfying (3.1), and $x_r \in \mathcal{A}(Q, R)$ for $r = 1, \ldots, t$.

Let $R_1(u)$ denote the number of solutions of (3.2) for which there is an $i \in \{1, \ldots, t\}$ such that for each $j \in \{1, \ldots, t\}$ with $j \ne i$, we have $p_i \ne p_j$.

Let $R_2(u)$ denote the number of solutions of (3.2) such that for each $i \in \{1,\dots,t\}$, there is a $j \neq i$ in $\{1,\dots,t\}$ such that $p_i = p_j$.

Then plainly

$$I_2(u) = R_1(u) + R_2(u) ,$$

and

$$I_2(u)^2 \ll R_1(u)^2 + R_2(u)^2 .$$

But then, by considering the underlying diophantine equations, we deduce that

$$I_1 = \sum_u I_2(u)^2 \ll I_3 + I_4 , \tag{3.3}$$

where

$$I_3 = \sum_u R_1(u)^2 \text{ and } I_4 = \sum_u R_2(u)^2.$$

We first consider I_4. We have $I_4 \ll \sum_{2s\le t} I_5(s)$, where $I_5(s)$ is the number of solutions to the equation

$$p_1^k(x_{1,1}^k+\dots+x_{1,r_1}^k) + \dots + p_s^k(x_{s,1}^k+\dots+x_{s,r_s}^k)$$
$$= p_1'^k(x_{1,1}'^k+\dots+x_{1,r_1}'^k) + \dots + p_s'^k(x_{s,1}'^k+\dots+x_{s,r_s}'^k)$$

with $r_1,\dots,r_s$ integers satisfying $r_i \ge 2$ for each i, and $r_1 + \dots + r_s = t$, and with the p_r and p_r' satisfying (3.1), and $x_{i,j}$, $x_{i,j}' \in \mathcal{A}(Q,R)$ for $i = 1,\dots,s$ and $j = 1,\dots,r_i$. Then by using the trivial inequality

$$|z_1\dots z_m| \le |z_1|^m + \dots + |z_m|^m , \tag{3.4}$$

we deduce, by considering the corresponding integral representation of the number of solutions to this last equation, that

$$I_5(s) \ll \sum_{p_1}\dots\sum_{p_s}\sum_{p_1}'\dots\sum_{p_s}' \int_0^1 \left(\sum_{i=1}^{s}\left(|g(\alpha p_i^k)|^{2t} + |g(\alpha p_i'^k)|^{2t}\right)\right) d\alpha$$
$$\ll M^{2s}\int_0^1 |g(\alpha)|^{2t} d\alpha .$$

Thus, by considering the underlying diophantine equation, and recalling (1.2) and (1.6), we have

$$\sum_{2s\le t} I_5(s) \ll M^t S_t(Q,R) \ll M^t Q^{\lambda+\varepsilon} \ll P^{\lambda+\varepsilon}, \tag{3.5}$$

since by (1.3) we have $\lambda > t$.

We now consider I_3. We have $I_3 \ll I_6$, where I_6 is the number of solutions to the equation

$$x^k + p_1^k y_1^k + \ldots + p_{t-1}^k y_{t-1}^k = x'^k + p_1'^k y_1'^k + \ldots + p_{t-1}'^k y_{t-1}'^k,$$

with the p_r satisfying (3.1) and

$$(x, p_1 \ldots p_{t-1}) = (x', p_1' \ldots p_{t-1}') = 1,$$

and $x, x' \le 2P$,

$$y_s, y_s' \in \mathcal{A}(Q,R) \text{ for } s = 1, \ldots, t-1.$$

Let

$$F(\alpha; m; L) = \sum_{\substack{x\le L \\ (x,m)=1}} e(\alpha x^k). \tag{3.6}$$

Then

$$I_6 \le \int_0^1 \Big(\sum_{p_1} \ldots \sum_{p_{t-1}} |F(\alpha; p_1 \ldots p_{t-1}; 2P)| \cdot \prod_{s=1}^{t-1} |g(\alpha p_s^k)| \Big)^2 d\alpha .$$

By Cauchy's inequality and inequality (3.4), the integrand is

$$\ll M^{t-1} \sum_{p_1} \ldots \sum_{p_{t-1}} |F(\alpha; p_1 \ldots p_{t-1}; 2P)|^2 \prod_{s=1}^{t-1} |g(\alpha p_s^k)|^2$$

$$\ll M^{t-1} \sum_{p_1} \ldots \sum_{p_{t-1}} |F(\alpha; p_1 \ldots p_{t-1}; 2P)|^2 \sum_{s=1}^{t-1} |g(\alpha p_s^k)|^{2t-2}.$$

Then, by considering the underlying diophantine equations, we have

$$I_6 \ll M^{2t-3} \int_0^1 \sum_p |F(\alpha; p; 2P)|^2 |g(\alpha p^k)|^{2t-2} \, d\alpha$$

$$\ll M^{2t-3} I_7 , \tag{3.7}$$

where I_7 is the number of solutions to the equation

$$x^k - y^k = p^k (x_1^k - y_1^k + \ldots + x_{t-1}^k - y_{t-1}^k), \tag{3.8}$$

with p satisfying (3.1),

$$1 \le x, y \le 2P, \ (xy, p) = 1,$$

and

$$x_j, y_j \in \mathcal{A}(Q,R) \quad (j = 1,\dots,t-1) .$$

By (1.6), the solutions with $x = y$ contribute

$$\ll PM.Q^{\mu+\varepsilon} .$$

From (3.8) we have $x^k \equiv y^k \pmod{p^k}$, so that since $p \equiv -1 \pmod{k}$ and $(xy,p) = 1$, we have

$$\begin{cases} x^2 \equiv y^2 \pmod{p^k} & \text{when } k \text{ is even} \\ x \equiv y \pmod{p^k} & \text{when } k \text{ is odd.} \end{cases}$$

When k is even we therefore have $x \equiv \pm y \pmod{p^k}$. If $x \equiv y \pmod{p^k}$, then we may write $h = (x-y)p^{-k}$ and $z = x+y$. If $x \equiv -y \pmod{p^k}$, then we may write $h = (x+y)p^{-k}$ and $z = x-y$. When k is odd only the first of these cases arises.

In this way we see that the number of solutions with $x \neq y$ is at most $4I_8$, where I_8 is the number of solutions to the equation

$$\Phi(z,h,p) = 2^k(x_1^k - y_1^k + \dots + x_{t-1}^k - y_{t-1}^k) \tag{3.9}$$

with $h \leq 4PM^{-k}$, $z \leq 4P$ and p, x_j, y_j as before, and where we have put

$$\Phi(z,h,p) = p^{-k}((z+hp^k)^k - (z-hp^k)^k) . \tag{3.10}$$

Let

$$F_1(\alpha) = \sum_{M<m\leq 2M} \sum_{h\leq 4H} \sum_{z\leq 4P} e(\alpha\Phi(z,h,m)) .$$

Then by considering the underlying diophantine equations, we have

$$I_8 \ll \int_0^1 F_1(\alpha).|g(2^k\alpha)|^{2t-2}\, d\alpha .$$

Now F_1 is as in Vaughan [1989a], definition (2.35), but with P replaced by $2P$, H by $4H$, and R by 2. Therefore the argument of Vaughan [1989a], Lemma 3.6 together with (1.6) now gives

$$I_7 \ll (PM + PHM(PM)^{-2\sigma})Q^{\mu+\varepsilon} + P^{1+\varepsilon}HMQ^{-k/t}Q^{(1-1/t)\lambda+\varepsilon}. \tag{3.11}$$

But from the definition of θ, we have

$$PHM(PM)^{-2\sigma} \leq PM , \tag{3.12}$$

and from (1.3) and (1.4), we have

$$PM^{2t-2}Q^{\mu+\varepsilon} \ll P^{\lambda+k\tau+\varepsilon} . \tag{3.13}$$

Also,

$$(PHM.Q^{-k/t+(1-1/t)\lambda})M^{2t-3} = (P^{(2t-k-\lambda)/t}\ M^{(1-1/t)(2t-k-\lambda)})P^{\lambda}$$

$$\ll P^{\lambda} \tag{3.14}$$

since by (1.3), we have $2t-k-\lambda < 0$. We conclude from (3.7), and (3.11)-(3.14) that $I_6 \ll P^{\lambda+k\tau+\varepsilon}$, and hence from (3.3) and (3.5) that $I_1 \ll P^{\lambda+k\tau+\varepsilon}$.

This completes the proof of the lemma.

Lemma 3.2. *Let*

$$J_1 = \int_0^1 \Big(\sum_{p} |g_s(\alpha p^k)| \Big)^{2t} d\alpha \ ,$$

where the summation is over all primes p satisfying

$$p \equiv -1 \pmod{k} \text{ and } M_s < p \le 2M_s \ .$$

Then we have

$$J_1 \ll M_s P^{\lambda+k\tau+\varepsilon}.$$

Proof: Adopt the same notation as in the proof of Lemma 3.1. By Cauchy's inequality, we have

$$J_1 \ll M^t J_2 \ , \tag{3.15}$$

where

$$J_2 = \int_0^1 \Big(\sum_{p} |g(\alpha p^k)|^2 \Big)^{t} d\alpha \ .$$

We now observe that J_2 is the number of solutions of the equation

$$\sum_{i=1}^{t} p_i^k (x_i^k - y_i^k) = 0 \tag{3.16}$$

with the p_i satisfying (3.1), and $x_i, y_i \in \mathcal{A}(Q,R)$.

Let J_3 denote the number of solutions of (3.16) for which there is an $i \in \{1,\dots,t\}$ such that for each $j \in \{1,\dots,t\}$ with $j \neq i$, we have $p_i \neq p_j$. Also, let J_4 denote the number of solutions of (3.16) such that for each $i \in \{1,\dots,t\}$, there is a $j \neq i$ in $\{1,\dots,t\}$ such that $p_i = p_j$. Then plainly

$$J_2 = J_3 + J_4 \ . \tag{3.17}$$

By an argument akin to that used to estimate I_4 in the proof of Lemma 3.1, we have

$$J_4 \ll M^{t/2}\int_0^1 |g(\alpha)|^{2t}d\alpha \ll M^{t/2}Q^{\lambda+\varepsilon} \leq P^{\lambda+\varepsilon}M^{k-3t/2},$$

since $\lambda > 2t-k$, and hence, as $t \geq 2k-2$, we deduce that

$$J_4 \ll M^{1-t}P^{\lambda+\varepsilon}. \tag{3.18}$$

Also, we have $J_3 \ll J_5$, where J_5 is the number of solutions of the equation

$$x^k-y^k + p_1^k(x_1^k-y_1^k) + \ldots + p_{t-1}^k(x_{t-1}^k-y_{t-1}^k) = 0$$

with the p_r satisfying (3.1) and

$$(xy, p_1\ldots p_{t-1}) = 1,\ x,y \leq 2P \text{ and } x_i, y_i \in \mathcal{A}(Q,R) \text{ for } i = 1,\ldots,t-1.$$

Then by an argument similar to that applied to I_6 in the proof of Lemma 3.1, we find that

$$J_5 \ll M^{t-2}I_7,$$

where I_7 is as in the proof of Lemma 3.1. Thus, by (3.11)-(3.14), we find that

$$M^{t-1}J_3 \ll P^{\lambda+k\tau+\varepsilon}. \tag{3.19}$$

Collecting together (3.15), (3.17), (3.18) and (3.19) completes the proof of the lemma.

4. THE MINOR ARCS.

We now introduce a suitable choice for the minor arcs $\mathfrak{m}$ and proceed to estimate

$$I(\mathfrak{m}) = \int_{\mathfrak{m}} |\mathcal{F}(\alpha)|\, d\alpha\ .$$

Our plan is to follow the argument of Vaughan [1989b] §5, and to attempt to convert the integral over $\mathfrak{m}$ to one in which we may apply arguments similar to those used in the proof of Lemma 3.1. However,

we have less room to spare than in Vaughan [1989b], and so in general a more precise argument is required. We shall give most of the details for the sake of completeness.

We write $C_k = k^2 3^{k^2}$. For p a rational prime, write $d = p^k$, and define

$$G(\alpha;p) = \sum_{h \le P/d} \sum_{\substack{hd < z \le 2P-hd \\ z \equiv h \pmod 2}} e(2^{-k}\alpha\Phi(z,h,p))$$

where $\Phi(z,h,p)$ is defined as in (3.10). It transpires that $G(\alpha;p)$ plays a fundamental rôle in the treatment of the minor arcs. We provide estimates for this sum for later use.

Let

$$M = P^{\theta+\tau}, \quad Q = PM^{-1}, \tag{4.1}$$

and when M_s satisfies (2.2), write

$$H_s = PM_s^{-k}. \tag{4.2}$$

Our first step is to simplify the ranges of summation in the expression for $G(\alpha;p)$. This we achieve by using a variant of the argument of §5.2 of Vaughan [1981b].

Lemma 4.1. *Let a and d be natural numbers satisfying $1 \le a \le d$, let U and V be real numbers satisfying $1 \le Ud+a \le V$, and suppose that $a_v \in \mathbb{C}$ for $1 \le v \le V$. Then*

$$\Big| \sum_{u=1}^{U} a_{ud+a} \Big| \ll d \log V \sup_{\alpha \in [0,1]} \Big| \sum_{v=1}^{V} a_v e(\alpha v) \Big| .$$

Proof: We have

$$\begin{aligned} \sum_{u=1}^{U} a_{ud+a} &= \int_0^1 \Big(\sum_{v=1}^{V} a_v e(\alpha v) \Big) \sum_{u=1}^{U} e(-\alpha(ud+a))\, d\alpha \\ &\ll \int_0^1 \Big| \sum_{v=1}^{V} a_v e(\alpha v) \Big| \min\{ U , \|\alpha d\|^{-1} \}\, d\alpha , \end{aligned}$$

and the lemma now follows.

By applying Lemma 4.1, we deduce that

$$\sum_{p_s} G(\alpha; p_s) \ll P^{\varepsilon} F_1^*(\alpha) , \tag{4.3}$$

where

$$F_1^*(\alpha) = \sum_{M_s < m \le 2M_s} \sum_{h \le H_s} \sup_{\beta \in [0,1]} |T(\alpha, \beta, h, m)| , \tag{4.4}$$

and

$$T(\alpha, \beta, h, m) = \sum_{z \le 2P} e(2^{-k}\alpha\Phi(z, h, m) + \beta z) .$$

Lemma 4.2. *Suppose that α is a real number in $(0,1]$ with the property that whenever $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and $|\alpha q - a| \le C_k^{-1} P Q_s^{-k}$, then one has $q > 2^{-k}P$. Then we have*

$$|\sum_{p_s} G(\alpha; p_s)| \ll (PM_s)^{1-2\sigma+\varepsilon} H_s ,$$

where the summation is over p_s satisfying (2.5).

Proof: It is seen rather easily that we may treat $F_1^*(\alpha)$ by using virtually the same refined differencing argument as is applied in Vaughan [1989a], §§2,3 to the exponential sum

$$F_1(\alpha) = \sum_{M < m \le MR} \sum_{h \le H} \sum_{z \le 2P} e(\alpha\Phi(z, h, m)) .$$

In order to convince ourselves of this, we shall require some notation. Let Δ_1 denote the forward difference operator

$$\Delta_1(f(x); h) = f(x+h) - f(x) ,$$

and define Δ_j recursively by

$$\Delta_{j+1}(f(x); h_1, \dots, h_{j+1}) = \Delta_1(\Delta_j(f(x); h_1, \dots, h_j); h_{j+1}) .$$

We define

$$\Psi_j = \Psi_j(z; h, h_2, \dots, h_j; m) = m^{-k}\Delta_j(f(z); 2hm^k, h_2, \dots, h_j)$$

where $f(z) = (z - hm^k)^k$.

Now define the exponential sum $F_2^*(\alpha)$ by

$$F_2^*(\alpha) = \sum_{M_s < m \le 2M_s} \sum_{h \le H_s} \sum_{h_2 \le 2P} \sup_{\beta \in [0,1]} |T_2| ,$$

where

$$T_2 = T_2(\alpha,\beta,h,h_2,m) = \sum_{0<z\le 2P-h_2} e(2^{-k}\alpha\Psi_2 + \beta h_2) ,$$

and define $F_j^*(\alpha)$ for $j \ge 3$ by

$$F_j^*(\alpha) = \sum_{M_s<m\le 2M_s} \sum_{h\le H_s} \sum_{h_2\le 2P} \cdots \sum_{h_j\le 2P} T_j ,$$

where

$$T_j = T_j(\alpha,h,h_2,\dots,h_j,m) = \sum_{z\in\mathcal{B}} e(2^{-k}\alpha\Psi_j) ,$$

and $\mathcal{B} = \mathcal{B}(h_2,\dots,h_j)$ is the set of z satisfying

$$0 < z \le 2P - h_2 - \dots - h_j .$$

Then by combining Hölder's inequality with the standard Weyl differencing technique (see, for example, Lemma 2.3 of Vaughan [1981b]), we deduce that for $j \ge 2$ we have

$$F_1^*(\alpha) \ll P^{1-2J}H_sM_s + P^{1-2JJ}(H_sM_s)^{1-2J}|F_j^*(\alpha)|^{2J},$$

where we have written $J = 2^{-j}$. Here, the implicit constant may depend on j.

Notice, in particular, that the differencing procedure eliminates all dependency on the auxiliary variable β for $j \ge 3$. Thus our functions $F_j^*(\alpha)$ are the same as Vaughan's functions $F_j(\alpha)$ for $j \ge 3$ (see Vaughan [1989a], equation (2.35)), save that we have $R = 2$.

We therefore deduce that

$$F_1^*(\alpha) \ll P^{1-4\sigma}H_sM_s + P^{1-4(k-2)\sigma}(H_sM_s)^{1-4\sigma}|F_{k-2}^*(\alpha)|^{4\sigma}, \tag{4.5}$$

where by Lemmata 3.1 and 3.2 of Vaughan [1989a], we have

$$|F_{k-2}^*(\alpha)|^2 \le D(\alpha)E(\alpha) , \tag{4.6}$$

and $D(\alpha)$ is an exponential sum which, whenever $|\alpha-a/q| \le q^{-2}$ and $(a,q) = 1$, satisfies

$$D(\alpha) \ll P^{\varepsilon}\left(\frac{P^{k-1}H_s}{q+Q_s^k|\alpha q-a|} + P^{k-2}H_s + q + Q_s^k|\alpha q-a| \right) , \tag{4.7}$$

and $E(\alpha)$ is an exponential sum which, whenever $M_s^k \le X \le Q_s^k M_s^{-k}$, $(a,q) = 1$, $q \le X$, and $|\alpha - a/q| \le q^{-1}X^{-1}$, satisfies

$$E(\alpha) \ll P^{\varepsilon}\left(\frac{P^{k-3}H_s M_s^2}{(q+Q_s^k|\alpha q - a|)^{1/k}} + P^{k-3}H_s M_s \right) . \tag{4.8}$$

Suppose that α satisfies the conditions of the lemma. By Dirichlet's theorem we may choose a and q with $(a,q) = 1$ and $q \le C_k P^{-1}Q_s^k$ so that $|\alpha q - a| \le C_k^{-1}PQ_s^{-k}$. Then $q > 2^{-k}P$, and by (4.5), (4.6), (4.7) and (4.8), we have

$$\begin{aligned} F_1^*(\alpha) &\ll P^{\varepsilon}(P^{1-4\sigma}M_s H_s + P^{1-(k-2)4\sigma}(M_s H_s)^{1-4\sigma}(P^{k-2}H_s . P^{k-3}M_s H_s)^{2\sigma}) \\ &\ll (PM_s)^{1-2\sigma+\varepsilon}H_s . \end{aligned}$$

The lemma follows on noting (4.3) and (4.4).

Lemma 4.3. *Suppose that* $(a,q) = 1$, $\beta = \alpha - a/q$, *and*

$$k(k-1)3^k qP^{k-2}H_s . 2^{k(k-1)}|\beta| \le 1 .$$

Then we have

$$\left|\sum_{p_s} G(\alpha; p_s)\right| \ll \frac{PH_s M_s q^{\varepsilon}}{(q+Q_s^k|\alpha q - a|)^{1/(k-1)}} + H_s M_s q^{\frac{k-2}{k-1}+\varepsilon} .$$

Proof: We are able to use essentially the same Van der Corput analysis for the sum in question as was used in Vaughan [1989a] for the sum $F_1(\alpha)$ defined in proof of the previous lemma. We write $G(\alpha; p_s)$ in the form

$$\sum_{h \le H/d} S(\alpha, h, p_s) ,$$

where

$$S(\alpha, h, m) = \sum_{\substack{hd < z \le 2P-hd \\ z \equiv h \pmod 2}} e(\alpha\Phi(z, h, m)) .$$

It is the exponential sum $S(\alpha, h, m)$ to which we apply the Van der Corput analysis, making a substitution of the form $z = 2z' + a$ to account for the parity condition in the summation over z. This leads to a power of 2 appearing in our analysis, relative to that of

Vaughan, which we have accounted for in the condition on $|\beta|$. The summations over h and p_s in $|\sum_{p_s} G(\alpha;p_s)|$ cause no further difficulties in the end-game analysis. The result now follows from Lemma 3.5 of Vaughan [1989a].

We let $\mathfrak{m}$ denote the set of real numbers α in $(C_k^{-1}P^{1-\sigma}Q^{-k}, 1 + C_k^{-1}P^{1-\sigma}Q^{-k}]$ with the property that whenever $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and $|\alpha q - a| \le C_k^{-1}P^{1-\sigma}Q^{-k}$, then one has $q > M^kP^{1-\sigma}$. Plainly, we have $\mathfrak{m} \subset (0,1)$. By (2.14) we have

$$|\mathcal{F}(\alpha)| = \prod_{s=1}^{2t} \sum_{p_s} |F(\alpha)|^{1/t} |g_s(\alpha p_s^k)| ,$$

and by (2.10), (2.12) and (3.6), we have

$$F(\alpha) = F(\alpha;p_s) + H(\alpha p_s^k; P/p_s) ,$$

where for the sake of brevity we have written $F(\alpha;p_s)$ for $F(\alpha;p_s;P)$. Therefore, by Hölder's inequality we have

$$I(\mathfrak{m}) \le \prod_{s=1}^{2t} \Big(\int_{\mathfrak{m}} \Big(\sum_{p_s} |F(\alpha)|^{1/t} |g_s(\alpha p_s^k)| \Big)^{2t} d\alpha \Big)^{1/2t} ,$$

and

$$\Big(\sum_{p_s} |F(\alpha)|^{1/t} |g_s(\alpha p_s^k)| \Big)^{2t} \ll \Big(\sum_{p_s} |F(\alpha;p_s)|^{1/t} |g_s(\alpha p_s^k)| \Big)^{2t} + \Big(\sum_{p_s} |H(\alpha p_s^k; P/p_s)|^{1/t} |g_s(\alpha p_s^k)| \Big)^{2t} .$$

Thus, by Hölder's inequality,

$$I(\mathfrak{m}) \ll \prod_{s=1}^{2t} \Big(M_s^{2t-1} I_s + J_s \Big)^{1/2t} \tag{4.9}$$

where

$$I_s = \sum_{p_s} \int_{\mathfrak{m}} |F(\alpha;p_s)|^2 |g_s(\alpha p_s^k)|^{2t} \, d\alpha , \tag{4.10}$$

and

$$J_s = \int_{\mathfrak{m}} \Big(\sum_{p_s} |H(\alpha p_s^k; P/p_s)|^{1/t} |g_s(\alpha p_s^k)| \Big)^{2t} d\alpha . \tag{4.11}$$

First consider the integral J_s. Suppose that $\alpha \in \mathfrak{m}$ and $M_s < p \le 2M_s$, and choose b, r with $(b,r) = 1$, $1 \le r \le Q_s^{k-1}$, and $|\alpha p^k r - b| \le Q_s^{1-k}$. Then by Weyl's inequality (see Vaughan [1981b], Lemma 2.4) and a standard major arc estimate (Vaughan [1981a], Theorem 2), we have

$$H(\alpha p^k; P/p) \ll Q_s^{1-\sigma+\varepsilon} \tag{4.12}$$

when $r > Q_s^{k\sigma}$ or $|\alpha p^k r - b| > Q_s^{-k(1-\sigma)}$. Moreover, when $r \le Q_s^{k\sigma}$ and $|\alpha p^k r - b| \le Q_s^{-k(1-\sigma)}$, we have by (2.1), (2.2) and (4.1),

$$rp^k \le 2^k Q_s^{k\sigma} M_s^k < M^k P^{1-\sigma},$$

and

$$|\alpha p^k r - b| \le C_k^{-1} P^{1-\sigma} Q^{-k},$$

since for $k \ge 5$ we have $k+1 < 2^{k-1}$, and hence $Q_s^{k\sigma} \le C_k^{-1} P^{1-\sigma}$. So in this latter case there are a, q with $(a,q) = 1$, $q < M^k P^{1-\sigma}$ and $|\alpha q - a| \le C_k^{-1} P^{1-\sigma} Q^{-k}$. This contradicts the definition of $\mathfrak{m}$, so we may conclude that (4.12) holds uniformly for all $\alpha \in \mathfrak{m}$ and all primes p with $M_s < p \le 2M_s$. Thus by (4.11), and Lemma 3.2, we have

$$J_s \ll Q_s^{2-2\sigma+2\varepsilon} M_s P^{\lambda+k\tau+\varepsilon}.$$

Hence, by Lemma 2.1(i),

$$J_s \ll P^{2t+2-k-\delta} \tag{4.13}$$

for some $\delta > 0$.

Now consider the integral I_s. Let $\mathfrak{n}$ denote the set of real numbers α in $(0,1]$ with the property that whenever $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and $|\alpha q - a| \le C_k^{-1} P^{1-\sigma} Q^{-k}$, then one has $q > 2^{-k} P^{1-\sigma}$, and put

$$\mathfrak{n}_c = \{ \alpha : \alpha - c \in \mathfrak{n} \}.$$

For a given prime number p with $M_s < p \le 2M_s$ and $p \equiv -1 \pmod{k}$, we write

$$d = p^k,$$

$$\mathcal{B}_d = \bigcup_{c=0}^{d-1} \mathfrak{n}_c,$$

$$\mathcal{C}_d = \{ \alpha : \alpha d \in \mathcal{B}_d \}.$$

We use an argument of Vaughan [1989b] to show that $\mathfrak{m} \subset \mathfrak{C}_d$. Let $\alpha \in \mathfrak{m}$ and choose c so that $0 \le c < \alpha d \le c+1 \le d$. Suppose that $\alpha d \notin \mathfrak{n}_c$. Since $0 < \alpha d - c \le 1$ there exist b, r with $(b,r) = 1$, $1 \le r \le 2^{-k}P^{1-\sigma}$, and

$$|(\alpha d - c)r - b| \le C_k^{-1}P^{1-\sigma}Q^{-k}.$$

Thus

$$|\alpha dr - (cr+b)| \le C_k^{-1}P^{1-\sigma}Q^{-k},$$

and so there exist a, q with $(a,q) = 1$, $|\alpha q - a| \le C_k^{-1}P^{1-\sigma}Q^{-k}$ and $q \le dr \le (2M_s)^k P^{1-\sigma}/2^k \le P^{1-\sigma}M^k$. This contradicts the definition of $\mathfrak{m}$, and hence we conclude that $\mathfrak{m} \subset \mathfrak{C}_d$.

We therefore have

$$\int_{\mathfrak{m}} |F(\alpha;p)|^2 |g_s(\alpha d)|^{2t}\, d\alpha \le \int_{\mathfrak{C}_d} |F(\alpha;p)|^2 |g_s(\alpha d)|^{2t} d\alpha ,$$

and by a change of variables,

$$\begin{aligned}\int_{\mathfrak{C}_d} |F(\alpha;p)|^2 |g_s(\alpha d)|^{2t} d\alpha &= d^{-1}\int_{\mathfrak{B}_d} |F(\beta/d;p)|^2 |g_s(\beta)|^{2t} d\beta \\ &= d^{-1}\sum_{c=0}^{d-1} \int_{\mathfrak{n}_c} |F(\beta/d;p)|^2 |g_s(\beta)|^{2t} d\beta \\ &= d^{-1}\sum_{c=0}^{d-1} \int_{\mathfrak{n}} |F((\alpha+c)/d;p)|^2 |g_s(\alpha+c)|^{2t} d\alpha .\end{aligned}$$

Thus

$$\int_{\mathfrak{m}} |F(\alpha;p)|^2 |g_s(\alpha d)|^{2t} d\alpha \le \int_{\mathfrak{n}} d^{-1}\sum_{c=0}^{d-1} |F((\alpha+c)/d;p)|^2 |g_s(\alpha)|^{2t} d\alpha .$$

By (3.6) and the orthogonality of the additive characters $e(cx/d)$ modulo d, we have

$$d^{-1}\sum_{c=0}^{d-1} |F((\alpha+c)/d;p)|^2 = \sum_x \sum_y e(\alpha(x^k - y^k)/d) ,$$

where the double sum is over x, y with $(xy,p) = 1$, $x,y \le P$, $x^k \equiv y^k \pmod d$. Since $d = p^k$ and $p \equiv -1 \pmod k$ we have $x^2 \equiv y^2 \pmod d$ if k is even, and otherwise $x \equiv y \pmod d$. Thus,

adopting the notation of writing χr for r if k is odd, and for $\pm r$ if k is even, the double sum in question is at most

$$\sum_{\substack{r=1\\(r,p)=1}}^{d}\Big|\sum_{w=\chi r}\sum_{\substack{x\le P\\ x\equiv w\ (\mathrm{mod}\ d)}}e(\alpha x^k/d)\Big|^2 \le \sum_{r=1}^{d}2\sum_{w=\chi r}\Big|\sum_{\substack{x\le P\\ x\equiv w\ (\mathrm{mod}\ d)}}e(\alpha x^k/d)\Big|^2$$

$$\le 4[P] + 8\mathrm{Re}\sum_{\substack{y<x\le P\\ x\equiv y\ (\mathrm{mod}\ d)}}\sum e(\alpha(x^k-y^k)/d).$$

We may make the substitutions $h = (x - y)/d$ and $z = x+y$, and then the last double sum becomes

$$G(\alpha;p) = \sum_{h\le P/d}\ \sum_{\substack{hd<z\le 2P-hd\\ z\equiv h\ (\mathrm{mod}\ 2)}} e(2^{-k}\alpha\Phi(z,h,p)) .$$

Therefore, by (4.10),

$$I_s \le 4\int_{\mathfrak{n}}\sum_{p_s}([P] + 2\ \mathrm{Re}\ G(\alpha;p_s))|g_s(\alpha)|^{2t}\,d\alpha$$

$$\ll M_s PQ_s^{\lambda+\varepsilon} + \int_{\mathfrak{n}}|\sum_{p_s}G(\alpha;p_s)|.|g_s(\alpha)|^{2t}\,d\alpha . \qquad (4.14)$$

We subdivide $\mathfrak{n}$ into two sets $\mathfrak{P}_s$ and $\mathfrak{l}_s$. We let $\mathfrak{P}_s$ denote the set of real numbers α in $\mathfrak{n}$ with the property that whenever $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and $|\alpha q - a| \le C_k^{-1}PQ_s^{-k}$, then one has $q > 2^{-k}P$. Plainly $\mathfrak{P}_s \subset \mathfrak{n}$. We then put $\mathfrak{l}_s = \mathfrak{n}\backslash\mathfrak{P}_s$.

If $\alpha \in \mathfrak{P}_s$, then by Lemma 4.2 we have

$$|\sum_{p_s}G(\alpha;p_s)| \ll (PM_s)^{1-2\sigma+\varepsilon}H_s . \qquad (4.15)$$

If $\alpha \in \mathfrak{l}_s$, then $\alpha \notin \mathfrak{P}_s$, and so there exist integers a and q satisfying $1 \le q \le 2^{-k}P$ and $(a,q) = 1$ with $|\alpha q - a| \le C_k^{-1}PQ_s^{-k}$. Then since $\alpha \in \mathfrak{n}$, either

$$|\alpha q - a| > C_k^{-1}P^{1-\sigma}Q^{-k},$$

or else

$$q > 2^{-k}P^{1-\sigma}.$$

We may therefore deduce from Lemma 4.3 that

$$|\sum_{p_s} G(\alpha;p_s)| \ll P^{1+\varepsilon}H_sM_s(P^{1-\sigma-k\tau})^{-1/(k-1)}$$

$$\ll (PM_s)^{1-2\sigma+\varepsilon}H_s$$

since for $k \geq 5$ we have $(1-\sigma)/(k-1) > 2\sigma(1+\theta)$.

Then by (1.6) and (4.15), on recalling that $\mathfrak{n} = \mathfrak{f}_s \cup \mathfrak{l}_s$, we have

$$\int_{\mathfrak{n}} |\sum_{p_s} G(\alpha;p_s)| . |g_s(\alpha)|^{2t}\, d\alpha \ll (PM_s)^{1-2\sigma+\varepsilon}H_s . Q_s^{\lambda+\varepsilon} .$$

Then by (4.14),

$$I_s \ll (PM_s + (PM_s)^{1-2\sigma+\varepsilon}H_s)Q_s^{\lambda+\varepsilon} ,$$

and so by (4.13) and the definition of θ, for some $\delta > 0$ we have

$$M_s^{2t-1}I_s + J_s \ll P^{2t+2-k-\delta}.$$

We may then conclude from (4.9) that

$$\int_{\mathfrak{m}} |\mathcal{F}(\alpha)|\, d\alpha \ll P^{2t+2-k-\delta}.$$

5. THE MAJOR ARCS.

Thus far, we have concluded that

$$r(n;\underline{\underline{M}}) = I(\mathfrak{M}) + O(P^{2t+2-k-\delta}) , \tag{5.1}$$

for some $\delta > 0$, where $I(\mathfrak{M})$ is the contribution to

$$\int_{C_k^{-1}P^{1-\sigma}Q^{-k}}^{1 + C_k^{-1}P^{1-\sigma}Q^{-k}} \mathcal{F}(\alpha)e(-\alpha n)\, d\alpha$$

from the major arcs $(C_k^{-1}P^{1-\sigma}Q^{-k}, 1 + C_k^{-1}P^{1-\sigma}Q^{-k}]\backslash\mathfrak{m}$; that is, from the union $\mathfrak{M}$ of the intervals

$$\mathfrak{M}(q,a) = \{\alpha : |\alpha q - a| \leq C_k^{-1}P^{1-\sigma}Q^{-k}\}$$

with $1 \leq a \leq q \leq P^{1-\sigma}M^k$ and $(a,q) = 1$. We now obtain a suitable estimate for this contribution, using ideas from Vaughan [1989a, b].

Define $V(\alpha)$ on $\mathfrak{M}$ by taking

$$V(\alpha) = q^{-1}S(q,a)v(\alpha-a/q) \quad (\alpha \in \mathfrak{M}(q,a))$$

where

$$v(\beta) = \sum_{x\le n} k^{-1}x^{1/k-1}e(\beta x) ,$$

and

$$S(q,a) = \sum_{r=1}^{q} e(ar^k/q) .$$

Then if $\Delta(\alpha) = F(\alpha) - V(\alpha)$, we have by Vaughan [1981a], Theorem 2 that

$$\Delta(\alpha) \ll q^{\varepsilon}(q + P^k|\alpha q-a|)^{1/2} \quad (\alpha \in \mathfrak{M}(q,a)) . \tag{5.2}$$

Hence, for $\alpha \in \mathfrak{M}$ we have $\Delta(\alpha) \ll P^{\varepsilon}(P^{1-\sigma}M^k)^{1/2}$. Then by Lemmata 2.1(ii) and 3.1, we have

$$\int_{\mathfrak{M}} |\Delta(\alpha)|^2 \Big| \sum_{p_s} g_s(\alpha p_s^k) \Big|^{2t} d\alpha \ll P^{1-\sigma+2\varepsilon}M^k . P^{\lambda+k\tau+\varepsilon}$$
$$\ll P^{2t+2-k-2\delta}, \tag{5.3}$$

for some $\delta > 0$.

Next, by appealing to Vaughan [1981b], Lemma 4.6, we obtain

$$V(\alpha) \ll P(q + P^k|\alpha q-a|)^{-1/k} \quad (\alpha \in \mathfrak{M}(q,a)),$$

and hence, by (5.2),

$$V(\alpha)\Delta(\alpha) \ll P^{1+\varepsilon}(P^{1-\sigma+\varepsilon}M^k)^{1/2 - 1/k} \quad (\alpha \in \mathfrak{M}(q,a)) .$$

Therefore, by Lemmata 2.1(iii) and 3.1, we have

$$\int_{\mathfrak{M}} |V(\alpha)\Delta(\alpha)| \Big| \sum_{p_s} g_s(\alpha p_s^k) \Big|^{2t} d\alpha \ll P^{1+\varepsilon}(P^{1-\sigma+\varepsilon}M^k)^{1/2-1/k}P^{\lambda+k\tau+\varepsilon}$$
$$\ll P^{2t+2-k-2\delta} \tag{5.4}$$

for some $\delta > 0$. Thus, by using (5.1), (5.3), (5.4) and Hölder's inequality, we have

$$\sum_{\underline{\underline{M}}} r(n;\underline{\underline{M}}) = \int_{\mathfrak{M}} V(\alpha)^2 \prod_{s=1}^{2t} \Big(\sum_{M_s} \sum_{p_s} g_s(\alpha p_s^k) \Big) e(-\alpha n)d\alpha + O(P^{2t+2-k-\delta}) \tag{5.5}$$

for some $\delta > 0$.

We are forced to hard prune the major arcs. Let W denote a parameter to be chosen later, and let $\mathfrak{N}$ denote the union of the intervals

$$\mathfrak{N}(q,a) = \{\alpha : |\alpha q - a| \le WP^{-k}\} ,$$

with $(a,q) = 1$ and $1 \le a \le q \le W$. We assume that

$$1 \le W \le P^{1/2},$$

so that $\mathfrak{N} \subset \mathfrak{M}$. Let

$$\mathfrak{P} = \mathfrak{M} \setminus \mathfrak{N} .$$

Then, by the methods of Vaughan [1981b] §4.4 (cf. Lemma 5.1 of Vaughan [1989a]) it can be shown that there is a positive constant δ such that

$$\int_{\mathfrak{P}} |V(\alpha)|^{t+2} \, d\alpha \ll P^{t+2-k} W^{-\delta}. \tag{5.6}$$

In addition, if we write

$$K = \int_0^1 |\sum_{M_s} \sum_{p_s} g_s(\alpha p_s^k)|^{2t+4} \, d\alpha ,$$

then by considering the underlying diophantine equations, we have

$$K \ll \int_0^1 |H(\alpha;2P)|^4 . |\sum_{M_s} \sum_{p_s} g_s(\alpha p_s^k)|^{2t} \, d\alpha .$$

But now we may once again apply the Hardy-Littlewood method.

Define

$$\mathfrak{B}(q,a) = \{\alpha : |q\alpha - a| \le (2k)^{-1} P^{1-k}\}$$

for $1 \le a \le q \le P$ and $(a,q) = 1$, and define $\mathfrak{B}$ to be the union of these arcs, and $\mathfrak{w} = ((2k)^{-1}P^{1-k}, 1 + (2k)^{-1}P^{1-k}] \setminus \mathfrak{B}$. Then by Weyl's inequality, we deduce that for $\alpha \in \mathfrak{w}$ we have $|H(\alpha;2P)| \ll P^{1-\sigma+\varepsilon}$. Hence, by Lemma 3.1 we have

$$K \ll P^{4-4\sigma+4\varepsilon} . P^{\lambda+k\tau+2\varepsilon} + \int_{\mathfrak{B}} |H(\alpha;2P)|^4 . |\sum_{M_s} \sum_{p_s} g_s(\alpha p_s^k)|^{2t} d\alpha .$$

But from Vaughan [1989a], Lemma 5.1 and an application of Hölder's

inequality, we deduce that

$$\int_{\mathfrak{M}} |H(\alpha;2P)|^4 . \Big|\sum_{M_s}\sum_{P_s} g(\alpha p_s^k)\Big|^{2t} d\alpha$$

$$\ll \left(\int_{\mathfrak{M}} |H(\alpha;2P)|^{2t+4} d\alpha\right)^{\frac{2}{t+2}} \left(\int_0^1 \Big|\sum_{M_s}\sum_{P_s} g(\alpha p_s^k)\Big|^{2t+4} d\alpha\right)^{\frac{t}{t+2}}$$

$$\ll (P^{2t+4-k})^{2/(t+2)} . K^{t/(t+2)}.$$

Then by Lemma 2.1(iv), we deduce that $K \ll P^{2t+4-k}$. Hence by Hölder's inequality and (5.6), we have

$$\int_{\mathfrak{m}} \left|V(\alpha)^2 \prod_{s=1}^{2t}\Big(\sum_{M_s}\sum_{P_s} g_s(\alpha p_s^k)\Big)\right| d\alpha \ll (P^{t+2-k}W^{-\delta})^{\frac{2}{t+2}} . (P^{2t+4-k})^{\frac{t}{t+2}}$$

$$\ll P^{2t+2-k}W^{-\nu},$$

for some $\nu > 0$.

By the methods of Vaughan [1989a], §5, when $W \le \log P$, $q \le \log P$, and $(a,q) = 1$, we have

$$\sum_{P_s} g_s(\alpha p_s^k) = q^{-1}S(q,a)u_s(\alpha - a/q) + O\left(\frac{P}{\log P}(q+P^k|\alpha q - a|)\right)$$

and

$$u_s(\beta) \ll \frac{P}{\log P}(1+P^k\|\beta\|)^{-1/k},$$

where

$$u_s(\beta) = \sum_{x \le (2P)^k} \frac{\mathrm{Min}\{\log(2Px^{-1/k}), \log 2\}}{\phi(k)\ \log M_s} k^{-1}x^{1/k-1}\rho\left(\frac{\log(x^{1/k}/M_s)}{\log R}\right)e(\beta x)$$

and ρ is Dickman's function, defined by (4.4.5).

In particular, when $x \ge 0$, $\rho(x)$ is positive and decreasing. Thus if we take ϕ sufficiently small, and

$$W = (\log P)^{\phi},$$

then in the usual way we obtain

$$\int_{\mathfrak{N}} V(\alpha)^2 \prod_{s=1}^{2t}\Big(\sum_{P_s} g_s(\alpha p_s^k)\Big)e(-\alpha n)d\alpha = \mathfrak{S}(n)J(n) + O(P^{2t+2-k}(\log P)^{-2t-\nu})$$

where $\mathfrak{S}(n)$ is the usual singular series in Waring's problem,

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} (q^{-1}S(q,a))^{2t+2} e(-an/q),$$

and

$$J(n) = \int_0^1 v(\beta)^2 \Big(\prod_{s=1}^{2t} u_s(\beta) \Big) e(-\beta n)\, d\beta ,$$

and ν is a positive constant. Now by hypothesis, $\Theta > 0$ (see, for example, Vaughan [1981b], Theorem 4.6), and a simple counting argument shows that

$$J(n) \gg n^{(2t+2-k)/k} (\log n)^{-2t}.$$

Thus

$$\sum_{\underline{\underline{M}}} r(n; \underline{\underline{M}}) = \sum_{\underline{\underline{M}}} \int_{\mathfrak{N}} V(\alpha)^2 \prod_{s=1}^{2t} \Big(\sum_{p_s} g_s(\alpha p_s^k) \Big) e(-\alpha n) d\alpha + O(P^{2t+2-k} (\log P)^{-\nu})$$

$$\gg n^{(2t+2-k)/k},$$

and this completes the proof of Theorem 1.3.

6. THE PROOF OF THEOREM 1.1.

The proof of Theorem 1.1 follows directly from Theorem 1.3 on checking that the inequalities (1.3)-(1.6) hold. We use existing methods of Vaughan [1989a, b] to provide values of λ_s which, when R is no larger than a small power of P, give bounds of the form

$$S_s(P, R) \ll P^{\lambda_s + \varepsilon}.$$

For this purpose, the methods implicit in the inequalities (j1), (j2) and (k-2) of §4 of Vaughan [1989a] are available to us. For $s = 3$ and 4 the estimates given by Theorem 1.4 of Vaughan [1989c] (which we shall call method "*") are always superior to the estimates given by Vaughan [1989a]. When $s = 5$, the second of the estimates of Lemma 2.2 of Vaughan [1989c] (which we denote by "j-2", where j is the parameter used in the method) sometimes leads to an improvement over the methods of Vaughan [1989a].

In Table 6.1 we list the optimal values of λ_s for those s with $\lambda_s > 2s-k$ which arise from one of the above mentioned methods when $k = 5$ and 6. We also list the corresponding values of θ_s used in these methods, and in column j we list the method giving the optimal choice of λ_s. The table of values was computed to 16 significant figures by using an electronic computer, the final significant figure being rounded up.

Table 6.1.

k	s	j	θ	λ_s
5	3	*	0.0681285335	3.136258
	4	*	0.1055957693	4.438657
	5	3	0.1516954796	5.978896
	6	3	0.1655871346	7.644740
	7	$(k-2)$	0.1707317074	9.388321
	8	$(k-2)$	0.1707317074	11.175681
6	3	*	0.0454545455	3.090910
	4	*	0.0833333334	4.333334
	5	3-2	0.1204041621	5.774816
	6	3	0.1346926697	7.343917
	7	4	0.1467086118	9.027005
	8	4	0.1510124478	10.777989
	9	$(k-2)$	0.1546391753	12.585517
	10	$(k-2)$	0.1546391753	14.422808
	11	$(k-2)$	0.1546391753	16.285260
	12	$(k-2)$	0.1546391753	18.168983
	13	$(k-2)$	0.1546391753	20.070687

We now complete the proof of Theorem 1.1. Let

$$t = t(k) = \begin{cases} 8 & k = 5 \\ 13 & k = 6 \end{cases}$$

and

$$\theta = \theta(k) = \begin{cases} 7/41 & k = 5 \\ 15/97 & k = 6 \end{cases}.$$

Then for $k = 5$ or 6, we take $\mu = \lambda_{t-1}$ to be as given by Table 6.1.

Then on noting that

$$\lambda_t = \lambda_{t-1}(1-\theta) + 1 + (2t-2)\theta ,$$

the conditions of Theorem 1.3 are easily verified, and we conclude that

$$G(k) \le 2t+2 .$$

7. A MAJOR ARC ESTIMATE.

As our first step towards the proof of Theorem 1.4 we establish a "major arc" estimate for $g(\alpha)$. We shall first require some notation.

We suppose that ε is a sufficiently small positive number, and that P is sufficiently large in terms of k and ε. We shall use the letter p to denote a rational prime. Implicit constants will depend at most on k and ε.

Lemma 7.1. *Suppose that* $X > 0$, $J > 0$, *and* $0 < \theta < 1$. *Then*

$$\sum_{j \le J} (1 + jX)^{-\theta} \le 2^{\theta}(1-\theta)^{-1}J(1 + XJ)^{-\theta}.$$

Proof: When $0 < X \le J^{-1}$, the sum in question is bounded above by

$$J \le 2^{\theta}J(1 + XJ)^{-\theta},$$

and when $J^{-1} < X$, it is bounded above by

$$X^{-\theta}\sum_{j \le J} j^{-\theta} < X^{-\theta}(1-\theta)^{-1}J^{1-\theta} \le 2^{\theta}(1-\theta)^{-1}J(1 + XJ)^{-\theta}.$$

This completes the proof of the lemma.

Define

$$g(\alpha) = \sum_{x \in \mathcal{A}(P,R)} e(\alpha x^k) .$$

Also, for the sake of conciseness, write

$$\mathcal{L} = \log P \text{ and } \mathcal{L}_2 = \log\log P .$$

Lemma 7.2. *Suppose that* $2 \le R \le M \le P$, $|\alpha - a/q| \le M/(kq(2P)^k R)$, *and* $(a,q) = 1$. *Then*

$$g(\alpha) \ll \mathcal{L}^3 q^{\epsilon}\left[P(q + P^k|\alpha q - a|)^{-1/2k} + (PMR)^{1/2} + q^{1/4}P(R/M)^{1/2}\right] .$$

Proof: For each $p \le R$, let

$$\mathcal{B}_p = \{ v : M < v \le Mp ;\ p|v ;\ p'|v \text{ implies } p \le p' \le R \} .$$

Then by Lemma 10.1 of Vaughan [1989a], we have

$$g(\alpha) = S_0 + O(M)$$

where

$$S_0 = \sum_{p \le R} \sum_{v \in \mathcal{B}_p} \sum_{u \in \mathcal{A}(P/v,\, p)} e(\alpha u^k v^k) .$$

Let

$$W = \{(2^j, 2^i M) : i = 0,1,2,\ldots,\ j = -1,0,1,2,\ldots,$$
$$2^i < R \text{ and } 2^{i+j}M < P\} .$$

Then

$$|S_0| \le \sum_{(U,V) \in W} S_1(U,V)$$

where

$$S_1(U,V) = \sum_{p \le R} \sum_{\substack{V < v \le 2V \\ p|v}} \Big| \sum_{\substack{u \in \mathcal{A}(P/v,\, p) \\ U < u \le 2U}} e(\alpha u^k v^k) \Big| .$$

Clearly $\mathrm{card}(W) \ll \mathcal{L}^2$. Thus on choosing $(U,V) \in W$ so that $S_1(U,V)$ is maximal, we obtain

$$g(\alpha) \ll M + \mathcal{L}^2 S_1(U,V) . \tag{7.1}$$

For future reference, note that

$$M \le V < MR \text{ and } 1/2 \le U < P/V. \tag{7.2}$$

On noting that

$$\sum_{p \le R} \sum_{\substack{V < v \le 2V \\ p|v}} 1 \ll V \sum_{p \le R} p^{-1} \ll V.\log\log R ,$$

we obtain by Cauchy's inequality

$$S_1(U,V)^2 \ll V\mathcal{L}_2 S_2 \tag{7.3}$$

where

$$S_2 = \sum_{p\le R} \sum_{\substack{V<v\le 2V \\ p\mid v}} \Big| \sum_{\substack{u\in\mathcal{A}(P/v,\,p) \\ U<u\le 2U}} e(\alpha u^k v^k)\Big|^2$$

$$= \sum_{p\le R} \sum_{u_1} \sum_{u_2} \sum_{\substack{V/p<w\le 2V/p \\ w\le P/(u_1 p) \\ w\le P/(u_2 p)}} e(\alpha p^k (u_2^k - u_1^k) w^k) \ . \qquad (7.4)$$

$$u_1\in\mathcal{A}(P/V,\,p),\quad U<u_1\le 2U$$

For each p, u_1, u_2 under consideration we have

$$p^k|u_2^k - u_1^k|.|\alpha - a/q| \le \frac{p^k(P/V)^k M}{kq(2P)^k R} = \frac{(p/R)(M/V)}{2kq(2V/p)^{k-1}} \ .$$

Thus, on writing

$$D = (q, p^k(u_2^k - u_1^k)) \ ,$$

and

$$a' = ap^k(u_2^k - u_1^k)/D \ , \quad q' = q/D \ , \quad \alpha' = \alpha p^k(u_2^k - u_1^k) \ ,$$

we have

$$|\alpha' - a'/q'| \le (2kq'(2V/p)^{k-1})^{-1}.$$

Hence, by Lemma 2.8 and Theorems 4.1 and 4.2 of Vaughan [1981b], the innermost sum in the expression for S_2 in (7.4) is

$$\ll (V/p).(q' + (V/p)^k|\alpha' q' - a'|)^{-1/k} + q'^{1/2+\varepsilon}.$$

Hence, by (7.4) we have

$$S_2 \ll RU^2 q^{1/2+\varepsilon} + \mathcal{L}_2 UV + S_3 \qquad (7.5)$$

where

$$S_3 = \sum_{p\le R} \sum_{U<u_1<u_2\le 2U} \frac{V(q, p^k(u_2^k - u_1^k))^{1/k}}{p(q + V^k(u_2^k - u_1^k)|\alpha q - a|)^{1/k}} \ .$$

By considering the cases $p|q$ and $p\nmid q$ separately and by bounding $u_2^k - u_1^k$ in the denominator in terms of $(u_2 - u_1)U^{k-1}$, we obtain

$$S_3 \ll q^{\varepsilon/2}\mathcal{L}_2 \sum_{U<u_1<u_2\le 2U} \frac{V(q, u_2^k - u_1^k)^{1/k}}{(q + U^{k-1}V^k(u_2 - u_1)|\alpha q - a|)^{1/k}} \ . \qquad (7.6)$$

So by (7.1), and (7.3)-(7.6), we deduce that

$$g(\alpha) \ll M + \mathcal{L}^2(\mathcal{L}_2 RU^2 Vq^{1/2+\varepsilon} + \mathcal{L}_2^2 UV^2 + \mathcal{L}_2^2 V^2 q^{\varepsilon/2} S_4)^{1/2} \qquad (7.7)$$

where

$$S_4 = \sum_{U<u_1<u_2\le 2U} \sum \frac{(q,u_2^k-u_1^k)^{1/k}}{(q + U^{k-1}V^k(u_2-u_1)|\alpha q-a|)^{1/k}} .$$

On writing $d = (q,u_2^k-u_1^k)$, $\beta = |\alpha-a/q|$, and then putting $m = (u_1,u_2)$ and $w_i = u_i/m$ $(i = 1,2)$, we obtain

$$S_4 \le \sum_{d|q} (d/q)^{1/k} \sum_{m\le 2U} \sum_{\substack{U/m<w_1<w_2\le 2U/m \\ (w_2,w_1)=1 \\ d|m^k(w_2^k-w_1^k)}} (1 + U^{k-1}V^k m(w_2-w_1)\beta)^{-1/k} .$$

For a given d and m let $d_0 = (d,m^k)$ and let $e_0 = d/d_0$. Thus

$$S_4 \le \sum_{d|q} (d/q)^{1/k} \sum_{\substack{d_0,e_0 \\ d_0e_0=d}} \sum_{\substack{m\le 2U \\ d_0|m^k}} S_5 \tag{7.8}$$

where

$$S_5 = \sum_{U/m<w_1<w_2\le 2U/m} \sum_{\substack{(w_2,w_1)=1 \\ e_0|(w_2^k-w_1^k)}} (1 + U^{k-1}V^k m(w_2-w_1)\beta)^{-1/k}.$$

Now we put $w = w_1$, $h = w_2 - w_1$, $e_1 = (h,e_0)$, $f_1 = e_0/e_1$ and $j = h/e_1$. Thus

$$S_5 \le \sum_{\substack{e_1,f_1 \\ e_1f_1=e_0}} \sum_{\substack{j\le U/me_1 \\ (j,f_1)=1}} \sum_{\substack{U/m<w\le 2U/m \\ (w,je_1)=1 \\ e_1f_1|((w+je_1)^k-w^k)}} (1 + U^{k-1}V^k mje_1\beta)^{-1/k}. \tag{7.9}$$

Suppose that $\pi|(e_1f_1,w)$. Then $\pi|(f_1,w)$ and $\pi|w+je_1$ so $\pi|w$ and $\pi|je_1$ which is impossible. Thus $(w,e_1f_1) = 1$. Also,

$$(w+je_1)^k \equiv w^k \pmod{e_1f_1}.$$

Hence

$$e_1^{-1}((w+je_1)^k-w^k) \equiv 0 \pmod{f_1}.$$

Choose $\bar{w}$ so that $w\bar{w} \equiv 1 \pmod{f_1}$. Then

$$((1+je_1\bar{w})^k - 1)/e_1 \equiv 0 \pmod{f_1}$$

and so

$$(1+je_1\bar{w})^k \equiv 1 \pmod{e_1f_1} .$$

The congruence $y^k \equiv 1 \pmod{e_1f_1}$ has at most v solutions, $y_1,\dots,y_v$ say, with $v \ll (e_1f_1)^{\varepsilon/4}$. Hence

$$1 + je_1\bar{w} \equiv y_i \pmod{e_1f_1}$$

for some i. Therefore $y_i \equiv 1 \pmod{e_1}$ and

$$j\bar{w} \equiv (y_i-1)/e_1 \pmod{f_1} .$$

Moreover $(j,f_1) = 1$, so there are at most v choices for $\bar{w}$ modulo f_1, and hence likewise for w. Thus the number of choices for w in the innermost sum in (7.9) is

$$\ll (U/mf_1 + 1)(e_1f_1)^{\varepsilon/4}.$$

Hence

$$S_5 \ll q^{\varepsilon/4} \sum_{e_1f_1=e_0} \sum_{j\le U/me_1} (U/mf_1)(1 + U^{k-1}V^k mje_1\beta)^{-1/k} + q^{\varepsilon/2}U/m .$$

By Lemma 7.1, we have

$$S_5 \ll q^{\varepsilon/4} \sum_{e_1f_1=e_0} (U/mf_1)(U/me_1)(1 + U^{k-1}V^k m(U/me_1)e_1\beta)^{-1/k} + q^{\varepsilon/2}U/m$$

$$\ll q^{\varepsilon/2}(U^2/m^2e_0)(1 + U^kV^k\beta)^{-1/k} + q^{\varepsilon/2}U/m .$$

Thus, by (7.8) we have

$$S_4 \ll q^{\varepsilon/2} \sum_{d\mid q} (d/q)^{1/k} \sum_{\substack{d_0,e_0 \\ d_0e_0=d}} \sum_{\substack{m\le 2U \\ d_0\mid m^k}} U^2(m^2e_0)^{-1}(1+U^kV^k\beta)^{-1/k} + q^{\varepsilon}U\mathcal{L} .$$

By writing $d_0 = d_1d_2^2\dots d_k^k$, where $d_1,\dots,d_{k-1}$ are square-free, we see that $d_1\dots d_k \mid m$, whence

$$S_4 \ll q^{\varepsilon-1/k}U^2(1 + U^kV^k\beta)^{-1/k} + q^{\varepsilon}U\mathcal{L} .$$

Therefore, by (7.7) we deduce that

$$g(\alpha) \ll M + \mathcal{L}^3q^{\varepsilon}\left[RU^2Vq^{1/2} + UV^2 + U^2V^2(q + U^kV^k|\alpha q-a|)^{-1/k}\right]^{1/2}.$$

On noting that $U^2(1 + U^kX)^{-1/k}$ is an increasing function of U, we

deduce that

$$g(\alpha) \ll M + q^{\varepsilon}\mathcal{L}^3\left(RP^2V^{-1}q^{1/2} + PV + P^2(q + P^k|\alpha q-a|)^{-1/k}\right)^{1/2}.$$

Therefore

$$g(\alpha) \ll q^{\varepsilon}\mathcal{L}^3\left(P(q + P^k|\alpha q-a|)^{-1/(2k)} + (PMR)^{1/2} + PM^{-1/2}R^{1/2}q^{1/4}\right).$$

This completes the proof of the lemma.

8. ANOTHER MAJOR ARC ESTIMATE.

Here we establish a further major arc estimate, useful on a rather narrower set than the previous estimate. We shall first require a few preliminary lemmata.

Lemma 8.1. *Let*

$$N_q(X) = \sum_{\substack{x\le X\\(x,q)=1}} X^{-1}, \text{ and } \Lambda_q(X,Y) = \sum_{\substack{y\in\mathcal{A}(X,Y)\\(y,q)=1}} 1 .$$

Then

$$\begin{cases} \Lambda_q(X,Y) = X.N_q(X) + X\int_0^{u-1} \rho'(u-v).N_q(Y^v)\,dv & (X \in \mathbb{N}) \\ \Lambda_q(X,Y) = \Lambda_q(X+,Y) & (X \notin \mathbb{N}) \end{cases}$$

where

$$u = (\log X)/(\log Y)$$

and ρ is Dickman's function, defined by (4.4.5).

Proof: Suppose that $X \notin \mathbb{N}$. Let $R_q(X) = N_q(X) - \phi(q)/q$. Then by equation (1.15) of Fouvry and Tenenbaum [1990], we have

$$\Lambda_q(X,Y) = XI(X,Y),$$

where

$$I(X,Y) = \int_{-\infty}^{\infty} \rho(u-v)\,dR_q(Y^v) .$$

By the definition of ρ, we have

$$I(X,Y) = \int_{(u-1)-}^{u+} dR_q(Y^v) + \int_{0-}^{(u-1)-} \rho(u-v)\, dR_q(Y^v)$$

and on integrating by parts we obtain

$$I(X,Y) = R_q(X) + \rho(u)\phi(q)/q + \int_0^{u-1} \rho'(u-v) R_q(Y^v)\, dv$$

$$= N_q(X) + \int_0^{u-1} \rho'(u-v) N_q(Y^v)\, dv\ .$$

When $X \in \mathbb{N}$ the result follows by one-sided continuity, and this completes the proof of the lemma.

Lemma 8.2. *Suppose that* $0 \le X \le Y$ *and* $\psi : [X,Y] \longrightarrow \mathbb{R}$ *is monotonic and differentiable on* (X,Y), ψ' *is continuous on* (X,Y), *and* $\psi'(X+)$ *and* $\psi'(Y-)$ *exist. Then*

$$\int_X^Y \psi(w) e(\gamma w^k)\, dw \ll (|\psi(X)| + |\psi(Y)|).Min\{\, Y-X,\ |\gamma|^{-1/k} \}\ .$$

Proof: It suffices to show that when $\gamma > 0$ we have

$$\int_X^Y \psi(w) e(\gamma w^k)\, dw \ll (|\psi(X)| + |\psi(Y)|).\gamma^{-1/k}.$$

Moreover it suffices to treat the situation when $X > \gamma^{-1/k}$. Now by integrating by parts the integral on the left is

$$\left[\frac{\psi(w)}{2\pi i \gamma k w^{k-1}}\, e(\gamma w^k) \right]_X^Y$$

$$+ \int_X^Y \left[(1-1/k) w^{-k} \psi(w) - k^{-1} w^{1-k} \psi'(w) \right] e(\gamma w^k)/(2\pi i\gamma)\, dw\ .$$

This completes the proof of the lemma.

Let

$$\Psi(X,Y;q,a) = \sum_{\substack{x \in \mathcal{A}(X,Y) \\ x \equiv a \pmod q}} 1\ . \qquad (8.1)$$

Lemma 8.3. *Let A be a real number with $A > 0$. There are absolute constants c_1 and c_2 such that if*

$$\exp(c_1(\log\log(3+X))^2) \le Y \le X ,$$

$$q \le (\log Y)^A, \text{ and } (a,q) = 1,$$

then

$$\Psi(X,Y;q,a) = \phi(q)^{-1}\Lambda_q(X,Y).(1 + O_A(\exp(-c_2(\log Y)^{1/2}))) .$$

Proof: This is Theorem 5 of Fouvry and Tenenbaum [1990].

Lemma 8.4. *Let*

$$W(q,a) = \sum_{\substack{r=1 \\ (q,r)=1}}^{q} e(ar^k/q) . \tag{8.2}$$

If $(q_1,q_2) = 1$, then $W(q_1q_2,a) = W(q_1,aq_2^{k-1}).W(q_2,aq_1^{k-1})$, and further, if $p \nmid a$, then

$$W(p^t,a) \ll p^{t/2}.$$

Proof: This follows from the proof of Lemma 8.5 of Hua [1965] .

We now prove the lemma alluded to at the start of this section.

Lemma 8.5. *Suppose that $R = P^\eta$ with $0 < \eta < 1/2$. Suppose also that a and q are integers with $(a,q) = 1$ and $q \le (\log P)^A$. Then*

$$g(\alpha) \ll q^\varepsilon P(q + P^k|\alpha q - a|)^{-1/k} + P.\exp(-c_3(\log P)^{1/2})(1 + P^k|\alpha - a/q|)$$

where c_3 may depend on η and A, and the implicit constant may depend on η, ε and A.

Proof: For brevity we write $\beta = \alpha - a/q$. Then

$$g(\alpha) = \sum_{\substack{d \mid q \\ q/d \in \mathcal{A}(P,R)}} \sum_{\substack{t=1 \\ (t,d)=1}}^{d} e((q/d)^{k-1}t^k a/d).\Psi(Pd/q, R; d, t, \beta(q/d)^k)$$

where

$$\Psi(Q,R;d,t,\gamma) = \sum_{\substack{m \in \mathcal{A}(Q,R) \\ m \equiv t \pmod d}} e(\gamma m^k) .$$

Write $Q = Pd/q$ and $\gamma = \beta(q/d)^k$. Then by (8.1) we have

$$\Psi(Q,R;d,t,\gamma) = e(\gamma Q^k).\Psi(Q,R;d,t) - \int_0^Q 2\pi i\gamma k X^{k-1} e(\gamma X^k)\Psi(X,R;d,t)\ dX$$

and so by Lemma 8.3, when $Q \geq P^{1/2}$ we have

$$\phi(d)\Psi(Q,R;d,t,\gamma) = \mathcal{M}_d + \mathcal{E} ,$$

where

$$\mathcal{M}_d = e(\gamma Q^k).\Lambda_d(Q,R) - \int_0^Q 2\pi i\gamma k X^{k-1} e(\gamma X^k)\Lambda_d(X,R)\ dX$$

and

$$\mathcal{E} \ll Q(1 + Q^k|\gamma|)\exp(-c_4(\log Q)^{1/2}) .$$

The main term $\mathcal{M}_d$ is independent of t. Thus, by (8.2),

$$g(\alpha) = \sum_{\substack{d \\ d \mid q \\ q/d \in \mathcal{A}(P,R)}} W(d, a(q/d)^{k-1}).\mathcal{M}_d/\phi(d) + O(P(1+P^k|\beta|)\exp(-c_5(\log P)^{1/2})) \qquad (8.3)$$

Moreover, by Lemma 8.1, we have

$$\mathcal{M}_d = e(\gamma Q^k)\left[QN_d(Q) + Q\int_{-\infty}^{\infty} \rho_1\left[\frac{\log Q}{\log R} - v\right]N_d(R^v)dv - \int_0^Q 2\pi i\gamma k X^{k-1} e(\gamma X^k)\left[X.N_d(X) + X\int_{-\infty}^{\infty} \rho_1\left[\frac{\log X}{\log R} - v\right]N_d(R^v)dv\right] dX$$

where

$$\rho_1(w) = \begin{cases} 0 & \text{for } w \leq 1 \\ -\rho(w-1)/w & \text{for } w > 1 . \end{cases}$$

Thus

$$\mathcal{M}_d = \mathcal{N}_1 - \mathcal{N}_2 + \mathcal{N}_3 \qquad (8.4)$$

where

$$\mathcal{N}_1 = e(\gamma Q^k)QN_d(Q) ,$$

$$\mathcal{N}_2 = \int_0^Q 2\pi i\gamma k X^{k-1} e(\gamma X^k)XN_d(X)\ dX ,$$

and

$$\mathcal{N}_3 = \int_{-\infty}^{\infty} N_d(R^v)\left[e(\gamma Q^k)Q\rho_1\left[\frac{\log Q}{\log R} - v\right] - \int_0^Q 2\pi i\gamma k X^{k-1} e(\gamma X^k)X\rho_1\left[\frac{\log X}{\log R} - v\right] dX\right] dv .$$

Since $N_d(X) = \phi(d)/d + O(d^{\varepsilon}X^{-1})$, we have

$$N_1 - N_2 = d^{-1}\phi(d)\int_0^Q e(\gamma X^k)\,dX + O(d^{\varepsilon}(1 + Q^k|\gamma|)) ,$$

and by Lemma 8.2 we find that

$$N_1 - N_2 \ll Q(1 + Q^k|\gamma|)^{-1/k} + d^{\varepsilon}(1 + Q^k|\gamma|) . \tag{8.5}$$

Now consider the integral

$$\int_0^Q 2\pi i\gamma k X^{k-1} e(\gamma X^k) X \rho_1\left(\frac{\log X}{\log R} - v\right) dX .$$

When $X < R^{v+1}$ the integrand is zero. Thus the integral is zero when $Q < R^{v+1}$, and when $Q \geq R^{v+1}$ it is

$$J(Q,R) = \int_{R^{v+1}}^Q 2\pi i\gamma k X^{k-1} e(\gamma X^k) X \rho_1\left(\frac{\log X}{\log R} - v\right) dX .$$

Let

$$\rho_2(w) = \begin{cases} 0 & w \leq 1 \\ w^{-2} & 1 < w \leq 2 \\ \dfrac{\rho(w-2)}{w(w-1)} + \dfrac{\rho(w-1)}{w^2} & w > 2 . \end{cases}$$

Then on integrating this last integral by parts we obtain

$$J(Q,R) = \left[e(\gamma X^k) X \rho_1\left(\frac{\log X}{\log R} - v\right)\right]_{R^{v+1}+}^{Q} - N_4(v) ,$$

where

$$N_4(v) = \int_{R^{v+1}+}^Q e(\gamma X^k)\left(\rho_1\left(\frac{\log X}{\log R} - v\right) + (\log R)^{-1}\rho_2\left(\frac{\log X}{\log R} - v\right)\right) dX .$$

Thus

$$J(Q,R) = e(\gamma Q^k) Q \rho_1\left(\frac{\log Q}{\log R} - v\right) + e(\gamma R^{kv+k}) R^{v+1} - N_4(v) ,$$

and hence

$$N_3 = \int_0^{\frac{\log(Q/R)}{\log R}} N_d(R^v)\left(N_4(v) - e(\gamma R^{kv+k}) R^{v+1}\right) dv .$$

By Lemma 8.2, we have

$$N_4(v) \ll Q(1 + Q^k|\gamma|)^{-1/k} .$$

Also, $0 \leq N_d(R^v) \leq 1$, and $(\log(Q/R))/\log R \ll 1$. Thus

$$N_3 = -N_5 + O(Q(1 + Q^k|\gamma|)^{-1/k}) \tag{8.6}$$

where

$$N_5 = \int_0^{\frac{\log(Q/R)}{\log R}} N_d(R^v).e(\gamma R^{kv+k})R^{v+1}\,dv\ .$$

Again using the fact that $N_d(X) = \phi(d)/d + O(d^{\varepsilon}X^{-1})$, we obtain

$$N_5 = \frac{\phi(d)}{d.\log R}\int_0^Q e(\gamma w^k)\,dw + O\left[\frac{Rd^{\varepsilon}\log(Q/R)}{\log R}\right]$$

provided that $Q \geq R$. Moreover we can once more estimate the integral above by Lemma 8.2. Therefore, by (8.4), (8.5) and (8.6) we have

$$M_d \ll Q(1 + Q^k|\gamma|)^{-1/k} + Rd^{\varepsilon}(1 + Q^k|\gamma|)\ .$$

Hence, by (8.3) we have

$$g(\alpha) \ll P.\mathcal{S}(q,a).(1 + P^k|\beta|)^{-1/k} + P(1 + P^k|\beta|)\exp(-c_5(\log P)^{1/2})$$

where

$$\mathcal{S}(q,a) = \sum_{d|q}\frac{d}{q\phi(d)}\,|W(d,a(q/d)^{k-1})|\ .$$

It remains to estimate $\mathcal{S}(q,a)$. Given d dividing q, for each prime p define $i = i(p,d)$ and $j = j(p,q)$ by $p^i \| d$ and $p^j \| q$. Then by Lemma 8.4, for some constant c_k depending at most on k, we have

$$d(\phi(d)q)^{-1}|W(d,a(q/d)^{k-1})|$$

$$\leq \prod_p c_k(p^i, p^{(j-1)(k-1)})^{1/2}p^{i/2 - j}.p/(p-1)\ .$$

When $ik \leq j(k-1)$, the general term in the product is

$$c_k p^{i-j}.p/(p-1)\ ,$$

and when $ik > j(k-1)$ it is

$$c_k p^{j(k-3)/2 - i(k-2)/2}p/(p-1)\ .$$

In either case this does not exceed $c_k p^{-j/k}.p/(p-1)$. Thus

$$\mathcal{S}(q,a) \ll q^{\varepsilon - 1/k}.$$

This completes the proof of the lemma.

9. THE PROOF OF THEOREM 1.4.

Suppose that ϕ, $\mathfrak{m}$, σ, $\mathcal{E}_{\varepsilon,k}$, η, k, t and u satisfy the conditions of Theorem 1.4 and the suppositions preceding it. We consider the representation of a large natural number n in the form

$$n = x_1^k + \dots + x_{2t}^k + y_1^k + \dots + y_u^k$$

with $x_i \in \mathcal{A}(P, P^\eta)$ $(1 \le i \le 2t)$, $y_j \in \mathcal{E}_{\varepsilon,k}$ $(1 \le j \le u)$, and $P = n^{1/k}$.

Assume that $\lambda_t = \lambda_t(k)$ has the property that for each $\varepsilon > 0$ and each η with $0 < \eta \le \eta_0(\varepsilon, k)$, whenever $P > P_0(\eta, \varepsilon, k, t)$ we have

$$S_t(P, P^\eta) < P^{\lambda_t + \varepsilon}.$$

Let $\mathfrak{M}(q,a)$ denote the set of $\alpha \in \mathbb{R}$ with $|\alpha - a/q| \le q^{-1}P^{\phi-k}$ and let $\mathfrak{M}$ denote the union of the $\mathfrak{M}(q,a)$ with $1 \le a \le q \le P^\phi$ and $(a,q) = 1$. Further put $\mathfrak{U} = (P^{\phi-k}, 1+P^{\phi-k}]$. By Dirichlet's theorem on diophantine approximation, $\mathfrak{U}\setminus\mathfrak{M} \subset \mathfrak{m}$. Hence, on choosing ε so that $2(u+1)\varepsilon < 2t-k-\lambda_t+\sigma u$ and $\eta = \eta_0(\varepsilon,k)/2$, we have

$$\left| \int_{\mathfrak{U}\setminus\mathfrak{M}} g(\alpha)^{2t} h(\alpha)^u e(-\alpha n)\, d\alpha \right| < P^{\lambda_t + \varepsilon + u(1-\sigma+\varepsilon)} < P^{2t+u-k-\delta}$$

where δ is a sufficiently small but fixed positive number.

Now let $X = (\log P)^{10kt}$, $\mathfrak{P}(q,a) = \{ \alpha : |\alpha q - a| \le XP^{-k} \}$, and define $\mathfrak{P}$ to be the union of the $\mathfrak{P}(q,a)$ with $1 \le a \le q \le X$ and $(a,q) = 1$. Clearly $\mathfrak{P} \subset \mathfrak{M}$. We take $\mathfrak{p} = \mathfrak{M}\setminus\mathfrak{P}$, and $M = k2^k RP^\phi$ in Lemma 7.2. Thus for $|\alpha - a/q| \le q^{-1}P^{\phi-k}$, $(a,q) = 1$ and $q \le P^\phi$ we obtain

$$g(\alpha) \ll (\log P)^3 q^{\varepsilon/(2t)} \left(P(q + P^k|\alpha q - a|)^{-1/(2k)} + P^{1 - 1/(2k+2)} R \right).$$

Thus

$$\int_{\mathfrak{p}} |g(\alpha)|^{2t} d\alpha \ll (\log P)^{6t} \Bigg(\sum_{q \le X} q^{1-t/k+\varepsilon} P^t \int_{q^{-1}XP^{-k}}^{\infty} \beta^{-t/k} d\beta + \sum_{X < q \le P^\phi} q^{\varepsilon+1-t/k} P^{2t-k} + P^{2t-2t/(2k+2)+\varepsilon} R^{2t} \sum_{q \le P^\phi} P^{\phi-k} \Bigg).$$

Then since $t \ge 2k+1$, we obtain

$$\begin{aligned}\int_{\mathfrak{p}} |g(\alpha)|^{2t} d\alpha &\\ &\ll P^{2t-k}(\log P)^{6t}\Big(\sum_{q\le X} X^{1-t/k+\varepsilon} + X^{\varepsilon+2-t/k} + P^{2\phi-t/(k+1)+\varepsilon} R^{2t}\Big)\\ &\ll P^{2t-k}(\log P)^{6t}\Big((\log P)^{20kt+10kt\varepsilon-10t^2} + P^{-1/(k+1)+\varepsilon} R^{2t}\Big)\\ &\ll P^{2t-k}(\log P)^{-10}.\end{aligned}$$

Let $\tau > 0$ be sufficiently small in terms of ε and η, and let $W = (\log P)^{2k\tau}$. Let $\mathfrak{N}(q,a) = \{ \alpha : |\alpha q - a| \le WP^{-k} \}$, and define $\mathfrak{N}$ to be the union of the $\mathfrak{N}(q,a)$ with $1 \le a \le q \le W$ and $(a,q) = 1$. Finally, let $\mathfrak{n} = \mathfrak{P}\setminus\mathfrak{N}$. Then by Lemma 8.5, much as we used Lemma 7.2 above, we deduce that

$$\int_{\mathfrak{n}} |g(\alpha)|^{2t} d\alpha \ll P^{2t-k}(\log P)^{-\tau}.$$

Thus we obtain

$$\int_{\mathfrak{U}\setminus\mathfrak{N}} |g(\alpha)^{2t} h(\alpha)^u|\, d\alpha \ll P^{2t-k} h(0)^u (\log P)^{-\tau}.$$

Finally, by applying the method of the latter part of §5 of Vaughan [1989a], we obtain

$$\int_{\mathfrak{N}} g(\alpha)^{2t} h(\alpha)^u e(-\alpha n)\, d\alpha \gg P^{2t-k} h(0)^u.$$

This completes the proof of Theorem 1.4.

10. USE OF VINOGRADOV'S MEAN VALUE THEOREM.

Let ε be a sufficiently small positive number, and let θ be a real number with $0 < \theta < 1/k$, and $\eta = \eta(\varepsilon)$ be a small positive number (a matter on which we shall elaborate later). Let P be a large real number, and define

$$M = P^{\theta},\quad H = PM^{-k},\quad Q = PM^{-1},\quad R = P^{\eta}.$$

Define

$$\Psi_1(z,h,m) = m^{-k}((z+hm^k)^k - (z-hm^k)^k) .$$

Then for some integers c_j with $c_j \ll 1$, we have

$$\Psi_1(z,h,m) = \sum_{j=0}^{k-1} c_j z^j h(hm^k)^{k-j-1}.$$

Note that $c_{k-1} = 2k$.

Let $T_s(P,R,\theta)$ denote the number of solutions of the equation

$$x^k + m^k(x_1^k + \ldots + x_{s-1}^k) = y^k + m^k(y_1^k + \ldots + y_{s-1}^k)$$

with

$$x,y \le P,\ x \equiv y \pmod{m^k},\ M < m \le \mathrm{Min}\{P,MR\} ,$$

$$x_j, y_j \in \mathcal{A}(Q,R)\ (j = 1,\ldots,s-1) .$$

A bound for $S_s(P,R)$ may be obtained from a bound for $T_s(P,R,\theta)$ by using Lemma 2.1 of Vaughan [1989a] (see Vaughan [1989a], §4). Let

$$F_1(\alpha) = \sum_{h\le H} \sum_{M<m\le MR} S(\alpha,h,m) ,$$

where

$$S(\alpha,h,m) = \sum_{z\le 2P} e(\alpha\Psi_1(z,h,m)) .$$

Then by Vaughan [1989a], equation (2.36), we have

$$T_s(P,R,\theta) \ll PMR.S_{s-1}(Q,R) + \int_0^1 F_1(\alpha).|f(2^k\alpha;Q,R)|^{2s-2}d\alpha . \qquad (10.1)$$

One approach to bounding $T_s(P,R,\theta)$ is through estimates for

$$I(s) = \int_0^1 |F_1(\alpha)|^{2s}d\alpha .$$

In Vaughan [1989c] §2, estimates for $I(2^{j-1})$ are obtained by adopting a Hua's lemma approach, giving

$$I(2^{j-1}) \ll P^{2^j-j+\varepsilon}(MRH)^{2^j-1}(MR)^{e_j} \quad \text{for } j = 1,\ldots,k-2,$$

in which for $k \ge 9$ we have

$$e_j = \begin{cases} 0 & \text{for } k \text{ odd and } j = 2 \\ 1 & \text{otherwise.} \end{cases}$$

Motivated by this idea, we now use estimates arising from Vinogradov's mean value theorem to improve on this estimate for $k \ge 12$ and s not too small.

Let $J_{s,k}(P)$ denote the number of solutions to the simultaneous diophantine equations

$$x_1^j + \ldots + x_s^j = y_1^j + \ldots + y_s^j \quad (1 \le j \le k)$$

with $1 \le x_i, y_i \le P$ $(1 \le i \le s)$. Estimates available for $J_{s,k}(P)$ are of the form

$$J_{s,k}(P) \ll P^{2s-k(k+1)/2+\delta}$$

where $\delta = \delta(s,k)$ is small for s large compared with k. We shall explore the available estimates after the following lemma.

Lemma 10.1. *Suppose that* $k \ge 3$. *Then*

$$I(s) \ll P^{2s-(k-1)+\delta(s,k-1)+\varepsilon}(MRH)^{2s}H^{f_s},$$

where

$$f_s = \begin{cases} -1 & \text{if } s \text{ is even and } 2\delta(s/2,k-1)+1-k\theta \le \delta(s,k-1)+k-1 \\ 0 & \text{otherwise .} \end{cases}$$

Proof: Suppose first that $f_s = -1$, say with $s = 2t$. Then by Hölder's inequality, we have

$$\begin{aligned} I(s) &= \int_0^1 \Big| \sum_{M<m\le MR} \sum_{h\le H} \sum_{z\le 2P} e(\alpha\Psi_1(z,h,m)) \Big|^{4t} d\alpha \\ &\le (MR)^{2s-1}H^{2s-2} \sum_{M<m\le MR} J_1(t,m) \end{aligned} \tag{10.2}$$

where

$$J_1(t,m) = \int_0^1 \Big(\sum_{h\le H} \Big| \sum_{z\le 2P} e(\alpha\Psi_1(z,h,m)) \Big|^{2t} \Big)^2 d\alpha .$$

Now $J_1(t,m)$ is the number of solutions of the diophantine equation

$$\sum_{i=1}^{t} \Big(\Psi_1(z_{2i-1},h,m) - \Psi_1(z_{2i},h,m) \Big)$$

$$= \sum_{i=1}^{t} \Big(\Psi_1(z'_{2i-1},h',m) - \Psi_1(z'_{2i},h',m) \Big)$$

with $z_i, z'_i \le 2P$ $(1 \le i \le t)$ and $1 \le h, h' \le H$.

For a given integer n, let $R(n,h,m)$ denote the number of solutions to the equation

$$\sum_{i=1}^{t} \Big(\Psi_1(z_{2i-1},h,m) - \Psi_1(z_{2i},h,m) \Big) = n ,$$

with $z_i \le 2P$ for $i = 1,\dots,2t$. For each i we have $h|\Psi_1(z_i,h,m)$, so that $h|n$, and hence we have

$$J_1(t,m) \le \Big(\sum_{h\le H} R(0,h,m)\Big)^2 + 2\sum_{1\le n\le N}\Big(\sum_{h|n} R(n,h,m)\Big)^2, \tag{10.3}$$

where

$$N = t \max_{z,h,m} \Psi_1(z,h,m) \ll H(PR^k)^{k-1}.$$

By Cauchy's inequality, and by using the well-known estimate for the divisor function, we have

$$\sum_{1\le n\le N}\Big(\sum_{h|n} R(n,h,m)\Big)^2 \ll H^{\varepsilon}\sum_{h\le H}\sum_{1\le n\le N} R(n,h,m)^2. \tag{10.4}$$

But

$$\sum_{1\le n\le N} R(n,h,m)^2$$

is bounded above by the number of solutions to the equation

$$\sum_{i=1}^{t}\Big(\Psi_1(z_{2i-1},h,m) - \Psi_1(z_{2i},h,m)\Big) = \sum_{i=1}^{t}\Big(\Psi_1(z'_{2i-1},h,m) - \Psi_1(z'_{2i},h,m)\Big)$$

with $z_i, z'_i \le 2P$. Thus

$$\sum_{1\le n\le N} R(n,h,m)^2 \le \sum_{\underline{\underline{A}}} J_2(t,m,h,\underline{\underline{A}}), \tag{10.5}$$

where $J_2(t,m,h,\underline{\underline{A}})$ is the number of solutions of the simultaneous equations

$$\sum_{i=1}^{2t}(x_i^j - y_i^j) = A_j \quad (1 \le j \le k-1)$$

with $1 \le x_i, y_i \le 2P$, and where $\sum_{\underline{\underline{A}}}$ denotes summation over all A_j satisfying

$$|A_j| \le 2t(2P)^j \ (1 \le j \le k-2), \text{ and}$$

$$c_{k-1}A_{k-1} = -\sum_{j=1}^{k-2} c_j (hm^k)^{k-j-1} A_j .$$

Then we have

$$J_2(t,m,h,\underline{\underline{A}}) = \int_{\mathcal{U}_{k-1}^*} |G(\underline{\underline{\alpha}})|^{2s} e(-A_1\alpha_1 - \dots - A_{k-1}\alpha_{k-1})\, d\underline{\underline{\alpha}},$$

where $\mathcal{U}_{k-1}^* = [0,1]^{k-1}$, and

$$G(\underline{\underline{\alpha}}) = \sum_{x\le 2P} e(\alpha_1 x + \alpha_2 x^2 + \dots + \alpha_{k-1}x^{k-1}).$$

Then

$$J_2(t,m,h,\underline{\underline{A}}) \le \int_{\mathfrak{U}_{k-1}^*} |G(\underline{\underline{\alpha}})|^{2s}\, d\underline{\underline{\alpha}} = J_{s,k-1}(2P) ,$$

and hence

$$\sum_{\underline{\underline{A}}} J_2(t,m,h,\underline{\underline{A}}) \ll P^{(k-1)(k-2)/2} . J_{s,k-1}(2P) . \tag{10.6}$$

We may treat $R(0,h,m)$ in a similar way, obtaining

$$R(0,h,m) \ll P^{(k-1)(k-2)/2} J_{t,k-1}(2P) . \tag{10.7}$$

Thus, by (10.3)-(10.7) we have

$$J_1(t,m) \ll (HP^{(k-1)(k-2)/2} J_{t,k-1}(2P))^2 + H^{1+\varepsilon} P^{(k-1)(k-2)/2} J_{s,k-1}(2P) ,$$

and hence by (10.2),

$$I(s) \ll (MR)^{2s} H^{2s-1+\varepsilon} P^{(k-1)(k-2)/2} J_{s,k-1}(P) ,$$

provided that

$$H.P^{(k-1)(k-2)/2} (J_{t,k-1}(2P))^2 \ll J_{s,k-1}(2P) .$$

The lemma now follows in the case that $f_s = -1$.

When $f_s = 0$, we use Hölder's inequality to obtain

$$I(s) \le (HMR)^{2s-1} \sum_{M<m\le MR} \sum_{h\le H} \int_0^1 \Big| \sum_{z\le 2P} e(\alpha\Psi_1(z,h,m)) \Big|^{2s} d\alpha ,$$

and the result now follows in a similar, though simpler manner. This completes the proof of the lemma.

We now turn to the matter of the available estimates for $\delta(u,k\text{-}1)$. The simplest result to use is what might be termed the "classical" mean value theorem of Vinogradov. Thus, by Theorem 5.1 of Vaughan [1981b] it is permissable to use

$$\delta(wk,k) = \tfrac{1}{2}k^2(1-1/k)^w \text{ for } w \in \mathbb{N} .$$

A slight improvement may be obtained by using some of the more recent refinements of this estimate. For example, Turina [1987], Theorem 1, gives the following scheme.

Recursively define the sequences r_n, s_n and $\Delta_n = \Delta(s_n)$ as follows. Set $s_1 = \Delta_1 = k$. For $n \geq 1$, we define r_n to be the integer nearest to the number

$$(2\Delta_n + k(k+1))/(2k+1) ,$$

and define

$$\Delta_{n+1} = \Delta_n(1 - 1/r_n) + r_n - (2r_n-k-1)(2r_n-k)/(2r_n) ,$$

$$s_{n+1} = s_n + r_n .$$

We are then able to take $\delta(s_n,k) = \frac{1}{2}k(k+1) - \Delta_n$ for each $n \geq 1$.

When η is sufficiently small, the bound given by Lemma 10.1 may be converted, as in §4 of Vaughan [1989a] and §2 of Vaughan [1989c], into an inequality for λ_s, where λ_s satisfies

$$S_s(P,R) \ll P^{\lambda_s+\varepsilon} .$$

We find that λ_s is permissable whenever for some r with $r \geq s$, there is a θ satisfying $0 < \theta \leq 1/k$ such that

$$\lambda_s \geq (2s-2)\theta + 1 + \lambda_{s-1}(1-\theta)$$

$$\lambda_s \geq \frac{(2s-2-k-kf_r/2r)\theta + 2 - ((k-1)-\delta-f_r)/2r + \lambda_{s-1}(1-\theta)(1-s/2r)}{1 - (1-\theta)(s-1)/2r} ,$$

where $\delta = \delta(r,k-1)$.

Define $\mathfrak{m}$ to be the set of points α in $[0,1]$ with the property that whenever $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and

$$k(k-1)3^k qP^{k-2}HR^{k(k-2)}|\alpha-a/q| \leq 1 ,$$

then we have $q > P$. Further, let

$$\mathfrak{M} = [0,1]\backslash\mathfrak{m} .$$

We now give a minor arc estimate for $F_1(\alpha)$ which for large k improves upon that of Vaughan [1989a], §3. We do this by using estimates given by Vinogradov's Mean Value Theorem.

Lemma 10.2 (*see Vaughan* [1981b], *Theorems* 5.2 *and* 5.3). *Define*

$$g(\underline{\alpha}) = \sum_{1 \le x \le N} e(\alpha_1 x + \ldots + \alpha_k x^k) .$$

Suppose that there exist j, a, q *with* $2 \le j \le k$, $|\alpha_j - a/q| \le q^{-2}$, $(a,q) = 1$, $q \le N^j$. *Then for each natural number* w, *we have*

$$g(\underline{\alpha}) \ll N(N^{\delta}(qN^{-j} + N^{-1} + q^{-1}))^{1/(2(k-1)w)} \log(2N) ,$$

where

$$\delta = \tfrac{1}{2}(k-1)^2 \left(\frac{k-2}{k-1} \right)^w .$$

We define

$$\sigma(k) = \operatorname*{Max}_w \frac{1}{2(k-1)w} \left(1 - \tfrac{1}{2}(k-1)^2 \left(\frac{k-2}{k-1} \right)^w \right) ,$$

and write σ for $\sigma(k-1)$.

Lemma 10.3. *Suppose that* $k \ge 3$, *and* $\alpha \in \mathfrak{m}$. *Then*

$$F_1(\alpha) \ll P^{1-\sigma+\varepsilon} H M R^{k^2} .$$

Proof: Let $C = \operatorname*{Max}_{0 \le j \le k-1} |c_j|$, and let $L = C.2k^2 P^{k-2} R^{k(k-1)}$. We define

$$\mathfrak{M}(q,a) = \{ \alpha : |\alpha q - a| \le L^{-1} \}$$

and we write $\mathfrak{M}$ for the union of the $\mathfrak{M}(q,a)$ with $(a,q) = 1$ and $1 \le a \le q \le P$. Also, we denote the minor arcs by $\mathfrak{n} = [0,1] \setminus \mathfrak{M}$.

We start by considering $S(\alpha,h,m)$ for fixed h and m satisfying $1 \le h \le H$ and $M < m \le MR$.

Suppose first that $h\alpha \in \mathfrak{n}$. By Dirichlet's theorem on diophantine approximation, we may choose a and q with $(a,q) = 1$ and $1 \le q \le L$ so that $|2khq\alpha - a| \le L^{-1}$. Then since $h\alpha \in \mathfrak{n}$ we must have $2kq > P$, so that by Lemma 10.2, we have

$$S(\alpha,h,m) \ll P^{1-\sigma+\varepsilon} R^{k(k-1)} .$$

Suppose now that $h\alpha \in \mathfrak{M}$. Then there exist integers a and q with $(a,q) = 1$, $1 \le q \le P$, and

$$|h\alpha - a/q| \le (qL)^{-1} .$$

But then, for $j = 1,\dots,k-1$ we have

$$|c_j(hm^k)^{k-j-1}.h\alpha - c_j(hm^k)^{k-j-1}a/q| \le |c_j|.(qL)^{-1}(PR^k)^{k-j-1}$$
$$\le (2k^2)^{-1}q^{-1}P^{1-j}.$$

On noting that

$$(q,2ka,c_{k-2}(hm^k)a,\dots,c_1(hm^k)^{k-2}a) \le (q,2ka) \ll 1$$

we deduce from R.Baker [1986], Lemma 4.4 that

$$S(\alpha,h,m) = q^{-1}S(q,a)v(\beta) + O(q^{1-1/(k-1)+\varepsilon}) ,$$

where

$$\beta = \alpha - a/q ,$$

$$S(q,a) = \sum_{r=1}^{q} e(\Psi_1(r,h,m)a/q)$$

and

$$v(\beta) = \int_0^{2P} e(\Psi_1(\gamma,h,m)\beta)\, d\gamma .$$

By Theorem 2 of Hua [1965] and Theorem 7.3 of Vaughan [1981b], we have

$$S(q,a) \ll q^{1-1/(k-1)+\varepsilon}$$

and

$$v(\beta) \ll P(1 + |2kh\beta|P^{k-1})^{-1/(k-1)}.$$

Thus

$$S(\alpha,h,m) \ll P^{1-1/(k-1)+\varepsilon} + P^{1+\varepsilon}(q + P^{k-1}\|2khq\alpha\|)^{-1/(k-1)}.$$

Then on noting that for $k \ge 3$ we have $\sigma \le 1/(k-1)$, we have uniformly for $\alpha \in [0,1]$ the estimate

$$S(\alpha,h,m) \ll P^{1-\sigma+\varepsilon}R^{k(k-1)} + P^{1+\varepsilon}\Big(\sum_{q\le P}(q + P^{k-1}\|2khq\alpha\|)^{-1}\Big)^{1/(k-1)}.$$

By using Hölder's inequality we deduce that

$$\sum_{h\le H}|S(\alpha,h,m)| \ll HP^{1-\sigma+\varepsilon}R^{k(k-1)} + P^{1+\varepsilon}H^{1-1/(k-1)}T^{1/(k-1)}, \qquad (10.8)$$

where

$$T = \sum_{h\le H}\sum_{q\le P}(q + P^{k-1}\|2khq\alpha\|)^{-1}.$$

Suppose that $\alpha \in \mathfrak{m}$. Choose b and r with $(b,r) = 1$,

$$1 \le r \le k(k-1)3^k P^{k-2} H R^{k(k-2)}$$

and

$$|\alpha r - b| . k(k-1)3^k P^{k-2} H R^{k(k-2)} \le 1 .$$

Then by the definition of $\mathfrak{m}$, we have $r > P$. But by using Lemma 2.2 of Vaughan [1981b], we obtain

$$P^{k-1}T \ll P^{\varepsilon} \sum_{x \le 2kPH} \text{Min } \{ 2kP^{k-1}H.x^{-1}, \|\alpha x\|^{-1} \}$$

$$\ll P^{k-1+2\varepsilon} H . (r^{-1} + P^{2-k} + r(P^{k-1}H)^{-1}) .$$

Thus, because $P < r \ll P^{k-2}HR^{k(k-2)}$, the last expression is

$$\ll P^{k-2+2\varepsilon} H R^{k(k-2)},$$

and hence

$$T \ll P^{-1+2\varepsilon} H R^{k(k-2)}.$$

Then by (10.8), we have

$$\sum_{M < m \le MR} \sum_{h \le H} |S(\alpha, h, m)| \ll P^{1-\sigma+\varepsilon} H M R^{k^2} + P^{1-1/(k-1)+3\varepsilon} H M R^{2k}.$$

On noting that for $k \ge 3$ we have $\sigma \le 1/(k-1)$, this completes the proof of the lemma.

It may now be readily verified, as in §4 of Vaughan [1989a], that when η is sufficiently small, the minor arc estimate given by Lemma 10.3 may be converted into a bound for λ_s, where λ_s satisfies

$$S_s(P,R) \ll P^{\lambda_s + \varepsilon}.$$

We find that when $s \ge 2k-2$, we have

$$\lambda_s = \text{Max } \{ (2s-2)\theta + 1 + \lambda_{s-1}(1 - \theta) , \ 2s-k \} ,$$

in which $\theta = (1-\sigma)/k$.

11. THE PROOF OF THEOREM 1.2.

In Table 11.1 we list the optimal values of λ_s, for selected values of s, which arise from one of the methods of §10 of this chapter, or from the methods of Vaughan [1989a,c], when $10 \le k \le 20$. We note that Turina's improvement on Vinogradov's estimate was used only in the case $k = 14$. This was the only case in which the improvements in the estimates proved worthwhile, Turina's improvement being in all cases rather marginal. The additional saving obtained by taking $f_s = -1$ instead of $f_s = 0$, when permissable, in Lemma 10.1 also proved to be inconsequential, and so was dropped so as to simplify the computation.

It may be helpful to sketch the course of the iteration process when $k > 12$. For each value of s, each method available to us was tested in turn to find the optimal permissable choice for λ_s. For $s = 3$ and 4, the estimates given by Theorem 1.4 of Vaughan [1989c] are always the best available. For s relatively small (up to about $\frac{1}{2}k$), the estimates of Lemma 2.2 of Vaughan [1989c] are usually superior to the estimates given in §4 of Vaughan [1989a]. The latter estimates take over for s up to about k, after which estimates stemming from Lemma 10.1 prove superior. In the very final stages of the iteration process, estimates arising from Lemma 10.3 prove the best available to us. The table of values was computed to 16 significant figures by using an electronic computer, the final significant figure being rounded up.

Table 11.1

k	s	λ_s	k	s	λ_s
10	32	54.226509	11	38	65.213049
	39	68.086265		45	79.098077
12	44	76.201666	13	48	83.224720
	52	92.094246		58	103.096585
14	54	94.210152	15	60	105.199880
	64	114.096388		70	125.097201
16	66	116.191810	17	72	127.185564
	77	138.091507		83	149.092824
18	76	134.203397	19	82	145.197918
	90	162.088867		96	173.090628
20	88	156.193804			
	103	186.087721			

Let $P = n^{1/k}$, $X = P^{k/(2k-1)}$, $Z = PX^{-1}$, and define the generating function h by

$$\mathcal{C} = \{ x : x = pz,\ X/2 < p \le X,\ z \in \mathcal{A}(Z, Z^{\eta}) \} ,$$

$$h(\alpha) = \sum_{x \in \mathcal{C}} e(\alpha x^k) .$$

Table 11.2.

k	$s(k)$	$\sigma(k,s)$	k	$s(k)$	$\sigma(k,s)$
10	32	0.00654723	16	66	0.00320695
11	38	0.00555735	17	72	0.00295264
12	44	0.00483284	18	76	0.00273350
13	48	0.00429307	19	82	0.00254408
14	54	0.00386420	20	88	0.00237729
15	60	0.00350623			

Define $s = s(k)$ as in Table 11.2. Since s is even we may write $s = 2r$ for some integer r. Now let $\mathfrak{n}$ denote the set of real numbers α with the property that whenever $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and $|\alpha - a/q| \le q^{-1}X^{1-k}(rZ^k)^{-1}$, one has $q > X$. Then as in Vaughan [1989a], §9, we have for each $\eta > 0$ sufficiently small,

$$\operatorname*{Sup}_{\alpha \in \mathfrak{n}} |h(\alpha)| \ll P^{1-\sigma+\varepsilon},$$

where

$$\sigma = \sigma(k) = \frac{(k-1)(2s-\lambda_s) - k(k-2)}{2s(2k-1)} .$$

Further, by choosing $\phi = k/(2k-1)$, we have that m (as defined for the purposes of Theorem 1.4) satisfies

$$m \leq n .$$

The values of σ given by the choices of s listed in Table 11.2 are also listed in Table 11.2, rounded down in the final decimal place. The $s(k)$ were chosen so as to give the maximum values of σ.

Table 11.3.

k	$u(k)$	$t(k)$	k	$u(k)$	$t(k)$	k	$u(k)$	$t(k)$
10	39	14	14	64	25	18	90	33
11	45	18	15	70	28	19	96	36
12	52	20	16	77	29	20	103	37
13	58	23	17	83	32			

Now let $u = u(k)$ and $t = t(k)$ be as given by Table 11.3. Then for $10 \leq k \leq 20$, by Tables 11.2 and 11.3 we have

$$\lambda_u + t(1-\sigma) < 2u + t - k ,$$

and

$$u \geq 2k+1 .$$

Thus the conditions of Theorem 1.4 are satisfied, and we may conclude that

$$G(k) \leq 2u+t .$$

This completes the proof of Theorem 1.2.

CHAPTER 6.

LARGE IMPROVEMENTS IN WARING'S PROBLEM.

1. INTRODUCTION.

Let $G(k)$ denote the smallest number s such that every sufficiently large natural number is the sum of at most s kth powers of natural numbers.

In this chapter we extend the new iterative method of Vaughan [1989a, 1989c]. The new method permits us to improve substantially all previous upper bounds for $G(k)$ when $k \geq 6$.

Table 1.1

k	$F(k)$	k	$F(k)$	k	$F(k)$	k	$F(k)$
6	27	10	63	14	95	18	129
7	36	11	70	15	103	19	138
8	47	12	79	16	112	20	146
9	55	13	87	17	120		

We obtain

Theorem 1.1. *When* $6 \leq k \leq 20$ *we have* $G(k) \leq F(k)$, *where* $F(k)$ *is given by Table* 1.1.

This may be compared with the respective bounds $G(6) \leq 28$ (Chapter 5), $G(7) \leq 41$ (Vaughan [1989a]), $G(8) \leq 57$ (Vaughan [1989c]), $G(9) \leq 75$ (Vaughan [1989a]), $G(10) \leq 92$, $G(11) \leq 108$, $G(12) \leq 124$, $G(13) \leq 139$, $G(14) \leq 153$, $G(15) \leq 168$, $G(16) \leq 183$, $G(17) \leq 198$, $G(18) \leq 213$, $G(19) \leq 228$, $G(20) \leq 243$ (Chapter 5).

We remark that the above results are intended as something of a demonstration of the power of the method, and improvements may be

obtained by refining the method, especially for smaller k. This is a matter we intend to return to in papers subsequent to this thesis.

For larger k, it is apparent that the new method almost halves the previous upper bounds, and it is natural to ask whether or not this phenomenon persists for very large k. The answer is in the affirmative, so that at last we can substantially improve on the upper bound $G(k) \le (2+o(1))k\log k$ due to Vinogradov [1959]. We now obtain:

Theorem 1.2. *Suppose that* $k \ge 3$. *Define the real numbers* θ_s, $\Delta(s)$, $\rho(s)$ $(s = 2,3,\dots)$ *successively by*

$$\theta_2 = 0,\ \Delta(2) = k-2,\ \rho(2) = 0,$$

and for $s > 2$ *by*

$$\theta_s = \frac{1}{k+\Delta(s-1)} + \left[\frac{1}{k} - \frac{1}{k+\Delta(s-1)}\right]\left[\frac{k-\Delta(s-1)}{2k}\right]^{k-1}, \tag{1.1}$$

$$\Delta(s) = \Delta(s-1).(1 - \theta_s) + k\theta_s - 1\ , \tag{1.2}$$

$$\rho(s) = \frac{1}{4s}\,(1 - \Delta(s))\ . \tag{1.3}$$

Let

$$\rho = \max_{s>2} \rho(s)\ . \tag{1.4}$$

Then

$$G(k) \le 3 + \min_{\substack{v \in \mathbb{N} \\ v \ge k+1}} \left\{ 2v + 2\left\lceil \frac{\Delta(v)}{2\rho} \right\rceil \right\}. \tag{1.5}$$

Corollary 1.2.1. *We have* $G(k) \le (1 + o(1))k\log k$. *To be more precise, as* $k \longrightarrow \infty$ *we have*

$$G(k) < k(\log k + \log\log k + O(1))\ .$$

Theorem 1.2 may be regarded as an improvement on Theorem 1.6 of Vaughan [1989a]. The latter gives, as $k \longrightarrow \infty$,

$$G(k) < 2k(\log k + \log\log k + 1 + \log 2 + O(\log\log k/\log k)).$$

Let $G^+(k)$ denote the smallest number s such that almost all sufficiently large natural numbers are the sum of at most s kth powers of natural numbers. Previously, the list of essentially best possible results in Waring's problem has consisted of the results $G(2) = 4$ (Lagrange), $G^+(3) = 4$ (Davenport [1939a]), $G(4) = 16$ (Davenport [1939b]), $G^+(4) = 15$ (Hardy and Littlewood [1925]), and $G^+(8) = 32$ (Vaughan [1986b]). We are able to add to this list.

Theorem 1.3. *We have* $G^+(16) = 64$ *and* $G^+(32) = 128$.

As a simple deduction from the results of our method, we are also able to improve a constant in a result of Vaughan:

Theorem 1.4. *Suppose that* $0 < \delta < 1/(2k)$, *and let* m *denote the set of real numbers* α *with the property that whenever* $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ *and*

$$|\alpha - a/q| \le q^{-1}P^{1/2 + \delta k - k}$$

one has $q > P^{1/2 + \delta k}$. *Further, let* $\mathcal{A}(P,R)$ *denote the set of natural numbers not exceeding* P *with no prime divisor exceeding* R. *Then there is a real number* $\rho(k)$, *given by* (1.4), *such that for each positive number* ε, *there is a positive number* η *such that whenever* $2 \le R \le P^{\eta}$ *the exponential sum*

$$S(\alpha) = \sum_{x \in \mathcal{A}(P,R)} e(\alpha x^k)$$

satisfies

$$\operatorname*{Sup}_{m} |S(\alpha)| \ll P^{1+\varepsilon}(P^{-\delta} + P^{-\rho(k)}) .$$

Further, as $k \longrightarrow \infty$ *we have*

$$2\rho(k)k\log k \ge 1 + o(1).$$

The concluding remark may be compared with Theorem 1.8 of Vaughan [1989a], where the concluding remark is $4\rho(k)k\log k \longrightarrow 1$ as $k \longrightarrow \infty$.

As is usual with results on Waring's problem, the methods we introduce can be applied equally well to diagonal forms and simultaneous additive equations (see the introduction of Vaughan [1989a]). The methods of this chapter may be used for simultaneous additive equations of differing degree, a matter we intend to return to in a paper subsequent to this thesis (but see Part II of this thesis).

As has become standard in applications of the Hardy-Littlewood method, bounding $G(k)$ depends fundamentally on estimates for the number of solutions of auxiliary equations of the form

$$x_1^k + \ldots + x_s^k = y_1^k + \ldots + y_s^k \tag{1.6}$$

with the x_j and y_j lying in certain subsets $\mathcal{A}_j$ of $[1,P]\cap\mathbb{Z}$. Vaughan's substantial innovation was to set each of the $\mathcal{A}_j$ equal to a set $\mathcal{A}$ with suitable arithmetic properties, and to use these properties to relate the number of solutions of the equation (1.6) to the number of solutions of the equation

$$x^k + m^k(u_1^k + \ldots + u_{s-1}^k) = y^k + m^k(v_1^k + \ldots + v_{s-1}^k) \tag{1.7}$$

with $(xy,m) = 1$, $x,y \le P$, $M < m \le M'$, and $u_j,v_j \in \mathcal{B}$, where $\mathcal{B}$ has similar properties to $\mathcal{A}$, but with $\mathcal{B} \subset [1,P/M]\cap\mathbb{Z}$. The condition $(xy,m) = 1$ enables us, roughly speaking, to assume that $x \equiv y \pmod{m^k}$, and we are effectively able to take first differences in a very efficient manner, by considering $\Psi(x,y) = m^{-k}(x^k-y^k)$ (which is, of course, an integer).

A constraint on the method is the absence of estimates for exponential sums depending on $\Psi(x,y)$ substantially better than

either Weyl differencing or Hua-type lemmata. Our new idea is to further exploit the arithmetic properties of $\mathcal{A}$ so as to continue taking differences, each difference being taken nearly as efficiently as the first. This new method puts no obstacles in the way of previous applications of Vaughan's iterative method, since the improvements effected do not change the character of the auxiliary equations used.

In §2 we establish the fundamental inequality between the numbers of solutions of equations (1.6) and (1.7). In §3 we demonstrate how this relationship may be used to take successive differences, and go on to derive a simple method for bounding the number of solutions of the equation (1.6). The situation where k is large is then considered in §4, thereby proving Theorems 1.2 and 1.4. In §§5-7 we consider k of intermediate size, proving Theorem 1.1, and in §8 we prove the "almost all" results which give Theorem 1.3.

Notation: Unless otherwise stated upper case Latin letters denote real numbers exceeding 2, lower case Latin letters denote integers, and lower case Greek letters denote positive real numbers. We use p to denote a prime number, and k, n and s to denote positive integers. Throughout, ε denotes a sufficiently small positive number, and implicit constants depend at most on k, s, t and ε. Unless otherwise stated, k is considered to be fixed.

2. THE FUNDAMENTAL LEMMA.

Let

$$\mathcal{A}(P,R) = \{\ n : n \le P,\ p|n \text{ implies } p \le R\ \}\ , \tag{2.1}$$

and let $S_s(P,R)$ denote the number of solutions of the equation

$$x_1^k + \ldots + x_s^k = y_1^k + \ldots + y_s^k \tag{2.2}$$

with

$$x_j, y_j \in \mathcal{A}(P,R) \quad (1 \le j \le s). \tag{2.3}$$

Let $\Psi(z,\underline{c})$ denote a polynomial with integer coefficients in the variables $z,\ c_1,\ldots,c_t$ of degree in z at least one, and write $\Psi'(z,\underline{c})$ for $\frac{\partial\Psi}{\partial z}(z,\underline{c})$. Let P, Q and R be positive real numbers with $R \le Q \le P$, and C_i, C_i' $(i = 1,\ldots,t)$ be real numbers with $1 \le C_i' \le C_i \ll P$. Denote by

$$S_s(P,Q,R) = S_s(P,Q,R;\Psi;\underline{C},\underline{C}')$$

the number of solutions of the equation

$$\Psi(z,\underline{c}) + x_1^k + \ldots + x_s^k = \Psi(z',\underline{c}') + y_1^k + \ldots + y_s^k \tag{2.4}$$

with

$$x_j, y_j \in \mathcal{A}(Q,R) \quad (1 \le j \le s), \tag{2.5}$$

$$1 \le z, z' \le P, \text{ and } C_i' < c_i, c_i' \le C_i \quad (i = 1,\ldots,t)\ . \tag{2.6}$$

For a given real number θ with $1 < P^\theta < Q$, let

$$T_s(P,Q,R;\theta) = T_s(P,Q,R;\theta;\Psi;\underline{C},\underline{C}')$$

denote the number of solutions of the equation

$$\Psi(z,\underline{c}) + w^k(u_1^k + \ldots + u_s^k) = \Psi(z',\underline{c}) + w^k(v_1^k + \ldots + v_s^k) \tag{2.7}$$

with $z, z', \underline{c}$ as in (2.6), and

$$P^\theta < w \le \text{Min}\{Q, P^\theta R\}\ ,\ u_j, v_j \in \mathcal{A}(QP^{-\theta},R) \quad (j = 1,\ldots,s)\ , \tag{2.8}$$

$$\text{and } z \equiv z' \pmod{w^k}\ . \tag{2.9}$$

Finally, let

$$N_s(P,Q,R) = N_s(P,Q,R;\Psi;\underline{C},\underline{C}')$$

denote the number of solutions of the equation (2.4) subject to (2.5), (2.6), and also $\Psi'(z,\underline{c}) = \Psi'(z',\underline{c}') = 0$.

We shall require a lemma on the density of integers with a given square-free kernel. Given an integer v with canonical prime factorisation

$$\prod_{i=1}^{t} p_i^{r_i},$$

denote by $s_0(v)$ the square-free kernel of v, that is

$$\prod_{i=1}^{t} p_i .$$

Lemma 2.1. *Suppose that L is a positive real number, and r is a positive integer with $\log r \ll \log L$. Then for each $\varepsilon > 0$ we have*

$$\mathrm{card}\ \{\ y \le L : s_0(y) = s_0(r)\ \} \ll L^{\varepsilon}.$$

Proof: Suppose that $s_0(r) = p_1 \ldots p_n$. Then

$$n \ll \frac{\log r}{\log\log r} \ll \frac{\log L}{\log\log L},$$

(see, for example, Hardy and Wright [1979], §22.10).

Let $L > e^e$ and

$$U = \mathrm{card}\ \{\ y \le L : s_0(y) = s_0(r)\ \}.$$

Then U is bounded above by the number of solutions of the inequality

$$u_1 \log p_1 + \ldots + u_n \log p_n \le \log L,$$

with $u_i \in \mathbb{N}$ $(1 \le i \le n)$. But for a given integer m, the number of solutions of the equation

$$u_1 + \ldots + u_n = m$$

with $u_i \in \mathbb{N} \cup \{0\}$ $(1 \le i \le n)$, is

$$(-1)^m \binom{-n}{m} \le (m+n)^{n-1}/(n-1)!,$$

(see for example Vaughan [1981b], Exercise 1 of Chapter 1). Thus

$$U \ll \frac{(\log L/\log 2)^n}{(n-1)!}.$$

Then by Stirling's formula,

$$\log U \ll \mathcal{U}(n),$$

where

$$\mathcal{U}(t) = t(\text{loglog}L - \log t) + t .$$

Now, $\mathcal{U}(t)$ is an increasing function of t in an interval $[1,M]$ with $M \gg \log L$. But $n \ll \log L/\text{loglog}L$, and so

$$\log U \ll \log L.\,\text{logloglog}L/\text{loglog}L .$$

That is, $U \ll L^{\varepsilon}$ for each $\varepsilon > 0$.

This completes the proof of the lemma.

The lemma below relates S_s to T_s.

Lemma 2.2. *Let* $\theta = \theta(s,k;\Psi)$ *satisfy* $1 < P^{\theta} < Q$. *Then*

$$S_s(P,Q,R) \ll S_s(P,P^{\theta},R) + N_s(P,Q,R) + QP^{\theta+\varepsilon}S_{s-1}(P,Q,R)$$
$$+ P^{\varepsilon}\Big(\prod_{i=1}^{t} C_i\Big)(P^{\theta}R)^{2s-1}.T_s(P,Q,R;\theta) .$$

The implicit constant may depend on Ψ.

Proof: Write

$$x\ \mathcal{D}(L)\ y$$

to denote that there is some divisor d of x, with $d \le L$, such that x/d has all of its prime divisors amongst those of y.

Let S' denote the number of solutions of (2.4) satisfying (2.5) and (2.6) for which

$$\text{Min}\{x_j, y_j\} \le P^{\theta} \tag{2.10}$$

for at least one j, let S'' denote the number for which

$$\Psi'(z,\underline{c}) = 0 \text{ or } \Psi'(z',\underline{c}') = 0 , \tag{2.11}$$

let S''' denote the number for which $\text{Min}\{x_j, y_j\} > P^{\theta}$ for every j, (2.11) does not hold, and

$$x_j\ \mathcal{D}(P^{\theta})\ \Psi'(z,\underline{c}) \quad \text{or} \quad y_j\ \mathcal{D}(P^{\theta})\ \Psi'(z',\underline{c}') \tag{2.12}$$

for at least one j, and let S'''' denote the number for which $x_j > P^{\theta}$

and $y_j > P^\theta$ for every j, (2.11) does not hold, and (2.12) holds for no j.

Then

$$S_s(P,Q,R) \le 4.\mathrm{Max}\{S',S'',S''',S''''\} .$$

We divide into cases.

(i) Suppose that $S' \ge \mathrm{Max}\{S'',S''',S''''\}$, so that $S_s(P,Q,R) \le 4S'$.

Let

$$f(\alpha;L,R) = \sum_{x \in \mathcal{A}(L,R)} e(\alpha x^k) ,$$

and

$$F(\alpha;P,R) = \sum_{z,\underline{c}} e(\alpha\Psi(z,\underline{c})) ,$$

where the summation is over z and $\underline{c}$ satisfying (2.6). Then

$$S' \ll \int_0^1 |F(\alpha;P,R)^2 f(\alpha;P^\theta,R).f(\alpha;Q,R)^{2s-1}| \, d\alpha .$$

Then by Hölder's inequality, we have

$$S_s(P,Q,R) \ll (S_s(P,P^\theta,R))^{1/(2s)}(S_s(P,Q,R))^{1-1/(2s)},$$

and the result follows in the first case.

(ii) Suppose that $S'' \ge \mathrm{Max}\{S''',S'''',S'\}$, so that $S_s(P,Q,R) \le 4S''$.

Let

$$G(\alpha;P,R) = \sum_{z,\underline{c}} e(\alpha\Psi(z,\underline{c})) ,$$

where the summation is over z, $\underline{c}$ satisfying (2.6) subject to $\Psi'(z,\underline{c}) = 0$. Then we have

$$S'' \ll \int_0^1 |G(\alpha;P,R).F(\alpha;P,R).f(\alpha;Q,R)^{2s}| \, d\alpha .$$

Then by Schwarz's inequality,

$$S_s(P,Q,R) \ll (S_s(P,Q,R))^{1/2}(N_s(P,Q,R))^{1/2},$$

and so the result follows in the second case.

(iii) Suppose that $S''' \ge \mathrm{Max}\{S'''',S',S''\}$, so that $S_s(P,Q,R) \le 4S'''$.

Given z and $\underline{c}$ satisfying (2.6) with $\Psi'(z,\underline{c}) \ne 0$, denote by $\mathcal{Y}(z,\underline{c})$ the set of integers x such that $x \le Q$, and x has a divisor d, with

$d \le P^{\theta}$, such that x/d has all of its prime divisors amongst those of $\Psi'(z,\underline{c})$. Let

$$H(\alpha;P,Q,R) = \sum_{z,\underline{c}} \sum_{x \in \mathcal{P}(z,\underline{c})} e(\alpha(x^k + \Psi(z,\underline{c}))) ,$$

where the summation is over $z,\underline{c}$ satisfying (2.6) and $\Psi'(z,\underline{c}) \neq 0$. Then we have

$$S''' \ll \int_0^1 |H(\alpha;P,Q,R).F(\alpha;P,R).f(\alpha;Q,R)^{2s-1}| \, d\alpha ,$$

so that by Schwarz's inequality,

$$S''' \ll (S_s(P,Q,R))^{1/2}.\left\{\int_0^1 |H(\alpha;P,Q,R)^2.f(\alpha;Q,R)^{2s-2}| \, d\alpha\right\}^{1/2}.$$

Then

$$S_s(P,Q,R) \ll \sum_{r,r'} V(r,r') , \tag{2.13}$$

where $V(r,r')$ denotes the number of solutions of the equation

$$\Psi(z,\underline{c}) + d^k x^k + x_1^k + \ldots + x_{s-1}^k = \Psi(z',\underline{c}') + e^k y^k + y_1^k + \ldots + y_{s-1}^k ,$$

with z, $\underline{c}$, z', $\underline{c}'$, x_j and y_j satisfying (2.5), (2.6) subject to

$$\Psi'(z,\underline{c}) \neq 0, \ \Psi'(z',\underline{c}') \neq 0, \ r|\Psi'(z,\underline{c}), \ r'|\Psi'(z',\underline{c}') ,$$

$$1 \le d,e \le P^{\theta} , \ x \le Q/d, \ y \le Q/e, \ s_0(x) = r, \ s_0(y) = r'.$$

Let

$$G_r(\alpha;P,R) = \sum_{z,\underline{c}}^{*} e(\alpha\Psi(z,\underline{c})) , \tag{2.14}$$

where the summation is over z, $\underline{c}$ satisfying (2.6), and subject to $\Psi'(z,\underline{c}) \neq 0$ and $r|\Psi'(z,\underline{c})$. Then by (2.13), for some constants K and ω (the total degree of Ψ) we have

$$S_s(P,Q,R) \ll \int_0^1 |\mathcal{G}(\alpha)^2.f(\alpha;Q,R)^{2s-2}| \, d\alpha , \tag{2.15}$$

where

$$\mathcal{G}(\alpha) = \sum_{r \le KP^{\omega}} G_r(\alpha;P,R) \sum_{d \le P^{\theta}} \sum_{\substack{y \le Q/d \\ s_0(y)=r}} e(\alpha d^k y^k) . \tag{2.16}$$

Here, if r is not square-free, we understand the third summation in (2.16) to be empty. Now by Cauchy's inequality and (2.16), we have

$$|\mathcal{G}(\alpha)|^2 \le \Big(\sum_{r \le KP^{\omega}} |G_r(\alpha; P, R)|^2 \Big) \Big(\sum_{r \le KP^{\omega}} \Big| \sum_{d \le P^{\theta}} \sum_{\substack{y \le Q/d \\ s_0(y)=r}} e(\alpha d^k y^k) \Big|^2 \Big) \qquad (2.17)$$

Also, by interchanging the order of summation,

$$\sum_{r \le KP^{\omega}} \Big| \sum_{d \le P^{\theta}} \sum_{\substack{y \le Q/d \\ s_0(y)=r}} e(\alpha d^k y^k) \Big|^2 = \sum_{r \le KP^{\omega}} \Big| \sum_{\substack{y \le Q \\ s_0(y)=r}} \sum_{\substack{d \le P^{\theta} \\ d \le Q/y}} e(\alpha d^k y^k) \Big|^2,$$

and by Cauchy's inequality combined with Lemma 2.1, the latter expression is

$$\begin{aligned} &\ll P^{\varepsilon} \sum_{r \le KP^{\omega}} \sum_{\substack{y \le Q \\ s_0(y)=r}} \Big| \sum_{\substack{d \le P^{\theta} \\ d \le Q/y}} e(\alpha d^k y^k) \Big|^2 \\ &\le P^{\varepsilon} \sum_{r \le KP^{\omega}} \sum_{\substack{y \le Q \\ s_0(y)=r}} P^{\theta} Q/y \\ &\ll QP^{\theta+\varepsilon} \sum_{y \le Q} \frac{1}{y} \\ &\ll QP^{\theta+2\varepsilon} . \end{aligned} \qquad (2.18)$$

Then by (2.15)-(2.18), we have

$$S_s(P,Q,R) \ll QP^{\theta+2\varepsilon} \int_0^1 \Big(\sum_{r \le KP^{\omega}} |G_r(\alpha; P, R)|^2 \Big) . |f(\alpha; Q, R)|^{2s-2} \, d\alpha .$$

But by considering the underlying diophantine equation, the integral on the right hand side of the last inequality is

$$\ll P^{\varepsilon} . S_{s-1}(P,Q,R) ,$$

by using estimates for the divisor function. The result now follows in the third case.

(iv) Suppose that $S'' \ge \mathrm{Max}\{S', S'', S'''\}$, so that $S_s(P,Q,R) \le 4S''$. Then for a given solution of (2.4) satisfying (2.5) and (2.6) counted by S'', we have for every j

$$x_j > P^{\theta} \quad \text{and} \quad y_j > P^{\theta},$$

$$\Psi'(z, \underline{c}) \ne 0 \quad \text{and} \quad \Psi'(z', \underline{c}') \ne 0 ,$$

and neither

$$x_j \mathcal{D}(P^\theta)\ \Psi'(z,\underline{c}) \quad \text{nor} \quad y_j \mathcal{D}(P^\theta)\ \Psi'(z',\underline{c}') \ . \tag{2.19}$$

Let w be the greatest divisor of x_j with the property that $(w,\Psi'(z,\underline{c})) = 1$. If $w \le P^\theta$, then $x_j \mathcal{D}(P^\theta)\ \Psi'(z,\underline{c})$, contradicting (2.19). So $w > P^\theta$, and since each prime divisor of x_j is at most R, we may find a divisor w_j of x_j with

$$P^\theta < w_j \le \text{Min}\{Q, P^\theta R\}$$

and satsifying $(w_j,\Psi'(z,\underline{c})) = 1$. We may do likewise with the y_j.

We therefore deduce that

$$S''' \ll V_1 \ ,$$

where V_1 denotes the number of solutions of the equation

$$\Psi(z,\underline{c}) + \sum_{j=1}^{s} (w_j u_j)^k = \Psi(z',\underline{c}') + \sum_{j=1}^{s} (w'_j v_j)^k,$$

with $z,z',\underline{c},\underline{c}'$ satisfying (2.6), and for $j = 1,\dots,s$,

$$u_j \in \mathcal{A}(Q/w_j,R) \ , \ v_j \in \mathcal{A}(Q/w'_j,R) \ ,$$
$$P^\theta < w_j,w'_j \le \text{Min}\{Q,P^\theta R\} \ , \tag{2.20}$$

and

$$(w_j,\Psi'(z,\underline{c})) = (w'_j,\Psi'(z',\underline{c}')) = 1 \ .$$

Let

$$F_w(\alpha;P,R) = \sum_{z,\underline{c}}{}' e(\alpha\Psi(z,\underline{c})) \ ,$$

where the summation is over all $z,\underline{c}$ satisfying (2.6) subject to $(w,\Psi'(z,\underline{c})) = 1$, and let

$$F_j(\alpha) = f(w_j^k\alpha;Q/w_j,R).f(-w'^k_j\alpha;Q/w'_j,R) \ .$$

Then

$$V_1 \le \int_0^1 \sum_{\underline{w},\underline{w}'} F_w(\alpha;P,R).F_{w'}(-\alpha;P,R) \prod_{j=1}^{s} F_j(\alpha)\ d\alpha \ ,$$

where the summation is over $\underline{w},\underline{w}'$ satisfying (2.20), and where we have written $W = w_1\dots w_s$, and likewise for W'.

Let

$$X_j(\alpha) = |F_w(\alpha; P, R)^2 . f(w_j^k \alpha; Q/w_j, R)^{2s}| ,$$

and let $Y_j(\alpha)$ denote the analogous function appropriate to the w_j'. Then

$$S'' \ll \sum_{\underline{w}, \underline{w}'} \int_0^1 \prod_{j=1}^{s} \left[X_j(\alpha) . Y_j(\alpha) \right]^{1/(2s)} d\alpha .$$

By Hölder's inequality, we have

$$\int_0^1 \prod_{j=1}^{s} \left[X_j(\alpha) . Y_j(\alpha) \right]^{1/(2s)} d\alpha$$

$$\ll \prod_{j=1}^{s} \left(\int_0^1 X_j(\alpha)\ d\alpha \right)^{1/(2s)} \left(\int_0^1 Y_j(\alpha)\ d\alpha \right)^{1/(2s)} .$$

Now we observe that

$$\int_0^1 X_j(\alpha)\ d\alpha \le W(P, Q, R, w_j) ,$$

where $W(P, Q, R, w)$ denotes the number of solutions of the equation

$$\Psi(z, \underline{c}) + w^k(u_1^k + \ldots + u_s^k) = \Psi(z', \underline{c}') + w^k(v_1^k + \ldots + v_s^k)$$

with $z, z', \underline{c}, \underline{c}'$ as in (2.6),

$$u_j, v_j \in \mathcal{A}(QP^{-\theta}, R) \quad (j = 1, \ldots, s) ,$$

and

$$(\Psi'(z, \underline{c}), w) = (\Psi'(z', \underline{c}'), w) = 1 .$$

Therefore, by Hölder's inequality we have

$$\sum_{\underline{w}, \underline{w}'} \int_0^1 \prod_{j=1}^{s} \left[X_j(\alpha) . Y_j(\alpha) \right]^{1/(2s)} d\alpha$$

$$\ll \sum_{\underline{w}, \underline{w}'} \prod_{j=1}^{s} \left[W(P, Q, R, w_j) . W(P, Q, R, w_j') \right]^{1/(2s)}$$

$$\ll \left(\sum_{\underline{w}, \underline{w}'} 1 \right)^{\frac{2s-1}{2s}} \left(\sum_{\underline{w}, \underline{w}'} \prod_{j=1}^{s} \left[W(P, Q, R, w_j) . W(P, Q, R, w_j') \right] \right)^{1/(2s)}$$

$$\ll \prod_{j=1}^{2s} \left[(P^{\theta} R)^{2s-1} V(P, Q, R; \theta) \right]^{1/(2s)}$$

where $V(P,Q,R;\theta)$ denotes the number of solutions of the equation

$$\Psi(z,\underline{c}) + w^k(u_1^k + \ldots + u_s^k) = \Psi(z',\underline{c}') + w^k(v_1^k + \ldots + v_s^k) \qquad (2.21)$$

with (2.6), (2.8) and

$$(\Psi'(z,\underline{c}),w) = (\Psi'(z',\underline{c}'),w) = 1 .$$

It is now sufficient to show that

$$V(P,Q,R;\theta) \ll P^{\varepsilon}\Big(\prod_{i=1}^{t} C_i\Big).T_s(P,Q,R;\theta) .$$

For a given w satisfying (2.8), let $\mathcal{B}(u,\underline{c})$ denote the set of solutions of the congruence

$$\Psi(z,\underline{c}) \equiv u \pmod{w^k}$$

with $(\Psi'(z,\underline{c}),w) = 1$. Now if p is any prime divisor of w, and $p^r \| w^k$, then the number of solutions of

$$\Psi(z,\underline{c}) \equiv u \pmod{p^r}$$

with

$$(\Psi'(z,\underline{c}),p) = 1$$

is $O(1)$, since all such solutions are non-singular (mod p). Thus, by using the Chinese Remainder Theorem we deduce that

$$\text{card } \mathcal{B}(u,\underline{c}) \ll w^{\varepsilon} . \qquad (2.22)$$

Plainly in (2.21) we have

$$\Psi(z,\underline{c}) \equiv \Psi(z',\underline{c}') \pmod{w^k} .$$

Thus each solution of (2.21) may be classified according to the common residue class (mod w^k) of $\Psi(z,\underline{c})$ and $\Psi(z',\underline{c}')$. Let

$$g_w(\alpha;z,\underline{c}) = \sum_{\substack{x\le P \\ x\equiv z \pmod{w^k}}} e(\alpha\Psi(x,\underline{c})) .$$

Then by (2.21),

$$V(P,Q,R;\theta) \le \sum_{P^{\theta}<w\le \mathrm{Min}\{Q,P^{\theta}R\}} V_w ,$$

where

$$V_w = \int_0^1 G_w(\alpha).|f(w^k\alpha;QP^{-\theta},R)|^{2s}\,d\alpha , \qquad (2.23)$$

and

$$G_w(\alpha) = \sum_{u=1}^{w^k} \Big| \sum_{\underline{c}} \sum_{z \in \mathcal{B}(u,\underline{c})} g_w(\alpha; z, \underline{c}) \Big|^2 .$$

Hence, by Cauchy's inequality and (2.22) we have

$$G_w(\alpha) \ll w^{\varepsilon}\Big(\prod_{i=1}^{t} C_i \Big) \sum_{u=1}^{w^k} \sum_{\underline{c}} \sum_{z \in \mathcal{B}(u,\underline{c})} |g_w(\alpha; z, \underline{c})|^2$$

$$= w^{\varepsilon}\Big(\prod_{i=1}^{t} C_i \Big) \sum_{\underline{c}} \sum_{\substack{z=1 \\ (\Psi'(z,\underline{c}),w)=1}}^{w^k} |g_w(\alpha; z, \underline{c})|^2 .$$

The result now follows from (2.23) on considering the underlying diophantine equation.

This completes the proof of the lemma.

We now define the modified forward difference operator, Δ_1^*, by

$$\Delta_1^*(f(x); h; m) = m^{-k}(f(x+hm^k) - f(x)) ,$$

and define Δ_j^* recursively by

$$\Delta_{j+1}^*(f(x); h_1, \ldots, h_{j+1}; m_1, \ldots, m_{j+1})$$
$$= \Delta_1^*(\Delta_j^*(f(x); h_1, \ldots, h_j; m_1, \ldots, m_j); h_{j+1}; m_{j+1}) .$$

We also adopt the convention that

$$\Delta_0^*(f(x); h; m) = f(x) .$$

For $0 \le j \le k$ let

$$\Psi_j = \Psi_j(z; h_1, \ldots, h_j; m_1, \ldots, m_j) = \Delta_j^*(f(z); 2h_1, \ldots, 2h_j; m_1, \ldots, m_j)$$

where $f(z) = (z - h_1 m_1^k - \ldots - h_j m_j^k)^k$.

We now consider the effect of substituting $\Psi_j(z; \underline{h}; \underline{m})$ for $\Psi(z, \underline{c})$ in Lemma 2.2. We shall first require some notation.

Let $\phi_i = \phi_i(s,k)$ satisfy $0 < \phi_i \le 1/k$ for $i = 1, \ldots, k$, and write

$$\Phi_j = \phi_1 + \ldots + \phi_j , \quad M_j = P^{\phi_j}, \quad H_j = PM_j^{-k}$$

$$\text{and } Q_j = P^{1-\Phi_j} \ (1 \le j \le k) .$$

Given a constant $\kappa > 0$, substitute the conditions

$$M_i < m_i \le M_i R \ , \ 1 \le h_i \le \kappa H_i \quad (1 \le i \le j) \tag{2.24}$$

for

$$C_i' < c_i \le C_i \quad (1 \le i \le t)$$

in (2.6), and then write

$$S_s(P,Q,R;\Psi_j) = S_s(P,Q,R;\Psi_j;\kappa)$$

for $S_s(P,Q,R;\Psi;\underline{\underline{C}},\underline{\underline{C}}')$, and do likewise with T_s and N_s .

Lemma 2.3. *Suppose that for some positive real number, η, we have*

$$\exp((\log\log P)^2) < R \le P^{\eta} \ .$$

Then for $j = 0,\dots,k-1$, and any constant $\kappa > 0$, we have

$$S_s(P,Q_j,R;\Psi_j;\kappa) \ll P^{\varepsilon}\Big(\prod_{i=1}^{j} H_i M_i R\Big)(M_{j+1}R)^{2s-1}.T_s(P,Q_j,R;\phi_{j+1};\Psi_j;\kappa) \ .$$

(Here we adopt the convention that $Q_0 = P$) .

Proof: Write $\phi = \phi_{j+1}$, and consider the estimate given by Lemma 2.2.

Notice first that for the permissable values of z, $\underline{\underline{h}}$ and $\underline{\underline{m}}$ we have

$$\Psi_j'(z;\underline{\underline{h}};\underline{\underline{m}}) > 0 \ ,$$

and hence for $j = 0,\dots,k-1$ we have

$$N_s(P,Q,R;\Psi_j) = 0 \ . \tag{2.25}$$

By considering $S_s(P,P^{\phi},R;\Psi_j)$ in integral form, we may make a trivial estimate to obtain

$$S_s(P,P^{\phi},R;\Psi_j) \le P^{2\phi}S_{s-1}(P,Q_j,R;\Psi_j) \le Q_j M_{j+1} S_{s-1}(P,Q_j,R;\Psi_j) \ . \tag{2.26}$$

Thus far we have established, on combining (2.25) and (2.26) with Lemma 2.2, that

$$\begin{aligned} S_s(P,Q_j,R;\Psi_j) &\ll P^{\varepsilon}Q_j M_{j+1} S_{s-1}(P,Q_j,R;\Psi_j) \\ &\quad + P^{\varepsilon}\Big(\prod_{i=1}^{j} H_i M_i R\Big)(M_{j+1}R)^{2s-1}.T_s(P,Q_j,R;\phi;\Psi_j) \ . \end{aligned} \tag{2.27}$$

We now show inductively that for $s = 1,2,\ldots$, we have

$$Q_j M_{j+1} S_{s-1}(P,Q_j,R;\Psi_j)$$
$$\ll P^{2s\varepsilon}.\left(\prod_{i=1}^{j} H_i M_i R\right)(M_{j+1}R)^{2s-1}.T_s(P,Q_j,R;\phi;\Psi_j) \ . \qquad (2.28)$$

This will complete the proof of the lemma.

We first note that by using standard estimates from prime number theory, when P is sufficiently large we have for $j = 0,\ldots,k-1$,

$$\mathrm{card}(\mathcal{A}(Q_j,R)) > Q_j^{1-\varepsilon}, \qquad (2.29)$$

since $R > \exp((\log\log P)^2)$. Then for $s = 1$ we have

$$S_{s-1}(P,Q_j,R;\Psi_j) \ll P\left(\prod_{i=1}^{j} H_i M_i R\right)^2 ,$$

whilst by (2.29),

$$T_s(P,Q_j,R;\phi;\Psi_j) \gg PQ_j^{1-\varepsilon}\left(\prod_{i=1}^{j} H_i M_i R\right) ,$$

by considering diagonal solutions alone. Then (2.28) follows in the case $s = 1$. Now suppose that the inductive hypothesis holds for $t < s$. Then by combining the inductive hypothesis with (2.27), we have

$$S_{s-1}(P,Q_j,R;\Psi_j)$$
$$\ll P^{(2s-1)\varepsilon}\left(\prod_{i=1}^{j} H_i M_i R\right)(M_{j+1}R)^{2s-3}.T_{s-1}(P,Q_j,R;\phi;\Psi_j) \ . \qquad (2.30)$$

But by considering solutions of (2.7) in which $u_s = v_s$, we have from (2.29),

$$T_s(P,Q_j,R;\phi;\Psi_j) \gg (Q_j/M_{j+1})^{1-\varepsilon}.T_{s-1}(P,Q_j,R;\phi;\Psi_j) \ , \qquad (2.31)$$

and (2.28) now follows from (2.30) and (2.31). Thus (2.28) holds for each $s \geq 1$, and the lemma now follows.

3. SUCCESSIVE DIFFERENCING.

Lemma 2.3 permits us to relate $S_s(P,Q,R;\Psi_j)$ to $T_s(P,Q,R;\phi;\Psi_j)$. We now consider the simplest method of relating $T_s(P,Q,R;\phi;\Psi_j)$ to $S_s(P,Q',R;\Psi_{j+1})$. This allows us to obtain a relatively simple estimate for $S_s(P,P,R;\Psi_0)$, and hence for $S_{s+1}(P,R)$.

Lemma 3.1. *Adopting the same notation as that preceding Lemma 2.3, we have for* $j = 0,1,\dots,k-1$, *and any constant* $\kappa \geq 1$,

$$T_s(P,Q_j,R;\phi_{j+1};\Psi_j;\kappa) \ll PM_{j+1}R\Big(\prod_{i=1}^{j} H_iM_iR\Big).S_s(Q_{j+1},R) + (S_s(2Q_{j+1},R).S_s(2P,2Q_{j+1},R;\Psi_{j+1};2\kappa))^{1/2}.$$

Proof: Consider the equation (2.7) with $\theta = \phi_{j+1}$. We put $x = z+z'$ and $h = (z-z')w^{-k}$. Thus $2z = x + hw^k$ and $2z' = x - hw^k$. Hence by (2.6), (2.8) and (2.9), we have

$$T_s(P,Q_j,R;\theta;\Psi_j;\kappa) \leq U_0 + 2U_1 , \tag{3.1}$$

where U_0 is the number of solutions of (2.7) with (2.6), (2.8), (2.9) and $z = z'$, and U_1 is the number of solutions of the equation

$$\Psi_j(x+hw^k;2\underline{h};\underline{m}) + (2w)^k(x_1^k+\dots+x_s^k) = \Psi_j(x-hw^k;2\underline{h};\underline{m}) + (2w)^k(y_1^k+\dots+y_s^k) \tag{3.2}$$

with

$$\left.\begin{aligned} x \leq 2P,\ 1 \leq h \leq H_{j+1},\ M_{j+1} < w \leq M_{j+1}R, \\ x_i,y_i \in \mathcal{A}(Q_{j+1},R)\quad (1 \leq i \leq s), \\ 1 \leq h_r \leq \kappa H_r\ ,\ M_r < m_r \leq M_rR\quad (1 \leq r \leq j). \end{aligned}\right\} \tag{3.3}$$

Plainly,

$$U_0 \ll PM_{j+1}R.\Big(\prod_{i=1}^{j} H_iM_iR\Big)S_s(Q_{j+1},R) . \tag{3.4}$$

Also, U_1 is the number of solutions of the equation

$$\Psi_{j+1}(x;2\underline{h},h;\underline{m},w) + 2^k(x_1^k + \dots + x_s^k - y_1^k - \dots - y_s^k) = 0 , \tag{3.5}$$

with the variables satisfying (3.3).

Now write

$$F_j(\alpha;\lambda) = \sum_{M_1<m_1\le M_1R} \cdots \sum_{M_j<m_j\le M_jR} \sum_{h_1\le\lambda H_1} \cdots \sum_{h_j\le\lambda H_j} \sum_{z\le 2P} e(\alpha\Psi_j) . \tag{3.6}$$

Then by (3.5) and (3.3), and by considering the underlying diophantine equation, we have

$$U_1 \ll \int_0^1 F_{j+1}(\alpha;2\kappa).|f(\alpha;2Q_{j+1},R)|^{2s}\,d\alpha , \tag{3.7}$$

and so by Schwarz's inequality,

$$U_1 \ll \left\{ \int_0^1 |f(\alpha;2Q_{j+1},R)|^{2s} d\alpha \right\}^{1/2} \times \left\{ \int_0^1 |F_{j+1}(\alpha;2\kappa)^2.f(\alpha;2Q_{j+1},R)^{2s}|\,d\alpha \right\}^{1/2}.$$

Thus

$$U_1 \ll (S_s(2Q_{j+1},R).S_s(2P,2Q_{j+1},R;\Psi_{j+1};2\kappa))^{1/2}. \tag{3.8}$$

Collecting together (3.1), (3.4) and (3.8) completes the proof of the lemma.

Combining the conclusions of Lemmata 2.3 and 3.1 gives us a means of relating $S_s(P,Q_j,R;\Psi_j)$ to $S_s(P,Q_{j+1},R;\Psi_{j+1})$, and thus we are effectively able to "difference" Ψ_0 repeatedly to obtain estimates for $S_{s+1}(P,R)$. We aim to establish bounds of the form

$$S_s(P,R) \ll P^{\lambda_s+\varepsilon}$$

when R is no larger than a small power of P. The next lemma supplies us with such a bound.

Lemma 3.2. *Suppose that t is a positive integer, and μ a positive real number with $2t-k < \mu \le 2t$, and satisfying the property that given $\varepsilon > 0$, there is a positive number $\eta_0 = \eta_0(k,\varepsilon)$ such that whenever $0 < \eta < \eta_0$, we have $S_t(P,P^\eta) \ll P^{\mu+\varepsilon}$.*

Define the real numbers λ_s, θ_s, $\Delta(s)$ $(s = t, t+1, \dots)$ successively by $\lambda_t = \mu$, $\theta_t = 0$, $\Delta(t) = \mu - 2t + k$, and for $s > t$ by

$$\theta_s = \frac{1}{k+\Delta(s-1)} + \left[\frac{1}{k} - \frac{1}{k+\Delta(s-1)}\right]\left[\frac{k-\Delta(s-1)}{2k}\right]^{k-1}, \tag{3.9}$$

$$\Delta(s) = \Delta(s-1).(1 - \theta_s) + k\theta_s - 1 , \tag{3.10}$$

$$\lambda_s = 2s - k + \Delta(s) . \tag{3.11}$$

Then given t' *and* $\varepsilon' > 0$, *there is an* $\eta_1 = \eta_1(k,\varepsilon',t')$ *such that whenever* $0 < \eta < \eta_1$ *and* $t \le s \le t'$, *we have*

$$S_s(P,P^\eta) \ll P^{\lambda_s+\varepsilon'} .$$

Proof: We prove the result by induction, the case $s = t$ being assumed. So suppose that the result holds for $s' \le s$, and consider $S_s(P,P,R;\Psi_0)$ with $R = P^\eta$ and $0 \le \eta \le \eta_0/k$.

For $j = 1,\dots,k$, let

$$\phi_j = \frac{1}{k+\Delta(s)} + \left[\frac{1}{k} - \frac{1}{k+\Delta(s)}\right]\left[\frac{k-\Delta(s)}{2k}\right]^{k-j},$$

and adopt the same notation as that preceding Lemma 2.3. Also, write λ for λ_s . We shall now prove inductively that for $j = 0,\dots,k-1$, and any constant $\kappa \ge 1$, we have

$$T_s(P,Q_j,R;\phi_{j+1};\Psi_j;\kappa) \ll P^{1+(k-j)\varepsilon}M_{j+1}R^{2s(k-j)}\left(\prod_{i=1}^{j} H_iM_iR\right).Q_{j+1}^{\lambda+\varepsilon} . \tag{3.12}$$

By making a trivial estimate, we have

$$S_s(2P,2Q_k,R;\Psi_k;2\kappa) \ll P^2\left(\prod_{i=1}^{k} H_iM_iR\right)^2 S_s(2Q_k,R) .$$

Then by Lemma 3.1, we have

$$T_s(P,Q_{k-1},R;\phi_k;\Psi_{k-1};\kappa) \ll T_1 + T_2 ,$$

where

$$T_1 = PM_kR\left(\prod_{i=1}^{k-1} H_iM_iR\right).S_s(Q_k,R) ,$$

and

$$T_2 = \left[S_s(2Q_k,R).P^2\left(\prod_{i=1}^{k} H_iM_iR\right)^2 S_s(2Q_k,R)\right]^{1/2}.$$

But we have $\phi_k = 1/k$, and hence $H_k = 1$, and so the result now follows in the case $j = k-1$.

Now suppose that the result is true for $j' \geq j$. Then we may assume that $j \leq k-1$, and hence that

$$\phi_1 + \ldots + \phi_j < j/k \leq 1 - 1/k .$$

So by (3.12) and Lemma 2.3 we have

$$S_s(P, Q_j, R; \Psi_j; 2\kappa) \ll P^{1+(k-j+1)\varepsilon}(M_{j+1}R)^{2s}\left(\prod_{i=1}^{j} H_i M_i R\right)^2 R^{2s(k-j+1)} Q_{j+1}^{\lambda+\varepsilon} .$$

We then deduce from Lemma 3.1 that

$$T_s(P, Q_{j-1}, R; \phi_j; \Psi_{j-1}; \kappa) \ll T_3 + T_4 , \tag{3.13}$$

where

$$T_3 = PM_jR\left(\prod_{i=1}^{j-1} H_i M_i R\right).Q_j^{\lambda+\varepsilon}, \tag{3.14}$$

$$T_4 = P^{(k-j)\varepsilon}R^{2s(k-j+1)}\left(\prod_{i=1}^{j-1} H_i M_i R\right).T_5^{1/2}, \tag{3.15}$$

and

$$T_5 = PM_{j+1}^{2s}(H_jM_j)^2(Q_jQ_{j+1})^{\lambda+\varepsilon}. \tag{3.16}$$

We now note that

$$\begin{aligned} 1 + 2s\phi_{j+1} + \lambda(1-\phi_{j+1}-\Phi_j) + 2(1-(k-1)\phi_j) - 2(1+\phi_j) - \lambda(1-\Phi_j) &= 1 + (k-\Delta(s))\phi_{j+1} - 2k\phi_j \\ &= 0 \end{aligned}$$

and hence

$$T_5^{1/2} \ll (PM_j).Q_j^{\lambda+\varepsilon}.$$

Then from (3.13), (3.14) and (3.15), we have

$$T_s(P, Q_{j-1}, R; \phi_j; \Psi_{j-1}; \kappa) \ll P^{1+(k-j+1)\varepsilon}.M_jR^{2s(k-j+1)}\left(\prod_{i=1}^{j-1} H_i M_i R\right).Q_j^{\lambda+\varepsilon},$$

and so our second assertion follows for $j' = j-1$.

We have shown that (3.12) holds for $j = 0, \ldots, k-1$, and hence that

$$T_s(P, Q_0, R; \phi_1; \Psi_0) \ll P^{1+k\varepsilon}M_1R^{2ks}.Q_1^{\lambda+\varepsilon},$$

so that by Lemma 2.3,

$$S_s(P, P, R; \Psi_0) \ll P^{1+(k+1)\varepsilon}.M_1^{2s}R^{2(k+1)s}.Q_1^{\lambda+\varepsilon}.$$

But since $\phi_1 = \theta_{s+1}$, from (3.10) and (3.11) we have

$$\lambda(1-\phi_1) + 1 + 2s\phi_1 = \lambda_{s+1},$$

and hence on writing λ' for λ_{s+1}, we have

$$S_{s+1}(P,R) \ll S_s(P,P,R;\Psi_0) \ll P^{\lambda'+(k+2)\varepsilon}.R^{2(k+1)s}.$$

Then given $\varepsilon' > 0$, on taking $\varepsilon(\varepsilon')$ and $\eta(\varepsilon')$ to be sufficiently small positive real numbers, we have

$$S_{s+1}(P,P^\eta) \ll P^{\lambda_{s+1}+\varepsilon'}.$$

Thus the inductive hypothesis follows for s+1 in place of s, and this completes the proof of the lemma.

In order to make use of Lemma 3.2 we require a suitable estimate of the form $S_t(P,R) \ll P^{\mu+\varepsilon}$. This may be obtained either from classical estimates ($s = 1,2$), or from results of Vaughan [1989a, 1989c].

We note that the proof of Lemma 3.2 shows that the bound for $S_s(P,P^\eta)$ given also applies to $S_{s-1}(P,P,R;\Psi_0)$, and so the conclusion

$$S_s(P,P^\eta) \ll P^{\lambda_s+\varepsilon}$$

can be replaced, by considering the underlying diophantine equation, by the bound

$$\int_0^1 |\sum_{x\le P} e(\alpha x^k)|^2 \, |\sum_{x \in \mathcal{A}(P,P^\eta)} e(\alpha x^k)|^{2s-2} \, d\alpha \ll P^{\lambda_s+\varepsilon}.$$

In fact the new "fundamental lemma" (Lemma 2.2) in the iterative method enables us to draw such conclusions in general. In particular, by combining this observation with Theorem 4.4 of Vaughan [1989a], we have

$$\int_0^1 |\sum_{x\le P} e(\alpha x^3)|^2 \, |\sum_{x \in \mathcal{A}(P,P^\eta)} e(\alpha x^3)|^4 \, d\alpha \ll P^{13/4+\varepsilon}$$

whenever the positive number $\eta = \eta(\varepsilon)$ is sufficiently small.

4. THE ESTIMATION OF G(K) WHEN K IS LARGE.

We now investigate the consequences of the estimate given by Lemma 3.2 when k is large.

The proof of Theorem 1.2.

Adopt the same notation as in the statement of Theorem 1.2, and define

$$\lambda_s = 2s - k + \Delta(s) \tag{4.1}$$

for $s > 2$. Then on noting the classical estimate

$$S_2(P,P) \ll P^{2+\varepsilon},$$

we deduce from Lemma 3.2 that given s_0 and $\varepsilon > 0$, there is an $\eta_1 = \eta_1(k,\varepsilon,s_0)$ such that whenever $0 < \eta < \eta_1$ and $2 \le s \le s_0$, we have

$$S_s(P,P^\eta) \ll P^{\lambda_s+\varepsilon}. \tag{4.2}$$

Let m denote the set of real numbers α with the property that whenever $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and $|\alpha - a/q| \le q^{-1}P^{1/2-k}$ one has $q > P^{1/2}$. Let

$$X = P^{1/2}$$

and for a sufficiently small positive number η let

$$\mathcal{B} = \{ x : x = py,\ X/2 < p \le X,\ y \in \mathcal{A}(X,X^\eta) \}$$

and

$$h(\alpha) = \sum_{x \in \mathcal{B}} e(\alpha x^k).$$

Then by substituting the conclusion (4.2) (noting (4.1)) in place of Theorem 7.1 of Vaughan [1989a] in the argument of §7 of Vaughan [1989a], we have

$$\sup_{\alpha \in m} |h(\alpha)| \ll P^{1-\rho+\varepsilon}, \tag{4.3}$$

where for $k \geq 3$ we have

$$\rho = \rho(k) = \operatorname*{Max}_{s > 2} \rho(k,s) , \tag{4.4}$$

in which

$$\rho(k,s) = \frac{1}{4s} (1 - \Delta(s)) .$$

Since our auxiliary equations are identical to those used in the applications of the new iterative method in Vaughan [1989a], we may conclude that every large positive integer n may be written as the sum of $1 + 2u + 2v$ kth powers of positive integers, where as in §8 of Vaughan [1989a] we have

$$v \geq k+1 \text{ and } 2u\rho(k) > \Delta(v) .$$

We take

$$u = 1 + \left\lceil \frac{\Delta(v)}{2\rho(k)} \right\rceil .$$

Thus, by the argument of Vaughan [1989a] we have

$$G(k) \leq 3 + 2v + 2\left\lceil \frac{\Delta(v)}{2\rho(k)} \right\rceil .$$

This completes the proof of Theorem 1.2.

We note that the argument of §9 of Chapter 5 would enable us to reduce 3 to 2 in the above estimate, at the cost of a slightly more elaborate analysis.

The proof of Theorem 1.4 and Corollary 1.2.1.

All but the last two lines of Theorem 1.4 follow by the argument of §10 of Vaughan [1989a] on noting (1.3), (1.4), (4.3) and (4.4).

Suppose now that k is a large positive integer, and adopt the same notation as in the statement of Theorem 1.2. Denote by $\log_i k$ the i-fold iterated logarithm, and take $N = N(k)$ to be the positive integer with $2 \leq \log_{N+1} k < e^2$.

For $i = 1,\ldots,N$ define

$$t(i) = [k\log_{i+1}k] + 1,$$

and for $j = 1,\ldots,N$ define

$$s(j) = t(N) + \ldots + t(N-j+1).$$

Now, for $s > 2$, provided that $0 < \Delta(s-1) \le k-2$, we have by (1.1),

$$\frac{1}{k+\Delta(s-1)} < \theta_s \le \frac{1 + 2^{1-k}\Delta(s-1)/k}{k+\Delta(s-1)} . \tag{4.5}$$

and by (1.2) we have

$$\Delta(s-1)\left[1 - \frac{2 + 2^{1-k}\Delta(s-1)/k}{k+\Delta(s-1)}\right] < \Delta(s) < \Delta(s-1)\left[1 - \frac{2 - 2^{1-k}}{k+\Delta(s-1)}\right] . \tag{4.6}$$

On noting that $\Delta(2) = k-2$, we may proceed inductively to deduce that for $s > 2$ we have $0 < \Delta(s) < \Delta(s-1) \le k-2$. Then for $s > 2$, we have that $0 < \theta_s < 1/k$, and $\Delta(s)$ is a decreasing function of s. In particular, for $s > s(1)$,

$$\Delta(s-1) \le \Delta(t(N)) < (k-2)(1 - 1/k)^{t(N)-2}.$$

Now for $x \ge 1$,

$$(1 - 1/x)^x < e^{-1}, \tag{4.7}$$

and hence for $s > s(1)$ we have

$$\Delta(s-1) < (k-2)(1 - 1/k)^{-2}.\exp(-\log_{N+1}k) < k/\log_N k . \tag{4.8}$$

Then for $s > s(1)$,

$$\begin{aligned}\frac{2 - 2^{1-k}}{k+\Delta(s-1)} &> \frac{2}{k}(1 - 2^{-k})\left[1 - \frac{1}{\log_N k} + (\log_N k)^{-2}(1 + 1/\log_N k)^{-1}\right] \\ &> \frac{2}{k}(1 - 1/\log_N k)\end{aligned} \tag{4.9}$$

since

$$(1 - 2^{-k})(\log_N k)^{-2}(1 + 1/\log_N k)^{-1} > 2^{-k}(1 - 1/\log_N k) .$$

This last inequality follows on observing that owing to the assumption $\log_{N+1}k \ge 2$ we have

$$2^k > (\log_N k)^2.$$

We therefore deduce from (4.6), (4.8) and (4.9) that for $s > s(1)$, we have

$$\Delta(s) < \frac{k}{\log_N k}\left(1 - \frac{2}{k}(1 - 1/\log_N k)\right)^{s-s(1)}. \tag{4.10}$$

We now apply an inductive argument to show that when $1 < j \le N$, we have that for $s > s(j)$,

$$\Delta(s) < \frac{k}{\log_N k}\mathcal{L}_j^{-2}\left(1 - \frac{2}{k}(1 - (\log_{N+1-j} k)^{-2})\right)^{s-s(j)}, \tag{4.11}$$

in which we have written

$$\mathcal{L}_j = e^{-1}\log_{N-1} k\left(\prod_{i=2}^{j-1}\exp(t(N-i)/k - 2(\log_{N+1-i} k)^{-1})\right). \tag{4.12}$$

Here we adopt the convention that when the product in the last expression is empty, it is taken to have value one.

We first show that the inductive hypothesis holds for $j = 2$. From (4.10), for $s > s(2)$ we have

$$\Delta(s-1) \le \Delta(s(2)) < \frac{k}{\log_N k}\left(1 - \frac{2}{k}(1 - (\log_N k)^{-1})\right)^{t(N-1)}, \tag{4.13}$$

and by (4.7),

$$\left(1 - \frac{2}{k}(1 - (\log_N k)^{-1})\right)^{t(N-1)} < \exp(2 - 2\log_N k) = e^2(\log_{N-1} k)^{-2}. \tag{4.14}$$

Then in particular,

$$\Delta(s-1) < k(\log_{N-1} k)^{-2},$$

since $\log_N k \ge e^2$, owing to our assumption $\log_{N+1} k \ge 2$, and a similar argument to that leading to (4.9) yields

$$\frac{2 - 2^{1-k}}{k+\Delta(s-1)} > \frac{2}{k}\left(1 - (\log_{N-1} k)^{-2}\right).$$

Thus, by (4.6), (4.13) and (4.14) we deduce that (4.11) holds with $j = 2$.

Suppose now that the inductive hypothesis holds for $j = J > 1$. Then for $s > s(J+1)$ we have, in a similar manner to the argument for $j = 2$,

$$\Delta(s-1) \le \Delta(s(J+1))$$

$$< \frac{k}{\log_N k}\, \ell_J^{-2} \exp\left\{4(\log_{N+1-J} k)^{-1} - 2t(N-J)/k\right\} \qquad (4.15)$$

$$= \frac{k}{\log_N k}\, \ell_{J+1}^{-2} .$$

Then in particular,

$$\Delta(s-1) < k(\log_{N-J} k)^{-2},$$

and a similar argument to that leading to (4.9) yields

$$\frac{2 - 2^{1-k}}{k+\Delta(s-1)} > \frac{2}{k}\left\{1 - (\log_{N-J} k)^{-2}\right\} .$$

Thus, by (4.6) and (4.15) we deduce that (4.11) holds with $J+1$ replacing J. This completes the induction.

Now consider the substitution of (4.11) into the expression

$$\rho(s) = \frac{1}{4s}\,(1 - \Delta(s))$$

for $s > s(N)$. The maximum value of this expression as s varies is attained for a value of s satisfying

$$\left| s - \lambda\left\{ \log\left(\frac{k}{k - 2(1 - (\log k)^{-2})} \right)\right\}^{-1} \right| < 1$$

where λ is the larger root of the transcendental equation

$$(\lambda+1).\frac{k}{\log_N k}\, \ell_N^{-2}\left\{ 1 - \frac{2}{k}\,(1 - (\log k)^{-2})\right\}^{-s(N)} = e^{\lambda}.$$

Thus, on using the estimate

$$\left\{ 1 - \frac{2}{k}\,(1 - (\log k)^{-2})\right\}^{-s(N)}$$

$$= \exp\left\{ 2\sum_{i=0}^{N-1} t(N-i)/k\right\}.\left\{ 1 + O\left\{ \frac{\log\log k}{(\log k)^2} \right\}\right\} ,$$

we obtain

$$\lambda = \log k + \log\log k + \log_{N+1} k + 4\sum_{i=2}^{N-1} (\log_{N-i+1} k)^{-1} + O(1) . \qquad (4.16)$$

But by recalling that $N(k)$ has been chosen with $\log_{N+1} k \ge 2$, we have

$$\sum_{i=2}^{N-1} (\log_{N+1-i} k)^{-1} < \sum_{j=1}^{\infty} e^{-2j} < 1 ,$$

and hence the optimising choice for s satisfies

$$s = \lambda\left\{ \log(1 + 2/k) + O\left\{ \frac{1}{k\log k} \right\}\right\}^{-1} + O(1)$$

$$= \frac{k}{2} \left(\log k + \log\log k + O(1)\right) .$$

Hence, by (4.11),

$$2k\log k.\rho(k) = 1 + O(\log_2 k/\log k) . \tag{4.17}$$

The last two lines of Theorem 1.4 now follow by the argument of §10 of Vaughan [1989a] on noting (1.3), (1.4), (4.3), (4.4), and (4.17).

Now consider the substitution of (4.11) into the expression

$$2v + 2\left[\frac{\Delta(v)}{2\rho}\right] .$$

The optimising choice for v occurs with

$$|v - s(N) - \mu| < 1 ,$$

where

$$\mu \log\left\{ \frac{k}{k - 2(1 - (\log k)^{-2})} \right\} = \log(M(k)) ,$$

and

$$M(k) = \frac{k.\mathcal{L}_N^{-2}}{2\rho.\log_N k} \log\left\{ \frac{k}{k - 2(1 - (\log k)^{-2})} \right\} .$$

Thus, by (1.5), (4.11) and (4.17), we have

$$G(k) \le 2s(N) + 2\log(eM(k))\left\{\log\left\{ \frac{k}{k - 2(1 - (\log k)^{-2})} \right\}\right\}^{-1} + O(\log k)$$

$$= 2s(N) + k.\log\left\{ \frac{e.\mathcal{L}_N^{-2}}{\rho(k)\log_N k} \right\} + O(k/\log k) .$$

Then by (4.12) and (4.17),

$$G(k) \le k(\log k + \log\log k + \log_{N+1} k + 3 + \log 2 + O\left\{\frac{\log\log k}{\log k}\right\})$$

$$+ 4k \sum_{i=2}^{N-1} (\log_{N-i+1} k)^{-1}.$$

Then by the same argument as above, we deduce that

$$G(k) \le k(\log k + \log\log k + O(1)) \text{ as } k \longrightarrow \infty.$$

This completes the proof of Corollary 1.2.1.

5. BOUNDS FOR THE NUMBER OF SOLUTIONS OF THE AUXILIARY EQUATIONS.

When k is of moderate size, existing methods of Vaughan [1989a, 1989c] can be used to improve estimates arising from Lemma 3.2. Here we outline the methods used to calculate the values of λ_s used in §§6-8 which, when R is no larger than a small power of P, give bounds of the form

$$S_s(P,R) \ll P^{\lambda_s+\varepsilon}.$$

For $s = 3$ and 4, the estimates given by Theorem 1.4 of Vaughan [1989c] (which we shall call method "$b1$") are always superior to those arising from Lemma 3.2. When $s = 5$, the second of the estimates of Lemma 2.2 of Vaughan [1989c] (which we shall call method "j", where j is the parameter used in the method) is sometimes superior to Lemma 3.2. Finally, for smaller k the inequality (k-2) of Vaughan [1989a], §4 proves superior to Lemma 3.2 for large s.

In Table 5.1 we list the optimal values of λ_s for those s with $\lambda_s > 2s - k$ which arise from one of the above mentioned methods when $6 \le k \le 8$. We also list the corresponding values of θ_s, and in column j we list the method giving the optimal choice of λ_s. In the latter column we use "*" to denote that it is Lemma 3.2 which gives that choice. The table of values was computed to 16 significant figures by using an electronic computer, the final significant figure being rounded up.

In Table 5.2 we extend Table 5.1 for certain values of s when $9 \le k \le 20$ and $k = 32$.

Table 5.1.

k	s	j	θ	λ_s
6	3	$b1$	0.0454545455	3.090910
	4	$b1$	0.0833333334	4.333334
	5	*	0.1201242962	5.773790
	6	*	0.1288434456	7.318309
	7	*	0.1369149650	8.959303
	8	*	0.1439930252	10.685128
	9	*	0.1498768941	12.481705
	10	*	0.1545351314	14.334475
	11	$(k-2)$	0.1546391753	16.210587
	12	$(k-2)$	0.1546391753	18.105857
	13	$(k-2)$	0.1546391753	20.017323
7	3	$b1$	0.0319595422	3.063920
	4	$b1$	0.0681855611	4.264118
	5	*	0.0974431925	5.628154
	6	*	0.1038982326	7.082381
	7	*	0.1101648099	8.624130
	8	*	0.1160399777	10.247946
	9	*	0.1213462902	11.945936
	10	*	0.1259617078	13.708517
	11	*	0.1298348055	15.525370
	12	*	0.1329814175	17.386376
	13	*	0.1354674555	19.282307
	14	*	0.1373867321	21.205229
	15	$(k-2)$	0.1377777778	23.141397
	16	$(k-2)$	0.1377777778	25.086360
	17	$(k-2)$	0.1377777778	27.038906
8	3	$b1$	0.0248055265	3.049612
	4	$b1$	0.0607775378	4.228929
	5	*	0.0817750602	5.537309
	6	*	0.0866803646	6.924136
	7	*	0.0915512425	8.388838
	8	*	0.0962759134	9.929058
	9	*	0.1007419941	11.540657
	10	*	0.1048498730	13.217918
	11	*	0.1085249793	14.953943
	12	*	0.1117254184	16.741167
	13	*	0.1144431325	18.571891
	14	*	0.1166991500	20.438745
	15	*	0.1185352535	22.335020
	16	*	0.1200048163	24.254855
	17	*	0.1211648790	26.193294
	18	*	0.1220704489	28.146262
	19	*	0.1227710977	30.110475
	20	$(k-2)$	0.1228070176	32.079364
	21	$(k-2)$	0.1228070176	34.052073
	22	$(k-2)$	0.1228070176	36.028135
	23	$(k-2)$	0.1228070176	38.007136

Table 5.2.

k	s	λ_s	k	s	λ_s
9	20	31.185933	10	24	38.147771
	23	37.088830		26	42.095236
11	26	41.183092	12	30	48.154488
	30	49.082996		33	54.089993
13	32	51.187003	14	36	58.163644
	37	61.081900		41	68.076269
15	40	65.146397	16	42	68.174119
	44	73.082998		48	80.078679
17	46	75.158678	18	50	82.146823
	52	87.075299		56	94.072760
19	52	85.171222	20	56	92.160009
	60	101.070744		63	106.076886
32	100	168.143478			
	112	192.066307			

6. ESTIMATING G(K) FOR SMALLER K.

Following the analyses of §5 of Vaughan [1989a] and Chapter 5, it is now a simple and routine matter to cross from the estimates contained in Tables 5.1 and 5.2 to an upper bound for $G(k)$. We shall therefore be rather brief in the remainder of the proof of Theorem 1.1.

For $k = 6$ or 8, let $\eta = \eta(k)$ be a sufficiently small positive number (in the context of Lemma 3.2). Let

$$t = t(k) = \begin{cases} 13 & k = 6 \\ 23 & k = 8 \ . \end{cases}$$

We now consider the number, $R(n)$, of representations of a large natural number n in the form

$$x_1^k + \ldots + x_{2t}^k + y^k = n$$

with

$$x_i \in \mathcal{A}(P, P^\eta) \quad (1 \le i \le 2t) \text{ and } 1 \le y \le P ,$$

in which $P = n^{1/k}$.

By putting $u = v = 1$ in the argument of §5 of Vaughan [1989a], and noting that

$$u \le v \le 2u, \quad t+u \ge k+2,$$

and

$$\lambda_t + v(1 - 2^{1-k}) < 2t + v - k ,$$

we deduce that in each of the cases $k = 6$ and $k = 8$ we have

$$R(n) \gg \mathfrak{S}(n) P^{2t+v-k},$$

in which $\mathfrak{S}(n)$ is the usual singular series in Waring's problem, that is

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} (q^{-1}S(q,a))^{2t+v} e(-an/q) , \tag{6.1}$$

where

$$S(q,a) = \sum_{r=1}^{q} e(ar^k/q) . \tag{6.2}$$

Then by Theorem 4.6 of Vaughan [1981b] we have in each case

$$G(k) \le 2t+v .$$

For the case $k = 7$ the above analysis would give $G(7) \le 37$. However, we may obtain the result stated in Theorem 1.1 by applying Theorem 1.3 of Chapter 5. We set $t = 17$, $\mu = \lambda_{16}$, $\lambda = \lambda_{17}$, and

$$\theta = \frac{2^5-1}{7.2^5+1} = \frac{31}{225} .$$

Then the conditions of the aforementioned theorem are satisfied, since for $k = 7$ we have

$$t \ge 2k-2 ,$$

$$S_{t-1}(P, P^\eta) \ll_\varepsilon P^{\mu+\varepsilon}, \quad S_t(P, P^\eta) \ll_\varepsilon P^{\lambda+\varepsilon},$$

and

$$S_{t+1}(P, P^\eta) \ll_\varepsilon P^{2t+2-k+\varepsilon}$$

when $\eta > 0$ is sufficiently small, and

$$\mu > 2t-2-k \ , \ \lambda > 2t-k \ ,$$

$$\mu(1-\theta) + 1 + (2t-2)\theta = \lambda \ ,$$

$$\lambda(1-\theta) + 1 + 2t\theta < 2t+2-k \ .$$

Thus we conclude that $G(7) \le 2t+2 = 36$.

7. THE ESTIMATION OF G(K) FOR INTERMEDIATE VALUES OF K.

When $9 \le k \le 20$ we make use of Theorem 1.4 of Chapter 5. We note that we could instead use the analysis of §9 of Vaughan [1989a], at the cost of adding one to the upper bounds for $G(k)$ given in Theorem 1.1 for $9 \le k \le 20$. This defect may be remedied by making use of the comment at the end of §3 in the latter analysis.

Let

$$P = n^{1/k}, \ X = P^{k/(2k-1)}, \ Z = PX^{-1},$$

and define the generating function h by

$$\mathcal{C} = \{ \ x : x = pz, \ X/2 < p \le X, \ z \in \mathcal{A}(Z, Z^{\eta}) \ \} \ , \tag{7.1}$$

$$h(\alpha) = \sum_{x \in \mathcal{C}} e(\alpha x^k) \ . \tag{7.2}$$

Table 7.1.

k	$s(k)$	$\sigma(k)$	k	$s(k)$	$\sigma(k)$
8	16	0.01295004	15	40	0.00558208
9	20	0.01104785	16	42	0.00514140
10	24	0.00950665	17	46	0.00476322
11	26	0.00839659	18	50	0.00442971
12	30	0.00746422	19	52	0.00413669
13	32	0.00672248	20	56	0.00388274
14	36	0.00610732	32	100	0.00218668

Define $s = s(k)$ as in Table 7.1. Since s is even we may write $s = 2r$ for some integer r. Now let n denote the set of real numbers

α with the property that whenever $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and $|\alpha - a/q| \le q^{-1}X^{1-k}(rZ^k)^{-1}$ one has $q > X$. Then as in Vaughan [1989a] §9, we have for each $\eta > 0$ sufficiently small,

$$\sup_{\alpha \in \mathfrak{n}} |h(\alpha)| \ll P^{1-\sigma+\varepsilon},$$

where

$$\sigma = \sigma(k) = \frac{(k-1)(2s-\lambda_s) - k(k-2)}{2s(2k-1)}.$$

The values of σ given by the choices of s listed in Table 7.1 are also listed in Table 7.1, rounded down in the final decimal place ($\sigma(8)$ has also been calculated, this being superior to the estimate following from Weyl's inequality). The $s(k)$ were chosen so as to give the maximum value of σ.

Table 7.2.

k	$u(k)$	$t(k)$	k	$u(k)$	$t(k)$	k	$u(k)$	$t(k)$
9	23	9	14	41	13	18	56	17
10	26	11	15	44	15	19	60	18
11	30	10	16	48	16	20	63	20
12	33	13	17	52	16	32	112	31
13	37	13						

Now let $u = u(k)$ and $t = t(k)$ be as given by Table 7.2. Then for $9 \le k \le 20$, by Tables 5.2, 7.1 and 7.2 we have

$$\lambda_u + t(1-\sigma) < 2u+t-k,$$

and

$$u \ge 2k+1.$$

Thus the conditions of Theorem 1.4 of Chapter 5 are satisfied, and we may conclude that

$$G(k) \le 2u + t.$$

This completes the proof of Theorem 1.1.

8. AN UPPER BOUND FOR $G^+(K)$.

We now apply standard methods to give upper bounds for $G^+(16)$ and $G^+(32)$. By considering the subsistent 2-adic equations, we have $G^+(k) \geq 4k$ when k is a power of 2 (or see the argument of Hardy and Wright [1979] Theorem 396), and so the upper bound we deduce here will be sufficient to prove Theorem 1.3. We note also that the methods we apply for the cases $k = 16$ and $k = 32$ apply equally well to each integer $k \geq 8$.

Let n be a large natural number, and

$$P = n^{1/k}.$$

Let W be a parameter to be chosen later with $2 \leq W \leq P$, and define

$$\mathfrak{N}(q,a) = \{ \alpha : |\alpha - a/q| \leq (2kq)^{-1}WP^{-k} \} ,$$

and let $\mathfrak{N}$ denote the union of the $\mathfrak{N}(q,a)$ with $1 \leq a \leq q \leq W$, $(a,q) = 1$.

Let

$$\mathfrak{n} = ((2k)^{-1}P^{1-k}, 1 + (2k)^{-1}P^{1-k}]\setminus\mathfrak{N} .$$

Also, let $\mathfrak{C}$ and $h(\alpha)$ be defined as in (7.1) and (7.2) respectively, and let $\sigma(k)$ be as given in Table 7.1.

Table 8.1.

k	$u(k)$	$t(k)$
16	48	16
32	112	16

Take $u = u(k)$ and $t = t(k)$ as in Table 8.1, and let $R(n)$ denote the number of representations of the positive integer n in the form

$$x_1^k + \ldots + x_u^k + y_1^k + \ldots + y_t^k = n , \tag{8.1}$$

with

$$x_i \in \mathfrak{A}(P,P^\eta) \quad (1 \leq i \leq u) \text{ and } y_j \in \mathfrak{C} \quad (1 \leq j \leq t).$$

Then

$$R(n) = \int_0^1 g(\alpha)^u\, h(\alpha)^t\, e(-\alpha n)\, d\alpha$$

where

$$g(\alpha) = \sum_{x \in \mathcal{A}(P, P^\eta)} e(\alpha x^k) .$$

Also, we define

$$F(\alpha) = \begin{cases} g(\alpha)^u\, h(\alpha)^t & \alpha \in \mathfrak{n} \\ 0 & \alpha \notin \mathfrak{n} . \end{cases}$$

Then if we write

$$R_{\mathfrak{n}}(n) = \int_0^1 F(\alpha) e(-\alpha n)\, d\alpha ,$$

we have that $R_{\mathfrak{n}}(n)$ is the nth Fourier coefficient of the function $F(\alpha)$.

Let

$$R_{\mathfrak{N}}(n) = \int_{\mathfrak{N}} g(\alpha)^u\, h(\alpha)^t\, e(-\alpha n)\, d\alpha .$$

Then by Bessel's inequality, for each positive integer N we have

$$\begin{aligned} \sum_{n \le N} |R(n) - R_{\mathfrak{N}}(n)|^2 &= \sum_{n \le N} |R_{\mathfrak{n}}(n)|^2 \\ &\le \int_0^1 |F(\alpha)|^2\, d\alpha \\ &= \int_{\mathfrak{n}} |g(\alpha)^{2u}\, h(\alpha)^{2t}|\, d\alpha \end{aligned}$$

But as in §7 we have

$$\operatorname*{Sup}_{\alpha \in n} |h(\alpha)| \ll P^{1-\sigma+\varepsilon},$$

so that by Tables 8.1 and 7.1 we have

$$\lambda_u + 2t(1-\sigma) < 2u+2t-k$$

and

$$u > 2k+1 .$$

Thus we may apply the arguments of §9 of Chapter 5 to deduce that for some $\delta > 0$, when W is no larger than a small power of $\log P$, we have

$$\int_{\mathfrak{n}} |g(\alpha)^{2u}\, h(\alpha)^{2t}|\, d\alpha \ll P^{2u+2t-k}\, W^{-\delta} (\log P)^{-2t}.$$

But by the argument of §5 of Vaughan [1989a], we have

$$\int_{\mathfrak{N}} g(\alpha)^{u}\, h(\alpha)^{t}\, e(-\alpha n)\, d\alpha \gg \mathfrak{S}(n).P^{u+t-k}(\log P)^{-t}$$

when W is a sufficiently small power of $\log P$. Here $\mathfrak{S}(n)$ denotes the usual singular series in Waring's problem, namely that given by (6.1) with $t+u$ replacing $2t+v$. Now when $t+u \geq 4k$ we have that

$$1 \ll \mathfrak{S}(n) \ll 1,$$

and hence by partial summation, with V a sufficiently small power of $\log N$, and N sufficiently large, we have

$$\sum_{N/2<n\leq N} \left| \frac{R(n) - R_{\mathfrak{N}}(n)}{R_{\mathfrak{N}}(n)} \right|^2 \ll \frac{N^{(2u+2t-k)/k}\, V^{-\delta} (\log N)^{-2t}}{(N^{(u+t-k)/k}\, (\log N)^{-t})^2}.$$

Then there is a $\nu > 0$ such that the number of integers, $E_s(N)$, in $[1,N]$ which are not the sum of s kth powers of positive integers, satisfies

$$E_{u+t}(N) \ll N.(\log N)^{-\nu}.$$

Thus almost all numbers are represented in the form (8.1), and hence we deduce that

$$G^*(k) \leq u+t$$

for $k = 16$ and $k = 32$.

This completes the proof of Theorem 1.3.

APPENDIX A.

ON A PROBLEM RELATED TO ONE OF LITTLEWOOD AND OFFORD.

1. INTRODUCTION.

Let $z_1, z_2, \ldots, z_n$ be complex numbers of modulus at least one. Denote by $N(\underline{z}) = N(z_1, \ldots, z_n)$ the number of sums of the form

$$\sum_{i=1}^{n} \varepsilon_i z_i ,$$

with $\varepsilon_i = 1$ or -1, lying in the interior of a given disc of unit radius.

From their investigations on "random polynomials", Littlewood and Offord were led to consider bounds for $N(\underline{z})$, and gave one adequate for their purposes (see Littlewood and Offord [1943], Theorem 1). Erdős [1945] applied a Sperner's Theorem argument (see Bollobás [1986], §§3,4) to deduce that when the z_i are all real, we have

$$N(\underline{z}) \leq \binom{n}{[n/2]}$$

with equality holding when $z_1 = \ldots = z_n = 1$. Later, Kleitman [1965, 1970] and Katona [1966] extended this result to the complex field, and even to an analogous result on vectors in an arbitrary archimedean normed space. More recently, Griggs [1980] has elaborated on these results and arguments.

It would appear that all results thus far have been on archimedean spaces of some sort, and indeed this would appear to be essential for the Sperner's Theorem argument to succeed. We now give a result for a non-archimedean example:

Theorem 1.1. *Let $\alpha_1, \ldots, \alpha_n$ be reduced residues (mod q), and let $N(q;\underline{\alpha})$ denote the number of choices of $\varepsilon_1, \ldots, \varepsilon_n$, with $\varepsilon_i = 0$ or 1, such that*

$$\sum_{i=1}^{n} \varepsilon_i \alpha_i \equiv 0 \ (mod\ q) .$$

If $(q,\alpha_i) = 1$ *for every* i, *and* $q > (n+1)/2$, *then*

$$N(q;\underline{\alpha}) \le \binom{n}{[n/2]} .$$

Further, putting

$$\gamma_i^{(n)} = \begin{cases} 1 & \text{for } 1 \le i \le [n/2] \\ -1 & \text{for } [n/2] < i \le n . \end{cases}$$

we have $N(q;\underline{\gamma}^{(n)}) = \binom{n}{[n/2]}$.

Corollary 1.1.1. *Let* $\alpha_1,\dots,\alpha_n \in \mathbb{Q}_p$ *satisfy* $|\alpha_i|_p = 1$. *Denote by*

$$N^{(r)}(p;\underline{\alpha}) = N^{(r)}(p;\alpha_1,\dots,\alpha_n)$$

the number of choices of $\varepsilon_1,\dots,\varepsilon_n$ *with* $\varepsilon_i = 0$ *or* 1, *such that*

$$\left|\sum_{i=1}^{n}\varepsilon_i\alpha_i\right|_p \le p^{-r}.$$

If $p^r > (n+1)/2$, *then*

$$N^{(r)}(p;\underline{\alpha}) \le \binom{n}{[n/2]} .$$

The corollary is immediate from the theorem on considering congruences (mod p^r). Results less precise than the theorem have recently been used in investigations on the local solubility of simultaneous additive equations (see, for example, Lemma 3.4 of Chapter 2).

Our proof is divided into many cases. When n is even, the use of exponential sums makes the result almost immediate. However, life is rather harder when n is odd, and we must take care to exploit all available asymmetries present in the residue system, these tending to deflate $N(q;\underline{\alpha})$.

2. PROOF OF THE THEOREM.

Let

$$S(\beta) = 1 + e(\beta/q) \ .$$

Then, by considering the underlying exponential sums, we have

$$N(q;\underline{\underline{\alpha}}) = q^{-1} \sum_{r=1}^{q} \prod_{i=1}^{n} S(r\alpha_i) \ . \tag{2.1}$$

We divide into cases.

(A) Suppose that n is even.

Applying Hölder's inequality to (2.1), we have

$$N(q;\underline{\underline{\alpha}}) \le \prod_{i=1}^{n} \left(q^{-1} \sum_{r=1}^{q} |S(r\alpha_i)|^n \right)^{1/n}$$
$$= \prod_{i=1}^{n} \left(q^{-1} \sum_{r=1}^{q} |S(r)|^n \right)^{1/n}$$

by a change of variable. Thus, since n is even and $q > (n+1)/2$, we have

$$N(q;\underline{\underline{\alpha}}) \le N(q;\underline{\underline{\gamma}}^{(n)}) = \sum_{j=0}^{[n/2]} \binom{[n/2]}{j}^2 = \binom{n}{[n/2]} \ ,$$

and the result holds in case (A).

(B) Suppose that n is odd.

We write $n = 2k+1$ with k a positive integer (the case $n = 1$ is trivial). We divide into cases according to the value of $\eta = \alpha_1 + \ldots + \alpha_n$.

(i) Suppose that $(\eta, q) = 1$.

For λ a given reduced residue (mod q), let $N_\lambda(q;\underline{\underline{\alpha}})$ denote the number of choices of $\varepsilon_1, \ldots, \varepsilon_n$, with $\varepsilon_i = 0$ or 1, such that

$$\sum_{i=1}^{n} \varepsilon_i \alpha_i \equiv \lambda \pmod{q} \ . \tag{2.2}$$

Then $N_\eta(q;\underline{\underline{\alpha}}) = N(q;\underline{\underline{\alpha}})$, since whenever (2.2) holds with $\lambda = \eta$, we have

$$\sum_{i=1}^{n} (1-\varepsilon_i)\alpha_i \equiv 0 \pmod{q} \ .$$

Then we have

$$N(q;\alpha_1, \ldots, \alpha_n, -\eta) = N(q;\underline{\underline{\alpha}}) + N_\eta(q;\underline{\underline{\alpha}}) = 2.N(q;\underline{\underline{\alpha}}) \ . \tag{2.3}$$

But by part (A), we have

$$N(q;\alpha_1,\dots,\alpha_n,-\eta) \le \binom{2k+2}{k+1} = 2\binom{2k+1}{k} ,$$

and the result follows in case (B)(i).

(ii) Suppose that (η,q) is a proper divisor of q.

Let $d = (\eta,q)$. By applying Hölder's inequality to (2.1), we have

$$N(q;\alpha_1,\dots,\alpha_n,-\eta) \le \prod_{i=1}^{k}\Big(q^{-1}\sum_{r=1}^{q}|S(r\alpha_i)|^{2k}.|S(r\eta)|^2\Big)^{1/(2k)} \times \prod_{i=k+1}^{n}\Big(q^{-1}\sum_{r=1}^{q}|S(r\alpha_i)|^{2k+2}\Big)^{1/(2k+2)} .$$

We now observe that

$$q^{-1}\sum_{r=1}^{q}|S(r\alpha_i)|^{2k+2} = N(q;\underline{\underline{\gamma}}^{(n+1)}) = \binom{2k+2}{k+1} = 2\binom{2k+1}{k} ,$$

and, by a change of variable,

$$q^{-1}\sum_{r=1}^{q}|S(r\alpha_i)|^{2k}.|S(r\eta)|^2 = N(q;\underline{\underline{\gamma}}^{(n-1)},\xi,-\xi) ,$$

for some ξ with $(\xi,q) = d$.

We establish now a lemma which is useful both here and later.

Lemma 2.1. *Suppose that $k \ge 1$, $q \ge k+2$ and $\alpha \not\equiv \pm 1,0 \pmod q$. Then*

$$N(q;\underline{\underline{\gamma}}^{(2k)},\alpha,-\alpha) < \frac{2k+3}{4k+8}\binom{2k+3}{k+1} \quad \text{for } k \ne 2,$$

and

$$N(q;\underline{\underline{\gamma}}^{(4)},\alpha,-\alpha) \le 16 .$$

Proof: We have

$$N(q;\underline{\underline{\gamma}}^{(2k)},\alpha,-\alpha) = N_0(q;\underline{\underline{\gamma}}^{(2k)}) + N_{\alpha-\alpha}(q;\underline{\underline{\gamma}}^{(2k)}) + N_{\alpha}(q;\underline{\underline{\gamma}}^{(2k)}) + N_{-\alpha}(q;\underline{\underline{\gamma}}^{(2k)}) .$$

Since $\alpha \not\equiv \pm 1,0 \pmod q$ we have

$$N(q;\underline{\underline{\gamma}}^{(2)},\alpha,-\alpha) = 2N_0(q;\underline{\underline{\gamma}}^{(2)}) = 4 < 25/6 = \frac{5}{12}\binom{5}{2} ,$$

and

$$N(q;\underline{\underline{\gamma}}^{(4)},\alpha,-\alpha) \le 2N_0(q;\underline{\underline{\gamma}}^{(4)}) + 4 = 16 .$$

Thus we may suppose that $k \geq 3$. Choose u and t so that $0 \leq u < q$, $0 \leq t < q$, $\alpha + u \equiv 0 \pmod q$, $-\alpha + t \equiv 0 \pmod q$. Then, as $\alpha \not\equiv \pm 1, 0 \pmod q$ we have $u \geq 2$, $t \geq 2$.

When $u \leq k$ the congruence

$$\sum_{i=1}^{2k} \varepsilon_i \gamma_i^{(2k)} \equiv \alpha \pmod q$$

has

$$\sum_{r=0}^{k-u} \binom{k}{r}\binom{k}{r+u} = \binom{2k}{k-u}$$

solutions of the type $r.1 + (r+u)(-1) \equiv \alpha \pmod q$, when $t \leq k$ it has $\binom{2k}{k-t}$ of the type $(r+t).1 + r(-1) \equiv \alpha \pmod q$, and it has no other solutions since $q \geq k+2$. There is a concomitant conclusion when α is replaced by $-\alpha$. Thus, on using part (A) to estimate $N_0(q; \underline{\gamma}^{(2k)})$, we deduce that

$$N(q; \underline{\gamma}^{(2k)}, \alpha, -\alpha) \leq 2\binom{2k}{k} + 2\binom{2k}{k-u} + 2\binom{2k}{k-t} . \tag{2.4}$$

Since $u \equiv -\alpha \equiv -t \pmod q$ we have $u+t = q$, and without loss of generality we may suppose now that $t \geq u$, whence $t \geq \frac{1}{2}q \geq \frac{1}{2}k + 1$. Thus we may now assume that

$$N(q; \underline{\gamma}^{(2k)}, \alpha, -\alpha) \leq 2\binom{2k}{k} + 2\binom{2k}{k-2} + 2\binom{2k}{k-t} .$$

holds with $k \geq 3$ and $t \geq \frac{1}{2}k + 1$. Hence

$$N(q; \underline{\gamma}^{(2k)}, \alpha, -\alpha) \leq \frac{2k+3}{4k+8}\binom{2k+3}{k+1}\lambda$$

where λ is given by

$$(\lambda-1)\frac{(2k+1)(2k+3)^2}{2k+4} = 1 - 2k + \frac{3}{2k+4} + 2k(k-1)\prod_{m=3}^{t}\frac{k+1-m}{k+m} .$$

When $k = 3, 5, 6$ the right hand side does not exceed $-27/10$, $-9 + 3/14 + 10/3$, $-11 + 3/16 + 8$ respectively, so $\lambda < 1$. When $k = 4$ and $t \geq 4$ it does not exceed $-7 + 1/4 + 6/7$, so again $\lambda < 1$. When $k = 4$ and $t = 3$, we have $3 \geq q/2 \geq 3$, so $q = 6$, $u = 3$. Thus

$$N(q; \underline{\gamma}^{(8)}, \alpha, -\alpha) \leq 2\binom{8}{4} + 4\binom{8}{1} = 172 < \frac{11}{24}\binom{11}{5} .$$

Therefore we may suppose that $k \geq 7$. Now the product on the right hand side is

$$\exp\left[\sum_{m=3}^{t} \log\left(\frac{1-\frac{2m-1}{2k+1}}{1+\frac{2m-1}{2k+1}}\right)\right] < \exp\left[-\sum_{m=3}^{t}\frac{4m-2}{2k+1}\right]$$

and when $k \geq 7$ we have

$$\sum_{m=3}^{t}\frac{4m-2}{2k+1} = \frac{2t^2-8}{2k+1} \geq \frac{(k+2)^2-16}{4k+2} = \frac{k}{4}+\frac{7}{8}-\frac{55}{16k+8} > \frac{k+1}{4}$$

and $\exp((k+1)/4) > k$. The last inequality may be established by observing that $f(x) = \exp((x+1)/4) - x$ is strictly increasing for $x \geq 8\log 2 - 1$, and that $7 > 8\log 2 - 1$ and $e^2 > 7$, and so $f(k) \geq f(7) > 0$.

It now follows that

$$(\lambda-1)\frac{(2k+1)(2k+3)^2}{2k+4} < 2 - 2k + 2(k-1) = 0,$$

whence $\lambda < 1$ once more.

This completes the proof of the lemma.

Returning to the proof of the theorem, by Lemma 2.1,

$$2.N(q;\underline{\alpha}) = N(q;\alpha_1,\ldots,\alpha_n,-\eta) < \left[\frac{1}{2}\left[\frac{2k+3}{k+2}\right]^2\left[\begin{matrix}2k+1\\k\end{matrix}\right]\right]^{1/2}\left[2\left[\begin{matrix}2k+1\\k\end{matrix}\right]\right]^{1/2}$$
$$< 2\left[\begin{matrix}2k+1\\k\end{matrix}\right].$$

Thus the result follows in the case (B)(ii).

(iii) Suppose that $q|\eta$.

Then we have

$$\alpha_n \equiv -(\alpha_1 + \ldots + \alpha_{n-1}) \pmod{q},$$

and hence

$$N(q;\alpha_1,\ldots,\alpha_n) = N(q;\alpha_1,\ldots,\alpha_{n-1},-(\alpha_1+\ldots+\alpha_{n-1}))$$
$$= 2.N(q;\alpha_1,\ldots,\alpha_{n-1}). \qquad (2.5)$$

We now divide into further cases.

(a) Suppose that there is a β with $\alpha_i \equiv \pm\beta$ for $i = 1,\dots,n$.

Suppose that there are m values of i with $\alpha_i \equiv \beta$, and $n-m$ values of i with $\alpha_i \equiv -\beta$. Thus

$$m\beta - (n-m)\beta \equiv 0 \pmod{q} .$$

Without loss of generality we may assume that $n-m \geq m$. Then since $(\beta,q) = 1$ and $q > (n+1)/2$, we have either $n = 2m$ or $n = 2m + q$. The first case cannot occur, since we have supposed n to be odd. Then we must have $1 \leq m = (n-q)/2 < k/2$.

We therefore have

$$\begin{aligned} N(q;\underline{\underline{\alpha}}) &= \sum_{r=0}^{m} \begin{bmatrix} n-m \\ r \end{bmatrix} \begin{bmatrix} m \\ r \end{bmatrix} + \sum_{r=0}^{m} \begin{bmatrix} n-m \\ q+r \end{bmatrix} \begin{bmatrix} m \\ r \end{bmatrix} \\ &= \begin{bmatrix} n \\ m \end{bmatrix} + \begin{bmatrix} n \\ q+m \end{bmatrix} \\ &= \begin{bmatrix} 2k+1 \\ k \end{bmatrix} 2 \prod_{i=1}^{k-m} \frac{m+i}{k+1+i} . \end{aligned}$$

Then on noting that $2(m+1) < k+2$, we deduce that

$$N(q;\underline{\underline{\alpha}}) < \begin{bmatrix} 2k+1 \\ k \end{bmatrix} ,$$

and the result follows once again.

(b) Suppose that there is no β such that for $i = 1,\dots,n$ we have $\alpha_i \equiv \pm\beta$.

By a rearrangement of variables we may suppose that there is no β such that for $i = 1,\dots,n-1$ we have $\alpha_i \equiv \pm\beta$. The case $k = 1$ is trivial. Suppose then that $k > 1$, and let $S_\zeta \subseteq \{1,\dots,n-1\}$ denote the set of indices for which $\alpha_i \equiv \pm\zeta \pmod{q}$. There are three cases:

(α) We may choose ζ with $1 < \mathrm{card}(S_\zeta) \leq 2k-2$.

Choose ζ so that $s = \mathrm{card}(S_\zeta)$ is minimal amongst those ζ satisfying $1 < \mathrm{card}(S_\zeta) \leq 2k-2$. Then we may rearrange variables so that $S_\zeta = \{\, 2k-s+1,\dots,2k \,\}$, and by Hölder's inequality, for $k > 2$ we have

$$N(q;\alpha_1,\dots,\alpha_{n-1}) \le \prod_{i=1}^{2k-s}\Bigg(\Big[q^{-1}\sum_{r=1}^{q}|S(r\alpha_i)|^{2k-2}.|S(r\zeta)|^2\Big]^{\frac{2k-2-s}{(2k-4)(2k-s)}} \times \Big[q^{-1}\sum_{r=1}^{q}|S(r\alpha_i)|^{2}.|S(r\zeta)|^{2k-2}\Big]^{\frac{s-2}{(2k-4)(2k-s)}}\Bigg)$$

For $k = 2$ (and $s = 2$), by Cauchy's inequality we have

$$N(q;\alpha_1,\dots,\alpha_{n-1}) \le \prod_{i=1}^{2}\Big[q^{-1}\sum_{r=1}^{q}|S(r\alpha_i)|^{2}.|S(r\zeta)|^2\Big]^{1/2}.$$

(β) For every ζ, $\mathrm{card}(S_\zeta) \le 1$.

Then we have

$$N(q;\alpha_1,\dots,\alpha_{n-1}) \le \prod_{i=1}^{k}\Bigg(\Big[q^{-1}\sum_{r=1}^{q}|S(r\alpha_{2i-1})|^{2k-2}.|S(r\alpha_{2i})|^2\Big]^{1/(2k)} \times \Big[q^{-1}\sum_{r=1}^{q}|S(r\alpha_{2i-1})|^{2}.|S(r\alpha_{2i})|^{2k-2}\Big]^{1/(2k)}\Bigg)$$

(γ) There are reduced residues ζ and ξ with $\mathrm{card}(S_\zeta) = 1$ and $\mathrm{card}(S_\xi) = 2k-1$.

We may rearrange variables so that $S_\zeta = \{2k\}$. Then by Cauchy's inequality, we have

$$N(q;\alpha_1,\dots,\alpha_{n-1}) \le \Big[q^{-1}\sum_{r=1}^{q}|S(r\xi)|^{2k-2}|S(r\zeta)|^2\Big]^{1/2} \times \Big[q^{-1}\sum_{r=1}^{q}|S(r\xi)|^{2k}\Big]^{1/2}.$$

We now observe that given α and β with $(\alpha\beta,q) = 1$ and $\alpha \not\equiv \pm\beta \pmod q$, there is a β' with $\beta' \not\equiv \pm1,0 \pmod q$ such that

$$q^{-1}\sum_{r=1}^{q}|S(r\alpha)|^{2k-2}.|S(r\beta)|^2 = N(q;\underline{\underline{\gamma}}^{(2k-2)},\beta',-\beta') .$$

The premiss of Lemma 2.1 is satisfied with k replaced by $k-1$, since $k \ge 2$. Thus the above does not exceed $\frac{2k+1}{4k+4}\binom{2k+1}{k}$ when $k \ne 3$, and 16 when $k = 3$. Hence in cases (α) and (β) we have

$$2.N(q;\alpha_1,\dots,\alpha_{n-1}) \le \binom{2k+1}{k}$$

and the result follows from (2.5).

In the case (γ), when $k \neq 3$ we obtain, via part (A),

$$2.N(q;\alpha_1,\ldots,\alpha_{n-1}) \leq 2\left[\frac{2k+1}{4k+4}\binom{2k+1}{k}\binom{2k}{k}\right]^{1/2} = \binom{2k+1}{k}$$

and again the result follows from (2.5)

When $k = 3$ we obtain, in the same way,

$$N(q;\alpha_1,\ldots,\alpha_{n-1}) \leq \left[16\binom{6}{3}\right]^{1/2} = (320)^{1/2} < 18.$$

Thus, by (2.5),

$$N(q;\underline{\underline{\alpha}}) \leq 34 < 35 = \binom{7}{3} .$$

This completes the proof of the theorem.

REFERENCES.

APOSTOL, T.M. (1957). Mathematical analysis. *Addison-Wesley.*

ARKHIPOV, G.I. & KARATSUBA, A.A. (1981). Local representation of zero by a form. *Izv. Akad. Nauk SSSR Ser. Mat.* 45, 5, 948-961 *and* 1198.

ARTIN, E. (1965). Collected Papers. *Addison-Wesley, Reading, Mass.*.

ATKINSON, O.D. & COOK, R.J. (1989). Pairs of additive congruences to a large prime modulus. *J. Austral. Math. Soc.*, 46*A*, 438-455.

BAKER, A. (1975). Transcendental number theory. *Cambridge University Press, Cambridge.*

BAKER, R.C. (1986). Diophantine inequalities. *L.M.S. Monographs New Series* 1, *O.U.P.*.

BAKER, R.C. (1990). Diagonal cubic equations II. *Acta Arith.* (*to appear*).

BAKER, R.C. & BRÜDERN, J. (1988). On pairs of additive cubic equations. *J. reine angew. Math.* 391, 157-180.

BALASUBRAMANIAN, Ram. (1985). On Waring's problem: $g(4) \leq 20$. *Hardy-Ramanujan J.* 8, 1-40.

BALASUBRAMANIAN, Ram., DESHOUILLERS, J.-M. & DRESS, F. (1986a). Problème de Waring pour les bicarrés. I: Schéma de la solution. *C.R. Acad. Sci., Paris*, Sér. I. 303, 85-88.

BALASUBRAMANIAN, Ram., DESHOUILLERS, J.-M. & DRESS, F. (1986b). Problème de Waring pour les bicarrés. II: résultats auxiliaires pour le théorème asymptotique. *C.R. Acad. Sci., Paris*, Sér. I. 303, 161-163.

BOLLOBÁS, B. (1986). Combinatorics. *Cambridge University Press.*

BRÜDERN, J. (1990). On Waring's problem for fifth powers and some related topics. *Proc. Lond. Math. Soc.* (3) (*to appear*).

CASSELS, J.W.S. (1986). Local fields. *London Mathematical Society Student Texts* 3, *Cambridge University Press.*

CHOWLA, S., MANN, H.B. & STRAUS, E.G. (1959). Some applications of the Cauchy-Davenport theorem. *Kon. Norske Vidensk. Selsk. Forh.*, 32, 74-80.

COOK, R.J. (1971). Simultaneous quadratic equations. *J. Lond. Math. Soc.*, (2), 4, 319-326.

COOK, R.J. (1972). Pairs of additive equations. *Michigan Math. J.*, 19, 325-331.

COOK, R.J. (1985). Pairs of additive congruences: cubic congruences. *Mathematika*, 32, 286-300.

DAVENPORT, H. (1939a). On Waring's problem for cubes. *Acta Math.* 71, 123-143.

DAVENPORT, H. (1939b). On Waring's problem for fourth powers. *Ann. Math.*, 40, 731-747.

DAVENPORT, H. & LEWIS, D.J. (1963). Homogeneous additive equations. *Proc. Roy. Soc. Lond.*, 274*A*, 443-460.

DAVENPORT, H. & LEWIS, D.J. (1966). Cubic equations of additive type. *Philos. Trans. Roy. Soc. Lond.*, 261*A*, 97-136.

DAVENPORT, H. & LEWIS, D.J. (1967). Two additive equations. *Proc. Symp. in Pure Math., Vol. XII, Houston, Tex.*, 74-98.

DAVENPORT, H. & LEWIS, D.J. (1969). Simultaneous equations of additive type. *Philos. Trans. Roy. Soc. Lond.*, 264*A*, 557-595.

DEMJANOV, V.B. (1956). Pairs of quadratic forms over a complete field with discrete norm with a finite field of residue classes. *Izv. Akad. Nauk. SSSR. Ser. Mat.*, 20, 307-324.

DODSON, M.M. (1966). Homogeneous additive congruences. *Philos. Trans. Roy. Soc. Lond.*, 261*A*, 163-210.

ERDÖS, P. (1945). On a lemma of Littlewood and Offord. *Bull. Amer. Math. Soc.*, 51, 898-902.

FOUVRY, E. & TENENBAUM, G. (1990). Entiers sans grand facteur premier en progressions arithmétiques. (*to appear*).

GREENBERG, M.J. (1969). Lectures on forms in many variables. *W.A.Benjamin, Inc., New-York - Amsterdam.*

GRIGGS, J.R. (1980). The Littlewood-Offord problem: tightest packing and an M-part Sperner theorem. *Europ. Jl. Combin.* 1, 225-234.

HARDY, G.H. & LITTLEWOOD, J.E. (1922). Some problems of 'Partitio Numerorum': IV. The singular series in Waring's Problem and the value of the number G(k). *Math. Z.*, 12, 161-88.

HARDY, G.H. & LITTLEWOOD, J.E. (1925). Some problems of 'Partitio Numerorum'(VI): Further researches in Waring's problem. *Math. Z.*, 23, 1-37.

HARDY, G.H. & WRIGHT, E.M. (1979). An introduction to the Theory of Numbers. *Oxford, University Press, fifth edition.*

HEATH-BROWN, D.R. (1988). Weyl's inequality, Hua's inequality, and Waring's problem. *J. Lond. Math. Soc.* (2), 38, 216-230.

HILBERT, D. (1909). Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl n-ter Potenzen (Waringsches Problem). *Math. Annalen* 67, 281-300.

HUA, L.-K. (1965). Additive theory of prime numbers. *Providence, Rhode Island: American Mathematical Society.*

KATONA, G.O.H. (1966). On a conjecture of Erdős and a stronger form of Sperner's theorem. *Studia Sci. Math. Hungar.* 1, 59-63.

KLEITMAN, D.J. (1965). On a lemma of Littlewood and Offord on the distribution of certain sums. *Math. Z.* 90, 251-259.

KLEITMAN, D.J. (1970). On a lemma of Littlewood and Offord on the distributions of linear combinations of vectors. *Adv. Math.* 5, 155-157.

LEWIS, D.J. (1957). Cubic congruences. *Michigan Math. J.*, 4, 85-95.

LINNIK, Ju. V. (1943). On the representation of large numbers as sums of seven cubes. *Rec. Math. (Mat. Sbornik) N.S.*, 12 (54), 218-224.

LITTLEWOOD, J.E. & OFFORD, A.C. (1943). On the number of real roots of a random algebraic equation, III. *Mat. Sbornik* 12, 277-286.

LOW, L., PITMAN, J. & WOLFF, A. (1988). Simultaneous diagonal congruences. *J. Number Theory* 29, 1, 31-59.

MATIJASEVIČ, Ju.V. (1970). The diophantineness of enumerable sets. *Dokl. Akad. Nauk SSSR* 191, 279-282.

MATIJASEVIČ, Ju.V. (1971). Diophantine representation of enumerable predicates. *Izv. Akad. Nauk SSSR Ser. Mat.* 35, 3-30.

NARASIMHAMURTI, V. (1941). On Waring's problem for 8^{th}, 9^{th}, and 10^{th} powers. *J. Indian Math. Soc. (N.S.)* 5, 122.

SCHMIDT, W.M. (1976). Equations over finite fields. An elementary approach. *Lecture Notes in Mathematics*, 536, *Springer-Verlag, Berlin.*

TERJANIAN, G. (1966). Un contre-exemple à une conjecture d'Artin. *C.R. Acad. Sci. Paris Sér. A-B* 262, A612.

TURINA, O.V. (1987). A new estimate for a trigonometric integral of I.M.Vinogradov. *Izv. Akad. Nauk SSSR, Ser. Mat.* 51, 2. Translated in *Math. USSR Izvestiya* 30 (1988), 2, 337-351.

VAUGHAN, R.C. (1977). On pairs of additive cubic equations. *Proc. Lond. Math. Soc.* (3), 34, 354-364.

VAUGHAN, R.C. (1981a). Some remarks on Weyl sums. *Topics in classical number theory, Colloquia Mathematica Societatis János Bolyai* 34 (*Budapest*, 1981).

VAUGHAN, R.C. (1981b). The Hardy-Littlewood Method. *Cambridge University Press, Cambridge.*

VAUGHAN, R.C. (1986a). On Waring's problem for cubes. *J. reine angew. Math.*, 365, 122-170.

VAUGHAN, R.C. (1986b). On Waring's problem for smaller exponents. *Proc. Lond. Math. Soc.* (3), 52, 445-463.

VAUGHAN, R.C. (1989a). A new iterative method in Waring's problem. *Acta Math.*, 162, 1-71.

VAUGHAN, R.C. (1989b). On Waring's Problem for cubes II. *J. Lond. Math. Soc.* (2), 39, 205-218.

VAUGHAN, R.C. (1989c). A new iterative method in Waring's problem II. *J. Lond. Math. Soc.* (2), 39, 219-230.

VINOGRADOV, I.M. (1934). On the upper bound G(n) in Waring's problem. *Izv. Akad. Nauk SSSR*, 10, 1455-1469.

VINOGRADOV, I.M. (1959). On an upper bound for G(n). *Izv. Akad. Nauk SSSR*, 23, 637-642.