Simultaneous Diophantine Approximation of Rationals by Rationals

J. C. Lagarias

Bell Laboratories
Murray Hill, NJ 07974

1. Introduction

This paper studies the behavior of simultaneous Diophantine approximations of vectors $\boldsymbol{\alpha} = (\frac{a_1}{B}, \ldots, \frac{a_n}{B})$ of rational numbers by vectors $\boldsymbol{\xi} = (\frac{x_1}{x}, \ldots, \frac{x_n}{x})$ of rational numbers with a smaller denominator $1 \le x < B$. To exclude the case that $\boldsymbol{\xi}$ perfectly approximates $\boldsymbol{\alpha}$ we suppose that $\boldsymbol{\alpha} = (\frac{a_1}{B}, \ldots, \frac{a_n}{B})$ is *primitive* in the sense that B is the minimal common denominator for the entries of $\boldsymbol{\alpha}$, i.e. that

$$g.c.d.\ (a_1, \ldots, a_n, B) = 1\ . \tag{1.1}$$

In general it seems difficult to determine the set of good simultaneous Diophantine approximations to an individual vector $\boldsymbol{\alpha}$, and such sets may be badly behaved, cf. [2],[3]. However it is possible to say more about the general behavior of such sets of approximations when averaged over appropriately chosen ensembles of vectors $\boldsymbol{\alpha}$, and this is the object of this paper. In particular we show (Theorem 1.5) a quantitative version of the assertion: Most primitive vectors $\boldsymbol{\alpha}$ that have at least one "unusually good" simultaneous Diophantine approximation $\boldsymbol{\xi}$ do not have very many such "unusually good" approximations. Here the vectors $\boldsymbol{\xi}$ are *not* required to be primitive but must have a denominator x with $1 \le x < B$.

In order to state our results we introduce some definitions and notation. First, we need a measure of goodness of approximation. We say that a vector $\boldsymbol{\xi}$ is a Δ-*good* *approximation* to a vector $\boldsymbol{\alpha}$ if

$$\left| \frac{a_i}{B} - \frac{x_i}{x} \right| \le \frac{\Delta}{Bx} \ ; \quad \text{for} \quad 1 \le i \le n \ . \tag{1.3}$$

This is the usual sup norm measure of approximation since (1.3) may be rewritten as

$$\{\{x\,\boldsymbol{\alpha}\}\} \le \frac{\Delta}{B}$$

where

$$\{\{\boldsymbol{\beta}\}\} = \underset{y \in \mathbf{Z}^n}{\mathrm{MIN}} \left(\underset{1 \le i \le n}{\mathrm{MAX}} \ |\beta_i - y_i| \right) \ .$$

We let $N(\boldsymbol{\alpha}, \Delta)$ denote the number of Δ-good approximation vectors to $\boldsymbol{\alpha}$, i.e. the number of solutions $(x, x_1, \ldots, x_n)$ to (1.3) with $1 \le x < B$. By clearing denominators in (1.3) we see that $N(\boldsymbol{\alpha}, \Delta)$ may alternatively be interpreted as the number of solutions $(x, x_1, \ldots, x_n)$ of the integer programming problem:

$$-\Delta \le a_i x - B x_i \le \Delta ; \quad 1 \le i \le n \ , \tag{1.4a}$$

$$1 \le x < B \ , \tag{1.4b}$$

$$x, x_1, \ldots, x_n \quad \text{integers} \ . \tag{1.4c}$$

Second, in order to formulate assertions about the behavior of "most" vectors $\boldsymbol{\alpha}$, we need to specify the sets of such vectors we are studying. We note that if $\mathbf{y} \in \mathbf{Z}^n$ is an integer lattice point, then we have

$$N(\boldsymbol{\alpha}, \Delta) = N(\boldsymbol{\alpha} + \mathbf{y}, \Delta) \ ,$$

by examining (1.4). Hence we may without loss of generality restrict our attention to primitive vectors $\boldsymbol{\alpha}$ lying in the half-open unit cube $[0, 1)^n$ in $\mathbf{R}^n$, i.e. we suppose that $\boldsymbol{\alpha} = \left(\frac{a_1}{B}, \ldots, \frac{a_n}{B} \right)$ has

$$0 \le a_i < B; \quad 1 \le i \le n \ . \tag{1.5}$$

We study the sets $S_n(B)$ consisting of all primitive vectors $\boldsymbol{\alpha}$ with denominator B, which lie in $[0,\ 1)^n$, i.e.

$$S_n(B) = \{\boldsymbol{\alpha} = (\frac{a_1}{B},\ldots,\ \frac{a_n}{B}):\ 0 \le a_i < B \ \text{ and } \ g.c.d.\ (a_1,\ldots,\ a_n,\ B) = 1\} \ . \tag{1.6}$$

Our object is to analyze the behavior of the function $N(\boldsymbol{\alpha},\ \Delta)$ viewed as a random variable on the sets $S_n(B)$ which we treat as discrete probability spaces with the uniform distribution. Our first result is an estimate for the mean value of $N(\boldsymbol{\alpha},\ \Delta)$. Here $d(B)$ denotes the number of divisors of B.

Theorem 1.1. For $n \ge 2$,

$$\sum_{\boldsymbol{\alpha} \in S_n(B)} N(\boldsymbol{\alpha},\ \Delta) = 2^n \psi_n(B) B \Delta^n + O(n 3^{n+1} d(B)^2 B \Delta^{n-1}) \ , \tag{1.7}$$

where $\psi_n(B)$ *is the multiplicative function defined by*

$$\psi_n(B) = \prod_{p|B} (1 - p^{-n}) \ , \tag{1.8}$$

and the constant implied by the O-symbol is independent of n and B.

Note that $\psi_n(B)$ is bounded away from zero, and in fact for $n \ge 2$

$$1 \ge \psi_n(B) \ge [\zeta(n)]^{-1} \ ,$$

where ζ denotes Riemann's zeta function.

It is easy to verify that the set $S_n(B)$ contains exactly $\psi_n(B) B^n$ elements, and combining this result with Theorem 1.1 immediately gives the following corollary.

Corollary 1.2. For $n \geq 2$,

$$\text{Prob}\ \{\boldsymbol{\alpha} \in S_n(B) \ \text{ has } \ N(\boldsymbol{\alpha}, \Delta) \geq 1\} \leq 2^n \frac{\Delta^n}{B^{n-1}} + O(n 3^{n+1} d(B)^2 (\frac{\Delta}{B})^{n-1}) \ . \tag{1.9}$$

Dirichlet's theorem for simultaneous Diophantine approximation implies that for $\Delta = B^{1 - \frac{1}{n}}$, $N(\boldsymbol{\alpha}, \Delta) \geq 1$ for all $\boldsymbol{\alpha}$, so that in this case

$$\text{Prob}\ \{\boldsymbol{\alpha} \in S_n(B) : \ N(\boldsymbol{\alpha}, B^{1 - \frac{1}{n}}) \geq 1\} = 1 \ . \tag{1.10}$$

Corollary 1.2 implies that when the dimension n is held fixed

$$\text{Prob}\ \{\boldsymbol{\alpha} \in S_n(B) : \ N(\boldsymbol{\alpha}, \Delta) \geq 1\} \to 0 \quad \text{as} \quad B \to \infty \ , \tag{1.11}$$

provided $\Delta = o(B^{1 - \frac{1}{n}})$ as $B \to \infty$; this is a quantitative version for rationals in $S_n(B)$ of the assertion that "most" vectors do not have simultaneous Diophantine approximations significantly better than those guaranteed to exist by Dirichlet's theorem.

Our main result is an upper bound for the second moment of $N(\boldsymbol{\alpha}, \Delta)$ on the set $S_n(B)$.

Theorem 1.3. For each $n \geq 2$ there are positive constants $c_{n,1}$, $c_{n,2}$ and $c_{n,3}$ such that

$$\sum_{\boldsymbol{\alpha} \in S_n(B)} N(\boldsymbol{\alpha}, \Delta)^2 \leq c_{n,1} B^2 (\frac{\Delta^2}{B})^n + c_{n,2} B \Delta^n + R_n(B, \Delta) \tag{1.12}$$

where the remainder term $R_n(B, \Delta)$ is given by

$$R_n(B, \Delta) = \begin{cases} c_{n,3} \; (d(B)^2 B \Delta^{n-1}) & \text{for} \;\; n \geq 4, \\ c_{3,3} \; (d(B)^2 B \Delta^3) & \text{for} \;\; n = 3, \\ c_{2,3} \; (d(B)^3 B \Delta^2) & \text{for} \;\; n = 2 \; . \end{cases} \tag{1.13}$$

For *prime B* the bounds (1.12) and (1.13) simplify for all $n \geq 2$ to

$$\sum_{\boldsymbol{\alpha} \in S_n(B)} N(\boldsymbol{\alpha}, \Delta)^2 \ll_n B^2 (\frac{\Delta^2}{B})^n + B\Delta^n \; . \tag{1.14}$$

Here $\ll_n$ is the Vinogradov notation, which says that the left side of (1.14) is less than the right side of (1.14) times a positive constant depending on n. This bound for prime B is the correct order of magnitude in that it can be shown that

$$\sum_{\boldsymbol{\alpha} \in S_n(B)} N(\boldsymbol{\alpha}, \Delta)^2 \gg_n B^2 (\frac{\Delta^2}{B})^n + B\Delta^n \; , \tag{1.15}$$

see the discussion at the end of Section 4. In (1.15) the term $B\Delta^n$ comes from a small set of $\boldsymbol{\alpha}$ having many Δ-good approximations, while the term $c_{n,3} \; B^2 (\frac{\Delta^2}{B})^n$ is associated with the contribution of the "average" value of $N(\boldsymbol{\alpha}, \Delta)$ on the set of those $\boldsymbol{\alpha}$ with $N(\boldsymbol{\alpha}, \Delta) \geq 1$. The $B^2 (\frac{\Delta^2}{B})^n$ term dominates (1.12) when $\Delta \geq B^{1 - \frac{1}{n}}$, while the $B\Delta^n$ term dominates (1.12) when $\Delta \leq B^{1 - \frac{1}{n}}$.

It is necessary to restrict ourselves to primitive vectors $\boldsymbol{\alpha}$ with denominator B in the ensemble $S_n(B)$ in Theorem 1.3, because the inequality (1.12) does not hold in general for composite B if the left side of (1.12) is enlarged to sum over all $\boldsymbol{\alpha}$, primitive and imprimitive, in $[0, 1)^n$ having denominator B. The inequality fails in this case because there is a large contribution from perfect approximations with smaller denominators.

The proof of Theorem 1.3 involves an auxiliary problem of independent interest. This concerns the distribution of the number of solutions of homogeneous linear congruences in two variables with bounds on the variables. We consider the linear congruence

$$\lambda x_1 \equiv x_2 \pmod{B} , \tag{1.15}$$

subject to the constraints

$$|x_1| \le \Delta_1 , \tag{1.16}$$

$$|x_2| \le \Delta_2 . \tag{1.17}$$

Let $f(\lambda, B, \Delta_1, \Delta_2)$ denote the number of solutions of (1.15), (1.16) and (1.17). Note that $f(\lambda, B, \Delta_1, \Delta_2) \ge 1$ because $x_1 = x_2 = 0$ is always a solution. We study the sums

$$Q_n(B, \Delta_1, \Delta_2) = \sum_{(\lambda, B) = 1} f(\lambda, B, \Delta_1, \Delta_2)^n$$

Let $T(B, \Delta_1, \Delta_2)$ denote the number of $f(\lambda, B, \Delta_1, \Delta_2)$ equal to 1 with $1 \le \lambda \le B$ with $(\lambda, B) = 1$. We prove the following result, where $\phi(B)$ is Euler's ϕ-function.

Theorem 1.4. For $n \ge 2$ and $\Delta_1 \le \Delta_2$, there are positive constants c_n and c_n^* such that

$$Q_n(B, \Delta_1, \Delta_2) - T(B, \Delta_1, \Delta_2) \le c_n^* \phi(B) \left(\frac{\Delta_1 \Delta_2}{B}\right)^n + R_n(B, \Delta_1, \Delta_2) \tag{1.18}$$

where the remainder term $R_n^*(B, \Delta_1, \Delta_2)$ is

$$R_n^*(B, \Delta_1, \Delta_2) = \begin{cases} c_n \ \Delta_1^{n-1}\Delta_2 & \text{if} \quad n \geq 3, \qquad (1.19) \\ c_2 \ \Delta_1\Delta_2 \ \log \Delta_1 (\sum_{d|B} d^{-1}) & \text{if} \quad n = 2 \ . \qquad (1.20) \end{cases}$$

The important feature of Theorem 1.4 for applications is the asymmetric form of the remainder term $\Delta_1^{n-1}\Delta_2$ for $n \geq 3$ in (1.19), where $\Delta_1 \leq \Delta_2$. The inequality (1.18) is best possible when B is prime and $n \geq 3$, in the sense that

$$Q_n(B, \Delta_1, \Delta_2) - T(B, \Delta_1, \Delta_2) \gg_n B\left(\frac{\Delta_1\Delta_2}{B}\right)^n + \Delta_1^{n-1}\Delta_2$$

in that case, cf. equation (3.4).

We combine Theorems 1.1. and 1.2 to get information about the distribution of $N(\boldsymbol{\alpha}, \Delta)$ on those α for which $N(\boldsymbol{\alpha}, \Delta) \geq 1$. We prove the following result.

Theorem 1.5. For each $n \geq 4$ there is a positive constant c_n^{**} such that for all $B \geq 2$ and all Δ satisfying

$$c_n^{**} d(B)^2 \leq \Delta \leq B^{1 - \frac{1}{n}} ,$$

we have

$$\text{Prob } \{\boldsymbol{\alpha} \in S_n(B) \ \ \textit{has} \ \ N(\boldsymbol{\alpha}, \Delta) \geq k \,|\, N(\boldsymbol{\alpha}, \Delta) \geq 1\} \leq \frac{c_n^{**}}{k^2} , \qquad (1.21)$$

for all $k \geq 1$. The same result holds for $n = 2$ and 3 provided B is restricted to be prime.

It is possible that (1.21) gives the correct size of the tail of this conditional probability distribution as a function of k, apart from the size of the constant c_n^{**}; I am unable to prove this.

Theorem 1.5 has applications to the cryptanalysis of public key cryptosystems of knapsack type. In particular it can be used to show that Shamir's attack [5] on the Basic Merkle-Hellman knapsack scheme succeeds in polynomial time for "almost all" knapsacks which encrypt at a fixed information rate R, as the number of knapsack items $n \to \infty$, for any R with $0 < R < 1$, see [2],[3]. (Shamir [5] showed this for $\frac{1}{2} < R < 1$.) These cryptanalytic applications motivated my study of the questions in this paper.

Finally we remark that the proofs of the theorems are substantially complicated by the inclusion-exclusion arguments needed to treat the case of composite B. The proofs simplify considerably in the special case that B is prime.

2. Bounding the mean value of $N(\boldsymbol{\alpha}, \Delta)$

Proof of Theorem 1.1. We let

$$G_n(B, \Delta) = \sum_{\alpha \in S_n(B)} N(\boldsymbol{\alpha}, \Delta) , \tag{2.1}$$

i.e. $G_n(B, \Delta)$ denotes the number of approximation pairs $(x, \boldsymbol{\alpha})$ where $\alpha = (\frac{a_1}{B}, \ldots, \frac{a_n}{B})$ such that

$$\left| \frac{a_i}{B} - \frac{x_i}{x} \right| \le \frac{\Delta}{Bx} \tag{2.2}$$

subject to

$$1 \le x \le B-1 \ , \tag{2.3}$$

$$0 \le a_i \le B-1 \quad \text{for} \quad 1 \le i \le n \ , \tag{2.4}$$

$$g.c.d.\ (a_1, \ldots, a_n, B) = 1 \ . \tag{2.5}$$

Our goal is to estimate $G_n(B, \Delta)$.

We use inclusion-exclusion to reduce the problem to the case where the relative primality condition (2.5) is omitted. For $d | B$ let $H_n(B, \Delta, d)$ denote the number of solutions to (2.2)-(2.4) with

$$d | (a_1, \ldots, a_n, B) \ . \tag{2.6}$$

Then by Mobius inversion

$$G_n(B, \Delta) = \sum_{d|B} \mu(d) H_n(B, \Delta, d) \ . \tag{2.7}$$

To estimate $H_n(B, \Delta, d)$, we set $a_i = da_i^*$ and $x_i = dy_i^*$ for $1 \le i \le n$, and rewrite the conditions defining $H_n(B, \Delta, d)$ as

$$a_i^* x \equiv y_i^* \quad (\text{mod } \frac{B}{d}) \ , \tag{2.8a}$$

$$|y_i^*| \le \frac{\Delta}{d} \quad \text{for} \quad 1 \le i \le n \ , \tag{2.8b}$$

subject to

$$1 \le x \le B-1 \ , \tag{2.8c}$$

$$0 \le a_i^* \le \frac{B}{d} - 1 \quad \text{for} \quad 1 \le i \le n \ . \tag{2.8d}$$

Next we let $H_n(B, \Delta, d, k)$ denote the number of solutions to (2.8) for which $(x, \frac{B}{d}) = k$, and we have

$$H_n(B, \Delta, d) = \sum_{k \mid \frac{B}{d}} H_n(B, \Delta, d, k) \ . \tag{2.9}$$

To estimate $H_n(B, \Delta, d, k)$, we set $x = x^* k$, $y_i^* = y_i^{**} k$ and rewrite the conditions defining $H_n(B, \Delta, d, k)$ as

$$a_i^* x^* \equiv y_i^{**} \quad (\mathrm{mod}\ \frac{B}{kd}) \ , \tag{2.10a}$$

$$|y_i^{**}| \le \frac{\Delta}{kd} \ , \tag{2.10b}$$

subject to

$$1 \le x^* \le \frac{B}{k} - 1 \ , \tag{2.10c}$$

$$0 \le a_i^* \le \frac{B}{d} - 1 \ , \tag{2.10d}$$

$$(x^*, \frac{B}{kd}) = 1 \ . \tag{2.10e}$$

We can count solutions to (2.10) directly by observing that the number of choices of x^* satisfying (2.10c) and (2.10e) is

$$\begin{cases} d\phi\left(\frac{B}{kd}\right) \quad \text{if} \quad kd < B \ , & (2.11a) \\ \frac{B}{k} - 1 \quad \text{if} \quad kd = B \ . & (2.11b) \end{cases}$$

For each choice of x^* there are $1 + 2[\frac{\Delta}{kd}]$ choices for each y_i^{**} in (2.10b), which determines a_i^* $(\mathrm{mod}\ \frac{B}{kd})$ uniquely, and there are then k choices for a_i^* $(\mathrm{mod}\ \frac{B}{d})$ satisfying (2.10d). Hence

$$H_n(B,\ \Delta,\ d,\ k)\ =\ \begin{cases} d\phi\ (\frac{B}{kd})\ k^n\ (1\ +\ 2[\frac{\Delta}{kd}])^n & \text{if}\quad kd\ <\ B, \\ (d-1)k^n\ (1\ +\ 2[\frac{\Delta}{B}])^n\ . & \text{if}\quad kd\ =\ B\ . \end{cases}$$

We may rewrite this as

$$H_n(B,\ \Delta,\ d,\ k)\ +\ [d\phi\ (\frac{B}{kd})\ -\ \delta(\frac{B}{kd})]k^n(1\ +\ 2[\frac{\Delta}{kd}])^n\ , \tag{2.12}$$

where we define $\phi(1)\ =\ 1$ and

$$\delta(n)\ =\ \begin{cases} 1\ \ \text{if}\quad n\ =\ 1, \\ 0\ \ \text{if}\quad n\ >\ 1\ . \end{cases} \tag{2.13}$$

Substituting this expression for $H_n(B,\ \Delta,\ d,\ k)$ into (2.7) and (2.9) gives

$$G_n(B,\ \Delta)\ =\ \sum_{d|B}\mu(d)\ [\sum_{k|\frac{B}{d}}\ [d\phi\ (\frac{B}{kd})\ -\ d(\frac{B}{kd})]^n k(1\ +\ 2[\frac{\Delta}{kd}])^n\ . \tag{2.14}$$

To estimate this sum, we use the approximation

$$k^n(1\ +\ 2[\frac{\Delta}{kd}])^n\ =\ k^n\ +\ 2^n(\frac{\Delta}{d})^n\ +\ \varepsilon_n(\Delta,\ k,\ d)\ , \tag{2.15}$$

where $\varepsilon_n(\Delta,\ k,\ d)$ is a remainder term. This decomposes the sum (2.14) for $G_n(B,\ \Delta)$ into three sums, which we call $T_1,\ T_2,\ T_3$ respectively and estimate separately.

For the first sum

$$T_1\ =\ \sum_{d|B}\mu(d)\ \sum_{k|\frac{B}{d}}\ [d\phi\ (\frac{B}{kd})\ -\ \delta\ (\frac{B}{kd})]k^n\ ,$$

we interchange the order of summation to get

$$T_1 = \sum_{k|B} k^n \{ \sum_{d|\frac{B}{k}} \mu(d)\ d\phi\ (\frac{B}{kd}) - \mu(\frac{B}{k}) \} . \tag{2.16}$$

We now use the identity valid for all $M \geq 1$ that

$$\sum_{d|M} \mu(d)\ d\phi\ (\frac{M}{d}) = \mu(M) \tag{2.17}$$

to conclude that all the inner sums in (2.16) are zero, and hence that

$$T_1 = 0 . \tag{2.18}$$

To verify (2.17), use the identity

$$\phi(M) = \sum_{d|M} \mu(d)\ \frac{M}{d} \tag{2.19}$$

to get

$$\begin{aligned} \sum_{d|M} \mu(d)\ d\phi\ (\frac{M}{d}) &= \sum_{d|M} \sum_{e|\frac{M}{d}} \mu(d)\mu(e)\ \frac{M}{e} \\ &= \sum_{e|M} \mu(e)\ \frac{M}{e}\ (\sum_{d|\frac{M}{e}} \mu(d)) . \end{aligned} \tag{2.20}$$

Now the inner sum in (2.20) is zero except when $e = M$, so the right side of (2.20) is just $\mu(B)$, proving (2.17).

For the second sum,

$$T_2 = \sum_{d|B} \mu(d) \sum_{k|\frac{B}{d}} [d\phi\ (\frac{B}{kd}) - \delta\ (\frac{B}{kd})]2^n\ (\frac{\Delta}{n})^n ,$$

we have

$$T_2 = 2^n \Delta^n \sum_{d|B} \mu(d) d^{-n+1} \left(\sum_{k|\frac{B}{d}} \phi(\frac{B}{kd}) \right)$$

$$+ 2^n \Delta^n \sum_{d|B} \mu(d) \ . \tag{2.21}$$

For $B > 1$ the second sum in (2.21) is zero, and using the identity

$$M = \sum_{d|M} \phi\,(\frac{M}{d})$$

the sum (2.21) becomes

$$T_2 = 2^n \Delta^n B \sum_{d|B} \mu(d) d^{-n} = 2^n \psi_n(B) B \Delta^n \ , \tag{2.22}$$

for $B > 1$.

To estimate the third sum

$$T_3 = \sum_{d|B} \mu(d) \sum_{k|\frac{B}{d}} [d\phi\,(\frac{B}{kd}) - \delta(\frac{B}{kd})] \varepsilon_n(\Delta, d, k) \tag{2.23}$$

we need bounds for the remainder terms $\varepsilon_n(\Delta, d, k)$. We collect these bounds in the following lemma.

Lemma 2.1. If $k > \Delta$,

$$\varepsilon_n(\Delta, d, k) = -2^n \, (\frac{\Delta}{n})^n \ . \tag{2.24}$$

If $k \le \Delta$ *then*

$$|\varepsilon_n(\Delta, d, k)| \le nk(k + 2\,\frac{\Delta}{d})^{n-1} + 3^n \Delta^{n-1} \ . \tag{2.25}$$

Proof. To prove the first part, if $k > \Delta$ then $[\frac{\Delta}{kd}] = 0$ so that

$$k^n (1 + 2[\frac{\Delta}{dk}])^n = k^n$$

Hence

$$\varepsilon_n(\Delta, k, d)^n = -2^n (\frac{\Delta}{d})^n .$$

For the second part,

$$|\varepsilon_n(\Delta, k, d)| \le |k^n (1 + 2[\frac{\Delta}{kd}])^n - k^n (1 + 2\frac{\Delta}{kd})^n|$$

$$+ |k^n (1 + 2\frac{\Delta}{kd})^n - k^n - 2^n (\frac{\Delta}{d})^n| . \qquad (2.26)$$

Now

$$|k^n (1 + 2[\frac{\Delta}{kd}])^n - k^n (1 + 2\frac{\Delta}{kd})^n| \le k^n n (1 + 2\frac{\Delta}{kd})^{n-1} (\frac{\Delta}{kd} - [\frac{\Delta}{kd}])$$

$$\le n k^n (1 + 2\frac{\Delta}{kd})^{n-1} , \qquad (2.27)$$

where we used

$$|\frac{x_1^n - x_2^n}{x_1 - x_2}| \le n x_1^{n-1} \quad \text{when} \quad x_1 \ge x_2 \ge 1 .$$

Next, using $k \le \Delta$, we have

$$|k^n (1 + 2\frac{\Delta}{kd})^n - k^n - 2^n (\frac{\Delta}{d})^n| = |\sum_{i=1}^{n-1} \binom{n}{i} k^i (\frac{\Delta}{kd})^{n-1}|$$

$$\le \Delta^{n-1} |\sum_{i=1}^{n-1} \binom{n}{i} 2^i (\frac{1}{kd})^{n-1}| \le 3^n \Delta^{n-1} . \qquad (2.28)$$

Combining (2.27) and (2.28) gives (2.25). ■

To estimate T_3 we split the sum into the sum S_1 of the terms with $k > \Delta$ and the sum S_2 of the terms with $k \leq \Delta$. Now

$$S_1 = \sum_{d|B} \mu(d) \sum_{\substack{k|\frac{B}{d} \\ k>\Delta}} [d\phi(\frac{B}{kd}) - \delta(\frac{B}{kd})](-2^n(\frac{\Delta}{d})^n) . \tag{2.24}$$

Using $\phi(M) \leq M$ we have

$$|d\phi(\frac{B}{kd}) - \delta(\frac{B}{kd})| \leq \frac{B}{k} , \tag{2.25}$$

Applying this in (2.24) and using $\frac{B}{k} < \frac{B}{\Delta}$ since $k > \Delta$ gives

$$|S_1| \leq 2^n \Delta^{n-1} B |\sum_{d|B} \mu(d) d^{-n}| \leq \zeta(2) 2^n B \Delta^{n-1} \tag{2.26}$$

for $n \geq 2$. Next Lemma 2.1 gives

$$\begin{aligned} |S_2| \leq & \sum_{d|B} \sum_{\substack{k|\frac{B}{d} \\ k<\Delta}} |d(\phi(\frac{B}{kd}) - \delta(\frac{B}{kd})| nk(k+2\frac{\Delta}{d})^{n-1} \\ & + \sum_{d|B} \sum_{\substack{k|\frac{B}{d} \\ k<\Delta}} |d\phi(\frac{B}{kd}) - \delta(\frac{B}{kd})| 3^n \Delta^{n-1}| \end{aligned} \tag{2.29}$$

Using (2.25) and the fact that $k \leq \Delta$ implies $k + 2\frac{\Delta}{d} \leq 3\Delta$ we get

$$|S_2| \leq \sum_{d|B} \sum_{k|\frac{B}{d}} (n+1) B 3^n \Delta^{n-1}$$

which implies

$$|S_2| \leq (n+1) 3^n B d(B)^2 \Delta^{n-1} \tag{2.28}$$

where $d(B)$ is the number of divisors of B. Combining (2.26) and (2.28) we obtain

$$|T_3| \le n 3^{n+1} B d(B)^2 \Delta^{n-1} \ . \tag{2.29}$$

Theorem 1.1 follows from (2.18), (2.22) and (2.29). ■

Proof of Corollary 1.2. The cardinality $|S_n(B)|$ of the set $S_n(B)$ is

$$|S_n(B)| = \psi_n(B) B^n \ . \tag{2.30}$$

To see this, note that for each $q|B$ the number of elements $\left(\frac{a_1}{B}, \ldots, \frac{a_n}{B}\right)$ for which $q | g.c.d.\ (a_1, \ldots, a_n)$ is exactly $q^{-n} B^n$. Hence by inclusion-exclusion

$$\begin{aligned} |S_n(B)| &= \sum_{q|B} \mu(q) q^{-n} B^n \\ &= \prod_{p|B} (1 - p^{-n}) B^n = \psi_n(B) B^n \ . \end{aligned}$$

Then

$$\text{Prob } \{\boldsymbol{\alpha} \in S_n(B) \text{ has } N(\boldsymbol{\alpha}, \Delta) \ge 1\} \le \frac{\sum_{\boldsymbol{\alpha} \in S_n(B)} N(\boldsymbol{\alpha}, \Delta)}{|S_n(B)|} \ ,$$

and the corollary follows from Theorem 1.1 and (2.30). ■

3. Small solutions of linear congruences

We consider the problem of counting the number $f(\lambda; B, \Delta_1, \Delta_2)$ of solutions to the linear congruence

$$\lambda x_1 \equiv x_2 \pmod{B} \tag{3.1}$$

subject to the constraints

$$|x_1| \le \Delta_1 \qquad (3.2)$$
$$|x_2| \le \Delta_2 \ .$$

This number fluctuates as a function of λ in an irregular way, related to the Diophantine approximation properties of the number $\frac{\lambda}{B}$. We are interested in the behavior of the numbers $f(\lambda;\, B,\, \Delta_1,\, \Delta_2)$ averaged over all λ with $1 \le \lambda < B$ and $(\lambda,\, B) = 1$, as measured by the quantities

$$Q_n(B;\, \Delta_1,\, \Delta_2) = \sum_{\substack{1 \le \lambda < B \\ (\lambda,\, B) = 1}} f(\lambda;\, B,\, \Delta_1,\, \Delta_2)^n \ . \qquad (3.3)$$

Our goal is to estimate the sums $Q_n(B,\, \Delta_1,\, \Delta_2)$. How large do we expect them to be? First, if B is prime then

$$\sum_{\lambda=1}^{B} f(\lambda;\, B,\, \Delta_1,\, \Delta_2) \sim 4\, \Delta_1 \Delta_2 \ ,$$

since each pair $(x_1,\, x_2)$ satisfying (3.2) with $x_1 x_2 \ne 0$ determines a unique λ in (3.1). In this case the average size of $f(\lambda;\, B,\, \Delta_1,\, \Delta_2)$ is $\frac{4 \Delta_1 \Delta_2}{B}$, so we must get a contribution to $Q_n(B,\, \Delta_1,\, \Delta_2)$ of at least of $4^n \phi(B)\, (\frac{\Delta_1 \Delta_2}{B})^n$. Second, there is a large contribution from certain λ's. For example for $\Delta_1 \le \Delta_2$ and $1 \le \lambda \le \frac{\Delta_2}{\Delta_1}$ we have

$$f(\lambda;\, B,\, \Delta_1,\, \Delta_2) = 1 + 2[\Delta_1] \ ,$$

so this range of λ contributes at least $2^n \Delta_1^{n-1} \Delta_2$ to the sum. Third, all $f(\lambda;\, B,\, \Delta_1,\, \Delta_2) \ge 1$ since $x_1 = x_2 = 0$ is always a solution. Let $T(B,\, \Delta_1,\, \Delta_2)$

denote the number of λ with $(\lambda, B) = 1$ for which $f(\lambda; B, \Delta_1, \Delta_2) = 1$. This discussion implies that, *for prime B*, we have

$$Q_n(B; \Delta_1, \Delta_2) - T(B; \Delta_1, \Delta_2) \underset{n}{\gg} B\left(\frac{\Delta_1 \Delta_2}{B}\right)^n + \Delta_1^{n-1}\Delta_2 \ . \qquad (3.4)$$

Theorem 1.4 asserts that these three contributions dominate the sum.

Proof of Theorem 1.4. We shall treat B, Δ_1, Δ_2 as fixed, and abbreviate $f(\lambda; B, \Delta_1, \Delta_2)$ to $f(\lambda)$. Since $f(\lambda) = f(B-\lambda)$ we assume that $0 \le \lambda \le B/2$.

The size of $f(\lambda)$ is determined by the continued fraction expansion of $\frac{\lambda}{B}$. Write

$$\frac{\lambda}{B} = [0; a_1, a_2, \ldots, a_m] \ ,$$

where $a_1 \ge 2$ since $0 \le \frac{\lambda}{B} < \frac{1}{2}$. Let $\{\frac{p_k}{q_k}: \ 0 \le k \le m\}$ denote the convergents to $\frac{\lambda}{B}$, with $q_0 = 1, \ q_1 \ge 2$. Define k by

$$q_{k-1} < q_k \le \Delta_1 < q_{k+1} \ . \qquad (3.5)$$

and set

$$\lambda q_j \equiv s_j \pmod{B}; \quad -\frac{B}{2} < s_j \le \frac{B}{2} \ . \qquad (3.6)$$

We recall the following facts from the basic theory of ordinary continued fractions [cf. Lang [6], pp. 1-6].

(i) The $\{s_j\}$ alternate in sign, with $s_0 > 0$ and

$$s_j = \lambda q_j - Bp_j \ . \qquad (3.7)$$

(ii) $|s_0| > |s_1| > |s_2| > \dots > |s_m| = 0.$

We also have

$$|s_j| \geq \frac{B}{2q_{j+1}} \ . \tag{3.8}$$

To prove (3.8), we observe that

$$\begin{aligned} |s_j| &= |\lambda q_j - Bp_j| = q_j B \left| \frac{\lambda}{B} - \frac{p_j}{q_j} \right| \\ &\geq \frac{q_j B}{2} \left| \frac{p_j}{q_j} - \frac{p_{j+1}}{q_{j+1}} \right| = \frac{B}{2q_{j+1}} \ . \end{aligned}$$

We transform the problem slightly, using the hypothesis $(\lambda,\ B) = 1$, to get

$$f(\lambda) = 1 + 2H(\lambda) \tag{3.9}$$

where $H(\lambda)$ is the number of solutions of

$$\left.\begin{array}{c} \lambda x_1 \equiv x_2 \quad (\mathrm{mod}\ B) \\ 1 \leq x_1 \leq \Delta_1 \\ 1 \leq |x_2| \leq \Delta_2 \end{array}\right\} \tag{3.10}$$

To proceed further, we will use classical results of Halton [1], see also Slater ([8], eqns. (30),(31),(34)), concerning the distribution of the least nonnegative residues $\lambda x_1 \pmod B$ for $1 \leq x_1 \leq \Delta_1$. He showed that these points partition the interval $[0,\ B)$ into subintervals of at most three different lengths, and that the shortest of these three lengths is $|s_k|$. Correspondingly, the solutions x_2 of (3.10) plus $x_2 = 0$ partition the interval $[-\Delta_2,\ \Delta_2]$ into subintervals of at most three different lengths, excluding the two subintervals containing the endpoints. In this partition, we label each such interval by the solution to (3.10) that is its left endpoint if $s_k < 0$, and its right endpoint if $s_k > 0$. We will call intervals of length $|s_k|$ *short subintervals* and all other intervals *long subintervals.* Let $S(\lambda) = S(\lambda; B, \Delta_1, \Delta_2)$ denote the number of short subintervals, and let $L(\lambda) = L(\lambda; B, \Delta_1, \Delta_2)$ denote the number of long subintervals. See Figure 1 for an example of these definitions. The arrows on each subinterval indicate the labelled endpoint. Since each solution to (3.10) is assigned a subinterval, plus $x = 0$, we have

$$1 + H(\lambda) = S(\lambda) + L(\lambda) \ . \tag{3.11}$$

We will use two different sets of bounds to estimate $f(\lambda)$ via (3.9). The first bound is the direct estimate

$$H(\lambda) \leq 2 \ \frac{\lceil \Delta_2 \rceil}{|s_k|} \ . \tag{3.12}$$

This holds because the interval $[-\Delta_2, \Delta_2]$ has length $2\Delta_2$ and all subintervals having a labelled endpoint have length $\geq |s_k|$ with at most one exception. The second set of bounds is given by the following lemma.

Lemma 3.1

$$\text{(i)} \qquad L(\lambda) \leq \frac{2\Delta_2}{|s_{k-1}|} + 1 \tag{3.13}$$

$$\text{(ii)} \qquad S(\lambda) \leq \frac{\Delta_1}{q_k} \ \left(2 \ \frac{\Delta_2}{|s_{k-1}|} + 1\right) \ . \tag{3.14}$$

Proof of Lemma 3.1. To prove (i), we observe that the long subintervals associated to $1 \leq x_1 \leq \Delta_1$ are in 1-1 correspondence with the complete set of subintervals present for $1 \leq x_1 \leq q_k - 1$. (Each step of x_1 after that creates a new short subinterval on $[0, B]$ of

Figure 1. Steps for $\lambda / B = 8/19$.

$$B = 19, \quad \lambda = 8, \quad \Delta_2 = 7$$

$$\text{Continued fraction of } \theta = \frac{8}{19} = [0, 2, 2, 1, 2].$$

j	a_j	p_j	q_j	s_j
0	0	0	1	8
1	2	1	2	-3
2	2	2	5	2
3	1	3	7	-1
4	2	8	19	0

Case 1. $\Delta_1 = 6$.

Length of short subinterval = 2 $\quad S(8;\ 11,\ 6,\ 7) = 1$
Length of long subinterval = 3 $\quad L(8;\ 11,\ 6,\ 7) = 4$

Case 2. $\Delta_1 = 3$.

Length of short subinterval = 3 $\quad S(8;\ 11,\ 3,\ 7) = 1$
Length of long subintervals = 5,8 $\quad L(8;\ 11,\ 3,\ 7) = 2$

length $|s_k|$) Since at step $q_k - 1$ all the short subintervals at that time have length $\geq |s_{k-1}|$ there are at most $\dfrac{2\Delta_2}{|s_{k-1}|} + 1$ of them, using (3.12).

To prove (ii), we look at what happens to the subintervals present at step $q_k - 1$. We claim that each such subinterval can contain no more than $\dfrac{\Delta_1}{q_k}$ short subintervals at step Δ_1. This is because the first subinterval to get filled is the one which s_k occupies, and it gets filled by $s_k, 2s_k, 3s_k, \ldots, js_k$. But the step corresponding to js_k is jq_k so $jq_k \leq \Delta_1$, whence the bound. ■

We continue the proof of Theorem 1.4. By Dirichlet's theorem there exists $1 \leq q \leq \Delta_1$ such that $\{\{q \dfrac{\lambda}{B}\}\} \leq \Delta_1^{-1}$. Hence

$$|s_k| \leq \frac{B}{\Delta_1} \ . \tag{3.15}$$

Now we are ready to estimate the contribution to $Q_n(B, \Delta_1, \Delta_2)$ of the various λ, according to the behavior of their continued fraction expansion. We start from

$$Q_n(B, \Delta_1, \Delta_2) - T(B, \Delta_1, \Delta_2) = \sum_{\substack{(\lambda, B) = 1 \\ f(\lambda) \geq 2}} f(\lambda)^n \ .$$

Case 1. $\dfrac{\lambda}{B}$ has $q_k = 1$.

This case occurs when $q_k = q_0 = 1$, and $q_1 > \Delta_1$. Then

$$s_0 \equiv \lambda q_0 \equiv \lambda \pmod B \ . \tag{3.16}$$

Since

$$q_1 = \lceil \frac{B}{\lambda} \rceil > \Delta_1$$

we have $\lambda \in [0, \frac{B}{\Delta_1})$. We use the bound (3.12) to obtain

$$f(\lambda) = 1 + 2H(\lambda) \le 1 + 4[\frac{\Delta_2}{|s_0|}] \le 1 + 4[\frac{\Delta_2}{\lambda}] \ . \qquad (3.17)$$

We have also the trivial bound

$$f(\lambda) \le 1 + 2\Delta_1 \ ,$$

which we use whenever $\Delta_1 \le 2 \frac{\Delta_2}{\lambda}$. By symmetry we may suppose $\Delta_1 \le \Delta_2$. Now

$$\sum_{\substack{\text{Case } 1 \\ f(\lambda) \ge 2}} f(\lambda)^n = \sum_{\lambda=1}^{B/\Delta_1} f(\lambda)^n$$

$$\le 2[\frac{\Delta_2}{\Delta_1}](1+2\Delta_2)^n + \sum_{\lambda - \frac{\Delta_2}{\Delta_1}}^{\mathrm{MIN}(\frac{B}{\Delta_1},\, \Delta_2)} (1+4[\frac{\Delta_2}{\lambda}])^n$$

$$\le 2[\frac{\Delta_2}{\Delta_1}] \ (1+2\Delta_1)^n + \frac{1}{n-1}(1+4\Delta_1)^{n-1}$$

$$\le c_n \Delta_1^{n-1} \Delta_2 \ . \qquad (3.18)$$

Case 2. $\frac{\lambda}{B}$ *has* $q_k > 1$ *and* $|s_{k-1}| \le \Delta_2$.

Note that (3.8) yields

$$2q_k|s_{k-1}| \ge B \ , \tag{3.19}$$

which with the hypothesis gives

$$2\Delta_1\Delta_2 \ge 2q_k|s_{k-1}| \ge B$$

so this case only occurs when $2\Delta_1\Delta_2 \ge B$. Now

$$\frac{\Delta_1}{q_k} \ \frac{\Delta_2}{|s_{k-1}|} \ge \frac{\Delta_1}{q_k} \ge 1 \ .$$

Hence using (3.11), and the claim,

$$\begin{aligned} f(\lambda) &= 1+2H(\lambda) = 1+2(S(\lambda)+L(\lambda)) \\ &\le 3+6\ \frac{\Delta_1\Delta_2}{q_k|s_{k+1}|} \le 9\ \frac{\Delta_1\Delta_2}{q_k|s_{k-1}|}\ . \end{aligned} \tag{3.20}$$

Using (3.19) gives

$$f(\lambda) \le 18\ \frac{\Delta_1\Delta_2}{B}\ .$$

Hence

$$\begin{aligned} \sum_{\substack{\text{Case 2} \\ f(\lambda)\ge 2}} f(\lambda)^n &\le \sum_{(\lambda,\ B)\ =\ 1} (18\ \frac{\Delta_1\Delta_2}{B})^n \\ &\le 18^n\ \phi(B)\ (\frac{\Delta_1\Delta_2}{B})^n\ . \end{aligned} \tag{3.21}$$

Case 3. $\frac{\lambda}{B}$ *has* $q_k > 1$ *and* $|s_{k-1}| > \Delta_2$.

Now we use the bounds (3.12):

$$H(\lambda) \le 2[\frac{\Delta_2}{|s_k|}] \ .$$

Next, since $|s_{k-1}| > \Delta_2$ the only values of x_2 satisfying (3.10) are multiples of $|s_k|$ and $|s_{k-1}| - j|s_k|$. The number of such multiples is at most $[\frac{\Delta_1}{q_k}]$ so that

$$H(\lambda) \le 2[\frac{\Delta_1}{q_k}] \ .$$

Combining these two inequalities gives

$$H(\lambda) \le 2 \ \mathrm{MIN} \ (\frac{\Delta_1}{q_k}, \ \frac{\Delta_2}{|s_k|}) \ . \tag{3.22}$$

Now we count the contribution over all pairs $(q_k, |s_k|)$. A given pair $(q_k, |s_k|)$ can occur with at most $2(q_k, B)$ values of λ, using (3.6), and with no values of λ unless $(q_k, B) = (|s_k|, B)$. We have

$$\sum_{\substack{\text{Case } 3 \\ f(\lambda) \ge 2}} f(\lambda)^n \le \sum_{d|B} d \sum_{\substack{q_k \le \Delta_1 \\ |s_k| \le \Delta_2 \\ (q_k, B) = (|s_k|, B) = d}} [1 + 4 \ \mathrm{MIN} \ (\frac{\Delta_1}{q_k}, \ \frac{\Delta_2}{|s_k|}]^n \tag{3.23}$$

We break this sum into two pieces (I) and (II), according as

(I) $\quad \dfrac{\Delta_1}{q_k} < \dfrac{\Delta_2}{|s_k|}$,

(II) $\quad \dfrac{\Delta_1}{q_k} \ge \dfrac{\Delta_2}{|s_k|}$.

In case (I), we have $|s_k| < q_k \dfrac{\Delta_2}{\Delta_1}$. This restricts the range of summation of $|s_k|$ in (3.23). Set $q_k = dq_k^*$, $|s_k| = ds_k^*$. Then $|s_k^*| < q_k^* \dfrac{\Delta^2}{\Delta_1}$ so that

$$\sum_{(I)} \le \sum_{d|B} d \sum_{q_k^* = 1}^{\frac{\Delta_1}{d}} (q_k^* \frac{\Delta_2}{\Delta_1}) \, [1 + 4 \frac{\Delta_1}{dq_k^*}]^n$$

$$\le \sum_{d|B} d \sum_{q_k^* = 1}^{\frac{\Delta_1}{d}} 5^n d^{-n} q_k^* \Delta_1^{n-1} \Delta_2 \qquad (3.24)$$

If $n \ge 3$ this implies

$$\sum_{(I)} \le (\sum_{d|B} d^{-(n-1)}) \zeta(n-1) 5^n \Delta_1^{n-1} \Delta_2$$

$$\le c_{n,\,1} \Delta_1^{n-1} \Delta_2 \ , \qquad (3.28)$$

where $c_{n,\,1} = \zeta(n-1)^2 5^n$. If $n = 2$ we obtain

$$\sum_{(I)} \le (\sum_{\substack{d|B \\ d \le \Delta_1}} d^{-1} \log (\frac{\Delta_1}{d})) 5^2 \Delta_1 \Delta_2 \ ,$$

$$\le 5^2 \Delta_1 \Delta_2 (\log \Delta_1) (\sum_{d|B} d^{-1}) \qquad (3.26)$$

In case (II) we have

$$|s_k| \ge q_k \frac{\Delta_2}{\Delta_1} \ .$$

Setting $q_k = q_k^* d$, $s_k = s_k^* d$, we have

$$\sum_{(II)} \le \sum_{d|B} d \sum_{q_k^* = 1}^{\frac{\Delta_1}{d}} \{ \sum_{|s_k^*| = q_k^* \frac{\Delta_1}{\Delta_2}}^{\frac{\Delta_2}{d}} [1+4 \frac{\Delta_2}{d|s_k^*|}]^n \} \ . \qquad (3.27)$$

We bound the inner sum by

$$\sum_{|s_k^*| = q_k^* \frac{\Delta_1}{\Delta_2}}^{\frac{\Delta_2}{d}} [1 + 4 \frac{\Delta_2}{d|s_k^*|}]^n \le \sum_{|s_k^*| = q_k^* \frac{\Delta_1}{\Delta_2}}^{\infty} 5^n \Delta_2^n d^{-n} (|s_k^*|)^{-n}$$

$$\le 2 \cdot \frac{5^n}{n-1} d^{-n} (q_k^*)^{-n} \Delta_1^{n-1} \Delta_2 .$$

Inserting this in (3.27) yields

$$\sum_{(II)} \le (\sum_{d|B} d^{-(n-1)}) \zeta(n) 2 \cdot \frac{5^n}{n-1} \Delta_1^{n-1} \Delta_2 \tag{3.28}$$

Hence if $n \ge 3$, we obtain

$$\sum_{(II)} \le c_{n,\,2} \Delta_1^{n-1} \Delta_2 \tag{3.29}$$

with $c_{n,\,2} = 2\zeta(n)\zeta(n-1) \frac{5^n}{n-1}$. If $n = 2$, we obtain

$$\sum_{(II)} \le c_{n,\,2}^* \Delta_1^{n-1} \Delta_2 (\log B) \tag{3.30}$$

where $c_{n,\,2^*} = 2\zeta(n) \frac{5^n}{n-1}$.

Combining these contributions yields Theorem 1.4. ■

Remarks on the proof of Theorem 1.4.

(1) The proof showed that we can take $c_n = 5^{n+1}$ and $c_n^* = (18)^n$. I believe that the sharpest possible value for c_n is $c_n \sim 2^n$ and that $c_n^* \ge 4^n$.

(2) The proof placed no restrictions on the size of Δ_1 and Δ_2, aside from $\Delta_1 \le \Delta_2$. In particular either of Δ_1 and Δ_2 may be $\ge B$.

4. Bounding the second moment of $N(\boldsymbol{\alpha}, \Delta)$.

Proof of Theorem 1.3. Let

$$D_n(B, \Delta) = \frac{1}{2} \sum_{\boldsymbol{\alpha} \in S_n(B)} [N(\boldsymbol{\alpha}, \Delta)^2 - N(\boldsymbol{\alpha}, \Delta)] \tag{4.1}$$

Now $D_n(B, \Delta)$ is exactly the number of solutions $(\frac{a_1}{B}, \ldots, \frac{a_n}{B}, x_1, x_2)$ to the system of inequalities:

$$|\frac{a_i}{B} - \frac{b_{1,i}}{x_1}| \le \frac{\Delta}{Bx_1} , \tag{4.2a}$$

$$|\frac{a_i}{B} - \frac{b_{2,i}}{y_2}| \le \frac{\Delta}{Bx_2} , \tag{4.2b}$$

$$1 \le x_1 < x_2 < B , \tag{4.2c}$$

$$0 \le a_i \le B-1 , \tag{4.2d}$$

$$g.c.d\ (a_1, \ldots, a_n, B) = 1 . \tag{4.3}$$

We are going to show that for $n \ge 4$ that

$$D_n(B,\Delta) \le c_{n,1}^* B^2 (\frac{\Delta^2}{B})^n + c_{n,2}^* B\Delta^n , \tag{4.4a}$$

for $n = 3$ that

$$D_3(B, \Delta) \le c_{3,1}^* B^2 (\frac{\Delta^2}{B})^3 + c_{3,2}^* B\Delta^3 (\sum_{d|B} d^{-\frac{1}{2}}) \tag{4.4b}$$

and for $n = 2$ that

$$D_2(B, \Delta) \le c_{2,1}^* B^2 (\frac{\Delta^2}{B})^2 + c_{2,2}^* B\Delta^2 d(B)^3 . \tag{4.4c}$$

Assuming these inequalities are proved, Theorem 1.3 follows immediately by observing that

$$\sum_{\boldsymbol{\alpha} \in S_n(B)} N(\boldsymbol{\alpha},\ \Delta)^2 \ = \ D_n(B,\ \Delta) \ + \sum_{\boldsymbol{\alpha} \in S_n(B)} N(\boldsymbol{\alpha},\ \Delta)$$

and using Theorem 1.1.

The conditions (4.2) are equivalent to the conditions

$$\begin{aligned} a_i x_1 &\equiv k_{1,\ i} \quad (\mathrm{mod}\ B) \\ a_i x_2 &\equiv k_{2,\ i} \quad (\mathrm{mod}\ B) \end{aligned} \tag{4.5a}$$

subject to

$$\begin{aligned} |k_{1,\ i}| &\le \Delta \\ |k_{2,\ i}| &\le \Delta \end{aligned} \tag{4.5b}$$

for $1 \le i \le h$. The relative primality condition $g.c.d.\ (a_1, \ldots,\ a_n,\ B) \ = \ 1$ implies that at least one of the $k_{1,\ i}$ and one of the $k_{2,\ i}$ is nonzero.

Now view x_1 and x_2 as fixed, and let $H(x_1,\ x_2;\ B,\ \Delta)$ denote the number of solutions $(a,\ k_1,\ k_2)$ of

$$\begin{aligned} a x_1 &\equiv k_1 \quad (\mathrm{mod}\ B) \\ a x_2 &\equiv k_2 \quad (\mathrm{mod}\ B)\ , \end{aligned} \tag{4.6a}$$

subject to

$$|k_1| \le \Delta, \quad |k| \le \Delta\ . \tag{4.6b}$$

Let $g.c.d.\ (x_1,\ x_2,\ B) \ = \ d$. There are d solutions to (4.6) having $k_1 \ = \ k_2 \ = \ 0$. All of the n variables a_i satisfy (4.6a), and at least one does not have $k_1 \ = \ k_2 \ = \ 0$ by the relative primality condition (4.3), hence

$$D_n(B, \Delta) = \sum_{\substack{1 \le x_1 < x_2 < B \\ d = g.c.d.\ (x_1,\ x_2,\ B)}} H(x_1, x_2; B, \Delta)^{n-1} [H(x_1, x_2, B, \Delta) - d] \quad (4.7)$$

Our general approach is to count solutions to (4.6) by eliminating a from these congruences. To see the idea, suppose

$$(x_1, B) = (x_2, B) = 1 , \quad (4.8)$$

and set

$$\lambda \equiv x_1^{-1} x_2 \quad (\mathrm{mod}\ B) . \quad (4.9)$$

Then (4.6) implies that

$$\lambda k_1 \equiv k_2 \quad (\mathrm{mod}\ B) ,$$
$$|k_1| \le \Delta, \quad |k_2| \le \Delta . \quad (4.10)$$

Conversely, each solution to (4.10) with (x_1, x_2) fixed gives rise to a unique solution of (4.6), using the relations

$$a_i \equiv k_{1,\ i} x_1^{-1} \quad (\mathrm{mod}\ B) .$$

Note that $x_1 \ne x_2$ is equivalent to $\lambda \not\equiv 1 \quad (\mathrm{mod}\ B)$. In this way, we have

$$H(x_1, x_2; B, \Delta) = f(\lambda; B, \Delta, \Delta) \quad (4.11)$$

defined by (3.1) and (3.2). Furthermore a given λ arises from exactly $\phi(B)$ different pairs (x_1, x_2) with $(x_1, B) = (x_2, B) = 1$. Hence

$$S_{1,\,1}(n) \;=\; \sum_{\substack{1 \le x_1 < x_2 \le B \\ (x_1 x_2,\, B) \,=\, 1}} H(x_{1,} x_2;\, B,\, \Delta)^n \;=\; \frac{1}{2}\;\phi(B) \sum_{\substack{\lambda=2 \\ (\lambda,\, B) \,=\, 1}}^{B} f(\lambda;\, B,\, \Delta,\, \Delta)^n$$

$$= \phi(B)(\frac{1}{2}\; Q_n(B;\, \Delta,\, \Delta) \,-\, f(1,\, B,\, \Delta,\, \Delta)^n)\;. \qquad (4.12)$$

We can in general estimate the whole sum $D_n(B,\ \Delta)$ in terms of sums of this kind. Let

$$S_{d_1,\, d_2}(n) \;=\; \sum_{\substack{1 \le x_1 < x_2 < B \\ (x_1,\, B) \,=\, d_1 \\ (x_2,\, B) \,=\, d_2}} H(x_1,\, x_2;\, B,\, \Delta)^n\;. \qquad (4.13)$$

Let

$$(d_1,\, d_2) \;=\; d$$

so that

$$[d_1,\, d_2] \;=\; \frac{d_1 d_2}{d}\;. \qquad (4.14)$$

We will show that

$$S_{d_1,\, d_2}(n) \;\le\; \frac{1}{2}\;\phi(\frac{B}{[d_1,\, d_2]})\, d_1^* d_2^* d^n Q_n(\frac{B}{[d_1,\, d_2]},\, \frac{\Delta}{d_1},\, \frac{\Delta}{d_2}) \qquad (4.15)$$

if $d_1 \ne d_2$, and

$$S_{d_1,\, d_1}(n) \;\le\; \frac{1}{2}\;\phi(\frac{B}{d_1})\, d_1^n Q_n(\frac{B}{d_1},\, \Delta_{d_1},\, \frac{\Delta}{d_1})$$

$$-\; \phi(\frac{B}{d_1})\, d_1^n f(1;\, \frac{B}{d_1},\, \frac{\Delta}{d_1},\, \frac{\Delta}{d_1})^n \qquad (4.16)$$

Indeed, let x_1^* and x_2^* be defined by

$$x_1 = x_1^* d_1 \ , \ (x_1^*, \frac{B}{d_1}) = 1 \ ,$$
$$x_2 = x_2^* d_2 \ , \ (x_2^*, \frac{B}{d_2}) = 1 \ . \tag{4.17}$$

Then (4.6a) implies that

$$k_1 = k_1^* d_1 \ ,$$
$$k_2 = k_2^* d_2 \ .$$

We now define d_1^* and, d_2^* by

$$d_1 = dd_1^*$$
$$d_2 = dd_2^* \tag{4.12}$$

where $d = (d_1, d_2)$ so that $(d_1^*, d_2^*) = 1$. Then (4.6a) implies that

$$ax_1^* \equiv k_1^* \pmod{\frac{B}{d_1}}$$
$$ax_2^* \equiv k_2^* \pmod{\frac{B}{d_2}} \tag{4.19a}$$

and

$$|k_1^*| < \frac{\Delta}{d_1} \ , \quad |k_2^*| < \frac{\Delta}{d_2} \tag{4.19b}$$

Now (4.19a) implies that

$$ax_1^* \equiv k_1^* \pmod{\frac{B}{[d_1, d_2]}}$$
$$ax_2^* \equiv k_2^* \pmod{\frac{B}{[d_1, d_2]}} \tag{4.20}$$

and $(x_1^* x_2^*, \frac{B}{[d_1, d_2]}) = 1$. Hence

$$\lambda k_1^* \equiv k_2^* \quad (\bmod \frac{B}{[d_1, d_2]}) \tag{4.21a}$$

subject to

$$|k_1^*| \le \frac{\Delta}{d_1}, \quad |k_2^*| \le \frac{\Delta}{d_2} \tag{4.21b}$$

where

$$\lambda \equiv (x_1^*)^{-1}(x_2^*) \quad (\bmod \frac{B}{[d_1, d_2]}) \ . \tag{4.22}$$

has $(\lambda, \frac{B}{[d_1, d_2]}) = 1$. Thus from each solution to (4.6) we derive a solution to (4.22). To reverse this, we must give an upper bound for how many solutions to (4.6) give rise to the same solution of (4.22). Now the solutions to (4.21), because of the bound (4.21b) determine k_1^* and k_2^* in (4.19a) uniquely. (Note that $\frac{\Delta}{d_1}, \frac{\Delta}{d_2}$ may be $\ge \frac{B}{[d_1, d_2]}$ in (4.21).) Then the equations (4.19a) determine a (mod $\frac{B}{(d_1, d_2)}$) so there are at most d choices for a.

Now (4.21) has by definition $f(\lambda;\ B,\ \frac{\Delta_1}{d_1},\ \frac{\Delta_2}{d_2})$ solutions. We obtain the bound

$$H(x_1, x_2; B, \Delta) \le df(\lambda; \frac{B}{[d_1, d_2]}, \frac{\Delta_1}{d_1}, \frac{\Delta_2}{d_2}) \ . \tag{4.23}$$

The number of pairs (x_1, x_2) with $(x_1, B) = d_1, (x_2\ B) = d_2$, giving rise to the same λ in (4.22) is at most

$$\phi(\frac{B}{[d_1, d_2]})\ d_1^* d_2^* \ . \tag{4.24}$$

Consequently (4.13) yields

$$S_{d_1,\,d_2}(n) \le \frac{1}{2}\ \phi(\frac{B}{[d_1,\,d_2]})\,d_1^* d_2^* d^n Q_n(\frac{B}{[d_1,\,d_2]},\ \frac{\Delta_1}{d_1},\ \frac{\Delta_2}{d_2})\ . \qquad (4.25)$$

which is (4.15). If $d_1 = d_2 = d$, then we may drop the $\lambda = 1$ term coming from the cases $x_1 = x_2$, and we obtain (4.16).

Now (4.7) gives

$$D_n(B,\,\Delta) \le \sum_{\substack{d_1 | B \\ d_2 | B \\ d = (d_1,\,d_2)}} [S_{d_1,\,d_2}(n) - dS_{d_1,\,d_2}(n-1)]\ . \qquad (4.26)$$

Hence (4.25) yields

$$D_n(B,\,\Delta) \le \sum_{\substack{d_1 | B \\ d_2 | B \\ d = (d_1,\,d_2)}} \frac{1}{2}\ \phi[\frac{B}{[d_1,\,d_2]})\,d_1^* d_2^* d^n [Q_n(\frac{B}{[d_1,\,d_2]},\ \frac{\Delta}{d_1},\ \frac{\Delta}{d_2}) - Q_{n-1}(\frac{B}{[d_1,\,d_2]},\ \Delta_{d_1},\ \frac{\Delta}{d_2})]\ .$$

Now we apply Theorem 1.4. We have for $n \ge 3$ and $d_1 \le d_2$ that

$$\begin{aligned} &Q_n(\frac{B}{[d_1,\,d_2]},\ \frac{\Delta}{d_1},\ \frac{\Delta}{d_2}) - Q_{n-1}(\frac{B}{[d_1,\,d_2]},\ \frac{\Delta}{d_1},\ \frac{\Delta}{d_2}) \\ &\qquad \le c_n^* \phi(\frac{B}{[d_1,\,d_2]})(\frac{\Delta^2}{Bd_1 d_2})^n + c_n(\frac{\Delta}{d_2})^{n-1}\frac{\Delta}{d_1}\ . \end{aligned} \qquad (4.28)$$

Substituting this inequality in (4.27) yields

$$\begin{aligned} D_n(B,\,\Delta) \le &\sum_{\substack{d_1 | Bd_2 | B \\ d_1 \le d_2}} c_n^* \phi(\frac{B}{[d_1,\,d_2]})^2 d_1^* d_2^* d^n (\frac{\Delta^2}{Bd_1 d_2})^n \\ &+ \sum_{\substack{d_1 | B,\ d_2 | B \\ d_1 \le d_2}} c_n \phi(\frac{B}{[d_1,\,d_2]})\,d_1^* d_2^* d_2^{-(n-1)}\,d_1^{-1}\,d^n \Delta^n \end{aligned} \qquad (4.29)$$

We bound the two terms $T_1(n)$ and $T_2(n)$ in this sum separately. We simplify the first term using

$$\phi\left(\frac{B}{[d_1, d_2]}\right) \le \frac{B}{[d_1, d_2]}) \tag{4.30}$$

to obtain

$$\begin{aligned} T_1(n) &= c_n^* B \sum_{d_1 \le d_2} \phi\left(\frac{B}{[d_1, d_2]}\right) d^{n-1}\left(\frac{\Delta^2}{Bd_1 d_2}\right)^n \\ &\le c^{n^*} B^2 \left(\frac{\Delta^2}{B}\right)^n \sum_{\substack{d_1 \le d_2 \\ d_1 | B \\ d_2 | B}} d^{-1}[d_1, d_2]^{-n-1} \ . \end{aligned} \tag{4.31}$$

The inner sum in (4.31) is bounded, since

$$\begin{aligned} \sum_{\substack{d_1 \le d_2 \\ d_1 | B \\ d_2 | B}} d^{-1}[d_1, d_2]^{-n-1} &\le \sum_{d | B} d^{-(n+2)} \sum_{d_1^*, d_2^*} (d_1^* d_2^*)^{-(n+1)} \\ &\le \zeta(n+1)^2 \zeta(n+2) \ . \end{aligned} \tag{4.32}$$

Hence we obtain

$$T_1(n) \le c_n^* \zeta(n+1)^2 \zeta(n+2) B^2 \left(\frac{\Delta^2}{B}\right)^n \ . \tag{4.33}$$

Note that this bound holds also for $n = 2$. For the second summation in (4.29), we have for $n \ge 3$

$$T_2(n) \le c_n \Delta^n \sum_{\substack{d_1 | B \\ d_2 | B \\ d_1 \le d_2}} \phi\left(\frac{B}{[d_1, d_2]}\right)(d_2^*)^{-(n-2)} \tag{4.34}$$

where $d_2^* = \dfrac{d_2}{(d_1, d_2)}$. We simplify this expression by summing over $d^* = [d_1, d_2]$

to obtain

$$\sum_{\substack{d_1 | B \\ d_2 | B \\ d_1 \le d_2}} \phi\left(\frac{B}{[d_1, d_2]}\right)(d_2^*)^{-(n-2)} = \sum_{d^* | B} \phi\left(\frac{B}{d^*}\right) \Omega_{n-2}^*(d^*) \tag{4.35}$$

where

$$\Omega_n^*(d^*) = \sum_{\substack{d_1, d_2 \\ [d_1, d_2] = d^* \\ d_1 \le d_2}} \left(\frac{d_2}{(d_1, d_2)}\right)^{-n} \tag{4.36}$$

We are going to show that for $n \ge 4$

$$T_2 (n) \le c_n^* B \Delta^n \tag{4.37}$$

and that for $n = 3$

$$T_2 (3) \le c_3^* B \Delta^n \left(\sum_{d | B} \frac{1}{d}\right) . \tag{4.38}$$

To prove (4.37) it suffices to show that for $n \ge 2$ there is a bound C_0 such that

$$\Omega_n^*(d^*) \le C_0 \tag{4.29}$$

for all d^*. If this is proved, then (4.34), (4.35) and (4.39) give

$$T_2 (n) \le C c_n \Delta^n \sum_{d^* | B} \phi\left(\frac{B}{d}\right) = C c_n B \Delta^n$$

which is (4.37).

To show the functions $\Omega_n^*(d_1^*)$ are bounded above for all $n \ge 2$, it suffices to prove the result for $n = 2$, since $\Omega_2(d^*) \ge \Omega_n(d^*)$ for all $n \ge 2$ by (4.36). We rewrite (4.36) using $d^* = [d_1, d_2] = e^2 d_1^* d_2^*, d_1 = e d_1^*, d_2 = d e_2^*$ to get

$$\Omega_n^*(d^*) \; = \sum_{e^2|d^*} \; \sum_{\substack{d_1,\, d_2 \\ (d_1,\, d_2) \,=\, e \\ [d_1,\, d_2] \,=\, d^* \\ d_1 \,\le\, d_2}} (\frac{d_2}{e})^{-n} \tag{4.40}$$

Now since $d_1^* \le d_2^*$ we must have

$$d^* \; = \; e^2 d_1^* d_2^* \le e^2 (d_2^*)^2 \quad \text{so that}$$

$$\frac{d_2}{e} \; = \; d_2^* \ge \frac{\sqrt{d^*}}{e} \; .$$

Applying this in (4.40) gives

$$\Omega_n^*(d^*) \le \sum_{e^2|d^*} d(\frac{d^*}{e^2}) \left[\frac{e^2}{d^*} \right]^{1/2} \tag{4.41}$$

which for $n \; = \; 2$ is

$$\Omega_2^*(d^*) \le \sum_{e^2|d^*} d(\frac{d^*}{e^2})(\frac{e^2}{d^*}) \tag{4.42}$$

To bound the right side of (4.42), we use the submultiplicativity of the divisor function, i.e. that

$$d(m_1 m_2) \le d(m_1) d(m_2) \tag{4.43}$$

for *all* $m_1 ,\; m_2$. We set $d^* \; = \; f^2 d^{**}$ where d^{**} is squarefree, and using (4.43) get

$$\begin{aligned} \sum_{e^2|d^*} d(\frac{d^*}{e^2}) \frac{e^2}{d^*} &\le \frac{d(d^{**})}{d^{**}} \sum_{h|f} d((\frac{f}{h})^2)(\frac{f}{h})^{-2} \\ &\le \sum_{m=1}^{\infty} d(m^2) m^{-2} \end{aligned} \tag{4.44}$$

using $\dfrac{d(m)}{m} \le 1$ for all $m \ge 1$. Then (4.42) and (4.44) give

$$\Omega_2^*(d^*) \le \sum_{m=1}^{\infty} d(m^2)m^{-2} = \prod_p (1 + \frac{3}{p^2} + \frac{5}{p^4} + \ldots) = C_0 .$$

which is the desired bound (4.39).

To prove (4.38) we proceed to establish the inequality

$$\Omega_1^*(d^*) \le C_1 (\sum_{d|d^*} d^{-1/2}) . \tag{4.45}$$

Then (4.34), (4.35) and (4.45) imply for $n = 3$ that

$$T_2\ (3) \le C_1 c_n B \Delta^3 (\sum_{d|B} d_1^{-1/2}) ,$$

which is (4.38).

To prove (4.38), we use (4.41) for $n = 1$, which gives

$$\Omega_1^*(d^*) \le \sum_{e^2|d^*} d(\frac{d^*}{e^2})(\frac{e^2}{d^*})^{1/2}$$

$$\le \sum_{d|d^*} (\frac{d}{d^*})^{1/2} = \sum_{d|d^*} d^{-1/2} .$$

Finally we treat the case $n = 2$. In this case we apply Theorem 1.4 for $n = 2$, and (4.29) is replaced by

$$D_2(B, \Delta) \le \sum_{\substack{d_1|B \\ d_2|B \\ d_1 \le 2}} c_2^* \phi(\frac{B}{[d_1, d_2]})^2 d_1^* d_2^* d^2 (\frac{\Delta^2}{Bd_1 d_2})^2$$

$$+ \sum_{\substack{d_1|B \\ d_2|B \\ d_1 \le d_2}} c_2 \phi(\frac{B}{[d_1, d_2]}) d_1^* d_2^* d_2^{-1} d_1^{-1} d^2 \Delta^2 (\sum_{f|B} \frac{1}{f}) , \tag{4.46}$$

where $d = (d_1, d_2)$. The first of the sums on the right side of (4.46) is bounded by

(4.33). The second term becomes, letting $d^* = [d_1, d_2]$,

$$T_2(2) \le c_2 \Delta^2 \Big(\sum_{f|B} \frac{1}{f} \Big) \sum_{d^*|B} \Big\{ \phi\Big(\frac{B}{d^*}\Big) \Big(\sum_{\substack{d_1, d_2 \\ [d_1, d_2] = d^* \\ d_1 \le d_2}} 1 \Big) \Big\} \tag{4.47}$$

We use the crude bound

$$\sum_{d^*|B} \Big\{ \phi\Big(\frac{B}{d^*}\Big) \Big(\sum_{\substack{d_1, d_2 \\ [d_1, d_2] = d^* \\ d_1 \le d_2}} 1 \Big) \Big\} \le \sum_{d^*|B} \phi\Big(\frac{B}{d^*}\Big)^2$$

$$\le B d(B)^2$$

to obtain in (4.48)

$$T_2(2) \le c_2 B d(B)^3 \Delta^2 . \tag{4.49}$$

Combining (4.46), (4.31) and (4.49) proves (4.4c). □

Now we show that when B is prime,

$$\sum_{\alpha \in S_n(B)} N(\alpha, \Delta)^2 \gg_n B^2 \Big(\frac{\Delta^2}{B}\Big)^n + B\Delta^n . \tag{4.50}$$

We have

$$\sum_{\alpha \in S_n(B)} N(\alpha, \Delta)^2 \ge D_n(B, \Delta)$$

$$\ge S_{1,1}(n) - S_{1,1}(n-1)$$

using (4.1), (4.7) and (4.12), since $(x_1, x_2, B) = 1$ always holds when B is prime. When B is prime (4.12) gives

$$S_{1,1}(n) - S_{1,1}(n-1) \ \geq\ \frac{1}{2}(B-1)[Q_n(B;\ \Delta,\ \Delta) - f(1;\ B,\ \Delta,\ \Delta)^n]$$

$$+\ O(BQ_{n-1}(B;\ \Delta,\ \Delta))\ . \tag{4.52}$$

Now for prime *B*,

$$\begin{aligned} Q_n(B;\ \Delta,\ \Delta) - f(1,\ B,\ \Delta,\ \Delta)^n &\geq f(2;\ B,\ \Delta,\ \Delta)^n \\ &\geq 2^{-n}\Delta^n \end{aligned} \tag{4.53}$$

by direct calculation, when *B* is odd. Also (3.4) implies that

$$Q_n(B;\ \Delta,\ \Delta) \underset{n}{\gg} B\left(\frac{\Delta^2}{B}\right)^n\ . \tag{4.54}$$

Since $f(1,\ B,\ \Delta,\ \Delta)\ \leq\ 2\Delta$, combining (4.53) and (4.54) gives

$$\begin{aligned} Q_n(B;\ \Delta,\ \Delta) - f(1,\ B,\ \Delta,\ \Delta)^n &\underset{n}{\gg} \mathrm{MAX}\left(\Delta^n,\ B\left(\frac{\Delta^2}{B}\right)^n\right) \\ &\underset{n}{\gg} \Delta^n\ +\ B\left(\frac{\Delta^2}{B}\right)^n\ . \end{aligned} \tag{4.55}$$

Now (4.51), (4.52) and (4.55) imply (4.50) as required.

5. Bounding the tail of a conditional probability distribution

Proof of Theorem 1.5. We study the conditional probabilities

$$p_k(B,\ \Delta,\ n)\ =\ \text{Prob}\ \{\boldsymbol{\alpha} \in S_n(B)\ \text{ has }\ N(\boldsymbol{\alpha},\ \Delta)\ \geq\ k \mid N(\boldsymbol{\alpha},\ \Delta)\ \geq\ 1\}\ .$$

We use the following elementary lemma.

Lemma 5.1. Let η *be a real-valued random variable with finite mean and variance, and suppose that there is a positive constant* α *such that*

$$E[\eta^2] \le \alpha |E[\eta]| \ . \tag{5.1}$$

Then

$$\text{Prob } [|\eta| \ge k] \le (\frac{\alpha}{k})^2 \ . \tag{5.2}$$

Proof. Now

$$\alpha |E[\eta]| \ge E[\eta^2] \ge (E[\eta])^2$$

so that $0|E[\eta]| \le \alpha$. Hence

$$k^2 \text{ Prob } [|\eta| \ge k] \le E[\eta^2] \le \alpha E[\eta] \le \alpha^2 \ . \ ■$$

We apply Lemma 5.1 to the discrete random variable $\eta = N(\boldsymbol{\alpha},\Delta)$ restricted to the subset

$$S_n^*(B) = \{\boldsymbol{\alpha} \in S_n(B): \ N(\boldsymbol{\alpha}, \Delta) \ge 1\}$$

of $S_n(B)$, with the uniform probability measure on $S_n^*(B)$. Let Prob^* denote this probability measure, and observe that

$$p_k(B, \Delta, n) = \text{Prob}^*\{\eta \ge k\} \ . \tag{5.3}$$

Now we have

$$E[\eta] = |S_n^*(A)|^{-1} (\sum_{\boldsymbol{\alpha} \in S_n(B)} N(\boldsymbol{\alpha}, \Delta))$$

$$E[\eta^2] = |S_n^*(A)|^{-1} \ (\sum_{\boldsymbol{\alpha} \in S_n(B)} N(\boldsymbol{\alpha}, \Delta)^2) \ .$$

so that

$$\frac{E[\eta^2]}{|E[\eta]|} \le \frac{\sum_{\boldsymbol{\alpha}\in S_n(B)} N(\boldsymbol{\alpha}, \Delta)^2}{\sum_{\boldsymbol{\alpha}\in S_n(B)} N(\boldsymbol{\alpha}, \Delta)} . \tag{5.4}$$

Now for $n \ge 2$ and $\Delta \ge c_{n,\,1} d(B)^2$ for a sufficiently large positive $c_{n,\,1}$, Theorem 1.1 gives

$$\sum_{\boldsymbol{\alpha}\in S_n(B)} N(\boldsymbol{\alpha}, \Delta) \ge c_{n,\,2} B\Delta^n \tag{5.5}$$

for some positive constant $c_{n,\,2}$. Next for $n \ge 4$ and $c_{n,\,1} d(B)^2 \le \Delta \le B^{1 - \frac{1}{n}}$ Theorem 1.3 gives

$$\sum_{\boldsymbol{\alpha}\in S_n(B)} N(\boldsymbol{\alpha}, \Delta)^2 \le c_{n,\,3} B\Delta^n . \tag{5.6}$$

for some positive constant $c_{n,\,3}$. Also for $n = 2$ or 3 and B prime, for $c_{n,\,1} d(B)^2 \le \Delta \le B^{1 - \frac{1}{n}}$ Theorem 1.3 gives (5.6) also. Now applying (5.6) and (5.5) in (5.4) gives

$$\frac{E[\eta^2]}{E[\eta]} \le \frac{c_{n,\,3}}{c_{n,\,2}} .$$

The theorem follows from Lemma 5.1 on choosing $\alpha = \frac{c_{n,\,3}}{c_{n,\,2}}$ and $c_n^{**} = \mathrm{MAX}\ (c_{n,\,1}, \alpha^2)$. ■

References

[1] J. H. Halton, The distribution of the sequence $\{n\xi\}$ ($n = 0, 1, 2, \ldots$), Proc. Cambridge Phil. Soc. *61* (1965) 665-670.

[2] J. C. Lagarias, Best simultaneous Diophantine approximations II. Behavior of consecutive best approximations, Pacific J. Math. *102* (1982), 61-88.

[3] J. C. Lagarias, The Computational Complexity of Simultaneous Diophantine Approximation Problems, SIAM J. Computing, to appear. (Preliminary version in: Proc. 23rd Annual IEEE Conference as the Foundations of Computer Science (1982), 32-39).

[4] J. C. Lagarias, Knapsack-type public key cryptosystems and Diophantine approximation, (Extended Abstract), in: *Advances in Cryptology* (D. Chaum, Ed.), Plenum Publ. Co., 1984, 3-24.

[5] J. C. Lagarias, Performance analysis of Shamir's attack on the basic Merkle-Hellman knapsack public key cryptosystem, (Extended Abstract), in: *Automata, Languages and Programming*, Springer Lecture Notes in Computer Science No. 172, 1984, 312-323.

[6] S. Lang, Introduction to Diophantine Approximations, Addison-Wesley Publ. Co.: Reading, Massachusetts 1966.

[7] A. Shamir, A polynomial-time algorithm for breaking the basic Merkle-Hellman cryptosystem, Proc. 23rd Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society Press 1982, 145-152.

[8] N. Slater, Gaps and steps for the sequence $n\theta$ mod 1, Proc. Cambridge Phil. Soc. *63* (1967), 1115-1123.

Simultaneous Diophantine Approximation of Rationals by Rationals

J. C. Lagarias

Bell Laboratories
Murray Hill, NJ 07974

ABSTRACT

Let $\boldsymbol{\alpha} = (\frac{a_1}{B}, \ldots, \frac{a_n}{B})$ be a vector of rational numbers satisfying the primitivity condition $g.c.d.\ (a_1, \ldots, a_n, B) = 1$. This paper studies the number $N(\boldsymbol{\alpha}, \Delta)$ of simultaneous Diophantine approximations to $\boldsymbol{\alpha}$ with denominators $x < B$ of a given degree of approximation measured by Δ, i.e. $N(\boldsymbol{\alpha}, \Delta)$ is the number of vectors $\boldsymbol{\xi} = (\frac{x_1}{x}, \ldots, \frac{x_n}{x})$ with $1 \le x < B$ such that $\left| \frac{a_i}{B} - \frac{x_i}{x} \right| \le \frac{\Delta}{Bx}$ for $1 \le i \le n$. It gives estimates for the first and second moments of $N(\boldsymbol{\alpha}, \Delta)$ over the ensemble $S_n(B)$ consisting of all primitive vectors $\boldsymbol{\alpha}$ in the unit n-cube having denominator B. As a consequence it shows for $n \ge 4$ that "most" vectors in $S_n(B)$ that have one "unusually good" simultaneous Diophantine approximation have a bounded number of such approximations. The paper also estimates the moments of the number of solutions to homogenous linear congruences $\lambda x_1 \equiv x_2 \pmod{B}$ with bounds $|x_1| \le \Delta_1$, $|x_2| \le \Delta_2$ on the variables, taken over the set of λ with $(\lambda, B) = 1$. These results have applications to the analysis of cryptanalytic attacks on knapsack-type public key cryptosystems.