Invent. math. 111, 171–195 (1993)

Inventiones mathematicae

The size of Selmer groups for the congruent number problem

D.R. Heath-Brown

Magdalen College, Oxford OX1 4AU, United Kingdom

Oblatum 18-I-1992 & 20-VII-1992

1 Introduction

The oldest problem in the theory of elliptic curves is to determine which positive integers D can be the common difference of a three term arithmetic progression of squares of rational numbers. Such integers D are known as congruent numbers. Equivalently one may ask which elliptic curves

$$E_D\colon\ y^2 = x^3 - D^2 x$$

have positive rank. Clearly one may, and we shall, restrict attention to square-free numbers D. At present there is no known algorithm for deciding whether or not a given integer is a congruent number. However the conjecture of Birch and Swinnerton-Dyer [1], if true, would provide such a procedure. One defines

$$L_D(s) = \prod_{p \nmid 2D} (1 - a_p p^{-s} + p^{1-2s})^{-1}, \qquad a_p = p + 1 - N_p,$$

where N_p is the number of solutions of the congruence $y^2 \equiv x^3 - D^2 x \pmod{p}$. Then $L_D(s)$ has an analytic continuation as an entire function on the complex plane. The conjecture of Birch and Swinnerton-Dyer then states, in particular, that the rank $r(D)$ of E_D is equal to the order $R(D)$ of $L_D(s)$ at $s = 1$, this being the so-called analytic rank. While we cannot at present find $R(D)$ in all cases, we can at least determine whether or not $L_D(1) = 0$, and hence, conjecturally, whether or not $r(D) = 0$. Moreover one has a functional equation for $L_D(s)$ which relates its values at s and $2-s$, via a sign change $\varepsilon_D = \pm 1$. One may deduce that $(-1)^{R(D)} = \varepsilon_D$. It would then follow from the conjecture of Birch and Swinnerton-Dyer that the rank is positive whenever $\varepsilon_D = -1$. According to the calculations of Birch and Stephens [2] one has

$$\varepsilon_D = \begin{cases} +1, & D \equiv 1, 2, 3 \pmod 8, \\ -1, & D \equiv 5, 6, 7 \pmod 8, \end{cases}$$

which would imply that D is congruent whenever $D \equiv 5$, 6 or 7 (mod 8). We know from the work of Coates and Wiles [5], Gross and Zagier [7], and Rubin

[11] that, for our curves, $r(D)=R(D)$ whenever $R(D)=0$ or 1, but little can be said when $R(D)\geqq 2$.

A straightforward approach to these questions is provided by the use of descents. We shall be concerned with the "full 2-descent", which can be done over $\mathbb{Q}$ for our curves. The process will be described in detail in the next section. However what is of interest for the present discussion is that the number of 2-descents is the order of the Selmer group $S^{(2)}$. This is a power of 2, and will be a multiple of 4, on account of the rational points of order 2 on E_D. We shall therefore write $\#S^{(2)}=2^{2+s(D)}$. The exponent $s(D)$ has sometimes been refered to as the 'Selmer rank' of the curve E_D. According to the Selmer conjecture, $s(D)$ and $r(D)$ should have the same parity. It therefore seems likely, in view of the conjecture of Birch and Swinnerton-Dyer, that $s(D)$ and $R(D)$ always have the same parity.

It should be noted that our terminology differs from that of Birch and Swinnerton-Dyer [1]. In their notation the 'number of first descents' is $\lambda+\lambda_1-2$ which is often much larger than $s(D)$. Indeed, $\lambda+\lambda_1-2$ is 'usually' of order $\log\log D$ at least, whereas $s(D)$ is 'usually' of order 1, as we shall see. For $D=3.11.59$ one may calculate that $\lambda=4$, $\lambda_1=0$ whereas $s(D)=0$. While it is known that $\lambda+\lambda_1-2$ must have the same parity as $R(D)$, see Birch and Stephens [2] or Lagrange [9], the corresponding statement for $s(D)$ has yet to be settled.

The purpose of this paper is to investigate $s(D)$ on average. We prove the following results.

Theorem 1 *For any odd integer h let*

$$S(X,h)=\{D\equiv h \pmod 8:\ 1\leqq D\leqq X,\ D \text{ square-free}\}.$$

Then

$$\sum_{D\in S(X,h)} 2^{s(D)}=3\#S(X,h)+O(X(\log X)^{-1/4}(\log\log X)^8). \tag{1}$$

Of course $\#S(X,h)$ is of order X, so that we have an asymptotic formula, with a relative saving of $O((\log X)^{-1/4}(\log\log X)^8)$. We can immediately deduce the following.

Corollary 1 *For any odd integer h we have*

$$\sum_{D\in S(X,h)} 2^{r(D)}\leqq 3\#S(X,h)+O(X(\log X)^{-1/4}(\log\log X)^8).$$

When $D\equiv 1$ or 3 (mod 8) we expect $s(D)$ to be even, so that

$$s(D)\leqq\tfrac{2}{3}(2^{s(D)}-1).$$

Similarly when $D\equiv 5$ or 7 (mod 8) we expect $s(D)$ to be odd, and

$$s(D)\leqq\tfrac{1}{3}(2^{s(D)}+1).$$

Without any assumption we can only say that

$$s(D)\leqq\tfrac{1}{2}2^{s(D)}.$$

We therefore deduce the following average bound for $s(D)$.

Corollary 2 *Assume that $s(D)$ and $R(D)$ always have the same parity. Then for any odd integer h we have*

$$\sum_{D \in S(X,h)} s(D) \leqq \tfrac{4}{3} \#S(X,h) + O(X(\log X)^{-1/4}(\log\log X)^8).$$

Hence $s(D)=0$ for at least one third of all $D \equiv 1$ or $3 \pmod 8$, and $s(D)=1$ for at least five-sixths of all $D \equiv 5$ or $7 \pmod 8$. Unconditionally we have

$$\sum_{D \in S(X,h)} s(D) \leqq \tfrac{3}{2} \#S(X,h) + O(X(\log X)^{-1/4}(\log\log X)^8).$$

We automatically deduce average bounds for $r(D)$.

Corollary 3 *Assume that $r(D)$ and $R(D)$ always have the same parity. Then for any odd integer h we have*

$$\sum_{D \in S(X,h)} r(D) \leqq \tfrac{4}{3} \#S(X,h) + O(X(\log X)^{-1/4}(\log\log X)^8).$$

Hence $r(D)=0$ for at least one third of all $D \equiv 1$ or $3 \pmod 8$, and $r(D)=1$ for at least five-sixths of all $D \equiv 5$ or $7 \pmod 8$. Unconditionally we have

$$\sum_{D \in S(X,h)} r(D) \leqq \tfrac{3}{2} \#S(X,h) + O(X(\log X)^{-1/4}(\log\log X)^8).$$

A couple of remarks should be made.

1. Our proof could be applied with only minor changes to the residue classes $D \equiv 2$ or $6 \pmod 8$. Other curves with 3 rational points of order two could be handled the same way. However a certain amount of extra work would be needed to determine whether or not the all important constant 3 on the right hand side of (1) remains the same. One might also ask whether a corresponding result can be obtained when there is only one rational point of order 2. In this case the descent must be done over a quadratic number field, but one still has good control over the 2-part of the class group.

2. Averages of the analytic rank have been estimated for a number of classes of curves. Thus for example, Brumer and Heath-Brown [3] show that, for twists of any given modular elliptic curve, $R(D)$ has average at most $\frac{3}{2}$, providing that the corresponding L-functions satisfy the Riemann Hypothesis. Corollary 3 gives the same bound unconditionally, or a better bound under a far weaker hypothesis, but holds in a more restricted setting.

It has been conjectured that the rank of an elliptic curve can be arbitrarily large, but it is not clear how frequent large ranks might be. Moreover it is unclear whether one should expect arbitrarily large ranks when one restricts attention to a family of twists of a fixed curve. For our curves the work of Gouvêa and Mazur [6] shows that

$$\#\{D \in S(X,1):\ R(D) \geqq 2\} \gg X^{1/2-\varepsilon},$$

for any $\varepsilon > 0$. Indeed, it is evident that one can in fact obtain

$$\#\{D \in S(X,1):\ R(D) \geqq 2\} \gg X^{1/2},$$

but we expect that much more is true. The corresponding problem for $s(D)$ is far more tractable however, and we prove the following estimate.

Theorem 2 *For any constant $\theta<1$ there exists $c_\theta>0$ such that*

$$\#\{D\in S(X,1):\ s(D)>c_\theta\sqrt{\log D}\}\gg_\theta X^\theta.$$

Of course it is immediate from the work of § 2 that

$$s(D)\leqq 2\omega(D)\ll\frac{\log D}{\log\log D},$$

where $\omega(D)$ is the number of prime factors of D. Thus the correct maximum order for $s(D)$ is still in some doubt.

While it is uncertain whether each fixed rank occurs for a positive proportion of elliptic curves, for the Selmer rank this seems rather likely. Here we shall prove the following rather trivial result.

Theorem 3 *Let a non-negative integer n be given. Then*

$$\#\{D\in S(X,h):\ s(D)=n\}\gg_n\frac{X}{\log X},$$

for $h=1$ *or* 3, *for n even, or $h=5$ or* 7, *for n odd.*

This is certainly not the strongest result of its type, but to achieve a lower bound of order X appears to be beyond our reach at present. Such a bound would of course allow the constant 4/3 in the corollaries to Theorem 1 to be improved.

2 Counting 2-descents

We begin by describing the familiar descent process. Various accounts of this are available in the literature, see for example Serf [12], but the arguments for removing the contribution of the torsion points, and for dismissing the 2-adic conditions, seem to justify inclusion of a full description.

We start from the well-known fact that the homomorphism

$$\theta:\ \frac{E_D(\mathbb{Q})}{2E_D(\mathbb{Q})}\to G\times G\times G\left(G=\frac{\mathbb{Q}^\times}{\mathbb{Q}^{\times 2}}\right)$$

given at the non-torsion points by

$$(x,y)\to(x,x+D,x-D)(\mathrm{mod}\,\mathbb{Q}^{\times 2})$$

is injective. The only torsion points of $E_D(\mathbb{Q})$ are the points of order two. We now observe that the coset of a non-torsion point in $E(\mathbb{Q})/\mathrm{Tors}\,(E(\mathbb{Q}))$, consisting of

$$(x,y),\qquad (x,y)+(0,0)=\left(\frac{-D^2}{x},\frac{D^2y}{x^2}\right),\qquad (x,y)+(D,0)=\left(D\frac{x+D}{x-D},\frac{-2D^2y}{(x-D)^2}\right)$$

and

$$(x, y)+(-D, 0)=\left(D\frac{D-x}{D+x}, \frac{-2D^2y}{(D+x)^2}\right),$$

contains exactly one element (x', y') for which $x'>0$ and $|x'|_2 \neq 1$. Consequently, if we restrict our points (x, y) to have $x>0$ and $|x|_2 \neq 1$, the image under the map θ will have size $2^{r(D)}$.

To analyze $\mathrm{Im}(\theta)$ we write $x=r/s$, $y=t/u$ with $(r, s)=(t, u)=1$ and where $r, s, u>0$ and r, s have opposite parities. Then

$$r(r+sD)(r-sD)u^2=t^2s^3,$$

and since $(t, u)=1$ we have $u^2|s^3$. Similarly, since $(r, s)=1$ we see that s^3 is coprime to $r(r+sD)(r-sD)$, so that $s^3|u^2$. Thus $s^3=u^2$, and $s=W^2$, $u=W^3$ for some integer W. It follows that

$$r(r+sD)(r-sD)=t^2.$$

We now write $(r, D)=D_0$ and $r=D_0r'$, whence $D_0^3|t^2$ and therefore $D_0^2|t$, since D_0 must be square-free. Thus

$$r'\left(r'+s\frac{D}{D_0}\right)\left(r'-s\frac{D}{D_0}\right)=D_0\left(\frac{t}{D_0^2}\right)^2. \tag{2}$$

Since $(r', sD/D_0)=1$ and $r'+sD/D_0$ is odd, because D and $r+s$ are odd, it follows that the three factors on the left of (2) are coprime in pairs. We may therefore write

$$D_0=D_1D_2D_3, \qquad tD_0^{-2}=XYZ,$$

and

$$r'=D_1X^2, \qquad r'+sD/D_0=D_2Y^2, \qquad r'-sD/D_0=D_3Z^2.$$

On setting $D/D_0=D_4$ we obtain the system

$$D_1X^2+D_4W^2=D_2Y^2, \qquad D_1X^2-D_4W^2=D_3Z^2. \tag{3}$$

Since $r>0$ and $D_0>0$ we automatically have $D_1, D_4>0$ and therefore we see that $D_2>0$ if the first of the equations (3) is to have non-trivial real solutions. Then, as $D=D_1D_2D_3D_4>0$ we see that D_3 must also be positive. We have therefore proved the following.

Lemma 1 *There are exactly $2^{r(D)}$ systems* (3) *with non-trivial integer solutions. Moreover there are $2^{s(D)}$ systems* (3) *which are everywhere locally solvable.*

Of course the second assertion is just the definition of $s(D)$.

Our insistence that $D_j>0$ for $j=1, \ldots, 4$ already ensures that (3) have real solutions. Moreover it is an easy exercise to show that there are p-adic solutions whenever $p\nmid 2D$. For primes $p|D_1$ it is clearly necessary and sufficient for D_4D_2 and $-D_4D_3$ to be squares modulo p. Similarly, when $p|D_4$ we require D_1D_2 and D_1D_3 to be squares modulo p. In case $p|D_2$ we write (3) as

$$D_1X^2+D_4W^2=D_2Y^2, \qquad 2D_1X^2=D_2Y^2+D_3Z^2,$$

which is solvable in $\mathbb{Q}_p$ if and only if $-D_1D_4$ and $2D_1D_3$ are squares modulo p. Finally, when $p|D_3$ the condition is that D_1D_4 and $2D_1D_2$ are squares modulo p.

Fortunately, for the prime $p=2$ no further condition is required, as the following result shows.

Lemma 2 *If the system* (3) *has solutions in* $\mathbb{R}$ *and in* $\mathbb{Q}_p$ *for every odd prime* p, *then there are also solutions in* $\mathbb{Q}_2$.

To prove this we observe that the equation

$$2D_1X^2=D_2Y^2+D_3Z^2$$

has solutions in $\mathbb{R}$ and in $\mathbb{Q}_p$ for every odd prime p, by our hypothesis about the solvability of (3). By the product formula for the Hasse norm residue symbol there are also solutions in $\mathbb{Q}_2$. We may assume that such a solution involves 2-adic integers, at least one of which is a unit. Since D_1, D_2 and D_3 are odd, we see that Y and Z are 2-adic units, and therefore we must have

$$D_2+D_3\equiv 2D_1 \text{ or } 0 \pmod 8.$$

A similar argument applied to the equation

$$2D_4W^2=D_2Y^2-D_3Z^2$$

shows that

$$D_2-D_3\equiv 2D_4 \text{ or } 0 \pmod 8.$$

It now follows that either $D_2+D_3\equiv 0 \pmod 8$ and $D_2\equiv D_4 \pmod 4$ or $D_2\equiv D_3 \pmod 8$ and $D_2\equiv D_1 \pmod 4$. Either possibility suffices for the 2-adic solubility of the system (3). For example, in the first case we can take $W=1$ and $X=0$ or 2 according as $D_2\equiv D_4$ or $4+D_4 \pmod 8$. This ensures that

$$D_2^{-1}(D_1X^2+D_4W^2)\equiv D_3^{-1}(D_1X^2-D_4W^2)\equiv 1 \pmod 8$$

so that Y and Z can be determined appropriately. A similar argument applies when $D_2\equiv D_3 \pmod 8$ and $D_2\equiv D_1 \pmod 4$.

We are now in a position to write down our formula for $2^{s(D)}$. When $p|D_1$, for example, the expression

$$\frac{1}{4}\left\{1+\left(\frac{D_2D_4}{p}\right)\right\}\left\{1+\left(\frac{-D_3D_4}{p}\right)\right\}=\frac{1}{4}\left\{1+\left(\frac{D_2D_4}{p}\right)+\left(\frac{-D_3D_4}{p}\right)+\left(\frac{-D_2D_3}{p}\right)\right\}$$

takes the values 1 or 0 according as D_2D_4 and $-D_3D_4$ are both squares (mod p) or not. Thus, on setting

$$\Pi_1=\prod_{p|D_1}\left(1+\left(\frac{D_2D_4}{p}\right)+\left(\frac{-D_3D_4}{p}\right)+\left(\frac{-D_2D_3}{p}\right)\right)$$

$$\Pi_2=\prod_{p|D_2}\left(1+\left(\frac{-D_1D_4}{p}\right)+\left(\frac{2D_1D_3}{p}\right)+\left(\frac{-2D_3D_4}{p}\right)\right)$$

$$\Pi_3=\prod_{p|D_3}\left(1+\left(\frac{2D_1D_2}{p}\right)+\left(\frac{D_1D_4}{p}\right)+\left(\frac{2D_2D_4}{p}\right)\right)$$

$$\Pi_4=\prod_{p|D_4}\left(1+\left(\frac{D_1D_2}{p}\right)+\left(\frac{D_1D_3}{p}\right)+\left(\frac{D_2D_3}{p}\right)\right),$$

we see that the product

$$4^{-\omega(D)}\Pi_1\Pi_2\Pi_3\Pi_4,$$

where $\omega(D)$ is the number of prime factors of D, will be 1 if the system (3) is everywhere locally solvable, and 0 otherwise. We can expand Π_1, for example, as

$$\Pi_1=\sum\left(\frac{D_2D_4}{D_{13}}\right)\left(\frac{-D_3D_4}{D_{12}}\right)\left(\frac{-D_2D_3}{D_{14}}\right),$$

where the sum is over all factorizations

$$D_1=D_{10}D_{12}D_{13}D_{14}.$$

For brevity we shall write the sum as $\sum f_1$. We shall expand the other factors Π_i in the same way, and write $\prod f_i=f(\mathbf{D})$. Here $\mathbf{D}$ represents the 16-tuple of elements D_{ij} with $1\leqq i\leqq 4$, $0\leqq j\leqq 4$ and $i\neq j$. According to Lemma 1, if we now sum over all quadruples D_1, D_2, D_3, D_4 with $D=D_1D_2D_3D_4$ we will get a total of $4^{\omega(D)}2^{s(D)}$. We therefore conclude as follows.

Lemma 3 *We have*

$$2^{s(D)}=\sum_{\mathbf{D}} g(\mathbf{D}),$$

where the sum is taken over all factorizations

$$D=\prod_{i,j} D_{ij},$$

and where

$$g(\mathbf{D})=\left(\frac{-1}{\alpha}\right)\left(\frac{2}{\beta}\right)\prod_i 4^{-\omega(D_{i0})}\prod_{j\neq 0}4^{-\omega(D_{ij})}\prod_{k\neq i,j}\prod_l\left(\frac{D_{kl}}{D_{ij}}\right)$$

with

$$\alpha=D_{12}D_{14}D_{23}D_{21},\qquad \beta=D_{24}D_{21}D_{34}D_{31}.$$

3 Averaging over D; Linked variables

In this section we begin our estimation of

$$\sum_{D\in S(X,h)} 2^{s(D)}.$$

Instead of summing over D we sum over the 16 variables D_{ij}, subject to the conditions that each D_{ij} is square-free, that they are coprime in pairs, and that their product D satisfies

$$D\leqq X,\qquad D\equiv h\,(\mathrm{mod}\ 8).$$

We divide the range of each variable D_{ij} into intervals $(A_{ij}, 2A_{ij}]$ where A_{ij} runs over powers of 2. This will give us $O(\log^{16}X)$ non-empty subsums, which

we shall write as $S(\mathbf{A})$, where $\mathbf{A}$ is the 16-tuple of numbers A_{ij}. Here we may suppose that

$$1 \ll \prod A_{ij} \ll X. \tag{4}$$

We shall describe the variables D_{ij} and D_{kl} as being 'linked' if $i \neq k$, and precisely one of the conditions $l \neq 0$, i or $j \neq 0$, k holds. This means that exactly one of the Jacobi symbols

$$\left(\frac{D_{kl}}{D_{ij}}\right), \quad \left(\frac{D_{ij}}{D_{kl}}\right)$$

occurs in the expression for $g(\mathbf{D})$. Let us suppose that the variables D_{ij} and D_{kl} are linked, and that it is the first of the above Jacobi symbols which occurs. We can then write $g(\mathbf{D})$ in the form

$$g(\mathbf{D}) = \left(\frac{D_{kl}}{D_{ij}}\right) a(D_{ij})\, b(D_{kl}),$$

where the function $a(D_{ij})$ depends on all the other variables D_{uv}, say, as well as D_{ij}, but is independent of D_{kl}, and similarly for the function $b(D_{kl})$. Moreover we have

$$|a(D_{ij})|, \ |b(D_{kl})| \leqq 1.$$

We can now write

$$|S(\mathbf{A})| \leqq \sum_{D_{uv}} \Big| \sum_{D_{ij},\, D_{kl}} \left(\frac{D_{kl}}{D_{ij}}\right) a(D_{ij})\, b(D_{kl}) \Big|.$$

The conditions that D_{ij} and D_{kl} should be coprime to each of the D_{uv} may be expressed by taking the functions a and b to vanish at appropriate values. Moreover the Jacobi symbol is automatically zero if the D_{ij} and D_{kl} are not coprime. The remaining conditions on these two variables may therefore be expressed by insisting that they are square-free and satisfy

$$D_{ij} D_{kl} \equiv h' \pmod 8, \qquad D_{ij} D_{kl} \leqq X',$$

where h' and X' will depend on the other variables D_{uv}. We now call on the following estimate which we shall prove in § 6.

Lemma 4 *Let a_m, b_n be complex numbers of modulus at most 1. Let an odd number h be given and let M, N, $X \geqslant 1$. Then*

$$\sum_{m,n} \left(\frac{n}{m}\right) a_m b_n \ll MN \{\min(M, N)\}^{-1/32},$$

uniformly in X, where the sum is for square-free m, n satisfying $M < m \leqq 2M$, $N < n \leqq 2N$, $mn \leqq X$, and $mn \equiv h \pmod 8$.

It immediately follows that

$$S(\mathbf{A}) \ll (\prod_{uv} A_{uv}) A_{ij} A_{kl} \{\min(A_{ij}, A_{kl})\}^{-1/32} \ll X \{\min(A_{ij}, A_{kl})\}^{-1/32},$$

by (4), and we deduce as follows.

Lemma 5 *We have*

$$S(\mathbf{A}) \ll X(\log X)^{-17}$$

whenever there is a pair of linked variables with

$$A_{ij}, A_{kl} \geqq \log^{544} X.$$

We now examine the case in which $A_{ij} \geqq \log^{544} X$, but every variable D_{kl} to which D_{ij} is linked has $A_{kl} < \log^{544} X$. We write D' for the product of these variables D_{kl}. Using the law of quadratic reciprocity we can now put $g(\mathbf{D})$ into the shape

$$4^{-\omega(D_{ij})}\left(\frac{D_{ij}}{D'}\right)\chi(D_{ij})c,$$

where χ is a character modulo 8, which may depend on the variables D_{uv} other than D_{ij}, and the where the remaining factor c is independent of D_{ij} and satisfies $|c| \leqq 1$. It follows that

$$|S(\mathbf{A})| \leqq \sum_{D_{uv}} \Big|\sum_{D_{ij}} 4^{-\omega(D_{ij})}\left(\frac{D_{ij}}{D'}\right)\chi(D_{ij})\Big|, \tag{5}$$

where the inner sum is restricted by the conditions that D_{ij} must be square-free and coprime to all the other variables D_{uv}, and that

$$D_{ij} \equiv h' (\mathrm{mod}\ 8), \qquad A_{ij} < D_{ij} \leqq \min(2A_{ij}, X'),$$

where h' and X' depend on the variables D_{uv} other than D_{ij}. We now apply the following result, which we shall prove in § 6.

Lemma 6 *Let $N > 0$ be given. Then for arbitrary positive integers q, r and any non-principal character $\chi(\mathrm{mod}\ q)$, we have*

$$\sum_{n \leqq x,\,(n,r)=1} \mu^2(n) 4^{-\omega(n)} \chi(n) \ll x d(r) \exp(-c\sqrt{\log x})$$

with a positive constant $c = c_N$, uniformly for $q \leqq \log^N x$.

To use this result we remove the condition $D_{ij} \equiv h' (\mathrm{mod}\ 8)$ from the inner sum on the right of (5) and insert instead a factor

$$\frac{1}{4} \sum_{\psi (\mathrm{mod}\ 8)} \psi(D_{ij}) \overline{\psi(h')}.$$

Taking

$$q = 8D' \ll (\log^{544} X)^{15}$$

and $r = \prod D_{uv}$, we conclude that

$$S(\mathbf{A}) \ll A_{ij} \exp(-c\sqrt{\log A_{ij}}) \sum_{D_{uv}} d(r),$$

providing that $D' \neq 1$. Since the variables D_{uv} are coprime in pairs we have $d(r) = \prod\prod d(D_{uv})$. Moreover for a single variable D_{kl} we will have

$$\sum_{D_{kl}} d(D_{kl}) \ll A_{kl} \log A_{kl} \ll A_{kl} \log X,$$

whence (4) yields

$$S(\mathbf{A}) \ll X(\log X)^{15} \exp(-c\sqrt{\log A_{ij}}),$$

providing that $D' \neq 1$. We can now summarize as follows.

Lemma 7 *There is an absolute constant $\kappa > 0$ such that*

$$S(\mathbf{A}) \ll X(\log X)^{-17}$$

whenever there are linked variables D_{ij} and D_{kl} for which

$$A_{ij} \geqq \exp\{\kappa(\log\log X)^2\} \tag{6}$$

and $D_{kl} > 1$.

We end this section with a straightforward estimate to handle the case in which at most three of the variables D_{ij} lie in ranges satisfying (6). For brevity we shall write

$$C = \exp\{\kappa(\log\log X)^2\}$$

and assume that C is a power of 2. Then if $\sum'$ indicates the condition that at most three of the A_{ij} satisfy $A_{ij} \geqq C$, we have

$$\sum_{A_{ij}}{}' |S(\mathbf{A})| \leqq \sum_{n_1 \ldots n_{16} \leqq X} 4^{-\omega(n_1)} \ldots 4^{-\omega(n_{16})},$$

where the n_i square-free and coprime in pairs, and at most three of the n_i have $n_i \geqq 2C$. We write

$$m = \prod_{n_i < 2C} n_i, \qquad n = \prod_{n_i \geqq 2C} n_i,$$

so that $m \leqq (2C)^{16}$ and $n \leqq X/m$. Moreover we see that each value of m can arise at most $16^{\omega(m)}$ times, and each value of n can arise at most $\binom{16}{3} 3^{\omega(n)}$ times. We may therefore deduce that

$$\sum_{A_{ij}}{}' |S(\mathbf{A})| \ll \sum_m 4^{\omega(m)} \sum_n \left(\tfrac{3}{4}\right)^{\omega(n)}.$$

We now use the bound

$$\sum_{n \leqq N} \gamma^{\omega(n)} \ll N(\log N)^{\gamma - 1}, \tag{7}$$

which is valid for any fixed $\gamma > 0$. Since

$$X/m \gg XC^{-16} \gg X^{1/2},$$

we have $\log X/m \gg \log X$, and we therefore find that

$$\sum_{A_{ij}}' |S(\mathbf{A})| \ll X(\log X)^{-1/4} \sum_{m} 4^{\omega(m)} m^{-1}.$$

A second application of (7), together with partial summation, shows that

$$\sum_{m \leqq M} 4^{\omega(m)} m^{-1} \ll \log^4 M,$$

whence

$$\sum_{A_{ij}}' |S(\mathbf{A})| \ll X(\log X)^{-1/4} \log^4 C \ll X(\log X)^{-1/4} (\log\log X)^8.$$

In view of Lemma 7 we may now summarize as follows.

Lemma 8 *We have*

$$\sum_{\mathbf{A}} |S(\mathbf{A})| \ll X(\log X)^{-1/4} (\log\log X)^8,$$

where the sum over $\mathbf{A}$ *is for all sets in which either there are at most three elements* $A_{ij} > C$, *or there are linked variables* D_{ij} *and* D_{kl} *with* $A_{ij} \geqq C$ *and* $D_{kl} > 1$.

4 Averaging over D; Characters modulo 8

We must now identify those sums $S(\mathbf{A})$ which are not eliminated by Lemma 8. There must be four or more elements $A_{ij} \geqq C$. If these include A_{10} and A_{20}, say, then we must have

$$D_{13} = D_{14} = D_{23} = D_{24} = D_{31} = D_{32} = D_{34} = D_{41} = D_{42} = D_{43} = 1,$$

since these variables are all linked to either A_{10} or A_{20}, or both. It follows that two or more of A_{12}, A_{21}, A_{30} or A_{40} must be at least C. If $A_{12} \geqq C$, then $D_{30} = D_{40} = 1$, since these variables are linked to D_{12}, and similarly if $A_{12} \geqq C$. On the other hand, if A_{30} or A_{40} is at least C, then we will have $D_{12} = D_{21} = 1$. We therefore conclude that when A_{10}, $A_{20} \geqq C$, we must have either A_{12}, $A_{21} \geqq C$ and the remaining variables all equal to 1, or A_{30}, $A_{40} \geqq C$ and the remaining variables all equal to 1. Of course an analogous conclusion holds whenever A_{i0}, $A_{j0} \geqq C$.

Now let us suppose that exactly one element A_{i0} satisfies $A_{i0} \geqq C$. Let us take this to be A_{10}. Then

$$D_{23} = D_{24} = D_{32} = D_{34} = D_{42} = D_{43} = 1,$$

these variables being linked to A_{10}. If A_{12}, $A_{21} \geqq C$, say, then also

$$D_{13} = D_{14} = D_{30} = D_{31} = D_{40} = D_{41} = 1,$$

and there is no fourth element A_{ij} which can be greater than or equal to C. A similar argument applies if A_{13}, $A_{31} \geqq C$ or A_{14}, $A_{41} \geqq C$. Hence we must have either A_{12}, A_{13}, $A_{14} \geqq C$, or A_{21}, A_{31}, $A_{41} \geqq C$, and in either case we

see that the remaining variables D_{ij} must all be equal to 1, since each one will be linked to a variable D_{kl} with $A_{kl} \geqq C$.

Finally we examine the case in which all of the variables A_{i0} are below C. If, say $A_{12}, A_{13} \geqq C$, then

$$D_{20} = D_{21} = D_{23} = D_{24} = D_{30} = D_{31} = D_{32} = D_{34} = D_{40} = D_{41} = 1.$$

If also $A_{14} \geqq C$ then $D_{42} = D_{43} = 1$, so that there cannot be four elements $A_{ij} \geqq C$. We must therefore have $A_{42}, A_{43} \geqq C$, whence all the remaining variables D_{ij} will be 1. An analogous argument applies whenever $A_{ij}, A_{ik} \geqq C$ with i, j, k distinct. There remains the possibility that the four elements for which $A_{ij} \geqq C$ have four different values for i. If one of these is A_{12}, say, then

$$D_{20} = D_{23} = D_{24} = 1,$$

since these are linked to D_{12}, and so $A_{21} \geqq C$. The only variables linked to neither of D_{12}, D_{21} are D_{10}, D_{20}, D_{34} and D_{43}. It follows that $A_{34}, A_{43} \geqq C$, and hence that all remaining variables D_{ij} must be 1.

We summarize our conclusions as follows.

Lemma 9 *A sum $S(\mathbf{A})$ which is not considered by Lemma 8 must have exactly four elements $A_{ij} \geqq C$, and the remaining variables D_{kl} must take the value 1. The possible sets of indices ij are*

$$10, 20, 30, 40,$$
$$i0, j0, ij, ji,$$
$$i0, ij, ik, il,$$
$$i0, ji, ki, li,$$
$$ij, ik, lj, lk,$$

and

$$ij, ji, kl, lk,$$

where i, j, k, l denote different non-zero indices.

It remains to handle these 24 types of sum. We shall rename the variables D_{ij} which occur non-trivially as $n_1, \ldots, n_4$, and write $N_1, \ldots, N_4$ for the corresponding A_{ij}. We shall describe the variables N_i, N_j as being 'joined' if both Jacobi symbols

$$\left(\frac{N_i}{N_j}\right), \quad \left(\frac{N_j}{N_i}\right)$$

occur in the definition of $g(\mathbf{D})$. Thus D_{ij}, D_{kl} are joined if $i \neq k$ and $j, l \neq i, k, 0$. If two variables are not joined we shall say they are 'independent'. By abuse of terminology we shall also refer to the indices ij and kl as being joined or independent, as appropriate. For each $\mathbf{A}$ occurring in Lemma 9 we may now write $S(\mathbf{A})$ in the form

$$\sum_{n_1, \ldots, n_4} \chi_1(n_1) \ldots \chi_4(n_4) PQ, \qquad Q = 4^{-\omega(n_1 \ldots n_4)}, \tag{8}$$

where the variables are square-free, coprime in pairs, and satisfy $N_i < n_i \leqq 2N_i$. Here the characters χ_i are to modulus 8, and are exactly those arising from the terms $\left(\frac{-1}{\alpha}\right)$ and $\left(\frac{2}{\beta}\right)$ in the definition of $g(\mathbf{D})$. The factor P is the result of applying the law of quadratic reciprocity, to produce a product of expressions

$$(-1)^{(n_i-1)(n_j-1)/4},$$

one for each pair of joined variables.

For each type of sum in Lemma 9 there is at least one pair of independent variables. Thus, by re-labeling the variables n_i as necessary we can estimate the expression (8) as

$$\sum_{n_1,n_2} \Big| \sum_{n_3,n_4} \chi_3(n_3)\chi_4(n_4) PQ \Big| \tag{9}$$

with n_3, n_4 independent. It follows that P can be written as a product of characters $\psi_3(n_3)$, $\psi_4(n_4)$ modulo 4, depending on n_1, n_2, together with a factor depending on n_1, n_2 alone. We claim that, except for the indices

$$10, 20, 30, 40; \qquad 40, 41, 42, 43; \qquad i0, ji, ki, li,$$

we can choose the labeling so that $\psi_3 = \psi_4$ and $\chi_3 \neq \chi_4$. To justify the first of these conditions we shall arrange that n_1, n_3 are joined if and only if n_1, n_4 are joined, and similarly for n_2, n_3 and n_2, n_4.

To justify our claim we first observe that the character χ corresponding to D_{12}, D_{14} and D_{23} is $\left(\frac{-1}{*}\right)$, the character corresponding to D_{24}, D_{31} and D_{34} is $\left(\frac{2}{*}\right)$, and the character corresponding to D_{21} is $\left(\frac{-2}{*}\right)$. The remaining variables have the trivial character. We begin by considering sums with indices $i0$, $j0$, ij, ji. Here one or other of ij or ji, say ij, automatically corresponds to a nontrivial character χ. We may then take

$$n_1 = D_{i0}, \quad n_2 = D_{ji}, \quad n_3 = D_{j0}, \quad n_4 = D_{ij}$$

since every pair of variables here is independent. Next we examine sums with indices $i0$, ij, ik, il. If $i \neq 4$ then at least one of ij, ik, il, say ij, corresponds to a non-trivial character χ. Again each pair of variables is independent, and we can take

$$n_1 = D_{ik}, \quad n_2 = D_{il}, \quad n_3 = D_{i0}, \quad n_4 = D_{ij}.$$

For sums with indices ij, ji, kl, lk, we observe that ij, ji necessarily correspond to different characters χ, and have independent variables associated to them. Moreover D_{ij} is joined to both D_{kl} and D_{lk}, as is D_{ji}. In this case we may therefore take

$$n_1 = D_{kl}, \quad n_2 = D_{lk}, \quad n_3 = D_{ij}, \quad n_4 = D_{ji}.$$

Finally, for sums with indices ij, ik, lj, lk we observe that we can assume D_{ij}, D_{ik} to correspond to different characters χ. This is clearly true if $i = 2$, or by

interchanging the labels i and l, if $l=2$. We may therefore suppose that j, say is 2. Now, if $i=3$, then D_{32} and D_{3k} will have different associated characters χ, whether $k=1$ or 4. A similar argument applies if $l=3$, so we may take $k=3$, whence D_{12} and D_{13} will be variables with different associated characters χ. Finally, if D_{ij}, D_{ik} correspond to different characters χ, we can take

$$n_1=D_{lj}, \quad n_2=D_{lk}, \quad n_3=D_{ij}, \quad n_4=D_{ik},$$

since ij, ik are independent, whereas ij and ik are both joined to lj and lk.
We have now verified that, for the sums in question, (9) may be put into the shape

$$\sum_{n_1,n_2} \Big| \sum_{n_3,n_4} \psi_3\chi_3(n_3)\psi_4\chi_4(n_4)Q\Big|, \tag{10}$$

with $\psi_3\chi_3\neq\psi_4\chi_4$. In fact this is also true for sums with indices $i0$, ji, ki, li when $i=1, 4$. Here there is always at least one variable, ji, say, whose associated character χ is $\left(\frac{2}{*}\right)$. We may then choose

$$n_1=D_{ki}, \quad n_2=D_{0i}, \quad n_3=D_{i0}, \quad n_4=D_{ji},$$

which makes n_3 and n_4 independent. Moreover, since ψ_3 and ψ_4 are characters modulo 4, we automatically have $\psi_3\chi_3\neq\psi_4\chi_4$.

We may proceed to apply the following lemma which we shall establish in § 6.

Lemma 10 *Let $X>0$ and $M, N\geqq C>0$ be given. Then for an arbitrary positive integer r, any odd integer h, and any distinct characters $\chi_1, \chi_2 \pmod 8$, we have*

$$\sum_{m,n} \mu^2(m)\mu^2(n)4^{-\omega(m)-\omega(n)}\chi_1(m)\chi_2(n) \ll d(r)X\exp(-c\sqrt{\log C})\log X,$$

for some positive absolute constant c, where the sum is over coprime variables satisfying the conditions

$$M<m\leqq 2M, \quad N<n\leqq 2N, \quad mn\leqq X, \quad mn\equiv h \pmod 8, \quad (mn,r)=1.$$

It follows that the sums $S(\mathbf{A})$ in question are all $O(X(\log X)^{-17})$, since the constant κ in Lemma 7 may be taken sufficiently large. The total contribution of these sums is therefore $O(X(\log X)^{-1})$, which is satisfactory. We summarize as follows.

Lemma 11 *We have*

$$\sum_{\mathbf{A}} S(\mathbf{A}) \ll X(\log X)^{-1/4}(\log\log X)^8,$$

where the sum over $\mathbf{A}$ is for all sets other than those corresponding to indices

$$10, 20, 30, 40; \quad 40, 41, 42, 43; \quad 20, 12, 32, 42; \quad 30, 13, 23, 43.$$

5 The leading terms

For sums with indices 10, 20, 30, 40 or 40, 41, 42, 43 the function $g(\mathbf{D})$ merely reduces to $4^{-\omega(D)}$, where D is the product of the variables D_{ij}. The contribution of all sums with indices 10, 20, 30, 40 and $A_{i0} \geqq C$, is therefore

$$\sum_{D_{i0}} 4^{-\omega(D)},$$

where the sum is subject to the conditions

$$D_{i0} > C, \quad D \leqq X, \quad D \equiv h (\mathrm{mod}\ 8), \quad D \text{ square-free.}$$

We can remove the condition $D_{i0} > C$ with an error

$$\begin{aligned}
&\ll \sum_{abcd \leqq X,\, a \leqq C} \mu^2(abcd) 4^{-\omega(a)-\omega(b)-\omega(c)-\omega(d)} \\
&= \sum_{ae \leqq X,\, a \leqq C} \mu^2(ae) 4^{-\omega(a)} (\tfrac{3}{4})^{w(e)} \\
&\ll \sum_{a \leqq C} 4^{-\omega(a)} \sum_{e \leqq X/a} (\tfrac{3}{4})^{\omega(e)} \\
&\ll X(\log X)^{-1/4} \sum_{a \leqq C} 4^{-\omega(a)} a^{-1} \\
&\ll X(\log X)^{-1/4} (\log\log X)^2,
\end{aligned}$$

by (7). Since D is square-free it factorizes as $D_{10}D_{20}D_{30}D_{40}$ in exactly $4^{\omega(D)}$ different ways. We therefore obtain

$$\sum_{D \leqq X} 1 + O(X(\log X)^{-1/4}(\log\log X)^2) = \# S(X,h) + O(X(\log X)^{-1/4}(\log\log X)^2).$$

Precisely the same argument applies to sums with indices 40, 41, 42, 43.

The situation for sums with indices 20, 12, 32, 42 is slightly more complicated. Here we can compute, using the law of quadratic reciprocity that $g(\mathbf{D})$ reduces to

$$\frac{1}{2}\left\{1 + \left(\frac{-1}{D_{12}D_{32}}\right) + \left(\frac{-1}{D_{12}D_{42}}\right) - \left(\frac{-1}{D_{32}D_{42}}\right)\right\} 4^{-\omega(D)}.$$

The term involving $\left(\frac{-1}{D_{12}D_{32}}\right)$, for example, may be handled using Lemma 10 as before, by putting the relevent part of the sum into the form (10) with

$$n_1 = D_{20}, \quad n_2 = D_{12}, \quad n_3 = D_{32}, \quad n_4 = D_{34}.$$

The terms containing $\left(\frac{-1}{D_{12}D_{42}}\right)$ and $\left(\frac{-1}{D_{32}D_{42}}\right)$ may be handled in precisely the same way, while the leading term, when summed over all appropriate vectors $\mathbf{A}$, yields

$$\tfrac{1}{2} \# S(X,h) + O(X(\log X)^{-1/4}(\log\log X)^2), \tag{11}$$

by exactly the same argument as above. Finally we observe that sums with indices 30, 13, 23, 43 behave in precisely the same way, and again contribute a total of the form (11). Theorem 1 now follows.

6 Lemmas on character sums

It remains to give the proof of Lemmas 4, 6 and 10. We start with Lemma 4. Results of this type appear to have their origins in work of Heilbronn [8]. We begin by writing our sum as

$$\sum_{i,j(\mathrm{mod}\ 8)} \sum_{m\equiv i(\mathrm{mod}\ 8)} \sum_{n\equiv j(\mathrm{mod}\ 8)} \left(\frac{n}{m}\right) a_m b_n,$$

where m, n are square-free, and i, j are restricted to satisfy $ij\equiv h(\mathrm{mod}\ 8)$. The restrictions on m and n mean that the summand is now essentially symmetrical between m and n, by the law of quadratic reciprocity. We may therefore suppose that $N\geqq M$, whence it suffices to prove that

$$\sum_{N<n\leqq 2N} \left| \sum_{M<m\leqq 2M} a_m \left(\frac{n}{m}\right)\right| \ll NM^{31/32}.$$

Here we have dropped all the conditions on n, but we have to retain the restrictions on m. By Cauchy's inequality the sum on the left is at most

$$N^{1/2}\left\{\sum_n \left| \sum_{M<m\leqq 2M} a_m \left(\frac{n}{m}\right)\right|^2\right\}^{1/2},$$

and on expanding the sum, and inverting the order of summations we get at most

$$N^{1/2}\left\{\sum_{m_1,m_2} \left|\sum_n \left(\frac{n}{m_1 m_2}\right)\right|\right\}^{1/2},$$

with m_1, m_2 still restricted to be square-free. The innermost sum is therefore trivial only when $m_1 = m_2$, contributing

$$N^{1/2}\{MN\}^{1/2},$$

which is satisfactory. When $m_1 \neq m_2$ we may estimate the inner sum as

$$\sum_{N<n\leqq 2N} \left(\frac{n}{m_1 m_2}\right) \ll N^{1/2}(m_1 m_2)^{3/16+\varepsilon}$$

for any $\varepsilon>0$, by Burgess' bound [4]. Taking $\varepsilon = 1/32$, we get a total contribution

$$N^{1/2}\{M^2 N^{1/2} M^{7/16}\}^{1/2} \leqq NM^{31/32},$$

since $M\leqq N$, and this also is satisfactory. This proves Lemma 4.

We turn now to Lemma 6. For the proof we introduce the Dirichlet series

$$f(s)=\prod_{p\nmid r}\left(1+\frac{\chi(p)}{4p^s}\right)=\sum_{(n,r)=1}\mu^2(n)\,4^{-\omega(n)}\chi(n)\,n^{-s}$$

and

$$g(s)=\prod_{p\mid r}\left(1-\frac{\chi(p)}{p^s}\right)\prod_{p\nmid r}\left\{\left(1-\frac{\chi(p)}{p^s}\right)\left(1+\frac{\chi(p)}{4p^s}\right)^4\right\}.$$

The products for $f(s)$ and $g(s)$ converge absolutely for $\Re(s)>1$ and $\Re(s)>\frac{1}{2}$ respectively. Moreover, since $f(s)^4=g(s)\,L(s,\chi)$, the function $f(s)$ has an analytic continuation into any region $\sigma\geqq\sigma_0>\frac{1}{2}$, $|t|\leqq T$ free of zeros of $L(s,\chi)$. We now recall that there are constants $c_1>0$ and $c_2(\varepsilon)>0$ such that $L(s,\chi)$ has no complex zeros for

$$\sigma\geqq 1-\frac{c_1}{\log qT},\ |t|\leqq T,$$

for $T\geqq 2$, and, by Siegel's Theorem, no real zeros for $\sigma\geqq 1-c_2(\varepsilon)\,q^{-\varepsilon}$. On taking $\varepsilon=1/2N$ and $T=\exp(\sqrt{\log x})$, the condition $q\leqq(\log x)^N$ gives us a zero-free region

$$R=\{s:\sigma\geqq 1-\delta,|t|\leqq T\},\qquad \delta=\frac{c_3}{\log T},$$

for an appropriate $c_3=c_3(N)>0$. Moreover for such s we have

$$L(s,\chi)\ll(1+(qT)^{(1-\sigma)/2})\log T\ll\log T.$$

We also have the trivial bound $g(s)\ll d(r)$ in the region R. We may therefore conclude that $f(s)$ has an analytic continuation in the region R and satisfies $f(s)\ll d(r)\log x$ there.

We now apply Perron's formula (see Titchmarsh [13, Lemma 3.19]) to give

$$\sum_{n\leqq x,(n,r)=1}\mu^2(n)\,4^{-\omega(n)}\chi(n)=\frac{1}{2\pi i}\int_{\alpha-iT}^{\alpha+iT}f(s)\,x^s\frac{ds}{s}+O\left(\frac{x\log x}{T}\right)+O(1),$$

where

$$\alpha=1+\frac{1}{\log x},$$

and $T=\exp(\sqrt{\log x})$ as before. We shall replace the path of integration by three line segments from $\alpha-iT$ to $1-\delta-iT$ to $1-\delta+iT$ to $\alpha+iT$. From the first

and third of these we get a contribution $O(xd(r)/T)$, and from the second, a contribution $O(x^{1-\delta}d(r)\log^2 x)$. It follows that

$$\sum_{n\leqq x,(n,r)=1}\mu^2(n)4^{-\omega(n)}\chi(n)\ll\frac{x\log x}{T}+1+\frac{xd(r)}{T}+x^{1-\delta}d(r)\log^2 x$$
$$\ll xd(r)\exp(-c\sqrt{\log x})$$

for an appropriate constant $c>0$. This completes the proof of Lemma 6.

Finally we consider Lemma 10. This is in fact a straightforward deduction from Lemma 6. We shall suppose that $MN\leqq X$, for otherwise the sum in question is empty. We remove the condition $mn\equiv h(\mathrm{mod}\ 8)$ from the summation and instead introduce the factor

$$\frac{1}{4}\sum_{\psi(\mathrm{mod}\ 8)}\psi(mn)\overline{\psi(h)}.$$

We shall estimate individually the sums corresponding to each character ψ. Since $\psi\chi_1\neq\psi\chi_2$, we may suppose that $\psi\chi_1$, say, is non-principal. Then the double sum under consideration is at most

$$\sum_n\Big|\sum_m\mu^2(m)4^{-\omega(m)}\psi(m)\chi_1(m)\Big|.$$

The inner sum here is subject to the conditions $(m, nr)=1$ and

$$M<m\leqq\min(2M, X/n).$$

Thus Lemma 6 provides an estimate

$$\ll Md(r)d(n)\exp(-c\sqrt{\log M})$$

for each of the inner sums. On summing over n we now obtain a bound

$$\ll Md(r)N(\log N)\exp(-c\sqrt{\log M}),$$

which is satisfactory. This completes the proof of the lemma.

7 Proof of Theorem 2

In order to prove Theorem 2 we shall construct numbers $D=p_1\ldots p_k$ with distinct prime factors $p_i\equiv 1(\mathrm{mod}\ 8)$ for which

$$\left(\frac{p_i}{p_j}\right)=1,\qquad i\neq j.$$

According to Lemma 1 we will then have $s(D)=2k$. We shall take P to be a sufficiently large parameter and restrict the prime factors p_i to lie in the range $P/2<p_i\leqq P$. We shall be interested in a certain subset, $\mathscr{S}(P, k)$ say, of these

numbers D. We shall say that a character χ is 'good' if either its conductor is a divisor of 8, or if the Dirichlet L-function $L(s, \chi)$ has no zeros in the region

$$\Re(s) \geqq \tfrac{15}{16}, \quad |\Im(s)| \leqq P.$$

We shall also say that D is 'good' if every real character to modulus $8D$ is good. We shall define $\mathscr{S}(P, k)$ to be the set of numbers D, of the form already described, which are good. We shall also write $\mathscr{S}(P, k, q)$ for those elements of $\mathscr{S}(P, k)$ which are multiples of q.

We shall put $\omega(q)=j$. Using induction on k, we shall show that

$$\#\mathscr{S}(P, k) \geqq \frac{1}{k!}\left(\frac{P}{8 \log P}\right)^k 2^{-3k-k(k-1)/2} \tag{12}$$

provided that

$$2^k < P^{1/40}, \tag{13}$$

and

$$\#\mathscr{S}(P, k+j, q) \leqq \frac{1}{k!}\left(\frac{P}{8 \log P}\right)^k 2^{2k-kj-k(k-1)/2}, \tag{14}$$

provided that $q \in \mathscr{S}(P, j)$ and

$$2^{k+j} \leqq P^{1/40}. \tag{15}$$

We begin by observing that when $k=0$ we will have $\mathscr{S}(P, 0)=\{1\}$, and $\mathscr{S}(P, j, q)=\{q\}$, so that (12) and (14) are certainly true. This establishes the base step for our induction. In getting from the case k to the case $k+1$ we shall first prove (14) and then (12). While doing this we shall use the fact that $P \geqq P_0$, say, is sufficiently large, and we must be careful to ensure that P_0 is independent of k and q.

To prove (12) and (14) we shall take $D \in \mathscr{S}(P, k)$ or $D \in \mathscr{S}(P, k+j, q)$ respectively, and write $l=k$ or $k+j$ as appropriate. In both cases we begin by counting the number of primes p in the set

$$\mathscr{D}(D)=\left\{P/2 < p \leqq P : p \equiv 1 \pmod 8, \left(\frac{p}{p_i}\right)=1 \text{ for } 1 \leqq i \leqq l\right\}.$$

If we let χ run over all real characters modulo $8D$ we see that

$$2^{-2-l}\sum_{\chi} \chi(p)$$

takes the value 1 if $p \equiv 1 \pmod 8$ and

$$\left(\frac{p}{p_i}\right)=1, \quad 1 \leqq i \leqq l,$$

and 0 otherwise. We now define

$$A(\chi)=\sum_{P/2<p\leqq P}\chi(p)\,(P-|4p-3P|)\log p$$

and

$$B(\chi)=\sum_{P/2<n\leqq P}\chi(n)\,(P-|4n-3P|)\,\Lambda(n).$$

Then if p runs over $\mathscr{D}(D)$ we will have

$$(P\log P)\#\mathscr{D}(D)\geqq\sum_{p}(P-|4p-3P|)\log p=2^{-2-l}\sum_{\chi}A(\chi). \tag{16}$$

We observe moreover that

$$A(\chi)=B(\chi)+O(P^{3/2})=B(\chi^*)+O(lP\log P)+O(P^{3/2}), \tag{17}$$

where χ^* is the primitive character which induces χ. When the conductor of χ^* does not divide 8, we handle $B(\chi^*)$ by means of the integral representation

$$B(\chi^*)=\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\left\{-\frac{L'}{L}(s,\chi^*)\right\}P^{s+1}w(s)\,ds,$$

where

$$w(s)=\frac{4(1+(1/2)^{s+1}-2(3/4)^{s+1})}{s(s+1)}.$$

We move the line of integration to $\Re(s)=-\frac{1}{2}$, where the integral is $O(P^{1/2}\log D)$, and obtain the formula

$$B(\chi^*)=-\sum_{\rho}P^{\rho+1}w(\rho)+O(lP^{1/2}\log P). \tag{18}$$

Here ρ runs over all non-trivial zeros of $L(s,\chi^*)$.

Since χ is 'good', all such zeros satisfy either $\Re(\rho)\leqq\frac{15}{16}$ or $|\Im(\rho)|\geqq P$. The symmetry of the zeros about the critical line then ensures that $|\rho|\geqq\frac{1}{16}$ for each zero. We can then conclude that

$$\sum|w(\rho)|^{-1}\ll\log D,$$

and

$$\sum_{|\Im(\rho)|\geqq P}|w(\rho)|^{-1}\ll P^{-1}\log D,$$

since there are $O(T\log DT)$ zeros up to height T. It now follows from (18) that

$$B(\chi^*)\ll lP^{31/16}\log P.$$

If the conductor of χ^* divides 8, the Prime Number Theorem for arithmetic progressions modulo 8 yields

$$B(\chi^*)=\varepsilon\frac{P^2}{4}+O(P^2(\log P)^{-1}),$$

where $\varepsilon=1$ if χ^* is identically 1, and $\varepsilon=0$ otherwise. A comparison with (16) and (17) now reveals that

$$\#\mathscr{D}(D)\geqq 2^{-2-l}\frac{P}{4}(\log P)^{-1}+O(lP^{15/16})+O(2^{-l}P(\log P)^{-2}).$$

There is therefore an absolute constant C_1 such that

$$\#\mathscr{D}(D)\geqq 2^{-5-l}\frac{P}{\log P}, \tag{19}$$

providing that

$$2^l\leqq C_1\,P^{1/17}.$$

This last condition is a consequence of (13) or (15), if P is large enough. A precisely similar argument, based on the fact that

$$(P\log P/2)\#\mathscr{D}(D)\leqq\sum_{P/4<p\leqq 5P/4}(2P-|4p-3P|)\log p,$$

shows that

$$\#\mathscr{D}(D)\leqq 2^{-1-l}\frac{P}{\log P}, \tag{20}$$

subject similarly to the conditions (13) and (15).

We are now ready to prove (14) in the case $k+1$. Each $D'\in\mathscr{S}(P,k+j+1,q)$ can be written in exactly $k+1$ ways as Dp with $q|D$. Moreover we must have $p\in\mathscr{D}(D)$. Not all products Dp with $D\in\mathscr{S}(P,k+j,q)$ and $p\in\mathscr{D}(D)$ will be good. However it certainly follows that

$$\begin{aligned}\#\mathscr{S}(P,k+j+1,q)&\leqq\frac{1}{k+1}\sum_{D\in\mathscr{S}(P,k+j,q)}\#\mathscr{D}(D)\\&\leqq\frac{1}{k+1}\cdot\frac{1}{k!}\left(\frac{P}{8\log P}\right)^k 2^{2k-kj-k(k-1)/2}\cdot 2^{-1-k-j}\frac{P}{\log P}\\&=\frac{1}{(k+1)!}\left(\frac{P}{8\log P}\right)^{k+1}2^{2(k+1)-(k+1)j-k(k+1)/2},\end{aligned}$$

by means of (20) and the case k of (14). This establishes (14) for $k+1$.

For the proof of (12) we observe that each $D'\in\mathscr{S}(P,k+1)$ can be written in exactly $k+1$ ways as $D'=Dp$, and in each such representation we have

$p \in \mathscr{D}(D)$. We shall show that at least half of the numbers Dp formed in this way are good. We will then have

$$\begin{aligned}\#\mathscr{S}(P, k+1) &\geqq \frac{1}{2k+2} \sum_{D \in \mathscr{S}(P,k)} \#\mathscr{D}(D) \\ &\geqq \frac{1}{2k+2} \cdot \frac{1}{k!} \left(\frac{P}{8 \log P}\right)^k 2^{-3k-k(k-1)/2} \cdot 2^{-5-k} \frac{P}{\log P} \\ &= \frac{1}{(k+1)!} \left(\frac{P}{8 \log P}\right)^{k+1} 2^{-3(k+1)-k(k+1)/2},\end{aligned}$$

by means of (19) and the case k of (12). This establishes (12) for $k+1$.

We must now see how many values of Dp fail to be good. There is then some character χ to modulus $8Dp$ which is not good. If the conductor of χ is q, say, then $q \nmid 8D$, since D is good. Thus $p \mid q$, and $q = pq'$, $4pq'$ or $8pq'$ with $q' \mid D$. It follows that, to each such q with $\omega(q') = j \leqq k$, there correspond at most $j+1$ choices for p, and at most

$$\#\mathscr{S}(P, k, q') \leqq \frac{1}{(k-j)!} \left(\frac{P}{8 \log P}\right)^{k-j} 2^e$$

choices for D, where

$$\begin{aligned}e &= 2(k-j) - (k-j)j - (k-j)(k-j-1)/2 \\ &\leqq 3(k+1) - k(k+1)/2 + (j+1)^2/2.\end{aligned}$$

We therefore see that each conductor q will divide at most

$$(k+1)^{j+2} 2^{(j+1)^2} \left(\frac{8 \log P}{P}\right)^{j+1} \frac{1}{(k+1)!} \left(\frac{P}{8 \log P}\right)^{k+1} 2^{3(k+1)-k(k+1)/2}$$

numbers $8Dp$. Since $q \leqq 8P^{j+1}$, we have

$$\begin{aligned}(k+1)^{j+2} 2^{(j+1)^2} \left(\frac{8 \log P}{P}\right)^{j+1} &\leqq \tfrac{1}{4} \left\{(k+1)^2 2^{j+1} \frac{32 \log P}{P}\right\}^{j+1} \\ &\leqq \tfrac{1}{4} \{P^{-2/3}\}^{j+1} \\ &\leqq q^{-2/3},\end{aligned}$$

providing that

$$(k+1)^2 2^{k+1} \leqq \frac{P^{1/3}}{32 \log P}.$$

This condition follows from (13) if P is sufficiently large.

According to the zero-density theorem of Montgomery [10] the total number of zeros in

$$\Re(s) \geqq \sigma, \quad |\Im(s)| \leqq T,$$

for all primitive L-functions of conductor at most Q, is

$$\ll (Q^2 T)^{2(1-\sigma)/\sigma} (\log QT)^{13}, \quad \tfrac{4}{5} \leqq \sigma \leqq 1.$$

For $\sigma = \frac{15}{16}$, $T = P$ and $Q \gg P$, the number of conductors which can occur is therefore $O(Q^{1/2})$. Since the relevent conductors q are all at least $P/2$, we see that the number of products Dp which are not good is at most

$$\sum_{q > P/2} q^{-2/3} \frac{1}{(k+1)!} \left(\frac{P}{8 \log P} \right)^{k+1} 2^{3(k+1) - k(k+1)/2}$$

$$\ll P^{-1/6} \frac{1}{(k+1)!} \left(\frac{P}{8 \log P} \right)^{k+1} 2^{3(k+1) - k(k+1)/2}.$$

It follows that at least half the available values of Dp are good, providing that

$$2^{6(k+1)} \leqq C_2 P^{1/6},$$

with a suitable constant C_2. This inequality follows from (13) if P is large enough. This completes our proof of (12).

It remains to deduce Theorem 2 from the estimate (12). We shall take

$$k = \left[\sqrt{\frac{(1-\theta) \log X}{\log 4}} \right]$$

and

$$P = X^{1/k}.$$

It follows that both k and P tend to infinity with X. We will then have a set $\mathscr{S}(P, k)$ of numbers $D \leqq X$, for which

$$s(D) = 2k \gg \sqrt{\log X}$$

as required. If θ is sufficiently close to 1, we will have

$$2^{k^2} \leqq X^{(1-\theta)/2} \leqq X^{1/40},$$

whence (13) will hold. We will then have

$$\# \mathscr{S}(P, k) \geqq \left(\frac{P}{8k\, 2^{3+(k-1)/2} \log P} \right)^k \geqq (P^\theta)^k = X^\theta,$$

since

$$8k\, 2^{3+(k-1)/2} \leqq 2^k \leqq X^{(1-\theta)/2k} = P^{(1-\theta)/2} \leqq \frac{P^{1-\theta}}{\log P},$$

for sufficiently large X. This yields the estimate claimed in Theorem 2.

8 Proof of Theorem 3

When $h=1, n=0$ we observe that $s(3p)=0$ for any prime $p \equiv 3 \pmod 8$. Otherwise we write

$$k=\begin{cases} n/2-1, & h=1, \\ n/2, & h=3, \\ (n-1)/2, & h=5 \text{ or } 7. \end{cases}$$

We then use the construction of the previous section to produce a product $D_0=p_1 \dots p_k$ with $p_i \equiv 1 \pmod 8$ and

$$\left(\frac{p_i}{p_j}\right)=1, \quad i \neq j.$$

We shall keep D_0 fixed[1], and consider numbers of the form $D=D_0 p^*$, with $p^* \equiv h \pmod 8$ and

$$\left(\frac{p^*}{p_i}\right)=1, \quad 1 \leqq i \leqq k.$$

The system (3) now has p-adic solutions for each p_i. Moreover, if $p^* | D_1$ there will be p^*-adic solutions if and only if -1 is a quadratic residue of p^*. For $p^* | D_2$ the condition is that -1 and 2 are both quadratic residues, but for $p^* | D_3$ we only need 2 to be a quadratic residue. Finally if $p^* | D_4$ the system automatically has solutions. Thus for $h=1$ each of the systems (3) is everywhere locally solvable. For $h=3$ only one quarter of them are everywhere locally solvable, and for $h=5$ or 7 one half of them are admissable. Lemma 1 now shows that $s(D)=n$ in each case, in view of our choice of k. It remains to observe that our conditions on p^* can be achieved by requiring p^* to lie in a suitable congruence class modulo $8D_0$, so that the number of available primes p^* is $\gg X/\log X$. The theorem then follows.

Acknowledgement. The present paper was prepared while the author was enjoying the hospitality and financial support of the University of Hong Kong. This assistance is gratefully acknowledged.

References

1. Birch, B.J., Swinnerton-Dyer, H.P.F.: Notes on elliptic curves, II. J. Reine Angew. Math. **218**, 79–108 (1965)
2. Birch, B.J., Stephens, N.M.: The parity of the rank of the Mordell-Weil group. Topology **5**, 295–299 (1966)
3. Brumer, A., Heath-Brown, D.R.: Average ranks of elliptic curves, II (to appear)
4. Burgess, D.A.: On character sums and L-series, II. Proc. Lond. Math. Soc., III. Ser. **13**, 524–536 (1963)

[1] By extending the argument, to allow D_0 to vary, one could produce more admissable values of D, thereby improving the bound in Theorem 3 somewhat.

5. Coates, J., Wiles, A.: On the conjecture of Birch and Swinnerton-Dyer. Invent. Math. **39**, 223–251 (1977)
6. Gouvêa, F., Mazur, B.: The square-free sieve and the rank of elliptic curves. J. Am. Math. Soc. **4**, 1–23 (1991)
7. Gross, B.H., Zagier, D.: Heegner points and derivatives of L-series. Invent. Math. **84**, 225–320 (1986)
8. Heilbronn, H.A.: On the averages of some arithmetic functions of two variables. Mathematika **5**, 1–7 (1958)
9. Lagrange, J.: Nombres congruents et courbes elliptiques. In: Sém. Delange-Pisot-Poitou (Théorie des nombres), no 16, 16e ennée 1974/75
10. Montgomery, H.L.: Zeros of L-functions. Invent. Math. **8**, 346 354 (1969)
11. Rubin, K.: Tate-Shafarevich groups and L-functions of elliptic curves with complex multiplication. Invent Math. **89**, 527 566 (1987)
12. Serf, P.: Congruent numbers and elliptic curves. In: Pethö, A. et al. (eds.) Computational number theory, pp. 227 238. Berlin: Gruyter 1991
13. Titchmarsh, E.C.: The theory of the Riemann Zeta-function, 2nd ed., revised by D.R. Heath-Brown. Oxford: Clarendon Press 1986