

On the mean squared modulus of a Dirichlet L -function over a short segment of the critical line

by

N. WATT (Al Ain)

Contents

1. Introduction	307
2. The approximation step of the Bombieri–Iwaniec method	316
3. Major arcs	323
4. Minor arcs: the first Poisson summation	325
5. The second Poisson summation	329
6. The Bombieri–Iwaniec Double Large Sieve	340
7. The First Spacing Problem with a Dirichlet character factor	347
8. The Second Spacing Problem	379
9. Bounds for $S(H, M)$	385
10. Application to $I(t, U, \chi)$: the proof of Theorem 1	397
References	401

1. Introduction. This paper is mainly about our recent work on the mean-value

$$(1.1) \quad I(t, U, \chi) = \frac{1}{2U} \int_{t-U}^{t+U} \left| L\left(\frac{1}{2} + i\tau, \chi\right) \right|^2 d\tau,$$

where $1 \leq U \leq t$ and $L(s, \chi)$ is the Dirichlet L -function for a Dirichlet character χ that is primitive to the modulus r . The new results we have obtained are of interest in some “local” cases, where $U = o(t)$. In the “global” case where $U = t = T/2$ (say), the mean-value is over the interval $[0, T]$, so we write $I_1(T, \chi)$ for $I(T/2, T/2, \chi)$. In [25] Motohashi has given an asymptotic formula for $I_1(T, \chi)$ with an error-term estimate sharp enough to imply that, if one restricts to cases where r is prime and $(rt)^{1/3+\varepsilon} \leq U \leq t^{1-\varepsilon}$ (with ε

2000 *Mathematics Subject Classification*: Primary 11L07; Secondary 11M06, 11P21, 11L15, 11L40, 11N99, 11L03.

being any positive absolute constant), then

$$(1.2) \quad I(t, U, \chi) \sim \frac{\phi(r)}{r} \left(\log \left(\frac{rt}{2\pi} \right) + 2\gamma + 2 \sum_{p|r} \frac{\log p}{p-1} \right) \quad (t \rightarrow \infty),$$

where p runs over the primes, γ is Euler’s constant and $\phi(r)$ is Euler’s function. In unpublished work [26] Motohashi extends his results in [25] to cover the case of composite r .

It is desirable to have information about $I(t, U, \chi)$ for U smaller than $(rt)^{1/3}$, since this can tell us more about the local behaviour of $L(1/2+it, \chi)$, or even about its size at an individual value of t . This is the motivation for our main result, which we now state.

THEOREM 1. *Let ε be an arbitrary positive constant. Let $t > 0$ and $r \ll t^{1/14}$, where either r is an odd prime, or $r = 1$. Suppose that*

$$(1.3) \quad t/2 \geq U \geq \max_{j=1, \dots, 5} U_j(r, t) \geq 1,$$

where

$$(1.4) \quad U_j(r, t) = \begin{cases} r^{51/146+\varepsilon} t^{23/73} (\log t)^{E'} & \text{if } j = 1, \\ r^{5743/12910} t^{2021/6455+\varepsilon} & \text{if } j = 2, \\ r^{1474/3151} t^{985/3151+\varepsilon} & \text{if } j = 3, \\ r^{788/1193} t^{365/1193+\varepsilon} & \text{if } j = 4, \\ r^{47/57} t^{17/57+\varepsilon} & \text{if } j = 5, \end{cases}$$

with E' sufficiently large (in absolute terms). Then, when χ is a primitive Dirichlet character modulo r ,

$$(1.5) \quad I(t, U, \chi) \ll \log t.$$

We make two initial observations. Firstly, the bound (1.5) is, in the context of Theorem 1, of the same order of magnitude as the right-hand side of (1.2). Secondly, when $r = t^{1/14}$, the lower bound on U in (1.3) is $r^{47/57} t^{17/57+\varepsilon} = r^{1/3} t^{1/3+\varepsilon}$, which shows that, outside of cases where $r \geq t^{1/14-o(1)}$, any result provided by Theorem 1 must penetrate to some extent beyond the reach of Motohashi’s aforementioned work. We are not the first to achieve this, since Dzhabbarov [5] essentially established that (1.2) holds, for $r \leq t^{1/21}$, when $U = r^{65/82} t^{27/82+\varepsilon}$ with ε positive and absolute. Theorem 1 is much stronger than this last result, and we know of no other previous work of this sort. There are, however, many more precedents to Theorem 1 for the special case $r = 1$, in which one has $\chi = \chi_0$, the (trivial) principal character, and $L(s, \chi) = \zeta(s)$, Riemann’s zeta-function. In this one case we write $I(t, U)$ for $I(t, U, \chi_0)$ and $I_1(T)$ for $I_1(T, \chi_0)$. Prior to Theorem 1, Huxley (in [17]) had sharpened the bound for the error-term in $I_1(T)$ ’s asymptotic expansion, succeeding to the extent that (1.2)

is implied for $U \geq t^{137/432}(\log t)^A$, some $A > 0$, $r = 1$ and $\chi = \chi_0$. Note that $137/432 > 0.3171$, while Theorem 1 provides the upper bound (1.5) for $t/2 \geq U \geq t^{23/73}(\log t)^A$ (some $A > 1$, $r = 1$ and $\chi = \chi_0$), where $23/73 < 0.3151$.

The number $23/73$ also occurs as an exponent in results of Huxley’s paper [11] on a class of lattice-point problems including the classical circle and divisor problems. It is no accident that it resurfaces in this paper, for Theorem 1 derives from bounds for certain exponential sums, $S(H, M)$, that are strongly analogous (despite two key differences) to the sums that had to be estimated in [11]. Moreover, this paper also follows [11] in establishing the crucial exponential sum estimates by an adaptation of Iwaniec and Mozzochi’s variant [21] of the Bombieri–Iwaniec method [1]. For the record, the sums $S(H, M)$ we shall need to consider have the shape

$$(1.6) \quad \sum_{H_1 < h \leq H} \sum_{M_1 < k \leq M} \chi\left(\frac{k+h}{k-h}\right) e\left(TF\left(\frac{k+h}{M}\right) - TF\left(\frac{k-h}{M}\right)\right),$$

where

$$(1.7) \quad H/2 \leq H_1 \leq H, \quad M/2 \leq M_1 \leq M,$$

$$(1.8) \quad H \leq M/U,$$

$$(1.9) \quad M \leq \sqrt{2rT},$$

$T \asymp t$, $F(X) = -\log X$, $e(Y) = \exp(2\pi iY)$ and the Dirichlet character factor is our notation for $\chi(k+h)\bar{\chi}(k-h)$. We shall see that the case $M \asymp \sqrt{rT}$ is a “worst case” when the maximum in (1.3) is attained at $j = 1$, or $j = 5$, but not when that maximum is attained at $j = 2$, $j = 3$, or $j = 4$. The greater part of our work, including the intermediate results most central to this paper, does apply to a fairly general class of functions $F(X)$. It is only in the penultimate stages (after Lemma 9.2) that we begin to assume $F(X) = -\log X$.

Apart from the Dirichlet character factors, the sum (1.6) is the same sum considered in the series of papers, [9], [16] and [17], on $I_1(T)$, whereas the sum one meets with in the circle, divisor, or lattice-point problems is

$$(1.10) \quad S^*(H, M) = \sum_{H_1 < h \leq H} \sum_{M_1 < k \leq M} e\left(\frac{hT}{M} F^*\left(\frac{k}{M}\right)\right)$$

(see [21], [10], [11] and [13]). As

$$F\left(\frac{k+h}{M}\right) - F\left(\frac{k-h}{M}\right) \sim \frac{2h}{M} F'\left(\frac{k}{M}\right) \quad (h = o(M)),$$

the analogy between (1.6) and (1.10) is clear: $S(H, M)$ is a bit like the sum $S^*(H, M)$ with $F^*(x) = 2F'(x)$. In practice, as was found in [9], this analogy is somewhat disrupted by third-order terms that have quite a marked

effect on the subsequent analysis of $S(H, M)$. Whereas, with $S^*(H, M)$, one ultimately needs to bound the number of integer solutions of a certain fairly simple system of equations and inequalities, \mathcal{S}^* (for which see [21] or [32]), it turns out that for $S(H, M)$ one has to consider a perturbed version \mathcal{S}^{**} of the system \mathcal{S}^* (see [14, Section 13.3]). Counting the integer solutions of \mathcal{S}^* , or of \mathcal{S}^{**} , is usually referred to as the “First Spacing Problem”. In [16] and [17] the perturbation in \mathcal{S}^{**} was handled trivially (treated as an extra error-term in the relevant inequality of the system). Consequently the error-term bounds of those papers involve larger exponents than the error-term bounds in [11] and [13], despite the analogy between (1.6) and (1.10). This phenomenon is not evident between [10] and [9] since these two papers were written before the advances made in [12], which led to the perturbation error-term becoming significant (as some parameters changed size relative to one another).

In Section 7 of this paper we make use of an alternative formulation of the perturbed First Spacing Problem. Indeed, it turns out that \mathcal{S}^{**} is equivalent (under a simple change of variables) to a system \mathcal{S} similar to that discussed in [2] and [31], but involving an extra constraint of the form $|g_i - h_i| \asymp L$ ($i = 1, \dots, 4$), where $g_1, h_1, \dots, g_4, h_4 \in [K, 2K]$ are the variables of the system, and $L = o(K)$ (see the discussion between Lemmas 7.10 and 7.11). Note that this means there is essentially the same analogy between the systems \mathcal{S} and \mathcal{S}^* as there is between the sums $S(H, M)$ and $S^*(H, M)$ in (1.6) and (1.10). The equivalence of \mathcal{S}^{**} and \mathcal{S} seems not to have been noticed before. Given the results of [31], the above new insight turns out to be enough, when $r = 1$, to essentially close any existing gap between our bound for the number of integer solutions of the perturbed system \mathcal{S}^{**} and the similar bound in respect of the simple system \mathcal{S}^* . This is why our exponent of t , in case $j = 1$ of (1.4) in Theorem 1, is the same as the exponent of T , or of M^2 , in [11, Theorems 4, 5]. In [33] we establish new results on the number of integer solutions of the system \mathcal{S} . These turn out to be useful for larger r , and are influential in making $U_3(r, t)$ of (1.4) smaller than it would otherwise be.

In applying the Bombieri–Iwaniec method to bound the sum $S(H, M)$ of (1.6) we take the same approach as in our earlier paper [20] with Huxley, and carry the Dirichlet characters through all stages of the calculation up to and including the First Spacing Problem (Section 7). At that point there arises naturally an opportunity to exploit the oscillations of the character, through the application of Weil’s bound for sums

$$\sum_{x \bmod r} \chi((x + a_1) \dots (x + a_n)) \overline{\chi}((x + b_1) \dots (x + b_n)) e(cx/r),$$

with $n = 4$ and $c = 0$. Weil’s bound for the case $n = 1$ (with any $c \in \mathbb{Z}$) is

used for the major arc estimates (Section 3) and also, in the case of minor arcs, for bounding error-terms in the first and second Poisson summation steps (Sections 4 and 5). It is these Poisson summation errors that give rise to the term $U_5(r, t)$ in (1.3)–(1.4) of Theorem 1. Despite the chronology [20] is in fact the sequel to this paper, which has undergone the longer period of gestation, and both were inspired by [19] in their handling of the Dirichlet character.

In bounding the order of $L(1/2 + it, \chi)$, Theorem 1 can be put to good use. We begin by observing that, for $t \geq (2 + \log r)^3$ and $\Delta = \log(rt/2\pi)$,

$$(1.11) \quad \left| L\left(\frac{1}{2} + it, \chi\right) \right|^2 \ll 1 + \Delta \int_{-\infty}^{\infty} \left| L\left(\frac{1}{2} + i(t + \eta), \chi\right) \right|^2 \operatorname{sinc}^2\left(\frac{\Delta}{\pi} \eta\right) d\eta.$$

This may be deduced from the bound

$$\begin{aligned} \left| \sum_{1 \leq m \leq e^\Delta} \frac{\chi(m)}{\sqrt{m}} m^{-it} \right|^2 &\leq \int_{-\infty}^{\infty} \left| \sum_{\max(0, x-1) \leq (\log m)/\Delta \leq \min(1, x+1)} \frac{\chi(m)}{\sqrt{m}} m^{-it} \right|^2 dx \\ &= \frac{2\Delta}{\pi} \int_{-\infty}^{\infty} \left| \sum_{1 \leq m \leq e^\Delta} \frac{\chi(m)}{\sqrt{m}} m^{-i(t+\eta)} \right|^2 \operatorname{sinc}^2\left(\frac{\Delta}{\pi} \eta\right) d\eta, \end{aligned}$$

with help from the Mean Value Theorem for Dirichlet polynomials [24], Rane’s results [27] on $L(s, \chi)$ ’s Approximate Functional Equation, and (for partial sums where $m^{i(t+\eta)}$ barely oscillates) Poisson summation and the bounds of the First and Second Derivative Tests for exponential integrals (see [14, Lemmas 5.4.4, 5.1.2, 5.1.3], the first of which is missing a factor $1/q$). Bounds such as (1.11) are not new: see, for example [6, Lemma 3].

Applying the uniform bound $\operatorname{sinc}^2(x) \leq 1$ for $|\eta| \leq U$, and the “convexity” bound $L(1/2 + i\tau, \chi) \ll (r(1 + |\tau|))^{1/4}$ otherwise, we deduce from (1.11) that, for $t \geq \max(U, (2 + \log r)^3)$ and $U > 0$,

$$(1.12) \quad |L(1/2 + it, \chi)|^2 \ll 1 + I(U, t, \chi)U \log(rt) + \sqrt{rt}/U.$$

As one always has $t/2 \geq U \geq (rt)^{1/4}$ in Theorem 1, it follows from the above that

$$L(1/2 + it, \chi) \ll U^{1/2} \log t,$$

if Theorem 1 applies and $t \geq 30^3$. From this we deduce (with no need of further proof):

THEOREM 2. *Let ε, r and t satisfy the hypotheses of Theorem 1. Then*

$$L(1/2 + it, \chi) \ll \max_{j=1, \dots, 5} \sqrt{U_j(r, t)},$$

where $U_j(r, t)$ is as in (1.4), with E' sufficiently large (in absolute terms).

This theorem is weaker than the results of [20] whenever the maximum is attained for $j \leq 2$. For $\alpha = (\log r)/\log t \geq 8213/287317 = 0.028585\dots$, however, Theorem 2 does yield an improvement. We sharpen [20, Theorem 1.1] by at least a factor $(rt)^{0.001}$ whenever $r^{15.3} \leq t \leq r^{29.5}$, and the improvement is by a factor greater than $(rt)^{0.0017}$ when $\alpha = 524/11155 = 1/21.288\dots$, which is the point at which certain Poisson summation errors first become significant (as α increases).

We follow [20] in presenting the combined results of Theorem 2 and [20, Theorem 1.1] in the form of a list of points $(\beta_j, E(j))$ for which

$$(1.13) \quad L(1/2 + it, \chi) \ll (rt)^{E(j)+\varepsilon} \quad \text{if } \log r = \beta_j \log(rt).$$

Points at which (1.13) is now established are:

$$\begin{aligned} (\beta_0, E(0)) &= \left(0, \frac{89}{570}\right) = (0, 0.15614\dots), \\ (\beta_1, E(1)) &= \left(\frac{2}{755}, \frac{1179}{7550}\right) = (0.00264\dots, 0.15615\dots), \\ (\beta_2, E(2)) &= \left(\frac{17333}{729920}, \frac{2296757}{14598400}\right) = (0.02374\dots, 0.15732\dots), \\ (\beta_3, E(3)) &= \left(\frac{8213}{295530}, \frac{31219}{197020}\right) = (0.0277907\dots, 0.158455\dots) \end{aligned}$$

from [20, Theorem 1.1], and

$$\begin{aligned} (\beta_4, E(4)) &= \left(\frac{245}{7348}, \frac{2335}{14696}\right) = (0.033342406\dots, 0.158886\dots), \\ (\beta_5, E(5)) &= \left(\frac{524}{11679}, \frac{1253}{7786}\right) = (0.0448668\dots, 0.160929\dots), \\ (\beta_6, E(6)) &= \left(\frac{1}{15}, \frac{1}{6}\right) = (0.06666\dots, 0.16666\dots) \end{aligned}$$

from Theorem 2. The advantage of this form of presentation is that if, say, $\beta_7 \in (\beta_{j-1}, \beta_j)$, with $j \in \{1, \dots, 6\}$, then (1.13) holds for $j = 7$ if the three points, $(\beta_{j-1}, E(j-1))$, $(\beta_7, E(7))$ and $(\beta_j, E(j))$, are collinear.

To put Theorem 2 in context, we now briefly survey something (not all) of what is known beyond the cramped range $r \ll t^{1/14}$. For $t \in \mathbb{R}$, $r \in \mathbb{N}$ and χ primitive modulo r there is Heath-Brown’s hybridization (for which see [3], [8]) of Burgess’ bound:

$$(1.14) \quad L(1/2 + it, \chi) \ll_{\varepsilon} (r(1 + |t|))^{3/16+\varepsilon} \quad (\varepsilon > 0).$$

For r prime, χ primitive modulo r , and $t \geq 2$, the asymptotic formula (1.2), in the range implied by Motohashi’s work [25], can be used in (1.12) to show,

$$(1.15) \quad L(1/2 + it, \chi) \ll ((rt)^{1/6} + r^{1/4})(\log(rt))^3,$$

which is (in such cases) essentially of the same strength as Heath-Brown’s other hybrid bound in [7] (Heath-Brown obtained stonger results for suitable composite r). More recently, Conrey and Iwaniec, in [4], have made advances in the theory of central values of modular forms twisted by a real primitive character χ modulo r . These include, in the non-holomorphic case, a strikingly elegant and powerful result about $L(s, \chi)$ on its critical line. This result is a kind of sixth-power moment (difficult to state here), from which it follows that, if N and ε are sufficiently small positive absolute constants, then

$$(1.16) \quad L(1/2 + it, \chi) \ll r^{1/6+\varepsilon}(1 + |t|)^{1/N},$$

for $t \in \mathbb{R}$, $r > 1$ odd, χ real and non-principal modulo r . No admissible value of N is given in [4], although it is stated that one should easily be computable.

It is worth discussing what prevents us from improving (1.15) outside of the range $t \gg r^{14}$. Note first that, in order to achieve a noticeable improvement over (1.15), we must have $U = o((rt)^{1/3})$ in (1.12), and in Theorem 1. As is clear from the proof of Theorem 1 given in Section 10, it follows from this requirement that we must (at the very least) show

$$(1.17) \quad S(H, M) \ll M$$

(where $S(H, M)$ is the sum of (1.6)), for some $M < \sqrt{2rT} \asymp \sqrt{rt}$ and some

$$(1.18) \quad H \asymp M/U > (rT)^{-1/3}M.$$

Unfortunately, even after Lemma 5.4, our bounds for the Poisson summation errors (see (5.3) and (5.1)) make a total

$$E_P \gg \frac{M}{N} \sqrt{rHN} = M \sqrt{\frac{rH}{N}},$$

when account is taken of all the minor arcs (which are of uniform length N). Therefore, in order to obtain (1.17), one is forced to accept the constraint

$$(1.19) \quad H \ll N/r,$$

which, after (1.18), implies

$$N \gg r^{2/3}T^{-1/3}M, \quad R \ll (rT)^{-1/3}M \ll H,$$

where R is the integer parameter given by (2.6).

To understand the constraint $t \gg r^{14}$, it will help to discuss the Second Spacing Problem, which involves counting pairs of minor arcs that are in some sense “coincident” (see Section 8 for details). There are $O((M/N)^2)$ pairs of minor arcs altogether, and coincidence conditions (8.2) and (8.6) ensure that, apart from the $O(M/N)$ pairs on the diagonal, $O(1)$ similar sets of pairs of the same cardinality, and some pairs for which the expression bounded by (8.2) vanishes, there are no more than $O(Z)$ other coincident

pairs, where

$$(1.20) \quad Z = r^{-2} \Delta_1 \Delta_2 (M/N)^2 \asymp \frac{R^6 M^2}{r^2 H^2 N^4 Q^2} \ll \frac{R^4 M^2}{H^2 N^4 r^2},$$

when the parameter Q is in the “worst” range, $Q \asymp R$. Using other conditions we can do better than (1.20) if $R = o(H)$. Yet (see Lemma 8.2), when we choose N to optimise the result in the Second Spacing Problem it turns out that, for $M > (rT)^{4/9}$ (at least), the bound for the total number of all coincident pairs of arcs is essentially the quantity $\beta(Q)$ of (8.17), which is itself $O(Z)$ for $Q \asymp R \ll H$. As this includes about M/N pairs on the diagonal, so it follows by (1.20) and (2.6) that

$$1 \ll \frac{Z}{M/N} \ll \frac{R^4 M}{H^2 N^3 r^2} \asymp \frac{M^7}{r^2 T^2 H^2 N^5}.$$

From this and (1.19) it follows that

$$H \ll r^{-1} T^{-2/7} M,$$

which contradicts (1.18) unless $T \gg r^{14}$. The conclusion we draw from this is that, in order to progress beyond $t \gg r^{14}$, one must adjust the optimal choice of N mentioned above, leading to a worse result in the Second Spacing Problem. This need not always conflict with the need to obtain (1.17), as it is the Poisson summation error (not the “main-term”) which is dominant for $t \ll r^{11155/524} = r^{21.288\dots}$.

Because of the diagonal contributions in both the First and Second Spacing Problems, the best bound one could possibly get is:

$$(S(H, M))^4 \ll r^3 M^3 N H^6 / R^6$$

(see Lemma 9.1). Therefore, in order to obtain (1.17), we require $r^3 N H^6 \ll M R^6$, which, by (2.6) and (1.19), implies

$$(1.21) \quad H \ll r^{-7/10} T^{-3/10} M.$$

As (1.18) and (1.21) together imply $T \gg r^{11}$, we conclude that any worthwhile extension of Theorem 1 to cases with $t \ll r^{11}$ would require either a solution for the problem of Poisson summation errors, or some radical alteration of the Bombieri–Iwaniec method. It may also be of interest to note that, if the constraint (1.19) could be relaxed to just $H = O(N)$, or (what is more likely) $H = o(N)$, then in place of (1.21) we would have $H/M \ll (rT)^{-3/10}$. Apart from any other considerations, one can show, by bounding below the mean-square contribution to $S(H, M)$ of a single arc (averaged over χ and T), that (1.17) is definitely only attainable if $N \gg H$, so that the constraint $H = O(N)$ is mandatory. Therefore the case $U \asymp (rt)^{3/10}$ of (1.5) is at or beyond the limit of what the Bombieri–Iwaniec method might achieve.

In [13] Huxley has obtained new results on lattice-point problems, in which the exponent $23/73$ is improved to $131/416 = 0.31490\dots$. This improvement can be attributed to a new iterative method for treating the Second Spacing Problem. The new method is also used in [15] and [17]. We have chosen not to use it here, on the grounds that the extent of the improvement would likely be small, and, as it is, the Dirichlet character factors lead to complications enough. The application of the iterative method to the case $r = 1$ is something worth returning to at a later date. Other improvements that we expect could be made to Theorem 1 are the removal of the constraint that r be prime, and the replacement of the bound (1.5) with a bound of $O(UT^{-1} \log T)$ for the error-term in $I_1(T, \chi)$'s asymptotic formula. There is also the possibility of lowering the lower bound (1.4) for U , provided that better solutions could be found for certain technical problems, such as the Second Spacing Problem, the First Spacing Problem for large r , and the problem of obtaining a useful uniform bound for just the single sum over k in (1.6), or for similar sums with $\chi(k+h)\overline{\chi}(k-h)$ replaced by $\chi(f(k))\overline{\chi}(g(k))$, where the polynomials f and g factorise completely over \mathbb{Z} . Progress on the last problem might lead to a better version of Lemma 9.3, and so improve cases $j = 2, 3, 4$ of (1.4).

We leave the proof of Theorem 1 until Section 10, where it can be deduced as a corollary of our results on the sum $S(H, M)$. The groundwork for these results is laid in Sections 2–8, which cover all the steps of the Bombieri–Iwaniec method, in its application to $S(H, M)$. The outcomes are worked out in Section 9, where they are combined with a lemma derived from [12, Theorems 1, 3] in order to get the needed results on $S(H, M)$ of (1.6) in the case $F(X) = -\log X$.

The author is pleased to acknowledge the profound influence of Prof. M. N. Huxley. He instigated our collaboration on [19], which opened the way to [20], and to the work presented in this paper. Moreover, his continuing interested correspondence, offering advice, encouragement and criticism, has been a great stimulus.

Notation and conventions. Throughout Sections 2–10, r shall denote either the number 1, or an odd prime number. We shall hardly ever need to make a distinction between the case where $r = 1$ and the cases where r is an odd prime. By χ we shall mean a Dirichlet character that is primitive to the modulus r . The associated Gauss sum is $\tau(\chi)$ (see (3.1)). The mean-value $I(t, U, \chi)$ and sum $S(H, M)$ will be as in (1.1) and (1.6)–(1.9). In the Introduction we have used α for $(\log r)/\log t$.

The Farey sequence of order q is $\mathbb{F}(q) = \bigcup_{n=1}^q \{x \mid nx \in \mathbb{Z}\}$. In expressions such as “ $\overline{a} \pmod{b}$ ”, or “ $e(c\overline{a}/b)$ ”, the \overline{a} denotes a solution x for $ax \equiv 1 \pmod{b}$ (it being implicit that $(a, b) = 1$); where this usage does not apply

we use \bar{y} to denote the complex-conjugate of any complex-valued y . The latter usage should be applied to $\bar{\chi}$ “modulo r ”, since the former usage only applies in the sense that $\bar{\chi}(a) = \chi(\bar{a})$ for $a\bar{a} \equiv 1 \pmod{r}$.

The Weil sum $W(x, y)$ is given by (3.3). Note that a term $\chi(f)$ will denote $\chi(x)\bar{\chi}(y)$, whenever f is given as the quotient x/y . The algebraic function $\varrho_l(\xi)$ is given by either of the equivalent definitions (5.7), (5.8).

By $a \mid b$ we mean that a divides b : when the letter “ p ” is used with this notation it denotes a prime. The “divisor” and “sum of divisor” functions are $d(n) = \sum_{d \mid n} 1$ and $\sigma(n) = \sum_{d \mid n} d$, respectively.

The greatest integer not exceeding x is $[x]$, and the distance from x to the nearest integer is $\|x\| = \min(\{x\}, \{-x\})$, where the fractional part, $\{x\} = x - [x]$, should be distinguished from the singleton set by context. We write $e(x)$, $\text{sinc}(x)$ and $\Lambda(x)$, respectively, for the functions $\exp(2\pi ix)$, $(\pi x)^{-1} \sin(\pi x)$ and $\max(1 - |x|, 0)$. For $z \in \mathbb{C}$ we use $\Re(z)$ to denote the real part of z . By $\log x$ we mean the natural logarithm.

The $o(x)$ and $O(x)$ notation is standard. We use $a \ll b$, $a \asymp b$ and $a \sim b$ to mean $a = O(b)$, $a \ll b \ll a$ and $a/b = 1 + o(1)$, respectively. The implicit constants in these notations depend (at most) upon parameters explicitly declared to be constant, or parameters appended to the notation as subscripts. We use $E, E', C_0, C_2, C_3, \dots, C^*, \varepsilon, \varepsilon_0$ and (sometimes) $\delta, \eta, \lambda_0, \Lambda_0$ to denote positive constants. The constants E and E' are supposed to be absolute, and, except in Theorem 2 (where E' should be greater by 2 than the E' of Theorem 1), one may take $E = 2^{6562}$ and $E' = 2^{6560} + 2$ throughout. For our proofs we assume that ε_0 is sufficiently small in terms of C_3 . Our ε 's need only be sufficiently small in absolute terms: unlike that of ε_0 , the value of ε may change from lemma to lemma. In Lemma 9.4 and in Section 10 (while finishing the proof of Theorem 1), we shall find it convenient to assume that $\varepsilon \geq 8E'\varepsilon_0$. Apart from these constraints, the positive constants ε and ε_0 are arbitrary. The constants C_2, \dots, C_5 depend on the choice of function $F(X)$ in (1.6). As a matter of convenience (for the proof of Lemma 8.3), we make the harmless assumption throughout our working that $C_3 \geq 2^{7/2}$. In the application of our work to $L(s, \chi)$ one has $F(X) = -\log X$, so that C_2, \dots, C_5 , and therefore also ε and ε_0 , may be chosen absolutely (with $C_3 = (2!)3^3 > 2^{7/2}$, for example).

2. The approximation step of the Bombieri–Iwaniec method. In the sum $S(H, M)$ of (1.6)–(1.9) we write

$$(2.1) \quad TF\left(\frac{x+y}{M}\right) - TF\left(\frac{x-y}{M}\right) = yf(x, y),$$

where $F(X)$ is now taken to be an arbitrary real function, five times con-

tinuously differentiable and such that, for $1/3 \leq X \leq 3$,

$$(2.2) \quad |F^{(j)}(X)| \leq C_j \quad (j = 2, 3, 4, 5),$$

$$(2.3) \quad |F^{(j)}(X)| \geq 1/C_j \quad (j = 2, 3, 4).$$

Either by considering $\overline{S(H, M)}$ in place of $S(H, M)$, or by substituting

$$k_1 = \left[\frac{M + M_1}{r} \right] r - k, \quad F_1(X) = -F \left(\left[\frac{M + M_1}{r} \right] \frac{r}{M} - X \right),$$

or by a combination of both tactics, one can see that, provided

$$(2.4) \quad M \geq 4r,$$

it is enough to consider only those cases where both $F^{(2)}(X)$ and $F^{(3)}(X)$ are positive-valued. Then, by (2.1), we have

$$(2.5) \quad \frac{\partial^i}{\partial x^i} f(x, 0) = \frac{2T}{M^{i+1}} F^{(i+1)} \left(\frac{x}{M} \right) \asymp \frac{T}{M^{i+1}},$$

for $M/3 < x < 3M$ and $i = 1, 2, 3$.

As in Section 9.1 of [14], we suppose it possible to choose integers N and R with $2 \leq R \leq H \leq N \leq M$ and

$$(2.6) \quad 2(R - 1)^2 NT < C_3 M^3 \leq 2R^2 NT,$$

$$(2.7) \quad 4C_3 \sqrt{H} \leq R \leq \varepsilon_0 H,$$

$$(2.8) \quad 64C_3 H \leq N \leq M/36,$$

$$(2.9) \quad HN^3 T \leq M^4.$$

Following the original approach of Iwaniec and Mozzochi [21], the interval $[M_1 - 2N, M]$ is covered by a minimal set of consecutive disjoint intervals I_0, I_1, \dots, I_l , each of length N . To each interval I_i there corresponds a short arc J_i , which is the interval that $\frac{\partial f}{\partial x}(x, 0)$ runs over as x runs over I_i . By (2.2), (2.3) and (2.5)–(2.7),

$$(2.10) \quad 1/R^2 \leq |J_i| < (2C_3/R)^2 < 1.$$

Given i , let b/s be the (unique) rational of least denominator lying in J_i . Then J_i is a minor arc if and only if

$$(2.11) \quad \frac{R^2}{12C_3 Hr} < s \leq 24C_3 Hr.$$

LEMMA 2.1. *If $s > 24C_3 Hr$, then the midpoint, g , of J_i satisfies*

$$\left| g - \frac{c}{d} \right| \leq \frac{1}{24C_3 r H d},$$

for some reduced rational c/d with

$$(2.12) \quad 1 \leq d \leq \frac{R^2}{12C_3 Hr}.$$

Proof. This is an easy consequence of Dirichlet’s approximation theorem (see [14, p. 169]).

To clarify matters we observe here that, if $12C_3Hr > R^2$, then there can be no reduced rational c/d satisfying (2.12), and so the upper bound of (2.11) will hold for all arcs J_i . More trivially, as $s \geq 1$, it follows that the lower bound of (2.11) will also hold for all J_i if $12C_3Hr > R^2$. We conclude that when $12C_3Hr > R^2$ all arcs are minor arcs.

The rational c/d in Lemma 2.1 will be a value of $\frac{\partial f}{\partial x}(x, 0)$ for some $x \in [M/3, 3M]$. Given such a rational c/d , we fuse all the short arcs J_i with midpoint g satisfying

$$(2.13) \quad \left| g - \frac{c}{d} \right| \leq \frac{C_3}{rHd},$$

so that together they become one long major arc $J(c/d)$. The union of the corresponding intervals I_i will be $I(c/d)$. Note that, by virtue of (2.7), major arcs are pairwise disjoint. A minor arc that becomes part of a major arc in this way is still a minor arc in its own right.

The inequality (2.13) has been made weaker than that in Lemma 2.1. By (2.10), this ensures that if J_i is not now part of a major arc, then either $i \leq 1$, or both J_i and J_{i-2} are minor arcs. In the latter case we look again at the rationals lying in J_{i-2} :

Suppose that b/s is the reduced rational of least denominator lying in the minor arc J_{i-2} . In typical applications of the Bombieri–Iwaniec method one uses (in place of b/s) the reduced rational b_1/s_1 of least denominator $s_1 \geq R$ lying in J_{i-2} (see “Modification 1” [14, p. 168]). Here either $b_1/s_1 = b/s$, or (from (2.10)) $s_1 \geq R^2/4C_3^2s$. That is,

$$(2.14) \quad s_1 \geq \max(R, s, R^2/4C_3^2s).$$

However, for our application of the Bombieri–Iwaniec method it is advantageous to have a rational whose denominator is coprime to r .

LEMMA 2.2. *Let u be a positive integer. Then there is a rational a/q in J_{i-2} with $(q, u) = 1$ and*

$$R \leq q \leq (\xi + 2) \max(s, 2R^2/s),$$

where

$$\xi = \xi(u) = \prod_{p|u} \frac{2p}{p-1}.$$

Furthermore, one of the neighbours of a/q in $\mathbb{F}(q)$ will be b/s .

Proof. See [19, Lemma 2.3].

Since $\xi(1) = 1$ and $\xi(r) = 2r/(r-1) \leq 3$, for prime $r \geq 3$, this lemma and (2.14) show that we can find a/q in J_{i-2} with $(q, r) = 1$, $|as - qb| = 1$

and

$$(2.15) \quad s_1 \leq q \leq 40C_3^2 s_1.$$

By (2.11), the above lemma also shows that

$$(2.16) \quad R \leq q \leq 120C_3 H r.$$

Let $J(c/d)$ be a major arc. The corresponding portion of the sum $S(H, M)$ is

$$(2.17) \quad \mathcal{B}(c/d) = \sum_{H_1 < h \leq H} \mathcal{A}(c/d; h),$$

where

$$(2.18) \quad \begin{aligned} \mathcal{A}(c/d; h) &= \sum_{\substack{M_1 < k \leq M \\ k \in I(c/d)}} \chi\left(\frac{k+h}{k-h}\right) e(hf(k, h)) \\ &= \sum_{A \leq n \leq B} \chi\left(\frac{m+n+h}{m+n-h}\right) e(hf(m+n, h)), \end{aligned}$$

say, with m being a nearest integer to the solution of $\frac{\partial f}{\partial x}(x, 0) = c/d$,

$$(2.19) \quad \left| \frac{\partial f}{\partial x}(m, 0) - \frac{c}{d} \right| \leq \frac{2C_3^2}{R^2 N},$$

and with

$$(2.20) \quad \max(|A|, |B|) \leq \frac{2C_3 R^2 N}{r H d} > 24C_3^2 N.$$

Let J_i be a minor arc not belonging to any major arc. If $i \geq 2$, then there is a corresponding portion of the sum $S(H, M)$ in (1.6):

$$(2.21) \quad \mathcal{B}_i = \sum_{H_1 < h \leq H} \mathcal{A}_i(h),$$

where

$$(2.22) \quad \begin{aligned} \mathcal{A}_i(h) &= \sum_{\substack{M_1 < k \leq M \\ k \in I_i}} \chi\left(\frac{k+h}{k-h}\right) e(hf(k, h)) \\ &= \sum_{N_1 < n \leq N_0} \chi\left(\frac{m+n+h}{m+n-h}\right) e(hf(m+n, h)), \end{aligned}$$

say, with m being a nearest integer to the solution of $\frac{\partial f}{\partial x}(x, 0) = a/q \in J_{i-2}$ (so that (2.19) holds with a/q in place of c/d), and with

$$(2.23) \quad N \leq N_1 \leq N_0 \leq 3N, \quad N_0 - N_1 \leq N.$$

To analyse the sums (2.18) and (2.22), we use:

LEMMA 2.3. *Suppose that $M/3 \leq m \leq 3M$, $M/2 \leq m + u \leq M$ and $0 \leq v \leq H$. Put $\Phi_i = \Phi_i(m)$, where*

$$\Phi_i(x) = \frac{\partial^i f}{\partial x^i}(x, 0) = \frac{2T}{M^{i+1}} F^{(i+1)}\left(\frac{x}{M}\right) \quad (i = 0, 1, 2, 3, 4).$$

Then there exist $w_1, \dots, w_4 \in [\min(0, u - v), \max(0, u + v)]$ such that:

$$\begin{aligned} f(m + u, v) &= \Phi_0 + u\Phi_1 + \left(\frac{1}{2}u^2 + \frac{1}{6}v^2\right)\Phi_2 + \left(\frac{1}{6}u^3 + \frac{1}{6}uv^2\right)\Phi_3 \\ &\quad + \left(\frac{1}{24}u^4 + \frac{1}{12}u^2v^2 + \frac{1}{120}v^4\right)\Phi_4(m + w_1), \\ \frac{\partial f}{\partial x}(m + u, v) &= \Phi_1 + u\Phi_2 + \left(\frac{1}{2}u^2 + \frac{1}{6}v^2\right)\Phi_3(m + w_2), \\ \frac{\partial f}{\partial y}(m + u, v) &= \frac{1}{3}v\Phi_2 + \frac{1}{3}uv\Phi_3 + \left(\frac{1}{6}u^2v + \frac{1}{30}v^3\right)\Phi_4(m + w_3), \\ \frac{\partial^2 f}{\partial x \partial y}(m + u, v) &= \frac{1}{3}v\Phi_3(m + w_4). \end{aligned}$$

Proof. This follows by use of elementary calculus, as in [9, Section 3].

For the major arc $J(c/d)$ we simply write

$$(2.24) \quad \mathcal{A}(c/d; h) = \sum_{A \leq n \leq B} \chi\left(\frac{m + n + h}{m + n - h}\right) e\left(\frac{hcn}{d} + g(n)\right),$$

with

$$g(u) = hf(m + u, h) - hcu/d.$$

By Lemma 2.3, (2.2) and (2.3) (and since $F^{(3)}(X) > 0$),

$$(2.25) \quad \begin{aligned} g^{(1)}(u) &= h(\Phi_1 - c/d + u\Phi_2) + O((u^2 + h^2)HTM^{-4}), \\ g^{(2)}(u) &= h \frac{\partial^2 f}{\partial x^2}(m + u, u) \asymp HTM^{-3} \asymp H/NR^2. \end{aligned}$$

By (2.19), (2.2) and then (2.6), (2.8), it follows that

$$\begin{aligned} g^{(1)}(u) &\ll H \left(\frac{1}{R^2N} + |u| \frac{T}{M^3} + (u^2 + H^2) \frac{T}{M^4} \right) \\ &\ll H \left(\frac{|u|}{R^2N} \left(1 + \frac{|u|}{M} \right) + \frac{1}{R^2} \right). \end{aligned}$$

For $A \leq u \leq B$, (2.20), (2.6), (2.12) and (1.9) show

$$\frac{|u|}{R^2N} \ll \frac{1}{rHd} \gg \frac{1}{R^2}, \quad \frac{|u|}{M} \ll \frac{R^2N}{rHM} \asymp \frac{M^2}{TrH} \ll \frac{1}{H} \ll 1.$$

Therefore

$$(2.26) \quad g^{(1)}(u) \ll 1/rd \quad (A \leq u \leq B).$$

We rewrite the minor arc sum $\mathcal{A}_i(h)$ in (2.22) as

$$(2.27) \quad \sum_{N_1 < n \leq N_0} \chi\left(\frac{m+n+h}{m+n-h}\right) \times e\left(hf(m,0) + \frac{han}{q} + h\mu n^2 + \frac{1}{3}h^3\mu + g(n,h)\right),$$

where

$$\mu = \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(m,0)$$

and (in the notation of Lemma 2.3)

$$\begin{aligned} g(u,v) &= vf(m+u,v) - vf(m,0) - uva/q - u^2v\mu - v^3\mu/3 \\ &= vf(m+u,v) - v\Phi_0 - uva/q - (u^2v/2 + v^3/6)\Phi_2. \end{aligned}$$

Note that (2.5), (2.2), (2.3) and (2.6) imply

$$(2.28) \quad 1/2NR^2 \leq \mu < 2C_3^2/NR^2.$$

By Lemma 2.3 and (2.2),

$$\begin{aligned} \frac{\partial g}{\partial u} &= v\left(\frac{\partial f}{\partial x}(m+u,v) - \frac{a}{q} - u\Phi_2\right) \\ &= v\left(\Phi_1 - \frac{a}{q} + O\left((u^2+v^2)\frac{T}{M^4}\right)\right), \\ \frac{\partial g}{\partial v} &= f(m+u,v) + v\frac{\partial f}{\partial y}(m+u,v) - \Phi_0 - u\frac{a}{q} - \frac{1}{2}(u^2+v^2)\Phi_2 \\ &= u\left(\Phi_1 - \frac{a}{q}\right) + \left(\frac{1}{6}u^3 + \frac{1}{2}uv^2\right)\Phi_3 + O\left((u^4+v^4)\frac{T}{M^5}\right), \\ \frac{\partial^2 g}{\partial u\partial v} &= \frac{1}{v}\frac{\partial g}{\partial u} + v\frac{\partial^2 f}{\partial x\partial y}(m+u,v) = \frac{1}{v}\frac{\partial g}{\partial u} + O\left(v^2\frac{T}{M^4}\right). \end{aligned}$$

Therefore, for $N_1 \leq u \leq N_0$ and $H_1 \leq v \leq H$, (2.23), (2.19), (2.2) and (2.7)–(2.9) show:

$$\begin{aligned} \frac{\partial g}{\partial u} &\ll H\left(\frac{1}{R^2N} + \frac{N^2T}{M^4}\right) \ll \frac{1}{N}, \\ \frac{\partial g}{\partial v} &\ll \frac{1}{R^2} + \frac{N^3T}{M^4} + \frac{N^4T}{M^5} \ll \frac{1}{H}, \\ \frac{\partial^2 g}{\partial u\partial v} &\ll \frac{1}{HN} + \frac{H^2T}{M^4} \ll \frac{1}{HN}. \end{aligned}$$

From this and (2.27), it follows by partial summation [14, Lemma 5.2.2], that, for some choice of integers H_2, H_3, N_2, N_3 , satisfying $N_3 - N_2 \leq N$,

$$(2.29) \quad H/2 \leq H_2 < H_3 \leq H, \quad N \leq N_2 < N_3 \leq 3N,$$

one has

$$(2.30) \quad \mathcal{B}_i = \sum_{H_1 < h \leq H} \mathcal{A}_i(h) \ll |\mathcal{B}'_i|,$$

where \mathcal{B}'_i is the sum

$$(2.31) \quad \sum_{H_2 \leq h \leq H_3} \sum_{N_2 \leq n \leq N_3} \chi\left(\frac{m+n+h}{m+n-h}\right) \times e\left(hf(m,0) + \frac{han}{q} + h\mu n^2 + \frac{1}{3}h^3\mu\right).$$

In anticipation of the Bombieri–Iwaniec Double Large Sieve (Section 6), we seek to introduce at this stage a little variation into the value of the coefficient of h in the exponential factor of \mathcal{B}'_i above. To this end we observe that, since $\text{sinc}^2(x) = \pi^{-2}x^{-2} \sin^2(\pi x)$ is decreasing, for $0 \leq x \leq 1/2$, and is (consequently) bounded below there by $4\pi^{-2}$, we may use partial summation to bound \mathcal{B}'_i in terms of a similar sum, which may have h running over only a subinterval of its original range, but otherwise differs only in having $\text{sinc}^2(h/2H)$ as a factor in each of its terms. As $\text{sinc}^2(x)$ is the Fourier transform of $\Lambda(x) = \max(0, 1 - |x|)$, we are able to replace (2.30) with the bound

$$(2.32) \quad \mathcal{B}_i = \sum_{H_1 < h \leq H} \mathcal{A}_i(h) \ll \int_{-\infty}^{\infty} \Lambda(2\eta) |\mathcal{B}'_{i,\eta}| d\eta,$$

where $\mathcal{B}'_{i,\eta}$ has the form of the sum \mathcal{B}'_i in (2.30)–(2.31), but with $f(m, 0)$ replaced by

$$(2.33) \quad f_\eta(m, 0) = f(m, 0) + \eta/H,$$

and with a possible change in the values of H_2, H_3 , which values nevertheless remain independent of η and continue to satisfy (2.29).

We shall use (2.30) for the minor arcs with $q \gg H$ and (2.32) for the remainder (with $R \ll q \ll H$). The treatment of the former case is completed early in Section 5, while the latter case requires a further three sections. Our analysis of minor arcs deals with $\mathcal{B}'_{i,\eta}$ throughout, since \mathcal{B}'_i may be regarded as the special case $\eta = 0$ of this sum (the particular values of H_2, H_3 in (2.29) having no bearing on our arguments). We shall find it convenient to write

$$(2.34) \quad qf_\eta(m, 0) = b_0 + \kappa_0,$$

where

$$(2.35) \quad b_0 = b_{0,i} = [qf(m, 0)], \quad \kappa_0 = \kappa_{0,i}(\eta) = \{qf(m, 0)\} + q\eta/H.$$

We shall never need to consider cases where $|q\eta| > H$, so it may be assumed throughout that

$$(2.36) \quad |\kappa_0| < 2.$$

3. Major arcs. As χ is primitive modulo r ,

$$(3.1) \quad \bar{\chi}(n) = \frac{1}{\tau(\chi)} \sum_{v \pmod r} \chi(v)e(vn/r)$$

for all integers n . Since r is odd, it follows that, for $k \in \mathbb{Z}$,

$$(3.2) \quad \begin{aligned} \chi\left(\frac{k+h}{k-h}\right) &= \frac{1}{|\tau(\chi)|^2} \sum_{u \pmod r} \sum_{v \pmod r} \chi\left(\frac{v}{u}\right) e\left(\frac{(v-u)k - (v+u)h}{r}\right) \\ &= \frac{1}{r} \sum_{u \pmod r} \sum_{v \pmod r} \chi\left(\frac{u+v}{u-v}\right) e\left(\frac{vk - uh}{r}\right) \\ &= \frac{1}{r} \sum_{v \pmod r} W(v, -h) e\left(\frac{kv}{r}\right), \end{aligned}$$

where

$$(3.3) \quad W(x, y) = \sum_{u \pmod r} \chi\left(\frac{u+x}{u-x}\right) e\left(\frac{y}{r}u\right).$$

By [28, Theorems 2G, 2C and Lemma 2C],

$$(3.4) \quad |W(x, y)| \leq \begin{cases} r & \text{if } x \equiv y \equiv 0 \pmod r, \\ 2\sqrt{r} & \text{otherwise.} \end{cases}$$

Using (3.2), with $k = m + n$, we rewrite (2.24) as

$$(3.5) \quad \mathcal{A}(c/d; h) = \frac{1}{r} \sum_{v \pmod r} W(v, -h) \mathcal{A}(c/d; h, v) e\left(\frac{mv}{r}\right),$$

where

$$\mathcal{A}(c/d; h, v) = \sum_{A \leq n \leq B} e\left(\left(\frac{hc}{d} + \frac{v}{r}\right)n + g(n)\right).$$

By [14, Lemma 5.4.3],

$$(3.6) \quad \mathcal{A}(c/d; h, v) = \sum_{\substack{(\alpha-1/4)rd \leq l \leq (\beta+1/4)rd \\ l \equiv -(rhc+dv) \pmod{rd}}} \int_A^B e\left(g(u) - \frac{lu}{rd}\right) du + O(\log(\beta - \alpha + 2)),$$

where $\alpha = g^{(1)}(A)$ and $\beta = g^{(1)}(B)$ (note that (2.25) implies that $g^{(1)}(u)$ is both increasing and continuous). By (2.25) and (2.20),

$$(3.7) \quad \beta - \alpha \ll (B - A)H/NR^2 \ll 1/rd.$$

Using the Second Derivative Test [14, Lemma 5.1.3] with (2.25), one finds

$$E_l = \left| \int_A^B e\left(g(u) - \frac{lu}{rd}\right) du \right| \ll R\sqrt{\frac{N}{H}}.$$

Furthermore, as (2.26) implies that

$$g^{(1)}(u) - l/rd = -(l + O(1))/rd \quad (A \leq u \leq B),$$

it also follows from the First Derivative Test [14, Lemma 5.1.2] that

$$(3.8) \quad E_l \ll \sum_{|\lambda| \leq A_0} \min\left(R\sqrt{\frac{H}{N}}, \frac{rd}{|\lambda - l|}\right),$$

for some positive absolute constant A_0 .

By (3.4)–(3.6) we may write:

$$(3.9) \quad \mathcal{A}(c/d; h) = \mathcal{A}_0(c/d; h) + \mathcal{A}_1(c/d; h/r) + O(\sqrt{r}),$$

where

$$|\mathcal{A}_0(c/d; h)| \leq \sum_{\substack{l \equiv -rhc \pmod{d} \\ \alpha - 1/4 \leq l/rd \leq \beta + 1/4}} E_l \frac{1}{r} \sum_{\substack{v \pmod{r} \\ v \equiv -(rhc+l)/d \pmod{r}}} 2\sqrt{r}$$

and

$$|\mathcal{A}_1(c/d; k)| \leq \begin{cases} \sum_{\substack{l \equiv -r^2kc \pmod{rd} \\ \alpha - 1/4 \leq l/rd \leq \beta + 1/4}} E_l & \text{if } k \in \mathbb{Z}, \\ 0 & \text{otherwise.} \end{cases}$$

In the bound for $\mathcal{A}_0(c/d; h)$ one has, for some integer j and any $\lambda \in \mathbb{R}$,

$$l = -rhc + jd, \quad \frac{\lambda}{d} - \frac{l}{d} = \frac{rhc}{d} + \frac{\lambda}{d} - j.$$

Therefore

$$\left| \frac{\lambda}{d} - \frac{l}{d} \right| \geq \left\| \frac{rhc}{d} + \frac{\lambda}{d} \right\| \quad (\lambda \in \mathbb{R})$$

and, by (3.8) and (3.7),

$$\begin{aligned} \mathcal{A}_0(c/d; h) &\ll \sqrt{r} \sum_{|\lambda| \leq A_0} \sum_{\substack{l \equiv -rhc \pmod{d} \\ r\alpha - r/4 \leq l/d \leq r\beta + r/4}} \min\left(\frac{R}{r} \sqrt{\frac{H}{N}}, \left| \frac{\lambda}{d} - \frac{l}{d} \right|^{-1}\right) \\ &\ll \sqrt{r} \sum_{|\lambda| \leq A_0} \left(\min\left(\frac{R}{r} \sqrt{\frac{H}{N}}, \left\| \frac{rhc}{d} + \frac{\lambda}{d} \right\|^{-1}\right) + \log r \right). \end{aligned}$$

Estimating $\mathcal{A}_1(c/d; h)$ similarly, and applying (3.9), we deduce that $\mathcal{B}(c/d)$ (the major arc sum in (2.17)) satisfies a formula of the shape

$$(3.10) \quad \mathcal{B}(c/d) = \sqrt{r} \mathcal{B}_0(c/d) + \mathcal{B}_1(c/d) + O(H\sqrt{r} \log r),$$

where, for some integer $\lambda \in [-\Lambda_0, \Lambda_0]$,

$$(3.11) \quad \mathcal{B}_0(c/d) \ll \sum_{H_1 < h \leq H} \min\left(\frac{R}{r} \sqrt{\frac{N}{H}}, \left\| \frac{rch}{d} + \frac{\lambda}{d} \right\|^{-1}\right),$$

$$(3.12) \quad \mathcal{B}_1(c/d) \ll \sum_{H_1/r < k \leq H/r} \min\left(R \sqrt{\frac{N}{H}}, \left\| \frac{rck}{d} + \frac{\lambda}{rd} \right\|^{-1}\right).$$

LEMMA 3.1. *Suppose that $X \geq 1, Y \geq 1, (a', q') = 1, |\theta - a'/q'| \leq (q')^{-2}$. Then*

$$\sum_{1 \leq x \leq X} \min(Y, \|\theta x\|^{-1}) \ll XY/q' + (X + q') \log(2q'),$$

and, for real ϕ ,

$$\sum_{1 \leq x \leq X} \min(Y, \|\theta x - \phi\|^{-1}) \ll XY/q' + Y + (X + q') \log(2q').$$

Proof. This is due originally to Vinogradov (see [30, Lemma 1] for a reference).

By (3.10)–(3.12) and the second part of the lemma above,

$$\mathcal{B}(c/d) \ll \frac{R}{\sqrt{r}} \sqrt{HN} \frac{(r, d)}{d} + R \sqrt{\frac{N}{H}} + \sqrt{r} (H + d) \log(2d) = G_d \quad (\text{say}).$$

Hence, and by (2.12), the total contribution $S_\infty(H, M)$, from all the major arcs to $S(H, M)$ in (1.6), satisfies a bound:

$$S_\infty(H, M) \ll \sum_{1 \leq d \ll R^2/Hr} \sum_{c \asymp dT/M^2} G_d \ll \frac{M}{NR^2} \sum_{1 \leq d \ll R^2/Hr} dG_d.$$

As $2 \leq R < H < N$ (see (2.7)–(2.8)), this gives

$$(3.13) \quad \begin{aligned} S_\infty(H, M) &\ll \frac{M}{NR^2} \sum_{1 \leq d \ll R^2/Hr} \frac{R}{\sqrt{r}} \sqrt{HN} (r, d) \log R \\ &\ll \frac{M}{NR^2} \frac{R}{\sqrt{r}} \sqrt{HN} \frac{R^2}{Hr} d(r) \log R \ll \frac{MR}{\sqrt{r^3 HN}} \log R. \end{aligned}$$

4. Minor arcs: the first Poisson summation. From here on until Section 9 we shall assume that

$$(4.1) \quad 240C_3 Hr \geq 2Q \geq q \geq Q \in \{R, 2R, 2^2R, \dots\}, \quad (q, r) = 1$$

(see (2.15)–(2.16)). After (2.31)–(2.33) and (3.1)–(3.2), we may write

$$(4.2) \quad \mathcal{B}'_{i,\eta} = \frac{1}{r} \sum_{u \bmod r} \sum_{v \bmod r} \chi\left(\frac{u+v}{u-v}\right) \times \sum_{H_2 \leq h \leq H_3} \omega_h e\left(\frac{vm - uh}{r}\right) \mathcal{A}_i(h, v),$$

where

$$\omega_h = e(hf_\eta(m, 0) + \frac{1}{3}h^3\mu)$$

and

$$(4.3) \quad \mathcal{A}_i(h, v) = \sum_{N_2 \leq n \leq N_3} e\left(\left(\frac{v}{r} + \frac{ha}{q}\right)n + h\mu n^2\right).$$

LEMMA 4.1. Let $\mathcal{E}_i(h, v) = \mathcal{A}_i(h, v) - \mathcal{A}_i^*(h, v)$, where

$$\mathcal{A}_i^*(h, v) = \sqrt{\frac{1}{2\mu h}} \sum_{\substack{2\mu r q h N_2 < l < 2\mu r q h N_3 \\ l \equiv -(har + vq) \pmod{rq}}} e\left(\frac{1}{8} - \frac{l^2 h^{-1}}{4\mu r^2 q^2}\right).$$

Then

$$\mathcal{E}_i(h, v) \ll 1 + \sum_{j=2}^3 \min\left(\frac{1}{\sqrt{\mu h}}, \left\| \frac{ha}{q} + \frac{v}{r} + 2\mu h N_j \right\|^{-1}\right).$$

Proof. By (4.3) and [14, Lemma 5.4.3],

$$(4.4) \quad \mathcal{A}_i(h, v) = \sum_{\substack{r q (\alpha_2 - 1/4) \leq l \leq r q (\alpha_3 + 1/4) \\ l \equiv -(q v + r h a) \pmod{r q}}} \int_{N_2}^{N_3} e\left(h\mu x^2 - \frac{l x}{r q}\right) dx + O(\log(\alpha_3 - \alpha_2 + 2)),$$

where $\alpha_j = 2h\mu N_j$. By (2.29), (2.28) and (2.7),

$$\alpha_3 - \alpha_2 = 2h(N_3 - N_2)\mu < 4C_3^2 H/R^2 \leq 1/4,$$

so that, on the right of (4.4), the O -term is $O(1)$ and the sum has at most one term.

By writing

$$h\mu x^2 - \frac{l x}{r q} = h\mu \left(x^2 - 2x \frac{l/h}{2\mu r q}\right) = \beta(x^2 - 2x x_{l/h}) \quad (\text{say}),$$

the factor $e(-\beta x_{l/h}^2)$ may be taken out of the integral in (4.4), leaving just:

$$\int_{N_2 - x_{l/h}}^{N_3 - x_{l/h}} e(\beta x^2) dx = \frac{\delta}{\sqrt{2\beta}} e(1/8) + O\left(\min\left(\frac{1}{\sqrt{\beta}}, \sum_{j=2}^3 \frac{1/\beta}{|N_j - x_{l/h}|}\right)\right),$$

with $\delta = 1$ for $N_2 < x_{l/h} < N_3$, and $\delta = 0$ otherwise (a straightforward application of the First and Second Derivative Tests [14, Lemmas 5.1.2 and

5.1.3], together with the well known evaluation of the integral from minus to plus infinity). To complete the proof observe that, when l is as in (4.4), one has

$$2\beta|N_j - x_{i/h}| = \left| 2\mu h N_j - \frac{l}{rq} \right| = \left| \left(2\mu h N_j + \frac{v}{r} + \frac{ha}{q} \right) - k \right|$$

for some integer k .

By Lemma 4.1, (4.2)–(4.3) and (3.3),

$$(4.5) \quad \mathcal{B}'_{i,\eta} = \mathcal{B}^*_{i,\eta} + \mathcal{E}_{i,\eta},$$

where

$$(4.6) \quad \begin{aligned} \mathcal{B}^*_{i,\eta} &= \frac{1}{r} \sum_{u \bmod r} \sum_{v \bmod r} \chi\left(\frac{u+v}{u-v}\right) \\ &\quad \times \sum_{H_2 \leq h \leq H_3} \omega_h e\left(\frac{vm - uh}{r}\right) \mathcal{A}_i^*(h, v) \end{aligned}$$

and

$$(4.7) \quad \begin{aligned} \mathcal{E}_{i,\eta} &= \frac{1}{r} \sum_{v \bmod r} e\left(\frac{vm}{r}\right) \sum_{H_2 \leq h \leq H_3} \omega_h W(v, -h) \mathcal{E}_i(h, v) \\ &= O(D_i) + O(E_i(N_2) + E_i(N_3)), \end{aligned}$$

with

$$(4.8) \quad D_i = \frac{1}{r} \sum_{v \bmod r} \sum_{H_2 \leq h \leq H_3} |W(v, -h)|,$$

and $E_i(N_j)$ equal to

$$(4.9) \quad \frac{1}{r} \sum_{v \bmod r} \sum_{H_2 \leq h \leq H_3} |W(v, -h)| \min\left(\frac{1}{\sqrt{\mu h}}, \left\| \frac{ha}{q} + \frac{v}{r} + 2\mu h N_j \right\|^{-1}\right).$$

In the sum $\mathcal{A}_i^*(h, v)$ of Lemma 4.1 the congruence condition on l is equivalent to two simultaneous conditions:

$$l + qv \equiv 0 \pmod{r}, \quad l + rah \equiv 0 \pmod{q}.$$

As $(q, ra) = 1$, this allows us to bring the sums over v and h of (4.6) inside the summation over l of $\mathcal{A}_i^*(h, v)$, and then to write $v \equiv -\bar{q}l \pmod{r}$,

$$\chi\left(\frac{u+v}{u-v}\right) = \chi\left(\frac{qu-l}{qu+l}\right), \quad e\left(\frac{vm}{r}\right) = e\left(-\frac{\bar{q}lm}{r}\right),$$

where \bar{q} and $\bar{r}a$ are integers satisfying

$$(4.10) \quad \bar{q}q + \bar{r}ara = 1.$$

Therefore, making the substitution

$$u \equiv -\bar{q}s \pmod{r} \quad (0 \leq s \leq r - 1),$$

we find that, after recalling the definitions made in (4.2)–(4.3) and Lemma 4.1, one has

$$\begin{aligned} (4.11) \quad \mathcal{B}_{i,\eta}^* &= \frac{e(1/8)}{r} \sum_{s=0}^{r-1} \sum_{L_2 < l < L_3} \chi\left(\frac{s+l}{s-l}\right) e\left(-\frac{\bar{q}ml}{r}\right) \\ &\times \sum_{\substack{H_4(l) \leq h \leq H_5(l) \\ h \equiv -\bar{r}al \pmod{q}}} \sqrt{\frac{1}{2\mu h}} e\left(\left(\frac{\bar{q}s}{r} + f_\eta(m, 0)\right)h\right. \\ &\left. + \frac{1}{3}\mu h^3 - \frac{l^2 h^{-1}}{4\mu r^2 q^2}\right) \end{aligned}$$

where

$$(4.12) \quad L_j = 2\mu r q H_j N_j, \quad H_j(l) = \max(H_{j-2}, l/2\mu r q N_{7-j}).$$

By (2.28), (2.29), (4.1), (2.16) and (2.7), we find that $L_2 < l < L_3$ implies

$$(4.13) \quad H/2 \leq H_4(l) < H_5(l) \leq H, \quad r \ll L = rQH/R^2 < 2l < 48C_3^2 L.$$

Generalising (2.34)–(2.36), we shall write

$$(4.14) \quad q(f_\eta(m, 0) + \bar{q}s/r) = b_s + \kappa_s,$$

where b_s is an integer and $|\kappa_s| < 3$. As it follows from (4.10) that

$$\frac{q\bar{q}s}{r} = \frac{(1 - \bar{r}ara)s}{r} = \frac{s}{r} - \bar{r}asa,$$

where $0 \leq s/r < 1$, it is possible to arrange it so that

$$(4.15) \quad \kappa_s = \kappa_0 + s/r, \quad b_s = b_0 - \bar{r}asa \quad (0 \leq s \leq r - 1).$$

For the treatment of the error-terms in (4.7)–(4.9) we proceed as follows. First (3.4) is applied to show that

$$(4.16) \quad D_i \leq \frac{1}{r} \sum_{v \pmod{r}} \sum_{H_2 \leq h \leq H_3} 2\sqrt{r} + \frac{1}{r} \sum_{H_2/r \leq k \leq H_3/r} r \ll H\sqrt{r},$$

$$(4.17) \quad E_i(N_j) \leq \sum_{H_2/r \leq k \leq H_3/r} F_k + \sqrt{r} \sum_{H_2 \leq h \leq H_3} F'_h \quad (j = 2, 3),$$

where

$$(4.18) \quad F_k = \min \left((\mu H)^{-1/2}, \left\| \left(\frac{a}{q} + 2\mu N_j \right) rk \right\|^{-1} \right),$$

$$F'_h = \frac{1}{r} \sum_{v \bmod r} \min \left((\mu H)^{-1/2}, \left\| \frac{ha}{q} + 2\mu h N_j + \frac{v}{r} \right\|^{-1} \right).$$

Now, for all integers v ,

$$\left\| \left(\frac{a}{q} + 2\mu N_j \right) rk \right\| = \left\| r \left(\frac{ha}{q} + 2\mu h N_j + \frac{v}{r} \right) \right\| \leq r \left\| \frac{ha}{q} + 2\mu h N_j + \frac{v}{r} \right\|.$$

Using this one can deduce:

$$(4.19) \quad F'_h \ll \min(r^{-1}(\mu H)^{-1/2}, \|(a/q + 2\mu N_j)rk\|^{-1}) + \log r.$$

By Dirichlet's approximation theorem, there exists a reduced rational a'/q' with

$$(4.20) \quad \left| \left(\frac{a}{q} + 2\mu N_j \right) r - \frac{a'}{q'} \right| \leq \frac{1}{3qq'}, \quad 1 \leq q' \leq 3q.$$

Here $a'/q' \neq ra/q$, since otherwise one would have the contradiction

$$r/R^2 \leq 2rN\mu \leq 2\mu N_j r \leq 1/3q^2 \leq 1/3R^2$$

(by (2.28), (2.29) and (4.1)). Therefore (4.20), (2.28) and (2.29) show that

$$\frac{1}{qq'} \leq \left| \frac{ra}{q} - \frac{a'}{q'} \right| \leq \frac{1}{3qq'} + 2r\mu N_j < \frac{1}{3qq'} + \frac{12C_3^2 r}{R^2},$$

from which it follows that

$$(4.21) \quad q' > R^2/18C_3^2 r q.$$

By (4.17)–(4.20), the first part of Lemma 3.1 applies and shows:

$$E_i(N_j) \ll \sqrt{r} (Hr^{-1}(\mu H)^{-1/2}(q')^{-1} + (H + q') \log(2rq')).$$

By (2.28), (4.20) and (4.21), this implies

$$E_i(N_j) \ll \sqrt{r} \left(\frac{\sqrt{HN}}{R} q + (H + q) \log(2rq) \right) \quad (j = 2, 3),$$

and we conclude (using (4.7), (4.16), and (4.1)) that $\mathcal{E}_{i,\eta}$ in (4.5) satisfies

$$(4.22) \quad \mathcal{E}_{i,\eta} \ll \sqrt{r} \left(\frac{Q}{R} \sqrt{HN} + (H + Q) \log(rQ) \right).$$

5. The second Poisson summation. By (3.3) and (3.4) we may estimate the minor arc sum, $\mathcal{B}_{i,\eta}^*$ from (4.11), using:

$$|\mathcal{B}_{i,\eta}^*| \leq \frac{1}{r} \sum_{L_2 < l < L_3} \sum_{\substack{H_4(l) \leq h \leq H_5(l) \\ h \equiv -\bar{r}al \pmod{q}}} \frac{1}{\sqrt{2\mu h}} |W(l, \bar{q}h)|$$

$$\begin{aligned} &\leq \sum_{L_2 < l < L_3} \sum_{\substack{H_4(l) \leq h \leq H_5(l) \\ h \equiv -\bar{r}\bar{a}l \pmod{q}}} \frac{2}{\sqrt{2r\mu h}} \\ &\quad + \sum_{L_2/r < \lambda < L_3/r} \sum_{\substack{H_4(l)/r \leq k \leq H_5(l)/r \\ k \equiv -\bar{r}\bar{a}\lambda \pmod{q}}} \frac{1}{\sqrt{2r\mu k}} \\ &\ll \frac{L}{\sqrt{r}} \left(\frac{H}{q} + 1\right) R \sqrt{\frac{N}{H}} + \frac{L}{r} \left(\frac{H}{rq} + 1\right) R \sqrt{\frac{N}{H}} \\ &\ll \sqrt{r} \frac{Q}{R} \sqrt{HN} \left(1 + \frac{H}{q}\right) \end{aligned}$$

(by (4.13) and (2.28)). Hence (and by (2.31)–(2.33) and (4.1)),

$$(5.1) \quad \mathcal{B}_{i,\eta}^* \ll \frac{Q}{R} \sqrt{rHN} \quad (\eta \in \mathbb{R} \text{ and } Q \gg H).$$

With reference to (4.1), suppose now that $Q \leq q \leq 2Q$ with given

$$(5.2) \quad Q \in [R, \varepsilon_0 H/2].$$

Note that as we effectively have $|\eta| \leq 1/2$ in (2.32), the assumption of (5.2) (with $\varepsilon_0 < 2$) guarantees $|\eta q| < H$, which is what we needed for (2.36).

Applying Poisson summation and stationary phase estimation of exponential integrals to the sum over h in (4.11), one can quite easily show that, for $\delta = 1$,

$$(5.3) \quad \mathcal{B}_{i,\eta}^* = \overline{\mathcal{B}_i^{**}(\kappa_0)} + O\left(\frac{Q}{R} r^\delta \sqrt{HN} \log(rN)\right),$$

where

$$(5.4) \quad \mathcal{B}_i^{**}(\kappa) = \frac{1}{2\sqrt{\mu r q}} \sum_{L_2 < l < L_3} \mathcal{C}_{i,l}(\kappa) e\left(\left(\frac{\bar{q}m}{r} + \frac{\bar{r}ab_0}{q}\right)l\right),$$

$$(5.5) \quad \begin{aligned} \mathcal{C}_{i,l}(\kappa) &= \sum_{K_4 \leq k - r\kappa - \sqrt{(k - r\kappa)^2 - l^2} \leq K_5} ((k - r\kappa)^2 - l^2)^{-1/4} \\ &\quad \times \chi\left(\frac{k + l}{k - l}\right) e\left(\frac{\bar{r}^2 alk}{q} + \frac{\varrho_l(k - r\kappa)}{\sqrt{\mu r^3 q^3}}\right), \end{aligned}$$

$$(5.6) \quad K_j = K_j(l) = 2\mu r q H_j^2, \quad H_j = H_j(l),$$

$$(5.7) \quad \varrho_l(\xi) = \frac{\sqrt{2}}{3} \left(\sqrt{1 + \sqrt{1 - (l/\xi)^2}} + \frac{1}{\sqrt{1 + \sqrt{1 - (l/\xi)^2}}} \right) l \sqrt{\xi},$$

with b_0, κ_0 as in (2.32)–(2.36).

Rather than giving the proof mentioned above, we shall establish (5.3) through the reciprocal procedure of applying the Poisson summation and

stationary phase estimation to the sums $\mathcal{C}_{i,l}(\kappa_0)$. By doing so we shall obtain a slight improvement of the O -term in (5.3), something that seems to us troublesome to achieve by the other mode of proof. The proof is divided into four lemmas. Before giving these lemmas we will mention that (5.7) gives $\varrho_l(\xi)$ as we first found it. Somewhat late in the course of this work we realised that it simplifies to an expression which one might guess:

$$(5.8) \quad \varrho_l(\xi) = \frac{1}{3}((\xi + l)^{3/2} - (\xi - l)^{3/2}).$$

For a proof examine each expression in turn, verifying that the domain, sign and square are the same. As (5.7) has its advantages, in places, we stick with it through to the end of this section, switching to the form (5.8) after that.

LEMMA 5.1 (Crudely truncated Poisson summation). *Let $f(x)$ be a real function with $f'(x)$ monotone and satisfying $A \leq f'(x) \leq B$, for $a < x < b$. Let $g(x)$ be a twice continuously differentiable function on $[a, b]$, and, for $\kappa \in \mathbb{R}$, define $G(\kappa)$ by*

$$G(\kappa) = \sum_{a \leq v + \kappa \leq b} g(v + \kappa)e(f(v + \kappa)).$$

Then, for $W_1 \geq W_0 = 2 \max(|A|, |B|)$ and $\kappa \in \mathbb{R}$, one has

$$G(\kappa) = \sum_{-W_1 < w < W_1} c_w e(w\kappa) + O\left(\frac{W_0 V_g + V_{g'}}{W_1} + \sum_{x \in \{a, b\}} \frac{|g(x)|}{1 + \|\kappa - x\| W_1}\right)$$

where

$$(5.9) \quad c_w = \int_0^1 G(\xi)e(-w\xi) d\xi = \int_a^b g(x)e(f(x) - wx) dx$$

and V_h denotes the total variation of $h(x)$ on $[a, b]$ plus the maximum modulus of $h(x)$ on $[a, b]$.

Proof. We first observe that Fourier's Theorem may be applied to $G(\kappa)$, since $G(\kappa)$ has period 1 and is of bounded total variation on $[0, 1]$: the latter observation follows from the bounds

$$V_g \leq \int_a^b |g'(x)| dx + \max_{a \leq x \leq b} |g(x)|, \quad V_{e(f)} \leq 2\pi \int_a^b |f'(x)| dx + 1,$$

on noting that

$$V_{h+k} \leq V_h + V_k, \quad V_{hk} \leq V_h V_k$$

and that the only discontinuities of $G(\kappa)$ occur where one, or both, of $a - \kappa$, $b - \kappa$ is an integer. Fourier's Theorem shows that the limit

$$(5.10) \quad F(\kappa) = \lim_{W \rightarrow \infty} \sum_{-W < w < W} c_w e(w\kappa)$$

(where c_w is as given by the lemma) is equal to the sum

$$\sum'_{a \leq v + \kappa \leq b} g(v + \kappa)e(f(v + \kappa)),$$

where the ' next to the summation sign indicates that the sum is to be interpreted as the arithmetic mean of the sum as given (over $v + \kappa \in [a, b]$) and the similar sum restricted to $v + \kappa \in (a, b)$. Therefore, for $W_1 > 0$ and $\kappa \in \mathbb{R}$,

$$(5.11) \quad G(\kappa) = F(\kappa) + O\left(\sum_{x \in \{a, b\}} \frac{|g(x)|}{1 + \|\kappa - x\|W_1}\right).$$

Integrating by parts in (5.9),

$$c_w = c_w(b) - c_w(a) - c_w^*(a, b),$$

where

$$c_w(x) = (-2\pi iw)^{-1}e(-wx)g(x)e(f(x)),$$

$$c_w^*(a, b) = (-2\pi iw)^{-1} \int_a^b \gamma(x)e(f(x) - wx) dx$$

with

$$\gamma(x) = 2\pi i f'(x)g(x) + g'(x).$$

For $|w| \geq W_0$, one has $f'(x) - w$ monotone and

$$|f'(x) - w| \geq |w| - \frac{1}{2}W_0 \geq |w|/2,$$

so the First Derivative Test [14, Lemma 5.1.2] may be applied to show that, for $|w| \geq W_0$,

$$|c_w^*(a, b)| \leq V_\gamma/\pi^2|w|^2 \leq (2\pi V_{f'}V_g + V_{g'})/\pi^2|w|^2$$

$$\leq (3\pi W_0V_g + V_{g'})/\pi^2|w|^2 < |w|^{-2}(W_0V_g + V_{g'}).$$

Therefore, for $W > W_1 \geq W_0$, one has

$$\sum_{W_1 \leq |w| < W} c_w e(w\kappa) = \lambda_b(W, W_1) - \lambda_a(W, W_1) + O(W_1^{-1}(W_0V_g + V_{g'})),$$

where

$$\lambda_x(W, W_1) = \sum_{W_1 \leq |w| < W} c_w(x)e(w\kappa)$$

$$= g(x)e(f(x)) \sum_{W_1 \leq |w| < W} \frac{e(-w(x - \kappa))}{-2\pi iw}$$

$$= g(x)e(f(x)) \sum_{W_1 \leq w < W} \frac{\sin(2\pi w(x - \kappa))}{\pi w}.$$

If $x - \kappa \in \mathbb{Z}$, then (trivially) $\lambda_x(W, W_1) = 0$, for all $W > W_1$. If $x - \kappa \notin \mathbb{Z}$, then, appealing once more to Fourier's Theorem,

$$\begin{aligned} \lim_{W \rightarrow \infty} \lambda_x(W, W_1) &= g(x)e(f(x)) \left(\frac{1}{2} - \{x - \kappa\} - \sum_{0 < w < W_1} \frac{\sin(2\pi w(x - \kappa))}{\pi w} \right) \\ &= g(x)e(f(x)) \cdot O((1 + \|\kappa - x\|W_1)^{-1}), \end{aligned}$$

where the last step needs only trivial estimates if $\|x - \kappa\| \ll 1/W_1$, and otherwise follows by a partial summation of the sum from W_1 to W (shown a few lines above) and the summation formula for geometric series. It follows that, for $W_1 \geq W_0$,

$$\lim_{W \rightarrow \infty} \sum_{W_1 \leq |w| < W} c_w e(w\kappa) \ll \frac{W_0 V_g + V_{g'}}{W_1} + \sum_{x \in \{a, b\}} \frac{|g(x)|}{1 + \|\kappa - x\|W_1}.$$

By (5.10) and (5.11), this gives the lemma.

LEMMA 5.2 (Stationary phase integrals). *Let $f(x)$ be a real function, four times continuously differentiable for $a \leq x \leq b$. Let $g(x)$, a complex-valued function of the real variable x , be twice continuously differentiable for $a \leq x \leq b$, and define V_g and $V_{g'}$ as in Lemma 5.1. Suppose that there are positive parameters X, Z with $X \geq b - a$, and positive constants C_r such that*

$$\begin{aligned} |f^{(r)}(x)| &\leq C_r Z X^{-r} && (a \leq x \leq b, r = 2, 3, 4), \\ f''(x) &\leq -C_2^{-1} Z X^{-2} && (a \leq x \leq b). \end{aligned}$$

Put $A = f'(b)$ and $B = f'(a)$. Then, provided only that Z be sufficiently large in terms of C_2, C_3, C_4 , we have, for $y < A$ or $y > B$,

$$\int_a^b g(x)e(f(x) - xy) dx = O\left(\frac{V_g}{\sqrt{Z}/X + \max(A - y, y - B)}\right),$$

and, for $A \leq y \leq B$,

$$\begin{aligned} \int_a^b g(x)e(f(x) - xy) dx &= g(x_y)(-f''(x_y))^{-1/2} e(f(x_y) - yx_y - 1/8) \\ &\quad + O\left(\frac{V_g}{\sqrt{Z}/X + \min(y - A, B - y)} + \frac{V_{g'}}{Z/X^2}\right), \end{aligned}$$

where $x = x_y$ is the unique solution of the equation $f'(x) = y$ in the interval $[a, b]$, and x_y^* is some member of the same interval $[a, b]$. The last O -term above is redundant if $V_{g'}/V_g \ll 1/X$.

Proof. The first result of the lemma follows from the First and Second Derivative Tests [14, Lemmas 5.1.2, 5.1.3].

Turning now to the lemma's second result, we consider first the special case in which $g(x)$ is identically equal to 1. As Z is assumed to be sufficiently large, this special case is an immediate consequence of [14, Lemma 5.1.3] (the Second Derivative Test) and [14, Lemma 5.5.2] (stationary phase integration), with the latter result only being called upon where

$$\min(y - A, B - y) > \sqrt{Z}/X.$$

We shall obtain the general case by application of the special case to the second term in the identity,

$$\begin{aligned} \int_a^b g(x)e(f(x) - xy) dx \\ = g(x_y) \int_a^b e(f(x) - xy) dx + \frac{1}{2\pi i} \int_a^b G(x) de(f(x) - xy), \end{aligned}$$

where

$$G(x) = \frac{g(x) - g(x_y)}{f'(x) - y} = \frac{g(x) - g(x_y)}{f'(x) - f'(x_y)},$$

which, l'Hôpital's rule permits us to ascertain, is a continuously differentiable real function on the interval $[a, b]$. As $g(x_y) = O(V_g)$ and since integration by parts shows the last integral above to be $O(V_G)$, the general case will follow if only we can show

$$V_G \ll \frac{V_{g'}}{ZX^{-2}}.$$

Using the truncated Taylor series, centred at x , for $f'(x_y)$ in

$$G'(x) = \frac{g'(x)(f'(x) - f'(x_y)) - (g(x) - g(x_y))f''(x)}{(f'(x) - f'(x_y))^2},$$

one can show that, for $a \leq x \leq b$ and $x \neq x_y$, there exist $x_0, x_1 \in [a, b]$ such that

$$G'(x) = -\frac{f'''(x_1)}{2(f''(x_0))^2} g'(x) - \frac{f''(x)}{(f''(x_0))^2} g_2(x),$$

where

$$g_2(x) = \frac{g(x) - g(x_y) - g'(x)(x - x_y)}{(x - x_y)^2} = \frac{-1}{(x - x_y)^2} \int_{x_y}^x (u - x_y)g''(u) du.$$

Since

$$\int_a^b |g_2(x)| dx \leq \int_a^b \frac{1}{(x - x_y)^2} \int_{x_y}^x |g''(u)|(u - x_y) du dx$$

$$\begin{aligned}
 &= \int_{x_y}^b |g''(u)|(u - x_y) \int_u^b \frac{1}{(x - x_y)^2} dx du \\
 &\quad + \int_a^{x_y} |g''(u)|(x_y - u) \int_a^u \frac{1}{(x - x_y)^2} dx du \\
 &\leq \int_{x_y}^b |g''(u)|(u - x_y) \int_u^\infty \frac{1}{(x - x_y)^2} dx du \\
 &\quad + \int_a^{x_y} |g''(u)|(x_y - u) \int_{-\infty}^u \frac{1}{(x - x_y)^2} dx du \\
 &= \int_a^b |g''(u)| du,
 \end{aligned}$$

it follows that we may use the conditions bounding $f^{(r)}(x)$ to show

$$\int_a^b |G'(x)| dx \ll \frac{1}{ZX^{-1}} \int_a^b |g'(x)| dx + \frac{1}{ZX^{-2}} \int_a^b |g''(x)| dx.$$

Bounding the first integral on the right by X times the maximum of its integrand, and applying mean-value theorem (or l'Hôpital's rule) to see that, for $a \leq x \leq b$,

$$|G(x)| \leq \frac{\max_{a \leq u \leq b} |g'(u)|}{\min_{a \leq v \leq b} |f''(v)|} \ll \frac{1}{ZX^{-2}} \max_{a \leq u \leq b} |g'(u)|,$$

we obtain the desired bound on V_G , so completing the proof of the second part of the lemma.

All that remains to be explained is the final remark of the lemma. As $Z \gg 1$, this is a consequence of the bounds

$$\min(y - A, B - y) < B - A = \int_b^a f''(x) dx \ll X \cdot ZX^{-2} = ZX^{-1}.$$

Our next lemma will deal with the particular sums $C_{i,l}(\kappa)$ found in (5.4) and defined in (5.5)–(5.7). Before stating the lemma we find it convenient to introduce some simplifying notation.

Given $l \in \mathbb{Z}$ satisfying

$$(5.12) \quad L_2 < l < L_3,$$

we have it from (5.5)–(5.7) (and see (5.22) below) that

$$(5.13) \quad C_{i,l}(\kappa) = \sum_{\alpha \leq \phi'(k-r\kappa) \leq \beta} \psi(k)\gamma(k - r\kappa)e(\phi(k - r\kappa) + nk/q),$$

where

$$(5.14) \quad \psi(k) = \chi\left(\frac{k+l}{k-l}\right),$$

$$(5.15) \quad n \equiv \overline{r^2al} \pmod{q},$$

$$(5.16) \quad \gamma(\xi) = (\xi^2 - l^2)^{-1/4}, \quad \phi(\xi) = \varrho_l(\xi)/\sqrt{\mu r^3 q^3},$$

$$(5.17) \quad \alpha = H_4/rq, \quad \beta = H_5/rq.$$

LEMMA 5.3. *Let i be the index of a minor arc for which $q = q_i$ satisfies (5.2). Let l be an integer satisfying (5.12). Then the sum $C_{i,l}(\kappa_0)$, given by (5.13)–(5.17) (with $\kappa = \kappa_0$), is equal to*

$$\sum_{\substack{rq\alpha \leq h \leq rq\beta \\ h \equiv -nr \pmod{q}}} \frac{W(-l, -\bar{q}h)}{r} \frac{\gamma(\xi_h)}{\sqrt{-\phi''(\xi_h)}} e\left(\phi(\xi_h) - \frac{h}{rq} (\xi_h + r\kappa_0) - \frac{1}{8}\right) + O\left(\frac{|\gamma(\xi_*)|}{\sqrt{-\phi''(\xi_*)}} \left(1 + \left(\frac{H}{N}\right)^{1/2} \log \frac{rH}{R}\right)\right),$$

where $W(x, y)$ is as in (3.3), $\xi = \xi_h$ is the unique solution in $[l, \infty)$ of

$$(5.18) \quad \phi'(\xi_h) = h/rq,$$

$\xi_* = \xi_y$ for some $y \in [H_4, H_5]$, and, for $l \not\equiv 0 \pmod{r}$, it is allowed that the O -term be made \sqrt{r} times smaller.

Proof. Writing $k = -rv - s$, where $v \in \mathbb{Z}$ and $s \in \{0, 1, \dots, r - 1\}$, we have (from (5.13) and (5.14)):

$$(5.19) \quad C_{i,l}(\kappa_0) = \sum_{s=0}^{r-1} \psi(-s)e(-sd/r + \theta\kappa_0)G(\kappa_s),$$

with

$$(5.20) \quad d = [nr/q], \quad \theta = \{nr/q\},$$

$$(5.21) \quad G(\kappa) = \sum_{\alpha \leq \phi'(-r(v+\kappa)) \leq \beta} \gamma(-r(v+\kappa))e(\phi(-r(v+\kappa)) - (v+\kappa)\theta).$$

The real function

$$(5.22) \quad \phi'(\xi) = \frac{l/\sqrt{2\mu r^3 q^3}}{\sqrt{\xi + \sqrt{\xi^2 - l^2}}} = \frac{\sqrt{\xi - \sqrt{\xi^2 - l^2}}}{\sqrt{2\mu r^3 q^3}}$$

has the domain $[l, \infty)$, on which it decreases strictly monotonically (tending to 0 in the limit as $\xi \rightarrow \infty$). Therefore $\phi'(\xi)$ has the range $(0, \sqrt{l/2\mu r^3 q^3}]$.

Note that α, β in (5.17) do lie properly in this range, since it follows from (5.12), (4.12), (2.29), (2.8) and (4.13) that

$$(5.23) \quad \sqrt{l/2\mu r^3 q^3} > \sqrt{H_2 N_2 / r q} \geq \sqrt{HN} / 2r q \geq H / r q \geq \beta > \alpha.$$

This shows that $G(\kappa)$ in (5.21) has the form of the sum in Lemma 5.1, with

$$(5.24) \quad g(x) = \gamma(-rx), \quad f(x) = \phi(-rx) - \theta x,$$

$$(5.25) \quad a < b < -l/r, \quad \phi'(-ra) = \alpha, \quad \phi'(-rb) = \beta,$$

so that we may take, in Lemma 5.1,

$$(5.26) \quad A = -r\beta - \theta, \quad B = -r\alpha - \theta$$

with, by (5.23) and (5.2),

$$(5.27) \quad W_0 = 2 \max(|A|, |B|) \leq 2(H/q + 1) < 4H/Q.$$

By (5.24) and (5.16),

$$(5.28) \quad g(x) = (r^2 x^2 - l^2)^{-1/4},$$

which is positive-valued, increasing and concave-up on the interval $[a, b]$ (see (5.25)). Therefore, with reference to Lemma 5.1, we have

$$(5.29) \quad V_g \leq 2g(b), \quad V_{g'} \leq 2g'(b).$$

By (5.22), (5.25),

$$\alpha \leq \frac{l/\sqrt{2\mu r^3 q^3}}{\sqrt{-rx}}, \quad \frac{l/\sqrt{2\mu r^3 q^3}}{\sqrt{-2rx}} \leq \beta \quad (a \leq x \leq b),$$

from which it follows, by (5.17), (4.12), (2.29) and (2.8), that

$$(5.30) \quad \frac{6N}{H} \geq \frac{N_3}{H_2} \geq \frac{-rx}{l} \geq \frac{N_2}{2H_3} \geq \frac{N}{2H} \geq 32C_3 \quad (a \leq x \leq b).$$

By (5.28)–(5.30) and (4.13), (4.1), (2.7) and (2.8),

$$(5.31) \quad V_g \asymp r^{-1/2} X^{-1/2} \asymp X V_{g'} \quad \text{with} \quad X = \frac{6Nl}{Hr} \asymp \frac{QN}{R^2} > 1.$$

Hence, and by virtue of (5.27), application of Lemma 5.1 yields the result that, for

$$(5.32) \quad W_1 \geq 4H/Q,$$

the sum $G(\kappa_s)$, given by (5.21), will equal

$$(5.33) \quad \sum_{-W_1 < w < W_1} c_w e(w\kappa_s) + O\left(\frac{H V_g}{Q W_1} + \sum_{x \in \{a, b\}} \frac{V_g}{1 + \|\kappa_s - x\| W_1}\right),$$

where c_w is as in (5.9).

By (4.15) and (5.14), putting (5.33) into (5.19) gives

$$(5.34) \quad \mathcal{C}_{i,l}(\kappa_0) = \sum_{-W_1 < w < W_1} c_w W(-l, w - d) e((w + \theta)\kappa_0) + O\left(V_g \left(\frac{rH}{QW_1} + 1 + \sum_{s=1}^{(r-1)/2} \frac{1}{W_1 s/r} \right)\right),$$

where $W(x, y)$ is the Weil sum given by (3.3). We choose

$$(5.35) \quad W_1 = 4r(H/Q + \log r) \geq 2$$

(see (5.2)), so that (5.32) is satisfied. This choice reduces the O -term in (5.34) to just $O(V_g)$.

By (5.22),

$$(5.36) \quad \phi''(\xi) = -\frac{\phi'(\xi)}{2\sqrt{\xi^2 - l^2}},$$

so it follows from (5.24)–(5.26), (5.30), (5.17), (4.13), (2.8), (4.1) and (5.2) that the conditions of Lemma 5.2 hold with X as in (5.31) and

$$(5.37) \quad Z = \frac{r^2(H/2rq)}{3\sqrt{r^2 X^2}} X^2 = \frac{HX}{6q} = \frac{Nl}{qr} > \frac{NL}{4Qr} = \frac{HN}{4R^2} \geq \left(\frac{H}{2Q}\right)^2 \geq \varepsilon_0^{-2}.$$

By (4.1) and (4.13), one also has

$$(5.38) \quad Z = Nl/qr \leq 24C_3^2 HN/R^2.$$

A sufficiently small choice of the positive constant ε_0 in (5.37) will ensure that Z is large enough for Lemma 5.2 to apply, yielding estimates for the exponential integral c_w given by (5.9). By substituting these estimates into (5.34), and appealing to no more than the trivial bound of order $O(r)$ for $W(-l, w - d)$ (see (3.3)), one finds that $\mathcal{C}_{i,l}(\kappa_0)$ is equal to

$$(5.39) \quad \sum_{A \leq w \leq B} W(-l, w - d) \frac{g(x_w)}{\sqrt{-f''(x_w)}} e\left(f(x_w) - wx_w + (w + \theta)\kappa_0 - \frac{1}{8}\right) + O(V_g) + O\left(V_g \left(\frac{X}{\sqrt{Z}} + \sum_{u=1}^{2W_1} \frac{1}{u} \right) r\right)$$

(by (5.31) the last remark of Lemma 5.2 applies).

When $l \not\equiv 0 \pmod{r}$, Weil’s bound (3.4) for $W(-l, w - d)$ is of order $O(\sqrt{r})$, which leads to a sharpening of the last O -term in (5.39) by a factor of \sqrt{r} . Even with this sharpening, the second O -term will still dominate the first O -term, which is therefore redundant in all cases. It follows that the

O -terms in (5.39) may be subsumed within a single term:

$$(5.40) \quad O\left(rV_g\left(\frac{X}{\sqrt{Z}} + \log W_1\right)\right) \ll \frac{g(b)}{\sqrt{-f''(b)/r^2}} \left(1 + \sqrt{\frac{H}{N}} \log \frac{Hr}{R}\right)$$

(see (5.29), (5.31), (5.35), (5.38), (2.7), (2.8) and (4.1)). Since $(q, r) = 1$, the lemma will follow after (5.24), (5.26), (5.20), (3.3) and the substitutions

$$-qw - q\theta = h \equiv -nr \pmod{q}, \quad -rx_w = \xi_h$$

are used to rewrite (5.40) and (5.39) (in which, by Lemma 5.2, $f'(x_w) = w$). The sharper O -terms for the case $l \not\equiv 0 \pmod{r}$ are inferred from the remarks made at the start of this paragraph.

LEMMA 5.4. *Let i be the index of a minor arc for which $q = q_i$ satisfies (5.2). Then (5.3) holds with $\delta = 1/2$.*

Proof. Our starting point is the result of Lemma 5.3, which we prepare for substitution into (5.4) by carrying out a reversal of the notational substitutions made in (5.14)–(5.17).

First note that, by (5.16), (5.36) and (5.18),

$$(5.41) \quad \frac{\gamma(\xi_h)}{\sqrt{-\phi''(\xi_h)}} = \frac{1}{\sqrt{-\phi''(\xi_h)\sqrt{\xi_h^2 - l^2}}} = \sqrt{\frac{2}{\phi'(\xi_h)}} = \sqrt{\frac{2rq}{h}}.$$

By (5.18) and (5.22),

$$(5.42) \quad \sqrt{\xi_h + \sqrt{\xi_h^2 - l^2}} = \frac{l}{h\sqrt{2\mu rq}}, \quad \sqrt{\xi_h - \sqrt{\xi_h^2 - l^2}} = h\sqrt{2\mu rq}.$$

From this, (5.16) and (5.7), it follows that

$$\phi(\xi_h) = \frac{\sqrt{2}l}{3\sqrt{\mu r^3 q^3}} \left(\frac{l}{h\sqrt{2\mu rq}} + \frac{h\sqrt{2\mu rq}}{l} \xi_h \right) = \frac{l^2}{3\mu r^2 q^2 h} + \frac{2h}{3rq} \xi_h$$

and

$$(5.43) \quad \phi(\xi_h) - \frac{h}{rq} \xi_h = \frac{l^2}{3\mu r^2 q^2 h} - \frac{h}{3rq} \xi_h.$$

By (5.42),

$$\xi_h = \frac{1}{2} \left(\frac{l^2}{2\mu r q h^2} + 2\mu r q h^2 \right) = \frac{l^2}{4\mu r q h^2} + \mu r q h^2.$$

Substituting this into (5.43) yields

$$\phi(\xi_h) - \frac{h}{rq} \xi_h = \frac{l^2}{4\mu r^2 q^2 h} - \frac{\mu h^3}{3}.$$

Using this, (5.41), (5.15) and (5.17) with the result of Lemma 5.3, we find the following approximate formula for $\mathcal{C}_{i,l}(\kappa_0)$:

$$\sum_{\substack{H_4 \leq h \leq H_5 \\ h \equiv -\bar{r}al \pmod{q}}} \frac{W(-l, -\bar{q}h)}{r} \sqrt{\frac{2rq}{h}} e\left(\frac{l^2}{4\mu r^2 q^2 h} - \frac{\mu h^3}{3} - r\kappa_0 \frac{h}{rq} - \frac{1}{8}\right) + O\left(\sqrt{\frac{rq}{H}} \left(1 + \sqrt{\frac{H}{N}} \log \frac{rH}{R}\right)\right),$$

with the O -term smaller by a factor of \sqrt{r} when $l \not\equiv 0 \pmod{r}$.

By (4.13) there are no more than $O(L/r)$ choices for l in (5.4) with $l \equiv 0 \pmod{r}$. Therefore substitution of the last result (above) into (5.4) shows us that $\mathcal{B}_i^{**}(\kappa_0)$ is equal to

$$(5.44) \quad \sum_{L_2 < l < L_3} e\left(\frac{\bar{q}ml}{r}\right) \sum_{\substack{H_4(l) \leq h \leq H_5(l) \\ h \equiv -\bar{r}al \pmod{q}}} \frac{W(-l, -\bar{q}h)}{r} \frac{1}{\sqrt{2\mu h}} \times e\left(\frac{l^2}{4\mu r^2 q^2 h} - \frac{\mu h^3}{3} - \frac{(b_0 + \kappa_0)h}{q} - \frac{1}{8}\right) + O\left(\frac{L}{\sqrt{\mu r q}} \sqrt{\frac{q}{H}} \left(1 + \sqrt{\frac{H}{N}} \log \frac{rH}{R}\right)\right).$$

By (3.3) and (4.14) (for $s = 0$), the sums in (5.44) are the complex conjugates of those defining $\mathcal{B}_{i,\eta}^*$ in (4.11). To complete the proof of the lemma, it suffices to observe that (4.12), (4.13) and (2.29) imply

$$\frac{L}{\sqrt{\mu r q}} \asymp \sqrt{LHN} \asymp \sqrt{\frac{rQH^2N}{R^2}} = \frac{Q}{R} \sqrt{r} H \sqrt{\frac{N}{Q}},$$

so that, by (4.1) and (2.8), the O -term in (5.44) is

$$O\left(\frac{Q}{R} \sqrt{rHN} \left(1 + \sqrt{\frac{H}{N}} \log \frac{rH}{R}\right)\right) = O\left(\frac{Q}{R} \sqrt{rHN} \log(rN)\right).$$

6. The Bombieri–Iwaniec Double Large Sieve. By (5.4)–(5.5), the sum $\mathcal{B}_i^{**}(\kappa)$ can be written as

$$(6.1) \quad \sum_{L_2 < l < L_3} \sum_{\substack{k \\ K_4(l) \leq k - r\kappa - \sqrt{(k - r\kappa)^2 - l^2} \leq K_5(l)}} \frac{\psi_{k,l}(\kappa)}{\sqrt{4\mu r q}} e(\mathbf{y}^{(i)} \cdot \mathbf{x}^{(k,l)}(\kappa)),$$

where

$$(6.2) \quad \psi_{k,l}(\kappa) = ((k - r\kappa)^2 - l^2)^{-1/4} \chi\left(\frac{k+l}{k-l}\right),$$

$$(6.3) \quad \mathbf{x}^{(k,l)}(\kappa) = (kl, l, \varrho_l(k - r\kappa)),$$

$$(6.4) \quad \mathbf{y}^{(i)} = \left(\left\{ \frac{r^2 a}{q} \right\}, \left\{ \frac{\bar{q}m}{r} + \frac{r\bar{a}b_0}{q} \right\}, \frac{1}{\sqrt{\mu r^3 q^3}} \right).$$

Before we can apply the Double Large Sieve it is necessary to dissolve all but one of the ties linking the ranges of summation for k and l to the minor arc index i . By (4.12) and (5.6), the restrictions on k and l in (6.1) are that

$$(6.5) \quad 2\mu r q H_2 N_2 < l < 2\mu r q H_3 N_3$$

and that $k - r\kappa - ((k - r\kappa)^2 - l^2)^{1/2}$ should lie in the interval

$$[2\mu r q \max(H_2^2, (l/2\mu r q N_3)^2), 2\mu r q \min(H_3^2, (l/2\mu r q N_2)^2)].$$

The last constraint is satisfied if and only if both

$$2\mu r q H_2^2 \leq k - r\kappa - \sqrt{(k - r\kappa)^2 - l^2} \leq 2\mu r q H_3^2$$

and

$$2\mu r q N_2^2 \leq k - r\kappa + \sqrt{(k - r\kappa)^2 - l^2} \leq 2\mu r q N_3^2,$$

or equivalently (given (6.5)) if both

$$(6.6) \quad 2H_2 \sqrt{\mu r q} \leq \sqrt{k - r\kappa + l} - \sqrt{k - r\kappa - l} \leq 2H_3 \sqrt{\mu r q}$$

and

$$(6.7) \quad 2N_2 \sqrt{\mu r q} \leq \sqrt{k - r\kappa + l} + \sqrt{k - r\kappa - l} \leq 2N_3 \sqrt{\mu r q}.$$

Note that (6.6) and (6.7) multiply together to give (6.5). Recalling (5.13), (5.25) and (5.30) (from Lemma 5.3 and its proof), we see that in (5.4)–(5.5) the ratio $(k - r\kappa)/l$ will always lie between half and six times N/H . Hence, and by (4.13) and (5.30), we may write

$$(6.8) \quad \mathcal{B}_i^{**}(\kappa_0) = \sum_{K_0} \sum_{L_0} \mathcal{B}_i^{**}(K_0, L_0; \kappa_0),$$

with K_0, L_0 running over the bounded number of integer powers of 2 such that

$$(6.9) \quad \frac{1}{8} \leq \frac{K_0}{K}, \frac{L_0}{L} \leq 144C_3^2, \quad \frac{12N}{H} \geq \frac{K_0}{L_0} \geq \frac{N}{4H} \geq 16,$$

where, by (5.2),

$$(6.10) \quad K = \frac{N}{H} L = \frac{rQN}{R^2} \geq \frac{rN}{Q} \geq r \frac{64H}{Q} > 64r,$$

and with

$$(6.11) \quad \mathcal{B}_i^{**}(K_0, L_0; \kappa) = \sum_{L_0 < l \leq 2L_0} \sum_{K_0 < k - r\kappa \leq 2K_0}^{(i)} \frac{\psi_{k,l}(\kappa)}{\sqrt{4\mu r q}} e(\mathbf{y}^{(i)} \cdot \mathbf{x}^{(k,l)}(\kappa)),$$

where the superfix (i) to the summation sign indicates that (6.6) and (6.7) are conditions of summation: note that we can drop (6.5) here as it is implicit.

LEMMA 6.1. *Let n be a positive integer constant. Let \mathcal{F} be any finite set and $\mathbf{c}(x)$ and $f(x)$ any functions from \mathcal{F} into (respectively) \mathbb{R}^n and \mathbb{C} . Suppose also that $a_i \leq b_i$ for $i = 1, \dots, n$. Put*

$$\mathcal{S} = \sum_{\substack{x \in \mathcal{F} \\ \mathbf{c}(x) \in [a_1, b_1] \times \dots \times [a_n, b_n]}} f(x)$$

and, for $\mathbf{u} \in \mathbb{R}^n$,

$$\mathcal{S}(\mathbf{u}) = \sum_{x \in \mathcal{F}} f(x) e(-\mathbf{c}(x) \cdot \mathbf{u}).$$

Then for $U_1, \dots, U_n > 0$,

$$\mathcal{S} \ll \int_{-U_1}^{U_1} \dots \int_{-U_n}^{U_n} |\mathcal{S}(\mathbf{u})| \lambda(\delta_1, u_1) du_1 \dots \lambda(\delta_n, u_n) du_n + \sum_{i=1}^n E_i,$$

where $\lambda(\delta, u) = \delta / (1 + \delta|u|)$, $\delta_i = b_i - a_i$ and

$$E_i = \sum_{x \in \mathcal{F}} |f(x)| \left(\frac{1}{1 + U_i |c_i(x) - a_i|} + \frac{1}{1 + U_i |c_i(x) - b_i|} \right).$$

Proof. This is a standard type of result which follows by application of the asymptotic formula

$$\frac{1}{2\pi i} \int_{-U}^U (e(vu) - 1) \frac{du}{u} = O\left(\frac{1}{1 + U|v|}\right) + \begin{cases} 1/2 & \text{if } v > 0, \\ 0 & \text{if } v = 0, \\ -1/2 & \text{if } v < 0 \end{cases}$$

(with $U = U_i > 0$ and $v = b_i - c_i(x)$, or $v = a_i - c_i(x)$) to estimate the integral

$$\int_{-U_1}^{U_1} \dots \int_{-U_n}^{U_n} \mathcal{S}(\mathbf{u}) \frac{(e(b_1 u_1) - e(a_1 u_1)) du_1}{2\pi i u_1} \dots \frac{(e(b_n u_n) - e(a_n u_n)) du_n}{2\pi i u_n},$$

while (separately) using the inequality

$$\left| \frac{e(vu) - 1}{2\pi i u} \right| = |v \operatorname{sinc}(vu)| \leq \frac{|v|}{1 + \pi|vu|} \leq \lambda(|v|, u) \quad (v, u \in \mathbb{R})$$

(with $v = \delta_i$) to bound the same integral.

We apply Lemma 6.1 to bound the sum $\mathcal{B}_i^{**}(K_0, L_0; \kappa)$ in (6.11) in terms of a similar sum not subject to the conditions (6.6) and (6.7). To this end the lemma is applied with $n = 2$,

$$\mathcal{F} = \{(k, l) \in \mathbb{Z}^2 \mid K_0 < k - r\kappa \leq 2K_0, L_0 < l \leq 2L_0\},$$

$$(6.12) \quad \mathbf{c}(k, l) = (\sqrt{k - r\kappa + l} - \sqrt{k - r\kappa - l}, \sqrt{k - r\kappa + l} + \sqrt{k - r\kappa - l}),$$

$$f(k, l) = \psi_{k,l}(\kappa) e(\mathbf{y}^{(i)} \cdot \mathbf{x}^{(k,l)}(\kappa)),$$

$$(6.13) \quad \mathbf{a} = 2\sqrt{\mu r q} (H_2, N_2), \quad \mathbf{b} = 2\sqrt{\mu r q} (H_3, N_3), \quad \mathbf{U} = K_0^{3/2} (1, 1).$$

With a linear change of variable in the integral, Lemma 6.1 yields

$$(6.14) \quad \mathcal{B}_i^{**}(K_0, L_0; \kappa) \ll \int_{-1}^1 \int_{-1}^1 |\mathcal{B}_{i,\mathbf{w}}^{**}(K_0, L_0; \kappa)| \lambda_1(w_1) dw_1 \lambda_2(w_2) dw_2 + (E_1 + E_2) / \sqrt{4\mu r q},$$

where $\mathcal{B}_{i,\mathbf{w}}^{**}(K_0, L_0; \kappa)$ is

$$(6.15) \quad \frac{1}{\sqrt{4\mu r q}} \sum_{L_0 < l \leq 2L_0} \sum_{K_0 < k - r\kappa \leq 2K_0} \psi_{k,l}(K_0^{3/2} \mathbf{w}; \kappa) e(\mathbf{y}^{(i)} \cdot \mathbf{x}^{(k,l)}(\kappa))$$

with

$$(6.16) \quad \psi_{k,l}(\mathbf{u}; \kappa) = \psi_{k,l}(\kappa) e(-\mathbf{u} \cdot \mathbf{c}(k, l)),$$

and where, by (2.29), (2.8), (2.28), (5.2), (6.10), (6.9), (6.13) and the monotonicity (with respect to δ) of $\lambda(\delta, u)$,

$$(6.17) \quad 0 \leq \lambda_i(w) = \lambda(U_i(b_i - a_i), w) \leq \lambda(O(K_0^2), w) \ll \lambda(K_0^2, w).$$

Furthermore, as one can show that, for $k - r\kappa > |l|$,

$$\frac{\partial}{\partial k} c_1(k, l) \leq -\frac{l}{2(k - r\kappa)^{3/2}}, \quad \frac{\partial}{\partial k} c_2(k, l) \geq \frac{1}{\sqrt{2(k - r\kappa)}},$$

it follows that, for fixed integer l in $[L_0, 2L_0]$ and $i \in \{1, 2\}$, the function $c_i(k, l)$, restricted to the strip $K_0 \leq k - r\kappa \leq 2K_0$, is strictly monotonic in k , with values at consecutive integer values of k always an interval of length greater than $(4K_0)^{-3/2}$ apart. Therefore, for $i = 1, 2$,

$$E_i \ll \sum_{L_0 < l \leq 2L_0} \left(1 + \sum_{0 < k_1 \leq K_0} \frac{1}{(4K_0)^{-3/2} k_1 U_i} \right) \max_{K_0 \leq k - r\kappa \leq 2K_0} |f(k, l)|$$

$$\ll (L_0 \log K_0) \max_{\substack{L_0 < l \leq 2L_0 \\ K_0 \leq k - r\kappa \leq 2K_0}} |\psi_{k,l}(\kappa)| \ll (L_0 \log K_0) (K_0^2 - 4L_0^2)^{-1/4}$$

and

$$\frac{E_i}{\sqrt{\mu r q}} \ll (K/N^2)^{-1/2} L_0 K_0^{-1/2} \log K_0 \ll N \frac{L}{K} \log K \ll H \log(rN)$$

(see (2.28), (5.2), (6.2), (6.9) and (6.10)).

This completes the process of simplifying the dependence on arc index (i) of the sums over k, l in (6.11). By the last bound, (2.32), (2.35), (4.5),

(4.22), Lemma 5.4, (6.8)–(6.10), (6.14)–(6.17), (5.2) and (2.8),

$$\begin{aligned} \mathcal{B}_i &\ll \int_{-1/2}^{1/2} \left(\sum_{K_0} \sum_{L_0} \int_{-1}^1 \int_{-1}^1 |\mathcal{B}_{i,\mathbf{w}}^{**}(K_0, L_0; \kappa_{0,i}(\eta))| \right. \\ &\quad \times \lambda(K_0^2, w_1) dw_1 \lambda(K_0^2, w_2) dw_2 \Big) \Lambda(2\eta) d\eta \\ &\quad + \frac{Q}{R} \sqrt{rHN} \log(rN). \end{aligned}$$

Hence, if we let $S(H, M; Q)$ be the total contribution to the sum $S(H, M)$, in (1.6), from all those minor arcs J_i that belong to no major arc and for which the chosen rational a/q has $Q \leq q \leq 2Q$, then (see (2.21)):

$$\begin{aligned} S(H, M; Q) &= \sum_i^{(Q)} \mathcal{B}_i \quad (\text{say}) \\ &\ll \sum_{K_0} \sum_{L_0} \int_{-1}^1 \int_{-1}^1 \sum_i^{(Q)} \int_{-1/2}^{1/2} |\mathcal{B}_{i,\mathbf{w}}^{**}(K_0, L_0; \kappa_{0,i}(\eta))| \\ &\quad \times \Lambda(2\eta) d\eta \lambda(K_0^2, w_1) dw_1 \lambda(K_0^2, w_2) dw_2 \\ &\quad + \sum_i^{(Q)} \frac{Q}{R} \sqrt{rHN} \log(rN). \end{aligned}$$

As there are $O(1)$ choices of K_0, L_0 as integer powers of 2 satisfying (6.9), and since, in Lemma 6.1,

$$\int_{-1}^1 \lambda(\delta, w) dw = 2 \log(1 + \delta) \quad (\delta \geq 0),$$

application of the First Mean-Value Theorem of integral calculus now permits us to conclude that either

$$(6.18) \quad S(H, M; Q) \ll \left(\frac{Q}{R} \sqrt{rHN} \log(rN) \right) W(Q),$$

where

$$(6.19) \quad W(Q) = \sum_i^{(Q)} 1,$$

or, for some K_0, L_0 satisfying (6.9), and some $\mathbf{w} \in \mathbb{R}^2$,

$$(6.20) \quad S(H, M; Q) \ll (\log^2 K_0) \sum_i^{(Q)} \int_{-1/2}^{1/2} |\mathcal{B}_{i,\mathbf{w}}^{**}(K_0, L_0; \kappa_{0,i}(\eta))| \Lambda(2\eta) d\eta.$$

As it follows from (2.15) and (2.5) that $q \geq Q$ implies both $s_1 \gg Q$ and $b_1 \gg P$ (where $P/Q \asymp T/M^2$ and b_1/s_1 is as described above (2.14)), one may choose either to apply [18, Lemma 2.1] in order to bound the number

of relevant subintervals $J_{i-2} \ni b_1/s_1$ directly, or to apply it to bound the number of “reciprocal” subintervals, $J_{i-2}^{-1} = \{1/x \mid x \in J_{i-2}\} \ni s_1/b_1$. The resulting two bounds combine to show

$$W(Q) \ll \frac{MR^2}{NQ^2} + \min\left(\frac{R^2}{Q}, \frac{M}{NQ}\right) \ll \frac{MR^2}{NQ^2} + \sqrt{\frac{MR^2}{NQ^2}}.$$

Therefore, given that $W(Q) \in \mathbb{Z}$, one has

$$(6.21) \quad W(Q) \ll \frac{MR^2}{NQ^2}.$$

Assuming (6.20) (and recalling (6.15) and (6.3)), the Cauchy–Schwarz inequality and (2.28) show that, for some $\mathbf{u} \in \mathbb{R}^2$,

$$(6.22) \quad (S(H, M; Q))^2 \ll (\log N)^4 W(Q) \frac{NR^2}{rQ} \sigma,$$

where

$$(6.23) \quad \sigma = \sum_i^{(Q)} \int_{-\infty}^{\infty} \sum_{\mathbf{j}} \Psi_{\mathbf{j}}(\kappa_{0,i}(\eta)) e(\mathbf{y}^{(i)} \cdot \mathbf{x}^{(\mathbf{j})}(\kappa_{0,i}(\eta))) \Lambda(2\eta) d\eta$$

with $\mathbf{j} = (k, k', l, l')$ running over \mathbb{Z}^4 ,

$$(6.24) \quad \Psi_{\mathbf{j}}(\kappa) = \begin{cases} \psi_{j_1, j_3}(\mathbf{u}; \kappa) \overline{\psi_{j_2, j_4}(\mathbf{u}; \kappa)} \\ \text{if } (j_h - r\kappa)/K_0, j_{h+2}/L_0 \in (1, 2] \text{ for } h = 1, 2, \\ 0 \text{ otherwise} \end{cases}$$

and

$$(6.25) \quad \mathbf{x}^{(\mathbf{j})}(\kappa) = (j_1 j_3 - j_2 j_4, j_3 - j_4, \varrho_{j_3}(j_1 - r\kappa) - \varrho_{j_4}(j_2 - r\kappa)).$$

By (6.25), (5.7), (2.35), (2.36), (6.9), (6.10), (6.4), (2.28) and (5.2), we find that, for all \mathbf{j}, i and η relevant to (6.23)–(6.24),

$$-X_h/2 \leq x_h^{(\mathbf{j})}(\kappa_{0,i}(\eta)) \leq X_h/2, \quad -Y_h/2 \leq y_h^{(i)} \leq Y_h/2 \quad (h = 1, 2, 3),$$

where

$$(6.26) \quad \frac{\mathbf{X}}{C_0} = \left(KL, L, \frac{H\sqrt{r^3 Q^3 N}}{R^3} \right), \quad \frac{\mathbf{Y}}{C_0} = \left(1, 1, \frac{R}{Q} \sqrt{\frac{N}{r^3 Q}} \right)$$

for some $C_0 \geq 1$, with $C_0 = O(1)$.

Recalling (2.33)–(2.36), we may rewrite (6.23) by substitution of the variable of integration

$$\eta = (\kappa_{0,i}(\eta) - \kappa_{0,i}(0))H/q = (\kappa_0 - \kappa_{0,i})H/q \quad (\text{say}),$$

where

$$(6.27) \quad \kappa_{0,i} = \{qf(m, 0)\} \in [0, 1).$$

We then have (see (5.2)),

$$\begin{aligned}
 (6.28) \quad \sigma &= \sum_i^{(Q)} \int_{-\infty}^{\infty} \Lambda\left(\frac{\kappa_0 - \kappa_{0,i}}{q/2H}\right) \sum_j \Psi_j(\kappa_0) e(\mathbf{y}^{(i)} \cdot \mathbf{x}^{(j)}(\kappa_0)) \frac{d\kappa_0}{q/H} \\
 &= \int_{-\varepsilon_0/2}^{1+\varepsilon_0/2} \sigma(\kappa_0) d\kappa_0,
 \end{aligned}$$

where

$$\sigma(\kappa) = \sum_i^{(Q)} \sum_j \mathbf{e}(\mathbf{y}^{(i)} \cdot \mathbf{x}^{(j)}(\kappa)) \Phi_i(\kappa) \Psi_j(\kappa),$$

with

$$(6.29) \quad \Phi_i(\kappa) = \frac{H}{q} \Lambda\left(\frac{\kappa - \kappa_{0,i}}{q/2H}\right).$$

Applying the Double Large Sieve [14, Lemma 5.6.6] yields the result that, for $-2 < \kappa < 2$,

$$(6.30) \quad (\sigma(\kappa))^2 \ll A(\kappa)B(\kappa) \prod_{h=1}^3 (X_h Y_h + 1),$$

where

$$(6.31) \quad A(\kappa) = \sum_j \sum_{j'} \Psi_j(\kappa) \overline{\Psi_{j'}(\kappa)} \prod_{h=1}^3 \Lambda\left(\frac{x_h^{(j)}(\kappa) - x_h^{(j')}(\kappa)}{X_h/Z_h}\right),$$

$$(6.32) \quad B(\kappa) = \sum_i^{(Q)} \sum_{i'}^{(Q)} |\Phi_i(\kappa) \Phi_{i'}(\kappa)| \prod_{h=1}^3 \Lambda\left(\frac{y_h^{(i)} - y_h^{(i')}}{Y_h/(1 + Z_h)}\right),$$

with

$$(6.33) \quad Z_h = X_h Y_h \quad (h = 1, 2, 3).$$

Note that, by (6.26), (4.13), (6.10) and (2.7)–(2.8),

$$(6.34) \quad Z_1 \asymp KL \gg 1, \quad Z_2 \asymp L \gg 1, \quad Z_3 \asymp \frac{HN}{R^2} \gg 1.$$

Hence (6.30) implies

$$\sigma(\kappa) \ll \frac{L}{R} \sqrt{HNK} \sqrt{A(\kappa)} \sqrt{B(\kappa)} \quad (-2 < \kappa < 2).$$

Using this bound in (6.28) and applying the Cauchy–Schwarz inequality, we find

$$(6.35) \quad \sigma^2 \ll \frac{HNKL^2}{R^2} AB,$$

where, recall of definitions pertinent to $A(\kappa)$ reveals,

$$(6.36) \quad A \leq \int_{-1}^2 A(\kappa) d\kappa = 3 \int_0^1 A(\kappa) d\kappa = 3A_1 \quad (\text{say})$$

and where, by (6.29) and (6.32),

$$(6.37) \quad \begin{aligned} B &= \int_{-\infty}^{\infty} B(\kappa) d\kappa \\ &= \sum_i^{(Q)} \sum_{i'}^{(Q)} \int_{-\infty}^{\infty} \Phi_i(\kappa) \Phi_{i'}(\kappa) d\kappa \prod_{i=1}^3 \Lambda\left(\frac{y_h^{(i)} - y_h^{(i')}}{Y_h/(1 + Z_h)}\right) \\ &\leq \sum_i^{(Q)} \sum_{i'}^{(Q)} \frac{2H}{Q} \Lambda\left(\frac{\kappa_{0,i} - \kappa_{0,i'}}{2Q/H}\right) \prod_{i=1}^3 \Lambda\left(\frac{y_h^{(i)} - y_h^{(i')}}{Y_h/(1 + Z_h)}\right) \\ &= \frac{2H}{Q} B_1 \quad (\text{say}). \end{aligned}$$

7. The First Spacing Problem with a Dirichlet character factor. By (6.33) and (6.26), $X_1/Z_1 = 1/Y_1 = X_2/Z_2 = 1/Y_2 = 1/C_0 \leq 1$. Therefore, and by (6.23)–(6.25) and (6.31),

$$(7.1) \quad A(\kappa) = \sum_{\mathbf{p}} \sum'_{\mathbf{k}} \Lambda_{\mathbf{p},\mathbf{k}}(\kappa) \Pi_{\mathbf{p},\mathbf{k}}(\kappa),$$

where the ' following the summation sign indicates that the sum is over integer quadruplets, $\mathbf{p} = (l_1, \dots, l_4)$ and $\mathbf{k} = (k_1, \dots, k_4)$, satisfying both

$$(7.2) \quad l_1 + l_2 = l_3 + l_4$$

and

$$(7.3) \quad l_1 k_1 + l_2 k_2 = l_3 k_3 + l_4 k_4,$$

while

$$(7.4) \quad \Lambda_{\mathbf{p},\mathbf{k}}(\kappa) = \Lambda(Y_3(x_3^{\mathbf{j}}(\kappa) - x_3^{\mathbf{j}'}(\kappa))), \quad \Pi_{\mathbf{p},\mathbf{k}}(\kappa) = \Psi_{\mathbf{j}}(\kappa) \overline{\Psi_{\mathbf{j}'}}(\kappa),$$

with $\mathbf{j} = (k_1, k_4, l_1, l_4)$ and $\mathbf{j}' = (k_3, k_2, l_3, l_2)$.

The real solutions of (7.2) and (7.3) may be grouped into equivalence classes, called *families*, with two solutions, (\mathbf{p}, \mathbf{k}) and $(\mathbf{p}', \mathbf{k}')$ (say), belonging to the same class (family) if and only if

$$(7.5) \quad \mathbf{p}' = \mathbf{p}, \quad k'_i - k'_4 = k_i - k_4 \quad (i = 1, 2, 3).$$

We only need to consider those families that contain an integer point. Such families automatically contain an infinite number of integer points, since the system (7.2), (7.3) is invariant under the addition of the “diagonal” vector

(1, 1, 1, 1) to \mathbf{k} . By (6.24), (7.4) and (7.1), the contribution to $A(\kappa)$ coming from one such family, \mathcal{F} (say), is zero unless

$$(7.6) \quad \mathbf{p} = (l_1, \dots, l_4) \in (L_0, 2L_0]^4,$$

in which case it is

$$(7.7) \quad A_{\mathcal{F}}(\kappa) = \sum_t \Lambda_{\mathbf{p}, \mathbf{k}(t)}(\kappa) \Pi_{\mathbf{p}, \mathbf{k}(t)}(\kappa),$$

where $\mathbf{k}(t) = (k_1(t), \dots, k_4(t))$ and, for $i = 1, \dots, 4$, $k_i(t)$ denotes a real linear function of t , with

$$(7.8) \quad k_i(0) \in \mathbb{Z}, \quad k'_i(t) = 1 \quad \text{for } t \in \mathbb{R},$$

so that, as t runs over \mathbb{R} , the point $(\mathbf{p}, \mathbf{k}) = (\mathbf{p}, \mathbf{k}(t))$ runs over all the real solutions of (7.2), (7.3) in the family.

By (7.2), (6.24), (6.16), (6.2) and (6.12),

$$(7.9) \quad \Pi_{\mathbf{p}, \mathbf{k}(t)}(\kappa) = g(t - r\kappa) \chi(\alpha(t)) e(\Xi(t - r\kappa)),$$

with

$$(7.10) \quad g(t) = \begin{cases} \prod_{i=1}^4 (k_i^2(t) - l_i^2)^{-1/4} & \text{if } K_0 < k_i(t) \leq 2K_0 \text{ for } i = 1, \dots, 4, \\ 0 & \text{otherwise,} \end{cases}$$

$$(7.11) \quad \alpha(t) = \frac{(k_1(t) + l_1)(k_2(t) + l_2)(k_3(t) - l_3)(k_4(t) - l_4)}{(k_1(t) - l_1)(k_2(t) - l_2)(k_3(t) + l_3)(k_4(t) + l_4)}$$

and

$$(7.12) \quad \Xi(t) = - \sum_{\sigma=\pm 1} (\sigma u_1 + u_2) \sum_{i=1}^2 (\sqrt{k_i(t) + \sigma l_i} - \sqrt{k_{i+2}(t) + \sigma l_{i+2}})$$

(where $u_1, u_2 \in \mathbb{R}$), while

$$(7.13) \quad \Lambda_{\mathbf{p}, \mathbf{k}(t)}(\kappa) = \Lambda(Y_3 J(t - r\kappa)),$$

with

$$(7.14) \quad J(t) = \varrho_{l_1}(k_1(t)) + \varrho_{l_2}(k_2(t)) - \varrho_{l_3}(k_3(t)) - \varrho_{l_4}(k_4(t)).$$

By (7.7)–(7.14),

$$A_{\mathcal{F}}(\kappa) = \sum_t \delta(t - r\kappa) \chi(\alpha(t)),$$

where

$$(7.15) \quad \delta(t) = g(t) e(\Xi(t)) \Lambda(Y_3 J(t)).$$

Therefore, for $\tau \in \mathbb{R}$, $s \in \mathbb{Z}$,

$$A_{\mathcal{F}}((s + \tau)/r) = \sum_t \delta(t - s - \tau) \chi(\alpha(t)) = \sum_t \delta(t - \tau) \chi(\alpha(s + t))$$

and it follows that

$$\begin{aligned} \int_0^1 A_{\mathcal{F}}(\kappa) d\kappa &= \frac{1}{r} \sum_{s=0}^{r-1} \int_0^1 A_{\mathcal{F}}\left(\frac{s+\tau}{r}\right) d\tau \\ &= \frac{1}{r} \sum_{s=0}^{r-1} \int_0^1 \sum_t \delta(t-\tau) \chi(\alpha(s+t)) d\tau \\ &= \frac{1}{r} \sum_t \sum_{s=t}^{r-1+t} \chi(\alpha(s)) \int_{t-1}^t \delta(\tau) d\tau = \frac{1}{r} \mathcal{W}_{\mathcal{F}} \mathcal{D}_{\mathcal{F}}, \end{aligned}$$

where

$$(7.16) \quad \mathcal{W}_{\mathcal{F}} = \sum_{s \bmod r} \chi(\alpha(s)), \quad \mathcal{D}_{\mathcal{F}} = \int_{-\infty}^{\infty} \delta(t) dt.$$

Hence, and by (6.36),

$$(7.17) \quad A_1 = \frac{1}{r} \sum'_{\mathcal{F}} \mathcal{W}_{\mathcal{F}} \mathcal{D}_{\mathcal{F}},$$

where the summation is over all families \mathcal{F} containing at least one integer solution of (7.2), (7.3) and (7.6).

For the estimation of $\mathcal{W}_{\mathcal{F}}$ we have recourse to the following lemma, which, by virtue of permitting an exponential factor in the summand, covers some cases superfluous to this paper. Note that this is the ‘‘Lemma 9.3’’ referred to in [20, page 406].

LEMMA 7.1. *Let*

$$W_c = \sum_{x \bmod r} \chi\left(\frac{(x+a_1)\dots(x+a_4)}{(x+b_1)\dots(x+b_4)}\right) e\left(\frac{cx}{r}\right),$$

where $a_1, \dots, a_4, b_1, \dots, b_4$ and c are integers. Then either

$$|W_c| \leq 12\sqrt{r},$$

or, after some renumbering of a_1, \dots, a_4 and of b_1, \dots, b_4 , we are in one of the following four cases:

$$(7.18) \quad a_i \equiv b_i \pmod{r} \quad (i = 1, \dots, 4);$$

$$(7.19) \quad a_1 \equiv a_2 \not\equiv b_1 \equiv b_2 \pmod{r}, \quad a_i \equiv b_i \pmod{r} \quad (i = 3, 4);$$

$$(7.20) \quad a_1 \equiv a_2 \equiv a_3 \not\equiv b_1 \equiv b_2 \equiv b_3 \pmod{r}, \quad a_4 \equiv b_4 \pmod{r};$$

$$(7.21) \quad a_1 \equiv a_2 \not\equiv b_1 \equiv b_2 \not\equiv a_3 \equiv a_4 \not\equiv b_3 \equiv b_4 \not\equiv a_1 \pmod{r}.$$

Proof. We begin by writing

$$W_c = \sum_{x \bmod r} \chi((x+\alpha_1)\dots(x+\alpha_4)(x+\beta_1)^{r-2}\dots(x+\beta_4)^{r-2}) e(cx/r),$$

where $\alpha_1, \dots, \alpha_4$ and β_1, \dots, β_4 are integers satisfying

$$0 \leq \alpha_i, \beta_i \leq r - 1, \quad \alpha_i \equiv a_i \pmod{r}, \quad \beta_i \equiv b_i \pmod{r}$$

for $i = 1, \dots, 4$. Suppose $\alpha_i = \beta_j$. Then we have

$$(x + \alpha_i)(x + \beta_j)^{r-2} \equiv (x + \alpha_i)^{r-1} \equiv 1 \pmod{r},$$

except when $x \equiv -\alpha_i \pmod{r}$. Repeating this step if possible, we find that there exist permutations (A_1, \dots, A_4) and (B_1, \dots, B_4) of $(\alpha_1, \dots, \alpha_4)$ and $(\beta_1, \dots, \beta_4)$, respectively, and $n \in \{0, 1, \dots, 4\}$ such that:

$$\begin{aligned} A_i = B_i \quad (i = n + 1, \dots, 4), \quad A_i \neq B_j \quad (1 \leq i, j \leq n), \\ A_i \leq A_{i+1}, \quad B_i \leq B_{i+1} \quad (i = 1, \dots, n - 1) \end{aligned}$$

and

$$|W_c| \leq \left| \sum_{x \pmod{r}} \chi(\mathcal{P}(x)) e\left(\frac{cx}{r}\right) \right| + 4,$$

where

$$\mathcal{P}(x) = (x + A_1) \dots (x + A_n)(x + B_1)^{r-2} \dots (x + B_n)^{r-2}$$

($\mathcal{P}(x) = 1$ if $n = 0$).

If $n = 0$, then we are in case (7.18) and have nothing more to prove. For $n = 1, 2, 3, 4$, we first observe that $Y^d - \mathcal{P}(X)$ (where d is the order of χ) is absolutely irreducible unless there is a non-trivial factor dividing both d and each of the multiplicities of the roots of $\mathcal{P}(X)$ (see [28, Lemma 2C(iii)]). Secondly, the polynomial $Z^r - Z - X$ is, in all cases, absolutely irreducible. We conclude, by appealing to [28, Theorems 2C and 2G(ii)], that either we have

$$|W_c| \leq 8\sqrt{r} + 4 \leq 12\sqrt{r},$$

or we have $n = 2, A_1 = A_2, B_1 = B_2$ and $2 \mid d$, or we have $n = 3, A_1 = A_2 = A_3, B_1 = B_2 = B_3$ and $3 \mid d$, or we have $n = 4, A_1 = A_2, A_3 = A_4, B_1 = B_2, B_3 = B_4$ and $2 \mid d$. Recalling our conclusion for $n = 0$, these are exactly the alternatives listed by the lemma.

Referring to the definition of $\alpha(t)$ in (7.11), we can see that Lemma 7.1 is of no avail when dealing with a family containing integer solutions of (7.2) and (7.3) such that one of (7.18)–(7.21) holds with a_1, \dots, a_4 and b_1, \dots, b_4 given, modulo independent renumberings of both quadruples, by

$$(7.22) \quad a_i = k_i(0) + \lambda_i, \quad b_i = k_i(0) - \lambda_i \quad (i = 1, \dots, 4),$$

where

$$(7.23) \quad (\lambda_1, \dots, \lambda_4) = (l_1, l_2, -l_3, -l_4).$$

We boldly label such awkward families “exceptional”.

For any non-exceptional family \mathcal{O} the case $c = 0$ of Lemma 7.1 does apply:

$$(7.24) \quad \mathcal{W}_{\mathcal{O}} \ll \sqrt{r}.$$

Using this in (7.17) when possible, and the trivial universal bound of $O(r)$ otherwise, one obtains

$$(7.25) \quad A_1 \ll \frac{1}{\sqrt{r}} A'_1 + A_1^*,$$

where

$$(7.26) \quad A'_1 = \sum'_{\mathcal{F}} |\mathcal{D}_{\mathcal{F}}|, \quad A_1^* = \sum^*_{\mathcal{E}} |\mathcal{D}_{\mathcal{E}}|,$$

with the ' having the same effect as in (7.17) and the * signifying its sum is over the exceptional families \mathcal{E} containing at least one integer solution of (7.2), (7.3) and (7.6).

We now take some steps aimed at clearing away some non-essential complicating features of these sums (in (7.26)), before starting the estimation proper. Suppose that \mathcal{F} is a family containing an integer solution of (7.2), (7.3) and (7.6). Then, by (7.16), (7.15), (7.10), (7.6) and (6.9),

$$(7.27) \quad \begin{aligned} \mathcal{D}_{\mathcal{F}} &\ll K_0^{-2} \int_{-\infty}^{\infty} \Delta_{\mathcal{F}}(K_0, 1/Y_3; t) dt \\ &= K_0^{-2} \int_0^1 \left(\sum_{t=-\infty}^{\infty} \Delta_{\mathcal{F}}(K_0, 1/Y_3; t - \tau) \right) d\tau, \end{aligned}$$

where

$$(7.28) \quad \Delta_{\mathcal{F}}(X, \theta; t) = \begin{cases} 1 & \text{if } |J(t)| \leq \theta \text{ and } \mathbf{k}(t) \in (X, 2X]^4, \\ 0 & \text{otherwise.} \end{cases}$$

Here we should recall that, by (7.14) and (5.8),

$$(7.29) \quad J(t) = \frac{1}{3} \sum_{i=1}^4 ((k_i(t) + \lambda_i)^{3/2} - (k_i(t) - \lambda_i)^{3/2}),$$

where $\lambda_1, \dots, \lambda_4$ are as in (7.23), and that, by (6.26), (6.33), (6.34) and (6.9),

$$(7.30) \quad Y_3^{-1} = X_3/Z_3 = Z_3^{-1} C_0 L \sqrt{K} = \Gamma L_0 \sqrt{K_0} \quad (\text{say}),$$

where

$$(7.31) \quad \Gamma \asymp 1/Z_3 \asymp R^2/HN \ll 1.$$

LEMMA 7.2. *Let \mathcal{F} be a family of solutions of (7.2) and (7.3). Let $\Delta_{\mathcal{F}}(X, \theta; t)$ be as defined in (7.28)–(7.29), with $X, \theta > 0$, and suppose that*

$m \in \mathbb{N}$ and $\tau \in \mathbb{R}$. Then

$$\sum_{t=-\infty}^{\infty} \Delta_{\mathcal{F}}(X, \theta; t - \tau) = O(m) + \sum_{\substack{t=-\infty \\ t \equiv 0 \pmod{m}}}^{\infty} m \Delta_{\mathcal{F}}(X, \theta; t).$$

Proof. Apart from the factor $1/3$ the function $J(t)$ in (7.28)–(7.29) is the same function named as $M_{3/2}(\mathcal{F}, t)$ in [31]. Reprising an argument used in that paper, we observe that the product of the expression $x - J(t)$ together with (see (7.29)) all $2^8 - 1$ algebraically conjugate expressions, such as $x + J(t)$, is a polynomial of degree at most 384 in t , the coefficients of which polynomial are themselves polynomials, this time in x , that may, or may not, all have x as a factor, but certainly cannot have (as such a common factor) any other one of the real monic linear polynomials (this follows by considering the behaviour of $x - J(t)$, or any conjugate, in the limit as $t \rightarrow \infty$). As θ is real and non-zero, and since the constraints on \mathbf{k} in (7.28) restrict t to an interval where $J(t)$ is continuous, it follows from the observations made in the last paragraph that the support of $\Delta_{\mathcal{F}}(X, \theta; t)$ is a union of no more than 386 disjoint intervals. Let $\mathcal{I}_1, \dots, \mathcal{I}_n$ be the relevant intervals, and let $\mu(\mathcal{I}_1), \dots, \mu(\mathcal{I}_n)$ be their respective lengths. Then we have the equalities

$$\begin{aligned} \sum_{t=-\infty}^{\infty} \Delta_{\mathcal{F}}(X, \theta; t - \tau) &= \sum_{i=1}^n \sum_{\substack{t=-\infty \\ t-\tau \in \mathcal{I}_i}}^{\infty} 1 = \sum_{i=1}^n (\mu(\mathcal{I}_i) + O(1)) \\ &= O(1) + \sum_{i=1}^n \mu(\mathcal{I}_i) \end{aligned}$$

and

$$\begin{aligned} \sum_{\substack{t=-\infty \\ t \equiv 0 \pmod{m}}}^{\infty} \Delta_{\mathcal{F}}(X, \theta; t) &= \sum_{i=1}^n \sum_{\substack{s=-\infty \\ ms \in \mathcal{I}_i}}^{\infty} 1 = \sum_{i=1}^n (m^{-1} \mu(\mathcal{I}_i) + O(1)) \\ &= O(1) + \frac{1}{m} \sum_{i=1}^n \mu(\mathcal{I}_i). \end{aligned}$$

The lemma follows on using the second of these equalities to eliminate the last sum over i from the first equality.

Using the case $m = 1$ of Lemma 7.2 in (7.27) we have, after (7.30), (7.31),

$$\begin{aligned} \mathcal{D}_{\mathcal{F}} &\ll \frac{1}{K_0^2} \int_0^1 \left(O(1) + \sum_{t=-\infty}^{\infty} \Delta_{\mathcal{F}}(K_0, \Gamma L_0 \sqrt{K_0}; t) \right) d\tau \\ &= \frac{1}{K_0^2} \left(O(1) + \sum_{t=-\infty}^{\infty} \Delta_{\mathcal{F}}(K_0, \Gamma L_0 \sqrt{K_0}; t) \right), \end{aligned}$$

which, when applied to the sums of (7.26), yields

$$(7.32) \quad A_1' \ll K_0^{-2} \left(P(\Gamma, K_0, L_0) + \sum_{\mathcal{F}}' 1 \right),$$

$$(7.33) \quad A_1^* \ll K_0^{-2} \left(P_r^*(\Gamma, K_0, L_0) + \sum_{\mathcal{E}}^* 1 \right),$$

where, while it is now implicit that there must be at least one member (\mathbf{p}, \mathbf{k}) satisfying

$$(7.34) \quad K_0 < k_1, \dots, k_4 \leq 2K_0$$

in the relevant family (\mathcal{F} or \mathcal{E}), the significance of the $'$ and the $*$ is otherwise what it was in (7.17) and (7.26), and where, for $0 < \eta \ll 1$ and $0 < Y \leq X/2$, we let $P(\eta, X, Y)$ be the number of integer solutions (\mathbf{p}, \mathbf{k}) of (7.2) and (7.3) satisfying both

$$(7.35) \quad X < k_1, \dots, k_4 \leq 2X, \quad Y < l_1, \dots, l_4 \leq 2Y$$

and the condition

$$(7.36) \quad |\varrho_{l_1}(k_1) + \varrho_{l_2}(k_2) - \varrho_{l_3}(k_3) - \varrho_{l_4}(k_4)| \leq \eta Y \sqrt{X},$$

while defining $P_r^*(\eta, X, Y)$ to be the number of the same solutions present in exceptional families (for which see the paragraph containing (7.22) and (7.23)).

Turning to the estimation of the sums over \mathcal{F} and \mathcal{E} in (7.32) and (7.33), we begin with a lemma classifying the families for which the sets $\{k_1(t), k_2(t)\}$ and $\{k_3(t), k_4(t)\}$ are equal (the “symmetric families”) as either “totally symmetric” or “semi-symmetric”. All other families are to be styled “asymmetric”. The distinction between symmetric and asymmetric here is exactly the same as between the “trivial” and “non-trivial” families in [32], but we reserve the latter two adjectives for making the slightly different distinction that they marked in [31].

LEMMA 7.3. *There are only two sorts of symmetric families of solutions of (7.2), (7.3):*

(i) the “totally symmetric” families, for which

$$\{(k_1(t), l_1), (k_2(t), l_2)\} = \{(k_3(t), l_3), (k_4(t), l_4)\};$$

(ii) the “semi-symmetric” families, for which

$$k_1(t) = k_2(t) = k_3(t) = k_4(t), \quad \{l_1, l_2\} \neq \{l_3, l_4\}.$$

Proof. This is evident from the analysis into cases given in the proof of [32, Lemma 1].

LEMMA 7.4. *Let $X, Y > 0$. The number of totally symmetric families containing at least one integer solution of (7.2), (7.3) and (7.35) is $O(XY^2)$.*

Proof. In order to fully determine a family it suffices to know the values of l_1, \dots, l_4 and $k_1(t) - k_2(t)$, $k_3(t) - k_2(t)$ and $k_4(t) - k_2(t)$ (all are quantities independent of t). As we have $\{k_3(t) - k_2(t), k_4(t) - k_2(t)\} = \{k_1(t) - k_2(t), k_2(t) - k_2(t)\} = \{k_1(t) - k_2(t), 0\}$, and also $\{l_3, l_4\} = \{l_1, l_2\}$, it follows that, after one has chosen integer values for l_1, l_2 (there are $O(Y^2)$ choices) and $k_1(t) - k_2(t)$ (another $O(X)$ choices), there remain just $O(1)$ choices for $l_3, l_4, k_3(t) - k_2(t)$ and $k_4(t) - k_2(t)$, and hence for the family.

LEMMA 7.5. *Let $X, Y > 0$. The number of semi-symmetric families containing at least one integer solution of (7.2) and (7.35) is $O(Y^3)$.*

Proof. The bound is just the number of choices for integers l_1, l_2, l_3 satisfying (7.35). By (7.2) these values determine that of l_4 , which completes determination of the family, since (given that it is singularly trivial) all differences of the form $k_i(t) - k_j(t)$ are zero.

LEMMA 7.6. *Let $X, Y > 0$. The number of asymmetric families containing at least one integer solution of (7.2), (7.3) and (7.35) is $O(X^2Y^2)$.*

Proof. See [32, Lemma 3].

LEMMA 7.7. *The sums of (7.32), (7.33) satisfy*

$$\sum_{\mathcal{E}}^* 1 \leq \sum_{\mathcal{F}}^l 1 \ll K_0L_0^2 + L_0^3 + K_0^2L_0^2 \ll K_0^2L_0^2$$

and, as a consequence of this, we have

$$A'_1 \ll K_0^{-2}P(\Gamma, K_0, L_0), \quad A^*_1 \ll K_0^{-2}P_r^*(\Gamma, K_0, L_0).$$

Proof. The first two bounds follow by (6.9) from Lemmas 7.3–7.6, with $X = K_0$ and $Y = L_0$ (the first sum is a subsum of the second sum, in which the family \mathcal{F} must contain at least one integer solution of (7.2), (7.3), (7.6) and (7.34)). The conclusion follows since both $P(\Gamma, K_0, L_0)$ and $P_r^*(\Gamma, K_0, L_0)$ count no less than $K_0^2L_0^2$ totally symmetric integer solutions of (7.2), (7.3), (7.6) and (7.34) (see Lemma 7.3(i), (7.11) and (7.18) of Lemma 7.1).

LEMMA 7.8. *Let $\eta > 0$ and $0 < Y \leq X/4$. Then*

$$P(\eta, X, Y) \leq B(\eta + O(Y^2/X^2), X, Y),$$

where $B(\delta, X, Y)$ is the number of integer solutions (\mathbf{p}, \mathbf{k}) of (7.2) and (7.3) satisfying (7.35) and (in place of (7.36)) the condition

$$(7.37) \quad |l_1\sqrt{k_1} + l_2\sqrt{k_2} - l_3\sqrt{k_3} - l_4\sqrt{k_4}| \leq \delta Y\sqrt{X}.$$

Proof. Let $i \in \{1, 2, 3, 4\}$. By (7.35) and the hypothesis that $0 < Y \leq X/4$ we may suppose that $0 < l_i/k_i < 2Y/X \leq 1/2$. It follows that we may

make use of the obvious binomial expansion of (5.8),

$$\begin{aligned} \varrho_{l_i}(k) &= \frac{1}{3} k^{3/2} \sum_{r=0}^{\infty} \binom{3/2}{r} \left(\left(\frac{l_i}{k} \right)^r - \left(-\frac{l_i}{k} \right)^r \right) \\ &= l_i k^{1/2} \sum_{s=0}^{\infty} \frac{2}{3} \binom{3/2}{2s+1} \left(\frac{l_i}{k} \right)^{2s}, \end{aligned}$$

from which we deduce

$$\varrho_{l_i}(k_i) = (1 + O(Y^2/X^2)) l_i \sqrt{k_i}.$$

The lemma follows by substituting this in (7.36): just observe that $l_i \sqrt{k_i} \asymp Y \sqrt{X}$ for $i = 1, \dots, 4$.

LEMMA 7.9. *Let $0 \leq \delta \leq 1$ and $X, Y > 0$. Then the number of integer solutions of (7.2), (7.3), (7.35) and (7.37) that lie in asymmetric families is $O((X^2Y^2 + \delta X^3Y^2) \log(1 + X))$.*

Proof. By [32, Lemmas 2 and 3] and the first displayed inequality below the middle of [32, page 59], one can see that it follows from the first of the two bounds given in [32, Theorem, page 64] that the number of integer solutions of (7.2), (7.3), (7.35) and (7.37) is

$$\begin{aligned} &\ll \sum_{1 \leq d < X} \sum_{1 \leq e < 2Y} \left(\frac{X^2Y^2}{de^2} + d \left(\left(\frac{X}{d} \right)^2 \left(\frac{Y}{e} \right) \log^3 \left(2 + \frac{Y}{e} \right) \right. \right. \\ &\quad \left. \left. + \delta \left(\frac{X}{d} \right)^3 \left(\frac{Y}{e} \right)^2 \log \left(2 + \frac{X}{d} \right) \right) \right) \end{aligned}$$

and the lemma follows.

LEMMA 7.10 (Lemma 13.3.1 of [14]). *Let $\eta > 0$ and $0 < Y \leq X/4$. Then*

$$P(\eta, X, Y) \ll (X^2Y^2 + XY^4 + \eta X^3Y^2) \log(1 + X).$$

Proof. This follows by an application of Lemma 7.8, in conjunction with Lemmas 7.9, 7.4, 7.5 and the trivial observation that there cannot be more than X integer solutions of (7.35) in any one family.

While Lemma 7.10 is entirely satisfactory when $Y^2 \ll X$, we shall need better bounds for other cases. To this end we recast the problem of bounding $P(\eta, X, Y)$ as a variant of the problem on semicubical powers treated in [2] and [31]. The first step is to observe that (7.2), (7.3) and (see (5.8)) (7.36) are merely the cases $\beta = 1, 3/2, 2$ of the inequality

$$\left| \sum_{i=1}^4 ((k_i + \lambda_i)^\beta - (k_i - \lambda_i)^\beta) \right| \leq 12(2 - \beta)(\beta - 1)\eta Y \sqrt{X},$$

where the λ_i 's are as in (7.23). By making the substitutions

$$(7.38) \quad k_i + \lambda_i = g_i, \quad k_i - \lambda_i = h_i \quad (i = 1, \dots, 4),$$

we find that the relevant three cases of the last inequality are equivalent to the conditions that

$$(7.39) \quad \sum_{i=1}^4 (g_i^2 - h_i^2) = \sum_{i=1}^4 (g_i - h_i) = 0$$

and

$$(7.40) \quad \left| \sum_{i=1}^4 (g_i^{3/2} - h_i^{3/2}) \right| \leq 3\Delta X^{3/2},$$

where

$$(7.41) \quad \Delta = \frac{Y}{X} \eta.$$

If we suppose that $0 < Y \leq X/12$, then (7.23), (7.35) and (7.38) imply

$$(7.42) \quad 5X/6 < g_1, \dots, g_4, h_1, \dots, h_4 \leq 13X/6$$

and

$$(7.43) \quad 2Y < |g_i - h_i| \leq 4Y \quad (i = 1, \dots, 4).$$

Note also that, since $-\lambda_i \equiv \lambda_i \pmod{2}$, we will always have

$$(7.44) \quad g_i \equiv h_i \pmod{2} \quad (i = 1, \dots, 4).$$

From these observations one may deduce the following lemma.

LEMMA 7.11. *Let $\eta > 0$ and $0 < Y \leq X/12$. If Δ is given by (7.41), then for some X_1 satisfying*

$$(7.45) \quad 5X/6 \leq X_1 \leq 2X$$

one has

$$P(\eta, X, Y) \ll I_8(X_1, Y, \Delta) \ll P(\eta, X_1, Y),$$

where $I_8(X_1, Y, \Delta)$ is the number of integer solutions (\mathbf{g}, \mathbf{h}) of (7.39), (7.43) and (7.44) satisfying both

$$(7.46) \quad \left| \sum_{i=1}^4 (g_i^{3/2} - h_i^{3/2}) \right| \leq \Delta X_1^{3/2}$$

and

$$(7.47) \quad X_1 < g_1, h_1, \dots, g_4, h_4 \leq 2X_1.$$

Proof. As the substitution (7.38) gives a one-to-one correspondence between the (\mathbf{p}, \mathbf{k}) 's and the (\mathbf{g}, \mathbf{h}) 's (see (7.6) and (7.23)), and since $Y \leq X/12$, it follows that we may establish the claimed upper bound on

$P(\eta, X, Y)$ by showing that, for some X_1 satisfying (7.45), the number of integer solutions (\mathbf{g}, \mathbf{h}) of (7.39), (7.40) and (7.42)–(7.44) is $O(I_8(X_1, Y, \Delta))$. Since the Fourier transform of $\text{sinc}^2(x) = (\sin(\pi x)/\pi x)^2$ equals $\Lambda(x) = \max(0, 1 - |x|)$, which is bounded below by $1/2$ on the interval $[-1/2, 1/2]$, this number of integer solutions is a lower bound for the quantity

$$12\Delta \int_{-\infty}^{\infty} \int_0^1 \int_0^1 \text{sinc}^2(6\Delta\gamma) S^4(\alpha, \beta, \gamma) d\alpha d\beta d\gamma,$$

where $S(\alpha, \beta, \gamma)$ is the sum

$$\sum_{\substack{5X/6 < g, h \leq 13X/6 \\ 2Y < |g-h| \leq 4Y \\ h \equiv g \pmod{2}}} e(\alpha(g^2 - h^2) + \beta(g - h) + \gamma((g/X)^{3/2} - (h/X)^{3/2})).$$

Therefore, the observations that $\text{sinc}^2(x) \leq \text{sinc}^2(x/n)$, for $x \in \mathbb{R}$ and $n \in \mathbb{N}$, and that

$$S(\alpha, \beta, \gamma) = \sum_{m=5}^{12} T_m(\alpha, \beta, \gamma_1),$$

where $\gamma_1 = X^{-3/2}\gamma$ and $T_m(\alpha, \beta, \gamma_1)$ is the sum

$$\sum_{g=[mX/6]+1}^{[(m+1)X/6]} \sum_{\substack{5X/6 < h \leq 13X/6 \\ 2Y < |h-g| \leq 4Y \\ h \equiv g \pmod{2}}} e(\alpha(g^2 - h^2) + \beta(g - h) + \gamma_1(g^{3/2} - h^{3/2})),$$

show that a simple application of the Cauchy–Schwarz inequality enables us to infer the existence of $m \in \{5, \dots, 12\}$ for which

$$P(\eta, X, Y) \ll \Delta \int_{-\infty}^{\infty} \int_0^1 \int_0^1 \text{sinc}^2(\Delta\gamma/2) |T_m(\alpha, \beta, X^{-3/2}\gamma)|^4 d\alpha d\beta d\gamma.$$

As $|T|^4 = T^2(\overline{T})^2$, the last quantity may be evaluated by multiplying out the product of four sums (two T_m 's and two conjugate sums) and then integrating term by term. Providing the labelling of variables of summation is done appropriately (with “g” and “h” exchanging places in the two conjugate sums), the result appears as a sum trivially bounded above by the number of integer solutions (\mathbf{g}, \mathbf{h}) of (7.39), (7.43) and (7.44) satisfying

$$\left| \sum_{i=1}^4 (g_i^{3/2} - h_i^{3/2}) \right| \leq \frac{\Delta}{2} X^{3/2}$$

and, for $i = 1, \dots, 4$,

$$\max(5X/6, mX/6 - 4Y) < g_i, h_i \leq (m + 1)X/6 + 4Y.$$

Since $1/2 < (5/6)^{3/2}$ and

$$\frac{(m + 1)X/6 + 4Y}{\max(5X/6, mX/6 - 4Y)} \leq \frac{m + 3}{\max(5, m - 2)} \leq 2,$$

it follows that (7.46) and (7.47) hold with $X_1 = \max(5X/6, mX/6 - 4Y)$ (satisfying (7.45)). This completes the proof of the lemma's upper bound for $P(\eta, X, Y)$.

The upper bound claimed for $I_8(X_1, Y, \Delta)$ follows directly from a reversal of the substitutions (7.38). One only has to observe that $Y \leq X_1/10$ and that, given (5.8), (7.23), (7.41) and (7.45), all except one of the required conditions on (\mathbf{p}, \mathbf{k}) are consequences of (7.39), (7.43), (7.44), (7.46) and (7.47). The single deficiency is that one obtains just $Y < |\lambda_1|, \dots, |\lambda_4| \leq 2Y$ instead of the desired condition $Y < l_1, \dots, l_4 \leq 2Y$. This is easily remedied, since the equation $\lambda_1 + \dots + \lambda_4 = 0$ (which is (7.2)) enables us to deduce from the bounds on the λ_i that exactly two of them are positive. There are 6 possible cases (depending on which two of the four λ_i are positive). For any one case there is a fixed permutation of the subscripts $1, \dots, 4$ that, when applied to both $\lambda_1, \dots, \lambda_4$ and k_1, \dots, k_4 , leads to the condition $Y < l_1, \dots, l_4 \leq 2Y$ being recovered, but does not disturb any of the other conditions. Therefore there are at most $P(\eta, X_1, Y)$ solutions counted by $I_8(X_1, Y, \Delta)$ in each of the 6 cases, so the bound claimed for $I_8(X_1, Y, \Delta)$ is obtained and the proof of the lemma is complete.

The fact that (7.38) gives a correspondence that is both one-to-one and onto implies that the partitioning of the set of (\mathbf{p}, \mathbf{k}) into equivalence classes (families) by the relation (7.5) will carry over to the set of (\mathbf{g}, \mathbf{h}) counted by $I_8(X_1, Y, \Delta)$ in Lemma 7.11. The equivalence relation (7.5) holds if and only if $\mathbf{p}' = \mathbf{p}$ and $\mathbf{k}' - \mathbf{k}$ lies on the diagonal of \mathbb{R}^4 . Since translation of \mathbf{k} parallel to the diagonal of \mathbb{R}^4 , with $\mathbf{p} = (l_1, \dots, l_4) = (\lambda_1, \lambda_2, -\lambda_3, -\lambda_4)$ fixed, corresponds to translation of (\mathbf{g}, \mathbf{h}) parallel to the diagonal in \mathbb{R}^8 (see (7.38)) it is clear that, as in [31], a family of the points (\mathbf{g}, \mathbf{h}) is a set of real solutions of (7.39) of the form $\{(g_1 + t, \dots, g_4 + t, h_1 + t, \dots, h_4 + t) \mid t \in \mathbb{R}\}$. We follow [31] in defining a family to be *trivial* if and only if it contains (\mathbf{g}, \mathbf{h}) such that the coordinates (g_1, \dots, g_4) of \mathbf{g} are equal to some permutation of those of \mathbf{h} .

LEMMA 7.12. *Let $X_1, Y > 0$. The number of trivial families containing an integer solution (\mathbf{g}, \mathbf{h}) of (7.43) and (7.47) is $O(X_1 Y^2)$.*

Proof. First note that (7.43) and (7.47) do not have integer solutions unless $X_1, Y > 1/4$, so this may be assumed to hold.

Each of the families we are counting will contain a unique member $(\mathbf{g}, \mathbf{h}) \in \mathbb{Z}^8$ such that $h_4 = 0$. This member will satisfy (7.43) and, in place of (7.47), the condition that $\max(|g_1|, \dots, |g_4|) \leq X_1$ (as the family is trivial

(h_1, \dots, h_4) will be a permutation of (g_1, \dots, g_4) . To count the families we only have to count these members.

Since $h_4 = 0$ the number of these members with $\max(|g_1|, \dots, |g_4|) \leq 8Y$ is $O((\min(X_1, Y))^3) = O(X_1 Y^2)$. Suppose that there exists $x \in \{g_1, \dots, g_4\}$ with $|x| > 8Y$. Then $\{0, x\} \subseteq \{g_1, \dots, g_4\} = \{h_1, \dots, h_4\}$. It follows by (7.43) that there exist $y, z \in \{g_1, \dots, g_4\}$ with $2Y < |y - 0| \leq 4Y$ and $2Y < |z - x| \leq 4Y$. Plainly $0 \neq |y| \leq 4Y$, $z \neq x$ and (since $|x| > 8Y$) $|z| > 8Y - 4Y = 4Y$. The four numbers $0, x, y, z$ are distinct. It follows that $\{g_1, \dots, g_4\} = \{h_1, \dots, h_4\} = \{0, x, y, z\}$. To complete the proof of the lemma we have merely to observe that there are $O(X_1)$ choices for x with $|x| \leq X_1$, then $O(Y)$ choices for y with $|y| \leq 4Y$ and $O(Y)$ choices for z with $|z - x| \leq 4Y$, followed by $O(1)$ choices of two permutations of $(0, x, y, z)$ to form \mathbf{g} and \mathbf{h} .

LEMMA 7.13. *The number of families containing an integer solution of (7.39), (7.43), (7.44) and (7.47) is $O(X_1^2 Y^2)$.*

Proof. We refer to the proof of the upper bound on $I_8(X_1, Y, \Delta)$ in Lemma 7.11, and to the comments which follow that proof, where it was established that there is a 6-to-1 mapping from the set of integer solutions (\mathbf{g}, \mathbf{h}) of (7.39), (7.43), (7.44) and (7.47) into the set of integer solutions (\mathbf{p}, \mathbf{k}) of (7.2), (7.3) and (7.35). This mapping maps families to families and, since one member of a family determines the whole family, it corresponds to a 6-to-1 mapping from the set of families of points (\mathbf{g}, \mathbf{h}) into the set of families of points (\mathbf{p}, \mathbf{k}) . Therefore, by applying Lemmas 7.4–7.6, we may conclude that the number of families containing an integer solution of (7.39), (7.43), (7.44) and (7.47) is at most $O(X_1 Y^2) + O(Y^3) + O(X_1^2 Y^2)$ multiplied by 6. As one may assume that $X_1 \geq Y + 1$, the lemma follows.

LEMMA 7.14. *Let $\Delta, X_1, Y > 0$. Suppose that (\mathbf{g}, \mathbf{h}) is a member of a non-trivial family \mathcal{F} that contains $\mu_{\mathcal{F}}$ integer solutions of (7.39), (7.43), (7.46) and (7.47). Then at least one of the eight products*

$$(g_1 - h_i) \dots (g_4 - h_i), \quad (h_1 - g_i) \dots (h_4 - g_i) \quad (i = 1, \dots, 4)$$

is non-zero, and either $\mu_{\mathcal{F}} = O(1)$, or all these products are $O(\Delta X_1^5 / \mu_{\mathcal{F}})$.

Proof. The first assertion is taken from [31, Lemma 7]. To verify the second assertion, we observe that, if $\mu_{\mathcal{F}}$ is sufficiently large, then the argument given as proof of [31, Theorem 1] shows that the conditions (3.6), (3.7) of [31, Lemma 9] must hold with the H and δ there satisfying $H = 2X_1$ and $\delta \asymp (X_1 / \mu_{\mathcal{F}}) \Delta$ (see [31, page 139]). By [31, (3.5)] the last mentioned conditions, taken with (7.39), permit the application of [31, Lemma 8] with the parameters K, N, r there being given by $K = \sqrt{2X_1}$, $N = 4$ and $r = 2$ (the x_i and y_i there are of the form $\sqrt{g_i + t}$ or $\sqrt{h_i + t}$). This yields the desired

bound on $(g_1 - h_i) \dots (g_4 - h_i)$ for $i = 1, \dots, 4$. The case of the other four products follows by symmetry.

LEMMA 7.15. *Let $\mathbf{u} \in \mathbb{Z}^4$ and $i \in \{1, 2, 3, 4\}$ be given. The number of families of integer solutions (\mathbf{g}, \mathbf{h}) of (7.39) that have either $(g_1 - h_i, \dots, g_4 - h_i)$, or $(h_1 - g_i, \dots, h_4 - g_i)$, equal to \mathbf{u} does not exceed $12d^*(F(\mathbf{u}))$, where*

$$d^*(n) = \begin{cases} \sum_{d|n} 1 & \text{if } n \neq 0, \\ 1 & \text{if } n = 0, \end{cases}$$

and

$$F(\mathbf{u}) = u_1^2 + \dots + u_4^2 - u_1u_2 - u_1u_3 - u_1u_4 - u_2u_3 - u_2u_4 - u_3u_4.$$

Proof. This follows by using [31, Lemma 1] to count the number of choices for the three not identically zero factors x_1, x_2, x_3 either from the expression $(h_1 - h_i) \dots (h_4 - h_i)$, or from the expression $(g_1 - g_i) \dots (g_4 - g_i)$.

LEMMA 7.16. *Let $\eta > 0$ and $0 < Y \leq X/12$. Then*

$$P(\eta, X, Y) \ll X^2Y^2 + \eta X^4Y \log^6(1 + X).$$

Proof. By Lemmas 7.11, 7.12 and the trivial upper bound, X_1 , for the number of integer solutions of (7.47) lying in one family, we only have to count the integer solutions of (7.39), (7.43), (7.44), (7.46) and (7.47) from non-trivial families \mathcal{F} (for a given X_1 satisfying (7.45)). Let $\mu_{\mathcal{F}}$ be as in Lemma 7.14. By Lemma 7.13 there are at most $O(X_1^2Y^2) = O(X^2Y^2)$ integer solutions in those families with $\mu_{\mathcal{F}} \ll 1$, and all that remains is to treat the cases where all the eight products in that lemma are $O(\Delta X_1^5/\mu_{\mathcal{F}})$. We may assume

$$(7.48) \quad 0 \neq (g_1 - h_4) \dots (g_4 - h_4) \ll \Delta X_1^5/\mu_{\mathcal{F}},$$

since the other seven cases are all similar. By (7.47) we may also assume $\Delta \ll 1$ in (7.46).

Consider now the subcase where $M \leq \mu_{\mathcal{F}} < 2M$, for some given $M \gg 1$. By Lemma 7.15, (7.43) and (7.48), the number of such families is

$$(7.49) \quad \ll \sum_{0 \neq u_1 u_2 u_3 \ll \Delta X_1^5/MY} \sum_{|u_4| \asymp Y} d^*(F(u_1, \dots, u_4)).$$

We can divide this sum into $O(\log^3 X_1)$ sums of the form

$$(7.50) \quad \sum_{U_1 \leq |u_1| < 2U_1} \dots \sum_{U_4 \leq |u_4| < 2U_4} d^*(F(u_1, \dots, u_4)),$$

where $U_4 \asymp Y$. Supposing that $\max(U_1, \dots, U_4) = U_i$ (say), it is a simple exercise to show that, if $U_j \leq u_j < 2U_j$ for $j \in \{1, 2, 3, 4\} \setminus \{i\}$, then

$$\sum_{U_i \leq |u_i| < 2U_i} d^*(F(u_1, \dots, u_4)) \ll U_i \log^3 U_i.$$

It follows that the sum (7.50) is $O(U_1 \dots U_4 \log^3 X_1)$. Summing this over values of U_1, \dots, U_4 which are integer powers of 2 satisfying $1 \ll U_1 U_2 U_3 \ll \Delta X_1^5 / MY$ and $U_4 \asymp Y$ we find that the sum (7.49) is

$$\begin{aligned} &\ll \sum_{U_4 \asymp Y} \sum_{1 \leq U_1, U_2 \ll X_1} \sum_{U_3 \ll \Delta X_1^5 / MY U_1 U_2} U_1 \dots U_4 \log^3 X_1 \\ &\ll \sum_{1 \leq U_1, U_2 \ll X_1} \frac{\Delta X_1^5}{M} \log^3 X_1 \ll \frac{\Delta X^5}{M} \log^5 X. \end{aligned}$$

Since each family contains less than $2M$ solutions, we conclude that there are just $O(\Delta X^5 \log^5 X)$ solutions of (7.39), (7.43), (7.44), (7.46) and (7.47) from the families with $M \leq \mu_{\mathcal{F}} < 2M$. As $\Delta = \eta Y / X$ in (7.41), and since $O(\log X)$ intervals $[M, 2M)$ are sufficient to cover the range $[2, X_1]$ for $\mu_{\mathcal{F}}$, we obtain the claimed result of the lemma.

The power of the logarithm in Lemma 7.16 can be halved, as was shown (for $Y \asymp X$) in [31]. We accept this small unnecessary loss as it is desirable to have the simpler proof here, where it may serve as an uncomplicated starting point for a useful generalisation (see Lemma 7.23 below).

The upper bounds of Lemmas 7.16 and 7.10 should be compared to the lower bound

$$P(\eta, X, Y) \gg X^2 Y^2 + \eta X^3 Y^2 = X^2 Y^2 + \Delta X^4 Y,$$

which is an easy corollary of the Cauchy–Schwarz inequality. The comparison makes it clear that Lemma 7.16 is only a bounded power of $\log X$ weaker than the best-possible bound for the case $\Delta \ll Y^2 / X^3$. The same observation holds true of Lemma 7.10 in respect of both the case where $Y \ll \sqrt{X}$ and the case where $\Delta \gg Y^3 / X^3$. In the case that remains (complementary to the union of the three just discussed) we shall apply the following lemma, established in [33] by an argument involving an iteration adapted to the conditions (7.39), (7.43), (7.46) and (7.47).

LEMMA 7.17. *Let $\eta > 0$, $\sqrt{X} \ll Y \leq X/12$ and $Y^2 / X^3 \ll \Delta \ll Y^3 / X^3$, where $\Delta = (Y/X)\eta$ (as in (7.41)). Then*

$$\begin{aligned} &P(\eta, X, Y) \\ &\ll X^\epsilon \min(X^{3/2} Y^3 + \Delta^{1/3} X^2 Y^3, \Delta X^{11/2} / Y + \Delta X^{14/3} Y^{1/3}) + X^2 Y^2. \end{aligned}$$

Proof. This follows by way of Lemma 7.11 from [33, Theorem 1.1].

LEMMA 7.18. *Let $\Delta = \Gamma L_0 / K_0$. Then*

$$\begin{aligned} P(\Gamma, K_0, L_0) &\ll K_0^2 L_0^2 \log^6 K_0 + K_0^\epsilon \min(\Delta^{1/3} K_0^2 L_0^3, \Delta K_0^{14/3} L_0^{1/3}) \\ &\quad + \Gamma K_0^3 L_0^2 \log K_0 + K_0^\epsilon \min(K_0^{3/2} L_0^3, \Delta K_0^{11/2} L_0^{-1}) \end{aligned}$$

with

$$(7.51) \quad \Gamma \asymp \left(\frac{Q}{R}\right)^2 \frac{r^2}{K_0 L_0}, \quad \Delta \asymp \left(\frac{Q}{R}\right)^2 \frac{r^2}{K_0^2} \asymp \left(\frac{Qr}{RK_0}\right)^2 \asymp \frac{R^2}{N^2}.$$

Proof. By (5.2) and (6.9), $r \geq 1$, $Q \geq R$ and $16L_0 \leq K_0$, so it follows by the discussion preceding Lemma 7.17 that we only have to consider the cases where that lemma applies, with $\eta = \Gamma$, $X = K_0$ and $Y = L_0$. Examining Lemma 7.17, we can observe that $\Delta X^{11/2} Y^{-1} \ll X^{3/2} Y^3$ if and only if $\Delta \ll (Y/X)^4$, and that the same is true of the inequality $\Delta X^{14/3} Y^{1/3} \ll \Delta^{1/3} X^2 Y^3$. Therefore the bound claimed by this lemma follows from that given by Lemma 7.17. The bounds given for Γ and Δ in (7.51) are derived from (7.31), (6.9) and (6.10).

Because of the cancellation from the character (see (7.24), (7.25)) we can accept a weaker bound for $P(\Gamma, K_0, L_0)$ than we can for $P_r^*(\Gamma, K_0, L_0)$. In order to bound the latter quantity effectively we shall need to use the fact that, in light of (7.38) and the remarks around (7.22) and (7.23), an integer solution (\mathbf{g}, \mathbf{h}) of (7.39) comes from an exceptional family if and only if there exists a permutation $\mathbf{a} = (a_1, \dots, a_4)$ of (g_1, \dots, g_4) and a permutation $\mathbf{b} = (b_1, \dots, b_4)$ of (h_1, \dots, h_4) such that (\mathbf{a}, \mathbf{b}) satisfies one of the congruence conditions (7.18)–(7.21). These congruence conditions have in common the property that, given any one solution (\mathbf{a}, \mathbf{b}) , there is a union of four or less residue classes modulo r that contains all the eight integers $a_1, b_1, \dots, a_4, b_4$. It follows that \mathbf{g}, \mathbf{h} can come from an exceptional family only if there exist $u_1, \dots, u_4 \in \{0, \dots, r - 1\}$ and some pair of functions $s, t : \{1, \dots, 4\} \rightarrow \{1, \dots, 4\}$ such that, for $i = 1, \dots, 4$,

$$(7.52) \quad g_i \equiv u_{s(i)} \pmod{r}, \quad h_i \equiv u_{t(i)} \pmod{r}.$$

Therefore, appealing (as in the proof of Lemma 7.11) to the one-to-one nature of the correspondence given by (7.38), one certainly has

$$(7.53) \quad P_r^*(\eta, X, Y) \leq \sum_{u_1=0}^{r-1} \dots \sum_{u_4=0}^{r-1} \sum_{s,t} I_8^*(X, Y, \Delta; r, \mathbf{u}, s, t) \\ \leq 4^8 \max_{s,t} \left(\sum_{u_1=0}^{r-1} \dots \sum_{u_4=0}^{r-1} I_8^*(X, Y, \Delta; r, \mathbf{u}, s, t) \right),$$

where $I_8^*(X, Y, \Delta; r, \mathbf{u}, s, t)$ is the number of integer solutions of (7.39), (7.40), (7.42)–(7.44) and (7.52).

LEMMA 7.19. *Let $\eta > 0$ and $0 < Y \leq X/12$. If Δ is as in (7.41), then for some X_1 satisfying (7.45),*

$$P_r^*(\eta, X, Y) \ll r^3 I_8'(X_1, Y, \Delta, r) + r^2 I_8''(X_1, Y, \Delta, r),$$

where $I_8^!(X_1, Y, \Delta, r)$ is the number of integer solutions of (7.39), (7.43), (7.44), (7.46) and (7.47) satisfying

$$(7.54) \quad g_1 \equiv h_1 \equiv g_2 \equiv h_2 \equiv g_3 \equiv h_3 \equiv g_4 \equiv h_4 \pmod{r},$$

while $I_8''(X_1, Y, \Delta, r)$ counts the number of the same solutions that instead satisfy

$$(7.55) \quad g_1 \equiv g_2 \equiv h_3 \equiv h_4 \not\equiv h_1 \equiv h_2 \equiv g_3 \equiv g_4 \pmod{r}.$$

Proof. Arguing as in the proof of Lemma 7.11, we find that the number $I_8^*(X, Y, \Delta; r, \mathbf{u}, s, t)$ in (7.53) is a lower bound for

$$12\Delta \int_{-\infty}^{\infty} \int_0^1 \int_0^1 \text{sinc}^2(6\Delta\gamma) \left(\prod_{i=1}^4 S_i(\alpha, \beta, \gamma) \right) d\alpha d\beta d\gamma,$$

where $S_i(\alpha, \beta, \gamma)$ is similar to the sum $S(\alpha, \beta, \gamma)$ (from the proof of Lemma 7.11), the only difference being that, for $S_i(\alpha, \beta, \gamma)$, there are two additional conditions,

$$(7.56) \quad g \equiv u_{s(i)} \pmod{r}, \quad h \equiv u_{t(i)} \pmod{r},$$

further restricting summation over g and h . By the arithmetic-geometric mean inequality,

$$\prod_{i=1}^4 |S_i(\alpha, \beta, \gamma)| \leq \frac{1}{4} \sum_{j=1}^4 |S_j(\alpha, \beta, \gamma)|^4.$$

Using subdivision and the Cauchy–Schwarz inequality, as in the proof of Lemma 7.11, one finds that

$$|S_j(\alpha, \beta, \gamma)|^4 \leq 8^3 \sum_{m=5}^{12} |T_{j,m}(\alpha, \beta, X^{-3/2}\gamma)|^4,$$

where $T_{j,m}(\alpha, \beta, \gamma_1)$ is the subsum of the sum $T_m(\alpha, \beta, \gamma_1)$ (from the proof of Lemma 7.11) in which the variables of summation, g, h , satisfy the extra constraints in (7.56). Finally (see the proof of Lemma 7.11 again), for $j = 1, \dots, 4$ and $m = 5, \dots, 12$,

$$\begin{aligned} & \Delta \int_{-\infty}^{\infty} \int_0^1 \int_0^1 \text{sinc}^2(6\Delta\gamma) |T_{j,m}(\alpha, \beta, X^{-3/2}\gamma)|^4 d\alpha d\beta d\gamma \\ & \leq \Delta \int_{-\infty}^{\infty} \int_0^1 \int_0^1 \text{sinc}^2(\Delta\gamma/2) T_{j,m}^2 \overline{T}_{j,m}^2(\alpha, \beta, X^{-3/2}\gamma) d\alpha d\beta d\gamma \\ & \ll I_8^{\sim}(X_{1,m}, Y, \Delta; r, u_{s(j)}, u_{t(j)}), \end{aligned}$$

where $X_{1,m} = \max(5X/6, mX/6 - 4Y)$ and $I_8^\sim(X_1, Y, \Delta; r, a, b)$ is the number of integer solutions of (7.39), (7.43), (7.44), (7.46) and (7.47) with

$$g_1 \equiv g_2 \equiv h_3 \equiv h_4 \equiv a \pmod{r}, \quad h_1 \equiv h_2 \equiv g_3 \equiv g_4 \equiv b \pmod{r}.$$

Applying (7.53) and the last five inequalities in sequence, we find that there exist functions s and t such that

$$P_r^*(\eta, X, Y) \ll \sum_{j=1}^4 \sum_{m=5}^{12} \sum_{u_1=0}^{r-1} \dots \sum_{u_4=0}^{r-1} I_8^\sim(X_{1,m}, Y, \Delta; r, u_{s(j)}, u_{t(j)}).$$

Therefore, for some $\sigma, \tau \in \{1, \dots, 4\}$ and some $X_1 \in [5X/6, 2X]$,

$$\begin{aligned} P_r^*(\eta, X, Y) &\ll \sum_{u_1=0}^{r-1} \dots \sum_{u_4=0}^{r-1} I_8^\sim(X_1, Y, \Delta; r, u_\sigma, u_\tau) \\ &= \begin{cases} r^3 \sum_{a \pmod{r}} I_8^\sim(X_1, Y, \Delta; r, a, a) & \text{if } \sigma = \tau, \\ r^2 \sum_{a \pmod{r}} \sum_{b \pmod{r}} I_8^\sim(X_1, Y, \Delta; r, a, b) & \text{if } \sigma \neq \tau. \end{cases} \end{aligned}$$

The lemma follows directly from this, as the single sum over $a \pmod{r}$ above is $I_8'(X_1, Y, \Delta, r)$, while the double sum over $a \pmod{r}$ and $b \pmod{r}$ is $I_8''(X_1, Y, \Delta, r) + I_8''(X_1, Y, \Delta, r)$.

LEMMA 7.20. *Let $\eta > 0$ and $0 < Y \leq X/12$. If Δ is as in (7.41) and X_1 as in (7.45), then*

$$I_8'(X_1, Y, \Delta, r) \ll P_r'(\eta, X_1, Y), \quad I_8''(X_1, Y, \Delta, r) \ll P_r''(\eta, X_1, Y),$$

where $P_r'(\eta, X, Y)$ is the number of integer solutions $(l_1, \dots, l_4, k_1, \dots, k_4) = (\mathbf{p}, \mathbf{k})$ (say) for (7.2), (7.3), (7.35) and (7.36) that, in addition, satisfy

$$(7.57) \quad k_i - k_4 \equiv l_i \equiv 0 \pmod{r} \quad (i = 1, \dots, 4),$$

while $P_r''(\eta, X, Y)$ is the number of the same solutions that, in place of (7.57), satisfy

$$(7.58) \quad l_1^2 \equiv \dots \equiv l_4^2 \not\equiv 0 \pmod{r}, \quad k_1 \equiv \dots \equiv k_4 \pmod{r}.$$

Proof. Since r is odd, reversing the substitution (7.38) in (7.54) yields the conditions that $\lambda_i \equiv 0 \pmod{r}$ ($i = 1, \dots, 4$) and $k_1 \equiv \dots \equiv k_4 \pmod{r}$, which (by (7.23)) imply (7.57). Similarly, reversing the substitution (7.38) in (7.55) yields $\lambda_1 \equiv \lambda_2 \equiv -\lambda_3 \equiv -\lambda_4 \not\equiv 0 \pmod{r}$ and $k_1 \equiv \dots \equiv k_4 \pmod{r}$, which imply (7.58). As permutation of the subscripts $1, \dots, 4$ does not change the effect either of (7.57) or of (7.58), it follows that the same argument as was used in the last paragraph of the proof given for Lemma 7.11 may be used here to establish 6-to-1 mappings that imply the claims of this lemma.

LEMMA 7.21. *Let $\eta > 0$ and $0 < Y \leq X_1/10$. Then*

$$P_r'(\eta, X_1, Y) \ll r^{-3} X_1^2 Y^2 + rP(\eta, X_1/r, Y/r).$$

Proof. First note that, by (7.35) (with $X = X_1$) and (7.57), it follows that $P'_r(\eta, X_1, Y) = 0$ unless $r \leq 2Y$. Since $Y \leq X_1/10$, we may therefore assume that $r \leq X_1/5$.

By (7.35) (with $X = X_1$) and (7.57), any family contributing to $P'_r(\eta, X_1, Y)$ contains some member $(\mathbf{p}, \mathbf{k}) = (l_1, \dots, l_4, k_1, \dots, k_4)$ with $k_1 \equiv \dots \equiv k_4 \equiv 0 \pmod{r}$ and $2X_1 < k_1, \dots, k_4 \leq 3X_1 + 1 + r < 4X_1$. Given that, put $(\mathbf{p}', \mathbf{k}') = (\mathbf{p}/r, \mathbf{k}/r) \in \mathbb{Z}^4 \times \mathbb{Z}^4$ (see (7.57)). Then (7.2), (7.3) and (7.35) will all hold with $\mathbf{p}, \mathbf{k}, X$ and Y replaced by $\mathbf{p}', \mathbf{k}', 2X_1/r$ and Y/r . The mapping from the family of (\mathbf{p}, \mathbf{k}) to the family of $(\mathbf{p}', \mathbf{k}')$ is well-defined and injective, so the number of families contributing to $P'_r(\eta, X, Y)$ does not exceed the number of families containing an integer solution of (7.2), (7.3) and (7.35) with X and Y replaced by $2X_1/r$ and Y/r , respectively. By Lemmas 7.3–7.6, this number is $O((2X_1/r)(Y/r)^2) + O((Y/r)^3) + O((2X_1/r)^2(Y/r)^2) = O((X_1/r)^2(Y/r)^2) = O(r^{-4}X_1^2Y^2)$.

Applying Lemma 7.2 with $X = X_1, \theta = \eta Y \sqrt{X_1}$ and $m = r$, we find (see (7.28), (7.29) and (5.8)) that the number of solutions of (7.35) and (7.36) (with $X = X_1$) that lie in any one family \mathcal{F} is

$$\mu(\mathcal{F}) = r\mu_r(\mathcal{F}) + O(r),$$

where $\mu_r(\mathcal{F})$ is the number of those solutions that also satisfy $k_4 \equiv 0 \pmod{r}$ (say). Summing over the relevant families \mathcal{F} (see (7.57)) and applying our bound from the last paragraph, we find that

$$P'_r(\eta, X_1, Y) = rP_r^\times(\eta, X_1, Y) + O(r^{-3}X_1^2Y^2),$$

where $P_r^\times(\eta, X, Y)$ is the number of integer solutions $(l_1, \dots, l_4, k_1, \dots, k_4) = (\mathbf{p}, \mathbf{k})$ of (7.2), (7.3), (7.35) and (7.36) with $k_i \equiv l_i \equiv 0 \pmod{r}$ for $i = 1, \dots, 4$. Since (see (5.8)) all the conditions (7.2), (7.3), (7.35) and (7.36) are homogeneous in $k_1, \dots, k_4, l_1, \dots, l_4, X, Y$, it follows that $P_r^\times(\eta, X_1, Y) = P_1^\times(\eta, X_1/r, Y/r) = P(\eta, X_1/r, Y/r)$, and that our last approximate formula is all that was claimed by the lemma.

LEMMA 7.22. *Let $\eta > 0$ and $0 < Y \leq X_1/12$. Then*

$$P'_r(\eta, X_1, Y) \ll (r^{-3}X_1^2Y^2 + r^{-4}\eta X_1^3Y^2) \log(1 + X_1/r) + r^{-4}\eta^{1/2}(X_1Y)^{5/2} \log^{7/2}(1 + X_1/r).$$

Proof. Since the lesser of the terms $\eta(X_1/r)^4(Y/r) \log^6(1 + X_1/r)$ and $(X_1/r)(Y/r)^4 \log(1 + X_1/r)$ does not exceed the geometric mean, the bound claimed by the lemma follows by Lemma 7.21 from whichever is the stronger of the two Lemmas 7.10 and 7.16.

LEMMA 7.23. *Let $\Delta > 0, X_1 > 0$ and $Y \gg r$. Then*

$$I''_8(X_1, Y, \Delta, r) \ll r^{-2}X_1^2Y^2 + r^{-3}\Delta X_1^{5+\varepsilon}.$$

Proof. By (7.43) and (7.47), $I_g''(X_1, Y, \Delta, r) = 0$ unless $Y < X_1/2$. We may therefore assume that $X_1 \gg Y \gg r$.

It is evident from the proof of Lemma 7.12 that if a trivial family contributes to $I_g''(X_1, Y, \Delta, r)$, then it contains a member $(\mathbf{g}, \mathbf{h}) = (g_1, \dots, g_4, h_1, \dots, h_4)$ with $\{g_1, \dots, g_4\} = \{h_1, \dots, h_4\} = \{0, x, y, z\}$, for some integers x, y, z that are either all $O(Y)$, or satisfy the conditions that $|x| \leq X_1$ and $|y|, |z - x| \leq 4Y$. By (7.55), the elements of $\{0, x, y, z\}$ coprime to r all lie in the same residue class modulo r . Having chosen this reduced residue class ($O(r)$ choices), there then remain $O((Y/r)^3) + O((X_1/r)(Y/r)^2) = O(r^{-3}X_1Y^2)$ choices for x, y, z , followed by just $O(1)$ choices for (\mathbf{g}, \mathbf{h}) . Therefore at most $O(r^{-2}X_1Y^2)$ trivial families contribute anything to $I_g''(X_1, Y, \Delta, r)$, and their combined contribution is (trivially) a total of no more than $O(r^{-2}X_1^2Y^2)$ trivial integer solutions of (7.47).

Before proceeding further we need a bound on the total number of all families that may contribute to $I_g''(X_1, Y, \Delta, r)$. Defining $k_i = (g_i + h_i)/2$ ($i = 1, \dots, 4$), $l_i = (g_i - h_i)/2$ ($i = 1, 2$), and $l_i = (h_i - g_i)/2$ ($i = 3, 4$), we set up a one-to-one correspondence between the integer solutions (\mathbf{g}, \mathbf{h}) of (7.39) and (7.44) and the integer solutions (\mathbf{k}, \mathbf{l}) of [32, formulas (1), (2)] (equations (7.2) and (7.3) in this paper). This also sets up a one-to-one correspondence between families of (\mathbf{g}, \mathbf{h}) 's and families of (\mathbf{k}, \mathbf{l}) 's (where the notion of a family is as defined in [32]), so it follows from (7.43), (7.47) and (7.55) that the number of families contributing to $I_g''(X_1, Y, \Delta, r)$ is at most the number of families of (\mathbf{k}, \mathbf{l}) 's containing at least one integer solution of (7.2), (7.3) with $X_1 < k_i \leq 2X_1$ and $Y < |l_i| \leq 2Y$ ($i = 1, \dots, 4$), $k_1 \equiv k_2 \equiv k_3 \equiv k_4 \pmod{r}$ and $l_1 \equiv l_2 \equiv l_3 \equiv l_4 \pmod{r}$ (r being odd). As the last congruence implies that $r|d = (k_1 - k_4, k_2 - k_4, k_3 - k_4)$ it follows that the same argument used in the proof of [32, Lemma 3] permits us to conclude here that the relevant number of “non-trivial” (asymmetric) families is

$$\ll \sum_{\substack{1 \leq d \leq X_1 \\ d \equiv 0 \pmod{r}}} \sum_{1 \leq e < 2Y} \left(\frac{X_1 Y}{de} \right)^2 \ll r^{-2} X_1^2 Y^2.$$

As for families of (\mathbf{k}, \mathbf{l}) 's that are trivial (in the sense of [32]), the proof of [32, Lemma 1] makes it clear that the number of these families is at most $O(Y(Y/r)^2 + (X_1/r)Y(Y/r)) = O(r^{-2}XY^2)$. Therefore the combined contribution to $I_g''(X_1, Y, \Delta, r)$ from all the families \mathcal{F} with $\mu_{\mathcal{F}}$ (of Lemma 7.14) bounded above by $O(1)$ is just $O(r^{-2}X_1^2Y^2)$.

We model the remainder of the proof on that of Lemma 7.16 (the argument from around (7.48) onwards). It suffices to consider the single case represented by (7.48), since neither it nor any of the seven alternative cases holds any special position in relation to the congruence conditions (7.55). We obtain a bound similar to (7.49), differing only in that there are extra

congruence conditions, $u_1 \equiv u_2 \equiv 0 \not\equiv u_3 \equiv u_4 \pmod{r}$, for the summation on the right of (7.49). It therefore suffices to bound sums similar to (7.50), but with u_1, \dots, u_4 there also subject to the additional congruence conditions just mentioned. Noting the definitions made in Lemma 7.15, we have a crude uniform bound, $d^*(F(u_1, \dots, u_4)) = O(X_1^{\varepsilon/2})$, for each term of the relevant subsum of (7.50). After employing this bound it becomes possible to use the extra congruence conditions to save by a factor of r in respect of each one of the first two summations shown in (7.50), and also (since $Y \gg r$) in respect of the last summation over $u_4 \asymp Y$ there. As a result we find that the non-trivial families \mathcal{F} with $\mu_{\mathcal{F}} \gg 1$ contribute altogether at most $O(X_1^{\varepsilon/2} r^{-3} \Delta X_1^5 \log^3 X_1) = O(r^{-3} \Delta X_1^{5+\varepsilon})$ to the number of integer solutions counted by $I''_8(X_1, Y, \Delta, r)$. In light of the earlier bounds for cases with \mathcal{F} trivial or $\mu_{\mathcal{F}} = O(1)$, the lemma is proved.

The one substantial task remaining, before we can bring our work in this section to its conclusion, is the generalisation of Lemma 7.10 so as to provide a good bound for $P''_r(\eta, X, Y)$ (and therefore also for $I''_8(X, Y, \Delta, r)$) of Lemma 7.20. By (7.58), we need only consider the integer solutions $(\mathbf{p}, \mathbf{k}) = (l_1, \dots, l_4, k_1, \dots, k_4)$ of (7.2), (7.3), (7.35) and (7.36) that satisfy

$$(7.59) \quad k_1 \equiv \dots \equiv k_4 \pmod{r}$$

and, for some $\mathbf{c} \in \{-1, 1\}^4$ and some integer $b \not\equiv 0 \pmod{r}$,

$$(7.60) \quad l_i \equiv c_i b \pmod{r} \quad (i = 1, \dots, 4).$$

We shall say that a family of solutions of (7.2) and (7.3) is “congruent modulo r ” if and only if its members (\mathbf{p}, \mathbf{k}) have $\mathbf{p} = (l_1, \dots, l_4)$ satisfying (7.60) for some $\mathbf{c} \in \{-1, 1\}^4$ and some integer $b \not\equiv 0 \pmod{r}$. Therefore what we seek is an upper bound for the number of integer solutions (\mathbf{p}, \mathbf{k}) of (7.2), (7.3), (7.35), (7.36) and (7.59) from families that are congruent modulo r . By definition no family is congruent modulo 1, so whenever families congruent modulo r are discussed there is an implicit assumption that r is an odd prime.

Since (7.60) is unchanged if \mathbf{c} and b are replaced by $-\mathbf{c}$ and $-b$ it follows that we may assume always that $c_1 = 1$. Furthermore, by (7.2) and (7.60) it follows that $\mathbf{c} \in \{-1, 1\}^4$ must satisfy $c_1 + c_2 \equiv c_3 + c_4 \pmod{r}$ and, since r is an odd prime, this implies that

$$(7.61) \quad c_1 + c_2 = c_3 + c_4.$$

Therefore, for any family which is congruent modulo r , $\mathbf{p} = (l_1, \dots, l_4)$ will satisfy (7.60) for some integer $b \not\equiv 0 \pmod{r}$, with either $\mathbf{c} = (1, 1, 1, 1)$, or $\mathbf{c} = (1, -1, 1, -1)$, or $\mathbf{c} = (1, -1, -1, 1)$.

In addition to the totally symmetric and semi-symmetric families (defined earlier in this section), we shall find it useful to reserve a special name

for families congruent modulo r whose members (\mathbf{p}, \mathbf{k}) satisfy

$$(7.62) \quad c_1k_1 + c_2k_2 = c_3k_3 + c_4k_4,$$

where \mathbf{c} is the same as in (7.60) and (7.61): these families are “singularly congruent modulo r ”. The families which are congruent modulo r but whose members do not satisfy (7.62) are “non-singularly congruent modulo r ”.

LEMMA 7.24. *Let $\eta > 0$ and $0 < Y \leq X/4$. Then*

$$P_r''(\eta, X, Y) \leq B_r''(\eta + O(Y^2/X^2), X, Y),$$

where $B_r''(\delta, X, Y)$ is the number of integer solutions (\mathbf{p}, \mathbf{k}) of (7.2), (7.3), (7.35), (7.37) and (7.59) belonging to families that are congruent modulo r .

Proof. This follows from the above discussion and the argument given as proof of Lemma 7.8.

For our treatment of the families non-singularly congruent modulo r we shall follow the approach used in [32]. As in [32, formula (8)], we note that (7.2) and (7.3) imply

$$(7.63) \quad u_1l_1 + u_2l_2 = u_3l_3,$$

where $u_i = k_i - k_4$ ($i = 1, 2, 3$), so that u_1, u_2, u_3 are constants of the family. We also follow [32] in writing $\mathbf{l} = (l_1, l_2, l_3)$, $\mathbf{u} = (u_1, u_2, u_3)$, while $d = (u_1, u_2, u_3)$ and $e = (l_1, l_2, l_3)$ denote the relevant greatest common divisors. As in [32], a family is “primitive” if and only if it contains integer solutions of (7.2) and (7.3) satisfying the condition $d = e = 1$ (such families become important after Lemma 7.28 below).

As we shall be referring often to [32] it is important to reemphasise one discontinuity in terminology, which is that the families referred to in [32] as “trivial” are here referred to as “symmetric”, so that the families here termed “asymmetric” would be “non-trivial” in [32].

LEMMA 7.25. *Let $X, Y > 0$. For an asymmetric family that is non-singularly congruent modulo r and contains an integer solution of (7.2), (7.3) and (7.35), the greatest common divisors d and e defined above must satisfy*

$$1 \leq d < 6X/r, \quad 1 \leq e \leq 4Y/r, \quad (e, r) = 1.$$

To any given pair d, e , there correspond no more than $O(r^{-1}(XY/de)^2)$ families.

Proof. As a first step we shall establish the constraints on $e = (l_1, l_2, l_3)$. The congruences (7.60) imply that $(e, r) = 1$ and that $l_i/e \equiv c_i b \bar{e} \pmod{r}$ for $i = 1, \dots, 4$. As $\mathbf{c} \in \{-1, 1\}^4$ it follows that $c_i l_i/e \equiv c_j l_j/e \pmod{r}$ for $1 \leq i < j \leq 4$, and therefore that, for $j = 2, 3, 4$, we have $m_j = (c_1 l_1 - c_j l_j)/er \in \mathbb{Z}$. If there exists $j \in \{2, 3, 4\}$ for which $c_j = -1$, then,

since $c_1 = 1$, it follows from (7.35) that the integer m_j must lie in the interval $(2Y/er, 4Y/er]$, so that we must have $4Y/er \geq 1$. If on the other hand $\mathbf{c} = (1, 1, 1, 1)$, then we find from (7.35) that $|m_j| < Y/er$ for $j = 2, 3, 4$, so that either $Y/er > 1$ or $m_2 = m_3 = m_4 = 0$. The latter eventuality can be ruled out, since it would imply that $l_1 = \dots = l_4 > Y$, leading, by way of (7.3), to the contradictory conclusion that a family, supposed to be non-singularly congruent modulo r , has members that satisfy (7.62) for the same \mathbf{c} as in (7.60). We conclude that either $4Y/er \geq 1$, or $Y/er > 1$, so that, in any event, $e \leq 4Y/r$.

To bound the number of families, for given d, e , we shall follow the proof of [32, Lemma 3], which considers the case where $U \leq |u| < 2U$ for given U . It is shown there that, if $U/d \leq Y/e$, then, given \mathbf{u} , there are $O((Y/e)^2/(U/d))$ choices for \mathbf{l} , while, if $U/d > Y/e$ and \mathbf{l} is given, then there are $O((U/d)^2/(Y/e))$ choices for \mathbf{u} pertaining to asymmetric families. We shall not improve these bounds. Instead we shall use the property of “congruence modulo r ” to cut down on the number of choices for whichever one of \mathbf{u}, \mathbf{l} is to be chosen first. We shall assume that \mathbf{c} , satisfying (7.61), has already been chosen.

Since $(e, r) = 1$, the congruences (7.60) imply that the choice of l_1 completely determines the residue classes of l_2/e and l_3/e modulo r . The number of choices for \mathbf{l} (when \mathbf{u} is not given) is therefore at most

$$O((Y/e)(Y/re)^2) = O(r^{-2}(Y/e)^3).$$

Now consider the choice of \mathbf{u} when \mathbf{l} is not given. We may observe that (7.3) implies $l_1(u_1/d) + l_2(u_2/d) = l_3(u_3/d)$, which, together with (7.60), goes to show that

$$(7.64) \quad c_1(u_1/d) + c_2(u_2/d) \equiv c_3(u_3/d) \pmod{r}.$$

If it were the case that $d \geq 2\sqrt{3}U/r$, then the bound

$$|c_1(u_1/d) + c_2(u_2/d) - c_3(u_3/d)| \leq \sqrt{3}|\mathbf{u}|/d < 2\sqrt{3}U/d$$

would now imply $c_1u_1 + c_2u_2 - c_3u_3 = 0$. This, however, would imply that the family is singularly congruent modulo r (see (7.61) and (7.62)), contradicting the hypotheses of the lemma. Therefore we conclude that, of necessity, $d < 2\sqrt{3}U/r$, which, together with the bounds $|u_i| < X$ ($i = 1, 2, 3$) from (7.35), leads us to a confirmation of the lemma’s claim that $d < 6X/r$. Moreover, as we found $U \gg rd$, so it follows from the congruence (7.64) that the number of choices for \mathbf{u} is at most $O((U/d)^2(U/rd)) = O(r^{-1}(U/d)^3)$.

In conclusion, when $U/d \leq Y/e$ the number of families is at most $O(r^{-1}(U/d)^3(Y/e)^2/(U/d)) = O(r^{-1}(U/d)^2(Y/e)^2)$, while, for $U/d > Y/e$, it is at most $O(r^{-2}(Y/e)^3(U/d)^2/(Y/e)) = O(r^{-2}(U/d)^2(Y/e)^2)$. Since there were just three choices for \mathbf{c} , the lemma follows from the last two bounds

applied to all the cases where U is an integer power of 2 satisfying $rd/2\sqrt{3} < U < \sqrt{3}X$.

LEMMA 7.26. *Let $\delta > 0$ and let $X \geq Y \geq r$. Then the number of families that are singularly congruent modulo r and contain an integer solution (\mathbf{p}, \mathbf{k}) of (7.2), (7.3), (7.35), (7.37) and (7.59) is not more than $O(r^{-2}XY(Y + \delta X) \log(2 + X/r))$.*

Proof. Any such family must contain some member $(\mathbf{p}, \mathbf{k}^*)$ with $k_1^* \equiv \dots \equiv k_4^* \equiv 0 \pmod{r}$,

$$(7.65) \quad 3X < k_1^*, \dots, k_4^* \leq 4X + 1 + r, \quad Y < l_1, \dots, l_4 \leq 2Y.$$

By (7.60) it follows that $\mathbf{p} = (l_1, \dots, l_4)$ and \mathbf{k}^* here are given by

$$(7.66) \quad l_i = rl'_i + c_i b, \quad k_i^* = rk'_i \quad (i = 1, \dots, 4),$$

where $\mathbf{c} \in \{-1, 1\}^4$ satisfies (7.61), $(l'_1, \dots, l'_4) = \mathbf{p}' \in \mathbb{Z}^4$, $(k'_1, \dots, k'_4) = \mathbf{k}' \in \mathbb{Z}^4$ and b is an integer satisfying $0 < |b| \leq (r - 1)/2$. It will suffice to count how many families there can be for a given choice of \mathbf{c} and b , so, in what follows, we shall assume that the values of b and c_1, \dots, c_4 have already been chosen.

As $(\mathbf{p}, \mathbf{k}^*)$ comes from a singularly congruent family modulo r , it follows that \mathbf{k}^* must satisfy the analog of (7.62). This, together with (7.61) and (7.66), means that the analogs of (7.2) and (7.3) for \mathbf{p} and \mathbf{k}^* imply the analogs of (7.2) and (7.3) for \mathbf{p}' and \mathbf{k}' . For given \mathbf{c} and b , the mapping from the family of $(\mathbf{p}, \mathbf{k}^*)$ to the family of $(\mathbf{p}', \mathbf{k}')$ is well-defined and injective. It therefore suffices to bound the number of choices for the latter family. To this end we observe that, since $X, Y \geq r$, the conditions (7.65) and (7.66) imply that

$$(7.67) \quad Y/2r < l'_i \leq 5Y/2r, \quad 3X/r < k'_i \leq 6X/r \quad (i = 1, \dots, 4).$$

By (7.66), the analog of (7.62) for \mathbf{k}' will hold. Therefore, it follows by (7.61) that

$$(7.68) \quad c_1 u'_1 + c_2 u'_2 = c_3 u'_3,$$

where $u'_i = k'_i - k'_4$ ($i = 1, 2, 3$). Likewise, it follows from the analogs of (7.2) and (7.3) for \mathbf{p}' and \mathbf{k}' that $l'_1 u'_1 + l'_2 u'_2 = l'_3 u'_3$. Assuming (as we may) that $c_1 = 1$, the elimination, through (7.68), of u'_1 from the last equation reveals that $a_1 b_1 = a_2 b_2$, where

$$(7.69) \quad \mathbf{a} = (l'_2 - c_2 l'_1, l'_3 - c_3 l'_1), \quad \mathbf{b} = (u'_2, u'_3).$$

Let us consider first the case where $\mathbf{a}, \mathbf{b} \neq \mathbf{0}$. In this case either $a_1/a_2 = b_2/b_1$ or $a_2/a_1 = b_1/b_2$. As both eventualities represent similar subcases, we shall assume the former. Let $\alpha = (a_1, a_2)$ and $\beta = (b_1, b_2)$ (greatest common divisors). Then $a'_1/a'_2 = b'_2/b'_1$ with $\mathbf{a}' = \mathbf{a}/\alpha$, $\mathbf{b}' = \mathbf{b}/\beta$ and $(a'_1, a'_2) = (b'_1, b'_2) = 1$, so that $\mathbf{a}' = \pm(b'_2, b'_1)$ (equivalently $\mathbf{b}' = \pm(a'_2, a'_1)$).

In light of this, and recalling (7.67), (7.69) and the definitions under (7.68), we find that $0 \neq |\mathbf{a}'| = |\mathbf{b}'| \ll \min(Y/r\alpha, X/r\beta)$ and that the number of choices for \mathbf{a} and \mathbf{b} is

$$\ll \sum_{1 \leq \alpha \ll Y/r} \sum_{1 \leq \beta \ll X/r} \min \left(\left(\frac{Y/r}{\alpha} \right)^2, \left(\frac{X/r}{\beta} \right)^2 \right) \ll \frac{XY}{r^2} \log \left(2 + \frac{Y}{r} \right).$$

Once \mathbf{a} and \mathbf{b} are chosen, then, by (7.69) and (7.68), each of u'_2, u'_3 and u'_1 is known. There will remain just $O(Y/r)$ available choices for l'_1 satisfying (7.67). Given \mathbf{a}, \mathbf{b} and l'_1 , the values of l'_2, l'_3 and l'_4 can be found from (7.69) and the analog of (7.2), thereby determining the family of $(\mathbf{p}', \mathbf{k}')$ completely. We conclude that at most $O((X/r)(Y/r)^2 \log(2 + Y/r))$ families belong in the case we have been discussing.

Now consider the case where $\mathbf{b} = \mathbf{0}$. Here (7.68) and (7.69) show that $u'_1 = u'_2 = u'_3 = 0$, so it follows that the family is either semi-symmetric, or totally symmetric (see Lemma 7.3). As the proofs of Lemmas 7.4 and 7.5 are easily adapted to work with (7.67) rather than an exact analog of (7.35), so we may conclude that this case (where $\mathbf{b} = \mathbf{0}$) covers at most $O((X/r)(Y/r)^2) + O((Y/r)^3)$ families.

Finally, consider the case where $\mathbf{a} = \mathbf{0}$. By (7.69) and (7.67), this case can only occur when $c_2, c_3 > 0$. By (7.61) and the remarks following it we conclude that $c_i = 1$ for $i = 1, \dots, 4$. By (7.69) and the analog of (7.2) it follows that l'_1, \dots, l'_4 are all equal. The last two deductions, together with (7.66), show that $l_1 = \dots = l_4 \equiv b \pmod{r}$. It is given us by the lemma that the family of $(\mathbf{p}, \mathbf{k}^*)$ contains another integer member (\mathbf{p}, \mathbf{k}) , which is a solution of (7.35) and (7.37), as well as (7.2), (7.3) and (7.59). Since $\mathbf{p} = (l_1, \dots, l_4)$ is a multiple of $(1, \dots, 1)$, we have it by (7.35), (7.37) and (7.3) that the \mathbf{k} here satisfies $k_1 + k_2 = k_3 + k_4$ and $|\sqrt{k_1} + \sqrt{k_2} - \sqrt{k_3} - \sqrt{k_4}| < \delta\sqrt{X}$. Comparing the expansions of $(x - \sqrt{k_1})(x - \sqrt{k_2})$ and $(x - \sqrt{k_3})(x - \sqrt{k_4})$ at $x = \sqrt{k_4}$, one deduces, by way of (7.35), (7.59) and the above conditions, that $\mathbf{u} = (k_1 - k_4, k_2 - k_4, k_3 - k_4)$ satisfies $u_1 u_2 \ll \delta X^2$, $u_3 = u_1 + u_2 \ll X$ and $u_1, u_2, u_3 \equiv 0 \pmod{r}$. Since $X \gg r$, the number of choices for \mathbf{u} is at most $O(X/r) + O(\delta(X/r)^2 \log(2 + X/r))$. Since $Y \gg r$, there are also at most $O(Y/r)$ choices for $\mathbf{p} = l(1, \dots, 1)$ with $Y < l \leq 2Y$ and $l \equiv b \pmod{r}$. We conclude that there are at most $O((X/r)(Y/r)) + O(\delta(X/r)^2(Y/r) \log(2 + X/r))$ families belonging in the final one of our three cases. Note that, since $Y/r \gg 1$, the first O -term in our last bound is dominated by terms in the bounds for the previous two cases.

Collecting the results from the three cases considered, eliminating redundant terms, and noting that there were only $O(r)$ choices available for the b and \mathbf{c} that we have taken as given, we arrive at the bound claimed by the lemma.

LEMMA 7.27. *Let $X \geq Y \geq r$. The number of symmetric families that are congruent modulo r and contain an integer solution (\mathbf{p}, \mathbf{k}) of (7.2), (7.3), (7.35) and (7.59) is $O(r^{-2}XY^2)$.*

Proof. Such a family will contain a unique member $(\mathbf{p}, \mathbf{k}^*)$ satisfying analogs of (7.2), (7.3), and with $0 = k_1^* \equiv \dots \equiv k_4^* \pmod{r}$, $l_1^2 \equiv \dots \equiv l_4^2 \not\equiv 0 \pmod{r}$, $|k_i^*| < X$ and $Y < l_i \leq 2Y$, for $i = 1, \dots, 4$. Since the member determines the family, it suffices to bound the number of such $(\mathbf{p}, \mathbf{k}^*)$.

If the family is totally symmetric, then we have $O(X/r)$ choices for k_2^* , then $O(Y)$ choices for l_1 , followed by $O(Y/r)$ choices for l_2 , then $O(1)$ choices for k_3^* , l_3 , k_4^* and l_4 as in Lemma 7.3(i). It follows that there are at most $O(XY^2/r^2)$ such totally symmetric families.

If the family is semi-symmetric, then $0 = k_1^* = \dots = k_4^*$, and, by (7.2) and the bounds and congruence conditions for l_1, \dots, l_4 , there are just $O(Y(Y/r)^2)$ choices for $\mathbf{p} = (l_1, \dots, l_4)$, and (therefore) for the family.

As $Y \leq X$, the lemma follows from our bounds in the two cases.

Two concepts found useful in [32] are “renumbering” and “orderedness”. They help to reduce the number of cases that need to be considered.

A “renumbering” is a permutation of the subscripts $1, \dots, 4$ that fixes the set of sets $\{\{1, 2\}, \{3, 4\}\}$. Note that conditions (7.2), (7.3), (7.35), (7.37) and (7.59)–(7.62) are invariant under the action of the group of 8 renumberings (it being assumed that the same renumbering is applied to each of $\mathbf{c} = (c_1, \dots, c_4)$, $\mathbf{p} = (l_1, \dots, l_4)$ and $\mathbf{k} = (k_1, \dots, k_4)$). The invariance of (7.60)–(7.62) under renumbering implies that the property of being non-singularly congruent modulo r is conserved when a family is transformed by any renumbering acting upon its members. Although renumbering may change \mathbf{c} , it remains an element of $\{-1, 1\}^4$ satisfying (7.61) so we may continue to assume that \mathbf{c} takes one of the three forms given under (7.61) (if $c_1 = -1$ replace b and \mathbf{c} in (7.60)–(7.62) by $-b$ and $-\mathbf{c}$).

The family of (\mathbf{p}, \mathbf{k}) is “ordered” if and only if u_1, u_2, u_3 in (7.63) satisfy either $0 \leq u_2 \leq u_1, u_3$, or $u_1, u_3 \leq u_2 \leq 0$. It is shown in [32, Lemma 2] that every integer solution of (7.2), (7.3), (7.35) and (7.37) from an asymmetric family is a renumbering of one which is from an ordered asymmetric family, and that the correspondence is at most 8-to-1.

A third piece of terminology from [32] is “stability”. A solution, $(\mathbf{p}, \mathbf{k}) = (\mathbf{p}^*, \mathbf{k}^*)$ (say), for (7.2), (7.3), (7.35) and (7.37) is “stable” if and only if $(\mathbf{p}, \mathbf{k}) = (\mathbf{p}^*, \mathbf{k}^* + t(1, \dots, 1))$ represents a solution of both (7.35) and (7.37) for all $t \in (-1/2, 1/2)$ (not just at $t = 0$).

After the next lemma it will be important to establish a good upper bound on how many of the stable integer solutions (\mathbf{p}, \mathbf{k}) of (7.2), (7.3), (7.35) and (7.37) are members of families that are (simultaneously) primitive, ordered, asymmetric and non-singularly congruent modulo r : whatever

the exact number of these solutions may be, we are momentarily content just to denote it by $B_r^+(\delta, X, Y)$.

LEMMA 7.28. *Let $\delta > 0$ and $X \geq Y \geq r$. Then*

$$B_r''(\delta, X, Y) \ll r^{-2}(X^2Y^2 + \delta X^3Y) \log(2 + X/r) + \sum_{\substack{1 \leq d < 6X/r \\ d \equiv 0 \pmod r}} \sum_{\substack{1 \leq e \leq 4Y/r \\ (e,r)=1}} B_r^+(O(\delta), X/d, Y/e)d.$$

Proof. As no family can contain more than X integer solutions of (7.35), it follows from Lemmas 7.26 and 7.27 that the contribution from families which are either symmetric or singularly congruent modulo r is covered by the first term in the upper bound of the lemma. Therefore, and by the discussion preceding the lemma, it will suffice to bound the contribution to $B_r''(\delta, X, Y)$ from the ordered asymmetric families that are non-singularly congruent modulo r . Given $d, e \in \mathbb{N}$, let $B_r^\times(\delta, X, Y; d, e)$ stand for how much (in absolute terms) of this last contribution arises from families with $(u_1, u_2, u_3) = d$ and $(l_1, l_2, l_3) = e$, where $u_i = k_i - k_4$, as in (7.63). To bound $B_r^\times(\delta, X, Y; d, e)$ we apply essentially the argument used in the proof of [32, Lemma 4], but with Lemma 7.25 replacing [32, Lemma 3] in the penultimate step. This gives $B_r^\times(\delta, X, Y; d, e) \leq B_r^+(O(\delta), X/d, Y/e)d + O(r^{-1}X^2Y^2/de^2)$.

Since (7.59) implies $u_i \equiv 0 \pmod r$ for $i = 1, 2, 3$, we have $d \equiv 0 \pmod r$, as well as all the other conditions that d, e are known to satisfy (after Lemma 7.25). The constraint $(e, r) = 1$, in particular, plays a rôle in establishing the last bound for $B_r^\times(\delta, X, Y; d, e)$, since it facilitates our adaptation of the proof of [32, Lemma 4] by permitting us to show that $(\mathbf{p}/e, \mathbf{k}/d)$ is from a non-singularly congruent family modulo r if (\mathbf{p}, \mathbf{k}) is (see (7.60) and (7.62)). Given its hypotheses, the lemma follows by summing the bound for $B_r^\times(\delta, X, Y; d, e)$ over the available choices for d, e .

LEMMA 7.29. *Let $\delta > 0$ and $X, Y \gg r$. Then*

$$B_r^+(\delta, X, Y) \ll r^{-1}X^2Y^2 + r^{-1}\delta X^3Y^2 \log^2(1 + XY).$$

Proof. This is, more or less, a generalisation of the first half of the Theorem in [32], the proof of which is divided between [32, Lemma 7] and [32, Lemma 8].

The proofs given in [32] for those lemmas are an essential reference for our proof here. They succeed by dividing into cases and employing the results of [32, Lemma 5] to bound the total number of choices for the pair of \mathbf{u}, \mathbf{l} from (7.63), subject to certain conditions (determined by the case): this number being itself a bound for the number of families pertinent to the case. Each case depends on a pair of parameters, U, W , which are assumed to be integer powers of 2 lying in the interval $[1, X]$ and satisfying (at least)

$W \ll U$. Given U, W , the case is determined by a set of conditions which (it turns out) always imply $|u_1|, |u_3| \asymp U$, $|u_1 - u_3| \asymp W$ and $u_2 \ll W$. In the proof of [32, Lemma 7] one has $UW \ll \delta X^2$, and the bound for $F(U, W)$, the number of families in the case, is whichever one is the lesser of the two sums:

$$S(\Delta, U; Y) = \sum_{l_1, l_2, l_3 \asymp Y} (d(l_3) + \Delta U^2 \sigma(l_3))$$

with $\Delta = 2W/U$, and

$$S^\perp(\Delta, Y; U, W) = \sum_{\substack{|u_1|, |u_3| \asymp U, |u_2| \ll W \\ |u_1 - u_3| \asymp W}} (d(|u_3|) + \Delta Y^2 \sigma(|u_3|)/u_3^2)$$

with $\Delta = 3/2$. Apart from new choices for Δ 's, the situation in [32, Lemma 8] is similar, but $UW \gg \delta X^2$, and $S(\Delta, U; Y)$ becomes $S'(\Delta, U, W, Y)$, by virtue of an extra condition $0 < l_3 - l_1 \ll (W/U)Y$ constraining the summation. Another difference (only important later) is that one actually has $|u_2| \asymp W$ in all cases relevant to [32, Lemma 8].

The sums we have to consider here are the subsums $S_r(\Delta, U; Y)$, $S'_r(\Delta, U, W, Y)$ and $S_r^\perp(\Delta, Y; U, W)$ of $S(\Delta, U; Y)$, $S'(\Delta, U, W, Y)$ and $S^\perp(\Delta, Y; U, W)$, respectively, where the former sums differ from the latter only in that the summations are made additionally contingent upon compatibility of the variable of summation (\mathbf{u} or \mathbf{l}) with its supposed rôle as one ‘‘half’’ of an integer solution (\mathbf{u}, \mathbf{l}) for (7.63) determining a family of solutions of (7.2), (7.3) that is non-singularly congruent modulo r . Recall that a family is non-singularly congruent modulo r if and only if its members (\mathbf{p}, \mathbf{k}) satisfy (7.60), with some b, \mathbf{c} as described there, and do not satisfy (7.62) (with the same \mathbf{c}). Therefore, by (7.60) it follows that in $S_r(\Delta, U; Y)$ and $S'_r(\Delta, U, W, Y)$ one must sum only over \mathbf{l} that are solutions of $c_1 l_1 \equiv c_2 l_2 \equiv c_3 l_3 \not\equiv 0 \pmod{r}$, so that $l_j \equiv \pm l_1 \pmod{r}$ for $j = 2, 3$. On the other hand, it follows from (7.60) and (7.63) that \mathbf{u} in $S_r^\perp(\Delta, Y; U, W)$ must satisfy $c_1 u_1 + c_2 u_2 \equiv c_3 u_3 \pmod{r}$. As (7.62) must not hold, it follows by (7.61) that one cannot have $c_1 u_1 + c_2 u_2 = c_3 u_3$ in $S_r^\perp(\Delta, Y; U, W)$.

Using the bounds $d(n) = O(n^\varepsilon)$, $\sigma(n) = O(n \log(1 + n))$ and $Y \geq r$, it is a trivial matter to show

$$(7.70) \quad \begin{aligned} S_r(\Delta, U; Y) &\ll (Y^\varepsilon + \Delta U^2 Y^{-1} \log(1 + Y)) Y (Y/r)^2 \\ &\ll r^{-2} (Y^{3+\varepsilon} + \Delta U^2 Y^2 \log(1 + Y)), \end{aligned}$$

$$(7.71) \quad \begin{aligned} S'_r(\Delta, U, W, Y) &\ll (Y^\varepsilon + \Delta U^2 Y^{-1} \log(1 + Y)) (W/U) Y^3 / r \\ &\ll r^{-1} (U^{-1} W Y^{3+\varepsilon} + \Delta U W Y^2 \log(1 + Y)) \end{aligned}$$

and

$$S_r^\perp(\Delta, Y; U, W) \ll (U^\varepsilon + \Delta Y^2 U^{-1} \log(2 + U)) T_r(U, W),$$

where $T_r(U, W)$ is the number of $\mathbf{u} \in \mathbb{Z}^3$ satisfying the conditions

$$\begin{aligned} |u_1|, |u_3| &\asymp U, & u_2 &\ll W, & |u_1 - u_3| &\asymp W, \\ 0 &\neq c_1u_1 + c_2u_2 - c_3u_3 &\equiv 0 & \pmod{r}. \end{aligned}$$

In order to bound $T_r(U, W)$ we note that $(c_1u_1 + c_2u_2 - c_3u_3)/r = m$ (say) is restricted to non-zero integer values. Recall that $c_2, c_3 \in \{-1, 1\}$, and we may assume $c_1 = 1$. Hence, if $c_3 = 1$, then $|m| \leq (|u_1 - u_3| + |u_2|)/r \ll W/r$, while if $c_3 = -1$, then $|m| \leq (|u_1| + |u_2| + |u_3|)/r \ll (U + W)/r \ll U/r$ and $|u_1 - rm/2| = |u_1 - u_2 - u_3|/2 \leq |u_1 - u_3| + |u_2| \ll W$. In the former case we have $O(W/r)$ choices for m , $O(U)$ independent choices for u_1 and $O(W)$ independent choices for u_2 ; u_3 is then determined by $c_3u_3 = c_1u_1 + c_2u_2 - rm$. In the latter case we have $O(U/r)$ choices for m , then $O(W)$ choices for u_1 (sufficiently near $rm/2$), followed by $O(W)$ independent choices for u_2 . In any event, we find that $T_r(U, W) = O(UW^2/r)$ and so

$$\begin{aligned} (7.72) \quad S_r^\perp(\Delta, Y; U, W) &\ll (U^\varepsilon + \Delta Y^2 U^{-1} \log(2 + U)) UW^2/r \\ &\ll r^{-1}(U^{1+\varepsilon}W^2 + \Delta Y^2 W^2 \log(2 + U)). \end{aligned}$$

By (7.70) and (7.72), the number of families in the case considered in the proof of [32, Lemma 7] is

$$\begin{aligned} (7.73) \quad F_r(U, W) &\ll \min(S_r(2W/U, U; Y), S_r^\perp(3/2, Y; U, W)) \\ &\ll \min(r^{-2}(Y^{3+\varepsilon} + UWY^2 \log(1 + Y)), \\ &\quad r^{-1}(U^{1+\varepsilon}W^2 + W^2Y^2 \log(2 + U))) \\ &\ll r^{-1}UWY^2 \log(1 + XY) + r^{-1} \min(Y^{3+\varepsilon}, U^{1+\varepsilon}W^2) \\ &\ll r^{-1}UWY^2 \log(1 + XY), \end{aligned}$$

where the last step follows on noting that

$$\min(Y^{3+\varepsilon}, U^{1+\varepsilon}W^2) < \min(Y^4, U^2W^2) \ll \sqrt{Y^4U^2W^2} = UWY^2$$

(we may assume that $0 < \varepsilon < 1/3$). By (7.71) and (7.72), the number of families in the case considered in the proof of [32, Lemma 8] is

$$\begin{aligned} (7.74) \quad F'_r(U, W) &\ll \min\left(S'_r\left(\frac{\delta X^2}{U^2} + \frac{W}{X}, U, W, Y\right), S_r^\perp\left(\frac{\delta X^2}{UW} + \frac{U}{X}, Y; U, W\right)\right) \\ &\ll r^{-1} \min\left(\frac{W}{U} Y^{3+\varepsilon} + \frac{W}{U} \left(\frac{\delta X^2}{U^2} + \frac{W}{X}\right) U^2 Y^2 \log(1 + Y), \right. \\ &\quad \left. U^{1+\varepsilon}W^2 + \left(\frac{\delta X^2}{UW} + \frac{U}{X}\right) W^2 Y^2 \log(U + 2)\right) \end{aligned}$$

$$\begin{aligned}
 &\ll r^{-1} \min\left(\frac{W}{U} Y^{3+\varepsilon} + \left(\frac{W}{U} \delta X^2 Y^2 + UW^2 \frac{Y^2}{X}\right) \log(1 + Y),\right. \\
 &\qquad \left.U^{1+\varepsilon} W^2 + \left(\frac{W}{U} \delta X^2 Y^2 + UW^2 \frac{Y^2}{X}\right) \log(1 + U)\right) \\
 &\ll r^{-1} \frac{W}{U} \delta X^2 Y^2 \log(1 + XY) + r^{-1} UW^2 \frac{Y^2}{X} \log(1 + XY) \\
 &\quad + r^{-1} \min\left(\frac{W}{U} Y^{3+\varepsilon}, U^{1+\varepsilon} W^2\right) \\
 &\ll r^{-1} \frac{W}{U} \delta X^2 Y^2 \log(1 + XY) + r^{-1} UW^2 \frac{Y^2}{X} \log(1 + XY) \\
 &\quad + r^{-1} \frac{W}{U} U^{1/9} X^{8/9} Y^2,
 \end{aligned}$$

where the last step is justified by the observation that either $U^{1+\varepsilon} W^2 \leq UW^2 Y^2 / X$, or $Y^2 < U^\varepsilon X$, which would imply that $Y \ll X^{(1+\varepsilon)/2} < X^{2/3}$ and (therefore) that $(W/U)Y^{3+\varepsilon} \ll (W/U)Y^2 Y^{4/3} \ll (W/U)Y^2 X^{8/9} \ll (W/U)U^{1/9} X^{8/9} Y^2$.

By (7.35) and [32, Lemma 6], each family has at most $O(X)$ integer members, and, if $u_1 u_2 u_3 \neq 0$, then at most $O(\delta X^4 / |u_1 u_2 u_3|)$ are stable solutions of (7.2), (7.3), (7.35) and (7.37). As in [32] we use only the former of these two bounds for cases in the proof of [32, Lemma 7], but choose between them for cases in the proof of [32, Lemma 8], where one actually has $|u_2| \asymp W$ as well as $|u_1|, |u_3| \asymp U$. In light of (7.73) and (7.74) we find that, regarding the above two sorts of case, there are the two respective upper bounds of

$$O(r^{-1} U W X Y^2 \log(1 + XY))$$

and

$$O\left(r^{-1} \frac{W}{U} \delta X^3 Y^2 \log(1 + XY) + r^{-1} \frac{W}{U} U^{1/9} X^{17/9} Y^2\right)$$

for the number of stable integer solutions (\mathbf{p}, \mathbf{k}) of (7.2), (7.3), (7.35) and (7.37) that are members of families simultaneously primitive, ordered, asymmetric and non-singularly congruent modulo r . Summing these bounds, with U, W running over integer powers of 2 satisfying $UW \ll \delta X^2$ in respect of the former bound, but $UW \gg \delta X^2$ in respect of the latter (and $1 \ll W \ll U \ll X$ in respect of both), one finds that the result is the upper bound given for $B_r^+(\delta, X, Y)$ by the lemma. Since it was shown in [32] that all contributing cases are covered in the last summation, we conclude that the lemma is proved.

LEMMA 7.30. *Let $\delta > 0$ and $X \geq Y \geq r$. Then*

$$B_r''(\delta, X, Y) \ll r^{-2} X^2 Y^2 \log(1 + X/r) + r^{-3} \delta X^3 Y^2 \log^2(1 + X).$$

Proof. The bound can be obtained by applying Lemma 7.29 to bound the terms of the sum in the bound for $B_r''(\delta, X, Y)$ given by Lemma 7.28. After carrying out the relevant summations over d, e , one only has to note that $r^{-2}\delta X^3 Y \leq r^{-2}\delta X^3 Y(Y/r) = r^{-2}\delta X^3 Y^2$.

LEMMA 7.31. *Let $\eta > 0$ and $r \leq Y \leq X_1/4$. Then*

$$P_r''(\eta, X_1, Y) \ll r^{-2} X_1^2 Y^2 \log(1 + X_1/r) + r^{-3} \eta X_1^3 Y^2 \log^2(1 + X_1) + r^{-3} X_1 Y^4 \log^2(1 + X_1).$$

Proof. This bound follows by applying Lemma 7.30, with $X = X_1$ and $\delta = \eta + O(Y^2/X_1^2)$, in order to bound the bound given for $P_r''(\eta, X_1, Y)$ by Lemma 7.24.

LEMMA 7.32. *Let $\eta > 0$ and $r \leq Y \leq X/16$. If Δ is as in (7.41) and X_1 as in (7.45), then*

$$\begin{aligned} I'_8(X_1, Y, \Delta, r) &\ll (r^{-3} X_1^2 Y^2 + r^{-4} \Delta X_1^4 Y) \log(X_1/r) \\ &\quad + r^{-4} \Delta^{1/2} X_1^3 Y^2 \log^{7/2}(X_1/r), \\ I''_8(X_1, Y, \Delta, r) &\ll r^{-2} X_1^2 Y^2 \log(X_1/r) + r^{-3} \Delta X_1^4 Y \log^2(X_1) \\ &\quad + r^{-3} X_1^\epsilon \min(\Delta X_1^5, X_1 Y^4). \end{aligned}$$

Proof. Note first that conditions (7.41) and (7.45) imply $X_1 \asymp X$, $\eta \asymp (X_1/Y)\Delta$ and (since $Y \leq X/16$) $Y \leq 6X_1/80 < X_1/12$. Therefore the lemma's claim concerning $I'_8(X_1, Y, \Delta, r)$ follows by Lemmas 7.20 and 7.22, while out of Lemmas 7.23, 7.20 and 7.31 comes the bound

$$I''_8(X_1, Y, \Delta, r) \ll r^{-2} X_1^2 Y^2 \log(X_1/r) + r^{-3} \min(\Delta X_1^{5+\epsilon}, (\Delta X_1^4 Y + X_1 Y^4) \log^2(X_1)),$$

which contains the lemma's second claim.

LEMMA 7.33. *Let $\eta > 0$ and $r \leq Y \leq X/16$. Then*

$$P_r^*(\eta, X, Y) \ll X^2 Y^2 \log(X/r) + r^{-1} \eta^{1/2} X^{5/2} Y^{5/2} \log^{7/2}(X/r) + r^{-1} \eta X^3 Y^2 \log^2 X + r^{-1} X^\epsilon \min(\eta X^4 Y, X Y^4).$$

Proof. The consecutive application of Lemmas 7.19 and both parts of Lemma 7.32 produces the bound

$$P_r^*(\eta, X, Y) \ll X_1^2 Y^2 \log(1 + X_1/r) + r^{-1} \Delta^{1/2} X_1^3 Y^2 \log^{7/2}(1 + X_1/r) + r^{-1} \Delta X_1^4 Y \log^2(1 + X_1) + r^{-1} X_1^\epsilon \min(\Delta X_1^5, X_1 Y^4),$$

where Δ and X_1 are as in (7.41) and (7.45). This is, give or take a bounded factor, just what the lemma claims.

LEMMA 7.34.

$$P_r^*(\Gamma, K_0, L_0) \ll \frac{Q}{R} K_0^2 L_0^2 \log^{7/2} K_0 + r^{-1} K_0^\epsilon \min(\Gamma K_0^4 L_0, K_0 L_0^4).$$

Proof. By (7.51), $\Gamma > 0$. By (6.9), (6.10) and (5.2), we have $L_0 \leq K_0/16$ and

$$L_0 \geq \frac{L}{4} = \frac{rQH}{4R^2} > \frac{rQ^2}{2\varepsilon_0 R^2} > \left(\frac{Q}{R}\right)^2 r \geq r.$$

Therefore Lemma 7.33 applies to give

$$\begin{aligned} P_r^*(\Gamma, K_0, L_0) &\ll K_0^2 L_0^2 \log(K_0/r) + r^{-1} \Gamma K_0^3 L_0^2 \log^2 K_0 \\ &\quad + r^{-1} \Gamma^{1/2} K_0^{5/2} L_0^{5/2} \log^{7/2}(K_0/r) \\ &\quad + r^{-1} K_0^\varepsilon \min(\Gamma K_0^4 L_0, K_0 L_0^4). \end{aligned}$$

By (7.51), we have $r^{-1} \Gamma^{1/2} K_0^{5/2} L_0^{5/2} \asymp (Q/R) K_0^2 L_0^2$ and (see the above lower bound for L_0) $r^{-1} \Gamma K_0^3 L_0^2 \asymp r(Q/R)^2 K_0^2 L_0 < K_0^2 L_0^2$, so the last bound for $P_r^*(\Gamma, K_0, L_0)$ implies the lemma.

LEMMA 7.35. *Suppose that A_1 is as in (6.36) and that $\Delta = \Gamma L_0/K_0$, where Γ is as in (7.30)–(7.31) (and (7.51)). Then*

$$\begin{aligned} K_0^2 A_1 &\ll \frac{Q}{R} K_0^2 L_0^2 \log^6 K_0 + r^{-1/2} K_0^\varepsilon \min(\Delta^{1/3} K_0^2 L_0^3, \Delta K_0^{14/3} L_0^{1/3}) \\ &\quad + r^{-1/2} \Gamma K_0^3 L_0^2 \log K_0 + r^{-1} K_0^\varepsilon \min(\Delta^{1/4} K_0^2 L_0^3, \Delta K_0^5). \end{aligned}$$

Proof. By Lemmas 7.7, 7.34 and 7.18, we have

$$\begin{aligned} K_0^2 A_1^* &\ll \frac{Q}{R} K_0^2 L_0^2 \log^{7/2} K_0 + r^{-1} K_0^\varepsilon \min(\Gamma K_0^4 L_0, K_0 L_0^4), \\ K_0^2 A_1' &\ll K_0^2 L_0^2 \log^6 K_0 + K_0^\varepsilon \min(\Delta^{1/3} K_0^2 L_0^3, \Delta K_0^{14/3} L_0^{1/3}) \\ &\quad + \Gamma K_0^3 L_0^2 \log K_0 + K_0^\varepsilon \min(K_0^{3/2} L_0^3, \Delta K_0^{11/2} L_0^{-1}). \end{aligned}$$

As (7.25) implies $K_0^2 A_1 = O(K_0^2 A_1^*) + O(r^{-1/2} K_0^2 A_1')$, the lemma will follow from the above bounds (and (5.2)) if we can only show

$$\begin{aligned} r^{-1} \min(\Delta^{1/4} K_0^2 L_0^3, \Delta K_0^5) \\ \gg r^{-1} \min(K_0 L_0^4, \Gamma K_0^4 L_0) + r^{-1/2} \min(K_0^{3/2} L_0^3, \Delta K_0^{11/2} L_0^{-1}). \end{aligned}$$

As $\Gamma = \Delta K_0/L_0$, the right-hand side above is

$$r^{-1} (1 + r^{1/2} K_0^{1/2} L_0^{-1}) \min(K_0 L_0^4, \Delta K_0^5),$$

which, after (7.51) and (5.2), is

$$\begin{aligned} &\ll r^{-1} \left(1 + \Delta^{1/4} \frac{K_0}{L_0}\right) \Delta K_0^5 \left(1 + \Delta^{1/4} \frac{K_0}{L_0}\right)^{-4} \\ &\ll r^{-1} \Delta K_0^5 \min\left(1, \Delta^{-3/4} \frac{L_0^3}{K_0^3}\right). \end{aligned}$$

As the last bound is all that was required, the lemma is proved.

8. The Second Spacing Problem. The term B_1 in (6.37) is bounded above by the number of pairs of minor arc indices i, j , occurring in (6.20), for which

$$(8.1) \quad Q \leq q_i, q_j \leq 2Q,$$

$$(8.2) \quad \left\| \frac{r^2 a_i}{q_i} - \frac{r^2 a_j}{q_j} \right\| \leq \frac{\Delta_1}{r^2} = \frac{Y_1}{1 + Z_1} \asymp \frac{1}{X_1} \asymp \frac{R^4}{r^2 H N Q^2},$$

$$(8.3) \quad \left\| \frac{\overline{q_i} m_i}{r} + \frac{r a_i b_{0,i}}{q_i} - \frac{\overline{q_j} m_j}{r} - \frac{r a_j b_{0,j}}{q_j} \right\| \leq \frac{Y_2}{1 + Z_2} \asymp \frac{1}{X_2} \asymp \frac{R^2}{r H Q},$$

$$(8.4) \quad \left| \frac{1}{\sqrt{\mu_i r^3 q_i^3}} - \frac{1}{\sqrt{\mu_j r^3 q_j^3}} \right| \leq \frac{Y_3}{1 + Z_3} \asymp \frac{1}{X_3} \asymp \frac{R^2}{H N} Y_3$$

(see (6.4), (6.26), (6.33), (6.34)) and

$$(8.5) \quad |\kappa_{0,i} - \kappa_{0,j}| \leq \frac{2Q}{H} = \Delta_4 \quad (\text{say}).$$

By (2.28), (8.1) and (6.26), we may substitute a condition

$$(8.6) \quad \left| \frac{\mu_i q_i^3}{\mu_j q_j^3} - 1 \right| \leq \Delta_2 \asymp \frac{R^2}{H N},$$

in place of (8.4). Having no idea how to use the full strength of (8.3), we multiply through it by r to obtain a weaker, but usable, bound:

$$(8.7) \quad \|\overline{a_i} b_{0,i}/q_i - \overline{a_j} b_{0,j}/q_j\| \leq \Delta_3 \asymp R^2/HQ.$$

As seen in [19, Section 4], conditions (8.1) and (8.2) imply a “magic matrix relation”:

$$(8.8) \quad \begin{pmatrix} a_i \\ q_i \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} a_j \\ q_j \end{pmatrix},$$

where A, B, C and D are integers with

$$(8.9) \quad AD - BC = 1, \quad r^2 |C|$$

and

$$(8.10) \quad |C| \leq \Delta_1 q_i q_j \asymp \Delta_1 Q^2 \asymp R^4/HN.$$

Note that, by the Cauchy–Schwarz inequality, one has

$$(8.11) \quad B_1 \ll V B_1(\Delta_1/V) \quad (V \geq 1)$$

where $B_1(\Delta)$ is the number of pairs of minor arc indices i, j for which (8.1) and (8.3)–(8.5) hold, while the leftmost inequality of (8.2) holds with Δ_1 replaced by Δ . For the arc-index pairs i, j counted by $B_1(\Delta_1/V)$ there is, in place of (8.10), the stronger bound

$$(8.12) \quad |C| \leq V^{-1} \Delta_1 q_i q_j \asymp \Delta_1 Q^2/V \asymp R^4/VHN.$$

The rôle of V here is exactly the same as in [19], although it was there found convenient to introduce it while applying the Double Large Sieve (the effect being similar to a later application of the Cauchy–Schwarz inequality). It is a trivial observation that $B_1(\Delta)$ is a non-decreasing function. On the other hand,

$$(8.13) \quad B_1(\Delta) = B_1(0) \quad (\Delta \leq r^2/4Q^2),$$

since it follows from (8.1) that the left-hand side of (8.2) cannot take any positive value less than $1/2Q(2Q - 1)$.

LEMMA 8.1. *Suppose that (2.1)–(2.9) hold (with $N, R \in \mathbb{N}$), and that Q satisfies (5.2). Suppose also that $P/Q \asymp T/M^2 \geq 1/2r$, that $\Delta_1, \dots, \Delta_4$ are as in (8.2) and (8.5)–(8.7), and that*

$$(8.14) \quad H^2 N^3 \leq MR^4.$$

Then

$$B_1(\Delta_1/V) \ll \frac{R^2}{Q^2} \left(PQ + \Delta_2 \Delta_4^{2/3} P^2 + \Delta_1 \Delta_2 \Delta_4^{2/3} \frac{P^2 Q^2}{r^{2-\varepsilon} V} \log^E R \right)$$

for

$$(8.15) \quad V \geq C^*/\Delta_4,$$

where C^* is some positive constant determined by C_2, \dots, C_5 in (2.2), (2.3) and (2.6)–(2.8).

If

$$(8.16) \quad V \geq 4\Delta_1 Q^2/r^2,$$

then, irrespective of whether or not (8.15) holds,

$$B_1(\Delta_1/V) \ll \frac{R^2}{Q^2} (PQ + \Delta_2 \Delta_4^{2/3} P^2).$$

Proof. Conditions (8.1), (8.5)–(8.9) and (8.12) are, with one exception, the self same conditions considered in [19, Sections 4 and 5] (see [19, (4.13)–(4.16)]). The one exception is that here we have $r^2 \mid C$, whereas [19] had just $r \mid C$. As [19] allows r to be any positive integer, this means that the situation here is simply a special case of that dealt with in [19]. There is also a difference of appearance, since the Δ_1 of [19] is Δ_1/V in our notation here. Note that $a_i/q_i, m_i, \mu_i, b_{0,i}$ and $\kappa_{0,i}$ (after (6.27) and (2.35)) are chosen in exactly the same way as their counterparts in [19] (although in terms of $f(x, 0)$ rather than $f(x)$): even the parameters, T, M, H, N, R and Q , are here subject to all the same constraints as in [19]. The first result of our lemma is essentially just the case $U \asymp \Delta_4^{-2/3}$ of [19, Lemma 5.3] (with r^2 in place of r) although, in point of fact, one must do just a little more work to adapt [19, Lemma 5.2] which is used in the counting of the magic

matrices with $K \leq |C| \leq 2K$ and $r^2 |C$. The number of these matrices is $O(r^{-2}(KP/Q)^2)$ if $P \gg Q$, or $O(r^{-2}(KP/Q)^2 \log^E(2 + K/r))$ if $P \ll Q$ and $K > r^4$, or $O_\varepsilon(r^{-2}(KP/Q)^{2+\varepsilon/4}) = O_\varepsilon(r^{\varepsilon-2}(KP/Q)^2)$ if $P \ll Q$ and $K \leq r^4$. Note that (8.10), (2.7) and (2.8) show one may assume $K \ll R^2$, so that $\log^E(1 + K/r) \ll \log^E R$.

To obtain the second result of the lemma we observe that it follows by (8.13) and (8.16) that $B_1(\Delta_1/V) = B_1(0) = B_1(\Delta_1/V')$ for $V' \geq V$. It therefore follows from the first result of the lemma that

$$B_1(\Delta_1/V) \ll \frac{R^2}{Q^2} \left(PQ + \Delta_2 \Delta_4^{2/3} P^2 + \Delta_1 \Delta_2 \Delta_4^{2/3} \frac{P^2 Q^2}{r^{2-\varepsilon} V'} \log^E R \right)$$

for $V' \geq \max(V, C^*/\Delta_4)$. Taking the limit as $V' \rightarrow \infty$ gives the second result of the lemma.

Finally, we mention that as this proof is ultimately an application of deep results from [14, Sections 14–16] we need to be assured that $f(x, 0)$ has the requisite properties. The only property we might want which does not already follow from (2.5) is [14, (14.3.3)]. This property, however, has no advantage over the simpler [14, (14.3.2)] (implied by (2.5)) except when one deals with the case of lower triangular magic matrices other than the identity. Our lower triangular matrices have $C = (q_i - q_j)/a_j \ll Q/P \asymp M^2/T \ll r$, so that there are at most a bounded number of them with $r^2 |C$. Therefore, as in [19], we have adjusted the scheme of [14, Section 14.2], by reclassifying all non-identity lower triangular matrices as Type 1 (rather than Type 2) of that scheme. Since the Type 1 matrices (including the identity) will remain bounded in number, we estimate their total contribution to $B_1(\Delta_1/V)$ as $O(1)$ times the contribution from the identity matrix. The assumption of [14, (14.3.3)] would offer no significant advantage here, so we assume only (2.5) with its corollary, [14, (14.3.2)]. This suffices for the estimation of the contributions to $B_1(\Delta_1/V)$ arising from the identity matrix, from the Type 3 matrices, and from the remaining (upper triangular) Type 2 matrices.

LEMMA 8.2. *Suppose the hypotheses of Lemma 8.1 (up to and including (8.14)) are true. Then*

$$(8.17) \quad r^{-2} \Delta_1 \Delta_2 \Delta_4^{2/3} P^2 R^2 \asymp \frac{M^2 R^4}{r^2 H^2 N^4} \left(\frac{Q}{H} \right)^{2/3} = \beta(Q) \quad (\text{say}).$$

If

$$(8.18) \quad \frac{N}{2} < N_1 = \frac{M}{T} \left(\frac{M}{H} \right)^{22/35} \left(\frac{T^3}{r^2} \right)^{6/35} \leq N,$$

then

$$B_1 \ll \beta(Q) \left(1 + \frac{r^2}{Q^2 \Delta_1 \Delta_4} \right) r^\varepsilon \log^E R \asymp \beta(Q) \left(1 + \frac{r^4 T^4}{H M^9} \right)^{1/5} r^\varepsilon \log^E R.$$

If

$$(8.19) \quad N_{II} = (M^6/TH^5)^{1/7}, \quad N_{III} = (M^6/r^2T^2H)^{1/3}$$

and

$$(8.20) \quad N/2 < \min(N_{II}, N_{III}) \leq N,$$

then

$$B_1 \ll \beta(Q).$$

Proof. The first result, (8.17), is trivial.

For the first bound on B_1 we use (8.11) and the first bound of Lemma 8.1, taking $V = C^*/\Delta_4$ in both. Note that (8.5) and (5.2) imply that we will have $V \geq 1$ here. The result obtained is

$$B_1 \ll \Delta_4^{-1} \frac{P}{Q} R^2 + \Delta_4^{-1} \frac{r^2}{\Delta_1 Q^2} \beta(Q) + \beta(Q)r^\varepsilon \log^E R.$$

The desired result follows since, after (8.17) and (5.2),

$$(8.21) \quad \begin{aligned} \Delta_4^{-1} \frac{P}{Q} R^2 / \beta(Q) &\asymp \frac{H}{Q} \frac{M}{N} \frac{r^2 H^2 N^4}{M^2 R^4} \left(\frac{H}{Q}\right)^{2/3} \\ &\ll \frac{r^2 H^2 N^3}{MR^4} \left(\frac{H}{R}\right)^{5/3} \asymp \left(\frac{N}{N_I}\right)^{35/6} \asymp 1, \end{aligned}$$

and

$$\frac{r^2}{Q^2 \Delta_1 \Delta_4} \asymp \frac{r^2 H^2 N}{R^4 Q} \ll \frac{r^2 H^2 N}{R^5} \asymp \left(\frac{r^4 T^4}{HM^9}\right)^{1/5},$$

when $N \asymp N_I$ and R is as in (2.6).

For the lemma's second bound on B_1 we apply (8.11) and the second bound of Lemma 8.1 with

$$V = 1 + 4\Delta_1 Q^2 / r^2.$$

Note that by (8.20) we have $N \ll N_{III}$, so it follows from (8.2), (2.6) and (8.19) that

$$\Delta_1 Q^2 / r^2 \asymp R^4 / r^2 H N \asymp M^6 / r^2 T^2 H N^3 = N_{III}^3 / N^3$$

and

$$(8.22) \quad V \asymp \Delta_1 Q^2 / r^2 \asymp R^4 / r^2 H N \gg 1.$$

This application of (8.11) and Lemma 8.1 yields the bound

$$B_1 \ll \left(\frac{PQ}{\Delta_2 \Delta_4^{2/3} P^2} + 1 \right) \beta(Q).$$

The claim of the lemma follows since

$$\Delta_2^{-1} \Delta_4^{-2/3} Q/P \asymp \frac{HN}{R^2} \left(\frac{H}{Q}\right)^{2/3} \frac{N}{M} R^2 \ll \frac{HN^2}{M} \left(\frac{H}{R}\right)^{2/3} \asymp \left(\frac{N}{N_{II}}\right)^{7/3} \ll 1.$$

LEMMA 8.3. *Let*

$$(8.23) \quad 1 \leq r \leq \varepsilon_0^3 T^{1/14}$$

and

$$(8.24) \quad M/(rT)^{1/3} \leq H \leq \varepsilon_0^3 N/r,$$

where either

$$(8.25) \quad r^{-37/57} T^{23/57} \leq M \leq \sqrt{2rT}$$

and $N \in \mathbb{N}$ satisfies (8.18), or

$$(8.26) \quad M \leq r^{11/8} T^{7/16}$$

and $N \in \mathbb{N}$ satisfies (8.19)–(8.20). Then, for some $R \in \mathbb{N}$, every one of the conditions (2.4), (2.6)–(2.9) and (8.14) holds.

Proof. We begin by showing that in order to verify (2.4), (2.6)–(2.9) and (8.14) one need only check that

$$(8.27) \quad 2R_0/\varepsilon_0 \leq H \leq R_0^2/16C_3^2$$

and

$$(8.28) \quad H^2 N^3 \leq MR_0^4$$

when

$$(8.29) \quad R_0 = (C_3 M^3 / 2NT)^{1/2}.$$

Indeed, assuming (8.27)–(8.29) and the hypotheses of the lemma, we can note that (2.6) merely specifies that R be the unique integer in the positive interval $[R_0, R_0 + 1)$, so that (2.7) is implied by (8.27) (in which one must have $2R_0 \geq R_0 + 1$), and (8.14) is implied by (8.28). The lower bound on N in (2.8) is a direct consequence of our hypotheses, (8.23) and (8.24), which ensure that $H \leq rH \leq \varepsilon_0^3 N$. Moreover, it follows from (8.27) that we must have $R_0 \geq 32C_3^2/\varepsilon_0$ and $2R_0 \leq \varepsilon_0 H$, so that we may assume that $32 \leq R_0 \leq \varepsilon_0 H \leq \varepsilon_0^4 N$. It therefore follows from (8.28) that $N \leq \varepsilon_0^{10} M$, so the upper bound of (2.8) may be assumed to hold. The same is true of (2.4), since (2.8) implies $M > N$ and we have already established that $N \geq \varepsilon_0^{-3} rH$ and $H \geq \varepsilon_0^{-1} R_0 \geq 32/\varepsilon_0$. As for (2.9), it is equivalent to the bound $C_3 H N^2 \leq 2MR_0^2$, which follows by one application each, in (8.28), of the established inequalities $R_0 \leq \varepsilon_0 H$ and $R_0 \leq \varepsilon_0^4 N$.

Now suppose that we are in the case described by the lemma where (8.23)–(8.25) are supposed to hold with $N_I \leq N < 2N_I$. In the course of proving Lemma 8.2 we already found that

$$H^2 N^3 / MR_0^4 \asymp r^{-2} (R_0/H)^{5/3},$$

with implicit constant independent of ε_0 (see (8.21)), so it may be assumed that (8.28) holds if (8.27) does. The latter condition will hold provided that

$$\varepsilon_0^{-35/24}(r^2/T^3)^{1/8}M \ll H \ll (r^2/T^3)^{6/13}M^{48/13},$$

with suitable implicit constants independent of ε_0 . By virtue of (8.23), one can show that the lower bound on H here is weaker than that in (8.24). Likewise, (8.25) permits one to establish that the upper bound is weaker than the one in (8.24), provided that ε_0 is sufficiently small in terms of C_3 and absolute constants.

It only remains to consider the second case covered by the lemma, so (contrary to what was assumed for the last paragraph) we shall henceforth suppose that (8.24) and (8.26) hold with $N/2 < \min(N_{II}, N_{III}) \leq N$. As was already observed in (8.22), while proving Lemma 8.2, we will have $r^2HN \ll R_0^4$ when $N \ll N_{III}$. Therefore, and since (8.24) implies $N \geq \varepsilon_0^{-3}rH$, it will follow that $H^2 \ll \varepsilon_0^3R^4/r^3$, which, by (8.23) and our assumptions about ε_0 , will imply the upper bound on H in (8.27). Before turning to the lower bound there, it will be helpful to observe that (8.24) has no solutions, H , unless

$$M/(rT)^{1/3} \ll \min(\varepsilon_0^{7/4}r^{-7/12}T^{-1/12}M^{1/2}, \varepsilon_0^{9/4}r^{-5/4}T^{-1/2}M^{3/2}),$$

where any dependence on ε_0 is explicit. This means that we have nothing to prove unless

$$(8.30) \quad \varepsilon_0^{-9/2}r^{11/6}T^{1/3} \ll M \ll \varepsilon_0^{7/2}r^{-1/2}T^{1/2}.$$

By (8.29), the lower bound on H in (8.27) holds subject to a condition of the form

$$H \gg \varepsilon_0^{-1}T^{-1/2}M^{3/2} \max(N_{II}^{-1/2}, N_{III}^{-1/2}),$$

which is equivalent to

$$H \gg \max(\varepsilon_0^{-14/9}T^{-2/3}M^{5/3}, \varepsilon_0^{-6/5}r^{2/5}T^{-1/5}M^{3/5}).$$

This is weaker than the lower bound of (8.24) when M satisfies a condition of the form

$$\varepsilon_0^{-3}r^{11/6}T^{1/3} \ll M \ll \varepsilon_0^{7/3}r^{-1/2}T^{1/2}.$$

As dependence on ε_0 has been kept explicit, both in the last condition and in (8.30), one can be sure that the latter implies the former, provided only that ε_0 be sufficiently small in terms of the implicit constants (which depend only on C_3). As (8.30) may be assumed, we conclude that the lower bound on H in (8.27) does hold.

To complete the proof we need only establish (8.28). By (8.29) this may be rewritten as the condition $N^5 \leq C_3^2M^7/4T^2H^2$. Therefore, and by (8.19)

and (8.20), we find that if (8.28) does not hold, then H must lie in the range

$$2^{-21}C_3^6r^{10}T^4M^{-9} < H < 2^{49/11}C_3^{-14/11}T^{9/11}M^{-19/11}.$$

If this range is not empty, then $M > 2^{-7/2}C_3r^{11/8}T^{7/16}$. Given that $C_3 \geq 2^{7/2}$ (see “Notation and conventions”), the last bound contradicts (8.26), so we conclude that (8.28) must hold. We have, in all cases, checked and verified those conditions, (8.27)–(8.29), that we earlier showed were sufficient, so the lemma is proved.

LEMMA 8.4. *Suppose that (2.1)–(2.3), (8.23) and (8.24) hold, and that either*

$$(8.31) \quad (M/H)^{1/10}(rT)^{2/5} \leq M \leq \sqrt{2rT}$$

and $N \in \mathbb{N}$ satisfies (8.18), or

$$(8.32) \quad M < (M/H)^{1/10}(rT)^{2/5}$$

and $N \in \mathbb{N}$ satisfies (8.19)–(8.20). Then, for some $R \in \mathbb{N}$, every one of the conditions (2.4), (2.6)–(2.9) and (8.14) holds and, for Q satisfying (5.2), we have

$$B_1 \ll \beta(Q)r^\varepsilon \log^E R,$$

with $\beta(Q)$ given by (8.17).

Proof. For the first assertion, about the conditions (2.4), (2.6)–(2.9) and (8.14), two observations suffice. The first observation is that (8.18) and (8.24) may be assumed to imply

$$(M/H)^{57/35} \geq rT(r^2/T^3)^{6/35} = r^{47/35}T^{17/35},$$

so that when (8.31) holds we will have

$$M \geq (r^{47/57}T^{17/57})^{1/10}(rT)^{2/5} = r^{55/114}T^{49/114},$$

which is stronger than (8.25). The second observation is that (8.24) alone implies $M/H \leq (rT)^{1/3}$, so that when (8.32) holds we will have

$$M < (rT)^{1/30}(rT)^{2/5} = (rT)^{13/30},$$

which is stronger than (8.26). These observations show that the hypotheses of the lemma imply those of Lemma 8.3, which has the assertion we seek to justify as its conclusion.

The bound for B_1 follows from Lemma 8.2 (and the assertion just justified) on noting that the lower bound for M in (8.31) implies $r^4T^4 \leq HM^9$.

9. Bounds for $S(H, M)$

LEMMA 9.1. *Suppose that the hypotheses of Lemma 8.4 hold. Then*

$$S(H, M) \ll M \left(\frac{rH}{N} \right)^{1/2} \log(rN) + \frac{M^{1/2}H}{r^{1/4}R^{1/2}} \alpha^{1/4}(\beta(R))^{1/4}r^\varepsilon \log^{E'} N,$$

where $R \gg 1$ is the integer satisfying (2.6) and

$$\alpha = r^4 \left(\frac{N^2 H^2}{R^4} \log^6(rN) + (rN)^\varepsilon \min \left(\frac{N^{3/2} H^3}{R^{9/2}}, \frac{N^3}{R^3} \right) \right) \\ + r^{9/2} \left(\frac{N^2 H}{R^3} \log(rN) + (rN)^\varepsilon \min \left(\frac{N^{4/3} H^3}{R^{13/3}}, \frac{N^{8/3} H^{1/3}}{R^3} \right) \right), \\ \beta(Q) = \frac{M^2 R^4}{r^2 H^2 N^4} \left(\frac{Q}{H} \right)^{2/3}.$$

Proof. By the construction in Section 2,

$$(9.1) \quad S(H, M) = S_\infty(H, M) + S_\pi(H, M),$$

where $S_\infty(H, M)$ denotes the contribution from major arcs $J(c/d)$, while $S_\pi(H, M)$ is the contribution from all minor arcs J_i that belong to no major arc (see (1.6), (2.1), (2.17)–(2.18) and (2.21)–(2.22)). By (3.13), (2.7) and (2.8),

$$(9.2) \quad S_\infty(H, M) \ll \frac{MR}{\sqrt{r^3 H N}} \log R \ll M \sqrt{\frac{H}{r^3 N}} \log N.$$

By (2.16),

$$(9.3) \quad S_\pi(H, M) = \sum_{R \leq Q = 2^k R \leq 120C_3 H r} S(H, M; Q),$$

where k runs over the integers and $S(H, M; Q)$ is the sum $\sum_i^{(Q)} \mathcal{B}_i$ defined a few lines above (6.18). By (2.30), (2.33), (4.5), (4.22), (5.1), (2.7) and (2.8), one has

$$\mathcal{B}_i \ll |\mathcal{B}_{i,0}^*| + \sqrt{r} \left(\frac{Q}{R} \sqrt{HN} + Q \log(rQ) \right) \ll \frac{Q}{R} \sqrt{rHN} \log(rN),$$

when (4.1) holds with $H \ll Q \ll rH$. Therefore, and by (6.18)–(6.22) and (6.35)–(6.37), each sum $S(H, M; Q)$ in (9.3) is such that either it satisfies

$$(9.4) \quad S(H, M; Q) = \sum_i^{(Q)} \mathcal{B}_i \ll \left(\frac{Q}{R} \sqrt{rHN} \log(rN) \right) W(Q) \\ \ll \frac{R}{Q} M \sqrt{rH/N} \log(rN),$$

or Q is in the range given by (5.2) and

$$(9.5) \quad (S(H, M; Q))^4 \ll (\log N)^8 (W(Q))^2 \frac{N^2 R^4}{r^2 Q^2} \frac{H N K L^2}{R^2} A_1 \frac{H}{Q} B_1 \\ \ll \frac{M^2 H^4 R^4}{r Q^6} K^2 A_1 \beta(Q) r^\varepsilon \log^{E+8} N$$

(see (6.10), Lemma 8.4 and (8.17)).

By (6.9), (6.10), (7.51) and (2.7), one finds that Lemma 7.35 implies $K^2 A_1 \ll (Q/R)^5 \alpha$. Therefore it follows from (9.4), (9.5) and the definition of $\beta(Q)$ that, for Q relevant to (9.3), the sum $S(H, M; Q)$ is $O((R/Q)^{1/12})$ times the bound the lemma claims for $S(H, M)$. As $\sum_{k=0}^\infty 2^{-k/12}$ converges, we may now deduce the claimed bound from (9.1)–(9.3) and our last observation, so the lemma is proved.

LEMMA 9.2. *Suppose that (2.2) and (2.3) hold, and that $1 \leq r \leq \varepsilon_0^3 T^{1/14}$, where $T \geq \exp(\varepsilon_0^{-2})$. Let*

$$(9.6) \quad U \geq \begin{cases} r^\varepsilon (\log T)^{E'} r^{51/146} T^{23/73} & \text{if } r \leq T^{2/93 - \sqrt{\varepsilon}}, \\ r^{149/384} T^{181/576 + \sqrt{\varepsilon}} & \text{if } T^{2/93 - \sqrt{\varepsilon}} < r < T^{10/273}, \\ r^{47/57} T^{17/57 + \sqrt{\varepsilon}} & \text{if } T^{10/273} < r \leq \varepsilon_0^3 T^{1/14} \end{cases}$$

and suppose furthermore that

$$(9.7) \quad \max_{i=1, \dots, 5} L_i(U) \leq M \leq \sqrt{2rT},$$

where

$$(9.8) \quad L_i(U) = \begin{cases} \frac{r^{5/2} T}{U^2} (\log T)^3 & \text{if } i = 1, \\ \frac{r^{\sqrt{\varepsilon} + 17/8} T^{7/4}}{U^{17/4}} (\log T)^{9(2E' + 3)/4} & \text{if } i = 2, \\ \frac{r^{\sqrt{\varepsilon} + 49/22} T^{25/22}}{U^{53/22}} (\log T)^{9(4E' + 1)/11} & \text{if } i = 3, \\ T^{\sqrt{\varepsilon}} \min \left(\frac{r^{37/19} T^{71/38}}{U^{175/38}}, \frac{r^{14/5} T^{13/10}}{U^{29/10}} \right) & \text{if } i = 4, \\ T^{\sqrt{\varepsilon}} \min \left(\frac{r^{49/22} T^{37/22}}{U^{89/22}}, \frac{r^{41/14} T^{17/14}}{U^{37/14}} \right) & \text{if } i = 5. \end{cases}$$

Then, for

$$(9.9) \quad M/(rT)^{1/3} \leq H \leq M/U,$$

we have

$$S(H, M)/M \ll (UH/M)^\lambda \quad (0 \leq \lambda \leq \lambda_0),$$

where λ_0 is some positive absolute constant.

Proof. We shall first treat the case $\lambda = 0$, postponing discussion of the cases where $\lambda > 0$ until the end. By (2.6), (2.8) and (9.7), the bound given for $S(H, M)$ by Lemma 9.1 (when it applies) may be written in the form

$$(9.10) \quad \frac{S(H, M)}{M} \ll \sqrt{\frac{rH}{N}} \log T + \frac{r^{1/4} T^{1/6} H^{5/6}}{M^{1/2} N^{1/3}} r^\varepsilon (\log T)^{E' + 3/2}$$

$$\begin{aligned}
 &+ \frac{r^{3/8} T^{1/24} H^{7/12}}{M^{1/8} N^{11/24}} r^\varepsilon (\log T)^{E'+1/4} \\
 &+ r^{1/4} \min \left(\frac{T^{11/48} H^{13/12}}{M^{11/16} N^{19/48}}, \frac{T^{1/24} H^{1/3}}{M^{1/8} N^{5/24}} \right) T^{\varepsilon/4} \\
 &+ r^{3/8} \min \left(\frac{T^{5/24} H^{13/12}}{M^{5/8} N^{11/24}}, \frac{T^{1/24} H^{5/12}}{M^{1/8} N^{7/24}} \right) T^{\varepsilon/4}.
 \end{aligned}$$

We assert that, in order to prove the case $\lambda = 0$ of the lemma, it will suffice to show that, subject to the stated hypotheses, every term on the right of (9.10) is less than or equal to 1 if (8.31) holds and $N = N_I$ (from (8.18)), or if (8.32) holds and N equals the minimum of N_{II} and N_{III} (from (8.19)). This assertion is justified since the first term on the right of (9.10) is greater than 1 unless $N > rH \log^2 T$, which (as $\log T \geq \varepsilon_0^{-2}$) implies the upper bound on H in (8.24), thereby verifying that, for some integer N , the hypotheses of Lemma 8.4 will hold and Lemma 9.1 will apply (note that (8.24) implies $N \geq \varepsilon_0^{-3}$, so that both (8.18) and (8.20) allow that a non-integer N may be replaced by the nearest greater integer).

Since all the powers of N in (9.10) are negative, and since (8.32) implies $N_I < N_{II}$, it follows that, for $\lambda = 0$, all we need do now is verify that, if (9.7) and (9.9) hold and $N \in \{N_I, N_{III}\}$, then no term on the right of (9.10) is greater than 1. The further observation that the powers of H in (9.10) are positive, while N_I and N_{III} from (8.18) and (8.19) involve only negative powers of H , leads us to the conclusion that it is sufficient to consider only the “worst” case, in which $H = M/U$. Substituting N_I (from (8.18)) for N in (9.10), the right hand side becomes:

$$\begin{aligned}
 &r^{47/70} T^{17/70} (H/M)^{57/70} \log T + r^{51/140} T^{23/70} (H/M)^{73/70} r^\varepsilon (\log T)^{E'+3/2} \\
 &+ r^{149/280} T^{37/140} (H/M)^{61/90} r^\varepsilon (\log T)^{E'+1/4} \\
 &+ T^{\varepsilon/4} \min(r^{27/70} T^{59/140} (H/M)^{373/280}, r^{9/28} T^{1/7} (H/M)^{13/28}) \\
 &+ T^{\varepsilon/4} \min(r^{149/280} T^{181/420} (H/M)^{48/35}, r^{19/40} T^{11/60} (H/M)^{3/5}).
 \end{aligned}$$

Given that $H = M/U$, all of the above terms will be less than or equal to 1, provided only that

$$U \geq \max_{j=1, \dots, 5} U_j^*,$$

where

$$U_j^* = \begin{cases} r^{47/57} T^{17/57} (\log T)^{70/57} & \text{if } j = 1, \\ r^\varepsilon (\log T)^{E'} r^{51/146} T^{23/73} & \text{if } j = 2, \\ r^{2\varepsilon} (\log T)^{2E'} r^{149/244} T^{37/122} & \text{if } j = 3, \\ T^\varepsilon \min(r^{9/13} T^{4/13}, r^{108/373} T^{118/373}) & \text{if } j = 4, \\ T^\varepsilon \min(r^{149/384} T^{181/576}, r^{19/24} T^{11/36}) & \text{if } j = 5. \end{cases}$$

When $r = T^{2/93}$ the product $r^{51/146}T^{23/73}$ and each of the four similar products involved in the two minima above equate to $T^{10/31}$, while the products $r^{47/57}T^{17/57}$ and $r^{149/244}T^{37/122}$ both equate to smaller powers of T . Therefore a comparison of the relative sizes of the powers of r occurring in the above lower bound for U allows us to conclude that, for $r \leq T^{2/93-\sqrt{\varepsilon}}$, we require only $U \geq r^\varepsilon(\log T)^{E'}r^{51/146}T^{23/73}$, while for $r > T^{2/93-\sqrt{\varepsilon}}$ it is sufficient that

$$U \geq T^{\sqrt{\varepsilon}} \max(r^{47/57}T^{17/57}, r^{149/244}T^{37/122}, r^{149/384}T^{181/576}).$$

As $47/57 > 149/244 > 149/384$, and since the first and third terms within the last maximum are equal to and greater than the second when $r = T^{10/273}$, the requirements here are nothing other than (9.6).

Recalling the observations made earlier, we may now complete the proof of the case $\lambda = 0$ by treating the cases where $H = M/U$ and $N = N_{III}$ (from (8.19)). This is a trivial matter, since the left half of (9.7), with $L_1(U), \dots, L_5(U)$ as stated, is exactly what is required to guarantee that no term on the right of (9.10) exceeds 1 in those cases.

For the case $0 < \lambda \leq \lambda_0$, we recall our earlier implicit recognition that (9.10) involves only positive powers of H when either N_I or N_{III} is substituted for N . As we always substituted M/U for H in the case $\lambda = 0$, it follows that in doing so we unnecessarily weakened our resulting bounds for $S(H, M)$ by some factor $(M/UH)^{\lambda_0}$, where λ_0 is a positive absolute constant. It follows that we could actually show $S(H, M)/M \ll (UH/M)^{\lambda_0}$, which, by (9.9), implies the bound claimed by the lemma for any $\lambda \leq \lambda_0$.

LEMMA 9.3. *Let δ and η be arbitrary positive constants. Let*

$$(9.11) \quad S = \sum_{H_1 < h \leq H} \sum_{M_1 < k \leq M} \chi\left(\frac{k+h}{k-h}\right) \left(\frac{k+h}{k-h}\right)^{-2\pi iT},$$

where (1.7) holds, $T \geq 1$ and $H/M \leq 1/4$. Then

$$(9.12) \quad S/M \ll (rT)^{1/2}(H/M)^{3/2} + r^{1/2}T^{-1/2}(H/M)^{1/2}M.$$

Suppose additionally that H/M is sufficiently small in terms of δ , and that

$$(9.13) \quad V = \frac{H}{M}T, \quad \omega = \frac{\log M}{\log V} \in \left[\frac{49}{114}, 1 - \delta\right].$$

Then

$$(9.14) \quad \frac{S}{M} \ll \sqrt{r} \frac{H}{M} V^\eta Z(M, V),$$

where

$$(9.15) \quad Z(M, V) = \begin{cases} M^{1/2}V^{89/570} & \text{if } 49/114 \leq \omega \leq 65/114, \\ M^{7/20}V^{29/120} & \text{if } 65/114 \leq \omega \leq 7/12, \\ M^{13/20}V^{1/15} & \text{if } 7/12 \leq \omega \leq 517/873, \\ M^{25/128}V^{43/128} & \text{if } 517/873 \leq \omega \leq 19/31, \\ M^{107/280}V^{31/140} & \text{if } 19/31 \leq \omega \leq 374/601, \\ M^{187/641}V^{178/641} & \text{if } 374/601 \leq \omega \leq 872/1295, \\ M^{53/244}V^{20/61} & \text{if } 872/1295 \leq \omega \leq 188/275, \\ M^{26/73}V^{17/73} & \text{if } 188/275 \leq \omega \leq 1 - \delta. \end{cases}$$

Proof. By (1.6) and (3.2),

$$S = \frac{1}{r} \sum_{b \bmod r} \sum_{H_1 < h \leq H} W(b, -h)S_h(M; b/r),$$

where $W(x, y)$ is given by (3.3) and where, for $\theta \in \mathbb{R}$,

$$(9.16) \quad S_h(M, \theta) = \sum_{M_1 < k \leq M} e(T \log(k - h) - T \log(k + h) + \theta k).$$

Using the bound (3.4) we deduce that

$$(9.17) \quad |S| \leq \sum_{H_1 < h \leq H} \left(|S_h(M; 0)| + \frac{2}{\sqrt{r}} \sum_{b=1}^{r-1} |S_h(M; b/r)| \right) \leq 3\sqrt{r}H \max_{\substack{0 \leq \theta < 1 \\ h \in [H/2, H]}} |S_h(M; \theta)|.$$

As for $S_h(M; \theta)$ in (9.16)–(9.17), we note first that the application of [29, Theorem 5.9], with

$$-f''(x) = \frac{T}{(k - h)^2} - \frac{T}{(k + h)^2} = \frac{4khT}{(k^2 - h^2)^2} \asymp \lambda_2 = \frac{HT}{M^3},$$

yields us the bound

$$S_h(M; \theta) \ll M\lambda_2^{1/2} + \lambda_2^{-1/2} = (HT/M)^{1/2} + (HT/M)^{-1/2}M$$

for $H/2 \leq h \leq H \leq M_1/2$ and $\theta \in \mathbb{R}$. Taken together with (9.17), this last bound for $S_h(M; \theta)$ immediately implies (9.12), the lemma’s first bound for S/M .

The second part of the lemma (the bound (9.14)) will follow from the uniform bound

$$(9.18) \quad S_h(M; \theta) \ll V^\eta Z(M, V) \quad (H/2 \leq h \leq H, 0 \leq \theta < 1),$$

which is what we next establish. In doing so we may, and shall, assume that all the stated conditions for the second part of the lemma do hold.

Applying [14, Lemma 5.5.3] (inversion) to the sum in (9.16), we find that, for $\varrho = (h/M)^2 < 1/9$,

$$\begin{aligned} \overline{S_h(M; \theta)} &= \sum_{A \leq n - \theta \leq B} \gamma(n - \theta) e\left(\phi(n - \theta) + \frac{1}{8}\right) \\ &\quad + O\left(\frac{M}{\sqrt{V}} + \log\left(1 + \frac{\sqrt{V}}{M}\right)\right), \end{aligned}$$

where $A = 2hT/(M^2 - h^2)$, $B = 2hT/(M_1^2 - h^2)$,

$$\gamma(x) = \frac{\sqrt{hT}}{x} \left(\frac{2hT}{x} + h^2\right)^{-1/4}, \quad \phi(x) = \int_0^x \sqrt{\frac{2hT}{y} + h^2} dy.$$

As $\gamma(x)$ is monotone in x we may readily apply partial summation and a division to deduce that if

$$(9.19) \quad \log^2 V \leq M \leq V/2,$$

then one has

$$(9.20) \quad S_h(M; \theta) \ll \frac{M}{\sqrt{V}} \left(\left| \sum_{N < n - \theta \leq N_1} e(\phi(n - \theta)) \right| + \sqrt{N} \right)$$

for some pair $N, N_1 \in [A, B]$ with $N \leq N_1 \leq 3N/2$. Since $\varrho < 1/9$, for some $M_2 \in [M/3, M]$ we have here $N = V_2/M_2$, with $V_2 = 2hT/M_2$, and $\phi(x) = V_2\Phi(x/N)$, where

$$\Phi(z) = \int_0^z \sqrt{u^{-1} + \varrho_2} du,$$

with $\varrho_2 = h/M_2$. Note that $M_2 \asymp M$, $V_2 \asymp V$, $\varrho_2 \asymp \varrho$ and $N \asymp V/M$.

Now write the sum over n in (9.20) as

$$(9.21) \quad \sum_{N < n \leq N_1 + \theta} e\left(V_2\Phi\left(\frac{n}{N} - \frac{\theta}{N}\right)\right) - \sum_{N < n - \theta \leq N + \theta} e\left(V_2\Phi\left(\frac{n}{N} - \frac{\theta}{N}\right)\right).$$

We may assume M large enough that (9.19) guarantees $2 \leq N \leq V_2/2$. As $0 \leq \theta < 1 \leq N/2$ it follows that the last two sums above are suitable candidates for the application of either one of [12, Theorems 1, 3], subject only to the verification of certain conditions involving derivatives of $\Phi(x - \theta/N)$ on an interval $[1 - \Delta, 2 + \Delta] \supset [1, 2]$ for x . Any positive absolute constant Δ will do, so take $\Delta = 1/6$ in our application. As translation commutes with differentiation, these conditions may be reformulated as conditions to be satisfied by derivatives of $\Phi(x)$ on the interval $[1 - 1/6 - \theta/N, 2 + 1/6 - \theta/N] \subseteq [1/3, 3]$. From the observation that

$$\Phi'(x) = \sqrt{x^{-1} + \varrho_2} = x^{-1/2} + \frac{\sqrt{1 + \varrho_2 x} - 1}{\sqrt{x}}$$

(where $\varrho_2 \asymp H^2/M^2$ is assumed to be sufficiently small), it follows that, in respect of the interval $[1/3, 3]$, for x , $\Phi(x)$ does satisfy the “virial” condition discussed in [12, pages 38–39]. The utility of the virial condition is that conditions involving derivatives of $\Phi(x)$ are implied by the corresponding conditions involving derivatives of x^{-s} , where s is some positive constant ($s = 1/2$ in our application). Verification of the relevant conditions of [12, Theorems 1, 3] therefore reduces to checking that certain polynomials do not vanish at s . It so happens that this non-vanishing is guaranteed for $s > 0$, it being evident from their factorisations over \mathbb{Q} that all real roots of those polynomials lie in $(-\infty, 0]$.

Having established the conditions we may bound the sums in (9.21) using the results of [12, Theorems 1, 3] summarised on [12, page 39] within rows 3–10 of the table given there (it being observed from [12, Theorem 3] itself that the last of these results holds in a wider range of cases than is indicated in the table). The reader may verify that, as $N \asymp V/M$, the resulting bounds for the sum over n in (9.20) do imply the desired bounds for $S_h(M; \theta)$, set out in (9.18) and (9.15). Points to note are that the term \sqrt{N} in (9.20) is always dominated by the bound for the adjacent sum, and that (9.19) is a valid assumption to make, given that M is large enough, for it is then a consequence of our hypothesis (9.13) (that we may assume M is so large follows from the hypothesis that M/H be sufficiently large and the observation that $S = 0$ unless $H \geq 1$).

LEMMA 9.4. *Let $\varepsilon, \varepsilon_0$ be sufficiently small positive absolute constants satisfying $\varepsilon \geq 8E'\varepsilon_0$. Let S be as in (9.11), where $T \in \mathbb{R}$ and (1.7) holds. Suppose that $T \geq \exp(\varepsilon_0^{-2})$, $1 \leq r \leq \varepsilon_0^3 T^{1/14}$ and*

$$(9.22) \quad U \geq \max_{j=1, \dots, 5} U'_j,$$

where

$$(9.23) \quad U'_j = U_j(r, T)$$

and $U_j(r, t)$ is as defined in (1.4). Then, for

$$(9.24) \quad M \leq \sqrt{2rT}, \quad M/(rT)^{1/3} \leq H \leq M/U,$$

we have

$$S/M \ll (UH/M)^\lambda \quad (0 \leq \lambda \leq \lambda_0),$$

where λ_0 is some positive absolute constant.

Proof. With the possible exception of the lower bound for M in (9.7), we have all the right conditions for applying Lemma 9.2. Indeed, the sum S in (9.11) is merely the case $F(X) = -\log X$ of the sum $S(H, M)$ in (1.6) and Lemma 9.2, so (2.2) and (2.3) hold with $C_j = ((j - 1)!)3^j$ for $j = 2, \dots, 5$, and it is therefore enough that ε_0 be sufficiently small in absolute terms. As

for (9.6), we may there replace each ε with ε^2 (provided that we do likewise in (9.8)). It is then seen that (9.6) holds by virtue of the bounds

$$U \geq U'_3 \geq r^{149/384} T^{181/576+\varepsilon} \quad (r \geq T^c),$$

where $c = 5942/289551 < 2/93 - 10^{-4}$ (see (9.22), (9.23)). We conclude from the above that, in cases where (9.7) does hold, an appeal to Lemma 9.2 yields the desired bound, $S/M \ll (UH/M)^\lambda$, provided that λ is (in absolute terms) sufficiently small.

As the case $H < 1$ is trivial, it follows from (9.9), the last paragraph and (9.7) that the case yet to be settled is that where

$$(9.25) \quad U \leq M < \max_{i=1,\dots,5} L_i(U),$$

with the $L_i(U)$'s as in (9.8). As the constraints on M and H can only weaken if U be decreased (see (9.24) and (9.8)), while the lemma's bound for S would strengthen, so it suffices to treat just the special case where (9.22) holds with equality. We make a division of this special case into narrower cases, numbered 1 to 5, where it is supposed of Case j that $U = U'_j$ and $U'_j \geq U'_k$ for $k = 1, \dots, 5$.

By finding what values of r result in equalities amongst U'_1, \dots, U'_5 , we are able to put limits on the range for r in each of Cases 1, \dots , 5. For $j = 1, \dots, 5$, we obtain both upper and lower bounds for M in Case j by substituting U'_j for U in (9.25). Comparison of $L_1(U), \dots, L_5(U)$, in each one of the five cases $U = U'_1, \dots, U'_5$, leads to simplification of the upper bound for M given by (9.25). Indeed it turns out that, if ε 's and the powers of $\log T$ could be ignored, then we would have (in Case j)

$$M < \begin{cases} L_2(U'_j) & \text{if } j = 1, 2, \\ L_5(U'_j) & \text{if } j = 3, \\ L_1(U'_j) & \text{if } j = 4, 5. \end{cases}$$

As ε 's and powers of $\log T$ cannot safely be ignored the actual bounds are marginally different. We pass over the details of the calculations to state the following valid conclusions.

In Case 1,

$$(9.26) \quad 1 \leq r \leq T^{932/45017}$$

and

$$(9.27) \quad U = U'_1 \leq M < K_1 = r^{187/292-\varepsilon} T^{30/73} (\log T)^{E'}.$$

In Case 2,

$$(9.28) \quad T^{932/45017-11\varepsilon} \leq r \leq T^{392/18297}$$

and

$$(9.29) \quad U = U'_2 \leq M < K_2 = r^{1513/6455+\varepsilon} T^{2707/6455-\varepsilon} (\log T)^{5E'}$$

In Case 3,

$$(9.30) \quad T^{392/18297} \leq r \leq T^{245/7103}$$

and

$$(9.31) \quad U = U'_3 \leq M < K_3 = r^{23213/69322+\varepsilon} T^{14461/34661-\varepsilon} (\log T)^{5E'}$$

In Case 4,

$$(9.32) \quad T^{245/7103} \leq r \leq T^{524/11155}$$

and

$$(9.33) \quad U = U'_4 \leq M < K_4 = r^{2813/2386+\varepsilon} T^{463/1193-\varepsilon} (\log T)^{5E'}$$

In Case 5,

$$(9.34) \quad T^{524/11155} \leq r \leq \varepsilon_0^3 T^{1/14}$$

and

$$(9.35) \quad U = U'_5 \leq M < K_5 = r^{97/114+\varepsilon} T^{23/57-\varepsilon} (\log T)^{5E'}$$

In order that Lemma 9.3 may be applied it suffices that

$$(9.36) \quad (HT/M)^{49/114} \leq M \leq (HT/M)^{1-\delta}$$

for some positive absolute constant δ . By (9.24) and (9.25), we are assured of (9.36) if

$$(T/U)^{49/114} \leq U, \quad M \leq (T/(rT)^{1/3})^{1-\delta}.$$

Therefore Lemma 9.3 will apply provided that

$$(9.37) \quad U \geq T^{49/163}, \quad M \leq r^{(\delta-1)/3} T^{2(1-\delta)/3}.$$

The first condition here is immediate (on comparing the powers of T involved) for all but the last of Cases 1, . . . , 5; and in Case 5 we have merely to note from (9.34) that $r > T^{1/22}$ and therefore that (9.35) and (9.23) imply

$$\frac{\log U}{\log T} > \frac{47/22 + 17}{57} > \frac{19}{57} = \frac{1}{3} > \frac{49}{163}.$$

As for the upper bound on M in (9.37), we may set $\delta = 3\varepsilon$ and write it in the form

$$(9.38) \quad Mr^{1/3-\varepsilon} T^{-2/3} \leq T^{-2\varepsilon}.$$

By (9.26)–(9.35), we find that in none of the Cases 1, . . . , 5 does the left-hand side of (9.38) exceed $(\log T)^{5E'} T^{-5/28-\varepsilon}$, which is the bound calculated in respect of Case 5. Given the constraints imposed on T , ε and ε_0 by the lemma, this last point implies that (9.38) holds throughout each of Cases

1, . . . , 5. As this completes our verification of (9.36) for Cases 1, . . . , 5, we may conclude that throughout each of those cases Lemma 9.3 will apply to give the bound

$$\frac{S}{M} \ll \sqrt{r} \frac{H}{M} \left(\frac{HT}{M}\right)^\eta Z\left(M, \frac{HT}{M}\right),$$

where η is any positive absolute constant. By (9.15) one may observe that the bound on S/M is $H^{1+\eta}$ times a continuous increasing function of H , so (9.24) implies that, provided that $0 \leq \lambda \leq 1 + \eta$, the last upper bound for S/M is smaller, always by at least a factor $(M/UH)^\lambda$, than $\sqrt{r}U^{-1}(T/U)^\eta Z(M, T/U)$. The latter expression represents a continuous increasing function of M , so we may conclude after (9.27), (9.29), (9.31), (9.33) and (9.35) that, for $j = 1, 2, 3, 4, 5$ and $\lambda \leq 1$, one has $S/M \ll (UH/M)^\lambda$ throughout Case j , provided only that

$$(9.39) \quad \Omega_j(r) = \frac{\sqrt{r}}{U'_j} Z\left(K_j, \frac{T}{U'_j}\right)$$

satisfy a bound

$$(9.40) \quad \Omega_j(r) \ll T^{-\eta}$$

(where η is some positive absolute constant), for all r satisfying whichever one of (9.26), (9.28), (9.30), (9.32), (9.34) corresponded to Case j .

Turning first to Case 1, we need only discuss $Z(M, V)$ with

$$(9.41) \quad \begin{aligned} \omega &= \frac{\log M}{\log V} = \frac{\log K_1}{\log(T/U'_1)} \\ &= \frac{(187/292 - \varepsilon)x + 30/73 + E'(\log \log T)/\log T}{1 - (51/146 + \varepsilon)x - 23/73 - E'(\log \log T)/\log T} \\ &= (1 + O(\varepsilon))\gamma(x), \end{aligned}$$

where $x = (\log r)/\log T$ and

$$\gamma(x) = \frac{187x + 120}{200 - 102x}$$

(see (9.39), (9.27) and (9.23)). As $\gamma(x)$ is increasing it follows by (9.41) and (9.26) that, in respect of Case 1, the validation of condition (9.39) only requires evaluation of the function $Z(M, V)$ in cases where

$$(9.42) \quad \frac{3}{5} + O(\varepsilon) \leq \omega = \frac{\log M}{\log V} \leq \frac{19097}{30508} + O(\varepsilon).$$

This means that only the fourth, fifth and sixth rows of (9.15) are relevant for (9.39) and (9.40) in Case 1. Given (9.23) and (9.27), trial of each of the three relevant forms for $Z(M, V)$ in (9.39) shows each to result in an expression for $\Omega_1(r)$ (on a subinterval of its domain) that is an increasing function of r . As $Z(M, V)$ is continuous in both M and V , we must conclude

that $\Omega_1(r)$ is, like ω , an increasing function of r throughout the range (9.26). It follows that the “worst” subcases of Case 1 are those where ω is at the upper end of the range given in (9.42). In such subcases it is the sixth row of (9.15) which applies, by way of (9.39), (9.27) and (9.23), to yield

$$\Omega_1(r) \ll \frac{r^{45017/187172-\varepsilon}}{T^{932/187172}(\log T)^{E'/2}}.$$

We conclude from this that if $j = 1$ then (9.40) holds with

$$(9.43) \quad \eta = \frac{932}{45017} \varepsilon > 0$$

for all r satisfying (9.26).

In Cases 2, 3 and 4 our argument initially parallels that used for Case 1. Taking all three of those cases together one has, in place of (9.42),

$$\frac{19097}{30508} + O(\varepsilon) \leq \omega = \frac{\log M}{\log V} = \frac{\log K_j}{\log(T/U'_j)} \leq \frac{4947}{7396} + O(\varepsilon) \quad (j = 2, 3, 4)$$

(see (9.28)–(9.33) and (9.23)). Given that ε is sufficiently small, this means that only the sixth row of (9.15) is relevant for (9.39) in Cases 2, 3 and 4. Therefore, and by (9.39), (9.29), (9.31), (9.33), (9.23) and row six of (9.15) (as well as the lemma’s conditions upon r, T, ε and ε_0), we find that

$$(9.44) \quad \Omega_j(r) \ll r^\varepsilon T^{-\varepsilon} (\log T)^{2E'} \leq T^{\varepsilon/14-\varepsilon+\varepsilon/2} = T^{-(3/7)\varepsilon}$$

for $j = 2, 3, 4$ and all r satisfying whichever of (9.28), (9.30), or (9.32) is appropriate for Case j . The reason that r and T only appear raised to the powers plus or minus ε in (9.44) is that U'_2, U'_3 and U'_4 are essentially the minimal choices for U such that every one of the cases in which we need to prove $S = O(M)$ will be covered by one or other of Lemmas 9.2, 9.3. In contrast to this U'_1 and U'_5 were chosen with reference only to the lower bound (9.6) required in Lemma 9.2.

In Case 5, as in Case 1, the relevant $\omega = (\log K_5)/\log(T/U'_5)$ is again clearly increasing with respect to r (see (9.35) and (9.23)). Using (9.35), (9.23) and (9.34) one finds that

$$(9.45) \quad \frac{4947}{7396} + O(\varepsilon) \leq \omega \leq \frac{13}{18} + O(\varepsilon),$$

so only the last three rows of (9.15) are relevant for (9.39) in Case 5. Using any one of the relevant three forms for $Z(M, V)$ in (9.39), together with (9.23) and (9.35), produces an expression for $\Omega_5(r)$ (on some interval) that is decreasing in r . Therefore, we may argue as in Case 1 that the continuity of $Z(M, V)$ implicit in (9.15) must lead us to the conclusion that $\Omega_5(r)$ is decreasing in r over the whole range (9.34). It follows that the “worst” subcases are where r and (therefore) ω are at the lower ends of their respective ranges (9.34), (9.45). Therefore we need only consider what happens when

$\omega = 4947/7396 + O(\varepsilon)$. As ε is assumed to be sufficiently small it will be the sixth row of (9.15) which applies in such cases, leading us by way of (9.39), (9.35) and (9.23) to the bound

$$\Omega_5(r) \ll \frac{T^{524/36537-\varepsilon} (\log T)^{2E'}}{r^{11155/36537-\varepsilon}}.$$

We conclude from this that if $j = 5$, then (9.40) holds with

$$(9.46) \quad \eta = \left(\frac{1}{2} - \frac{524}{11155}\right)\varepsilon > 0$$

for all r satisfying (9.34).

As (9.43), (9.44) and (9.46) together complete the verification of (9.40), for r in whichever of the ranges (9.26), (9.28), (9.30), (9.32) or (9.34) corresponds to Case j , and for $j = 1, \dots, 5$, so it follows that, if $\lambda \leq 1$, then $S/M \ll (UH/M)^\lambda$ throughout each of Cases $1, \dots, 5$. In view of our remarks just before and after (9.25), this last conclusion makes the proof of the lemma complete.

10. Application to $I(t, U, \chi)$: the proof of Theorem 1. We begin with three lemmas to show that $I(t, U, \chi)$ may be bounded in terms of a sum of sums similar to the sum S of Lemmas 9.3 and 9.4, which is the case $F(X) = -\log X$ of the sum $S(H, M)$ in (1.6)–(1.9). The proof of Theorem 1 follows after the lemmas.

LEMMA 10.1. *Suppose $\ell(r, t) = r[\sqrt{t/2\pi r}] > 0$. Then*

$$L(1/2 + it, \chi) = \sum_{n=1}^{\ell(r,t)} \chi(n)n^{-1/2-it} + \theta(t, \chi) \sum_{n=1}^{\ell(r,t)} \bar{\chi}(n)n^{-1/2+it} + O(t^{-1/4}r^{3/4} \log(2r)),$$

where, with $a = (1 - \chi(-1))/2$,

$$\theta(t, \chi) = i^{-a} \frac{\tau(\chi)}{\sqrt{r}} \left(\frac{\pi}{r}\right)^{it} \frac{\Gamma((a + 1/2 - it)/2)}{\Gamma((a + 1/2 + it)/2)}.$$

Proof. This is a corollary of the more precise approximate functional equation worked out in [22] and [23] (some errors in the former paper are corrected in the latter paper).

LEMMA 10.2. *Let $t \geq 4\pi r$ and $0 < U \leq t/2$. Put $L(r, t) = \sqrt{rt/2\pi}$. Then*

$$I(t, U, \chi) \ll J(t, U, \chi) + t^{-1/2}r^{3/2} \log^2(2r) + U^2t^{-3/2}r^{1/2},$$

where

$$J(t, U, \chi) = \frac{1}{2U} \int_{t-U}^{t+U} \left| \sum_{n \leq L(r,t)} \chi(n)n^{-1/2-i\tau} \right|^2 d\tau.$$

Proof. This follows from (1.1) and Lemma 10.1, after sums with n running between $L(r, t)$ and $\ell(r, \tau)$ are estimated trivially.

LEMMA 10.3. *Let $t, r, U, L(r, t)$ and $J(t, U, \chi)$ be as in Lemma 10.2. Then there exists $\tau \in [t/2, 3t/2]$ such that*

$$J(t, U, \chi) \ll \sum_{\substack{m \leq L(r,t) \\ (m,r)=1}} \frac{1}{m} + \sum_{i=0}^{\infty} |S_i(\tau)| + \frac{L(r, t)}{U^2} + \frac{L(r, t)}{t},$$

where

$$S_i(\tau) = \sum_{K_{i+1} < k \leq K_i} \frac{1}{k} \sum_{1 \leq d \leq D_i} \chi\left(\frac{k+d}{k-d}\right) \left(\frac{k+d}{k-d}\right)^{-i\tau},$$

$K_i = 2^{-i}L(r, t)$ and $D_i = 2^{-1}(e^{1/U} - 1)K_i$.

Proof. Note first that

$$J(t, U, \chi) \ll \frac{1}{U} \int_{-\infty}^{\infty} \sum_{a=0}^1 \left| \sum_{\substack{n \leq L(r,t) \\ 2|(n-a)}} \chi(n)n^{-1/2-i(\eta+t)} \right|^2 \operatorname{sinc}^2\left(\frac{\eta}{2\pi U}\right) d\eta.$$

Multiplying out the squared absolute values and integrating term-by-term, we conclude that $J(t, U, \chi)$ is

$$\ll \sum_{\substack{m \leq L(r,t) \\ (m,r)=1}} \frac{1}{m} + 2\Re\left(\sum_{\substack{1 \leq n < m \leq L(r,t) \\ n \equiv m \pmod{2}}} \frac{\chi(m)\bar{\chi}(n)}{\sqrt{mn}} \left(\frac{m}{n}\right)^{-it} \Lambda\left(U \log \frac{m}{n}\right) \right).$$

By substituting $m = k + d, n = k - d$, we rewrite the double sum here as

$$\sum_{d \geq 1} \sum_{d < k \leq L(r,t)-d} \frac{\chi(k+d)\bar{\chi}(k-d)}{\sqrt{k^2-d^2}} \left(\frac{k+d}{k-d}\right)^{-it} \Lambda\left(U \log \frac{k+d}{k-d}\right).$$

For $U \gg 1$, it is implicit here that

$$d < 2^{-1}(e^{1/U} - 1)(k - d) \ll k/U \ll L(r, t)/U.$$

As $U \geq 1/\log 2$ may be assumed,

$$(k^2 - d^2)^{-1/2} = k^{-1}(1 + O(U^{-2})),$$

and the last sum above is

$$\begin{aligned} & \sum_{k \leq L(r,t)} \frac{1}{k} \sum_{1 \leq d < k} \chi\left(\frac{k+d}{k-d}\right) \left(\frac{k+d}{k-d}\right)^{-it} \Lambda\left(U \log \frac{k+d}{k-d}\right) + O(E) \\ &= \sum_{i=0}^{\infty} \sum_{K_{i+1} < k \leq K_i} \frac{1}{k} \sum_{d=1}^{[D_i]} \chi\left(\frac{k+d}{k-d}\right) \left(\frac{k+d}{k-d}\right)^{-it} \Lambda\left(U \log \frac{k+d}{k-d}\right) + O(E), \end{aligned}$$

where $E = L(r, t)/U^2$. The proof is completed by first writing here

$$A\left(U \log \frac{k+d}{k-d}\right) = \frac{1}{2\pi U} \int_{-\infty}^{\infty} \left(\frac{k+d}{k-d}\right)^{-i\eta} \operatorname{sinc}^2\left(\frac{\eta}{2\pi U}\right) d\eta,$$

then bringing the summations over i, k and d inside the integration over η , estimating the parts of the integral where $|\eta| \geq t/2$ trivially and (finally) using

$$\int_{-t/2}^{t/2} \left| \sum_{i=0}^{\infty} S_i(\eta+t) \right| \operatorname{sinc}^2\left(\frac{\eta}{2\pi U}\right) d\eta \leq 2\pi U \max_{t/2 \leq \tau \leq 3t/2} \left| \sum_{i=0}^{\infty} S_i(\tau) \right|.$$

The sum $S_i(\tau)$ from Lemma 10.3 can be further simplified. Indeed, partial summation shows that

$$S_i(\tau) \ll \frac{1}{K_i} \left| \sum_{K'_i < k \leq K_i} \sum_{1 \leq d \leq D_i} \chi\left(\frac{k+d}{k-d}\right) \left(\frac{k+d}{k-d}\right)^{-i\tau} \right|$$

for some $K'_i \in [K_{i+1}, K_i]$. Hence,

$$(10.1) \quad S_i(\tau) \ll \frac{1}{K_i} \sum_{j=0}^{\infty} |S_{i,j}(\tau)|,$$

where

$$(10.2) \quad S_{i,j}(\tau) = \sum_{K'_i < k \leq K_i} \sum_{D_{i+j+1} < d \leq D_{i+j}} \chi\left(\frac{k+d}{k-d}\right) \left(\frac{k+d}{k-d}\right)^{-i\tau}.$$

Proof of Theorem 1. It suffices to establish the theorem for all $\varepsilon > 0$ that are sufficiently small (in absolute terms). With Lemma 9.4 in mind, we define ε_0 to be equal to $\varepsilon/8E'$. If $t < 4\pi r \exp(\varepsilon_0^{-2})$, then the hypothesis that $r \ll t^{1/14}$ implies $r \asymp 1$, and it follows, by (1.3) and (1.4), that $2 \leq t \ll 1$ and $U \asymp 1$. Under such conditions (1.5) is equivalent to the bound $I(t, U, \chi) \ll 1$, which is then a trivial consequence of the continuity of the Dirichlet L -functions on the critical line. It therefore only remains to discuss the cases where

$$(10.3) \quad t \geq 4\pi r \exp(\varepsilon_0^{-2}).$$

As it follows from the formula (1.1) that $I(t, U, \chi)$ is the arithmetic mean of $I(t+U/2, U/2, \chi)$ and $I(t-U/2, U/2, \chi)$, and as each $U_j(r, t')$ varies only by a bounded factor when t' runs over $[t/2, 3t/2]$, one can show by induction that what is left to prove of the theorem will follow if only it be established in all the cases where

$$U \asymp \max_{j=1, \dots, 5} U_j(r, t).$$

It therefore follows by (1.3), (1.4) and (10.3) that we may assume henceforth

$$(10.4) \quad 4 \leq (rt)^{1/4} \leq U \ll rt^{1/3}.$$

By (10.3) again, we may suppose that Lemmas 10.2 and 10.3 apply, showing that, for some $i \in \mathbb{N} \cup \{0\}$, one has

$$I(t, U, \chi) \ll (1 + |S_i(\tau)| + E(U; r, t)) \log(rt),$$

where τ and $S_i(\tau)$ are as in Lemma 10.3 and, after (10.4),

$$\begin{aligned} E(U; r, t) &\asymp (rt)^{1/2}U^{-2} + r^{3/2}t^{-1/2} \log(t) + r^{1/2}t^{-3/2}U^2 \\ &\ll 1 + r^{3/2}t^{-1/2} \log(t) + r^{5/2}t^{-5/6} \ll 1 \end{aligned}$$

for $r \ll t^{1/3}/\log(t)$. Therefore, by using (10.1) in the last bound for $I(t, U, \chi)$, we may conclude that the theorem will follow if it can be shown that

$$(10.5) \quad \sum_{j=0}^{\infty} |S_{i,j}(\tau)| \ll K_i \quad (i = 0, 1, \dots),$$

where $S_{i,j}(\tau)$ is as in (10.2).

Referring back to Lemmas 10.2 and 10.3, to the paragraph of (10.1)–(10.2), and to (10.4), we find that the sum $S_{i,j}(\tau)$ either has no terms (and so equals zero), or $S_{i,j}(\tau) = S$ where S is as in (9.11) with

$$(10.6) \quad T = \tau/2\pi \in [t/4\pi, 3t/4\pi],$$

$$(10.7) \quad M = K_i = 2^{-i} \sqrt{rt/2\pi} \leq 2^{-i} \sqrt{2rT},$$

$$(10.8) \quad H = D_{i+j} = 2^{-1}(e^{1/U} - 1)K_{i+j} \leq 2^{-j}M/U,$$

$$(10.9) \quad M_1 = K'_1 \in [K_{i+1}, K_i] = [M/2, M], \quad H_1 = D_{i+j+1} = H/2.$$

As $i, j \geq 0$ in (10.5), it follows by (10.8) and (10.4) that the conclusion (9.12) of Lemma 9.3 applies to $S = S_{i,j}(\tau)$, and shows

$$(10.10) \quad \begin{aligned} S_{i,j}(\tau)/K_i &\ll ((rT)^{1/2}H/M + r^{1/2}T^{-1/2}M)(H/M)^{1/2} \\ &\ll (rT)^{1/6}(H/K_i)^{1/2}, \end{aligned}$$

provided that $H = D_{i+j} \ll M/(rT)^{1/3}$, $M = K_i \ll (rT)^{1/2}$ and $r \ll T^{1/5}$. As Theorem 1 assumes $r \ll t^{1/14}$, while (10.7) shows that $M = K_i \leq \sqrt{2rT}$ for $i \geq 0$, so it follows from (10.6) and (10.10) that

$$\sum_{\substack{j \geq 0 \\ D_{i+j} < K_i/(rT)^{1/3}}} |S_{i,j}(\tau)| \ll K_i \sum_{k=0}^{\infty} 2^{-k/2} \ll K_i \quad (i = 0, 1, \dots).$$

Therefore (10.5) (and the theorem) will follow once we show

$$(10.11) \quad \sum_{\substack{j \geq 0 \\ D_{i+j} \geq K_i / (rT)^{1/3}}} |S_{i,j}(\tau)| \ll K_i \quad (i = 0, 1, \dots).$$

Reverting to the notation introduced with (10.6)–(10.9), observe that the conditions of summation and constraints upon the index i in (10.11) together imply the condition (9.24) of Lemma 9.4. As that condition cannot possibly hold unless $U \leq (rT)^{1/3}$, it follows by (1.3) and case $j = 5$ of (1.4) that the bound (10.11) is trivial (by virtue of the sum there being empty) unless

$$r^{47/57} t^{17/57+\varepsilon} \leq r^{1/3} T^{1/3}.$$

From this and (10.6) we may conclude that (10.11) only requires further consideration in those cases where

$$(10.12) \quad r \leq t^{1/14-2\varepsilon}.$$

As (10.3) holds, with $\varepsilon_0 = \varepsilon/8E'$ and $r \geq 1$, we have here

$$t \geq 4\pi \exp(\varepsilon_0^{-2})$$

and

$$t^{2\varepsilon} \geq \exp(2\varepsilon\varepsilon_0^{-2}) \geq \exp(16E'\varepsilon_0^{-1}) > 2\varepsilon_0^{-3},$$

so it follows, by (10.6), that (10.3) and (10.12) imply the two conditions, $T \geq \exp(\varepsilon_0^{-2})$ and $r \leq \varepsilon_0^3 T^{1/14}$, appearing in Lemma 9.4. Similarly (10.6), (9.23) and (1.4) show that (9.22) is implied by (1.3). As (9.24) was earlier remarked to hold, we have now shown that either (10.11) is trivial, or the sums $S = S_{i,j}(\tau)$ there involved are such that all conditions of Lemma 9.4 hold. That lemma therefore implies, for $\lambda = \lambda_0$,

$$\sum_{\substack{j \geq 0 \\ D_{i+j} \geq K_i / (rT)^{1/3}}} |S_{i,j}(\tau)| \ll K_i \sum_{j=0}^{\infty} 2^{-j\lambda} \quad (i = 0, 1, \dots)$$

(see (10.8), (10.7)). As λ_0 is positive and absolute this establishes (10.11), completing our proof.

References

[1] E. Bombieri and H. Iwaniec, *On the order of $\zeta(1/2 + it)$* , Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 13 (1986), 449–472.
 [2] —, —, *Some mean-value theorems for exponential sums*, ibid. (4) 13 (1986), 473–486.

- [3] D. A. Burgess, *On character sums and L-series II*, Proc. London Math. Soc. (3) 13 (1963), 524–536.
- [4] J. B. Conrey and H. Iwaniec, *The cubic moment of central values of automorphic L-functions*, Ann. of Math. 151 (2000), 1175–1216.
- [5] I. Sh. Dzhabbarov, *Mean values of Dirichlet L-functions on short intervals of the critical line and their applications*, Izv. Akad. Nauk Azerbaïdzan. SSR Ser. Fiz-Tekhn. Mat. Nauk 9 (1988), no. 1, 3–9.
- [6] D. R. Heath-Brown, *The twelfth power moment of the Riemann zeta function*, Quart. J. Math. Oxford (2) 29 (1978), 443–462.
- [7] —, *Hybrid bounds for Dirichlet L-functions*, Invent. Math. 47 (1978), 149–170.
- [8] —, *Hybrid bounds for Dirichlet L-functions II*, Quart. J. Math. Oxford Ser. (2) 31 (1980), 157–167.
- [9] D. R. Heath-Brown and M. N. Huxley, *Exponential sums with a difference*, Proc. London Math. Soc. (3) 61 (1990), 227–250.
- [10] M. N. Huxley, *Exponential sums and lattice points*, ibid. 60 (1990), 471–502.
- [11] —, *Exponential sums and lattice points II*, ibid. 66 (1993), 279–301.
- [12] —, *Exponential sums and the Riemann zeta function IV*, ibid. 66 (1993), 1–40.
- [13] —, *Exponential sums and lattice points III*, ibid. 87 (2003), 591–609.
- [14] —, *Area, Lattice Points, and Exponential Sums*, London Math. Soc. Monogr. (N.S.) 13, Clarendon Press, Oxford, 1996.
- [15] —, *Exponential sums and the Riemann zeta function V*, preprint.
- [16] —, *A note on exponential sums with a difference*, Bull. London Math. Soc. 26 (1994), 325–327.
- [17] —, *Integer points, exponential sums and the Riemann zeta function*, in: Number Theory for the Millennium, II (Urbana, IL, 2000), Vol. 2, B. Berndt *et al.* (eds.), A K Peters, Natick, MA, 2002, 275–290.
- [18] M. N. Huxley and N. Watt, *Exponential sums and the Riemann zeta function*, Proc. London Math. Soc. (3) 57 (1988), 1–24.
- [19] —, —, *The number of ideals in a quadratic field, II*, Israel J. Math. 120 (2000), part A, 125–153.
- [20] —, —, *Hybrid bounds for Dirichlet’s L-function*, Math. Proc. Cambridge Philos. Soc. 129 (2000), 385–415.
- [21] H. Iwaniec and C. J. Mozzochi, *On the divisor and circle problems*, J. Number Theory 29 (1988), 60–93.
- [22] K. Matsumoto, *The mean square of Dirichlet L-functions*, Proc. Japan Acad. Ser. A Math. Sci. 58 (1982), 443–446.
- [23] —, *Corrections to: “The mean square of Dirichlet L-functions”*, ibid. 65 (1989), 344.
- [24] H. L. Montgomery and R. C. Vaughan, *Hilbert’s inequality*, J. London Math. Soc. (2) 8 (1974), 73–82.
- [25] Y. Motohashi, *A note on the mean value of the zeta and L-functions II*, Proc. Japan Acad. Ser. A Math. Sci. 61 (1985), 313–316.
- [26] —, *On the mean-square of L-functions*, unpublished notes, 1986.
- [27] V. V. Rane, *On an approximate functional equation for Dirichlet L-series*, Math. Ann. 264 (1983), 137–145.
- [28] W. M. Schmidt, *Equations over Finite Fields. An Elementary Approach*, Lecture Notes in Math. 536, Springer, Berlin, 1976.

- [29] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, Clarendon Press, Oxford, 1986.
- [30] R. C. Vaughan, *On the distribution of αp modulo 1*, *Mathematika* 24 (1977), 135–141.
- [31] N. Watt, *A problem on semicubical powers*, *Acta Arith.* 52 (1989), 119–140.
- [32] —, *A problem on square roots of integers*, *Period. Math. Hungar.* 21 (1990), 55–64.
- [33] —, *On differences of semicubical powers*, *Monatsh. Math.*, to appear.

Math. Program
P.O. Box 17172
U.G.R.U.
U.A.E. University
Al Ain
United Arab Emirates

Current address:
45 Charles Way
Limekilns
Fife
Scotland KY11 3LH
Great Britain
E-mail: elgin_awtt@hotmail.com

Received on 5.2.2002
and in revised form on 6.11.2002

(4213)