Integers Represented as a Sum of Primes and Powers of Two

D.R. Heath-Brown and J.-C. Puchta
Mathematical Institute, Oxford

1 Introduction

It was shown by Linnik [10] that there is an absolute constant K such that every sufficiently large even integer can be written as a sum of two primes and at most K powers of two. This is a remarkably strong approximation to the Goldbach Conjecture. It gives us a very explicit set $\mathcal{K}(x)$ of integers $n \leq x$ of cardinality only $O((\log x)^K)$, such that every sufficiently large even integer $N \leq x$ can be written as $N = p+p'+n$, with p, p' prime and $n \in \mathcal{K}(x)$. In contrast, if one tries to arrange such a representation using an interval in place of the set $\mathcal{K}(x)$, all known results would require $\mathcal{K}(x)$ to have cardinality at least a positive power of x.

Linnik did not establish an explicit value for the number K of powers of 2 that would be necessary in his result. However, such a value has been computed by Liu, Liu and Wang [12], who found that $K = 54000$ is acceptable. This result was subsequently improved, firstly by Li [8] who obtained $K = 25000$, then by Wang [18], who found that $K = 2250$ is acceptable, and finally by Li [9] who gave the value $K = 1906$. One can do better if one assumes the Generalized Riemann Hypothesis, and Liu, Liu and Wang [13] showed that $K = 200$ is then admissible.

The object of this paper is to give a rather different approach to this problem, which leads to dramatically improved bounds on the number of powers of 2 that are required for Linnik's theorem.

Theorem 1 *Every sufficiently large even integer is a sum of two primes and exactly 24 powers of 2.*

Theorem 2 *Assuming the Generalized Riemann Hypothesis, every sufficiently large even integer is a sum of two primes and exactly 9 powers of 2.*

In fact we shall sketch an argument, in the final section, which improves Theorem 1 so as to allow 21 powers of 2. Indeed it seems likely that further improvements are possible.

Previous workers have based their line of attack on a proof of Linnik's theorem due to Gallagher [3]. Let ϖ be a small positive constant. Set

$$S(\alpha) = \sum_{\varpi N < p \leq N} e(\alpha p), \tag{1}$$

where $e(x) := \exp(2\pi i x)$, and

$$T(\alpha) = \sum_{1\le\nu\le L} e(\alpha 2^\nu), \quad L = [\frac{\log N/2K}{\log 2}].$$

Earlier proofs of Linnik's Theorem have used estimates for $\mathrm{meas}(\mathcal{A}_\lambda)$, where

$$\mathcal{A}_\lambda = \{\alpha \in [0,1] : |T(\alpha)| \ge \lambda L\}.$$

In contrast we shall investigate a high power moment

$$I(q) = \int_0^1 |T(\alpha)|^{2q} d\alpha.$$

We write $r(n,q,N)$ for the number of representations of an integer n as a sum of q terms $2^\nu \le N/2K$. Thus

$$T(\alpha)^q = \sum_n r(n,q,N) e(\alpha n),$$

and

$$I(q) = \sum_n r(n,q,N)^2.$$

We set

$$r(q) = \max_{n,N} \ r(n,q,N)$$

and note that

$$\sum_n r(n,q,N) = T(0)^q = L^q,$$

whence

$$I(q) \le r(q) L^q. \tag{2}$$

In §§7 and 8 we shall bound $r(q)$ by a combinatorial argument, and prove the following estimate.

Lemma 1 *We have*

$$r(q) \le (1.753)^q q!$$

for all $q \in \mathbb{N}$.

We shall use the resulting inequality for $I(q)$ directly, without passing via an estimate for $\mathrm{meas}(\mathcal{A}_\lambda)$. However it may be of interest to see what bound Lemma 1 implies. We have

$$\begin{aligned}\mathrm{meas}(\mathcal{A}_\lambda) &\le (\lambda L)^{-2q} I(q)\\ &\le (\lambda L)^{-2q} (1.753)^q L^q q!\\ &\ll (\frac{1.753q}{eL\lambda^2})^q \sqrt{q}.\end{aligned}$$

We minimize this by taking $q = [L\lambda^2/1.753]$, whence

$$\begin{aligned}\mathrm{meas}(\mathcal{A}_\lambda) &\ll \exp(-L\lambda^2/1.753)\sqrt{L}\\ &\ll N^{-\lambda^2/(1.753\log 2)}\sqrt{L}\\ &\ll N^{-0.822\lambda^2}\sqrt{\log N}.\end{aligned}$$

It is easy to see that one cannot have a bound of the form $O(N^{-c\lambda^2})$ with $c > 1$, so it is natural to ask whether

$$\mathrm{meas}(\mathcal{A}_\lambda) \ll_\varepsilon N^{-\lambda^2+\varepsilon}$$

for every fixed $\varepsilon > 0$. We expect that the constant 1.753 in Lemma 1 is essentially optimal. However there is a potential loss, in our application, in deriving (2). This arises because we take q to be a constant multiple of $\log N$ in §6, whereas $r(q)$ might be determined by values of N larger than this implies.

The best bound for $\mathrm{meas}(\mathcal{A}_\lambda)$ in the literature is due to Liu, Liu and Wang [11; Lemma 3], and states that $\mathrm{meas}(\mathcal{A}_{1-\eta}) \ll N^{-\theta}(\log N)^{5/2}$ for $\eta < (7e)^{-1}$, with

$$\theta = 1 - F(\frac{2+\sqrt{2}}{4}\eta) - F(1 - \frac{2+\sqrt{2}}{4}\eta)$$

and $F(x) = x(\log x)/(\log 2)$. This is distinctly less elegant than our bound and holds for a shorter range. None the less, for small values of η it is stronger than is achieved by our method. It should be stressed however, that in the present paper the critical size for $|S(\alpha)|$ corresponds to a value of η considerably larger than $(7\eta)^{-1}$.

The estimate provided by Lemma 1 will be injected into the circle method, where it will be crucial in bounding the minor arc contribution. On the major arcs we shall improve on Gallagher's analysis so as to show that hypothetical zeros close to $\sigma = 1$ play no rôle. Thus, in contrast to previous workers, we will have no need for explicit numerical zero-free regions for L-functions. Naturally this produces a considerable simplification in the computational aspects of our work. Thus it is almost entirely the value of the constant 1.753 in Lemma 1 which determines the number of powers of 2 appearing in Theorems 1 and 2.

The paper naturally divides into two parts, one of which is analytic, involving the circle method and zeros of L-functions, and the other of which is combinatorial, devoted to the proof of Lemma 1. We begin with the former.

One remark about notation is in order. At various stages in the proof, numerical upper bounds on ϖ will be required. Since we shall always take ϖ to be sufficiently small, we shall assume that any such bound is satisfied. Moreover, since ϖ is to be thought of as fixed, we will allow the implied constants in the $O(\ldots)$ and $\ll$ notations to depend on ϖ.

2 The Major Arcs

We shall follow the method of Gallagher [3; §1] closely. We choose a parameter P in the range $1 \le P \le N^{2/5}$ and define the major arcs $\mathfrak{M}$ as the set of $\alpha \in [0,1]$ for which there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ such that $q \le P$ and

$$|\alpha - \frac{a}{q}| \le \frac{P}{qN}.$$

If χ is a character to modulus q, we write

$$c_n(\chi) = \sum_{a=1}^{q} \chi(a) e(\frac{an}{q})$$

and

$$\tau(\chi)=\sum_{a=1}^{q}\chi(a)e(\frac{a}{q}).$$

Moreover we put

$$A(\chi,\beta)=\sum_{\varpi N<p\leq N}\chi(p)e(\beta p)$$

and

$$I_{n,s}(\chi,\chi')=\int_{-P/sN}^{P/sN}A(\chi,\beta)A(\chi',\beta)e(-\beta n)d\beta.$$

If χ is a character to a modulus $r|q$ we also write χ_q for the induced character modulo q, and if χ,χ' are characters to moduli r and r' respectively, we set

$$J_n(\chi,\chi')=\sum_{\substack{q\leq P\\ [r,r']|q}}\frac{1}{\phi(q)^2}c_n(\chi_q\chi'_q)\tau(\overline{\chi_q})\tau(\overline{\chi'_q})I_{n,q}(\chi,\chi').$$

Then, by a trivial variant of the argument leading to Gallagher [3; (3)], we find that

$$\int_{\mathfrak{M}}S(\alpha)^2e(-\alpha n)d\alpha=\sum_{\chi,\chi'}J_n(\chi,\chi')+O(P^{5/2}),\tag{3}$$

for any integer n, the sum being over primitive characters χ,χ' to moduli r,r' for which $[r,r']\leq P$. In what follows we shall take $1\leq n\leq N$.

To estimate the contribution from a particular pair of characters χ,χ' we put

$$A_q(\chi)=\{\int_{-P/qN}^{P/qN}|A(\chi,\beta)|^2d\beta\}^{1/2}$$

and

$$C_n(\chi,\chi')=\sum_{\substack{q\leq P\\ [r,r']|q}}\frac{1}{\phi(q)^2}|c_n(\chi_q\chi'_q)\tau(\overline{\chi_q})\tau(\overline{\chi'_q})|.$$

Note that what Gallagher calls $||A(\chi)||$ is our $A_1(\chi)$. We have $A_q(\chi)\leq A_m(\chi)$ whenever $m\leq q$. Then, as in Gallagher [3; (4)] we find

$$|J_n(\chi,\chi')|\leq C_n(\chi,\chi')A_{[r,r']}(\chi)A_{[r,r']}(\chi').\tag{4}$$

It is in bounding $C_n(\chi,\chi')$ that there is a loss in Gallagher's argument. Let r'' be the conductor of $\chi\chi'$, and write $m=[r,r']$. moreover, for any positive integers a and n we wite

$$a_n=\frac{a}{(a,n)}.$$

Then Gallagher shows that

$$C_n(\chi,\chi')\leq(rr'r'')^{1/2}\sum_{q\leq P,\,m|q}(\phi(q)\phi(q_n))^{-1},$$

where q/m is square-free and coprime to m. Moreover we have $r''|m_n$. It follows that

$$C_n(\chi,\chi') \le \frac{(rr'r'')^{1/2}}{\phi(m)\phi(m_n)} \sum_{(s,m)=1} \mu^2(s)/\phi(s)\phi(s_n).$$

The sum on the right is

$$\prod_{p\nmid mn}(1+\frac{1}{(p-1)^2})\prod_{p|n,p\nmid m}(1+\frac{1}{(p-1)}) \ll \prod_{p|n,p\nmid m}\frac{p}{(p-1)},$$

and

$$\frac{m}{\phi(m)}\prod_{p|n,p\nmid m}\frac{p}{(p-1)} \le \frac{n}{\phi(n)}\frac{m_n}{\phi(m_n)}.$$

We therefore deduce that

$$C_n(\chi,\chi') \ll \frac{(rr'r'')^{1/2}}{m}\frac{m_n}{\phi^2(m_n)}\frac{n}{\phi(n)}.$$

Now if $p^e||r$ and $p^f||r'$, then $p^{|e-f|}|r''$, since r'' is the conductor of $\chi\chi'$. (Here the notation $p^e||r$ means, as usual, that $p^e|r$ and $p^{e+1}\nmid r$.) We therefore set

$$h=(r,r') \quad \text{and} \quad r=hs,\ r'=hs', \tag{5}$$

so that $ss'|r''$ and $m=hss'$. Since

$$\frac{m_n}{\phi^2(m_n)} \ll m_n^{\varpi-1}$$

we therefore have

$$\frac{(rr'r'')^{1/2}}{m}\frac{m_n}{\phi^2(m_n)} \ll (ss')^{-1/2}r''^{1/2}m_n^{\varpi-1}.$$

Now, using the bounds $r''\le m_n$ and $ss'\le r''$, we find that

$$\begin{aligned}\frac{(rr'r'')^{1/2}}{m}\frac{m_n}{\phi^2(m_n)} &\ll (ss')^{-1/2}r''^{1/2}r''^{\varpi-1}\\ &= (ss')^{-1/2}r''^{\varpi-1/2}\\ &\ll (ss')^{\varpi-1}.\end{aligned}$$

Alternatively, using only the fact that $m_n\ge r''$, we have

$$\begin{aligned}\frac{(rr'r'')^{1/2}}{m}\frac{m_n}{\phi^2(m_n)} &\ll (ss')^{-1/2}m_n^{1/2}m_n^{\varpi-1}\\ &\ll m_n^{\varpi-1/2}.\end{aligned}$$

These estimates produce

$$C_n(\chi,\chi') \ll \min\{(ss')^{\varpi-1}\,,\,m_n^{\varpi-1/2}\}\frac{n}{\phi(n)}.$$

On combining this with the bounds (3) and (4) we deduce the following result.

Lemma 2 *Suppose that $P \le N^{2/5-\varpi}$. Then*

$$\int_{\mathfrak{M}} S(\alpha)^2 e(-\alpha n)d\alpha = J_n(1,1) + O(\frac{n}{\phi(n)}S_n) + O(N^{1-\varpi}),$$

where

$$S_n = \sum_{\chi,\chi'} A_{[r,r']}(\chi)A_{[r,r']}(\chi')\min\{(ss')^{\varpi-1}\,,\,m_n^{-1/3}\},$$

the sum being over primitive characters, not both principal, of moduli r, r', with $[r, r'] \le P$.

We have next to consider $A_m(\chi)$. According to the argument of Montgomery and Vaughan [15; §7] we have

$$A_m(\chi) \ll N^{1/2} \max_{\varpi N<x\le N}\max_{0<h\le x}(h+mN/P)^{-1}|\sum_x^{x+h}\chi(p)|.$$

Note that we have firstly taken account of the restriction in (1) to primes $p > \varpi N$, and secondly replaced $(h + N/P)^{-1}$ as it occurs in Montgomery and Vaughan, by the smaller quantity $(h+mN/P)^{-1}$. The argument of [15; §7] clearly allows this.

By partial summation we have

$$\sum_x^{x+h}\chi(p) \ll (\log x)^{-1}\max_{0<j\le h}\sum_x^{x+j}\chi(p)\log p.$$

Moreover, a standard application of the 'explicit formula' for $\psi(x, \chi)$ produces the estimate

$$\sum_x^{x+j}\chi(p)\log p \ll N^{1/2+3\varpi}(\log N)^2 + \sum_\rho |\frac{(x+j)^\rho}{\rho} - \frac{x^\rho}{\rho}|,$$

where the sum over ρ is for zeros of $L(s, \chi)$ in the region

$$\beta \ge \frac{1}{2} + 3\varpi, \quad |\gamma| \le N.$$

When χ is the trivial character we shall include the pole $\rho = 1$ amongst the 'zeros'. Since $j \le h$ and

$$\frac{(x+j)^\rho}{\rho} - \frac{x^\rho}{\rho} \ll \min\{jN^{\beta-1}\,,\,N^\beta|\gamma|^{-1}\},$$

we find that

$$A_m(\chi) \ll \frac{P}{m}N^{4\varpi} + \frac{N^{1/2}}{\log N}\{\max_{0<h\le N}(h+mN/P)^{-1}\sum_\rho N^{\beta-1}\min\{h\,,\,N|\gamma|^{-1}\}.$$

However we have

$$\min\{\frac{h}{h+H}\,,\,\frac{A}{h+H}\} \le \min\{1\,,\,\frac{A}{H}\}$$

whenever $h, H, A > 0$. Applying this with $H = mN/P$ and $A = N|\gamma|^{-1}$, we deduce that

$$A_m(\chi) \ll \frac{P}{m}N^{4\varpi} + \frac{N^{1/2}}{\log N}\sum_\rho N^{\beta-1}\min\{1\,,\,Pm^{-1}|\gamma|^{-1}\}. \tag{6}$$

3 The Sum S_n

In order to investigate the sum S_n we decompose the available ranges for r, r' and the corresponding zeros ρ, ρ' into (overlapping) ranges

$$\left\{ \begin{array}{cc} R \le r \le RN^{\varpi}, & R' \le r' \le R'N^{\varpi}, \\ T-1 \le |\gamma| < TN^{\varpi}, & T'-1 \le |\gamma'| < T'N^{\varpi}. \end{array} \right. \tag{7}$$

Clearly $O(1)$ such ranges suffice to cover all possibilities, so it is enough to consider the contribution from a fixed range of the above type. Throughout this section we shall follow the convention that $\rho = 1$ is to included amongst the 'zeros' corresponding to the trivial character.

Let $N(\sigma, \chi, T)$ denote as usual, the number of zeros ρ of $L(s, \chi)$, in the region $\beta \ge \sigma$, $|\gamma| \le T$, and let $N(\sigma, r, T)$ be the sum of $N(\sigma, \chi, T)$ for all characters χ of conductor r. Since

$$N^{\beta-1} = N^{3\varpi-1/2} + \int_{1/2+3\varpi}^{\beta} N^{\sigma-1}(\log N)d\sigma$$

for $\beta \ge 1/2 + 3\varpi$, we find that

$$\sum_{\rho} N^{\beta-1} \ll N^{6\varpi-1/2}RT + I(r)\log N, \tag{8}$$

where the sum is over zeros of $L(s, \chi)$ for all χ of conductor r, subject to $T-1 \le |\gamma| \le TN^{\varpi}$, and were

$$I(r) = \int_{1/2+3\varpi}^{1} N^{\sigma-1}N(\sigma, r, TN^{\varpi})d\sigma.$$

We now insert (8) into (6) so that, for given r, r', the range (7) contributes to

$$\sum_{\chi \pmod r} A_m(\chi)$$

a total

$$\begin{aligned} &\ll \phi(r)\frac{P}{m}N^{4\varpi} + \frac{N^{1/2}}{\log N}m(R,T)N^{6\varpi-1/2}RT + N^{1/2}m(R,T)I(r) \\ &\ll PN^{6\varpi} + N^{1/2}m(R,T)I(r). \end{aligned} \tag{9}$$

Similarly, for the double sum

$$\sum_{\chi \pmod r}\sum_{\chi' \pmod{r'}} A_m(\chi)A_m(\chi')$$

the contribution is

$$\begin{aligned} \ll\ & P^2N^{12\varpi} + PN^{1/2+6\varpi}m(R,T)I(r) \\ & + PN^{1/2+6\varpi}m(R',T')I(r') + Nm(R,T)m(R',T')I(r)I(r'). \end{aligned} \tag{10}$$

We then sum over r, r' using the following lemma.

Lemma 3 *Let*

$$\max_{r\le R} N(\sigma,r,T)=N_1(R),\quad \max_{r'\le R'} N(\sigma',r',T')=N_1(R'),$$

and

$$\sum_{r\le R} N(\sigma,r,T)=N_2(R),\quad \sum_{r'\le R'} N(\sigma',r',T')=N_2(R').$$

In the notation of (5) we have

$$\sum_{r\le R}\sum_{r'\le R'} N(\sigma,r,T)N(\sigma',r',T')(ss')^{\varpi-1} \tag{11}$$
$$\ll \quad \{N_1(R)N_2(R)N_1(R')N_2(R')\}^{1/2+2\varpi},$$

for $1/2\le\sigma,\sigma'\le 1$.

Moreover, if

$$P\le N^{45/154-4\varpi},$$

then

$$\sum_{r,r'} m(R,T)m(R',T')N(\sigma,r,TN^{\varpi})N(\sigma',r',T'N^{\varpi})(ss')^{\varpi-1}, \tag{12}$$

$$\ll N^{(1-\varpi)(1-\sigma)+(1-\varpi)(1-\sigma')} \tag{13}$$

for $1/2+3\varpi\le\sigma,\sigma'\le 1$, *where the summation is for* $R\le r\le RN^{\varpi}$ *and* $R'\le r'\le R'N^{\varpi}$.

We shall prove this at the end of this section. Henceforth we shall assume that $P\le N^{45/154-4\varpi}$.

For suitable values of η in the range

$$0\le\eta\le\log\log N \tag{14}$$

we shall define $\mathcal{B}(\eta)$ to be the set of characters χ of conductor $r\le P$, for which the function $L(s,\chi)$ has at least one zero in the region

$$\beta>1-\frac{\eta}{\log N},\quad |\gamma|\le N.$$

According to our earlier convention the trivial character is always in $\mathcal{B}(\eta)$. Now, if we restrict attention to pairs χ,χ' for which $\chi\notin\mathcal{B}(\eta)$ we have

$$\begin{aligned}
\sum_{R\le r\le RN^{\varpi}}\ \sum_{R'\le r'\le R'N^{\varpi}} & Nm(R,T)m(R',T')I(r)I(r')(ss')^{\varpi-1}\\
&\ll \int_{1/2+3\varpi}^{1-\eta/\log N}\int_{1/2+3\varpi}^{1} N^{1-\varpi(1-\sigma)-\varpi(1-\sigma')}d\sigma' d\sigma\\
&\ll N^{1-\varpi\eta/\log N}(\log N)^{-2}\\
&= e^{-\varpi\eta}N(\log N)^{-2}.
\end{aligned}$$

Terms for which $\chi\in\mathcal{B}(\eta)$ but $\chi'\notin\mathcal{B}(\eta)$ may be handled similarly. This concludes our discussion of the final term in (10) for the time being.

To handle the third term in (10) we use the zero density estimate

$$\sum_{r\le R} N(\sigma,r,T) \ll (R^2T)^{\kappa(\sigma)(1-\sigma)}, \tag{15}$$

where

$$\kappa(\sigma) = \begin{cases} \frac{3}{2-\sigma} + \varpi, & \frac{1}{2} \le \sigma \le \frac{3}{4} \\ \frac{12}{5} + \varpi, & \frac{3}{4} \le \sigma \le 1. \end{cases} \tag{16}$$

This follows from results of Huxley [5], Jutila [7; Theorem 1] and Montgomery [14; Theorem 12.2]. For each fixed value of r' we have

$$\begin{aligned} \sum_r (ss')^{\varpi-1} &\le \sum_{h|r'} (r'/h)^{\varpi-1} \sum_{s\le P/h} s^{\varpi-1} \\ &\ll \sum_{h|r'} (r'/h)^{\varpi-1} (P/h)^{\varpi} \\ &\ll N^{\varpi}. \end{aligned}$$

The contribution of the third term in (10) to S_n is therefore

$$\ll PN^{1/2+5\varpi} m(R',T') \sum_{r'} I(r').$$

However the bound (15) shows that

$$m(R',T') \sum_{r'} N(\sigma,r',TN^{\varpi}) \ll \max\{1\,,\ \frac{P}{R'T'}\}(R'^2N^{2\varpi}T'N^{\varpi})^{\kappa(\sigma)(1-\sigma)}.$$

Since

$$0 \le \kappa(\sigma)(1-\sigma) \le 1$$

in the range $1/2+\varpi \le \sigma \le 1$, this is

$$\ll (P^2N^{3\varpi})^{\kappa(\sigma)(1-\sigma)}.$$

Moreover, if $P \le N^{45/154-4\varpi}$, then

$$(P^2N^{3\varpi})^{\kappa(\sigma)(1-\sigma)} N^{\sigma-1} \le N^{f(\sigma)}$$

with

$$\begin{aligned} f(\sigma) &= (\frac{45}{77}\kappa(\sigma)-1)(1-\sigma) \\ &\le (\frac{45}{77}\{\frac{12}{5}+\varpi\}-1)(1-\sigma) \\ &\le (\frac{31}{77}+\varpi)(1-\sigma) \\ &\le (\frac{31}{77}+\varpi)\frac{1}{2} \\ &\le \frac{31}{154}+\varpi. \end{aligned}$$

It follows that the contribution of the third term in (10) to S_n is

$$\ll PN^{1/2+6\varpi}.N^{31/154+\varpi} \ll N^{1-\varpi}.$$

The second term may of course be handled similarly.

Finally we deal with the first term of (10) which produces a contribution to S_n which is

$$\begin{aligned}
&\ll P^2 N^{12\varpi} \sum_{r,r'} (ss')^{\varpi-1} \\
&\ll P^2 N^{12\varpi} \sum_{ss'h\le P} (ss')^{\varpi-1} \\
&\ll P^2 N^{12\varpi} \sum_{ss'\le P} P(ss')^{\varpi-2} \\
&\ll P^3 N^{12\varpi} \\
&\ll N^{1-\varpi},
\end{aligned}$$

for $P \le N^{45/154-4\varpi}$.

We summarize our conclusions thus far as follows.

Lemma 4 *If $P \le N^{45/154-4\varpi}$ then*

$$S_n \le \sum_{\chi,\chi'\in\mathcal{B}(\eta)} A_m(\chi)A_m(\chi')m_n^{-1/3} + O(e^{-\varpi\eta}N(\log N)^{-2}).$$

To handle the characters in $\mathcal{B}(\eta)$ we use the zero-density estimate

$$N(\sigma,r,T) \ll (rT)^{\kappa(\sigma)(1-\sigma)}, \tag{17}$$

with $\kappa(\sigma)$ given by (16). This also follows from work of Huxley [5], Jutila [7; Theorem 1] and Montgomery [14; Theorem 12.1]. Thus

$$\begin{aligned}
m(R,T)N(\sigma,r,TN^{\varpi}) &\ll \max\{1\,,\,\frac{P}{RT}\}(rTN^{\varpi})^{\kappa(\sigma)(1-\sigma)} \\
&\ll (PN^{2\varpi})^{\kappa(\sigma)(1-\sigma)} \\
&\ll (PN^{2\varpi})^{(12/5+\varpi)(1-\sigma)} \\
&\ll N^{(1-\varpi)(1-\sigma)}
\end{aligned}$$

for $P \le N^{45/154-4\varpi}$. We deduce that

$$m(R,T)I(r) \ll (\log N)^{-1}.$$

It follows from (9) that

$$A_m(\chi) \ll N^{1/2}(\log N)^{-1}.$$

We also note that

$$\#\mathcal{B}(\eta) \ll \sum_r N(1-\frac{\eta}{\log N},r,N) \ll (P^2N)^{3\eta/\log N} \ll e^{6\eta},$$

by (15), since $\kappa(\sigma) \le 3$ for all σ. We therefore have the following facts.

Lemma 5 *If $\chi \in \mathcal{B}(\eta)$, we have $A_m(\chi) \ll N^{1/2}(\log N)^{-1}$. Moreover, we have $\#\mathcal{B}(\eta) \ll e^{6\eta}$.*

We end this section by establishing Lemma 3. We shall suppose, as we may by the symmetry, that

$$N_2(R)N_1(R') \le N_2(R')N_1(R). \tag{18}$$

Let $U \ge 1$ be a parameter whose value will be assigned in due course, see (19). For those terms of the sum (11) in which $ss' \ge U$ we plainly have a total

$$\le \sum_{r\le R}\sum_{r'\le R'} N(\sigma,r,T)N(\sigma',r',T')U^{\varpi-1} \ll N_2(R)N_2(R')U^{\varpi-1}.$$

On the other hand, when $ss' < U$ we observe that, for fixed s, s' we have

$$\begin{aligned}\sum_h N(\sigma,hs,T)N(\sigma',hs',T') &\ll \sum_h N(\sigma,hs,T)N_1(R')\\ &\ll \sum_r N(\sigma,r,T)N_1(R')\\ &\ll N_2(R)N_1(R').\end{aligned}$$

On summing over s and s' we therefore obtain a total

$$\ll N_2(R)N_1(R')\sum_{ss'\le U}(ss')^{\varpi-1} \ll N_2(R)N_1(R')U^{2\varpi}.$$

It follows that the sum (11) is

$$\ll N_2(R)\{N_2(R')U^{2\varpi-1} + N_1(R')U^{2\varpi}\}.$$

We therefore choose

$$U = N_2(R')/N_1(R'), \tag{19}$$

whence the sum (11) is

$$\begin{aligned}&\ll N_2(R)N_1(R')U^{2\varpi}\\ &\ll N_2(R)N_1(R')\{N_1(R)N_2(R)N_1(R')N_2(R')\}^{2\varpi}\\ &\ll \{N_2(R)N_1(R')N_2(R')N_1(R)\}^{1/2}\{N_1(R)N_2(R)N_1(R')N_2(R')\}^{2\varpi}\end{aligned}$$

in view of (18). This produces the required bound.

To establish (13) we shall bound $N_1(R)$ and $N_1(R')$ using (17). Moreover to handle $N_2(R)$ and $N_2(R')$ we shall use the estimate

$$\sum_{r\le R} N(\sigma,r,T) \ll \begin{cases} (R^2T)^{\kappa(\sigma)(1-\sigma)}, & \frac{1}{2}+\varpi \le \sigma \le \frac{23}{38}\\ (R^2T^{6/5})^{\lambda(1-\sigma)}, & \frac{23}{38} < \sigma \le 1,\end{cases}$$

where

$$\lambda = \frac{20}{9} + \varpi.$$

This follows from (15) along with Heath-Brown [4; Theorem 2] and Jutila [7; Theorem 1].

We now see that the sum (12) may be estimated as

$$\ll m(R,T)R^aT^c.m(R',T')R'^bT'^d.N^e, \tag{20}$$

say, where

$$a = \begin{cases} 3\kappa(\sigma)(1-\sigma)(\frac{1}{2}+2\varpi), & \frac{1}{2}+3\varpi \le \sigma \le \frac{23}{38} \\ \{\kappa(\sigma)+2\lambda\}(1-\sigma)(\frac{1}{2}+2\varpi), & \frac{23}{38} < \sigma \le 1, \end{cases}$$

and

$$c = \begin{cases} 2\kappa(\sigma)(1-\sigma)(\frac{1}{2}+2\varpi), & \frac{1}{2}+3\varpi \le \sigma \le \frac{23}{38} \\ \{\kappa(\sigma)+6\lambda/5\}(1-\sigma)(\frac{1}{2}+2\varpi), & \frac{23}{38} < \sigma \le 1, \end{cases}$$

and similarly for b and d. Moreover we may take

$$e = 6\varpi(1-\sigma) + 6\varpi(1-\sigma').$$

It therefore follows that $0 \le c, d < 1$, whence (20) is maximal for $T = P/R$ and $T' = P/R'$. Similarly we have $a \ge c$ and $b \ge d$. Thus, after substituting $T = P/R$ and $T' = P/R'$ in (20), the resulting expression is increasing with respect to R and R', and hence is maximal when $R = R' = P$. We therefore see that (20) is

$$\ll P^{a+b}N^e.$$

Finally one can check that

$$(\frac{45}{154} - 4\varpi)a \le (1-7\varpi)(1-\sigma),$$

and similarly for b. This suffices to establish the bound (13) for $P \le N^{45/154-4\varpi}$.

4 Summation Over Powers of 2

In this section we consider the major arc integral

$$\int_{\mathfrak{M}} S(\alpha)^2 T(\alpha)^K e(-\alpha N) d\alpha,$$

where we now assume N to be even. According to Lemmas 2 and 4 we have

$$\begin{aligned} \int_{\mathfrak{M}} S(\alpha)^2 T(\alpha)^K e(-\alpha N) d\alpha &= \Sigma_0 + O(e^{-\varpi\eta} N (\log N)^{-2} \Sigma_1) \\ &\quad + O(N(\log N)^{-2}\Sigma_2), \end{aligned} \tag{21}$$

where

$$\Sigma_0 = \sum_n J_n(1,1),$$

$$\Sigma_1 = \sum_n \frac{n}{\phi(n)}$$

and

$$\Sigma_2 = \sum_{\chi,\chi' \in \mathcal{B}(\eta)} \sum_n \frac{n}{\phi(n)} m_n^{-1/3}.$$

In each case the sum over n is for values

$$n = N - \sum_{j=1}^{K} 2^{\nu_j}. \tag{22}$$

We begin by considering the main term Σ_0. We put

$$T(\beta) = \sum_{\varpi N < m \le N} \frac{e(\beta m)}{\log m}$$

and

$$R(\beta) = S(\beta) - T(\beta).$$

We also set

$$||R|| = \int_{-P/N}^{P/N} |R(\beta)|^2 d\beta$$

and

$$J(n) = \sum_{\substack{\varpi < m_1, m_2 < N \\ m_1 + m_2 = n}} (\log m_1)^{-1} (\log m_2)^{-1}.$$

Then, as in Gallagher [3; (11)], we have

$$\begin{aligned} J_n(1,1) &= J(n)\mathcal{S}(n) + O(N(\log N)^{-2} \frac{n}{\phi(n)} d(n) \frac{\log P}{P}) \\ &\quad + O(\frac{n}{\phi(n)} \{N^{1/2} (\log N)^{-1} ||R|| + ||R||^2\}), \end{aligned} \tag{23}$$

where

$$\mathcal{S}(n) = \prod_{p|n} (\frac{p}{p-1}) \prod_{p \nmid n} (1 - \frac{1}{(p-1)^2}).$$

In analogy to (6) we have

$$||R|| \ll PN^{4\varpi} + \frac{N^{1/2}}{\log N} \sum_{\rho} N^{\beta - 1} \min\{1\,,\, P|\gamma|^{-1}\},$$

where the sum over ρ is for zeros of $\zeta(s)$ in the region

$$\beta \ge \frac{1}{2} + 3\varpi, \quad |\gamma| \le N.$$

We split the range for $|\gamma|$ into $O(1)$ overlapping intervals

$$T - 1 \le |\gamma| \le TN^{\varpi},$$

and find, as in (8) that each range contributes

$$\ll PN^{4\varpi} + N^{1/2} \min\{1\,,\, \frac{P}{T}\}\{N^{6\varpi - 1/2} T + \int_{1/2+3\varpi}^{1} N^{\sigma - 1} N(\sigma, 1, TN^{\varpi}) d\sigma\}$$

to $||R||$. Using the case $R = 1$ of (15), together with Vinogradov's zero-free region

$$\sigma \ge 1 - \frac{c_0}{(\log T)^{3/4} (\log\log T)^{3/4}}$$

(see Titchmarsh [16; (6.15.1)]), we find that this gives

$$||R|| \ll N^{1/2}(\log N)^{-10},$$

say, for $P \le N^{45/154-4\varpi}$. The error terms in (21) are therefore $O(N(\log N)^{-9})$.

We also note that

$$\begin{aligned} J(n) &= (\log N)^{-2}\#\{m_1, m_2 : \varpi N < m_1, m_2 \le N,\, m_1+m_2=n\} \\ &\qquad +O(N(\log N)^{-3}) \\ &= (\log N)^{-2}R(n) + O(N(\log N)^{-3}), \end{aligned}$$

where

$$R(n) = \begin{cases} 2N-n, & (1+\varpi)N \le n \le 2N, \\ n-2\varpi N, & 2\varpi N \le n \le (1+\varpi)N, \\ 0, & \text{otherwise.} \end{cases}$$

In particular, we have $R(N-m) = (1-2\varpi)N(\log N)^{-2} + O(m(\log N)^{-2})$ for $1 \le m \le N$. Since

$$\mathcal{S}(n) \ll \frac{n}{\phi(n)} \ll \log\log N,$$

we find, on taking n of the form (22), that

$$\sum_n J(n)\mathcal{S}(n) = (1-2\varpi)N(\log N)^{-2}\sum_n \mathcal{S}(n) + O(N(\log N)^{K-5/2})$$

for $K \ge 2$, whence

$$\Sigma_0 = (1-2\varpi)N(\log N)^{-2}\sum_n \mathcal{S}(n) + O(N(\log N)^{K-5/2}).$$

Since the numbers n are all even, we have

$$\mathcal{S}(n) = 2C_0 \prod_{p|n,p\ne 2}\frac{p-1}{p-2} = 2C_0\sum_{d|n}k(d),$$

where

$$C_0 = \prod_{p\ne 2}(1-\frac{1}{(p-1)^2}) \tag{24}$$

and $k(d)$ is the multiplicative function defined by taking

$$k(p^e) = \begin{cases} 0, & p=2 \text{ or } e \ge 2, \\ (p-2)^{-1}, & \text{otherwise.} \end{cases} \tag{25}$$

For any odd integer d we shall define $\varepsilon(d)$ to be the order of 2 in the multiplicative group modulo d, and we shall set

$$H(d;N,K) = \#\{(\nu_1,\ldots,\nu_K) : 1 \le \nu_i \le \varepsilon(d),\, d|N-\sum 2^{\nu_i}\}.$$

Then for any fixed D we have

$$\begin{aligned}\sum_n \mathcal{S}(n) &= 2C_0\sum_d k(d)\#\{n : d|n\}\\ &\geq 2C_0\sum_{d\leq D} k(d)\#\{n : d|n\}\\ &\geq 2C_0\sum_{d\leq D} k(d)H(d;N,K)[L/\varepsilon(d)]^K\\ &\geq \{1+O((\log N)^{-1})\}2C_0L^K\sum_{d\leq D} k(d)H(d;N,K)\varepsilon(d)^{-K}.\end{aligned}$$

We shall take $D=5$. We trivially have $\varepsilon(1)=1$ and $H(1;N,K)=1$ for all N and K. When $d=3$ or $d=5$ the powers of 2 run over all non-zero residues modulo d, and it is an easy exercise to check that

$$H(d;N,K)=\left\{\begin{array}{cc}\frac{1}{d}\{(d-1)^K-(-1)^K\}, & d\nmid N\\ \frac{1}{d}\{(d-1)^K+(-1)^K(d-1)\}, & d|N.\end{array}\right.$$

Thus if $K\geq 9$ we have

$$H(3;N,K)\varepsilon(3)^{-K}\geq\frac{1}{3}(1-2^{-8})$$

and

$$H(5;N,K)\varepsilon(5)^{-K}\geq\frac{1}{5}(1-4^{-8}),$$

whence

$$2\sum_{d\leq D} k(d)H(d;N,K)\varepsilon(d)^{-K}\geq 2.7973$$

for any choice of N. We therefore conclude that

$$\Sigma_0\geq 2.7973(1-2\varpi)C_0N(\log N)^{-2}L^K+O(N(\log N)^{K-5/2}), \tag{26}$$

providing that $K\geq 9$.

To bound Σ_1 we note that

$$\frac{n}{\phi(n)}\ll\prod_{p|n,\,p\neq 2}(1+\frac{1}{p})=\sum_{q|n,\,2\nmid q}\frac{\mu^2(q)}{q}.$$

We deduce that

$$\Sigma_1\ll\sum_{q\leq N,\,2\nmid q}\frac{\mu^2(q)}{q}\#\{n:\,q|n\}.$$

However, if q is odd, then

$$\#\{\nu : 0\leq\nu\leq L,\,2^\nu\equiv m \pmod{q}\}\ll 1+\frac{L}{\varepsilon(q)}.$$

It follows that

$$\#\{n:\,q|n\}\ll L^{K-1}+L^K/\varepsilon(q),$$

whence

$$\Sigma_1\ll(\log N)^K+(\log N)^K\sum_{q\leq N,\,2\nmid q}\frac{\mu^2(q)}{q\varepsilon(q)}.$$

To bound the final sum we call on the following simple result of Gallagher [3; Lemma 4]

Lemma 6 *We have*

$$\sum_{\varepsilon(q)\le x} \frac{\mu^2(q)}{\phi^2(q)} q \ll \log x.$$

From this we deduce that

$$\sum_{x/2<\varepsilon(q)\le x} \frac{\mu^2(q)}{q\varepsilon(q)} \ll \frac{\log x}{x}. \tag{27}$$

We take x to run over powers of 2 and sum the resulting bounds to deduce that

$$\sum_{q\le N,\, 2\nmid q} \frac{\mu^2(q)}{q\varepsilon(q)} \ll 1,$$

and hence that

$$\Sigma_1 \ll (\log N)^K. \tag{28}$$

Turning now to Σ_2, we fix a particular pair of characters $\chi, \chi' \in \mathcal{B}(\eta)$, and investigate

$$\sum_n \frac{n}{\phi(n)} m_n^{-1/3} = \Sigma_2(\chi, \chi'),$$

say. Let $m = [r, r']$ as usual, and write $m = 2^\mu f$, with f odd. Put $g = (f, n)$ so that

$$m_n \ge f_n = f/g, \tag{29}$$

and consider

$$\sum_{g|n} \frac{n}{\phi(n)}.$$

As before we have

$$\frac{n}{\phi(n)} \ll \sum_{q|n,\, 2\nmid q} \frac{\mu^2(q)}{q}.$$

Terms q with $q \ge d(n)$ can contribute at most 1 in total, so that in fact

$$\frac{n}{\phi(n)} \ll \sum_{q|n,\, 2\nmid q, q\le d(n)} \frac{\mu^2(q)}{q}.$$

Thus, if

$$D = \max_{1\le n\le N} d(n),$$

we deduce as before that

$$\begin{aligned}\sum_{g|n} \frac{n}{\phi(n)} &\ll \sum_{q\le D,\, 2\nmid q} \frac{\mu^2(q)}{q} \#\{n : [g,q]|n\} \\ &\ll \sum_{q\le D,\, 2\nmid q} \frac{\mu^2(q)}{q} \{(\log N)^{K-1} + \frac{(\log N)^K}{\varepsilon([g,q])}\}.\end{aligned}$$

Here we note that

$$\sum_{q\le D} q^{-1} \ll \log D \ll \frac{\log N}{\log\log N}.$$

To deal with the remaining terms let ξ be a positive parameter. Then

$$\begin{aligned}\sum_{\varepsilon(q)>\xi}\frac{\mu^2(q)}{q\varepsilon([g,q])} &\le \sum_{\varepsilon(q)>\xi}\frac{\mu^2(q)}{q\varepsilon(q)}\\ &\ll \frac{\log\xi}{\xi},\end{aligned}$$

by (27). If $\varepsilon(q)\le\xi$ we note that

$$q\le 2^{\varepsilon(q)}-1,\quad\text{for}\quad q>1, \tag{30}$$

so that $q\le 2^{\xi}$. Thus

$$\begin{aligned}\sum_{\varepsilon(q)\le\xi}\frac{\mu^2(q)}{q\varepsilon([g,q])} &\le \sum_{q\le 2^{\xi}}\frac{\mu^2(q)}{q\varepsilon(g)}\\ &\le \frac{\xi}{\varepsilon(g)}.\end{aligned}$$

On choosing $\xi=\sqrt{\varepsilon(g)}$ we therefore conclude that

$$\sum_{2\nmid q}\frac{\mu^2(q)}{q\varepsilon([g,q])}\ll\frac{\log\varepsilon(g)}{\sqrt{\varepsilon(g)}},$$

and hence that

$$\sum_{g|n}\frac{n}{\phi(n)}\ll(\log N)^K\{(\log\log N)^{-1}+\varepsilon(g)^{-1/3}\}.$$

It follows from (30) that $\varepsilon(g)\gg\log g$, and we now conclude that

$$\sum_{g|n}\frac{n}{\phi(n)}\ll(\log N)^K\{(\log\log N)^{-1}+(\log g)^{-1/3}\}.$$

We now observe from (29) that

$$\Sigma_2(\chi,\chi')\le\sum_n\frac{n}{\phi(n)}(\frac{f}{(f,n)})^{-1/3}.$$

Let $\tau\ge 1$ be a parameter to be fixed in due course. Then terms in which $(f,n)\le f/\tau$ contribute

$$\le\tau^{-1/3}\sum_n\frac{n}{\phi(n)}=\tau^{-1/3}\Sigma_1\ll\tau^{-1/3}(\log N)^K,$$

by (28). The remaining terms contribute

$$\begin{aligned}&\le \sum_{g|f,\,g\ge f/\tau}(f/g)^{-1/3}\sum_{g|n}\frac{n}{\phi(n)}\\ &\ll \sum_{g|f,\,g\ge f/\tau}(f/g)^{-1/3}(\log N)^K\{(\log\log N)^{-1}+(\log g)^{-1/3}\}\\ &\ll \sum_{g|f,\,g\ge f/\tau}(\log N)^K\{(\log\log N)^{-1}+(\log f)^{-1/3}\}\\ &\ll \sum_{j|f,\,j\le\tau}(\log N)^K\{(\log\log N)^{-1}+(\log f)^{-1/3}\}\\ &\ll \tau(\log N)^K\{(\log\log N)^{-1}+(\log f)^{-1/3}\}.\end{aligned}$$

We deduce that

$$\Sigma_2(\chi,\chi') \ll \tau^{-1/3}(\log N)^K + \tau(\log N)^K\{(\log\log N)^{-1} + (\log f)^{-1/3}\}.$$

We therefore choose

$$\tau = \{(\log\log N)^{-1} + (\log f)^{-1/3}\}^{-3/4},$$

whence

$$\Sigma_2(\chi,\chi') \ll (\log N)^K\{(\log\log N)^{-1/4} + (\log f)^{-1/12}\}. \tag{31}$$

In order to bound f from below we note that, since χ,χ' are not both trivial, we may suppose that χ, say, is non-trivial. We then use a result of Iwaniec [6; Theorem 2]. This shows that if $L(\beta+i\gamma,\chi)=0$, with $|\gamma|\le N$, and χ of conductor $r\le N$, then either χ is real, or

$$1-\beta \gg \{\log d + (\log N\log\log N)^{3/4}\}^{-1},$$

where d is the product of the distinct prime factors of r. In our application we clearly have $f\ge d/2$, so that if χ, say, is in $\mathcal{B}(\eta)$ we must have

$$\frac{\eta}{\log N} \gg \{\log f + (\log N\log\log N)^{3/4}\}^{-1}$$

if χ is not real. Thus, if we insist that $\eta\le(\log N)^{1/5}$ it follows that either

$$\log f \gg \eta^{-1}\log N \gg (\log N)^{4/5},$$

or χ is real. Of course if χ is real we will have $16\not| r$, whence $f\gg r$. Moreover we will also have

$$(\log N)^{4/5} \gg \frac{\eta}{\log N} \gg 1-\beta \gg r^{\varpi-1/2},$$

so that $f\gg r\gg(\log N)^{3/2}$. Thus in either case we find that $\log f\gg\log\log N$, so that (31) yields

$$\Sigma_2(\chi,\chi') \ll (\log N)^K(\log\log N)^{-1/12}.$$

In view of the bound for $\#\mathcal{B}(\eta)$ given in Lemma 5, we conclude that

$$\Sigma_2 \ll e^{12\eta}(\log N)^K(\log\log N)^{-1/12}. \tag{32}$$

We may now insert the bounds (26), (28) and (32) into (21) to deduce that

$$\begin{aligned}\int_{\mathfrak{M}} S(\alpha)^2T(\alpha)^K e(-\alpha N)d\alpha \quad\ge\quad & 2.7973(1-2\varpi)C_0N(\log N)^{-2}L^K\\ &+O(N(\log N)^{K-5/2})\\ &+O(e^{-\varpi\eta}N(\log N)^{K-2})\\ &+O(e^{12\eta}N(\log N)^{K-2}(\log\log N)^{-1/12}).\end{aligned}$$

We therefore define η by taking

$$e^\eta = (\log\log N)^{1/145},$$

so that η satisfies the condition (14), and conclude as follows.

Lemma 7 *If $p\le N^{45/154-4\varpi}$ and $K\ge 9$ we have*

$$\int_{\mathfrak{M}} S(\alpha)^2T(\alpha)^K e(-\alpha N)d\alpha \ge 2.7973(1-3\varpi)C_0N(\log 2)^{-2}L^{K-2}$$

for large enough N.

5 A Mean Square Estimate

In this section we shall estimate the mean square

$$J(\mathfrak{m}) = \int_{\mathfrak{m}} |S(\alpha)T(\alpha)|^2 d\alpha,$$

where $\mathfrak{m} = [0,1] \setminus \mathfrak{M}$ is the set of minor arcs. Instead of this integral, previous researchers have worked with the larger integral

$$J = \int_0^1 |S(\alpha)T(\alpha)|^2 d\alpha.$$

Thus it was shown by Li [9; Lemma 6], building on work of Liu, Liu and Wang [13; Lemma 4] that

$$J \le (24.95 + o(1))\frac{C_0}{\log^2 2} N,$$

In this section we shall improve on this bound, and give a lower bound for the corresponding major arc integral

$$J(\mathfrak{M}) = \int_{\mathfrak{M}} |S(\alpha)T(\alpha)|^2 d\alpha.$$

By subtraction we shall then obtain our bound for $J(\mathfrak{m})$.

We begin by observing that

$$J = \sum_{\mu,\nu \le L} r(2^\mu - 2^\nu),$$

where

$$r(n) = \#\{\varpi N < p_1, p_2 \le N : n = p_1 - p_2\}.$$

Moreover, by Theorem 3 of Chen [2] we have

$$r(n) \le C_0 C_1 h(n) \frac{N}{(\log N)^2},$$

for $n \ne 0$ and N sufficiently large, where C_0 is given by (24),

$$C_1 = 7.8342, \tag{33}$$

and

$$h(n) = \prod_{p|n,\, p>2} (\frac{p-1}{p-2}).$$

Observe that our notation for the constants that occur differs from that used by Liu, Liu and Wang, and by Li. Since $h(2^\mu - 2^\nu) = h(2^{\mu-\nu} - 1)$ for $\mu > \nu$ we conclude, as in Liu, Liu and Wang [13; §3] and Li [9; §4] that

$$\sum_{\mu \ne \nu \le L} r(2^\mu - 2^\nu) \le 2C_0 C_1 \frac{N}{(\log N)^2} \sum_{1 \le l \le L} (L-l) h(2^l - 1), \tag{34}$$

while the contribution for $\mu = \nu$ is $L\pi(N) - L\pi(\varpi N) \le LN(\log N)^{-1}$, for large N. Now

$$h(n) = \sum_{d|n} k(d),$$

where $k(d)$ is the multiplicative function defined in (25). Thus

$$\begin{aligned}\sum_{1\leq j\leq J} h(2^j-1) &= \sum_{d=1}^{\infty} k(d)\#\{j\leq J : d|2^j-1\} \\ &= \sum_{d=1}^{\infty} k(d)[\frac{J}{\varepsilon(d)}].\end{aligned}$$

However $[\theta]=\theta+O(\theta^{1/2})$ for any real $\theta>0$, whence

$$\sum_{1\leq j\leq J} h(2^j-1) = C_2J+O(J^{1/2}) \tag{35}$$

with

$$C_2=\sum_{d=1}^{\infty}\frac{k(d)}{\varepsilon(d)}. \tag{36}$$

Here we use the observation that the sum

$$\sum_{d=1}^{\infty}\frac{k(d)}{\varepsilon(d)^{1/2}}$$

is convergent, since Lemma 6 implies that

$$\sum_{x/2<\varepsilon(d)\leq x}\frac{k(d)}{\varepsilon(d)^{1/2}} \ll x^{-1/2}\sum_{x/2<\varepsilon(d)\leq x}\frac{\mu^2(d)d}{\phi^2(d)} \ll \frac{\log x}{x^{1/2}} \tag{37}$$

for any $x\geq 2$.

We may now use partial summation in conjunction with (35) to deduce that

$$\sum_{1\leq l\leq L}(L-l)h(2^l-1) = C_2\frac{L^2}{2}+O(L^{3/2}),$$

Thus, using (34) we reach the following result.

Lemma 8 *We have*

$$J\leq\{\frac{C_0C_1C_2}{\log^2 2}+\frac{1}{\log 2}+o(1)\}N,$$

with the constants given by (24), (33) and(36).

We now turn to the integral $J(\mathfrak{M})$. According to Lemma 3.1 of Vaughan [17], if

$$|\alpha-\frac{a}{q}|\leq\frac{\log x}{x},\quad (a,q)=1,$$

and $q\leq 2\log x$, we have

$$\sum_{p\leq x} e(\alpha p)\log p = \frac{\mu(q)}{\phi(q)}v(\alpha-\frac{a}{q})+O(x(\log x)^{-3}),$$

with

$$v(\beta)=\sum_{m\leq x} e(\beta m).$$

It follows by partial summation that

$$S(\alpha)=\frac{\mu(q)}{\phi(q)}w(\alpha-\frac{a}{q})+O(N(\log N)^{-4}),$$

with

$$w(\beta)=\sum_{\varpi N<m\le N}\frac{e(\beta m)}{\log m},$$

providing that

$$|\alpha-\frac{a}{q}|\le\frac{\log N}{N},\quad (a,q)=1 \tag{38}$$

and $q\le\log N$. Then if $\mathfrak{a}$ denotes the set of $\alpha\in[0,1]$ for which such a,q exist, we easily compute that

$$\begin{aligned} J(\mathfrak{M}) &\ge J(\mathfrak{a}) \\ &= \int_{\mathfrak{a}}|\frac{\mu(q)}{\phi(q)}w(\alpha-\frac{a}{q})T(\alpha)|^2d\alpha+O(N(\log N)^{-1}), \end{aligned}$$

where, for each $\alpha\in\mathfrak{a}$, we have taken a/q to be the unique rational satisfying (38). By partial summation we have

$$w(\beta)\ll(||\beta||\log N)^{-1},$$

whence

$$\int_{-(\log N)/N}^{(\log N)/N}|w(\beta)T(\frac{a}{q}+\beta)|^2d\beta=\int_{-1/2}^{1/2}|w(\beta)T(\frac{a}{q}+\beta)|^2d\beta+O(N(\log N)^{-1}).$$

It follows that

$$J(\mathfrak{a})=\sum_{q\le\log N}\sum_{(a,q)=1}\frac{\mu^2(q)}{\phi^2(q)}\int_0^1|w(\beta)T(\frac{a}{q}+\beta)|^2d\beta+O(N(\log N)^{-1}\log\log N).$$

The integral on the right is

$$\sum_{0\le\mu,\nu\le L}e(a(2^\mu-2^\nu)/q)S(2^\mu-2^\nu),$$

where

$$\begin{aligned} S(n) &= \sum_{\substack{\varpi N<m_1,m_2\le N\\ m_1-m_2=n}}(\log m_1)^{-1}(\log m_2)^{-1} \\ &= (\log N)^{-2}\#\{m_1,m_2:\varpi N<m_1,m_2\le N,\, m_1-m_2=n\} \\ &\qquad +O(N(\log N)^{-3}) \\ &= (\log N)^{-2}\max\{N(1-\varpi)-|n|\,,\,0\}+O(N(\log N)^{-3}). \end{aligned}$$

Thus

$$S(n)=(1-\varpi)N(\log N)^{-2}+O(|n|(\log N)^{-2})+O(N(\log N)^{-3}) \tag{39}$$

for $n \ll N$. On summing over a we now obtain

$$J(\mathfrak{a}) = \sum_{0 \le \mu,\nu \le L} \sum_{q \le \log N} \frac{\mu^2(q)}{\phi^2(q)} c_q(2^\mu - 2^\nu) S(2^\mu - 2^\nu) + O(N(\log N)^{-1} \log\log N),$$

where $c_q(n)$ is the Ramanujan sum. When q is square-free we have $c_q(n) = \mu(q)\mu((q,n))\phi((q,n))$. Thus the error terms in (39) make a total contribution $O(N(\log N)^{-1}\log\log N)$ to $J(\mathfrak{a})$. Moreover

$$\mu^2(q)c_q(n) = \mu(q) \sum_{d|(q,n)} \mu(d)d,$$

whence

$$\sum_{0 \le \mu,\nu \le L} \mu^2(q)c_q(n) = \mu(q) \sum_{d|q} \mu(d)d\#\{\mu,\nu:\, 1 \le \mu,\nu \le L,\, d|2^\mu - 2^\nu\}.$$

If d is odd we have

$$\#\{\mu,\nu:\, 1 \le \mu,\nu \le L,\, d|2^\mu - 2^\nu\} = L^2\varepsilon(d)^{-1} + O(L),$$

while if d is even, of the form $2e$ with e odd, we have

$$\#\{\mu,\nu:\, 1 \le \mu,\nu \le L,\, d|2^\mu - 2^\nu\} = L^2\varepsilon(e)^{-1} + O(L).$$

The error terms contribute $O(N(\log N)^{-1}\log\log N)$ to $J(\mathfrak{a})$, by (39), so that

$$J(\mathfrak{a}) = \frac{(1-\varpi)N}{(\log N)^2} L^2 \sum_{q \le \log N} \frac{\mu(q)}{\phi^2(q)} \sum_{d|q} \mu(d)d\varepsilon(d)^{-1} + O(N(\log N)^{-1}\log\log N),$$

where $\varepsilon(d)$ is to be interpreted as $\varepsilon(e)$ when $d = 2e$. Now

$$\begin{aligned}
\sum_{q \le \log N} \frac{\mu(q)}{\phi^2(q)} \sum_{d|q} \frac{\mu(d)d}{\varepsilon(d)} &= \sum_{d \le \log N} \frac{\mu(d)d}{\varepsilon(d)} \sum_{\substack{q \le \log N \\ d|q}} \frac{\mu(q)}{\phi^2(q)} \\
&= \sum_{d \le \log N} \frac{\mu(d)d}{\varepsilon(d)} \sum_{j \le (\log N)/d} \frac{\mu(jd)}{\phi^2(jd)} \\
&= \sum_{d \le \log N} \frac{\mu^2(d)d}{\varepsilon(d)\phi^2(d)} \sum_{\substack{j \le (\log N)/d \\ (j,d)=1}} \frac{\mu(j)}{\phi^2(j)} \\
&= \sum_{d \le \log N} \frac{\mu^2(d)d}{\varepsilon(d)\phi^2(d)} \{ \sum_{\substack{j=1 \\ (j,d)=1}}^{\infty} \frac{\mu(j)}{\phi^2(j)} + O(\frac{d}{\log N}) \} \\
&= \sum_{d \le \log N} \frac{\mu^2(d)d}{\varepsilon(d)\phi^2(d)} \prod_{p \nmid d} \{1 - (p-1)^{-2}\} \\
&\qquad + O((\log N)^{-1} \sum_{d \le \log N} \frac{\mu^2(d)d^2}{\varepsilon(d)\phi^2(d)}). \qquad (40)
\end{aligned}$$

If $d = 2e$ with e odd, we have

$$\frac{\mu^2(d)d}{\varepsilon(d)\phi^2(d)} \prod_{p \nmid d} \{1 - (p-1)^{-2}\} = 2C_0 k(e)/\varepsilon(d),$$

while if d is odd we have

$$\prod_{p \nmid d} \{1 - (p-1)^{-2}\} = 0,$$

since the factor with $p = 2$ vanishes. Moreover

$$\sum_{d \gg \log N} \frac{k(d)}{\varepsilon(d)} \ll \frac{\log N}{\log\log N}$$

by Lemma 6, applied as in (37). The leading term in (40) is therefore $2C_0C_2 + o(1)$, with C_0 and C_2 as in (24) and (36).

To bound the error term we use Lemma 6, which shows that

$$\sum_{\substack{X < d \le 2X \\ x < \varepsilon(d) \le 2x}} \frac{\mu^2(d)d^2}{\varepsilon(d)\phi^2(d)} \ll \frac{X \log x}{x}.$$

According to (30) we must have $x \gg \log X$, so on summing as x runs over powers of 2 we obtain

$$\sum_{X < d \le 2X} \frac{\mu^2(d)d^2}{\varepsilon(d)\phi^2(d)} \ll \frac{X \log\log X}{\log X}.$$

Now, summing as X runs over powers of 2 we conclude that

$$\sum_{d \le \log N} \frac{\mu^2(d)d^2}{\varepsilon(d)\phi^2(d)} \ll \frac{(\log N)(\log\log\log N)}{\log\log N}.$$

We may therefore summarize our results as follows.

Lemma 9 *We have*

$$J(\mathfrak{M}) \ge \{\frac{2(1-\varpi)C_0C_2}{\log^2 2} + o(1)\}N,$$

and hence

$$J(\mathfrak{m}) \le \{\frac{C_0(C_1 - 2 + 2\varpi)C_2}{\log^2 2} + \frac{1}{\log 2} + o(1)\}N,$$

by Lemma 8.

It remains to compute the constants. We readily find

$$\prod_{2 < p \le 200000} (1 - (p-1)^{-2}) = 0.6601...$$

Since
$$\prod_{p>K}(1-(p-1)^{-2}) \geq \prod_{n=K}^{\infty}(1-n^{-2}) = 1-K^{-1},$$
we deduce that
$$C_0 \geq 0.999995 \times 0.6601 \geq 0.66. \tag{41}$$
However the estimation of C_2 is more difficult. We set
$$m = \prod_{e \leq x}(2^e - 1)$$
and
$$s(x) = \sum_{\varepsilon(d) \leq x} k(d),$$
whence
$$\begin{aligned} s(x) &\leq \sum_{d|m} k(d) \\ &= h(m) \\ &= \prod_{p|m,\, p>2}\left(\frac{p-1}{p-2}\right) \\ &\leq \prod_{p>2}\left(\frac{(p-1)^2}{p(p-2)}\right)\prod_{p|m}\left(\frac{p}{p-1}\right) \\ &= C_0^{-1}\frac{m}{\phi(m)}. \end{aligned}$$
Moreover we have $m/\phi(m) \leq e^{\gamma}\log x$ for $x \geq 9$, as shown by Liu, Liu and Wang [13; (3.9)]. It then follows that
$$\begin{aligned} C_2 &= \int_1^{\infty} s(x)\frac{dx}{x^2} \\ &= \int_1^{M} s(x)\frac{dx}{x^2} + \int_M^{\infty} s(x)\frac{dx}{x^2} \\ &\leq \sum_{\varepsilon(d) \leq M}\int_{\varepsilon(d)}^{M} k(d)\frac{dx}{x^2} + C_0^{-1}e^{\gamma}\int_M^{\infty}\log x\frac{dx}{x^2} \\ &\leq \sum_{\varepsilon(d) < M} k(d)\left(\frac{1}{\varepsilon(d)} - \frac{1}{M}\right) + 2.744\left(\frac{1+\log M}{M}\right) \end{aligned}$$
for any integer $M \geq 9$.

We now set
$$\sum_{\varepsilon(d)=e} k(d) = \kappa(e)$$
so that
$$\sum_{e|d}\kappa(e) = \sum_{\varepsilon(e)|d} k(e).$$

However $\varepsilon(e)|d$ if and only if $e|2^d-1$. Thus

$$\sum_{e|d}\kappa(e)=\sum_{e|2^d-1}k(e)=h(2^d-1).$$

We therefore deduce that

$$\kappa(e)=\sum_{d|e}\mu(e/d)h(2^d-1).$$

This enables us to compute

$$\sum_{\varepsilon(d)<M}k(d)(\frac{1}{\varepsilon(d)}-\frac{1}{M})=\sum_{m<M}\kappa(m)(\frac{1}{m}-\frac{1}{M})$$

by using information on the prime factorization of 2^d-1 for $d<M$. In particular, taking $M=20$ we find that

$$\sum_{m<20}\kappa(m)(\frac{1}{m}-\frac{1}{20})=1.6659\ldots,$$

and hence that

$$C_2\le\sum_{m<20}\kappa(m)(\frac{1}{m}-\frac{1}{20})+2.744(\frac{1+\log 20}{20})=2.2141\ldots \tag{42}$$

For comparison with this upper bound for C_2 we note that

$$C_2\ge\sum_{d\le 10000}k(d)/\varepsilon(d)=1.9326\ldots$$

This latter figure is probably closer to the true value, but the discrepancy is small enough for our purposes.

From (33), (41) and (42) we calculate that

$$(C_1-2)C_2+C_0^{-1}\log 2\le 13.967,$$

so that Lemma 9 yields the following bound.

Lemma 10 *We have*

$$J(\mathfrak{m})\le\{13.968+o(1)\}C_0\frac{N}{\log^2 2}.$$

6 Completion of the Proof

Let $R(N)$ denote the number of representations of N as a sum of two primes and K powers of 2 in the ranges under consideration, so that

$$R(N)=\int_0^1 S(\alpha)^2T(\alpha)^Ke(-\alpha N)d\alpha.$$

We estimate the minor arc contribution to $R(N)$ by using Hölder's inequality as follows. We have

$$
\begin{aligned}
&|\int_{\mathfrak{m}} S(\alpha)^2 T(\alpha)^K e(-\alpha N) d\alpha| \\
&\qquad \leq \int_{\mathfrak{m}} |S(\alpha)^2 T(\alpha)^K| d\alpha \\
&\qquad \leq (\sup_{\alpha\in\mathfrak{m}} |S(\alpha)|)^{2\mu} (\int_{\mathfrak{m}} |T(\alpha)|^{2q} d\alpha)^{\mu} (\int_{\mathfrak{m}} |S(\alpha)T(\alpha)|^2 d\alpha)^{\nu} \\
&\qquad \leq (\sup_{\alpha\in\mathfrak{m}} |S(\alpha)|)^{2\mu} I(q)^{\mu} J(\mathfrak{m})^{\nu}, \qquad (43)
\end{aligned}
$$

where

$$\mu = \frac{K-2}{2q-2}, \quad \nu = \frac{2q-K}{2q-2}.$$

We shall apply this bound with

$$q = [\xi^{-1} \log N]$$

for a suitably chosen positive constant ξ, see (44).

According to Theorem 3.1 of Vaughan [17] we have

$$\sum_{p\leq x} e(\alpha p) \log p \ll (\log x)^4 \{x q^{-1/2} + x^{4/5} + x^{1/2} q^{1/2}\}$$

if $|\alpha - a/q| \leq q^{-2}$ with $(a,q) = 1$. Thus if $\alpha \in \mathfrak{m}$ we may take $P \ll q \ll N/P$ to deduce that

$$S(\alpha) \ll (\log N)^3 \{N^{4/5} + N P^{-1/2}\}.$$

Taking $P = N^{45/154 - 4\varpi}$, we obtain

$$S(\alpha) \ll N^{263/308 + 3\varpi}.$$

If one assumes the Generalized Riemann Hypothesis, we may apply Lemma 12 of Baker and Harman [1], which implies that

$$\sum_{n\leq x} \Lambda(n) e((\frac{a}{q} + \beta)n) \ll (\log x)^2 \{q^{-1} \min(x, |\beta|^{-1}) + x^{1/2} q^{1/2} + x(q|\beta|)^{1/2}\}$$

when $|\beta| \leq x^{-1/2}$. It follows by partial summation that

$$S(\frac{a}{q} + \beta) \ll (\log N) \{q^{-1} \min(N, |\beta|^{-1}) + N^{1/2} q^{1/2} + N(q|\beta|)^{1/2}\}$$

for $|\beta| \leq N^{-1/2}$. According to Dirichlet's Approximation Theorem, we can find a and q with

$$|\alpha - \frac{a}{q}| \leq \frac{1}{qN^{1/2}}, \quad q \leq N^{1/2}.$$

Thus

$$S(\alpha) \ll (\log N) N^{3/4}$$

unless $q \leq N^{1/4}$ and $|\alpha - a/q| \leq q^{-1} N^{-3/4}$. Since $\alpha \in \mathfrak{m}$ and $P = N^{45/154 - 4\varpi} \geq N^{1/4}$, these latter conditions cannot hold.

We therefore conclude that

$$S(\alpha) \ll N^{\theta+o(1)}$$

for $\alpha \in \mathfrak{m}$, where we take $\theta = 263/308$ in general, and $\theta = 3/4$ under the Generalized Riemann Hypothesis. Thus

$$(\sup_{\alpha\in\mathfrak{m}} |S(\alpha)|)^{2\mu} \leq (1+o(1))\exp(\theta(K-2)\xi).$$

From (2) and Lemma 1 we see that

$$\begin{aligned} I(q)^{\mu} &\leq (1+o(1))\{(\frac{q}{e})^{q}(1.753)^{q}L^{q}\}^{\mu} \\ &= (1+o(1))\{\frac{q}{e}(1.753)L\}^{q\mu} \\ &= (1+o(1))\{\frac{q}{e}(1.753)L\}^{(K-2)/2} \\ &= (1+o(1))\{\frac{1.753\log 2}{\xi e}\}^{(K-2)/2}L^{K-2}. \end{aligned}$$

Moreover Lemma 10 yields

$$\begin{aligned} J(\mathfrak{m})^{\nu} &\leq (1+o(1))13.968\frac{C_0}{\log^2 2}N^{\nu} \\ &= (1+o(1))13.968\frac{C_0}{\log^2 2}N\exp(-\xi(K-2)/2). \end{aligned}$$

On combining these estimates we see from (43) that

$$|\int_{\mathfrak{m}} S(\alpha)^2T(\alpha)^K e(-\alpha N)d\alpha| \leq (1+o(1))13.968\frac{C_0}{\log^2 2}NL^{K-2}\lambda^{(K-2)/2},$$

where

$$\lambda = e^{(2\theta-1)\xi}\frac{1.753\log 2}{\xi e}.$$

We minimize λ by choosing

$$\xi = (2\theta-1)^{-1}, \tag{44}$$

so that

$$\lambda = (2\theta-1)1.753\log 2$$

and

$$\begin{aligned} &|\int_{\mathfrak{m}} S(\alpha)^2T(\alpha)^K e(-\alpha N)d\alpha| \\ &\leq (1+o(1))13.968\frac{C_0}{\log^2 2}NL^{K-2}\{(2\theta-1)1.753\log 2\}^{(K-2)/2}. \end{aligned}$$

Finally we compare this with the estimate for the major arc integral, given by Lemma 7, and conclude that

$$\int_0^1 S(\alpha)^2T(\alpha)^K e(-\alpha N)d\alpha > 0$$

providing that N is large enough, ϖ is small enough, and

$$13.968\{(2\theta-1)1.753\log 2\}^{(K-2)/2} < 2.7973. \tag{45}$$

When $\theta = 263/308$ this is satisfied for $K > 23.4$, so that $K = 24$ is admissible. Similarly, when $\theta = 3/4$ one can take any $K > 8.5$, so that $K = 9$ is admissible. This completes the proof of our theorems, subject to Lemma 1.

7 The Proof of Lemma 1 — First Steps

We begin by recalling that $r(n,q,N)$ is the number of representations of an integer n as a sum of q terms $2^\nu \le N/2K$, and that

$$r(q) = \max_{n,N} \, r(n,q,N).$$

It might seem that the above maximum should be attained for an integer n which is a sum of q distinct powers of 2, and that one should therefore have $r(q) = q!$. Unfortunately this simple conjecture is false, since for $q = 4$ one has

$$36 = 32 + 2 + 1 + 1 = 16 + 8 + 8 + 4 = 16 + 16 + 2 + 2,$$

whence $r(36, 4, 32) = 30$, after allowing for re-orderings of the above representations.

In fact the behaviour of $r(q)$ is determined by the properties of a slightly simpler counting function, $R(q)$, defined as the number of representations of the number 1 as a sum of q negative powers of 2. We adopt the convention that $R(0) = 0$. Numerical computation shows that $R(1) = 1$, $R(2) = 1$, $R(3) = 3$, $R(4) = 13, \ldots$, and it is not hard to see that $R(q)$ is always finite.

The following result expresses the connection between the two functions.

Lemma 11 *We have*

$$r(q) = \max_{h \le q} \sum_{q_1+\ldots+q_h=q} R(q_1)\ldots R(q_h) \begin{pmatrix} q \\ q_1 \ldots q_h \end{pmatrix},$$

where

$$\begin{pmatrix} q \\ q_1 \ldots q_h \end{pmatrix} = \frac{q!}{q_1! \ldots q_h!}.$$

To prove Lemma 11 we first suppose that the maximum on the right is attained at a particular value for h, and that all the relevant negative powers of 2 which occur in the corresponding counting functions $R(q_i)$ take the form $2^{-\nu}$ with $\nu < d$, say. We now consider the integer

$$n = 2^d + 2^{2d} + \ldots + 2^{hd}. \tag{46}$$

For each set of values $q_1, \ldots, q_h$ for which $q_1 + \ldots + q_h = q$, there are

$$\begin{pmatrix} q \\ q_1 \ldots q_h \end{pmatrix}$$

ways to split the integers $1, \ldots, q$ into sets $S_1, \ldots, S_h$ of cardinalities $q_1, \ldots, q_h$. Moreover each power 2^{jd} may be decomposed into q_j non-negative powers of 2 in $R(q_j)$ ways. If

$$2^{jd} = \sum_{1 \le i \le q_j} 2^{\nu_{ij}} \tag{47}$$

is such a representation, we may define the sequence $\mu_1, \ldots, \mu_q$ by taking $\mu_k = \nu_{ij}$ when $k \in S_j$ and k is the i-th largest element of S_j. We then have

$$n = \sum 2^{\mu_k}. \tag{48}$$

Moreover different representations (47) will lead to different solutions of (48). It therefore follows that

$$r(q) \geq \max_{h \leq q} \sum_{q_1+\ldots+q_h=q} R(q_1)\ldots R(q_h) \left(\begin{array}{c} q \\ q_1 \ldots q_h \end{array} \right).$$

To prove the reverse inequality we use the following fact.

Lemma 12 *Let*

$$n = 2^a + 2^b + \ldots + 2^c \quad (a > b > \ldots > c \geq 0)$$

be the binary representation of a positive integer n, and let

$$n = \sum_{\nu \in S} 2^{\nu}$$

for some finite set of non-negative integers S. Then there is a subset T of S such that

$$2^a = \sum_{\nu \in T} 2^{\nu}.$$

We shall prove this later in this section. Now suppose that $r(n, q, N) = r(q)$, and let

$$n = 2^{a_1} + \ldots + 2^{a_h}$$

be the binary representation of n. It then follows via repeated applications of Lemma 12 that, given any expression

$$n = 2^{\nu_1} + \ldots + 2^{\nu_q},$$

we may partition the integers $\nu_1, \ldots, \nu_q$ into sets $T_1, \ldots, T_h$ such that

$$2^{a_j} = \sum_{i \in T_j} 2^i. \tag{49}$$

If we set $q_j = \#T_j$ we see that each such relation contributes a solution to $R(q_j)$. It is thus easy to see that

$$r(q) = r(n, q, N) \leq \sum_{q_1+\ldots+q_h=q} R(q_1)\ldots R(q_h) \left(\begin{array}{c} q \\ q_1 \ldots q_h \end{array} \right).$$

This completes the proof of Lemma 11, subject to Lemma 12.

We remark that there is potentially some loss in our argument in employing Lemma 11. In order to achieve equality in the lemma we have considered an integer (46) for which the exponents are well-spaced. In general we might expect a certain amount of overlap between the representations (49) with nearby values of a_j.

We now establish Lemma 12. Since $n < 2^{a+1}$ we have $\nu \leq a$ for all $\nu \in S$. Choose $T \subseteq S$ so that $\sum_{\nu \in T} 2^{\nu} = \Sigma$, say, is minimal, subject to having $\Sigma \geq 2^a$. Then if μ is a minimal element of T we will have $\Sigma - 2^{\mu} < 2^a \leq \Sigma$. Thus

$$\sum_{\nu \in T} 2^{\nu-\mu} - 1 < 2^{a-\mu} \leq \sum_{\nu \in T} 2^{\nu-\mu}.$$

However the expressions above are all integers, so that we must have $2^{a-\mu} = \sum_{\nu \in T} 2^{\nu-\mu}$. The lemma then follows.

Our next result shows how information on the size of $R(q)$ leads to a bound for $r(q)$.

Lemma 13 *Suppose that ρ is a positive real root of the equation*

$$\sum_{q=1}^{\infty} \frac{R(q)}{q!} x^{-q} = 1. \tag{50}$$

Then $r(q) \le \rho^q q!$ for all $q \in \mathbb{N}$.

For the proof it will be convenient to write

$$f(q) = \frac{R(q)}{q!}.$$

From Lemma 11 we have

$$r(q) \le \sum_{1 \le h \le q} \sum_{q_1+\ldots+q_h=q} R(q_1)\ldots R(q_h) \left(\begin{array}{c} q \\ q_1 \ldots q_h \end{array} \right),$$

whence $r(q) \le r_0(q) q!$, where

$$r_0(q) = \sum_{1 \le h \le q} \sum_{q_1+\ldots+q_h=q} f(q_1)\ldots f(q_h).$$

We may rewrite this as

$$\begin{aligned} r_0(q) &= \sum_{1 \le h \le q} \sum_{q_h} f(q_h) \sum_{q_1+\ldots+q_{h-1}=q-q_h} f(q_1)\ldots f(q_{h-1}) \\ &= f(q) + \sum_{m=1}^{q-1} f(m) r_0(q-m). \end{aligned}$$

Here the term $f(q)$ arises from the case $h = 1$. If we define $r_0(0) = 1$ we therefore have a recurrence relation

$$r_0(q) = \sum_{m=1}^{q} f(m) r_0(q-m), \tag{51}$$

valid for all $q \ge 1$.

We may now prove that $r_0(q) \le \rho^q$, using induction on q. This will suffice for Lemma 13. The claim is trivial for $q = 0$. Assuming that $r_0(n) \le \rho^n$ for $0 \le n < q$ we deduce from (51) that

$$\begin{aligned} r_0(q) &\le \sum_{m=1}^{q} f(m) \rho^{q-m} \\ &\le \rho^q \sum_{m=1}^{\infty} f(m) \rho^{-m} \\ &= \rho^q, \end{aligned}$$

by definition of ρ. This completes the induction, and with it, the proof of Lemma 13.

8 Proof of Lemma 1 — Bounds for $f(m)$

It is clear from Lemma 13 that our remaining task is to estimate from above the positive root of the equation (50). This will require upper bounds for the numbers $f(m)$. We will consider two different ranges for m. For $m \le 20$, we shall compute $f(m)$ precisely, and for $m \ge 21$ we shall use a recursive estimate to bound $f(m)$.

Our primary tool in studying the function $R(q)$ will be the use of "splitting sequences". Suppose that we have an ordered representation

$$1 = 2^{-e_1} + \ldots + 2^{-e_q}.$$

From this we produce the corresponding un-ordered representation

$$1 = 2^{-f_1} + \ldots + 2^{-f_q}, \quad f_1 \le \ldots \le f_q. \tag{52}$$

We introduce the notation $(a_0, a_1, \ldots, a_k)$ for this un-ordered representation, where

$$a_i = \#\{j : f_j = i\}$$

and $k = f_q$. Clearly we have $q = \sum a_i$. Moreover we have $a_k \ge 2$ except for the trivial representation $1 = 2^{-0}$ given by the sequence (1). (There should be no confusion between this notation for a sequence, and references to equation (1) in §1.) From any decomposition $(a_0, \ldots, a_k)$ with $q \ge 2$ we can produce a "derived" representation $(a_0, \ldots, a_{k-2}, a_{k-1}+1, a_k-2)$ (if $a_k > 2$) or $(a_0, \ldots, a_{k-2}, a_{k-1}+1)$ (if $a_k = 2$). Set $g_{q-1} = 1$ in the first case and $g_{q-1} = 0$ in the second. We now repeat the process with the derived representation, and set $g_{q-2} = 1$ or 0 accordingly. We continue this process until we reach the trivial representation (1). The sequence $[g_1, \ldots, g_{q-1}]$ will be called the "splitting sequence" of the un-ordered decomposition (52).

As an example we may start with

$$\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16}; \quad (0, 0, 3, 0, 4). \tag{53}$$

The derived representation is

$$\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{16}; \quad (0, 0, 3, 1, 2),$$

whence $g_6 = 1$. The next derived representation is

$$\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8}; \quad (0, 0, 3, 2),$$

whence $g_5 = 0$. This is followed by

$$\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4}; \quad (0, 0, 4),$$

with $g_4 = 0$. We then have

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{4}; \quad (0, 1, 2),$$

where $g_3 = 1$. Next is

$$\frac{1}{2} + \frac{1}{2}; \quad (0, 2),$$

giving $g_2 = 0$. Finally we come down to

$$\frac{1}{1}; \quad (1),$$

so that $g_1 = 0$. Thus the splitting sequence is $[0, 0, 1, 0, 0, 1]$.

A key fact is that this process may be reversed, knowing the splitting sequence, since the value of g_i tells us whether to split the final term or a term of penultimate size. Thus in our example, given that $g_3 = 1$, the sequence from which

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{4}; \quad (0, 1, 2)$$

was derived must be the result of splitting $1/2$ into $1/4 + 1/4$. It follows that un-ordered decompositions are uniquely determined by their splitting sequences. However not every sequence of 0's and 1's is a splitting sequence. Clearly we must have $g_1 = g_2 = 0$, but there are other necessary conditions. Thus, for example, $[0, 0, 1, 1]$ is not a splitting sequence.

We are now ready to explain our procedure for computing $R(q)$ explicitly. We shall define $R(q, a, b)$ as the number of ordered representations for which the corresponding un-ordered decomposition takes the form $(a_0, a_1, \ldots, a_k)$ with $a_{k-1} = a$ and $a_k = b$. When $q \geq 2$ and $b = 2$ the derived sequence will contribute to $R(q-1, c, a+1)$ for some c. Since we are counting ordered representations we must allow for $q(q-1)/2$ possible locations for the two powers 2^{-a_k}, and also $q-1$ locations for the newly created term $2^{-a_{k-1}}$. Finally, this new term is one of $a+1$ such terms in the representation counted by $R(q-1, c, a+1)$. It therefore follows that

$$\begin{aligned} R(q, a, 2) &= \frac{q(q-1)}{2} \frac{a+1}{q-1} \sum_{c=0}^{q-a-2} R(q-1, c, a+1) \\ &= \frac{q(a+1)}{2} \sum_{c=0}^{q-a-2} R(q-1, c, a+1). \end{aligned} \tag{54}$$

We may analyse $R(q, a, b)$ when $b > 2$ in a similar way, to derive the recurrence formula

$$\begin{aligned} R(q, a, b) &= \frac{q(q-1)}{2} \frac{2}{b(b-1)} \frac{a+1}{q-1} R(q-1, a+1, b-2) \\ &= \frac{q(a+1)}{b(b-1)} R(q-1, a+1, b-2), \quad (b > 2). \end{aligned} \tag{55}$$

Finally we note that $R(q)$ is the sum of $R(q, a, b)$ for all available a and b.

Using these recursion formulae we may calculate $R(q, a, b)$ for all a, b, and for all $q \leq 20$. Summing the resulting values we obtain the following table.

Table 1

$f(1) = 1.000\ldots,$	$f(2) = 0.500\ldots,$	$f(3) = 0.500\ldots,$
$f(4) = 0.541\ldots,$	$f(5) = 0.625\ldots,$	$f(6) = 0.729\ldots,$
$f(7) = 0.862\ldots,$	$f(8) = 1.022\ldots,$	$f(9) = 1.217\ldots,$
$f(10) = 1.449\ldots,$	$f(11) = 1.727\ldots,$	$f(12) = 2.059\ldots,$
$f(13) = 2.455\ldots,$	$f(14) = 2.928\ldots,$	$f(15) = 3.492\ldots,$
$f(16) = 4.165\ldots,$	$f(17) = 4.968\ldots,$	$f(18) = 5.924\ldots,$
$f(19) = 7.066\ldots,$	$f(20) = 8.428\ldots.$	

We turn next to the large values of q. Let $[g_1, \ldots, g_{q-1}]$ be the splitting sequence of the un-ordered decomposition (52), and suppose that the sequence contains exactly l elements for which $g_i = 0$, namely

$$g_{i_1} = \ldots = g_{i_l} = 0, \quad i_1 < \ldots < i_l.$$

We shall write $h_j = i_{j+1} - i_j - 1$ for $1 \le j \le l-1$, and $h_l = q-1-i_l$. Thus in our example (53) we have the splitting sequence $[0,0,1,0,0,1]$, whence $h_1 = 0$, $h_2 = 1$, $h_3 = 0$ and $h_4 = 1$. It is then apparent that for an un-ordered decomposition $(a_0, a_1, \ldots, a_k)$ we have $k = l$. Moreover we have $a_j = 2h_j - h_{j+1} + 1$ for $1 \le j \le l-1$, and $a_l = 2h_l + 2$. We note in particular that any splitting sequence must satisfy $h_1 = 0$ and $h_{j+1} \le 2h_j + 1$ for $1 \le j \le l-1$, since we must have $a_j \ge 0$.

The un-ordered decomposition $(a_0, a_1, \ldots, a_k)$ produces

$$\frac{q!}{a_0! \ldots a_k!}$$

ordered representations, so that the unordered decompositions with splitting sequence $[g_1, \ldots, g_{q-1}]$ will contribute exactly

$$\frac{1}{(2h_k+2)!} \prod_{j=1}^{k-1} \frac{1}{(2h_j - h_{j+1} + 1)!}$$

to $f(q)$. When $q \ge 3$ we shall write $S(q)$ for the total contribution from all splitting sequences in which $h_k = 0$, so that

$$S(q) = \frac{1}{2} \sum_{k=2}^{q-1} \sum_{h_1 + \ldots + h_{k-1} = q-1-k} \prod_{j=1}^{k-1} \frac{1}{(2h_j - h_{j+1} + 1)!}.$$

Here we adopt the natural convention that $1/m! = 0$ if m is a negative integer. Moreover we take $h_1 = h_k = 0$ and $h_j \ge 0$ for $2 \le j \le k-1$. Alternatively we may write $S(q)$ in terms of the function $R(q,a,b)$ defined in the previous section. This produces

$$S(q) = \frac{1}{q!} \sum_{a \ge 0} R(q, a, 2). \tag{56}$$

Similarly, when $q \ge 3$ we shall write $T(q)$ for the contribution to $f(q)$ corresponding to splitting sequences for which $h_k > 0$, so that

$$f(q) = S(q) + T(q). \tag{57}$$

The definition of $T(q)$ gives us the formula

$$T(q) = \sum_{k=2}^{q-1} \sum_{h_k>0} \frac{1}{(2h_k+2)!} \sum_{h_1+\ldots+h_k=q-1-k} \prod_{j=1}^{k-1} \frac{1}{(2h_j - h_{j+1}+1)!}.$$

Here we take $h_1 = 0$. We now assume that $r \geq 4$ and put $k = l-1$ and $q = r-1$ in the above expression, whence

$$T(r-1) = \sum_{l=3}^{r-1} \sum_{\substack{h_1+\ldots+h_{l-1}=r-1-l \\ h_{l-1}>0}} \frac{1}{(2h_{l-1}+2)!} \prod_{j=1}^{l-2} \frac{1}{(2h_j - h_{j+1}+1)!}.$$

Since $2h_{l-1}+2 \geq 4$ for $h_{l-1} \geq 1$ it follows that

$$\begin{aligned}
T(r-1) &\leq \frac{1}{4} \sum_{l=3}^{r-1} \sum_{\substack{h_1+\ldots+h_{l-1}=r-1-l \\ h_{l-1}>0}} \frac{1}{(2h_{l-1}+1)!} \prod_{j=1}^{l-2} \frac{1}{(2h_j - h_{j+1}+1)!} \\
&\leq \frac{1}{4} \sum_{l=2}^{r-1} \sum_{h_1+\ldots+h_{l-1}=r-1-l} \frac{1}{(2h_{l-1}+1)!} \prod_{j=1}^{l-2} \frac{1}{(2h_j - h_{j+1}+1)!} \\
&= \frac{1}{4} \sum_{l=2}^{r-1} \sum_{h_1+\ldots+h_{l-1}=r-1-l} \prod_{j=1}^{l-1} \frac{1}{(2h_j - h_{j+1}+1)!}
\end{aligned}$$

where we now assume that $h_1 = h_l = 0$. It therefore follows that

$$T(r-1) \leq \frac{1}{2} S(r), \quad r \geq 4. \tag{58}$$

We shall now estimate $R(q,a,2)$ from above, so that $f(q)$ can be bounded using (57), (58) and (56). We begin by noting that $R(q,a,b) = 0$ unless b and $a + b/2$ are even. Now (54) becomes

$$R(q, 2h+1, 2) = q(h+1) \sum_{c=0}^{q-2h-3} R(q-1, c, 2h+2),$$

while repeated application of (55) yields

$$R(q-1, c, 2h+2) = \frac{(q-1)!}{(q-1-h)!} \frac{(c+h)!}{c!} \frac{2!}{(2h+2)!} R(q-1-h, c+h, 2).$$

If we now set

$$s(q,h) = \frac{(2h+1)!}{q!} R(q, 2h+1, 2) \tag{59}$$

we deduce that

$$\begin{aligned}
s(q,h) &= \sum_{2|c+h-1} \frac{1}{c!} s(q-1-h, \frac{c+h-1}{2}) \\
&= \sum_{l\geq 0} \frac{1}{(2l-h+1)!} s(q-1-h, l), \qquad (60)
\end{aligned}$$

with the usual convention that $1/m! = 0$ for $m < 0$.

We proceed to prove an estimate of the form

$$s(q,h) \le C\mu^q\nu^h, \quad (0 \le h \le q), \tag{61}$$

for all $q \ge 20$, using induction on q. We note at once that $R(q,a,b) = 0$ unless $a+b \le q$, so that $s(q,h) = 0$ unless $h \le (q-3)/2$. Now suppose that $k \ge 39$, and that (61) holds for $20 \le q < k$. We shall prove (61) for $s(k,h)$, and it clearly suffices to suppose that $h \le (k-3)/2$, whence $k-1-h \ge (k+1)/2 \ge 20$. Thus, by our induction assumption, the recursion formula (60) yields

$$\begin{aligned} s(k,h) &\le \sum_{l\ge 0} \frac{1}{(2l-h+1)!} C\mu^{k-1-h}\nu^l \\ &= C\mu^k\nu^h \sum_{l\ge 0} \frac{\mu^{-1-h}\nu^{l-h}}{(2l-h+1)!}. \end{aligned}$$

To estimate the final sum we consider separately the cases $h = 2t$ and $h = 2t+1$. When $h = 2t$ we have

$$\frac{\mu^{-1-2t}\nu^{l-2t}}{(2l-2t+1)!} = 0$$

unless $l \ge t$, and putting $l = t+j$ we obtain

$$\sum_{l\ge 0} \frac{\mu^{-1-2t}\nu^{l-2t}}{(2l-2t+1)!} = \mu^{-1-2t}\nu^{-t} \sum_{j\ge 0} \frac{\nu^j}{(2j+1)!} = \mu^{-1-2t}\nu^{-t}\nu^{-1/2}\sinh \nu^{1/2}.$$

Thus (61) holds for $s(k,2t)$ providing that

$$\nu^{-1/2}\sinh \nu^{1/2} \le \mu \quad \text{and} \quad \mu^2\nu \ge 1.$$

Similarly for $h = 2t+1$ we find that only terms with $l \ge t$ can contribute, whence

$$\sum_{l\ge 0} \frac{\mu^{-1-h}\nu^{l-h}}{(2l-h+1)!} = \mu^{-2-2t}\nu^{-t-1}\cosh \nu^{1/2}.$$

Thus (61) holds for $s(k,2t+1)$ providing that

$$\nu^{-1}\cosh \nu^{1/2} \le \mu^2 \quad \text{and} \quad \mu^2\nu \ge 1.$$

We now choose ν so that

$$\cosh \nu^{1/2} = (\sinh \nu^{1/2})^2,$$

by taking $\nu = 1.126304\ldots$. This then allows us to use

$$\mu = \frac{\sinh \nu^{1/2}}{\nu^{1/2}} = 1.198576\ldots$$

We may now conclude as follows, using (56) and (59).

Lemma 14 *Let $\mu = 1.198576\ldots$ and $\nu = 1.126304\ldots$. Suppose that (61) holds for $20 \le q \le 38$. Then (61) holds for all $q \ge 20$, and*

$$S(q) \le C\nu^{-1/2}(\sinh \nu^{1/2})\mu^q = C\mu^{q+1}$$

for $q \ge 20$.

We therefore compute $s(q,h)$ for $h \le q$ and $20 \le q \le 38$, using the recursion (60). In this range we discover that

$$C = 0.172345$$

is admissible in (61).

It now follows from Lemma 14 taken in conjunction with (57) and (58) that

$$f(q) \le C(\mu + \frac{\mu^2}{2})\mu^q \le C'\mu^q, \quad C' = 0.330363,$$

for $q \ge 20$. We shall use this upper bound for $q > 20$, together with the numerical values in Table 1. We then find that the value of ρ, as given by Lemma 13, satisfies $\rho > \mu$ and

$$1 \le \sum_{q=1}^{20} f(q)\rho^{-q} + C' \sum_{q=21}^{\infty} (\mu\rho^{-1})^q = \sum_{q=1}^{20} f(q)\rho^{-q} + C' \frac{\mu^{21}}{\rho^{20}(\rho-\mu)}.$$

However the polynomial

$$X^{20}(X-\mu) - \sum_{q=1}^{20} f(q) X^{20-q} - C'\mu^{21}$$

has $X = 1.752\ldots$ as its only root greater than μ. We therefore conclude that

$$\rho \le 1.753$$

which suffices for Lemma 1. Finally we remark that there is little to be gained from pushing the precise evaluation of $f(q)$ further than $q = 20$, since a lower bound for ρ is given by the positive root of the polynomial

$$X^{20} - \sum_{q=1}^{21} f(q) X^{20-q},$$

which one can compute as $1.7519\ldots$.

9 An Improvement to Theorem 1

In this section we sketch a variant of our argument which leads to a sharpened version of Theorem 1. The same variant may be used under the Generalized Riemann Hypothesis, but leads to a result involving 10 powers of 2.

We replace the bound (43) by

$$|\int_{\mathfrak{m}} S(\alpha)^2 T(\alpha)^K e(-\alpha N) d\alpha| \le (\sup_{\alpha\in\mathfrak{m}} |S(\alpha)|)^{2\mu} I(q)^{\mu} J^*(\mathfrak{m})^{\nu},$$

where

$$J^*(\mathfrak{m}) = \int_{\mathfrak{m}} |S(\alpha)|^2 |T(\alpha)|^4 d\alpha)$$

and

$$\mu = \frac{K-4}{2q-4}, \quad \nu = \frac{2q-K}{2q-4}.$$

The argument of §5 must then be adapted so as to give an upper bound for $J^*(\mathfrak{m})$. In analogy to our previous work we first consider

$$J^* = \int_0^1 |S(\alpha)|^2 |T(\alpha)|^4 d\alpha) = \sum_{a,b,c,d\le L} r(2^a + 2^b - 2^c - 2^d).$$

When $2^a + 2^b \neq 2^c + 2^d$ the analysis depends upon the behaviour of the sum

$$\sum_{\substack{a,b,c,d\le J \\ 2^a+2^b\neq 2^c+2^d}} h(|2^a + 2^b - 2^c - 2^d|),$$

for which one may establish an asymptotic formula $(C_3 + o(1))J^4$. Here

$$C_3 = \sum_{d=1}^{\infty} k(d)N(d)\varepsilon(d)^{-3} \le 1.683,$$

where

$$N(d) = \#\{(a,b,c) : 1 \le a,b,c \le \varepsilon(d),\ 2^a + 2^b \equiv 2^c + 1 \pmod{d}\}.$$

However the key point is that terms for which $2^a+2^b = 2^c+2^d$ make a negligible contribution to J^*. In this way we find that

$$J^* \le \{\frac{C_0 C_1 C_3}{\log^2 2} + o(1)\}NL^2.$$

The corresponding calculation of a lower bound for

$$J^*(\mathfrak{M}) = \int_{\mathfrak{M}} |S(\alpha)|^2 |T(\alpha)|^4 d\alpha$$

proceeds much as in §5 and produces

$$J^*(\mathfrak{M}) \ge \{\frac{2(1-\varpi)C_0 C_3}{\log^2 2} + o(1)\}NL^2,$$

and hence

$$J^*(\mathfrak{m}) \le \{\frac{C_0(C_1 - 2 + 2\varpi)C_3}{\log^2 2} + o(1)\}NL^2 \le \{9.819 + o(1)\}C_0 \frac{N}{\log^2 2} L^2.$$

If this estimate is injected into the method one then finds that one requires

$$9.819\{(2\theta - 1)1.753 \log 2\}^{(K-4)/2} < 2.7973,$$

which one should compare with (45). Since the above inequality certainly holds with $\theta = 31/36$ and $K \ge 21$, we are lead to the improved version of Theorem 1.

References

[1] R.C. Baker and G. Harman, Diophantine approximation by prime numbers, *J. London Math. Soc. (2),* 25 (1982), 201-215.

[2] J.-R. Chen, On Goldbach's problem and the sieve methods, *Scientia Sin.,* 21 (1978), 701-739.

[3] P.X. Gallagher, Primes and powers of 2, *Invent. Math.,* 29 (1975), 125-142.

[4] D.R. Heath-Brown, The density of zeros of Dirichlet's L-functions, *Can. J. Math.,* 31 (1979), 231-240.

[5] M.N. Huxley, Large values of Dirichlet polynomials. III, *Acta Arith.,* 26 (1974/75), 435-444.

[6] H. Iwaniec, On zeros of Dirichlet's L-series, *Invent. Math.,* 23 (1974), 97-104.

[7] M. Jutila, On Linnik's constant, *Math. Scand.,* 41 (1977), 45-62.

[8] H.-Z. Li, The number of powers of 2 in a representation of large even integers by sums of such powers and of two primes, *Acta Arith.,* 92 (2000), 229-237.

[9] H.-Z. Li, The number of powers of 2 in a representation of large even integers by sums of such powers and of two primes, II, *Acta Arith.,* 96 (2001), 369-379.

[10] Yu. V. Linnik, Addition of prime numbers with powers of one and the same number. (Russian) Mat. Sbornik N.S., 32(74), (1953), 3-60.

[11] J.-Y. Liu, M.-C. Liu and T.-Z. Wang, The number of powers of 2 in a representation of large even numbers (I), *Sci. China Ser. A,* 41 (1998), 386-398.

[12] J.-Y. Liu, M.-C. Liu and T.-Z. Wang, The number of powers of 2 in a representation of large even numbers (II), *Sci. China Ser. A,* 41 (1998), 1255-1271.

[13] J.-Y. Liu, M.-C Liu and T.-Z. Wang, On the almost Goldbach problem of Linnik, *J. Théor. Nombres Bordeaux,* 11 (1999), 133-147.

[14] H.L. Montgomery, *Topics in multiplicative number theory,* Lecture Notes in Mathematics, Vol. 227. (Springer-Verlag, Berlin-New York, 1971).

[15] H.L. Montgomery and R.C. Vaughan, The exceptional set in Goldbach's problem, *Acta Arith.,* 27 (1975), 353-370.

[16] E.C. Titchmarsh, *The theory of the Riemann Zeta-function,* 2nd Edition, revised by D.R. Heath-Brown. (Oxford University Press, Oxford, 1986).

[17] R.C. Vaughan, *The Hardy-Littlewood method*, Cambridge tracts in mathematics, 125 (Cambridge University Press, Cambridge, 1981).

[18] T.-Z. Wang, On Linnik's almost Goldbach theorem, *Sci. China Ser. A,* 42 (1999), 1155-1172.