On Linnik's approximation to Goldbach's problem, I

by

J. Pintz and I. Z. Ruzsa (Budapest)

1. Introduction. As an approximation to Goldbach's problem Linnik proved in 1951 [Lin1] under the assumption of the Generalized Riemann Hypothesis (GRH), and later in 1953 [Lin2] unconditionally, that every sufficiently large even integer can be written as the sum of two primes and K powers of 2, where K is an unspecified but absolute constant. The first explicit value for K, $K = 770$, was proved in 1998 by J. Y. Liu, M.-C. Liu and T. Z. Wang [LLW1], supposing GRH. The first explicit unconditional result $K = 54\,000$ is due to the same authors [LLW2] and it was shown also in 1998. The subsequent unconditional results were the following:

$$\begin{aligned}
K &= 25\,000 & \text{(Li [Li1])}, \\
K &= 2\,250 & \text{(Wang [Wan])}, \\
K &= 1\,906 & \text{(Li [Li2])}.
\end{aligned}$$

Assuming the Generalized Riemann Hypothesis (GRH) the corresponding improvements were as follows:

$$\begin{aligned}
\text{GRH implies } K &= 200 & \text{(Liu, Liu, Wang [LLW3])}, \\
\text{GRH implies } K &= 160 & \text{(Wang [Wan])}.
\end{aligned}$$

We have to mention yet an important work of Gallagher [Gal] who did not give any explicit value for K, but significantly simplified the original proofs of Linnik and in this way opened the way for the later works mentioned above, which provided explicit values of K.

In the present work we will investigate the distribution of values of

$$(1.1) \quad r''_K(N) = \#\{(p_1, p_2, \nu_1, \ldots, \nu_K) : N = p_1 + p_2 + 2^{\nu_1} + \ldots + 2^{\nu_K}\}$$

for $2 \mid N$ and

$$(1.2) \quad r'_K(N) = \#\{(p, \nu_1, \ldots, \nu_K) : N = p + 2^{\nu_1} + \ldots + 2^{\nu_K}\}$$

for $2 \nmid N$, where p_i and p will always denote odd (positive) primes. We will prove the following result.

THEOREM 1. *Suppose GRH. Let K be a fixed natural number, $K \geq 7$. Then*

(1.3) $r_K''(N) > 0$ *if* $2 \mid N$, $N > N_0(K)$

where $N_0(K)$ is an explicit constant, depending on K.

We will show in Part II that the same result also holds unconditionally, under the slightly weaker assumption $K \geq 8$.

We mention that independently of us the same result $K = 7$ under GRH, and the value $K = 13$ unconditionally were proved by D. R. Heath-Brown and J. C. Puchta [HP].

2. Results about primes. In this section we will summarize some results about the Goldbach problem, which are mostly proved or referred to in [LLW3]. For convenience we will use the same notations as in [LLW3].

Let N be a large integer, and P, Q parameters satisfying

(2.1) $2 \leq P < Q \leq N.$

By Dirichlet's lemma any $\alpha \in [1/Q, 1 + 1/Q]$ can be written as

(2.2) $\alpha = \dfrac{a}{q} + \lambda, \quad |\lambda| \leq \dfrac{1}{qQ},$

with positive integers a, q with $1 \leq a \leq q$, $(a, q) = 1$. Denote by $\mathcal{M}(a, q)$ the set of α's satisfying (2.2) and let

(2.3) $\mathcal{M} = \bigcup_{\substack{q \leq P}} \bigcup_{\substack{a=1 \\ (a,q)=1}}^{q} \mathcal{M}(a, q), \quad C(\mathcal{M}) = [1/Q, 1 + 1/Q] \setminus \mathcal{M}.$

Let

(2.4) $e(\alpha) = e^{2\pi i \alpha}, \quad S(\alpha) = S(\alpha, N) = \sum_{p \leq N} e(p\alpha), \quad L = [\log_2 N],$

where $\log_2 N$ denotes the logarithm to base 2.

On the major arcs we will use the following estimate.

LEMMA 1. *Assume GRH. Let $h \neq 0$ be an even number. Choose*

(2.5) $P = \sqrt{N}\, L^{-8}, \quad Q = \sqrt{N}.$

Then

(2.6) $\int_{\mathcal{M}} |S^2(\alpha)| e(h\alpha) \, d\alpha = \sigma(h) \dfrac{\max(N - |h|, 0)}{\log^2 N} + O(NL^{-3}),$

where

$$(2.7) \quad \sigma(h) = 2C_0 \prod_{\substack{p|h \\ p>2}} \left(\frac{p-1}{p-2} \right), \quad C_0 = \prod_{p>2} \left(1 - \frac{1}{(p-1)^2} \right) = 0.6601\ldots$$

We remark that for $|h| \ll N/L$ we can write (2.6) as

$$(2.8) \qquad \int_{\mathcal{M}} |S^2(\alpha)| e(h\alpha) \, d\alpha = \sigma(h) \frac{N}{\log^2 N} \left(1 + O\left(\frac{1}{L} \right) \right).$$

The proof of Lemma 1 is, apart from slight changes, the same as that of Theorem 2 of [LLW3] where the analogous result is shown for $S^2(\alpha)$ in place of $|S(\alpha)|^2$.

On the minor arcs the following estimate (see [BH, Lemma 12])—basically due to Hardy–Littlewood, [HL, Lemma 9]—will be needed.

LEMMA 2. *Assume GRH. If P and Q satisfy (2.1) then for $\alpha \in C(\mathcal{M})$ we have*

$$(2.9) \qquad S(\alpha) \ll \left(\frac{N}{P} + \sqrt{NQ} + \frac{N}{\sqrt{Q}} \right) L^2.$$

With the choice of P and Q in (2.5) we obtain

$$(2.10) \qquad S(\alpha) \ll N^{3/4} L^2.$$

The third important theorem we include here deals with the closely connected generalized twin-prime problem. The following result of Chen [Che, Theorem 3] was proved by sieve methods and it holds unconditionally, unlike Lemmas 1 and 2.

LEMMA 3. *Let h be any even integer, and N sufficiently large. Then the number of solutions of the equation $h = p_1 - p_2$ with $p_j \leq N$ is*

$$(2.11) \quad R(h) = \#\{h = p_1 - p_2 : p_j \leq N, \, j = 1, 2\} < C^* \cdot 2C_0 \, f(h) \frac{N}{\log^2 N}$$

where h is a positive even integer and

$$(2.12) \qquad f(h) = \prod_{\substack{p|h \\ p>2}} \left(1 + \frac{1}{p-2} \right), \quad C^* = 3.9171.$$

We note here that Hardy–Littlewood conjectured that if h is much smaller than N, then

$$(2.13) \qquad R(h) \sim 2C_0 f(h) \frac{N}{\log^2 N} \quad \text{as } N \to \infty.$$

Thus the factor C^* of Chen represents the loss compared to the conjectured value (at least for $h = o(N)$).

3. Estimates for large deviations of certain exponential sums.
We shall consider the problem of finding estimates for the measure of the
set of numbers $x \in [0, 1]$ for which an inequality of type

$$\left| \sum_{j=1}^{L} e(g_j x) \right| > \lambda L$$

holds. Our primary interests lie in the case $g_j = g^j$, especially for $g = 2$.

These estimates will be deduced from estimates of certain exponential
moments, as is often done in probability theory.

Our main interest lies in finding upper estimates, which we need for the
work on primes.

4. Exponential moments of $\sum e(g^j x)$. We assume that g is a fixed
positive integer, and put

$$G_L(x) = \sum_{j=0}^{L-1} e(g^j x).$$

We shall estimate exponential moments of $|G_L(x)|$ and of the related sum

$$S_L(x) = \operatorname{Re} G_L(x) = \sum_{j=0}^{L-1} \cos(2\pi g^j x).$$

THEOREM 2. (a) *For every real λ there exists a real $\psi(\lambda)$ such that* (*as*
$L \to \infty$ *with λ fixed*)

(4.1) $$\int_0^1 e^{\lambda S_L(x)} \, dx = e^{L(\psi(\lambda)+o(1))}.$$

(b) *If $\lambda \geq 0$, then for every L we have*

(4.2) $$\int_0^1 e^{\lambda S_L(x)} \, dx \leq e^{L\psi(\lambda)} \leq g^{-L} \sum_{k=0}^{g^L-1} \exp(\lambda S_L(kg^{-L})).$$

(c) *If $\lambda \geq 0$, then*

(4.3) $$\int_0^1 e^{\lambda |G_L(x)|} \, dx = e^{L(\psi(\lambda)+o(1))}.$$

Inequality (4.2) forms the basis of further calculations. Indeed, to find
an upper estimate of $\psi(\lambda)$, it is sufficient to apply (4.2) for some L, and
initially we did this, for $g = 2$ and $L = 20$. However, the convergence of this
process is very slow; the error is of size $O(1/L)$, so to get n digits we need
$\exp \exp n$ operations. In Sections 5–7 we describe a more complicated but
much faster procedure, in fact polynomial in n. Instead of fewer than 2 exact

digits with a direct application of (4.2), we calculated 15 (and could have calculated more if necessary); with the low precision available by the direct approach we could establish only weaker estimates for the K in Theorem 1.

It seems difficult (and quite irrelevant for our purposes) to estimate the left side of (4.3) for negative λ. It is easy to see that (4.3) fails for $\lambda < 0$.

Proof of Theorem 2. (a) We consider λ to be fixed. Write

$$E_L = \int_0^1 e^{\lambda S_L(x)} \, dx.$$

Take now two positive integers $M < L$. (M will be fixed as $L \to \infty$.) We have

$$S_L(x) = S_M(x) + S_{L-M}(g^M x),$$

and hence

$$(4.4) \qquad E_L = \sum_{k=1}^{g^M} \int_{(k-1)/g^M}^{k/g^M} e^{\lambda S_M(x)} e^{\lambda S_{L-M}(g^M x)} \, dx.$$

Since $S_{L-M}(g^M x)$ is periodic with period g^{-M}, replacing $g^M x$ by x we can rewrite (4.4) as

$$(4.5) \qquad E_L = g^{-M} \int_0^1 e^{\lambda S_{L-M}(x)} \sum_{k=1}^{g^M} e^{\lambda S_M((x+k)g^{-M})} \, dx.$$

Define

$$\overline{E}_M = \max_{0 \le x \le 1} g^{-M} \sum_{k=1}^{g^M} e^{\lambda S_M((x+k)g^{-M})},$$

$$\underline{E}_M = \min_{0 \le x \le 1} g^{-M} \sum_{k=1}^{g^M} e^{\lambda S_M((x+k)g^{-M})}.$$

Then (4.5) implies

$$\underline{E}_M E_{L-M} \le E_L \le \overline{E}_M E_{L-M}.$$

By induction we get

$$E_L \le \overline{E}_M^{[L/M]} E_{L_0}, \qquad E_L \ge \underline{E}_M^{[L/M]} E_{L_0},$$

where L_0 is the residue of L modulo M. Hence

$$(4.6) \qquad \limsup_{L \to \infty} \frac{\log E_L}{L} \le \frac{\log \overline{E}_M}{M},$$

$$(4.7) \qquad \liminf_{L \to \infty} \frac{\log E_L}{L} \ge \frac{\log \underline{E}_M}{M}$$

for every fixed M. Since

$$S'_M(x) = - \sum_{j=0}^{M-1} 2\pi g^j \sin(2\pi g^j x),$$

we have

$$|S'_M(x)| \leq 2\pi \sum_{j=0}^{M-1} g^j < \frac{2\pi}{g-1} g^M \quad \text{for all } x,$$

and consequently

$$\left| S_M\left(\frac{x+k}{g^M}\right) - S_M\left(\frac{k}{g^M}\right) \right| \leq \frac{2\pi}{g-1} \quad \text{for } |x| \leq 1.$$

Thus with

$$E_M^* = g^{-M} \sum_{k=0}^{g^M-1} e^{\lambda S_M(kg^{-M})}$$

we have

(4.8) $$\overline{E}_M \leq e^{2\pi\lambda/(g-1)} E_M^*,$$

(4.9) $$\underline{E}_M \geq e^{-2\pi\lambda/(g-1)} E_M^*.$$

This shows that the bounds on the right sides of (4.6) and (4.7) differ by at most $O(1/M)$, hence the upper and lower limits are equal and the existence of $\psi(\lambda)$ is proved.

To prove part (b), we show that for $\lambda \geq 0$, the bound (4.8) can be improved to

(4.10) $$\overline{E}_M = E_M^*.$$

In order to show this observe that the expansion of S_M has nonnegative coefficients, hence so does the expansion of S_M^n for every positive integer n, and by the power-series expansion

$$e^{\lambda S_M} = \sum \frac{\lambda^n}{n!} S_M^n$$

we see that so does $e^{\lambda S_M}$, provided $\lambda \geq 0$. Thus we have

(4.11) $$e^{\lambda S_M(x)} = \sum a_n \cos(2\pi n x)$$

with coefficients $a_n \geq 0$, whence

$$g^{-M} \sum_{k=1}^{g^M} e^{\lambda S_M((x+k)g^{-M})} = g^{-M} \sum a_{ng^M} \cos(2\pi n x).$$

This function assumes its maximum at $x = 0$, and this is precisely the content of (4.10). This suffices to show the second inequality of (b).

To prove the first inequality we cannot proceed analogously via $\underline{E}_M \geq E_M$, since this is clearly false. Instead, we shall show directly that

$$(4.12) \qquad E_L \geq E_M E_{L-M}.$$

Indeed, similarly to (4.11), we have

$$e^{\lambda S_{L-M}(g^M x)} = \sum b_n \cos(2\pi nx)$$

with coefficient $b_n \geq 0$. Hence

$$E_L = \int\limits_0^1 \left(\sum a_n \cos(2\pi nx) \right) \left(\sum b_n \cos(2\pi nx) \right) dx$$

$$= a_0 b_0 + \frac{1}{2} \sum_{n=1}^{\infty} a_n b_n \geq a_0 b_0 = E_M E_{L-M}.$$

The supermultiplicativity property (4.12) clearly implies

$$\lim \frac{\log E_L}{L} = \sup \frac{\log E_L}{L},$$

which is equivalent to the first inequality of (b).

Now we prove part (c). Since clearly

$$\int\limits_0^1 e^{\lambda(G_L(x))} \, dx \geq \int\limits_0^1 e^{\lambda S_L(x)} \, dx,$$

it suffices to find an inequality in the other direction. This will be

$$(4.13) \qquad \int\limits_0^1 e^{\lambda(G_L(x))} \, dx \leq c(1 + \sqrt{\lambda L}) \int\limits_0^1 e^{\lambda S_L(x)} \, dx \quad \text{ for } \lambda \geq 0,$$

and it will be proved in the following more general form:

LEMMA 4. *Let a_n be complex numbers such that $A = \sum_{n=-\infty}^{\infty} |a_n| < \infty$. Let*

$$G(x) = \sum_n a_n e(nx), \qquad S(x) = \sum_n |a_n| \cos(2\pi nx).$$

We have

$$(4.14) \qquad \int\limits_0^1 e^{|G(x)|} \, dx \leq c(1 + \sqrt{A}) \int\limits_0^1 e^{S(x)} \, dx$$

with an absolute constant c.

(4.13) follows by putting $G(x) = \lambda G_L(x)$.

Proof. We have

$$e^{|G(x)|} \leq e^{|G(x)|} + e^{-|G(x)|} = 2 \sum_{j=0}^{\infty} \frac{|G(x)|^{2j}}{(2j)!},$$

and so

$$(4.15) \qquad \int_0^1 e^{|G(x)|} \, dx \le \sum_{j=0}^\infty \frac{2}{(2j)!} \int_0^1 |G(x)|^{2j} \, dx.$$

Write

$$G_0(x) = \sum_{n=-\infty}^\infty |a_n| e(nx).$$

Clearly

$$\int_0^1 |G(x)|^{2j} \, dx = \sum_{m_1+\ldots+m_j=n_1+\ldots+n_j} a_{m_1} \ldots a_{m_j} \overline{a}_{n_1} \ldots \overline{a}_{n_j}$$

$$\le \sum |a_{m_1} \ldots a_{m_j} a_{n_1} \ldots a_{n_j}| = \int_0^1 (G_0(x))^{2j} \, dx.$$

Furthermore $S(x) = \operatorname{Re} G_0(x) = (G_0(x) + \overline{G_0(x)})/2$, so

$$\int_0^1 S(x)^{2j} \, dx = 2^{-2j} \sum_{\nu=0}^{2j} \binom{2j}{\nu} \int_0^1 G_0(x)^\nu \overline{G_0(x)}^{2j-\nu} \, dx.$$

Each integral is nonnegative, being a positive polynomial in the coefficients, and retaining only the middle term we obtain

$$\int_0^1 S(x)^{2j} \, dx \ge 2^{-2j} \binom{2j}{j} \int_0^1 |G_0(x)|^{2j} \, dx,$$

or

$$(4.16) \qquad \int_0^1 |G_0(x)|^{2j} \, dx \le c\sqrt{j+1} \int_0^1 S(x)^{2j} \, dx$$

with a certain absolute constant c.

We split the sum in (4.15) into two parts. For $j < K$ we apply (4.16), and for $j \ge K$ we use the trivial estimate $|G(x)| \le A$. This gives us

$$(4.17) \qquad \int_0^1 e^{|G(x)|} \, dx \le c_1 \sum_{j=0}^{K-1} \frac{\sqrt{j+1}}{(2j)!} \int_0^1 S(x)^{2j} \, dx + \sum_{j=K}^\infty \frac{2A^{2j}}{(2j)!}.$$

We estimate the first term as follows:

$$\le c_1 \sqrt{K} \sum_{j=0}^{K-1} \frac{1}{(2j)!} \int_0^1 S(x)^{2j} \, dx \le c_1 \sqrt{K} \sum_{j=0}^\infty \frac{1}{j!} \int_0^1 S(x)^j \, dx = c_1 \sqrt{K} \int_0^1 e^{S(x)} \, dx,$$

since $\int_0^1 S(x)^j \, dx \ge 0$ for every j.

The second term of (4.17) is

$$\frac{2A^{2K}}{(2K)!}\left(1 + \frac{A^2}{(2K+1)(2K+2)} + \ldots\right) \le \frac{4A^{2K}}{(2K)!},$$

if K is so large that $A^2/((2K+1)(2K+2)) \le 1/2$, which will hold if $K \ge A$. Hence

$$\int\limits_0^1 e^{|G(x)|}\,dx \le c_1\sqrt{K}\int\limits_0^1 e^{S(x)}\,dx + \frac{4A^{2K}}{(2K)!}.$$

The choice of $K = 2[A] + 1$ makes the second term bounded by an absolute constant independent of A and we are done.

5. A recurrence for certain moments. In this section we show how to estimate

$$\int\limits_0^{2\pi} \prod_{j=0}^{L-1} \varphi(g^j x)\,dx$$

and the related finite sum for trigonometric polynomials

(5.1) $$\varphi(x) = b_0 + b_1\cos x + \ldots + b_k\cos kx.$$

Since for given λ and ε we can find $\underline{\varphi}$ and $\overline{\varphi}$ such that

(5.2) $$0 < \underline{\varphi}(x) \le e^{\lambda\cos x} \le \overline{\varphi}(x), \quad \overline{\varphi}(x)/\underline{\varphi}(x) \le 1 + \varepsilon$$

for all x, this can be used to estimate our exponential moments. A choice of $\underline{\varphi}$ and $\overline{\varphi}$ will be described in the next section.

We consider the integral and the finite sum jointly. In what follows let μ_n be one of the following two sequences of measures on $[0, 2\pi]$:

(a) $\mu_n^{(1)} = \mu = $ the Lebesgue measure, normed by the factor $1/2\pi$ on $[0, 2\pi]$, independently of n;

(b) $\mu_n^{(2)}$ is the discrete measure supported at the points $(j/g^n)2\pi$, $j = 0, \ldots, g^n - 1$, each having weight g^{-n}.

Write

$$v_{n,L} = \int\limits_0^{2\pi} \cos nx \prod_{j=0}^{L-1} \varphi(g^j x)\,d\mu_L(x).$$

LEMMA 5. *We have*

$$v_{n,L+1} = \frac{1}{2}\left(\sum_{\substack{j=0 \\ g|j+n}}^k b_j v_{(j+n)/g,L} + \sum_{\substack{j=0 \\ g|j-n}}^k b_j v_{|j-n|/g,L}\right).$$

Proof. Note that

$$v_{n,L+1} = \int\limits_{0}^{2\pi} \underbrace{(\varphi(x)\cos nx)}_{h(x)} \prod_{j=1}^{L} \varphi(g^j x)\, d\mu_{L+1}(x).$$

As the product has period $2\pi/g$, we can rewrite this as

$$v_{n,L+1} = \int\limits_{0}^{2\pi/g} \prod_{j=1}^{L} \varphi(g^j x) \sum_{\nu=0}^{g-1} h\!\left(x + \frac{2\pi\nu}{g}\right) d\mu_{L+1}(x).$$

Now

$$h(x) = \sum_{j=0}^{k} b_j \cos jx \cos nx = \frac{1}{2} \sum_{j=0}^{k} b_j(\cos(j+n)x + \cos(j-n)x),$$

hence

$$\sum_{\nu=0}^{g-1} h\!\left(x + \frac{2\pi\nu}{g}\right) = \frac{g}{2}\left(\sum_{g|j+n} b_j \cos(j+n)x + \sum_{g|j-n} b_j \cos(j-n)x \right)$$

$$= gh^*(gx)$$

with

$$h^*(x) = \frac{1}{2}\left(\sum_{\substack{j=0 \\ g|j+n}}^{k} b_j \cos\frac{j+n}{g}x + \sum_{\substack{j=0 \\ g|j-n}}^{k} b_j \cos\frac{j-n}{g}x \right).$$

So

$$v_{n,L+1} = g \int\limits_{0}^{2\pi/g} \prod_{j=1}^{L} \varphi(g^j x) h^*(gx)\, d\mu_{L+1}(x).$$

If we change the variable x to x/g, the interval changes to $[0, 2\pi]$ and the measure $g\mu_{L+1}$ becomes μ_L in both cases. Hence

$$v_{n,L+1} = \int\limits_{0}^{2\pi} \prod_{j=0}^{L-1} \varphi(g^j x) h^*(x)\, d\mu_L(x)$$

$$= \frac{1}{2}\left(\sum_{\substack{j=0 \\ g|j+n}}^{k} b_j v_{(j+n)/g,L} + \sum_{\substack{j=0 \\ g||j-n|}}^{k} b_j v_{|j-n|/g,L} \right).$$

We are interested in $v_{0,L}$, and this recurrence involves all $v_{n,L}$. Fortunately we can restrict our attention to a few values of n. If we take m so that

$$\left[\frac{m+k}{g}\right] \le m,$$

then the combined recurrence for $v_{0,L}, \ldots, v_{m,L}$ will involve only the same values. The smallest such m is given by

$$m = \left[\frac{k-1}{g-1}\right].$$

So the above lemma implies the following:

LEMMA 6. *Write*

$$\underline{v}_L = (v_{0,L}, v_{1,L}, \ldots, v_{m,L}).$$

Then $\underline{v}_{L+1} = U\underline{v}_L$, where U is an $(m+1) \times (m+1)$ matrix, whose entries are given by

$$u_{nl} = \frac{1}{2}(b_{gl-n} + b_{gl+n} + b_{n-gl}), \quad 0 \leq n, l \leq m,$$

where we extend $b_j = 0$ for $j < 0$ and $j > k$.

We have the same recurrence for the continuous and the discrete measures. The initial values are, however, different. In the continuous case we have

$$v_{0,0} = 1, \quad v_{n,0} = 0 \quad \text{for } n \geq 1,$$

while in the discrete case

$$v_{n,0} = 1 \quad \text{for all } n.$$

Hence we have the following:

LEMMA 7. *With the matrix U as given above, and writing $U^L = (u_{nl}^{(L)})$, we have*

$$\frac{1}{2\pi} \int_0^{2\pi} \prod_{j=0}^{L-1} \varphi(g^j x)\, dx = u_{00}^{(L)},$$

$$g^{-L} \sum_{\nu=0}^{g^L-1} \prod_{j=0}^{L-1} \varphi\left(2\pi g^j \frac{\nu}{g^L}\right) = u_{00}^{(L)} + u_{01}^{(L)} + \ldots + u_{0m}^{(L)}.$$

The asymptotic behaviour of these sequences depends on the largest eigenvalue of U. In practice we rather calculated U^L for some large L, which can be done in $O(\log L)$ matrix multiplications, and used the above formulas.

6. Polynomial approximations to exponential functions. In this section we will describe those cosine polynomials $\underline{\varphi}, \overline{\varphi}$ of degree k, for a given λ, which satisfy

(6.1) $$\underline{\varphi}(x) \leq \gamma(x) = e^{\lambda \cos x} \leq \overline{\varphi}(x),$$

and which are nearest to $\gamma(x)$ in L_1 norm. The "best" polynomial for the calculations of the preceding section would optimize the largest eigenvalue of the described matrix U. We cannot determine these, but one can show

that our polynomials are also quite good. Indeed, it follows immediately
from Bernstein's inequality that $\|\underline{\varphi}(x) - \overline{\varphi}(x)\|_\infty \leq k\|\underline{\varphi}(x) - \overline{\varphi}(x)\|_1$, and
this quantity can be shown to be of order $c^k/k!$ with some c depending on λ.

A cosine polynomial can also be written as $\varphi(x) = f(\cos x)$ with an
algebraic polynomial f, so (6.1) is related to the description of polynomials
$\underline{f}$, $\overline{f}$, satisfying

(6.2) $$\underline{f}(x) \leq g(x) = e^{\lambda x} \leq \overline{f}(x).$$

From now on we think of λ as fixed, and reserve the notation $g(x) = e^{\lambda x}$.

Let f be a polynomial of degree $\leq k$. We say that f is an *extremal upper
approximation* of g if

$$f(x) \geq g(x) \qquad \text{for } |x| \leq 1,$$

and there is no polynomial $f^* \neq f$ of degree $\leq k$ such that

$$f(x) \geq f^*(x) \geq g(x).$$

Extremal lower approximations are defined analogously.

Let f be a polynomial, and let

$$-1 = a_0 < a_1 < \ldots < a_l = 1$$

be the sequence of roots of $f - g$ and the endpoints.

Let d_i be the order of a_i as a zero of $f - g$ (so d_0 and d_l may be 0).

LEMMA 8. (a) f is an extremal approximation if and only if $\sum_{i=0}^{l} d_i = k + 1$ and $2 \,|\, d_i$ for $1 \leq i \leq l - 1$.

(b) *For every upper (resp. lower) approximation f one can always find
an extremal upper (resp. lower) approximation f^* such that $f \geq f^* \geq g$
resp. $f \leq f^* \leq g$.*

(c) *An extremal approximation is upper if d_l is odd, and lower if d_l is
even.*

(d) *For a sequence a_i, d_i with $\sum d_i = k + 1$ there exists a unique f such
that $f - g$ has roots of order d_i at a_i.*

Proof. First observe that $f(x) - g(x)$ has at most $k + 1$ real roots alto-
gether (counted with multiplicities). Indeed,

$$(f(x) - g(x))^{(k+1)} = -g^{(k+1)}(x) = -\lambda^{k+1} e^{\lambda x}$$

has no root and the claim follows by calculus. In particular, $\sum d_i \leq k + 1$.
If $f \geq g$ and f is not extremal, then the inequality

$$f(x) \geq f^*(x) \geq g(x)$$

shows that $f^* - f$ has roots of order $\geq d_i$ at a_i. As this is a nonzero poly-
nomial of degree $\leq k$, we have $\sum d_i \leq k$. This shows the "if" part of (a)
for upper approximations; for lower approximations the proof works with
obvious modifications.

The condition $2 \mid d_i$ for $1 \le i \le l-1$ is clearly equivalent to the property that $f - g$ does not change sign, that is, f is either an upper or a lower approximation.

Now we prove the "only if" part of (a), and claim (b) together. Consider a polynomial $f \ge g$. We show that if $\sum d_i \le k$, then there is an f_1 such that $f \ge f_1 \ge g$ and the total number of roots of $f_1 - g$ is $\ge 1 + \sum d_i$. If we repeat this process, in at most $k + 1 - \sum d_i$ steps we stop at a function with value $k + 1$, which must be extremal by claim (a).

Consider the function

$$\frac{f(x) - g(x)}{\prod (x - a_i)^{d_i}}.$$

It has no root, so it is of constant sign.

Let c be its minimum if it is positive, its maximum if it is negative, and suppose it is assumed at a. Then

$$f_1(x) = f(x) - c \prod (x - a_i)^{d_i}$$

satisfies $f \ge f_1 \ge g$, and $f_1 - g$ has at least one more root than $f - g$, namely at a.

For claim (c), observe that $f(x) - g(x) \to -\infty$ as $x \to \infty$. As $f - g$ has already $k+1$ roots in $[-1, 1]$ and cannot have more, it has no root in $(1, \infty)$. Thus it is negative in $(1, \infty)$; if it is positive in $(-1, 1)$, it must change sign at 1, and if it is negative, it cannot.

To prove (d), observe that a polynomial satisfying $f^{(\nu)}(a_i) = g^{(\nu)}(a_i)$ for $0 \le \nu \le d_i$ can be constructed by Hermite's interpolation. The order of roots will then be exactly d_i, for otherwise we would have $> k + 1$ roots altogether (counted with multiplicities), contradicting our starting remark.

Now we consider $\gamma(x) = e^{\lambda \cos x}$ and a cosine polynomial $\varphi(x) = f(\cos x)$. Let

$$0 = \alpha_0 < \alpha_1 < \ldots < \alpha_l = \pi$$

be the sequence of roots of $\varphi - \gamma$ in $[0, \pi]$ and the endpoints. Assume that α_i has order δ_i. Then $f - g$ has roots at the points $a_i = \cos \alpha_i$ (which, however, are now in descending order), and the orders of the roots satisfy

$$\delta_i = \begin{cases} d_i & \text{for } 1 \le i \le l - 1, \\ 2d_i & \text{for } i = 0 \text{ or } i = l. \end{cases}$$

Hence Lemma 8 can be reformulated as follows:

LEMMA 9. (a) φ *is an extremal approximation to* γ *if and only if*

(6.3)
$$\sum_{i=1}^{l-1} \delta_i + \frac{\delta_0 + \delta_l}{2} = k + 1$$

and $2 \mid \delta_i$ *for* $1 \le i \le l - 1$.

(b) *For every upper (resp. lower) approximation φ one can always find an extremal upper (resp. lower) approximation φ^* such that*

$$\varphi \geq \varphi^* \geq \gamma \quad resp. \quad \varphi \leq \varphi^* \leq \gamma.$$

(c) *An extremal approximation is upper if $\delta_0 \equiv 2 \pmod 4$ and lower if $\delta_0 \equiv 0 \pmod 4$.*

(d) *For a sequence α_i, δ_i satisfying (6.3) there exists a unique φ such that $\gamma - \varphi$ has roots of order δ_i at α_i.*

Write

$$\varphi(x) = b_0 + b_1 \cos x + \ldots + b_k \cos kx.$$

Since for an upper or lower approximation φ we have

$$\|\varphi - \gamma\|_1 = \frac{1}{2\pi} \left| \int_0^{2\pi} \varphi - \gamma \right| = \left| b_0 - \frac{1}{2\pi} \int_0^{2\pi} \gamma \right|,$$

finding the best (in L_1 norm) upper and lower approximations is equivalent to minimizing (resp. maximizing) b_0 among the upper (resp. lower) approximations.

For any φ of degree k, any $K \geq k+1$ and any real t we have

$$b_0 = \frac{1}{K} \sum_{j=0}^{K-1} \varphi\left(t + \frac{2\pi j}{K}\right),$$

hence

(6.4)
$$b_0 \begin{array}{c} \geq \\ \leq \end{array} \frac{1}{K} \sum_{j=0}^{K-1} \gamma\left(t + \frac{2\pi j}{K}\right)$$

for upper (resp. lower) approximations.

If we find φ such that there is equality in (6.4) for some t, it must be the extremal function and we must have

$$\varphi\left(t + \frac{2\pi j}{K}\right) = \gamma\left(t + \frac{2\pi j}{K}\right) \quad \text{for all } j.$$

However, as the functions are even, the same equality holds for the numbers $-t + 2\pi j/K$. One can easily see that this means too many roots for $\varphi - \gamma$, so these must coincide, which happens only if $t \equiv 0 \pmod{2\pi/K}$ or $t \equiv \frac{\pi}{K} \pmod{2\pi/K}$, and even then the only possibility is $K = k + 1$. We shall see that these cases do actually occur and they yield the best upper and lower approximations.

THEOREM 3. *Let $K = k+1$, $t = 0$ or $t = \pi/K$. There exists a unique cosine polynomial $\varphi = \varphi_t$ of degree k satisfying*

(6.5) $$\varphi\left(t + \frac{2\pi j}{K}\right) = \gamma\left(t + \frac{2\pi j}{K}\right), \quad \varphi'\left(t + \frac{2\pi j}{K}\right) = \gamma'\left(t + \frac{2\pi j}{K}\right)$$

for all j. This polynomial is given by the formula

$$(6.6) \qquad \varphi(x) = \sigma_0 + 2 \sum_{j=1}^{k} \frac{(K-j)\sigma_j - \varrho_j}{K} \cos jx,$$

where

$$(6.7) \qquad \sigma_j = \frac{1}{K} \sum_{m=0}^{K-1} \gamma\left(t + \frac{2\pi m}{K}\right) \cos j\left(t + \frac{2\pi m}{K}\right),$$

$$(6.8) \qquad \varrho_j = \frac{1}{K} \sum_{m=0}^{K-1} \gamma'\left(t + \frac{2\pi m}{K}\right) \sin j\left(t + \frac{2\pi m}{K}\right).$$

$\overline{\varphi} = \varphi_0$ *is the best upper approximation, and* $\underline{\varphi} = \varphi_{\pi/k}$ *is the best lower approximation in* L_1 *norm.*

Proof. Since our functions are even, if equation (6.5) holds for $j \leq K/2$, it will hold for all j. (Here we use the fact that t is of the prescribed special form.) If $t = 0$ and $K = 2l$, this means roots of order 2 at the points

$$0, \frac{2\pi}{2l}, \frac{4\pi}{2l}, \dots, \frac{2l\pi}{2l};$$

if $t = \pi/K$, at

$$\frac{\pi}{2l}, \frac{3\pi}{2l}, \dots, \frac{(2l-1)\pi}{2l};$$

if $t = 0$ and $K = 2l + 1$, at

$$0, \frac{2\pi}{2l+1}, \frac{4\pi}{2l+1}, \dots, \frac{2l\pi}{2l+1};$$

if $t = \pi/K$, at

$$\frac{\pi}{2l+1}, \frac{3\pi}{2l+1}, \dots, \frac{(2l+1)\pi}{2l+1}.$$

In all four cases we can easily check that (6.3) holds, so existence and uniqueness follows from Lemma 9(d).

Now we find the coefficients in

$$\varphi(x) = \sum_{i=0}^{k} b_i \cos ix.$$

We have

$$(6.9) \qquad \varphi(x) \cos jx = \sum_{i=0}^{k} b_i \cos ix \cos jx$$

$$= \sum_{i=0}^{k} \frac{b_i}{2} (\cos(i+j)x + \cos(i-j)x),$$

(6.10) $\qquad \varphi'(x) \sin jx = -\sum_{i=0}^{k} ib_i \sin ix \sin jx$

$$= \sum_{i=0}^{k} \frac{ib_i}{2} (\cos(i+j)x - \cos(i-j)x).$$

Moreover,

$$\frac{1}{K} \sum_{m=0}^{K-1} \cos l\left(t + \frac{2\pi m}{K}\right) = \begin{cases} \cos lt & \text{if } K \mid l, \\ 0 & \text{if } K \nmid l. \end{cases}$$

We apply this to (6.9) and (6.10) to express σ_j and ϱ_j by the coefficients b_i. Recall that $K = k + 1$, so $K \mid i - j$ is equivalent to $i = j$, and $K \mid i + j$ is equivalent to $i = j = 0$ or $i + j = K$. Hence we obtain

(6.11) $\qquad \sigma_0 = b_0, \quad \sigma_j = \frac{b_j}{2} + \frac{b_{K-j}}{2} \cos Kt \quad (1 \le j \le k),$

(6.12) $\qquad \varrho_0 = 0, \quad \varrho_j = -\frac{jb_j}{2} + \frac{(K-j)b_{K-j}}{2} \cos Kt \quad (1 \le j \le K).$

Multiplying (6.11) by $K - j$ and subtracting (6.12) we obtain

$$\frac{(K-j)\sigma_j - \varrho_j}{K} = \frac{b_j}{2}$$

as claimed.

7. The estimate of large deviations. Our aim was to find for certain positive numbers c another number d with the following property: for

$$G_L(x) = \sum_{j=0}^{L-1} e(2^j x),$$

we have

$$\Delta = \mu(x \in [0,1] : |G_L(x)| > dL) < e^{-cL}.$$

Since

$$\int_0^1 e^{\lambda |G_L(x)|} \, dx > \Delta e^{\lambda dL}$$

and we know that the left side is $e^{\psi(\lambda)L + o(L)}$, we have

$$\Delta < e^{(\psi(\lambda) - \lambda d + o(1))L}.$$

So d is good if there is a λ such that $\psi(\lambda) - \lambda d < -c$, that is, if

(7.1) $\qquad d > \inf_{\lambda} \frac{\psi(\lambda) + c}{\lambda}.$

That is, to exhibit an acceptable value of d it is sufficient to take any value of λ and calculate an upper estimate of $\psi(\lambda)$. Such an estimate can be found

by taking any trigonometric polynomial $\overline{\varphi}$ satisfying $\overline{\varphi}(x) \geq e^{\lambda \cos x}$ for all x, calculating the matrix U described in Section 5, calculating U^L for some value of L and applying Lemma 7. For $\overline{\varphi}$ we took the polynomial described in Section 6, using degree 19, and for L we took powers of 2, so that U^L could be determined by repeated squaring. We also calculated lower estimates via the corresponding polynomial $\underline{\varphi}$ to reassure ourselves about the order of error. To find a good value of λ we observed that ψ is convex by Hölder's inequality, and it follows easily that the function $\psi_1(\lambda) = (\psi(\lambda) + c)/\lambda$ is first decreasing, then increasing for any fixed value of c. Hence if for some $\lambda_1 < \lambda_2 < \lambda_3$ we have

$$\psi_1(\lambda_1) > \psi_1(\lambda_2), \quad \psi_1(\lambda_2) < \psi_1(\lambda_3),$$

then the minimum is in $[\lambda_1, \lambda_3]$. By calculating values at midpoints, smaller and smaller intervals can be found.

We will formulate two corollaries to be applied later: (1) in the case when GRH is supposed to be true, (2) unconditionally in Part II.

COROLLARY 1. *We have*

$$|G_L(x)| = \left| \sum_{j=0}^{L-1} e(2^j x) \right| \leq 0.7163435444776661\, L$$

if $x \in [0,1] \setminus \mathcal{E}$ where $\mu(\mathcal{E}) = |\mathcal{E}| = O(N^{-1/2} L^{-100})$.

COROLLARY 2. *We have*

$$|G_L(x)| \leq 0.7894008449792309\, L$$

if $x \in [0,1] \setminus \mathcal{E}$ where $\mu(\mathcal{E}) = |\mathcal{E}| = O(N^{-3/5} L^{-100})$.

The actual calculations were done using U-basic, a multiprecision language written by and available for free from Professor Yuju Kida, Rikkyo University.

We remark that it is also possible to find lower estimates for Δ for a given d, or for d for a given c. To do this one takes three numbers $\lambda_1 < \lambda_2 < \lambda_3$ and applies Hölder's inequality on the set where $|G_L(x)| > dL$. In this way it is possible to prove that inequality (7.1) is an equality, assuming that the function ψ is differentiable. This seems to be very likely, though we have not proved it.

8. On numbers of the form $p + 2^\nu$. During the proof the distribution of the numbers

$$(8.1) \qquad r_1(n) = \#\{(p, m) : n = p + 2^m, p \leq N, 1 \leq m \leq L\}.$$

will play an important role. These numbers determine the important function

(8.2) $s(N) = \#\{(p_1, p_2, m_1, m_2) : p_1 - p_2 = 2^{m_2} - 2^{m_1},$

$$p_j \leq N,\ 1 \leq m_j \leq L,\ j = 1, 2\}$$

through the relation

(8.3) $$s(N) = \sum_n r_1^2(n).$$

Romanov used upper estimates for this quantity to show that a positive proportion of odd integers can be written in the form $p + 2^k$.

Using the generating functions of primes,

$$S(\alpha) = \sum_{p \leq N} e(p\alpha),$$

and of powers of 2,

$$G(\alpha) = \sum_{m \leq L} e(2^m \alpha),$$

we can express $s(N)$ also by the following integral:

(8.4) $$s(N) = \int_0^1 |S(\alpha)G(\alpha)|^2 \, d\alpha.$$

Similarly to previous works we will use the important theorem of Chen (our Lemma 3) to prove an upper bound for the integral (8.4) via estimating $s(N)$ from above. This problem was dealt with by Romanov [Rom] using Brun's sieve (see also [Pra, Satz 8.1]) with an unspecified constant in the upper bound.

LEMMA 10. *We have unconditionally*

(8.5) $$s(N) = \int_0^1 |S(\alpha)G(\alpha)|^2 \, d\alpha \leq \frac{2}{\log^2 2} C_2 N$$

where $C_2 < 5.3636$.

This is a significant improvement over the estimate $C_2 < 16.1214$ given implicitly in [LLW2, Lemma 4]. Further, it turns out later that actually we only need an estimate of the integral $|S(\alpha)G(\alpha)|^2$ on $C(\mathcal{M})$ instead of $[0, 1]$ (cf. (10.5)). Assuming GRH we can prove a better bound for the integral on $C(\mathcal{M})$ as follows.

LEMMA 11. *Assume GRH. Then*

(8.6) $$s_2(N) = \int_{C(\mathcal{M})} |S(\alpha)G(\alpha)|^2 \, d\alpha \leq \frac{2}{\log^2 2} C_2' N$$

where $C_2' < 3.9095$.

In [LLW2] this integral is estimated directly by the whole integral on $[0, 1]$, so our final gain is our final value $C_2' < 3.9095$ compared to 16.1214 in [LLW2]. Unlike the case of Romanov, the direct arithmetic information about $s(N)$ and $r_1(n)$ is irrelevant here. What we need is the estimate of the integral $|SG|^2$ and only on $C(\mathcal{M})$, which yields further savings as compared to the previous treatments.

In the proofs of Lemmas 10 and 11 we will work with the function $f(h)$ in (2.12) instead of the function $g(h) = \prod_{p|h}(1 + 1/p)$ of [LLW2]. Using the definitions of $s(N), R(h), f(h)$ in (8.2) and (2.11), (2.12), taking into account $R(0) = \pi(N)$ we obtain, by Chen's theorem (2.11),

$$(8.7) \quad s(N) \leq 2 \sum_{1 \leq m_1 < m_2 \leq L} f(2^{m_2} - 2^{m_1}) \cdot 2C_0 C^* \frac{N}{\log^2 N} + \pi(N)L$$

$$= 2 \sum_{l=1}^{L-1} f(2^l - 1)(L - l) \cdot 2C_0 C^* \frac{N}{\log^2 N} + \pi(N)L$$

(with $l = m_2 - m_1$). We shall investigate the quantity

$$(8.8) \quad F(L) = \frac{2}{L} \sum_{l=1}^{L-1} \left(1 - \frac{l}{L}\right) f(2^l - 1) \quad \text{as } L \to \infty.$$

Denote the order of 2 for an odd modulus d by

$$(8.9) \quad \xi(d) = \min_{\nu \geq 1}\{\nu : 2^\nu \equiv 1 \pmod{d}\}.$$

Further let $\sum'$ mean summation over squarefree odd integers and let

$$(8.10) \quad k(d) = \prod_{\substack{p>2 \\ p|d}} \frac{1}{p-2} \quad \text{so} \quad f(h) = {\sum_{d|h}}' k(d) \quad (f(1) = k(1) = 1).$$

With these notations we have

$$(8.11) \quad F(L) = \frac{2}{L} \sum_{l=1}^{L-1} \left(1 - \frac{l}{L}\right) {\sum_{d|2^l-1}}' k(d) = \frac{2}{L} {\sum_{d=1}^{2^L}}' k(d) \cdot \sum_{\substack{l=1 \\ \xi(d)|l}}^{L-1} \left(1 - \frac{l}{L}\right)$$

$$= \frac{2}{L} {\sum_{d=1}^{2^L}}' k(d) \sum_{1 \leq m \leq L/\xi(d)} \left(1 - \frac{m}{L/\xi(d)}\right).$$

Using the simple relation

$$(8.12) \quad \sum_{m \leq M} \left(1 - \frac{m}{M}\right) = [M] - \frac{[M]([M]+1)}{2M} \leq [M] - \frac{[M]}{2} \leq \frac{M}{2},$$

from (8.11) we obtain

$$(8.13) \qquad F(L) \le \sideset{}{'}\sum_{1 \le d \le 2^L} \frac{k(d)}{\xi(d)} =: F_0(L).$$

We know already from Romanov's work [Rom] that the positive series $F_0(L)$ converges to a constant R_0 as $L \to \infty$. One can show by calculations the estimate (see [KP, Corollary 1])

$$(8.14) \qquad 1.936 < R_0 < 1.94.$$

This implies by the prime number theorem, for any $\varepsilon > 0$ and $N > N_0(\varepsilon)$,

$$(8.15) \qquad s(N) \le L^2 \cdot R_0 2 C_0 C^* \frac{N}{\log^2 N} + \frac{NL(1+\varepsilon)}{\log N}$$

$$\le \frac{2N}{\log^2 2} \left(C_0 R_0 C^* + \frac{\log 2}{2} + \varepsilon \right).$$

Taking into account the values

$$(8.16) \qquad C_0 < 0.6602, \quad R_0 < 1.94, \quad C^* = 3.9171,$$

we obtain the proof of Lemma 10.

In order to show Lemma 11 we can write

$$(8.17) \qquad s_2(N) = \sum_{m_1=1}^{L} \sum_{m_2=1}^{L} R_2(2^{m_2} - 2^{m_1})$$

$$= 2 \sum_{1 \le m_1 < m_2 \le L} R_2(2^{m_2} - 2^{m_1}) + L R_2(0),$$

where

$$(8.18) \qquad R_2(h) = \int_{C(\mathcal{M})} |S^2(\alpha)| \, e(-h\alpha) \, d\alpha = \int_0^1 - \int_{\mathcal{M}} .$$

Using Lemma 1 and (8.7)–(8.14) we have

$$(8.19) \quad 2 \sum_{1 \le m_1 < m_2 \le L} R_2(2^{m_2} - 2^{m_1})$$

$$= 2 \sum_{1 \le m_1 < m_2 \le L} R(2^{m_2} - 2^{m_1}) - 2C_0 \frac{N}{\log^2 N} 2 \sum_{1 \le m_1 < m_2 \le L} f(2^{m_2} - 2^{m_1})$$

$$+ O\left(\frac{\log \log N}{\log^2 N} \right) \sum_{1 \le m_1 < m_2 \le L} (2^{m_2} - 2^{m_1}) + O(N/L)$$

$$\le \frac{2C_0 N}{\log^2 2} (R_0 C^* - R_0) + O\left(\frac{L \log \log N \cdot N}{\log^2 N} \right).$$

Further supposing GRH we have, in view of (2.5),

$$(8.20) \qquad R_2(0) = \pi(N) - \int_{\mathcal{M}} |S^2(\alpha)| \, d\alpha$$

$$\sim \operatorname{li} N - \sum_{q \leq P} \frac{\mu^2(q)}{\varphi^2(q)} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \int_{-1/qQ}^{1/qQ} \left| \sum_{n \leq N} \frac{e(n\eta)}{\log n} \right|^2 d\eta$$

$$\sim \frac{N}{\log N} - \sum_{q \leq P} \frac{\mu^2(q)}{\varphi(q)} \left(\frac{N}{\log^2 N} + O(qQ) \right)$$

$$\sim \frac{N}{\log N} - \frac{N \log P}{\log^2 N} + O(NL^{-8}).$$

Taking into account $\log P \sim (\log N)/2$, from (8.20) we get

$$(8.21) \qquad LR_2(0) \sim L \frac{N}{2 \log N} \sim \frac{N}{2 \log 2}$$

and this together with (8.17)–(8.19) implies

$$(8.22) \qquad s_2(N) \leq \frac{2N}{\log^2 2} \left(C_0 R_0 (C^* - 1) + \frac{\log 2}{4} + o(1) \right),$$

which proves Lemma 11, in view of (8.16).

9. Sums of k powers of 2. In this section we will investigate the function

$$(9.1) \quad r_{k,k}(m) = \#\{m = 2^{\nu_1} + \ldots + 2^{\nu_k} - 2^{\mu_1} - \ldots - 2^{\mu_k} : \nu_i, \mu_j \in [1, L]\}.$$

First we state the simple

LEMMA 12. $r_{k,k}(0) \leq 2L^{2k-2}$.

The proof (see Lemma 5 of [Gal]) follows from the fact that choosing all ν_i and $\mu_1, \ldots, \mu_{k-2}$ arbitrarily we have at most one representation for the remaining n as $2^{\mu_{k-1}} + 2^{\mu_k}$ if $\mu_{k-1} \leq \mu_k$.

As we will see later in the final result, a crucial role will be played by the upper estimation of

$$(9.2) \qquad S_1(k, N) = \int_{\mathcal{M}} |S^2(\alpha) G^{2k}(\alpha)| \, d\alpha.$$

Lemma 1 implies that the above quantity is

$$(9.3) \qquad S_1(k, N) = \sum_{m \leq N} r_{k,k}(m) \sigma(m) \frac{N - |m|}{\log^2 N} + O(NL^{2k-3}),$$

which can be estimated from above by

$$(9.4) \qquad S_1(k, N) \leq \frac{N}{\log^2 N} S(k, L) + O(NL^{2k-3}),$$

where $S(k, L) = \sum_m r_{k,k}(m)\sigma(m)$. Although it has no significance for the final result we remark that actually

$$(9.5) \qquad S_1(k, N) \sim \frac{N}{\log^2 N} S(k, L),$$

because $S(k, L) \geq L^{2k} \cdot 2C_0$ and

$$(9.6) \qquad \sum_{m \leq N} r_{k,k}(m)\sigma(m)|m|$$

$$\leq c \log \log N \left\{ (L - \log L)^{2k} \frac{N}{2^{\log L}} + L^{2k-1} \log L \cdot N \right\}.$$

An evaluation or even a good upper estimate of $S(k, L)$ in (9.4) turns out to be a rather difficult task. Using theoretical arguments and extensive computation the following result was proved in [KP] (cf. Theorems 1 and 2):

THEOREM 4. *For any given $k \geq 1$ the limit*

$$(9.7) \qquad A(k) := \lim_{L \to \infty} \left(\frac{S(k, L)}{2L^{2k}} - 1 \right)$$

exists, it strictly decreases with k, and

$$(9.8) \qquad A(k) > 2^{-2k-1} \quad \text{for every } k,$$

$$(9.9) \qquad \lim_{k \to \infty} A(k) = 0.$$

Further,

$$(9.10) \qquad 0.0126 < A(3) < 0.0136, \quad 0.003 < A(4) < 0.004.$$

We remark that the value $A(k)$ can be determined for any given k with a precision of ε in a finite time $T(k, \varepsilon)$ using a finite amount $M(R, \varepsilon)$ of memory with the methods applied in [KP].

10. Proof of Theorem 1. In order to prove Theorem 1 the crucial estimate is, similarly to [Gal] and [LLW1]:

LEMMA 13. *Assume GRH. Let $\eta = 0.283656$. For $k \geq 1$ and any $\delta > 0$ there exists $N_{k,\delta}$, depending on k and δ only, such that for $N \geq N_{k,\delta}$ we have*

$$(10.1) \qquad \sum_{m \leq N} (r'_k(m))^2 \leq \frac{2NL^{2k}}{\log^2 N} \{1 + A(k) + C'_2(1 - \eta)^{2k-2} + \delta\},$$

where $A(k)$ is defined by Theorem 4, and $C_2' = 3.9095$ by Lemma 11. In particular

(10.2) $A(3) + C_2'(1 - \eta)^4 < 1.0431, \quad A(4) + C_2'(1 - \eta)^6 < 0.5323.$

In general, we will follow [LLW1], since the method of [LLW3] is not so efficient in our case.

Proof of Lemma 13. Ignoring the terms with $m > N$ we obtain by Parseval's identity

(10.3) $\displaystyle\sum_{m \leq N} (r_k'(m))^2 \leq \int_0^1 |S(\alpha)G^k(\alpha)|^2 \, d\alpha = \int_{\mathcal{M}} + \int_{C(\mathcal{M}) \cap \mathcal{E}} + \int_{C(\mathcal{M}) \cap C(\mathcal{E})}$

where $\mathcal{E}$ is defined in Corollary 1 and $\mathcal{M}$ is given by (2.3) with P and Q as in (2.5). For the integral on the major arcs we have, by Lemmas 1, 12 and Theorem 4,

(10.4) $\displaystyle\int_{\mathcal{M}} |SG^k|^2 = \sum_m r_{k,k}(m) \int_{\mathcal{M}} |S(\alpha)|^2 e(m\alpha) \, d\alpha$

$\displaystyle\leq r_{k,k}(0) \int_0^1 |S(\alpha)|^2 \, d\alpha + \sum_{m \neq 0} r_{k,k}(m) \int_{\mathcal{M}} |S(\alpha)|^2 e(m\alpha) \, d\alpha$

$\displaystyle\leq 2L^{2k-2} \frac{2N}{\log N} + \sum_{m \neq 0} r_{k,k}(m)\sigma(m) \frac{N}{\log^2 N}\left(1 + O\left(\frac{1}{L}\right)\right)$

$\displaystyle\leq \frac{2NL^{2k}}{\log^2 N}\left(1 + A(k) + \frac{\delta}{3}\right).$

Using Lemma 11 and Corollary 1 we obtain

(10.5) $\displaystyle\int_{C(\mathcal{M}) \cap C(\mathcal{E})} \leq ((1-\eta)L)^{2k-2} \int_{C(\mathcal{M})} |S(\alpha)G(\alpha)|^2 \, d\alpha$

$\displaystyle\leq \frac{2C_2'N((1-\eta)L)^{2k-2}}{\log^2 2}.$

Finally, since $|\mathcal{E}| = o(N^{-1/2}L^{-6})$ (see Corollary 1) and by (2.10) we infer

(10.6) $\displaystyle\int_{C(\mathcal{M}) \cap \mathcal{E}} \leq C|\mathcal{E}| \cdot N^{3/2}L^4 L^{2k} \leq \frac{\delta}{3} NL^{2k-2}.$

Now the three estimates (10.4)–(10.6) prove our lemma.

From now on we will exactly follow [LLW1] but our estimate will be significant for much smaller values of k. We remark that [LLW1] follows basically [Gal], for the special case $g = 2$. Our next lemma is the following simple relation.

Lemma 14 ([Gal, Lemma 14]). *For $N \to \infty$,*

$$(10.7) \qquad \sum_{n \leq N} r_k'(n) \sim \frac{NL^k}{\log N} = \frac{N}{2} \cdot \frac{2L^k}{\log N}.$$

Proof of Theorem 1. The dispersion method and Lemma 14 yield

$$(10.8) \qquad E = \sum_{\substack{n \leq N \\ 2 \nmid n}} (r_k'(n) - \lambda_k)^2$$

$$= \sum_{\substack{n \leq N \\ 2 \nmid n}} (r_k'(n))^2 - 2\lambda_k \sum_{\substack{n \leq N \\ 2 \nmid n}} r_k'(n) + \lambda_k^2 \sum_{\substack{n \leq N \\ 2 \nmid n}} 1$$

$$\leq \sum_{\substack{n \leq N \\ 2 \nmid n}} (r_k'(n))^2 - \frac{2NL^{2k}}{\log^2 N} (1 + o(1)),$$

where

$$(10.9) \qquad \lambda_k := \frac{2L^k}{\log N} \sim \frac{2L^{k-1}}{\log 2}.$$

Put $K = i + j$ with either $i = j$ or $i = j + 1$ according as K is even or not. Then

$$(10.10) \qquad r_K''(N) = \sum_{\substack{m+n=N \\ 2 \nmid m, \, 2 \nmid n}} r_i'(m) r_j'(n).$$

Let

$$(10.11) \qquad r_i'(m) = \lambda_i + s_i(m) = \frac{2L^i}{\log N} + s_i(m) \quad \text{for } 2 \nmid m.$$

Then

$$(10.12) \quad r_K''(N)$$
$$= \frac{4L^{i+j}}{\log^2 N} \sum_{\substack{m+n=N \\ 2 \nmid m, \, 2 \nmid n}} 1 + 2 \left\{ \frac{L^i}{\log N} \sum_{\substack{m+n=N \\ 2 \nmid m, \, 2 \nmid n}} s_j(n) + \frac{L^j}{\log N} \sum_{\substack{m+n=N \\ 2 \nmid m, \, 2 \nmid n}} s_i(m) \right\}$$
$$+ \sum_{\substack{m+n=N \\ 2 \nmid m, \, 2 \nmid n}} s_i(m) s_j(n).$$

The 2nd and 3rd terms are $o(NL^{i+j-2}) = o(NL^{K-2})$, since by Lemma 14,

$$(10.13) \qquad \sum_{\substack{m \leq N \\ 2 \nmid m}} s_k(m) = \sum_{\substack{m \leq N \\ 2 \nmid m}} r_k'(m) - \frac{2L^k}{\log N} \cdot \frac{N}{2} = o(NL^{k-1}).$$

By Cauchy's inequality, (10.8) and Lemma 13 the last sum is bounded by

$$(10.14) \quad \Big\{\sum_{\substack{m \leq N \\ 2 \nmid m}} s_i^2(m)\Big\}^{1/2} \Big\{\sum_{\substack{n \leq N \\ 2 \nmid n}} s_j^2(n)\Big\}^{1/2}$$

$$\leq \frac{2NL^K}{\log^2 N} (A(i) + C_2'(1-\eta)^{2i-2} + \delta)^{1/2}(A(j) + C_2'(1-\eta)^{2j-2} + \delta)^{1/2}.$$

If we choose $K = 7$, $i = 4$, $j = 3$, the above term is

$$(10.15) \qquad\qquad \leq \frac{2NL^7}{\log^2 N} \cdot 0.7452,$$

by (10.2). Since the first term in (10.12) is

$$(10.16) \qquad\qquad \frac{2NL^K}{\log^2 N} = \frac{2NL^7}{\log^2 N},$$

the statement of Theorem 1 is proved for $K = 7$. For $K \geq 7$ the result follows either directly from this (using $\nu_8 = \ldots = \nu_K = 1$ for example) or we could prove better bounds for the sum in (10.14) using the fact that $A(k)$ is strictly increasing and tends to 0 as $k \to \infty$.

References

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

[Che]   J. R. Chen, *On Goldbach's problem and the sieve methods*, Sci. Sinica Ser. A 21 (1978), 701–738.

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

[HL]   G. H. Hardy and J. E. Littlewood, *Some problems of "Partitio Numerorum":* *III, On the expression of a number as a sum of primes*, Acta Math. 44 (1923), 1–70.

[HP]   D. R. Heath-Brown and J.-C. Puchta, *Integers represented as a sum of primes and powers of two*, manuscript.

[KP]   A. Khalfalah and J. Pintz, *On the representation of Goldbach numbers by a bounded number of powers of two*, manuscript.

[Li1]   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.

[Li2]   —, *The number of powers of 2 in a representation of large even integers by sums of such powers and of two primes (II)*, ibid. 96 (2001), 369–379.

[Lin1] Yu. V. Linnik, *Prime numbers and powers of two*, Trudy Mat. Inst. Steklov. 38 (1951), 152–169 (in Russian).

[Lin2] —, *Addition of prime numbers and powers of one and the same number*, Mat. Sb. (N.S.) 32 (1953), 3–60 (in Russian).

[LLW1] J. Y. Liu, M. C. Liu and T. Z. Wang, *The number of powers of 2 in a representation of large even integers (I)*, Sci. China Ser. A 41 (1998), 386–397.

[LLW2] —, —, —, *The number of powers of 2 in a representation of large even integers (II)*, ibid., 1255–1271.

[LLW3]   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.
[Pra]   K. Prachar, *Primzahlverteilung*, Springer, Berlin, 1957.
[Rom]   N. P. Romanov, *Über einige Sätze der additiven Zahlentheorie*, Math. Ann. 109 (1934), 668–678.
[Wan]   T. Z. Wang, *On Linnik's almost Goldbach theorem*, Sci. China Ser. A 42 (1999), 1155–1172.

Rényi Institute of the Hungarian Academy of Sciences
H-1364 Budapest, Pf. 127, Hungary
E-mail: pintz@renyi.hu
        ruzsa@renyi.hu