Disproof of the Mertens conjecture

By *A. M. Odlyzko* at Murray Hill and *H. J. J. te Riele* at Amsterdam

1. Introduction

Let $\mu(n)$ denote the Möbius function, so that

$$\mu(n)=\begin{cases} 1, & n=1, \\ (-1)^k, & n=\prod_{i=1}^{k} p_i, \quad p_i \text{ distinct primes}, \\ 0, & p^2 \mid n \text{ for some prime } p, \end{cases}$$

and let

$$M(x)=\sum_{n\leqq x}\mu(n). \tag{1.1}$$

Then $M(x)$ is the difference between the number of squarefree positive integers $n\leqq x$ with an even number of prime factors and of those with an odd number of prime factors. In 1885, T. J. Stieltjes claimed in a letter to Hermite [43] to have a proof that $M(x)\,x^{-\frac{1}{2}}$ always stays between two fixed bounds, no matter how large x may be. In parentheses, Stieltjes added that one could probably take $+1$ and -1 for these bounds. Stieltjes never published his "proof," but his claim to have it was apparently known to quite a few mathematicians, as were the important consequences that would follow from it. Thus, for example, Hadamard in his paper proving the Prime Number Theorem [16] mentioned that Stieltjes had much stronger results than Hadamard on the zeros of the zeta function, but that the new results of Hadamard might still be of interest because of their simpler proofs! In retrospect it seems likely that Stieltjes was wrong in his assertion, since, as will be explained later, it seems very probable that

$$\limsup_{x\to\infty}|M(x)|x^{-\frac{1}{2}}=\infty. \tag{1.2}$$

This conjecture remains unproved.

The motivation for Stieltjes' work on $M(x)$ was that, as will be explained in Section 2, the size of $M(x)$ is closely connected to the distribution of the non-trivial zeros of the Riemann zeta function, and the boundedness of $M(x)\,x^{-\frac{1}{2}}$ would imply the Riemann hypothesis. This same motivation inspired the work of other mathematicians (cf. [9]) and it led Mertens to publish in 1897 a paper [27] with a 50-page table of $\mu(n)$ and $M(n)$ for $n=1, 2, \ldots, 10000$. On the basis of the evidence in the table, Mertens concluded that the inequality

$$(1.3) \qquad |M(x)| < x^{\frac{1}{2}}, \quad x > 1,$$

is "very probable." The inequality (1. 3), which was first conjectured in the letter of Stieltjes we mentioned above, is now known as the *Mertens conjecture*.

In a series of papers [39]–[42], von Sterneck published additional values of $M(n)$ for $n \leqq 5 \times 10^6$, and on the basis of that evidence he conjectured that

$$(1.4) \qquad |M(x)| < \frac{x^{\frac{1}{2}}}{2} \quad \text{for} \quad x > 200.$$

He stated [42] that (1. 4) is a "yet unproved, but extremely probable number-theoretic law." However, in 1960 W. Jurkat [19], [20] found a disproof of (1. 4) that involved very little computation. Jurkat's method, which did not produce a specific counterexample to (1. 4), is described in Section 2. The first counterexample to (1. 4) that was found is due to Neubauer [29], who computed all $M(n)$ for $n \leqq 10^8$ and for various values of n in the interval $(10^8, 10^{10})$. Near 7.77×10^9 he found values of n for which $M(n) > \frac{n^{\frac{1}{2}}}{2}$. However, Neubauer's computations as well as the later ones of Yorinaga [45] (who computed $M(n)$ for all $n \leqq 4 \times 10^8$) and of Cohen and Dress [8] (who computed $M(n)$ for all $n \leqq 7.8 \times 10^9$ and found that the smallest n for which $M(n) > \frac{n^{\frac{1}{2}}}{2}$ is $n = 7{,}725{,}038{,}629$ with $M(7{,}725{,}038{,}629) = 43947$) did not find any values of n for which the Mertens conjecture is violated. The inequality $|M(n)| < 0.6\, n^{\frac{1}{2}}$ holds for all the values of n for which $M(n)$ has been computed.

In this paper we will disprove the Mertens conjecture by showing that

$$\limsup_{x \to \infty} M(x)\, x^{-\frac{1}{2}} > 1.06,$$

$$\liminf_{x \to \infty} M(x)\, x^{-\frac{1}{2}} < -1.009.$$

Our disproof is indirect, and does not produce any single value of x for which $|M(x)| > x^{\frac{1}{2}}$. In fact, we suspect that there are no counterexamples to the Mertens conjecture for $x \leqq 10^{20}$ or perhaps even 10^{30}. (Section 5 explains the reasons for this belief.)

The disproof of the Mertens conjecture closes off another possible road to proving the Riemann hypothesis. The Riemann hypothesis would also follow from any inequality of the form $|M(x)| \leqq c x^{\frac{1}{2}}$ for any fixed c. Our disproof provides some additional evidence that no such inequality holds, and that (1. 2) is correct, since our method can undoubtedly be used to produce larger values for $\limsup |M(x)|\, x^{-\frac{1}{2}}$ than 1.06 with the use of more computer time.

While the Mertens conjecture was known to imply the Riemann hypothesis, the converse is definitely not the case. Hence our disproof of the Mertens conjecture does not imply anything about the possible falsity of the Riemann hypothesis (which has just been verified for the first 1.5×10^9 zeros [26]). In fact, as is explained in Section 2, the Mertens conjecture has been expected to be false for a long time.

No good conjectures about the rate of growth of $M(x)$ are known. Certainly $M(x)\,x^{-\frac{1}{2}}$ is expected to be unbounded, and the Riemann hypothesis is known [44] to be equivalent to $|M(x)| = O(x^{\frac{1}{2}+\varepsilon})$ for every $\varepsilon > 0$. The assumption of certain random features in the behavior of the sequence $\{\mu(n)\}$ led Good and Churchhouse [13] to conjecture that

$$\limsup_{x\to\infty} |M(x)|(x \log\log x)^{-\frac{1}{2}} = \frac{\sqrt{12}}{\pi},$$

and a similar remark was made by Paul Lévy in a comment on a paper by Saffari [37]. However, these conjectures seem quite questionable, since, as will be explained in Section 2, the behavior of $M(x)$ is determined by the zeros of the zeta function. Various rigorous results about sign changes of $M(x)$, for example, can be found in [31] and the references listed there.

Conjectures analogous to the Mertens conjecture, but for coefficients of cusp forms, have also been made [12]. Many instances of those conjectures have been disproved by indirect methods [1], [15], but the analogues of the conjecture (1. 2) remain unproved.

In Section 2 we survey previous work on the Mertens conjecture, and in particular the reasons why it was thought to be false and the possible methods of disproving it. Section 3 describes the lattice basis reduction algorithm of Lenstra, Lenstra, and Lovász [25], which was the main new ingredient that allowed us to obtain much stronger results than those of previous authors. Section 4 describes the numerical computations used in our disproof, which consisted mainly of computing the first 2000 zeros of the zeta function to about 100 significant decimal digits and of applying the lattice basis reduction algorithm. Finally, Section 5 concludes with some remarks about the possible locations of counterexamples to the Mertens conjecture, the complexity of computing $M(x)$, the random behavior of zeros of the zeta function, and possible extensions of our work.

2. The Mertens conjecture and diophantine approximation properties of zeros of the zeta function

It is easy to see that the Mertens conjecture implies the Riemann hypothesis. For $\sigma = \mathrm{Re}(s) > 1$, we have

$$\begin{aligned} \frac{1}{\zeta(s)} &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \sum_{n=1}^{\infty} \frac{M(n)-M(n-1)}{n^s} \\ &= \sum_{n=1}^{\infty} M(n)\left\{\frac{1}{n^s} - \frac{1}{(n+1)^s}\right\} = \sum_{n=1}^{\infty} M(n) \int_n^{n+1} \frac{s\,dx}{x^{s+1}} \\ &= s \sum_{n=1}^{\infty} \int_n^{n+1} \frac{M(x)\,dx}{x^{s+1}} = s \int_1^{\infty} \frac{M(x)\,dx}{x^{s+1}}, \end{aligned} \tag{2. 1}$$

since $M(x)$ is constant on each interval $[n, n+1)$. If the Mertens conjecture were true, then the last integral in (2. 1) would define a function analytic in $\sigma > \frac{1}{2}$, and this would give an analytic continuation of $\frac{1}{\zeta(s)}$ to $\sigma > \frac{1}{2}$. In particular, this would imply that $\zeta(s)$ has no zeros in $\sigma = \mathrm{Re}(s) > \frac{1}{2}$, which is exactly the statement of the Riemann hypothesis. Furthermore, the integral representation (2. 1) would then show that for $\sigma > \frac{1}{2}$,

$$\left|\frac{1}{\zeta(s)}\right| \leqq |s| \int_1^\infty \frac{x^{\frac{1}{2}}\,dx}{x^{\sigma+1}} = \frac{|s|}{\sigma - \frac{1}{2}}. \tag{2. 2}$$

This would imply that the zeta function does not have any multiple zeros, since if $\sigma = \frac{1}{2} + i\gamma$ were a zero of multiplicity k, then for some constant $a > 0$,

$$\left|\zeta\left(\frac{1}{2} + u + i\gamma\right)\right| \sim a u^k \quad \text{as} \quad u \to 0^+,$$

which is inconsistent with (2. 2) for $k \geqq 2$.

The above proof that the Mertens conjecture implies both the Riemann hypothesis and the simplicity of the zeros of the zeta function does not depend in any way on the constant in the conjecture; the assumption that $|M(x)| \leqq A x^{\frac{1}{2}}$ for any fixed A and all $x \geqq 1$ would have sufficed. Furthermore, it has been shown [18], [44] that the Riemann hypothesis and the simplicity of the zeros, as well as some other results, follow from any one of the following three weaker hypotheses:

i) $\limsup\limits_{x \to \infty} M(x)\, x^{-\frac{1}{2}} \leqq A$ for some constant A;

ii) $\liminf\limits_{x \to \infty} M(x)\, x^{-\frac{1}{2}} \geqq -A$ for some constant A;

iii) $\int_1^y M(x)^2 x^{-2}\,dx = O(\log y)$ as $y \to \infty$.

The Riemann hypothesis and the simplicity of the zeros of the zeta function are quite widely expected to hold, so the fact that they follow from the Mertens conjecture did not cast any special doubt on the latter. What did raise overwhelming skepticism about the truth of the Mertens conjecture was a series of completely unexpected results about the zeros of the zeta function that were deduced from it. We next explain these results.

The "exact formulas" of prime number theory, which express functions such as $\pi(x)$ in terms of zeros of the zeta function, are well known. Titchmarsh [44] has obtained a similar formula for $M(x)$. He showed that if the Riemann hypothesis holds and if there are no multiple zeros of the zeta function, then there is a sequence T_k, $k \leqq T_k \leqq k+1$, such that

$$M_0(x) = \lim_{k \to \infty} \sum_{\substack{\rho \\ |\gamma| < T_k}} \frac{x^\rho}{|\rho\zeta'(\rho)|} - 2 + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}\left(\frac{2\pi}{x}\right)^{2n}}{(2n)!\,n\zeta(2n+1)}, \tag{2. 3}$$

where $M_0(x) = M(x) - \frac{\mu(x)}{2}$ if $x \in \mathbb{Z}^+$ and $M_0(x) = M(x)$ otherwise, and $\rho = \frac{1}{2} + i\gamma$ runs over the nontrivial zeros of the zeta function. The formula (2. 3) has to be modified if there are multiple zeros of the zeta function, but since we are interested in the consequences of the Mertens conjecture, we will be assuming from now on that all the nontrivial zeros are simple and on the critical line $\left(\rho = \frac{1}{2} + i\gamma\right)$.

The second series on the right side of (2. 3) converges very rapidly. That is not the case with the first series. In fact, since $M_0(x)$ has jump discontinuities at the square-free integers, we must have

(2. 4) $$\sum_\rho \frac{1}{|\rho\zeta'(\rho)|} = \infty .$$

Aside from (2. 4), very little is known (cf. [43]) about the sizes of the $\rho\zeta'(\rho)$, whether the Mertens conjecture is assumed or not. Write

(2. 5) $$x = e^y, \qquad -\infty < y < \infty,$$

and note that with this notation (neglecting for the moment the question of convergence of the series involved),

$$\sum_\rho \frac{x^\rho}{\rho\zeta'(\rho)} = e^{\frac{y}{2}} \sum_\rho \frac{e^{i\gamma y}}{\rho\zeta'(\rho)},$$

which (aside from the $\exp\left(\frac{y}{2}\right)$ factor) looks like a general harmonic series. If we define

(2. 6) $$m(y) = M(x)\, x^{-\frac{1}{2}} = M(e^y)\, e^{-\frac{y}{2}},$$

then (2. 3) shows that if

(2. 7) $$h(y) = \lim_{k \to \infty} \sum_{\substack{\rho \\ |\gamma| < T_k}} \frac{e^{i\gamma y}}{\rho\zeta'(\rho)},$$

then

(2. 8) $$m(y) = h(y) + O\left(\min\left(1, e^{-\frac{y}{2}}\right)\right).$$

To prove that $\limsup m(y) > A$ as $y \to \infty$, it therefore suffices to prove that $\limsup h(y) > A$ as $y \to \infty$. Unfortunately we just don't know enough about the series (2. 7) defining $h(y)$ to do this directly.

Ingham [18] reduced the problem of the behavior of $h(y)$ to that of a somewhat more tractable function. (For other applications of Ingham's method to number theoretic conjectures see, for example, [3].) The problem with the series in (2. 7) is that it is infinite, and very little is known about the sizes of the coefficients. Ingham's solution was to study finite series of that kind. In engineering language, in order to remove particular frequency components from a signal (such as $h(y)$), one passes the signal through an appropriate filter, which corresponds to convolving the signal with another function. More precisely, if $K(y)$ is a suitably behaved function, and

(2. 9) $$k(t) = \int_{-\infty}^{\infty} K(y)\, e^{-ity}\, dy,$$

then, neglecting questions about the convergence of the series (2. 7) for $h(y)$ and the validity of various interchanges of summation and integration below, we obtain

$$h_K(y) = \int_{-\infty}^{\infty} h(y-t)\, K(t)\, dt = \sum_{\rho} \frac{e^{i\gamma y}}{\rho \zeta'(\rho)} \int_{-\infty}^{\infty} K(t)\, e^{-i\gamma t}\, dt = \sum_{\rho} k(\gamma) \frac{e^{i\gamma y}}{\rho \zeta'(\rho)}. \tag{2. 10}$$

If $K(y)$ is chosen such that $k(t)$ is of bounded support, then the last sum in (2. 10) is finite, and there are no problems about its convergence. That sum no longer gives us $h(y)$, but $h_K(y)$, which is a weighted average of $h(y)$. However, if $h_K(y_0)$ is large, then $h(y)$ must be large for some value of y;

$$\sup_y |h(y)| \int_{-\infty}^{\infty} |K(t)| dt \geqq |h_K(y_0)|. \tag{2. 11}$$

In fact, if $K(t) \geqq 0$ for all t and $k(-\gamma) = k(\gamma)$ is real, so that $h_K(y)$ is real, then we even obtain

$$\sup_y h(y) \int_{-\infty}^{\infty} K(t)\, dt \geqq h_K(y_0), \tag{2. 12}$$

$$\inf_y h(y) \int_{-\infty}^{\infty} K(t)\, dt \leqq h_K(y_0). \tag{2. 13}$$

Given any y_0, one can actually draw conclusions stronger than (2. 12)—(2. 13). The function $h_K(y)$ is almost-periodic in the sense of Bohr: it follows that given any y_0 and any $\varepsilon > 0$ there is an unbounded sequence of values of y such that

$$|h_K(y) - h_K(y_0)| < \varepsilon.$$

In our case, where the sum in (2. 10) that equals $h_K(y)$ is finite, this result follows almost trivially from Kronecker's theorem about simultaneous diophantine approximations, since that result implies that for any $\delta > 0$, we can find arbitrarily large values of y^* such that

$$|\gamma y^* - 2\pi m_\gamma| < \delta \tag{2. 14}$$

for some $m_\gamma \in \mathbb{Z}$ and all γ such that $k(\gamma) \neq 0$, and then $y = y_0 + y^*$ gives the desired result if δ is small enough. Therefore we find that

$$\limsup_{y \to \infty} h(y) \int_{-\infty}^{\infty} K(t)\, dt \geqq h_K(y_0), \tag{2. 15}$$

$$\liminf_{y \to \infty} h(y) \int_{-\infty}^{\infty} K(t)\, dt \leqq h_K(y_0). \tag{2. 16}$$

Of course, we really wish to study $m(y)$, not $h(y)$. However, (2. 8) shows that we can replace $h(y)$ by $m(y)$ in (2. 15) and (2. 16) and still obtain a valid result.

The above discussion was only meant to provide a heuristic explanation of the Ingham method. The technical difficulties involved in carrying out these ideas can be overcome in several ways [18], [20], [21], and it is possible to obtain the following result.

Theorem. *Suppose that* $K(y) \in C^2(-\infty, \infty)$, $K(y) \geqq 0$, $K(-y) = K(y)$,

$$K(y) = O\left((1+y^2)^{-1}\right)$$

as $y \to \infty$, *and that* $k(t)$, *defined by* (2. 9), *satisfies* $k(t) = 0$ *for* $|t| \geqq T$ *for some* T, *and* $k(0) = 1$. *If the zeros* $\rho = \beta + i\gamma$ *of the zeta function with* $0 < \beta < 1$ *and* $|\gamma| < T$ *satisfy* $\beta = \frac{1}{2}$ *and are simple, then for any* y_0,

$$\limsup_{y \to \infty} m(y) \geqq h_K(y_0), \tag{2. 17}$$

$$\liminf_{y \to \infty} m(y) \leqq h_K(y_0), \tag{2. 18}$$

where

$$h_K(y) = \sum_{\rho} k(\gamma) \frac{e^{i\gamma y}}{\rho \zeta'(\rho)}.$$

Perhaps the simplest function $k(t)$ that satisfies the conditions of the theorem is the Fejer kernel used by Ingham:

$$k(t) = \begin{cases} 1 - \frac{|t|}{T}, & |t| \leqq T, \\ 0, & |t| > T. \end{cases} \tag{2. 19}$$

(Later on we will use a somewhat better kernel.) For this choice

$$\begin{aligned} h_K(y) &= \sum_{|\gamma| < T} \left(1 - \frac{|\gamma|}{T}\right) \frac{e^{i\gamma y}}{\rho \zeta'(\rho)} \\ &= 2 \sum_{0 < \gamma < T} \left(1 - \frac{\gamma}{T}\right) \frac{\cos(\gamma y - \psi_\gamma)}{|\rho \zeta'(\rho)|}, \end{aligned} \tag{2. 20}$$

where

$$\psi_\gamma = \operatorname{Arg} \rho \zeta'(\rho). \tag{2. 21}$$

Since the sum of all the $|\rho \zeta'(\rho)|^{-1}$ diverges, we can make the sum of the coefficients in (2. 20) arbitrarily large by choosing T very large. If we could then find values of y such that all of the $\gamma y - \psi_\gamma$ were close to integer multiples of 2π, we could make $h_K(y)$ arbitrarily large, contradicting (by the theorem above) the Mertens conjecture. If the γ's were linearly independent over the rationals, then by Kronecker's theorem for any given $\varepsilon > 0$, there would even be integer values of y which satisfy

$$|\gamma y - \psi_\gamma - 2\pi m_\gamma| < \varepsilon \tag{2. 22}$$

for all $\gamma \in (0, T)$ and some integers m_γ. That would show that $h_K(y)$ can be made arbitrarily large. Furthermore, it is easy to deduce [18] that even if there are only a bounded number of linearly independent relations of the form

$$\sum_{\gamma} c_\gamma \gamma = 0, \quad c_\gamma \in \mathbb{Z}, \tag{2. 23}$$

with only finitely many c_γ being nonzero, then $h_K(y)$ is unbounded. If relations of the form (2. 23) do exist, however, the behavior of $h_K(y)$ could possibly be quite arbitrary (cf. [35]).

Since there did not seem to be any reasonable explanation as to why the γ's ought to satisfy any linear relations with integral coefficients, most experts concluded from Ingham's observation that the Mertens conjecture was unlikely to be true. Further doubt on the validity of the Mertens conjecture, and of the weaker conjecture that $m(y)$ is bounded, was generated by the work of Bateman et al. [4]. Using a technique developed by Bohr and Jessen [6] to prove Kronecker's theorem, they showed that if $m(y)$ is bounded, then there exist infinitely many relations among the γ's of the form (2. 23), where the $c_\gamma = 0$, ± 1, or ± 2, and at most one of the c_γ satisfies $|c_\gamma| = 2$. This was even more surprising and helped deepen skepticism about the Mertens conjecture even further, especially since Bateman et al. [4] looked at linear combinations of the first few γ's with coefficients of the above form and did not find anything that might suggest the existence of linear relations of the required type.

The Ingham [18] and Bateman et al. [4] results not only provided grounds for disbelieving in the Mertens conjecture, but in addition suggested ways to disprove it. One way to disprove this conjecture, of course, is to simply compute $M(x)$ for various values of x until a counterexample is found. This is basically how Neubauer [29] disproved the von Sterneck conjecture that $|m(y)| < \frac{1}{2}$ for $y \geqq 5.3$. However, we suspect (for reasons that will be explained in the last section) that there are no counterexamples to the Mertens conjecture for $x \leqq 10^{30}$, so this approach does not look very promising.

Another way to disprove the Mertens conjecture, which is due to Jurkat [19], [20] is to use the second sum in (2. 3),

$$g(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}\left(\frac{2\pi}{x}\right)^{2n}}{(2n)!\, n\zeta(2n+1)}. \tag{2. 24}$$

For example, as x passes 1, $M_0(x)$ jumps by 1, while $g(x)$ is continuous there. Hence $h(y)$ has a jump of 1 at $y = 0$, and so its absolute value has to be at least $\frac{1}{2}$ on at least one side of $x = 1$. Then, by using the almost-periodicity of $h(y)$ (or its averages, to be rigorous), we find even without any computation that as $y \to \infty$, $\limsup |h(y)| \geqq \frac{1}{2}$, which disproves the von Sterneck conjecture. (Anderson [2] has recently shown that $\limsup |m(y)| \geqq \frac{1}{2}$ as $y \to \infty$ by a somewhat different method.) By computing the value of $g(1)$, Jurkat [20] showed by this approach that

$$\liminf_{y \to \infty} m(y) \leqq -0.5054. \tag{2. 25}$$

One could hope to obtain results better than (2. 25) by finding very small values of x for which $g(x)$ is large, but so far no good way for finding such values of x has been proposed. Our computations do produce some candidate values for such x, but they require impractically large amounts of computation to test.

Another method for disproving the Mertens conjecture was developed from the work of Bateman et al. [4]. It was shown that the Mertens conjecture implies that there exist relations of the form (2. 23) in which the c_γ are not too large, and where only relatively small γ's can have $c_\gamma \neq 0$ [10], [14], [36], [37]. This reduces the problem of disproving the Mertens conjecture to verifying that none of a finite number of linear relations holds. Quantitatively the best result of this kind is due to Grosswald [14], who showed that the Mertens conjecture implies that there is a relation of the form (2. 23) with all $|c_\gamma| \leqq 13$ and $c_\gamma \neq 0$ for no more than the first 75 γ's. With presently known algorithms, though, it does not seem feasible to disprove the Mertens conjecture this way; we would need to show that none of the $27^{75} \approx 10^{107}$ possible relations holds, and no method is known for doing this in fewer than about 10^{54} operations, which is much higher than the 10^{10} to 10^{15} operations that one can realistically expect to be able to perform with present and foreseeable computers. (We do not specify precisely what we mean by an operation since it is not very important in the present context, given the huge numbers involved.)

The final method of disproving the Mertens conjecture that we discuss is the one that had given the best results in the past and enabled us to carry out the disproof. It is based on the Ingham approach and proceeds by finding values of y for which $h_K(y)$ is large in absolute value. The simplest way to carry out this idea is to simply evaluate $h_K(y)$ at various values of y. A slightly more sophisticated approach is to start evaluating the series for $h_K(y)$, and if the partial sums seem too small, to terminate the evaluation and go on to the next value of y. In this way Spira [38] showed that

$$\limsup_{y \to \infty} m(y) \geqq 0.5355,$$

$$\liminf_{y \to \infty} m(y) \leqq -0.6027.$$

He used $k(t)$ of the form (2. 19) with $T=1000$.

Jurkat and Peyerimhoff [21] improved on Spira's results by using a more sophisticated approach. While we know very little theoretically about the sizes of the coefficients $(\rho\zeta'(\rho))^{-1}$, numerically they appear to be typically on the order of ρ^{-1} or of $(\rho \log|\rho|)^{-1}$. In particular, these coefficients decrease quite rapidly, and so the size of $h_K(y)$ is determined largely by the first few terms. To make the first few terms large (and positive, say), one needs to find a y that solves the inhomogeneous diophantine approximation problem of making

$$|\gamma y - \psi_\gamma - 2\pi m_\gamma| < \varepsilon \tag{2. 26}$$

for the chosen γ's, where $m_\gamma \in \mathbb{Z}$ and ε is not too large. Jurkat and Peyerimhoff invented an algorithm for finding such values of y. For each value of y produced by this algorithm, $h_K(y)$ was evaluated (for $K(y)$ of the kind we will describe later, and with $k(t)=0$ for $|t| \gtrsim 900$), and the best values gave

$$\limsup_{y \to \infty} m(y) \geqq 0.779,$$

$$\liminf_{y \to \infty} m(y) \leqq -0.638.$$

The Jurkat-Peyerimhoff computations were carried out on a programmable desk calculator. Te Riele [33] implemented the Jurkat-Peyerimhoff algorithm (together with a few improvements) on a high speed computer, and proved that

$$\limsup_{y \to \infty} m(y) \geqq 0.860,$$
$$\liminf_{y \to \infty} m(y) \leqq -0.843.$$

(The kernel $k(t)$ used in these computations was of the same form as that of Jurkat and Peyerimhoff, but it was nonzero at the first 15,000 zeros instead of the first 536.) The computations took several hundred hours, and te Riele concluded that with the use of the Jurkat-Peyerimhoff algorithm and then current technology, a disproof of the Mertens conjecture was unlikely to be achieved.

Our disproof of the Mertens conjecture is due not to advances in computer technology (since we used much less computer time than was used by te Riele in the earlier work), but to a major breakthrough in diophantine approximation methods which was made recently, and which is described in the next section.

3. Inhomogeneous diophantine approximation

In order to find a y which solves (2.26) for a subset of small γ's, call them $\gamma_1, \gamma_2, \ldots, \gamma_n$ (which in general are not the first n γ's, in contrast to the notation of Section 4.2) and a small ε, we have used a remarkable new algorithm due to Lenstra, Lenstra, and Lovász [25], which we will refer to as the L^3 algorithm. This algorithm was designed to find short vectors in lattices, and since many computational problems can be reduced to finding short vectors in lattices, it has since found widespread applications in polynomial factorizations [25] and public key cryptography (cf. [23]). The problem of finding the shortest nonzero vector in a lattice appears to be very hard. The L^3 algorithm is not guaranteed to find the shortest vector, but it does run in polynomial time (in the length of the input) and finds quite short vectors. More precisely, if $\underline{v}_1, \ldots, \underline{v}_m$ is a set of basis vectors of an m-dimensional lattice L in $\mathbb{R}^m$, then the L^3 algorithm finds another basis, $\underline{v}_1^*, \ldots, \underline{v}_m^*$, called *reduced* [25], which satisfies

$$\|\underline{v}_1^*\| \leqq \left(\frac{4}{4u-1}\right)^{\frac{m-1}{2}} \min_{\substack{\underline{v} \in L \\ \underline{v} \neq \underline{0}}} \|\underline{v}\|,$$

$$\prod_{i=1}^{m} \|\underline{v}_i^*\| \leqq \left(\frac{4}{4u-1}\right)^{\frac{m(m-1)}{4}} d(L),$$

where $u \in \left(\frac{1}{4}, 1\right)$ is a parameter chosen beforehand, $d(L)$ is the determinant of the lattice, and $\|\underline{v}\|$ denotes the euclidean norm of the vector $\underline{v}$. What is perhaps most remarkable about the L^3 algorithm is that in practice it performs much better than it is guaranteed to. This is important in our case because we have used it in situations it was not designed to deal with, and so there was no *a priori* guarantee that it would find the desired solution.

Results of extensive experiments with the L^3 algorithm and descriptions of various modifications to it which make it run faster and find better solutions are described in [23]. Right now we describe how the problem of finding a y such that each of

$$\eta_j = \gamma_j y - \psi_j - 2\pi m_j, \quad 1 \leqq j \leqq n, \tag{3. 1}$$

is small, where

$$\psi_j = \operatorname{Arg}\left(\frac{1}{2} + i\gamma_j\right)\zeta'\left(\frac{1}{2} + i\gamma_j\right), \tag{3. 2}$$

was transformed into a problem about short vectors in lattices. The lattice L we used to obtain the values of y which make each of the terms in (3. 1) small is generated by the columns $\underline{v}_1, \ldots, \underline{v}_{n+2}$ of the following $(n+2)\times(n+2)$ matrix (here $[x]$ means the greatest integer $\leqq x$):

$$\begin{array}{cccccc}
-[\alpha_1\psi_1 2^v] & [\alpha_1\gamma_1 2^{v-10}] & [2\pi\alpha_1 2^v] & 0 & & 0 \\
-[\alpha_2\psi_2 2^v] & [\alpha_2\gamma_2 2^{v-10}] & 0 & [2\pi\alpha_2 2^v] & & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
-[\alpha_n\psi_n 2^v] & [\alpha_n\gamma_n 2^{v-10}] & 0 & 0 & & [2\pi\alpha_n 2^v] \\
2^v n^4 & 0 & 0 & 0 & & 0 \\
0 & 1 & 0 & 0 & & 0
\end{array} \tag{3. 3}$$

where v is an integer (usually $2n \leqq v \leqq 4n$) and

$$\alpha_j = \left|\left(\frac{1}{2} + i\gamma_j\right)\zeta'\left(\frac{1}{2} + i\gamma_j\right)\right|^{-\frac{1}{2}}.$$

The L^3 algorithm produces a reduced basis $\underline{v}_1^*, \ldots, \underline{v}_{n+2}^*$ for the lattice L. This reduced basis usually contains some very short vectors. However, we are actually interested in the longest vector in the reduced basis. Since the reduced basis is a basis for L, it has to contain at least one vector $\underline{w}$ which has a nonzero coordinate in the $(n+1)$-st position. Since that coordinate is a multiple of $2^v n^4$, it is very large compared to all the other entries in the original basis, and this makes $\underline{w}$ quite long. Therefore in order to obtain a set of short basis vectors, a good basis transformation algorithm ought to contain exactly one vector $\underline{w}$ with a nonzero $(n+1)$-st coordinate, and that coordinate then has to be $\pm 2^v n^4$. As it turns out, in all the tests that we ran, the L^3 algorithm did indeed behave in this desirable fashion. Given that there is a single vector $\underline{w}$ in the reduced basis with nonzero $(n+1)$-st coefficient, which we may take to be $2^v n^4$ without loss of generality, its j-th coordinate for $1 \leqq j \leqq n$ equals

$$z[\alpha_j\gamma_j 2^{v-10}] - [\alpha_j\psi_j 2^v] - m_j[2\pi\alpha_j 2^v]$$

and the $(n+2)$-nd coordinate is z, for some integers $z, m_1, \ldots, m_n$. To minimize the length of $\underline{w}$, these all have to be small which means that all of the

$$z\alpha_j\gamma_j 2^{\nu-10} - \alpha_j\psi_j 2^{\nu} - m_j 2\pi\alpha_j 2^{\nu}$$

have to be small, so all of the

$$\beta_j = \alpha_j(\gamma_j y - \psi_j - 2\pi m_j)$$

have to be very small, where $y = \frac{z}{1024}$. In practice, the vectors $\underline{w}$ produced by the L^3 algorithm did indeed have this desired property.

The reason for the presence of the α_j in the basis is that we wish to make

$$\sum_{j=1}^{n} \alpha_j^2 \cos(\gamma_j y - \psi_j - 2\pi m_j)$$

large. Now if all of the $\gamma_j y - \psi_j - 2\pi m_j$ are small, this sum approximately equals

$$\sum_{j=1}^{n} \alpha_j^2 - \frac{1}{2}\sum_{j=1}^{n} \alpha_j^2 (\gamma_j y - \psi_j - 2\pi m_j)^2,$$

and we wish to have the second sum above small. That, however, corresponds to minimizing the euclidean norm of the vector $(\beta_1, \ldots, \beta_n)$, which is what the L^3 algorithm attempts to do.

In order to obtain values of y for which the chosen zeros contribute negative amounts, so that $h_K(y)$ will hopefully be negative, we used similar lattices. The only change was that the ψ_j were replaced by $\psi_j + \pi$.

The above discussion explains why we chose the lattice L the way we did. It is clear, though, that the choice was made on heuristic grounds, since the L^3 algorithm was not guaranteed to find the solutions we were looking for. In the end, though, that algorithm did fulfill our expectations and enabled us to disprove the Mertens conjecture.

4. Numerical computations

4. 1 Preliminary considerations

If the first 400, say, of the γ's are numbered $\gamma_1, \gamma_2, \ldots$ so that the quantities $|\rho_j\zeta'(\rho_j)|^{-1}$ for $\rho_j = \frac{1}{2} + i\gamma_j$ are decreasing, then

$$2\sum_{j=1}^{n} |\rho_j\zeta'(\rho_j)|^{-1}$$

exceeds 1 for $n \geqq 54$ and equals 1.0787... for $n = 70$. This suggested to us that a disproof of the Mertens conjecture might be obtained if we could use the L^3 algorithm to find a y that made each of the quantities η_j in (3. 1) quite small for $n = 70$. Any such value of y was likely to be quite large, since if we wish to make each of the $|\eta_j| \leqq \frac{\pi}{10}$, for example, than under the assumption that the γ_j behave like random numbers with respect to inhomogeneous diophantine approximation, we can expect that the smallest y that has the desired properties is of the order of 10^{70} in size. Therefore it was clear that the γ's had to be known with great accuracy. Moreover, the number T which governs the length of the finite sum $h_K(y)$ (cf. (2. 20)) should be so large that the cosine-values in that sum which come from the chosen 70 zeros (and so are close to 1) should have a weight factor $k(\gamma)$ which is also close to 1. We chose $T = 2515.286\ldots$, the height of the 2000-th zero, and the accuracy of the first 2000 γ's to be at least 100 decimal digits. As it turned out, we only needed about 75 digit accuracy, and with more careful choice of the parameters perhaps even less. Since the running time was not expected to be very high, however, we did not attempt to choose the most efficient set of parameters.

The function $k(t)$ used in our computations is of the form $k(t) = g\left(\frac{t}{T}\right)$, where $T = 2515.286\ldots$ is the height of the 2000-th zero and

$$(4.1) \qquad g(t) = \begin{cases} (1 - |t|) \cos(\pi t) + \pi^{-1} \sin(\pi |t|), & |t| \leqq 1, \\ 0, & |t| \geqq 1. \end{cases}$$

This function was introduced into the work on the Mertens conjecture by Jurkat and Peyerimhoff [21], and by Odlyzko in the work on discriminants of number fields (see [32]). What is needed in both contexts is a function $f(t)$ which has support in $[-1, 1]$, has nonnegative Fourier transform, and is as close to 1 as possible in a neighborhood of 0, since it is desired to make the contributions of the initial zeros (which are lined up by the inhomogeneous diophantine approximation algorithm) as large as possible. Among all such functions $f(t)$ with $f(0) = 1$, the minimum of $-f''(0)$ is attained by $f(t) = g(t)$. This was proved under some smoothness assumptions on $\hat{f}(u)$ by Jurkat and Peyerimhoff [21] and under somewhat different assumptions by Poitou [32]. However, this result follows in full generality from the work of Boas and Kac [5], who proved that all functions $f(t)$ with support in $[-1, 1]$, $f(0) = 1$, and nonnegative Fourier transforms satisfy $|f(u)| \leqq w(u)$ for $|u| < 1$, where

$$(4.2) \qquad w(u) = \cos \frac{\pi}{[|u|^{-1}] + 1}.$$

The bound $|f(u)| \leqq w(u)$ is best possible in the sense that for every u with $|u| < 1$, there is a function f satisfying all the required properties for which $|f(u)| = w(u)$, but there is no single function f for which equality holds for all $|u| < 1$. In applications to the disproof of the Mertens conjecture, the function $g(t)$ is somewhat better than the Féjer kernel (2. 19) used by Ingham, and not far from the bound $w(u)$. In fact, the sum

$$(4.3) \qquad 2 \sum_{\rho \in V} k(\gamma) |\rho \zeta'(\rho)|^{-1},$$

where V denotes the 70 zeros of the zeta function out of the first 400 with $\gamma>0$ for which $|\rho\zeta'(\rho)|^{-1}$ are largest, equals 1.0787... if $k(u)=1$, equals 1.0482... if $k(u)=1-\frac{|u|}{2500}$, equals 1.0524... if $k(u)=g\left(\frac{u}{2500}\right)$, and equals 1.0566... if $k(u)=w\left(\frac{u}{2500}\right)$. Thus even if we could find a better function f, this would not by itself improve our results by more than 0.5%. Finally, it is conceivable that one could obtain a slight improvement by using kernels $k(t)$ for which $K(y)$ is allowed to be negative, but that is unlikely, since we would then obtain bounds of the form

$$\limsup_{y\to\infty} |m(y)| \geqq |h_K(y)| \left(\int_{-\infty}^{\infty} |K(u)|du\right)^{-1},$$

and the fact that $\int |K(u)|du > k(0)$ would be working against us.

4. 2 Computation of the first 2000 γ's to at least 100 decimal digits

Experience with 28D-computation (i.e., 28 decimal digit computation) of the γ's was gained already in the work described in [33], [34]. The program for those computations, which was written in double precision FORTRAN for a CDC CYBER 73/173 computer, was converted to multiple-precision for a CDC CYBER 750 computer (which is about ten times as fast as the 73/173), with the help of Brent's multiple-precision package MP [7]. The array-length of the multiple-precision numbers corresponding to an accuracy of 100 decimal digits allowed us to obtain a slightly higher accuracy of 105 decimal digits, without extra computing costs. The γ's were computed with the Newton process, starting from the 28D values obtained in [34]. For $\zeta(s)$, we used the Euler-Maclaurin summation formula

$$\zeta(s)=\sum_{j=1}^{N-1} j^{-s}+\frac{1}{2}N^{-s}+\frac{N^{1-s}}{s-1}+\sum_{k=1}^{M} T_{k,N}(s)+E_{M,N}(s), \tag{4. 4}$$

where

$$T_{k,N}(s)=\frac{B_{2k}}{(2k)!}N^{1-s-2k}\prod_{j=0}^{2k-2}(s+j),$$

$\left(B_2=\frac{1}{6},\ B_4=-\frac{1}{30},\ldots \text{ are the Bernoulli numbers}\right)$ and

$$|E_{M,N}(s)| < \left| T_{M+1,N}(s)\,\frac{s+2M+1}{\mathrm{Re}(s)+2M+1}\right| \tag{4. 5}$$

for all $M\geqq 0$, $N\geqq 1$, and $\mathrm{Re}(s)>-(2M+1)$. By taking N and M large enough, this formula gives $\zeta(s)$ to any desired accuracy. The precise choice of N and M was derived from the following heuristic estimate of the error $|E_{M,N}(s)|$. The number $\frac{B_{2M+2}}{(2M+2)!}$ for large M is about $2(2\pi)^{-2M-2}$. Moreover, we took $s=\frac{1}{2}+it$, replaced

$$(s+j)\,(s+2M+1-j) \quad \text{by} \quad \left(s+M+\frac{1}{2}\right)^2$$

and assumed that $\frac{N^{\frac{1}{2}}}{M+\frac{3}{4}}<1$. In this way we obtained the approximate upper bound $\left(\frac{M+1+t}{2\pi N}\right)^{2M+2}$ for the error $|E_{M,N}(s)|$ in (4. 5). For this to be approximately equal to 10^{-A}, we must have

$$N \approx (2\pi)^{-1} 10^{\frac{A}{2M+2}} (M+1+t). \tag{4. 6}$$

This still leaves freedom to choose one of either N or M, given t and A. In our program, storage has to be reserved for the numbers $\log(j)$ and $j^{-\frac{1}{2}}$, $j=1, 2,\ldots, N$ and $\frac{B_{2j}}{(2j)!}$, $j=1, 2,\ldots, M$. N and M were chosen to satisfy (4. 6), given t and A, and such that the storage and computing costs were minimal. The precise derivation depends on the accounting formula of the computer used, and will be omitted. In Table 1, we give the numbers N and M as they were chosen for various values of $t=\gamma_j$ (where γ_j now denotes the j-th zero, in contrast to sections 3 and 4. 1) and $A=105$.

Table 1

Some values of N and M used in the computation of $\zeta(s)$ for $s=\frac{1}{2}+it$, $t=\gamma_j$

j	γ_j (approx.)	N	M
1	14.13	67	80
10	49.77	82	95
100	236.5	139	145
200	396.4	183	170
300	541.8	220	190
500	811.2	284	220
1000	1419.	418	275
1500	1981.	536	315
2000	2515.	646	345

During the computation of $\zeta\left(\frac{1}{2}+i\gamma_j\right)$ the actual error was checked by computing the quantity

$$\left|N^s T_{M,N}(s) \frac{s+2M-1}{\mathrm{Re}(s)+2M-1}\right| \tag{4. 7}$$

after the computation of the right hand side of (4. 4) without, of course, $E_{M,N}(s)$. Note that $T_{M,N}(s)$ is the last term of the second sum in (4. 4). In view of (4. 5), this quantity (4. 7) is a safe upper bound for the error committed in (4. 4). Its value was always smaller than 10^{-125}.

The 105D approximations of γ_j, $j=1,\ldots,2000$ were computed with the following Newton process $\left(\text{which used the fact that the zeros of } \zeta \text{ to be computed have real part } \frac{1}{2}\right)$:

$$\gamma_j^{(i+1)} = \gamma_j^{(i)} - \operatorname{Im} \frac{\zeta\left(\frac{1}{2}+i\gamma_j^{(i)}\right)}{\zeta'\left(\frac{1}{2}+i\gamma_j^{(i)}\right)}, \quad i=0,1,\ldots,$$

where for $\gamma_j^{(0)}$ we took the (about) 28D-approximation of γ_j from [34]. The value of the derivative of ζ was computed simultaneously with ζ from the derivative with respect to s of the right hand side of (4. 4). The iteration process was terminated as soon as the absolute value of the Newton correction term was smaller than 10^{-105}. This bound was achieved always after three or four iterations.

In the first Newton step, the values in (4. 4) of

$$j^{-\frac{1}{2}-i\gamma_j^{(0)}} = j^{-\frac{1}{2}}\{\cos(\gamma_j^{(0)}\log j) - i\sin(\gamma_j^{(0)}\log j)\}, \quad j=1,2,\ldots,N,$$

were computed with help of the cosine-routine MPCOS from [7] and with the $(1-\cos^2)^{\frac{1}{2}}$-formula (this turned out to be cheaper than using the MPSIN or the MPCIS-routines!). The cos- and sin-values were stored. In the next Newton-step we used the fact that already $\gamma_{(j)}^{(0)}$ was such a good approximation to γ_j that $\delta_j := \gamma_j^{(1)} - \gamma_j^{(0)}$ satisfies

$$(4.8) \qquad |\delta_j| \lesssim 10^{-28}|\gamma_j|.$$

This allowed us to compute $\cos\left(\gamma_j^{(1)}\log(j)\right)$ from the formula

$$\begin{aligned}\cos(\gamma_j^{(1)}\log j) &= \cos\left((\gamma_j^{(0)}+\delta_j)\log j\right)\\ &= \cos(\gamma_j^{(0)}\log j)\cos(\delta_j\log j) - \sin(\gamma_j^{(0)}\log j)\sin(\delta_j\log j),\end{aligned}$$

where $\cos(\delta_j\log j)$ and $\sin(\delta_j\log j)$ were very cheaply computed by using 3 resp. 2 terms of the series expansions of the cos- and the sin-functions. The $\sin(\gamma_j^{(1)}\log j)$-term was computed similarly. In this way the time needed for the second, third, etc., Newton step was only about $\frac{1}{5}$ of the time needed for the first step.

The numbers ψ_j and $|\rho_j\zeta'(\rho_j)|^{-1}$ were computed together with γ_j.

The computations were carried out on the CDC CYBER 750 computer system of SARA (Academic Computer Centre Amsterdam) and consumed about 40 hours CPU-time for the first 2000 zeros of ζ. All the computed quantities can be obtained on tape from the second author.

4. 3 Computations with the L^3 algorithm

The first author programmed the L^3 algorithm on the CRAY-1 computer at AT&T Bell Laboratories in Murray Hill and applied it to various numbers of zeros γ, in the way described in Section 3. In Table 2 we give 21 values of z $(=z(i),\ i=1, 2, \ldots, 21)$ obtained with the L^3 algorithm for various combinations of n and v. The fifth column gives the total contribution

$$2 \sum_{\rho \in V_n} \operatorname{Re} \frac{\exp\left(\dfrac{i\gamma z}{1024}\right)}{\rho \zeta'(\rho)},$$

where V_n is the set of n zeros of the zeta function out of the first 400 with $\gamma > 0$ for which the $|\rho\zeta'(\rho)|^{-1}$ are largest. The final column shows the maximum sum that is attainable, namely

$$2 \sum_{\rho \in V_n} |\rho\zeta'(\rho)|^{-1}.$$

It follows that $z(i)$ for $i = 14$, 15 and 21 are promising candidates for disproving the Mertens conjecture, the first two on the positive, the last one on the negative side.

The total time on the CRAY-1 was about 10 hours. Programming was in FORTRAN using the Brent MP package [7].

Table 2

i	$z(i)$	n	v	Total Contrib.	Maxim. Possible
1	1558740347670	20	50	0.654437	0.714787
2	11152996741251888040	20	65	0.698869	0.714787
3	3038088714793971066628	20	75	0.708483	0.714787
4	7884496876200	25	50	0.709868	0.780108
5	22512628597332611084	25	70	0.754569	0.780108
6	30423753191565158754037305	25	90	0.765205	0.780108
7	-35184499366749	30	50	0.685344	0.829928
8	-37766814051167995908	30	70	0.768417	0.829928
9	30173551610132642824712844	30	90	0.808450	0.829928
10	45811847622307	40	50	0.684944	0.910707
11	-27950995863621785302277841555417	40	110	0.867289	0.910707
12	-11547227804089875278215723173686203	40	120	0.884580	0.910707
13	-4948696983958480235838716850780012170210818	50	150	0.953197	0.977090
14	1174322091443909775800331523627991861701053793090667	60	180	0.983071	1.033091
15	-14382376632927229999913330309529375874458206207784691752925143820823	70	230	1.048646	1.078718
16	-64838544414151	25	50	-0.683989	0.780108
17	-36925065810626800521	25	70	-0.749338	0.780108
18	-6442382518920661025945199	30	90	-0.808116	0.829928
19	-42125186087757776278560297731223425	40	120	-0.864651	0.910707
20	15186801174602568083509602134954646365019863	50	150	-0.942832	0.977090
21	32867354391472799610613760190680378470192276009317440187192649371338	70	230	-1.029400	1.078718

Next, for all these 21 z-values a local maximum of $h_K(y)$ with $k(t)=g\left(\frac{t}{T}\right)$ as given in (4. 1), with y in the neighborhood of $\frac{z}{1024}$, and $T=\gamma_{2000}$, was computed. The results are given in Table 3.

Table 3

i	$y\left(\text{near } \frac{z(i)}{1024}\right)$	$h_K(y)$
1	1522207370.767659	0.641948
2	1089159838012878.942885	0.723633
3	29668835105409873 6.943226	0.757590
4	7699703980.663424	0.744898
5	21984988864582628.010504	0.753561
6	29710696476137850345739.553492	0.752062
7	-34359862662.834675	0.735926
8	-3688165434684374 5.996274	0.763238
9	29466358994270159008508.638083	0.789869
10	44738132443.656341	0.658563
11	-272958943980681497092557046 43.959226	0.905585
12	-1127658965243151882638254216 1802.931681	0.939414
13	-48327118983969533553112469245898556 34971.501850	0.978293
14	1146798917425693140430011253542960802442435344815.103296	0.996988
15	-14045289680592998046790361630399781127400591999789738039965960762.521505	1.061545
16	-63318891029.439972	-0.712989
17	-360596345806902 34.890490	-0.740998
18	-629138917863345803 3149.611855	-0.811204
19	-41137877038825953397031540753147.877254	-0.846405
20	148308605220728203940523458349166468 40839.710921	-0.925911
21	32097025772922655869740000186211307099797144540349062682805321651.697419	-1.009749

Consequently, the Mertens conjecture is false, as is shown on the positive side by the result on line 15 and on the negative side by the result on line 21.

5. Final Remarks

5. 1 Behavior of the function $h_K(y)$

The main reason the Mertens conjecture took so long to be disproved is that the functions $h_K(y)$ (and presumably also $h(y)$) are seldom large. Heuristics suggest that the sum

$$\sum_{\rho} \frac{1}{|\rho\zeta'(\rho)|^2}$$

converges, and numerical evidence suggests it converges to 0.029. Hence we can expect that the L_2-norms of $h(y)$ and $h_K(y)$ over large intervals might be on the order of 0.17. In fact, $h_K(y)$ is usually of about that size, exceeding even 0.5 very rarely. In Figures 1 and 2 we present graphs (on different scales) of $h_K(y)$ for y near to the value in Table 3 for $i=15$, which is the value that led to the lim sup $M(x)\,x^{-\frac{1}{2}} > 1.06$ result. Figure 1 shows just how atypical large values of $h_K(y)$ are.

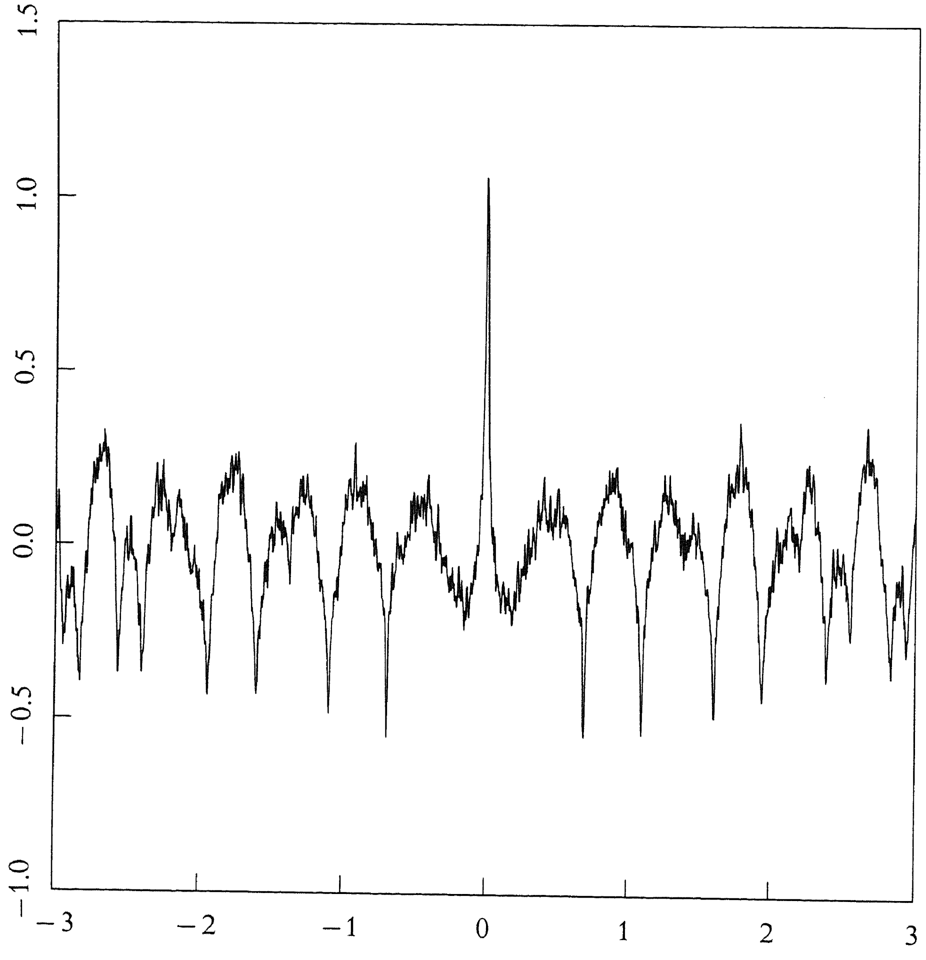

Fig. 1. Graph of the function $h_K(y_0+t)$, for y_0 given in the y-entry for $i=15$ in Table 3

The function $h_K(y)$ for the kernel $K(t)$ that we have been using is derived (at least for y large and positive) from averaging $M(u)\,u^{-\frac{1}{2}}$ over an infinite interval, but with most of the weight of the average concentrated on

$$x\left(1-\frac{1}{2500}\right) \leqq u \leqq x\left(1+\frac{1}{2500}\right), \quad x=\exp(y).$$

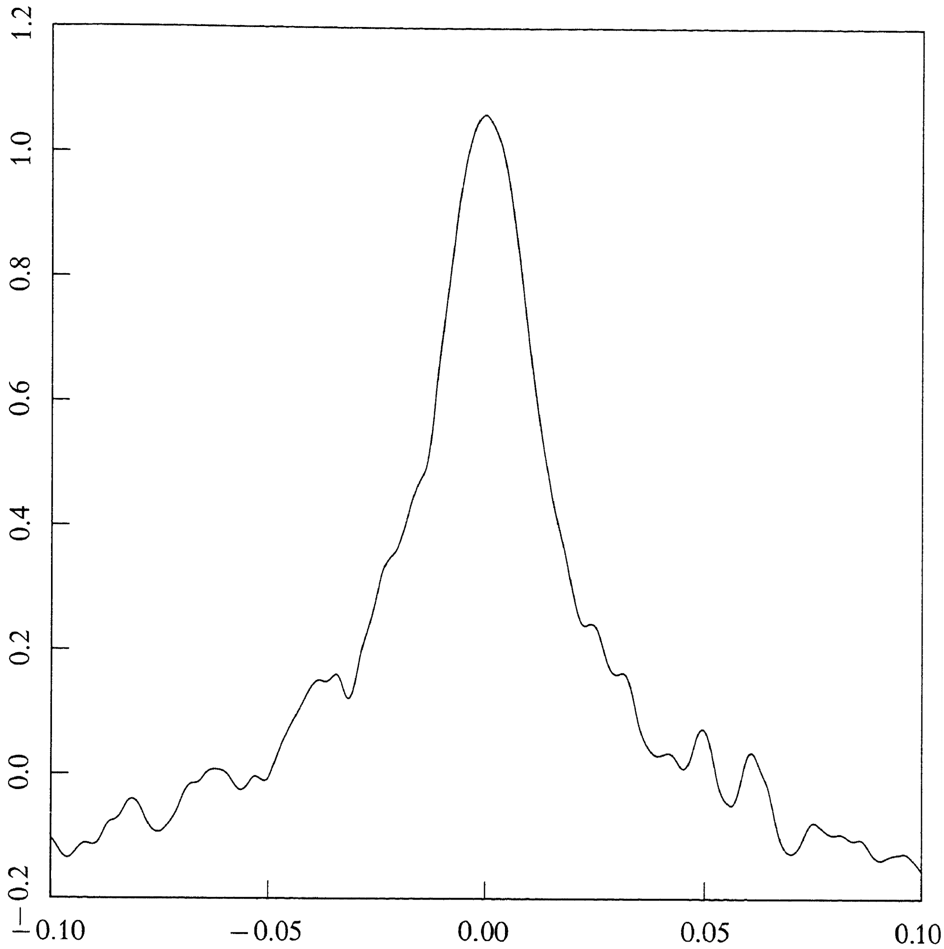

Fig. 2. Enlargement of the central part of Figure 1

The computations of $M(x)$ by Neubauer [29] and Yorinaga [45] show that in the ranges investigated, every large value of $m(y) = M(e^y)\, e^{-\frac{y}{2}}$ was part of a relatively long range of large values of $m(y)$. Therefore we might expect that as long as $x = \exp(y)$ does not get too big, $h_K(y)$ might provide a fairly good approximation to $m(y)$. In view of the computations of $h_K(y)$ by Spira [38] and Jurkat and Peyerimhoff [21], we therefore do not expect $|M(x)\, x^{-\frac{1}{2}}| > 1$ to occur for $x < 10^{20}$ and maybe not even for $x < 10^{30}$. For larger values of x, however, the interval over which $M(x)\, x^{-\frac{1}{2}}$ is being averaged to obtain $h_K(y)$ is so broad that our $h_K(y)$ may no longer be a good representation of $m(y)$.

5.2 Counterexamples to Mertens' Conjecture

Our results do not provide explicit counterexamples to the Mertens conjecture. However, since large values of $h_K(y)$ come from averages of $m(y)$, our results do suggest quite strongly that $M(x) < -x^{\frac{1}{2}}$ for some x close to $\exp(t_0)$, where t_0 is given by the y-entry for $i = 21$ in Table 3. Unfortunately we cannot compute any values of $M(x)$ in that range. R. S. Lehman [24] found an algorithm for computing $M(x)$ that takes $O(x^{\frac{2}{3}+\varepsilon})$ bit operations. A somewhat faster algorithm can be obtained by adapting the Lagarias-Odlyzko algorithm for computing $\pi(x)$ [22], but even that method requires on the order of $O(x^{\frac{3}{5}+\varepsilon})$ bit operations to compute $M(x)$. For x on the order of $\exp(10^{65})$, such algorithms are far too slow. It is probably possible to adapt the Lagarias-Odlyzko algorithm to produce approximations to $M(x)$ somewhat faster than in time $x^{\frac{3}{5}}$, but even such variations apparently would have running times that are fractional powers of x, and so again would be too slow. Therefore to be able to exhibit a specific example of $|M(x)| > x^{\frac{1}{2}}$, we have to find candidate values of x much smaller than $\exp(10^{65})$.

In the case of Pólya's conjecture, which states that the summatory function $L(x)$ of Liouville's function $\lambda(n)$ is $\leqq 0$ for $x \geqq 2$, the first disproof was achieved by Haselgrove [17], using Ingham's method [18]. That disproof did not provide a specific counterexample, since it basically showed that the function corresponding to our $h_K(y)$ (see Section 2) is >0 for some y. The value of y found by Haselgrove probably corresponds to violations of Pólya's conjecture, but it was too large to allow him to compute $L(x) = L(\exp(y))$ directly. Lehman [24] later found a specific counterexample by finding a much smaller value of y for which the function analogous to our $h_K(y)$ was negative but small and by actually computing $L(x)$ for x close to $\exp(y)$. A similar strategy might work for the Mertens conjecture. However, as was mentioned in Section 5. 1, it appears likely that no counterexamples occur for $x < 10^{20}$ and maybe not even for $x < 10^{30}$. Therefore this approach is not likely to be successful until much faster algorithms for computing $M(x)$ are found.

5. 3 Random behavior of zeros of the zeta function

Inspection of Tables 2 and 3 shows that in most of the cases that were tried, it was easier to obtain large positive values of $h_K(y)$ than large negative values. This situation is similar to that in the work of Jurkat and Peyerimhoff [21] and te Riele [33], who also obtain better bounds for $\limsup M(x)\, x^{-\frac{1}{2}}$ than for $\liminf M(x)\, x^{-\frac{1}{2}}$. Whether this phenomenon is due to chance or not is not clear. It is possible that there are some strange diophantine relations among the zeros which make it easier to find y that makes $h_K(y)$ large and positive, and that the phenomenon we are observing is due to the influence of such relations. Even if such relations exist, it is not clear whether their influence would still be noticeable if we were to work with much larger numbers of zeros.

There is an interesting conjecture about the random behavior of the zeros of the zeta function. It is derived from, and motivated by, the work of Montgomery [28], and it says that statistically, the zeros of the zeta function behave like eigenvalues of a random hermitian matrix of unitary type. There is substantial numerical evidence in favor of this conjecture [30]. This conjecture does not say much about the diophantine approximation properties of the zeros, but since it does predict that the spacings between the zeros ought to be more regular than in the case of numbers drawn uniformly and independently from an interval, it might help to explain why the sums of the form

$$\sum c_\gamma \gamma, \quad c_\gamma \in Z, \quad |c_\gamma| \text{ small},$$

that were investigated by Bateman et al. [4] often were quite small.

5.4 Possible further extensions

We have shown that $\limsup |M(x)|x^{-\frac{1}{2}} > 1.06$. What is generally expected, of course, is that the true value of this limes superior is $+\infty$. The method we use cannot in principle yield such a result, but it can almost certainly be used to improve on the 1.06 constant. The sum of $2|\rho\zeta'(\rho)|^{-1}$ over the best 100 zeros out of the first 1600 (i.e., the 100 zeros that give the largest contribution) is 1.18, over the best 200 zeros is 1.43, over the best 500 zeros is 1.77, and over the best 1000 zeros is 2.03. It appears therefore that with the method we have used we could hope to improve the 1.06 of our result to 1.5 with the use of hundreds of hours of time on computers that either already exist or are likely to become available in the near future. To reach 2, however, appears to require either special purpose processors or better inhomogeneous diophantine approximation algorithms.

References

[1] *R. J. Anderson*, On the Mertens conjecture for cusp forms, Mathematika **26** (1979), 236—249 and **27** (1980), 261.

[2] *R. J. Anderson*, A note on the Möbius sum function, in preparation.

[3] *R. J. Anderson* and *H. M. Stark*, Oscillation theorems, in: M. I. Knopp (ed.), Analytic Number Theory, Lecture Notes in Math. **899**, Berlin-Heidelberg-New York 1981, 79—106.

[4] *P. T. Bateman*, *J. W. Brown*, *R. S. Hall*, *K. E. Kloss* and *R. M. Stemmler*, Linear relations connecting the imaginary parts of the zeros of the zeta function, in: A. O. L. Atkin and B. J. Birch (editors), Computers in Number Theory, New York 1971, 11—19.

[5] *R. P. Boas, Jr.*, and *M. Kac*, Inequalities for Fourier transforms of positive functions, Duke Math. J. **12** (1945), 189—206.

[6] *H. Bohr* and *B. Jessen*, One more proof of Kronecker's Theorem, J. London Math. Soc. **7** (1932), 274—275.

[7] *R. P. Brent*, A Fortran multiple-precision arithmetic package, ACM Trans. Math. Software **4** (1978), 57—70.

[8] *H. Cohen*, Arithmétique et informatique, Astérisque **61** (1979), 57—61.

[9] *C. Couturier*, Problem 641, L'Intermédiaire des Mathématiciens **2** (1895), 314—315.

[10] *H. G. Diamond*, Two oscillation theorems, in: A. A. Gioia and D. L. Goldsmith (editors), The Theory of Arithmetic Functions, Lecture Notes in Math. **251**, Berlin-Heidelberg-New York 1972, 113—118.

[11] *H. M. Edwards*, Riemann's Zeta Function, New York 1974.

[12] *L. J. Goldstein*, A necessary and sufficient condition for the Riemann hypothesis for zeta functions attached to eigenfunctions of the Hecke operator, Acta Arith. **15** (1968/69), 205—215.

[13] *I. J. Good* and *R. F. Churchhouse*, The Riemann hypothesis and pseudorandom features of the Möbius sequence, Math. Comp. **22** (1968), 857—861.

[14] *E. Grosswald*, Oscillation theorems, in: A. A. Gioia and D. L. Goldsmith (editors), The Theory of Arithmetic Functions, Lecture Notes in Math. **251**, Berlin-Heidelberg-New York 1972, 141—168.

[15] *F. Grupp*, On the Mertens conjecture for cusp forms, Mathematika **29** (1982), 213—226.

[16] *J. Hadamard*, Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques, Bull. Soc. Math. France **24** (1896), 199—220.

[17] *C. B. Haselgrove*, A disproof of a conjecture of Pólya, Mathematika **5** (1958), 141—145.

[18] *A. E. Ingham*, On two conjectures in the theory of numbers, Amer. J. Math. **64** (1942), 313—319.

[19] *W. B. Jurkat*, Eine Bemerkung zur Vermutung von Mertens, Nachrichten der Österr. Math. Ges., Sondernummer Ber. V. Österr. Mathematikerkongress, Vienna 1961, 11.

[20] *W. B. Jurkat*, On the Mertens conjecture and related general Ω-theorems, in: H. Diamond (editor), Analytic Number Theory, Providence 1973, 147—158.

[21] *W. Jurkat* and *A. Peyerimhoff*, A constructive approach to Kronecker approximations and its application to the Mertens conjecture, J. reine angew. Math. **286/287** (1976), 332—340.

[22] *J. C. Lagarias* and *A. M. Odlyzko*, Computing $\pi(x)$: An analytic method, preprint.

[23] *J. C. Lagarias* and *A. M. Odlyzko*, Solving low-density subset sum problems, in: Proc. 24th Found. Computer Sci. Symp., New York 1983, 1—10, revised version to appear in J. ACM.

[24] *R. S. Lehman*, On Liouville's function, Math. Comp., **14** (1960), 311—320.
[25] *A. K. Lenstra, H. W. Lenstra, Jr.* and *L. Lovász*, Factoring polynomials with rational coefficients, Math. Ann. **261** (1982), 515—534.
[26] *J. van de Lune* and *H. J. J. te Riele*, On the zeros of the Riemann zeta function in the critical strip. IV, in preparation.
[27] *F. Mertens*, Über eine zahlentheoretische Funktion, Sitzungsberichte Akad. Wiss. Wien IIa **106** (1897), 761—830.
[28] *H. L. Montgomery*, On the pair-correlation of zeros of the zeta function, in: H. Diamond (editor), Analytic Number Theory, Providence 1973, 181—194.
[29] *G. Neubauer*, Eine empirische Untersuchung zur Mertensschen Funktion, Numer. Math. **5** (1963), 1—13.
[30] *A. M. Odlyzko*, Distribution of zeros of the zeta function, in preparation.
[31] *J. Pintz*, Oscillatory properties of $M(x) = \sum_{n \leq x} \mu(n)$. III, Acta Arith. **43** (1984), 105—113.
[32] *G. Poitou*, Sur les petits discriminants, Sém. Delange-Pisot-Poitou **18** (1976/77), Paris 1977.
[33] *H. J. J. te Riele*, Computations concerning the conjecture of Mertens, J. reine angew. Math. **311/312** (1979), 356—360.
[34] *H. J. J. te Riele*, Tables of the first 15000 zeros of the Riemann zeta function to 28 significant digits, and related quantities, Math. Centrum Report **NW 67/79**, Amsterdam 1979.
[35] *L. A. Rubel* and *E. G. Straus*, Special trigonometric series and the Riemann hypothesis, Math. Scand. **18** (1966), 35—44.
[36] *B. Saffari*, Sur les oscillations des fonctions sommatoires des fonctions de Möbius et de Liouville, C. R. Acad. Sc. Paris **A**, **271** (1970), 578—580.
[37] *B. Saffari*, Sur la fausseté de la conjecture de Mertens (with comment by Paul Lévy), C. R. Acad. Sc. Paris **A**, **271** (1970), 1097—1101.
[38] *R. Spira*, Zeros of sections of the zeta function. II, Math. Comp. **22** (1966), 163—173.
[39] *R. D. von Sterneck*, Empirische Untersuchung über den Verlauf der zahlentheoretischen Funktion $\sigma(n)$ im Intervalle von 0 bis 150000, Akad. d. Wiss. Wien, Math.-Nat. Cl. IIa **106** (1897), 835—1024.
[40] *R. D. von Sterneck*, Empirische Untersuchung über den Verlauf der zahlentheoretischen Funktion $\sigma(n)$ im Intervalle 150000 bis 500000, Akad. d. Wiss. Wien, Math.-Nat. Cl. IIa **110** (1901), 1053—1102.
[41] *R. D. von Sterneck*, Die zahlentheoretische Funktion $\sigma(n)$ bis zur Grenze 5000000, Akad. d. Wiss. Wien. Math.-Nat. Cl. IIa **121** (1912), 1083—1096.
[42] *R. D. von Sterneck*, Neue empirische Daten über die zahlentheoretischen Funktion $\sigma(n)$, Proc. 5th Int. Congress of Math., Cambridge 1913, 341—343.
[43] *T. J. Stieltjes*, Lettre à Hermite de 11 juillet 1885, Lettre 79, in: B. Baillaud et H. Bourget, Correspondance d'Hermite et Stieltjes, Paris 1905, 160—164.
[44] *E. C. Titchmarsh*, The Theory of the Riemann Zeta-Function, Oxford 1951.
[45] *M. Yorinaga*, Numerical investigation of sums of the Möbius function, Math. J. Okayama Univ. **21** (1979), 41—47.

AT & T Bell Laboratories, Murray Hill, N. J. 07974, USA
Centre for Mathematics and Computer Science, Kruislaan 413, NL-1098 SJ Amsterdam

Eingegangen 19. Juni 1984