Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Number Theory 98 (2003) 10–33

JOURNAL OF Number Theory

http://www.elsevier.com/locate/jnt

Short effective intervals containing primes

Olivier Ramaré[a,*] and Yannick Saouter[b]

[a] UMR 8524, Université Lille I, 59 655 Villeneuve d'Ascq Cedex, France
[b] IRIT, Batiment 1R1, 118 route de Narbonne, 31062 Toulouse Cedex 4, France

Received 9 March 2000; revised 20 March 2002

Abstract

We prove that every interval $]x(1-\Delta^{-1}), x]$ contains a prime number with $\Delta = 28\,314\,000$ and provided $x \geqslant 10\,726\,905\,041$. The proof combines analytical, sieve and algorithmical methods.

MSC: 11N05

Keywords: Distribution of primes

1. Introduction

We consider the determination of numerical intervals containing at least one prime number. The history of this problem can be divided into three parts: asymptotical properties, conjectures (mainly results assuming Riemann hypothesis) and numerical results, a part which we shall see as bridging the other two.

The story seems to start in 1845 when Bertrand conjectured after numerical trials that the interval $]n, 2n-3]$ contains a prime for $n \geqslant 4$. This was proved by Čebyšev in 1852 in a famous work where he got the first good quantitative estimates for the number of primes less than a given bound, say x. By now, analytical means combined with sieve methods (and the joint efforts of Baker et al. [4]) ensures us that each of the intervals $[x, x+x^{0.525}]$ for $x \geqslant x_0$ contains at least one prime. This statement concerns only for the (very) large integers.

*Corresponding author.

E-mail addresses: ramare@agat.univ-lille1.fr (O. Ramaré), saouter@enst-bretagne.fr (Y. Saouter).

URLs: http://www.informatik.uni-giessen.de/staff/richstein/ca/Goldbach.html, http://www.utm.edu/research/primes/notes/gaps.html, http://www.lynchburg.edu/public/academic/math/nicely/gaps/gaps.htm.

PII: S0022-314X(02)00029-X

It falls very close to what we can get under the assumption of the Riemann hypothesis: the interval $[x - K\sqrt{x}\log x, x]$ contains a prime, where K is an effective large constant and x is sufficiently large (cf. [25] for an account on this subject and Theorem 1). A theorem of Schoenfeld [18] also tells us that the interval

$$[x - \sqrt{x}\log^2 x/(4\pi), x] \tag{1}$$

contains a prime for $x \geqslant 599$ under the Riemann hypothesis. These results are still far from the conjecture Cramer [8] made in 1936 on probabilistic grounds: the interval $[x - K\log^2 x, x]$ contains a prime for any $K > 1$ and $x \geqslant x_0(K)$. Note that this statement has been proved for almost all intervals in a quadratic average sense by Selberg [19] in 1943 assuming the Riemann hypothesis and replacing K by a function $K(x)$ tending arbitrarily slowly to infinity.

From a numerical point of view, the Riemann hypothesis is known to hold up to a very large height. It is possible to use this fact to obtain estimates for $\pi(x)$, a development initiated by Rosser [21] in 1941, followed by Rosser and Schoenfeld's improvements [18,22] in 1975–1976. Wedeniwski [24] announced in November 2001 that he and others have verified the Riemann hypothesis up to height 3.33×10^9, thus extending the work of Van de Lune et al. [23] who had conducted such a verification in 1986 till height 5.45×10^8. Since then, the height 6.75×10^9 has been reached but we will refrain from using this result since it has been subject to no formal announcement. However updating our Table 1 below is a simple matter (see Theorem 4). Such computational results are substantially better than what was available to Rosser and Schoenfeld, namely that the first 3.6×10^6 zeroes verify the Riemann hypothesis. Furthermore, Rosser and Schoenfeld concentrated on proving effective bounds for $\pi(x)$, while we are interested only in proving that short intervals

Table 1

$\log x_0$	Δ	α_1	α_2	α_4	a	m	T
46	81 353 847	0.81	11	22/40	0.39890	48	1 361 250 000
47	127 680 085	0.71	12	22/40	0.40890	48	1 930 500 000
48	160 366 248	0.47	19	22/40	0.42373	48	2 095 500 000
49	178 274 183	0.28	31	22/40	0.43200	48	2 062 500 000
50	190 341 073	0.17	51	23/40	0.43606	48	2 062 500 000
51	197 073 687	0.105	85	23/40	0.43948	50	2 095 500 000
52	201 243 345	0.065	135	23/40	0.44187	52	2 128 500 000
53	204 879 661	0.040	217	24/40	0.44320	53	2 161 500 000
54	206 405 270	0.024	355	24/40	0.44366	53	2 145 000 000
55	207 313 717	0.016	573	24/40	0.44492	55	2 343 000 000
56	207 833 950	0.010	943	24/40	0.44507	55	2 409 000 000
57	208 988 147	0.006	1547	25/40	0.44522	55	2 392 500 000
58	209 026 205	0.004	2548	25/40	0.44522	55	2 607 000 000
59	209 257 759	0.002	4199	25/40	0.44529	55	2 178 000 000
60	209 267 308	0.002	6920	25/40	0.44529	55	3 300 000 000
150	212 215 384	0.00033	23×10^{23}	34/40	0.44536	55	3 300 000 000

contain primes. This leads to additional improvements. The third ingredient we use is new and comes from sieve theory via the Brun–Titchmarsh inequality. To explain it roughly, let us say that we estimate from below a sum over primes $\sum_p f(p/x)$ where f is a non-negative smooth function with support a compact interval containing 1, and we seek to find the smallest such interval for which a positive lower bound is obtainable. The smoothness implies that f and a good number of its derivatives are 0 on the boundary of this interval which in turn implies that f will be small near to this boundary. Analytical methods can however not tell us that all the primes detected by this lower estimate are not clustering precisely near this boundary (since we would otherwise choose a smaller interval) but sieve methods can do that up to some (here) minor loss. With the test function we chose this phenomenon is at its peak when x is large when compared to the height (say T_0) up to which the Riemann hypothesis is known to hold. Numerically, the sieve argument shortens the interval by a factor 5 while asymptotically in T_0 it yields

$$\max\{\varDelta \geqslant 1 \mid \forall x \geqslant (T_0 \log T_0)^2,\ \exists p \in \,]x(1-\varDelta^{-1}), x]\} \gg T_0(\log T_0)^{-1/2}. \qquad (2)$$

A similar approach but without the sieve argument gives the lower bound $T_0(\log T_0)^{-1}$. For so small x's, the infinite zero-free region is of course of no use. In order to offer a comparison with (1), we mention the following:

Theorem 1. *Under the Riemann hypothesis, the interval* $]x - \frac{8}{5}\sqrt{x}\log x, x]$ *contains a prime for* $x \geqslant 2$.

Let us recall here that a second line of approach following the original work of Čebyšev is still under examination though it does not give results as good as analytical means (see [7] for the latest result).

We prove the following:

Theorem 2. *Let* x *be a real number* $\geqslant x_0 \geqslant 10^{20}$. *Then the interval* $]x(1-1/\varDelta), x]$ *contains at least one prime, where* $\varDelta$ *is a function of* x_0 *defined by Table* 1.

Note that $\log(10^{20}) = 46.051\ldots$ and that all prime gaps have been computed up to 10^{15} by Nicely [15], extending a result of Young and Potler [26]. The other parameters given by Table 1 are explained in Section 4.

We thus have to cover the range 10^{15}–10^{20}, which we did via algorithmic means. In fact, we covered the whole range 10^{10}–10^{20}, so as to offer an independant verification. The problem of the prime certificate is a crucial issue since, while general primality checkers can easily establish or disprove the primality of numbers having 20 decimal digits, they are much too slow for our purpose. Indeed, approximatively 400 millions of prime numbers were necessary for the derivation of Theorem 3 and each one would have required about 1 s with the elliptic curve primality checker ECPP [2], amounting to a quite unrealistic global running time of about 10 years. Other methods lead to a similar situation as explained in Section 7.

We thus decided to use prime generation techniques [12]: we only look at families of numbers whose primality can be established with one or two Fermat-like or Pocklington's congruences. This kind of technique has been already used in a quite similar problem [17]. The generation technique we use relies on a theorem proven by Brillhart, Lehmer, and Selfridge [6] and enables us to generate dense enough families for the upper part of the range to be investigated. For the remaining range, we use theorems of Jaeschke [11] that yield a fast primality test (for this limited range). Finally, we obtain:

Theorem 3. *Let x be a real number larger than* 10 726 905 041. *Then the interval*

$$\left]x\left(1-\frac{1}{28\,314\,000}\right),x\right]$$

contains at least one prime.[1]

Our theorem should be compared with Schoenfeld's result [18] which states that the interval $]x, x(1+1/16\,597)]$ contains a prime, for $x \geqslant 2\,010\,760$. Also note that the largest prime gap before 10 726 905 041 (which is itself a prime) is 381 by Young and Potler's result and that $10\,726\,905\,041/28\,314\,000 = 378.8\ldots$.

The value 10 726 905 041 in the corollary is optimal. In case the reader would want another kind of interval, let us mention that the above theorem says also that the interval $]x, x(1+1/28\,313\,999)]$ contains a prime for $x \geqslant 10\,726\,905\,041$.

As a corollary, and recalling that Richstein [20] extending a work of Deshouillers, te Riele and Saouter proved that every even integer not more than 4×10^{14} is a sum of two primes at most, we get:

Corollary 1. *Every odd integer* >1 *not more than* 1.13256×10^{22} *is a sum of at most three primes.*

We deduce this corollary from the preceding theorem by using a greedy algorithm. This result is an improvement of [17] where the second named author with co-authors proved a similar statement to hold true, but with 10^{20} instead of 1.13256×10^{22}.

As usual, we define for any real number X

$$\psi(X) = \sum_{n \leqslant X} \Lambda(n), \text{ where } \Lambda \text{ is Von Mangold's function,}$$

$$\vartheta(X) = \sum_{p \leqslant X} \log p, \text{ where } p \text{ denotes a prime,}$$

$$\rho = \beta + i\gamma \text{ is a non-trivial zero of } \zeta(s) \text{ i.e. with } 0 < \beta < 1.$$

[1] The reader may wonder on the link between 28 314 000 and 81 353 847 given by Table 1. A first version of this paper relying only on [23] had the first value instead of the second one, which explains why the algorithmic part was driven as it was.

An unusual notation: We assume the Riemann hypothesis to hold up to T_0. By Wedeniwski [24], we can take $T_0 = 3.33 \times 10^9$, but we shall only use $T_0 = 3.3 \times 10^9$ in numerical applications. However Theorem 4 is valid for any T_0.

2. Lemmas

We set

$$N(T) = \sum_{\substack{\rho \\ 0<\gamma\leqslant T}} 1. \tag{3}$$

In the next two lemmas, by $f(x) = \mathcal{O}^*(g(x))$ we mean $|f(x)| \leqslant g(x)$.

Lemma 1. *If T is a real number $\geqslant 10^3$ then*

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + \frac{7}{8} + \mathcal{O}^*\left(0.67\log\frac{T}{2\pi}\right).$$

Proof. This follows easily from Theorem 19 of [21]. □

Lemma 2. *If $m \geqslant 1$ and $T \geqslant 10^3$ then*

$$\sum_{\substack{\rho \\ |\gamma|>T}} \frac{1}{|\gamma|^{m+1}} = \frac{1}{m\pi T^m}\left(\log(T/2\pi) + \frac{1}{m}\right) + \mathcal{O}^*\left(\frac{1.34}{T^{m+1}}(2\log(T/2\pi)+1)\right).$$

Proof. We have

$$\begin{aligned}\sum_{\substack{\rho \\ |\gamma|>T}} \frac{1}{|\gamma|^{m+1}} &= 2\sum_{\substack{\rho \\ \gamma>T}} \frac{1}{\gamma^{m+1}} \\ &= 2\sum_{\substack{\rho \\ \gamma>T}} (m+1)\int_\gamma^\infty \frac{dt}{t^{m+2}} \\ &= 2(m+1)\int_T^\infty N(t)\frac{dt}{t^{m+2}} - 2\frac{N(T)}{T^{m+1}},\end{aligned}$$

and an appeal to Lemma 1 concludes. □

We set

$$s_m(T) = \left\{\frac{1}{m\pi}\left(\log(T/2\pi) + \frac{1}{m}\right) + \frac{1.34}{T}(2\log(T/2\pi) + 1)\right\}\{1 + 10^{-5}\}; \tag{4}$$

the factor $1 + 10^{-5}$ is explained in the next lemma.

Lemma 3. *If $m \geqslant 1$, $T \geqslant 10^3$ and $x \geqslant 10^5$ then*

$$\sum_{\substack{\rho \\ |\gamma| > T}} \frac{x^\beta}{|\gamma|^{m+1}} \leqslant s_m(T)\frac{\sqrt{x}}{T^m} + s_m(T_0)\frac{x}{2T_0^m}.$$

Note: Though the assumption $T \leqslant T_0$ is not necessary, the lemma has been patterned for this case.

Proof. The functional equation implies that if $\beta + i\gamma$ is a non-trivial zero of ζ, then so is $1 - \beta + i\gamma$. Hence

$$\sum_{\substack{\rho \\ |\gamma| > T}} \frac{x^\beta}{|\gamma|^{m+1}} = \frac{1}{2}\sum_{\substack{\rho \\ |\gamma| > T}} \frac{x^\beta + x^{1-\beta}}{|\gamma|^{m+1}}.$$

Now when $\gamma > T_0$, we use the inequality $x^\beta + x^{1-\beta} \leqslant 1 + x \leqslant (1 + 10^{-5})x$ for $x \geqslant 1$ and Lemma 2 concludes. □

Lemma 4. *Let g be a continuously differentiable function on $[a, b]$ with $2 \leqslant a \leqslant b < \infty$. We have*

$$\begin{aligned}\int_a^b \psi(t)g(t)\,dt = \int_a^b tg(t)\,dt - \sum_\rho \int_a^b \frac{t^\rho}{\rho}g(t)\,dt \\ + \int_a^b \left(\log 2\pi - \frac{1}{2}\log(1 - t^{-2})\right)g(t)\,dt.\end{aligned}$$

Proof. It is enough to prove this lemma when no integer lies between a and b, a hypothesis we shall henceforth make. We recall that for $a < y < b$ and $T > 2$,

$$\begin{aligned}\psi(y) = y - \sum_{\substack{\rho \\ |\gamma| \leqslant T}} \frac{y^\rho}{\rho} + \log 2\pi - \frac{1}{2}\log(1 - y^{-2}) \\ + \mathcal{O}\left(\frac{y\log^2 yT}{T} + \frac{y\log y}{\langle y\rangle T}\right),\end{aligned} \tag{5}$$

where $\langle y \rangle = \min(y-a, b-y)$ (see [9, Chapter 17, formulae (9) and (10)]). We have for $0 < \varepsilon < (b-a)/2$

$$\begin{aligned}\int_{a+\varepsilon}^{b-\varepsilon} \psi(t)g(t)\,dt = \int_{a+\varepsilon}^{b-\varepsilon} tg(t)\,dt - \sum_{\substack{\rho \\ |\gamma| \leqslant T}} \int_{a+\varepsilon}^{b-\varepsilon} \frac{t^\rho}{\rho} g(t)\,dt \\ + \int_{a+\varepsilon}^{b-\varepsilon} \left(\log 2\pi - \frac{1}{2}\log(1-t^{-2})\right) g(t)\,dt \\ + \mathcal{O}\left(\frac{\log^2 T}{T} + \frac{\log(1/\varepsilon)}{T}\right),\end{aligned}$$

where the error term depends on a, b and g. Now an integration by parts yields

$$\int_{a+\varepsilon}^{b-\varepsilon} t^\rho g(t)\,dt \ll 1/|\rho|$$

uniformly in ε. Our lemma follows by letting T tend to ∞ and then ε to 0. □

Lemma 5. *Let X, u, δ be real numbers satisfying*

$$X \geqslant 10^{12}, \quad 0 \leqslant \delta \leqslant 0.0001, \quad 0 \leqslant u \leqslant 0.0001.$$

We have for all t in $[0, 1]$

$$\vartheta(Xe^u(1+\delta t)) - \vartheta(X(1+\delta t)) \leqslant 2.0004\, uX \frac{\log X}{\log(uX)}.$$

Proof. Define $Y = X(1+\delta t)$. The Brun–Titchmarsh inequality of Montgomery and Vaughan [14, Theorem 2] gives us

$$\begin{aligned}\vartheta(Ye^u) - \vartheta(Y) &\leqslant 2\log(Ye^u)\frac{Y(e^u-1)}{\log(Y(e^u-1))} \\ &\leqslant 2.00012\, Yu \frac{u + \log Y}{\log(e^u-1) + \log Y} \\ &\leqslant 2.00012\, Yu \frac{u + \log X}{\log(e^u-1) + \log X} \\ &\leqslant 2.00012\, Yu \left(1 + \frac{u - \log(e^u-1)}{\log(e^u-1) + \log X}\right) \\ &\leqslant 2.0004\, Yu \left(1 - \frac{\log u}{\log(u) + \log X}\right)\end{aligned}$$

since $e^u - 1 \leqslant 1.00006u$ and $u - \log(e^u - 1) \leqslant -1.00005 \log u$ for $u \in]0, 0.0001]$. This proves Lemma 5. □

3. The general principle

Let m be a positive integer. A function f over $[0, 1]$ is called m-admissible if the following properties hold:

(1) f is a m-times differentiable function.
(2) $f^{(k)}(0) = f^{(k)}(1) = 0$ for $0 \leqslant k \leqslant m-1$.
(3) $f \geqslant 0$.
(4) f is non-identically zero.

For such a function, we define

$$\mu_m^*(f) = ||f^{(m)}||_1/||f||_1, \quad \mu_m(f) = ||f^{(m)}||_2/||f||_1, \tag{6}$$

where as usual

$$||g||_1 = \int_0^1 |g(t)|\, dt \quad \text{and} \quad ||g||_2 = \left(\int_0^1 |g(t)|^2\, dt\right)^{\frac{1}{2}}. \tag{7}$$

We shall use

$$v(f, a) = \int_0^a f(t)\, dt + \int_{1-a}^1 f(t)\, dt$$

for $0 \leqslant a \leqslant 1/2$, as well as

$$w(f, a) = v(f, a)/||f||_1. \tag{8}$$

We have the following result:

Theorem 4. *Let m be an integer $\geqslant 2$. Let $\alpha_1 > 0$, $\alpha_2 \geqslant 1$, $\alpha_4 \in [\frac{1}{2}, 1]$ and $a \in [0, \frac{1}{2}]$ be four real parameters. Let us select an m-admissible function f. We set*

$$\alpha_3^{-1} = \left(0.99 - 1.0007\alpha_1 - \frac{2.001}{\alpha_4} w(f, a) - 0.0326/\alpha_2\right)/2.0012$$

and assume it to be positive. Let $Y \geqslant 2 \times 10^{12}$ and $T \geqslant 10^3$ be real numbers such that

$$2N(T) \leqslant \alpha_1 \sqrt{Y}.$$

We put

$$u = \frac{1.0001}{T}\left(\alpha_3 \frac{\mu_m^*(f_m)}{m^m} T\left(\frac{s_m(T)}{\sqrt{Y}} + \frac{s_m(T_0)}{2}\left(\frac{T}{T_0}\right)^m\right)\right)^{\frac{1}{m+1}}$$

and we assume that

$$mu \leqslant 0.0001, \quad \alpha_2 \leqslant u\sqrt{Y}, \quad Y^{\alpha_4} \leqslant uY, \quad T \leqslant T_0.$$

We have

$$\vartheta(Y) - \vartheta\left(Y\frac{1+mua}{1+mu(1-a)}e^{-u}\right) \geqslant \frac{uY}{101(1-w(f,a))}.$$

Notice that if parameters α_1, α_2, a and T satisfy the hypothesis of the above theorem for a given value of Y, then they satisfy it also for any larger value of Y and u is decreased.

Proof. We put $\delta = mu$ and $e^u X(1+\delta(1-a)) = Y$ so that $X \geqslant 0.999Y$. We also set

$$Y' = Y\frac{1+mua}{1+mu(1-a)}e^{-u}. \tag{9}$$

The proof begins with

$$\int_a^{1-a} (\vartheta(e^u(1+\delta(1-a))X) - \vartheta(X(1+\delta a)))f(t)\,dt \geqslant \int_a^{1-a} (\vartheta(e^u(1+\delta t)X) - \vartheta((1+\delta t)X))f(t)\,dt.$$

We extend the latter integral to $[0,1]$ by using Lemma 5 getting

$$(||f||_1 - v(f,a))(\vartheta(Y) - \vartheta(Y')) \geqslant \int_0^1 (\vartheta(e^u(1+\delta t)X) - \vartheta((1+\delta t)X))f(t)\,dt - 2.0004uXv(f,a)\frac{\log X}{\log(uX)}.$$

We then use $\log X/\log(uX) \leqslant \log Y/\log(uY)$ and the inequality involving α_4 to get the lower bound

$$(||f||_1 - v(f,a))(\vartheta(Y) - \vartheta(Y')) \geqslant \int_0^1 (\vartheta(e^u(1+\delta t)X) - \vartheta((1+\delta t)X))f(t)\,dt - \frac{2.001}{\alpha_4}uXv(f,a). \tag{10}$$

Let us call I this integral divided by $||f||_1$ and J a similar expression but with ϑ replaced by ψ. We go from ψ to ϑ by using the following estimates taken from Schoenfeld's work [18]:

$$\begin{cases} \psi(Z) - \vartheta(Z) \geqslant 0.998\,697\sqrt{Z} & (Z \geqslant 121), \\ \psi(Z) - \vartheta(Z) \leqslant 1.001\,093\sqrt{Z} + 3Z^{1/3} & (Z > 0), \end{cases} \tag{11}$$

so that for $t \in [0, 1]$:

$$\begin{aligned} &\vartheta(e^u(1+\delta t)X) - \vartheta((1+\delta t)X) - (\psi(e^u(1+\delta t)X) - \psi((1+\delta t)X)) \\ &\quad \geqslant -\sqrt{X}(1.001\,093\sqrt{e^u(1+\delta)} + 3(e^u(1+\delta))^{1/3}X^{-1/6} - 0.998\,697) \\ &\quad \geqslant -0.0325\sqrt{X}. \end{aligned}$$

Thus

$$I \geqslant J - 0.0325\sqrt{X} \tag{12}$$

and we are left with evaluating J. We apply Lemma 4 and the elementary inequality

$$\begin{aligned} &\log\left(1 - \frac{1}{(e^u(1+\delta t)X)^2}\right) - \log\left(1 - \frac{1}{((1+\delta t)X)^2}\right) \\ &\quad \leqslant \left(1 - \frac{1}{(e^u(1+\delta t)X)^2}\right) \Big/ \left(1 - \frac{1}{((1+\delta t)X)^2}\right) - 1 \leqslant \frac{u}{X} \end{aligned}$$

to get

$$J \geqslant u\left(X - \frac{1}{X}\right) - \frac{1}{||f||_1} \sum_\rho \frac{e^{u\rho} - 1}{\rho} X^\rho \int_0^1 (1+\delta t)^\rho f(t)\, dt. \tag{13}$$

To treat the sum over the zeroes, we distinguish whether $|\gamma|$ is larger or smaller than T. When $|\gamma| \leqslant T$, we have $\beta = \frac{1}{2}$ and we use

$$\left|\int_0^1 (1+\delta t)^\rho f(t)\, dt\right| \leqslant (1+\delta)^{1/2}||f||_1,$$
$$\left|\frac{e^{u\rho} - 1}{\rho}\right| = \left|\int_0^u e^{x\rho}\, dx\right| \leqslant ue^{u/2} \leqslant u(1+\delta)^{1/2}$$

while, if $|\gamma| > T$, m integrations by parts give

$$\int_0^1 (1+\delta t)^\rho f(t)\, dt = \frac{(-1)^m}{(\rho+1)\dots(\rho+m)\delta^m} \int_0^1 (1+\delta t)^{\rho+m} f^{(m)}(t)\, dt$$

from which we infer

$$\left|\int_0^1 (1+\delta t)^\rho f(t)\,dt\right| \leqslant \frac{(1+\delta)^{m+1}}{|\gamma|^m\delta^m}\,||f^{(m)}||_1. \tag{14}$$

By (13) we have thus proved that $J/(uX)$ is not less than

$$1-\frac{1}{X^2}-(1+\delta)\frac{2N(T)}{\sqrt{X}}-(1+\delta)^{m+1}\mu_m^*(f)\frac{e^u+1}{u\delta^m}\sum_{\substack{\rho\\ |\gamma|>T}}\frac{X^{\beta-1}}{|\gamma|^{m+1}}. \tag{15}$$

Using Lemma 3, the bound for u and the hypothesis on α_1, we can replace this lower bound by

$$1-\frac{1}{X^2}-\frac{1+\delta}{\sqrt{0.999}}\alpha_1-(1+\delta)^{m+1}\mu_m^*(f)\frac{2.0001}{u\delta^m}\left(\frac{s_m(T)}{\sqrt{X}T^m}+\frac{s_m(T_0)}{2T_0^m}\right). \tag{16}$$

We infer that $I/(uX)$ is not less than

$$\begin{aligned}&1-\frac{1}{X^2}-1.0007\alpha_1-\frac{0.0325}{u\sqrt{0.999Y}}\\&-\frac{2.0001}{\sqrt{0.999}}\frac{\mu_m^*(f)}{m^m}\left(\frac{1.0001}{uT}\right)^{m+1}\left(\frac{s_m(T)T}{\sqrt{Y}}+\frac{s_m(T_0)T^{m+1}}{2T_0^m}\right).\end{aligned} \tag{17}$$

Recalling the hypothesis on α_2, (10) and the definition of u, we get

$$\begin{aligned}(1-w(f,a))(\vartheta(Y)-\vartheta(Y'))&\geqslant I-\frac{2.001}{\alpha_4}uXw(f,a)\\&\geqslant uX\left(1-\frac{2.001}{\alpha_4}w(f,a)-1.0007\alpha_1-\frac{0.0326}{\alpha_2}-\frac{2.0012}{\alpha_3}\right)-\frac{u}{X}\end{aligned} \tag{18}$$

and this lower bound is at least $(uX/101)$ by the hypothesis on α_3. This proof implies that $1-w(f,a)>0$ (it follows from $\alpha_3>0$). □

4. The test-function

We now have to choose the m-admissible function in order to apply Theorem 4. We choose

$$f(t)=f_m(t)=(4t(1-t))^m \tag{19}$$

which can be shown to be optimal among the $(2m+1)$-times differentiable functions when we impose $a=0$ (we however do not make any claim). We have

Lemma 6.

$$||f_m||_1 = \frac{2^{2m}m!^2}{(2m+1)!}, \quad ||f_m^{(m)}||_2 = 2^{2m}\frac{m!}{\sqrt{2m+1}}, \tag{20}$$

$$\mu_m^*(f_m) \leqslant \mu_m(f_m) = \frac{(2m+1)!}{m!\sqrt{2m+1}} \leqslant \sqrt{4m+2}\, e^{\frac{1}{24m}}(4m/e)^m, \tag{21}$$

$$(2m+1)v(f_m,a) = 2mv(f_{m-1},a) + (2a-1)f_m(a). \tag{22}$$

Proof. The first result is given by integration by parts. The second one follows from

$$\int_0^1 f_m^{(m)}(t)^2\,dt = (-1)^m \int_0^1 f_m(t) f_m^{(2m)}(t)\,dt. \tag{23}$$

The third one follows by using the estimates

$$(n/e)^n\sqrt{2\pi n} \leqslant n! \leqslant (n/e)^n \sqrt{2\pi n}\, e^{\frac{1}{12n}} \tag{24}$$

valid for $n \geqslant 1$. Finally the recurrence formula for $v(f_m,a)$ comes from the relation $(2m+1)f_m = 2mf_{m-1} + ((t-\frac{1}{2})f_m)'$. □

Remarks concerning Tables 1 and 2. (1) For $46 \leqslant \log X \leqslant 60$, we have given the values of all the parameters we chose so as to make our results easily verified.

(2) The quantity $\mu_m^*(f)$ is the one appearing in the proof but $\mu_m(f)$ is easier to handle and is the one we have used. We have computed values of $\mu_m^*(f)$ and checked numerically that $\mu_m^*(f)^{1/(m+1)}$ and $\mu_m(f)^{1/(m+1)}$ were very close one to another.

(3) Riemann hypothesis is known to hold up to $T_0 = 3.33 \times 10^9$ [25], so our results can be improved.

(4) We looked for optimal values of $(m, \alpha_1, T, a, \alpha_4, \alpha_2)$ in this order (recursively: with the first datas fixed, find the best next one). The parameters T, a, α_4 and α_2 were easy to get, but α_1 turned out to be quite troublesome, due to the fact that Δ as a function of α_1, with m fixed and the other parameters chosen almost optimally, is *not* clearly first increasing and then decreasing. We thus decided to scan a large range of values for α_1. We have assumed that as a function of m, Δ was increasing and then decreasing. This optimization process has been carried out with 28 digits of precision and final datas recomputed to 1000 digits. We can thus claim that the values of Δ we give are correct (up to a big blunder) but not that they are optimal,

Table 2
$a = 0$

$\log X$	Δ	m
46	34 063 443	21
47	35 425 690	22
48	35 958 929	22
49	36 217 784	22
50	36 359 809	22
56	36 390 432	22

though we of course expect them to be close to it. We modified this program several times, ran it on different machines and the outputs were consistent.

(5) Lowering the value of α_1 means lowering the value of T but also enables us to take a larger a. The way these two mechanisms interfere is highly non-trivial. For large values of m, we found that it was slightly better to gain on a but a much larger value of α_1 would have yielded results only marginally inferior. This accounts for the seemingly high instability of the near-optimal parameters we took.

(6) To see that the sieve effect is indeed noticeable, we produced Table 2 where we took $a = 0$: the intervals are about 5 times smaller.

5. Proof of (2)

In this section we examine the strength of Theorem 4 asymptotically in T_0 and prove (2).

We evaluate the size of the interval in the following lemma:

Lemma 7. *For $a \in [0, \frac{1}{2}]$, we have*

$$1 - \frac{1 + mua}{1 + mu(1-a)} e^{-u} \leqslant u(1 + (1-2a)m).$$

Proof.

$$\begin{aligned} 1 - \frac{1 + mua}{1 + mu(1-a)} e^{-u} &= \frac{(1 + mua)}{1 + mu(1-a)}(1 - e^{-u}) + mu \frac{1 - 2a}{1 + mu(1-a)} \\ &\leqslant u + mu(1-2a), \end{aligned}$$

which concludes the proof. □

We choose

$$T = T_0, \quad \alpha_1 = 1/2, \quad \alpha_2 = 1, \quad \alpha_4 = 1/2, \quad m = \log T_0 + \mathcal{O}(1),$$

$$a = \frac{1}{2} - \frac{b}{\sqrt{m}} \text{ for a large enough positive } b.$$

We first need to evaluate $w(f_m, a)$, and we prove below that $w(f_m, a) \leqslant e^{-4b^2}/(2b)$. Assuming this estimate, we choose b so that $w(f_m, a) \leqslant 1/16$ and for instance $b = 1$ is enough. We find that α_3 is of order 1, hence

$$u \asymp \frac{1}{T_0}\left(T_0 \frac{\log T_0}{m}\right)^{1/(m+1)} \asymp 1/T_0$$

(where $f \asymp g$ means $f \ll g$ and $g \ll f$), from which we infer

$$\Delta^{-1} \ll \frac{1}{T_0}\left(1 + m\frac{2b}{\sqrt{m}}\right) \ll \sqrt{\log T_0}\,/T_0$$

as claimed.

We are thus left to prove the inequality concerning $w(f_m, a)$. Note first that $||f_m||_1 \geqslant (2m)^{-1/2}$. Elementary transformations yields

$$\begin{aligned} v(f_m, a) &= 2\int_0^{\frac{1}{2}-\frac{b}{\sqrt{m}}} (4t(1-t))^m \, dt \\ &= \int_{\frac{2b}{\sqrt{m}}}^1 (1-h^2)^m dh \qquad (t = (1-h)/2) \qquad (25) \end{aligned}$$

$$v(f_m, a) = \frac{1}{2\sqrt{m}} \int_{4b^2}^m \left(1 - \frac{k}{m}\right)^m \frac{dk}{\sqrt{k}} \qquad (h^2 = k/m) \qquad (26)$$

$$\leqslant \frac{1}{2\sqrt{m}} \int_{4b^2}^m e^{-k} \frac{dk}{\sqrt{k}} \leqslant \frac{1}{2\sqrt{m}} \int_{4b^2}^\infty e^{-k} \frac{dk}{\sqrt{k}} \leqslant \frac{e^{-4b^2}}{4b\sqrt{m}} \qquad (27)$$

from which the estimate concerning $w(f_m, a)$ follows readily.

6. A result under Riemann hypothesis

If $Y \in [599, 10^8]$, the result follows from (1); if $Y \in [10^8, 10^{11}]$ it follows from the theorem of Schoenfeld cited after Theorem 3, while for $Y \in [10^8, 10^{18}]$, it follows from Theorem 3. A direct verification covers the lower range. We could also have employed Nicely's result [15]. We thus restrict our attention to the case $Y \geqslant 10^{15}$.

We use Theorem 4 with $T_0 = \infty$, $m = 2$, $\alpha_1 = 0.63$, $\alpha_2 = 0.5 \log Y$, $\alpha_4 = 1/2$ and $a = 2/25$. The list of inequalities below clearly amounts to an algorithm and these parameters are close to the optimal ones. We use

$$T = 2\alpha_1 \pi \sqrt{Y}/\log Y \tag{28}$$

and prove that the interval $[Y - \frac{8}{5}\sqrt{Y}\log Y, Y]$ contains a prime. A short calculation also shows that $\mu_2^*(f_2) = 40\sqrt{3}/3$.

Here is the list of inequalities we verify

$$\begin{aligned}
&\alpha_3 \in [6.189, 6.226],\\
&2N(T) \leqslant \alpha_1 \sqrt{Y},\\
&S_2 = \frac{T s_2(T)}{\sqrt{Y}} \in [0.25, 0.321],\\
&u = \frac{1.0001 \log Y}{2\alpha_1 \pi \sqrt{Y}} (\alpha_3 10\sqrt{3} S_2/3)^{1/3} \in \left[0.525 \frac{\log Y}{\sqrt{Y}}, 0.570 \frac{\log Y}{\sqrt{Y}}\right],\\
&2u \leqslant 0.0001,\\
&0.5 \log Y \leqslant u\sqrt{Y},\\
&\frac{1 + 2(1-a)u - (1 + 2au)e^{-u}}{1 + 2(1-a)u} \leqslant 2.68u.
\end{aligned}$$

Their proof relies only on standard numerical analysis and is of no great interest. We display below some of the steps for the reader to be able to check our result.

Bounds for α_3: We have $w(f_2, a) = 2a^3(10 - 15a + 6a^2)$. Furthermore,

$$\alpha_3^{-1} = (0.99 - 1.0007\alpha_1 - 4.002 w(f_2, a) - 0.0326/\alpha_2)/2.0012$$

which lies between $1/6.189$ and $1/6.226$.

$2N(T)$ *versus* $\alpha_1\sqrt{Y}$: Consider

$$h(Y) = 2g(T/(2\pi)) - \alpha_1 \sqrt{Y}, \tag{29}$$

where $g(u) = u\log u - u + \frac{7}{8} + 0.67\log u$. Recall that $Y \geqslant 10^{15}$. We have

$$\begin{aligned} h(Y) &= 2\alpha_1\frac{\sqrt{Y}}{\log Y}(\log(\alpha_1) + \tfrac{1}{2}\log Y - \log\log Y - 1) + \tfrac{7}{8} \\ &\quad + 0.67\log\frac{\alpha_1\sqrt{Y}}{\log Y} - \alpha_1\sqrt{Y} \\ &= 2\alpha_1\frac{\sqrt{Y}}{\log Y}(\log(\alpha_1) - \log\log Y - 1) + \tfrac{7}{8} + 0.67\log\frac{\alpha_1\sqrt{Y}}{\log Y} \\ &\leqslant -6.3\frac{\sqrt{Y}}{\log Y} + 0.58 + 0.67\log\frac{\sqrt{Y}}{\log Y} \end{aligned}$$

which is easily seen to be non-positive.

Bounds for S_2: We have

$$\begin{aligned} \frac{Ts_2(T)}{\sqrt{Y}} = \frac{2\alpha_1\pi}{\log Y}\Bigg\{&\frac{1}{2\pi}\left(\log(\alpha_1\sqrt{Y}) - \log\log Y + \frac{1}{2}\right) \\ &+ \frac{1.34\log Y}{2\alpha_1\pi\sqrt{Y}}(2\log(\alpha_1\sqrt{Y}) - 2\log\log Y + 1)\Bigg\}\{1 + 10^{-5}\} \end{aligned}$$

which is easily shown to be between 0.255 and 0.324 (asymptotically $\alpha_1(1+10^{-5})/2$).

Bounds for u:

$$u = \frac{1.0001\log Y}{2\pi\alpha_1\sqrt{Y}}(\alpha_3 10\sqrt{3}S_2/3)^{1/3} \tag{30}$$

is between $0.525(\log Y)/\sqrt{Y}$ and $0.570(\log Y)/\sqrt{Y}$.

An upper bound for the size of the interval: We next have to find an upper bound for the size of the interval, for which we simply use Lemma 5:

$$\frac{1 + 2(1-a)u - (1+2au)e^{-u}}{1 + 2(1-a)u}Y \leqslant 2.68uY.$$

We conclude by noticing that $0.570 \times 2.68 \leqslant 8/5$.

7. The algorithm in the large

While the main result of this paper is established for numbers greater than 10^{20}, numerical computations have been necessary to compute the least value x_0 for which the property is not true. Informally, the verification amounts to exhibiting a sequence of prime number $p_0, p_1, \ldots, p_N$ with $p_0 > 10^{20}$ and $p_N = x_0 + 1$ and such that for any integer X, $x_0 < X \leqslant 10^{20}$, there exists a prime number p_k of the sequence such that $p_k \in]X.(1-1/\varDelta), X]$. Informally, our algorithm was the following one:

```
Step 1.  Set $X = 10^{20}$.
Step 2.  Set $X' = \lfloor X.(1-1/\Delta) \rfloor + 1$
Step 3.  For $Y$=$X'$ to $X$ do
           If $Y$ is prime,
             then Output($Y$);
             Set $X = Y - 1$; Goto Step 2
           Endif.
Step 4.  Output("Property fails for "); Output($X$). End.
```

The problem of the prime certificate is a crucial issue. Indeed, while general primality checker can easily establish or disprove the primality of numbers with around 20 decimal digits, they are much too slow for our purpose. Indeed, approximatively 400 millions of prime numbers were necessary for the derivation and each one would have required about 1 s with the elliptic curve primality checker ECPP [2], thus we would have a global running time of about 10 years which is quite unrealistic except perhaps by using a large amount of machines. Certificates using cyclotomic fields [1,5] have nearly the same time performances. We also investigated the problem with a program using Lehmer's certificate. Although the global throughput of the program was better, it was regularly lowered with numbers whose primality was more difficult (in terms of elementary operations) to establish. The reason for this is that to establish the primality of a number p with this certificate the entire factorization of $p-1$ has to be known and this factorization can be costly if it contains only quite large factor. It has to be noted also that Pocklington's certificate which is a refinement of Lehmer's certificate but with only incomplete factorization of $p-1$ is here useless. Indeed, this certificate requires factorization of $p-1$ with a factored part greater or equal to $\sqrt{p}$. But in our implementation, problematic prime numbers p are the ones such that $p-1 = 2q_1q_2$ where q_1 and q_2 are prime numbers of approximatively the same size. Since a factorization to an extent of a least $\sqrt{p}$ is necessary, one of the factors q_1 or q_2 has to be known to apply Pocklington's theorem, but in this case the other one is also immediately known and then Lehmer's criterion could be used as well. There exists also a refinement of Pocklington's theorem which has been given by Brillhart, Lehmer, and Selfridge [6]. Although, this certificate is here quite useless, we will see above that it is the cornerstone of our verification.

As we have seen, a systematic prime certificate is too costly to give an efficient implementation. However, it is also well known that most of the prime numbers can be certified quickly. Thus we decided to make implementations which only take into account the prime numbers which require not too much elementary operations. This was made by limiting the effort in the factorization of $p-1$. Although it gives better results it is still insufficient since it requires the verification of many congruences, similar to the ones of Fermat little's theorem. Thus we decided to use prime generation techniques [12]: we interest ourselves only to family of numbers whose primality can be established with one or two Fermat-like or Pocklington's congruences. This kind of technique has been already used in a quite similar problem [17].

In fact, our generation techniques rely on a theorem proven by Brillhart, Lehmer, and Selfridge [6] which states:

Lemma 8. *Let $N = RF + 1$ an odd integer for which the entire factorization of F is known, F is even and $\gcd(R, F) = 1$. We suppose that there exists an integer a such that $a^{N-1} \equiv 1 \pmod{N}$ and, for all prime factor p_i of F, $\gcd(a^{(N-1)/p_i} - 1, N) = 1$. We pose then $R = 2Fs + r$ with $0 \leqslant r < 2F$. We suppose $N < 2F^3$, then N is a prime number if and only if either $s = 0$ or $r^2 - 8s$ is not a perfect square.*

This lemma in fact uses Pocklington's property. Indeed this latter theorem, under those assumptions, states that any divisor of N is of the form $kF + 1$. We have $N = 2F^2s + Fr + 1$ and since $N < 2F^3$, it can have at most two factors. If $s = 0$, N is too small to have two factors and thus is prime and in the general case if N has two factors, we have $N = (k_1F + 1)(k_2F + 1) = k_1k_2F^2 + (k_1 + k_2)F + 1$ we then have $2s = k_1k_2$ and $r = (k_1 + k_2)$. In this case we have $(k_1 - k_2)^2 = (k_1 + k_2)^2 - 4k_1k_2 = r^2 - 8s$ which is a perfect square, whence the result.

8. Primes generation techniques

Our verification was split in three intervals. Each step is detailed in subsequent paragraph and we only give here an overview of the situation. The first interval starts at 10^{20} and ends at 18 723 898 090 586 113. There we generate primes of the shape $N = 2^kR + 1$ whose primality will be ensured by Pocklington's test. Only one congruence is required, and if furthermore $N \equiv 3$ [8] then the base $a = 5$ is enough.

The second interval goes from 18 723 898 090 586 113 to 877 803 410 503. Numbers of the previous shape are too sparse and we generate primes of the shape $N = 2pR + 1$ where p is a prime ranges [209 519, 399 989], and the primality of N is still ensured by a Pocklington's test. For such numbers, two congruences are required while only one was enough in the previous method. To avoid looking for a proper base, we only try $a = 3$. Furthermore we explore all the arithmetic progressions $2 \times 209\,519 \times R + 1, \ldots, 2 \times 399\,989 \times R + 1$ successively. It would have been possible to explore all these sequences simultaneously, and that would have amounted to a lesser number of generated primes since the present method somehow overdoes the work. It was however simpler to supervise and control the results the way we chose. Also, the limitation to $a = 3$ is questionable, since a prime out of 2 ought to miss this test (i.e. $3((N - 1)/2) \equiv 1[N]$ and not $\equiv -1$) while the second congruence will always be satisfied. But the probability of a composed N satisfying these conditions looks very small. Thus a number in this situation has still a good chance to be a prime, and will most probably validate the test with a being some fairly small number. These remarks are being made to mention that we could most probably generate a denser sequence with similar techniques. The value 877 803 410 503 arose from practical considerations: the storage disk was full and the third method was applicable.

This third method is essentially exact and goes from 877 803 410 503 to 10 726 905 041 and uses results of Jaeschke [11].

The softwares were written in C language, using the GMP multiprecision library [10]. All the computations were performed on the Power Challenge Array of the Loria (http://www.loria.fr), Nancy, France.

8.1. Upper interval

In this part, intervals in which primes have to be searched are large and then quite easy to found. As in [17], we take $F = 2^k$ and set $R = 2^{k+1}s + r$ in Lemma 8. We note that since $N \equiv 3 \pmod{10}$ it is enough to consider the base 5: indeed 5 is not a square modulo N if N happens to be prime. In this way we derive:

Lemma 9. *Let* $N = 2^k.R + 1$ *with* $N \leqslant 2^{3k+1}$, $N \equiv 3 \pmod{10}$ *and* R *odd. Suppose that* $5^{(N-1)/2} \equiv -1 \pmod{N}$. *Then* N *is prime if and only if either* $s = 0$ *or* $r^2 - 8s$ *is not a perfect square*, r *and* s *defined as above.*

At the beginning since $2^{67} > 10^{20}$, the first execution was made with $k = 22$. With this value, the search failed after that 168 624 289 prime numbers were generated. The last prime number generated was 292 579 660 678 561 793. Since this number is less than 2^{61}, a value $k = 20$ is satisfactory to continue. After 39 738 963 prime numbers, the search failed again on 75 527 775 596 838 913. We used then $k = 19$ for 23 508 467 prime numbers down to 35 449 980 491 661 313 and $k = 18$ for 18 667 460 prime numbers down to 18 723 898 090 586 113. Since this number is greater than 2^{52}, it is impossible to continue with $k = 17$: while it is quite certain that the obtained numbers would be prime, they cannot be certified by the previous lemma. Fig. 1 illustrates the execution time of the algorithm for series of 1 million generated primes. While it is decreasing in a very regular fashion almost everywhere, we observe a great decrease of execution time near the 50 millionth prime number. The reason for this is that the R10000 is a 64-bit processor and around 10^{20}, computed numbers with GMP requires two 64-bit words to be represented, while smaller numbers encountered further in the verification need only one 64-bit word. Thus near the 50 millionth prime number, the decrease of execution time corresponds to the transition between two and one 64-bit words. Other low execution times were observed for transitions on k values, when an execution ended up with less than 1 million prime numbers.

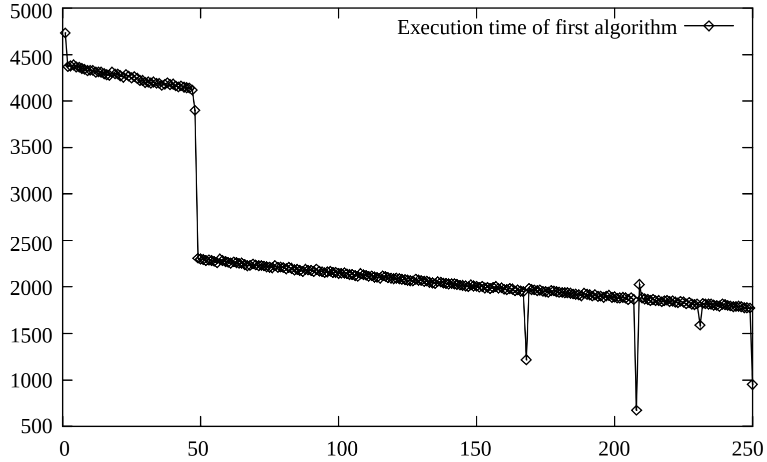

Fig. 1. Execution times for the first algorithm by steps of 1 000 000 prime numbers generated. x-axis denotes the total number of primes generated since the beginning. y-axis is the execution time of a given step in seconds on a single R10000 processor of the PCA.

8.2. Middle interval

At this point, prime generation techniques can still be used but, since possible series have a low density with respect to the length of the intervals where prime numbers have to be found, we needed to consider a large family of series. We focused on series of the form $2pR+1$ where p is a prime number and $\gcd(2p, R) = 1$. Those prime numbers, if not too large can be certified with two Pocklington's congruences. Here we take $F = 2p$ and set $R = 4ps + r$ in Lemma 8 to derive:

Lemma 10. *Let $N = 2pR + 1$ with p prime, $\gcd(2p, R) = 1$ and $N \leqslant 16p^3$. Suppose that there exists an integer a such that*

- $a^{(N-1)/2} \equiv -1 \pmod N$,
- $\gcd(a^{(N-1)/p} - 1, N) = 1$.

Then N is prime if and only if either $s = 0$ or $r^2 - 8s$ is not a perfect square, r and s defined as above.

Since we wanted to limit the search of an eventual primitive root a, we only focused on prime numbers of this kind which may be certified by $a = 3$. For the values of p, the least one was necessarily at least 209 519: this is the least prime number p such that $16p^3$ is greater than the value at which the preceding step ends. As for the greatest value, we arbitrarily decide to consider only prime numbers less than 400 000, thus 399 989 was the greatest possible prime number p. In our implementation, we consider sequentially each of the possible arithmetic progression.

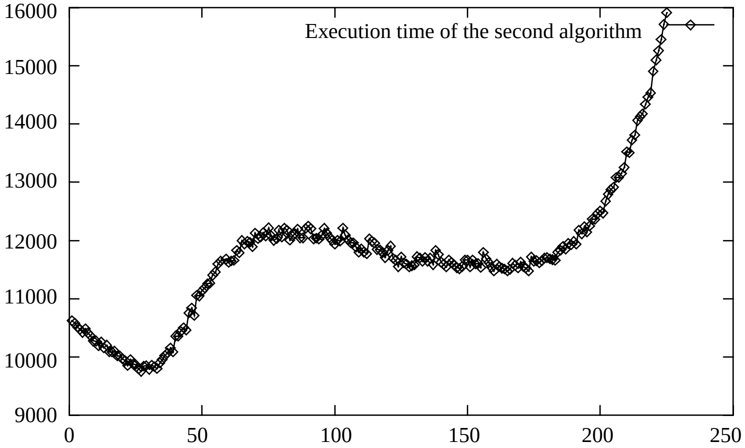

Fig. 2. Execution times for the second algorithm by steps of 1 000 000 prime numbers generated. x-axis denotes the total number of primes generated since the beginning of this phase. y-axis is the execution time of a given step in seconds on a single R10000 processor of the PCA.

A more efficient way might have been to consider all the arithmetic progressions at the same time and to explore them concurrently in depth. We considered the first method for the sake of simplicity, all the more that the second solution requires a more complicated control which might have degraded performances of the implementation. Fig. 2 illustrates the execution time of each run of this second phase. The execution was intentionally stopped at 877 803 410 503 for practical reasons (disk occupation) and because of the increase of execution time for the runs. We can see, firstly that the execution time is larger than for the first phase and secondly that the execution time is no more decreasing. The first point is easily explained by the facts that the criterion for the second phase is more complicated and that several candidate families are considered together instead of a single one in the first phase. The relative growth of execution time is explained, as for itself, by the decrease of density for candidate families: at the beginning of the phase, the intervals in which prime numbers have to be found are large and thus the first family $(2.209519k + 1)$ whose progression step is fixed can easily give a successful candidate. At the end of the phase, intervals become shorter and thus more and more family have to be considered to find a prime number.

8.3. Lower interval

In the last interval of verification, the numbers were below 10^{12}. Many prime certificates might have been successfully used, in regard to the small number of digits as well as the small number of primes to certify. We decided to use the Jaeschke's results [11] about pseudoprimality, which give easy and fast certificate for such numbers.

Definition 1. Let p be an odd integer and a be an integer. Let h be such that $p = 1 + 2^h d$ with d being odd. Then p is a *strong pseudoprime* for base a if we have either $a^d \equiv 1 \pmod{p}$, or there exists k such that $0 \leqslant k < h$ with $a^{2^k.d} \equiv -1 \pmod{p}$.

This notion was introduced by Miller and Rabin [13]. Every prime number is strong pseudoprime for any base but there exists also composite numbers that pass some tests: for instance $2047 = 23 \times 89$ is the smallest composite number strong pseudoprime for base 2. However, it is known that a composite number cannot be pseudoprime for all base and even if the GRH is true then it is possible to sketch a polynomial primality certificate with those tests [3]. Pomerance, Selfridge and Wagstaff [16] made the exhaustive list of strong pseudoprimes for bases 2, 3, 5 and 7 up to 25×10^9. Jaeschke's designed a new algorithm to search pseudoprimes and tabulated them up to 3.4×10^{14}. He obtained in particular two useful results:

Theorem 5. *Let p be an odd integer. If p is strong pseudoprime for bases* 2, 3, 5, 7, 11, 13 *and* 17 *and* $p < 3.4 \times 10^{14}$, *then p is a prime number.*

Theorem 6. *Let p be an odd integer. If p is strong pseudoprime for bases* 2, 13, 23 *and* 1 662 803 *and* $p < 10^{12}$, *then p is a prime number.*

We used this latter theorem to design our primality certificate but we firstly sieve the candidates with small prime numbers in order to eliminate most of the composite numbers. The algorithm produced an output of more than 126 millions of prime number before ending with 10 726 905 041 for which the search failed. This latter value is thus the optimal x_0 value of the main theorem of this article. Execution times

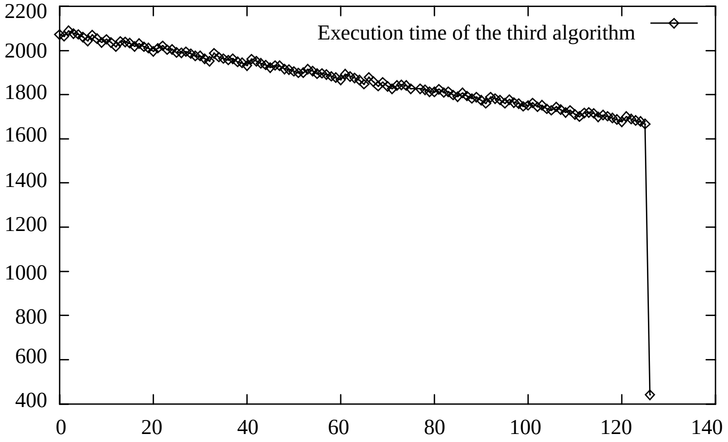

Fig. 3. Execution times for the third algorithm by steps of 1 000 000 prime numbers generated. x-axis denotes the total number of primes generated since the beginning of this phase. y-axis is the execution time of a given step in seconds on a single R10000 processor of the PCA.

are depicted in Fig. 3. We can see that execution times regularly decrease. This is easily explained by the fact that the computational cost of a Miller–Rabin test (this test is in fact due to Selfridge in the early 1970s but is known as the Miller–Rabin test), decrease with the size of the integer to check and also since the density of prime numbers increase while test numbers decrease in size and thus it leads to less unsuccessful checks.

Acknowledgments

I express my thanks first of all to my advisor, Jean-Marc Deshouillers, since this work (and my interest to it) started under his guidance, and then to Andrew Granville, Henryk Iwaniec and Carl Pomerance for helpful discussions on this subject, and for their hospitality. This last thanks extends to the Institute for Advanced Study where a good part of this work was completed.

Thanks are also due to the referee for his careful reading of the paper.

References

[1] L.M. Adleman, C. Pomerance, R.S. Rumely, On distinguishing prime numbers from composite numbers, Ann. Math. 117 (1983) 173–206.

[2] A.O.L. Atkin, F. Morain, Elliptic curves and primality proving, Math. Comput. 61 (203) (1993) 29–68.

[3] E. Bach, Analytic methods in the analysis and design of number-theoretic algorithms, ACM Distinguished Dissertations, 1985.

[4] R. Baker, G. Harman, J. Pintz, The difference between consecutive primes, II, Proc. London Math. Soc. (3) 2001, to appear.

[5] W. Bosma, M.-P. van der Hulst, Faster primality testing, in: H. Hartmanis, G. Goos (Eds.), Advances in Cryptology—EUROCRYPT '89, Lecture Notes in Computer Science, Vol. 434, Springer, Berlin, 1990, pp. 652–656.

[6] J. Brillhart, D.H. Lehmer, J.L. Selfridge, New primality criteria and factorizations for $2^m \pm 1$, Math. Comput. 29 (130) (1975) 620–647.

[7] N. Costa Pereira, Elementary estimates for the Chebyshev function $\psi(x)$ and for the Möbius function $M(x)$, Acta Arith. 52 (1989) 307–337.

[8] H. Cramer, On the order of magnitude of the difference between consecutive prime numbers, Acta Arith. 2 (1936) 23–46.

[9] H. Davenport, Multiplicative Number Theory, 3rd Edition, Graduate texts in Mathematics, Vol. 74, Springer, Berlin, 2000.

[10] T. Grandlung, The GNU multiple precision arithmetic library, Technical Documentation, 1993.

[11] G. Jaeschke, On strong pseudoprimes to several bases, Math. Comput. 61 (204) (1993) 915–926.

[12] U. Maurer, Fast generation of prime numbers and secure public-key cryptographic parameters, J. Cryptol. 8 (3) (1995) 123–156.

[13] G.L. Miller, Riemann's hypothesis and tests for primality, J. Comput. System Sci. 13 (1976) 300–317.

[14] H. Montgomery, R.C. Vaughan, The large sieve, Mathematika 20 (1973) 119–134.

[15] T.R. Nicely, New maximal primes gaps and first occurences, Math. Comput. 24 (1999) 1311–1315.

[16] C. Pomerance, J.L. Selfridge, S.S. Wagstaff, The pseudoprimes up to 25×10^9, Math. Comput. 35 (1980) 1003–1026.

[17] Y. Saouter, Checking the odd Goldbach conjecture up to 10^{20}, Math. Comput. 35 (1998) 863–866.

[18] L. Schoenfeld, Sharper bounds for the Chebyshev functions $\psi(x)$ ii, Math. Comput. 30 (134) (1976) 337–360.

[19] A. Selberg, On the normal density of primes in small intervals, and the difference between consecutive primes, Arch. Math. Naturv. B 47 (6) (1943) 82–105.

[20] J. Richstein, Verifying the Goldbach conjecture up to 4×10^{14}, Math. Comput. 70 (236) (2001) 1745–1749.

[21] J.B. Rosser, Explicit bounds for some functions of prime numbers, Amer. J. Math. 63 (1941) 211–232.

[22] J.B. Rosser, L. Schoenfeld, Sharper bounds for the Chebyshev functions $\vartheta(x)$ and $\psi(x)$, Math. Comput. 29 (129) (1975) 243–269.

[23] van de Lune, H.J.J. te Riele, D.T. Winter, On the zeros of the Riemann zeta-function in the critical strip. iv, Math. Comput. 46 (174) (1986) 667–681.

[24] S. Wedeniwski, Announcement, see http://www.hipilib.de/zeta/index.html.

[25] D. Wolke, On the explicit formula of Riemann–von Mangoldt, ii, J. London Math. Soc. 2 (28) (1983) 406–416.

[26] J. Young, A. Potler, First occurrence prime gaps, Math. Comput. 52 (185) (1989) 221–224.