MATHEMATICS OF COMPUTATION
Volume 79, Number 269, January 2010, Pages 513–533
S 0025-5718(09)02262-5
Article electronically published on May 6, 2009

A DETERMINISTIC VERSION OF POLLARD'S $p-1$ ALGORITHM

BARTOSZ ŹRAŁEK

Abstract. In this article we present applications of smooth numbers to the unconditional derandomization of some well-known integer factoring algorithms.

We begin with Pollard's $p-1$ algorithm, which finds in random polynomial time the prime divisors p of an integer n such that $p-1$ is smooth. We show that these prime factors can be recovered in deterministic polynomial time. We further generalize this result to give a partial derandomization of the k-th cyclotomic method of factoring ($k \geq 2$) devised by Bach and Shallit.

We also investigate reductions of factoring to computing Euler's totient function φ. We point out some explicit sets of integers n that are completely factorable in deterministic polynomial time given $\varphi(n)$. These sets consist, roughly speaking, of products of primes p satisfying, with the exception of at most two, certain conditions somewhat weaker than the smoothness of $p-1$. Finally, we prove that $O(\ln n)$ oracle queries for values of φ are sufficient to completely factor any integer n in less than $\exp\Big((1+o(1))(\ln n)^{\frac{1}{3}}(\ln\ln n)^{\frac{2}{3}}\Big)$ deterministic time.

1. Introduction

A fundamental question of algorithmic number theory, in particular, and complexity theory, in general, asks whether there are computational problems which cannot be solved efficiently without the use of randomness. If the answer is no, then we would say that every algorithm can be derandomized. The issue surely has a philosophical flavour, but above all is essential for the development of mathematics. As a rule, derandomization presupposes making the most of the rich mathematical structures involved. It gives rise to new ideas, subtle refinements of existing ones, or, in the worst case, generates fascinating open problems. One of these problems, determining the complexity of primality testing, has been brilliantly solved in [2]: primes are recognizable in deterministic polynomial time.

In this article we present applications of smooth numbers to the unconditional derandomization of some well-known integer factoring algorithms. Recall that a smooth number is a product of small primes (small relative to, say, n meaning polynomial in the size of n).

Received by the editor November 26, 2007 and, in revised form, October 3, 2008 and January 1, 2009.

2000 *Mathematics Subject Classification.* Primary 11Y16; Secondary 11Y05, 68Q10.

Key words and phrases. Pollard's $p-1$ method, derandomization, Euler's φ-function and factorization.

In sections 3 and 4 we analyze Pollard's $p-1$ method [21], important both in theory and practice [23, 17]. Pollard's algorithm finds in random polynomial time those prime divisors p of an integer n for which $p-1$ is smooth. We show that such prime factors can be recovered in deterministic polynomial time (Corollary 4.6). Let us merely indicate the two ingredients of the proof. The first comes from Fürer [12], Fellows and Koblitz [11], and also Konyagin and Pomerance [14]: take small integers or, what amounts to the same, small primes to generate a large subgroup G of $\mathbb{Z}_n^*$. The second is a novel idea inspired by the Pohlig-Hellman algorithm [20] for computing discrete logarithms. Namely, let H be the group generated by two elements a and b of $\mathbb{Z}_n^*$, both having smooth order. Then given a, b and their orders, we can compute a generator of H or a nontrivial divisor of n in deterministic polynomial time. This result is easily extended by induction to any number of given generators for H (Corollary 4.3). We apply it with $H = G$.

In section 5 we give a partial derandomization of the k-th cyclotomic method of factoring devised by Bach and Shallit [6]. This method is used to find in random polynomial time such prime factors p of an integer n that the value at p of the k-th cyclotomic polynomial is smooth. For the reader's convenience, we first treat the simpler case $k = 2$ (Theorem 5.1), corresponding to Williams' $p+1$ method [26], then that of an arbitrary k, $k \geq 2$ (Theorem 5.5). The arguments involve more than the derandomization of the $p-1$ algorithm: some elementary algebraic number theory and a lemma proved in [27].

In the last three sections, we attempt to make some progress on a famous open problem: is factoring reducible in deterministic polynomial time to computing Euler's totient function φ? (Cf. problem 23 of [1].)

In section 6 we discuss the current state of the art. Miller [19] found a reduction whose correctness depends on the Extended Riemann Hypothesis (ERH). Rabin [22] obtained an unconditional reduction at the cost of giving up determinism. A relatively recent result of Burthe [8] yields a reduction for almost all integers, but these cannot be simply described.

In section 7 we point out some explicit sets of integers n that are completely factorable in deterministic polynomial time given $\varphi(n)$ (Theorem 7.1). These sets consist, roughly speaking, of products of primes p satisfying, with the exception of at most two, certain conditions somewhat weaker than the smoothness of $p-1$.

In section 8 we study the deterministic complexity of factoring given an oracle for the function φ. Suppose that we want to factor into primes the integer n. Our idea is first to query the oracle for the iterations $\varphi(n)$, $\varphi^2(n)$, $\varphi^3(n)$, etc. until $\varphi^k(n) = 1$. Then to come back up to the complete factorization of n ($n = \varphi^0(n)$) by a recursive procedure, which recovers the prime factorization of $\varphi^{l-1}(n)$ from the prime factorization of $\varphi^l(n)$, starting with $l = k$. We are basically left with the task of finding the prime factorization of an integer n given the complete factorization of $\varphi(n)$. In the hard case, all the prime divisors of n are congruent to 1 modulo a large integer A that we compute; we further retrieve the missing information either by a direct search or by factoring the polynomial whose coefficients are the coefficients of n in base A (Lemma 8.5). The resulting algorithm runs in less than $\exp((1+o(1))(\ln n)^{\frac{1}{3}}(\ln\ln n)^{\frac{2}{3}})$ deterministic time (Theorem 8.1). Consequently, factoring is reducible in deterministic subexponential time to computing φ (Corollary 8.2).

2. Notation

Throughout the text n is an odd integer, and p, q, s are prime numbers.

The greatest common divisor, respectively the least common multiple, of the integers a, b is denoted by (a, b), respectively $\mathrm{LCM}(a, b)$.

We let $v_s(m)$ be the exponent of the highest power of s dividing m.

For G a group, $\mathcal{B} \subset G$, $b \in G$, we should denote by $\langle \mathcal{B} \rangle_G$ the subgroup of G generated by $\mathcal{B}$, and denote by $\mathrm{ord}_G(b)$ the order of b in G. However, if $G = \mathbb{Z}_d^*$, respectively $G = \mathbb{Z}_d[\sqrt{m}]^*$, we will just write $\langle \mathcal{B} \rangle_d$ and $\mathrm{ord}_d(b)$, respectively $\langle \mathcal{B} \rangle_{d,m}$ and $\mathrm{ord}_{d,m}(b)$. The cyclic group with m elements is denoted by C_m. The symbol $\mathbb{P}$ stands for the set of all prime numbers. We denote by $p_-(m)$, respectively $p_+(m)$, the least, respectively the largest, prime dividing m. We use a_i to represent the i-th coordinate of $a \in \mathbb{Z}_n^* = \bigoplus_{q|n} \mathbb{Z}^*_{q^{v_q(n)}}$. We recall the definitions of the familiar number-theoretic functions appearing in the text:

$$\begin{aligned}
&\varphi(m) = \#\{d \le m : \ (d, m) = 1\} \text{ (Euler's totient function)},\\
&\omega(m) = \sum_{p|m} 1 \text{ and } \Omega(m) = \sum_{p|m} v_p(m),\\
&\psi(x, y) = \#\{m \le x : \ p_+(m) \le y\}.
\end{aligned}$$

We will make frequent use of the following theorem proved in [14]:

Theorem 2.1 (Konyagin, Pomerance)**.** *If $n \ge 4$ and $2 \le (\ln n)^c \le n$, then $\psi(n, (\ln n)^c) > n^{1-\frac{1}{c}}$.*

We will always assume that its hypotheses are satisfied when c is fixed (this is natural in the task of factoring n). In the last section another estimation of ψ will be applied.

Theorem 2.2 (Canfield et al.)**.** *There is an effective, positive constant C such that for $x, y \ge 1$ and $u := \frac{\ln x}{\ln y} \ge 3$ we have*

$$\psi(x, y) \ge x \exp\left[-u\left\{\ln(u \ln u) - 1 + \frac{\ln\ln u - 1}{\ln u} + C\left(\frac{\ln\ln u}{\ln u}\right)^2\right\}\right].$$

3. Pollard's $p-1$ factoring algorithm

We first sketch the ideas behind the probabilistic version of Pollard's $p-1$ factorization method. Let n be an odd integer, not a prime power. Assume that we are given an integer M such that $p - 1 \mid M$ for some $p \mid n$ (for the moment we do not consider the issue of finding a suitable M). Choose $b \in \mathbb{Z}_n^*$. By Fermat's little theorem we have $b^M \equiv 1 (\mathrm{mod}\ p)$ and thus $d := (b^M - 1, n) > 1$. If additionally $d < n$, then d is a nontrivial divisor of n. But what if $d = n$, i.e. $b^M = 1$? We can pick another element of $\mathbb{Z}_n^*$. We can also hope to find a nontrivial factor of n in the sequence $(b^{\frac{M}{2^l}} - 1, n)_{l=1,\ldots,v_2(M)}$, as all square roots of 1 in $\mathbb{Z}_n^*$ are of the form $(\pm 1, \ldots, \pm 1) \in \mathbb{Z}_n^* = \bigoplus_{q|n} \mathbb{Z}^*_{q^{v_q(n)}}$. It turns out that the expected number of random $b \in \mathbb{Z}_n^*$ needed to split n does not exceed 2.

Theorem 3.1 (Rabin). *Let n be odd, $n > 2$, M be even,*

$$\mathcal{F}(M) = \{b \in \mathbb{Z}_n^* : \ b^M \neq 1\},$$
$$\mathcal{S}(M) = \{b \in \mathbb{Z}_n^* \setminus \mathcal{F}(M) : \ \exists_{1 \leq l \leq v_2(M)} \ 1 < (b^{\frac{M}{2^l}} - 1, n) < n\}.$$

Then $\frac{\#(\mathcal{F}(M) \cup \mathcal{S}(M))}{\varphi(n)} \geq \ 1 - 2^{1-\omega(n)}$.

Note that we want not only M to be a multiple of $p-1$ for some (a priori unknown) $p \mid n$, but also $\ln M$ to be relatively small (e.g., bounded by a fixed power of $\ln n$), so that raising to the power M (or $\frac{M}{2^l}$) modulo n does not take too much time. Suppose that n has a prime divisor p such that $p-1$ is smooth, say $p_+(p-1) \leq (\ln n)^u$. Set $M = \prod_{q \leq (\ln n)^u} q^{\left[\frac{\ln n}{\ln q}\right]}$. Then M satisfies the two conditions, since $\ln M \leq \sum_{q \leq (\ln n)^u} \frac{\ln n}{\ln q} \ln q = \pi((\ln n)^u) \ln n = O\big(\frac{(\ln n)^{u+1}}{u \ln \ln n}\big)$ from Chebyshev's theorem. By contrast, there is no efficient method of finding M if n is not divisible by a prime p as above.

As before suppose that n is odd, divisible by at least two different primes p and q. It is well known that if a multiple M of $p-1$ is given, then the previously described search for a nontrivial factor of n can be derandomized under the ERH. Without loss of generality assume that $b^M \equiv 1 (\mathrm{mod}\ n)$ for all $b < 2(\ln n)^2$.

Theorem 3.2 (Bach). *Suppose that the ERH is true. Let $n \geq 3$, χ be a nonprincipal character modulo n. There is an integer $b < 2(\ln n)^2$ such that $\chi(b) \neq 1$.*

Using this theorem, we can easily prove the existence of $b < 2(\ln n)^2$ such that for some l, $b^{\frac{M}{2^l}} - 1$ is divisible by q or p, but not both. We apply it with χ induced by the quadratic character $\left(\frac{\cdot}{p}\right), \left(\frac{\cdot}{q}\right), \left(\frac{\cdot}{pq}\right)$ when $v_2(p-1) > v_2(q-1), v_2(p-1) < v_2(q-1), v_2(p-1) = v_2(q-1)$, respectively.

4. A deterministic variant of Pollard's $p-1$ factoring algorithm

Our basic framework is as follows. Let $\mathcal{B} = \{2, 3, \ldots, [(\ln n)^2]\}$. Assume that we are given an integer M together with its complete factorization such that $b^M \equiv 1 (\mathrm{mod}\ n)$ for every $b \in \mathcal{B}$. We want to find a simple and not restrictive condition on n under which n is factorable in deterministic polynomial time in $\ln n$ and $\ln M$. The starting point is a reformulation of the primality criterion from [11]. We restate the argument for completeness and clarity of exposition.

Theorem 4.1 (Fellows-Koblitz). *Let $\mathcal{B} = \{2, 3, \ldots, [(\ln n)^2]\}$, $\mathcal{B} \subset \mathbb{Z}_n^*$. Then n is prime if and only if the following conditions are satisfied:*

(i) $\mathrm{ord}_p(b) = \mathrm{ord}_n(b)$ *for every* $b \in \mathcal{B}$ *and* $p \mid n$,
(ii) $\mathrm{LCM}_{b \in \mathcal{B}}(\mathrm{ord}_n(b)) > \sqrt{n}$.

Proof. Suppose n is prime. Condition (i) is then a tautology. We check condition (ii). The group $\langle \mathcal{B} \rangle_n$ is cyclic, since n is prime. Therefore

$$\mathrm{LCM}_{b \in \mathcal{B}}(\mathrm{ord}_n(b)) = \#\langle \mathcal{B} \rangle_n \geq \psi(n, (\ln n)^2) > \sqrt{n},$$

where the last inequality follows from Theorem 2.1.

Assume now that conditions (i) and (ii) are satisfied. Let $p = p_-(n)$. We then have $\mathrm{ord}_p(b) = \mathrm{ord}_n(b)$ for all $b \in \mathcal{B}$ and thus

$$\mathrm{LCM}_{b\in\mathcal{B}}(\mathrm{ord}_p(b)) = \mathrm{LCM}_{b\in\mathcal{B}}(\mathrm{ord}_n(b)) > \sqrt{n}.$$

However $\mathrm{LCM}_{b\in\mathcal{B}}(\mathrm{ord}_p(b)) \mid p-1$. Consequently $p > \sqrt{n}$; hence $n \in \mathbb{P}$. □

Let $b \in \mathbb{Z}_n^*$, $p \mid n$. Recall that $\mathrm{ord}_p(b) < \mathrm{ord}_n(b)$ is equivalent to $p \mid b^{\frac{\mathrm{ord}_n(b)}{s}} - 1$ for some $s \mid \mathrm{ord}_n(b)$. If $(b^{\frac{\mathrm{ord}_n(b)}{s}} - 1, n) > 1$ for some $s \mid \mathrm{ord}_n(b)$, then we will say that b is a *Fermat-Euclid witness for* n. Checking conditions (i) and (ii) therefore reduces to factoring the orders of the elements of $\mathcal{B}$, which can be done efficiently under our assumption on M. Taking $M = n-1$ yields a deterministic polynomial time algorithm for deciding the primality of integers n such that $n-1$ is smooth. Actually, a stronger test, in which only a part of $n-1$ exceeding $n^{\frac{1}{2}+\varepsilon}$ ($\varepsilon > 0$) is assumed to be smooth, was first discovered by Fürer [12]. Konyagin and Pomerance [14] further reduced the exponent $\frac{1}{2}+\varepsilon$ to ε. The key point is that beside searching some other appropriately chosen "small" subset $\mathcal{B}$ of $\mathbb{Z}_n^*$ for Fermat-Euclid witnesses for n, one can also check the cyclicity of $\langle\mathcal{B}\rangle_n$. The authors verify this stringent condition by applying the classic Pohlig-Hellman technique [20] of discrete logarithm computation in a prime field. Here we will in a sense extend this technique for the purpose of splitting the integer n.

Suppose for greater generality that $\mathcal{B}$ is any subset of $\mathbb{Z}_n^*$. We will describe below a deterministic algorithm that finds a generator of $\langle\mathcal{B}\rangle_n$ or, particularly in the case when $\langle\mathcal{B}\rangle_n$ is not cyclic, a nontrivial divisor of n. This algorithm runs in polynomial time if $\mathcal{B}$ consists of elements having smooth orders in $\mathbb{Z}_n^*$. By induction, it is sufficient to restrict our attention to the case $\#\mathcal{B} = 2$, say $\mathcal{B} = \{a, b\}$.

We assume temporarily that $\mathrm{ord}_n(a) = s^v$, $b^{s^v} = 1$ with $s \in \mathbb{P}$, $v \in \mathbb{N}$. Let $n = p_1^{e_1} \cdot \ldots \cdot p_k^{e_k}$ be the complete factorization of n. There exist an i, $1 \le i \le k$, such that $\mathrm{ord}_{p_i^{e_i}}(a_i) = s^v$. Since $b_i^{s^v} = 1$ and $\mathbb{Z}^*_{p_i^{e_i}}$ is cyclic, we have $a_i^l = b_i$ for some uniquely determined, less than s^v, natural number l. Write l in base s: $l = \sum\limits_{0\le r<v} l_r s^r$. Set $l_{-1} = 0$ and reason by induction. Assume we have computed $l_{-1}, \ldots, l_m$, where $-1 \le m \le v-2$. Put $c = ba^{-\sum_{-1\le r\le m} l_r s^r}$. Then $c_i = a_i^{\sum_{m<r<v} l_r s^r}$. Therefore $c_i^{s^{v-m-2}} = a_i^{l_{m+1}s^{v-1}}$. Denote $(c^{s^{v-m-2}} - a^{js^{v-1}}, n)$ by d_j. We successively compute $d_0, d_1, \ldots$, until we get $d_j > 1$ for some $j \le s-1$. This will happen, because $p_i^{e_i} \mid d_{l_{m+1}}$. If moreover $d_j < n$, then d_j is a nontrivial factor of n. Otherwise, $d_j = n$. In particular, $c_i^{s^{v-m-2}} = a_i^{js^{v-1}}$. Hence $j = l_{m+1}$. Eventually, if $m = v-2$, then $d_{l_{m+1}} = n$ implies $b = a^l$. More formally we use the ensuing algorithm.

PH(n, a, b, s, v, w) $\{a, b \in \mathbb{Z}_n^*, s \in \mathbb{P}, \mathrm{ord}_n(a) = s^v, \mathrm{ord}_n(b) = s^w\}$

(1) If $w > v$, then interchange a and b.
(2) For $j = 1$ to $s-1$ compute $a^{js^{v-1}}$.
(3) Let $c = b$.
(4) For $m = -1$ to $v-2$ do
  (a) Let $j = 0$.
  (b) While $(c^{s^{v-m-2}} - a^{js^{v-1}}, n) = 1$ do $j = j+1$.
  (c) Let $d = (c^{s^{v-m-2}} - a^{js^{v-1}}, n)$. If $d \ne n$, then return d.
  (d) Let $c = ca^{-js^{m+1}}$.

Theorem 4.2. *Let $a, b \in \mathbb{Z}_n^*, s \in \mathbb{P}, \mathrm{ord}_n(a) = s^v, \mathrm{ord}_n(b) = s^w$. If the algorithm PH($n, a, b, s, v, w$) does not find a nontrivial divisor of n, then $\langle a, b\rangle_n$ is cyclic. This algorithm uses $O((s + u \ln s)u(\ln n)^2)$ operations, where $u = \max(v, w)$.*

Proof. The correctness of PH(n, a, b, s, v, w) follows from the preceding discussion. Step 2 requires $O(u(\ln n)^2 \ln s + s(\ln n)^2)$ operations. The total number of operations used by step 4b in the loop 4 is $O(u^2(\ln n)^2 \ln s + us(\ln n)^2)$. Step 4d takes on the whole in the loop 4, $O(u^2(\ln n)^2 \ln s)$ operations, hence the stated running time. □

Now suppose that $\mathcal{B} = \{a, b\}$ with $\mathrm{ord}_n(a)$ and $\mathrm{ord}_n(b)$ arbitrary. Let $A = \mathrm{ord}_n(a)$, $B = \mathrm{ord}_n(b)$. For $s \in \mathbb{P}$, set $g_s = a^{\frac{A}{s^{v_s(A)}}}$ if $v_s(A) \geq v_s(B)$, else $g_s = b^{\frac{B}{s^{v_s(B)}}}$. We follow the procedures PH($n, a^{\frac{A}{s^{v_s(A)}}}, b^{\frac{B}{s^{v_s(B)}}}, s, v_s(A), v_s(B)$), s running through the set of primes dividing (A, B). The group $\langle a, b\rangle_n$ is a direct sum of its s-primary parts $\langle a^{\frac{A}{s^{v_s(A)}}}, b^{\frac{B}{s^{v_s(B)}}}\rangle_n$. Therefore, either a nontrivial factor of n will be found, or $\langle a, b\rangle_n$ is cyclic, generated by $\prod_{s|AB} g_s$.

Corollary 4.3. *Assume we are given a subset $\mathcal{B}$ of $\mathbb{Z}_n^*$ and the complete factorization of all the integers $\mathrm{ord}_n(b)$ for $b \in \mathcal{B}$. Then we can find a generator of $\langle\mathcal{B}\rangle_n$ or a nontrivial factor of n in $O(\#\mathcal{B} \cdot (p + \ln n)(\ln n)^3)$ deterministic time, where p is the greatest prime dividing the order of at least two distinct $b_1, b_2 \in \mathcal{B}$ (put $p = 0$ if there is no such prime).*

Proof. Again, the correctness has been already discussed. We obtain the run-time bound by summing $(s + v_s(\varphi(n)) \ln s)v_s(\varphi(n))(\ln n)^2$ over $s \mid \varphi(n)$, $s \leq p$, and multiplying by $\#\mathcal{B}$. □

Remark 4.4. The number p in the O symbol above could be replaced by $\sqrt{p} \ln p$. To achieve this, one uses FFT techniques, well known from Pollard's [21] or Strassen's [24] algorithms for factoring n in $O(n^{\frac{1}{4}+\varepsilon})$ deterministic time. The hardest part of the PH() algorithm is finding j, $0 \leq j < s$, such that $d_j > 1$. The integer j is of the form $j = j_0 + j_1\lceil\sqrt{s}\rceil$ for some integers j_0, j_1, $0 \leq j_0, j_1 < \lceil\sqrt{s}\rceil$. Let $a' = a^{s^{v-1}}$, $c' = c^{s^{v-m-2}}$. We introduce the polynomial $h = \prod_{0 \leq i_0 < \lceil\sqrt{s}\rceil} (c' - a'^{i_0}X)$ and compute $h(a'^{i_1\lceil\sqrt{s}\rceil})$ for $i_1 = 0, 1, \ldots, \lceil\sqrt{s}\rceil - 1$. By Theorem 4 of [25] it can be done in $O(\sqrt{s}(\ln s)^2(\ln n)^2)$ deterministic time. Next we find j_1 satisfying $(h(a'^{j_1\lceil\sqrt{s}\rceil}), n) > 1$. Afterwards we find j_0 such that $(c' - a'^{j_0+j_1\lceil\sqrt{s}\rceil}, n) > 1$. The computational cost of these last two steps is $O(\sqrt{s}(\ln n)^2)$, thus negligible.

Turning back to our main question, we propose the following deterministic algorithm for splitting n given an integer M as in the beginning of this section.

Split($n, M, s_1, v_1, \ldots, s_t, v_t$) $\{M = s_1^{v_1} \cdot \ldots \cdot s_t^{v_t}$ is the complete factorization of $M\}$

(1) For every $b \in \mathcal{B}$, compute b^M modulo n, and:
   (a) If $(b^M - 1, n) = 1$, then report failure and stop.
   (b) If $(b^M - 1, n) < n$, then output this gcd and stop.
(2) Using the complete factorization of M, compute $\mathrm{ord}_n(b)$ for each $b \in \mathcal{B}$.
(3) For every $b \in \mathcal{B}$ and each prime $s \mid \mathrm{ord}_n(b)$, compute $(b^{\frac{\mathrm{ord}_n(b)}{s}} - 1, n)$. If one of these gcds is a nontrivial factor of n, then stop.

(4) Using the algorithm associated with Corollary 4.3, check whether $\langle \mathcal{B} \rangle_n$ is cyclic. If a nontrivial divisor of n is found during these computations, then stop.
(5) State that n is prime.

Theorem 4.5. *Let $\mathcal{B} = \{2, 3, \ldots, [(\ln n)^2]\}$, $M = s_1^{v_1} \cdot \ldots \cdot s_t^{v_t}$ be the complete factorization of the integer M, $s_0 = \max\{s \mid M : \forall_{q|n}\ s \mid q-1\} \cup \{0\}$. Suppose that $b^M \equiv 1 (\mathrm{mod}\ n)$ for all $b \in \mathcal{B}$. Then the algorithm Split$(n, M, s_1, v_1, \ldots, s_t, v_t)$ finds a nontrivial divisor (or a proof of the primality) of n in $O((s_0 \ln n + (\ln M)(\ln \ln M) + (\ln n)^2)(\ln n)^4)$ deterministic time.*

Proof. For the correctness assume that we have reached step 5 of the algorithm. Step 3 implies that $\mathcal{B}$ contains no Fermat-Euclid witness for n and step 4 that $\langle \mathcal{B} \rangle_n$ is cyclic. Therefore n is indeed prime in the light of the Fellows-Koblitz primality criterion. We proceed to the running time analysis. Step 1 requires $O((\ln M)(\ln n)^4)$ operations. Step 2 can be done in $O((\ln M)(\ln \ln M)(\ln n)^4)$ time (see [14]–the analysis of the runtime of algorithm 3.1). Step 3 costs $O\big(\frac{(\ln n)^6}{\ln \ln n}\big)$ operations. When we get to step 4, the exponent of $\langle \mathcal{B} \rangle_n$ divides $q-1$ for every prime factor q of n. By Corollary 4.3, the remaining computations thus take $O((s_0 + \ln n)(\ln n)^5)$ time. □

There might be inputs n for which the runtime of Split$(n, M, s_1, v_1, \ldots, s_t, v_t)$ is not polynomial in $\ln n$ and $\ln M$, but it actually is if the integer s_0 defined in Theorem 4.5 is small, say bounded by a polynomial B in $\ln n$. This is obviously satisfied whenever n has a prime divisor p such that $p-1$ is B-smooth.

Corollary 4.6 (deterministic version of Pollard's $p-1$ algorithm)**.** *Let $B \geq \ln n$.*

(i) *Assume n has a prime divisor p such that $p-1$ is B-smooth. Then we can find a nontrivial divisor (or a proof of the primality) of n in $O(B(\ln n)^5)$ deterministic time.*
(ii) *Suppose in addition that n has at most one prime divisor p such that $p-1$ is not B-smooth. Then we can obtain the complete factorization of n, together with a primality proof for each of the prime factors, in $O(B(\ln n)^6)$ deterministic time.*

Proof. Put $M = \prod_{q \leq B} q^{\left[\frac{\ln n}{\ln q}\right]}$ in Theorem 4.5. Part (i) follows, since $\ln M = O\big(\frac{B}{\ln B} \ln n\big)$ and $\ln \ln M = O(\ln B)$. For part (ii), simply consider the iteration of the algorithm corresponding to part (i), combined with the Lenstra-Pomerance variant of the AKS primality test [18], which runs in $O((\ln n)^6 (\ln \ln n)^c)$ deterministic time for some constant c. □

Let us briefly compare the running times of the original Pollard $p-1$ algorithm with the new version. The original algorithm finds a nontrivial divisor of n in $O\big(\frac{B}{\ln B}(\ln n)^3\big)$ random time under the assumption of Corollary 4.6 (i). Our deterministic version is slower (though not as much as we would expect) and thus rather of theoretical than practical interest.

Of course, the obtained running time bound of Split$(n, M, s_1, v_1, \ldots, s_t, v_t)$ is polynomial in $\ln n$ and $\ln M$ for more inputs n than those considered in Corollary 4.6, with B a polynomial in $\ln n$. Let $D(n, u) = \max_{q > (\ln n)^u} \#\{p \mid n : \ q \mid p-1\}$, $u > 0$.

We should expect that the integers n for which $D(n,u) > 1$ (with u fixed) are rare. This is in fact true. We prove slightly more than needed to motivate the ideas of section 7.

Theorem 4.7. *Let $l \in \mathbb{N}$. The number $B(x,u,l)$ of integers $n \le x$ such that $D(n,u) > l$ is bounded above by $cx\frac{2^{lu}(\ln\ln x)^{l+1}}{(\ln x)^{lu}}$, where the constant c does not depend upon u.*

Proof. We have:

$$B(x,u,l) \le \sqrt{x} + \sum_{\sqrt{x}<n\le x}\sum_{q>(\ln n)^u}\sum_{\substack{p_1<\ldots<p_{l+1}\\ p_i|n\\ p_i\equiv 1(\mathrm{mod}\, q)}} 1 \le \sqrt{x} + \sum_{q>2^{-u}(\ln x)^u}\sum_{\sqrt{x}<n\le x}\sum_{\substack{p_1<\ldots<p_{l+1}\\ p_i|n\\ p_i\equiv 1(\mathrm{mod}\, q)}} 1,$$

$$\sum_{n\le x}\sum_{\substack{p_1<\ldots<p_{l+1}\\ p_i|n\\ p_i\equiv 1(\mathrm{mod}\, q)}} 1 = \sum_{\substack{p_1<\ldots<p_{l+1}\le x\\ p_i\equiv 1(\mathrm{mod}\, q)}}\left[\frac{x}{p_1\cdot\ldots\cdot p_{l+1}}\right] \le x\sum_{\substack{p_1<\ldots<p_{l+1}\le x\\ p_i\equiv 1(\mathrm{mod}\, q)}}\frac{1}{p_1\cdot\ldots\cdot p_{l+1}}$$

$$\le x\Big(\sum_{\substack{p\le x\\ p\equiv 1(\mathrm{mod}\, q)}}\frac{1}{p}\Big)^{l+1} \le \frac{c_1 x(\ln\ln x)^{l+1}}{(q-1)^{l+1}},$$

where the last inequality follows from the uniform bound

$$\sum_{\substack{p\le x\\ p\equiv 1(\mathrm{mod}\, d)}}\frac{1}{p} \le \frac{c_0}{\varphi(d)}\ln\ln x$$

(use summation by parts and apply the Brun-Titchmarsh inequality). Hence

$$\sum_{q>2^{-u}(\ln x)^u}\sum_{\sqrt{x}<n\le x}\sum_{\substack{p_1<\ldots<p_{l+1}\\ p_i|n\\ p_i\equiv 1(\mathrm{mod}\, q)}} 1 \le c_1 x(\ln\ln x)^{l+1}\sum_{q>2^{-u}(\ln x)^u}\frac{1}{(q-1)^{l+1}}$$

$$\le c_2 x\frac{2^{lu}(\ln\ln x)^{l+1}}{(\ln x)^{lu}}.$$

Thus

$$B(x,u,l) \le c_3 x\frac{2^{lu}(\ln\ln x)^{l+1}}{(\ln x)^{lu}}.$$ □

5. Generalization to the $p+1$ and other cyclotomic methods

Williams [26] designed a method of factoring analogous to Pollard's $p-1$ algorithm, the $p+1$ method. It splits in random polynomial time integers n having a prime divisor p such that $p+1$ is smooth. Traditionally, it is described in terms of Lucas sequences, but the analogy with the $p-1$ method becomes clear if one works, modulo n, in some quadratic extension of $\mathbb{Z}$, as we will do. This section is mainly devoted to the proof of

Theorem 5.1. *Let n and m be odd, coprime integers, $n > 2$, m squarefree. Let $B \ge \ln n$. Suppose that n has a prime factor p such that $p+1$ is B-smooth and $\left(\frac{m}{p}\right) = -1$. Then we can find a nontrivial divisor (or a proof of the primality) of n in $O_{c,m}(B(\ln n)^{ch+3})$ deterministic time, where h is the class number of $\mathbb{Q}(\sqrt{m})$ and c is any constant greater than 4.*

The obtained derandomization of the $p+1$ algorithm is only partial, because of the requirement $\left(\frac{m}{p}\right) = -1$, m being fixed. We should therefore talk about deterministic $p+1$ methods (for varying m) instead of one deterministic $p+1$ algorithm. We need some auxiliary results in the spirit of [27], the extension of the Pohlig-Hellman algorithm for the group $\mathbb{Z}_n[\sqrt{m}]^*$ to begin with.

Theorem 5.2. *Suppose that m modulo p is a quadratic nonresidue for some prime p dividing n. Let a subset $\mathcal{B}$ of $\mathbb{Z}_n[\sqrt{m}]^*$ and the complete factorization of all the integers $\mathrm{ord}_{n,m}(b)$ for $b \in \mathcal{B}$ be given. Then a generator of $\langle \mathcal{B} \rangle_{n,m}$ or a nontrivial factor of n can be computed in $O_m(\#\mathcal{B} \cdot (q + \ln n)(\ln n)^3)$ deterministic time, where q is the greatest prime dividing the order of at least two distinct $b_1, b_2 \in \mathcal{B}$ (set $q = 0$ if there is no such prime).*

Proof. As in Corollary 4.3 the argument reduces to the case of $\mathcal{B} = \{a, b\}$ with $\mathrm{ord}_{n,m}(a)$ and $\mathrm{ord}_{n,m}(b)$ equal to the powers of some prime s, say $\mathrm{ord}_{n,m}(a) = s^v$, $\mathrm{ord}_{n,m}(b) = s^w$, $v \geq w$. Let $a^{s^{v-1}} = a_1 + a_2\sqrt{m}$. We can also assume that $\mathrm{ord}_{p,m}(a) = s^v$, for otherwise $(a_1 - 1, n)$ or (a_2, n) would be a nontrivial divisor of n. The rest of the proof follows the lines of section 4, since $\mathbb{Z}_p[\sqrt{m}]^*$ is, by assumption, isomorphic to $\mathbb{F}^*_{p^2}$, hence cyclic. □

We introduce the standard integral basis of the ring of integers in $\mathbb{Q}(\sqrt{m})$, letting $y = \sqrt{m}$ if $m \equiv 2, 3 (\mathrm{mod}\, 4)$, and $y = \frac{1+\sqrt{m}}{2}$ if $m \equiv 1 (\mathrm{mod}\, 4)$. The next theorem is well known in the context of solving generalized Pell equations (norm equations in $\mathbb{Z}[y]$).

Theorem 5.3. *There is an effective, positive constant c_1 depending upon m and having the following property. For any nonzero $a \in \mathbb{Z}[y]$, there exists $b \in \mathbb{Z}[y]$, $b = b_1 + b_2 y$, such that $\frac{b}{a} \in \mathbb{Z}[y]^*$ and $|b_i| \leq c_1\sqrt{|N(a)|}$, where $N(a)$ is the norm of a and $i = 1, 2$.*

Finally, we formulate some kind of analogue of Theorem 2.1 for the ring $\mathbb{Z}[y]$.

Theorem 5.4. *Let n be odd, $n > 2$. Also, let $c > 1$. Adopting the above notation, define*

$$\mathcal{A} = \{a_1 + a_2 y : \ |a_i| \leq c_1 (\ln n)^{\frac{ch}{2}}, 1 \leq i \leq 2\},$$
$$\mathcal{S} = \{v \cdot \alpha_1 \cdot \ldots \cdot \alpha_t : \ v \in \mathbb{Z}[y]^*, t \in \mathbb{N}, \alpha_i \in \mathcal{A}, 1 \leq i \leq t\},$$

and $\pi_n : \ \mathbb{Z}[y] \to \mathbb{Z}_n[\sqrt{m}]$ as the obvious projection. Then $\#\pi_n(\mathcal{S}) > n^{2 - \frac{2}{c} - \varepsilon} + 1$ for any $\varepsilon > 0$ and $n \geq n_0$, $n_0 = n_0(m, c, \varepsilon)$.

Proof. This is in fact a special case of Lemma 3.5 from [27]. □

Let f_n be the endomorphism

$$a_1 + a_2\sqrt{m} \mapsto (a_1 - a_2\sqrt{m})(a_1 + a_2\sqrt{m})^{-1}$$

of $\mathbb{Z}_n[\sqrt{m}]^*$. Let $\mathcal{U}$ be a set of generators of the group of units $\mathbb{Z}[y]^*$, $\#\mathcal{U} \leq 2$ ($\mathcal{U}$ could be written explicitly), and let

$$\mathcal{B}_n = \pi_n(\mathcal{U} \cup \mathcal{A}) \setminus \{0\}.$$

The algorithm below is a deterministic version of the $p+1$ factorization method. We justify the correctness in the proof of Theorem 5.1.

Split2$(n, c, m, M, s_1, v_1, \ldots, s_t, v_t)$ {$c > 4$, $M = s_1^{v_1} \cdot \ldots \cdot s_t^{v_t}$ is the complete factorization of M}

(1) If n is a nontrivial power d^k, then output d and stop.
(2) Let n_0 be as in Theorem 5.4, with $\varepsilon = \frac{1}{2} - \frac{2}{c}$. If n has a prime factor below n_0, then output such one and stop.
(3) For each $a \in \mathcal{A}$, compute $N(a)$, let $\pi_n(a) = a_1 + a_2\sqrt{m}$, and:
  (a) If $1 < (N(a), n) < n$, then output $(N(a), n)$ and stop.
  (b) If $n \mid N(a)$, then:
    (i) If $(a_1, n) = 1$ or $(a_2, n) = 1$, then output failure and stop.
    (ii) If $(a_1, n) < n$, then output this gcd and stop. Do the same with (a_2, n).
(4) For every $b \in f_n(\mathcal{B}_n)$, compute b^M, $b^M = b_1 + b_2\sqrt{m}$, and:
  (a) If $(b_1 - 1, n) = 1$ or $(b_2, n) = 1$, then report failure and stop.
  (b) If $(b_1 - 1, n) < n$, then output this gcd and stop. Do the same with (b_2, n).
(5) Using the complete factorization of M, compute $\mathrm{ord}_{n,m}(b)$ for each $b \in f_n(\mathcal{B}_n)$.
(6) For every $b \in f_n(\mathcal{B}_n)$ and each prime $s \mid \mathrm{ord}_{n,m}(b)$, compute $b^{\frac{\mathrm{ord}_{n,m}(b)}{s}}$, $b^{\frac{\mathrm{ord}_{n,m}(b)}{s}} = b_1 + b_2\sqrt{m}$, and the gcds $(b_1 - 1, n)$, (b_2, n). If one of these gcds is a nontrivial factor of n, then stop.
(7) Using the algorithm associated with Theorem 5.2, check whether $\langle f_n(\mathcal{B}_n)\rangle_{n,m}$ is cyclic. If a nontrivial divisor of n is found during these computations, then stop.
(8) State that n is prime.

Proof of Theorem 5.1. Set $M = \prod_{q \le B} q^{\left[\frac{\ln(n+1)}{\ln q}\right]}$. First, we have to show that under our assumptions the algorithm will not report any failure. This could happen only in step 3b(i) or 4a. Let $n \mid N(a)$ in step 3b. Then, in particular, $p \mid N(a)$ and thus the element $\pi_p(a)$ is not invertible. Moreover, $\mathbb{Z}_p[\sqrt{m}]$ is isomorphic to the field $\mathbb{F}_{p^2}$, since $\left(\frac{m}{p}\right) = -1$. We conclude that $\pi_p(a)$ must be zero, that is to say, $p \mid a_1$ and $p \mid a_2$. Consequently, the algorithm cannot terminate in step 3b(i). Now let $b \in f_n(\mathcal{B}_n)$ in step 4. From step 3, $\mathcal{B}_n \subset \mathbb{Z}_n[\sqrt{m}]^*$, so b is correctly defined. The conjugation modulo p is easily seen to be nothing but the Frobenius map. The endomorphism f_p thus raises the elements of $\mathbb{Z}_p[\sqrt{m}]^*$ to the power of $p-1$. As M is a multiple of $p+1$, it follows that b^M modulo p must be equal to 1. Therefore no failure can be reported in step 4a.

Second, we should prove that n is prime when step 8 is reached. Let us assume the contrary and seek a contradiction. Denote by q the least prime factor of n, and by n' the squarefree part of n. Define A as $\mathrm{LCM}_{b \in f_{n'}(\mathcal{B}_{n'})} \mathrm{ord}_{n',m}(b)$. From step 6, we have $A = \mathrm{LCM}_{b \in f_q(\mathcal{B}_q)} \mathrm{ord}_{q,m}(b)$. By step 7, $\langle f_n(\mathcal{B}_n)\rangle_{n,m}$ is cyclic; so are its homomorphic images $\langle f_{n'}(\mathcal{B}_{n'})\rangle_{n',m}$ and $\langle f_q(\mathcal{B}_q)\rangle_{q,m}$. Thus

$$\#\langle f_{n'}(\mathcal{B}_{n'})\rangle_{n',m} = A = \#\langle f_q(\mathcal{B}_q)\rangle_{q,m}.$$

Hence

$$\#\langle f_{n'}(\mathcal{B}_{n'})\rangle_{n',m} \mid \#f_q(\mathbb{Z}_q[\sqrt{m}]^*).$$

Furthermore, $\#\langle f_{n'}(\mathcal{B}_{n'})\rangle_{n',m} \geq \frac{\#\langle \mathcal{B}_{n'}\rangle_{n',m}}{\#\ker f_{n'}}$. From step 2, $n' \geq n_0$, which by Theorem 5.4 yields $\#\langle \mathcal{B}_{n'}\rangle_{n',m} > n'^{\frac{3}{2}}$. We will evaluate $\#\ker f_{n'}$. Let s be a prime dividing n'. If $\left(\frac{m}{s}\right) = -1$, then we already know that $\#\ker f_s = s-1$. In the case when $\left(\frac{m}{s}\right) = 1$, it is not hard to show that f_s acts like the endomorphism $(a,b) \mapsto (ba^{-1}, ab^{-1})$ of $\mathbb{Z}_s^* \oplus \mathbb{Z}_s^*$, and therefore $\#\ker f_s = s-1$. Consequently, $\#\ker f_{n'} = \prod_{s|n'} \#\ker f_s = \prod_{s|n'}(s-1)$. Combining all the above, we get

$$\#\langle f_{n'}(\mathcal{B}_{n'})\rangle_{n',m} > \#f_q(\mathbb{Z}_q[\sqrt{m}]^*) \cdot \frac{q^{\frac{3}{2}}}{\#f_q(\mathbb{Z}_q[\sqrt{m}]^*) \cdot \#\ker f_q} \cdot (q^{-1}n')^{\frac{1}{2}}.$$

By the isomorphism theorem, $\#f_q(\mathbb{Z}_q[\sqrt{m}]^*) \cdot \#\ker f_q = \#\mathbb{Z}_q[\sqrt{m}]^*$, which is less than q^2. From step 1, $q < n'^{\frac{1}{2}}$. Hence

$$\#\langle f_{n'}(\mathcal{B}_{n'})\rangle_{n',m} > \#f_q(\mathbb{Z}_q[\sqrt{m}]^*) \cdot q^{-1}n'^{\frac{1}{2}} > \#f_q(\mathbb{Z}_q[\sqrt{m}]^*).$$

This contradicts the previously obtained inequality $\#\langle f_{n'}(\mathcal{B}_{n'})\rangle_{n',m} \leq \#f_q(\mathbb{Z}_q[\sqrt{m}]^*)$.

The running time analysis is similar to that of algorithm Split; the role of the "base set" $\mathcal{B}$ is played here by $f_n(\mathcal{B}_n)$, whose cardinality is $O_m((\ln n)^{ch})$. □

Pollard's $p-1$ and William's $p+1$ algorithms are part of a family of factoring algorithms called the cyclotomic methods. These were introduced by Bach and Shallit [6], who proved, conditionally on the generalized Riemann hypothesis (GRH), the following. Let Φ_k be the k-th cyclotomic polynomial. An integer n can be split in random polynomial time whenever $\Phi_k(p)$ is smooth for some prime p dividing n, and integer k polynomial in the size of n. If we fix k and strengthen (reasonably, of course) the condition on p, it will eventually appear that neither GRH nor randomness are necessary.

Theorem 5.5. *Let F be a monic, irreducible polynomial of degree k in $\mathbb{Z}[X]$, $k \geq 2$, such that the extension K of $\mathbb{Q}$, obtained by adjoining a root θ of F, is cyclic. Let $m \mid k$, $m \geq 2$, and $B \geq \ln n$. Assume that n is divisible by a prime p with the property that $\Phi_m(p)$ is B-smooth and F modulo p is irreducible in $\mathbb{Z}_p[X]$. Then a nontrivial factor (or a proof of the primality) of n can be computed in $O_{c,\theta}(B(\ln n)^{ch+3})$ deterministic time, where h is the class number of K and c is any constant greater than $2k$.*

In the proof we will adopt two more pieces of notation. We will write $\mathcal{O}_K$ for the ring of integers of K. Furthermore, let G be a group (written multiplicatively), $a \in G$, $\eta : G \to G$, $V = \sum v_i X^i \in \mathbb{Z}[X]$. The expression $V(\eta)(a)$ will stand for $\prod \eta^i(a^{v_i})$, η^i being the i-th iteration of η (η^0 the identity).

Proof. There is no loss of generality in supposing that n is coprime to the discriminant of F. The rings $\mathcal{O}_K/(n)$ and $\mathbb{Z}_n[\theta]$ are then isomorphic; we identify them for convenience. The Galois group of K over $\mathbb{Q}$ consists of, say, $\psi_1, \ldots, \psi_k$. Denote by $\psi_{i,n}$ the automorphism of $\mathbb{Z}_n[\theta]$ induced by ψ_i. Let $f_{i,n}$ be the endomorphism

$$a \mapsto \prod_{l|k,\ l\neq m} \Phi_l(\psi_{i,n})(a)$$

of $\mathbb{Z}_n[\theta]^*$. The prime p remains prime in $\mathcal{O}_K$; let ψ_j be the Frobenius over (p). Then $f_{j,p}$ acts like $\mathbb{F}_{p^k}^* \ni a \mapsto a^{\frac{p^k-1}{\Phi_m(p)}} \in \mathbb{F}_{p^k}^*$. Consequently, setting $M = \prod_{q\leq B} q^{\left[\frac{m\ln n}{\ln q}\right]}$

yields $f_{j,p}(a)^M = 1$ for any $a \in \mathbb{Z}_p[\theta]^*$. Up to now, we followed [6]. However, in order to compute deterministically a nontrivial factorization of n, we define a "base set" of the form $f_{j,n}(\mathcal{B}_n)$. We do not know j a priori, but in practice we can work in turn with each endomorphism $f_{i,n}$, $i = 1, \ldots, k$. An integral basis $\omega = (\omega_1, \ldots, \omega_k)$ of $\mathcal{O}_K$ and a finite set $\mathcal{U}$ of generators for $\mathcal{O}_K^*$ should be constructed independently of n, in a precomputation phase. Consider

$$\mathcal{A} = \{a_1\omega_1 + \ldots + a_k\omega_k : \ |a_i| \leq c_1(\ln n)^{\frac{ch}{k}}, 1 \leq i \leq k\},$$

where c_1 is the constant c_3 from Theorem 3.4 of [27]. Let π_n be the projection $\mathcal{O}_K \to \mathbb{Z}_n[\theta]$. Similarly to the proof of Theorem 5.1, we can assume that $\pi_n(\mathcal{U} \cup \mathcal{A}) \setminus \{0\} \subset \mathbb{Z}_n[\theta]^*$ and put $\mathcal{B}_n = \pi_n(\mathcal{U} \cup \mathcal{A}) \setminus \{0\}$. Again, let $q = p_-(n)$ and let n' be the squarefree part of n. Here also we can force $f_{j,n}(\mathcal{B}_n)^M = \{1\}$ and further

$$\#\langle f_{j,n'}(\mathcal{B}_{n'})\rangle_{\mathbb{Z}_{n'}[\theta]^*} \mid \# f_{j,q}(\mathbb{Z}_q[\theta]^*).$$

This would follow from appropriate generalizations of steps 4-7 of algorithm Split2. Still, the extension of Theorem 5.4 to $\mathcal{O}_K$ gives

$$\#\langle f_{j,n'}(\mathcal{B}_{n'})\rangle_{\mathbb{Z}_{n'}[\theta]^*} \geq \frac{\#\langle \mathcal{B}_{n'}\rangle_{\mathbb{Z}_{n'}[\theta]^*}}{\#\ker f_{j,n'}} > \frac{n'^{k-\frac{k}{c}-\varepsilon}}{\prod\limits_{s|n'} \#\ker f_{j,s}}$$

if $\varepsilon > 0$ and n' exceeds some constant n_0 independent of n. We have finally reached the interesting part of the proof, which is bounding $\#\ker f_{j,s}$ for s a prime factor of n. There are two cases to treat:

(i) s stays prime in $\mathcal{O}_K$,
(ii) s splits in $\mathcal{O}_K$: $(s) = S_1 \cdot \ldots \cdot S_e$, where the S_i are distinct primes of degree d, $d = \frac{k}{e}$, $e \geq 2$.

Before we do this, note that ψ_j has order k (because $\psi_{j,p}$ has order k). Suppose that (i) holds. The automorphism ψ_j generates the Galois group of K over $\mathbb{Q}$, isomorphic by reduction modulo s to the Galois group of $\mathcal{O}_K/(s)$ over $\mathbb{F}_s$, and so $\psi_{j,s}$ is raising to the power of s^r for some r relatively prime to k, $r < k$. Therefore $f_{j,s}$ acts as $\mathbb{F}_{s^k}^* \ni a \mapsto a^{\prod\limits_{l|k,\ l\neq m} \Phi_l(s^r)} \in \mathbb{F}_{s^k}^*$. It is easy to show that $\prod\limits_{l|k,\ l\neq m} \Phi_l(X^r) = \prod\limits_{l|k,\ l\neq m} \prod\limits_{t|r} \Phi_{tl}$. This product is coprime to Φ_m, since $m \mid k$. We apply Bézout's identity for polynomials to see that $(\Phi_m(s), \prod\limits_{l|k,\ l\neq m} \Phi_l(s^r))$ is bounded by a constant c_2 depending solely on k. Hence

$$\#\ker f_{j,s} = (s^k - 1, \prod_{l|k,\ l\neq m} \Phi_l(s^r)) \leq c_2 \frac{s^k - 1}{\Phi_m(s)} \leq c_3 s^{k-1},$$

where c_3 also depends only upon k.

Now assume that s satisfies (ii). We want to bound the number of solutions $(a_1, \ldots, a_e) \in (\mathcal{O}_K/S_1)^* \oplus \ldots \oplus (\mathcal{O}_K/S_e)^*$ to the equation $f_{j,s}(a_1, \ldots, a_e) = 1$. The automorphism ψ_j acts on the set $\{S_1, \ldots, S_e\}$ as a cyclic permutation. In particular, ψ_j^e generates the decomposition group of S_1, which is known to be isomorphic (by reduction modulo S_1) to the Galois group of $\mathcal{O}_K/S_1$ over $\mathbb{F}_s$. Consequently, there is an r coprime to d, such that $\psi_j^e(a) + S_1 = a^{s^r} + S_1$ for every $a \in \mathcal{O}_K$. Thus

$f_{j,s}(a_1,\dots,a_e)+S_1$ is of the form

$$ba_1^{-1+\sum\limits_{0<i\le d-1} u_i s^{ir}},$$

with b independent of a_1, and u_i integers depending just on k and m. The -1 in the exponent of a_1 corresponds actually to the free term of $\prod\limits_{l|k,\ l\neq m}\Phi_l$ $(m\ge 2)$. Since $(r,d)=1$, we have

$$a_1^{-1+\sum\limits_{0<i\le d-1} u_i s^{ir}} = a_1^{-1+\sum\limits_{0<i\le d-1} v_i s^{i}},$$

where the v_i are a permutation of the u_i. In the field $\mathcal{O}_K/S_1$ there are at most $|-1+\sum\limits_{0<i\le d-1} v_i s^i|$ solutions to the equation

$$ba_1^{-1+\sum\limits_{0<i\le d-1} v_i s^{i}} = 1$$

of unknown a_1. Therefore

$$\#\ker f_{j,s} \le (s^d-1)^{e-1}\cdot|-1+\sum_{0<i\le d-1} v_i s^i| \le c_4 s^{k-1}$$

for a constant c_4 depending only upon k.

Proceeding along the same lines as the proof of Theorem 5.1, we get, if $\varepsilon>0$ and $n'\ge n_0$, the inequality

$$\#\langle f_{j,n'}(\mathcal{B}_{n'})\rangle_{\mathbb{Z}_{n'}[\theta]^*} > \#f_{j,q}(\mathbb{Z}_q[\theta]^*)\cdot c_5 q^{-1} n'^{1-\frac{k}{c}-\varepsilon},$$

where the (positive) constant c_5 depends solely on k. Take $\varepsilon=\frac14-\frac{k}{2c}$. Since $\#\langle f_{j,n'}(\mathcal{B}_{n'})\rangle_{\mathbb{Z}_{n'}[\theta]^*}\le \#f_{j,q}(\mathbb{Z}_q[\theta]^*)$, we conclude that n is divisible by a prime less than $\max(n_0, c_5^{\frac{4c}{2k-c}})$, or n is a prime power. □

Remark 5.6. According to Frobenius' theorem, if F is as in Theorem 5.5, then the set of primes p, such that F modulo p is irreducible in $\mathbb{Z}_p[X]$, has density $\frac{\varphi(k)}{k}$. This set consists in fact of primes lying in residue classes, which can be explicitly determined. It suffices to express the root θ of F as an element of a cyclotomic field (here we appeal to the Kronecker-Weber theorem) and examine the order of the Frobenius automorphism in $\mathbb{Z}_p[\theta]$ (for p not dividing the discriminant of F). As an example, $F=X^3-3X+1$ (a correct choice) is irreducible in $\mathbb{Z}_p[X]$ if and only if $p\equiv\pm2 \pmod 9$ or $p\equiv\pm4 \pmod 9$. We could thus reformulate Theorem 5.5 in completely elementary terms for specific polynomials F. We highly recommend that the reader interested in the theoretical setting of cyclotomic factoring algorithms, and willing to compare in detail our result with the classic method of Bach and Shallit, consult [6].

6. Some known reductions of factoring to computing φ

Taking $M=\varphi(n)$ in Theorem 3.1 we get the following classical result.

Theorem 6.1 (Rabin)**.** *Given $\varphi(n)$ we can completely factor n in $O((\ln n)^4)$ expected time.*

For reasons already explained at the end of section 3, substituting $M=\varphi(n)$ also gives

Theorem 6.2 (Miller)**.** *If the ERH holds, then given $\varphi(n)$ we can completely factor n in $O((\ln n)^6)$ deterministic time.*

Define $G(n)$ as the least integer m such that $\mathbb{Z}_n^*$ is generated by integers less than or equal to m and coprime to n. In [8], Burthe proved that $\frac{1}{x}\sum_{n\le x} G(n) = O((\ln x)^{97})$. In particular, $G(n) < (\ln n)^{97+\varepsilon}$ for almost all integers n. Now recall that any nonprincipal character modulo n takes a value different from 1 for an integer less than or equal to $G(n)$. It follows by a similar argument to the one used after Theorem 3.2 that given $\varphi(n)$ we can completely factor n in $O((\ln n)^{101+\varepsilon})$ deterministic time for almost all n.

While it is an open problem whether factoring unconditionally reduces in deterministic polynomial time to computing Euler's φ function, for some integers such a reduction is particularly easy. The simplest nontrivial examples are integers n with exactly two prime factors. Suppose first that $n = pq$. Then $p + q = n - \varphi(n) + 1$. Given $\varphi(n)$ we compute the right-hand side of this equality and find p and q by solving a quadratic equation. Now turn to the general case $n = p^\alpha q^\beta$, say $p < q$. If $p \nmid q - 1$, then $\frac{n}{(n,\varphi(n))} = pq$ and $\frac{\varphi(n)}{(n,\varphi(n))} = (p-1)(q-1) = \varphi(pq)$; thus the previous method applies. If $p \mid q - 1$, then $\frac{n}{(n,\varphi(n))} = q$ and therefore q, β, α, p will be obtained one after the other.

Landau [15] showed that computing the equal order factorization of any integer n, that is, the sequence $n_i := \prod_{p:\ v_p(n)=i} p$ $(i \ge 1)$, can be done in deterministic polynomial time given a "φ-oracle" (this oracle finds instantly the values of Euler's φ function for $O(\ln n)$-bit inputs). In fact, if $\omega(n) \ge 3$, then $O(\Omega(n)(\ln n)^2)$ bit operations and at most $\omega(n) - 2$ oracle calls (including $\varphi(n)$) are needed. Notice that if $\omega(n_i) \le 2$ for all i, then the additional calls $\varphi(n_i)$ will lead to the complete factorization of n. For instance every integer $n = p^\alpha q^\beta s^\gamma$, where p, q, s are distinct primes and α, β, γ integers not all equal, can be, given $\varphi(n)$, completely factored in $O((\ln n)^3)$ deterministic time.

7. Some subsets of the graph of φ recognizable in deterministic polynomial time

In section 4 we have described in simple, arithmetic terms a set of integers of density 1 in $\mathbb{N}$ (the set $\{n :\ D(n,u) \le 1\}$ with u fixed) whose elements n are all factorable in deterministic polynomial time if $\varphi(n)$ is given in a fully factored form. The ideas presented there are extended here to get a much more concrete result: exhibit a possibly large set of integers n that are factorable in deterministic polynomial time given $\varphi(n)$ and only a part of its factorization, which in turn can be obtained in polynomial time with the deterministic Pollard $p-1$ method.

Let B and δ be positive real numbers. First define the following subsets of $\mathbb{P}$.

- $\mathcal{P}_B$ is the set of primes q such that $p - 1$ is B-smooth for every prime p dividing $q - 1$.
- $\mathcal{Q}_{B,\delta}$ is the set of primes q such that the B-smooth part of $q - 1$ is not less than q^δ.

Now consider, for k an integer, u, δ, η positive real numbers, $\delta < 1$, $\eta \le 1$, the set $\mathcal{N}_{k,u,\delta,\eta}$ of integers that can be written in the form $n = n_1 n_2 n_3$, where the n_i are pairwise coprime, and:

(1) n_1 has exactly k distinct prime factors, all belonging to $\mathcal{P}_{(\ln n)^u}$.
(2) n_2 is a product of primes from $\mathcal{Q}_{(\ln n)^u,\delta}$.
(3) n_3 has at most two distinct prime factors. Furthermore, if $\omega(n_3)=2$ and $n_2\neq 1$, then $p_-(n_2)>p_-(n_3)^\eta$.

We will prove

Theorem 7.1. *Let $\mathcal{N}_{k,u,\delta,\eta}$ be as above. Given the pair $(n,\varphi(n))$, with $n\in\mathcal{N}_{k,u,\delta,\eta}$, we can completely factor n in $O((\ln n)^C)$ deterministic time for some constant C depending only on k,u,δ,η. In particular, the set $\{(n,\varphi(n)):\ n\in\mathcal{N}_{k,u,\delta,\eta}\}$ is recognizable in deterministic polynomial time (k,u,δ,η being fixed).*

We prepare the proof with some lemmas, keeping the notation of the theorem and assuming, without loss of generality, that $p_-(n)>(\ln n)^{\max(\frac{2}{\delta},\frac{2+\eta}{\delta\eta},k+3)}$.

Lemma 7.2. *Let d be a factor of n, M the $(\ln n)^u$-smooth part of $\varphi(n)$, $\mathcal{B}=\{2,3,\dots,[(\ln d)^{\frac{2}{\delta}}]\}$ and $\mathcal{G}=\mathcal{B}^{\frac{\varphi(n)}{M}}$ modulo d. Assume that d is divisible by two distinct primes q_1,q_2 from $\mathcal{Q}_{(\ln n)^u,\delta}$. Then $\mathcal{G}$ contains a Fermat-Euclid witness for d or $\langle\mathcal{G}\rangle_d$ is not cyclic.*

Proof. Without loss of generality we let $q_1<q_2$. Suppose, on the contrary, that there is no Fermat-Euclid witness for d among the elements of $\mathcal{G}$ and that $\langle\mathcal{G}\rangle_d$ is cyclic. Then $\langle\mathcal{G}\rangle_{q_1q_2}$ is also cyclic, as a homomorphic image of $\langle\mathcal{G}\rangle_d$, and so $\#\langle\mathcal{G}\rangle_{q_1q_2}=\mathrm{LCM}_{g\in\mathcal{G}}\mathrm{ord}_{q_1q_2}(g)$. Moreover,

$$\mathrm{LCM}_{g\in\mathcal{G}}\mathrm{ord}_{q_1q_2}(g)=\mathrm{LCM}_{g\in\mathcal{G}}\mathrm{ord}_d(g)=\mathrm{LCM}_{g\in\mathcal{G}}\mathrm{ord}_{q_1}(g).$$

Therefore $\#\langle\mathcal{G}\rangle_{q_1q_2}$ divides (q_1-1,M), which equals, say M_1. We will show that $\#\langle\mathcal{G}\rangle_{q_1q_2}>M_1$ to derive a contradiction. Denote by h the endomorphism raising every element of $\mathbb{Z}^*_{q_1q_2}$ to the power of $\frac{\varphi(n)}{M}$. We have $\langle\mathcal{G}\rangle_{q_1q_2}=h(\langle\mathcal{B}\rangle_{q_1q_2})$; hence $\#\langle\mathcal{G}\rangle_{q_1q_2}\geq\frac{\#\langle\mathcal{B}\rangle_{q_1q_2}}{\#\ker h}$. The numerator $\#\langle\mathcal{B}\rangle_{q_1q_2}\geq\psi(q_1q_2,(\ln q_1q_2)^{\frac{2}{\delta}})>(q_1q_2)^{1-\frac{\delta}{2}}$. The denominator $\#\ker h=(q_1-1,\frac{\varphi(n)}{M})(q_2-1,\frac{\varphi(n)}{M})=\frac{(q_1-1)(q_2-1)}{M_1M_2}$, where we let $M_2=(q_2-1,M)$. Also, $q_2\in\mathcal{Q}_{(\ln n)^u,\delta}$ and $q_2>q_1$; thus $M_2\geq q_2^\delta>(q_1q_2)^{\frac{\delta}{2}}$. Putting all together gives $\#\langle\mathcal{G}\rangle_{q_1q_2}>M_1M_2\frac{(q_1q_2)^{1-\frac{\delta}{2}}}{(q_1-1)(q_2-1)}>M_1$, as required. □

Lemma 7.3. *Let d be a factor of n, M the $(\ln n)^u$-smooth part of $\varphi(n)$, $\mathcal{B}'=\{2,3,\dots,[(\ln d)^{\frac{2+\eta}{\delta\eta}}]\}$ and $\mathcal{G}'=\mathcal{B}'^{\frac{\varphi(n)}{M}}$ modulo d. Suppose that d is divisible by two distinct primes p and q, $q\in\mathcal{Q}_{(\ln n)^u,\delta}$, $q>p^\eta$. Then $\mathcal{G}'$ contains a Fermat-Euclid witness for d or $\langle\mathcal{G}'\rangle_d$ is not cyclic.*

Proof. Suppose that neither element of $\mathcal{G}'$ is a Fermat-Euclid witness for d. We are to explain why then $\langle\mathcal{G}'\rangle_d$ cannot be cyclic. Let $A=\mathrm{LCM}_{g\in\mathcal{G}'}\mathrm{ord}_p(g)$. By assumption, A also equals $\mathrm{LCM}_{g\in\mathcal{G}'}\mathrm{ord}_q(g)$, which is $\#\langle\mathcal{G}'\rangle_q$. Write M_1 for the $(\ln n)^u$-smooth part of $q-1$. Similarly to the proof of Lemma 7.2, we obtain

$$\#\langle\mathcal{G}'\rangle_q\geq M_1\frac{\#\langle\mathcal{B}'\rangle_q}{q-1}>q^{\frac{2\delta}{2+\eta}}.$$

Therefore $A>q^{\frac{2\delta}{2+\eta}}>p^{\frac{2\delta\eta}{2+\eta}}$. Since A divides $(p-1,M)$, it follows that $p\in\mathcal{Q}_{(\ln n)^u,\frac{2\delta\eta}{2+\eta}}$. Furthermore, $q\in\mathcal{Q}_{(\ln n)^u,\delta}\subset\mathcal{Q}_{(\ln n)^u,\frac{2\delta\eta}{2+\eta}}$. Replacing δ by $\frac{2\delta\eta}{2+\eta}$ in Lemma 7.2, we conclude that $\langle\mathcal{G}'\rangle_d$ is not cyclic. □

Lemma 7.4. *Let d be a factor of n, $M' = \prod p^{v_p(\varphi(n))}$, where the product ranges over the primes p such that $p-1$ is $(\ln n)^u$-smooth, $\mathcal{B}'' = \{2, 3, \ldots, [(\ln d)^{k+3}]\}$. Assume that d has a prime divisor $q \in \mathcal{P}_{(\ln n)^u}$ and that $\omega(d) \leq k+2$. Then one of the following conditions holds.*

(i) $1 < (b^{M'} - 1, d) < d$ *for some* $b \in \mathcal{B}''$.
(ii) $b^{M'} \equiv 1 (\mathrm{mod}\ d)$ *for all* $b \in \mathcal{B}''$ *and* $\mathcal{B}''$ *contains a Fermat-Euclid witness for* d.
(iii) $b^{M'} \equiv 1 (\mathrm{mod}\ d)$ *for every* $b \in \mathcal{B}''$ *and, setting* $A = \mathrm{LCM}_{b \in \mathcal{B}''} \mathrm{ord}_d(b)$, *we have* $p_-(d) \equiv 1 (\mathrm{mod}\ A)$, $A > d^\alpha$, *with* $\alpha > \frac{1}{\omega(d)} - \frac{1}{\omega(d)^2}$.

Proof. The definitions of M' and q imply that $q - 1 \mid M'$. Consequently, $b^{M'} \equiv 1 (\mathrm{mod}\, q)$ for any $b \in \mathcal{B}''$. We shall therefore suppose that $b^{M'} \equiv 1 (\mathrm{mod}\, d)$ for every $b \in \mathcal{B}''$, that there is no Fermat-Euclid witness for d in $\mathcal{B}''$, and verify the properties of A. Under the latter assumption, $A \mid p - 1$ for all primes p dividing d, for $p_-(d)$ in particular. That forces $(A, d) = 1$ and so $(\#\langle \mathcal{B}'' \rangle_d, d) = 1$. Hence $\langle \mathcal{B}'' \rangle_d \leq \bigoplus_{p \mid d} C_{p-1} \leq \mathbb{Z}_d^*$. Therefore $\langle \mathcal{B}'' \rangle_d$ contains, for each prime factor q of A, at most $\omega(d)$ linearly independent elements of order dividing $q^{v_q(A)}$. It follows that $A^{\omega(d)} \geq \#\langle \mathcal{B}'' \rangle_d$. Thus $A > \psi(d, (\ln d)^{k+3})^{\frac{1}{\omega(d)}} > d^\alpha$, where $\alpha = \frac{1}{\omega(d)}(1 - \frac{1}{k+3})$. Checking that $\alpha > \frac{1}{\omega(d)} - \frac{1}{\omega(d)^2}$ is straightforward. □

Lemma 7.5 (Coppersmith et al.)**.** *Assume we are given integers $h > v > 0$ and reals α, β, γ satisfying $0 < \alpha < 1$, $0 \leq \beta < \gamma \leq 1 - \alpha$, $v(v+1) + \gamma h(h-1) - 2(\alpha + \beta)vh < 0$. If d is larger than some effectively computable constant, then all the divisors of d of the form $sx + r$, where $0 < r < s < d$, $s \geq d^\alpha$, $(r, s) = (s, d) = 1$, $d^\beta \leq x \leq d^\gamma$, can be found in deterministic polynomial time in $v, h, \ln d$.*

Lemma 7.6. *Let r, s, d, l be integers and α a real number. Suppose that $0 < r < s < d$, $s \geq d^\alpha$, $(r, s) = (s, d) = 1$, $\alpha > \frac{1}{l} - \frac{1}{l^2}$ and d is sufficiently large. Then all the divisors of d of the form $sx + r$ and less than $d^{\frac{1}{l}}$ can be found in $O_\varepsilon((\ln d)^3)$ deterministic time, where $\varepsilon = \alpha - \frac{1}{l} + \frac{1}{l^2}$.*

Proof. This is achieved by partitioning $[1, d^{\frac{1}{l} - \alpha}]$, the range of x, into intervals to which Lemma 7.5 can be applied. We refer the reader to [10] for the details of the algorithm. For the running time, just follow closely the proof of Lemma 7.5 therein. □

Proof of Theorem 7.1. We describe an algorithm to compute the complete factorization of n.

(1) Let $L_1 = \{n\}$, $L_2 = \emptyset$.
(2) Use the AKS primality test to check whether L_1 consists exclusively of prime numbers. If so or $L_1 = \emptyset$, then:
   (a) If $L_2 = \emptyset$, then output $n = \prod_{p \in L_1} p^{v_p(n)}$ as the complete factorization of n and stop.
   (b) If $\#L_2 > 1$, then report failure and stop. In the contrary case, try to factor the only element m of L_2 into a product of two primes, $m = p^\alpha q^\beta$, assuming that $\varphi(m) = \frac{\varphi(n)}{\prod_{s \in L_1} s^{v_s(n)-1}(s-1)}$. If this works,

then output $n = p^\alpha q^\beta \prod_{s \in L_1} s^{v_s(n)}$ as the complete factorization of n and stop. Otherwise report failure and stop.

(3) Choose $d \in L_1 \setminus \mathbb{P}$.
(4) If d is a prime power p^α, then replace d by p in L_1. Return to step 2.
(5) Attempt to split d by means of the factoring algorithms corresponding evidently to Lemmas 7.2, 7.3, 7.4 and 7.6. If this produces a nontrivial factorization $d = d_1 d_2$, then further apply a factor refinement procedure (cf. [4]) to get a nontrivial factorization $d = d_1' d_2'$ with $(d_1', d_2') = 1$. Also, remove d from L_1, adjoin d_1', d_2' to L_1, and return to step 2.
(6) Remove d from L_1 and adjoin it to L_2. Return to step 2.

The algorithm obviously terminates. All we need to show is that when it does, $L_2 = \emptyset$ or $L_2 = \{m\}$, with $\omega(m) = 2$. Let d be an integer chosen in step 3 of the algorithm, d not equal to a prime power. Then d must have one of the following forms:

(i) d divisible by two distinct primes from $\mathcal{Q}_{(\ln n)^u,\delta}$;
(ii) d divisible by a prime from $\mathcal{P}_{(\ln n)^u}$, at most one prime from $\mathcal{Q}_{(\ln n)^u,\delta}$ and at most one prime factor of n_3;
(iii) d divisible by a prime q from $\mathcal{Q}_{(\ln n)^u,\delta}$ and the prime $p_-(n_3)$, $\omega(n_3) = 2$;
(iv) $d = p^{v_p(n)} q^{v_q(n)}$, where $q \in \mathcal{Q}_{(\ln n)^u,\delta}$, $p = p_+(n_3)$, $\omega(n_3) = 2$;
(v) $d = n_3$, $\omega(n_3) = 2$;
(vi) $d = n_3 q^{v_q(n)}$, where $q \in \mathcal{Q}_{(\ln n)^u,\delta}$, $\omega(n_3) = 1$.

The integer d will be split in deterministic polynomial time:

- In case (i) by Lemma 7.2.
- In case (ii) by Lemmas 7.4 and 7.6, since then $\omega(d) \le k+2$.
- In case (iii) by Lemma 7.3, because then $q > p_-(n_3)^\eta$.

Clearly, d can be adjoined to L_2 only in cases (iv)-(vi), and if it is, no other element will. □

Remarks. In part 1 of the definition of $\mathcal{N}_{k,u,\delta,\eta}$, assuming that the prime factors of n_1 belong to $\mathcal{P}_{(\ln n)^u}$ is assuming that the part of $\varphi(n)$, which can be completely factored in deterministic polynomial time with the $p-1$ method, is a multiple of $\prod_{q \mid n_1} (q-1)$. This assumption could be slightly relaxed by considering other deterministic factoring methods, such as the $p+1$ methods of section 5. Also, the condition $\omega(n_1) = k$ could be replaced by the weaker: if $q_1, \ldots, q_{k+1}$ are $k+1$ distinct primes dividing n_1, then the gcd of $q_1 - 1, \ldots, q_{k+1} - 1$, is $(\ln n)^u$-smooth.

Primality testing is a special case of the problem of testing for membership in $\{(n, \varphi(n)) : \ n \in \mathcal{N}_{k,u,\delta,\eta}\}$ or, more generally, in $\{(n, \varphi(n)) : \ n \in \mathbb{N}\}$. Indeed, the set of primes can be identified with the subset $\{(n, n-1) : \ n \in \mathbb{P}\}$ of the graph of φ. Before primality was known to be decidable in deterministic polynomial time [2], Konyagin and Pomerance [14] showed that for any fixed, positive u and δ, the set $\{q : \ q \in \mathcal{Q}_{(\ln q)^u,\delta}\}$ is recognizable in deterministic polynomial time. Some of their ideas are used in this article, but in a more synthetic way.

To conclude this section, we shall state without proof a result similar to Theorem 7.1 for the sum of divisors function σ (for a random polynomial time reduction of factoring to computing σ, cf. [4]). Let $\mathcal{R}$ be a finite subset of $\mathbb{Z}$, and let $\mathcal{R}'$ be the set of primes q such that $\left(\frac{m}{q}\right) = -1$ for some $m \in \mathcal{R}$. Moreover, let

- $\mathcal{P}_{\mathcal{R},B}$ be the subset of $\mathcal{R}'$ of such primes q that for each prime p dividing $q+1$:
  $p-1$ is B-smooth or
  $p+1$ is B-smooth and $p \in \mathcal{R}'$.
- $\mathcal{Q}_{\mathcal{R},B,\delta}$ be the subset of $\mathcal{R}'$ of such primes q that the B-smooth part of $q+1$ is not less than q^δ.

To define $\mathcal{N}_{\mathcal{R},k,u,\delta,\eta}$, replace in the definition of $\mathcal{N}_{k,u,\delta,\eta}$ the set $\mathcal{P}_{(\ln n)^u}$ by $\mathcal{P}_{\mathcal{R},(\ln n)^u}$, the set $\mathcal{Q}_{(\ln n)^u,\delta}$ by $\mathcal{Q}_{\mathcal{R},(\ln n)^u,\delta}$, and add a fourth condition:

(4) $v_q(n_1 n_2)$ is odd for all primes q dividing $n_1 n_2$.

Then the following analogue of Theorem 7.1 holds.

Theorem 7.7. *Given the pair $(n, \sigma(n))$, with $n \in \mathcal{N}_{\mathcal{R},k,u,\delta,\eta}$, the complete factorization of n can be computed in $O((\ln n)^{C'})$ deterministic time, where C' is some constant depending only upon $\mathcal{R}, k, u, \delta, \eta$. In particular, membership in $\{(n, \sigma(n)) : n \in \mathcal{N}_{\mathcal{R},k,u,\delta,\eta}\}$ is decidable in deterministic polynomial time (for $\mathcal{R}, k, u, \delta, \eta$ fixed and $\mathcal{R}$ finite).*

8. A subexponential reduction of factoring to computing φ

We shall abbreviate any expression of the form $\exp\Big((\ln x)^a (\ln\ln x)^{1-a}\Big)$ as $L(x, a)$. In this section we will first prove

Theorem 8.1. *Suppose that $\varphi(n)$ is given in a completely factored form. Then the complete factorization of n can be found in less than $L(n, \frac{1}{3})^{1+o(1)}$ deterministic time.*

Then deduce

Corollary 8.2. *Let $k = \min\{l \in \mathbb{N} : \varphi^l(n) = 1\}$. There is a deterministic algorithm that, given the sequence $(n, \varphi(n), \varphi^2(n), \dots, \varphi^k(n))$, outputs the complete factorization of n in less than $L(n, \frac{1}{3})^{1+o(1)}$ time.*

Proof. Let $1 \le m \le k$. Once we have found the complete factorization of $\varphi^m(n)$, we can compute, from Theorem 8.1, the complete factorization of $\varphi^{m-1}(n)$ in less than $L(n, \frac{1}{3})^{1+o(1)}$ deterministic time. Since $\varphi^k(n) = 1$ and $k \le 1 + \log_2 n$, the corollary follows by induction. □

In the proof of Theorem 8.1 we will exhibit a procedure that factors n recursively, that is, splits any previously computed, reducible divisor d of n further. Let $p = p_-(d)$. Additionally, let α, β, γ be real numbers from the interval $(0, 1)$, parameters to be optimally chosen. Assume that $p > L(d, 1-\alpha)$. Define $\mathcal{B}$ as $\{2, 3, \dots, [L(d, 1-\alpha)]\}$, and denote $\mathrm{LCM}_{b\in\mathcal{B}}(\mathrm{ord}_d(b))$ by A.

Lemma 8.3. *Let $(1-\beta)(1-\gamma) \le 1-\alpha$. Suppose that $\mathcal{B}$ contains no Fermat-Euclid witness for d and that $\omega(d) > \big(\frac{\ln d}{\ln\ln d}\big)^\beta$. Then $p = mA + 1$ for some integer $m < L(d, (1-\beta)\gamma)$ if p is sufficiently large.*

Proof. We have

$$L(p, 1-\gamma) \le \exp\Big(\big(\frac{1}{\omega(d)} \ln d\big)^{1-\gamma} (\ln\ln d)^\gamma\Big) < L(d, (1-\beta)(1-\gamma)) \le L(d, 1-\alpha),$$

where the last inequality holds if d is large enough. Assume that d is indeed such. As $\mathcal{B}$ contains no Fermat-Euclid witness for d, it follows that $A = \mathrm{LCM}_{b\in\mathcal{B}}(\mathrm{ord}_p(b))$.

Consequently, $A = \#\langle\mathcal{B}\rangle_p \geq \psi(p, L(p, 1-\gamma))$. By Theorem 2.2, we obtain $A \geq pL(p,\gamma)^{-1}$ if p is sufficiently large. We can write $p = mA + 1$ for some $m \in \mathbb{N}$, because $A \mid p-1$. Therefore $mA < p \leq AL(p,\gamma)$. Hence

$$m < L(p,\gamma) \leq \exp\Big((\frac{1}{\omega(d)}\ln d)^{\gamma}(\ln\ln d)^{1-\gamma}\Big) < L(d,(1-\beta)\gamma). \qquad \square$$

Lemma 8.4. *Let $\beta \leq \frac{1}{2}$, $1-\beta \geq \alpha$. Assume that there is no Fermat-Euclid witness for d in $\mathcal{B}$ and that $\omega(d) \leq \big(\frac{\ln d}{\ln\ln d}\big)^{\beta}$. Then $A^{\omega(d)+1} > d\binom{\omega(d)}{[\omega(d)/2]}$ if d is sufficiently large.*

Proof. Just as in the proof of Lemma 7.4, we have $A^{\omega(d)} \geq \#\langle\mathcal{B}\rangle_d \geq \psi(d, L(d, 1-\alpha))$. Hence $A^{\omega(d)+1} \geq \psi(d, L(d,1-\alpha))^{\frac{\omega(d)+1}{\omega(d)}}$. Let $-1 < \varepsilon < -\alpha$. It follows from Theorem 2.2 that $A^{\omega(d)+1} \geq d^{1+\frac{1}{\omega(d)}} L(d,\alpha)^{\varepsilon\frac{\omega(d)+1}{\omega(d)}}$ if d is large enough. It is sufficient to show that $d^{\frac{1}{\omega(d)}} L(d,\alpha)^{\varepsilon\frac{\omega(d)+1}{\omega(d)}} > \binom{\omega(d)}{[\omega(d)/2]}$ for large d. This is clear when $\varepsilon\frac{\omega(d)+1}{\omega(d)} \leq -1$, because then $\omega(d)$ is bounded from above. Suppose therefore that $\varepsilon\frac{\omega(d)+1}{\omega(d)} > -1$. For sufficiently large d we get

$$\begin{aligned} d^{\frac{1}{\omega(d)}} L(d,\alpha)^{\varepsilon\frac{\omega(d)+1}{\omega(d)}} &\geq L(d,1-\beta)L(d,\alpha)^{\varepsilon\frac{\omega(d)+1}{\omega(d)}} \geq L(d,1-\beta)^{1+\varepsilon\frac{\omega(d)+1}{\omega(d)}} \\ &> \exp\Big(\Big(\frac{\ln d}{\ln\ln d}\Big)^{\beta}\ln 2\Big) \geq \exp(\omega(d)\ln 2) = 2^{\omega(d)} \\ &> \binom{\omega(d)}{[\omega(d)/2]}. \end{aligned}$$

$\square$

The case $k = 3$ of the ensuing lemma was proved in [14].

Lemma 8.5. *Let $d = p_1^{e_1}\cdot\ldots\cdot p_k^{e_k}$. Assume A divides $p_i - 1$ for $i = 1,\ldots,k$; $p_i = b_iA + 1$. Suppose in addition that $A^{k+1} > \binom{k}{[k/2]}d$. Write d in base A: $d = 1 + a_1A + \ldots + a_kA^k$. Let $g = 1 + a_1X + \ldots + a_kX^k$. Then $g = (b_1X+1)\cdot\ldots\cdot(b_kX+1)$ in $\mathbb{Z}[X]$. Furthermore, this factorization can be obtained with the Hensel-Berlekamp algorithm in $O((\ln d)^5(\ln\ln d)^2)$ deterministic time.*

Proof. We have $d = p_1^{e_1}\cdot\ldots\cdot p_k^{e_k} = (b_1A+1)^{e_1}\cdot\ldots\cdot(b_kA+1)^{e_k}$. Since $A^{k+1} > d$, it follows that $e_1 = \ldots = e_k = 1$. Hence

$$1 + a_1A + \ldots + a_kA^k = (b_1A+1)\cdot\ldots\cdot(b_kA+1) = 1 + \sum_{j=1}^{k}\sigma_{k,j}(b_1,\ldots,b_k)A^j,$$

where $\sigma_{k,j}(b_1,\ldots,b_k) = \sum_{1\leq i_1<\ldots<i_j\leq k} b_{i_1}\cdot\ldots\cdot b_{i_j}$. It is therefore sufficient to show that $0 \leq \sigma_{k,j}(b_1,\ldots,b_k) < A$ for every j, $1 \leq j \leq k$. By assumption, $A^{k+1} > \binom{k}{[k/2]}d$ and thus $b_1\cdot\ldots\cdot b_k\binom{k}{[k/2]}d < b_1\cdot\ldots\cdot b_kA^{k+1} < dA$. Hence $b_1\cdot\ldots\cdot b_k\binom{k}{[k/2]} < A$ and it follows that $0 \leq \sigma_{k,j}(b_1,\ldots,b_k) \leq \binom{k}{j}b_1\cdot\ldots\cdot b_k \leq \binom{k}{[k/2]}b_1\cdot\ldots\cdot b_k < A$.

It remains to prove that g can be completely factored in the stated time. We first need a "small" prime p not dividing a_k and such that g_p is squarefree, g_p being the reduction of g modulo p. An upper bound for such a p is given in [16] (3.9): $p = O(k\ln k + k\ln|g|)$, where $|g| := (1 + \sum_{i=1}^{k} a_i^2)^{\frac{1}{2}}$. Verifying that $p = O((\ln d)^2)$ is straightforward. Let $\alpha = a_k^{-1}(p)$, $e = \lceil\frac{\ln d}{\ln p}\rceil$. We factor completely αg_p with

the Berlekamp algorithm in $O(k(k+p)(k\ln p)^2) = O((\ln d)^5(\ln\ln d)^2)$ deterministic time (cf. Theorem 7.4.5 of [5]). Then we lift this factorization to the factorization $\prod_{1\le i\le k}(x+b_i^{-1})$ modulo p^e with the Hensel algorithm in $O(ke(k\ln p)^2) = O((\ln d)^4(\ln\ln d)^2)$ deterministic time (cf. Theorem 7.7.2 of [5]). Finally, we compute b_i (p^e) for every i. This finishes the proof, as each b_i is less than p^e. □

Proof of Theorem 8.1. We find the complete factorization of n using the algorithms associated with Lemmas 8.3, 8.4 and 8.5. The running time bound of our recursive procedure is obviously less than $L(n, \max(1-\alpha,(1-\beta)\gamma))^{1+o(1)}$. It remains to minimize $\max(1-\alpha,(1-\beta)\gamma)$ over the set

$$\{(\alpha,\beta,\gamma):\ 0<\alpha<1, 0<\beta\le\frac{1}{2}, 0<\gamma<1, 1-\beta\ge\alpha, (1-\beta)(1-\gamma)\le 1-\alpha\}.$$

Some easy calculations show that the minimum is $\frac{1}{3}$, reached for $\alpha=\frac{2}{3}$, $\beta=\frac{1}{3}$, $\gamma=\frac{1}{2}$. □

Remark 8.6. The above method reduces the factorization of Carmichael numbers n to the factorization of $n-1$ in less than $L(n,\frac{1}{3})^{1+o(1)}$ deterministic time.

Acknowledgements

This paper contains part of the author's doctoral dissertation, written under the supervision of Dr. Jacek Pomykała. It is a pleasure to thank him for all his help, encouragement and kindness.

References

1. L. M. Adleman, K. S. McCurley, *Open problems in number-theoretic complexity. II*, Algorithmic Number Theory Symposium **I** (1994), 291-322. MR1322733 (95m:11142)
2. M. Agrawal, N. Kayal, N. Saxena, *PRIMES is in P*, Annals of Mathematics (2), **160** (2004), 781–793. MR2123939 (2006a:11170)
3. E. Bach, *Explicit bounds for primality testing and related problems*, Mathematics of Computation, **55** (1990), 355-380. MR1023756 (91m:11096)
4. E. Bach, G. Miller, J. O. Shallit, *Sums of divisors, perfect numbers and factoring*, SIAM Journal on Computing, **15** (1986), 1143-1154. MR861378 (87k:11139)
5. E. Bach, J. O. Shallit, *Algorithmic number theory, Volume 1: Efficient algorithms*, MIT Press, 1996. MR1406794 (97e:11157)
6. E. Bach, J. O. Shallit, *Factoring with cyclotomic polynomials*, Mathematics of Computation, **52** (1989), 201-219. MR947467 (89k:11127)
7. E. R. Berlekamp, *Factoring polynomials over finite fields*, Bell Systems Technical Journal, **46** (1967), 1853-1859. MR0219231 (36:2314)
8. R. J. Burthe, Jr., *The average least witness is* 2, Acta Arithmetica, **80** (1997), 327-341. MR1450927 (98h:11118)
9. E. R. Canfield, P. Erdös, C. Pomerance, *On a problem of Oppenheim concerning "factorisatio numerorum"*, Journal of Number Theory, **17** (1983), 1-28. MR712964 (85j:11012)
10. D. Coppersmith, N. Howgrave-Graham, S. V. Nagaraj, *Divisors in residue classes, constructively*, Mathematics of Computation, **77** (2008), 531-545. MR2353965 (2009b:11221)
11. M. R. Fellows, N. Koblitz, *Self-witnessing polynomial-time complexity and prime factorization*, Designs, Codes and Cryptography, **2** (1992), 231-235. MR1181730 (93e:68032)
12. M. Fürer, *Deterministic and Las Vegas primality testing algorithms*, Lecture Notes in Computer Science, **194** (1985), 199-209. MR819255 (87c:11123)
13. K. Hensel, *Neue Grundlagen der Arithmetic*, Journal für die Reine und Angewandte Mathematik, **127** (1904), 51-84.
14. S. Konyagin, C. Pomerance, *On primes recognizable in deterministic polynomial time*, The Mathematics of Paul Erdös, R. L. Graham, J. Nesetril, eds., Springer-Verlag, 1997, 176-198. MR1425185 (98a:11184)

15. S. Landau, *Some remarks on computing the square parts of integers*, Information and Computation, **78**, No. 3 (1988), 246-253. MR959811 (89k:11003)
16. A. K. Lenstra, H. W. Lenstra, Jr., L. Lovász, *Factoring polynomials with rational coefficients*, Mathematische Annalen, **261** (1982), 515-534. MR682664 (84a:12002)
17. H. W. Lenstra, Jr., *Factoring integers with elliptic curves*, Annals of Mathematics (2), **126** (1987), 649-673. MR916721 (89g:11125)
18. H. W. Lenstra, Jr., C. Pomerance, *Primality testing with Gaussian periods*, preliminary version, July 20, 2005.
19. G. L. Miller, *Riemann's Hypothesis and tests for primality*, Journal of Computer and System Sciences, **13** (1976), 300-317. MR0480295 (58:470a)
20. S. C. Pohlig, M. E. Hellman, *An improved algorithm for computing logarithms over $GF(p)$ and its cryptographic significance*, IEEE Transactions on Information Theory, **24** (1978), 106-110. MR0484737 (58:4617)
21. J. M. Pollard, *Theorems on factorization and primality testing*, Proceedings of the Cambridge Philosophical Society, **76** (1974), 521-528. MR0354514 (50:6992)
22. M. O. Rabin, *Probabilistic algorithm for testing primality*, Journal of Number Theory, **12** (1980), 128-138. MR566880 (81f:10003)
23. R. L. Rivest, A. Shamir, L. M. Adleman, *A method for obtaining digital signatures and public-key cryptosystems*, Communications of the ACM, **21** (1978), 120-126. MR700103 (83m:94003)
24. V. Strassen, *Einige Resultate über Berechnungskomplexität*, Jahresbericht der Deutschen Mathematiker-Vereinigung, **78** (1976), 1-8. MR0438807 (55:11713)
25. J. W. M. Turk, *Fast arithmetic operations on numbers and polynomials*, Computational Methods in Number Theory **I** (1982), 43-54. MR700257 (84f:10006)
26. H. C. Williams, *A $p+1$ method of factoring*, Mathematics of Computation, **39** (1982), 225-234. MR658227 (83h:10016)
27. B. Źrałek, *Using the smoothness of $p-1$ for computing roots modulo p*, submitted, preliminary version available on http://arxiv.org/abs/0803.0471.

Institute of Mathematics, Polish Academy of Sciences, 00-956 Warsaw, Poland
E-mail address: b.zralek@impan.gov.pl