A deterministic version of Pollard's $p-1$ algorithm

Bartosz Źrałek *

May 13, 2009

Abstract

We present a deterministic version of Pollard's $p-1$ integer factoring algorithm. More precisely, we prove that the integers factorable (partially or completely) in random polynomial time by the $p-1$ algorithm are in fact factorable in deterministic polynomial time. We also point out two large classes of integers n (almost all) that have simple arithmetic characterization in terms of multiplicative structure and are completely factorable in deterministic polynomial time if $\varphi(n)$ is given and partially factored (φ denotes Euler's totient function). Finally, we show that $O(\ln n)$ oracle calls 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

The design and implementation of probabilistic algorithms confronts us not only with mathematical and philosophical, but also practical questions. Such algorithms can in theory run forever, failing to find a proper result. Moreover we do not even know if truly random events occur in a real life. Finally, generating "long" pseudorandom number sequences is a difficult task on standard hardware. That is why the derandomization of probabilistic algorithms (hopefully without assuming hypotheses such as the Extended Riemann Hypothesis - ERH) so much occupies computational complexity theory and, if we restrict ourselves to number-theoretic algorithms, computational number theory. Admittedly the most spectacular result in this

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

direction is the Agrawal-Kayal-Saxena deterministic polynomial time algorithm for primality proving (AKS).

In this article we show how to derandomize Pollard's $p-1$ integer factoring algorithm [15]. More precisely, we show a deterministic polynomial time algorithm that splits (or proves the primality of) any integer divisible by a prime p such that $p-1$ is smooth. By iterating this algorithm and using the AKS test (to be precise the the Lenstra-Pomerance variant [12]), we get a procedure that factors completely in deterministic polynomial time integers having at most one prime divisor p such that $p-1$ is not smooth (corollary 4.6). The proofs of correctness rely on techniques devised by Pohlig and Hellman [14] (theorem 4.2) and Fellows and Koblitz [8] (theorem 4.5). The latter was developed by Konyagin and Pomerance [10].

We also attempt to make some progress on the open problem of reducing factoring in deterministic polynomial time to computing Euler's totient function φ (see [1]). Miller [13] found a reduction whose correctness depends on the ERH. Rabin [16] obtained an unconditional reduction at the cost of giving up determinism. A relatively recent result of Burthe [5] yields a reduction for almost all integers, but these integers cannot be simply described. The above reductions require only the single value of $\varphi(n)$, where n is the integer to be factored. Here we consider a relaxed problem: is there a deterministic algorithm that given $\varphi(n)$ and its (partial or complete) factorization factors completely n in polynomial time? It should be noted in the first place that this question is not trivial. Indeed, Erdös and Pomerance [7] have proved that the normal number of distinct prime divisors of $\varphi(n)$ is $\frac{1}{2}\ln^2 \ln n$. Therefore, a 'naive' combinatorial search for a prime factor of n among all the elements of the form $d+1$, where d is an even divisor of $\varphi(n)$, does not lead to a polynomial time algorithm. We point out two large classes of integers n (almost all) that have simple arithmetic characterization in terms of multiplicative structure and for which the answer to our question is 'yes' (theorems 6.4, 6.7, see also 6.8). Notice that a positive answer for all n would imply a solution to the original reduction problem. One could compute $\varphi(n)$ and all the iterations $\varphi^2(n), \dots, \varphi^k(n)$, where k is the least integer such that $\varphi^k(n) = 1$ ($k \leq 1 + \log_2 n$); then factor $n = \varphi^0(n)$ by induction, that is, by factoring $\varphi^{k-1}(n)$ given the factorization of $\varphi^k(n)$ and so on. Although we leave our question open, we prove that every n can be, given $\varphi(n)$ and its complete factorization, completely factored in less than $\exp\Big((1+o(1))(\ln n)^{\frac{1}{3}}(\ln\ln n)^{\frac{2}{3}}\Big)$ deterministic time (theorem 7.3). Consequently factoring is reducible in deterministic subexponential time to computing φ (corollary 7.5). The result is far from being satisfactory, still seems to be of some interest, as we are unaware of any previously proposed non-exponential deterministic reduction in this context.

2 Notation

Throughout the text n is an odd integer, 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.
By $\langle\mathcal{B}\rangle_G$ we mean the subgroup of the group G generated by the set $\mathcal{B}$, by $\mathrm{ord}_G(b)$ the order of the element b in G. If $G = \mathbb{Z}_m^*$ we simplify these notations and write $\langle\mathcal{B}\rangle_m$, respectively $\mathrm{ord}_m(b)$. 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 $a \in \mathbb{Z}_n^* = \bigoplus_{q|n} \mathbb{Z}^*_{q^{v_q(n)}}$.
The Lenstra-Pomerance variant of the AKS test is abbreviated as $\mathrm{Isprime}(\cdot)$.
We recall the definitions of the familiar number theoretic functions appearing in the text:

$$\varphi(n) = \#\{m \le n : \ (m, n) = 1\} \text{ (Euler's totient)},$$

$$\omega(n) = \sum_{p|n} 1 \text{ and } \Omega(n) = \sum_{p|n} v_p(n),$$

$$\psi(x, z) = \#\{m \le x : \ p_+(m) \le z\}.$$

We will make frequent use of the following theorem proved in [10]

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

and 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.) $\psi(x, x^{\frac{1}{u}}) = xu^{-u+o(u)}$ *uniformly for* $u = u(x) \to +\infty$ *and* $u < (1 - \varepsilon)\frac{\ln x}{\ln \ln x}$.

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 $b^M = 1(p)$ 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. $b^M = 1(n)$? 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 $n > 2$ be odd, M even, $\mathcal{F}(M) = \{b \in \mathbb{Z}_n^* : b^M \neq 1(n)\}$, $\mathcal{S}(M) = \{b \in \mathbb{Z}_n^* \setminus \mathcal{F}(M) : \exists_{1 \le l \le v_2(M)}\ 1 < (b^{\frac{M}{2^l}} - 1, n) < n\}$. 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 $p_+(p-1) \le \ln^u n$. Set $M = \prod_{q \le \ln^u n} q^{\left[\frac{\ln n}{\ln q}\right]}$. Then M satisfies the two conditions, since $\ln M \le \sum_{q \le \ln^u n} \frac{\ln n}{\ln q} \ln q = \pi(\ln^u n) \ln n = O\left(\frac{\ln^{u+1} n}{u \ln \ln n}\right)$ 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 $b^M = 1(n)$ for all $b < 2\ln^2 n$.

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

Using this theorem, we can easily prove the existence of $b < 2\ln^2 n$ 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 $\mathcal{B} = \{2, 3, \ldots, [\ln^2 n]\}$. Assume that we are given an integer M together with its complete factorization such that $b^M = 1(n)$ 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 [8]. We restate the argument for the completeness and clarity of exposition.

Theorem 4.1 (Fellows-Koblitz) *Let $\mathcal{B} = \{2, 3, \ldots, [\ln^2 n]\}$, $\mathcal{B} \subset \mathbb{Z}_n^*$. Then n is prime if and only if the following conditions are satisfied.*
(1) $ord_p(b) = ord_n(b)$ for every $b \in \mathcal{B}$ and $p \mid n$.
(2) $LCM_{b\in\mathcal{B}}(ord_n(b)) > \sqrt{n}$.

Proof. Suppose n is prime. Condition (1) is then a tautology. We check condition (2). Let $H = \langle \mathcal{B} \rangle_n$. H is cyclic, since n is prime. Therefore $\mathrm{LCM}_{b\in\mathcal{B}}(\mathrm{ord}_n(b)) = \#H \geq \psi(n, \ln^2 n) > \sqrt{n}$, where the last inequality follows from theorem 2.1.
Assume now that conditions (1) and (2) 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 $n \in \mathbb{P}$. $\square$

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 say that b is a Fermat-Euclid witness for n. Checking conditions (1) and (2) therefore reduces to factoring the orders of the elements of $\mathcal{B}$, which can be done efficiently under our assumption on M. Now our task being not primality testing, but factorization, we would like conditions (1) and (2) to be effective, that is violation of one of them to lead to a nontrivial divisor of n. As noted above condition (1) is effective, unfortunately condition (2) is not. However it is easy to see that if $\mathrm{LCM}_{b\in\mathcal{B}}(\mathrm{ord}_n(b)) \leq \sqrt{n}$ then $\langle \mathcal{B} \rangle_n$ is not cyclic.
Suppose for greater generality that $\mathcal{B}$ is any "small" subset of $\mathbb{Z}_n^*$ whose

elements have "smooth" orders in $\mathbb{Z}_n^*$. We will describe below an efficient deterministic algorithm that finds a nontrivial divisor of n if $\langle \mathcal{B} \rangle_n$ is not cyclic or a generator of $\langle \mathcal{B} \rangle_n$ otherwise. By induction, it is sufficient to restrict our attention to the case $\#\mathcal{B} = 2$, say $\mathcal{B} = \{a, b\}$.
We assume temporarily that $\mathrm{ord}_n(a) = s^v$, $b^{s^v} = 1$, where $s \in \mathbb{P}$. 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 $\mathrm{ord}_n(a) = s^v = \mathrm{ord}_{p_i^{e_i}}(a_i)$. Since $\mathbb{Z}_{p_i^{e_i}}^*$ is cyclic, $b_i \in \langle a_i \rangle_{p_i^{e_i}}$. Write $b_i = a_i^{l_0 + l_1 s + \ldots + l_{v-1} s^{v-1}}$ for some (uniquely determined) l_m, $0 \le l_m < s$. Then $b_i^{s^{v-1}} = a_i^{l_0 s^{v-1}}$, that is $p_i^{e_i} \mid b^{s^{v-1}} - a^{l_0 s^{v-1}}$. We can thus find a j, $0 \le j < s$, such that $d := (b^{s^{v-1}} - a^{j s^{v-1}}, n) > 1$. If $d < n$ then d is a nontrivial divisor of n, so suppose that $d = n$. Therefore $b^{s^{v-1}} = a^{j s^{v-1}}$. In particular $b_i^{s^{v-1}} = a_i^{j s^{v-1}}$ and $j = l_0$. Now reason by induction. Let $m < v - 2$, $c = b a^{-(l_0 + \ldots + l_m s^m)}$; assume that $c^{s^{v-m-1}} = 1$. Then $c_i = a_i^{l_{m+1} s^{m+1} + \ldots + l_{v-1} s^{v-1}}$, hence $p_i^{e_i} \mid c^{s^{v-m-2}} - a^{l_{m+1} s^{v-1}}$. As before, either we find a j, $0 \le j < s$, such that $1 < (c^{s^{v-m-2}} - a^{j s^{v-1}}, n) < n$, or $c^{s^{v-m-2}} = a^{l_{m+1} s^{v-1}}$. If this procedure does not lead to a nontrivial factor of n then $b = a^{l_0 + \ldots + l_{v-1} s^{v-1}}$ and thus $b \in \langle a \rangle_n$. 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) $c = b$
(3) For $j = 1$ to $s - 1$ compute $a^{j s^{v-1}}$
(4) For $m = 0$ to $v - 1$ do
  (5) $j = 0$
  (6) While $(a^{j s^{v-1}} - c^{s^{v-m-1}}, n) = 1$ do $j = j + 1$
  (7) If $(a^{j s^{v-1}} - c^{s^{v-m-1}}, n) \neq n$ then return $(a^{j s^{v-1}} - c^{s^{v-m-1}}, n)$
  (8) Let $c = c a^{-j s^m}$

Theorem 4.2 *Let $a, b \in \mathbb{Z}_n^*, s \in \mathbb{P}$, $ord_n(a) = s^v$, $ord_n(b) = s^w$. 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 $O((\max(v, w) + s) \max(v, w) \ln^2 n \ln s)$ operations.*

Proof. The correctness of PH(n, a, b, s, v, w) follows from the precedent discussion. Step (3) requires $\sum_{j=1}^{s-1} O(\ln^2 n \ln(j s^{v-1})) = O(v s \ln^2 n \ln s)$ operations. The total number of operations used by step (6) in the loop (4) is

$O(v^2 \ln^2 n \ln s)$. Step (8) takes on the whole in the loop (4) $\sum_{m=0}^{v-1} O(\ln^2 n \ln s^{m+1}) = O(v^2 \ln^2 n \ln s)$ operations. Hence the stated running time. $\square$

Remark 4.3 *If we compute in step (3) $a^{js^{v-1}}$, $a^{js^{v-1}\lceil\sqrt{s}\rceil}$ for $j = 1, \dots, \lceil\sqrt{s}\rceil$ and use Shanks's baby-steps, giant-steps method in (6) to find j such that $(a^{js^{v-1}} - c^{s^{v-m-1}}, n) > 1$ then the running time bound of PH(n, a, b, s, v, w) can be improved to $O((\max(v, w) + \sqrt{s} \ln s) \max(v, w) \ln^2 n \ln s)$.*

Corollary 4.4 *Let $a, b \in \mathbb{Z}_n^*$, $A = ord_n(a)$, $B = ord_n(b)$. If none of the procedures PH$(n, a^{\frac{A}{s^{v_s(A)}}}, b^{\frac{B}{s^{v_s(B)}}}, s, v_s(A), v_s(B))$, where $s \in \mathbb{P}$ and $s \mid (A, B)$, finds a nontrivial divisor of n then $\langle a, b\rangle_n$ is cyclic. The overall time used by these procedures is $O((\ln n + p) \ln^3 n)$, where $p = \max\{s \mid (A, B)\}$.*

Proof. Suppose that none of the procedures PH$(n, a^{\frac{A}{s^{v_s(A)}}}, b^{\frac{B}{s^{v_s(B)}}}, s, v_s(A), v_s(B))$, $s \mid (A, B)$, finds a nontrivial factor of n. We see from theorem 4.2 that

$$\forall_{s\in\mathbb{P}, s|(A,B)} \quad b^{\frac{B}{s^{v_s(B)}}} \in \langle a^{\frac{A}{s^{v_s(A)}}}\rangle_n \text{ or } a^{\frac{A}{s^{v_s(A)}}} \in \langle b^{\frac{B}{s^{v_s(B)}}}\rangle_n \tag{4.1}$$

Set $c_s = b^{\frac{B}{s^{v_s(B)}}}$ if $v_s(A) \le v_s(B)$, else $c_s = a^{\frac{A}{s^{v_s(A)}}}$ $(s \in \mathbb{P}, s \mid AB)$. We claim that $\langle a, b\rangle_n = \langle c_s : s \mid AB\rangle_n$. The inclusion $\langle c_s : s \mid AB\rangle_n \subset \langle a, b\rangle_n$ being obvious, we justify the reverse. Let $s \mid A$. If $s \nmid B$ then $v_s(A) \ge 1 > 0 = v_s(B)$ and $a^{\frac{A}{s^{v_s(A)}}} = c_s$. If $s \mid B$ then (4.1) implies $a^{\frac{A}{s^{v_s(A)}}} = c_s^{\alpha(s)}$ for some $\alpha(s)$. As the elements $a^{\frac{A}{s^{v_s(A)}}}$ $(s \mid A)$ have pairwise relatively prime orders in $\mathbb{Z}_n^*$, we get $\text{ord}_n(\prod_{s|A} a^{\frac{A}{s^{v_s(A)}}}) = \prod_{s|A} s^{v_s(A)} = A$. Therefore $\langle a\rangle_n = \langle\prod_{s|A} a^{\frac{A}{s^{v_s(A)}}}\rangle_n \subset \langle c_s : s \mid AB\rangle_n$. In a similar fashion $b \in \langle c_s : s \mid AB\rangle_n$. Hence $\langle a, b\rangle_n = \langle c_s : s \mid AB\rangle_n$. Since the elements $c_s \in \mathbb{Z}_n^*$ have pairwise relatively prime orders, it follows that the group $\langle a, b\rangle_n$ is cyclic, generated by $c := \prod_{s|AB} c_s$.

The procedures PH() above on the whole cost, by theorem 4.2,
$\sum_{s|(A,B)} O((\max(v_s(A), v_s(B)) + s) \ln^2 n \ln(s^{\max(v_s(A),v_s(B))})) =$
$O((\ln n + p) \ln^2 n \ln(\text{LCM}(A, B))) = O((\ln n + p) \ln^3 n)$ operations. $\square$

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$(n, M, s_1, v_1, \dots, s_t, v_t)$ $\{M = s_1^{v_1} \cdot \ldots \cdot s_t^{v_t}$ is the complete factorization of $M\}$
(1) For $b = 2$ to $[\ln^2 n]$ do
(2) If $b^M \neq 1(n)$ then return ('failure') and stop
(3) Compute $B = \mathrm{ord}_n(b)$
(4) For every prime s dividing B do
(5) If $(b^{\frac{B}{s}} - 1, n) > 1$ then return $(b^{\frac{B}{s}} - 1, n)$
(6) If $\mathrm{LCM}(\mathrm{ord}_n(2), \dots, \mathrm{ord}_n([\ln^2 n])) > \sqrt{n}$ then return (n 'is prime')
(7) Let $a = 2$, $A = \mathrm{ord}_n(2)$
(8) For $b = 3$ to $[\ln^2 n]$ do
(9) For every prime s dividing (A, B), where $B = \mathrm{ord}_n(b)$, do
(10) $\mathrm{PH}(n, a^{\frac{A}{s^{v_s(A)}}}, b^{\frac{B}{s^{v_s(B)}}}, s, v_s(A), v_s(B))$
(11) If a factor d of n has been found then return d
(12) Let $a = \prod_{s|AB} c_s$, where
$c_s = b^{\frac{B}{s^{v_s(B)}}}$ if $v_s(A) \leq v_s(B)$, else $c_s = a^{\frac{A}{s^{v_s(A)}}}$,
let $A = \mathrm{LCM}(A, B)$

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

Proof. Step (3) requires $O\big(\ln^2 n \frac{\ln^2 M}{\ln\ln M}\big)$, step (4) $O\big(\frac{\ln^4 n}{\ln\ln n}\big)$ operations. Note that if step (6) is reached then $\{2, 3, \dots, [\ln^2 n]\}$ contains no Fermat-Euclid witness for n. Hence in step (9) $B \mid q-1$ for every $q \mid n$. By corollary 4.4, step (9) requires therefore $O((\ln n + s_0)\ln^3 n)$ operations. The time used by the remaining steps is negligible. The stated running time follows. □

We have not proven that the running time bound of Split$(n, M, s_1, v_1, \dots, s_t, v_t)$ is polynomial in $\ln n$ and $\ln M$, but it actually is if the integer s_0 defined in theorem 4.5 is small, say $s_0 \leq \ln^u n$ for some fixed u, $u \geq 1$. This inequality is obviously satisfied whenever n has a prime divisor p such that $p - 1$ is $\ln^u n$ smooth.

Corollary 4.6 (deterministic version of Pollard's $p-1$ algorithm) *Let $M = \prod_{q \leq \ln^u n} q^{\left[\frac{\ln n}{\ln q}\right]}$. (i) Assume n has a prime divisor p such that $p - 1 \mid M$.*

Then we can find a nontrivial divisor (or a proof of the primality) of n in $O(\ln^{2u+6} n)$ deterministic time.
(ii) Suppose in addition that n has at most one prime divisor p such that $p-1 \nmid M$. Then we can obtain the complete factorization of n in $O(\ln^{2u+7} n)$ deterministic time.

Proof. (i) The ensuing deterministic algorithm, whose correctness follows from theorem 4.5, finds a nontrivial divisor (or a proof of the primality) of n.

Split2$(n, M, s_1, v_1, \ldots, s_t, v_t)$ $\{M = s_1^{v_1} \cdot \ldots \cdot s_t^{v_t}$ is the complete factorization of $M\}$
(1) If n is a nontrivial power then return $n^{1/m}$, where m is the largest integer such that $2 \leq m \leq \ln_2 n$ and $n^{1/m} \in \mathbb{N}$
(2) Isprime(n)
(3) If n is prime then return (n 'is prime')
(4) For $b = 2$ to $[\ln^2 n]$ do
    (5) If $(b, n) \neq 1$ then return (b, n)
    (6) Let $d = (b^M - 1, n)$
    (7) If $d = 1$ then return ('failure') and stop
    (8) If $d < n$ then return d
(9) Split$(n, M, s_1, v_1, \ldots, s_t, v_t)$

Step (2) requires, as was shown in [12], $O(\ln^6 n \ln^c \ln n)$ operations for some constant c. We have $\ln M = O\big(\frac{\ln^{u+1} n}{u \ln\ln n}\big)$. From theorem 4.5 the running time of Split2$(n, M, s_1, v_1, \ldots, s_t, v_t)$ is therefore $O(\ln^{2u+6} n)$ operations.
(ii) We obtain the complete factorization of n by iterating Split2$(n, M, s_1, v_1, \ldots, s_t, v_t)$.

Factor$(n, M, s_1, v_1, \ldots, s_t, v_t)$
(1) Let d =Split2$(n, M, s_1, v_1, \ldots, s_t, v_t)$
(2) If d=(x 'is prime') then return x
(3) Factor$(d, M, s_1, v_1, \ldots, s_t, v_t)$
(4) Factor$(n/d, M, s_1, v_1, \ldots, s_t, v_t)$

Factor$(n, M, s_1, v_1, \ldots, s_t, v_t)$ calls Split2$(n, M, s_1, v_1, \ldots, s_t, v_t)$ $O(\ln n)$ times. Hence Factor$(n, M, s_1, v_1, \ldots, s_t, v_t)$ runs in $O(\ln^{2u+7} n)$ time. □

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 $O(\ln^{u+3} n)$ random time under the assumption of corollary 4.6 (i). Our deterministic version also runs in polynomial time, but is considerably

slower and is thus rather of theoretical than practical interest.
Of course, the obtained running time bound of $\mathrm{Split}(n, M, s_1, v_1, \ldots, s_t, v_t)$ is polynomial in $\ln n$ and $\ln M$ for more integers n than those considered in corollary 4.6. Let $D(n,u) = \max_{q > \ln^u n} \#\{p \mid n : \ q \mid p-1\}$. 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 6.

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

Proof. We have:

$$B(x,u,l) \leq \sqrt{x} + \sum_{\sqrt{x} < n \leq x} \sum_{q > \ln^u n} \sum_{\substack{p_1 < \ldots < p_{l+1} \\ p_i \mid n \\ p_i = 1(q)}} 1 \leq \sqrt{x} + \sum_{q > 2^{-u} \ln^u x} \sum_{\sqrt{x} < n \leq x} \sum_{\substack{p_1 < \ldots < p_{l+1} \\ p_i \mid n \\ p_i = 1(q)}} 1$$

$$\sum_{n \leq x} \sum_{\substack{p_1 < \ldots < p_{l+1} \\ p_i \mid n \\ p_i = 1(q)}} 1 = \sum_{\substack{p_1 < \ldots < p_{l+1} \leq x \\ p_i = 1(q)}} \left[\frac{x}{p_1 \cdot \ldots \cdot p_{l+1}} \right] \leq x \sum_{\substack{p_1 < \ldots < p_{l+1} \leq x \\ p_i = 1(q)}} \frac{1}{p_1 \cdot \ldots \cdot p_{l+1}}$$

$$\leq x \Bigg(\sum_{\substack{p \leq x \\ p = 1(q)}} \frac{1}{p} \Bigg)^{l+1} \leq \frac{c_1 x \ln^{l+1} \ln x}{(q-1)^{l+1}},$$

where the last inequality follows from the uniform bound

$$\sum_{\substack{p \leq x \\ p = 1(d)}} \frac{1}{p} \leq \frac{c_0}{\varphi(d)} \ln \ln x$$

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

$$\sum_{q > 2^{-u} \ln^u x} \sum_{\sqrt{x} < n \leq x} \sum_{\substack{p_1 < \ldots < p_{l+1} \\ p_i \mid n \\ p_i = 1(q)}} 1 \leq c_1 x \ln^{l+1} \ln x \sum_{q > 2^{-u} \ln^u x} \frac{1}{(q-1)^{l+1}}$$

$$\leq c_2 x \frac{2^{lu} \ln^{l+1} \ln x}{\ln^{lu} x}$$

Thus

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

□

5 Some known reductions of factoring to computing φ

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

Theorem 5.1 (Rabin) *Given $\varphi(n)$ we can completely factor n in $O(\ln^4 n)$ expected time.*

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

Theorem 5.2 (Miller) *If the ERH holds, then given $\varphi(n)$ we can completely factor n in $O(\ln^6 n)$ deterministic time.*

Define $G(n)$ as the least integer m such that $\mathbb{Z}_n^*$ is generated by integers less than or equal m and coprime to n. In [5] Burthe proved that $\frac{1}{x} \sum_{n \leq x} G(n) = O(\ln^{97} x)$. In particular, $G(n) < \ln^{97+\varepsilon} n$ 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 $G(n)$. It follows by a similar argument to the one used after theorem 3.2 that given $\varphi(n)$ we can completely factor n in $O(\ln^{101+\varepsilon} n)$ 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. It is worth making here a general observation. The value $\varphi(n)$ can be used to check whether a nontrivial factorization of n is a factorization into primes.

Lemma 5.3 *(i) For every $n \geq 2$ we have $\varphi(n^\alpha) \leq n^{\alpha-1}(n-1)$ with equality if and only if $n \in \mathbb{P}$.*

(ii) Suppose that $n = \prod_{i=1}^{k} n_i^{\alpha_i}$, where the n_i are pairwise coprime integers greater than 1. Then $\varphi(n) = \prod_{i=1}^{k} n_i^{\alpha_i-1}(n_i-1)$ if and only if each n_i is prime.

Proof. Showing (i) is straightforward. (ii) is a consequence of (i). □

Landau [11] showed that computing the equal order factorization of any integer n, that is the sequence $n_i := \prod_{p:\ v_p(n)=i} p$, 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 $\omega(n) \geq 3$ then $O(\Omega(n)\ln^2 n)$ bit operations and at most $\omega(n)-2$ oracle calls (including $\varphi(n)$) are needed. Notice that if $\omega(n_i) \leq 2$ 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 $O(\ln^3 n)$ deterministic time.

6 Factoring almost all n in deterministic polynomial time when $\varphi(n)$ is given and (partially) factored

Recall the definition of $D(n,u)$ from section 4. Fix the parameter $u \geq 1$. We have seen that $D(n,u) \leq 1$ for almost all integers n. Moreover, n is factorable in deterministic polynomial time if $D(n,u) \leq 1$ and the integer $\varphi(n)$ as well as its complete factorization are known. It is natural to ask whether these undeniably strong assumptions can be weakened. To answer this question we first introduce some supplementary notations.
Let M be a divisor of $\varphi(n)$ and $0 < y \leq 1$ a parameter. Set

- $F(n,M,y) = \#\{p \mid n :\ (M, p-1) \geq (p-1)^y\}$,
- $D(n,M,u) = \max_{q > \ln^u n} \#\{p \mid n :\ q \mid (M, p-1)\}$.

From now on suppose that the integer M is given in a completely factored form. Then $F(n,M,y)$ measures the part of the factorization of $\varphi(n)$ that we know. If d is factor of n then $D(n,M,u)$ measures the efficiency of the algorithm PH(d, a, b, s, v, w): it runs in polynomial time in $\ln n$ whenever $\omega(d) > D(n, M, u)$, $s \mid M$ and $s \mid p-1$ for every $p \mid d$. In the three subsections below we will consider the problem of factoring n in deterministic polynomial time for different values of y, $F(n, M, y)$ and $D(n, M, u)$.

6.1 $F(n, M, 1) \geq \omega(n) - 2$, $D(n, M, u) \leq l$ $(l \in \mathbb{N})$

Let $\eta > 0$ be arbitrary. Assume that n has no prime divisor below $\ln^{l+1+\eta} n$. For simplicity, we first suppose that $F(n, M, 1) = \omega(n)$. Assume that we have found a divisor $d > 1$ of n. We now describe a procedure that leads to a nontrivial divisor or a proof of the primality of d in deterministic polynomial time (in $\ln n$). The complete factorization of n can be thus obtained by induction.

Let $\mathcal{B}_1 = \{2, 3, \ldots, [\ln^{1+\frac{1}{l}} d]\}$, $\mathcal{B}_2 = \{2, 3, \ldots, [\ln^{l+1+\eta} d]\}$. By our assumptions on n and M we have $(b^M - 1, d) > 1$ for every $b \in \mathcal{B}_2$. We can suppose that $b^M = 1(d)$ for all $b \in \mathcal{B}_2$, because in the contrary case a nontrivial divisor of d is found.

We first check whether the set $\mathcal{B}_1$ contains a Fermat-Euclid witness for d. If this is the case then we get a nontrivial divisor of d, so assume the contrary. Let $A_1 = \mathrm{LCM}_{b \in \mathcal{B}_1}(\mathrm{ord}_d(b))$. Then $A_1 \mid p-1$ for every $p \mid d$. Suppose that $p_+(A_1) \leq \ln^u n$; note that this inequality is satisfied if $\omega(d) > l$, since $D(n, M, u) \leq l$. We follow the procedure from section 4 to test if $\langle \mathcal{B}_1 \rangle_d$ is cyclic. We can suppose that this is the case, for otherwise we find a nontrivial divisor of d.

Lemma 6.1 *Let $\mathcal{B}_1 = \{2, 3, \ldots, [\ln^{1+\frac{1}{l}} d]\}$, $\mathcal{B}_1 \subset \mathbb{Z}_d^*$. Assume that $\langle \mathcal{B}_1 \rangle_d$ is cyclic and none of the elements of $\mathcal{B}_1$ is a Fermat-Euclid witness for d. Then $\omega(d) \leq l$.*

Proof. A similar argument to that in the proof of theorem 4.1 shows that $p_-(d) > d^{\frac{1}{l+1}}$. This yields $\omega(d) \leq l$. $\square$

By the above lemma there is no loss of generality in assuming that $\omega(d) \leq l$. We test whether there is a Fermat-Euclid witness for d among the elements of $\mathcal{B}_2 \setminus \mathcal{B}_1$. For the same reasons as above, suppose that there is no such witness. We compute $A_2 = \mathrm{LCM}_{b \in \mathcal{B}_2}(\mathrm{ord}_d(b))$. Therefore $A_2 \mid p-1$ for every $p \mid d$. In particular, $(A_2, d) = 1$.

Lemma 6.2 *Let* $\mathcal{B}_2 = \{2, 3, \ldots, [\ln^{l+1+\eta} d]\}$, $\mathcal{B}_2 \subset \mathbb{Z}_d^*$, $A_2 = LCM_{b \in \mathcal{B}_2}(ord_d(b))$. *Assume that* $\omega(d) \le l$ *and* $(A_2, d) = 1$. *Then* $A_2^{\omega(d)+1} > d^{1+\frac{\eta}{l(l+1+\eta)}}$.

Proof. Set $H_2 = \langle \mathcal{B}_2 \rangle_d$. Since $(A_2, d) = 1$, we have also $(\#H_2, d) = 1$. Hence $H_2 \le \bigoplus_{p|d} C_{p-1} \le \mathbb{Z}_d^*$. Therefore H_2 contains at most $\omega(d)$ linearly independent elements of order dividing $q^{v_q(A_2)}$ for each $q \mid A_2$. It follows that $A_2^{\omega(d)} \ge \#H_2 \ge \psi(d, \ln^{l+1+\eta} d) > d^{1-\frac{1}{l+1+\eta}}$. Thus $A_2^{\omega(d)+1} > d^{\frac{\omega(d)+1}{\omega(d)}(1-\frac{1}{l+1+\eta})} = d^{1+\frac{\eta+l-\omega(d)}{\omega(d)(l+1+\eta)}} \ge d^{1+\frac{\eta}{l(l+1+\eta)}}$. $\square$

If $d < \binom{l}{[l/2]}^{\frac{l(l+1+\eta)}{\eta}}$ we factor d by trial division. Suppose that the opposite inequality holds. By lemma 6.2 we have $A_2^{\omega(d)+1} > d^{1+\frac{\eta}{l(l+1+\eta)}} \ge d\binom{l}{[l/2]} \ge d\binom{\omega(d)}{[\omega(d)/2]}$. In the ensuing lemma (whose case $k = 3$ was considered in the article of Konyagin and Pomerance [10]) we show that the factorization of d can be found by factoring a suitable polynomial in $\mathbb{Z}[x]$. We use the Hensel-Berlekamp algorithm [9], [4] to this end.

Lemma 6.3 *Let* $d = p_1^{e_1} \cdot \ldots \cdot p_k^{e_k}$. *Assume* $A \in \mathbb{N}$ *divides* $p_i - 1$ *for* $i = 1, \ldots, k$; $p_i = b_i A + 1$. *Suppose in addition that* $A^{k+1} > \binom{k}{[k/2]} d$. *Write* d *in base* A: $d = 1 + a_1 A + \ldots + a_k A^k$. *Let* $g(x) = 1 + a_1 x + \ldots + a_k x^k$. *Then* $g(x) = (b_1 x + 1) \cdot \ldots \cdot (b_k x + 1)$ *in* $\mathbb{Z}[x]$. *Furthermore, this factorization can be obtained with the Hensel-Berlekamp algorithm in* $O(\ln^6 d \ln^2 \ln d)$ *deterministic time.*

Proof. We have $d = p_1^{e_1} \cdot \ldots \cdot p_k^{e_k} = (b_1 A + 1)^{e_1} \cdot \ldots \cdot (b_k A + 1)^{e_k}$. Since $A^{k+1} > d$, it follows that $e_1 = \ldots = e_k = 1$. Hence $1 + a_1 A + \ldots + a_k A^k = (b_1 A + 1) \cdot \ldots \cdot (b_k A + 1) = 1 + \sum_{j=1}^{k} \sigma_{k,j}(b_1, \ldots, b_k) A^j$, where $\sigma_{k,j}(b_1, \ldots, b_k) = \sum_{1 \le i_1 < \ldots < i_j \le k} b_{i_1} \cdot \ldots \cdot b_{i_j}$. It is therefore sufficient to show that $0 \le \sigma_{k,j}(b_1, \ldots, b_k) < A$ for every j, $1 \le j \le k$. By assumption, $A^{k+1} > \binom{k}{[k/2]} d$ and thus $b_1 \cdot \ldots \cdot b_k \binom{k}{[k/2]} d < b_1 \cdot \ldots \cdot b_k A^{k+1} < dA$. Hence $b_1 \cdot \ldots \cdot b_k \binom{k}{[k/2]} < A$ and it follows that $0 \le \sigma_{k,j}(b_1, \ldots, b_k) \le \binom{k}{j} b_1 \cdot \ldots \cdot b_k \le \binom{k}{[k/2]} b_1 \cdot \ldots \cdot b_k < A$.
It remains to prove that g can be completely factored in the stated time. We first need a 'small' prime p such that $\deg g_p = \deg g$ and g_p is squarefree, where g_p denotes the reduction of g modulo p. Let $D = a_k \prod_{1 \le i < j \le k} (b_i - b_j)$. It is enough to show that there is a 'small' prime p not dividing D. Since $a_k < A < d$ and $2b_i < p_i \le d$ for each i, we have $D < d^{1+\frac{k(k-1)}{2}} \le d^{k^2}$.

Moreover, $k = \omega(d) \le \ln d$ (d is odd), thus $D < e^{\ln^3 d}$. Set $P(y) = \prod_{p \le y} p$. Then, by Chebyshev's theorem, $P(y) \ge e^{cy}$ for some constant c. Take $y = \frac{1}{c}\ln^3 d$. Now suppose that $p \mid D$ for all $p \le y$. It follows that $P(y) \mid D$, in particular $P(y) \le D$. Hence $e^{cy} < e^{\ln^3 d}$ i.e. $y < \frac{1}{c}\ln^3 d$ - a contradiction. Consequently, there exists a prime $p \le y$ such that $p \nmid D$. Let $\alpha = a_k^{-1}(p)$, $e = \lceil \frac{\ln d}{\ln p} \rceil$. We factor completely αg_p with the Berlekamp algorithm in $O(k(k+p)(k \ln p)^2) = O(\ln^6 d \ln^2 \ln d)$ deterministic time (cf. theorem 7.4.5 of [3]). 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 $O(ke(k \ln p)^2) = O(\ln^4 d \ln^2 \ln d)$ deterministic time (cf. theorem 7.7.2 of [3]). Finally, we compute $b_i(p^e)$ for every i. This finishes the proof, as each b_i is less than p^e. □

It remains to treat the slightly more general case $F(n, M, 1) \ge \omega(n) - 2$. We then have $\#\{p \mid n : \; p - 1 \nmid M\} \le 2$. It is therefore easy to see that iterated use of the preceding algorithm will lead either to the complete factorization of n, or a factorization $n = mp^e q^f$, where $(m, pq) = 1$, the complete factorization of m is known, but p, q, e, f are not. In the latter case we can easily compute $\varphi(m)$ and $\varphi(p^e q^f) = \frac{\varphi(n)}{\varphi(m)}$. Finally we get p, q, e, f with the method described in section 5. Recall that, by lemma 5.3 (ii), we never had to call Isprime($\cdot$).

Theorem 6.4 *Let $l \in \mathbb{N}$, $\eta > 0$. Suppose that $\varphi(n)$ is given together with a divisor M of $\varphi(n)$ in a completely factored form, such that $F(n, M, 1) \ge \omega(n) - 2$ and $D(n, M, u) \le l$. Then the complete factorization of n can be found in $O_\eta(\ln^{5+\max(l+1+\eta, u+\frac{1}{l})} n)$ deterministic time.*

Proof. Such an algorithm has just been discussed. We now justify the running time analysis. Let $d > 1$ be a divisor of n. Define $\mathcal{B}_1$, $\mathcal{B}_2$, A_1 as in the above discussion. Computing $\mathrm{ord}_d(b)$ for every $b \in \mathcal{B}_2$ and checking whether $\mathcal{B}_2$ contains a Fermat-Euclid witness for d takes $O(\ln^{l+1+\eta} n \ln^4 n) = O(\ln^{l+5+\eta} n)$ time. Testing, in the case when $p_+(A_1) \le \ln^u n$, if $\langle \mathcal{B}_1 \rangle_d$ is cyclic takes $O(\ln^{1+\frac{1}{l}} n (\ln n + p_+(A_1)) \ln^3 n) = O(\ln^{4+u+\frac{1}{l}} n)$ time (see corollary 4.4). The time required to factor the polynomial from lemma 6.3 is $O(\ln^6 d \ln^2 \ln d)$, thus negligible (since $l \ge 1$). Therefore the described procedure leads to a nontrivial divisor or a proof of the primality of d in $O_\eta(\ln^{4+\max(l+1+\eta, u+\frac{1}{l})} n)$ time (we make the big-O constant depend upon η to include the time of the trial division step). It has to be iterated $O(\ln n)$ times, hence the stated running time bound. □

Remark 6.5 *Along the same lines, we see that for every Carmichael number n it can be decided whether $\omega(n) \leq l$, where $l \geq 3$ is some fixed integer, in deterministic polynomial time (depending upon l) given the complete factorization of $n-1$; in the case when $\omega(n) \leq l$ the complete factorization of n is obtained.*

6.2 $F(n, M, \frac{1}{3}+\varepsilon) = \omega(n)$, $D(n, M, u) \leq 1$

Choose $\eta > 0$ arbitrary. Assume that n is squarefree and has no prime divisor below $\ln^{\frac{1}{\varepsilon}+\eta} n$. Let $d > 1$ be a divisor of n. We show how to find a nontrivial factor or a proof of the primality of d in deterministic polynomial time (in $\ln n$). Set $M_p = (M, p-1)$, $p-1 = M_p N_p$ for all $p \mid d$. There is no loss of generality in assuming that $(M, N_p) = 1$ and consequently $(M_p, N_p) = 1$ for each $p \mid d$. Hence $\mathbb{Z}_d^* \simeq G_1 \oplus G_2$, where $G_1 = \bigoplus_{p \mid d} C_{M_p}$ and $G_2 = \bigoplus_{p \mid d} C_{N_p}$. Let $\mathcal{B} = \{2, 3, \ldots, [\ln^{\frac{1}{\varepsilon}+\eta} d]\}$. As we know $\varphi(n)$ and every prime factor of M we can compute $B(b) := \prod_{s \mid M} s^{v_s(\mathrm{ord}_d(b))}$ for all $b \in \mathcal{B}$. It is easy to see that $B(b) = \mathrm{ord}_{G_1}(b)$. We check whether there are $b \in \mathcal{B}$ and $s \mid B(b)$ such that $(b^{\frac{\varphi(n) s^{v_s(B(b))-1}}{s^{v_s(\varphi(n))}}} - 1, d) > 1$. We can suppose that there are no such, since otherwise we get a nontrivial divisor of d. Let $A = \mathrm{LCM}_{b \in \mathcal{B}}(B(b))$. Then $A \mid p-1$ for all $p \mid d$. The following analogue of corollary 4.4 for the group G_1 will be of use.

Theorem 6.6 *Let $\mathbb{Z}_d^* = G_1 \oplus G_2$, $(\#G_1, \#G_2) = 1$, $a, b \in \mathbb{Z}_d^*$, α be a multiple of $ord_d(a)$ and $ord_d(b)$. If none of the procedures $PH(d, a^{\frac{\alpha}{s^{v_s(\alpha)}}}, b^{\frac{\alpha}{s^{v_s(\alpha)}}}, s, v_s(ord_d(a)), v_s(ord_d(b)))$, where $s \in \mathbb{P}$ and $s \mid (ord_d(a), ord_d(b), \#G_1)$, finds a nontrivial divisor of d then $\langle a, b \rangle_{G_1}$ is cyclic.*

Proof. Similar to that of corollary 4.4. In particular, if the procedures PH() above fail to find a nontrivial divisor of d then $\langle a, b \rangle_{G_1}$ is generated by $\prod_{s \mid (\mathrm{ord}_d(a)\mathrm{ord}_d(b), \#G_1)} c_s$, where $c_s = b^{\frac{\alpha}{s^{v_s(\alpha)}}}$ if $v_s(\mathrm{ord}_d(a)) \leq v_s(\mathrm{ord}_d(b))$, else $c_s = a^{\frac{\alpha}{s^{v_s(\alpha)}}}$. $\square$

Assume that $p_+(A) \leq \ln^u n$. This inequality holds when d is not a prime power, since $D(n, M, u) \leq 1$. Then we can take $\alpha = \varphi(n)$ in the above theorem and test if $\langle\mathcal{B}\rangle_{G_1}$ is cyclic by using an inductive procedure as in section 4. We suppose that $\langle\mathcal{B}\rangle_{G_1}$ is cyclic, for in the contrary case we get a nontrivial divisor of d. If d is a prime power then G_1 is a fortiori cyclic. Therefore

$$A = \#\langle\mathcal{B}\rangle_{G_1} = \#\frac{\langle\mathcal{B}\rangle_d}{\langle\mathcal{B}\rangle_d \cap G_2} \geq \frac{\psi(d, \ln^{\frac{1}{\varepsilon}+\eta} d)}{\prod_{p|d} N_p} > \frac{d^{1-(\frac{1}{\varepsilon}+\eta)^{-1}}}{\prod_{p|d} p^{1-(\frac{1}{3}+\varepsilon)}} = d^{\frac{1}{3}+\frac{\eta\varepsilon^2}{1+\eta\varepsilon}}.$$

However $A \mid p-1$ for all $p \mid d$, thus $A < p_-(d)$. Hence $\omega(d) \leq 2$. If $d < 2^{\frac{1+\eta\varepsilon}{3\eta\varepsilon^2}}$ then we factor d by trial division, so assume that the reverse inequality holds. Therefore $A^3 > d^{1+\frac{3\eta\varepsilon^2}{1+\eta\varepsilon}} \geq 2d$. If $\omega(d) = 2$ then the factorization of d is obtained by factoring the polynomial from lemma 6.3. Otherwise, d is a prime power.

Theorem 6.7 *Let $0 < \varepsilon \leq \frac{2}{3}, \eta > 0$. Assume that n is squarefree. Suppose furthermore that $\varphi(n)$ is given together with a divisor M of $\varphi(n)$ in a completely factored form, such that $F(n, M, \frac{1}{3}+\varepsilon) = \omega(n)$ and $D(n, M, u) \leq 1$. Then the complete factorization of n can be found in $O_{\eta,\varepsilon}(\ln^{4+u+\frac{1}{\varepsilon}+\eta} n)$ deterministic time.*

Proof. A suitable algorithm has been described above. It remains to verify the running time analysis. Let $d > 1$ be a divisor of n. Recall the above definitions of G_1, $\mathcal{B}$, $B(b)$, A. Computing $B(b)$ for each $b \in \mathcal{B}$ and checking whether $(b^{\frac{\varphi(n) s^{v_s(B(b))-1}}{s^{v_s(\varphi(n))}}} - 1, d) > 1$ for some $b \in \mathcal{B}$, $s \mid B(b)$ requires $O(\ln^{\frac{1}{\varepsilon}+\eta} n \ln^4 n) = O(\ln^{4+\frac{1}{\varepsilon}+\eta} n)$ time. Testing, in the case when $p_+(A) \leq \ln^u n$, if $\langle\mathcal{B}\rangle_{G_1}$ is cyclic can be done in $O(\ln^{\frac{1}{\varepsilon}+\eta} n(\ln n + p_+(A)) \ln^3 n) = O(\ln^{3+u+\frac{1}{\varepsilon}+\eta} n)$ time. As $u \geq 1$, it follows that the previous procedure finds a nontrivial factor or a proof of the primality of d in $O_{\eta,\varepsilon}(\ln^{3+u+\frac{1}{\varepsilon}+\eta} n)$ time (the time of the trial division step has been included by making the big-O constant depend upon η and ε). This procedure must be iterated $O(\ln n)$ times, which yields the stated running time bound. $\square$

6.3 $F(n, M, \frac{1}{4}+\varepsilon) = \omega(n)$, $D(n, M, u) \leq 2$

Theorem 6.8 *Let $0 < \varepsilon \leq \frac{3}{4}, \eta > 0$. Suppose that $\varphi(n)$ is given together with a divisor M of $\varphi(n)$ in a completely factored form, such that $F(n, M, \frac{1}{4}+\varepsilon) = \omega(n)$ and $D(n, M, u) \leq 2$. Then there is an algorithm that calls a 'φ-oracle' $O(\ln n)$ times to output the complete factorization of n in $O_{\eta,\varepsilon}(\ln^{4+u+\frac{1}{\varepsilon}+\eta} n)$ deterministic time.*

Proof. We follow a similar inductive procedure to the one presented in subsection 6.2. But this time, if $d > 1$ is a divisor of n we first compute $\varphi(d)$ using the given 'φ-oracle'. Therefore we can assume that d is squarefree (since otherwise $(d, \varphi(d))$ is a nontrivial factor of d) and $\omega(d) \geq 3$ (for in the contrary case the factorization of d can be found with the elementary method from section 5). □

7 Factoring all n in deterministic subexponential time when a 'φ-oracle' is given

In this section we point out a simple deterministic subexponential time reduction of factoring to computing φ. Let n be any odd integer. We first show that n can be completely factored in deterministic subexponential time provided that $\varphi(n)$ and its complete factorization are known. At the end of the following discussion, some parameters (α, β, γ) will be optimally chosen from the interval $(0, 1)$. We shall abbreviate any expression of the form $\exp\Big((\ln x)^a (\ln\ln x)^{1-a}\Big)$ as $L(x, a)$. Assume that n has no prime factor below $L(n, 1-\alpha)$. Let $d > 1$ be a divisor of n. Our goal is to find a nontrivial divisor or a proof of the primality of d. Let $\mathcal{B} = \{2, 3, \ldots, [L(d, 1-\alpha)]\}$. We first check whether there is a Fermat-Euclid witness for d among the elements of $\mathcal{B}$. Suppose that this is not the case, for otherwise we are done. We compute $A = \mathrm{LCM}_{b\in\mathcal{B}}(\mathrm{ord}_d(b))$. Assume that $(1-\beta)(1-\gamma) \leq 1-\alpha$. We further test if d has a nontrivial factor of the form $mA + 1$, where $m < L(d, (1-\beta)\gamma)$.

Lemma 7.1 *Let $\alpha, \beta, \gamma \in (0, 1)$, $(1-\beta)(1-\gamma) \leq 1-\alpha$, $\mathcal{B} = \{2, 3, \ldots, [L(d, 1-\alpha)]\}$, $\mathcal{B} \subset \mathbb{Z}_d^*$, $A = LCM_{b\in\mathcal{B}}(ord_d(b))$. Suppose that $\omega(d) > \big(\frac{\ln d}{\ln\ln d}\big)^\beta$ and none of the elements of $\mathcal{B}$ is a Fermat-Euclid witness for d. Then $p_-(d) = mA + 1$ for some integer $m < L(d, (1-\beta)\gamma)$ if $p_-(d)$ is sufficiently large.*

Proof. Let $p = p_-(d)$. We have $L(p, 1-\gamma) \leq \exp\Big((\frac{1}{\omega(d)}\ln d)^{1-\gamma}(\ln\ln d)^\gamma\Big) < L(d, (1-\beta)(1-\gamma)) \leq L(d, 1-\alpha)$, 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 $A = \mathrm{LCM}_{b\in\mathcal{B}}(\mathrm{ord}_d(b)) = \mathrm{LCM}_{b\in\mathcal{B}}(\mathrm{ord}_p(b)) = \#\langle\mathcal{B}\rangle_p \geq \psi(p, L(p, 1-\gamma))$. By theorem 2.2, we obtain $A \geq pL(p, \gamma)^{-\gamma+o(1)}$. Suppose that p is sufficiently large, so that the term $-\gamma + o(1)$ in the last expression is bigger than -1. We can write $p = mA + 1$ for some $m \in \mathbb{N}$, since $A \mid p - 1$. Therefore $mA < p \leq AL(p, \gamma)$. Hence $m < L(p, \gamma) \leq \exp\Big((\frac{1}{\omega(d)}\ln d)^\gamma(\ln\ln d)^{1-\gamma}\Big) < L(d, (1-\beta)\gamma)$. □

If d has no such divisor then, from the above lemma, either d is divisible by a 'small' prime (less than some constant), or $\omega(d) \leq \left(\frac{\ln d}{\ln \ln d}\right)^{\beta}$. Assume the latter. Suppose moreover that $\beta \leq \frac{1}{2}$ and $1-\beta \geq \alpha$.

Lemma 7.2 *Let $0 < \alpha < 1$, $0 < \beta \leq \frac{1}{2}$, $1-\beta \geq \alpha$, $\mathcal{B} = \{2, 3, \ldots, [L(d, 1-\alpha)]\}$, $\mathcal{B} \subset \mathbb{Z}_d^*$, $A = LCM_{b \in \mathcal{B}}(ord_d(b))$. Assume that $\omega(d) \leq \left(\frac{\ln d}{\ln \ln d}\right)^{\beta}$ and $(A, d) = 1$. 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 6.2, we have $A^{\omega(d)} \geq \#\langle\mathcal{B}\rangle_d \geq \psi(d, L(d, 1-\alpha))$. Hence $A^{\omega(d)+1} \geq \psi(d, L(d, 1-\alpha))^{\frac{\omega(d)+1}{\omega(d)}}$. It follows from theorem 2.2 that $A^{\omega(d)+1} \geq d^{1+\frac{1}{\omega(d)}} L(d, \alpha)^{\frac{\omega(d)+1}{\omega(d)}(-\alpha+o(1))}$. Let $-1 < \varepsilon < -\alpha$; assume that d is large enough, so that the above term $(-\alpha + o(1))$ is greater than ε. 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

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

□

We finally obtain the complete factorization of d by factoring the polynomial from lemma 6.3. The running time bound of our inductive factorization 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 $\{(\alpha, \beta, \gamma) : 0 < \alpha < 1, 0 < \beta \leq \frac{1}{2}, 0 < \gamma < 1, 1-\beta \geq \alpha, (1-\beta)(1-\gamma) \leq 1-\alpha\}$. Some easy calculations show that the minimum is $\frac{1}{3}$, reached for $\alpha = \frac{2}{3}$, $\beta = \frac{1}{3}$, $\gamma = \frac{1}{2}$. We have thus proved

Theorem 7.3 *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.*

Remark 7.4 *The above algorithm reduces the factorization of Carmichael numbers n to the factorization of $n-1$ in $L(n, \frac{1}{3})^{1+o(1)}$ deterministic time.*

Corollary 7.5 *There is a deterministic algorithm that calls a 'φ-oracle' $O(\ln n)$ times to output the complete factorization of n in less than $L(n, \frac{1}{3})^{1+o(1)}$ time.*

Proof. We use the given 'φ-oracle' to compute the sequence $\varphi(n), \varphi^2(n), \ldots, \varphi^k(n)$, where $k \leq 1 + \log_2 n$ is the least integer such that $\varphi^k(n) = 1$. Let $1 \leq m \leq k$. If we have found the complete factorization of $\varphi^m(n)$ then, by theorem 7.3, we can obtain the complete factorization of $\varphi^{m-1}(n)$ in less than $L(n, \frac{1}{3})^{1+o(1)}$ deterministic time. The corollary follows by induction. □

References

[1] L. M. Adleman, K. S. McCurley, *Open problems in number theoretic complexity II*, Algorithmic Number Theory Symposium **I** (1994), 291-322.

[2] E. Bach, *Explicit bounds for primality testing and related problems*, Mathematics of Computation, **55** (1990), 355-380.

[3] E. Bach, J. O. Shallit, *Algorithmic Number Theory, Volume 1: Efficient Algorithms*, MIT Press, 1996.

[4] E. R. Berlekamp, *Factoring Polynomials Over Finite Fields*, Bell Systems Technical Journal, **46** (1967), 1853-1859.

[5] R. J. Burthe Jr., *The average least witness is* 2, Acta Arithmetica, **80** (1997), 327-341.

[6] E. R. Canfield, P. Erdös, C. Pomerance, *On a problem of Oppenheim concerning 'Factorisatio Numerorum'*, Journal of Number Theory, **17** (1983), 1-28.

[7] P. Erdös, C. Pomerance, *The normal number of prime factors of* $\varphi(n)$, Rocky Mountain Journal of Mathematics, **15** (1985), 343-352.

[8] M. R. Fellows, N. Koblitz, *Self-Witnessing Polynomial-Time Complexity and Prime Factorization*, Designs, Codes and Cryptography, **2** (1992), 231-235.

[9] K. Hensel, *Neue Grundlagen der Arithmetic*, Journal für die Reine und Angewandte Mathematik, **127** (1904), 51-84.

[10] S. Konyagin, C. Pomerance, *On Primes Recognizable in Deterministic Polynomial Time*, The Mathematics of Paul Erdos, R. L. Graham, J. Nesetril, eds., Springer-Verlag, 1997, 176-198.

[11] S. Landau, *Some Remarks on computing the Square Parts of Integers*, Information and Computation, **78**, No. 3 (1988), 246-253.

[12] H. W. Lenstra, Jr., C. Pomerance, *Primality Testing with Gaussian Periods*, preliminary version July 20, 2005.

[13] G. L. Miller, *Riemann's Hypothesis and Tests for Primality*, Journal of Computer and System Sciences, **13** (1976), 300-317.

[14] 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.

[15] J. M. Pollard, *Theorems on factorization and primality testing*, Proceedings of the Cambridge Philosophical Society, **76** (1974), 521-528.

[16] M. O. Rabin, *Probabilistic algorithm for testing primality*, Journal of Number Theory, **12** (1980), 128-138.