International Mathematical Forum, 1, 2006, no. 33, 1599-1605

Integer Factorization and Twin Primes Verification Algorithms

Zvi Retchkiman Konigsberg

Mineria 17-2, Col. Escandon, Mexico D.F 11800, Mexico
e-mail: mzvi@cic.ipn.mx

Abstract

In this paper two algorithms based on the the divisibility properties of binomial expressions are introduced. The mathematical foundation lies in the connection that exists between binomial expressions and the number of carries that result in the sum in different bases, of the variables that form the binomial expression. The first one decomposes an integer into its prime factorization, while the second one answers if given two twin integers whether they are twin primes or not by just verifying one condition. The proposed procedures have the inconvenience of requiring the knowledge of the primes up to some fix number however, by a slight modification of the first algorithm this can be overcome. The mathematical approach applied is novice, and both algorithms are new.

Mathematics Subject Classification: 11Y05, 11Y11, 11Y16

Keywords: Prime Numbers, Twin Prime Numbers, Binomial Coefficients, Algorithms, Factorization, Verification

1 Introduction

There are many algorithms for integer factorization however, they are not efficient (deterministic or random polynomial-time) in fact they are computationally intractable. Most of the factoring algorithms fall into one of the two following classes: The running time depending mainly on the size of N the number to be factored or on the size of p the factor of N. However its running time is at best sub-exponential (Polynomial $\subset$ Super-polynomial $\subset$ Sub-exponential $\subset$ Exponential) therefore why should we care in studying them. The main reasons (to mention some) are: they can be of some aid for further algorithm developments and for the mathematics involved in their creation which can also lead to new ideas. In this paper two algorithms based on

the the divisibility properties of binomial expressions are presented. The first one decomposes an integer into its prime factorization while the second one answers if given two twin integers whether they are twin primes or not by just verifying one condition. Their mathematical justification results from the work done by Kummer in 1852 [1] in relation to the connection that exists between binomial expressions and the number of carries that result in the sum in different bases, of the variables that form the binomial expression. The necessary and sufficient conditions provided for twin prime verification were inspired in the work presented in [2]. The proposed procedures have the inconvenience of requiring the knowledge of the primes up to some fix number however by a slight modification of the first algorithm this can be overcome. The mathematical approach applied is novice, and both algorithms are new. The paper is organized as follows. Section 1, gives the mathematical preliminaries needed to understand the rest of the paper. Section 2, deals with the prime factorization algorithm while section 3, with the twin prime one. Finally, some concluding remarks are given.

2 Preliminaries

Definition 2.1 *Let n and p be integers, the $p-$adic expansion of n (which is the representation of n in base p) is given by,*

$$n = a_0 + a_1 p + a_2 p^2 + ... + a_m p^m \tag{1}$$

where the digits $a_i \in \{o...., p-1\}$ and m is an integer. Alternatively n is said to have the p.adic expansion,

$$n = (a_m a_{m-1} \cdots a_1 a_0)_p \tag{2}$$

Definition 2.2 *Let n and k be integers, the $p-$adic addition of n and k consists in, the addition of its respective $p-$adic representations in base p. The number of carries in the $p-$adic addition of n and k will be denoted by $\tau = c_p(n,k)$.*

The main result of this section, which is next stated, is due to Kummer 1852 [1] , (for completeness purposes the proof is supplied).

Theorem 2.3 *Let $\tau = c_p(n,k)$ be the number of carries in the $p-$adic addition of n and k then, $\begin{pmatrix} n+k \\ k \end{pmatrix}$ is divisible by the prime power p^τ but not by $p^{\tau+1}$.*

In order to derive this beautiful theorem the following result, called Legendre's formula (1808), is used.

Lemma 2.4 *Let $\mu(n)$ be the largest exponent of the prime power $p^{\mu(n)}$ which divides $n!$ then,*

$$\mu(n) = \frac{n-\sigma}{p-1} \tag{3}$$

where σ is the sum of the $p-$adic coefficients of $a_i \in \{0....,p-1\}$ of n.

Proof 2.5 *From the identity $\mu(n) = \sum_{i=1}^{\infty} \left\lfloor \frac{n}{pi} \right\rfloor$ Legendre's formula is equivalent to $\sum_{i=1}^{\infty} \left\lfloor \frac{n}{pi} \right\rfloor (p-1) = n - \sigma$ which is next established. Using the $p-$adic representation of n and the definition of the floor function $\left\lfloor \frac{n}{pi} \right\rfloor = a_i + a_{i+1}p + a_2p^2 + ... + a_mp^{m-i}, i \leq m$. Next, computing the two sums one gets*

$$\begin{aligned}
\sum_{i=1}^{\infty} \left\lfloor \frac{n}{pi} \right\rfloor (p) &= \sum_{i=1}^{\infty} (a_ip + a_{i+1}p^2 + ... + a_mp^{(m-i+1)}) \\
&= a_1p + a_2p^2 + a_3p^3 + \cdots + a_mp^m \\
&\quad + a_2p + a_3p^2 + \cdots + a_mp^{m-1} \\
&\quad + a_3p + \ldots + a_mp^{m-2} \\
&\quad \vdots \\
&\quad + a_mp
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor &= \sum_{i=1}^{\infty} (a_i + a_{i+1}p + ... + a_mp^{(m-i)}) \\
&= a_1 + a_2p + a_3p^2 + \cdots + a_mp^{m-1} \\
&\quad + a_2 + a_3p + \cdots + a_mp^{m-2} \\
&\quad + a_3 + \ldots + a_mp^{m-3} \\
&\quad \vdots \\
&\quad + a_m.
\end{aligned}$$

Finally performing the difference between the two sums $\sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor (p-1) = (a_1p + a_2p^2 + \cdots + a_mp^m) - (a_1 + a_2 + \cdots + a_m) = (n - a_0) - (\sigma - a_0) = n - \sigma$, which proves the formula.

Next, the theorem is proved.

Proof 2.6 *Let the $p-$adic expansion of n and k be $n = a_0 + a_1p + a_2p^2 + ... + a_mp^m$, and $k = b_0 + b_1p + b_2p^2 + ... + b_mp^m$ where $a_i, b_i \in \{o....,p-1\}$. Now if p^τ is the largest prime power which divides $\begin{pmatrix} n+k \\ k \end{pmatrix}$ then $\nu = \mu(n+k) - \mu(n) - \mu(k)$. Therefore, it remains to prove that the following identity holds*

$$c_p(n,k) = \mu(n+k) - \mu(n) - \mu(k). \tag{4}$$

Carrying out the $p-$adic addition of n and k produces carries $\epsilon_0, \epsilon_1, ...$ (obtained from $\epsilon_0 = \left\lfloor \frac{a_0+b_0}{p} \right\rfloor$ and $\epsilon_i = \left\lfloor \frac{a_i+b_i+\epsilon_{i-1}}{p} \right\rfloor, i = 1, 2, ...$), therefore, the

sum of carries takes the form $c_p(n,k)=\sum_{i=0}^{\infty}\epsilon_i$. *On the other hand, the* $p-$*adic representation of the sum* $n+k$ *can be expressed as* $n+k=\sum_{i=0}^{\infty}c_ip^i$ *where* $c_i\in\{0....,p-1\}$. *Moreover, the* c_i *digits of this addition in terms of those of* n *and* k *and the carries* ϵ_i *is given by the formula* $c_i=a_i+b_i+\epsilon_{i-1}-\epsilon_ip, \epsilon-1=0$ *for* $i=0,1,...$
Finally, employing this last formula and Legendre's identity, 4 is shown to be true as can be seen in the next computation,

$\nu=\mu(n+k)-\mu(n)-\mu(k)=\frac{n+k-\sum_{i=0}^{\infty}c_i}{p-1}-\frac{n-\sum_{i=0}^{\infty}a_i}{p-1}-\frac{k-\sum_{i=0}^{\infty}b_i}{p-1}=$
$\frac{1}{p-1}\left(\sum_{i=0}^{\infty}a_i+\sum_{i=0}^{\infty}b_i-\sum(a_i+b_i+\epsilon_{i-1}-\epsilon_ip)\right)=\frac{1}{p-1}\left(\sum_{i=0}^{\infty}\epsilon_ip-\epsilon_{i-1}\right)=$
$\frac{1}{p-1}\left(\sum_{i=0}^{\infty}\epsilon_i(p-1)\right)=\sum_{i=0}^{\infty}\epsilon_i=c_p(n,k)$. *Therefore, theorem 2.3 is established.*

Corollary 2.7 *Let* n *be any integer and* p *a prime number* $(\leq n)$. *Set* $\tau=c_p(n-1,1)$ *then,* n *is divided by the prime power* p^{τ} *but not by* $p^{\tau+1}$.

The next result is related to twin prime numbers.

Theorem 2.8 *Let* $2n-1$ *and* $2n+1$ *be any two numbers with* $n\geq 3$ *and consider the set* $\Pi_T=\{p:p \text{ is a prime such that } 3\leq p\leq\sqrt{2n+1}\}$. *Then the pair* $(2n-1,2n+1)$ *is a twin prime pair if and only if* $\forall p\,\varepsilon\,\Pi_T, p\mid\begin{pmatrix}n+\frac{p-3}{2}\\ p-2\end{pmatrix}$. *Moreover if* p *does not divide* $\begin{pmatrix}n+\frac{p-3}{2}\\ p-2\end{pmatrix}$ *then* p *divides either* $2n-1$ *or* $2n+1$.

Proof 2.9 *First, assume that the pair* $(2n-1,2n+1)$ *is a twin prime pair we must show that* $\forall p\varepsilon\Pi, p\mid\begin{pmatrix}n+\frac{p-3}{2}\\ p-2\end{pmatrix}$. *The following result borrowed from [2] is used in order to prove the claim.*

Lemma 2.10 *The expression* $\begin{pmatrix}n+\frac{p-3}{2}\\ p-2\end{pmatrix}\left[\frac{(2n-1)(2n+1)}{p}\right]$ *is an integer whenever* $n\geq 1$ *and* $(0\leq\frac{p-3}{2}\leq n-1)$.

Since $(2n-1),(2n+1)$ *and* p *are primes (*p *and* n *satisfying the bound condition)* $\frac{(2n-1)(2n+1)}{p}$ *is a rational number therefore for the whole expression to become an integer it is needed that,* $p\mid\begin{pmatrix}n+\frac{p-3}{2}\\ p-2\end{pmatrix}$.
Now let us prove the converse. Assume that the pair $(2n-1,2n+1)$ *is not a twin prime. If* $p\mid(2n-1)\Rightarrow 2n-1=p(2m+1)$ *with* $m\geq 1$. *Then* $n+\frac{p-3}{2}=p-1+mp$ *and the binomial coefficient becomes*

$$\frac{(p-1+mp)(p-2+mp)\cdots(2+mp)}{(p-2)!}.$$

In case $p \mid (2n+1)$ *we get that*

$$\frac{(p-2+mp)(p-1+mp)\cdots(1+mp)}{(p-2)!}.$$

In either case p *does not divide* $\begin{pmatrix} n+\frac{p-3}{2} \\ p-2 \end{pmatrix}$.

Finally, if p *does not divide* $\begin{pmatrix} n+\frac{p-3}{2} \\ p-2 \end{pmatrix}$ *since* $\begin{pmatrix} n+\frac{p-3}{2} \\ p-2 \end{pmatrix}\left[\frac{(2n-1)(2n+1)}{p}\right]$ *is an integer,* p *must divide either* $2n-1$ *or* $2n+1$.

3 Integer Factorization

This section provides an algorithm for integer factorization whose proof follows directly from corollary 2.7. The time taken to compute the factorization for several sizes of n and different computer platform systems is depicted in two tables.

Algorithm

Step 1. Enter n.

Step 2. Do $\forall p \varepsilon\, \Pi_F = \{p : p \text{ is a prime} \leq \sqrt{n}\}$ (where a numeration of the elements of Π_F will be denoted by $\{q_{k_i}\}_{i=1}^{\#(\Pi_F)}$) the q_{k_i}adic addition of $n-1$ and 1, compute the number of carries $c_{q_{k_i}}$. Set $k = q_{k_1}^{c_{q_{k_1}}} \cdot q_{k_2}^{c_{q_{k_2}}} \cdot \ldots \cdot q_{k_{\sqrt{n}}}^{c_{q_{k_{\#(\Pi_F)}}}}$.

Step 3. If $k = 1$ then n is prime otherwise, n is composite and its prime factorization is given by $n = kq$ with $q = \frac{n}{k}$.

Remark 3.1 *Notice that the procedure has the inconvenience of requiring the knowledge of the set* Π_F *however, the same algorithm (with a slight modification) can be used to generate it.*

Assuming that n is such that all additions up to $\pi(\sqrt{n})$ have to be computed, (where $\pi(x)$ is the prime counting function), the next two tables summarize the time taken by the algorithm to make the factorization. The first one provides a comparison between a system running under a Pentium 4, 2GHz processor and the Bluegene, 700MHz (the most powerful system available today) while the second one makes a comparison between different number cf computers connected in parallel.

n	*Pentium 4*	*Bluegene*
10^{30}	11.93 min	.5 min
10^{38}	65 days	3.5 days
10^{40}	153 years	35 days

Table 1

n	*1 Serial computer*	*15 Parallel μp*	*50 Parallel μp*	*100 Parallel μp*
10^{16}	1 min	4 sec	1.2sec	.6sec
10^{20}	1.26hr	6min	1.8min	54sec
10^{30}	9 years	219 days	66 days	33 days

Table 2

Remark 3.2 *As can be seen the time consumed grows exponentially in terms of the size of n and it becomes quite large for relatively small n's.*

4 Twin Primes

In this section an algorithm for twin primes based on theorems 2.3 and 2.8 is presented.

Algorithm

Step 1. Enter n.
Step 2. Do $\forall p\varepsilon\, \Pi_T = \{p : p \text{ is a prime such that } 3 \leq p \leq \sqrt{2n+1}\}$ (where a numeration of the elements of Π_T will be denoted by $\{q_{k_i}\}_{i=1}^{\#(\Pi_T)}$) the q_{k_i}adic addition of $n + \frac{p-3}{2} - (p-2)$ and $p-2$, compute the number of carries $c_{q_{k_i}}$.
Step 3. If the number of carries is such that $c_{q_{k_i}} \geq 1\ \forall p\varepsilon\, \Pi_F$ then the pair of integers $(2n-1, 2n+1)$ is a twin prime pair otherwise is not and p divides either $2n-1$ or $2n+1$.

Remark 4.1 *Notice that the procedure has the inconvenience of requiring the knowledge of the set Π_T however, the algorithm proposed in section 3 (with a slight modification) can be used to generate it.*

Example 4.2

- *1. Take $n = 12 \Rightarrow \Pi_T = \{3, 5\}$. Next computing the carries of the 3 and 5 addition of (11,1) and (10,3) we obtain carry in the first one but no carry for the second one therefore the pair is not a twin prime pair. Even more $5 \mid (2n+1 = 25)$*
- *2. Take $n = 15 \Rightarrow \Pi_T = \{3, 5\}$. Next computing the carries of the 3 and 5 addition of (14,1) and (13,3) we obtain carries for both of them from where we conclude that the twin pair is a twin prime pair.*

References

[1] E.E. Kummer, Über Erngänzungssätze den allgemeinen Reziprozitätsgetzen, *J*ornal für die reine und angewandte Mathematik, 44:93–146, 1852.

[2] K. Dilcher and K. Stolarsky, A Pascal-Type Triangle Charactrizing Twin Primes, (*T*he Mathematical Association of America Monthly), 673:681, 2005.

Received: January 25, 2006