Malaya J. Mat. 2(3)(2014) 215–221

Malaya Journal of Matematik

MJM
an international journal of mathematical sciences with computer applications...

www.malayajournal.org

ISSN : 2319-3786

Conversion of number systems and factorization

Janusz Wlodarczyk,[a] Djilali Behloul,[b] and Sui Sun Cheng[c]*

[a]*Space Research Centre, Polish Academy of Sciences, Bartycka 18A, Warsaw, 00-716, Poland.*

[b]*Department of Computer Science, USTHB, BP32 El Alia, Bab Ezzouar, 16111 Algiers, Algeria.*

[c]*Department of Mathematics, Tsing Hua University, Hsinchu, Taiwan 30043, R. O. China.*

Abstract

In this paper one can see a new method for conversion of number systems. As an application we give an algorithm of factorization of an integer n with arithmetic complexity $O(\sqrt{n}\ln^2 n)$.

Keywords: conversion, number systems, factorization.

2010 MSC: 11Y05 11Y16.

1 Introduction

Let us first start with a whole number n described in the number system with base $\mathbf{p}$:

$$n = a_{m,\mathbf{p}} \cdots {}_{1,\mathbf{p}}\, a_{0,\mathbf{p}} = a_{m,\mathbf{p}}\mathbf{p}^m + \cdots + a_{1,\mathbf{p}}\mathbf{p} + a_{0,\mathbf{p}} \tag{1.1}$$

where $a_{i,\mathbf{p}}$ is the digit at position i. If the least significant number $a_{0,\mathbf{p}} = 0$, then $\mathbf{p}$ is a divisor of n.

In this note, we are interested in converting a number in the number system with base $\mathbf{p}$ to that with base $\mathbf{p}+2$. As a consequence, we are able to design an algorithm for factorizing a number n with arithmetic complexity $O(\sqrt{n}\ln^2 n)$. Here we use arithmetic complexity models, where cost is measured by the number of machine instructions performed on a single processor with addition and substraction of m-bit integers that costs $O(m)$ (see [1]).

2 Conversion

The conversion of a number n in the $\mathbf{p}$ base number system to the $\mathbf{p}+2$ base number system uses Horner's scheme 'illustrated' as follows :

$$\begin{aligned}
n &= \cdots + (a\mathbf{p}+b)\mathbf{p}+c \\
&= \cdots + (a(\mathbf{p}{+}2) + (-2a+b))\mathbf{p}+c \\
&= \cdots + (a(\mathbf{p}{+}2)\mathbf{p} + (-2a+b)\mathbf{p}) + c \\
&= \cdots + (a(\mathbf{p}{+}2)(\mathbf{p}{+}2) - 2a(\mathbf{p}{+}2) + (-2a+b)(\mathbf{p}{+}2) + (-2a+b)(-2)) + c \\
&= \cdots + (a(\mathbf{p}{+}2) - 2a + (-2a+b))(\mathbf{p}{+}2) + (-2a+b)(-2) + c.
\end{aligned}$$

Note that the conversion only employs additions and/or substractions. This idea of conversion is first announced by Walter Soden (see [2, p. 320]), but expressed in special cases. Knuth [2] also mentions this idea for numbers, not for digits.

*Corresponding author.
E-mail addresses: u_mnicha@inetia.pl (Janusz Wlodarczyk), dbehloul@yahoo.fr (Djilali Behloul), suisuncheng@gmail.com (Sui Sun Cheng).

Lemma 2.1. *Let $n = a_{m,\mathbf{p}}\mathbf{p}^m + \cdots + a_{1,\mathbf{p}}\mathbf{p} + a_{0,\mathbf{p}}$ be an integer written in base $\mathbf{p}$. If $a_{m,\mathbf{p}} \neq 0$, then $m = \left[\log_{\mathbf{p}} n\right]$.*

Proof. All the digits $a_{i,\mathbf{p}}$, except $a_{m,\mathbf{p}}$, are between 0 and $\mathbf{p}-1$. Hence

$$\mathbf{p}^m \leq n \leq (\mathbf{p}-1)\cdot\mathbf{p}^m + \cdots + (\mathbf{p}-1)\cdot\mathbf{p} + (\mathbf{p}-1).$$

But

$$(\mathbf{p}-1)\cdot\mathbf{p}^m + \cdots + (\mathbf{p}-1)\cdot\mathbf{p} + (\mathbf{p}-1) = (\mathbf{p}^{m+1}-1).$$

Hence

$$\mathbf{p}^m \leq n \leq \mathbf{p}^{m+1} - 1 < \mathbf{p}^{m+1}.$$

Taking $\log_{\mathbf{p}}$ on both sides, we see that

$$m \leq \log_{\mathbf{p}} n < m+1,$$

which implies $m = \left[\log_{\mathbf{p}} n\right]$. □

Lemma 2.2. *Let $n = a\mathbf{p} + b$ be an integer written in base $\mathbf{p}$. Then n can be written in base $\mathbf{p}$+2 as $n = a'(\mathbf{p}+2) + b'$, where*

1) if $b - 2a \geq 0$ then $a' = a$ and $b' = b - 2a$,

2) if $b - 2a < 0$ and $b - 2a + (\mathbf{p}+2) \geq 0$ then $a' = a - 1$ and $b' = b - 2a + \mathbf{p} + 2$, and

3) if $b - 2a + \mathbf{p} + 2 < 0$, then $a' = a - 2$ and $b' = b - 2a + \mathbf{p} + 2 + \mathbf{p} + 2$.

The arithmetic complexity is at most $O(\log_2 \mathbf{p})$.

Proof. We have,

$$\begin{aligned} n &= a(\mathbf{p}+2-2) + b \\ &= a(\mathbf{p}+2) + (b-2a). \end{aligned}$$

It is easy to see that $b - 2a \leq b < \mathbf{p} + 2$.

First case: if $b - 2a \geq 0$ then $a' = a$ and $b' = b - 2a$.

Second case: if $b - 2a < 0$ and $b - 2a + (\mathbf{p}+2) \geq 0$ then we substract 1 from the digit a and we add $(\mathbf{p}+2)$ to the number $(b-2a)$,

$$n = (a-1)(\mathbf{p}+2) + (b - 2a + \mathbf{p} + 2).$$

Then $a' = a - 1$ and $b' = b - 2a + \mathbf{p} + 2$

Third case: if $b - 2a + \mathbf{p} + 2 < 0$ then we substract 1 from the digit $a - 1$ and we add $(\mathbf{p}+2)$ to the number $(b - 2a + \mathbf{p} + 2)$, we obtain

$$n = (a-2)(\mathbf{p}+2) + (b - 2a + \mathbf{p} + 2 + \mathbf{p} + 2).$$

It is easy to see that $(b - 2a + \mathbf{p} + 2 + \mathbf{p} + 2) \geq 0$. Then $a' = a - 2$ and $b' = b - 2a + \mathbf{p} + 2 + \mathbf{p} + 2$.

It is easy to see that the number of additions or substractions manipulating the numbers $1, 2, a, b$ and $\mathbf{p}$ is 9. We have 5 additions, 4 substractions where we consider $b - 2a$ as $b - a - a$. The lengths of a, b and $\mathbf{p}$ are $\leq \log_2 \mathbf{p}$, the complexity (i.e., the number of binary arithmetic operations) is then $9\log_2 \mathbf{p} \in O(\log_2 \mathbf{p})$. □

3 Transformation I

Let $n = (a(\mathbf{p}+2) + b)\mathbf{p} + c$, where $0 \leq a, b < \mathbf{p} + 2$ and $0 \leq c < \mathbf{p}$. Then n can be written in base $\mathbf{p}$+2 as

$$n = a'(\mathbf{p}+2)^2 + b'(\mathbf{p}+2) + c'.$$

The transformation will be achieved in two steps: transform step and correction step.

Transform step: Write

$$\begin{aligned} n &= (a(\mathbf{p}+2) + b)\mathbf{p} + c \\ &= (a(\mathbf{p}+2) + b)(\mathbf{p}+2-2) + c \\ &= a(\mathbf{p}+2)^2 + (b-2a)(\mathbf{p}+2) + c - 2b \\ &= A(\mathbf{p}+2)^2 + B(\mathbf{p}+2) + C \end{aligned}$$

where $A = a, B = b - 2a$ and $C = c - 2b$.

Correction step:

1) If $C \geq 0$ then $c' = C$.

2) If $C < 0$ and $C + \mathbf{p} + 2 \geq 0$ then we substract 1 from B and we add $\mathbf{p} + 2$ to C. Then $c' = C + \mathbf{p} + 2$.

3) If $C + \mathbf{p} + 2 < 0$ then we substract 1 from $B - 1$ and we add $\mathbf{p} + 2$ to $C + \mathbf{p} + 2$. Then $c' = C + \mathbf{p} + 2 + \mathbf{p} + 2$.

Now we assume that C is corrected. Then

$$n = A(\mathbf{p}+2)^2 + \tilde{B}(\mathbf{p}+2) + c'$$

where $\tilde{B} = B$ or $B - 1$ or $B - 2$.

1) If $\tilde{B} \geq 0$ then $b' = \tilde{B}$ and $a' = A$.

2) If $\tilde{B} < 0$ and $\tilde{B} + \mathbf{p} + 2 \geq 0$, we substract 1 from A and we add $\mathbf{p} + 2$ to $\tilde{B}$, then $b' = \tilde{B} + \mathbf{p} + 2$ and $a' = A - 1$.

3) If $\tilde{B} + \mathbf{p} + 2 < 0$, we substract 1 from $A - 1$ and we add $\mathbf{p} + 2$ to $\tilde{B} + \mathbf{p} + 2$, then $b' = \tilde{B} + \mathbf{p} + 2 + \mathbf{p} + 2$ and $a' = A - 2$.

It is easy to see that the number of additions or substractions involving $1, 2, \tilde{a}, \tilde{b}, \mathbf{p}$ and c is 16. We have 8 additions and 8 substractions. The lengths of a, b, c and $\mathbf{p}$ are $\leq \log_2(\mathbf{p}+1)$. Evaluation of a', b', c' involves $16\log_2(\mathbf{p}+1)$ binary arithmetic operations.

4 Transformation II

Let

$$\Delta_k = (...((a_k(\mathbf{p}+2) + a_{k-1})(\mathbf{p}+2) + a_{k-2})(\mathbf{p}+2) + \cdots + a_1)\mathbf{p} + a_0$$

where $0 \leq a_i < \mathbf{p} + 2$ for $i = 1, 2..., k$ and $0 \leq a_0 < \mathbf{p}$. Then Δ_k can be written in base $\mathbf{p}$+2 in the form

$$\Delta_k = a'_k(\mathbf{p}+2)^k + a'_{k-1}(\mathbf{p}+2)^{k-1} + \cdots + a'_1(\mathbf{p}+2) + a'_0.$$

Again, this can be done in two steps : transform step and correction step.

Transform step: Write

$$\begin{aligned}
\Delta_k &= (a_k(\mathbf{p}+2)^{k-1} + a_{k-1}(\mathbf{p}+2)^{k-2} + \cdots + a_2(\mathbf{p}+2) + a_1)\mathbf{p} + a_0 \\
&= (a_k(\mathbf{p}+2)^{k-1} + a_{k-1}(\mathbf{p}+2)^{k-2} + \cdots + a_2(\mathbf{p}+2) + a_1)(\mathbf{p}+2-2) + a_0 \\
&= a_k(\mathbf{p}+2)^k + (a_{k-1} - 2a_k)(\mathbf{p}+2)^{k-1} + \cdots + (a_1 - 2a_2)(\mathbf{p}+2) + a_0 - 2a_1.
\end{aligned}$$

Put $A_k = a_k$ and $A_{i-1} = a_{i-1} - 2a_i$ for $i = 1, ..., k$, then

$$\Delta_k = A_k(\mathbf{p}+2)^k + A_{k-1}(\mathbf{p}+2)^{k-1} + \cdots + A_1(\mathbf{p}+2) + A_0$$

Correction step:

1) If $A_0 \geq 0$ then $a'_0 = C$.

2) If $A_0 < 0$ and $A_0 + \mathbf{p} + 2 \geq 0$, we substract 1 to A_1 and we add $\mathbf{p} + 2$ to A_0 then $a'_0 = A_0 + \mathbf{p} + 2$

3) If $A_0 + \mathbf{p} + 2 < 0$, we substract 1 to $A_1 - 1$ and we add $\mathbf{p} + 2$ to $A_0 + \mathbf{p} + 2$ then $a'_0 = A_0 + \mathbf{p} + 2 + \mathbf{p} + 2$.

Now we assume that A_i is corrected, inductively, we will correct A_{i+1}:

$$\Delta_k = A_k(\mathbf{p}+2)^k + \cdots + A_{i+2}(\mathbf{p}+2)^{i+2} + \tilde{A}_{i+1}(\mathbf{p}+2)^{i+1} + a'_i(\mathbf{p}+2)^i + \cdots + a'_0$$

where $\tilde{A}_{i+1} = A_{i+1}$ or $A_{i+1} - 1$ or $A_{i+1} - 2$.

1) If $\tilde{A}_{i+1} \geq 0$ then $a'_{i+1} = \tilde{A}_{i+1}$.

2) If $\tilde{A}_{i+1} < 0$ and $\tilde{A}_{i+1} + \mathbf{p} + 2 \geq 0$, we substract 1 from A_{i+2} and we add $\mathbf{p} + 2$ to $\tilde{A}_{i+1}$, then $a'_{i+1} = \tilde{A}_{i+1} + \mathbf{p} + 2$.

3) If $\tilde{A}_{i+1} + \mathbf{p} + 2 < 0$, we substract 1 from $A_{i+2} - 1$ and we add $\mathbf{p} + 2$ to $\tilde{A}_{i+1} + \mathbf{p} + 2$, then $a'_{i+1} = \tilde{A}_{i+1} + \mathbf{p} + 2 + \mathbf{p} + 2$.

Number of operations :

1) The transform step needs $2k$ substractions

2) Correction step needs at most 4 additions, 2 substractions to correct A_0; 4 additions, 2 substractions to correct $\tilde{A}_1$; 4 additions and 2 substractions to correct $\tilde{A}_{k-1}$.

In total we need $2k + 6k = 8k$ operations.

The lengths of a_i and $\mathbf{p}$ are $\leq \log_2(\mathbf{p}+1)$. Evaluation of a'_i, $i = 0,..,k$, involves at most $8k\log_2(\mathbf{p}+1)$ binary arithmetic operations.

Theorem 4.1. *Let $n = a_{m,\mathbf{p}}\mathbf{p}^m + \cdots + a_{1,\mathbf{p}}\mathbf{p} + a_{0,\mathbf{p}}$ be an integer written in base $\mathbf{p}$. Then we can write n in the base $\mathbf{p}+2$ in a systematic manner, as $n = a'_{m',\mathbf{p}+2}(\mathbf{p}+2)^{m'} + \cdots + a'_{1,\mathbf{p}+2}(\mathbf{p}+2) + a'_{0,\mathbf{p}+2}$ where $m' = \left[\log_{\mathbf{p}+2} n\right]$. Furthermore, the arithmetic complexity is at most $O\left(\frac{\log_2^2 n}{\log_2 \mathbf{p}}\right)$.*

Proof. The numbers $a'_{i,\mathbf{p}+2}$ are determined by the Lemma 2.2, Transforms I and II described above (and is implemented in the conversion algorithm below). The total number $T(n,\mathbf{p})$ of operations is given by:

$$\begin{aligned} T(n,\mathbf{p}) &= 9\log_2 \mathbf{p} + 16\log_2(\mathbf{p}+1) + ... + 8k\log_2(\mathbf{p}+1) + ... + 8m\log_2(\mathbf{p}+1) \\ &= 9\log_2 \mathbf{p} + 8(2+3+...+m)\log_2(\mathbf{p}+1) \\ &= 9\log_2 \mathbf{p} + 8\log_2(\mathbf{p}+1)\left(\frac{m(m+1)}{2} - 1\right) \\ &= O(m^2 \log_2 \mathbf{p}). \end{aligned}$$

But $m = \left[\log_{\mathbf{p}} n\right]$ and $\log_{\mathbf{p}}^2 n \log_2 \mathbf{p} = \dfrac{\log_2^2 n}{\log_2 \mathbf{p}}$, hence the complexity is $O(\log_2^2 n / \log_2 \mathbf{p})$. □

We may now summarize our previous discussions by means of the following

Conversion Algorithm:

INPUT: number $n = a_{m,\mathbf{p}} \dots a_{1,\mathbf{p}} a_{0,\mathbf{p}}$ in the number system with base $\mathbf{p}$.

OUTPUT: number n in the number system with base $\mathbf{p}+2$ expressed in the form $n = a_{m,\mathbf{p}+2} \dots a_{1,\mathbf{p}+2} a_{0,\mathbf{p}+2}$.

for $i = 0$ to m step 1 $\{a_i \leftarrow a_{i,\mathbf{p}}\}$ end for;

for $k = m-1$ to 0 step -1

  borrow index $b \leftarrow 0$

  for $i = k$ to m step 1 $\{a_{m+1} \leftarrow 0; a_i \leftarrow a_i - 2a_{i+1} - b; b \leftarrow 0;\}$

    if ($a_i < 0$) { $b \leftarrow b+1; a_i \leftarrow a_i + \mathbf{p} + 2;$ }

    if ($a_i < 0$) { $b \leftarrow b+1; a_i \leftarrow a_i + \mathbf{p} + 2;$ }

  end for

end for

$m \leftarrow \left[\log_{\mathbf{p}+2} n\right]$ which is the actual number of digits n;

for $i = 0$ to m step 1 $\{a_{i,\mathbf{p}+2} \leftarrow a_i\}$ end for;

We remark that at the end of our algorithm, we correct the length of our number and the output number often has less digits. Thanks to lemma 2.1, the number of digits is related to the roots of the number n : when the current base $\mathbf{p}$ is greater than $\sqrt[k]{n}$, then $\ln \mathbf{p} > \ln \sqrt[k]{n}$ which implies $\log_{\mathbf{p}} n < k$, we have only k digits at most.

There are now numerous conversion algorithms, the present one has one interesting consequence.

5 Factorization

Factorization Algorithm:
INPUT: positive number $n = a_{m,2} \cdots a_{1,2}a_{0,2}$.
OUTPUT: table of divisors of the number n.

1. while ($a_{0,2} = 0$) { delete the least significant digit; table.insert(2) } end while;
2. $n \leftarrow$ the actual number n in which the least significant digit > 0;
3. conversion n into tertiary number $n = a_{m,3} \ldots a_{1,3}a_{0,3}$ the base of system $\mathbf{p} \leftarrow 3$;
4. while ($a_{0,3} = 0$) { delete the least significant digit; table.insert(3) } end while;
5. $n \leftarrow$ the actual number n in which the least significant digit > 0;
6. **while** number of digits of number $n > 1$
7. convert the number $n = a_{m,\mathbf{p}} \ldots a_{1,\mathbf{p}}a_{0,\mathbf{p}}$ into a number in the number system with base $\mathbf{p}+2$ with the form $n = a_{m,\mathbf{p}+2} \ldots a_{1,\mathbf{p}+2}a_{0,\mathbf{p}+2}$ by means of the algorithm given by Theorem 4.1;
8. while ($a_{0,\mathbf{p}+2} = 0$),{delete the least significant digit;table.insert($\mathbf{p}+2$)} end while;
9. $n \leftarrow$ the actual number n in which the least significant digit > 0;
10. if $(\mathbf{p}+2)^2 > n$ then {table.insert(n); exit;}
11. $\mathbf{p} \leftarrow \mathbf{p}+2$;
12. end **while;**

Theorem 5.2. *The complexity of the factorization algorithm is* $O(\sqrt{n}\ln^2 n)$

Proof. In the above algorithm, we look for divisors by checking the least significant number. We delete zeros if necessary and call the conversion algorithm repeatedly. This algorithm starts with the base $\mathbf{p} = 3$ and terminates when the base $\mathbf{p}$ is greater than $\sqrt{n}$.

The total number $T(n)$ of operations is given by

$$\begin{aligned} T(n) &= O\left(\frac{\log_2^2 n}{\log_2 3}\right) + O\left(\frac{\log_2^2 n}{\log_2 5}\right) + \cdots + O\left(\frac{\log_2^2 n}{\log_2 \sqrt{n}}\right) \\ &\leq \frac{\sqrt{n}}{2} O\left(\frac{\log_2^2 n}{\log_2 3}\right) \\ &= O(\sqrt{n}\ln^2 n). \end{aligned}$$

□

As an example, we factorize the number 2525. It is even and in the number system with base 3, it has the form $1011\,0112_3$. The least significant number 3 is not equal to 0. Let us start converting it as a number in the number system with base 5 :

base 3:	1	0	1	1	0	1	1	2
	1	0						
	1	-2						
correct	0	3	1					
		3	-5					
correct		2	0	1				
		2	-4	1				
correct		1	1	1	0			
		1	-1	-1	-2			
correct		0	3	3	3	1		
			3	-3	-3	-5		
correct			2	1	1	0	1	
			2	-3	-1	-2	1	
correct			1	1	3	3	1	2
			1	-1	1	-3	-5	0
base 5:				4	0	1	0	0

The number in the number system with base 5 has the form 40100_5. The least significant digit is 0, so there exists a divisor which is equal to the base 5. After removing the least significant digit, the resulting number also has the digit 0 as its least significant digit. Hence, we have 2 divisors 5 and 5^2. Then the divisor 5 can be placed on a stack twice. Removing the least significant digit again, we have 401_5. Repeating the conversion procedure, we have

base 5:	4	0	1
	4	0	
	4	-8	
correct	2	6	1
	2	2	-11
base 7:	2	0	3

Continuing, we have

base 7:	2	0	3
	2	0	
	2	-4	
correct	1	5	3
	2	3	-7
base 9:	1	2	2

We continue with the number 122_9:

base 9:	1	2	2
	1	2	
	1	0	2
	1	-2	2
base 11:		9	2

We can stop the algorithm now, because $92_{11} < 100_{11}$. The number 2525 does not have any more divisors. The last is $9 \cdot 11 + 2 = 101$ in decimal system. After placing the divisor 101 on our stack, we see that all divisors can be obtained from our stack which contains $5, 5, 101$.

References

[1] R. P. Brent and P. Zimmerann, *Modern Computer Arithmetic*, arXiv 1004.4710.

[2] D. Knuth, *The Art of Computer Programming*, Vol. 2, Addison-Wesley 1997, 1998

Received: November 30, 2013; *Accepted*: April 2, 2014

UNIVERSITY PRESS

Website: http://www.malayajournal.org/