Numeration and Computer Arithmetic
Some Examples

JC Bajard

LIRMM, CNRS UM2
161 rue Ada, 34392 Montpellier cedex 5, France

April 2007

Computer Arithmetic

Compromise:

- Speed
- Accuracy
- Cost

Heart:

- Number representations
- Associated algorithms

Approaches:

- Theory
- Software
- Hardware

Contents

Function Evaluation

an example of numeration

Briggs Algorithm (1561-1630)

- Evaluation of the logarithm, constructions of the first tables (15 decimal digits, 1624).
- In radix 2: digits $d_k = -1, 0, 1$, such that for a given x we have

$$x \prod_{k=1}^{n} (1 + d_k 2^{-k}) \simeq 1$$

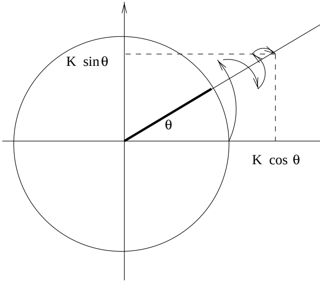
- The logarithm of x is

$$\ln(x) \simeq -\sum_{k=1}^{n} \ln(1 + d_k 2^{-k})$$

CORDIC Algorithm (COrdinate Rotation DIgital Computer, VOLDER 1959)

Basic step $d_n \in \{-1, 1\}$ (sign of z).

$$\begin{cases} x_{n+1} & = & x_n - d_n y_n 2^{-n} \\ y_{n+1} & = & y_n + d_n x_n 2^{-n} \\ z_{n+1} & = & z_n - d_n \arctan(2^{-n}) \end{cases}$$



For cosine and sine:
$x_0 = 1, y_0 = 0, z_0 = \theta (= \sum_{n \geq 0} d_n \arctan(2^{-n}))$
Constant factor
$K = \prod_{n=0}^{\infty} \sqrt{1 + 2^{-2n}} = 1.646760...$

Complex algorithm (BKM 1993)

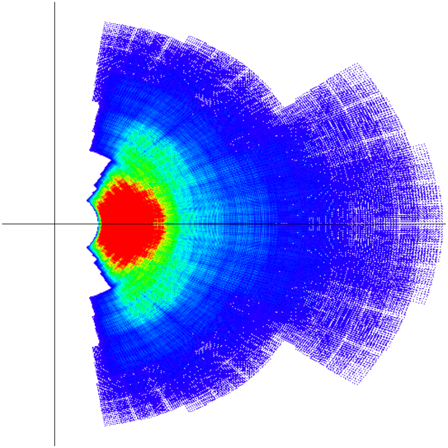
Basic step of the complex algorithm:

$$\begin{cases} E_{k+1} &= E_k(1 + d_k 2^{-k}) \\ L_{k+1} &= L_k - \ln(1 + d_k 2^{-k}) \end{cases}$$

with $d_k = d_k^r + i d_k^i$, and $d_k^r, d_k^i = -1, 0, 1$.
Two evaluation modes

- L-mode : $\begin{array}{lcl} E_n & \rightarrow & 1 \\ L_n & \rightarrow & L_1 + \ln(E_1) \end{array}$
- E-mode : $\begin{array}{lcl} L_n & \rightarrow & 0 \\ E_n & \rightarrow & E_1 e^{L_1} \end{array}$

26 12



$$\frac{1}{E_1} = \prod_{i=1}^{n}(1 + d_i 2^{-i}) \rightarrow L_n = -\sum_{i=1}^{n} \ln(1 + d_i 2^{-i}) = \ln(E_1)$$

Redundant Number Systems

Avizienis (1961)

- Redundant Number Systems
  Signed digits: $x_i \in \{-a, \ldots, -1, 0, 1, \ldots, a\}$ Radix β with $a \leq \beta - 1$.
- Properties
  - If $2a + 1 \geq \beta$, then each integer has at least one representation. An integer X, with $-a\frac{\beta^n - 1}{\beta - 1} \leq X < a\frac{\beta^n - 1}{\beta - 1}$, admits a unique representation

$$X = \sum_{i=0}^{n-1} x_i \beta^i \quad \text{with } x_i \in \{-a, \cdots -1, 0, 1, \ldots, a\}$$

  - If $2a \geq \beta + 1$, then we have a carry free algorithm. 25
- Borrow-save (Duprat, Muller 1989): extension to radix 2.

Example: radix 10, $a = 9$

$$
\begin{array}{rl}
\bar{2}359\bar{4}2 & (= -164138) \\
+\ 461\bar{6}7 & (= 46047) \\
\hline
011\bar{1}10 & (= t) \\
\bar{2}7100\bar{1} & (= w) \\
\hline
\bar{2}82\bar{1}1\bar{1} & (= s = -118091)
\end{array}
$$

Properties of the signed digits redundant systems

- **Advantages:**
  - Constant time carry-free addition
  - Large radix: parallelisation
  - Small radix: fast circuits
  - Increasing of the performances of the algorithms based on the addition [7]
- **Drawbacks:** comparisons, sign...

Non-Adjacent Form

- This representation is inspired from **Booth recoding (1951)** used in multipliers.
- **Definition of NAF_w recoding:** (Reitwiesner 1960) Let k be an integer and $w \geq 2$. The non-adjacent form of weight w of k is given by $k = \sum_{i=0}^{l-1} k_i 2^i$ where $|k_i| < 2^{w-1}$, $k_{l-1} \neq 0$ and each w-bit word contains at most one non-zero digit.

1. For a given k, $NAF_w(k)$ is unique.
2. For a given $w \geq 2$, the length of $NAF_w(k)$ is at most equal to the length of k plus one.
3. The average density of non-zero digits is $1/(w+1)$.

27 28

NAF_w Examples

We consider $k = 31415592$.

$k_2 =$	1	1 1 0 1	1 1 1 1	0 1 0 1	1 1 0 1	0 0 1 0	1 0 0 0
$NAF_2(k) =$	1 0	0 0 $\bar{1}$ 0	0 0 0 0	$\bar{1}$ 0 $\bar{1}$ 0	0 $\bar{1}$ 0 1	0 0 1 0	1 0 0 0
$NAF_3(k) =$	1 0	0 0 $\bar{1}$ 0	0 0 0 $\bar{1}$	0 0 3 0	0 0 $\bar{1}$ 0	0 $\bar{3}$ 0 0	$\bar{3}$ 0 0 0
$NAF_4(k) =$	1 0	0 0 $\bar{1}$ 0	0 0 0 0	0 0 $\bar{5}$ 0	0 0 0 $\bar{3}$	0 0 0 0	5 0 0 0
$NAF_5(k) =$		15 0	0 0 0 0	0 0 $\bar{5}$ 0	0 0 0 $\bar{3}$	0 0 0 0	5 0 0 0
$NAF_6(k) =$		15 0	0 0 0 0	$\bar{1}$ 0 0 0	0 0 $\overline{17}$ 0	0 0 0 0	$\overline{27}$ 0 0 0

Number Systems for Modular Arithmetic

Lattices and Modular Systems

- Number system: radix β and a set of digits $\{0, ..., \beta - 1\}$.

  $0 \leq A < \beta^n$ is expanded as: $A = \sum_{i=0}^{n-1} a_i \beta^i$.
- We denote by P the modulo, with $P < \beta^n$,

  $\beta^n \pmod{P} = \sum_{i=0}^{n-1} \epsilon_i \beta^i$ with $\epsilon_i \in \{0, ..., \beta - 1\}$
- A modular operation (for example: a modular multiplication):
  1. Polynomial operation: $W(X) = A(X) \bigotimes B(X)$
  2. Polynomial reduction : $V(X) = W(X) \bmod (X^n - \sum_{i=0}^{n-1} \epsilon_i X^i)$
  3. Coefficient reduction : $M(X) = \text{Reductcoeff}(V(X))$

Lattices and Modular Systems

Lattice approach

In a classical system "Reductcoeff" is equivalent to a combination of the carry propagation and the modular reduction:

$$\begin{pmatrix} -\beta & 1 & \dots & 0 & 0 \\ 0 & -\beta & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & -\beta & 1 \\ P & 0 & \dots & 0 & 0 \end{pmatrix} \begin{array}{c} \leftarrow \text{lattice} \\ \text{sublattice} \rightarrow \end{array} \begin{pmatrix} -\beta & 1 & \dots & 0 & 0 \\ 0 & -\beta & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & -\beta & 1 \\ \epsilon_0 & \epsilon_1 & \dots & \epsilon_{n-2} & (\epsilon_{n-1} - \beta) \end{pmatrix}$$

Lattices and Modular Systems

Example

For $P = 97$ and $\beta = 10$, we have $10^2 \equiv 3 \pmod{P}$. We consider the lattice:

$$\begin{pmatrix} B_0 \\ B_1 \end{pmatrix} = \begin{pmatrix} -10 & 1 \\ 3 & -10 \end{pmatrix}$$

Let $V(25, 12) = 25 + 12\beta$.

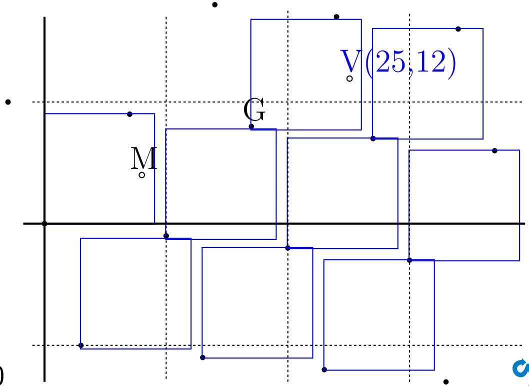
For reducing V, we determine $G(17, 8) = -2B_0 - B_1$ a vector of the lattice close to V.

Thus , $V(25, 12) \equiv M(8, 4) = V(25, 12) - G(17, 8)$.
We verify that $25 + 120 = 145 \equiv 48 \pmod{97}$

Lattices and Modular Systems

Example

The reduction is equivalent with finding a close vector.
Let $G(X)$ be this vector, then $M(X) = V(x) - G(X)$



$P = 97$ $\beta = 10$

Lattices and Modular Systems

A new system

- Polynomial reduction depends of the representation of $\beta^n \pmod{P}$
- In Thomas Plantard's PhD (2005), β can be as large as P, but with a set of digits $\{0, ..., \rho - 1\}$ where ρ is small.

Example: Let us consider a MNS defined with $P = 17, n = 3, \beta = 7, \rho = 2$. Over this system, we represent the elements of $\mathbb{Z}_{17}$ as polynomials in β, of degree at most 2, with coefficients in $\{-1, 0, 1\}$

Lattices and Modular Systems

A new system

0	1	2	3	4	5
0	1	$-\beta^2$	$1-\beta^2$	$-1+\beta+\beta^2$	$\beta+\beta^2$
6	7	8	9	10	11
$-1+\beta$	β	$1+\beta$	$-1-\beta$	$-\beta$	$1-\beta$
12	13	14	15	16	
$-\beta-\beta^2$	$1-\beta-\beta^2$	$-1+\beta^2$	β^2	$1+\beta^2$	

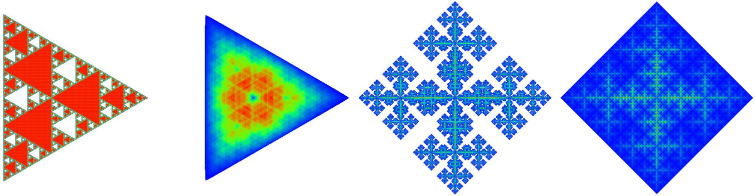
The system is clearly redundant.
For example: $6 = 1+\beta+\beta^2 = -1+\beta$, or
$9 = 1-\beta+\beta^2 = -1-\beta$.

Lattices and Modular Systems

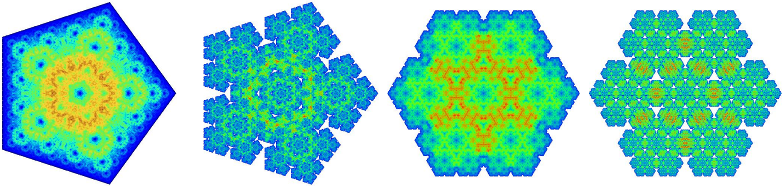
Construction of Plantard Systems

- In a first approach, n and $\rho = 2^k$ are fixed. The lattice is constructed from the representation of ρ in the number system. P and β are deduced. Efficient algorithm for finding a close vector. 31
- In a general approach, where P, β and n are given, the determination of ρ is obtained by reducing with LLL (Lenstra Lenstra Lovasz, 1982). No efficient algorithm for finding a close vector. 29

Conclusion



Thank you!



Annexes

Annexe: Avizienis Algorithm [10]

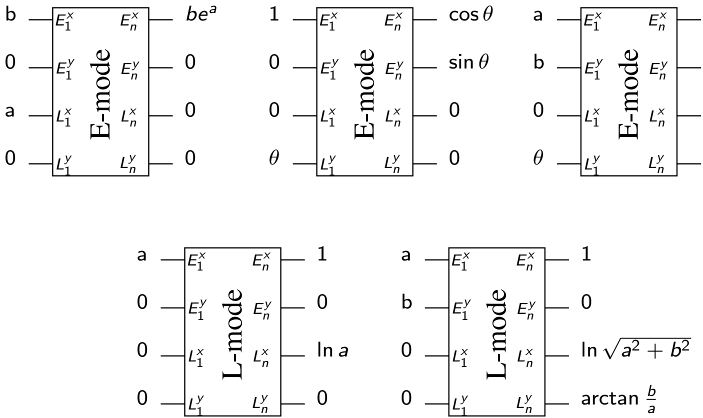
- We note $S = X + Y$ with
  $X = x_{n-1} \ldots x_0$
  $Y = y_{n-1} \ldots y_0$
  $S = s_n \ldots s_0$
- **Step 1:** For $i = 1$ to n in parallel,

$$
\begin{array}{rrcl}
 & t_{i+1} = & \bar{1} & \text{if, } x_i + y_i < -a + 1 \\
 & & 1 & \text{if, } x_i + y_i > a - 1 \\
 & & 0 & \text{if, } -a + 1 \leq x_i + y_i \leq a - 1 \\
\text{and} & w_i & = & x_i + y_i - \beta * t_{i+1} \\
\text{with} & w_n & = & t_0 = 0
\end{array}
$$

- **Step 2:** for $i = 0$ to n in parallel,

$$s_i = w_i + t_i$$

Annexe: Functions computable using one mode of BKM 7



$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

Annexe: NAF_w Computing [13]

Data: Two integers $k \geq 0$ and $w \geq 2$.
Result: $NAF_w(k) = (k_{l-1}k_{l-2}\dots k_1k_0)$.

```
l ← 0;
while k ≥ 1 do
    if k is odd then
        k_l ← k mod 2^w;
        if k_l > 2^(w-1) then
            k_l ← k_l − 2^w ;
        end
        k ← k − k_l;
    else
        k_l ← 0;
    end
    k ← k/2, l ← l + 1;
end
```

Annexe: Double and Add with NAF_w 13

Data: $P \in E$, $k = \in \mathbb{N}$ et $w \geq 2$, $NAF_w(k) = (k_{l-1} k_{l-2} \ldots k_1 k_0)$
$P_i = [i]P$ pour $i \in \{1, 3, 5, \ldots, 2^{w-1} - 1\}$
Result: $Q = [k]P \in E$.

```
begin
  Q ← P_{k_{l-1}};
  pour i = l − 2 ... 0 faire
    Q ← [2]Q;
    si k_i ≠ 0 alors
      si k_i > 0 alors
        Q ← Q + P_{k_i};
      sinon
        Q ← Q − P_{−k_i}
      fin
    fin
  fin
end
```

Lattices and Modular Systems

Annexe: Examples of Plantard System [22]

Example1: $P = 53$, $n = 7$, $\beta = 14$, $\rho = 2$.
We have $\beta^7 \equiv 2 \pmod{P}$. In this number system, integers have at least two representations, the total number of representations is 128.
The lattice could be defined by (vectors in row):

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \\ V_6 \\ V_7 \end{pmatrix} = \begin{pmatrix} -14 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -14 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -14 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -14 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -14 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -14 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -14 & 1 \\ 53 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Lattices and Modular Systems

Annexe: Examples of Plantard System 22

We can remark that there is a short vector : $(1,1,0,0,0,0,1) = V_6 + 14 * V_5 + 14^2 * V_4 + 14^3 * V_3 + 14^4 * V_2 + (14^5 + 1) * V_1 + V_7$.
From this vector we can construct a reduced basis of a sublattice, using that: $\beta^7 \equiv 2 \pmod{P}$

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 2 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 & 1 \end{pmatrix}$$

Lattices and Modular Systems

Annexe: Examples of Plantard System 22

Example #2: This example is proposed in PhD of Thomas Plantard. He gives some conditions that number system must verify: $\beta^8 \equiv 2 \pmod{P}$ and $\rho = 2^{32}$.

P is the determined:

$P = 115792089021636622262124715160334756877804245386980633020041035952359812890593$

Then β is deduced

$\beta = 144740111277045777827655893952245323141792170589214883950498277337595903999 96$