Proc. Indian Acad. Sci., Vol. 81 A, No. 2, 1975, pp. 58 – 79

Finite segment p-adic number systems with applications to exact computation

E. V. KRISHNAMURTHY, F.A.SC., T. MAHADEVA RAO AND K. SUBRAMANIAN†

Department of Applied Mathematics, Indian Institute of Science, Bangalore 560012

MS received 5 December 1974

ABSTRACT

A fractional weighted number system, based on Hensel's p-adic number system, is proposed for constructing a unique code (called Hensel's code) for rational numbers in a certain range. In this system, every rational number has an exact representation. The four basic arithmetic algorithms that use the code for the rational operands, proceed in one direction, giving rise to an exact result having the same code-word-length as the two operands. In particular, the division algorithm is deterministic (free from trial and error). As a result, arithmetic can be carried out exactly and much faster, using the same hardware meant for p-ary systems.

This new number system combines the best features and advantages of both the p-ary and residue number systems. In view of its exactness in representation and arithmetic, this number system will be a very valuable tool for solving numerical problems involving rational numbers, exactly.

* * * *

"*There still remain three studies suitable for free man. Arithmetic is one of them.*"—Plato

1. INTRODUCTION

THIS paper is concerned with the problem of exact or error-free computation in digital computers. It is well known that digital computers work with finite length numbers in a p-ary scale, where p is any positive integer called 'radix' or 'base'; usually either a binary or decimal scale is chosen for this purpose. Such a system, known as the decimal or p-ary system enables us to measure any numerical quantity in steps of p^{-r} where r is the length or the number of digits available for calculations. It is obvious that the p-ary system has the disadvantage, *viz.*, many of the rational numbers do not have exact representations, since they may not terminate within r-digits or may be recurring or periodic. Hence it is not possible to represent every

† School of Automation, Indian Institute of Science.

rational number of the form a/b (a and b are both integers, $b \neq 0$) exactly, if b has factors which are relatively prime to the base of choice p. This results in an error in representing every rational quantity. Also we face the difficulty of accumulated round-off or truncation errors. These are unavoidable since even with such simple repetitive operations such as addition, multiplication or division the number of digits of the result can increase so much that the results cannot be held fully in the registers available in the computer; accordingly, we are compelled to discard a certain number of digits. Such errors accumulate one after another from operation to operation, originating fresh errors. Thus error analysis and propagation of errors became important aspects of study in computational mathematics. The developments in this area have been very significant and error analysis is available for most of the widely used algorithms; in fact, we are cautioned that no algorithm should be considered valid unless an estimate of the errors involved is available.

It is also well-known that the numerical problems are classified as well-conditioned or ill-conditioned. In well-conditioned problems one can obtain reliable results in spite of small round-off errors committed in the computational algorithm. But most of the practical problems do not turn out to be well-conditioned. These are called ill-conditioned in the sense, for very small errors in computation the errors get amplified and propagate so wildly that the computed solution bears little or no resemblance to the actual solution. This is particularly so in matrix computations. For such problems residue-arithmetic procedures have been suggested (Young and Gregory, 1973) in recent years as a means to perform exact computation, thereby totally circumventing the problem of round-off errors. In these procedures every rational number is transformed into an integer by proper scaling which is then represented in a residue number system with one or more prime moduli. While such a representation is useful for exact computation, since division is not an easy operation while using several prime moduli (Szabo and Tanaka 1967) the computational problems get involved. Also when numbers involved are in a wide range many primes have to be chosen since the residue number system does not provide a range of the form p^r (where r is the number of digits and p is the base of representation) as in p-ary base.

These difficulties motivated us to look for an alternative number system which possesses the best features as well as advantages of both the p-ary and residue number systems. Such a number system is the p-adic number system which was first proposed by Hensel in 1900 (Knuth 1969). A detailed introduction to p-adic numbers is found in Bachman (1964). It is surprising that the potentialities and the application of the p-adic numbers for computer arithmetic have not been explored. This may be partly due

to the fact that every rational number other than integers or radix fractions (those fractions whose denominators are powers of the radix p) has only an infinite recurring p-adic expansion and a truncation of this expansion to a finite length is catastrophic in the sense, the value of this truncrated number bears no resemblance to the convergent value, unlike the p-ary representations, where the truncated number is almost always a close approximation to its actual value.

In this paper, we will show that a segmented p-adic expansion of a rational number can be used as a code for the representation of the rational number. We will call this code as the Hensel Code (in honour of Hensel) denoted by $H(p, r)$, where p is the prime base and r is the number of digits used. The conditions for construction of this code are:

(i) The numerator and denominator of the rational numbers to be represented have a prescribed bound.

(ii) The p-adic expansions are terminated at the right such that r is even.

The above two conditions are sufficient for designing a very efficient p-adic computer arithmetic system. This will be shown in a later section.

2. PROPERTIES OF $H(p, r)$ CODES

The Hensel Codes mentioned above have the following important properties:

(i) Every rational number a/b (where a and b are integers with their greatest common divisor† $(a, b) = 1$, $b \neq 0$) in the range $0 \leq a \leq N$ and $0 < b \leq N$ where N is a positive integer chosen such that $N \leqslant p^{r/2}/\sqrt{2}$, can be uniquely coded as an r-digit ordered sequence where each digit assumes values from $0, 1, \ldots, (p-1)$, where p is the prime chosen.

(ii) $H(p, r)$ code satisfies certain weight properties which enable us to convert them into rationals. The individual digit positions have the following positive negative fractional weights:

$$p^0, p^1, p^2, \ldots, p^{r/2-1}, \frac{-p^{r/2}}{p^{r/2}-1}, \frac{-p^{r/2+1}}{p^{r/2}-1}, \quad \cdots \frac{-p^{r-1}}{p^{r/2}-1}.$$

Thus positive and negative rational numbers are representable without an explicit sign. The positive integers occupy the positions $0, 1, \ldots, r/2-1$ followed by a sequence of zeros. The negative integers occupy the same positions followed by a sequence of digits of value $(p-1)$. The representation of radix fractions (both positive and negative) is exactly similar to positive and negative integers except that the p-adic point can move to the right up to the $(r/2-1)$-th position.

Other rationals, however, can occupy all the r-digits. Among these, those which have denominators which are divisors of $(p^{r/2}-1)$ satisfy

† denoted as GCD.

the weight properties, in the sense the weighted digit sum using the above weights, gives rise to their actual value. Such fractions will be called as " soft fractions ".

For the other fractions (called " hard fractions ") the weighted sum is unique for each one of them, but will not coincide with their actual value; in fact this sum holds only modulo p^r.

(iii) The negative rational numbers have a valid radix complement representation (Richards 1955) as in p-ary arithmetic; however, unlike in positive p-ary arithmetic the weight structure is retained. In this sense, it resembles the polarization operation in negative p-ary arithmetic (Sankar *et al* 1973 *a*). In view of this, multiplication and also division can be carried out directly without corrections that are required (Richards 1955) in positive p-ary complement number representations.

(iv) All the four basic arithmetic operations using $H(p, r)$ codes proceed in one direction (left to right); also the length of the resulting code remains constant, provided the resulting rational number is in the range specified in (i).

(v) The arithmetic using $H(p, r)$ codes is very much simpler than rational arithmetic, where every rational number is represented as an ordered pair of integers (Knuth 1969). For instance, addition/subtraction in rational arithmetic requires three multiplications, one addition and a reduction of the fraction to the minimal form, by dividing both the numerator and denominator by their greatest common divisor; multiplication and division involve two multiplications and also the reduction of the result to the minimal form.

The $H(p, r)$ code, uses a unified representation for all rationals (as in p-ary system) and arithmetic operations are performed with a constant length word somewhat similar to the residue number system. Unlike the residue system, where division cannot be performed easily, the p-adic system permits us to do an exact division operation as in rational arithmetic.

(vi) The division operation using $H(p, r)$ codes is deterministic in the sense, it does not involve the comparison of relative magnitudes of the operands (Sankar *et al.* 1973 *b*). This operation is exactly similar to multiplication; this is not surprising since in rational arithmetic using an ordered pair of integers, the multiplication and division operations are very similar.

(vii) The essential distinction between using a truncated p-ary code and a truncated p-adic code [$H(p, r)$ code] using r-digits is that in the p-ary system the numbers are exactly representable only if they are either integers or radix fractions; hence all other rationals can only be approximately represented.

In the $H(p, r)$ code, however, every rational number a/b which is an element of the Farey sequence (Niven and Zuckerman 1966) with numerators and denominators less than N (we call such fractions as order N Farey fractions) are exactly representable. The relation between these two systems of measurement reminds us of the scale of equal temperament and mean-tone temperament used in music.

(viii) The conversion from a rational to its $H(p, r)$ code (subject to choice of p and r) is straightforward involving a recursive solution of congruences; this can also be realized as a division process using $H(p, r)$ code, if the numerator and denominator are represented in p-adic form.

The conversion from $H(p, r)$ code to rationals is a little more involved. In fact, without this conversion it seems not possible to know the sign and magnitude of the numbers. The conversion algorithm for integers, radix and soft fractions (both positive and negative) is direct involving the formation of the weighted sum. However, the hard fractions can only be converted to rationals through a solution of a congruence or a diophantine equation or by a direct read out.

(ix) Since all the basic arithmetic operations are performed exactly for a given range of the rationals, the $H(p, r)$ codes can be used for exact numerical computations.

(x) Although one could think of a $H(p, r)$ code where r is odd, such a code does not possess unique weights for digit positions; hence such a code cannot be easily converted to rationals by numerical computations.

3. p-ADIC NUMBER REPRESENTATIONS

3.1. *Principle*

Any rational number $a = a/b \cdot p^n$ where n, a, b are positive or negative integers $b \neq 0$, and GCD $(a, b) = 1$, can be written in the p-adic form

$$a = \sum_{j=n}^{\infty} a_j p^j \tag{1}$$

where, the coefficients a_j are integers such that $0 \leq a_j \leq (p-1)$. The infinite series (1) converges to a, with respect to the p-adic norm.

For example 1/3 has the p-adic expansion

$$1/3 = \cdot 2313131\ldots \quad (\text{for } p = 5)$$
$$= 2 + 3p(1 + p^2 + p^4 + \ldots) + p^2(1 + p^2 + p^4 + \ldots) \tag{2}$$

The infinite series $1 + p^2 + p^4 + \ldots$ converges to $1/1 - p^2$ (Bachman 1964) with respect to the p-adic norm. Thus for $p = 5$, the series (2) converges to $8/24 = 1/3$.

Note: In the conventional sense the series (2) under usual norm is divergent. However, in the sense of p-adic norm it is convergent.

For convenience the expansion (1) will be denoted by the sequence

$a_n a_{n+1}$ $\cdot a_0 a_1 \ldots$ for n negative

$\cdot a_0 a_1 \ldots$ for $n = 0$

$\cdot 00 \ldots 0 a_n a_{n+1} \ldots$ for n positive.

It is easy to see that integers as well as radix fractions have representations which are identical to the p-ary representations except for the fact that the sequence is written here from left to right in the ascending powers of p, instead of right to left (as a mirror reflection); however, other rationals have a totally different representation, with the actual rational being realized only as a sum of the infinite series.

It is interesting to observe that negative rationals occur as true-complement (left-to-right) of the positive number.

e.g., $-\frac{1}{3} = \cdot 313131 \ldots$ for $p = 5$.

3.2. *Algorithm for conversion of rational number to* p-*adic form*

Given a rational number $\alpha = a/b$ with GCD $(a, b) = 1$ and $b \neq 0$, its p-adic expansion can be obtained by the following algorithm.

Step 0. Set $\beta = \alpha$. Find n such that

$\alpha = \frac{c}{d} \cdot p^n$. If $n > 0$ set

$a_0, a_1, \ldots, a_{n-1} = 0$.

Go to Step 2.

Step 1. Find n such that $\beta = \frac{c}{d} \cdot p^n$

$\left(e.g., \frac{25}{41} = \frac{1}{41} \cdot 5^2, \text{ for } p = 5\right)$.

Step 2. Solve the congruence $dx \equiv 1 \pmod{p}$

If x_n is a solution, then

$a_n = c \cdot x_n \pmod{p}$.

Step 3. Set $\gamma = \beta - a_n p^n$

If $\gamma = 0$, set $a_i = 0$, for $i > n$ and stop.

Otherwise, set $\beta = \gamma$ and go to step 1.

Note: This algorithm is nonterminating since the convergence to the given rational can be achieved only for infinite number of digits, if the rational is neither an integer nor a radix fraction.

4. SEGMENTED WEIGHTED *p*-ADIC CODES

4.1. *Word format*

In the last section we mentioned that a general rational number (other than integers and radix fractions) does not terminate in a *p*-adic expansion and the convergence to the actual value is obtained only for infinite terms. This is in contrast to the conventional *p*-ary representation where one can terminate with a finite number of digits and yet closely approximate every non-terminating rational. In the case of *p*-adic expansion truncation of a number is mathematically meaningless, since the truncated series represents a "large" integer. Therefore, if *p*-adic number system is to be of any practical use one needs a finite expansion. In such a case one can only construct a finite length code which is a segment derived from the *p*-adic expansion of the rational. In order that this code be unique and weighted, the following conditions are to be satisfied:

(i) Given the range of numerators and denominators of rational numbers, every rational number in the Farey sequence has a unique code.

(ii) All the rationals contained in the given range can be represented using *r*-digits in a base *p*.

(iii) The digit positions in the code have some definite weights so that conversion from the code to the rational is possible.

The above three conditions set up a constraint on the construction of $H(p, r)$ for a given p and r.

It will be shown below that for a given prime p, the following logical organization satisfies all the above three conditions:

(*a*) The number of digits r is even.

(*b*) The numerator and denominator of rational numbers are less than or equal to N. This means there are exactly $r/2$ $(0, 1, \ldots r/2 - 1)$ digits allocated for the representation of positive and negative integers, assuming that the *p*-adic point is to the left of the zeroth digit. By allowing the *p*-adic point to move to the right of $0, 1, 2, \ldots r/2 - 1$ positions the above integers become radix fractions and hence all radix fractions whose denominators do not exceed N become representable. All other rationals will in general occupy all the *r*-digits. In particular, the word format for integers and radix fractions will be as follows.

Positive integers: These are of the form

$$\bullet a_0\, a_1 \ldots a_{r/2-1} \underbrace{00 \ldots 0}_{r/2}$$

consisting of a sequence of zero digits from $(r/2)$-th to $(r-1)$-th position.

e.g., $p = 5$, $r = 8$.

$\cdot 44430000$ is the positive integer 499.

Negative integers: These are in the complement form, occupying $0, 1, \ldots, (r/2) - 1$ digits followed by a sequence of digits of value $(p-1)$ from $r/2, (r/2)+1, \ldots, r-1$ digits.

$$\cdot a_0 a_1 \ldots a_{r/2-1} \underbrace{(p-1), (p-1), (p-1)\ldots(p-1)}_{r/2}$$

e.g., $p = 5$, $r = 8$

$\cdot 02144444$ is the negative integer -90.

Radix fractions: These are obtained by shifting the p-adic point of integers (positive or negative) by less than or equal to $(r/2-1)$ digits.

e.g., $421 \cdot 00000$ is the radix fraction of value $39/125$

$421 \cdot 40000$ is the radix fraction of value $539/125$

$421 \cdot 44444$ is the negative radix fraction $-86/125$.

Soft and hard fractions: These are of the form

$a_{-n} a_{-n+1} \ldots a_{-1} . a_0 a_1 \ldots a_m$ where $n + m + 1 = r$ and $n \leq r/2 - 1$ and the digits have general values from 0 to $(p-1)$.

Note: As before we mean, by an order N Farey fraction, a fraction of the form $\pm a/b$, where $0 \leqslant a,\ b \leqslant N,\ b \neq 0$.

It is easy to prove that every order N Farey fraction has a unique representation within r-digits, if $N \leqslant p^{r/2}/\sqrt{2}$. The proof is based on arguments involving congruence relations.

We will also show that this choice of N would permit us to code all the order N Farey fractions.

It is known that the number of order N Farey fractions in the interval [0, 1] is asymptotically equal to $3N^2/\pi^2 \approx N^2/3$ (Beiler 1964, Abramowitz and Stegun 1965). Since there are an equal number of their reciprocals, we have a total of $2N^2/3$ positive rationals in this range. Since the negative rationals in the same range have also to be represented, we have approximately a total of $4N^2/3$ rationals in this range.

The logical structure of the format we have chosen (including the p-adic point) can represent more than p^r different sequences. Therefore, the inequality $4N^2/3 \leqslant p^r$ will always be satisfied for $N \leqslant p^{r/2}/\sqrt{2}$. (3)

4.2. *Weight assignment*

We will now discuss the assignment of weights for the digit positions in the $H(p, r)$ code assuming r is even. Since by construction the positive integers occupy the first $r/2$ positions and followed by a sequence of zeros, the weights of this positions have to be necessarily the p-ary weights:

$$p^0, p^1, \ldots, p^{r/2-1}$$

Also the negative integers which occur in complement form occupy the first $r/2$ digits followed by a sequence of digits of value $(p-1)$. Using this as well as the fact that all other negative rationals which occur in complement form should have a weight opposite in sign to that of the corresponding positive rationals, we can show that the weights for the remaining $r/2$ positions have to be negative and fractional. These are respectively

$$-\frac{p^{r/2}}{p^{r/2}-1}, \quad -\frac{p^{r/2+1}}{p^{r/2}-1}, \quad \ldots, \quad -\frac{p^{r-1}}{p^{r/2}-1}.$$

Thus we arrive at a new class of positional number systems with fractional weights, as a consequence of truncating the p-adic expansion. Since each order N Farey fraction has a unique Hensel Code the assignment of the above weights will result in a unique weighted sum for each one of them.

e.g., A table of weights of a few rational numbers is provided below (Table 1).

Table 1

No.	Rational number	$H(5, 4)$	Weight
1	1/8	·2414	3/24
2	1/16	·1234	−311/24
3	−1/16	·4210	311/24
4	1/9	·4201	211/24
5	1/7	·3302	182/24

The proof that each one of the order N Farey fractions will get a unique weight follows from an argument involving the solution of a diophantine equation; it is omitted here. It is clear from the assignment of weights that every weighted sum will have the denominator $p^{r/2}-1$; for negative integers the numerator of the weighted sum will be the 'that' integral multiple of $p^{r/2}-1$. For rationals a/b where b is a divisor of $p^{r/2}-1$ the weighted sum W will be equal to their actual value thus satisfying the equation

$$a/b = \frac{W}{p^{r/2}-1}$$

or

$$a(p^{r/2} - 1) - bW = 0. \tag{4}$$

Such fractions are called soft fractions since their conversion from p-adic to rational involves finding only the weighted sum.

Other fractions, however, will not have code weights equal to their actual value (*see* Table 1); so we call them as pseudo weights. These rationals are called hard fractions, and they will only satisfy the congruential relation

$$a(p^{r/2} - 1) - bW \equiv 0 \pmod{p^r} \tag{5}$$

Hence for converting hard fractions we have to solve for (5) in the range prescribed

e.g.,

$p = 5 \quad r = 4$

$1/9 = \cdot 4201$

pseudo weight $= 211/24$

Thus we have to solve for

$$24\,a - 211\,b \equiv 0 \pmod{625}$$

which has a solution for right hand side equal to -1875, for which $a = 1$, $b = 9$.

Thus a p-adic system with r-digits in which r is even and $r/2$ digits are reserved for integers exhibits features which are common to both the p-ary as well as the residue number systems. In particular, operations with purely integers and soft fractions work analogous to the p-ray system while operations with hard fractions are analogous to using residue arithmetic. The price paid for this is of course in terms of the loss of information of sign and magnitude of hard fractions. To determine these one has to solve (5).

Since all the four basic arithmetic operations can be carried out and zero can be detected all numerical algorithms can be implemented to obtain exact results. We will discuss the application of the $H(p, r)$ code for exact solution of linear equations in Section 7.

5. Arithmetic Algorithms

In this section, we will describe the four basic arithmetic algorithms using Hensel codes. As we will observe all these algorithms proceed from the lower index (power) position of the radix and proceed towards the higher positions. Even the division operation proceeds in this way. This is in

contrast with the arithmetic in p-ary or any general weighted systems known so far. In the case of a non-weighted system like the residue system the division process does not exist. In addition, the division operation using p-adic numbers is deterministic for all possible divisors which permits us to carry out a very fast division algorithm unlike the p-ary system.

5.1. *Addition*

Addition in a p-adic system is similar to that in positive p-ary base. In general, the algorithm for addition of any two numbers α and β (positive or negative) given by

$$\alpha = a_{-m}a_{-m+1} \ldots a_{-1}.a_0a_1 \ldots a_n$$

$$\beta = b_{-m}b_{-m+1} \ldots b_{-1}.b_0b \ldots b_n$$

(For convenience we take that α and β have identical number of digits) aligns the p-adic point and proceeds finding the sum digit s_i and carry digit c_{i+1} from a knowledge of a_i, b_i and c_i (the suffix here indicates the digit position to which a, b, c, s belong. The carry c_{n+1} arising from the addition of $a_n + b_n + c_n$ is ignored.

Thus

$$s_i = a_i + b_i + c_i \qquad (\text{mod } p)$$

for

$$i = -m, -m+1, \ldots n$$

with

$$c_{-m} = 0$$

and

$$c_{i+1} = 1 \quad \text{if} \quad a_i + b_i + c_i \geq p$$
$$= 0 \text{ otherwise}$$

Ignore c_{n+1}.

e.g.,

$$p = 5 \quad r = 4$$
$$\alpha = 4/9 = \cdot 1124$$
$$\beta = 8/9 = \cdot 2243$$

and

$$\sigma = \alpha + \beta = 12/9 = 4/3 = \cdot 3313.$$

Table 2

i	0	1	2	3
a_i	1	1	2	4
b_i	2	2	4	3
c_i	0	0	0	1
s_i	3	3	1	3

For convenience a table of p-adic representation for rationals whose numerator (a) and denominator (b) do not exceed $N = 17$ in 5-adic base for $r = 4$ is provided (Table 3).

5.2. *Complementation algorithm and subtraction*

Subtraction using p-adic numbers can be done exactly similar to that in positive p-ary system taking the borrow from the higher digit position. However, it is more convenient to realize it as a complemented addition, since the negative numbers naturally occur in this form.

The complementation algorithm is exactly similar to taking true or radix complement of positive p-ary numbers except that it is taken from the lower index position.

Let

$$\alpha = a_{-m} \cdot a_{-m+1} \ldots a_{-1} \cdot a_0 a_1 \ldots$$

and $$\bar{\alpha} = \bar{a}_{-m}\bar{a}_{-m+1} \ldots \bar{a}_{-1} \cdot \bar{a}_0\bar{a}_1 \ldots$$

This algorithm consists of the following rules

Rule 1

If

$$a_i \neq 0, \quad \text{for} \quad i = -m, -m+1, \ldots n$$

then

$$\bar{a}_i = p - a_i \quad \text{for} \quad i = -m$$

and

$$\bar{a}_i = (p-1) - a_i \quad \text{for} \quad -m+1 \leq i \leq n$$

Table 3. Table of Hensel codes

$H(5, 4)$

b \ a	1	2	3	4	5	6	7	8
1	·1000	·2000	·3000	·4000	·0100	·1100	·2100	·3100
2	·3222	·1000	·4222	·2000	·0322	·3000	·1322	·4000
3	·2313	·4131	·1000	·3313	·0231	·2000	·4313	·1231
4	·4333	·3222	·2111	·1000	·0433	·4222	·3111	·2000
5	1·000	2·000	3·000	4·000	·1000	1·100	2·100	3·100
6	·1404	·2313	·3222	·4131	·0140	·1000	·2404	·3313
7	·3302	·1214	·4021	·2423	·0330	·3142	·1000	·4302
8	·2414	·4333	·1303	·3222	·0241	·2111	·4030	·1000
9	·4201	·3012	·2313	·1124	·0420	·4131	·3432	·2243
10	3·222	1·000	4·222	2·000	·3222	3·000	1·322	4·000
11	·1332	·2120	·3403	·4240	·0133	·1411	·2204	·3041
12	·3424	·1404	·4333	·2313	·0342	·3222	·1202	·4131
13	·2034	·4014	·1143	·3123	·0203	·2232	·4212	·1341
14	·4101	·3302	·2013	·1214	·0410	·4021	·3222	·2423
15	2·313	4·131	1·000	3·313	·2313	2·000	4·313	1·231
16	·1234	·2414	·3104	·4333	·0123	·1303	·2042	·3222
17	·3043	·1132	·4121	·2210	·0304	·3342	·1431	·4420

b \ a	9	10	11	12	13	14	15	16	17
1	·4100	·0200	·1200	·2200	·3200	·4200	·0300	·1300	·2300
2	·2322	·0100	·3322	·1100	·4322	·2100	·0422	·3100	·1422
3	·3000	·0413	·2231	·4000	·1413	·3231	·0100	·2413	·4231
4	·1433	·0322	·4111	·3000	·2433	·1322	·0211	·4000	·3433
5	4·100	·2000	1·200	2·200	3·200	4·200	·3000	1·300	2·300
6	·4222	·0231	·1140	·2000	·3404	·4313	·0322	·1231	·2140
7	·2214	·0121	·3423	·1330	·4142	·2000	·0402	·3214	·1121
8	·3414	·0433	·2303	·4222	·1241	·3111	·0130	·2000	·4414
9	·1000	·0301	·4012	·3313	·2124	·1420	·0231	·4432	·3243
10	2·322	·1000	3·322	1·100	4·322	2·100	·4222	3·100	1·422
11	·4324	·0212	·1000	·2332	·3120	·4403	·0340	·1133	·2411
12	·2111	·0140	·3020	·1000	·4424	·2404	·0433	·3313	·1342
13	·3321	·0401	·2430	·4410	·1000	·3034	·0114	·2143	·4123
14	·1134	·0330	·4431	·3142	·2343	·1000	·0201	·4302	·3013
15	3·000	·4131	2·231	4·000	1·413	3·231	·1000	2·413	4·231
16	·4402	·0241	·1421	·2111	·3340	·4030	·0310	·1000	·2234
17	·2024	·0113	·3102	·1240	·4234	·2323	·0412	·3401	·1000

Rule 2

If $a_i = 0$ for $-m \leq i \leq j$,

then $\bar{a}_i = 0$ for $-m \leq i \leq j$

and $\bar{a}_{j+1} = p - a_{j+1}$

$\bar{a}_i = (p-1) - a_i$ for $i = j+2 \leq i \leq n$.

e.g., $p = 5, \quad r = 4$

Let $\alpha = 5/4 = \cdot 0433$

then $\dot{\alpha} = -5/4 = \cdot 0111$.

5.3. *Multiplication*

Let α and β be respectively the multiplicand and multiplier (positive or negative).

$$\alpha = a_{-m}a_{-m+1} \ldots a_{-1} \cdot a_0 a_1 \ldots a_n$$

$$\beta = b_{-m}b_{-m+1} \ldots b_{-1} \cdot b_0 b_1 \ldots b_n.$$

The multiplication algorithm is similar to the multiplication in p-ary system except that the product is developed only to the same length as the multiplier and multiplicand. The multiplication algorithm consists in forming $P_{ij} = b_i a_j$ for $i = -m, -m+1, \ldots, n$ and for the values of $j = -m, -m+1, \ldots n-i$ and then forming the partial products P_i and the final product P by shifts and additions as specified by the following recursions.

For each i form

$$P_{ij} = b_i \cdot a_j \quad \text{for} \quad j = -m, -m+1, \ldots n-i.$$

Since this is a multiplication of two single digits, the product in general will consist of two digits.

Then obtain

$$P_i = \sum_{j=-m}^{n} P_{ij} \triangle (m+j)$$

where $\triangle (x)$ indicates right shift of the number by x digits and

$$\alpha \cdot \beta = P = \sum_{j=-m}^{n} P_i \triangle (m+i).$$

To place the p-adic point for P, we use the following rule:

Replace the index k (which runs through $-m, -m+1, \ldots, n$) of each digit of P by $(k-m)$; here m is the number of digits in multiplier and multiplicand to the left of the p-adic point.

Note that the p-adic point is assumed to be between indices $k = -1$ and 0.

e.g., Take

$$\alpha = 1/4 = \cdot 4333$$
$$\beta = 1/3 = \cdot 2313 \qquad \text{for} \qquad p = 5, \quad r = 4$$
$$\alpha \,.\, \beta = 1/12 = \cdot 3424$$

Table 4

i, j	0	1	2	3
a_j ..	4	3	3	3
b_i ..	2	3	1	3
$P_{0,0}$	3	1		
$P_{0,1}$		1	1	
$P_{0,2}$			1	1
$P_{0,3}$				1
P_0 ..	3	2	2	2
P_1 ..		2	1	1
P_2 ..			4	3
P_3 ..				2
P ..	3	4	2	4

5.4. *Division*

Let $\alpha = \cdot a_0 a_1 a_2 \ldots a_n$ be the dividend and $\beta = \cdot b_0 b_1 \ldots b_n$ be the divisor and $\gamma = \cdot q_0 q_1 \ldots q_n$ be the quotient. The division operation is similar to multiplication in p-adic arithmetic, but for choosing the multiplicative inverse of b_0, the leading digit of β. Thus in spite of the fact that the rational numbers have been mapped into the p-adic form, they continue to retain the similarity of operations involved in multiplication and division in rational arithmetic.

As already mentioned, the division process is deterministic and does not require the trial and error process encountered in a p-ray system. Also as in multiplication the operation proceeds from the left to right. This again is in contrast to the p-ray division schemes. The division algorithm

consists of the following steps. We assume here $b_0 \neq 0$ (for $b_0 = 0$ *see* remark below).

Set $R_0 = \alpha$ = Zeroth partial remainder.

Let R_i denote the partial remainder at the i-th stage and R_{ii} denote its i-th digit. Then

$$q_i = R_{ii} \cdot b_0^{-1} \pmod{p} \quad \text{for} \quad i = 0, 1, \ldots n$$

where b_0^{-1} is the multiplicative inverse of b_0, the leading digit of β.

The next partial remainder R_{i+1} is then formed using

$$R_{i+1} = R_i + q_i \cdot \bar{\beta} \cdot \triangle (i)$$

where $\bar{\beta}$ = complement of β, and $\triangle (i)$ denotes right shift by i digits. The algorithm terminates when q_n is obtained.

Remark. If $b_0 = 0$, shift the divisor left keeping count until the first digit is non-zero and suitably adjust the quotient. It is assumed that by shifting the divisor, it does not go outside the range of the order N Farey fraction for the given $H(p, r)$; in such a case the division operation is invalid.

e.g., $p = 5, \quad r = 4$

$\alpha = 8/9 = \cdot 2243$

$\beta = \frac{1}{2} = \cdot 3222$

$\gamma = 16/9 = \cdot 4432$

$b_0^{-1} = 2, \quad \bar{\beta} = \cdot 2222$

Table 5

i	0	1	2	3
R_0 ..	2	2	4	3
$q_0\bar{\beta}\triangle(0)$ ..	3	4	4	4
R_1 ..	0	2	4	3
$q_1\bar{\beta}\triangle(1)$	0	3	4	4
R_2 ..	0	0	4	3
$q_2\bar{\beta}\triangle(2)$	0	0	1	2
R_3 ..	0	0	0	1
$q_3\bar{\beta}\triangle(3)$	0	0	0	4
R_4 ..	0	0	0	0
γ ..	4	4	3	2

Remark. All the above algorithms can be conveniently implemented by expressing the segmented p-adic numbers having p-adic points, in the exponent-mantissa form assuming the point to be at the left end; the mantissa used here is of fixed length and the exponents are only used as a means for locating the p-adic point.

6. CONVERSION ALGORITHMS

We will describe three different procedures for conversion of p-adic order N Farey fractions into rational form.

Before proceeding for conversion, the positive and negative integers and radix fractions which have well defined word format (*see* Section 4.1) are filtered-out and converted using the weights described earlier.

For other rationals, if the weighted sum of the code (reduced to the minimal rational form) is an order N Farey fraction, then it is a soft fraction with value equal to the weighted sum.

For hard fractions one of the following three procedures can be used.

6.1. *Successive addition or multiplication method*

Since integers and soft fractions are easily tested out, we can use the following procedure for conversion of the hard fractions. The given hard fraction can be added to itself 1, 2, ..., b times until it becomes an integer. This can also be realized by successive multiplication. To speed up, one could also carry out the same process simultaneously with the reciprocal of the given code; in such a case we will be able to convert the code in s-steps where $s = \min(a, b)$. Also instead of just testing for integers one could check after each addition whether the result is a soft fraction. If so it can be converted directly.

6.2. *Method of congruences*

Here we solve the congruence (5)

$$a(p^{r/2} - 1) - bW \equiv 0 \pmod{p^r} \tag{5}$$

by reformulating the problem as a solution of the diophantine equation

$$a(p^{r/2} - 1) - bW = k \quad p^r \quad \text{for} \quad k = \pm 1, \pm 2, \quad \text{etc.} \tag{6}$$

It is well known that a diophantine equation of the form (Gelfond 1961, Khinchin 1964, Mordell 1969, Hardy and Wright 1960)

$$ma - nb = 1 \tag{7}$$

has a solution $a = a_0$ and $b = b_0$ which are obtained by expanding m/n as a continued fraction and taking the last but one convergent which equals b_0/a_0. Since there are many solutions to the problem one has to also try the other solutions.

$$a_0 \pm tn, \; b_0 \pm tm \quad \text{for} \quad t = 0, 1, 2, \ldots$$

and take those which are acceptable according to a certain criterion.

Thus (6) can be solved for various values of k and that solution which gives an order N Farey fraction is taken.

Since the above procedure involves searching through two variables k and t it is somewhat slow. So we suggest here an alternate procedure in which one can obtain the value of k. For this purpose we use the weight $W'/p^{r/2} - 1$ of the reciprocal of the given p-adic number whose weight $W/p^{r/2} - 1$ is known.

Then the value of k to be taken is

$$k = \frac{WW' - (p^{r/2} - 1)^2}{p^r}. \tag{8}$$

Using this value of k a search is made for that value of t which will yield the desired order N Farey fraction.

e.g., Consider the conversion of ·3423 in $H(5, 4)$. Let its rational form be equal to b/a. Pseudo weight of ·3423 is 127/24. The pseudo weight of the reciprocal (·2204) of ·3423 is $-212/24$.

Hence

$$k = \frac{-212 \times 127 - 576}{625} = -44 \tag{9}$$

The following diophantine equation has to be solved.

$$127a - 24b = -44 \times 625. \tag{10}$$

The solution of

$$127a - 24b = 1 \tag{11}$$

is given by the last but one convergent of the continued fraction expansion of 127/24.

$$127/24 = 5 + \cfrac{1}{3 + \cfrac{1}{2 + \frac{1}{3}}}. \tag{12}$$

The convergents are

$$5,\quad 16/3,\quad 37/7,\quad 127/24.$$

Thus $b_c/a_0 = 37/7$.

(10) can now be written as

$$127\,a - 24\,(b - 44 \times 26) = -44. \tag{13}$$

Hence the solution of (10) is

$$a = -44 \times 7 = -308$$
$$b - 44 \times 26 = -44 \times 37 \quad \text{or} \quad b = -484 \tag{14}$$

or

$$a = 7,\quad b = 11 \text{ or the fraction is } 11/7.$$

6.3. *Direct table-look-up*

The third procedure for converting or decoding the $H(p, r)$ codes is by a direct high-speed table look-up. Here one stores the rational numbers as an ordered pair corresponding to each one of the possible combinations of the $H(p, r)$ codes, which are ordered lexicographically.

7. Applications to Exact Matrix Computations

We will now illustrate the application of $H(p, r)$ codes for exact computation. For the sake of illustration we consider the Gaussian elimination procedure (Young and Gregory 1973) to obtain the solution of the following linear system.

$$\begin{bmatrix} 3 & 1 & 3 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 16 \\ 8 \\ 12 \end{bmatrix} \tag{15}$$

The entries of the augmented matrix are converted into p-adic form. Then the elimination and back substitution procedures are done exactly as in real arithmetic and finally the p-adic entries of the solution vector are converted into rational number system.

Let $p = 11$ and $r = 4$.

Then the system is

$$\begin{bmatrix} \cdot 3000 & \cdot 1000 & \cdot 3000 \\ \cdot 1000 & \cdot 3000 & \cdot 1000 \\ \cdot 1000 & \cdot 1000 & \cdot 3000 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \cdot 5100 \\ \cdot 8000 \\ \cdot 1100 \end{bmatrix} \tag{16}$$

The eliminated augmented matrix is

$$\begin{bmatrix} \cdot 1000 & \cdot 4737 & \cdot 1000 & \cdot 9737 \\ \cdot 0000 & \cdot 1000 & \cdot 0000 & \cdot 1000 \\ \cdot 0000 & \cdot 0000 & \cdot 1000 & \cdot 3000 \end{bmatrix} \tag{17}$$

Then back substitution yields

$$x_3 = \cdot 3000$$

$$x_2 = \cdot 1000$$

and

$$x_1 = \cdot 9737 - x_3 - .4737\, x_2 \tag{18}$$

$$= \cdot 2000.$$

Hence the solution vector is

$$\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

We now present the solution of a larger system of highly ill-conditioned linear equations using Gaussian Elimination in a digital computer, in which the p-adic arithmetic has been implemented.

e.g., Solve $Ax = b$, where

$$A = \begin{bmatrix} 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 9 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 8 & 8 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 7 & 7 & 7 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 6 & 6 & 6 & 6 & 6 & 5 & 4 & 3 & 2 & 1 \\ 5 & 5 & 5 & 5 & 5 & 5 & 4 & 3 & 2 & 1 \\ 4 & 4 & 4 & 4 & 4 & 4 & 4 & 3 & 2 & 1 \\ 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 2 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$b^t = [1,\ 2,\ -5,\ 9,\ 15,\ 1,\ 6,\ 14,\ 3,\ 1]$$

Solution vector in 8209-adic system for $r = 8$:

x_1 =	8208	8208	8208	8208	8208	8208	8208	8208
x_2 =	8	0	0	0	0	0	0	0
x_3 =	8188	8208	8208	8208	8208	8208	8208	8208
x_4 =	8	0	0	0	0	0	0	0
x_5 =	20	0	0	0	0	0	0	0
x_6 =	8190	8208	8208	8208	8208	8208	8208	8208
x_7 =	8206	8208	8208	8208	8208	8208	8208	8208
x_8 =	19	0	0	0	0	0	0	0
x_9 =	8200	8208	8208	8208	8208	8208	8208	8208
x_{10} =	8208	8208	8208	8208	8208	8208	8208	8208

which corresponds to

$$x^T = [-1,\ 8,\ -21,\ 8,\ 20,\ -19,\ -3,\ 19,\ -9,\ -1].$$

It is evident, p-adic system can be used for implementing other algorithms for obtaining exact inverses of nonsingular matrices and Generalized Inverses of singular matrices with rational entries (Mahadeva Rao *et al.* 1975).

A complete FORTRAN program (for IBM 360 series) for p-adic arithmetic (for general p and r) and for solution of linear equations is available with the authors.

8. CONCLUDING REMARKS

The following problems remain to be solved before using p-adic arithmetic for all computations.

(i) Simplification of conversion of hard fractions to rationals.

(ii) Detection of sign and magnitude of $H(p, r)$ codes directly without explicit conversion;

(iii) Detection of overflow when the rationals exceed the range under consideration.

We believe that these problems will be solved in the near future.

ACKNOWLEDGEMENT

Two of the authors are thankful to CSIR, New Delhi, for providing research fellowships.

REFERENCES

Abramovitz, M. and Stegun, I. A., *Handbook of mathematical functions*, Dover Publications, Inc., New York (1965).

Bachman, G., *Introduction to* p-*adic numbers and valuation theory*, Academic Press, New York (1964).

Beiler, A. H., *Recreations in the theory of numbers—The queen of mathematics entertains*, Dover Publications, Inc., New York (1964).

Gelfond, A. O., *The solution of equations in integers*, W. H. Freeman & Company, London (1961).

Hardy, G. H. and Wright, E. M., *An introduction to theory of numbers*, Oxford, at the Clarendon Press (1960).

Khinchin, A. Ya., *Continued fractions*, The University of Chicago Press (1964).

Knuth, D. E., *The art of computer programming*, 2: *Semi numerical algorithms*, Addison Weslly Reading, Mass. (1969).

Mahadeva Rao, T., Subramanian, K. and Krishnamurthy, E. V., *SIAM J. Numerical Analysis*, (1975) (to appear).

Mordell, L. J., *Diophanline equations*, Academic Press. London (1969).

Niven, I. and Zuckerman, H. S., *An introduction to the theory of numbers*, John Wiley & Sons, Inc., New York (1966).

Richards, R. K., *Arithmetic operations in digital computers*, D. Van Nostrand Company, Inc. Princeton, New Jersey (1955).

Sankar, P. V., Chakrabarti, S. and Krishnamurthy, E. V., *IEEE Trans. Computers*, **22** 120–125, (1973 *a*)

Sankar, P. V., Chakrabarti, S. and Krishnamurthy, E. V., *IEEE Trans. Computers*, **22** 125–128, (1973 *b*).

Szabo, N. S. and Tanaka, R. I., *Residue arithmetic and its applications to computer technology*, McGraw-Hill Book Company, New York (1967).

Young, D. M. and Gregory, R. T., *A survey of numerical mathematics*, Addison Wesley. Reading, Mass, **2** (1973).