Canonical systems in the ring of integers

By B. KOVÁCS and A. PETHŐ (Debrecen)

Introduction

Let R be a ring, $\alpha \in R$, and $N_0 = \{0, 1, \ldots, n\}$. $\{\alpha, N_0\}$ is called a canonical number system (CNS) in R if every $\gamma \in R$ has a unique representation

$$\gamma = a_0 + a_1\alpha + \ldots + a_m\alpha^m, \quad a_i \in N_0, \quad a_i \neq 0. \tag{1}$$

If there exists a CNS in R, then R is called CNS ring.

Denote by $N = \{\beta_0, \beta_1, \ldots, \beta_s\}$ a subset of R and let α be an element of R. $\{\alpha, N\}$ is called a canonical system in R if every $\gamma \in R$ has a unique representation of the form (1), where $a_i \in N$. R is called a CS ring, if there exists a canonical system in R.

From the definitions it follows immediately that every CNS ring is a CS ring. The converse of this assertion is not true as we shall show in Theorem 1.

W. PENNEY determined all integers α, such that $\{\alpha, N\}$ is a CNS in Z with $N_0 = \{0, 1, \ldots, |\alpha| - 1\}$. His result was extended to the ring of Gaussian integers by I. KÁTAI and J. SZABÓ [2]. The same question has been answered for all quadratic number fields by B. KOVÁCS and I. KÁTAI [3], [4]. The first author has given a necessary and sufficient condition for the ring of integers of an algebraic number field to have CNS [5]. In [6] he characterized all CNS rings.

The concept of CS rings has been introduced by B. KOVÁCS [7]. There he characterized the structure of the additive group of such rings.

The question of determining all the CS in some rings seems to be difficult. We were not able to solve this problem even in the simplest case, when the ring in question is Z. We shall describe an algorithm after Theorem 4, by using it we can decide if $\{\alpha, N\}$, $\alpha \in Z$, $N \subseteq Z$ is a CS or not. In Theorem 5 we prove that for every $\alpha \in Z$ $\alpha < -2$ there exist infinitely many $N \subseteq Z$, such that $\{\alpha, N\}$ is a canonical system in Z.

2. Connection between CNS and CS rings

As we have pointed out in the introduction every CNS ring is a CS ring. The converse of this assertion is not true, as we shall show by an example.

Theorem 1. *Let $n \geqq 2$ be a natural number, and denote by $Z_{n\times n}$ the ring of $n \times n$ quadratic matrices over Z. Then $Z_{n\times n}$ is a* CS, *but not a* CNS *ring.*

PROOF. First we prove that $Z_{n\times n}$ is not a CNS ring. Suppose the contrary, i.e. that there exists a matrix $A\in Z_{n\times n}$, and $N_0=\{0, 1, \ldots, k\}$, such that $\{A, N_0\}$ is a CNS in $Z_{n\times n}$. Then every $B\in Z_{n\times n}$ can be written in the form

$$B=\sum_{i=0}^{s} b_i A^i, \quad b_i\in N_0, \quad b_s\neq 0. \tag{2}$$

Let $Q_{n\times n}$ be the ring of the $n\times n$ quadratic matrices over Q — the field of the rational numbers — and denote by $L(A)$ the subspace of $Q_{n\times n}$ generated by $E, A, A^2, \ldots$. Then the dimension of $L(A)$ is at most n. This is obvious because the degree of the characteristic polynomial $p(\lambda)$ of A is n, $p(\lambda)$ has integer coefficients and because by the theorem of Cayley and Hamilton $p(A)=0$ holds.

If $B\in Z_{n\times n}$ has a representation in the form (2) then B is contained in $L(A)$. But $Z_{n\times n}$ is not contained in any subspace of $Q_{n\times n}$ of dimension less than n^2. Further $n^2>n$, thus there exists a $C\in Z_{n\times n}\setminus L(A)$, which does not have a representation (2). Consequently $Z_{n\times n}$ is not a CNS ring.

To complete the proof, we construct a canonical system in $Z_{n\times n}$. Let $k\geqq 2$ be a natural number and $-kI=K$, where I denotes the identity matrix of $Z_{n\times n}$. Let N be the set of all matrices $A_i=(a_{l,m}^{(i)})$, with $0\leqq a_{l,m}^{(i)}<k$ for any $1\leqq l, m\leqq n$. We show that $\{K, N\}$ is a canonical system in $Z_{n\times n}$.

Let $B=(b_{r,s})$ be an arbitrary element of $Z_{n\times n}$. $b_{r,s}$ may be written uniquely in the form

$$b_{r,s}=c_{r,s}^{(0)}+c_{r,s}^{(1)}(-k)+\ldots+c_{r,s}^{t_{r,s}}(-k)^{t_{r,s}}, \quad \text{with} \quad 0\leqq c_{r,s}^{(j)}<k, \quad j=0, \ldots, t_{r,s}, \tag{3}$$

since $\{-k, \{0, 1, \ldots, k-1\}\}$ is a CNS in Z ([1]). Put $T=\max\limits_{1\leqq r,s\leqq n} t_{r,s}$ and $C_j=(c_{r,s}^{(j)})$, with $c_{r,s}^{(j)}=0$ for $T\geqq j>t_{r,s}$. Then $c_j\in N$ and

$$B=C_0+C_1K+\ldots+C_TK^T \tag{4}$$

can be easily seen. $Z_{n\times n}\cdot K$ is a normal subgroup in the additive group of $Z_{n\times n}$, further N is a complete residue system of $Z_{n\times n}/Z_{n\times n}\cdot K$. From this follows the uniqueness of the representation of B in the form (4). This completes our proof.

We need the following lemma.

Lemma 1. *Let R be a ring with at least two elements. If $\{\alpha, N\}$ is a* CS *in R, then $0\in N$ and α is not a right zero divisor in R.*

PROOF. Let $\{\alpha, N\}$ be a canonical system in R and assume $0\notin N$. Let

$$0=a_0+a_1\alpha+\ldots+a_m\alpha^m, \quad a_i\in\mathcal{N}, \quad a_m\neq 0$$

be the representation of 0. By the assumption there exists a nonzero element x of R. Let us write x in the form

$$x=b_0+b_1\alpha+\ldots+b_s\alpha^s, \quad b_i\in N, \quad b_s\neq 0.$$

Then x has another representation too, namely

$$x=b_0+b_1\alpha+\ldots+b_s\alpha^s+a_0\alpha^{s+1}+\ldots+a_m\alpha^{m+s+1}$$

in contradiction with the uniqueness of the representation.

Further assume $\{\alpha, N\}$ to be a CS in R with α a right zero divisor. Let $0 \neq x \in R$ be an element with $x\alpha = 0$. Write $x = a_0 + a_1\alpha + \ldots + a_m\alpha^m$ with $a_i \in N$, $a_m \neq 0$. Then

$$0 = x\alpha = a_0\alpha + a_1\alpha^2 + \ldots + a_m\alpha^{m+1}$$

is a representation of 0. But by the first assertion $0 \in N$. Thus 0 has at least two distinct representations, contrary to the uniqueness.

As we mentioned in the introduction every CNS ring is a CS ring. By Theorem 1 $Z_{n\times n}$ is such a CS ring which is not a CNS ring. We find the following question interesting: Do there exist in a CNS ring such canonical systems, which are not CNS? In the following we shall prove some results concerning this problem.

Theorem 2. *Let* $\{\alpha, N_0\}$ *be a* CNS *and* ε *be a unit in the ring* R. *Then* $\{\alpha, \varepsilon \cdot N_0\}$ *is a* CS *in* R.

PROOF. According to [6] every CNS ring is commutative. Let $0 \neq \gamma \in R$. Write

$$\gamma \cdot \varepsilon^{-1} = a_0 + a_1\alpha + \ldots + a_m\alpha^m, \quad a_i \in \mathcal{N}_0, \ a_m \neq 0.$$

Then

$$\gamma = (a_0 + a_1\alpha + \ldots + a_m\alpha^m)\varepsilon = (a_0\varepsilon) + (a_1\varepsilon)\alpha + \ldots + (a_m\varepsilon)\alpha^m.$$

The theorem is proved.

3. Canonical systems in Z

If $\{\alpha, N\}$ is a CS in Z, then $|\alpha| \geqq 2$, as one can easily show. Further N must form a complete residue system mod α and by Lemma 1 $0 \in N$. In the sequel N will denote a subset of Z with these properties.

Theorem 3. *Let* $\alpha \in Z$, $N \subseteqq Z$. *Put* $K = \max\limits_{b_j \in N} |b_j|$. $\{\alpha, N\}$ *is a* CS *in* Z *if and only if any* $0 \neq \gamma \in Z$, *with* $|\gamma| \leqq \dfrac{K}{|\alpha| - 1}$ *can be written in the form*

$$\gamma = a_0 + a_1\alpha + \ldots + a_k \cdot \alpha^k, \quad a_i \in \mathcal{N}, \quad a_k \neq 0. \tag{5}$$

PROOF. Let $m \in Z$ and

$$m = a_0 + m_1\alpha, \ m_1 = a_1 + m_2\alpha, \ \ldots \tag{6}$$

with $a_i \in N$. It is clear that m does not have any representation (5) if and only if all m_i of the sequence (6) do not.

Assume there exists an integer which can not be written in the form (5). Let m denote one of the — in absolute value — smallest integers with this property. Then

$$|m| \leqq |m_1| = \left|\frac{m - a_0}{\alpha}\right| \leqq \frac{|m|}{|\alpha|} + \frac{K}{|\alpha|}.$$

Thus $|m| \leqq \dfrac{K}{|\alpha| - 1}$. Therefore if $\{\alpha, N\}$ is not a CS in Z, then there exists an $m \in Z$, with $|m| \leqq \dfrac{K}{|\alpha| - 1}$ which has not any representation (5).

If all integers $|m| \leqq \frac{K}{|\alpha|-1}$ have a representation (5), then all integers have and this is unique because N is a complete residue system mod α. This proves completely the assertion.

Remarks: 1. One can effectively decide of a given $\{\alpha, N\}$ whether it is a CS in Z in the following way: First examine whether N is a complete residue system in Z. Then test all integers with $|m| \leqq \frac{K}{|\alpha|-1}$, whether they have a representation (5). This can be done in finitely many steps. Of course start with an m, from the sequence (6). If $|m_i| \leqq \frac{K}{|\alpha|-1}$ for an i, then

$$|m_{i+1}| = \left|\frac{m_i - a_i}{\alpha}\right| \leqq \frac{|m_i|}{|\alpha|} + \frac{K}{|\alpha|} \leqq \frac{K}{|\alpha|-1}.$$

Thus all members of (6) lie in absolute value below $\frac{K}{|\alpha|-1}$. If m has a representation (5), then this procedure breaks off in at most $\frac{2K}{|\alpha|-1}$ steps. Otherwise there exists an m', with $|m'| \leqq \frac{K}{|\alpha|-1}$ and

(7) $$m' = a_0 + a_1\alpha + \ldots + m'\alpha^k \quad \text{for some} \quad k \geqq 1 \quad \text{and} \quad a_i \in N.$$

With the described procedure one can decide in $0(K^2)$ steps whether $\{\alpha, N\}$ is a CS in Z.

2. From (7) and from Theorem 3 follows the condition: $\{\alpha, N\}$ is a CS in Z if and only if, there do not exist integers r with $0 \leqq r \leqq \frac{2K}{|\alpha|-1}$ such that the congruence

$$x_1 + x_2\alpha + \ldots + x_r\alpha^r \equiv 0 \pmod{1-\alpha^{r+1}}$$

is solvable in integers $x_1, x_2, \ldots, x_r \in N$, $x_r \neq 0$.

To prove one further general result on canonical systems in Z we need the following lemma:

Lemma 2. *Let the real number β be a root of the polynomial with real coefficients $P(x) = b_0 + b_1 x + \ldots + b_s x^s$. Let $h(P) = h = \max\{|b_0|, \ldots, |b_s|\}$. Then*

$$|\beta| \leqq \frac{h}{|b_s|} + 1.$$

PROOF. See for example [8] page 5.

Theorem 4. *Let $\{\alpha, N\}$ CS in Z. Put $K = \max_{b \in N} |b|$ and $k = \min_{0 \neq b \in N} |b|$. Then either $|K| < |\alpha|$ or $K/k \geqq |\alpha| - 1$.*

PROOF. Let $N=\{0, b_1, \ldots, b_{|\alpha|-1}\}$. For simplicity assume $K=|b_1|$ and $k=|b_2|$. Further assume $K\geqq|\alpha|$. N is a complete residue system mod α and $0\in N$, thus $b_1\not\equiv 0 \pmod{\alpha}$ and $K>|\alpha|$. So there exists an integer $0\neq\gamma$, with $|\gamma|<|\alpha|$, $\gamma\equiv b_1 \pmod{\alpha}$ and γ and b_1 have the same sign. Write

$$\gamma = b_1+c_1\alpha+\ldots+c_n\alpha^n, \quad c_i\in N, \quad c_n\neq 0.$$

Thus α is a real root of the polynomial $b_1-\gamma+c_1x+\ldots+c_1x^n=P(x)$. By the assumption $h(P)=\max\{|b_1-\gamma|, |c_1|, \ldots, |c_n|\}\leqq K$ and $|c_n|\geqq k$. From Lemma 2 $|\alpha|\leqq\frac{K}{|c_n|}+1\leqq\frac{K}{k}+1$ follows which proves the theorem.

4. Infinite families of canonical systems in Z

In 3. we have proved general theorems on CS in Z. The following question remains open: How many sets N exists for a given integer α such that $\{\alpha, N\}$ is a CS in Z. We shall prove in this section, that if $\alpha<-2$, then there exist infinitely many. Indeed, the following theorems are true:

Theorem 5. *Let* $\alpha<-2$ *and* $N_b=\{0, 1, \ldots, t-1, b, t+1, \ldots, |\alpha|-1\}$ *with integers* b, u *where* $b=t-u\alpha^k$, $2\leqq t\leqq|\alpha|-1$, $k\geqq 1$, $0\leqq u<|\alpha|-1$, $u\neq t$. *Then* $\{\alpha, N_b\}$ *is a* CS *in* Z.

Theorem 6. *Let* $\alpha<-2$ *and* $N_a=\{0, a, 2, \ldots, |\alpha|-1\}$ a, v *integers with* $a=1-v\alpha^k$, $k\geqq 1$, $0\leqq v\leqq|\alpha|-1$, $v\neq 1$. *Then* $\{\alpha, N_a\}$ *is a* CS *in* Z.

PROOF OF THEOREM 5. Let $N=\{0, 1, \ldots, |\alpha|-1\}$. It is well known that $\{\alpha, N\}$ is CNS in Z. If $u=0$ then $b=t$ and $N_b=N$. So we may assume $u>0$. In this case $|b|>|\alpha|$, therefore $|b|=\max\{x|x\in N_b\}$. By Theorem 3 it is enough to show that all integers m, with $|m|\leqq\frac{|b|}{|\alpha|-1}=\frac{|t-u\alpha^k|}{|\alpha|-1}$ are representable by $\{\alpha, N_b\}$. Put

$$n = \sum_{i=0}^{s} a_i\alpha_i \quad \text{with} \quad a_i\in N, \quad a_s\neq 0. \tag{8}$$

Then

$$|n| = \left|\sum_{i=0}^{[s/2]} a_{2i}\alpha^{2i} - \sum_{i=1}^{[s/2]-\varepsilon} a_{2i-1}\alpha^{2i-1}\right|,$$

where $\varepsilon=0$ or 1 according as s is even or odd, and $[x]$ denotes the integer part of x. All coefficients in (8) are non-negative, thus

$$|n| \geqq |\alpha|^s-|\alpha|(|\alpha|-1)(1+|\alpha|^2+\ldots+|\alpha|^{s-2}) = \frac{|\alpha|^s+|\alpha|}{|\alpha|+1} \tag{9}$$

if s is even and

$$|n| \geqq |\alpha|^s-(|\alpha|-1)(1+|\alpha|^2+\ldots+|\alpha|^{s-1}) = \frac{|\alpha|^s+1}{|\alpha|+1} \tag{10}$$

if s is odd. Consequently

$$|n| \geqq \frac{|\alpha|^s+1}{|\alpha|+1}$$

holds in both cases.

In the sequel $L(n)$ will denote the length of n, i.e. $L(n)=v$ if $n=\sum_{i=0}^{v} n_i\alpha^i$, with $n_i\in N$, $n_v\neq 0$.

Now we shall show that if $|m|\leqq\frac{|t-u\alpha^k|}{|\alpha|-1}$, then $L(m)\leqq k$. Of course, let k be even, then

$$\frac{|t-u\alpha^k|}{|\alpha|-1} < \frac{u|\alpha|^k}{|\alpha|-1} < \frac{|\alpha|^{k+1}+1}{|\alpha|+1} \tag{11}$$

holds because of $u<|\alpha|-1$, and $|\alpha|>2$. Similarly if k is odd, then

$$\frac{|t-u\alpha^k|}{|\alpha|-1} < \frac{u|\alpha|^k+|\alpha|}{|\alpha|-1} < \frac{|\alpha|^{k+1}+|\alpha|}{|\alpha|+1}. \tag{12}$$

Comparing (11) and (12) with (9) and (10) we see that the smallest integer with length $k+1$ is in absolute value greater then $\frac{|t-u\alpha^k|}{|\alpha|-1}$. Thus by Theorem 3 we must establish the representability of integers with length at most k by $\{\alpha, N_b\}$.

In the sequel let $|m|\leqq\frac{|t-u\alpha^k|}{|\alpha|-1}$,

$$m = \sum_{i=0}^{v} m_i\alpha^i, \quad m_i\in N, \quad m_v \neq 0. \tag{13}$$

Case 1. Let $L(m)<k$. If $m_j\neq t$ for $j=0, 1, \ldots, v$, then all $m_j\in N_b$, and we have nothing to do. Otherwise let $j\geqq 0$ denote the smallest index with $m_j=t$; then

$$m = \sum_{i=0}^{v} m_i\alpha^i = \sum_{i=0}^{j-1} m_i\alpha^i + b\alpha^j + \sum_{i=j+1}^{v} m_i\alpha^i + u\alpha^{j+k}.$$

m_i, $i=0, 1, \ldots, j-1$, b and u belong to N_b, furthermore $j+k>v$. Changing all coefficients of (13) which are equal to t to $b+u\alpha^k$ we receive the representation of m by $\{\alpha, N_b\}$.

Case 2. Let $L(m)=k$, $m_0\neq t$. Take $M=\frac{m-m_0}{\alpha}$. Then $L(M)<k$, thus by Case 1 M is representable by $\{\alpha, N_b\}$, consequently m too.

Case 3. Let $L(M)=k$, $m_0=t$, $m_k<|\alpha|-u$. Replace m_0 by $b+u\alpha^k$, then we have $m_0^{(1)}=b$, $m_i^{(1)}=m_i$, $i=1, 2, \ldots, k-1$, $m_k^{(1)}=m_k+u<|\alpha|$. Here and in the sequel $m_j^{(n)}$ will denote the j-th coefficient in the representation of m after the j-th substitution. $m_j^{(1)}\in N$ for $j=1, 2, \ldots, k$. Take $M=\frac{m-b}{\alpha}$, then $L(M)<k$. M fulfills the assumption of Case 1, therefore it is and so m too is representable by $\{\alpha, N_b\}$.

Case 4. $L(M)=k$, $m_0=t$, $m_k\geqq|\alpha|-u$. Replacing m_0 by $b+u\alpha^k$ we have $m_0^{(1)}=b$, $m_i^{(1)}=m_i$, $i=1, 2, \ldots, k-1$, $m_k^{(1)}=m_k+u\geqq|\alpha|$. Let $m_k+u=w+(-1)\alpha$ and take $m_i^{(2)}=m_i^{(1)}$, $i=0, 1, \ldots, k-1$, $m_k^{(2)}=w\in N$, $m_{k+1}^{(2)}=-1$. We distinguish two cases:

a) $m_1^{(2)}=m_1=t$. Then replacing $m_1^{(2)}$ by $b+u\alpha^k$ we have

$$m_0^{(3)}=m_0^{(2)},\quad m_1^{(3)}=b,\quad m_i^{(3)}=m_i^{(2)},\quad i=2,3,\ldots,k,\quad m_{k+1}^{(3)}=u-1\in N.$$

a) $m_1^{(2)}=m_1=t$. Then replacing $m_1^{(2)}$ by $b+u\alpha^k$ we have

$$m_0^{(3)}=m_0^{(2)},\quad m_1^{(3)}=b,\quad m_i^{(3)}=m_i^{(2)},\quad i=2,3,\ldots,k,\quad m_{k+1}^{(3)}=u-1\in N.$$

Take $M=\dfrac{m-m_0^{(3)}-m_1^{(3)}\alpha}{\alpha^2}$, then $M=m_2^{(3)}+\ldots+m_{k+1}^{(3)}\alpha^{k-1}$, i.e. $L(M)<k$, and we get again case 1.

b) $m_1^{(2)}\neq t$. Write $-1=(|\alpha|-1)+\alpha$, and $m_i^{(3)}=m_i^{(2)}$, $i=0,1,\ldots,k$, $m_{k+1}^{(3)}=|\alpha|-1\in N$, $m_{k+2}^{(3)}=1\in N$. Put

$$m_2^{(4)}=\begin{cases}m_2^{(3)}, & \text{if } m_2^{(3)}\neq t\\ b, & \text{if } m_2^{(3)}=t\end{cases}$$

and $m_i^{(4)}=m_i^{(1)}$ for $i=0,1,3,\ldots,k+1$. Then

$$m_{k+2}^{(4)}=\begin{cases}1, & \text{if } m_2^{(3)}\neq t\\ u+1, & \text{if } m_2^{(3)}=t\end{cases}$$

$m_{k+2}^{(4)}\in N$ follows from the assumption $u<|\alpha|-1$. In both cases the length of $M=\dfrac{m-m_0^{(4)}+m_1^{(4)}\alpha+m_2^{(4)}\alpha^2}{\alpha^3}$ is less than k. We can apply again case 1, and this proves the theorem.

PROOF OF THEOREM 6. Now one must check the representability of integers with length at most $k+1$ by $\{\alpha, N_a\}$. This may be done with the same method as in the proof of Theorem 5.

References

[1] W. PENNEY, A "binary" system for complex numbers, J. ACM **12** (1965) 247—248.
[2] I. KÁTAI and J. SZABÓ, Canonical number systems for complex integers, *Acta Sci. Math. (Szeged)* **37** (1975), 255—260.
[3] I. KÁTAI and B. KOVÁCS, Canonical number systems in imaginary quadratic fields, *Acta Math. Acad. Sci. Hungar.*, **37** (1981), 159—164.
[4] I. KÁTAI und B. KOVÁCS, Kanonische Zahlensysteme in der Theorie der quadratischen algebraischen Zahlen, *Acta Sci. Math. Szeged*, **42** (1980), 99—107.
[5] B. KOVÁCS, Canonical number systems in algebraic number fields, *Acta Math. Sci. Hungar.*, **37** (1981), 405—407.
[6] B. KOVÁCS, CNS rings, *in print*.
[7] B. KOVÁCS, CNS rings, *Coll. Math. Soc. János Bolyai Budapest*, 1981 *(to appear)*.
[8] TH. SCHNEIDER, Einführung in die transzendenten Zahlen, *Springer Verlag Berlin*, 1957.

(Received March 22, 1980)