To appear, Theoretical Computer Science, Nov. 1992

The Ring of k-Regular Sequences

Jean-Paul Allouche *
C. N. R. S. (U. R. A. 226)
Mathématiques et Informatique
33405 Talence Cedex
France
allouche%frbdx11.bitnet

Jeffrey Shallit §
Dept. of Computer Science
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada
shallit@watdragon.waterloo.edu

Abstract.

The *automatic sequence* is the central concept at the intersection of formal language theory and number theory. It was introduced by Cobham, and has been extensively studied by Christol, Kamae, Mendès France and Rauzy, and other writers. Since the range of automatic sequences is finite, however, their descriptive power is severely limited.

In this paper, we generalize the concept of automatic sequence to the case where the sequence can take its values in a (possibly infinite) ring R; we call such sequences *k–regular*. (When R is finite, we obtain automatic sequences as a special case.) We argue that k–regular sequences provide a good framework for discussing many "naturally–occurring" sequences, and we support this contention by exhibiting many examples of k–regular sequences from numerical analysis, topology, number theory, combinatorics, analysis of algorithms, and the theory of fractals.

We investigate the closure properties of k–regular sequences. We prove that the set of k–regular sequences forms a ring under the operations of term-by-term addition and convolution. Hence the set of associated formal power series in $R[[X]]$ also forms a ring.

We show how k–regular sequences are related to $\mathbb{Z}$–rational formal series. We give a machine model for the k–regular sequences. We prove that all k–regular sequences can be computed quickly.

Let the *pattern sequence* $e_P(n)$ count the number of occurrences of the pattern P in the base–k expansion of n. Morton and Mourant showed that every sequence over $\mathbb{Z}$ has a unique expansion as a sum of pattern sequences. We prove that this "Fourier" expansion maps k–regular sequences to k–regular sequences. (This can be viewed as a generalization of results of Choffrut and Schützenberger, and previous results of Allouche, Morton, and Shallit.) In particular, the coefficients in the expansion of $e_P(an+b)$ form a k–automatic sequence.

Many natural examples and some open problems are given.

* Research supported in part by "PICS: Théorie des nombres et ordinateurs"
§ Research supported in part by NSF Grant CCR-8817400, the Wisconsin Alumni Research Foundation, and a Walter Burke award from Dartmouth College.

I. Introduction.

Let $\{S(n)\}_{n\geq 0}$ be a sequence with values chosen from a finite set Σ. Then $\{S(n)\}_{n\geq 0}$ is said to be k–automatic if, informally speaking, $S(n)$ is a finite-state function of the base–k expansion of n.

Automatic sequences have been studied by Cobham [Cob], Christol, Kamae, Mendès France and Rauzy [CKMR], and others. (For example, see [DMFP], [Mau], and the survey paper of Allouche [A1].) There are many other ways to characterize automatic sequences. For example, consider the following

Definition 1.1.

The *k–kernel* of a sequence is the set of all subsequences of the form $\{S(k^e n + a)\}_{n\geq 0}$, where $e \geq 0$ and $0 \leq a < k^e$.

Cobham [Cob] proved the following

Theorem 1.2.

A sequence is k–automatic if and only if its k–kernel is finite.

Unfortunately, the range of automatic sequences is necessarily finite, and this restricts their descriptive power.

In this paper, we are concerned with a natural generalization of automaticity to the case where the sequence $\{S(n)\}_{n\geq 0}$ takes its values in a (possibly infinite) ring; we call such sequences k–regular. (Another generalization of automatic sequences was already given by Allouche [A4].) We use an analogue of Theorem 1.2 as our *definition*. We show that k–regular sequences provide an excellent framework for describing many "naturally occurring" sequences, such as the numerators of the left endpoints of the Cantor set, the sequence $\{\nu_p(n!)\}_{n\geq 0}$, which counts the number of times a prime p divides a factorial, binary Gray code, numerators of entries of the Stern-Brocot tree, multiplicative-cost addition chains, etc.

We prove that k–regular sequences have nice closure properties. By associating a formal power series with each sequence, we prove that the set of k–regular sequences forms a ring, but not a field, under the usual power series operations.

We explore the connection with a machine model of Schützenberger [Sch], which includes finite automata with counters as a special case. This allows us to prove that the n-th term of a k–regular sequence can be computed in time polynomial in $\log n$.

We introduce the *pattern sequences* $e_P(n)$, which count the number of occurrences of the string P in the base–k expansion of n. Morton and Mourant [MM] showed that every sequence $\{S(n)\}_{n\geq 0}$ over $\mathbb{Z}$ has a unique expansion as a sum of pattern sequences. In analogy with the Fourier transform, we call this sequence of coefficients $\{\hat{S}(n)\}_{n\geq 0}$ the *pattern transform* of $\{S(n)\}_{n\geq 0}$. We show that a sequence is k–regular if and only if its pattern transform is k–regular. This can be viewed as a generalization of results of Choffrut and Schützenberger [CS] and previous results of the authors and P. Morton [AMS].

Finally, we give many examples and some open problems.

II. k–regular sequences: definition and properties.

Let R' be a commutative Noetherian ring, i. e. a ring in which every ideal is finitely generated. (Examples of such rings include all finite rings, $\mathbb{Z}$, all fields K, and the polynomial rings $K[X]$.) Let R be a ring containing R'.

Let $\mathcal{S}(R)$ denote the set of sequences with values in R. Let $\{S(n)\}_{n\geq 0}$ be a sequence with values in R, and let k be an integer ≥ 2.

Definition 2.1.

We say $\{S(n)\}_{n\geq 0}$ is (R',k)*–regular* if there exist a finite number of sequences $S_1, S_2, \ldots, S_j$ with values in R, such that each sequence in the k–kernel of $\{S(n)\}_{n\geq 0}$ is an R'–linear combination of the S_i.

Let $\mathcal{K}$ denote the k–kernel of $\{S(n)\}_{n\geq 0}$. Then $\{S(n)\}_{n\geq 0}$ is (R',k)–regular means that

$$\langle \mathcal{K} \rangle \subseteq \langle S_1, S_2, \ldots S_n \rangle,$$

i. e. $\langle \mathcal{K} \rangle$ is a sub-module of a finitely generated R'–module. By a well-known theorem (see, e. g., [Lan, pp. 142–144]), it follows that $\langle \mathcal{K} \rangle$ itself is finitely generated.

Thus Definition 2.1 can be restated as follows: a sequence $\{S(n)\}_{n\geq 0}$ with values in R is (R',k)–regular if the R'–module generated by its k–kernel is a *finitely generated* R'–submodule of $\mathcal{S}(R)$.

If the context is clear, we usually write just k–regular.

Note that if R' is a finite ring, then we recover the case of k–automatic sequences. For if every subsequence in the k–kernel can be written as an R'–linear combination of a finite set of sequences, then there are only a finite number of distinct elements of the k–kernel. In fact, the same holds for sequences that take on only finitely many values (see Theorem 2.3 below).

The reader may now wish to look at Section VII for some examples of k–regular sequences.

Our first theorem gives several alternative characterizations of k–regular sequences:

Theorem 2.2.

The following are equivalent:

(a) $\{S(n)\}_{n\geq 0}$ *is* (R',k)*–regular;*

(b) The R'*–module generated by the* k*–kernel of* $\{S(n)\}_{n\geq 0}$ *is generated by a finite number of its subsequences of the form* $S(k^{f_i}n + b_i)$ *where* $0 \leq b_i < k^{f_i}$*;*

(c) There exists an integer E *such that for all* $e_j > E$*, each subsequence* $S(k^{e_j}n + a_j)$ *with* $0 \leq a_j < k^{e_j}$ *can be expressed as an* R'*–linear combination*

$$S(k^{e_j}n + a_j) = \sum_i c_{ij} S(k^{f_{ij}}n + b_{ij}),$$

where $f_{ij} \leq E$ *and* $0 \leq b_{ij} < k^{f_{ij}}$*;*

(d) There exist an integer r *and* r *sequences* $S = S_1, S_2, \ldots, S_r$*, such that for* $1 \leq i \leq r$*, the* k *sequences* $\{S_i(kn + a)\}_{n\geq 0}$*,* $0 \leq a < k$*, are* R'*–linear combinations of the* S_i*;*

(e) There exist an integer r, r sequences $S = S_1, S_2, \ldots, S_r$, and k matrices $B_0, B_1, \ldots, B_{k-1}$ in $M_{r,r}(R')$ such that if

$$V(n) = \begin{pmatrix} S_1(n) \\ \vdots \\ S_r(n) \end{pmatrix},$$

one has $V(kn+a) = B_a V(n)$ for $0 \le a < k$.

Proof.

(a) $\Rightarrow$ (b): Let $\mathcal{K}$ denote the k–kernel of $S(n)$. Then $\langle \mathcal{K} \rangle$, the module generated by $\mathcal{K}$, is finitely generated, so there exist sequences $S_1, S_2, \ldots, S_k$ such that

$$\langle \mathcal{K} \rangle = \langle S_1, S_2, \ldots, S_k \rangle.$$

But then each S_i is necessarily a finite linear combination of elements from $\mathcal{K}$, and there are only finitely many S_i, so $\langle \mathcal{K} \rangle$ is generated by only finitely many members of $\mathcal{K}$.

(b) $\Rightarrow$ (c): Let the k–kernel of $\{S(n)\}_{n \ge 0}$ be generated by a finite set of its subsequences of the specified form, say

$$S(k^{f_i} n + b_i)$$

for $1 \le i \le i'$. Let $E = \max_{1 \le i \le i'} f_i$. Then for all $e_j > E$, we can write

$$S(k^{e_j} n + a_j) = \sum_i c_{ij} S(k^{f_{ij}} n + b_{ij}),$$

where $f_{ij} \le E$ and $0 \le b_{ij} < k^{f_{ij}}$.

(c) $\Rightarrow$ (d): Take as the r sequences the set $\mathcal{K}$ of subsequences $S_i(n) = S(k^{f_i} n + b_i)$ with $0 \le f_i \le E$ and $0 \le b_i < k^{f_i}$. Then

$$S_i(kn+a) = S(k^{f_i}(kn+a) + b_i) = S(k^{f_i+1} n + ak^{f_i} + b_i),$$

which, if $f_i + 1 \le E$, is an element of $\mathcal{K}$, and if $f_i + 1 > E$, is a linear combination of elements of $\mathcal{K}$.

(d) $\Rightarrow$ (e): Follows trivially.

(e) $\Rightarrow$ (a): We need to see that $S(k^e n + a)$ is a linear combination of the S_i. Express a in base k (possibly with leading zeroes) as

$$\sum_{0 \le i < e} a_i k^i;$$

then it is easy to see that

$$V(k^e n + a) = B_{a_0} B_{a_1} \cdots B_{a_{e-1}} V(n),$$

and this expresses $S(k^e n + a)$ as a linear combination of the S_i. ∎

Remarks.

- Note that in parts (d) and (e) of the theorem, the sequences S_i can be taken to be in the k–kernel of S.
- Part (e) of the theorem gives a substitution-like definition, which can be compared to the linear k–substitutions of Liardet [Li], which generate exactly the k–automatic sequences.
- The *dimension* of the R'–module generated by the k–kernel is an invariant that may be interpreted as a measure of complexity of the sequence $\{S(n)\}_{n\geq 0}$.
- We note that every ultimately periodic sequence is (R', k)–regular for all R' and k.

Our next theorem illustrates a connection between k–regular sequences and k–automatic sequences:

Theorem 2.3.

A sequence is (R', k)–regular and takes on only finitely many values if and only if it is k–automatic.

Proof.

If a sequence is k–automatic, it is by definition finitely valued, and since its k–kernel is finite, it generates a finitely generated module.

Now suppose $S(n)$ is k–regular and takes on finitely many values. From Theorem 2.2 (e), there exist sequence $S = S_1, S_2, \ldots, S_r$ (which can be taken in the k–kernel of S) and matrices $B_0, B_1, \ldots B_{k-1}$ such that

$$V(n) = \begin{pmatrix} S_1(n) \\ S_2(n) \\ \vdots \\ S_r(n) \end{pmatrix}$$

satisfies $V(kn + a) = B_a V(n)$ for $0 \leq a < k$ and $n \geq 0$. Let $\mathcal{V}$ be the (finite) set of values of $\{V(n)\}_{n\geq 0}$, and define the k–homomorphism σ by $\sigma(v) = w_0 w_1 \cdots w_{k-1}$, where $v \in \mathcal{V}$ and $w_a = B_a v$ for $0 \leq a < k$. Then the infinite word

$$V(0)V(1)V(2)\cdots$$

is a fixed point of σ and $S_1(n)$ is an image of this fixed point. Hence $S(n)$ is k–automatic. ■

Corollary 2.4.

If $S(n)$ is $(\mathbb{Z}, k)$–regular, then for all $m \geq 1$, $\{S(n) \bmod m\}_{n\geq 0}$ is k-automatic.

Remark.

The converse does not hold. Let $S(n) = 2^n$ and use Theorem 2.11 below.

We now investigate the closure properties of k–regular sequences:

Theorem 2.5.

Let $\{S(n)\}_{n\geq 0}$ and $\{T(n)\}_{n\geq 0}$ be k–regular sequences. Then so are $S+T=\{S(n)+T(n)\}_{n\geq 0}$, $\alpha S=\{\alpha S(n)\}_{n\geq 0}$, and $ST=\{S(n)T(n)\}_{n\geq 0}$.

Proof.

Let $S_1=S, S_2, \ldots, S_r$ (respectively $T_1=T, T_2, \ldots, T_r$) be a system of generators for the module generated by the k–kernel of S (respectively T). Then it is easy to see that the $r+r'$ sequences $S_1, \ldots, S_r, T_1, \ldots, T_{r'}$ generate the module generated by the k–kernel of $S+T$. Similarly, the rr' sequences S_iT_j, $1\leq i\leq r$, $1\leq j\leq r'$ generate the module generated by the k–kernel of ST. Finally, the sequences αS_i, $1\leq i\leq r$, generate the module generated by the k–kernel of αS. ■

Remarks.

We observe that some simple transformations do *not* preserve k–regularity.

• Let $S(n), T(n)$ be $(\mathbb{Z}, k)$–regular sequences with $T(n)\neq 0$ for all n. Then the sequence $S/T=\{S(n)/T(n)\}_{n\geq 0}$ need not even be $(\mathbb{Q}, k)$–regular.

For example, define $T(2n)=n+1$, $T(2n+1)=T(n)+1$ for $n\geq 0$. Define $T_j(n)=T(2^jn+2^{j-1}-1)$. Then it is easy to see that $T_j(n)=n+j$ for $j\geq 1$.

Suppose $1/T(n)$ were $(\mathbb{Q}, 2)$–regular. Then the module generated by the sequences

$$1/T_1(n), 1/T_2(n), 1/T_3(n), \ldots$$

would have finite rank. Then for some $m\geq 1$, the $m\times m$ matrix M_{ij} defined by

$$M_{ij}=1/T_j(i-1)=1/(i+j-1),$$

$1\leq i, j\leq m$, would have determinant 0. But M_{ij} is a Hilbert matrix and is well-known to have nonzero determinant, a contradiction, and the conclusion follows.

• We note that k–regular sequences are not closed under absolute value (and hence not closed under max and min). Consider the function $f(n)=e_0(n)-e_1(n)$, where $e_0(n)$ counts the number of 0's in the binary expansion of n, and $e_1(n)$ counts the number of 1's in the binary expansion of n. It is easily verified that $e_0(n)$ and $e_1(n)$ are k–regular; hence so is $f(n)$. But $|f(n)|$ is not k–regular. For we have

$$f(2^jn)=|e_0(n)-e_1(n)+j|$$

for $n\geq 1$ and $j\geq 0$. Now suppose there were a linear dependency among these subsequences; i. e. there exist a, b such that

$$|n+a|=\sum_{a+1\leq i\leq b} c_i|x+i|$$

for all integers n. For $n\geq -(a+1)$ the right side is of the form $An+B$ and hence monotone; but the left side is not, a contradiction.

• We also note that k–regular sequences are not closed under composition. As mentioned above, $e_1(n)$, the number of 1's in the binary expansion of n, is 2–regular, as is the function $f(n) = n^2$. However, the composition $e_1(f(n)) = e_1(n^2)$ is not 2–regular; if it were, then by Corollary 2.4, $e_1(n^2)$ mod 2 would be 2–automatic. However, $e_1(n^2)$ mod 2 is not 2–automatic, by results of Allouche [A2].

In the next theorem, we show that if a sequence is k–regular, then so is the subsequence obtained by periodic indexing:

Theorem 2.6.

Let $\{S(n)\}_{n\geq 0}$ be a k–regular sequence. Then for $a \geq 1, b \geq 0$, the sequence $\{S(an+b)\}_{n\geq 0}$ is k–regular.

Proof.

Define $T(n) = S(an+b)$.

Suppose $S(n)$ is k–regular. Then the module generated by its k–kernel is generated by $S_1(n), S_2(n), \ldots, S_r(n)$. We claim that each sequence in the k–kernel of $T(n)$ can be expressed as a linear combination of $S_i(an+c)$, for $1 \leq i \leq r$ and $0 \leq c < a+b$.

Proof: Take an element of the k–kernel of $T(n)$, say $T(k^e n + j)$, $0 \leq j < k^e$. Write $ja + b = d \cdot k^e + f$, where $0 \leq f < k^e$. Then

$$\begin{aligned} T(k^e n + j) &= S(a(k^e n + j) + b) \\ &= S(k^e(an + d) + f), \end{aligned}$$

Notice that since $0 \leq j < k^e$, we have $0 \leq d < a+b$. Now the module generated by the k–kernel of $\{S(n)\}_{n\geq 0}$ is finitely generated, so $S(k^e m + f) = \sum_j c_j S_j(m)$ for constants c_j. Hence it follows that

$$S(k^e(an+d)+f) = \sum_j c_j S_j(an+d),$$

and the result follows. ■

Remark.

Let us define S indexed by negative arguments to be 0. For example, $\{S(n-1)\}_{n\geq 0}$ is the sequence $\{S(n)\}_{n\geq 0}$ with a 0 tacked on the front.

Then it is easy to see that the preceding theorem holds even when $b < 0$.

Theorem 2.7.

Let $\{S(n)\}_{n\geq 0}$ be a sequence such that there exists an $a \geq 2$ such that $\{S(an+i)\}_{n\geq 0}$ is k–regular for $0 \leq i < a$. Then $\{S(n)\}_{n\geq 0}$ is k–regular.

Proof.

For $0 \leq i < a$, define

$$T_i(n) = \begin{cases} S(n), & \text{if } n \equiv i \pmod{a}; \\ 0, & \text{if } n \not\equiv i \pmod{a}. \end{cases}$$

Also, write $S_i(n) = S(an+i)$. Then it is easy to see that each sequence $T_i(n)$ is k–regular; indeed, $T_i(k^j n + c)$ is either the 0-sequence or the sequence $S_i(k^j n + c')$ interspersed with groups of $a/\gcd(a,k^j)-1$ zeros. Hence the k–kernel of $T_i(n)$ is finitely generated. Finally, we see that

$$S(n) = \sum_{0\le i<a} T_i(n),$$

which shows that $\{S(n)\}_{n\ge 0}$ is k–regular. ■

Remark.

From Theorems 2.6 and 2.7 it follows that if $S(n)$ is k–regular, and r is a rational number, then $S(\lceil rn\rceil)$ and $S(\lfloor rn\rfloor)$ are also k-regular.

Many sequence transformations from the literature preserve regularity. For example, let $\{S(n)\}_{n\ge 0}$ be a sequence, and consider its *Toeplitz transformation* $\{S'(n)\}_{n\ge 0}$ defined by $S'(2n) = S(n)$ and $S'(2n+1) = S'(n)$ for $n \ge 0$. (See [JK], [Pro]). Then we have the following, which generalizes the case of automatic sequences [A3]:

Theorem 2.8.

$\{S(n)\}_{n\ge 0}$ is 2–regular if and only if $\{S'(n)\}_{n\ge 0}$ is 2–regular.

Proof.

Suppose $\{S(n)\}_{n\ge 0}$ is 2–regular. Then the module generated by its 2–kernel is finitely generated, say by $S_1(n),\ldots,S_k(n)$. Now consider the module

$$M = \langle S'(n), S_1(n), \ldots, S_k(n)\rangle.$$

Note that $S_i(2n)$ and $S_i(2n+1)$ are linear combinations of the S_j. Also, $S'(2n) = S(n)$ and $S'(2n+1) = S'(n)$. Thus by Theorem 2.2 (d), $\{S'(n)\}_{n\ge 0}$ is 2–regular.

Now assume $\{S'(n)\}_{n\ge 0}$ is 2–regular. Then by Theorem 2.6, $S'(2n)$ is 2–regular. But $S'(2n) = S(n)$, and the result follows. ■

Theorem 2.9

Let f be an integer ≥ 1. Then $\{S(n)\}_{n\ge 0}$ is k^f–regular if and only if $\{S(n)\}_{n\ge 0}$ is k–regular.

Proof.

Suppose $\{S(n)\}_{n\ge 0}$ is k–regular. Then the module generated by its k–kernel is finitely generated and contains its k^f–kernel. Hence the module generated by its k^f–kernel is also finitely generated.

To prove the other direction, assume $\{S(n)\}_{n\ge 0}$ is k^f–regular.

We now show there exists a B such that for all $b > B$, each subsequence $S(k^b n + c)$ can be expressed as a linear combination

$$S(k^b n + c) = \sum_i d_i S(k^{b_i} n + c_i)$$

with $b_i \le B$ and $0 \le c_i < k^{b_i}$. The result will then follow from Theorem 2.2 (c).

For let us write $b = fr + s$, $0 \le s < f$ and $c = qk^{fr} + t$, $0 \le t < k^{fr}$. Then, by Theorem 2.2 (c), there exists E such that for all $r > E$ we can express

$$S((k^f)^r m + t) = \sum_i d_i S((k^f)^{r_i} m + t_i),$$

where $r_i \le E$ and $0 \le t_i < k^{fr_i}$.

Now put $m = k^s n + q$; we find

$$S((k^f)^r m + t) = S(k^b n + c) = \sum_i d_i S(k^{fr_i + s} n + qk^{fr_i} + t_i) = \sum_i d_i S(k^{b_i} n + c_i),$$

where $b_i = fr_i + s$ and $c_i = qk^{fr_i} + t_i$. Notice that $b_i < fE + f$. Also, $q \le k^s - 1$, so

$$\begin{aligned} c_i &= qk^{fr_i} + t_i \le (k^s - 1)k^{fr_i} + t_i \\ &\le k^{fr_i + s} - k^{fr_i} + t_i < k^{fr_i + s} = k^{b_i}; \end{aligned}$$

thus we may take $B = f(E+1)$. Hence $\{S(n)\}_{n \ge 0}$ is also k–regular. ■

C. Choffrut and C. Reutenauer have pointed out that we may obtain alternative proofs of Theorems 2.6–2.9 using the notion of rational transduction [SS] and Theorem 4.3 below.

Theorem 2.10.

Let $\{S(n)\}_{n \ge 0}$ be a k–regular sequence with values in $\mathbb{C}$, the complex numbers. Then there exists a constant c such that $S(n) = O(n^c)$.

Proof.

We use the characterization of Theorem 2.2 (e). Let the base–k expansion of n be

$$a_{j-1} a_{j-2} \cdots a_1 a_0;$$

then $j \le 1 + \log_k n$. Then

$$V(n) = B_{a_0} B_{a_1} \cdots B_{a_{j-1}} V(0).$$

If v is a d-dimensional vector, define

$$\|v\| = \sum_{1 \le i \le d} |v_i|;$$

if M is a $d \times d$-matrix, define

$$\|M\| = \max_{1 \le i \le d} \sum_{1 \le j \le d} |M_{ij}|.$$

Then it is easy to see that $\|Mv\| \le \|M\| \|v\|$.

Thus we see

$$S(n) \leq \|V(n)\| \leq \|B_{a_0}\| \|B_{a_1}\| \cdots \|B_{a_{j-1}}\| \|V(0)\|.$$

Now let $c = \max_{0 \leq i \leq k-1} \|B_i\|$, and $d = \|V(0)\|$. Then we have

$$S(n) \leq c^{1+\log_k n} d \leq d' n^{c'},$$

and the result follows. ■

Thus we see, for example, that $\{2^n\}_{n \geq 0}$ is not $(\mathbb{Z}, k)$–regular.

Theorem 2.11.

Let R be a Noetherian ring without zero divisors, and let $a \in R$. Then the sequence of powers $\{a^n\}_{n \geq 0}$ is (R, k)–regular if and only if $a = 0$ or a is a root of unity.

Proof.

One direction is simple, since if a is 0 or a root of of unity, then the sequence of powers is ultimately periodic, hence k–regular.

Now assume $\{a^n\}_{n \geq 0}$ is (R, k)–regular. Then there exist $r < \infty$ and λ_j, $0 \leq j < r$ such that

$$\sum_{0 \leq j < r} \lambda_j a^{k^j \cdot n} = 0$$

for all $n \geq 0$.

Now recall the following identity for the Vandermonde determinant:

$$\begin{pmatrix} 1 & b_0 & b_0^2 & \ldots & b_0^m \\ 1 & b_1 & b_1^2 & \ldots & b_1^m \\ \vdots & \vdots & \vdots & \ddots & \ldots \\ 1 & b_m & b_m^2 & \ldots & b_m^m \end{pmatrix} = \prod_{i > j} (b_i - b_j).$$

From this, we see that the sequences $\{b_j^n\}_{n \geq 0}$ are linearly independent if and only if the numbers $b_1, b_2, \ldots, b_m$ are distinct.

Hence the numbers $1, a^k, a^{k^2}, \ldots, a^{k^r}$ are not all distinct and we must have

$$a^{k^j} = a^{k^l}$$

for some $j \neq l$. Since R has no zero-divisors, either $a = 0$ or a is a root of unity. ■

III. The ring of k–regular sequences.

Associated to every k–regular sequence $\{S(n)\}_{n \geq 0}$ is the formal power series in $R[[X]]$ defined by

$$\sum_{n \geq 0} S(n) X^n,$$

where X is an indeterminate. We call such a power series *k–regular*. In this section we show that the set of all k–regular power series forms a ring (but not a field).

Recall that the *convolution* $S \star T$ of two sequences $S(n)$ and $T(n)$ is defined as follows:

$$(S \star T)(n) = \sum_{i+j=n} S(i)T(j).$$

Theorem 3.1.

The set of k–regular sequences is closed under convolution.

Proof.

For simplicity we prove this only in the case $k = 2$.

Let us agree to write $\{A(2n)\}$ as shorthand for the sequence $\{A(2n)\}_{n\geq 0}$.

Let A and B be 2–regular sequences. The modules generated by their 2–kernels are generated by sequences $a_1, a_2, \ldots, a_{i'}$ and $b_1, b_2, \ldots, b_{j'}$, respectively. We want to find a basis for C, the module generated by the 2–kernel of $A \star B$. We write $u_{ij} = a_i \star b_j$ for $1 \leq i \leq i', 1 \leq j \leq j'$. We claim that the set $\mathcal{M}$ of $2i'j'$ sequences $\{u_{ij}(n)\}_{n\geq 0}$ and $\{u_{ij}(n-1)\}_{n\geq 0}$ generates the module C. (As in the previous section, we define $u_{ij}(-1) = 0$.)

It is clear that $\mathcal{M}$ contains all sequences of the form

$$(\{A(2^e n + i)\} \star \{B(2^f n + j)\})(n) \tag{1}$$

and

$$(\{A(2^e n + i)\} \star \{B(2^f n + j)\})(n-1). \tag{2}$$

Thus it suffices to show how to write all the sequences of the form

$$\{(A \star B)(2^g n + a)\}$$

as a linear combination of the sequences in (1) and (2).

This is done using the following formula:

$$\begin{aligned}(A \star B)(2^g n + a) = &\sum_{0\leq i\leq a} (\{A(2^g n + i)\} \star \{B(2^g n + a - i)\})(n) \\ &+ \sum_{a<j<2^g} (\{A(2^g n + j)\} \star (\{B(2^g n + 2^g + a - j)\})(n-1).\end{aligned}$$

Hence the result follows. ■

(Note: it is apparently impossible to obtain Theorem 3.1 using the standard tools of rational series, such as rational transductions.)

It follows from Theorem 3.1 that if the sequence $\{S(n)\}_{n\geq 0}$ is k–regular, then so is its running sum $\{\sum_{0\leq j\leq n} S(j)\}_{n\geq 0}$.

Since the convolution of sequences is equivalent to (ordinary) multiplication of the associated power series, we have:

Corollary 3.2.

The set of k–regular power series forms a ring.

Remark.

The set of k–regular power series does not form a field. This follows from the identity

$$\frac{1}{1-2X} = 1 + 2X + 4X^2 + 8X^3 + \cdots$$

and the fact that $\{2^n\}_{n\geq 0}$ is not k–regular (Theorem 2.10).

Theorem 3.3.

Let F be an algebraically closed field (e. g., $\mathbb{C}$). Let $\{S(n)\}_{n\geq 0}$ be a sequence with values in F. Let $f(X) = \sum_{n\geq 0} S(n)X^n$ be a formal power series in $F[[X]]$. Assume that $f(X)$ represents a rational function of X; i. e. there exist polynomials $p(X)$, $q(X)$ such that $f(X) = p(X)/q(X)$. Then $\{S(n)\}_{n\geq 0}$ is k–regular if and only if the poles of f are roots of unity.

Proof.

Note that by assumption, 0 is not a pole of f.

Suppose the poles of f are roots of unity. Then using expansion by partial fractions, we can write

$$f(X) = \sum_i \frac{c_i}{(1-\zeta_i X)^{e_i}}$$

where $c_i \in F$, the e_i are non-negative integers, and each ζ_i is a root of unity. To prove the coefficients of f form a k–regular sequence, it clearly suffices to show that $(1-\zeta_i X)^{-1}$ is k–regular. But this power series has periodic coefficients and so is k–regular.

Now suppose $f(X) = p(X)/q(X)$ for polynomials p, q, and f is k–regular. Let $1/\zeta$ be one of the poles of f; we may assume $\zeta \neq 0$. We can then write

$$f(X) = \frac{p(X)}{q(X)} = \frac{r(X)}{s(X)(1-\zeta X)^e},$$

where $r(x), s(X)$ are polynomials and $r(X)$ and $1-\zeta X$ are relatively prime. Then there exist two polynomials $u(X), v(X)$ such that

$$u(X)r(X) + v(X)(1-\zeta X)^e = 1.$$

Now $u(X)f(X)s(X) + v(X)$ is also a k–regular power series, and we have

$$u(X)f(X)s(X) + v(X) = (1-\zeta X)^{-e}. \tag{3}$$

Thus $(1-\zeta X)^{-e}$ is k–regular. But $(1-\zeta X)^{e-1}$ is a polynomial and hence a k–regular power series, so its product with (3) is k–regular and thus $(1-\zeta X)^{-1}$ is k–regular. But the coefficients of this power series are ζ^n, which by Theorem 2.11 is k–regular if and only if ζ is a root of unity.

This completes the proof. ■

Remarks.

• We note that Theorem 3.3 gives us the following characterization of k–regular sequences associated with rational formal power series: they must be linear recurrences whose characteristic polynomial is a product of cyclotomic polynomials. See, for example, Section VII, Example 18.

• Also note that if $R = \mathbb{Q}$, then (using Corollary 4.2 below) the radius of convergence of a k–regular power series is 1, and such a series either represents a rational function or has the unit circle as a natural boundary.

IV. Rational series and k–regular sequences.

At first glance, it might seem that there is no relationship between k–regular power series and the theory of $\mathbb{Z}$–rational formal series, as described in [SS], [BR] [E, Chap. V]. For $\sum_{n\geq 0} 2^n X^n$ is $\mathbb{Z}$–rational, but is not k–regular. Similarly, $\sum_{n\geq 0} e_1(n)X^n$ is k–regular, but is not $\mathbb{Z}$–rational. (Here $e_1(n)$ counts the number of 1's in the base–k expansion of n).

Nevertheless, there *is* a relationship which can be roughly described as follows: 2–regular power series are the "binary" analogue of $\mathbb{Z}$–rational formal series in one variable. Alternatively, $\mathbb{Z}$–rational series in one variable are the "unary" analogue of k–regular power series.

In this section, we develop this relationship between k–regular sequences and $\mathbb{Z}$–rational formal series. From this, we get a machine model for the k–regular sequences. This model plays the same role as the ordinary finite automaton does for k–automatic sequences. We also prove that all k–regular sequences can be computed quickly.

We introduce some notation that will be used throughout this section. Let k be fixed and define $\Sigma = \{0, 1, \ldots, k-1\}$. We need a way to uniquely associate integers with strings giving their base–k representation. If

$$n = \sum_{0\leq i<e} a_i k^i,$$

and $a_{e-1} \neq 0$, then we say that the string $a_{e-1}a_{e-2}\ldots a_1 a_0$ is the *standard base–k representation* of n. Note that the standard representation of 0 is the empty string. The set of *all* standard representations is just $\epsilon + (\Sigma - 0)\Sigma^*$.

First, we prove a useful lemma:

Lemma 4.1.

Let $\{S(n)\}_{n\geq 0}$ be a sequence with entries in R. Then $\{S(n)\}_{n\geq 0}$ is (R', k)–regular if and only if there exist matrices M_0, $M_1, \ldots, M_{k-1}$ with entries in R' and vectors λ, κ with entries in R such that

$$S(n) = \lambda M_{a_0} M_{a_1} \ldots M_{a_{e-1}} \kappa,$$

where $a_{e-1}a_{e-2}\ldots a_1 a_0$ is the standard base–k representation of n.

Proof.

Suppose $S(n)$ is k–regular. Then by Theorem 2.2 (e), we know that there exist matrices $M_0, \ldots M_{k-1}$ such that

$$V(kn+a) = M_a V(n),$$

where

$$V(n) = \begin{pmatrix} S_1(n) \\ \vdots \\ S_r(n) \end{pmatrix},$$

and $S(n) = S_1(n)$. Hence by setting $\kappa = V(0)$ and $\lambda = [1\ 0\ 0 \cdots 0]$, we see that

$$V(n) = \lambda M_{a_0} M_{a_1} \ldots M_{a_{e-1}} \kappa$$

for all $n \geq 0$.

Now suppose $S(n) = \lambda M_{a_0} \ldots M_{a_{e-1}} \kappa$ for all $n \geq 0$, where $a_{e-1} \cdots a_0$ is the standard base–k representation of n. Define $V(n) = M_{a_0} \cdots M_{a_{e-1}} \kappa$ and

$$V(n) = \begin{pmatrix} v_1(n) \\ \vdots \\ v_r(n) \end{pmatrix}.$$

Then

$$V(kn+a) = M_a M_{a_0} \cdots M_{a_{e-1}} \kappa = M_a V(n),$$

except possibly when $n = 0$ and $a = 0$. (This special case arises because the standard representation of kn is the string $a_{e-1} \cdots a_1 a_0 0$, for $n \geq 1$, but not for $n = 0$.) In this case, by setting $v' = V(0) - M_0 V(0)$ we see

$$V(kn) = M_0 V(n) + U(n) v'$$

for all $n \geq 0$, where $U(n)$ denotes the sequence that is 1 when $n = 0$ and 0 otherwise.

Then by Theorem 2.2 (d), we see that each of the sequences $v_1(n), \ldots, v_r(n)$ is k–regular. But then $S(n) = \lambda V(n)$ is k–regular, by Theorem 2.5. ■

Corollary 4.2. *Suppose $\{S(n)\}_{n\geq 0}$ is a $(\mathbb{Z}, k)$–regular sequence with values in $\mathbb{Q}$. Then there exist an integer r and a $(\mathbb{Z}, k)$–regular sequence $\{T(n)\}_{n\geq 0}$ with values in $\mathbb{Z}$ such that $S(n) = T(n)/r$.*

Proof. By Lemma 4.1, we have

$$S(n) = \lambda M_{a_0} M_{a_1} \cdots M_{a_{e-1}} \kappa$$

where $a_{e-1} \cdots a_1 a_0$ is the standard base–k representation of n. The matrices M_i have integral entries, and the vectors λ and κ have rational entries. Let g be the least common multiple of the denominators of entries in λ, and g' be the least common multiple of the denominators of entries in κ. Then $T(n) = (g\lambda)M_{a_0}M_{a_1}\cdots M_{a_{e-1}}(g'\kappa)$ is a $(\mathbb{Z}, k)$–regular sequence with values in $\mathbb{Z}$. The result follows by putting $r = gg'$. ■

Now we show how k–regular sequences are related to $\mathbb{Z}$–rational formal series. Let $x_0, x_1, \ldots, x_{k-1}$ be non-commuting variables. If $w = w_1 \cdots w_r \in \Sigma^*$, then define $x_w = x_{w_1} \cdots x_{w_r}$. Let τ be the map that sends n to $x_{a_0}x_{a_1} \ldots x_{a_{e-1}}$, where the standard base–k representation of x is the string $a_{e-1} \cdots a_1 a_0$.

Theorem 4.3.
$\{S(n)\}_{n\geq 0}$ is k–regular if and only if the formal series

$$\sum_{n\geq 0} S(n)\tau(n)$$

is $\mathbb{Z}$–rational.

For example, in the case $k = 2$ we have

$$\sum_{n\geq 0} S(n)\tau(n) = S(0) + S(1)x_1 + S(2)x_0x_1 + S(3)x_1x_1 + S(4)x_0x_0x_1 + \cdots .$$

Proof. Suppose $\{S(n)\}_{n\geq 0}$ is k–regular. Then by Lemma 4.1, there exist matrices $M_0, \ldots, M_{k-1}$ such that

$$S(n) = \lambda M_{a_0} \cdots M_{a_{e-1}} \kappa.$$

But by the fundamental theorem for $\mathbb{Z}$–rational formal series (see, e.g. [SS, Theorem 2.3]),

$$T = \sum_{w\in\Sigma^*} \lambda M_w \kappa x_w$$

is $\mathbb{Z}$–rational. This is essentially the series $\sum_{n\geq 0} S(n)\tau(n)$, but it also contains terms that correspond to non-standard base–k representations of n. Let A be the set of standard base–k representations (e. g. those not beginning with a 0). Then as above, $A = \epsilon + (\Sigma - 0)\Sigma^*$, and so A is regular. Let A^R denote the set of reversals of strings in A; then A^R is also regular. Now

$$U = \text{char } A^R = \sum_{w\in A^R} x_w$$

is a $\mathbb{Z}$–rational formal series (see, e. g. [SS, Corollary 5.4 (iii)]). Then $T \odot U$ (the Hadamard product) is equal to $\sum_{n\geq 0} S(n)\tau(n)$, and since $\mathbb{Z}$–rational series are closed under $\odot$ (see, e. g. [SS, Theorem 4.4]), the result follows.

Now suppose $\sum_{n\geq 0} S(n)\tau(n)$ is $\mathbb{Z}$–rational. Then again by the definition of τ and the fundamental theorem we have $S(n) = \lambda M_w \kappa$, where $w = a_0 a_1 \cdots a_{e-1}$, and $a_{e-1} \cdots a_1 a_0$ is the standard base–k representation of n. This completes the proof. ■

Theorem 4.3 allows us to use the well-developed theory of $\mathbb{Z}$–rational series to discuss the properties of k–regular sequences, at least in some cases. We continue this below in Section V. Now, however, we sketch a description of our machine model.

This model is essentially the same as that first given by Schützenberger [Sch]. However, we repeat the description for completeness.

Let us define what we call a *matrix machine*. It is a finite-state machine with auxiliary storage in the form of a column vector $v \in R_{j1}$ for some $j > 0$. Here is how the machine operates: Suppose we are in state q. Upon reading a symbol a from the input, the machine first replaces v with Mv, where $M = M(q, a)$ is a $j \times j$ matrix. Then the machine moves to a new state $\delta(q, a)$. The output is determined as follows: when the last input symbol is read, we are in state q'. There is a row vector $\lambda(q')$, and the output is the scalar $\lambda(q')v$.

Now consider the case where the input is the base–k representation of an integer n, starting with the most significant digit, and the matrix machine computes $S(n)$. We claim this is precisely the class of k–regular sequences. By Lemma 4.1, this equivalence is easily seen in the case of 1-state machines. Thus to prove the equivalence it suffices to prove the following

Theorem 4.4 (Schützenberger).

A matrix machine with r states can be simulated by a matrix machine with 1 state.

Proof.

To simplify the exposition we show how to do this in the case where j, the size of the vectors and matrices involved, equals 1.

The idea is to replace the single element v by a vector v' of size r. All of the entries of v' will be zero, except for a single entry which equals v. We code the current state by the position of v inside v'; if it is in position i, we are currently in state i. Instead of multiplying by $M(q, a)$ we multiply by the matrix PQ, where $Q_{ii} = M(q_i, a)$, $0 \leq i \leq r-1$, and P is a permutation matrix defined as follows:

$$P_{ij} = \begin{cases} 1, & \text{if } \delta(q_j, a) = q_i \\ 0, & \text{otherwise.} \end{cases}$$

Finally, $\lambda(q_i)$ is the vector consisting of all ones.

The correctness of the construction is left to the reader. To extend this proof to the case $j > 1$, we replace all entries by block matrices. ∎

Corollary 4.5.

The n-th term of a k–regular sequence can be computed using $O(\log n)$ operations, where an operation is an addition or multiplication of elements in the ring R.

Corollary 4.6.

The n-th term of a k–regular sequence over $\mathbb{Z}$ can be computed in time polynomial in $\log n$.

Remarks.

• At first glance, our matrix machines would also seem to be similar to the linear sequential machines (LSM) of Harrison [Har1]. This is not the case, however. Our input symbols a are chosen from an arbitrary alphabet Σ, while the LSM model uses k-tuples chosen from a field. Our model allows a different $n \times n$ matrix $M(q, a)$ for every state q and input symbol a, whereas the LSM model uses exactly two matrices A and B and defines a transition by

$$\delta(q, a) = Aq + Ba.$$

Our model allows the matrices to contain arbitrary ring elements, whereas the LSM model uses a field. Finally, in our model we are only interested in the output associated with the final state, rather than the string of outputs associated with each state visited.

• We mention a connection between $(\mathbb{Z}, k)$–regular sequences with values in $\mathbb{Z}$ and the group $\Gamma_k(\mathbb{Z})$ of Morton and Mourant [MM]. Indeed, every sequence $\{S(n)\}_{n\geq 0}$ in $\Gamma_k(\mathbb{Z})$ is k–regular, as it is easily seen that $\{S(n)\}_{n\geq 0} \in \Gamma_k(\mathbb{Z})$ if and only if the sequence $\{S(n) - S(\lfloor n/k \rfloor)\}_{n\geq 0}$ is periodic.

V. The zero-set of a k–regular sequence.

Let $\{S(n)\}_{n\geq 0}$ be a k–regular sequence. In this section, we discuss the set

$$Q = \{n \mid S(n) = 0\},$$

or, more precisely, the set $Z(S)$ of strings of the standard base–k representations of elements of Q. We call this set the *zero-set* of the sequence $\{S(n)\}_{n\geq 0}$.

We also discuss the set $\overline{Z}(S)$, the set of strings of the standard base–k representations of n such that $S(n) \neq 0$. (This set is essentially the *support* of the associated $\mathbb{Z}$–rational power series.) Note that

$$Z(S) + \overline{Z}(S) = \epsilon + (\Sigma - 0)\Sigma^*,$$

where $\Sigma = \{0, 1, \ldots, k-1\}$.

Theorem 5.1. *The set $Z(S)$ is simultaneously in logarithmic space and polynomial time. The set $Z(S)$ is also in the complexity class* `NC`.

Proof.

The first statement follows immediately from results of Lipton and Zalcstein [LZ].

The second statement is left to the reader. ■

Theorem 5.2. *For fixed $k \geq 2$, it is undecidable if a given k–regular sequence $\{S(n)\}_{n\geq 0}$ has a zero term. In other words, it is undecidable if $Z(S)$ is nonempty.*

Proof.

To specify the k–regular sequence $S(n)$, it is necessary to agree on a representation. We assume we have been given the matrices in Lemma 4.1 or Theorem 2.2 (e).

As in [SS, Theorem 12.1], we reduce the problem of determining whether or not an arbitrary multivariate polynomial equation

$$p(x_1, x_2, \ldots x_r)$$

has a solution in non-negative integers (Hilbert's tenth problem) to the problem of whether $Z(S)$ is nonempty. The result will then follow by the celebrated result of Davis-Matijacevič-Putnam-Robinson [Dav].

Suppose we are given $p(x_1, x_2, \ldots, x_r)$. We encode this equation as a k–regular sequence as follows. First, we choose f such that $k^f \geq r+1$. We now represent the variable x_j by $e_j(n)$, the number of j's in the base–k^f expansion of n. Clearly for each r-tuple of non-negative integers $(b_1, b_2, \ldots, b_r)$, there exists an n for which

$$(e_1(n), \ldots, e_r(n)) = (b_1, \ldots, b_r).$$

Now $S(n) = p(e_1(n), e_2(n), \ldots, e_r(n))$ is k^f–regular and its matrix representation can be computed with a recursive algorithm. But by Theorem 2.9, $S(n)$ is also k–regular; furthermore, the corresponding matrices are effectively determinable. Clearly $Z(S)$ is nonempty if and only if $p(x_1, x_2, \ldots, x_r)$ has a solution in non-negative integers. ∎

As in [SS, p. 124], we can also give a more explicit example:

Theorem 5.3. *There exists a k–regular sequence $\{S(n)\}_{n \geq 0}$ such that neither $Z(S)$ nor $\overline{Z}(S)$ are context-free.*

Proof.

Define $S(n) = e_1(n)^2 - e_0(n)$. It is not hard to verify that $\{S(n)\}_{n \geq 0}$ is k–regular. (Indeed, it will follow from Theorem 6.1.)

Now suppose $Z(S)$ is context-free. Then $Z(S) \cap 1^+0^* = \{1^n 0^{n^2} \mid n \geq 1\}$ would also be context-free. But this can easily seen to be false, using the pumping lemma.

Now suppose $\overline{Z}(S)$ is context-free. Then $L_1 = Z(S) \cap 1^+0^* = \{1^n 0^r \mid n \geq 1, r \neq n^2\}$ would be context-free. By a theorem of Ginsburg and Spanier [GS, Theorem 6.2, Corollary 2], $L_2 = 1^*0^* - L$ would be context free. But $L_2 \cap 1^+0^+ = \{1^n 0^{n^2} \mid n \geq 1\}$, which is not context-free, a contradiction. ∎

VI. Some "Fourier" expansions.

For simplicity, all results and proofs in this section assume $k = 2$.

We introduce some notation that will be used throughout this section. Let $n_{(2)}$ denote the string in $A = \epsilon + 1(0+1)^*$ that represents n in base 2. If s is a string in A, let $v(S)$ denote the integer represented by s. Let $|s|$ denote the length of the string s. Let $\lambda(n)$ be the integer obtained from n by deleting the most significant bit of its base–2 expansion. Let m and n be integers; we write m suff n for the relation: the string $m_{(2)}$ is a suffix of the string $n_{(2)}$. Define $E = 1(0+1)^*$. Let $P \in E$, and let $e_P(n)$ denote the number of (possibly overlapping) occurrences of P in the base-k expansion of n. Let $x_P(n)$ be the function that takes the value 1 if P is a suffix of $n_{(2)}$, and 0 otherwise.

Morton and Mourant proved [MM] that any sequence $\{S(n)\}_{n\geq 0}$ taking values in $\mathbb{Z}$ has a unique expansion as an infinite sum, as follows:

$$S(n) = S(0) + \sum_{P\in E} \hat{S}(v(P))e_P(n).$$

Here the "Fourier" coefficients $\hat{S}(m)$ are integers. We define $\hat{S}(0) = S(0)$, and call the sequence $\{\hat{S}(n)\}_{n\geq 0}$ the *pattern transform* of $\{S(n)\}_{n\geq 0}$.

In this section, we prove the following result: a sequence is 2–regular if and only if its pattern transform is 2–regular. First, however, we show that the sequences e_P themselves are 2–regular.

Theorem 6.1.

The sequence $\{e_P(n)\}_{n\geq 0}$ is 2–regular for any pattern $P \in E$.

Proof.

Let us introduce the following notation: if $w = w_1 w_2 \cdots w_{j'}$ is a string and $j \leq j'$, then

$$\text{take}(j, w) = w_1 w_2 \cdots w_j.$$

We claim that each element of the 2–kernel can be written as a linear combination of the sequences $e_P(2^f n + a)$ for $0 \leq f < |P|$ and $0 \leq a < 2^f$ and the constant sequence 1.

Proof: Consider an element of the 2–kernel, $e_P(2^f n + a), 0 \leq a < 2^f$. Then if $f \leq |P| - 1$, this sequence is already in the list above. Otherwise, $f \geq |P|$. Then $2^f n + a$ can be written in base 2 as

$$n_{(2)} a'$$

where $|a'| = f$ and $v(a') = a$. Then

$$e_P(2^f n + a) = e_P(2^{|P|-1} n + c) + e_P(a),$$

where $c = v(\text{take}(|P| - 1, a'))$.

Now the first term on the right is in the list above, and the second term is a constant multiple of the constant sequence 1. Hence $e_P(2^f n + a)$ is a $\mathbb{Z}$–linear combination of elements in the list, and this completes the proof. ■

Corollary 6.2.

$\{e_P(an + b)\}_{n\geq 0}$ is 2–regular for all $a, b \geq 0$.

Theorem 6.3.

$\{S(n)\}_{n\geq 0}$ is 2–regular if and only if $\{\hat{S}(n)\}_{n\geq 0}$ is 2–regular.

First we prove two lemmas.

Lemma 6.4.

For all $n \geq 0$ we have

$$S(2n) = S(n) + \sum_{\substack{m \geq 1 \\ m \text{ suff } n}} \hat{S}(2m)$$

and

$$S(2n+1) = S(n) + \hat{S}(1) + \sum_{\substack{m \geq 1 \\ m \text{ suff } n}} \hat{S}(2m+1).$$

Proof.

$$\begin{aligned}
S(2n) &= S(0) + \sum_P \hat{S}(v(P)) e_P(2n) \\
&= S(0) + \hat{S}(v(1)) e_1(2n) + \sum_P \hat{S}(v(P0)) e_{P0}(2n) + \sum_P \hat{S}(v(P1)) e_{P1}(2n) \\
&= S(0) + \hat{S}(1) e_1(n) + \sum_P \hat{S}(v(P0)) e_{P0}(n) + \sum_P \hat{S}(v(P0)) x_P(n) + \sum_P \hat{S}(v(P1)) e_{P1}(n) \\
&= S(0) + \sum_P \hat{S}(v(P)) e_P(n) + \sum_P \hat{S}(v(P0)) x_P(n) \\
&= S(n) + \sum_P \hat{S}(v(P0)) x_P(n) \\
&= S(n) + \sum_{\substack{m \geq 1 \\ m \text{ suff } n}} \hat{S}(2m).
\end{aligned}$$

The formula for $S(2n+1)$ is proved similarly; the extra term $\hat{S}(1)$ comes from the fact that

$$\hat{S}(v(1)) e_1(2n+1) = \hat{S}(1)(e_1(n) + 1) = \hat{S}(1) e_1(n) + \hat{S}(1).$$

This completes the proof. ■

Lemma 6.5.

For all $n \geq 1$ we have

$$\hat{S}(2n) = S(2n) - S(n) - S(2\lambda(n)) + S(\lambda(n)).$$

For all $n \geq 1$ we have

$$\hat{S}(2n+1) = S(2n+1) - S(n) - S(2\lambda(n) + 1) + S(\lambda(n)).$$

Proof.

Notice first that

$$\begin{aligned}\{m \geq 1 \mid m \text{ suff } n\} &= \{m \geq 1 \mid (m \text{ suff } n) \text{ and } (m \neq n)\} \cup \{n\} \\ &= \{m \geq 1 \mid m \text{ suff } \lambda(n)\} \cup \{n\},\end{aligned}$$

the unions being disjoint.

Hence, using Lemma 6.4, we find

$$\begin{aligned}S(2n) - S(n) &= \sum_{\substack{m \geq 1 \\ m \text{ suff } n}} \hat{S}(2m) \\ &= \hat{S}(2n) + \sum_{\substack{m \geq 1 \\ m \text{ suff } \lambda(n)}} \hat{S}(2m),\end{aligned}$$

which can be rewritten as:

$$\hat{S}(2n) = (S(2n) - S(n)) - (S(2\lambda(n)) - S(\lambda(n))).$$

The second formula is obtained in a slightly different manner:

$$\begin{aligned}S(2n+1) - S(n) &= \hat{S}(1) + \sum_{\substack{m \geq 1 \\ m \text{ suff } n}} \hat{S}(2m+1) \\ &= \hat{S}(2n+1) + \hat{S}(1) + \sum_{\substack{m \geq 1 \\ m \text{ suff } \lambda(n)}} \hat{S}(2m+1),\end{aligned}$$

which can be rewritten as:

$$\begin{aligned}\hat{S}(2n+1) &= (S(2n+1) - S(n) - \hat{S}(1)) - (S(2\lambda(n)+1) - S(\lambda(n)) - \hat{S}(1)) \\ &= S(2n+1) - S(n) - S(2\lambda(n)+1) + S(\lambda(n)).\end{aligned}$$

This completes the proof of Lemma 6.5. ■

We are now ready to prove Theorem 6.3:

Proof.

Suppose first that S is 2–regular and let $\{S_1 = S, S_2, \ldots, S_r\}$ be a finite set of generators for the $\mathbb{Z}$–module generated by its 2–kernel. Define

$$U(n) = \begin{cases} 1, & \text{if } n = 0; \\ 0, & \text{otherwise.} \end{cases}$$

Consider the $\mathbb{Z}$–module $\mathcal{M}$ generated by the $S_j(n)$, the $S_j(\lambda(n))$, and $U(n)$; i. e.

$$\mathcal{M} = \langle S_1, S_2, \ldots, S_r, S_1 \circ \lambda, S_2 \circ \lambda, \ldots, S_r \circ \lambda, U, \hat{S} \rangle.$$

To prove that $\hat{S}$ is 2–regular, it suffices to prove that for each sequence $V(n)$ contained in the list of generators for $\mathcal{M}$, the subsequences $\{V(2n)\}_{n\geq 0}$ and $\{V(2n+1)\}_{n\geq 0}$ are in $\mathcal{M}$.

For $S_1, \ldots, S_r$, this follows from the 2–regularity of S. By Lemma 6.5,

$$\hat{S}(2n) = S(2n) - S(n) - S(2\lambda(n)) + S(\lambda(n)) + \hat{S}(0)U(n),$$

and so $\mathcal{M}$ contains $\{\hat{S}(2n)\}_{n\geq 0}$. Similarly, Lemma 6.5 implies that

$$\hat{S}(2n+1) = S(2n+1) - S(n) - S(2\lambda(n)+1) + S(\lambda(n)) + \hat{S}(1)U(n);$$

hence $\mathcal{M}$ contains $\{\hat{S}(2n+1)\}_{n\geq 0}$.

For $S_i \circ \lambda$, we have $S_i(\lambda(2n)) = S_i(2\lambda(n))$. Similarly,

$$\begin{aligned} S_i(\lambda(2n+1)) &= \begin{cases} S_i(2\lambda(n)+1), & \text{if } n \geq 1; \\ S_i(0), & \text{if } n = 0, \end{cases} \\ &= S_i(2\lambda(n)+1) + (S_i(0) - S_i(1))U(n), \end{aligned}$$

showing that $S_i(2\lambda(n)+1)$ can be written as a linear combination of generators.

Finally, we see that $U(2n) = U(n)$, and $U(2n+1) = 0$, for all $n \geq 0$.

Now suppose that $\hat{S}$ is 2–regular. We wish to see that S is 2–regular.

Let $\hat{S}_1 = \hat{S}, \hat{S}_2, \ldots, \hat{S}_t$ be a finite set of generators for the $\mathbb{Z}$–module generated by the 2–kernel of $\hat{S}$. Then there exist integers a_{ij} and b_{ij} such that

$$\hat{S}_i(2n) = \sum_{1\leq j\leq t} a_{ij}\hat{S}_j(n),$$

and

$$\hat{S}_i(2n+1) = \sum_{1\leq j\leq t} b_{ij}\hat{S}_j(n).$$

Define

$$T_i(n) = \sum_{\substack{m\geq 1 \\ m \text{ suff } n}} \hat{S}_i(m),$$

and consider the $\mathbb{Z}$–module $\mathcal{N}$ generated by $S, T_1, T_2, \ldots, T_t$, and the constant sequence 1.

This module contains S. We must prove for each of the generators V, the sequences $\{V(2n)\}_{n\geq 0}$ and $\{V(2n+1)\}_{n\geq 0}$ are in $\mathcal{N}$.

For S, this follows from Lemma 6.4:

$$\begin{aligned} S(2n) &= S(n) + \sum_{\substack{m\geq 1 \\ m \text{ suff } n}} \hat{S}(2m) \\ &= S(n) + \sum_{\substack{m\geq 1 \\ m \text{ suff } n}} \sum_{1\leq j\leq t} a_{1j}\hat{S}(m) \\ &= S(n) + \sum_{1\leq j\leq t} a_{1j}T_j(n); \end{aligned}$$

similarly,

$$S(2n+1) = S(n) + \hat{S}(1) + \sum_{1 \le j \le t} b_{1j} T_j(n).$$

For T_i, we have:

$$\begin{aligned} T_i(2n) &= \sum_{\substack{m \ge 1 \\ m \text{ suff } 2n}} \hat{S}_i(m) \\ &= \sum_{\substack{k \ge 1 \\ k \text{ suff } n}} \hat{S}_i(2k) \\ &= \sum_{1 \le j \le t} a_{ij} T_j(n), \end{aligned}$$

and

$$\begin{aligned} T_i(2n+1) &= \sum_{\substack{m \ge 1 \\ m \text{ suff } 2n+1}} \hat{S}_i(m) \\ &= \hat{S}_i(1) + \sum_{\substack{k \ge 1 \\ k \text{ suff } n}} \hat{S}_i(2k+1) \\ &= \hat{S}_i(1) + \sum_{1 \le j \le t} b_{ij} T_j(n). \end{aligned}$$

The result for the constant sequence 1 is left to the reader! ■

Remarks.

• C. Reutenauer has pointed out the following simple proof of Theorem 6.3: Let $A = \{\epsilon\} \cup \{1, 2, \ldots, k-1\}\Sigma^*$. Then

$$S = (S, \epsilon) + \sum_{P \in E} (\hat{S}, P) \underline{A}\ \underline{P}\ \underline{\Sigma}^*,$$

where $\underline{L}$ is the characteristic series of L. We have

$$\begin{aligned} S - (S, \epsilon) &= \underline{A}\ \hat{S}\ \underline{\Sigma}^* \\ &= \underline{A}\hat{S}(\epsilon - \Sigma)^{-1}, \end{aligned}$$

and so

$$\hat{S} = \underline{A}^{-1}(S - (S, \epsilon))(\epsilon - \Sigma).$$

However, it is not immediately clear how to obtain the explicit formula in Lemma 6.5 from this observation.

• It is possible to view Theorem 6.3 as a generalization of results of Choffrut and Schützenberger [CS]. They discussed counting functions similar to our sum

$$\sum_{P \in E} \hat{S}(v(P)) e_P(n).$$

However, because they restricted their attention to finite automata with counters, they were forced to put restrictions on the set E.

• Theorem 6.3 is also a generalization of previous results of Allouche, Morton, and Shallit [AMS].

Our last result concerns the pattern transform of $\{e_P(an+b)\}_{n\geq 0}$. We prove that, in this case, the coefficients $\hat{S}(m)$ are bounded and in fact, are k–automatic.

Theorem 6.6.
Let

$$e_P(an+b) = \hat{S}(0) + \sum_{P\in E} \hat{S}(v(P))e_P(n).$$

Then $\hat{S}(m)$ is a 2–automatic sequence.

Proof.
By Corollary 6.2, we know that $S(n) = e_P(an+b)$ is 2–regular. Hence by Theorem 6.3, $\hat{S}(n)$ is 2–regular. By Theorem 2.3, it suffices to show that $\hat{S}$ takes only finitely many values. By Lemma 6.5 it suffices to prove that $S(n) - S(\lambda(n))$ takes only finitely many values.

If $n \neq 0$ and $s = |n_{(2)}|$, one has $(an+b) - (a\lambda(n)+b) = a(n-\lambda(n)) = a2^{s-1}$. Hence $an+b$ and $a\lambda(n)+b$ have the same $s-1$ final digits. Let x be fixed such that $\max(a,b) < 2^x$; then $an+b < 2^{x+s} + 2^x < 2^{x+s+1}$; hence $an+b$ has at most $x+s+1$ digits.

Finally, the numbers $an+b$ and $a\lambda(n)+b$ differ in at most $(x+s+1)-(s-1) = x+2$ digits. Hence, for every P, $|e_P(an+b) - e_P(a\lambda(n)+b)|$ is bounded by $x+2$, and the result follows. ■

VII. Some examples.

Unless otherwise indicated, we assume $k = 2$ in the examples that follow. Sequence numbers refer to Sloane's book [Sl].

Example 1.
By Theorem 6.1, we know the sequence $\{e_1(n)\}_{n\geq 0}$ is 2–regular. In fact, it satisfies the relations $e_1(2n) = e_1(n)$; $e_1(2n+1) = e_1(n)+1$. Hence its 2–kernel is generated by $e_1(n)$ and the constant sequence 1. (This is Sloane's sequence #41.)

Example 2.
Define $A(n) = \sum_{1\leq j<n} e_1(j)$, the total number of 1's in the base–2 expansion of the first n integers. Then $A(n)$ is 2–regular by the remark after Theorem 3.1. $A(n)$ has been extensively studied in the literature ([BS], [CL], [CY]). It is Sloane's sequence #360.

Example 3.

Consider the sequence

$$\{c(n)\}_{n\geq 0} = 0, 2, 6, 8, 18, 20, 24, 26, 54, 56, \ldots,$$

which lists the numerators of the left endpoints of the Cantor set. (Alternatively, these are the integers whose base–3 representations contain no 1's; see [MFP].) Then it is easy to see that $c(2n) = 3c(n)$ and $c(2n+1) = 3c(n) + 2$. Hence it is 2–regular. (Note, however, that its characteristic sequence $101000101\cdots$ is actually 3–automatic.)

For a more general perspective on such sequences, see [Mah2].

Example 4.

The sequence $e_1(3n)$ has been studied extensively by Newman, Slater, and Coquet ([N], [NS], [Coq]). By Corollary 5.2 it is 2–regular. By Theorem 6.6 it has a 2–automatic pattern transform. In fact, we find

$$\begin{aligned} e_1(3n) &= 2e_1(n) - 2e_{11}(n) + e_{111}(n) - 2e_{1011}(n) + e_{11011}(n) - 2e_{101011}(n) + e_{1101011}(n) - \cdots \\ &= 2e_1(n) - 2\sum_{i\geq 0} e_{(10)^i 11}(n) + \sum_{i\geq 0} e_{11(01)^i 1}(n). \end{aligned}$$

This expansion gives an alternative explanation to the observation [N] that the first few values of $e_1(3n)$ are almost all even. See [AMS].

Example 5.

Let j be an integer ≥ 0. The sequence $\{n^j\}_{n\geq 0}$ is 2–regular, as the module generated by its 2–kernel is generated by the constant sequence 1 and the sequences $\{n\}_{n\geq 0}$, $\{n^2\}_{n\geq 0}$, $\ldots$, $\{n^j\}_{n\geq 0}$.

From Theorem 6.3, we know the corresponding pattern transforms are 2–regular. Using Lemma 6.5, we find:

$$\begin{aligned} n &= e_1(n) + e_{10}(n) + e_{11}(n) + 2(e_{100}(n) + \cdots + e_{111}(n)) \\ &\quad + 4(e_{1000} + \cdots + e_{1111}(n)) + 8(e_{10000}(n) + \cdots + e_{11111}(n)) + \cdots \end{aligned}$$

Example 6.

Let w^R denote the reverse of the string w. Consider the map which takes every integer to the integer represented by the reverse of its base–2 representation, i. e. $r(n) = v(n_{(2)}^R)$. Then it is not difficult to show that [IMO] $r(2n) = r(n)$, $r(4n+3) = 3r(2n+1) - 2r(n)$, $r(8n+1) = 3r(4n+1) - 2r(2n+1)$, and $r(8n+5) = 5r(2n+1) - 4r(n)$. Hence it follows that the module generated by the 2–kernel of $\{r(n)\}_{n\geq 0}$ is generated by its subsequences $\{r(n)\}_{n\geq 0}$, $\{r(2n+1)\}_{n\geq 0}$, and $\{r(4n+1)\}_{n\geq 0}$.

Example 7.

Let $d(0) = 0$, $d(1) = 1$, $d(2n) = d(n)$, and $d(2n+1) = d(n) + d(n+1)$. This sequence forms the numerators of the entries in the *Stern-Brocot tree* (see [St], [GKP]). It was also studied by de Rham [R] and is Sloane's sequence #56. The first few terms are

$$0, 1, 1, 2, 1, 3, 2, 3, 1, 4, 3, 5, 2, 5, 3, 4, \ldots.$$

It is easy to see that $d(4n+1) = d(n) + d(2n+1)$, and $d(4n+3) = 2d(2n+1) - d(n)$, and it follows that d is 2–regular. Also see [Dij, pp. 215–216, 230–232].

A similar sequence is given by $a(0) = 0$, $a(1) = 1$, $a(2n) = a(n)$, and $a(2n+1) = a(n+1) - a(n)$. It satisfies $a(4n+1) = a(2n+1) - a(n)$ and $a(4n+3) = a(n)$ and hence is 2–regular. See [Rez1], [Rez2].

Example 8.

Define $\nu_2(n)$ to be the exponent of the highest power of 2 that divides n. (This is essentially Sloane's sequence #51.) Then if $h(n) = \nu_2(n+1)$, we see that $h(2n) = 0$ and $h(2n+1) = h(n) + 1$. Thus $\{h(n)\}_{n\geq 0}$ is 2–regular.

Using Lemma 6.4, we find

$$\nu_2(n+1) = e_1(n) - (e_{10}(n) + e_{110}(n) + e_{1110}(n) + \cdots).$$

Example 9.

Using the remark after Theorem 3.1, we see that $\nu_2(n!) = \sum_{1\leq j\leq n} \nu_2(j)$ is 2–regular.

Example 10.

Let the binary expansion of an integer n be written as

$$\sum_{i\geq 0} b_i(n) 2^i,$$

where $b_i \in \{0,1\}$. Define $g(n) = \sum_{i\geq 0}(i+1)b_i(n)$. Then it is easy to see that $g(2n) = g(n) + e_1(n)$ and $g(2n+1) = g(n) + e_1(n) + 1$. Hence $\{g(n)\}_{n\geq 0}$ is 2–regular.

Using Lemma 6.4, we can compute the pattern transform of $g(n)$. We find

$$g(n) = \sum_{P \in 1(0+1)^*} e_P(n).$$

Example 11.

Let $f(n) = |n_{(2)}|$, i. e.

$$f(n) = \begin{cases} 0, & \text{if } n = 0; \\ 1 + \lfloor \log_2 n \rfloor, & \text{if } n \geq 1. \end{cases}$$

Then we easily see that $f(2n+1) = f(n)+1$, $f(4n) = 2f(2n) - f(n)$, and $f(4n+2) = f(n) + 2$. Hence the module generated by its 2-kernel is generated by $\{f(n)\}_{n\geq 0}$, $\{f(2n)\}_{n\geq 0}$, and the constant sequence 1. Using Lemma 6.4, we find

$$f(n) = e_1(n) + e_{10}(n) + e_{100}(n) + e_{1000}(n) + \cdots.$$

Example 12.

Let $\{B(n)\}_{n\geq 0}$ be the sequence $0, 3, 5, 6, 9, 10, 12, 15, \ldots$, the integers whose base–2 representation contains an even number of 1's. (This is Sloane's sequence #952). Let $u(n) = e_1(n) \bmod 2$; then $u(n)$ is the classical Thue-Morse sequence and $1 - u(n)$ is the characteristic sequence of $B(n)$. We easily prove that $B(2n) = 2B(n) - u(n)$ and $B(2n+1) = 2B(n) + 3(1 - u(n))$; hence B is 2–regular.

Example 13.

Let $\{C(n)\}_{n\geq 0}$ be the sequence of *Moser-de Bruijn* ([Mos], [B2]): $0, 1, 4, 5, 16, 17, 20, 21, \ldots$ It consists of integers that can be written as the sum of distinct powers of 4. This is Sloane's sequence #1315. Note that $C(2n) = 4C(n)$ and $C(2n+1) = 4C(n) + 1$; hence C is 2–regular. See [LMP]. In [BM] it is shown that the characteristic sequence of $C(n)$ gives a binary number such that its binary expansion and the binary expansion of its reciprocal are explicitly known. Its continued fraction is also explicitly known.

Similarly, the sequence of *Loxton-van der Poorten* [LP1]

$$0, 1, 3, 4, 5, 11, 12, 13, 15, 16, 17, 19, 20, 21, 43, 44, \ldots$$

of positive integers that can be represented in base 4 using only the digits $-1, 0, 1$ is 3–regular.

Example 14.

Let $G(n) = 2^{e_1(n)}$. This is *Gould's sequence* [G], and Sloane's sequence #109. It satisfies $G(2n) = G(n)$; $G(2n+1) = 2G(n)$ and hence is 2–regular.

Glaisher [Gl] showed that $G(n)$ counts the number of odd binomial coefficients in row n of Pascal's triangle.

More generally, let p be a prime and let $G_p(n)$ be the number of binomial coefficients in row n of Pascal's triangle which are not divisible by p. Then Fine [Fi] showed that

$$G_p(n) = \prod_{0\leq i\leq e} (a_i + 1)$$

where the base–p expansion of n is $a_e a_{e-1} \cdots a_1 a_0$. Of course, $G_p(n)$ is p–regular.

Now put $H_p(n) = \sum_{0\leq k\leq n} G_p(k)$. Then $H_p(n)$ is also p–regular. The sequences $H_2(n)$,

$$1, 3, 5, 9, 11, 15, 19, 27, 29, 33, 37, 45, 49, 57, \ldots$$

and $H_3(n)$,

$$1, 3, 6, 8, 12, 18, 21, 27, 36, 38, 42, 48, 52, 60, 72, \ldots$$

appear in [LM]. Also see [HLVVM], [LMVV].

Example 15.

Let $\{b(n)\}_{n\geq 0}$ be the sequence of numbers represented by binary Gray code [Gr], [Gi]:

$$0, 1, 3, 2, 6, 7, 5, 4, 12, 13, 15, 14, 10, 11, 9, 8, \cdots$$

Then it is easy to see that $b(4n) = 2b(2n)$, $b(4n+1) = 2b(2n)+1$, $b(4n+2) = 2b(2n+1)+1$, and $b(4n+3) = 2b(2n+1)$. Hence $\{b(n)\}_{n\geq 0}$ is 2–regular.

Similarly, if $\gamma(n)$ denotes the sum of the bits in the Gray code representation of n, then we find $\gamma(2n+1) = 2\gamma(n) - \gamma(2n) + 1$; $\gamma(4n) = \gamma(2n)$; and $\gamma(4n+2) = \gamma(2n+1)+1$. Hence $\{\gamma(n)\}_{n\geq 0}$ is 2–regular. See [FR].

Example 16.

Consider the sequence of lattice points $(x(n), y(n))$ traced out by paperfolding curves with an ultimately periodic sequence of unfolding instructions [DMFP, MFS]. Then $\{x(n)\}_{n\geq 0}$ and $\{y(n)\}_{n\geq 0}$ are 2–regular.

For example, consider the sequence of lattice points $(x(n), y(n))$ traced out by the space-filling curve with unfolding instructions `RLRLRL` $\cdots$.

$$\begin{array}{rcl} n & = & 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14\ 15\ \cdots \\ x(n) & = & 0\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 2\ 2\ \ 3\ \ \ 3\ \ \ 2\ \ \ 2\ \ \ 1\ \ \ 1\ \ \cdots \\ y(n) & = & 0\ 1\ 1\ 2\ 2\ 3\ 3\ 2\ 2\ 3\ \ 3\ \ \ 4\ \ \ 4\ \ \ 3\ \ \ 3\ \ \ 4\ \ \cdots \end{array}$$

Then the sequences satisfy the identities $x(0) = 0$, $x(2) = 1$, $x(2n+1) = x(2n)$, $x(4n) = 2x(n)$, $x(8n+2) = -2x(n) + 2x(2n) + x(4n+2)$, $x(16n+6) = 2x(n) + x(4n+2)$, $x(16n+14) = 2x(2n) + 2x(4n+2) - x(8n+6)$, and $y(0) = 0$, $y(1) = 1$, $y(4n) = 2y(n)$, $y(4n+1) = y(4n+2) = 2y(n) - y(2n) + y(2n+1)$, $y(8n+3) = y(8n+7) = 2y(2n+1)$.

Example 17.

Van der Corput's sequence $\varphi_2(n)$ is defined as follows [Cor]: if

$$n = \sum_{i\geq 0} b_i(n) 2^i,$$

where $b_i \in \{0, 1\}$, then

$$\varphi_2(n) = \sum_{i\geq 0} b_i(n) 2^{-i-1}.$$

We see that $\varphi_2(0) = 0$, $\varphi_2(2n) = \frac{1}{2}\varphi_2(n)$, and $\varphi_2(2n+1) = \frac{1}{2} + \frac{1}{2}\varphi_2(n)$. Hence the sequence of rational numbers $\varphi_2(n)$ is $(\mathbb{Q}, 2)$–regular.

Also note that $\varphi_2(n) = r(n)/2^{f(n)}$, where $r(n)$ is the sequence of Example 6 and $f(n)$ is the sequence of Example 11.

Halton [Hal] generalized van der Corput's sequence to bases $b \geq 2$.

Example 18.

Let

$$\frac{1}{(1-X)(1-X^2)\cdots(1-X^j)} = \sum_{n\geq 0} P_j(n) X^n.$$

Then $P_j(n)$ enumerates the number of partitions of n into j or fewer parts. The sequence $P_3(n)$ is Sloane's sequence #186; $P_4(n)$ is sequence #229; $P_5(n)$ is sequence #237, and $P_6(n)$ is sequence #243.

By Theorem 3.3, $P_j(n)$ is k–regular for all $j \geq 1$ and all $k \geq 2$. Note, however, that the function $P_\infty(n) = \lim_{j\to\infty} P_j(n)$, which counts the number of unrestricted partitions, is not k–regular, as it grows too quickly. Hence k–regular sequences are not closed under taking simple limits.

Example 19.

Let $w = w_0 w_1 w_2 \cdots$ be an infinite word over a finite alphabet, and define $s_w(n)$ to be the number of distinct subwords (European terminology: factors) of length n in w. Then $s_w(n) - s_w(n-1)$ is frequently k–automatic, and hence in these cases, $s_w(n)$ is k–regular. For example, this is true when w is the fixed point of the Toeplitz substitution given by $0 \to 0010$ and $1 \to 1010$ [Rau]; when w is the infinite word of Thue-Morse, the fixed point of the substitution given by $0 \to 01$ and $1 \to 10$ [Brl] [LV]; and in a more general class of infinite words given by iterated homomorphisms discussed by Tapsoba [Tap].

Example 20.

It is well-known that n is a sum of three squares if and only if n is not of the form $4^a(8k+7)$. It is easily seen that the sequence

$$t(n) = \begin{cases} 0, & \text{if } n = 4^a(8k+7); \\ 1, & \text{otherwise,} \end{cases}$$

is 2–automatic. Hence the sequence

$$Q(n) = \sum_{1 \leq k \leq n} t(k),$$

which counts the number of positive integers $\leq n$ that are the sum of three squares, is 2–regular. See [Sh], [OS], [W].

Example 21.

An *addition chain to* n is a sequence of pairs of positive integers

$$(a_1, b_1), (a_2, b_2), \ldots, (a_r, b_r)$$

where (i) $a_r + b_r = n$ and (ii) for all s, either $a_s = 1$, or $a_s = a_i + b_i$ for some $i < s$, and the same holds for b_s. The cost of the addition chain is $\sum_{1 \leq i \leq r} a_i b_i$. Denote the cost of the minimum cost addition chain to n as $c(n)$. Then it can be shown [GYY] that $c(1) = 0$, and $c(2n) = c(n) + n^2$, $c(2n+1) = c(n) + n^2 + 2n$ for $n \geq 1$. Hence $c(n)$ is 2–regular.

Example 22.

Define $b(d; n)$ as the number of representations

$$n = \sum_{i \geq 0} \epsilon_i 2^i,$$

where $0 \leq \epsilon_i < d$. Then $b(2;n) = 1$, $b(3;n) = d(n+1)$, where $d(n)$ is the sequence of Example 7, and $b(4;n) = 1 + \lfloor n/2 \rfloor$. (See [Rez3]). It is possible to show that $b(d;n)$ is 2–regular for all $d \geq 1$. However, $b(\infty;n) = \lim_{d\to\infty} b(d;n)$ is not k–regular, as it is known that

$$\log b(\infty;n) \sim \frac{1}{\log 4}(\log n)^2.$$

(See [Mah], [B1], [Kn1]). Note that if $f(X) = \sum_{n\geq 0}(-1)^{e_1(n)}X^n$, and $g(X) = \sum_{n\geq 0} b(\infty;n)X^n$, then $1/f(X) = g(X^2)(1+X)$, which shows that $f(X)$ is not invertible in the ring of 2–regular power series. P. Dumas has pointed out [Dum] that the sequence $\{b(\infty;n) \bmod 2^M\}_{n\geq 0}$ is 2–automatic for all $M \geq 0$.

Example 23.

Let $\nu_3(n)$ denote the exponent of the highest power of 3 that divides n, and $s_3(n)$ denote the sum of the digits of n when expressed in base 3.

Define $r(n) = \sum_{0\leq i<n} \binom{2i}{i}$. Then $\nu_3(r(n))$ is 3–regular. This follows from the (not-so-trivial) fact that

$$\nu_3(r(n)) = \nu_3(\binom{2n}{n}) + 2\nu_3(n)$$

and the (trivial) fact that

$$\nu_3(\binom{2n}{n}) = s_3(n) - \frac{1}{2}s_3(2n).$$

See [SS2].

Example 24.

As in Section VI, let

$$\lambda(n) = \begin{cases} 0, & \text{if } n = 0; \\ n - 2^{\lfloor \log_2 n \rfloor}, & \text{if } n > 0. \end{cases}$$

Then A. Liao (personal communication) asked for the solution $T(n)$ to the recurrence

$$T(n) = \lambda(n) + T(\lambda(n)),$$

where $f(0) = 0$. We see that $T(n)$ is 2–regular, as the identities $T(2n) = 2T(n)$, $T(4n+1) = 2T(n) + T(2n+1)$, and $T(4n+3) = -2T(n) + 3T(2n+1) + 1$ can easily be verified by induction.

$T(n)$ also has the following pleasant expansion as a sum of pattern sequences:

$$T(n) = \sum_{v(P)\geq 3} \left\lceil \frac{\lambda(P)}{2} \right\rceil e_P(n).$$

Example 25.

The earliest reference to a non-trivial class of k–regular sequences we have found is from an 1822 paper of Charles Babbage [Bab], in which he discusses sequences such as

$$\Delta^2 u_n = u_{n+1} \bmod 10,$$

"which is one of a class of equations never hitherto integrated." By considering both sides (mod 10), we see that the sequence $\{u_n \bmod 10\}_{n\geq 0}$ is ultimately periodic and therefore u_n is k–regular for all $k \geq 2$. This sequence was deemed to have "no intrinsic mathematical significance" by Dubbey [Dub, pp. 182].

Example 26.

Let $e(0) = 0$, $e(1) = 1$, and define $e(n)$ to be the least integer greater than $e(n-1)$ such that the sequence $e(0), \ldots, e(n)$ contains no three terms in arithmetic progression. The first few terms of this sequence are

$$0, 1, 3, 4, 9, 10, 12, 13, 27, 28, 30, 31, 36, 37, 39, 40, 81, \ldots$$

and in general the sequence consists of numbers that can be written as distinct powers of 3. (Compare Example 13.) We have $e(2n) = 3e(n)$ and $e(2n+1) = 3e(n) + 1$, and so $e(n)$ is 2–regular. See [ET] and [Guy, p. 114].

Example 27.

Let k be an integer ≥ 2, and put

$$f_k(n) = \sum_{1\leq i\leq n} \lfloor \log_k i \rfloor.$$

Then $f_k(n)$ is k–regular. In fact, we have

$$f_k(n) = (n+1)\lfloor \log_k n \rfloor - \frac{k^{\lfloor \log_k n \rfloor + 1} - k}{k-1}.$$

See [Kn3, Section 1.2.4, Exercise 42 (b)].

The number of comparisons required to sort n items in many sorting algorithms forms a 2–regular sequence. The following examples illustrate this:

Example 28.

Merge sort, given a list of n integers, proceeds as follows: first the left half of the list is sorted (recursively), then the right half is sorted, and finally the two halves are merged together. The number of comparisons needed to merge sort n items is given by $T(1) = 0$, and

$$T(n) = T(\lfloor n/2 \rfloor) + T(\lceil n/2 \rceil) + n - 1,$$

for $n \geq 2$, and it is not difficult to see that $T(n)$ is a 2–regular sequence.

The resulting sequence

$$0, 1, 3, 5, 8, 11, 14, 17, 21, \ldots$$

is Sloane's sequence #963. It was discussed by Levitt, Green, and Goldberg [LGG], who gave the following closed form:

$$T(n) = n\lceil \log_2 n \rceil - 2^{\lceil \log_2 n \rceil} + 1.$$

Also see [Kn2, Section 5.3.1, Equation (3)].

Example 29.

Let $c(n)$ denote the number of key comparisons used to sort n elements by Batcher's method (see, for example, [Knu2, Section 5.2.2]). The first few terms of this sequence are

$$0, 1, 3, 5, 9, 12, 16, 19, 26, 31, 37, 41, 48, 53, 59, 63, \ldots$$

Define $a(n) = c(n+1) - c(n)$. Then it is shown in [Knu2, Section 5.2.2, Exercises 14, 15] that $a(2n) = a(n) + \lfloor \log_2(2n) \rfloor$; $a(2n+1) = a(n) + 1$ and hence $a(n)$ is 2–regular. Hence $c(n)$ is 2–regular.

Example 30.

Let $F(n)$ denote the number of key comparisons in Ford-Johnson sorting. Here are the first few values of this sequence:

$$0, 1, 3, 5, 7, 10, 13, 16, 19, 22, 26, 30, 34, 38, 42, 46, 50, \ldots$$

It is Sloane's sequence #954. A. Hadian showed that

$$F(n) = \sum_{1 \le k \le n} \lceil \log_2 \frac{3n}{4} \rceil;$$

see [Kn2, Section 5.3.1]. It is easy to show that $\lceil \log_2 n \rceil$ is a 2–regular sequence. Then by Theorem 2.6 $\lceil \log_2 3n \rceil - 2 = \lceil \log_2 3n/4 \rceil$ is 2–regular. Finally, by Theorem 3.1, $F(n)$ must be 2–regular. Knuth [Kn2, Section 5.3.1, Exercise 14] gives the following "closed form" for $F(n)$:

$$F(n) = n\lceil \log_2 \frac{3n}{4} \rceil - \lfloor 2^{\lfloor \log_2 6n \rfloor}/3 \rfloor + \lfloor \frac{1}{2} \log_2 6n \rfloor.$$

Example 31.

Let $k(n)$ denote the maximum number of key comparisons used by list-merge sorting; see [Kn2, Section 5.2.4]. Here are the first few terms of this sequence

$$0, 1, 3, 5, 9, 11, 14, 17, 25, 27, 30, 33, 38, 41, 45, 49, \ldots$$

Then it is known that if the binary representation of n is

$$2^{e_1} + 2^{e_2} + \cdots + 2^{e_t},$$

then

$$k(n) = 1 - 2^{e_t} + \sum_{1 \le k \le t} (e_k + k - 1)2^{e_k}.$$

(See [Kn2, Section 5.2.4, Exercises 14]).

From this it is easy to see that $k(2n) = 2k(n) + 2n - 1$ for $n \ge 1$ and $k(2n+1) = k(2n) + e_1(n) + 2^{\nu_2(n)+1} - 1$ for $n \ge 1$. Hence $k(n)$ is a 2–regular sequence.

Example 32.

Similarly, the number of comparisons needed in many merging algorithms forms a 2–regular sequence. For example, let $M(m, n)$ denote the minimum number of comparisons to merge m things with n. Then

$$M(1, n) = \lceil \log_2(n+1) \rceil$$

and

$$M(2, n) = \lceil \log_2 \frac{7}{12}(n+1) \rceil + \lceil \log_2 \frac{14}{17}(n+1) \rceil.$$

(See [Kn2, Section 5.3.2].) While $M(1, n)$ is easily seen to be 2–regular, we can prove that $M(2, n)$ is 2–regular using Theorems 2.6 and 2.7.

Example 33.

In analysis of a greedy heuristic for a matching problem, Reingold and Tarjan [RT] define a function $f(n)$ for positive even arguments, and write

$$f(n) = \min_{\substack{2 \le t \le n-2 \\ t \text{ even} \\ \alpha \ge 1-\alpha-\beta > 0 \\ \beta \ge 1-\alpha-\beta > 0}} \{\alpha f(t) + \beta f(n-t)\}.$$

Later they show that

$$f(2n) = \begin{cases} \frac{2}{3} f(n), & \text{if } n \text{ is even;} \\ \frac{1}{3} f(n+1) + \frac{1}{3} f(n-1), & \text{if } n \text{ is odd.} \end{cases}$$

They also give the following explicit form for $f(2n)$:

$$f(2n) = 1 - \sum_{2 \le i \le n} 3^{-\lceil \log_2 i \rceil}.$$

It follows from this that $f(2n)$ is a $(\mathbb{Q}, 2)$–regular sequence. The first few values of this sequence are:

$$1, 2/3, 5/9, 4/9, 11/27, 10/27, 1/3, 8/27, \cdots$$

Example 34.

The *Josephus problem* is as follows: the numbers from 1 to n are written in a circle. Starting with the number 1, every $2nd$ number that remains is crossed off until only one is left. The "survivor" is denoted $J(n)$. The first few values of $J(n)$ are as follows:

$$1, 1, 3, 1, 3, 5, 7, 1, 3, 5, 7, 9, 11, 13, 15, \cdots.$$

This problem was discussed by Graham, Knuth, and Patashnik [GKP, pp. 8–16], who observed that $J(2n) = 2J(n) - 1$ and $J(2n+1) = 2J(n) + 1$ for $n \geq 1$. It follows that $J(n)$ is 2–regular.

The same problem, where 2 is replaced by k and the result is the first uncrossed-off number encountered when there are only $k-1$ numbers left, does not appear to be k-regular in general. See [GKP, pp. 79–81].

We are grateful to P. Dumas for pointing out this example.

Example 35.

We show that the sequence of primes $\{p(n)\}_{n\geq 0}$

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, \ldots$$

is not k–regular. Suppose it were. Then using Lemma 4.1, we see that $\{p(k^n)\}_{n\geq 0}$ must satisfy a linear recurrence. Then if

$$\lim_{n\to\infty} \frac{p(k^n)}{nk^n}$$

exists, it must be an algebraic number. But from the prime number theorem,

$$\lim_{n\to\infty} \frac{p(k^n)}{nk^n} = \log k,$$

which is transcendental, a contradiction.

VIII. Some Open Problems.

1. Let $R' = \mathbb{Q}$. Prove that if $\{S(n)\}_{n\geq 0}$ is k_1-regular and k_2-regular, and k_1 and k_2 are multiplicatively independent, then the associated power series $\sum_{n\geq 0} S(n)X^n \in \mathbb{Q}[[X]]$ is a rational function. In the case where R' is finite, this is a result of Cobham [Cob2].

2. Determine all the units of the ring of k-regular power series.

3. Obtain transcendence results for the real numbers $\sum_{n\geq 0} S(n)p^{-n}$, where $S(n)$ is p-regular and $\sum_{n\geq 0} S(n)X^n$ is not a rational function. See [LP2].

IX. Acknowledgments.

We are grateful to O. Salon for suggesting the term k–kernel [Sa].

Part of this work was done while the first author was visiting Dartmouth College.

Part of this work was done while the second author was a visiting scientist at the University of Waterloo and a visiting professor at the University of Wisconsin, Madison.

A preliminary version of this paper was presented at the Symposium on Theoretical Aspects of Computer Science (STACS) in Rouen, France, on February 24, 1990.

The second author acknowledges with thanks conversations with E. Bach, C. Choffrut, L. Dickey, J. Driscoll, G. Frandsen, D. Joseph, S. Kurtz, R. Lipton, A. Lubiw, D. Passman, E. Reingold, N. J. A. Sloane, and C. Reutenauer.

References

[A1] J.-P. Allouche, Automates finis en théorie des nombres, *Expo. Math.* **5** (1987), 239-266.

[A2] J.-P. Allouche, Somme des chiffres et transcendance, *Bull. Soc. Math. France* **110** (1982), 279-285.

[A3] J.-P. Allouche, Suites infinies à répétitions bornées, *Séminaire de Théorie des Nombres de Bordeaux*, 1983-1984, Exposé no. 20.

[A4] J.-P. Allouche, Note sur un article de Sharif et Woodcock, *Sém. de Théorie des Nombres de Bordeaux*, 2^e série, **1** (1989), 163-187.

[AMS] J.-P. Allouche, P. Morton, and J. Shallit, Pattern spectra, substring enumeration, and automatic sequences, preprint.

[B1] N. G. de Bruijn, On Mahler's partition problem, *Indag. Math.* **10** (1948), 210-220

[B2] N. G. de Bruijn, Some direct decompositions of the set of integers, *Math. Tables Aids Comput.* **18** (1964), 537-546.

[Bab] C. Babbage, Observations on the application of machinery to the computation of mathematical tables, *Memoirs of the Astronomical Society* **1** (1822), 311-314; reprinted in *Babbage's Calculating Engines*, Tomash Publishers, Los Angeles, 1982.

[BM] A. Blanchard and M. Mendès France, Symétrie et transcendance, *Bull. Sci. Math.* **106** (1982), 325-335.

[BR] J. Berstel and C. Reutenauer, *Rational Series and Their Languages*, Springer-Verlag, 1988.

[Brl] S. Brlek, Enumeration of factors in the Thue-Morse word, *Disc. Appl. Math.* **24** (1989), 83-96.

[BS] R. Bellman and H. N. Shapiro, On a problem in additive number theory, *Ann. Math.* **49** (1948), 333-340.

[CKMR] G. Christol, T. Kamae, M. Mendès France and G. Rauzy, Suites algébriques, automates et substitutions, *Bull. Soc. Math. France* **108** (1980), 401-419.

[CL] G. Clements and B. Lindström, A sequence of (± 1)-determinants with large values, *Proc. Amer. Math. Soc.* **16** (1965), 548-550.

[Cob] A. Cobham, Uniform tag sequences, *Math. Systems Theory* **6** (1972), 164-192.

[Cob2] A. Cobham, On the base-dependence of sets of numbers recognizable by finite automata, *Math. Systems Theory* **3** (1969), 186-192.

[Coq] J. Coquet, A summation formula related to the binary digits, *Invent. Math.* **73** (1983), 107-115.

[Cor] J. C. van der Corput, Verteilungsfunktionen, *Proc. Ned. Akad. v. Wet.* **38** (1935), 813-821.

[CS] C. Choffrut and M. P. Schützenberger, Counting with rational functions, *Theor. Comput. Sci.* **58** (1988), 81-101.

[CY] P. Cheo and S. Yien, A problem on the k-adic representations of positive integers, *Acta. Math. Sinica* **5** (1955), 433-438.

[Dav] M. Davis, Hilbert's tenth problem is unsolvable, *Amer. Math. Monthly* **80** (1973), 233-269.

[Dij] E. W. Dijkstra, *Selected Writings on Computing: a Personal Perspective*, *Springer-Verlag*, New York, 1982.

[DMFP] M. Dekking, M. Mendès France, and A. van der Poorten, FOLDS!, *Math. Intell.* **4** (1982), 130-138; 173-181; 190-195. (Errata in *Math. Intell.* **5** (1983), 5.)

[Dub] J. M. Dubbey, *The Mathematical Work of Charles Babbage*, Cambridge University Press, 1978.

[Dum] P. Dumas, Suite automatiques à valeurs dans un anneau commutatif, preprint.

[E] S. Eilenberg, *Automata, Languages, and Machines*, Volume A, Academic Press, 1974.

[ET] P. Erdös and P. Turán, On some sequences of integers, *J. Lond. Math. Soc.* **11** (1936), 261-264.

[Fi] N. J. Fine, Binomial coefficients modulo a prime, *Amer. Math. Monthly* **54** (1947), 589-592.

[FR] P. Flajolet and L. Ramshaw, A note on Gray code and odd-even merge, *SIAM J. Comput.* **9** (1980), 142-158.

[G] H. W. Gould, Exponential binomial coefficient series, Technical Report 4, Department of Mathematics, W. Virginia Univ., September 1961.

[Gi] E. Gilbert, Gray codes and paths on the n-cube, *Bell Sys. Tech. J.* **37** (1958), 815-826.

[GKP] R. Graham, D. Knuth, and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, 1989.

[Gl] J. W. L. Glaisher, On the residue of a binomial-theorem coefficient with respect to a prime modulus, *Quart. J. Pure Appl. Math.* **30** (1899), 150-156.

[Gr] F. Gray, U. S. patent 2,632,058, March 17, 1953 (Filed November 13, 1947).

[GS] S. Ginsburg and E. H. Spanier, Bounded ALGOL-like languages, *Trans. Amer. Math. Soc.* **113** (1964), 333-368.

[Guy] R. K. Guy, *Unsolved Problems in Number Theory*, Springer-Verlag, 1981.

[GYY] R. Graham, A. Yao, and F. Yao, Addition chains with multiplicative cost, *Disc. Math.* **23** (1978), 115-119.

[Hal] J. H. Halton, On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals, *Numer. Math.* **2** (1960), 84-90.

[Har1] M. Harrison, *Lectures on Linear Sequential Machines*, Academic Press, 1969.

[HLVVM] N. S. Holter, A. Lakhtakia, V. K. Varadan, V. V. Varadan, and R. Messier, On a new class of planar fractals: the Pascal-Sierpinski gaskets, *J. Phys. A: Math. Gen.* **19** (1986), 1753-1759.

[IMO] 29th International Math Olympiad–Solutions, *Math. Mag.* **62** (1989), 212.

[JK] K. Jacobs and M. Keane, 0-1-Sequences of Toeplitz type, *Z. Wahrscheinlichkeitstheorie verw. Geb.* **13** (1969), 123-131.

[Kn1] D. E. Knuth, An almost linear recurrence, *Fib. Quart.* **4** (1966), 117-128.

[Kn2] D. E. Knuth, *Sorting and Searching*, The Art of Computer Programming, V. 3, Addison-Wesley, 1973.

[Kn3] D. E. Knuth, *Fundamental Algorithms*, The Art of Computer Programming, V. 1, Addison-Wesley, 1973.

[Lan] S. Lang, *Algebra*, Addison-Wesley, 1971.

[LGG] K. N. Levitt, M. W. Green, and J. Goldberg, A study of the data commutation problems in a self-repairable multiprocessor, *Proc. AFIPS Conf.* **32** (1968), 515-527.

[Li] P. Liardet, Automata and generalized Rudin-Shapiro sequences, *Sem. Salzburg Universität*, 1986; Publication 23, U. R. A. #225, C. N. R. S., Marseille, 1989.

[LM] A. Lakhtakia and R. Messier, Self-similar sequences and chaos from Gauss sums, *Comput. & Graphics* **13** (1989), 59-62.

[LMP] D. H. Lehmer, K. Mahler, and A. J. van der Poorten, Integers with digits 0 or 1, *Math. Comp.* **46** (1986) 683-689.

[LMVV] A. Lakhtakia, R. Messier, V. K. Varadan, and V. V. Varadan, Fractal sequences derived from the self-similar extensions of the Sierpinski gasket, *J. Phys. A: Math. Gen.* **21** (1988), 1925-1928.

[LP1] J. H. Loxton and A. J. van der Poorten, An awful problem about integers in base four, *Acta Arithmetica* **49** (1987), 193-203.

[LP2] J. H. Loxton and A. J. van der Poorten, Arithmetic properties of automata: regular sequences, *J. reine angew. Math.* **392** (1988), 57-69.

[LV] A. de Luca and S. Varricchio, Some combinatorial properties of the Thue-Morse sequence and a problem in semigroups, *Theor. Comput. Sci.* **63** (1989), 333-348.

[LZ] R. Lipton and Y. Zalcstein, Word problems solvable in logspace, *J. ACM* **24** (1977), 522-526.

[Mah] K. Mahler, On a special functional equation, *J. Lond. Math. Soc.* **15** (1940), 115-123.

[Mah2] K. Mahler, On the generating function of the integers with a missing digit, *J. Indian Math. Soc.* **15** (1951), 33-40.

[Mau] C. Mauduit, Substitutions et ensembles normaux, Habilitation, Marseille, 1989.

[MFP] M. Mendès France and A. J. van der Poorten, From geometry to Euler identities, *Theor. Comput. Sci.* **65** (1989), 213-220.

[MFS] M. Mendès France and J. Shallit, Wire bending, *J. Combinatorial Theory, A* **50** (1989) 1-23.

[MM] P. Morton and W. Mourant, Paper folding, digit patterns, and groups of arithmetic fractals, *Proc. Lond. Math. Soc.* **59** (1989), 253-293.

[Mos] L. Moser, An application of generating series, *Math. Mag.* **35** (1962), 37-38.

[N] D. J. Newman, On the number of binary digits in a multiple of three, *Proc. Amer. Math. Soc.* **21** (1969), 719-721.

[NS] D. J. Newman and M. Slater, Binary digit distribution over naturally defined sequences, *Trans. Amer. Math. Soc.* **213** (1975), 71-78.

[OS] A. H. Osbaldestin and P. Shiu, A correlated digital sum problem associated with sums of three squares, *Bull. Lond. Math. Soc.* **21** (1989), 369-374.

[Pro] H. Prodinger, Non-repetitive sequences and Gray code, *Disc. Math.* **43** (1983), 113-116.

[R] G. de Rham, Un peu de mathématiques à propos d'une courbe plane, *Elem. Math.* **2** (1947), 73-77; 89-97.

[Rau] G. Rauzy, Suites à termes dans un alphabet fini, *Sém. Théorie des Nombres de Bordeaux*, 2^e série, 1982-3, 25.01–25.16.

[Rez1] B. Reznick, A new sequence with many properties, *Abs. Amer. Math. Soc.* **5** (1984), 16.

[Rez2] B. Reznick, Some extremal problems for continued fractions, *Ill. J. Math.* **29** (1985), 261-279.

[Rez3] B. Reznick, Some binary partition functions, in B. C. Berndt, H. G. Diamond, H. Halberstam, and A. Hildebrand, eds., *Analytic Number Theory (Proceedings of a Conference in Honor of Paul T. Bateman)*, Birkhäuser, Boston, 1990, pp. 451-477.

[RT] E. M. Reingold and R. E. Tarjan, On a greedy heuristic for complete matching, *SIAM J. Comput.* **10** (1981), 676-681.

[Sa] O. Salon, Quelles tuiles! (Pavages apériodiques du plan et automates bidimensionnels), *Sém. de Théorie des Nombres de Bordeaux*, (2) **1** (1989), 1-25.

[Sch] M. P. Schützenberger, On the definition of a family of automata, *Information and Control* **4** (1961), 245-270.

[Sh] P. Shiu, Counting sums of three squares, *Bull. Lond. Math. Soc.* **20** (1988), 203-208.

[Sl] N. J. A. Sloane, *A Handbook of Integer Sequences*, Academic Press, 1973.

[SS] A. Salomaa and M. Soittola, *Automata-theoretic Aspects of Formal Power Series*, Springer-Verlag, 1978.

[SS2] N. Strauss and J. Shallit, Advanced Problem 6625, *Amer. Math. Monthly* **97** (1990), 252.

[St] M. A. Stern, Über eine zahlentheoretische Funktion, *J. reine angew. Math.* **55** (1858), 193-220.

[Tap] T. Tapsoba, Complexité de suites automatiques, Thèse de troisième cycle, Université d'Aix-Marseille II, 1987.

[W] S. S. Wagstaff, Jr., The Schnirelmann density of the sums of three squares, *Proc. Amer. Math. Soc.* **52** (1975), 1-7.

Last revision: May 21, 1991