On Kaprekar's Junction Numbers

Max A. Alekseyev
The George Washington University,
Department of Mathematics,
Washington, DC 20052, USA
Email: maxal@gwu.edu

and

N. J. A. Sloane[1]
The OEIS Foundation, Inc.,
11 So. Adelaide Ave.,
Highland Park, NJ 08904, USA
Email: njasloane@gmail.com

Abstract

A base b junction number u has the property that there are at least two ways to write it as $u = v + s(v)$, where $s(v)$ is the sum of the digits in the expansion of the number v in base b. For the base 10 case, Kaprekar in the 1950's and 1960's studied the problem of finding $K(n)$, the smallest u such that the equation $u = v + s(v)$ has exactly n solutions. He gave the values $K(2) = 101$, $K(3) = 10^{13} + 1$, and conjectured that $K(4) = 10^{24} + 102$. In 1966 Narasinga Rao gave the upper bound $10^{1111111111124} + 102$ for $K(5)$, as well as upper bounds for $K(6)$, $K(7)$, $K(8)$, and $K(16)$. In the present work, we derive a set of recurrences which determine $K(n)$ for any base b and in particular imply that these conjectured values of $K(n)$ are correct. The key to our approach is an apparently new recurrence for $F(u)$, the number of solutions to $u = v + s(v)$. We illustrate our method by computing the values of $K(n)$ for $n \leq 16$ and bases $b \leq 10$, and show that for each base $K(n)$ grows as a tower of height proportional to $\log_2(n)$. Rather surprisingly, the values of $K(n)$ for the base 5 problem are determined by the classical Thue–Morse sequence, which leads us to define generalized Thue–Morse sequences for other bases.

Keywords: Junction numbers, self-numbers, Colombian numbers, Kaprekar, Thue–Morse sequence.

2010 *Mathematics Subject Classification*: Primary 11A63; Secondary: 11B37.

1 Introduction

For a fixed base $b \geq 2$, let $s(v)$ denote the sum of the digits in the base b expansion of $v \in \mathbb{N} = \{0, 1, 2, \ldots\}$, and let $f(v) := v + s(v)$. Sequences

[1]Corresponding author.

that arise by iterating f have a long history [1, 15] (the latter reference has an extensive bibliography). In the 1950's and 1960's, Dattaraya Ramchandra Kaprekar in a series of self-published booklets [5, 7, 8, 9, 10] studied the inverse mapping to f in the base 10 case. Let $\mathrm{Gen}(u) := \{v \in \mathbb{N} \mid f(v) = u\}$ and $F(u) := |\mathrm{Gen}(u)|$. Kaprekar called the elements of $\mathrm{Gen}(u)$ the *generators* of u, and in 1956 [6] defined a *self-number* to be any number u with $F(u) = 0$. The first few self-numbers (in base 10) are

$$1, 3, 5, 7, 9, 20, 31, 42, 53, 64, 75, 86, 97, 108, 110, 121, 132, 143, 154, 165, 176, \ldots \quad (1)$$

(A003052).[2] Self-numbers are also known as *Colombian numbers*, after a problem proposed by Recamán in 1973 [13].

Kaprekar called numbers with at least two generators *junction numbers*. The smallest junction number (again in base 10) is 101, which has generators 91 and 100, and the first few junction numbers are

$$101, 103, 105, 107, 109, 111, 113, 115, 117, 202, 204, 206, 208, 210, 212, 214, \ldots$$

(A230094).

Kaprekar was particularly interested in finding what we will call $K(n)$, the smallest number with n generators, which is the subject of the present paper. We will show that the sequence $(K(n))_{n\geq 1}$ begins

$$\begin{aligned} &0,\ 101,\ 10^{13}+1,\ 10^{24}+102,\ 10^{1111111111124}+102,\ 10^{2222222222224}+10^{13}+2,\\ &\quad 10^{(10^{24}+10^{13}+115)/9}+10^{13}+2,\ 10^{(2\cdot 10^{24}+214)/9}+10^{24}+103,\\ &\quad 10^{(10^{1111111111124}+10^{24}+214)/9}+10^{24}+103,\ \ldots \end{aligned} \quad (2)$$

(A006064).[3] It is easy to check by hand that $K(2) = 101$, and with today's computers it is easy to verify $K(3) = 10^{13}+1$ by direct search. As to what was known by Kaprekar and his colleagues more than fifty years ago, the various accounts given by Kaprekar [8], Narasinga Rao [11], and Gardner [3] do not quite agree, but it seems that Kaprekar believed that he had proved that $K(3) = 10^{13} + 1$, that the putative value $10^{24} + 102$ for $K(4)$ was discovered independently by Kaprekar and Professor Gunjikar in 1961, and that Kaprekar was convinced that it was the true value of $K(4)$ and not just an upper bound. Gardner [3] reports in 1975 that Kaprekar told him that he had also found what he conjectured to be the values of $K(5)$ and $K(6)$. Kaprekar's work on this problem is also discussed by Schorn [14].

However, Narasinga Rao, writing in 1963 [11] (although not published until 1966), states things slightly differently. He gives a recipe for finding junction numbers with a specified number of generators, improving on an earlier recipe

[2] Throughout this article, six-digit numbers prefixed by A refer to entries in the OEIS [12].

[3] The reader may detect a pattern in these numbers, but should be warned that it breaks down after a while. See Tables 2 and 11.

of Kaprekar's, and gives Kaprekar's value of $K(3) = 10^{13} + 1$. He then *conjectures* that $K(4) = 10^{24} + 102$, and gives as a candidate for $K(5)$ the value $10^{1111111111124} + 102$, remarking that no much smaller value is likely to exist. (Narasinga Rao's recipe does not necessarily produce the smallest junction number with a given number of generators.) He also gives upper bounds for $K(6)$, $K(7)$, $K(8)$, and $K(16)$. Remarkably enough, all of Narasinga Rao's upper bounds turn out to be the true values for these $K(n)$. We return to the base 10 case in Section 9.

The principal goal of this paper is to present a set of recurrences which generate the sequence $(K(n))_{n\geq 1}$ for any base b (see Section 6, and in particular Theorems 18 and 20). These recurrences depend upon an apparently new recurrence for $F(u)$, discussed in Section 3.

Because the values of $K(n)$ for small b and n are both easy to determine and somewhat exceptional, we start by discussing the cases $n \leq 3$ in Section 4 and bases $b = 2$ and $b = 3$ in Section 5. Bases 2 and 3 are also exceptional since for them the recurrences for $K(n)$ are quite simple and can be obtained without the machinery developed in Section 6. Tables 1 and 2 in Section 3 collect the numerical values of $K(n)$ for $n \leq 7$ and bases $b \leq 10$.

Section 6 gives the recurrences for a general base b. Sections 8 and 9 apply the results of Section 6 to bases $b \in \{4, 5, 7, 10\}$. In general, the calculation of $K(n)$ involves a subsidiary sequence $(\tau(n))_{n\geq 1}$ of integers in the range $[0, b-2]$ when b is even, or $[0, \frac{b-3}{2}]$ when b is odd. For $b = 5$, it turns out that $(\tau(n))_{n\geq 1}$ is essentially the classical Thue–Morse sequence[4] A010060, and so, for any base, we refer to $(\tau(n))_{n\geq 1}$ as a "generalized Thue–Morse sequence". For example, for both bases $b = 4$ and $b = 7$ we obtain the ternary sequence shown in (81), and for base $b = 10$ the sequence (84).

Section 10 (and in particular Conjecture 28) discusses the rate of growth of $K(n)$ as a function of n. We begin by applying the recurrence for $F(u)$ from Section 3 to establish some general bounds on $K(n)$. Then we consider the representation of $K(n)$ as a tower of exponentials. For bases $b \neq 3$, it appears that $K(n)$ is a tower

$$K(n) \;=\; b^{b^{\cdot^{\cdot^{\cdot^{b^{\omega(n)}}}}}} \,, \tag{3}$$

with $0 < \omega(n) \leq 1$, of height[5] $\lceil \log_2(n) \rceil + \lambda$, where $\lambda = 3$ if $b = 2$, $\lambda = 2$ if $b \geq 4$ is even, and $\lambda = 1$ if $b \geq 5$ is odd (base $b = 3$ is slightly exceptional).

Notation. We will always work in a fixed base $b \geq 2$. If the base b expansion of $v \in \mathbb{N} := \{0, 1, 2, \ldots\}$ is

$$v \;=\; \sum_{i=0}^{k-1} v_i b^i \tag{4}$$

[4]This provides yet another illustration of the ubiquity of this sequence [2].
[5]The height of a tower is defined in (87) in Section 10.

(where $0 \le v_i < b$, $k = \lceil \log_b(v+1) \rceil$, and $v_{k-1} \neq 0$ unless $v = 0$), we refer to the v_i as "digits", even if $b \neq 10$, and say that v has "length" k. We will also write v as

$$v = [v_{k-1}, v_{k-2}, \ldots, v_1, v_0]_b\,, \tag{5}$$

where we omit the commas between the digits if there is no possibility of confusion. In this notation, $s(v) = \sum_{i=0}^{k-1} v_i$ and $f(v) = \sum_{i=0}^{k-1} v_i(b^i + 1)$.

As already mentioned, for $u \in \mathbb{N}$, we let $\mathrm{Gen}(u) := \{v \in \mathbb{N} \mid f(v) = u\}$, $F(u) := |\,\mathrm{Gen}(u)\}|$, and for $u < 0$ we set $\mathrm{Gen}(u) := \emptyset$ (the empty set) and $F(u) := 0$. For $n \ge 1$, $K(n)$ is defined to be the smallest $u \in \mathbb{N}$ such that $F(u) = n$. For $n \ge 2$, it will turn out that the leading base-b digit of $K(n)$ is 1 (see Theorem 8), and we define $E(n)$ for $n \ge 2$ by $K(n) = b^{E(n)} +$ terms of smaller order (in other words, $E(n) + 1$ is the length of $K(n)$ in base b). In Tables 1 and 2 we write $E_b(n)$ to indicate the value of b. For use in Section 6, in Section 4 we also define $K_i(n)$ to be the smallest $u \in \mathbb{N}$ such that $F(u) = n$ and $u \equiv i \pmod{b-1}$. Of course, the functions s, f, F, Gen, K, K_i all depend on the value of b, but to indicate this with subscripts would have made the equations unnecessarily complicated, so we hope the value of b will always be clear from the context.

In summary, the principal symbols are

b	:	base, ≥ 2,
$s(v)$	:	sum of base-b digits of v,
$f(v)$	:	$v + s(v)$,
$\mathrm{Gen}(u)$	:	set of v such that $f(v) = u$,
$F(u)$	:	number of v such that $f(v) = u$,
$K(n)$	:	smallest u such that $F(u) = n$,
$K_i(n)$	:	smallest u such that $F(u) = n$ and $u \equiv i \pmod{b-1}$,
$E(n)$	:	for $n \ge 2$, $K(n) = b^{E(n)} +$ terms of smaller order,
$\mathcal{I}$	:	subset of the residue classes modulo $b - 1$ defined by (18),
$J(n)$	:	subset of $\mathcal{I}$ defined by (70),

where $b, u, v, n \in \mathbb{N}$.

2 Preliminary results

We begin with some elementary lemmas.

Lemma 1. *(i) If b is odd, then $f(v)$ is even for any $v \in \mathbb{N}$, and so $F(u) = 0$ if u is odd.*

(ii) If b is even, then $F(u) = 0$ if u is odd and $u < b$.

Proof. (i) If b is odd and v is given by (4) then $f(v) = \sum_i v_i(b^i + 1)$ is even.

(ii) Numbers below b have at most one generator, v (say), for which $f(v) = 2v$ is even. □

The next lemma says that if $u = f(v)$, then v is smaller than u, but not too much smaller. This is useful when making computer searches.

Lemma 2. *If $u \geq 2, v \in \mathbb{N}$ satisfy $v + s(v) = u$, then*

$$u - (b-1)\lceil \log_b(u) \rceil \leq v \leq u - 1\,. \tag{6}$$

Proof. Since $u \geq 2$, $v \geq 1$, $s(v) \geq 1$, and so $v \leq u - 1$. Since the length of v is $k = \lceil \log_b(v+1) \rceil \leq \lceil \log_b(u) \rceil$, we have $s(v) \leq (b-1)\lceil \log_b(u) \rceil$. □

The third lemma is a generalization of the observation that the Hamming weight of $2^m - 1 - v$ is equal to $m -$ Hamming weight(v). We omit the proof.

Lemma 3. *If $m \geq 0$, $1 \leq c \leq b-1$, and $0 \leq v \leq cb^m - 1$, then*

$$s(cb^m - 1 - v) \;=\; (b-1)m + c - 1 - s(v)\,. \tag{7}$$

For example, in base 10, $s(281) = s(3 \cdot 10^2 - 1 - 18) = 9 \cdot 2 + 2 - 9 = 11$.

Lemma 4. *Let $n \geq 2$ be an integer.*

(i) If $(a(i))_{i\geq 1}$ is a sequence of nonnegative real numbers such that $a(m+1) \geq 2a(m)$ for all $m = 1, 2, \ldots, n-2$, then

$$\min_{1\leq i\leq n-1} a(i) + a(n-i) \;=\; a\left(\left\lceil \tfrac{n}{2} \right\rceil\right) + a\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right) \quad \text{for } n \geq 2\,. \tag{8}$$

Moreover, if $a(2) > 0$ then $i \in \left\{\left\lfloor \frac{n}{2} \right\rfloor, \left\lceil \frac{n}{2} \right\rceil\right\}$ are the only values of i that attain the minimum in (8).

(ii) If $(a_1(i))_{i\geq 1}$ and $(a_2(i))_{i\geq 1}$ are a pair of sequences of nonnegative real numbers such that

$$a_1(m+1) \geq a_1(m) + a_2(m) \quad \text{and} \quad a_2(m+1) \geq a_1(m) + a_2(m) \tag{9}$$

for all $m = 1, 2, \ldots, n-2$, then

$$\begin{aligned}&\min_{1\leq i\leq n-1} a_1(i) + a_2(n-i)\\ &\qquad = \min\left\{a_1\left(\left\lceil \tfrac{n}{2} \right\rceil\right) + a_2\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right), a_1\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right) + a_2\left(\left\lceil \tfrac{n}{2} \right\rceil\right)\right\}\,.\end{aligned} \tag{10}$$

Moreover, if $a_1(2) > 0$ and $a_2(2) > 0$ then $i = \left\lfloor \frac{n}{2} \right\rfloor$ and $i = \left\lceil \frac{n}{2} \right\rceil$ are the only values of i that attain the minimum in the left-hand side of (10).

Proof. (i) Suppose $n = 2t$. Then $2a(t) \leq a(t-1) + 2a(t) \leq a(t-1) + a(t+1) \leq 2a(t+1) \leq a(t+2) \leq a(t-2) + a(t+2) \leq \cdots \leq a(1) + a(n-1)$. If $n = 2t+1$ is odd, then $a(t) + a(t+1) \leq 2a(t+1) \leq a(t+2) \leq a(t-1) + a(t+2) \leq \cdots \leq a(1) + a(n-1)$. If $a(2) > 0$, then every second inequality in the inequality chains above is strict, implying that the minimum of $a(i) + a(n-i)$ is attained only at $i = t$ when $n = 2t$, and only at $i = t$ or $i = t+1$ when $n = 2t+1$.

Part (ii) is proved similarly. □

We call a sequence $(a(i))_{i\geq 1}$ satisfying (8) for all $n \geq 2$ a *sequence of exponential type*, and we say that sequences $(a_1(i))_{i\geq 1}$ and $(a_2(i))_{i\geq 1}$ satisfying (10) for all $n \geq 2$ form a *pair of sequences of exponential type*.

The following lemma introduces a representation of integers that plays a central role in our study.

Lemma 5. *Let $b \geq 2$ be a fixed integer.*

(i) Every integer $u > b$ has a unique representation in the form

$$u \;=\; c(b^m + 1) + k\,, \tag{11}$$

where $m \geq 1$, $1 \leq c \leq b-1$, and

$$\begin{cases} 0 \leq k \leq b^m, & \text{if } c < b-1; \\ 0 \leq k \leq b^m - b + 1, & \text{if } c = b-1. \end{cases} \tag{12}$$

(ii) Let $u = c(b^m + 1) + k$ and $u' = c'(b^{m'} + 1) + k'$ be positive integers represented as in (11). Suppose that $u \leq u'$. Then $m \leq m'$. Furthermore, if $m = m'$, then $c \leq c'$. Finally, if $m = m'$ and $c = c'$, then $k \leq k'$.[6]

Proof. It is not hard to see that the representation (11) uniquely defines $m = \lfloor \log_b(u-1) \rfloor$, $c = \left\lfloor \frac{u}{b^m+1} \right\rfloor$, and $k = u - c(b^m + 1)$. The value of m is a non-decreasing function of u, and so is c when m is fixed, and so is k when m and c are fixed. □

The following examples illustrate how numbers are represented in the form (11):

- for $b + 1 \leq u \leq b^2$, we have $m = 1$, $c = \left\lfloor \frac{u}{b+1} \right\rfloor$, and $k = u - (b+1)c$;
- for $u = b^r$ with $r \geq 2$, we have $m = r-1$, $c = b-1$, and $k = b^{r-1} - b + 1$.

We will also need the following technical lemma.

Lemma 6. *For any integers $b \geq 2$ and $m \geq 2$,*

$$(b-1)m - 2 \leq b^m - b + 1$$

and

$$b^m \geq \frac{1}{2} b^m + 2(b-1) \geq b^{m-1} + 2(b-1)\,.$$

We omit the elementary proof.

[6] In other words, if $u \leq u'$, the triple (m, c, k) is lexicographically smaller than the triple (m', c', k').

3 The recurrence for $F(u)$

The following result, in particular the recurrence (14), is the key to the whole paper.

Theorem 7. *We have* $\mathrm{Gen}(u) = \emptyset$ *and* $F(u) = 0$ *when* $u < 0$ *or* $u = 1$, *and* $\mathrm{Gen}(0) = \{0\}$, $F(0) = 1$. *For* $u \geq 2$, *consider its representation* $u = c(b^m + 1) + k$ *defined in Lemma 5. Then*[7]

$$\mathrm{Gen}(u) \;=\; \{cb^m + v \mid v \in \mathrm{Gen}(k)\} \cup \{cb^m - 1 - v \mid v \in \mathrm{Gen}((b-1)m - k - 2)\} \tag{13}$$

and

$$F(u) \;=\; F(k) \;+\; F((b-1)m - k - 2)\,. \tag{14}$$

Proof. The first assertion is clear (there is no $v \in \mathbb{N}$ such that $f(v) = 1$), and (14) follows at once from (13). To prove (13), let $u = c(b^m + 1) + k$ as in Lemma 5. We will show that (i) any element of the right-hand side of (13) is a generator of u, and (ii) every generator of u is an element of the right-hand side of (13).

(i) Suppose $v \in \mathrm{Gen}(k)$. Since $k \leq b^m$, $v < b^m - 1$ (by (6)), we have $s(cb^m + v) = c + s(v)$, and $f(cb^m + v) = cb^m + v + c + s(v) = c(b^m + 1) + k = u$. On the other hand, suppose $v \in \mathrm{Gen}((b-1)m - k - 2)$. Let $w = cb^m - 1 - v$. By Lemma 3, $s(w) = (b-1)m + c - 1 - s(v)$ (the condition $v \leq cb^m - 2$ follows from $v < (b-1)m - k - 2$). Then $f(w) = w + s(w) = cb^m - 1 - v + (b-1)m + c - 1 - s(v) = c(b^m + 1) + k = u$.

(ii) Suppose w is a generator for $u = c(b^m + 1) + k$. Clearly, $u \leq b^{m+1}$, and $u = b^{m+1}$ only when $c = b - 1$ and $k = b^m - b + 1$. Trivially, either $w \geq cb^m$ or $w < cb^m$.

First, suppose $w \geq cb^m$ and write it as $w = cb^m + v$. If $v < b^m$ then $s(w) = c + s(v)$, and $w + s(w) = u$ implies $v + s(v) = k$ and $v \in \mathrm{Gen}(k)$. If $v \geq b^m$ then if $c = b - 1$, $w \geq b^{m+1} \geq u$, contradicting (6). So $c \leq b - 2$ and $w = (c + 1)b^m + \mu$, where $\mu = v - b^m \geq 0$, which implies $u \geq (c + 1)b^m$, that is, $u = (c + 1)b^m + \lambda$, where $\lambda = c + k - b^m \leq c$. But $w + s(w) = u$ implies $\mu + s(\mu) + c + 1 \;=\; \lambda$, a contradiction. So $v \geq b^m$ cannot happen.

Second, suppose $w < cb^m$, that is, $w = cb^m - 1 - v$ with $0 \leq v \leq cb^m - 2$. By Lemma 3, $s(w) = (b-1)m + c - 1 - s(v)$, and $w + s(w) = u$ implies $v + s(v) = (b-1)m - k - 2$ and $v \in \mathrm{Gen}((b-1)m - k - 2)$. ☐

For example, in base 10, if $u = 10^{13} + 1$, we have $c = 1$, $m = 13$, $k = 0$, so $F(10^{13} + 1) = F(0) + F(115)$. Now $115 = 10^2 + 1 + 14$, so $F(115) = F(14) + F(2) = 1 + 1$, and therefore $F(10^{13} + 1) = 3$. In Section 4 we will confirm Kaprekar's result that there is no smaller number with three inverses.

Theorem 7 can be used for computing the values of $K(n)$ for small n. In Appendix A, we provide a PARI/GP program that implements the formula

[7] We remark that the identity (14) does not mention c and holds even when the argument $(b-1)m - k - 2$ is negative.

(13). We used it to compute the entries below 10^{10} in Tables 1 and 2, which show the values of $K(n)$ for $n \le 7$ and bases $2 \le b \le 10$. The values of $K(2)$ and $K(3)$ for any base b will be derived in the next section, and the values of $K(n)$ for any n and bases 2 and 3 in Section 5. The values of $K(n)$ in Tables 1 and 2 for $n \ge 4$ and bases $b \ge 4$ are included here for convenience, but they will not be officially established until we have the recurrences of Section 6.[8]

b	2	3	4	5	6
$K(1)$	0	0	0	0	0
$K(2)$	2^2+1	$3+1$	4^2+1	$5+1$	6^2+1
$K(3)$	2^7+1	3^3+1	4^7+1	5^2+1	6^9+1
$K(4)$	$2^{12}+6$	3^5+5	$4^{12}+18$	5^4+7	$6^{16}+38$
$K(5)$	$2^{136}+6$	$3^{17}+5$	$4^{5468}+18$	5^9+9	$6^{(6^9+44)/5}+38$
$K(6)$	$2^{260}+130$	$3^{29}+29$	$4^{10924}+4^7+2$	$5^{15}+27$	$6^{(2 \cdot 6^9+8)/5}+6^9+2$
$K(7)$	$2^{4233}+130$	$3^{139}+29$	$4^{E_4(7)}+4^7+21$, $E_4(7)=(4^{12}+4^7+40)/3$	$5^{165}+27$	$6^{E_6(7)}+6^9+2$, $E_6(7)=(6^{16}+6^9+43)/5$

Table 1: Values of $K(1), \ldots, K(7)$ for bases $b = 2, \ldots, 6$ (the columns are A230303, A230640, A230638, A230867, A238840).

b	7	8	9	10
$K(1)$	0	0	0	0
$K(2)$	$7+1$	8^2+1	$9+1$	10^2+1
$K(3)$	7^2+1	$8^{11}+1$	9^2+1	$10^{13}+1$
$K(4)$	7^3+9	$8^{20}+66$	9^3+11	$10^{24}+102$
$K(5)$	$7^{10}+9$	$8^{E_8(5)}+66$, $E_8(5)=(8^{11}+76)/7$	$9^{12}+11$	$10^{E_{10}(5)}+102$, $E_{10}(5)=(10^{13}+116)/9$
$K(6)$	$7^{17}+51$	$8^{E_8(6)}+8^{11}+2$, $E_8(6)=(2 \cdot 8^{11}+12)/7$	$9^{21}+83$	$10^{E_{10}(6)}+10^{13}+2$, $E_{10}(6)=2(10^{13}+8)/9$
$K(7)$	$7^{67}+51$	$8^{E_8(7)}+8^{11}+2$, $E_8(8)=(8^{20}+8^{11}+75)/7$	$9^{103}+83$	$10^{E_{10}(7)}+10^{13}+2$, $E_{10}(7)=(10^{24}+10^{13}+115)/9$

Table 2: Values of $K(1), \ldots, K(7)$ for bases $b = 7, \ldots, 10$ (the columns are A238841, A238842, A238843, A006064).

[8]The values of $K(2)$ and $K(3)$ could also be obtained from the recurrences of Section 6, but it seems more informative to calculate them directly.

4 Properties of $K(n)$

When $u = K(n)$, the smallest number with n generators in base b, Theorem 7 allows us to make a stronger assertion than (11). (Note that $K(1) = 0$ for any base.)

Theorem 8. *Let $b \geq 2$ and $n \geq 2$.*

(i) $K(n)$ has the following representation in the form (11)*:*

$$K(n) \;=\; b^{E(n)} + 1 + k\,, \tag{15}$$

where the exponent $E(n)$ is at least 1 and[9] $0 \leq k \leq (b-1)E(n) - 2$.

(ii) If b is odd, then both $K(n)$ and k are even.

Proof. Let $n \geq 2$. We notice that no integer from 1 to b can have more than one generator, and hence $K(n) > b$.

(i) Let $u := K(n)$, and write $u = c(b^m + 1) + k$ as in Lemma 5. If $c > 1$, then we let $u' := b^k + 1 + k$ and notice that, by (14), $F(u') = F(u) = n$ while $u' < u$, a contradiction. Hence $c = 1$.

If $k > (b-1)m - 2$, we apply (14) and obtain $n = F(u) = F(k) + F((b-1)m - k - 2)$. Since the argument of the last term is negative, we have $F(u) = F(k)$ while $k < u$, a contradiction. Hence $k \leq (b-1)m - 2$.

(ii) If b is odd, $K(n)$ is even by Lemma 1, and therefore k is even. □

In Theorems 9 and 10 we compute the values of $K(2)$ and $K(3)$ and the corresponding values of $E(2)$ and $E(3)$.

Theorem 9. *For any $b \geq 2$, we have*

$$K(2) \;=\; \begin{cases} b^2+1, & \text{if } b \text{ is even;} \\ b+1, & \text{if } b \text{ is odd;} \end{cases} \qquad E(2) \;=\; \begin{cases} 2, & \text{if } b \text{ is even;} \\ 1, & \text{if } b \text{ is odd.} \end{cases} \tag{16}$$

Proof. Suppose b is even. The number $b^2 + 1 = [101]_b$ has the two generators $b^2 = [100]_b$ and $b^2 - b + 1 = [b-1, 1]_b$, so $K(2) \leq b^2 + 1$. However, it is easy to check by hand that the values of $f(v)$ for $0 \leq v \leq b^2$ are all distinct, so $K(2) = b^2 + 1$. The case when b is odd is even easier to verify, and we omit the details. □

Theorem 10. *For any $b \geq 2$, we have*

$$K(3) \;=\; \begin{cases} 129, & \text{if } b = 2; \\ 28, & \text{if } b = 3; \\ b^{b+3}+1, & \text{if } b \geq 4 \text{ is even;} \\ b^2+1, & \text{if } b \geq 5 \text{ is odd;} \end{cases} \qquad E(3) \;=\; \begin{cases} 7, & \text{if } b = 2; \\ 3, & \text{if } b = 3; \\ b+3, & \text{if } b \geq 4 \text{ is even;} \\ 2, & \text{if } b \geq 5 \text{ is odd.} \end{cases} \tag{17}$$

[9]By Lemma 6, the upper bound for k here is better than (12) when $E(n) \geq 2$, which holds for $n \geq 3$ (as we will show below).

Proof. For $b \in \{2,3,4\}$, we refer to Table 1. Suppose first that $b \geq 6$ is even. Certainly $b^{b+3}+1$ has three generators, namely $b^{b+3}-1-[1,0,b-3]_b$, $b^{b+3}-1-[b-1,b-2]_b$, and b^{b+3} (this is easily checked using Lemma 3). So $K(3) \leq b^{b+3}+1$. If $u := K(3) < b^{b+3}+1$, then by Theorem 8 we would have $u = b^m+1+k$ with $m \leq b+2$ and $0 \leq k \leq (b-1)m-2$. By (14) we have $3 = F(u) = F(k) + F((b-1)m-k-2)$. So either $F(k)=1$ and $F((b-1)m-k-2)=2$, or $F(k)=2$ and $F((b-1)m-k-2)=1$. We discuss only the first possibility, the second being similar. From $k \geq K(1)=0$ and $(b-1)m-k-2 \geq K(2) = b^2+1$, we find that m must equal $b+2$, and $0 \leq k \leq b-5$. By Lemma 1, k is even. Let $\lambda := b-5-k$, which is odd, and $0 \leq \lambda \leq b-5$. But now $F(b^2+\lambda+1) = 2 = F(\lambda)+F(\lambda-4)$ implies (again by Lemma 1) that λ is even, a contradiction. This completes the proof of $K(3) = b^{b+3}+1$ for even $b \geq 6$.

Now, suppose that $b \geq 5$ is odd. Certainly b^2+1 has three generators, $b^2-1-[1,(b-5)/2]_b$, $b^2-1-[b-2]_b$, and b^2, and we can easily check that at most two of the values $f(v)$ for $0 \leq v \leq b^2$ can coincide. ☐

Theorem 10 confirms Kaprekar's result that $K(3) = 10^{13}+1$ in base 10.

When we come to study the case of a general base b in Section 6, we will need to know the values of a refined version of $K(n)$. We define $K_i(n)$ to be the smallest number $v \equiv i \pmod{b-1}$ for which $F(v)=n$, where i is a residue class modulo $b-1$. If b is odd, by Lemma 1 we need only consider even values of i, so we can say more precisely that $i \in \mathcal{I}$, where $\mathcal{I}$ is a subset of the residue classes modulo $b-1$ given by

$$\mathcal{I} := \begin{cases} \{0,1,2,3,\ldots,b-2\}, & \text{if } b \text{ is even,} \\ \{0,2,4,6,\ldots,b-3\}, & \text{if } b \text{ is odd.} \end{cases} \tag{18}$$

Then we have:

$$K(n) \;=\; \min_{i \in \mathcal{I}} K_i(n)\,. \tag{19}$$

There is an analog of Theorem 8 for $K_i(n)$.

Theorem 11. *For any $b \geq 2$, $i \in \mathcal{I}$, and $n \geq 2$, $K_i(n)$ has the following representation in the form* (11)*:*

$$K_i(n) \;=\; c(b^{E(n)}+1)+k\,, \tag{20}$$

where $E(n)$ is as in (15),[10] *for some integers c and k satisfying $1 \leq c \leq b-1$ and $0 \leq k \leq (b-1)E(n)-2$. Furthermore, if b is odd then $c \leq \frac{b-1}{2}$.*

Proof. Theorem 8 states that $K(n) = b^{E(n)}+1+k'$ for some k', and so we have $F(b^{E(n)}+1+k') = n$. Theorem 7 (applied twice) implies that for *any* c' and the same k',

$$F(c'(b^{E(n)}+1)+k') = F(k')+F((b-1)E(n)-k'-2) = F(b^{E(n)}+1+k') = n\,.$$

[10]However, the k in (20) is not the same as the k in (15).

Let us show that there exists a value $c' := c_i$ such that $c_i(b^{E(n)} + 1) + k' \equiv i \pmod{b-1}$. Since $c_i(b^{E(n)}+1)+k' \equiv k'+2c_i \pmod{b-1}$, we want $k'+2c_i \equiv i \pmod{b-1}$. For even b, this congruence is trivially solvable for c_i in the interval $1 \le c_i \le b-1$. For odd b, $i \in \mathcal{I}$ is even and so is k' (by Theorem 8), so the congruence reduces to $\frac{k'}{2} + c_i \equiv \frac{i}{2} \pmod{\frac{b-1}{2}}$, which is solvable for c_i in the interval $1 \le c_i \le \frac{b-1}{2}$.

By the definition of $K_i(n)$, we have

$$b^{E(n)} + 1 + k' = K(n) \le K_i(n) \le c_i(b^{E(n)} + 1) + k'. \tag{21}$$

By Theorem 7, $K_i(n) = c(b^m + 1) + k$ for some integers c, k, $1 \le c \le b-1$, $0 \le k \le b^m$. From the inequality (21) and Lemma 5, we conclude that $m = E(n)$ and $c \le c_i$, and thus $c \le \frac{b-1}{2}$ if b is odd. We further apply (14) to obtain $n = F(c(b^{E(n)} + 1) + k) = F(k) + F((b-1)E(n) - k - 2)$. If $k > (b-1)E(n) - 2$, then $F(k) = n$ and thus $k \ge K(n) = b^{E(n)} + 1 + k'$, which contradicts $k \le b^{E(n)}$. Hence $k \le (b-1)E(n) - 2$. □

For bases $b = 2$ and 3, we have $\mathcal{I} = \{0\}$, so there is only one $K_i(n)$, which is $K_0(n) = K(n)$. For bases $b \ge 4$ and $n \le 3$, we can give $K_i(n)$ explicitly. In the following three theorems, the subscripts i in $K_i(n)$ are elements of $\mathcal{I}$, and in particular are to be read modulo $b-1$. We omit the proofs, which are similar to those of Theorems 9 and 10. Table 3 illustrates these theorems.

Theorem 12. *For even $b \ge 2$,*

$$\begin{aligned} K_{2\lambda}(1) &= 2\lambda \quad \textit{for} \quad 0 \le \lambda \le \frac{b-2}{2}, \\ K_{2\lambda+1}(1) &= b + 2\lambda \quad \textit{for} \quad 0 \le \lambda \le \frac{b-4}{2}\,; \end{aligned} \tag{22}$$

for odd $b \ge 3$,

$$K_{2\lambda}(1) = 2\lambda \quad \textit{for} \quad 0 \le \lambda \le \frac{b-3}{2}\,. \tag{23}$$

Theorem 13. *For even $b \ge 2$,*

$$K_{2+2\lambda}(2) = b^2 + 1 + 2\lambda \quad \textit{for} \quad 0 \le \lambda \le b-2\,; \tag{24}$$

for odd $b \ge 3$,

$$K_{2+2\lambda}(2) = b + 1 + 2\lambda \quad \textit{for} \quad 0 \le \lambda \le \frac{b-3}{2}\,. \tag{25}$$

Theorem 14. *For even $b \ge 4$,*

$$K_0(3) = b^{b+3} + b^2 + 2b - 4\,, \tag{26}$$

$$K_{2+2\lambda}(3) = b^{b+3} + 1 + 2\lambda \quad \textit{for} \quad 0 \le \lambda \le b-3\,; \tag{27}$$

for odd $b \ge 5$,

$$K_0(3) = b^2 + 2b - 3\,, \tag{28}$$

$$K_{2+2\lambda}(3) = b^2 + 1 + 2\lambda \quad \textit{for} \quad 0 \le \lambda \le \frac{b-5}{2}\,. \tag{29}$$

$b=6$				$b=9$			
$i \setminus n$	1	2	3	$i \setminus n$	1	2	3
0	0	45	6^9+44	0	0	16	96
1	6	41	6^9+5	2	2	10	82
2	2	37	6^9+1	4	4	12	84
3	8	43	6^9+7	6	6	14	86
4	4	39	6^9+3				

Table 3: Values of $K_i(n)$ ($n \le 3$) for bases $b=6$ (left) and $b=9$ (right), illustrating Theorems 12-14.

For $b \ge 4$, the minimal values of $K_i(2)$ and $K_i(3)$ occur when $\lambda = 0$, that is, when $i=2$, and confirm (via (19)) the values of $K(2)$ and $K(3)$ given in Theorems 9 and 10.

The following bounds on $K_i(n)$ will be used in the proof of Theorem 18.

Theorem 15. *For any $b \ge 2$, $i \in \mathcal{I}$, and $n \ge 2$, we have*[11]

$$b^{E(n)} \; < \; K(n) \; \le \; K_i(n) \; < \; \beta\, b^{E(n)+1} ,$$

where $\beta := 1$ if b is even, and $\beta := \frac{1}{2}$ if b is odd.

Proof. The lower bound for any b and the upper bound for an even b follow directly from Theorem 11. Let us prove that for odd b, $K_i(n) \le \frac{1}{2} b^{E(n)+1}$.

For $n=2$, the bound can be verified directly using Table 1 (for $b=3$) and Theorem 13 (for odd $b \ge 5$). Suppose $n \ge 3$. By Theorem 11, we have $K_i(n) = c(b^m+1)+k$, where $m = E(n)$, $1 \le c \le \frac{b-1}{2}$ and $0 \le k \le (b-1)m-2$. Furthermore, for $b=3$ we have $m \ge 3$ (see Table 1); while for odd $b \ge 5$, we have $m \ge 2$ by Theorem 14. Hence

$$K_i(n) = c(b^m+1)+k \le \frac{b-1}{2}(b^m+1) + (b-1)m - 2 < \frac{b-1}{2}(b^m+2m+1)\,.$$

It is easy to check that for $b=3$, $m \ge 3$ and $b \ge 5$, $m \ge 2$, we have $2m+1 \le \frac{1}{b-1} b^m$, implying that

$$K_i(n) < \frac{b-1}{2}\left(b^m + \frac{1}{b-1} b^m\right) = \frac{1}{2} b^{m+1}$$

as required. □

5 $K(n)$ for bases 2 and 3

We first discuss the base 2 case for general n. The initial values of $f(u)$ and $F(u)$ are shown in Table 4. We see that the smallest numbers with 1 and

[11] From Theorem 8 it follows that $K(n) \le 2b^{E(n)}$, which is a stronger upper bound for $K(n)$ when $b \ge 4$.

u	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$f(u)$	0	2	3	5	5	7	8	10	9	11	12	14	14	16	17	19	17	19	20
$F(u)$	1	0	1	1	0	2	0	1	1	1	1	1	1	0	2	0	1	2	0

Table 4: Values of $f(u)$ and $F(u)$ in base 2 (A092391, A228085)

2 generators are $K(1) = 0$ and $K(2) = 5$, respectively. Direct search by computer gives $K(3) = 129$ and $K(4) = 4102$, as we have already seen in Table 1 (although $K(5) = 2^{136} + 6$ is out of reach). The general solution is given by the following pair of recurrences.

Theorem 16. *For $b = 2$ and any $n \geq 2$, we have*

$$E(n) \ = \ K\left(\left\lceil \tfrac{n}{2} \right\rceil\right) + K\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right) + 2 \tag{30}$$

and

$$K(n) \ = \ 2^{E(n)} + 1 + K\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right) . \tag{31}$$

Also

$$K(n) \ > \ 2K(n-1) . \tag{32}$$

Proof. The proof is by induction on n. The results are true for $n \leq 3$, so we assume $n \geq 4$. As in Theorem 8, let $u := K(n) = 2^m + 1 + k$, where $m := E(n)$ and $0 \leq k \leq m-2$. By (14), $n = F(u) = F(k) + F(m-k-2)$. Let $x := F(k)$, $y := F(m-k-2)$ so that $x + y = n$. Then $k \geq K(x)$, $m - k - 2 \geq K(y)$, and thus

$$m \geq K(x) + K(y) + 2 . \tag{33}$$

We know from (32) that the sequence $(K(n))_{n \geq 1}$ is of exponential type, so by Lemma 4 the right-hand side of (33) is minimized only when either $x = \left\lceil \frac{n}{2} \right\rceil$, $y = \left\lfloor \frac{n}{2} \right\rfloor$ or $x = \left\lfloor \frac{n}{2} \right\rfloor$, $y = \left\lceil \frac{n}{2} \right\rceil$ (there is no difference if n is even). From Lemma 5(ii), it follows that the value of $E(n)$ is given by (30), and that k equals the smaller of $K(\left\lceil \frac{n}{2} \right\rceil)$ and $K(\left\lfloor \frac{n}{2} \right\rfloor)$, which is $K(\left\lfloor \frac{n}{2} \right\rfloor)$ by induction and (32). This proves (31). The proof of (32) is now a routine calculation; we omit the details. ☐

Remark. The proof of Theorem 16 also shows that

$$\mathrm{Gen}(K(n)) \ = \ \{2^{E(n)} + v \mid v \in \mathrm{Gen}\left(K\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right)\right)\} \ \cup \ \{2^{E(n)} - 1 - v \mid v \in \mathrm{Gen}\left(K\left(\left\lceil \tfrac{n}{2} \right\rceil\right)\right)\} . \tag{34}$$

Table 5 extends the $b = 2$ column of Table 1 to $n = 16$. (The first 100 terms of $E(n)$ and $K(n)$ are given in the entries A230302 and A230303 in [12]).

There is a similar pair of recurrences in the base 3 case.

n	$E(n)$	$K(n)$
8	8206	$2^{8206} + 4103$
9	$2^{136} + 4110$	$2^{E(9)} + 4103$
10	$2^{137} + 14$	$2^{E(10)} + 2^{136} + 7$
11	$2^{260} + 2^{136} + 138$	$2^{E(11)} + 2^{136} + 7$
12	$2^{261} + 262$	$2^{E(12)} + 2^{260} + 131$
13	$2^{4233} + 2^{260} + 262$	$2^{E(13)} + 2^{260} + 131$
14	$2^{4234} + 262$	$2^{E(14)} + 2^{4233} + 131$
15	$2^{8206} + 2^{4233} + 4235$	$2^{E(15)} + 2^{4233} + 131$
16	$2^{8207} + 8208$	$2^{E(16)} + 2^{8206} + 4104$

Table 5: Base 2: $E(n)$ and $K(n)$ for $n = 8, \ldots, 16$, extending Table 1.

Theorem 17. *For $b = 3$ and any $n \geq 2$, we have*

$$E(n) \ = \ \frac{K\left(\left\lceil \frac{n}{2} \right\rceil\right) + K\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + 2}{2} \tag{35}$$

and

$$K(n) \ = \ 3^{E(n)} + 1 + K\left(\left\lfloor \frac{n}{2} \right\rfloor\right) . \tag{36}$$

Also

$$K(n) \ > \ 3K(n-1) . \tag{37}$$

Proof. The proof is similar to that of Theorem 16, except at one step. Again we use induction on $n \geq 4$ and let $u := K(n) = 3^m + 1 + k$, where $m := E(n)$ and $0 \leq k \leq 2m - 2$. Then $n = F(u) = F(k) + F(2m - k - 2) = x + y$, say, with $x + y = n$. Then $k \geq K(x)$, $2m - k - 2 \geq K(y)$, so

$$2m \geq K(x) + K(y) + 2 . \tag{38}$$

The difference from (33) in the base 2 case lies in the presence of the factor of 2 (in general it will be $b - 1$) on the left-hand side of this inequality. So now we must minimize the sum $K(x) + K(y)$ subject to the additional requirement that the sum is even. Here that does not cause any difficulty, because all values of K are even (by Lemma 1). We complete the proof as in the base 2 case, by taking $x = \left\lfloor \frac{n}{2} \right\rfloor$, $y = \left\lceil \frac{n}{2} \right\rceil$. □

The first seven terms of $E(n)$ and $K(n)$ for base 3 are shown in Table 1; the first 100 terms may be found in A230639 and A230640.

6 $K(n)$ for a general base b

In this section we give a set of recurrences that determine $K(n)$ for a general base $b \geq 2$. The divisibility requirement that we encountered in (38) for the base 3 case makes the recurrences in the general case considerably more complicated.

We know from Theorem 8 that $K(n)$ has the form

$$K(n) = b^{E(n)} + 1 + k\,, \tag{39}$$

where $0 \leq k \leq (b-1)E(n) - 2$. Then by (14),

$$n = F(K(n)) = F(k) + F((b-1)E(n) - k - 2) = x + y\,,$$

where $x := F(k)$ and $y := F((b-1)E(n) - k - 2)$. Since both k and $(b-1)E(n) - k - 2$ are smaller than $K(n)$ and thus cannot have n generators, the values of x, y must be in the range from 1 to $n-1$.

The definitions of x, y imply $k \geq K(x)$, $(b-1)E(n) - k - 2 \geq K(y)$, and therefore

$$(b-1)E(n) \geq K(x) + K(y) + 2\,. \tag{40}$$

Since in general $K(x)+K(y)+2$ will not be a multiple of $b-1$, the implied lower bound on $E(b)$ cannot always be attained. We therefore refine the inequality (40) using the functions $K_i(n)$ introduced in Section 4, and replace (40) with an inequality where the implied lower bound on $E(n)$ *can* be attained. If $k \equiv i \pmod{b-1}$ for $i \in \mathcal{I}$, then $(b-1)E(n) - k - 2 \equiv -i - 2 \pmod{b-1}$ and so $k \geq K_i(x)$, $(b-1)E(n) - k - 2 \geq K_{-i-2}(y)$, and

$$(b-1)E(n) \geq K_i(x) + K_{-i-2}(y) + 2\,. \tag{41}$$

Now, in contrast to (40), the right-hand side is divisible by $b-1$, and so we obtain an integer-valued lower bound on $E(n)$ (for some i and $x + y = n$). Namely, (41) implies

$$E(n) \geq \frac{\min_{i\in\mathcal{I}} \min_{1\leq x\leq n-1}\ K_i(x) + K_{-i-2}(n-x) + 2}{b-1}\,. \tag{42}$$

We will show by induction that for any $i \in \mathcal{I}$, the sequences $(K_i(n))_{n\geq 1}$ and $(K_{-i-2}(n))_{n\geq 1}$ form a pair of sequences of exponential type. Then by Lemma 4, we will be able to replace the inner minimum in (42) with

$$K'_i(n) \ := \ \min\left\{K_i\left(\left\lceil \tfrac{n}{2}\right\rceil\right) + K_{-i-2}\left(\left\lfloor \tfrac{n}{2}\right\rfloor\right),\ K_i\left(\left\lfloor \tfrac{n}{2}\right\rfloor\right) + K_{-i-2}\left(\left\lceil \tfrac{n}{2}\right\rceil\right)\right\}\,. \tag{43}$$

In fact, we will prove that equality holds in (42), i.e., $E(n) = \hat{E}(n)$, where

$$\hat{E}(n) \ := \ \frac{\min_{i\in\mathcal{I}} K'_i(n) + 2}{b-1} \ = \ \frac{\min_{i\in\mathcal{I}} K_i\left(\left\lceil \tfrac{n}{2}\right\rceil\right) + K_{-i-2}\left(\left\lfloor \tfrac{n}{2}\right\rfloor\right) + 2}{b-1}\,, \tag{44}$$

where the latter expression follows from the symmetry between i and $-i-2$.

For $n \geq 2$ and $i \in \mathcal{I}$, we define

$$c_{i,n} \ := \ \text{smallest integer } c \geq 1 \text{ such that } K'_{i-2c}(n) = \min_{j\in\mathcal{I}} K'_j(n) \tag{45}$$

and

$$h_{i,n} := \left\lceil \frac{n}{2} \right\rceil \quad \text{if } K_{i-2c_{i,n}}\left(\left\lceil \tfrac{n}{2} \right\rceil\right) + K_{2c_{i,n}-i-2}\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right) < K_{i-2c_{i,n}}\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right) + K_{2c_{i,n}-i-2}\left(\left\lceil \tfrac{n}{2} \right\rceil\right);$$
$$h_{i,n} := \left\lfloor \frac{n}{2} \right\rfloor \quad \text{otherwise}. \tag{46}$$

These definitions allow us to express $\hat{E}(n)$ as

$$\hat{E}(n) = \frac{K'_{i-2c_{i,n}}(n) + 2}{b-1} = \frac{K_{i-2c_{i,n}}(h_{i,n}) + K_{2c_{i,n}-i-2}(n - h_{i,n}) + 2}{b-1}, \tag{47}$$

which holds for any $i \in \mathcal{I}$.

Theorem 18. *For all $n \geq 2$,*

$$\min_{1 \leq j \leq n-1} K_i(j) + K_{-i-2}(n-j) = K'_i(n) \quad \textit{for all } i \in \mathcal{I}, \tag{48}$$

and

$$E(n) \;=\; \hat{E}(n). \tag{49}$$

Furthermore, for all $n \geq 3$,[12]

$$E(n) \;\geq\; \begin{cases} E(n-1) + 1, & \textit{if } b \textit{ is odd and } n \in \{3, 4\}; \\ E(n-1) + 2, & \textit{otherwise}. \end{cases} \tag{50}$$

Proof. We will prove all three statements (48), (49), and (50) together by induction on n. We write $(48)_j$, $(49)_j$, $(50)_j$, to refer to the statements (48), (49), (50) for $n = j$. We divide the proof into the following four parts:

(I) $(48)_2$, $(49)_2$, $(49)_3$, $(50)_3$ are true.

(II) For $n \geq 3$, $(48)_n$ follows from $(50)_j$ for $3 \leq j < n$.

(III) For $n \geq 4$, $(49)_n$ follows from $(48)_n$ and $(50)_j$ for $3 \leq j < n$.

(IV) For $n \geq 4$, $(50)_n$ follows from $(49)_n$, $(49)_{n-1}$, $(50)_j$ for $3 \leq j < n$.

Proof of (I). This is easily verified using the values of $K(n)$ and $K_i(n)$ from Theorems 9, 10, 12, 13, and noticing that $c_{i,2} = 1$ and $h_{i,2} = 1$ for all $i \in \mathcal{I}$.

Proof of (II). Let $n \geq 3$ and $i \in \mathcal{I}$. To establish $(48)_n$, we will first show that

$$K_i(m+1) \;>\; K_i(m) + K_{-i-2}(m) \tag{51}$$

holds for all $m = 1, 2, \ldots, n-2$, and then apply Lemma 4(ii). The inequality (51) can be verified directly for $m = 1$ using Theorems 12 and 13. If $m \geq 2$, we consider two cases depending on the parity of b, and use Theorem 15 to

[12]Note that $E(1)$ is not defined, which is why we start (50) at $n = 3$.

bound the terms in (51). For even b, we have $K_i(m+1) > b^{E(m+1)} \geq b^{E(m)+2}$ by $(50)_{m+1}$, while the right-hand side of (51) is at most $2b^{E(m)+1} \leq b^{E(m)+2}$ (since $b \geq 2$). For odd b, we have $K_i(m+1) > b^{E(m+1)} \geq b^{E(m)+1}$ by $(50)_{m+1}$, while the right-hand side of (51) is at most $b^{E(m)+1}$. This proves (51), which by Lemma 4(ii) (taking $a(m) = K_i(m)$, $b(m) = K_{-i-2}(m)$) implies $(48)_n$.

Proof of (III). Let $n \geq 4$. We fix an arbitrary $i \in \mathcal{I}$. To prove $(49)_n$, we first use Theorem 11 to write $K_i(n) = c(b^{E(n)} + 1) + k$ for some integers c and k satisfying $1 \leq c \leq b-1$ and $0 \leq k \leq (b-1)E(n) - 2$. Then (14) implies $n = F(K_i(n)) = F(k) + F((b-1)E(n) - k - 2)$.

Since $K_i(n) \equiv i \pmod{b-1}$, we have $k \equiv i - 2c \pmod{b-1}$. Let $x := F(k)$. Then $k \geq K_{i-2c}(x)$ and $(b-1)E(n) - k - 2 \geq K_{2c-i-2}(n-x)$, thus

$$E(n) \overset{\text{(i)}}{\geq} \frac{K_{i-2c}(x) + K_{2c-i-2}(n-x) + 2}{b-1} \overset{\text{(ii)}}{\geq} \frac{K'_{i-2c}(n) + 2}{b-1} \overset{\text{(iii)}}{\geq} \hat{E}(n)\,, \quad (52)$$

where inequality (i) is immediate, (ii) follows from $(48)_n$, and (iii) follows from (44).

Conversely, we now prove that $E(n) \leq \hat{E}(n)$. First, we notice that $(50)_j$ for $3 \leq j \leq \lceil \frac{n}{2} \rceil$ and Theorem 9 imply that $E\left(\lceil \frac{n}{2} \rceil\right) \geq E(2) \geq 1$. Together with (47) and Theorem 15, this further gives

$$\hat{E}(n) \geq \frac{K\left(\lceil \frac{n}{2} \rceil\right) + 2}{b-1} > \frac{b^{E\left(\lceil \frac{n}{2} \rceil\right)} + 2}{b-1} \geq \frac{b+2}{b-1} > 1\,. \quad (53)$$

Now, let us define $L := c_{i,n}(b^{\hat{E}(n)} + 1) + K_{i-2c_{i,n}}(h_{i,n})$, which has the form (11) since

$$K_{i-2c_{i,n}}(h_{i,n}) \leq (b-1)\hat{E}(n) - 2 \leq b^{\hat{E}(n)} - b + 1$$

as follows from (47), (53), and Lemma 6. Using (14) and (47), we have

$$\begin{aligned} F(L) &= F(K_{i-2c_{i,n}}(h_{i,n})) + F((b-1)\hat{E}(n) - K_{i-2c_{i,n}}(h_{i,n}) - 2) \\ &= F(K_{i-2c_{i,n}}(h_{i,n})) + F(K_{2c_{i,n}-i-2}(n - h_{i,n})) \\ &= h_{i,n} + n - h_{i,n} \\ &= n\,. \end{aligned} \quad (54)$$

From (54) and $L \equiv 2c_{i,n} + (i - 2c_{i,n}) \equiv i \pmod{b-1}$, it follows that $K_i(n) \leq L$. Then Lemma 5 implies that $E(n) \leq \hat{E}(n)$, which together with (52) establishes $(49)_n$.

Proof of (IV). For $n = 4$, we use the identity $(49)_4$ and Theorem 13 to obtain

$$E(4) = \begin{cases} 12, & \text{if } b = 2\,; \\ 5, & \text{if } b = 3\,; \\ 4, & \text{if } b = 5\,; \\ 2b+4, & \text{if } b \geq 4 \text{ is even}\,; \\ 3, & \text{if } b \geq 7 \text{ is odd}\,. \end{cases} \quad (55)$$

Comparing these values to those of $E(3)$ given in Theorem 10, we conclude that $(50)_4$ holds.

Let $n \geq 5$. To prove $(50)_n$, we consider two cases depending on the parity of n.

First, suppose that n is even, i.e., $n = 2t$ for some $t \geq 3$. The identity $(49)_n$ gives

$$E(n) \;=\; \frac{K_{i-2c_{i,n}}(t) + K_{2c_{i,n}-i-2}(t) + 2}{b-1} \;\geq\; \frac{2K(t)+2}{b-1}\,, \tag{56}$$

while $(49)_{n-1}$ gives

$$E(n-1) \;=\; \frac{\min_j K_j(t) + K_{-j-2}(t-1) + 2}{b-1}\,. \tag{57}$$

We obtain an upper bound on the right-hand side of (57) if we choose any particular value of j, so let us choose $j = \ell$ such that $K_\ell(t) = K(t)$. Then

$$E(n-1) \;\leq\; \frac{K(t) + K_{-\ell-2}(t-1) + 2}{b-1}\,.$$

The inequality $(50)_n$ will follow if we show that

$$\frac{2K(t)+2}{b-1} \geq \frac{K(t) + K_{-\ell-2}(t-1)+2}{b-1} + 2\,,$$

or, equivalently,

$$K(t) \geq K_{-\ell-2}(t-1) + 2(b-1)\,. \tag{58}$$

In fact, (58) holds for any t such that $(50)_t$ holds and any $\ell \in \mathcal{I}$. To prove it, we consider two cases depending on the parity of b and use Lemma 6, Theorem 15, and $(50)_t$. For even b, the inequality (58) follows from

$$K(t) > b^{E(t)} \geq b^{E(t-1)+2} \geq b^{E(t-1)+1} + 2(b-1) > K_{-\ell-2}(t-1) + 2(b-1)\,.$$

For odd b, the inequality (58) follows from

$$K(t) > b^{E(t)} \geq b^{E(t-1)+1} \geq \frac{1}{2} b^{E(t-1)+1} + 2(b-1) > K_{-\ell-2}(t-1) + 2(b-1)\,.$$

This completes the proof of (58) and thus $(50)_n$ for even n.

Second, suppose that n is odd, i.e., $n = 2t-1$ for some $t \geq 3$. Arguing as before, we have

$$E(n) \geq \frac{K(t) + K(t-1) + 2}{b-1} \text{ and } E(n-1) \leq \frac{K(t-1) + K_{-\ell-2}(t-1) + 2}{b-1}$$

for some $\ell \in \mathcal{I}$. We know from the previous paragraph that the inequality (58) holds, and this again implies $(50)_n$.

This completes the proof of the theorem. □

From (48) we immediately have

Corollary 19. *For any $i \in \mathcal{I}$, the sequences $(K_i(n))_{n\geq 1}$ and $(K_{-i-2}(n))_{n\geq 1}$ form a pair of sequences of exponential type.*

Theorem 20. *For any $b \geq 2$, $i \in \mathcal{I}$, and $n \geq 2$, we have*

$$K_i(n) \;=\; c_{i,n}(b^{E(n)}+1) + K_{i-2c_{i,n}}(h_{i,n})\,. \tag{59}$$

Moreover, the representation (59) *has the form* (20).

Proof. For $n = 2$ and $n = 3$, the statement follows from Theorems 9, 10, 12, 13, and 14. It can be verified that $c_{i,2} = c_{i,3} = 1$ and $h_{i,2} = h_{i,3} = 1$ for all $i \in \mathcal{I}$, except for $h_{0,3} = 2$ when $b \geq 4$.

For $n \geq 4$, the proof emerges from the proof of Part (III) of Theorem 18. We use Theorem 11 to write $K_i(n) = c(b^{E(n)} + 1) + k$ for some integers c and k satisfying $1 \leq c \leq b-1$ and $0 \leq k \leq (b-1)E(n) - 2$. Since $E(n) = \hat{E}(n)$ (by Theorem 18), the chain of inequalities (i), (ii), and (iii) in (52) are all equalities.

By (44), the equality (iii) is equivalent to $K'_{i-2c}(n) = \min_{i\in\mathcal{I}} K'_i(n)$, implying that $c \geq c_{i,n}$ (by the definition of $c_{i,n}$). Again, we consider $L := c_{i,n}(b^{\hat{E}(n)} + 1) + K_{i-2c_{i,n}}(h_{i,n})$, for which we proved that $K_i(n) \leq L$. Since $E(n) = \hat{E}(n)$, Lemma 5 now implies $c \leq c_{i,n}$, and thus $c = c_{i,n}$.

Since $c = c_{i,n}$, the equality (i) implies $k = K_{i-2c_{i,n}}(x)$, while the equality (ii) implies that $K_{i-2c_{i,n}}(x)+K_{2c_{i,n}-i-2}(n-x)$ equals its minimum value $K'_{i-2c_{i,n}}(n)$. Then from Lemma 4 it follows that $x = \left\lfloor \frac{n}{2} \right\rfloor$ or $x = \left\lceil \frac{n}{2} \right\rceil$. Furthermore, $K_i(n) \leq L$ implies that $k = K_{i-2c_{i,n}}(x) \leq K_{i-2c_{i,n}}(h_{i,n})$ (by Lemma 5) and thus $x \leq h_{i,n}$. When $h_{i,n} = \left\lfloor \frac{n}{2} \right\rfloor$ (in particular, when n is even), we must have $x = h_{i,n}$. On the other hand, if n is odd and $h_{i,n} = \left\lceil \frac{n}{2} \right\rceil$, then $x = \left\lfloor \frac{n}{2} \right\rfloor \neq h_{i,n}$ does not produce the minimum of $K_{i-2c_{i,n}}(x) + K_{2c_{i,n}-i-2}(n-x)$ as follows from (46). Hence in all cases we have $x = h_{i,n}$ and thus $k = K_{i-2c_{i,n}}(h_{i,n})$. in which cases $x = h_{i,n}$ by (46). In the case when n is odd and the two sums above are equals, both $x = \left\lfloor \frac{n}{2} \right\rfloor = h_{i,n}$ and $x = \left\lceil \frac{n}{2} \right\rceil = n - h_{i,n}$ deliver the minimum of $K_{i-2c_{i,n}}(x) + K_{2c_{i,n}-i-2}(n-x)$. However, here $K_i(n) \leq L$ implies that $k = K_{i-2c_{i,n}}(x) \leq K_{i-2c_{i,n}}(h_{i,n})$ (by Lemma 5) and thus $x = h_{i,n}$. Therefore, in all cases we have $k = K_{i-2c_{i,n}}(h_{i,n})$, which completes the proof. □

Corollary 21.

$$K(n) \;\geq\; b^{E(n)} + 1 + K\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right)\,. \tag{60}$$

Proof. From (19) and Theorem 20 it follows that $K(n) = K_i(n) = c_{i,n}(b^{E(n)} + 1) + K_{i-2c_{i,n}}(h_{i,n})$ for some $i \in \mathcal{I}$, for which we also have $c_{i,n} = 1$ by Theorem 8 and $h_{i,n} \geq \left\lfloor \frac{n}{2} \right\rfloor$ by (46). Hence

$$K(n) = b^{E(n)} + 1 + K_{i-2}(h_{i,n}) \geq b^{E(n)} + 1 + K\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right)\,.$$

□

Remarks on Theorems 18 and 20.

(1) Since, by definition, $K_i(n) \equiv i \pmod{b-1}$, we have $K_i(n) \neq K_j(n)$ for $i \neq j$ from $\mathcal{I}$. It follows that the choice of $i \in \mathcal{I}$ in (19) is unique, and so we may define a "generalized Thue–Morse sequence" $(\tau(n))_{n\geq 1}$ for base b by:

$$K(n) = \begin{cases} K_{\tau(n)}(n), & \text{if } b \text{ is even}, \\ K_{2\tau(n)}(n), & \text{if } b \text{ is odd}, \end{cases} \tag{61}$$

where $0 \leq \tau(n) \leq b-2$ if b is even, and $0 \leq \tau(n) \leq \frac{b-3}{2}$ if b is odd. The name comes from the fact that in base $b = 5$ this is the classical Thue–Morse sequence (see Section 8).

From (61) and Theorems 8 and 20 it follows that

$$c_{\tau(n),n} = 1 \text{ if } b \text{ is even}, \quad c_{2\tau(n),n} = 1 \text{ if } b \text{ is odd}. \tag{62}$$

(2) The $K'_i(n)$ are not all distinct. It follows at once from (43) that:

(i) if b is even, the distinct $K'_i(n)$ are

$$K'_i(n), \text{ for } 0 \leq i \leq \frac{b-4}{2}, \text{ and } K'_{b-2}(n), \tag{63}$$

where the remaining values are given by $K'_i(n) = K'_{b-i-3}(n)$;

(ii) if $b \equiv -1 \pmod 4$, the distinct $K'_i(n)$ are

$$K'_{2i}(n), \text{ for } 0 \leq i \leq \frac{b-7}{4}, \text{ and } K'_{\frac{b-3}{2}}(n), \tag{64}$$

where the remaining values are given by $K'_{2i}(n) = K'_{b-2i-3}(n)$; and

(iii) if $b \equiv 1 \pmod 4$, the distinct $K'_i(n)$ are

$$K'_{2i}(n), \text{ for } 0 \leq i \leq \frac{b-5}{4}, \tag{65}$$

where the remaining values are given by $K'_{2i}(n) = K'_{b-2i-3}(n)$.

In fact, the calculations for the case $b = 4m-1$ are (apart from a relabeling of the variables) essentially the same as the calculations for the case $b = 2m$. For $m = 1$, this can be seen from Theorems 9 and 10, which establish similar recurrences for $K(n)$ in bases $b = 2$ and $b = 3$. We will formally prove this similarity for general m in Theorem 24 in the next section.

(3) If n is even, no minimization is needed in the formula (43) for $K'_i(n)$, since the two terms inside the braces are the same, and also $h_{i,n} = \frac{n}{2}$ for all $i \in \mathcal{I}$. If n is odd, the two terms inside the braces in (43) may still coincide, including the case of $K'_{i-2c_{i,n}}(n)$ (which happens always when $b = 2, 3$, quite frequently when $b \geq 4$ is even, and sometimes when $b \geq 5$ is odd). While both $\lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{2} \rceil$ in this case may serve as $h_{i,n}$ in (47), the choice of $h_{i,n} = \lfloor \frac{n}{2} \rfloor$ is dictated by Theorem 20, which expects the contribution of $K_{i-2c_{i,n}}(h_{i,n})$ be as small as possible.

(4) We initially thought that the minimization in (44) would be determined by choosing the index i to be either j or $-j-2$, where $K_j\left(\left\lceil \frac{n}{2} \right\rceil\right) = K\left(\left\lceil \frac{n}{2} \right\rceil\right)$. This would imply that

$$\min_{i \in \mathcal{I}} K'_i(n) \stackrel{?}{=} K\left(\left\lceil \tfrac{n}{2} \right\rceil\right) + K_\ell\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right) , \tag{66}$$

for some $\ell \in \mathcal{I}$. To prevent others from falling into this trap, we mention that (66) is false. There are counter-examples when $b = 7$ and $n = 13$, and when $b = 9$ and $n = 9$ (see details in Appendix B).

Computing $E(n)$ and $K(n)$. It may be helpful to summarize the steps involved in using the recurrences to compute $E(n)$ and $K(n)$:

Step 1: For every $i \in \mathcal{I}$, compute $K'_i(n)$ from (43), omitting the duplicates mentioned in Remark (2) above.

Step 2: For every $i \in \mathcal{I}$, compute $c_{i,n}$ and $h_{i,n}$ using (45) and (46), or using the equivalent formulas (71) and (72) given below.

Step 3: Compute $E(n)$ with the formula

$$E(n) = \frac{K'_{i-2c_{i,n}}(n) + 2}{b-1} , \tag{67}$$

which follows from (47) and (49), and holds for any $i \in \mathcal{I}$.

Step 4: For every $i \in \mathcal{I}$, compute $K_i(n)$ using Theorem 20.

Step 5: Finally, compute $K(n)$ from (19).

Below we illustrate the computations for even $b \geq 4$, while in the next two sections we provide further information about bases 4, 5, 7, and 10. Additional illustrations of the computation flow are given in Appendix B.

Examples. We illustrate the computations using Theorems 18 and 20 in the case when $b \geq 4$ is even and $n = 2, 3, 4$.

For $n = 2$, we find that $E(2) = (2b-2)/(b-1) = 2$, and $K'_i(2) = 2b-4$, $c_{i,2} = 1$ and $h_{i,2} = 1$ for all i. From this we obtain the values of $K_i(2)$ that we saw in Theorem 13.

For $n = 3$, we find that $E(3) = (b^2+2b-3)/(b-1) = b+3$, $K'_i(3) = b^2+2b-5$ and $c_{i,3} = 1$ for all i, and $h_{0,3} = 2$, $h_{1,3} = h_{2,3} = h_{3,3} = \cdots = 1$. From this we obtain the values of $K_i(3)$ that we saw in Theorem 14.

For $n = 4$, we find that $K'_0(4) = 2b^2+2b-8$, $K'_i(4) = 2b^2+2b-6$ for $i \geq 1$, $E(4) = (2b^2+2b-4)/(b-1) = 2b+4$, $c_{0,4} = 2$, $c_{1,4} = 1$, $c_{2,4} = 3$, $c_{i,4} = 1$ for

$3 \le i \le b-2$, and all $h_{i,4} = 2$. Then

$$\begin{aligned} K_0(4) &= 2b^{2b+4} + b^2 + 2b - 5\,, \\ K_1(4) &= b^{2b+4} + b^2 + b - 2\,, \\ K_2(4) &= 3b^{2b+4} + b^2 + 2b - 4\,, \\ K_3(4) &= b^{2b+4} + b^2 + b\,, \\ K_4(4) &= b^{2b+4} + b^2 + 2\,, \end{aligned} \tag{68}$$

and thus

$$K(4) = K_4(4) = b^{2b+4} + b^2 + 2\,. \tag{69}$$

This confirms Kaprekar's conjecture of $10^{24} + 102$ in base 10.

Table 6 summarizes the results from the recurrence for $n \le 7$ and even bases $b \ge 4$ and odd bases $b \ge 7$. (For smaller values of b, see Tables 1, 2, 5.)

	even $b \ge 4$		odd $b \ge 7$	
n	$E(n)$	$K(n)$	$E(n)$	$K(n)$
1	–	0	–	0
2	2	b^2+1	1	$b+1$
3	$b+3$	$b^{b+3}+1$	2	b^2+1
4	$2b+4$	$b^{2b+4}+b^2+2$	3	b^3+b+2
5	$\frac{b^{b+3}+b^2+2b-4}{b-1}$	$b^{E(5)}+b^2+2$	$b+3$	$b^{b+3}+b+2$
6	$\frac{2b^{b+3}+2b-4}{b-1}$	$b^{E(6)}+b^{b+3}+2$	$2b+3$	$b^{2b+3}+b^2+2$
7	$\frac{b^{2b+4}+b^{b+3}+b^2+2b-5}{b-1}$	$b^{E(7)}+b^{b+3}+2$	b^2+2b+4	$b^{E(7)}+b^2+2$

Table 6: Values of $E(n)$ and $K(n)$ for $n \le 7$ and even bases $b \ge 4$, odd bases $b \ge 7$. See also Fig. 3 and Figs. 5–9 in Appendix B.

Let us define the set

$$J(n) := \{j \in \mathcal{I} \mid \min_{\ell \in \mathcal{I}} K_\ell\left(\left\lceil \tfrac{n}{2} \right\rceil\right) + K_{-\ell-2}\left(\left\lfloor \tfrac{n}{2} \right\rfloor\right) \text{ is attained at } \ell = j\}. \tag{70}$$

Then $c_{i,n}$ and $h_{i,n}$ can be equivalently expressed as

$$c_{i,n} \;=\; \text{smallest integer } c \ge 1 \text{ such that } i - 2c \in J(n) \text{ or } 2c - i - 2 \in J(n) \tag{71}$$

and

$$h_{i,n} = \begin{cases} \left\lceil \frac{n}{2} \right\rceil, & \text{if } 2c_{i,n} - i - 2 \notin J(n); \\ \left\lfloor \frac{n}{2} \right\rfloor, & \text{otherwise}. \end{cases} \tag{72}$$

In bases $b = 2$ and $b = 3$, we trivially have $J(n) = \mathcal{I} = \{0\}$ for all $n \ge 2$. In bases $b = 4$ and $b \ge 6$, from Theorems 12, 13, and 10 it can be verified that

$$\begin{aligned} J(2) &= \mathcal{I}, \\ J(3) &= \mathcal{I} \setminus \{0\}, \\ J(4) = J(5) = J(6) &= \mathcal{I} \setminus \{0, b-3\}. \end{aligned} \tag{73}$$

In particular, this implies that in bases $b = 4$ and $b \ge 6$,

- $c_{i,2} = c_{i,3} = 1$ and $h_{i,2} = h_{i,3} = 1$ for all $i \in \mathcal{I}$, with the exception of $h_{0,3} = 2$ (as was already noted in the proof of Theorem 20);

- $c_{i,4} = c_{i,5} = c_{i,6} = 1$, $h_{i,4} = h_{i,5} = 2$ and $h_{i,6} = 3$ for all $i \in \mathcal{I}$, with the exception of $c_{0,4} = c_{0,5} = c_{0,6} = 2$ and $c_{2,4} = c_{2,5} = c_{2,6} = 3$.

For base $b = 5$, we have $J(2) = J(4) = J(6) = \{0, 2\} = \mathcal{I}$ and $J(3) = J(5) = \{2\}$,[13] implying that $c_{i,2} = c_{i,3} = c_{i,4} = c_{i,5} = c_{i,6} = 1$, $h_{i,2} = h_{i,3} = 1$, $h_{i,4} = h_{i,3} = 2$, and $h_{i,6} = 3$ for all $i \in \mathcal{I}$, with the exception of $h_{0,3} = 2$ and $h_{0,5} = 3$. More details can be found in Section 8.

7 Quasi-positional representation for $K(n)$

Theorem 20 suggests that $K_i(n)$ can be expressed as a linear combination of terms

$$\mathcal{B}(m) := b^{E(m)} + 1 \tag{74}$$

and a single term $K_\beta(1)$ for some $\beta \in \mathcal{I}$. In this section we investigate properties of this representation and show that it resembles a conventional positional numeral system.

Theorem 22. *For any $n \geq 1$ and $i \in \mathcal{I}$, $K_i(n)$ is uniquely represented as*

$$K_i(n) = \alpha_1 \mathcal{B}(n_1) + \alpha_2 \mathcal{B}(n_2) + \cdots + \alpha_t \mathcal{B}(n_t) + K_\beta(1) \tag{75}$$

for some $t \geq 0$, where

- *$\alpha_1, \alpha_2, \ldots, \alpha_t$ are integers from the interval $[1, b-1]$ if b is even, or the interval $[1, \frac{b-1}{2}]$ if b is odd;*

- *$\beta \in \mathcal{I}$; and*

- *if $t > 0$, then $n_1 > n_2 > \cdots > n_t$ with $n_j \in \left\{\left\lfloor \frac{n_{j-1}}{2} \right\rfloor, \left\lceil \frac{n_{j-1}}{2} \right\rceil\right\}$ for $j = 2, 3, \ldots, t$, and $n_t \in \{2, 3\}$.*

Proof. If $n = 1$, we set $t := 0$ and $\beta := i$. If $n \geq 2$, we set $n_1 := n$. From Theorem 20 it follows that $K_i(n_1) = \alpha_1 \mathcal{B}(n_1) + K_{i-2\alpha_1}(n_2)$, where $\alpha_1 := c_{i,n_1}$ and $n_2 := h_{i,n_1} \in \left\{\left\lfloor \frac{n_1}{2} \right\rfloor, \left\lceil \frac{n_1}{2} \right\rceil\right\}$. If $n_2 > 1$, we represent $K_{i-2\alpha_1}(n_2)$ in a similar way, and continue the process until we get a representation (75), where $n_j := h_{i-2(\alpha_1+\alpha_2+\cdots+\alpha_{j-2}),n_{j-1}}$ and $\alpha_j := c_{i-2(\alpha_1+\alpha_2+\cdots+\alpha_{j-1}),n_j}$ for $j \geq 2$, and $\beta := i - 2(\alpha_1 + \alpha_2 + \cdots + \alpha_t)$. The properties of α_j and n_j easily follow from this construction.

We prove that the representation (75) is unique by induction on t. For $t = 0$, the statement follows from Theorem 12 as all $K_\beta(1)$ for $\beta \in \mathcal{I}$ are distinct. For $t \geq 1$, we let $k := K_i(n) - \alpha_1 \mathcal{B}(n_1)$ and show that $\alpha_1(b^{E(n_1)}+1)+k$ has the form (11). Indeed, if $t = 1$, then

[13] The apparent pattern does not continue here as $J(7) = \{0\}$.

- when b is even, $k = K_\beta(1) \leq 2b-4$ (by Theorem 12) and $E(n_1) \geq E(2) = 2$ (by Theorem 9), implying that $k \leq 2b-4 \leq b^2-b+1 \leq b^{E(n_1)}-b+1$ satisfying (12);

- when b is odd, $k = K_\beta(1) \leq b-3$ (by Theorem 12), $\alpha_1 \leq \frac{b-1}{2} < b-1$, and $E(n_1) \geq E(2) = 1$ (by Theorem 9), implying that $k \leq b-3 \leq b \leq b^{E(n_1)}$ satisfying (12).

Finally, if $t \geq 2$, then $n_1 \geq 4$ and $n_1 - n_2 \geq 2$, implying by (50) that $E(n_1) \geq E(n_2) + 2$. Then by induction $k = \alpha_2(b^{E(n_2)} + 1) + \cdots$ has the form (11) and since $\alpha_2 \leq b-1$, we get

$$k \leq \alpha_2(b^{E(n_2)} + 1) + b^{E(n_2)} \leq b^{E(n_2)+1} + b - 1 \leq b^{E(n_1)} - b + 1,$$

satisfying (12). That is, we have proved that $\alpha_1(b^{E(n_1)}+1)+k$ has the form (11), and then the uniqueness of the representation (75) follows from that of (11). □

Our proof suggests that the representation (75) can be viewed as the result of iterative application of Lemma 5, which enables positional comparison of such representations. In particular, for bounding purposes we will find it convenient to introduce a generic notation $O_{\mathcal{B}}(n_1)$ for the right-hand side of (75) (when $t \geq 1$).

We will need the following lemma.

Lemma 23. *Let $n \geq 3$. Then for any $x, y \in \{1, 2, \ldots, n-1\}$ and any $i, j \in \mathcal{I}$,*

$$K_i(x) + K_j(y) < \mathcal{B}(n).$$

Proof. From Theorems 12 and 10, it follows that for any $i \in \mathcal{I}$

$$K_i(1) \leq 2b - 4 < \frac{b^2+1}{2} \leq \frac{1}{2}\mathcal{B}(3) \leq \frac{1}{2}\mathcal{B}(n).$$

For $x > 1$ and any $i \in \mathcal{I}$, from Theorems 15 and 18, it follows that for odd b

$$K_i(x) < \frac{1}{2}b^{E(x)+1} < \frac{1}{2}b^{E(n-1)+1} \leq \frac{1}{2}b^{E(n)} < \frac{1}{2}\mathcal{B}(n),$$

while for even b

$$K_i(x) < b^{E(x)+1} < b^{E(n-1)+1} \leq \frac{1}{2}b^{E(n-1)+2} \leq \frac{1}{2}b^{E(n)} < \frac{1}{2}\mathcal{B}(n).$$

Hence in all cases,

$$K_i(x) + K_j(y) < \frac{1}{2}\mathcal{B}(n) + \frac{1}{2}\mathcal{B}(n) = \mathcal{B}(n).$$

□

The following theorem shows that the quasi-positional representations for $K(n)$ in bases $b_1 = 2m$ and $b_2 = 4m - 1$ are essentially the same. We use superscripts (b_1) and (b_2) to distinguish between the bases.

Theorem 24. *For an integer $m \geq 1$, let $b_1 := 2m$ and $b_2 := 4m - 1$. We identify the additive groups $\mathcal{I}^{(b_1)}$ (consisting of the residues modulo $2m - 1$) and $\mathcal{I}^{(b_2)}$ (consisting of the even residues modulo $2(2m - 1)$).*[14] *Then for any $n \geq 1$ and any $i \in \mathcal{I} = \mathcal{I}^{(b_1)} \cong \mathcal{I}^{(b_2)}$, $K_i^{(b_1)}(n)$ has a representation in the form* (75)

$$K_i^{(b_1)}(n) = \alpha_1 \mathcal{B}^{(b_1)}(n_1) + \alpha_2 \mathcal{B}^{(b_1)}(n_2) + \cdots + \alpha_t \mathcal{B}^{(b_1)}(n_t) + K_\beta^{(b_1)}(1)$$

if and only if $K_i^{(b_2)}(n)$ has a representation:

$$K_i^{(b_2)}(n) = \alpha_1 \mathcal{B}^{(b_2)}(n_1) + \alpha_2 \mathcal{B}^{(b_2)}(n_2) + \cdots + \alpha_t \mathcal{B}^{(b_2)}(n_t) + K_\beta^{(b_2)}(1).$$

Proof. The proof is by induction on n. For $n = 1$, the statement follows from Theorem 12, which gives $K_\beta^{(b_1)}(1) = K_\beta^{(b_2)}(1)$ for all $\beta \in \mathcal{I}$.

Let $n \geq 2$. For any $i \in \mathcal{I}$, Theorem 20 gives

$$K_i^{(b_1)}(n) = c_{i,n}^{(b_1)} \mathcal{B}^{(b_1)}(n) + K^{(b_1)}_{i - 2c_{i,n}^{(b_1)}}(h_{i,n}^{(b_1)})$$

and

$$K_i^{(b_2)}(n) = c_{i,n}^{(b_2)} \mathcal{B}^{(b_2)}(n) + K^{(b_2)}_{i - 2c_{i,n}^{(b_2)}}(h_{i,n}^{(b_2)}).$$

The statement will follow by induction if we show that $c_{i,n}^{(b_1)} = c_{i,n}^{(b_2)}$ and $h_{i,n}^{(b_1)} = h_{i,n}^{(b_2)}$. In each base, the values of $c_{i,n}$ and $h_{i,n}$ are entirely determined by the set of indices $J(n) \subset \mathcal{I}$ defined in (70). It is therefore sufficient to show that $J^{(b_1)}(n) = J^{(b_2)}(n)$. This equality holds for $n \leq 6$ as established by (73) (notice that $b_1 - 3 = b_2 - 3$ as residues in $\mathcal{I}$). For the rest of the proof we assume that $n \geq 7$.

The set $J(n)$ consists of the indices j producing the minimum of $K_j\left(\left\lceil \frac{n}{2} \right\rceil\right) + K_{-j-2}\left(\left\lfloor \frac{n}{2} \right\rfloor\right)$. Hence we need to show that comparing two such sums gives the same result ("$<$", "$=$", or "$>$") in the two bases. More generally, we will prove the following statement:

($\star$) For any positive integers u, v, u', v' less than n and any $i, j, i', j' \in \mathcal{I}$, the result of comparing $X^{(b)} := K_i^{(b)}(u) + K_j^{(b)}(v)$ and $Y^{(b)} := K_{i'}^{(b)}(u') + K_{j'}^{(b)}(v')$ is the same in bases $b = b_1$ and $b = b_2$.

Without loss of generality we assume that $\max\{u, v, u', v'\} = u$, and prove the statement by induction on u. For $u = 1, 2$, ($\star$) follows from Theorems 12 and 13. Suppose now that $2 < u < n$.

[14] More formally, we may define an isomorphism of additive groups $\pi : \mathcal{I}^{(4m-1)} \to \mathcal{I}^{(2m)}$ by $\pi([i \bmod 2(2m-1)]) = [i \bmod (2m-1)]$.

Let $(x, k) \in \{(u, i), (v, j), (u', i'), (v', j')\}$. If $x = u$, we have $x < n$, and by induction on n, $c_{k,x}^{(b_1)} = c_{k,x}^{(b_2)} =: \alpha_{k,x}$ and $h_{k,x}^{(b_1)} = h_{k,x}^{(b_2)} =: w_{k,x}$. Then by Theorem 20,

$$K_k^{(b)}(x) = \alpha_{k,x}\mathcal{B}^{(b)}(u) + K_{k-2\alpha_{k,x}}^{(b)}(w_{k,x}), \qquad b \in \{b_1, b_2\}. \tag{76}$$

If $x < u$, we obtain the same representations (76) by setting $\alpha_{k,x} := 0$ and $w_{k,x} := x$. Notice that in all cases, we have $w_{k,x} < u$.

Using the representations (76), we get

$$X^{(b)} = (\alpha_{i,u} + \alpha_{j,v})\mathcal{B}^{(b)}(u) + K_{i-2\alpha_{i,u}}^{(b)}(w_{i,u}) + K_{j-2\alpha_{j,v}}^{(b)}(w_{j,v})$$

and

$$Y^{(b)} = (\alpha_{i',u'} + \alpha_{j',v'})\mathcal{B}^{(b)}(u) + K_{i'-2\alpha_{i',u'}}^{(b)}(w_{i',u'}) + K_{j'-2\alpha_{j',v'}}^{(b)}(w_{j',v'}).$$

By Lemma 23, the sum of two K's in $X^{(b)}$ is smaller than $\mathcal{B}^{(b)}(u)$, and so is the sum of two K's in $Y^{(b)}$. Hence, if $\alpha_{i,u} + \alpha_{j,v}$ and $\alpha_{i',u'} + \alpha_{j',v'}$ are not equal, then their comparison completely determines the result of comparison of $X^{(b)}$ and $Y^{(b)}$ in each base b.

On the other hand, if $\alpha_{i,u}+\alpha_{j,v} = \alpha_{i',u'}+\alpha_{j',v'}$, then the result of comparison of $X^{(b)}$ and $Y^{(b)}$ is determined by comparison of $K_{i-2\alpha_{i,u}}^{(b)}(w_{i,u}) + K_{j-2\alpha_{j,v}}^{(b)}(w_{j,v})$ and $K_{i'-2\alpha_{i',u'}}^{(b)}(w_{i',u'}) + K_{j'-2\alpha_{j',v'}}^{(b)}(w_{j',v'})$, which is the same for $b = b_1$ and $b = b_2$ by induction (since all w's are smaller than u). This completes the proof of statement $(\star)$, which further implies that $J(n)$ is the same for $b = b_1$ and $b = b_2$, and thus proves the theorem by induction for all $n \geq 1$. □

Theorem 24 explains why the flow-charts in Appendix B for $b = 2$ and $b = 3$ (Figs. 1 and 2) are the same apart from the labels, as are the flow-charts for $b = 4$ and $b = 7$ (Figs. 3 and 6).

8 $K(n)$ for bases 4, 5, and 7

We discuss base 5 first, since this turns out to be simpler than bases 4 or 7.

For $b = 5$, the index set is $\mathcal{I} = \{0, 2\}$. From (63), there is only one $K_i'(n)$ to consider, namely

$$K_0'(n) \;=\; \min\; \left\{K_0\left(\left\lceil \tfrac{n}{2}\right\rceil\right) + K_2\left(\left\lfloor \tfrac{n}{2}\right\rfloor\right),\; K_0\left(\left\lfloor \tfrac{n}{2}\right\rfloor\right) + K_2\left(\left\lceil \tfrac{n}{2}\right\rceil\right)\right\}\,. \tag{77}$$

Then $E(n) = (K_0'(n) + 2)/4$, $c_{0,n} = c_{2,n} = 1$ for all n,

$$h_{0,n} \;=\; \begin{cases} \left\lceil \frac{n}{2}\right\rceil, & \text{if } K_0\left(\left\lfloor \frac{n}{2}\right\rfloor\right) + K_2\left(\left\lceil \frac{n}{2}\right\rceil\right) < K_0\left(\left\lceil \frac{n}{2}\right\rceil\right) + K_2\left(\left\lfloor \frac{n}{2}\right\rfloor\right), \\ \left\lfloor \frac{n}{2}\right\rfloor, & \text{otherwise}, \end{cases} \tag{78}$$

and $h_{2,n} = n - h_{0,n}$. Also

$$K_0(n) \;=\; 5^{E(n)} + K_2(h_{0,n}) + 1, \quad K_2(n) \;=\; 5^{E(n)} + K_0(h_{2,n}) + 1\,, \tag{79}$$

n	$K_0'(n)$	$E(n)$	$h_{0,n}$	$h_{2,n}$	$K_0(n)$		$K_2(n)$
1	–	–	–	–	$\mathbf{0}$	$\searrow$	2
2	2	1	1	1	$5+3$	$\searrow$	$\mathbf{5+1}$
3	6	2	2	1	5^2+7	$\swarrow$	$\mathbf{5^2+1}$
4	14	4	2	2	$\mathbf{5^4+7}$	$\searrow$	5^4+9
5	34	9	3	2	5^9+27	$\swarrow$	$\mathbf{5^9+9}$
6	58	15	3	3	$\mathbf{5^{15}+27}$	$\swarrow$	$5^{15}+33$
7	658	165	3	4	$\mathbf{5^{165}+27}$	$\searrow$	$5^{165}+633$
8	1266	317	4	4	$5^{317}+635$	$\searrow$	$\mathbf{5^{317}+633}$
9	5^9+5^4+16	488442	5	4	$5^{E(9)}+5^9+10$	$\swarrow$	$\mathbf{5^{E(9)}+5^4+8}$
10	$2\cdot 5^9+36$	976572	5	5	$\mathbf{5^{E(10)}+5^9+10}$	$\swarrow$	$5^{E(10)}+5^9+28$

Table 7: Base 5: $K(n)$ (shown in bold font) is the smaller of the entries in the last two columns. The meaning of the arrows is explained in the text. See also Fig. 4 in Appendix B.

and $K(n) = \min\{K_0(n), K_2(n)\}$. The initial values of these variables are shown in Table 7. The value of $K(n)$ is shown in bold font (the first 100 values of $E(n)$ and $K(n)$ can be found in A230868 and A230867). The symbol in the penultimate column of the table indicates the choice made in (77) when calculating $K_0'(n)$ for odd n. An arrow $\searrow$ in row i indicates that $K_0'(2i+1) = K_0(i) + K_2(i+1)$, while an arrow $\swarrow$ indicates that $K_0'(2i+1) = K_0(i+1) + K_2(i)$.[15]

The values of the generalized Thue–Morse sequence $\tau(n)$ (see (61)) are shown in Table 8.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\tau(n)$	0	1	1	0	1	0	0	1	1	0	0	1	0	1	1	0

Table 8: Initial values of $\tau(n)$ (this is essentially the Thue–Morse sequence A010060).

In this case $\tau(n)$ actually *is* the classical Thue–Morse sequence, except shifted by one step.[16] We prove this in the next theorem.

Theorem 25. *For $n \geq 1$,*

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

Proof. (Sketch.) The basis for the inductive proof are the following observations.

[15] The arrows are intended to suggest, for the four elements $K_0(i), K_2(i), K_0(i+1), K_2(i+1)$, whether it is better to pair up the North West and South East entries, or the North East and South West entries.

[16] The classical sequence is $\tau(n-1)$.

(i) If n is even, then (79) implies that $K_0(n) < K_2(n)$ if and only if $K_2\left(\frac{n}{2}\right) < K_0\left(\frac{n}{2}\right)$, and hence $\tau(n) = 1 - \tau(\frac{n}{2})$.

(ii) Suppose on the other hand that $n = 2i + 1$ is odd. There are two possibilities. If

$$K_0(i) + K_2(i+1) \; < \; K_0(i+1) + K_2(i) \tag{80}$$

(the $\searrow$ case in Table 7), then $K_2(i+1) < K_0(i+1)$, $\tau(i+1) = 2$, $K(n) = K_2(n)$, and hence $\tau(n) = 1 = \tau(\lceil\frac{n}{2}\rceil)$. If the inequality in (80) is reversed, we similarly find that $\tau(n) = 2 = \tau(\lceil\frac{n}{2}\rceil)$.

For the induction to work, we need to also show that the values of $E(n)$, $K_0(n)$, and $K_2(n)$ are considerably larger than the values of $E(n-1)$, $K_2(n-1)$, and $K_0(n-1)$, respectively, but this follows from Theorem 15 and (50). We omit the details of the proof. $\square$

The situation is more complicated in base 4. Here the index set is $\mathcal{I} = \{0, 1, 2\}$ (modulo 3), $K(n)$ is the minimum of the three terms $K_0(n)$, $K_1(n)$, $K_2(n)$, and so is specified by a ternary sequence $\tau(n) \in \mathcal{I}$. The first 100 values of $E(n)$ and $K(n)$ can be found in A230637 and A230638, and the first 100 terms of $\tau(n)$ are:

$$\begin{aligned}&0,2,2,1,1,1,2,0,2,0,2,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,2,0,2,0,2,0,2,0,2,\\&0,2,0,2,0,2,0,2,0,2,0,2,0,2,0,2,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,\\&1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,\ldots\end{aligned} \tag{81}$$

(A239110). If we write $(2,0)^{13}$ to denote a run of 13 copies of $2,0$, etc., then the first 1000 terms of this sequence are

$$0, 2^2, 1^3, (2,0)^3, 0, 1^{13}, (2,0)^{13}, 0, 1^{53}, (2,0)^{53}, 0, 1^{213}, (0,2)^{213}, 0, 1^{\geq 147}\,, \tag{82}$$

which suggests that after the initial three terms, there is a repeating pattern

$$1^{\delta_j}, (2,0)^{\delta_j}, 0\,,$$

where $\delta_j = \frac{10\cdot 4^j - 1}{3}$ (see A072197). The next theorem shows that this pattern continues for ever.

Theorem 26. *Let $b = 4$ and let $d \geq 0$ be an integer. For an integer n,*

(i) if $\frac{10\cdot 4^d+2}{3} \leq n < \frac{20\cdot 4^d+1}{3}$, then

$$\begin{aligned}K_0(n) &= 2\mathcal{B}(n) + \mathcal{B}(\lfloor\tfrac{n}{2}\rfloor) + O_{\mathcal{B}}(\lfloor\tfrac{n}{4}\rfloor)\,,\\ K_1(n) &= \mathcal{B}(n) + \mathcal{B}(\lfloor\tfrac{n}{2}\rfloor) + O_{\mathcal{B}}(\lfloor\tfrac{n}{4}\rfloor)\,,\\ K_2(n) &= 3\mathcal{B}(n) + \mathcal{B}(\lfloor\tfrac{n}{2}\rfloor) + O_{\mathcal{B}}(\lfloor\tfrac{n}{4}\rfloor)\,,\end{aligned}$$

and thus $\tau(n) = 1$;

(ii) *if* $\frac{20\cdot 4^d+2}{3} \le n < \frac{10\cdot 4^{d+1}-1}{3}$, *then*

$$\begin{aligned} K_0(n) &= \mathcal{B}(n) + \mathcal{B}(\lceil \tfrac{n}{2} \rceil) + O_{\mathcal{B}}(\lceil \tfrac{n}{4} \rceil)\,,\\ K_1(n) &= 2\mathcal{B}(n) + (2-(n \bmod 2))\mathcal{B}(\lfloor \tfrac{n}{2} \rfloor) + O_{\mathcal{B}}(\lfloor \tfrac{n}{4} \rfloor)\,,\\ K_2(n) &= \mathcal{B}(n) + (2-(n \bmod 2))\mathcal{B}(\lfloor \tfrac{n}{2} \rfloor) + O_{\mathcal{B}}(\lfloor \tfrac{n}{4} \rfloor)\,, \end{aligned}$$

and thus $\tau(n) = 2$ *when* n *is odd, and* $\tau(n) = 0$ *when* n *is even;*

(iii) *if* $n = \frac{10\cdot 4^{d+1}-1}{3}$, *then*

$$\begin{aligned} K_0(n) &= \mathcal{B}(n) + \mathcal{B}(\lfloor \tfrac{n}{2} \rfloor) + O_{\mathcal{B}}(\lceil \tfrac{n}{4} \rceil)\,,\\ K_1(n) &= 2\mathcal{B}(n) + \mathcal{B}(\lceil \tfrac{n}{2} \rceil) + O_{\mathcal{B}}(\lfloor \tfrac{n}{4} \rfloor)\,,\\ K_2(n) &= \mathcal{B}(n) + \mathcal{B}(\lceil \tfrac{n}{2} \rceil) + O_{\mathcal{B}}(\lfloor \tfrac{n}{4} \rfloor)\,, \end{aligned}$$

and thus $\tau(n) = 0$.

Proof. We prove the statement by induction on d. For $d = 0$ (i.e., $4 \le n \le 13$), the statement can be verified directly (e.g., see Fig. 3). Let $d > 0$.

(i) Let n belong to the interval $\frac{10\cdot 4^d+2}{3} \le n < \frac{20\cdot 4^d+1}{3}$. Then both $m = \lceil \frac{n}{2} \rceil$ and $n - m = \lfloor \frac{n}{2} \rfloor$ are in the interval (ii). If n is even, the values of $K_j(m) + K_{-2-j}(m)$ are

$$\begin{aligned} K_0(m) + K_1(m) &= 3\mathcal{B}(m) + O_{\mathcal{B}}(\lceil \tfrac{m}{2} \rceil)\,,\\ K_2(m) + K_2(m) &= 2\mathcal{B}(m) + O_{\mathcal{B}}(\lceil \tfrac{m}{2} \rceil)\,, \end{aligned}$$

implying that $J(n) = \{2\}$. If n is odd, the values of $K_j(m) + K_{-2-j}(m-1)$ are

$$\begin{aligned} K_0(m) + K_1(m-1) &= \mathcal{B}(m) + 2\mathcal{B}(m-1) + O_{\mathcal{B}}(\lceil \tfrac{m}{2} \rceil)\,,\\ K_1(m) + K_0(m-1) &= 2\mathcal{B}(m) + \mathcal{B}(m-1) + O_{\mathcal{B}}(\lceil \tfrac{m}{2} \rceil)\,,\\ K_2(m) + K_2(m-1) &= \mathcal{B}(m) + \mathcal{B}(m-1) + O_{\mathcal{B}}(\lceil \tfrac{m}{2} \rceil)\,, \end{aligned}$$

also implying that $J(n) = \{2\}$. From (71) and (72), it follows that $c_{0,n} = 2$, $c_{1,n} = 1$, $c_{2,n} = 3$ and $h_{0,n} = h_{1,n} = h_{2,n} = \lfloor \frac{n}{2} \rfloor$. Statement (i) now follows by induction from formula (59).

Statements (ii) and (iii) are proved similarly. We omit the details. □

Theorem 24 implies that essentially the same sequence $(\tau(n))_{n\ge 1}$ arises in base 7, with the only difference being that 1s and 2s are interchanged.

9 $K(n)$ for base 10

In base $b = 10$, the case studied by Kaprekar and others, the index set is $\mathcal{I} = \{0, 1, 2, \ldots, 8\}$ (modulo 9), and, from (63), there are five distinct $K'_i(n)$, namely $K'_0(n) = K'_7(n)$, $K'_1(n) = K'_6(n)$, $K'_2(n) = K'_5(n)$, $K'_3(n) = K'_4(n)$, and

$K_8'(n)$. There are nine variables $c_{i,n}$, $h_{i,n}$, and $K_i(n)$, with $0 \le i \le 8$. The values of $K_i(n)$ for $n \le 7$ are shown in Table 9. Then

$$K(n) \;=\; \min_{0\le i\le 8} K_i(n) \;=\; 10^{E(n)} + \text{ terms of smaller order}\,. \tag{83}$$

We have already seen $E(n)$ and $K(n)$ for $n \le 7$ in Table 1. Tables 10 and 11 extend these values to $n = 16$, going far enough that we can see – and confirm! – the values for $K(4), \ldots, K(8)$, and $K(16)$ found by Kaprekar and Narasinga Rao more than fifty years ago (see the discussion in the Introduction). The first 100 terms of these two sequences can be seen in entries A230857 and A006064 in [12].

	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$
$K_i(1)$	$\mathbf{0}$	10	2	12	4
$K_i(2)$	117	109	$\mathbf{101}$	111	103
$K_i(3)$	$10^{13}+116$	$10^{13}+9$	$\mathbf{10^{13}+1}$	$10^{13}+11$	$10^{13}+3$
$K_i(4)$	$2\cdot 10^{24}+115$	$10^{24}+108$	$3\cdot 10^{24}+116$	$10^{24}+110$	$\mathbf{10^{24}+102}$
$K_i(5)$	$2\cdot 10^{E(5)}+115$	$10^{E(5)}+108$	$3\cdot 10^{E(5)}+116$	$10^{E(5)}+108$	$\mathbf{10^{E(5)}+102}$

	$i=5$	$i=6$	$i=7$	$i=8$
$K_i(1)$	14	6	16	8
$K_i(2)$	113	105	115	107
$K_i(3)$	$10^{13}+13$	$10^{13}+5$	$10^{13}+15$	$10^{13}+7$
$K_i(4)$	$10^{24}+112$	$10^{24}+104$	$10^{24}+114$	$10^{24}+106$
$K_i(5)$	$10^{E(5)}+112$	$10^{E(5)}+104$	$10^{E(5)}+114$	$10^{E(5)}+106$

Table 9: Base 10: $K_i(n)$ for $n \le 5$; the value of $K(n)$ (A006064) is shown in bold font. In the $K_i(5)$ rows, $E(5) = (10^{13}+116)/9 = 1111111111124$ as in Table 1. See also Fig. 9 in Appendix B.

n	$E(n)$
8	$(2\cdot 10^{24}+214)/9$
9	$(10^{(10^{13}+116)/9}+10^{24}+214)/9$
10	$(2\cdot 10^{(10^{13}+116)/9}+214)/9$
11	$(10^{(2\cdot 10^{13}+16)/9}+10^{(10^{13}+116)/9}+10^{13}+114)/9$
12	$(2\cdot 10^{(2\cdot 10^{13}+16)/9}+2\cdot 10^{13}+14)/9$
13	$(10^{E(7)}+10^{(2\cdot 10^{13}+16)/9}+2\cdot 10^{13}+14)/9$
14	$(2\cdot 10^{E(7)}+2\cdot 10^{13}+14)/9$
15	$(10^{E(8)}+10^{E(7)}+E(7)-2)/9$
16	$(2\cdot 10^{E(8)}+E(8)-2)/9$

Table 10: Base 10: $E(n)$ for $8 \le n \le 16$, extending Table 1; $E(7) = (10^{24}+10^{13}+115)/9$. See also Fig. 9 in Appendix B.

n	$K(n)$
8	$10^{E(8)} + 10^{24} + 103$
9	$10^{E(9)} + 10^{24} + 103$
10	$10^{E(10)} + 10^{(10^{13}+116)/9} + 103$
11	$10^{E(11)} + 10^{(10^{13}+116)/9} + 103$
12	$10^{E(12)} + 10^{(2\cdot 10^{13}+16)/9} + 10^{13} + 3$
13	$10^{E(13)} + 10^{(2\cdot 10^{13}+16)/9} + 10^{13} + 3$
14	$10^{E(14)} + 10^{(10^{24}+10^{13}+115)/9} + 10^{13} + 3$
15	$10^{E(15)} + 10^{(10^{24}+10^{13}+115)/9} + 10^{13} + 3$
16	$10^{E(16)} + 10^{(2\cdot 10^{24}+214)/9} + 10^{24} + 104$

Table 11: Base 10: $K(n)$ for $8 \le n \le 16$, extending Table 1. See also Fig. 9 in Appendix B.

The first 100 terms of the base 10 generalized Thue–Morse sequence $\tau(n)$ are as follows:

$$0, 2, 2, 4, 4, 4, 4, 6, 6, 6, 6, 6, 6, 6, 6, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 1, 1, 1,\\ 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 8, 3, 8, 3, 8, 3,\\ 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, 8, 3, \ldots \tag{84}$$

(A239896). This can be rewritten as $0, 2^2, 4^4, 6^8, 8^{16}, 1^{31}, (8,3)^{\ge 19}$, but now, unlike the base 4 case, there is no obvious pattern.

10 Growth of $K(n)$

In this section we discuss the rate of growth of $K(n)$ for a fixed b.

The following theorem generalizes the inequalities (32) and (37). It implies that, for any base b, $\{K(n), n \ge 1\}$ is a sequence of exponential type (cf. Lemma 4).

Theorem 27. *For $b \ge 2$ and $n \ge 1$,*

$$K(n+1) \ > \ bK(n)\,, \tag{85}$$

except that for odd $b \ge 5$, we only have

$$K(3) \ > \ (b-1)K(2)\,. \tag{86}$$

Proof. For $n = 1$ and 2, the statement can be verified directly from Theorems 9 and 10.

For $n = 3$ and $b = 3$ or 5, the statement follows from Table 1. For $n = 3$ and odd $b \ge 7$, Corollary 21, Theorem 9, and (55) imply that $K(4) \ge b^{E(4)} + 1 + K(2) = b^3 + b + 2$. Since $K(3) = b^2 + 1$ (by Theorem 10), we have $K(4) \ge bK(3)$.

For $n \geq 4$ and any b, as well as for $n = 3$ and even b, the inequality (50) and Theorem 15 imply $K(n+1) \geq b^{E(n+1)} \geq b^{E(n)+2} \geq bK(n)$. □

Since $K(n)$ grows rapidly, it is appropriate to describe its value by a tower of exponentials. For $b \geq 2$, any number $u \geq 1$ can be written in a unique way as a "tower"

$$u \;=\; b^{b^{\cdot^{\cdot^{\cdot^{b^{\omega}}}}}}\,, \tag{87}$$

with $0 < \omega \leq 1$. If this tower contains $h-1$ b's and one ω, we call h the base b *height* of u, denoted by $\mathrm{ht}(u)$. Then $\mathrm{ht}(u)$ is one more than the number of times one has to take logarithms to the base b of u until reaching a number $\omega \leq 1$.

Examination of the data in Tables 1, 2, 5, 7, 11 (and in the more extended tables in [12]) suggests the following conjecture.

Conjecture 28. *It appears that:*
(i) If $b = 2$ *and* $n \geq 2$*, then* $\mathrm{ht}(n) = \lceil \log_2(n) \rceil + 3$*;*
(i) If $b = 3$ *and* $n \geq 3$*, then* $\mathrm{ht}(n) = \lceil \log_2(n/5) \rceil + 4$*;*
(iii) If $b \geq 4$ *is even and* $n \geq 2$*, then* $\mathrm{ht}(n) = \lceil \log_2(n) \rceil + 2$*;*
(iii) If $b \geq 5$ *is odd and* $n \geq 2$*, then* $\mathrm{ht}(n) = \lceil \log_2(n) \rceil + 1$*.*

For example, in base $b = 10$,

$$K(3) \;=\; 10^{13} + 1 \;=\; 10^{10^{10^{0.04686\ldots}}}\,, \tag{88}$$

which has height 4. The heights of $K(2)$ through $K(16)$ in base 10 (see (2) and Tables 1, 11) are $3, 4, 4, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 6$, respectively, in agreement with the conjecture.

There are two reasons for believing the conjecture. First, it is true in every case that we have checked. Second, from Section 6, $K(n)$ is very roughly equal to

$$b^{(K(\lceil \frac{n}{2} \rceil)+K(\lfloor \frac{n}{2} \rfloor))/(b-1)}\,, \tag{89}$$

which suggests that the height of the tower for $K(n)$ is one greater than the height of the tower for $K(\lceil \frac{n}{2} \rceil)$, which would leads to the formulas in the conjecture. However, two difficulties arise when trying to make this argument rigorous. One is the fact that if u in (87) has height h, and $\omega(u)$ is very close to 1, b^u can have height $h+2$ instead of $h+1$. This seems not to happen with $K(n)$, but we cannot rule out that possibility, even for base 2. The second difficulty is that (89) ignores the choices that must be made (for $b \geq 4$) among the $K_i(n)$ when determining $K(n)$.

The following example shows the first of these difficulties in a simpler setting. Consider the sequence defined by the recurrence

$$a(1) = 0, \;\; a(n) = 2^{a(\lceil \frac{n}{2} \rceil)+a(\lfloor \frac{n}{2} \rfloor)} \quad \text{for } n \geq 2\,. \tag{90}$$

This is similar to the recurrence for $K(n)$ in base 2 given in (31) and (30), except that the additive terms on the right-hand sides of those equations are missing. The initial values of $a(n)$ for $n = 1, 2, 3, \ldots$ are

$$0, 1, 2, 4, 8, 16, 64, 2^8, 2^{12}, 2^{16}, 2^{24}, 2^{32}, 2^{80}, 2^{128}, 2^{320}, 2^{512}, 2^{4352}, \ldots,$$

(A230863). The heights of $a(n)$ for $n = 2, ..., 10$ are $1, 1, 2, 3, 4, 4, 5, 5, 5, 5$. For $11 \le n \le 40$, if $9 \cdot 2^{i-1} < n \le 9 \cdot 2^i$ then $\mathrm{ht}(a(n)) = i + 5$, although here we do not know if this will hold for all n.[17]

For small values of n, of course, there is no difficulty in computing the height of $K(n)$. From Theorems 9 and 10, for example, we have $\mathrm{ht}(K(2)) = 4$ if $b = 2$, or 3 if $b \ge 3$, and

$$\mathrm{ht}(K(3)) = \begin{cases} 5 & \text{if } b = 2, \\ 4, & \text{if } b = 3, \\ 4, & \text{if } b \ge 4 \text{ is even}, \\ 3, & \text{if } b \ge 5 \text{ is odd}. \end{cases} \quad (91)$$

Acknowledgments

The authors are grateful to the late Donovan Johnson and Reinhard Zumkeller for carrying out computations during the early stages of this investigation.

References

[1] N. Agronomof. Problem 4421. *L'Intermédiaire des mathématiciens*, 21:147, 1914.

[2] J.-P. Allouche and J. Shallit. The ubiquitous Prouhet–Thue–Morse sequence. In C. Ding, T. Helleseth, and H. Niederreiter, editors, *Sequences and Their Applications: Proceedings of SETA '98*, pages 1–16. Springer, 1999.

[3] M. Gardner. *Time Travel and Other Mathematical Bewilderments*, chapter 9, Section 6, Self-numbers, pages 115–117, 122. W. H. Freeman, New York, 1988. Based on his *Mathematical Games* columns in the *Scientific American* for March and April, 1975.

[4] J. Janyška, M. Modugno, and R. Vitolo. An algebraic approach to physical scales. *Acta Appl. Math.*, 110:1249–1276, 2010.

[5] D. R. Kaprekar. Cycles of recurring decimals (from $N = 3$ to 161 and some other numbers). Published by the author, 311 Devlali Camp, Devlali, India, 1951.

[17] The difficulty lies in the fact that terms $a(n)$, for which the top entry in the tower, ω, is 1 or very close to 1, could disrupt the pattern of the heights.

[6] D. R. Kaprekar. Self-numbers. *Scripta Math.*, 22:80–81, 1956.

[7] D. R. Kaprekar. Puzzles of the self-numbers. Published by the author, 311 Devlali Camp, Devlali, India, 1959.

[8] D. R. Kaprekar. Self numbers and numbers with 3 and 4 generators. Published by the author, 311 Devlali Camp, Devlali, India, 1962.

[9] D. R. Kaprekar. The mathematics of the new self-numbers. Published by the author, 311 Devlali Camp, Devlali, India, 1963.

[10] D. R. Kaprekar. The mathematics of the new self-numbers (Part V). Published by the author, 311 Devlali Camp, Devlali, India, 1967.

[11] A. Narasinga Rao. On a technique for obtaining numbers with a multiplicity of generators. *Math. Student*, 34:79–84, 1966.

[12] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences. Published electronically at https://oeis.org. Accessed: 2022-01-01.

[13] B. Recamán Santos (proposer) and D. W. Bragg (solver). Problem E2408. *Amer. Math. Monthly*, 80, 81:434, 407, 1973, 1974.

[14] R. Schorn. Kaprekar's Sequence and his "Selfnumbers". *DERIVE Newsletter*, 53:30–32, 2004.

[15] K. B. Stolarsky. The sum of a digitaddition series. *Proc. Amer. Math. Soc.*, 59:1–5, 1976.

Appendix A Practical Computations

The most difficult task in programming the algorithm in Section 6 is in handling the very large numbers that appear. As we saw in Section 10, to compute $K(100)$ in base 10, for example, we have to work with numbers that are of the order of a tower of 10's of height 9. This problem was solved using C++ by defining a special type of object (we called it the "sparse radix representation"), which represents an integer in the algebraic form:

$$\frac{1}{\gamma}\left(\alpha_1 b^{d_1} + \alpha_2 b^{d_2} + \cdots + \alpha_k b^{d_k}\right), \tag{92}$$

where γ and the α_i are integers, with $\gamma \geq 1$ (typically we have $\gamma = 1$ or $\gamma = b - 1$) and $1 \leq \alpha_i < b$, and d_i are objects of the same type, satisfying $d_1 > d_2 > \cdots > d_k$. These objects support the operations of comparison and addition, as well as multiplication by positive rational numbers (so these objects form a semi-vector space [4] over the semi-field of positive rationals $\mathbb{Q}^+$).

Comparison of two objects with $\gamma = 1$ is done recursively, starting by comparing the highest-order terms, and if these are tied, comparing the next-to-highest terms, and so on. Similarly, addition of two objects with $\gamma = 1$ is done by first combining powers of b with equal exponents, and then reducing the coefficients into the required range (e.g., βb^d with $\beta \geq b$ is replaced with $(\beta \bmod b)\, b^{d+\lfloor \beta/b \rfloor}$, where the addition in the exponents is performed recursively). If the denominator γ of either of the two numbers is not 1, the objects (i.e., the coefficients α_i) are first multiplied by L, the least common multiple of their γ's, the coefficients are then reduced into the required range, and the resulting objects are compared or added as above (in the case of addition we set γ for the sum equal to L). We remark that although the representation of an integer in the form (92) is not unique (e.g., $\frac{1}{b-1}\,(b^1 + (b-2)b^0)$ and $2b^0$ represent the same integer 2), any two such representations can be efficiently compared and tested for equality.

We also made extensive use of the following PARI/GP program for computing Gen(u) and $F(u)$. The procedure `Gen(u,b)` uses the recurrence in (13) to compute Gen(u) in base b for $u \in \mathbb{N}$. For example, `Gen(10^13+1,10)` would return the three generators 9999999999892, 9999999999901, 10000000000000 of $10^{13} + 1$ in base 10. Correspondingly, the number of generators may be obtained as `#Gen(u,b)`.

```
/* The PARI/GP procedure Gen(u,b) */

{ Gen(u,b=10) = my(d,m,k);
  if(u<0 || u==1, return([]); );
  if(u==0, return([0]); );

  d = #digits(u,b)-1;
  m = u\b^d;
  while( sumdigits(m,b) > u - m*b^d,
    m--;
    if(m==0, m=b-1; d--; );
  );
  k = u - m*b^d - sumdigits(m,b);

  vecsort( concat( apply(x->x+m*b^d,Gen(k,b)),
                   apply(x->m*b^d-1-x,Gen((b-1)*d-k-2,b)) ) );
}
```

Appendix B Flow-charts of computations for bases 2 through 10

To illustrate how the computation of the $K_i(n)$ and $E(n)$ proceeds, and to show the complexity of their recursive structure, in Figures 1–9 we provide

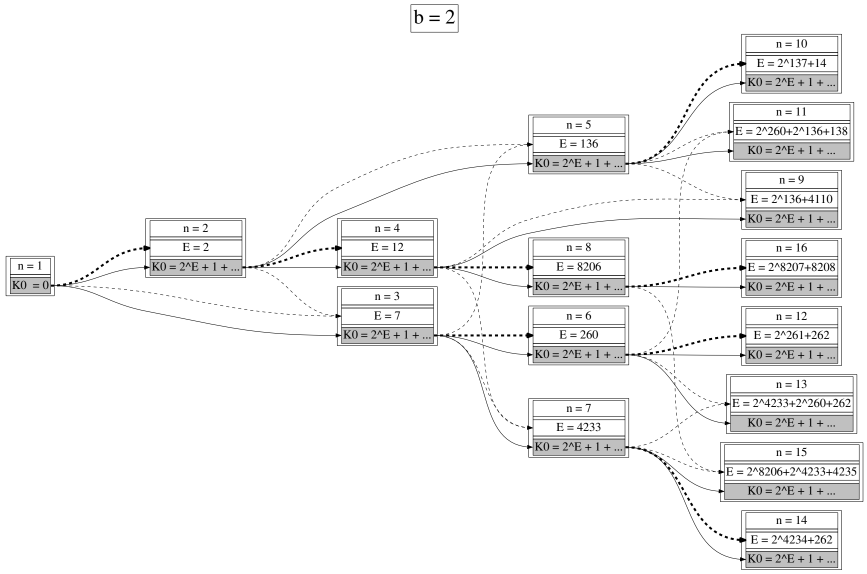

Figure 1: Flow-chart illustrating the calculation of $E(n)$, $K_i(n)$, and $K(n)$ for $n \leq 16$ in base $b = 2$. For a description of these flow-charts see Appendix B.

flow-charts of the computation for bases $b = 2, 3, \ldots, 10$.

For each $n \leq 16$, there is a stack of boxed nodes containing the values of $\mathrm{E} := E(n)$ and $\mathrm{K}i := K_i(n)$ for $i \in \mathcal{I}$. The shaded node denotes the value of $K(n)$ (having the minimum value among all K-nodes in the stack).

Each arc between K-nodes corresponds to an instance of the formula (59) and connects the K-nodes corresponding to $K_{i-2c_{i,n}}(h_{i,n})$ and $K_i(n)$. So, for $n > 1$, each $K_i(n)$ node has exactly one incoming arc. The value of $K_i(n)$ is given in the form $c_{i,n}(b^{\mathrm{E}} + 1) + \ldots$, where "…" stands for the corresponding value of $K_{i-2c_{i,n}}(h_{i,n})$ (located at the starting node of the incoming arc).

Each E-node for $n > 1$ has two incoming arcs, shown dashed (which may coincide and be shown in bold when n is even), illustrating the formula

$$E(n) = \frac{K_j\left(\left\lceil \frac{n}{2} \right\rceil\right) + K_{-j-2}\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + 2}{b - 1},$$

which holds for some $j \in \mathcal{I}$ as follows from (67). The incoming arcs start at the K-nodes corresponding to $K_j\left(\left\lceil \frac{n}{2} \right\rceil\right)$ and $K_{-j-2}\left(\left\lfloor \frac{n}{2} \right\rfloor\right)$. Since there may be several equally good choices for j, we assume that j is such that $K_j\left(\left\lceil \frac{n}{2} \right\rceil\right) = K\left(\left\lceil \frac{n}{2} \right\rceil\right)$ whenever equality holds in (66); otherwise j is taken to be the smallest of the possible values. So if equality holds in (66), there exists an incoming arc from the shaded K-node with the value $K\left(\left\lceil \frac{n}{2} \right\rceil\right)$.

It can be seen that for $b = 4$ (Fig. 3) and $b = 7$ (Fig. 6) there are no incoming arcs to the E-node from shaded K-nodes when $n \in \{13, 15, 16\}$. Similarly, there no such arcs for $b = 9$ (Fig. 8) and $n \in \{8, 9, 10, 11, 12, 13, 14, 16\}$. That is, for these b and n, the formula (66) does not hold.

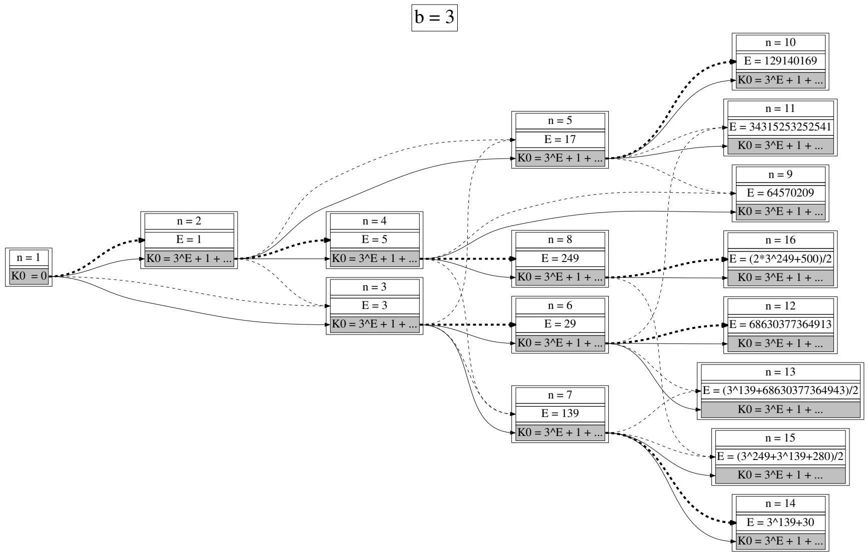

Figure 2: Flow-chart illustrating the calculation of $E(n)$, $K_i(n)$, and $K(n)$ for $n \leq 16$ in base $b = 3$. Note that apart from the labels, this is the same flow-chart as in Fig. 1.

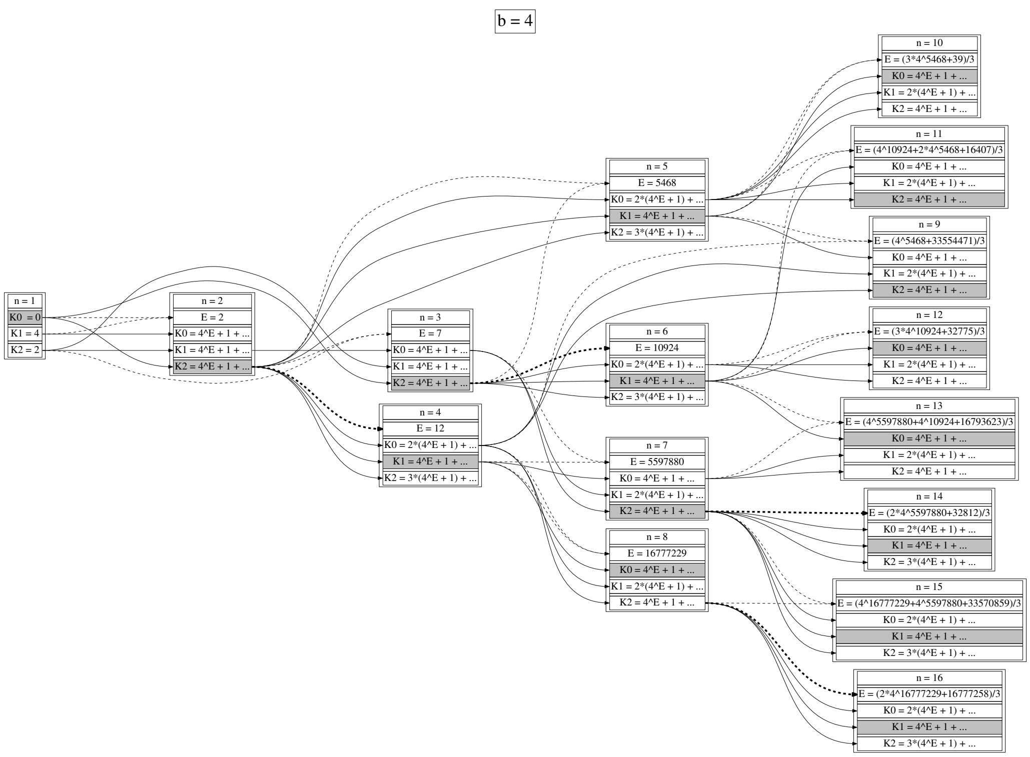

Figure 3: Flow-chart illustrating the calculation of $E(n)$, $K_i(n)$, and $K(n)$ for $n \leq 16$ in base $b = 4$.

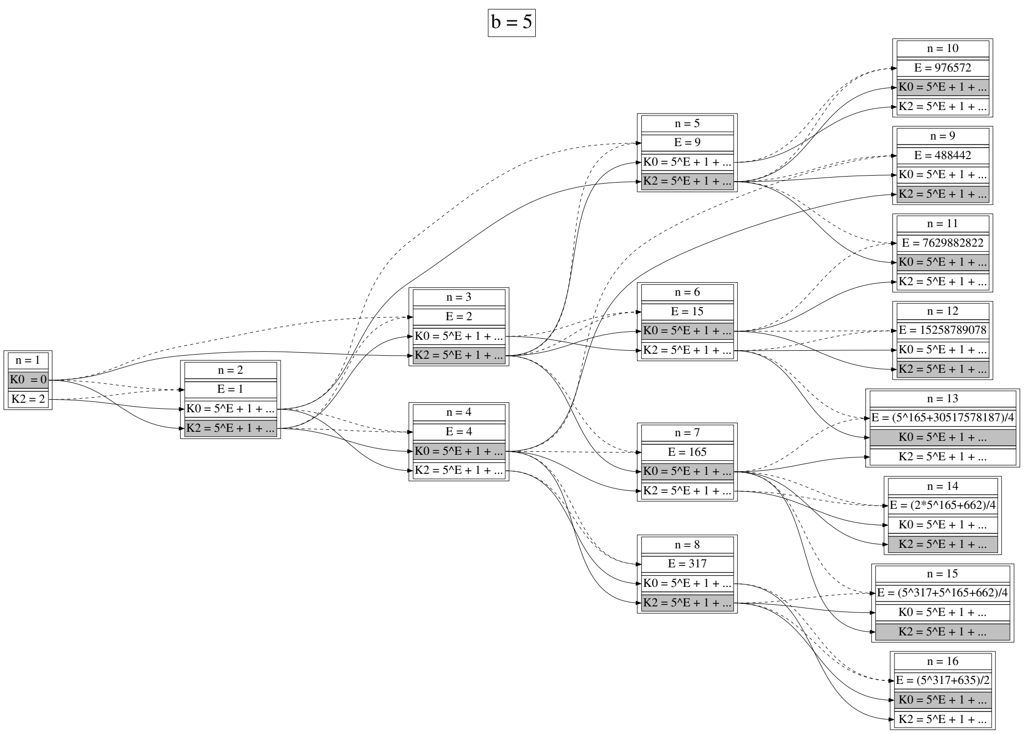

Figure 4: Flow-chart illustrating the calculation of $E(n)$, $K_i(n)$, and $K(n)$ for $n \leq 16$ in base $b = 5$.

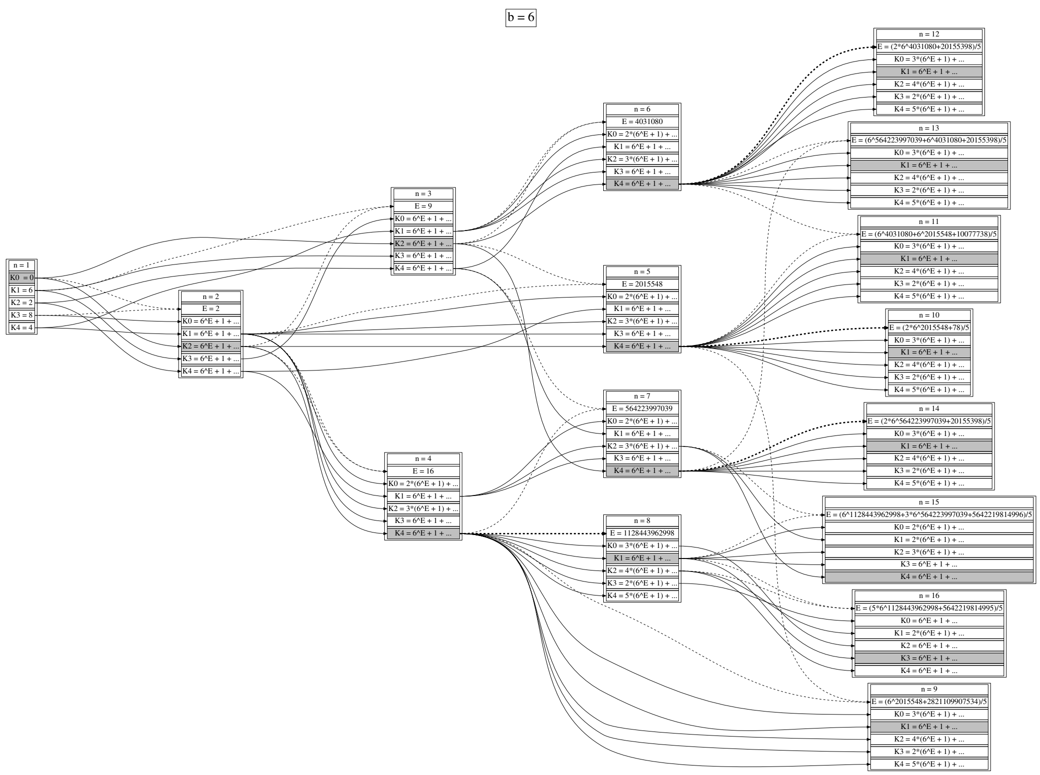

Figure 5: Flow-chart illustrating the calculation of $E(n)$, $K_i(n)$, and $K(n)$ for $n \leq 16$ in base $b = 6$.

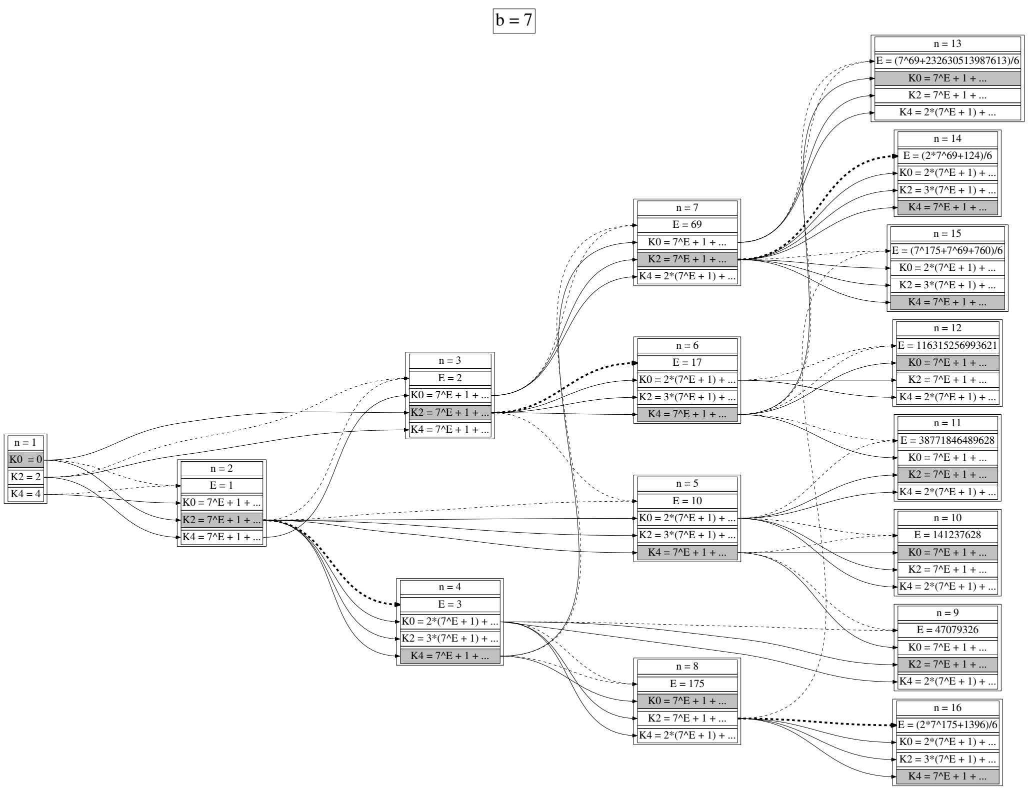

Figure 6: Flow-chart illustrating the calculation of $E(n)$, $K_i(n)$, and $K(n)$ for $n \leq 16$ in base $b = 7$. Note that apart from the labels, this is the same flow-chart as in Fig. 3.

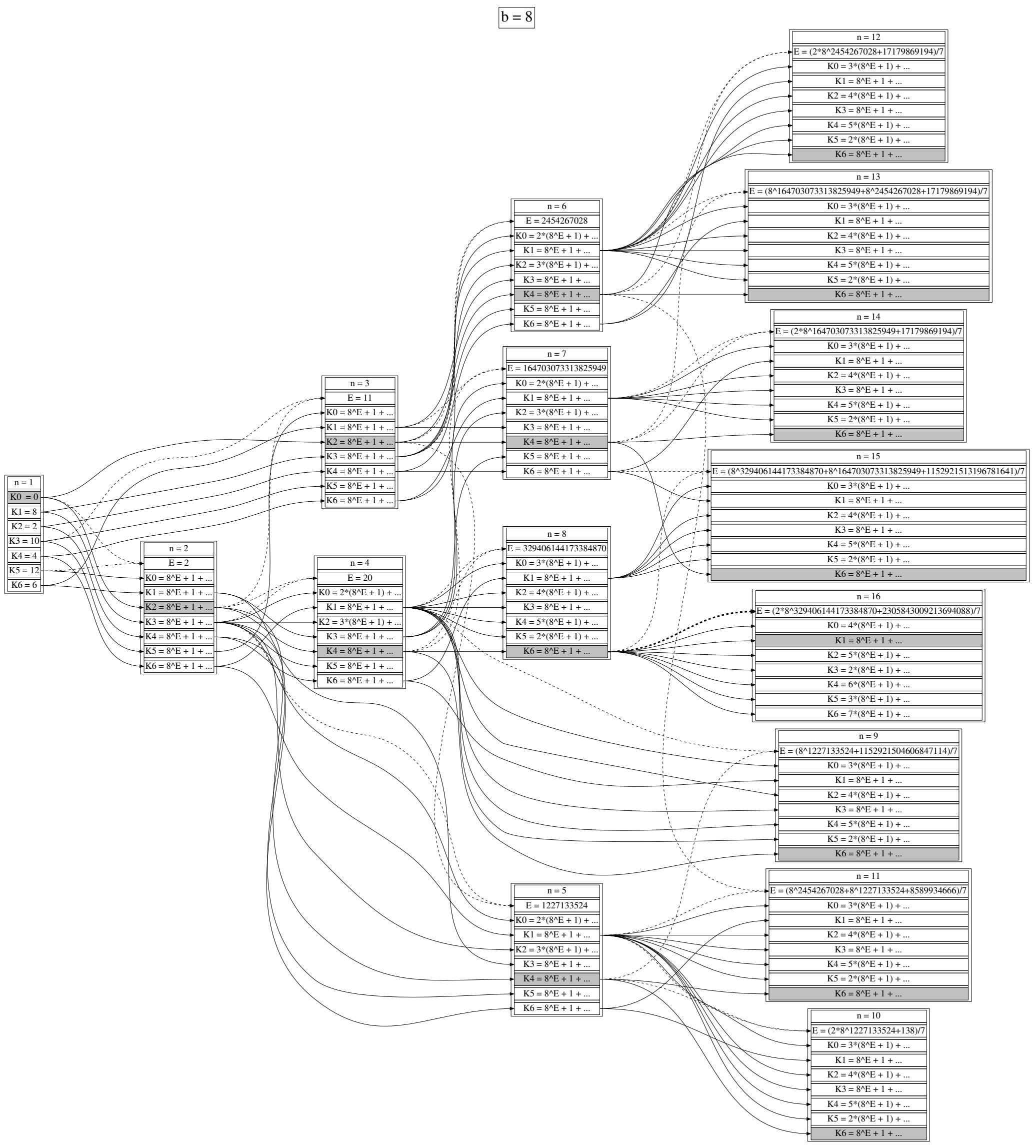

Figure 7: Flow-chart illustrating the calculation of $E(n)$, $K_i(n)$, and $K(n)$ for $n \leq 16$ in base $b = 8$.

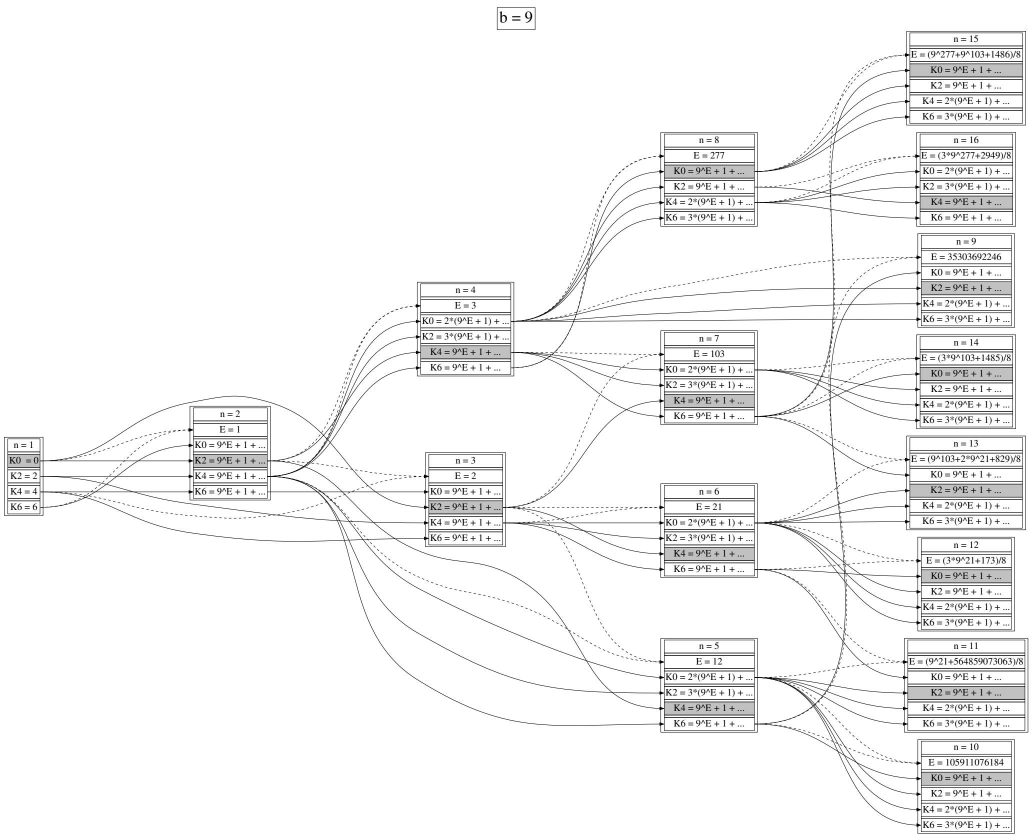

Figure 8: Flow-chart illustrating the calculation of $E(n)$, $K_i(n)$, and $K(n)$ for $n \leq 16$ in base $b = 9$.

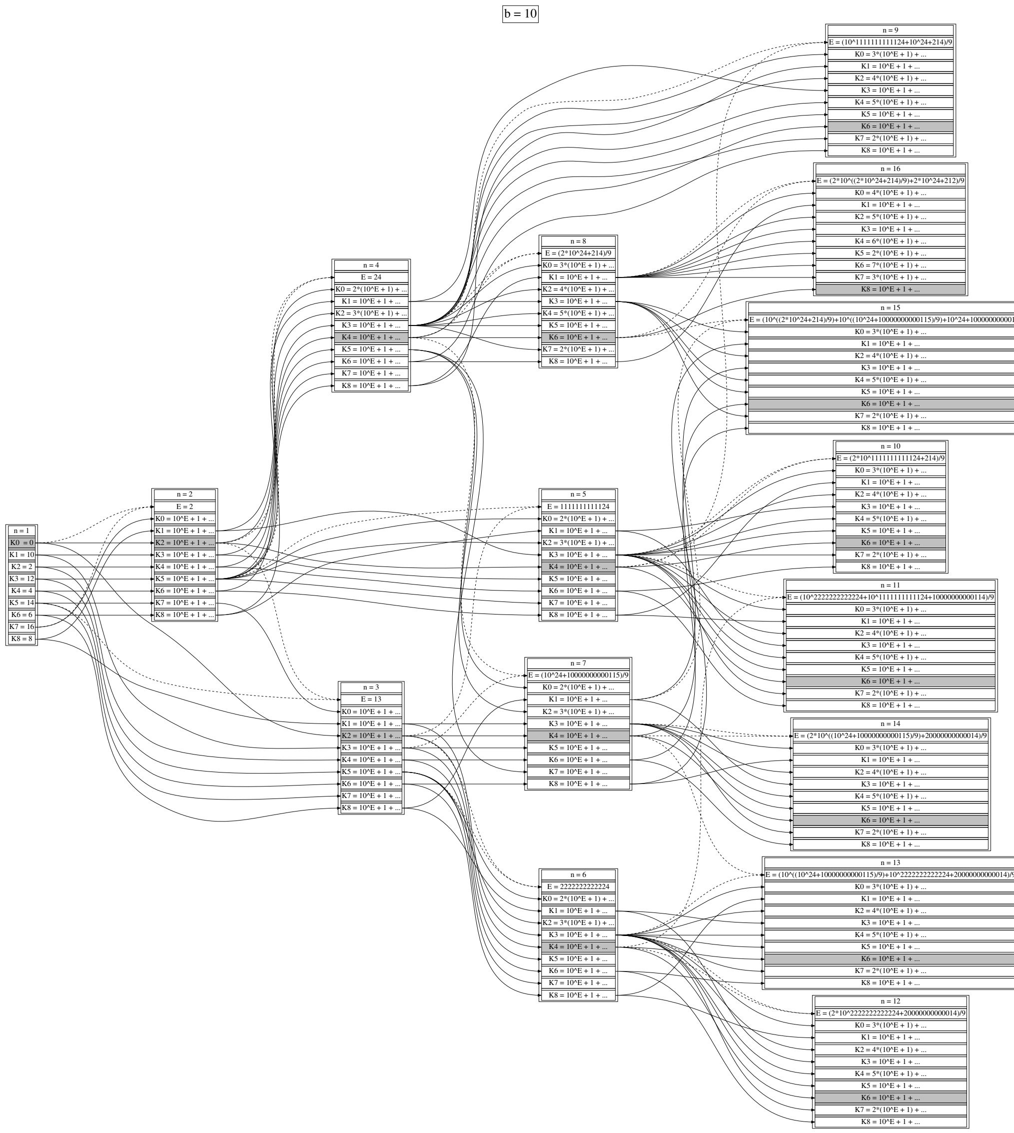

Figure 9: Flow-chart illustrating the calculation of $E(n)$, $K_i(n)$, and $K(n)$ for $n \leq 16$ in base $b = 10$.