Open Journal of Mathematical Sciences

PSR Press

Article

Colombian numbers, Kaprekar's test, and generator sequences

Juan Gabriel Triana

Department of Basic Science, Universitaria Agustiniana. Bogotá, Colombia; juang.triana@uniagustiniana.edu.co

Received: 23 Apr 2026; Revised: 20 Jun 2026; Accepted: 25 Jun 2026; Published: 01 Sep 2026

Abstract: Colombian numbers, also known as self-numbers, are natural numbers that do not have any generator under the digit-addition process $\phi(n) = n + s(n)$, where $s(n)$ denotes the sum of digits of n. Here, a generator of n is any integer x such that $\phi(x) = n$. In this paper we study both structural and computational aspects of these numbers, including Kaprekar's test for their identification and the concept of junction numbers. We present computational counts of numbers with zero, one, and two generators in intervals of the form $[10^k + 1, 10^{k+1}]$, revealing highly regular digit patterns. Furthermore, we analyze the sequences X and Y, defined by selecting the largest and smallest generators respectively, and compare their lengths across the same intervals. The results show a remarkable stability in the relative frequencies of the three possible outcomes (X shorter, Y shorter, or equal length) suggesting the existence of asymptotic regularities. Finally, we prove that every non-Colombian number can be obtained by successive applications of ϕ starting from a Colombian number.

Keywords: Colombian numbers, digit-addition map, generators of integers

MSC: 11A63, 11B99.

1. Introduction

Let n be a natural number. We define $\phi(n) = n + s(n)$, where $s(n)$ denotes the sum of digits of n. The process described by ϕ is known as digit-addition, term coined by Dattatreya Kaprekar (1905-1986) [1]. If $N = \phi(n)$ then n is a generator of N. A natural number without any generator is called a Colombian number [2], also known as self number [3]. The first Colombian numbers, cf. (OEIS A003052), are:

$$1, 3, 5, 7, 9, 20, 31, 42, 53, 64, 75, 86, 97, 108, \ldots.$$

In [4], problem E2408, the name base b Colombian numbers was coined. In [5], a recurrence relation for generating base b Colombian numbers was introduced in the form $C_k = (b-2)b^{k-1} + C_{k-1} + b - 2$. It is easy to verify that not all base b Colombian numbers are generated by the recurrence, but it is enough to conclude that for each base b there are infinitely many base b Colombian numbers. In [6], it was proved that for every $b > 2$, the numbers $2b$, $4b + 2$ and $(b + 1)^2$ are base b Colombian numbers, in [7] were introduced the universal generated numbers as the numbers generated in every base.

We may consider families of Colombian numbers with an additional property, such as Colombian prime numbers (OEIS A006378). For instance, we can consider numbers such that n, n^2, $\ldots n^m$ are Colombian numbers, it is easy to verify that 1 is the first number with the identity for each m. but the second number in the list for $m = 2$, $m = 3$ and $m = 4$ are 3, 714 and 81826 respectively. As an interesting fact the 41^{st} Mersenne prime $2^{24036583} - 1$ is a Colombian number.

The literature on Colombian numbers has developed along several complementary directions. A first line of work concerns their basic arithmetic properties and the problem of deciding whether a given integer has a generator under the digit-addition map $\phi(n)$. A second line studies generalizations to other numeration systems, leading to the notion of base b Colombian numbers and to recurrence formulas that generate infinite families of such numbers. A third line focuses on numbers with multiple generators, especially numbers with at least two generators, called junction numbers, and the problem of determining the smallest integer with exactly m generators (denoted by $K(m)$).

Despite these advances, less attention has been given to the inverse dynamical structure induced by the digit-addition process, namely, to the iterative study of generator chains leading from a given number back to a Colombian number. Likewise, although Kaprekar's test provides an efficient way to detect generators, systematic computational counts according to the exact number of generators in large decimal intervals remain limited.

The contribution of this paper is threefold. First, we show that Kaprekar's test yields all generators of a given non-Colombian number, and we use this fact to obtain computational counts for numbers with 0, 1, and 2 generators over decimal intervals of the form $[10^k + 1, 10^{k+1}]$. Second, we introduce and compare the lengths of the sequences X and Y, defined by selecting the largest and smallest generators respectively. Finally, we prove that every non-Colombian number belongs to a finite inverse orbit of the digit-addition map ending at a Colombian number.

2. Finding Colombian numbers

In [8], Kaprekar gave a characterization of certain Colombian numbers. Consider the number $n = \sum_{k=0}^{m} d_k 10^k$, with $d_m, d_1, d_0 \neq 0$, if the digit sum satisfies $s(n) = 4, 15, 26$ or 37 then n is a Colombian number. In general, determining whether a number is Colombian requires extensive computational search, and for this reason we introduce several auxiliary lemmas that reduce the search space and provide theoretical support for the algorithmic classification of these numbers.

Lemma 1. *For any natural number n, let $s(n)$ denote the sum of the digits of n. Then $s(n) \equiv n \ (mod\ 9)$.*

Proof. Let $n = \sum_{k=0}^{m} d_k 10^k$. Since $10^k \equiv 1 \ (mod\ 9)$ we have

$$\begin{aligned} n &\equiv \sum_{k=0}^{m} d_k 10^k \ (mod\ 9) \\ &\equiv \sum_{k=0}^{m} d_k \ (mod\ 9). \end{aligned}$$

Hence, $s(n) \equiv n \ (mod\ 9)$. □

The following lemma establishes bounds to search for possible generators of a number n.

Lemma 2. *Let n be an m-digit number, if x is a generator of n then $n - 9m \leq x \leq n - 1$.*

Proof. Let $x = \sum_{k=0}^{r} d_k 10^k$ a generator of n, therefore $\phi(x) = x + \sum_{k=0}^{r} d_k = n$. Since $\phi(x) = n$ we have $x < n$.

As n is an m-digit number and $x < n$, it follows that $r + 1 \leq m$ where $r + 1$ denotes the number of digits of x. Moreover, since $d_k \leq 9$ for each k, we have $\phi(x) = n \leq x + 9(r + 1) \leq x + 9m$. Thus, $n - 9m \leq x \leq n - 1$. □

Kaprekar introduces a test to find out if a given number n is a Colombian number or not [9], by the following steps:

1. Let $s(n)$ be sum of the digits of n.
2. Compute $c = \begin{cases} \dfrac{s(n)}{2} & \text{if} \quad s(n) \text{ is even} \\ \dfrac{s(n)+9}{2} & \text{if} \quad s(n) \text{ is odd} \end{cases}$
3. Consider the finite sequence $\{n - c - 9k\}_{k=0}^{m}$ where m is the number of digits of n.
4. If any of the numbers in the sequence generate n then the number is not a Colombian number, otherwise it is a Colombian number.

The numbers in the sequence $\{n-c-9k\}_{k=0}^{m}$ are candidates to generators of n, but they must be tested individually to determine whether they generate n.

If we want to perform an empirical study in which millions of numbers are classified according to their number of generators, the computational cost will be unfeasible, especially when n is large. The Lemma 2 provides an upper and lower bound for the potential generators of a number n, restricting the search to an interval of length $9m-1$, where m is the number of digits of n. Kaprekar's test, described above, reduces this search to just $m+1$ numbers, the following proposition ensures that the test is not only efficient but also complete obtaining all the generators of a number n.

Proposition 1. *Let n be a natural number. If n is a non-Colombian number then all the generators can be found by the test of Kaprekar.*

Proof. Let n be a natural number and x a generator of n, therefore $n = x + s(x)$ where $s(x)$ is the sum of digits of x. Since $s(x) \equiv x (mod\ 9)$, by Lemma 1, we have

$$\begin{aligned} s(x) &\equiv x \pmod 9 \\ s(x) &\equiv n - s(x) \pmod 9 \\ 2s(x) &\equiv n \pmod 9. \end{aligned}$$

Since 5 is the inverse of 2 modulo 9we get

$$s(x) \equiv 5n \pmod 9. \tag{1}$$

On the other hand, let $s(n)$ be the sum of the digits of n. If $s(n)$ is odd, there is c such that $2c = s(n) + 9$, therefore $2c \equiv s(n) \pmod 9$ and $2c \equiv n \pmod 9$. If $s(n)$ is even we get $c = \frac{s(n)}{2}$, therefore $2c = s(n)$. Since $s(n) \equiv n \pmod 9$ we obtain that $2c \equiv n \pmod 9$ thus

$$c \equiv 5n \pmod 9. \tag{2}$$

By Eqs. 1 and 2 we have $s(x) \equiv c\ (mod\ 9)$ thus $s(x)$ can be written in the form $s(x) = c + 9k$. Therefore, x can be expressed in the form $x = n - c - 9k$, by Lemma 2 we conclude that $0 \leq k \leq m$ where m is the number of digits of n. □

3. Empirical distribution of generators

In [8], Kaprekar introduces special cases of numbers where a generator can be found rapidly. For instance, consider the number $n = \sum_{k=0}^{m} d_k 10^k$, with $d_m, d_1, d_0 \neq 0$. if $\sum_{k=0}^{m} d_k = 11$, i.e. the sum of the digits of n is 11, then a generator of n is $n - 10$.

A natural number may have zero, one or multiple generators. Numbers with at least two generators are called junction numbers. For example, 3502 and 3493 are generators of 3512 but with the method above we just obtain 3502. We can state that 3502 and 3493 are co-generators of 3512, and 3512 is a junction. This rule to obtain a generator also applies to numbers such that the sum of the digits is 11 and ends in only one zero. However, we do not have a method to directly obtain all the generators of a number.

Kaprekar defines the sequence $n, \phi(n), \phi^2(n), \ldots$, to study junction numbers and co-generators [8], he was particularly interested in finding the smallest junction number with m generators, known as $K(m)$. The first numbers that can be generated by exactly two generators are (OEIS A230094):

$$101, 103, 105, 107, 109, 111, 113, 115, 117, 202, 204, \ldots.$$

Hence, the smallest number with two generators is $K(2) = 101$. In [10], $K(m)$ numbers (OEIS A006064) were widely studied and $K(3) = 10^{13} + 1$ was obtained, currently we know several numbers with three generators cf. (OEIS A230100). In [2], recurrences to find $K(m)$ for several bases b were introduced.

In this section we present computational counts of numbers with 0, 1 and 2 generators in intervals in the form $[10^k + 1, 10^{k+1}]$. The counts were obtained by implementing Kaprekar's test and verifying all the candidates.

Table 1. Quantity of numbers with 0,1 and 2 generators between the given values

interval $[10^k + 1, 10^{k+1}]$	Colombians	1 generator	2 generators
$k = 1$	8	82	0
$k = 2$	89	730	81
$k = 3$	881	7246	873
$k = 4$	8801	72406	8793
$k = 5$	88002	724005	87993
$k = 6$	880001	7240006	879993
$k = 7$	8800001	72400006	8799993
$k = 8$	88000001	724000006	87999993
$k = 9$	880000001	7240000006	879999993

In row 5 of Table 1 we observe one more Colombian number than expected and one fewer number with a single generator. This discrepancy may be explained by the fact that 10^6 is the only Colombian number of the form 10^k included in the table, whereas all other numbers of the form 10^k have exactly one generator.

Since the first number with three generators is $K(3) = 10^{13} + 1$ cf. [2], in the Table above we just consider until 2 generators. It is easy to verify that 10^{13}, $10^{13} - 99$ and $10^{13} - 108$ are co-generators of $10^{13} + 1$. Since $K(2) = 10^2 + 1$ and $K(3) = 10^{13} + 1$ One might conjecture that $K(4)$ is of the form $10^k + 1$ but this is false; in fact, $K(4) = 10^{24} + 102$. cf. [2]. It can be verified that $10^{24} - 107$, $9 \times 10^{24} - 98$, $10^{24} + 91$, $10^{24} + 100$ are co-generators of $10^{24} + 102$.

Let $Z(k)$, $O(k)$ and $T(k)$ be the number of numbers in the interval of the form $[10^k + 1, 10^{k+1}]$ with zero, one, and two generators respectively. The empirical patterns observed in Table 1 suggest a highly structured distribution of Colombian numbers and numbers with one or two generators. This regularity is consistent with asymptotic results on the density of Colombian numbers presented in [11]. The computational evidence leads to the following conjectures:

- Numbers with exactly one generator are abundant compared to those with zero, two, or more generators within intervals of the form $[10^k + 1, 10^{k+1}]$, for each k.
- Extending the table to numbers with three, four, and more generators may reveal patterns analogous to those already observed in $Z(k), O(k)$ and $T(k)$.

4. Results

In this section we introduce the sequence of natural numbers X by setting $x_0 = n > 0$ and x_{k+1} as the largest generator of x_k, i.e. $\phi(x_{k+1}) = x_k$. We may verify that X is a one-term sequence if and only if $x_0 = n$ is a Colombian number. For instance, if $x_0 = 2026$ we have that $x_{18} = 1783$ is the last term in the sequence.

Lemma 3. *X is a strictly decreasing finite sequence whose last term is a Colombian number.*

Proof. If $x_0 = n$ is a Colombian number, then x_0 is the sequence's first and last term.

Assuming that $x_k = n$ is not a Colombian number, there is a number v such that $\phi(v) = x_k$, therefore $x_{k+1} = v$. Since $\phi(v) > v$ we get $x_{k+1} < x_k$ with $x_k \in \mathbb{N}$ for each k. Thus, we may conclude that X is a strictly decreasing finite sequence.

Let $x_m = v$ be the last term of the sequence where v is not a Colombian number; therefore, there is w such that $\phi(w) = v$, but by definition of X we get $x_{m+1} = w$ which is a contradiction because x_m is the last term of the sequence. Hence, the last term of the sequence is a Colombian number. □

Similar to the sequence of Lemma 3, we define Y by $y_0 = n > 0$ and y_{k+1} as the smallest integer such that $\phi(y_{k+1}) = y_k$. This construction also yields a strictly decreasing finite sequence whose last term is a Colombian

number. For instance, if $y_0 = 2026$ the last term is $y_2 = 1996$. Since $K(2) = 101$, for each $n \leq 100$ we have $X = Y$.

In order to compare the lengths of the sequences X and Y, we perform a computational study over intervals in the form $[10^k + 1, 10^{k+1}]$. We denote by $len(X)$ and $len(Y)$ the lengths of sequences X and Y respectively. The results in Table 2 reveal a stability in the relative frequencies of the three outcomes Although the absolute counts grow rapidly with k, the proportions remain nearly constant across successive intervals, suggesting the existence of an asymptotic distribution.

Table 2. Computational counts, for each interval $[10^k + 1, 10^{k+1}]$, of the number of cases in which sequence X is shorter, sequence Y is shorter, or both sequences have the same length

$[10^k + 1, 10^{k+1}]$	$len(X) < len(Y)$	$len(Y) < len(X)$	$len(X) = len(Y)$
$k = 1$	0	0	90
$k = 2$	286	190	424
$k = 3$	2510	2270	4220
$k = 4$	25613	22956	41431
$k = 5$	256923	228000	415077
$k = 6$	2540816	2313025	4146159
$k = 7$	25072722	23340600	41586678

Lemma 3 shows that every integer lies in a finite inverse orbit of the digit-addition map, with Colombian numbers as terminal elements. Thus, we have that each non-Colombian number can be generated by successive applications of the digit-addition process starting from a Colombian number.

Corollary 1. *If n is a non-Colombian number then there exists a Colombian number c and $m \in \mathbb{N}$ such that $n = \phi^m(c)$.*

Proof. By Proposition 3, there exists a sequence $(x_k)_{k=0}^m$ with $x_0 = n$ and $x_m = c$ such that x_{k+1} is a generator of x_k, thus we have $x_k = \phi^{m-k}(c)$. Hence, $x_0 = n = \phi^m(c)$. □

5. Conclusions

In this paper, we investigated both structural and computational properties of Colombian numbers. We proved that every non-Colombian number belongs to a finite inverse orbit of the digit-addition map, and established that Kaprekar's test yields all generators of a given number.

Furthermore, the computational evidence presented in Table 1 and Table 2 highlights structural regularities. Table 1 shows that the distribution of numbers with zero, one, and two generators across decimal intervals exhibits a highly regular pattern. Table 2 compares the lengths of the sequences X and Y, revealing a remarkable stability in the relative frequencies of the three possible outcomes. These computational findings open new directions for theoretical investigation.

Future work includes establishing rigorous bounds for the length of inverse generator sequences and determining whether the observed counting functions admit exact closed-form expressions. Another important direction is to formalize the conjectured asymptotic distribution of the comparative lengths of the sequences X and Y, thereby connecting empirical stability with theoretical results.

Future work will focus on formalizing the conjectures proposed in §3, namely that numbers with exactly one generator are abundant compared to those with zero, two, or more generators, and that extending the table to numbers with three, four, or more generators may reveal patterns analogous to those already observed in $Z(k)$, $O(k)$ and $T(k)$.

Conflicts of Interest: The author declares no conflict of interest.

Data Availability: No data is required for this research.

Funding Information: No funding is available for this research.

Acknowledgments: The author thanks the reviewers and the editor for their careful reading and valuable comments and suggestions.

References

[1] Stolarsky, K. B. (1976). The sum of a digitaddition series. *Proceedings of the American Mathemati3cal Society, 59*(1), 1–5.

[2] Alekseyev, M. A., & Sloane, N. J. A. (2022). On Kaprekar's junction numbers. *Journal of Combinatorics and Number Theory, 12*(3), 115–155.

[3] Gupta, S. S. (2025). On some marvellous numbers of Kaprekar. In *Exploring the Beauty of Fascinating Numbers* (pp. 275–315). Springer.

[4] Recamán, B., Boyd, A. V., Wolk, B., Barnes, F. W., Goldberg, M., & Grosch, C. B. (1973). Elementary problems: E2408–E2413. *The American Mathematical Monthly, 80*(4), 434–435.

[5] Bange, D. W. (1974). Colombian numbers [Solution to Problem E2408]. *The American Mathematical Monthly, 81*(4), 407.

[6] Patel, R. B. (1991). Some tests for *k*-self-numbers. *The Mathematics Student, 56*(1–4), 206–210.

[7] Cai, T. (1996). On *k*-self-numbers and universal generated numbers. *The Fibonacci Quarterly, 34*(2), 144–146.

[8] Kaprekar, D. R. (1963). *The Mathematics of the New Self Numbers*. Privately printed.

[9] Athmaraman, R. (Ed.). (2004). *The Wonder World of Kaprekar Numbers*. The Association of Mathematics Teachers of India.

[10] Narasinga Rao, A. (1966). On a technique for obtaining numbers with a multiplicity of generators. *The Mathematics Student, 34*(2), 79–84.

[11] Zannier, U. (1982). On the distribution of self-numbers. *Proceedings of the American Mathematical Society, 85*(1), 10–14.