

Descriptions of Cantor Sets: A Set-Theoretic Survey and Open Problems

Mohsen Soltanifar ^{*,1}

*Biostatistics Division, Dalla Lana School of Public Health, University of Toronto, 620-155 College Street, Toronto, ON M5T 3M7, Canada.

ABSTRACT This paper synthesizes the principal descriptive set-theoretic perspectives on deterministic Cantor sets on the real line and charts directions for future study. After recounting their historical genesis and compiling an up-to-date taxonomy, we review the Borel hierarchy and four hierarchically ordered representations, general, nested, iterated-function-system (IFS), and q -ary expansion, presented from the most general to the most specific set-theoretic description of deterministic Cantor sets. We then present explicit closed-form descriptions for two thin families of measure-zero Cantor sets and, for the augmented “thick” family of positive Lebesgue measure, a new closed recursive formula for the endpoint maps governing every retained interval at every finite stage; we further show that the classical middle-third set lies in the intersection of all three families, providing a unifying example across the four representations. The survey closes by isolating several open problems in four directions, aiming to provide mathematicians with a coherent platform for further descriptive set-theoretic investigations into Cantor-type sets on the real line.

KEYWORDS

Cantor set
 Borel operations
 Formulas
 Set theory
 Fractal geometry

INTRODUCTION

The conceptual roots of the Cantor Sets lie in Georg Cantor’s 1872 introduction of the *derived set* (the set of all limit points of a given subset of $\mathbb{R}$), which laid the groundwork for perfect, nowhere-dense subsets of $\mathbb{R}$. Shortly thereafter, H. J. S. Smith’s 1875 note provided the first printed construction of a “Cantor-type” set by iteratively excising subintervals of $[0, 1]$, thereby demonstrating that a closed set can be uncountable yet possess arbitrarily small outer content, although the significance of his example went largely unnoticed at the time. Cantor returned to the theme in his six-part memoir *Über unendliche lineare Punktmannichfaltigkeiten* (1879–1884); Part V (written October 1882, published 1883) contains the now-classical *3-ary or ternary expansion set* (later called the middle-third Cantor set), presented as a perfect set that is nowhere dense.

In a letter dated November 1883 Cantor extends this construction of the Cantor set by defining the *Cantor function*, a continuous, non-decreasing map that is constant on the removed intervals and thus furnishes a counter-example to naïve forms of the Fundamental Theorem of Calculus. His 1884 paper *De la puissance des ensembles parfaits de points* in *Acta Mathematica* analyzed the cardinalities of perfect sets, cementing the role of the ternary Cantor set as both a set-theoretic and analytic paradigm. These milestones,

Smith’s overlooked precursor, Cantor’s derived-set framework, the explicit ternary construction, and the analytic Cantor function, together mark the emergence of what is now simply known as the Cantor set (Fleron 1994). Various classes and approaches to the construction of Cantor sets have been extensively studied, each generating distinct subclasses and mathematical structures. As of the date of this survey (31 January 2026), mathematicians have identified numerous specific types of Cantor sets, including Affine, Autonomous, Balanced, Bunched, Central, Cookie-Cutter, Deterministic, Discrete, Distorted, Dusts, Fat, Generalized, Genus- g , Geometric, Hairy, High Dimensional, Homogeneous, Hyperbolic, Iteration Function System (IFS) type, Invariant, Kinetic, Linear, Minimal, Multi-model, Nearly-affine, Non autonomous, Non-homogeneous, Non-archimedean (p -adic), Non-central, Non-Symmetric, Nonstandard, p -Adic Heisenberg, Random, Regular, Scrawny, Self-Similar, Semi-bounded, Smooth, Sticky, Superior, Smith-Volterra-Cantor(SVC) type, Symbolic, Symmetric, Tame, Thick, Thin, Twofold, Ultrametric, Uniform, Universal, Visible, and Wild. Comprehensive references for these subclasses can be found in (Edgar 2007; Vallin 2013; Falconer 2014; Gorodetski and Northrup 2015; Lin *et al.* 2024; McLoughlin and Krizan 2014; Jumha *et al.* 2024; Nassiri and Bidaki 2025). In this paper, however, our primary focus is on deterministic Cantor sets defined explicitly on the real line, and we direct the interested reader to relevant literature for more general or abstract variants.

Manuscript received: 11 March 2026,
Revised: 03 May 2026,
Accepted: 14 May 2026.

¹mohsen.soltanifar@alumni.utoronto.ca (Corresponding author)

Since Georg Cantor's original construction in the 1880s, the Cantor set has become a touchstone throughout mathematics and beyond. On the real line it provides one of the simplest explicit exemplars of a fractal: it is nowhere dense (its closure has empty interior), totally disconnected (its only connected subsets are singletons), perfect (closed with no isolated points), and uncountable, yet its Lebesgue measure can be tuned from zero (the classical middle-third set) to any prescribed value in $(0, 1)$. Such extremal combinations of topological and measure-theoretic properties make Cantor sets a natural testing ground for "pathological" phenomena, most famously the Cantor–Lebesgue function, a continuous, non-decreasing map that is constant on a set of full measure.

In constructive and descriptive settings the Cantor set reappears as a paradigmatic building block: for example, basic Cantor-type function blocks span $5/28$ of the canonical representatives in the function space $F(R, R)$, (Soltanifar 2023) and they appear to prove the existence of aleph-two deterministic as well as aleph-two random fractals on the real line (Soltanifar 2021, 2022). Outside pure mathematics, Cantor-type patterns have been cited in art, early Egyptian column motifs (Frame *et al.* 2025; Kainth 2023), in modeling financial markets (Jin 2021), and in nature, such as ring-like stratification of Saturn when idealized as repeated circular products (Mandelbrot 1982).

Beyond the illustrations just listed, Cantor-type sets recur across mathematics as both structural prototypes and technical tools, and a brief tour of these connections clarifies why a unified descriptive survey is timely. The principal connections between them can be summarized as follows:

- Hyperbolic invariant sets of smooth dynamical systems frequently carry the topology of a Cantor set, and the conjugacy between such systems and full shifts on finite alphabets is the foundational technique of symbolic dynamics (Lind and Marcus 1995). The horseshoe construction of Smale and the wild hyperbolic sets of Newhouse (1979) are paradigmatic examples, and the IFS attractors discussed in Section are their simplest one-dimensional analogues.
- The set of badly approximable real numbers in $[0, 1]$ is, up to measure zero, a countable union of Cantor-type sets indexed by partial-quotient bounds. The arithmetic theory of sums and differences of Cantor sets, initiated by Newhouse and developed extensively by (Palis and Takens 1993; Moreira and Yoccoz 2001), asks when $\Gamma + \Gamma'$ or $\Gamma - \Gamma'$ contains an interval, with applications to homoclinic bifurcations and to the Markov and Lagrange spectra (Moreira *et al.* 2020).
- Self-similar measures supported on Cantor sets serve as test cases and counterexamples in Fourier analysis. Salem sets, compact sets whose Hausdorff and Fourier dimensions coincide, are constructed using randomized Cantor-type schemes (Salem 1951), and Cantor measures appear centrally in modern work on arithmetic progressions in sparse sets (Laba and Pramanik 2009).
- Cantor sets are the canonical examples in the theory of Hausdorff dimension, self-similar measures, and rectifiability (Mattila 1995; Falconer 2014). The Hausdorff-Dimension Theorem (Soltanifar 2006b) is itself most transparently proved by exhibiting an explicit one-parameter family of Cantor sets, and the recent extensions in (Soltanifar 2021, 2022) place such constructions in a broader cardinality-theoretic framework.
- Brouwer's classical theorem characterizes the Cantor space $\{0, 1\}^{\mathbb{N}}$ as the unique (up to homeomorphism) non-empty perfect, compact, totally disconnected, metrizable space (Brouwer 1910; Kechris 1995). Every Cantor set on the real line is there-

fore homeomorphic to every other, which makes the descriptive distinctions surveyed in this paper, arising not from topology but from Lebesgue measure, Hausdorff dimension, and explicit arithmetization, the more striking.

- Random analogues of the constructions in Section give rise to random recursive fractals (Mauldin and Williams 1986; Soltanifar 2022) and to determinantal point processes whose intensity measures are concentrated on generalized Cantor sets (Lin *et al.* 2024). These constructions are the stochastic counterpart of the deterministic survey pursued here.
- In addition to the art-, finance-, and physics-related references cited above, Cantor-type structures have been invoked as descriptive models in signal processing and information theory (Schroeder 2009).

Several of these directions, particularly the arithmetic of intersections in dynamical systems and the harmonic-analytic study of Cantor measures, motivate the open problem (OP5) on common intersections.

Despite this wide relevance, no unified survey has yet cataloged the diverse set-theoretic descriptions of Cantor sets, nor identified which directions are fully developed and which remain open. The present paper closes that gap by reviewing known explicit and implicit set-theoretic formulae on the real line for the deterministic cases, classifying them by structural features, and highlighting avenues for further investigation. While the book-length treatment of Vallin (2013) catalogs many properties and applications of Cantor sets, it does not organize their descriptions into a hierarchical set-theoretic framework, nor does it supply explicit endpoint recursions for positive-measure middle- α families; the present work is, to our knowledge, the first survey to do so.

Beyond consolidating known material, this survey advances the descriptive set-theoretic study of Cantor sets in four concrete directions viewed here:

- We organize the general, nested, iterated-function-system, and q -ary descriptions under a single set-theoretic language, making translation between representations routine.
- Theorem 0.3 supplies a closed recursive formula for the endpoint maps $a_{n,k}(p/q)$, $b_{n,k}(p/q)$ of the augmented thick family $\Gamma_3(p/q, 2)$ with $0 < p/q \leq 1/3$; to the best of our knowledge this is the first such explicit recursion in the literature.
- Corollary 0.4 identifies the classical Cantor set as a member of all three families simultaneously, with equivalence across all four representations made explicit.
- The accompanying R implementation (Appendix B) realizes the recursion in (N2) for exact finite-stage computation, and The present framework leads naturally to five open problems (OP1-OP5), each of which is amenable to computational experimentation.

These contributions distinguish the present work from prior catalogs and position it as a platform for further descriptive set-theoretic research on Cantor-type sets of the real line.

This paper focuses exclusively on deterministic Cantor sets on the real line, examined through a set-theoretic lens, and highlights several descriptive perspectives that remain undiscovered. The remainder of this paper is organized as follows. We first review the necessary background in descriptive set theory, recalling the Borel hierarchy and introducing four nested levels of set-theoretic representation, ranging from the broadest class of Cantor sets to the narrowest class of construction. We then present descriptive formulas for three families of Cantor sets, showing that the middle-third Cantor set belongs to the intersection of all three and thus serves as

a unifying example. Finally, we discuss current challenges, formulate open problems, and outline directions for future research. An appendix collects intermediate background, and interim results, and provides annotated computer code that automates some of the constructions, offering a platform for subsequent mathematical investigations.

SET-THEORETIC FOUNDATIONS

This section deals with the set theoretic foundations for description of Cantor sets. We start with fundamental definitions of F_σ , G_δ and the Borel Hierarchy (Moschovakis 2009; Kechris 1995). Then, we present the proposed set theoretic representations of Cantor sets in hierarchy from macro-level to micro-level in four levels: (i) The general representation, (ii) the nested representation, (iii) the Iterated Function System representation, and, (iv) the q -ary expansion. Recurring specialized terms are collected in Appendix C for the convenience of non-specialists.

The F_σ and G_δ Properties and the Borel Hierarchy

Let $\mathcal{B}([0, 1])$ denote the Borel σ -algebra on the closed unit interval. Set

$$\Sigma_1^0 := \{U \subseteq [0, 1] \mid U \text{ is open}\},$$

$$\Pi_1^0 := \{F \subseteq [0, 1] \mid F^c \in \Sigma_1^0\} (= \text{all closed sets}),$$

and define the *Borel hierarchy* inductively for $n \geq 1$ by:

$$\Sigma_{n+1}^0 := \left\{ \bigcup_{k=1}^{\infty} A_k \mid A_k \in \Pi_n^0 \right\},$$

$$\Pi_{n+1}^0 := \left\{ \bigcap_{k=1}^{\infty} B_k \mid B_k \in \Sigma_n^0 \right\}.$$

The classes Σ_2^0 and Π_2^0 are traditionally denoted:

$$F_\sigma := \Sigma_2^0 \quad (\text{countable unions of closed sets}),$$

$$G_\delta := \Pi_2^0 \quad (\text{countable intersections of open sets}).$$

Thus, every closed subset of $[0, 1]$ is F_σ , and every open subset of $[0, 1]$ is G_δ . Also, the middle-third Cantor set is simultaneously F_σ and G_δ (Kechris 1995).

General Representation

At the broadest level, a Cantor set is identified with whatever closed remainder is left in $[0, 1]$ after excising a countable family of pairwise disjoint open intervals; no further constraint is placed on how those intervals are chosen, distributed, or scaled. This representation therefore captures the largest class of objects that the term ‘‘Cantor set’’ can reasonably denote on the real line. A precise set-theoretic formulation of this viewpoint was given by Astels (1999). To begin with, let $I_0 = [0, 1]$ be the closed unit interval and let $\{I_n^*\}_{n \geq 1}$ be a (finite or countable) family of pairwise-disjoint open sub-intervals contained in I_0 . The general definition of Cantor set Γ is given by:

$$\Gamma = I_0 - \bigcup_{n=1}^{+\infty} I_n^*. \quad (1)$$

Considering closed sets $I_n = I_0 - I_n^*$ ($n \in \mathbb{N}$) we have the following equivalent for the most general representation of Cantor set:

$$\Gamma = \bigcap_{n=0}^{+\infty} I_n. \quad (2)$$

This definition covers many non-fractal subsets of the closed unit interval. Furthermore, by definition Γ is closed subset of $[0, 1]$ and hence a F_σ set. Also, given that for the Euclidean distance d_E we have $\Gamma = \bigcap_{n=1}^{+\infty} \Gamma^{(n)}$ where $\Gamma^{(n)} := \{x \in [0, 1] \mid d_E(x, \Gamma) < \frac{1}{n}\}$ ($n \in \mathbb{N}$) are open sets, it follows that Γ is a G_δ set.

Nested Representation

The nested representation refines the general one by demanding that the construction proceed in stages: at every step, the surviving set is a finite union of closed intervals, each of which is subdivided in the next step, with diameters tending uniformly to zero. This staged structure, absent from the general representation, is what enables stage-by-stage geometric and measure-theoretic analysis. A formal version of this more structured viewpoint was introduced by Falconer (2014) where a nested structure is imposed on their construction process as follows. Let $(s_n)_{n \geq 0}$ be a strictly increasing sequence of natural numbers with $s_0 = 1$ and $s_n \rightarrow \infty$. Construct closed sets $I_n \subset [0, 1]$ for given $n \geq 0$ recursively as follows:

1. Base step. Set $I_0 = [0, 1]$ and denote its single component by $I_{0,1}$.
2. Inductive step. Assume $I_{n-1} = \bigcup_{k=1}^{s_{n-1}} I_{n-1,k}$ is the disjoint union of s_{n-1} closed intervals. Subdivide each $I_{n-1,k}$ into $m_{n,k} \geq 1$ pairwise-disjoint closed sub-intervals. From the entire collection of these sub-intervals choose exactly s_n of them (at least one descendant from every parent), relabel them $I_{n,1}, \dots, I_{n,s_n}$, and set:

$$I_n = \bigcup_{k=1}^{s_n} I_{n,k} \subseteq I_{n-1} \quad (n \in \mathbb{N}). \quad (3)$$

Assume moreover that the intervals shrink, i.e. $\max_{1 \leq k \leq s_n} |I_{n,k}| \downarrow 0$. Then $I_0 \supset I_1 \supset I_2 \supset \dots$, with $I_n = \bigcup_{k=1}^{s_n} I_{n,k}$ and $s_n \uparrow +\infty$. The associated Cantor-set limit set is:

$$\Gamma = \bigcap_{n=0}^{+\infty} I_n. \quad (4)$$

This definition encompasses a broad class of real-line fractals, characterized by their symmetry, Lebesgue measure, and key topological attributes, most notably perfectness and nowhere denseness.

Iterated Function System (IFS) Representation

The IFS representation further specializes the nested construction by requiring that the surviving sub-intervals at every stage arise as images of $[0, 1]$ under a fixed finite family of affine contractions. The resulting Cantor set is then the unique non-empty compact attractor of the system, and its self-similarity becomes a structural identity rather than an asymptotic property. This more rigid formulation is naturally expressed in the language of iterated-function systems, and appears as a special instance of the nested representation in (Falconer 2014). In this case, at each stage of the construction the surviving sub-intervals arise as images of the unit interval under a fixed finite family of affine maps. To see construction, we recall the relevant terminology as follows. First, we consider the following special maps:

1. A map $T : [0, 1] \rightarrow [0, 1]$ is *affine* if $T(tx_1 + (1-t)x_2) = tT(x_1) + (1-t)T(x_2)$ for all $t, x_1, x_2 \in [0, 1]$.
2. It is a *contraction* if $|T(x_1) - T(x_2)| \leq r|x_1 - x_2|$ for some $r \in (0, 1)$ and all $x_1, x_2 \in [0, 1]$.

Second, every affine contraction can be written $T(x) = rx + s$ with $r, s \in [0, 1]$ and $r < 1$ (Edgar 2007); we call $r = \text{cont. coeff}(T)$ its *contraction coefficient*. Third, let $\mathcal{T} = \{T_k\}_{k=1}^K$ ($K \geq 2$) be a finite family of affine contractions with $\max_{1 \leq k \leq K} \text{cont. coeff}(T_k) < 1$. Such a family constitutes an IFS. For a non-empty compact set $I \subset [0, 1]$ define:

$$T(I) = \bigcup_{k=1}^K T_k(I). \quad (5)$$

The operator T is itself a contraction with ratio $r = \max_{1 \leq k \leq K} \text{cont. coeff}(T_k)$ (Falconer 2014).

Hence the sequence $\{T^{\circ n}([0, 1])\}_{n=0}^{+\infty}$, where $T^{\circ n}$ denotes the n -fold composition and $T^{\circ 0} = \text{id}$, converges in the Hausdorff metric to a unique non-empty compact set, called the *attractor* of the system, and we set the Cantor set as:

$$\Gamma = \bigcap_{n=0}^{+\infty} T^{\circ n}([0, 1]). \quad (6)$$

Here, each $T^{\circ n}([0, 1])$ is the disjoint union of K^n closed intervals of length at most r^n , so their diameters tend to 0 as $n \rightarrow +\infty$. Furthermore, the Cantor set satisfies the invariance relation:

$$\Gamma = T(\Gamma) = \bigcup_{k=1}^K T_k(\Gamma). \quad (7)$$

Thus, the affine-IFS description furnishes the more specific and informative representation in the nested hierarchy of set-theoretic representations of Cantor sets.

q-Ary Expansion

The q -ary expansion is the most rigid level of the hierarchy: every affine contraction in the IFS is required to share the common slope $1/q$, so each retained interval at stage n has length exactly q^{-n} and each point of the limit set admits a digit-string address in $\mathcal{A}^{\mathbb{N}}$. This arithmetization is what makes explicit closed-form descriptions possible. The q -Ary expansion of Cantor sets is indeed an special case of their IFS representation when the involved affine maps have equal slopes. In details, let $q \in \mathbb{N} + 2$, and $A \subset \{0, 1, \dots, q-1\}$, $|A| = K$ ($2 \leq K \leq q$). For every digit $a \in A$ define the a^{th} strict contraction T_a with slope $\frac{1}{q}$ and y-intercept $\frac{a}{q}$:

$$\begin{aligned} T_a &: [0, 1] \longrightarrow [0, 1] \\ T_a(x) &= \frac{1}{q}x + \frac{a}{q}. \end{aligned} \quad (8)$$

Now, the IFS representation of the Cantor set Γ is given by:

$$\Gamma = \bigcap_{n=0}^{+\infty} I_n : I_0 = [0, 1], I_n = \bigcup_{a \in A} T_a(I_{n-1}), \quad (n \in \mathbb{N}). \quad (9)$$

This special case has the following properties:

1. Each component of I_n is a closed interval of length q^{-n} .
2. There are K^n such components with one for each finite digit block $\vec{a} = (a_1, \dots, a_n) : a_k \in A$ ($1 \leq k \leq n$).
3. The system is nested: $I_n \supset I_{n-1}$ ($n \in \mathbb{N}$).

A straightforward verification by mathematical induction shows that:

$$\begin{aligned} I_n &= T_{a_1} \circ \dots \circ T_{a_n}([0, 1]) \\ &= \left[\sum_{k=1}^n \frac{a_k}{q^k}, \sum_{k=1}^n \frac{a_k}{q^k} + \frac{1}{q^n} \right], \quad (n \in \mathbb{N}). \end{aligned} \quad (10)$$

Also, the map $F_n = T_{a_1} \circ \dots \circ T_{a_n}$ is q^{-n} -Lipschitz continuous ($n \in \mathbb{N}$) (Searcoid 2006). Consequently, $x \in \Gamma$ if and only if there is a unique (up to the standard endpoint ambiguity) infinite digit string $\vec{a}^* = (a_1, a_2, a_3, \dots) \in A^{\mathbb{N}}$ such that $x = \sum_{n=1}^{+\infty} \frac{a_n}{q^n}$. Hence, we have the q -ary expansion of the Cantor set:

$$\Gamma = \{x \in [0, 1] | x = \sum_{n=1}^{+\infty} \frac{a_n}{q^n}, a_n \in A\}. \quad (11)$$

Notation: Henceforth, we denote $\Gamma(\alpha, K)$ middle α -Cantor set defined by K affine maps.

In the upcoming subsections, we study, explicit and recursive representation of three families of Cantor sets given their associated q -ary expansions.

Hierarchical Relationships and Practical Differences

The four representations introduced in the previous four sections form a strict descending chain of structural specificity:

$$\text{General} \supsetneq \text{Nested} \supsetneq \text{IFS} \supsetneq q\text{-ary}. \quad (12)$$

Each inclusion is proper, and at each step a concrete structural commitment is added to the preceding level:

- *General* $\rightarrow$ *Nested*. The arbitrary countable family of deleted open intervals is reorganized into a stagewise construction in which the surviving set at each stage is a finite union of closed intervals with diameters tending to zero. The strictness of the inclusion follows from the existence of closed nowhere-dense subsets of $[0, 1]$ that admit a representation of the form (1) but not a stagewise nested decomposition with $s_n \uparrow \infty$ (see (Astels 1999)).
- *Nested* $\rightarrow$ *IFS*. The stagewise construction is required to be generated by a fixed finite family of affine contractions, so that the limit set satisfies the self-similarity identity $\Gamma = \bigcup_{k=1}^K T_k(\Gamma)$. Not every nested Cantor set is an IFS attractor: nested constructions in which the contraction ratios vary across stages or across components fail this identity (cf. (Hutchinson 1981; Falconer 2014)).
- *IFS* $\rightarrow$ *q-ary*. All affine contractions are forced to share the common slope $1/q$ for some integer $q \geq 2$. IFS attractors with unequal contraction ratios—e.g. the standard examples with $r_1 \neq r_2$ used to realize prescribed Hausdorff dimensions—are excluded.

The price of descending the hierarchy is generality; the reward is descriptive power. Table 1 summarizes the practical trade-offs.

The three families treated in Section illustrate the lower end of this hierarchy concretely: Γ_1 and Γ_2 are presented through their q -ary expansions and admit closed-form gap descriptions, while $\Gamma_3(p/q, 2)$, although IFS-generated, requires the recursive endpoint treatment of Theorem 0.3 because its retained-interval lengths do not collapse to a single power of a fixed base.

■ **Table 1** Practical comparison of the four set-theoretic representations of Cantor Sets on the real line

#	Item	General	Nested	IFS	q -ary
1	Operative Parameter	Family $\{I_n^s\}$ of deleted open intervals	Sequence (s_n) and per-stage subdivisions	Affine maps $\{T_k\}_{k=1}^K$	Base q and digit set $\mathcal{A} \subseteq \{0, \dots, q-1\}$
2	Stage- n Structure	Not stage-defined in general	s_n closed intervals of varying length	K^n closed intervals, length $\leq r^n$	K^n closed intervals of length exactly q^{-n}
3	Self-Similarity	None required	None required	Exact: $\Gamma = \bigcup_k T_k(\Gamma)$	Exact, with uniform ratio $1/q$
4	Closed-form Points	Not available in general	Not available in general	Available via fixed-point address	Available via digit expansion $\sum a_n q^{-n}$
5	Typical Use	Existence and Borel-class arguments	Stagewise measure / dimension analysis	Self-similarity, dimension theorems	Explicit formulas, arithmetic structure

DESCRIPTIVE FORMULAS FOR CANTOR SETS

Thin Family 1

The earliest set-theoretic explicit description of Cantor sets appears in (Soltanifar 2006b), where the author treats the symmetric thin family generated by at least two affine maps in an iterated-function system. This construction yields a fully constructive proof of the Hausdorff-Dimension Theorem (which asserts that for every $s > 0$ there exists a continuum of compact subsets of $\mathbb{R}^n$ ($n \geq \lceil s \rceil$) with Hausdorff dimension of exactly s), thereby establishing the existence of the corresponding fractals for any prescribed Hausdorff dimension. Subsequent work extends the argument to broader deterministic and stochastic settings (Soltanifar 2021, 2022).

Theorem 0.1 Let $\Gamma_1(\frac{1}{q}, \frac{q+1}{2}) := \{x \in [0, 1] | x = \sum_{n=1}^{+\infty} \frac{a_n}{q^n}, a_n = 0, 2, 4, \dots, q-1\}$ where $q \in 2\mathbb{N} + 1$. Then:

$$\Gamma_1(\frac{1}{q}, \frac{q+1}{2}) = [0, 1] - \bigcup_{n=1}^{+\infty} \bigcup_{k=0}^{q^{n-1}-1} \bigcup_{r=1}^{\frac{q-1}{2}} (\frac{qk + (2r-1)}{q^n}, \frac{qk + 2r}{q^n}) : (q \in 2\mathbb{N} + 1). \quad (13)$$

In particular, for the case $q = 3$ we obtain the explicit formulae for the middle-third Cantor set $C = \Gamma_1(\frac{1}{3}, 2)$ as in the second formulae in Corollary 0.4.

Thin Family 2

In the same year, the explicit set-theoretic description of a second family of Cantor sets was presented (Soltanifar 2006a). As with the earlier construction, the sets obtained are thin and symmetric; however, their underlying iterated-function system involves only two affine contractions, in contrast to generally larger number used for the first family.

Theorem 0.2 Let $\Gamma_2(\frac{q-2}{q}, 2) := \{x \in [0, 1] | x = \sum_{n=1}^{+\infty} \frac{a_n}{q^n}, a_n = 0, q-1\}$ where $q \in \mathbb{N} + 2$. Then:

$$\Gamma_2(\frac{q-2}{q}, 2) = [0, 1] - \bigcup_{n=1}^{+\infty} \bigcup_{k=0}^{q^{n-1}-1} (\frac{qk+1}{q^n}, \frac{qk+2}{q^n}) : (q \in \mathbb{N} + 2). \quad (14)$$

Similar to the Thin Family Γ_1 , for the case $q = 3$ we obtain the explicit formulae for the middle-third Cantor set $C = \Gamma_2(\frac{1}{3}, 2)$ as in the second formulae in Corollary 0.4.

Augmented Thick Family 1

In both preceding cases we provided closed-form descriptions for two families of measure-zero Cantor sets. This naturally raises the question: Does an analogous explicit representation exist for Cantor sets of positive Lebesgue measure? Current evidence suggests a negative answer; no such closed formula is known at the time of writing. Nevertheless, a constant-recursive representation remains feasible (Elaydi 2005), and the requisite background together with interim results is presented in Appendix A for interested readers.

Theorem 0.3 Let $\Gamma_3(\frac{p}{q}, 2) := \{x \in [0, 1] | x = \sum_{n=1}^{+\infty} \frac{a_n}{(2q)^n}, a_n = 0, 1, \dots, q-p-1, q+p, \dots, 2q-1\}$ where $p, q \in \mathbb{N}, (p, q) = 1, 0 < \frac{p}{q} \leq \frac{1}{3}$. Then:

$$\begin{aligned} \Gamma_3(\frac{p}{q}, 2) &= \bigcap_{n=0}^{+\infty} \bigcup_{k=1}^{2^n} [a_{n,k}(\frac{p}{q}), b_{n,k}(\frac{p}{q})] : \\ a_{0,1}(\frac{p}{q}) &= 0; \\ a_{n,k}(\frac{p}{q}) &= 1_{\text{odd}}(k) \left(a_{n-1, \frac{k+1}{2}}(\frac{p}{q}) \right) \\ &\quad + 1_{\text{even}}(k) \left(a_{n-1, \frac{k}{2}}(\frac{p}{q}) + \frac{(1-3\frac{p}{q}) + (1-\frac{p}{q})(2\frac{p}{q})^n}{(1-2\frac{p}{q})2^n} \right); \\ b_{0,1}(\frac{p}{q}) &= 1; \\ b_{n,k}(\frac{p}{q}) &= 1_{\text{even}}(k) \left(b_{n-1, \frac{k}{2}}(\frac{p}{q}) \right) \\ &\quad + 1_{\text{odd}}(k) \left(b_{n-1, \frac{k+1}{2}}(\frac{p}{q}) - \frac{(1-3\frac{p}{q}) + (1-\frac{p}{q})(2\frac{p}{q})^n}{(1-2\frac{p}{q})2^n} \right) : \\ 1 \leq k \leq 2^n, n \in \mathbb{N}. \end{aligned} \quad (15)$$

The Computer software code to calculate the end points $a_{n,k}(\cdot), b_{n,k}(\cdot)$ are given in Appendix B with associated examples

for $\alpha_1 = \frac{1}{3}$, and $\alpha_2 = \frac{1}{4}$ (R Core Team 2022). It is noteworthy to mention that the middle-third Cantor set belongs to all three discussed families. Figure 1 summarizes our above discussions in terms of hierarchy of descriptive representations.

Finally, for the case $p = 1, q = 3$ we obtain the first order recursive formulae for the middle-third Cantor set $C = \Gamma_3(\frac{1}{3}, 2)$ as in the third formulae in the following Corollary 0.4:

Corollary 0.4 *The middle-third Cantor set has three equivalent set theoretic descriptive representations:*

$$C = \{x \in [0, 1] | x = \sum_{n=1}^{+\infty} \frac{a_n}{3^n}, a_n = 0, 2\} \quad (16)$$

$$= [0, 1] - \bigcup_{n=1}^{+\infty} \bigcup_{k=0}^{3^{n-1}-1} \left(\frac{3k+1}{3^n}, \frac{3k+2}{3^n} \right) \quad (17)$$

$$= \bigcap_{n=0}^{+\infty} \bigcup_{k=1}^{2^n} [a_{n,k}, b_{n,k}] : \quad (18)$$

$$a_{0,1} = 0; a_{n,k} = 1_{\text{odd}}(k) \left(a_{n-1, \frac{k+1}{2}} \right) + 1_{\text{even}}(k) \left(a_{n-1, \frac{k}{2}} + \frac{2}{3^n} \right);$$

$$b_{0,1} = 1; b_{n,k} = 1_{\text{odd}}(k) \left(b_{n-1, \frac{k+1}{2}} - \frac{2}{3^n} \right) + 1_{\text{even}}(k) \left(b_{n-1, \frac{k}{2}} \right);$$

$$1 \leq k \leq 2^n, n \in \mathbb{N}.$$

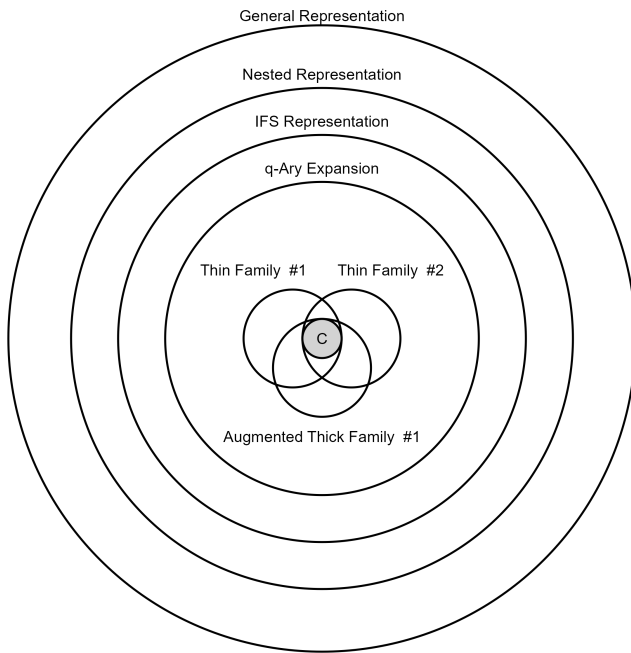


Figure 1 Diagram of descriptive formulas for the Cantor sets from set-theoretic perspective: C refers to the middle-third Cantor set

DISCUSSION

The four descriptions surveyed here, general, nested, IFS, and q -ary, form a strict hierarchy in which each level imposes additional structure on its predecessor, and the three families $\Gamma_1, \Gamma_2, \Gamma_3$ populate this hierarchy at successively finer levels of specificity. The recursive endpoint formula for $\Gamma_3(p/q, 2)$ closes a gap in the positive-measure case: where the thin families admit closed-form

gap descriptions, the augmented thick family had previously resisted such an explicit treatment, and the recursion presented in Theorem 0.3 now makes every retained interval at every finite stage exactly computable. The middle-third Cantor set, sitting in the intersection of all three families, illustrates how a single canonical object can serve as a stress test for the entire descriptive framework. The remainder of this section explores three consequences of this synthesis: a continuum-theoretic reading of Γ_3 via embeddings into standard continua, a measure-theoretic construction of canonical self-similar measures supported on $\Gamma_3(\alpha)$, and an outlook tying the framework to the open problems of Section ??.

Continuum-theoretic perspective: Although Cantor sets are totally disconnected, they occur naturally inside many *continua* (compact connected metric spaces) as closed, perfect, nowhere-dense subsets (Nadler 1992). The explicit endpoint recursion for Γ_3 supplies fine control of gap geometry at every stage. This facilitates: (i) explicit embeddings of $\Gamma_3(\alpha)$ into standard continua (e.g., as closed subsets of *arcs* (homeomorphic images of $[0, 1]$), *dendrites* (locally connected continua containing no simple closed curve), or *Peano continua* (locally connected continua, equivalently continuous images of $[0, 1]$)) with prescribed gap patterns; (ii) constructive “fill-in” procedures that yield continua whose complements are modeled by a chosen $\Gamma_3(\alpha)$; and (iii) numerical investigation of intersection phenomena (useful in studying when two embedded Cantor-type sets intersect), where computable gap data and thickness proxies are central. In this sense the paper’s formulas act as a bridge, turning qualitative continuum-theory questions into explicit, checkable computations.

Measure-theoretic perspective (a canonical Γ_3 measure): For each $\Gamma_3(\alpha)$ we can *canonically* equip the set with a Borel probability measure that generalizes the Cantor–Lebesgue measure (Rosenthal 2006). At stage n there are q^n retained intervals; assign either uniform weights q^{-n} or, more generally, a fixed weight vector $w = (w_1, \dots, w_q)$ with $\sum_{i=1}^q w_i = 1$, propagated multiplicatively along the recursion. This produces a self-similar probability measure $\mu_{\alpha, w}$ supported on $\Gamma_3(\alpha)$ whose cumulative distribution function is a Devil’s-staircase generalization of the Cantor function: it is continuous, nondecreasing, constant on deleted gaps, and increases exactly on $\Gamma_3(\alpha)$. The explicit endpoints $(a_{n,k}, b_{n,k})$ allow fast, exact evaluation of cylinder-set masses and numerical plots of the associated distribution function. This construction is immediate from the framework developed here and is implemented by minor extensions of Appendix B.

Because the geometry of Γ_3 is explicit at every finite stage, the framework supports systematic computation of gap-length statistics, empirical thickness proxies, and intersection behavior under translation, tools that are relevant both to continuum-theoretic embeddings and to questions in additive/fractal geometry. We conclude with open problems on asymmetry, irrational endpoints, variable-gap deletions, the role of symmetry, and common intersections; these problems are now posed in a setting where algorithmic experiments and exact finite-stage calculations are straightforward.

The preceding survey has shown that three natural families of deterministic Cantor sets, two of measure zero with explicit descriptions and one of non-negative measure defined recursively, share a common symmetric architecture and possess rational construction end-points; indeed, the classical middle-third Cantor set lies in the intersection of all three. These parallel features highlight several directions in which the present set-theoretic framework remains incomplete and invite further research. We organize the open problems into three thematic clusters reflecting their logical

interdependence: *structural generalization* (Theme A), *variable-scale constructions* (Theme B), and *interaction phenomena* (Theme C), with the last depending on progress in the first two.

Theme A (Structural Generalization)

The descriptive formulas of Section 3 rely on two structural restrictions—bilateral symmetry of the deletion pattern and rationality of the retained-interval endpoints. The following two problems ask whether either restriction can be dropped without losing explicitness.

- OP1: *Asymmetric constructions.* Devise explicit or recursive set-theoretic formulas for Cantor sets generated by iterated-function systems with $K \geq 2$ maps that do not preserve bilateral symmetry.
- OP2: *Irrational endpoints.* Find descriptive formulas when the retained intervals have irrational endpoints, again allowing $K \geq 2$ IFS components.

Rationale and impact: Symmetric, rationally-located Cantor sets form a measure-zero subfamily within the space of all IFS attractors, yet they are essentially the only ones for which closed or recursive descriptive formulas are currently known. A resolution of OP1 would extend the techniques of Section 3 to the asymmetric self-similar sets that arise as hyperbolic invariant sets in low-dimensional dynamics (Palis and Takens 1993; Moreira and Yoccoz 2001), where asymmetry is the rule rather than the exception. A resolution of OP2 would connect the descriptive framework to the algebraic theory of self-similar measures with non-rational contraction ratios, where deep recent work by Hochman (Hochman 2014) and Shmerkin (Shmerkin 2019) has clarified the dimension theory but explicit set-theoretic formulas remain absent.

Tools and feasibility: The endpoint recursion of Theorem 0.3 is candidate machinery for OP1: removing the symmetry assumption requires tracking left- and right-endpoints under independent contraction ratios $r_L \neq r_R$, which extends the recursion in Lemmas 4.1–4.3 (Appendix A) at the cost of a more elaborate index scheme. OP2 is genuinely harder, since the arithmetic identity $p/q = 2p/(2q)$ that drives the $2q$ -ary expansion in Appendix A has no analogue for irrational ratios; symbolic-dynamic coordinates (Lind and Marcus 1995) or continued-fraction expansions of the endpoints offer alternative descriptive routes that have not yet been systematically explored.

Theme B (Variable-Scale Constructions)

The middle- α family $\Gamma_3(\alpha, 2)$ removes a gap of length α^n at stage n , so the deletion sequence is geometric. The following two problems concern non-geometric deletion schemes and the role of symmetry within them.

- OP3: *Variable-gap deletions.* Given a sequence $(\alpha_n)_{n \geq 1}$ with $0 < \alpha_n < 1$ satisfying $1 - \sum_{n=1}^{\infty} 2^{n-1} \alpha_n \geq 0$, provide a recursive descriptive formula for the Cantor set obtained by removing at stage n the open middle interval of length α_n .
- OP4: *Role of symmetry.* Determine how the solution of OP3 varies when bilateral symmetry of the deletion is enforced or relaxed.

Rationale and impact: Variable-scale constructions interpolate between the self-similar Cantor sets of Section 3 and the broader nested representations of (4). Smith–Volterra–Cantor sets and their generalizations belong to this class (Vallin 2013), and their Lebesgue measure can be tuned arbitrarily within $[0, 1]$ by choice of (α_n) . A resolution of OP3 would yield exact endpoint formulas

for these positive-measure constructions, parallel to what Theorem 0.3 achieves in the self-similar case, and would supply concrete examples for testing fractal-uncertainty principles in the variable-scale setting (Laba and Pramanik 2009). OP4 is significant because asymmetric variable-gap constructions exhibit qualitatively different harmonic-analytic behavior—in particular, the Fourier decay of the associated self-similar measures depends sensitively on whether deletions are symmetric.

Tools and feasibility: The recursion of Lemma 4.3 generalizes directly to OP3 by replacing α^n with α_n in the increment $\Delta_{n,k}$; the closed form of $\Delta_{n,k}$ no longer collapses but remains explicit as a finite sum at each stage. OP4 admits a tractable first case (left/right asymmetry only, with all other parameters symmetric), which would be a natural starting point.

Theme C (Interaction Phenomena)

Themes A and B describe families of Cantor sets indexed by structural parameters (asymmetry, irrationality, deletion sequence). The next problem asks how families from these themes interact.

- OP5: *Common ground.* Establish conditions under which the families arising in OP1–OP4 admit a non-empty mutual intersection, and characterize the intersection when it exists.

Rationale and impact: The middle-third Cantor set is, in the present survey, the unique explicit point of intersection across the three families $\Gamma_1, \Gamma_2, \Gamma_3$. Whether analogous canonical objects exist across the broader families of OP1–OP4 is a question of both descriptive and dynamical interest. Common intersections are the deterministic shadow of the Newhouse phenomenon and the Palis conjecture (Newhouse 1979; Palis and Takens 1993; Moreira and Yoccoz 2001): when do two Cantor sets meet, and when does the meeting persist under perturbation? A resolution of OP5 would supply concrete test cases for these conjectures and would bridge the explicit set-theoretic framework of this survey to the geometric measure-theoretic literature on intersections (Mattila 1995).

Tools and feasibility: The thickness theory of Astels (Astels 1999, 2000) provides sufficient conditions for non-empty intersection of two Cantor sets in terms of their gap structures, and the explicit endpoint data supplied by Theorem 0.3 (and its Theme-A and Theme-B extensions) make these thickness computations exact rather than asymptotic. Moreira–Yoccoz (Moreira and Yoccoz 2001) provides the deeper machinery for *stable* intersections under perturbation, which becomes relevant once OP1–OP4 have been resolved.

Resolving these questions would extend the taxonomy of Cantor sets beyond the symmetric, rational, geometrically-scaled cases treated here and clarify how set-theoretic representations interact with geometric asymmetry, irrational numbers, variable deletion schemes, and the intersection phenomena that link them. The thematic grouping above is intended to make explicit the logical dependence, Theme C builds on Themes A and B, and to facilitate parallel progress on the structural and variable-scale fronts.

CONCLUSION

This survey brings four set-theoretic perspectives, general, nested, IFS, and q -ary, under one framework and demonstrates how they meet in three representative Cantor-set families containing the middle-third classic. The synthesis clarifies which descriptive formulas are now complete and which remain missing, especially for asymmetric, irrational-endpoint, and variable-gap constructions.

By distilling these gaps into a concise list of open problems, the paper sets a clear agenda for future work in descriptive set theory and fractal analysis on the real line.

Abbreviations

The following abbreviations are used in this manuscript: ACDC: Applications and Cross-Disciplinary Connections; IFS: Iterated Function System; NC: Novelty and Contribution; OP: Open Problem; SVC: Smith–Volterra–Cantor.

Acknowledgments

The author is grateful to the reviewers for their constructive comments, which significantly improved the presentation of this manuscript.

Ethical standard

The author has no relevant financial or non-financial interests to disclose.

Availability of data and material

Not applicable.

Conflicts of interest

The author declares that there is no conflict of interest regarding the publication of this paper.

Funding

This research received no external funding.

APPENDIX

Appendix A: Interim Results

Let $\Gamma_3(\alpha, 2)$ where $0 < \alpha \leq \frac{1}{3}$ be the middle $-\alpha$ Cantor set defined by two affine maps. Here, its nested representation is given by $\Gamma_3(\alpha, 2) = \bigcap_{n=0}^{+\infty} I_n(\alpha)$ where $I_0(\alpha) = [0, 1]$ and $I_n(\alpha)$ is defined recursively by removing 2^{n-1} middle open intervals each of the length α^n from $I_{n-1}(\alpha)$ ($n \in \mathbb{N}$). We have the following interim results:

- 1. Family Name** By construction, the Lebesgue measure is given by $\lambda(\Gamma_3(\alpha, 2)) = 1 - \sum_{n=1}^{+\infty} 2^{n-1} \alpha^n = \frac{1-3\alpha}{1-2\alpha}$ ($0 < \alpha \leq \frac{1}{3}$). Hence, we augment the "Thick Family" with $0 < \alpha < \frac{1}{3}$ with the "Thin case" $\alpha = \frac{1}{3}$; and, call it the "Augmented Thick Family" with $0 < \alpha \leq \frac{1}{3}$.
- 2. 2q-Ary Expansion** When $0 < \alpha = \frac{p}{q} \leq \frac{1}{3}$, $p, q \in \mathbb{N}$, $(p, q) = 1$, by above construction, the equality $\frac{p}{q} = \frac{2^p}{2^q}$, and mathematical induction, a straightforward verification shows the following 2q-Ary expansion:

$$\Gamma_3(\alpha, 2) = \{x \in [0, 1] | x = \sum_{n=1}^{+\infty} \frac{a_n}{(2q)^n}, \\ a_n = 0, 1, \dots, q-p-1, q+p, \dots, 2q-1\}.$$

This representation facilitates the n-Ary expansion for cases with even q . Furthermore, the middle-third Cantor set has the secondary 6-Ary expansion as $C = \{x \in [0, 1] | x = \sum_{n=1}^{+\infty} \frac{a_n}{6^n}, a_n = 0, 1, 4, 5\}$.

- 3. Proof of Theorem 0.3** We prove the Theorem in steps (a)–(e).

- (a) Notation** Given above recursive construction it follows that, $\Gamma_3(\alpha, 2) = \bigcap_{n=0}^{+\infty} I_n(\alpha)$ where $I_n(\alpha) = \bigcup_{k=1}^{2^n} I_{n,k}(\alpha)$ and $I_{n,k}(\alpha) = [a_{n,k}(\alpha), b_{n,k}(\alpha)]$, ($1 \leq k \leq 2^n, n \in \mathbb{N}_0$).

- (b) Lemma 0.5** $\lambda(I_{n,k}(\alpha)) = \frac{(2\alpha)^{n+1} + 2(1-3\alpha)}{2^{n+1}(1-2\alpha)} : (1 \leq k \leq 2^n, n \in \mathbb{N}_0)$.

Proof. The case for $n = 0$ is trivial. Let $n \in \mathbb{N}$ and $J_{n,k}(\alpha)$ be the middle open interval of length α^n removed from $I_{n-1,k}(\alpha)$ in the nested construction process i.e. $I_{n-1,k}(\alpha) = I_{n,2k-1}(\alpha) \cup J_{n,k}(\alpha) \cup I_{n,2k}(\alpha)$. Then, given, $\lambda(I_{n,2k-1}(\alpha)) = \lambda(I_{n,2k}(\alpha))$ we have: $\lambda(I_{n-1,k}(\alpha)) = 2\lambda(I_{n,k}(\alpha)) + \alpha^n$. Consequently, we have a first order recursive relation:

$$\lambda(I_{n,k}(\alpha)) = \frac{1}{2}\lambda(I_{n-1,k}(\alpha)) - \frac{\alpha^n}{2}, \lambda(I_{0,1}(\alpha)) = 1.$$

Finally, the assertion follows by mathematical induction on $n \in \mathbb{N}$.

- (c) Lemma 0.6** Let $\Delta_{n,k}(\alpha) = \lambda(I_{n,k}(\alpha)) + \alpha^n$ ($1 \leq k \leq 2^n, n \in \mathbb{N}_0$). Then,

$$\Delta_{n,k}(\alpha) = \frac{(1-3\alpha) + (1-\alpha)(2\alpha)^n}{(1-2\alpha)2^n} \quad (1 \leq k \leq 2^n, n \in \mathbb{N}_0).$$

Proof. This is straightforward result from Lemma 0.5.

- (d) Lemma 0.7** Let $a_{n,k}(\alpha), b_{n,k}(\alpha)$, ($1 \leq k \leq 2^n, n \in \mathbb{N}_0$) be defined as in above recursive construction. Then:

$$\begin{aligned} a_{0,1}(\alpha) &= 0, \\ a_{n,k}(\alpha) &= 1_{\text{odd}}(k) \left(a_{n-1, \frac{k+1}{2}}(\alpha) \right) \\ &\quad + 1_{\text{even}}(k) \left(a_{n-1, \frac{k}{2}}(\alpha) + \Delta_{n,k}(\alpha) \right), \\ b_{0,1}(\alpha) &= 1, \\ b_{n,k}(\alpha) &= 1_{\text{even}}(k) \left(b_{n-1, \frac{k}{2}}(\alpha) \right), \\ &\quad + 1_{\text{odd}}(k) \left(b_{n-1, \frac{k+1}{2}}(\alpha) - \Delta_{n,k}(\alpha) \right) \\ &\quad (1 \leq k \leq 2^n, n \in \mathbb{N}_0). \end{aligned}$$

Proof. This is straightforward result from definition of $\Delta_{n,k}(\alpha)$ in Lemma 0.6, and comparing the end points of closed intervals in the following equations:

$$\begin{aligned} [a_{n-1,k}(\alpha), b_{n-1,k}(\alpha)] &= I_{n-1,k}(\alpha) \\ &= I_{n,2k-1}(\alpha) \cup J_{n,k}(\alpha) \cup I_{n,2k}(\alpha) \\ &= [a_{n,2k-1}(\alpha), b_{n,2k-1}(\alpha)] \\ &\quad \cup (b_{n,2k-1}(\alpha), a_{n,2k}(\alpha)) \\ &\quad \cup [a_{n,2k}(\alpha), b_{n,2k}(\alpha)], \\ &\quad (1 \leq k \leq 2^n, n \in \mathbb{N}_0). \end{aligned}$$

- (e) Completion** Finally, by considering $0 < \alpha = \frac{p}{q} \leq \frac{1}{3}$ in Lemma 0.6 and substituting corresponding $\Delta_{n,k}$ in the equations presented in Lemma 0.7, the proof for the Theorem 0.3 is completed.

Appendix B: R Software Code

The following function in R software code produces the end points of the intervals of the recursive representation of "Augmented Thick Family" $\Gamma_3(\frac{p}{q}, 2)$ for given pair $(n, k) : 1 \leq k \leq 2^n, n \in \mathbb{N}$. It presents two examples of calculations one for the thin Cantor set and another for the thick Cantor set:

```
## =====
## Recursive evaluator for lower endpoint a_{n,k}(p/q)
## =====
a_nk <- local({
  ## simple memorization (saves repeated work)
  cache <- new.env(parent = emptyenv())
  f <- function(n, k, p, q) {
    if (n == 0L) { # base level
      if (k != 1L) stop("For n = 0 you must take k = 1")
      return(0)
    }
    if (k < 1 || k > 2^n)
      stop("k must satisfy 1 <= k <= 2^n")
    ## build a unique key for memorization
    key <- paste(n, k, p, q, sep = "-")
    if (exists(key, envir = cache, inherits = FALSE))
      return(cache[[key]])
    r <- p / q # the ratio p/q
    ## recursive step
    if (k %% 2L == 1L) { # k is odd
      val <- f(n - 1L, (k + 1L) / 2L, p, q)
    } else { # k is even
      prev <- f(n - 1L, k / 2L, p, q)
      term <- ((1 - 3 * r) +
        (1 - r) * (2 * r)^n) / ((1 - 2 * r) * 2^n)
      val <- prev + term
    }
    cache[[key]] <- val # store and return val
  }
  f
})

## =====
## Recursive evaluator for upper endpoint b_{n,k}(p/q)
## =====
b_nk <- local({
  cache <- new.env(parent = emptyenv()) #memorization
  f <- function(n, k, p, q) {
    if (n == 0L) { # base case
      if (k != 1L) stop("For n = 0 you must take k = 1")
      return(1)
    }
    if (k < 1L || k > 2L^n)
      stop("k must satisfy 1 <= k <= 2^n")
    key <- paste(n, k, p, q, sep = "-")
    if (exists(key, envir = cache, inherits = FALSE))
      return(cache[[key]])
    r <- p / q
    term <- ((1 - 3 * r) +
      (1 - r) * (2 * r)^n) / ((1 - 2 * r) * 2^n)
    ## recursive step
    if (k %% 2L) { # k is odd
      val <- f(n - 1L, (k + 1L) %% 2L, p, q) - term
    } else { # k is even
      val <- f(n - 1L, k %% 2L, p, q)
    }
    cache[[key]] <- val
    val
  }
  f
})
```

```
})

## =====
## Intervals Endpoints Calculator
## =====
EndpointCalculator <- function(n, p, q) {
  ## ----- base-level endpoints -----
  a0 <- a_nk(0L, 1L, p, q)
  b0 <- b_nk(0L, 1L, p, q)
  ## keep the original "too-large" denominator
  denom <- q^(n + 1L)
  ## ----- level-n endpoints -----
  res1 <- sapply(1:2^n, function(k) a_nk(n, k, p, q))
  res2 <- sapply(1:2^n, function(k) b_nk(n, k, p, q))
  num1 <- c(a0 * denom, round(res1 * denom))
  num2 <- c(b0 * denom, round(res2 * denom))
  ## ----- LABELS (divide out one factor of q) -----
  display_denom <- denom / q # q^(n+1) / q -> q^n
  col_labs <- paste0("k=", 0:2^n)
  fractions1 <- paste0(num1 / q, "/", display_denom)
  fractions2 <- paste0(num2 / q, "/", display_denom)
  ## ----- NEW ROW -----
  combo <- paste0("[", fractions1, ",", fractions2, "]")
  out1 <- rbind("a_{n,k}(\alpha)" = fractions1,
    "b_{n,k}(\alpha)" = fractions2,
    "[a_{n,k}(\alpha), b_{n,k}(\alpha)]" = combo)
  colnames(out1) <- col_labs
  out2 <- t(out1)
  return(out2)
}

## =====
## Examples
## =====
## Example (1) : Thin Cantor Set
EndpointCalculator(n=2, p=1, q=3)
# a_{n,k}(\alpha) b_{n,k}(\alpha) [a_{n,k}(\alpha), b_{n,k}(\alpha)]
# k=0 "0/9" "9/9" "[0/9,9/9]"
# k=1 "0/9" "1/9" "[0/9,1/9]"
# k=2 "2/9" "3/9" "[2/9,3/9]"
# k=3 "6/9" "7/9" "[6/9,7/9]"
# k=4 "8/9" "9/9" "[8/9,9/9]"

## Example (2) : Thick Cantor Set
EndpointCalculator(n=2, p=1, q=4)
# a_{n,k}(\alpha) b_{n,k}(\alpha) [a_{n,k}(\alpha), b_{n,k}(\alpha)]
# k=0 "0/16" "16/16" "[0/16,16/16]"
# k=1 "0/16" "2.5/16" "[0/16,2.5/16]"
# k=2 "3.5/16" "6/16" "[3.5/16,6/16]"
# k=3 "10/16" "12.5/16" "[10/16,12.5/16]"
# k=4 "13.5/16" "16/16" "[13.5/16,16/16]"
```

Appendix C: Glossary

This appendix includes notational conventions and key terms used throughout the survey for the convenience of readers outside descriptive set theory and fractal geometry. Specialists may skip this subsection without loss.

The terms in the paper that a general mathematical reader (e.g., an analyst or topologist) may pause fall into four clusters:

1. **Topological / set-theoretic descriptors:** *Perfect set, nowhere dense, totally disconnected, derived set, F_σ , G_δ , Borel hierarchy.*

2. **Measure-theoretic descriptors:** *Lebesgue measure, outer content, measure-zero vs. positive measure, self-similar measure, Cantor-Lebesgue function, Devil's staircase.*
3. **Fractal / dynamical terminology:** *Hausdorff dimension, attractor, IFS, contraction coefficient, affine map, thickness.*
4. **Continuum-theoretic terms:** *Continuum, arc, dendrite, Peano continuum, embedding*

The concise mathematical definitions and descriptions are as in follows:

- *Perfect set* — a closed set in which every point is a limit point of the set; equivalently, a closed set with no isolated points (Munkres 2000).
- *Nowhere dense set* — a set whose closure has empty interior; informally, a set that is “full of holes” in every sub-interval.
- *Totally disconnected set* — a set whose only connected subsets are singletons; equivalently, between any two distinct points one can interpose a gap.
- F_σ set — a countable union of closed sets.
- G_δ set — a countable intersection of open sets.
- *Borel hierarchy* — the transfinite stratification Σ_n^0, Π_n^0 of Borel sets by alternating countable unions and intersections, beginning with open and closed sets (Kechris 1995).
- *Affine map on $[0, 1]$* — a function of the form $T(x) = rx + s$; a contraction if $|r| < 1$, with $r = \text{cont. coeff}(T)$ its contraction coefficient.
- *Iterated function system (IFS)* — a finite family $\{T_k\}_{k=1}^K$ of contractions; its attractor is the unique non-empty compact set Γ satisfying $\Gamma = \bigcup_{k=1}^K T_k(\Gamma)$ (Falconer 2014).
- *Hausdorff dimension* — the critical exponent $\dim_H(E)$ at which the s -dimensional Hausdorff measure of E transitions from $+\infty$ to 0; for self-similar Cantor sets generated by K maps with common ratio r , $\dim_H(E) = \log K / \log(1/r)$.
- *Self-similar measure* — a Borel probability measure μ on the attractor of an IFS $\{T_k\}$ satisfying $\mu = \sum_{k=1}^K w_k (\mu \circ T_k^{-1})$ for a probability weight vector $(w_1, \dots, w_K)$.
- *Cantor–Lebesgue function (Devil's staircase)* — the continuous, non-decreasing function $[0, 1] \rightarrow [0, 1]$ that is constant on every gap of the middle-third Cantor set and increases only on the set itself.

LITERATURE CITED

- Akhadkulov, H. and Y. Jiang, 2019 Volume of the hyperbolic cantor sets. *Dynamical Systems* **35**: 185–196.
- Astels, S., 1999 Thickness measures for cantor sets. *Electronic Research Announcements of the American Mathematical Society* **5**: 108–111.
- Astels, S., 2000 Cantor sets and numbers with restricted partial quotients. *Transactions of the American Mathematical Society* **352**: 133–170.
- Babich, A., 1992 Scrawny cantor sets are not definable by tori. *Proceedings of the American Mathematical Society* **115**: 829.
- Bartoszewicz, A., M. Filipczak, S. Glab, F. Prus-Wisniowski, and J. Swaczyna, 2017 On generating regular cantorvals connected with geometric cantor sets. *arxiv - Cornell university. Pre-print.*
- Batakis, A. and G. Havard, 2024 Dimensions of harmonic measures on non-autonomous cantor sets. *arxiv - Cornell university. Pre-print.*
- Batakis, A. and A. Zdunik, 2015 Hausdorff and harmonic measures on non-homogeneous cantor sets. *Annales Academiae Scientiarum Fennicae Mathematica* **40**: 279–303.
- Bellissard, V., J. and A. Julien, 2013 Bi-lipshitz embedding of ultrametric cantor sets into euclidean spaces. *arxiv - Cornell university. Pre-print.*
- Brouwer, L. E. J., 1910 On the structure of perfect sets of points. *Proceedings of the Royal Netherlands Academy of Arts and Sciences* **12**: 785–794.
- Cabrelli, C., U. Darji, and U. Molter, 2011 Visible and invisible cantor sets. *arxiv - Cornell university Pre-print.*
- Cheraghi, D. and M. Pedramfar, 2019 Hairy cantor sets. *arxiv - Cornell university. Pre-print.*
- Cockerill, E., 2002 Multi-model cantor sets. *arxiv - Cornell university. Pre-print.*
- Denette, E., 2016 *Minimal Cantor Sets: The Combinatorial Construction of Ergodic Families and Semi-conjugations*. Ph.D. thesis, University of Rhode Island, Open Access Dissertations. Paper 436.
- Edgar, G., 2007 *Measure, Topology, and Fractal Geometry*. Springer, New York, NY, USA.
- Elaydi, S., 2005 *An Introduction to Difference Equations*. Springer, New-york, NY, USA, fourth edition.
- Eswarathan, S. and X. Han, 2023 Fractal uncertainty principle for discrete cantor sets with random alphabets. *Mathematical Research Letters* **30**: 1657–1679.
- Falconer, K., 2014 *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley & Sons, Chichester, West Sussex, UK.
- Fillman, J. and S. H. Tidwell, 2023 On sums of semibounded cantor sets. *Rocky Mountain Journal of Mathematics* **53**: 737–754.
- Fleron, J. F., 1994 A note on the history of the cantor set and cantor function. *Mathematics Magazine* **67**: 136–140.
- Fletcher, A., D. Stoertz, and V. Vellis, 2022 Genus g cantor sets and germane julia sets. *arxiv - Cornell university Pre-print.*
- Fletcher, A. and J. Wu, 2014 Julia sets and wild cantor sets. *arxiv - Cornell university. Pre-print.*
- Frame, M., B. Mandelbrot, and N. Neger, 2025 Ancient egyptian columns [web page]. in *fractals in african art*. Department of Mathematics. Yale University, Retrieved June 15, 2025.
- Garcia, I., 2011 A family of smooth cantor sets. *Annales Fennici Mathematici* **36**: 21–45.
- Gorodetski, A. and S. Northrup, 2015 On sums of nearly affine cantor sets. *arxiv - Cornell university. Pre-print.*
- Hare, K. and K.-S. Ng, 2015 Hausdorff and packing measures of balanced cantor sets. *Real Analysis Exchange* **40**: 113.
- Hare, K. G. and N. Sidorov, 2024 The minkowski sum of linear cantor sets. *Acta Arithmetica* **212**: 173–193.
- Hassan, M. K., M. Z. Hassan, and N. I. Pavel, 2018 Extensive analytical and numerical investigation of the kinetic and stochastic cantor set. *arxiv - Cornell university Pre-print:0907.4837.*
- Hieronymi, P., 2018 A tame cantor set. *Journal of the European Mathematical Society* **20**: 2063–2104.
- Hochman, M., 2014 On self-similar sets with overlaps and inverse theorems for entropy. *Annals of Mathematics* **180**: 773–822.
- Honary, B., M. Pourbarat, and M. R. Velayati, 2006 On the arithmetic difference of affine cantor sets. *Journal of Dynamical Systems and Geometric Theories* **4**: 139–146.
- Hu, M. and S. Wen, 2008 Quasisymmetrically minimal uniform cantor sets. *Topology and Its Applications* **155**: 515–521.
- Hutchinson, J. E., 1981 Fractals and self-similarity. *Indiana University Mathematics Journal* **30**: 713–747.
- Jin, Z. A., 2021 Cantor sets in topology, analysis, and financial markets. Department of Mathematics. University of Chicago, Retrieved June 15, 2025.
- Jumha, S. H., I. O. Hamad, and A. O. Hassan, 2024 Cantor set from a nonstandard viewpoint. *Polytechnic Journal* **14**, Article 12.

- Kainth, S. P. S., 2023 *A Comprehensive Textbook on Metric Spaces*. Springer, Singapore.
- Kamalutdinov, K. and A. Tetenov, 2018 Twofold cantor sets in $\mathbb{R}$. arxiv - Cornell university. Pre-print.
- Kechris, A. S., 1995 *Classical Descriptive Set Theory*. Springer, New York, NY, USA.
- Kong, D., 2014 On the self similarity of generalized cantor sets. *SCIENTIA SINICA Mathematica* **44**: 945–956.
- Krushkal, V., 2018 Sticky cantor sets in $\mathbb{R}^d$. *Journal of Topology and Analysis* **10**: 477–482.
- Łaba, I. and M. Pramanik, 2009 Arithmetic progressions in sets of fractional dimension. *Geometric and Functional Analysis* **19**: 429–456.
- Lapidus, M. L. and H. Lu, 2008 Nonarchimedean cantor set and string. *Journal of Fixed Point Theory and Applications* **3**: 181–190.
- Li, W. and Y. Yao, 2008 The pointwise densities of non-symmetric cantor sets. *International Journal of Mathematics* **19**: 1121–1135.
- Lin, Z., Y. Qiu, and K. Wang, 2024 Number rigid determinantal point processes induced by generalized cantor sets. arxiv - Cornell university. Pre-print.
- Lind, D. and B. Marcus, 1995 *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, Cambridge, UK.
- Mandelbrot, B. B., 1982 *The Fractal Geometry of Nature*. W. H. Freeman, New York, NY, USA.
- Mantica, G., 2015 Numerical computation of the isospectral torus of finite gap sets and of ifs cantor sets. arxiv - Cornell university. Pre-print.
- Martens, M., 1994 Distortion results and invariant cantor sets of unimodal maps. *Ergodic Theory and Dynamical Systems* **14**: 331–349.
- Mattila, P., 1995 *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability*, volume 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, UK.
- Mauldin, R. D. and S. C. Williams, 1986 Random recursive constructions: Asymptotic geometric and topological properties. *Transactions of the American Mathematical Society* **295**: 325–346.
- McLoughlin, M. P. M. M. and C. R. Krizan, 2014 Some results on distorted cantor sets. *International Journal of Mathematical Analysis* **8**: 35–49.
- Moreira, C. G., M. Pacifico, and S. Rom  a Ibarra, 2020 Hausdorff dimension, lagrange and markov dynamical spectra for geometric lorenz attractors. *Bulletin of the American Mathematical Society* **57**: 269–292.
- Moreira, C. G. and J.-C. Yoccoz, 2001 Stable intersections of regular cantor sets with large hausdorff dimensions. *Annals of Mathematics* **154**: 45–96.
- Moschovakis, Y. N., 2009 *Descriptive Set Theory*, volume 155. American Mathematical Society, Santa Monica, CA, USA.
- Munkres, J. R., 2000 *Topology*. Prentice Hall, second edition.
- Nadler, S. B., Jr., 1992 *Continuum theory: An introduction*. Marcel Dekker.
- Nakata, T., 1997 An approximation of hausdorff dimensions of generalized cookie-cutter cantor sets. *Hiroshima mathematical journal* **27**: 467–475.
- Nassiri, M. and M. Z. Bidaki, 2025 Cantor sets in higher dimension i : Criterion for stable intersections. arxiv - Cornell university. Pre-print.
- Newhouse, S. E., 1979 The abundance of wild hyperbolic sets and non-smooth stable sets for diffeomorphisms. *Publications Math  matiques de l’IH  S* **50**: 101–151.
- Nowakowski, P., 2021 The family of central cantor sets with packing dimension zero. *Tatra Mountains Mathematical Publications* **78**: 1–8.
- Palis, J. and F. Takens, 1993 *Hyperbolicity and Sensitive Chaotic Dynamics at Homoclinic Bifurcations: Fractal Dimensions and Infinitely Many Attractors*, volume 35 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, UK.
- R Core Team, 2022 R: A language and environment for statistical computing (version 4.2.0). [Computer software]. R Foundation for Statistical Computing. Vienna, Austria.
- Rani, M. and S. K. Prasad, 2010 Superior cantor sets and superior devil staircases. *International Journal of Artificial Life Research* **1**: 78–84.
- Rosenthal, J. S., 2006 *A First Look at Rigorous Probability Theory*. World Scientific Publishing, Singapore, Singapore, second edition.
- Salem, R., 1951 On singular monotonic functions whose spectrum has a given hausdorff dimension. *Arkiv f  r Matematik* **1**: 353–365.
- Schroeder, M. R., 2009 *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*. Dover Publications, Mineola, NY, USA.
- Searcoid, M., 2006 Lipschitz functions. In *Metric Spaces*, Springer undergraduate mathematics series, Springer-Verlag, New York, NY, USA.
- Semmes, S., 2011 p -adic heisenberg cantor sets, 2. arxiv - Cornell university. Pre-print.
- Shmerkin, P., 2019 On furstenberg’s intersection conjecture, self-similar measures, and the l^q norms of convolutions. *Annals of Mathematics* **189**: 319–391.
- Soltanifar, M., 2006a A different description of a family of middle- α cantor sets. *American Journal of Undergraduate Research* **5**.
- Soltanifar, M., 2006b On a sequence of cantor fractals. *Rose-Hulman Undergraduate Mathematics Journal* **7**: 9.
- Soltanifar, M., 2021 A generalization of the hausdorff dimension theorem for deterministic fractals. *Mathematics* **9**: 1546.
- Soltanifar, M., 2022 The second generalization of the hausdorff dimension theorem for random fractals. *Mathematics* **10**: 706.
- Soltanifar, M., 2023 A classification of elements of function space $f(\mathbb{R}, \mathbb{R})$. *Mathematics* **11**: 3715.
- Tresser, C., 1991 Fine structure of universal cantor sets. In *Instabilities and Nonequilibrium Structures III*, edited by E. Tirapegui and W. Zeller, volume 64 of *Mathematics and Its Applications*, Springer, Dordrecht, Netherlands.
- Vallin, R. W., 2013 *The Elements of Cantor Sets: With Applications*. John Wiley & Sons, Hoboken, NJ, USA.
- Zeng, Y., 2017 Self-similar subsets of symmetric cantor sets. *Fractals* **25**: 1750003.
- Zhou, Q., X. Li, and Y. Li, 2019 Deformations on symbolic cantor sets and ultrametric spaces. arxiv - Cornell university. Pre-print.

How to cite this article: Soltanifar, M. Descriptions of Cantor Sets: A Set-Theoretic Survey and Open Problems. *Chaos and Fractals*, 3(2), 65-75, 2026.

Licensing Policy: The published articles in CHF are licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

