

Haros graphs: an exotic representation of real numbers

Jorge Calero-Sanz^{1,2,*}, Bartolo Luque^{1,†} and Lucas Lacasa^{3‡}

¹*Departamento de Matemática Aplicada a la Ingeniería Aeroespacial,
ETSIAE, Universidad Politécnica de Madrid, Madrid, Spain;*

²*Signal and Communications Theory and Telematic Systems and Computing, Rey Juan Carlos University, Madrid, Spain*

³*Institute for Cross-Disciplinary Physics and Complex Systems IFISC (CSIC-UIB), Palma de Mallorca, Spain.*

(Dated: August 19, 2022)

This paper introduces Haros graphs, a construction which provides a graph-theoretical representation of real numbers in the unit interval reached via paths in the Farey binary tree. We show how the topological structure of Haros graphs yields a natural classification of the real numbers into a hierarchy of families. To unveil such classification, we introduce an entropic functional on these graphs and show that it can be expressed, thanks to its fractal nature, in terms of a generalised de Rham curve. We show that this entropy reaches a global maximum at the reciprocal of the Golden number and otherwise displays a rich hierarchy of local maxima and minima that relate to specific families of irrationals (noble numbers) and rationals, overall providing an exotic classification and representation of the reals numbers according to entropic principles. We close the paper with a number of conjectures and outline a research programme on Haros graphs.

Keywords: Number theory, graph entropy, fractals, complex systems, representation of real numbers, Golden number

I. INTRODUCTION

The structure of real numbers has been a traditional topic of study in mathematics [1–3], tackled from several angles, both from a computational perspective –e.g. the quest of exact arithmetic for the approximation of rationals and reals [4, 5] – and from a theoretical one, e.g. the characterisation of different families of irrationals such as transcendental or normal numbers [1, 6]. Among other techniques [3], representing numbers in terms of a continued fraction expansion [7, 8] has been a very fruitful avenue. A classical approach is to consider the Farey binary tree, where rationals acquire a natural ordering [9]. Some works [9, 10] present the relationship between Farey sequences and continued fractions. Moreover, other objects can be related such as the modular group [11], and the dynamic and thermodynamic perspectives are highlighted [12].

In parallel and also recently, a new technique for representing sequences of numbers in terms of graphs has been proposed [13, 14]: Visibility Graphs—either Natural Visibility Graphs (VGs) or Horizontal Visibility Graphs (HVGs)—are combinatorial representations of sequences of numbers, and a number of works [15–17] have made it clear that the structure of such graphs inherit several properties of the associated sequences, thereby making them a useful technique for time series analysis and feature-based classification of complex signals. In addition to their applicability for signal processing, a range of theoretic works have studied how the trajectories of dynamical systems are mapped in graph space. In this sense, different routes to low-dimensional chaos have been explored, including the Feigenbaum scenario [18], the Pomeau-Manneville scenario [19], and the quasiperiodic route [20], and a number of dynamical properties such as Lyapunov exponents or renormalization-group features have been shown to be inherited in the associated graphs. Interestingly, the Horizontal Visibility Graph extracted from the circle map certifies that the mode locking frequencies –that relate to specific rationals– have a very concrete graph-theoretical footprint, and similarly some selected irrationals –i.e. the Golden number, and others– acquire a natural graph-theoretic representation.

Motivated by these facts, in this work, we build on the dynamical association between rationals and HVGs and proceed to construct a graph-theoretical representation of rational and irrational numbers in the Farey binary tree. As we will show, such representation can be defined independently of HVGs –and thus no actual time series is

*Electronic address: jorge.calero@urjc.es

†Electronic address: bartolome.luque@upm.es

‡Electronic address: lucas@ifisc.uib-csic.es

required—, although some results from HVG theory will be used, e.g. [20]. Our contention is that the topological structure of such graphs inherits combinatorially important properties of numbers. We will show that such an exotic graph-theoretical representation allows for a classification of rationals and irrationals numbers, as well as for finer classifications inside the set of irrationals in the unit interval $[0, 1]$, which unveils an extremely rich and self-similar structure.

The rest of the paper goes as follows: In Section II we present the necessary background on Farey sequences, the Farey binary tree, and continued fractions. In Section III, we introduce Haros graphs, which are graph-theoretical representations of rationals and irrationals numbers. Note at this point that Haros graphs are different from the so-called Farey graphs [10, 21], hence our labelling ¹. In this section, we also explore both analytically and numerically the degree distribution of these graphs in relation to the number they represent, and we also study properties of the degree sequence in relation to Khinchin's constant. In Section IV, we proceed to study a graph-entropy function, calculated from the degree distribution. We find that the such entropy has a fractal shape which can be characterised in closed-form in terms of a generalised de Rham curve: a self-affine function, continuous but not differentiable at any point. By further exploring the extrema of such curve, we find that it has an infinite number of local minima, which can be classified into families of rational numbers, and in turn, the infinite number of local maxima are related to irrationals that can be naturally parametrised into families of noble numbers. In Section V, we conclude, state some conjectures, and outline several open questions that can form the basis for a research programme on Haros graphs.

II. PRELIMINARIES

A. Farey sequences

In 1816, the geologist John Farey [22] proposed to order from the smallest to the largest the irreducible fractions with a denominator less than or equal to n belonging to the interval $[0, 1]$, hence defining the so-called *Farey sequence* of order n , \mathcal{F}_n , as the ordered set of all irreducible fractions $0 < p/q < 1$, $p, q \in \mathbb{N}^+$, whose denominators do not exceed n . The first four sequences are:

$$\mathcal{F}_1 = \left\{ \frac{0}{1}, \frac{1}{1} \right\}; \quad \mathcal{F}_2 = \left\{ \frac{0}{1}, \frac{1}{2}, \frac{1}{1} \right\}; \quad \mathcal{F}_3 = \left\{ \frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1} \right\}; \quad \mathcal{F}_4 = \left\{ \frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1} \right\},$$

and, in general: $\mathcal{F}_n = \{p/q \in [0, 1] : 0 \leq p \leq q \leq n, (p, q) = 1\}$.

The mediant \oplus of two rational numbers p/q and p'/q' is defined as:

$$\frac{p}{q} \oplus \frac{p'}{q'} := \frac{p+p'}{q+q'}.$$

For two consecutive rationals in a Farey sequences \mathcal{F}_n , the mediant of these terms is indeed another irreducible fraction:

$$\frac{p}{q} < \frac{p+p'}{q+q'} < \frac{p'}{q'},$$

that does not appear in \mathcal{F}_n but, provided that $q+q' \leq n+1$, appears in \mathcal{F}_{n+1} [1, 2]. Thus, one can iteratively construct Farey sequences, using \mathcal{F}_n and the mediant operator to construct \mathcal{F}_{n+1} . For instance, using the mediant sum over \mathcal{F}_1 we get:

$$\frac{0}{1} \oplus \frac{1}{1} = \frac{1}{2},$$

which is the middle term in \mathcal{F}_2 . Similarly:

$$\mathcal{F}_3 = \left\{ \frac{0}{1}, \frac{0}{1} \oplus \frac{1}{2} = \frac{1}{3}, \frac{1}{2}, \frac{1}{2} \oplus \frac{1}{1} = \frac{2}{3}, \frac{1}{1} \right\}.$$

¹ Interestingly, the classical concept of Farey graph [21] is indeed similar to the concept of Feigenbaum graph, HVGs extracted from the Feigenbaum scenario in the period-doubling bifurcation cascade [18]

Observe, however, that some of the terms resulting from the mediant sum over \mathcal{F}_3 ($\frac{1}{3} \oplus \frac{1}{2} = \frac{2}{5}$ and $\frac{1}{2} \oplus \frac{2}{3} = \frac{3}{5}$) are not in \mathcal{F}_4 as the denominator exceeds $n = 4$, in this sense \mathcal{F}_n is a subset of the set formed by sequentially applying the mediant sum to adjacent pairs of terms in \mathcal{F}_{n-1} .

Each Farey sequence \mathcal{F}_n induces a partition of $[0, 1]$ into $\sum_{j=1}^n \varphi(j)$ subintervals of different sizes (where $\varphi(j)$ is the Euler totient function of j), which we will call Farey subintervals of order n . For example, \mathcal{F}_3 induces the partition:

$$[0, 1] = \left[0, \frac{1}{3}\right) \cup \left[\frac{1}{3}, \frac{1}{2}\right) \cup \left[\frac{1}{2}, \frac{2}{3}\right) \cup \left[\frac{2}{3}, 1\right].$$

As n grows, it is easy to see that \mathcal{F}_n provides finer approximations for the ordering of rational numbers, so that all rational numbers will eventually be ordered and enumerated in the limit set $\mathcal{F} := \lim_{n \rightarrow \infty} \mathcal{F}_n$. Therefore, all rational numbers in the unit interval are contained in \mathcal{F} [2], something conjectured by Farey and independently proved by Haros and Cauchy [23]. At this point, we should emphasise that, while popularly known as Farey sequences, \mathcal{F}_n were originally and independently studied by Charles Haros, publishing similar results (including some proofs) some 15 years before Farey [24]. In tribute, our graph-theoretic construction (Section III) is called Haros graphs.

B. Farey binary tree, binary sequences and continued fractions

A classical way of visualising the different elements of \mathcal{F}_n is using the so-called *Farey binary tree* (see Fig. 1 for an illustration). This is an infinite complete binary tree (adding the nodes $0/1$ and $1/1$ as special cases) where each node corresponds to a given rational, computed using the mediant sum over specific numbers above in the tree. This tree is hierarchically organised in levels ℓ_n , such that the first 4 levels are:

$$\ell_1 = \left\{ \frac{0}{1}, \frac{1}{1} \right\}; \ell_2 = \left\{ \frac{1}{2} \right\}; \ell_3 = \left\{ \frac{1}{3}, \frac{2}{3} \right\}, \ell_4 = \left\{ \frac{1}{4}, \frac{2}{5}, \frac{3}{5}, \frac{3}{4} \right\}.$$

For $n \leq 3$, we have $\mathcal{F}_n = \cup_{i=1}^n \ell_i$, i.e. the Farey sequence corresponds to projecting all ancestors in the Farey binary tree up to that level. However, this is no longer true for $n > 3$, for which we have:

$$\mathcal{F}_n \subset \bigcup_{i=1}^n \ell_i.$$

For instance: $\cup_{i=1}^4 \ell_i = \mathcal{F}_4 \cup \{2/5, 3/5\}$ (the mismatch is due to the definition of the Farey sequence, which imposes the denominator $q \leq n$). While the offset:

$$\left| \bigcup_{i=1}^n \ell_i \right| - |\mathcal{F}_n|,$$

grows without bounds, in the limit $n \rightarrow \infty$, one can prove that:

$$\text{card} \left(\bigcup_{i=1}^{\infty} \ell_i \right) = \text{card}(\mathcal{F}),$$

and, therefore, the Farey binary tree orders all real numbers [4].

In this work, we consider the ordering as provided by the Farey binary tree rather than by the Farey sequences. However, for computational reasons, the numerical calculus used to represent some figures has been computed for the Farey sequence \mathcal{F}_{1000} . There are exponentially many different downstream paths of a certain depth in the Farey tree starting from $1/1$ and one can enumerate these by assigning to each path a unique binary sequence of symbols L/R indicating descending to the Left/Right, respectively. For example, it can be observed in Fig. 1 that the rational number $8/13$ follows the symbolic path $LRLRL$ in the Farey Binary Tree. Since any finite path has an ending element p/q and such rational is reached by a single finite path in the Farey binary Tree, one can establish a bijection between rational numbers in the unit interval and finite binary sequences [9, 12].

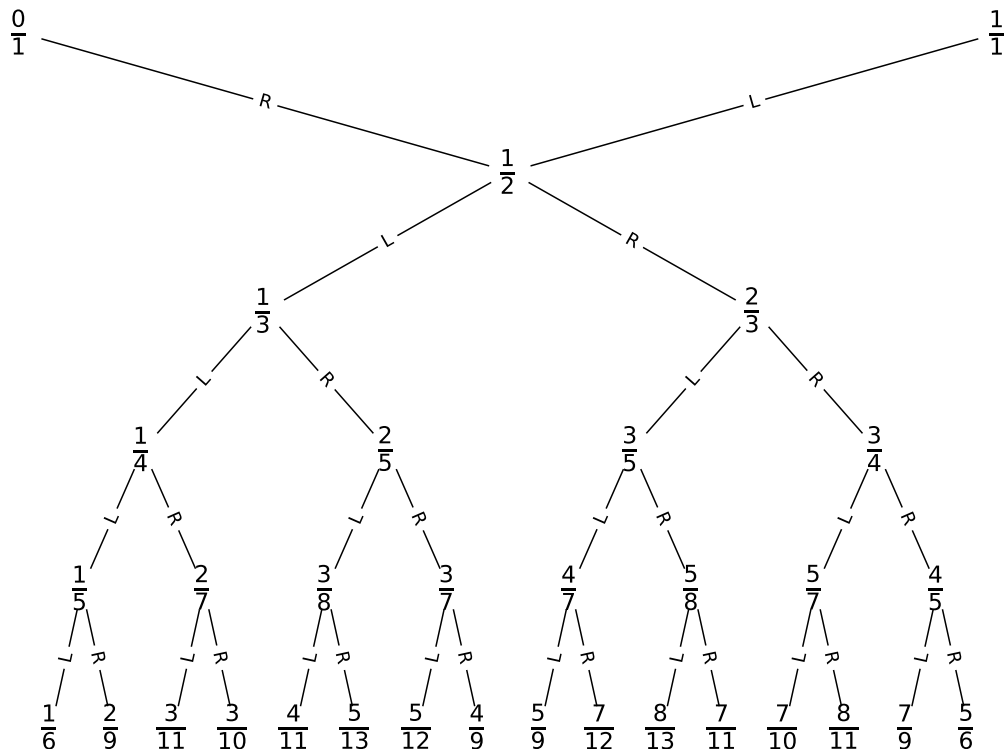


Fig. 1. Six levels of the Farey binary tree, where the special nodes $0/1$ and $1/1$ form the first level, ℓ_1 . Every irreducible rational number appears only once in the tree and can be reached following a path which is encoded in a symbolic binary sequence Left/Right. By convention, paths start in $1/1$; hence, the first symbol in every path will be L . The Farey sequence \mathcal{F}_n is a subset of the rationals appearing in $\cup_{i=1}^n \ell_i$.

Another way of representing rationals is via continued fractions. It is well known that irrational numbers in $[0, 1]$ admit a description in terms of an infinite continued fraction, whose finite truncations (called convergents) are rational numbers. For instance, the reciprocal of the Golden number $\phi^{-1} = 0.618\dots$ is a solution of the quadratic equation $x^2 + x - 1 = 0$. Therefore, it fulfils $\phi^{-1} = \frac{1}{1+\phi^{-1}}$ which, when interpreted recursively, leads us to:

$$\phi^{-1} = \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \dots}}}$$

The above equation determines that the reciprocal of the Golden number admits an expression as a continued fraction as: $\phi^{-1} = [1, 1, \dots] = [\bar{1}]$. The successive convergents of this continued fraction are: $[1] = 1$, $[1, 1] = 1/2$, $[1, 1, 1] = 2/3$, $[1, 1, 1, 1] = 3/5$, and so on. Interestingly, note that the Fibonacci sequence is generated by the recurrence equation $\mathfrak{F}_n = \mathfrak{F}_{n-1} + \mathfrak{F}_{n-2}$ and the initial conditions $\mathfrak{F}_0 = \mathfrak{F}_1 = 1$. The Fibonacci quotients $\{\mathfrak{F}_{n-1}/\mathfrak{F}_n\}_{n \geq 1}$ correspond to the convergents of ϕ^{-1} and: $\lim_{n \rightarrow \infty} \mathfrak{F}_{n-1}/\mathfrak{F}_n = \phi^{-1}$.

Now, for any given rational (or irrational) number, there exists a close relationship between its continued fraction expansion and its representation in terms of a binary sequence in the Farey binary tree [9]. Indeed, a continued fraction $[a_1, a_2, \dots]$ has a correspondent symbolic path in the Farey binary tree as $L^{a_1} R^{a_2} L^{a_3} \dots$, where L^q is interpreted as a sequence of q L s (in the rational case, the last symbol has an index $a_n - 1$). For instance, $\phi^{-1} = [1, 1, 1, 1, 1, \dots] = LRLRLR\dots = \overline{LR}$, that is, an infinite zigzag.

As a result, and since any two subsequent convergents can be reached by a specific path in the Farey binary tree, it turns out that any irrational number in $[0, 1]$ can be reached asymptotically via a particular infinite path. In other words, any irrational in $[0, 1]$ can be approximated with arbitrary precision via convergents [6], in such a way that in the limit the irrationals are reached asymptotically and therefore $\cup_{i=1}^{\infty} \ell_i = \mathcal{F} = [0, 1]$.

III. HAROS GRAPHS

A. Definition and basic properties

We start by considering a graph concatenation operator \oplus (note that this is different from the median operator \oplus defined in the previous section, but the abuse of notation is convenient and will become evident later). This operator was originally defined with a different notation (a so-called inflation operator) in [25], and we refer the reader to that paper for a precise definition. Here, instead, we provide a visual description, which we believe is more intuitive than the formal definition. Visually, the concatenation of two graphs consists on merging two extreme nodes (so-called boundary nodes) and connect the first and last nodes of the new graph with a new link, see Fig. 2 for an illustration. The boundary nodes of the two graphs that are involved in the merging are called merging nodes. It is easy to see that \oplus is not commutative.

We are now ready to define Haros graphs:

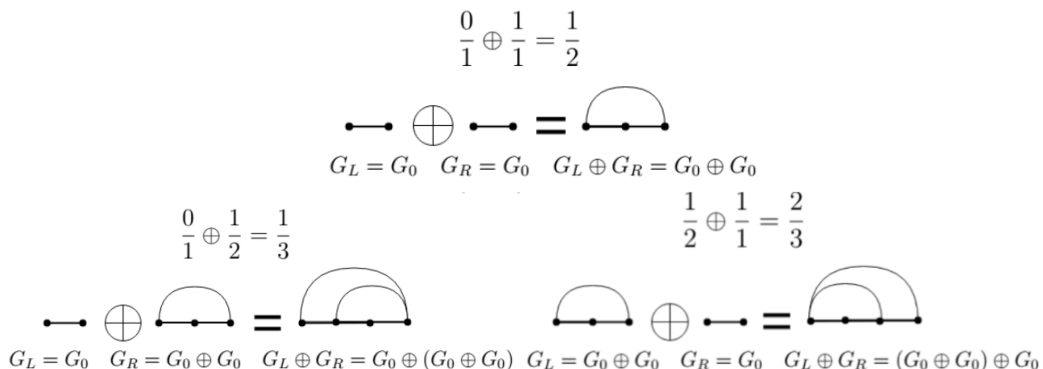


Fig. 2. (a) First line: Construction of $G_1 = G_0 \oplus G_0$. The graph concatenation operator merges the boundary nodes: labelling the nodes of G_0 as $[1, 2]$, the boundary nodes are 2 for G_L and 1 for G_R . The graph obtained by the concatenation has three nodes and the external nodes (1 and 3 in $G_L \oplus G_R$) are linked by a new vertex. (b) Second line: Construction of the two possible concatenations between G_0 and G_1 .

Definition (Haros graph) *Haros graphs are built by iteratively applying the concatenation operator \oplus on a previous (ancestor) Haros graph, where the oldest ancestor G_0 consists of two nodes linked by an edge (see Fig. 3 for an illustration). Since \oplus is not commutative, one can apply \oplus in two ways to produce two offspring Haros graphs from the same ancestor Haros graph G , by either ‘left’ concatenating ($X_\ell \oplus G$) and ‘right’ concatenating ($G \oplus X_r$). X_ℓ is the closest ‘left’ ancestor of G and X_r is the closest ‘right’ ancestor of G (Fig. 3 clarifies this).*

The successive applications of \oplus thus induce a binary sequence (left/right) very much like the one which enumerates paths in the Farey binary tree, actually Fig. 3 is in one-to-one correspondence to the Farey binary tree. Since binary sequences describing paths in such tree are in a one-to-one correspondence with rationals, we can thus make a one-to-one correspondence between rationals and Haros graphs.

We illustrate now how to relate a given rational $x = p/q$ (where p/q is an irreducible fraction) to its Haros graph G_x (see again Figs. 2 and 3 for an illustration): by definition we let $G_0 = G_{0/1} = G_{1/1}$ to be associated at the same time to rational numbers $0/1$ and $1/1$. Then we build $G_0 \oplus G_0$ and associate it with $G_{1/2}$, the Haros graph related to the rational number $1/2$. At this point, we can now build two new possible Haros graphs, depending on how we concatenate again: $G_0 \oplus G_{1/2}$, which is identified with $G_{1/3}$, or $G_{1/2} \oplus G_0$, which is identified with $G_{2/3}$. In other words, we have the following equation:

$$G_{p/q} \oplus G_{p'/q'} = G_{p/q \oplus p'/q'}, \quad (1)$$

i.e. the concatenation operator between Haros graphs is identified with the median sum between rationals. As a result, we obtain a one-to-one correspondence between the set of Haros graphs and the set of Farey sequences, and accordingly every number $x \in [0, 1]$ can now be represented by a unique Haros graph G_x . If x is rational, there exists a finite binary sequence in the Farey binary tree that reaches x , and equivalently there is a finite sequence of concatenations that reaches the associated Haros graph; thus, G_x is a finite graph that we denote a rational Haros

graph. On the contrary, if x is irrational, then G_x is a countable infinite graph that denote an irrational Haros graph.

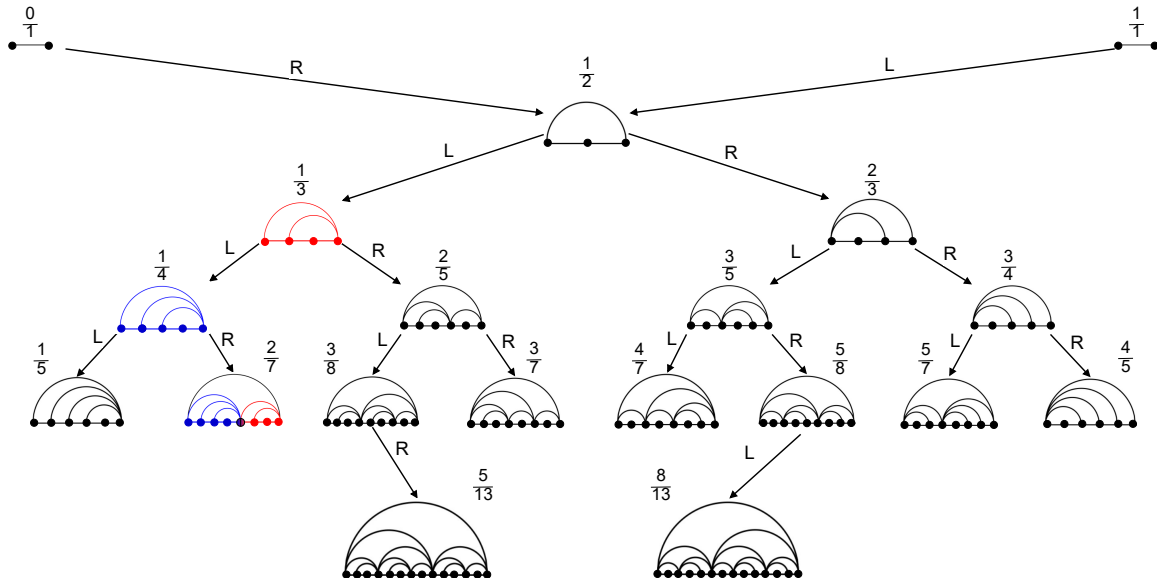


Fig. 3. Six levels of the Haros graph binary tree with the graphs associated with the corresponding rational fractions p/q (for space reasons only two of these are shown at the sixth level). On the left, we highlight the Haros graph $G_{2/7}$, built as a concatenation of $G_{1/4}$ (blue) and $G_{1/3}$ (red).

Boundary node convention. By construction, $G_{p/q}$ is an Haros graph of $q + 1$ nodes, where we have two extreme nodes (respectively, the rest of the nodes are called *inner* nodes). Some results presented in the following sections will establish the relationship between the number p/q and the proportion of nodes in $G_{p/q}$ with a specific property. For convenience and to make analytical calculations cleaner in what follows, we assume that the two extreme nodes of an Haros graph (linked by construction) are identified as a single ‘boundary’ node, while maintaining both incident edges. This means that under this convention the Haros graph $G_{p/q}$ has effectively q nodes. Furthermore, as we point out in lemma 5, under this convention allows a relationship between the degree of the boundary node and the level of Haros graph tree. For instance, if the first and last nodes have degree 2, then our convention is to say that the boundary node has degree 4. Such a boundary node is depicted at the end of the node sequence.

For example, the degree sequence of $G_{1/2}$ is:

$$[2, 2, 2] \rightarrow [2, 2 + 2] = [2, 4],$$

and for:

$$G_{1/3} = [2, 3, 2, 3] \rightarrow [3, 2, 2 + 3] = [3, 2, 5].$$

Similarly:

$$G_{1/3} \oplus G_{1/2} = G_{2/5},$$

with degree sequences:

$$[2, 3, 2, 3] \oplus [2, 2, 2] = [2 + 1, 3, 2, 3 + 2, 2, 2 + 1] = [3, 3, 2, 5, 2, 3],$$

which is then presented as $[3, 2, 5, 2, 6]$.

B. Degree distribution

While trivially rational numbers $x \in \mathbb{Q}$ have G_x with a finite vertex set and irrational numbers $x \in \mathbb{I}$ have an infinite vertex set, we wonder if the structure of these graphs allows us to further distinguish among classes of irrational numbers. Among the plethora of different graph properties at our disposal, we shall concentrate on the degree sequence,

i.e. the ordered sequence whose i -th term provides the degree (number of edges) of the i -th node in the graph. This choice is indeed informed by the fact that Haros graphs form a subset of Horizontal Visibility Graphs (HVG), and a recent theorem [26] proved that HVGs are unigraphs, and thus they are uniquely determined by their degree sequence, i.e., the degree sequence is a ‘maximally informative’ property of the graph [27]. In this section, we consider the probability degree distribution $P(k, x)$, that provides the probability that a node chosen at random in G_x has degree k .

By construction, Haros graphs do not have isolated nodes, so $P(0, x) = 0, \forall x$. Similarly, and with the exception of G_0 and G_1 , $P(1, x) = 0, \forall x$, so the subsequent analysis is on $k \geq 2$. We start by focusing on the interval $x \in [0, 1/2)$ and $P(k, x)$ for $k \leq 4$. We can state the following theorem:

Theorem 1. $\forall x \in (0, 1/2)$, the first values of the degree distribution $P(k, x)$ of G_x are:

$$P(k, x) = \begin{cases} x, & k = 2 \\ 1 - 2x, & k = 3 \\ 0, & k = 4. \end{cases} \quad (2)$$

See Appendix VIA for a proof. Now, an important but obvious observation is that there exists a mirror symmetry with respect to $1/2$ ($x \rightarrow 1 - x$) in the construction of Haros graphs (see Fig. 3) which directly implies $P(k, x) = P(k, 1 - x)$. Accordingly, for $x \in (1/2, 1)$ the values in Eq. 2 must be replaced by $P(2, x) = 1 - x$, $P(3, x) = 2x - 1$, and $P(4, x) = 0$.

Note that, conditioned on a generic but fixed degree $k = \kappa$, $P(\kappa, x)$ is a real-valued function with domain $[0, 1]$. In this sense, by construction, the boundary cases $G_{0/1}, G_{1/2}$ and $G_{1/1}$ are associated with isolated discontinuities of this function in $k = 2, 3, 4$. In Fig. 4 we show the numerical values for the functions $P(2, x), P(3, x)$ and $P(4, x)$ computed for all Haros graphs G_x with $x \in \mathcal{F}_{1000}$, in full agreement with the previous theorem and highlighting the above mentioned mirror symmetry. Consequently, from now on, we will restrict the analysis to the interval $x \in [0, 1/2]$.

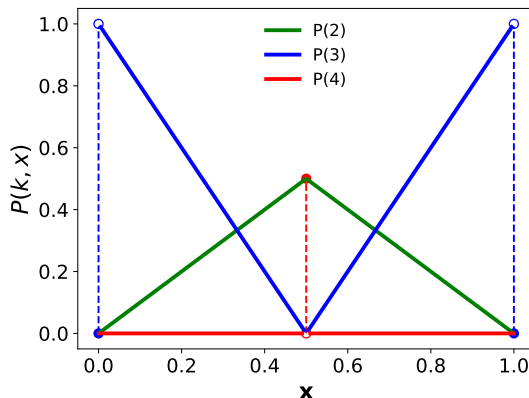


Fig. 4. Numerical plot of the degree distribution $P(2, x)$ (green), $P(3, x)$ (blue) and $P(4, x)$ (red) of G_x as a function of x , for all $x \in \mathcal{F}_{1000}$, showing a perfect good agreement with Eq. 2. Note that $P(3, x)$ presents two discontinuities in the graphs $G_{0/1}$ and $G_{1/1}$ (blue dots), and likewise $P(4, x)$ presents a discontinuity at $G_{1/2}$.

It is not easy to find in closed form an expression for $P(k, x)$ for a generic k, x . Let us consider the case $k = 5$. One can prove by induction that:

$$P(5, x) = \begin{cases} 3x - 1, & x \in (1/3, 1/2) \\ -3x + 2, & x \in (1/2, 2/3) \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Notice that the degree $k = 5$ appears with a non-null probability for graphs G_x where x belongs to the subintervals induced by $\cup_{i=1}^3 \ell_i = \mathcal{F}_3$: $(1/3, 1/2)$ and $(1/2, 2/3)$. It turns out that a generic degree k only ‘emerges’ (i.e. leading to a non-null probability in the degree distribution) after a sufficiently long and specific descent in Farey’s binary tree, i.e., for specific subintervals of x . In this sense, the following theorem can be stated:

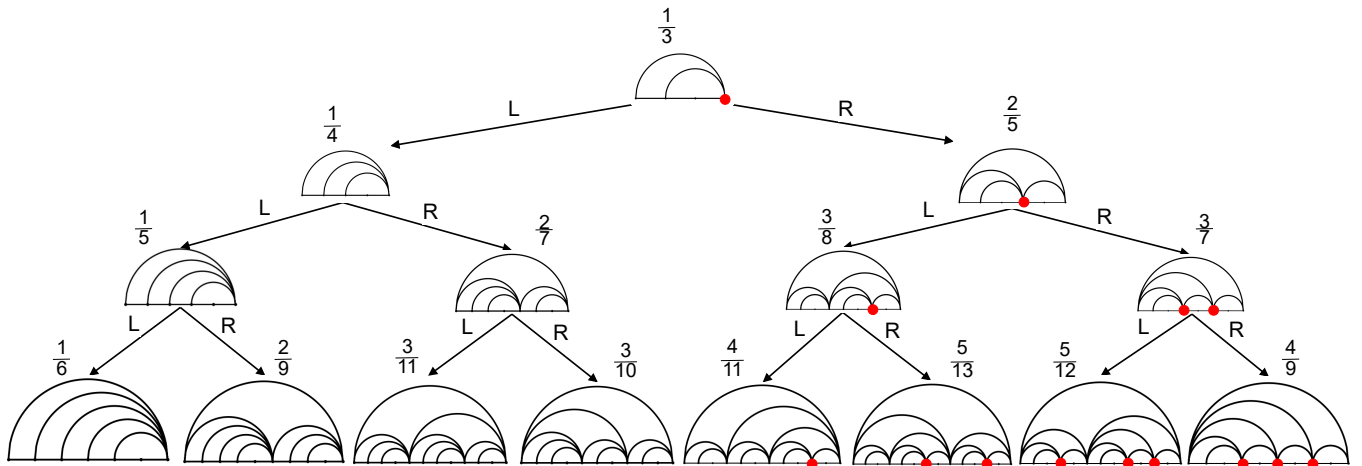


Fig. 5. Further visual aid of paths in the Haros graph binary tree. The symbolic path in the Haros graph tree of $G_{1/3} = G_{0/1} \oplus G_{1/2}$ is LL and its degree sequence is $[3, 2, 2 + 3 = 5]$, i.e., disregarding the extreme node of eventual degree $2 + 3 = 5$, then only degrees $k = 2, 3$ emerge. A further descent to the left L^3 generates the graph $G_{1/4}$ whose degree sequences is $[3, 3, 2, 2 + 4 = 6]$, where degree $k = 5$ does not appear; while if the descent is to the right LLR , the graph obtained is $G_{2/5}$ whose degree sequence is $[3, 2, 5, 2, 3 + 3 = 6]$. In the subtree starting at LLR (level $n = 3$), there are Haros graphs with nodes with degree $k = n + 2 = 5$ (red dots). However, in the Haros graphs whose paths start at LLL , no nodes with degree $k = 5$ appear. Nodes with degree $k = 5$ will not appear if they have not been generated in the descent starting at $1/3$.

Theorem 2. For all $\kappa \geq 5$, $P(k = \kappa, x) = 0$ if and only if there is a symbol repetition in the binary sequence of G_x (either LL or RR) at the position $\kappa - 2$, or equivalently, a LL or RR on the downstream path at the level $\kappa - 2$ of the Haros graph binary tree.

Proof of this theorem is available in Appendix VIB. This theorem has several implications:

- The emergence of a 'new' degree κ occurs for the first time at a certain level of the Haros graph binary tree (see Fig. 5 for an illustration). It can then be observed that if the last downstream step was 'L' (analogously 'R'), a subsequent step 'L' will not generate an Haros graph with inner nodes with degrees larger than κ , as its only effect is to replicate the degree sequences (except in the boundary node). On the contrary, if the subsequent downstream step makes a symbol change ($L \rightarrow R$ or $R \rightarrow L$), then the resulting Haros graph inherits an inner node with degree ($k > \kappa$).

An example will help us illustrate this observation: Consider $G_{1/3}$, generated as $G_{0/1} \oplus G_{1/2}$. Its associated path in the Haros graph binary tree is LL and its degree sequence is $[3, 2, 2 + 3 = 5]$, i.e., disregarding the boundary node of degree $2 + 3 = 5$, only the degrees $k = 2$ and $k = 3$. A further descent to the left (LLL) generates the graph $G_{1/4}$ whose degree sequence is $[3, 3, 2, 2 + 4 = 6]$, where the degree $k = 5$ does not appear, whereas if the descent was towards the right (LLR), the resulting graph is $G_{2/5}$ with degree sequence $[3, 2, 5, 2, 3 + 3 = 6]$, having an inner node with $k = 5$.

In figure 5, we can see how all Haros graphs whose binary sequence starts with LLR have inner nodes with $k = 5$, and this is not true for those graphs with binary sequence starting with LLL . In short, the degree $k = 5$ will not appear if it has not been generated in the descent starting from $G_{1/3}$. This fact is confirmed in Fig. 6, which illustrates that degree $k = 5$ has a non-zero probability only for Haros graphs in the interval $[1/3, 2/3]$ (except the case $x = 1/2$), which are those whose path in the tree starts with LLR or LRL .

- The hole patterns (the null values in $P(k, x)$) can be completely described by the binary sequence in the Farey binary tree, or equivalently, by the continued fraction of the number x . Hence, the construction of Haros graphs is intimately related to the subintervals induced by $\cup_{i=1}^n \ell_n$, see also Figure 6 for an illustration.

One can now use theorem 2 and its implications to discuss the properties of $P(k, x)$ for selected families of irrational numbers. For instance, consider the reciprocal of the Golden number. Its symbolic path in the Farey binary tree consists of an infinite zigzag \overline{LR} . Theorem 2 establishes that the degree distribution of its associated Haros graph is necessarily holeless, and, in fact, this is the only irrational number with such property. Furthermore, two classical results –the Lagrange's theorem and Galois' theorem– establish that the periodic continued fractions are, precisely,

the quadratic irrationals [1]. Therefore, by virtue of theorem 2, the Haros graphs of quadratic irrationals will have a degree distribution with a periodic pattern of zeros.

1. A conjecture for a closed expression of $P(k, x)$

Observe that we started the preceding section stating that a closed-form expression for a generic $P(k, x)$ is difficult to obtain. However, based on theorem 2 and on the triangular-like dependence on x observed for some values of k (Fig. 6) we can now state the following conjecture.

Conjecture 1. *Let ℓ_n be the set of Farey fractions emerging at level $n \geq 1$ of the Farey binary tree. For instance, $\ell_3 = \{1/3, 2/3\}$ and $\ell_4 = \{1/4, 2/5, 3/5, 3/4\}$, and, in general, at level ℓ_n there are $\text{Card}(\ell_n) = 2^{n-2}$ new fractions added. Then, for all real number x and degree $k \geq 5$ in G_x , we have:*

$$P(k, x) = \begin{cases} q_i \cdot x - p_i, & \text{if } x \in \left(\frac{p_i}{q_i}, \frac{a}{b}\right), \\ -q_{i+1} \cdot x + p_{i+1}, & \text{if } x \in \left(\frac{a}{b}, \frac{p_{i+1}}{q_{i+1}}\right), \\ 1/q_i, & \text{if } x = \frac{p_i}{q_i}, \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

with:

$$\frac{p_i}{q_i}, \frac{p_{i+1}}{q_{i+1}} \in \ell_{k-2}, \frac{a}{b} \in \ell_{k-3} \quad \text{and} \quad \frac{p_i}{q_i} < \frac{a}{b} < \frac{p_{i+1}}{q_{i+1}}.$$

We have numerically verified the correctness of this conjecture up to $k \leq 60$, and Fig. 6 illustrates it for $k \leq 8$. Equation 4 implies that the degree distribution is a piecewise-linear function in some subintervals defined by the elements at a certain level of the Farey binary tree. Moreover, the figure also points out several important aspects of the degree distribution:

- The emergence of different degrees is related to whether G_x is located in a specific subinterval defined by Farey fractions, in agreement with theorem 2 and its implications discussed above.
- $P(k, x)$ seems to have a self-similar structure.
- $P(k, x)$ seems to be continuous except on a null-measure set of points (red points), where we have removable discontinuities.

The next two subsections address these last two observations.

2. Self-similarity of $P(k, x)$

To dig into the particular scaling properties of $P(k, x)$, it can be observed that the right panel of Fig. 6 depicts the cumulative degree distributions, showing Cantor's staircase function shapes. This further suggests that $P(k, x)$ has a self-similar structure, where $P(k = \kappa + 1, x)$ is found by adequately scaling and shifting the base of the 'triangle' in $P(k = \kappa, x)$ (left panel of the same figure).

To explore the scaling equations, let us consider the Farey binary tree and define T_n to be the subtree whose asymptotic end nodes densely populate the interval $\mathcal{I}_n = (1/(n+1), 1/n]$ (see Fig.7 for an illustration of T_2, T_3 and T_4). Each T_n has a root node at the rational number $2/(2n+1)$ obtained as the mediant sum of $1/(n+1) \oplus 1/n$. In terms of symbolic paths in the Farey binary tree, the roots of T_n 's have symbolic paths $L^n R$ (Fig. 7 highlights the root nodes of T_2, T_3 and T_4 in green colour). The descendants of such roots have a generic symbolic path $L^n R\mathcal{P}$, where \mathcal{P} is an arbitrary binary sequence. For illustration, in Fig.7 we depict in magenta the nodes reached by the symbolic paths $L^n RL$, for $n \geq 2$, associated to rational numbers $3/(3n+2)$; and we depict in orange the nodes reached by the symbolic paths $L^n R^2$, for $n \geq 2$, associated to rational numbers $3/(3n+1)$.

In terms of continued fractions, the subintervals \mathcal{I}_n , or equivalently the subtree T_n starting in the root nodes $2/(2n+1)$, are formed by the real numbers which continued fractions have as the first element $a_1 = n$. The families of nodes described above have the following continued fraction expansions: blue nodes $\rightarrow [n]$, green nodes $\rightarrow [n, 2]$, magenta

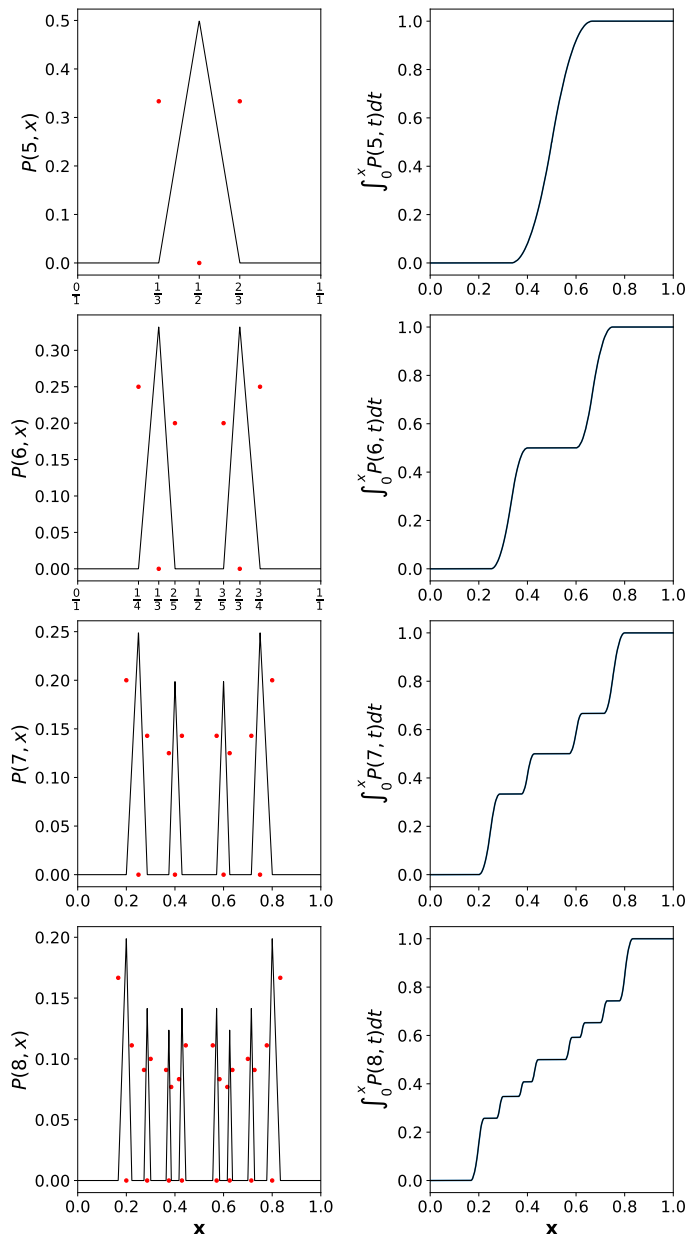


Fig. 6. (Left panels) Degree distribution $P(k, x)$ as a function of x , computed numerically for $k = 5, 6, 7, 8$, for all Haros graphs G_x with $x \in \mathcal{F}_{1000}$. Red points represent removable discontinuities. Solid lines are in perfect agreement with Eq.4. (Right panels): Cumulative distributions. As k grows, the cumulative distributions approach a Cantor's staircase function, suggesting self-similarity of $P(k, x)$. This figure supports the study of the behaviour of $P(k, x)$ established in Subsection III B. The 'peak' or triangle where the degree distribution is positive is located between two Farey fractions. These fractions are the nodes that form the level $n = \kappa - 2$ of the Farey binary tree (e.g. the nodes $1/3$ and $2/3$ are the level 3). For example, see the plot of $P(k = 5, x)$. The degree sequence is zero for Haros graphs G_x with $x < 1/3$ or $x > 2/3$, that is, Haros graphs whose symbolic path starts with LLL or LR . In the subinterval $(1/3, 1/2)$, $P(5, x) = 3x + 1$ whereas $P(5, x) = -3x + 2$ in $(1/2, 2/3)$. Moreover, the degree distribution $P(k + 1, x)$ can be obtained as a duplicated and scaled version of $P(k, x)$ through the function $F(x)$ defined in Eq. 5.

nodes $\rightarrow [n, 1, 2]$, and orange nodes $\rightarrow [n, 3]$. Now, since $[0, 1/2] = \cup_{i=2}^{\infty} \mathcal{I}_n$, any arbitrary (real) number x in $[0, 1/2]$ belongs to a given subtree T_k . Suppose for the sake of argument that $x \in \mathcal{I}_k$, and consider the function:

$$F(x) = \frac{x}{1+x}. \quad (5)$$

It is easy to see that the function $F(x)$ acts on $x = [a_1 = k, a_2, a_3, \dots]$, resulting in $F(x) = [k + 1, a_2, a_3, \dots]$ [28], hence, $F(x) \in \mathcal{I}_{k+1}$. In general, composing such a function $n > 0$ times, $F^{(n)}(x) \in \mathcal{I}_{k+n}$. Moreover,

if a number x is reached by the symbolic path $L^n R P$, the number $F(x)$ has the symbolic path $L^{n+1} R P$. Hence, the families depicted in Fig.7 can be described as $\{x, F(x), F^{(2)}(x), \dots, F^{(n)}(x), \dots\}$. With a little abuse of notation, we can then write $F(T_n) = T_{n+1}$, i.e. the subtree T_n is completely mapped into T_{n+1} by the action of F [29].

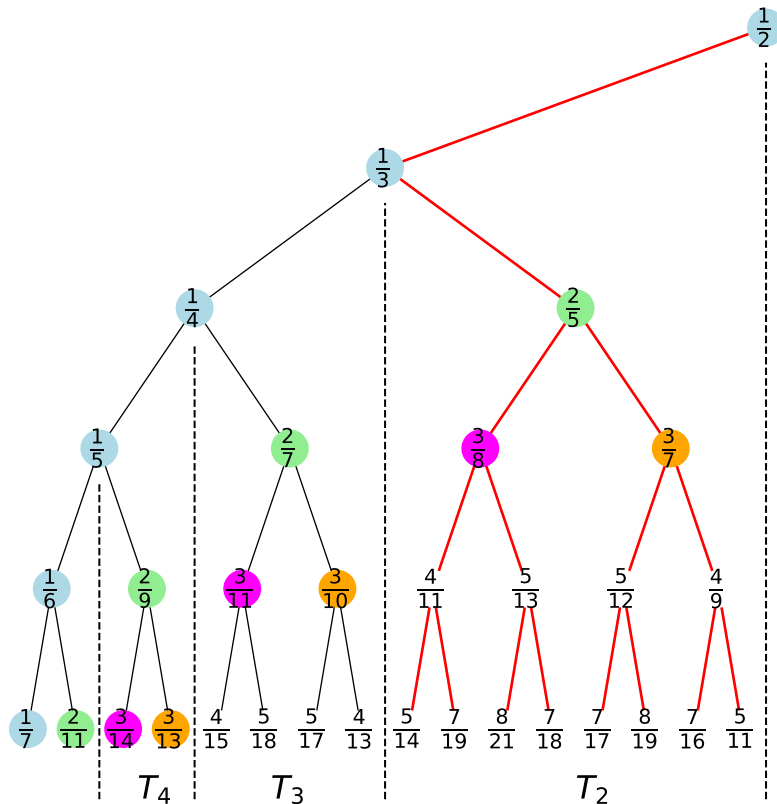


Fig. 7. Farey binary subtree starting at $1/3$. The coloured nodes are the dots in figures 10 and 12. The blue nodes are the extremes of the intervals \mathcal{I}_n . The green nodes are the roots of the subtrees restricted to two consecutive blue nodes. The magenta and orange dots form different rational families: in magenta the symbolic paths $L^n RL$ and in orange the paths $L^n R^2$. Equivalently, in continued fractions, the magenta dots are $[n, 1, 2]$ and the orange dots are $[n, 3]$.

To conclude, let us translate this discussion into Haros graphs. Consider the Haros graph G_x associated with $x \in T_n$. It can be proved that if G_x has α nodes of degree k , then the Haros graph $G_{F(x)}$, associated with $F(x) \in T_{n+1}$, has α nodes of degree $k + 1$. See Appendix VIC for a complete proof of this fact. For instance, the Haros graph $G_{2/5}$, whose sequential degree is $[3, 2, 5, 2, 6]$, has a single node with degree $k = 5$, whereas the Haros graph $G_{F(2/5)} = G_{2/7}$ has a single node with degree $k + 1 = 6$ (see also Fig. 8 for a visual proof for the three first elements of the families depicted in green, magenta and orange). Altogether, these facts provide the following scaling equation:

$$P(k, x) = (1 + x) \cdot P\left(k + 1, \frac{x}{1 + x}\right). \quad (6)$$

In general, $\forall m \geq 1$:

$$P(k, x) = (1 + mx) \cdot P\left(k + m, \frac{x}{1 + mx}\right). \quad (7)$$

The scaling property reported in Eq.7 will be exploited later when we explore the structure of the degree distribution entropy.



Fig. 8. First elements of the Haros graph families $G_{2/(2n+1)}$, $G_{3/(3n+2)}$ and $G_{3/(3n+1)}$, $n = 2, 3, 4$. The graphs of each family are reached by symbolic paths of the shape $L^n R^p$. The connectivity structure is similar in each family: the number of nodes with degrees $k \geq 5$ is preserved, even though the degrees change. For example, the family $G_{2/(2n+1)}$ has two degrees $k = n + 3$ and $k = n + 4$.

3. On the continuity of $P(k, x)$

To conclude our study of the behaviour of $P(k, x)$, we focus on the continuity over the variable x , when k is fixed. Consider again Figure 5, where red dots illustrate how degree $k = 5$ ‘emerges’ along the tree. Although $P(5, x)$ follows Eq.3, nothing is yet said about the boundaries $x = 1/3, 1/2, 2/3$. First, notice in Fig. 6 that at $x = 1/3$, we have $P(5, 1/3) = 1/3$. Now, $G_{x \rightarrow 1/3^-}$ is reached asymptotically through the paths $L^3 R^n$, where the degree $k = 5$ does not appear by construction, and thus $\lim_{x \rightarrow 1/3^-} P(5, x) = 0$. Analogously, following the paths $L^2 R L^n$ we have $\lim_{x \rightarrow 1/3^+} P(5, x) = 0$. We conclude that $P(5, 1/3)$ constitutes a removable discontinuity. A similar phenomenon occurs at $x = 1/2$ and $x = 2/3$. A similar reasoning demonstrates the continuity of the degree distribution close to discontinuities in $x = \xi$: taking the sequence starting at $G_{p/q}$ and descending LR^n or RL^n , we have $\lim_{x \rightarrow \xi^-} P(k, \xi) = \lim_{x \rightarrow \xi^+} P(k, \xi)$. In general, $P(k, x)$ has removable discontinuities at the rational numbers located at the level of the Farey binary tree where degree k first appears (ℓ_{k-3}) and at the level immediately above (ℓ_{k-2}), and is continuous otherwise.

C. Arithmetic and geometric mean degree: Khinchin constant

Once we have established the properties of $P(k, x)$, in this section we further study two additional aspects of the degree sequence: its arithmetic mean degree $\bar{k}(x)$ and the geometric mean degree $\bar{k}_g(x)$. For a G_x where k_i correspond to the degree of the node i we have:

$$\bar{k}(x) = \langle k \rangle = \sum_{k=2}^{\infty} k P(k, x) = \sum_{i=1}^q k_i / q$$

$$\bar{k}_g(x) = \exp(\langle \ln(k) \rangle) = \exp\left(\sum_{k=2}^{\infty} \ln k P(k, x)\right) = \left(\prod_{i=1}^q k_i\right)^{1/q}.$$

In his introduction to continued fractions, A. Ya. Khinchin [8] proves that, for almost all real numbers, the set of coefficients $\{a_i\}$ of the continued fraction expansion of $x = [a_1, a_2, \dots]$ have a finite geometric mean independent of the value of x . Such a constant was coined as the Khinchin constant:

$$K_0 = \prod_{r=1}^{\infty} \left(1 + \frac{1}{r(r+2)} \right)^{\log_2 r} = 2.68545\dots$$

While this is true for almost all numbers, not much is known about specific numbers fulfilling the theorem, other than the fact that both rational and quadratic irrational numbers (with infinite periodic continued fractions) are sets of null measure that do not verify the result. On the other hand, it is also well known that the arithmetic mean of a number's continued fraction expansion diverges for almost all numbers in the unit interval.

Here we study whether the structure of Haros graphs G_x inherits such universality by interpreting the graph degree sequence (k_1, k_2, \dots) as the graph analogue of the coefficients of the continued fraction expansion $[a_1, a_2, \dots]$ of x . In the

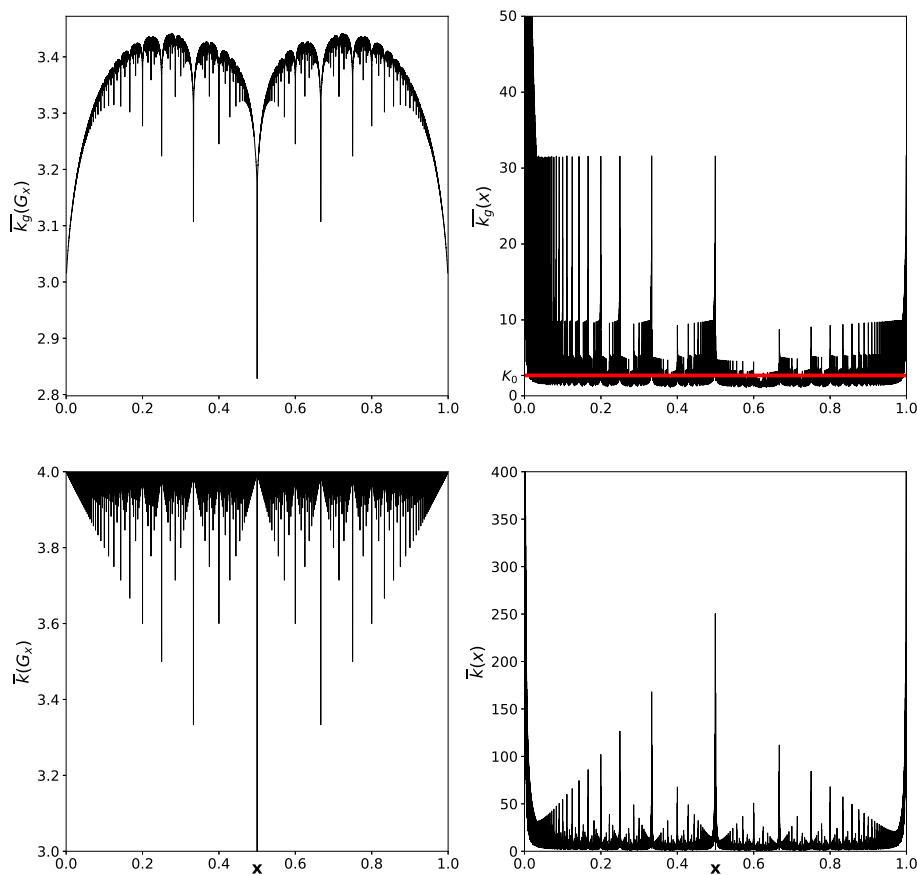


Fig. 9. (Upper panels) Geometric mean degree function $\bar{k}_g(G_x)$ computed for all Haros graphs G_x with $x \in \mathcal{F}_{1000}$ (left) and geometric mean $\bar{k}_g(x)$ for continued fraction terms $x = [a_1, \dots, a_n]$, $\forall x \in \mathcal{F}_{1000}$ (right). The upper right panel is unbounded when $x = 1/n \rightarrow 0$. The horizontal red line denotes Khinchin's constant K_0 . (Bottom panels) Arithmetic mean degree of G_x (left) and $x = [a_1, \dots, a_n]$ (right), for the same values of x . The asymmetric behaviour of the right panels occurs because the continued fraction expansions are different for elements x and $1-x$.

top panels of Fig. 9 we plot the geometric mean of the continued fraction expansion of x for all $x \in \mathcal{F}_{1000}$ (right panel) and $\bar{k}_g(G_x)$, the geometric mean of the degree sequence Haros graph G_x , for the same set (left panel). Notice that the former seems to fluctuate around K_0 , represented by the horizontal red line in the panel, which is an indication that, while in rigour \mathcal{F}_{1000} does not contain irrationals (and thus K_0 is not supposed to show up), the geometric mean

of approximants to rational numbers seems to converge to K_0 . On the other hand, the structure of such convergence is not clear, and there are no other noticeable patterns. The results for the Haros graphs reveal in turn a self-affine shape, similar to a Takagi curve [30]. This curve does not seem to concentrate its measure close to any particular constant, but shows a richer structure.

In the bottom panels of Fig. 9, we then plot the arithmetic mean of the continued fraction expansion of x for all $x \in \mathcal{F}_{1000}$ (right panel), and $\bar{k}(G_x)$, the arithmetic mean of the degree sequence of the Haros graph G_x , for the same set (left panel). The former does not show the exact mirror symmetry of the latter as the continued fraction expansion of x and $1-x$ are different. Furthermore, the arithmetic mean of continued fraction expansion is known to diverge for almost all real numbers, however in the case of Haros graphs we find that their arithmetic mean accumulates at $\bar{k} = 4$, with self-affine fluctuations. To make sense of this finding, we resort to a result in the HVG theory [18, 31] that states that an aperiodic HVG with q nodes –and thus an Haros graph $G_{p/q}$ – has an arithmetic mean degree $\bar{k} = 4 - 2/q$. Since for irrationals $q \rightarrow \infty$, we then have

$$\bar{k}(G_x) = \begin{cases} 4 - \frac{2}{q}, & \text{if } x = p/q \text{ a reduced fraction} \\ 4, & \text{if } x \text{ irrational,} \end{cases} \quad (8)$$

i.e. a Thomae’s function that reflects the fractal structure of the partitions of the interval $[0, 1]$ that generate the successive Farey sequences [32].

To conclude, we found that the universality of Khinchin’s constant is not found in Haros graphs –a self-affine structure with no clear concentration of the measure is found instead, pointing to a richer structure. Such universality is, however, found in the arithmetic mean, which concentrates at $\bar{k} = 4$, a quantity that allows us to discriminate between rational and irrational numbers but fails to further distinguish e.g. between quadratic and non-quadratic irrationals as K_0 does. In the next section, we proceed to construct yet another property of the degree sequence which we will show to possess a richer structure capable of classifying different families of rationals and irrationals.

IV. ENTROPY

It is suggestive to explore the topological properties of the Haros graphs in relation to the real numbers to which they correspond. The previous section suggests that the structure of G_x –in particular, its degree sequence– is inheriting the properties of the number x to which it corresponds, i.e., the properties associated with the path followed in the Farey binary tree. To further explore in more detail such structure, we propose to interpret an Haros graph’s G_x degree sequence as a ‘signal’, and study its informational content by computing the entropy $S(x)$:

$$S(x) = - \sum_{k \geq 2} P(k, x) \log P(k, x). \quad (9)$$

In information-theoretic terms, $S(x)$ is the block-1 approximant to the Shannon entropy of the signal [33]. In graph-theoretic terms, it quantifies the heterogeneity of the graph degree distribution, that is, it measures the amount of disorder contained in the wiring architecture of G_x . Our contention is that the rich and intertwined structure of rational and irrational numbers can be unveiled by studying the structure of $S(x)$.

To give a flavour of the shape of $S(x)$, in Fig. 10 we show the results from a numerical computation for all elements of the Farey sequence \mathcal{F}_{1000} , that includes all irreducible fractions p/q with $q \geq 1000$. As $P(k, x) = P(k, 1-x)$, the curve is symmetric around $x = 1/2$ and $S(x) = S(1-x)$. At first sight, the entropy curve seems self-affine. For instance, $S(x)$ over $(1/4, 1/3]$ seems to be a rescaled copy of the function at $(1/3, 1/2]$ and, in general, every interval $\mathcal{I}_n = \left(\frac{1}{n+1}, \frac{1}{n}\right]$, $n = 2, 3, \dots$ shows a similar shape. A numerical estimation of its box-counting dimension [34] suggests $D_0 \approx 1.43$. In the following subsections, we dig deeper into the shape of $S(x)$ and aim to relate the structure with the properties of x .

A. Continuity of $S(x)$

Visually, it is not clear whether $S(x)$ is continuous on $[0, 1]$. Observe that the arithmetic mean of the degree sequence is a Thomae’s function (Eq. 8) and thus discontinuous at all rationals. This, together with the fact that $P(k, x)$ shows discontinuities, might give the false impression that $S(x)$ might also be discontinuous at all rational numbers. However, a subtler analysis shows that this is not the case. The main reason is based on the fact that

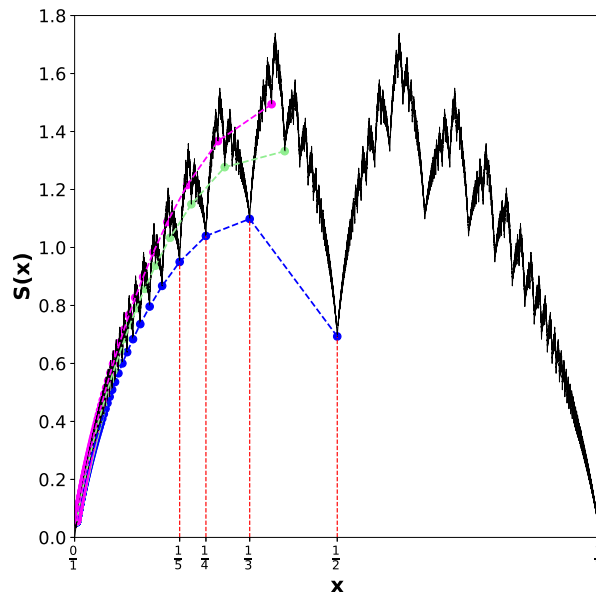


Fig. 10. Entropy function $S(x)$, computed for all Haros graphs G_x with $x \in \mathcal{F}_{1000}$. Blue dots highlight the values $S(1/n)$ and define the intervals $\mathcal{I}_n = \left(\frac{1}{n+1}, \frac{1}{n}\right]$ for $n \geq 2$ (vertical red lines). $S(x)$ appears to have self-affine behaviour in every interval \mathcal{I}_n . Two examples of different families of rational numbers are represented in which each element appears in a single subinterval \mathcal{I}_n , for $n \geq 2$: green dots are the fractions $\frac{2}{2n+1}$, and magenta dots are the fractions $\frac{3}{3n+2}$.

the discontinuities found in $P(k, x)$ are systematically compensated for by the values of $P(k+1, x)$. To showcase this, consider again Fig. 4: for $k = 4$ and close to $x = 1/2$ we have a discontinuity, where $P(4, 1/2) = 1/2$ and $\lim_{x \rightarrow 1/2^-} P(4, x) = \lim_{x \rightarrow 1/2^+} P(4, x) = 0$. This discontinuity is indeed ‘balanced-out’ with the discontinuity at $k = 5$ and close to $x = 1/2$, where we have $P(5, 1/2) = 0$ and $\lim_{x \rightarrow 1/2^-} P(5, x) = \lim_{x \rightarrow 1/2^+} P(5, x) = 1/2$ (see Fig. 6). Accordingly, $P(4, x) + P(5, x)$ is continuous at $x = 1/2$, and so is $P(4, x) \log P(4, x) + P(5, x) \log P(5, x)$ (where we use the convention $0 \log 0 = 0$). Similarly, close to any x that shows a removable discontinuity in $P(k, x)$, such a discontinuity is exactly balanced with one at $P(k+1, x)$. When summing over k in the computation of $S(x)$ these discontinuities are systematically removed, and the resulting function $S(x)$ is continuous.

B. Disentangling Entropy: from rational Haros graphs being families of local minima to a generalised de Rham curve

In this subsection, we shall produce two results: first, by exploring families of Haros graphs emerging naturally in the entropy function, we unveil a hierarchical structure of local minima associated with a partition of rational numbers into different families. Leveraging on some of the patterns that will emerge in the first part, the second objective of this subsection is to obtain a functional equation for the entropy function, which we will show takes the form of a generalised de Rham curve, i.e., certifying that $S(x)$ is indeed a self-affine function, continuous but not differentiable at any point.

Let us start by considering rationals of the form $x = 1/n$, $n \in \mathbb{N}^+$. In Fig.11 we show the first Haros graphs $G_{1/n}$ for $n = 2, 3, 4$, and their respective symbolic paths in the Farey binary tree. Their degree distributions $P(k, 1/n)$ have the following expression for $n \geq 2$:

$$P(k, 1/n) = \begin{cases} 1/n & , k = 2 \\ 1 - 2/n & , k = 3 \\ 1/n & , k = n + 2 \\ 0 & , \text{otherwise.} \end{cases}$$

Accordingly, their graph entropy values are: $S(1/n) = -2/n \cdot \log(1/n) - (1 - 2/n) \cdot \log(1 - 2/n)$, and these values appear as local minima (blue dots in Fig.10). Taking advantage of the analytical expression, we can then define a

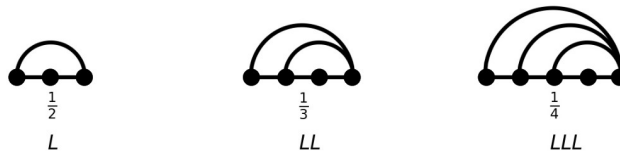


Fig. 11. The first three elements of the Haros graph family $G_{1/n}$. These Haros graphs, represented in blue in figures 10, 12 and 7, are reached by the symbolic paths L^{n-1} . Each element of this family maintains the connectivity structure $[3, \dots, 3, 2, n+2]$, where the number of nodes with degree $k=3$ increases as n grows.

reduced graph entropy $H(x)$ as follows:

$$H(x) \equiv \begin{cases} S(x) + 2x \cdot \log(x) + (1-2x) \cdot \log(1-2x) & , x \in [0, 1/2] \\ S(x) + 2(1-x) \cdot \log(1-x) + (2x-1) \cdot \log(2x-1) & , x \in [1/2, 1]. \end{cases} \quad (10)$$

The reduced entropy preserves the original symmetry $H(x) = H(1-x)$, and is identically null for all rationals of the form $x = 1/n, n > 1$: see Fig.12 (left panel) for a plot of $H(x)$, where $H(1/n)$ is highlighted as a dashed blue line connecting all rationals of the form $1/n$. In the same plot, other sets of minima of the reduced graph entropy seem to emerge naturally and connected by straight lines of different slopes, e.g. $G_{2/(2n+1)}$ (green dots connected by a green dashed line), $G_{3/(3n+2)}$ (magenta dots connected by a magenta dashed line), or $G_{3/(3n+1)}$ (orange dots connected by an orange dashed line). The first elements of these respective sets of Haros graphs, along with their respective symbolic sequence, are depicted in Fig. 8. We now show that this is indeed the case and that such families are indeed aligned local minima of $H(x)$. It can indeed be proved that, in the reduced entropy space, every such family can be interpolated by a straight line $H(x) = ax$ with a different slope a (a proof is presented in Appendix VID).

Furthermore, if we now define a reduced entropy density $H(x)/x$, the numbers families are now classified in constant levels of reduced entropy density, since $H(x)/x = H(F^{(m)}(x))/F^{(m)}(x)$, with $x \in \mathcal{I}_2$ and $F^{(m)}(x) \in \mathcal{I}_{m+2}$ (see the right panel of Fig. 12 to visually verify this fact). Accordingly and with Eq. 7, we have that the reduced entropy fulfils a scaling equation:

$$H(x) = (1+mx) \cdot H\left(\frac{x}{1+mx}\right). \quad (11)$$

Note however that this scaling holds for $x < 1/2$, and the scaling is indeed different if $x > 1/2$ as shown in Appendix VIC. If we now apply the change of variable $z = \frac{1}{x} - 2$, in Eq. 11, the interval $[0, 1/2]$ is deformed into $[0, \infty)$, and every rational number $x = 1/n$ is transformed into a natural number $z = n - 2$ for $n \geq 2$. Thus, the intervals \mathcal{I}_n deform into intervals of equal size $[n-2, n-1)$, $H(z)/z$ becomes periodic (see Fig. 13) and we can write for $z \in [0, 1)$ and $m = 0, 1, 2, \dots$:

$$(z+2) \cdot H\left(\frac{1}{z+2}\right) = (z+m+2) \cdot H\left(\frac{1}{z+m+2}\right). \quad (12)$$

Moreover, from Fig. 14, we visually observe that $(z+2)H(1/(z+2))$ is exactly the graph entropy defined in Eq. 9 for $z \in [1/2, 1]$:

$$S(z) = (z+2) \cdot H\left(\frac{1}{z+2}\right). \quad (13)$$

or, equivalently, by the definition of reduced entropy 10:

$$S(z) = (z+2) \cdot S\left(\frac{1}{z+2}\right) + z \log z - (z+2) \log(z+2), \text{ for } z \in [1/2, 1]. \quad (14)$$

A rigorous proof of this scaling equation is given in the proof of Theorem 16 in Appendix VIC.

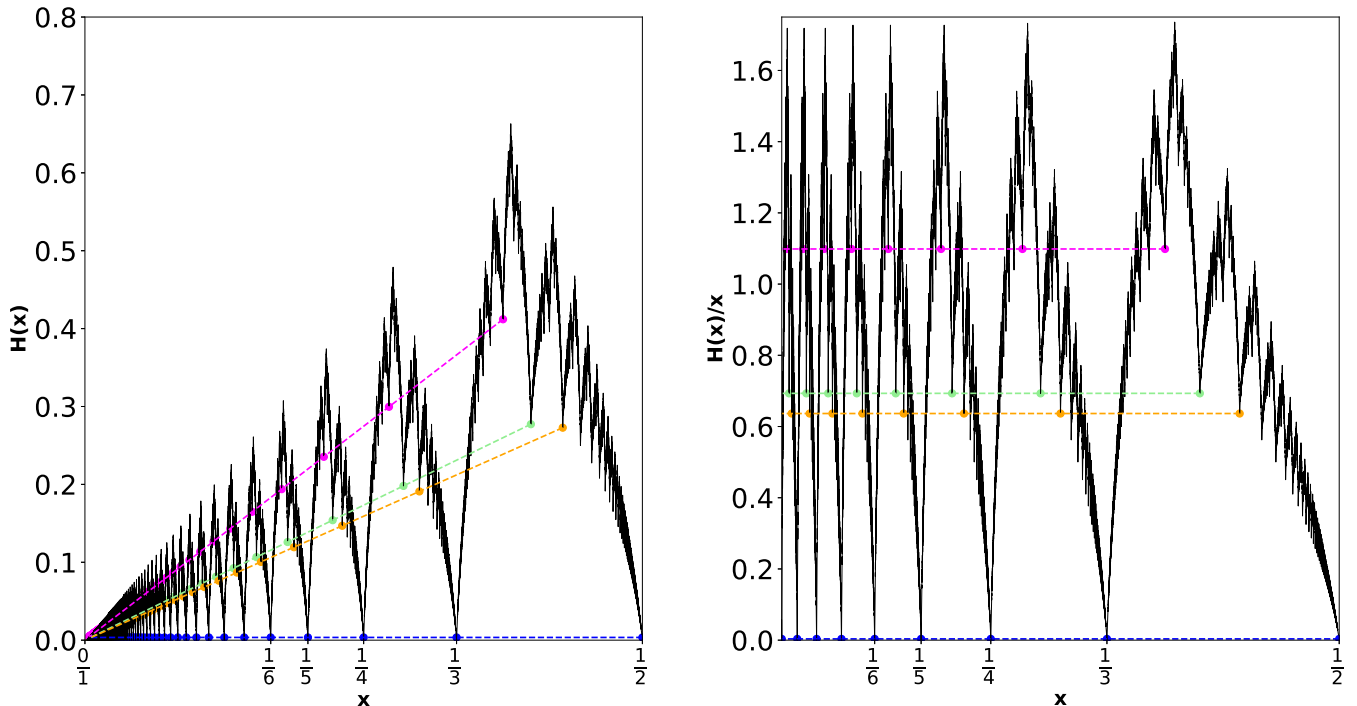


Fig. 12. Reduced entropy function $H(x) = S(x) + 2x \cdot \log(x) + (1 - 2x) \cdot \log(1 - 2x)$ for $x \leq 1/2$ (left panel) and reduced entropy density $H(x)/x$ for the same values (right panel). In both panels, different families of rational numbers are highlighted. The blue dots are the fractions $1/n$ reached by the paths L^{n-1} ; blue dashed line has a null slope in both panels. The green dots are the fractions $\frac{2}{2n+1}$ reached by paths $L^n R$; in the left panel, the slope of the green line is $\log(2)$. Magenta dots are the fractions $\frac{3}{3n+2}$ reached by paths $L^n RL$; in the left panel, the slope of the magenta line is $\log(3)$. The orange dots are the fractions $\frac{3}{3n+1}$ reached by paths $L^n R^2$; in the left panel, the slope of the orange line is $\log(3) - \frac{2}{3} \cdot \log(2)$. In the right panel, these three families form different constant levels.

This equation has the form of a generalised Rham curve [30], where the graph of $S\left(\frac{1}{z+2}\right)$ in $[1/3, 2/5]$ is a self-affine copy of $S(z)$ in $[1/2, 1]$ scaled by a factor $(z+2)$ and displaced $z \log z - (z+2) \log(z+2)$. In Fig. 14, we numerically illustrate this self-affinity. Observe that in the last step, the interval $[1/3, 2/5]$ is the interval $[1/2, 1]$ transformed by $1/(z+2)$. As a family of de Rham curves, $S(x)$ is indeed a continuous function with no derivative at any point.

To conclude this section, we observe how Figure 10 supports visually the idea that the rational numbers form envelope curves under the graph entropy function. We sketch a proof of local minima as follows: Clearly we have that the global minima are reached in $x = 0, 1$. The local minima in a given interval will be reached at a point obtained as a finite composition of affine transformations $\frac{1}{z+2}$, where z are rational numbers, so the local minima will be rational numbers.

C. Entropy maxima: from the Golden mean to noble numbers

As we showed in the previous section, the entropy function reaches local minima at every rational number, and such minima are organised into families. Since the global minimum value of $S(x)$ is trivially reached for $x = 0/1, 1/1$, where $S(0) = S(1) = 0$, the analysis of minima is concluded and now we turn to explore the structure of the maxima of $S(x)$.

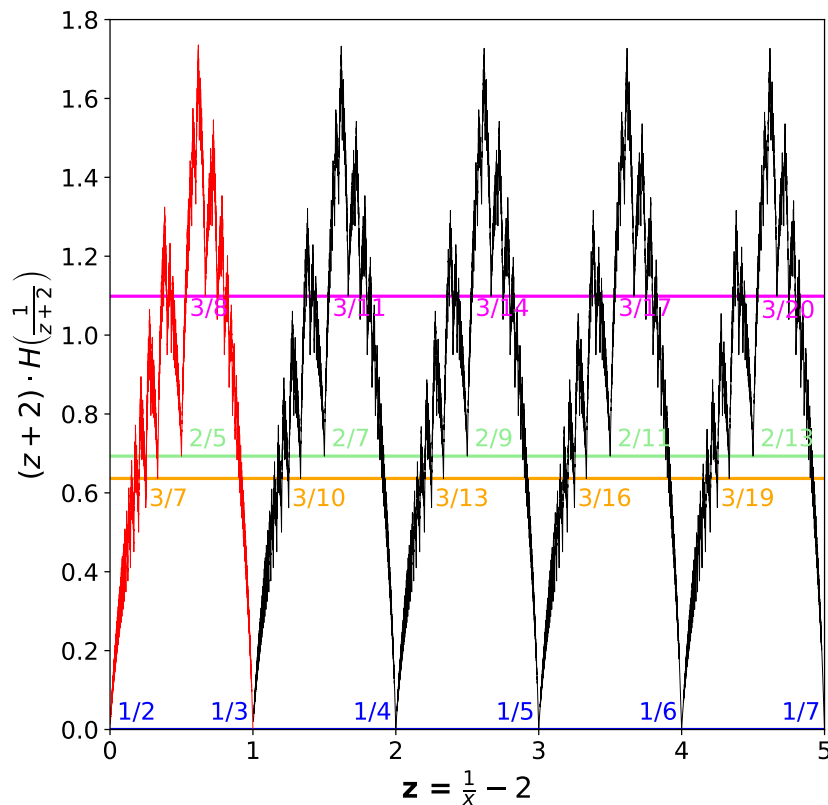


Fig. 13. Plot of $(z+2) \cdot H(1/(z+2))$. The change of variable $z = 1/x - 2$ transforms the interval $[0, 1/2]$ into $[0, \infty)$. Furthermore, the fractions $1/n$ are equispaced and we have a periodic pattern in the intervals $[n, n+1]$. The interval \mathcal{I}_2 is highlighted in red in correspondence with the red-coloured subtree in Fig. 7.

1. Global maximum: ϕ^{-1}

Here we make use of technique of Lagrange multipliers to prove that $S(x)$ reaches its global maximum at $x = \phi^{-1} = 0.618\dots$, and by symmetry at $x' = 1 - \phi^{-1}$, where $\phi = (1 + \sqrt{5})/2 = 1.618033\dots$ is the so-called Golden number. To fix the constraints of the Lagrange multipliers technique, we consider again the degree distribution $P(k, x)$. Since all Haros graphs in $(1/2, 1)$ fulfil $P(2, x) = x$, $P(3, x) = 1 - 2x$ and $P(4, x) = 0$ (Eq.2 and the mirror symmetry $P(k, x) = P(k, 1 - x)$), normalisation implies $\mathcal{Q}_0 = \sum_{k=5}^{\infty} P(k, x) = 1 - x$. Second, observe that $S(x)$ reaches minima at the rationals, and thus the global maximum should take place at an irrational number. According to Eq.8, all irrational Haros graphs have arithmetic mean degree $\bar{k} = 4$, hence the second constraint is $\mathcal{Q}_1 = \sum_{k=5}^{\infty} k \cdot P(k, x) = 5 - 4x$. Altogether, for irrational numbers, the following Lagrangian functional can be defined:

$$\mathcal{L}[\{P(k, x)\}] = - \sum_{k=5}^{\infty} P(k, x) \log P(k, x) - (\lambda_0 - 1) \left(\sum_{k=5}^{\infty} P(k, x) - \mathcal{Q}_0 \right) - \lambda_1 \left(\sum_{k=5}^{\infty} k P(k, x) - \mathcal{Q}_1 \right),$$

which will take an extreme in the solution $\delta\mathcal{L}[P^*(k, x)] = 0$ that also extremizes the entropy $S(x)$. One can prove that such extrema are reached for a specific shape of the degree distribution $P(k, x)$ (see Appendix VI E for details):

$$P(k, x) = \begin{cases} 1 - x & k = 2 \\ 2x - 1 & k = 3 \\ 0 & k = 4 \\ x^{k-1} & k \geq 5. \end{cases} \quad (15)$$

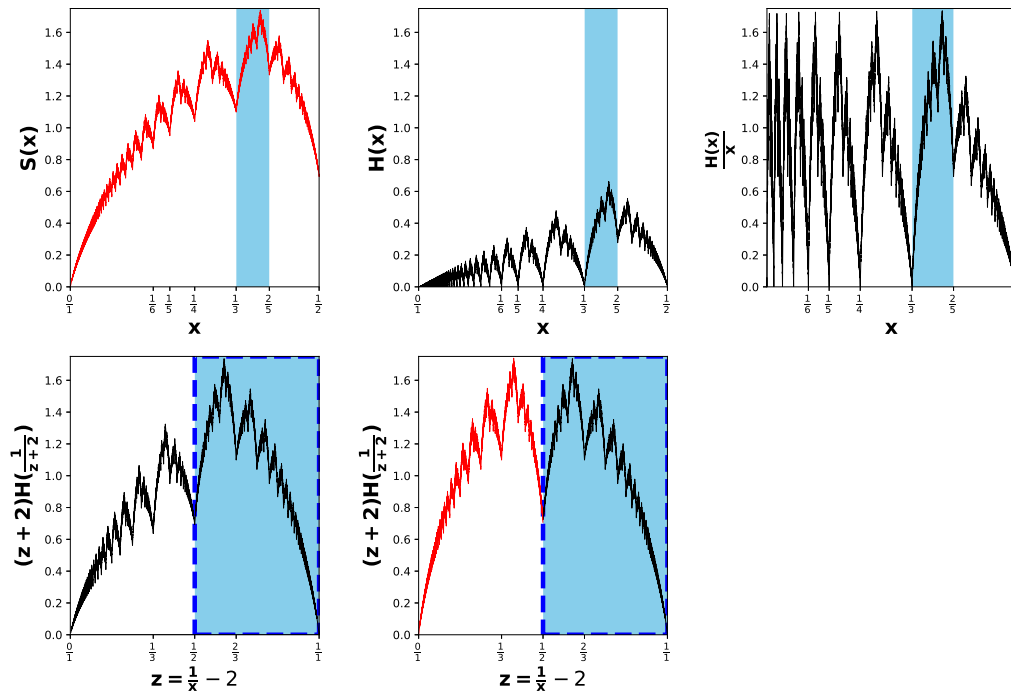


Fig. 14. Compositions of functions to visually describe the self-affine structure of $S(x)$. The blue shaded zone is the interval $[\frac{1}{3}, \frac{2}{5}]$. (Upper, left panel) Graph entropy $S(x)$. (Upper, middle panel) The transformation $H(x) = S(x) + 2x \cdot \log(x) + (1 - 2x) \cdot \log(1 - 2x)$ removes the entropy of rational numbers $1/n$. (Upper, right panel) Transformation $H(x)/x$. (Bottom, left panel) Change of variable $z = 1/x - 2$. (Bottom, middle panel) The resulting function matches the original function $S(x)$ over $[1/2, 1]$.

Note that this technique does not determine the value of x for us, only the shape of the degree distribution. But according to Eq.15, whatever x is, G_x will necessarily contain nodes spanning all possible degrees (except $k = 4$), i.e., the degree distribution does not have any holes. According to the analysis in Section III, this implies that the symbolic path that identifies G_x in the binary tree cannot have adjacent repeated symbols, i.e. it needs to follow an infinite zigzag, hence $x = \phi^{-1}$, the reciprocal of the Golden number, and by symmetry $1 - x = 1 - \phi^{-1}$.

We can also independently determine $P(k, \phi^{-1})$ from the successive degree distributions of Haros graphs $G_{\mathfrak{F}_{n-1}/\mathfrak{F}_n}$ associated with rational convergents $\mathfrak{F}_{n-1}/\mathfrak{F}_n$ of ϕ^{-1} obtained via the Fibonacci sequence (\mathfrak{F}_n) , see Fig. 15. From a previous work [20], we have:

$$P\left(k, \frac{\mathfrak{F}_{n-1}}{\mathfrak{F}_n}\right) = \begin{cases} 1 - \frac{\mathfrak{F}_{n-1}}{\mathfrak{F}_n} = \frac{\mathfrak{F}_{n-2}}{\mathfrak{F}_n} & k = 2 \\ 2 \cdot \frac{\mathfrak{F}_{n-1}}{\mathfrak{F}_n} - 1 = \frac{\mathfrak{F}_{n-3}}{\mathfrak{F}_n} & k = 3 \\ 0 & k = 4 \\ \frac{\mathfrak{F}_{n+k-1}}{\mathfrak{F}_n} & 5 \leq k \leq n + 1. \end{cases} \quad (16)$$

Now, using Binet's formula and taking the limit $\lim_{n \rightarrow \infty} \mathfrak{F}_{n-1}/\mathfrak{F}_n = \phi^{-1}$, we recover Eq.15 with $x = \phi^{-1}$.

If additional restrictions are imposed in the degree distribution, which Haros graphs will be maximally entropic accordingly? In what follows, we show that further restrictions can be parametrised by holes in the degree distribution, i.e. $P(\kappa, x) = 0$ for $4 \leq \kappa \leq f(n)$ (where n denotes an ordinal). We will prove that if the restrictions are a finite number of zeros of $P(k, x)$, the emergent entropy local maxima coincide with a family of affine transformations of the Golden number.

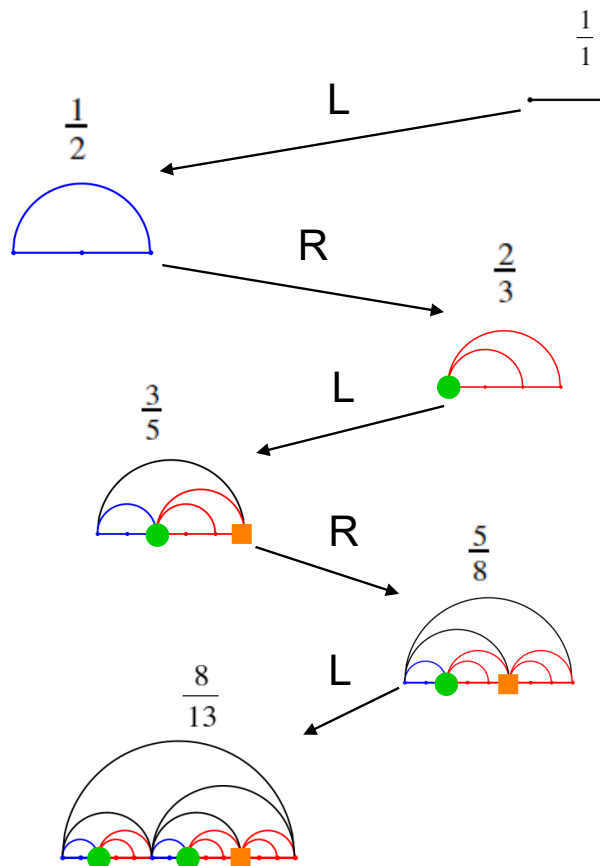


Fig. 15. We depict the first six steps of the path of the Haros graph tree to $G_{\phi^{-1}}$. The rational numbers $1/1, 1/2, 2/3, 3/5, 5/8$ and $8/13$ are quotients $\frac{\tilde{\delta}_{n-1}}{\tilde{\delta}_n}$ of Fibonacci elements. The green points highlight the nodes with degree $k = 5$ and illustrate how the zigzag descent increase the degree frequency following a Fibonacci sequence, whereas the orange square points remark the nodes with degree $k = 6$. Due to lack of space, only six of these convergents are shown. However, it is easy to verify that the next convergent $8/13 \oplus 5/8 = 8/13 = 13/21$ has 3 green nodes and 2 orange square node, and, in general, the frequency of green and orange square nodes follow a Fibonacci sequence.

2. Local maxima

We have proved that the entropy function $S(x)$ finds its global maximum at $x = \phi^{-1}$ (and by symmetry at $1 - x = 1 - \phi^{-1} = 1/(2 + \phi^{-1})$). Here, we further explore the hierarchy of local maxima that emerge in the rich structure of $S(x)$. For simplicity, from now on we focus only on $x \in [0, 1/2]$, since the results are then extended by symmetry to $[0, 1]$. Now, let us again consider the graphic representation of $S(x)$, see Fig.16. For $x \in [0, 1/2]$, one can extract a set of local maxima's positions \mathcal{M}_1 that can be connected by a certain envelope curve $\mathcal{C}_1(n)$, highlighted by red dots connected by a red dashed line in 16). Numerically, \mathcal{M}_1 appears to coincide with the set of numbers

$$\mathcal{M}_1 = \left\{ \frac{1}{2 + \phi^{-1}}, \frac{1}{3 + \phi^{-1}}, \frac{1}{4 + \phi^{-1}}, \frac{1}{5 + \phi^{-1}}, \dots \right\},$$

whose elements are naturally ordered by the discrete function $\mathcal{C}_1(n)$:

$$\mathcal{C}_1(n) = \frac{1}{n + \phi^{-1}}, \quad n \geq 2.$$

First, we study in some detail the elements of \mathcal{M}_1 . Elements in this set are a subset of the so-called *noble numbers*: irrationals whose continued fraction expansion eventually reaches an infinite sequence of 1's: $[a_1, a_2, \dots, a_m, 1, 1, 1, \dots]$. We denote the initial sequence of symbols $[a_1, a_2, \dots, a_m]$ as the transient sequence. The reciprocal of the Golden number ϕ^{-1} is a very particular case of a noble number with no transient sequence ($m = 0$). In particular, the elements in \mathcal{M}_1 (and ordered by $\mathcal{C}_1(n)$) have a transient sequence $[a_1]$, with $m = 1$ and $a_1 = n$.

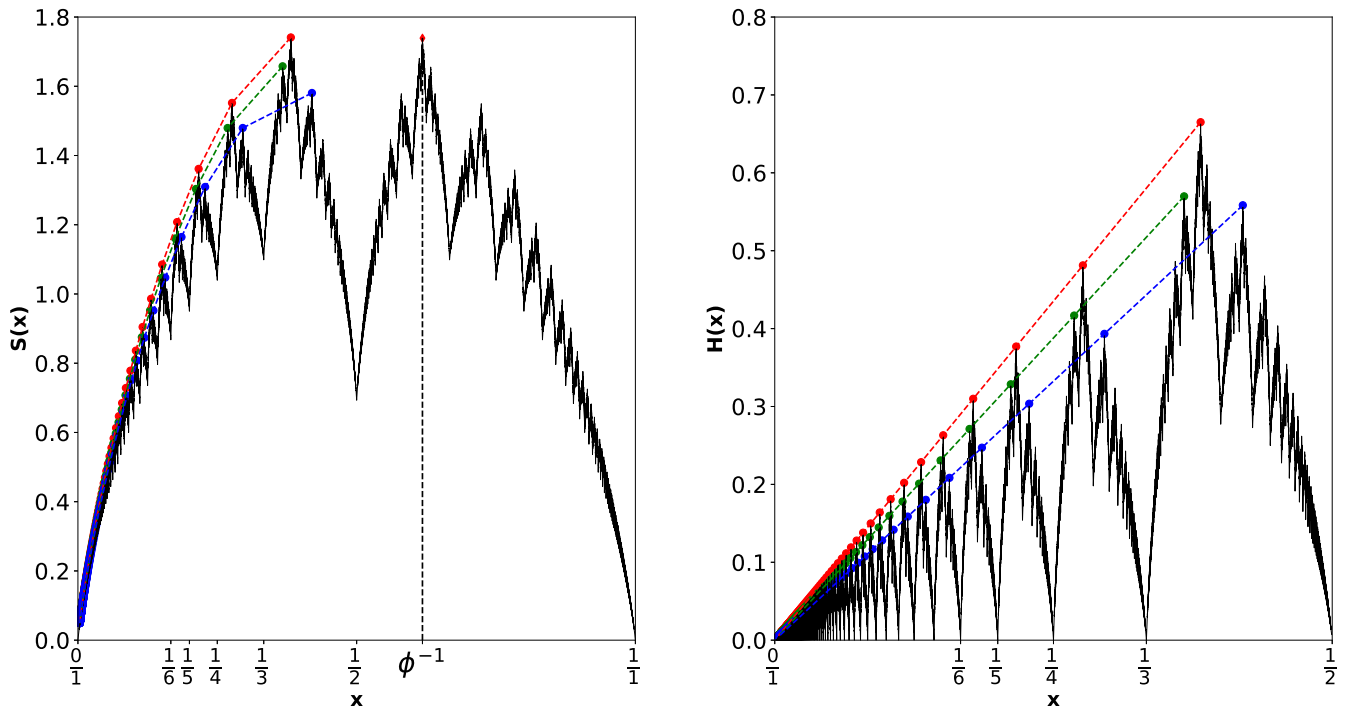


Fig. 16. Entropy $S(x)$ (left panel) and reduced entropy $H(x)$ (right panel), computed for all Haros graphs G_x with $x \in \mathcal{F}_{1000}$. These functions have been shown in Fig. 10 12, but in this case, they have been pointed out families of noble numbers. In both panels, the red dots correspond to the local maxima $\mathcal{C}_1(n) = \frac{1}{n+\phi-1}$ for $n \geq 2$. The global maxima are $\mathcal{C}_1(2) = 1 - \phi^{-1}$ and its mirror symmetry ϕ^{-1} . Green and blue dots represent other families of noble numbers. The right panel helps to visualise how these families display a linearly increasing reduced entropy with a slope that depends on the family and where each element appears only once in each interval \mathcal{I}_n .

The noble numbers can also be coded as an infinite path in the Farey binary tree, which, after a transient sequence of L 's and R 's, eventually reaches an infinite zigzag $(LR)^\infty$. For example, $\mathcal{C}_1(3)$ has a continued fraction expansion $[3, \bar{1}]$, i.e., its transient sequence therefore consists only of the number 3, and its symbolic path is $L^2(LR)^\infty$. The convergents in the transient are thus $0/1$ and $1/3$, and once the infinite zigzag sets in, the path continues and leads to the successive approximants $1/4, 2/7, 3/11, 5/18, \dots$. Observe that this sequence of approximants can also be generated using suitable combinations of Fibonacci numbers, e.g. for this case

$$\left(\frac{1}{4}, \frac{2}{7}, \frac{3}{11}, \frac{5}{18}, \dots \right) = \left(\frac{\mathfrak{F}_n}{\mathfrak{F}_{n-1} + 3\mathfrak{F}_n} \right)_{n \geq 1}, \quad (17)$$

where the right-hand side of Eq.17 is just the particular case of the general expression

$$\left(\frac{A\mathfrak{F}_{n-1} + B\mathfrak{F}_n}{C\mathfrak{F}_{n-1} + D\mathfrak{F}_n} \right)_{n \geq 1}, \quad (18)$$

where A/C and B/D are the last two convergents of the transient sequence, i.e. $A = 0$, $C = 1$, $B = 1$, $D = 3$ in our example $x = \mathcal{C}_1(3)$. Taking the limit $n \rightarrow \infty$ in the rhs of Eq.17, we correctly recover the irrational number $\mathcal{C}_1(3) = \frac{1}{3+\phi^{-1}}$. Likewise, taking the limits in Eq.18 provides an expression for all noble numbers as

$$\frac{A\phi^{-1} + B}{C\phi^{-1} + D},$$

where it is easy to see that for a general $\mathcal{C}_1(n)$, the convergents are $0/1$ and $1/n$ i.e. $A = 0$, $B = 1$, $C = 1$, $D = n$.

Let us consider then the topological structure of Haros graphs $G_{\mathcal{C}_1(n)}$. Since noble numbers $\mathcal{C}_1(n)$ have paths in the Farey binary tree that start with a sequence of n consecutive L 's, by theorem 2, it turns out that degrees $4 \leq k \leq n+2$

do not appear in $G_{\mathcal{C}_1(n)}$. The degrees that do appear can be calculated using $F(\mathcal{C}_1(n)) = \mathcal{C}_1(n+1)$, the expression of $P(k, \phi^{-1}) = P(k, 1 - \phi^{-1} = \mathcal{C}_1(2))$ and the scaling equation 7. In the same manner, taking as additional the restriction over the interval $[0, 1/n]$ into consideration and using similar techniques as those leading to Eq.15, after a bit of algebra and using Eqs. 17 and 18, it can be proved that

$$P(k, \mathcal{C}_1(n)) = \begin{cases} \mathcal{C}_1(n) & k = 2 \\ 1 - 2\mathcal{C}_1(n) & k = 3 \\ 0 & 4 \leq k \leq 2 + n \\ \mathcal{C}_1(n)(\phi^{-1})^{k-(n+1)} & k \geq 3 + n. \end{cases} \quad (19)$$

Interestingly, the degree distribution for all $\mathcal{C}_1(n)$ has a universal exponential tail with the same exponent, based on the Golden number. It turns out that in the intervals $[0, \frac{1}{n}]$, the graph entropy function $S(x)$ necessarily reaches its maximum precisely at $x = \mathcal{C}_1(n)$ (see Fig. 16 for an illustration and Appendix VIF for a proof). We can also prove that the reduced entropy $H(x)$ linearly interpolates the elements of \mathcal{M}_1 , more concretely:

$$H(\mathcal{C}_1(n)) = (-\log(\phi^{-1}) \cdot (3 + \phi^{-1})) \cdot \mathcal{C}_1(n). \quad (20)$$

Hence, the slope of the red dashed line in Fig. 16 is precisely $-\log(\phi^{-1}) \cdot (3 + \phi^{-1})$ (See Appendix VIG for more details).

One can now define the full set of Noble numbers via continued fractions as:

$$\mathcal{C}_m(n_1, \dots, n_m) = [n_1, n_2, \dots, n_m, \bar{1}],$$

that is, these are numbers with an arbitrary transient, eventually followed by an infinite period-1 pattern (or, in terms of the binary sequence, followed by an infinite zigzag). Noble numbers are those quadratic irrationals which are quotients of affine functions of the Golden number, i.e.

$$\mathcal{C}_m(n_1, \dots, n_m) = \frac{p_m + \phi^{-1}p_{m-1}}{q_m + \phi^{-1}q_{m-1}},$$

where p_i/q_i is the i -th convergent of the rational reached after the transient sequence $p/q = [n_1, \dots, n_m]$. In Fig.16 we highlight two specific families: $\mathcal{C}_3(n, 1, 2)$ (green line) and $\mathcal{C}_2(n, 2)$ (blue line) (we remind that the red line interpolates $\mathcal{C}_1(n)$). For example, it can be proved (the proof is however cumbersome and not reported) that the degree distribution $P(k, x \in \mathcal{C}_3(n, 1, 2))$, with $x = \frac{3+\phi^{-1}}{(3n+2)+(n+1)\cdot\phi^{-1}}$ is:

$$P(k, x) = \begin{cases} x & k = 2 \\ 1 - 2x & k = 3 \\ 0 & 4 \leq k \leq n + 2 \\ \frac{1}{3+\phi^{-1}} \cdot x & k = n + 3 \\ \frac{1+\phi^{-1}}{3+\phi^{-1}} \cdot x & k = n + 4 \\ 0 & k = n + 5 \\ \frac{1}{3+\phi^{-1}} \cdot x \cdot (\phi^{-1})^{k+n-4} & k \geq n + 6. \end{cases} \quad (21)$$

In the same way, with the theoretical degree distribution, we can calculate the reduced entropy of $H(\mathcal{C}_3(n, 1, 2)) = \mathcal{C}_3(n, 1, 2) \cdot \left(\log(3 + \phi^{-1}) - \frac{1}{3+\phi^{-1}} \cdot \log(1 - \phi^{-1}) \right)$. Again, the reduced entropy has the form $H(x) = a \cdot x$, as we have shown visually with the green dashed line in Fig.16.

Interestingly, we can see that while the specific transient of the continued fraction has a specific echo in the first values of the degree distribution, its tail is still exponential with the same base as for the family $\mathcal{C}_1(n)$:

Theorem 3. *Haros graphs G_x where x is an arbitrary noble number $\mathcal{C}_m(n_1, \dots, n_m)$ have a degree distribution with an exponential tail, with the Golden number as a base.*

Here we report a sketch of the proof: The degree distribution $P(k, x = \mathcal{C}_m(n_1, \dots, n_m))$ is constructed in a zigzag process starting at the degree distribution of p/q . After the transient sequence, the infinite zigzag reaches the values:

$$\frac{p + p_{m-1}}{q + q_{m-1}}, \frac{2p + p_{m-1}}{2q + q_{m-1}}, \frac{3p + 2p_{m-1}}{3q + 2q_{m-1}} \rightarrow \frac{p + \phi^{-1}p_{m-1}}{q + \phi^{-1}q_{m-1}},$$

where p_{m-1}/q_{m-1} is the $m - 1$ convergent of p/q . At this point, the generation of new degrees follows the process illustrated by the green and orange points in Fig. 15. In each descent, a new degree appears at the boundary node, which becomes an inner node in the next two descents. After that, its value increases as a Fibonacci sequence: 2, 3, 5, 8, 13, ..., whereas the denominators increase as $q + q_{m-1}, 2q + q_{m-1}, 3q + 2q_{m-1}, \dots$. Hence, the tail of the degree distribution of the noble number will have follow:

$$\frac{\mathfrak{F}_{n-(k-\alpha)}}{\mathfrak{F}_n \cdot q + \mathfrak{F}_{n-1} \cdot q_{m-1}},$$

where the value α depends on the concrete noble number x . Using Binet's formula, from the last expression we easily obtain an exponential tail, where the base is indeed the reciprocal of the Golden number (see Fig. 17).

To recap, we have proved the Haros graphs associated with Noble numbers –a subset of the family of quadratic irrationals– have a degree distribution with an exponential tail of base ϕ^{-1} . This result has been obtained via an entropic-maximization process, where the infinite zigzag period-1 expansion $[\bar{1}]$ of Noble numbers plays an important role. The question arises naturally: What happens if we then consider a periodic patterns of larger period ≥ 2 ? As a matter of fact, the so-called family of Metallic ratios ϕ_b^{-1} are the positive solutions of the equation $x^2 + bx - 1 = 0$, and their continued fraction expression is $[\bar{b}]$, i.e., metallic ratios have a period- b pattern. For example, the reciprocal of the Golden number is the Metallic ratio for $b = 1$, while $\sqrt{2} - 1$ is the so-called Silver ratio, or Metallic ratio for $b = 2$. From [20], the theoretical expression of $P(k, \phi_b^{-1})$ is:

$$P(k, \phi_b^{-1}) = \begin{cases} \phi_b^{-1} & k = 2 \\ 1 - 2\phi_b^{-1} & k = 3 \\ 0 & k = 4 \\ (1 - \phi_b^{-1}) \cdot (\phi_b^{-1})^{\frac{k-3}{b}} & k = bn + 3, n \geq 1 \\ 0 & \text{otherwise.} \end{cases} \quad (22)$$

The proof of Eq.22 is similar to others presented in this paper and proceeds to use Lagrange multiplier techniques with the constraints $P(k, x) = 0$ for values $k \neq bn + 3$, i.e., we search for local maxima establishing an infinite number of periodic restrictions. We conclude that Haros graphs associated with metallic ratios have a degree distribution with a periodic pattern of zeros and an exponential tail with base the metallic ratio ϕ_b^{-1} , and are also local maxima of $S(x)$.

We close this section by further generalizing our previous results: we can define the full set of generalised metallic ratios (note that this family contains all Noble numbers) as

$$\mathcal{C}_{b,m}(n_1, \dots, n_m) = [n_1, n_2, \dots, n_m, \bar{b}],$$

We can state the following theorem (the proof is cumbersome and is not presented, but its essence is the same as that seen in previous theorems):

Theorem 4. *Haros graphs associated with generalised metallic ratios $\mathcal{C}_{b,m}(n_1, \dots, n_m)$ have a degree distribution with a periodic pattern of zeros and an exponential envelope with a base ϕ_b^{-1} , and are thus also local maxima of $S(x)$.*

V. DISCUSSION AND OPEN PROBLEMS

Let us first summarise the paper's highlights. In this work, we have introduced Haros graphs G_x , that provide a graph-theoretic representation of real numbers $x \in [0, 1]$. The structure of the Farey binary tree is replicated using the set of Haros graphs and a concatenation operator that plays the role of the mediant sum operator in graph space. The degree sequence of Haros graphs encapsulates rich information of the number they represent. We have provided analytical results on the degree distribution $P(k, x)$ of the Haros graph G_x associated to $x \in [0, 1]$ and provide insight into such topological structure in relation to the continued fraction representation of x . We have further explored the arithmetic and geometric mean degree of Haros graphs for both rational and irrational numbers, revealing an intricate

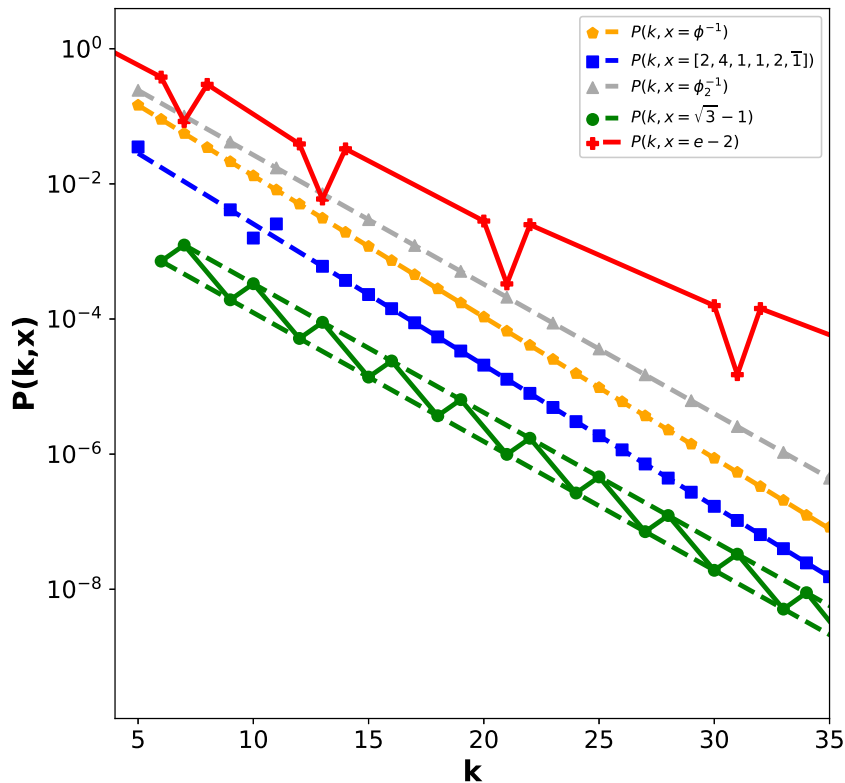


Fig. 17. Semi-log plot of the numerical degree distributions $P(k, x)$ for values $5 \leq k \leq 30$ of Haros graphs associated with irrational numbers of different families: the reciprocal of the Golden number (golden circles), the general noble number $[0; 2, 4, 1, 1, 2, 2, \bar{1}]$ (blue squares), the silver mean or metallic ratio with $b = 2$, $\sqrt{3} - 1 = [\bar{1}, 2]$, a quadratic irrational with continued fraction of period greater than 1, and the number e , a non-algebraic irrational. The dots represent the numerical $P(k, x)$ obtained by the large convergents of each irrational number. The dashed lines correspond to theoretical tails of the degree distribution of each quadratic irrational. In the case of $\sqrt{3} - 1$, the degrees $k_1 = 3n + 1$ and $k_2 = 3n + 2$ form two parallel exponential tails. It can be observed that the tails of the degree distribution of the reciprocal of Golden number and the noble numbers are parallels and their slope matches the the reciprocal of the Golden number.

fractal structure which is absent in the continued fraction expansion.

In a second step, we have defined a graph entropy $S(x)$ in terms of the entropy over the degree distribution of the Haros graph G_x . We have proved that such an entropy is a self-affine function, which can be expressed in terms of a generalised de Rham curve. The local minima of such a function relate to rational numbers, whereas the local maxima are related to specific families of irrationals. We have proved that the subset of noble numbers $\mathcal{C}_1(n)$ are local maxima of $S(x)$ and have an Haros graph whose degree distribution has an exponential tail with base equal to the reciprocal of the Golden number, and additional analytical and numerical evidence supports the fact that this is also the case for the whole set of noble numbers $\mathcal{C}_m(n_1, \dots, n_m)$. Moreover, the Noble numbers and metallic ratios are also local maxima of $S(x)$, and their Haros graphs have a degree distribution with an exponential tail with base the reciprocal of the Golden number or metallic ratio, respectively. In other words, the infinite zigzag period- b expansion $[\bar{b}]$ that emerges after the transient of the continued fraction expansion of any Noble number or metallic ratio is mapped into an Haros graph with a degree distribution with an exponential tail of base ϕ_b^{-1} . Observe that Noble numbers and metallic ratios are subsets of quadratic irrationals, this larger set being formed by numbers whose continued fraction –after a generic transient– reaches a generic periodic pattern of period b (Nobles have period 1 pattern). At this point we can state the following conjectures, that will form the basis of future work:

Conjecture 2. *All quadratic irrationals have associated Haros graphs with a periodic pattern of zeros in the degree distribution and an exponential envelope in the tail whose base is related to the periodic pattern of zeros (i.e., to the periodic pattern of the continued fraction expansion).*

Conjecture 3. Only quadratic irrationals have the properties displayed in the previous conjecture. Accordingly, non-quadratic irrationals have associated Haros graphs with a degree distribution which is not maximally entropic, i.e. does not have an exponential tail.

For illustration of these conjectures, in Fig. 17 we plot $P(k, x)$ for $x = \phi^{-1}$ (Golden number, period 1), $x = [0; 2, 4, 1, 1, 2, 2, \bar{1}]$ (Noble number, period 1), $x = \phi_2^{-1}$ (Silver ratio, period 2), $x = \sqrt{3} - 1$ (non-noble, nonmetallic, quadratic, period 1, 2) and $x = e - 2$ (non-algebraic). A proof of conjecture 3 would naturally yield a classification of irrational numbers into quadratic and non-quadratic irrationals.

Furthermore, additional research should also be undertaken to explore the ability of Haros graphs to provide constructive classifications of other sets of irrationals, e.g., transcendental or normal numbers, this being an important open problem in mathematics. In this sense, see Fig. 18 where we plot the degree distributions of Haros graphs associated with different transcendental, algebraic and unclassified numbers, giving a preliminary hint of the apparent differences between algebraic and transcendental numbers in the Haros graph representation.

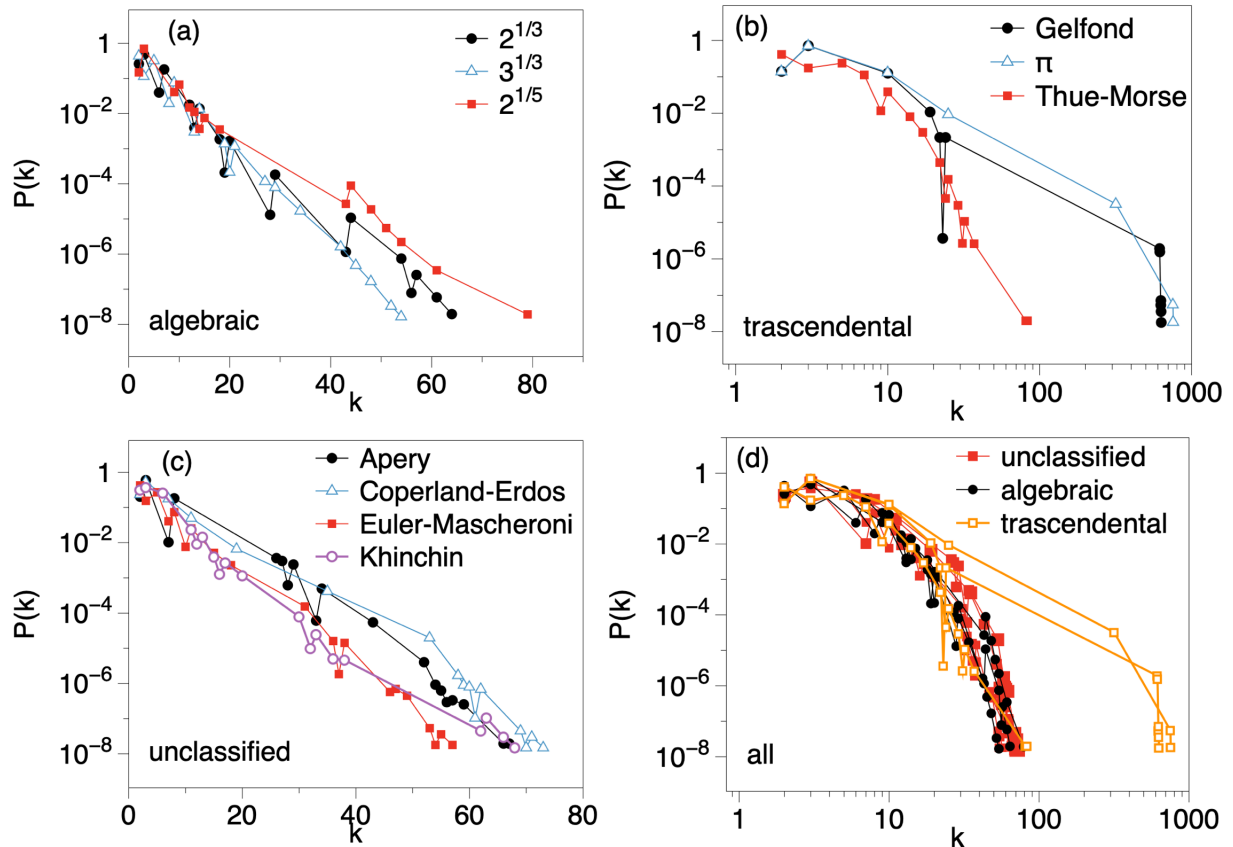


Fig. 18. Degree distributions $P(k, x)$ of Haros graphs associated with irrational numbers of different families: (a) (non-quadratic) algebraics (fractional parts of $\sqrt[3]{2}$, $\sqrt[5]{2}$ and $\sqrt[3]{3}$), (b) transcendentals (Gelfond constant e^π , the fractional part of π and Prouhet–Thue–Morse τ), (c) unclassified (Euler-Mascheroni constant γ , fractional part of Apery constant $\zeta(3)$, fractional part of Khinchin K_0 and Copeland-Erdős constant $C_{CE} = 0.2357111317\dots$). Some plots are in semi-log, revealing approximate yet not perfect exponential trends for (non quadratic) algebraic numbers, while others are in log-log, revealing longer tails. Panel (d) merges all of them, and suggest that transcendentals and algebraics typically display different structural signatures.

The above-mentioned open problems and conjectures relate to a proposed research programme on the *structural* properties of Haros graphs. Complementary to this, we envision two additional research avenues. The first is about defining *dynamical* rules on the set of Haros graphs by means of a graph renormalization operator [20, 31] and accordingly studying how the dynamics attractors relate to the underlying number system. The second stems from the fact that Noble numbers fulfil a maximum entropy principle, which suggests the development of a *statistical-mechanical* formalism underlying the set of Haros graphs. All in all, we sincerely hope that the community will find these open problems of interest and that the research programme on Haros graphs will get momentum in the years to come.

VI. APPENDIX

A. Proof of the three first values of $P(k, x)$

In what follows, we assume $p/q < 1/2$ without loss of generality. We first prove that, out of a total of q nodes (the period), $G_{p/q}$ will have p nodes with degree $k = 2$ and $q - 2p$ nodes with degree $k = 3$, i.e. $P(2, p/q) = p/q$ and $P(3) = 1 - 2 \cdot p/q$. We prove the results by induction over the level n of $\cup_n \ell_n$:

Theorem: $P(2, p/q) = p/q$.

Proof. The root of the Haros graph tree $G_{1/2}$ has $p = 1$ nodes with degree $k = 2$. Let us assume now that two given graphs are *adjacents* in the Farey binary tree at a given level n , with the property that their degree distributions fulfill the induction hypothesis $P(2, p_1/q_1) = p_1/q_1$ and $P(2, p_2/q_2) = p_2/q_2$ for G_{p_1/q_1} and G_{p_2/q_2} , respectively. By virtue of the properties of the concatenation operation, we have $G_{p_1/q_1} \oplus G_{p_2/q_2} = G_{p/q}$, where $p = p_1 + p_2$ and $q = q_1 + q_2$. Since, by construction in all Haros graphs, but $G_{0/1}$, the initial and final nodes have degree $k \geq 2$ (Fig. 3 illustrates this matter in the first level of Haros graph Tree), an Haros graph resulting from the concatenation of two other Haros graphs does not produce new nodes with degree $k = 2$, hence the total number of nodes with degree $k = 2$ in $G_{p/q}$ is simply $p = p_1 + p_2$, hence $P(2, p/q) = p/q$. ■

Theorem: $P(3, p/q) = 1 - 2 \cdot p/q$.

Proof. We must distinguish the special case $p/q = 1/q$. Trivially, we find that $G_{1/2}$ has $2 - 2 \cdot 1 = 0$ nodes with degree $k = 3$ in agreement with the statement. Let us assume that $G_{1/(q-1)}$ has $(q-1) - 2 = q - 3$ nodes of degree $k = 3$. Then, the graph $G_{1/q} = G_{0/1} \oplus G_{1/(q-1)}$ will have $q - 3$ nodes with $k = 3$ plus one node obtained by the concatenation of $G_{0/1}$; i.e., $P(3, 1/q) = (q-2)/q = 1 - 2 \cdot (1/q)$. For all other cases, since by construction, the initial and final nodes in each Haros graph have degrees $k \geq 2$, the concatenation of two graphs ($G_{p_1/q_1} \oplus G_{p_2/q_2} = G_{p/q}$) does not produce new nodes with $k = 3$, hence the total number of nodes with $k = 3$ in the concatenated graph is simply the sum $(q_1 - 2p_1) + (q_2 - 2p_2) = (q_1 + q_2) - 2(p_1 + p_2) = q - 2p$. ■

Theorem: $P(4, p/q) = 0$.

Proof. Finally, every Haros graph $G_{p/q}$ with $p/q \neq 1/2$ has no nodes with $k = 4$. The proof is based on an earlier fact: It is not possible for nodes of order 4 to appear, as none of them appear in the interior of any graph. It does not appear in the union node, as there are no unions of nodes with degree 1 and 3, nor 2 and 2 that can generate it. ■

Observe, to conclude, that we have focused on $G_{p/q}$ for $p/q < 1/2$. In the symmetric case $p/q > 1/2$, we end up with the same result if we change $p \rightarrow q - p$: $P(2, p/q) = 1 - p/q$, $P(3, p/q) = 2 \cdot p/q - 1$ and $P(4, p/q) = 0$.

B. Proof of distribution of holes in $P(k, x)$

In order to prove the theorem, we want to study how do the Haros graphs acquire their structure as they grow during the descent through the Haros graph Tree. The proof presents a slight technical drawback. The boundary node convention expressed in Section III.A is applied after the construction of the whole Haros graph tree, i.e., to obtain a new Haros graph, the ancestors do not have the boundary node identification.

Lemma 5. *The boundary node has degree $\kappa + 2$ at the level κ , $\forall \kappa \geq 3$.*

Proof of Lemma B.5. By induction over κ :

For $\kappa = 3$, the graphs $G_{1/3}$ and $G_{2/3}$ have degree $\kappa = 5$ at the boundary nodes.

Induction hypothesis: Suppose the result for $\kappa \leq n$. Let be an Haros graph $G = G_L \oplus G_R$, where G, G_L and G_R are at level $\kappa \leq n$ as we see in the figure 19.

The induction hypothesis gives us the result that the boundary nodes (the joining of two extreme nodes in the picture, as we mentioned in *Boundary node convention*) are:

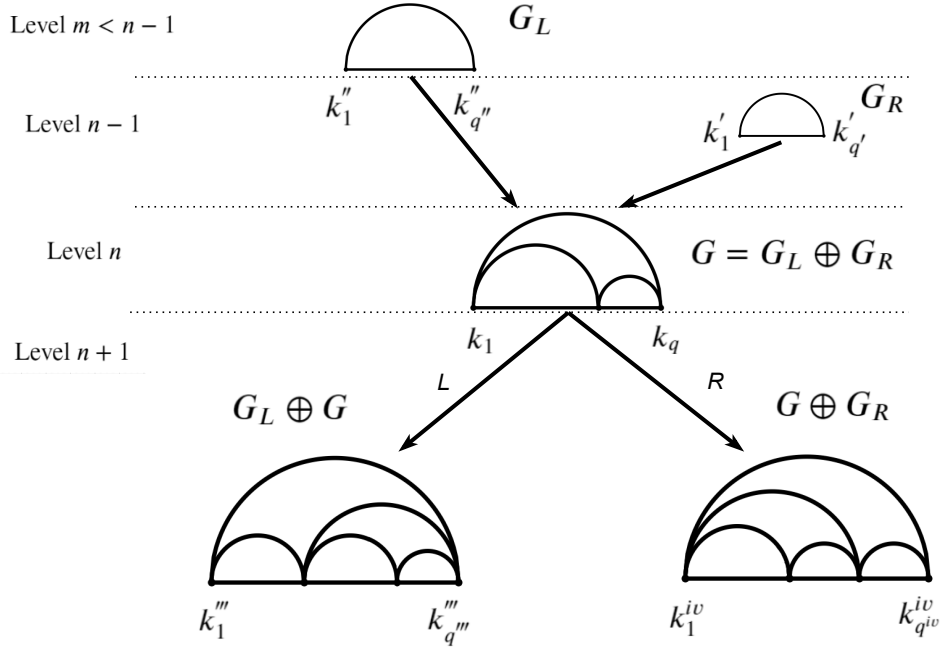


Fig. 19. Diagram of boundary node generation. Given the boundary node of Haros graph G located at level n and the boundary nodes of their ancestors G_L and G_R , the diagram shows the boundary node of descendants of G at level $n + 1$.

$$k_1'' + k_{q''}'' = m + 2, \quad (23)$$

$$k_1' + k_{q'}' = (n - 1) + 2 = n + 1, \quad (24)$$

$$k_1 + k_q = n + 2. \quad (25)$$

The construction of Haros graphs implies the following relationships:

$$k_1 = k_1'' + 1, \quad (26)$$

$$k_q = k_q' + 1. \quad (27)$$

Therefore, the descendant Haros graphs of G are $G_L \oplus G$ and $G \oplus G_R$, located at the level $n + 1$. The boundary node of $G_L \oplus G$ has degree:

$$k_1''' + k_{q''' }''' = k_1 + (k_q + 1) = n + 3,$$

whereas for $G \oplus G_R$:

$$k_1^{iv} + k_{q^{iv}}^{iv} = (k_1 + 1) + k_q = n + 3.$$

■

Corollary 6. *The merging node has a smaller degree than the boundary node.*

Proof of Corollary B.6. The merging node of G has degree (using equations [23, 25, 26, 27]):

$$\begin{aligned} k_M &= k_{q''}'' + k_1' = (m + 2 - k_1'') + (n + 1 - k_{q'}') \\ &= (m - 2 - (k_1 - 1)) + (n + 1 - (k_q - 1)) \\ &= m + 3 + n + 2 - (k_1 + k_q) = m + 3 \\ &< n + 2 = k_1 + k_q. \end{aligned}$$

■

It is easy to verify the following result:

Corollary 7. *The boundary node has the highest degree in the Haros graph.*

Proof of Corollary B.7.

A new induction exercise over the level in the Haros graph tree κ requires checking that, an internal node at level $n+1$ has lower degree than the boundary node, which by the lemma 5 has degree $n+3$. Now, we have two possibilities:

- The internal node is the merging node. Then, by the corollary 6 we conclude the proof.
- That internal node is not the merging node. In this case, the node was an internal node of one of its ancestors (let us suppose that of G). Then, applying the induction hypothesis we will have that the degree k_α of that node verifies:

$$k_\alpha < n + 2 < n + 3$$

■

The lemma 5 and the corollaries 6 and 7 lead us to the following conclusion: The degree κ may appear, at first, at the level $\kappa - 2$.

Lemma 8. *The symbol repetition in a path in the Haros graph Tree (either LL or RR) will not generate a new degree except in the boundary node.*

Proof of Lemma B.8. We start by considering the diagram of the lemma 5. If $G = G_L \oplus G_R$ is located at level $\kappa = n$ and G_R is located at $\kappa = n - 1$, then G is the left descendant of G_R . Hence, the last symbol is L (the proof is analogous for R).

The right descendant $G \oplus G_R$ has a merging node with degree:

$$k_M^{iv} = k_q + k'_1 = k'_{q'} + 1 + k'_1 = n + 2,$$

where we use $k'_1 + k'_{q'} = n + 1$ according to the lemma 5. Hence, the degree $n + 2$ is enclosed in the Haros graph, and the subtree with root $G \oplus G_R$ will contain this degree in each Haros graph.

The left descendant $G_L \oplus G$ has a merging node with degree:

$$k_M''' = k''_{q''} + k_1 = k''_{q''} + k'_1 + 1 = m + 3,$$

Using again the lemma 5 to highlight $k''_{q''} + k'_1 = m + 2$. However, this is the same degree as the merging node of G . Using equations 23, 24 and $k_1 + k_q = n + 2$:

$$k''_{q''} + k'_1 = (m + 2 - k''_1) + (n + 1 - k'_{q'}) = (m + 2 - k_1 + 1) + (n + 1 - k_q + 1) = m + 3 + n + 2 - (k_1 + k_q) = m + 3.$$

To conclude the proof of the theorem, we need to ensure that degree $\kappa = n + 2$ cannot appear. But the merging node in $(G_L \oplus G) \oplus G$ has degree:

$$k_{q'''} + k_1 = k_q + 1 + k_1 = n + 3.$$

Then, the degrees generated in this subtree will be greater than $n + 2$. Analogously, it is easy to verify that $G_L \oplus (G_L \oplus G)$ will also not contain degree $n + 2$. ■

C. Proof of scaling equations of $P(k, x)$, $H(x)$ and $S(x)$ equations 6, 7, 11 and 13)

In order to prove the different scaling equations, we first analyse the operator $F(x) = \frac{x}{1+x}$ as real number function and as a Haros graph operator. Therefore, the first result is easy to prove:

Lemma 9.

$$F\left(\frac{p_1}{q_1} \oplus \frac{p_2}{q_2}\right) = F\left(\frac{p_1}{q_1}\right) \oplus F\left(\frac{p_2}{q_2}\right) \quad (28)$$

Proof of Lemma C.9. Let be $\frac{p}{q} = \frac{p_1}{q_1} \oplus \frac{p_2}{q_2}$. By definition, we have:

$$F\left(\frac{p}{q}\right) = \frac{\frac{p}{q}}{1 + \frac{p}{q}} = \frac{p}{q+p}$$

Then, the left side of the equation has a value:

$$F\left(\frac{p_1}{q_1} \oplus \frac{p_2}{q_2}\right) = \frac{p_1 + p_2}{q_1 + q_2 + p_1 + p_2}$$

, whereas the right side of the equation is equal to :

$$F\left(\frac{p_1}{q_1}\right) \oplus F\left(\frac{p_2}{q_2}\right) = \frac{p_1}{q_1 + p_1} \oplus \frac{p_2}{q_2 + p_2} = \frac{p_1 + p_2}{q_1 + q_2 + p_1 + p_2}$$

■

By virtue of the distributive property between the operator F and the mediant sum via \oplus given in the previous lemma, we can prove, by a simple induction exercise, that the operator F takes each number x one level down in the Farey Tree:

Lemma 10. $x \in \ell_n \Rightarrow F(x) \in \ell_{n+1}$.

Next, we need a result that is technically cumbersome but necessary for the proof of the following theorem. The lemma 5 determines that the boundary node of the Haros graph G_x at level n has degree $n + 2$. In addition, let us recall that the boundary node arises from identifying the extreme nodes, that is, that this value $n + 2$ arises from the sum of the degrees $k_0 + k_q$, the degrees of the extremes. Combining the lemmas 5 and 10 leads us to the conclusion that the boundary node of $G_{F(x)}$ has degree $n + 3$. Now, we do not know a priori the relation between the degree of the extreme node k_0 of G_x and the degree k_0^f of the extreme node of $G_{F(x)}$ (and the same for the other extreme node). Therefore, we state the following lemma:

Lemma 11. Let k_0 and k_q be the degrees of the extreme nodes of the Haros graph G_x , and let k_0^f and k_q^f be the degrees of the extreme nodes of the Haros graph $G_{F(x)}$. Then, we have the equalities:

$$k_0^f = k_1 \quad (29)$$

$$k_{q_f}^f = k_q + 1 \quad (30)$$

Proof of Lemma C.11.

A new induction on level κ in the Haros graph tree, proving that the result is immediate for level $\kappa = 2$, where we have only the graph $G_{1/2}$. Let us assume the result to be true up to level $\kappa = n - 1$, and take a graph G at level $\kappa = n$. Let us again make use of the diagram 19 and its notation. We can clearly see that, by construction, we obtain $k_0 = k_0'' + 1$ and $k_q = k_q'' + 1$. Now, applying the induction hypothesis to the ancestors of $G = G_L \oplus G_R$, we obtain:

$$\begin{aligned} k_0^{f''} &= k_0'' \\ k_{q_f}^{f''} &= k_{q_f}'' + 1 \\ k_0^{f'} &= k_0' \\ k_{q_f}^{f'} &= k_{q_f}' + 1 \end{aligned}$$

where $k_0^{f''}, k_{q_f}^{f''}$ y $k_0^{f'}, k_{q_f}^{f'}$ are the degree of extreme nodes of Haros graphs $F(G_L)$ y $F(G_R)$, respectively. Then, we obtain the degree of the extreme nodes of $F(G)$. Again, by construction, we get $k_0^f = k_1^{f''} + 1$ y $k_{q_f}^f = k_{q_f}^{f'} + 1$. In this manner, we have:

$$k_0^f = k_0^{f''} + 1 = k_0'' + 1 = k_1$$

and

$$k_{q_f}^f = k_{q_f}^{f'} + 1 = (k_{q_f}' + 1) + 1 = k_q + 1$$

■

With all these previous results, we can proceed to state and prove our next lemma, whose purpose is to show that the frequency of nodes of degree $k \geq 5$ in the graph $G_{F(x)}$ is the same as for the Haros graph G_x , except that there is a shift with respect to the degree:

Lemma 10 tells us that the image through the operator $F(x)$ is one level lower than x . The following lemma shows that the frequencies of degrees $k \geq 5$ of $G_{F(x)}$ are the same as the frequencies of degrees of G_x but with displacement:

Lemma 12. *Let be an Haros graph G_x and let $m_{k,x}$ be the number of nodes with degree $k \geq 5$. Then, $m_{k,x} = m_{k+1,F(x)}$ where $F(x) = \frac{x}{1+x}$.*

Proof of Lemma B.12. By induction over the levels of the Farey binary tree ℓ_n :

- In case $n = 3$, the proof is trivially true because there is only one degree k greater than 5, where we have $m_{5,1/3} = m_{6,1/4}$.
- Induction hypothesis: The result is true for every $G_{p/q}$ with $p/q \in \cup_{i \leq n} \ell_i$.
- Let us consider $G_{p/q}$ an Haros graph with $p/q \in \ell_{n+1}$. Then $G_{p/q} = G_{p_1/q_1} \oplus G_{p_2/q_2}$, with $p_1/q_1, p_2/q_2 \in \cup_{i \leq n} \ell_i$. Consider an arbitrary non-null k in the degree distribution of $G_{p/q}$. Therefore, we have 3 possibilities:
 - The boundary node has degree k : The lemma 5 and the corollary 7 imply that $k = n + 3$ and $m_{k,p/q} = 1$. Hence, by the lemma 10 we get $F(p/q) \in \ell_{n+2}$ and its boundary node has degree $k = n + 4$. Again, by lemma 5 and corollary 7, we get $m_{k+1,F(p/q)} = 1$ and the result is verified.
 - The merging node has degree k : To clarify this case, we must distinguish between two cases:
 - * $m_{k,p/q} = 1$, i.e., the merging node is the only node with degree k . Therefore, an ancestor has the boundary node of degree k . This is due to the fact that there are no internal nodes of $G_{p/q}$ with that degree, so its ancestors cannot have internal nodes either; and the degree k cannot appear for the first time at the merging node. Thus, we can assume without loss of generality, that the left ancestor G_{p_1/q_1} has the boundary node with degree k . The corollary 5 implies that $m_{k,p_1/q_1} = 1$. On the other hand, the right ancestor p_2/q_2 cannot have nodes of degree k , since we have seen that its internal nodes do not have this degree, and the boundary node cannot have the same degree, since both ancestors would then be at the same level of the Haros graph tree. Therefore, we have that: $m_{k,p_2/q_2} = 0$. Applying the induction hypothesis on the ancestors, we obtain: $m_{k+1,F(p_1/q_1)} = 1$ and $m_{k+1,F(p_2/q_2)} = 0$. Thanks to the lemma 9, we know that the ancestors of $F(p/q)$ are $F(p_1/q_1)$ and $F(p_2/q_2)$. Then, we conclude the study of the sub-case with equality:

$$m_{k+1,F(p/q)} = m_{k+1,F(p_1/q_1)} + m_{k+1,F(p_2/q_2)} = 1 + 0 = 1 = m_{k,p/q}.$$

- * $m_{k,p/q} > 1$, i.e., in this case there are internal nodes with degree k apart from the merging node. While in the previous sub-case, the degree of the boundary node of an ancestor was inherited to the merging node of its right descendant (i.e., obtained by an R descent), this case would be obtained by continuing to perform R descents, that is, adding in each descent a single internal node with degree k resulting from the merging node of the previous step. In fact, we first check that, as before, one of the two ancestors of $G_{p/q}$ does not have nodes of degree k . If so, it would follow that both would come from an ancestor with boundary node k , but this is impossible, since two Haros graphs of different subtrees cannot be concatenated. So, we again assume $m_{k,p_2/q_2} = 0$, while the other ancestor has the $m_{k,p/q} - 1$

internal nodes of degree k , and does not have any more, since the boundary node cannot have degree k , i.e. $m_{k,p_1/q_1} = m_{k,p/q} - 1$. Now, applying the induction hypothesis on the ancestor p_1/q_1 and p_2/q_2 , it follows that $m_{k+1,F(p_1/q_1)} = m_{k,p/q} - 1$ and $m_{k+1,F(p_2/q_2)} = 0$. Finally, since the concatenation of the extreme nodes of p_1/q_1 and p_2/q_2 produced a node of degree $k_{q_1}^1 + k_0^2 = k$, by virtue of the lemma 11, the concatenation of $F(p_1/q_1)$ and $F(p_2/q_2)$ produces a merging node of degree

$$k_{q_{F(1)}}^{F(1)} + k_0^{F(2)} = k_{q_1}^1 + k_1^2 + 1 = k + 1,$$

and we can conclude that:

$$m_{k+1,F(p/q)} = m_{k+1,F(p_1/q_1)} + m_{k+1,F(p_2/q_2)} + 1 = m_{k,p/q} - 1 + 0 + 1 = m_{k,p/q}.$$

* $m_{k,p/q} > 1$, i.e, other inner nodes have a degree k apart from the merging node. In that subcase, we again find that one ancestor has no degree k . The other ancestor, supposing again p_1/q_1 , has exactly the same number of nodes with degree k as inner nodes of degree k in p/q , i.e, $m_{k,p_1/q_1} = m_{k,p/q} - 1$. Applying the induction hypothesis and the concatenation of $F(p_1/q_1) \oplus F(p_2/q_2)$, we have $m_{k,p/q} = m_{k+1,F(p/q)}$.

– Only some inner nodes have degree k . In that case, we must distinguish two subcases:

* The number of nodes with degree k in $G_{p/q}$ is the sum of nodes with degree k in his ancestors, that is:

$$m_{k,p/q} = m_{k,p_1/q_1} + m_{k,p_2/q_2}.$$

Therefore, by hypothesis induction:

$$m_{k,p_1/q_1} = m_{k+1,F(p_1/q_1)},$$

$$m_{k,p_2/q_2} = m_{k+1,F(p_2/q_2)},$$

and, we clearly have:

$$m_{k+1,F(p/q)} = m_{k+1,F(p_1/q_1)} + m_{k+1,F(p_2/q_2)} = m_{k,p_1/q_1}.$$

* The other case occurs when $m_{k,p/q} = m_{k,p_1/q_1}$ and $m_{k,p_2/q_2} = 1$ (analogously if $m_{k,p/q} = m_{k,p_2/q_2}$ and $m_{k,p_1/q_1} = 1$). Hence, $m_{k+1,F(p_2/q_2)} = 1$ and $m_{k+1,F(p_1/q_1)} = m_{k,p_1/q_1}$, concluding that:

$$m_{k+1,F(p/q)} = m_{k+1,F(p_1/q_1)} = m_{k,p_1/q_1} = m_{k,p/q}.$$

■

The lemma 12 leads us to the first scaling equation. This self-similar behaviour can be observed in Fig. 6

Theorem 13. *For all $k \geq 5$, we have the following scaling equation:*

$$P\left(k+1, F(x) = \frac{x}{x+1}\right) = \frac{1}{x+1} \cdot P(k, x).$$

Proof. Let us $x = p/q$ and $m_{k,p/q}$ be the number of nodes with degree k in $G_{p/q}$. In virtue of Lemma 12, we have $m_{k,x} = m_{k+1,F(p/q)}$. Using that $F\left(\frac{p}{q}\right) = \frac{p}{p+q}$ then:

$$P\left(k, \frac{p}{q}\right) = \frac{m_{k,p/q}}{q} = \frac{m_{k+1,F(p/q)}}{q} = \frac{p+q}{q} \cdot \frac{m_{k+1,F(p/q)}}{p+q} = \left(1 + \frac{p}{q}\right) \cdot P(k+1, F(p/q)).$$

■

Theorem 13 shows that if we know the following degree values of G_x : $P(5, x), P(6, x), \dots$, we know the degree values of $G_{F(x)}$: $P(6, F(x)), P(7, F(x)), \dots$. Then, the only unknown value of $P(k, F(x))$ is $k = 5$. The following corollary shows this value to complete the degree distribution of $G_{F(x)}$:

Corollary 14. *Given the degree distribution $P(k, x)$ of the Haros graph G_x , the degree distribution $P(k, F(x))$ of the Haros graph $G_{F(x)}$ is known. In particular, if $x > 1/2$, then $P(k = 5, F(x)) = \frac{2x-1}{x+1}$ and if $x < 1/2$, then $P(k = 5, F(x)) = 0$.*

Proof of Corollary B.14.

If $x < 1/2$, we know that:

$$\sum_{k \geq 5} P(k, x) = 1 - P(2) - P(3) - P(4) = 1 - x - (1 - 2x) = x.$$

Moreover, as $F(x) < 1/2$, we have:

$$\sum_{k \geq 5} P(k, F(x)) = \frac{x}{x+1}.$$

By Theorem 13, we note the following:

$$\begin{aligned} \frac{x}{x+1} &= \sum_{k \geq 5} P(k, F(x)) = P(5, F(x)) + \sum_{k \geq 6} P(k, F(x)) = \\ &= P(5, F(x)) + \frac{1}{x+1} \sum_{k \geq 5} P(k, x) = P(5, F(x)) + \frac{x}{x+1}. \end{aligned}$$

Hence, $P(5, F(x)) = 0$.

Now, if $x > 1/2$, then:

$$\sum_{k \geq 5} P(k, x) = 1 - x,$$

while again we have:

$$\sum_{k \geq 5} P(k, F(x)) = \frac{x}{x+1}.$$

Using Theorem 13, we observe:

$$\begin{aligned} \frac{x}{x+1} &= \sum_{k \geq 5} P(k, F(x)) = P(5, F(x)) + \sum_{k \geq 6} P(k, F(x)) = \\ &= P(5, F(x)) + \frac{1}{x+1} \sum_{k \geq 5} P(k, x) = P(5, F(x)) + \frac{1-x}{x+1}. \end{aligned}$$

Hence, $P(5, F(x)) = \frac{2x-1}{x+1}$.

■

To illustrate the meaning of the corollary, we first show the degree distribution of the Haros graph $G_{2/5}$ and the Haros graph of $F(2/5) = 2/7$:

$$P(k, 2/5) = \begin{cases} 2/5, & k = 2 \\ 1/5, & k = 3 \\ 0, & k = 4 \\ 1/5, & k = 5, 6 \end{cases} \quad P(k, 2/7) = \begin{cases} 2/7, & k = 2 \\ 3/7, & k = 3 \\ 0, & k = 4, 5 \\ 1/7, & k = 6, 7. \end{cases}$$

It can be observed that the degrees $k = 6, 7$ of $G_{2/7}$ that appear are the same degrees as those of $G_{2/5}$ that displaced one position, i.e., $k = 5, 6$. Moreover, the degree distribution follows the scaling equation given in Theorem 13. However, if we now consider $x = 3/5 > 1/2$ and his image by $F(x)$, $3/8$, we have the degree distributions:

$$P(k, 3/5) = \begin{cases} 2/5, & k = 2 \\ 1/5, & k = 3 \\ 0, & k = 4 \\ 1/5, & k = 5, 6 \end{cases} \quad P(k, 3/8) = \begin{cases} 3/8, & k = 2 \\ 2/8, & k = 3 \\ 0, & k = 4 \\ 1/8, & k = 5, 6, 7. \end{cases}$$

In that case, the degrees $k = 5, 6$ appearing in $G_{3/5}$ correspond to the degrees $k = 6, 7$ in $G_{3/8}$ through the scaling property. However, also the degree $k = 5$ is not generated via the scaling equation given in Theorem 13.

Theorem 15. For $x \in [0, 1]$, we have

$$H(x) = \begin{cases} (1+x) \cdot H(F(x)) & x \in [0, 1/2] \\ (1+x) \cdot H(F(x)) - x \log x + (2x-1) \log(2x-1) + (1-x) \log(1-x) & x \in [1/2, 1] \end{cases}$$

that is, the scaling of $H(x)$ is different depending on if $x < 1/2$ or $x > 1/2$.

Proof. Let us $x \in [0, 1/2]$. By the definition of $H(x)$ given in 10, it is clear that $H(x) = -\sum_{k \geq 4} P(k, x) \log(P(k, x)) + x \cdot \log x$ (remember that $P(4, x) = 0$). Therefore, using Theorem 13 :

$$\begin{aligned} H(x) &= -\sum_{k \geq 4} P(k, x) \log(P(k, x)) + x \cdot \log x \\ &= -\sum_{k \geq 4} (1+x)P(k+1, F(x)) \log((1+x)P(k+1, F(x))) + x \cdot \log x \\ &= (1+x) \left(-\sum_{k \geq 5} P(k, F(x)) \log(1+x) - \sum_{k \geq 5} P(k, F(x)) \log(P(k, F(x))) + \frac{x}{1+x} \cdot \log x \right) \\ &= (1+x) \left(-\frac{x}{1+x} \log(1+x) - \sum_{k \geq 5} P(k, F(x)) \log(P(k, F(x))) + \frac{x}{1+x} \cdot \log x \right) \\ &= (1+x) \left(-\sum_{k \geq 5} P(k, F(x)) \log(P(k, F(x))) + \frac{x}{1+x} \cdot \log \frac{x}{1+x} \right) \\ &= (1+x) \left(-\sum_{k \geq 5} P(k, F(x)) \log(P(k, F(x))) + F(x) \cdot \log F(x) \right) \\ &= (1+x) \cdot H(F(x)). \end{aligned}$$

Let us consider $x \in [1/2, 1]$. In that case, we have $H(x) = -\sum_{k \geq 4} P(k, x) \cdot \log(P(k, x)) + (1-x) \cdot \log(1-x)$. Moreover, using the scaling of $P(k, x)$ provided in Theorem 13 and the degree distribution of $P(k, F(x))$ given in Corollary 14, we have the following:

$$H(x) = - \sum_{k \geq 4} P(k, x) \cdot \log(P(k, x)) + (1 - x) \cdot \log(1 - x).$$

$$\text{Adding and subtracting the term } P(k_0, F(x)) = \frac{2x - 1}{x + 1}:$$

$$\begin{aligned} &= - \sum_{k \geq 5} (1 + x) \cdot P(k, F(x)) \cdot \log((1 + x) \cdot P(k, F(x))) + (1 + x) \cdot \frac{2x - 1}{x + 1} \cdot \log\left((1 + x) \cdot \frac{2x - 1}{x + 1}\right) + \\ &+ (1 - x) \cdot \log(1 - x) \\ &= (1 + x) \cdot \left[- \sum_{k \geq 5} P(k, F(x)) \cdot \log(1 + x) - \sum_{k \geq 5} P(k, F(x)) \cdot \log(P(k, F(x))) + \frac{2x - 1}{x + 1} \cdot \log(2x - 1) + \right. \\ &+ \left. \frac{1 - x}{1 + x} \cdot \log(1 - x) \right] \\ &= (1 + x) \cdot \left[- \frac{x}{1 + x} \cdot \log(1 + x) - \sum_{k \geq 5} P(k, F(x)) \cdot \log(P(k, F(x))) + \frac{2x - 1}{x + 1} \cdot \log(2x - 1) + \right. \\ &+ \left. \frac{1 - x}{1 + x} \cdot \log(1 - x) \right] \\ &= (1 + x) \cdot \left[- \sum_{k \geq 5} P(k, F(x)) \cdot \log(P(k, F(x))) + \frac{x}{1 + x} \cdot \log\left(\frac{x}{1 + x}\right) - \frac{x}{1 + x} \cdot \log(x) + \right. \\ &+ \left. \frac{2x - 1}{x + 1} \cdot \log(2x - 1) + \frac{1 - x}{1 + x} \cdot \log(1 - x) \right] \\ &= (1 + x) \cdot H(F(x)) - x \cdot \log(x) + (2x - 1) \cdot \log(2x - 1) + (1 - x) \cdot \log(1 - x). \end{aligned}$$

■

Theorem 16. For $x \in [1/2, 1]$, we have the following functional scaling equation:

$$S(x) = (x + 2) \cdot H\left(\frac{1}{x + 2}\right).$$

Proof. For $x > 1/2$, we have:

$$S(x) = H(x) - 2 \cdot (1 - x) \cdot \log(1 - x) - (2x - 1) \cdot \log(2x - 1).$$

Applying the scaling equation of $H(x)$ when $x > 1/2$:

$$\begin{aligned} S(x) &= (1 + x) \cdot H\left(\frac{x}{x + 1}\right) - x \cdot \log(x) + \underbrace{(2x - 1) \cdot \log(2x - 1)}_{\text{I)}} + (1 - x) \cdot \log(1 - x) - \\ &- 2(1 - x) \cdot \log(1 - x) - \underbrace{(2x - 1) \cdot \log(2x - 1)}_{\text{I)}} \end{aligned}$$

$$= (1 + x) \cdot H\left(\frac{x}{x + 1}\right) - x \cdot \log(x) - (1 - x) \cdot \log(1 - x)$$

By symmetry of H , we have $H(x) = H(1 - x)$:

$$= (1 + x) \cdot H\left(\frac{1}{x + 1}\right) - x \cdot \log(x) - (1 - x) \cdot \log(1 - x)$$

As $F(x) = \frac{x}{1 + x} < \frac{1}{2}$, then $\frac{1}{x + 1} = 1 - \frac{x}{x + 1} > \frac{1}{2}$.

Therefore, we apply again the scaling equation of $H(x)$ when $x > 1/2$:

$$\begin{aligned} &= (1 + x) \cdot \left[\left(1 + \frac{1}{x + 1}\right) \cdot H\left(\frac{1}{x + 2}\right) - \frac{1}{x + 1} \cdot \log\left(\frac{1}{x + 1}\right) + \left(\frac{2}{x + 1} - 1\right) \cdot \log\left(\frac{2}{x + 1} - 1\right) \right. \\ &+ \left. \left(1 - \frac{1}{x + 1}\right) \cdot \log\left(1 - \frac{1}{x + 1}\right) \right] - x \cdot \log(x) - (1 - x) \cdot \log(1 - x) \\ &= (x + 2) \cdot H\left(\frac{1}{x + 2}\right) + \log(x + 1) + (1 - x) \cdot \log\left(\frac{1 - x}{1 + x}\right) + x \cdot \log\left(\frac{x}{1 + x}\right) - x \cdot \log(x) - (1 - x) \cdot \log(1 - x) \\ &= (x + 2) \cdot H\left(\frac{1}{x + 2}\right) + \underbrace{\log(x + 1)}_{\text{II)}} + \underbrace{(1 - 2x) \cdot \log(1 - x)}_{\text{III)}} - \underbrace{(1 - x) \cdot \log(x + 1)}_{\text{II)}} + \underbrace{x \cdot \log(x)}_{\text{I)}} - \\ &- \underbrace{x \cdot \log(x + 1)}_{\text{II)}} - \underbrace{x \cdot \log(x)}_{\text{I)}} - \underbrace{(1 - x) \cdot \log(1 - x)}_{\text{III)}} \\ &= (x + 2) \cdot H\left(\frac{1}{x + 2}\right). \end{aligned}$$

■

D. Some aspects of $H(x)$

The entropy function $H(x)$ evaluated in $x = p/q$ is:

$$\begin{aligned} H(p/q) &= - \sum_{i \geq 5} \frac{n_i}{q} \cdot \log\left(\frac{n_i}{q}\right) + \frac{p}{q} \cdot \log\left(\frac{p}{q}\right) \\ &= - \frac{1}{q} \sum_{i \geq 5} n_i \cdot \log(n_i) + \frac{\log(q)}{q} \cdot \sum_{i \geq 5} n_i + \frac{p}{q} \cdot \log\left(\frac{p}{q}\right). \end{aligned}$$

Using that

$$\sum_{i \geq 5} n_i = q - p - q + 2p = p,$$

we obtain the following

$$H(p/q) = -\frac{1}{q} \sum_{i \geq 5} n_i \cdot \log(n_i) + \frac{p}{q} \log(q) + \frac{p}{q} \cdot \log\left(\frac{p}{q}\right) = -\frac{1}{q} \sum_{i \geq 5} n_i \cdot \log(n_i) + \frac{p}{q} \cdot \log(p).$$

Then:

$$\frac{q}{p} \cdot H(p/q) = \log(p) - \frac{1}{p} \sum_{i \geq 5} n_i \cdot \log(n_i).$$

For example, fractions $\frac{2}{q} = \frac{2}{2n+1} < \frac{1}{2}$ with $n \geq 2$ always have a node k_1 with frequency $n_1 = 1$ and the outer node k_2 with frequency $n_2 = 1$. This leads to the fact that $\frac{q}{2} \cdot H\left(\frac{2}{q}\right) = \log(2)$, as corroborated in Fig. 8.

If we now take the fractions of the numerator $p = 3$, we get two distinct possibilities: $\left\{\frac{3}{3n+1}\right\}_{n \geq 2}$ and $\left\{\frac{3}{3n+2}\right\}_{n \geq 2}$. It can be checked, as Fig. 8 illustrates for the 3 first elements, how the rational numbers $\frac{3}{3n+1}$ always have $n_1 = 1$ and $n_2 = 2$, while $\frac{3}{3n+2}$ have $n_1 = n_2 = n_3 = 1$. We obtain that this rational family always follows:

$$\begin{aligned} \frac{3n+1}{3} \cdot H\left(\frac{3}{3n+1}\right) &= \log(3) - \frac{2}{3} \cdot \log(2), \\ \frac{3n+2}{3} \cdot H\left(\frac{3}{3n+2}\right) &= \log(3). \end{aligned}$$

The explanation for this fact is as follows: the roots of the intervals \mathcal{I}_n are the concatenation of $G_{\frac{1}{n+1}}$ and $G_{\frac{1}{n}}$. The degree distribution is made up of 2 nodes with $k = 2$, one internal link of $k = n + 3$, the extreme node $k = n + 4$. The other nodes have degree $k = 3$. In summary, each graph reached by paths $L^n R P$ will have the same nodes with degree $p = 2$ and identical frequencies n_i , even if their degree values are different.

E. Proof of global maxima of $S(x)$

We shall now prove that S defined over the set of Haros graphs is indeed maximal when the degree distribution coincides with the degree distribution of $G_{\phi^{-1}}$. For symmetry, we shall then consider the interval $(1/2, 1]$ (we proceed analogously for $x' = 1 - \phi^{-1} \in [0, 1/2]$). We will use the technique of Lagrange multipliers with two constraints for $P(k, x)$: normalisation of the degree distribution and arithmetic mean degree distribution (that is closely related to q by construction). The Lagrangian functional has the form

$$\mathcal{L}[P(k, x)] = S(x) - (\text{normalization}) - (\text{arithmetic mean degree}),$$

which will take an extremum at the solution $\delta \mathcal{L}[P^*(k, x)] = 0$.

The normalisation constraints read

$$\mathcal{Q}_0 = 1 - P(2, x) - P(3, x) - P(4, x) = 1 - (1 - x) - (2x - 1) - 0 = 1 - x.$$

We know that if the number x is irrational, G_x will necessarily have an arithmetic mean degree $\bar{k} = 4$. In such a case, the second constraint is enclosed in the constant

$$\mathcal{Q}_1 = 4 - 2P(2, x) - 3P(3, x) - 4P(4, x) = 5 - 4x.$$

If x is instead a rational number $x = p/q$, the arithmetic mean degree $\bar{k} = 4(1 - 1/2q)$ and the restriction on the mean degree are slightly different.

For irrational numbers, the Lagrangian functional reads

$$\mathcal{L}[P(k, x)] = -\sum_{k=5}^{\infty} P(k, x) \log P(k, x) - (\lambda_0 - 1) \left(\sum_{k=5}^{\infty} P(k, x) - \mathcal{Q}_0 \right) - \lambda_1 \left(\sum_{k=5}^{\infty} kP(k, x) - \mathcal{Q}_1 \right),$$

for which the extreme condition is

$$\frac{\partial \mathcal{L}}{\partial P(k, x)} = -\log P(k, x) - \lambda_0 - \lambda_1 k = 0,$$

has the solution

$$P(k, x) = e^{-\lambda_0 - \lambda_1 k}.$$

From this, and using the definition of \mathcal{Q}_0 we get

$$\mathcal{Q}_0 = \sum_{k \geq 5} P(k, x) = \sum_{k=5}^{\infty} e^{-\lambda_0 - \lambda_1 k} = e^{-\lambda_0} \sum_{k=5}^{\infty} e^{-\lambda_1 k} = 1 - x.$$

As the infinite sum satisfies

$$\sum_{k=5}^{\infty} e^{-\lambda_1 k} = \frac{e^{-4\lambda_1}}{e^{\lambda_1} - 1}, \quad (1)$$

we have the following relationship between the Lagrange multipliers

$$e^{-\lambda_0} = \frac{(1-x)(e^{\lambda_1} - 1)}{e^{-4\lambda_1}}. \quad (2)$$

For the reduced mean connectivity \mathcal{Q}_1 we get:

$$\mathcal{Q}_1 = \sum_{k \geq 5} kP(k, x) = \sum_{k=5}^{\infty} k e^{-\lambda_0 - \lambda_1 k} = e^{-\lambda_0} \sum_{k=5}^{\infty} k e^{\lambda_1 k} = 5 - 4x. \quad (3)$$

To calculate the sum, we can differentiate Eq. 1 with respect to λ_1 , to obtain:

$$\sum_{k=5}^{\infty} k e^{-\lambda_1 k} = e^{-4\lambda_1} \frac{-4 + 5e^{\lambda_1}}{(e^{\lambda_1} - 1)^2}.$$

Substituting this sum into Eq. 3 and using Eq. 2 we have

$$5 - 4x = \left(\frac{(1-x)(e^{\lambda_1} - 1)}{e^{-4\lambda_1}} \right) \left(e^{-4\lambda_1} \frac{-4 + 5e^{\lambda_1}}{(e^{\lambda_1} - 1)^2} \right) = \frac{(1-x)(-4 + 5e^{\lambda_1})}{e^{\lambda_1} - 1},$$

which, after some algebra, yields for the second Lagrange multiplier

$$x = e^{-\lambda_1}.$$

Introducing the above result in Eq. 2 we obtain for the first Lagrange multiplier

$$e^{-\lambda_0} = \frac{(1-x)(x^{-1} - 1)}{x^4} = (1-x)^2 x^{-5}.$$

Therefore, the degree distribution maximising the graph entropy is given by

$$P(k, x) = \begin{cases} 1-x & k=2 \\ 2x-1 & k=3 \\ 0 & k=4 \\ (1-x)^2 x^{k-5} & k \geq 5. \end{cases}$$

But this is the distribution of $G_{\phi^{-1}}$, since

$$\begin{aligned} (1-\phi^{-1})^2 (\phi^{-1})^{k-5} &= (\phi^{-1})^{k-5} \left(\frac{\phi-1}{\phi} \right)^2 \\ &= (\phi^{-1})^{k-5} (\phi^{-1})^2 (1-\phi^{-1}) \\ &= (\phi^{-1})^{k-5} (\phi^{-1})^2 (\phi^{-1})^2 \\ &= (\phi^{-1})^{k-5+4} \\ &= \phi^{1-k}. \end{aligned}$$

Therefore, we conclude the proof. ■

F. Proof of local maxima

We now prove that S defined over the set of Haros graphs with $x \leq 1/n$ is maximal when the degree distribution coincides with the degree distribution of the graphs associated with noble numbers $\mathcal{C}_1(n) = \frac{1}{n+\phi^{-1}}$. As before, we establish two constraints for $P(k, x)$: normalisation of the degree distribution and the arithmetic mean degree distribution. An immediate consequence of Theorem 2 tells us that if $x \leq 1/n$, then $P(k, x) = 0$ for $4 \leq k \leq n+2$. The normalisation constraint reads

$$\mathcal{Q}_0 = 1 - P(2, x) - P(3, x) - \sum_{k \geq 4}^{n+2} P(k, x) = 1 - x - (1 - 2x) - 0 = x,$$

and so, the arithmetic mean degree constraint now is

$$\mathcal{Q}_1 = 4 - 2P(2, x) - 3P(3, x) - \sum_{k \geq 4}^{n+2} kP(k, x) = 4 - 2x - 3(1 - 2x) - 0 = 1 + 4x.$$

Then, for irrational numbers, the Lagrangian functional

$$\mathcal{L}[P(k, x)] = - \sum_{k=n+3}^{\infty} P(k, x) \log P(k, x) - (\lambda_0 - 1) \left(\sum_{k=n+3}^{\infty} P(k, x) - \mathcal{Q}_0 \right) - \lambda_1 \left(\sum_{k=n+3}^{\infty} kP(k, x) - \mathcal{Q}_1 \right),$$

for which the extremum condition reads

$$\frac{\partial \mathcal{L}}{\partial P(k, x)} = -\log P(k, x) - \lambda_0 - \lambda_1 k = 0,$$

has the solution

$$P(k, x) = e^{-\lambda_0 - \lambda_1 k}.$$

For constraint \mathcal{Q}_0 we have

$$\mathcal{Q}_0 = \sum_{k=n+3}^{\infty} P(k, x) = \sum_{k=n+3}^{\infty} e^{-\lambda_0 - \lambda_1 k} = e^{-\lambda_0} \sum_{k=n+3}^{\infty} e^{-\lambda_1 k} = x.$$

As the infinite sum gives,

$$\sum_{k=n+3}^{\infty} e^{-\lambda_1 k} = \frac{e^{-\lambda_1(n+2)}}{e^{\lambda_1} - 1}, \quad (31)$$

we obtain

$$e^{-\lambda_0} = x(e^{\lambda_1} - 1)e^{\lambda_1(n+2)}. \quad (32)$$

The second constraint satisfies

$$\mathcal{Q}_1 = \sum_{k=n+3}^{\infty} kP(k, x) = \sum_{k=n+3}^{\infty} ke^{-\lambda_0 - \lambda_1 k} = e^{-\lambda_0} \sum_{k=n+3}^{\infty} ke^{-\lambda_1 k} = 1 + 4x. \quad (33)$$

Again, to calculate the infinite sum, we can differentiate equation 31 with respect to λ_1 and obtain the following result

$$\sum_{k=n+3}^{\infty} ke^{-\lambda_1 k} = e^{-\lambda_1(n+2)} \cdot \frac{(n+3)e^{\lambda_1} - (n+2)}{(e^{\lambda_1} - 1)^2}.$$

And from equations 33 and 32

$$1 + 4x = e^{-\lambda_0} \sum_{n+3}^{\infty} ke^{-\lambda_1 k} =$$

$$\left(\frac{x(e^{\lambda_1} - 1)}{e^{-\lambda_1(n+2)}} \right) \cdot \left(e^{-\lambda_1(n+2)} \frac{(n+3)e^{\lambda_1} - (n+2)}{(e^{\lambda_1} - 1)^2} \right) = \frac{x((n+3)e^{\lambda_1} - (n+2))}{e^{\lambda_1} - 1}.$$

Simplification of the above yields for the second Lagrange multiplier

$$e^{\lambda_1} = \frac{1 - (n-2)x}{1 - (n-1)x}.$$

Introducing the above equality in equation 32, we get for the first Lagrange multiplier

$$e^{-\lambda_0} = \frac{x^2}{1 - (n-1)x} \left(\frac{1 - (n-2)x}{1 - (n-1)x} \right)^{n+2}.$$

Therefore, the degree distribution maximising graph entropy is given by:

$$P(k, x) = \begin{cases} x & k = 2 \\ 1 - 2x & k = 3 \\ 0 & 4 \leq k \leq n+2 \\ \frac{x^2}{1 - (n-1)x} \left(\frac{1 - (n-2)x}{1 - (n-1)x} \right)^{n+2-k} & k \geq n+3. \end{cases} \quad (34)$$

This is the probability distribution of $\mathcal{C}_1(n)$ and using these three identities

$$\begin{aligned} \frac{\mathcal{C}_1(n)}{\mathcal{C}_1(n+1)} &= 1 + \mathcal{C}_1(n), \\ \frac{1 - (n-2)\mathcal{C}_1(n)}{1 - (n-1)\mathcal{C}_1(n)} &= \phi, \\ \frac{\mathcal{C}_1(n)^2}{1 - (n-1)\mathcal{C}_1(n)} &= \mathcal{C}_1(n)\phi^{-1}, \end{aligned}$$

the expression has simplified: $P(k, x) = \mathcal{C}_1(n)(\phi^{-1})^{k-(n+1)}$ ■

G. Slope of $H(\mathcal{C}_1(n))$

$$\begin{aligned}
H(\mathcal{C}_1(n)) &= - \sum_{k=n+3}^{\infty} \mathcal{C}_1(n)(\phi^{-1})^{k-(n+1)} \cdot \log \left(\mathcal{C}_1(n)(\phi^{-1})^{k-(n+1)} \right) + \mathcal{C}_1(n) \cdot \log(\mathcal{C}_1(n)) \\
&= -\mathcal{C}_1(n) \left[\sum_{k=n+3}^{\infty} (\phi^{-1})^{k-(n+1)} \cdot \log(\mathcal{C}_1(n)) + \sum_{k=n+3}^{\infty} (\phi^{-1})^{k-(n+1)} \cdot (k - (n+1)) \cdot \log(\phi^{-1}) \right] + \\
&\quad + \mathcal{C}_1(n) \cdot \log(\mathcal{C}_1(n)) \\
&= -\mathcal{C}_1(n) \left[\log(\mathcal{C}_1(n)) \cdot \frac{(\phi^{-1})^2}{1 - \phi^{-1}} + \log(\phi^{-1}) \cdot \frac{(2 - \phi^{-1}) \cdot (\phi^{-1})^2}{(\phi^{-1} - 1)^2} \right] + \mathcal{C}_1(n) \cdot \log(\mathcal{C}_1(n)) \\
&= -\mathcal{C}_1(n) [\log(\mathcal{C}_1(n)) + \log(\phi^{-1}) \cdot (3 + \phi^{-1})] + \mathcal{C}_1(n) \cdot \log(\mathcal{C}_1(n)) \\
&= -\mathcal{C}_1(n) \cdot \log(\phi^{-1}) \cdot (3 + \phi^{-1}).
\end{aligned}$$

Hence:

$$\frac{H(\mathcal{C}_1(n))}{\mathcal{C}_1(n)} = -\log(\phi^{-1}) \cdot (3 + \phi^{-1}).$$

H. Slope of $H(\mathcal{C}_3(n, 1, 2))$

Let us $x = [n, 1, 2, \bar{1}] = \frac{3+\phi^{-1}}{(3n+2)+(n+1)\cdot\phi^{-1}}$, for $n \geq 2$.

To shorten the equations, we denote $\mathcal{Q} = \frac{1}{(3n+2)+(n+1)\cdot\phi^{-1}}$. We have the following:

$$\begin{aligned}
H(\mathcal{C}_3(n, 1, 2)) &= -\frac{1}{\mathcal{Q}} \cdot \log\left(\frac{1}{\mathcal{Q}}\right) - \frac{1+\phi^{-1}}{\mathcal{Q}} \cdot \log\left(\frac{1+\phi^{-1}}{\mathcal{Q}}\right) - \\
&\quad - \sum_{k=n+6}^{\infty} \left[\frac{1-\phi^{-1}}{\mathcal{Q}} \cdot (\phi^{-1})^{k-(n+6)} \cdot \log\left(\frac{1-\phi^{-1}}{\mathcal{Q}} \cdot (\phi^{-1})^{k-(n+6)}\right) \right] + \frac{3+\phi^{-1}}{\mathcal{Q}} \cdot \log\left(\frac{3+\phi^{-1}}{\mathcal{Q}}\right) \\
&= -\frac{1}{\mathcal{Q}} \cdot \left[(1+\phi^{-1}) \cdot \log(1+\phi^{-1}) - (2+\phi^{-1}) \cdot \log(\mathcal{Q}) - \right. \\
&\quad \left. + (1+\phi^{-1}) \cdot \left(\log\left(\frac{1-\phi^{-1}}{\mathcal{Q}}\right) \cdot (2+\phi^{-1}) + \log(\phi^{-1}) \cdot \left(1 + \frac{2}{\phi^{-1}}\right) \right) - \right. \\
&\quad \left. - (3+\phi^{-1}) \cdot \log\left(\frac{3+\phi^{-1}}{\mathcal{Q}}\right) \right] \\
&= -\frac{1}{\mathcal{Q}} \cdot \left[(1+\phi^{-1}) \cdot \log(1+\phi^{-1}) - (2+\phi^{-1}) \cdot \log(\mathcal{Q}) + \log\left(\frac{1-\phi^{-1}}{\mathcal{Q}}\right) \right. \\
&\quad \left. + \phi \cdot \log(\phi^{-1}) - (3+\phi^{-1}) \cdot \log\left(\frac{3+\phi^{-1}}{\mathcal{Q}}\right) \right] \\
&= -\frac{1}{\mathcal{Q}} \cdot \left[(1+\phi^{-1}) \cdot \log(1+\phi^{-1}) + \log(1-\phi^{-1}) + \phi \log(\phi^{-1}) + (3+\phi^{-1}) \cdot \log(3+\phi^{-1}) \right. \\
&\quad \left. - \log(\mathcal{Q}) \cdot (-2 - \phi^{-1} - 1 + 3 + \phi^{-1}) \right] \\
&= -\frac{1}{\mathcal{Q}} \cdot \left[(1+\phi^{-1}) \cdot \log(1+\phi^{-1}) + \log(1-\phi^{-1}) + \phi \log(\phi^{-1}) + (3+\phi^{-1}) \cdot \log(3+\phi^{-1}) \right] = \\
&= -\frac{1}{\mathcal{Q}} \cdot \left[\log(1-\phi^{-1}) - (3+\phi^{-1}) \cdot \log(3+\phi^{-1}) \right].
\end{aligned}$$

Thus:

$$H(\mathcal{C}_3(n, 1, 2)) = \mathcal{C}_3(n, 1, 2) \cdot \left(\log(3 + \phi^{-1}) - \frac{1}{3 + \phi^{-1}} \cdot \log(1 - \phi^{-1}) \right).$$

Acknowledgments

JC and BL acknowledge funding from Spanish Ministry of Science and Innovation under project M2505 (PID2020-113737GB-I00). LL acknowledges funding from projects DYNDEEP (EUR2021-122007), MISLAND (PID2020-114324GB-C22) and through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (MDM-2017-0711), all of them funded by the Spanish Ministry of Science and Innovation via AEI.

- [1] G. H. Hardy, E. M. Wright, D. R. Heath-Brown, and J. H. Silverman, An Introduction to the Theory of Numbers. Oxford University Press, 6th ed., 2008.
- [2] R. L. Graham, D. E. Knuth, and O. Patashnik, Concrete Mathematics: A Foundation for Computer Science. USA: Addison-Wesley Longman Publishing Co., Inc., 2nd ed., 1994.
- [3] D. Angell, Irrationality and Transcendence in Number Theory. Chapman and Hall/CRC, 2022.
- [4] M. Niqui, “Exact arithmetic on the stern-brocot tree,” Journal of Discrete Algorithms, vol. 5, no. 2, pp. 356–379, 2007. 2004 Symposium on String Processing and Information Retrieval.
- [5] J. Vuillemin, “Exact real computer arithmetic with continued fractions,” IEEE Transactions on Computers, vol. 39, no. 8, pp. 1087–1105, 1990.
- [6] I. Niven, H. Zuckerman, and H. Montgomery, An Introduction to the Theory of Numbers. Wiley, 5th ed., 1991.
- [7] B. Adamczewski and Y. Bugeaud, “On the complexity of algebraic numbers, ii. continued fractions,” Acta Mathematica, vol. 195, no. 1, pp. 1–20, 2005.
- [8] A. Y. Khinchin, Continued Fractions. Dover Publications, translated from the third (1961) russian edition. reprint of the 1964 translation. ed., 1997.
- [9] C. Bonnano and S. Isola, “Orderings of the rationals and dynamical systems,” Colloquium Mathematicum, vol. 116, pp. 165 – 189, 2009.
- [10] W. K. Jones G. A, Singerman D, “The modular group and generalized farey graphs,” Groups St. Andrews 1989, vol. 2, p. 316, 1991.
- [11] L. Vepstas, “The minkowski question mark, $psl(2, z)$ and the modular group,” 2004.
- [12] S. Isola, “Continued fractions and dynamics,” Applied Mathematics, vol. 2014, -04-15 2014.
- [13] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, “From time series to complex networks: The visibility graph,” Proceedings of the National Academy of Sciences, vol. 105, no. 13, pp. 4972–4975, 2008.
- [14] B. Luque, L. Lacasa, F. Ballesteros, and J. Luque, “Horizontal visibility graphs: Exact results for random time series,” Physical Review E, vol. 80, no. 4, p. 046103, 2009.
- [15] L. Lacasa, A. Nunez, É. Roldán, J. M. Parrondo, and B. Luque, “Time series irreversibility: a visibility graph approach,” The European Physical Journal B, vol. 85, no. 6, pp. 1–11, 2012.
- [16] A. Nunez, L. Lacasa, E. Valero, J. P. Gómez, and B. Luque, “Detecting series periodicity with horizontal visibility graphs,” International Journal of Bifurcation and Chaos, vol. 22, no. 07, p. 1250160, 2012.
- [17] R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks,” Rev. Mod. Phys., vol. 74, pp. 47–97, Jan 2002.
- [18] B. Luque, L. Lacasa, F. J. Ballesteros, and A. Robledo, “Feigenbaum graphs: A complex network perspective of chaos,” PLOS ONE, vol. 6, pp. 1–8, 09 2011.
- [19] A. M. Núñez, B. Luque, L. Lacasa, J. P. Gómez, and A. Robledo, “Horizontal visibility graphs generated by type-i intermittency,” Physical Review E, vol. 87, no. 5, p. 052801, 2013.
- [20] B. Luque, F. J. Ballesteros, A. Nunez, and A. Robledo, “Quasiperiodic graphs: structural design, scaling and entropic properties,” Journal of Nonlinear Science, vol. 23, no. 2, pp. 335–342, 2013.
- [21] Z. Zhang, B. Wu, and Y. Lin, “Counting spanning trees in a small-world farey graph,” Physica A: Statistical Mechanics and its Applications, vol. 391, no. 11, pp. 3342–3349, 2012.
- [22] J. Farey, “On a curious property of vulgar fractions,” London and Edinburgh Philosophical Magazine and Journal of Science, vol. 47, pp. 385–386, 1816.
- [23] S. B. Guthery, A Motif of Mathematics: History and Application of the Mediant and the Farey Sequence. Docent Press, 6th ed., 2010.
- [24] A. H. Beiler, Recreations in the theory of numbers: The queen of mathematics entertains. Courier Corporation, 1964.
- [25] R. Flanagan, L. Lacasa, and V. Nicosia, “On the spectral properties of feigenbaum graphs,” Journal of Physics A: Mathematical and Theoretical, vol. 53, no. 2, p. 025702, 2019.
- [26] B. Luque and L. Lacasa, “Canonical horizontal visibility graphs are uniquely determined by their degree sequence,” The European Physical Journal Special Topics, vol. 226, no. 3, pp. 383–389, 2017.
- [27] L. Lacasa and W. Just, “Visibility graphs and symbolic dynamics,” Physica D: Nonlinear Phenomena, vol. 374, pp. 35–44, 2018.
- [28] S. Isola, “On the spectrum of farey and gauss maps,” Nonlinearity, vol. 15, no. 5, p. 1521, 2002.
- [29] L. Kocic, L. Stefanovska, and S. Gegovska-Zajkova, “Iterative operators for farey tree,” Kragujevac Journal of Mathematics, vol. 30, pp. 253–262, 2007.
- [30] J. C. Lagarias, “The takagi function and its properties,” 2012.
- [31] B. Luque, L. Lacasa, F. J. Ballesteros, and A. Robledo, “Analytical properties of horizontal visibility graphs in the feigenbaum scenario,” Chaos: An Interdisciplinary Journal of Nonlinear Science, vol. 22, p. 013109, March 1, 2012.

- [32] V. Trifonov, L. Pasqualucci, R. Dalla-Favera, and R. Rabadan, “Fractal-like distributions over the rational numbers in high-throughput biological and clinical data,” Scientific Reports, vol. 1, no. 1, p. 191, 2011.
- [33] B. Luque, F. J. Ballesteros, A. Robledo, and L. Lacasa, Entropy and Renormalization in Chaotic Visibility Graphs, ch. 1, pp. 1–39. John Wiley & Sons, Ltd, 2016.
- [34] I. G. Torre, R. J. Heck, and A. Tarquis, “Multifrac: An imagej plugin for multiscale characterization of 2d and 3d stack images,” SoftwareX, vol. 12, p. 100574, 2020.