HAL
open science

Counting walks by their last erased self-avoiding polygons using sieves

P.-L Giscard

To cite this version:

P.-L Giscard. Counting walks by their last erased self-avoiding polygons using sieves. Discrete Mathematics, Elsevier, 2020, 344 (4), pp.112305. 10.1016/j.disc.2021.112305 . hal-03597248

HAL Id: hal-03597248

https://hal.archives-ouvertes.fr/hal-03597248

Submitted on 4 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Counting walks by their last erased self-avoiding polygons using sieves

Pierre-Louis Giscard[a]

[a] *Université du Littoral Côte d'Opale, EA2597-LMPA*
Laboratoire de Mathématiques Pures et Appliquées Joseph Liouville,
Calais, France.

Abstract

Let G be an infinite, vertex-transitive lattice with degree λ and fix a vertex on it. Consider all cycles of length exactly l from this vertex to itself on G. Erasing loops chronologically from these cycles, what is the fraction $F_p/\lambda^{\ell(p)}$ of cycles of length l whose last erased loop is some chosen self-avoiding polygon p of length $\ell(p)$, when $l \to \infty$? We use combinatorial sieves to prove an exact formula for $F_p/\lambda^{\ell(p)}$ that we evaluate explicitly. We further prove that for all self-avoiding polygons p, $F_p \in \mathbb{Q}[\chi]$ with χ an irrational number depending on the lattice, e.g. $\chi = 1/\pi$ on the infinite square lattice. In stark contrast we current methods, we proceed via purely deterministic arguments relying on Viennot's theory of heaps of pieces seen as a semi-commutative extension of number theory. Our approach also sheds light on the origin of the difference between exponents stemming from loop-erased walk and self-avoiding polygon models, and suggests a natural route to bridge the gap between both.

Keywords: Self-avoiding polygons, Heaps of pieces, Sieves
2000 MSC: 05C38, 05A15, 11N36

⋆This work is supported by the French National Research Agency (ANR) throught the ANR-19-CE40-0006 project ALgebraic COmbinatorics of Hikes On Lattices (ALCOHOL).
Email address: mailto:giscard@univ-littoral.fr (Pierre-Louis Giscard)
URL: http://www-lmpa.univ-littoral.fr/~plgiscard/ (Pierre-Louis Giscard)

1. Context

1.1. Self-avoiding objects and loop-erased walks

The "widely open problem of counting Self-Avoiding Walks (SAWs) and Self-Avoiding Polygons (SAPs) on lattices" (quoting Flajolet & Sedgewick) was first conceived from the study of polymer chemistry, but these objects arise in a wide range of physical and mathematical problems; e.g. as phase boundaries [1] and in percolation clusters [2, 3, 4] or, as B. Mandelbrot observed [5], from the outer frontier of Brownian motion [6]. Because in such models SAWs and SAPs are invariably realised through random processes, the problem of studying and counting them has so far only been attacked with tools from statistical physics and probability theory. Works along these directions have yielded deep insights into renormalisation and conformal mappings such as the relation between self-avoiding curves and the Schramm-Loewner Evolution (SLE) [7, 8], and the result of H. Duminil-Copin and S. Smirnov who proved the value of the connective constant of the honeycomb lattice [9].

One of the most active subfields of this research concerns loop-erased random walks (LERWs) introduced by G. Lawler [10]. Lawler's original motivation was to produce yet another model of random generation of self-avoiding objects. Indeed, removing loops from an ordinary random walk in chronological order–the so called *loop-erasing* procedure–yields the self-avoiding 'skeleton' of the walk. LERWs are now the object of much research on conformal loop ensembles and loop soups [11, 12, 13]. At the heart of this activity are conformally invariant loop measures, one of which [13] associates to a SAP p the proportion of all cycles–also known as closed walks–whose last erased loop is the SAP p. This measure, studied within the framework of probability theory, remains very difficult to evaluate explicitly. Indeed, this requires a complicated mapping to Abelian sand-pile models followed by arduous calculations which were completed only for the shortest SAPs (length less than 6). But the straightforward combinatorial meaning of the measure as a proportion means that a purely deterministic counting method should be able to evaluate it as well. This is indeed so, using number theoretic sieves and heaps of cycles.

1.2. Heaps of cycles and number theory

The theory of heaps of pieces [14, 15, 16], which describes the combinatorics of piles of arbitrary pieces, has now found a remarkable number of

applications throughout mathematics, from combinatorics to group theory. Studying specifically heaps of cycles on graphs, called *hikes*, P.-L. Giscard and P. Rochet showed that these constitute a very natural semi-commutative extension of number theory [17]. This extension comprises all the fundamental objects of number theory (zeta functions, sieve methods) as well as most of the relations between these. In this context, the hikes themselves extend the natural integers, the walks (a.k.a. paths) extend the integers of the form p^k with p prime and SAWs and SAPs extend the primes. This last observation implies that in the semi-commutative framework of the hikes, *the extension of the prime number theorem will give the asymptotic number of SAPs on regular lattices*. What is even more remarkable here is that the set of prime factors of a walk as dictated by the extension of number theory precisely coincides with the loops erased from a walk in Lawler's procedure. The fundamental premise of this work is that the tools offered by number theory are also effective in the study of self-avoiding objects.

1.3. Layout

In this self-contained contribution (definitions in §3), we use purely deterministic sieve techniques for evaluating the asymptotic fraction of all closed walks whose last erased loop is any given SAP p on any finite graph (§4) or infinite vertex-transitive lattice (§5). We arrive at an exact, closed-form, easy to evaluate formula for this fraction, immediately presented in the next section. Within the framework of probability-theory, this fraction is a conformally-invariant loop-measure known and studied by Lawler [13]. The sieves also make it clear that the loop-measure comes with seemingly hitherto unnoticed error-terms (Appendix E), which explain combinatorially the difference between LERW and SAP exponents. A route for overcoming this final hurdle using 1990s work is suggested (§6).

2. Statement of the main result and illustrations

2.1. Main theorem

The main result, which we prove in the remainder of this work is as follows:

Infinite Sieve Theorem. *Let G be an infinite vertex transitive graph of bounded degree λ and let A_G be its adjacency matrix. Let p be a self-avoiding polygon on G. Let $\{G_N^{Tor}\}$ be a the small tori sequence of vertex-transitive*

graphs on N vertices converging to G as $N \to \infty$ (see Appendix A). Let $R(z)$ be the ordinary generating function of all closed walks on G with fixed initial vertex. Then, on G_N^{Tor}, the fraction of all <u>*hikes*</u>*, i.e. heaps of cycles, which are closed walks whose last erased loop is p is given asymptotically for $N \gg 1$ by*

$$\frac{\alpha^N}{N}\frac{F_p}{\lambda^{\ell(p)}}.$$

In this expression $\alpha := \lim_{z\to 1/\lambda^-} \exp\big(\int \frac{1}{z}\big(R(z)-1\big)dz\big)$, $\alpha \in]0,1[$, is well defined and $F_p\lambda^{-\ell(p)}$ designates the fraction of all <u>*closed walks*</u> *defined up to translation on G whose last erased loop is p. This fraction is explicitly given by*

$$\frac{F_p}{\lambda^{\ell(p)}} = \frac{1}{\lambda^{\ell(p)+1}}\ \mathsf{deg}^{\mathrm{T}}.\,\mathrm{adj}\left(\mathsf{I} + \mathsf{C}_G\big|_p.\mathsf{B}_p\right).\,\mathsf{1}, \tag{1}$$

where adj(.) *designates the adjugate operator, B_p is the adjacency matrix of the graph G_p induced by p and its immediate neighbours on G, $\mathsf{1}$ designates the vector full of 1 and $\mathsf{deg} = \mathrm{diag}(\mathsf{B}_p^2)$ is the vector of vertex-degrees on G_p. Finally, $\mathsf{C}_G|_p$ is the restriction to G_p of the matrix $\mathsf{C}_G := \lim_{z\to 1/\lambda^-}(\mathsf{I}-\mathsf{P}_\lambda)\mathsf{R}(z)$, with P_λ the projector onto the eigenspace associated with the dominant eigenvalue and $\mathsf{R}(z) := (\mathsf{I} - z\mathsf{A}_G)^{-1}$ is the resolvent of graph G.*

Corollary 2.1. *Let G be an infinite vertex-transitive lattice of degree λ. Let p be a self-avoiding polygon on G and let $F_p\,\lambda^{-\ell(p)}$ be the fraction of all closed walks whose last erased loop is the SAP p.*

- *If G is a d-dimensional hypercubic lattice, then $F_p \in \mathbb{Q}[1/\pi^{d-1}]$.*
- *If G is the triangular, hexagonal or Kagomé lattice, then $F_p \in \mathbb{Q}[\sqrt{3}/\pi]$.*

The error terms generated by the sieve on infinite graphs are given in Appendix E together with a discussion of their relevance for analytical estimates of the asymptotic growth of the number of SAPs of length ℓ as $\ell \to \infty$. See also the discussion of §6.

From a practical point of view, we observe that the matrix B_p is of size $e(p) \times e(p)$, where $e(p)$ is the number of vertices of G at distance at most 1 from p. Since clearly $\ell(p) \le e(p) \le \lambda\ell(p)$, computing the fraction $F_p/\lambda^{\ell(p)}$ costs $O\big(\ell(p)^3\big)$ operations.

The lattice constant α relates the *densities* of walks and of hikes on the infinite lattice G. On the square lattice, $\alpha = \frac{1}{4}e^{\frac{4C}{\pi}} \simeq 0.8025...$ with C Catalan's constant.

The matrix C_G is easy to obtain on regular graphs because its entries obey the same recursion relations as the graph resolvent. More precisely, let $\left(\mathsf{C}_G\right)_{m,n}$ designate the entry of the matrix corresponding to jumping from vertex m to vertex n. Then $\lambda\left(\mathsf{C}_G\right)_{m,n} = \sum_{i\in\mathcal{N}(n)}\left(\mathsf{C}_G\right)_{m,i} + \sum_{j\in\mathcal{N}(m)}\left(\mathsf{C}_G\right)_{j,n}$, where $\mathcal{N}(n)$ and $\mathcal{N}(m)$ designate the set of vertices that are neighbours to n and m on G, respectively. On the square lattice this implies that C_G has the following explicit expression:

$$\left(\mathsf{C}_G\right)_{ij} = -\frac{1}{\pi}\int_0^\infty \frac{1}{\tau}\left(1 - \left(\frac{\tau - \mathbb{i}}{\tau + \mathbb{i}}\right)^{x_{ij}-y_{ij}}\left(\frac{\tau-1}{\tau+1}\right)^{x_{ij}+y_{ij}}\right)d\tau$$

where $\mathbb{i}^2 = -1$, x_{ij} and y_{ij} are the distance along x and y between vertices i and j of G_p, respectively. In particular if $x_{ij} = y_{ij} = m$ then

$$\left(\mathsf{C}_G\right)_{ij} = -\frac{4}{\pi}\sum_{k=1}^{m-1}\frac{1}{2k+1} = -\frac{2}{\pi}\left(H_{m-\frac{1}{2}} + \log(4)\right),$$

with H_m the mth harmonic number. Explicit expressions for C_G have already been determined on many more lattices owing to its relation with lattice Green's functions and the resistor problem [18, 19].

2.2. Illustrations

We may now illustrate the Infinite Sieve Theorem with concrete results on the infinite square lattice.

▶ The fraction of closed walks whose last erased loop is a given edge e is

$$\frac{F_e}{4^2} = \frac{1}{8} = 0.125.$$

Since a point is connected to 4 edges, this means that an edge is the last erased loop of 1/2 of all closed walks on the square lattice. In the language of the semi-commutative extension of number theory that holds for walks and hikes on graph, this indicates that 1/2 of all closed walks on the square

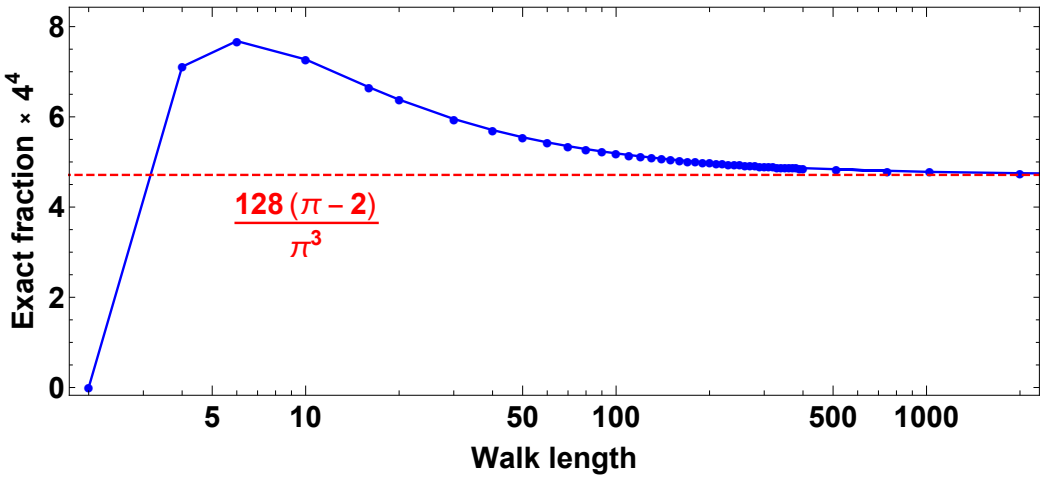


Figure 1: Exact fraction of closed walks of length ℓ on the infinite square lattice whose last erased loop is a 1×1 square as a function of the length of these walks. The exact fraction was obtained from the extension of Viennot's lemma to infinite graphs Proposition 5.1 and Eq. (3). The exact fraction converges to its asymptotic value proportionally with the inverse of the walk length, as dictated by an analysis of the error terms associated with the Infinite Sieve Theorem, see Appendix E.

lattice whose unique right prime divisor is an edge cycle.

► The fraction of closed walks whose last erased loop is a 1×1 square is

$$\frac{F_{1\times 1}}{4^4} = \frac{128(\pi - 2)}{4^4 \pi^3} \simeq 0.0184. \tag{2}$$

This result was first obtained via a complicated mapping between F_p and Abelian sand-pile models [20, 21]. According to the authors this process "becomes very tedious" for longer self-avoiding polygons, so that very few explicit values for F_p have been published in the literature so far, and all pertain to self-avoiding polygons of length $\ell \leq 6$.

See Fig. (1) for an illustration of the convergence of the fraction of closed walks on the infinite square lattice whose last erased loop is a 1×1 square to the above number. Here the extension of Viennot's lemma to infinite graphs yields the ordinary generating function of closed walks whose last erased loop

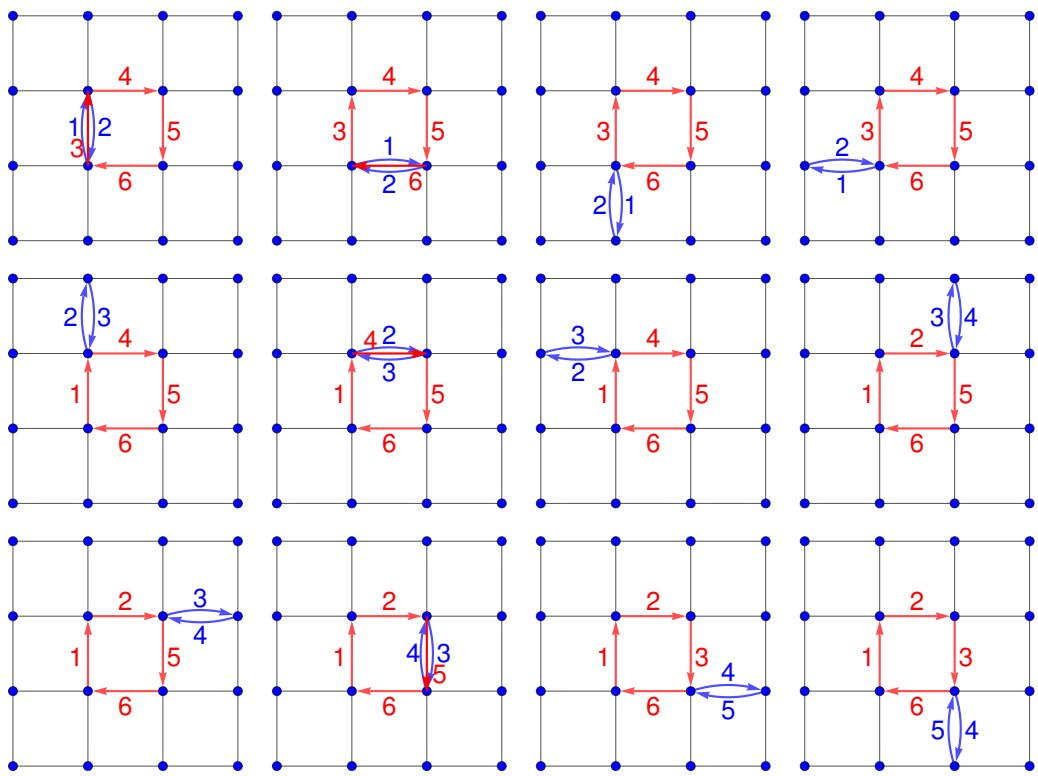


Figure 2: Illustration of the 12 walks of length 6 on the square lattice whose last erased loop is right-oriented 1×1 square cycle (highlighted in red). The numbers next to the highlighted edges indicate the order in which these edges are traversed.

is an oriented 1×1 square, which we denote $R_{1\times1}(z)$, as

$$R_{1\times1}(z) = \frac{1}{256\,\pi^4 z^4}\Big(\big(16z^2-1\big)\;K(16z^2)+E(16z^2)\Big)^2\times \tag{3}$$
$$\Big(\big(1-16z^2\big)\,K(16z^2)^2+2K(16z^2)\,\big(8\pi z^2-E(16z^2)\big)-4\pi^2z^2+E(16z^2)^2\Big),$$
$$= z^4+12\,z^6+144\,z^8+1804\,z^{10}+23464\,z^{12}+\cdots\,,$$

where $K(x) := \int_0^{\pi/2}\big(1-x\sin^2(\theta)\big)^{-1/2}d\theta$ and $E(x) := \int_0^{\pi/2}\big(1-x\sin^2(\theta)\big)^{1/2}d\theta$ are the complete elliptic integrals of the first and second kind, respectively. In Fig. (2) we illustrate the 12 walks of length 6 on the square lattice whose last erased loop is a 1×1 oriented square, as correctly counted by the coefficient of z^6 in $R_{1\times1}(z)$, denoted $[z^6]R_{1\times1}(z) = 12$. Eq. (2) establishes that

asymptotically

$$[z^{2n}]R_{1\times 1}(z) \sim \frac{128(\pi-2)}{4^4\pi^3}\binom{2n}{n}^2, \text{ as } n\to\infty.$$

▶ The fraction of closed walks whose last erased loop is a 1×2 rectangle is

$$\frac{F_{1\times 2}}{4^6} = \frac{32(\pi-8)(\pi-4)(3\pi-8)(3\pi-4)}{4^6\pi^4} \simeq 0.002585.$$

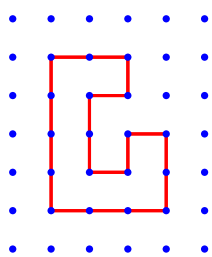
▶ The fraction of closed walks whose last erased loop is a 1×3 rectangle is

$$\frac{F_{1\times 3}}{4^8} = \frac{1024(16-3\pi)(64+3(\pi-12)\pi)(64+27(\pi-4)\pi)(128+\pi(27\pi-124))}{4^8\times 81\pi^7},$$
$$\simeq 0.00035499.$$

▶ The fraction of closed walks whose last erased loop is a 2×2 square is

$$\frac{F_{2\times 2}}{4^8} = \frac{32768(\pi-8)^2(\pi-4)(3\pi-8)^3(9\pi-32)}{4^8\times 81\pi^7} \simeq 0.00044623.$$

▶ As an example of longer SAP, consider:



Then the fraction of all closed walks whose last erased loop is this SAP is

$$\begin{aligned}\frac{F_p}{4^{18}} =& \frac{8388608}{4^{18}\times 8303765625\,\pi^{12}}\Big(172151036713123194478159462 4\\ &- 673302912063441661102915584 0\pi + 1200172504512664753714652774 4\pi^2\\ &- 1289567574563800792193984102 4\pi^3 + 9303982639359984674575220736\pi^4\\ &- 4748903115679537036020154368\pi^5 + 1758418560456019196044640256\pi^6\\ &- 4759107232844883759700377 60\pi^7 + 9343026756136228129413120 0\pi^8\\ &- 1297345994115522517270800 0\pi^9 + 1209211981439562793530000\pi^{10}\\ &- 6790636334966358352500 0\pi^{11} + 1736896666805181140625\pi^{12}\Big),\\ \simeq& 7.7644\times 10^{-9}.\end{aligned}$$

This result is well beyond what is realistically achievable from conformally invariant measures mapped to Abelian sandpile models.

Fractions of walks whose last erased loop is a much longer SAP are easily obtained numerically, costing no more than $O(\ell(p)^3)$ to compute as outlined earlier. For example, the fraction of closed walks whose last erased loop in a 70×70 square is

$$\frac{F_{70\times 70}}{4^{280}} \simeq 1.5236 \times 10^{-108}. \tag{4}$$

Analytically speaking these fractions become very involved very quickly as a function of SAP length and there is no reason to believe that there exists a simpler expression for them than that given by Eq. (1) of the Infinite Sieve Theorem. For example, the analytical expression for the fraction of closed walks whose last erased loop is a 6×6 square already involves sums and products of up to 16-digits prime integers. In fully expanded form this fraction involves a 67-digits prime integer (!). Overall, we have calculated the fractions for over 100,700,000 SAPs *analytically*, and for more than 3,480,000,000 SAPs numerically on the square lattice. These results as well as the accompanying algorithm will soon be presented in a separate contribution. The hexagonal and triangular lattices will be also be treated.

3. Definitions

3.1. Hikes and related objects

In the general setting, we consider (weighted di)graphs $G = (\mathcal{V}; \mathcal{E})$ with $N = |\mathcal{V}|$ nodes and $M = |\mathcal{E}|$ edges, both of which may be infinite but the degree of G must be bounded. The ordinary adjacency matrix of G is denoted A_G or simply A. If G is weighted then the entry A_{ij} is the weight of the edge e_{ij} from i to j if this edge exists, and 0 otherwise. The labelled adjacency matrix of G is denoted W and its entries are formal variables belonging to the Cartier-Foata monoid, $\mathsf{W}_{ij} = e_{ij}$.

A *induced subgraph* H of G, also called simply a *subgraph* of G and denoted $H \prec G$, is a set of vertices $\mathcal{V}_H \subseteq \mathcal{V}$ together with the set of all edges linking these vertices in G, $\mathcal{E}_H = \{e_{ij} \in \mathcal{E} : i, j \in \mathcal{V}_H\}$.

A *walk* w of length $\ell(w)$ from v_i to v_j on G is a sequence $w = e_{ii_1} e_{i_1 i_2} \cdots e_{i_{\ell-1} j}$ of ℓ contiguous edges. The walk w is *open* if $i \neq j$ and *closed* otherwise.

A *simple cycle*, also known in the literature under the names *loop*, *cycle*, *elementary circuit* and *Self-Avoiding Polygon* or *SAP*, is a closed walk

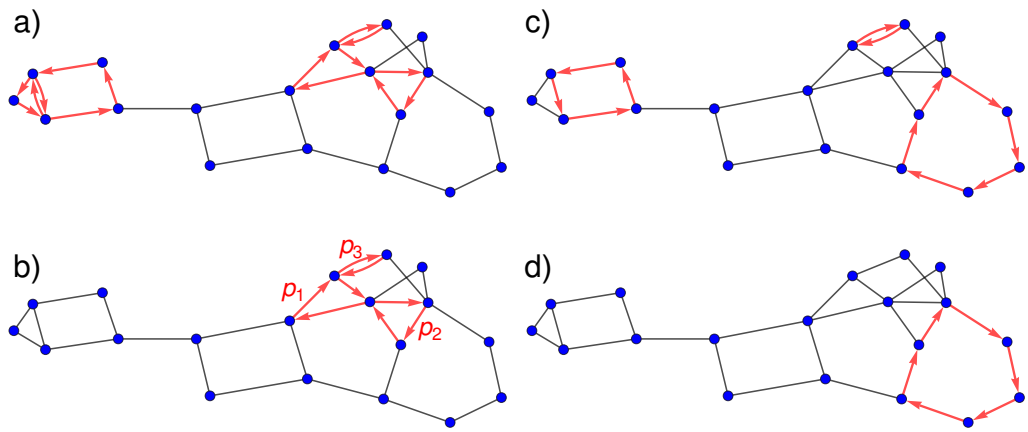


Figure 3: Illustration of the main objects of the hike family on a graph G: **a)** a hike comprising two walks; **b)** a walk $w = p_2p_3p_1 \equiv p_3p_2p_1$; **c)** a self-avoiding hike comprising three vertex disjoint simple cycles; **d)** a simple cycle. The walk w has a unique right prime divisor, here the triangle p_1. In the Cartier-Foata monoids, the starting point of this walk can be any of the three vertices of p_1 and p_1 will always be the last erased loop of the walk.

$w = e_{ii_1}e_{i_1i_2}\cdots e_{i_{\ell-1}i}$ which does not cross the same vertex twice, that is, the indices $i, i_1, \ldots, i_{\ell-1}$ are all different. Two simple cycles differing only by orientation are considered distinct (Rule 1), but two simple cycles differing only by their starting point are taken to be identical (Rule 2). The necessity of these choices and of the definition of hikes below can be found in Cartier and Foata's foundational work [14]. From now on, we employ the letter $\mathcal{P}$ to designate sets of simple cycles, in particular $\mathcal{P}_G$ will be the set of all simple cycles on a graph G.

The central objects of the present work are *hikes*, a hike h being an unordered collection of vertex-disjoint closed walks. Hikes can be also be seen as equivalence classes on words $W = p_{i_1}p_{i_2}\cdots p_{i_n}$ over the alphabet of simple cycles p_i of a graph. Two words W and W' are equivalent if and only if W' can be obtained from W through allowed permutations of consecutive simple cycles. In this context, two simple cycles are allowed to commute if and only if they are vertex disjoint $\mathcal{V}(p_i) \cap \mathcal{V}(p_j) = \emptyset \iff p_ip_j = p_jp_i$.

For example, if p_1 and p_2 commute but neither commute with p_3, then p_1p_2 and p_2p_1 represent the same hike, but $p_1p_3p_2$ and $p_2p_3p_1$ are distinct hikes.

The letters $p_{i_1}, \cdots, p_{i_n}$ found in a hike h are called its *prime* divisors. This terminology is due to the observation that simple cycles obey the defining property of prime elements in the semi-commutative Cartier-Foata monoid $\mathcal{H}$ of hikes equipped with the concatenation. In the context of Viennot's theory of heaps of pieces, hikes are heaps of simple cycles modulo Rules 1 and 2. In this work, we use the terminologies "primes" and "SAPs" interchangeably.

Two special types of hikes will be important for our purpose here:

A *self-avoiding hike* is a hike all prime factors of which commute with one another. In other terms, it is collection of vertex-disjoint simple cycles. If $\mathcal{P}$ designates a set of simple cycles, we designate $\mathcal{P}^{s.a.}$ the set of all <u>s</u>elf-<u>a</u>voiding hikes that can be built from cycles of $\mathcal{P}$.

In Viennot's terminology, a hike is a *walk* if and only if it is a pyramid of cycles, i.e. a heap of cycle with a unique top cycle. Equivalently, this means that a hike is a *walk* if and only if it has a unique right prime divisor p [17]. In this situation we say that p right divides the walk w, denoted $p|_r w$ or that w is a (left) multiple of p. Remarkably, these notions are also identical with those produced by G. Lawler's loop erasing procedure [10]: in this framework the unique right prime divisor of a closed walk is the last erased loop of this walk. As an example, let p_1, p_2 and p_3 be three simples cycles such that p_2 and p_3 commute with one another, but p_1 commutes with neither p_2 nor p_3. Then the hike $h = p_2 p_3 p_1 \equiv p_3 p_2 p_1$ is a walk whose unique right prime divisor is p_1, i.e. p_1 is the last erased loop of h. This is illustrated in Fig. (3).

3.2. Functions on hikes

A function on hikes is a complex-valued function $f : \mathcal{H} \mapsto \mathbb{C}$. The most important example here will be that of *rank function*.

A rank function, is a function $\rho : \mathcal{H} \mapsto \mathbb{R}$ that is totally additive over the hikes, $\rho(hh') = \rho(h) + \rho(h')$ and which respects the divisibility order, i.e. $h \leq h' \Rightarrow \rho(h) \leq \rho(h')$. The reverse implication does not hold in general. Examples of rank function include the length and the number of self-crossings of a hike. In general, we will denote $\rho(h)$ the rank of h as per the rank function $\rho(.)$ and will denote ϱ any given rank, i.e. value taken by the function ρ. In the case of the length function, $\ell(h)$ is the length of hike and l denotes a length. The set of hikes with a given rank ϱ is $\mathcal{H}_\varrho := \{h \in \mathcal{H},\ \rho(h) = \varrho\}$. The cardinality of this set is denoted $|\mathcal{H}_\varrho| := \mathrm{card}(\mathcal{H}_\varrho)$, if the graph is weighted this is understood to mean the total weight carried by hikes of rank ϱ.

A function $f : \mathcal{H} \mapsto \mathbb{C}$ on hikes is associated with a formal series $\sum_{h\in\mathcal{H}} f(h)h$. This series is rarely accessible as such, rather linear algebra provides tools to access a related ordinary generating function $F(z) := \sum_{h\in\mathcal{H}} f(h)z^{\ell(h)}$. Important examples of functions on hikes and their related series include:

The *zeta function* on hikes is the identity function over the hikes $\forall h \in \mathcal{H}, \zeta(h) = 1$. The associated ordinary generating function will be denoted $\zeta(z)$, it is given by

$$\zeta(z) = \sum_{h\in\mathcal{H}} z^{\ell(h)} = \frac{1}{\det\big(\mathsf{I} - z\mathsf{A}_G\big)},$$

where A_G is the adjacency matrix of the graph G on which the hikes live.

The *Möbius function* on hikes is the convolution inverse of the zeta function. We have

$$\forall h \in \mathcal{H},\ \mu(h) = \begin{cases} (-1)^{\Omega(h)}, & \text{if } h \text{ is self-avoiding} \\ 0, & \text{otherwise} \end{cases}$$

Here $\Omega(h)$ is the *prime factor counting function*, its count the number of prime right-divisors of h. The associated ordinary generating function will be denoted $\mu(z)$, it is given by $\mu(z) = \det\big(\mathsf{I} - z\mathsf{A}_G\big)$.

The *walk von Mangoldt* function on hikes Λ is defined as the number $\Lambda(h)$ of contiguous is the number of possible contiguous rearrangements of the edges in h, obtained without permuting two edges with the same starting point. This is equivalent to

$$\forall h \in \mathcal{H},\ \Lambda(h) = \begin{cases} \ell(p), & \text{if } h \text{ is a walk with unique right prime divisor } p, \\ 0, & \text{otherwise.} \end{cases}$$

On a graph with N vertices, the ordinary generating function associated with the walk von Mangoldt function is given by $\Lambda(z) = \mathrm{Tr}\big(\mathsf{I} - z\mathsf{A}_G\big) - N$.

4. The asymptotics of hikes and walks on finite graphs

The aim of this section is to develop sieving tools to asymptotically count hikes satisfying certain properties on finite graphs. The main results here will

be the Finite Sieve Theorem and its length corollary. Before we state and prove these results, there is an important precedent to be found in Viennot's work [15, 22], which provides the ordinary generating functions of hikes which are closed walk multiples of any chosen prime p, i.e. whose last erased loop is the SAP p. Since the asymptotic expansion of this result is among the results obtained below, we start by recalling Viennot's result.

Viennot's lemma [15, 22]. *Let G be a finite graph. Let p be a prime on this graph and let $W_p := \sum_{w:\, p|_r w} w$ be the formal series of closed walks whose unique right prime divisor is p. Then*

$$W_p = \frac{\det\left(\mathsf{I} - \mathsf{W}_{G\backslash p}\right)}{\det\left(\mathsf{I} - \mathsf{W}_G\right)} p.$$

where $\mathsf{W}_{G\backslash p}$ and W_G designate the labelled adjacency matrices of $G\backslash p$ and G, respectively.

Viennot gave a beautiful bijective proof of this result in [15]. There are at least four more proofs, one of which is the spirit of sieves and is provided below. The extension of Viennot's lemma to infinite graphs is obtained in Section 5.1.

Proof. Let $\mathcal{H}_{[p,.]\neq 0}$ be the set of hikes none of whose connected components commute with p. Clearly, for all $h \in \mathcal{H}_{[p,.]\neq 0}$, hp is a walk multiple of p and we need only determine $\mathcal{H}_{[p,.]\neq 0}$ to obtain the lemma. But this set is the set of all hikes minus the set of hikes such that at least one connected component commutes with p. Such a component must be divisible on the right by prime(s) p' commuting with p, hence $\mathcal{V}(p') \subseteq G\backslash p$. Let $\mathcal{P}_{G\backslash p}$ be the set of all such primes. The exclusion-inclusion principle then yields $\sum_{h\in\mathcal{H}_{[p,.]\neq 0}} h = \sum_{d\in\mathcal{P}^{s.a}_{G\backslash p}} \mu(d) M(d)$, with the convention that $1 \in \mathcal{P}^{s.a}_{G\backslash p}$ and $M(d)$ is the formal series of the left-multiples of d. For any hike this is $M(h) = \det(\mathsf{I} - \mathsf{W})^{-1} h$ since all hikes multiplied by h on the right are left multiples of h. Then

$$\sum_{h\in\mathcal{H}_{[p,.]\neq 0}} h = \frac{\sum_{d\in\mathcal{P}^{s.a}_{G\backslash p}} \mu(d) d}{\det(\mathsf{I} - \mathsf{W})} = \frac{\det(\mathsf{I} - \mathsf{W}_{G\backslash p})}{\det(\mathsf{I} - \mathsf{W})}.$$

The series W_p is obtained upon multiplying the above by p on the right. □

The sieve based proof of Viennot's lemma suggests a wider family of results to count exactly or asymptotically families of hikes satisfying chosen properties on finite graphs. Concentrating on asymptotic expansions, we have:

Finite Sieve Theorem. *Let G be a finite (weighted di)graph with adjacency matrix A. Let H be an induced subgraph of G and let $\mathcal{P}_H$ be the set of primes on H. Let $\rho : \mathcal{H} \mapsto \mathbb{R}$ be a rank function on hikes such that $|\mathcal{H}_\varrho| = \lambda^\varrho f(\varrho)$ with λ a real constant and $f(.)$ a bounded function such that $\lim_{\varrho\to\infty} f(\varrho)$ exists.*

Then the number (weight)[1] *$S(\mathcal{H}_\varrho, \mathcal{P}_H)$ of <u>hikes</u> of rank $\rho(h) = \varrho$ which are not multiples of primes on H is asymptotically given by*

$$\frac{S(\mathcal{H}_\varrho, \mathcal{P}_H)}{|\mathcal{H}_\varrho|} \sim \sum_{d\in\mathcal{P}_H^{s.a}} \mu(d)\, \lambda^{-\rho(d)}, \quad as \quad \varrho \to \infty.$$

The Finite Sieve Theorem's most important application here will be with the length rank-function and sieving subgraph $H = G\backslash p$ for p a prime. This provides the asymptotic expansion of Viennot's lemma:

Length corollary. *Let G be a finite (weighted di)graph with adjacency matrix A and dominant eigenvalue λ, which we assume to be unique.*[2] *Let p be a simple cycle or a simple path on G of length $\ell(p)$ and let $S(\mathcal{H}_l, \mathcal{P}_{G\backslash p})$ be defined as in the Finite Sieve Theorem.*

Then $S(\mathcal{H}_l, \mathcal{P}_{G\backslash p})$ is equal to the number (weight) of <u>closed walks</u> of length $\ell(w) = l$ on G whose unique right prime divisor is p and is asymptotically given by

$$\frac{S(\mathcal{H}_l, \mathcal{P}_{G\backslash p})}{|\mathcal{H}_l|} \sim \frac{1}{\lambda^{\ell(p)}} \det\left(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_{G\backslash p}\right) \quad as \quad \ell \to \infty.$$

Let $\mathrm{Err}(\mathcal{H}_l, \mathcal{P}_{G\backslash p})$ be the difference between the two terms above. Let $f(l) := [z^l]\zeta(z/\lambda)$ be the coefficient of z^l in the expansion of $\zeta(z/\lambda)$. Then f is

[1] The notation $S(\mathcal{H}_\rho, \mathcal{P}_H)$ for this quantity is employed in keeping with conventions from sieve theory.

[2] The theorem extends if λ is not unique upon replacing λ^{-1} by λ^{-g} with g its multiplicity.

bounded, $\lim_{l\to\infty} f(l)$ *exists, and*

$$\mathrm{Err}(\mathcal{H}_l, \mathcal{P}_{G\backslash p}) = \frac{1}{\lambda^{\ell(p)}} \sum_{k\geq 0}^{\infty} \left(\frac{\nabla^k[f]\big(l-\ell(p)\big)}{f(l)\,\lambda^k\,k!} - \delta_{k,0} \right) \det{}^{(k)}\Big(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_{G\backslash p}\Big).$$

with $\delta_{k,0}$ *the Kronecker delta and* ∇ *is the backward difference operator*[3]*. Here,* $\det^{(k)}(\mathsf{I}-\frac{1}{\lambda}\mathsf{A}_{G\backslash p})$ *stands for the* k*th derivative of* $\det(\mathsf{I}-z\mathsf{A}_{G\backslash p})$ *evaluated in* $z = 1/\lambda$*.*

Proof of the Finite Sieve Theorem. The proof relies on an inclusion-exclusion principle in the poset of the hikes ordered by right-divisibility, see [17] for an overview of this poset. Let $\mathcal{P} \subsetneq \mathcal{H}$ be a set of primes and $\mathcal{P}^{\text{s.a.}}$ the set of all self-avoiding hikes constructible from $\mathcal{P}$. Let $S(\mathcal{H}_\varrho, \mathcal{P})$ be the number (weight) of hikes in $\mathcal{H}_\varrho$ which are not right-divisible by any prime of $\mathcal{P}$. The inclusion-exclusion principle–here the extension to hikes of the sieve of Erathostenes-Legendre–yields

$$S(\mathcal{H}_\varrho, \mathcal{P}) = \sum_{d\in\mathcal{P}^{\text{s.a.}}} \mu(d)|\mathcal{M}_d|,$$

with $|\mathcal{M}_d|$ the number of multiples of d in $\mathcal{H}_\rho$ and $\mu(d)$ is the Möbius function on hikes.

In order to progress, we seek a multiplicative function prob(.) such that $|\mathcal{M}_d| = \mathrm{prob}(d)|\mathcal{H}_\varrho| + r(d)$. In this expression, $\mathrm{prob}(d)$ approximates the probability that a hike taken uniformly at random in $\mathcal{H}_\varrho$ is right-divisible by d. If edge-weights are present, the hikes are not all uniformly probable but follow a distribution dependent on these weights. No knowledge of this distribution is required here. Similarly, $m(d) = \mathrm{prob}(d)|\mathcal{H}_\varrho|$ is the expected number of multiples of d in $\mathcal{H}_\varrho$. Finally, $r(d)$ is the associated error term, arising from the fact that $|\mathcal{M}_d|$ is not truly multiplicative. Supposing that we can identify the $m(.)$ function, we would obtain

$$S(\mathcal{H}_\varrho, \mathcal{P}) = \sum_{d\in\mathcal{P}^{\text{s.a.}}} \mu(d)m(d) + \sum_{d\in\mathcal{P}^{\text{s.a.}}} \mu(d)r(d).$$

[3]This operator acts on a function F of a variable x as $\nabla[F](x) := F(x) - F(x-1)$, and $\nabla^k[F](x)$ designates the kth iteration of operator ∇, i.e. $\nabla^k[F](x) = \sum_{j=0}^{k}(-1)^j\binom{k}{j}F(x-j)$

Contrary to number theory, the first term does not admit any simpler form without further assumptions on $\mathcal{P}$. This is because of the possible lack of commutativity between some elements of $\mathcal{P}$. We note however that since $\mu(d)$ is non-zero if and only if d is self-avoiding, and since we have required that $m(.)$ be multiplicative,[4] then it follows that the first term is determined solely from the values of $m(.)$ over the primes of $\mathcal{P}$.

We therefore turn to determining $m(p)$ for p prime. The set of left-multiples of p in $\mathcal{H}$ is $\mathcal{M}_p := \{hp,\ h \in \mathcal{H}\}$, hence in bijection with the set $\mathcal{H}$. Thus, the number of left-multiples of p in $\mathcal{H}_\varrho$, is exactly $|\mathcal{H}_{\varrho-\rho(p)}|$. Then

$$\text{prob}(p) + \frac{r(p)}{|\mathcal{H}_\varrho|} = \frac{|\mathcal{H}_{\varrho-\rho(p)}|}{|\mathcal{H}_\varrho|}.$$

Seeking the best possible probability function $\text{prob}(\rho)$, let us suppose that we can choose this function such that the error term of the above equation vanishes in the limit $\rho \to \infty$. If this is true, then we obtain

$$\text{prob}(\rho) = \lim_{\rho\to\infty} \frac{|\mathcal{H}_{\varrho-\rho(p)}|}{|\mathcal{H}_\varrho|}.$$

In order to progress, we have to make an important assumption regarding the cardinality of the set $\mathcal{H}_\varrho$:

Assumption 4.1. *There exists a scaling constant λ and bounded function $f : \mathbb{R} \mapsto \mathbb{R}$ such that $\lim_{\varrho\to\infty} f(\varrho)$ exists and for $\varrho \in \mathbb{N}^*$*

$$|\mathcal{H}_\varrho| = \lambda^\varrho f(\varrho).$$

In the case of the length rank function, this assumption is actually a proposition:

Proposition 4.2. *Let G be a finite (weighted di)graph with dominant eigenvalue λ of multiplicity g. Let $\mathcal{H}_\ell := \{h \in \mathcal{H} :\ \ell(h) = \ell\}$ be set of all hikes on G of length ℓ. Then, there exists a bounded function $f : \mathbb{N} \mapsto \mathbb{R}$ such that $\lim_{\ell\to\infty} f(\ell)$ exists and for $\ell \in \mathbb{N}^*$ we have exactly*

$$|\mathcal{H}_\ell| = \lambda^{g\ell} f(\ell).$$

[4]But not necessarily totally multiplicative.

Proof. This follows directly from the ordinary zeta function on hikes $\zeta(z) = \det(\mathsf{I} - z\mathsf{A})^{-1}$, from which we have

$$|\mathcal{H}_\ell| = [z^\ell]\left(\frac{1}{\det(\mathsf{I} - z\mathsf{A})}\right) = \sum_{i_1,\cdots,i_N \vdash \ell} \lambda_1^{i_1}\lambda_2^{i_2}\cdots\lambda_N^{i_N} = \lambda^\ell \sum_{i_1,\cdots,i_N \vdash \ell} \lambda^{i_1-\ell}\lambda_2^{i_2}\cdots\lambda_N^{i_N}$$

where the sums run over all non-negative values of $i_j \geq 0$ such that $\sum_j i_j = \ell$ and $\lambda \equiv \lambda_1$ is the eigenvalue of the graph with the largest absolute value. We assume for the moment that λ is unique and let $f(\ell) := \sum_{i_1,\cdots,i_N \vdash \ell} \lambda^{i_1-\ell}\lambda_2^{i_2}\cdots\lambda_N^{i_N} = [z^\ell]\zeta(z/\lambda)$. This function is clearly bounded and

$$\lim_{\ell\to\infty} f(\ell) = \lim_{z\to 1/\lambda^-} (1 - z\lambda)\zeta(z),$$

exists and is finite. If $|\lambda|$ is not unique and has multiplicity g, then the scaling constant for the number of hikes becomes λ^g and then $f(\ell) = [z^\ell]\zeta(z/\lambda^g)$. $\square$

Proceeding with Assumption 4.1–or in the case of the length rank function Proposition 4.2–the existence of the limit for f gives

$$\text{prob}(p) = \lim_{\varrho\to\infty} \frac{\lambda^{\varrho-\rho(p)} f\big(\varrho - \rho(p)\big)}{\lambda^\varrho f(\varrho)} = \lambda^{-\rho(p)}.$$

The prob(.) function is multiplicative over the primes as desired and yields $m(p) = |\mathcal{H}_\varrho|\lambda^{-\rho(p)}$. The associated error term is

$$r(\varrho) = |\mathcal{H}_{\varrho-\rho(p)}| - |\mathcal{H}_\varrho|\lambda^{-\rho(p)} = \lambda^{\varrho-\rho(p)}\Big(f\big(\varrho - \rho(p)\big) - f(\varrho)\Big).$$

To establish the validity of these results, we need only verify that they are consistent with our initial supposition concerning the error term, namely that $r(p)/|\mathcal{H}_\varrho|$ vanishes in the limit $\varrho \to \infty$. The existence of the limit of f implies $\lim_{\varrho\to\infty} |f\big(\varrho - \rho(p)\big) - f(\varrho)| = 0$ and therefore that

$$\lim_{\varrho\to\infty} \frac{r(p)}{|\mathcal{H}_\varrho|} = \lim_{\varrho\to\infty} \lambda^{-\rho(p)}\Big(f\big(\varrho - \rho(p)\big) - f(\varrho)\Big) = 0,$$

as required.

We are now ready to proceed with general self-avoiding hikes. Let $d = p_1 \cdots p_{\Omega(d)}$ be self-avoiding. Since m is multiplicative and the rank function

is totally additive over $\mathcal{H}$, $m(d) = \prod_i m(p_i) = \lambda^{-\sum_i \rho(p_i)} = \lambda^{-\rho(d)}$. The associated error term follows as

$$r(d) = |\mathcal{H}_{\varrho-\rho(d)}| - |\mathcal{H}_\varrho|\lambda^{-\rho(d)} = \lambda^{\varrho-\rho(d)}\big(f\big(\varrho-\rho(d)\big) - f(\varrho)\big).$$

Inserting these forms for $m(d)$ and $r(d)$ in the sieve yields

$$S(\mathcal{H}_\varrho, \mathcal{P}) = |\mathcal{H}_\varrho| \sum_{d\in\mathcal{P}^{\text{s.a.}}} \mu(d)\lambda^{-\rho(d)} + \lambda^\varrho \sum_{d\in\mathcal{P}^{\text{s.a.}}} \mu(d)\lambda^{-\rho(d)}\big(f(\varrho-\rho(d)) - f(\varrho)\big). \tag{5}$$

We can now progress much further on making an additional assumption concerning the nature of the prime set $\mathcal{P}$. We could consider two possibilities: i) that $\mathcal{P}$ is the set of all primes on an induced subgraph $H \prec G$; or ii) that $\mathcal{P}$ is a cut-off set, e.g. one disposes of all the primes of length $\ell(p) \leq \Theta$.

In the situation where all primes commute with one-another–i.e. when Viennot's theory of heaps of pieces reduces to number theory [17]–then one may have both i) and ii) simultaneously. This is because in the semi-commutative extension of number theory that holds on the monoid of hikes, coprimality extends to being vertex-disjoint [17]. Therefore, requiring all primes to commute is equivalent to forcing the graph G to be made of disjoint oriented simple cycles. Consequently, we can choose H to be the induced subgraph of G comprising all simple cycles of length up to some cut-off Θ and both situations i) and ii) are realised. For this reason, number theoretic sieves benefit from the advantages of both situations: ii) guarantees that sieves can be used to obtain estimates on the number of primes, while i) allows these estimates to be computable. On general graphs however, i) and ii) are not compatible and while ii) could be used to obtain direct estimates for the number of primes of any length, a problem of great interest, this actually makes the sieve NP-hard to implement. We therefore focus on the first situation.

Let $H \prec G$ be an induced subgraph of the graph G and let that $\mathcal{P} \equiv \mathcal{P}_H$ be the set of all primes (here simple cycles) on H. To conclude the proof we need only show that the error term of Eq. (5) is asymptotically dominated by the first term $\sum_{d\in\mathcal{P}_H^{\text{s.a.}}} \mu(d)\lambda^{-\rho(d)}$. To this end, we note that since H is

finite[5]

$$\lambda^{\varrho} \sum_{d \in \mathcal{P}_H^{\text{s.a.}}} \mu(d) \lambda^{-\rho(d)} \big(f(\varrho - \rho(d)) - f(\varrho) \big),$$

is a sum involving finitely many self-avoiding hikes d. In addition, given that $\lim_{\varrho \to \infty} f(\varrho)$ exists (either by Assumption 4.1 or by Proposition 4.2 for the length rank function), $\lim_{\varrho \to \infty} f(\varrho - \rho(d)) - f(\varrho) = 0$ as long as $\rho(d)$ is finite, which is guaranteed by the finiteness of H. We have consequently established that the error term comprises finitely many terms, each of which vanishes in the $\varrho \to \infty$ limit. As a corollary, the first term is asymptotically dominant:

$$\frac{S(\mathcal{H}_\varrho, \mathcal{P}_H)}{|\mathcal{H}_\varrho|} \sim \sum_{d \in \mathcal{P}_H^{\text{s.a.}}} \mu(d) \lambda^{-\rho(d)} \quad \text{as} \quad \varrho \to \infty,$$

where we assume that $|\mathcal{H}_\varrho| \neq 0$.

We now turn to establishing the length corollary of the Finite Sieve Theorem. We are specifically looking for the number of closed walks which are multiples of a prime p. To this end, we need only choose H correctly. Let h be a hike, for $w = hp$ to be a walk of length l, then h must have length $l - \ell(p)$ and be such that none of its right-prime divisor commutes with p. The sieve must thus eliminate all hikes h which are left-multiples of primes *commuting* with p. Observe that all such primes are on $H = G \backslash p$. Consequently the Finite Sieve Theorem yields, for $|\mathcal{H}_{l-\ell(p)}| \neq 0$,

$$\begin{aligned} S(\mathcal{H}_l, \mathcal{P}_{G\backslash p}) = |\mathcal{H}_{l-\ell(p)}| \sum_{d \in \mathcal{P}^{\text{s.a.}}} \mu(d) \lambda^{-\ell(d)} \\ + \lambda^{l-\ell(p)} \sum_{d \in \mathcal{P}^{\text{s.a.}}} \mu(d) \lambda^{-\ell(d)} \big(f(l - \ell(p) - \ell(d)) - f(l - \ell(p)) \big), \end{aligned}$$

where λ is now the graph dominant eigenvalue per Proposition 4.2. The asymptotically dominant term is a sum over all the self-avoiding hikes on $G \backslash p$, each with coefficient $\mu(d)\lambda^{-\ell(d)}$ and is equal to $\det(\mathsf{I} - \lambda^{-1} \mathsf{A}_{G\backslash p})$. Since furthermore $|\mathcal{H}_{l-\ell(p)}| = |\mathcal{H}_l| \lambda^{-\ell(p)} f\big(l - \ell(p)\big) / f(l)$, we have asymptotically for $l \gg 1$

$$\frac{S(\mathcal{H}_l, \mathcal{P}_{G\backslash p})}{|\mathcal{H}_l|} \sim \lambda^{-\ell(p)} \det \left(\mathsf{I} - \frac{1}{\lambda} \mathsf{A}_{G\backslash p} \right),$$

[5] G is finite and so are all its induced subgraphs.

while the error terms is

$$\begin{aligned}\mathrm{Err}(\mathcal{H}_l, \mathcal{P}_{G\backslash p}) &:= \frac{S(\mathcal{H}_l, \mathcal{P}_{G\backslash p})}{|\mathcal{H}_l|} - \lambda^{-\ell(p)} \det\left(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_{G\backslash p}\right), \\ &= \lambda^{-\ell(p)}\big(f\big(l-\ell(p)\big)/f(l) - 1\big) \det\left(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_{G\backslash p}\right) + \\ &\qquad \frac{\lambda^{-\ell(p)}}{f(l)} \sum_{d\in\mathcal{P}^{\text{s.a.}}} \mu(d)\lambda^{-\ell(d)}\big(f(l-\ell(p)-\ell(d)) - f(l-\ell(p))\big).\end{aligned}$$

The last line can be brought in determinantal form as well, since

$$f(l-\ell(p)-\ell(d)) - f(l-\ell(p)) = \sum_{k\geq 1}^{\ell(d)} \frac{\nabla^k[f]\big(l-\ell(p)\big)}{k!}\big(\ell(d)\big)_{(k)},$$

with $(a)_{(k)} := \prod_{i=0}^{k-1}(a-i)$ the falling factorial and ∇ the backward difference operator. This operator acts on a function F of a variable x as $\nabla[F](x) := F(x) - F(x-1)$, and $\nabla^k[F](x)$ designates the kth iteration of operator ∇, i.e. $\nabla^k[F](x) = \sum_{j=0}^{k}(-1)^j\binom{k}{j}F(x-j)$. Then

$$\begin{aligned}&\frac{\lambda^{-\ell(p)}}{f(l)} \sum_{d\in\mathcal{P}^{\text{s.a.}}} \mu(d)\lambda^{-\ell(d)}\big(f(l-\ell(p)-\ell(d)) - f(l-\ell(p))\big). \\ &\qquad = \lambda^{-\ell(p)} \sum_{d\in\mathcal{P}_H^{\text{s.a.}}} \mu(d)\lambda^{-\ell(d)} \sum_{k\geq 1}^{\ell(d)} \frac{\nabla^k[f]\big(l-\ell(p)\big)}{f(l)\,k!}\big(\ell(d)\big)_{(k)}.\end{aligned}$$

The upper limit of the inner sum over k can be extended to ∞ since all terms with $k > \ell(d)$ are nul. Noting that $\big(\ell(d)\big)_{(k)} z^{\ell(d)} = z^k \frac{d^k}{dz^k} z^{\ell(d)}$, this allows us to invert the two sums, yielding

$$\lambda^{-\ell(p)} \sum_{k\geq 1}^{\infty} \frac{\nabla^k[f]\big(l-\ell(p)\big)}{f(l)\lambda^k\,k!}\ \det{}^{(k)}\Big(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_{G\backslash p}\Big).$$

Now setting $k=0$ in the above recovers $\lambda^{-\ell(p)}\big(f\big(l-\ell(p)\big)/f(l)-1\big)\det\big(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_{G\backslash p}\big)$ with the exception of the -1 in the parenthesis, which can be introduced as $-\delta_{k,0}$. This establishes the Finite Sieve Theorem and its length corollary. $\square$

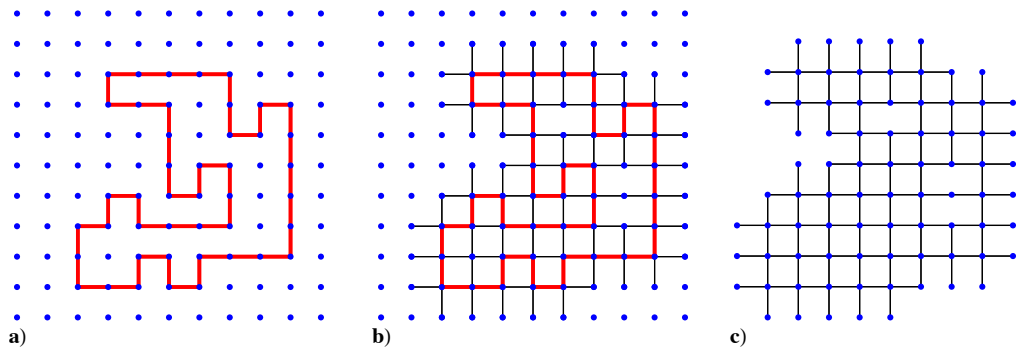


Figure 4: **a)** A self-avoiding polygon p on the square lattice; **b)** Polygon p and all the edges that have at least one endpoint on p; **c)** Graph G_p.

5. Infinite graphs

5.1. Viennot's lemma on infinite graphs

As we have seen, when counting closed walk multiples of a prime according to their length on finite graphs[6], the Finite Sieve Theorem produces the asymptotics of Viennot's lemma. To put the infinite graphs results in context we thus start by proving that this lemma extends to infinite graphs with bounded degree. Here, we state only the ordinary generating function version of the extension here, that for formal series on hikes is provided in Appendix B.

Proposition 5.1. *Let G be an infinite (weighted di)graph with bounded degree. Let p be a prime on G, with support $\mathcal{V}(p)$ and neighborhood $\mathcal{N}(p)$. Let $G_p \prec G$ be the induced subgraph of G with vertex set $\mathcal{V}(G_p) = \mathcal{V}(p) \cup \mathcal{N}(p)$ (see Fig. 4 for an example) and B_p its adjacency matrix. Let $\mathsf{R}(z) = \big(\mathsf{I} - z\mathsf{A}\big)^{-1}\big|_{G_p}$ be the restriction of the resolvent of G to G_p.*

Then the ordinary generating function $R_p(z)$ of closed walks whose last erased loop is p is given by

$$R_p(z) = z^{\ell(p)} \det\big(\mathsf{I} + z\mathsf{R}(z)\mathsf{B}_p\big). \tag{6}$$

Observe that neither the series $\det(\mathsf{I} - z\mathsf{A}_{G\setminus p})$ nor $\det(\mathsf{I} - z\mathsf{A})$ appearing in Viennot's original lemma are well defined on infinite graphs—e.g. all their

[6]Recall that this is the same thing as counting closed walk whose last erased loop is a certain simple cycle.

finite order coefficients can be infinite. Their ratio evaluated on a sequence of finite graphs converging to G (as defined in Appendix A) nonetheless gives rise to a well defined series in the sense of the Proposition above.

Proof. Let $\{G_N^{\mathrm{Tor}}\}_{N\in\mathbb{N}}$ be the small torus sequence of graphs converging to G as defined in Appendix A. For any SAP p, define $N(p)\in\mathbb{N}$ such for all $N\geq N(p)$, then G_p is an induced subgraph of G_N.[7] The result follows by using Viennot's lemma on G_N^{Tor} and grouping all terms into a single determinant. Let p be a prime of finite support, hence finite length $\ell(p)$ on G. Let $N\geq N(p)$, $\mathsf{R}_N(z):=(\mathsf{I}-z\mathsf{A}_{G_N^{\mathrm{Tor}}})^{-1}$ and $\mathsf{B}_N(p):=\mathsf{A}_{G_N^{\mathrm{Tor}}}-\mathsf{A}_{G_N^{\mathrm{Tor}}\setminus p}$. Then Viennot's lemma on G_N^{Tor} yields the ordinary generating function $R_{N,p}(z)$ of closed walks whose last erased loop is p on G_N^{Tor} as

$$\begin{aligned}R_{N,p}(z)=\frac{\det(\mathsf{I}-z\mathsf{A}_{G_N^{\mathrm{Tor}}\setminus p})}{\det(\mathsf{I}-z\mathsf{A}_{G_N^{\mathrm{Tor}}})}&=\det\big(\mathsf{R}_N(z)\big)\det\big(\mathsf{I}-z\mathsf{A}_{G_N^{\mathrm{Tor}}}+z\mathsf{B}_{N,p}\big),\\&=\det\big(\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_{N,p}\big).\end{aligned}$$

Given that the sequence of small tori G_N^{Tor} converges to G, $\lim_{N\to\infty}R_{N,p}(z)=R_p(z)$ provided we can show that $\det\big(\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_{N,p}\big)$ is well defined under this limit.

Now since $\big(\mathsf{B}_{N,p}\big)_{ij}=0$ unless both $i,j\in\mathcal{V}(p)\cup\mathcal{N}(p)$ and since p is of finite length, then $\det\big(\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_{N,p}\big)$ is equal to the determinant of the finite $|\mathcal{V}(p)\cup\mathcal{N}(p)|\times|\mathcal{V}(p)\cup\mathcal{N}(p)|$ matrix $\mathsf{Q}_{ij}:=\big(\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_{N,p}\big)_{ij}$, for i,j in p and its neighborhood. Given that $\mathsf{B}_{N,p}\big|_{G_p}=\mathsf{B}_p$ for all $N\geq N(p)$ and since the graph has bounded degree $\lim_{N\to\infty}\mathsf{R}_N(z)=\mathsf{R}(z)$ is well defined. We consequently have

$$R_p(z)=\lim_{N\to\infty}R_{N,p}(z)=\lim_{N\to\infty}\det\Big(\big(\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_{N,p}\big)|_{G_p}\Big)=\det\big(\mathsf{I}+z\mathsf{R}(z)\mathsf{B}_p\big).$$

We emphasise that the determinant $\det\big(\mathsf{I}+z\mathsf{R}(z)\mathsf{B}_p\big)$ is equal to that of the finite matrix Q and no considerations pertaining to the determinants of infinite matrices is needed. □

We illustrate Viennot's lemma on infinite graphs with the ordinary generating function $R_e(z)$ of closed walks whose last erased loop is an edge cycle

[7] Existence of $N(p)$ is guaranteed for finite length SAPs as the small torus $G_{N:=n^2}^{\mathrm{Tor}}$ contains the ball of radius $n/2$ centred on the starting point of p on G.

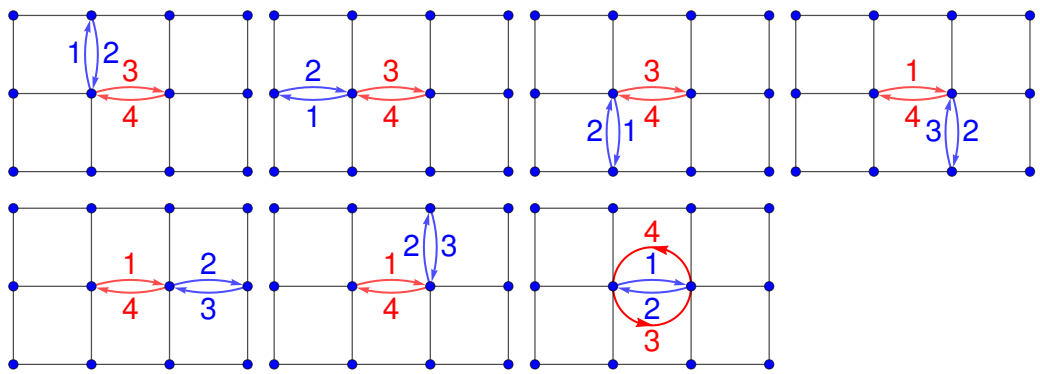


Figure 5: Illustration of the 7 walks of length 4 on the square lattice whose last erased loop is an edge cycle (highlighted in red). The numbers next to the highlighted edges indicate the order in which these edges are traversed.

e on the square lattice. Direct application of Eq. (6) gives

$$\begin{aligned} R_e(z) &= \frac{\pi}{4} - \frac{1}{16} + \frac{1}{16\pi^2}\big(64z^2-4\big)K\big(16z^2\big)^2 + \frac{1}{4\pi}\Big(K\big(16z^2\big)-\pi^2\Big), \qquad (7)\\ &= z^2 + 7\,z^4 + 70\,z^6 + 807\,z^8 + 10,046\,z^{10} + 131,206\,z^{12} + \cdots \end{aligned}$$

where $K(x) := \int_0^{\pi/2}\big(1 - x\sin^2(\theta)\big)^{-1/2}d\theta$ is the complete elliptic integral of the first kind. In Fig. (5) we illustrate the 7 walks of length 4 on the square lattice whose last erased loop is an edge cycle, as correctly counted by $[z^4]R_e(z) = 7$. Typically, generating functions for walks whose last erased loop is a certain self-avoiding polygon p, such as $R_e(z)$ above, are not meromorphic as functions of z. Consequently, the asymptotic growth of their coefficients cannot be determined with the traditional tools of meromorphic asymptotics [23]. In the next section we develop a generic method as a replacement that allows us to determine this asymptotic growth for any prime p.

5.2. Asymptotic expansion of Viennot's lemma

We now aim at establishing the formula for the fraction of walks whose last erased loop is any chosen self-avoiding polygon on any infinite vertex-transitive graph.

Proof of Corollary 2.1. The first difficulty in extending the Finite Sieve Theorem to infinite graphs comes from the proliferation of hikes on such graphs:

there are either exactly 0 or infinitely many hikes of any given length. Furthermore, the number of hikes increases uncontrollably with the length as there are also infinitely many more hikes of any length $L' > L$ than of length L. These observations continue to be true even when hikes are considered up to translation. To make matters worse, the fraction of all hikes which are walks is exactly 0; and even with edge weights uniformly set to $1/\lambda$, the total weight carried by all walks whose last erased loop is any SAP p is still divergent.

To resolve these serious difficulties requires us to separate the finite sieve results into two contributions, the first of which relates hikes to walks and the second relates closed walks to walks with fixed last erased loop p. This second contribution must itself be dealt with carefully to cure divergences stemming from the non-meromorphic nature of the generating functions produced by the extension of Viennot's Lemma to infinite graphs. We illustrate every step of the proof with explicit results on the square lattice.

Let us consider the dominant term of the asymptotic expansion of Viennot's lemma on the sequence of finite graphs G_N^{Tor}. Fixing N, consider p a self-avoiding polygon on G_N^{Tor}. Using the transformation presented in the extension of Viennot's lemma to infinite graphs, we can express the asymptotic fraction of hikes of length l that are walks whose last erased loop is p as $l \to \infty$ as

$$\begin{aligned}\frac{S(\mathcal{H}_l, \mathcal{P}_{G_N^{\mathrm{Tor}}\setminus p})}{|\mathcal{H}_l|} &\sim \lambda^{-\ell(p)} \det\left(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_{G_N^{\mathrm{Tor}}\setminus p}\right) \\ &= \lambda^{-\ell(p)} \lim_{z\to 1/\lambda^-} \zeta_N(z)^{-1} \det\left(\mathsf{I} + z\mathsf{R}_N(z)\mathsf{B}_p\right).\end{aligned}$$

This suggests a strategy consisting of proving separate convergence in $z \to 1/\lambda^-$ of the two terms in the limit above. This naive strategy ultimately fails, but the procedure that works is best understood once the nature of this failure is made apparent and several results we will obtain along the way are necessary to implement the correct proof strategy. In this spirit, we pretend to follow the naive approach and thus first examine the behaviour of the limit $\lim_{z\to\lambda^{-1}} \zeta_N(z)^{-1}$ asymptotically in N:

Lemma 5.2. *Let $\{G_N^{Tor}\}_{N\in\mathbb{N}}$ be the small tori sequence of vertex-transitive graphs converging to the infinite bounded-degree vertex-transitive graph G with maximum eigenvalue λ. Let $\zeta_N(z)$ be the zeta function of hikes on G_N^{Tor}*

and let $R(z) = \mathsf{R}(z)_{ii}$ be the ordinary generating function of closed walks on G. Then,

$$\lim_{N\to\infty} \zeta_N(z)^{1/N} = \tilde{\zeta}(z) = \exp\left(\int \frac{1}{z}\big(R(z)-1\big)dz\right). \tag{8}$$

Furthermore $\alpha := \lim_{z\to 1/\lambda^-} \tilde{\zeta}^{-1}(z)$ is well defined. All the coefficients $[z^n]\tilde{\zeta}(z)$ are positive integers, while the coefficients $[z^n]\tilde{\zeta}^{-1}(z)$ are integers.

Proof. Since the log-derivative of the hike zeta function is the trace of the resolvent, the log derivative of $\zeta(z)^{1/N}$ is, on vertex-transitive graphs, a single diagonal entry of the resolvent, yielding Eq. (8). Assuming that the dominant eigenvalue is unique, $R(z)$ diverges at worse as $1/(1-z\lambda)$ around $z \sim 1/\lambda$ so that necessarily $\int \frac{1}{z}\big(R(z)-1\big)dz$ converges in $1/\lambda$.

To understand the coefficients of $[z^n]\tilde{\zeta}(z)$, let h be a hike and let $w_1, \cdots w_n$ be vertex-disjoint walks making up hike h, i.e. $h = w_1 w_2 \cdots w_n$ modulo the fact that all these walks commute with one another. A *rooted hike* h_{root} is the object obtained from h on translating all walks w_i so that the origin lies on their unique right prime divisor but retaining the fact that they commute. Then $[z^n]\tilde{\zeta}(z)$ counts the number of rooted hikes of length n, $\tilde{\zeta}(z)$ is the zeta function of rooted hikes, while $\tilde{\mu}(z) := \tilde{\zeta}^{-1}(z)$ is the Möbius function on rooted hikes. □

Corollary 5.3. *On the square lattice, the generating function of rooted hikes is*

$$\begin{aligned}\tilde{\zeta}(z) &= \exp\left(\sum_{n\geq 1}\binom{2n}{n}^2 \frac{z^{2n}}{2n}\right),\\ &= 1 + 2z^2 + 11z^4 + 86z^6 + 805z^8 + 8402z^{10} + 94306z^{12} + \ldots,\end{aligned}$$

This is illustrated on Fig. (6). The corresponding Möbius function is

$$\begin{aligned}\tilde{\mu}(z) &:= 1/\tilde{\zeta}(z),\\ &= 1 - 2z^2 - 7z^4 - 50z^6 - 456z^8 - 4728z^{10} - 53095z^{12} + \ldots,\end{aligned}$$

and $\alpha := \tilde{\zeta}^{-1}(1/\lambda) = \frac{1}{4}e^{\frac{4C}{\pi}} \simeq 0.8025...$ with C Catalan's constant.

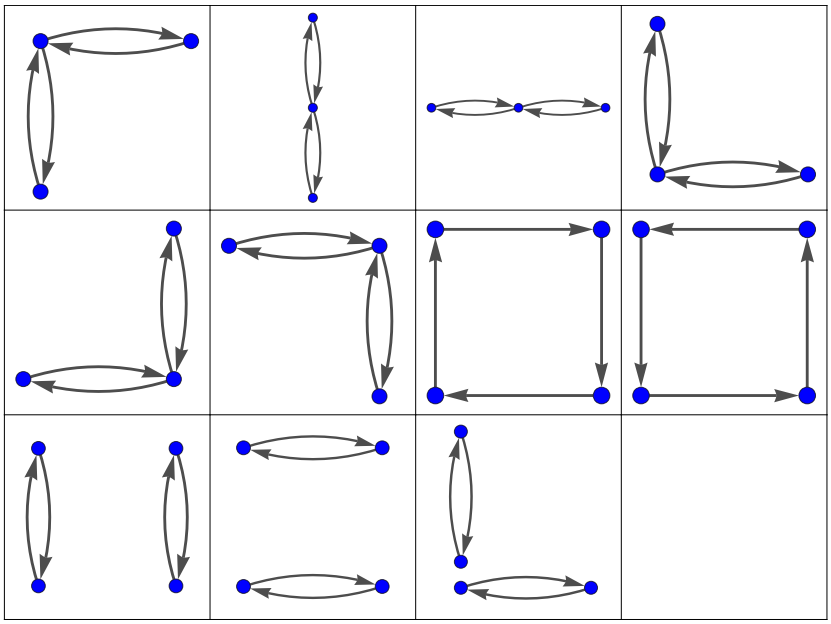


Figure 6: The $[z^4]\tilde{\zeta}(z) = 11$ objects of length 4 giving rise to distinct rooted hikes of length 4 on the infinite square lattice. Rigorously speaking, rooted hikes are obtained on considering the above objets up to translation of each walk.

Proof. While this result is an immediate corollary of the precedent lemma, we can prove it directly by considering the sequence of small $n \times n$ square lattices $G^{\mathrm{Sq}}_{n^2}$ converging to G as described in Appendix A. The eigenvalues of $G^{\mathrm{Sq}}_{N:=n^2}$ are $\lambda(i,k) = 2\cos\left(\pi\frac{j}{n+1}\right) + 2\cos\left(\pi\frac{k}{n+1}\right)$, where j and k are two integers between 1 and n. Product-integration yields

$$\lim_{N\to\infty} \zeta_N(z)^{1/N} = \exp\left(\int_S \log\big(1 - 2z\cos(\pi x) - 2z\cos(\pi y)\big)dxdy\right).$$

with S the unit square $[0,1]\times[0,1]$. We now obtain the relation to $R(z)$ directly. Observe that

$$\int_S \log\big(1 - 2z\cos(\pi x) - 2z\cos(\pi y)\big)dxdy = \\ -\sum_{n\geq 1}\frac{(2z)^{2n}}{2n}\int_S \big(\cos(\pi x) + \cos(\pi y)\big)^{2n}dxdy,$$

where we used the fact that odd powers of the sum of the cosines must have 0 integral over S since both cosines are symmetric functions on S. Then we

have

$$\int_S \big(\cos(\pi x)+\cos(\pi y)\big)^{2n}dxdy = \sum_{h=0}^{n}\binom{2n}{2h}\int_0^1 \cos(\pi x)^{2n-2h}dx\int_0^1 \cos(\pi y)^{2h}dy,$$
$$= \sum_{h=0}^{n}\binom{2n}{2h}\frac{\Gamma\left(n-h+\frac{1}{2}\right)}{\sqrt{\pi}(n-h)!}\frac{\Gamma\left(h+\frac{1}{2}\right)}{\sqrt{\pi}h!},$$
$$= \frac{4^n\Gamma\left(n+\frac{1}{2}\right)^2}{\pi(n!)^2} = \binom{2n}{n}^2 4^{-n}.$$

It follows that $\int_S \log\big(1-2z\cos(\pi x)-2z\cos(\pi y)\big)dxdy = -\sum_{n=1}^{\infty}\binom{2n}{n}^2\frac{z^{2n}}{2n} = -\int\frac{1}{z}\big(R(z)-1\big)dz$. □

Lemma 5.2 implies that asymptotically, when $N \gg 1$, $\lim_{z\to 1/\lambda^-}\zeta_N(z)^{-1} \sim \alpha^N$ is well defined. We will see later on that α^N relates the density of hikes to that of walks on an infinite graphs. More precisely, it can be interpreted as follows: every time a vertex is added to a graph G_N^{Tor} or G_N^{Sq}, $N \gg 1$, the fraction of hikes which are closed walks is multiplied by α.

Returning to the fraction of hikes on G_N^{Tor} which are walk whose last erased loop is p, Lemma 5.2 and our naive strategy suggest that we express this fraction asymptotically for $N \gg 1$ as

$$\det\left(\mathsf{I}-\frac{1}{\lambda}\mathsf{A}_{G_N^{\mathrm{Tor}}\setminus p}\right) \stackrel{?}{=} \alpha^N \lim_{z\to\lambda^{-1}}\det\big(\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_p\big). \tag{9}$$

In fact the right hand side is divergent under the limit $z \to 1/\lambda^-$, but we will see below that its divergence is generic: it is the same for any SAP p as well as for all closed walks. In other terms, the right hand side diverges only because the series of all walks diverges at the point $z \to 1/\lambda^-$, hinting that the fraction of walks whose last erased loop is p with respect to all walks might itself be well defined. To put this observation on firm foundations, we need to show that $\lim_{z\to 1/\lambda^-}\det\big(\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_p\big)/R(z)$ is well defined, and that the fraction of all walks with respect to all hikes is well defined as well, in spite of the generic divergence. These two results rely on different techniques. We start with proving of the second fact without which the first would be useless.

Lemma 5.4. *Let G be an infinite vertex-transitive graph of bounded degree. Let $\{G_N^{Tor}\}_{N\in\mathbb{N}}$ be the small tori sequence of vertex-transitive graphs converging to G. Let $\bullet$ be any vertex of G. Then the fraction $F_\bullet$ of hikes which are closed walks from $\bullet$ to itself is asymptotically given by*

$$F_\bullet \sim \frac{\alpha^N}{N}, \quad N \to \infty.$$

In this expression α is the constant defined in Lemma 5.2.

The $1/N$ factor in $F_\bullet$ originates from that we have fixed the vertex $\bullet$. If instead we consider translation invariant quantities, i.e. we consider all closed walks irrespectively of their starting point, then the fraction of hikes which are closed walks is asymptotically α^N, $N \gg 1$.

Remark 5.1. Let $|W|_N(\ell)$ and $|H|_N(\ell)$ be the *total* number of walks and of hikes up to length $\ell \geq 0$ on G_N^{Tor}, respectively. Then, Lemma (5.4) is equivalent to

$$\lim_{N\to\infty}\lim_{\ell\to\infty}\frac{1}{N}\log\left(|W|_N(\ell)/|H|_N(\ell)\right) = \log(\alpha),$$

that is the density of the logarithm of the fraction of hikes which are closed walks is well defined on G.

Proof. On the finite graphs G_N^{Tor}, the fraction of all closed walks off $\bullet$ with respect to hikes is given by $F_\bullet = \det(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_{G_N^{\mathrm{Tor}}\setminus\bullet})$. Let $\mathsf{B}_{N,\bullet}$ be the $N\times N$ identically zero matrix except on edges adjacent to $\bullet$ where its value is 1, i.e. so that $\mathsf{A}_{G_N^{\mathrm{Tor}}\setminus\bullet} = \mathsf{A}_{G_N^{\mathrm{Tor}}} - \mathsf{B}_{N,\bullet}$. Then $F_\bullet = \det(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_{G_N^{\mathrm{Tor}}} + \frac{1}{\lambda}\mathsf{B}_{N,\bullet})$ and we determine this determinant by expansion around $\mathsf{A}_{G_N^{\mathrm{Tor}}}$, hereafter denoted A_N in order to alleviate the equations. This gives

$$\begin{aligned} F_\bullet = \det\left(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_N\right) + \frac{1}{\lambda}\operatorname{Tr}\left(\operatorname{adj}\left(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_N\right)\mathsf{B}_{N,\bullet}\right) \\ - \frac{1}{2\lambda^2}\lim_{z\to\lambda^{-1}}\det\left(\mathsf{I} - z\mathsf{A}_N\right)\left[\operatorname{Tr}^2\big(\mathsf{R}_N(z)\mathsf{B}_{N,\bullet}\big) - \operatorname{Tr}\big(\mathsf{R}_N(z)\mathsf{B}_{N,\bullet}\big)^2\right]. \end{aligned}$$

All orders of degree higher than two are exactly zero. In this expression $\mathsf{R}_N(z) := \big(\mathsf{I} - z\mathsf{A}_N\big)^{-1}$, $\operatorname{Tr}(\mathsf{M})$ designates the trace of a matrix M and $\operatorname{adj}(\mathsf{M})$ its adjugate.

Order zero of the expansion is null since λ is an eigenvalue of G_N^{Tor}. By the same token, the adjugate matrix of the first order is proportional to the projector $\mathsf{P}_{N,\lambda}$ onto the eigenvector associated with λ on G_N^{Tor}

$$\mathrm{adj}\left(\mathsf{I}-\frac{1}{\lambda_N}\mathsf{A}_N\right)=\alpha_N\,\mathsf{P}_{N,\lambda},$$

where $\alpha_N := \prod_{\substack{\lambda_i\in\mathrm{Sp}(\mathsf{A}_N)\\ \lambda_i\neq\lambda_N}}(1-\lambda_i/\lambda_N)$. Since G_N^{Tor} is regular $\mathsf{P}_{\lambda_N}=\frac{1}{N}\mathsf{J}$, where $\mathsf{J}_{ij}=1$. Then

$$\frac{1}{\lambda}\,\mathrm{Tr}\left(\mathrm{adj}\left(\mathsf{I}-\frac{1}{\lambda}\mathsf{A}_N\right)\mathsf{B}_{N,\bullet}\right)=\frac{\alpha_N}{\lambda}\times\frac{2\lambda}{N}=2\frac{\alpha_N}{N},$$

since $\mathsf{B}_{N,\bullet}$ has exactly 2λ non-zero entries.

The second order of the expansion is always well defined, as is readily seen from the equivalent form

$$\begin{aligned}-\frac{1}{2\lambda^2}\lim_{z\to\lambda^{-1}}\det\left(\mathsf{I}-z\mathsf{A}_N\right)&\left[\mathrm{Tr}^2\big(\mathsf{R}_N(z)\mathsf{B}_{N,\bullet}\big)-\mathrm{Tr}\big(\mathsf{R}_N(z)\mathsf{B}_{N,\bullet}\big)^2\right]\\ &=\frac{\alpha_N}{\lambda^2}\left[\mathrm{Tr}\big(\mathsf{P}_{N,\lambda}\mathsf{B}_{N,\bullet}\big)\mathrm{Tr}\big(\mathsf{C}_N\mathsf{B}_{N,\bullet}\big)-\mathrm{Tr}\big(\mathsf{P}_{\lambda_N}\mathsf{B}_{N,\bullet}\mathsf{C}_N\mathsf{B}_{N,\bullet}\big)\right].\end{aligned}\tag{10}$$

where

$$\mathsf{C}_N:=\lim_{z\to1/\lambda^-}\big(\mathsf{I}-\mathsf{P}_{N,\lambda}\big)\mathsf{R}_N(z),\tag{11}$$

and

$$\alpha_N:=\lim_{z\to1/\lambda^-}(1-z\lambda)^{-1}\det\left(\mathsf{I}-z\mathsf{A}_N\right)=\prod_{\substack{\lambda_i\in\mathrm{Sp}(\mathsf{A}_N)\\ \lambda_i\neq\lambda_N}}(1-\lambda_i/\lambda_N),$$

which, by Lemma 5.2, is asymptotically given by α^N for $N\gg1$ and $\alpha=\tilde{\zeta}(1/\lambda)^{-1}$. The second line of Eq. (10) stems directly from the observation that $\mathsf{P}_{N,\lambda}$ is a rank one projector. Indeed for any such projector Q and matrix M of bounded norm we have $\mathrm{Tr}\big((\mathsf{QM})^2\big)=\mathrm{Tr}^2\big(\mathsf{QM}\big)$.

Each term of the second line of Eq. (10) can be evaluated generically. For convenience, let c_0 be the diagonal entry of C_G, c_1 the entry relating first neighbours on G and $c_{2,i}$, all the distinct entries of C_G relating second neighbours on G. Then $\mathrm{Tr}\big(\mathsf{P}_{N,\lambda}\mathsf{B}_{N,\bullet}\big)=2\lambda/N$, $\mathrm{Tr}\big(\mathsf{C}_N\mathsf{B}_{N,\bullet}\big)=2\lambda c_1$ and

$$\begin{aligned}N\,\mathrm{Tr}\big(\mathsf{P}_{\lambda_N}\mathsf{B}_{N,\bullet}\mathsf{C}_N\mathsf{B}_{N,\bullet}\big)&=\sum_{\substack{\text{ordered}\\ \text{vertex pairs } v_i,v_j}}\deg(v_i)\deg(v_j)\,\mathsf{C}_{v_iv_j},\\ &=(\lambda^2+\lambda)c_0+2\lambda^2c_1+\sum_i n_{2,i}\,c_{2,i},\end{aligned}$$

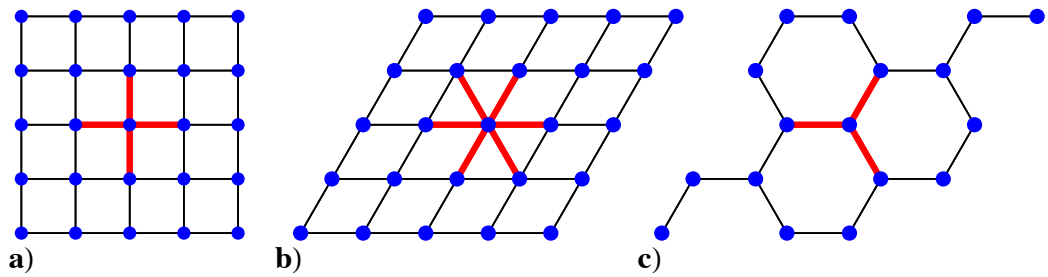


Figure 7: In thick red edges, corolla graphs $G_\bullet$ on: **a)** the square lattice; **b)** the triangular lattice; **c)** the hexagonal lattice. The same construction and proof given in this section applies to any vertex transitive graph, in particular it is not limited to planar lattices and corollas can be considered in any dimension.

where $\deg(v_i)$ is the degree of v_i on the corolla $G_\bullet$ (see Fig. 7 for an illustration of corolla graphs) and $n_{2,i}$ is the number of times entry $c_{2,i}$ appears. Recursion relations between the entries of C_G presented in Appendix C give $c_0 = 0$, $c_1 = -1$, $\sum_{i=1}^{\lfloor \lambda/2 \rfloor} n_{2,i} c_{2,i} = -\lambda^2$. The second order then generically evaluates to

$$\frac{\alpha_N}{N\lambda^2}\left(-4\lambda^2 - (-2\lambda^2 - \lambda^2)\right) = -\frac{\alpha_N}{N}.$$

We can finally put the zeroth, first and second orders of the determinant expansion together, yielding

$$\det\left(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_N\right) = 0 + 2\frac{\alpha_N}{N} - \frac{\alpha_N}{N} = \frac{\alpha_N}{N}.$$

Asymptotically, for $N \gg 1$, $\alpha_N \sim \alpha^N$ as per Lemma 5.2, which gives the result. □

The method employed in the proof of Lemma 5.4 to calculate the fraction of hikes which are closed walks extends to any self-avoiding polygon. In this more general situation, the expansion of the determinant always terminates at a finite order which grows with the polygon's length. The formulas to be evaluated become very involved however, to the point of being effectively impractical even for short SAPs. We nonetheless provide the general expression of the formal expansion at all orders for any SAP in Appendix D.

To further exemplify this determinant expansion it is worth considering the calculation of the fraction F_e of hikes which are walk whose last erased loop is an edge e on the square lattice. In this case, the expansion of F_e around $\det(\mathsf{I} - \frac{1}{\lambda}\mathsf{A}_G)$ has exactly four non-zero orders each of which involves

complicated sums of entries of the graph resolvent $\mathsf{R}(z)$ For example, the second order is asymptotically equal to

$$\lim_{z\to 1/\lambda^-} \frac{\alpha_N}{\lambda^2 N}\Big(112\mathsf{R}_{10}(z) - 38\mathsf{R}_{00}(z) - 40\mathsf{R}_{11}(z) \\ - 20\mathsf{R}_{20}(z) - 12\mathsf{R}_{21}(z) - 2\mathsf{R}_{30}(z)\Big) = \frac{\alpha^N}{\lambda^2 N} \times -10,$$

for $N \gg 1$. Although $\lambda = 4$ on the square lattice, we left it unevaluated to help see that this is the second order of the expansion. Here $\mathsf{R}_{ij}(z)$ designates the entry of the resolvent relating two vertices with distance i along x and j along y. Taken together, the four orders of the determinant expansion give asymptotically for $N \gg 1$

$$F_e = \frac{1}{\lambda^2}\left(14\frac{\alpha^N}{\lambda N} - 10\frac{\alpha^N}{\lambda^2 N} - 80\frac{\alpha^N}{\lambda^3 N} + 96\frac{\alpha^N}{\lambda^4 N}\right) = \frac{2\alpha^N}{N\lambda^2} = \frac{\alpha^N}{8N},$$

where the overall $1/\lambda^2$ factor in front arises from the fact that the edge has length 2 and would simply be $1/\lambda^{\ell(p)}$ for a general SAP. Lemma 5.4 indicates that α^N/N relates hikes to closed walks, while the remaining 1/8 factor is thus the asymptotic fraction of closed walks which are walk whose last erased loop is an edge cycle. This means in particular that, whenever $[z^\ell]R(z) \neq 0$,

$$\lim_{\ell\to\infty} \frac{[z^\ell]R_e(z)}{[z^\ell]R(z)} = \frac{1}{8},$$

where $R_e(z)$ is the generating function given in Eq. (7).

It might struck the reader that a remarkable number of simplifications must have taken place to yield such a simple answer as 1/8 from the complicated orders of the expansion. This hints at the existence of a much simpler calculation procedure, and this is precisely the procedure in the spirit of Eq. (9). Now however, we will be able to remove the generic divergence of the naive approach, having separately obtained an asymptotic expression for the fraction of walks with respect to hikes by Lemma 5.4.

Lemma 5.5. *Let G be an infinite vertex-transitive graph of bounded-degree and let λ be the supremum of its spectrum. Let $\{G_N^{Tor}\}_{N\in\mathbb{N}}$ be the small tori sequence of vertex-transitive graphs converging to G. Then the asymptotic*

fraction of closed walks which are walk whose last erased loop is p is well defined and given by

$$\begin{aligned}\frac{F_p}{\lambda^{\ell(p)}} &= \lim_{N\to\infty}\lim_{z\to 1/\lambda^-} z^{\ell(p)}\frac{\det\big(\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_p\big)}{\det\big(\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_\bullet\big)},\\ &= \frac{1}{\lambda^{\ell(p)+1}}\,\mathsf{deg}^{\mathrm{T}}.\operatorname{adj}\left(\mathsf{I}+\mathsf{C}_G\big|_p.\mathsf{B}_p\right).\mathbf{1},\end{aligned}$$

where $\mathsf{C}_G\big|_p$ is the restriction to G_p of $\mathsf{C}_G := \lim_{z\to 1/\lambda^-}(\mathsf{I}-\mathsf{P}_\lambda)\mathsf{R}(z)$.

Proof. According to the Finite Sieve Theorem, the asymptotic fraction of hikes which are closed walks whose last erased loop is p on G_N^{Tor} is $\lim_{z\to 1/\lambda^-} z^{\ell(p)}\det(\mathsf{I}-z\mathsf{A}_{G_N^{\mathrm{Tor}}\setminus p})$. Since all the limits taken here are finite and well defined (as everything takes place on G_N^{Tor}), the asymptotic fraction of closed walks whose last erased loop is p on these finite graphs is

$$\begin{aligned}\frac{\lim_{z\to 1/\lambda^-} z^{\ell(p)}\det(\mathsf{I}-z\mathsf{A}_{G_N^{\mathrm{Tor}}\setminus p})}{\lim_{z\to 1/\lambda^-} z^{0}\det(\mathsf{I}-z\mathsf{A}_{G_N^{\mathrm{Tor}}\setminus \bullet})} &= \lim_{z\to 1/\lambda^-} z^{\ell(p)}\frac{\zeta_N(z)\det(\mathsf{I}-z\mathsf{A}_{G_N^{\mathrm{Tor}}\setminus p})}{\zeta_N(z)\det(\mathsf{I}-z\mathsf{A}_{G_N^{\mathrm{Tor}}\setminus \bullet})},\\ &= \lim_{z\to 1/\lambda^-} z^{\ell(p)}\frac{\det\big(\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_p\big)}{\det\big(\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_\bullet\big)}.\end{aligned}$$

We now turn to studying the behaviour of the right hand side as $N\to\infty$. To this end, we expand $\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_p$ around $1/\lambda$ with $z<1/\lambda$.

We need to distinguish behaviours based on the dimensionality $d>1$ of the lattice under study. We ignore the trivial 1D case (for which the only SAP is the edge, and the fraction of closed walks whose last erased loop is the left or right edge attached to any vertex is 1/2). On $d>1$ dimensional lattices we have,

$$\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_p = \mathsf{C}_N\mathsf{B}_p - \frac{1}{\lambda\pi}\mathsf{P}_{N,\lambda}\mathsf{B}_p\,\mathrm{Li}_{d/2}(1-z\lambda)+O(1-z\lambda), \tag{12}$$

with C_N as defined in Eq. (11) and $\mathrm{Li}_a(x) := \sum_{n>1}\frac{x^n}{n^a}$ is the a-polylogarithm function. Combinatorially, it arises here from summations over closed walks weighted by $\lambda^{-\ell}$, which leaves a residual total weight asymptotically given by $\ell^{-d/2}$ for all closed walks of length $\ell\gg 1$. The generic nature of the behaviour exhibited by $\mathsf{I}+z\mathsf{R}_N(z)\mathsf{B}_p$ is now readily apparent: 1) divergence occurs only on 2D lattices, where it is logarithmic; 2) it is the same for all SAPs; and 3) it is also the same for all closed walks (which are readily recovered upon taking

p to be length 0, i.e. $\mathsf{B}_p \equiv \mathsf{B}_\bullet$ is a corolla). Thanks to these observations, the determinant expansion at $1/\lambda^-$ is

$$\begin{aligned}\det\big(\mathsf{I} + z\mathsf{R}_N(z)\mathsf{B}_p\big) &= \det\left(\mathsf{C}_N\mathsf{B}_p - \frac{1}{\lambda\pi}\mathsf{P}_{N,\lambda}\mathsf{B}_p\,\mathrm{Li}_{d/2}(1-z\lambda) + O(1-z\lambda)\right),\\ &= -\frac{1}{\pi N}\mathrm{Li}_{d/2}(1-z\lambda)\frac{1}{\lambda}\,\mathsf{deg}^{\mathrm{T}}.\,\mathrm{adj}\,(\mathsf{I}+\mathsf{C}_N.\mathsf{B}_p)\;.\,\mathsf{1}\\ &\qquad\qquad + o\big(\mathrm{Li}_{d/2}(1-z\lambda)\big),\end{aligned}$$

where we used the matrix-determinant lemma and the QR decomposition

$$-\frac{1}{\lambda\pi}\mathsf{P}_{N,\lambda}\mathsf{B}_p = -\frac{1}{\pi N}\times\frac{1}{\lambda}\times\mathsf{1}.\mathsf{deg}^{\mathrm{T}}.$$

This decomposition relies on the observation that $\mathsf{P}_{N,\lambda}$ is the projector onto $\mathsf{1}$, i.e. that all G_N^{Tor} are regular.[8] We recall that in the above expression, $\mathsf{deg} = \mathsf{B}_p.\mathsf{1} = \mathrm{diag}(\mathsf{B}_p^2)$ is the degree of vertices on G_p. Similarly, at $1/\lambda^-$,

$$\det\big(\mathsf{I} + z\mathsf{R}_N(z)\mathsf{B}_\bullet\big) = -\frac{1}{\pi N}\mathrm{Li}_{d/2}(1-z\lambda) + o\big(\mathrm{Li}_{d/2}(1-z\lambda)\big),$$

and finally

$$z^{\ell(p)}\frac{\det\big(\mathsf{I} + z\mathsf{R}_N(z)\mathsf{B}_p\big)}{\det\big(\mathsf{I} + z\mathsf{R}_N(z)\mathsf{B}_\bullet\big)} = \frac{1}{\lambda^{\ell(p)+1}}\mathsf{deg}^{\mathrm{T}}.\,\mathrm{adj}\,(\mathsf{I}+\mathsf{C}_N.\mathsf{B}_p)\;.\,\mathsf{1} + o(1),$$

which yields the result after taking the limits $z \to 1/\lambda^-$ and $N \to \infty$ now both clearly well-defined, even when $d = 2$. Combinatorially, the divergence curing on 2D lattices effected here comes from relating walk whose last erased loop is a fixed SAP to all closed walks rather than directly to the hikes. The relation between closed walks and hikes is performed separately through Lemma 5.4.

Remark 5.2. Recall that $R(z)$ and $R_p(z)$ are the ordinary generating functions of all closed walks on G and of closed walk whose last erased loop is p on G, respectively. The finite sieve here indicated that the limit in $z \to 1/\lambda^-$ of

[8] This step can be adapted should we consider lattices which are not vertex-transitive but for which $G/\mathrm{Aut}(G)$ has finitely many vertices. This is beyond the scope of this work.

the ratio of power series $R_p(z)/R(z)$ yields the asymptotic behaviour of the term-by-term ratio $[z^\ell]R_p(z)/[z^\ell]R(z)$ as $\ell \to \infty$, which rather corresponds to an Hadamard division of $R_p(z)$ by $R(z)$, i.e.

$$\lim_{z\to 1/\lambda^-} \frac{R_p(z)}{R(z)} = \lim_{\ell\to\infty} \frac{[z^\ell]R_p(z)}{[z^\ell]R(z)}.$$

This is a corollary of the fact that the leading divergence of both $R_p(z)$ and $R(z)$ located at $z = 1/\lambda$. For a general discussion on the relation between Hadamard products and singularity analysis, we refer to [24].

□

This concludes the proof of the Infinite Sieve Theorem. □

5.3. Probabilistic interpretations

The results of the Infinite Sieve Theorem have a probabilistic interpretation which motivated the study of the fraction $F_p/\lambda^{\ell(p)}$ in the literature ultimately leading to SLE_2, albeit without the relation to hikes and the explicit form of Eq. (1). Here the interpretation of $F_p/\lambda^{\ell(p)}$ as a probability distribution over the SAPs is first manifested in the following:

$$\sum_{p:\,\mathrm{SAP}} \frac{F_p}{\lambda^{\ell(p)}} = 1. \tag{13}$$

Combinatorially, this trivially states that any walk has a unique right prime factor [17] and consequently the total fraction of all closed walks whose last erased loop is *any* SAP is exactly 1.

Remark 5.3. All *simple cycles* passing through some chosen vertex of G are present in the sum of Eq. (13). Following the rules in the Cartier-Foata monoid of the hikes, a SAP of length $\ell(p)$ thus appears at least $2\ell(p)$ times in the sum. Here, the factor of 2 accounts for the orientation; and $\ell(p)$ reflects all the valid starting points for the SAP as a closed walk.[9] For example, on

[9]Rigorously, what matters is the size of the equivalence class of all words on edges that define the same hike h in the original Cartier-Foata monoid. As mentioned earlier, this size is the value of the von Mangoldt function $\Lambda(h)$, which is the length of the unique right factor of a walk if h is a walk and 0 otherwise [17]. Seen as a closed walk, a SAP p is its own unique right-prime divisor and thus $\Lambda(p) = \ell(p)$.

the square lattice, the 1×1 square gets a factor of $8 = 2 \times 4 \times 1$ and the 1×2 rectangle has a factor of $12 = 2 \times 6$. Strictly speaking, the sum Eq. (13) therefore runs over simple cycles with fixed starting point and the SAP index is a (harmless) notational abuse.

We can go further in the probabilistic interpretation using purely combinatorial arguments:

Proposition 5.6. *Let G be an infinite vertex transitive planar graph and let p be a SAP on it. Let w be a random walk with uniform edge-transition probability $1/\lambda$. Run the walk until it comes back to its starting point. Then the probability $\mathbb{P}(w \mapsto p)$ that the last erased loop of w be p is equal to the fraction of all closed walks (including those passing an arbitrary number of times through the origin) whose last erased loop is p, $\mathbb{P}(w \mapsto p) = F_p/\lambda^{\ell(p)}$.*

Proof. Let G_N be a family of finite vertex-transitive graphs of degree λ converging to G as $N \to \infty$. Let $W_{N,p}(z)$ be the ordinary generating function of closed walks on G_N with right prime divisor p and such that these walks never revisit their starting point except on their final step. By construction, we have $\mathbb{P}(w \mapsto p) = \lim_{z \to 1/\lambda^-} \lim_{N \to \infty} W_{N,p}(z)$ provided both limits are well defined. But

$$\begin{aligned} W_{N,p}(z) &= z^{\ell(p)} \frac{\det\left(\mathsf{I} - z\mathsf{A}_{G_N \setminus p}\right)}{\det\left(\mathsf{I} - z\mathsf{A}_{G_N \setminus \bullet}\right)}, \\ &= z^{\ell(p)} \Big(\zeta_N(z) \det\left(\mathsf{I} - z\mathsf{A}_{G_N \setminus p}\right)\Big) \Big(\zeta_N(z) \det\left(\mathsf{I} - z\mathsf{A}_{G_N \setminus \bullet}\right)\Big)^{-1}. \end{aligned}$$

This has a well defined limit when $N \to \infty$ as per the extension of Viennot's lemma to infinite graphs

$$W_p(z) := \lim_{N \to \infty} W_{N,p}(z) = z^{\ell(p)} \frac{\det\left(\mathsf{I} + z\mathsf{R}(z)\mathsf{B}_p\right)}{\det\left(\mathsf{I} + z\mathsf{R}(z)\mathsf{B}_\bullet\right)},$$

and by the Infinite Sieve Theorem the limit $z \to 1/\lambda^-$ of the above is also well defined. It is given by $\lim_{z \to 1/\lambda^-} W_p(z) = F_p/\lambda^{\ell(p)}$ with F_p as per Eq. (1). $\square$

The logarithm of F_p also has a probabilistic interpretation that has not appeared in the literature so far. It is based on results by Espinasse and Rochet [25]:

Proposition 5.7 (Espinasse and Rochet, 2019)**.** *Let G be an infinite vertex-transitive graph and let p be a SAP on it. Let $\mathbb{E}_w(.)$ designate the expectation value of a random variable with respect to the closed random walks defined up to translation and weighted with probability $\lambda^{-\ell(w)}$. Then*

$$\log(F_p) = \sum_{w:\, walk} \frac{\Lambda_p(w)}{\ell_p(w)} \lambda^{-\ell(w)} = \mathbb{E}_w \left(\frac{\Lambda_p(w)}{\ell_p(w)} \right).$$

Here $\Lambda_p(w)$ counts the vertices that are both in p and the unique right divisor of w and $\ell_p(w)$ is the number of vertices of p visited by w, counted with multiplicity.

The reason for the notation $\Lambda_p(w)$ is that this function is similar to the von Mangoldt function Λ defined on hikes [17]. Defining $\Lambda_p(h)$ to be 0 when the hike h is not a walk and otherwise $\Lambda_p(h)$ is as in the Proposition above, we have $\Lambda_p(h) = \Lambda(h)$ when p is the unique right prime divisor of h.

The combination of Propositions 5.6 and 5.7 implies the rather uncommon result that the entropy of the distribution of $F_p\lambda^{-\ell(p)}$ values itself has a probabilistic interpretation. Consequences of this observation will be explored in a separate work.

6. Discussion

6.1. Extension to SAWs and further lattices

In this contribution, we presented fully deterministic combinatorial arguments based on number-theoretic sieves for counting walk multiples of SAPs on any finite or infinite vertex-transitive graphs. This is equivalent to counting all the walks whose last erased loop following Lawler's loop erasing procedure is some chosen self-avoiding polygon p on such graphs. In fact, all the results presented here are immediately valid for self-avoiding walks (SAWs) as well. In particular, Eq. (1) of the Infinite Sieve Theorem giving the fraction F_p is immediately correct. For a self-avoiding walk p, $F_p/\lambda^{\ell(p)}$ is the fraction of all open walks with the same fixed starting and end points for which p is the self-avoiding skeleton remaining after loop-erasing.

Finally, the arguments presented here should extend without fundamental changes to infinite graphs that are not vertex transitive as long as $G/\mathrm{Aut}(G)$ is finite. This nonetheless requires further work and is beyond the scope of the present contribution.

6.2. Counting the self-avoiding polygons

We recall that $R(z)$ and $R_p(z)$ are the ordinary generating functions of closed walks and of walks whose last erased loop is a SAP p, respectively.

The research presented here suggests a natural strategy to tackle the open problem of asymptotically counting SAWs and SAPs. First, observe that we know the exact number $R(z)[L]$ of closed walks of length L defined up to translation. Then, if we could determine the exact number $R_p(z)[L]$ of closed walks of length exactly L whose last erased loop is p, it would be sufficient to sum this over all SAPs of length strictly less than L and subtract the result from $R(z)[L]$ to determine the number $\pi(L)$ of SAPs of length exactly L:

$$\pi(L) = R(z)[L] - \sum_{\substack{p:\,\mathrm{SAP}\\ \ell(p)<L}} R_p(z)[L].$$

While such a precise count is not feasible in practice as $L \to \infty$, an asymptotic estimate of the number of walks whose last erased loop is any chosen SAP may seem, at first, to be sufficient to gain an insight into the number of SAPs themselves. Following this idea, we would rather write

$$\frac{\pi(L)}{R(z)[L]} = 1 - \sum_{\substack{p:\,\mathrm{SAP}\\ \ell(p)<L}} \frac{R_p(z)[L]}{R(z)[L]}. \tag{14}$$

and use $R_p(z)[L]/R(z)[L] \sim F_p\lambda^{-\ell(p)}$ for $L \gg 1$. Thus, we would only need to estimate sums like

$$S(L) := \sum_{\substack{p:\,\mathrm{SAP}\\ \ell(p)\le L}} \frac{F_p}{\lambda^{\ell(p)}},$$

for $L \gg 1$, in order to work out an asymptotic expansion for $\pi(L)$. Such an estimate can already be determined from R. Kenyon's seminal results [26], we find

$$S(L) = 1 - L^{-3/5} + O(L^{-3/5}). \tag{15}$$

See also Fig 8 for a numerical illustration. This result of course wildly differs from the $(\mu/\lambda)^L L^{-1/2}$ expected here from the numerically conjectured scaling for $\pi(L)$.[10] From the point of view of probability theory, the origin of this

[10]The correction term is $L^{-1/2}$ here because we count simple cycles rather than SAPs.

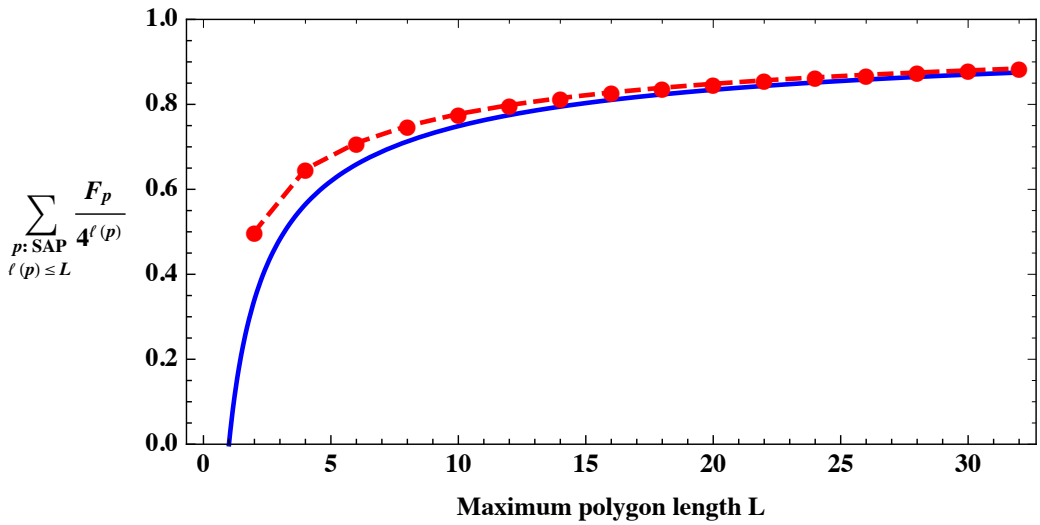


Figure 8: In red points and dashed red line, the sum over all self-avoiding polygons with fixed starting point on the square lattice of length at most L of the fraction $F_p/4^{\ell(p)}$, as a function of L. The corresponding data table is presented in Appendix F. The solid blue line is the first term of asymptotic expansion for this quantity, that is $1 - L^{-3/5}$. The discrepancy between $-3/5$ and the $-1/2$ expected from SAP counting is due to an accumulation of error terms in the Infinite Sieve Theorem, which require us to include terms beyond the dominant $F_p/4^{\ell(p)}$ or use stronger two-sided sieves.

discrepancy is clear: the law governing Lawler's loop erased random walks essentially converges to SLE_2 rather than the conjectured $\mathrm{SLE}_{8/3}$ for SAP and SAW models. From the point of view of sieve techniques however, the chasm between these results originates from an uncontrolled accumulation of error terms affecting the estimate $R_p(z)[L]/R(z)[L] \sim F_p\lambda^{-\ell(p)}$.[11]

It is important to recall that $F_p\lambda^{-\ell(p)}$ is only the first, asymptotically dominant term of the asymptotic expansion of the number of walks whose last erased loop is p. In particular $F_p\lambda^{-\ell(p)}$ is a good approximation to $R_p(z)[L]/R(z)[L]$ only when $L \gg \ell(p)$ (see Appendix E). Yet, when we subtract walks with fixed last erased loops from all closed walks of length

This is responsible for a factor of L in front of the $L^{-5/2}$. Since in addition, $R(z)[L] \sim \lambda^L/(\pi L)$ for $L \gg 1$, this accounts for another factor of L and finally we get $L \times L \times L^{-5/2} = L^{-1/2}$.

[11] These errors have the same origin as those affecting the Eratosthenes-Legendre sieve in number theory!

L, we must consider the walks whose last erased loop p is of length up to $\ell(p) = L - 1$. Given the exponential growth in the number of SAPs, this means that *most* of our estimates are affected by large, uncontrolled errors, and it is impossible to exploit Eq. (14) using solely $S(L)$.

This problem has two potential solutions. The first idea is to take into account some error terms $\mathrm{Err}_p(L)$ in the asymptotic expansion of $R_p(z)[L]/R(z)[L]$ so as to determine this quantity more precisely. Since all the error terms are exactly available[12], it seems possible that an extension to Kenyon's arguments would allow us to estimate sums of such error terms generically; just as Eq. (15) does for the dominant term. This idea suffers from a major drawback: error terms actually *grow* with L if we consider classes of SAPs for which $L-\ell(p)$ is fixed. Since most SAPs of length up to $L-1$ are close in length to L, this means that the overall error term affecting Eq. (14) grows uncontrollably with L. Thus, an increasingly (and unrealistically) detailed knowledge of the errors is needed as $L \to \infty$, so that this strategy collapses completely with respect to rigorous arguments.

6.3. The path to rigorous progress

The second approach relies on a crucial foundational work by M. Bousquet-Mélou regarding the enumeration of heaps of pieces satisfying both left and right constraints [27, 28]. This work opens the way for two-sided sieves in the same manner as Viennot's Lemma relates to the Finite and Infinite Sieve Theorems: they give control over both the left and right prime divisors of a walk. Consequently, the maximum length of the primes to be considered in Eq. (14) is reduced to only $L/2$. The "sieving gap" between L and $L/2$ dramatically reduces the importance of the error terms to the extend that, in accordance with Appendix E, we expect them all to vanish under the limit $L \to \infty$.[13] At the same time, the dominant contribution is not $F_p/\lambda^{\ell(p)}$ anymore. The resulting calculations will be presented in a separate work.

[12] In fact $R_p(z)[L]$ itself is in principle exactly available from the extention of Viennot's Lemma to infinite graphs. In this situation however, it is a precise estimate for the sum over SAPs of $R_p(z)[L]$ which is utterly lacking.

[13] Seeing heaps of pieces as an extension of number theory as in [17], shows that this $L/2$ is the extension of the $\sqrt{x}$ gap present in all standard number-theoretic sieves. We can similarly show that the fraction $F_p\lambda^{-\ell(p)}$ extends the quantity $\log(x)/\log(p)$ and all identities given here extend (and hence reduce to) valid number-theoretic identities. Non-trivial (novel) results on partial sums of the Möbius function also follow heuristically.

Appendix A Converging sequences of graphs

We here recall the notion of a converging sequence of graphs. We follow directly the work of [29]:

Let G and G' be two bounded degree graphs and let r and r' be vertices called the roots on G and G', respectively. A topology on the space $\mathcal{X}$ of isomorphism classes of rooted connected graphs is induced by the following metric. Let $B_G(r,n)$, $n \in \mathbb{N}^*$ be the ball of radius n centred on the root r in graph G. Let k be the supremum of all n such that $(B_G(r,n),r)$ and $(B_{G'}(r',n),r')$ be isomorphic as rooted graphs and define the distance d between (G,r) and (G',r') as 2^{-k}. Then d is a metric on $\mathcal{X}$. Now we say that a sequence (G_N,r_N) of rooted graphs converges to a rooted graph (G,r) if and only if $\lim_{N\to\infty} d\big((G_N,r_N),(G,r)\big) = 2^{-\infty} := 0$.

In this work we specifically deal with two sequences converging to infinite vertex-transitive graphs (termed *lattices*):

i) Small graph sequence: where G_n is a finite cut-out of the lattice G that includes $B_G(r,n)$ but not $B_G(r,n+1)$.

In the context of planar lattices we construct a specific small graph sequence as follows. Define G_N^{Sq} the induced subgraph of G which is the cut-out of G contained within the square of side length $n \in \mathbb{N}^*$ centred on r (edges being given a length of 1). Here N stands for the number of vertices of this square cut-out, the precise relation between n and N depending on the underlying lattice G. We can now also define:

ii) Small torus sequence: where the small graphs G_N^{Sq} are wrapped around a torus. The resulting graphs G_N^{Tor} are vertex-transitive tori with the same degree as G.

Since all tori G_N^{Tor} as well as G are vertex-transitive, the roots r_N and r are irrelevant when considering convergence of this sequence. We therefore simply say that the sequence of graphs G_N^{Tor} converges to the lattice G.

The small torus sequence generalises to non-planar lattices by wrapping around genus $g > 2$ tori the small graphs G_N^{Hyp} defined as the induced subgraph of G which is the cut-out of G contained within the d-hypercube of

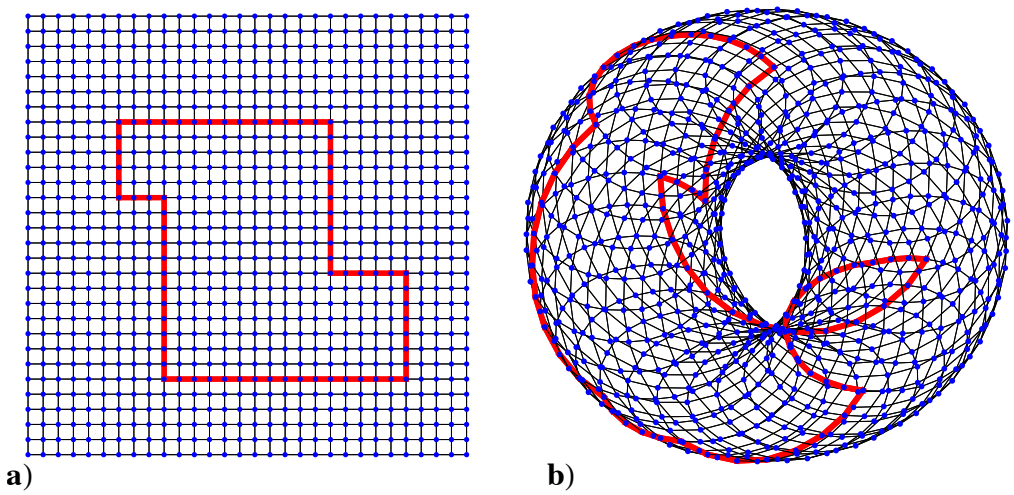


Figure 9: Examples of finite graphs part of the sequences converging to the (infinite) square lattice: **a)** small graph G_{30}^{Sq} with a self-avoiding polygon on it; **b)** corresponding small torus G_{30}^{Tor} with the same self-avoiding polygon.

side length $n \in \mathbb{N}^*$ centred on r. In this case, the dimension $d > 2$ is the smallest dimension of the space $\mathbb{R}^d$ such that G can be embedded in $\mathbb{R}^d$ in a way that no two edges cross.

Appendix B Viennot's lemma on infinite graphs with formal series

The proof is entirely similar to that given in the case of ordinary generating functions Proposition 5.1. We have:

Proposition B.1. *Let G be an infinite (di)graph with bounded degree with labelled adjacency matrix W. Let p be a prime on G, with support $\mathcal{V}(p)$ and neighborhood $\mathcal{N}(p)$. Let $G_p \prec G$ be the induced subgraph of G with vertex set $\mathcal{V}(G_p) = \mathcal{V}(p) \cup \mathcal{N}(p)$ and W_p its labelled adjacency matrix. Let $\mathsf{R} = \big(\mathsf{I} - \mathsf{W}\big)^{-1}\big|_{G_p}$ be the restriction of the formal resolvent of G to G_p.*

Then the formal series of closed walks whose last erased loop is p is given by

$$\sum_{\substack{w \ \ walk \\ p|_r w}} w = \det\big(\mathsf{I} + \mathsf{R}\,\mathsf{W}_p\big)p. \tag{16}$$

Appendix C Recursion relations for C and $\mathsf{R}(z)$

Recursion relations between entries of the resolvent naturally arise from the basic observations that walks of length ℓ going from vertex v_i to vertex v_j were walks of length $\ell-1$ from v_i to a neighbours of v_j. We have a similar relation relating walks of length ℓ with those of length $\ell-2$. In terms of generating functions, these read

$$\begin{aligned}
&\lambda z \mathsf{R}(z)_1 = R(z) - 1, \\
&z^2 \sum_i n_{2,i} \mathsf{R}(z)_{2,i} + \lambda z^2 R(z) = R(z) - 1
\end{aligned}$$

where $R(z)$, $\mathsf{R}(z)_1$ and $\mathsf{R}(z)_{2,i}$ designate a diagonal entry of the resolvent, the entry of the resolvent relating first neighbours, and entries of the resolvent relating second neighbours, respectively. The quantity $n_{2,i}$ is the number of times $\mathsf{R}(z)_{2,i}$ appears. Similar recursion relations are thus obeyed by the entries of C

$$c_1 = c_0 - 1, \quad \frac{1}{\lambda} c_0 + \sum_{i=1}^{\lfloor \lambda/2 \rfloor} n_{2,i} c_{2,i} = \lambda^2 (c_0 - 1).$$

where c_0, c_1 and $c_{2,i}$ are defined similarly to $R(z)$, $\mathsf{R}(z)_1$ and $\mathsf{R}(z)_{2,i}$ respectively.

These recursion relations simplify further on noting that $c_0 = 0$:

Lemma C.1. *Let G be an infinite vertex transitive graph with bounded degree λ and resolvent $\mathsf{R}(z)$. Let P_λ be the projector onto the eigenspace associated with eigenvalue λ and let $\mathsf{C}_G := \lim_{1/\lambda^-} \big(\mathsf{I} - \mathsf{P}_\lambda\big)\mathsf{R}(z)$. Then all diagonal entries c_0 of C_G are 0.*

Proof. Since G is regular $\big(\mathsf{I} - \mathsf{P}_\lambda\big)_{ij} = (1 - 1/N)\delta_{i,j} + (1/N)\delta_{i \neq j}$. Then

$$\Big(\big(\mathsf{I} - \mathsf{P}_\lambda\big)\mathsf{R}(z)\Big)_{kk} = \left(1 - \frac{1}{N}\right) \mathsf{R}(z)_{k,k} + \frac{1}{N} \sum_{j \neq k} \mathsf{R}(z)_{j,k} = 0.$$

□

Appendix D All orders of the determinant expansion

The expansion of determinants such as $\det(\mathsf{M}_1 + \epsilon \mathsf{M}_2)$ around $\det(\mathsf{M}_1)$ is well known. Here we report the slight variant valid when M_1 is singular. Indeed in the case of interest here $\mathsf{M}_1 \equiv \mathsf{I} - z\mathsf{A}_G$ which is singular in $z = 1/\lambda$. Taking into consideration that the adjugate matrix of a singular matrix with a kernel of dimension 1 is the projector P_λ onto that kernel times a constant which, for us, is asymptotically α^N, $N \gg 1$, the kth order of the expansion of $\det(\mathsf{I} - \lambda^{-1}\mathsf{A}_{G\backslash p})$ for an arbitrary SAP p is:

$$\mathrm{Order}^{(k)} = \frac{\alpha^N}{\lambda^k} \sum_{\pi \in \Pi(k)} (-1)^{|\pi|} \sum_{j=1}^{k} \left\{ \frac{1}{(\pi_j - 1)! j^{\pi_j - 1}} \mathrm{Tr}\left[\big(- z\mathsf{P}_\lambda \mathsf{B}_p \big) \big(- z\mathsf{C}\mathsf{B}_p \big)^{j-1} \right] \right.$$
$$\left. \times \mathrm{Tr}\left[\big(- z\mathsf{C}_G|_p \mathsf{B}_p \big)^{j} \right]^{\pi_j - 1} \prod_{\substack{m=1 \\ m \neq j}}^{k} \mathrm{Tr}\left[\big(- z\mathsf{C}_G|_p \mathsf{B}_p \big)^{m} \right]^{\pi_m} \frac{1}{m!\, m^{\pi_m}} \right\}.$$

In this expression, $\Pi(k)$ designates the set of partitions $\pi := \{\pi_i\}_{1 \leq i \leq k}$ of k such that $\sum_i i\pi_i = k$ and $|\pi| = \sum_i \pi_i$. The matrices $\mathsf{C}_G|_p$ and B_p are as defined in the Infinite Sieve Theorem.

Appendix E Error terms in the Infinite Sieve Theorem

The Infinite Sieve Theorem establishes the dominant term in the asymptotic expansion of the fraction of closed walks whose unique right prime factor is a given SAP p. This term is dominant only under the limit $\ell \to \infty$, where the closed walks under consideration have diverging length. In this section we evaluate the accompanying error terms which come into consideration for finite length walks. Concretely, these error terms cannot be neglected when this finite length is 'too close' to that of the SAP p, in a sense which will be made precise below.

The Infinite Sieve Theorem produces the asymptotic expansion of the extension of Viennot's Lemma to infinite graphs which gives, for the generating function $R_p(z)$ of walks whose last erased loop is the SAP p,

$$R_p(z) = z^{\ell(p)} \det\big(\mathsf{I} + z\mathsf{R}(z)\mathsf{B}_p\big).$$

The various divergences encountered in the asymptotic expansion were specifically avoided with the following form,

$$R_p(z) = R(z) \left(\frac{z^{\ell(p)}}{R(z)} \det\big(\mathsf{I} + z\mathsf{R}(z)\mathsf{B}_p\big) \right). \tag{17}$$

This presentation of $R_p(z)$ is also supported by its combinatorial meaning: with Lemma 5.4 the first term relates to the density of walks amongst the hikes, while the terms in parentheses give rise to the finite asymptotic fraction of walks whose last erased loop is p among all closed walks.

As we have seen, we should not expect any of the generating functions appearing here to be meromorphic in general. For this reason, we should only use properties of products of generating functions rather than ratios. One such standard property is recalled below:

Lemma E.1. *Let $F(z) = \sum_{\ell \geq 0} \lambda^\ell f(\ell) z^\ell$ with λ real and $f(\ell)$ a function such that $\lim_{\ell \to \infty} f(\ell)$ exists and is finite. Then for any function $G(z)$ of z such that $G(1/\lambda)$ is finite,*

$$\big(F(z)G(z)\big)[L] = \left(G(1/\lambda) + \sum_{k>0} \frac{\nabla^k[f](L)}{\lambda^k\, k!\, f(L)} G^{(k)}(1/\lambda) \right) \times F(z)[L],$$

where $G^{(k)}(1/\lambda)$ stands for the kth derivative of $G(z)$ evaluated in $1/\lambda$.

Returning to Eq. (17), we introduce the walk correction function $f_w(\ell) := R(z/\lambda)$, so that the number of closed walks of length ℓ on G is exactly $\lambda^\ell f_w(\ell)$. On the square lattice, we have

$$f_w(2\ell) = \binom{2\ell}{\ell}^2 4^{-2\ell} \sim \frac{1}{\ell\pi} \quad \text{for } \ell \gg 1,$$

while $f_w(2\ell+1) = 0$, and $\lim_{\ell\to\infty} f_w(\ell) = 0$. In general, on a d dimensional lattice $f_w(\ell)$ is on the order of $\ell^{-d/2}$ for $\ell \gg 1$.

We may now use Lemma E.1 on Eq. (17). Together with the Infinite Sieve Theorem, this gives

$$[z^\ell]R_p(z) = \left(\frac{F_p}{\lambda^{\ell(p)}} + \sum_{k\geq 1}^{\infty} \frac{\nabla^k[f_w](\ell - \ell(p))}{f_w(\ell)\lambda^k\, k!} \operatorname{Frac}_p^{(k)}(1/\lambda) \right) [z^\ell]R(z),$$

with $\operatorname{Frac}_p(z) := R_p(z)/R(z)$ and F_p as per Eq. (1). Let us denote,

$$\operatorname{Err}_p(2\ell) := \sum_{k\geq 1}^{\infty} \frac{\nabla^k[f_w](2\ell - \ell(p))}{f_w(2\ell)\lambda^k\, k!} \operatorname{Frac}_p^{(k)}(1/\lambda),$$

and observe that since $\operatorname{Frac}_p^{(k)}(1/\lambda)$ does not depend on ℓ, we may now estimate the decay of the correction terms with respect to the dominant contribution $F_p/\lambda^{\ell(p)}$ as $\ell \to \infty$. Taking $f_w(\ell) \sim \ell^{-d/2}$ as a guide on a d dimensional lattice, two situations arise:

Length L	2*	4*	6*	8*	10*	12*	14*
$S(L)$	0.5	0.6473	0.7093	0.7493	0.7774	0.7984	0.8149
Length L	16*	18*	20*	22*	24*	26*	28*
$S(L)$	0.8282	0.8392	0.8485	0.8565	0.8635	0.8696	0.8751
Length L	30	32					
$S(L)$	0.8799	0.8843					

Table 1: Table of value for the sum $S(L)$ as a function of L. The asterisk $*$ indicates that all calculations for this length were analytical.

i) If $\ell \gg 1$ and $\ell - \ell(p) \gg 1$, then the error terms decrease proportionally to $\big(\ell - \ell(p)\big)^{-1}$ on all d-dimensional vertex transitive lattices;

ii) If $\ell \gg 1$ but $\ell - \ell(p)$ is on the order of 1, then the error terms *actually increase* with ℓ, as $\ell^{d/2}$.

This analysis indicates that $[z^\ell]R(z) \times F_p/\lambda^{\ell(p)}$ is a good approximation to $[z^\ell]R_p(z)$ when ℓ is much larger than $\ell(p)$. This confirms the discussion of §6: using $F_p/\lambda^{\ell(p)}$ to obtain an estimate $\pi(L)$ via Eq. (14) is misguided precisely because most of the SAPs fall in situation ii), where using $F_p/\lambda^{\ell(p)}$ to estimate $[z^\ell]R_p(z)$ is outright wrong.

Appendix F Data table

Table 1 gives all computed values of

$$S(L) := \sum_{\substack{p:\,\mathrm{SAP} \\ \ell(p)\le L}} \frac{F_p}{4^{\ell(p)}}$$

as a function of L and evaluated on the square lattice. Although most tabulated values of $S(L)$ were *computed analytically*, we here report only numerical results rounded at 10^{-4} owing to length concerns. At length 32, $S(L)$ requires computing F_p values for $3,484,564,613$ self-avoiding polygons.

References

[1] S. Smirnov, Conformal invariance in random cluster models. I. Holomorphic fermions in the Ising model, Annals of Mathematics 172 (2010) 1435–1467. doi:10.4007/annals.2010.172.1435.

[2] W. Werner, The conformally invariant measure on self-avoiding loops, Journal of the American Mathematical Society 21 (2000) 137–169. doi:10.1090/S0894-0347-07-00557-7.
URL https://doi.org/10.1090/S0894-0347-07-00557-7

[3] G. Lawler, W. Wendelin, Universality for conformally invariant intersection exponents, Journal of the European Mathematical Society 2 (2000) 291–328. doi:10.1007/s100970000024.

[4] V. Beffara, H. Duminil-Copin, Planar percolation with a glimpse of Schramm-Loewner evolution, Probability Surveys 10 (2013) 1–50. doi:10.1214/11-PS186.
URL https://doi.org/10.1214/11-PS186

[5] B. Mandelbrot, The fractal geometry of nature, Freeman, San Francisco, CA, 1982.
URL https://cds.cern.ch/record/98509

[6] G. Lawler, O. Schramm, W. Werner, The dimension of the planar brownian frontier is 4/3, Mathematical Research Letters 8 (2001) 401–411. doi:10.4310/MRL.2001.v8.n4.a1.
URL http://dx.doi.org/10.4310/MRL.2001.v8.n4.a1

[7] G. Lawler, O. Schramm, W. Werner, On the scaling limit of planar self-avoiding walk, in: Proceedings of Symposia in Pure Mathematics, Vol. 72, 2004, pp. 339–364.

[8] V. Beffara, The dimension of the SLE curves, The Annals of Probability 36 (4) (2008) 1421–1452. doi:10.1214/07-AOP364.
URL https://doi.org/10.1214/07-AOP364

[9] H. Duminil-Copin, S. Smirnov, The connective constant of the honeycomb lattice equals $\sqrt{2+\sqrt{2}}$, Annals of Mathematics 175 (3) (2012) 1653–1665.
URL http://www.jstor.org/stable/23234646

[10] G. Lawler, A self-avoiding random walk, Duke Mathematical Journal 47 (3) (1980) 655–693. doi:10.1215/S0012-7094-80-04741-9.
URL https://doi.org/10.1215/S0012-7094-80-04741-9

[11] G. Lawler, W. Werner, The Brownian loop soup, Probability Theory and Related Fields 128 (4) (2004) 565–588. doi:10.1007/s00440-003-0319-6.
URL https://doi.org/10.1007/s00440-003-0319-6

[12] G. Lawler, J. Trujillo Ferreras, Random walk loop soup, Transactions of the American Mathematical Society 359 (2007) 767–787. doi:10.1090/S0002-9947-06-03916-X.
URL https://doi.org/10.1090/S0002-9947-06-03916-X

[13] G. Lawler, Conformally invariant loop measures, in: Proceedings of the International Congress Of Mathematicians, Rio de Janeiro, Vol. 1, 2018, pp. 669–704.

[14] P. Cartier, D. Foata, Problèmes combinatoires de commutation et réarrangements, Vol. 85 of Lecture Notes in Mathematics, Springer-Verlag Berlin Heidelberg, 1969.

[15] G. X. Viennot, Heaps of pieces, i : Basic definitions and combinatorial lemmas, in: G. Labelle, P. Leroux (Eds.), Combinatoire énumérative, Springer Berlin Heidelberg, Berlin, Heidelberg, 1986, pp. 321–350.

[16] C. Krattenthaler, The theory of heaps and the Cartier-Foata monoid, complement to the electronic republication to Cartier's and Foata's *Problèmes combinatoires de commutation et réarrangements*. (2006).

[17] P.-L. Giscard, P. Rochet, Algebraic combinatorics on trace monoids: Extending number theory to walks on graphs, SIAM Journal on Discrete Mathematics 31 (2) (2017) 1428–1453. arXiv:https://doi.org/10.1137/15M1054535, doi:10.1137/15M1054535.
URL https://doi.org/10.1137/15M1054535

[18] D. Atkinson, F. J. van Steenwijk, Infinite resistive lattices, American Journal of Physics 67 (6) (1999) 486–492. arXiv:https://doi.org/10.1119/1.19311, doi:10.1119/1.19311.
URL https://doi.org/10.1119/1.19311

[19] J. Cserti, G. Széchenyi, G. Dávid, , Journal of Physics A: Mathematical and Theoretical 44 (21) (2011) 215201. doi:10.1088/1751-8113/44/21/215201.
URL https://doi.org/10.1088%2F1751-8113%2F44%2F21%2F215201

[20] S. N. Majumdar, D. Dhar, Height correlations in the Abelian sandpile model, Journal of Physics A: Mathematical and General 24 (7) (1991) L357–L362. doi:10.1088/0305-4470/24/7/008.

[21] S. S. Manna, D. Dhar, S. N. Majumdar, Spanning trees in two dimensions, Physical Review A 46 (1992) R4471–R4474. doi:10.1103/PhysRevA.46.R4471.
URL https://link.aps.org/doi/10.1103/PhysRevA.46.R4471

[22] G. X. Viennot, Heaps of pieces, I: Basic definitions and combinatorial lemmas, Annals of the New York Academy of Sciences 576 (1) (1989) 542–570.

[23] P. Flajolet, R. Sedgewick, Analytic combinatorics, Cambridge University Press, 2009.

[24] J. A. Fill, P. Flajolet, N. Kapur, Singularity analysis, Hadamard products, and tree recurrences, Journal of Computational and Applied Mathematics 174 (2) (2005) 271–313. doi:10.1016/j.cam.2004.04.014.

[25] T. Espinasse, P. Rochet, A coupling of the spectral measures at a vertex, The Electronic journal of Combinatorics 26.

[26] R. Kenyon, The asymptotic determinant of the discrete Laplacian, Acta Mathematica 185 (2) (2000) 239–286. doi:10.1007/BF02392811.
URL https://doi.org/10.1007/BF02392811

[27] M. Bousquet-Mélou, Q-Énumération de Polyominos Convexes, Journal of Combinatorial Theory, Series A 64 (2) (1993) 265–288. doi:10.1016/0097-3165(93)90098-S.

[28] M. Bousquet-Mélou, q-Énumération de Polyominos Convexes, Ph.D. thesis, Université du Québec à Montréal, Département de mathématiques et d'informatique (1991).

[29] I. Benjamini, O. Schramm, Recurrence of distributional limits of finite planar graphs, Electron. J. Probab. 6.