CYCLE-CENTRALITY

COUNTING GRAPH WALKS WITH NUMBER THEORY SIEVES
OUTLINE

- Motivations
  - On the notion of importance in networks
  - A 70 year-old problem
  - Extending number theory to walks on graphs

- Counting walks with sieves
  - What is a sieve ?!
  - Finite graphs results
  - Infinite graphs results and $O(1)$ computations

- Outlook
ON THE NOTION OF IMPORTANCE IN NETWORKS

- Vertex centrality: how important is a node in a network?

\[ \text{Importance} = \text{degree} \]
ON THE NOTION OF IMPORTANCE IN NETWORKS

- Vertex centrality: how important is a node in a network?

Lightbulb: \[ \text{degree} = \# \text{neighbours} \]

Box: \[ \text{Importance} = \# \text{neighbours} + \# \text{2nd neighbours} \]
ON THE NOTION OF IMPORTANCE IN NETWORKS

- Vertex centrality: how important is a node in a network?

💡 # N-neighbours ≡ counting walks

\[
\text{Importance} = \sum_{\text{walks: } v \rightarrow v} 1
\]

doesn’t converge!
ON THE NOTION OF IMPORTANCE IN NETWORKS

- Vertex centrality: how important is a node is a network?

💡 # N-neighbours = counting walks

\[
\text{Importance} = \sum_{\text{walks: } v \rightarrow v} \alpha^{\ell(w)} \quad \text{converges} \quad \alpha < \lambda_{\text{max}}
\]

\[
\sum_{\text{walks: } v \rightarrow v} \frac{1}{\ell(w)!} \quad \text{always converges}
\]

- Centrality question: how many walks start at \( v \)?

- Kills off the long walks!
A 70 YEAR OLD PROBLEM
A 70 YEAR OLD PROBLEM

Polymer chains

Self-Avoiding Walk (SAW)

Self-Avoiding Polygon (SAP)
A 70 YEAR OLD PROBLEM

Polymer chains

Self-Avoiding Walk (SAW)

Self-Avoiding Polygon (SAP)

QUESTION (~1957)
How many SAPs of length $\ell$ are there as $\ell \to \infty$?

$\mu^\ell \ell^{-5/2}$

« A widely open problem » Flajolet & Sedgewick
A 70 YEAR OLD PROBLEM

- SAWs and SAPs are ‘everywhere’

- Works of Schramm, Smirnov, Werner, Lawler…

- In all random models, SAWs & SAPs should obey a single probability law : SLE$_K$
A 70 YEAR OLD PROBLEM

- SAWs and SAPs are ‘everywhere’

- Works of Schramm, Smirnov, Werner, Lawler…. 

- In all random models, SAWs & SAPs should obey a single probability law: SLE$K$ 

STILL A CONJECTURE WIZ 1957-PROBLEM 

2019 Wolf Prize
Gregory Lawler (2019 Wolf Prize, 1987 work)

SAWs generation from loop-erasing

Lawler’s question: how many walks give the SAW $\rho$?

Generalises the centrality question
For any two closed walks \( w, w' \) if a loop \( \gamma \) is erased from \( w \cdot w' \) then it is erased from either \( w \) or \( w' \)
For any two closed walks $w, w'$ if a loop $\gamma$ is erased from $w.w'$ then it is erased from either $w$ or $w'$

$$\gamma | w.w' \Rightarrow \gamma | w \text{ or } \gamma | w'$$

A number theory for walks ?!
OUTLINE

‣ Motivations
  ‣ On the notion of importance in networks
  ‣ A 70 year-old problem
  ‣ Extending number theory to walks on graphs

‣ Counting walks with sieves
  ‣ What is a sieve ?!
  ‣ Finite graphs results
  ‣ Infinite graphs results and O(1) computations

‣ Outlook
EXTENDING NUMBER THEORY

- Monoid: words on simple cycles & ordinary multiplication
  \[ \gamma_1 \gamma_2 \cdots \gamma_n \]

  **Rule 1**
  *Orientation of cycles is retained*
  *Starting point is disregarded*

  **Rule 2**
  *Two words commute if they have no vertices in common*
  \[ h \cdot h' = h' \cdot h \iff \mathcal{V}(h) \cap \mathcal{V}(h') = \emptyset \]

- Hikes: words on cycles modulo \( R1 \) and \( R2 \)

Cartier & Foata (1969)
Viennot (1987)
Various types of hikes

**Hike**
Anything made up of closed walks!

**Self-avoiding hike**
Disjoint cycles

**Closed Walk**
Walks are hikes with a unique right prime divisor $w = hp$

**Prime**
One cycle
How to build a theory in 3 steps!

1. Simple cycles are **prime** in the monoid of hikes

\[
\gamma | h.h' \Rightarrow \gamma | h \text{ or } \gamma | h'
\]

2. We can consider the poset of hikes **ordered by divisibility**

\[
h | h' \Rightarrow h \leq h'
\]

3. Formal series on hikes form an algebra obeying a **plethora** of relations

\[
\sum_{\text{Hikes}} f(h)h
\]

Rota et al. (1964-1974)
EXTENDING NUMBER THEORY

**THEOREM**
There exist graphs such that\[\sum_{\text{Hikes}} f(h)h\] reduces to\[\sum_{n>0} \frac{a(n)}{n^s}\]

*Number theory recovered as a special case*
## EXTENDING NUMBER THEORY

<table>
<thead>
<tr>
<th>Hikes</th>
<th>Number Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$ hike</td>
<td>$n$ integer</td>
</tr>
<tr>
<td>$h$, $h'$ disjoint</td>
<td>$n$, $m$ coprime</td>
</tr>
<tr>
<td>$h$ self-avoiding</td>
<td>$n$ square-free</td>
</tr>
<tr>
<td>$p$ simple cycle</td>
<td>$p$ prime</td>
</tr>
<tr>
<td>$w$ walk</td>
<td>$n = p^k$</td>
</tr>
</tbody>
</table>
## Extending Number Theory

<table>
<thead>
<tr>
<th></th>
<th>Hikes</th>
<th>Number theory</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Zeta</strong></td>
<td>$\zeta = S1 = \sum_{h \in \mathcal{H}} h$</td>
<td>$\zeta_R(s) = \sum_{n &gt; 0} \frac{1}{n^s}$</td>
</tr>
<tr>
<td></td>
<td>$\zeta = \frac{1}{\det(I-W)}$</td>
<td></td>
</tr>
<tr>
<td><strong>Möbius</strong></td>
<td>$\mu(h) = \begin{cases} (-1)^{\Omega(h)}, &amp; h \text{ self-avoiding} \ 0, &amp; \text{otherwise.} \end{cases}$</td>
<td>$\mu(n) = \begin{cases} (-1)^{\Omega(n)}, &amp; n \text{ square-free} \ 0, &amp; \text{otherwise.} \end{cases}$</td>
</tr>
<tr>
<td><strong>von Mangoldt</strong></td>
<td>$\Lambda(h) = \begin{cases} \ell(p), &amp; h \text{ walk, } p</td>
<td>h \ 0, &amp; \text{otherwise.} \end{cases}$</td>
</tr>
<tr>
<td></td>
<td>$S\Lambda = \text{Tr} \left[ (I - W)^{-1} \right]$</td>
<td></td>
</tr>
<tr>
<td><strong>Liouville</strong></td>
<td>$\lambda(h) = (-1)^{\Omega(h)}$</td>
<td>$\lambda(n) = (-1)^{\Omega(n)}$</td>
</tr>
<tr>
<td></td>
<td>$S\lambda = \frac{1}{\text{perm}(I-W)}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>Number of divisors</td>
<td>$\zeta^2$</td>
<td>$\zeta_R(s)^2$</td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td>Log Zeta</td>
<td>$\log \zeta = \sum_h \frac{\Lambda(h)}{\ell(h)} h$</td>
<td>$\log \zeta_R(s) = \sum_n \frac{\Lambda(n)}{\log(n)} \frac{1}{n^s}$</td>
</tr>
<tr>
<td>Log-Mangoldt</td>
<td>$\ell(h) = \sum_{h'</td>
<td>h} \Lambda(h')$</td>
</tr>
<tr>
<td>Totally multiplicative functions</td>
<td>$f^{-1} = \sum_h \mu(h) f(h) h$</td>
<td>$f^{-1} = \sum_n \frac{\mu(n) f(n)}{n^s}$</td>
</tr>
<tr>
<td></td>
<td>$f'/f = - \sum_h \Lambda(h) f(h) h$</td>
<td>$f'/f = \sum_n \frac{\Lambda(n) f(n)}{n^s}$</td>
</tr>
<tr>
<td>Totally additive functions</td>
<td>$(f * \mu)(h) = \begin{cases} f(p), &amp; \text{h walk, } p</td>
<td>h \ 0, &amp; \text{otherwise.} \end{cases}$</td>
</tr>
<tr>
<td>$\zeta'/\zeta$ from the primes</td>
<td>$- \sum_{\gamma: \text{simple cycle}} \ell(\gamma) \frac{\det(I - W_\gamma)}{\det(I - W)}$</td>
<td>$- \sum_p \log p \frac{p^{-s}}{1 - p^{-s}}$</td>
</tr>
<tr>
<td>Number $\Omega$ of prime factors</td>
<td>$\sum_{w: \text{walk}} w = \det(I - W) \sum_{h \in \mathcal{H}} \Omega(h) h$</td>
<td>$\sum_{p,n} \frac{1}{p^{-ns}} = \zeta_R(s)^{-1} \sum_n \frac{\Omega(n)}{n^s}$</td>
</tr>
</tbody>
</table>
EXTENDING NUMBER THEORY

QUESTION (~1957)
How many SAPs of length $\ell$ are there as $\ell \to \infty$ ?

$\mu^\ell \ell^{-5/2}$

« A widely open problem » Flajolet & Sedgewick
EXTENDING NUMBER THEORY

**QUESTION (≈1957)**
How many SAPs of length \( \ell \) are there as \( \ell \to \infty \) ?

\[
\mu^\ell \ell^{-5/2}
\]

« *A widely open problem* » Flajolet & Sedgewick

**It is the extension of the prime number theorem!**

- Lawler’s question: how many walks give a SAP \( y \) ? SAW \( \rho \) ?

*How many integers are of the form \( p^k \) ?*
Motivations
- On the notion of importance in networks
- A 70 year-old problem
- Extending number theory to walks on graphs

Counting walks with sieves
- What is a sieve ?!
- Finite graphs results
- Infinite graphs results and O(1) computations

Outlook
WHAT IS A SIEVE?!

Hikes that are multiples of primes we know

Those primes we know

Natural choice for $\mathcal{P}$: all cycles on a subgraph $H$ of $G$

$\mathcal{P}$ unknown

Walks multiples

Big green bag: all hikes up to length $\ell$

$\mathcal{H}_\ell$

Hikes that are not multiples of the primes we know $S(\mathcal{H}_\ell, \mathcal{P})$
What is a sieve?!

What a sieve counts.

\( \mathcal{P} \text{ unknown} \)

Hikes that are not multiples of the primes we know \( S(\mathcal{H}_\ell, \mathcal{P}) \).
WHAT IS A SIEVE ?!

A sieve counts *hikes* which are *not* multiples of any cycle on subgraph $H$

- Pick $H = G \setminus p$
- So $h$ is counted by the sieve if it is not a multiple of cycles on $G \setminus p$
WHAT IS A SIEVE?!

A sieve counts **hikes** which are not multiples of any cycle on subgraph $H$

- Pick $H = G \setminus p$
- So $h$ is counted by the sieve if it is not a multiple of cycles on $G \setminus p$
- Hence $h$ is a multiple of cycles that are on $G$ but not $G \setminus p$ : such cycles cross $p$
WHAT IS A SIEVE ?!

A sieve counts hits which are not multiples of any cycle on subgraph $H$

- Pick $H = G \setminus p$
- So $h$ is counted by the sieve if it is not a multiple of cycles on $G \setminus p$
- Hence $h$ is a multiple of cycles that are on $G$ but not $G \setminus p$ : such cycles cross $p$
- Thus $p$ commutes with none of the right divisors of $h$
i.e. $w = hp$ is a walk!

On $G \setminus p$ the sieve counts the walk multiples of $p$
The sieve is just an inclusion-exclusion in the poset of hikes ordered by divisibility.

\[ S(\mathcal{H}_\ell, \mathcal{P}_{G\setminus p}) = \sum_{d \in \mathcal{P}_{G\setminus p}^{s,a}} \mu(d) |M(d)_\ell| \]

- **Möbius function on hikes**
- **Number of multiples of d of length up to \( \ell \)**
- **Self-avoiding hikes on \( G\setminus p \)**
- **Sieve of Eratosthenes**
SIEVING ON FINITE GRAPHS

The sieve is just an inclusion-exclusion in the poset of hikes ordered by divisibility

\[ S(\mathcal{H}_\ell, \mathcal{P}_{G \setminus p}) = \sum_{d \in \mathcal{P}_{G \setminus p}^{s,a}} \mu(d) |\mathcal{M}(d)_{\ell}| \]

Skipping a lot of technical details ...
SIEVING ON FINITE GRAPHS

The sieve is just an inclusion-exclusion in the poset of hikes ordered by divisibility

\[
S(\mathcal{H}_\ell, \mathcal{P}_{G\setminus p}) = \sum_{d \in \mathcal{P}_{G\setminus p}^{s,a}} \mu(d) \left| \mathcal{M}(d)_\ell \right|
\]

Skipping a lot of technical details ...
SIEVING ON FINITE GRAPHS

The sieve is just an *inclusion-exclusion* in the poset of hikes ordered by divisibility

\[ S(\mathcal{H}_\ell, \mathcal{P}_{G\backslash p}) = \sum_{d \in \mathcal{P}_G^{s,a}} \mu(d) |\mathcal{M}(d)_\ell| \]

Skipping a **lot** of technical details …

\[
S(\mathcal{H}_\ell, \mathcal{P}_{G\backslash p}) = |\mathcal{H}_\ell| \det \left( I - \frac{1}{\lambda} A_{G\backslash p} \right) + \lambda^\ell \sum_{k \geq 1} \frac{\nabla^k[f](\ell)}{\lambda^k k!} \det^{(k)} \left( I - \frac{1}{\lambda} A_{G\backslash p} \right)
\]

\('Nice bit' \quad \text{\textcolor{orange}{`Nasty bit'}}
SIEVING ON FINITE GRAPHS

\[ S(H_\ell, P_{G\setminus p}) = |H_\ell| \det \left( I - \frac{1}{\lambda} A_{G\setminus p} \right) + \lambda^\ell \sum_{k \geq 1} \frac{\nabla^k[f](\ell)}{\lambda^k k!} \det^{(k)} \left( I - \frac{1}{\lambda} A_{G\setminus p} \right) \]

- **Nice bit** is between 0 and 1 and **dominates asymptotically** as \( \ell \to \infty \)

\[ \frac{S(H_\ell, P_{G\setminus p})}{|H_\ell|} \sim \det \left( I - \frac{1}{\lambda} A_{G\setminus p} \right) \quad \ell \to \infty \]

- **Nasty bit** is made of correction terms, vanishingly small for \( \ell \gg N \)

- Determinant
- Identity matrix
- Adjacency matrix of subgraph H

- Dominant eigenvalue of the entire graph G
SIEVING ON FINITE GRAPHS

\[
\frac{S(\mathcal{H}_\ell, \mathcal{P}_{G\setminus p})}{|\mathcal{H}_\ell|} \sim \det \left( I - \frac{1}{\lambda} A_{G\setminus p} \right) \quad \ell \to \infty
\]

**Meaning?**

Total fraction of hikes \( h \) such that \( hp \) is a walk \( w \)

- **Cycle** \( p \) is the unique prime divisor of \( w \) on the right \( w = hp \)
- **Walk** \( w \) can be started from any vertex of \( p \)
- **Cycle** \( p \) is the last erased loop of \( w \)
SIEVING ON FINITE GRAPHS

\[
\det \left( I - \frac{1}{\lambda} A_{G \setminus p} \right)
\]

Centrality of the cycle \( p \)

Huge success in network analysis!

- Case \( p \equiv \bullet \) gives the fraction of hikes which are walks starting at \( \bullet \)
- Induces eigenvector centrality (PageRank) on vertices
- Applications in biology, machine learning, economy, sociology...
Problem: which proteins are targeted by pathogens?

25% increase in the quality of pathogen targets predictions as compared to a degree-based model

OUTLINE

- Motivations
  - On the notion of centrality
  - A 70 year-old problem
  - Extending number theory to walks on graphs

- Counting walks with sieves
  - What is a sieve ?!
  - Finite graphs results
  - Infinite graphs results and O(1) computations

- Outlook
SIEVING ON INFINITE GRAPHS

\[ \zeta(z) = \exp \left( -\frac{i\pi}{2} - \int \frac{1}{z} \left( 1 - \frac{2}{\pi} K(16z^2) \right) dz \right) \]

\[ 1 + zR(z)B_p = C B_k - \frac{1}{\lambda\pi} P\lambda B_p \log(1 - z\lambda) + \cdots \]

Well it’s tough
Sieve still works! **Nice bit** is still **dominant**!

*Fraction of walks which are multiples of $p$ with respect to all walks*

\[
Frac_p = \frac{1}{\lambda^\ell(p)} v^T \cdot \text{adj}(C.B_p).u
\]
Sieving on Infinite Graphs

- Sieve still works! **Nice bit** is still **dominant**!

*Fraction of walks which are multiples of p with respect to all walks*

\[
\text{Frac}_p = \frac{1}{\lambda^{\ell(p)}} v^T \cdot \text{adj}(C.B_p) \cdot u
\]

- Dominant eigenvalue of the graph
- Length of cycle p
- Matrix encoding the `shape' of p

\[
u^T \cdot \lambda B_p = \frac{1}{\lambda} P^\lambda \cdot B_p
\]

C depends on the graph in a complicated manner, e.g. on square lattice:

\[
C_{m,n} = -\frac{1}{\pi} \int_0^\infty \frac{1}{\tau} \left(1 - \left(\frac{\tau - i}{\tau + i}\right)^{m-n} \left(\frac{\tau - 1}{\tau + 1}\right)^{m+n}\right) \, d\tau
\]
SIEVING ON INFINITE GRAPHS

- Answers Lawler’s question!

\[
\frac{128(\pi - 2)}{4^4 \times \pi^3} \approx 0.0184
\]

1.8% of all walks give the square as last erased loop.
SIEVING ON INFINITE GRAPHS

Answers Lawler’s question!

\[
\frac{128(\pi - 2)}{4^4 \times \pi^3} \approx 0.0184
\]

1.8% of all walks give the square as last erased loop \(\gamma\)

**QUESTION (~1957)**

How many SAPs of length \(\ell\) are there as \(\ell \to \infty\)?

\[\mu^\ell \ell^{-5/2}\]

Gives a series for \(\mu\)!

\[
8 \ln\left(\frac{\mu}{\lambda}\right) + \ln(8) = \ln\left(1 - 4 \frac{2}{\lambda^2} - 8 \frac{128(\pi - 2)}{\pi^3 \lambda^4} - 24 \frac{32(\pi - 8)(\pi - 4)(3\pi - 8)(3\pi - 4)}{\pi^4 \lambda^6} + \cdots\right)
\]
OUTLOOK

Motivations

- On the notion of importance in networks
- A 70 year-old problem
- Extending number theory to walks on graphs

Counting walks with sieves

- What is a sieve ?!
- Finite graphs results
- Infinite graphs results and $O(1)$ computations

Outlook
OUTLOOK

QUESTION (~1957)
How many SAPs of length $\ell$ are there as $\ell \to \infty$?

$\mu^\ell \ell^{-5/2}$

Still out of reach

- How fast do we converge to $\text{Frac}_p$?
- How can we prove that $\sum_{\ell(p) \leq \Lambda} \text{Frac}_p \sim \Lambda^{-3/2}$?
MERCI

THANK YOU