Extending number theory to walks on graphs

Why? how? and so what?

P.-L. Giscard, P. Rochet, T. Espinasse, R. Wilson

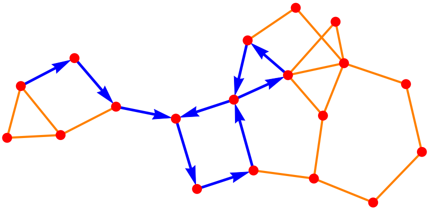
Keble College Networks Cluster

25th April 2015

Outline

1. Why?
   - Network analysis
2. How?
   - From walks to hikes
   - Posets and trace monoids
   - Hopf and Lie algebras
3. So what?
   - Getting the primes
   - Graph characterisation
   - Cycle Double Cover Conjecture
4. Conclusion

Network analysis

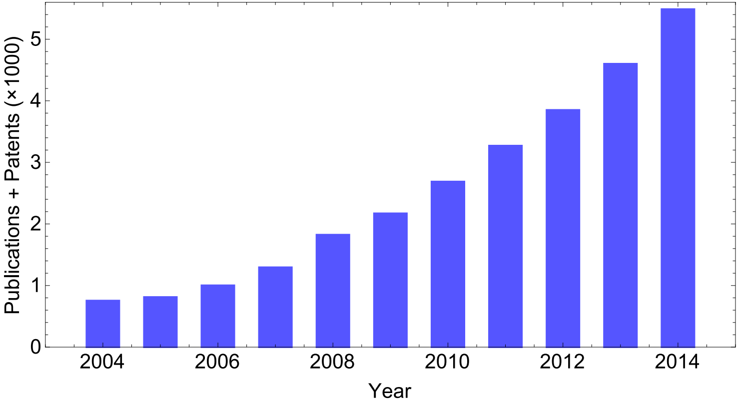


- Network G: relations (edges) between entities (nodes)
- Walk: trajectory on G, dynamics of the system

Graph modelling		*Complex systems*
Network / walk	$\Longleftrightarrow$	topology / dynamics

Network analysis

- Walk-based methods



Data: Google scholar

⟶ 2016: around 23 papers / day

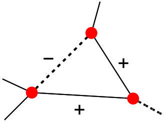
⟶ Biology, social sciences, economy, chemistry, computer science, physics, mathematics

Walk-based network analysis

- Theoretical underpinnings poorly understood
- Lack of tools to answer simple questions

Example: frustration (balance) in networks

$\hookrightarrow$ Are (social) networks balanced or not?



i) Yes! G. Facchetti, G. Iacono, and C. Altafini, Proc. Natl. Acad. Sci. USA **108**, 20953 (2011)

ii) No! E. Estrada, B. Benzi, Phys. Rev. E **90**, 042802 (2014)

Simple cycles **HARD** to count:

i) ground state Ising spin glass $\Longleftrightarrow$ number of changes to get perfect balance

ii) approx. by walks $\mathrm{Tr}(\exp \beta A) = \sum_{walks} \mathrm{sign}(w) \frac{\beta^{\ell}}{\ell(w)!}$

Back to the drawing board

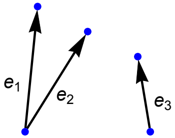
1. Why?
   - Network analysis
2. How?
   - From walks to hikes
   - Posets and trace monoids
   - Hopf and Lie algebras
3. So what?
   - Getting the primes
   - Graph characterisation
   - Cycle Double Cover Conjecture
4. Conclusion

From walks to hikes

Cartier-Foata 1960s: edge-based description

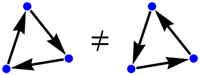
- All sequences of edges $e_1 e_3 e_2 e_6 \ldots$.
- Commutation rule: $e_1 e_2 \neq e_2 e_1$

  $e_1 e_3 = e_3 e_1$

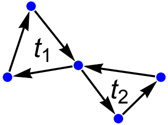


Orientation preserved

Starting point irrelevant



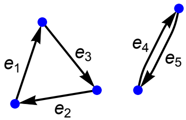
Walk-traversal order preserved $t_1 t_2 \neq t_2 t_1$



P. Cartier, D. Foata, Problemes combinatoires de commutation et rearrangements, Lec. Math. **85** (1969)

Hikes

Hikes equivalence classes on sequences of edges

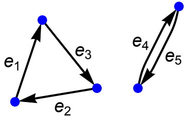


This talk:
collections of cycles

$$h = [e_1 e_2 e_3 e_4 e_5] = [e_2 e_4 e_1 e_3 e_5] = [e_5 e_4 e_3 e_2 e_1] = \ldots$$

Hikes

Hikes equivalence classes on sequences of edges



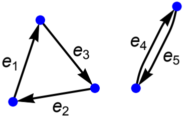
This talk:
collections of cycles

$$h = [e_1 e_2 e_3 e_4 e_5] = [e_2 e_4 e_1 e_3 e_5] = [e_5 e_4 e_3 e_2 e_1] = ...$$

- Multiply by concatenation
- h, h' commute $\iff$ disjoint
- Bunch of semi-commutative objects, *trace monoid*

Hikes

Hikes equivalence classes on sequences of edges

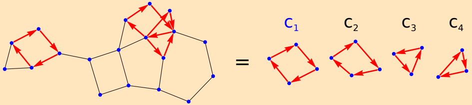


This talk:
collections of cycles

$$h = [e_1 e_2 e_3 e_4 e_5] = [e_2 e_4 e_1 e_3 e_5] = [e_5 e_4 e_3 e_2 e_1] = \ldots$$

- Multiply by concatenation
- h, h' commute $\iff$ disjoint
- Bunch of semi-commutative objects, *trace monoid*

Hikes form a unique factorization trace monoid

$h =$ 

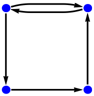
$h = c_1 c_2 c_3 c_4 = c_2 c_1 c_3 c_4 = \ldots$ primes! $c \,|\, h.h' \Rightarrow c \,|\, h$ or $c \,|\, h'$

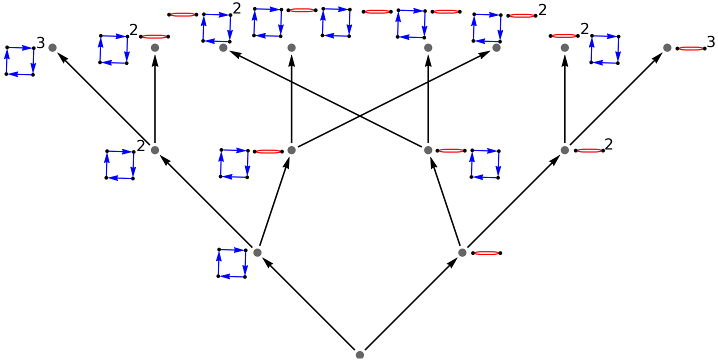
Posets of hikes

Construct a prime-based representation of hikes

- Sets of hikes ordered by left divisibility

$$h \mid h' \iff h \leq h'$$

- Example: poset of hikes $P_G = (\mathcal{H}, \leq)$ on $G =$ 



Hikes: structure

How can we exploit P_G?

$\hookrightarrow$ Matrix encoding P_G, $Z(h,h') = 1 \iff h \leq h'$

$\hookrightarrow$ Functions on P_G, $M(h,h') \neq 0 \iff h \leq h'$

Hikes: structure

How can we exploit P_G?

$\longrightarrow$ Matrix encoding P_G, $Z(h,h') = 1 \iff h \leq h'$

$\longrightarrow$ Functions on P_G, $M(h,h') \neq 0 \iff h \leq h'$

Theorem

All matrices on $P_G \simeq$ series on hikes $\sum_{hikes} f(h)h$

Hikes: structure

How can we exploit P_G?

$\hookrightarrow$ Matrix encoding P_G, $Z(h,h') = 1 \iff h \leq h'$

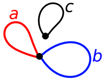
$\hookrightarrow$ Functions on P_G, $M(h,h') \neq 0 \iff h \leq h'$

Theorem

All matrices on $P_G \simeq$ series on hikes $\sum_{hikes} f(h)h$

$$Z \longrightarrow \zeta = \sum_{hikes} h = \frac{1}{\det(\mathrm{I} - \mathrm{A}_G)}$$

⚠ ζ semi-commutative

G = 

$\longrightarrow \zeta = 1 + a + b + c$

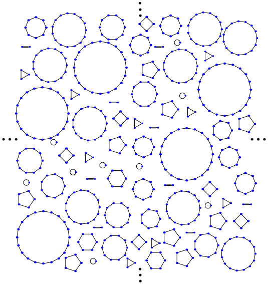
$+ab + ba + ac + \cdots$

Graph $G_{\mathbb{N}}$: number theory

Algebra series on hikes on $G_{\mathbb{N}}$ $\Longrightarrow$ Dirichlet series $\sum_n a_n/n^s$

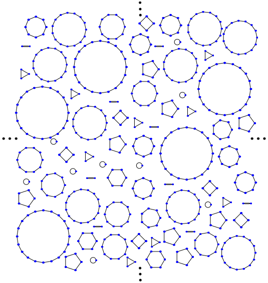
Number theory from hikes!

- Hikes $\longrightarrow n$
- Walks $\longrightarrow p^k$
- Walk length $\longrightarrow \log n$
- h_1, h_2 disjoint $\longrightarrow n, m$ coprime



Graph $G_{\mathbb{N}}$: number theory

Algebra series on hikes on $G_{\mathbb{N}}$ $\Longrightarrow$ Dirichlet series $\sum_n a_n/n^s$



Number theory from hikes!

- Hikes $\rightarrow n$
- Walks $\rightarrow p^k$
- Walk length $\rightarrow \log n$
- h_1, h_2 disjoint $\rightarrow n, m$ coprime

$$\zeta \Longrightarrow \text{Riemann } \zeta_R$$

$$1/\zeta = \det(\mathsf{I} - z\mathsf{A}_G) \Longrightarrow \text{Möbius function } \mu$$

$$\Lambda[h] = \mathrm{Tr}\,(\mathsf{I} - z\mathsf{A}_G)^{-1}[h] \Longrightarrow \text{von Mangoldt } \Lambda(n)$$

$$\big(1/\mathrm{perm}(\mathsf{I} + z\mathsf{A}_G)\big)[h] \Longrightarrow \text{Liouville } \lambda[n] = (-1)^{\Omega[n]}$$

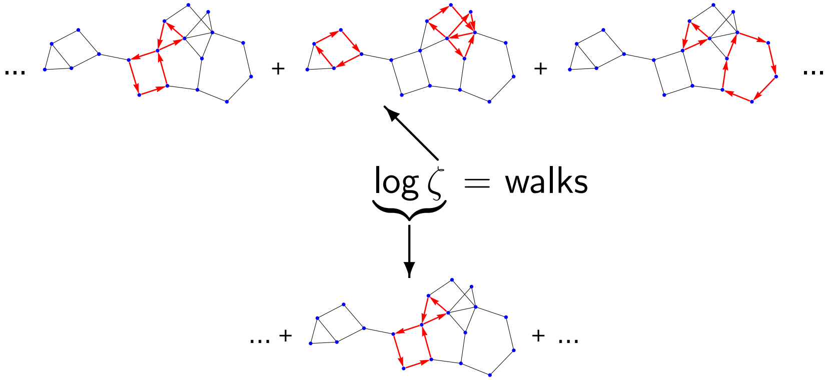
Extension of number theory

	"Hike theory", any G	Number theory $G_{\mathbb{N}}$
Number of divisors	ζ^2	ζ_R^2
Log zeta	$\log\zeta = \sum_h \frac{\Lambda(h)}{\ell(h)} h$	$\log\zeta_R = \sum_n \frac{\Lambda(n)}{\log(n)} \frac{1}{n^s}$
Log-Mangoldt	$\ell(h) = \sum_{h'\mid h} \Lambda(h')$	$\log(n) = \sum_{d\mid n} \Lambda(n)$
Totally mult. functions $f(ab) = f(a)f(b)$	$f^{-1} = \sum_h \mu(h) f(h) h \,\star$	$f^{-1} = \sum_n \frac{\mu(n)f(n)}{n^s}$
	$f'/f = -\sum_h \Lambda(h) f(h) h$	$f'/f = \sum_n \frac{\Lambda(n)f(n)}{n^s}$
$\vdots$		

$\star$ Gives rise to and generalize MacMahon master theorem

Algebraic Walk Theory

A strange fact

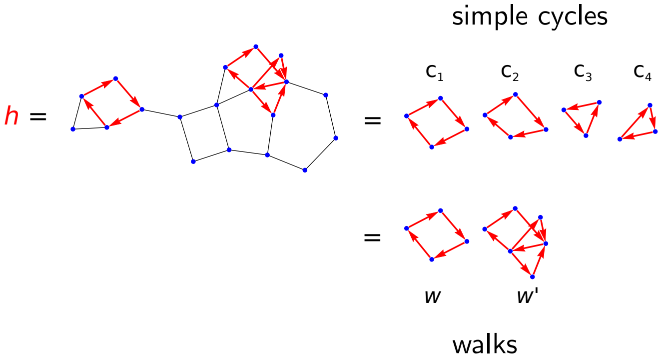


Same weirdness in number theory

$\log \zeta(s)$ non-zero only if $n = p^k$

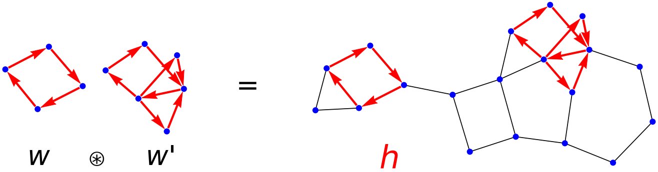
Hikes from walks

Generating hikes from:



Hikes from walks

Multiplication: $w \circledast w' = h$ iff w and w' vertex disjoint

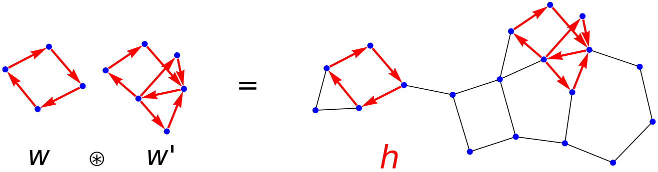


$\hookrightarrow$ Series on hikes: subgraph convolution $*_G$

$$\sum_{h,h'} s_1(h)s_2(h')\, h \circledast h' = \sum_{H \prec G} S_1|_H \; S_2|_{G-H} = S_1 *_G S_2$$

Hikes from walks

Multiplication: $w \circledast w' = h$ iff w and w' vertex disjoint



$\hookrightarrow$ Series on hikes: subgraph convolution $*_G$

$$\sum_{h,h'} s_1(h)s_2(h')\, h \circledast h' = \sum_{H \prec G} S_1|_H \; S_2|_{G-H} = S_1 *_G S_2$$

Theorem

With $\circledast$, hikes can be made into a Hopf algebra

Primes gets the structure of a free Lie algebra .... so what?

Algebraic Walk Theory

1. Why?
   - Network analysis
2. How?
   - From walks to hikes
   - Posets and trace monoids
   - Hopf and Lie algebras
3. So what?
   - Getting the primes
   - Graph characterisation
   - Cycle Double Cover Conjecture
4. Conclusion

Lie magic

Theorem

Any Lie idempotent $\imath$ projects:

$$\text{i) Hikes} \xrightarrow{\imath} \text{walks}$$

$$\text{ii) Self-avoiding hikes} \xrightarrow{\imath} \text{primes}$$

<u>Concretely:</u> a PLETHORA of formulas!!

$$P(z) = \sum_{Primes} z^{\ell(p)} p$$

Lie magic

Theorem

Any Lie idempotent $\boldsymbol{\imath}$ projects:

$$i)\ \text{Hikes} \xrightarrow{\boldsymbol{\imath}} \text{walks}$$

$$ii)\ \text{Self-avoiding hikes} \xrightarrow{\boldsymbol{\imath}} \text{primes}$$

Concretely: a PLETHORA of formulas!!

$$P(z) = \sum_{Primes} z^{\ell(p)} p$$

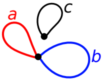
Problem is NP-hard $\longrightarrow$ cost at least 2^N

Culprit: $*_G$

Getting the primes

- Eulerian idempotent

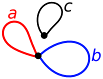
$$P(z) = -\log_{*_G} \det(\mathsf{I} - z\mathsf{A}_G)$$

Let's try it! $\mathsf{G} =$ 

Getting the primes

- Eulerian idempotent

$$P(z) = -\log_{*_G} \det(\mathsf{I} - z\mathsf{A}_G)$$

Let's try it! G = 

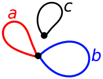
$\longrightarrow \det = 1 - a - b - c + ac + bc$

$-\log_{*_G} \det(\mathsf{I} - \mathsf{A}_G) =$

Getting the primes

- Eulerian idempotent

$$P(z) = -\log_{*_G} \det(\mathrm{I} - z\mathrm{A}_G)$$

Let's try it! $\quad \mathrm{G} =$ 

$\longrightarrow \det = 1 - a - b - c + ac + bc$

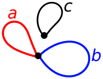
$$-\log_{*_G} \det(\mathrm{I} - \mathrm{A}_G) =$$

$$a + b + c - ac - bc - \frac{1}{2}\underbrace{(a+b+c-ac-bc) *_G (a+b+c-ac-bc)} + \ldots$$

Getting the primes

- Eulerian idempotent

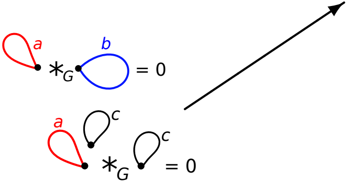
$$P(z) = -\log_{*_G} \det(\mathsf{I} - z\mathsf{A}_G)$$

Let's try it! G = 

$\longrightarrow \det = 1 - a - b - c + ac + bc$

$$-\log_{*_G} \det(\mathsf{I} - \mathsf{A}_G) =$$

$$a + b + c - ac - bc - \frac{1}{2}\underbrace{(a+b+c-ac-bc) *_G (a+b+c-ac-bc)}_{\underbrace{ac + ca + bc + cb}_{2ac + 2bc}} + \dots$$

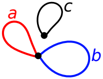


etc.

Getting the primes

- Eulerian idempotent

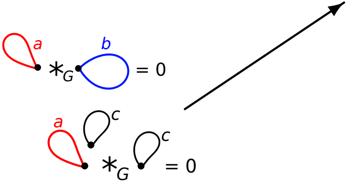
$$P(z) = -\log_{*_G} \det(\mathrm{I} - z\mathrm{A}_G)$$

Let's try it! G = 

$\longrightarrow \det = 1 - a - b - c + ac + bc$

$$-\log_{*_G} \det(\mathrm{I} - \mathrm{A}_G) =$$

$$a + b + c - ac - bc - \frac{1}{2}\underbrace{(a+b+c-ac-bc) *_G (a+b+c-ac-bc)}_{\underbrace{ac + ca + bc + cb}_{2ac + 2bc}} + \ldots$$



etc.

$$P = -\log_{*_G} \det(\mathrm{I} - \mathrm{A}_G) = a + b + c$$

Getting the primes

$$P(z) = \sum_{H \prec G} \int dz \, \det(z\mathsf{A}_H) \, \frac{d}{dz} \mathrm{perm}(\mathsf{I} + z\mathsf{A}_{G-H}) \qquad (\star)$$

$$P(z) = \sum_{H \prec G} \int dz \, \frac{1}{z} \mathrm{Tr}\Big((z\mathsf{A}_H)^{|H|} \, (\mathsf{I} - z\mathsf{A}_H)^{N-|H|} \Big) \qquad (\star\star)$$

... *and many more*

Full P: exponential cost in N, but still the best

Getting the primes

Always valid in number theory

- Euler idempotent (here $n *_G m = 0$ if (n, m) not coprime)

$$\log_{*_G} \frac{\zeta_R(s)}{\zeta_R(2s)} = \sum_{p \textbf{ prime}} p^{-s}$$

- $(\star)$ formula

$$\int ds \sum_n \frac{|\mu(n)|}{n^s} \sum_{m|n} \mu(m) \log(m) |\mu(n/m)| = \sum_{p \textbf{ prime}} p^{-s}$$

- $(\star\star)$ formula

$$\sum_n \frac{|\mu(n)|}{n^s} \frac{\Lambda(n)}{\log(n)} = \sum_{p \textbf{ prime}} p^{-s}$$

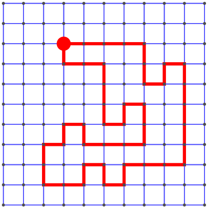
Short primes – long primes

- Short primes: check balance on large networks
  $\longrightarrow$ Primes length ℓ cost $O(N^\ell) \rightarrow$ *network balance*
  $\longrightarrow$ Amenable to Monte-Carlo

Short primes – long primes

- Short primes: check balance on large networks
  $\hookrightarrow$ Primes length ℓ cost $O(N^\ell) \rightarrow$ *network balance*
  $\hookrightarrow$ Amenable to Monte-Carlo

- Long primes: how many self-avoiding polygons of length $\ell \gg 1$? numerics $\mu^\ell \, \ell^{11/32}$



Open since ~70 years

- μ related to the max. eigenvalues of subgraphs
  $\hookrightarrow$ *And more!* *(functional equation)* *In progress...*

Algebraic Walk Theory

Graph characterisation

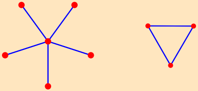
Proposition

Poset P_G + prime length = G

Perfect characterisation of undirected graphs

In progress

Poset $P_G = G$, with one exception



$\longrightarrow$ Full, semi-commutative, ζ perfect graph invariant

What if we use a simpler commutative version of ζ?

(Abelianization)

Ihara zeta function

The Abelianization of ζ is the Ihara zeta function ζ_I

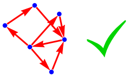
$$\zeta \xrightarrow[\text{commute}]{\text{Everything}} \zeta_{\text{Ab}} = \zeta_I \zeta_b$$

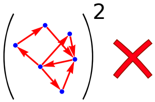
Ihara what?

$\hookrightarrow$ J.P. Serre: similarities with number theory

$$\zeta_I = \prod_{p_o} \left(1 - z^{\ell(p_o)}\right)^{-1} \text{ looks like } \zeta_R(s) = \prod_p \left(1 - p^{-s}\right)^{-1}$$

$p_o \neq w^k$



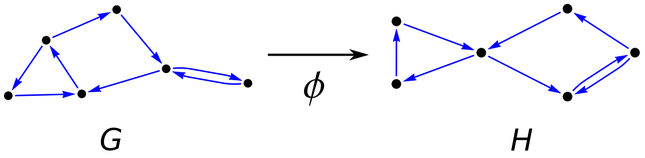


A. Terras, *Zeta Functions of Graphs*, Cambridge studies in adv. math., 2011

Directed graphs

There exists graph transformations with $P_G = \phi(P_G)$

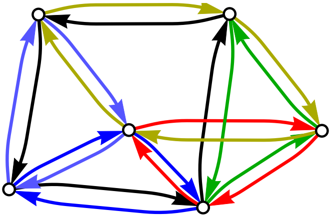
Hikes do not characterize digraphs, even strongly connected ones



- G, H cospectral
- Same traces $\mathrm{Tr}\big(f(G)\big) = \mathrm{Tr}\big(f(H)\big)$
- Same Weisfeiler-Lehman colors
- All immanents identical
- ...

Algebraic Walk Theory

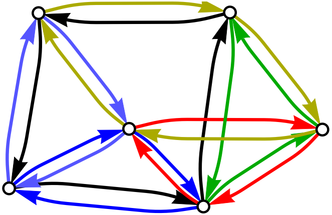
Cycle Double Cover Conjecture



Strong CDC

"Bridgeless graphs admit a cycle cover with all oriented edges covered exactly once"

Cycle Double Cover Conjecture



Strong CDC

"Bridgeless graphs admit a cycle cover with all oriented edges covered exactly once"

- Square-free hikes h_c, no backtrack factors, length $2|E|$
- Coeff of h_c in ζ non-zero

$\hookrightarrow$ Purely algebraic question

Conclusion

Conclusion

Main Message

Hikes obey an extremely rich algebraic structure

Results

- Many exact & approx. formulas for counting simple cycles
- Perfect cycle-based graph-invariants
- **HUGE** reservoir of algebraic machinery

Open problems

- Additive properties of hikes (CDC)
- Long prime approximations (in progress)

Thank You!

Collaborators

Prof. P. Rochet (University of Nantes)

Prof. T. Espinasse (Claude Bernard University Lyon 1)

Prof. R. Wilson (University of York)

Funding

Royal commission for the exhibition of 1851

"Algebraic Combinatorics on Trace Monoids: Extending Number Theory to Walks on Graphs", Giscard, Rochet, arXiv:1601.01780 (under review)