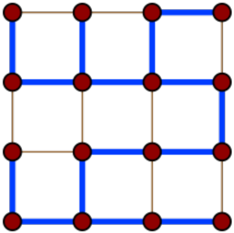
Spanning Trees and Continued Fractions

Swee Hong Chan

joint with Alex Kontorovich and Igor Pak



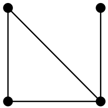
$$\cfrac{4}{1+\cfrac{1}{2+\cfrac{9}{2+\ddots}}}$$

What is a spanning tree?

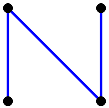
Let $G = (V, E)$ be a simple graph.

A spanning tree is a subset of edges of G that

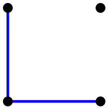
- includes all vertices (spanning),
- has no cycles (tree).



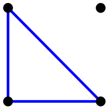
G



spanning tree

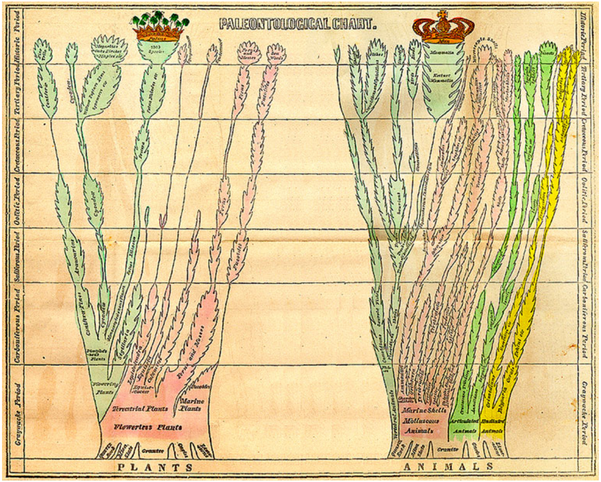


not spanning



not tree

Examples: Phylogenetic tree



From *Elementary Geology* (1840),

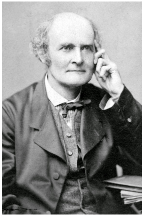
by Edward Hitchcock

How many spanning trees can a graph have?

Cayley's formula

Theorem (Borchardt 1860, Cayley 1889)
The number of spanning trees of a complete graph with n vertices is n^{n-2}.

Carl W. Borchardt

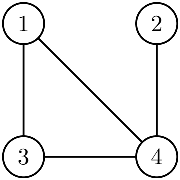


Arthur Cayley

Matrix tree theorem

Theorem (Kirchhoff 1847)

The number of spanning trees $t(G)$ of G is equal to the determinant of a minor of its Laplacian matrix.



G

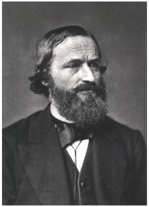
$$\begin{bmatrix} 2 & 0 & -1 & -1 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}$$

Laplacian

Matrix tree theorem

Theorem (Kirchhoff 1847)

The number of spanning trees $t(G)$ of G is equal to the determinant of a minor of its Laplacian matrix.



Gustav Kirchhoff

Note: Kirchhoff's paper has neither matrices nor trees.

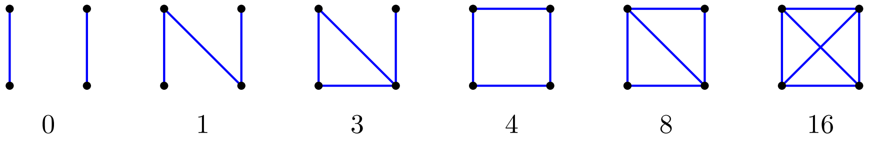
Sedláček's Problems

Set of spanning tree numbers

For $n \geq 1$, let

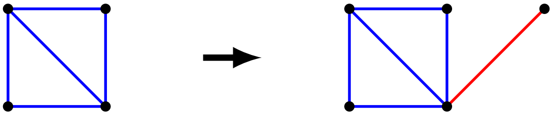
$$\mathcal{G}_n := \{\text{all simple graphs with } n \text{ vertices}\},$$

$$t(\mathcal{G}_n) := \left\{\begin{array}{l}\text{number of spanning trees of all} \\ \text{simple graphs with } n \text{ vertices}\end{array}\right\}.$$

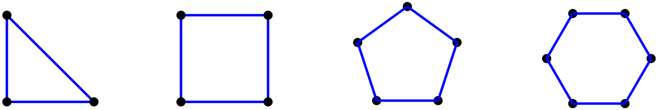


Properties of $t(\mathcal{G}_n)$

- We have $t(\mathcal{G}_n) \subseteq t(\mathcal{G}_{n+1})$.



- We have $\bigcup_{n \geq 1} t(\mathcal{G}_n) = \{0, 1, \cancel{2}, 3, 4, 5, \dots\}$.



Sedláček's First Problem

Problem (Sedláček 1966)

Describe the set of spanning tree numbers $t(\mathcal{G}_n)$.

O KOSTRÁCH KONEČNÝCH GRAFŮ

JIŘÍ SEDLÁČEK, Praha

(Došlo dne 28. srpna 1965)

ON THE SPANNING TREES OF FINITE GRAPHS

JIŘÍ SEDLÁČEK, Praha

Sedláček's First Problem

Theorem (Sedláček 1966)

For $n \geq 3$,

$$n^2 \;\leq\; |t(\mathcal{G}_n)| \;\leq\; n^{n-2}.$$

It is clear that the lower bound is not tight.

Conjecture

$$|t(\mathcal{G}_n)| \;\geq\; C^n \quad \textit{for some} \;\; C > 1.$$

Motivation: Inverse counting problem

Input: Integer $T \geq 3$.

Problem: Construct graph G with $t(G) = T$ and

$$|V(G)| \leq c \log T \quad \text{for some } c > 0.$$

Understanding the problem was needed for studying the computational complexity of counting number of bases of matroids.

Motivation: Inverse counting problem

Input: Integer $T \geq 3$.

Problem: Construct graph G with $t(G) = T$ and

$$|V(G)| \leq c \log T \quad \text{for some} \quad c > 0.$$

Solution to the above problem would imply,

$$|t(\mathcal{G}_n)| \geq (e^{1/c})^n,$$

and would solve Sedláček's First Problem.

What was known

First super-polynomial lower bound was due to Azarija (2014):

$$|t(\mathcal{G}_n)| \ \geq \ e^{\Omega(\sqrt{n/\log n})}.$$

Best lower bound prior to our work was due to Stong (2022):

$$|t(\mathcal{G}_n)| \ \geq \ e^{\Omega(n^{2/3})}.$$

Sedláček's Second Problem

For $n \geq 1$, let

$$\mathcal{P}_n := \{\text{all simple planar graphs with } n \text{ vertices}\}.$$

Problem (Sedláček 1966)

Describe the set of spanning tree numbers of planar graphs $t(\mathcal{P}_n)$.

Note: Four-color theorem was proved in 1976.

What was known

It follows from Euler's formula that

$$|t(\mathcal{P}_n)| \;\le\; 2^{|E|} \;\le\; 8^n.$$

The best bounds prior to our work were:

$$e^{\Omega(n^{2/3})} \;\le\; |t(\mathcal{P}_n)| \;\le\; (5.2852)^n.$$

Upper bound was due to Buchin-Schulz (2010),
lower bound was due to Stong (2022).

First main result

Theorem 1 (C.–Kontorovich–Pak 2024+)

For sufficiently large n,

$$|t(\mathcal{P}_n)| \;\geq\; (1.1103)^n.$$

Note that this implies

$$|t(\mathcal{G}_n)| \;\geq\; |t(\mathcal{P}_n)| \;\geq\; (1.1103)^n.$$

This is the first exponential lower bound for Sedláček's First Problem, and a tight lower bound for Second Problem.

Recall: Inverse counting problem

Input: Integer $T \geq 3$.

Problem: Construct graph G with $t(G) = T$ and

$$|V(G)| \leq c \log T \quad \text{for some} \quad c > 0.$$

Can we also solve this problem?

Recall: Inverse counting problem

Input: Integer $T \geq 3$.

Problem: Construct graph G with $t(G) = T$ and

$$|V(G)| \leq c \log T \quad \text{for some } c > 0.$$

Can we also solve this problem?

Yes we (mostly) **can!**

Almost all integers

A set $S \subseteq \mathbb{N}$ contains almost all integers if

$$\lim_{N \to \infty} \frac{|S \cap \{1, \ldots, N\}|}{N} = 1.$$

This is a weaker notion than requiring S to contain all but finitely many integers.

Second main result

Theorem 2 (C.–Kontorovich–Pak 2024+)

For almost all integers T, there exists a planar graph G with $t(G) = T$ and

$$|V(G)| \leq 56 \log_\varphi T,$$

where $\varphi := \frac{1+\sqrt{5}}{2} = 1.618$ is the golden ratio.

This implies the exponential lower bound

$$|t(\mathcal{P}_n)| \geq (\varphi^{1/56})^n \approx (1.008)^n,$$

which can be further improved to $(1.1103)^n$.

Connections to continued fractions

Continued fractions

For integers $a_1, \ldots, a_k \geq 1$,

$$[a_1, \ldots, a_k] := \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + a_k}}}.$$

Every rational number $\frac{t}{u} \leq 1$ can be written as a finite continued fraction using Euclidean algorithm.

Furthermore, we have $k \leq \log_\varphi u$.

Connect spanning trees to continued fractions

Input: $b_1, \ldots, b_\ell \geq 1$.

Output: Planar graph G and edge e with

$$\frac{t(G-e)}{t(G/e)} = \left[b_1, 1, b_2, 1, \ldots, b_\ell, 1\right],$$

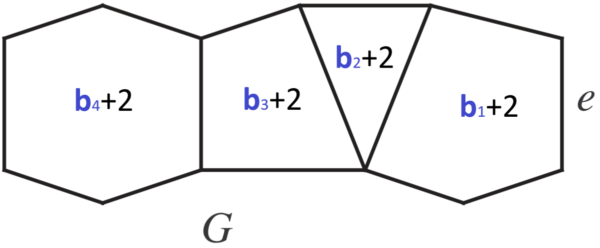
$t(G-e)$ and $t(G/e)$ are coprime,

$|V(G)| = b_1 + \cdots + b_\ell + 2.$

Here $G - e$ is graph deletion,

and G/e is graph contraction.

The silkworm graph



The i-th cycle has $b_i + 2$ vertices.

The example above has

$$[3, 1, 1, 1, 2, 1, 4, 1] = \frac{63}{229},$$

$$t(G - e) = 63, \qquad t(G/e) = 229.$$

Zaremba's conjecture

Zaremba's conjecture

Conjecture (Zaremba 1972)

For every integer u, there exists coprime $t < u$ with

$$\frac{t}{u} = [a_1, \ldots, a_k],$$

$$a_1, \ldots, a_k \leq 5.$$

Conjecture is false if 5 is replaced with 4,

with $u = 54$.

Note $a_1 + \ldots + a_k \leq 5 \log_\varphi u$

by Euclidean's algorithm.

Bourgain–Kontorovich theorem

Theorem (Bourgain–Kontorovich 2014)

For almost all integers u, that there exists coprime $t < u$ with

$$\frac{t}{u} = [a_1, \ldots, a_k],$$

$$a_1, \ldots, a_k \leq 50.$$

Huang (2015) has since improved the bound from 50 to 5.

Bourgain–Kontorovich theorem

Theorem (Bourgain–Kontorovich 2014)

For almost all integers u, that there exists coprime $t < u$ with

$$\frac{t}{u} = [a_1, \ldots, a_k],$$

$$a_1, \ldots, a_k \leq 50.$$

Huang (2015) has since improved the bound from 50 to 5.

This is almost what we need for Sedláček's Problem.

Alternating BK theorem

Theorem (C.–Kontorovich–Pak 2024+)

For almost all integers t, there exists coprime $u > t$ with

$$\frac{t}{u} = [b_1, 1, b_2, 1, \ldots, b_\ell, 1],$$

$$b_1, \ldots, b_\ell \le 110.$$

This is exactly what we need!

Back to inverse counting problem

Input: Integer $T \geq 3$.

Goal: Construct graph G with $t(G) = T$ and

$$|V(G)| \ \leq \ 56 \log_{\varphi} T.$$

We now give a construction that is guaranteed to work 99% of the time.

Solution to inverse counting problem

For $u \in \{T, \ldots, 110T\}$ coprime to T:

- Compute continued fraction $\frac{T}{u} = [a_1, \ldots, a_k]$.
- If $(a_1, \ldots, a_k) = (b_1, 1, \ldots, b_{k/2}, 1)$, $b_i \leq 110$:

  construct silkworm $(b_1, \ldots, b_{k/2})$ graph (G, e).
  Note that

$$t(G - e) = T,$$
$$|V(G)| = b_1 + \ldots + b_{k/2} + 2 \leq 56 \log_\varphi T.$$

Output: graph $G - e$.

Solution to inverse counting problem

Input: Integer $T \geq 3$.

Goal: Construct graph G with $t(G) = T$ and

$$|V(G)| \ \leq \ 56 \log_{\varphi} T.$$

Alternating BK theorem thus guarantees our construction works 99% of the time...

Solution to inverse counting problem

Input: Integer $T \geq 3$.

Goal: Construct graph G with $t(G) = T$ and

$$|V(G)| \ \leq \ 56 \log_{\varphi} T.$$

Alternating BK theorem thus guarantees our construction works 99% of the time...
and might still work for the other 1%.

Alternating Zaremba's conjecture

Conjecture

There exists an absolute constant $A > 0$, such that for every integer t, there exists coprime $u > t$ with

$$\frac{t}{u} = [b_1, 1, b_2, 1, \dots, b_\ell, 1],$$

$$b_1, \dots, b_\ell \leq A.$$

If this conjecture is true,

then our construction will always work.

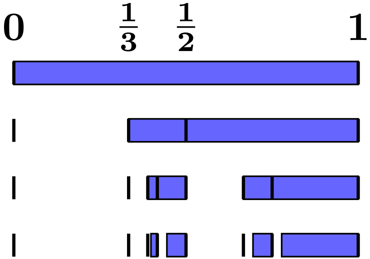
Sketch of proof of
original and alternating BK theorem

Cantor-like fractals

For $A \geq 2$,

$$\mathfrak{C}_A := \left\{ \left[a_1, a_2, \ldots \right] \mid a_i \leq A \right\},$$

limit set of rational numbers in Zaremba's conjecture.



We would like to measure this set.

Hausdorff dimension

For $S \subseteq \mathbb{R}$ and $d \in \mathbb{R}_+$, Hausdorff measure $H^d(S)$ is

$$\lim_{\delta \to 0} \inf \left\{ \sum_{i=1}^{\infty} (b_i - a_i)^d \ : \ \bigcup_{i=1}^{\infty} (a_i, b_i) \ \supseteq \ S, \ b_i - a_i < \delta \right\}.$$

The Hausdorff dimension of S is

$$\mathrm{Hdim}(S) \ := \ \inf \left\{ d \geq 0 \ : \ H^d(S) = 0 \right\}.$$

Note that, for Cantor-like fractals,

$$0 \ < \ \mathrm{Hdim}(\mathfrak{C}_A) \ < \ 1, \qquad \mathrm{Hdim}(\mathfrak{C}_A) \nearrow 1 \ \text{ as } \ A \to \infty.$$

Black box: Orbital circle method

Theorem (Bourgain–Kontorovich 2014)

Let $A \geq 2$. Then, for almost all integers u, there exists coprime $t < u$ with

$$\frac{t}{u} = [a_1, \ldots, a_k], \qquad a_1, \ldots, a_k \leq A$$

if

$$\mathrm{Hdim}(\mathfrak{C}_A) > 0.984.$$

This reduces density one version of Zaremba's conjecture to computing Hausdorff dimension.

Original BK theorem

Bourgain–Kontorovich (2014) computed that

$$\mathrm{Hdim}(\mathfrak{C}_{50}) = 0.986\ldots > 0.984.$$

Theorem (Bourgain–Kontorovich 2014)

For almost all integers u, there exists coprime $t < u$ with

$$\frac{t}{u} = [a_1, \ldots, a_k], \qquad a_1, \ldots, a_k \leq 50.$$

Improvements have since been made on orbital circle method and computing Hausdorff dimensions.

Improvements to orbital circle method

Frolenkov–Kan (2014) improves to

$$\mathrm{Hdim}(\mathfrak{C}_A) > 0.83 \implies (\text{positive proportion}).$$

Huang (2015) improves to

$$\mathrm{Hdim}(\mathfrak{C}_A) > 0.83 \implies (\text{density one}).$$

Kan (2015, 2017, 2021) improves to

$$\mathrm{Hdim}(\mathfrak{C}_A) > 0.7749 \implies (\text{positive proportion}) \implies (\text{density one}).$$

Improvements to computing Hausdorff dimension

The state of the art algorithm to compute Hausdorff dimension is due to Pollicott–Vytnova (2022):

$$\mathrm{Hdim}(\mathfrak{C}_5) = 0.836829443680\ldots$$

Recall the result of Huang (2015):

$$\mathrm{Hdim}(\mathfrak{C}_A) > 0.83 \implies (\text{density one}),$$

which gives the current best result for Zaremba's conjecture.

Back to alternating BK theorem

For $A \geq 2$, the Hančl–Turek fractal is

$$\mathrm{Hdim}(\mathfrak{D}_A) \ := \ \{[b_1, 1, b_2, 1, \dots] \mid b_i \leq A\}.$$

Based on Kan (2021), we need to find A satisfying

$$\mathrm{Hdim}(\mathfrak{D}_A) \ > \ 0.7749.$$

If such A exists, we get everything.

If such A does not exist, we get nothing.

Luck is on our side

Assisted by Pollicott–Vytnova and computers,

$$\mathrm{Hdim}(\mathfrak{D}_{110}) = 0.7750\ldots > 0.7749.$$

Theorem (C.–Kontorovich–Pak 2024+)

For almost all integers t, there exists coprime $u > t$ with

$$\frac{t}{u} = [b_1, 1, \ldots, b_\ell, 1], \quad b_1, \ldots, b_\ell \leq 110.$$

Pollicott (2025+) has since shown that

$$\mathrm{Hdim}(\mathfrak{D}_{109}) = 0.774902739\ldots > 0.7749.$$

Open problem

Improvement for Sedláček's First Problem

Conjecture

There exists $c > 0$ so that

$$|t(\mathcal{G}_n)| \;\geq\; 2^{c n \log n}.$$

Contrast this with the trivial upper bound

$$|t(\mathcal{G}_n)| \;\leq\; n^{n-2} \;\leq\; e^{n \log n}.$$

Solving this problem would most likely require new graph constructions.

Improvement for Sedláček's Second Problem

Alon–Bucić–Gishboliner (2025+) recently improved our lower bound from $(1.1103)^n$ to

$$|t(\mathcal{P}_n)| \;\geq\; (1.49)^n.$$

Problem

Does there exist $C > 1$ so that

$$\lim_{n\to\infty} \frac{1}{n} \log |t(\mathcal{P}_n)| \;=\; C.$$

If C exists, then it must satisfy

$$1.49 \quad < \quad C \quad < \quad 5.2852.$$

THANK YOU!

Preprint: www.arxiv.org/abs/2411.18782

Webpage: www.math.rutgers.edu/~sc2518/

Email: sweehong.chan@rutgers.edu