

Approaching Hofstadter's Q -sequence through a stable variant

Benoît Cloitre

Experimental Mathematics Seminar

Rutgers University

8 October 2026

Plan

I. Proof of well-definedness and asymptotic behavior of $\tilde{Q}$

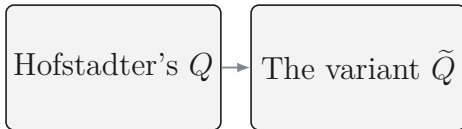
Plan

I. Proof of well-definedness and asymptotic behavior of $\tilde{Q}$

Hofstadter's Q

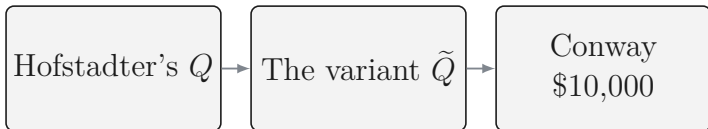
Plan

I. Proof of well-definedness and asymptotic behavior of $\tilde{Q}$



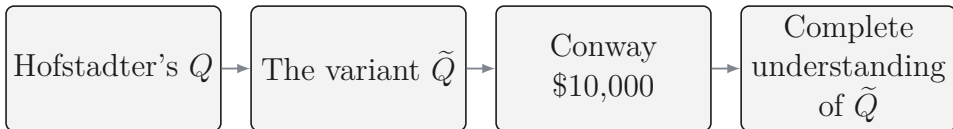
Plan

I. Proof of well-definedness and asymptotic behavior of $\tilde{Q}$



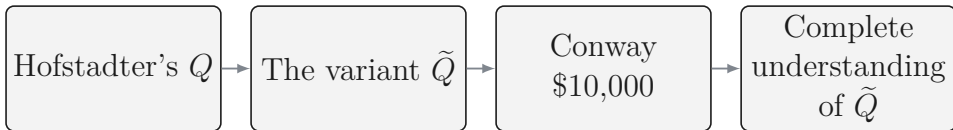
Plan

I. Proof of well-definedness and asymptotic behavior of $\tilde{Q}$

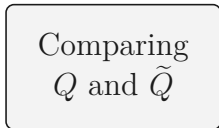


Plan

I. Proof of well-definedness and asymptotic behavior of $\tilde{Q}$

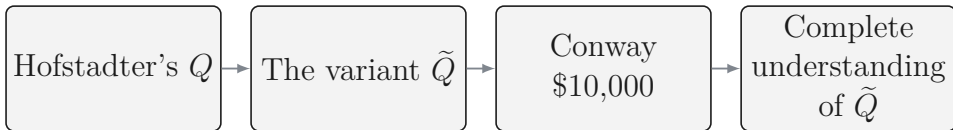


II. Back to Q

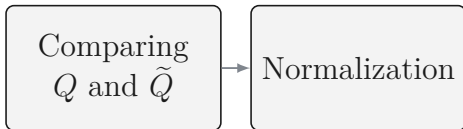


Plan

I. Proof of well-definedness and asymptotic behavior of $\tilde{Q}$

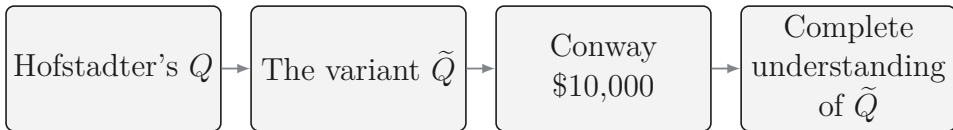


II. Back to Q



Plan

I. Proof of well-definedness and asymptotic behavior of $\tilde{Q}$

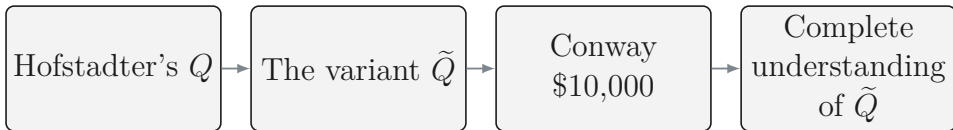


II. Back to Q

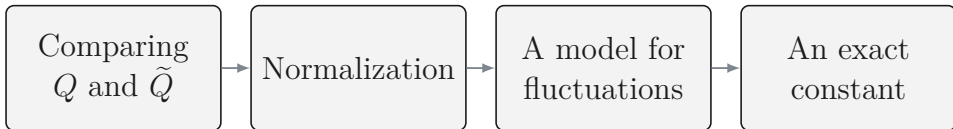


Plan

I. Proof of well-definedness and asymptotic behavior of $\tilde{Q}$



II. Back to Q



Hofstadter's Q -sequence

$$Q(n) = Q(n - Q(n - 1)) + Q(n - Q(n - 2))$$

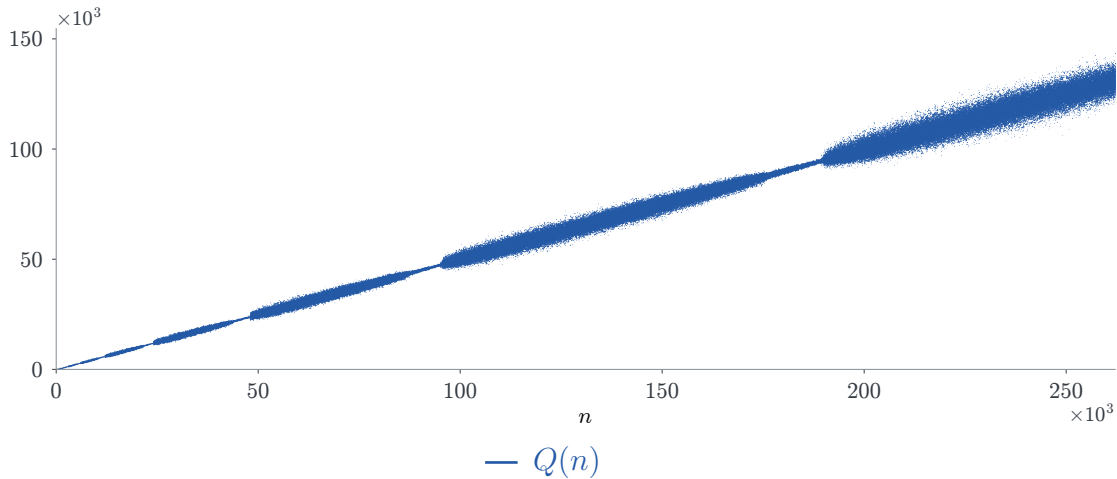
$$Q(1) = Q(2) = 1$$

Is $Q(n)$ defined for every n ?

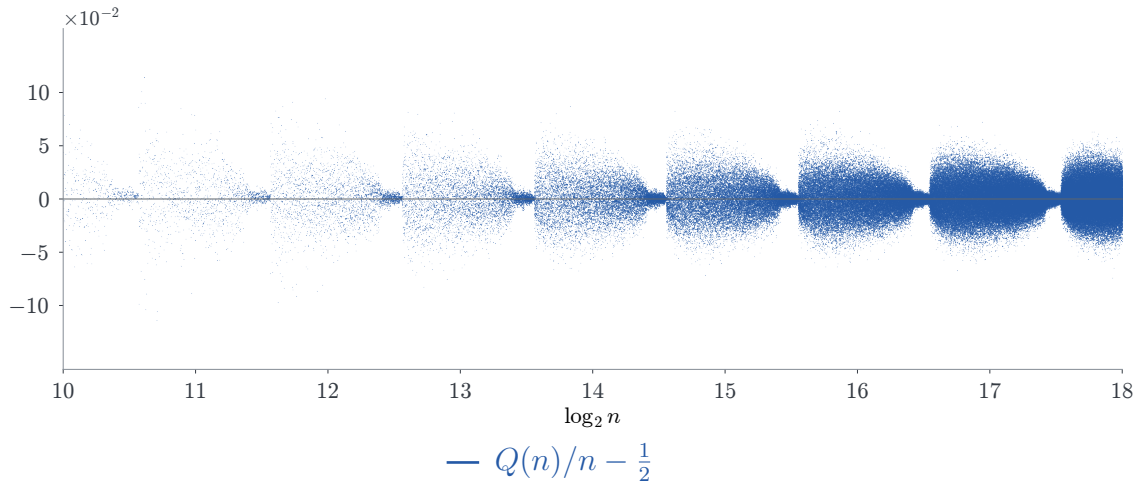
If so, does $Q(n)/n$ converge to $1/2$?

D. R. Hofstadter, *Gödel, Escher, Bach* (1979). OEIS A005185.

Hofstadter's Q -sequence



Hofstadter's Q -sequence



Mantovanelli's perturbation

$$\tilde{Q}(n) = \tilde{Q}(n - \tilde{Q}(n - 1)) + \tilde{Q}(n - \tilde{Q}(n - 2)) + (-1)^n$$

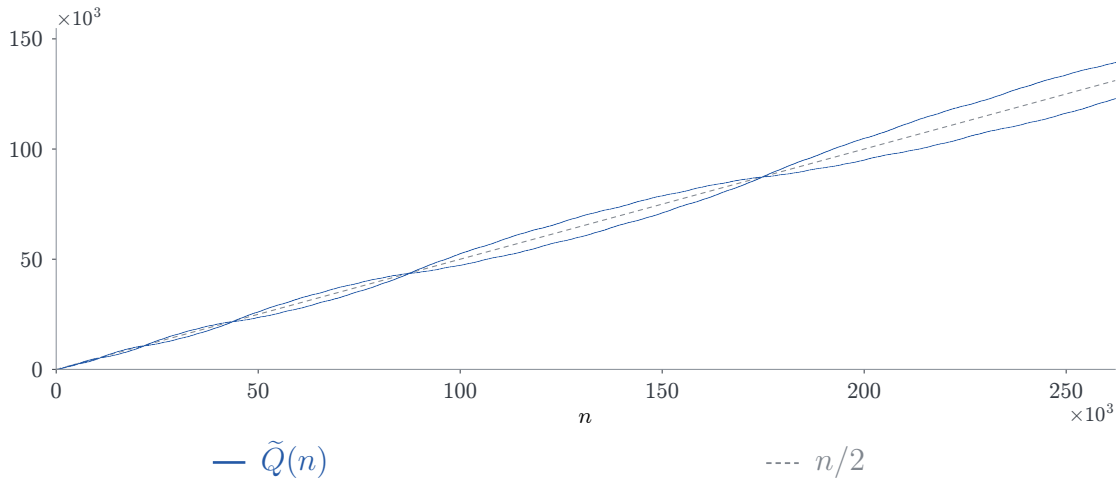
$$\tilde{Q}(1) = \tilde{Q}(2) = 1$$

+1 at even indices

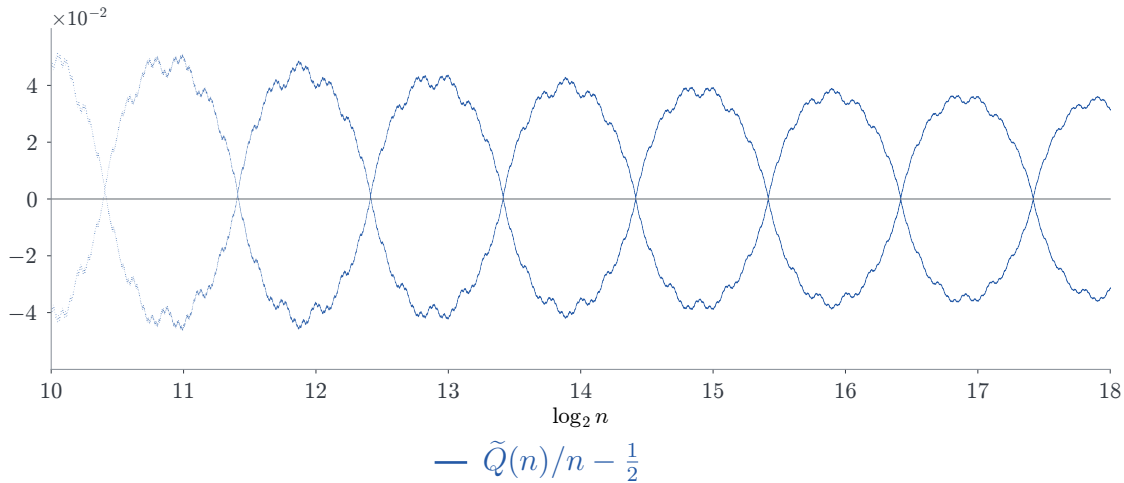
−1 at odd indices

M. Mantovanelli, arXiv:2603.16111 (2026). OEIS A394051.

The perturbed sequence $\tilde{Q}$



The perturbed sequence $\tilde{Q}$



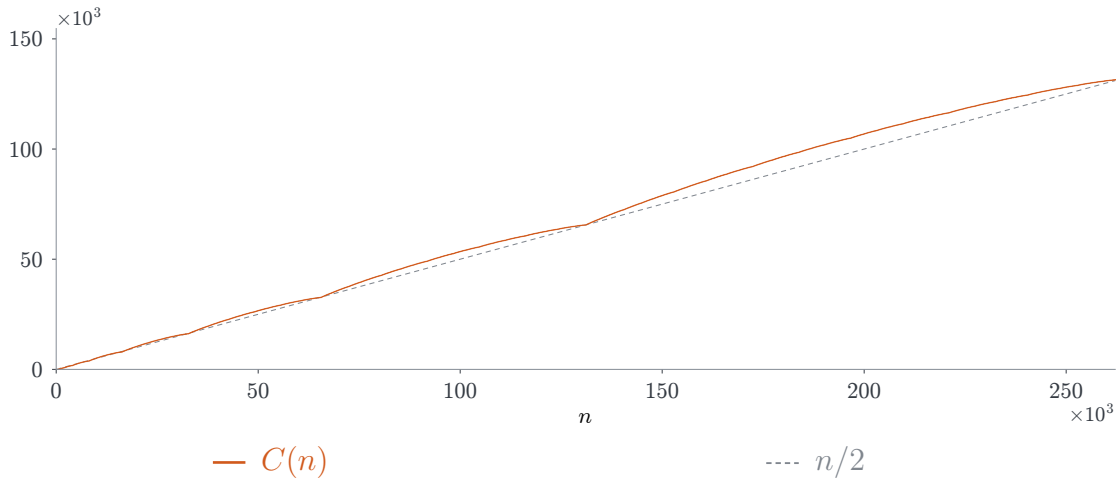
Conway's sequence

$$C(n) = C(C(n-1)) + C(n - C(n-1))$$

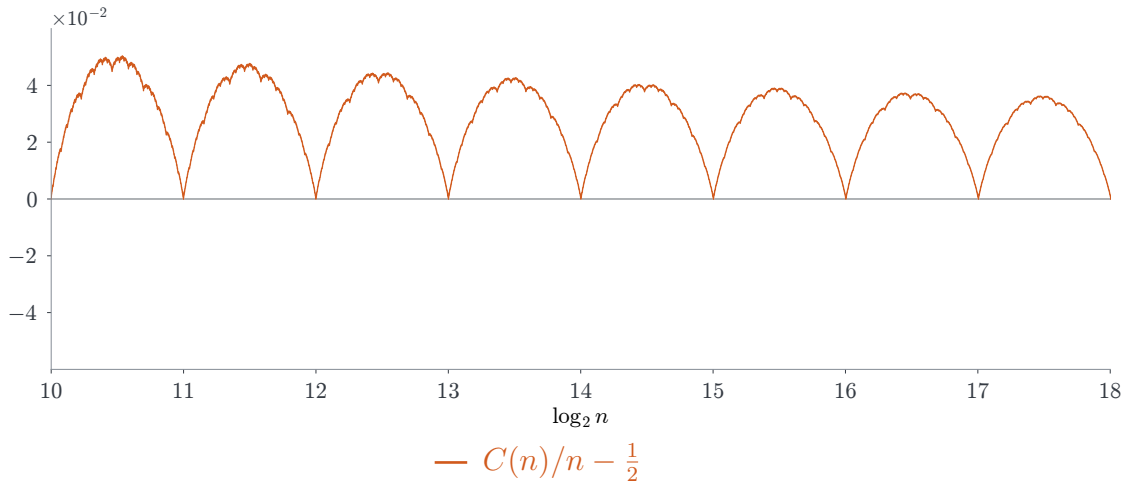
$$C(1) = C(2) = 1$$

OEIS A004001. C.L. Mallows (1991). T. Kubo, R. Vakil (1996).

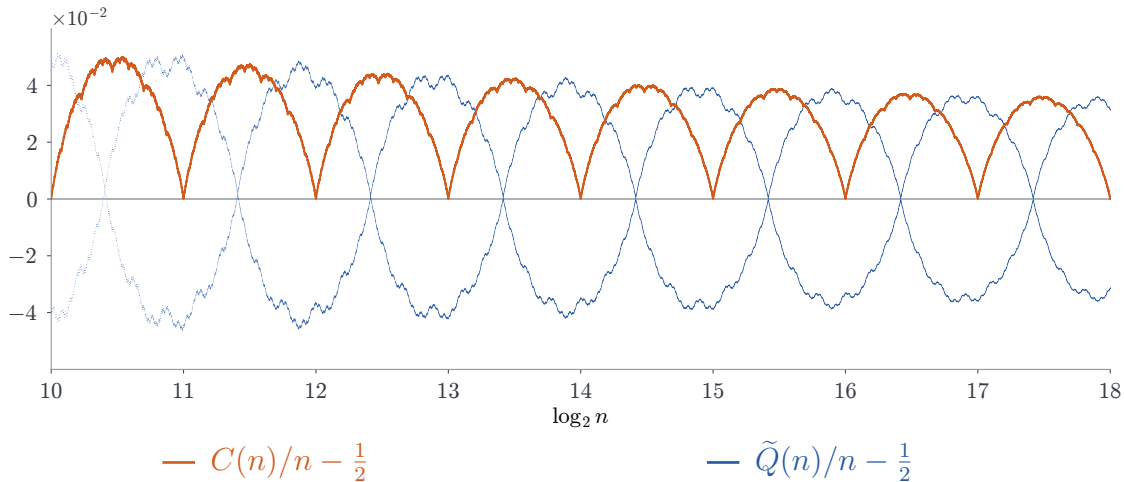
Conway's sequence



Conway's sequence



Conway's sequence and $\tilde{Q}$



How to understand $\tilde{Q}$?

Conway's sequence

Mantovanelli's sequence

How to understand $\tilde{Q}$?

Conway's sequence

Mantovanelli's sequence

Kubo and Vakil (1996)

$n \longmapsto S(n)$: a finite set of integers.

How to understand $\tilde{Q}$?

Conway's sequence

Mantovanelli's sequence

Kubo and Vakil (1996)

$n \mapsto S(n)$: a finite set of integers.

For sets of size ≥ 2 :

drop the smallest element,
subtract 1 from the others.

$$S(C(n)) = \text{compression}(S(n))$$

How to understand $\tilde{Q}$?

Conway's sequence

Kubo and Vakil (1996)

$n \mapsto S(n)$: a finite set of integers.

For sets of size ≥ 2 :

drop the smallest element,
subtract 1 from the others.

$$S(C(n)) = \text{compression}(S(n))$$

Mantovanelli's sequence

Two crossing curves.

Reduce the study to one curve.

How to understand $\tilde{Q}$?

Conway's sequence

Kubo and Vakil (1996)

$n \mapsto S(n)$: a finite set of integers.

For sets of size ≥ 2 :

drop the smallest element,
subtract 1 from the others.

$$S(C(n)) = \text{compression}(S(n))$$

Mantovanelli's sequence

Two crossing curves.

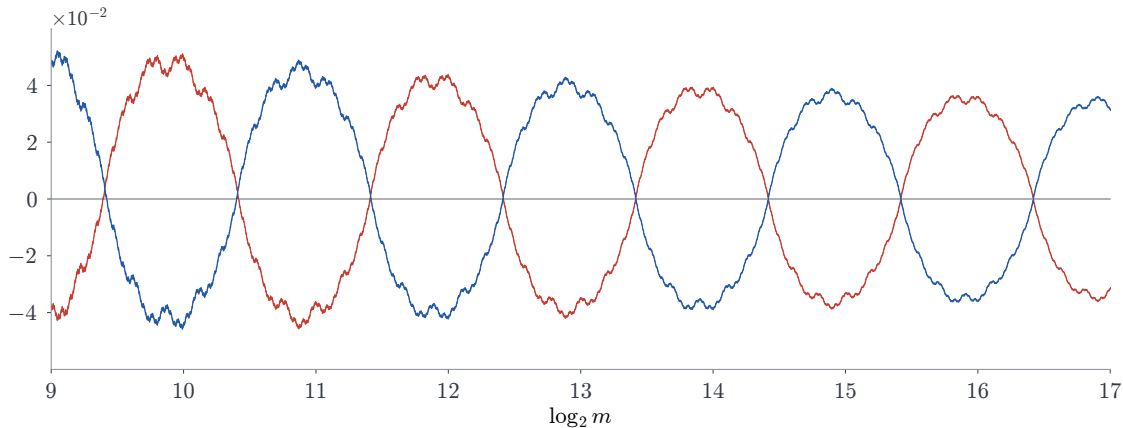
Reduce the study to one curve.

Arches as finite binary words.

Words and paths lead to
binomial coefficients and
asymptotics.

Compression also appears on record heights.

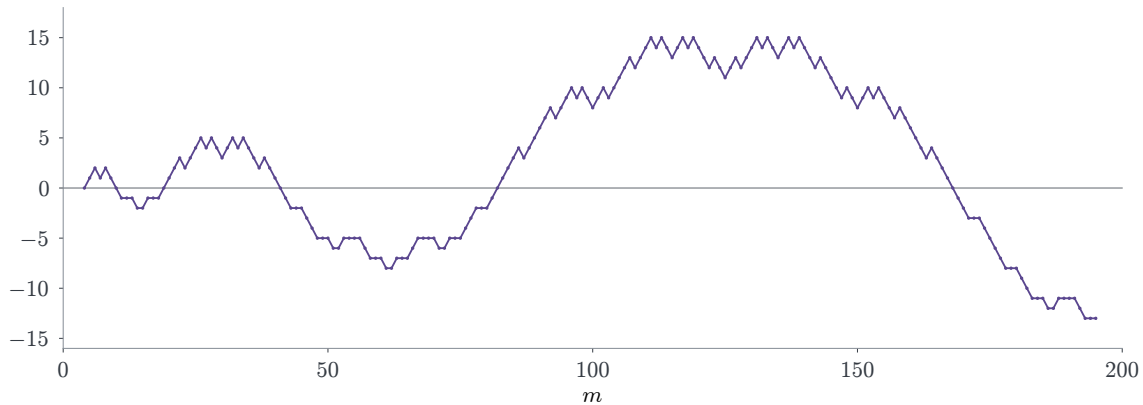
Even and odd indices



— $\mathfrak{a}(m)/m - \frac{1}{2}$, odd
 $\mathfrak{a}(m) = \frac{1}{2}(\tilde{Q}(2m-1) + 1)$

— $\mathfrak{b}(m)/m - \frac{1}{2}$, even
 $\mathfrak{b}(m) = \frac{1}{2}(\tilde{Q}(2m) + 1)$

The difference $\delta = \mathfrak{b} - \mathfrak{a}$



First observations

Positive arch r	0	1	2	3	4
Starting index u_r	4	19	82	337	1360
Maximum V_r	2	5	15	50	176
Increase $V_r - V_{r-1}$	—	3	10	35	126
Next minimum	-2	-8	-28	-98	-350

First observations

Positive arch r	0	1	2	3	4
Starting index u_r	4	19	82	337	1360
Maximum V_r	2	5	15	50	176
Increase $V_r - V_{r-1}$	—	3	10	35	126
Next minimum	-2	-8	-28	-98	-350

OEIS: maxima A024718; increases A001700.

Absolute values of the next minima: A066796.

First observations

Positive arch r	0	1	2	3	4
Starting index u_r	4	19	82	337	1360
Maximum V_r	2	5	15	50	176
Increase $V_r - V_{r-1}$	—	3	10	35	126
Next minimum	-2	-8	-28	-98	-350

OEIS: maxima A024718; increases A001700.

Absolute values of the next minima: A066796.

$$V_r - V_{r-1} = \binom{2r+1}{r} = (2r+1) \text{Cat}_r \quad (r \geq 1)$$

Catalan numbers: A000108. Next minimum = $-2(V_r - 1)$.

First observations

Positive arch r	0	1	2	3	4
Starting index u_r	4	19	82	337	1360
Maximum V_r	2	5	15	50	176
Increase $V_r - V_{r-1}$	—	3	10	35	126
Next minimum	-2	-8	-28	-98	-350

OEIS: maxima A024718; increases A001700.

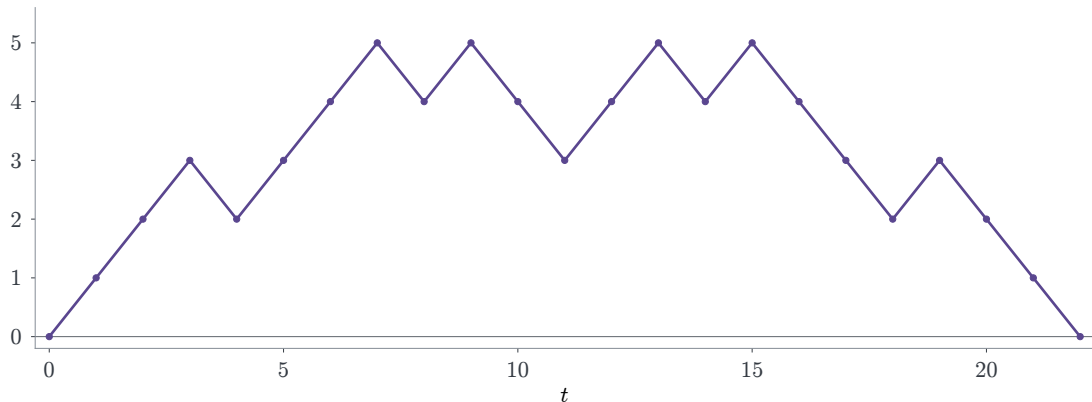
Absolute values of the next minima: A066796.

$$V_r - V_{r-1} = \binom{2r+1}{r} = (2r+1) \text{Cat}_r \quad (r \geq 1)$$

Catalan numbers: A000108. Next minimum = $-2(V_r - 1)$.

What combinatorial structure explains these numbers?

Arches as binary words



P_r : the word of positive arch r ($r = 0, 1, \dots$). Here $t = m - 19$.

$P_1 = 0001000101100101110111$ 0: up 1: down

Binary increments

$\Delta f(m) = f(m+1) - f(m)$. The positive arch $19 \leq m < 41$.

$\Delta \mathfrak{a}$ 0001000101100101110111

$\Delta \mathfrak{b}$ 1110111010011010001000

Binary increments

$\Delta f(m) = f(m+1) - f(m)$. The positive arch $19 \leq m < 41$.

$\Delta \mathbf{a}$ 0001000101100101110111

$\Delta \mathbf{b}$ 1110111010011010001000

$$\Delta \delta = \Delta \mathbf{b} - \Delta \mathbf{a} = 1 - 2 \Delta \mathbf{a}$$

$\Delta \mathbf{a} = 0$: up

$\Delta \mathbf{a} = 1$: down

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X

Y

Output

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X $\textcircled{0}$

Y

Output

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X $\textcircled{0}$

Y

Output 0

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X 0 $\textcircled{0}$

Y

Output 0

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X 0 0

Y

Output 0 0

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X 0 0 ①

Y

Output 0 0

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X 0 0 ①

Y

Output 0 0 1

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1	
Y				0
Output	0	0	1	

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1	
Y				0
Output	0	0	1	0

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1	0
Y				0
Output	0	0	1	0

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0
Y				0	
Output	0	0	1	0	0

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0	①
-----	---	---	---	--	---	---

Y				0		
-----	--	--	--	---	--	--

Output	0	0	1	0	0
--------	---	---	---	---	---

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0	①
Y				0		
Output	0	0	1	0	0	1

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0	1
Y				0		①
Output	0	0	1	0	0	1

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0	1	
Y				0			①
Output	0	0	1	0	0	1	1

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0	1	
Y				0		1	0
Output	0	0	1	0	0	1	1

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0	1		
Y				0			1	0
Output	0	0	1	0	0	1	1	0

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0	1		①
Y				0			1	0
Output	0	0	1	0	0	1	1	0

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0	1			①
Y				0			1	0	
Output	0	0	1	0	0	1	1	0	1

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0	1			1
Y				0			1	0	①
Output	0	0	1	0	0	1	1	0	1

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0	1				1
Y				0			1	0		1
Output	0	0	1	0	0	1	1	0	1	1

Interleaving

Start with X . After each output letter, read X if it is 0 and Y if it is 1.

$$X = 0W1 = 001011, \quad Y = W = 0101$$

X	0	0	1		0	1			1	
Y				0			1	0		1
Output	0	0	1	0	0	1	1	0	1	1

$$T(W) = \text{Interleave}(0W1, W)$$

Construction and well-definedness

$P_r = 0A_r1$: remove the first 0 and the last 1. $A_0 = 0101$.

$$T(A_r) = 0N_r, \quad T(0N_r) = P_{r+1}$$

N_r : increments of $\mathfrak{b}$ along the negative arch.

Construction and well-definedness

$P_r = 0A_r1$: remove the first 0 and the last 1. $A_0 = 0101$.

$$T(A_r) = 0N_r, \quad T(0N_r) = P_{r+1}$$

N_r : increments of $\mathfrak{b}$ along the negative arch.

$$T(A_0) = 0010011011, \quad T^2(A_0) = P_1$$

Both input words are used completely
The construction never needs a missing letter.

Construction and well-definedness

$P_r = 0A_r1$: remove the first 0 and the last 1. $A_0 = 0101$.

$$T(A_r) = 0N_r, \quad T(0N_r) = P_{r+1}$$

N_r : increments of $\mathfrak{b}$ along the negative arch.

$$T(A_0) = 0010011011, \quad T^2(A_0) = P_1$$

Both input words are used completely

The construction never needs a missing letter.

The constructed sequences satisfy the recurrence, with $1 \leq \mathfrak{a}(m), \mathfrak{b}(m) \leq m$.

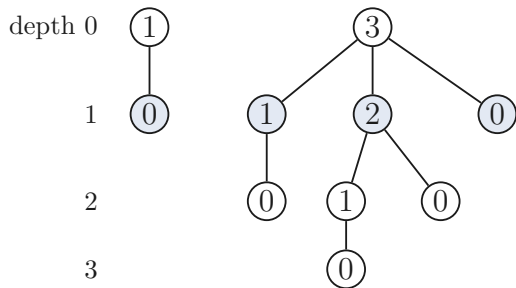
$$1 \leq n - \widetilde{Q}(n - j) \leq n - 1 \quad (j = 1, 2, n \geq 3)$$

$$\widetilde{Q}(n) \text{ exists for every } n \geq 1.$$

A different, computer-certified proof: M. Mantovanelli, arXiv:2603.29622.

Forests and binomial layers

Each zero of A_r gives one vertex.



Forest of A_1

Labels give the number of children.

A vertex labelled L has one child of each label $0, 1, \dots, L-1$.

At depth k below this root:
 $\binom{L}{k}$ vertices.

Roots of A_r : $1, 3, \dots, 2r+1$.

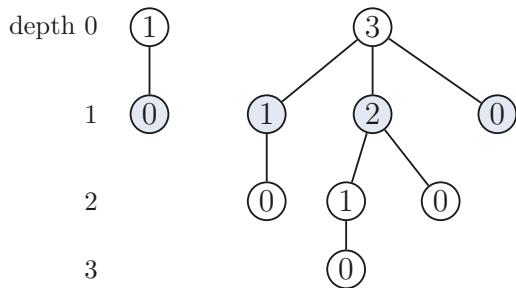
$$B_{r,k} = \sum_{i=0}^r \binom{2i+1}{k}$$

Number of vertices at depth k .

For A_1 : 2, 4, 3, 1.

Forests and binomial layers

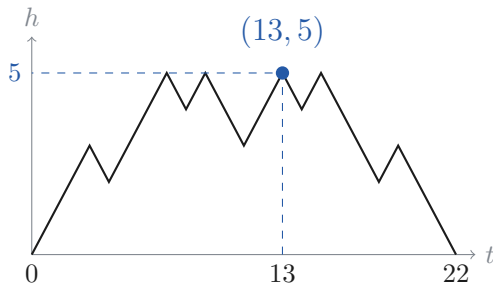
Each zero of A_r gives one vertex.



Forest of A_1

Labels give the number of children.

The corresponding point of P_1



Layer size 4; arch height $4 + 1 = 5$.
The first 0 in $P_1 = 0A_11$ adds one up-step.

Record heights and fibers

A record cut: the first position where the arch reaches height h .

Count j ones before the cut; move just after zero $j + 1$.

The image is again a record cut. Its height is $g(h)$.

Height h	0	1	2	3	4	5
First position $\rho(h)$	0	1	2	3	6	7
Image height $g(h)$	1	1	1	1	2	2

$h = 4$: prefix 000100; one 1; after the second 0, height 2.

Record heights and fibers

A record cut: the first position where the arch reaches height h .

Count j ones before the cut; move just after zero $j + 1$.

The image is again a record cut. Its height is $g(h)$.

Height h	0	1	2	3	4	5
First position $\rho(h)$	0	1	2	3	6	7
Image height $g(h)$	1	1	1	1	2	2

$h = 4$: prefix 000100; one 1; after the second 0, height 2.

A fiber groups record heights with the same image.

$$\{0, 1, 2, 3\} \quad \{4, 5\} \quad f_1 = (4, 2)$$

$$4 + 2 = 6 = V_1 + 1.$$

From fibers to Catalan numbers

Components: maximal intervals of consecutive image heights.

Fiber sizes give their lengths at the next generation.

Only the first fiber loses one element: the height-zero parameter.

$$f_1 = (4, 2) \quad \longmapsto \quad \{1, 2, 3\}, \{6, 7\} \text{ in } P_2$$

Two components, with lengths 3 and 2.

From fibers to Catalan numbers

Components: maximal intervals of consecutive image heights.

Fiber sizes give their lengths at the next generation.

Only the first fiber loses one element: the height-zero parameter.

$$f_1 = (4, 2) \longmapsto \{1, 2, 3\}, \{6, 7\} \text{ in } P_2$$

Two components, with lengths 3 and 2.

O_L : a component of length L not containing height 1.

$$O_L \longmapsto O_2 + O_3 + \cdots + O_{L+1}$$

The component containing height 1 follows a separate rule.

From fibers to Catalan numbers

Components: maximal intervals of consecutive image heights.

Fiber sizes give their lengths at the next generation.

Only the first fiber loses one element: the height-zero parameter.

$$f_1 = (4, 2) \longmapsto \{1, 2, 3\}, \{6, 7\} \text{ in } P_2$$

Two components, with lengths 3 and 2.

O_L : a component of length L not containing height 1.

$$O_L \longmapsto O_2 + O_3 + \cdots + O_{L+1}$$

The component containing height 1 follows a separate rule.

V_r : arch maximum. $W_r = V_r - V_{r-1}$. $\mathscr{W}(t) = 1 + \sum_{r \geq 1} W_r t^r$.

$\text{Cat}(t) = \sum_{r \geq 0} \text{Cat}_r t^r$: the Catalan generating function.

$$\mathscr{W} = \text{Cat} + 2t \text{Cat} \mathscr{W}$$

$$W_r = (2r + 1) \text{Cat}_r = \binom{2r+1}{r} \quad (r \geq 1)$$

Limit and error term

Maximum of the positive arch P_r

$$V_r = 2 + \sum_{j=1}^r \binom{2j+1}{j}$$

$$V_r \asymp \frac{4^r}{\sqrt{r}}$$

$$\frac{|\delta(m)|}{m} \leq \frac{V_r}{u_r} = O(r^{-1/2}), \quad r \sim \log_4 m$$

Positive arches; the following negative arches are controlled by domination.

Starting index of P_r

$$u_0 = 4, \quad u_1 = 19, \quad u_2 = 82 \\ u_r \asymp 4^r$$

Limit and error term

Maximum of the positive arch P_r

$$V_r = 2 + \sum_{j=1}^r \binom{2j+1}{j}$$

$$V_r \asymp \frac{4^r}{\sqrt{r}}$$

$$\frac{|\delta(m)|}{m} \leq \frac{V_r}{u_r} = O(r^{-1/2}), \quad r \sim \log_4 m$$

Positive arches; the following negative arches are controlled by domination.

$$\tilde{Q}(2m) = m + \delta(m) + O(\log m), \quad \tilde{Q}(2m-1) = m - \delta(m) + O(\log m)$$

Starting index of P_r

$$u_0 = 4, \quad u_1 = 19, \quad u_2 = 82 \\ u_r \asymp 4^r$$

Limit and error term

Maximum of the positive arch P_r

$$V_r = 2 + \sum_{j=1}^r \binom{2j+1}{j}$$

$$V_r \asymp \frac{4^r}{\sqrt{r}}$$

$$\frac{|\delta(m)|}{m} \leq \frac{V_r}{u_r} = O(r^{-1/2}), \quad r \sim \log_4 m$$

Positive arches; the following negative arches are controlled by domination.

$$\tilde{Q}(2m) = m + \delta(m) + O(\log m), \quad \tilde{Q}(2m-1) = m - \delta(m) + O(\log m)$$

$$\tilde{Q}(n) = \frac{n}{2} + O\left(\frac{n}{\sqrt{\log n}}\right), \quad \frac{\tilde{Q}(n)}{n} \longrightarrow \frac{1}{2}$$

The exact limsup constant

$$F_{\tilde{Q}}(n) = \left(\frac{\tilde{Q}(n)}{n} - \frac{1}{2} \right) \sqrt{\log_2 n}$$

The exact limsup constant

$$F_{\tilde{Q}}(n) = \left(\frac{\tilde{Q}(n)}{n} - \frac{1}{2} \right) \sqrt{\log_2 n}$$

$$\limsup_{n \rightarrow \infty} F_{\tilde{Q}}(n) = \limsup_{n \rightarrow \infty} |F_{\tilde{Q}}(n)| = C_{\text{layer}}$$

$$C_{\text{layer}} = \frac{z_0}{2} = 0.138014902399 \dots$$

$$z_0[1 + \Phi(-z_0)] = \varphi(z_0), \quad z_0 > 0.$$

The exact limsup constant

$$F_{\tilde{Q}}(n) = \left(\frac{\tilde{Q}(n)}{n} - \frac{1}{2} \right) \sqrt{\log_2 n}$$

$$\limsup_{n \rightarrow \infty} F_{\tilde{Q}}(n) = \limsup_{n \rightarrow \infty} |F_{\tilde{Q}}(n)| = C_{\text{layer}}$$

$$C_{\text{layer}} = \frac{z_0}{2} = 0.138014902399 \dots$$

$$z_0[1 + \Phi(-z_0)] = \varphi(z_0), \quad z_0 > 0.$$

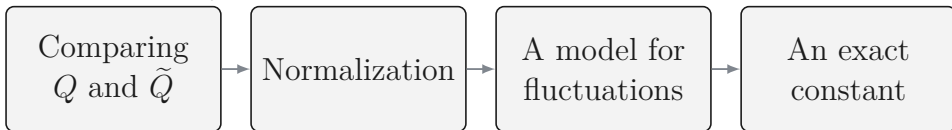
The order $n/\sqrt{\log n}$ is optimal.

Back to Q

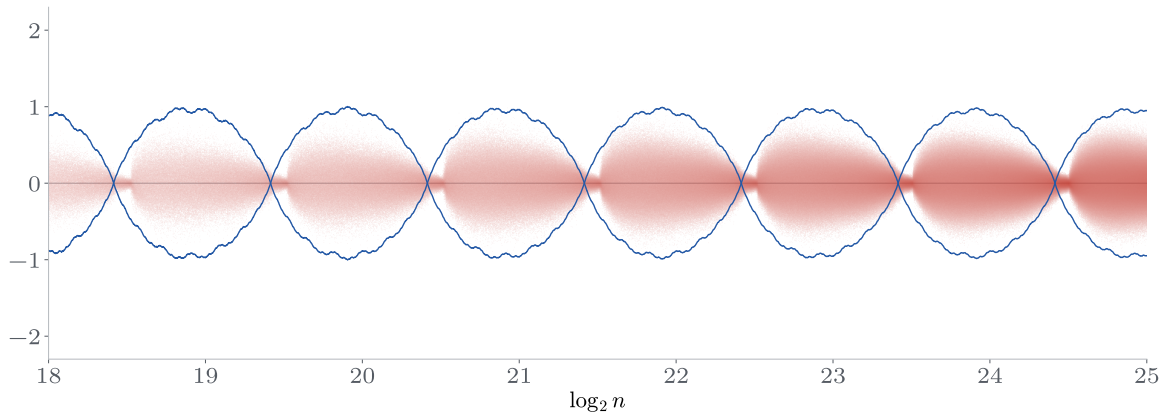
I. Proof of well-definedness and asymptotic behavior of $\tilde{Q}$



II. Back to Q



Back to Q : the cages of $\tilde{Q}$



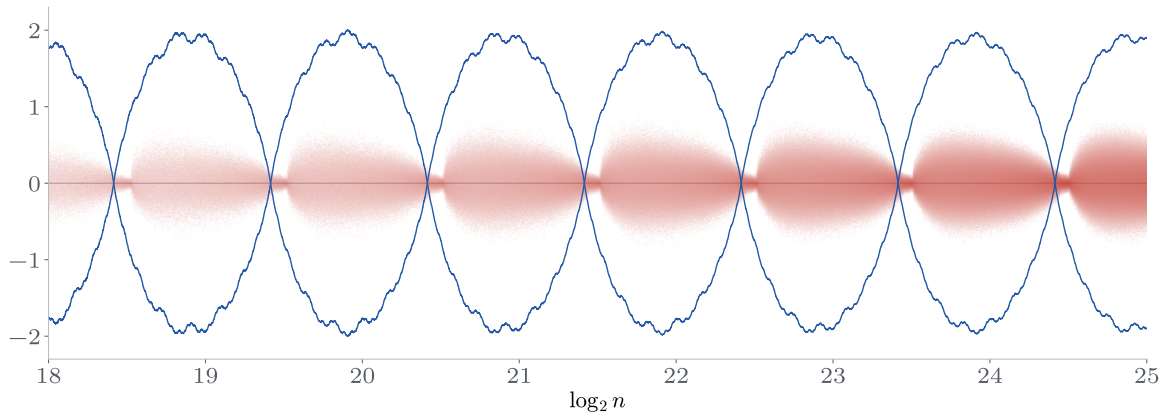
— Q : normalized deviation

— $\tilde{Q}$: normalized deviation

Deviation $X(n)/n - \frac{1}{2}$. Q : divided by its interval maximum. $\tilde{Q}$: divided by $0.150/\sqrt{\log_2 n}$.

Here, intervals lie between consecutive zeros of δ , mapped to the n -axis.

Back to Q : the cages of $\tilde{Q}$



— Q : normalized deviation

— $\tilde{Q}$: normalized deviation $\times 2$

Deviation $X(n)/n - \frac{1}{2}$. Q : divided by its interval maximum. $\tilde{Q}$: divided by $0.150/\sqrt{\log_2 n}$.

Here, intervals lie between consecutive zeros of δ , mapped to the n -axis.

Correlations between nearby fluctuations

$$e(n) = Q(n) - \frac{1}{2}n$$

Compare $e(n)$ and $e(n + m)$ in short blocks.

$$x_n = e(n) - \text{local mean}, \quad y_n = e(n + m) - \text{local mean}$$

$$\hat{\rho}_m = \frac{\sum x_n y_n}{\sqrt{\sum x_n^2 \sum y_n^2}}$$

Positive: same direction. Negative: opposite directions.

Separate local means for the two samples; centered sums pooled over the blocks.

Measured correlations

m	Measured correlation
1	-0.414202
2	-0.292911
3	$+0.353547$
4	-0.030360

$2^{22} \leq n < 2^{28}$. Blocks of 2^{16} terms; local means removed.

Measured correlations

m	Measured correlation	A guess
1	-0.414202	$1 - \sqrt{2}$
2	-0.292911	
3	+0.353547	
4	-0.030360	

$2^{22} \leq n < 2^{28}$. Blocks of 2^{16} terms; local means removed.

Measured correlations

m	Measured correlation	A guess
1	-0.414202	$1 - \sqrt{2}$
2	-0.292911	$\frac{1}{\sqrt{2}} - 1$
3	+0.353547	$\frac{1}{2\sqrt{2}}$
4	-0.030360	$\frac{1}{2} - \frac{3\sqrt{2}}{8}$

$2^{22} \leq n < 2^{28}$. Blocks of 2^{16} terms; local means removed.

An empirical correlation law

$$\rho_0 = 1, \quad \rho_1 = 1 - \sqrt{2}$$

The guessed values satisfy

$$-2\rho_2 = \rho_1 + \rho_0, \quad -2\rho_3 = \rho_2 + \rho_1.$$

An empirical correlation law

$$\rho_0 = 1, \quad \rho_1 = 1 - \sqrt{2}$$

The guessed values satisfy

$$-2\rho_2 = \rho_1 + \rho_0, \quad -2\rho_3 = \rho_2 + \rho_1.$$

$$\rho_m = -\frac{1}{2}(\rho_{m-1} + \rho_{m-2}) \quad (m \geq 2)$$

An empirical correlation law

$$\rho_0 = 1, \quad \rho_1 = 1 - \sqrt{2}$$

The guessed values satisfy

$$-2\rho_2 = \rho_1 + \rho_0, \quad -2\rho_3 = \rho_2 + \rho_1.$$

$$\rho_m = -\frac{1}{2}(\rho_{m-1} + \rho_{m-2}) \quad (m \geq 2)$$

Does this describe the correlations at every distance?

Can we explain it from the recurrence of Q ?

What the recurrence tells us

$$i_n = n - Q(n - 1), \quad j_n = n - Q(n - 2)$$

The two positions read to compute $Q(n)$.

$$e(n) + \frac{1}{2}(e(n - 1) + e(n - 2)) = e(i_n) + e(j_n) + \frac{3}{4}$$

If $Q(n) \approx n/2$, parents are read near $n/2$ and produce terms near n .

What the recurrence tells us

$$i_n = n - Q(n - 1), \quad j_n = n - Q(n - 2)$$

The two positions read to compute $Q(n)$.

$$e(n) + \frac{1}{2}(e(n - 1) + e(n - 2)) = e(i_n) + e(j_n) + \frac{3}{4}$$

Compute	First position	Second position
$Q(n)$	i_n	j_n
$Q(n + 1)$	i_{n+1}	$j_{n+1} = i_n + 1$

This pair of neighboring positions is forced by Q .

If $Q(n) \approx n/2$, parents are read near $n/2$ and produce terms near n .

Building a covariance model

Parents (indices near $n/2$): variance V , neighboring covariance R .

U_n models the centered sum of the two terms read at step n .

Building a covariance model

Parents (indices near $n/2$): variance V , neighboring covariance R .

U_n models the centered sum of the two terms read at step n .

$$\begin{aligned}\mathrm{Var}(U_n) &= 2V, & \mathrm{Cov}(U_n, U_{n+1}) &= R \\ \mathrm{Cov}(U_n, U_{n+m}) &= 0 & (m \geq 2)\end{aligned}$$

Building a covariance model

Parents (indices near $n/2$): variance V , neighboring covariance R .

U_n models the centered sum of the two terms read at step n .

$$\begin{aligned}\text{Var}(U_n) &= 2V, & \text{Cov}(U_n, U_{n+1}) &= R \\ \text{Cov}(U_n, U_{n+m}) &= 0 & (m \geq 2)\end{aligned}$$

Two terms contribute to the variance; one forced pair contributes at distance 1.

Assumption: the forced pair samples the parent covariance R ;
all other covariance contributions are ignored.

Recovering the correlation law

Y_n : centered stationary fluctuations produced at the new scale.

$$Y_n + \frac{1}{2}(Y_{n-1} + Y_{n-2}) = U_n$$

Y_n is the stable response to present and past inputs U .

Recovering the correlation law

Y_n : centered stationary fluctuations produced at the new scale.

$$Y_n + \frac{1}{2}(Y_{n-1} + Y_{n-2}) = U_n$$

Y_n is the stable response to present and past inputs U .

For $m \geq 2$, U_n is uncorrelated with Y_{n-m} .

Take covariance with Y_{n-m} and divide by $\text{Var}(Y_n)$.

Recovering the correlation law

Y_n : centered stationary fluctuations produced at the new scale.

$$Y_n + \frac{1}{2}(Y_{n-1} + Y_{n-2}) = U_n$$

Y_n is the stable response to present and past inputs U .

For $m \geq 2$, U_n is uncorrelated with Y_{n-m} .

Take covariance with Y_{n-m} and divide by $\text{Var}(Y_n)$.

$$\rho_m = -\frac{1}{2}(\rho_{m-1} + \rho_{m-2}) \quad (m \geq 2)$$

This recovers the guessed recurrence. We still need ρ_1 .

From one scale to the next

$$(V, R) \longmapsto (V', R')$$

Parents near $n/2$

New terms near n

$$V' = \text{Var}(Y_n), \quad R' = \text{Cov}(Y_n, Y_{n-1}).$$

From one scale to the next

$$(V, R) \longmapsto (V', R')$$

Parents near $n/2$

New terms near n

$$V' = \text{Var}(Y_n), \quad R' = \text{Cov}(Y_n, Y_{n-1}).$$

Take covariances in the linear equation:

$$3V' + R' = 8V - 2R, \quad V' + 3R' = 2R.$$

From one scale to the next

$$(V, R) \longmapsto (V', R')$$

Parents near $n/2$

New terms near n

$$V' = \text{Var}(Y_n), \quad R' = \text{Cov}(Y_n, Y_{n-1}).$$

Take covariances in the linear equation:

$$3V' + R' = 8V - 2R, \quad V' + 3R' = 2R.$$

$$V' = 3V - R, \quad R' = R - V$$

Where does $1 - \sqrt{2}$ come from?

At the parent scale, write the half-variance of neighboring differences as

$$D = V - R \geq 0.$$

The covariance transfer becomes $V' = 2V + D$, $D' = 2V + 2D$.

Where does $1 - \sqrt{2}$ come from?

At the parent scale, write the half-variance of neighboring differences as

$$D = V - R \geq 0.$$

The covariance transfer becomes $V' = 2V + D$, $D' = 2V + 2D$.

Normalize by the variance: $r = D/V = 1 - \rho_1$.

$$r' = \frac{2+2r}{2+r}, \quad r' = r \iff r^2 = 2$$

Where does $1 - \sqrt{2}$ come from?

At the parent scale, write the half-variance of neighboring differences as

$$D = V - R \geq 0.$$

The covariance transfer becomes $V' = 2V + D$, $D' = 2V + 2D$.

Normalize by the variance: $r = D/V = 1 - \rho_1$.

$$r' = \frac{2+2r}{2+r}, \quad r' = r \iff r^2 = 2$$

$$r \longrightarrow \sqrt{2}, \quad \rho_1 \longrightarrow 1 - \sqrt{2}$$

Every starting ratio $r \geq 0$ converges to this same value.

How large are the fluctuations?

$$e(n) = Q(n) - \frac{1}{2}n.$$

Klaus Pinn: an observed exponent near 0.88.

$$\sigma_k = \sqrt{\langle e^2 \rangle_k - \langle e \rangle_k^2}, \quad \frac{\sigma_k}{\sigma_{k-1}} \approx 2^\alpha, \quad \alpha = 0.88(1)$$

$\langle \cdot \rangle_k$: average over generation k , at indices of order 2^k .

An empirical fluctuation scale $n^{0.88}$ for $Q(n) - n/2$.

For $Q(n)/n - 1/2$, the corresponding scale is $n^{-0.12}$.

K. Pinn, *Order and Chaos in Hofstadter's Q(n) Sequence*, 1999, section 4.

From variance growth to an exponent

$$n \mapsto 2n : \quad V \mapsto (2 + \sqrt{2})V, \quad \sqrt{V} \mapsto \sqrt{2 + \sqrt{2}}\sqrt{V}$$

$$\alpha_* = \frac{\log(2 + \sqrt{2})}{2 \log 2} = 0.88577665 \dots$$

$$2^{\alpha_*} = \sqrt{2 + \sqrt{2}}$$

Model prediction: a standard deviation of order n^{α_*} .

From variance growth to an exponent

$$n \mapsto 2n : \quad V \mapsto (2 + \sqrt{2})V, \quad \sqrt{V} \mapsto \sqrt{2 + \sqrt{2}}\sqrt{V}$$

$$\alpha_* = \frac{\log(2 + \sqrt{2})}{2 \log 2} = 0.88577665 \dots$$

$$2^{\alpha_*} = \sqrt{2 + \sqrt{2}}$$

Model prediction: a standard deviation of order n^{α_*} .

A stronger question: for every $\varepsilon > 0$,

$$Q(n) = \frac{n}{2} + O_\varepsilon(n^{\alpha_* + \varepsilon}) \quad ?$$

Testing the prediction on Q

Compare blocks of 1024 terms with the local variances of their parents.

Generation	$\hat{\lambda}$	$\hat{\alpha} = \frac{1}{2} \log_2 \hat{\lambda}$
26	3.41395684	0.88572241
27	3.41446218	0.88582918
28	3.41417809	0.88576916
Model	3.41421356	0.88577665

$$\text{Model: } \lambda = 2 + \sqrt{2}, \quad \alpha_* = \log_2(2 + \sqrt{2})/2.$$

Q computed to 2^{29} . Generation g : $2^g \leq n < 2^{g+1}$. Recomputed 7 October 2026.

The honeycomb constant

In our model, doubling the scale multiplies the standard deviation by

$$2^{\alpha_*} = \sqrt{2 + \sqrt{2}}$$

H. Duminil-Copin and S. Smirnov, *Annals of Mathematics* **175** (2012), 1653–1665.

The honeycomb constant

In our model, doubling the scale multiplies the standard deviation by

$$2^{\alpha_*} = \sqrt{2 + \sqrt{2}} = \mu_{\text{hex}}$$

μ_{hex} : the connective constant of the honeycomb lattice.

It is the exponential growth rate of the number of self-avoiding walks.

H. Duminil-Copin and S. Smirnov, *Annals of Mathematics* **175** (2012), 1653–1665.

The honeycomb constant

In our model, doubling the scale multiplies the standard deviation by

$$2^{\alpha_*} = \sqrt{2 + \sqrt{2}} = \mu_{\text{hex}}$$

μ_{hex} : the connective constant of the honeycomb lattice.

It is the exponential growth rate of the number of self-avoiding walks.

Is there a connection between Q and self-avoiding walks
on the honeycomb lattice?

H. Duminil-Copin and S. Smirnov, *Annals of Mathematics* **175** (2012), 1653–1665.

From the model to Q

The model explains both the observed recurrence and its initial value.

From the model to Q

The model explains both the observed recurrence and its initial value.

For Q , justify the covariance assumption at large scales.

Do the forced reading pairs reproduce the parent covariance?

Are the remaining covariance contributions negligible?

From the model to Q

The model explains both the observed recurrence and its initial value.

For Q , justify the covariance assumption at large scales.

Do the forced reading pairs reproduce the parent covariance?

Are the remaining covariance contributions negligible?

The numerical agreement supports this mechanism.

A proof for the actual sequence still requires this step.

The correlations and the OEIS

$$\rho_m = \frac{a_{m+1} + (\sqrt{2} - 1)^2 a_m}{2^m}$$

$$a_0 = 0, \quad a_1 = 1, \quad a_m = -a_{m-1} - 2a_{m-2}$$

$$0, 1, -1, -1, 3, -1, -5, 7, 3, -17, \dots$$

OEIS A001607 A sequence related to Lehmer numbers

OEIS A001607. The formula concerns the correlations predicted by the model.

The correlations and the OEIS

$$\rho_m = \frac{a_{m+1} + (\sqrt{2} - 1)^2 a_m}{2^m}$$

$$a_0 = 0, \quad a_1 = 1, \quad a_m = -a_{m-1} - 2a_{m-2}$$

$$0, 1, -1, -1, 3, -1, -5, 7, 3, -17, \dots$$

OEIS A001607 A sequence related to Lehmer numbers

Multiplying ρ_m by 2^m gives the same recurrence as a_m .

One integer sequence describes the whole correlation law.

OEIS A001607. The formula concerns the correlations predicted by the model.

A three-term recurrence

Why trust this model?

Let us test the same approach on a three-term recurrence.

$$a(n) = a(n - a(n - 1)) + a(n - a(n - 2)) + a(n - a(n - 3))$$

Three-term Hofstadter recurrence. Examples studied by Nathan Fox.

Seven examples: initial conditions matter

OEIS	Initial conditions	Recurrence starts at
A292351	$(1, 1, 3, 4, 4)$	$n = 6$
A296440	$(2, 3, 3, 4, 6, 6)$	$n = 7$
A373227	$(1, 2, \dots, 7)$	$n = 8$
A373228	$(1, 2, \dots, 8)$	$n = 9$
A373229	$(1, 2, \dots, 9)$	$n = 10$
A373230*	$(1, 2, \dots, 10)$	$n = 11$
A373231*	$(1, 2, \dots, 11)$	$n = 12$

*These two sequences also use $a(n) = 0$ for $n \leq 0$.

OEIS entries A292351, A296440, A373227–A373231.

A covariance model for three terms

Y_j models local fluctuations of $a(n) - \frac{2}{3}n$.

$$V = \text{Var}(Y_j), \quad D_m = \frac{1}{2} \text{Var}(Y_{j+m} - Y_j) \quad (m = 1, 2).$$

Keep the forced neighboring pairs; ignore the other covariance terms.

A covariance model for three terms

Y_j models local fluctuations of $a(n) - \frac{2}{3}n$.

$$V = \text{Var}(Y_j), \quad D_m = \frac{1}{2} \text{Var}(Y_{j+m} - Y_j) \quad (m = 1, 2).$$

Keep the forced neighboring pairs; ignore the other covariance terms.

$$Z = \begin{pmatrix} V \\ D_1 \\ D_2 \end{pmatrix}, \quad Z' = B_3 Z, \quad B_3 = \begin{pmatrix} 3 & 8/3 & 4/3 \\ 3 & 6 & 0 \\ 4 & 0 & 5 \end{pmatrix}$$

The prime denotes the next scale: $n \mapsto 3n$.

A closed form for λ_3

λ_3 : largest eigenvalue of B_3 , the variance multiplier when $n \mapsto 3n$.

$$3\lambda^3 - 42\lambda^2 + 149\lambda - 54 = 0$$

$$\lambda_3 = \frac{14 + 2\sqrt{47} \cos\left[\frac{1}{3} \arccos\left(-\frac{142}{47\sqrt{47}}\right)\right]}{3} = 8.23276848 \dots$$

Measured on seven sequences: from 8.2210 to 8.2351.

$$\alpha_3 = \frac{\log \lambda_3}{2 \log 3} = 0.95944783 \dots$$

Model prediction: $a(n) - \frac{2}{3}n$ has standard deviation of order n^{α_3} .

The predicted correlations

$$\rho_m = \text{Corr}(Y_j, Y_{j+m}), \quad \rho_0 = 1.$$

$$D_m = V(1 - \rho_m).$$

$$\rho_1 = 1 - \frac{3}{\lambda_3 - 6}, \quad \rho_2 = 1 - \frac{4}{\lambda_3 - 5}$$

The predicted correlations

$$\rho_m = \text{Corr}(Y_j, Y_{j+m}), \quad \rho_0 = 1.$$

$$D_m = V(1 - \rho_m).$$

$$\rho_1 = 1 - \frac{3}{\lambda_3 - 6}, \quad \rho_2 = 1 - \frac{4}{\lambda_3 - 5}$$

$$\rho_m = -\frac{2}{3}(\rho_{m-1} + \rho_{m-2} + \rho_{m-3}) \quad (m \geq 3)$$

Three-term correlations: experiment and model

OEIS	ρ_1	ρ_2	ρ_3	ρ_4
A292351	-0.343888	-0.236906	-0.279457	+0.573486
A296440	-0.343060	-0.238278	-0.279094	+0.573636
A373227	-0.344612	-0.235734	-0.279737	+0.573466
A373228	-0.344420	-0.235992	-0.279703	+0.573356
A373229	-0.343973	-0.236804	-0.279456	+0.573488
A373230	-0.343525	-0.237406	-0.279420	+0.573460
A373231	-0.343938	-0.236628	-0.279462	+0.573074
Model	-0.343623	-0.237330	-0.279365	+0.573545

All seven sequences computed to 2^{28} . $e(n) = a(n) - \frac{2}{3}n$.
 Measurements on $2^{22} \leq n < 2^{28}$; blocks of 2^{16} terms; local means removed.

The same approach can be investigated for recurrences with more terms.

Main results

$\tilde{Q}$: proved	Q : model and experiments
Combinatorial construction and well-definedness	A precise correlation law
$\tilde{Q}(n)/n \longrightarrow 1/2$	Candidate exponent α_*
Optimal error order and exact lim-sup constant	Close numerical agreement, also for three-term recurrences

$$\alpha_* = \frac{\log(2 + \sqrt{2})}{2 \log 2} = 0.88577665 \dots$$

For Q , well-definedness and the limit remain open.

Thank you

Questions?

B. Cloitre, arXiv:2604.06237. M. Mantovanelli, arXiv:2603.16111 and arXiv:2603.29622.
D.R. Hofstadter (1979). C.L. Mallows (1991). T. Kubo and R. Vakil (1996).
K. Pinn (1999). H. Duminil-Copin and S. Smirnov, *Ann. of Math.* **175** (2012).
N. Fox, three-term recurrence, *J. Integer Seq.* **20** (2017) and **27** (2024).
OEIS A005185, A394051, A004001, A292351, A296440.