# Approximate Counting

Professor Thomas Prellberg

School of Mathematical Sciences, Queen Mary University of London

Inaugural Lecture, March 30, 2017

# Topic Outline

# Pathway to Professorship

| | |
|---|---|
| 1983 | Undergraduate Student, *TU Braunschweig*, Germany |
| 1988 | Graduate Student, *Virginia Tech*, VA, USA |
| 1989 | Visiting Scientist, *Weizmann Institute*, Israel |
| 1991 | PhD in Mathematical Physics, MSc in Mathematics, MSc in Physics |
| 1991 | Research Fellow, *University of Melbourne*, Australia |
| 1994 | EU Postdoctoral Fellow, *University of Oslo*, Norway |
| 1996 | Research Associate, *University of Manchester*, UK |
| 1999 | Visiting Assistant Professor, *Syracuse University*, NY, USA |
| 2000 | Assistent, *Clausthal University of Technology*, Germany |
| 2002 | Habilitation, Promotion to Oberassistent |
| 2004 | Senior Lecturer, *Queen Mary University of London*, UK |
| 2005 | Promotion to Reader |
| 2008 | Libra Visiting Professor of Diversity, *University of Maine*, ME, USA |
| 2014 | Promotion to Professor |

# Outline

1. A Brief History of Counting
2. Counting and Combinatorics
3. Exact Counting
   - Lattice Paths and Counting Functions
   - Walks in a Triangle
   - Lattice Path Models of Polymers
   - Pulling Polymers off a Surface
4. Approximate Counting
   - Sampling of Simple Random Walks
   - Sampling of Self-Avoiding Walks
   - Applications

# Counting and Language

- Pirahã (Amazon):
  *hói* = one/small/less,
  *hoí* = two/many/large/more

# Counting and Language

- Pirahã (Amazon):
  *hói* = one/small/less,
  *hoí* = two/many/large/more
- Warlpiri (Australia):
  one, two, many

# Counting and Language

- Pirahã (Amazon):
  *hói* = one/small/less,
  *hoí* = two/many/large/more
- Warlpiri (Australia):
  one, two, many
- Gumulgal (Australia):
  one, two, two-one, two-two,
  two-two-one, . . .

# Counting and Language

- Pirahã (Amazon):
  *hói* = one/small/less,
  *hoí* = two/many/large/more
- Warlpiri (Australia):
  one, two, many
- Gumulgal (Australia):
  one, two, two-one, two-two,
  two-two-one, . . .
- Roman numerals:
  I, II, III, IV, V, VI, . . . , LXXXIX,
  XC, XCI, . . .

# Counting and Language

- Pirahã (Amazon):
  *hói* = one/small/less,
  *hoí* = two/many/large/more
- Warlpiri (Australia):
  one, two, many
- Gumulgal (Australia):
  one, two, two-one, two-two,
  two-two-one, . . .
- Roman numerals:
  I, II, III, IV, V, VI, . . . , LXXXIX,
  XC, XCI, . . .
- India: zero, decimal system
  1, 2, . . . , 9, 10, 11, . . .

# Counting and Society

- 20,000 BC: Ishango bone (Congo)
  tally marks on baboon fibula

# Counting and Society

- 20,000 BC: Ishango bone (Congo)
  tally marks on baboon fibula
- 4000 BC: Sumeria
  livestock $\equiv$ tokens

# Counting and Society

- 20,000 BC: Ishango bone (Congo)
  tally marks on baboon fibula
- 4000 BC: Sumeria
  livestock $\equiv$ tokens
- 3000 BC: Egypt

  hieroglyph for one million 

# Counting and Society

- 20,000 BC: Ishango bone (Congo)
  tally marks on baboon fibula
- 4000 BC: Sumeria
  livestock $\equiv$ tokens
- 3000 BC: Egypt
  hieroglyph for one million 
- 500 BC: Pythagoras (Greece)
  "of all things numbers are the first"

# Counting and Society

- 20,000 BC: Ishango bone (Congo)
  tally marks on baboon fibula
- 4000 BC: Sumeria
  livestock $\equiv$ tokens
- 3000 BC: Egypt
  hieroglyph for one million 
- 500 BC: Pythagoras (Greece)
  "of all things numbers are the first"
- Roman Empire:
  Mathematics only for bookkeeping

# Counting and Mathematics

What are numbers (natural numbers, positive integers)?

# Counting and Mathematics

What are numbers (natural numbers, positive integers)?

- Peano axioms (1889):
  formalisation of the "obvious"
  1. 1 is a number
  2. every number has a successor
  3. 1 is not a successor

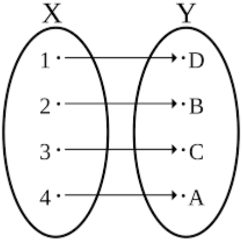
# Counting and Mathematics

What are numbers (natural numbers, positive integers)?

- Peano axioms (1889):
  formalisation of the "obvious"
  1. 1 is a number
  2. every number has a successor
  3. 1 is not a successor

  and the "less obvious"
  4. two numbers with the same successors are themselves equal
  5. if a set $\mathcal{S}$ of numbers contains 1 and also the successor of every number in $\mathcal{S}$, then every number is in $\mathcal{S}$

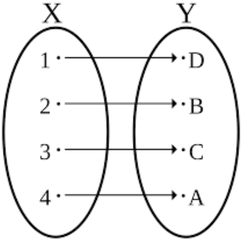
## Counting and Mathematics




- counting: matching to numbers

$$\{\star, \bullet, \diamond\} \equiv \{1, 2, 3\}$$

# Counting and Mathematics




- counting: matching to numbers

$$\{\star, \bullet, \diamond\} \equiv \{1, 2, 3\}$$

A set of objects has size $n$ if there exists a *one-to-one map* between the objects of this set and the set of natural numbers from 1 to $n$.

# Outline

# Counting and Combinatorics

- Combinatorics is *the study of finite or countable discrete structures*

# Counting and Combinatorics

- Combinatorics is *the study of finite or countable discrete structures*
- Two basic types of questions
  - Do there exist structures of a given kind and size?

# Counting and Combinatorics

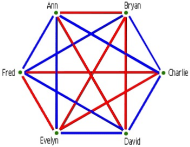
- Combinatorics is *the study of finite or countable discrete structures*
- Two basic types of questions
  - Do there exist structures of a given kind and size?

    *Given any six members of* `linkedin.com`*, does there exist a collection of three of them who are either all connected to each other or don't share any connections?*

# Counting and Combinatorics

- Combinatorics is *the study of finite or countable discrete structures*
- Two basic types of questions
  - Do there exist structures of a given kind and size?

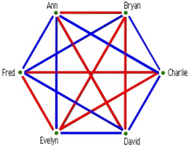
*Given any six members of* `linkedin.com`*, does there exist a collection of three of them who are either all connected to each other or don't share any connections?* Yes.



# Counting and Combinatorics

- Combinatorics is *the study of finite or countable discrete structures*
- Two basic types of questions
  - Do there exist structures of a given kind and size?

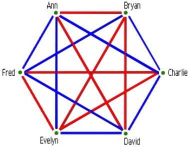
    *Given any six members of* `linkedin.com`*, does there exist a collection of three of them who are either all connected to each other or don't share any connections?* Yes.

    
    
  - How many structures of a given kind and size are there?

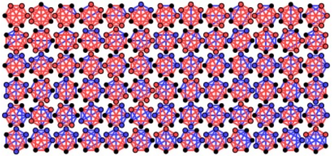
# Counting and Combinatorics

- Combinatorics is *the study of finite or countable discrete structures*
- Two basic types of questions
  - Do there exist structures of a given kind and size?

    *Given any six members of* `linkedin.com`, *does there exist a collection of three of them who are either all connected to each other or don't share any connections?* Yes.

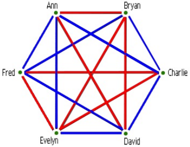
  - How many structures of a given kind and size are there?

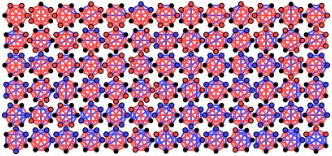
# Counting and Combinatorics

- Combinatorics is *the study of finite or countable discrete structures*
- Two basic types of questions
  - Do there exist structures of a given kind and size?

    *Given any six members of* `linkedin.com`*, does there exist a collection of three of them who are either all connected to each other or don't share any connections?* Yes.

  - How many structures of a given kind and size are there?

    There are $2^{15} = 32768$ different friends-strangers graphs on six labelled vertices (ignoring labels and change of colour, one gets 78 different graphs)

# Outline

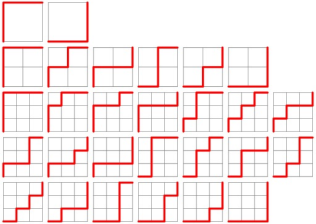
# Outline

# Lattice Paths and Counting Functions

- How many directed lattice paths with $n$ up-steps and $n$ east-steps are there?

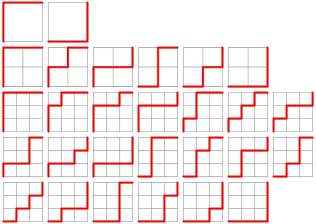
# Lattice Paths and Counting Functions

- How many directed lattice paths with $n$ up-steps and $n$ east-steps are there?



# Lattice Paths and Counting Functions

- How many directed lattice paths with $n$ up-steps and $n$ east-steps are there?



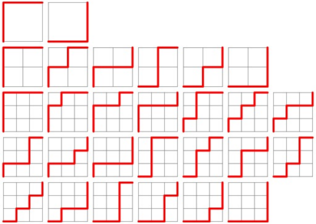
2 paths of length 2,

6 paths of length 4,

20 paths of length 6, ...

# Lattice Paths and Counting Functions

- How many directed lattice paths with $n$ up-steps and $n$ east-steps are there?



2 paths of length 2,

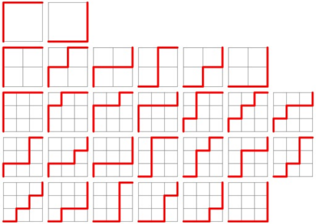
6 paths of length 4,

20 paths of length 6, ...

- The number of $2n$-step paths is $c_n = \binom{2n}{n}$

# Lattice Paths and Counting Functions

- How many directed lattice paths with $n$ up-steps and $n$ east-steps are there?



2 paths of length 2,

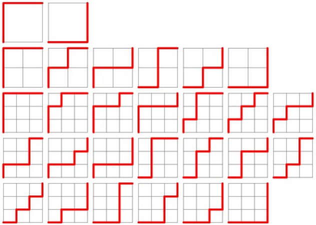
6 paths of length 4,

20 paths of length 6, ...

- The number of $2n$-step paths is $c_n = \binom{2n}{n}$
- The counting function $Z(x) = c_0 + c_1 x + c_2 x^2 + \ldots$

# Lattice Paths and Counting Functions

- How many directed lattice paths with $n$ up-steps and $n$ east-steps are there?



2 paths of length 2,

6 paths of length 4,

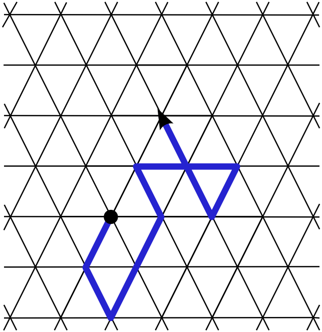
20 paths of length 6, ...

- The number of $2n$-step paths is $c_n = \binom{2n}{n}$
- The counting function $Z(x) = c_0 + c_1 x + c_2 x^2 + \ldots$ is

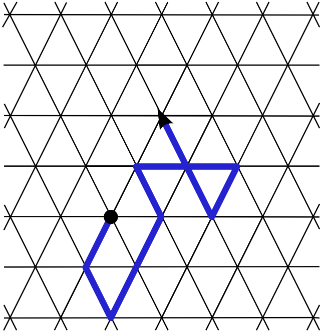
$$Z(x) = \frac{1}{\sqrt{1-4x}}$$

# Outline

# Walks on the Triangular Lattice

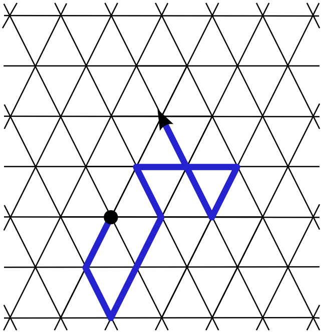


# Walks on the Triangular Lattice



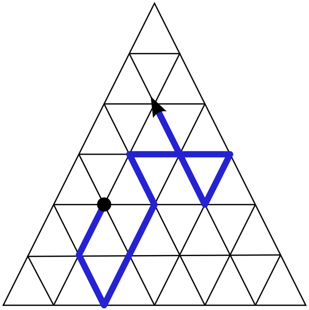
$c_n = 6^n$

# Walks on the Triangular Lattice



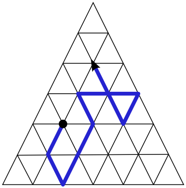
$$c_n = 6^n \,, \quad Z(x) = 1/(1-6x)$$

# Walks in a Triangle



Restrict to a triangular domain

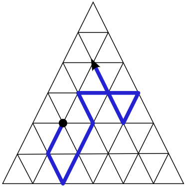
# Counting Walks



Parameters

- Side-length $L$
- Number of steps $n$
- Starting point $\mathbf{a}$
- End point $\mathbf{b}$

# Counting Walks



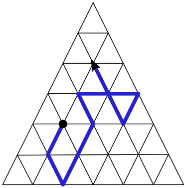
Parameters

- Side-length $L$
- Number of steps $n$
- Starting point $\mathbf{a}$
- End point $\mathbf{b}$

Number of $n$-step walks from $\mathbf{a}$ to $\mathbf{b}$ within triangle of side-length $L$

$$c_{n,L}^{\mathbf{a},\mathbf{b}}$$

# Counting Walks



Parameters

- Side-length $L$
- Number of steps $n$
- Starting point $\mathbf{a}$
- End point $\mathbf{b}$

Number of $n$-step walks from **a** to **b** within triangle of side-length $L$

$$c_{n,L}^{\mathbf{a},\mathbf{b}}$$

No general closed form known for $c_{n,L}^{\mathbf{a},\mathbf{b}}$ or associated counting function

$$Z_L^{\mathbf{a},\mathbf{b}}(t) = \sum_n c_{n,L}^{\mathbf{a},\mathbf{b}} t^n$$

# A Special Case

Starting the walks in a corner of the triangle, we find

**Theorem (Mortimer, Prellberg, 2015)**

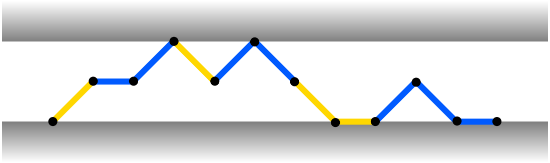
The counting function which counts $n$-step walks in a triangle of side-length $L$ starting at a chosen corner with no restrictions on the endpoint is given by

$$\frac{(1-p^3)(1-p^{1+L})}{(1-p)(1-p^{3+L})}$$

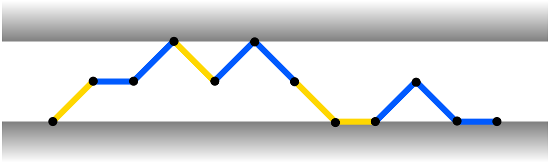
where

$$p=\frac{1-2t-\sqrt{(1+2t)(1-6t)}}{4t}$$

# Bi-Colored Motzkin Paths



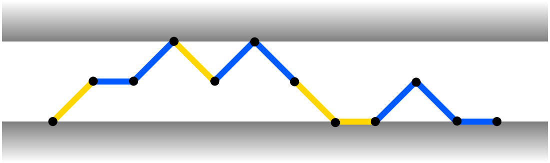
# Bi-Colored Motzkin Paths



**Corollary (Mortimer, Prellberg, 2015)**

$n$-step walks starting in a corner of a triangle of odd side-length $L = 2H + 1$ with arbitrary endpoint are in one-to-one correspondence with bi-colored $n$-step Motzkin paths in a strip of height $H$.

# Bi-Colored Motzkin Paths



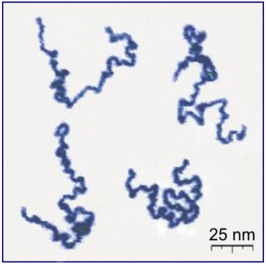
**Corollary (Mortimer, Prellberg, 2015)**

$n$-step walks starting in a corner of a triangle of odd side-length $L = 2H + 1$ with arbitrary endpoint are in one-to-one correspondence with bi-colored $n$-step Motzkin paths in a strip of height $H$.

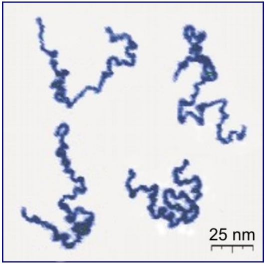
There is only a counting function proof, and no direct mapping is known.

# Outline

# Lattice Path Models of Polymers



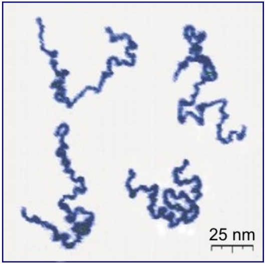
# Lattice Path Models of Polymers




## Lattice Paths

- Physical space $\rightarrow$ cubic lattice
- Ghost polymer $\rightarrow$ random walk

# Lattice Path Models of Polymers



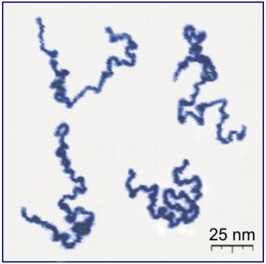

## Lattice Paths

- Physical space $\rightarrow$ cubic lattice
- Ghost polymer $\rightarrow$ random walk

## Self-Avoiding Walks (SAW)

- Polymer with Excluded Volume $\rightarrow$ self-avoiding random walk

# Lattice Path Models of Polymers




### Lattice Paths

- Physical space $\rightarrow$ cubic lattice
- Ghost polymer $\rightarrow$ random walk

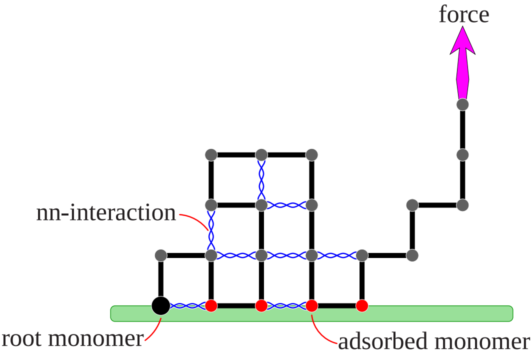
### Self-Avoiding Walks (SAW)

- Polymer with Excluded Volume $\rightarrow$ self-avoiding random walk

### Interacting Self-Avoiding Walks (ISAW)

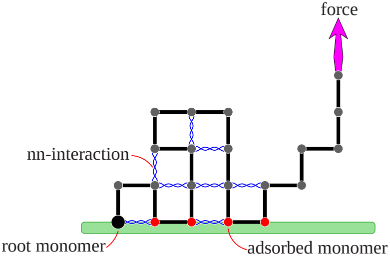
- Quality of solvent $\rightarrow$ interactions
- Model for the collapse of polymers

# “Realistic” Lattice Models of Polymers



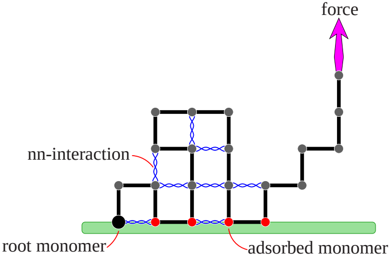

A self-avoiding walk lattice model of an interacting polymer tethered to a sticky surface under the influence of a pulling force

# Counting and Density of States




- Combinatorial question: How many $n$-step lattice paths are there with $m$ nearest-neighbour interactions, $k$ contacts with the surface, and ending at distance $h$ from the surface?

# Counting and Density of States



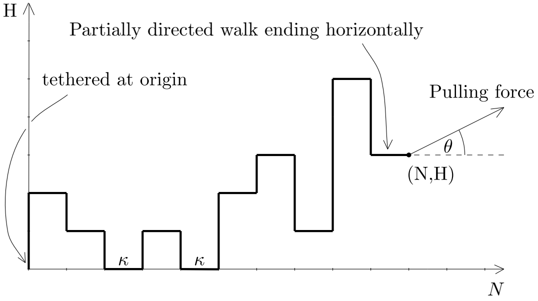

- Combinatorial question: How many $n$-step lattice paths are there with $m$ nearest-neighbour interactions, $k$ contacts with the surface, and ending at distance $h$ from the surface?

- Physicists relate this to the Density of States and can extract from this lots of interesting thermodynamic information

# Outline

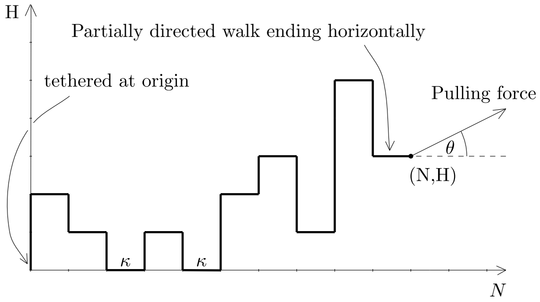
# Pulling Polymers off a Surface

- A partially directed walk model of a polymer tethered to a sticky surface under the influence of a pulling force



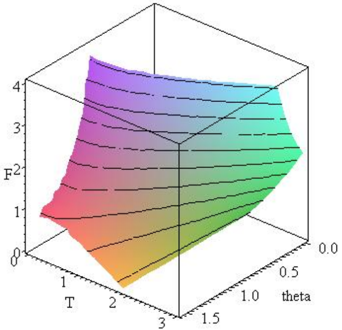
# Pulling Polymers off a Surface

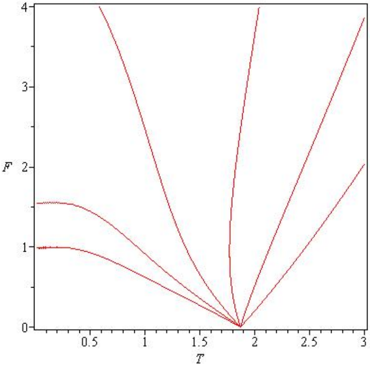
- A partially directed walk model of a polymer tethered to a sticky surface under the influence of a pulling force




- This model is exactly solvable (Osborn, Prellberg, 2010)

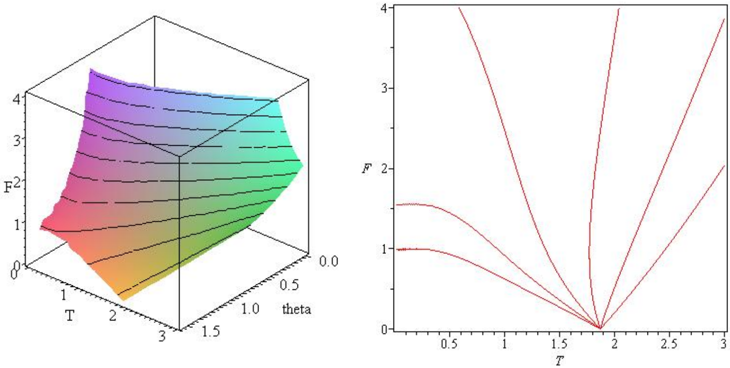
# Varying the Pulling Angle



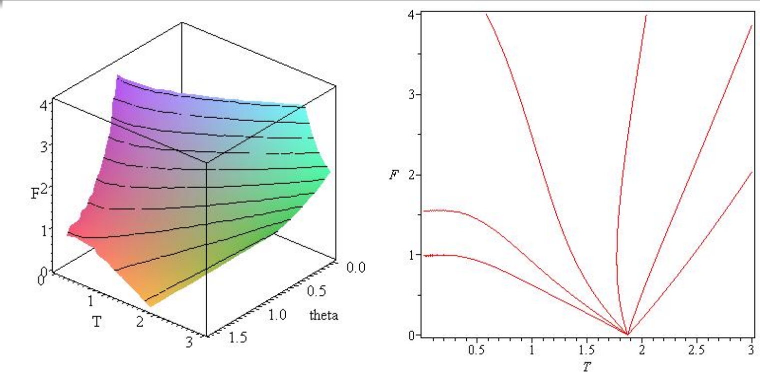

- Thermal desorption at $T = 1/\log(1+\sqrt{2}/2) \approx 1.87$

# Varying the Pulling Angle



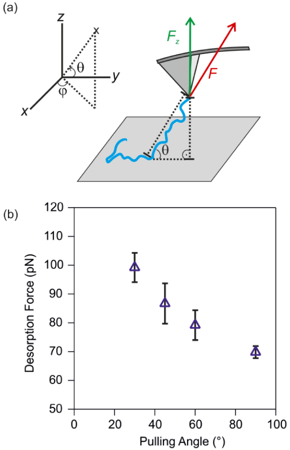

- Thermal desorption at $T = 1/\log(1+\sqrt{2}/2) \approx 1.87$
- Vertical pulling, $\theta = 90^o$ (left curve):
  Increasing $F$ favours desorption

# Varying the Pulling Angle




- Thermal desorption at $T = 1/\log(1+\sqrt{2}/2) \approx 1.87$
- Vertical pulling, $\theta = 90^o$ (left curve):
  Increasing $F$ favours desorption
- Horizontal pulling, $\theta = 0^o$ (right curve):
  Increasing $F$ disfavours desorption

# Varying the Pulling Angle




REVIEW OF SCIENTIFIC INSTRUMENTS **88**, 033705 (2017)

## Pulling angle-dependent force microscopy

L. Grebíková, H. Gojzewski, B. D. Kieviet,[a] M. Klein Gunnewiek, and G. J. Vancso[b]
*Materials Science and Technology of Polymers, MESA+, Institute of Nanotechnology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands*

(Received 22 November 2016; accepted 28 February 2017; published online 20 March 2017)

In this paper, we describe a method allowing one to perform three-dimensional displacement control in force spectroscopy by atomic force microscopy (AFM). Traditionally, AFM force curves are measured in the normal direction of the contacted surface. The method described can be employed to address not only the magnitude of the measured force but also its direction. We demonstrate the technique using a case study of angle-dependent desorption of a single poly(2-hydroxyethyl methacrylate) (PHEMA) chain from a planar silica surface in an aqueous solution. The chains were end-grafted from the AFM tip in high dilution, enabling single macromolecule pull experiments. Our experiments give evidence of angular dependence of the desorption force of single polymer chains and illustrate the added value of introducing force direction control in AFM. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4978452]

### IV. CONCLUSIONS

We describe in this study the design and implementation of performing directional force spectroscopy experiments by AFM. By modifying the built-in functions of a standard AFM instrument, we were able to control the cantilever trajectory. This approach has been demonstrated by a case study on the angle-dependent desorption of an end-grafted polymer chain. The polymer response to the external force exerted at various pulling angles with respect to the substrate has been addressed. By employing this technique in single molecule force spectroscopy experiments, we obtained experimental evidence for theoretically predicted enhancement of polymer adhesion while decreasing the pulling angle with respect to planar substrate surfaces.[18,33]

# Outline

# Outline

# Simple Random Walk in One Dimension

Galton Board

# Simple Random Walk in One Dimension

Galton Board

- Start at origin and go to left or right with equal probability (fair coin-toss)
- $2^n$ possible random walks with $n$ steps
- Endpoint position follows binomial distribution
- Trajectories are directed lattice paths

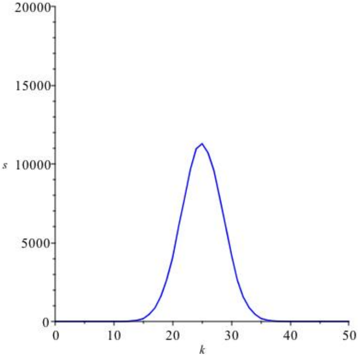
# Simple Random Walk in One Dimension

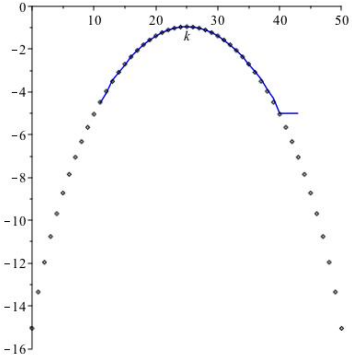
Galton Board

- Start at origin and go to left or right with equal probability (fair coin-toss)
- $2^n$ possible random walks with $n$ steps
- Endpoint position follows binomial distribution
- Trajectories are directed lattice paths

Model of a directed polymer in two dimensions

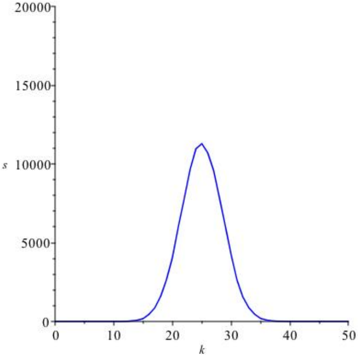
# Simple Sampling

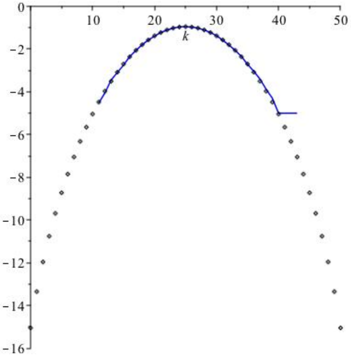



Simple sampling of simple random walk for $n = 50$ steps. For each simulation, 100000 samples were generated.

# Simple Sampling




Simple sampling of simple random walk for $n = 50$ steps. For each simulation, 100000 samples were generated.

How can we tweak the algorithm to reach the tails?

# Pruned and Enriched Sampling

- Smart idea: change sampling rate to achieve uniform sampling

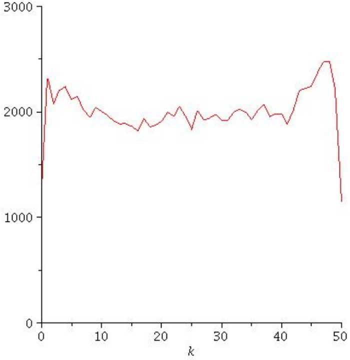
# Pruned and Enriched Sampling

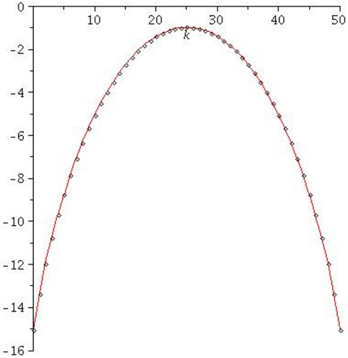
- Smart idea: change sampling rate to achieve uniform sampling

## Pruning and Enrichment Strategy

- **Pruning** If sampling rate is too large, remove the configuration probabilistically
- **Enrichment** If sampling rate is too small, make several copies of the configuration and continue growing each

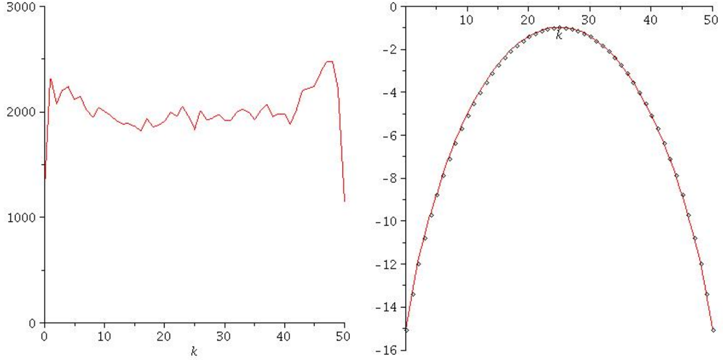
# Pruned and Enriched Sampling




- Uniform sampling with genuinely blind algorithm

# Pruned and Enriched Sampling




- Uniform sampling with genuinely blind algorithm

Can be applied to a large class of growth processes

# Outline

# Simple Sampling of Self-Avoiding Walk

Consider Self-Avoiding Walks (SAW) on the square lattice $\mathbb{Z}^2$

- Simple sampling of SAW works like simple sampling of random walks
- But now walks get removed if they self-intersect

# Simple Sampling of Self-Avoiding Walk

Consider Self-Avoiding Walks (SAW) on the square lattice $\mathbb{Z}^2$

- Simple sampling of SAW works like simple sampling of random walks
- But now walks get removed if they self-intersect

Generating SAW with simple sampling is very inefficient

- There are $4^n$ $n$-step random walks, but only about $2.638^n$ $n$-step SAW

# Simple Sampling of Self-Avoiding Walk

Consider Self-Avoiding Walks (SAW) on the square lattice $\mathbb{Z}^2$

- Simple sampling of SAW works like simple sampling of random walks
- But now walks get removed if they self-intersect

**Generating SAW with simple sampling is very inefficient**

- There are $4^n$ $n$-step random walks, but only about $2.638^n$ $n$-step SAW
- The probability of successfully generating an $n$-step SAW decreases exponentially fast

This is called exponential attrition

# Algorithm Development

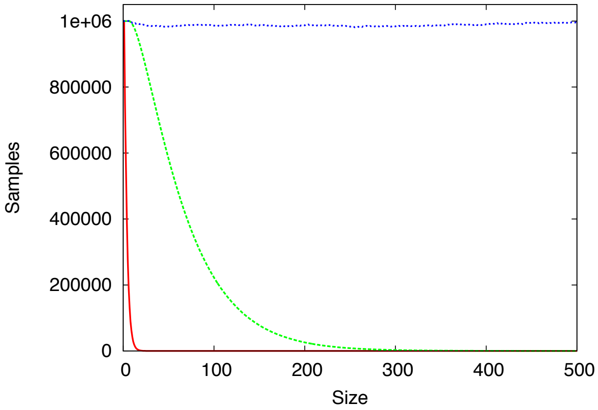
- Rosenbluth Method (Rosenbluth, Rosenbluth, 1956):
  Only take steps that avoid intersections

# Algorithm Development

- Rosenbluth Method (Rosenbluth, Rosenbluth, 1956):
  Only take steps that avoid intersections
- PERM (Grassberger, 1997):
  Add Pruning and Enrichment to Rosenbluth Method

# Algorithm Development

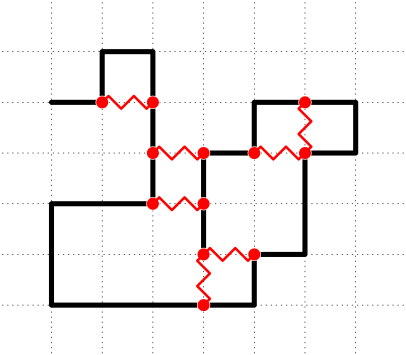
- Rosenbluth Method (Rosenbluth, Rosenbluth, 1956):
  Only take steps that avoid intersections
- PERM (Grassberger, 1997):
  Add Pruning and Enrichment to Rosenbluth Method
- FlatPERM (Prellberg, Krawczyk, 2004):
  Add Uniform Sampling Strategy to PERM



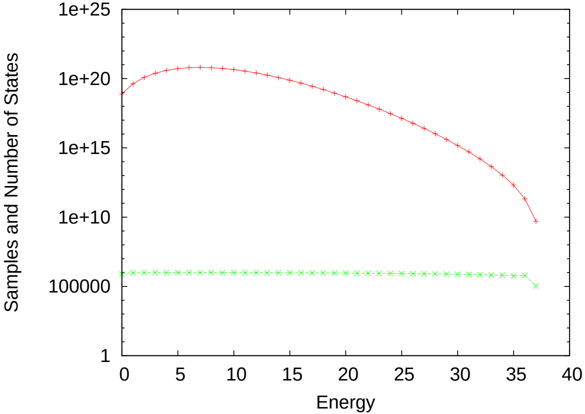

Attrition for Simple Sampling, Rosenbluth Sampling, and PERM

# Interacting Self-Avoiding Walks

Consider sampling with respect to an extra parameter, for example the number of nearest-neighbour contacts



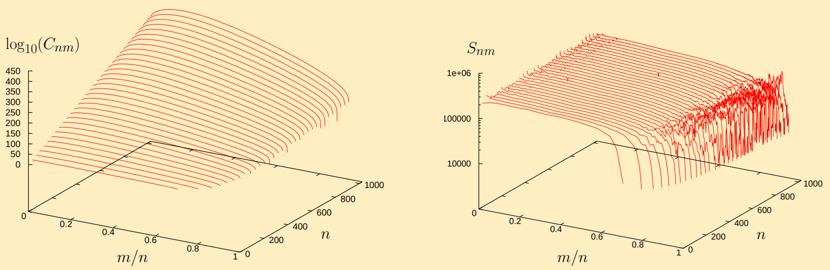
An interacting self-avoiding walk on the square lattice with $n = 26$ steps and $m = 7$ contacts.



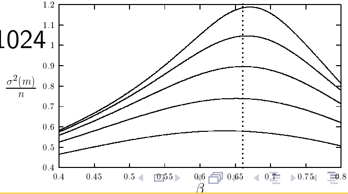

Generated samples and estimated number of states for ISAW with 50 steps estimated from $10^6$ flatPERM tours.

# ISAW simulations

Prellberg, Krawczyk, 2004




- Square lattice ISAW up to $n = 1024$
- One simulation suffices
- 400 orders of magnitude



# Outline

# A Pure Mathematics Application

In homage to last week's Mathematics Colloquium:

"On the cogrowth of Thompson's $F$ group"
Rechnitzer, Elder, Wong (2012)

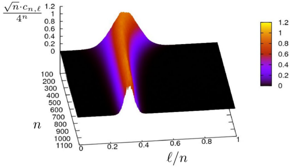



Figure 3: A plot of the normalised distribution of the number of words $c_{n,\ell}$ of length $n$ and geodesic length $\ell$ in Thompson's group $F$. Notice that the peak position is quite stable, indicating that the mean geodesic length grows roughly linearly with word length.

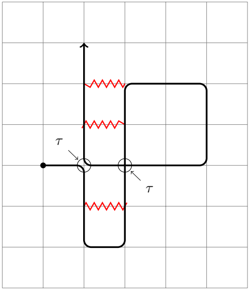
We will proceed along a similar line but using a more powerful random sampling method based on flat-histogram ideas used in the FlatPERM algorithm [18, 19]. Each sample word is grown in a similar manner to simple sampling — append one generator at a time chosen uniformly at random. The weight of a word of $n$ symbols is simply 1, so that the total weight of all possible words at any given length is just $4^n$. As the word grows we keep track of its geodesic length. We now deviate from simple sampling by "pruning" and "enriching" the words.

The mean geodesic length of the amenable groups studied grow sublinearly, while those of $\mathbb{Z} \wr F_2$ and Thompson's group are observed to grow linearly. Using simple sampling we estimate that the mean geodesic length of Thompson's group does indeed grow linearly and that the rate of escape is $0.27 \pm 0.01$.

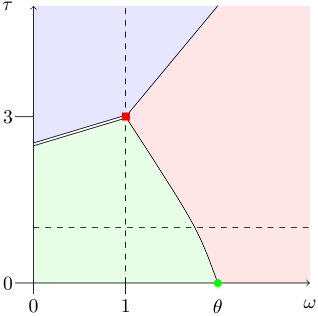
Indication that Thompson's $F$ group is not amenable[1]

[1] does not have a finitely-additive left-invariant probability measure

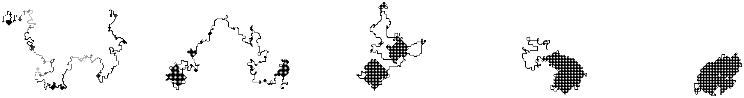
# Lattice polymers with two competing interactions




Bedini, Owczarek, Prellberg (2014)

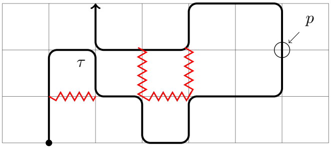


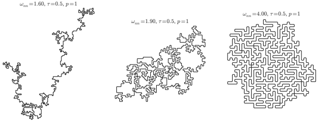

$\omega = 0.5$, increasing $\tau$:

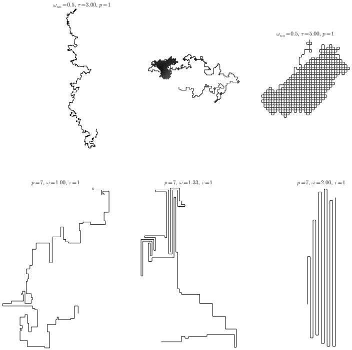


# Self-attracting polymers in two dimensions with three low-temperature phases

Bedini, Owczarek, Prellberg (2017)

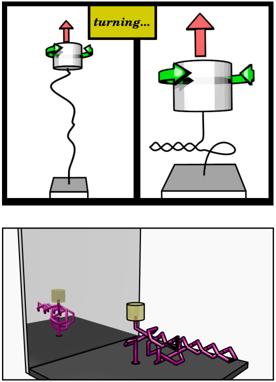


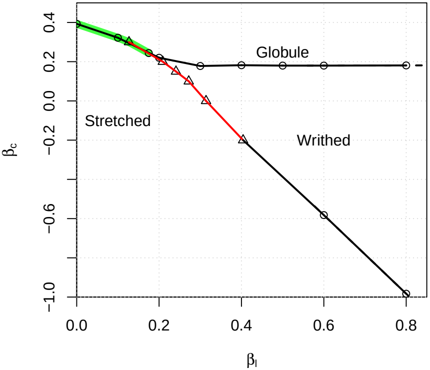



# Supercoiling in a lattice polymer

Dagrosa, Owczarek, Prellberg (preprint)



# A big THANK YOU to all co-authors

R. F. Bailey, O. F. Bandtlow, J. Bastian, A. Bedini, D. Bennett-Wood, S. Berg, R. Brak, K. M. Briggs, P. J. Cameron, S. Corteel, E. Dagrosa, W. Dantas, J. Doukas, B. Drossel, P. Dyke, J. Fiala, A. J. Guttmann, H. Harborth, N. Haug, G. Iliev, E. J. Janse van Rensburg, D. Johannsen, D. Johannsmann, M. Josuat-Vergès, W. Just, P. Kleban, J. Krawczyk, J. Lee, M. C. Marchetti, A. A. Middleton, A. Narros, P. R. G. Mortimer, P. Oertel, A. Olde Daalhuis, T. J. Oliveira, J. Osborn, A. L. Owczarek, A. Rechnitzer, M. Rubey, P. Schweitzer, G. Siudem, J. Slawny, L. Song, D. Stanton, D. Stark, J. Stilck, H. Touchette, C. Stump, S. G. Whittington