Exact solution of some friendly directed walker problems

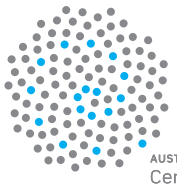
[†] Aleks Owczarek, [‡] Andrew Rechnitzer, [†] Rami Tabbara and [‡] Thomas Wong

[†] MASCOS and Department of Mathematics and Statistics, The University of Melbourne

[‡] Department of Mathematics and Statistics, University of British Columbia

January, 2013

Workshop on Random Polymers, EURANDOM



AUSTRALIAN RESEARCH COUNCIL
Centre of Excellence for Mathematics and Statistics of Complex Systems

THE UNIVERSITY OF MELBOURNE

DIRECTED WALKS

- Exact solutions of single and multiple directed walk models
- Recurrence and functional equation
- Rational, algebraic, Differentially-finite and non D-finite solutions
- Multiple walks: Bethe Ansatz & Lindström-Gessel-Viennot
- LGV Lemma: multiple walks = determinant of single walks
- LGV problems result in generating functions that are D-finite

SOLVING FUNCTIONAL EQUATIONS

- Functional equation for an expanded generating function
- Uses an extra catalytic variable
- Answer is a 'boundary' value
- Fix catalytic variable $\rightarrow$ 'bulk' term disappears (Kernel method)
- Obstinate kernel method: multiple values of catalytic variable
- See Bousquet-Mélou *Math. and Comp. Sci 2* (2002)), Bousquet-Mélou, Mishna *Contemp. Math.* **520** (2010)
- Solutions are not always D-finite

Polymer Adsorption

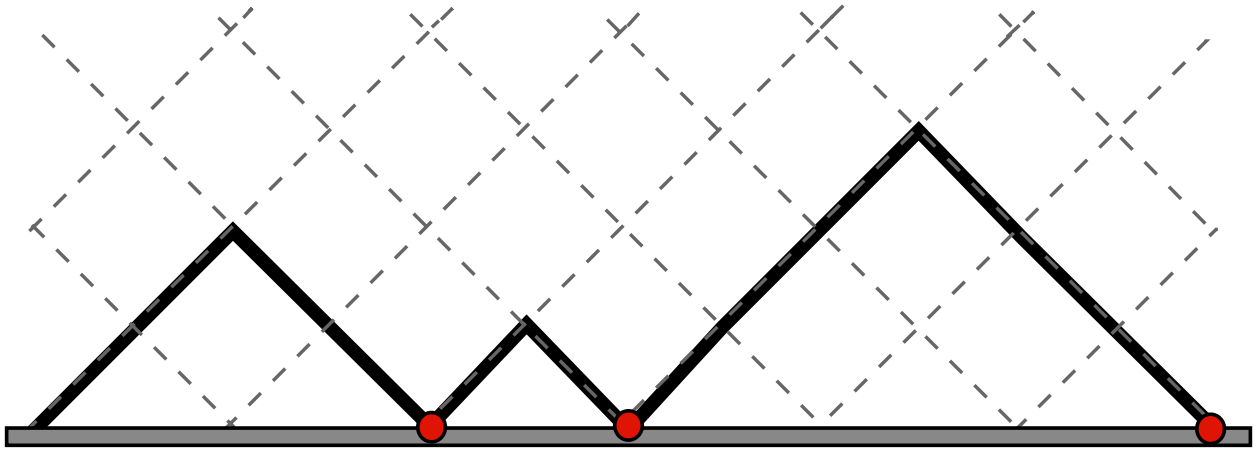
The physical motivation is the adsorption phase transition

- Second order phase transition with jump in specific heat
- Crossover exponent $\phi = 1/2$ for directed walks and SAW
- Order parameter is density of visits to surface by the polymer

ADSORPTION: ONE DIRECTED WALK

Exact solution and analysis of single and multiple directed walk models exist

- Single Dyck path in a half space
- Energy $-\varepsilon_a$ for each time (number m_a) it visits the surface
- Boltzmann weight $a = e^{\varepsilon_a / k_B T}$



ADSORPTION: ONE DIRECTED WALK

A complete solution exists and the generating function is algebraic

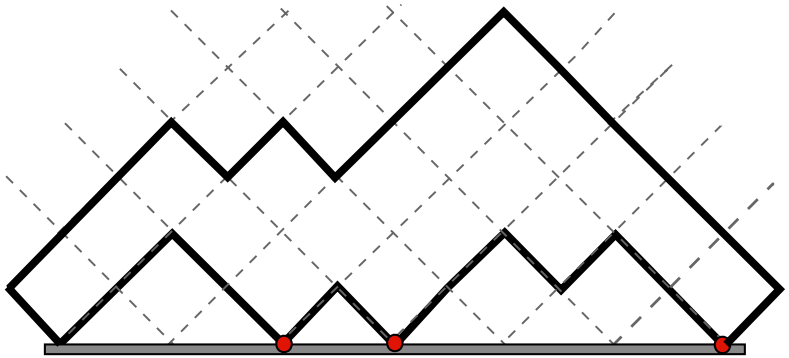
Consider the density of visits

$$\mathcal{A} = \lim_{n\to\infty} \frac{\langle m_a \rangle}{n}$$

There exists a phase transition at a temperature T_a given by $a = 2$:

- For $T > T_a$ the walk moves away entropically and $\mathcal{A} = 0$
- For $T < T_a$ the walk is adsorbed onto the surface and $\mathcal{A} > 0$

VESICLE ADSORPTION



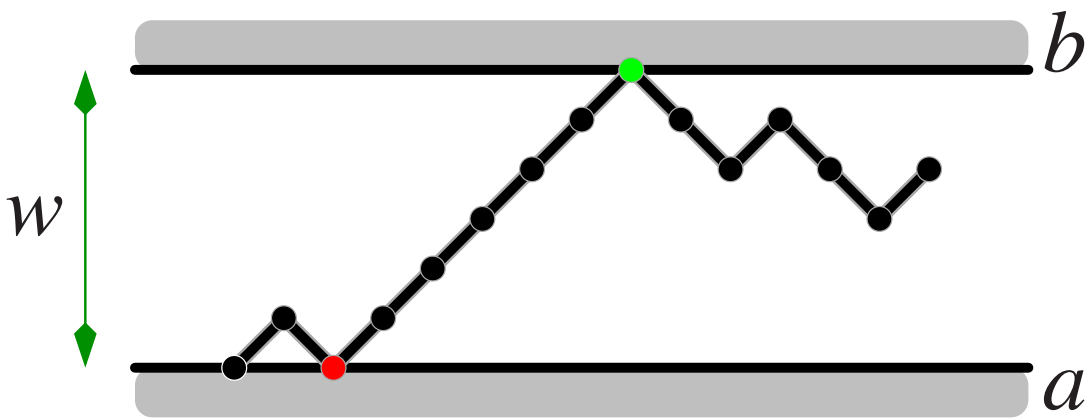
- Exact solution of two directed walks joined making a simple "vesicle" (Brak, Essam, Owczarek , *J. Stat. Phys.* **93**, 155 (1998))
- Vesicles with interactions for visits of the *bottom* walk to height 0 and height 1 (Lonsdale, Brak, Essam, Owczarek, Rechnitzer, *J. Phys. A.: Math. and Theor.* **42** 1, (2009).)

Single second order transition — similar to the single walk adsorption transition

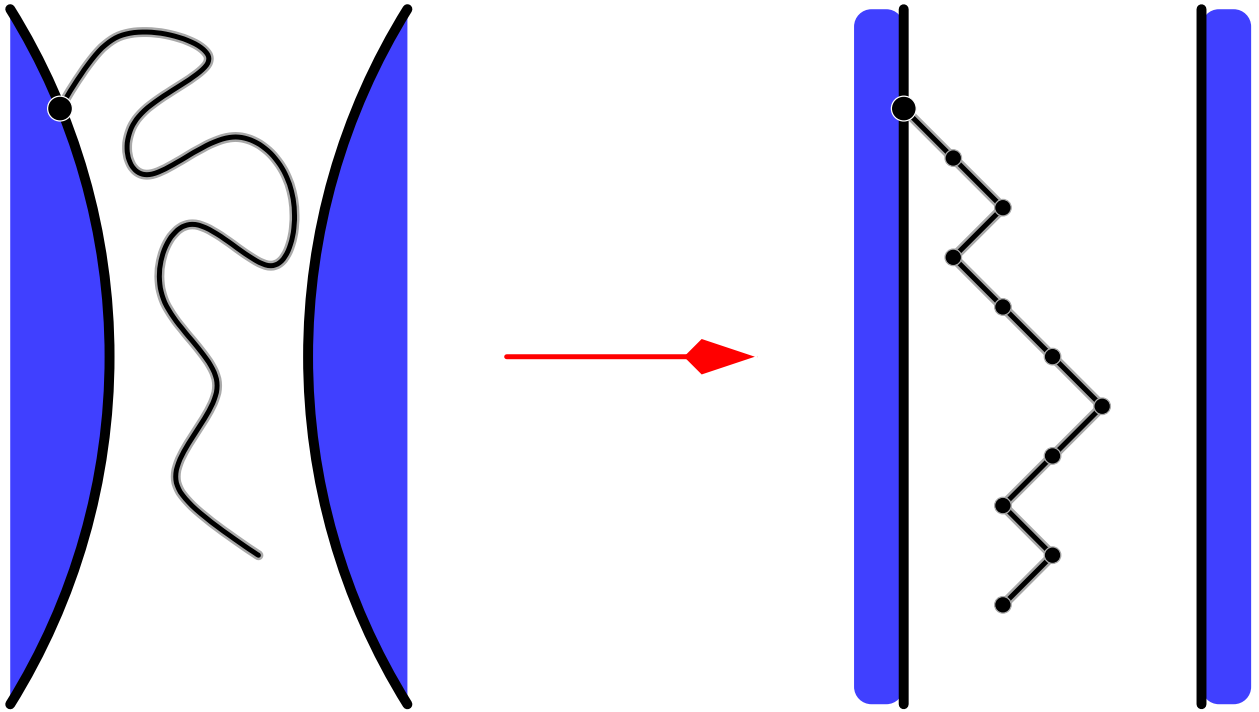
More motivation: polymers in a slit

Paths in a slit

- A directed path of length n in a slit of width w.
- Boltzmann weights a and b for interactions with lines (walls).
- Brak, Owczarek, Rechnitzer and Whittington *J. Phys. A*, **38** (2005) 4309–4325
- Subsequent works on undirected SAW

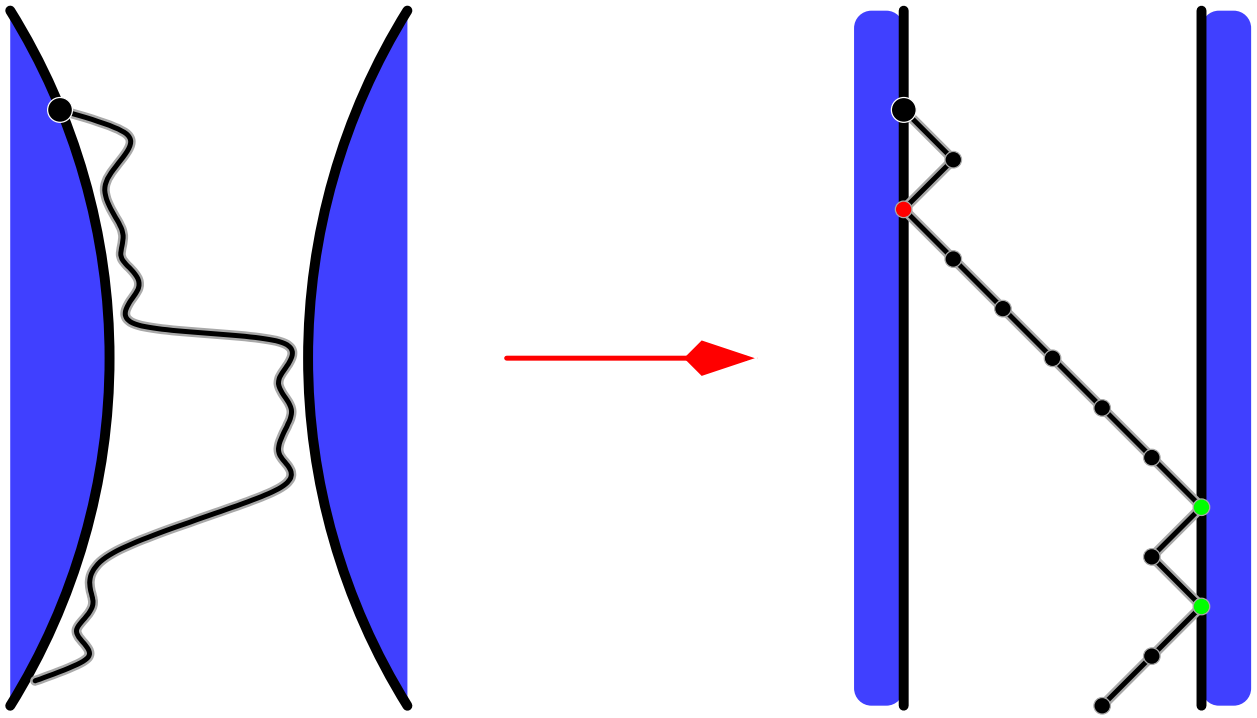


WHAT IS STERIC STABILISATION?



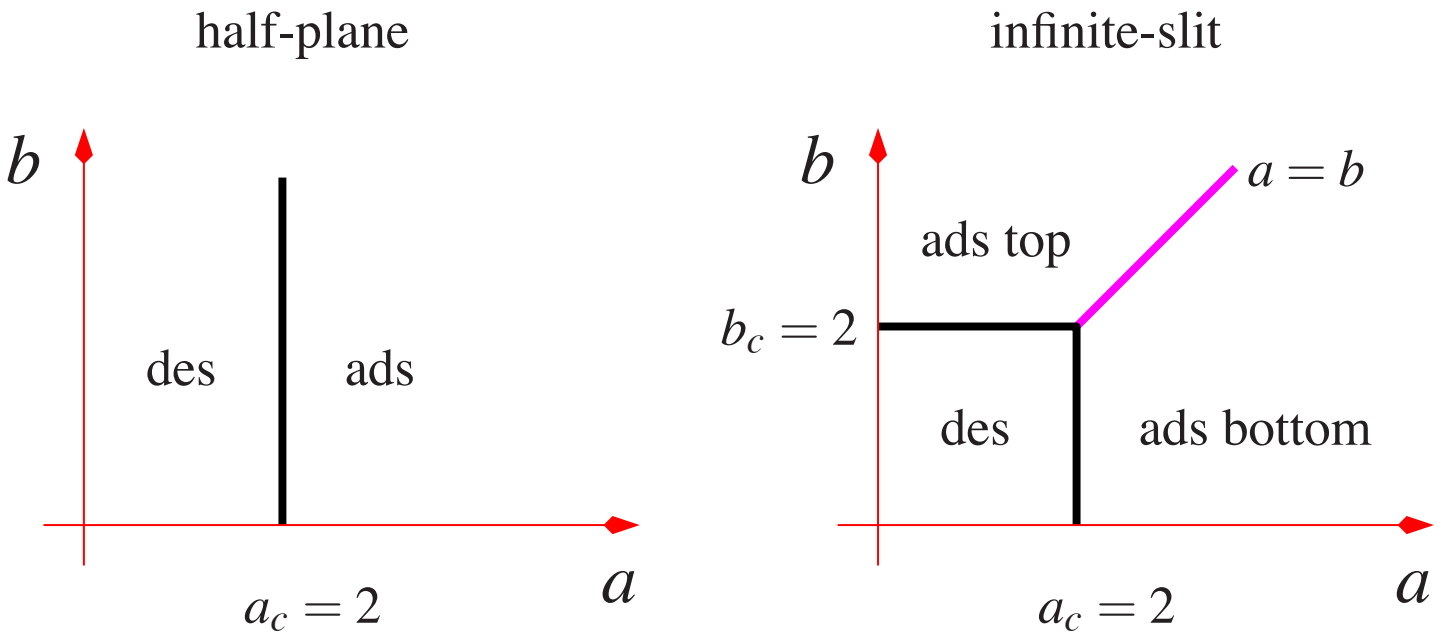
- Entropic repulsion between the colloidal particles.

WHAT IS SENSITISED FLOCCULATION?



- Polymer adsorbs and pulls the particles together.

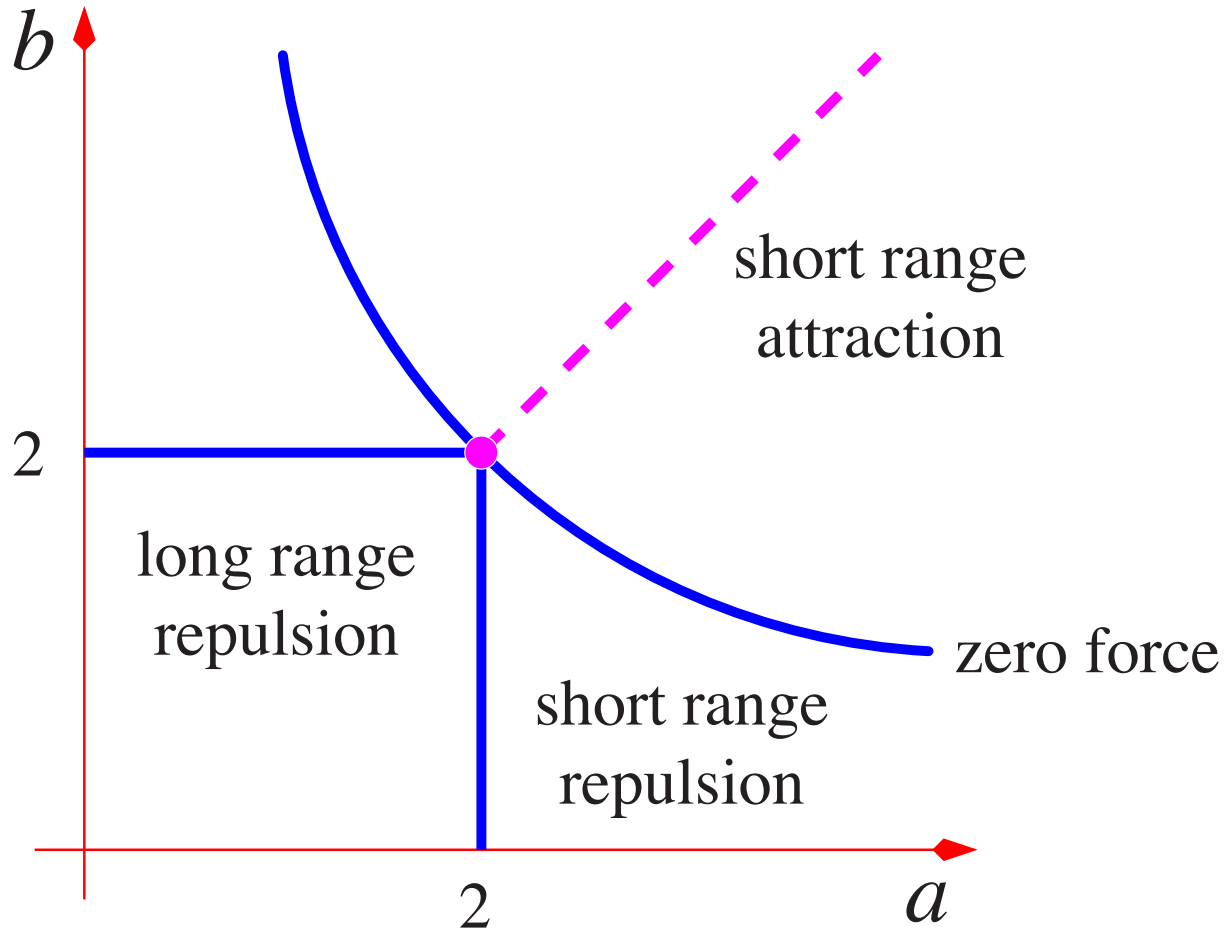
THE PHASE DIAGRAM POLYMERS IN A SLIT



Phase diagram depends on order of limits

- Black = 2^{nd} order and pink = 1^{st} order.
- Left-hand diagram: $w \to \infty$ and then $n \to \infty$.
- Right-hand diagram: $n \to \infty$ and then $w \to \infty$.
- The order of n and w limits really does matter!

More motivation: polymers in a slit



RING POLYMERS IN A SLIT

- Our motivation is a recent Monte Carlo study of ring polymers in a slit
- Here both sides of the polygon interact with the surfaces of the slit Alvarez, Janse van Rensburg, Soteros and Whittington *J. Phys. A.: Math. and Theor.* **41**, 185004 (2008)
- Novel zero force curve

(Our Model)

Directed vesicle where both walks can interact with a single surface

Model

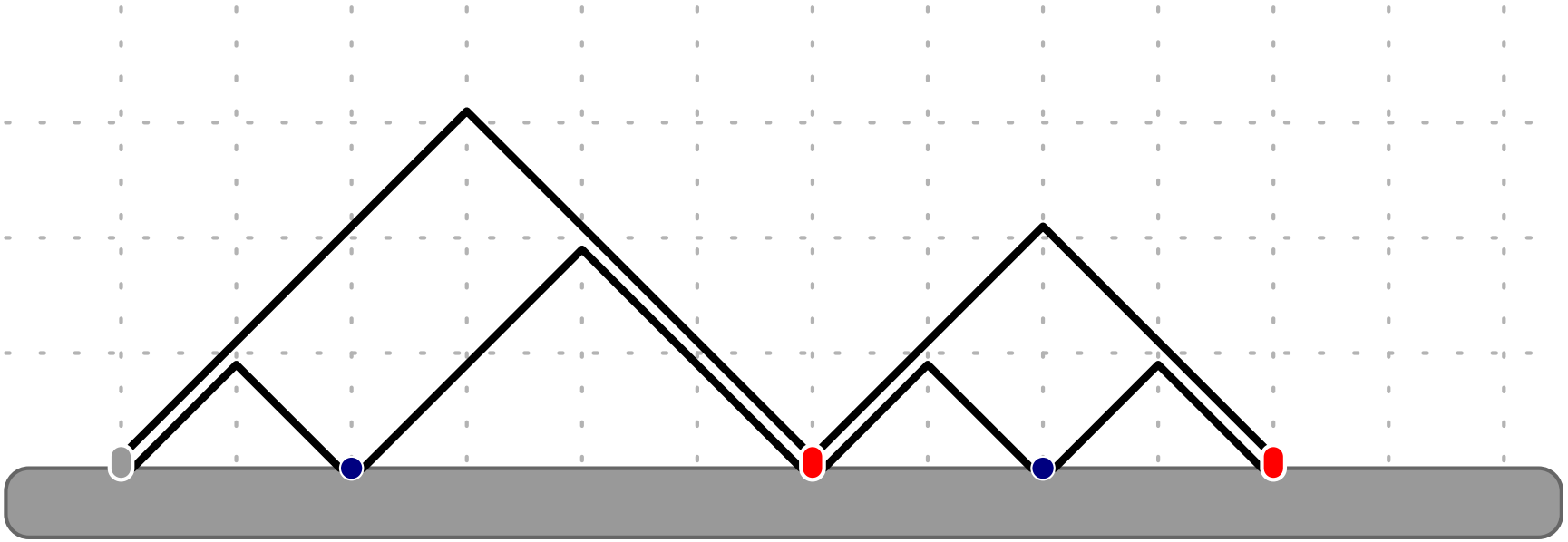


Figure: Two directed walks with single and "double" visits to the the surface.

- *energy* $-\varepsilon_a$ *for visits of the bottom walk only (single visits) to the wall,*
- *energy* $-\varepsilon_d$ *when both walks visit a site on the wall (double visits)*

MODEL

- number of *single visits* to the wall will be denoted m_a,
- number of *double visits* will be denoted m_d.

The partition function:

$$Z_n(a,d) = \sum_{\widehat{\varphi} \ni |\widehat{\varphi}|=n} e^{(m_a(\widehat{\varphi})\cdot\varepsilon_a + m_d(\widehat{\varphi})\cdot\varepsilon_d)/k_BT}$$

where $a = e^{\varepsilon_a/k_BT}$ and $d = e^{\varepsilon_d/k_BT}$.

The thermodynamic reduced free energy:

$$\kappa(a,d) = \lim_{n\to\infty} n^{-1} \log\left(Z_n(a,d)\right).$$

Generating Function

To find the free energy we will instead solve for the generating function

$$G(a,d;z) = \sum_{n=0}^{\infty} Z_n(a,d) z^n.$$

The radius of convergence of the generating function $z_c(a,d)$ is directly related to the free energy via

$$\kappa(a,d) = \log(z_c(a,d)^{-1}).$$

Two order parameters:

$$\mathcal{A}(a,d) = \lim_{n\to\infty} \frac{\langle m_a \rangle}{n} \qquad \text{and} \qquad \mathcal{D}(a,d) = \lim_{n\to\infty} \frac{\langle m_d \rangle}{n},$$

FUNCTIONAL EQUATION

We consider walks φ in the larger set, where each walk can end at any possible height.

The expanded generating function

$$F(r,s;z) \equiv F(r,s) = \sum_{\varphi\in\Omega} z^{|\varphi|} r^{\lfloor\varphi\rfloor} s^{\lceil\varphi\rceil/2} a^{m_a(\varphi)} d^{m_d(\varphi)},$$

where

- z is conjugate to the length $|\varphi|$ of the walk,
- r is conjugate to the distance $\lfloor\varphi\rfloor$ of the bottom walk from the wall and
- s is conjugate to *half* the distance $\lceil\varphi\rceil$ between the final vertices of the two walks.

and we recover $G(a,d;z) = F(0,0)$.

FUNCTIONAL EQUATION

Consider adding steps onto the ends of the two walks

This gives the following functional equation

$$\begin{aligned} F(r,s) =& 1 + z\left(r + \frac{1}{r} + \frac{s}{r} + \frac{r}{s}\right) \cdot F(r,s) \\ & - z\left(\frac{1}{r} + \frac{s}{r}\right) \cdot [r^0]F(r,s) - z\frac{r}{s} \cdot [s^0]F(r,s) \\ & + z(a-1)(1+s) \cdot [r^1]F(r,s) + z(d-a) \cdot [r^1 s^0]F(r,s). \end{aligned}$$

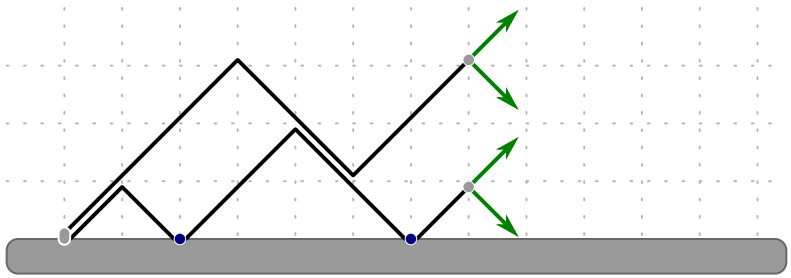


Figure: Adding steps to the walks when the walks are away from the wall.

The Kernel

Rewrite equation as "Bulk = Boundary"

$$K(r,s)\cdot F(r,s) = \frac{1}{d} + \left(1 - \frac{1}{a} - \frac{zs}{r} - \frac{z}{r}\right)\cdot F(0,s) - \frac{zr}{s}\cdot F(r,0) + \left(\frac{1}{a} - \frac{1}{d}\right)\cdot F(0,0)$$

where the *kernel* K is

$$K(r,s) = \left[1 - z\left(r + \frac{1}{r} + \frac{s}{r} + \frac{r}{s}\right)\right].$$

Recall, we want $F(0,0)$ so we try to find values that kill the kernel

Symmetries of the Kernel

The kernel is symmetric under the following two transformations:

$$(r,s) \mapsto \left(r, \frac{r^2}{s}\right), \qquad (r,s) \mapsto \left(\frac{s}{r}, s\right)$$

Transformations generate a family of 8 symmetries ('group of the walk')

$$(r,s), \left(r, \frac{r^2}{s}\right), \left(\frac{s}{r}, \frac{s}{r^2}\right), \left(\frac{r}{s}, \frac{1}{s}\right), \left(\frac{1}{r}, \frac{1}{s}\right), \left(\frac{1}{r}, \frac{s}{r^2}\right), \left(\frac{r}{s}, \frac{r^2}{s}\right), \text{ and } \left(\frac{s}{r}, s\right)$$

We make use of 4 of these which only involve positive powers of r.

This gives us four equations.

Magic combination

Following Bousquet-Mélou when $a = 1$ we form the simple alternating sum

$$Eqn1 - Eqn\ 2 + Eqn\ 3 - Eqn\ 4.$$

- When $a \neq 1$ one needs to generalise that approach
- Multiply by rational functions,

The form of the Left-hand side of the final equation being

$$a^2 r K(r,s)\left(sF(r,s) - \frac{r^2}{s}F\left(r, \frac{r^2}{s}\right) + \frac{Lr^2}{s^2}F\left(\frac{r}{s}, \frac{r^2}{s}\right) - \frac{L}{s^2}F\left(\frac{r}{s}, \frac{1}{s}\right)\right)$$

where

$$L = \frac{zas - ars + rs + zar^2}{zas - ar + r + zar^2}.$$

EXTRACTING THE SOLUTION $a = 1$

$$K(r,s) \cdot (\text{linear combination of } F) = \frac{r(s-1)(s^2+s+1-r^2)}{s^2}\left(1+(d-1)F(0,0)\right) - zd(1+s)sF(0,s) + \frac{zd(1+s)}{s^2}F\left(0,\frac{1}{s}\right).$$

- The kernel has two roots
- choose the one which gives a positive term power series expansion in z
- with Laurent polynomial coefficients in s:

$$\hat{r}(s;z) \equiv \hat{r} = \frac{s\left(1-\sqrt{1-4\frac{(1+s)^2z^2}{s}}\right)}{2(1+s)z} = \sum_{n\geq 0} C_n \frac{(1+s)^{2n+1}z^{2n+1}}{s^n},$$

where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is a Catalan number.

EXTRACTING THE SOLUTION $a = 1$

- *Make the substitution* $r \mapsto \hat{r}$
- *rewrite to remove* z*:* $z = (\hat{r} + 1/\hat{r} + \hat{r}/s + s/\hat{r})^{-1}$.

Setting $r \mapsto \hat{r}$ gives

$$0 = ds^4 F(0,s) - dsF\left(0, \frac{1}{s}\right) - (s-1)(s^2 + s + 1 - \hat{r}^2)(s + \hat{r}^2)\left(1 + (d-1)F(0,0)\right)$$

Note coefficients of $F(0,s)$ and $F(0,1/s)$ are independent of $\hat{r}$.

If we divide by equation by s *— then* $F(0,0)$ *is the constant term in* s.

SOLUTION FOR $a = 1$

Now extract the coefficient of s^1:

$$0 = -\left(1 + \sum_{n=0}^{\infty} \frac{12(2n+1)}{(n+2)^2(n+3)} C_n^2 z^{2n+2}\right) \cdot (1 + (d-1)F(0,0)) - d \cdot F(0,0).$$

Solving the above when $d = 1$ gives

$$G(1,1;z) = 1 + \sum_{n=0}^{\infty} \frac{12(2n+1)}{(n+2)^2(n+3)} C_n^2 z^{2n+2},$$

and hence for general d we have

$$F(0,0) = G(1,d;z) = \frac{G(1,1;z)}{d + (1-d)G(1,1;z)}.$$

$a = d$

Need to extract coefficients term by term in a to give

$$[a^k z^{2n}]F(0,0) = \sum_{k'=0}^{k} \frac{k'(k'+1)(2+4n-k'n-2k')}{(k'-1-n)(n+1)^2(-2n+k')(n+2)} \binom{2n-k'}{n} \binom{2n}{n}$$
$$= \frac{k(k+1)(k+2)}{(2n-k)(n+1)^2(n+2)} \binom{2n-k}{n} \binom{2n}{n}$$

which gives us

$$G(a,a) = \sum_{n\geq 0} z^{2n} \sum_{k=0}^{n} a^k \frac{k(k+1)(k+2)}{(n+1)^2(n+2)(2n-k)} \binom{2n}{n} \binom{2n-k}{n}.$$

Agrees with Brak *et al.* (1998) that used LGV

One can now consider $d \neq a$:

$$G(a,d;z) = \frac{aG(a,a;z)}{d+(a-d)G(a,a;z)}.$$

COMBINATORIAL STRUCTURE

- Combinatorial structure the underlying the functional equation.
- Breaking up our configurations into pieces between double visits gives

$$G(a,d;z) = \frac{1}{1 - dP(a;z)}$$

  where $P(a;z)$ is the generating function of so-called primitive factors.
- Rearranging this expression gives

$$P(a;z) = \frac{G(a,d;z) - 1}{dG(a,d;z)} = \frac{G(a,a;z) - 1}{aG(a,a;z)}.$$

- This allows us to calculate $P(a;z)$ from a known expression for $G(a,a;z)$

PHASES

The phases determined by dominant singularity of the generating function

The singularities of $G(a,d;z)$ are

- *those of $P(a;z)$ which are related to those of $G(a,a;z)$ and*
- *the simple pole at $1-dP(a;z)=0$.*

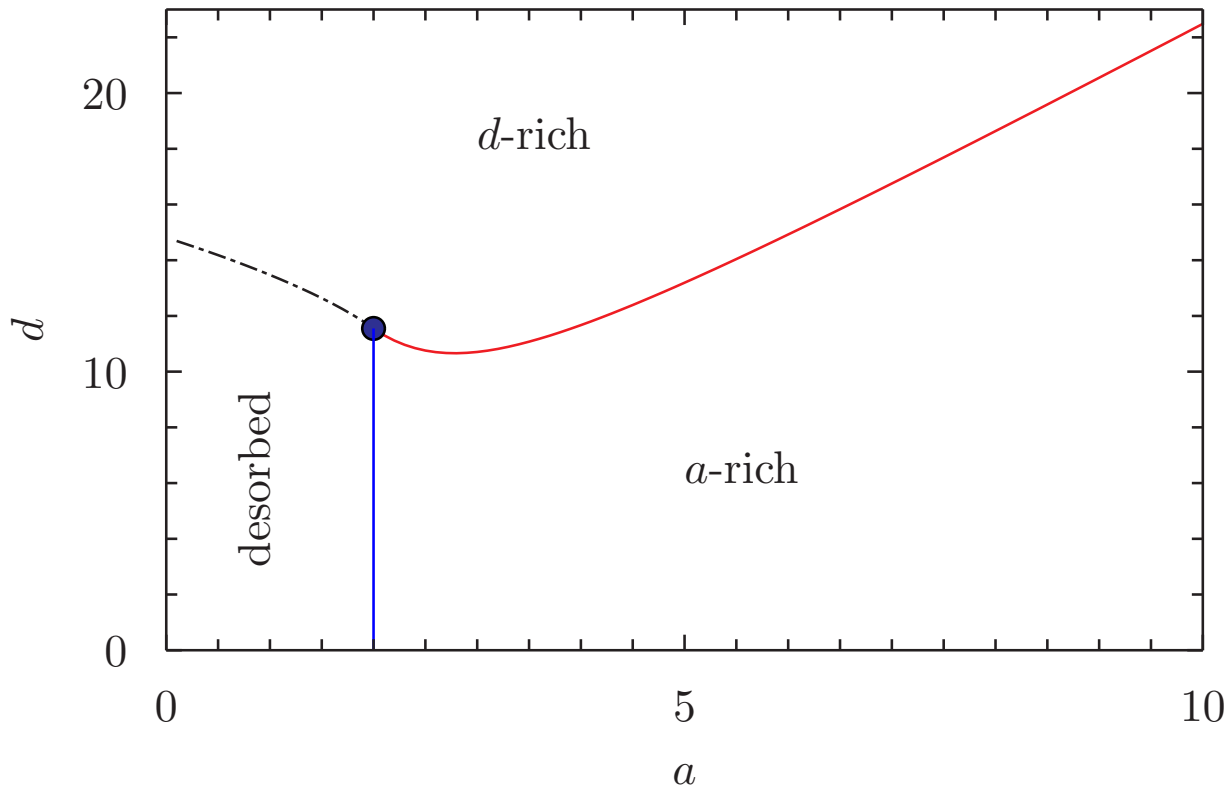
There are two singularities of $G(a,a;z)$ giving rise to two phases:

- A desorbed phase: $\mathcal{A}=\mathcal{D}=0$
- The bottom walk is adsorbed (an a-rich phase): $\mathcal{A}>0$ with $\mathcal{D}=0$

The simple pole in $1-dP(a;z)=0$ gives rise to the third phase

- Both walks are adsorbed and this is a d-rich phase: $\mathcal{D}>0$, and $\mathcal{A}>0$

PHASE DIAGRAM



The first-order transition is marked with a dashed line, while the two second-order transitions are marked with solid lines. The three boundaries meet at the point $(a, d) = (a^*, d^*) = (2, 11.55\ldots)$.

PHASE TRANSITIONS

- The Desorbed to a-rich transition is
  - the standard second order adsorption transition
  - on the line $a = 2$ for $d < d^*$
- On the other hand the Desorbed to d-rich transition is first order
- While the a-rich to d-rich transition is also second order.

The other two phase boundaries are solutions to equations involving $G(a, a)$

The point where the three phase boundaries meet can be computed as

$$(a^*, d^*) = \left(2, \frac{16(8 - 3\pi)}{64 - 21\pi}\right)$$

Note that d^* is not algebraic.

NATURE OF THE SOLUTION

Desorbed to *d*-rich transition occurs at a value of $d_c(a)$ for $a < 2$.
We found

$$d_c(1) = \frac{8(512 - 165\pi)}{4096 - 1305\pi}$$

which is not algebraic.

- *If generating function were D-finite then $d_c(1)$ must be algebraic*
- *Hence generating function is not D-finite*
- *though it is calculated in terms of one.*

FIXED ENERGY RATIO MODEL FAMILY

Family of models parameterised by $-\infty < r < \infty$ *where*

$$\varepsilon_d = r\varepsilon_a \qquad \textit{and so} \qquad d = a^r$$

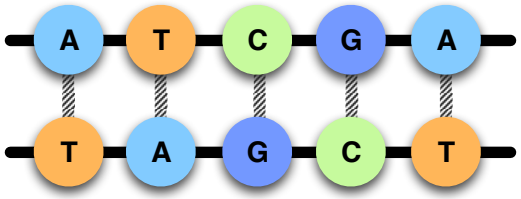
- $r = 2$ model has *two* phase transitions as temperature changed .
- At very low temperatures the model is in a d-rich phase
- while at high temperatures the model is in the desorbed state.
- At intermediate temperatures the system is in an a-rich phase.
- Both transitions are second-order with jumps in specific heat.

Conclusions for Double Interaction model

- Vesicle above a surface — both sides of the vesicle can interact
- Exact solution of generating function
- Obstinate kernel method with a minor generalisation
- Solution is not D-finite — LGV lemma does not apply directly
- There are two low temperature phases
- Line of first order transition and usual second order adsorption.
- Owczarek, Rechnitzer, and Wong, *J. Phys. A*, 45 425002, (2012)

DNA

- DNA is a polymer consisting of four repeating nucleic bases A,C,G,T.
- Two strands entwined with a helix structure
- Denaturation: At high T, strands pulled apart



Work with Rami Tabbara

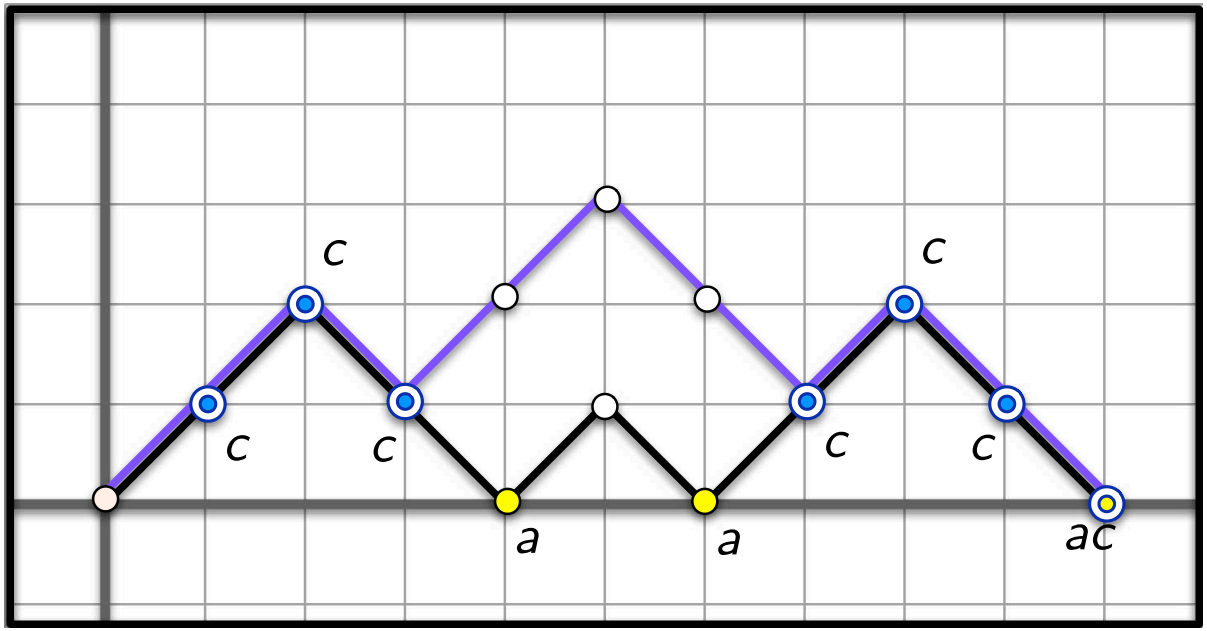
(Our model — Adsorption and Unzipping)

- *double DNA strand in a solvent*
- *near attractive surface*
- *assume aligned base sequence*
- *Use (∞) - friendly directed walks again*

UNZIPPING ADSORPTION MODEL

Let T be the system temperature, k_B the Boltzmann constant.

- surface visit step: $a \equiv e^{\varepsilon_a / k_B T}$
- shared site contact: $c \equiv e^{\varepsilon_c / k_B T}$
- trivial walk consisting of zero steps has weight 1.



An allowed configuration of length 10. The overall weight is $a^3 c^7$

GENERATING FUNCTION

- Partition function: $Z_n(a,c) = \sum_{\widehat{\varphi} \ni |\widehat{\varphi}|=n} a^{m_a(\widehat{\varphi})} c^{m_c(\widehat{\varphi})}$
- Generating function: $G(a,c) \equiv G(a,c;z) = \sum_{n\geq 1} Z_n(a,c) z^n$
- Reduced free energy:

$$\kappa(a,c) = \lim_{n\to\infty} n^{-1} \log Z_n(a,c) = \log z_s(a,c)$$

where $z_s(a,c)$ is dominant singularity of G w.r.t. z

Two order parameters:

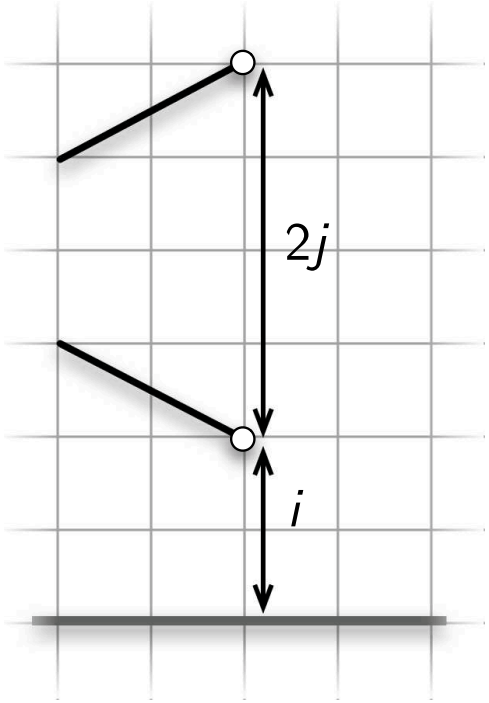
$$\mathcal{A}(a,c) = \lim_{n\to\infty} \frac{\langle m_a \rangle}{n} \qquad \text{and} \qquad \mathcal{C}(a,c) = \lim_{n\to\infty} \frac{\langle m_c \rangle}{n},$$

Generalised generating function

We consider walks φ in the larger set, where each walk can end at any possible height.

- To find $G(a,c)$, consider larger class of configs.
- Generalised generating function:

$$F(r,s) \equiv F(r,s,a,c;z) = \sum_{\varphi\in\Omega} a^{m_a(\varphi)} c^{m_c(\varphi)} r^i s^j z^n$$



FUNCTIONAL EQUATION

As before:

$$K(r,s)F(r,s) = \frac{1}{ac} + \left(C(c) - \frac{zr}{s}\right) F(r,0) + \left[A(a) - \frac{z}{r}(s+1)\right] F(0,s) - A(a)C(c)F(0,0)$$

where

$$A(a) = \frac{a-1}{a}, \quad C(c) = \frac{c-1}{c}$$

and the same kernel $K(r,s)$ is

$$K(r,s) \equiv K(r,s;z) = \left(1 - z\left[r + \frac{s}{r} + \frac{r}{s} + \frac{1}{r}\right]\right).$$

Magic decomposition of $G(a,c)$

We can re-write $G(a,c)$ in terms of $G(a,1)$ and $G(1,c)$:

$$G(a,c) = \frac{1}{(a-1)(c-1)} + \frac{p_1(a,c,z)}{p_2(a,c,z) + p_3(a,c,z)G(a,1) + p_4(a,c,z)G(1,c)}$$

SOLUTION FOR $G(a,1)$

Exact solution for $G(a,1)$ is known from first part of talk!

Specifically:

$$G(a,1) = \sum_{n\geq 0} z^{2n} \sum_{k=0}^{n} a^k \frac{k(k+1)(k+2)}{(2n-k)(n+1)^2(n+2)} \binom{2n-k}{n} \binom{2n}{n}.$$

SOLUTION FOR $G(1, c)$

No known previous solution for $G(1, c)$

- We have: $G(1, c) = [r^1]\text{mess}(r, a, c; z) \equiv [r^1]H(r; z)$
- Traditionally, get series rep. for H. Sometimes an art — see first part of this talk!
- Alternative — find differential equation satisfied by generating function
- Difficult to solve, but DE just as useful for analysis!
- Use Zeilberger-Gosper algorithm: Maple: DETools package, Zeilberger hyperexp. implementation
- Result: DE for $G(1, c)$ is order 6 with poly. coeff of $\deg_z = 12$

SINGULARITIES OF $G(a,1)$ & $G(1,c)$

- Recall, free energy $\kappa(a,c) = \log z_s(a,c)$
- For $G(a,1)$, prev. known:

$$z_s(a,1) = \begin{cases} z_b \equiv 1/4, & a \le 2 \\ z_a \equiv \frac{\sqrt{a-1}}{2a}, & a > 2 \end{cases}$$

- For $G(1,c)$, we use the DE (roots of leading poly. coeff.):

$$z_s(1,c) = \begin{cases} z_b \equiv 1/4, & c \le 4/3 \\ z_c \equiv \frac{1-c+\sqrt{c^2-c}}{c}, & c > 4/3 \end{cases}$$

ORDER PARAMETERS AND PHASES

Recall lim. avg. surface and shared site contacts resp.

$$\mathcal{A}(a,c) = \lim_{L\to\infty} \frac{\langle m_a \rangle}{L} = a\frac{\partial \kappa}{\partial a}, \qquad \mathcal{C}(a,c) = \lim_{L\to\infty} \frac{\langle m_c \rangle}{L} = c\frac{\partial \kappa}{\partial c}$$

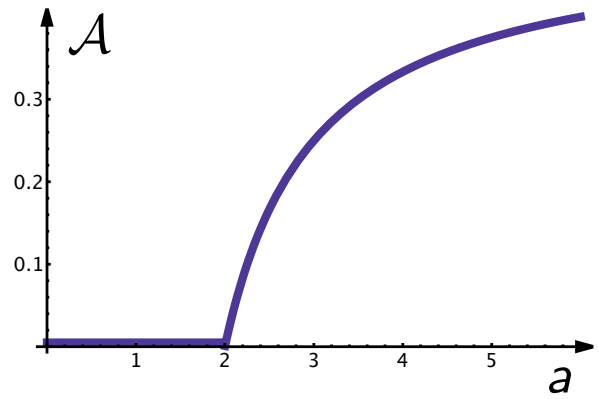
Four Phases:

- Desorbed: $\mathcal{A} = \mathcal{C} = 0$
- Adsorbed: (a-rich) $\mathcal{A} > 0, \mathcal{C} = 0$
- Zipped: (c-rich) $\mathcal{A} = 0, \mathcal{C} > 0$
- Zipped & Adsorbed: (ac-rich) $\mathcal{A} > 0, \mathcal{C} > 0$

TRANSITIONS OF $G(a,1)$ & $G(1,c)$

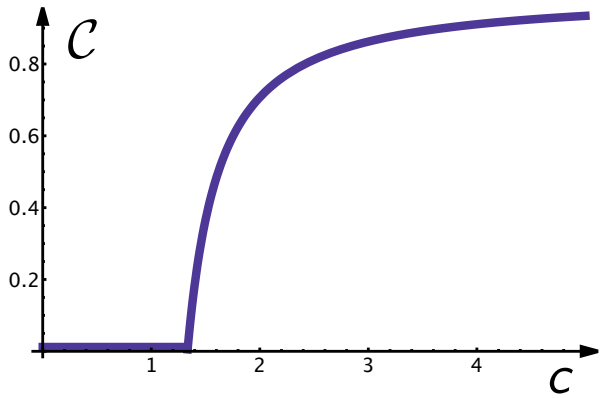
- For $G(a,1)$:

$$\mathcal{A}(a,1) = \begin{cases} 0, & a \leq 2 \\ \frac{a-2}{2(a-1)}, & a > 2 \end{cases}$$



- For $G(1,c)$:

$$\mathcal{C}(1,c) = \begin{cases} 0, & c \leq 4/3 \\ \frac{c-2+\sqrt{c(c-1)}}{2(c-1)}, & c > 4/3 \end{cases}$$



- Second-order adsorption and zipping phase trans. resp.

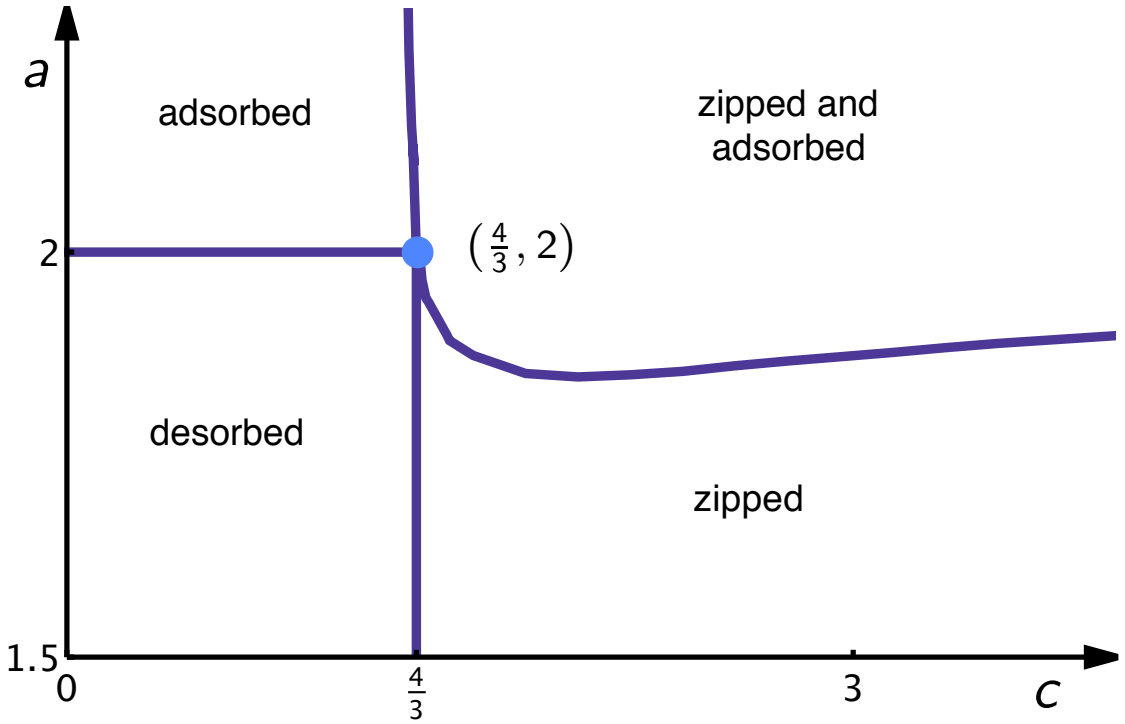
Analysing $G(a, c)$

Recall

$$G(a,c) \sim \frac{p_1(a,c,z)}{p_2(a,c,z) + p_3(a,c,z)G(a,1) + p_4(a,c,z)G(1,c)}$$

- $\Rightarrow$ Singularities: Look at $G(a,1)$, $G(1,c)$ and root of above denom.
- Can find location of pole precisely

PHASE DIAGRAM



Low-temp argument

- $c \to \infty$, crit. $a = 2$
- $a \to \infty$, crit. $c = \sqrt{5} - 1$

Conclusion

- Simple model of DNA as two friendly walks near a boundary
- Used obstinate kernel method to produce two sets refined fn. eqns
- Combined these eqn to relate $G(a, c)$ to both $G(a, 1)$ and $G(1, c)$
- Used Zeilberger-Gosper algorithm to find lin. DE for $G(1, c)$
- Started analysis of phase diagram

FUTURE WORKS

- Finish analysis of (a, c)- model
- Combine single, double surface and unzipping interactions
- Also in progress is work in a slit
- Extend these to various interactions