Topological Entanglement Complexity of Systems of Polygons and Walks in Tubes

A Thesis Submitted to the
College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the degree of Doctor of Philosophy
in the Department of Department of Mathematics and Statistics
University of Saskatchewan
Saskatoon

By

Mahshid Atapour

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Mathematics and Statistics
106 Wiggins Road
University of Saskatchewan
Saskatoon, Saskatchewan
Canada
S7N 5E6

Abstract

In this thesis, motivated by modelling polymers, the topological entanglement complexity of systems of two self-avoiding polygons (2SAPs), stretched polygons and systems of self-avoiding walks (SSAWs) in a tubular sublattice of $\mathbb{Z}^3$ are investigated. In particular, knotting and linking probabilities are used to measure a polygon's self-entanglement and its entanglement with other polygons respectively. For the case of 2SAPs, it is established that the homological linking probability goes to one at least as fast as $1 - O(n^{-1/2})$ and that the topological linking probability goes to one exponentially rapidly as n, the size of the 2SAP, goes to infinity. For the case of stretched polygons, used to model ring polymers under the influence of an external force f, it is shown that, no matter the strength or direction of the external force, the knotting probability goes to one exponentially as n, the size of the polygon, goes to infinity. Associating a two-component link to each stretched polygon, it is also proved that the topological linking probability goes to unity exponentially fast as $n \to \infty$. Furthermore, a set of entangled chains confined to a tube is modelled by a system of self- and mutually avoiding walks (SSAW). It is shown that there exists a positive number γ such that the probability that an SSAW of size n has entanglement complexity (EC), as defined in this thesis, greater than γn approaches one exponentially as $n \to \infty$. It is also established that EC of an SSAW is bounded above by a linear function of its size. Using a transfer-matrix approach, the asymptotic form of the free energy for the SSAW model is also obtained and the average edge-density for span m SSAWs is proved to approach a constant as $m \to \infty$. Hence, it is shown that EC is a "good" measure of entanglement complexity for dense polymer systems modelled by SSAWs, in particular, because EC increases linearly with system size, as the size of the system goes to infinity.

Acknowledgements

First I would like to express my gratitude to my supervisor, Prof. Chris Soteros, for her support, patience and attempt to generously transmit the best of her knowledge to me. My special thanks goes to my husband Dr. Vahid Anvari who went through a lot of difficulties because of my studies abroad but never gave up his encouragement, guidance and full support on me, my mother Mrs Tahereh Abdali who sacrificed a lot of things of her own life to provide me with a good education, and finally my father Mr Gholamreza Atapour who taught me how to stand up at the times of difficulty without sacrificing my dignity.

I would like to thank the members of my advisory committee and my external examiner, Prof. Yuanan Diao, for reading my thesis and their valuable comments. I would also like to thank all the professors with whom I took at least one course and acknowledge their share in my PhD training. I am grateful to the head of the Department of Mathematics and Statistics, Prof. Raj Srinivasan, for his continual support during the course of my PhD. I would like to extend my thanks to the former head secretary of the department, Mrs Cheryl Piche, who was a great comfort for me and my husband during our times of difficulty in a foreign country. I wish her a very happy retirement days. I would also like to thank the current department's secretaries, Mrs Charlotte Hill and Mrs Karen Ruston, and my friends Manuela Haias, Alice Li, Keivan Mohajer, Yaling Yin and Jiezhi Qi who all had a share in creating a friendly work environment for me.

Based on the results of Chapter 4 of this thesis, in collaboration with my supervisor, Prof. S.G. Whittington and Prof. C. Ernst, a paper has been submitted to the Journal of Knot Theory and Its Ramifications. I would like to acknowledge the help of my collaborators in preparation of this chapter. I would also like to particularly thank Prof. S.G. Whittington for his valuable suggestions regarding Chapter 5 of my thesis. Thanks to the University of Saskatchewan for providing me financial assistance and to the NSERC of Canada for indirectly supporting me financially, through my supervisor's grant, during my stay in Saskatoon.

This thesis is dedicated to three people whose existence defines my life: my husband, my mother and my father.

Contents

List of Figures

List of Abbreviations

SAW	an acronym for self-avoiding walk
USAW	an acronym for undirected self-avoiding walk
SAP	an acronym for self-avoiding polygon
$\mathbb{Z}^d$	the d-dimensional hypercubic lattice
$\mathbb{R}^d$	d-dimensional Euclidean space
c_n	the number of n-edge SAWs starting at the origin in $\mathbb{Z}^3$
p_n	the number of n-edge SAPs in $\mathbb{Z}^3$ (up to translation)
κ	the connective constant for $\mathbb{Z}^3$
2SAP	a system of two self-avoiding polygons
SSAW	a system of self-avoiding walks
$\hat{i}$	the unit vector in $\mathbb{R}^3$ along the positive direction of x-axis
$\hat{j}$	the unit vector in $\mathbb{R}^3$ along the positive direction of y-axis
$\hat{k}$	the unit vector in $\mathbb{R}^3$ along the positive direction of z-axis
$V(G)$	the vertex set of the graph G
$E(G)$	the edge set of the graph G
$x(v)$	the x-coordinate of the vertex v
$y(v)$	the y-coordinate of the vertex v
$z(v)$	the z-coordinate of the vertex v
$\mathbb{N}$	the set of positive integer numbers
κ_b	the connective constant for bridges in $\mathbb{Z}^3$
$T(N, M)$	an (N, M)-tube in $\mathbb{Z}^3$
$\kappa(N, M)$	the connective constant for SAWs confined to an (N, M)-tube
$\kappa_p(N, M)$	the connective constant for SAPs confined to an (N, M)-tube
$\kappa_b(N, M)$	the connective constant for bridges confined to an (N, M)-tube
$\mathbb{S}^d$	the d-dimensional sphere
$Lk(K_1, K_2)$	linking number of a given two-component link (K_1, K_2)
τ	a link
$\tau^{\#}$	an augmented link obtained from τ
κ_o	the connective constant for unknotted SAPs in $\mathbb{Z}^3$

Chapter 1

Introduction

Long ring (linear) polymers have been modelled mathematically by self-avoiding polygons (walks) in lattices [55]. A self-avoiding walk (SAW) is a path on a lattice that does not visit the same site more than once. Similarly, a self-avoiding polygon (SAP) is a closed undirected path on a lattice which does not intersect itself. Simple examples of lattices are hypercubic lattices such as the square and simple cubic lattice, with which we are mainly concerned in this thesis. For any integer number $d \geq 2$, the *hypercubic lattice* $\mathbb{Z}^d$ consists of all the integer points in $\mathbb{R}^d$, as vertices, and all the unit length edges joining two integer points in $\mathbb{R}^d$. A *polymer* is a molecule consisting of several small units called monomers joined together by chemical bonds. The number of monomers which bond to a monomer represents the functionality of the monomer. A polymer which is composed of a string of bonded monomers, all of which have functionality two is referred to as a ring polymer. In the SAP lattice model of a ring polymer, each n-step self-avoiding polygon (or walk) represents a possible polymer conformation, and the self-avoidance represents that there is an excluded volume around each monomer. Let c_n represent the number of n-edge SAWs in $\mathbb{Z}^3$ starting at the origin. Let also p_n denote the number of n-edge SAPs in $\mathbb{Z}^3$ (up to translation). The asymptotic behaviour of c_n (p_n) is of interest to mathematicians, in particular, because there is no exact formula for c_n or p_n. In 1954 Hammersley [18] proved that

$$\kappa \equiv \lim_{n\to\infty} n^{-1} \log c_n, \tag{1.0.1}$$

called the *connective constant* for $\mathbb{Z}^3$, exists and is finite. A similar statement has also been proved by Hammersley for self-avoiding polygons in $\mathbb{Z}^3$, i.e. it has been proved that

$$\kappa_p \equiv \lim_{n\to\infty} n^{-1} \log p_n, \tag{1.0.2}$$

the connective constant for SAPs in $\mathbb{Z}^3$, exists and is exactly equal to κ [19].

Polymers in solution are highly flexible objects which can assume many different configurations. These configurations can result in self-entanglement or entanglement with other polymer molecules. Measurement of entanglements is important to chemists, physicists and molecular biologists since it is believed that the entanglement complexity of polymers is related to their crystallization behaviour [17] and rheological properties [14] and is important in understanding cellular processes (replication, transcription, etc) involving biopolymers such as DNA [56, 60, 61]. In particular, if a ring closure reaction occurs, the entanglement can appear as knots or links in a solution of ring polymer molecules. In this case, the topology of any knot or link (i.e. knot or link type) produced cannot change without breaking chemical bonds in the polymer [42]. For investigating the topological entanglement of such polymer solutions, so-called "good" measures for characterizing a polymer's topological entanglement with other polymer molecules are needed. Generally speaking, a good measure of topological entanglement should have the property that the larger the size of a random polymer system and the more complex its topology, the greater the topological entanglement [50]. For instance, knotting probability was proved, using a lattice model of polymers, to be a good measure of topological self-entanglement [55]. In this thesis, we consider lattice models of polymers confined to tubes and investigate their topological entanglement. We will rigorously show that the measures used here, which are based on "linking", are good measures for characterizing topological entanglement between lattice polymers. We will investigate the topological entanglement of polymer systems modelled by a set of self-avoiding polygons (or walks) confined to a tube (a sublattice of the simple cubic lattice $\mathbb{Z}^3$). The models which are considered in this thesis are systems of two self-avoiding polygons (2SAPs), stretched polygons, loops and systems of self-avoiding walks (SSAWs). We will focus on measures based on linking, such as linking probability and one we refer to as the entanglement complexity (EC), in order to characterize the topological entanglement of these models.

As mentioned before, the measures used in this thesis are based on the topological concept of linking. To explain the measures further, we first review the following topological terminology. Note that the definitions presented here are informal and the precise definitions are given in Chapter 3. Any simple closed curve in $\mathbb{R}^3$ is called a knot. Two knots K_1 and K_2 are said to have the same knot type if $\mathbb{R}^3$ can be continuously deformed such that K_1 is taken into K_2 (as will be explained in Chapter 3, mathematically this means that K_1 and K_2 are ambient isotopic). In particular, K is said to be unknotted if $\mathbb{R}^3$ can be continuously deformed such that K is taken into the unit circle (i.e. K is ambient isotopic to $\mathbb{S}^1$). Otherwise K is said to be knotted [7]. Knots in closed polymer chains are significant to physicists, chemists and molecular biologists. For instance, knots in closed circular DNA give information about the mechanism of enzyme action on DNA molecules [60, 61]. Knots have been detected in circular DNA and determined by electron microscopy [8, 62], then the knotting probability has been measured experimentally [47]. In addition to numerical investigations of knotting probabilities and knot distributions [23, 29, 35, 59] of self-avoiding polygons, some rigorous results are known. In particular, it has been proved that all but exponentially few sufficiently long self-avoiding polygons are knotted and thus the probability of knotting of lattice polygons approaches one as the size of a polygon goes to infinity [43, 55]. A "pattern theorem" due to Kesten [28] for self-avoiding walks was a key ingredient to this proof.

Polymers are often confined to restricted spaces; for instance the presence of some large molecules in the cell-nucleus confines the nuclear DNA to a reduced space, with corresponding effects on the topological properties of the DNA. The effects of such geometrical constraints on the knot probability of ring polymers, modelled by self-avoiding polygons in $\mathbb{Z}^3$, have been investigated by restricting the polygons to a tube. An (N, M)-tube is a sublattice of $\mathbb{Z}^3$ bounded by the four planes $y = 0$, $y = N$, $z = 0$ and $z = M$. Note that in this thesis we will be mainly concerned with lattice objects confined to a tube. It has been proved rigorously that the probability of knotting for a polygon confined to a tube also approaches one exponentially as the length of the polygon goes to infinity [49]. In a tube, SAPs and SAWs have different asymptotic properties, in the sense that the connective constant for SAPs is not the same as that for SAWs anymore. Instead it is strictly less than the connective constant for SAWs [49]; this means that asymptotically the number of SAPs and SAWs behave differently. So, unlike the situation for SAPs and SAWs in $\mathbb{Z}^3$, in this case a pattern theorem for SAWs does not necessarily work for SAPs. An appropriate pattern theorem for SAPs in a tube has been established and used to compute the probability of knotting as $n \to \infty$ [51]. More details on SAWs and SAPs and their respective asymptotic behaviour are reviewed in Chapter 2. In particular, the existence proof of the connective constant and pattern theorems for SAWs and SAPs in $\mathbb{Z}^3$ as well as tubes are reviewed.

Polygonal self-entanglements have been investigated using knotting probability [25, 42]. One can ask similar questions regarding the topological entanglement of two ring polymers. A pair of polygons in a lattice can be considered as a two-component link, which is defined to be a disjoint union of two knots. Roughly speaking, similar to the case for knots, two links L_1 and L_2 are said to have the same link type if $\mathbb{R}^3$ can be continuously deformed such that L_1 is carried to L_2. We say two disjoint knots K_1 and K_2 are topologically unlinked (splittable) if $\mathbb{R}^3$ can be continuously deformed such that K_1 and K_2 are carried to a pair of knots which are separated by a two-dimensional plane. On the other hand, K_1 and K_2 are said to be homologically unlinked if K_1 bounds an orientable surface which is disjoint from K_2. One way to determine whether a two-component link L is homologically linked is by calculating the linking number and showing that it is non-zero [7] (note that the precise definition of linking number is given in Chapter 3). Chapter 3 provides background information about knots and links and reviews some previous works on the knotting and linking probability of SAPs in $\mathbb{Z}^3$ as well as tubes.

The topological entanglement of two ring ploymers, modelled by a pair of self-avoiding polygons, has been measured by investigating the linking probability of the polygon pairs [39, 58]. In Chapter 4 of this thesis, in particular, the following question is addressed: *under what conditions are all but exponentially few sufficiently* *long pairs of self-avoiding polygons linked?* For pairs of mutually avoiding n-edge self-avoiding polygons in $\mathbb{Z}^3$ where each polygon is constrained to be unknotted and where the two polygons are constrained to form a non-splittable link, it has been shown [52] that the number of embeddings grows exponentially with n and that the exponential growth rate is independent of the link type. If the two polygons are constrained to have a pair of edges, one from each polygon, which are within a fixed distance from each other, we use arguments similar to those of Orlandini *et al* [39] and Soteros *et al* [52] to establish that the exponential growth rate of the number of topologically linked polygon pairs (up to translation) is equal to that of the number of topologically unlinked polygon pairs. We prove this statement in Chapter 3 of the thesis. So, unlike the situation with knotting, we cannot say that all but exponentially few sufficiently long pairs of self-avoiding polygons are linked, even with this distance constraint. It is possible (although not proved) that the linking probability goes to one as n goes to infinity, but it will not go to one exponentially rapidly.

Tesi *et al* [58] investigated the same question for pairs of mutually avoiding self-avoiding polygons confined to tubes and came to the same conclusion. In a tube, as in $\mathbb{Z}^3$, even if two edges, with one from each polygon, are forced to be close, the rest of each polygon has a considerable amount of freedom so that their centres of mass can be very far apart. In Chapter 4 of this thesis we consider a much more severe distance constraint. The two polygons are constrained in such a way that (roughly speaking) each edge of one polygon is forced to be "close" to some edge of the other polygon. In this case one might expect that for large enough values of n the two polygons would be linked with high probability. Herein, we consider a system of two self-avoiding polygons (2SAP) confined to a lattice tube with dimensions $(\infty \times N \times M)$ with an added constraint that forces each edge of one polygon to be close to some edge of the other polygon and prove theoretical results about the rate that the linking probabilities (both homologically and topologically) go to 1. Specifically we consider a pair of mutually avoiding self-avoiding polygons each confined to and spanning a tube. Such a pair is referred to as a System of two Self-avoiding Polygons (2SAP) and n is used to denote the total number of edges in the pair. We establish that the homological linking probability goes to one at least as fast as $1 - O(n^{-1/2})$ and that the topological linking probability goes to one exponentially rapidly. Furthermore we prove for this model that the linking number grows (with probability one) faster than any function that is $o(\sqrt{n})$; i.e. for any function $f(n) = o(\sqrt{n})$, there exists $A \geq 0$ such that as $n \to \infty$ the probability that $|Lk(\omega_1, \omega_2)| \geq f(n)$, with (ω_1, ω_2) the component polygons of an n-edge 2SAP, satisfies

$$\mathbb{P}\big(|Lk(\omega_1, \omega_2)| \geq f(n)\big) = 1 - \frac{A}{\sqrt{n}} + o\big(\frac{1}{\sqrt{n}}\big), \tag{1.0.3}$$

where $Lk(\omega_1, \omega_2)$ represents the linking number of (ω_1, ω_2). Hence

$$\lim_{n\to\infty} \mathbb{P}\big(|Lk(\omega_1, \omega_2)| \geq f(n)\big) = 1. \tag{1.0.4}$$

Note that here, given a pair of functions $f(n)$ and $g(n)$, we write $f(n) = O(g(n))$ if there exist constants A and B, and a positive integer N such that $Ag(n) \leq f(n) \leq Bg(n)$ for any $n \geq N$, and write $f(n) = o(g(n))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$. We also show that the linking number cannot grow faster than linearly in n because of the tube constraint; i.e. there exist constants a and b such that for any n-edge 2SAP

$$|Lk(\omega_1, \omega_2)| \leq an + b. \tag{1.0.5}$$

We give a simple example to show that the upper bound in equation 1.0.5 for 2SAPs can be realized.

A polymer's topological entanglement may be affected by being subject to some external forces. One important question in this regard is: *How does the topological entanglement change when a polymer is compressed or stretched under the influence of an external force f?* In Janse van Rensburg *et al* 2008 [27], a ring polymer confined between two parallel walls (planes) and pulled by an external force in the direction perpendicular to the wall is modelled by a lattice polygon subjected to an applied force f along the z-direction of the lattice and perpendicular to the walls, a so-called stretched polygon. This model explains the situation where a ring polymer, such as circular DNA, is subject to a force in the presence of a topoisomerase which mediates strand passages and may change the knot type of the polymer. In [27], a pattern theorem is proved for lattice polygons in the presence of any sufficiently large applied force $f > 0$. This theorem is then used to examine the incidence of entanglements such as knotting when the polygons are influenced by a large force. Here, in Chapter 5 of this thesis, we add the tube confinement constraint to this model and address similar questions for polygons confined to a tube. Unlike the situation for polygons in $\mathbb{Z}^3$, in a tube, polygons under the influence of the external force f do not have much freedom and must stay inside the tube. In particular, the tube constraint will allow us to prove the pattern theorem for any arbitrary value of f, not just for large values of f. We also prove a pattern theorem for loops (an undirected self-avoiding walk in which both endpoints have the same x-coordinate). Furthermore, in addition to investigating the knotting probability, we associate a two-component link to each polygon (loop) in a tube and examine the incidence of topological linking when the polygon (loop) is under the influence of a force f. Specifically, for any f, we prove that both the knotting and topological linking probabilities go to one exponentially as the number of edges tends to infinity. Using a transfer-matrix approach, we also determine the asymptotic form of the free energy for stretched polygons. We use this to show that, the average span per edge of a randomly chosen n-edge stretched polygon approaches a positive constant as $n \to \infty$ and is non-decreasing in f almost everywhere. We also establish that the average number of occurrences per edge of a tight trefoil SAP configuration (precisely defined in Section 5.4) in any n-edge stretched polygon approaches a positive constant (independent of n but dependent on f) as $n \to \infty$.

Another area of interest for polymer chemists is characterizing entanglements for polymers in dense systems such as melts [40, 41]. Most results about entanglement complexity of polymers in melts are obtained using numerical studies [13, 40, 41]. However, at least one open question still remains: *What is the best measure for characterizing the entanglement complexity of polymers in dense systems and how does this measure depend on various properties of the system?* It is difficult to find the "best" measure for this purpose but we can at least look for a "good" measure. This is discussed in Chapter 6 of this thesis. One way to characterize this entanglement complexity, proposed by Orlandini *et al* in 2000 [41], is as follows. A polymer melt is considered as a set of entangled chains, modelled by a set of self- and mutually avoiding walks. Fixing the chain conformations, imagine cutting a cube or tube out of the system. The conformations of the parts of the chains which are in the interior of the cube or tube are considered. Using the simple cubic lattice model, they investigate this numerically by studying a number of self- and mutually avoiding walks confined to a cube. For each pair of self-avoiding walks, joining the two ends (vertices of degree one) of each walk outside the cube, a two-component link is formed. And they take the sum of the absolute value of the linking numbers, over all the possible SAW pairs, as a measure of the entanglement complexity (EC) of this polymer system. Then the properties of the complete melt can be inferred by investigating the properties of these chains in cubes.

In order to investigate Orlandini *et al*'s proposal further we address the problem of the entanglement complexity of a polymer system modelled by a system of self-avoiding walks (SSAW) (the precise definition is given in Chapter 6). The goal is to build a theoretical framework that will allow us to apply the available mathematical techniques towards developing and understanding the Orlandini *et al* 2004 model and answering the following questions:

- How does the entanglement complexity change with respect to the total number of monomers in the system?

- How does the entanglement complexity change with respect to the span of the system along the tube?

- How does the entanglement complexity change with respect to the number of chains in the system?

- How does the entanglement complexity change with respect to the system's density?

- How does the entanglement complexity change with respect to the size of the tube to which the polymers are confined?

- How does the total number of monomers, the span, the number of chains, the density and the size of the tube depend on each other?

Specifically, in Chapter 6 we investigate, under some constraints, the entanglement complexity of several self- and mutually avoiding walks confined to an (N, M)-tube. We rigorously prove that the entanglement complexity (EC), as measured in [41], of a polymer system with "size" n (size of an SSAW G can be measured by the number of edges $(f(G, 3))$, span $(f(G, 1))$ or the number of degree one vertices $(f(G, 2))$) is asymptotically (with probability one) bounded below by a linear function of n; i.e. there exists a positive number γ such that the probability that a polymer system of size n has entanglement complexity greater than γn approaches 1 as n goes to infinity. This supports the idea that EC is a good measure of topological entanglement in polymer systems modelled by SSAWs. It is also shown that the entanglement complexity of SSAWs of size n is bounded above by a linear function of n. Furthermore, measuring size by the number of edges, for $N \geq 2$ and $M \geq 2$, the connective constant for SSAWs in an (N, M)-tube is compared with the connective constant for self-avoiding walks in an $(N-2, M-2)$-tube and is shown to be strictly greater than that for SAWs. Given $Y_3 = [x, y, z]^T$, for $j = 1, 2, 3$, let $X_n^{\#j}(Y_3)$ be a random variable taking its values from the set of SSAWs with size $n = f(G, j)$ (in C_n^*) and with the probability distribution

$$\mathbb{P}(X_n^{\#j}(Y_3) = G) = \frac{x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}}{Z_n^{\#j}(N, M, Y_3)}, \tag{1.0.6}$$

where

$$Z_n^{\#j}(N, M, Y_3) = \sum_{G \in C_n^*} x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}. \tag{1.0.7}$$

Ultimately, based on our theoretical results on the SSAW model we will conclude that the following statements (equations) show how EC depends on various properties of SSAWs such as the number of edges, span and the number of degree one vertices and also indicate how these properties are related to each other:

- There exists $\gamma^{\#1}(Y_3) > 0$ such that the probability that the EC of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ is bounded below by $\gamma^{\#1}(Y_3)n$ and bounded above by $a'(N,M)a(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}(\gamma^{\#1}(Y_3)n < EC(X_n^{\#1}(Y_3)) \leq a'(N,M)a(N,M)n) = \lim_{n\to\infty} [1 - e^{-\alpha'(1,Y_3)n+o(n)}] = 1, \qquad (1.0.8)$$

where

$$a(N,M) = (1/2)(M+2N+1)[(N+1)(4N-2)+(9/2)(4N-3)] \qquad (1.0.9)$$

and

$$a'(N,M) = 2N(M+1)+2M(N+1)+(M+1)(N+1). \qquad (1.0.10)$$

In other words, as the span of SSAWs increases one expects EC to be bounded linearly in the span with probability one.

We also show that the limit inferior of the average EC of span m SSAWs per span is bounded below by a positive constant, i.e.

$$\liminf_{m\to\infty} \frac{E_Y\big(EC(X_m^{\#1}(Y_3))\big)}{m} \geq \gamma_l. \qquad (1.0.11)$$

- There exists $\gamma^{\#3}(Y_3) > 0$ such that the probability that the EC of a randomly chosen n-edge SSAW $X_n^{\#3}(Y_3)$ is bounded below by $\gamma^{\#3}(Y_3)n$ and bounded above by $a(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}(\gamma^{\#3}(Y_3)n < EC(X_n^{\#3}(Y_3)) \leq a(N,M)n) = \lim_{n\to\infty} [1 - e^{-\alpha'(3,Y_3)n+o(n)}] = 1. \qquad (1.0.12)$$

In other words, as the number of edges of SSAWs increases one expects EC to be bounded linearly in the number of edges with probability one.

- There exists $\epsilon_P^{\#1}(Y_3) > 0$ such that the probability that the number of degree one vertices of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ is bounded below by

$\epsilon_P^{\#1}(Y_3)n$ and bounded above by $a'(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#1}(Y_3)n \le f(X_n^{\#1}(Y_3),2) \le a'(N,M)n\big) = \lim_{n\to\infty}[1-e^{-\alpha''(1,2,Y_3)n+o(n)}] = 1. \tag{1.0.13}$$

In other words, as the span of SSAWs increases one expects the number of degree one vertices (the number of disjoint walks) to be bounded linearly in the span with probability one.

- There exists $\epsilon_P^{\#1}(Y_3) > 0$ such that the probability that the number of edges of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ is bounded below by $\epsilon_P^{\#1}(Y_3)n$ and bounded above by $a'(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#1}(Y_3)n \le f(X_n^{\#1}(Y_3),3) \le a'(N,M)n\big) = \lim_{n\to\infty}[1-e^{-\alpha''(1,3,Y_3)n+o(n)}] = 1. \tag{1.0.14}$$

In other words, as the span of SSAWs increases one expects the number of edges to be bounded linearly in the span with probability one.

Using the transfer-matrix method, we also establish that the average number of edges per unit volume of a randomly chosen span m SSAW, $\frac{E_Y(f(X_m^{\#1}(Y_3),3))}{mNM}$, approaches a positive constant as $m \to \infty$ and is non-decreasing in z almost everywhere. However, it still needs further investigation to see how EC changes with respect to Y_3.

- There exists $\epsilon_P^{\#3}(Y_3) > 0$ such that the probability that span of a randomly chosen SSAW $X_n^{\#3}(Y_3)$ with n edges is bounded below by $\epsilon_P^{\#3}(Y_3)n$ and bounded above by $2n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#3}(Y_3)n \le f(X_n^{\#3}(Y_3),1) \le 2n\big) = \lim_{n\to\infty}[1 - e^{-\alpha''(3,1,Y_3)n+o(n)}] = 1. \tag{1.0.15}$$

In other words, as the number of edges of SSAWs increases one expects the span to be bounded linearly in the number of edges with probability one.

- There exists $\epsilon_P^{\#3}(Y_3) > 0$ such that the probability that the number of degree one vertices of a randomly chosen SSAW $X_n^{\#3}(Y_3)$ with n edges is bounded below by $\epsilon_P^{\#3}(Y_3)n$ and bounded above by $2n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#3}(Y_3)n \leq f(X_n^{\#3}(Y_3), 2) \leq 2n\big) = \lim_{n\to\infty} [1 - e^{-\alpha''(3,2,Y_3)n+o(n)}] = 1. \tag{1.0.16}$$

In other words, as the number of edges of SSAWs increases one expects the number of degree one vertices (the number of disjoint walks) to be bounded linearly in the number of edges with probability one.

Furthermore, we obtain the asymptotic form of the free energy for the SSAW model. We also investigate $\rho(N, M; \epsilon)$, the growth constant for SSAWs with limiting edge-density $\frac{\epsilon}{NM}$, as a function of ϵ. We establish the existence of this function and show that it is a continuous and concave function of ϵ and is differentiable almost everywhere in $(\epsilon_{min}, \epsilon_{max})$. However, in order to see how EC changes in terms of the density, we still need to know more about $\rho(N, M; \epsilon)$. In particular, proving a pattern theorem for this function may lead to some results regarding the connection between EC and the density.

Chapter 2

The Lattice Model of SAWs and SAPs

In this chapter the lattice model and the lattice objects under investigation, such as SAWs and SAPs, are introduced. The existence of the connective constant for these lattice objects in $\mathbb{Z}^3$ is reviewed in Section 2.1. In Section 2.2, the extension of these results to the case where SAWs and SAPs are confined to a tube is reviewed [49].

A "Pattern theorem" is the key for investigating the knotting and linking probability of lattice polygons. Kestern's pattern theorem for SAWs in $\mathbb{Z}^3$, [28], and pattern theorems for SAPs and SAWs in tubes are reviewed in sections 2.3 and 2.4 respectively. There have been different approaches towards proving pattern theorems for SAWs and SAPs. Here, in particular, we present in Section 2.5 the so called *pattern insertion* approach which was first introduced in [33, 53]. The pattern insertion strategy will later be employed to obtain pattern theorems for the lattice objects under consideration in the next chapters, i.e. 2SAPs, SSAWs and stretched polygons. Furthermore, in Section 2.6, a transfer-matrix argument, similar to that presented in [51], is established. This will allow us to investigate the asymptotic behaviour of SSAWs and stretched polygons in the next chapters.

2.1 The Connective Constant for SAWs (SAPs) in $\mathbb{Z}^3$

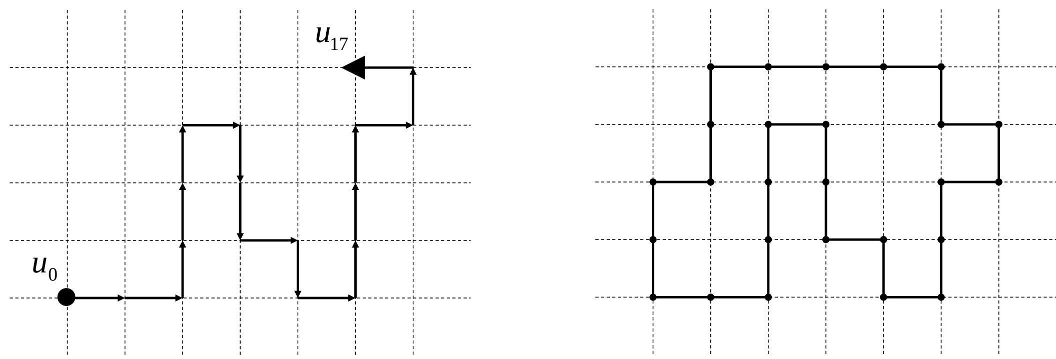
In this section definitions of the simple cubic lattice, $\mathbb{Z}^3$, a self-avoiding walk and a self-avoiding polygon are reviewed. Some arguments regarding the existence of the connective constant for these lattice objects are also discussed; note that graph theory is used for the definitions.

The three dimensional integer lattice is defined to be the infinite graph embedded in $\mathbb{R}^3$ with vertex set $\mathbb{Z}^3$ and edge set $E(\mathbb{Z}^3) = \{\{u,v\} | \ u,v \in \mathbb{Z}^3, \ |u-v| = 1\}$, where $|u-v|$ is the Euclidean distance between u and v. Note that in this thesis by a "subgraph" of $\mathbb{Z}^3$ we mean an embedding of a graph in $\mathbb{Z}^3$. An n-edge *self-avoiding polygon* (SAP) is an n-edge connected subgraph of $\mathbb{Z}^3$ with each vertex having degree two; note that $n \geq 4$. An n-step *self-avoiding walk* (SAW) w in $\mathbb{Z}^3$ is a sequence of distinct vertices

$$w := u_0, u_1, ..., u_{n-1}, u_n \tag{2.1.1}$$

such that $u_i \in \mathbb{Z}^3$, for $i = 0, ..., n$, and for each $i = 0, ..., n-1$ the directed edge (u_i, u_{i+1}) joins two nearest neighbour vertices in $\mathbb{Z}^3$, i.e. $|u_i - u_{i+1}| = 1$. u_0 and u_n are called respectively the *starting point* and *end point* of w. If the direction of a SAW is ignored then it is called an undirected self-avoiding walk; in other words, an *undirected self-avoiding walk* (USAW) is a finite connected subgraph of $\mathbb{Z}^3$ which has exactly two vertices with degree one and every other vertex with degree two. Examples of SAWs and SAPs are illustrated in Figure 2.1. Let $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ be the unit vectors in $\mathbb{R}^3$ along the x-axis, y-axis and z-axis respectively. For convenience, we will sometimes refer to w according to its starting point u_0 and the sequence of steps of unit length which form the walk, for example the SAW in Figure 2.1 (a) is $(2\hat{\imath}, 3\hat{k}, \hat{\imath}, -2\hat{k}, \hat{\imath}, -\hat{k}, \hat{\imath}, 3\hat{k}, \hat{\imath}, \hat{k}, -\hat{\imath})$ starting at $u_0 = (0,0,0)$. We will also refer to a self-avoiding walk (self-avoiding polygon) as a walk (polygon) for short.

In the following two paragraphs some necessary definitions and terminology are reviewed for any general finite subgraph of $\mathbb{Z}^3$. However, we will frequently use these definitions specifically for the lattice objects of interest such as SAPs, 2SAPs and SSAWs. Lexicographical order on vertices of $\mathbb{Z}^3$ is defined as follows. For any pair of vertices $u := (a_1, a_2, a_3)$ and $v := (b_1, b_2, b_3)$ in $\mathbb{Z}^3$, we say $u < v$ *lexicographically* if and only if there exists an integer $1 \leq k \leq 3$ such that $a_i = b_i$ for any $1 \leq i \leq k-1$ and $a_k < b_k$. Given any finite subgraph G of $\mathbb{Z}^3$, let $V(G) \subseteq \mathbb{Z}^3$ and $E(G) \subseteq E(\mathbb{Z}^3)$ represent the set of vertices and edges of G respectively. v_b (v_t) is said to be the



(a) An example of a 17-edge SAW $(2\hat{\imath}, 3\hat{k}, \hat{\imath}, -2\hat{k}, \hat{\imath}, -\hat{k}, \hat{\imath}, 3\hat{k}, \hat{\imath}, \hat{k}, -\hat{\imath})$ starting at u_0 in $\mathbb{Z}^2$.

(b) An example of a 26-edge SAP in $\mathbb{Z}^2$.

Figure 2.1: Examples of SAWs and SAPs in $\mathbb{Z}^2$.

bottom (top) vertex of G if it is the lexicographically smallest (largest) vertex amongst all the vertices in $V(G)$. $e_b = \{v_b, v\}$ ($e_t = \{v_t, v\}$) is said to be the *bottom (top) edge* of G if v is the lexicographically smallest (largest) vertex amongst all the vertices in $V(G) \setminus \{v_b\}$ ($V(G) \setminus \{v_t\}$) that are connected by an edge to v_b (v_t). Let $x = x_1$ ($x = x_2$) be the plane which contains the lexicographically smallest (largest) vertex of G, i.e. v_b (v_t). $x = x_1$ ($x = x_2$) is called the *left-most plane* (*right-most plane*) of G and $b_G := x_2 - x_1$ is said to be the *span* of G.

Similarly, a bottom (top) edge can be defined for any set of edges, E, from $\mathbb{Z}^3$. $e_b = \{v_b, v\} \in E$ ($e_t = \{v_t, v\} \in E$) is said to be the *bottom (top) edge* of E if and only if v_b (v_t) is the lexicographically smallest (largest) vertex amongst all the end vertices of the edges in E and v is the lexicographically smallest (largest) vertex amongst all the vertices that are connected by an edge in E to v_b (v_t). Note that the bottom (top) edge of a given finite graph G is the bottom (top) edge of $E(G)$, the set of edges in G. An edge $e \in E$ is said to be a *horizontal edge* (*vertical edge*) if it is parallel to the y-axis (z-axis). The cartesian coordinates of a vertex $v = (x, y, z)$ are represented respectively by $x(v)$, $y(v)$ and $z(v)$ unless otherwise stated.

Let c_n and p_n be, respectively, the number of n-step SAWs starting at the origin and n-edge SAPs (up to translation) in $\mathbb{Z}^3$. For example, $c_1 = 6$, $c_2 = 30$, $c_3 = 150$,

$c_4 = 726$ and $p_4 = 3$, $p_6 = 22$, $p_8 = 207$. These are known up to $n = 30$ for SAWs and $n = 32$ for SAPs [5]. The quantity

$$\kappa_n := n^{-1} \log c_n \tag{2.1.2}$$

represents the entropy per monomer for the SAW model of a linear polymer, and the limit (as $n \to \infty$) of this quantity, $\kappa := \lim_{n\to\infty} \kappa_n$, has been proved to exist [18]. This entropy reflects the conformational freedom of the polymer. κ is also known as the *connective constant* for SAWs in $\mathbb{Z}^3$ [18]. A numerical estimate for κ is $\log(4.68404) \approx 1.54416$ [5, Table 14].

Because a similar approach will be later used to show the existence of the connective constant for some other lattice objects such as 2SAPs and SSAWs introduced in the next chapters, we discuss next a strategy for the proof of the existence of the connective constant for SAWs. We start with reviewing some definitions and properties related to sequences of real numbers, i.e. $\{a_n\}_{n=1}^{\infty}$, where $a_n \in \mathbb{R}$ for any $n \in \mathbb{N}$.

A sequence $\{a_n\}_{n=1}^{\infty}$ is said to be *sub-additive* (*super-additive*) if for any $n \in \mathbb{N}$ and $m \in \mathbb{N}$,

$$a_{n+m} \leq a_n + a_m \;\; (a_{n+m} \geq a_n + a_m). \tag{2.1.3}$$

Similarly, a sequence $\{a_n\}_{n=1}^{\infty}$ is said to be *sub-multiplicative* (*super-multiplicative*) if for any $n \in \mathbb{N}$ and $m \in \mathbb{N}$,

$$a_{n+m} \leq a_n a_m \;\; (a_{n+m} \geq a_n a_m). \tag{2.1.4}$$

The following lemma is a standard result and is the main ingredient used for the proof of the existence of the connective constant for SAWs in $\mathbb{Z}^3$. A proof of this lemma is given in [32, Lemma 1.2.2].

Lemma 2.1.1. *Let $\{a_n\}_{n=1}^{\infty}$ be a sub-additive sequence of real numbers, then*

$$\lim_{n\to\infty} \frac{a_n}{n} \tag{2.1.5}$$

exists in $[-\infty, \infty)$ and

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\geq 1} \frac{a_n}{n}. \tag{2.1.6}$$

Moreover, if the sequence $\{\frac{a_n}{n}\}_{n=1}^{\infty}$ *is bounded below then the limit* $\lim_{n\to\infty}\frac{a_n}{n}$ *is a finite number.*

Note that the sequence $\{a_n\}_{n=1}^{\infty}$ is super-additive if and only if the sequence $\{-a_n\}_{n=1}^{\infty}$ is sub-additive. Thus, the following result is an immediate consequence of Lemma 2.1.1.

Lemma 2.1.2. *Let* $\{a_n\}_{n=1}^{\infty}$ *be a super-additive sequence of real numbers, then*

$$\lim_{n\to\infty}\frac{a_n}{n} \tag{2.1.7}$$

exists in $(-\infty,\infty]$ *and*

$$\lim_{n\to\infty}\frac{a_n}{n}=\sup_{n\geq 1}\frac{a_n}{n}. \tag{2.1.8}$$

Moreover, if the sequence $\{\frac{a_n}{n}\}_{n=1}^{\infty}$ *is bounded above then the limit* $\lim_{n\to\infty}\frac{a_n}{n}$ *is a finite number.*

Note that the sequence $\{a_n\}_{n=1}^{\infty}$ of real positive numbers is sub-multiplicative (super-multiplicative) if and only if the sequence $\{\log a_n\}_{n=1}^{\infty}$ is sub-additive (super-additive). Hence, the following result is a generalization of Lemma 2.1.2 which will be used for showing the existence of the connective constant for some lattice objects discussed in the next chapters.

Lemma 2.1.3 (Wilker and Whittington 1979 [64])**.** *Suppose* $\{a_n\}_{n=1}^{\infty}$ *is a non-decreasing sequence of positive numbers such that* $\{n^{-1}\log a_n\}_{n=1}^{\infty}$ *is bounded above and* $a_{n+f(m)}\geq a_n a_m$ *for some positive function* f *which satisfies* $\lim_{m\to\infty}m^{-1}f(m)=1$. *Then the limit* $\lim_{n\to\infty}n^{-1}\log a_n$ *exists and is finite.*

The following lemma will be also used later, in Section 6.6, to show that the super-additive sequence of real numbers associated to some lattice objects is bounded above.

Lemma 2.1.4 (Madras *et al* 1988 [34])**.** *For positive rational numbers* a *and* b *such that* $a>b$,

$$\lim_{n\to\infty}n^{-1}\log\binom{an}{bn}=a\log a-b\log b-(a-b)\log(a-b). \tag{2.1.9}$$

In order to prove the existence of the connective constant for SAWs we need to show that the sequence $\{\log c_n\}_{n=1}^{\infty}$ is sub-additive, or equivalently show that $\{c_n\}_{n=1}^{\infty}$ is sub-multiplicative. The *concatenation argument* for SAWs, introduced below, will be used to prove the super-multiplicative property of $\{c_n\}_{n=1}^{\infty}$.

Let w^1 and w^2 be two SAWs in $\mathbb{Z}^3$, where

$$w^1 := u_0^1, u_1^1, u_2^1, ..., u_{n-1}^1, u_n^1 \tag{2.1.10}$$

and

$$w^2 := u_0^2, u_1^2, u_2^2, ..., u_{m-1}^2, u_m^2. \tag{2.1.11}$$

The *concatenation* of w^2 to w^1, $w^1 \circ w^2$, is the $(n+m)$-step walk

$$w := u_0, u_1, u_2, ..., u_{n+m-1}, u_{n+m}, \tag{2.1.12}$$

where

$$u_k = u_k^1, \qquad k = 0, ..., n, \tag{2.1.13}$$

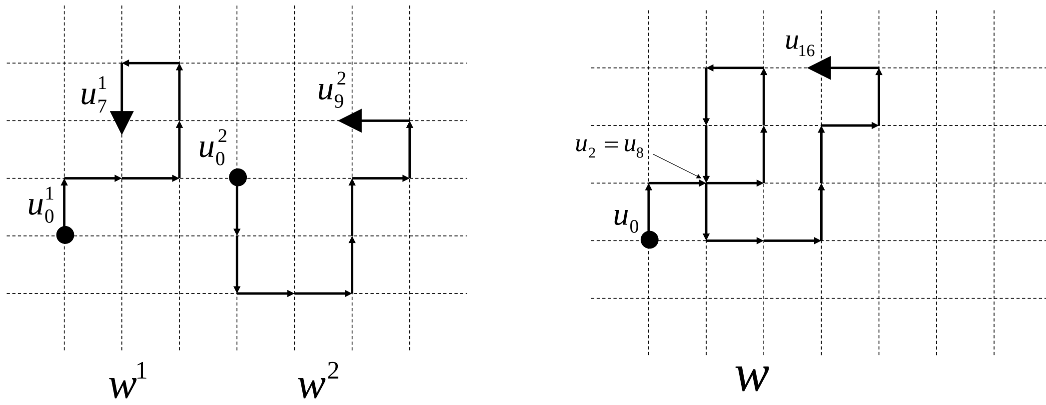
$$u_k = u_n^1 + u_{(k-n)}^2 - u_0^2, \qquad k = n+1, ..., n+m. \tag{2.1.14}$$

Note that $w := w^1 \circ w^2$ is not necessarily a self-avoiding walk. However, due to the construction, it is self-avoiding for the initial n steps and the final m steps. Figure 2.2 shows an example of concatenating two self-avoiding walks in $\mathbb{Z}^2$. Note that many of the results presented in this section hold also for $\mathbb{Z}^d$, $d \geq 2$, but here we will focus on $d = 3$ except that sometimes for illustration $\mathbb{Z}^2$ will be used.

The existence of the connective constant for SAWs is now a result of the above concatenation argument and Lemma 2.1.1.

Theorem 2.1.5 (Hammersley and Morton 1954 [18])**.** *The following limit exists and is finite.*

$$\kappa \equiv \lim_{n\to\infty} n^{-1} \log c_n. \tag{2.1.15}$$



(a) The SAWs w^1 and w^2 in $\mathbb{Z}^2$.

(b) The walk $w := w^1 \circ w^2$ in $\mathbb{Z}^2$.

Figure 2.2: An example of the concatenation of two SAWs w^1 and w^2 in $\mathbb{Z}^2$ which results in the walk $w := w^1 \circ w^2$ that is not a self-avoiding walk.

Proof. The concatenation process for SAWs in $\mathbb{Z}^3$ implies that the product $c_n c_m$ is equal to the number of elements in the set of $(n+m)$-step walks which are self-avoiding for the initial n steps and the final m steps, but are not necessarily self-avoiding walks. Moreover, every $(n+m)$-step SAW can be obtained by the concatenation of an m-step SAW to an n-step SAW. Hence

$$c_{n+m} \leq c_n c_m, \tag{2.1.16}$$

i.e. the sequence $\{c_n\}_{n=1}^{\infty}$ is sub-multiplicative. Taking logarithms in this inequality results in

$$\log c_{n+m} \leq \log c_n + \log c_m, \tag{2.1.17}$$

which shows that the sequence $\{\log c_n\}_{n=1}^{\infty}$ is sub-additive. By Lemma 2.1.1, this proves that $\lim_{n\to\infty} n^{-1} \log c_n < \infty$; however this limit cannot equal $-\infty$ since $c_n \geq 1$, for any $n \geq 1$, implies that $\frac{\log c_n}{n} \geq 0$. $\square$

An n-step *bridge* is an n-step self-avoiding walk $w := u_0, u_1, ..., u_{n-1}, u_n$ whose vertices satisfy the inequality

$$x(u_0) < x(u_i) \leq x(u_n) \tag{2.1.18}$$

for $1 \leq i \leq n$ [32]. Let b_n denote the number of n-step bridges starting at the origin in $\mathbb{Z}^3$. The following theorem holds.

Theorem 2.1.6 (Hammersley and Welsh 1961 [20])**.**

$$\kappa_b = \kappa \tag{2.1.19}$$

where

$$\kappa_b \equiv \lim_{n\to\infty} n^{-1} \log b_n \tag{2.1.20}$$

is the connective constant for bridges in $\mathbb{Z}^3$*.*

A similar statement has also been proved for SAPs in $\mathbb{Z}^3$. It has been proved that the connective constant for SAPs exists and is exactly equal to κ:

Theorem 2.1.7 (Hammersley 1961 [19])**.**

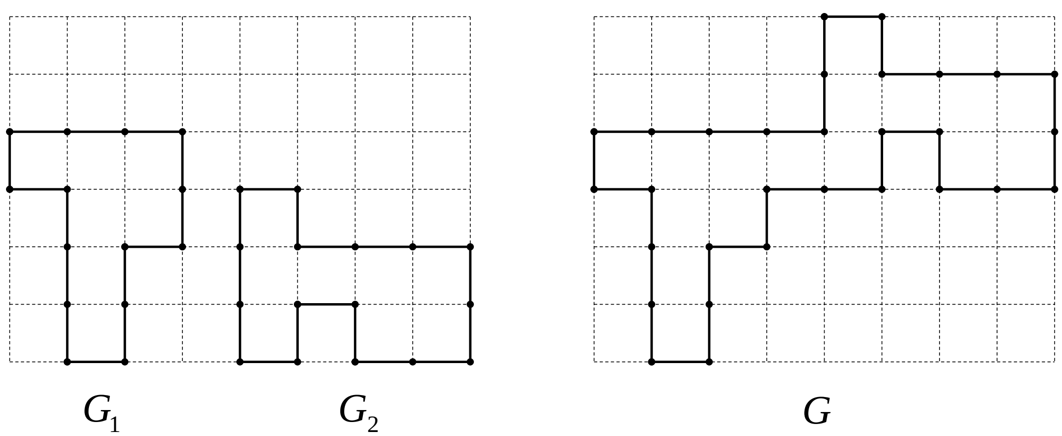
$$\lim_{n\to\infty} n^{-1} \log p_n = \kappa, \tag{2.1.21}$$

where the limit is taken over all the non-zero values of p_n*; i.e.* $n \geq 4$ *and even.*

The strategy for the proof of the existence of the connective constant for SAPs is similar to that presented in Theorem 2.1.5 for SAWs. In fact, the proof results from introducing an appropriate concatenation argument for SAPs in $\mathbb{Z}^3$ which then leads to the super-multiplicative property for the sequence $\{\frac{p_n}{2}\}_{n=4}^{\infty}$. Given an n-edge polygon G_1 and an m-edge polygon G_2, the concatenation of these two polygons $G := G_1 \circ G_2$ is defined as follows (as an example see Figure 2.3). Note first that the top edge of G_1, $e_t = \{v_t, v\}$ (v_t is the top vertex of G_1), and the bottom edge of G_2, e_b, each lie in a 2-dimensional space perpendicular to the x-axis. So e_t (e_b) is either a horizontal edge or a vertical edge. Translate and rotate (if necessary) the polygon G_2 about the x-axis so that its bottom edge is parallel to the top edge of G_1 and its bottom vertex satisfies $x(v_b) = x(v) + 1$, $y(v_b) = y(v)$ and $z(v_b) = z(v)$. Then remove the top edge of G_1 and the bottom edge of G_2 and add back two edges to join the two polygons into a single new polygon. Note that if originally e_t and e_b were parallel then G_2 only needs to be translated and not rotated. Otherwise, G_2 has to be rotated first and then concatenated to G_1. So there are two distinct choices for G_2, G' and G'', where G'' is a rotated version of G' and $G_1 \circ G' = G_1 \circ G''$. Thus, two distinct polygons G' and G'' result in the same concatenated polygon and hence there are $p_m/2$ choices for G_2 and p_n choices for G_1. Therefore,

$$p_n p_m/2 \leq p_{n+m} \tag{2.1.22}$$

which implies that the sequence $\{\frac{p_n}{2}\}_{n=4}^{\infty}$ is super-multiplicative.



(a) The 14-edge polygon G_1 and the 16-edge polygon G_2 in $\mathbb{Z}^2$.

(b) The 30-edge polygon $G = G_1 \circ G_2$ in $\mathbb{Z}^2$.

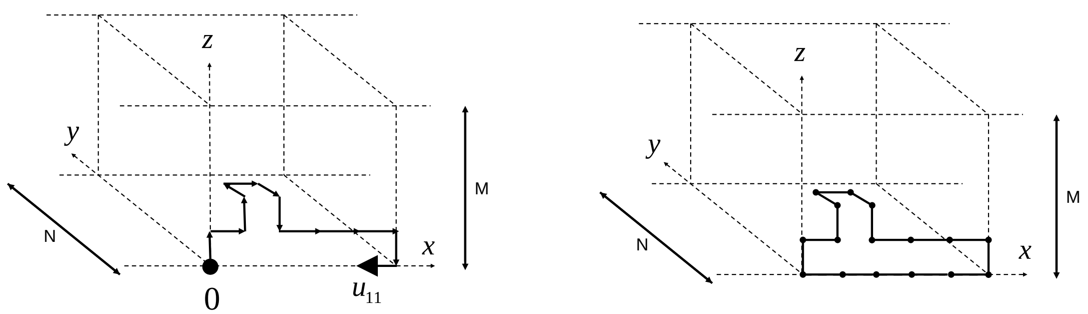
Figure 2.3: An example of concatenating two polygons in $\mathbb{Z}^2$. Note that in $\mathbb{Z}^2$ no rotations are required.

2.2 The Connective Constant for SAWs (SAPs) in Tubes

In this section we will present some results regarding the connective constant for SAWs and SAPs in a tube, considered as a sublattice of $\mathbb{Z}^3$, and discuss the effects of the confinement to a tube on the asymptotic behaviour of SAWs and SAPs.

An (N, M)-*tube*, $T(N, M)$, is the subgraph of $\mathbb{Z}^3$ bounded by the planes $y = 0$, $y = N$, $z = 0$ and $z = M$, i.e. it is the subgraph induced by the vertex set

$\{(x, y, z) \in \mathbb{Z}^3 | \ 0 \leq y \leq N, \ 0 \leq z \leq M\}$. SAPs (SAWs) behave differently in a tube than in $\mathbb{Z}^3$ in the sense that the connective constant for SAPs is no longer the same as that for SAWs. Let $c_n(N, M)$ be the number of n-edge SAWs in an (N, M)-tube (up to translation). Let also $p_n(N, M)$ be the number of n-edge SAPs confined to an (N, M)-tube (up to x-translation); see Figure 2.4 for examples of SAWs and SAPs in a tube. The proof of the existence of the connective constant for SAWs and SAPs in a tube is also based on an appropriate concatenation argument. Note that concatenations introduced for SAWs and SAPs in $\mathbb{Z}^3$ do not necessarily work for SAWs and SAPs in a tube since the lattice objects are constrained to stay within a tube. The following lemmas establish the existence of the connective constant for SAWs and SAPs in $T(N, M)$.



(a) An example of a 12-edge SAW in an (N, M)-tube. Note that the end point of the walk is marked by a large arrow.

(b) An example of a 16-edge SAP in an(N, M)-tube.

Figure 2.4: Examples of SAWs and SAPs in a tube.

Theorem 2.2.1 (Soteros and Whittington 1989 [49]). *The following limit exists and is finite:*

$$\kappa(N, M) \equiv \lim_{n\to\infty} n^{-1} \log c_n(N, M). \tag{2.2.1}$$

Theorem 2.2.2 (Soteros and Whittington 1989 [49]). *Given* $(N, M) \neq (0, 0)$*, the* *following limit exists and satisfies the following inequality:*

$$\kappa_p(N,M) \equiv \lim_{n\to\infty} n^{-1} \log p_n(N,M) < \kappa(N,M) = \lim_{n\to\infty} n^{-1} \log c_n(N,M). \qquad (2.2.2)$$

Let $b_n(N,M)$ denote the number of n-step bridges starting at the origin in $T(N,M)$. The following theorem shows that a similar result to Theorem 6.7 is also satisfied in a tube.

Theorem 2.2.3 (Whittington 1983 [63]).

$$\kappa_b(N,M) = \kappa(N,M) \qquad (2.2.3)$$

where

$$\kappa_b(N,M) \equiv \lim_{n\to\infty} n^{-1} \log b_n(N,M) \qquad (2.2.4)$$

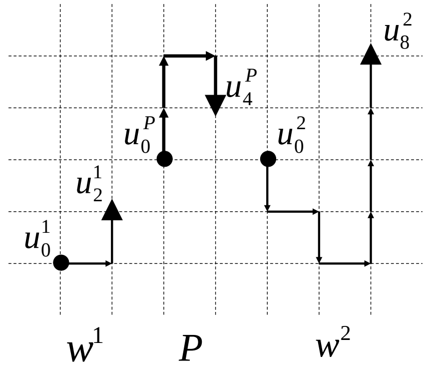
is the connective constant for bridges in $T(N,M)$.

2.3 Kesten's Pattern Theorem for SAWs (SAPs) in $\mathbb{Z}^3$

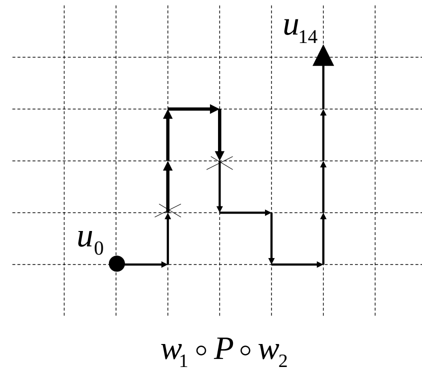
Here we will introduce the concept of Kesten patterns and discuss Kesten's pattern theorem which will allow us to review results on the knotting probability of lattice polygons in the next chapter.

Given any integer number $d \geq 2$, any (relatively short) self-avoiding walk in the lattice $\mathbb{Z}^d$ is called a *SAW pattern*. Given positive integer numbers m and n such that $m \leq n$, let P be an m-edge SAW pattern (or equivalently a SAW) and w be an n-edge SAW. We say the pattern P appears in the SAW w if $w = w_1 \circ P \circ w_2$ for some SAWs w_1 and w_2. Let $n_1 \geq 0$ ($n_2 \geq 0$) denote the number of edges in w_1 (w_2) then $n = n_1 + m + n_2$. Note that w_1 (w_2) is also allowed to be an *empty walk* defined as a single-vertex walk with no edge in it, i.e. $n_1 = 0$ ($n_2 = 0$). In particular, we say P has occurred at the *start* (*end*) of w if w_1 (w_2) is an empty walk, i.e. $n_1 = 0$

($n_2 = 0$). We also say P has occurred in the middle of w if both w_1 and w_2 are non-empty walks (e.g. see Figure 2.5).



(a) The 2-step SAW w_1, P and the 8-step SAW w_2.



(b) The SAW $w = w_1 \circ P \circ w_2$ resulted from the concatenation of w_1, P and w_2.

Figure 2.5: An example of an occurrence of the pattern P in the middle of a 14-step SAW.

A SAW pattern in $\mathbb{Z}^d$ is called a *Kesten Pattern* if it appears at least three times in a SAW in $\mathbb{Z}^d$, or equivalently appears in the middle of a SAW (e.g. see Figure 2.6 (a)). Three times occurrence of the pattern is needed to exclude those patterns which can occur only at the start (end) of SAWs; examples of such patterns for $\mathbb{Z}^2$ are illustrated in Figure 2.6 (b), (c) and (d).

Let $c_n(P)$ ($c_n(\bar{P})$) be the number of n-step SAWs starting at the origin in $\mathbb{Z}^3$ which contain (do not contain) the pattern P. The following pattern theorem for SAWs is due to Kesten.

Theorem 2.3.1 (Kesten 1963 [28]). *Let P be a Kesten pattern, then*

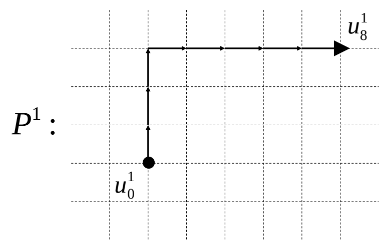
$$\kappa(\bar{P}) < \kappa, \tag{2.3.1}$$

where

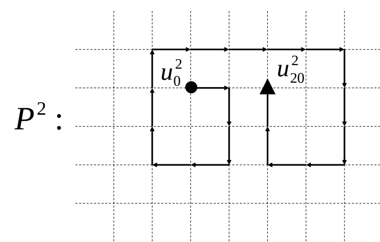
$$\kappa(\bar{P}) \equiv \lim_{n\to\infty} n^{-1} \log c_n(\bar{P}). \tag{2.3.2}$$

Hence, for n sufficiently large, all but exponentially few SAWs contain the pattern P.

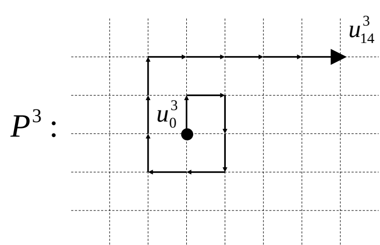
By Theorem 2.1.7, the connective constant for SAPs in $\mathbb{Z}^3$ is the same as that for SAWs. Note also that every SAP with its bottom vertex fixed at the origin and its bottom edge removed and an orientation added on the edges starting at the bottom vertex results in a SAW starting at the origin. Therefore, Kesten's pattern theorem introduced in Theorem 2.3.1 also yields a pattern theorem for SAPs in $\mathbb{Z}^3$ [55].



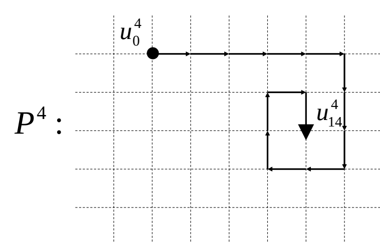
(a) P^1 is an example of a Kesten pattern in $\mathbb{Z}^2$.



(b) P^2 is an example of a non-Kesten pattern which cannot occur in the middle of any SAW; note that this pattern also cannot occur at the start or end of any SAW other than itself.



(c) P^3 is an example of a non-Kesten pattern which cannot occur in the middle of any SAW; note that this pattern also cannot occur at the end of any SAW other than itself but it can occur at the start of SAWs.



(d) P^4 is an example of a non-Kesten pattern which cannot occur in the middle of any SAW; note that this pattern also cannot occur at the start of any SAW other than itself but it can occur at the end of SAWs.

Figure 2.6: Examples of Kesten and non-Kesten patterns in $\mathbb{Z}^2$.

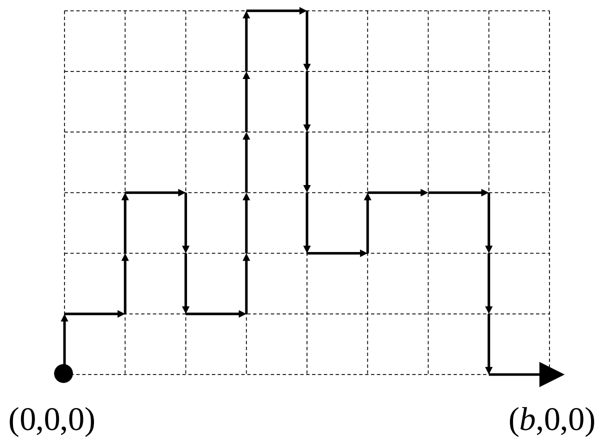


Figure 2.7: An example of a $\mathcal{K}_8$ pattern in $T(0,6)$; $b = 8$.

2.4 Pattern Theorems for SAWs (SAPs) in Tubes

In this section we will present pattern theorems for SAWs and SAPs in a tube $T(N, M)$. In order to do this, we first need to introduce some special SAW patterns in $T(N, M)$, which will correspond to Kesten patterns in the lattice $\mathbb{Z}^3$.

Let $\mathcal{C}_b(N, M)$ be the set of SAWs contained in $\mathcal{D}_b(N, M) = \{(x, y, z) \in \mathbb{Z}^3 \mid\ 0 \leq x \leq b,\ 0 \leq y \leq N,\ 0 \leq z \leq M\}$ which have one endpoint at the origin and the other at $(b, 0, 0)$. For $b > 0$, any element of $\mathcal{C}_b(N, M)$ is called a $\mathcal{K}_b$ pattern (e.g. see Figure 2.7). Let $c_n(N, M; \bar{P})$ denote the number of n-edge SAWs in $T(N, M)$ (up to translation) which do not contain the pattern P. The following result holds.

Theorem 2.4.1 (Soteros and Whittington 1989 [49])**.** *For any $b > 0$, let P be a $\mathcal{K}_b$ pattern, then*

$$\kappa(N, M; \bar{P}) < \kappa(N, M), \tag{2.4.1}$$

where

$$\kappa(N, M; \bar{P}) \equiv \lim_{n\to\infty} n^{-1} \log c_n(N, M; \bar{P}). \tag{2.4.2}$$

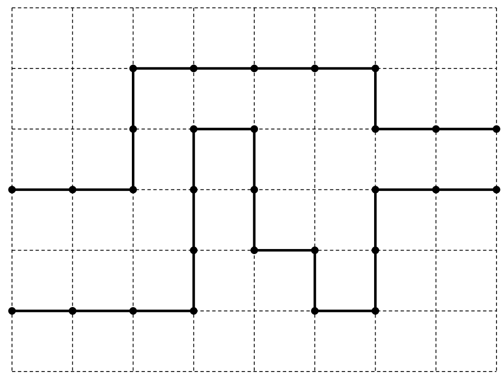


Figure 2.8: An example of a $\tilde{K}_8$ pattern in $T(0,6)$.

Unlike the situation for SAWs and SAPs in $\mathbb{Z}^3$, the pattern theorem for SAWs in a tube cannot be used for SAPs in a tube. This is because the connective constant for SAPs is different from that for SAWs in a tube. The required pattern theorem for SAPs is introduced as follows.

Define $\mathcal{P}(N,M)$ to be the set of SAPs in an (N,M)-tube. Let $\mathcal{P}_b(N,M) \subseteq \mathcal{P}(N,M)$ be the set of SAPs with left-most plane $x=0$ and right-most plane $x=b$. For any $b>0$, a $\tilde{K}_b$ pattern is defined to be a configuration (including both occupied and unoccupied edges) of any element of $\mathcal{P}_b(N,M)$ with the edges in the $x=0$ and $x=b$ planes excluded (e.g. see Figure 2.8). Let $p_n(N,M;\bar{P})$ denote the number (up to x-translation) of n-edge SAPs in $T(N,M)$ which do not contain the pattern P.

Theorem 2.4.2 (Soteros 1998 [51])**.** *For any integer $b \geq 2$ and any $\tilde{K}_b$ pattern P,*

$$\kappa_p(N,M;\bar{P}) < \kappa_p(N,M) < \kappa(N,M), \tag{2.4.3}$$

where

$$\kappa_p(N,M;\bar{P}) \equiv \lim_{n\to\infty} n^{-1}\log p_n(N,M;\bar{P}). \tag{2.4.4}$$

(The last inequality in equation 2.4.4 follows from Theorem 2.2.2.)

The proof of this theorem in [51] relied on a transfer-matrix argument; an alternative proof is also given in [53]. The transfer-matrix and its main properties are introduced in Appendix A, and a similar argument to that given for the proof of Theorem 2.4.2 in [51] will be later used to generalize the transfer-matrix argument for certain types of clusters, including SSAWs and stretched polygons, in $T(N, M)$.

2.5 The Pattern Insertion Strategy

In this section the so called *pattern insertion* argument used to obtain pattern theorems for certain types of clusters, such as SAPs, in an (N, M)-tube is reviewed from [33, 53]. This strategy, in the next chapters, will then be employed to establish pattern theorems for some lattice objects under consideration such as 2SAPs, SSAWs and stretched polygons.

In order to obtain the required pattern theorem for clusters in tubes via [33, Theorem 2.1], the cluster axioms (CA1), (CA2) and (CA4) of [33] must hold for any cluster under investigation. These axioms are reviewed here as follows.

Cluster axiom (CA1) basically defines the types of clusters that can be considered. For the purpose here, given any integers $N \geq 0$ and $M \geq 0$ let the lattice $T = T(N, M)$. For each positive integer n, let C_n be the set of all clusters of size n. Examples of "cluster" include SAPs, 2SAPs and SSAWs. Also examples of "size" are the number of edges or the span of a cluster. Let

$$S^* \equiv \{u \in \mathbb{R}^3 \mid T + u = T\}, \tag{2.5.1}$$

where $T + u$ is the translation of T by the vector $u \in \mathbb{R}^3$. Clearly for $T = T(N, M)$,

$$S^* = \{u \in \mathbb{R}^3 \mid T(N, M) + u = T(N, M)\} = \{(x, 0, 0) \mid x \in \mathbb{Z}\}. \tag{2.5.2}$$

The following is the statement of (CA1) as presented in [33]:

(CA1): *C_n is a collection of finite subgraphs of T that is invariant under translation (i.e., if $G \in C_n$ and $u \in S^*$, then $G + u \in C_n$). The C_n's are pairwise disjoint (i.e., $C_n \cap C_m = \emptyset$ whenever $n \neq m$).*

Before stating the cluster axiom (CA2) the definition of a weight function which assigns positive weight to each cluster needs to be reviewed. Given $C := \bigcup_{n=1}^{\infty} C_n$, the weight function

$$wt : C \to (0, \infty), \tag{2.5.3}$$

is defined so that it is invariant under translation, i.e.

$$wt(G) = wt(G + u) \quad \text{for every } u \in S^* \text{ and } G \in C. \tag{2.5.4}$$

The cluster axiom (CA2), as given in [33], is as follows:

(CA2): *For each positive integer k, there is a finite positive constant γ_k with the property that*

$$\frac{1}{\gamma_k} wt(G) \leq wt(G') \leq \gamma_k wt(G), \tag{2.5.5}$$

whenever G and G' differ by at most k vertices and edges (i.e. whenever $\mid E(G) \triangle E(G') \mid + \mid V(G) \triangle V(G') \mid \leq k$).

Note that $\triangle$ here denotes a symmetric difference; i.e. for any pair of sets A and B, $A \triangle B = (A \setminus B) \cup (B \setminus A)$. An example of a weight function satisfying (CA2) is the constant function, i.e. $wt(G) = 1$ for any cluster G.

For each positive integer n, let C_n^* be the set of all clusters with size n whose lexicographically smallest vertex is in the plane $x = 0$, i.e. $x_1 = 0$. Define $C^* = \bigcup_{n<\infty} C_n^*$. Given any cluster $G \in C_n$, note that $x_1 \in \mathbb{Z}$ is such that the translation of G along the x-axis, $G + (-x_1, 0, 0)$, gives an element of C_n^*, i.e. $G + (-x_1, 0, 0) \in C_n^*$. For any finite set $A \subset C$, define

$$\mathcal{G}(A) = \sum_{G \in A} wt(G). \tag{2.5.6}$$

In particular, let $\mathcal{G}_n = \mathcal{G}(C_n^*)$. Let also

$$\lambda \equiv \limsup_{n\to\infty} (\mathcal{G}_n)^{\frac{1}{n}}. \tag{2.5.7}$$

The third cluster axiom is as follows:

(CA3): *The limit $\lim_{n\to\infty}(\mathcal{G}_n)^{\frac{1}{n}}$, called the growth constant for the clusters under consideration, exists and is finite (and equals λ).*

Note that (CA3) is not required for proving the pattern theorem here. However, if it holds, the connective constant for the clusters under consideration exists and is finite.

In order to state the cluster axiom (CA4), some definitions and terminology about patterns need to be introduced. For a non-negative integer b, define V_b to be the subgraph of $T(N, M)$ generated by the vertex set $\{(x, y, z) \in \mathbb{Z}^3 \mid 0 \leq x \leq b,\ 0 \leq y \leq N,\ 0 \leq z \leq M\}$; this is an example of a *section* with span b of $T(N, M)$. We next define the meaning of the term *pattern* as in [53]. Given any non-negative integer b, any subgraph of V_b with at least one vertex in each of the planes $x = 0$ and $x = b$ is referred to as a pattern with *span* b, or *b-pattern* for short. We say a cluster G *contains* a b-pattern P if P is the subgraph of G generated by the vertices in V_b, i.e. $G \cap V_b = P$, or if there exists a translate of V_b, $(V_b + (x, 0, 0))$, such that $G \cap (V_b + (x, 0, 0)) = P + (x, 0, 0)$; in the latter case we say P *occurs* at $(x, 0, 0)$ in G. A b-pattern P is said to be a *proper b-pattern* (see [53]) with respect to C^* if

i) there are infinitely many values of n such that P is contained in some cluster in C^*_n, and

ii) there exists a cluster G in which P occurs at some $t \in S^*$ and such that $G \setminus (V_b + t)$ still contains the left-most and right-most planes of G.

Condition (ii) is needed to exclude patterns which can only occur at the left-most or the right-most plane of a cluster and nowhere else. A pattern P for which there exists a cluster G having left-most plane $x_1 = 0$ and containing P at (0,0,0) is referred to as a *start pattern*. If P is a start pattern for G, then P is said to occur at the start or at the left-most plane of G. Similarly, a b-pattern P for which there exists a cluster G having right-most plane $x_2 = b$ and containing P at $(0, 0, 0)$ is referred to as an *end pattern*. If P is an end pattern for G, then P is said to occur at the end or at the right-most plane of G.

The fourth cluster axiom says that any part of any cluster can be locally changed to create an occurrence of some translate of a given proper pattern.

(CA4): *For every proper pattern P there exists an integer $b \geq 0$ and translation vector $b' \in S^*$ such that: For every cluster $G \in C$ and every vertex v of G, there is another cluster $G' \in C$ and a translation vector $t = t(v) \in S^*$ such that $v \in V_b + b' + t$, G' contains P at t, and $G' \setminus (V_b + b' + t) = G \setminus (V_b + b' + t)$.*

Next we shall review the two tube axioms (TA1) and (TA2) which were introduced in [53] specifically for applying the pattern insertion strategy to clusters in tubes. In fact, it is shown in [53] that for clusters in tubes the cluster axiom (CA4) can be replaced by the tube axioms (TA1) and (TA2).

The following are the tube axioms as introduced in [53]:

(TA1) (CONCAT): *There exists a concatenation process for C in $T = T(N, M)$ and associated integers $t_T \geq 0$, $c_T \geq 0$ such that: Given $G_1 \in C^*_n$ and $G_2 \in C^*_m$ with spans b_1 and b_2, concatenating G_1 to the translate of G_2, $G_2 + (t_T + b_1, 0, 0)$, forms $G \in C^*_{n+m+c_T}$ such that $G \cap V_{b_1-1} = G_1 \cap V_{b_1-1}$ and $G \cap (V_{b_2-1} + (t_T + b_1 + 1, 0, 0)) = G_2 \cap (V_{b_2-1} + (1, 0, 0))$ (i.e. only the right-most plane of G_1 and the left-most plane of G_2 can be altered in the concatenation process).*

(TA2) (CAPOFF): *There exists an integer $m_T > 0$ such that: For any integer $b \geq 0$ and any b-pattern (not necessarily proper) P that occurs at $(0, 0, 0)$ in some finite size cluster in C^* with span $s \geq b + 1$ (i.e. P occurs at the start of some cluster but is not itself a cluster), there exists a cluster $G \in C^*$ with span $b + m_T$ which also contains P at $(0, 0, 0)$ (i.e. P is also at the start of G). Similarly, given any b-pattern P' that occurs at $(s - b, 0, 0)$ in some finite size cluster in C^* with span $s \geq b + 1$ (i.e. P' occurs at the end of some cluster but is not itself a cluster), there exists a cluster $G' \in C^*$ with span $b + m_T$ which contains P' at $(m_T, 0, 0)$ (i.e. P' also ends G').*

(CONCAT) and (CAPOFF) together allow one to insert any proper pattern P into an arbitrary cluster at an arbitrary location. Figure 2.9 illustrates an example where proper 2SAP pattern P is inserted into 2SAP G; note that this procedure will be explained in detail later when we discuss 2SAPs in Chapter 4. In particular (CONCAT) and (CAPOFF) combined with (CA1) and (CA2) immediately give us the following proposition.

Proposition 2.5.1 (Soteros 2006 [53])**.** *Given $T = T(N, M)$ and a set of clusters C of T which satisfy (CA1) of [33] and (TA1) and (TA2), for every proper pattern P there exists an integer $b \geq 0$ and translation vector $b' \in S^*$ such that:*

For every cluster $G \in C$ and every vertex v of G, there is another cluster $G' \in C$ and a translation vector $t = t(v) \in S^*$ such that $v \in V_b + b' + t$, G' contains P at t, and $G' \setminus (V_b + b' + t) = G \setminus (V_b + b' + t)$ (see Figure 2.9).

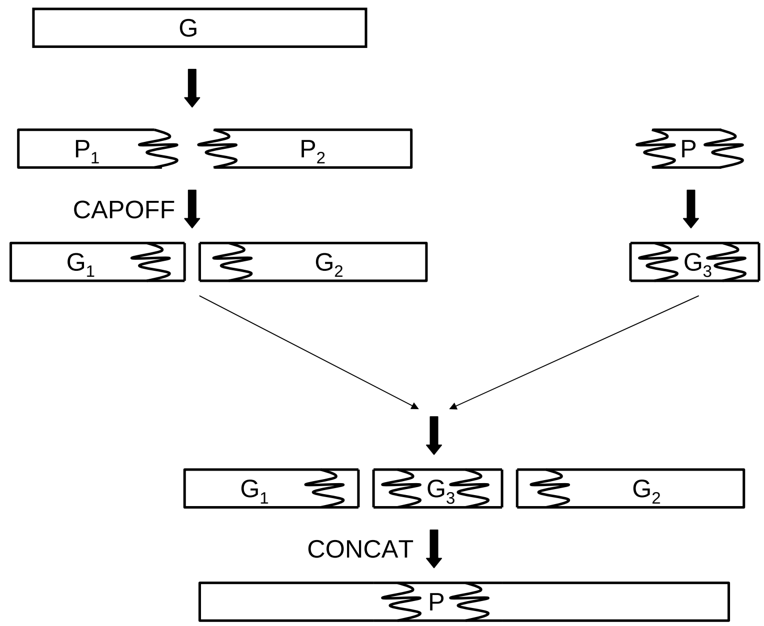


Figure 2.9: For cluster G (top left) and proper pattern P (top right), a sketch of the pattern insertion algorithm for Proposition 2.5.1 is shown: cluster G is broken into the start pattern P_1 and the end pattern P_2. P_1 and P_2 are capped off using (CAPOFF) to create clusters G_1 and G_2. Then G_1, G_3 (the cluster shown below P on right), and G_2 are concatenated using (CONCAT) to create a cluster which contains P.

Let $C_n^*[\leq m, P] \subseteq C_n^*$ denote the set of clusters in C_n^* that contain at most m translates of the pattern P. Then define

$$\mathcal{G}_n[\leq m, P] = \sum_{G \in C_n^*[\leq m, P]} wt(G). \tag{2.5.8}$$

Theorem 2.5.2 (Madras 1999 [33]). *Assume that Cluster Axioms (CA1), (CA2) and (CA4) hold. Let P be a proper pattern. Then there exists an $\epsilon_P > 0$ such that*

$$\limsup_{n\to\infty} (\mathcal{G}_n[\leq \epsilon_P n, P])^{\frac{1}{n}} < \lambda. \tag{2.5.9}$$

The following corollary is an immediate result of Proposition 2.5.1 and Theorem 2.5.2, and yields a pattern theorem for clusters in a tube.

Corollary 2.5.3 (Soteros 2006 [53]). *Assume that Cluster Axioms (CA1), (CA2) and Tube Axioms (TA1) and (TA2) hold. Let P be a proper pattern. Then there exists an $\epsilon_P > 0$ such that*

$$\limsup_{n\to\infty}(\mathcal{G}_n[\leq \epsilon_P n, P])^{\frac{1}{n}} < \lambda. \tag{2.5.10}$$

2.6 Transfer-Matrix Results for Clusters in Tubes

In this section the transfer-matrix argument, used by Soteros [51] to prove a pattern theorem for SAPs in a tube, is generalized to enable the exploration (in chapters 5 and 6) of the asymptotic behaviour of certain types of clusters such as stretched polygons and SSAWs in tubes. Given a positive integer t, a set of clusters C and a vector of variables $Y = [y_1, ..., y_t]^T$, let

$$Z_n(N, M; Y) = \sum_{G\in C_n^*} \prod_{j=1}^{t} y_j^{d(G,j)} \tag{2.6.1}$$

be a partition function for clusters in C, where n represents the size of the clusters in C and the $d(G, j)$'s are non-negative integers associated with a cluster G in C. Note that "size" of a cluster can be measured in different ways; for example, it can be measured by the number of edges or the span of the cluster. The $d(G, j)$'s also can be defined as some quantities, such as the number of edges, the span or the number of vertices with degree one, associated with a cluster G. The goal is to find the asymptotic form of $Z_n(N, M; Y)$ as $n \to \infty$.

An (N, M)-tube is equivalent to $\mathbb{Z} \times H(N, M)$, where $H(N, M)$ is the finite subgraph of $\mathbb{Z}^2$ induced by the vertex set $\{(y, z) \in \mathbb{Z}^2 | 0 \leq y \leq N, 0 \leq z \leq M\}$, and hence, in the terminology of [1], an (N, M)-tube is a one-dimensional lattice. So an (N, M)-tube can also be considered as an alternating sequence of *hinges* and *sections* where for any $i \in \mathbb{Z}$ the ith hinge $H_i(N, M)$ is defined to be the subgraph of the tube induced by the vertex set $\{(i, y, z) \in \mathbb{Z}^3 | 0 \leq y \leq N, 0 \leq z \leq M\}$ and the ith section, $S_i(N, M)$, is defined to be the set of edges which join $H_{i-1}(N, M)$ to $H_i(N, M)$ in the tube.

Given a cluster G with left-most (right-most) plane $x = 0$ ($x = m$) in an (N, M)-tube, a *cluster configuration* with span k (k-cluster config) of G can be considered as G's *configuration* in a sublattice of the form $H_{i-1}(N, M) \cup S_i(N, M) \cup H_i(N, M) \cup ... \cup S_{i+k-1}(N, M) \cup H_{i+k-1}(N, M)$ for some $1 \leq i \leq m - k + 1$. We also say the cluster config of span k corresponding to G's configuration in $H_{i-1}(N, M) \cup S_i(N, M) \cup H_i(N, M) \cup ... \cup S_{i+k-1}(N, M) \cup H_{i+k-1}(N, M)$ *occurs* at the ith section of G. G's *configuration* in such a sublattice of the tube consists of the sublattice and a specific assignment of an ordering and labelling on the edges of G in the sublattice. An appropriate ordering and labelling of the edges for SAPs (SSAWs) is considered and explained in detail in Chapter 5 (Chapter 6) so that a k-cluster config at the ith section of G is defined by not just the edges of G in the section but also their labelling and relative ordering, according to the order on the edges in G. Two k-cluster configs are considered *equivalent* if they have the same set of occupied vertices and edges, and have the same labelling and relative ordering on their edges. Note that ignoring the ordering and labelling on the edges of a k-cluster config results in a k-pattern (as defined in Section 2.5). In fact, depending on the ordering and labelling on a pattern's edges, a pattern can correspond to more than one distinct cluster configuration. The extra information on the edges is needed in order to ensure that the Cluster Configuration Axiom, defined below, is satisfied. Note that start, end and proper cluster configurations are also defined in the exact same way that start, end and proper patterns were defined in the previous section.

Given $k \geq 2$, a set of clusters C and three sets of k-cluster configurations $\Pi(k)$, $\Pi_1(k)$ and $\Pi_2(k)$ corresponding to the proper, start and end k-cluster configs respectively, define a digraph $D_k^0 = (V_k^0, A_k^0, \phi_k^0)$ as follows: Let the k-cluster configs in $\Pi(k) \cup \Pi_1(k) \cup \Pi_2(k)$ be the vertices of the digraph, i.e.

$$V_k^0 = \Pi(k) \cup \Pi_1(k) \cup \Pi_2(k) = \{P_1^0, P_2^0, ..., P_{|V_k^0|}^0\}. \tag{2.6.2}$$

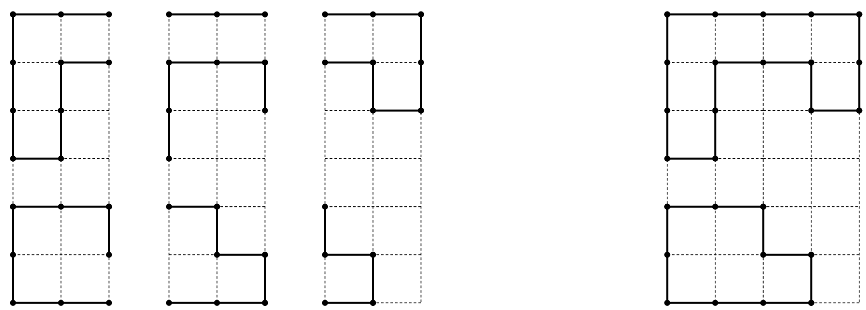
An arc from P_i^0 to P_j^0 belongs to A_k^0 if and only if the configuration of the k-cluster config P_i^0 on $H_1(N,M) \cup S_2(N,M) \cup ... \cup S_k(N,M) \cup H_k(N,M)$ is equivalent to the configuration of the k-cluster config P_j^0 on $H_0(N,M) \cup S_1(N,M) \cup ... \cup S_{k-1}(N,M) \cup H_{k-1}(N,M)$. In this case the function $\phi_k^0 : A_k^0 \to V_k^0 \times V_k^0$ is one to one so for simplicity we can represent any arc from P_i^0 to P_j^0 by the ordered pair (P_i^0, P_j^0). $\Pi(k)$, $\Pi_1(k)$ and $\Pi_2(k)$ are then said to satisfy the Cluster Configuration Axiom if the following holds:

Cluster Configuration Axiom: *Given $r \geq 1$, $1 \leq i_j \leq |V_k^0|$ for $1 \leq j \leq r$, consider a walk of the digraph with length r and of the form $P_{i_1}^0, P_{i_2}^0, P_{i_3}^0, ..., P_{i_{r-1}}^0, P_{i_r}^0$, where $P_{i_1}^0 \in \Pi_1(k)$, $P_{i_r}^0 \in \Pi_2(k)$ and $P_{i_j}^0 \in \Pi(k)$ for $2 \leq j \leq r-1$ (this is also called a sequence of correctly connected k-cluster configs). D_k^0 has the property that every such walk defines a span $r+k-1$ cluster $G \in C^*$ starting (ending) with the cluster config $P_{i_1}^0$ ($P_{i_r}^0$) and in which cluster config $P_{i_j}^0$ occurs at the jth section, for $j = 2, ..., r-1$. Moreover, any span $r+k-1$ cluster $G \in C^*$ starting with cluster config $P_{i_1}^0$ and ending with cluster config $P_{i_r}^0$ corresponds to a walk of length k of D_k^0 as above.*

Note that it is essential to define the cluster configs so that the above axiom is satisfied. For example, in the case of SAPs in tubes, if no ordering or labelling is assigned to the cluster configs' edges then the above axiom does not hold. Figure 2.6 shows an example where a sequence of three correctly connected SAP 2-patterns does not correspond to a SAP and instead leads to the construction of a pair of polygons. Hence, the Cluster Configuration Axiom does not hold if SAP 2-patterns are used without assigning additional labelling.

Each k-cluster config can be considered as a configuration of $H_0(N,M) \cup S_1(N,M) \cup H_1(N,M) \cup ... \cup H_{k-1}(N,M) \cup S_k(N,M) \cup H_k(N,M)$. Since these cluster configs are contained in a finite subgraph of the lattice and the edges are ordered and labelled in finitely many ways, there is a finite number of such cluster configs. Thus $\Pi(k)$, $\Pi_1(k)$ and $\Pi_2(k)$ are all finite sets. Hence, let

$$\Pi(k) = \{P_1, P_2, ..., P_{|\Pi(k)|}\}, \tag{2.6.3}$$



(a) A sequence of three SAP 2-patterns in $T(0,6)$.

(b) The resulting cluster is a pair of SAPs not a single SAP.

Figure 2.10: An example of a sequence of correctly connected 2-SAP configs which does not correspond to a SAP; instead it results in a pair of polygons.

$$\Pi_1(k) = \{P_1', P_2', ..., P_{|\Pi_1(k)|}'\} \tag{2.6.4}$$

and

$$\Pi_2(k) = \{P_1'', P_2'', ..., P_{|\Pi_2(k)|}''\}. \tag{2.6.5}$$

Given a set of clusters C, let a weight function $wt_s : C^* \to \mathbb{R}[x]$ be given such that

$$wt_s(G) = x^{e(G)}, \tag{2.6.6}$$

where $e(G) = n$ represents the size of the cluster G. Let the functions $h : \Pi(k) \to [0, \infty)$, $h^1 : \Pi_1(k) \to [0, \infty)$ and $h^2 : \Pi_2(k) \to [0, \infty)$ be given such that, for any $G \in C$ with span m, $wt_s(G)$ can be written as

$$wt_s(G) = x^{e(G)} = x^{h^1(P_{l_1}')+h^2(P_{l_{m-k+1}}'')+\sum_{l=2}^{m-k} h(P_{l_i})}, \tag{2.6.7}$$

where $P_{l_i} \in \Pi(k)$ occurs at the ith section of G, for $2 \leq i \leq m-k$ and P_{l_1}' ($P_{l_{m-k+1}}''$) is the start (end) cluster config for G. Furthermore, given a set of clusters C, let a multi-variable weight function $wt : C^* \to \mathbb{R}[y_1, ..., y_t]$ be given such that

$$wt(G) = \prod_{j=1}^{t} y_j^{d(G,j)}, \tag{2.6.8}$$

where $d(G, j)$ is a non-negative integer associated with G, for $1 \leq j \leq t$. Let the multi-variable functions $f : \Pi(k) \to [0, \infty)^t$, $f^1 : \Pi_1(k) \to [0, \infty)^t$ and $f^2 : \Pi_2(k) \to [0, \infty)^t$ be given such that, for any $G \in C$ with span m, $wt(G)$ can be written as

$$wt(G) = \prod_{j=1}^{t} y_j^{d(G,j)} = \prod_{j=1}^{t} y_j^{f_j^1(P'_{l_1})+f_j^2(P''_{l_{m-k+1}})+\sum_{l=2}^{m-k} f_j(P_{l_i})}, \tag{2.6.9}$$

where $P_{l_i} \in \Pi(k)$ occurs at the ith section of G, for $2 \leq i \leq m-k$ and P'_{l_1} ($P''_{l_{m-k+1}}$) is the start (end) cluster config for G. Note that $f = (f_1, ..., f_t)$, $f^1 = (f_1^1, ..., f_t^1)$ and $f^2 = (f_1^2, ..., f_t^2)$.

As an example, take C to be the set of SAPs confined to an (N, M)-tube. Consider $e(G)$ ($h^2(P''_i)$) to be the total number of edges in the SAP G (end cluster config P''_i) and $h(P_i)$ ($h^1(P'_i)$) to be the number of edges in the first hinge and section of P_i (P'_i). Let $t = 2$ and fix the two cluster configs P_1 and P_2 in $\Pi(k)$. Then define $d(G, 1)$ ($d(G, 2)$) to be the number of occurrences of the cluster config P_1 (P_2) in G. Let also $f_j(P_i) = \delta_{j,i}$, $f_j^1(P'_i) = \delta_{j,i}$ and $f_j^2(P''_i) = \delta_{j,i}$ for $j = 1, 2$ and $i = 1, 2$. Then equations 2.6.7 and 2.6.9 are satisfied.

For convenience, for any $1 \leq i \leq |\Pi(k)|$, define $e_i = h(P_i)$ and, for any $1 \leq i \leq |\Pi_1(k)|$ ($1 \leq i \leq |\Pi_2(k)|$), define $e'_i = h^1(P'_i)$ ($e''_i = h^2(P''_i)$). Also, for any $1 \leq i \leq |\Pi(k)|$, define $D_i = f(P_i)^T$ and $d_i(j) = f_j(P_i)$ thus $D_i = [d_i(1), ..., d_i(t)]^T$ is a vector of non-negative integers associated to $P_i \in \Pi(k)$. For any $1 \leq i \leq |\Pi_1(k)|$ ($1 \leq i \leq |\Pi_2(k)|$), define also $D'_i = f^1(P'_i)^T$ ($D''_i = f^2(P''_i)^T$) and $d'_i(j) = f_j^1(P'_i)$ ($d''_i(j) = f_j^2(P''_i)$) thus $D'_i = [d'_i(1), ..., d'_i(t)]^T$ ($D''_i = [d''_i(1), ..., d''_i(t)]^T$) is a vector of non-negative integers associated to $P'_i \in \Pi_1(k)$ ($P''_i \in \Pi_2(k)$).

Let

$$Z_n(N, M; \bar{P}, Y) = \sum_{G \in C_n^*(\bar{P})} \prod_{j=1}^{t} y_j^{d(G,j)}, \tag{2.6.10}$$

where $C_n^*(\bar{P}) \subset C_n^*$ is the set of those clusters in C_n^* which do not contain the cluster config P. Next we prove the following theorem which yields the asymptotic form of $Z_n(N, M; Y)$ and leads to stronger results about the asymptotic properties of clusters confined to an (N, M)-tube. Note that this result is a generalization of Theorem 6.1 in [51].

Theorem 2.6.1. *Given $k \geq 2$, a set of clusters C in $T(N, M)$ and three sets of k-cluster configs $\Pi(k)$, $\Pi_1(k)$ and $\Pi_2(k)$, suppose that the Cluster Configuration Axiom and (CONCAT) hold for C. Let P be any k-cluster config in $\Pi(k)$. Given weight functions $wt_s : C^* \to \mathbb{R}[x]$ and $wt : C^* \to \mathbb{R}[y_1, ..., y_t]$, if there exist functions $h : \Pi(k) \to [0, \infty)$, $h^1 : \Pi_1(k) \to [0, \infty)$, $h^2 : \Pi_2(k) \to [0, \infty)$, $f : \Pi(k) \to [0, \infty)^t$, $f^1 : \Pi_1(k) \to [0, \infty)^t$ and $f^2 : \Pi_2(k) \to [0, \infty)^t$ such that equations 2.6.7 and 2.6.9 are satisfied, then there exist non-negative values $x_0(Y)$ and α_Y such that*

$$Z_n(N, M; Y) = \alpha_Y (x_0(Y))^{-n} + o((x_0(Y))^{-n}) \quad \text{as } n \to \infty. \tag{2.6.11}$$

Moreover, there exist non-negative values $\bar{x}_0(Y) > x_0(Y)$ and $\bar{\alpha}_Y$ such that

$$Z_n(N, M; \bar{P}, Y) = \bar{\alpha}_Y (\bar{x}_0(Y))^{-n} + o((\bar{x}_0(Y))^{-n}) \quad \text{as } n \to \infty. \tag{2.6.12}$$

Proof. Given $k \geq 2$, let $D_k = (V_k, A_k, \phi_k)$ be the sub-digraph of D_k^0 (the digraph for which the Cluster Configuration Axiom holds) induced by the set of vertices $V_k = \Pi(k) = \{P_1, P_2, ..., P_{|\Pi(k)|}\} \subset V_k^0$. In addition, associate the weight $x^{e_i} \prod_{j=1}^{t} y_j^{d_i(j)}$ to each arc $(P_i, P_j) \in A_k$. The transfer matrix $G(x, Y) = (g_{i,j}(x, Y))$ is then defined as follows:

$$g_{i,j}(x, Y) = \begin{cases} x^{e_i} \prod_{j=1}^{t} y_j^{d_i(j)} & \text{if } (P_i, P_j) \in A_k \\ 0 & \text{otherwise.} \end{cases} \tag{2.6.13}$$

Define also the $|\Pi_1(k)| \times |\Pi(k)|$ matrix $A(x, Y)$ as follows: For $1 \leq i \leq |\Pi_1(k)|$ and $1 \leq j \leq |\Pi(k)|$, if the configuration of the start cluster config $P_i' \in \Pi_1(k)$ on $H_1(N, M) \cup S_2(N, M) \cup ... \cup S_k(N, M) \cup H_k(N, M)$ is equivalent to the configuration of the proper cluster config $P_j \in \Pi(k)$ on $H_0(N, M) \cup S_1(N, M) \cup ... \cup S_{k-1}(N, M) \cup H_{k-1}(N, M)$ then $A_{i,j}(x, Y) = x^{e_i'} \prod_{j=1}^{t} y_j^{d_i'(j)}$; otherwise $A_{i,j}(x, Y) = 0$. Similarly, define the $|\Pi(k)| \times |\Pi_2(k)|$ matrix $B(x, Y)$ as follows: For $1 \leq i \leq |\Pi(k)|$ and $1 \leq j \leq |\Pi_2(k)|$, if the configuration of the proper cluster config $P_i \in \Pi(k)$ on $H_1(N, M) \cup S_2(N, M) \cup ... \cup S_k(N, M) \cup H_k(N, M)$ is equivalent to the configuration of the end cluster config $P''_j \in \Pi_2(k)$ on $H_0(N,M) \cup S_1(N,M) \cup ... \cup S_{k-1}(N,M) \cup H_{k-1}(N,M)$ then $B_{i,j}(x,Y) = x^{e''_i} \prod_{j=1}^{t} y_j^{d''_i(j)}$; otherwise $B_{i,j}(x,Y) = 0$.

Given $r \geq 1$, $1 \leq i_1 \leq |\Pi_1(k)|$, $1 \leq i_r \leq |\Pi_2(k)|$ and $1 \leq i_j \leq |\Pi(k)|$ for $2 \leq j \leq r-1$, consider a sequence of r cluster configs of the form $P'_{i_1}, P_{i_2}, P_{i_3}, ..., P_{i_{r-1}}, P''_{i_r}$ such that $A_{i_1,i_2}(x,Y) \neq 0, B_{i_{r-1},i_r}(x,Y) \neq 0$, and $g_{i_j,i_{j+1}}(x,Y) \neq 0$ for $2 \leq j \leq r-2$. Since the Cluster Configuration Axiom holds, this sequence defines a span $r+k-1$ cluster, G, starting (ending) with cluster config P'_{i_1} (P''_{i_r}) and in which cluster config P_{i_j} occurs at the jth section, for $j = 2, ..., r-1$. The weight associated to this cluster in $(A(x,Y)G(x,Y)^{r-1}B(x,Y))_{i_1,i_r}$ is

$$x^{e'_{i_1}+e''_{i_r}\sum_{j=2}^{r-1} e_{i_j}} \prod_{j=1}^{t} y_j^{d'_{i_1}(j)+d''_{i_r}(j)+\sum_{l=2}^{r-1} d_{i_l}(j)} = x^{e(G)} \prod_{j=1}^{t} y_j^{d(G,j)}. \tag{2.6.14}$$

Also by the Cluster Configuration Axiom, any span $r+k-1$ cluster starting with cluster config P'_{i_1} and ending with cluster config P''_{i_r} can be decomposed into a sequence of r cluster configs as above. Thus the generating function $F(x,Y) = \sum_{n\geq 1} Z_n(N,M;Y)x^n$ (summed over all the values of n for which $Z_n(N,M;Y) \neq 0$) satisfies the following:

$$\begin{aligned} F(x,Y) &= \sum_{h=0}^{\infty} \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} (A(x,Y)G(x,Y)^h B(x,Y))_{i,j} \\ &= \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} (A(x,Y)(I-G(x,Y))^{-1}B(x,Y))_{i,j} \\ &= \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} \Big[A_{i,l}(x,Y)((I-G(x,Y))^{-1})_{l,o}B_{o,j}(x,Y)\Big]. \end{aligned} \tag{2.6.15}$$

Note also that by Theorem A.0.2,

$$((I-G(x,Y))^{-1})_{l,o} = \frac{\det((I-G(x,Y))^{-1};o,l)}{\det(I-G(x,Y))}, \tag{2.6.16}$$

where $(A;o,l)$ represents the matrix obtained by removing the lth row and oth column of a given matrix A. Thus

$$\begin{aligned} F(x,Y) &= \frac{1}{\det(I-G(x,Y))} \\ &\times \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} \Big[A_{i,l}(x,Y) \det((I-G(x,Y))^{-1}; o,l) B_{o,j}(x,Y) \Big]. \end{aligned} \tag{2.6.17}$$

Given any $x > 0$, it can be shown that for any pair of proper cluster configs $P_i, P_j \in \Pi(k)$ there exists an integer m such that $(G(x,Y)^m)_{i,j} > 0$ (to see this start with a cluster in C in which proper cluster config P_i occurs and use (CONCAT) to concatenate it to a cluster in C in which proper cluster config P_j occurs. This yields a cluster in C in which P_i occurs at some section and cluster config P_j occurs at a later section. From this obtain a sequence of m correctly connected cluster configs starting with cluster config P_i and ending in cluster config P_j). Thus, by Theorem A.0.4, $G(x,Y)$ is an irreducible and aperiodic matrix and Frobenius theory (Theorem A.0.5) implies that: the spectral radius, $\rho(x,Y)$, of $G(x,Y)$ is a simple root of $\det(\lambda I - G(x,Y))$; $G(x,Y)$ has a strictly positive eigenvector associated with $\rho(x,Y)$; and $\rho(x,Y)$ is the only eigenvalue of modulus $\rho(x,Y)$. Since $\rho(0,Y) = 0$ and, by Theorem A.0.7, $\rho(x,Y)$ is an unbounded, increasing, continuous function on $[0,\infty)$, hence there exists a unique $x_0(Y) > 0$ such that $\rho(x_0(Y),Y) = 1$. From equation 2.6.17, $F(x,Y)$ has poles only when $\frac{1}{\det(I-G(x,Y))}$ has poles, that is, when 1 is an eigenvalue of $G(x,Y)$. Thus based on the results and arguments of [1, Lemma 9 and Theorem 3] $F(x,Y)$ is analytic for $|x| < x_0(Y)$ and has one simple pole when $|x| = x_0(Y)$, namely, $x = x_0(Y)$. In particular, Theorem A.0.8 implies that as

$x \to x_0(Y)$

$$F(x,Y) = \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} \Big[A_{i,l}(x,Y)((I-G(x,Y))^{-1})_{l,o} B_{o,j}(x,Y)\Big] \to$$

$$\sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} \Big[A_{i,l}(x_0(Y),Y)(x_0(Y)-x)^{-1} x_0(Y) \beta_Y^{-1} (\eta_Y)_l (\varsigma_Y)_o^T B_{o,j}(x_0(Y),Y)\Big], \tag{2.6.18}$$

where $\beta_Y = x_0(Y) \varsigma_Y^T G'(x_0(Y),Y)\eta_Y$, and η_Y and ς_Y^T are, respectively, strictly positive left and right eigenvectors of $G(x_0(Y),Y)$ associated with $\rho(x_0(Y),Y) = 1$ and normalized so that $\varsigma_Y^T \eta_Y = 1$ (note that $G'(x,Y)$ denotes the derivative of $G(x,Y)$ with respect to x). Thus, differentiating both sides of the above equation n times with respect to x, dividing by $n!$ and setting $x = 0$ implies that

$$Z_n(N,M;Y) = \alpha_Y (x_0(Y))^{-n} + o((x_0(Y))^{-n}) \quad \text{as } n \to \infty, \tag{2.6.19}$$

where

$$\alpha_Y = \beta_Y^{-1} \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} \Big[A_{i,l}(x_0(Y),Y)(\eta_Y)_l (\varsigma_Y)_o^T B_{o,j}(x_0(Y),Y)\Big] > 0. \tag{2.6.20}$$

Let $P \in \Pi(k)$ be a specific proper cluster config. Let also P represent the label of cluster config P in $\Pi(k)$. Consider the generating function $\bar{F}(x,Y) = \sum_{n\geq 1} \hat{Z}_n(N,M;\bar{P},Y)x^n$. Then

$$\begin{aligned} \bar{F}(x,Y) &= \sum_{h=0}^{\infty} \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} (\bar{A}(x,Y)\bar{G}(x,Y)^h \bar{B}(x,Y))_{i,j} \\ &= \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} (\bar{A}(x,Y)(I-\bar{G}(x,Y))^{-1}\bar{B}(x,Y))_{i,j} \end{aligned} \tag{2.6.21}$$

where $\bar{G}(x,Y)$ is obtained from $G(x,Y)$ by deleting its Pth row and column, and $\bar{A}(x,Y)$ and $\bar{B}(x,Y)$ are defined as follows. If $P \notin \Pi_1(k)$ ($P \notin \Pi_2(k)$), $\bar{A}(x,Y)$ ($\bar{B}(x,Y)$) is defined to be the matrix obtained from $A(x,Y)$ ($B(x,Y)$) by deleting its Pth column (row). Otherwise, $\bar{A}(x,Y)$ ($\bar{B}(x,Y)$) is defined to be the matrix obtained from $A(x,Y)$ $(B(x,Y))$ by deleting its Pth column (row) and lth row (column) where l satisfies $P'_l = P$ $(P''_l = P)$. The argument used after equation 2.6.18 applies again so that there exist $\bar{x}_0(Y) > 0$ and $\bar{\alpha}_Y > 0$ such that

$$Z_n(N,M;\bar{P},Y) = \bar{\alpha}_Y(\bar{x}_0(Y))^{-n} + o((\bar{x}_0(Y))^{-n}) \quad \text{as } n \to \infty \tag{2.6.22}$$

and the spectral radius, $\bar{\rho}(\bar{x}_0(Y),Y)$, of $\bar{G}(\bar{x}_0(Y),Y)$ equals 1. Now consider the matrix $G_P(x,Y)$ obtained from $G(x,Y)$ by replacing the Pth row and column by a row and column of zeros. Then the spectral radius, $\rho_P(x,Y)$, of $G_P(x,Y)$ equals $\bar{\rho}(x,Y)$. Furthermore, $G_P(x,Y) \le G(x,Y)$ and at least one element of $G_P(x,Y)$ is strictly less than the corresponding element of $G(x,Y)$. Theorem A.0.7 then implies that $\rho_P(x,Y) < \rho(x,Y)$ and hence for $x = x_0(Y)$, $\rho_P(x_0(Y),Y) < \rho(x_0(Y),Y) = 1$ and therefore $\bar{x}_0(Y) > x_0(Y)$. $\square$

A function $\psi : C^* \to \mathbb{N} \cup \{0\}$ is called an *additive functional* if, for any $G \in C^*$ with span m, $\psi(G)$ can be written as

$$\psi(G) = d'(P'_{h_1}) + d''(P''_{f_1}) \sum_{i=2}^{m-k} \psi(P_{l_i}), \tag{2.6.23}$$

where $\psi(P_{l_i})$ is a non-negative integer associated with the proper cluster config $P_{l_i} \in \Pi(k)$ occurring at the ith section of G, for $2 \le i \le m-k$, and $d'(P'_{h_1})$ $(d''(P''_{f_1}))$ is also a non-negative integer associated with the start (end) cluster config P'_{h_1} (P''_{f_1}) for G [1]. Note that $\psi(P_{l_i})$ depends only on P_{l_i} not on G. Let $\Lambda_\psi(x,Y)$ be a $\Pi(k) \times \Pi(k)$ matrix with (i,j)th elements

$$\Lambda_\psi(x,Y)(i,j) = \begin{cases} \psi(P_i)g_{i,j}(x,Y) & \text{if } \psi(P_i) \ne 0 \\ 0 & \text{otherwise} \end{cases}. \tag{2.6.24}$$

Define $\Lambda_A(x,Y)$ to be a $\Pi_1(k) \times \Pi(k)$ matrix with (i,j)th elements

$$\Lambda_A(x,Y)(i,j) = \begin{cases} d'(P'_i)A_{i,j}(x,Y) & \text{if } d'(P'_i) \ne 0 \\ 0 & \text{otherwise} \end{cases}, \tag{2.6.25}$$

and $\Lambda_B(x,Y)$ to be a $\Pi(k) \times \Pi_2(k)$ matrix with (i,j)th elements

$$\Lambda_B(x,Y)(i,j) = \begin{cases} d''(P''_i)B_{i,j}(x,Y) & \text{if } d''(P''_i) \ne 0 \\ 0 & \text{otherwise} \end{cases}. \tag{2.6.26}$$

Let X_n denote a random variable which takes clusters from C_n^* at random with probabilities determined by the partition function $Z_n(N, M; Y)$, i.e.

$$\mathbb{P}(X_n = G) = \frac{\prod_{j=1}^{t} y_j^{d(G,j)}}{Z_n(N, M; Y)}. \tag{2.6.27}$$

Next, we investigate the asymptotic behaviour of the expected value of $\psi(X_n)$, $E_Y(\psi(X_n))$. Note that this result is a generalization of Theorem 7 in [1].

Theorem 2.6.2. *Let ψ be an additive functional. Let $Y = [y_1, ..., y_t]^T$, $x_0(Y)$, ς_Y, η_Y, $G(x(Y), Y)$ and β_Y be given as introduced in the proof of Theorem 2.6.1. Then there exists a $\gamma_Y > 0$ such that as $n \to \infty$*

$$E_Y(\psi(X_n)) = \gamma_Y n + O(1). \tag{2.6.28}$$

Proof. For any $1 \leq i \leq \Pi(k)$, $1 \leq j \leq \Pi_1(k)$ and $1 \leq l \leq \Pi_2(k)$, let $d_i(t+1) = \psi(P_i)$, $d'_j(t+1) = d'(P'_j)$ and $d''_l(t+1) = d''(P''_l)$. For any cluster G, let $d(G, t+1) = \psi(G)$. Define $D_i^1 = [d_i(1), ..., d_i(t), d_i(t+1)]^T$, $D_i'^1 = [d'_i(1), ..., d'_i(t), d'_i(t+1)]^T$ and $D_i''^1 = [d''_i(1), ..., d''_i(t), d''_i(t+1)]^T$ where $D_i = [d_i(1), ..., d_i(t)]$, $D'_i = [d'_i(1), ..., d'_i(t)]$ and $D''_i = [d''_i(1), ..., d''_i(t)]$. Set $Y^1 = [Y, e^s]^T$, where $y_{t+1} = e^s$, i.e. $s = \ln y_{t+1}$. Define $G(x, Y; s) = G(x, Y^1)$, $A(x, Y; s) = A(x, Y^1)$ and $B(x, Y; s) = B(x, Y^1)$. Thus

$$G(x, Y; 0) = G(x, Y), \quad A(x, Y; 0) = A(x, Y) \text{ and } B(x, Y; 0) = B(x, Y). \tag{2.6.29}$$

Note that

$$\frac{\partial}{\partial s} G(x(Y), Y; s)|_{s=0} = \Lambda_\psi(x(Y), Y), \tag{2.6.30}$$

$$\frac{\partial}{\partial s} A(x(Y), Y; s)|_{s=0} = \Lambda_A(x(Y), Y) \tag{2.6.31}$$

and

$$\frac{\partial}{\partial s} B(x(Y), Y; s)|_{s=0} = \Lambda_B(x(Y), Y). \tag{2.6.32}$$

Let also $\Lambda_\psi = \Lambda_\psi(x_0(Y), Y)$, $\Lambda_1 = \Lambda_A(x_0(Y), Y)$ and $\Lambda_2 = \Lambda_B(x_0(Y), Y)$.

We have

$$E_Y(e^{s\psi(X_n)}) = \frac{1}{Z_n(N,M;Y)} \sum_{G\in C_n^*} \big(e^{sd(G,t+1)} \prod_{j=1}^{t} y_j^{d(G,j)}\big). \tag{2.6.33}$$

Hence, we obtain

$$\begin{aligned}
&\sum_{n\geq 1} E_Y(e^{s\psi(X_n)}) Z_n(N,M;Y)x^n = \\
&\quad \sum_{n\geq 1}\sum_{G\in C_n^*} e^{sd(G,t+1)} x^n \prod_{j=1}^{t} y_j^{d(G,j)} = \\
&\quad \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} (A(x,Y;s)(I-G(x,Y;s))^{-1}B(x,Y;s))_{i,j} = \\
&\quad \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} \Big[A_{i,l}(x,Y;s)((I-G(x,Y;s))^{-1})_{l,o} B_{o,j}(x,Y;s)\Big].
\end{aligned} \tag{2.6.34}$$

Differentiating with respect to s and setting $s=0$ gives

$$\begin{aligned}
&\sum_{n\geq 1} E_Y(\psi(X_n)) Z_n(N,M;Y)x^n = \\
&\quad \frac{\partial}{\partial s}\Big[\sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} A_{i,l}(x,Y;s)((I-G(x,Y;s))^{-1})_{l,o} B_{o,j}(x,Y;s)\Big]_{s=0} = \\
&\quad \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} \frac{\partial}{\partial s}\Big[A_{i,l}(x,Y;s)((I-G(x,Y;s))^{-1})_{l,o} B_{o,j}(x,Y;s)\Big]_{s=0} = \\
&\quad \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} \Big[\big[\frac{\partial}{\partial s} A_{i,l}(x,Y;s)\big] B_{o,j}(x,Y;s)((I-G(x,Y;s))^{-1})_{l,o} + \\
&\qquad\qquad A_{i,l}(x,Y;s)\big[\frac{\partial}{\partial s} B_{o,j}(x,Y;s)\big]((I-G(x,Y;s))^{-1})_{l,o} + \\
&\qquad\qquad A_{i,l}(x,Y;s)B_{o,j}(x,Y;s)\big[\frac{\partial}{\partial s}((I-G(x,Y;s))^{-1})_{l,o}\big]\Big]_{s=0} = \\
&\quad \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} \Big[[\Lambda_A(x,Y)_{i,l} B_{o,j}(x,Y) + A_{i,l}(x,Y)\Lambda_B(x,Y)(o,j)] \times \\
&\qquad\qquad ((I-G(x,Y))^{-1})_{l,o} + A_{i,l}(x,Y)B_{o,j}(x,Y)\big((I-G(x,Y))^{-1} \times \\
&\qquad\qquad \Lambda_\psi(x,Y)(I-G(x,Y))^{-1}\big)_{l,o}\Big],
\end{aligned} \tag{2.6.35}$$

where the last line of the equation is obtained by equations 2.6.29 to 2.6.32. The right side of the above equation is analytic for $| x |< x_0(Y)$ and Theorem A.0.8 yields

$$\sum_{n\geq 1} E(\psi(X_n))Z_n(N,M;Y)x^n =$$
$$\sum_{i=1}^{\Pi_1(k)}\sum_{j=1}^{\Pi_2(k)}\sum_{o=1}^{\Pi(k)}\sum_{l=1}^{\Pi(k)}\Big[[\Lambda_1(i,l)B_{o,j}(x_0(Y),Y)+A_{i,l}(x_0(Y),Y)\Lambda_1(o,j)]\times$$
$$(\eta_Y)_l(\varsigma_Y^T)_o(\varsigma_Y^T G'(x_0(Y),Y)\eta_Y)^{-1}(x_0(Y)-x)^{-1}+$$
$$A_{i,l}(x_0(Y),Y)B_{o,j}(x_0(Y),Y)(\eta_Y\varsigma_Y^T\Lambda_\psi\eta_Y\varsigma_Y^T)_{l,o}\times$$
$$(\varsigma_Y^T G'(x_0(Y),Y)\eta_Y)^{-2}(x_0(Y)-x)^{-2}\Big]+O((x_0(Y)-x)^{-1}) \tag{2.6.36}$$

as $x \rightarrow x_0(Y)$.

Therefore, differentiating both sides of the above equation n times with respect to x, dividing by $n!$ and setting $x = 0$ gives

$$E_Y(\psi(X_n))Z_n(N,M;Y) =$$
$$\sum_{i=1}^{\Pi_1(k)}\sum_{j=1}^{\Pi_2(k)}\sum_{o=1}^{\Pi(k)}\sum_{l=1}^{\Pi(k)}\Big[[\Lambda_1(i,l)B_{o,j}(x_0(Y),Y)+A_{i,l}(x_0(Y),Y)\Lambda_2(o,j)]\times$$
$$(\eta_Y)_l(\varsigma_Y^T)_o\beta_Y^{-1}(x_0(Y))^{-n}+A_{i,l}(x_0(Y),Y)B_{o,j}(x_0(Y),Y)\times$$
$$(\eta_Y\varsigma_Y^T\Lambda_\psi\eta_Y\varsigma_Y^T)_{l,o}\beta_Y^{-2}n(x_0(Y))^{-n}\Big]+O((x_0(Y))^{-n})$$
$$=\gamma'_Y n(x_0(Y))^{-n}+O((x_0(Y))^{-n}), \tag{2.6.37}$$

where

$$\gamma'_Y = \beta_Y^{-2}\sum_{i=1}^{\Pi_1(k)}\sum_{j=1}^{\Pi_2(k)}\sum_{o=1}^{\Pi(k)}\sum_{l=1}^{\Pi(k)}\Big[A_{i,l}(x_0(Y),Y)B_{o,j}(x_0(Y),Y)(\eta_Y\varsigma_Y^T\Lambda_\psi\eta_Y\varsigma_Y^T)_{l,o}\Big]. \tag{2.6.38}$$

Similarly, by the proof of Theorem 2.6.1, for any $\epsilon > 0$ such that $0 < x_0(Y)^{-1}-\epsilon <$ 1

$$Z_n(N,M;Y) = \alpha_Y(x_0(Y))^{-n}+O((x_0(Y)^{-1}-\epsilon)^n), \tag{2.6.39}$$

as $n \to \infty$.

Therefore, equations 2.6.37 and 2.6.39 imply that as $n \to \infty$

$$E_Y(\psi(X_n)) = \gamma_Y n + O(1), \tag{2.6.40}$$

where

$$\gamma_Y = \frac{\gamma'_Y}{\alpha_Y} = \frac{\sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} \left[A_{i,l}(x_0(Y), Y) B_{o,j}(x_0(Y), Y) (\eta_Y \varsigma_Y^T \Lambda_\psi \eta_Y \varsigma_Y^T)_{l,o}\right]}{\beta_Y \sum_{i=1}^{\Pi_1(k)} \sum_{j=1}^{\Pi_2(k)} \sum_{o=1}^{\Pi(k)} \sum_{l=1}^{\Pi(k)} \left[A_{i,l}(x_0(Y), Y)(\eta_Y)_l (\varsigma_Y)_o^T B_{o,j}(x_0(Y), Y)\right]}. \tag{2.6.41}$$

□

Corollary 2.6.3.

$$\frac{1}{n} E_Y(\psi(X_n)) \to \gamma_Y \tag{2.6.42}$$

as $n \to \infty$.

Chapter 3

Polygonal Knots and Links in $\mathbb{Z}^3$

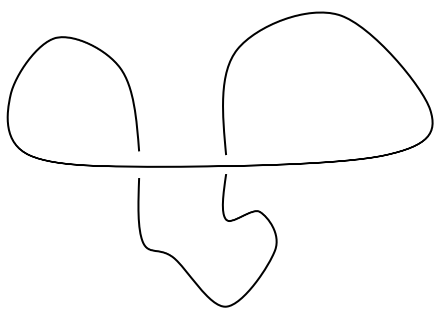
The entanglement complexity of polymers has been explored using a variety of rigorous [43, 55] and numerical methods [23, 29, 35, 59]. In particular, knotting and linking have been frequently used to measure the topological entanglement of ring polymers, modeled by self-avoiding polygons [25, 42]. The main advantage of this approach is that knots and links are mathematically well defined and they are of special interest to chemists and molecular biologists [56, 60, 61].

In order to discuss polygonal topological entanglement via knotting and linking we need to introduce some definitions and results from topology; this background information is given in Section 3.1. Then some theorems and arguments concerned with the knotting and linking probability of lattice polygons in $\mathbb{Z}^3$ are reviewed in Section 3.2. In order to investigate the influence of the tube constraint on a polygon's topological entanglement complexity, knotting and linking probabilities for polygons in tubes are also considered in Section 3.3.

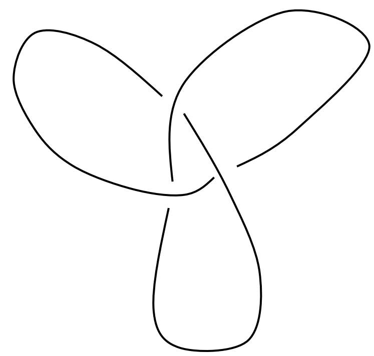
3.1 Knots, Links and 2-String Tangles

A quick review of the main definitions and results related to knots, links and 2-string tangles is given in this section. We will be using these results in the next sections of this chapter as well as the next chapters to discuss the knotting and linking probability of lattice polygons. Unless stated otherwise, the discussion in this section is based on the presentation in [7]. Note that some of the definitions and theorems of this section can be stated for arbitrary topological spaces X and Y. However, for our purpose, here we will take X and Y to be subsets of the three-dimensional Euclidean space $\mathbb{R}^3$ with the usual topology.

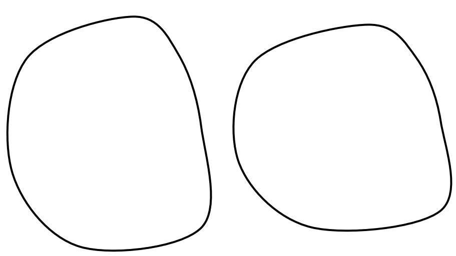
The two topological spaces $X \subseteq \mathbb{R}^3$ and $Y \subseteq \mathbb{R}^3$ are said to be *homeomorphic* if there exists a bijection $f : X \to Y$ such that both f and its inverse, f^{-1}, are continuous functions. f then will be called a *homeomorphism*. Any subset K of $\mathbb{R}^3$ which is homeomorphic to the unit circle $\mathbb{S}^1 = \{(x_1, x_2) \in \mathbb{R}^2 \mid \sqrt{x_1^2 + x_2^2} = 1\} \subset \mathbb{R}^2$ is said to be a *knot* (e.g. see Figure 3.1 (a) and (b)). A *polygonal knot* in $\mathbb{R}^3$ is a knot $K \subset \mathbb{R}^3$ that is the union of a finite number of closed straight-line segments, called edges, in $\mathbb{R}^3$ (e.g. see Figure 3.2 (a)); the point of intersection of any two edges of a polygonal knot is called a *vertex*. So a self-avoiding polygon in $\mathbb{Z}^3$ can be considered as a polygonal knot.



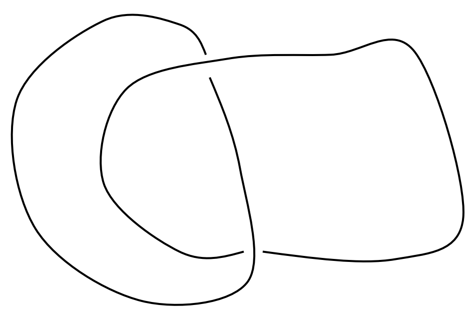
(a) Unknot



(b) Trefoil knot



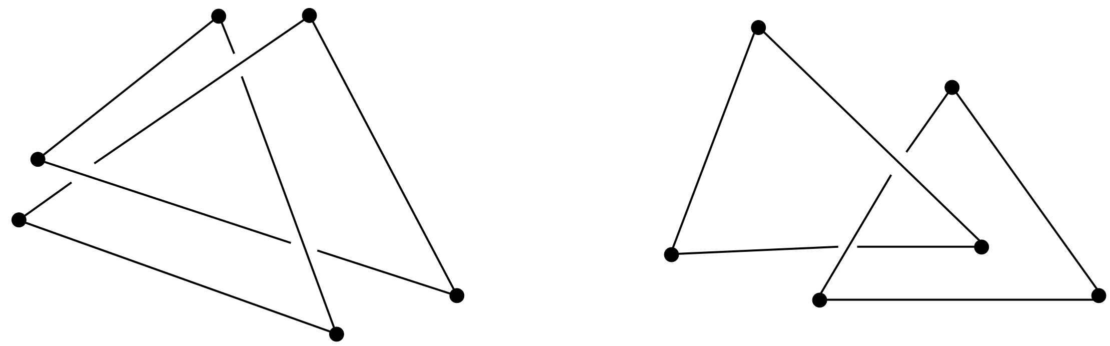
(c) Unlink



(d) A linked link

Figure 3.1: Examples of knots and links.

A *homotopy* of a topological space $X \subseteq \mathbb{R}^3$ is a continuous map $h : X \times [0, 1] \to \mathbb{R}^3$ such that $h(x, 0) = x$ for any $x \in X$. The restriction of h to level $t \in [0, 1]$ is defined



(a) An example of a polygonal knot (b) An example of a polygonal link

Figure 3.2: Polygonal knots and links.

to be $h_t : X \to \mathbb{R}^3$ given by $h_t(x) = h(x,t)$. Note that here t is chosen to indicate time, and the images $h_t(X)$ for increasing values of t show the evolution of X in $\mathbb{R}^3$. An *isotopy* of a topological space $X \subseteq \mathbb{R}^3$ is a homotopy of X, $h : X \times [0,1] \to \mathbb{R}^3$, such that $h_t : X \to h_t(X)$ is a homeomorphism for any $t \in [0,1]$. Two topological spaces $X \subseteq \mathbb{R}^3$ and $Y \subseteq \mathbb{R}^3$ are said to be *ambient isotopic* if there is an isotopy of $\mathbb{R}^3$, $h : \mathbb{R}^3 \times [0,1] \to \mathbb{R}^3$, such that it carries X to Y, i.e. $h(X,1) = h_1(X) = Y$.

Two ambient isotopic knots are said to have the same *knot type*. Any knot ambient isotopic to the unit circle $\mathbb{S}^1$ is called an *unknot* (see Figure 3.1 (a)); otherwise it is called *knotted* (e.g. see Figure 3.1 (b)).

A *link* L is defined to be a finite disjoint union of knots, i.e. $L = K_1 \cup ... \cup K_n$ for some $n \in \mathbb{N}$. Each knot in the union is called a *component* of the link and $L = K_1 \cup ... \cup K_n$ is called an n-component link (e.g. see Figure 3.1 (c) and (d)). Note that a knot can also be considered as a one-component link. A *polygonal link* is a finite disjoint union of polygonal knots (e.g. see Figure 3.2 (b)). Here we focus on two-component links (a disjoint union of two knots). Two links L_1 and L_2 are said to have the same *link type* if they are ambient isotopic, i.e. there is an isotopy $h : \mathbb{R}^3 \times [0,1] \to \mathbb{R}^3$ such that $h(L_1,0) = h_0(L_1) = L_1$ and $h(L_1,1) = h_1(L_1) = L_2$. Hereafter we restrict our discussion to links (knots) which are ambient isotopic to a polygonal link (knot). Hence, in the remainder of thesis the term "link" ("knot") refers only to links (knots) which are ambient isotopic to a polygonal link (knot).

The following definitions and results are given for both links and knots, viewed as one-component links. Let $L \subset \mathbb{R}^3$ be a link (knot) and let $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ be a projection map. A point $x \in \pi(L)$ is *regular* if $\pi^{-1}(x)$ is a single point, and is *singular* otherwise. If $| \pi^{-1}(x) |= 2$ then x is called a *double point*. If $\pi(L)$ has a finite number of singular points and they are all transverse double points, the projection is said to be *regular*.

Theorem 3.1.1 (Theorem 3.2.1 [7])**.** *Every polygonal link (polygonal knot) L has a regular projection.*

A link (knot) *diagram D* is a regular projection of a link that has relative height information added to it at each of the double points. The convention is to make breaks in the line corresponding to the strand that passes underneath (e.g. see Figure 3.1). The double points in the projection are called *crossings* in the diagram.

Theorem 3.1.2 (Theorem 3.3.2 [7])**.** *Every polygonal link (polygonal knot) L has a regular diagram.*

Informally, the following theorem from [31, Theorem 1, Chapter 2] says that, for any given projection plane P and any given polygonal knot K, we can always find another polygonal knot in the neighborhood of K which has a regular projection into P and has the same knot type as K.

Theorem 3.1.3 (Theorem 1, Chapter 2 [31])**.** *Given any projection plane P, let K be a polygonal knot determined by the ordered set of vertices $(v_1, ..., v_n)$. For every number $\delta > 0$ there is a polygonal knot K' determined by an ordered set of vertices $(v'_1, ..., v'_n)$ such that the distance from v_i to v'_i is less than δ for all $1 \leq i \leq n$, K' and K have the same knot type, and the projection of K' in P is regular.*

This theorem leads to the following corollary which will be particularly important later, in sections 4.8 and 6.8, when we investigate the topological entanglement of polygons by looking at their regular diagrams.

Corollary 3.1.4. *Given any plane P, any positive real number δ, positive integers N, M, and any vertex-disjoint finite union G of SAPs in $T(N,M)$, let $v_1, ..., v_n$ be the vertices of G. Then, there exists a polygonal link L in $\mathbb{R}^3$ with vertices $l_1, ..., l_n$ that has the same link type as G, has $|l_i - v_i| \leq \delta$ for each $i = 1, ..., n$, and has a regular projection into the plane P. In particular we can choose P as the (x,y)- or (x,z)-plane to generate regular projections D_G^z and D_G^y respectively. Moreover, let $\{v_i, v_j\}$ be any edge in G where $v_i = (x_i, y_i, z_i)$ and $v_j = (x_j, y_j, z_j)$. Then there is a corresponding edge $\{w_i, w_j\}$ in D_G^y with $w_i = (\hat{x}_i, 0, \hat{z}_i)$ and $w_j = (\hat{x}_j, 0, \hat{z}_j)$ and with $\hat{x}_i - \delta \leq x_i \leq \hat{x}_i + \delta$ and $\hat{z}_i - \delta \leq z_i \leq \hat{z}_i + \delta$. Similarly there is a corresponding edge $\{w_i', w_j'\}$ in D_G^z with $w_i' = (x_i', y_i', 0)$ and $w_j = (x_j', y_j', 0)$ and with $x_i' - \delta \leq x_i \leq x_i' + \delta$ and $y_i' - \delta \leq y_i \leq y_i' + \delta$.*

Hereafter we take $\delta = 1/6$. Given any vertex-disjoint union G of SAPs in $T(N,M)$, we fix a polygonal link $L_G = L$ (L') as prescribed by the lemma and define the regular diagram D_G^y (D_G^z) to be L_G's regular projection in the (x,z)-plane ((x,y)-plane).

Next we would like to discuss a significant invariant of links, called the linking number. Before that we need to give a brief introduction to orientable and compact surfaces.

The upper half-space in $\mathbb{R}^2$ is the set $\mathbb{R}^2_+ = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \geq 0\}$. Its boundary is $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = 0\}$ which is homeomorphic to $\mathbb{R}^1$. A *neighborhood* of $x \in X$ is a set $U \subseteq X$, which contains an open set V containing x. Any $X \subseteq \mathbb{R}^3$ is called a *surface* if each point $x \in X$ has a neighborhood $U \ni x$ which is homeomorphic to $\mathbb{R}^2$ or to $\mathbb{R}^2_+$. If $x \in U$ maps to a point in the boundary of $\mathbb{R}^2_+$ then x is a boundary point of X. The set of all boundary points, denoted by ∂X, is called the *boundary* of X.

A *cover* of a set X is a collection of sets in X, $C = \{U_\alpha\}_{\alpha \in A}$ where $U_\alpha \subseteq X$, such that X is a subset of the union of the sets in C, i.e. $X \supseteq \bigcup_{\alpha \in A} U_\alpha$. Any subset of C which is a cover of X is called a *subcover* of C. A surface $X \subseteq \mathbb{R}^3$ is said to be *compact* if for any given covering of X by open sets, we can find a finite subcover.

Here since $X \subset \mathbb{R}^3$ inherits the usual topology, X is compact if and only if it is closed and bounded.

Informally, a surface $X \subset \mathbb{R}^3$ is orientable if a two-dimensional figure (for example a disc with two distinguishable sides) cannot be moved around the surface and back to where it started so that it looks like its own mirror image. Otherwise the surface is non-orientable. More precisely, a surface X is orientable if there is no continuous map $f : B^2 \times [0, 1] \to X$, where $B^2 = \{(x_1, x_2) \in \mathbb{R}^2 \mid \sqrt{x_1^2 + x_2^2} \leq 1\} \subset \mathbb{R}^2$, such that $f(b, t) = f(c, t)$ only if $b = c$ for any t in $[0, 1]$, and $f(b, 0) = f(r(b), 1)$ for every b in B^2, where $r : B^2 \to B^2$ is a reflection map. i.e. $r(x_1, x_2) = (-x_1, -x_2)$ for any $(x_1, x_2) \in B^2$.

There are three types of linking defined in topology: topological linking, homotopic linking and homological linking. Here we focus only on topological and homological linking. We say two disjoint knots K_1 and K_2 are *topologically unlinked* (splittable) if there is a homeomorphism, f, of $\mathbb{R}^3$ onto itself such that the images $f(K_1)$ and $f(K_2)$ are separated by a two-dimensional plane; any link ambient isotopic to a pair of splittable unknots is called an *unlink*. On the other hand, K_1 and K_2 are said to be *homologically unlinked* if K_1 bounds an orientable surface which is disjoint from K_2. K_1 is said to be *homotopically unlinked* from K_2 if there is a homotopy h, from K_1 to the constant map (i.e. $h_0(K_1) = K_1$ and $h_1(K_1)$ is a point) such that $h_t(K_1)$ is disjoint from K_2 for any $t \in [0, 1]$. Note that homotopic linking is not a symmetric relation; i.e. it is possible for K_1 to be homotopically unlinked from K_2 but for K_2 to be homotopically linked to K_1. The following theorem shows the relation between the different types of linking:

Theorem 3.1.5 ([44]). *Homological linking implies homotopic linking and homotopic linking implies topological linking.*

One way to determine whether a two-component link L is homologically linked is by calculating the *linking number* as follows. In a regular diagram D of L with n crossings $c_1, c_2, \ldots, c_n$, let $c_1, c_2, \ldots, c_l$, $l \leq n$, be the crossings that involve both components of L. We choose an orientation of each of the two components of L

and assign a *sign*, $\sigma_i \in \{-1, 1\}$, $1 \leq i \leq l$ to each of the l crossings by applying a *right-hand-rule* (see Figure 3.3 (a)). The linking number of D is then computed by the following formula (e.g. see Figure 3.3):

$$Lk(K_1, K_2) = \frac{1}{2} \sum_{i=1}^{l} \sigma_i. \tag{3.1.1}$$

Note that the absolute value of this linking number is independent of the choice of the diagram D and is also independent of the orientation given to the two components. The following theorem relates homological linking to the linking number.

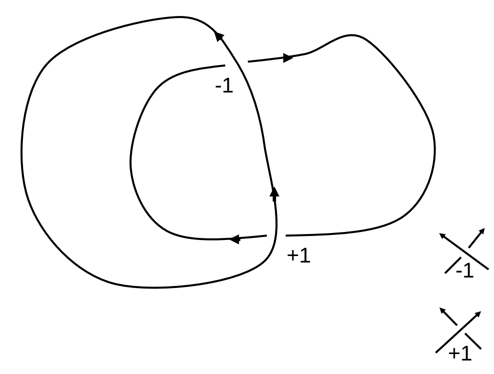
Theorem 3.1.6 (Corollary 5.7.4 [7])**.** *Two knots K_1 and K_2 are homologically linked if and only if*

$$Lk(K_1, K_2) \neq 0. \tag{3.1.2}$$

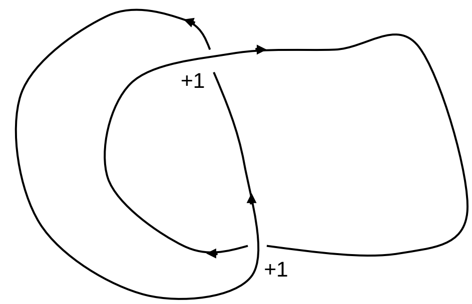
Note that, by Theorems 3.1.5 and 3.1.6, any 2-component link with non-zero linking number is non-splittable; however there are non-splittable links with linking number zero (for example the Whitehead link [44] shown in Figure 3.3 (c)).

There is an operation, called the connected sum, defined on knots which allows one to connect two knots in order to create a new knot. This operation can be described as follows: For any pair of knots K_1 and K_2, consider a regular diagram of these knots and suppose these diagrams are disjoint. Find a pair of distinct points on K_1, (a, b), and a pair of distinct points on K_2, (c, d), such that: 1) The arc in K_1 (K_2) joining a (c) to b (d), $\widehat{ab}$ ($\widehat{cd}$), does not contain any double point of the diagram. 2) There exist two disjoint arcs $\widehat{ac}$ and $\widehat{bd}$ in $\mathbb{R}^2$ connecting a to c and b to d in the diagram, respectively, so that $\widehat{ac}$ ($\widehat{bd}$) intersects K_1 and K_2 only at a (b) and c (d) (e.g. see Figure 3.4 (c)). Now join the two knots together by deleting the two arcs $\widehat{ab}$ and $\widehat{cd}$ from the knots and adding the arcs $\widehat{ac}$ and $\widehat{bd}$ (e.g. see Figure 3.4 (d)). This procedure results in a diagram of a new knot which is called the *connected sum* of the original knots K_1 and K_2 and is denoted by $K_1 \# K_2$ (e.g. see Figure 3.4).

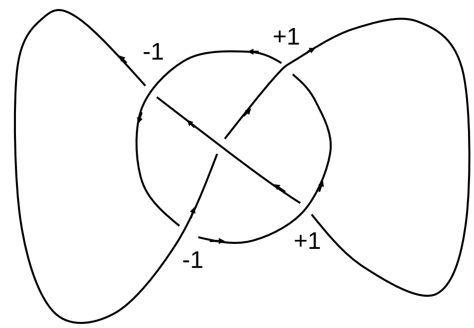
Theorem 3.1.7 ([44])**.** *For any pair of knots K_1 and K_2, the connected sum $K_1 \# K_2$*



(a) An example of a splittable link with linking number zero



(b) An example of a link with linking number one



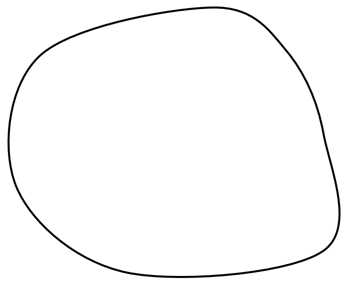
(c) Whitehead Link; an example of a non-splittable link with linking number zero

Figure 3.3: Examples of links with different linking numbers.

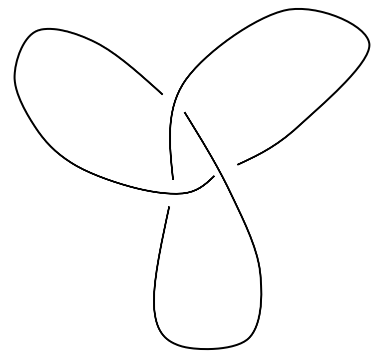
is well-defined, i.e. it is independent of the choice of diagrams for K_1 *and* K_2 *and also the choice of arcs on* K_1 *and* K_2.

Given a link τ, we can also consider the *augmented link* $\tau^\#$, defined as the union of all links which can be obtained from τ by taking the connected sum of each of its components with any knot (including the unknot). Any link in this union is said to have *augmented link type* $\tau^\#$ (e.g see Figure 3.5). Note that regarding the construction of $\tau^\#$, roughly speaking, the knot components K_1 and K_2 of any link $L = (K_1, K_2)$ in $\tau^\#$ are linked in the same way as the components of the other links in $\tau^\#$; however, the links in $\tau^\#$ do not necessarily have the same link type.

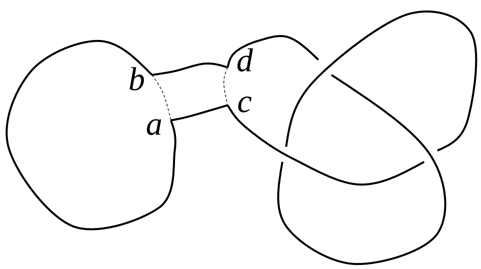
Next, we give a brief introduction to 2-string tangles. Then we state a lemma,



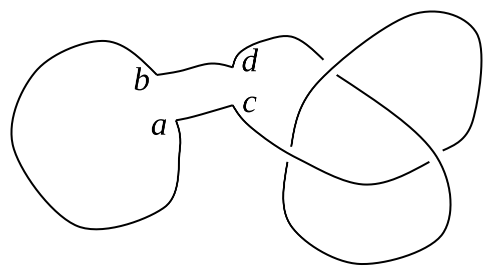
(a) Unknot



(b) Trefoil knot



(c)

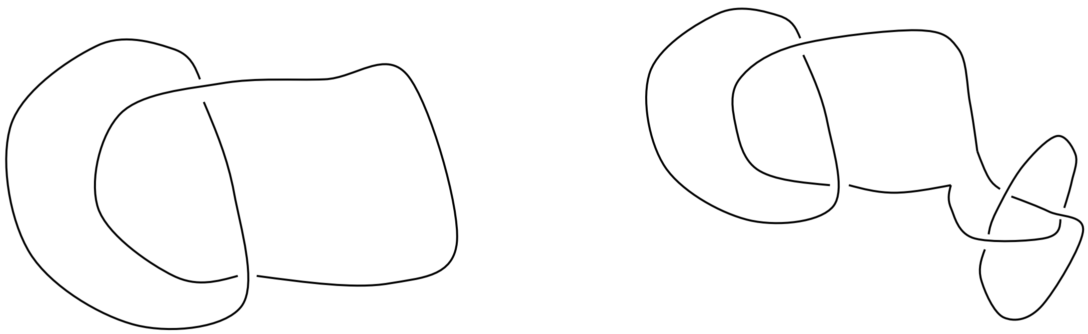


(d) The connected sum of unknot and trefoil knot

Figure 3.4: The connected sum of an unknot with a trefoil knot resulting in another trefoil knot.

from [37, 38], indicating the conditions under which a two-component link is non-splittable (topologically linked). We will use this result later, in chapters 4 and 5, to define a pattern that guarantees topological linking of polygons.

The following definitions are presented from [38]. A *2-string tangle* is a pair (B, t) where B is a 3-ball, i.e. it is homeomorphic to $\{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid \sqrt{x_1^2 + x_2^2 + x_3^2} \leq 1\} \subset \mathbb{R}^3$, and t is a set of two disjoint arcs properly embedded in B. Any subset E of $\mathbb{R}^3$ that is homeomorphic to $[0, 1]$ is called an *arc*. An arc E is said to be *properly embedded* in B if there is a continuous function $f : [0, 1] \to B$ such that $E = f([0, 1])$ is homeomorphic to $[0, 1]$ and the two endpoints of the arc, $f(0)$ and $f(1)$, are the only points in the boundary of B. Note that S^2 is the boundary of B. Any surface D that is homeomorphic to $D_1 \equiv \{(x_1, x_2, 0) \in \mathbb{R}^3 \mid \sqrt{x_1^2 + x_2^2} \leq 1\} \subset \mathbb{R}^3$ is called



(a) τ (b) τ'

Figure 3.5: τ' is an augmented link made by connecting a trefoil knot to one of τ's components. τ' has augmented link type $\tau^{\#}$ but doesn't have the same link type as τ.

a *disc*. A disc D is said to be *properly embedded* in B if there is a continuous function $f : D_1 \to B$ such that $D = f(D_1)$ is homeomorphic to D_1 and ∂D is a subset of the boundary of B and no other point of D intersects the boundary of B. Two tangles are considered *equivalent* if there is an ambient isotopy of one tangle to the other keeping the boundary of the 3-ball fixed. A tangle (B, t) is *locally trivial* if any 2-sphere in B which meets t transversely in two points bounds in B the unknotted tangle, i.e. the trivial tangle on one string. A 2-string tangle (B, t) is *inseparable* if the two arcs cannot be separated by a disk properly embedded in B. A tangle is *prime* if it is locally trivial and inseparable, see [30, 38] for more details on tangles. Examples of trivial, prime and separable tangles are illustrated in Figure 3.1.

Let S^3 denote the three dimensional sphere, i.e. $S^3 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2} = 1\} \subset \mathbb{R}^4$. The following is shown in [37, 38] (actually it is shown in more general terms involving tangles on n strings):

Lemma 3.1.8 ([37, 38])**.** *(i) Let (C, v) be a 2-string tangle and let D be a disk properly embedded in C that intersects both arcs of (C, v) in a single point each and separates (C, v) into two 2-string tangles (A, t) and (B, u). Let (B, u) be an inseparable tangle. Suppose for any disk D' properly embedded in A with $D' \cap \partial D = \emptyset$ and $D' \cap t = \emptyset$ that D' does not separate t, then (C, v) is an inseparable tangle.*

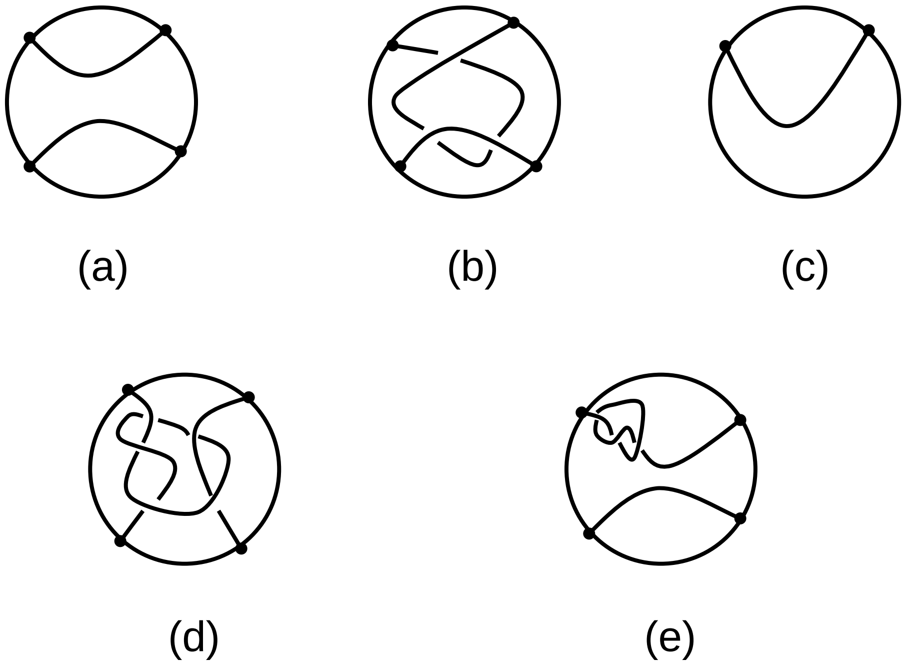


Figure 3.6: (a) is the trivial 2-string tangle, (b) is equivalent to the trivial 2-string tangle, (c) is the trivial 1-string tangle, (d) is a prime tangle and (e) is a separable tangle that is not locally trivial so it is not a prime tangle.

(ii) Let L be a link in S^3. Suppose that S is a 2-sphere in S^3 meeting L transversely in 4 points and dividing (S^3, L) into two 2-string tangles $(A, A \cap L)$ and $(B, B \cap L)$. Then if the two tangles $(A, A \cap L)$ and $(B, B \cap L)$ are inseparable, then L is non-splittable.

3.2 Asymptotic Behaviour of Polygonal Knots and Links in $\mathbb{Z}^3$

In this section some results on the asymptotic behaviour of the number of polygonal knots and links, with some specific topological properties, in $\mathbb{Z}^3$ are reviewed. In particular, the knotting probability of lattice polygons, considered as a knot or as a component of a two-component link in $\mathbb{Z}^3$, is discussed.

Let p_n^o be the number of n-edge unknotted SAPs in $\mathbb{Z}^3$ up to translation. The following theorem holds.

Theorem 3.2.1 (Sumners and Whittington 1988 [55])**.**

$$\lim_{n\to\infty} n^{-1}\log p_n^o \equiv \kappa_o < \kappa, \tag{3.2.1}$$

hence the probability that an n-edge self-avoiding polygon is knotted, $\frac{p_n - p_n^o}{p_n}$, goes to unity as $1 - e^{-\alpha n + o(n)}$ when $n \to \infty$, with $\alpha \equiv \kappa - \kappa_o$.

(a) A tight trefoil pattern

(b) An occurrence of a tight trefoil pattern in a knot

Figure 3.7: A tight trefoil pattern such that its occurrence in any polygon guarantees that the polygon is knotted.

The proof of Theorem 3.2.1 is based on the pattern theorem for SAPs in $\mathbb{Z}^3$, Theorem 2.3.1, and the fact that any lattice polygon containing a tight trefoil pattern (see Figure 3.7) must be knotted. We will later use a similar strategy to discuss the topological linking probability of two self-avoiding polygons. Note that the existence of the connective constant for unknotted SAPs is mainly based on the fact that concatenating any two unknotted SAPs, as described in Section 2.2, results in another unknotted SAP.

Given a non-splittable two-component link τ, let $p_n^{(2)}(\tau^\#)$ denote the number of polygonal links (ω_1, ω_2) in $\mathbb{Z}^3$ (i.e. (ω_1, ω_2) has its vertices in $\mathbb{Z}^3$) having the same augmented link type as $\tau^{\#}$ and consisting of two mutually avoiding self-avoiding polygons each with n edges (up to translation). The following theorem explores the asymptotic behaviour of the number of such polygonal links in $\mathbb{Z}^3$.

Theorem 3.2.2 (Orlandini *et al.* 1994 [39])**.** *Let τ be a non-splittable two-component link. Then*

$$\lim_{n\to\infty} \frac{\log p_n^{(2)}(\tau^{\#})}{2n} = \kappa. \tag{3.2.2}$$

Given a two-component link τ, let $p_n^{(2)}(M,\tau^{\#})$ denote the number of polygonal links (ω_1,ω_2) in a cube of side M, having the same augmented link type as $\tau^{\#}$ and consisting of two mutually avoiding self-avoiding polygons each with n edges.

Theorem 3.2.3 (Orlandini *et al.* 1994 [39])**.** *Let τ be a two-component link. Then*

$$\lim_{n,M\to\infty} \frac{\log p_n^{(2)}(M,\tau^{\#})}{2n} = \kappa \tag{3.2.3}$$

provided that n and M both go to infinity such that $M \geq n+q$ and $M = e^{o(n)}$. Here q is independent of n and M, but may depend on τ.

In order to investigate the asymptotic behaviour of polygonal links, one must somehow constrain the two polygons so that there is a finite number of configurations with n-edges. In Theorem 3.2.2, one type of constraint is considered by assuming τ to be a non-splittable link. Confinement of polygonal links to a cube of side M is also another type of constraint which is assumed in Theorem 3.2.3. Next, we will consider a similar but different type of constraint and investigate the asymptotic behaviour of both splittable and non-splittable polygonal links in $\mathbb{Z}^3$. Given any pair of edges e and f in $E(\mathbb{Z}^3)$, the *distance between these two edges*, $d(e,f)$, is defined to be the Euclidean distance between the midpoints of e and f. Given a non-negative integer k, let $p_n^{(2)}(k)$ be the number (up to translation) of polygonal links (ω_1,ω_2) in $\mathbb{Z}^3$ consisting of two mutually avoiding self-avoiding polygons each with n edges and having a pair of edges, one from each polygon, within distance k of each other.

Let $p_n^{(2)}(u;k)$ denote the number (up to translation) of such polygonal links in $\mathbb{Z}^3$ which are non-splittable. Hence $p_n^{(2)}(s;k) \equiv p_n^{(2)}(k) - p_n^{(2)}(u;k)$ is the number of those polygonal links in $\mathbb{Z}^3$ which are spittable. The following theorem holds.

Theorem 3.2.4.

$$\lim_{n\to\infty} \frac{\log p_n^{(2)}(k)}{2n} = \lim_{n\to\infty} \frac{\log p_n^{(2)}(u;k)}{2n} = \lim_{n\to\infty} \frac{\log p_n^{(2)}(s;k)}{2n} = \kappa. \qquad (3.2.4)$$

Proof. The proof presented here is a straightforward modification of the proof of Theorem 3.2.2 given in [39, Theorem 2.2].

Let (ω_1, ω_2) be a polygonal link in $\mathbb{Z}^3$ consisting of two mutually avoiding self-avoiding polygons ω_1 and ω_2 each with n edges and having a pair of edges, one from each polygon, within distance k of each other. Let v be the bottom vertex of ω_1. ω_1 is confined to a cube, C_1, of side n with bottom vertex $v - (\frac{n}{2}, \frac{n}{2}, \frac{n}{2})$, in particular, because the span of an n-edge polygon cannot exceed $\frac{n}{2}$. Let C_2 be a cube, containing C_1, with side length $2n+2k$ and bottom vertex $v - (n+k, n+k, n+k)$. Note that (ω_1, ω_2) cannot contain a pair of edges, one from each polygon, k units apart if the bottom vertex of the other polygon, ω_2, is not within the cube C_2. So we can construct each polygon in p_n ways and translate one relative to the other in at most $(2n+2k)^3$ positions. Hence the following upper bound is obtained for $p_n^{(2)}(k)$:

$$p_n^{(2)}(k) \leq (2n+2k)^3 p_n^2. \qquad (3.2.5)$$

Taking logarithms, dividing both sides of the above equation by $2n$ and taking the limit superior as $n \to \infty$ gives

$$\limsup_{n\to\infty} \frac{\log p_n^{(2)}(k)}{2n} \leq \kappa. \qquad (3.2.6)$$

Since $p_n^{(2)}(u;k) \leq p_n^{(2)}(k)$ and $p_n^{(2)}(s;k) \leq p_n^{(2)}(k)$, the above equation also implies

$$\limsup_{n\to\infty} \frac{\log p_n^{(2)}(u;k)}{2n} \leq \kappa \qquad (3.2.7)$$

and

$$\limsup_{n\to\infty} \frac{\log p_n^{(2)}(s;k)}{2n} \leq \kappa. \tag{3.2.8}$$

Define the two-component polygonal link (ω_1', ω_2') as follows (e.g. see Figure 3.8 (a)): Let ω_1' be the polygon which is the undirected version of the closed walk $((1+k)\hat{j}, 2\hat{i}, \hat{k}, -(1+k)\hat{j}, -\hat{k}, -2\hat{i})$ starting at the origin. Let also ω_2' be the polygon which is the undirected version of the closed walk $(\hat{k}, (1+k)\hat{j}, -\hat{k}, 2\hat{i}, -(1+k)\hat{j}, -2\hat{i})$ starting at $(1,1,0)$. Note that each polygon has $2k+8$ edges, and (ω_1', ω_2') is a non-splittable link containing the pair of edges $e_1' = \{(1, 1+k, 0), (2, 1+k, 0)\}$ and $e_2' = \{(1,1,0), (2,1,0)\}$, respectively from ω_1' and ω_2', exactly k units apart. Given any pair of $(n-2k-8)$-edge polygons G_1 and G_2 in $\mathbb{Z}^3$, using the concatenation argument for SAPs in $\mathbb{Z}^3$, one can easily concatenate G_1 to ω_1' and ω_2' to G_2 to obtain a new pair of polygons $G_1 \circ \omega_1'$ and $\omega_2' \circ G_2$. An example is illustrated in Figure 3.8. Regarding the construction, $(G_1 \circ \omega_1', \omega_2' \circ G_2)$ is a non-splittable link containing a pair of edges exactly k units apart. Note that, by the concatenation argument for SAPs in $\mathbb{Z}^3$, both G_1 (G_2) and the polygon obtained by rotating G_1 (G_2) give rise to the same polygon $G_1 \circ \omega_1'$ $(\omega_2' \circ G_2)$. So there are exactly $p_{n-2k-8}/2$ choices for G_1 (G_2). Hence the following lower bound is obtained for $p_n^{(2)}(u;k)$.

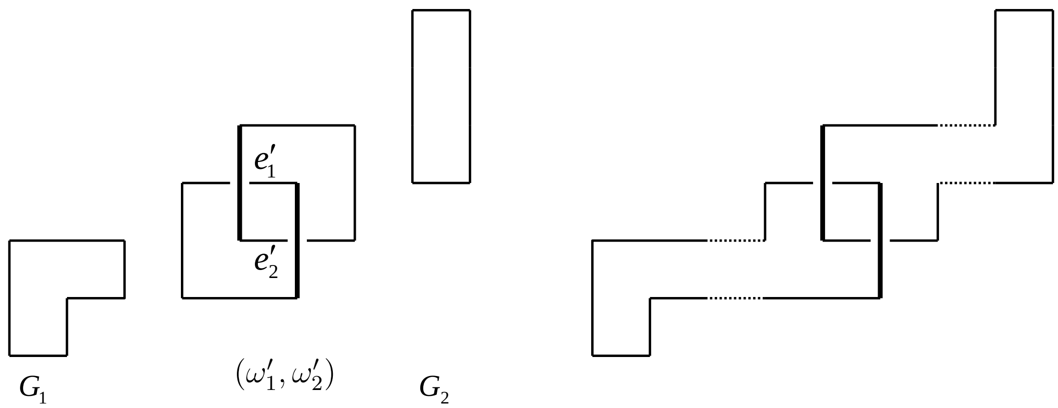
$$p_{n-2k-8}^2/4 \leq p_n^{(2)}(u;k). \tag{3.2.9}$$

Taking logarithms, dividing both sides of the above equation by $2n$ and taking the limit inferior as $n \to \infty$ gives

$$\kappa \leq \liminf_{n\to\infty} \frac{\log p_n^{(2)}(u;k)}{2n}. \tag{3.2.10}$$

Therefore, equations 3.2.7 and 3.2.10 together yield

$$\lim_{n\to\infty} \frac{\log p_n^{(2)}(u;k)}{2n} = \kappa. \tag{3.2.11}$$



(a) The 8-edge polygons G_1 and G_2, and the 20-edge link (ω_1', ω_2').

(b) The 36-edge link $(G_1 \circ \omega_1', \omega_2' \circ G_2)$.

Figure 3.8: An example of constructing $(G_1 \circ \omega_1', \omega_2' \circ G_2)$ from G_1, G_2 and (ω_1', ω_2'). Note that the edges shown with thicker lines are in the plane $z = 1$ and the rest are in the plane $z = 0$. The four edges added to concatenate G_1 to ω_1' and ω_2' to G_2 are shown with dashed lines. Note that in this example $k = 1$, $n = 18$ and the two edges $e_1' = \{(1, 2, 0), (2, 2, 0)\}$ and $e_2' = \{(1, 1, 0), (2, 1, 0)\}$ are exactly 1 unit apart.

A similar argument to that presented above works also for splittable polygons. Define the two-component polygonal link (ω_1'', ω_2'') as follows: Let ω_1'' be the polygon which is the undirected version of the closed walk $(\hat{\jmath}, \hat{\imath}, -\hat{\jmath}, -\hat{\imath})$ starting at the origin. Let also ω_2'' be the polygon which is the undirected version of the closed walk $(\hat{\jmath}, \hat{\imath}, -\hat{\jmath}, -\hat{\imath})$ starting at $(k + 1, 0, 0)$. Note that each polygon has 4 edges and (ω_1'', ω_2'') is a splittable link with the pair of edges $e_1'' = \{(1, 0, 0), (1, 1, 0)\}$ and $e_2'' = \{(1 + k, 0, 0), (1 + k, 1, 0)\}$, respectively from ω_1'' and ω_2'', exactly k units apart. Given any pair of $(n - 4)$-edge polygons G_1 and G_2 in $\mathbb{Z}^3$, using the concatenation argument for SAPs in $\mathbb{Z}^3$, one can easily concatenate G_1 to ω_1'' and ω_2'' to G_2 to obtain a new pair of polygons $G_1 \circ \omega_1''$ and $\omega_2'' \circ G_2$. Regarding the construction, $(G_1 \circ \omega_1'', \omega_2'' \circ G_2)$ is a splittable link containing a pair of edges exactly k units apart. Note that, by the concatenation argument for SAPs in $\mathbb{Z}^3$, both G_1 (G_2) and the polygon obtained by rotating G_1 (G_2) give rise to the same polygon $G_1 \circ \omega_1''$ ($\omega_2'' \circ G_2$).

So there are exactly $p_{n-4}/2$ choices for G_1 (G_2). Hence the following lower bound is obtained for $p_n^{(2)}(s;k)$.

$$p_{n-4}^2/4 \le p_n^{(2)}(s;k). \tag{3.2.12}$$

Taking logarithms, dividing both sides of the above equation by $2n$ and taking the limit inferior as $n \to \infty$ gives

$$\kappa \le \liminf_{n\to\infty} \frac{\log p_n^{(2)}(s;k)}{2n}. \tag{3.2.13}$$

Therefore, equations 3.2.8 and 3.2.13 together yield

$$\lim_{n\to\infty} \frac{\log p_n^{(2)}(s;k)}{2n} = \kappa. \tag{3.2.14}$$

Furthermore, the fact that $p_n^{(2)}(u;k) \le p_n^{(2)}(k)$, equation 3.2.6 and 3.2.10 imply that

$$\lim_{n\to\infty} \frac{\log p_n^{(2)}(k)}{2n} = \kappa. \tag{3.2.15}$$

Therefore, equations 3.2.11, 3.2.14 and 3.2.15 yield

$$\lim_{n\to\infty} \frac{\log p_n^{(2)}(k)}{2n} = \lim_{n\to\infty} \frac{\log p_n^{(2)}(u;k)}{2n} = \lim_{n\to\infty} \frac{\log p_n^{(2)}(s;k)}{2n} = \kappa. \tag{3.2.16}$$

□

Like the knotting probability of a single polygon, one can ask under what conditions are all but exponentially few sufficiently long pairs of polygons linked? If the two polygons are constrained to have a pair of edges, one from each polygon, within a fixed distance from each other, Theorem 3.2.4 establishes that the exponential growth rate of the number of topologically linked polygon pairs (up to translation) is equal to that of the number of topologically unlinked polygon pairs. So, unlike the situation with knotting, we cannot say that all but exponentially few sufficiently long pairs of self-avoiding polygons are linked, even with this distance constraint. It is possible (although not proved) that the linking probability goes to one as n goes to infinity, but it will not go to one exponentially rapidly.

The following theorem investigates the knotting probability of the knot components of the polygonal links in $\mathbb{Z}^3$ which have the same augmented link type as $\tau^{\#}$, for a non-splittable link τ.

Theorem 3.2.5 (Orlandini *et al.* 1994 [39])**.** *Let τ be a non-splittable two-component link. The probability $\mathbb{P}^{(2)}(n, \tau^{\#})$ that both components (each with n edges) of a polygonal link (ω_1, ω_2) in $\mathbb{Z}^3$ with augmented link type $\tau^{\#}$ are knotted goes to unity as*

$$\mathbb{P}^{(2)}(n, \tau^{\#}) = 1 - e^{-\alpha n + o(n)} \tag{3.2.17}$$

when $n \to \infty$. α is the value introduced in Theorem 3.2.1.

Given a non-splittable two-component link τ with each component the unknot, let $p_n^{(2)}(\tau)$ denote the number of polygonal links (ω_1, ω_2) in $\mathbb{Z}^3$ having the same link type as τ and consisting of two mutually avoiding unknotted self-avoiding polygons each with n edges (up to translation). The following theorem shows that $p_n^{(2)}(\tau)$ grows exponentially with n and the exponential growth rate is independent of the link type τ [39, 52].

Theorem 3.2.6 (Orlandini *et al.* 1994 [39])**.** *Let τ be a non-splittable two-component link with each component the unknot. Then*

$$\lim_{n \to \infty} \frac{\log p_n^{(2)}(\tau)}{2n} = \kappa_o. \tag{3.2.18}$$

Note that theorems 3.2.2, 3.2.4 and 3.2.5 can be easily extended to obtain similar results for k-component links, i.e. disjoint unions of k self-avoiding polygons [52].

3.3 Asymptotic Behaviour of Polygonal Knots and Links in Tubes

The goal of this section is to discuss the effects of the tube constraint on the topological entanglement of polygons measured by knotting and linking probability.

Let $p_n^o(N,M)$ denote the number (up to x-translation) of n-edge unknotted SAPs in an (N,M)-tube. The following theorem shows that a similar result to Theorem 3.2.1 holds for lattice polygons in tubes, i.e.

Theorem 3.3.1 (Soteros 1998 [51])**.** *For N and M such that the (N,M)-tube can contain a tight trefoil $\tilde{K}_b$ pattern for some $b > 0$*

$$\lim_{n\to\infty} n^{-1} \log p_n^o(N,M) \equiv \kappa_o(N,M) < \kappa_p(N,M) \tag{3.3.1}$$

and hence the probability that a self-avoiding polygon in an (N,M)-tube is knotted goes to unity as $1 - e^{-\alpha(N,M)n+o(n)}$ when $n \to \infty$, with $\alpha(N,M) = \kappa_p(N,M) - \kappa_o(N,M)$.

The idea of the proof of this theorem is similar to that given for Theorem 3.2.1 except that here the pattern theorem for SAPs in tubes, Theorem 2.4.2, is used instead of that for SAPs in $\mathbb{Z}^3$.

Given a non-splittable two-component link τ, let $p_n^{(2)}((N,M),\tau^\#)$ denote the number of polygonal links (ω_1,ω_2) in $T(N,M)$ having the same augmented link type as $\tau^\#$ and consisting of two mutually avoiding self-avoiding polygons each with n edges (up to x-translation). The following theorem is a straightforward modification of Theorem 3.2.2 for polygonal links in tubes.

Theorem 3.3.2 (Tesi *et al.* [58])**.** *Let τ be a non-splittable two-component link. Then*

$$\lim_{n\to\infty} \frac{\log p_n^{(2)}((N,M),\tau^\#)}{2n} = \kappa_p(N,M). \tag{3.3.2}$$

Let $p_n^{(2)}((N,M),u;k)$ ($p_n^{(2)}((N,M),s;k)$) denote the number (up to x-translation) of non-splittable (splittable) polygonal links (ω_1,ω_2) in $T(N,M)$ consisting of two mutually avoiding self-avoiding polygons each with n edges and having a pair of edges, one from each polygon, within a fixed distance k from each other. The following theorem explores the asymptotic behaviour of the number of such polygonal links in $T(N,M)$.

Theorem 3.3.3.

$$\lim_{n\to\infty} \frac{\log p_n^{(2)}((N,M),u;k)}{2n} = \lim_{n\to\infty} \frac{\log p_n^{(2)}((N,M),s;k)}{2n} = \kappa_p(N,M). \qquad (3.3.3)$$

We can use an argument similar to that given for the proof of Theorem 3.2.4 to prove Theorem 3.3.3, except that the concatenation for SAPs in a tube will be needed instead of that for SAPs in $\mathbb{Z}^3$. Theorem 3.3.3 shows that in a tube, as in $\mathbb{Z}^3$, even if two edges, with one from each polygon, are forced to be close, the rest of each polygon has a considerable amount of freedom so that their centres of mass can be very far apart. Therefore, like the situation in $\mathbb{Z}^3$, we cannot say that in a tube all but exponentially few sufficiently long pairs of self-avoiding polygons are linked. However, later in Chapter 4, we will consider a much more severe distance constraint and will prove that in this case for large enough values of n the two polygons would be linked with high probability (except here n will be the total number of edges).

Given a non-splittable two-component link τ with each component the unknot, let $p_n^{(2)}((N,M),\tau)$ denote the number of polygonal links (ω_1,ω_2) in $T(N,M)$ having the same link type as τ and consisting of two mutually avoiding unknotted self-avoiding polygons each with n edges (up to x-translation). A straightforward modification of Theorem 3.2.6 implies the following result for polygons in a tube.

Theorem 3.3.4 (Tesi *et al.* [58])**.** *Let τ be a non-splittable two-component link with each component the unknot. Then*

$$\lim_{n\to\infty} \frac{\log p_n^{(2)}((N,M),\tau)}{2n} = \kappa_o(N,M). \qquad (3.3.4)$$

Chapter 4

The Linking Probability of Two SAPs Confined To and Spanning a Tube

4.1 Introduction

Polygonal self-entanglements have been investigated using topological measures such as the probability of knotting. In particular, as discussed in Section 3.2, it has been proved, both in $\mathbb{Z}^3$ and in a tube, that all but exponentially few sufficiently long self-avoiding polygons are knotted and thus the probability of knotting of lattice polygons approaches one as the size of a polygon goes to infinity [43, 51, 55]. One can ask similar questions regarding the entanglement complexity of two self-avoiding polygons. For example, under what conditions are all but exponentially few sufficiently long pairs of polygons linked? If the two polygons are constrained to have a pair of edges, one from each polygon, which are within a fixed distance from each other, Theorem 3.2.4 established that the exponential growth rate of the number of topologically linked polygon pairs (up to translation) is equal to that of the number of topologically unlinked polygon pairs. So, unlike the situation with knotting, we cannot say that all but exponentially few sufficiently long pairs of self-avoiding polygons are linked, even with this distance constraint. It is possible (although not proved) that the linking probability goes to one as n goes to infinity, but it will not go to one exponentially rapidly.

As discussed previously in Theorem 3.3.2, Tesi *et al* [58] investigated the same question for pairs of mutually avoiding self-avoiding polygons confined to tubes and came to the same conclusion. In this chapter we consider a much more severe distance constraint. The two polygons are constrained in such a way that (roughly speaking) each edge of one polygon is forced to be "close" to some edge of the other polygon. In this case one might expect that for large enough values of n the two polygons would be linked with high probability. This type of constraint was investigated numerically by Orlandini *et al* [39] and Tesi *et al* [58] who considered two n-edge polygons confined to a lattice cube with side length M and used Monte Carlo simulation to estimate the average linking number as a function of n and M; their results (see for example [39, Fig. 3] or [58, Fig. 7]) show that the probability that two polygons in a cube are homologically linked (i.e. have non-zero linking number) increases with increasing n for fixed M, and with decreasing M for fixed n. Thus, as expected, more confined polygon pairs have a higher linking probability. Their results also suggested that topologically linked pairs are also homologically linked with high probability for configurations in which the two polygons are strongly interpenetrating, whereas this probability becomes smaller for configurations in which the two components are, on average, further apart [39, p.342]. Off the lattice, Arsuaga *et al* [2] investigated both theoretically and numerically uniform random polygons in the unit cube in $\mathbb{R}^3$. For a fixed simple closed curve S, they proved that (as $n \to \infty$) the probability that the linking number, $Lk(S, R_n)$, is non-zero for any two-component link (S, R_n), with R_n an n-edge uniform random polygon, approaches 1 at least as fast as $1 - O(1/\sqrt{n})$. Their numerical results indicate that the probability that $Lk(R_n, R_m)$ is non-zero, with (R_n, R_m) a pair of uniform random polygons having n and m edges respectively, goes to 1 like $1 - O(1/\sqrt{nm})$. Herein, we do not confine the polygons to a cube but rather we consider two self-avoiding polygons confined to a lattice tube ($\infty \times N \times M$) with an added constraint that forces each edge of one polygon to be close to some edge of the other polygon and prove theoretical results about the linking probability (both homologically and topologically).

Specifically we consider a pair of mutually avoiding self-avoiding polygons each confined to and spanning a tube, i.e. each with the same span in the tube direction (see Figure 4.1). Such a pair is referred to as a System of two Self-avoiding Polygons and n is now used to denote the total number of edges in the pair. We establish that the homological linking probability goes to one at least as fast as $1-O(n^{-1/2})$ and that the topological linking probability goes to one exponentially rapidly. Furthermore we prove for this model that the linking number grows (with probability one) faster than any function that is $o(\sqrt{n})$; i.e. for any function $f(n) = o(\sqrt{n})$, there exists $A \geq 0$ such that as $n \to \infty$ the probability that $|Lk(\omega_1, \omega_2)| \geq f(n)$, with (ω_1, ω_2) the component polygons of an n-edge 2SAP, satisfies

$$\mathbb{P}\big(|Lk(\omega_1, \omega_2)| \geq f(n)\big) = 1 - \frac{A}{\sqrt{n}} + o\big(\frac{1}{\sqrt{n}}\big). \tag{4.1.1}$$

Hence

$$\lim_{n\to\infty} \mathbb{P}\big(|Lk(\omega_1, \omega_2)| \geq f(n)\big) = 1. \tag{4.1.2}$$

We also show that the linking number cannot grow faster than linearly in n because of the tube constraint; i.e. there exist constants a and b such that for any n-edge 2SAP

$$|Lk(\omega_1, \omega_2)| \leq an + b. \tag{4.1.3}$$

We give a simple example to show that the upper bound in equation (4.1.3) for 2SAPs can be realized. Thus the homological linking probability goes to one at least as fast as $1 - O(n^{-1/2})$ and at most linearly in n.

This chapter is organized as follows. In Section 4.2, we introduce a precise definition of our model. Then, in Section 4.3, two lemmas are proved to show that (CONCAT) and (CAPOFF) are satisfied for 2SAPs. We first use these, in Section 4.4, to establish the existence of the connective constant for 2SAPs and show that it is strictly less than that of SAPs. Then the lemmas of Section 4.3 are used in Section 4.5 to prove a pattern theorem for 2SAPs. In Section 4.6, the pattern theorem and a pattern proved by C. Ernst to guarantee topological linking is used to establish the results on topological linking of 2SAPs. In Section 4.7, the pattern theorem combined with some techniques similar to the ones presented in [24] implies a lower bound (with probability one) on the rate of increase of the linking number of 2SAPs and proves that the linking probability goes to one as the length of a 2SAP goes to infinity. In Section 4.8, we give an upper bound for the linking number of 2SAPs in a tube which is of smaller order than the known upper bound for links in $\mathbb{Z}^3$ due to the tube constraint. Finally, the results presented in this chapter are summarized in Section 4.9.

4.2 2SAPs

In this chapter, we are going to investigate the linking probability of two mutually avoiding SAPs confined to a tube under some constraints. The model for the configurations of a pair of SAPs is defined as follows:

Definition 4.2.1. A *System of two Self-avoiding Polygons* (2SAP) of size n in an (N, M)-tube is a finite subgraph G of $T(N, M)$ satisfying the following conditions:

(i) The total number of edges in G is n (n even).

(ii) Each vertex of G has degree two and G has exactly two connected components ω_1 and ω_2 (hence ω_1 and ω_2 are mutually avoiding self-avoiding polygons), where ω_1 is the component that contains the lexicographically smallest vertex of G. We write $G := (\omega_1, \omega_2)$.

(iii) There exists $x_1, x_2 \in \mathbb{Z}$ such that the components ω_1 and ω_2 each *span* the length (in the x-direction) of the subtube with vertex set $\{(x, y, z) \in \mathbb{Z}^3 | x_1 \leq x \leq x_2, 0 \leq y \leq N, 0 \leq z \leq M\}$. That is,

$$\min\{x|(x, y, z) \in \omega_1\} = \min\{x|(x, y, z) \in \omega_2\} = x_1$$

and

$$\max\{x|(x, y, z) \in \omega_1\} = \max\{x|(x, y, z) \in \omega_2\} = x_2.$$

Consistent with the definitions of the span and the left-most (right-most) plane for any finite subgraph of $\mathbb{Z}^3$, presented in Section 2.1, we say $x_2 - x_1$ is the *span* of the 2SAP and that $x = x_1$ ($x = x_2$) is the left-most (right-most) plane of the 2SAP. See Figure 4.1 for an example of a 2SAP. Note that in these figures (and all other figures in this chapter), edges of one component polygon are illustrated with thicker edges than those of the other component polygon. Also note that by Definition 4.2.1, there exists no 2SAP in the (N, M)-tube for the choices of $(N, M) \in \{(0,0), (0,1), (1,0), (0,2), (2,0)\}$ and at most two n-edge 2SAPs for any n when $(N, M) = (1,1)$. For convenience and since we are focussing on three dimensions, when considering 2SAPs in an (N, M)-tube, we refer to the set of pairs of positive integers $(N, M) \neq (1,1)$ as *allowed* tube dimensions and, unless stated otherwise, assume that (N, M) is restricted to the allowed pairs.

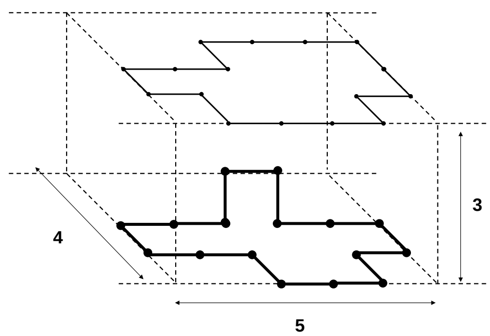


Figure 4.1: A 2SAP with size 32 and span 5 in a $(4,3)$-tube.

4.3 (CONCAT) and (CAPOFF) for 2SAPs

In order to prove our main results about linking probabilities we need to establish a pattern theorem for 2SAPs. As discussed in Section 2.5, this can be done by establishing that it is possible to insert any pattern at an arbitrary location in a 2SAP as depicted in Figure 4.6 and as described in Proposition 2.5.1 of Section 4.5. The two lemmas, (CONCAT) and (CAPOFF), presented in this section are useful for establishing Proposition 2.5.1. In addition, (CONCAT) will be used to establish the existence of the connective constant for 2SAPs in Section 4.4.

Following the terminology of Section 2.5, we adapt the arguments given in Section 2.5 for 2SAPs. For each positive integer n, let Q_n^* be the set of all n-edge 2SAPs whose lexicographically smallest vertex (bottom vertex) is in the plane $x = 0$. Define

$Q^* = \bigcup_{n<\infty} Q_n^*$. For any n-edge 2SAP G, the translation $(-x_1, 0, 0)$ along the x-axis gives an element of Q_n^*, i.e. $G + (-x_1, 0, 0) \in Q_n^*$. Let $b_G = x_2 - x_1$ denote the span of G. Hence $x = b_G + x_1$ is the right-most plane of G.

The following lemmas show that for 2SAPs there exist choices of c_T, t_T and m_T for which (CONCAT) and (CAPOFF) hold.

Lemma 4.3.1. *For any allowed tube dimensions (N, M), any choice of $t_T \geq 18 + 8(N + M)$ and $c_T = 4t_T + 8(N + M) + 4$, (CONCAT) holds for 2SAPs in $T(N, M)$.*

Proof. In particular we prove that there exists integers $t_T \geq 1$, $c_T \geq 1$ and a concatenation process defined for 2SAPs in $T(N, M)$ such that:

Given $G_1 = (\omega_1, \omega_2) \in Q_n^*$ and $G_2 \in Q_m^*$ with respective spans b_1 and b_2, concatenating G_1 to the translate, $G_2 + (t_T + b_1, 0, 0)$, forms $G \in Q_{n+m+c_T}^*$ such that $G \cap V_{b_1-1} = G_1 \cap V_{b_1-1}$ and $G \cap (V_{b_2-1} + (t_T + b_1 + 1, 0, 0)) = [G_2 \cap (V_{b_2-1} + (1, 0, 0))] + (t_T + b_1, 0, 0)$ (i.e. only the right-most plane of G_1 and the left-most plane of G_2 can be altered in the concatenation process).

Define $e_1 = \{v_1, v_2\}$ $(v_1 < v_2)$ and $e_2 = \{u_1, u_2\}$ $(u_1 < u_2)$ such that e_1 is the bottom edge in the plane $x = b_1$ of one of the polygons of G_1, e_2 is the bottom edge in the plane $x = b_1$ of the other polygon of G_1 and $v_1 < u_2$, where in the case of ambiguity v_1 is chosen such that $z(v_1) < z(u_1)$ (see for example Figure 4.2 (a)). Similarly, we can obtain the two edges $\hat{e}_1 = \{\hat{v}_1, \hat{v}_2\}$ and $\hat{e}_2 = \{\hat{u}_1, \hat{u}_2\}$ on the left-most plane $(x = 0)$ of G_2 with $\hat{v}_1 < \hat{u}_2$, where in the case of ambiguity $\hat{v}_1$ is chosen such that $z(\hat{v}_1) < z(\hat{u}_1)$.

The strategy for the proof of this lemma is to first show that there exists an integer t such that it is always possible to connect $G_1 \backslash \{e_1, e_2\}$ with $(G_2 \backslash \{\hat{e}_1, \hat{e}_2\}) + (t + b_1, 0, 0)$ by four mutually self-avoiding walks inside the tube; the walks connect the endpoints v_i, u_i, $i = 1, 2$, to the endpoints $\hat{v}_i + (t + b_1, 0, 0)$, $\hat{u}_i + (t + b_1, 0, 0)$, $i = 1, 2$, so that the right-most plane of G_1 is connected to the left-most plane of $G_2 + (t + b_1, 0, 0)$ and the result is a 2SAP. Then we give an argument for determining appropriate

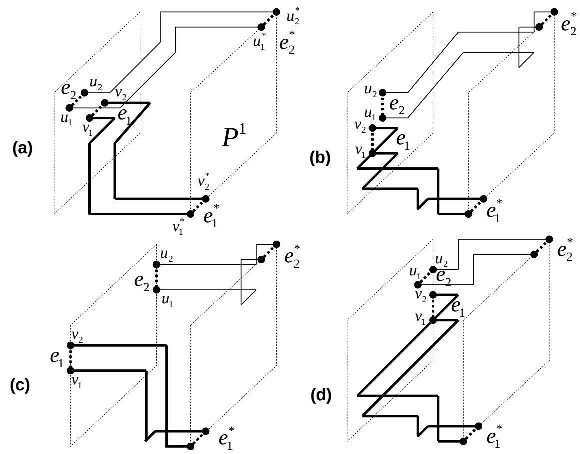


Figure 4.2: Converting (e_1, e_2) to (e_1^*, e_2^*) as needed for the proof of Lemma 4.3.1.

values of t_T and c_T for which (CONCAT) holds (see for example Figure 4.3).

Let $v_1^* = (b_1 + 5, 0, 0)$, $v_2^* = (b_1 + 5, 1, 0)$, $u_1^* = (b_1 + 5, N - 1, M)$ and $u_2^* = (b_1 + 5, N, M)$. Let $e_1^* = \{v_1^*, v_2^*\}$ and $e_2^* = \{u_1^*, u_2^*\}$, as illustrated in Figure 4.2. We will explicitly construct four paths that will connect the endpoints of e_i to the endpoints of e_i^* for $i = 1, 2$. Then by symmetry, there are also four paths that connect the endpoints of $\hat{e}_i + (b_1 + 10, 0, 0)$ to the endpoints of e_i^* for $i = 1, 2$. Therefore, taking the union of these paths, we obtain a connection of the endpoints of the edge e_i for $i = 1, 2$ to, respectively, the endpoints of $\hat{e}_i + (b_1 + 10, 0, 0)$ for $i = 1, 2$ (see Figure 4.3 (a)).

The following algorithm describes the paths connecting e_i to e_i^*, $i = 1, 2$ respectively.

1) Extend all the vertices v_1, v_2, u_1 and u_2 one unit along the positive x direction; i.e. add the edges $\{v_1, v_1 + \hat{\imath}\}$, $\{v_2, v_2 + \hat{\imath}\}$, $\{u_1, u_1 + \hat{\imath}\}$ and $\{u_2, u_2 + \hat{\imath}\}$.

2) If edge e_1 is horizontal then we do the following: Since the only occupied vertices in $\{(x, y, z) \in \mathbb{Z}^3 | b_1 + 1 \leq x \leq b_1 + 4, 0 \leq y \leq y(v_2), 0 \leq z \leq z(v_2)\}$ are the vertices created by the first step, continue by adding the walk $(-y(v_1)\hat{\jmath}, -z(v_1)\hat{k}, 4\hat{\imath})$ to v_1 to connect it to v_1^*. Similarly, add the walk $(\hat{\imath}, -(y(v_2) - 1)\hat{\jmath}, -z(v_2)\hat{k}, 3\hat{\imath})$ to v_2 to connect it to v_2^* (see Figure 4.2 (a)).

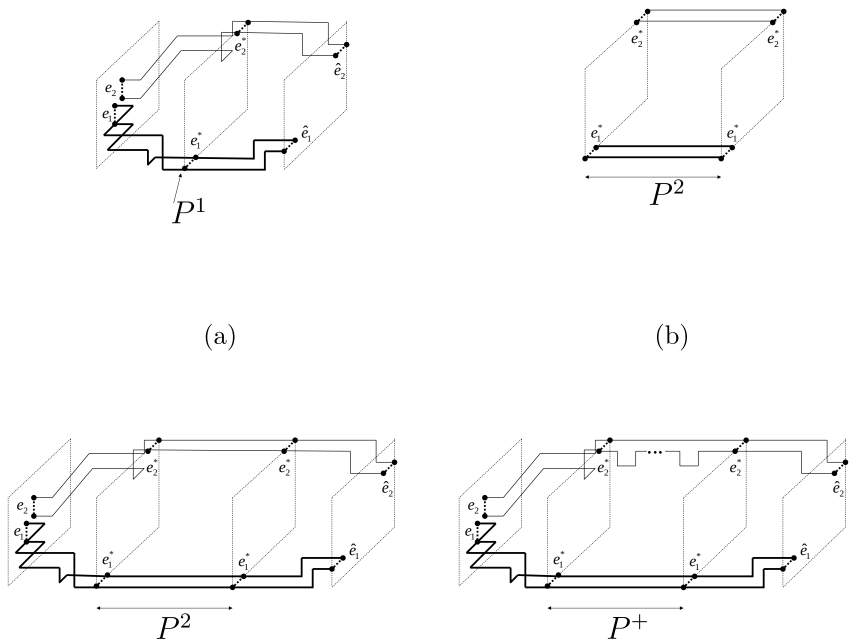
3) If edge e_1 was not horizontal then it is vertical. Instead of step two we do the following: As before, the only occupied vertices in $\{(x,y,z) \in \mathbb{Z}^3 | b_1+1 \leq x \leq b_1+4, 0 \leq y \leq y(v_1), 0 \leq z \leq z(v_2)\}$ are the vertices created by the first step. Hence continue by adding the walk $(-y(v_1)\hat{j}, 2\hat{i}, -z(v_1)\hat{k}, \hat{j}, 2\hat{i})$ to v_1 to connect it to v_2^*. Similarly, add the walk $(-y(v_2)\hat{j}, 3\hat{i}, -z(v_2)\hat{k}, \hat{i})$ to v_2 to connect it to v_1^* (see Figure 4.2 (b), (c) and (d)).

4) If edge e_2 is horizontal then we do the following: The fact that $v_1 < u_2$ guarantees that either $u_1 > v_2$ or $z(u_2) = z(u_1) > z(v_2) \geq z(v_1)$ (see Figures 4.2 (a) and (d)). Hence the previous constructions ensure that $\{(x,y,z) \in \mathbb{Z}^3 | b_1+1 \leq x \leq b_1+4, y(u_1) \leq y \leq N, z(u_1) \leq z \leq M\}$ contains only vertices created by the first step, and the following construction is possible. Continue by adding the walk $(\hat{i}, (N-1-y(u_1))\hat{j}, (M-z(u_1))\hat{k}, 3\hat{i})$ to u_1 to connect it to u_1^*. Similarly, add the walk $((N-y(u_2))\hat{j}, (M-z(u_2))\hat{k}, 4\hat{i})$ to u_2 to connect it to u_2^* (see Figure 4.2 (a) and (d)).

5) If edge e_2 was not horizontal then it is vertical. In this case, the fact that $v_1 < u_2$ guarantees that either $u_1 > v_2$ or $z(u_2) > z(u_1) > z(v_2) = z(v_1)$. Hence the previous constructions ensure that the only occupied vertices in $\{(x,y,z) \in \mathbb{Z}^3 | b_1+1 \leq x \leq b_1+4, y(u_1) \leq y \leq N, z(u_1) \leq z \leq M\}$ are vertices created by the first step. Instead of step three we now do the following: Continue by adding the walk $((N-y(u_1))\hat{j}, 3\hat{i}, -\hat{j}, (M-z(u_1))\hat{k}, \hat{i})$ to u_1 to connect it to u_1^*. Similarly, add the walk $((N-y(u_2))\hat{j}, 3\hat{i}, (M-z(u_2))\hat{k}, \hat{i})$ to u_2 to connect it to u_2^* (see Figure 4.2 (b), (c)).

The above construction results in four mutually avoiding SAWs with span 5 that connect the vertices of the edges (e_1, e_2) to the vertices of the edges (e_1^*, e_2^*). By symmetry, there are four mutually avoiding SAWs with span 5 that connect the vertices of the translated edges $(\hat{e}_1, \hat{e}_2) + (10+b_1, 0, 0)$ to the vertices of the edges (e_1^*, e_2^*), and hence we can concatenate G_1 with G_2 by inserting the four

SAWs of span 10 described by the above algorithm (see for example Figure 4.3 (a)). Let $c(N, M, e_1, e_2, \hat{e}_1, \hat{e}_2)$ be the number of edges added in this construction. Note that $40 \leq c(N, M, e_1, e_2, \hat{e}_1, \hat{e}_2) \leq 40 + 8(N + M + 1)$ and the exact value of $c(N, M, e_1, e_2, \hat{e}_1, \hat{e}_2)$ depends only on the edges $e_1, e_2, \hat{e}_1, \hat{e}_2$ from the 2SAPs G_1 and G_2.



(a) (b) (c) (d)

Figure 4.3: An example illustrating how to find t_T and c_T for the proof of Lemma 4.3.1. (a) Shows an example of the four SAWs of span 10 used to concatenate G_1 and G_2 in the proof. (b) Illustrates the $8(N + M + 1)-$pattern P^2. (c) Shows P^2 inserted at P^1. (d) Shows how to increase the number of edges by appropriately changing the pattern P^2 to P^+.

We now have to choose numbers $t_T \geq 10$ and c_T, such that (CONCAT) holds for all 2SAPs. To do this we need to insert an appropriate pattern into the concatenated 2SAPs. The appropriate pattern is constructed as follows (see for example Figure

4.3). Let P^2 be the $8(N+M+1)-$pattern illustrated in Figure 4.3 (b). P^2 can be inserted into the four SAWs by our construction. (The algorithm above describes the construction of a 5-pattern whose right-most-plane is the 0-pattern P^1 (see Figures 4.2 and 4.3 (a)); this allows for the insertion of P^2 starting at P^1, see Figure 4.3 (c).) Inserting P^2 increases the span size by $8(N+M+1)$ units and will allow a choice of $t_T = 18+8(N+M)$. Inserting P^2 also increases the number of edges by $32(N+M+1)$. Let $c = c(N, M, e_1, e_2, \hat{e}_1, \hat{e}_2)-40$ Now we want to increase the number of edges by a further $8(N+M+1)-c$ to obtain a total of $40+40(N+M+1)$ edges within the 4 SAWs, independent of the choice of the 4 SAWs. Note that $8(N+M+1)-c$ is a non-negative even integer since $40+c$ is the number of edges in a 2SAP minus 4. If $M>1$ set $\bar{u} = \hat{k}$; otherwise we have $M=1$ and $N>1$ so set $\bar{u} = \hat{j}$. We increase the number of edges by appropriately changing the pattern P^2 to P^+ (see Figure 4.3 (d)). More precisely, we replace the walk, $(8(N+M+1)\hat{i})$, starting at $u_1^* = (b_1+5, N-1, M)$, by the walk, $(\frac{1}{2}(8(N+M+1)-c)[\hat{i}, -\bar{u}, \hat{i}, \bar{u}], c\hat{i})$, which has length $16(N+M+1)-c$. Thus the final result is a 2SAP G with length $n+m+c_T$ where $c_T = 36+40(N+M+1)$ (recall that four edges must be deleted from G_1 and G_2) and with span $b_1+b_2+t_T$ where $t_T = 18+8(N+M)$. (Note that by inserting another P^2 with span $k \geq 1$, one can further increase the span by k and at the same time increase the number of edges by $4k$. Therefore we can generalize this construction by taking any $t_T \geq 18+8(N+M)$ and correspondingly $c_T = 4t_T + 4 + 8(N+M)$.) □

Lemma 4.3.2. *(CAPOFF) holds for 2SAPs in Q^*.*

Proof. We prove that there exists an integer $m_T > 0$ such that the following two cases hold:

Case 1: For any integer $b \geq 0$ and any b-pattern P that occurs at $(0,0,0)$ in some finite size 2SAP $H = (\omega_1, \omega_2) \in Q^*$ with span $s \geq b+1$ (i.e. P occurs at the start of some 2SAP but is not itself a 2SAP), there exists a 2SAP $G \in Q^*$ with span $b+m_T$

which also contains P at $(0,0,0)$ (i.e. P is also at the start of G).

Case 2: Similarly, given any b-pattern P' that occurs at $(s-b,0,0)$ in some finite size 2SAP in Q^* with span $s \geq b+1$ (i.e. P' occurs at the end of some 2SAP but is not itself a 2SAP), there exists a 2SAP $G' \in Q^*$ with span $b+m_T$ which contains P' at $(m_T,0,0)$ (i.e. P' also ends G').

We will only consider the first case and construct a 2SAP $G \in Q^*$ satisfying (CAPOFF) using the pattern P. A symmetry argument will give Case 2. In the right most plane of P ($x = b$), there are an even number, i, of vertices from the polygon ω_1 and an even number, j, of vertices from the polygon ω_2, where the next step in H is a positive step in the x-direction, i.e. where the vertices have degree one in P. In fact any capping of the pattern P to create a 2SAP must contain a total of $(i+j)/2$ mutually avoiding undirected SAWs to the right of $x = b$. Thus the set of all possible ways of capping P will be contained in the set $\cup_{(\sigma_1,\sigma_2)}\mathcal{T}(\sigma_1,\sigma_2)$ defined below.

Consider any subset W consisting of an even number of vertices from V_0 (recall that V_0 is the subgraph of the tube generated by the vertices with x-coordinate $x = 0$). Take any partition σ of W into two element subsets, and any partition of σ into two parts σ_1 and σ_2. Let $\mathcal{T}(\sigma_1,\sigma_2)$ be the set of all patterns P' occurring at the end of a 2SAP, with left-most plane consisting of exactly the vertices in W, with the vertices in σ_1 in one polygon and those in σ_2 in the other, and consisting of $(|\sigma_1|+|\sigma_2|)/2$ mutually avoiding undirected SAWs that join up pairwise the vertices in the left-most plane according to the pair partitions σ_1 and σ_2. For $\mathcal{T}(\sigma_1,\sigma_2) \neq \emptyset$, let $m(\sigma_1,\sigma_2) > 0$ denote the minimum span for any element of $\mathcal{T}(\sigma_1,\sigma_2)$. For $\mathcal{T}(\sigma_1,\sigma_2) = \emptyset$, let $m(\sigma_1,\sigma_2) = 0$. The number of ways to form σ_1,σ_2 is bounded above by a finite function of N and M. Hence $m_T = \max_{\sigma_1,\sigma_2}\{m(\sigma_1,\sigma_2)\} > 0$ exists and depends only on N and M.

Since there exists a 2SAP $G' \in Q^*$ which starts with pattern P, there exists an end pattern P' which caps P off. Clearly P' and the vertices of degree one in the right-most plane of P uniquely define a partition, (σ_1, σ_2), as described above. Now let P_1 be a member of $\mathcal{T}(\sigma_1, \sigma_2)$ with the minimum span. Then $G_1 = P \cup (P_1 + (b, 0, 0))$ is a 2SAP with span $b + m(\sigma_1, \sigma_2)$. G_1 can easily be extended to a 2SAP G with span m_T by concatenating a specific 2SAP with span $(m_T - m(\sigma_1, \sigma_2))$ starting at the plane $x = b + m(\sigma_1, \sigma_2)$. Therefore 2SAPs in Q^* satisfy (CAPOFF). $\square$

4.4 The Connective Constant for 2SAPs

Fix two positive integers N and M and let $p_n(N, M)$ be the number of n-edge SAPs in $T(N, M)$ (up to x-translation). Let $q_n^{(2)}(N, M) = |Q_n^*|$, i.e. the number of n-edge 2SAPs in $T(N, M)$ whose left-most-plane is $x = 0$. Theorem 2.2.2, proved by Soteros and Whittington [48, 49], shows the existence of the connective constant for SAPs in $T(N, M)$, i.e.

$$\kappa_p(N, M) \equiv \lim_{n\to\infty} (2n)^{-1} \log p_{2n}(N, M) < \infty. \tag{4.4.1}$$

In this section we prove, using lemma 4.3.1 (CONCAT), the existence of the connective constant for 2SAPs,

$$\kappa_p^{(2)}(N, M) \equiv \lim_{n\to\infty} (2n)^{-1} \log q_{2n}^{(2)}(N, M), \tag{4.4.2}$$

and show using the pattern theorem for SAPs in $T(N, M)$, Theorem 2.4.2, that

$$\kappa_p^{(2)}(N, M) < \kappa_p(N, M). \tag{4.4.3}$$

Theorem 4.4.1. *The following limit exists*

$$\kappa_p^{(2)}(N, M) \equiv \lim_{n\to\infty} (2n)^{-1} \log q_{2n}^{(2)}(N, M). \tag{4.4.4}$$

Proof. By (CONCAT) (Lemma 4.3.1), one obtains for any even $n > 7$ and $m > 7$

$$q_n^{(2)}(N, M) q_m^{(2)}(N, M) \le q_{(n+m+c_T)}^{(2)}(N, M). \tag{4.4.5}$$

Also, for even $n \geq 8$, $q_n^{(2)}(N,M) \leq \sum_{j=4}^{n-4} p_j(N,M)p_{n-j}(N,M)$. Hence, $\limsup_{n\to\infty}$ $(2n)^{-1}\log q_{2n}^{(2)}(N,M) \leq \kappa_p(N,M) < \infty$. The existence of the limit now follows from the generalized super-multiplicative arguments of Lemma 2.1.3. $\square$

Theorem 4.4.2. *For any allowed (N,M), the following inequality holds.*

$$\kappa_p^{(2)}(N,M) < \kappa_p(N,M). \tag{4.4.6}$$

Proof. Let $A = \{$ all proper SAP b-patterns ($b > 1$) which contain exactly two disjoint walks joining the left-most plane of the pattern to the right-most plane of the pattern $\}$. Note that these SAP patterns cannot appear in any 2SAP. For a given 2SAP we construct a SAP that will not contain any pattern in A either. Then using the pattern theorem for SAPs in tubes, Theorem 2.4.2, we can show that the connective constant for 2SAPs is strictly less than that for SAPs in a tube. Fix the allowed pair of non-negative integers $(N,M) \neq (1,1)$. Let G be an n-edge 2SAP in $T(N,M)$. We start by defining the patterns P^i for $i = 3,4,5,6$ (see Figure 4.4). P^3 is the 1-pattern that contains only the configurations of the two walks: $(\hat{\imath},\hat{\jmath},-\hat{\imath})$ starting at $(0,0,0)$, and $(\hat{\imath},\hat{\jmath},-\hat{\imath})$ starting at $(0,0,1)$. Similarly, P^4 is the 1-pattern that contains only the configurations of the two walks: $(\hat{\imath},\hat{k},-\hat{\imath})$ starting at $(0,0,0)$, and $(\hat{\imath},\hat{k},-\hat{\imath})$ starting at $(0,0,3)$.

The 2-patterns P^5 and P^6 are defined as follows. P^5 is the 2-pattern that contains the two walks: $(2\hat{\imath},\hat{k},-2\hat{\imath})$ starting at $(0,0,0)$, and $(2\hat{\imath},\hat{k},-2\hat{\imath})$ starting at $(0,1,0)$. P^6 is defined to be the 2-pattern that contains the two walks: $(2\hat{\imath},3\hat{k},-2\hat{\imath})$ starting at $(0,0,0)$, and $(\hat{\imath},\hat{k},-\hat{\imath})$ starting at $(0,0,1)$.

Let G_1 (see Figure 4.4) be the 8-edge 2SAP made of the two polygons composed of the edges in the walks: $(\hat{\imath},\hat{\jmath},-\hat{\imath},-\hat{\jmath})$ starting at $(0,0,0)$, and $(\hat{\imath},\hat{\jmath},-\hat{\imath},-\hat{\jmath})$ starting at $(0,0,1)$. Similarly, let G_2 (see Figure 4.4) be the 8-edge 2SAP made of the two polygons composed of the edges in the walks: $(\hat{\imath},\hat{k},-\hat{\imath},-\hat{k})$ starting at $(0,0,0)$, and

$(\hat{\imath}, \hat{k}, -\hat{\imath}, -\hat{k})$ starting at $(0,0,3)$. By Lemma 4.3.1 (CONCAT), we can concatenate any given 2SAP G to the 2SAP G_1 (or G_2 if $N = 0$) so that the resulting 2SAP, $\hat{G}$, ends with the 1-pattern P^3 (or P^4 if $N = 0$). Now, without changing the 2SAP except for the configuration of its right-most plane, we can associate a SAP to this 2SAP by connecting the two polygons of the 2SAP using the edges of two appropriately chosen mutually avoiding SAWs. The SAWs are chosen so that they lie on the right side of the 2SAP's right-most plane. More precisely, we remove the two edges in the right-most plane of $\hat{G}$ and add a translate of the 2-pattern P^5 (or P^6 if $N = 0$).

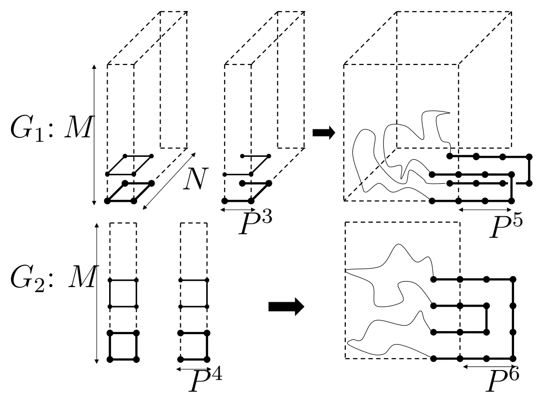


Figure 4.4: Converting a 2SAP to a SAP.

It is clear that the patterns from A cannot occur in a SAP that is constructed from a 2SAP as explained above. For $P \in A$, let $p_n(N, M; \bar{P})$ be the number (up to x-translation) of n-edge SAPs in $T(N, M)$ which do not contain P. Given a positive integer m, let also $p_n(N, M; < m, P)$ denote the number (up to x-translation) of n-edge SAPs in $T(N, M)$ which contain fewer than m copies of P. Hence for any $\epsilon > 0$

$$q_n^{(2)}(N, M) \leq p_{n+c}(N, M; \bar{P}) \leq p_{n+c}(N, M; < \lfloor \epsilon n \rfloor, P), \tag{4.4.7}$$

where (by Lemma 4.3.1) (CONCAT) $c = c_T + 16$ is a constant depending only on $T(N, M)$. (The value of 16 arises as follows: 8 obtained from G_1 is added to 8 from the pattern P^5 or P^6.) Taking logarithms, dividing by n and letting $n \to \infty$ through even n implies

$$\kappa_p^{(2)}(N,M) \le \kappa_p(N,M;<\epsilon,P), \tag{4.4.8}$$

where

$$\kappa_p(N,M;<\epsilon,P) \equiv \limsup_{n\to\infty} n^{-1}\log p_n(N,M;<\lfloor \epsilon n\rfloor,P). \tag{4.4.9}$$

Moreover, by the pattern theorem for SAPs in tubes, Theorem 2.4.2, we know that for $\epsilon = \epsilon_P$

$$\kappa_p(N,M;<\epsilon_P,P) < \kappa_p(N,M). \tag{4.4.10}$$

Therefore, we have $\kappa_p^{(2)}(N,M) \le \kappa_p(N,M;<\epsilon_P,P) < \kappa_p(N,M)$. $\square$

4.5 Pattern Theorem for 2SAPs

In this section, following the terminology of Section 2.5, we obtain a pattern theorem for 2SAPs in an (N,M)-tube.

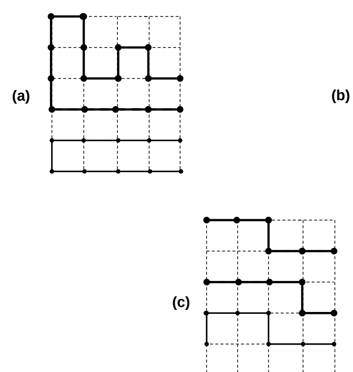


Figure 4.5: (a) A 2SAP start pattern in $T(0,5)$. (b) A 2SAP end pattern in $T(0,5)$. (c) A proper 2SAP pattern in $T(0,5)$.

In the definition of a proper pattern in Section 2.5, Condition (ii) is needed to exclude patterns which can only occur at the left-most or the right-most plane of a 2SAP and nowhere else. However, due to Lemma 4.3.1 (CONCAT) any pattern satisfying condition (ii) automatically satisfies condition (i). This is because once a pattern P occurs in the interior of a 2SAP G (i.e. not at the start or end of the 2SAP), by Lemma 4.3.1 (CONCAT), we can concatenate together any number of G's and construct an infinite number of 2SAPs containing the pattern P. Thus condition (ii) is sufficient for the definition of proper 2SAP patterns. (See Figure 4.5 for examples of start, end and proper 2SAP patterns.)

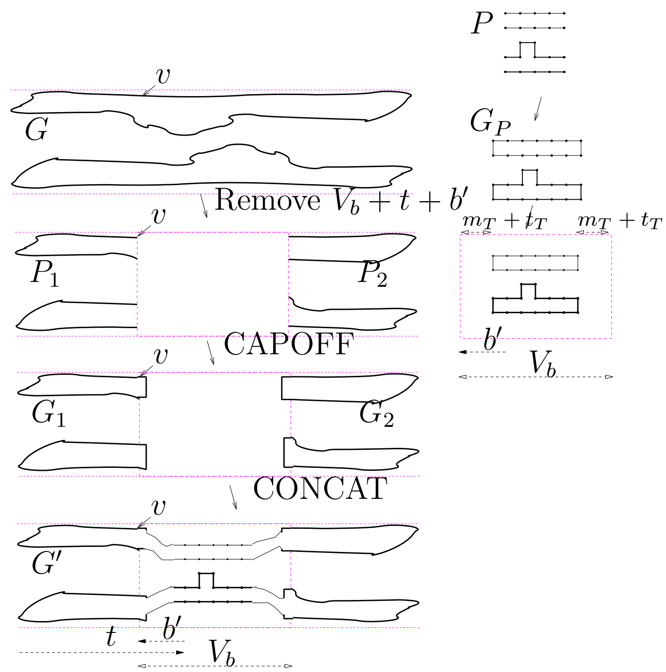


Figure 4.6: For 2SAP G (top left) and proper pattern P (top right), the pattern insertion algorithm (top-to-bottom on left) for Proposition 2.5.1 is shown: 2SAP G has edges removed (those in $V_b + t + b'$; V_b and b' as defined on right) to result in start pattern P_1 and end pattern P_2. P_1 and P_2 are capped off using Lemma 4.3.2 (CAPOFF) to create 2SAPs G_1 and G_2. Then G_1, G_P (2SAP shown below P on right), and G_2 are concatenated using Lemma 4.3.1 (CONCAT) to create 2SAP G' in which P occurs at t.

Define $q_n^{(2)}(N, M; < m, P)$ to be the number of 2SAPs in Q^* which contain less than m translates of P. The following result is the required pattern theorem and it is an immediate consequence of Corollary 2.5.3 and the fact that 2SAPs satisfy (CONCAT) and (CAPOFF).

Theorem 4.5.1. *Let P be any proper pattern for 2SAPs in $T(N, M)$. Then there exists an $\epsilon_P > 0$ such that*

$$\kappa_p^{(2)}(N, M; < \epsilon_P, P) < \kappa_p^{(2)}(N, M), \tag{4.5.1}$$

where

$$\kappa_p^{(2)}(N,M;<\epsilon_P,P) :\equiv \limsup_{n\to\infty} n^{-1}\log q_n^{(2)}(N,M;<\lfloor \epsilon_P n\rfloor,P). \tag{4.5.2}$$

Therefore, the following Corollary holds.

Corollary 4.5.2. *Let P be any proper pattern for 2SAPs in $T(N,M)$. Then there exists an $\epsilon_P > 0$ such that the probability that an n-edge 2SAP contains at least $\lfloor \epsilon_P n\rfloor$ copies of the pattern P approaches one exponentially fast as n goes to infinity.*

Proof. Given a proper 2SAP pattern P and $\epsilon_P > 0$ as prescribed by Theorem 4.5.1, let

$$g(n) = \big(\log q_n^{(2)}(N,M;<\lfloor\epsilon_P n\rfloor,P) - \kappa_p^{(2)}(N,M;<\epsilon_P,P)n\big) - \big(q_n^{(2)}(N,M) - \kappa_p^{(2)}(N,M)n\big). \tag{4.5.3}$$

By Theorem 4.4.1,

$$\lim_{n\to\infty}\frac{g(n)}{n} = 0 \tag{4.5.4}$$

hence $g(n) = o(n)$. So

$$\log\frac{q_n^{(2)}(N,M;<\lfloor\epsilon_P n\rfloor,P)}{q_n^{(2)}(N,M)} = \big(\kappa_p^{(2)}(N,M;<\epsilon_P,P) - \kappa_p^{(2)}(N,M)\big)n + o(n) \tag{4.5.5}$$

which leads to

$$\frac{q_n^{(2)}(N,M;<\lfloor\epsilon_P n\rfloor,P)}{q_n^{(2)}(N,M)} = e^{-\gamma_P n + o(n)}, \tag{4.5.6}$$

where $\gamma_P \equiv \kappa_p^{(2)}(N,M) - \kappa_p^{(2)}(N,M;<\epsilon_P,P)$.

Therefore, the probability that an n-edge 2SAP contains at least $\lfloor\epsilon_P n\rfloor$ copies of P is given by

$$\begin{aligned}\frac{q_n^{(2)}(N,M;\geq\lfloor\epsilon_P n\rfloor,P)}{q_n^{(2)}(N,M)} &= \frac{q_n^{(2)}(N,M) - q_n^{(2)}(N,M;<\lfloor\epsilon_P n\rfloor,P)}{q_n^{(2)}(N,M)}\\ &= 1 - \frac{q_n^{(2)}(N,M;<\lfloor\epsilon_P n\rfloor,P)}{q_n^{(2)}(N,M)}\\ &= 1 - e^{-\gamma_P n + o(n)}. \end{aligned} \tag{4.5.7}$$

By Theorem 4.5.1, $\gamma_P > 0$ hence the above probability goes to unity as $1 - e^{-\gamma_P n + o(n)}$ when $n \to \infty$. $\square$

4.6 Topological Linking of 2SAPs

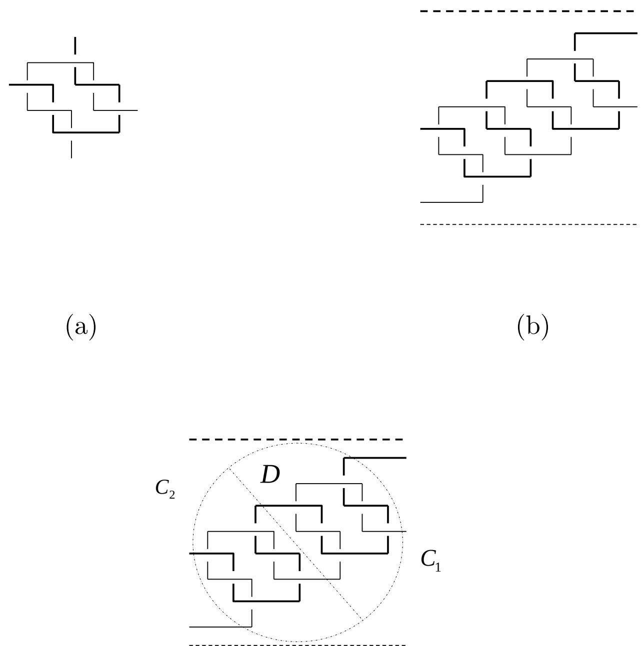
In this section we show that all but exponentially few sufficiently large 2SAPs are topologically linked by first defining the pattern for 2SAPs introduced by C. Ernst that guarantees topological linking.

Let (A, t) be the tangle as shown in Figure 4.7 (a) where the 3-ball A is not shown and only the set of arcs t is visible. Let $A + A = (B, s)$ be the sum of two copies of A, see the solid lines of Figure 4.7 (b). Note that the tangle (A, t) is prime [30] and therefore also inseparable.

The following theorem was proved by C. Ernst using Theorem 3.1.8.

Theorem 4.6.1 (Ernst 2008 [4])**.** *Let L be a 2-component link. Assume that there exists a 2-sphere S intersecting each component of L in 2 points such that S bounds the tangle $A + A$ on one side (see Figure 4.7 (c)), then L is a non-splittable link.*

Proof. Let B be a ball bounded by S and let (B, s) be the tangle $A + A$. Let D be a disk in the tangle (B, s) that cuts the tangle (B, s) into its two summands (C_1, t_1) and (C_2, t_2) (see Figure 4.7 (c)) each of which is equivalent to the tangle (A, t) shown in Figure 4.7 (a). Let S' be a 2-sphere bounding the 3-ball C_1. The goal is to show that S' divides (S^3, L) into two 2-string tangles each of which is inseparable. One of these tangles (C_1, t_1) is prime and therefore inseparable. The other is also inseparable by Lemma 3.1.8 (i). This can be seen as follows: The tangle $(cl(S^3 - C_1), cl(S^3 - C_1) \cap L)$ can be subdivided by the disk $D' = cl(S \cap (S^3 - C_1))$ into two tangles. One of these is (C_2, t_2) which is inseparable. The other is a two string tangle that if it contains a disk that separates its two arcs then this disk must intersect D'. Therefore by Lemma 3.1.8 (i) the tangle $(cl(S^3 - C_1), cl(S^3 - C_1) \cap L)$ is also inseparable. By Lemma 3.1.8 (ii) we conclude that L is non-splittable. $\square$



(a) (b)

(c)

Figure 4.7: a) The tangle (A, t). b) Projection into the (x, y)-plane of a tight 2SAP pattern of tangle $A + A$ on the cubic lattice. The thick dashed lines and thick solid lines are in one SAP and the thin dashed lines and thin solid lines are in the other SAP. c) The angled dotted-dashed line is the projection of the disk D, introduced in the proof of Theorem 4.6.1, that cuts tangle (B, s) into its two summands (C_1, t_1) and (C_2, t_2) each of which is equivalent to (A, t) in (a).

The solid lines in Figure 4.7 (b) show a tight lattice pattern, or rather its projection in the (x, y)-plane, which represents the tangle $A + A$. Once such a tangle has been created within a link L on the cubic lattice then the link can't be undone by the rest of the lattice link. It is easy to check that this pattern occurs in a (7,1)-lattice tube (i.e. $T(7, 1)$) involving both strings. If one adds two straight segments parallel to the x-axis (shown as dashed lines in Figure 4.7 (b)), then this constitutes the (x, y)-projection of a proper 10-pattern that can occur in a 2SAP in a (7,1)-lattice tube. It is easy to see that a topologically equivalent pattern can also occur in a $(2, 2)$, and $(3, 1)$-lattice tube (by increasing the b-span of the pattern).

Therefore we obtain the following corollary to Theorem 4.6.1 and Corollary 4.5.2.

Corollary 4.6.2. *For allowed (N, M) where $M + N \geq 4$ the probability that an n-edge 2SAP in $T(N, M)$ is a non-splittable link approaches one exponentially fast as n goes to infinity through even values of n.*

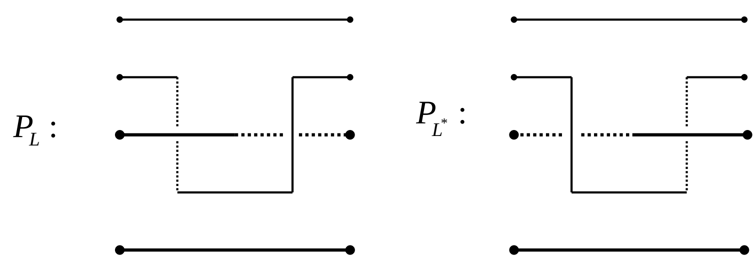
4.7 Homological Linking of 2SAPs

In this section, we find a lower bound (with probability one) for the linking number of 2SAPs. The arguments needed to show the high probability of homological linking are more complicated than those used for topological linking because there is no local pattern P that can guarantee a non-zero linking number. Any linking number contribution generated locally can always be undone in the 2SAP elsewhere. Nevertheless the probability that a 2SAP has a small linking number goes to zero as the length of the 2SAP increases. In fact we prove that the linking number asymptotically grows (with probability one) no slower than $\sqrt{n}$; i.e. for any n-edge 2SAP G_n and any function $f(n) = o(\sqrt{n})$

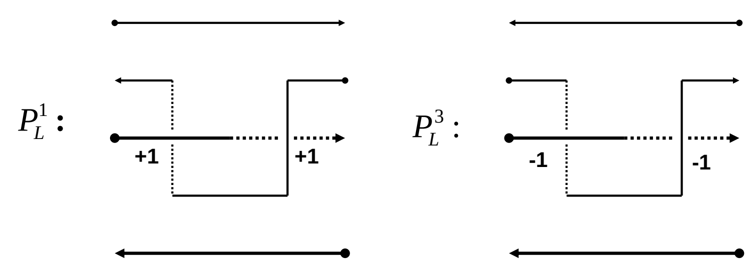
$$\lim_{n\to\infty} \mathbb{P}\big(|Lk(G_n)| \geq f(n)\big) = 1. \tag{4.7.1}$$

As a result, we show that the probability of a non zero linking number of an n-edge 2SAP goes to one as $n \to \infty$.

Let $q_n^{(2)}(N, M; Lk \leq m)$ denote the number of 2SAPs $G = (\omega_1, \omega_2)$ (up to x-translation) with linking number, $Lk(\omega_1, \omega_2)$, at most m. We will use two patterns P_L and P_{L^*} (with projections that are mirror images of each other) whose occurrence in any 2SAP increases or decreases the linking number by one, depending on the order induced on the pattern's edges by the 2SAP's orientation. Figures 4.8 (a) and (b) show diagrams of the projections of P_L and P_{L^*} into the (x, z)-plane. In particular, P_L and P_{L^*} are chosen so that if one of them occurs in a 2SAP G, then the corresponding sequence of edges in D_G^y (as defined in Lemma 3.1.4) creates exactly two crossings. Note that in the patterns P_L and P_{L^*}, the edges shown with the same thickness are in the same polygon. However, the thicker edges do not necessarily belong to ω_1, the first polygon of the 2SAP. In these diagrams a solid line of either thickness is in the plane $z = 1$ while a dashed line of either thickness between two 2SAP pattern vertices, u and v say, represents a sequence of edges in the pattern with u and v in the plane $z = 1$ and the intermediate vertices in the plane $z = 0$.



(a) The pattern P_L. (b) The pattern P_{L^*}.



(c) The occurrence type P_L^1. (d) The occurrence type P_L^3.

Figure 4.8: Projections of proper 2SAP patterns P_L, P_L^* are shown in (a) and (b) respectively. Each pattern is a 4-pattern in $T(5, 1)$. Lines shown with the same thickness correspond to edges in the same polygon. Dashed lines represent a sequence of edges in the pattern with the start and end vertices in the plane $z = 1$ and the intermediate vertices in the plane $z = 0$. Solid lines are in the plane $z = 1$. Orientations of P_L representing two of its occurrence types, P_L^1 and P_L^3, are shown in (c) and (d) respectively.

Let G be a 2SAP in which the pattern P_L (P_{L^*}) occurs. Depending on the orientations of P_L's (P_{L^*}'s) edges, as induced by an orientation of G, P_L (P_{L^*}) has four different occurrence types P_L^1 ($P_{L^*}^1$), P_L^2 ($P_{L^*}^2$), P_L^3 ($P_{L^*}^3$) and P_L^4 ($P_{L^*}^4$). Figures 4.8 (c) and (d) show P_L^1 and P_L^3. The projections of $P_{L^*}^1$ and $P_{L^*}^3$ are the mirror images of the projections of P_L^1 and P_L^3 respectively. P_L^2 (P_L^4, $P_{L^*}^2$, $P_{L^*}^4$) is obtained from

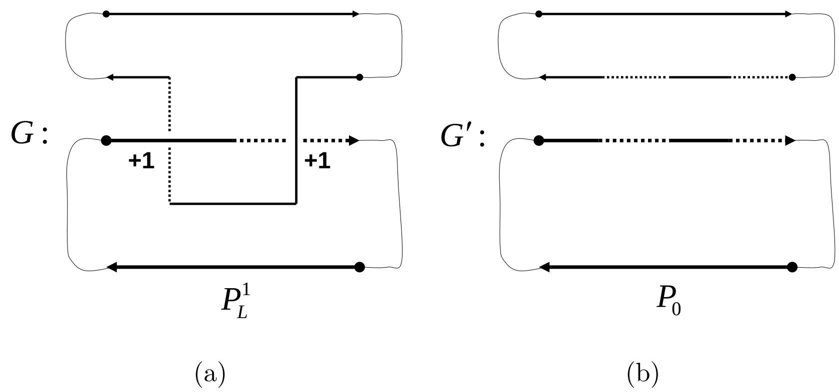


Figure 4.9: (a) An example of P_L^1 occurring in a 2SAP G. (b) G' formed by replacing the 24-edge pattern P_L^1 in G by the 24-edge 4-pattern P_0; in this case, $Lk(G) = Lk(G') + 1$. Dashed lines represent a sequence of edges in the pattern with the start and end vertices in the plane $z = 1$ and the intermediate vertices in the plane $z = 0$.

P_L^1 (P_L^3, $P_{L^*}^1$, $P_{L^*}^3$) by reversing the directions on all its edges. Note that wherever P_L (P_{L^*}) occurs in G, it can be replaced by the pattern P_0 as shown in Figure 4.9 (b) (with one of its orientations) which has the same span and the same number of edges as P_L (P_{L^*}). These definitions and Lemma 3.1.4 lead directly to the following lemma.

Lemma 4.7.1. *Let G be a 2SAP in which the pattern P_L (P_{L^*}) occurs at a specific location and let G' be the 2SAP constructed from G by replacing the pattern P_L (P_{L^*}) by the pattern P_0 as shown in Figure 4.9 (b). If P_L (P_{L^*}) is of type P_L^1 or P_L^2 ($P_{L^*}^3$ or $P_{L^*}^4$) then the following equation holds*

$$Lk(G) = Lk(G') + 1. \tag{4.7.2}$$

If P_L (P_{L^*}) is of type P_L^3 or P_L^4 ($P_{L^*}^1$ or $P_{L^*}^2$) then the following equation holds

$$Lk(G) = Lk(G') - 1. \tag{4.7.3}$$

We need one additional lemma to prove the main theorem of this section.

Lemma 4.7.2. *For $i = 1, 2, 3, 4$, both P_L^i and $P_{L^*}^i$ occur in 2SAPs with the same probability distribution.*

Proof. For $i = 1, 2, 3, 4$, it can be seen from Figure 4.8 that if P_L^i occurs in any n-edge 2SAP at a specific location, then we can replace P_L^i by $P_{L^*}^i$ without changing the number of edges and without changing the orientation of the 2SAP. Similarly, any occurrence of $P_{L^*}^i$ can be replaced by P_L^i. Therefore, P_L^i and $P_{L^*}^i$ appear in 2SAPs with the same probability, i.e. the number of n-edge 2SAPs in which the pattern P_L^i occurs at a given location $(j, 0, 0)$ is exactly the same as the number of n-edge 2SAPs in which the pattern $P_{L^*}^i$ occurs at location $(j, 0, 0)$. □

Theorem 4.7.3. *Let (N, M) be an allowed pair with $M + N \geq 4$. For every function $f(n) = o(\sqrt{n})$, the probability that the absolute value of the linking number of an n-edge 2SAP in $T(N, M)$ is less than $f(n)$ goes to zero as n goes to infinity.*

Proof. The proof is based on the pattern theorem for 2SAPs, Theorem 4.5.1, and also a coin tossing argument as in [10]. Note that the proof here uses the patterns P_L (P_{L^*}) and hence requires $N \geq 5$ and $M \geq 1$; modification of these patterns to fit in other tube sizes is possible and then the arguments below give the proof for any allowed (N, M) with $N + M \geq 4$. Applying Theorem 4.5.1 for the pattern P_L (P_{L^*}), we have a positive number ϵ_L (ϵ_{L^*}) and an integer $N_L > 0$ ($N_{L^*} > 0$) such that for even $n \geq N_L$ ($n \geq N_{L^*}$), all except exponentially few n-edge 2SAPs contain at least $\lfloor \epsilon_L n \rfloor$ ($\lfloor \epsilon_{L^*} n \rfloor$) pairwise disjoint translates of the pattern P_L (P_{L^*}). This implies that for $\epsilon = \min\{\epsilon_L, \epsilon_{L^*}\}$ and even $n \geq \max\{N_L, N_{L^*}\}$, all except exponentially few n-edge 2SAPs contain at least $\lfloor \epsilon n \rfloor$ pairwise disjoint copies of $P \in \{P_L, P_{L^*}\}$. Furthermore, by the proof of corollary 4.5.2, there exist $\gamma_L > 0$ and $\gamma_{L^*} > 0$ such that for $\gamma = \min\{\gamma_L, \gamma_{L^*}\}$

$$\frac{q_n^{(2)}(N, M; \geq \lfloor \epsilon n \rfloor, P)}{q_n^{(2)}(N, M)} \geq 1 - e^{-\gamma n + o(n)}, \text{ as } n \to \infty, n \text{ even}, \tag{4.7.4}$$

where $q_n^{(2)}(N, M; \geq \lfloor \epsilon n \rfloor, P)$ is the number of n-edge 2SAPs which contain at least $\lfloor \epsilon n \rfloor$ copies of $P \in \{P_L, P_{L^*}\}$.

Let G be a randomly chosen 2SAP and let l be the total number of copies of $P \in \{P_L, P_{L^*}\}$ that occur in G. By Lemma 4.7.1, the two patterns P_L and P_{L^*} occur independently in 8 different ways. Let $A = \{P_L^1, P_L^2, P_{L^*}^3, P_{L^*}^4\}$ and $B = \{P_{L^*}^1, P_{L^*}^2, P_L^3, P_L^4\}$. Hence, by Lemma 4.7.1, every pattern from A will add one to the linking number and every pattern from B reduces the linking number by one. We order the occurrences of $P \in \{P_L, P_{L^*}\}$ in G with the integers $i = 1, ..., l$ starting from the left-most plane of G. Now by Lemma 4.7.2, the probability that a chosen occurrence of P (amongst the l occurrences) in G belongs to A is the same as the probability that it belongs to B. So

$$\begin{aligned} \mathbb{P}\big(P \in A \mid P \;\; occurs\ at\ the\ location\ (j,0,0)\ in\ a\ 2SAP\big) &= \\ \mathbb{P}\big(P \in B \mid P \;\; occurs\ at\ the\ location\ (j,0,0)\ in\ a\ 2SAP\big) &= \frac{1}{2}. \end{aligned}$$

So their distribution is analogous to tossing a fair coin; i.e.

$$\begin{aligned} \mathbb{P}\,(P \in A\ exactly\ k\ times\ in\ first \quad t \quad & occurrences\ of\ P|\ P\ \ occurs\ \ \geq t\ times) \\ &= \begin{pmatrix} t \\ k \end{pmatrix} \big(\frac{1}{2}\big)^k \big(\frac{1}{2}\big)^{t-k} = \begin{pmatrix} t \\ k \end{pmatrix} \big(\frac{1}{2}\big)^t \quad (4.7.5) \\ &\leq \frac{2}{\sqrt{t}}, \quad (4.7.6) \end{aligned}$$

where the final inequality holds only for n sufficiently large, i.e. $n \geq N'$ for some $N' > 0$. This follows from Stirling's formula and the fact that the probability in equation (4.7.5) reaches its maximum at $k = \lfloor t/2 \rfloor$.

Let $Q^*_{n,\geq}(\epsilon, P) \subset Q^*_n$ denote the set of 2SAPs with at least $\lfloor \epsilon n \rfloor$ occurrences of either P_L or P_L*. For any 2SAP $G \in Q^*_{n,\geq}(\epsilon, P)$, let $\vec{P} =(P_i,\ i = 1, 2, \cdots, \lfloor \epsilon n \rfloor)$ denote the first $\lfloor \epsilon n \rfloor$ occurrences of P_L or P_{L^*} in G starting from G's left-most plane. By Lemma 4.7.1, the linking number of G can be written as the sum of two terms where one term only depends on $\vec{P}$, i.e.

$$
\begin{aligned}
Lk(G) &= Lk(G') + \sum_{i=1}^{\lfloor \epsilon n \rfloor} (I_A(P_i) - I_B(P_i)) && (4.7.7)\\
&= Lk(G') + \sum_{i=1}^{\lfloor \epsilon n \rfloor} (2I_A(P_i) - 1) && (4.7.8)\\
&= Lk(G') - \lfloor \epsilon n \rfloor + 2\sum_{i=1}^{\lfloor \epsilon n \rfloor} I_A(P_i), && (4.7.9)
\end{aligned}
$$

where G' is the 2SAP obtained by replacing each P_i by P_0 (as in Figure 4.9) and $I_A(P) = 1$ ($I_B(P) = 1$) if and only if $P \in A$ ($P \in B$) and otherwise $I_A(P) = 0$ ($I_B(P) = 0$). Note that G is completely determined by $(G', P_i, i = 1, ..., \lfloor \epsilon n \rfloor)$, and we express this by writing $G = (G', \vec{P})$.

We now have that $|Lk(G)| < f(n)$ if and only if $\sum_{i=1}^{\lfloor \epsilon n \rfloor} I_A(P_i)$ is an integer in the interval $I = (\frac{-f(n) - Lk(G') + \lfloor \epsilon n \rfloor}{2}, \frac{f(n) - Lk(G') + \lfloor \epsilon n \rfloor}{2})$. Note that, regardless of the value of $Lk(G')$, there are at most $\lfloor f(n) \rfloor + 1$ such integers. Moreover given that $G = (G', \vec{P}) \in Q^*_{n,\geq}(\epsilon, P)$, the probability that $|Lk(G)| < f(n)$ is equal to the probability that $\sum_{i=1}^{\lfloor \epsilon n \rfloor} I_A(P_i)$ is an integer in the interval I. By equation (4.7.6), for $n \geq N'$ and regardless of the choice of G', the probability that $\sum_{i=1}^{\lfloor \epsilon n \rfloor} I_A(P_i)$ takes on a particular integer value in I is bounded above by $\frac{2}{\sqrt{\lfloor \epsilon n \rfloor}}$. Hence we obtain, for $n \geq N'$,

$$
\mathbb{P}\big(|Lk(G)| < f(n) | G \in Q^*_{n,\geq}(\epsilon, P)\big) \leq \frac{2\lfloor f(n) \rfloor + 2}{\sqrt{\lfloor \epsilon n \rfloor}}. \tag{4.7.10}
$$

Therefore, for $n \geq N = \max\{N', N_L, N_{L^*}\}$ and G a random n-edge 2SAP

$$
\begin{aligned}
\mathbb{P}\big(|Lk(G)| < f(n)\big) \leq& \\
&\mathbb{P}\big(|Lk(G)| < f(n), G \in Q^*_{n,\geq}(\epsilon, P)\big) + \mathbb{P}\big(G \notin Q^*_{n,\geq}(\epsilon, P)\big)\\
\leq& \frac{2\lfloor f(n) \rfloor + 2}{\sqrt{\lfloor \epsilon n \rfloor}} \left(\frac{q_n^{(2)}(N, M; \geq \lfloor \epsilon n \rfloor, P)}{q_n^{(2)}(N, M)} \right) + \left(1 - \frac{q_n^{(2)}(N, M; \geq \lfloor \epsilon n \rfloor, P)}{q_n^{(2)}(N, M)} \right).
\end{aligned}
\tag{4.7.11}
$$

This along with equation (4.7.4) proves that $\mathbb{P}\big(|Lk(G)| < f(n)\big)$ goes to zero as n goes to infinity. $\square$

Corollary 4.7.4. *The probability that an n-edge 2SAP G is homologically linked approaches one as n goes to infinity.*

Proof. Since $|Lk(G)| \geq f(n)$ implies that G is homologically linked, Theorem 6.7.1 implies that for any $f(n) = o(\sqrt{n})$

$$\lim_{n\to\infty} \mathbb{P}\big(|Lk(G)| \geq f(n)\big) \leq \lim_{n\to\infty} \mathbb{P}\big(G \, is \ \text{homologically linked}\big) = 1. \tag{4.7.12}$$

□

4.8 An Upper Bound for The Linking Number of 2SAPs

In this section we find an upper bound for the linking number of a 2SAP.

Theorem 4.8.1. *Let G in $T(N, M)$ be a disjoint union of SAPs with a total length of n-edges. (G could be a knot or a link with any number of components.) For fixed N and M, the crossing number of the knot or link formed by G is bounded above by a linear function of n. Therefore the absolute value of the linking number between any two components is bounded above by a linear function as well.*

Proof. By Lemma 3.1.4 there is a polygonal link L' with regular projection D_G^z in the (x, y)-plane.

We will provide an upper bound for the crossing number of D_G^z by determining an upper bound on how many crossings each edge l_i of L' can generate. To do so we need to consider two cases depending on whether the edge l_i in L' corresponds to a horizontal or a vertical edge l_i^* in G. If we assume that l_i^* is horizontal with vertices (x_{i_1}, y_{i_1}, z_i) and (x_{i_2}, y_{i_2}, z_i) then, by Lemma 3.1.4, the only other horizontal edges of G that could generate edges in L' that subsequently generate crossings with l_i in

D_G^z, are the edges that have at least one end point incident on either (x_{i_1}, y_{i_1}, z) or (x_{i_2}, y_{i_2}, z) for $0 \leq z \leq M$. There are at most $7(M+1)$ of these. By Lemma 3.1.4 the only vertical edges of G that can generate crossings with l_i in D_G^z are the edges that have both endpoints with (x, y)-coordinates (x_{i_1}, y_{i_1}) or (x_{i_2}, y_{i_2}). There are $2M$ of these. Combining these we get that a horizontal edge can only generate less than $9(M+1)$ crossings.

Now we assume that l_i^* is vertical with vertices (x_i, y_i, z_i) and (x_i, y_i, z_i+1). Similar to the argument before we can argue that there are at most $4(M+1)$ horizontal edges (those with a vertex with coordinates (x_i, y_i, z) for some $0 \leq z \leq M$) and M vertical edges that can generate crossings with l_i. Combining the two cases gives us the upper bound $9(M+1)n/2$ on the crossing number of the knot or link type G. (It is $9(M+1)n/2$ since each crossing will be counted twice, i.e. if an edge l_i crosses an edge l_j then we counted the crossing both in the upper bound for l_i and in the upper bound for l_j.)

By Lemma 3.1.4 there is also a polygonal link L with regular projection D_G^y in the (x, z)-plane. A similar argument gives an upper bound $9(N+1)n/2$ on the crossing number of the knot or link type G. Combining the two bounds results in an upper bound on the crossing number of $9(\min(M, N)+1)n/2$. Therefore the absolute value of the linking number between any two components of G is bounded above by $9(\min(M, N)+1)n/4$. $\square$

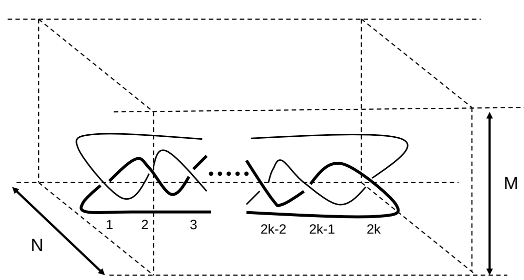


Figure 4.10: An example of a 2SAP with linking number of order n.

Remark 4.8.1. The following example shows that there exists a 2SAP with linking number linear in its length. Let G be a 2SAP with the link type of a $(2k, 2)$-torus link (see Figure 4.10). $Lk(G) = k$ and it is easy to see that such a link can be built with a total length not exceeding $20k + 6$ in a $(3, 1)$-tube. We also note that the linear growth rate of the number of crossings in terms of the overall length is an artifact of the tube constraint. No linear upper bound exists without the tube constraint and in fact the number of crossings can grow as fast as $O(n^{4/3})$ for a knot or link with length n. This was shown in [11] for knots and in [12] for links.

4.9 Summary

In this chapter, the following question was addressed regarding the linking probability of two self-avoiding polygons: *under what conditions does the "linking probability" of pairs of self-avoiding polygons go to one?* The answer can depend on how one defines linking probability. In order to approach this question, we introduced the 2SAP model. We showed that (CONCAT) and (CAPOFF) are satisfied for this model. We also established the existence of the connective constant for 2SAPs and showed that it is strictly less than that of SAPs. We proved a pattern theorem for 2SAPs and used it to investigate homological as well as topological linking of 2SAPs. We showed that the homological linking probability goes to one at least as fast as $1 - O(n^{-1/2})$ and that the topological linking probability goes to one exponentially rapidly, as $n \to \infty$. Furthermore we proved that the linking number grows (with probability one) faster than any function that is $o(\sqrt{n})$; i.e. for any function $f(n) = o(\sqrt{n})$, there exists $A \geq 0$ such that as $n \to \infty$ the probability that $|Lk(\omega_1, \omega_2)| \geq f(n)$, with (ω_1, ω_2) the component polygons of an n-edge 2SAP, satisfies

$$\mathbb{P}\big(|Lk(\omega_1, \omega_2)| \geq f(n)\big) = 1 - \frac{A}{\sqrt{n}} + o\big(\frac{1}{\sqrt{n}}\big). \tag{4.9.1}$$

Hence

$$\lim_{n\to\infty} \mathbb{P}\big(|Lk(\omega_1, \omega_2)| \geq f(n)\big) = 1. \tag{4.9.2}$$

We also showed that the linking number cannot grow faster than linearly in n because of the tube constraint; i.e. there exist constants a and b such that for any n-edge 2SAP

$$|Lk(\omega_1, \omega_2)| \leq an + b. \tag{4.9.3}$$

We gave a simple example to show that the upper bound in equation 4.9.3 for 2SAPs can be realized. Hence sufficient conditions for ensuring that the linking probability goes to one are established. Note that based on the results of this chapter, in collaboration with C.E. Soteros (my supervisor), S.G. Whittington and C. Ernst, a paper has been submitted to the Journal of Knot Theory and Its Ramifications.

Chapter 5

Stretched Polygons and Loops in a Tube

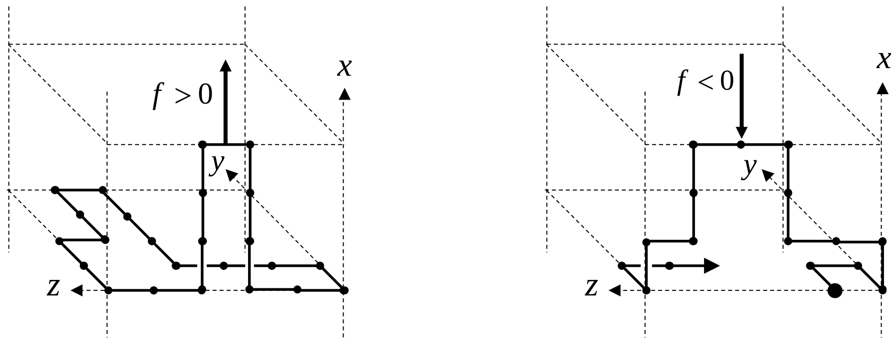
5.1 Introduction

A polymer's topological entanglement may be affected by being subject to some external forces. Knotting and linking probabilities have been frequently used to measure topological entanglements of polymers [27]. In Janse van Rensburg *et al* 2008 [27], a ring polymer confined between two parallel walls (planes) and pulled by an external force is modelled by a lattice polygon subjected to an applied force f along the z-direction of the lattice and perpendicular to the walls. This model explains the situation where a ring polymer, such as circular DNA, is subject to a force in the presence of a topoisomerase which mediates strand passages and may change the knot type of the polymer.

In [27], a pattern theorem is proved for lattice polygons in the presence of a large applied force $f > 0$. This theorem is then used to examine the incidence of entanglements such as knotting when the polygons are influenced by a large force. Here, we add the tube confinement constraint to this model and address similar questions for polygons confined to a tube. Unlike the situation for polygons in $\mathbb{Z}^3$, in a tube, polygons under the influence of the external force f do not have much freedom and must stay inside the tube. In particular, the tube constraint will allow us to prove the pattern theorem for any arbitrary value of f, not just for large values of f. Furthermore, in addition to investigating the knotting probability, we associate a two-component link to each polygon in a tube and examine the incidence of topological linking when the polygon is under the influence of force f.

We will establish the existence of the limiting free energy for stretched polygons and analyze the asymptotic form of the partition function for stretched polygons using the transfer-matrix approach. We show that the average span per edge of a randomly chosen n-edge stretched polygon approaches a positive constant as $n \to \infty$ and is non-decreasing in f almost everywhere. We also prove that the average number of occurrences of a given proper SAP config (per edge) in any n-edge stretched polygon approaches a positive constant as $n \to \infty$.

Given $m \geq 0$, here we will investigate polygons of span m which are confined to a tube. Such polygons are bounded by the tube walls ($y = 0$, $y = N$, $z = 0$ and $z = M$) as well as the two planes $x = x_1$ and $x = x_1 + m$, where $x = x_1$ is the left-most plane of the polygon. We assume that a force f parallel to the x-axis, perpendicular and incident to the plane $x = x_1 + m$ is applied to a single polymer modelled by a self-avoiding polygon. Figure 5.1 (a) illustrates an example of such a scenario, where we have rotated the tube 90 degrees for comparison with the model considered in [27]. The lattice model of polymers confined to a tube and subject to a force f is introduced in Section 5.2. The existence of the limiting free energy is also discussed in this section. In Section 5.3, we use the method presented in Section 2.5, the pattern insertion approach, to prove a pattern theorem for polygons in the presence of a given force f. In Section 5.4, the asymptotic behaviour of such polygons is analyzed, using the transfer-matrix approach, and some results are established regarding the average span of the randomly chosen n-edge stretched polygons. In Section 5.5, we use the results obtained for polygons to prove a pattern theorem for loops (an undirected self-avoiding walk which has both endpoints on its left-most plane; see Figure 5.1 (b) for an example of a loop). Then, in Section 5.6, the pattern theorems are used to investigate topological entanglements through the knotting probability for a polygon. Moreover, associating a two-component link to any polygon (loop), we also use the topological linking probability to measure a polygon's (loop's) topological entanglement complexity. Finally, in Section 5.7, we discuss the main conclusions of this chapter.



(a) A 24-edge polygon with span 3 in $T(4,5)$ which is under the influence of force $f > 0$; note that the tube is rotated 90 degrees unclockwise.

(b) A 17-edge loop with span 3 in $T(4,5)$ which is under the influence of force $f < 0$.

Figure 5.1: Examples of polygons and loops confined to a tube and subject to a force f.

5.2 Stretched Polygons in Tubes

The partition function of the *stressed ensemble* model is given by

$$Z_n(N, M; f) = \sum_{m=0}^{n/2-1} p_n(N, M; m) e^{fm}, \tag{5.2.1}$$

where $p_n(N, M; m)$ denotes the number of n-edge SAPs with span m in $T(N, M)$ (up to x-translation). If $f > 0$ then the force is a pulling force, tending to stretch the polygon in the x-direction and the polygons influenced by this force are called *stretched polygons* [27]. On the other hand, if $f < 0$, then the force tends to push the planes $x = x_1$ and $x = x_1 + m$ together. Here, for our convenience, regardless of the sign of f we call the polygons under the influence of f stretched polygons.

Let S_n^f be a random variable taking its values from the set of n-edge stretched polygons with left-most plane $x = 0$, and with the probability distribution

$$\mathbb{P}(S_n^f = G) = \frac{e^{fm}}{Z_n(N, M; f)}, \tag{5.2.2}$$

where $m = m(G)$, the span of the stretched polygon G. One goal of this chapter is to investigate the behaviour of the expected value of $m(S_n^f)$ as a function of f.

Let

$$F_n(N,M;f) \equiv \frac{1}{n} \log Z_n(N,M;f). \tag{5.2.3}$$

We will start analyzing the model by proving the existence of the limiting free energy for polygons confined to a tube and in the stressed ensemble. Note that the proof is a straightforward modification of the proof given in [27, Theorem 2.1] however using the tube concatenation for SAPs.

Theorem 5.2.1. *Given $f \in \mathbb{R}$, the limiting free energy of stretched polygons in the tube $T(N,M)$, defined by*

$$\mathcal{F}(N,M;f) \equiv \lim_{n\to\infty} n^{-1} \log Z_n(N,M;f), \tag{5.2.4}$$

exists where the limit is taken through even values of n.

In addition, there exist non-negative values c_T and t_T such that $Z_n(N,M;f) \le e^{(n+c_T)\mathcal{F}(N,M;f)+ft_T}$ and $\mathcal{F}(N,M;f)$ is convex in f and thus continuous. Moreover, its right- and left-derivatives exist everywhere, and they are non-decreasing functions of f. $\mathcal{F}(N,M;f)$ is also differentiable almost everywhere, and whenever $d\mathcal{F}(N,M;f)/df$ exists it is given by $\lim_{n\to\infty}(dF_n(N,M;f)/df)$. Finally, $f/2 \le \mathcal{F}(N,M;f) \le \log \mu_p(N,M) + f/2$ if $f \ge 0$ while $0 \le \mathcal{F}(N,M;f) \le \log \mu_p(N,M)$ if $f < 0$ ($\mu_p(N,M)$ is the growth constant for polygons confined to $T(N,M)$).

Proof. As discussed in [49, Section 4], there exist non-negative values c_T and t_T such that concatenating an n_1-edge polygon with span m_1 to an n_2-edge polygon with span $m - m_1$ results in an $(n_1+n_2+c_T)$-edge polygon with span $m+t_T$. Therefore, the following inequality holds

$$\sum_{m_1=0}^{n_1/2-1} p_{n_1}(N,M;m_1)p_{n_2}(N,M;m-m_1) \le p_{n_1+n_2+c_T}(N,M;m+t_T). \tag{5.2.5}$$

Multiplying both sides of this inequality by e^{fm} and summing over m gives rise to

$$\begin{aligned} &\sum_{m=0}^{(n_1+n_2)/2-1} \sum_{m_1=0}^{n_1/2-1} p_{n_1}(N,M;m_1)p_{n_2}(N,M;m-m_1)e^{fm} \\ \le &\sum_{m=0}^{(n_1+n_2)/2-1} p_{n_1+n_2+c_T}(N,M;m+t_T)e^{fm}. \end{aligned} \tag{5.2.6}$$

Moreover, the following inequalities hold since $c_T \geq 2t_T$, $n_1/2-1 < m$ and $n_1 \geq 2m_1$.

$$\begin{aligned} Z_{n_1+n_2+c_T}(N,M;f) &= \sum_{m=0}^{(n_1+n_2+c_T)/2-1} p_{n_1+n_2+c_T}(N,M;m)e^{fm} \\ &= \sum_{j=-t_T}^{(n_1+n_2+c_T)/2-1-t_T} p_{n_1+n_2+c_T}(N,M;j+t_T)e^{f(j+t_T)}, \end{aligned} \tag{5.2.7}$$

$$Z_{n_1}(N,M;f) = \sum_{m_1=0}^{n_1/2-1} p_{n_1}(N,M;m_1)e^{fm_1}, \tag{5.2.8}$$

$$\begin{aligned} Z_{n_2}(N,M;f) &= \sum_{m=m_1}^{n_2/2+m_1-1} p_{n_2}(N,M;m-m_1)e^{f(m-m_1)} \\ &\leq \sum_{m=0}^{(n_1+n_2)/2-1} p_{n_2}(N,M;m-m_1)e^{f(m-m_1)}. \end{aligned} \tag{5.2.9}$$

Now equations 5.2.6 to 5.2.9 imply that

$$Z_{n_1}(N,M;f)Z_{n_2}(N,M;f) \leq e^{ft_T} Z_{n_1+n_2+c_T}(N,M;f). \tag{5.2.10}$$

Letting $a_n = \log e^{-ft_T} Z_{n-c_T}(N,M;f)$, equation 5.2.10 implies that $\{a_n\}$ is a super-additive sequence. Therefore, by Lemma 2.1.2

$$\mathcal{F}(N,M;f) = \lim_{n\to\infty} n^{-1}a_n = \sup_{n\geq 1} n^{-1}a_n. \tag{5.2.11}$$

Furthermore, equation 5.2.11 implies that for any even integer n

$$n^{-1}a_n \leq \mathcal{F}(N,M;f). \tag{5.2.12}$$

Multiplying both sides of the inequality by n and replacing a_n by $\log e^{-ft_T} Z_{n-c_T}(N,M;f)$ gives

$$\log e^{-ft_T} Z_{n-c_T}(N,M;f) \leq n\mathcal{F}(N,M;f); \tag{5.2.13}$$

hence

$$Z_n(N,M;f) \leq e^{(n+c_T)\mathcal{F}(N,M;f)+ft_T}. \tag{5.2.14}$$

Furthermore,

$$\begin{aligned}\max\{1, e^{f(n-1)/2}\} \leq Z_n(N, M; f) &= \sum_{m=0}^{n/2-1} p_n(N, M; m) e^{fm} \\ &\leq \max\{1, e^{f(n-1)/2}\} p_n(N, M). \quad (5.2.15)\end{aligned}$$

So

$$f/2 \leq \mathcal{F}(N, M; f) \leq \log \mu_p(N, M) + f/2 \quad (5.2.16)$$

if $f \geq 0$ while

$$0 \leq \mathcal{F}(N, M; f) \leq \log \mu_p(N, M) \quad (5.2.17)$$

if $f < 0$.

The function $Z_n(N, M; f)$ is a convex function of f; this follows immediately from the Cauchy-Schwartz inequality:

$$\begin{aligned}Z_n(N, M; f_1) Z_n(N, M; f_2) &= \sum_{m_1=0}^{n/2} p_n(N, M; m_1) e^{f_1 m_1} \sum_{m_2=0}^{n/2} p_n(N, M; m_2) e^{f_2 m_2} \\ &\geq \Big(\sum_{m=0}^{n/2} p_n(N, M; m) e^{[(f_1+f_2)/2]m} \Big)^2 \\ &= \big(Z_n(N, M; ((f_1 + f_2)/2)) \big)^2. \quad (5.2.18)\end{aligned}$$

Taking logarithms of this and dividing by n yields

$$F_n(N, M; f_1) + F_n(N, M; f_2) \geq 2F_n(N, M; (f_1 + f_2)/2). \quad (5.2.19)$$

Hence, by Lemma B.0.10, $F_n(N, M; f)$ is convex in f. Moreover,

$$\{F_n(N, M; f)\} \to \mathcal{F}(N, M; f) \quad (5.2.20)$$

as $n \to \infty$ so, by Lemma B.0.12, $\mathcal{F}(N, M; f)$ is also a convex function of f. Lemma B.0.9 implies that $\mathcal{F}(N, M; f)$ is continuous and Lemma B.0.13 shows that $\mathcal{F}(N, M; f)$ is differentiable almost everywhere. By Lemma B.0.11, the right- and left-derivatives of $\mathcal{F}(N, M; f)$ exist everywhere and they are non-decreasing functions of f. Also, by Lemma B.0.14,

$$\lim_{n\to\infty} (dF_n(N, M; f)/df) = d\mathcal{F}(N, M; f)/df \quad (5.2.21)$$

almost everywhere. $\square$

The following corollary investigates the behaviour of the expected value of $m(S_n^f)$, $E_f(m(S_n^f))$, as a function of f.

Corollary 5.2.2. *Given $f \in \mathbb{R}$,*

$$E_f(m(S_n^f)) = n\frac{d}{df}F_n(N, M; f) \tag{5.2.22}$$

almost everywhere. Hence $E_f(m(S_n^f))$ is non-decreasing in f almost everywhere.

Proof. By Theorem 5.2.1, the followig equality holds almost everywhere:

$$\begin{aligned}
\frac{d}{df}F_n(N, M; f) &= \frac{d}{df}\left[n^{-1}\log Z_n(N, M; f)\right] \\
&= n^{-1}\frac{\sum_{m=0}^{n/2-1} mp_n(N, M; m)e^{fm}}{Z_n(N, M; f)} \\
&= n^{-1}E(m(S_n^f)).
\end{aligned} \tag{5.2.23}$$

Hence by Theorem 5.2.1

$$E_f(m(S_n^f)) = n\frac{d}{df}F_n(N, M; f) \tag{5.2.24}$$

is non-decreasing in f almost everywhere. $\square$

5.3 A Pattern Theorem for Stretched Polygons in Tubes

In this section, we use the so called pattern insertion strategy, introduced in Section 2.5, to obtain a pattern theorem for stretched polygons in a tube. The pattern theorem for polygons in a tube has been proved, using this approach, in [53]. So we know that (CONCAT) and (CAPOFF), as introduced in Section 2.5, hold for SAPs in a tube [49, 53]. In order to meet the requirements of Corollary 2.5.3, which gives the pattern theorem, we only need to check the satisfaction of (CA1) and (CA2). (CA1) basically gives the definition of the clusters under investigation. Here C_n is the set of all SAPs in $T(N,M)$ and $C_n^* \subseteq C_n$ represents the set of all n-edge SAPs in $T(N,M)$ which have their left-most plane at $x=0$. It can be easily seen that C_n is invariant under translation, i.e., given

$$u \in S^* = \{u \in \mathbb{R}^3 \mid T(N,M) + u = T(N,M)\} = \{(x,0,0) \mid x \in \mathbb{Z}\}, \tag{5.3.1}$$

$G \in C_n$ implies that $G + u \in C_n$. The C_n's are pairwise disjoint since $C_n \cap C_m = \emptyset$ for $n \neq m$. So (CA1) is satisfied for SAPs in $T(N,M)$. We also need to define an appropriate weight function that satisfies (CA2) and leads to the required pattern theorem. Given $f \in \mathbb{R}$ and $C := \bigcup_{n=1}^{\infty} C_n$, we consider the following weight function:

$$wt_f : C \to (0,\infty), \tag{5.3.2}$$

such that $wt_f(G) = e^{fm}$ for any $G \in C$, where m is the span of the polygon G. This function is invariant under translation since the span does not change by translating a polygon along the x-axis, i.e.

$$wt_f(G) = wt_f(G+u) \quad \text{for every } u \in S^* \text{ and } G \in C. \tag{5.3.3}$$

The following theorem also shows that (CA2) is satisfied for this choice of weight function.

Theorem 5.3.1. *Given $f \in \mathbb{R}$ and the weight function wt_f, (CA2) is satisfied for stretched polygons.*

Proof. Given $k \geq 0$, let the two SAPs G and G' differ by at most k vertices and edges, i.e.

$$\mid E(G) \vartriangle E(G') \mid + \mid V(G) \vartriangle V(G') \mid \leq k. \tag{5.3.4}$$

Then we claim that the difference of spans of G and G' is bounded above by k, i.e. $\mid m - m' \mid \leq k$ where m and m' are the spans of G and G' respectively. Suppose to the contrary that $\mid m - m' \mid > k$. Without loss of generality assume that $m > m'$ then there are exactly $\mid m - m' \mid$ sections with span 1 in which G has at least two edges (because G is a polygon) and G' does not have any edge thus $\mid E(G) \setminus E(G') \mid \geq 2 \mid m - m' \mid > 2k$. This implies that

$$\mid E(G) \vartriangle E(G') \mid + \mid V(G) \vartriangle V(G') \mid \geq \mid E(G) \setminus E(G') \mid > 2k \tag{5.3.5}$$

which contradicts equation 5.3.4. Therefore, the following inequality must hold

$$\mid m' - m \mid \leq k. \tag{5.3.6}$$

Let $\gamma_k := e^{|f|k}$. Then multiplying both sides of equation 5.3.6 by $\mid f \mid$ gives

$$\mid f(m' - m) \mid \leq \mid f \mid k \tag{5.3.7}$$

which implies

$$- \mid f \mid k \leq f(m' - m) \leq \mid f \mid k. \tag{5.3.8}$$

Hence

$$e^{-|f|k} \leq e^{f(m'-m)} \leq e^{|f|k}; \tag{5.3.9}$$

thus

$$\frac{1}{\gamma_k} \leq e^{f(m'-m)} \leq \gamma_k. \tag{5.3.10}$$

Multiplying both sides of the above inequality by e^{fm} gives

$$\frac{1}{\gamma_k} e^{fm} \leq e^{fm'} \leq \gamma_k e^{fm}. \tag{5.3.11}$$

Hence, having $wt_f(G) = e^{fm}$ and $wt_f(G') = e^{fm'}$, the following inequality is satisfied

$$\frac{1}{\gamma_k} wt_f(G) \leq wt_f(G') \leq \gamma_k wt_f(G). \tag{5.3.12}$$

□

Therefore, since (CONCAT), (CAPOFF), (CA1) and (CA2) are all satisfied for stretched polygons, Corollary 2.5.3 yields the pattern theorem for stretched polygons. Note that

$$\begin{aligned} Z_n(N, M; f) = \sum_{G \in C_n^*} wt_f(G) &= \sum_{G \in C_n^*} e^{fm} \\ &= \sum_{m=0}^{n/2-1} p_n(N, M; m) e^{fm}. \end{aligned} \tag{5.3.13}$$

We define

$$Z_n(N, M; \leq j, P, f) = \sum_{m=0}^{n/2-1} p_n(N, M; \leq j, P, m) e^{fm}, \tag{5.3.14}$$

where $p_n(N, M; \leq j, P, m)$ represents the number of n-edge polygons (up to x-translation) with span m in $T(N, M)$ which do not contain more than j copies of the pattern P. Given $\epsilon > 0$, we also define

$$\mathcal{F}(N, M; \epsilon, P, f) \equiv \limsup_{n \to \infty} n^{-1} \log Z_n(N, M; \leq \epsilon n, P, f) \tag{5.3.15}$$

to be the limit superior of the free energy for those n-edge stretched polygons which contain no more than ϵn translates of P.

Theorem 5.3.2. *Let P be a proper pattern for polygons in $T(N, M)$. Then there exists an $\epsilon_P > 0$ such that*

$$\mathcal{F}(N, M; \leq \epsilon_P, P, f) < \mathcal{F}(N, M; f). \tag{5.3.16}$$

5.4 Transfer-Matrix Results for Stretched Polygons in Tubes

In this section the transfer-matrix argument, introduced in Section 2.6, is applied to SAPs to explore the asymptotic behaviour of stretched polygons further.

An orientation on any SAP G is defined as follows. Let v_b and $e_b = \{v_b, v\}$ be respectively the bottom vertex and edge of G. We can assign a direction to e_b by directing the edge to go from v_b to v. This will naturally induce an orientation (a direction) on G starting with edge e_b as the first edge and ordering the other edges following the direction induced by that of e_b. Hence a k-cluster config (k-SAP config) of a SAP G is considered as G's configuration in a sublattice of the form $H_{i-1}(N, M) \cup S_i(N, M) \cup H_i(N, M) \cup ... \cup S_{i+k-1}(N, M) \cup H_{i+k-1}(N, M)$ for some $1 \leq i \leq m - k + 1$, where G's configuration in such a sublattice of the tube consists of the sublattice and the ordering of the edges of G, as introduced above. With this extra information on the edges, the Cluster Configuration Axiom, given in Section 2.6, is satisfied for the following sets of k-SAP configs:

$$\Pi(k) = \{P_1, P_2, ..., P_{|\Pi(k)|}\}, \tag{5.4.1}$$

$$\Pi_1(k) = \{P'_1, P'_2, ..., P'_{|\Pi_1(k)|}\} \tag{5.4.2}$$

and

$$\Pi_2(k) = \{P''_1, P''_2, ..., P''_{|\Pi_2(k)|}\} \tag{5.4.3}$$

which are the set of all distinct proper, start and end k-SAP configs, respectively. Figure 5.2 shows an example of a polygon in $T(0, 5)$ with it's associated 2-SAP configs. Note that in Figure 5.2 the numbers beside the vertices of each 2-SAP config represent the relative order of the walk starting at that vertex, induced by the ordering of the edges of the SAP in which the pattern occurs; the relative order on the walks also determines a relative order on the edges of the 2-SAP config.

For any SAP G, let $e(G)$ and $d(G, 1)$ denote the total number of edges and the span of G respectively. For any $1 \le i \le |\Pi(k)|$ $(1 \le i \le |\Pi_1(k)|)$, let e_i (e'_i) be the number of edges in the first hinge and section of P_i (P'_i) and, for any $1 \le i \le |\Pi_2(k)|$, let e''_i be the total number of edges in P''_i. Also, for any $1 \le i \le |\Pi(k)|$ $(1 \le i \le |\Pi_1(k)|)$, let $d_i(1) = 1$ $(d'_i(1) = 1)$. Thus $D_i = [d_i(1)]^T$ and $D'_i = [d'_i(1)]^T$ $(t = 1)$. For any $1 \le i \le |\Pi_2(k)|$, let also $d''_i(1) = k$ and thus $D''_i = [d''_i(1)]^T$. Then equations 2.6.7 and 2.6.9 are satisfied.

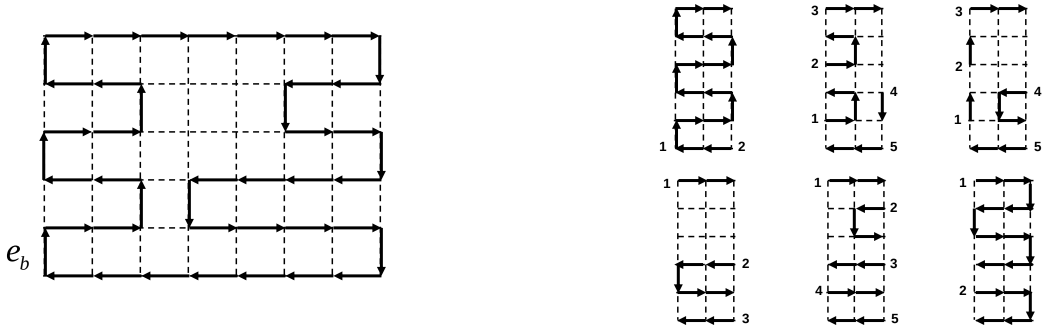
Let $y_1 = e^f$ and $Y = [y_1]^T$. Then

$$Z_n(N, M; Y) = \sum_{G \in C^*_n} y_1^{d(G,1)} = \sum_{m=0}^{n/2-1} p_n(N, M; m) e^f = Z_n(N, M; f) \tag{5.4.4}$$

and

$$Z_n(N, M; \bar{P}, Y) = \sum_{G \in C^*_n(\bar{P})} y_1^{d(G,1)} = \sum_{m=0}^{n/2-1} p_n(N, M; \bar{P}, m) e^f = Z_n(N, M; \bar{P}, f). \tag{5.4.5}$$

The definition of SAP configuration here is the same as that given in [51] except for the fact that the edges on the left- and right-most plane are not removed here.



(a) An example of a 44-edge oriented polygon with span 7 in $T(0,5)$. e_b represents the bottom edge of the polygon.

(b) The associated 2-SAP configs.

Figure 5.2: An example of a polygon in a tube and its associated 2-SAP configs. Note that the numbers beside the vertices represent the relative order of the walk starting at that vertex; the relative order on the walks also determines a relative order on the edges of the 2-SAP config.

[51, Theorem 6.1] establishes that the Cluster Configuration Axiom holds for SAP configurations as defined in [51]. Only minor adjustments to the proof of [51, Theorem 6.1] are needed to establish that the Cluster Configuration Axiom also holds for the revised SAP configuration definition used here. Theorem 2.6.1 and the fact that the Cluster Configuration Axiom and (CONCAT) hold for SAPs yield the following result regarding the asymptotic form of $Z_n(N,M;f)$ and $Z_n(N,M;\bar{P},f)$.

Theorem 5.4.1. *For any integer $k \geq 2$ and any proper SAP-config $P \in \Pi(k)$, there exist non-negative values $x_0(f)$ and α_f such that*

$$Z_n(N,M;f) = \alpha_f(x_0(f))^{-n} + o((x_0(f))^{-n}) \quad \text{as } n \to \infty. \tag{5.4.6}$$

Moreover, there exist non-negative values $\bar{x}_0(f) > x_0(f)$ and $\bar{\alpha}_f$ such that

$$Z_n(N,M;\bar{P},f) = \bar{\alpha}_f(\bar{x}_0(f))^{-n} + o((\bar{x}_0(f))^{-n}) \quad \text{as } n \to \infty. \tag{5.4.7}$$

Define $m : C^* \to \mathbb{N} \cup \{0\}$ such that $m(G)$ is the span of G, for any SAP $G \in C^*$. For $1 \leq i \leq |\Pi(k)|$ $(1 \leq i \leq |\Pi_1(k)|)$, let $m(P_i) = 1$ $(d_l'(P_i') = 1)$. Similarly, for

$1 \leq i \leq |\Pi_2(k)|$, let $d_l''(P_i'') = k$. Then equation 2.6.23 is satisfied and thus m is an additive functional.

Let $t = 1$ and $y_1 = e^f$ as in Theorem 5.4.1. Then $m(S_n^f)$ is a random variable, representing the span of a randomly chosen stretched polygon S_n^f, with the probability distribution

$$\mathbb{P}(m(S_n^f) = m) = \frac{p_n(N, M; m)e^{fm}}{Z_n(N, M; f)}. \tag{5.4.8}$$

Theorem 2.6.2 implies that there exists $\gamma_f > 0$ such that

$$E_f(m(S_n^f)) = \gamma_f n + O(1) \tag{5.4.9}$$

and thus

$$\frac{1}{n} E_f(m(S_n^f)) \to \gamma_f \tag{5.4.10}$$

as $m \to \infty$. Hence, the average span per edge of a randomly chosen n-edge stretched polygon, $E_f(m(S_n^f))$, approaches a positive constant as $n \to \infty$. Moreover, by Corollary 5.2.2, $E_f(m(S_n^f))$ is non-decreasing in f almost everywhere.

Given a proper k-SAP config $P_l \in \Pi(k)$, define $\psi_l : C^* \to \mathbb{N} \cup \{0\}$ such that $\psi_l(G)$ is the number of times the SAP config P_l appears in G, for any SAP $G \in C^*$. For $1 \leq i \leq |\Pi(k)|$, let $\psi_l(P_i) = \delta_{i,l}$ where

$$\delta_{i,l} = \begin{cases} 1 & \text{if } i = l \\ 0 & \text{if } i \neq l \end{cases}. \tag{5.4.11}$$

Similarly, for $1 \leq i \leq |\Pi_1(k)|$ ($1 \leq i \leq |\Pi_2(k)|$), let $d_l'(P_i') = \delta_{i,l}$ ($d_l''(P_i'') = \delta_{i,l}$). Then equation 2.6.23 is satisfied thus ψ_l is an additive functional.

Let $t = 1$ and $y_1 = e^f$ as in Theorem 5.4.1. Then $\psi_l(S_n^f)$ is a random variable, representing the number of times the SAP config P_l appears in a randomly chosen span m stretched polygon S_n^f, with the probability distribution

$$\mathbb{P}(\psi_l(S_n^f) = b) = \frac{\sum_{G:\psi_l(G)=b} e^{fm(G)}}{Z_n(N, M; f)}, \tag{5.4.12}$$

where the sum is over all the n-edge SAPs $G \in C_n^*$ with $\psi_l(G) = b$ and $m(G)$ denotes the span of G. Theorem 2.6.2 implies that there exists $\gamma_f^l > 0$ such that

$$E_f(\psi_l(S_n^f)) = \gamma_f^l n + O(1) \tag{5.4.13}$$

and thus

$$\frac{1}{n} E_f(\psi_l(S_n^f)) \to \gamma_f^l \tag{5.4.14}$$

as $m \to \infty$.

5.5 Loops in Tubes

A *loop* in $T(N, M)$ is an undirected self-avoiding walk with both endpoints (vertices of degree one) on its left-most plane. Let $l_n(N, M; m)$ be the number of loops in $T(N, M)$ with span m and n edges. We show next that loops have the same asymptotic behaviour as polygons in a tube. In fact, in a tube the pattern theorem for stretched polygons results in a pattern theorem for loops in the presence of an external force f.

Let

$$L_n(N, M; f) = \sum_{m=0}^{n/2-1} l_n(N, M; m) e^{fm}. \tag{5.5.1}$$

The following theorem shows that loops in the presence of force f have the same limiting free energy as stretched polygons.

Theorem 5.5.1. *Given $f \in \mathbb{R}$,*

$$\mathcal{F}_l(N, M; f) = \mathcal{F}(N, M; f), \tag{5.5.2}$$

where

$$\mathcal{F}_l(N, M; f) = \lim_{n \to \infty} n^{-1} \log L_n(N, M; f) \tag{5.5.3}$$

is the limiting free energy for loops in $T(N, M)$.

Proof. Given an n-edge polygon G with span m, let e_b be the bottom edge of G. We can obtain an $(n-1)$-edge loop G_l with span m from G just by deleting the edge e_b. So the following inequality holds

$$p_n(N, M; m) \le l_{n-1}(N, M; m). \tag{5.5.4}$$

Multiplying both sides by e^{fm} and summing over m gives

$$Z_n(N, M; m) \leq L_{n-1}(N, M; m). \tag{5.5.5}$$

Taking logarithms, multiplying both sides of this equation by n^{-1} and letting $n \to \infty$ implies

$$\mathcal{F}(N, M; f) \leq \mathcal{F}_l(N, M; f). \tag{5.5.6}$$

Furthermore, given an n-edge loop G_l in $T(N, M)$ with span m, if $M + N = 1$ then we can connect the two endpoints of G_l by a single edge and we will have

$$l_n(N, M; m) \leq p_{n+1}(N, M; m) \tag{5.5.7}$$

which leads to

$$\mathcal{F}_l(N, M; f) \leq \mathcal{F}(N, M; f). \tag{5.5.8}$$

If $N+M > 1$ then for all the even values (odd values) of n we show next that we can find fixed numbers c_T (c'_T) and t_T (t'_T) so that the endpoints of G_l can be connected in order to create an $(n + c_T)$-edge ($(n + c'_T)$-edge) polygon G with span $m + t_T$ ($m+t'_T$); an example is shown in Figure 5.3. Assume that n is an even number. Let $v_1 = (x(v_1), y(v_1), z(v_1))$ and $v_2 = (x(v_2), y(v_2), z(v_2))$ be the two endpoints of the loop G_l, where $v_1 < v_2$ lexicographically. We connect the two endpoints v_1 and v_2 using the walk $w : -\hat{\imath}, (y(v_2) - y(v_1))\hat{\jmath}, (z(v_2) - z(v_1))\hat{k}, \hat{\imath}$ starting at v_1. This results in an $(n + n(w))$ self-avoiding polygon G_1 with span $m + 1$, where $n(w)$ denotes the number of edges in w; note that $n(w)$ is an even number and $4 \leq n(w) \leq (2+N+M)$.

Let $e'_b = \{v_b, v\}$ be the bottom edge of G_1, where v_b is the bottom vertex. Since $M + N > 1$, we can choose $\bar{u} \in \{\pm\hat{k}, \pm\hat{\jmath}\}$ so that the walk w' (or w'' if w' does not work) introduced as follows will stay inside the tube $T(N, M)$. Let w' be the walk $-2(N + M + 2)\hat{\imath}, (y(v) - y(v_b))\hat{\jmath}, (z(v) - z(v_b))\hat{k}, ((2(N + M + 2) - n(w))/2)[\hat{u}, \hat{\imath}, -\hat{u}, \hat{\imath}], n(w)\hat{\imath}$ starting at v_b. Similarly, let w'' be the walk $-2(N + M +$

$2)\hat{\imath}, (y(v_b)-y(v))\hat{\jmath}, (z(v_b)-z(v))\hat{k}, ((2(N+M+2)-n(w))/2)[\hat{u}, \hat{\imath}, -\hat{u}, \hat{\imath}], n(w)\hat{\imath}$ starting at v. The polygon G with $(n+6(N+M+2))$ edges and span $(m+2(N+M+2)+1)$ can be obtained by deleting e_b' and adding the walk w' (or w'') to G_1.

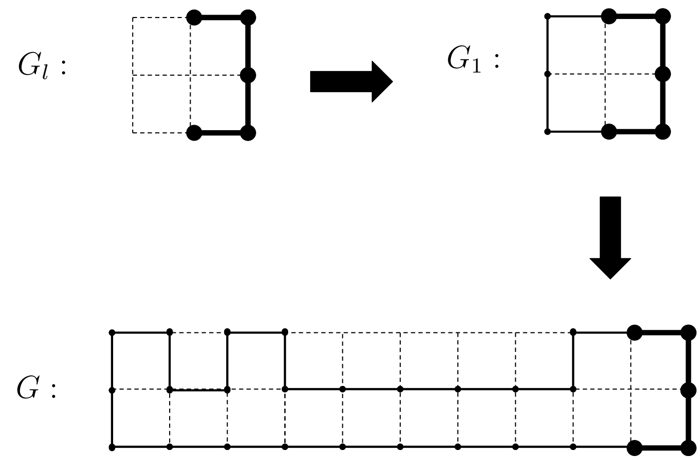


Figure 5.3: An example of converting a 4-edge loop G_l, shown with thicker edges, with span 1 in $T(0,2)$ to a 28-edge polygon G with span 10 in $T(0,2)$. Here for $T(0,2)$, $c_T = 24$, $t_T = 9$ and $n(w) = 4$. Note that for this specific loop G_l the walk w' is used with $u = \hat{k}$.

Therefore, choosing $c_T = 6(N+M+2)$ and $t_T = 2(N+M+2)+1$ implies

$$l_n(N,M;m) \leq p_{n+c_T}(N,M;m+t_T). \tag{5.5.9}$$

Multiplying both sides by $e^{f(m+t_T)}$ and summing over m gives

$$\begin{aligned} \sum_{m=0}^{n/2-1} l_n(N,M;m)e^{f(m+t_T)} &\leq \sum_{m=0}^{n/2-1} p_{n+c_T}(N,M;m+t_T)e^{f(m+t_T)} \\ &\leq \sum_{m=-t_T}^{(n+c_T)/2-t_T-1} p_{n+c_T}(N,M;m+t_T)e^{f(m+t_T)} \end{aligned} \tag{5.5.10}$$

thus

$$e^{ft_T}L_n(N,M;f) \leq Z_{n+c_T}(N,M;f). \tag{5.5.11}$$

Note first that the number of edges of a polygon has to be an even number. Hence, for the case that G_l had an even number of edges (n is even), an even number of edges was also required to construct the SAP G, i.e. c_T had to be even. On the other hand, for the case that G_l has an odd number of edges (n is odd), an odd number of edges (c'_T edges) is required to construct the polygon G. This can be done using a similar strategy that was used for loops with an even number of edges; i.e. for n odd

$$e^{ft'_T} L_n(N, M; f) \leq Z_{n+c'_T}(N, M; f). \tag{5.5.12}$$

Taking logarithms of both sides of equation 5.5.11 and 5.5.12, multiplying both sides by n^{-1} and letting $n \to \infty$ yields

$$\mathcal{F}_l(N, M; f) \leq \mathcal{F}(N, M; f). \tag{5.5.13}$$

Therefore, equations 5.5.6, 5.5.8 and 5.5.13 together imply

$$\mathcal{F}_l(N, M; f) = \mathcal{F}(N, M; f). \tag{5.5.14}$$

□

Note first that, by the definition of proper patterns stated in Chapter 2, P is a proper pattern for loops if 1) there are infinitely many values of n such that P is contained in an n-edge loop with left-most plane $x = 0$ and 2) there exists a loop G such that P occurs in the middle of G. The following argument shows that every proper SAP pattern is also a proper pattern for loops. By (CONCAT) and (CAPOFF) for SAPs, there are infinitely many values of n such that a proper SAP pattern P occurs in the middle of an n-edge SAP G_n with left-most plane $x = 0$; i.e. P occurs in G_n so that the left-most (right-most) plane of P is not the same as the left-most (right-most) plane of G_n. As was explained in the proof of Theorem 5.5.1, we can convert each G_n to a loop G'_n by just deleting the bottom edge e_b of G_n. Note that e_b lies on the left-most plane of G_n so it does not change the configuration of the pattern P. Thus every occurrence of the pattern P in the middle of a SAP G_n

leads to the occurrence of P in the middle of a loop G'_n. Therefore, P is also a proper pattern for loops. Similarly, it can be shown that every proper pattern for loops is also a proper SAP pattern. In particular, every proper pattern P for loops occurs at the start or in the middle of a loop G_l; i.e. P occurs in G_l so that the right-most plane of P is not the same as the right-most plane of G_l. As was explained in the proof of Theorem 5.5.1, we can convert G_l to a SAP G by increasing the span by $t_T \geq 1$ while keeping the configuration of G_l fixed. This will lead to the occurrence of P in the middle of G. Therefore, P is also a proper SAP pattern.

Let

$$L_n(N, M; \leq j, P, f) = \sum_{m=0}^{n/2-1} l_n(N, M; \leq j, P, m)e^{fm}, \tag{5.5.15}$$

where $l_n(N, M; \leq j, P, m)$ denotes the number of n-edge loops (up to x-translation) with span m which do not contain more than j copies of the proper SAP pattern P. Also define

$$\mathcal{F}(N, M; \leq \epsilon, P, f) \equiv \limsup_{n\to\infty} n^{-1} \log L_n(N, M; \leq \epsilon n, P, f) \tag{5.5.16}$$

to be the limit superior of the free energy for those n-edge loops which contain no more than ϵn translates of P. The following is a pattern theorem for loops in $T(N, M)$.

Theorem 5.5.2. *Let P be a proper SAP k-pattern in $T(N, M)$. Then there exists an $\epsilon_P > 0$ such that*

$$\mathcal{F}_l(N, M; \leq \epsilon_P, P, f) < \mathcal{F}_l(N, M; f). \tag{5.5.17}$$

Proof. Since P is a proper SAP pattern, by Theorem 5.3.2 there exists an $\epsilon'_P > 0$ such that

$$\mathcal{F}(N, M; \leq \epsilon'_P, P, f) < \mathcal{F}(N, M; f). \tag{5.5.18}$$

The above equation and Theorem 5.5.1 imply

$$\mathcal{F}(N, M; \leq \epsilon'_P, P, f) < \mathcal{F}_l(N, M; f). \tag{5.5.19}$$

Let t_T be the constant introduced in the proof of Theorem 5.5.1. Let N be the smallest positive integer number such that $\epsilon'_P - t_T/N > 0$. Note that such a number exists since $\lim_{n\to\infty} t_T/n = 0$. Set $\epsilon_P = \epsilon'_P - t_T/N$. Let G_l be an n-edge loop with span m which contains no more than $\epsilon_P n$ copies of the pattern P. Regarding the construction of the $(n+c_T)$-edge polygon G with span $m+t_T$ from G_l, as described in the proof of Theorem 5.5.1, in addition to the k-patterns occurring in G_l, G will contain at most t_T new k-patterns. This may lead to the occurrence of at most t_T extra copies of P in G. Hence G will contain in total at most $\epsilon_P n + t_T$ copies of P. Note that, for $n \geq N$, we have

$$\epsilon'_P - t_T/N \leq \epsilon'_P - t_T/n \qquad (5.5.20)$$

so

$$\epsilon_P n + t_T \leq \epsilon'_P n - t_T + t_T \leq \epsilon'_P n \leq \epsilon'_P(n+c_T). \qquad (5.5.21)$$

Therefore, G contains at most $\epsilon'_P(n+c_T)$ copies of P and the following inequality holds

$$l_n(N,M;\leq \epsilon_P n, P, m) \leq p_{n+c_T}(N,M;\leq \epsilon'_P(n+c_T), P, m+t_T). \qquad (5.5.22)$$

Using arguments similar to that presented in Theorem 5.5.1, the above equation leads to

$$\mathcal{F}_l(N,M;\leq \epsilon_P, P, f) \leq \mathcal{F}(N,M;\leq \epsilon'_P, P, f). \qquad (5.5.23)$$

Therefore, equations 5.5.19 and 5.5.23 together yield

$$\mathcal{F}_l(N,M;\leq \epsilon_P, P, f) < \mathcal{F}_l(N,M;f). \qquad (5.5.24)$$

□

5.6 Knotting and Topological Linking of Stretched Polygons and Loops in Tubes

Having the pattern theorem for both stretched polygons and loops in a tube, we can now discuss the knotting probability. In particular, we can take P to be a tight trefoil pattern (e.g. the pattern shown in Figure 3.7) in $T(N, M)$ and prove that the knotting probability goes to one as $n \to \infty$ for any arbitrary value of f.

Let

$$Z_n^\circ(N, M; f) = \sum_{m=0}^{n/2-1} p_n^\circ(N, M; m) e^{fm}, \tag{5.6.1}$$

where $p_n^\circ(N, M; m)$ is the number of unknotted n-edge SAPs with span m in $T(N, M)$. Recall that concatenating two unknotted polygons results in an unknotted polygon. So the proof of Theorem 5.2.1 can be modified in a straightforward fashion to show the existence of the limiting free energy for unknotted stretched polygons.

Theorem 5.6.1. *The following limit exists:*

$$\mathcal{F}^o(N, M; f) \equiv \lim_{n\to\infty} n^{-1} \log Z_n^o(N, M; f). \tag{5.6.2}$$

The following theorem discusses the asymptotic behaviour of unknotted stretched polygons.

Theorem 5.6.2. *Given $f \in \mathbb{R}$, for N and M such that the (N, M)-tube can contain the tight trefoil pattern P introduced in Figure 3.7, the following inequality holds:*

$$\mathcal{F}^o(N, M; f) < \mathcal{F}(N, M; f) \tag{5.6.3}$$

and the probability that a stretched polygon in an (N, M)-tube is knotted goes to one exponentially as $n \to \infty$.

Proof. For the tight trefoil proper SAP pattern P, as shown in Figure 3.7 (a), let ϵ_P be the positive number as prescribed by Theorem 5.3.2. An n-edge unknotted polygon cannot contain the tight trefoil pattern P, thus it contains at most $0 < \epsilon_P n$ copies of P. Hence

$$p_n^\circ(N, M; m) \leq p_n(N, M; \leq \epsilon_P n, P, m). \tag{5.6.4}$$

Multiplying both sides by e^{fm} and summing over m implies that

$$Z_n^o(N, M; f) \leq Z_n(N, M; \leq \epsilon_P n, P, f). \tag{5.6.5}$$

Taking logarithms, multiplying both sides by n^{-1} and letting $n \to \infty$ gives

$$\mathcal{F}^o(N, M; f) \leq \mathcal{F}(N, M; \epsilon_P, P, f). \tag{5.6.6}$$

Therefore, the pattern theorem for stretched polygons, Theorem 5.3.2, together with equation 5.6.6 show that

$$\mathcal{F}^o(N, M; f) < \mathcal{F}(N, M; f). \tag{5.6.7}$$

Furthermore, the probability that a stretched polygon is knotted is given by

$$\frac{Z_n(N, M; f) - Z_n^o(N, M; f)}{Z_n(N, M; f)} = 1 - \frac{Z_n^o(N, M; f)}{Z_n(N, M; f)}. \tag{5.6.8}$$

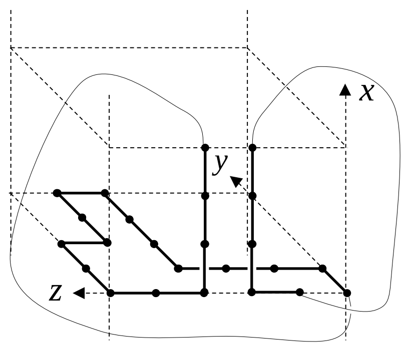
Using a similar argument to that given in the proof of Theorem 4.5.2, this fraction goes to one as $1 - e^{-\gamma_o(N,M;f)n+o(n)}$ when $n \to \infty$, with

$$\gamma_o(N, M; f) \equiv \mathcal{F}(N, M; f) - \mathcal{F}^o(N, M; f). \tag{5.6.9}$$

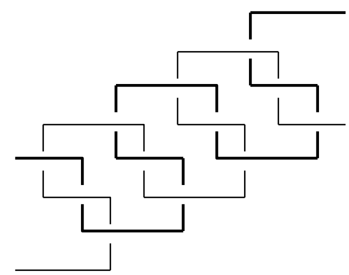
□

Removing the bottom edge and top edge of each polygon G (the top edge of a loop G_l) with span m results in a pair of mutually avoiding self-avoiding walks with span m, say (w_1, w_2), where w_1 contains the bottom vertex. One way of investigating the topological entanglement of polygons (loops) is to see how these two walks w_1 and w_2 are linked to each other. We can consider the configuration of these two walks as a 2-string tangle (B, t) where $t = \{w_1, w_2\}$ and B is the 3-ball $[x_1, x_2] \times [0, N] \times [0, M]$

with $x = x_1$ and $x = x_2$ being the left- and right-most planes of G (G_l) respectively. We can join the two ends of each walk from outside of the 3-ball and create a two-component link (K_1, K_2) and ask about the topological linking probability of (K_1, K_2). For this purpose, we will apply the pattern theorem to the pattern $A' + A'$ illustrated in Figure 5.4 (b).



(a) An example showing how to get the pair of self-avoiding walks (w_1, w_2) and the associated link (K_1, K_2).



(b) The pattern $A' + A'$.

Figure 5.4

Let

$$Z_n^u(N, M; f) = \sum_{m=0}^{n/2-1} p_n^u(N, M; m) e^{fm}, \tag{5.6.10}$$

where $p_n^u(N, M; m)$ represents the number of n-edge polygons with span m such that their associated link (K_1, K_2) is topologically linked (non-splittable).

Theorem 5.6.3. *Given $f \in \mathbb{R}$, for N and M such that the (N, M)-tube can contain the pattern $A' + A'$:*

$$\mathcal{F}^u(N, M; f) < \mathcal{F}(N, M; f), \tag{5.6.11}$$

where

$$\mathcal{F}^u(N, M; f) \equiv \limsup_{n\to\infty} n^{-1} \log Z_n^u(N, M; f). \tag{5.6.12}$$

Hence the probability that an n-edge polygonal link (K_1, K_2) is topologically linked (non-splittable) goes to one as $n \to \infty$.

Proof. By Theorem 4.6.1 the occurrence of the pattern $P_A = A' + A'$ in any polygon G guarantees the topological linking of the associated link (K_1, K_2). Let ϵ_{P_A} be the positive number associated with the pattern P_A as prescribed by Theorem 5.3.2. Hence

$$p_n^u(N, M; m) \leq p_n(N, M; \leq \epsilon_{P_A} n, P_A, m). \tag{5.6.13}$$

Multiplying both sides by e^{fm} and summing over m implies

$$Z_n^u(N, M; f) \leq Z_n(N, M; \leq \epsilon_{P_A} n, P_A, f). \tag{5.6.14}$$

Taking logarithms, multiplying both sides by n^{-1} and letting $n \to \infty$ gives

$$\mathcal{F}^u(N, M; f) \leq \mathcal{F}(N, M; \epsilon_{P_A}, P_A, f). \tag{5.6.15}$$

Thus using the pattern theorem for stretched polygons, Theorem 5.3.2, we have

$$\mathcal{F}^u(N, M; f) < \mathcal{F}(N, M; f). \tag{5.6.16}$$

Furthermore, the probability that the link (K_1, K_2) associated to a stretched polygon is topologically linked is given by

$$\frac{Z_n(N, M; f) - Z_n^u(N, M; f)}{Z_n(N, M; f)} = 1 - \frac{Z_n^u(N, M; f)}{Z_n(N, M; f)}. \tag{5.6.17}$$

This fraction goes to one as $1 - e^{-\gamma_u(N,M;f)n + o(n)}$ when $n \to \infty$, with $\gamma_u(N, M; f) \equiv \mathcal{F}(N, M; f) - \mathcal{F}^u(N, M; f)$. $\square$

Note that similar results can also be proved for loops.

5.7 Conclusions

In this chapter, we examined the topological entanglements of polygons (loops) confined to a tube and under the influence of an external force f. We proved a pattern theorem for such polygons, the so-called stretched polygons. We used this to also obtain a pattern theorem for loops in a tube. Note that the tube constraint allowed us to prove the pattern theorem for any arbitrary value of f, not just for large values of f as in [27]. The pattern theorem was then used to show that the knotting probability of an n-edge stretched polygon confined to a tube goes to one exponentially as $n \to \infty$. Furthermore, associating a two-component link to each polygon (loop) in a tube, the incidence of topological linking was examined when the polygon (loop) was under the influence of a force f. We proved that the probability that the link associated to an n-edge polygon (loop) is topologically linked approaches unity exponentially rapidly as $n \to \infty$. This implies that as $n \to \infty$ when polygons are influenced by an external force f, no matter its strength or direction, topological entanglements, as defined by knotting and topological linking, occur with high probability.

Furthermore, we proved the existence of the limiting free energy for stretched polygons and analyzed the asymptotic form of the partition function for stretched polygons, using the transfer-matrix approach. We showed that, the average span per edge of a randomly chosen n-edge stretched polygon S_n^f, $\frac{E_f(m(S_n^f))}{n}$, approaches a positive constant as $n \to \infty$. We also proved that $E_f(m(S_n^f))$ is non-decreasing in f almost everywhere. Let P_* be the proper SAP config obtained by ordering the edges of the tight trefoil pattern P introduced in Figure 3.7. Then, by proving equation 5.4.14, we established that the average number of occurrences of P_* per edge in any n-edge stretched polygon S_n^f approaches a positive constant as $n \to \infty$. Future work would involve investigating how this limiting value depends on f.

Chapter 6

The Entanglement Complexity of Polymer Systems Modelled by SSAWs

6.1 Introduction

Polymer molecules can become both self-entangled and entangled with other molecules. These entanglements can have significant effects on some properties of the polymer such as the crystallization behaviour of the polymer [17]. The self-entanglement of an isolated ring polymer, modelled by a single self-avoiding polygon in $\mathbb{Z}^3$, is well studied using topological measures such as the probability of knotting [25, 42]. However, many problems about entanglements in polymers in dense systems such as melts are left unanswered [40, 41]. Most results about the entanglement complexity of polymers in melts are obtained using numerical studies [13, 40, 41]. However, at least the following open questions still remain and need further investigation both rigorously and numerically:

- What is the best way to characterize the entanglement complexity of polymers in dense systems?
- How does the entanglement complexity depend on some properties of the system such as the total number of interpenetrating monomers in the system, the number of polymers in the system and the density (the number of monomers per unit volume)?

One way to characterize this entanglement complexity, proposed by Orlandini *et al* in 2000 [40], is as follows. A polymer melt is considered as a set of entangled chains, modelled by a set of self- and mutually avoiding walks. Fixing the chain conformations, imagine cutting a cube or tube out of the system. The conformations of the parts of the chains which are in the interior of the cube or tube are considered. Using the simple cubic lattice model, they investigated this numerically by studying a number of self- and mutually avoiding walks confined to a cube. For each pair of self-avoiding walks, by joining the two ends (vertices of degree one) of each walk outside the cube, a two-component link is formed. And they take the sum of the absolute value of the linking numbers, over all the possible SAW pairs, as a measure of the entanglement complexity of this polymer system. Then the properties of the complete melt can be inferred by investigating the properties of these chains in cubes.

In order to investigate Orlandini *et al*'s proposal further we address the problem of the entanglement complexity of a polymer system modelled by a system of self-avoiding walks (SSAW) confined to a tube, defined in Section 6.2. An example of a polymer system is sketched in Figure 6.1. Our goal is to build a theoretical framework that will allow us to apply the available mathematical techniques towards developing and understanding the Orlandini *et al* 2004 model and answering the following questions:

- How does the entanglement complexity change with respect to the total number of monomers in the system?
- How does the entanglement complexity change with respect to the span of the system along the tube?
- How does the entanglement complexity change with respect to the number of chains in the system?
- How does the entanglement complexity change with respect to the systems's density?
- How does the entanglement complexity change with respect to the size of the tube to which the polymers are confined?

- How do the total number of monomers, the span, the number chains, the density and the size of the tube depend on each other?

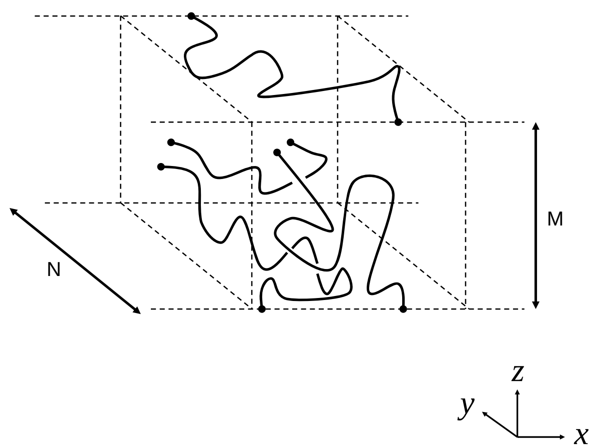


Figure 6.1: An example of a polymer system in an (N, M)-tube.

Specifically, we investigate, under some constraints, the entanglement complexity of several self- and mutually avoiding walks confined to an (N, M)-tube. We rigorously prove that the entanglement complexity, as measured in [41], of a polymer system with size n is asymptotically (with probability one) bounded below by a linear function of n; i.e. there exists a positive number γ such that the probability that a polymer system of size n has entanglement complexity greater than γn approaches 1 as n goes to infinity. It is also shown that the entanglement complexity of SSAWs of size n is bounded above by a linear function of n. Note that here "size" can be measured by the total number of edges, the span or the number of degree one vertices (or, equivalently, twice the number of disjoint walks) in SSAWs . Furthermore, measuring size by the number of edges, for $N \geq 2$ and $M \geq 2$, the connective constant for SSAWs in an (N, M)-tube is compared with the connective constant for self-avoiding walks in an $(N-2, M-2)$-tube and is shown to be strictly greater than that for SAWs.

For $Y_3 = [x, y, z]^T$, let $X_n^{\#1}(Y_3)$ be a random variable taking its values from the set of SSAWs with span m (in C_m^*) and with the probability distribution

$$\mathbb{P}(X_m^{\#1}(Y_3) = G) = \frac{x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}}{Z_m^{\#1}(N, M, Y_3)}, \tag{6.1.1}$$

where

$$Z_m^{\#1}(N, M, Y_3) = \sum_{G \in C_m^*} x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}. \tag{6.1.2}$$

We use the transfer-matrix results proved in Section 2.6 to obtain the asymptotic form of the free energy for SSAWs. Let $f(X_m^{\#1}(Y_3), 3)$ be a random variable representing the number of edges in $X_m^{\#1}(Y_3)$. We use the transfer-matrix approach to explore the asymptotic behaviour of the *average edge-density* of SSAWs defined by $E_Y(\frac{f(X_m^{\#1}(Y_3),3)}{NMm})$, i.e. the expected value of $\frac{f(X_m^{\#1}(Y_3),3)}{NMm}$.

Moreover, we investigate the asymptotic behaviour of SSAWs in $T(N, M)$ with a fixed edge-density, where the *edge-density* of an n-edge SSAW with a fixed span m is defined by $\rho(n) = \frac{n}{NMm}$. In particular, we look at span m SSAWs which contain $\lfloor \epsilon m \rfloor$ edges, for some $\epsilon > 0$. Let $q_m^s(N, M; \lfloor \epsilon m \rfloor)$ denote the number (up to x-translation) of such SSAWs in $T(N, M)$. In this case,

$$\rho(\lfloor \epsilon m \rfloor) = \frac{\lfloor \epsilon m \rfloor}{NMm} \tag{6.1.3}$$

and the *limiting edge-density* is defined by

$$\rho_* \equiv \lim_{m \to \infty} \frac{\lfloor \epsilon m \rfloor}{NMm} = \frac{\epsilon}{NM}. \tag{6.1.4}$$

We investigate the growth constant for SSAWs with a fixed limiting edge-density.

This chapter is organized as follows. In Section 6.2 we give a precise definition of SSAWs. In Section 6.3, (CONCAT) and (CAPOFF) are proved to be satisfied for SSAWs and the existence of the limiting free energy is shown in Section 6.4. Then a pattern theorem is proved for the limiting free energy in Section 6.5. In Section 6.6, it is shown that the connective constant of SSAWs in $T(N, M)$ (when size is measured by the number of edges) is strictly greater than that of SAWs in an $(N-2, M-2)$-tube, for $N \geq 2$ and $M \geq 2$. Then, in Section 6.7, we use the pattern theorem to investigate the asymptotic behaviour of the entanglement complexity of polymer systems. In Section 6.8, we prove that the entanglement complexity of any n-edge (span n) SSAW cannot grow faster than linearly in n. In Section 6.9, we use the transfer-matrix results obtained in Section 2.6 to determine the asymptotic form of the free energy for SSAWs. Then, in Section 6.10, some questions regarding the asymptotic behaviour of the average edge-density of SSAWs are addressed. Furthermore, in Section 6.11, we discuss the properties of the growth constant for SSAWs with a fixed limiting edge-density. Finally, we finish this chapter by giving a summary of the obtained results in Section 6.12.

6.2 SSAWs

Here we investigate the asymptotic behaviour of the entanglement complexity of polymer systems, measured using a quantity based on linking numbers. For this purpose a polymer system is modelled mathematically by a system of self-avoiding walks defined as follows:

Definition 6.2.1. A *System of Self-avoiding Walks* (SSAW) with n edges in an (N, M)-tube is a finite subgraph G of $T(N, M)$ satisfying the following conditions:

(i) The total number of edges in G is n.

(ii) Each connected component of G is an undirected self-avoiding walk which has both its vertices of degree one and no vertex of degree two on the boundaries of the tube (the four planes $y = 0$, $y = N$, $z = 0$ and $z = M$ are considered the boundaries of the (N, M)-tube).

(iii) For

$$x_1 \equiv \min\{x|(x, y, z) \text{ is a vertex of } G\},$$

$$x_2 \equiv \max\{x|(x, y, z) \text{ is a vertex of } G\}$$

and any integer $x_0 \in [x_1, x_2]$, G has at least one vertex (x, y, z) with $x = x_0$.

An example of an SSAW is illustrated in Figure 6.2. Consistent with the definitions of the span and the left-most (right-most) plane for any finite subgraph of $\mathbb{Z}^3$, presented in Section 2.1, $x_2 - x_1$ is defined to be the *span* of the SSAW and $x = x_1$ ($x = x_2$) is said to be the *left-most (right-most) plane* of the SSAW.

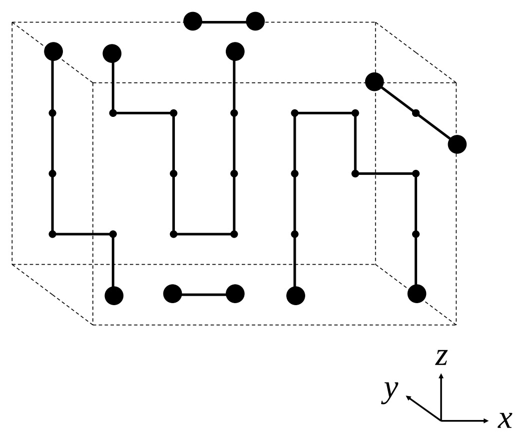


Figure 6.2: An example of a 25-edge SSAW in $T(2, 4)$ with span 6.

In order to measure the entanglement complexity of an SSAW G according to the Orlandini *et al* 2004 approach we need to associate a two-component link to each pair of walks in G. Next we will explain the construction of such a link. To do this nicely we first extend the walk endpoints in an SSAW so that they lie in either the plane $y = 0$ or $y = N$.

Let G be an SSAW composed of the sequence of self- and mutually avoiding undirected walks $w_1, w_2, ..., w_k$, where k is the total number of walks in G and the walks have been ordered according to the following algorithm: Among the two endpoints (vertices of degree one) of each undirected self-avoiding walk w_i, for $1 \leq i \leq k$, let v_i be the vertex which is lexicographically smallest. Then the walks are ordered from $i = 1, ..., k$ according to the lexicographical order of the v_i's. We will assume this order for the walks in any SSAW G after this, unless otherwise stated.

Lemma 6.2.1. *Given two non-negative integers N and M, let G be an n-edge SSAW in $T(N, M)$. There exists an SSAW $\bar{G}$ in $T(N, M+2N)$ such that $\bar{G} \cap (G + (0, 0, N)) =$*

$G + (0, 0, N)$, the number of edges in $\bar{G}$ is bounded above by a linear function $h(n)$ where $h(n) = (4N - 3)n$, and $\bar{G}$ has all its degree one vertices in the planes $y = 0$ and $y = N$. $\bar{G}$ is contained in the box

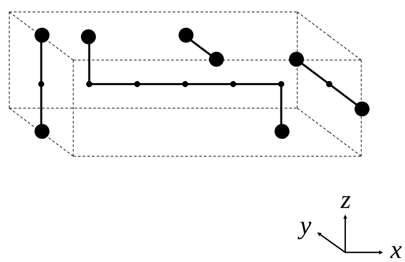
$$B_G := \{(x, y, z) \in \mathbb{Z}^3 \mid x_1 \leq x \leq x_2, 0 \leq y \leq N, 0 \leq z \leq M + 2N\}, \qquad (6.2.1)$$

where $x = x_1$ and $x = x_2$ are respectively the left- and right-most planes of G.

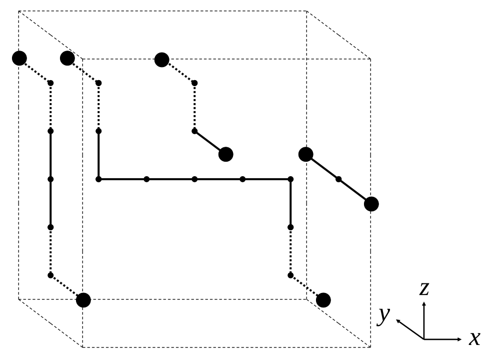
Proof. Given an n-edge SSAW G in $T(N, M)$, let $w_1, ..., w_k$ be the sequence of undirected self-avoiding walks composing G. For $i = 1, ..., k$, let n_i be the number of edges in w_i. For every USAW w_i in G we associate a new USAW $\bar{w}_i$ as follows (as an example, see Figure 6.3): For every x^* such that $x_1 \leq x^* \leq x_2$, we order lexicographically all the degree one vertices $(x^*, y, 0)$ in the plane $z = 0$ with $0 < y < N$ and $x = x^*$, if there are any. Let $v_1 < ... < v_r$ (for some $0 \leq r \leq (N-1)$) represent the sequence of these vertices. Then for $j = 1, ..., r$, we add the walk $(-j\hat{k}, -y(v_j)\hat{j})$ starting at v_j and then ignore the orientation on the edges. Similarly, we order lexicographically all the degree one vertices of G in the plane $z = M$ with $0 < y < N$ and $x = x^*$. Say $v_1 > ... > v_r$ (for some $0 \leq r \leq (N-1)$), where the vertices are labelled from largest to smallest. Then for $j = 1, ..., r$, we add the walk $(j\hat{k}, (N - y(v_j))\hat{j})$ starting at v_j and ignore the orientation on the edges. Under this transformation, any USAW in G will be either extended, at least from one of its endpoints, or it will remain unchanged. Now, we translate all the USAWs N units in the positive z-direction. For $i = 1, ..., k$, we denote by $\bar{w}_i$ such a USAW obtained from w_i; note that each $\bar{w}_i$ has $(n_i + m_i)$ edges where $m_i \leq 4(N-1)$. This results in a new SSAW, $\bar{G}$, in $T(N, M + 2N)$ with the property that the USAWs in $\bar{G}$ have all their endpoints on the boundary planes $y = 0$ and $y = N$. It is also clear from the construction that $\bar{G} \cap (G + (0, 0, N)) = G + (0, 0, N)$. Moreover, the total number of edges in $\bar{G}$, $n(\bar{G})$, satisfies the following inequality

$$\begin{aligned} n(\bar{G}) = \sum_{i=1}^{k}(n_i + m_i) \leq \sum_{i=1}^{k} n_i + \sum_{i=1}^{k} m_i &\leq n + 4k(N-1) \\ &\leq n + 4(N-1)n \\ &= (4N-3)n. \end{aligned} \tag{6.2.2}$$

□



(a) The 11-edge SSAW G in $T(2,2)$ with span 6. G is composed of the sequence of walks w_1, w_2, w_3, w_4 with $n_1 = 2$, $n_2 = 6$, $n_3 = 1$ and $n_4 = 2$.



(b) The 21-edge SSAW $\bar{G}$ in $T(2,6)$ with span 6. $\bar{G}$ is composed of the sequence of walks $\bar{w}_1, \bar{w}_2, \bar{w}_3, \bar{w}_4$ with $n_1 + m_1 = 6$, $n_2 + m_2 = 10$, $n_3 + m_3 = 3$ and $n_4 + m_4 = 2$.

Figure 6.3: An example showing how to construct $\bar{G}$ from the SSAW G in $T(2,2)$. Note that the edges added to G in order to obtain $\bar{G}$ are shown as dashed lines.

Given an n-edge SSAW G, let $w_1, ...w_k$ be the sequence of USAWs in G. By Lemma 6.2.1, any pair of USAWs (w_i, w_j) $(i < j)$ corresponds to a new pair of USAWs $(\bar{w}_i, \bar{w}_j)$ such that $\bar{w}_i$ and $\bar{w}_j$ are confined to the box B_G and have their endpoints in the boundary planes $y = 0$ and $y = N$. Let a_i (a_j) and b_i (b_j) be the lexicographically smallest and largest vertices of $\bar{w}_i$ $(\bar{w}_j)$ on the boundary respectively. The four points a_i, b_i, a_j and b_j lie in the boundary of B_G and, regardless of the configurations of $\bar{w}_i$ and $\bar{w}_j$, a_i (a_j) can be connected to b_i (b_j) using three closed straight-line segments lying outside B_G; for example, see Figure 6.4. In particular, this can be done so that the result is a two-component link $K_{ij} = (K_i, K_j)$ satisfying the following conditions: First, when K_{ij} is projected into the xy-plane, the six closed straight-line segments of K_{ij} which lie outside B_G create no crossing with the edges inside B_G and at most one crossing amongst themselves. Secondly, when K_{ij} is projected into the xy-plane, the only crossing amongst edges outside B_G (if there is any) must involve both components of K_{ij}. Such a crossing will consist of an over-crossing from one polygon edge and an under-crossing edge from the other polygon. The addition of the six closed straight-line segments is done so that the polygon containing w_i, K_i, contains the over-crossing. Figure 6.4 shows an example of such a link associated to a pair of walks. The conditions under which the projection of K_{ij} into the xy-plane creates (or does not create) a crossing outside B_G are discussed next. Without loss of generality we take $i = 1$ and $j = 2$ and we have $a_1 < a_2 < b_2$. Also, depending on the coordinates of b_1, one of the following cases holds:

1) If $a_1 < b_1 < a_2 < b_2$ then the following argument shows that K_{12} can be constructed such that its regular projection in the xy-plane does not create any crossing outside B_G. For $i = 1, 2$, every a_i (b_i) lies in $y = 0$ or $y = N$. This gives rise to a partition $\{A, B\}$ of the set $\{a_1, b_1, a_2, b_2\}$ satisfying $A \cap B = \emptyset$ and $A \cup B = \{a_1, b_1, a_2, b_2\}$ such that the elements in A lie in one boundary plane and the elements in B lie in the other boundary plane. For all the possible partitions $\{\{a_1, b_1, a_2, b_2\}, \emptyset\}$, $\{\{b_1, a_2, b_2\}, \{a_1\}\}$, $\{\{a_1, a_2, b_2\}, \{b_1\}\}$, $\{\{a_1, b_1, b_2\}, \{a_2\}\}$, $\{\{a_1, b_1, a_2\}, \{b_2\}\}$, $\{\{a_1, b_1\}, \{a_2, b_2\}\}$, $\{\{a_1, a_2\}, \{b_1, b_2\}\}$ and $\{\{a_1, b_2\}, \{b_1, a_2\}\}$, K_{12} can be constructed such that its regular projection in the xy-plane does not create any crossing outside B_G. This is because there is a plane which divides B_G so that a_1 and b_1 are on one side and the other two vertices, a_2 and b_2, are on the other side of the plane; so we can add three closed line segments joining a_1 and b_1 on one side of the plane and another three closed line segments joining a_2 and b_2 on the other side of the plane.

2) If $a_1 < a_2 < b_1 < b_2$ then, depending on the position of the points on the boundary planes, K_{12} can be constructed such that its regular projection in the xy-plane creates zero or one crossing outside B_G. For the partitions $\{\{a_1, a_2, b_1, b_2\}, \emptyset\}$ $(x(b_1) > x(a_2))$, $\{\{a_2, b_1, b_2\}, \{a_1\}\}$ $(x(b_1) > x(a_2))$, $\{\{a_1, a_2, b_1\}, \{b_2\}\}$ $(x(b_1) <$

$x(a_2)$) and $\{\{a_1, b_2\}, \{a_2, b_1\}\}$ ($x(b_1) > x(a_2)$) the extra crossing is needed (e.g. see Figure 6.4). On the other hand, for the partitions $\{\{a_1, a_2, b_1, b_2\}, \emptyset\}$ ($x(b_1) = x(a_2)$), $\{\{a_2, b_1, b_2\}, \{a_1\}\}$ ($x(b_1) = x(a_2)$), $\{\{a_1, a_2, b_1\}, \{b_2\}\}$ ($x(b_1) = x(a_2)$), $\{\{a_1, b_1, b_2\}, \{a_2\}\}$,$\{\{a_1, b_2\}, \{a_2, b_1\}\}$ ($x(b_1) = x(a_2)$), $\{\{a_1, a_2, b_2\}, \{b_1\}\}$, $\{\{a_1, a_2\}, \{b_1, b_2\}\}$ and $\{\{a_1, b_1\}, \{a_2, b_2\}\}$ no crossing is needed.

3) If $a_1 < a_2 < b_2 < b_1$ then, similar to the previous case, for the partitions $\{\{a_1, b_2, b_1\}, \{a_2\}\}$ ($x(b_1) > x(b_2)$), $\{\{a_1, a_2, b_1\}, \{b_2\}\}$ ($x(b_1) > x(a_2)$), $\{\{a_1, a_2\}, \{b_2, b_1\}\}$ ($x(b_1) > x(b_2)$) and $\{\{a_1, b_2\}, \{a_2, b_1\}\}$ ($x(b_1) > x(a_2)$) one crossing is needed while for the partitions $\{\{a_1, b_2, b_1\}, \{a_2\}\}$ ($x(b_1) = x(b_2)$), $\{\{a_1, a_2, b_1\}, \{b_2\}\}$ ($x(b_1) = x(a_2)$), $\{\{a_1, a_2\}, \{b_2, b_1\}\}$ ($x(b_1) = x(b_2)$), $\{\{a_1, b_2\}, \{a_2, b_1\}\}$ ($x(b_1) = x(a_2)$), $\{\{a_1, a_2, b_2, b_1\}, \emptyset\}$, $\{\{a_2, b_2, b_1\}, \{a_1\}\}$, $\{\{a_1, a_2, b_2\}, \{b_1\}\}$ and $\{\{a_1, b_1\}, \{a_2, b_2\}\}$ no crossing is needed.

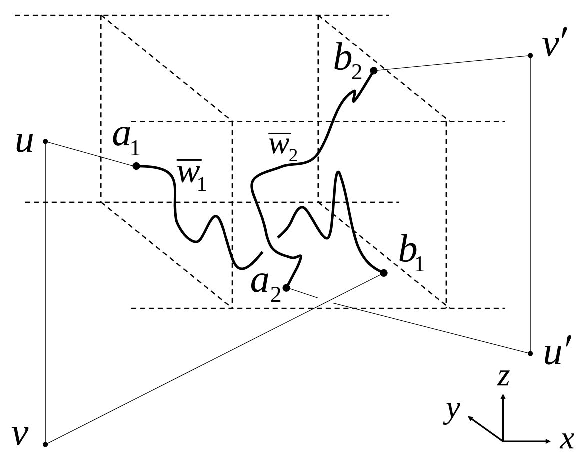


Figure 6.4: A two-component link K_{12} is associated to a pair of USAWs $(\bar{w}_1, \bar{w}_2)$. $a_1 = (x(a_1), N, z(a_1))$, $b_1 = (x(b_1), 0, z(b_1))$, $a_2 = (x(a_2), 0, z(a_2))$ and $b_2 = (x(b_2), N, z(b_2))$. $a_1 < a_2 < b_1 < b_2$ and the corresponding partition is $\{\{a_1, b_2\}, \{a_2, b_1\}\}$ ($x(b_1) > x(a_2)$). Note that in this case we can take $u = (x_1 - x, N + y_u, z_u)$, $v = (x_1 - x, -y_v, z_v)$, $u' = (x_2 + x', -y'_u, z'_u)$ and $v' = (x_2 + x', N + y'_v, z'_v)$ where $x, x', y_u, z_u, y_v, z_v, y'_v, z'_v$ are arbitrary positive integers and two positive integers y'_u and z'_u are chosen so that the line joining a_2 and u' lies under the line joining b_1 and v.

Thus we obtain the following corollary from Lemma 6.2.1:

Corollary 6.2.2. *Given an n-edge SSAW G, let $w_1, ...w_k$ be the sequence of USAWs in G. Let $\bar{G}$ be the SSAW associated to G as prescribed in Lemma 6.2.1. For any $1 \leq i < j \leq k$, there exists a two-component polygonal link $K_{ij} = (K_i, K_j)$ corresponding to each pair of USAWs (w_i, w_j) such that when it is projected into the xy-plane those edges of K_{ij} which lie outside B_G will create no crossing with the edges inside B_G and at most one crossing, involving both polygons K_i and K_j, amongst themselves.*

The *Entanglement Complexity*, $EC(G)$, of any SSAW G is now defined as follows [41]:

$$EC(G) = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} |Lk(K_{ij})|, \tag{6.2.3}$$

where k is the total number of disjoint undirected walks $w_1, w_2, ..., w_k$ contained in G, K_{ij} is the two-component link associated to the pair of USAWs (w_i, w_j) as prescribed in Corollary 6.2.2 and $Lk(K_{ij})$ is the linking number of K_{ij}.

6.3 (CONCAT) and (CAPOFF) for SSAWs

Given an SSAW G, let $f(G, 1)$, $f(G, 2)$ and $f(G, 3)$ represent the span, the number of degree one vertices (or, equivalently, twice the number of disjoint walks) and the total number of edges of G respectively. For $j = 1, 2, 3$, define

$$Z_n^{\#j}(N, M, Y_3) = \sum_G x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}, \tag{6.3.1}$$

where the sum is over all the SSAWs G with "size" n, measured by $f(G, j)$, and left-most plane $x = 0$; $Y_3 = [x, y, z]^T$ is a vector of variables. Note that here "size" of an SSAW G can be measured by the total number of edges ($f(G, 3)$), the span ($f(G, 1)$) or the number of degree one vertices of G ($f(G, 2)$ or, equivalently, twice the number of USAWs in G).

We first use the so called pattern insertion strategy, introduced in Section 2.5, to obtain a pattern theorem for SSAWs. In order to meet the requirements of Corollary

2.5.3, which gives the pattern theorem, in this section we will check the satisfaction of (CA1), (CA2), (CONCAT) and (CAPOFF) for SSAWs. (CA1) basically gives the definition of the clusters under investigation. Here $C_n := C_n(N, M)$ is the set of all SSAWs in $T(N, M)$ and $C_n^* \subseteq C_n$ represents the set of all SSAWs with "size" n in $T(N, M)$ which have their left-most plane at $x = 0$. It can be easily seen that C_n is invariant under translation, i.e., given

$$u \in S^* = \{u \in \mathbb{R}^3 \mid T(N, M) + u = T(N, M)\} = \{(x, 0, 0) \mid x \in \mathbb{Z}\}, \tag{6.3.2}$$

$G \in C_n$ implies that $G+u \in C_n$. The C_n's are also pairwise disjoint since $C_n \cap C_m = \emptyset$ for $n \neq m$. So (CA1) is satisfied for SSAWs in $T(N, M)$. Similar to the case for stretched polygons, given $C := \bigcup_{n=1}^{\infty} C_n$, we consider the weight function

$$wt : C \to (0, \infty) \tag{6.3.3}$$

such that $wt(G) = x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}$ for any $G \in C$, where x, y and z are positive real numbers. $f(G, 1)$, $f(G, 2)$ and $f(G, 3)$ all remain unchanged when an SSAW G is translated along the x-axis thus the weight function wt is invariant under translation, i.e.

$$wt(G) = wt(G + u) \quad \text{for every } u \in S^* \text{ and } G \in C. \tag{6.3.4}$$

The following theorem also shows that (CA2) is satisfied for this choice of weight function.

Theorem 6.3.1. *Given the weight function wt, (CA2) is satisfied for SSAWs.*

Proof. Given $k \geq 0$, let the two SSAWs G and G' differ by at most k vertices and edges, i.e.

$$\mid E(G) \vartriangle E(G') \mid + \mid V(G) \vartriangle V(G') \mid \leq k. \tag{6.3.5}$$

Then we claim that

$$\mid f(G, 1) - f(G', 1) \mid \leq k, \quad \mid f(G, 2) - f(G', 2) \mid \leq k \;\; \textit{and} \quad \mid f(G, 3) - f(G', 3) \mid \leq k. \tag{6.3.6}$$

Suppose to the contrary that $| f(G,1) - f(G',1) |> k$ ($| f(G,2) - f(G',2) |> k$). Without loss of generality assume that $f(G,1) > f(G',1)$, i.e. the span of G is greater than that of G', then there are exactly $f(G,1) - f(G',1)$ hinges in which G has at least one vertex (because G is an SSAW) and G' does not have any vertices thus

$$| V(G) \setminus V(G') |\geq| f(G,j) - f(G',j) |> k \text{ for } j = 1,2. \tag{6.3.7}$$

This implies that

$$| E(G) \vartriangle E(G') | + | V(G) \vartriangle V(G') |\geq| V(G) \setminus V(G') |> k \tag{6.3.8}$$

which contradicts equation 6.3.5. Therefore, the following inequality must hold

$$| f(G,j) - f(G',j) |\leq k \text{ for } j = 1,2. \tag{6.3.9}$$

Similarly, suppose to the contrary that for the number of edges,

$$| f(G,3) - f(G',3) |> k \tag{6.3.10}$$

then

$$| E(G) \setminus E(G') |\geq| f(G,3) - f(G',3) |> k. \tag{6.3.11}$$

This implies that

$$| E(G) \vartriangle E(G') | + | V(G) \vartriangle V(G') |\geq| E(G) \setminus E(G') |> k \tag{6.3.12}$$

which contradicts equation 6.3.5. Hence, the following inequality must hold

$$| f(G,3) - f(G',3) |\leq k. \tag{6.3.13}$$

Equations 6.3.9 and 6.3.13 yield equation 6.3.6. Thus,

$$-k \leq f(G',j) - f(G,j) \leq k \text{ for } j = 1,2,3, \tag{6.3.14}$$

which implies

$$x^{-k} \leq x^{f(G',1)-f(G,1)} \leq x^k, \ \ y^{-k} \leq y^{f(G',2)-f(G,2)} \leq y^k \ \ \textit{and} \ \ z^{-k} \leq z^{f(G',3)-f(G,3)} \leq z^k. \tag{6.3.15}$$

Hence, for $x > 0$, $y > 0$, $z > 0$,

$$(xyz)^{-k} \le x^{f(G',1)-f(G,1)} y^{f(G',2)-f(G,2)} z^{f(G',3)-f(G,3)} \le (xyz)^{k}. \tag{6.3.16}$$

Multiplying all sides of the above inequality by $x^{f(G,1)}y^{f(G,2)}z^{f(G,3)}$ gives

$$(xyz)^{-k} x^{f(G,1)} y^{f(G,2)} z^{f(G,3)} \le x^{f(G',1)} y^{f(G',2)} z^{f(G',3)} \le (xyz)^{k} x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}. \tag{6.3.17}$$

Let $\gamma_k := (xyz)^{-k}$. Having

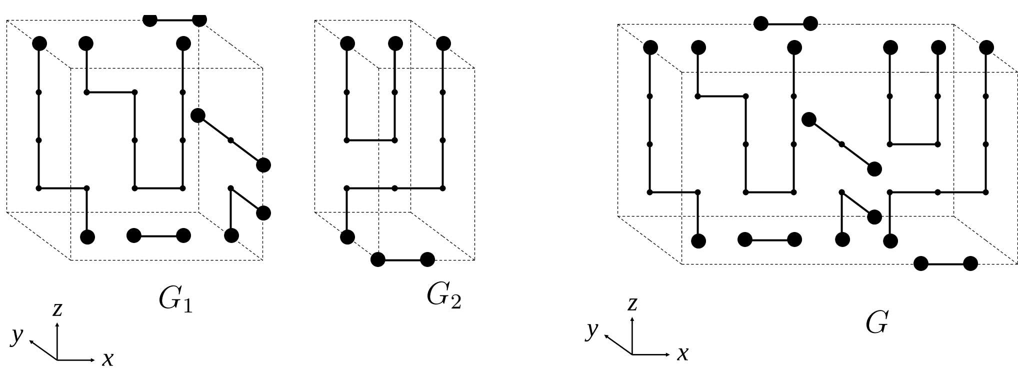
$$wt(G) = x^{f(G,1)} y^{f(G,2)} z^{f(G,3)} \text{ and } wt(G') = x^{f(G',1)} y^{f(G',2)} z^{f(G',3)}, \tag{6.3.18}$$

equation 6.3.17 yields

$$\frac{1}{\gamma_k} wt(G) \le wt(G') \le \gamma_k wt(G). \tag{6.3.19}$$

□

The following lemmas show that for SSAWs and various choices of size there exist choices of c_T, t_T and m_T for which (CONCAT) and (CAPOFF) hold.



(a) The two SSAWs G_1 and G_2 in $T(2,4)$.

(b) The SSAW G resulted from the concatenation of G_1 and G_2.

Figure 6.5: An example of concatenating the 12-edge SSAW G_2 with span 2 ($f(G_2, 2) = 6$) to the 19-edge SSAW G_1 with span 4 ($f(G_1, 2) = 12$) which results in the 31-edge SSAW G with span 7 ($f(G, 2) = 18$).

Lemma 6.3.2. *Given two non-negative integer numbers N and M, $t_T = 1$ and $c_T = 0$ ($c_T = 1$), (CONCAT) holds for SSAWs in $T(N, M)$ when the size of an*

SSAW G is measured either by the number of edges, $f(G,3)$, or by the number of degree one vertices, $f(G,2)$ (by the span, $f(G,1)$).

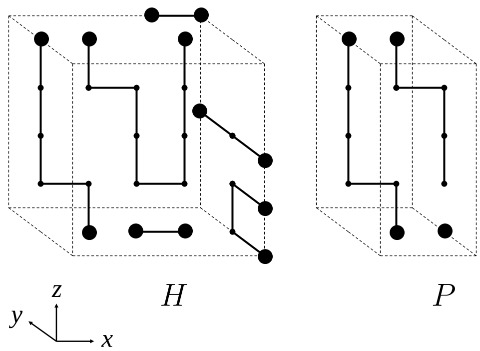
Proof. We prove that there exists a concatenation process defined for SSAWs on $T(N,M)$ and associated integers $t_T \geq 0$ and $c_T \geq 0$ such that:

Given $G_1 \in C_n^*$ with span b_1 and $G_2 \in C_m^*$ with span b_2, concatenating G_1 to the translate of G_2, $G_2+(t_T+b_1,0,0)$, forms $G \in C_{n+m+c_T}^*$ such that $G \cap V_{b_1-1} = G_1 \cap V_{b_1-1}$ and $G \cap (V_{b_2-1} + (t_T + b_1 + 1, 0, 0)) = G_2 \cap (V_{b_2-1} + (1,0,0))$ (i.e. only the right-most plane of G_1 and the left-most plane of G_2 can be altered in the concatenation process).

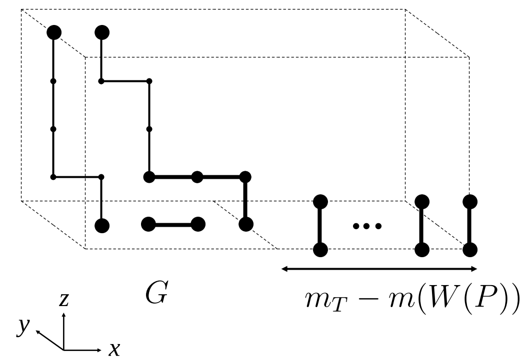
The proof of this lemma when the size of G is measured either by $f(G,2)$ or $f(G,3)$ is identical while slight modifications are required when the size is measured by $f(G,1)$; these modifications are set off in brackets. Let $G_1 \in C_n^*$ with span b_1 and $G_2 \in C_m^*$ with span b_2, *uniting* G_1 with the translate of G_2, $G_2 + (b_1+1,0,0)$, forms the concatenated SSAW G with span $b = b_1 + b_2 + 1$ (i.e. $f(G,1) = f(G_1,1) + f(G_2,1) + 1$) and satisfying $f(G,j) = f(G_1,j) + f(G_2,j)$ for $j = 2,3$. G is in C_{n+m}^* if size is measured by $f(G,3)$ or $f(G,2)$ (in C_{n+m+1}^* if size is measured by $f(G,1)$). An example of such a concatenation is shown in Figure 6.5. It is clear that this concatenation process is well defined since G automatically inherits all the required properties of SSAWs from G_1 and G_2. Letting $t_T = 1$ and $c_T = 0$ if size is measured by $f(G,3)$ or $f(G,2)$ ($c_T = 1$ if size is measured by $f(G,1)$) proves that (CONCAT) holds for SSAWs in $T(N,M)$. □

Note that larger SSAWs can be made by concatenating smaller ones; however, there are many SSAWs which cannot be made by concatenating two or more small ones. For example, SSAWs which contain only a single undirected self-avoiding walk have this property.

Lemma 6.3.3. *Given a pair of non-negative integer numbers $(N,M) \neq (0,0)$ and $m_T \geq (N-1)(M-1)+1$, (CAPOFF) holds for SSAWs in $T(N,M)$ regardless of*



(a) The 2-pattern P which has occurred in an SSAW H.



(b) The SSAW G with span $2+m_T$ which starts with the pattern P and ends with $m_T - m(W(P))$ copies of the 0-pattern P_0.

Figure 6.6: An example of a 2-pattern P which occurs at the start of an SSAW G with span $2 + m_T$; note that $m(W(P)) = 2$ and the edges added to P in order to obtain G are shown to be thicker than P's edges.

whether size is measured by the number of edges, the span or the number of degree one vertices.

Proof. We prove that there exists an integer $m_T > 0$ such that:

For any integer $b \geq 0$ and any b-pattern P that occurs at $(0, 0, 0)$ in some finite size SSAW in C^* with span $s \geq b + 1$ (i.e. P occurs at the start of some SSAW), there exists an SSAW $G \in C^*$ with span $b + m_T$ which also contains P at $(0, 0, 0)$ (i.e. P is also at the start of G). Similarly, given any b-pattern P' that occurs at $(s - b, 0, 0)$ in some finite size SSAW in C^* with span $s \geq b + 1$ (i.e. P' occurs at the end of some SSAW), there exists an SSAW $G' \in C^*$ with span $b + m_T$ which contains P' at $(m_T, 0, 0)$ (i.e. P' also ends G').

Without loss of generality consider the first case, i.e. let P be a b-pattern that occurs at the start of an SSAW with span $> b$. P may already be an SSAW or P must have a degree one vertex in (or a degree zero vertex in the boundary of) its right-most plane. In the latter case, to obtain an SSAW G satisfying (CAPOFF) any degree one (or degree zero) vertex must be extended to the boundary (one unit in the boundary). Thus the SSAW G satisfying (CAPOFF) can be constructed from the pattern P as follows (Figure 6.6 gives an example of such a construction):

Let $W(P) := \{v_1, ..., v_i\}$ be the set of all the degree one vertices (ordered lexicographically from smallest to largest) in the plane $x = b$ which are not on the boundaries of the tube; i.e. $v_j \in X_b := \{(b, y, z) \in \mathbb{Z}^3 | 1 \le y \le N-1, 1 \le z \le M-1\}$ for $j = 1, ..., i$ ($i \le (N-1)(M-1)$). Given $v_j = (x_j, y_j, z_j)$, we join each v_j to a point on the boundary of the tube using a USAW. Specifically, for each $1 \le j \le i$, we add the USAW $((j+1)\hat{\imath}, -z_j\hat{k})$ starting at v_j and then ignore the orientation on the edges. Next, we join any vertex u of P that has degree zero and lies on the boundary of the tube to the vertex $u + (1, 0, 0)$ by adding an edge. Note that such an edge must be on the boundary of the tube. The result will be an SSAW G_1 with span equal to $b + i + 1$. Clearly, the span of G_1 depends on $m(W(P)) := i + 1$ which depends on i, the number of degree one vertices in the configuration of the right-most plane of P, i.e. W. We will remove this dependence using the following strategy.

Let W be any non-empty subset of vertices in X_b. Let $W = \{v_1, ..., v_{i(W)}\}$, where the v_j's, $j = 1, ..., i(W)$, are ordered lexicographically (from smallest to largest) and $i(W)$ is the number of elements in W. By the construction above, corresponding to each W we get a set of USAWs which fit in a section of the tube with span $m(W) := i(W) + 1$, i.e. $V_{m(W)} + (b, 0, 0)$. Since $i(W) \le (N-1)(M-1)$, $m(W)$ will be bounded above by a function of N and M, i.e. $m(W) \le (N-1)(M-1)+1$. In particular, taking $W = W(P)$ from above, $m(W(P)) \le (N-1)(M-1)+1$. Thus we can remove the dependence of the span of G_1 on $W(P)$ by extending G_1 to create a new SSAW G having span at least $(N-1)(M-1)+1+b$. To see this, let m_T be an integer $\ge (N-1)(M-1)+1$. Let P_0 be the 0-pattern containing the single edge walk $(\hat{k})$ (or $(\hat{\jmath})$ if $M = 0$) starting at the origin with orientation ignored. G_1 can easily be extended to an SSAW G with span m_T by concatenating $\big(m_T - m(W(P))\big)$ copies of P_0 to the left-most plane of G_1; for example, see Figure 6.6. Furthermore, by symmetry a similar argument leads to the construction of the SSAW G' satisfying (CAPOFF) for an end pattern P'. Therefore (CAPOFF) is satisfied for SSAWs. $\square$

6.4 The Limiting Free Energy for SSAWs

In this section, we prove the existence of the limiting free energy for SSAWs when the size is measured by various measures. Let $q(N, M, a_1, a_2, a_3)$ denote the number of SSAWs G (up to x-translation) satisfying $f(G, j) = a_j$ for $j = 1, 2, 3$. Note that for any $n \in \mathbb{N}$, $Z_n^{\#j}(N, M, Y_3)$ from equation 6.3.1 can be written as

$$\begin{aligned} Z_n^{\#j}(N, M, Y_3) &= \sum q(N, M, a_1, a_2, a_3) x^{f(G,1)} y^{f(G,2)} z^{f(G,3)} \\ &= \sum q(N, M, a_1, a_2, a_3) x^{a_1} y^{a_2} z^{a_3}, \end{aligned} \tag{6.4.1}$$

where the sums are over all the non-negative integer triples (a_1, a_2, a_3) with $a_j = n$.

Theorem 6.4.1. *For each of $j = 1, 2, 3$, the following limit exists and is finite*

$$\mathcal{F}^{\#j}(N, M, Y_3) = \lim_{n \to \infty} n^{-1} \log Z_n^{\#j}(N, M, Y_3). \tag{6.4.2}$$

Proof. Note first that here we prove the theorem for the case where size of any SSAW G is measured by $f(G, 3)$, i.e. the number of edges in G. A similar argument works for the other two cases.

By (CONCAT) for SSAWs, Lemma 6.3.2, the following inequality holds

$$\sum_{m_1=0}^{2n_1} \sum_{c_1=0}^{2n_1} q(N, M, m_1, c_1, n_1) q(N, M, m - m_1 - 1, c - c_1, n_2) \leq q(N, M, m, c, n_1 + n_2). \tag{6.4.3}$$

Multiplying both sides of this inequality by $x^m y^c z^{n_1+n_2}$ and summing over m and c

gives rise to

$$\sum_{m=0}^{2(n_1+n_2)} \sum_{c=0}^{2(n_1+n_2)} \sum_{m_1=0}^{2n_1} \sum_{c_1=0}^{2n_1} q(N,M,m_1,c_1,n_1) q(N,M,m-m_1-1,c-c_1,n_2) x^m y^c z^{n_1+n_2}$$
$$\leq \sum_{m=0}^{2(n_1+n_2)} \sum_{c=0}^{2(n_1+n_2)} q(N,M,m,c,n_1+n_2) x^m y^c z^{n_1+n_2}$$
$$= Z_{n_1+n_2}(N,M,Y_3).$$

Moreover,

$$Z_{n_1}(N,M,Y_3) = \sum_{m_1=0}^{2n_1} \sum_{c_1=0}^{2n_1} q(N,M,m_1,c_1,n_1) x^{m_1} y^{c_1} z^{n_1}, \tag{6.4.4}$$

$$Z_{n_2}(N,M,Y_3) = \sum_{m=m_1+1}^{m_1+2n_2+1} \sum_{c=c_1}^{c_1+2n_2} q(N,M,m-m_1-1,c-c_1,n_2) x^{m-m_1-1} y^{c-c_1} z^{n_2}. \tag{6.4.5}$$

Now equations 6.4.3 to 6.4.5 imply that

$$Z_{n_1}^{\#3}(N,M,Y_3) Z_{n_2}^{\#3}(N,M,Y_3) \leq x^{-1} Z_{n_1+n_2}^{\#3}(N,M,Y_3). \tag{6.4.6}$$

Letting $a_n = \log x Z_n^{\#3}(N,M,Y_3)$, equation 6.4.6 implies that $\{a_n\}$ is a superadditive sequence. Therefore, by Lemma 2.1.2

$$\mathcal{F}^{\#3}(N,M,Y_3) = \lim_{n\to\infty} n^{-1} a_n = \sup_{n\geq 1} n^{-1} a_n. \tag{6.4.7}$$

Any n-edge SSAW with $x_1 = 0$ is a set of n edges from the edges in the box $\{(x,y,z) \in \mathbb{Z}^3 \mid 0 \leq x \leq 2n, 0 \leq y \leq N, 0 \leq z \leq M\}$. There are in total

$$2n(M+1)(N+1) + M(2n+1)(N+1) + N(M+1)(2n+1) \tag{6.4.8}$$

edges in this box. Hence, given $a = 4[(M+1)(N+1) + M(N+1) + N(M+1)]$,

$$q(N,M,a_1,a_2,n) \leq \binom{2n(M+1)(N+1) + M(2n+1)(N+1) + N(M+1)(2n+1)}{n}$$
$$\leq \binom{an}{n}. \tag{6.4.9}$$

Since, $0 \le a_1 \le 2n$, $0 \le a_2 \le n$ and $a_3 = n$, so there are at most $(2n+1)(n+1)$ choices for the triple (a_1, a_2, a_3). Thus

$$Z_n^{\#3}(N, M, Y_3) \le (2n+1)(n+1)\binom{an}{n}\max\{1, x^{2n}y^n z^n\}. \tag{6.4.10}$$

Therefore, using Lemma 2.1.4, we have

$$\limsup_{n\to\infty} n^{-1}\log Z_n^{\#3}(N, M, Y_3) \le \big[a\log a - (a-1)\log(a-1)\big] + \log(x^2yz) < \infty. \tag{6.4.11}$$

The existence and finiteness of the limit now follows from Lemma 2.1.2. $\square$

Furthermore, for $j = 1, 2, 3$, Theorem 6.4.1 and [26, Section 2.1] lead immediately to the following result for $\mathcal{F}^{\#j}(N, M, Y_3)$.

Theorem 6.4.2 (Janse van Rensburg 2000 [26])**.** *$\mathcal{F}^{\#j}(N, M, Y_3)$ is a convex function of $\log Y_3(j)$. Moreover, its right- and left-derivatives exist everywhere in $(0, \infty)$, and they are non-decreasing functions of $Y_3(j)$. $\mathcal{F}^{\#j}(N, M, Y_3)$ is also differentiable almost everywhere, and whenever*

$$d\mathcal{F}^{\#j}(N, M, Y_3)/dY_3(j) \tag{6.4.12}$$

exists it is given by

$$\lim_{n\to\infty}(dF_n^{\#j}(N, M, Y_3)/dY_3(j)), \tag{6.4.13}$$

where

$$F_n^{\#j}(N, M, Y_3) = n^{-1}\log Z_n^{\#j}(N, M, Y_3). \tag{6.4.14}$$

6.5 Pattern Theorem for the Limiting Free Energy of SSAWs

Given a pattern P for SSAWs, for $j = 1, 2, 3$, define

$$Z_n^{\#j}(N, M, Y_3; < m, P) = \sum_{G \in C_n^*[<m,P]} x^{f(G,1)}y^{f(G,2)}z^{f(G,3)} \tag{6.5.1}$$

to be the partition function for the SSAWs which contain less than m translates of P, when size of an SSAW G, n, is measured by $f(G,j)$. (CONCAT), (CAPOFF), (CA1) and (CA2) are all satisfied for SSAWs hence Corollary 2.5.3 yields the following pattern theorem for the limiting free energy.

Theorem 6.5.1. *Let P be any proper pattern for SSAWs in $T(N,M)$. Then, for $j=1,2,3$, there exists an $\epsilon_P^{\#j}(Y_3)>0$ such that*

$$\mathcal{F}^{\#j}(N,M,Y_3;<\epsilon_P^{\#j}(Y_3),P)<\mathcal{F}^{\#j}(N,M,Y_3), \tag{6.5.2}$$

where

$$\mathcal{F}^{\#j}(N,M,Y_3;<\epsilon_P^{\#j}(Y_3),P)\equiv\limsup_{n\to\infty} n^{-1}\log Z_n^{\#j}(N,M,Y_3;<\lfloor\epsilon_P^{\#j}(Y_3)n\rfloor,P). \tag{6.5.3}$$

Note that here any SSAW pattern is also a proper pattern. In fact, any SSAW start (end) pattern can also occur in the middle of an SSAW. This is due to the nature of (CONCAT) for SSAWs which leaves both the left- and right-most planes of a given pattern P unchanged. More precisely, let P be a start pattern which occurs at the start of an SSAW H. By Lemma 6.3.2 (CONCAT), we can concatenate together any number of copies of H without changing the right-most or left-most plane of H. This shows that P can occur in the middle of infinitely many SSAWs thus P is a proper SSAW pattern. By symmetry, a similar argument also works for end patterns.

6.6 The Connective Constant for SSAWS

Given $Y_3=[1,1,1]^T$, for $j=1,2,3$ let

$$q_n^{\#j}(N,M)\equiv Z_n^{\#j}(N,M,Y_3) \tag{6.6.1}$$

be the number of SSAWs of size n measured by $f(G,j)$. By Theorem 6.4.1, the connective constants exist for SSAWs, i.e.

$$\kappa_q^{\#j}(N,M)\equiv\mathcal{F}(N,M,[1,1,1]^T)=\lim_{n\to\infty} n^{-1}\log q_n^{\#j}(N,M), \tag{6.6.2}$$

for $j = 1, 2, 3$.

As was discussed in Section 2.2, Soteros and Whittington [48, 49] proved the existence of the connective constant for SAWs in $T(N, M)$, i.e.

$$\kappa(N, M) = \lim_{n\to\infty} n^{-1} \log c_n(N, M) < \infty, \tag{6.6.3}$$

where n denotes the number edges. In this section we show, using the pattern theorem for SSAWs in $T(N, M)$, that the connective constant for SAWs in $T(N - 2, M - 2)$ is strictly less than the connective constant for SSAWs in $T(N, M)$ when the size is measured by the number of edges.

Theorem 6.6.1. *For $N \geq 2$ and $M \geq 2$,*

$$\kappa(N - 2, M - 2) < \kappa_q^{\#3}(N, M). \tag{6.6.4}$$

Proof. The idea of the proof is to start with a bridge w in $T(N - 2, M - 2)$ and construct an SSAW in $T(N, M)$ by translating w and then extending its end vertices to the boundary plane $z = 0$ of the tube $T(N, M)$. See Figure 6.7 for an example of such a construction. Then we use the pattern theorem for SSAWs, Theorem 6.5.1, combined with the fact that bridges and SAWs in $T(N, M)$ have the same connective constant (Theorem 2.2.3) to complete the proof.

Given an n-step bridge $w := u_0, u_1, ..., u_{n-1}, u_n$ in $T(N-2, M-2)$, let w_1 and w_2 be respectively the walk $(-(z(u_0)+1)\hat{k})$ starting at u_0 and and the walk $(\hat{\imath}, -(z(u_n)+1)\hat{k})$ starting at u_n. Let $w_1 \circ w \circ w_2$ be the walk obtained from concatenation of w_1, w and w_2 (see Figure 6.7 (b)). Translating $w_1 \circ w \circ w_2$ by $\hat{k}$ and then by $\hat{\jmath}$ and ignoring the orientation on the edges results in an SSAW G' in $T(N, M)$ which contains only a single USAW. The number of edges in G', $n(G')$, satisfies the following inequality

$$\begin{aligned} n(G') &= n + n(w_1) + n(w_2) \\ &\leq n + (M - 1) + (M - 1) + 1 \\ &= n + 2M - 1, \end{aligned} \tag{6.6.5}$$

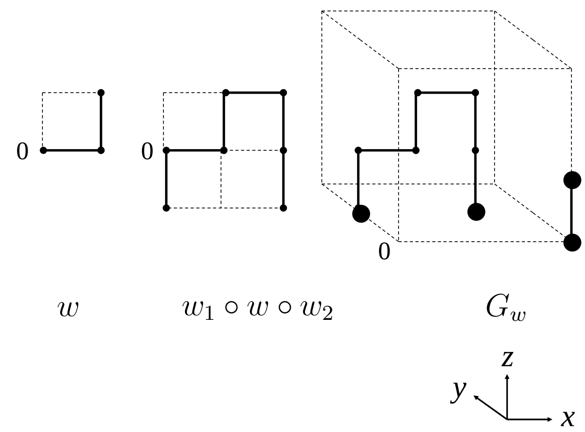


Figure 6.7: An example showing how to construct the SSAW G_w in $T(2,3)$ from the bridge w in $T(0,1)$. Note that here $n = 2$, $m_w = 4$ and $n(G_w) = 7$.

where $n(w_1)$ and $n(w_2)$ denote the number of edges in w_1 and w_2 respectively.

Note that $m_w := n(w_1) + n(w_2) \leq 2M - 1$. Let P_0 be the pattern introduced in the proof of Lemma 6.3.3. We can easily concatenate $(2M - 1 - m_w)$ copies of P_0 to G''s right-most plane and obtain a new SSAW G_w with the total number of edges $n(G_w) = n + 2M - 1$ (see Figure 6.7 (c)). Thus

$$\frac{b_n(N-2, M-2)}{2} \leq q^{\#3}_{n+2M-1}(N, M). \tag{6.6.6}$$

Taking logarithms, multiplying both sides by n^{-1} and letting $n \to \infty$ implies

$$\kappa_b(N-2, M-2) \leq \kappa^{\#3}_q(N, M). \tag{6.6.7}$$

Using Theorem 2.2.3, we obtain

$$\kappa(N-2, M-2) \leq \kappa^{\#3}_q(N, M). \tag{6.6.8}$$

It is easy to see from the construction of the SSAW G_w that there are many SSAW patterns that do not appear in any SSAW in $T(N, M)$ constructed, as explained above, from a USAW in $T(N-2, M-2)$. For example, each 1-span SSAW containing only one undirected walk, considered as a pattern P, cannot occur in any

G_w constructed from a walk w. This is because every SSAW (considered as a pattern) that occurs in G_w and contains only one walk has either span zero (like P_0) or has span greater than one. In particular, let P_* be the SSAW 1-pattern which contains a single edge joining the origin to $(1,0,0)$. We can use $P = P_*$ and the pattern theorem for SSAWs, Theorem 6.5.1, to show that for some $\epsilon_P \equiv \epsilon_P^{\#3}([1,1,1]^T) > 0$

$$\kappa(N-2, M-2) \le \kappa_q^{\#3}(N, M; < \epsilon_P, P) < \kappa_q^{\#3}(N, M), \tag{6.6.9}$$

where

$$\kappa_q^{\#3}(N, M; < \epsilon_P, P) \equiv \mathcal{F}^{\#3}(N, M, [1,1,1]^T; < \epsilon_P, P). \tag{6.6.10}$$

□

6.7 Entanglement Complexity and the Asymptotic Behaviour of SSAWs

In this section, we first use the pattern theorem for SSAWs to find a lower bound (with probability one) for the entanglement complexity of SSAWs. We prove that EC of SSAWs with size n, as defined in equation 6.2.3, asymptotically grows (with probability one) at least as fast as γn for some $\gamma > 0$.

For each $j = 1, 2, 3$, let

$$Z_n^{\#j}(N, M, Y_3; EC \le m) \equiv \sum_G x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}, \tag{6.7.1}$$

where the sum is over all the SSAWs in C_n^* with $n = f(G, j)$ and $EC(G) \le m$. We will show that there exists a positive number $\gamma^{\#j}(Y_3)$ such that all but exponentially few SSAWs with size $n \ge N^{\#j}(Y_3)$ (for some $N^{\#j}(Y_3) \in \mathbb{N}$) have EC greater than $\gamma^{\#j}(Y_3)n$. Let P be any SSAW containing two USAWs w_1 and w_2 so that the link K_{12} associated to this pair of walks has a non-zero linking number. In particular, for $N \ge 4$ and $M \ge 3$, we can take $P = P_1$ to be the 2-pattern in $T(N, M)$ that contains the following two walks with the orientation on the edges ignored:

$((N-3)\hat{j}, -2\hat{\imath}, 2\hat{j}, -\hat{k}, 2\hat{\imath}, \hat{k}, \hat{j})$ starting at $(2,0,2)$, and $(N-2)\hat{j}, \hat{k}, 2\hat{j}$ starting at $(1,0,1)$. The projection of P_1 into the xy-plane is illustrated in Figure 6.8.

Theorem 6.7.1. *For $j = 1,2,3$, there exists a positive number $\gamma^{\#j}(Y_3)$ such that*

$$\mathcal{F}^{\#j}(N, M, Y_3; EC \leq \gamma^{\#j}(Y_3)) < \mathcal{F}^{\#j}(N, M, Y_3), \tag{6.7.2}$$

where

$$\mathcal{F}^{\#j}(N, M, Y_3; EC \leq \gamma^{\#j}(Y_3)) = \limsup_{n\to\infty} n^{-1} \log Z_n^{\#j}(N, M, Y_3; EC \leq \gamma^{\#j}(Y_3)n). \tag{6.7.3}$$

Proof. For the pattern $P = P_1$, introduced above, by Theorem 6.5.1 there exists a real number $\epsilon_P^{\#j}(Y_3) > 0$ such that

$$\mathcal{F}^{\#j}(N, M, Y_3; \leq \epsilon_P^{\#j}(Y_3), P) < \mathcal{F}^{\#j}(N, M, Y_3). \tag{6.7.4}$$

Note that the pattern P can also be considered as an SSAW. P contains the two USAWs w_1 and w_2 such that $\mid Lk(K_{12}) \mid = 1$, where K_{12} is the two-component link associated to (w_1, w_2) as prescribed in Corollary 6.2.2. So each occurrence of P in an SSAW G adds one to EC(G). Thus P cannot appear more than $\epsilon_P^{\#j}(Y_3)n$ times in any SSAW with size n which has EC at most $\epsilon_P^{\#j}(Y_3)n$. Hence for $\gamma^{\#j}(Y_3) := \epsilon_P^{\#j}(Y_3)$ we have

$$\mathcal{F}^{\#j}(N, M, Y_3; EC \leq \gamma^{\#j}(Y_3)) \leq \mathcal{F}^{\#j}(N, M; \leq \epsilon_P^{\#j}(Y_3), P). \tag{6.7.5}$$

Therefore, combining equations 6.7.4 and 6.7.5, we have

$$\mathcal{F}^{\#j}(N, M, Y_3; EC \leq \gamma^{\#j}(Y_3)) < \mathcal{F}^{\#j}(N, M, Y_3). \tag{6.7.6}$$

□

For $j = 1,2,3$, let $X_n^{\#j}(Y_3)$ be a random variable taking its values from the set of SSAWs with size $n = f(G,j)$ (in C_n^*) with the probability distribution

$$\mathbb{P}(X_n^{\#j}(Y_3) = G) = \frac{x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}}{Z_n^{\#j}(N, M, Y_3)}. \tag{6.7.7}$$

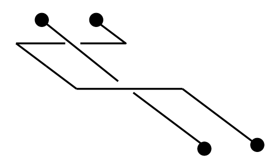


Figure 6.8: The regular projection into the xy-plane of the pattern P_1 with linking number 1.

Corollary 6.7.2. *For $\gamma^{\#j}(Y_3)$ $(j = 1, 2, 3)$ introduced in Theorem 6.7.1, the probability that an SSAW G with size $n = f(G, j)$ has $EC(G)$ greater than $\gamma^{\#j}(Y_3)n$ approaches one exponentially as n goes to infinity, i.e.*

$$\lim_{n\to\infty} \mathbb{P}(EC(X_n^{\#j}(Y_3)) > \gamma^{\#j}(Y_3)n) = \lim_{n\to\infty}[1 - e^{-\alpha'(j,Y_3)n+o(n)}] = 1, \tag{6.7.8}$$

where $\alpha'(j, Y_3) \equiv \mathcal{F}^{\#j}(N, M, Y_3) - \mathcal{F}^{\#j}(N, M, Y_3; \leq \gamma^{\#j}(Y_3)) > 0$.

Proof. Let

$$\begin{aligned} g(n, Y_3) &= \big(\log Z_n^{\#j}(N, M, Y_3; EC \leq \gamma^{\#j}(Y_3)n) - \mathcal{F}^{\#j}(N, M, Y_3; \leq \gamma^{\#j}(Y_3))n\big) \\ &- \big(\log Z_n^{\#j}(N, M, Y_3) - \mathcal{F}^{\#j}(N, M, Y_3)n\big). \end{aligned} \tag{6.7.9}$$

By Theorem 6.4.1,

$$\lim_{n\to\infty} \frac{g(n, Y_3)}{n} = 0 \tag{6.7.10}$$

hence $g(n, Y_3) = o(n)$. So

$$\log \frac{Z_n^{\#j}(N, M, Y_3; EC \leq \gamma^{\#j}(Y_3)n))}{Z_n^{\#j}(N, M, Y_3)} = \big(\mathcal{F}^{\#j}(N, M, Y_3; \leq \gamma^{\#j}(Y_3)) - \mathcal{F}^{\#j}(N, M, Y_3)\big)n + o(n) \tag{6.7.11}$$

which leads to

$$\frac{Z_n^{\#j}(N, M, Y_3; EC \leq \gamma^{\#j}(Y_3)n))}{Z_n^{\#j}(N, M, Y_3)} = e^{-\alpha'(j,Y_3)n+o(n)}. \tag{6.7.12}$$

Therefore, for $j = 1, 2, 3$, the probability that an SSAW with size n has EC greater than $\gamma^{\#j}(Y_3)n$ is given by:

$$\begin{aligned}\mathbb{P}(EC(X_n^{\#j}(Y_3)) > \gamma^{\#j}(Y_3)n) &= \frac{Z_n^{\#j}(N, M, Y_3; EC > \gamma^{\#j}(Y_3)n)}{Z_n^{\#j}(N, M, Y_3)} \\ &= \frac{Z_n^{\#j}(N, M, Y_3) - Z_n^{\#j}(N, M, Y_3; EC \leq \gamma^{\#j}(Y_3)n)}{Z_n^{\#j}(N, M, Y_3)} \\ &= 1 - \frac{Z_n^{\#j}(N, M, Y_3; EC \leq \gamma^{\#j}(Y_3)n)}{Z_n^{\#j}(N, M, Y_3)} \\ &= 1 - e^{-\alpha'(j, Y_3)n + o(n)}. \end{aligned} \tag{6.7.13}$$

By Theorem 6.7.1, $\alpha'(j, Y_3) > 0$ hence the above probability goes to unity as $1 - e^{-\alpha'(j,Y_3)n+o(n)}$ when $n \to \infty$. $\square$

For $k, j \in \{1, 2, 3\}$ and $k \neq j$, $f(X_n^{\#j}(Y_3), k)$ is a random variable with the probability distribution

$$\begin{aligned}\mathbb{P}\big(f(X_n^{\#j}(Y_3), k) = a\big) &= \frac{\sum_{G: f(G,k)=a, f(G,j)=n} x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}}{Z_n^{\#j}(N, M, Y_3)} \\ &= \frac{\sum_{(a_1,a_2,a_3), a_k=a, a_j=n} q(N, M, a_1, a_2, a_3) x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}}{Z_n^{\#j}(N, M, Y_3)}. \end{aligned} \tag{6.7.14}$$

Given $\gamma > 0$, let

$$Z_n^{\#j}(N, M, Y_3; f(k) \leq \gamma n) = \sum_G x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}, \tag{6.7.15}$$

where the sum is over all the SSAWs $G \in C_n^*$ with $n = f(G, j)$ and $f(G, k) \leq \gamma n$.

Theorem 6.7.3. *Given any pair j and k such that $k, j \in \{1, 2, 3\}$ and $k \neq j$, the following inequality is satisfied*

$$\mathcal{F}^{\#j}(N, M, Y_3; f(k) \leq \epsilon_P^{\#j}(Y_3)) < \mathcal{F}^{\#j}(N, M, Y_3), \tag{6.7.16}$$

where

$$\mathcal{F}^{\#j}(N, M, Y_3; f(k) \leq \epsilon_P^{\#j}(Y_3)) = \limsup_{n\to\infty} n^{-1} \log Z_n^{\#j}(N, M, Y_3; f(k) \leq \epsilon_P^{\#j}(Y_3)n). \tag{6.7.17}$$

Proof. For the pattern (SSAW) $P = P_1$ and $\epsilon_P^{\#j}(Y_3) > 0$ introduced in Theorem 6.7.1, we have

$$\mathcal{F}^{\#j}(N, M, Y_3; \leq \epsilon_P^{\#j}(Y_3), P) < \mathcal{F}^{\#j}(N, M, Y_3). \tag{6.7.18}$$

P has span 2, $2N+8 > 1$ edges and 4 degree one vertices on the tube's boundary hence each occurrence of P in an SSAW G increases $f(P,1)$, $f(P,3)$ and $f(P,2)$ at least by one. Therefore, the number of occurrences of P in an SSAW with size n is at most $f(G,i)$ for $i = 1,2,3$. Thus P cannot appear more than $\epsilon_P^{\#j}n$ times in any SSAW with size n satisfying $f(G,k) \leq \epsilon_P^{\#j}(Y_3)n$. Hence we have

$$\mathcal{F}^{\#j}(N, M, Y_3; f(k) \leq \epsilon_P^{\#j}(Y_3)) \leq \mathcal{F}^{\#j}(N, M, Y_3; \leq \epsilon_P^{\#j}(Y_3), P). \tag{6.7.19}$$

Combining equations 6.7.18 and 6.7.19, we have

$$\mathcal{F}^{\#j}(N, M, Y_3; f(k) \leq \epsilon_P^{\#j}(Y_3)) < \mathcal{F}^{\#j}(N, M, Y_3). \tag{6.7.20}$$

□

Therefore, a similar argument to Corollary 6.7.2, implies the following.

Corollary 6.7.4. *For $k, j \in \{1,2,3\}$ and $k \neq j$, the probability that an SSAW G with size n satisfies $f(G,k) > \epsilon_P^{\#j}(Y_3)n$ approaches one exponentially as n goes to infinity, i.e.*

$$\lim_{n\to\infty} \mathbb{P}\big(f(X_n^{\#j}(Y_3), k) > \epsilon_P^{\#j}(Y_3)n\big) = \lim_{n\to\infty}[1 - e^{-\alpha''(j,k,Y_3)n+o(n)}] = 1, \tag{6.7.21}$$

where $\alpha''(j,k,Y_3) \equiv \mathcal{F}^{\#j}(N, M, Y_3) - \mathcal{F}^{\#j}(N, M, Y_3; f(k) \leq \epsilon_P^{\#j}(Y_3)) > 0$.

6.8 An Upper Bound For The Entanglement Complexity of SSAWs

In this section we find an upper bound for the entanglement complexity of SSAWs with size n that is linear in n. Given an SSAW G, we first find an upper bound for

$EC(G)$ when the size of G is measured by the number of edges, then we will use this to obtain an upper bound for the case where the size of G is measured by its span.

Given an SSAW G, for $\delta = 1/6$ and for any pair of USAWs (w_i, w_j) in G, let K_{ij} be the two-component link associated to (w_i, w_j). We fix a polygonal link L_{ij} and its regular diagram in the (x, y)-plane, D_{ij}, as prescribed by Corollary 3.1.4. In order to get an upper bound for EC(G), the sum of the linking number of the K_{ij}'s, we will find an upper bound for the total number of crossings in the D_{ij}'s. There are two types of crossings in each D_{ij}; the crossings created by the edges of K_{ij} which are outside the box B_G and the crossings created by the edges inside B_G. Our strategy is to bound the total sum of each type of crossing by a linear function of n. The following lemma gives an upper bound on the total number of crossings in the D_{ij}'s that result from those edges of K_{ij}'s which lie outside B_G.

Lemma 6.8.1. *Let $w_1, ..., w_k$ be the sequence of USAWs in a given n-edge SSAW G in $T(N, M)$. Let $\bar{G}$ and $\bar{w}_i$ (for $1 \leq i \leq k$) be the SSAW and the USAW obtained from G as prescribed by the proof of Lemma 6.2.1. Let K_{ij}, $1 \leq i < j \leq k$, be the two-component link associated to the pair of USAWs (w_i, w_j) and L_{ij} be the polygonal link prescribed by Corollary 3.1.4. For fixed integers N and M, the total number of crossings made by those edges of L_{ij} which are outside B_G is bounded above by the linear function $f(n) = (N + 1)(M + 2N + 1)(4N - 2)n$.*

Proof. Let G be an n-edge SSAW containing the USAWs $w_1, ..., w_k$. Let $\bar{G}$ be the SSAW associated to G and $\bar{w}_1, ..., \bar{w}_k$ be the corresponding sequence of USAWs as prescribed in the proof of Lemma 6.2.1. Let n_i and $(n_i + m_i)$ $(m_i \leq 4(N - 1))$ denote respectively the number of edges in w_i and $\bar{w}_i$. For $1 \leq i \leq k$, let a_i and b_i represent the x-coordinates of the two vertices of $\bar{w}_i$ with degree one $(a_i \leq b_i)$. For $j \neq i$ and $1 \leq j \leq k$, L_{ij} will not create any crossing outside the box B_G if both a_j and b_j are strictly less than a_i or if both a_j and b_j are strictly greater than b_i. Now any walk which does not satisfy this must have at least one vertex in the set $\{(x, y, z) \in \mathbb{Z}^3 | a_i \leq x \leq b_i, 0 \leq y \leq N, 0 \leq z \leq M + 2N\}$. There are only

$(N+1)(M+2N+1)(b_i - a_i + 1)$ vertices in this set so there can be at most that many L_{ij}'s which will possibly result in a crossing outside B_G. Therefore, letting $f(n) = (N+1)(M+2N+1)(4N-2)n$, the total number of crossings, c, made by the L_{ij}'s outside B_G satisfies the following inequality:

$$
\begin{aligned}
c &\leq \sum_{i=1}^{k}(N+1)(M+2N+1)(b_i - a_i + 1) \\
&\leq (N+1)(M+2N+1)\sum_{i=1}^{k}(n_i + m_i + 1) \\
&\leq (N+1)(M+2N+1)\sum_{i=1}^{k}(n_i + 4(N-1) + 1) \\
&\leq (N+1)(M+2N+1)(n + 4(N-1)n + n) \\
&= (N+1)(M+2N+1)(4N-2)n \\
&= f(n). \qquad (6.8.1)
\end{aligned}
$$

□

Theorem 6.8.2. *Given non-negative integer numbers N and M, the entanglement complexity, $EC(G)$, of any n-edge SSAW G in $T(N,M)$ is bounded above by the linear function*

$$g(n) = a(N,M)n, \qquad (6.8.2)$$

where

$$a(N,M) = \lceil (1/2)((N+1)(M+2N+1)(4N-2)) + (9/4)(M+2N+1)(4N-3) \rceil. \qquad (6.8.3)$$

Proof. Let G be an n-edge SSAW composed of the USAWs $w_1, ..., w_k$, for some $0 \leq k \leq n$. By Corollary 3.1.4, corresponding to each two-component link K_{ij} associated to the pair of USAWs (w_i, w_j) $(i < j)$, there is a polygonal link L_{ij} with regular projection D_{ij} in the (x,y)-plane. By Lemma 6.8.1, we know that the total contribution of the edges of L_{ij}'s which are outside B_G to the total number of crossings is at most $f(n) = (N+1)(M+2N+1)(4N-2)n$. So let us investigate the contribution of the edges of L_{ij} that are inside B_G.

We will provide an upper bound for the total number of crossings of the D_{ij}'s by estimating how many crossings each edge l_r in some L_{ij}, say $L_{i_1j_1}$, can generate. To do so we need to consider two cases depending on whether the edge l_r in $L_{i_1j_1}$ corresponds to a horizontal or a vertical edge l_r^* in $K_{i_1j_1}$. If we assume that l_r^* is horizontal with vertices (x_{r_1}, y_{r_1}, z_r) and (x_{r_2}, y_{r_2}, z_r) then, by Corollary 3.1.4, the only other horizontal edges of the K_{ij}'s that could generate edges in the L_{ij}'s that subsequently generate crossings with l_r in the D_{ij}'s, are the edges that have at least one end point incident on either (x_{r_1}, y_{r_1}, z) or (x_{r_2}, y_{r_2}, z) for $0 < z < M+2N$. There are at most $7(M+2N+1)$ of these. The only vertical edges of the K_{ij}'s that can generate crossings with l_r in the D_{ij}'s are the edges that have both endpoints with (x, y)-coordinates (x_{r_1}, y_{r_1}) or (x_{r_2}, y_{r_2}). There are $2(M+2N)$ of these. Combining these we get that a horizontal edge generates fewer than $9(M+2N+1)$ crossings. Now we assume that l_r is vertical with vertices (x_r, y_r, z_r) and (x_r, y_r, z_r+1). Similar to the argument before we can argue that there are at most $4(M+2N+1)$ horizontal edges (those with a vertex with coordinates (x_r, y_r, z) for some $0 < z < M+2N$) and $M+2N$ vertical edges that can generate crossings with l_r. Combining the two cases gives us the upper bound $(9/2)(M+2N+1)(n_{ij}+m_{ij})$ ($n_{ij}+m_{ij}$ is the number of edges of K_{ij} inside B_G; $m_{ij} := m_i + m_j$ is the sum of the difference between the number of edges w_i and $\bar{w}_i$ and w_j and $\bar{w}_j$; note that $m_{ij} \le 8(N-1)$) on the contribution to the number of crossings by those edges of K_{ij} inside B_G. (It is $(9/2)(M+2N+1)(n_{ij}+m_{ij})$ since each crossing will be counted twice, i.e. if an edge l_r crosses an edge l_s then we counted the crossing both in the upper bound for l_r and in the upper bound for l_s.) The contribution to the linking number by those edges is now bounded by $(9/4)(M+2N+1)(n_{ij}+m_{ij})$.

For $1 \leq i \leq k$ and $1 \leq j \leq n_i + m_i$, let l_{ij} denote the contribution of the jth edge of the undirected walk $\bar{w}_i$ to $EC(G)$, i.e. to the sum of the linking number of the K_{ij}'s. We have proved that

$$l_{ij} \leq (9/4)(M + 2N + 1). \tag{6.8.4}$$

Therefore, letting $g(n) = (1/2)(M + 2N + 1)[(N + 1)(4N - 2) + (9/2)(4N - 3)]n$,

$$\begin{aligned}
EC(G) &\leq (1/2)f(n) + \sum_{i=1}^{k} \sum_{j=1}^{(n_i+m_i)} l_{ij} \\
&\leq (1/2)f(n) + \sum_{i=1}^{k} \sum_{j=1}^{(n_i+m_i)} (9/4)(M + 2N + 1) \\
&\leq (1/2)(N + 1)(M + 2N + 1)(4N - 2)n + (9/4)(M + 2N + 1) \sum_{i=1}^{k} \sum_{j=1}^{n_i+m_i} 1 \\
&\leq (1/2)(N + 1)(M + 2N + 1)(4N - 2)n \\
&+ (9/4)(M + 2N + 1)(n + 4(N - 1)n) \\
&= (1/2)(M + 2N + 1)[(N + 1)(4N - 2) + (9/2)(4N - 3)]n \\
&= a(N, M)n \\
&= g(n).
\end{aligned} \tag{6.8.5}$$

□

Remark 6.8.1. The following example shows that, given $N \geq 4$ and $M \geq 3$, for infinitely many values of n there exists an n-edge SSAW G with $EC(G)$ linear in n. Let G be an $n := (2N + 7)k'$-edge SSAW made of concatenating k' copies of the $(2N+7)$-edge pattern (SSAW) P_1 illustrated in Figure 6.8. Clearly, here, $EC(G) = k'$ is linear in n.

Note that the total number of edges in a slice with width 1 of an (N, M)-tube is

$$a'(N, M) = 2N(M + 1) + 2M(N + 1) + (M + 1)(N + 1), \tag{6.8.6}$$

which depends only on N and M. Hence the number of edges in any SSAW G, $f(G, 3)$, satisfies

$$f(G, 3) \leq a'(N, M)f(G, 1). \tag{6.8.7}$$

Thus Theorem 6.8.2 implies

$$EC(G) \leq a'(N,M)a(N,M)n \tag{6.8.8}$$

when the size of G is measured by its span, i.e. $n = f(G,1)$. Therefore, by Theorem 6.7.2 we conclude that:

- There exists $\gamma^{\#1}(Y_3) > 0$ such that the probability that EC of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ is bounded below by $\gamma^{\#1}(Y_3)n$ and bounded above by $a'(N,M)a(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}(\gamma^{\#1}(Y_3)n < EC(X_n^{\#1}(Y_3)) \leq a'(N,M)a(N,M)n) = \lim_{n\to\infty}[1 - e^{-\alpha'(1,Y_3)n+o(n)}] = 1. \tag{6.8.9}$$

  In other words, as the span of SSAWs increases one expects EC to grow linearly in the span with probability one.

- There exists $\gamma^{\#3}(Y_3) > 0$ such that the probability that EC of a randomly chosen SSAW $X_n^{\#3}(Y_3)$ with n edges is bounded below by $\gamma^{\#3}(Y_3)n$ and bounded above by $a(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}(\gamma^{\#3}(Y_3)n < EC(X_n^{\#3}(Y_3)) \leq a(N,M)n) = \lim_{n\to\infty}[1 - e^{-\alpha'(3,Y_3)n+o(n)}] = 1. \tag{6.8.10}$$

  In other words, as the number of edges of SSAWs increases one expects EC to grow linearly in the number of edges with probability one.

Furthermore, $f(G,2) \leq 2f(G,3) \leq a'(N,M)f(G,1)$ and $f(G,1) \leq 2f(G,3)$. So by Theorem 6.7.4 we conclude that:

- There exists $\epsilon_P^{\#1}(Y_3) > 0$ such that the probability that the number of degree one vertices of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ is bounded below by $\epsilon_P^{\#1}(Y_3)n$ and bounded above by $a'(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#1}(Y_3)n \leq f(X_n^{\#1}(Y_3),2) \leq a'(N,M)n\big) = \lim_{n\to\infty}[1-e^{-\alpha''(1,2,Y_3)n+o(n)}] = 1. \tag{6.8.11}$$

In other words, as the span of SSAWs increases one expects the number of degree one vertices (the number of disjoint walks) to grow linearly in the span with probability one.

- There exists $\epsilon_P^{\#1}(Y_3) > 0$ such that the probability that the number of edges of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ is bounded below by $\epsilon_P^{\#1}(Y_3)n$ and bounded above by $a'(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#1}(Y_3)n \leq f(X_n^{\#1}(Y_3),3) \leq a'(N,M)n\big) = \lim_{n\to\infty}[1-e^{-\alpha''(1,3,Y_3)n+o(n)}] = 1. \tag{6.8.12}$$

In other words, as the span of SSAWs increases one expects the number of edges to grow linearly in the span with probability one.

- There exists $\epsilon_P^{\#3}(Y_3) > 0$ such that the probability that span of a randomly chosen SSAW $X_n^{\#3}(Y_3)$ with n edges is bounded below by $\epsilon_P^{\#3}(Y_3)n$ and bounded above by $2n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#3}(Y_3)n \leq f(X_n^{\#3}(Y_3),1) \leq 2n\big) = \lim_{n\to\infty}[1-e^{-\alpha''(3,1,Y_3)n+o(n)}] = 1. \tag{6.8.13}$$

In other words, as the number of edges of SSAWs increases one expects the span to grow linearly in the number of edges with probability one.

- There exists $\epsilon_P^{\#3}(Y_3) > 0$ such that the probability that the number of degree one vertices of a randomly chosen SSAW $X_n^{\#3}(Y_3)$ with n edges is bounded below by $\epsilon_P^{\#3}(Y_3)n$ and bounded above by $2n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#3}(Y_3)n \leq f(X_n^{\#3}(Y_3),2) \leq 2n\big) = \lim_{n\to\infty}[1-e^{-\alpha''(3,2,Y_3)n+o(n)}] = 1. \tag{6.8.14}$$

In other words, as the number of edges of SSAWs increases one expects the number of degree one vertices (the number of disjoint walks) to grow linearly in the number of edges with probability one.

6.9 Transfer-Matrix Results for the SSAW Model

In this section, the transfer matrix argument, introduced in Section 2.6, is employed to explore the asymptotic form of SSAWs with a fixed span m, i.e. $Z_m^{\#1}(N, M, Y_3)$ where $Y_3 = [1, y, z]$.

In order to ensure that the Cluster Configuration Axiom, as stated in Section 2.6, is satisfied, we first define an ordering and labelling on the edges of a given SSAW G. Let G be an SSAW composed of the sequence of USAWs $w_1, w_2, ..., w_k$, where k is the total number of walks in G. We define a labelling on the edges of G such that, for $1 \leq i \leq k$, all the edges in w_i have the same label, and edges belonging to different walks have distinct labels. As discussed before, the walks are ordered according to the following algorithm: Among the two endpoints (vertices of degree one) of each USAW w_i, for $1 \leq i \leq k$, let v_i be the vertex which is lexicographically smallest. Then the walks are ordered from $i = 1, ..., k$ according to the lexicographical order of the v_i's. An ordering can also be defined on the edges of each w_i as follows: We can assume a direction on the edge incident on v_i, $e_i = \{v_i, u_i\}$, by directing the edge to go from v_i to u_i. This will naturally induce an ordering on the edges of w_i starting with edge e_i as the first edge and ordering the other edges following the direction induced by that of e_i. The ordering on the w_i's and their edges will naturally induce an ordering on the edges of G starting with edge e_1 as the first edge and ordering the other edges of G according to the ordering of the w_i's and their edges. Hence a k-cluster config (k-SSAW config) of an SSAW G is considered as G's configuration in a sublattice of the form $H_{i-1}(N, M) \cup S_i(N, M) \cup H_i(N, M) \cup ... \cup S_{i+k-1}(N, M) \cup H_{i+k-1}(N, M)$ for some $1 \leq i \leq m - k + 1$, where G's configuration in such a sublattice of the tube consists of the sublattice and the ordering as well as labelling of the edges of G, as introduced above. With this extra information on the edges, the Cluster Configuration Axiom, given in Section 2.6, is satisfied for the following sets of k-SSAW configs:

$$\Pi(k) = \{P_1, P_2, ..., P_{|\Pi(k)|}\}, \tag{6.9.1}$$

$$\Pi_1(k) = \{P'_1, P'_2, ..., P'_{|\Pi_1(k)|}\} \tag{6.9.2}$$

and

$$\Pi_2(k) = \{P_1'', P_2'', ..., P''_{|\Pi_2(k)|}\} \tag{6.9.3}$$

which are the set of all distinct proper, start and end k-SSAW configs, respectively. Figure 6.9 shows an example of an SSAW in $T(2,2)$ with it's associated 3-SSAW configs. Note that in Figure 6.9 the pairs of numbers beside the vertices of each 3-SSAW config represent respectively the labelling and the relative order of the walk starting at that vertex, induced by the ordering of the edges of the SSAW in which the pattern occurs; the relative order on the walks also determines a relative order on the edges of the 3-SSAW config.

For any SSAW G, let $e(G)$, $d(G,1)$ and $d(G,2)$ denote respectively the span, the number of edges and the number of degree one vertices of G. For any $1 \le i \le |\Pi(k)|$ ($1 \le i \le |\Pi_1(k)|$), let $e_i = 1$ ($e_i' = 1$) and, for any $1 \le i \le |\Pi_2(k)|$, let $e_i'' = k$. Also, for any $1 \le i \le |\Pi(k)|$ ($1 \le i \le |\Pi_1(k)|$), let $d_i(1)$ ($d_i'(1)$) and $d_i(2)$ ($d_i'(2)$) be respectively the number of edges and the number of vertices with degree one (on the tube's boundary) in the first hinge and section of P_i (P_i'). Thus $D_i = [d_i(1), d_i(2)]^T$ and $D_i' = [d_i'(1), d_i'(2)]^T$ ($t = 2$). For any $1 \le i \le |\Pi_2(k)|$, let also $d_i''(1)$ and $d_i''(2)$ be respectively the total number of edges and the total number of vertices with degree one (on the tube's boundary) in P_i''. Thus, $D_i'' = [d_i''(1), d_i''(2)]^T$. Then equations 2.6.7 and 2.6.9 are satisfied.

Given an SSAW config P, define

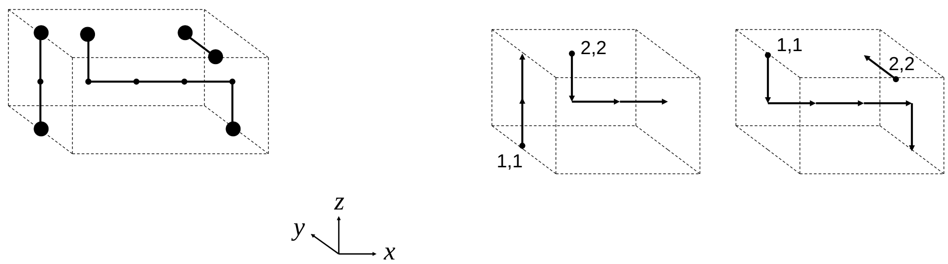
$$Z_m^{\#1}(N, M, Y_3; \bar{P}) \equiv \sum_G x^{f(G,1)} y^{f(G,2)} z^{f(G,3)}, \tag{6.9.4}$$

where the sum is over all the span m SSAWs in C_m^* which do not contain the SSAW config P. Also define

$$\mathcal{F}^{\#1}(N, M, Y_3; \bar{P}) \equiv \limsup_{m\to\infty} m^{-1} \log Z_m^{\#1}(N, M, Y_3; \bar{P}). \tag{6.9.5}$$

Moreover, let $y_1 = z$, $y_2 = y$, $Y = [y_1, y_2]^T$ and $Y_3 = [1, y, z]^T$. Then

$$Z_m(N, M; Y) = \sum_{G \in C_m^*} y_1^{d(G,1)} y_2^{d(G,2)} = Z_m^{\#1}(N, M, Y_3) \tag{6.9.6}$$



(a) An example of an 8-edge SSAW with span 4 in $T(2,2)$.

(b) The associated 3-SSAW configs.

Figure 6.9: An example of an SSAW in a tube and its associated 3-SSAW configs. Note that the pairs of numbers beside the vertices of each 3-SSAW config represent respectively the labelling and the relative order of the walk starting at that vertex; the relative order on the walks also determines a relative order on the edges of the 3-SSAW config.

and

$$Z_m(N, M; \bar{P}, Y) = \sum_{G \in C^*_m(\bar{P})} y_1^{d(G,1)} y_2^{d(G,2)} = Z^{\#1}_m(N, M, Y_3; \bar{P}). \tag{6.9.7}$$

Theorem 2.6.1 and the fact that the Cluster Configuration Axiom and (CONCAT) hold for SSAWs yield the following result regarding the asymptotic form of $Z^{\#1}_m(N, M, Y_3)$ and also the asymptotic form of $Z^{\#1}_m(N, M, Y_3; \bar{P})$.

Theorem 6.9.1. *For $Y_3 = [1, y, z]^T$, any integer $k \geq 2$ and any proper SSAW config $P \in \Pi(k)$, there exist non-negative values $x_0(Y)$ and α_Y such that*

$$Z^{\#1}_m(N, M, Y_3) = \alpha_Y (x_0(Y))^{-m} + o((x_0(Y))^{-m}) \quad \text{as } m \to \infty. \tag{6.9.8}$$

Moreover, there exist non-negative values $\bar{x}_0(Y) > x_0(Y)$ and $\bar{\alpha}_Y$ such that

$$Z^{\#1}_m(N, M, Y_3; \bar{P}) = \bar{\alpha}_Y (\bar{x}_0(Y))^{-m} + o((\bar{x}_0(Y))^{-m}) \quad \text{as } m \to \infty. \tag{6.9.9}$$

6.10 Asymptotic Behaviour of the Average Edge-Density of SSAWs

For any SSAW G, let $\psi(G) = f(G, 3)$, i.e. the number of edges of G. For $1 \leq i \leq |\Pi(k)|$ ($1 \leq i \leq |\Pi_1(k)|$), let $\psi(P_i)$ ($d'(P_i')$) be the number of edges in the first hinge and section of P_i (P_i'). For $1 \leq i \leq |\Pi_2(k)|$, let $d''(P_i'')$ be the total number of edges in P_i''. Then equation 2.6.23 is satisfied thus $f(G, 3)$ is an additive functional.

Let $t = 2$, $Y = [z, y]^T$ and $Y_3 = [1, y, z]^T$ as defined in Theorem 6.9.1. As discussed before, $X_m^{\#}(Y_3)$ is a random variable taking its values from the SSAWs in C_m^* and with the probability distribution

$$\mathbb{P}(X_m^{\#1}(Y_3) = G) = \frac{y^{f(G,2)} z^{f(G,3)}}{Z_m^{\#1}(N, M, Y_3)}. \tag{6.10.1}$$

hence $f(X_m^{\#1}(Y_3), 3)$ is a random variable, representing the number of edges in a randomly chosen SSAW $X_m^{\#1}(Y_3)$ with span m, with the probability distribution

$$\mathbb{P}(f(X_m^{\#1}(Y_3), 3) = a) = \frac{\sum_{G:f(G,1)=m} y^{f(G,2)} z^a}{Z_m^{\#1}(N, M, Y_3)}. \tag{6.10.2}$$

Also $\frac{E_Y(f(X_m^{\#1}(Y_3),3))}{NMm}$, the expected number of edges per unit volume, represents the *average edge-density* of SSAWs. Theorem 2.6.2 implies that there exists $\gamma_Y > 0$ such that

$$E_Y(f(X_m^{\#1}(Y_3), 3)) = \gamma_Y m + O(1) \tag{6.10.3}$$

and hence

$$\frac{1}{NMm} E_Y(f(X_m^{\#1}(Y_3), 3)) \to \frac{1}{NM}\gamma_Y \tag{6.10.4}$$

as $m \to \infty$.

The following Theorem is an immediate result of Theorem 6.4.2 and investigates the behaviour of the expected value of $f(X_m^{\#1}(Y_3), 3)$, $E_Y(f(X_m^{\#1}(Y_3), 3))$, with respect to $\log z$.

Corollary 6.10.1. *Given* $Y_3 = [1, y, z]^T$,

$$E_Y(f(X_m^{\#1}(Y_3), 3)) = m \frac{d}{d \log z} F_m^{\#1}(N, M, Y_3) \tag{6.10.5}$$

almost everywhere. Hence $E_Y(f(X_m^{\#1}(Y_3),3))$ *is non-decreasing in* $\log z$ *almost everywhere.*

Proof. By Theorem 6.4.2, the following equality holds almost everywhere:

$$\begin{aligned}
\frac{d}{d\log z}F_m^{\#1}(N,M,Y_3) &= \frac{d}{d\log z}\left[m^{-1}\log Z_m^{\#1}(N,M,Y_3)\right] \\
&= m^{-1}\frac{\sum_{G:f(G,1)=m} f(G,3)y^{f(G,2)}z^{f(G,3)}}{Z_m(N,M,Y_3)} \\
&= m^{-1}E_Y(f(X_m^{\#1}(Y_3),3)). \qquad (6.10.6)
\end{aligned}$$

Hence

$$E_Y(f(X_m^{\#1}(Y_3),3)) = m\frac{d}{d\log z}F_m^{\#1}(N,M,Y_3) \qquad (6.10.7)$$

is non-decreasing in $\log z$ almost everywhere. □

Hence, the average number of edges per unit volume of a randomly chosen span m SSAW, $\frac{E_Y(f(X_m^{\#1}(Y_3),3))}{mNM}$, approaches a positive constant as $m \to \infty$ and is non-decreasing in $\log z$ and thus non-decreasing in z almost everywhere.

Given a proper k-SSAW config $P_l \in \Pi(k)$, define $\psi_l : C^* \to \mathbb{N}\cup\{0\}$ such that $\psi_l(G)$ is the number of times the SSAW config P_l appears in G, for any SSAW $G \in C^*$. For $1 \le i \le |\Pi(k)|$, let $\psi_l(P_i) = \delta_{i,l}$ where

$$\delta_{i,l} = \begin{cases} 1 & \text{if } i = l \\ 0 & \text{if } i \neq l \end{cases}. \qquad (6.10.8)$$

Similarly, for $1 \le i \le |\Pi_1(k)|$ ($1 \le i \le |\Pi_2(k)|$), let $d'_l(P'_i) = \delta_{i,l}$ ($d''_l(P''_i) = \delta_{i,l}$). Then equation 2.6.23 is satisfied thus ψ_l is an additive functional.

Then $\psi_l(X_m^{\#1}(Y_3))$ is a random variable, representing the number of times the SSAW config P_l appears in a randomly chosen SSAW with span m, with the probability distribution

$$\mathbb{P}(\psi_l(X_m^{\#1}(Y_3)) = b) = \frac{\sum_{G:\psi_l(G)=b} y^{f(G,2)}z^{f(G,3)}}{Z_m^{\#1}(N,M,Y_3)}, \qquad (6.10.9)$$

where the sum is over all the span m SSAWs $G \in C_m^*$ with $\psi_l(G) = b$. Also $\frac{E_Y(\psi_l(X_m^{\#1}(Y_3)))}{m}$ represents the *average density of* P_l *per unit span* in SSAWs. Theorem

2.6.2 implies that there exists $\gamma_l > 0$ such that

$$E_Y(\psi_l(X_m^{\#1}(Y_3))) = \gamma_l m + O(1) \tag{6.10.10}$$

and thus

$$\frac{1}{m} E_Y(\psi_l(X_m^{\#1}(Y_3))) \to \gamma_l \tag{6.10.11}$$

as $m \to \infty$.

Let $P_l = P_+$ be the proper 2-SSAW config obtained by ordering and labelling the edges of the 2-pattern P_1 introduced in Theorem 6.7.1. Then the fact that $EC(G) \geq \psi_l(G)$, for any SSAW G, and equation 5.4.14 imply that

$$\liminf_{m\to\infty} \frac{E_Y\big(EC(X_m^{\#1}(Y_3))\big)}{m} \geq \gamma_l. \tag{6.10.12}$$

Hence the average value of EC per span is bounded below by a positive value, representing the density of the 2-SSAW config P_+ in span m SSAWs. However, it is still an open question how EC depends on the density of SSAWs and how γ_l depends on z.

6.11 Asymptotic Behaviour of the SSAWs with a Fixed Limiting Edge-Density

In order to explore the relation between EC and the density of SSAWs, in this section we investigate the asymptotic behaviour of the n-edge SSAWs in $T(N, M)$ with a fixed edge-density $\rho(n) = \frac{n}{NMm}$, where m represents the span. So we are primarily concerned with the number of n-edge SSAWs with a fixed span m, denoted by $q_m^s(N, M; n) \equiv \sum_{a_2} q(N, M, m, a_2, n)$. Note that these SSAWs are said to have edge-density $\rho(n)$. $q_m^s(N, M; \lfloor \epsilon m \rfloor)$ is the number of SSAWs with edge-density $\rho(\lfloor \epsilon m \rfloor)$ and limiting edge-density $\rho_* = \frac{\epsilon}{NM}$ as $m \to \infty$. Note also that $\epsilon = NM\rho_*$ so ϵ is proportional to ρ_*. The function $\rho(N, M; \epsilon)$ is defined to be the growth constant for SSAWs with limiting edge-density $\frac{\epsilon}{NM}$, i.e.

$$\log \rho(N, M; \epsilon) \equiv \lim_{m\to\infty} m^{-1} \log q_m^s(N, M; \lfloor \epsilon m \rfloor). \tag{6.11.1}$$

Here we will investigate the existence of this function. Following the terminology from [26], we will start with showing that Assumptions 3.1 from [26] are satisfied for the SSAW model:

Lemma 6.11.1. *The following statements hold:*

(1) There exists a constant $S > 0$ independent of n and m such that $0 \leq q^s_m(N, M; n) \leq S^m$ for each value of m and n.

(2) There exist a finite constant $C > 0$, and numbers A_m and B_m such that $0 \leq A_m \leq B_m \leq Cm$ and $q^s_m(N, M; n) \geq 0$ if $A_m \leq n \leq B_m$ with $q^s_m(N, M; A_m) > 0$ and $q^s_m(N, M; B_m) > 0$, and $q^s_m(N, M; n) = 0$ otherwise.

(3) $q^s_m(N, M; n)$ satisfies the following inequality:

$$q^s_{m_1}(N, M; n_1) q^s_{m_2}(N, M; n_2) \leq q^s_{m_1+m_2+1}(N, M; n_1 + n_2). \tag{6.11.2}$$

Proof. **(1)** Let $\mathcal{A}$ be the set of all SSAW 1-patterns in $T(N, M)$. Let S be the number of elements in $\mathcal{A}$. Any SSAW 1-pattern P is composed of a set of occupied edges, occupied vertices and unoccupied vertices from all the edges and vertices lying in $V_1(N, M) = \{(x, y, z) \in \mathbb{Z}^3 \mid \ 0 \leq x \leq 1,\ 0 \leq y \leq N,\ 0 \leq z \leq M\}$. There are in total $((N+1)(M+1)+2N(M+1)+2(N+1)M)$ edges and $2(N+1)(M+1)$ vertices in $V_1(N, M)$. Therefore, S is bounded above by the number of ways to choose the occupied edges and vertices from $V_1(N, M)$ which leads to

$$\begin{aligned} S &\leq \left(2^{(N+1)(M+1)+2N(M+1)+2(N+1)M}\right)\left(2^{2(N+1)(M+1)}\right) \\ &= 2^{3(N+1)(M+1)+2N(M+1)+2(N+1)M}. \end{aligned} \tag{6.11.3}$$

Moreover, every SSAW G with span m and left-most plane $x = x_1$ can be represented as a sequence of length m in $\mathcal{A}$, i.e. $\{P_1, P_2, ..., P_m\}$ where $P_i \in \mathcal{A}$ occurs at $x = x_1 + (i - 1)$ for $i = 1, 2, ..., m$. The total number of such sequences is S^m but not necessarily every sequence corresponds to an SSAW. Therefore,

$$0 \leq q_m^s(N, M; n) \leq S^m. \tag{6.11.4}$$

(2) Given an n-edge SSAW G with a fixed span m , let G be composed of the USAWs $w_1, ..., w_k$. For $1 \leq i \leq k$, let n_i and v_i denote respectively the number of edges and vertices of w_i. Then

$$\sum_{i=1}^{k} n_i = \sum_{i=1}^{k}(v_i - 1) = -k + \sum_{i=1}^{k} v_i. \tag{6.11.5}$$

Each w_i contains at least two vertices so

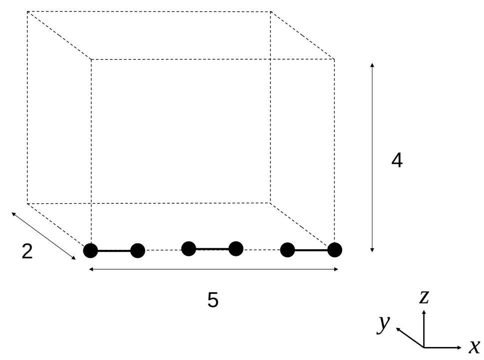
$$k \leq \lfloor \frac{v}{2} \rfloor, \tag{6.11.6}$$

where $v = \sum_{i=1}^{k} v_i$. Hence, equations 6.11.5 and 6.11.6 give

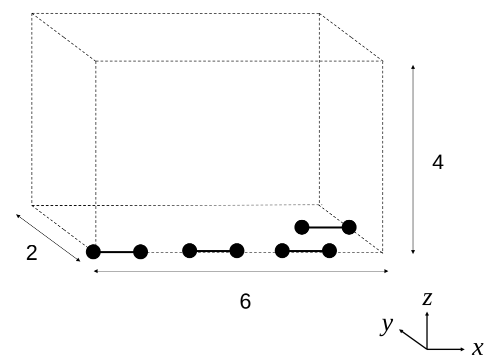
$$\begin{aligned} \sum_{i=1}^{k} n_i &\geq -\lfloor \frac{v}{2} \rfloor + v \\ &= \begin{cases} \lfloor \frac{v}{2} \rfloor & \text{if } v \text{ is even} \\ \lfloor \frac{v}{2} \rfloor + 1 & \text{if } v \text{ is odd} \end{cases} \\ &\geq \lfloor \frac{m}{2} \rfloor + 1, \end{aligned} \tag{6.11.7}$$

where the last inequality holds since for any span m SSAW G, by Definition 6.2.1, there has to be at least one vertex of G in the plane $x = x_0$ for any $x_1 \leq x_0 \leq x_1 + m$, where $x = x_1$ is the left-most plane of G; i.e. $v \geq m + 1$.

Next, we will show that this lower bound for the number of edges of G can be reached, i.e. $q_m^s(N, M; \lfloor \frac{m}{2} \rfloor + 1) > 0$. If m is odd then the SSAW G' composed of the edges $\{(x, 0, 0), (x + 1, 0, 0)\}$, for $x \in \mathbb{Z}$ and $0 \leq x \leq m - 1$, has span m and exactly $\lfloor \frac{m}{2} \rfloor + 1$ edges. Figure 6.10 (a) shows an example of an SSAW in $T(2, 4)$ which has the minimum number of edges for the fixed span $m = 5$. On the other hand, if m is even then the SSAW G'' composed of the edges $\{(x, 0, 0), (x+1, 0, 0)\}$, for $x \in \mathbb{Z}$ and $0 \leq x \leq m - 2$, and $\{(m - 1, 1, 0), (m, 1, 0)\}$ has span m and exactly $\lfloor \frac{m}{2} \rfloor + 1$ edges.



(a) An example of an SSAW in $T(2,4)$ with the minimum edge-density among all the SSAWs with the fixed span $m = 5$



(b) An example of an SSAW in $T(2,4)$ with the minimum edge-density among all the SSAWs with the fixed span $m = 6$

Figure 6.10: Examples of SSAWs with the minimum edge-density for a fixed span m.

Figure 6.10 (b) shows an example of an SSAW in $T(2,4)$ which has the minimum number of edges for the fixed span $m = 6$. Therefore, $A_m = \lfloor \frac{m}{2} \rfloor + 1$.

Moreover, take B_m to be the maximum number of edges that an SSAW with span m can have. If $N = 1$ (or $M = 1$) then we will show that $B_m = (M+1)(m+1)$ ($B_m = (N+1)(m+1)$). Suppose without loss of generality that $N = 1$, then every SSAW G in $T(1, M)$ is composed of only single-edge USAWs. Thus any G with span m and left-most plane $x = 0$ will have a maximum number of edges if it occupies all the vertices in a slice $V_m(1, M) = \{(x, y, z) \in \mathbb{Z}^3 \mid 0 \leq x \leq m,\ 0 \leq y \leq 1,\ 0 \leq z \leq M\}$ of width m in $T(1, M)$. There are in total $2(M+1)(m+1)$ vertices in such a slice. Hence $B_m \leq (M+1)(m+1)$. Moreover, this upper bound for B_m can be reached since the following SSAW G with span m has exactly $(M+1)(m+1)$ edges. Let G be an SSAW composed of $(M+1)(m+1)$ single-edge USAWs each joining the vertex $(i, 0, k)$ to the vertex $(i, 1, k)$ for $x_1 \leq i \leq x_2$ and $0 \leq k \leq M$, where $x = x_1$ ($x = x_2$) is the left-most (right-most) plane of G. Therefore, $B_m = (M+1)(m+1)$ if $N = 1$.

For $N > 1$ and $M > 1$, let G be an n-edge SSAW in $T(N, M)$ with span m and left-most plane $x = 0$. Let $w_1, ..., w_{k_1}$ be the USAWs in G which do not have any edge in the boundary planes of $T(N, M)$. Let $w'_1, ..., w'_{k_2}$ be the single edge USAWs in G which lie in the boundary planes of $T(N, M)$. For $1 \leq i \leq k_1$, let n_i denote the number of edges in w_i. Let also v_i denote the number of vertices of w_i in $T(N-2, M-2) + (0, 1, 1)$. Note that $n_i - 2 = v_i - 1$ and hence we have $n_i = v_i + 1$ since G is an SSAW. Thus the total number of edges in G can be evaluated as follows:

$$\begin{aligned} n &= k_2 + \sum_{i=1}^{k_1} n_i \\ &= k_2 + \sum_{i=1}^{k_1} (v_i + 1) \\ &= k_1 + k_2 + \sum_{i=1}^{k_1} v_i \\ &\leq k_1 + k_2 + (N-1)(M-1)(m+1), \end{aligned} \tag{6.11.8}$$

where $(N-1)(M-1)(m+1)$ is the total number of vertices in a slice of width m in $V(N-2, M-2) + (0, 1, 1)$.

Each w_i has its endpoints on the boundary planes so k_1 USAWs will occupy exactly $2k_1$ vertices of the total $2(N+M)(m+1)$ vertices of the boundary planes. Therefore, $2k_2 \leq 2(N+M)(m+1) - 2k_1$ so

$$\begin{aligned} n &\leq (N-1)(M-1)(m+1) + k_1 + k_2 \\ &\leq (N-1)(M-1)(m+1) + k_1 + (N+M)(m+1) - k_1 \\ &\leq (NM+1)(m+1). \end{aligned} \tag{6.11.9}$$

If N or M is odd then we show next that the above upper bound for the number of edges can be reached, i.e. $B_m = (NM+1)(m+1)$. If N is odd then a span m SSAW G_1 with $(NM+1)(m+1)$ edges can be constructed as follows: let G_1 be the SSAW composed of the USAWs $(N\hat{j})$ starting at $(x, 0, z)$, for $0 \leq x \leq m+1$ and $1 \leq z \leq M-1$, and the USAWs $(\hat{j})$ starting at $(x', 2y', z')$, for $0 \leq x' \leq m+1$, $0 \leq y' \leq \frac{N-1}{2}$ and $z' \in \{0, M\}$. Note that the edge orientations are ignored. Similarly, if M is odd, it can be shown that $B_m = (NM+1)(m+1)$. If both N and M are even but m is odd then again a similar argument shows that $B_m = (NM+1)(m+1)$. A span m SSAW G_2 with $(NM+1)(m+1)$ edges can be constructed as follows: let G_2 be the SSAW composed of the USAWs $(N\hat{j})$ starting at $(x, 0, z)$, for $0 \le x \le m+1$ and $1 \le z \le M-1$, and the USAWs $(\hat{i})$ starting at $(2x', y', z')$, for $0 \le x' \le \frac{m-1}{2}$, $0 \le y' \le N+1$ and $z' \in \{0, M\}$. However, for the case that N, M and m are all even the above construction leads only to a lower bound $B_m \ge (NM+1)(m+1)-1$. A span m SSAW G_3 with $(NM+1)(m+1)-1$ edges can be constructed as follows: let G_3 be the SSAW composed of the USAWs $(N\hat{j})$ starting at $(x, 0, z)$, for $0 \le x \le m+1$ and $1 \le z \le M-1$, the USAWs $(\hat{i})$ starting at $(2x', y', z')$, for $0 \le x' \le \frac{m-2}{2}$, $0 \le y' \le N+1$ and $z' \in \{0, M\}$, and the USAWs $(\hat{j})$ starting at $(m, 2y'', z'')$, for $0 \le y'' \le \frac{N-2}{2}$ and $z'' \in \{0, M\}$. Thus $q^s_m(N, M; B_m) > 0$.

We can also take $C = 10NM$ which satisfies the following inequality:

$$A_m \le B_m \le Cm. \tag{6.11.10}$$

Furthermore, for any $n < A_m$ or $n > B_m$,

$$q^s_m(N, M; n) = 0 \tag{6.11.11}$$

since A_m and B_m are respectively the minimum and maximum number of edges that an SSAW with span m can have.

(3) The concatenation process, Lemma 6.3.2, implies that

$$q^s_{m_1}(N, M; n_1) q^s_{m_2}(N, M; n_2) \le q^s_{m_1+m_2+1}(N, M; n_1 + n_2). \tag{6.11.12}$$

□

The maximal and minimal densities in the model can be obtained from A_m and B_m as follows:

$$\epsilon_{max} = \limsup_{m\to\infty} [B_m/m], \tag{6.11.13}$$

$$\epsilon_{min} = \limsup_{m\to\infty} [A_m/m]. \tag{6.11.14}$$

Note that for $N = 1$

$$\epsilon_{max} = \lim_{m\to\infty}[B_m/m] = \lim_{m\to\infty}((M+1)(m+1))/m = M+1 \tag{6.11.15}$$

and similarly for $M = 1$

$$\epsilon_{max} = \lim_{m\to\infty}[B_m/m] = \lim_{m\to\infty}((N+1)(m+1))/m = N+1. \tag{6.11.16}$$

Also for $N > 1$ and $M > 1$,

$$\epsilon_{max} = \lim_{m\to\infty}[B_m/m] = NM+1. \tag{6.11.17}$$

Furthermore,

$$\epsilon_{min} = \limsup_{m\to\infty}[A_m/m] = \lim_{m\to\infty}(\lfloor\frac{m}{2}\rfloor+1)/m = 1/2. \tag{6.11.18}$$

Now, by Lemma 6.11.1, Assumptions 3.1 from [26] are satisfied for the SSAW model so the arguments presented in [26] can be used in a very straightforward fashion to give a proof for the following theorems.

Lemma 6.11.2 (Janse van Rensburg 2000 [26]). *Given $\epsilon > 0$, there exists a function z_{m_1,m_2}, dependent on ϵ, and with $|z_{m_1,m_2}| \le 1$, such that*

$$q^s_{m_1}(N,M;\lfloor\epsilon m_1\rfloor)q^s_{m_2}(N,M;\lfloor\epsilon m_2\rfloor + z_{m_1,m_2}) \quad\le\quad q^s_{m_1+m_2+1}(N,M;\lfloor\epsilon(m_1+m_2+1)\rfloor) \; . \tag{6.11.19}$$

Let $\epsilon \in (\epsilon_{min}, \epsilon_{max})$. The following theorem proves the existence of $\rho(N,M;\epsilon)$, the growth constant for the SSAWs with the limiting edge-density $\frac{\epsilon}{NM}$.

Theorem 6.11.3 (Janse van Rensburg 2000 [26]). *Given any $\epsilon \in (\epsilon_{min}, \epsilon_{max})$, the following limit exists and is finite:*

$$\log\rho(N,M;\epsilon) = \lim_{m\to\infty} m^{-1}\log q^s_m(N,M;\lfloor\epsilon m\rfloor). \tag{6.11.20}$$

Moreover, there exists $\eta_m \in \{0,1\}$ such that for each value of m,

$$q^s_m(N,M;\lfloor\epsilon m\rfloor + \eta_m) \le [\rho(N,M;\epsilon)]^m. \tag{6.11.21}$$

Theorem 6.11.4 (Janse van Rensburg 2000 [26]). *$\log \rho(N,M;\epsilon)$ is a concave function of $\epsilon \in (\epsilon_{min}, \epsilon_{max})$. Therefore, $\log \rho(N,M;\epsilon)$ is continuous in $(\epsilon_{min}, \epsilon_{max})$, has right- and left-derivatives everywhere in $(\epsilon_{min}, \epsilon_{max})$, and is differentiable almost everywhere in $(\epsilon_{min}, \epsilon_{max})$.*

Theorem 6.11.5 (Janse van Rensburg 2000 [26]). *Let δ_m be a sequence of integers such that $A_m \le \delta_m \le B_m$ for all $m \ge M_0$, where M_0 is a fixed integer. Suppose that $\lim_{n\to\infty}[\delta_m/m] = \delta$ with $\epsilon_{min} < \delta < \epsilon_{max}$. Then*

$$\lim_{m\to\infty} m^{-1} \log q^s_m(N,M;\delta_m) = \log \rho(N,M;\delta). \qquad (6.11.22)$$

By Theorem 6.11.3, the function $\rho(N,M;\epsilon)$ is defined on $(\epsilon_{min}, \epsilon_{max})$. We can also define this function at the boundary point ϵ_{min} as follows:

Lemma 6.11.6. *The following limit exists and is finite for the SSAW model:*

$$\log \rho(N,M;\epsilon_{min}) \equiv \lim_{m\to\infty} m^{-1} \log q^s_m(N,M;A_m) = \frac{1}{2}\log(2N+2M). \qquad (6.11.23)$$

Proof. Let G be an SSAW with fixed span m and $A_m = \lfloor \frac{m}{2} \rfloor + 1$ edges. Let G be composed of the USAWs $w_1, ..., w_k$. For $1 \le i \le k$, let n_i and v_i denote respectively the number of edges and vertices of w_i. The value of $q^s_m(N,M;A_m)$ is calculated in the following distinct cases:

1) If m is odd then $A_m = \lfloor \frac{m}{2} \rfloor + 1$ implies that

$$v = \sum_{i=1}^{k} v_i \le 2(\lfloor \frac{m}{2} \rfloor + 1) = m+1 \qquad (6.11.24)$$

since each walk w_i contains at least two vertices. Moreover, by the definition of SSAWs, $v \ge m+1$ for any SSAW with span m. Therefore, $v = m+1$.

Furthermore,

$$\begin{aligned} A_m = \sum_{i=1}^{k} n_i &= \sum_{i=1}^{k} (v_i - 1) \\ &= -k + v \\ &= -k + m + 1. \end{aligned} \qquad (6.11.25)$$

Hence,

$$k = m + 1 - A_m = m + 1 - \lfloor \frac{m}{2} \rfloor - 1 = \lfloor \frac{m}{2} \rfloor + 1. \tag{6.11.26}$$

Therefore, G contains exactly $\lfloor \frac{m}{2} \rfloor + 1$ USAWs. Each w_i has exactly 2 vertices since otherwise $v = \sum_{i=1}^{k} v_i > 2k = m + 1$. Since G has span m, by the SSAW definition it must have at least one vertex in each of the $m+1$ planes $x = x_1, ..., x = x_1 + m$. Since $v = m + 1$, hence w_i must lie on a boundary plane with an edge from (x, y, z) to $(x+1, y, z)$, for some integer values of x, y and z. Because the vertex (x, y, z) lies on a boundary plane, for any $0 \leq x \leq m - 1$, there are exactly $2N + 2M$ choices for the pair (y, z). Therefore, in total there are $(2N + 2M)^{\lfloor \frac{m}{2} \rfloor + 1}$ choices for G. Since m is odd, $\lfloor \frac{m}{2} \rfloor = \frac{m-1}{2}$; hence

$$q_m^s(N, M; A_m) = (2N + 2M)^{\lfloor \frac{m}{2} \rfloor + 1} \tag{6.11.27}$$

and

$$\begin{aligned} \log \rho(N, M; \epsilon_{min}) = \log \rho(N, M; \frac{1}{2}) &= \lim_{m \to \infty} m^{-1} \log q_m^s(N, M; A_m) \\ &= \lim_{m \to \infty} m^{-1} \log[(2N + 2M)^{\lfloor \frac{m}{2} \rfloor + 1}] \\ &= \frac{1}{2} \log(2N + 2M), \end{aligned} \tag{6.11.28}$$

when the limit is taken through odd values of m.

2) If m is even then $A_m = \lfloor \frac{m}{2} \rfloor + 1$ implies that

$$v = \sum_{i=1}^{k} v_i \leq 2(\lfloor \frac{m}{2} \rfloor + 1) = m + 2. \tag{6.11.29}$$

If $v = m + 1$ then $k = m + 1 - \lfloor \frac{m}{2} \rfloor - 1 = \lfloor \frac{m}{2} \rfloor = \frac{m}{2}$. Each walk has at least two vertices so there has to be exactly one walk, w', with three vertices. By the definition of SSAWs, w' cannot lie in a boundary plane and hence because of the structure of $T(N, M)$ all the three vertices of w' must have the same x-coordinates. Therefore, there are $m + 1 - 1 = m$ more x-coordinates to fill in but there are left only $m + 1 - 3 = m - 2$ more vertices. This is a contradiction. Hence, $v = m + 2$

and $k = m + 2 - \lfloor \frac{m}{2} \rfloor - 1 = \lfloor \frac{m}{2} \rfloor + 1 = \frac{m}{2} + 1$. Each w_i has exactly 2 vertices since otherwise $v = \sum_{i=1}^{k} v_i > 2k = m + 2$. By the definition of SSAWs, w_i must lie on a boundary plane of G. Since G has span m, there is exactly one $x_1 \leq x_* \leq x_2$ so that there are two vertices u_1 and u_2 of G with $x(u_1) = x(u_2) = x_*$. Now, we consider the following two cases separately:

a) If all the edges of G are parallel to the x-axis then there exists an undirected walk, w_t where $1 \leq t \leq k$, in G with its bottom vertex in $\{u_1, u_2\}$. There are $\lfloor \frac{m}{2} \rfloor$ choices for the x-coordinate of the bottom vertex of w_t, x_*, so that $w_t = \{(x_*, y, z), (x_* + 1, y, z)\}$ for some integer values of y and z. There are exactly $R(N, M) = 2N + 2M - 1$ choices for the pair (y, z). Moreover, G contains exactly $\lfloor \frac{m}{2} \rfloor$ edges of the form $\{(x, y, z), (x + 1, y, z)\}$, for some integer values of x, y and z. There are exactly $2N + 2M$ choices for the pair (y, z). Therefore, in total there are $(2N + 2M)^{\lfloor \frac{m}{2} \rfloor}$ choices for such edges. Hence, for even m with G having all its edges parallel to the x-axis, the number of such SSAWs with A_m edges is

$$\lfloor \frac{m}{2} \rfloor R(N, M)(2N + 2M)^{\lfloor \frac{m}{2} \rfloor}. \tag{6.11.30}$$

b) Otherwise, there exists a walk, w'_t where $1 \leq t \leq k$, of G with a single edge $\{u_1, u_2\}$. There are $\lfloor \frac{m}{2} \rfloor$ choices for the x-coordinate of the vertices of w'_t, x_*, so that $w'_t = \{(x_*, y, z), (x_*, y + 1, z)\}$ or $w'_t = \{(x_*, y, z), (x_*, y, z + 1)\}$ for some integer values of x, y and z. There are exactly $R'(N, M) = 2N + 2M$ (or $R'(N, M) = 2N + 2M + M - 1$ if $N = 1$ and $M > 1$ or $M = 1$ and $N > 1$) choices for the pair (y, z). Moreover, G contains exactly $\lfloor \frac{m}{2} \rfloor$ edges of the form $\{(x, y, z), (x+1, y, z)\}$, for some integer values of x, y and z. There are also $2N+2M$ choices for the pair (y, z). Therefore, in total there are $(2N + 2M)^{\lfloor \frac{m}{2} \rfloor}$ choices for such edges. Hence, for m even,

$$q^s_m(N, M; A_m) = \lfloor \frac{m}{2} \rfloor \big(R(N, M) + R'(N, M)\big)(2N + 2M)^{\lfloor \frac{m}{2} \rfloor} \tag{6.11.31}$$

and

$$\begin{aligned}\log\rho(N,M;\epsilon_{min}) &= \log\rho(N,M;\frac{1}{2}) \\ &= \lim_{m\to\infty} m^{-1}\log q^s_m(N,M;A_m) \\ &= \lim_{m\to\infty} m^{-1}\log\left[\lfloor\frac{m}{2}\rfloor\big(R(N,M)+R'(N,M)\big)(2N+2M)^{\lfloor\frac{m}{2}\rfloor}\right] \\ &= \frac{1}{2}\log(2N+2M), \end{aligned} \tag{6.11.32}$$

where the limit is through even values of m. Combining equations 6.11.28 and 6.11.32 gives the result. □

$$q^{\#1}_m(N,M) = \sum_{n\geq 0} q^s_m(N,M;n) \tag{6.11.33}$$

is the number of SSAWs with span m. In Theorem 6.4.1 it was shown that

$$\mu^s(N,M) \equiv e^{\kappa^{\#1}_q(N,M)} = \lim_{m\to\infty}(q^{\#1}_m(N,M))^{1/m} \tag{6.11.34}$$

exists and is finite. The following theorem shows that the maximum value of $\rho(N,M;\epsilon)$ is in fact equal to $\mu^s(N,M)$.

Theorem 6.11.7 (Janse van Rensburg 2000 [26])**.** *There exist values ϵ_0 and ϵ_1 in $[\epsilon_{min},\epsilon_{max}]$ such that $\rho(N,M;\epsilon)=\mu^s(N,M)$ for all values of $\epsilon\in[\epsilon_0,\epsilon_1]$.*

Proof. Given fixed integer $m>0$, let δ_m be the minimum value of n which maximizes $q^s_m(N,M;n)$, i.e.

$$\delta_m = \min_{n\in\mathbb{N}}\{n \mid \ q^s_m(N,M;n) \geq q^s_m(N,M;n')\ \forall n'\}. \tag{6.11.35}$$

Then

$$q^s_m(N,M;\delta_m) \leq q^{\#1}_m(N,M) = \sum_{n=A_m}^{B_m} q^s_m(N,M;n) \leq (1+B_m-A_m)q^s_m(N,M;\delta_m). \tag{6.11.36}$$

Therefore, taking logarithms and dividing by m gives

$$\lim_{m\to\infty} m^{-1}\log q^s_m(N,M;\delta_m) = \log\mu^s(N,M). \tag{6.11.37}$$

Let also ς_m be the maximum value of n which maximizes $q^s_m(N,M;n)$, i.e.

$$\varsigma_m = \max_{n\in\mathbb{N}}\{n \mid \ q^s_m(N,M;n) \geq q^s_m(N,M;n') \ \forall n'\}, \tag{6.11.38}$$

and

$$\epsilon_0 = \liminf_{m\to\infty}[\delta_m/m], \quad \epsilon_1 = \limsup_{m\to\infty}[\varsigma_m/m]; \tag{6.11.39}$$

then by Theorem 6.11.5 $\rho(N,M;\epsilon) = \mu^s(N,M)$ for all values of $\epsilon \in [\epsilon_0, \epsilon_1]$. □

The following theorem also shows the relationship between $\rho(N,M;\epsilon)$ and the limiting free energy $\mathcal{F}^{\#1}(N,M,Y_3)$ when $Y_3 = [1,1,z]^T$.

Theorem 6.11.8 (Janse van Rensburg 2000 [26])**.**

$$\mathcal{F}^{\#1}(N,M,Y_3) = \sup_{\epsilon_{min}\leq\epsilon\leq\epsilon_{max}}\{\log\rho(N,M;\epsilon) + \epsilon\log z\} \tag{6.11.40}$$

and

$$\log\rho(N,M;\epsilon) = \sup_{0<z<\infty}\{\mathcal{F}^{\#1}(N,M,Y_3) - \epsilon\log z\}. \tag{6.11.41}$$

Also $\mathcal{F}^{\#1}(N,M,Y_3) \geq \max\{\epsilon_{min}\log z, \epsilon_{max}\log z\}$.

In this section we investigated $\rho(N,M;\epsilon)$, the growth constant for SSAWs with limiting edge-density $\frac{\epsilon}{NM}$, as a function of ϵ. However, in order to see how EC changes in terms of the density, we need to know more about $\rho(N,M;\epsilon)$. In particular, proving a pattern theorem for this function may lead to some results regarding the connection between EC and the density.

6.12 Summary

In this chapter, we considered the measure proposed in [41] and asked how its value depends on various properties of SSAWs such as the total number of edges, span, the number of degree one vertices (or, equivalently, twice the number of disjoint walks) and the density. We rigorously proved that the entanglement complexity, as measured in [41], of a polymer system with size n (e.g. the number of edges, span or the number of degree one vertices) is asymptotically (with probability one) bounded below by a linear function of n; i.e. there exists a positive number γ such that the probability that a polymer system of size n has entanglement complexity greater than γn approaches 1 as n goes to infinity. This supports the idea that EC is a good measure of topological entanglement in polymer systems modelled by SSAWs. We also showed that the entanglement complexity of SSAWs of size n is bounded above by a linear function of n. Furthermore, measuring the size by the number of edges, for $N \geq 2$ and $M \geq 2$, we compared the connective constant for SSAWs in an (N, M)-tube with the connective constant for self-avoiding walks in an $(N-2, M-2)$-tube and showed that it is strictly greater than that for SAWs.

Ultimately, based on our theoretical results on the SSAW model we conclude that the following statements (equations) show how EC depends on various properties of SSAWs such as the number of edges, span and the number of degree one vertices and also indicate how these properties are related to each other:

- There exists $\gamma^{\#1}(Y_3) > 0$ such that the probability that the EC of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ from the distribution in equation 6.7.7 (for $j = 1$) is bounded below by $\gamma^{\#1}(Y_3)n$ and bounded above by $a'(N, M)a(N, M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}(\gamma^{\#1}(Y_3)n < EC(X_n^{\#1}(Y_3)) \leq a'(N, M)a(N, M)n) = \lim_{n\to\infty}[1 - e^{-\alpha'(1,Y_3)n+o(n)}] = 1, \tag{6.12.1}$$

  where

$$a(N, M) = (1/2)(M + 2N + 1)[(N + 1)(4N - 2) + (9/2)(4N - 3)] \tag{6.12.2}$$

  and

$$a'(N, M) = 2N(M + 1) + 2M(N + 1) + (M + 1)(N + 1). \tag{6.12.3}$$

  In other words, as the span of SSAWs increases one expects EC to be bounded linearly in the span with probability one.

We also showed that the limit inferior of the average EC of span m SSAWs per span is bounded below by a positive constant, i.e.

$$\liminf_{m\to\infty} \frac{E_Y\big(EC(X_m^{\#1}(Y_3))\big)}{m} \geq \gamma_l. \tag{6.12.4}$$

- There exists $\gamma^{\#3}(Y_3) > 0$ such that the probability that the EC of a randomly chosen n-edge SSAW $X_n^{\#3}(Y_3)$ from the distribution in equation 6.7.7 (for $j = 3$) is bounded below by $\gamma^{\#3}(Y_3)n$ and bounded above by $a(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}(\gamma^{\#3}(Y_3)n < EC(X_n^{\#3}(Y_3)) \leq a(N,M)n) = \lim_{n\to\infty} [1 - e^{-\alpha'(3,Y_3)n+o(n)}] = 1. \tag{6.12.5}$$

  In other words, as the number of edges of SSAWs increases one expects EC to be bounded linearly in the number of edges with probability one.

- There exists $\epsilon_P^{\#1}(Y_3) > 0$ such that the probability that the number of degree one vertices of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ is bounded below by $\epsilon_P^{\#1}(Y_3)n$ and bounded above by $a'(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#1}(Y_3)n \leq f(X_n^{\#1}(Y_3),2) \leq a'(N,M)n\big) = \lim_{n\to\infty} [1-e^{-\alpha''(1,2,Y_3)n+o(n)}] = 1. \tag{6.12.6}$$

  In other words, as the span of SSAWs increases one expects the number of degree one vertices (the number of disjoint walks) to be bounded linearly in the span with probability one.

- There exists $\epsilon_P^{\#1}(Y_3) > 0$ such that the probability that the number of edges of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ is bounded below by $\epsilon_P^{\#1}(Y_3)n$ and bounded above by $a'(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#1}(Y_3)n \leq f(X_n^{\#1}(Y_3),3) \leq a'(N,M)n\big) = \lim_{n\to\infty} [1-e^{-\alpha''(1,3,Y_3)n+o(n)}] = 1. \tag{6.12.7}$$

  In other words, as the span of SSAWs increases one expects the number of edges to be bounded linearly in the span with probability one.

Using the transfer-matrix method, we also established that the average number of edges per unit volume of a randomly chosen span m SSAW, $\frac{E_Y(f(X_m^{\#1}(Y_3),3))}{mNM}$, approaches a positive constant as $m \to \infty$ and is non-decreasing in z almost everywhere. However, it still needs further investigation to see how EC changes with respect to Y_3.

- There exists $\epsilon_P^{\#3}(Y_3) > 0$ such that the probability that span of a randomly chosen SSAW $X_n^{\#3}(Y_3)$ with n edges is bounded below by $\epsilon_P^{\#3}(Y_3)n$ and bounded above by $2n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#3}(Y_3)n \le f(X_n^{\#3}(Y_3), 1) \le 2n\big) = \lim_{n\to\infty} [1 - e^{-\alpha''(3,1,Y_3)n+o(n)}] = 1. \tag{6.12.8}$$

  In other words, as the number of edges of SSAWs increases one expects the span to be bounded linearly in the number of edges with probability one.

- There exists $\epsilon_P^{\#3}(Y_3) > 0$ such that the probability that the number of degree one vertices of a randomly chosen SSAW $X_n^{\#3}(Y_3)$ with n edges is bounded below by $\epsilon_P^{\#3}(Y_3)n$ and bounded above by $2n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#3}(Y_3)n \le f(X_n^{\#3}(Y_3), 2) \le 2n\big) = \lim_{n\to\infty} [1 - e^{-\alpha''(3,2,Y_3)n+o(n)}] = 1. \tag{6.12.9}$$

  In other words, as the number of edges of SSAWs increases one expects the number of degree one vertices (the number of disjoint walks) to be bounded linearly in the number of edges with probability one.

Furthermore, we obtained the asymptotic form of the free energy for the SSAW model. We also investigated $\rho(N, M; \epsilon)$, the growth constant for SSAWs with limiting edge-density $\frac{\epsilon}{NM}$, as a function of ϵ. We established the existence of this function and showed that it is a continuous and concave function of ϵ and is differentiable almost everywhere in $(\epsilon_{min}, \epsilon_{max})$. However, in order to see how EC changes in terms of the density, we need to know more about $\rho(N, M; \epsilon)$. In particular, proving a pattern theorem for this function may lead to some results regarding the connection between EC and the density.

Chapter 7

Conclusions and Future Work

In this thesis, we investigated the topological entanglement of 2SAPs, stretched polygons and SSAWs confined to a lattice tube measured by homological (topological) linking probability, topological linking probability and EC respectively. We mainly addressed the following question: "Are these good measures for characterizing topological entanglement between polymers and how do these measures depend on various properties of the lattice object under consideration?" We rigorously showed that these are good measures for characterizing topological entanglement between polymers in the sense that the larger the "size" of a polymer system (where size can be measured by the number of edges, span or the number of degree one vertices), the greater its measure [50]. The main theoretical results obtained regarding the topological entanglement of the lattice objects under consideration (2SAPs, stretched polygons and SSAWs) as well as the important open questions raised from these models are summarized as follows.

In Chapter 4, the following question was addressed regarding the linking probability of two self-avoiding polygons: *under what conditions does the "linking probability" of pairs of self-avoiding polygons go to one?* The answer can depend on how one defines linking probability. In order to approach this question, we introduced the 2SAP model. We showed that (CONCAT) and (CAPOFF) are satisfied for this model. We also established the existence of the connective constant for 2SAPs and showed that it is strictly less than that of SAPs. We proved a pattern theorem for 2SAPs and used it to investigate homological as well as topological linking of 2SAPs. We showed that the homological linking probability goes to one at least as fast as $1 - O(n^{-1/2})$ and that the topological linking probability goes to one exponentially rapidly, as

$n \to \infty$. Furthermore we proved that the linking number grows (with probability one) faster than any function that is $o(\sqrt{n})$; i.e. for any function $f(n) = o(\sqrt{n})$, there exists $A \geq 0$ such that as $n \to \infty$ the probability that $|Lk(\omega_1, \omega_2)| \geq f(n)$, with (ω_1, ω_2) the component polygons of an n-edge 2SAP, satisfies

$$\mathbb{P}\big(|Lk(\omega_1, \omega_2)| \geq f(n)\big) = 1 - \frac{A}{\sqrt{n}} + o\big(\frac{1}{\sqrt{n}}\big). \tag{7.0.1}$$

Hence

$$\lim_{n \to \infty} \mathbb{P}\big(|Lk(\omega_1, \omega_2)| \geq f(n)\big) = 1. \tag{7.0.2}$$

We also showed that the linking number cannot grow faster than linearly in n because of the tube constraint; i.e. there exist constants a and b such that for any n-edge 2SAP

$$|Lk(\omega_1, \omega_2)| \leq an + b. \tag{7.0.3}$$

We gave a simple example to show that the upper bound in equation 7.0.3 for 2SAPs can be realized. Hence sufficient conditions for ensuring that the linking probability goes to one are established.

The future work needed regarding the 2SAP model is as follows: Although we proved that the homological linking probability goes to one as $n \to \infty$, unlike the situation for topological linking, we still don't know if it can go to one exponentially rapidly. This is an open question that needs further investigation. Furthermore, in the 2SAP model, in order to increase the likelihood of interpenetration between the two polygons, the polygons are confined to a tube and are constrained to have the same left- and right-most plane. One important question is whether we can weaken these conditions and still have the linking probability going to one. For instance, we may look at pairs of polygons confined to a tube and constrained so that the two polygons share a slice of span of order $n - o(n)$.

In Chapter 5, we were mainly concerned with the following question: *how does the topological entanglement change when a polymer is compressed or stretched under the influence of an external force f?* We proved a pattern theorem for stretched polygons. We used this to also obtain a pattern theorem for loops in a tube. The tube constraint allowed us to prove the pattern theorem for any arbitrary value of f, not just for large values of f as in [27]. The pattern theorem was then used to show that the knotting probability of an n-edge stretched polygon confined to a tube goes to one exponentially as $n \to \infty$. Furthermore, by associating a two-component link to each polygon (loop) in a tube, the incidence of topological linking was examined when the polygon (loop) was under the influence of a force f. We proved that the probability the link associated to an n-edge polygon (loop) is topologically linked approaches unity exponentially as $n \to \infty$. This implies that as $n \to \infty$ when polygons are influenced by an external force f, no matter its strength or direction, topological entanglements as defined by knotting and topological linking are highly probable. In addition, using a transfer-matrix approach, the asymptotic form of the free energy for stretched polygons confined to a tube was obtained. We used this to show that, the average span per edge of a randomly chosen n-edge stretched polygon S_n^f, $\frac{E_f(m(S_n^f))}{n}$, approaches a positive constant as $n \to \infty$ and is non-decreasing in f almost everywhere. We also established that the average number of occurrences of the tight trefoil SAP config P_* per edge in any n-edge stretched polygon S_n^f approaches a positive constant as $n \to \infty$.

The future work needed regarding stretched polygons is as follows: Like the 2SAP model, in addition to the topological linking, the homological linking probability of stretched polygons in a tube can be investigated. Other suitable link invariants can be investigated using the models and methods introduced here to analyze further the effect of an external force on the topological entanglement of ring polymers. We associated a 2-string tangle to a stretched polygon (loop). The tangle model of stretched polygons and suitable tangle invariants can also be used to find out more about the topological entanglement of stretched polygons (loops). Future work on stretched polygons would also involve investigating how the average number of occurrences of the tight trefoil SAP config P_* (per edge) depends on f.

The work presented in Chapter 6 was motivated by the following question: *What is the best measure for characterizing the entanglement complexity of polymers in dense systems and how does this measure depend on various properties of the system?* Needless to say, it is difficult to find the "best" measure for this purpose. So we approached this question by looking for a "good" measure of entanglement. We considered the measure proposed in [41] and asked how its value depends on various properties of SSAWs such as the total number of edges, span, the number of degree one vertices (or, equivalently, twice the number of disjoint walks) and the density. We rigorously proved that the entanglement complexity, as measured in [41], of a polymer system with size n (e.g. the number of edges, span or the number of degree one vertices) is asymptotically (with probability one) bounded below by a linear function of n; i.e. there exists a positive number γ such that the probability that a polymer system of size n has entanglement complexity greater than γn approaches 1 as n goes to infinity. This supports the idea that EC is a good measure of topological entanglement in polymer systems modelled by SSAWs. We also showed that the entanglement complexity of SSAWs of size n is bounded above by a linear function of n. Furthermore, measuring the size by the number of edges, for $N \geq 2$ and $M \geq 2$, we compared the connective constant for SSAWs in an (N, M)-tube with the connective constant for self-avoiding walks in an $(N-2, M-2)$-tube and showed that it is strictly greater than that for SAWs.

Ultimately, based on our theoretical results on the SSAW model we conclude that the following statements (equations) show how EC depends on various properties of SSAWs such as the number of edges, span and the number of degree one vertices and also indicate how these properties are related to each other:

- There exists $\gamma^{\#1}(Y_3) > 0$ such that the probability that the EC of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ from the distribution in equation 6.7.7 (for $j = 1$) is bounded below by $\gamma^{\#1}(Y_3)n$ and bounded above by $a'(N, M)a(N, M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}(\gamma^{\#1}(Y_3)n < EC(X_n^{\#1}(Y_3)) \leq a'(N, M)a(N, M)n) = \lim_{n\to\infty} [1 - e^{-\alpha'(1,Y_3)n+o(n)}] = 1, \tag{7.0.4}$$

where

$$a(N, M) = (1/2)(M + 2N + 1)[(N + 1)(4N - 2) + (9/2)(4N - 3)] \tag{7.0.5}$$

and

$$a'(N,M) = 2N(M+1) + 2M(N+1) + (M+1)(N+1). \tag{7.0.6}$$

In other words, as the span of SSAWs increases one expects EC to be bounded linearly in the span with probability one.

We also showed that the limit inferior of the average EC of span m SSAWs per span is bounded below by a positive constant, i.e.

$$\liminf_{m\to\infty} \frac{E_Y\big(EC(X_m^{\#1}(Y_3))\big)}{m} \geq \gamma_l. \tag{7.0.7}$$

- There exists $\gamma^{\#3}(Y_3) > 0$ such that the probability that the EC of a randomly chosen n-edge SSAW $X_n^{\#3}(Y_3)$ from the distribution in equation 6.7.7 (for $j = 3$) is bounded below by $\gamma^{\#3}(Y_3)n$ and bounded above by $a(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}(\gamma^{\#3}(Y_3)n < EC(X_n^{\#3}(Y_3)) \leq a(N,M)n) = \lim_{n\to\infty}[1 - e^{-\alpha'(3,Y_3)n+o(n)}] = 1. \tag{7.0.8}$$

  In other words, as the number of edges of SSAWs increases one expects EC to be bounded linearly in the number of edges with probability one.

- There exists $\epsilon_P^{\#1}(Y_3) > 0$ such that the probability that the number of degree one vertices of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ is bounded below by $\epsilon_P^{\#1}(Y_3)n$ and bounded above by $a'(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#1}(Y_3)n \leq f(X_n^{\#1}(Y_3),2) \leq a'(N,M)n\big) = \lim_{n\to\infty}[1 - e^{-\alpha''(1,2,Y_3)n+o(n)}] = 1. \tag{7.0.9}$$

  In other words, as the span of SSAWs increases one expects the number of degree one vertices (the number of disjoint walks) to be bounded linearly in the span with probability one.

- There exists $\epsilon_P^{\#1}(Y_3) > 0$ such that the probability that the number of edges of a randomly chosen span n SSAW $X_n^{\#1}(Y_3)$ is bounded below by $\epsilon_P^{\#1}(Y_3)n$ and bounded above by $a'(N,M)n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#1}(Y_3)n \leq f(X_n^{\#1}(Y_3),3) \leq a'(N,M)n\big) = \lim_{n\to\infty}[1-e^{-\alpha''(1,3,Y_3)n+o(n)}] = 1. \tag{7.0.10}$$

In other words, as the span of SSAWs increases one expects the number of edges to be bounded linearly in the span with probability one.

Using the transfer-matrix method, we also established that the average number of edges per unit volume of a randomly chosen span m SSAW, $\frac{E_Y(f(X_m^{\#1}(Y_3),3))}{mNM}$, approaches a positive constant as $m \to \infty$ and is non-decreasing in z almost everywhere. However, it still needs further investigation to see how EC changes with respect to Y_3.

- There exists $\epsilon_P^{\#3}(Y_3) > 0$ such that the probability that span of a randomly chosen SSAW $X_n^{\#3}(Y_3)$ with n edges is bounded below by $\epsilon_P^{\#3}(Y_3)n$ and bounded above by $2n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#3}(Y_3)n \leq f(X_n^{\#3}(Y_3),1) \leq 2n\big) = \lim_{n\to\infty}[1-e^{-\alpha''(3,1,Y_3)n+o(n)}] = 1. \tag{7.0.11}$$

In other words, as the number of edges of SSAWs increases one expects the span to be bounded linearly in the number of edges with probability one.

- There exists $\epsilon_P^{\#3}(Y_3) > 0$ such that the probability that the number of degree one vertices of a randomly chosen SSAW $X_n^{\#3}(Y_3)$ with n edges is bounded below by $\epsilon_P^{\#3}(Y_3)n$ and bounded above by $2n$ goes to one exponentially rapidly as $n \to \infty$; i.e.

$$\lim_{n\to\infty} \mathbb{P}\big(\epsilon_P^{\#3}(Y_3)n \leq f(X_n^{\#3}(Y_3),2) \leq 2n\big) = \lim_{n\to\infty}[1-e^{-\alpha''(3,2,Y_3)n+o(n)}] = 1. \tag{7.0.12}$$

In other words, as the number of edges of SSAWs increases one expects the number of degree one vertices (the number of disjoint walks) to be bounded linearly in the number of edges with probability one.

Furthermore, we obtained the asymptotic form of the free energy for the SSAW model. We also investigated $\rho(N,M;\epsilon)$, the growth constant for SSAWs with limiting edge-density $\frac{\epsilon}{NM}$, as a function of ϵ. We established the existence of this function and showed that it is a continuous and concave function of ϵ and is differentiable almost everywhere in $(\epsilon_{min}, \epsilon_{max})$.

The future work needed regarding the SSAW model is as follows: We were able to obtain theoretical results on how EC depends on various properties of SSAWs such as the number of edges, span and the number of degree one vertices. We also showed how these properties change with respect to each other. However, in order to see how EC changes in terms of the density, we need to know more about the function $\rho(N, M; \epsilon)$, the growth constant for SSAWs with limiting edge-density $\frac{\epsilon}{NM}$. In particular, proving a pattern theorem for this function may lead to some results regarding the connection between EC and the density. We may also investigate the relation between EC and the density of SSAWs by learning more about the z dependence of γ_l in equation 7.0.7.

Appendix A

Transfer-Matrix Method

In this appendix, the transfer-matrix method is briefly introduced. Note that the discussion here is based on the presentation in [54], unless stated otherwise. A *directed graph* or *digraph* D is a triple (V, A, ϕ), where V is a set of *vertices*, A is a set of (directed) edges or *arcs*, and ϕ is a map from A to $V \times V$. If $\phi(e) = (u, v)$, then e is called an edge from u to v, with *initial vertex* u and *final vertex* v. This is denoted by $u = int\ e$ and $v = fin\ e$. If $u = v$ then e is called a *loop*. A *walk* in D of length n from vertex $u \in V$ to vertex $v \in V$ is a sequence $e_1, e_2, ..., e_n$ of n edges such that $u = int\ e_1$, $v = fin\ e_n$, and $fin\ e_i = int\ e_{i+1}$ for $1 \leq i < n$. If also $u = v$, then Γ is called a *closed walk based at* u.

Let $w : A \to \mathbb{C}$ be a *weight function* on A. If $\Gamma = e_1, e_2, ..., e_n$ is a walk then the weight of Γ is defined by $w(\Gamma) = w(e_1)...w(e_n)$. Note that we assume D is *finite* so $V = \{v_1, ..., v_p\}$ and A are finite sets. For any $n \in \mathbb{N}$, let $[n]$ denote the set $\{1, 2, ..., n\}$. For any $i, j \in [n]$, define

$$A_{ij}(n) = \sum_{\Gamma} w(\Gamma), \tag{A.0.1}$$

where the sum is over all walks Γ in D of length n from v_i to v_j. In particular, for $n = 0$, let

$$A_{ij}(0) = \delta_{ij}, \tag{A.0.2}$$

where δ_{ij} is defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } \ i = j \\ 0 & \text{otherwise} \end{cases} . \tag{A.0.3}$$

Let $A(n) = (A_{ij}(n))$. The main concern of the transfer matrix method is the evaluation of $A_{ij}(n)$. Let $p = |V|$. Define a $p \times p$ matrix $B = (B_{ij})$ by

$$B_{ij} = \begin{cases} \sum_{e \in \phi^{-1}((v_i, v_j))} w(e) & \text{if } \ \phi^{-1}((v_i, v_j)) \neq \emptyset \\ 0 & \text{otherwise} \end{cases} . \tag{A.0.4}$$

Note that $B_{ij} = A_{ij}(1)$. The matrix B is called the *adjacency matrix* of D, with respect to the weight function w. The following theorem shows that, for any $n \in \mathbb{N}$, $A(n)$ can be obtained by evaluating the nth power of the matrix B.

Theorem A.0.1. *For any* $n \in \mathbb{N}$,

$$A_{ij}(n) = (B^n)_{ij}. \tag{A.0.5}$$

(Here we define $A^0 = I$ *even if* A *is not invertible.)*

The behaviour of the function $A_{ij}(n)$ can be analyzed through its generating function, i.e.

$$F_{ij}(D, x) = \sum_{n \geq 0} A_{ij}(n) x^n = \sum_{n \geq 0} (B^n)_{ij} x^n. \tag{A.0.6}$$

The following theorem relates the generating function of $A_{ij}(n)$ to the matrix B.

Theorem A.0.2. *The generating function $F_{ij}(D, x)$ is given by*

$$F_{ij}(D, x) = \frac{(-1)^{i+j} \det(I - xB : j, i)}{\det(I - xB)}, \tag{A.0.7}$$

where $(C : j, i)$ denotes the matrix obtained by removing the j-th row and i-th column of C. Thus in particular $F_{ij}(D, x)$ is a rational function of x whose degree is strictly less than n_0, the multiplicity of 0 as an eigenvalue of B.

A matrix or vector A is said to be *non-negative* (*non-positive*) if all its elements are non-negative (non-positive) and we write $A \geq 0$ ($A \leq 0$).

Theorem A.0.3 ([46])**.** *A non-negative matrix always has a non-negative real eigenvalue r such that the modulus of any other eigenvalue of the matrix does not exceed r. To this maximal eigenvalue corresponds an eigenvector with non-negative coordinates.*

A *permutation matrix* is a square matrix that has exactly one entry 1 in each row and each column and has 0's elsewhere. A matrix is called *reducible* if there is a permutation matrix P such that the matrix $P^{-1}AP$ is of the form $\begin{pmatrix} X & 0 \\ Y & Z \end{pmatrix}$, where X and Z are square matrices. Otherwise, we say A is *irreducible*.

Theorem A.0.4 ([46])**.** *A matrix A is irreducible if for each i, j there exists an $m \geq 1$ such that $(A^m)_{ij} > 0$.*

The *period* d of an irreducible matrix A is the greatest common divisor of the integers m for which $(A^m)_{ii} > 0$; A is said to be an *aperiodic* matrix if $d = 1$.

A digraph $D = (V, A, \phi)$ is called *strongly connected* if for each pair of vertices v_i and v_j in V, there exists a walk from v_i to v_j. Let $D = (V, A, \phi)$ be a strongly connected digraph with the weight function $w : A \to \mathbb{C}$ such that $w(a) > 0$ for any $a \in A$. Then B, the adjacency matrix of D, is non-negative and irreducible.

Theorem A.0.5 (Perron-Frobenius Theorem [46])**.** *An irreducible non-negative matrix A always has a positive eigenvalue r that is a simple root of the characteristic* *polynomial of A. The modulus of any other eigenvalue of A does not exceed r. To the maximal eigenvalue r corresponds a positive eigenvector. Moreover, if A has h eigenvalues of modulus r then they are all distinct roots of $x^h - r^h = 0$. Furthermore, if A is aperiodic then r is the only eigenvalue with modulus r.*

Note that the maximal eigenvalue of any matrix A is also called the *spectral radius* of A.

Theorem A.0.6 ([46]). *Increasing any element of a non-negative matrix A does not decrease the maximal eigenvalue. The maximal eigenvalue strictly increases if A is an irreducible matrix.*

In particular,

Theorem A.0.7 ([46]). *The maximal eigenvalue r' of every principle sub-matrix (obtained by removing one row and one column) of a non-negative matrix A does not exceed the maximal eigenvalue r of A. If A is irreducible, then $r' < r$ always holds.*

Theorem A.0.8 (Lemma 9 [1]). *Suppose that $M(x) \geq 0$ is a continuously differentiable matrix valued function of $x > 0$ and let $\rho(x)$ be the spectral radius $\rho(M(x))$. If $\rho(x_0) > 0$ is a simple eigenvalue of $M(x_0)$ and η and ς^T are corresponding eigenvectors, normalized such that $\varsigma^T \eta = 1$, then*

$$\rho'(x_0) = \varsigma^T M'(x_0)\eta \tag{A.0.8}$$

and provided $\rho'(x_0) \neq 0$,

$$(x_0 - x)(\rho(x_0)I - M(x))^{-1} \to \frac{1}{\rho'(x_0)}\eta\varsigma^T = (\varsigma^T M'(x_0)\eta)^{-1}\eta\varsigma^T \tag{A.0.9}$$

as $x \to x_0$.

Appendix B

Convex Functions

Convex and concave functions have a significant role in the study of the free energies of various models in statistical mechanics. If $f(x)$ is a concave function then $-f(x)$ is convex. So all the results presented here can be easily reworded to obtain similar results for concave function. In this appendix, the main properties of convex functions are reviewed. The discussion here is based on the presentation in [26]. More detailed information about the properties of convex functions can be found in [21] and [45].

A function is said to be *convex* on a closed interval $[a, b] \subset \mathbb{R}$ if

$$\lambda f(x) + (1-\lambda)f(y) \geq f(\lambda x + (1-\lambda)y), \tag{B.0.1}$$

for every $a \leq x < y \leq b$ whenever $0 \leq \lambda \leq 1$. Convex functions are continuous under some conditions; these conditions are explained in the next lemma.

Lemma B.0.9. *If $f(x)$ is a convex function in $[a, b]$, and $f(x)$ is bounded above in some open interval $I \subseteq (a, b)$, then there exists an interval $(c, d) \supseteq I$ such that $f(x)$ is bounded above and continuous in (c, d), and $f(x) = +\infty$ in $(a, b) \setminus (c, d)$.*

In order to prove that a function f is convex, it is usually easier to first show that f satisfies the relation

$$f(x) + f(y) \geq 2f((x + y)/2), \tag{B.0.2}$$

for any x and y. Then, the next lemma can be used to show that the function is convex.

Lemma B.0.10. *Suppose that $f(x)+f(y) \geq 2f((x+y)/2)$ for all $a \leq x \leq y \leq b$. If $f(x)$ is bounded above in some open interval $I \subset (a, b)$, then there exists an interval $(c, d) \supseteq I$ such that $f(x)$ is convex and bounded above in (c, d), and $f(x) = +\infty$ in $(a, b) \setminus (c, d)$.*

The following lemma shows that finite convex functions have left- and right-derivatives everywhere.

Lemma B.0.11. *If $f : [a, b] \to \mathbb{R}$ is a convex function and is finite in (a, b), then $d^- f(x)/dx$ and $d^+ f(x)/dx$ exist everywhere in (a, b). Moreover,*

$$-\infty < d^- f(x)/dx \leq d^+ f(x)/dx < \infty \tag{B.0.3}$$

and both the left- and right-derivatives are non-decreasing with increasing x.

Let $\{f_n\}_{n\geq 1}$ be a sequence of functions $f_n : [a, b] \to \mathbb{R}$, $n = 1, 2, ...$ The sequence is said to *converge pointwise* to a function $f : [a, b] \to \mathbb{R}$ if the sequence of real numbers $\{f_n(x)\}_{n\geq 1}$ converges to $f(x)$ for any $x \in [a, b]$. The following lemma shows that convergent sequences of convex functions have a convex limit.

Lemma B.0.12. *Suppose that $f_n : [a, b] \to \mathbb{R}$, $n = 1, 2, ...,$ is a sequence of convex functions converging pointwise to a limit $f : [a, b] \to \mathbb{R}$. Then $f(x)$ is convex in $[a, b]$.*

Convex functions have another property which is important in applications. They are differentiable almost everywhere, i.e. they are differentiable everywhere except, possibly, on a set with measure zero.

Lemma B.0.13. *Let f be a non-decreasing, real-valued function on $[a, b]$. Then f is differentiable almost everywhere.*

Lemma B.0.14. *Suppose that the sequence of convex functions $\{f_n\}_{n\geq 1}$ approaches a limit f. Then the right- and left-derivatives satisfy*

$$\frac{d^-}{dx} f(x) \leq \liminf_{n\to\infty} \frac{d^-}{dx} f_n(x) \leq \limsup_{n\to\infty} \frac{d^+}{dx} f_n(x) \leq \frac{d^+}{dx} f(x). \tag{B.0.4}$$

Moreover, this implies that

$$\lim_{n\to\infty} \frac{d}{dx} f_n(x) = \frac{d}{dx} f(x) \tag{B.0.5}$$

almost everywhere.

References

[1] S. E. Alm and S. Janson, Random self-avoiding walks on one-dimensional lattices, *Commun. Statist.-Stochastic Models* **6** (1990) 169-212.

[2] J. Arsuaga, T. Blackstone, Y. Diao, E. Karadayi and M Saito, Linking of uniform random polygons in confined spaces, *J. Phys. A: Math. Theor.* **40** (2007) 1925-1936.

[3] N. W. Ashcroft and N.D. Mermin, *Solid State Physics*, Rinehart and Winston (1976).

[4] M. Atapour, C. E. Soteros, C. Ernst and S. G. Whittington, The linking probability for 2-component links which span a lattice tube, *Journal of Knot Theory and Its Ramifications*, submitted (2008).

[5] N. Clisby, R. Liang and G. Slade, Self-avoiding walk enumeration via the lace expansion, *J. Phys. A: Math. Theor.* **40** (2007). 10973-11017.

[6] R. H. Crowell and R.H. Fox, Introduction to knot theory, Boston: Ginn (1963).

[7] P. R. Cromwell, Knots and Links, Cambridge University Press (2004).

[8] F. B. Dean and N. R. Cozzarelli, *J. Biol. Chem.* **260** (1985) 4984.

[9] M. Delbruck, Mathematical Problems in the Biological Sciences, *Proceeding of the Symposium on Applied Mathematics, AMS* Providence RI **14** (1962) 55-63.

[10] Y. Diao, N. Pippenger and D. W. Sumners, On random knots, *Journal of Knot Theory and Its Ramifications* **3** (1994) 419-429.

[11] Y. Diao and C. Ernst, The complexity of lattice knots, *Topology and Its Applications* **90** No. 1 (1998) 1-9.

[12] Y. Diao, C. Ernst and E. J. Janse van Rensburg, Upper bounds on linking numbers of thick links, *Journal of Knot Theory and Its Ramifications* **11** No. 2 (2002) 199-210.

[13] R. Dickman and C. K. Hall, *J. Chem. Phys.*, **89** (1988) 3168-3174.

[14] S. F. Edwards, Statistical mechanics with topological constraints : I, *Proc. Phys. Soc.* **91** (1967) 513-519.

[15] O. Farago, Y. Kantor and M. Kardar, Pulling knotted polymers, *Europhys. Lett.*, **60** (1) (2002) 5359.

[16] H. L. Frisch and E. Wasserman, Chemical topology, *J. Am Chem Soc.* **83** (1961) 3789-3795.

[17] P. G. de Gennes, Tight knots, *Macromol.* **17** (1984) 703-704.

[18] J. M. Hammersley and K. W. Morton, Poor man's Monte Carlo, *J. R. Stat. Soc. B* **16** (1954) 23-38.

[19] J. M. Hammersley, The number of polygons on a lattice, *Proc. Camb. Phil. Soc.* **57** (1961) 516-523.

[20] J.M. Hammersley, On the rate of convergence to the connective constant of the hypercubical lattice, *Quart. J. Math.* Oxford (2) **12** (1961), 250-256.

[21] G.H. Hardy, J.E. Littlewood and G. Pòlya, Inequalities, second edition, Cambridge University Press. (1952).

[22] B. Hudson and J. Vinogad, *Nature* **1** (1967) 216-647.

[23] E. J. Janse van Rensburg and S. G. Whittington, The knot probability in lattice polygons, *J. Phys. A: Math. Gen.* **23** (1990) 3573-3590.

[24] E. J. Janse van Rensburg, E Orlandini, D. W. Sumners, M. C. Tesi and S. G. Whittington, The writhe of a self-avoiding polygon, *J. Phys. A: Math. Gen.* **26** (1993) L981-L986.

[25] E. J. Janse van Rensburg, E. Orlandini, D. W. Sumners, M. C. Tesi and S. G. Whittington, Topology and geometry of biopolymers, *IMA Vol. Math. Appl.* eds. J. P. Mesirov, K. Schulten and D. W. Sumners **82** (Springer, New York, 1996) 21-37.

[26] E. J. Janse van Rensburg, The Statistical Mechanics of Interacting Walks, Polygons, Animals and Vesicles, Oxford University Press (2000).

[27] E. J. Janse van Rensburg, E. Orlandini, M. C. Tesi and S. G. Whittington, Knotting in stretched polygons, *J. Phys. A: Math. Theor.* **41** (2008).

[28] H. Kesten, On the number of self-avoiding walks, *J. Math. Phys.* **4** (1963) 960-969.

[29] K. Knoniaris and M. Muthukumar, Self-entanglement in ring polymers, *J. Chem. Phys.* **95** No. 4 (1991) 2873-2881.

[30] W. B. R. Lickorish, Prime knots and tangles, *Trans. AMS* **267** No. 1 (1981) 321-332.

[31] C. Livingston, Knot theory, Carus Mathematical Monographs **24** AMS (1993).

[32] N. Madras and G. Slade, The Self-Avoiding Walk, Boston: Birkhäuser (1993).

[33] N. Madras, A pattern theorem for lattice clusters, *Annals of Combinatorics* **3** (1999) 357-384.

[34] N. Madras, C. E. Soteros and S. G. Whittington, Statistics of lattice animals, *J. Phys. A: Math. Gen.*, **21** (1988) 4617-4635.

[35] J. P. J. Michels and F. W. Wiegel, On the topology of a polymer ring, *Proc. R. Soc. Lond. A* **403** No. 1825 (1986) 269-284.

[36] K. Murasugi, Knot theory and Its Applications, Birkhauser Boston (1996).

[37] Y. Nakanishi, Primeness of links, *Math. Sem. Notes Kobe Univ.* **9** (1981) 415-440.

[38] Y. Nakanishi, Prime and simple links, *Math. Sem. Notes Kobe Univ.* **11** (1983) 249-256.

[39] E. Orlandini, E. J. Janse van Rensburg, M. C. Tesi and S. G. Whittington, Random linking of lattice polygons, *J. Phys. A: Math. Gen.* **27** (1994) 335-345.

[40] E. Orlandini, M. C. Tesi and S. G. Whittington, Polymer entanglement in melts, *J. Phys. A: Math. Gen.* **33** (2000) L181-L186.

[41] E. Orlandini and S. G. Whittington, Entangled polymers in condensed phases, *J. Chem. Phys.* **121** (2004) 12094-12099.

[42] E. Orlandini and S. G. Whittington, Statistical topology of closed curves: Some applications in polymer physics, *Reviews of Modern Physics* **79** No. 2 (2007) 611-642.

[43] N. Pippenger, Knots in random walks, *Discrete Applied Math.* **25** (1989) 273-278.

[44] D. Rolfsen, Knots and Links (Publish or Perish, 1976) or (AMS Chelsea Publishing, 2003).

[45] H. L. Royden, Real Analysis, third edition, Prentice Hall (Englewood-Cliffs, New Jersey) (1987).

[46] H. H. Schaefer, Banach lattices and positive operators, Springer-Verlag, Berlin and New York (1974).

[47] S. Y. Shaw and J. C. Wang, *Science* **260** (1993) 533.

[48] C. E. Soteros and S. G. Whittington, Polygons and stars in a slit geometry, *J. Phys. A: Math. Gen.* **21** (1988) 857-861.

[49] C. E. Soteros and S. G. Whittington, Lattice models of branched polymers: effects of geometrical constraints, *J. Phys. A: Math. Gen.* **22** (1989) 5259-5270.

[50] C. E. Soteros, D. W. Sumners and S. G. Whittington, Entanglement complexity of graphs in $\mathbb{Z}^3$, *Math. Proc. Camb. Phil. Soc.* **111** (1992) 75-91.

[51] C. E. Soteros, Knots in graphs in subsets of $\mathbb{Z}^3$, *IMA Vol. in Math. and its Appl.* eds. S. G. Whittingon, D. W. Sumners and T. Lodge **103** (Springer, New York, 1998) 101-133.

[52] C. E. Soteros, D. W. Sumners and S. G. Whittington, Linking of p-spheres in $\mathbb{Z}^d$, *Journal of Knot Theory and Its Ramifications* **8** No. 1 (1999) 49-70.

[53] C. E. Soteros, Eulerian graph embeddings and trails confined to lattice tubes, *Journal of Physics: Conference Series* **42** (2006) 258–267.

[54] R. P. Stanley, *Enumerative Combinatorics*, The Wadsworth and Brooks/Cole Mathematics Series, Volume 1 (1986).

[55] D. W. Sumners and S. G. Whittington, Knots in self-avoiding walks, *J. Phys. A: Math. Gen.* **21** (1988) 1689-1694.

[56] D. W. Sumners, C. Ernst, S. J. Spengler and N. R. Cozzarelli, Analysis of the mechanism of DNA recombination using tangles, *Q Rev Biophys.* **28** No. 3 (1995) 253-313.

[57] O. Sundin and A. Varshavskv, *Cell* **25** (1981) 659.

[58] M. C. Tesi, E. J. Janse van Rensburg, E. Orlandini and S. G. Whittington, Topological entanglement complexity of polymer chains in confined geometries, *IMA Vol. in Math. and its Appl.* eds. S. G. Whittingon, D. W. Sumners and T. Lodge **103** (Springer, New York, 1998) 135-157.

[59] A. V. Vologodskii, A. V. Lukashin, M. D. Frank-Kamenetskii and V. V. Anshelevich, The knot problem in statistical mechanics of polymer chains, *Sov. Phys.-JETP* **39** (1974) 1059-1063.

[60] S. A. Wassetman, J. M. Dungan and N. R. Cozzarelli, Discovery of a predicted DNA knot substantiates a model for site-specific recombination, *Science* **229** (1985) 171-174.

[61] S. A. Wasserman and N. R. Cozzarelli, Biochemical topology: applications to DNA recombination and replication, *Science* **232** (1986) 951-960.

[62] S.A. Wasserman and N.R. Cozzarelli, *J. Biol. Chem.* **266** (1991) 20567.

[63] S.G. Whittington, Self-avoiding walks with geometrical constraints, *J. Stat. Phys.* **30** (1983) 449-456.

[64] J. B. Wilker and S. G. Whittington, Extension of a theorem on supermultiplicative functions, *J. Phys. A: Math. Gen.* **12** No. 10 (1979) L245-L247.