Topology of random linkages

MICHAEL FARBER

Betti numbers of configuration spaces of mechanical linkages (known also as polygon spaces) depend on a large number of parameters – the lengths of the bars of the linkage. Motivated by applications in topological robotics, statistical shape theory and molecular biology, we view these lengths as random variables and study asymptotic values of the average Betti numbers as the number of links n tends to infinity. We establish a surprising fact that for a reasonably ample class of sequences of probability measures the asymptotic values of the average Betti numbers are independent of the choice of the measure. The main results of the paper apply to planar linkages as well as for linkages in $\mathbb{R}^3$. We also prove results about higher moments of Betti numbers.

1 Introduction

Configuration spaces of mechanical systems which appear in topological robotics depend typically on many parameters which are only partially known and can be viewed as random variables. The configuration space, which depends essentially on the values of the parameters, can be viewed in such a situation as a *random topological space*. Betti numbers of the configuration space are then random functions and information about their mathematical expectations and other statistical characteristics may have practical importance in various applications.

Interesting examples of such random topological spaces are provided by configuration spaces of mechanical linkages, the main object of study in this paper. A *linkage* is a simple mechanism consisting of n bars in $\mathbb{R}^3$ having fixed lengths $l_1, \ldots, l_n$ which are cyclically connected by revolving joints forming a closed polygonal chain. Angles between bars of the linkage may vary, the only condition is that the links do not become disconnected from each other. We consider a pair of configurations of the linkage as being identical if one can be obtained from the other by a rigid motion of the space $\mathbb{R}^3$. The configuration space of the linkage, denoted in this paper by

$$(1) \qquad N_\ell = \left\{ (u_1, \ldots, u_n) \in S^2 \times \cdots \times S^2; \ \sum_{i=1}^{n} l_i u_i = 0 \in \mathbb{R}^3 \right\} / \mathrm{SO}(3),$$
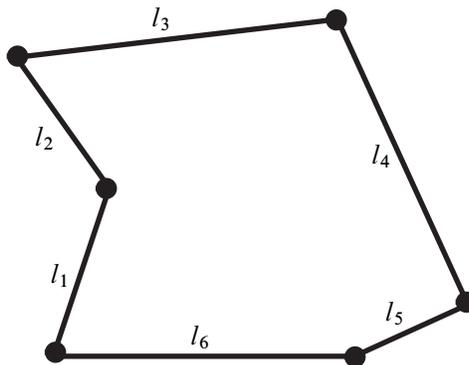
Figure 1: Linkage

parameterizes all possible configurations. Here $\ell = (l_1, l_2, \ldots, l_n) \in \mathbb{R}_+^n$ is the $n-$tuple of the bar lengths, called the *length vector*[1]. Spaces N_ℓ are also known as *polygon spaces* as they parameterize shapes of all $n-$gons in $\mathbb{R}^3$ having sides of length $l_1, \ldots, l_n$.

The spaces N_ℓ appear in molecular biology where they represent shapes of long molecules. Clearly, information about topological properties of these spaces may lead to interesting new effects in molecular and chemical design. Statistical shape theory (see eg Kendall et al [10]) is another area where the spaces N_ℓ play an interesting role: they describe the space of shapes having certain geometric properties with respect to the central point. Having in mind these applications it is quite natural to assume that the number of links n is large, $n \to \infty$, and that the numbers $l_i > 0$ are not entirely known or are known with some random error.

Let us now give some basic facts concerning the topology of N_ℓ and its dependence on the length vector $\ell \in \mathbb{R}_+^3$. For a *generic* ℓ (this term is explained below), N_ℓ is a compact smooth manifold of dimension $2(n-3)$. If ℓ is not generic then N_ℓ is a compact manifold with singularities. Clearly, $N_\ell = N_{t\ell}$ for any $t > 0$. Hence we may consider ℓ as lying in the quotient space Δ^{n-1} of $\mathbb{R}_+^n$ modulo the action of $\mathbb{R}_+$ by scalar multiplication. Clearly, Δ^{n-1} can be identified with the interior of the standard simplex, ie the set given by the inequalities $l_1 > 0, \ldots, l_n > 0$ and $\sum l_i = 1$.

It is easy to see that N_ℓ is diffeomorphic to $N_{\ell'}$ if ℓ' is obtained from ℓ by permuting coordinates. Let Σ_n denote the permutation group of n symbols. Clearly Σ_n acts on

---

[1]In this paper $\mathbb{R}_+^n \subset \mathbb{R}^n$ denotes the set of all points $(l_1, \ldots, l_n)$ having positive coordinates $l_1 > 0, \ldots, l_n > 0$.

$\mathbb{R}^n$ and on Δ^{n-1} permuting coordinates and the manifold N_ℓ depends only on the Σ_n–orbit of the vector ℓ.

To explain further the character of dependence of N_ℓ on ℓ we need to introduce the concept of a *chamber*. For any subset $J \subset \{1,\dots,n\}$ we denote by $H_J \subset \mathbb{R}^n$ the hyperplane defined by the equation

(2)
$$\sum_{i \in J} l_i = \sum_{i \notin J} l_i.$$

The connected components of the complements

$$\mathbb{R}^n_+ - \bigcup_J H_J$$

are called *chambers*. *Generic* length vectors are defined as those lying in chambers, not on hyperplanes H_J. It is an interesting combinatorial problem to find or estimate the number c_n of different Σ_n–orbits of chambers for large n. The numbers c_n are known for $n \le 9$ (see Hausmann and Rodriguez [8]):

n	3	4	5	6	7	8	9
c_n	2	3	7	21	135	2470	175428

We see that the number c_n grows very fast. The diffeomorphism type of N_ℓ is constant when ℓ varies inside a chamber. One of the main results of Farber, Hausmann and Schütz [4] states that for a given *n the map $\ell \mapsto N_\ell$ gives a one-to-one correspondence between the Σ_n–orbits of chambers and the diffeomorphism types of manifolds N_ℓ where $\ell \in \mathbb{R}^n_+$ is generic.*

The following picture summarizes our description of *the field of topological spaces* $\ell \mapsto N_\ell$ viewed as a single object. The open simplex Δ^{n-1} is divided into a huge number of tiny chambers, each representing a diffeomorphism type of manifolds N_ℓ. The symmetric group Σ_n acts on the simplex Δ^{n-1} mapping chambers to chambers and for $n \ne 4$ manifolds N_ℓ and $N_{\ell'}$ are diffeomorphic if and only if the vectors ℓ, ℓ' lie in chambers belonging the same Σ_n–orbit.

The main idea of this work is to use methods of probability theory and statistics in dealing with the variety of diffeomorphism types of configuration spaces N_ℓ for n large. In applications, different manifolds N_ℓ appear with different probabilities and our intention is to study the most "frequently emerging" manifolds N_ℓ and the mathematical expectations of their topological invariants. Formally, we view the length vector $\ell \in \Delta^{n-1}$ as a random variable whose statistical behavior is characterized by a probability measure ν_n. Topological invariants of N_ℓ become random functions and

their mathematical expectations might be very useful for applications. Thus, one is led to study the *average* or *expected* Betti numbers[2]

$$(3) \qquad \mathrm{E}(b_{2p}(N_\ell)) = \int_{\Delta^{n-1}} b_{2p}(N_\ell) dv_n$$

where the integration is understood with respect to ℓ. One of the main results of this paper states that for p fixed and n large this average $2p$–dimensional Betti number can be calculated explicitly up to an exponentially small error. More precisely, we prove that

$$\mathrm{E}(b_{2p}(N_\ell)) = \int_{\Delta^{n-1}} b_{2p}(N_\ell) dv_n \sim \sum_{i=0}^{p} \binom{n-1}{i}.$$

It might appear surprising that the asymptotic value of the average Betti number $b_{2p}(N_\ell)$ does not depend of the sequence of probability measures v_n which are allowed to vary in an ample class of *admissible* probability measures described in Section 2 below.

We also find the asymptotics of the average Betti numbers $b_p(M_\ell)$ of configuration spaces of *planar* polygon spaces

$$(4) \qquad M_\ell = \left\{ (u_1, \ldots, u_n) \in S^1 \times \cdots \times S^1; \ \sum_{i=1}^{n} l_i u_i = 0 \in \mathbb{R}^2 \right\} \Big/ \mathrm{SO}(2).$$

In the paper [5] we calculated the asymptotic values of the average Betti numbers

$$(5) \qquad \mathrm{E}(b_p(M_\ell)) = \int_{\Delta^{n-1}} b_p(M_\ell) dv_n$$

for two sequences of probability measures v_n on $\mathbb{R}_+^n$. It was discovered in [5] that for large n the answers for these two distinct measures were equal. The present paper explains this *universality phenomenon*. We also compute asymptotics of the moments

$$(6) \qquad \int_{\Delta^{n-1}} b_{2p}(N_\ell)^k dv_n, \qquad \int_{\Delta^{n-1}} b_p(M_\ell)^k dv_n,$$

where $k = 1, 2, 3, \ldots$ assuming that n tends to infinity.

In this paper we employ a method different from the one used in [5]: instead of dealing with explicit expressions for Betti numbers we specify a domain in the simplex of parameters where the behavior of Betti numbers can easily be described and, moreover, the volume of the complement of the domain is exponentially small. The proofs presented below are shorter than in [5] although theorems of the present paper are more general in several respects: (a) they allow more general class of measures, (b) in this

---

[2]It is well known that all odd-dimensional Betti numbers of N_ℓ vanish; see Klyachko [11].

paper we treat both cases of planar and spatial linkages and (c) the present approach is applicable to higher moments as well.

Here are a few comments on the history of the problem. Polygon spaces were studied by K Walker [13], M Kapovich and J Millson [9] and others. Betti numbers of the spaces N_ℓ were first described by A A Klyachko [11] who used methods of algebraic geometry. J-Cl Hausmann and A Knutson [7] applied methods of symplectic topology (toric varieties) to study the cohomology algebras $H^*(N_\ell)$. Betti numbers of planar polygon spaces M_ℓ as functions of the length vector ℓ were found in Farber and Schütz [6]; this result uses techniques of Morse theory of manifolds with involution. The recent preprint of the author, Hausmann and Schütz [4] gives a general classification of diffeomorphism types of polygon spaces M_ℓ and N_ℓ in terms of combinatorics of chambers and the action of the symmetric group Σ_n.

2 Statements of the main results

To state the main results of this paper we need to define what is meant by an *admissible sequence of measures*.

For a vector $\ell = (l_1, \ldots, l_n)$ we denote by $|\ell| = \max\{|l_1|, \ldots, |l_n|\}$ the maximum of absolute values of coordinates. The symbol Δ^{n-1} denotes the open unit simplex, ie the set of all vectors $\ell = (l_1, l_2, \ldots, l_n) \in \mathbb{R}^n$ such that $l_i > 0$ and $l_1 + \cdots + l_n = 1$. Let μ_n denote the Lebesgue measure on Δ^{n-1} normalized so that $\mu_n(\Delta^{n-1}) = 1$. In other words, for a Lebesgue measurable subset $A \subset \Delta^{n-1}$ one has

$$\mu_n(A) = \frac{\text{vol}(A)}{\text{vol}(\Delta^{n-1})}$$

where the symbol vol denotes the $(n-1)$–dimensional volume.

For an integer $p \geq 1$ we denote by Λ_p the set

$$(7) \qquad\qquad \Lambda_p = \{\ell \in \Delta^{n-1}; |\ell| \geq (2p)^{-1}\}.$$

Clearly, $\Lambda_p \subset \Lambda_q$ for $p \leq q$ and $\Lambda_p = \Delta^{n-1}$ for $2p \geq n$.

It will be helpful to think of p being fixed and of n being large, say, tending to ∞. The set Λ_p is shown on the left in Figure 2. It is the union of n domains Λ_p^i defined as $\Lambda_p^i = \{\ell = (l_1, \ldots, l_n) \in \Delta^{n-1}; l_i \geq \frac{1}{2p}\}$, where $i = 1, \ldots, n$. Each Λ_p^i is homothetic to Δ^{n-1} with coefficient $(1 - \frac{1}{2p})$ and hence $\mu_n(\Lambda_p^i) = (1 - \frac{1}{2p})^{n-1}$. It follows that

$$(8) \qquad\qquad \mu_n(\Lambda_p) \leq n \cdot \left(1 - \frac{1}{2p}\right)^{n-1},$$

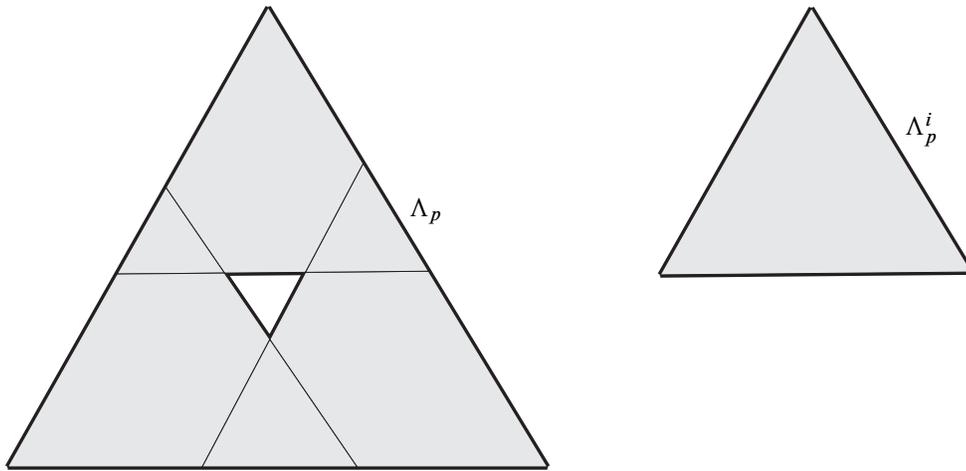ie the normalized Lebesgue measure of Λ_p is exponentially small for large n.

Figure 2

Definition 2.1 Consider a sequence of probability measures ν_n on Δ^{n-1} where $n = 1, 2, \ldots$. It is called *admissible* if $\nu_n = f_n \cdot \mu_n$ where $f_n : \Delta^{n-1} \to \mathbb{R}$ is a sequence of Lebesgue measurable functions satisfying: (i) $f_n \geq 0$, (ii) $\int_{\Delta^{n-1}} f_n d\mu_n = 1$, and (iii) for any $p \geq 1$ there exist constants $A > 0$ and $0 < b < 2$ such that

$$(9) \qquad\qquad f_n(\ell) \leq A \cdot b^n$$

for any n and any $\ell \in \Lambda_p \subset \Delta^{n-1}$.

Note that property (iii) imposes restrictions on the behavior of the sequence ν_n only in domains Λ_p^{n-1}.

Example Consider the unit cube $\square^n \subset \mathbb{R}_+^n$ given by the inequalities $0 \leq l_i \leq 1$ for $i = 1, \ldots, n$. Let χ_n be the probability measure on $\mathbb{R}_+^n$ supported on $\square^n \subset \mathbb{R}_+^n$ such that the restriction $\chi_n|_{\square^n}$ is the Lebesgue measure, $\chi_n(\square^n) = 1$. Consider the sequence of induced measures

$$(10) \qquad\qquad \nu_n = q_*(\chi_n)$$

on simplices Δ^{n-1} where $q : \mathbb{R}_+^n \to \Delta^{n-1}$ is the normalization map $q(\ell) = t\ell$ where $t = (l_1 + \cdots + l_n)^{-1}$. The measures ν_n have a very clear geometric meaning: it is the probability distribution in the case when the bar lengths l_i are independent and are uniformly distributed in the unit interval $[0, 1]$ [5]. The goal of the following arguments is to show that the sequence (10) is admissible.

It is easy to see that

$$\nu_n = f_n \mu_n$$

where $f_n: \Delta^{n-1} \to \mathbb{R}$ is given by

$$(11) \qquad f_n(\ell) = k_n \cdot |\ell|^{-n}, \quad \ell \in \Delta^{n-1}, \quad k_n \in \mathbb{R}.$$

Indeed, consider coordinates $0 \le y_1, \ldots, y_n \le 1$ in $\square^n$ and another coordinate system in $\square^n$ is given by variables $z_1, \ldots, z_{n-1}, t$ satisfying $z_i \ge 0$, $z_1 + \cdots + z_{n-1} \le 1$ and $0 \le t \le |z|^{-1}$, where $z = (z_1, \ldots, z_n)$ and z_n is an auxiliary variable given by $z_n = 1 - z_1 - \cdots - z_{n-1}$. These two coordinate systems are related by $y_i = t z_i$, where $i = 1, \ldots, n$. A simple calculation shows that

$$dy_1 \wedge dy_2 \wedge \ldots dy_n = t^{n-1} dz_1 \wedge \ldots dz_{n-1} \wedge dt$$

leading to formula (11). The constant k_n which appears in (11) can be found (using (ii) of Definition 2.1) from the equation

$$(12) \qquad k_n^{-1} = \int_{\Delta^{n-1}} |\ell|^{-n} d\mu_n.$$

If $\ell \in \Lambda_p^{n-1}$ then $f_n(\ell) \le k_n \cdot (2p)^n$. Using (8) and (12) we find that

$$k_n^{-1} \ge (2p)^n \cdot \left(1 - n \left(\frac{2p-1}{2p}\right)^{n-1}\right).$$

This shows that the sequence $(2p)^n k_n$ remains bounded as $n \to \infty$ implying (iii) of Definition 2.1. Hence, the sequence of measures $\{\nu_n\}$ is admissible.

Next we state main theorems of this paper.

Theorem 2.2 *Fix an admissible sequence of probability measures ν_n and an integer $p \ge 0$, and consider the $2p$–dimensional Betti number (3) of polygon spaces N_ℓ in $\mathbb{R}^3$ as a random variable on Δ^{n-1}, for large $n \to \infty$. Then there exist constants $C > 0$ and $0 < a < 1$ (depending on the sequence of measures ν_n and on the number p but independent of n) such that the average Betti numbers (3) satisfy*

$$(13) \qquad \left| \int_{\Delta^{n-1}} b_{2p}(N_\ell) d\nu_n - \sum_{i=0}^{p} \binom{n-1}{i} \right| < C \cdot a^n$$

for all n.

Theorem 2.3 *Fix an admissible sequence of probability measures ν_n and an integer $p \ge 0$, and consider the average p–dimensional Betti number (5) of planar polygon*

spaces for large $n \to \infty$. Then there exist constants $C > 0$ and $0 < a < 1$ (depending on the sequence of measures ν_n and on the number p but independent of n) such that

$$(14) \qquad \left| \int_{\Delta^{n-1}} b_p(M_\ell) d\nu_n - \binom{n-1}{p} \right| < C \cdot a^n$$

for all n.

3 Volume of the frustum of a simplex

In this section we obtain a formula for the volume of the intersection of a simplex with a half-space. General formulae of this kind are well known; see Curry and Schoenberg [2], Ali [1], Varsi [12] and Farber and Kappeler [5]. However in this paper we have to consider a highly degenerate situation when the functional determining the half-space takes only two distinct values of the set of vertices. We give an explicit formula and its proof in this special case.

Lemma 3.1 *Let $b_0, \ldots, b_n \in \mathbb{R}^n$ be vertices of a simplex $\Delta^n \subset \mathbb{R}^n$. Let $\phi \colon \mathbb{R}^n \to \mathbb{R}$ be an affine functional such that $\phi(b_i) = -1$ for $i = 0, \ldots, p-1$ and $\phi(b_i) = 1$ for $i = p, p+1, \ldots, n$. For $x \in \mathbb{R}$ denote by H_x the half-space $H_x = \{v \in \mathbb{R}^n ; \phi(v) \le x\}$. Then the ratio*

$$r(x) = \frac{\mathrm{vol}(H_x \cap \Delta)}{\mathrm{vol}(\Delta)}$$

for $x \in [-1, 1]$ is given by

$$(15) \qquad r(x) = \left(\frac{x+1}{2} \right)^q \cdot \sum_{k=0}^{p-1} \binom{q-1+k}{q-1} \cdot \left(\frac{1-x}{2} \right)^k.$$

Here $q = n - p + 1$ denotes the multiplicity of value $1 = \phi(b_i)$.

Proof The function $r(x)$ is closely related to spline functions which were introduced in Section 3 of Curry and Schoenberg [2]. Clearly $r(x)$ vanishes for $x \le -1$ and $r(x)$ is identically 1 for $x \ge 1$. Moreover, from geometrical considerations [2, Section 3] one knows that:

(a) $r(x)$ is a polynomial of degree n for $x \in [-1, 1]$,

(b) $r(x)$ is of class C^{n-p} near -1,

(c) $r(x)$ is of class C^{n-q} near 1.

Hence we obtain that

$$r^{(j)}(-1) = 0 \quad \text{for} \quad j = 0, 1, \ldots, n - p$$

and
$$(r(x) - 1)^{(j)}(1) = 0 \quad \text{for} \quad j = 0, 1, \ldots, n - q.$$

The first equation implies that $r(x) = (x + 1)^{n-p+1} g(x)$ where $g(x)$ is polynomial of degree $p - 1$. To satisfy the second equation, we must take for $g(x)$ the sum of terms of degree $\leq p - 1$ of Taylor expansion at $x = 1$ of the rational function $(x + 1)^{p-n-1}$. Computing this Taylor expansion we obtain (15). □

Corollary 3.2 *Under conditions of Lemma 3.1 the number* $r_{p,q} = r(0)$ *is given by*

$$(16) \qquad r_{p,q} = 2^{-q} \cdot \sum_{k=0}^{p-1} \binom{q-1+k}{k} \cdot 2^{-k}$$

The number $r_{p,q}$ has a clear geometric meaning. Divide the set of vertices of Δ^n into two subsets, one containing p elements and another containing q elements, here $p + q = n + 1$. Let F and F' denote simplices generated by each of these subsets. Any point of Δ^n lies on a unique segment connecting a point of F with a point of F'. Then $r_{p,q}$ is the relative volume of the set of points $x \in \Delta^n$ such that $x = (1-t)a + ta'$ where $a \in F$, $a' \in F'$ and $t \in [0, 1/2]$.

4   Proof of Theorem 2.2

We start by recalling the description of Betti numbers of polygon spaces N_ℓ given by A A Klyachko [11] and by J-Cl Hausmann and A Knutson [7]. A subset $J \subset \{1, \ldots, n\}$ is called *short* if

$$\sum_{i \in J} l_i < \sum_{i \notin J} l_i.$$

A subset is called *long* if its complement is short. Given a generic length vector we denote by $\alpha_j(\ell)$ the number of subsets $J \subset \{1, \ldots, n-1\}$ such that $|J| = j$ and the set $J \cup \{n\}$ is short. The result of Corollary 4.3 from [7] can be equivalently stated as

$$(17) \qquad b_{2p}(N_\ell) = \sum_{j=0}^{p} \left[\alpha_j(\ell) - \alpha_{n-j-2}(\ell) \right].$$

We will need the following upper bound.

Proposition 4.1 *For a generic length vector* $\ell = (l_1, \ldots, l_n) \in \Delta^{n-1}$ *with* $n \geq 3$ *one has*

(18) $$b_{2p}(N_\ell) \leq 2 \cdot (n-1)^p.$$

Proof From (17) one obtains

$$b_{2p}(N_\ell) \leq \sum_{j=0}^{p} \alpha_j(\ell) \leq \sum_{j=0}^{p} \binom{n-1}{j}$$

$$\leq \sum_{j=0}^{p} (n-1)^j = \frac{(n-1)^{p+1} - 1}{n-2} \leq \frac{(n-1)^{p+1}}{n-2} \leq 2 \cdot (n-1)^p.$$

On the last step we used the assumption $n \geq 3$. $\qquad\square$

Given an integer p we denote by

(19) $$\Gamma_p = \Gamma_p^{n-1} \subset \Delta^{n-1}$$

the set of all length vectors $\ell = (l_1, \ldots, l_n) \in \Delta^{n-1}$ such that any subset $J \subset \{1, \ldots, n\}$ of cardinality $|J| = p$ is short with respect to ℓ. Clearly $\Gamma_p \supset \Gamma_q$ if $p \leq q$ and, moreover, $\Gamma_p = \varnothing$ for $p \geq n/2$.

Proposition 4.2 *If* $\ell \in \Gamma_p \subset \Delta^{n-1}$ *is generic then for all* $q \leq p-2$ *one has*

(20) $$b_{2q}(N_\ell) = \sum_{j=0}^{q} \binom{n-1}{j}.$$

Proof If $\ell \in \Gamma_p$ and $j \leq p-1$ then $\alpha_j(\ell) = \binom{n-1}{j}$ since any subset of cardinality $j+1 \leq p$ is short with respect to ℓ. Similarly, for $j \leq p-2$ one has $\alpha_{n-2-j}(\ell) = 0$. Indeed, $\alpha_{n-2-j}(\ell)$ coincides with the number of long subsets $J \subset \{1, \ldots, n-1\}$ of cardinality $j+2$ and (as follows from our assumption $j+2 \leq p$) all such subsets are short. Formula (20) now follows from (17). $\qquad\square$

Our next goal is to show that for a fixed p the Lebesgue measure of the set Γ_p is exponentially close to the measure of the whole simplex Δ^{n-1} as $n \to \infty$. More precisely, we will prove the following statement:

Proposition 4.3 *For* $n > 4$ *one has*

(21) $$1 - n^{2p} \cdot 2^{-n} \leq \frac{\mathrm{vol}(\Gamma_p)}{\mathrm{vol}(\Delta^{n-1})} \leq 1.$$

Proof   Let $J \subset \{1, \ldots, n\}$ be a subset with $|J| = p$. Let $L \subset \mathbb{R}^n$ be the affine subspace given by the equation $l_1 + \cdots + l_n = 1$. Denote by $\phi_J \colon L \to \mathbb{R}$ the affine functional

$$\phi_J(l_1, \ldots, l_n) = \sum_{i \in J} l_i - \sum_{i \notin J} l_i. \tag{22}$$

Denote by $V_J \subset \Delta^{n-1}$ the set of vectors $\ell \in \Delta^{n-1}$ satisfying the inequality $\phi_J(\ell) \geq 0$. It is clear that

$$\Delta^{n-1} - \Gamma_p = \bigcup_{|J|=p} V_J$$

and therefore

$$\frac{\mathrm{vol}(\Delta^{n-1} - \Gamma_p)}{\mathrm{vol}(\Delta^{n-1})} \leq \sum_{|J|=p} \frac{\mathrm{vol}(V_J)}{\mathrm{vol}(\Delta^{n-1})}. \tag{23}$$

Inequality (21) now follows by combining (23) with the inequality of Proposition 4.4 (see below) which allows estimating each term of (23). The number of terms in the sum (23) equals $\binom{n}{p}$ which is less or equal than n^p.   □

Proposition 4.4   *Assume that $n > 4$. Then for any subset $J \subset \{1, \ldots, n\}$ with $|J| = p$ one has*

$$\frac{\mathrm{vol}(V_J)}{\mathrm{vol}(\Delta^{n-1})} < n^p \cdot 2^{-n}. \tag{24}$$

Proof   Consider the complement $\bar{J}$ of J and the functional $\phi_{\bar{J}} = -\phi_J \colon L \to \mathbb{R}$ (22). Clearly V_J can be represented as the intersection $H \cap \Delta^{n-1}$ where H is the half-space $H = \{\ell \in L; \phi_{\bar{J}}(\ell) \leq 0\}$. Comparing with notation of Lemma 3.1 and Corollary 3.2 we may write

$$\frac{\mathrm{vol}(V_J)}{\mathrm{vol}(\Delta^{n-1})} = r_{p,n-p} = 2^{-n} \cdot \sum_{k=0}^{p-1} \binom{n-p-1+k}{k} \cdot 2^{p-k}. \tag{25}$$

The binomial coefficients can be estimated by $\binom{n-p-1+k}{k} \leq n^k$ and therefore

$$\frac{\mathrm{vol}(V_J)}{\mathrm{vol}(\Delta^{n-1})} \leq 2^{p-n} \cdot \sum_{k=0}^{p-1} \left(\frac{n}{2}\right)^k \leq 2^{p-n} \cdot \frac{\left(\frac{n}{2}\right)^p}{\frac{n}{2}-1} < n^p \cdot 2^{-n}.$$

On the last step we used the assumption $n > 4$.   □

Proof of Theorem 2.2   Let ν_n be an admissible sequence of measures on Δ^{n-1}, where $n = 1, 2, \ldots$; see Definition 2.1. Using Proposition 4.2 we obtain

$$\left| \int_{\Delta^{n-1}} b_{2p}(N_\ell) d\nu_n - \sum_{i=0}^{p} \binom{n-1}{i} \right|$$

$$= \left| \int_{\Delta^{n-1} - \Gamma_{p+2}} b_{2p}(N_\ell) d\nu_n - \sum_{i=0}^{p} \binom{n-1}{i} \cdot \nu_n(\Delta^{n-1} - \Gamma_{p+2}) \right|$$

$$\leq \left[n^{p+1} + \sum_{i=0}^{p} \binom{n-1}{i} \right] \cdot \nu_n(\Delta^{n-1} - \Gamma_{p+2}).$$

The last inequality uses Proposition 4.1. If $\ell = (l_1, \ldots, l_n) \in \Delta^{n-1} - \Gamma_{p+2}$ then there exists a subset $J \subset \{1, \ldots, n\}$ with $|J| = p + 2$ such that $\sum_{i \in J} l_i \geq 1/2$ and hence $l_i \geq 1/(2p+4)$ for some $i \in J$. This shows that $\Delta^{n-1} - \Gamma_{p+2} \subset \Lambda_{p+2}$ (which was defined before Definition 2.1 in Section 2. Therefore we may use property (iii) from Definition 2.1 to continue the previous string of inequalities

$$\left[n^{p+1} + \sum_{i=0}^{p} \binom{n-1}{i} \right] \cdot \nu_n(\Delta^{n-1} - \Gamma_{p+2}) \leq 2n^{p+1} \cdot \mu_n(\Delta^{n-1} - \Gamma_{p+2}) \cdot M_n$$

where
$$M_n = \max_{\ell \in \Lambda_{p+2}} f_n(\ell) \leq Ab^n.$$

The constants $A > 0$ and $0 < b < 2$ appearing here are given by Definition 2.1; see (iii). Using Proposition 4.3 we have $\mu_n(\Delta^{n-1} - \Gamma_{p+2}) \leq n^{2p+4} \cdot 2^{-n}$. Hence we finally obtain

$$\left| \int_{\Delta^{n-1}} b_{2p}(N_\ell) d\nu_n - \sum_{i=0}^{p} \binom{n-1}{i} \right| \leq 2n^{p+1} \cdot Ab^n \cdot n^{2(p+2)} \cdot 2^{-n}$$

$$\leq 2A \cdot n^{3p+5} \cdot \left(\frac{b}{2} \right)^n \leq Ca^n.$$

Here a is any number satisfying $b/2 < a < 1$ and $C > 0$ is chosen accordingly.   □

5   Proof of Theorem 2.3

First we recall the result of [6] describing Betti numbers $b_p(M_\ell)$ of planar polygon spaces as functions of the length vector ℓ. Fix an index $1 \leq i \leq n$ such that l_i is maximal

among $l_1, \ldots, l_n$. Denote by $a_p(\ell)$ the number of short subsets $J \subset \{1, \ldots, n\}$ of cardinality $|J| = 1 + p$ containing i. A subset $J \subset \{1, \ldots, n\}$ is called *median* if

$$\sum_{i \in J} l_i = \sum_{i \notin J} l_i.$$

Denote by $\tilde{a}_p(\ell)$ the number of median subsets $J \subset \{1, \ldots, n\}$ containing i and such that $|J| = 1 + p$. Then one has

$$(26) \qquad b_p(M_\ell) = a_p(\ell) + \tilde{a}_p(\ell) + a_{n-3-p}(\ell).$$

for $p = 0, 1, \ldots, n - 3$ [6].

Proposition 5.1 *For a generic length vector $\ell = (l_1, \ldots, l_n) \in \Delta^{n-1}$ one has*

$$(27) \qquad b_p(M_\ell) \leq n^{p+2}.$$

Proof From the description of $b_p(M_\ell)$ given above one deduces that for a generic ℓ the number $\tilde{a}_p(\ell)$ vanishes and $a_p(\ell)$ (see (26)) is less or equal than the binomial coefficient $\binom{n-1}{p}$. Hence using (26) we obtain

$$b_p(M_\ell) \leq \binom{n-1}{p} + \binom{n-1}{n-3-p} \leq (n-1)^p + (n-1)^{p+2} \leq n^{p+2}. \qquad \square$$

Proposition 5.2 *If $\ell \in \Gamma_p \subset \Delta^{n-1}$ is generic then*

$$(28) \qquad b_j(M_\ell) = \binom{n-1}{j} \qquad \text{for all} \quad j \leq p - 2.$$

Proof As follows from (26), if ℓ is generic and lies in Γ_p then $\tilde{a}_j(\ell) = 0$ and for $j \leq p$ the number $a_j(\ell)$ equals the number of all subsets of $\{1, \ldots, n\}$ which do not contain the element with the maximal length l_i, ie $a_j(\ell) = \binom{n-1}{j}$. The number $a_{n-3-j}(\ell)$ is the number of all short subsets of cardinality $n - 2 - j$ containing the maximal element; the complements of these sets are short and their cardinality is $j + 2$. However if $j + 2 \leq p$ all such subsets must be short, ie $a_{n-3-j}(\ell) = 0$ for $j \leq p - 2$. $\qquad \square$

Proof of Theorem 2.3 The proof essentially repeats the arguments of the proof of Theorem 2.2 with Proposition 5.1 replacing Proposition 4.1 and with Proposition 5.2 replacing Proposition 4.2. $\qquad \square$

6 Normal length vectors

In the paper [4] we introduced the notion of a normal length vector. A vector $\ell = (l_1, \ldots, l_n) \in \mathbb{R}_+^n$ is called *normal* if the intersection of all subsets $J \subset \{1, \ldots, n\}$ of cardinality 3 which are long with respect to ℓ is not empty. A length vector ℓ with the property that all subsets of cardinality 3 are short with respect to ℓ is normal since then the intersection of all long subsets of cardinality 3 equals $\{1, \ldots, n\}$ as the intersection of the empty family.

The importance of normal length vectors stems from the following result proven in [4]:

Theorem 6.1 [4] *Suppose that $\ell, \ell' \in \Delta^{n-1}$ are two ordered length vectors such that there exists a graded algebra isomorphism between the integral cohomology algebras $H^*(M_\ell) \to H^*(M_{\ell'})$. Assume that one of the vectors ℓ, ℓ' is normal. Then the other vector is normal as well and ℓ and ℓ' lie in the same stratum of the simplex Δ^{n-1}; in particular, the polygon spaces M_ℓ and $M_{\ell'}$ are diffeomorphic.*

We apply the technique developed in this paper to show that the Lebesgue measure of the set of length vectors which are not normal is exponentially small for large $n \to \infty$.

Proposition 6.2 *Let $\mathcal{N}_n \subset \Delta^{n-1}$ denote the set of all normal length vectors. The relative volume of $\mathcal{N}_n$ satisfies the inequality*

$$(29) \qquad \frac{\mathrm{vol}(\Delta^{n-1} - \mathcal{N}_n)}{\mathrm{vol}(\Delta^{n-1})} < \frac{n^6}{2^n}.$$

Proof We observe that $\Gamma_3 \subset \mathcal{N}_n$ and applying (21) we find

$$\frac{\mathrm{vol}(\Delta^{n-1} - \mathcal{N}_n)}{\mathrm{vol}(\Delta^{n-1})} < \frac{\mathrm{vol}(\Delta^{n-1} - \Gamma_3)}{\mathrm{vol}(\Delta^{n-1})} < \frac{n^6}{2^n}. \qquad \square$$

7 Mathematical expectations of higher moments

Theorem 7.1 *Given an admissible sequence of probability measures ν_n and integers $p \geq 0$, and k, then there exist constants $C > 0$ and $0 < a < 1$ such that the k–th powers of the average Betti numbers (6) satisfy the following equations for all $n = 3, 4, \ldots$:*

$$(30) \qquad \left| \int_{\Delta^{n-1}} b_{2p}(N_\ell)^k d\nu_n - \left(\sum_{i=0}^{p} \binom{n-1}{i} \right)^k \right| < C \cdot a^n$$

$$(31) \qquad \left| \int_{\Delta^{n-1}} b_p(M_\ell)^k d\nu_n - \binom{n-1}{i}^k \right| < C \cdot a^n$$

Proof One simply repeats the arguments used in proofs of Theorem 2.2 and Theorem 2.3 with minor modifications. The scheme remains the same: on the central domain $\Gamma_{p+2} \subset \Delta^{n-1}$ the functions appearing in (30) and (31) coincide. The volume of the remaining part $\Delta^{n-1} - \Gamma_{p+2}$ is exponentially small (by Proposition 4.3) and the functions involved have polynomial upper bounds in n. $\qquad\square$

8 Some open questions

In this section I would like to mention a few interesting open questions. The tools of the present paper seem to be inadequate to give their solutions.

It would be interesting to find the average total Betti number

$$B(n, \mu) = \int_{\mathbb{R}^n_+} B(M_\ell)d\mu, \quad \text{where} \quad B(M_\ell) = \sum_{p=0}^{n-3} b_p(M_\ell)$$

for various natural probability measures μ on the spaces of parameters $\mathbb{R}^n_+$ (or on Δ^{n-1}) and to examine the behavior of $B(n, \mu)$ for $n \to \infty$. Although we know the behavior of the individual average Betti numbers $\int_{\mathbb{R}^n_+} b_p(M_\ell)d\mu$ for large n and for fixed p one cannot simply add the terms up.

By Theorem 2 from [6] one has an upper bound

$$(32) \qquad B(n, \mu) \leq 2^{n-1} - \binom{n-1}{r} \sim 2^{n-1} \cdot \left(1 - \sqrt{\frac{2}{n\pi}}\right),$$

where $r = [(n-1)/2]$ denotes the integer part of $(n-1)/2$. It is plausible that the RHS of (32) gives the right asymptotic for $B(n, \mu)$.

One can raise a similar question concerning the average total Betti numbers of spatial polygon spaces N_ℓ.

A homotopy invariant $\mathbf{TC}(X)$, introduced in [3], measures the complexity of the problem of navigation in a topological space X, viewed as the configuration space of a mechanical system. It is a challenging problem to compute $\mathbf{TC}(M_\ell)$ as a function of the length vector $\ell \in \mathbb{R}^n_+$ and then study its average

$$(33) \qquad \mathbf{TC}(n, \mu) = \int_{\mathbb{R}^n_+} \mathbf{TC}(M_\ell)d\mu$$

and behavior as $n \to \infty$ under different assumptions on the measure μ. Formally the invariant $\mathbf{TC}(X)$ is defined only when X is path-connected. This assumption may be violated in the case of spaces M_ℓ. If X is not path-connected it is natural to define

$\mathbf{TC}(X)$ as max $\mathbf{TC}(X_i)$ where X_i are path-connected components of X. If M_ℓ is disconnected then it is disjoint union of two tori T^{n-3} (see Kapovich and Millson [9]) and hence in this case $\mathbf{TC}(M_\ell) = \mathbf{TC}(T^{n-3}) = n - 2$ [3].

The similar question concerning spatial polygon spaces N_ℓ is much easier. One can show that $\mathbf{TC}(N_\ell) = 2n - 5$ assuming that $N_\ell \neq \varnothing$ and hence

$$(34) \qquad \int_{\mathbb{R}^n_+} \mathbf{TC}(N_\ell) d\nu_n \sim 2n - 5$$

for any admissible sequence of probability measures ν_n. The error in (34) is exponentially small for large n.

The results (33) and (34) concerning average topological complexity of polygon spaces M_ℓ and N_ℓ may have important applications in molecular biology, statistical shape theory and robotics.

References

[1] **M M Ali**, *Content of the frustum of a simplex*, Pacific J. Math. 48 (1973) 313–322 MR0394430

[2] **H B Curry**, **I J Schoenberg**, *On Pólya frequency functions. IV. The fundamental spline functions and their limits*, J. Analyse Math. 17 (1966) 71–107   MR0218800

[3] **M Farber**, *Topological complexity of motion planning*, Discrete Comput. Geom. 29 (2003) 211–221   MR1957228

[4] **M Farber**, **J-C Hausmann**, **D Schütz**, *On the conjecture of Kevin Walker*, preprint (2007)

[5] **M Farber**, **T Kappeler**, *Betti numbers of random manifolds*, to appear in Homology, Homotopy and Applications

[6] **M Farber**, **D Schütz**, *Homology of planar polygon spaces*, Geom. Dedicata 125 (2007) 75–92   MR2322541

[7] **J-C Hausmann**, **A Knutson**, *The cohomology ring of polygon spaces*, Ann. Inst. Fourier (Grenoble) 48 (1998) 281–321   MR1614965

[8] **J-C Hausmann**, **E Rodriguez**, *The space of clouds in Euclidean space*, Experiment. Math. 13 (2004) 31–47   MR2065566

[9] **M Kapovich**, **J Millson**, *On the moduli space of polygons in the Euclidean plane*, J. Differential Geom. 42 (1995) 133–164   MR1350697

[10] **D G Kendall**, **D Barden**, **T K Carne**, **H Le**, *Shape and shape theory*, Wiley Series in Probability and Statistics, John Wiley & Sons Ltd., Chichester (1999)   MR1891212

[11] **A A Klyachko**, *Spatial polygons and stable configurations of points in the projective line*, from: "Algebraic geometry and its applications (Yaroslavl', 1992)", Aspects Math., E25, Vieweg, Braunschweig (1994) 67–84 MR1282021

[12] **G Varsi**, *The multidimensional content of the frustum of the simplex*, Pacific J. Math. 46 (1973) 303–314 MR0322680

[13] **K Walker**, *Configuration spaces of linkages*, Undergraduate thesis, Princeton University (1985)

Department of Mathematical Sciences, University of Durham, UK

Michael.Farber@durham.ac.uk

http://maths.dur.ac.uk/~dma0mf/