The Growth of the Mean Average Crossing Number of Equilateral polygons in Confinement

J. Arsuaga*§, B. Borgo*, Y. Diao†‡ and R. Scharein*

*Department of Mathematics
San Francisco State University
1600 Holloway Ave
San Francisco, CA 94132
†Department of Mathematics and Statistics
University of North Carolina at Charlotte
Charlotte, NC 28223

Abstract. The physical and biological properties of collapsed long polymer chains as well as of highly condensed biopolymers (such as DNA in all organisms) are known to be determined, at least in part, by their topological and geometrical properties. With this purpose of characterizing the topological properties of such condensed systems equilateral random polygons restricted to confined volumes are often used. However very few analytical results are known. In this paper, we investigate the effect of volume confinement on the mean average crossing number (ACN) of equilateral random polygons. The mean ACN of knots and links under confinement provides a simple alternative measurement for the topological complexity of knots and links in the statistical sense. For an equilateral random polygon of n segments without any volume confinement constrain, it is known that its mean ACN $\langle \text{ACN} \rangle$ is of the order $\frac{3}{16} n \ln n + O(n)$. Here we model the confining volume as a simple sphere of radius R. We provide an analytical argument that shows that $\langle \text{ACN} \rangle$ of an equilateral random polygon of n segments under extreme confinement (meaning $R << n$) grows as $O(n^2)$. We propose to model the growth of $\langle \text{ACN} \rangle$ as $a(R)n^2 + b(R)n\ln(n)$ under a less extreme confinement condition, where $a(R)$ and $b(R)$ are functions of R with R being the radius of the confining sphere. Computer simulations performed show a fairly good fit using this model.

AMS classification scheme numbers: Primary 57M25, secondary 92B99.

1. Introduction

DNA and other biopolymers fold in space in ways that are usually difficult to quantify. This problem is severely magnified when the polymer is confined to small volumes. For instance DNA is highly condensed in all organisms and the spatial trajectory of the duplex remains a matter of study [12]. A simple measure to study the folding of polymers in space is by counting the number of crossings that can be perceived while observing a non-perturbed trajectory of a given polymer in a given orthogonal projection [15]. To avoid the dependence of this measure on a single

§ To whom correspondence regarding the numerical aspect of the paper should be addressed (jarsuaga@math.sfsu.edu)
‡ To whom correspondence regarding the analytical aspect of the paper should be addressed (ydiao@uncc.edu)

selected orthogonal projection, the average crossing number (ACN), defined as the average of crossing numbers over all orthogonal projections, is used. For practical purposes one considers the mean ACN which is defined as the average of ACN over the entire statistical ensemble of polymers with a fixed length, we denote this value by $\langle \text{ACN} \rangle$.

In the case of knotted DNA molecules the $\langle \text{ACN} \rangle$ can be experimentally measured since it correlates well with the experimentally observed speed of electrophoretic migration of knotted DNA molecules of the same size but of various knot types [28], with the expected sedimentation coefficient of different types of DNA knots [33], and with relaxation dynamics of modeled knotted polymers [13]. This observation is particularly relevant for the understanding of DNA knots extracted from bacteriophage P4 [19], which have been proposed to study DNA folding inside bacteriopages [3].

Some circular biopolymers and in particular DNA are often modeled in free solution as equilateral random polygons (also called freely jointed polygons). Some topological and geometrical properties are known about equilateral random polygons (e.g. [7, 11, 32]) and in particular it is known that an equilateral random polygon of n segments without any volume confinement constrain the $\langle \text{ACN} \rangle$ scales as $\frac{3}{16} n \ln n + O(n)$ [8]. One may model densely packed circular polymers confined to a tight space using equilateral random polygons confined into a given volume and aim at developing a theory that explains the scaling behavior of different topological and geometrical parameters. This however has proven to be a very difficult task. In most cases these results have been limited to numerical studies [3, 4, 20, 21, 22, 23]. Theoretical results on the other hand have been limited to simpler polymer models [2].

In this paper, we investigate the case of an equilateral random polygon confined to a sphere. This simple model can easily be generalized to other convex volumes and has biological relevance since it can be directly applied to the folding of DNA in bacteriophage capsids. First we revise some of the basic results for equilateral random polygons. Second we provide an analytic proof that shows that for tightly confined random equilateral polygons of length n, $\langle \text{ACN} \rangle$ scales as $O(n^2)$. This prompts us to model $\langle \text{ACN} \rangle$ as $a(R)n^2 + b(R)n\ln(n)$, where R is the radius of the confining sphere. The function $a(R)$ is positive when R is small and it decreases to 0 as R increases to $n/2$. The function $b(R) \approx 0$ for small values of R and increases to $3/16$ as R increases to $n/2$ (when the effect of confinement effect totally vanishes). Third we perform numerical simulations to determine the values of these scaling coefficients. We conclude by discussing the relevance of these results to the problem of DNA packing in bacteriophage P4.

2. Basic facts about equilateral random polygons

Let us formally define the equilateral random polygons first. Let Y_1, Y_2, ... , Y_n be n independent random vectors uniformly distributed on S^2 (so the joint probability density function of the three coordinates of each Y_j is simply $\frac{1}{4\pi}$ on the unit sphere and 0 otherwise). An equilateral random walk of n steps, denoted by EW_n, is defined as the sequence of points in the three dimensional space $\mathbf{R}^3$: $X_0 = O$, $X_k = Y_1 + Y_2 + \cdots + Y_k$, $k = 1, 2, ..., n$. Each X_k is called a vertex of the EW_n

and the line segment joining X_k and X_{k+1} is called an edge of EW_n (which is of unit length). If the last vertex X_n of EW_n is fixed, then we have a conditioned random walk $EW_n|X_n$. In particular, EW_n becomes a polygon if $X_n = O$. In this case, it is called an equilateral random polygon and is denoted by EP_n. The joint probability density function $f(X_1, X_2, ..., X_n)$ of the vertices of an EW_n is $f(X_1, X_2, ..., X_n) = \varphi(U_1)\varphi(U_2)\cdots\varphi(U_n) = \varphi(X_1)\varphi(X_2 - X_1)\cdots\varphi(X_n - X_{n-1})$.

Let X_k be the k-th vertex of an EW_n ($n \geq k > 1$), its density function is defined by

$$f_k(X_k) = \int\int\cdots\int \varphi(X_1)\varphi(X_2 - X_1)\cdots \varphi(X_k - X_{k-1})dX_1dX_2\cdots dX_{k-1} \quad (1)$$

and it has the closed form $f_k(X_k) = \frac{1}{2\pi^2 r}\int_0^\infty x \sin rx \left(\frac{\sin x}{x}\right)^k dx$ [25]. In the case of EP_n, the density function $h_k(X_k)$ of the vertex X_k can be approximated by the following Gaussian distribution (derived from the above formula) [7, 9, 10]:

$$h_k(X_k) \approx \left(\sqrt{\frac{3}{2\pi\sigma_{nk}^2}}\right)^3 \exp\left(-\frac{3|X_k|^2}{2\sigma_{nk}^2}\right),$$

where $\sigma_{nk}^2 = \frac{k(n-k)}{n}$ and the error of the estimation is at most of the order of $O\left(\frac{1}{k^{5/2}} + \frac{1}{(n-k)^{5/2}}\right)$. From this one can then derive some important results concerning equilateral random polygons. An application of this formula that is particularly of interest to us is the derivation of the mean ACN of EP_n, as stated in the following theorem.

Theorem 1 [8] *For an equilateral random walk of n steps,*

$$\langle \text{ACN} \rangle = \frac{3}{16} n \ln n + O(n).$$

Another relevant theoretical result regarding the topological aspects of EP_n is the following theorem (this had been observed in simulations by many independent research groups (for example [31]).

Theorem 2 [7] *Let $\mathcal{K}$ be any knot type, then there exists a positive constant ϵ such that EP_n contains $\mathcal{K}$ as a connected sum component with a probability at least $1-\exp(-n^\epsilon)$, provided that n is large enough.*

In the next section, we will establish the asymptotic growth rate of $\langle \text{ACN} \rangle$ for an equilateral random polygon confined in a tight volume.

3. Analytical Results

Let us first give a formal definition of equilateral random walks and polygons confined in a space. For simplicity, we will assume that the confining space is $B(O, r)$, namely the ball of radius $r > 1$ that is centered at the origin O. Furthermore, we will assume that the length of each step of the walks and polygons is one.

Definition 1 *Let $X_0 \in B(O,r)$ and let $S(X_0)$ be the unit sphere centered at X_0. $S'(X_0) = S(X_0) \cap B(O,r)$ is either $S(X_0)$ itself or is a spherical cap. If X_1 is a random point uniformly distributed on $S'(X_0)$, then we will say that $\overrightarrow{X_0X_1}$ is a confined random step (from X_0). An equilateral random walk of n steps confined in $B(O,r)$ (starting from a point $X_0 \in B(O,r)$) is then defined as a sequence of points X_0, X_1, X_2, ..., X_n such that $\overrightarrow{X_jX_{j+1}}$ is a confined random step (from X_j) for each $n-1 \geq j \geq 0$. An equilateral random polygon of n steps confined in $B(O,r)$ is simply a conditioned equilateral random walk of n steps confined in $B(O,r)$ under the condition that $X_n = X_0$.*

Of course, every equilateral random walk or polygon defined above is confined in $B(O,r)$. Here in the definition we are allowing the random walks and polygons to intersect the boundary of $B(O,r)$, but the probability of that is actually zero as one can check. From now on, we will use $W_n^c(X_0)$ to denote an equilateral random walk of n steps confined in $B(O,r)$ with starting point $X_0 \in B(O,r)$ and will use $P_n^c(X_0)$ to denote an equilateral random polygon of n steps confined in $B(O,r)$ with starting point $X_0 \in B(O,r)$. Furthermore, we will use $W_n^c(X',X'')$ to denote a conditioned equilateral random walk of n steps confined in $B(O,r)$ where the condition is that the starting and ending point of the random walk are X' and X'' respectively. Any specific configuration of a $W_n^c(X',X'')$ is called a *pattern*. A $W_n^c(X',X'')$ is said to be within an ϵ-neighborhood of a pattern if each vertex of $W_n^c(X',X'')$ is within an ϵ-distance from its corresponding vertex in the pattern. Figure 1 shows two special patterns of equilateral random walks confined in a sphere where they intersect each other at the center of the sphere.

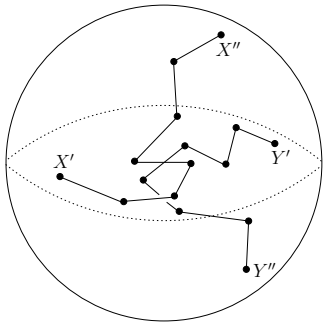


Figure 1: Two special patterns with their middle steps intersecting each other perpendicularly.

Lemma 1 *Let $m = 2\lceil r \rceil + 3$, then there exist positive constants $\alpha > 0$ and $\beta > 0$ (depending only on r) such that for any X', X'', Y', Y'' in $B(O,r)$, the probability that the ACN between the middle step of $W_m^c(X',X'')$ and the middle step of $W_m^c(Y',Y'')$ is greater than β is at least α.*

Proof. The choice of m ensures that no matter what X', X'', Y' and Y'' are, one could always draw two special patterns, one for $W_n^c(X',X'')$ and one for $W_n^c(Y',Y'')$ such that their middle steps form a perpendicular cross centered at the origin as shown in Figure 1. This is possible since the distance from any end point in this cross is of a distance at most $r + 1/2 < \lceil r \rceil + 1$ to any point in $B(O,r)$. If we perturb the vertices

of $W_n^c(X', X'')$ and $W_n^c(Y', Y'')$ within the 1/4-neighborhood U of these two patterns, then the ACN of the two middle steps would be greater than some positive constant β. Since the probability that the vertices of $W_m^c(X', X'')$ and $W_m^c(Y', Y'')$ fall into U (for any such special patterns) is bounded below by a positive constant (only depending on r), it follows that the probability that the two middle steps has an ACN $\geq \beta$ is bounded below by some positive constant $\alpha > 0$.

Theorem 3 *Let P_n^c be an equilateral random polygon of n edges confined in $B(O, r)$, then the mean ACN of P_n^c is of the order of $O(n^2)$.*

Proof. Let ℓ_1, ℓ_2, ..., ℓ_n be the consecutive edges of P_n^r and let a_{ij} be the average crossing number between ℓ_i and ℓ_j, then the ACN of P_n (written as χ_n) is simply the sum of the a_{ij}'s: $\sum_{1\leq i<j\leq n} a_{ij} = \frac{1}{2}\sum_{1\leq i,j\leq n} a_{ij}$, and $\langle \text{ACN} \rangle = \sum_{1\leq i<j\leq n} E(a_{ij}) = \frac{1}{2}\sum_{1\leq i,j\leq n} E(a_{ij})$. Since $a_{ij} \leq 1$ for any i, j, it follows that $\langle \text{ACN} \rangle$ is at most of the order $O(n^2)$. On the other hand, since we are only concerned with the asymptotic behavior of the ACN, we may assume that $n \gg r$ and we will only consider the a_{ij} terms where $i < j$, $j - i \geq 2\lceil r \rceil + 3$ and $n - j + i \geq 2\lceil r \rceil + 3$. Now take such a pair and let T_1 be the i-th edge and T_2 be the j-th edge. As defined in Lemma 1, let m be the unique odd positive integer such that $m = 2\lceil r \rceil + 3$. Let X' be the starting point of the $(m-1)/2$-edge (step) prior to T_1 (following the natural orientation of the edges inherited from the definition of P_n^c), X'' be the ending point of the $(m-1)/2$-edge after T_1. Similarly, let Y' be the starting point of the $(m-1)/2$-edge prior to T_2 and Y'' be the ending point of the $(m-1)/2$-edge after T_2. By the choice of i, j, X', X'', Y' and Y'' match the cyclic order they inherit from the polygon. Let $f(X', X'', Y', Y'')$ be the joint probability density function of X', X'', Y' and Y''. Let $P(X', X'', Y', Y'')$ be the (conditional) probability that the ACN between the middle step of $W_n^c(X', X'')$ (which is just T_1) and the middle step of $W_n^c(Y', Y'')$ (which is T_2) is greater than β. By Lemma 1, $P(X', X'', Y', Y'') \geq \alpha$. Thus we have

$$\begin{aligned} & P(a_{ij} \geq \beta) \\ = & \int P(X', X'', Y', Y'') f(X', X'', Y', Y'') dX' dX'' dY' dY'' \\ \geq & \alpha \int f(X', X'', Y', Y'') dX' dX'' dY' dY'' \\ = & \alpha. \end{aligned}$$

It follows that

$$E(a_{ij}) \geq P(a_{ij} \geq \beta)\beta \geq \alpha\beta.$$

That is, $E(a_{ij})$ is bounded below by a positive constant. The result of the theorem now follows from $E(\chi_n) = \sum_{1\leq i<j\leq n} E(a_{ij})$ and the fact that there are $O(n^2)$ pairs (i, j) satisfying the conditions $i < j$, $j - i \geq 2\lceil r \rceil + 3$ and $n - j + i \geq 2\lceil r \rceil + 3$.

4. Numerical Results

4.1. Methods

To generate ensembles of random equilateral polygons, we use the generalized hedgehog algorithm [32]. Confinement within spherical volumes was achieved by rejecting those polygons that had at least one vertex outside the specified radius.

Figure 2 shows three examples of random polygons inside spheres of radii $R = 2,\ 3$ and 4 (where the radius of confinement is measured in unit polygon segment length). Because the computational complexity of the generalized hedgehog algorithm grows linearly with the length of the polygon we managed to generate polygons of up to 400 segments confined to spheres of radius 4. Sample sizes ranged from 10^3 to 10^6 which is comparable to sample sizes from previous studies [8, 20].

Figure 2: Confined random polygons generated using the hedgehog method. Each polygon consists of 150 segments of equal length. From left to right: the radius of confining sphere is 2, 3 and 4, the radius of gyration is 2.23, 3.003, and 4.042 and the ACN is 234.73, 141.24, and 82.67 respectively. (The confining spheres have been re-scaled to unit radius so that the details of the complexity change can be compared easily.) The effect of confinement on the polygons complexity is apparent. The presented polygons have been "smoothed" using KnotPlot [30].

4.2. Individual ACN computation

Sample conformations were used to calculate both the $\langle \mathrm{ACN} \rangle$ and the radius of gyration. Each individual ACN was computed using the solid angle formulation (also known as the Gaussian integral formula) as explained in [17]. Briefly speaking, in the case that C is a polygon of n segments of unit length such that each segment is parameterized as $\gamma_j(s)$ where s is the arclength of the segment, then the ACN can be expressed as

$$\sum_{1 \le i \ne j \le n} \frac{1}{4\pi} \int_0^1 \int_0^1 \frac{|(\dot{\gamma}_i(t), \dot{\gamma}_j(s), \gamma_i(t) - \gamma_j(s))|}{|\gamma_i(t) - \gamma_j(s)|^3} dt ds. \tag{2}$$

4.3. Results

In order to model the behavior of the $\langle \mathrm{ACN} \rangle$ in confinement and in free space we fit the data with the function $\mathrm{ACN}(R, n) = a(R)n^2 + b(R)n \ln(n)$, where $\mathrm{ACN}(R, n)$ is the $\langle \mathrm{ACN} \rangle$ as a function of the number of segments, and the radius of the confining sphere. The first term of this formula accounts for the $O(n^2)$ growth rate of $\langle \mathrm{ACN} \rangle$ in the case that R is small (Theorem 4), and the second term accounts for the $O(n \ln n)$ growth rate of $\langle \mathrm{ACN} \rangle$ in the case that there is no volume confinement constrain (Theorem 2). Figure 4 shows the asymptotic behavior for random equilateral polygons for $R = 2,\ 3,\ 4$ and no confinement. By visual inspection one can easily appreciate the different behavior of the $\langle \mathrm{ACN} \rangle$ for confined polygons versus unconfined ones. Curves fit to our simulated data by non-linear least squares are in excellent agreement, with χ^2 values of 4.85 ($R = 2,\ a = .01119,\ b = .05358$), 3.84

($R = 3,\ a = .004871,\ b = .076516$), and 3.36 ($R = 4,\ a = .002495,\ b = .099431$). For large radii of confinement (> 4), significantly longer segments would be required to obtain similar results and are not considered necessary in this study.

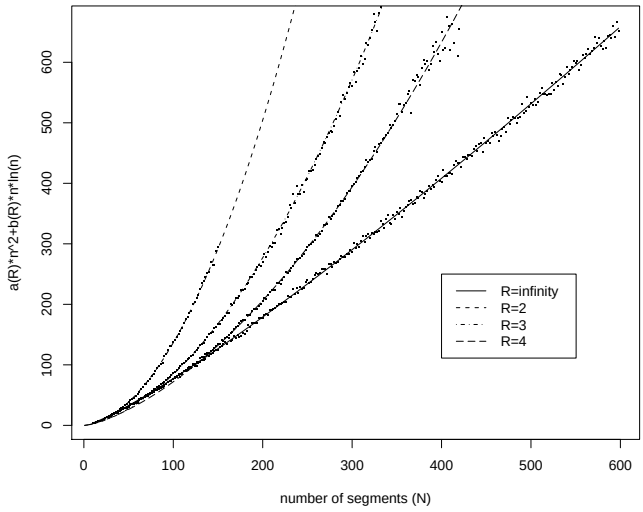
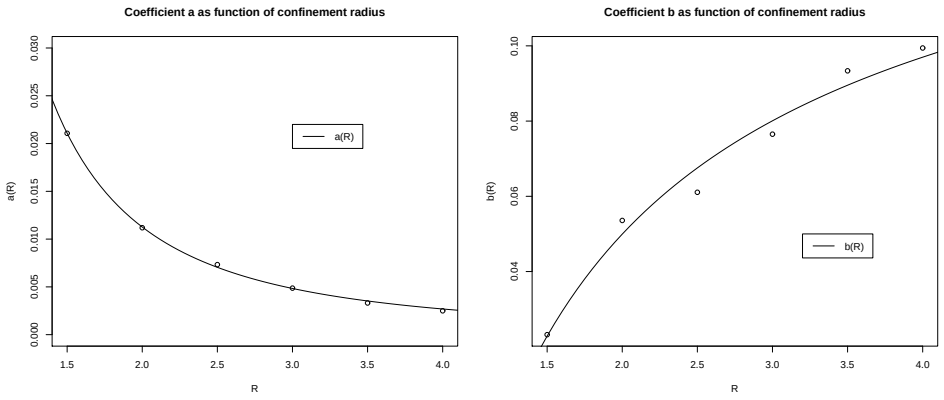


Figure 3: Relationship between mean ACN and the confining radius. The mean ACN is plotted against the number of segments and is fit as $a(R)n^2+b(R)n\ln(n)$. Deviation from free space scaling is apparent.

4.4. The Fitting of the Scaling Coefficients

Intuitively, one expects that $a(R)$ would be a decreasing function of R that starts as a positive number when R is a small and decreases to near 0 as R increases to pass the average radius of gyration of the polygons (that is when the confining effect disappears). Similarly, $b(R)$ should be an increasing function that starts from near 0 when R is small and increases to the constant $3/16$ as n passes the average radius of gyration. Figure 4 shows $a(R)$ and $b(R)$ as a function of R. In the plot, we have empirically fit $a(R) = \frac{a_1}{R^2+a_2}$, where $a_1 = .04235$ and $a_2 = -.24017$. On the other hand, we have fit $b(R)$ with the formula $b(R) = \frac{b_1+b_2\sqrt{R}}{R} + .1875$ where $b_1 = -.06521$ and $b_2 = -.14842$. This provided a fairly good fit for $b(R)$ as shown in Figure 4(b). Notice that while this fitting of $a(R)$ does not totally vanish when R goes to the order of n, it becomes of the order $O(1/n^2)$ hence the $a(R)n^2$ term becomes negligible. On the other hand, the fitting function used for $b(R)$ becomes $3/16 + O(1/\sqrt{n})$ when $R = O(n)$, so the corresponding $b(R)n\ln n$ term will capture the main term $3/16n\ln n$ in $\langle \mathrm{ACN} \rangle$.



(a) (b)

Figure 4: Coefficients $a(R)$ and $b(R)$ of $\mathrm{ACN}(n, R)$. In each plot, the horizontal axis is the radius of confinement, and the vertical axis is the value of each parameter. The data was fit using nonlinear least squares.

5. Conclusion and future work

Developing the theory that describes the topological and geometrical properties of polymers in confinement is important and remains an important challenge. Here we have shown that the growth of the $\langle \mathrm{ACN} \rangle$ for equilateral random polygons confined to spherical volumes follows a $O(n^2)$ formula. This result is important because it shows that the complexity of self-entanglement of a polymer (modeled as a freely jointed chain) in confinement is different from that in free space which is known to be $\frac{3}{16} n \ln n$ for $\langle \mathrm{ACN} \rangle$ [8]. Furthermore it provides a reference for models of polymer folding and in particular for DNA folding in confinement. Interestingly other models give similar results. For instance uniform random polygons (i.e. those generated by placing random points uniformly inside a cube) and spooling random polygons also show an $O(n^2)$ behavior [1, 2]. In the future we aim at applying these results to better characterize DNA knots extracted from bacteriophage P4 and use these results to provide new insights in the organization of DNA inside bacteriophages, which still remains a matter of debate [1, 6, 14, 24]. Furthermore we aim at investigating other polymer models as well as other topological properties such as the knotting probability and the mean writhe which so far has been only investigated using numerical methods.

Acknowledgments

J. Arsuaga was supported in part by NIH grant 2S06GM52588-12 and Y. Diao is partially supported by NSF grant DMS-0712958. B. Borgo was supported by a grant from CSUPERB and the Dorothy Howell Foundation, and partially supported by the NIH UBM program to E. Connor.

References

[1] Arsuaga J and Diao Y 2008, *Computational and Mathematical Methods in Medicine* **9**(3-4) 303–16.
[2] Arsuaga J, Blackstone T, Diao Y, Karadayi E and Saito M 2007, *J Phys A Math Gen* **40** 11697–711.
[3] ——et el 2002, *Proc Natl Acad Sci USA* **99** 5373–7.
[4] ——et al 2005, *Proc Natl Acad Sci USA* **102** 9165–9.
[5] Casjens S 1997, *Structural Biology of Viruses* (Chiu W, Burnett R and Garcea R Eds), Oxford University Press 1–37.
[6] Comolli L R et al 2008, *Virology* **371**(2) 267–77.
[7] Diao Y 1995, *J Knot Theory Ramifications* **4**(2) 189–96.
[8] ——et al 2003, *J Phys A Math Gen* **36**(46) 11561–74.
[9] ——, Nardo J and Sun Y 2001, *J Knot Theory Ramifications* **10**(4) 597–607.
[10] ——, Pippenger N and Sumners D W 1994, *J Knot Theory Ramifications* **3**(3) 419–29.
[11] Dobay A, Dubochet J, Millett K, Sottas PE, Stasiak A. *Proc Natl Acad Sci U S A.* **100** (10) 5611–5615
[12] Holmes V F and Cozzarelli N R 2000, *Proc Natl Acad Sci USA* **97** 1322-4.
[13] Huang J Y and Lai P Y 2001, *Phys Rev E* **63** 021506.1–6.
[14] Johnson J E and Chiu W 2007, *Curr Opinion Struct Biol* **17** 237–43.
[15] Katritch V et al 1996, *Nature* **384** 142–5.
[16] Millett K 2000, *Knots in Hellas'98 (Delphi), Series on Knots and Everything* **24**, World Scientific 306–34.
[17] Klenin K, Langowski J 2000, *Biopolymers* **54**(5) 307–17.
[18] Lepault J, Dubochet J, Baschong W and Kellenberger E 1987, *EMBO J* **6**(5) 1507–12.
[19] Liu L F, Davis J L and Calendar R 1981, *Nucleic Acids Res* **9** 3979–89.
[20] Micheletti C, Marenduzzo D, Orlandini E and Sumners D W 2006, *J Chem Phys* **124** 064903.1–10.
[21] J. P. J. Michels and F. W. Wiegel *Proc. R. Soc. London Ser A* **403** (1986), 269-284.
[22] ——and ——1982, *Phys Letts A* **90** 381–4.
[23] Millett KC, Plunkett P, Piatek M, Rawdon EJ, Stasiak A *J. Chem. Phys.* **130** (2009), 165104.
[24] Petrov A S, Boz M B and Harvey S C 2007, *J Struct Biol* **160** 241–8.
[25] Rayleigh L 1919, *Phil Mag S 6* **37**(220) 321–47.
[26] Rishovd S, Holzenburg A, Johansen B and Lindqvist B 1998, *Virology* **245** 11–7.
[27] Rybenkov V V, Cozzarelli N R and Vologodskii A V 1993, *Proc Natl Acad Sci USA* **90** 5307–11.
[28] Stasiak A, Katritch V, Bednar J, Michoud D and Dubochet J 1996, *Nature* **384** 122.
[29] Tzill S, Kindt J T, Gelbart W and Ben-Shaul A 2003, *Biophys J* **84** 1616–27.
[30] Scharein R, *KnotPlot*, software available online at `http://www.knotplot.com`.
[31] Shimamura M K and Deguchi T 2002, *Phys Rev E* **66** 040801–4.
[32] Varela R, Hinson K, Arsuaga J and Diao Y 2009, *J Phys A Math Gen* **42**(9) 095204.
[33] Vologodskii A V et al 1998, *J Mol Biol* **278** 1–3.