Curvature of random walks and random polygons in confinement

Y. Diao[†], C. Ernst[*], A. Montemayor[*], and U. Ziegler[*]

[*]Department of Mathematics and Computer Science
Western Kentucky University
Bowling Green, KY 42101, USA
[†]Department of Mathematics and Statistics
University of North Carolina Charlotte
Charlotte, NC 28223

Abstract. The purpose of this paper is to study the curvature of equilateral random walks and polygons that are confined in a sphere. Curvature is one of several basic geometric properties that can be used to describe random walks and polygons. We show that confinement affects curvature quite strongly, and in the limit case where the confinement diameter equals the edge length the unconfined expected curvature value doubles from $\pi/2$ to π. To study curvature a simple model of an equilateral random walk in spherical confinement in dimensions two and three is introduced. For this simple model we derive explicit integral expressions for the expected value of the total curvature in both dimensions. These expressions are functions that depend only on the radius R of the confinement sphere. We then show that the values obtained by numeric integration of these expressions agrees with numerical average curvature estimates obtained from simulations of random walks. Finally, we compare the confinement effect on curvature of random walks with random polygons.

1. Introduction

There have been numerous studies, both theoretical and numerical, on equilateral random polygons (also known as ideal random polygons), which are often used to model ring polymers under the θ-conditions where polymer segments that are not in a direct contact neither attract nor repel each other. In addition there have been studies on geometric properties of random polygons. For example, the overall dimensions such as the average end-to-end distance or the average radius of gyration is known to scale with the number of segments n as $\sqrt{n}$ [7, 8, 9, 18], the average crossing number grows as $O(n \ \ln \ n)$ [2] and the average squared writhe is believed (with numerical and partial analytical evidence) to grow as $O(n)$ [14]. Moreover, there are several generating methods that are well tested and efficient in generating random equilateral polygons such as the hedgehog method [6], the crankshaft method [12, 13], or the generalized hedgehog method [17]. These methods are all well tested and believed to generate non-correlated samples of equilateral random polygons. In the case of the generalized hedgehog method (developed by one of the PIs), a length n equilateral random polygon of length n can be generated in $O(n)$ time [17].

The theory of confined random polygons is much less well developed, in particular there are only very few analytic results and mostly numerical studies. One example is a numeric study of the average crossing number [1]. In [3, 4, 5] the authors introduced models of equilateral random polygons in spherical confinement that are based on explicit probability density functions for each step of the generation process. These methods could allow for some analytic analysis and in this paper such an analysis is attempted for curvature. The motivation of such an equilateral random polygon model is the well known fact of the highly compact packing of genomic material (long DNA chains) inside living organisms observed in macromolecular self-assembly processes in the complex network of interactions that take place in every organism. Even in the case of a simple organism such as viruses, the DNA packing is of high density. For example, in the prototypic case of the P4 bacteriophage virus, the $3\mu m$-long double-stranded DNA is packed within a viral capsid with a caliper size of about $50nm$, corresponding to a 70-fold linear compaction [11]. Unlike equilateral random polygons without confinement, the confined equilateral random polygons have not been thoroughly studied and there are many unanswered questions, see for example [15] for the data on curvature of unconfined polygons. In this article we study the effects of confinement on the curvature of walks and polygons.

It is easy to understand that a random walk/polygon in confinement has a larger expected curvature value than its unconfined counterpart: the physical confinement condition forces the random walk/polygon to make more turns in order to avoid hitting the confining sphere. In the extreme case where the diameter of the confinement sphere is close to the edge length of the random walk/polygon, the walk/polygon has to make near 180 degree turns at each step hence the average curvature would be close to $(n-1)\pi$ where n is the number of setments in the walk/polygon. On the other hand,

if the confinement radius is large (for example comparable to the total length of the walk/polygon) and the starting point of the walk/polygon is far away from the confining boundary, then the walk/polygon behaves as if it is unconfined. In the unconfined case, the average curvature per vertex is precisely $\pi/2$ for a random walk, except for the first and last vertices. Thus if the total length of the random walk is n, then its average curvature would be exactly $(n-1)\pi/2$. On the other hand, for an unconfined random polygon, its average curvature per vertex also approaches $\pi/2$ as n approaches infinity, although its total average curvature is larger. In fact it has been shown in [10] that for large n the average curvature of an unconfined random polygon is close to $n\pi/2+3\pi/8$.

The generation methods of equilateral confined random polygons introduced in [3, 4, 5] all rely on the use of explicit probability density functions that guide the generation of the polygons step-by-step. However these probability density functions are quite complicated. Thus an explicit expression of the expected curvature for random polygons will be hard to derive, if it is at all possible. In this paper we try to investigate the mean curvature for random polygons using an indirect approach. Since the mean curvature of random polygons behaves like the mean curvature of random walks when the random polygons and the random walks approach the same vertex density in the confinement sphere, it is sufficient to study random walks in confinement instead. Since the probability density functions involved in the case of random walks are simpler, both the theoretical and numerical studies on the mean curvature of confined random walks are easier. By the results in [5], these results can then be used to make inference about the mean curvature of confined random polygons.

The paper is organized as follows: In Section 2, a simple model of a two-dimensional random walk in confinement is introduced and an explicit expression of the expected curvature is derived that only depends on the radius of confinement R. In Section 3 we do the same for a three-dimensional random walk model. Numerical results are presented in Section 4. We conclude the paper with some discussions on the mean curvature of confined random polygons and some open questions for potential future research in Section 5.

2. The expected curvature of two-dimensional confined random walks

2.1. The vertex distribution.

Let $S \in \mathbb{R}^2$ be a confining circle with radius $R > 1/2$ and consider an equilateral random walk W_k of length k confined in S. Let X_0, X_1, ..., X_k be the (consecutive) vertices of the random walk defined as a Markov chain where each X_{j+1} depends only on X_j in the following way: once X_j is chosen, X_{j+1} is chosen uniformly over the portion of the unit circle centered at X_j that is contained within S. We have the following theorem.

Theorem 1 *Let $f(r)$ be a probability density function defined by*

$$f(r) = \begin{cases} ar, & 0 \leq r \leq R-1; \\ \frac{ar}{\pi}\cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right), & |R-1| < r \leq R; \\ 0, & \text{otherwise} \end{cases}$$

for $1/2 < R$ where the constant a is chosen so that $\int_0^R f(r)dr = 1$. If the initial vertex X_0 of a 2-dimensional random walk W_k is chosen with the distribution $f(|X_0|)/(2\pi|X_0|)$, then each vertex X_j of W_k follows the same distribution $f(|X_j|)/(2\pi|X_j|)$.

Notice that the constant a is defined by

$$1/a = \int_0^{R-1} r dr + \int_{R-1}^{R} \frac{r}{\pi}\cos^{-1}(\frac{1+r^2-R^2}{2r})dr$$

for $R \geq 1$ and

$$1/a = \int_{|R-1|}^{R} \frac{r}{\pi}\cos^{-1}(\frac{1+r^2-R^2}{2r})dr$$

for $1/2 < R < 1$. However in both cases it can be shown (through elementary integration and algebraic manipulations) that

$$a = \frac{4\pi}{4R^2\sec^{-1}(2R) - \sqrt{4R^2-1}}.$$

Proof. Given the way W_k is defined, it suffices to prove the following: Let X and Y be two random points in S that are a unit distance apart where X is chosen using the probability distribution $f(|X|)/(2\pi|X|)$ and where Y is chosen using a uniform distribution over the portion of the unit circle centered at X that is contained in S. It now suffices to show that Y follows the same probability distribution as X.

Let $g(u)$ be the probability density function of $u = |Y|$, $f(r)$ be the probability density function of $r = |X|$ and let $g(u|r)$ be the conditional probability density function of $|Y| = u$ under the condition that $|X| = r$. We have:

$$g(u) = \int_0^R g(u|r)f(r)dr.$$

Note that $g(u|r)$ for a fixed u is zero for values of $r \notin [|u-1|, min\{u+1, R\}]$ and thus

$$g(u) = \int_{|u-1|}^{min\{u+1,R\}} g(u|r)f(r)dr. \tag{1}$$

Let θ be the angle between $\overrightarrow{XO}$ and $\overrightarrow{XY}$, where O is the center of S, then $|Y|^2 = 1 + r^2 - 2r\cos\theta$. If $r \leq R-1$ the confining condition does not apply to Y and θ is uniformly distributed, that is $\theta \sim U[0,\pi]$. If $R-1 < r \leq R$ the confining

condition applies to Y by forcing $|Y| \leq R$ and thus $\theta \sim U[0, \cos^{-1}(\frac{1+r^2-R^2}{2r})]$. It follows that if $r \leq R-1$, then

$$\begin{aligned} P(|Y| \leq u|r) &= P(1+r^2-2r\cos\theta \leq u^2) \\ &= P(\cos\theta \geq \frac{1+r^2-u^2}{2r}) \\ &= \frac{1}{\pi}\cos^{-1}\left(\frac{1+r^2-u^2}{2r}\right). \end{aligned}$$

Differentiating with respect to u yields

$$g(u|r) = g_1(u|r) = \frac{2u}{\pi\sqrt{4r^2-(1+r^2-u^2)^2}}, \tag{2}$$

where $|r-1| \leq u \leq r+1$.

On the other hand, if $R-1 < r \leq R$, then

$$\begin{aligned} P(|Y| \leq u|r) &= P(1+r^2-2r\cos\theta \leq u^2) \\ &= P(\cos\theta \geq \frac{1+r^2-u^2}{2r}) \\ &= \frac{\cos^{-1}\left(\frac{1+r^2-u^2}{2r}\right)}{\cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right)}. \end{aligned}$$

It follows that if $R-1 < r \leq R$, then

$$g(u|r) = g_2(u|r) = \frac{2u}{\cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right)\sqrt{4r^2-(1+r^2-u^2)^2}}, \tag{3}$$

where $|r-1| \leq u \leq R$.

The integral (1) can now be written as the sum of two integrals of equations (2) and (3).

$$g(u) = \int_{|u-1|}^{min\{u+1,b_1\}} g_1(u|r)f(r)dr + \int_{min\{u+1,b_1\}}^{min\{u+1,R\}} g_2(u|r)f(r)dr,$$

where $b_1 = max\{R-1, |u-1|\}$.

Note that in the above equation if $|u-1| > R-1$, the the integral on the left is 0 since its upper and lower bound are identical. Similarly if $u+1 \leq R-1$ then the integral on the right is zero. Thus the intervals for the two integrals are lined up with the intervals of $f(r)$. This allows us to combine the two integrals into one as follows:

$$\begin{aligned} g(u) &= \int_{|u-1|}^{min\{u+1,b_1\}} g_1(u|r)f(r)dr + \int_{min\{u+1,b_1\}}^{min\{u+1,R\}} g_2(u|r)f(r)dr \\ &= \int_{|u-1|}^{min\{u+1,b_1\}} \frac{2u}{\pi\sqrt{4r^2-(1+r^2-u^2)^2}} ardr + \\ &\quad + \int_{min\{u+1,b_1\}}^{min\{u+1,R\}} \frac{2u\frac{ar}{\pi}\cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right)}{\cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right)\sqrt{4r^2-(1+r^2-u^2)^2}} dr \end{aligned}$$

$$= \int_{|u-1|}^{min\{u+1,b_1\}} \frac{2uar}{\pi\sqrt{4r^2-(1+r^2-u^2)^2}}dr + \\ + \int_{min\{u+1,b_1\}}^{min\{u+1,R\}} \frac{2uar}{\pi\sqrt{4r^2-(1+r^2-u^2)^2}}dr \\ = \int_{|u-1|}^{min\{u+1,R\}} \frac{2uar}{\pi\sqrt{4r^2-(1+r^2-u^2)^2}}dr. \quad (4)$$

To evaluate (4) we now consider two cases:

Case 1. $0 \leq u \leq R-1$. In this case, $min\{u+1, R\}$ is always equal to $u+1$. Thus we have

$$g(u) = \int_{|u-1|}^{u+1} \frac{2uar}{\pi\sqrt{4r^2-(1+r^2-u^2)^2}}dr = au = f(u).$$

Case 2. $|R-1| < u \leq R$. In this case, $min\{u+1, R\}$ is always equal to R. Thus we have

$$g(u) = \int_{|u-1|}^{R} \frac{2uar}{\pi\sqrt{4r^2-(1+r^2-u^2)^2}}dr \\ = \frac{au}{\pi}\cos^{-1}(\frac{1+u^2-R^2}{2u}) = f(u).$$

This finishes the proof. □

Note that the the vertex distribution of the random walk generated by the probability density function of Theorem 1 is non-uniform. For $R > 1$, it is an increasing function for $r < R-1$ and is a decreasing function for $R > r > R-1$ as shown in Figure 1.

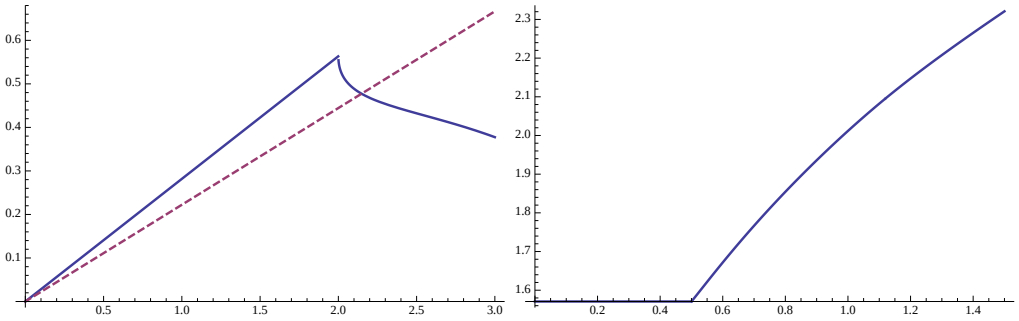

Figure 1. Left: The probability density function $f(r)$ given in Theorem 1 with the confinement radius $R = 3$. The probability density function $2r/R^2$ (corresponding to the uniform distribution in the confining circle) is plotted with the dashed line for comparison purposes. Right: The expected curvature between two consecutive edges as a function of the distance between the common vertex of the two edges to the center of the confining sphere (with $R = 1.5$), see Theorem 2.

2.2. The curvature dependence on the distance of a vertex to the boundary.

Let c_j be the curvature between the $(j-1)$-th edge and the j-th edge, *i.e.*, the angle between $\overrightarrow{X_{j-1}X_j}$ and $\overrightarrow{X_jX_{j+1}}$. Our next step is to derive an expression for $E(c_j)$, the average curvature at vertex $X = X_j$. Since the vertices share the same distribution, the two random vectors $\overrightarrow{X_{j-1}X_j}$ and $\overrightarrow{X_jX_{j+1}}$ can be generated by choosing X_j first following the distribution given in Theorem 1, then choosing X_{j-1} and X_{j+1} uniformly over the portion of the unit circle centered at X_j that is within the confining sphere S. For the sake of simplicity let $X = X_j$, $Y = X_{j-1}$ and $Z = X_{j+1}$. The corresponding mean curvature is actually label independent and only depends on $r = |X|$ so we denote it by c_r. We now consider several different cases.

Case 1. $R \geq 1$ and $|X| = r \leq R - 1$. In this case there are no restrictions on the possible angles and $E(c_r) = \pi/2$.

Case 2. $R \geq 1$ and $R - 1 < |X| = r \leq R$. In this case the unit circle centered at X intersects the confining sphere S at two points (which are denoted by Q_1 and Q_2). The portion of the unit circle that is contained in S has length $2\cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right)$ hence is spanned by an angle of the same measure. Since Y and Z are uniformly distributed on this part of the unit circle, they are uniquely determined by the angle measured from $\overrightarrow{XQ_1}$ to $\overrightarrow{XY}$ or $\overrightarrow{XZ}$ respectively, following the direction of the arc that is within S. If we call these two angles α and β respectively, then α, β are uniformly distributed over $[0, 2\theta_{\max}]$ where $\theta_{\max} = \cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right)$. Furthermore, $c_r = |\pi - |\beta - \alpha||$. There are two subcases that require different treatment.

Subcase 2A. $2\theta_{\max} \leq \pi$. In this case,$|\beta - \alpha| \leq \pi$ hence $c_r = \pi - |\beta - \alpha|$. Thus

$$\begin{aligned} E(c_r) &= E(\pi - |\beta - \alpha|) = \pi - E(|\beta - \alpha|) \\ &= \pi - \frac{1}{2\theta_{\max}} \int_0^{2\theta_{\max}} \left(\frac{1}{2\theta_{\max}} \int_0^{2\theta_{\max}} |\alpha - \beta| d\beta \right) d\alpha \\ &= \pi - \frac{2\theta_{\max}}{3} = \pi - \frac{2}{3} \cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right). \end{aligned}$$

Subcase 2B. $2\theta_{\max} > \pi$. This case is more complicated since it is possible that $|\beta - \alpha| > \pi$ and in which case $c_r = |\beta - \alpha| - \pi$ instead. Here we derive the probability distribution function of c_r directly. For $0 \leq t \leq 2\theta_{\max} - \pi$ (notice that $2\theta_{\max} - \pi \leq \pi$ since $\theta_{\max} \leq \pi$) we have

$$\begin{aligned} P(c_r \leq t) &= P(|\pi - |\beta - \alpha|| \leq t) = P(\pi - t \leq |\beta - \alpha| \leq \pi + t) \\ &= P(|\beta - \alpha| \leq \pi + t) - P(|\beta - \alpha| \leq \pi - t). \end{aligned}$$

On the other hand, for any $0 \leq s \leq 2\theta_{\max}$, we have

$$\begin{aligned} P(|\alpha - \beta| \leq s) &= P(\beta - s \leq \alpha \leq \beta + s) \\ &= 1 - P(\alpha \geq \beta + s) - P(\alpha \leq \beta - s) \\ &= 1 - P(\beta \geq \alpha + s) - P(\alpha \leq \beta - s) \end{aligned}$$

$$= 1 - 2P(\alpha \leq \beta - s)$$
$$= 1 - \frac{(2\theta_{\max} - s)^2}{4\theta_{\max}^2} = \frac{s}{\theta_{\max}} - \frac{s^2}{4\theta_{\max}^2}.$$

Thus

$$P(c_r \leq t) = P(|\beta - \alpha| \leq \pi + t) - P(|\beta - \alpha| \leq \pi - t)$$
$$= \frac{(\pi + t)}{\theta_{\max}} - \frac{(\pi + t)^2}{4\theta_{\max}^2} - \frac{(\pi - t)}{\theta_{\max}} + \frac{(\pi - t)^2}{4\theta_{\max}^2} = \frac{2t}{\theta_{\max}} - \frac{\pi t}{\theta_{\max}^2}. \quad (5)$$

For $2\theta_{\max} - \pi \leq t \leq \pi$, using the fact that $P(|\beta - \alpha| \leq 2\theta_{\max} + t) = 1$ and an argument similar to the above, we get

$$P(c_r \leq t) = 1 - \frac{\pi - t}{\theta_{\max}} + \frac{(\pi - t)^2}{4\theta_{\max}^2}. \quad (6)$$

Combining equations (5) and (6) we obtain the probability density function $h(t)$ for the curvature c:

$$h(t) = \begin{cases} \frac{2}{\theta_{\max}} - \frac{\pi}{\theta_{\max}^2}, & 0 \leq t \leq 2\theta_{\max} - \pi \\ \frac{1}{\theta_{\max}} - \frac{(\pi - t)}{2\theta_{\max}^2}, & 2\theta_{\max} - \pi \leq t \leq \pi. \end{cases}$$

It now follows that

$$E(c_r) = \int_0^{\pi} t \cdot h(t) dt = \frac{1}{3}(2\theta_{\max} + \frac{3\pi^2}{\theta_{\max}} - \frac{\pi^3}{2\theta_{\max}^2} - 3\pi).$$

Case 3. $\frac{1}{2} < R \leq 1$. If $R \leq 1$, for all vertices X_j of the random walk the following holds: $1 - R \leq |X_j| \leq R$ and as before the maximal angle which can be achieved between the y-axis and any possible edge connected to X is $\theta_{\max} = \cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right)$ and $\theta_{\max} \leq \frac{\pi}{2}$. Thus the derivation from case 2A applies.

We have now obtained the following theorem:

Theorem 2 *In a 2-dimensional random walk in a confining sphere of radius R, the expected curvature $E(c_r)$ of a vertex X with $|X| = r$ is given by:*
if $R \geq 1$

$$E(c_r) = \begin{cases} \frac{\pi}{2}, & 0 \leq r \leq R - 1 \\ \frac{1}{3}(2\theta_{\max} + \frac{3\pi^2}{\theta_{\max}} - \frac{\pi^3}{2\theta_{\max}^2} - 3\pi), & R - 1 < r \leq \sqrt{R^2 - 1} \\ \pi - \frac{2\theta_{\max}}{3}, & \sqrt{R^2 - 1} \leq r \leq R, \end{cases}$$

and if $1/2 < R < 1$

$$E(c_r) = \begin{cases} 0 & 0 \leq r \leq 1 - R \\ \pi - \frac{2\theta_{\max}}{3}, & 1 - R \leq r \leq R, \end{cases}$$

where $\theta_{\max} = \cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right)$.

The explicit expression of $E(c_r)$ given in Theorem 2 shows how the mean curvature at a given vertex is affected by the distance of the vertex to the boundary of the confining circle and it enables us to compute the mean total curvature $C_{2,n}(R)$ of a random walk of n edges confined in a circle of radius R:

$$C_{2,n}(R) = (n-1)\int_{|min\{0,R-1\}|}^{R} E(c_r)f(r)dr, \tag{7}$$

where n is the number of edges of the random walk. This expression is also valid for values $1/2 < R \leq 1$. In this case the boundary of the integral simplifies and we obtain

$$C_{2,n}(R) = (n-1)\int_{1-R}^{R} E(c_r)f(r)dr. \tag{8}$$

For the three dimensional case, unfortunately, an explicit expression like $E(c_r)$ is no longer feasible and a different approach needs to be taken. In the following we present a different way to compute $C_{2,n}(R)$. It is not as intuitive as the above approach, but can be easily extended to the three dimensional case.

Theorem 3 *Let $f(r)$ be as defined in Theorem 1, then the mean total curvature of a 2-dimensional walk with n edges in a confining sphere of radius R is*

$$C_{2,n}(R) = (n-1)\int_{c}^{R}\int_{a}^{b}\int_{-\pi}^{\pi} \cos^{-1}(\cos\,(\alpha-\beta))h(\beta|r)g(r|s)f(s)d\beta dr ds,$$

where $\alpha = \cos^{-1}\left(\frac{s^2-1-r^2}{2r}\right)$, $c = |min\{0, R-1\}|$, $a = |s-1|$, $b = \min\{s+1, R\}$, and

$$h(\beta|r) = \begin{cases} \frac{1}{2\pi}, & -\pi \leq \beta \leq \pi \text{ and } \frac{1+r^2-R^2}{2r} \leq -1 \\ \frac{1}{2\theta_{\max}}, & -\theta_{\max} \leq \beta \leq \theta_{\max} \text{ and } -1 < \frac{1+r^2-R^2}{2r} < 1 \end{cases}$$

where $\theta_{\max} = \cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right)$ as defined before.

Proof. Consider three points along the random walk X_{k-1}, X_k, and X_{k+1}. Set $|X_{k-1}| = s$, $|X_k| = r$, the angle between $\overrightarrow{X_{k-1}X_k}$ and $\overrightarrow{X_kO}$ as $\alpha \in [0, \pi]$, the angle between $\overrightarrow{X_kX_{k+1}}$ and $\overrightarrow{X_kO}$ as $\beta \in [-\pi, \pi]$, and finally the angle between $\overrightarrow{X_{k-1}X_k}$ and $\overrightarrow{X_kX_{k+1}}$ as $\gamma \in [0, \pi]$ (γ equals to the curvature at X_k), see Figure 2. We have $\cos(\alpha) = (s^2 - 1 - r^2)/(2r)$ and $\cos(\gamma) = \cos(\alpha - \beta)$.

So the expected curvature for a confinement radius R over the sample space V is then

$$\int_V \cos^{-1}(\cos(\alpha-\beta))h(\beta|r)g(r|s)f(s)d\beta dr ds$$

where $f(s)$ is given by Theorem 1 and $g(r|s)$ is given in equations (2) and (3). $h(\beta|r)$, the density function of β given the fixed r, is the uniform distribution over the interval of possible choices for β. That is, if $r \leq R-1$ then the confinement condition does not apply and $\beta \sim U[-\pi, \pi]$. If $r > R-1$ the confinement condition applies and $\beta \sim U[-\theta_{\max}, \theta_{\max}]$.

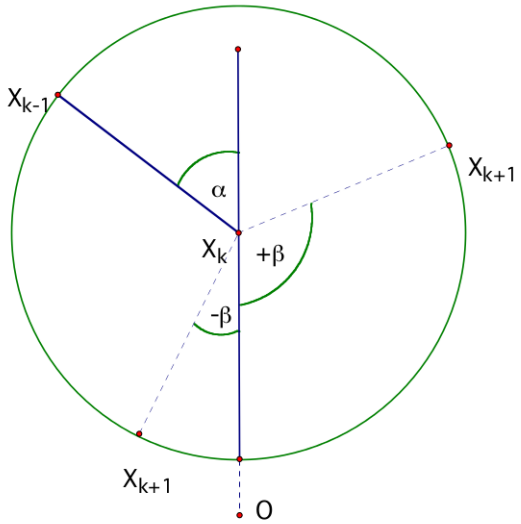

Figure 2. Shown are the three vertices X_{k-1}, X_k, and X_{k+1} together with the angles α and β. By rotational symmetry we can assume X_k is on the y-axis, and by reflectional symmetry we can assume that X_{k-1} is on the left side of the y-axis. For X_{k+1} two possibilities are shown one with a positive angle β and the second with a negative β indicated by the $\pm$ symbol.

3. The expected curvature of three-dimensional confined random walks

We now study the problem in three dimensions. Let $S \in \mathbb{R}^3$ be a confining sphere with radius $R > 1/2$ and consider an equilateral random walk W_k of length k confined in S. Let X_0, X_1, ..., X_k be the vertices of the random walk. Here, the random walk is defined as a Markov chain: each X_{j+1} depends only on X_j in the following way: once X_j is chosen, X_{j+1} is chosen uniformly over the portion of the unit sphere centered at X_j that is contained within S. In [5] the following Theorem is established:

Theorem 4 *Let $f(r)$ be a probability density function defined by*

$$f(r) = \begin{cases} ar^2, & 0 < r \leq R-1; \\ \frac{ar}{4}(R^2-(r-1)^2), & R-1 < r \leq R; \end{cases}$$

where $a = 48/(16R^3 - 12R^2 + 1) = 48/((2R-1)^2(4R+1))$. If the initial vertex X_0 of W_k is chosen with the distribution $f(|X_0|)/(4\pi|X_0|^2)$, then each vertex X_j of W_k follows the same distribution $f(|X_j|)/(4\pi|X_j|^2)$.

Note the vertex distribution of the random walk generated by the probability density function of Theorem 4 is non-uniform. The density declines when r is close to R, see Figure 3.

Let the center of confining sphere S be the origin O. Consider three consecutive vertices X_{k-1}, X_k, X_{k+1} along the random walk. For the purpose of computing the

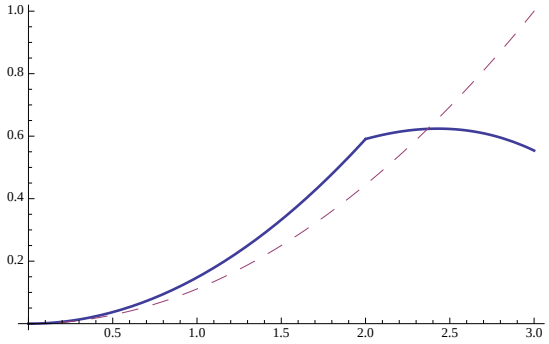

Figure 3. The probability density function $f(r)$ (of $r = |X|$) for $R = 3$ as defined in Theorem 4. The probability density function $3r^2/R^3$ of $r = |X|$ corresponding to the uniform distribution of X is shown with the dashed line for comparison.

mean curvature at X_k, without loss of generality we can assume that X_k is on the non-negative z axis and that the point X_{k-1} is in the xz plane (since this can be done through rotations and rotations do not change the curvature). Denote $|X_{k-1}| = s$, $|X_k| = r$, the angle between $\overrightarrow{X_{k-1}X_k}$ and $\overrightarrow{X_kO}$ as $\alpha \in [0, \pi]$, the angle between $\overrightarrow{X_kX_{k+1}}$ and $\overrightarrow{OX_k}$ as $\beta \in [0, \pi]$, and finally the angle between $\overrightarrow{X_kX_{k-1}}$ and $\overrightarrow{X_kX_{k+1}}$ as $\gamma' \in [0, \pi]$ (the curvature at X_k is then defined by $\gamma = \pi - \gamma'$), see Figure 4. Let τ be the angle between the plane $OX_{k-1}X_k$ and OX_kX_{k+1}, measured in a clockwise direction starting with the $OX_{k-1}X_k$ plane so $\tau \in [0, 2\pi)$. In fact, by a symmetry argument, it is easy to see that $\tau \sim U[0, 2\pi)$. (Note that if the four points O, X_{k-1}, X_k and X_{k+1} are collinear or coplanar then $\tau = 0$.) So if X_{k-1}, X_k have been chosen (with X_k on the z axis and X_{k-1} on the xz plane with positive x coordinate), then X_{k+1} can be determined (generated) by selecting τ (uniformly from $[0, 2\pi)$) and β (according to its distribution subject to the chosen r value). Let N be the north pole of the unit sphere $S_1(X_k)$ centered at X_k and consider the spherical triangle $\Delta X_{k-1}X_{k+1}N$ on $S_1(X_k)$: NX_{k-1} has arc length α, NX_{k+1} has arc length β, $X_{k-1}X_{k+1}$ has arc length γ' and the angle opposite to $X_{k-1}X_{k+1}$ is τ (see Figure 4 for an illustration). Applying the law of cosines for the sides of a spherical triangle, we have

$$\cos\gamma' = \cos\alpha\cos\beta + \sin\alpha\sin\beta\cos\tau. \tag{9}$$

The values $\cos\alpha$ and $\sin\alpha$ are determined by the values of s and r as

$$\cos\alpha = \frac{s^2 - 1 - r^2}{2r} \quad \text{and} \quad \sin\alpha = \sqrt{1 - \left(\frac{\mathrm{s}^2 - 1 - \mathrm{r}^2}{2\mathrm{r}}\right)^2}.$$

Substituting this into equation (9) yields

$$\gamma = \cos^{-1}\left(-\frac{s^2 - 1 - r^2}{2r}\cos\beta - \sqrt{1 - \left(\frac{s^2 - 1 - r^2}{2r}\right)^2}\sin\beta\cos\tau\right). \tag{10}$$

Assume that X_{k-1}, X_k, X_{k+1} are generated in the following order and manner. First X_{k-1} is generated by choosing $s = |X_{k-1}|$ according to the distribution $f(s)$. Then X_k

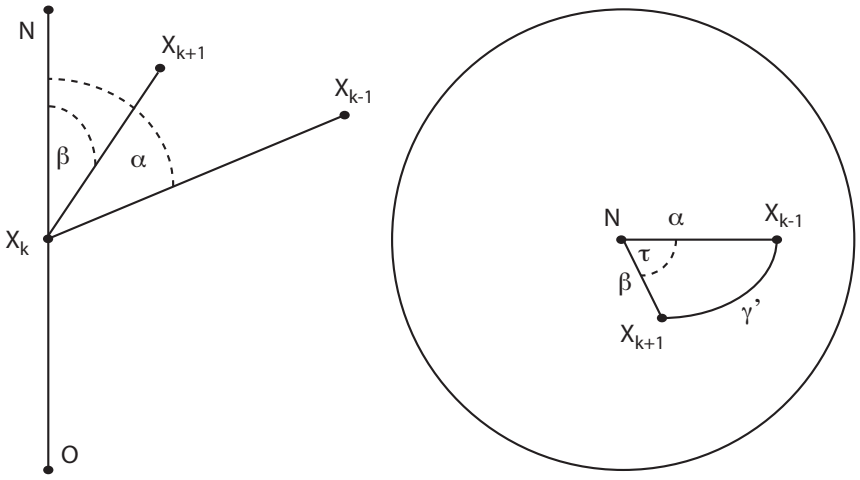

Figure 4. Left: An illustration of the angles α and β; Right: The top view of the spherical triangle $\Delta X_{k-1}X_{k+1}N$ where X_k is directly underneath N.

is chosen by selecting $r = |X_k|$ according to its distribution $p(r|s)$ (conditioned to the chosen $s = |X_{k-1}|$). By the reflection and rotation symmetry as we mentioned before we can assume that X_k is on the positive z axis and X_{k-1} is in the half xz plane with positive x axis direction. We then select the angle τ uniformly from $[0, 2\pi)$, and finally choose the angle β. β is independent of τ but depends on r, so its probability distribution $k(\beta|r)$ is conditioned by the given $r = |X_k|$. In order to obtain the integral form of the mean total curvature $C_{3,n}(R)$ of a random equilateral walk of n edges that is confined in a sphere of radius R, it is necessary for us to obtain the density functions $p(r|s)$ and $k(\beta|r)$ first.

For the unconfined case, it is well known that $k(\beta|r) = \sin(\beta)/2$ and $\cos(\beta) \sim U[-1, 1]$ [16]. For the confined case, denote the quantity $\frac{R^2-1-r^2}{2r}$ by z_r and we have the following result.

$$k(\beta|r) = \begin{cases} \frac{\sin\beta}{2}, & \beta \in [0, \pi] \text{ and } \mathrm{z_r} \geq 1; \\ \frac{\sin\beta}{1+z_r}, & \beta \in [\cos^{-1}(z_r), \pi] \text{ and } -1 < \mathrm{z_r} < 1, \end{cases} \tag{11}$$

To see this, assume that $-1 < z_r < 1$ (so it is necessary that $R < r + 1$), then $\beta \geq \beta_r = \cos^{-1}(z_r)$ since otherwise X_{k+1} would be outside of the confining sphere. The portion of the unit sphere centered at X_k that is within the confining sphere thus has area $2\pi(\cos\beta_r - \cos\pi) = 2\pi(z_r + 1)$. Similarly, the portion of the unit sphere centered at X_k that is within the confining sphere that corresponds to $\beta_0 \geq \beta \geq \beta_r$ for any $\beta_0 > \beta_r$ has area $2\pi(\cos\beta_r - \cos(\beta_0)) = 2\pi(z_r - \cos(\beta_0))$. Thus the probability for X_{k+1} (which is equivalent to $\beta \leq \beta_0$ under the condition that $\beta \geq \beta_r$) to fall into this area is $(z_r - \cos(\beta_0))/(1 + z_r)$ and the result follows by taking derivative of this with respect to β_0 and reset the notation β_0 to β.

Finally, the probability density function $p(r|s)$ is given by the following Lemma. The unconfined version of this Lemma has been shown in [3].

Lemma 1 *Let X_k be uniformly distributed on the portion of the unit sphere centered at $X_{k-1} \neq O$ (with $|X_{k-1}| = s$) that is within the confining sphere S, and let $r = |X_k|$. Then the probability density function $p(r|s)$ of r is given by*

$$p(r|s) = \begin{cases} r/(2s), & r \in [|s-1|, s+1] \text{ if } s \leq R-1; \\ (2r)/(R^2-(s-1)^2), & r \in [|s-1|, R] \text{ if } s > R-1. \end{cases}$$

Proof. As before, we assume that X_{k-1} is on the z-axis so $X_{k-1} = (0, 0, s)$. Let η be the angle between $\overrightarrow{X_{k-1}X_k}$ and the positive z-axis $\overrightarrow{OX_{k-1}}$, where O is the origin (also the center of the confining sphere S). As we mentioned before, $\cos\eta$ is uniformly distributed on $[-1, 1]$ when the confinement condition does not apply. Let us first consider the case $0 \leq s \leq R-1$. Simple trigonometric calculations lead to

$$\begin{aligned} P(|X_k| \leq r) &= P(\cos\eta \leq \frac{r^2-1-s^2}{2s}) \\ &= \begin{cases} 1, & r > s+1; \\ \frac{1}{2}(1+\frac{r^2-1-s^2}{2s}), & |s-1| \leq r \leq s+1; \\ 0, & r < |s-1|. \end{cases} \end{aligned}$$

It follows that

$$\frac{dP(|X_k| \leq r)}{dr} = \begin{cases} \frac{r}{2s}, & |s-1| \leq r \leq s+1; \\ 0, & \text{otherwise.} \end{cases}$$

In the case that $s > R-1$, the only difference is now that $\cos\eta$ is uniformly distributed on $[-1, \frac{R^2-1-s^2}{2s}]$ instead. So

$$\begin{aligned} P(|X_k| \leq r) &= P(\cos\eta \leq \frac{r^2-1-s^2}{2s}) \\ &= \begin{cases} 1, & r > R; \\ \frac{2s}{R^2-(1-s)^2}(1+\frac{r^2-1-s^2}{2s}), & |s-1| \leq r \leq R; \\ 0, & r < |s-1|. \end{cases} \end{aligned}$$

The result now follows.

We now have all the pieces needed to write down an integral expression for the curvature of a 3 dimensional random walk, which we now give in the following theorem:

Theorem 5 *The expected total curvature of an equilateral random walk (as defined in this section) of n edges that is confined in a sphere of radius R, denoted by $C_{3,n}(R)$, is given by the following integral:*

$$C_{3,n}(R) = \frac{n-1}{2\pi}\int_0^R \int_a^b \int_0^{2\pi} \int_0^{\pi} \mathrm{curv}(\beta, \tau, \mathrm{r}, \mathrm{s})\mathrm{k}(\beta|\mathrm{r})\mathrm{p}(\mathrm{r}|\mathrm{s})\mathrm{f}(\mathrm{s})\mathrm{d}\beta\mathrm{d}\tau\mathrm{d}\mathrm{r}\mathrm{d}\mathrm{s},$$

where $a = |s-1|$, $b = \min\{R, s+1\}$, $\mathrm{curv}(\beta, \tau, \mathrm{r}, \mathrm{s}) = \cos^{-1}(-\mathrm{z}_{\mathrm{r,s}}\cos\beta - \sqrt{1-\mathrm{z}_{\mathrm{r,s}}^2}\sin\beta\cos\tau)$ and $z_{r,s} = (s^2-1-r^2)/(2r)$.

4. Numerical results

In this section we want to obtain curvature values that are independent of the length of the random walks. Thus we divide all total curvature values by $n-1$ and only talk about the mean curvature per turn. Unfortunately the integral forms of $C_{2,n}(R)$ and $C_{3,n}(R)$ as given in equations (7), (8) and in Theorems 3 and 5 do not have explicit expressions for most values of R. An exception is the case of $R = 1$ where Mathematica was able to calculate the integral given in Theorem 3 exactly:

$$C_{2,n}(1)/(n-1) = \frac{18 - 15\sqrt{3}\pi + 22\pi^2}{9(4\pi - 3\sqrt{3})} \approx 2.31427.$$

For $R = 1$, a numerically simulated 2D random walk of $2,000,000$ steps yielded a mean curvature of 2.31419 which is an error of less than 10^{-4}. For $R = 3$, a numeric integration of the integrals given in equations (7), (8) and in Theorem 3 resulted in $C_{2,n}(3)/(n-1) \approx 1.6981$ and 1.6979 respectively. A numerically simulated random walk of $1,000,000$ steps yielded a mean curvature of 1.69945. The results of a simulation of random walks with $1,000,000$ steps are shown in Figure 5 for R values in the interval $[.55, 3]$ with increments of length $1/20$. The mean of the absolute value of the difference of the simulation and the numeric integration is approximately 0.00063 for the integral given in equations (7) and (8), and approximately 0.0015 for the integral in Theorem 3. The difference in error is most likely due to difficulty in carrying out the numerical integration. For both integrals we used standard integration methods built into Mathematica. However, we repeated simulations of the random walks of length $1,000,000$ for a few selected values of R and found that the difference in curvature is of the order ranging between 10^{-4} and 10^{-3} and we observe that the difference between the simulations and the integrals falls into a similar range.

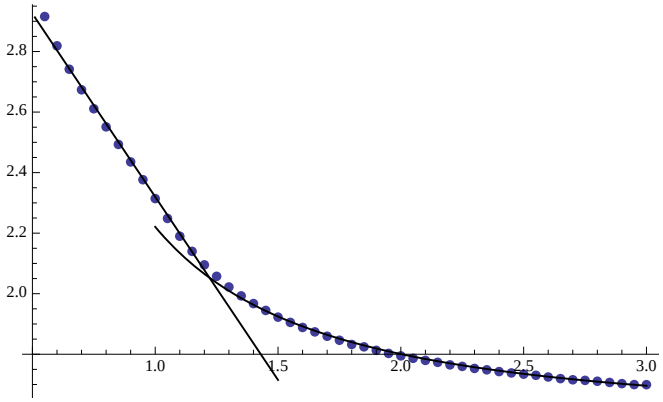

Figure 5. Average curvature per turn for radii $1/2+1/20 \leq R \leq 3$ with increments of length $1/20$. Each data point is obtained by a simulation of a single random walk with $1,000,000$ steps. The decrease in curvature seems to occur in two different regimes, represented by the linear function $3.53 - 1.21R$ and the function $\pi/2 + 0.65/R^{1.5}$. However the authors do not have a good explanation of this.

In Figure 6 we compare the curvature of two dimensional random walks with three dimensional random walks. As before each data point is based on a single random walk with $1,000,000$ steps. The curvature in two dimension is consistently higher than the curvature in three dimensions which is not surprising. Consider a point X which has distance $r = |X|$ from the origin. In both the two and three dimensional cases the range of possible curvature angles is the same: the minimal angle of curvature at X is zero for $r \leq \sqrt{R^2-1}$ and $\pi - 2\theta_{\max} = \pi - 2\cos^{-1}\left(\frac{1+r^2-R^2}{2r}\right)$ for $R \geq r > \sqrt{R^2-1}$; the maximal angle is of course π regardless of the distance r. Moreover, if we assume that X is close to the boundary of the confining sphere then we obtain a small angle of curvature at X if the two adjacent vertices along the random walk are also close to the boundary. The percentage of volume close to the boundary (of the total volume) is much smaller in the two dimensional case when compared with the three dimensional case. (For example the volume of points at least .9R away form the origin is 19% in the two dimensional case and 27.1% in the three dimensional case.) This means that the probability to have a small curvature angle at X is higher in the three dimensional case when compared with the two dimensional case.

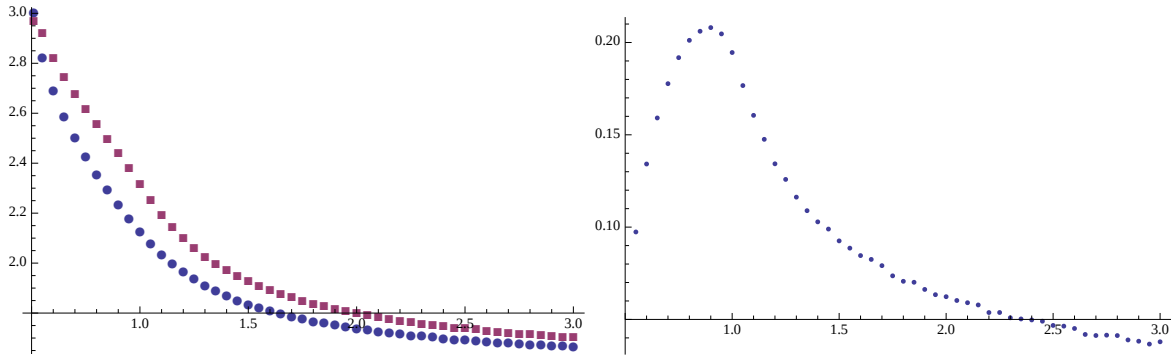

Figure 6. left: Average curvature per turn for radii $+1/20 \leq R \leq 3$ with increments of length $1/20$ (The 3D data contains additional values for $R = 0.501$ and $R = 0.505$). The curve with the square markers represents the two dimensional data and the curve with round markers represents the three dimensional data. Right: The difference between the two data sets.

Unfortunately, the multiple integral for $C_{3,n}(R)$ given in Theorem 5 is even more convoluted than its counterpart in the 2-dimensional case. However it can be numerically approximated by Mathematica using built-in "global adaptive" integration methods. We compared the curvature values obtained by numeric simulation of random walks with the results of numeric integration as shown in Figure 7 together with a graph of the difference between the simulation value (where each data point is based on a single random walk of one million steps) and the integration value. Note that the maximal difference between the two data sets shown in Figure 7 is reasonable except that there are larger differences close to radii of $1/2$ and of $3/2$. The mean of the absolute value of the difference over the data sets is approximately 0.003. While the larger error close to the minimal value of $R = 1/2$ may be due to numerical instability (both in the

integral and the random walk), the authors currently have no explanation why there is an instability for these R values close to $3/2$.

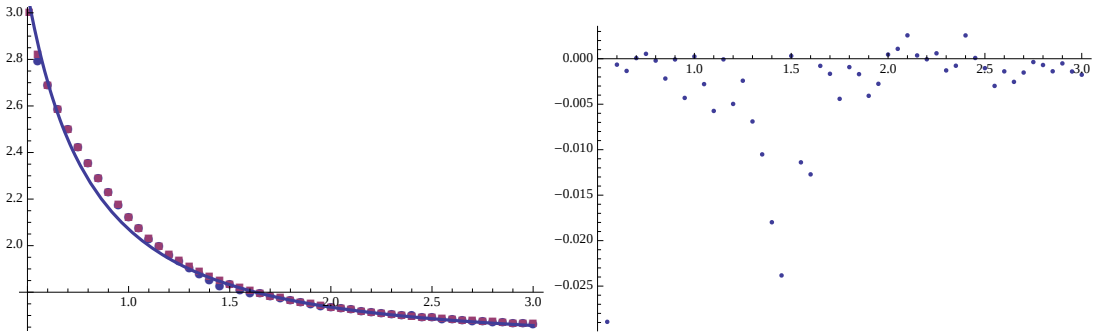

Figure 7. Left: The three dimensional data of Figure 6 together with data generated by numerical integration. Due to numerical instability the additional values of $R = 0.51$ and $R = 0.505$ are omitted for the integration. The squares represent the data points obtained by simulation of the walk and the circles represent the results of numerical integration. However, the difference can only be seen for R values around 1.5. Also shown is the function $\pi/2 + .5/x^{1.6}$ which fits the data reasonable well. Right: The difference between the simulation and the numerical integration.

5. Future work and open questions

We will conclude this paper with some rather detailed discussion concerning the mean curvature of a confined random polygon and some open questions for potential future studies. Having seen the results concerning the confined random walks, one cannot help but to ask: what about the confined random polygons? How does the average curvature per vertex of a confined random polygon compare with the average curvature per vertex of a confined random walk? In [10] it was shown that an unconfined random polygon of length n has a mean total curvature of $\pi n/2+3\pi/8$ for large n, thus the mean curvature per vertex of a long random polygon is about $\pi/2 + 3\pi/(8n)$, which is slightly larger than the mean curvature per vertex of a random walk (which is exactly $\pi/2$). Notice that a key fact that enabled us to derive the integral forms of the mean curvature for the confined random walks is that the random walks defined here have the property that every vertex has the same distribution hence the mean total curvature can be derived locally at any given vertex (except the first and the last). Unfortunately, for the several algorithms which created confined random polygons that have been introduced and studied (see [3, 4, 5]), that is not the case. Consequently, it is not possible to derive an integral form for the mean total curvature of these confined random polygons and we were only able to carry out some numerical studies. However, the results from the confined random walks provided excellent bench marks for comparison. Figure 8 shows the earlier data of Figure 7 compared with the curvature of a polygon. The polygon data is based on random polygons generated using the methods of [4]. Basically, in this method, the random polygons can be thought of as being generated with their starting

points fixed at the origin (that is, $X_0 = O$) and with no confinement, but only those that are contained in the confining sphere S are kept. Of course, it would be very time consuming to generate the confined random polygons in this naive way. Much faster algorithms were developed in [4]. Each of the data points are based on the average of $10,000$ polygons of 60 edges. However, only data for selected radii is available due to the computational difficulties in generating the random polygons. In addition, the algorithm of [4] assumes that the polygons start (and end) at the origin and thus the confinement radius cannot be less than one. For each random polygon let the total curvature be T_p. We then computed the mean of the values $(T_p - 3\pi/8)/60$ to compare with the mean curvature per vertex for the random walks under the corresponding confinement radii.

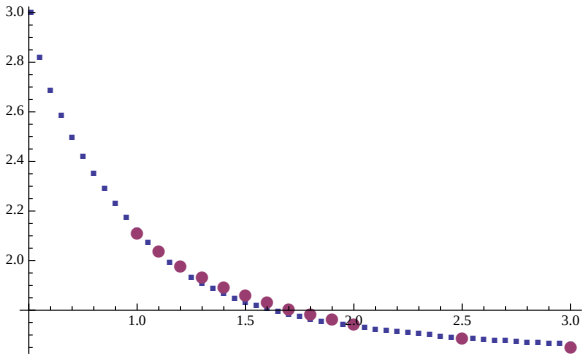

Figure 8. The data of Figure 7 of random walks (small circles) together with the data of length 60 random polygons (big circles).

As we can see the two data sets are surprisingly similar. It is surprising since we have good reasons to expect more significant differences. First, all random polygons start and end at the origin. Thus for larger radii polygons with a small number of segments might experience little or no confinement - which would not be the case for random starting points leading to some bias towards non-confinement for the larger radii. Second, for large length n the vertex distribution for the random polygons generated with the methods in [4] does not approach the vertex density of the random walks as shown in Figure 3. Thus it would be really interesting to see the more general $(T_p - 3\pi/8)/n$ behavior for much larger n values. Should $(T_p - 3\pi/8)/n$ behave like its counterparts in the confined random walks, then it may be telling us something intrinsic about the nature of confinement that is less dependent on how the polygons/walks are generated.

We shall now end our paper with several open questions for potential future work in this direction.

1. Torsion is another geometric quantity that in some sense is similar to curvature. For an unconfined random walk the average torsion value is also $\pi/2$ per vertex just as for curvature. Preliminary studies have shown that confinement decreases torsion

somewhat for R values close to $1/2$ to $\pi/3$. Is it possible to do a similar analysis with torsion as it is done here for curvature?

2. The functions we obtained here for the expected curvature of random walks are complicated due to the convoluted integrals. Is this due to the particular ways the random walks are defined, or is it something that is due to the nature of the confinement?

3. Random polygons can be generated in different ways, see [5]. Depending on the generation process we obtain random polygons that have different vertex densities and thus random polygons that may have slightly different behavior in their curvature. Is it possible to quantify these differences?

4. What is the effect of knotting on the curvature of random polygons. For example, if we consider the 60-step polygons generated with radius $R = 1$ (i.e the left most small data point in Figure 8), then these polygons are highly knotted. In fact almost all of them will represent knots outside the knot table (i.e. with more than 16 crossings). For these highly confined polygons is it true that the more complex knotted polygons have higher curvature?

Acknowledgments

This work is supported in part by NSF Grants #DMS-0920880 and #DMS-1016460 (Y Diao), and by NSF grant #DMS-1016420 (C. Ernst, A. Montemayor and U. Ziegler).

References

[1] Arsuaga J, Borgo B, Diao Y and Scharein R 2009, *The Growth of the Average Crossing Number of Equilateral polygons in Confinement*, J. Phys. A: Math. Theor **42** 465202.

[2] Diao Y, Dobay A, Kusner R, Millett K and Stasiak A 2003, *The Average Crossing Number of Equilateral Random Polygons*, J. Phys A: Math. Theor, **36**(46), 11561–74.

[3] Diao Y, Ernst C, Montemayor A and Ziegler U 2011, *Generating equilateral random polygons in confinement*, J. Phys. A: Math. Theor., **44** 405202.

[4] Diao Y, Ernst C, Montemayor A and Ziegler U 2012, *Generating equilateral random polygons in confinement II*, J. Phys. A: Math. Theor., **45** 275203.

[5] Diao Y, Ernst C, Montemayor A and Ziegler U 2012, *Generating equilateral random polygons in confinement III*, J. Phys. A: Math. Theor., **45** 465003.

[6] Dobay A, Dubochet J, Millett K, Sottas P-E and Stasiak A 2003, *Scaling behavior of random knots*, Proc Natl Acad Sci USA **100** (10) 5611–5.

[7] Doi M and Edwards S F, *The theory of polymer dynamics*, Oxford University Press, 1986.

[8] Flory P, *Principles of Polymer Chemistry*, Cornell University Press, 1953.

[9] De Gennes P G, *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca and London, 1979.

[10] Grosberg A Y 2008, *Total Curvature and Total Torsion of a Freely Jointed Circular Polymer with $n >> 1$ Segments*, Macromolecules **41** 4524–7.

[11] Jardine P J and Anderson D L 2006, *DNA packaging in double-stranded DNA phages,* The bacteriophages, Ed. Richard Calendar, Oxford University Press, 49–65.

[12] Klenin K V, Vologodskii A V, Anshelevich V V, Dykhne A M and Frank-Kamenetskii M D 1988, *Effect of excluded volume on topological properties of circular DNA*, J. Biomolec. Str. and Dyn. **5** 1173–85.

[13] Millett K 2000, *Monte Carlo Explorations of Polygonal Knot Spaces*, Knots in Hellas'98 (Delphi), Ser. Knots Everything (World Scientific) **24** 306–34.

[14] Portillo J, Arsuaga J, Diao Y, Scharein R and Vazquez M 2011, *On the Mean and Variance of the Writhe of Random Polygons*, J. Phys. A: Math. Theor, **44** 275004.

[15] Plunkett P et al 2007, *Total curvature and total torsion of knotted polymers*, Macromolecules **40** 3860–7.

[16] Rayleigh L 1919, *On the problems of random vibrations, and of random flights in one, two, or three dimensions*, Phil. Mag. S. 6. **37**(220) 321–47.

[17] Varela R, Hinson K, Arsuaga J and Diao Y 2009, *A Fast Ergodic Algorithm for Generating Ensembles of Equilateral Random Polygons*, J. Phys. A: Math. Theor. **42** 095204.

[18] Zirbel L and Millett K 2012, *Characteristics of shape and knotting in ideal rings*, J. Phys. A: Math. Theor. **45** 225001.