Central limit theorems for the winding and linking number

by Péter Simon

Supervisor: Viktor Harangi

May 2020

AN ESSAY PRESENTED TO CENTRAL EUROPEAN UNIVERSITY (CEU) IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF A MASTER OF SCIENCE IN MATHEMATICS AND ITS APPLICATIONS

CENTRAL EUROPEAN UNIVERSITY

Declaration

This work was carried out at Central European University in partial fulfilment of the requirements for a Master of Science in Mathematics and its Applications.

I hereby declare that except where due acknowledgement is made, this work has never been presented wholly or in part for the award of a degree at Central European University or any other university.

Student: Péter Simon
Supervisor: Viktor Harangi

Acknowledgements

First of all, I would like to wholeheartedly thank my supervisor Vikor Harangi, who was patient enough to lead me through this whole thesis. He gave me more help than a supervisee can ask for.

I am also thankful to Károly Böröczky, who was there for me throughout my whole academic life at CEU.

Moreover, I am grateful to my teachers and the department. They provided a truly friendly and flexible environment all along.

I am thankful and grateful to my fellow mentors and teachers, especially Lajos Pósa and Péter Juhász. There is not enough space here to even describe the degree they influenced me, and how much they taught me.

Last but not least, I am grateful to my friends. Without them, I would not be the person that I have become.

Abstract

We aim to study the winding number of certain random curves and the linking number of pairs of random curves. Our motivation comes from a work of Liu, Dehmamy, and Barabási [2], where the authors study the "tangledness" of graph embeddings. In order to test their hypothesis, they need to generate random embeddings of a given graph into $\mathbb{R}^3$ with many "self-linkings". Once the image of each vertex is fixed, they replace each edge with a polygonal path (broken line) whose intermediate points are IID points chosen from a uniform distribution. Then they measure the tangledness of the embedding by considering the linking numbers of pairs of disjoint cycles of the graph. They make several empirical observations for these random embeddings.

Our original goal was to study the distribution of the linking number of two random polygonal paths. Computer simulations suggested that, after proper normalization, it might converge to a normal distribution (as the number of IID intermediate points go to infinity). We started our investigations with a less complicated problem of similar flavor: the winding nuber of a random (closed) polygonal path on the plane. After expressing the winding number as the sum of a martingale difference sequence, we could rigorously prove that its distribution converges to a Gaussian by applying a Central Limit Theorem (CLT) for martingales. Then we turned to the linking number hoping to be able to prove a CLT for its distribution using similar tools. Now we believe that the limiting distribution is not quite normal. We do not have a rigorous proof at this point but our observations suggest that we may see an uncountable mixture of centered Gaussians in the limit.

Outline of the thesis

In Section 1 we formalize the winding number problem in a probabilistic language before proving a Central Limit Theorem for the problem using martingales in Section 2. Section 3 contains an inequality for anti-symmetric kernels that arose during the study of the previous sections. Then we investigate the role of the center point by analyzing the special case of the uniform distribution on the unit circle (Section 4). We present a possible generalization of the CLT result, along with an application, in Section 5. Finally, Section 6 is concerned with the linking number.

Contents

1 The winding number problem

As a warm-up, let us start with the following simple setup. We take independent uniform random points from the unit circle: $Z_0, Z_1, \ldots, Z_{N-1}$. Then we consider the closed polygonal path $Z_0 Z_1 \ldots Z_{N-1} Z_0$. It is easy to determine the winding number of a polygonal path: we simply need to add up the "signed angle increments" between neighboring points (Z_i and Z_{i+1}), viewed from the origin O for now, and divide this sum by 2π:

$$\big(\angle Z_0 O Z_1 + \angle Z_1 O Z_2 + \cdots + \angle Z_{N-2} O Z_{N-1} + \angle Z_{N-1} O Z_0\big)/2\pi,$$

where each signed angle $\angle Z_i O Z_{i+1}$ is meant to be in $(-\pi, \pi]$. See Figure 1 for an example. If we omit one of the summands, say, $\angle Z_{N-1} O Z_0$, then the remaining terms clearly form an IID sequence, chosen from the uniform distribution on the interval $(-\pi, \pi)$. So the Central Limit Theorem (CLT) tells us that the sum (and hence the winding number in question) normalized by $\sqrt{N}$ converges in distribution to a Gaussian.

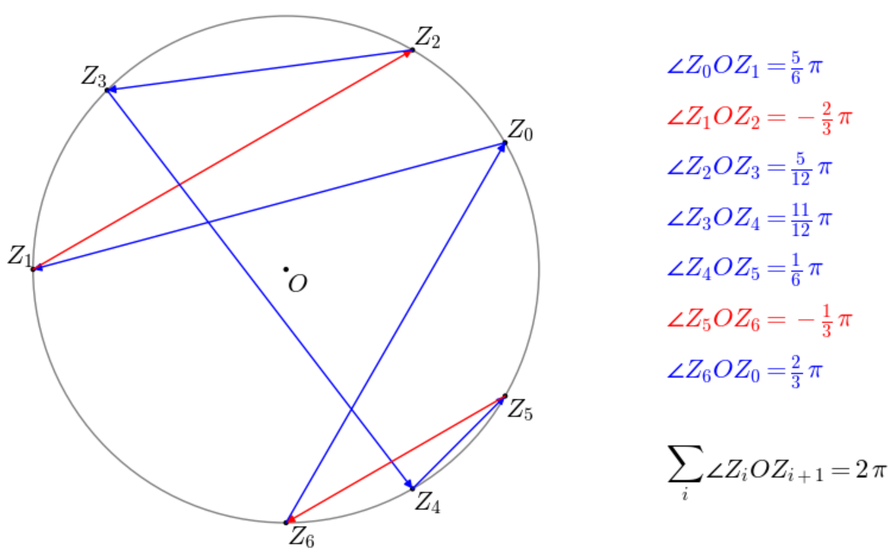


Figure 1: A closed polygonal path of seven intermediate points on the unit circle. The sum of signed angle increments around the origin O is $+2\pi$, and hence the winding number is $+1$. Blue/red arrows represent positive/negative angle increments.

The general setting

Now we consider the same problem but replace the uniform distribution of the unit circle with a probability measure ν on $\mathbb{R}^2$ and generate our IID sequence Z_i from this distribution. Then the consecutive signed angles $\angle Z_{i-1} O Z_i$ and $\angle Z_i O Z_{i+1}$ are not necessarily independent any more, and hence we cannot use classical CLT results that are only applicable to independent summands. Note that this issue already comes up if we work with the uniform distribution of the unit circle but use a center point different from the origin (but still inside the unit circle). (The question for general ν is meaningful for any probability measure as long as it holds with probability 1 that the segments $Z_{i-1}Z_i$ do not pass through the center point C. We will always assume this. It suffices to have $\nu(L) = 0$ for any line L passing through C.)

Let us fix some notations. Our center point is denoted by C. By $\arg_C(Z)$ we denote the argument of a point $Z \in \mathbb{R}^2$ w.r.t. the center C which is simply the signed angle between the positive x-axis and the ray CZ so that $\arg_C(Z) \in (-\pi, \pi]$. When the center coincides with the origin ($C = O$), we may omit the subscript and simply write $\arg(Z)$ for $\arg_O(Z)$. Note that $\arg(Z)$ is the argument of the corresponding complex number.

Given two arguments we need to consider their difference "modulo 2π". To be more precise, we define the function $\mathrm{mod}_{2\pi} \colon \mathbb{R} \to (-\pi, \pi]$ as follows. Let $\mathrm{mod}_{2\pi}(\theta) = \theta - 2k\pi$, where k is the unique integer such that $\theta - 2k\pi \in (-\pi, \pi]$. Then the argument difference, viewed from C, is

$$\Delta_C(Z_1, Z_2) := \mathrm{mod}_{2\pi}(\arg_C(Z_2) - \arg_C(Z_1)).$$

Again, we may omit the subscript when our center point is the origin. Note that Δ_C is essentially anti-symmetric: $\Delta_C(Z_2, Z_1) = -\Delta_C(Z_1, Z_2)$ unless the segment Z_1Z_2 contains C when both values are equal to π (but this will happen only with probability 0 under our conditions).

Now we can write the sum under consideration as follows (working with $C = O$ for the moment):

$$\sum_{i=0}^{N-2} \Delta(Z_i, Z_{i+1}). \tag{1}$$

If we divide this sum by 2π and round it to the closest integer, then we get the winding number of the closed polygonal path $Z_0Z_1 \dots Z_{N-1}Z_0$. (The rounding actually corrects the absence of the term $\Delta(Z_{N-1}, Z_0)$.)

We will prove that CLT holds for the winding number in this general setting as well.

Theorem 1.1. *Let ν be a Borel probability measure on $\mathbb{R}^2$ with the property that each line through the origin has zero measure. Take a random polygonal path with N independent intermediate points of distribution ν. Its winding number around the origin, divided by $\sqrt{N}$, converges in distribution to a Gaussian $N(0, \sigma^2)$. The variance σ^2 is zero if and only if ν is supported on an open half-plane not containing the origin.*

The summands corresponding to odd indices i are clearly independent. It follows that the standard CLT can be applied to this subsum. The same applies to the subsum corresponding to even indices i so we have normality for both subsums. There can be (a lot of) dependency between these two subsums, however, so the normality of the total sum does not follow directly from this observation. To prove Theorem 1.1 we need to use a dependent CLT result of some sort. As we will see in the next section, there is a well-developed theory of such results.

2 Central Limit Theorems

We have seen that the study of the winding number of a random polygonal path with IID intermediate points is equivalent to the study of a sum of the form (1) for an explicitly defined function Δ expressing the signed angle difference. At this point, it makes sense to move to greater generality and consider arbitrary functions in place of this specific Δ. In fact, we will work in a more abstract setting.

Theorem 2.1. *Given a probability measure space $(\Omega, \mathcal{A}, \nu)$ and a bounded measurable function $f\colon \Omega \times \Omega \to \mathbb{R}$, let us take an IID sequence $(Z_i)_{i\geq 0}$ with distribution ν. Then the normalized sum*

$$\left(\sum_{i=0}^{N-1} f(Z_i, Z_{i+1}) - N\mu_{f,\nu} \right) \Big/ \sqrt{N}$$

converges in distribution to $N(0, \sigma^2_{f,\nu})$ as $N \to \infty$, where

$$\mu_{f,\nu} = \mathbb{E}_{\omega_1,\omega_2} f(\omega_1, \omega_2) = \int f \,\mathrm{d}\nu^2$$

and

$$\begin{aligned}\sigma^2_{f,\nu} =& \mathbb{E}_{\omega_1,\omega_2} f^2(\omega_1, \omega_2) + 2\mathbb{E}_\omega \left(\mathbb{E}_{\omega_1} f(\omega_1, \omega) \mathbb{E}_{\omega_2} f(\omega, \omega_2) \right) = \\ & \int f^2 \,\mathrm{d}\nu^2 + 2 \int f(*, \omega) f(\omega, *) \,\mathrm{d}\nu(\omega).\end{aligned}$$

Here the $*$ notation means that the corresponding variable should be "integrated out" (w.r.t. ν) resulting in a function of one less variable.

In this formulation, we may consider functions of more than two variables. For an integer $k \geq 2$ let $f\colon \Omega^k \to \mathbb{R}$ be a bounded measurable function with zero expectation w.r.t. ν^k. Similarly, CLT holds for the normalized sum

$$\sum_{i=0}^{N-1} f(Z_i, Z_{i+1}, \ldots, Z_{i+k-1}) / \sqrt{N}.$$

See Section 5 for details.

Remark 2.2. As remarked for the case of winding number (that is, when $f = \Delta$), classical CLT can be applied for the "odd-index subsum" and for the "even-index subsum". These subsums can be far from independent, though. For example, let $h\colon \Omega \to \mathbb{R}$ be an arbitrary non-constant measurable function. If $f(x,y) = h(x)+h(y)$, then the correlation of the two subsums converges to 1, while the correlation goes to -1 for $f(x,y) = h(x) - h(y)$.

To prove Theorem 2.1 we need to turn our sum into a martingale. Next we explain how we can do this. After subtracting a constant from f, we may assume that f has zero expectation: $\mu_{f,\nu} = 0$.

Note that in the case when $f = \Delta$, this condition already holds, since we can deduct the equality $\mathbb{E}_{\omega_1,\omega_2}\Delta(\omega_1,\omega_2) = \mathbb{E}_{\omega_1,\omega_2}\Delta(\omega_2,\omega_1)$ from the independence, and we also have $\mathbb{E}_{\omega_1,\omega_2}\Delta(\omega_1,\omega_2) = -\mathbb{E}_{\omega_1,\omega_2}\Delta(\omega_2,\omega_1)$ due to the anti-symmetric property of Δ.

The k-th element of our filtration will be the σ-algebra $\mathcal{F}_k = \sigma(Z_0, Z_1, \dots, Z_k)$ generated by the the first $k+1$ elements of our IID sequence. This means that a random variable is $\mathcal{F}_k$-measurable if and only if it is a measurable function of $(Z_i)_{i\le k}$. Therefore

$$W_k := \sum_{i=0}^{k-1} f(Z_i, Z_{i+1})$$

is $\mathcal{F}_k$-measurable. (We say that the stochastic process $(W_k)_{k\ge1}$ is adapted to the filtration $(\mathcal{F}_k)_{k\ge1}$.) However, $(W_k)_{k\ge1}$ is not necessarily a martingale. For that, we would need $\mathbb{E}(W_{k+1} \mid \mathcal{F}_k) = W_k$ to hold for each k. For our filtration, given a random variable X (which is a measurable function of all $(Z_i)_{i\ge0}$) we can get the conditional expectation $\mathbb{E}(X \mid \mathcal{F}_k)$ by taking expectation of X in $(Z_i)_{i>k}$. Consequently, W_k is a martingale adapted to $\mathcal{F}_k$ if and only if $\mathbb{E}_{Z_{k+1}} f(Z_k, Z_{k+1}) = 0$, that is,

$$f(x,*) := \mathbb{E}_{Y\sim\nu} f(x,Y) = 0 \text{ for } \nu\text{-almost all } x.$$

In other words, $f(x,\cdot)$ must be centered for almost all x. However, we only know that $f(x,y)$ has zero expectation as a two-variable function. We can easily fix this by setting $\tilde{f}(x,y) = f(x,y) - f(x,*)$. Then $\widetilde{W}_k := \sum_{i=0}^{k-1} \tilde{f}(Z_i, Z_{i+1})$ is clearly a martingale. However, this sum is quite different from the original one that we want to study. We can fix this by considering the following function instead:

$$g(x,y) := f(x,y) - f(x,*) + f(y,*).$$

Then

$$X_k := \sum_{i=0}^{k-1} g(Z_i, Z_{i+1})$$

is still a martingale and is very close to the original sum W_k due to a telescopic sum that makes most of the extra terms cancel out. Indeed,

$$g(x,*) = \mathbb{E}_{Y\sim\nu}\big(f(x,Y) - f(x,*) + f(Y,*)\big) = f(x,*) - f(x,*) + 0 = 0$$

implies the martingale property, and

$$X_k = W_k - f(Z_0,*) + f(Z_k,*),$$

where the two extra terms will vanish in the limit because we will always normalize X_k and W_k by $\sqrt{k}$.

As we mentioned, there are generalizations of the classical CLT results for martingales and near-martingales. This theory was developed in the 1970s primarily by Brown, Dvoretzky, McLeish, Kłopotowski, Hall, Rebolledo, and Helland. There are many different (but very similar) sets of conditions under which CLT holds. Luckily, there is an excellent survey paper by Helland [1] that carefully explains how these results and conditions are related to each other. In our setting, we need to check two conditions to be able to conclude that $X_k/\sqrt{k}$ (and hence $W_k/\sqrt{k}$) converges in distribution to a Gaussian. At this point it is convenient to consider the corresponding martingale difference sequence $Y_k := X_k - X_{k-1} = g(Z_{k-1}, Z_k)$ for which we have $\mathbb{E}(Y_k \mid \mathcal{F}_{k-1}) = 0$. The following CLT result was proved by McLeish [3]. Also see [1, Theorem 2.5, conditions (b)] and [4, Theorem 2]. (Their setting, concerning martingale difference arrays, is more general.)

Theorem (McLeish, 1974 [3])**.** *Let $(Y_k)_{k\geq 1}$ be a martingale difference sequence adapted to a filtration $(\mathcal{F}_k)_{k\geq 1}$, and let $a_k \to \infty$ be a normalizing sequence. Suppose that, as $k \to \infty$, we have*

- $\max_{i\leq k} Y_i / a_k \to 0$ *in probability;*
- $\sum_{i\leq k} Y_i^2 / a_k^2$ *converges to σ^2 in probability for some real number $\sigma \geq 0$.*

Then $\sum_{i\leq k} Y_i / a_k$ converges in distribution to the Gaussian $N(0,\sigma^2)$ as $k \to \infty$.

We will apply the above theorem with the normalizing sequence $a_k = \sqrt{k}$. In our setting, the first condition is a trivial consequence of the fact that f (and hence g) is bounded. The second condition will follow using the weak law of large numbers.

Proof of Theorem 2.1. The only thing left to be done is verifying the second condition of McLeish's theorem.

For an arbitrary k, we have

$$\begin{aligned}
\mathbb{E}Y_k^2 =& \mathbb{E}_{\omega_1,\omega_2} g^2(\omega_1,\omega_2) \\
=& \mathbb{E}_{\omega_1,\omega_2}(f(\omega_1,\omega_2) - f(\omega_1,*) + f(\omega_2,*))^2 \\
=& \mathbb{E}_{\nu^2} f^2 + \mathbb{E}_{\omega_1} f(\omega_1,*)^2 + \mathbb{E}_{\omega_2} f(\omega_2,*)^2 - 2\mathbb{E}_{\omega_1,\omega_2}(f(\omega_1,\omega_2)f(\omega_1,*)) \\
&+ 2\mathbb{E}_{\omega_1,\omega_2}(f(\omega_1,\omega_2)f(\omega_2,*)) - 2\mathbb{E}_{\omega_1,\omega_2}(f(\omega_1,*)f(\omega_2,*)) \\
=& \mathbb{E}_{\nu^2} f^2 + 2\mathbb{E}_{\omega_1} f(\omega_1,*)^2 - 2\mathbb{E}_{\omega_1} f(\omega_1,*)^2 + 2\mathbb{E}_{\omega_2}(f(*,\omega_2)f(\omega_2,*)) - 2(\mathbb{E}_\nu f)^2 \\
=& \mathbb{E}_{\nu^2} f^2 + 2\mathbb{E}_{\omega}\big(\mathbb{E}_{\omega_1} f(\omega_1,\omega)\mathbb{E}_{\omega_2} f(\omega,\omega_2)\big) \\
=& \sigma_{f,\nu}^2
\end{aligned}$$

Since $a_k = \sqrt{k}$, we have

$$\sum_{i\le k} Y_i^2/a_k^2 = \frac{\lceil k/2\rceil}{k}\sum_{i=1}^{\lceil k/2\rceil} Y_{2i-1}^2/\lceil k/2\rceil + \frac{\lfloor k/2\rfloor}{k}\sum_{i=1}^{\lfloor k/2\rfloor} Y_{2i}^2/\lfloor k/2\rfloor$$

$\left(Y_{2i-1}^2\right)_{i=1}^{\lceil k/2\rceil}$ is an IID sequence with mean $\sigma_{f,\nu}^2$. It follows from the weak law of large numbers that $\sum_{i=1}^{\lceil k/2\rceil} Y_{2i-1}^2/\lceil k/2\rceil$ converges to $\sigma_{f,\nu}^2$ in probability. Since $\frac{\lceil k/2\rceil}{k} \to \frac{1}{2}$ as $k \to \infty$, we have

$$\frac{\lceil k/2\rceil}{k}\sum_{i=1}^{\lceil k/2\rceil} Y_{2i-1}^2/\lceil k/2\rceil \xrightarrow{p} \frac{\sigma_{f,\nu}^2}{2}$$

The same applies for the even indices, therefore

$$\sum_{i\le k} Y_i^2/k \xrightarrow{p} \frac{\sigma_{f,\nu}^2}{2} + \frac{\sigma_{f,\nu}^2}{2} = \sigma_{f,\nu}^2$$

□

Proof of Theorem 1.1. Take $f = \Delta$ and use Theorem 2.1. The occurrence of zero variance is covered in the following section.

□

3 Cauchy–Schwarz for anti-symmetric kernels

While studying the variance of the limiting normal distribution for the winding number described above, we stumbled upon the following inequality that we find interesting in its own right. It is certainly known in some form but we did not find it stated in the literature.

Let $f(x, y)$ be a *kernel*, that is, a two-variable function over some probability measure space. For a fixed x, one could apply the Cauchy–Schwarz inequality for $f(x, \cdot)$ and the constant 1 function to get the following (using our * notation from Theorem 2.1):

$$f(x,*) \le \sqrt{\int f^2(x,y)\,\mathrm{d}y}.$$

Squaring this and integrating w.r.t. x gives

$$\int f(x,*)^2\,\mathrm{d}x \le \iint f^2(x,y)\,\mathrm{d}x\,\mathrm{d}y. \tag{2}$$

Our inequality is essentially a strengthening of (2) for anti-symmetric kernels $f(x,y) = -f(y,x)$: we will show that (2) remains to be true even if we divide the right-hand side by 2.

3.1 General version

Proposition 3.1. *Given a probability measure space $(\Omega, \mathcal{A}, \nu)$ and a measurable function $f\colon \Omega\times\Omega \to \mathbb{R}$ with finite square integral, i.e. $f \in L^2(\nu^2)$, we have the following inequality:*

$$\mathbb{E}_{\nu^2} f^2 - \mathbb{E}_{\omega\sim\nu} f(\omega,*)^2 - \mathbb{E}_{\omega\sim\nu} f(*,\omega)^2 + \left(\mathbb{E}_{\nu^2} f\right)^2 \ge 0.$$

If f is anti-symmetric, i.e. $f(x,y) = -f(y,x)$, then the inequality simplifies to the following:

$$\mathbb{E}_{\omega\sim\nu} f(\omega,*)^2 \le \frac{1}{2}\mathbb{E}_{\nu^2} f^2. \tag{3}$$

As noted before, a simple Cauchy–Schwarz inequality would give (3) without the factor 1/2. In other words, by assuming that f is anti-symmetric, we get a bound twice as good as the standard one.

Before we give a concise proof of the inequality in Proposition 3.1, let us briefly sketch how we "discovered" it. As in the proposition, let $f\colon \Omega\times\Omega\to\mathbb{R}$ be a measurable function for some probability measure space $(\Omega,\mathcal{A},\nu)$. Take independent random variables X, Y, Z, each with distribution ν, and consider the variance of the sum $f(X,Y)+f(Y,Z)+f(Z,X)$ which must be non-negative:

$$0 \le \operatorname{var}\big(f(X,Y)+f(Y,Z)+f(Z,X)\big) = 3\operatorname{var}\big(f(X,Y)\big) + 6\operatorname{cov}\big(f(X,Y), f(Y,Z)\big).$$

When the expectation is zero ($\mathbb{E}_{\nu^2} f = 0$), this leads to the following inequality:

$$\mathbb{E}_{\nu^2} f^2 + 2\mathbb{E}_{\omega\sim\nu}(f(*,\omega)f(\omega,*)) \ge 0.$$

For anti-symmetric f we get the second inequality (3) in the proposition.

Proof of Proposition 3.1. Let X, $Y \sim \nu$. Then

$$\begin{aligned}
0 \leq &\operatorname{var}(f(X,Y) - f(X,*) - f(*,Y)) = \mathbb{E}_{\nu^2}(f(X,Y) - f(X,*) - f(*,Y))^2 - (\mathbb{E}_{\nu^2} f)^2 \\
=&\mathbb{E}_{\nu^2} f^2 + \mathbb{E}_\nu f(X,*)^2 + \mathbb{E}_\nu f(*,Y)^2 - 2\mathbb{E}_{\nu^2}(f(X,Y)f(X,*)) - 2\mathbb{E}_{\nu^2}(f(X,Y)f(*,Y)) \\
&+ 2\mathbb{E}_{\nu^2}(f(X,*)f(*,Y)) - (\mathbb{E}_{\nu^2} f)^2 \\
=&\mathbb{E}_{\nu^2} f^2 + \mathbb{E}_\nu f(X,*)^2 + \mathbb{E}_\nu f(*,Y)^2 - 2\mathbb{E}_\nu f(X,*)^2 - 2\mathbb{E}_\nu f(*,Y)^2 \\
&+ 2(\mathbb{E}_{\nu^2} f)^2 - (\mathbb{E}_{\nu^2} f)^2 \\
=&\mathbb{E}_{\nu^2} f^2 - \mathbb{E}_\nu f(X,*)^2 - \mathbb{E}_\nu f(*,Y)^2 + 2(\mathbb{E}_{\nu^2} f)^2 - (\mathbb{E}_{\nu^2} f)^2 \\
=&\mathbb{E}_{\nu^2} f^2 - \mathbb{E}_\nu f(X,*)^2 - \mathbb{E}_\nu f(*,Y)^2 + (\mathbb{E}_{\nu^2} f)^2
\end{aligned}$$

□

Next we show that equality occurs if and only if $f(x,y) = h(x)+l(y)$ almost everywhere for some functions $h, l \in L^2(\nu)$.

If we have equality then $f(X,Y)-f(X,*)-f(*,Y)$ is a constant ν^2 almost everywhere. This constant must be the expected value, thus $f(X,Y) - f(X,*) - f(*,Y) = -\mathbb{E}_{\nu^2} f$ a.e. Let $h(x) = f(x,*)$ and $l(y) = f(*,y) - \mathbb{E}_{\nu^2} f$. Then we have $f(x,y) = h(x) + l(y)$.

Now if we assume that $f(x,y) = h(x) + l(y)$ for some h and l. Then we have

$$\mathbb{E}_{\nu^2} f^2 = \mathbb{E}_{\nu^2}(h(x) + l(y))^2 = \mathbb{E}_\nu h^2 + 2\mathbb{E}_\nu h \mathbb{E}_\nu l + \mathbb{E}_\nu l^2$$

$$\mathbb{E}_\nu f(x,*)^2 = \mathbb{E}_\nu(h + \mathbb{E}_\nu l)^2 = \mathbb{E}_\nu h^2 + 2\mathbb{E}_\nu h \mathbb{E}_\nu l + (\mathbb{E}_\nu l)^2$$

Similarly,

$$\mathbb{E}_\nu f(*,y)^2 = (\mathbb{E}_\nu h)^2 + 2\mathbb{E}_\nu h \mathbb{E}_\nu l + \mathbb{E}_\nu l^2$$

$$(\mathbb{E}_{\nu^2} f)^2 = (\mathbb{E}_\nu h + \mathbb{E}_\nu l)^2 = (\mathbb{E}_\nu h)^2 + 2\mathbb{E}_\nu h \mathbb{E}_\nu l + (\mathbb{E}_\nu l)^2$$

From these we can finally deduct that

$$\mathbb{E}_{\nu^2} f^2 - \mathbb{E}_{\omega\sim\nu} f(\omega,*)^2 - \mathbb{E}_{\omega\sim\nu} f(*,\omega)^2 + \left(\mathbb{E}_{\nu^2} f\right)^2 = 0$$

Note that if f is anti-symmetric then $h(x) + l(y) = f(x,y) = -f(y,x) = -h(y) - l(x)$, hence $h(x) + l(x) = -h(y) - l(y)$ almost everywhere. Therefore $h(x) + l(x) = 0$ almost surely, and we conclude that $f(x,y) = h(x) - h(y)$.

Corollary 3.2. *The variance is zero in Theorem 1.1 if and only if ν is supported on an open half-plane not containing the origin.*

Proof. From the above discussion we get that the variance is zero if and only if $\Delta(x,y) = h(x) - h(y)$ holds ν^2 almost everywhere for some $h \in L^2(\nu)$. Then, since $-\pi < \Delta \le \pi$, we have $\operatorname{ess\,sup}_{\nu^2} \Delta = \operatorname{ess\,sup}_\nu h - \operatorname{ess\,inf}_\nu h \le \pi$. Hence the image of h is contained in an interval of length π. Choose y_0 such that $\nu(\{x | \Delta(x, y_0) \ne h(x) - h(y_0)\}) = 0$. Then as x varies, the image of $\Delta(x, y_0)$ also lands in an interval of length π, so ν is indeed supported on a half-plane. □

3.2 Matrix version

It is worth mentioning the special case when ν is the uniform measure on an n-element discrete set Ω, in which case we get an inequality for $n \times n$ matrices. Proving this would actually make a good maths competition problem for high-school students.

Corollary 3.3. *For arbitrary real numbers $(a_{i,j})_{1 \le i,j \le n}$ we have*

$$\sum_{i,j} a_{i,j}^2 - \frac{1}{n} \sum_i \left(\sum_j a_{i,j} \right)^2 - \frac{1}{n} \sum_j \left(\sum_i a_{i,j} \right)^2 + \frac{1}{n^2} \left(\sum_{i,j} a_{i,j} \right)^2 \ge 0.$$

If $a_{i,j} = -a_{j,i}$ holds for all i, j, then we get

$$\frac{1}{n} \sum_i \left(\sum_j a_{i,j} \right)^2 \le \frac{1}{2} \sum_{i,j} a_{i,j}^2.$$

Proof. By multiplying both sides by n^2, we get

$$\begin{aligned}
&n^2 \sum_{i,j} a_{i,j}^2 - n \sum_i \left(\sum_j a_{i,j} \right)^2 - n \sum_j \left(\sum_i a_{i,j} \right)^2 + \left(\sum_{i,j} a_{i,j} \right)^2 = \\
&(n-1)^2 \sum_{i,j} a_{i,j}^2 - n \sum_i \sum_{j \ne k} a_{i,j} a_{i,k} - n \sum_j \sum_{i \ne k} a_{i,j} a_{k,j} + \sum_{(i,j) \ne (p,r)} a_{i,j} a_{p,r} = \\
&(n-1)^2 \sum_{i,j} a_{i,j}^2 - (n-1) \sum_i \sum_{j \ne k} a_{i,j} a_{i,k} - (n-1) \sum_j \sum_{i \ne k} a_{i,j} a_{k,j} + \sum_{i \ne p, j \ne r} a_{i,j} a_{p,r} = \\
&\sum_{i<p, j<r} (a_{i,j} + a_{p,r} - a_{i,r} - a_{p,j})^2 \ge 0
\end{aligned}$$

We have equality if and only if $a_{i,j} + a_{p,r} = a_{i,r} + a_{p,j}$ for all i, j, p and r. We will prove that this is equivalent to saying, there exists $(b_i)_{i=1}^n$ and $(c_j)_{j=1}^n$ sequences such that $\forall i, j$, we have $a_{i,j} = b_i + c_j$.

If $a_{i,j} + a_{p,r} = a_{i,r} + a_{p,j}$, then $a_{1,1} + a_{p,r} = a_{1,r} + a_{p,1}$. Hence if we define $b_p = a_{p,1}$ and $c_r = a_{1,r} - a_{1,1}$, then we get $a_{p,r} = b_p + c_r$.

Now if $a_{i,j} = b_i + c_j$, for all i, j, for some sequences $(b_i)_{i=1}^n$ and $(c_j)_{j=1}^n$, then

$$\begin{aligned} a_{i,j} + a_{p,r} &= (b_i + c_j) + (b_p + c_r) \\ &= (b_i + c_r) + (b_p + c_j) \\ &= a_{i,r} + a_{p,j} \end{aligned}$$

□

4 Winding number for different center points

When studying the winding number of random polygonal paths, it is natural to look at the role of the center point. What can we say about the variance of the limiting Gaussian as the center point varies? In this section we will study this question in detail for the case when ν is the uniform distribution on the unit circle.

If our center point C is outside or on the circle, then the winding number is 0 with probability 1. So the variance is 0. If $C = 0$ (that is, it coincides with the center of the circle), then the angle increments are independent and uniformly distributed on $(-\pi, \pi)$ as we noted in Section 1. The variance is therefore $\mathrm{var}(U(-\pi,\pi)) = \frac{\pi^2}{3}$.

Next we will try to compute the variance around an arbitrary center point C inside the circle. Namely, we want the value of $\mathbb{E}_{\omega_1,\omega_2}\Delta_C^2(\omega_1,\omega_2) + 2\mathbb{E}_\omega\big(\mathbb{E}_{\omega_1}\Delta_C(\omega_1,\omega)\mathbb{E}_{\omega_2}\Delta_C(\omega,\omega_2)\big)$. Since Δ_C is anti-symmetric, the variance is $\mathbb{E}_{\omega_1,\omega_2}\Delta_C^2(\omega_1,\omega_2) - 2\mathbb{E}_{\omega_2}\big(\mathbb{E}_{\omega_1}\Delta_C(\omega_1,\omega_2)\big)^2$.

Because of rotational symmetry, the variance depends only on the distance r of C and O. So we assume that $C = (r, 0)$ for some real number $0 \le r < 1$. From Section 1 we know that

$$\Delta_C(\omega_1,\omega_2) = \mathrm{mod}_{2\pi}(\arg_C(\omega_2) - \arg_C(\omega_1)).$$

There is a measure-preserving bijection between the unit circle (with the uniform distribution) and the interval $(-\pi, \pi]$ with the Lebesgue measure normalized by 2π. Therefore, in what follows we will identify the unit circle with $(-\pi, \pi]$, that is, we will think of ω both as a signed angle and as the corresponding point on the circle. Thus, the integral of any function w.r.t. $\nu(\omega)$ is equal to its Lebesque integral from $-\pi$ to π, divided by 2π.

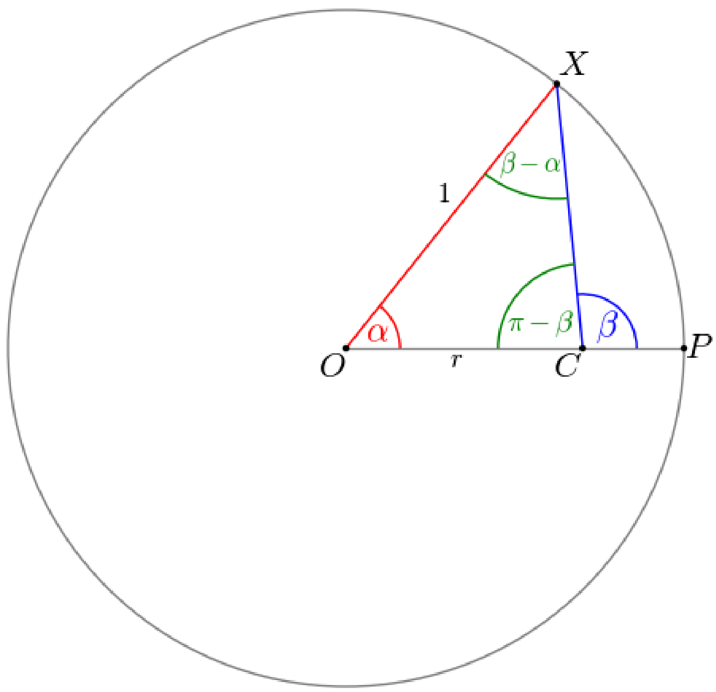
We will use the following change of variables to move between angles w.r.t. O and angles w.r.t. C. Let $x = \arg_C(\omega)$, and hence $\omega = \arg_C^{-1}(x)$. Then, setting $h(x) = \frac{\arg_C^{-1}(x)}{2\pi}$, we need to multiply with the density function $g(x) := h'(x)$ whenever we wish to integrate something as a function of x w.r.t. $\nu(\omega)$. Note that $\int_a^b g(x)\,\mathrm{d}x = h\big|_a^b$. In particalur, $\int_0^\pi g(x)\,\mathrm{d}x = \frac{1}{2}$.

Lemma 4.1. *$arg_C^{-1}(x) = x - \arcsin(r\sin(x))$*

Proof. Let $P=(1,0)$ and X be an arbitrary point on the unit circle. Furthermore let $\alpha=\angle COX$ and $\beta=\angle PCX$. We work with signed angles, ie. $-\pi<\alpha,\beta\leq\pi$. Now let us calculate α as a function of β. First suppose that $0<\alpha<\pi$. Since $\angle PCX=\beta$, we have $\angle XCO=\pi-\beta$ and then $\angle OXC=\beta-\alpha$. Using the law of sines on $OXC_\triangle$ we get $\frac{\sin(\beta-\alpha)}{r}=\frac{\sin(\pi-\beta)}{1}=\sin\beta$, and hence $\sin(\beta-\alpha)=r\sin(\beta)$.

Note that $\angle CPX=\frac{\pi-\alpha}{2}$, so $\beta<\pi-\frac{\pi-\alpha}{2}=\frac{\pi+\alpha}{2}$. Thus $\beta-\alpha<\frac{\pi-\alpha}{2}<\frac{\pi}{2}$.

This shows we can use the arcsin function without any problem on $\sin(\beta-\alpha)=r\sin(\beta)$ to get what we needed.



Finally let us recognise that $\beta-\arcsin(r\sin\beta)$ is an odd function so it works when β is negative as well. □

Let us define

$$f(x,y):=\Delta_C(\arg_C^{-1}(x),\arg_C^{-1}(y))=\mathrm{mod}_{2\pi}(y-x).$$

Now our task is equivalent to calculating $\mathbb{E}_{x,y}f^2(x,y)-2\mathbb{E}_y\big(\mathbb{E}_x f(x,y)\big)^2$. We always assume that $0\leq y\leq\pi$, and we just multiply the results by two. Note that in this case we have

$$f(x,y)=\begin{cases} y-x & \text{for} \quad y-\pi\leq x\leq\pi \\ y-x-2\pi & \text{for} \quad -\pi\leq x<y-\pi \end{cases}$$

Define

- $a = 4\int_0^\pi x^2 g(x)\,\mathrm{d}x$
- $b = -16\pi\int_0^\pi xg(x)\left(\frac{1}{2} + h(x-\pi)\right)\mathrm{d}x$
- $c = 8\pi^2\int_0^\pi g(x)\left(\frac{1}{2} + h(x-\pi)\right)\mathrm{d}x$

Lemma 4.2. $\mathbb{E}_{x,y} f^2(x,y) = a + b + c$

Proof. By above

$$\begin{aligned}
\mathbb{E}f^2 &= 2\int_0^\pi\int_{y-\pi}^\pi (y-x)^2 g(x)g(y)\,\mathrm{d}x\,\mathrm{d}y + 2\int_0^\pi\int_{-\pi}^{y-\pi}(y-x-2\pi)^2 g(x)g(y)\,\mathrm{d}x\,\mathrm{d}y\\
&= 2\int_0^\pi\int_0^\pi (y-x)^2 g(x)g(y)\,\mathrm{d}x\,\mathrm{d}y + 2\int_0^\pi\int_0^{\pi-y}(y+x)^2 g(x)g(y)\,\mathrm{d}x\,\mathrm{d}y\\
&\quad + 2\int_0^\pi\int_{\pi-y}^\pi (y+x-2\pi)^2 g(x)g(y)\,\mathrm{d}x\,\mathrm{d}y\\
&= 4\int_0^\pi\int_0^\pi (y^2+x^2) g(x)g(y)\,\mathrm{d}x\,\mathrm{d}y - 8\pi\int_0^\pi\int_{\pi-y}^\pi yg(x)g(y)\,\mathrm{d}x\,\mathrm{d}y\\
&\quad - 8\pi\int_0^\pi\int_{\pi-y}^\pi xg(x)g(y)\,\mathrm{d}x\,\mathrm{d}y + 8\pi^2\int_0^\pi\int_{\pi-y}^\pi g(x)g(y)\,\mathrm{d}x\,\mathrm{d}y\\
&= 8\int_0^\pi y^2 g(y)\,\mathrm{d}y\int_0^\pi g(x)\,\mathrm{d}x - 8\pi\int_0^\pi yg(y)\int_{\pi-y}^\pi g(x)\,\mathrm{d}x\,\mathrm{d}y\\
&\quad - 8\pi\int_0^\pi\int_{\pi-x}^\pi xg(x)g(y)\,\mathrm{d}y\,\mathrm{d}x + 8\pi^2\int_0^\pi g(y)\int_{\pi-y}^\pi g(x)\,\mathrm{d}x\,\mathrm{d}y\\
&= 4\int_0^\pi x^2 g(x)\,\mathrm{d}x - 8\pi\int_0^\pi yg(y)\left(\frac{1}{2} + h(y-\pi)\right)\mathrm{d}y\\
&\quad - 8\pi\int_0^\pi xg(x)\left(\frac{1}{2} + h(x-\pi)\right)\mathrm{d}x + 8\pi^2\int_0^\pi g(y)\left(\frac{1}{2} + h(y-\pi)\right)\mathrm{d}y\\
&= a + b + c
\end{aligned}$$

□

Lemma 4.3.

$$2\mathbb{E}_y\big(\mathbb{E}_x f(x,y)\big)^2 = d \coloneqq 4\int_0^\pi g(x)\left(2\pi\left(\frac{1}{2}h(x-\pi)\right) - x\right)^2\mathrm{d}x$$

.

Proof.

$$\begin{aligned}
2\mathbb{E}_y\big(\mathbb{E}_x f(x,y)\big)^2 &= 4\int_0^\pi g(y)\left(\int_{y-\pi}^{\pi}(y-x)g(x)\,\mathrm{d}x + \int_{-\pi}^{y-\pi}(y-x-2\pi)g(x)\,\mathrm{d}x\right)^2\mathrm{d}y \\
&= 4\int_0^\pi g(y)\left(y\int_{-\pi}^{\pi}g(x)\,\mathrm{d}x + \int_{\pi-y}^{\pi}2\pi g(x)\,\mathrm{d}x\right)^2\mathrm{d}y \\
&= 4\int_0^\pi g(y)\left(y - 2\pi\left(\frac{1}{2} + h(y-\pi)\right)\right)^2\mathrm{d}y \\
&= d
\end{aligned}$$

□

Now we are able to calculate the variance of the winding number of uniformly distributed polygonal path around any C inside the unit circle. It is simply $a+b+c-d$. Working around this sum, one may deduct a shorter form of the variance, namely

$$(2\pi)^2\int_0^\pi g(x)\Big(\big(1+2h(x-\pi)\big) - \big(1+2h(x-\pi)\big)^2\Big)\,\mathrm{d}x$$

Unfortunately, we could not find the primitive function for this last integral, and hence we do not have an explicit formula for the variance. However, this one-variable definite integral is now simple enough so that it can be computed numerically quickly and with great precision for any given value of r. We plotted the integral as a function of r using Code 1 in the Appendix, see Figure 2 for the result. (Note that to get the variance of the normalized winding number, we need to omit the factor $(2\pi)^2$ from the above integral, and this is how we plotted the result.)

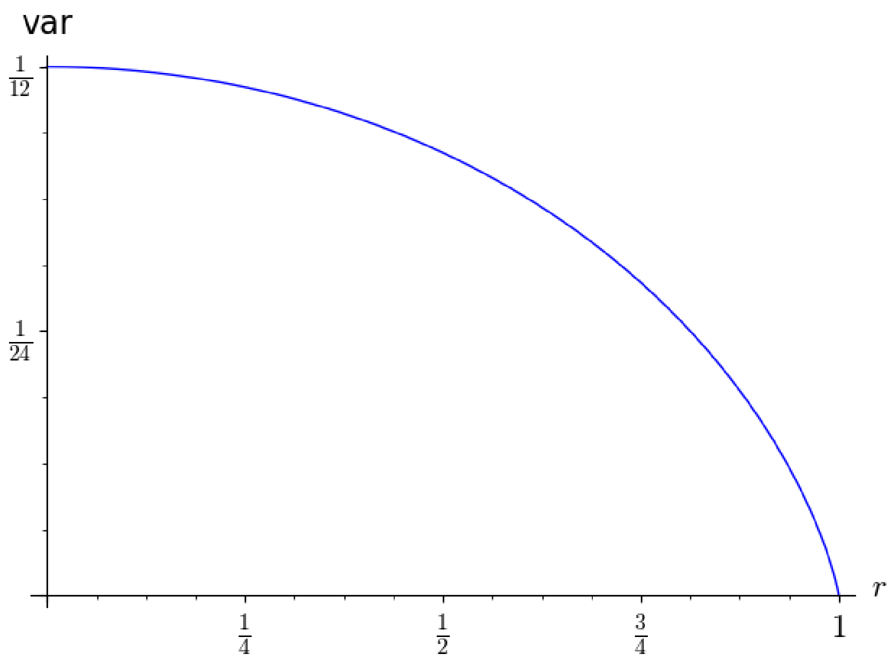


Figure 2: Variance of the normalized winding number around $C = (r, 0)$

4.1 An identity

While studying the variance of the winding number around center points inside the unit circle, we accidentally came across an identity between our integrals.

Proposition 4.4. $a + b + c + d = \frac{\pi^2}{3}$.

Note that $\pi^2/3$ is the variance around the origin.

Proof. First, let us recognise that

$$\begin{aligned}\frac{1}{2} + h(x - \pi) &= \frac{1}{2} + \frac{x - \pi - \arcsin(x - \pi)}{2\pi} \\ &= \frac{x + \arcsin(x)}{2\pi} \\ &= \frac{x}{\pi} - h(x)\end{aligned}$$

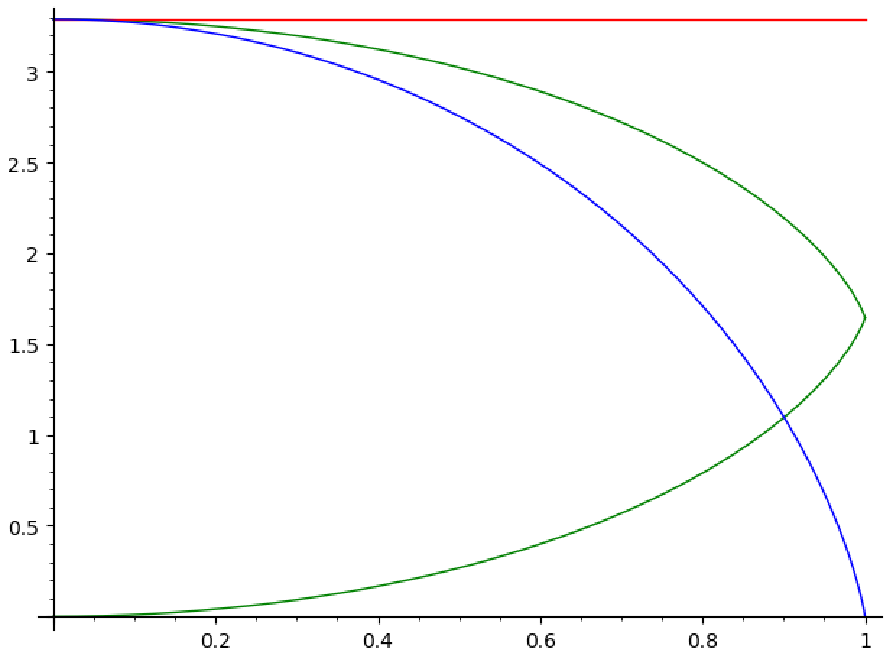


Figure 3: Two integrals (a+b+c and d) as functions of r (green curves). The sum is constant $\pi^2/3$ (red). The difference gives the variance in question (blue).

Hence

$$\begin{aligned}
a &= 4\int_0^\pi x^2 g(x)\,\mathrm{d}x \\
b &= -16\pi\int_0^\pi xg(x)\left(\frac{1}{2}+h(x-\pi)\right)\mathrm{d}x \\
&= -16\pi\int_0^\pi xg(x)\left(\frac{x}{\pi}-h(x)\right)\mathrm{d}x \\
&= -16\int_0^\pi x^2 g(x)\,\mathrm{d}x + 16\pi\int_0^\pi xg(x)h(x)\,\mathrm{d}x \\
c &= 8\pi^2\int_0^\pi g(x)\left(\frac{1}{2}+h(x-\pi)\right)\mathrm{d}x \\
&= 8\pi^2\int_0^\pi g(x)\left(\frac{x}{\pi}-h(x)\right)\mathrm{d}x \\
&= 8\pi\int_0^\pi xg(x)\,\mathrm{d}x - 8\pi^2\int_0^\pi g(x)h(x)\,\mathrm{d}x \\
d &= 4\int_0^\pi g(x)\left(2\pi\left(\frac{1}{2}+h(x-\pi)\right)-x\right)^2\mathrm{d} \\
&= 4\int_0^\pi g(x)\left(2\pi\left(\frac{x}{\pi}-h(x)\right)-x\right)^2\mathrm{d} \\
&= 4\int_0^\pi g(x)(x-2\pi h(x))^2 \\
&= 4\int_0^\pi x^2 g(x)\,\mathrm{d}x - 16\pi\int_0^\pi xg(x)h(x)\,\mathrm{d}x + 16\pi^2\int_0^\pi g(x)h^2(x)\,\mathrm{d}x
\end{aligned}$$

Therefore

$$\begin{aligned} a+b+c+d = &-8\int_0^\pi x^2 g(x)\,\mathrm{d}x + 8\pi\int_0^\pi xg(x)\,\mathrm{d}x - 8\pi^2\int_0^\pi g(x)h(x)\,\mathrm{d}x \\ &+16\pi^2\int_0^\pi g(x)h^2(x)\,\mathrm{d}x \end{aligned} \tag{4}$$

Now using integration by parts and chain rule, we have:

$$\begin{aligned} -8\int_0^\pi x^2 g(x)\,\mathrm{d}x &= -8x^2h(x)\Big|_0^\pi + 16\int_0^\pi xh(x)\,\mathrm{d}x \\ &= -4\pi^2 + \frac{8}{\pi}\int_0^\pi x^2\,\mathrm{d}x - \frac{8}{\pi}\int_0^\pi x\arcsin(r\sin(x))\,\mathrm{d}x \\ &= -\frac{4}{3}\pi^2 - \frac{8}{\pi}\int_0^\pi x\arcsin(r\sin(x))\,\mathrm{d}x \\ 8\pi\int_0^\pi xg(x)\,\mathrm{d}x &= 8\pi xh(x)\Big|_0^\pi - 8\pi\int_0^\pi h(x)\,\mathrm{d}x \\ &= 4\pi^2 - 4\int_0^\pi x\,\mathrm{d}x + 4\int_0^\pi \arcsin(r\sin(x))\,\mathrm{d}x \\ &= 2\pi^2 + 4\int_0^\pi \arcsin(r\sin(x))\,\mathrm{d}x \\ -8\pi^2\int_0^\pi g(x)h(x)\,\mathrm{d}x &= -4\pi^2h^2\Big|_0^\pi \\ &= -\pi^2 \\ 16\pi^2\int_0^\pi g(x)h^2(x)\,\mathrm{d}x &= 16\pi^2\frac{h^3}{3}\Big|_0^\pi \\ &= \frac{2}{3}\pi^2 \end{aligned}$$

Substituting these into (4), we get

$$a+b+c+d = \frac{\pi^2}{3} - \frac{8}{\pi}\int_0^\pi x\arcsin(r\sin(x))\,\mathrm{d}x + 4\int_0^\pi \arcsin(r\sin(x))\,\mathrm{d}x$$

Hence we need to prove that

$$\int_0^\pi x\arcsin(r\sin(x))\,\mathrm{d}x = \frac{\pi}{2}\int_0^\pi \arcsin(r\sin(x))\,\mathrm{d}x$$

We will prove this in a more general form.

□

Lemma 4.5. *Let $p : [a,b] \mapsto \mathbb{R}$ be integrable such that $p(a+x) = p(b-x)$ $\forall x \in [0, b-a]$. Then*

$$\int_a^b xp(x)\,\mathrm{d}x = \frac{(b+a)}{2}\int_a^b p(x)\,\mathrm{d}x$$

Proof. Let $s(y) = \int_a^y p(x)\,\mathrm{d}x$ for $y \in [a,b]$.

Then by symmetry we have $s(x) + s(b-x+a) = \int_a^b p(x)\,\mathrm{d}x$. By integration by parts

$$\begin{aligned}
\int_a^b xp(x)\,\mathrm{d}x &= xs(x)|_a^b - \int_a^b s(x)\,\mathrm{d}x \\
&= b\int_a^b p(x)\,\mathrm{d}x - \int_a^{\frac{b+a}{2}} s(x) + s(b-x+a)\,\mathrm{d}x \\
&= b\int_a^b p(x)\,\mathrm{d}x - \frac{b-a}{2}\int_a^b p(x)\,\mathrm{d}x \\
&= \frac{b+a}{2}\int_a^b p(x)\,\mathrm{d}x
\end{aligned}$$

□

5 CLT for functions of k consecutive variables

In this section we show how Theorem 2.1 generalizes to functions of more than two variables. The proof method is the same, only straightforward modifications are required that we will briefly sketch.

Theorem 5.1. *Given a probability measure space $(\Omega, \mathcal{A}, \nu)$ and a bounded measurable function $f\colon \Omega^k \to \mathbb{R}$, let us take an IID sequence $(Z_i)_{i\geq 0}$ with distribution ν. Then the normalized sum*

$$\left(\sum_{i=0}^{N-1} f(Z_i, Z_{i+1}, \ldots, Z_{i+k-1}) - N\mu_{f,\nu}\right) \Big/ \sqrt{N}$$

converges in distribution to $N(0, \sigma^2_{f,\nu})$ as $N \to \infty$, where

$$\mu_{f,\nu} = \mathbb{E}_{\nu^k} f = \int f\,\mathrm{d}\nu^k$$

and

$$\sigma^2_{f,\nu} = \operatorname{var}\left(f(X_1, \ldots, X_k) + \sum_{i=1}^{k-1}\left(f(X_{i+1}, \ldots, X_k, *, \ldots, *) - f(X_i, \ldots, X_{k-1}, *, \ldots, *)\right)\right).$$

(Here $X_1, \ldots, X_k$ denotes an IID sequence with distribution ν).

Proof. Once again, by subtracting a constant from f, we may assume that f has a zero expectation. Now let

$$g(x_1, \ldots, x_k) = \left(f(x_1, \ldots, x_k) + \sum_{i=1}^{k-1}\left(f(x_{i+1}, \ldots, x_k, *, \ldots, *) - f(x_i, \ldots, x_{k-1}, *, \ldots, *)\right)\right)$$

and

$$X_N := \sum_{i=0}^{N-1} g(Z_i, \dots, Z_{i+k-1})$$

.

Then $(X_i)_{i\geq 0}$ is a martingale adapted to the filtration $\mathcal{F}_i = \sigma(Z_0, \dots, Z_{k+i-2})$.

Note that due to the telescoping sum, the normalized X_i converges to the the same distribution as the normalized sum of f.

Considering the martingale difference sequence $Y_i := X_{i+1} - X_i$, and using McLeish's theorem (see Section 2 for more detail), we end up getting just what we wanted to prove. $\square$

5.1 Turning number

Next we present an application of Theorem 5.1 where the number of variables (of f) is greater than 2. The *turning number* of a closed directed curve is the number of full (signed) turns that we take when we walk it through.

When this closed curve is a polygonal path $Z_0Z_1 \dots Z_{N-1}Z_0$, then the turning number is simply the sum of signed angles between $\overrightarrow{Z_{i-1}Z_i}$ and $\overrightarrow{Z_iZ_{i+1}}$, divided by 2π. (The indices are meant in modulo N, i.e. $Z_N = Z_0$ and $Z_{N+1} = Z_1$.) So let $f(X,Y,Z)$ be the signed angle enclosed by $\overrightarrow{XY}$ and $\overrightarrow{YZ}$. Then the turning number of a polygonal path is

$$t(Z_0Z_1 \dots Z_{N-1}Z_0) = \frac{\sum_{i=0}^{N-3} f(Z_i, Z_{i+1}, Z_{i+2})}{2\pi} + O(1).$$

Now let ν be a measure on $\mathbb{R}^2$ and $(Z_i)_{i\geq 0}$ be an IID sequence with distribution ν. Then by Theorem 5.1, $\sum_{i=0}^{N-3} f(Z_i, Z_{i+1}, Z_{i+2}) \big/ \sqrt{N}$ converges in distribution to Gaussian with variance

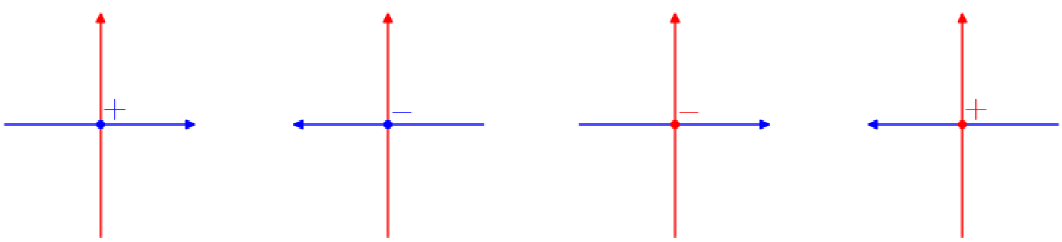
$$\begin{aligned}
&\overbrace{\int f(\omega_1,\omega_2,\omega_3) f(\omega_1,\omega_2,\omega_3) \,\mathrm{d}\nu(\omega_1)\,\mathrm{d}\nu(\omega_2)\,\mathrm{d}\nu(\omega_3)}^{\int f^2 \,\mathrm{d}\nu^3} + \\
2&\int f(*,\omega_1,\omega_2) f(\omega_1,\omega_2,*) \,\mathrm{d}\nu(\omega_1)\,\mathrm{d}\nu(\omega_2) + \\
2&\int f(*,*,\omega_1) f(\omega_1,*,*) \,\mathrm{d}\nu(\omega_1).
\end{aligned}$$

6 Linking number

Let us now turn to our original goal of studying the linking number of two closed polygonal paths. Given a red polygonal path and a blue one, both with a fixed orientation, their linking number can be intuitively defined as the number of times the red curve winds around the blue curve. There is a precise topological definition. For our purposes, it will suffice to know how one can compute the linking number from the so-called *linking diagram* (see Figure 4). Project the polygonal paths to a plane which is in "general position" with respect to the two curves. Then look at the crossings (i.e. points that both projected curves go through). Four types of crossings can be distinguished based on the following two properties.

- Which curve (red or blue) "goes over" the other at the crossing?
- What is the sign ($+$ or $-$) of the crossing? It is determined by the unit vectors representing the "direction" of the projected curves at the crossing. The sign of the crossing is $+$ (or $+1$) if we need to rotate the top directional vector (corresponding to the "overcrossing" curve) by some positive angle $0 < \alpha < \pi$ to get the bottom directional vector (corresponding to the "undercrossing" curve). If the angle of the rotation is negative ($-\pi < \alpha < 0$), then the sign of the crossing is defined to be $-$ (or -1).

So each crossing can be labelled by $+$, $-$, $+$, or $-$.



Lemma 6.1. *The sum of the red signs is always equal to the sum of the blue signs, and we define the linking number as this sum.*

Proof. It is enough to prove that for any two closed curves, we have $\#\{+,-\} = \#\{-,+\}$. Note that these numbers correspond to the numbers of crossings when we need to rotate the red vector by some positive and negative angle respectively to get the blue one.

Let $\gamma_b : [0,1] \mapsto \mathbb{R}^2$ be the blue curve. Then let $r = \inf\{l \big| \exists s < l, \gamma_b(l) = \gamma_b(s)\}$. Define $\gamma_b^1, \gamma_b^{(1)} : [0,1] \mapsto \mathbb{R}^2$ as $\gamma_b^1(x) = \gamma_b(s + (r-s)x)$, and $\gamma_b^{(1)} = \gamma_b \backslash \gamma_b^1$. Then repeat these steps with $\gamma_b^{(1)}$ to get γ_b^2, $\gamma_b^{(2)}$. Then repeat with $\gamma_b^{(2)}$, and so on.

Let $n = \inf\{k \big| |\gamma_b^{(k)}| = 0\}$. Then $\bigcup_{i=1}^n \gamma_b^i = \gamma_b$, and for all i, γ_b^i is a non-intersecting closed curve. Thus for all i, the red curve intersects γ_b^i an even number of times. On top of that, in half of the cases the red vector points "inward" of the curve, and it points "outward" in the other half cases. This means that for all i, when we consider the intersections on γ_b^i, it contributes the same number for $\#\{+,-\}$ and $\#\{-,+\}$, hence they are equal.

□

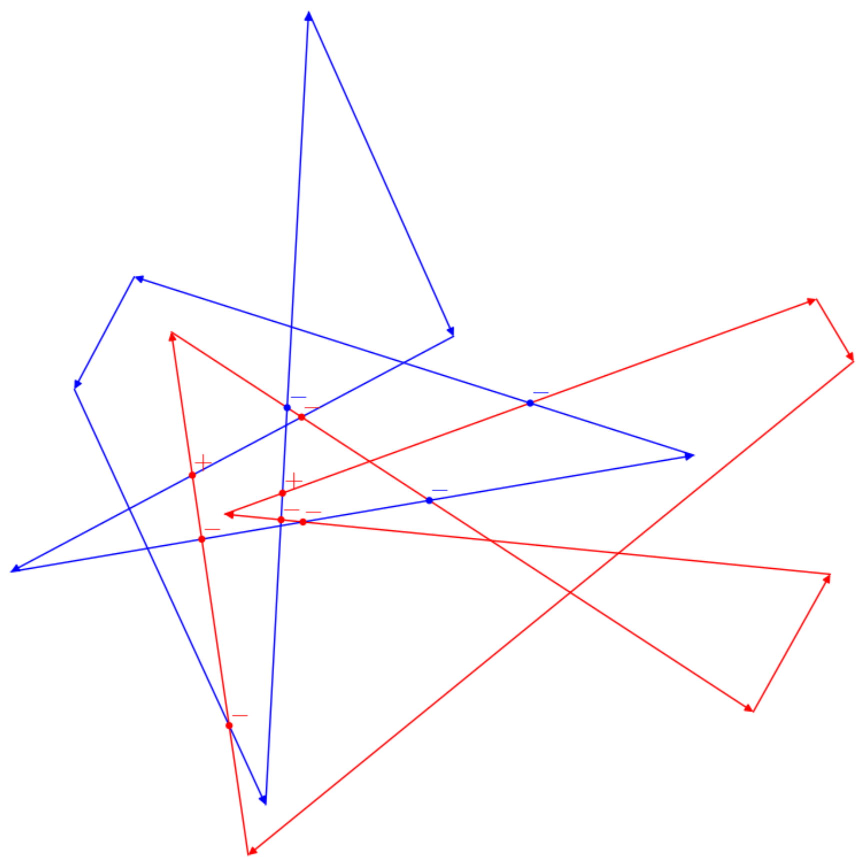


Figure 4: Linking diagram for two closed polygonal paths in $\mathbb{R}^3$, each with seven intermediate points: we take the 2D projection, then at each crossing we determine which curve is on top (indicated by the color of the intersection point), and assign a sign to the crossing based on the relative position of the overcrossing curve to the undercrossing curve. The sum of red signs and the sum of blue signs will always be the same, giving the linking number (-3 in this case).

6.1 Variance of the linking number

Now we can formulate our problem in a purely probabilistic language. For the sake of simplicity, assume that the red and the blue polygonal paths have the same number of IID intermediate points (N) and from the same distribution (ν) on $\mathbb{R}^3$. The red path will be denoted by $Z_0^{\mathrm{r}} Z_1^{\mathrm{r}} \dots Z_{N-1}^{\mathrm{r}} Z_0^{\mathrm{r}}$, while $Z_0^{\mathrm{b}} Z_1^{\mathrm{b}} \dots Z_{N-1}^{\mathrm{b}} Z_0^{\mathrm{b}}$ is the blue path. Indices are always meant modulo N: e.g. $Z_N^{\mathrm{r}} = Z_0^{\mathrm{r}}$ or $Z_{N+1}^{\mathrm{b}} = Z_1^{\mathrm{b}}$. Given a "red index" $0 \le i < N$ and a "blue index" $0 \le j < N$, we define the random variable $X_{i,j}$ to be $+1/-1$ if the red vector $Z_i^{\mathrm{r}} Z_{i+1}^{\mathrm{r}}$ "overcrosses" the blue vector $Z_j^{\mathrm{b}} Z_{j+1}^{\mathrm{b}}$ with a (red) $+/-$ sign, otherwise $X_{i,j}$ is 0. In other words, we only consider the red signs in our linking diagram. This way we get the linking number as an N^2-element sum: $\sum_{i,j} X_{i,j}$, where

$$X_{i,j} = f\big(Z_i^{\mathrm{r}}, Z_{i+1}^{\mathrm{r}}, Z_j^{\mathrm{b}}, Z_{j+1}^{\mathrm{b}}\big)$$

for some fixed $\{-1, 0, 1\}$-valued measurable function f. Note that f has the property that

$$f(x_0, x_1, y_0, y_1) = -f(x_1, x_0, y_0, y_1) \text{ and } f(x_0, x_1, y_0, y_1) = -f(x_0, x_1, y_1, y_0),$$

that is, f changes sign if we exchange the first two or the last two variables ("flipping" either the red vector, or the blue vector).

For the variance of the sum: we need to find the covariance of $X_{i,j}$ and $X_{i',j'}$. If $i' \neq i-1, i, i+1$ and $j' \neq j-1, j, j+1$, then they are independent and hence cov $= 0$. If $i' \neq i-1, i, i+1$, then one of the red vectors can be "flipped" and we would still have the same joint distribution for the endpoints but covariance gets multiplied with -1. Therefore the covariance must be 0. Similarly for the case $j' \neq j-1, j, j+1$.

The remaining cases are:

- $i = i'$ and $j = j'$. Then we get $a := \mathrm{cov}(X_{i,j}, X_{i,j}) = \mathbb{E}(X_{0,0}^2)$. When we consider the variance of $\sum_{i,j} X_{i,j}$ then a appears for all i, j once, hence altogether N^2 times.

- $|i - i'| + |j - j'| = 1$, then we have $b := \mathrm{cov}(X_{i,j}, X_{i',j'}) = \mathbb{E}(X_{0,0}X_{0,1})$. This one occurs $4N^2$ times.

- $|i - i'| = |j - j'| = 1$. In this case $c := \mathrm{cov}(X_{i,j}, X_{i',j'}) = \mathbb{E}(X_{0,0}X_{1,1})$, and this happens in $4N^2$ cases.

In conclusion, the variance of the linking number (expressed in terms of a, b, c) is

$$(a + 4b + 4c)N^2.$$

In what follows we will try to test our hypothesis (that the distribution of the linking number is close to a Gaussian with the above variance) by generating a random sample of linking numbers using a computer.

6.2 Simulations in the unit cube

We consider the following simple setup with the goal of generating random linking numbers. Let ν be the uniform measure of the unit cube $[0, 1]^3$. When creating a linking diagram, we will always project onto the xy plane (i.e. first two coordinates).

First we explain how $f(r_0, r_1, b_0, b_1)$ can be determined for four given points r_0, r_1, b_0, b_1.

Let $r_0 = (r_0^x, r_0^y, r_0^z)$, $r_1 = (r_1^x, r_1^y, r_1^z)$, $b_0 = (b_0^x, b_0^y, b_0^z)$ and $b_1 = (b_1^x, b_1^y, b_1^z)$. Project these points onto the usual plane, and let the image be r_0', r_1', b_0' and b_1' respectively. Now we would like to calculate the intersection of the lines $r_0'r_1'$ and $b_0'b_1'$. More precisely we are looking for the numbers R, B such that $R(r_0' - r_1') + r_1' = B(b_0' - b_1') + b_1'$. Considering only the first coordinates from this equation, we get $R(r_0^x - r_1^x) + r_1^x = B(b_0^x - b_1^x) + b_1^x$. Hence

$$R = B\frac{(b_0^x - b_1^x)}{(r_0^x - r_1^x)} + \frac{(b_1^x - r_1^x)}{(r_0^x - r_1^x)} \tag{5}$$

For the second coordinates, we have $R(r_0^y - r_1^y) + r_1^y = B(b_0^y - b_1^y) + b_1^y$. Substituting (5) into this, we get

$$\left(B\frac{(b_0^x - b_1^x)}{(r_0^x - r_1^x)} + \frac{(b_1^x - r_1^x)}{(r_0^x - r_1^x)}\right)(r_0^y - r_1^y) = B(b_0^y - b_1^y) + (b_1^y - r_1^y)$$

Thus

$$B = \frac{(b_1^y - r_1^y)(r_0^x - r_1^x) - (b_1^x - r_1^x)(r_0^y - r_1^y)}{(b_0^x - b_1^x)(r_0^y - r_1^y) - (b_0^y - b_1^y)(r_0^x - r_1^x)}$$

Similarly

$$R = -\frac{(r_1^y - b_1^y)(b_0^x - b_1^x) - (r_1^x - b_1^x)(b_0^y - b_1^y)}{(b_0^x - b_1^x)(r_0^y - r_1^y) - (b_0^y - b_1^y)(r_0^x - r_1^x)}$$

Using R and B, we can easily check whether the line segments $[r_0', r_1']$, $[b_0', b_1']$ intersect each other: they do if and only if $R, B \in [0, 1]$.

When $R, B \in [0, 1]$, then we need to check that at the intersection, we have the red segment "on top". We can calculate the heights of each segment from the intersection. For the red one it is $Rr_0^z + (1 - R)r_1^z$ and for the blue one it is $Bb_0^z + (1 - B)b_1^z$.

If $R \notin [0, 1]$ or $B \notin [0, 1]$ or $Rr_0^z + (1 - R)r_1^z < Bb_0^z + (1 - B)b_1^z$, then $f(r_0, r_1, b_0, b_1) = 0$.

Otherwise, we have to check in which direction (positive or negative) we should rotate $\overrightarrow{r_0'r_1'}$ to get $\overrightarrow{b_0'b_1'}$. We can do this by considering the third coordinate of the cross product $\overrightarrow{r_0'r_1'} \times \overrightarrow{b_0'b_1'}$. Thus we have

$$f(r_0, r_1, b_0, b_1) = \operatorname{sign}((r_0^x - r_1^x)(b_0^y - b_1^y) - (r_0^y - r_1^y)(b_0^x - b_1^x))$$

Using this formula, we can run various computer simulations. We used Code 4 in the appendix to get approximate values for a, b, c in the unit cube setup:

$$a \approx +0.1157$$

$$b \approx -0.0366$$

$$c \approx +0.0119$$

For the variance we get

$$(a + 4b + 4c)N^2 \approx 0.0169N^2 = (0.13N)^2.$$

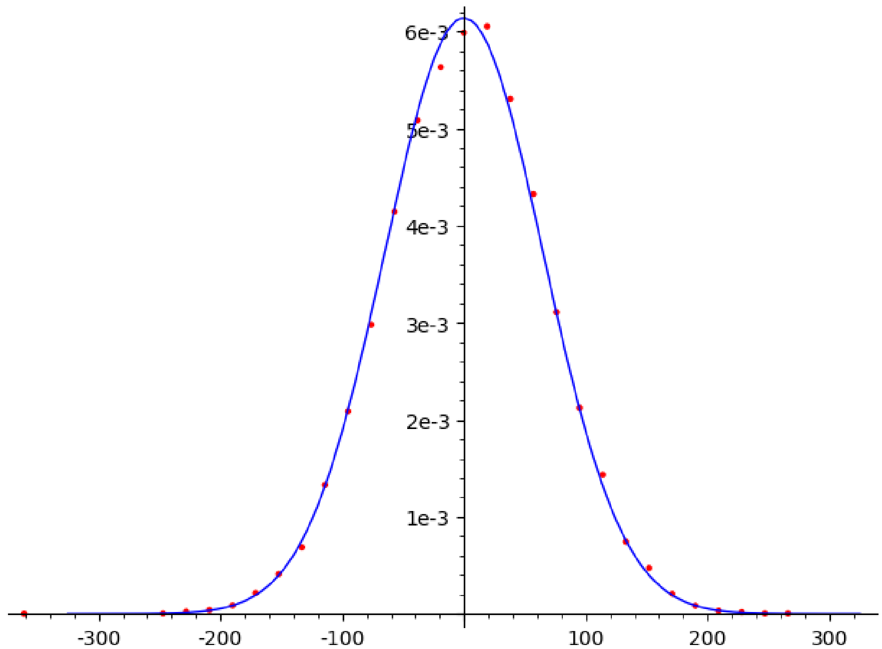


Figure 5: Histogram for a sample of random linking numbers, compared to normal distribution

Next we compared the empirical distribution of a random sample for the linking number to a normal distribution. We generated 20000 pairs of closed polygonal paths, each with length $N = 500$. Then we computed the linking number for each pair. The resulting sample of 20000 numbers is compared to the normal distribution of variance $(0.13N)^2$ in Figure 5. The code that we used is attached to the the appendix as Code 6.

6.3 True limit: a mixture of Gaussians?

It seems from Figure 5 that the linking number is very close to the normal distribution as we had originally expected. After further investigation, however, we now think that the limiting distribution is not quite normal. Instead, we expect it to be a mixture of Gaussians. To explain why we think that let us look at a similar problem with less dependency.

Problem 6.2. Let $g\colon \Omega \times \Omega \to \mathbb{R}$ be a function of two variables. Suppose that we have two IID sequences $Z_0^{\mathrm{r}}, Z_1^{\mathrm{r}}, \ldots$ and $Z_0^{\mathrm{b}}, Z_1^{\mathrm{b}}, \ldots$ from the same distribution ν on Ω. For the random variables $X_{i,j} = g(Z_i^{\mathrm{r}}, Z_j^{\mathrm{b}})$ we consider their sum

$$S_N := \sum_{i,j<N} X_{i,j} = \sum_{i,j<N} g(Z_i^{\mathrm{r}}, Z_j^{\mathrm{b}}).$$

What can we say about the distribution of S_N in the limit as $N \to \infty$?

Remark 6.3. Depending on g and ν, the nature of Problem 6.2 can be very different. For example, given independent $Z^{\mathrm{r}}, \hat{Z}^{\mathrm{r}}, Z^{\mathrm{b}}, \hat{Z}^{\mathrm{b}}$, each with distribution ν, it makes a big difference whether the covariances

$$\operatorname{cov}\left(g(Z^{\mathrm{r}}, Z^{\mathrm{b}}), g(\hat{Z}^{\mathrm{r}}, Z^{\mathrm{b}})\right) \text{ and } \operatorname{cov}\left(g(Z^{\mathrm{r}}, Z^{\mathrm{b}}), g(Z^{\mathrm{r}}, \hat{Z}^{\mathrm{b}})\right)$$

are zero or not. It is easy to see that these covariances are zero if and only if

$$g(*, x) = g(x, *) = 0 \text{ for almost all } x,$$

in which case S_N has variance of order N^2 (and not N^3 as in general), and hence we should consider the normalized sum S_N/N.

Example 6.4. Let $\Omega = [0,1]^3 \times [0,1]^3$ and $g((x_0, y_0); (x_1, y_1)) = f(x_0, y_0, x_1, y_1)$ for the function f defined in Section 6.1 for the linking number problem. So for N independent red vectors and N independent blue vectors, S_N is simply the sum of the signs of red overcrossings for all pairs of one red vector and one blue vector. (That is, the vectors here are not from closed paths of length N but each vector is chosen independently.) This problem captures the essence of the original linking number problem without having to deal with the dependencies between consecutive vectors. In this case the covariances mentioned in the remark are zero so the proper normalization will be S_N/N.

Now we make a few general comments regarding Problem 6.2. First let us think of the first sequence $Z_0^{\mathrm{r}}, Z_1^{\mathrm{r}}, \ldots, Z_{N-1}^{\mathrm{r}}$ as a fixed deterministic sequence, and let Z be a random variable with distribution ν. Then

$$Y = \sum_{i<N} g(Z_i^{\mathrm{r}}, Z)$$

has some distribution $\kappa_N = \kappa(Z_0^{\mathrm{r}}, \ldots, Z_{N-1}^{\mathrm{r}})$ depending on $(Z_i^{\mathrm{r}})_{i<N}$. Now for any given j, the sum

$$Y_j = \sum_{i<N} X_{i,j}$$

has the same distribution as Y (still thinking of $Z_0^{\mathrm{r}}, Z_1^{\mathrm{r}}, \ldots, Z_{N-1}^{\mathrm{r}}$ as a fixed deterministic sequence). So $S_N = \sum_{j<N} Y_j$ is actually the sum of N independent copies of Y, that is,

the sum of N IID samples from $\kappa_N = \kappa(Z_0^{\mathrm{r}}, \ldots, Z_{N-1}^{\mathrm{r}})$. If we think of the first sequence as random, then we need to consider a random κ_N in the above discussion. Therefore we first need to generate a probability distribution κ_N at random (by randomizing $Z_0^{\mathrm{r}}, \ldots, Z_{N-1}^{\mathrm{r}}$), then sampling N independent copies from κ_N and add them up. Standard CLT "suggests" that what we get is "close" to a normal distribution $N(0, \sigma^2)$ where $\sigma = \sqrt{N}\sigma(\kappa_N)$. Note that $\sigma(\kappa_N)$ is random a number here.

This approach would involve two steps. The first step is to understand the behavior of $\sigma(\kappa_N)$. The second step is to prove that the limiting distribution of S_N/N is the corresponding mixture of Gaussians. Note that we have to take the limit simultaneously in the index of κ_N and in the number N of independent samples of κ_N.

For the linking number, we still expect each Y_j to be close to a Gaussian but with a nonzero covariance between consecutive ones which further complicates the matter. Nevertheless, we think that S_N will converge to a mixture of centered Gaussians. The mixture distribution is quite close to the Gaussian distribution (with deterministic variance) that we previously considered. So close, in fact, that we would need a gigantic sample of random linking numbers to be able to distinguish between the two distributions, which explains why simulations suggested that CLT might hold.

7 Appendix: codes

We include the SageMath codes of our computations, simulations, and figures for the thesis. Most of these codes can be run in a “Sagecell”: go to `sagecell.sagemath.org` and copy-paste the code, then hit “Evaluate”. Some codes require longer running times and need to be run on a sage environment installed on a computer. (When copy-pasting, some PDF viewers may not include the spaces at the start of each line. So make sure to re-enter those spaces in the Sagecell as these indentations are vital for Python (and hence Sage) codes.)

7.1 Winding number

Code 1: Plotting the variance of the winding number (Figure 2)

```
def variance(r):
    h_inv(x)=(x-arcsin(r*sin(x)))/(2*pi)
    g=diff(h_inv(x),x)
    return numerical_integral(g*( (1+2*h_inv(x-pi))-(1+2*h_inv(x-pi))^2 ),0,pi)[0]
fig=plot(lambda r: variance(r),(r,0,1))
fig.show(ticks=[1/4,1/24],tick_formatter=[1,1/12],axes_labels=["$r$","var"],fontsize=16,axes_labels_size=1)
```

Code 2: Plotting “parts” of the variance (Figure 3)

```
def variance_parts(r):
    h_inv(x)=(x-arcsin(r*sin(x)))/(2*pi)
    g=diff(h_inv(x),x)
    # E f^2
    a=4*numerical_integral(x^2*g,0,pi)[0]
    b=-16*pi.n()*numerical_integral(x*g*(1/2+h_inv(x-pi)),0,pi)[0]
    c=8*pi.n()^2*numerical_integral(g*(1/2+h_inv(x-pi)),0,pi)[0]
    # 2*E_w f(w,*)^2
    d=4*numerical_integral(g*(2*pi*(1/2+h_inv(x-pi))-x)^2,0,pi)[0]
    return [a+b+c,d,a+b+c+d,a+b+c-d]

fig0=plot(lambda r: variance_parts(r)[0],(r,0,1),color='green')
fig1=plot(lambda r: variance_parts(r)[1],(r,0,1),color='green')
fig2=plot(lambda r: variance_parts(r)[2],(r,0,1),color='red')
fig3=plot(lambda r: variance_parts(r)[3],(r,0,1),color='blue')
show(fig0+fig1+fig2+fig3)
```

7.2 Linking number

Code 3: Crossing probabilities

```
def crossing(Rsgmt,Bsgmt):
    [[RstX,RstY,RstZ],[RendX,RendY,RendZ]]=Rsgmt
    [[BstX,BstY,BstZ],[BendX,BendY,BendZ]]=Bsgmt
    denom=(BendY-BstY)*(RendX-RstX)-(BendX-BstX)*(RendY-RstY)
    Rw=-( (BendY-RendY)*BstX-BendX*(BstY-RendY)-(BendY-BstY)*RendX )/denom
    Bw= ( (BendY-RstY)*RendX-BendX*(RendY-RstY)-(BendY-RendY)*RstX )/denom
    if Rw<0 or Rw>1 or Bw<0 or Bw>1: return 0
    if Rw*RstZ+(1-Rw)*RendZ > Bw*BstZ+(1-Bw)*BendZ: return 1
    return -1

def random_point(): return [random(),random(),random()]
def random_segment(): return [random_point(),random_point()]

@interact
def _(reps=10^5):
        count=[0,0,0]
        for _ in range(reps): count[1+crossing(random_segment(),random_segment())] += 1
        print([n(count[i]/reps,digits=6) for i in range(3)])
```

Code 4: Covariance for overlapping pairs

```
import numpy

def redcrossing_sign(Rsgmt,Bsgmt):
    [[RstX,RstY,RstZ],[RendX,RendY,RendZ]]=Rsgmt
    [[BstX,BstY,BstZ],[BendX,BendY,BendZ]]=Bsgmt
    denom=(BendY-BstY)*(RendX-RstX)-(BendX-BstX)*(RendY-RstY)
    Rw=-( (BendY-RendY)*BstX-BendX*(BstY-RendY)-(BendY-BstY)*RendX )/denom
    Bw= ( (BendY-RstY)*RendX-BendX*(RendY-RstY)-(BendY-RendY)*RstX )/denom
    if Rw<0 or Rw>1 or Bw<0 or Bw>1: return 0
    if Rw*RstZ+(1-Rw)*RendZ > Bw*BstZ+(1-Bw)*BendZ: return sgn(denom)
    return 0

def random_point(): return [random(),random(),random()]
def random_path(nr): return [random_point() for _ in range(nr)]

signs = numpy.full((3,3),0)
def sign_products():
    Rpath=random_path(4)
    Bpath=random_path(4)
    #signs=numpy.array([[0,0,0],[0,0,0],[0,0,0]])
    for Ri in range(3):
        for Bi in range(3):
            signs[Ri][Bi]=redcrossing_sign([Rpath[Ri],Rpath[Ri+1]],[Bpath[Bi],Bpath[Bi+1]])
    return signs[1][1]*signs

reps=10^5;
total = numpy.full((3,3),0) #total=numpy.array([[0,0,0],[0,0,0],[0,0,0]])
for _ in range(reps): total+=sign_products()
print(total/reps)
vr=numpy.sum(total)/reps
print("sigma^2: " + str(vr))
print("sigma: " + str(sqrt(vr)))
```

Code 5: Linking diagram (Figure 4)

```
def crossing_full_info(Rsgmt,Bsgmt):
    [[RstX,RstY,RstZ],[RendX,RendY,RendZ]]=Rsgmt
    [[BstX,BstY,BstZ],[BendX,BendY,BendZ]]=Bsgmt
    denom=(BendY-BstY)*(RendX-RstX)-(BendX-BstX)*(RendY-RstY)
    Rw=-( (BendY-RendY)*BstX-BendX*(BstY-RendY)-(BendY-BstY)*RendX )/denom
    Bw= ( (BendY-RstY)*RendX-BendX*(RendY-RstY)-(BendY-RendY)*RstX )/denom
    ptX=Rw*RstX+(1-Rw)*RendX
    ptY=Rw*RstY+(1-Rw)*RendY
    pos = -1;
    if Rw<0 or Rw>1 or Bw<0 or Bw>1: pos=0
    elif Rw*RstZ+(1-Rw)*RendZ > Bw*BstZ+(1-Bw)*BendZ: pos=1
    return [pos,sgn(denom),ptX,ptY]

def int_point(info):
    pt=(info[2],info[3])
    lbl="$+$"
    if (info[0]*info[1]==-1): lbl="$-$"
    if (info[0]==1): return point(pt,color='red',pointsize=20,zorder=10)+text(lbl, pt, horizontal_alignment=
         'left', vertical_alignment='bottom',color='red',fontweight='bold',fontsize=16,zorder=20)
    if (info[0]==-1): return point(pt,color='blue',pointsize=20,zorder=10)+text(lbl, pt,
         horizontal_alignment='left', vertical_alignment='bottom',color='blue',fontweight='bold',fontsize=16,
         zorder=20)
    return pt.plot(point={'color':'black'})
#red or blue labels may be removed if it is too "crowded"

def random_point(): return [random(),random(),random()]
def random_path(nr): return [random_point() for _ in range(nr)]
def random_closed_path(nr): path=random_path(nr); path.append(path[0]); return path

def proj2d(path): return [pt[:2] for pt in path]

def draw_linking_figure(Rn,Bn,min_lnk_nr=0):
    Rpath=random_closed_path(Rn); Rpath2d=proj2d(Rpath)
    Bpath=random_closed_path(Bn); Bpath2d=proj2d(Bpath)
    #fig=line(Rpath2d,color='red')+line(Bpath2d,color='blue')
    fig=Graphics()
    for i in range(Rn):
        fig+=arrow(Rpath2d[i],Rpath2d[i+1],color='red',width=1,arrowsize=2)
    for i in range(Bn):
        fig+=arrow(Bpath2d[i],Bpath2d[i+1],color='blue',width=1,arrowsize=2)
    lnk_nr=0
    for Ri in range(Rn):
        for Bi in range(Bn):
            info=crossing_full_info([Rpath[Ri],Rpath[Ri+1]],[Bpath[Bi],Bpath[Bi+1]])
            if (info[0]!=0): fig+=int_point(info)
            if (info[0]==1): lnk_nr+=info[1]
    if abs(lnk_nr)>=min_lnk_nr:
        print("Linking number: " + str(lnk_nr))
        fig.show(figsize=[8,8],axes=False)
        return True
    return False

@interact
def _(nr_of_red_points=7,nr_of_blue_points=7,min_lnk_nr=[0..4],auto_update=False):
        while not draw_linking_figure(nr_of_red_points,nr_of_blue_points,min_lnk_nr): 0
```

Code 6: Comparing linking number sample to Gaussian (Figure 5)

```
#VERY LONG running time

def redcrossing_sign(Rsgmt,Bsgmt):
    [[RstX,RstY,RstZ],[RendX,RendY,RendZ]]=Rsgmt
    [[BstX,BstY,BstZ],[BendX,BendY,BendZ]]=Bsgmt
    denom=(BendY-BstY)*(RendX-RstX)-(BendX-BstX)*(RendY-RstY)
    Rw=-( (BendY-RendY)*BstX-BendX*(BstY-RendY)-(BendY-BstY)*RendX )/denom
    Bw= ( (BendY-RstY)*RendX-BendX*(RendY-RstY)-(BendY-RendY)*RstX )/denom
    if Rw<0 or Rw>1 or Bw<0 or Bw>1: return 0
    if Rw*RstZ+(1-Rw)*RendZ > Bw*BstZ+(1-Bw)*BendZ: return sgn(denom)
    return 0

def random_point(): return [random(),random(),random()]
def random_path(nr): return [random_point() for _ in range(nr)]
def random_closed_path(nr): path=random_path(nr); path.append(path[0]); return path

def random_linking_number(Rn,Bn):
    Rpath=random_closed_path(Rn)
    Bpath=random_closed_path(Bn)
    link_nr=0
    for Ri in range(Rn):
        for Bi in range(Bn):
            link_nr+=redcrossing_sign([Rpath[Ri],Rpath[Ri+1]],[Bpath[Bi],Bpath[Bi+1]])
    return link_nr

def compare_to_gaussian(smpl,sigma,odd_len=9):
    total_nr = len(smpl)
    smpl_rounded=[odd_len*round(val/odd_len) for val in smpl]
    hash = {}
    for val in smpl_rounded:
        if(val in hash): hash[val]+=1
        else: hash[val] = 1
    for val in hash.keys(): hash[val] /= (odd_len*total_nr)
    fig1=list_plot(list(hash.items()),color='red')
    fig2=RealDistribution('gaussian', sigma).plot(x,-5*sigma,5*sigma,color='blue')
    show(fig1+fig2)

smpl=[random_linking_number(500,500) for _ in range(20000)]
compare_to_gaussian(smpl,0.13*500,19)
```

References

[1] Inge S. Helland. Central limit theorems for martingales with discrete or continuous time. *Scandinavian Journal of Statistics*, 9(2):79–94, 1982.

[2] Yanchen Liu, Nima Dehmamy, and Albert-László Barabási. Topology of tangledness of network embeddings. In *APS March Meeting Abstracts*, volume 2019 of *APS Meeting Abstracts*, page B56.010, January 2019.

[3] D. L. McLeish. Dependent central limit theorems and invariance principles. *The Annals of Probability*, 2(4):620–628, 1974.

[4] Sunder Sethuraman. A martingale central limit theorem. PDF accessible on author's website: `https://www.math.arizona.edu/ sethuram/notes/wi_mart1.pdf`.