Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2011, Vol. 47, No. 1, 130–147
DOI: 10.1214/10-AIHP365

ANNALES
DE L'INSTITUT
HENRI
POINCARÉ
PROBABILITÉS
ET STATISTIQUES
www.imstat.org/aihp

Windings of planar random walks and averaged Dehn function

Bruno Schapira[a] and Robert Young[b]

[a] *Département de Mathématiques, Bât. 425, Université Paris-Sud 11, F-91405 Orsay Cedex, France. E-mail: bruno.schapira@math.u-psud.fr*
[b] *Institut des Hautes Études Scientifiques, Le Bois Marie, 35 route de Chartres, F-91440 Bures-sur-Yvette, France. E-mail: rjyoung@ihes.fr*

Received 10 December 2009; revised 1 March 2010; accepted 9 March 2010

Abstract. We prove sharp estimates on the expected number of windings of a simple random walk on the square or triangular lattice. This gives new lower bounds on the averaged Dehn function, which measures the expected area needed to fill a random curve with a disc.

Résumé. Le principal résultat de cet article donne un équivalent précis de l'espérance du nombre total de tours effectués par la marche aléatoire simple sur $\mathbb{Z}^2$ ou sur le réseau triangulaire. Comme corollaire, nous obtenons une nouvelle borne inférieure de la fonction de Dehn moyennée sur $\mathbb{Z}^d$, $d \geq 2$, qui mesure l'aire moyenne du disque remplissant de manière optimale une courbe de longueur donnée.

MSC: 52C45; 60D05

Keywords: Simple random walk; Winding number; Averaged Dehn function

1. Introduction

The winding numbers of random curves have received much study, going back to Lévy [11] and Spitzer [14]. Lévy, in particular, studied the Lévy area of a planar Brownian motion, which is based in spirit on adding, with sign, the winding number of the curve around different regions of the plane. In this paper, we will study the total winding number of a random closed curve, that is, the integral of the magnitude of the winding number over the plane. The total winding number of a curve is always non-negative, unlike the Lévy area, and it is connected to problems of filling curves by discs or cycles.

Let $\theta_t(z)$ be the angular part of a Brownian motion with respect to z. Spitzer showed that $(\ln t)^{-1}\theta_t(z)$ converges to the Cauchy distribution, so the winding angle of a Brownian motion can be quite large. Further results on the winding number of Brownian motion suggest that the total winding number of a loop based on Brownian motion is infinite. Werner [16] showed that if $\theta_t(z)$ is the angular part of a Brownian motion with respect to z, then

$$k^2 \operatorname{area}\bigl\{z \in \mathbb{C} \mid \theta_t(z) - \theta_0(z) \in \bigl[2\pi k, 2\pi(k+1)\bigr)\bigr\} \to t/2\pi \tag{1}$$

in L^2 as $k \to +\infty$. Thus if γ is the loop formed by connecting the ends of a Brownian motion by a straight line and $i_\gamma(x)$ is the winding number of γ around x, then

$$\mathbb{E}\biggl[\int_{\mathbb{R}^2} \bigl|i_\gamma(x)\bigr| \,\mathrm{d}x\biggr] = \infty.$$

Similarly, Yor [17] explicitly computed the law of the index of a point $z \in \mathbb{C}$ with respect to a Brownian loop. Using this result, one can show that there is equality in (1) for all $k \neq 0$, so the expectation above is also infinite if γ is a Brownian loop.

This infinite area, however, is due to the presence of small regions with arbitrarily large winding number, which are not present if γ is the curve formed by connecting the ends of a random walk by a straight line. In fact, since γ has finite length, its total winding number is finite. Let $i_n(z)$ be the random variable corresponding to the number of times that a random walk of n steps winds around z. Windings of random walks have been studied less than windings of Brownian curves, but one result is an analogue of Spitzer's theorem for the random walk, due to Bélisle [2], who proved that for any fixed $z \in \mathbb{C}$, the distribution of $i_n(z)/\ln n$ converges to the hyperbolic secant distribution as $n \to +\infty$.

In this paper, we will prove the following theorem.

Theorem 1.1. *Let $(S_n, n \geq 0)$ be the simple random walk on the unit square lattice in the complex plane. Let $\mathcal{S}_n$ be the loop joining the points $S_0, S_1, \dots, S_n, S_0$ in order by straight lines and for $z \in \mathbb{C} \setminus \mathcal{S}_n$, let $i_n(z)$ be the index of z with respect to $\mathcal{S}_n$. Then*

$$\mathbb{E}\left[\int_{\mathbb{C}} \left|i_n(z)\right| \mathrm{d}z\right] \sim \frac{1}{2\pi} n \ln \ln n,$$

when $n \to +\infty$.

Note that the integral in the theorem is well defined since $\mathcal{S}_n$ has Lebesgue measure 0.

The basic idea of the proof is to use strong approximation to relate windings of the random walk to windings of Brownian motion. The main effect of replacing a Brownian curve with a random walk is to eliminate points with very high winding numbers, and the replacement does not affect points which are far from the Brownian motion. We thus find a lower bound on the expected winding number by considering points which are far from the random walk (Proposition 3.2). We find an upper bound by bounding the number of points with high winding numbers (Proposition 3.3). Windings around a point can be broken into classes depending on their distance from the point, and we bound the number of windings in each class separately. We will give some notation and preliminaries in Section 2, and prove Theorem 1.1 in Section 3.

In Section 4, we describe an application to geometric group theory. The isoperimetric problem is a classical problem in geometry which asks for the largest area that can be bounded by a loop of a given length. This question can be asked for a variety of spaces, and is particularly important in geometric group theory, where the growth rate of this area as a function of the length of the loop is known as the Dehn function and carries information on the geometry of a group (see [5] for a survey). Gromov [8], Chapter 5.A_6', proposed studying the distribution of areas of random curves as an alternative to studying the supremal area of a curve, and the total filling area of a curve in $\mathbb{R}^2$ is bounded below by its total winding number.

Finally, in Section 5, we give some extensions of Theorem 1.1, including a version that holds for the triangular lattice and a lower bound for Brownian bridges in the plane.

2. Preliminaries and notation

We start by laying out some of the background and notation for the rest of the paper. We will recall some standard notation and describe results of Zaitsev on approximating random walks by Brownian motion and of Werner on winding numbers of Brownian motion. Throughout this paper, ϕ will represent a positive L^1 function on $\mathbb{C}$; this function may change from one step to another, but we plan to provide some warning of these shifts.

We first describe some notation for the growth of functions. Recall that if g is positive, then $f(x) = \mathrm{O}(g(x))$ if and only if $\limsup_{x\to\infty} |f(x)/g(x)| < \infty$. We define a class of quickly-decaying functions:

$$\mathcal{G} = \left\{g : \mathbb{R}^+ \to \mathbb{R}^+ \;\middle|\; g(k) = \mathrm{O}\left(k^{-c}\right) \forall c > 0\right\}.$$

We next describe some notation for planar random walks and winding numbers. Throughout this paper, we will identify $\mathbb{R}^2$ with $\mathbb{C}$. Let $(S_i, i \geq 0)$ be a random walk on $\mathbb{R}^2$ with i.i.d. bounded increments, $\mathbb{E}[S_i - S_{i-1}] = 0$, and

$\mathrm{cov}(S_i - S_{i-1}) = \kappa I$. Except express mention of the contrary, we will assume that S is the simple random walk on the unit square lattice, for which $\kappa = 1/2$. For any n, consider the rescaled process $(X_t, 0 \le t \le 1)$ defined by

$$X_t := \frac{S_{\lfloor nt \rfloor} + (t - \lfloor nt \rfloor/n)(S_{\lfloor nt \rfloor+1} - S_{\lfloor nt \rfloor})}{\sqrt{\kappa n}}$$

for all $t \ge 0$ (the dependence on n in the notation will often be implicit). Here, $\lfloor x \rfloor$ represents the largest integer which does not exceed x. Denote by $\mathcal{C}_n$ the loop made of the curve $(X_t, 0 \le t \le 1)$ and the segment joining X_0 and X_1. This curve connects the points of the random walk in order and joins the endpoints. Note in comparison with the notation of Theorem 1.1 that $\mathcal{C}_n = \mathcal{S}_n/\sqrt{n/2}$. We will primarily work with X_t and $\mathcal{C}_n$ rather than S_i and $\mathcal{S}_n$.

If Y_t is a continuous function of $t \in [0,1]$ and z is not in its image, let $\psi_Y^z(t)$ be the unique continuous lift of $\mathrm{Im}(\ln(Y_t - z))$ such that $\psi_Y^z(0) \in [0, 2\pi)$. If $T = \bigcup_i I_i$ is a finite disjoint union of intervals I_i with endpoints x_i and y_i, let $w_Y^z(T) = \sum_i [\psi_Y^z(y_i) - \psi_Y^z(x_i)]$. If z is in the image of Y, set $w_Y^z(T) = 0$. For z not in the image of $\mathcal{C}_n$, let

$$j_n(z) := \left[\frac{w_X^z([0,1])}{2\pi}\right], \tag{2}$$

where $[x]$ represents the closest integer to x. Note that if $w_X^z([0,1])$ is an odd multiple of π, then z is on the line connecting X_0 and X_1. So if $z \notin \mathcal{C}_n$, then $j_n(z)$ is well defined, and this is the index of z with respect to $\mathcal{C}_n$. Actually it will be also convenient to define $j_n(z)$ when z is on the line connecting X_0 and X_1. In this case we set by convention

$$j_n(z) := \frac{w_X^z([0,1])}{2\pi} - \frac{1}{2}. \tag{3}$$

Observe now that

$$i_n\big(z\sqrt{n/2}\big) = j_n(z) \quad \text{for all } z \in \mathbb{C} \setminus \mathcal{C}_n.$$

One of our main tools is the fact that a random walk can be approximated by a Brownian motion. A strong approximation theorem due to Zaitsev [19], which improves bounds of Einmahl [6] and generalizes results of Komlós, Major and Tusnády [9], implies the following theorem.

Theorem 2.1 ([19]). *If $(X_t, 0 \le t \le 1)$ is defined as above, then there is a constant $c > 0$ such that for any $n > 1$, there exists a coupling of $(X_t, 0 \le t \le 1)$ with a Brownian motion $(\beta_t, 0 \le t \le 1)$ such that*

$$\mathbb{P}\left[\sup_{k \le n, k \in \mathbb{N}} |X_{k/n} - \beta_{k/n}| \ge c\frac{\ln n}{\sqrt{n}}\right] \le \frac{1}{n^4}.$$

Standard properties of the Brownian motion imply that there is a constant $c' > 0$ such that

$$\mathbb{P}\left[\sup_{t \le 1} \sup_{h \le 1/n} |\beta_t - \beta_{t+h}| \ge \frac{\ln n}{\sqrt{n}}\right] \le \frac{c'}{n^4}$$

for all $n > 1$, so there are constants $c_0 > 0$ and $c'' > 0$ such that

$$\mathbb{P}\left[\sup_{t \le 1} |X_t - \beta_t| \ge c_0\frac{\ln n}{\sqrt{n}}\right] \le \frac{c''}{n^4}$$

for all $n \ge 1$. Let

$$\varepsilon_n := c_0\frac{\ln n}{\sqrt{n}}$$

and let $\mathcal{U}$ be the event

$$\mathcal{U} = \mathcal{U}(n) := \left\{\sup_{t \le 1} |X_t - \beta_t| \le \varepsilon_n\right\}.$$

Let $(\beta_t, t \geq 0)$ be a complex Brownian motion starting from 0. As with the random walk, we can connect the endpoints of the Brownian motion by a line segment to form a closed curve $\tilde{\mathcal{C}}$. If $z \notin \tilde{\mathcal{C}}$, let

$$\tilde{j}(z) := \left[\frac{w_\beta^z([0,1])}{2\pi} \right]$$

be its index with respect to $\tilde{\mathcal{C}}$.

We will need to consider the number of times β winds around a point, especially when this occurs outside a ball of radius ε. Let $B(z,r)$ denote the closed disc with center $z \in \mathbb{C}$ and radius r. For $r \geq 0$, let

$$\begin{aligned} T_r(z) &:= \inf\{s \geq 0 \mid \beta_s \notin B(z,r)\}, \\ t_r(z) &:= \inf\{s \geq 0 \mid \beta_s \in B(z,r)\}. \end{aligned} \tag{4}$$

For any $z \neq 0$, we have the following skew-product representation (see, for instance, [10] or [13]):

$$\beta_t - z = \exp\left(\rho^z_{A^z_t} + \mathrm{i}\theta^z_{A^z_t}\right),$$

where $((\rho^z_t, \theta^z_t), t \geq 0)$ is a two-dimensional Brownian motion, and

$$A^z_t = \int_0^t \frac{1}{|\beta_s - z|^2}\, \mathrm{d}s \quad \text{for all } t \geq 0.$$

Note the intuition behind this representation; when β_t is far from z, then A^z increases slowly, so that $\arg(\beta_t - z)$ also varies slowly. For $\varepsilon > 0$ and $t \geq 0$, let

$$Z_\varepsilon(t) := \left(\int_0^t \frac{1_{\{|\beta_s - z| \geq \varepsilon\}}}{|\beta_s - z|^2}\, \mathrm{d}s \right)^{1/2}, \tag{5}$$

and let $Z_\varepsilon := Z_\varepsilon(1)$. This Z_ε controls the amount of winding around z which occurs while the Brownian motion is outside $B(z,\varepsilon)$. The next lemma is essentially taken from [15] and Lemma 2 and Corollary 3(ii) of [16]. It shows, among other things, that Z_ε is not likely to be much larger than $|\ln \varepsilon|$.

Lemma 2.2.

(i) *There is a* $g \in \mathcal{G}$, *such that for all* $\varepsilon \in (0, 1/2)$ *and all* $k \geq 1$,

$$\mathbb{P}[Z_\varepsilon \geq k] \leq g\left(k/|\ln \varepsilon|\right).$$

(ii) *There exists a function* $\phi \in L^1$, *such that for all* $\varepsilon \in (0, 1/2)$ *and all* $z \neq 0$,

$$\mathbb{P}\left[t_\varepsilon(z) \leq 1\right] \leq \frac{\phi(z)}{|\ln \varepsilon|}.$$

(iii) *There exists a function* $\phi \in L^1$, *such that for all* $\varepsilon \in (0, 1/2)$, *all* $z \neq 0$ *and all* $k \geq 1$,

$$\mathbb{P}[Z_\varepsilon \geq k] \leq \frac{\phi(z)}{k}.$$

Proof. We start with part (i). The proof is essentially contained in the proof of Lemma 2 in [16]. Let $M = \mathrm{e}^{\sqrt{k}}$. Consider first the case where $|z| \leq M/2$. The skew-product decomposition shows that $Z_\varepsilon(T_M(z))^2$ has the distribution of the exit time of $[0, \ln(M/\varepsilon)]$ by a reflected Brownian motion starting from $\ln|z| - \ln\varepsilon$ if $|z| > \varepsilon$, or from 0 if $|z| \leq \varepsilon$.

The Markov property implies that this exit time is dominated by the exit time $\sigma_{\ln(M/\varepsilon)}$ of $[0,\ln(M/\varepsilon)]$ by a reflected Brownian motion starting from 0. Thus

$$\begin{aligned}\mathbb{P}\big[Z_\varepsilon\big(T_M(z)\big)\geq k\big]&\leq \mathbb{P}\big[\sigma_{\ln(M/\varepsilon)}\geq k^2\big]\\&\leq \mathbb{P}\biggl[\sigma_1\geq \frac{k^2}{(\sqrt{k}+|\ln\varepsilon|)^2}\biggr]\\&\leq c\exp\biggl(-c'\frac{k^2}{(\sqrt{k}+|\ln\varepsilon|)^2}\biggr),\end{aligned}$$

where c and c' are positive constants. For the last inequality see, for instance, [12], Proposition 8.4. If $k/|\ln\varepsilon|>1$,

$$\begin{aligned}\mathbb{P}\big[Z_\varepsilon\big(T_M(z)\big)\geq k\big]&\leq c\exp\biggl(-c'\frac{k/|\ln\varepsilon|}{(|\ln\varepsilon|^{-1/2}+\sqrt{|\ln\varepsilon|/k})^2}\biggr)\\&\leq c\exp\biggl(-c'\frac{k}{10|\ln\varepsilon|}\biggr)\leq g\big(k/|\ln\varepsilon|\big)\end{aligned}$$

for some $g\in\mathcal{G}$.

On the other hand, since $|z|\leq M/2$, by the maximal inequality, there is a $g\in\mathcal{G}$ such that

$$\mathbb{P}\big[T_M(z)\leq 1\big]\leq \mathbb{P}\big[T_{M/2}(0)\leq 1\big]\leq g(k).$$

The case $|z|\leq M/2$ now follows from the inequality

$$\mathbb{P}[Z_\varepsilon\geq k]\leq \mathbb{P}\big[T_M(z)\leq 1\big]+\mathbb{P}\big[Z_\varepsilon\big(T_M(z)\big)\geq k\big].$$

Next assume that $|z|>M/2$. Then

$$\mathbb{P}[Z_\varepsilon\geq k]\leq \mathbb{P}\Bigl[\inf_{s\leq 1}|\beta_s-z|\leq k^{-1}\Bigr]\leq \mathbb{P}\big[T_{M/2}(0)\leq 1\big]\leq g'(k)$$

for some function $g'\in\mathcal{G}$, which concludes the proof of (i).

Part (ii) is essentially due to Spitzer [15] (see also [10] for a precise statement). Part (iii) is a special case of Corollary 3(ii) in [16]. □

Let ε_n be as in the remarks after Theorem 2.1. The consequence of (ii) in the previous lemma for the random walk is the

Corollary 2.3. *For any $c>0$, there exists a function $\phi\in L^1$, such that for all $z\neq 0$ and sufficiently large n,*

$$\mathbb{P}\bigl[\inf\bigl\{t\geq 0\mid X_t\in B(z,c\varepsilon_n)\bigr\}\leq 1\bigr]\leq \frac{\phi(z)}{\ln n}.$$

Proof. By Lemma 2.2(ii), there is a $\phi\in L^1$ and a $c'>0$ such that for sufficiently large n,

$$\begin{aligned}\mathbb{P}\bigl[\inf\bigl\{t\geq 0\mid X_t\in B(z,c\varepsilon_n)\bigr\}\leq 1\bigr]&\leq \mathbb{P}\bigl[t_{(c+1)\varepsilon_n}(z)\leq 1,\mathcal{U}\bigr]+\mathbb{P}\bigl[\mathcal{U}^c\bigr]\\&\leq \frac{\phi(z)}{\ln n}+c'n^{-4}.\end{aligned}$$

Since the length of a step of the random walk is bounded, if n is sufficiently large and $|z|>n+1$, then $\mathbb{P}[\inf\{t\geq 0\mid X_t\in B(z,c\varepsilon_n)\}\leq 1]=0$. We thus find

$$\mathbb{P}\bigl[\inf\bigl\{t\geq 0\mid X_t\in B(z,c\varepsilon_n)\bigr\}\leq 1\bigr]\leq \frac{\phi(z)}{\ln n}+c'\min\biggl\{n^{-4},\frac{|z|^{-3}}{n}\biggr\}\leq\frac{\phi'(z)}{\ln n},$$

as desired. □

3. Proof of Theorem 1.1

Let $D \subset \mathbb{C}$ be given by

$$D = \{z_1 + \mathrm{i}z_2 \mid z_1\sqrt{n/2}, z_2\sqrt{n/2} \notin \mathbb{Z} \text{ for any } n\}.$$

This is the set of points which are not on any of the edges of the rescaled lattices for the random walk, and $\mathbb{C} \setminus D$ has measure 0.

Theorem 1.1 follows by dominated convergence from the following proposition.

Proposition 3.1. *Let $(S_n, n \geq 0)$ be the simple random walk on the unit square lattice in the complex plane. For $z \in D$, let $j_n(z)$ be defined by (2) and (3). Then*

$$\lim_{n\to+\infty} \frac{\mathbb{E}[|j_n(z)|]}{\ln\ln n} = \frac{\int_0^1 p_s(0,z)\,\mathrm{d}s}{\pi} \quad \textit{for all } z \in D,$$

where $p_s(0,z) = (2\pi s)^{-1}\exp(-|z|^2/2s)$. Moreover, there is a function $\phi \in L^1$ such that for all sufficiently large n,

$$\frac{\mathbb{E}[|j_n(z)|]}{\ln\ln n} \leq \phi(z) \quad \textit{for all } z \in D.$$

3.1. *Sketch of proof*

The main idea of our proof is that there are no small windings in the rescaled random walk because the granularity of the walk makes it impossible to approach a point more closely than roughly $(\kappa n)^{-1/2}$. We thus expect the number of windings of $\mathcal{C}_n$ to be roughly the number of windings of the Brownian motion which stay far from z (relative to $(\kappa n)^{-1/2}$). This has the effect of eliminating points with large winding numbers, since points with many windings generally have many close windings. We will show that the number of points with winding number $\leq \ln n$ remains roughly the same, but there are many fewer points with winding number $> \ln n$.

Accordingly, we will bound the total winding number from below by considering just the points with small winding number.

Proposition 3.2 (Points with low index). *Let $(S_i, i \geq 0)$ be a random walk on $\mathbb{R}^2$ with i.i.d. bounded increments, $\mathbb{E}[S_i - S_{i-1}] = 0$, and $\operatorname{cov}(S_i - S_{i-1}) = \kappa I$. Let X_t and $j_n(z)$ be as in Section 2. Then*

$$\lim_{n\to+\infty} \frac{1}{\ln\ln n} \sum_{k=1}^{\ln n} \mathbb{P}[|j_n(z)| \geq k] = \frac{1}{\pi}\int_0^1 p_s(0,z)\,\mathrm{d}s \quad \textit{for all } z \in D,$$

and there is a $\phi \in L^1$ such that for all sufficiently large n,

$$\frac{1}{\ln\ln n} \sum_{k=1}^{\ln n} \mathbb{P}[|j_n(z)| \geq k] \leq \phi(z) \quad \textit{for all } z \in D.$$

Note that if $i_n(z)$ is the winding number of the unscaled random walk, as in Theorem 1.1, then

$$\mathbb{E}\left[\int_{\mathbb{C}} |i_n(z)|\,\mathrm{d}z\right] \geq \kappa n \sum_{k=1}^{\ln n} \mathbb{P}[|j_n(z)| \geq k],$$

so it follows as a corollary that

$$\liminf_{n\to\infty} \frac{1}{n\ln\ln n} \mathbb{E}\left[\int_{\mathbb{C}} |i_n(z)|\,\mathrm{d}z\right] \geq \frac{\kappa}{2\pi}.$$

We will use this in Section 4.

Werner's [16] results (see Lemma 3.4 in the next subsection) imply that the proposition also holds for the Brownian motion; that is, when $j_n(z)$ is replaced by $\tilde{j}(z)$, the number of times that the Brownian motion winds around z. This gives the lower bound required by Proposition 3.1.

In Section 3.2, we prove this proposition through strong approximation. Since the random walk is usually close to the Brownian motion, $j_n(z)$ and $\tilde{j}(z)$ can differ only if z is close to the Brownian motion. Using Spitzer's estimate of the area of the Wiener sausage, we will show that most of the points with winding numbers $\leq \ln n$ lie far from the curve.

This proposition relies mainly on strong approximation and not on properties of the random walk, so we can prove it for random walks with arbitrary increments. Furthermore, it can be proved using a weaker embedding theorem, such as the Skorokhod embedding theorem.

We get an upper bound by showing that there are few points with winding number $> \ln n$. In fact, we show that

Proposition 3.3. *Under the hypotheses of Proposition 3.1, there is a function $\phi \in L^1$ such that for all $k \geq \ln n$,*

$$\mathbb{P}\big[\big|j_n(z)\big| \geq k\big] \leq \frac{\phi(z)}{k} \quad \textit{for all } z \in D.$$

Furthermore, there is a $\phi' \in L^1$ such that for any $\varepsilon > 0$, any $k \geq (\ln n)^{1+\varepsilon}$ and n large enough,

$$\mathbb{P}\big[\big|j_n(z)\big| \geq k\big] \leq \frac{\ln n}{k^2}\phi(z) \quad \textit{for all } z \in D.$$

The proof of the proposition proceeds by decomposing the random walk into pieces that are close to and far from a given point z; this technique is similar to that used by Bélisle [2]. We bound the amount of winding accumulated by the faraway pieces using strong approximation and Lemma 2.2. We bound the winding of the nearby pieces by noting that the random walk usually spends little time near z, and so does not accumulate a large winding number while close to z. This is different from the case of the Brownian motion, in which most of the winding around a point with high winding number occurs very close to the point.

The proof of this proposition requires stronger machinery than the proof of the lower bound. In particular, it requires that S_n be the simple random walk on the square grid. Furthermore, we need the full power of Theorem 2.1, since Theorem 2.1 allows us to choose ε_n so that the ε_n-neighborhood of z contains roughly $(\ln n)^2$ grid points.

Assuming the two propositions, we can prove Proposition 3.1.

Proof of Proposition 3.1. Note that by summation by parts,

$$\begin{aligned}\mathbb{E}\big[\big|j_n(z)\big|\big] &= \sum_{k=-\infty}^{+\infty} |k|\mathbb{P}\big[j_n(z) = k\big] \\ &= \sum_{k=1}^{+\infty} \mathbb{P}\big[\big|j_n(z)\big| \geq k\big]. \qquad (6)\end{aligned}$$

Moreover, Proposition 3.3 implies that for n large enough,

$$\frac{1}{\ln\ln n}\sum_{k=\ln n}^{\infty} \mathbb{P}\big[\big|j_n(z)\big| \geq k\big] \leq \varepsilon\phi'(z) + \frac{\phi(z)}{\ln\ln n} \quad \text{for all } z \in D. \qquad (7)$$

Proposition 3.1 now immediately follows from (6), Proposition 3.2 and (7). □

We will prove Proposition 3.2 in the next subsection and Proposition 3.3 in Sections 3.3–3.6.

3.2. *Points with low index*

We concentrate here on Proposition 3.2. The proof combines strong approximation [19], Spitzer's estimate on the area of the Wiener sausage [15], and Werner's estimate of $\mathbb{P}[\tilde{j}(z)=k]$ [16]. Recall that Werner showed that

Lemma 3.4 ([16], Lemma 5).

(1) *For all $t>0$ and all $z\neq 0$,*

$$x^2\mathbb{P}\big[w_\beta^z\big([0,t]\big)\in[x,x+2\pi]\big]\stackrel{x\to\infty}{\longrightarrow}2\pi\int_0^t p_s(0,z)\,\mathrm{d}s.$$

(2) *For all $t>0$, there is a function $\phi_t\in L^1$ such that for all $z\neq 0$ and all x large enough,*

$$x^2\mathbb{P}\big[w_\beta^z\big([0,t]\big)\in[x,x+2\pi]\big]\le\phi_t(z).$$

Proof of Proposition 3.2. For $\varepsilon>0$, let W_ε be the ε-neighborhood of the loop $\tilde{\mathcal{C}}$ formed by connecting the endpoints of the Brownian motion. Since W_ε is mostly made up of the Wiener sausage of β, Spitzer's estimate [15] (see also [10]) shows that for all z,

$$\lim_{\varepsilon\to 0}|\ln\varepsilon|\mathbb{P}[z\in W_\varepsilon]=\pi\int_0^1 p_s(0,z)\,\mathrm{d}s,\tag{8}$$

and that there is a $\phi\in L^1$ such that

$$\mathbb{P}[z\in W_\varepsilon]\le\frac{\phi(z)}{|\ln\varepsilon|}.\tag{9}$$

Observe that on the set $\mathcal{U}$, the straight-line homotopy between $\tilde{\mathcal{C}}$ and $\mathcal{C}_n$ lies entirely in W_{ε_n}. So if $z\notin W_{\varepsilon_n}$, then $j_n(z)=\tilde{j}(z)$. Thus

$$\big|\mathbb{P}\big[|j_n(z)|\ge k,\mathcal{U}\big]-\mathbb{P}\big[|\tilde{j}(z)|\ge k,\mathcal{U}\big]\big|\le\mathbb{P}[z\in W_\varepsilon].$$

On the other hand, if $j_n(z)\neq\tilde{j}(z)$, then one of them is non-zero, so $|z|\le\sup_{0\le t\le 1}|X_t|$ or $|z|\le\sup_{0\le t\le 1}|\beta_t|$. Therefore

$$\begin{aligned}\big|\mathbb{P}\big[|j_n(z)|\ge k,\mathcal{U}^c\big]-\mathbb{P}\big[|\tilde{j}(z)|\ge k,\mathcal{U}^c\big]\big|&\le\mathbb{P}\Big[\sup_{0\le t\le 1}|X_t|>|z|,\mathcal{U}^c\Big]+\mathbb{P}\Big[\sup_{0\le t\le 1}|\beta_t|>|z|,\mathcal{U}^c\Big]\\&\le\mathbb{P}\big[\mathcal{U}^c\big]1_{\{|z|<n\}}+\min\Big\{\mathbb{P}\Big[\sup_{0\le t\le 1}|\beta_t|>|z|\Big],\mathbb{P}\big[\mathcal{U}^c\big]\Big\}.\end{aligned}$$

Since $\mathbb{P}[\mathcal{U}^c]=\mathrm{O}(n^{-4})$, there is a $\phi'\in L^1$ such that

$$\big|\mathbb{P}\big[|j_n(z)|\ge k,\mathcal{U}^c\big]-\mathbb{P}\big[|\tilde{j}(z)|\ge k,\mathcal{U}^c\big]\big|\le\frac{\phi'(z)}{n}.$$

Thus

$$\big|\mathbb{P}\big[|j_n(z)|\ge k\big]-\mathbb{P}\big[|\tilde{j}(z)|\ge k\big]\big|\le\mathbb{P}[z\in W_{\varepsilon_n}]+\frac{\phi'(z)}{n},\tag{10}$$

and by (8),

$$\left|\sum_{k=1}^{\ln n}\mathbb{P}\big[|j_n(z)|\ge k\big]-\sum_{k=1}^{\ln n}\mathbb{P}\big[|\tilde{j}(z)|\ge k\big]\right|\le(\ln n)\mathbb{P}[z\in W_{\varepsilon_n}]+\frac{\phi'(z)\ln n}{n}$$

is bounded. Thus

$$\lim_{n\to+\infty}\frac{1}{\ln\ln n}\sum_{k=1}^{\ln n}\mathbb{P}\big[\big|j_n(z)\big|\geq k\big]=\lim_{n\to\infty}\frac{1}{\ln\ln n}\sum_{k=1}^{\ln n}\mathbb{P}\big[\big|\tilde{j}(z)\big|\geq k\big]$$
$$=\frac{1}{\pi}\int_0^1 p_s(0,z)\,\mathrm{d}s,$$

where the last equality follows from Lemma 3.4. Finally, by (10), (9) and Lemma 3.4, there is a $\phi''\in L^1$ such that

$$\mathbb{P}\big[\big|j_n(z)\big|\geq k\big]\leq\frac{\phi''(z)}{k}+\frac{\phi''(z)}{\ln n},$$

and so

$$\sum_{k=1}^{\ln n}\mathbb{P}\big[\big|j_n(z)\big|\geq k\big]\leq\phi''(z)\ln\ln n+\phi''(z).$$

This concludes the proof of Proposition 3.2. □

3.3. *Decomposing the winding number*

In the next subsections, we will prove Proposition 3.3.

To bound the probability that a point will have large index, we will introduce a decomposition of the winding number into small, medium, and large windings, depending on their distance from the point. We will then use different methods to bound the number of windings. During large windings, the random walk stays away from the point and is well approximated by the Brownian motion. The number of medium and small windings is bounded by the fact that the random walk spends little time near the point, and while the bounds on the medium and large windings only use strong approximation, the bound on the small windings uses the full power of Theorem 2.1.

Define the stopping times $(\tau_i, i\geq 0)$ and $(\sigma_i, i\geq 1)$ by

$$\tau_0:=0,$$
$$\sigma_i:=\inf\big\{t\geq\tau_{i-1}\mid|\beta_t-z|<2\varepsilon_n\big\}\quad\forall i\geq 1,$$
$$\tau_i:=\inf\big\{t\geq\sigma_i\mid|\beta_t-z|>4\varepsilon_n\big\}\quad\forall i\geq 1.$$

These times divide the curve into pieces close to z and far from z. If $t\in[\sigma_i,\tau_i)$, then $|\beta_t-z|\leq 4\varepsilon_n$; if $t\in[\tau_i,\sigma_{i+1})$, then $|\beta_t-z|\geq 2\varepsilon_n$. Let

$$\mathcal{T}:=\bigcup_{i\geq 0}[\tau_i,\sigma_{i+1})\cap[0,1];$$

this is a.s. a finite union of intervals during which $|\beta_t-z|\geq 2\varepsilon_n$.

We also break the random walk into excursions with winding number $\pm 1/2$; this decomposition relies on the assumption that $(S_i, i\geq 0)$ is a nearest-neighbor walk on a square lattice, and this is the main step that requires this assumption. Let $\hat{z}$ be the center of the square in the rescaled lattice containing z; since $z\in D$, this is unique. Let

$$\Delta_z^{\pm}:=\hat{z}\pm(1+i)\mathbb{R}^+$$

be the halves of the diagonal line through $\hat{z}$. If the random walk hits Δ_z^+ before Δ_z^-, define

$$e_0:=\inf\big\{t\geq 0\mid X_t\in\Delta_z^+\big\},$$
$$e_{2i+1}:=\inf\big\{t\geq e_{2i}\mid X_t\in\Delta_z^-\big\},$$
$$e_{2i}:=\inf\big\{t\geq e_{2i-1}\mid X_t\in\Delta_z^+\big\};$$

otherwise, define the e_i's with $+$ and $-$ switched. Note that $e_i \in \mathbb{Z}/n$ for all i. We call intervals of the form $[e_i, e_{i+1}]$ excursions with respect to z, and let $\mathcal{E} = \mathcal{E}(z)$ be the set of all such excursions which are subsets of $[0, 1]$. If $e = [t_1, t_2]$ is an excursion, set $w(e) := w_X^z([t_1, t_2])/2\pi = \pm 1/2$. Then

$$\left| \frac{w_X^z([0,1])}{2\pi} - \sum_{e \in \mathcal{E}} w(e) \right| \le 2.$$

We will classify these excursions as small, medium or large. Let

$$\mathcal{E}_{\mathrm{sm}} = \mathcal{E}_{\mathrm{sm}}(z) := \bigl\{ [u, v] \in \mathcal{E} \mid |X_u - z| \le \varepsilon_n \bigr\},$$

be the set of excursions starting close to z; we call them *small excursions*. Let

$$\mathcal{E}_{\mathrm{lg}} = \mathcal{E}_{\mathrm{lg}}(z) := \bigl\{ [u, v] \in \mathcal{E} \mid [u, v] \subset \mathcal{T} \bigr\}.$$

These *large excursions* stay far from z. On the set $\mathcal{U}$, these two sets are disjoint. Let

$$\mathcal{E}_{\mathrm{md}} = \mathcal{E}_{\mathrm{md}}(z) := \mathcal{E} \setminus (\mathcal{E}_{\mathrm{sm}} \cup \mathcal{E}_{\mathrm{lg}})$$

be the set of *medium excursions*, so that every excursion is either small, medium or large.

Proposition 3.3 is a consequence of the following lemmas.

Lemma 3.5. *There exists a function $\phi \in L^1$, such that for all sufficiently large n, all $k \ge 1$, and all $z \in D$:*

(i) $\mathbb{P}\left[\left| \sum_{e \in \mathcal{E}_{\mathrm{sm}}} w(e) \right| \ge k \right] \le \frac{\ln n}{k^2} \phi(z).$

(ii) $\mathbb{P}\left[\left| \sum_{e \in \mathcal{E}_{\mathrm{md}}} w(e) \right| \ge k \right] \le \frac{\ln n}{k^2} \phi(z).$

Lemma 3.6.

(i) *There is a function $\phi \in L^1$, such that for any $\varepsilon > 0$, all sufficiently large n and $k \ge (\ln n)^{1+\varepsilon}$,*

$$\mathbb{P}\left[\left| \sum_{e \in \mathcal{E}_{\mathrm{lg}}} w(e) \right| \ge k \right] \le \frac{\ln n}{k^2} \phi(z) \quad \textit{for all } z \in D.$$

(ii) *There exists a function $\phi \in L^1$, such that for all sufficiently large n and all $k \ge 1$,*

$$\mathbb{P}\left[\left| \sum_{e \in \mathcal{E}_{\mathrm{lg}}} w(e) \right| \ge k \right] \le \frac{\phi(z)}{k} \quad \textit{for all } z \in D.$$

We will prove these lemmas in the next three subsections.

3.4. *Small excursions*

In this subsection, we will prove Lemma 3.5(i) by using results on the occupation times of points by a random walk on a lattice.

We first show that the symmetry of the random walk implies that excursions are equally likely to go clockwise or counterclockwise. For the simple random walk, this follows from the fact that reflecting an excursion across the line $\Delta_z^+ \cup \Delta_z^-$ results in an equally probable excursion which winds in the opposite direction, but there is also an argument which relies only on the random walk being symmetric. Let $[e_i, e_{i+1}]$ be an excursion from Δ_z^+ to Δ_z^- and let

$$t_0 = \max\bigl\{ t \in [e_i, e_{i+1}] \mid X_t \in \Delta_z^+ \bigr\}.$$

The portion of the excursion between t_0 and e_{i+1} travels between $p_1 = X_{t_0}$ and $p_2 = X_{e_{i+1}}$. Its time reversal travels from p_2 to p_1. Rotating this curve by 180 degrees and translating gives a curve going from p_1 to p_2 which winds around z in the opposite direction of the original. Replacing the section of the excursion between t_0 and e_{i+1} by this rotated curve is a measure-preserving map on the space of random walk paths. This map does not change $\mathcal{E}$ or $\mathcal{E}_{\mathrm{sm}}$ and changes the direction of the excursion from clockwise to counterclockwise or vice versa, so the variables $w([e_i, e_{i+1}])$ are independent and take values from $\{-1/2, 1/2\}$ with equal probability. Furthermore,

$$\mathbb{E}\left[\left(\sum_{e\in\mathcal{E}_{\mathrm{sm}}} w(e)\right)^2\right] = \frac{1}{4}\mathbb{E}[\#\mathcal{E}_{\mathrm{sm}}].$$

Many points have no small excursions; if the random walk never comes close to z, then $\mathcal{E}_{\mathrm{sm}} = \varnothing$. Let $\tau = \inf\{t \geq 0 \mid X_t \in B(z, 2\varepsilon_n)\}$. Then by Corollary 2.3,

$$\mathbb{P}[\tau \leq 1] \leq \frac{\phi(z)}{\ln n}$$

for some $\phi \in L^1$. Applying the Markov property, this gives

$$\mathbb{P}\left[\left|\sum_{e\in\mathcal{E}_{\mathrm{sm}}} w(e)\right| \geq k\right] \leq \frac{1}{4}\frac{\mathbb{E}[\#\mathcal{E}_{\mathrm{sm}}]}{k^2}\frac{\phi(z)}{\ln n},$$

where the expectation on the right-hand side is taken conditionally on the past before time τ. Since each small excursion starts at a point in $(\Delta_z^+ \cup \Delta_z^-) \cap B(z, \varepsilon_n)$, the number $\#\mathcal{E}_{\mathrm{sm}}$ of small excursions is bounded by the number of visits to such points; it is well known that the mean number of visits to each site is bounded by $c \ln n$ for some constant c (this follows from the fact that probability of return to the origin in k steps decays like $1/k$). Since the number of lattice points in $(\Delta_z^+ \cup \Delta_z^-) \cap B(z, \varepsilon_n)$ is of order $\ln n$, we get

$$\mathbb{E}[\#\mathcal{E}_{\mathrm{sm}}] \leq c(\ln n)^2$$

for some constant $c > 0$, which establishes Lemma 3.5(i).

3.5. *Medium excursions*

On the set $\mathcal{U}$, medium excursions start outside of $B(z, \varepsilon_n)$ and enter $B(z, 5\varepsilon_n)$ at some point. We will divide the medium excursions into two classes, depending on whether they start inside or outside the ball $B(z, 8\varepsilon_n)$, and bound the number in each class.

Let

$$\mathcal{E}'_{\mathrm{md}} = \mathcal{E}'_{\mathrm{md}}(z) := \big\{[t_1, t_2] \in \mathcal{E}_{\mathrm{md}} \mid |X_{t_1} - z| > 8\varepsilon_n\big\},$$

be the set of medium excursions starting outside $B(z, 8\varepsilon_n)$ and

$$\mathcal{E}''_{\mathrm{md}} = \mathcal{E}''_{\mathrm{md}}(z) := \big\{[t_1, t_2] \in \mathcal{E}_{\mathrm{md}} \mid |X_{t_1} - z| \leq 8\varepsilon_n\big\},$$

be its complement. Given $0 < a < b$, define the sequences of stopping times

$$\begin{aligned}
&\tau_0^{a,b} := 0,\\
&\sigma_i^{a,b} = \sigma_i^{a,b}(z) := \inf\big\{t \geq \tau_{i-1}^{a,b} \mid |\beta_t - z| < a\varepsilon_n\big\} \quad \forall i \geq 1,\\
&\tau_i^{a,b} = \tau_i^{a,b}(z) := \inf\big\{t \geq \sigma_i^{a,b} \mid |\beta_t - z| > b\varepsilon_n\big\} \quad \forall i \geq 1.
\end{aligned}$$

These stopping times record the number of times the Brownian motion crosses the annulus $B(z, b\varepsilon_n) \setminus B(z, a\varepsilon_n)$. Using the skew-product decomposition, we can bound the number of such crossings that occur before time 1.

Lemma 3.7. *For any $0 < a < b$, there are a constant $c > 0$ and functions $\phi \in L^1$ and $g \in \mathcal{G}$, such that for sufficiently large n and all $z \in \mathbb{C}$,*

$$\mathbb{P}\big[\sigma_i^{a,b} < 1\big] \le \frac{\phi(z)}{\ln n}\left(1 - \frac{c}{\ln n}\right)^i + g(n) \le \frac{\phi(z)}{\ln n}\mathrm{e}^{-ci/\ln n} + g(n).$$

Proof. Recall the definitions of $T_r(z)$ and $t_r(z)$ from (4). By the skew-product decomposition, there is a constant $c > 0$ (depending on a and b) such that a Brownian motion starting on the circle of radius $b\varepsilon_n$ around z has probability $p \le 1 - c/\ln n$ of entering $B(z, a\varepsilon_n)$ before escaping $B(z, n)$. Thus, if $\sigma_i^{a,b} < T_n(z)$, then a.s. $\tau_i^{a,b} < T_n(z)$ and

$$\mathbb{P}\big[\sigma_{i+1}^{a,b} < T_n(z) | \sigma_i^{a,b} < T_n(z)\big] = p.$$

Thus

$$\mathbb{P}\big[\sigma_i^{a,b} < T_n(z) | \sigma_1^{a,b} < T_n(z)\big] = p^i.$$

Furthermore it is well known that

$$\mathbb{P}[T_n(z) < 1] \le g(n)$$

for some $g \in \mathcal{G}$. Now $\sigma_1^{a,b} = t_{a\varepsilon_n}(z)$ and by Lemma 2.2(ii), there is a $\phi \in L^1$ such that for all sufficiently large n,

$$\mathbb{P}\big[\sigma_1^{a,b} < 1\big] = \mathbb{P}\big[t_{a\varepsilon_n}(z) < 1\big] \le \frac{\phi(z)}{\ln n}.$$

We conclude that

$$\mathbb{P}\big[\sigma_i^{a,b} < 1\big] \le \mathbb{P}\big[\sigma_i^{a,b} < T_n(z), \sigma_1^{a,b} < 1\big] + g(n) \le \frac{\phi(z)}{\ln n} p^i + g(n),$$

as desired. □

On $\mathcal{U}$, the Brownian motion crosses the annulus $B(z, 7\varepsilon_n) \setminus B(z, 6\varepsilon_n)$ during each excursion in $\mathcal{E}'_{\mathrm{md}}$. Thus, there are $\phi, \phi' \in L^1$ such that for all sufficiently large n,

$$\begin{aligned}\mathbb{P}\big[\#\mathcal{E}'_{\mathrm{md}} > k\big] &\le \mathbb{P}\big[\sigma_k^{6,7} < 1\big] + \mathbb{P}\big[\mathcal{U}^c\big] \\ &\le \frac{\phi(z)}{\ln n}\mathrm{e}^{-ck/\ln n} + \mathrm{O}\big(n^{-4}\big) \\ &\le \phi'(z)\frac{\ln n}{k^2} + \mathrm{O}\big(n^{-4}\big)\end{aligned} \tag{11}$$

for all k.

To bound the cardinality of $\mathcal{E}''_{\mathrm{md}}$, we will show that if $\mathcal{E}''_{\mathrm{md}}$ is large, then the rescaled random walk likely crosses the annulus

$$B(z, 11\varepsilon_n) \setminus B(z, 8\varepsilon_n)$$

many times before time 1. First note that, on $\mathcal{U}$,

$$\mathcal{E}''_{\mathrm{md}} \subset \bar{\mathcal{E}} := \big\{[t_1, t_2] \in \mathcal{E} \mid \varepsilon_n \le |X_{t_1} - z| \le 8\varepsilon_n\big\}.$$

Using strong approximation and the skew-product decomposition, one can show that the probability that a random walk started in

$$\big(B(z, 8\varepsilon_n) \setminus B(z, \varepsilon_n)\big) \cap \Delta_z^+$$

leaves $B(z, 11\varepsilon_n)$ before hitting Δ_z^- is bounded away from zero, say by $p > 0$. Let η be the number of times the random walk crosses from inside $B(z, 8\varepsilon_n)$ to outside $B(z, 11\varepsilon_n)$. Let $\{x_i\}$ be a sequence of independent random variables with $\mathbb{P}[x_i = 1] = p$, $\mathbb{P}[x_i = 0] = 1 - p$. Since the starting times of excursions in $\bar{\mathcal{E}}$ form a sequence of stopping times,

$$\mathbb{P}[\#\bar{\mathcal{E}} \geq 2k/p, \eta \leq k] \leq \mathbb{P}\left[\sum_{i=1}^{2k/p} x_i \leq k\right] \leq \mathrm{e}^{-ck}$$

for some $c > 0$. So by reproducing the steps of (11), we get

$$\mathbb{P}[\#\bar{\mathcal{E}} \geq 2k/p] \leq \phi(z)\frac{\ln n}{k^2} + \mathrm{O}\bigl(n^{-4}\bigr)$$

for some $\phi \in L^1$.

Combining the bounds on $\#\mathcal{E}'_{\mathrm{md}}$ and $\#\mathcal{E}''_{\mathrm{md}}$, we see that there is some $\phi \in L^1$ such that

$$\mathbb{P}\left[\left|\sum_{e\in\mathcal{E}_{\mathrm{md}}} w(e)\right| \geq \left(1+\frac{2}{p}\right)k\right] \leq \mathbb{P}\left[\#\mathcal{E}_{\mathrm{md}} \geq \left(1+\frac{2}{p}\right)k\right] \leq \phi(z)\frac{\ln n}{k^2} + \mathrm{O}\bigl(n^{-4}\bigr).$$

We can eliminate the $\mathrm{O}(n^{-4})$ term using an argument like the one in the proof of Corollary 2.3. That is, the number of medium excursions is at most n, and if $|z| > \sqrt{n/\kappa}$, then the random walk does not come near z. So

$$\begin{aligned}\mathbb{P}\left[\left|\sum_{e\in\mathcal{E}_{\mathrm{md}}} w(e)\right| \geq \left(1+\frac{2}{p}\right)k\right] &\leq \phi(z)\frac{\ln n}{k^2} + cn^{-4}1_{\{|z|\leq\sqrt{n/\kappa}\}}1_{\{(1+2/p)k\leq n\}} \\ &\leq \phi'(z)\frac{\ln n}{k^2}\end{aligned}$$

for some $c > 0$ and $\phi' \in L^1$, as desired. This proves Lemma 3.5(ii).

3.6. *Large excursions*

In this subsection, we will prove Lemma 3.6 using strong approximation.

Note first that on the set $\mathcal{U}$, the winding of the Brownian motion during $\mathcal{T}$ is approximated by the winding number of the large excursions. The difference between the two arises from excursions which intersect the beginning or end of an interval of $\mathcal{T}$; each such interval can thus increase the difference by at most 1. Thus on the set $\mathcal{U}$,

$$\left|\sum_{e\in\mathcal{E}_{\mathrm{lg}}} w(e)\right| \leq \left|\frac{w_\beta^z(\mathcal{T})}{2\pi}\right| + \#\{i \mid \tau_i \leq 1\} + 1$$

and

$$\mathbb{P}\left[\left|\sum_{e\in\mathcal{E}_{\mathrm{lg}}} w(e)\right| > 2k+1\right] \leq \mathbb{P}\left[\left|\frac{w_\beta^z(\mathcal{T})}{2\pi}\right| \geq k\right] + \mathbb{P}\bigl[\#\{i \mid \tau_i \leq 1\} > k\bigr] + \mathbb{P}\bigl[\mathcal{U}^c\bigr]. \tag{12}$$

By Lemma 3.7, we have

$$\mathbb{P}\bigl[\#\{i \mid \tau_i \leq 1\} > k\bigr] \leq \frac{\ln n}{k^2}\phi(z) + g(n) \tag{13}$$

for some $\phi \in L^1$ and $g \in \mathcal{G}$. To prove Lemma 3.6, it thus suffices to bound

$$\mathbb{P}\left[\left|\frac{w_\beta^z(\mathcal{T})}{2\pi}\right| \geq k\right].$$

Let

$$\widetilde{Z} := \left(\int_{\mathcal{T}} \frac{\mathrm{d}s}{|\beta_s - z|^2}\right)^{1/2}.$$

Since $\mathcal{T}$ is a finite union of intervals whose endpoints are stopping times for the Brownian motion, $w^z_\beta(\mathcal{T} \cap [0,t])$ is a continuous local martingale whose quadratic variation at time 1 is $\widetilde{Z}^2$. Thus $w^z_\beta(\mathcal{T})$ is equal in law to $\gamma\widetilde{Z}$, with γ a standard normal variable. If Z_{ε_n} is given by (5), then $\widetilde{Z} \le Z_{\varepsilon_n}$, and we get

$$\mathbb{P}\big[\big|w^z_\beta(\mathcal{T})\big| \ge k\big] \le \mathbb{P}\big[|\gamma| Z_{\varepsilon_n} \ge k\big].$$

We first show part (i) of the lemma. Let $\delta > 0$ be such that $(1-\delta)(1+\varepsilon) > 1$. Then by Lemma 2.2(i), for n sufficiently large depending on ε and $k \ge (\ln n)^{1+\varepsilon}$, there are $c > 0$, $c' > 0$, $\varepsilon' > 0$ and g', $g'' \in \mathcal{G}$ such that

$$\begin{aligned}\mathbb{P}\big[|\gamma| Z_{\varepsilon_n} \ge k\big] &\le \mathbb{P}\big[|\gamma| \ge k^\delta\big] + \mathbb{P}\big[Z_{\varepsilon_n} \ge k^{1-\delta}\big] \\ &\le c'\mathrm{e}^{-k^{2\delta}/2} + g\big(k^{1-\delta}/|\ln \varepsilon_n|\big) \\ &\le c'\mathrm{e}^{-k^{2\delta}/2} + g'\big(ck^{\varepsilon'}\big) \le g''(k).\end{aligned}$$

Thus for sufficiently large n and $k \ge (\ln n)^{1+\varepsilon}$,

$$\mathbb{P}\big[\big|w^z_\beta(\mathcal{T})\big| \ge k\big] \le k^{-4}.$$

Moreover, if $|w^z_\beta(\mathcal{T})| \ge \pi$, then we must have $T_{|z|/2}(0) \le 1$; this has probability less than $\mathrm{e}^{-d|z|^2}$ for some constant $d > 0$. By interpolation this gives

$$\mathbb{P}\big[\big|w^z_\beta(\mathcal{T})\big| \ge k\big] \le k^{-2}\mathrm{e}^{-d|z|^2/2}. \tag{14}$$

Thus by (12)–(14) there is a $\phi \in L^1$, independent of ε, such that for all sufficiently large n and $k \ge (\ln n)^{1+\varepsilon}$,

$$\mathbb{P}\left[\left|\sum_{e \in \mathcal{E}_{\mathrm{lg}}} w(e)\right| \ge k\right] \le \frac{\ln n}{k^2}\phi(z) + \mathrm{O}\big(n^{-4}\big).$$

As in the previous subsection, we can eliminate the $\mathrm{O}(n^{-4})$ term by changing ϕ. This proves part (i) of Lemma 3.6.

For part (ii), we use Lemma 2.2(iii). If n is sufficiently large, we have

$$\begin{aligned}\mathbb{P}\big[|\gamma| Z_{\varepsilon_n} \ge k\big] &= \int_0^\infty \frac{2}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\mathbb{P}\left[Z_{\varepsilon_n} \ge \frac{k}{x}\right]\mathrm{d}x \\ &\le \int_0^\infty \frac{2}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\frac{\phi(z)x}{k}\,\mathrm{d}x = \frac{2}{\sqrt{2\pi}}\frac{\phi(z)}{k}\end{aligned}$$

for some $\phi \in L^1$. Thus

$$\mathbb{P}\left[\left|\sum_{e \in \mathcal{E}_{\mathrm{lg}}} w(e)\right| > k\right] \le \frac{\phi'(z)}{k} + \mathrm{O}\big(n^{-4}\big)$$

for some $\phi' \in L^1$, and as above, the $\mathrm{O}(n^{-4})$ term can be absorbed into ϕ'.

4. Application to the study of Dehn functions

One motivation for studying the winding numbers of random walks comes from geometric group theory, namely the study of Dehn functions. For an introduction to Dehn functions with rigorous definitions, see [5]; we will only sketch the definitions. Given a closed curve γ in a space X, one can ask for the infimal area of a disc with boundary γ. We call this the filling area of γ, denoted $\delta(\gamma)$, and define a function δ so that $\delta(n)$ is the supremal filling area of curves in X with length at most n; this is the Dehn function of X.

Gromov noted that when a group G acts properly discontinuously and cocompactly on a connected, simply connected manifold or simplicial complex X, the filling area of curves can be described in terms of G [8]. Roughly, formal products of generators of G (words) correspond to curves in X and conversely, curves in X can be approximated by words. For example, if X is the universal cover of a compact manifold M and $G = \pi_1(M)$, generators of G correspond to closed curves in M. A formal product w of generators corresponds to a concatenation of these curves, and if w represents the identity in G, this concatenation lifts to a closed curve in X. A disc whose boundary is this curve corresponds to a way of using the relators of G to show that w represents the identity. We define the Dehn function of a group similarly to that of a space; if w is a word, let $\bar{w}$ be the element of G represented by w, and if $\bar{w}$ is the identity, let the *filling area* $\delta(w)$ of w be the minimal number of applications of relators necessary to reduce w to the identity. Define the *Dehn function* δ of G so that $\delta(n)$ is the maximal filling area of words with length at most n (that is, formal products with at most n terms). The Dehn function of a group depends a priori on the presentation of the group, but one can show that the growth rate of δ is a group invariant and that the Dehn function of G grows at the same rate as that of X.

Note that we are using two related notions of length: the length of a curve in a manifold and the length of a word, which is the number of generators in the word. We will denote word length by ℓ_w and curve length by ℓ_c; if γ is the curve corresponding to a word w, then $\ell_c(\gamma) \sim \ell_w(w)$.

The Dehn function measures maximal filling area; Gromov [8], Chapter 5.A_6', also proposed studying the distribution of $\delta(w)$ when w is a random word of length n. This has led to multiple versions of an averaged Dehn function or mean Dehn function; see [3] for some alternatives to our definition. We will give definitions based on random walks. Given a random walk on G supported on the generators, paths of the random walk correspond to words; in particular, the random walk induces a measure on the set of words of length n. We can thus define the *averaged Dehn function* $\delta_{\mathrm{avg}}(n)$ to be the expectation of $\delta(w)$ with respect to this measure, conditioned on the event that w represents the identity in G. If the random walk has a non-zero probability of not moving at each step, this is defined for all n. The averaged Dehn function is not known to be a group invariant and it is possible that its growth depends on the transition probabilities of the random walk.

The averaged Dehn function often behaves differently than the Dehn function. For example, in $\mathbb{Z}^2$, the Dehn function is quadratic, corresponding to the fact that a curve of length n in $\mathbb{R}^2$ encloses an area at most quadratic in n. On the other hand, random walks in $\mathbb{Z}^2$ enclose much less area, so if $G = \mathbb{Z}^2$, then

$$\delta_{\mathrm{avg}}(n) = \mathrm{O}(n \ln n)$$

for a wide variety of random walks [18]. A similar phenomenon occurs in nilpotent groups; if G is a nilpotent group and $\delta(n) = \mathrm{O}(n^k)$ for some $k > 2$, then $\delta_{\mathrm{avg}}(n) = \mathrm{O}(n^{k/2})$ for many random walks [18].

Since the averaged Dehn function involves bridges of random walks on groups, it is often difficult to work with. M. Sapir has proposed an alternative, the *random Dehn function* [3], which depends on a choice of a random walk and a choice of a word v_x for every element $x \in G$ so that v_x represents x^{-1}. If w is a word, $wv_{\bar{w}}$ is a word representing the identity; essentially, the v_x give a way to close up any path. One can then define

$$\delta_{\mathrm{rnd}}(n) = \mathbb{E}\big[\delta(wv_{\bar{w}})\big],$$

where w is chosen from the set of words of length n with measure corresponding to the random walk.

This function depends a priori on the choice of v_x, but if the length $\ell_w(v_x)$ of v_x satisfies $\ell_w(v_x) \sim \|x\|_2$, where $\|\cdot\|_2$ is the ℓ^2-norm in $\mathbb{Z}^d$, then v_x is usually much shorter than the random walk, and we will show that the choice of v_x has little effect on the asymptotics of $\delta_{\mathrm{rnd}}(n)$.

If $w = w_1 \cdots w_n$ is a word in $\mathbb{Z}^d$, let $\widetilde{w}$ be the curve in $\mathbb{R}^d$ connecting

$$\overline{w_1}, \overline{w_1 w_2}, \ldots, \overline{w_1 \cdots w_n} \in \mathbb{Z}^d \subset \mathbb{R}^d$$

by straight lines.

Proposition 4.1. *Let $(S_i, i \geq 0)$ be a random walk on $\mathbb{Z}^d$ with i.i.d. bounded increments and mean* 0. *Let $c > 0$, and for every $x \in \mathbb{Z}^d$, $d \geq 2$, let γ_x be a curve connecting x to* 0 *such that $\ell_c(\gamma_x) \leq c\|x\|_2$. There is a $c' > 0$ such that for sufficiently large n,*

$$\mathbb{E}\big[\delta(\widetilde{w}\gamma_{\bar{w}})\big] \geq c' n \ln\ln n,$$

where w is chosen from the set of words of length n with measure corresponding to the random walk.

Proof. Winding numbers provide a lower bound on the filling area of curves in the plane; if $\gamma : S^1 \to \mathbb{R}^2$, and $f : D^2 \to \mathbb{R}^2$ is a piecewise linear or smooth map with boundary γ, then $\# f^{-1}(x) \geq |i_n(x)|$ away from a set of measure zero and thus

$$\operatorname{area} f \geq \int_{\mathbb{R}^2} \big|i_\gamma(x)\big| \, \mathrm{d}x.$$

Indeed, $\int_{\mathbb{R}^2} |i_\gamma(x)| \, \mathrm{d}x$ is the abelianized area of $\widetilde{w}\gamma_{\bar{w}}$, a function defined in [1] as a lower bound for the filling area.

As in the 2-dimensional case, we can define a random closed curve $\mathcal{S}$ in $\mathbb{R}^d$ which consists of a random walk and a line connecting the endpoints. Let $p : \mathbb{R}^d \to \mathbb{R}^2$ be the projection onto the first two coordinates; since this map is area-decreasing,

$$\delta(\mathcal{S}) \geq \int_{\mathbb{R}^2} \big|i_{p(\mathcal{S})}(x)\big| \, \mathrm{d}x.$$

The curve $\mathcal{S}$ can be obtained from $\widetilde{w}\gamma_{\bar{w}}$ by replacing $\gamma_{\bar{w}}$ with a straight line. Define λ_x to be the concatenation of γ_x and the straight line connecting 0 and x; by the hypothesis on γ_x, this is a closed curve of length at most $(c+1)\|x\|_2$. It can be filled by a disc $f_x : D^2 \to \mathbb{R}^d$ whose boundary is λ_x and which satisfies $\operatorname{area}(f_x) \leq (c+1)^2\|x\|_2^2$.

If $h : D^2 \to \mathbb{R}^d$ is a disc with boundary $\widetilde{w}\gamma_{\bar{w}}$, we can construct a disc with boundary $\mathcal{S}$ by adjoining $f_{\bar{w}}$ to h, and so

$$\delta(\widetilde{w}\gamma_{\bar{w}}) \geq \int_{\mathbb{R}^2} \big|i_{p(\mathcal{S})}(x)\big| \, \mathrm{d}x - (c+1)^2\|\bar{w}\|_2^2.$$

Then Proposition 3.2 implies that there is a $c' > 0$ such that

$$\mathbb{E}\big[\delta(\widetilde{w}\gamma_{\bar{w}})\big] \geq c' n \ln\ln n - (c+1)^2 \mathbb{E}\big[\big(\|\bar{w}\|_2\big)^2\big].$$

Since $\bar{w}$ is the nth step of a random walk on $\mathbb{Z}^d$, $\bar{w}$ is the sum of n independent variables with bounded variance and mean 0, so there is a c'' such that $\mathbb{E}[\|\bar{w}\|_2^2] = c'' n$. Thus

$$\mathbb{E}\big[\delta(\widetilde{w} v_{\bar{w}})\big] \geq \frac{c'}{2} n \ln\ln n$$

for sufficiently large n, as desired. □

If v_x are as in the definition of the random Dehn function and $\ell_w(v_x) \sim \|x\|_2$, we can let $\gamma_x = \widetilde{v}_x$ in the proposition above to get a bound

$$\delta_{\mathrm{rnd}}(n) \geq cn \ln\ln n$$

for some c. If $\mathbb{Z}^d$ is given by the standard set of generators $z_1, \dots, z_d$, we can define an appropriate set of v_x by letting

$$v_{(x_1,\dots,x_d)} = z_1^{-x_1} \cdots z_d^{-x_d}.$$

Compare this result to the upper bound attributed to Sapir in [3], which states that under these conditions, there is a c' such that

$$\delta_{\mathrm{rnd}}(n) \le cn \ln n.$$

5. Extensions to bridges and other random walks

We can prove a similar result to Proposition 4.1 for the averaged Dehn function of $\mathbb{Z}^2$ by extending Proposition 3.2 to the random walk bridge. To prove Proposition 3.2, it suffices to have strong approximation of the random walk bridge by a Brownian bridge, estimates of $\mathbb{E}[\tilde{j}(z)]$, and estimates of the probability of lying near the loop; all of these results exist in the literature. Borisov [4] has proven a strong approximation theorem which applies to the simple random walk bridge in $\mathbb{Z}^2$, and the analogue of Werner's result for the Brownian loop is a consequence of Yor [17]. Note that for Brownian loop, the result is more precise since [17] gives an exact formula for the expected area of points with index $k \neq 0$. The case $k = 0$ has even been computed recently by Garban and Trujillo Ferreras [7]. Finally the upper bound on the probability of being close to the Brownian loop is a consequence of the fact that the law of the Brownian loop for $t \in [0, 1/2]$ or $t \in [1/2, 1]$ is absolutely continuous with respect to the law of the Brownian motion. More precisely, for $\varepsilon > 0$, denote by W'_ε the Wiener sausage associated to the Brownian loop. Denote by $\mathbb{P}$ or $\mathbb{E}$ the law of the Brownian motion and by $\mathbb{P}'$ the law of the Brownian loop. Now if $p_t(x, y)$ is the heat kernel, then one has

$$\mathbb{P}'\left[z \in W'_\varepsilon\right] \le 2\mathbb{E}\left[\frac{p_{1/2}(0, \beta_{1/2})}{p_1(0,0)} 1_{\{z \in W_\varepsilon\}}\right]$$

for all $\varepsilon > 0$. Since $p_{1/2}(0, x)$ is bounded, (8) implies that

$$\mathbb{P}'\left[z \in W'_\varepsilon\right] \le \phi(z)/|\ln \varepsilon|$$

for some $\phi \in L^1$ as desired. Thus there is a $c > 0$ such that $\delta_{\mathrm{avg}}(n) \ge cn \ln \ln n$ for sufficiently large n. We suspect that neither of these lower bounds is sharp and conjecture that in fact $\delta_{\mathrm{avg}}(n)$ and $\delta_{\mathrm{rnd}}(n)$ grow like $n \ln n$, corresponding to the upper bound in [18] for $\delta_{\mathrm{avg}}(n)$ and the upper bound on $\delta_{\mathrm{rnd}}(n)$ which is attributed to Sapir in [3]. The integral $\int_{\mathbb{R}^2} |i_\gamma(x)|\, \mathrm{d}x$ is the area necessary to fill γ with a chain or an arbitrary manifold with boundary, and requiring that the filling be a disc should increase the necessary area.

Extending Theorem 1.1 to more general situations seems difficult, because of the use of $\mathcal{E}$, but it can be extended to walks with certain symmetries. For instance, $\mathcal{E}$ can be defined for the simple random walk on the triangular lattice. If z is a point in the plane and X_t is the rescaled triangular random walk, let s be an edge of the triangle containing z, let Δ^+ and Δ^- be opposite sides of the straight line containing s and define e_i and $\mathcal{E}$ as before. If $e \in \mathcal{E}$, define $w(e)$ to be 0 if X_t traverses s during e. If this does not happen, then X_t goes around s in either the positive or negative direction; let $w(e) = \pm 1/2$ depending on the direction. Note that each direction is equally likely, just as in the case of the square lattice.

Then

$$\left| j_n(z) - \sum_{e \in \mathcal{E}} w(e) \right| \le 2 + \frac{g_n(s)}{2},$$

where $g_n(s)$ is the number of times that X_t traverses s. We can bound $\sum_{e \in \mathcal{E}} w(e)$ exactly as before, by breaking $\mathcal{E}$ into small, medium, and large excursions, so it remains only to bound $g_n(s)$. This is straightforward; since the random walk is n steps long, $\sum_s g_n(s) = n$, and there is a $\phi \in L^1$ such that $\mathbb{E}[g_n(s)] \le \phi(z)/n$. A similar proof holds for the simple random walk on the honeycomb lattice.

Acknowledgements

We would like to thank Anna Erschler, Jean-François Le Gall, David Mason, Pierre Pansu, Christophe Pittet, Oded Schramm and Wendelin Werner for useful discussions and suggestions.

References

[1] G. Baumslag, C. F. Miller III and H. Short. Isoperimetric inequalities and the homology of groups. *Invent. Math.* **113** (1993) 531–560. MR1231836

[2] C. Bélisle. Windings of random walks. *Ann. Probab.* **17** (1989) 1377–1402. MR1048932

[3] O. Bogopolski and E. Ventura. The mean Dehn functions of abelian groups. *J. Group Theory* **11** (2008) 569–586. MR2429356

[4] I. S. Borisov. On the rate of convergence in the "conditional" invariance principle. *Teor. Verojatn. Primen.* **23** (1978) 67–79. MR0471011

[5] M. R. Bridson. The geometry of the word problem. In *Invitations to Geometry and Topology*. *Oxf. Grad. Texts Math.* **7** 29–91. Oxford Univ. Press, Oxford, 2002. MR1967746

[6] U. Einmahl. Extensions of results of Komlós, Major, and Tusnády to the multivariate case. *J. Multivariate Anal.* **28** (1989) 20–68. MR0996984

[7] C. Garban and J. A. Trujillo Ferreras. The expected area of the filled planar Brownian loop is $\pi/5$. *Comm. Math. Phys.* **264** (2006) 797–810. MR2217292

[8] M. Gromov. Asymptotic invariants of infinite groups. In *Geometric Group Theory, Vol. 2 (Sussex, 1991)* 1–295. *London Math. Soc. Lecture Note Ser.* **182**. Cambridge Univ. Press, Cambridge, 1993. MR1253544

[9] J. Komlós, P. Major and G. Tusnády. An approximation of partial sums of independent RV's and the sample DF, I. *Z. Wahrsch. Verw. Gebiete* **32** (1975) 111–131. MR0375412

[10] J.-F. Le Gall. Some properties of planar Brownian motion. In *École d'Été de Probabilités de Saint-Flour XX, 1990* 111–235. *Lecture Notes in Math.* **1527**. Springer, Berlin, 1992. MR1229519

[11] P. Lévy. *Processus Stochastiques et Mouvement Brownien. Suivi d'une note de M. Loève*. Gauthier-Villars, Paris, 1948.

[12] S. C. Port and C. J. Stone. *Brownian Motion and Classical Potential Theory*. Academic Press, New York, 1978. MR0492329

[13] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* **293**. Springer, Berlin, 1999.

[14] F. Spitzer. Some theorems concerning 2-dimensional Brownian motion. *Trans. Amer. Math. Soc.* **87** (1958) 187–197. MR0104296

[15] F. Spitzer. Electrostatic capacity, heat flow, and Brownian motion. *Z. Wahrsch. Verw. Gebiete* **3** (1964) 110–121. MR0172343

[16] W. Werner. Sur les points autour desquels le mouvement Brownien plan tourne beaucoup. *Probab. Theory Related Fields* **99** (1994) 111–144. MR1273744

[17] M. Yor. Loi de l'indice du lacet Brownien, et distribution de Hartman–Watson. *Z. Wahrsch. Verw. Gebiete* **53** (1980) 71–95. MR0576898

[18] R. Young. Averaged Dehn functions for nilpotent groups. *Topology* **47** (2008) 351–367. MR2422531

[19] A. Y. Zaitsev. Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments. *ESAIM Probab. Statist.* **2** (1998) 41–108 (electronic). MR1616527