



HAL
open science

Windings of planar random walks and averaged Dehn function

Bruno Schapira, Robert Young

► **To cite this version:**

Bruno Schapira, Robert Young. Windings of planar random walks and averaged Dehn function. 2008.
hal-00296638v1

HAL Id: hal-00296638

<https://hal.science/hal-00296638v1>

Preprint submitted on 14 Jul 2008 (v1), last revised 7 Dec 2009 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

WINDINGS OF PLANAR RANDOM WALKS AND AVERAGED DEHN FUNCTION

BRUNO SCHAPIRA AND ROBERT YOUNG

ABSTRACT. We prove a sharp estimate on the expected value of the integral of the index 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.

1. INTRODUCTION

The study of the winding number of a planar random curve has a rather long history. Most of the results in the literature, however, concern complex Brownian motion or continuous time processes, for which stochastic calculus techniques apply. In this paper, we will prove some results on the winding number of a random walk, including the following:

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} be the loop joining the points $S_0, S_1, \dots, S_n, S_0$ in order by straight lines and for $z \in \mathbb{C}$, let $i_n(z)$ be the index of z with respect to \mathcal{S} . Then*

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

when $n \rightarrow +\infty$.

The same asymptotics hold for the simple random walk on the unit triangular lattice; we will sketch the proof for the triangular lattice in Section 4.

One of the first results on the winding number of a random curve is Spitzer's theorem [Spi58], which states that if θ_t is the angular part of the complex Brownian motion starting from 1, then $\theta_t / \ln t$ converges in law to a Cauchy distribution as $t \rightarrow +\infty$. This result has been extended in many ways; see for instance [BW96, LGY86, LGY87, LGY90, PY84, PY86, PY89]. Yor [Yor80] explicitly computed the law of the index of a Brownian loop with respect to a point $z \in \mathbb{C}$. Garban and Trujillo Ferreras recently used SLE techniques to calculate the expected area of the filled Brownian loop. Combining this with Yor's result one obtains an explicit formula for the expected area of the region with any given winding number (including 0). The asymptotics of these expected areas for the (unconditioned) Brownian motion were obtained by Werner [Wer94]: if $\theta_t(z)$ is the angular part of the Brownian motion seen from z , then

$$k^2 \text{ area} \{z \in \mathbb{C} \mid \theta_t(z) - \theta_0(z) \in [2\pi k, 2\pi(k+1))\} \rightarrow t/2\pi$$

in L^2 as $k \rightarrow +\infty$. Our Theorem 1.1 can be seen as a discrete counterpart of this result. There are fewer results on the winding number of a random walk. One is an

2000 *Mathematics Subject Classification.* 52C45; 60D05.

Key words and phrases. simple random walk, winding number, averaged Dehn function.

analogue of Spitzer's theorem for the random walk, due to B elisle [B el89]. B elisle proved that for any fixed $z \in \mathbb{C}$, the distribution of $i_n(z)/\ln n$ converges to the hyperbolic secant distribution as $n \rightarrow +\infty$.

One motivation for studying the winding numbers of random walks is to study their filling area, that is, the area necessary to fill them with a disc. If $\gamma : S^1 \rightarrow \mathbb{R}^2$ is a closed curve, $f : D^2 \rightarrow \mathbb{R}^2$ is a Lipschitz map such that $f|_{S^1} = \gamma$, and $i_\gamma(x)$ is the index of x with respect to γ , then $\#f^{-1}(x) \geq |i_\gamma(x)|$ except for a set of measure zero, so

$$\int_{\mathbb{R}^2} |i_\gamma(x)| dx \leq \text{area } f.$$

Theorem 1.1 thus gives a lower bound on the expected filling area of a random walk; this quantity is related to the averaged Dehn function from geometric group theory. We will discuss some of these applications in Section 4.

Acknowledgments: We would like to thank Anna Erschler, Jean-Fran ois Le Gall, David Mason, Pierre Pansu and Wendelin Werner for useful discussions.

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.

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

$$(1) \quad \mathcal{G} = \{g : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \mid g(k) = O(k^{-c}) \quad \forall c > 0\}.$$

We next describe some notation for planar random walks and winding numbers. Throughout this paper, we will identify \mathbb{R}^2 and \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 $\text{cov}(S_i - S_{i-1}) = \kappa I$, for example, the simple random walk on the unit square lattice. For any n , we can consider a process $(X_t, 0 \leq t \leq 1)$ defined by:

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

for all $t \geq 0$ (the dependence in n in the notations will often be implicit). Here, $\lfloor x \rfloor$ represents the largest integer which does not exceed x . Denote by \mathcal{C} the loop made of the curve $(X_t, 0 \leq t \leq 1)$ and the segment joining X_0 and X_1 . This curve connects the points of the random walk in order and joins the endpoints.

If Y_t is a continuous function of t and z is not in its image, let $\psi_Y^z(t)$ be the unique continuous lift of $\text{Im}(\ln(Y_t - z))$ 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 $w_Y^z(T) = \sum_i \psi_Y^z(y_i) - \psi_Y^z(x_i)$. Let

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

where $\lceil x \rceil$ represents the closest integer to x ; this is the index of z with respect to \mathcal{C} . 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 is not on \mathcal{C} , then $j_n(z)$ is well-defined. For $k \in \mathbb{Z}$, let

$$\begin{aligned} A_k &:= \{z \in \mathbb{C} \mid z \notin \mathcal{C}, j_n(z) = k\} \\ \mathcal{A}_k &:= \text{area}(A_k) = \int_{\mathbb{C}} 1_{\{z \in A_k\}} dz \\ A_{\geq k} &:= \{z \in \mathbb{C} \mid z \notin \mathcal{C}, j_n(z) \geq k\} \\ \mathcal{A}_{\geq k} &:= \text{area}(A_{\geq k}). \end{aligned}$$

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 [Zai98], which improves bounds of Einmahl [Ein89] and generalizes results of Komlós, Major and Tusnády [KMT75], implies the following:

Theorem 2.1 ([Zai98]). *If $(X_t, 0 \leq t \leq 1)$ is defined as above, then there is a constant $c > 0$ such that for any $n > 1$, there exists a coupling of $(X_t, 0 \leq t \leq 1)$ with a Brownian motion $(\beta_t, 0 \leq t \leq 1)$ such that*

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

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

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

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

$$(2) \quad \mathbb{P} \left[\sup_{t \leq 1} |X_t - \beta_t| \geq c_0 \frac{\ln n}{\sqrt{n}} \right] \leq \frac{c''}{n^4},$$

for all $n \geq 1$. Let

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

and let Ω be the event

$$(3) \quad \Omega := \left\{ \sup_{t \leq 1} |X_t - \beta_t| \leq \epsilon_n \right\}.$$

Observe that Ω depends on n .

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 $\tilde{\mathcal{C}}$. Then if we define

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

then $\tilde{j}(z)$ is the index of z with respect to $\tilde{\mathcal{C}}$. Define analogously \tilde{A}_k , $\tilde{A}_{\geq k}$, and $\tilde{A}_{\leq k}$ to be the regions of the plane around which $\tilde{\mathcal{C}}$ winds k (resp. $\geq k$ or $\leq k$) times and define $\tilde{\mathcal{A}}_k$, $\tilde{\mathcal{A}}_{\geq k}$, and $\tilde{\mathcal{A}}_{\leq k}$ to be their respective areas. Note that then

$$\tilde{A}_k = \text{area} \left\{ z \in \mathbb{C} \mid w_{\tilde{\beta}}^z([0, 1]) \in ((2k-1)\pi, (2k+1)\pi) \right\},$$

which differs slightly from Werner's definition, but satisfies the same asymptotics.

We will need to consider the number of times that a Brownian motion winds around a point, especially the windings which occur outside a ball of radius ϵ . Let

$(\beta_t, t \geq 0)$ be a complex Brownian motion starting from 0. Let $B(z, r)$ denote the closed disc with center $z \in \mathbb{C}$ and radius r . For $m \geq 0$, let

$$\begin{aligned} T_m(z) &:= \inf\{s \geq 0 \mid \beta_s \notin B(z, m)\} \\ t_m(z) &:= \inf\{s \geq 0 \mid \beta_s \in B(z, m)\}. \end{aligned}$$

For any $z \neq 0$, we have the following skew-product representation (see for instance [LG92] or [RY99]):

$$\beta_t - z = \exp(\rho_{A_t^z}^z + i\theta_{A_t^z}^z),$$

where $((\rho_t^z, \theta_t^z), t \geq 0)$ is a two-dimensional Brownian motion, and

$$A_t^z = \int_0^t \frac{1}{|\beta_s - z|^2} ds \quad \forall t \geq 0.$$

For $\epsilon > 0$ and $t \geq 0$, let

$$Z_\epsilon(t) := \left(\int_0^t \frac{1_{\{|\beta_s - z| \geq \epsilon\}}}{|\beta_s - z|^2} ds \right)^{1/2}$$

and let $Z_\epsilon := Z_\epsilon(1)$. This Z_ϵ controls the amount of winding around z which occurs while the Brownian motion is outside $B(z, \epsilon)$. The next lemma is essentially taken from [Spi64] and Lemma 2 and Corollary 3 (ii) of [Wer94]. It shows, among other things, that Z_ϵ is not likely to be much larger than $|\ln \epsilon|$.

Lemma 2.1.

- (1) *There exist positive constants $c > 0$ and $c' > 0$, and $g \in \mathcal{G}$, such that for all $\epsilon \in (0, 1/2)$ and all $k \geq 1$,*

$$\mathbb{P}[Z_\epsilon \geq k] \leq g(k) + c \exp\left(-c' \frac{k^2}{(\sqrt{k} + |\ln \epsilon|)^2}\right).$$

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

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

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

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

Proof. We start with part (1). The proof is essentially contained in the proof of Lemma 2 in [Wer94]. Let $M = e^{\sqrt{k}}$. Consider first the case where $|z| \leq M/2$. The skew-product decomposition shows that $Z_\epsilon(T_M(z))^2$ has the distribution of the exit time of $[0, \ln(M/\epsilon)]$ by a reflected Brownian motion starting from $\ln |z| - \ln \epsilon$ if $|z| > \epsilon$, or from 0 if $|z| \leq \epsilon$. The Markov property implies that this exit time is dominated by the exit time $\sigma_{\ln(M/\epsilon)}$ of $[0, \ln(M/\epsilon)]$ by a reflected Brownian motion

starting from 0. Thus

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

where c and c' are positive constants. For the last inequality see for instance [PS78] Proposition 8.4. Since $|z| \leq M/2$, by the maximal inequality, there is a $g \in \mathcal{G}$ such that

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

The case follows from the inequality

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

Now assume $|z| > M/2$. Then

$$\begin{aligned} \mathbb{P}[Z_\epsilon \geq k] &\leq \mathbb{P}\left[\inf_{s \leq 1} |\beta_s - z| \leq k^{-1}\right] \\ &\leq \mathbb{P}[T_{M/2}(0) \leq 1] \leq g(k), \end{aligned}$$

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

The second claim is essentially due to Spitzer [Spi64] (see also [LG92] for a precise statement). The third claim is a special case of Corollary 3 (ii) in [Wer94]. \square

Corollary 2.1. *For any $c > 0$, there exists a function $\phi \in L^1$, such that for all $z \neq 0$ and sufficiently large n ,*

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

Proof. By Lemma 2.1.(2), there is a $\phi \in L^1$ and a $c' > 0$ such that for sufficiently large n ,

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

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

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

as desired. \square

3. PROOF OF THEOREM

In this section, we will prove Theorem 1.1. Note that it suffices to prove that when $(S_i, i \geq 0)$ is the simple random walk on the unit square lattice,

$$\mathbb{E}\left[\int_{\mathbb{C}} |j_n(z)| dz\right] \sim \frac{1}{\pi} \ln \ln n.$$

Note that by summation by parts and symmetry,

$$\mathbb{E} \left[\int_{\mathbb{C}} |j_n(z)| dz \right] = \sum_{k=-\infty}^{+\infty} |k| \mathbb{E}(\mathcal{A}_k) = 2 \sum_{k=1}^{+\infty} \mathbb{E}(\mathcal{A}_{\geq k}),$$

and

$$\mathbb{E} \left[\int_{\mathbb{C}} |j_n(z)| dz \right] = 2 \sum_{k=1}^n \mathbb{E}(\mathcal{A}_{\geq k}),$$

since $j_n(z) < n/2 + 1$. It thus suffices to show that

$$2 \sum_{k=1}^n \mathbb{E}(\mathcal{A}_{\geq k}) \sim \frac{1}{\pi} \ln \ln n.$$

Werner showed

Theorem 3.1 ([Wer94]). $\frac{\tilde{\mathcal{A}}_k}{k^2} \xrightarrow[k \rightarrow \infty]{L^2} \frac{1}{2\pi}$

Thus

$$\sum_{k=-\ln n}^{\ln n} |k| \mathbb{E}(\tilde{\mathcal{A}}_k) \sim \frac{1}{\pi} \ln \ln n.$$

Werner's proof shows that large indices are generally due to many small windings around a point. We will see that the main difference between windings of random walks and of Brownian motion is the absence of these small windings and thus the absence of points with index larger than $(\ln n)(\ln \ln n)$.

Our proof has two parts. First, we will see that $\mathbb{E}(\tilde{\mathcal{A}}_{\geq k})$ and $\mathbb{E}(\mathcal{A}_{\geq k})$ are similar for $k \leq \ln n$. This uses strong approximation; since the random walk is close to the Brownian motion, the difference between the two regions must fall close to the Brownian motion. Using Spitzer's estimate of the area of the Wiener sausage, we will show that this difference is small.

Second, we will show that $\mathbb{E}(\mathcal{A}_{\geq k})$ drops sharply for $k > \alpha \ln n \ln \ln n$. This relies on 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 Belisle [Bel89]. We bound the amount of winding accumulated during the faraway pieces using strong approximation and Lemma 2.1. 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.

For the first part, we will show that

Proposition 3.1 (Points with low index). *Let $(S_i, i \geq 0)$ be a random walk on \mathbb{C} with i.i.d. bounded increments, $\mathbb{E}[S_i - S_{i-1}] = 0$, and $\text{cov}(S_i - S_{i-1}) = \kappa I$, and let $\mathcal{A}_{\geq k}$ be as in Section 2. Then*

$$\left| \mathbb{E}(\mathcal{A}_{\geq k}) - \frac{1}{2\pi k} \right| \leq \frac{c_1(k)}{k} + \frac{c_2(n)}{\ln n},$$

for some c_1, c_2 such that $\lim_{k \rightarrow \infty} c_1(k) = 0$ and $\lim_{n \rightarrow \infty} c_2(n) = \pi$.

Thus for any $\alpha > 0$,

$$\liminf_{n \rightarrow +\infty} \frac{1}{\ln \ln n} \sum_{k=1}^{\alpha(\ln n)(\ln \ln n)} \mathbb{E}(\mathcal{A}_{\geq k}) \geq \frac{1}{2\pi},$$

and

$$\limsup_{n \rightarrow +\infty} \frac{1}{\ln \ln n} \sum_{k=1}^{\alpha(\ln n)(\ln \ln n)} \mathbb{E}(\mathcal{A}_{\geq k}) \leq \frac{1}{2\pi} + \pi\alpha.$$

This proposition implies the lower bound in Theorem 1.1.

For the second part, we will show

Proposition 3.2. *Let $(S_i, i \geq 0)$ be the simple random walk on the unit square lattice in \mathbb{C} and $(X_t, t \geq 0)$ a rescaling of S_i as in Section 2. For any $\alpha > 0$, there is a function $\phi \in L^1$ such that for all z in the interior of a square of the rescaled lattice, if $k \geq \alpha(\ln n)(\ln \ln n)$, then*

$$\mathbb{P}[|j_n(z)| \geq k] \leq \frac{\sqrt{\ln \ln n}}{k} \phi(z).$$

Furthermore, there is a constant $C > 0$ and a function $\phi' \in L^1$ such that if $k \geq C(\ln n)(\ln \ln n)$,

$$\mathbb{P}[|j_n(z)| \geq k] \leq \frac{\ln n}{k^2} \phi'(z).$$

Therefore for any $\alpha > 0$,

$$\sum_{k=\alpha(\ln n)(\ln \ln n)}^n \mathbb{E}(\mathcal{A}_{\geq k}) = O(\sqrt{\ln \ln n}).$$

Given these propositions, we have

$$\liminf_{n \rightarrow +\infty} \frac{1}{\ln \ln n} \sum_{k=1}^n \mathbb{E}(\mathcal{A}_{\geq k}) \geq \frac{1}{2\pi},$$

and

$$\limsup_{n \rightarrow +\infty} \frac{1}{\ln \ln n} \sum_{k=1}^n \mathbb{E}(\mathcal{A}_{\geq k}) \leq \frac{1}{2\pi} + \pi\alpha$$

for any $\alpha > 0$, implying Theorem 1.1.

We will prove Proposition 3.1 in the next section, then prove Proposition 3.2 in sections 3.2–3.5.

3.1. Points with low index. We concentrate here on Proposition 3.1. The proof combines strong approximation [Zai98], Spitzer's estimate on the area of the Wiener sausage [Spi64], and Werner's estimate of the asymptotics of \mathcal{A}_k [Wer94].

Proof of Proposition 3.1: For $\epsilon > 0$, let W_ϵ be the ϵ -neighborhood of the loop $\tilde{\mathcal{C}}$ formed by connecting the endpoints of the Brownian motion. Let $\mathcal{W}_\epsilon = \text{area}(W_\epsilon)$. Since W_ϵ is mostly made up of the Wiener sausage of β , Spitzer's estimate [Spi64] (see also [LG92]) shows that:

$$(4) \quad \mathbb{E}[\mathcal{W}_\epsilon] \sim \frac{\pi}{|\ln \epsilon|},$$

when $\epsilon \rightarrow 0$, and thus

$$(5) \quad \mathbb{E}[\mathcal{W}_{\epsilon_n}] \sim \frac{\pi}{\ln n},$$

when $n \rightarrow +\infty$. Observe that on the set Ω , the straight-line homotopy between $\tilde{\mathcal{C}}$ and \mathcal{C} lies entirely in W_{ϵ_n} , so if $z \notin W_{\epsilon_n}$, then the index of z with respect to both curves is the same. Thus on the set Ω ,

$$\tilde{\mathcal{A}}_{\geq k} \setminus W_{\epsilon_n} \subset \mathcal{A}_{\geq k} \subset \tilde{\mathcal{A}}_{\geq k} \cup W_{\epsilon_n}$$

and so

$$|\mathbb{E}(\mathcal{A}_{\geq k} 1_{\Omega}) - \mathbb{E}(\tilde{\mathcal{A}}_{\geq k} 1_{\Omega})| \leq \mathbb{E}[\mathcal{W}_{\epsilon_n} 1_{\Omega}] \leq \mathbb{E}[\mathcal{W}_{\epsilon_n}].$$

On the other hand, for $k > 1$,

$$|\mathcal{A}_{\geq k} - \tilde{\mathcal{A}}_{\geq k}| \leq |\mathcal{A}_{\geq k}| + |\tilde{\mathcal{A}}_{\geq k}| \leq n + \pi \sup_{0 \leq t \leq 1} |\beta_t|^2,$$

so

$$\left| \mathbb{E}[\mathcal{A}_{\geq k} 1_{\Omega}] - \frac{1}{2\pi k} \right| \leq \left| \mathbb{E}[\tilde{\mathcal{A}}_{\geq k}] - \frac{1}{2\pi k} \right| + \mathbb{E}[\mathcal{W}_{\epsilon_n}] + \mathbb{E}[(n + \pi \sup_{0 \leq t \leq 1} |\beta_t|^2) 1_{\Omega^c}].$$

Let

$$\begin{aligned} c_1(k) &= k \left| \mathbb{E}(\tilde{\mathcal{A}}_{\geq k}) - \frac{1}{2\pi k} \right| \\ c_2(n) &= (\ln n)(\mathbb{E}[\mathcal{W}_{\epsilon_n}] + \mathbb{E}[(n + \pi \sup_{0 \leq t \leq 1} |\beta_t|^2) 1_{\Omega^c}]). \end{aligned}$$

Then

$$\left| \mathbb{E}(\mathcal{A}_{\geq k} 1_{\Omega}) - \frac{1}{2\pi k} \right| \leq \frac{c_1(k)}{k} + \frac{c_2(n)}{\ln n}.$$

By Theorem 3.1,

$$\lim_{k \rightarrow \infty} c_1(k) = 0.$$

Similarly, since $\mathbb{P}[\Omega^c] = O(n^{-4})$,

$$\begin{aligned} \mathbb{E}[n 1_{\Omega^c}] &= O(n^{-3}), \\ \mathbb{E} \left[\sup_{0 \leq t \leq 1} |\beta_t|^2 1_{\Omega^c} \right] &= O(n^{-2}) \end{aligned}$$

and thus

$$\lim_{n \rightarrow \infty} c_2(n) = \pi$$

as desired.

3.2. Decomposing the winding number. In the next subsections, we will prove Proposition 3.2; for the rest of the proof, $(S_i, i \geq 0)$ will denote the simple random walk on the unit square lattice.

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.

Let $z = z_1 + iz_2 \in \mathbb{C}$ be such that $z_1, z_2 \notin \mathbb{Z}/\sqrt{n/2}$, that is, z is in the interior of a square in the lattice corresponding to the rescaled random walk. Since the

random walk travels along lattice edges, the index of z differs from the index of the center of the square only if the line connecting the endpoints of the random walk travels through the square. We thus assume that z is the center of a square in the rescaled lattice.

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

$$\begin{aligned}\tau_0 &= 0, \\ \sigma_i &= \inf\{t \geq \tau_{i-1} \mid |\beta_t - z| < 2\epsilon_n\} \quad \forall i \geq 1, \\ \tau_i &= \inf\{t \geq \sigma_i \mid |\beta_t - z| > 4\epsilon_n\} \quad \forall i \geq 0.\end{aligned}$$

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

$$(6) \quad \mathcal{T} := \bigcup_{i \geq 0} [\tau_i, \sigma_{i+1}) \cap [0, 1];$$

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

We also break the random walk into excursions with winding number $\pm 1/2$; this decomposition relies on the fact that $(S_i, i \geq 0)$ is a nearest-neighbor walk on a square lattice. Let

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

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

$$\begin{aligned}e_0 &= \inf\{t \geq 0 \mid X_t \in \Delta_z^+\}, \\ e_{2i+1} &= \inf\{t \geq e_{2i} \mid X_t \in \Delta_z^-\}, \\ e_{2i} &= \inf\{t \geq e_{2i-1} \mid X_t \in \Delta_z^+\};\end{aligned}$$

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 and let \mathcal{E} be the set of all excursions which are subsets of $[0, 1]$. If $e = [t_1, t_2]$ is an excursion, then define $w(e) := w_X^z([t_1, t_2])/2\pi = \pm 1/2$. Then

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

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

$$\mathcal{E}_{\text{sm}} := \{[u, v] \in \mathcal{E} \mid |X_u - z| \leq \epsilon_n\}$$

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

$$\mathcal{E}_{\text{lg}} := \{[u, v] \in \mathcal{E} \mid [u, v] \subset \mathcal{T}\}.$$

These *large excursions* stay far from z . On the set Ω , these two sets are disjoint. Let

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

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

Proposition 3.2 follows from the following lemmas.

Lemma 3.1. *For any $\alpha > 0$, there exists a function $\phi \in L^1$, such that for all sufficiently large n , all $k \geq \alpha(\ln n)(\ln \ln n)$, and all $z \neq 0$,*

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

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

Lemma 3.2.

- (1) *There is a constant $C > 0$ and a function $\phi \in L^1$, such that for all sufficiently large n , all $k \geq C(\ln n)(\ln \ln n)$, and all $z \neq 0$,*

$$\mathbb{P} \left[\left| \sum_{e \in \mathcal{E}_{lg}} w(e) \right| \geq k \right] \leq \frac{\ln n}{k^2} \phi(z).$$

- (2) *For any $\alpha > 0$, there exists a function $\phi \in L^1$, such that for all sufficiently large n , all $\alpha(\ln n)(\ln \ln n) \leq k \leq \alpha^{-1}(\ln n)(\ln \ln n)$, and all $z \neq 0$,*

$$\mathbb{P} \left[\left| \sum_{e \in \mathcal{E}_{lg}} w(e) \right| \geq k \right] \leq \frac{\sqrt{\ln \ln n}}{k} \phi(z).$$

We will prove these lemmas in the next three sections.

3.3. Small excursions. In this section, we will prove Lemma 3.1.1 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

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

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 doesn't change \mathcal{E} or \mathcal{E}_{sm} and changes the direction of the excursion from clockwise to counterclockwise or vice versa, so the variables $w(e_i)$ are independent and take values from $\{-1/2, 1/2\}$ with equal probability. Furthermore,

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

Many points have no small excursions; if the random walk never comes close to z , then $\mathcal{E}_{sm} = \emptyset$. By Corollary 2.1,

$$\mathbb{P}[\inf\{t \geq 0 \mid X_t \in B(z, 2\epsilon_n)\} \leq 1] \leq \frac{\phi(z)}{\ln n}.$$

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

$$(7) \quad \mathbb{P} \left[\left| \sum_{e \in \mathcal{E}_{sm}} w(e) \right| \geq k \right] \leq \frac{1}{4} \frac{\mathbb{E}[\#\mathcal{E}_{sm}]}{k^2} \frac{\phi(z)}{\ln n},$$

where the expectation on the right hand side is taken conditionally on the past before time $t_{2\epsilon_n}(z)$. Since each small excursion starts at a point in $(\Delta_z^+ \cup \Delta_z^-) \cap$

$B(z, \epsilon_n)$, the number $\#\mathcal{E}_{\text{sm}}$ 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 . Since the number of lattice points in $(\Delta_z^+ \cup \Delta_z^-) \cap B(z, \epsilon_n)$ is of order $\ln n$, we get

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

for some constant c , which establishes the lemma.

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

Let

$$\mathcal{E}'_{\text{md}} := \{[t_1, t_2] \in \mathcal{E}_{\text{md}} \mid |X_{t_1} - z| > 8\epsilon_n\}$$

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

$$\mathcal{E}''_{\text{md}} := \{[t_1, t_2] \in \mathcal{E}_{\text{md}} \mid |X_{t_1} - z| \leq 8\epsilon_n\}$$

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

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

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

Lemma 3.3. *For any $0 < a < b$, there is a $c > 0$ and a function $\phi \in L^1$ such that for sufficiently large n ,*

$$\mathbb{P}[\sigma_i^{a,b} < 1, \Omega] \leq \frac{\phi(z)}{\ln n} \left(1 - \frac{c}{\ln n}\right)^i.$$

Proof. Let $r = n$. By the skew-product decomposition, there is a constant c such that a Brownian motion starting on the circle of radius $b\epsilon_n$ around z has probability $p \leq 1 - c/\ln n$ of entering $B(z, a\epsilon_n)$ before escaping $B(z, r)$. Thus, if $\sigma_i^{a,b} < T_r(z)$, then a.s. $\tau_i^{a,b} < T_r(z)$ and

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

Thus

$$\mathbb{P}[\sigma_i^{a,b} < T_r(z)] \leq p^i.$$

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

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

Since $T_r(z) \geq 1$ on Ω , by the Markov property applied at time $\sigma_1^{a,b}$,

$$\mathbb{P}[\sigma_i^{a,b} < 1, \Omega] \leq \mathbb{P}[\sigma_i^{a,b} < T_r(z), \sigma_1^{a,b} < 1] \leq \frac{\phi(z)}{\ln n} p^i$$

as desired. \square

On Ω , the Brownian motion crosses the annulus $B(z, 7\epsilon_n) \setminus B(z, 6\epsilon_n)$ during each excursion in \mathcal{E}'_{md} . Thus, for sufficiently large n ,

$$\mathbb{P}[\#\mathcal{E}'_{\text{md}} > k] \leq \mathbb{P}[\sigma_k^{6,7} < 1, \Omega] + \mathbb{P}[\Omega^c] \leq \frac{\phi(z)}{\ln n} \left(1 - \frac{c}{\ln n}\right)^k + O(n^{-4}) \leq \phi(z) \frac{\ln n}{k^2} + O(n^{-4})$$

for all $k > \alpha(\ln n)(\ln \ln n)$ as desired.

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

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

many times before time 1. First note that

$$\mathcal{E}''_{\text{md}} \subset \bar{\mathcal{E}} := \{[t_1, t_2] \in \mathcal{E} \mid \epsilon_n \leq |X_{t_1} - z| \leq 8\epsilon_n\}.$$

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

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

leaves $B(z, 11\epsilon_n)$ 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 $B(z, 8\epsilon_n)$ to outside $B(z, 11\epsilon_n)$. Since the starting times of excursions in $\bar{\mathcal{E}}$ form a sequence of stopping times, there is a $c > 0$ such that

$$\begin{aligned} \mathbb{P}[\bar{\mathcal{E}} \geq 2k/p] &= \mathbb{P}[\bar{\mathcal{E}} \geq 2k/p, \eta > k] + \mathbb{P}[\bar{\mathcal{E}} \geq 2k/p, \eta \leq k] \\ &\leq \mathbb{P}[\eta > k] + \mathbb{P}\left[\sum_{i=1}^{2k/p} x_i \leq k\right] \\ &\leq \mathbb{P}[\sigma_{k-1}^{9,10} < 1] + \mathbb{P}[\Omega^c] + e^{-ck} \end{aligned}$$

where the x_i 's are independent random variables with $\mathbb{P}[x_i = 1] = p$, $\mathbb{P}[x_i = 0] = 1 - p$. Thus for any $\alpha > 0$ and sufficiently large n , if $k > \alpha(\ln n)(\ln \ln n)$, then

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

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

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

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

We can eliminate the $O(n^{-4})$ term using an argument like that in Corollary 2.1. There is a c such that if $|z| > c\sqrt{n}$ or $k > n$, then $\mathbb{P}[\#\mathcal{E}_{\text{md}} \geq k] = 0$. There is a $\phi' \in L^1$ such that if $|z| \leq c\sqrt{n}$ and $k \leq np/2$, then

$$n^{-4} \leq \phi'(z) \frac{\ln n}{k^2},$$

so there is a $\phi''(z) \in L^1$ such that

$$\mathbb{P}\left[\left|\sum_{e \in \mathcal{E}_{\text{md}}} w(e)\right| \geq k\right] \leq \phi''(z) \frac{\ln n}{k^2},$$

as desired. This, combined with the results of the previous subsection, proves Lemma 3.1.

3.5. Large excursions. In this section, we will prove Lemma 3.2 using strong approximation.

Note first that on the set Ω , 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 Ω ,

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

and

$$(8) \quad \mathbb{P} \left[\left| \sum_{e \in \mathcal{E}_{1g}} w(e) \right| > 2k + 1 \right] \leq \mathbb{P} \left[\left| \frac{w_\beta^z(\mathcal{T})}{2\pi} \right| \geq k \right] + \mathbb{P} [\#\{i \mid \tau_i \leq 1\} > k] + \mathbb{P}[\Omega^c].$$

By Lemma 3.3, we have

$$\mathbb{P}[\#\{i \mid \tau_i \leq 1\} > k] \leq \frac{\ln n}{k^2} \phi(z) + \mathbb{P}[\Omega^c],$$

for some $\phi \in L^1$. To prove Lemma 3.2, it suffices to bound

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

Let

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

Since \mathcal{T} is a finite union of intervals whose endpoints are stopping times for the Brownian motion, $w_\beta^z(\mathcal{T} \cap [0, t])$ is a continuous local martingale whose quadratic variation at time 1 is \tilde{Z}^2 . Thus $w_\beta^z(\mathcal{T})$ is equal in law to $\gamma \tilde{Z}$, with γ a standard normal variable. Since $\tilde{Z} \leq Z_{\epsilon_n}$, we get

$$\mathbb{P} [|w_\beta^z(\mathcal{T})| \geq k] \leq \mathbb{P} [|\gamma| Z_{\epsilon_n} \geq k].$$

We first show part (1) of the lemma. Let $C > 0$. Then for n sufficiently large depending on C and $k \geq C(\ln n)(\ln \ln n)$, we have:

$$\begin{aligned} \mathbb{P} [|\gamma| Z_{\epsilon_n} \geq k] &\leq \mathbb{P} [|\gamma| \geq \sqrt{C \ln \ln n}] + \mathbb{P} [Z_{\epsilon_n} \geq (\ln n) \sqrt{C \ln \ln n}] \\ &\leq (\ln n)^{-C/2} + g \left((\ln n) \sqrt{C \ln \ln n} \right) + \\ &\quad \exp \left(-c' \frac{C(\ln n)^2 \ln \ln n}{(-\ln \epsilon_n + (\ln n)^{1/2} (C \ln \ln n)^{1/4})^2} \right) \\ &\leq 2(\ln n)^{-C/2} + \exp(-c'' C \ln \ln n) \\ &\leq 2(\ln n)^{-C/2} + (\ln n)^{-c'' C} \end{aligned}$$

for some constant $c'' > 0$.

Furthermore, if $k \geq (\ln n)^4$ and n is sufficiently large, we have

$$\begin{aligned} \mathbb{P} [|\gamma| Z_{\epsilon_n} \geq k] &\leq \mathbb{P} [|\gamma| \geq \sqrt{k}] + \mathbb{P} [Z_{\epsilon_n} \geq \sqrt{k}] \\ &\leq c''' e^{-k^{1/4}/c'''} , \end{aligned}$$

for some constant $c''' > 0$. This bound and the previous bound imply that for C such that $\min\{C/2, c''C\} > 20$ and sufficiently large n ,

$$\mathbb{P}[|w_\beta^z(\mathcal{T})| > k] \leq k^{-5}$$

for $k \geq C(\ln n)(\ln \ln n)$. Moreover if $|w_\beta^z(\mathcal{T})| \geq \pi$, then we must have $T_0(|z|/2) \leq 1$; this has probability less than $e^{-d|z|^2}$ for some constant $d > 0$. By interpolation this gives

$$\mathbb{P}[|w_\beta^z(\mathcal{T})| \geq k] \leq k^{-2}e^{-d|z|^2/2}.$$

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

$$\mathbb{P}\left[\left|\sum_{e \in \mathcal{E}_{1g}} w(e)\right| > k\right] \leq \frac{\ln n}{k^2}\phi(z) + O(n^{-4})$$

for $k \geq C(\ln n)(\ln \ln n)$. As in the previous section, we can eliminate the $O(n^{-4})$ term by changing ϕ . This proves part (1) of Lemma 3.2.

For part (2), we use Lemma 2.1.(3). If n is sufficiently large, we have

$$\begin{aligned} \mathbb{P}[|\gamma|Z_{\epsilon_n} \geq k] &\leq \mathbb{P}\left[|\gamma| \geq \sqrt{6 \ln \ln n}\right] + \mathbb{P}\left[Z_{\epsilon_n} \geq \frac{k}{\sqrt{6 \ln \ln n}}\right] \\ &\leq (\ln n)^{-3} + \frac{\phi(z)\sqrt{6 \ln \ln n}}{k}. \end{aligned}$$

As before, $\mathbb{P}[|w_\beta^z(\mathcal{T})| \geq \pi] \leq e^{-d|z|^2}$ for some constant $d > 0$, so by using interpolation again, one can show that there is a $\phi' \in L^1$ such that if n is sufficiently large and $\alpha(\ln n)(\ln \ln n) \leq k \leq \alpha^{-1}(\ln n)(\ln \ln n)$, then

$$\mathbb{P}[|\gamma|Z_{\epsilon_n} \geq k] \leq \frac{\phi'(z)}{k} + \frac{\phi(z)\sqrt{6 \ln \ln n}}{k},$$

and thus a $\phi'' \in L^1$ such that

$$\mathbb{P}\left[\left|\sum_{e \in \mathcal{E}_{1g}} w(e)\right| > k\right] \leq \frac{\phi''(z)\sqrt{\ln \ln n}}{k}$$

as desired.

4. OTHER RANDOM WALKS, BRIDGES AND APPLICATIONS

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 [Bri02]; 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 space X , the filling area of curves can be described in terms of G [Gro93]. 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 space 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 . 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 .

The Dehn function measures maximal filling area; Gromov [Gro93, 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 [VB] 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 $\delta_{\text{avg}}(n)$ 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 nonzero 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, and the averaged Dehn function is at most $O(n \ln n)$ for a wide variety of random walks. A similar phenomenon occurs in nilpotent groups; if $\delta(n) = O(n^k)$ for some $k > 2$, then $\delta_{\text{avg}}(n) = O(n^{k/2})$ for many random walks [You08].

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 [VB], 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

$$\delta_{\text{rd}}(n) = \mathbb{E}[\delta(wv_{\bar{w}})]$$

where w is chosen from the set of words of length n with measure corresponding to the random walk. This function depends on the choice of v_x , but if the v_x are short enough, the choice may not affect the function. In particular, Proposition 3.1 implies the following:

Proposition 4.1. *For every $x \in \mathbb{Z}^d$, $d \geq 2$, let v_x be a word of minimal length representing x^{-1} . Then there is a $c > 0$ such that the random Dehn function of \mathbb{Z}^d with respect to the v_x and any random walk on \mathbb{Z}^d with i.i.d. bounded increments, $\mathbb{E}[S_i - S_{i-1}] = 0$, and whose support generates \mathbb{Z}^d satisfies*

$$\delta_{\text{rd}}(n) \geq cn \ln \ln n$$

for sufficiently large n .

Proof. To use Proposition 3.1, we need to use the correspondence between words in \mathbb{Z}^d and curves in \mathbb{R}^d . If $w = w_1 \dots w_n$ is a word in some generating set for \mathbb{Z}^d , let \widetilde{w} be the curve in \mathbb{R}^d connecting

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

by straight lines.

It suffices to show that

$$\mathbb{E}[\delta(\widetilde{w\overline{v_{\widetilde{w}}})}] \geq cn \ln \ln n.$$

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

$$\text{area } f \geq \int_{\mathbb{R}^2} |i_\gamma(x)| dx.$$

Indeed, $\int_{\mathbb{R}^2} |i_\gamma(x)| dx$ is the abelianized area of $w\overline{v_{\widetilde{w}}}$, a function defined in [BMS93] 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 $p : \mathbb{R}^d \rightarrow \mathbb{R}^2$ be the projection onto the first two coordinates; since this map is area-decreasing,

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

The curve \mathcal{S} can be obtained from $\widetilde{w\overline{v_{\widetilde{w}}}}$ by replacing $\widetilde{v_{\widetilde{w}}}$ with a straight line. For every $x \in \mathbb{Z}^d$, let $f_x : D^2 \rightarrow \mathbb{R}^d$ be a disc whose boundary consists of v_x and a straight line connecting 0 and x and such that $\text{area}(f_x) \leq 4\ell(v_x)^2$. If $h : D^2 \rightarrow \mathbb{R}^d$ is a disc with boundary $\widetilde{w\overline{v_{\widetilde{w}}}}$, we can construct a disc with boundary \mathcal{S} by adjoining $f_{\widetilde{w}}$ to h , and so

$$\delta(\widetilde{w\overline{v_{\widetilde{w}}})} \geq \int_{\mathbb{R}^2} |i_{p(\mathcal{S})}(x)| dx - 4\ell(v_w)^2.$$

Then Proposition 3.1 implies that

$$\mathbb{E}[\delta(\widetilde{w\overline{v_{\widetilde{w}}})}] \geq cn \ln \ln n - 4\mathbb{E}[\ell(v_w)^2] \geq \frac{c}{2}n \ln \ln n$$

for sufficiently large n , as desired. \square

We can prove a similar result for the averaged Dehn function by extending Proposition 3.1 to the random walk bridge. To prove Proposition 3.1, it suffices to have strong approximation of the random walk bridge by a Brownian bridge, estimates of $\mathbb{E}(\mathcal{A}_k)$, and estimates of the area of the Wiener sausage; all of these results exist in the literature. Borisov [Bor78] has proven a strong approximation theorem which applies to the simple random walk bridge, and the analogue of Werner's result for the Brownian loop is a consequence of Yor [Yor80] (see [GTF06] for an explicit computation). As Werner has informed us, the upper bound on the area of the Wiener sausage for 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 $\epsilon > 0$, denote by \mathcal{W}'_ϵ the area of the Wiener sausage associated to the Brownian loop. Denote by \mathbb{E}

the law of the Brownian motion and by \mathbb{E}' the law of the Brownian loop. Now if $p_t(x, y)$ is the heat kernel, then one has

$$\mathbb{E}'[\mathcal{W}'_\epsilon] \leq 2\mathbb{E}\left[\frac{p_{1/2}(0, \beta_{1/2})}{p_1(0, 0)}\mathcal{W}_\epsilon\right],$$

for all $\epsilon > 0$. Since $p_{1/2}(0, x)$ is bounded, (4) implies that

$$\mathbb{E}'[\mathcal{W}'_\epsilon] \leq c/|\ln \epsilon|$$

as desired. Thus there is a $c > 0$ such that $\delta_{\text{avg}}(n) \geq cn \ln \ln n$ for sufficiently large n . We suspect that neither of these lower bounds is sharp and conjecture that in fact $\delta_{\text{avg}}(n)$ and $\delta_{\text{rnd}}(n)$ grow like $n \ln n$. The integral $\int_{\mathbb{R}^2} |i_\gamma(x)| dx$ 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, but it can be extended to the simple random walk on a triangular lattice. We will sketch the necessary changes to the proof. The main change is the definition of \mathcal{E} ; instead of letting z be in the center of a rescaled square and considering paths going between halves of a diagonal line, let z be in the center of a rescaled triangle and let s be an edge of the triangle. Let Δ^+ and Δ^- be opposite sides of the straight line containing s and define e_i and \mathcal{E} as before. 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. As before, one can show that $\mathbb{E}[w([e_i, e_{i+1}])] = 0$. The proof then gives bounds on $\mathbb{P}[\sum_{e \in \mathcal{E}} w(e) > k]$ and shows that

$$\mathbb{E}\left[\frac{\sqrt{3}}{4n} \sum_z \left| \sum_{e \in \mathcal{E}} w(e) \right| \right] \sim \frac{1}{\pi} \ln \ln n$$

where the first sum is taken over centers of triangles. This is a close approximation to the integral of the winding numbers. Indeed,

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

where $f(s)$ is the number of times that X_t traverses s . Since X_t takes n steps in all, $\sum_s f(s) = n$, and

$$\left| \frac{\sqrt{3}}{4n} \sum_z \left| \sum_{e \in \mathcal{E}} w(e) \right| - \int_z |j_n(z)| dz \right| \leq c \sup_t |X_t|^2 + \frac{c}{n} \sum_s f(s) \leq c \sup_t |X_t|^2 + c$$

for some constant c . Thus

$$\mathbb{E}\left[\int_z |j_n(z)| dz\right] \sim \frac{1}{\pi} \ln \ln n.$$

REFERENCES

- [Bél89] C. Bélisle, *Windings of random walks*, Ann. Probab. **17** (1989), no. 4, 1377–1402. MR **MR1048932** (91d:60112)
- [BMS93] G. Baumslag, C. F. Miller, III, and H. Short, *Isoperimetric inequalities and the homology of groups*, Invent. Math. **113** (1993), no. 3, 531–560. MR **MR1231836** (95b:57004)
- [Bor78] I. S. Borisov, *On the rate of convergence in the “conditional” invariance principle*, Teor. Veroyatnost. i Primenen. **23** (1978), no. 1, 67–79. MR **MR0471011** (57 #10754)

- [Bri02] M. R. Bridson, *The geometry of the word problem*, Invitations to geometry and topology, Oxf. Grad. Texts Math., vol. 7, Oxford Univ. Press, Oxford, 2002, pp. 29–91. MR **MR1967746 (2004g:20056)**
- [BW96] J. Bertoin and W. Werner, *Stable windings*, Ann. Probab. **24** (1996), no. 3, 1269–1279. MR **MR1411494 (98a:60105)**
- [Ein89] U. Einmahl, *Extensions of results of Komlós, Major, and Tusnády to the multivariate case*, J. Multivariate Anal. **28** (1989), no. 1, 20–68. MR **MR996984 (90g:60032)**
- [Gro93] M. Gromov, *Asymptotic invariants of infinite groups*, Geometric group theory, Vol. 2 (Sussex, 1991), London Math. Soc. Lecture Note Ser., vol. 182, Cambridge Univ. Press, Cambridge, 1993, pp. 1–295. MR **MR1253544 (95m:20041)**
- [GTF06] C. Garban and J. A. Trujillo Ferreras, *The expected area of the filled planar Brownian loop is $\pi/5$* , Comm. Math. Phys. **264** (2006), no. 3, 797–810. MR **MR2217292 (2007c:82035)**
- [KMT75] J. Komlós, P. Major, and G. Tusnády, *An approximation of partial sums of independent RV's and the sample DF. I*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **32** (1975), 111–131. MR **MR0375412 (51 #11605b)**
- [LG92] J.-F. Le Gall, *Some properties of planar Brownian motion*, École d'Été de Probabilités de Saint-Flour XX—1990, Lecture Notes in Math., vol. 1527, Springer, Berlin, 1992, pp. 111–235. MR **MR1229519 (94g:60156)**
- [LGY86] J.-F. Le Gall and M. Yor, *Étude asymptotique de certains mouvements browniens complexes avec drift*, Probab. Theory Relat. Fields **71** (1986), no. 2, 183–229. MR **MR816704 (87d:60072)**
- [LGY87] ———, *Étude asymptotique des enlacements du mouvement brownien autour des droites de l'espace*, Probab. Theory Related Fields **74** (1987), no. 4, 617–635. MR **MR876259 (88d:60073)**
- [LGY90] ———, *Enlacements du mouvement brownien autour des courbes de l'espace*, Trans. Amer. Math. Soc. **317** (1990), no. 2, 687–722. MR **MR946219 (90e:60099)**
- [PS78] S. C. Port and C. J. Stone, *Brownian motion and classical potential theory*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1978, Probability and Mathematical Statistics. MR **MR0492329 (58 #11459)**
- [PY84] J. W. Pitman and M. Yor, *The asymptotic joint distribution of windings of planar Brownian motion*, Bull. Amer. Math. Soc. (N.S.) **10** (1984), no. 1, 109–111. MR **MR722863 (85d:60154)**
- [PY86] ———, *Asymptotic laws of planar Brownian motion*, Ann. Probab. **14** (1986), no. 3, 733–779. MR **MR841582 (88a:60145)**
- [PY89] ———, *Further asymptotic laws of planar Brownian motion*, Ann. Probab. **17** (1989), no. 3, 965–1011. MR **MR1009441 (90j:60080)**
- [RY99] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR **MR1725357 (2000h:60050)**
- [Spi58] F. Spitzer, *Some theorems concerning 2-dimensional Brownian motion*, Trans. Amer. Math. Soc. **87** (1958), 187–197. MR **MR0104296 (21 #3051)**
- [Spi64] ———, *Electrostatic capacity, heat flow, and Brownian motion*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **3** (1964), 110–121. MR **MR0172343 (30 #2562)**
- [VB] E. Ventura and O. Bogopolski, *The mean Dehn function of abelian groups*, J. Group Theory, to appear.
- [Wer94] W. Werner, *Sur les points autour desquels le mouvement brownien plan tourne beaucoup*, Probab. Theory Related Fields **99** (1994), no. 1, 111–144. MR **MR1273744 (95a:60111)**
- [Yor80] M. Yor, *Loi de l'indice du lacet brownien, et distribution de Hartman-Watson*, Z. Wahrsch. Verw. Gebiete **53** (1980), no. 1, 71–95. MR **MR576898 (82a:60120)**
- [You08] R. Young, *Averaged Dehn functions for nilpotent groups*, Topology **47** (2008), no. 5, 351–367.
- [Zai98] A. Yu. 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). MR **MR1616527 (99k:60094)**

DÉPARTEMENT DE MATHÉMATIQUES, BÂT. 425, UNIVERSITÉ PARIS-SUD 11, F-91405 ORSAY,
CEDEX, FRANCE.

E-mail address: `bruno.schapira@math.u-psud.fr`

INSTITUT DES HAUTES ÉTUDES SCIENTIFIQUES. LE BOIS MARIE, 35 ROUTE DE CHARTRES,
F-91440 BURES-SUR-YVETTE, FRANCE

E-mail address: `rjyoung@ihes.fr`