The Hammersley-Welsh bound for self-avoiding walk revisited

Tom Hutchcroft

November 23, 2017

Abstract

The Hammersley-Welsh bound (Quart. J. Math., 1962) states that the number c_n of length n self-avoiding walks on \mathbb{Z}^d satisfies

$$c_n \le \exp\left[O(n^{1/2})\right] \mu_c^n,$$

where $\mu_c = \mu_c(d)$ is the connective constant of \mathbb{Z}^d . While stronger estimates have subsequently been proven for $d \geq 3$, for d = 2 this has remained the best rigorous, unconditional bound available. In this note, we give a new, simplified proof of this bound, which does not rely on the combinatorial analysis of unfolding. We also prove a small, non-quantitative improvement to the bound, namely

$$c_n \le \exp\left[o(n^{1/2})\right] \mu_c^n.$$

The improved bound is obtained as a corollary to the sub-ballisticity theorem of Duminil-Copin and Hammond (*Commun. Math. Phys.*, 2013). We also show that any quantitative form of that theorem would yield a corresponding quantitative improvement to the Hammersley-Welsh bound.

1 Introduction

Fix $d \geq 2$ and consider the hypercubic lattice \mathbb{Z}^d . A **self-avoiding walk** (SAW) is a simple path in \mathbb{Z}^d , that is, a path that does not visit any vertex more than once. Self-avoiding walk was introduced as a model of a linear polymer in a good solvent by Flory and Orr [7, 19]. The rigorous study of self-avoiding walk leads to many questions that are easy to state but difficult to solve, several of which are still open. See [17, 2] for detailed introductions to the theory.

Write Ω for the set of self-avoiding walks, and let c_n be the number of length-n self-avoiding walks starting at the origin. Hammersley and Morton [9] observed that the sequence $(c_n)_{n\geq 0}$ is **submultiplicative**, meaning that $c_{n+m} \leq c_n c_m$ for every $n, m \geq 0$. It follows by Fekete's Lemma [6] that there exists a constant $\mu_c = \mu_c(d)$, known as the **connective constant** of \mathbb{Z}^d , such that

$$\mu_c^n \le c_n \le \mu_c^{n+o(n)}$$

for every $n \ge 0$. Submultiplicativity arguments alone do not give any control of the subexponential correction to the growth of c_n , and it is a major open problem to determine the true asymptotics of

 c_n when d = 2, 3, 4.

It is believed that in fact the number of self-avoiding walks satisfies

$$c_n \approx \begin{cases} n^{\gamma_2 - 1} \mu_c^n & d = 2\\ n^{\gamma_3 - 1} \mu_c^n & d = 3\\ (\log n)^{1/4} \mu_c^n & d = 4\\ \mu_c^n & d \ge 5. \end{cases}$$
 $n \to \infty$

for some constants γ_2 and γ_3 . For $d \geq 5$ this conjecture was verified in the seminal work of Hara and Slade [12, 11]. Rapid progress is being made on the four-dimensional case of the conjecture, including most notably a proof of the analogous conjecture for four-dimensional weakly self-avoiding walk by Bauerschmidt, Brydges, and Slade [1]. For d = 2, 3 the conjecture is wide open.

The gap between what is conjectured and what is known for d=2 is very large. Nienhuis [18] used non-rigorous Coulomb gas methods to compute that $\gamma_2 = 43/32$. This conjectured value of γ_2 is strongly supported by numerical evidence [16, 13], non-rigorous conformal field theory arguments [4, 5], and by the theory of SLE [15] (see also [8]). In spite of all this, the best rigorous, unconditional estimate on c_n for d=2 was until now the following theorem of Hammersley and Welsh [10]. For $d \geq 3$, a similar stretched exponential bound with a better exponent was proven by Kesten [14], see also [17, Section 3.3].

Theorem 1.1 (Hammersley-Welsh). Let $d \geq 2$. Then

$$c_n \le \exp\left[\sqrt{\frac{2\pi^2 n}{3}} + o(n^{1/2})\right] \mu_c^n = \exp\left[O(n^{1/2})\right] \mu_c^n$$

as $n \to \infty$.

In this note, we prove the following slight improvement to the Hammersley-Welsh bound, which we show to be a corollary to the work of Duminil-Copin and Hammond [3] on sub-ballisticity of the self-avoiding walk.

Theorem 1.2. Let
$$d \geq 2$$
. Then $c_n \leq \exp\left[o(n^{1/2})\right] \mu_c^n$ as $n \to \infty$.

Along the way, we also present a simplified proof of the Hammersley-Welsh bound (with a suboptimal constant in the exponent) that does not rely on either the combinatorial analysis of 'unfolding' or the analysis of integer partitions. We believe that both have been used in all published proofs of Hammersley-Welsh to date. We also believe that this simplified method of proof will be useful for obtaining further improvements to the bound in the future.

Let us briefly discuss the theorem of Duminil-Copin and Hammond. We will not in fact use their main result, but rather an intermediate result of theirs concerning self-avoiding bridges [3, Corollary 2.4 and Theorem 2.5]. A self-avoiding walk ω is a **self-avoiding bridge** (SAB) if the dth coordinate of ω is uniquely minimized by its starting point, and is maximized (not necessarily uniquely) by its endpoint. Let b_n be the number of self-avoiding bridges of length n starting at the

origin. Probabilistically, Theorem 1.3 states that for every fixed $\varepsilon > 0$, a uniformly chosen length n SAB is exponentially unlikely to reach height εn , where the rate of exponential decay depends on ε (in some unknown way) but is always positive. Note that a length n SAB cannot reach height greater than n. We write $\omega : v \to A$ to mean that ω starts at the vertex v and ends in the set A.

Theorem 1.3 (Duminil-Copin and Hammond). Let $d \geq 2$. There exists an increasing¹ function $\phi = \phi_d : (0,1] \to (0,\infty)$ such that

$$\sum_{\omega \in \Omega} \mathbb{1}\left[\omega : 0 \to \mathbb{Z}^{d-1} \times \{m, m+1, \ldots\} \text{ is a length } n \text{ SAB}\right] \le b_n \exp\left[-\phi\left(\frac{m}{n}\right)n\right]$$
 (1.1)

for all $n \ge m > 0$.

The proof of Theorem 1.3 is not quantitative and does not give any estimates on the function ϕ . However, our derivation of Theorem 1.2 from Theorem 1.3 is quantitative, so that any quantitative version of Theorem 1.3 would yield a quantitative improvement to Hammersley-Welsh. Indeed, suppose that $\phi:(0,1]\to[0,\infty)$ satisfies (1.1). Define $\Phi:(0,1]\to[0,\infty)$ by

$$\Phi(\varepsilon) = \inf_{\varepsilon < \delta < 1} \delta^{-1} \phi(\delta)$$

and define $\psi:(0,1)\to[1,\infty)$ by

$$\psi(\varepsilon) = \sup \left\{ \lambda \ge 1 : \varepsilon \le 1 - \exp \left[-\frac{\Phi(\lambda^{-1})}{\lambda - 1} \right] \right\}.$$

If ϕ is positive and increasing then Φ is positive and increasing also, and it follows that $\psi(\varepsilon) \to \infty$ as $\varepsilon \to 0$. Define $\Psi : \{0, 1, \ldots\} \to [0, \infty)$ by

$$\Psi(n) = \inf_{0 < \varepsilon < 1} \left[2 \left[1 - (1 - \varepsilon)^{\psi(\varepsilon)} \right]^{-1} - (n+1) \log(1 - \varepsilon) \right].$$

The following is a quantitative version of Theorem 1.2.

Theorem 1.4. Let $d \geq 2$. Suppose that $\phi : (0,1] \to [0,\infty)$ is increasing and satisfies (1.1), and let Ψ be as above. Then

$$c_n \le \exp\left[\Psi(n) - 2\right] \mu_c^{n+1}$$

for every $n \geq 0$.

To deduce Theorem 1.2 from Theorem 1.4 it suffices to show that $\Psi(n) = o(n^{1/2})$ when ϕ is positive and increasing, which is straightforward. The classical Hammersley-Welsh bound (or rather the variant stated in Proposition 2.2 below) can also be deduced from Theorem 1.4 and the observation that $\phi \equiv 0$ trivially satisfies (1.1).

¹In [3] it is not stated that the function is increasing. However, if $\phi:(0,1]\to(0,\infty)$ satisfies (1.1), then $\tilde{\phi}(\varepsilon)=\sup\{\phi(\delta):0<\delta\leq\varepsilon\}$ is positive, increasing and also satisfies (1.1).

Finally, we remark that any polynomial estimate on the function ϕ appearing in Theorem 1.3 would yield an improved exponent in the Hammerlsey-Welsh bound. We do not treat values $\nu \leq 1$ as these cannot possibly occur; see Remark 2.5.

Corollary 1.5. Let $d \ge 2$, and suppose that $\phi(\varepsilon) = C\varepsilon^{\nu}$ satisfies (1.1) for some $\nu > 1$ and C > 0. Then

$$c_n \le \exp\left[O\left(n^{(\nu-1)/(2\nu-1)}\right)\right]\mu_c^n.$$

Remark 1.6. The central step in the proof of Theorem 1.4 is to show that the generating function $B(z) = \sum_{n=0}^{\infty} z^n b_n$ satisfies

$$B((1-\varepsilon)z_c) \le \left[1-(1-\varepsilon)^{\psi(\varepsilon)}\right]^{-1}.$$

In particular, we obtain that if $\phi(\varepsilon) = C\varepsilon^{\nu}$ satisfies (1.1) for some C > 0 and $\nu > 1$, then

$$B((1-\varepsilon)z_c) = O\left(\varepsilon^{(1-\nu)/\nu}\right) \quad \text{as } \varepsilon \to 0.$$
 (1.2)

This can be thought of as an inequality between the 'sub-ballisticity exponent' and the bridge counting exponent.

2 Proof

2.1 Proof of Hammersley-Welsh

Fix $d \geq 2$. In this section we give our new proof of the classical Hammersley-Welsh bound. Our starting point will be the following inequality between generating functions due to Madras and Slade [17, eq. 3.1.13]. We define

$$\chi(z) = \sum_{n\geq 0} z^n c_n$$
 and $B(z) = \sum_{n\geq 0} z^n b_n$

to be the generating functions of self-avoiding walks and self-avoiding bridges respectively. We define $z_c = \mu_c^{-1}$, which is the radius of convergence of $\chi(z)$ and hence also of B(z) by the following proposition.

Proposition 2.1 (Madras and Slade). $\chi(z) \leq z^{-1} \exp \left[2B(z) - 2\right]$ for every $z \geq 0$.

This inequality relies on similar ideas as the proof of the Hammersley-Welsh bound, but is easier to prove. It does not rely on the combinatorial analysis of the 'unfolding' of walks. We will prove that the following version of Hammersley-Welsh with a suboptimal constant can be deduced directly from Proposition 2.1 by elementary methods.

Proposition 2.2.
$$c_n \leq \exp \left| \sqrt{8n} + O\left(\sqrt{1/n}\right) \right| \mu_c^{n+1}$$
.

Let $L_n = \mathbb{Z}^{d-1} \times \{n\}$ for each $n \geq 0$. For each $z \geq 0$ and $n \geq 0$, define

$$a(z;n) = \sum_{\omega \in \Omega} z^{|\omega|} \mathbb{1}[\omega : 0 \to L_n \text{ is a SAB}].$$

If $\omega_1: 0 \to L_n$ and $\omega_2: 0 \to L_m$ are bridges, then we can form a bridge $\omega: 0 \to L_{n+m}$ by applying a translation to ω_2 so that it starts at the endpoint of ω_1 and then concatenating the two paths. This implies that the sequence a(z; n) is supermultiplicative for each $z \ge 0$, meaning that

$$a(z; n+m) \ge a(z; n)a(z; m)$$

for every $n, m \ge 0$. It follows by Fekete's lemma that for each $z \ge 0$ there exists $\xi(z)$ such that

$$\xi(z) = \lim_{n \to \infty} -\frac{1}{n} \log a(z; n) \in [-\infty, \infty]$$
(2.1)

and that

$$a(z;n) \le e^{-\xi(z)n} \tag{2.2}$$

for every $n \geq 0$.

Lemma 2.3. $\xi(z_c) \geq 0$.

Proof. For each $n \ge 0$, a(z;n) is expressible as a power series in z with non-negative coefficients, and is therefore left-continuous in z for z > 0. If $z < z_c$ then $B(z) = \sum_{n \ge 0} a(z;n) < \infty$, so that $\xi(z) \ge 0$ by the identity (2.1), and hence that $a(z;n) \le 1$ for every $n \ge 0$ by the inequality (2.2). Since this bound holds for all $z < z_c$, it also holds for $z = z_c$ by left continuity.

A similar analysis shows that $(b_n)_{n\geq 0}$ is supermultiplicative and that

$$b_n \le \mu_c^n \tag{2.3}$$

for every $n \ge 0$ [17, eq. 1.2.17].

Proof of Proposition 2.2. We have the trivial inequality

$$a((1-\varepsilon)z_c;n) \le (1-\varepsilon)^n a(z_c;n) \le (1-\varepsilon)^n.$$
(2.4)

It follows that $B((1-\varepsilon)z_c) \leq \varepsilon^{-1}$, and hence that

$$\chi((1-\varepsilon)z_c) \le \frac{1}{z} \exp\left[2\varepsilon^{-1} - 2\right].$$

To conclude, we apply the trivial inequality

$$z_c^n c_n \le (1 - \varepsilon)^{-n} \chi((1 - \varepsilon) z_c) \tag{2.5}$$

with $\varepsilon = (n/2)^{-1/2}$ to deduce that

$$z_c^{n+1}c_n \le \exp\left[\sqrt{2n} - 2\right] \left(1 - \sqrt{2/n}\right)^{-n-1} = \exp\left[\sqrt{8n} + O(\sqrt{1/n})\right]$$

as $n \to \infty$, where the equality on the right-hand side follows by calculus (the -2 has not been forgotten).

2.2 Proof of the improvement

The main idea behind Theorems 1.2 and 1.4 is that Theorem 1.3 allows us to improve upon the trivial inequality (2.4). This improvement is encapsulated in the following lemma.

Lemma 2.4.
$$\xi((1-\varepsilon)z_c) \ge -\psi(\varepsilon)\log(1-\varepsilon)$$
 for every $\varepsilon > 0$.

Proof. It suffices to prove that

$$\xi((1-\varepsilon)z_c) \ge \min\{-\lambda \log(1-\varepsilon), -\log(1-\varepsilon) + \Phi(\lambda^{-1})\}$$
(2.6)

for every $\varepsilon > 0$ and $\lambda \ge 1$, since the result then follows by optimizing over λ . Splitting the walks contributing to $a((1-\varepsilon)z_c; n)$ according to whether they have length more than λn or not yields that

$$a((1-\varepsilon)z_c;n) \leq \sum_{m \geq \lambda n} \sum_{\omega \in \Omega} z_c^m (1-\varepsilon)^m \mathbb{1} \left[\omega : 0 \to L_n \text{ a length } m \text{ SAB}\right]$$
$$+ \sum_{m=n}^{\lambda n} \sum_{\omega \in \Omega} z_c^m (1-\varepsilon)^m \mathbb{1} \left[\omega : 0 \to L_n \text{ a length } m \text{ SAB}\right]$$

and hence that

$$a((1-\varepsilon)z_c;n) \leq (1-\varepsilon)^{\lambda n} \sum_{m \geq \lambda n} \sum_{\omega \in \Omega} z_c^m \mathbb{1} \left[\omega : 0 \to L_n \text{ a length } m \text{ SAB} \right]$$
$$+ (1-\varepsilon)^n \sum_{m=n}^{\lambda n} \sum_{\omega \in \Omega} z_c^m \mathbb{1} \left[\omega : 0 \to L_n \text{ a length } m \text{ SAB} \right].$$

This implies that

$$a((1-\varepsilon)z_c;n) \le (1-\varepsilon)^{\lambda n} a(z;n) + (1-\varepsilon)^n \sum_{m=n}^{\lambda n} z_c^m b_m \exp\left[-\phi\left(\frac{n}{m}\right)m\right]$$

$$\le (1-\varepsilon)^{\lambda n} + \lambda n(1-\varepsilon)^n \exp\left[-\Phi(\lambda^{-1})n\right]$$

for every $n \geq 0$, where (2.2), Lemma 2.3, (2.3) and the definition of Φ were used in the second inequality. The claimed inequality (2.6) follows.

Proof of Theorem 1.4. Applying (2.6) and (2.2) yields the estimate

$$B((1-\varepsilon)z_c) = \sum_{n\geq 0} a((1-\varepsilon)z_c; n) \leq \sum_{n\geq 0} e^{\psi(\varepsilon)\log(1-\varepsilon)n} = \left[1 - (1-\varepsilon)^{\psi(\varepsilon)}\right]^{-1}$$

We deduce from Proposition 2.1 that

$$\chi((1-\varepsilon)z_c) \le \frac{1}{(1-\varepsilon)z_c} \exp\left[2\left[1-(1-\varepsilon)^{\psi(\varepsilon)}\right]^{-1}-2\right].$$

The claim now follows by applying the trivial inequality (2.5) as in the proof of Proposition 2.2. \square

Proof of Theorem 1.2. Let ϕ be as in Theorem 1.3. It suffices to prove that $\Psi(n) = o(n^{1/2})$. Since $\phi(\varepsilon)$ is increasing and $\phi(\varepsilon) > 0$ for all $0 < \varepsilon \le 1$, we have that $\Phi(\varepsilon) > 0$ for all $0 < \varepsilon \le 1$, and hence that $\psi(\varepsilon) \to \infty$ as $\varepsilon \to 0$. Fix $M \ge 1$ and $\delta > 0$. Then there exists $0 < \varepsilon_0 \le 1$ such that $\psi(\varepsilon) \ge M$ for every $0 < \varepsilon \le \varepsilon_0$. Setting $\varepsilon = \delta n^{-1/2}$ yields that

$$\Psi(n) \le 2 \left[1 - (1 - \delta n^{-1/2})^M \right]^{-1} - (n+1) \log(1 - \delta n^{-1/2})$$

for every sufficiently large n. It follows by calculus that

$$\Psi(n) \le \left(\delta + \frac{2}{\delta M}\right) n^{1/2} + o(n^{1/2})$$

as $n \to \infty$. The result follows since $M \ge 1$ and $\delta > 0$ were arbitrary (e.g. by taking $M = \delta^{-2}$ and sending $\delta \to 0$).

Proof of Corollary 1.5. Suppose that $\phi(\varepsilon) \geq C\varepsilon^{\nu}$ for some C > 0 and $\nu > 1$, and that ϕ satisfies (1.1). Then $\Phi(\varepsilon) \geq C\varepsilon^{\nu-1}$ and hence

$$\psi(\varepsilon) \ge \sup \left\{ \lambda \ge 1 : \varepsilon \le 1 - \exp \left[-C \frac{\lambda^{1-\nu}}{\lambda - 1} \right] \right\}.$$

A straightforward analysis then yields that

$$\psi(\varepsilon) \ge C' \varepsilon^{-1/\nu}$$

for some C' > 0 and every $\varepsilon > 0$ sufficiently small. Let $\alpha > 0$. Then

$$\Psi(n) \le 2 \left[1 - (1 - n^{-\alpha})^{C'n^{\alpha/\nu}} \right]^{-1} - (n+1)\log(1 - n^{-\alpha})$$

for every n sufficiently large. It follows by calculus that

$$\Psi(n) = O\left(n^{\alpha(\nu-1)/\nu} + n^{1-\alpha}\right)$$

as $n \to \infty$. Optimizing by taking $\alpha = \nu/(2\nu - 1)$ and applying Theorem 1.4 yields the claim.

Remark 2.5. We now explain why values of $\nu \leq 1$ are not possible in Corollary 1.5. Suppose that ϕ satisfies (1.1). An inequality of Madras and Slade [17, eq. 3.1.14] (which is a trivial consequence of

Proposition 2.1) states that

$$B((1-\varepsilon)z_c) \ge 1 + \frac{1}{2}\log[(1-\varepsilon)z_c] + \frac{1}{2}\log\varepsilon^{-1},$$

and it follows that

$$\left[1 - (1 - \varepsilon)^{\psi(\varepsilon)}\right]^{-1} \ge 1 + \frac{1}{2}\log[(1 - \varepsilon)z_c] + \frac{1}{2}\log\varepsilon^{-1}.$$

This rearranges to give

$$\psi(\varepsilon) \le \frac{2 + o(1)}{-\varepsilon \log \varepsilon}$$
 and hence $\phi\left(\frac{-\varepsilon \log \varepsilon}{2 + o(1)}\right) \le (1 + o(1))\varepsilon$

as $\varepsilon \to 0$. Probabilistically, this yields a lower bound on the asymptotic probability that an *n*-step SAB has height at least εn .

Acknowledgments

The author was supported by internships at Microsoft Research and a Microsoft Research PhD Fellowship. We thank Omer Angel, Hugo Duminil-Copin, Tyler Helmuth and Gordon Slade for comments on an earlier draft. Finally, we thank the anonymous referee for catching several errors in the preprint.

References

- [1] R. Bauerschmidt, D. C. Brydges, and G. Slade. Logarithmic correction for the susceptibility of the 4-dimensional weakly self-avoiding walk: a renormalisation group analysis. *Comm. Math. Phys.*, 337(2):817–877, 2015.
- [2] R. Bauerschmidt, H. Duminil-Copin, J. Goodman, and G. Slade. Lectures on self-avoiding walks. In *Probability and statistical physics in two and more dimensions*, volume 15 of *Clay Math. Proc.*, pages 395–467. Amer. Math. Soc., Providence, RI, 2012.
- [3] H. Duminil-Copin and A. Hammond. Self-avoiding walk is sub-ballistic. *Comm. Math. Phys.*, 324(2):401–423, 2013.
- [4] B. Duplantier. Intersections of random walks. A direct renormalization approach. *Comm. Math. Phys.*, 117(2):279–329, 1988.
- [5] B. Duplantier. Renormalization and conformal invariance for polymers. In Fundamental problems in statistical mechanics VII (Altenberg, 1989), pages 171–223. North-Holland, Amsterdam, 1990.
- [6] M. Fekete. über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten. Math. Z., 17(1):228–249, 1923.
- [7] P. J. Flory. Principles of polymer chemistry. Cornell University Press, 1953.
- [8] E. Gwynne and J. Miller. Convergence of the self-avoiding walk on random quadrangulations to $sle_{8/3}$ on $\sqrt{8/3}$ liouville quantum gravity, 2016.
- [9] J. M. Hammersley and K. W. Morton. Poor man's Monte Carlo. J. Roy. Statist. Soc. Ser. B., 16:23–38; discussion 61–75, 1954.
- [10] J. M. Hammersley and D. J. A. Welsh. Further results on the rate of convergence to the connective constant of the hypercubical lattice. *Quart. J. Math. Oxford Ser.* (2), 13:108–110, 1962.

- [11] T. Hara and G. Slade. The lace expansion for self-avoiding walk in five or more dimensions. *Rev. Math. Phys.*, 4(2):235–327, 1992.
- [12] T. Hara and G. Slade. Self-avoiding walk in five or more dimensions. I. The critical behaviour. *Comm. Math. Phys.*, 147(1):101–136, 1992.
- [13] I. Jensen. Enumeration of self-avoiding walks on the square lattice. J. Phys. A, 37(21):5503-5524, 2004.
- [14] H. Kesten. On the number of self-avoiding walks. II. J. Mathematical Phys., 5:1128-1137, 1964.
- [15] G. F. Lawler, O. Schramm, and W. Werner. On the scaling limit of planar self-avoiding walk. In *Fractal geometry and applications: a jubilee of Benoît Mandelbrot, Part 2*, volume 72 of *Proc. Sympos. Pure Math.*, pages 339–364. Amer. Math. Soc., Providence, RI, 2004.
- [16] B. Li, N. Madras, and A. D. Sokal. Critical exponents, hyperscaling, and universal amplitude ratios for two- and three-dimensional self-avoiding walks. *J. Statist. Phys.*, 80(3-4):661–754, 1995.
- [17] N. Madras and G. Slade. The self-avoiding walk. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2013. Reprint of the 1993 original.
- [18] B. Nienhuis. Exact critical point and critical exponents of O(n) models in two dimensions. *Phys. Rev. Lett.*, 49(15):1062-1065, 1982.
- [19] W. Orr. Statistical treatment of polymer solutions at infinite dilution. *Transactions of the Faraday Society*, 43:12–27, 1947.