The connective constant of the honeycomb lattice equals $\sqrt{2+\sqrt{2}}$

DUMINIL-COPIN, Hugo, SMIRNOV, Stanislav

Abstract

We provide the first mathematical proof that the connective constant of the hexagonal lattice is equal to $\sqrt{2+\sqrt{2}}$. This value has been derived non rigorously by B. Nienhuis in 1982, using Coulomb gas approach from theoretical physics. Our proof uses a parafermionic observable for the self avoiding walk, which satisfies a half of the discrete Cauchy-Riemann relations. Establishing the other half of the relations (which conjecturally holds in the scaling limit) would also imply convergence of the self-avoiding walk to SLE(8/3).

Reference

DUMINIL-COPIN, Hugo, SMIRNOV, Stanislav. The connective constant of the honeycomb lattice equals $\sqrt{2+\sqrt{2}}$. [Submitted to:] Annals of Mathematics

arxiv: 1007.0575v1
The connective constant of the honeycomb lattice equals $\sqrt{2 + \sqrt{2}}$

Hugo Duminil-Copin and Stanislav Smirnov

Abstract

We provide the first mathematical proof that the connective constant of the hexagonal lattice is equal to $\sqrt{2 + \sqrt{2}}$. This value has been derived non rigorously by B. Nienhuis in 1982, using Coulomb gas approach from theoretical physics. Our proof uses a parafermionic observable for the self avoiding walk, which satisfies a half of the discrete Cauchy-Riemann relations. Establishing the other half of the relations (which conjecturally holds in the scaling limit) would also imply convergence of the self-avoiding walk to SLE(8/3).

Introduction

P. Flory [2] proposed to consider self-avoiding (i.e. visiting every vertex at most once) walks on a lattice as a model for polymer chains. Self-avoiding walks turned out to be a very interesting object, leading to rich mathematical theories and challenging questions, see [3].

Denote by $c_n$ the number of $n$-step self-avoiding walks on the hexagonal lattice $H$ started from some fixed vertex, e.g. the origin. Elementary bounds on $c_n$ (for instance $\sqrt{2}^n \leq c_n \leq 3 \cdot 2^{n-1}$) guarantee that $c_n$ grows exponentially fast. Since a $(n+m)$-step self-avoiding walk can be uniquely cut into a $n$-step self-avoiding walk and a parallel translation of a $m$-step self-avoiding walk, we infer that

$$c_{n+m} \leq c_nc_m,$$

from which it follows that there exists $\mu \in (0, +\infty)$ such that

$$\mu := \lim_{n \to \infty} \frac{1}{n} c_n^\frac{1}{n}.$$

The positive real number $\mu$ is called the connective constant of the hexagonal lattice.

Using Coulomb gas formalism, B. Nienhuis [6] proposed physical arguments for $\mu$ to have the value $\sqrt{2 + \sqrt{2}}$. We rigorously prove this statement. While our methods are different from those harnessed by Nienhuis, they are similarly motivated by considerations of vertex operators in the $O(n)$ model.

Theorem 1 For the hexagonal lattice,

$$\mu = \sqrt{2 + \sqrt{2}}.$$
It will be convenient to consider walks between mid-edges $H$, i.e. centers of edges of $H$ (the set of mid-edges will be called $H$). We will write $\gamma : a \to E$ if a walk $\gamma$ starts at $a$ and ends at some mid-edge of $E \subset H$. In the case $E = \{b\}$, we simply write $\gamma : a \to b$.

The length $\ell(\gamma)$ of the walk is the number of vertices belonging to $\gamma$.

It will be convenient to work with the (decreasing in $x$) sum

$$Z(x) = \sum_{\gamma : a \to H} x^{-\ell(\gamma)} \in (0, +\infty].$$

This sum does not depend on the choice of $a$. Establishing $\mu = \sqrt{2 + \sqrt{2}}$ is equivalent to showing that $Z(x) = +\infty$ for $x < \sqrt{2 + \sqrt{2}}$ and $Z(x) < +\infty$ for $x > \sqrt{2 + \sqrt{2}}$. To this effect, we first restrict walks to bounded domains and weight them counting their winding. The vertex operator obtained leads to a parafermionic observable. Such observables can be used in other contexts, see [1, 7]. To simplify formulæ, below we set $x_c := \sqrt{2 + \sqrt{2}}$ and $j = e^{2\pi/3}$.

**Parafermionic observable** A (hexagonal lattice) domain $\Omega \subset H$ is a union of all mid-edges emanating from a given collection of vertices $V(\Omega)$ (see Fig. 1): a mid-edge $z$ belongs to $\Omega$ if at least one end-point of its associated edge is in $\Omega$, it belongs to $\partial \Omega$ if only one of them is in $\Omega$. We further assume $\Omega$ to be simply connected, i.e. having a connected complement.

![Figure 1](image)

**Definition 1** The winding $W_\gamma(a, b)$ of a self-avoiding walk $\gamma$ between mid-edges $a$ and $b$ (not necessarily the start and the end) is the total rotation of the direction in radians when $\gamma$ is traversed from $a$ to $b$, see Fig. [7].

The parafermionic observable is defined as follows: for $a \in \partial \Omega, z \in \Omega$, set

$$F(z) = F(a, z, x, \sigma) = \sum_{\gamma \subset \Omega : a \to z} e^{-i\sigma W_\gamma(a, z)} x^{-\ell(\gamma)}.$$
Lemma 1 If \( x = x_c \) and \( \sigma = \frac{5}{8} \), then \( F \) satisfies the following relation for every vertex \( v \in V(\Omega) \):

\[
(p - v)F(p) + (q - v)F(q) + (r - v)F(r) = 0,
\]

where \( p, q, r \) are the mid-edges of the three edges adjacent to \( v \).

Note that with \( \sigma = 5/8 \), the term \( e^{i\sigma W_{r(a,z)}} \) gives a weight \( \lambda \) or \( \bar{\lambda} \) per left or right turn of \( \gamma \), where

\[
\lambda = \exp(-i \cdot \frac{5}{8} \cdot \pi) = \exp(-i \frac{5\pi}{24}).
\]

Proof In this proof, we further assume that \( p, q \) and \( r \) are oriented counter-clockwise around \( v \). Note that \( (p - v)F(p) + (q - v)F(q) + (r - v)F(r) \) is a sum of contributions \( c(\gamma) \) over all possible walks \( \gamma \) finishing at \( p, q \) or \( r \). For instance, if the walk ends at the mid-edge \( p \), the contribution will be given by

\[
c(\gamma) = (p - v)e^{i\sigma W_{r(a,p)}}x_c^{-\ell(\gamma)}.
\]

One can partition the set of walks \( \gamma \) finishing at \( p, q \) or \( r \) into pairs and triplets of walks in the following way, see Fig 2:

- If a walk \( \gamma_1 \) visits all three mid-edges \( p, q, r \), it means that the edges belonging to \( \gamma_1 \) form a self-avoiding path plus (up to a half-edge) a self-avoiding loop from \( v \) to \( v \). One can associate to \( \gamma_1 \) the walk passing through the same edges, but exploring the loop from \( v \) to \( v \) in the other direction. Hence, walks visiting the three mid-edges can be grouped in pairs.

- If a walk \( \gamma_1 \) visits only one mid-edge, it can be associated to two walks \( \gamma_2 \) and \( \gamma_3 \) that visit exactly two mid-edges by prolonging the walk one step further (there are two possible choices). The reverse is true: a walk visiting exactly two mid-edges is naturally associated to a walk visiting only one mid-edge by erasing the last step. Hence, walks visiting one or two mid-edges can be grouped in triplets.

If one can prove that the sum of contributions for each pair and each triplet vanishes, then the total sum is zero.

Let \( \gamma_1 \) and \( \gamma_2 \) be two walks that are grouped as in the first case. Without loss of generality, we assume that \( \gamma_1 \) ends at \( q \) and \( \gamma_2 \) ends at \( r \). Note that \( \gamma_1 \) and \( \gamma_2 \) coincide up to the mid-edge \( p \) since \( (\gamma_1, \gamma_2) \) are matched together. We deduce

\[
\ell(\gamma_1) = \ell(\gamma_2) \quad \text{and} \quad \left\{ \begin{array}{l}
W_{\gamma_1(a,q)} = W_{\gamma_1(a,p)} + W_{\gamma_1(p,q)} = W_{\gamma_1(a,p)} + \frac{\pi}{2} \\
W_{\gamma_2(a,r)} = W_{\gamma_2(a,p)} + W_{\gamma_2(p,r)} = W_{\gamma_1(a,p)} + \frac{\pi}{2}.
\end{array} \right.
\]

In order to evaluate the winding of \( \gamma_1 \) between \( p \) and \( q \), we used the fact that \( a \) is on the boundary and \( \Omega \) is simply connected. Therefore,

\[
c(\gamma_1) + c(\gamma_2) = (q - v)e^{i\sigma W_{\gamma_1(a,q)}}x_c^{-\ell(\gamma_1)} + (r - v)e^{i\sigma W_{\gamma_2(a,r)}}x_c^{-\ell(\gamma_2)}
\]

\[
= (p - v)e^{i\sigma W_{\gamma_1(a,p)}}x_c^{-\ell(\gamma_1)} \left( j\lambda^4 + \bar{j}\lambda^4 \right) = 0
\]
where the last equality is due to the chosen value $\lambda = \exp(-i5\pi/24)$.

Let $\gamma_1, \gamma_2, \gamma_3$ be three walks matched as in the second case. Without loss of generality, we assume that $\gamma_1$ ends at $p$ and that $\gamma_2$ and $\gamma_3$ extend $\gamma_1$ to $q$ and $r$ respectively. As before, we easily find that

\[
\ell(\gamma_2) = \ell(\gamma_3) = \ell(\gamma_1) + 1 \quad \text{and} \quad \{ W_{\gamma_2}(a,r) = W_{\gamma_2}(a,p) + W_{\gamma_1}(a,p) - \frac{\pi}{2} \\
W_{\gamma_3}(a,r) = W_{\gamma_3}(a,p) + W_{\gamma_3}(p,r) - W_{\gamma_1}(a,p) + \frac{\pi}{2} \}.
\]

Following the same steps as above, we obtain

\[
c(\gamma_1) + c(\gamma_2) + c(\gamma_3) = (p - v) e^{-i\sigma} W_{\gamma_1}(a,p) x^{-\ell(\gamma_1)} \left( 1 + x^{-1} j \lambda + x^{-1} j \bar{\lambda} \right) = 0.
\]

Here is the only place where we use the crucial fact that $x_c = \sqrt{2 + \sqrt{2}} = 2 \cos \frac{\pi}{8}$.

The claim follows readily by summing over all pairs and triplets.

\[
\begin{align*}
\gamma_1 & \quad \gamma_2 \\
\gamma_1 & \quad \gamma_3
\end{align*}
\]

Figure 2: Left: a pair of walks visiting the three mid-edges and matched together. Right: a triplet of walks, one visiting one mid-edge, the two others visiting two mid-edges, which are matched together.

Remark 1 Coefficients above are three cube roots of unity multiplied by $p - v$, so that the left-hand side can be seen as a discrete integral along an elementary contour on the dual lattice. The fact that the integral of the parafermionic observable along discrete contours vanishes is a glimpse of conformal invariance of the model. Indeed, this observable should converge, when properly rescaled, to a holomorphic martingale, as explained in [7]. Establishing this convergence would pave the way for proving that the self-avoiding walk converges to Schramm’s SLE(8/3) in the scaling limit.

Counting argument in a strip domain. We consider a vertical strip domain $S_T$ composed of $T$ strips of hexagons, and its finite version $S_{T,L}$ cut at height $L$ at an angle of $\pi/3$, see Fig. 3. Namely, position a hexagonal lattice $\mathbb{H}$ of meshsize 1 in $\mathbb{C}$ so that there exists a horizontal edge $e$ with mid-edge $a$ being 0. Then

\[
V(S_T) = \{ z \in V(\mathbb{H}) : 0 \leq \text{Re}(z) \leq \frac{3T + 1}{2} \},
\]

\[
V(S_{T,L}) = \{ z \in V(S_T) : |\sqrt{3} \text{Im}(z) - \text{Re}(z)| \leq 3L \}.
\]
Denote by $\alpha$ the left boundary of $S_T$, by $\beta$ the right one. Symbols $\varepsilon$ and $\bar{\varepsilon}$ denote the top and bottom boundaries of $S_{T,L}$. Introduce the following positive quantities:

\[
A_{T,L}^x := \sum_{\gamma \subset S_{T,L} : a \rightarrow \alpha \setminus \{a\}} x^{-\ell(\gamma)},
\]
\[
B_{T,L}^x := \sum_{\gamma \subset S_{T,L} : a \rightarrow \beta} x^{-\ell(\gamma)},
\]
\[
E_{T,L}^x := \sum_{\gamma \subset S_{T,L} : a \rightarrow \varepsilon \cup \bar{\varepsilon}} x^{-\ell(\gamma)}.
\]

**Figure 3:** Domain $S_{T,L}$ and boundary parts $\alpha$, $\beta$, $\varepsilon$ and $\bar{\varepsilon}$.

**Lemma 2** When $x = x_c$, we have

\[
1 = c_\alpha A_{T,L}^{x_c} + B_{T,L}^{x_c} + c_\varepsilon E_{T,L}^{x_c},
\]

where $c_\alpha = \cos\left(\frac{3\pi}{8}\right)$ and $c_\varepsilon = \cos\left(\frac{\pi}{4}\right)$.

**Proof** Sum the relation (1) over all vertices in $V(S_{T,L})$. Values at interior half-edges disappear and we arrive at

\[
0 = -\sum_{z \in \alpha} F(z) + \sum_{z \in \beta} F(z) + \sum_{z \in \varepsilon} F(z) + \sum_{z \in \bar{\varepsilon}} F(z).
\]

Using the symmetry of the domain, we deduce $F(z) = \bar{F}(z)$. Observe that the winding of any self-avoiding walk from $a$ to the bottom part of $\alpha$ is $-\pi$ while the winding to the top part is $\pi$. We conclude

\[
\sum_{z \in \alpha} F(z) = F(a) + \sum_{z \in \alpha \setminus \{a\}} F(z) = 1 + \frac{e^{-i\sigma\pi} + e^{i\sigma\pi}}{2} A_{T,L}^x = 1 - \cos\left(\frac{3\pi}{8}\right) A_{T,L}^x = 1 - c_\alpha A_{T,L}^x.
\]

5
Above, we have used the fact that the only walk from \(a\) to \(a\) is of length 0. Similarly, the winding from \(a\) to any half-edge in \(\beta\) (resp. \(\varepsilon\) and \(\bar{\varepsilon}\)) is 0 (resp. \(\frac{2\pi}{T}\) and \(-\frac{2\pi}{T}\)), therefore

\[
\sum_{z \in \beta} F(z) = B^x_{T,L} \quad \text{and} \quad j \sum_{z \notin \beta} F(z) + \bar{j} \sum_{z \notin \bar{\beta}} F(z) = \cos \left( \frac{\pi}{4} \right) E^x_{T,L} = c_\varepsilon E^x_{T,L}.
\]

The lemma follows readily by plugging these three formulæ in \((3)\). \(\square\)

Observe that sequences \((A^x_{T,L})_{L>0}\) and \((B^x_{T,L})_{L>0}\) are increasing in \(L\) and are bounded for \(x \geq x_c\) thanks to \((2)\) and the monotonicity in \(x\). Thus they have limits

\[
A_T^x = \lim_{L \to \infty} A^x_{T,L} = \sum_{\gamma \subset S_T: \ a \to \alpha \in \{a\}} x^{-\ell(\gamma)},
\]

\[
B_T^x = \lim_{L \to \infty} B^x_{T,L} = \sum_{\gamma \subset S_T: \ a \to \beta} x^{-\ell(\gamma)}.
\]

When \(x = x_c\), via \((2)\) again, we conclude that \((E^x_{T,L})_{L>0}\) decreases and converges to a limit \(E^x_T = \lim_{L \to \infty} E^x_{T,L}\). Then, \((2)\) implies

\[
1 = c_\alpha A^x_T + B^x_T + c_\varepsilon E^x_T.
\]

**Proof of Theorem** Let us first prove that \(Z(x_c) = +\infty\), which implies \(\mu \leq \sqrt{2 + \sqrt{2}}\).

Suppose that for some \(T\), \(E^x_T > 0\). As noted before, \(E^x_{T,L}\) decreases in \(L\) and

\[
Z(x_c) \geq \sum_{L>0} E^x_{T,L} \geq \sum_{L>0} E^x_{T} = +\infty,
\]

which completes the proof. Assume on the contrary that \(E^x_T = 0\), then \((4)\) simplifies to

\[
1 = c_\alpha A^x_T + B^x_T.
\]

Observe that walks entering into the count of \(A^x_{T+1,L}\) and not in \(A^x_T\) have to visit some vertex adjacent to \(\beta\) for \(S_{T+1}\). Cutting such a walk at the first such point (and adding half-edges to the two halves), we obtain two walks of width \(T + 1\) in \(S_{T+1}\). We conclude that

\[
A^x_{T+1} - A^x_T \leq \frac{1}{x_c} (B^x_{T+1})^2.
\]

Combining \((5)\) for \(T\) and \(T + 1\) with \((6)\), we can write

\[
0 = 1 - 1 = (c_\alpha A^x_{T+1} + B^x_{T+1}) - (c_\alpha A^x_T + B^x_T)
\]

\[
= c_\alpha (A^x_{T+1} - A^x_T) + B^x_{T+1} - B^x_T
\]

\[
\leq \frac{c_\alpha}{x_c} (B^x_{T+1})^2 + B^x_{T+1} - B^x_T,
\]

so

\[
\frac{c_\alpha}{x_c} (B^x_{T+1})^2 + B^x_{T+1} \geq B^x_T.
\]
By induction, it is easy to check that

$$B_T^{ce} \geq \frac{\min(B_1^{xe}, x_c/c_\alpha)}{T}$$

for every $T \geq 1$, implying

$$Z(x_c) \geq \sum_{T>0} B_T^{ce} = +\infty.$$ 

This completes the proof of the inequality $\mu \leq x_c = \sqrt{2 + \sqrt{2}}$.

Let us turn to the other needed inequality $\mu \geq x_c$. An excursion of width $T$ is a self-avoiding walk in $S_T$ from one side to the opposite side, defined up to vertical translation. The partition function of excursions of width $T$ is $B_T^{xe}$. Using (4), we can bound $B_T^{ce}$ by 1. Noting that an excursion of width $T$ has length at least $T$, we obtain for $x > x_c$

$$B_T^{xe} \leq \left(\frac{x_c}{x}\right)^T, \quad B_T^{xe} \leq \left(\frac{x_c}{x}\right)^T.$$ 

Thus, the series $\sum_{T>0} B_T^{xe}$ converges and so does the product $\prod_{T>0} (1 + B_T^{xe})$. Let us assume the following fact: any self-avoiding walk can be canonically decomposed into a sequence of excursions of widths $T_i < \cdots < T_1$ and $T_0 > \cdots > T_j$. Furthermore, if one fixes the starting mid-edge and the first vertex visited, the decomposition uniquely determines the walk. This decomposition was first introduced by Hammersley and Welsh in [4] (for a modern treatment, see Section 3.1 of [3]). Applying the decomposition to walks starting at $a$ (the first visited vertex is 0 or -1), we conclude

$$Z(x) \leq 2 \sum_{T_i < \cdots < T_1} \left( \prod_{k=-i}^j B_{T_k}^{xe} \right) = \prod_{T>0} (1 + B_T^{xe})^2 < \infty.$$ 

The factor 2 is due to the fact that there are two possibilities for the first vertex once we fix the starting mid-edge. Therefore, $Z(x) < +\infty$ whenever $x > x_c$ and $\mu \geq x_c = \sqrt{2 + \sqrt{2}}$. To complete the proof of the theorem it only remains to prove that such a decomposition into excursions does exist. Once again, this fact is well-known [3, 4], we include the proof nevertheless.

First assume that $\tilde{\gamma}$ is a half-plane self-avoiding walk, meaning that the start of $\tilde{\gamma}$ has extremal real part: we prove by induction on the width $T_0$ that the walk admits a canonical decomposition into excursions of widths $T_0 > \cdots > T_j$. Without loss of generality, we assume that the start has minimal real part. Out of the vertices having the maximal real part, choose the one visited last, say after $n$ steps. The $n$ first vertices of the walk form an excursion $\tilde{\gamma}_1$ of width $T_0$, which is the first excursion of our decomposition when prolonged to the mid-edge on the right of the last vertex. We forget about the $(n+1)$-th vertex, since there is no ambiguity in its position. The consequent steps form a half-plane walk $\tilde{\gamma}_2$ of width $T_1 < T_0$. Using the induction hypothesis, we know that $\tilde{\gamma}_2$ admits a decomposition into excursions of widths $T_1 > \cdots > T_j$. The decomposition of $\tilde{\gamma}$ is created by adding $\tilde{\gamma}_1$ before the decomposition of $\tilde{\gamma}_2$. 

7
Figure 4: **Left**: Decomposition of a half-plane walk into four excursions with widths $8 > 3 > 1 > 0$. The first excursion corresponds to the maximal excursion containing the origin. Note that the decomposition contains one excursion of width 0. **Right**: The reverse procedure. If the starting mid-edge and the first vertex are fixed, the decomposition is unambiguous.

If the walk is a reverse half-plane self-avoiding walk, meaning that the end has extremal real part, we set the decomposition to be the decomposition of the reverse walk in the reverse order. If $\gamma$ is a self-avoiding walk in the plane, one can cut the trajectory into two pieces $\gamma_1$ and $\gamma_2$: the vertices of $\gamma$ up to the first vertex of maximal real part, and the remaining vertices. The decomposition of $\gamma$ is given by the decomposition of $\gamma_1$ (with widths $T_{-1} < \cdots < T_{-T_j}$) plus the decomposition of $\gamma_2$ (with widths $T_0 > \cdots > T_j$).

Once the starting mid-edge and the first vertex are given, it is easy to check that the decomposition uniquely determines the walk by exhibiting the reverse procedure, see Fig. 4 for the case of half-plane walks.

**Remark 2** The proof provides bounds for the number of excursions from a to the right side of the strip of width $T$, namely,

$$\frac{c}{T} \leq B_T^{x_c} \leq 1.$$  

In sections 3.4.2 and 3.4.3 of [5], precise behaviors are conjectured for the number of self-avoiding walks between two points on the boundary of a domain. It easily implies the following estimate:

$$\sum_{\gamma \in S_T: 0 \rightarrow T+i\eta T} x^{-\ell(\gamma)} \approx T^{-5/4} H(0, 1 + i\eta)^{5/4}$$

where $H$ is the boundary derivative of the Poisson kernel. Integrating with respect to $\eta$, we obtain that $B_T^{x_c}$ should decay as $T^{-1/4}$ when $T$ goes to infinity. Similar (conjectured) asymptotics are available for walks in $S_T$ from 0 to $\eta T$.

**Acknowledgements.** The authors would like to thank G. Slade for useful comments on the manuscript, and G. Lawler for suggesting Remark 2. The authors were supported
by the ANR grant BLAN06-3-134462, the EU Marie-Curie RTN CODY, the ERC AG CONFRA, as well as by the Swiss NSF.

References


Département de Mathématiques
Université de Genève
Genève, Switzerland

E-mail: hugo.duminil@unige.ch ; stanislav.smirnov@unige.ch