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

Hugo Duminil-Copin and Stanislav Smirnov

DRAFT as of June 2, 2010

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 $\mathbb{H}$ started from some fixed vertex, *e.g.* the origin. Elementary bounds on c_n (for instance $\sqrt{2}^n \le c_n \le 3\cdot 2^{n-1}$) 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

$$c_{n+m} \le c_n c_m,$$

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

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

The positive real number μ is called the *connective constant* of the hexagonal lattice.

Using Coulomb gas formalism, B. Nienhuis [4] proposed physical arguments for μ 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* of $\mathbb{H}$, *i.e.* centers of edges of $\mathbb{H}$ (the set of mid-edges will be called H). We will write $\gamma : a \to E$ if a walk γ 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 γ.

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

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

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, 5]. To simplify formulæ, below we set $x_c := \sqrt{2+\sqrt{2}}$ and $j = \mathrm{e}^{\mathrm{i}2\pi/3}$.

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 Ω if at least one end-point of its associated edge is in Ω, it belongs to $\partial\Omega$ if only one of them is in Ω. We further assume Ω to be simply connected, *i.e.* having a connected complement.

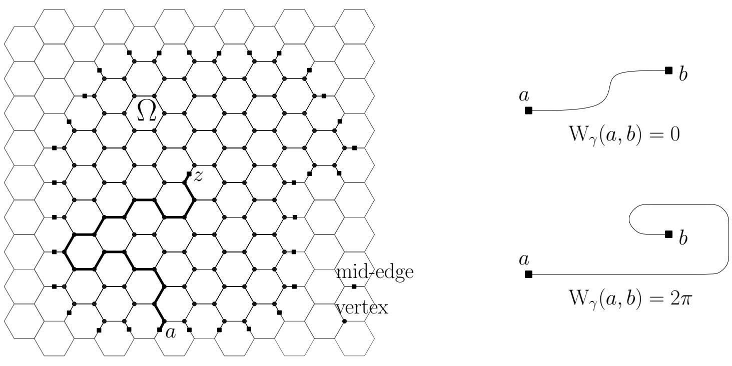

Figure 1: **Left.** A domain Ω whose mid-edges are pictured by small black squares. Vertices of $V(\Omega)$ correspond to circles. **Right.** Winding of a curve γ.

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

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

$$F(z) = F(a,z,x,\sigma) = \sum_{\gamma\subset\Omega:\ a\to z} \mathrm{e}^{-\mathrm{i}\sigma \mathrm{W}_\gamma(a,z)} x^{-\ell(\gamma)}.$$

Lemma 1 *If $x = x_c$ and $\sigma = \frac{5}{8}$, then F satisfies the following relation for every vertex $v \in V(\Omega)$:*

$$(p-v)F(p) + (q-v)F(q) + (r-v)F(r) = 0, \tag{1}$$

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

Note that with $\sigma = 5/8$, the term $\mathrm{e}^{-\mathrm{i}\sigma \mathrm{W}_\gamma(a,z)}$ gives a weight λ or $\bar\lambda$ per left or right turn of γ, where

$$\lambda = \exp(-\mathrm{i}\frac{5}{8}\cdot\frac{\pi}{3}) = \exp(-\mathrm{i}\frac{5\pi}{24}).$$

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 γ finishing at p,q or r. For instance, if the walk ends at the mid-edge p, the contribution will be given by

$$c(\gamma) = (p-v)\mathrm{e}^{-\mathrm{i}\sigma \mathrm{W}_\gamma(a,p)} x_c^{-\ell(\gamma)}.$$

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

- If a walk γ_1 visits all three mid-edges p,q,r, it means that the edges belonging to γ_1 form a self-avoiding path plus (up to a half-edge) a self-avoiding loop from v to v. One can associate to γ_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 γ_1 visits only one mid-edge, it can be associated to two walks γ_2 and γ_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 γ_1 and γ_2 be two walks that are grouped as in the first case. Without loss of generality, we assume that γ_1 ends at q and γ_2 ends at r. Note that γ_1 and γ_2 coincide up to the mid-edge p since (γ_1,γ_2) are matched together. We deduce

$$\ell(\gamma_1) = \ell(\gamma_2) \qquad \text{and} \qquad \begin{cases} \mathrm{W}_{\gamma_1}(a,q) = \mathrm{W}_{\gamma_1}(a,p) + \mathrm{W}_{\gamma_1}(p,q) = \mathrm{W}_{\gamma_1}(a,p) - \frac{4\pi}{3} \\ \mathrm{W}_{\gamma_2}(a,r) = \mathrm{W}_{\gamma_2}(a,p) + \mathrm{W}_{\gamma_2}(p,r) = \mathrm{W}_{\gamma_1}(a,p) + \frac{4\pi}{3} \end{cases}.$$

In order to evaluate the winding of γ_1 between p and q, we used the fact that a is on the boundary and Ω is simply connected. Therefore,

$$\begin{aligned} c(\gamma_1) + c(\gamma_2) &= (q-v)\mathrm{e}^{-\mathrm{i}\sigma \mathrm{W}_{\gamma_1}(a,q)} x_c^{-\ell(\gamma_1)} + (r-v)\mathrm{e}^{-\mathrm{i}\sigma \mathrm{W}_{\gamma_2}(a,r)} x_c^{-\ell(\gamma_2)} \\ &= (p-v)\mathrm{e}^{-\mathrm{i}\sigma \mathrm{W}_{\gamma_1}(a,p)} x_c^{-\ell(\gamma_1)} \left(j\bar{\lambda}^4 + \bar{j}\lambda^4\right) = 0 \end{aligned}$$

where the last equality is due to the chosen value $\lambda = \exp(-\mathrm{i}5\pi/24)$.

Let $\gamma_1, \gamma_2, \gamma_3$ be three walks matched as in the second case. Without loss of generality, we assume that γ_1 ends at p and that γ_2 and γ_3 extend γ_1 to q and r respectively. As before, we easily find that

$$\ell(\gamma_2) = \ell(\gamma_3) = \ell(\gamma_1) + 1 \qquad \text{and} \qquad \begin{cases} \mathrm{W}_{\gamma_2}(a,r) = \mathrm{W}_{\gamma_2}(a,p) + \mathrm{W}_{\gamma_2}(p,q) = \mathrm{W}_{\gamma_1}(a,p) - \frac{\pi}{3} \\ \mathrm{W}_{\gamma_3}(a,r) = \mathrm{W}_{\gamma_3}(a,p) + \mathrm{W}_{\gamma_3}(p,r) = \mathrm{W}_{\gamma_1}(a,p) + \frac{\pi}{3} \end{cases}.$$

Following the same steps as above, we obtain

$$c(\gamma_1) + c(\gamma_2) + c(\gamma_3) = (p-v)\mathrm{e}^{-\mathrm{i}\sigma \mathrm{W}_{\gamma_1}(a,p)} x_c^{-\ell(\gamma_1)} \left(1 + x_c^{-1} j\bar{\lambda} + x_c^{-1}\bar{j}\lambda\right) = 0.$$

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. □

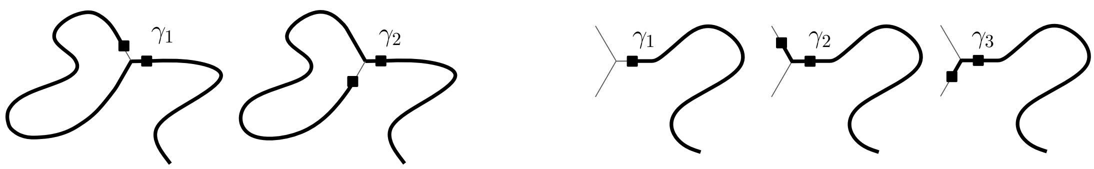

Figure 2: **Left:** a pair of walks visiting the three mid-edges and matched together. **Right:** a triplet of walks, one visiting one mid-edge, the two others visiting two mid-edges, which are matched together.

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 [5]. 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

$$V(S_T) = \{z \in V(\mathbb{H}) : 0 \le \mathrm{Re}(z) \le \frac{3T+1}{2}\},$$
$$V(S_{T,L}) = \{z \in V(S_T) : |\sqrt{3}\mathrm{Im}(z) - \mathrm{Re}(z)| \le 3L\}.$$

Denote by α the left boundary of S_T, by β the right one. Symbols ε and $\bar{\varepsilon}$ denote the top and bottom boundaries of $S_{T,L}$. Introduce the following positive quantities:

$$A^x_{T,L} := \sum_{\gamma \subset S_{T,L}:\ a \to \alpha \setminus \{a\}} x^{-\ell(\gamma)},$$
$$B^x_{T,L} := \sum_{\gamma \subset S_{T,L}:\ a \to \beta} x^{-\ell(\gamma)},$$
$$E^x_{T,L} := \sum_{\gamma \subset S_{T,L}:\ a \to \varepsilon \cup \bar{\varepsilon}} x^{-\ell(\gamma)}.$$

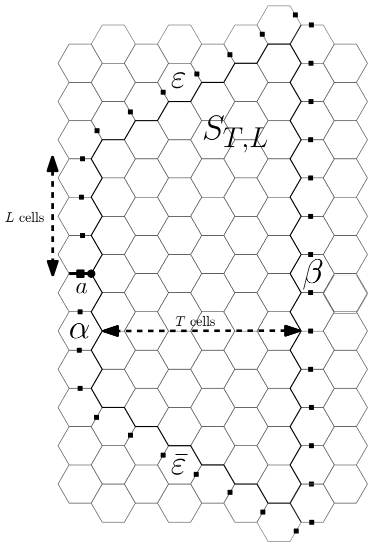

Figure 3: Domain $S_{T,L}$ and boundary parts α, β, ε and $\bar{\varepsilon}$.

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

$$1 = c_\alpha A^{x_c}_{T,L} + B^{x_c}_{T,L} + c_\varepsilon E^{x_c}_{T,L}, \tag{2}$$

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

$$0 = -\sum_{z\in\alpha} F(z) + \sum_{z\in\beta} F(z) + j\sum_{z\in\varepsilon} F(z) + \bar{j}\sum_{z\in\bar{\varepsilon}} F(z). \tag{3}$$

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

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

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 β (resp. ε and $\bar{\varepsilon}$) is 0 (resp. $\frac{2\pi}{3}$ and $-\frac{2\pi}{3}$), therefore

$$\sum_{z\in\beta} F(z) = B_{T,L}^x \quad \text{and} \quad j\sum_{z\in\varepsilon} F(z) + \bar{j}\sum_{z\in\bar{\varepsilon}} F(z) = \cos\left(\frac{\pi}{4}\right) E_{T,L}^x = c_\varepsilon E_{T,L}^x.$$

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

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

$$A_T^x = \lim_{L\to\infty} A_{T,L}^x = \sum_{\gamma\in S_T:\ a\to\alpha\smallsetminus\{a\}} x^{-\ell(\gamma)},$$
$$B_T^x = \lim_{L\to\infty} B_{T,L}^x = \sum_{\gamma\in S_T:\ a\to\beta} x^{-\ell(\gamma)}.$$

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

$$1 = c_\alpha A_T^{x_c} + B_T^{x_c} + c_\varepsilon E_T^{x_c}. \tag{4}$$

Proof of Theorem 1 Let us first prove that $Z(x_c) = +\infty$, which implies $\mu \leq \sqrt{2+\sqrt{2}}$. Suppose that for some T, $E_T^{x_c} > 0$. As noted before, $E_{T,L}^{x_c}$ decreases in L and

$$Z(x_c) \geq \sum_{L>0} E_{T,L}^{x_c} \geq \sum_{L>0} E_T^{x_c} = +\infty,$$

which completes the proof. Assume on the contrary that $E_T^{x_c} = 0$, then (4) simplifies to

$$1 = c_\alpha A_T^{x_c} + B_T^{x_c}. \tag{5}$$

Observe that walks entering into the count of $A_{T+1}^{x_c}$ and not in $A_T^{x_c}$ have to visit some vertex adjacent to β 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

$$A_{T+1}^{x_c} - A_T^{x_c} \leq \frac{1}{x_c}\left(B_{T+1}^{x_c}\right)^2. \tag{6}$$

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

$$\begin{aligned}0 = 1 - 1 &= \left(c_\alpha A_{T+1}^{x_c} + B_{T+1}^{x_c}\right) - \left(c_\alpha A_T^{x_c} + B_T^{x_c}\right)\\ &= c_\alpha\left(A_{T+1}^{x_c} - A_T^{x_c}\right) + B_{T+1}^{x_c} - B_T^{x_c}\\ &\leq \frac{c_\alpha}{x_c}\left(B_{T+1}^{x_c}\right)^2 + B_{T+1}^{x_c} - B_T^{x_c},\end{aligned}$$

so

$$\frac{c_\alpha}{x_c}\left(B_{T+1}^{x_c}\right)^2 + B_{T+1}^{x_c} \geq B_T^{x_c}.$$

By induction, it is easy to check that

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

for every $T \geq 1$, implying

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

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^x. Using (4), we can bound $B_T^{x_c}$ by 1. Noting that an excursion of width T has length at least T, we obtain for $x > x_c$

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

Thus, the series $\sum_{T>0} B_T^x$ converges and so does the product $\prod_{T>0}(1+B_T^x)$. 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. Applying this decomposition to walks starting at a (the first visited vertex is 0 or -1), we conclude

$$Z(x) \leq 2 \sum_{\substack{T_{-i}<\cdots<T_{-1} \\ T_j<\cdots<T_0}} \left(\prod_{k=-i}^{j} B_{T_k}^x\right) = \prod_{T>0}(1+B_T^x)^2 < \infty.$$

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. This fact is well-known, but 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 $\tilde{\gamma}_2$.

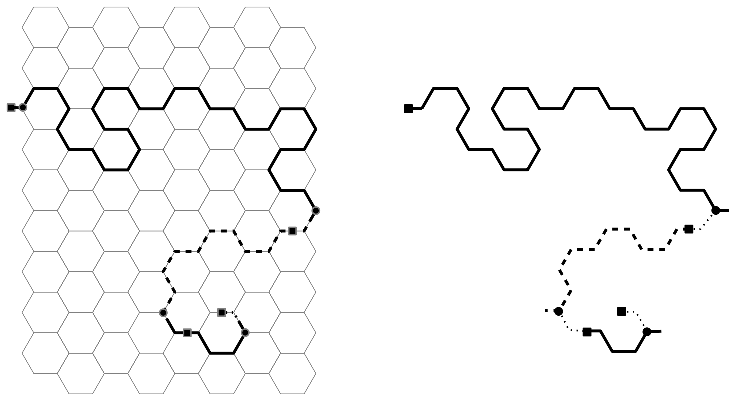

Figure 4: **Left:** Decomposition of a half-plane walk into four excursions with widths $8 > 3 > 1 > 0$. The first excursion corresponds to the maximal excursion containing the origin. Note that the decomposition contains one excursion of width 0. **Right:** The reverse procedure. If the starting mid-edge 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 γ is a self-avoiding walk in the plane, one can cut the trajectory into two pieces γ_1 and γ_2: the vertices of γ up to the first vertex of maximal real part, and the remaining vertices. The decomposition of γ is given by the decomposition of γ_1 (with widths $T_{-i} < \cdots < T_{-1}$) plus the decomposition of γ_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. □

Acknowledgements. 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

[1] J. Cardy and Y. Ikhlef, Discretely holomorphic parafermions and integrable loop models, *J. Phys. A* **42** (2009), *no. 10*, 102001, *11pp.*

[2] P. Flory, Principles of Polymer Chemistry, *Cornell University Press. ISBN: 0-8014-0134-8 (1953)*

[3] N. Madras and G. Slade, Self-avoiding walks, *Probability and its Applications. Birkhäuser Boston, Inc. Boston, MA* (1993).

[4] B. Nienhuis, Exact critical point and critical exponents of O(n) models in two dimensions, *Phys. Rev. Lett.* **49** (1982), 1062-1065.

[5] S. Smirnov, Towards conformal invariance of 2D lattice models *International Congress of Mathematicians, Eur. Math. Soc., Zürich,* **Vol. II** (2006) 1421–1451.

Département de Mathématiques
Université de Genève
Genève, Switzerland
E-mail: hugo.duminil@unige.ch ; stanislav.smirnov@unige.ch