Power series expansions for the planar monomer-dimer problem

Gleb Pogudin*
Institute for Algebra, Johannes Kepler University, Linz, Austria
(Received 22 July 2017; published 7 September 2017)

We compute the free energy of the planar monomer-dimer model. Unlike the classical planar dimer model, an exact solution is not known in this case. Even the computation of the low-density power series expansion requires heavy and nontrivial computations. Despite the exponential computational complexity, we compute almost three times more terms than were previously known. Such an expansion provides both lower and upper bounds for the free energy and makes it possible to obtain more accurate numerical values than previously possible. We expect that our methods can be applied to other similar problems.

DOI: 10.1103/PhysRevE.96.033303

I. INTRODUCTION

The exact solution of the close-packed dimer plane model obtained in [1–3] is a fundamental result in statistical mechanics and combinatorics. In particular, it implies that the number of tilings of an $m \times n$ rectangle using dimers grows as $e^{\frac{G}{\pi}mn}$, where $G \approx 0.916$ is the Catalan constant. Similar results were later obtained for other shapes (see [4] and references therein). Applications to physics suggest two natural further questions: What if the dimension of the lattice is higher (i.e., we compute the number of tilings of a hyperrectangle), and what if we consider tilings using both dimers and monomers with a fixed proportion. For the first question, we refer the reader to [5–7] and references therein. The second questions originates from the study of liquid mixtures on crystal surfaces in [8]; see also [9] for comparison with experimental data. Monomer-dimer systems also arise in connection with the Ising model and the Heisenberg model; see [10, Sec. 5]. For both these questions, the exact solution is out of reach so far. However, even finding the answer numerically leads to very challenging computational problems, because the underlying combinatorial counting problems are very hard, and even a small change of the parameters of the problem makes computations much harder or even unfeasible. For example, in [11] it is proved that the monomer-dimer tilings counting problem is #P complete in the sense of theoretical computer science.

In this paper we will focus on the second question, namely on the case of planar monomer-dimer tilings with fixed dimer density. Let us state the problem precisely. We denote by $a_p(m,n)$ the number of tilings of an $m \times n$ rectangle by monomers and dimers with exactly $\lfloor pmn/2 \rfloor$ dimers, where m,n are positive integers and $p \in [0,1]$. Then p is roughly the fraction of the area covered by dimers. We are interested in the limit

$$f_2(p) = \lim_{n,m\to\infty} \frac{\ln a_p(mn)}{mn}.$$

In other words, we want to determine the constant λ such that $a_p(m,n) \sim e^{\lambda mn}$ as a function of p. From the point of view of statistical mechanics, $f_2(p)$ is equal to the negative of the Helmholtz free energy per lattice site expressed in units of the thermal energy $k_B T$. Some lower and upper bounds for $f_2(p)$ were rigorously proved in [12,13]. However, these bounds are not very tight.

Another approach, taken in a series of papers [14–16] and independently in [17, IV.A], is to expand this function as a power series in p (in some of these papers also expansion with respect to the dimension was discussed). In the former papers, the authors look for a representation of $f_2(p)$ of the form

$$f_2(p) = \frac{1}{2}[(2\ln 2 - 1)p - p\ln p] - (1-p)\ln(1-p) + \sum_{j=2}^{\infty} a_j p^j. \quad (1)$$

In [17], the representation is of the form (see details in Sec. II)

$$f_2(p) = \frac{1}{2}[(2\ln 2 + 1)p - p\ln p] + \sum_{j=2}^{\infty} b_j p^j. \quad (2)$$

Expanding $(1-p)\ln(1-p)$ into a Taylor series in p, it is easy to move back and forth between (1) and (2) [see (18)]. An important observation is that $a_j > 0$ and $b_j < 0$ for all known a_j and b_j. Under the assumption that this pattern holds for all j, truncations of (1) and (2) provide lower and upper bounds for $f_2(p)$, respectively. Thus, computing more terms of these series would result in tighter bounds for $f_2(p)$.

Previously, the record in the number of computed terms in (1) and (2) was 23 (i.e., from a_2 until a_{24}) obtained in [16, Table II]. This result is highly nontrivial, because the underlying combinatorial problem has exponential complexity, so the cost of every next term is usually higher by some factor. In this paper, we compute 63 terms (from a_2 to a_{64}; see Table I), i.e., almost three times more than was previously possible.

The contribution of our paper is twofold. First, our approach allows us to compute significantly more terms for both series (1) and (2) than were previously known and, combining them, we obtain very accurate values of $f_2(p)$. Moreover, we provide additional support for the important conjecture that $a_j > 0$ and $b_j < 0$ for all $j \geq 2$. Second, we show how methods of computer algebra (guessing, modular computation, etc.) can be applied to study models in statistical mechanics. We expect that our methods can be used in other problems of this type

*pogudin.gleb@gmail.com

TABLE I. Values of a_k.

k	a_k	k	a_k	k	a_k
2	$\frac{1}{16}$	23	$\frac{4312434281365}{17803292276948992}$	44	$\frac{1848760120624441058217185 9}{29277281929099243501364287897 6}$
3	$\frac{1}{192}$	24	$\frac{5789230773063}{25895697857380352}$	45	$\frac{74150661042096992710148129}{1225560638892526472150132981760}$
4	$\frac{7}{1536}$	25	$\frac{69044819053441}{337769972052787200}$	46	$\frac{297604910587450946018199331}{5125071762641474338082374287360}$
5	$\frac{41}{10240}$	26	$\frac{272097812497681}{1463669878895411200}$	47	$\frac{1194303993371769853836734501}{21411410919479937234655252578304}$
6	$\frac{181}{61440}$	27	$\frac{1068966474984721}{6323053876828176384}$	48	$\frac{4789513328571295127284133845}{89369367316090172805517575979008}$
7	$\frac{757}{344064}$	28	$\frac{601281977474899}{3891110078048108544}$	49	$\frac{19188774086998950351884051009}{372689276467099444040030742380544}$
8	$\frac{3291}{1835008}$	29	$\frac{16672616519735441}{117021532717594968064}$	50	$\frac{76803645872757902332578961121}{1552871985279581016833461426585600}$
9	$\frac{14689}{9437184}$	30	$\frac{66545602395606901}{501520854503978434560}$	51	$\frac{307176141884436645170078617001}{6465018061163969947633186347417600}$
10	$\frac{64771}{47185920}$	31	$\frac{267471214350929957}{2144433998568735375360}$	52	$\frac{1228026136368811312663436458705}{26894475134442114982154055205257216}$
11	$\frac{276101}{230686720}$	32	$\frac{1080431496491179115}{9149585060559937601536}$	53	$\frac{4909003176336757275553467075425}{111796641735328007376797249088520192}$
12	$\frac{1132693}{1107296256}$	33	$\frac{4374403039126240385}{38959523483674573012992}$	54	$\frac{19627584575160129028816787257753}{464386050285208646026696265444622336}$
13	$\frac{4490513}{5234491392}$	34	$\frac{17705045340400677607}{165577974805616935305216}$	55	$\frac{78505240133588264624896189049521}{1927640208731054757091946762222960640}$
14	$\frac{17337685}{24427626496}$	35	$\frac{71484177460946258777}{702452014326859725537280}$	56	$\frac{314123632091141305526902518303973}{7996137162143634547936964346998947840}$
15	$\frac{65867621}{112742891520}$	36	$\frac{287529593953850293471}{2975090884207876484628480}$	57	$\frac{1257288843192384664389299749835521}{33147623144886339580538688565741092864}$
16	$\frac{249437227}{515396075520}$	37	$\frac{1151710503160001680385}{12580384310364734849286144}$	58	$\frac{1677695623930304081656255827713551}{45775289104843040373124855638404366336}$
17	$\frac{955110593}{2336462209024}$	38	$\frac{4596336312298962012663}{53117178199317769363652608}$	59	$\frac{20147683002193594117896886735926057}{568577275196997764634603470034917392384}$
18	$\frac{3740591431}{10514079940608}$	39	$\frac{18298456303802689186745}{223953508083610054614319104}$	60	$\frac{26879884904186172110556704720248631}{784244517513100365013246165565403299840}$
19	$\frac{15039656569}{47004122087424}$	40	$\frac{72784234597284215364691}{942962139299410756270817280}$	61	$\frac{322682332818808295011085893297500673}{9729948929145584189655867681252122296320}$
20	$\frac{61727254227}{208907209277440}$	41	$\frac{289698911730110389042529}{3965276688335983693036257280}$	62	$\frac{1290942327848947576849492154270349133}{40217122240468414650577586415842105491456}$
21	$\frac{255640084561}{923589767331840}$	42	$\frac{1155125274097244765650075}{16654162091011131510752280576}$	63	$\frac{5163832046366445947035366917883833877}{166142865649148204785992652078560829243392}$
22	$\frac{50273131919}{193514046488576}$	43	$\frac{4616317010648384103125561}{69866240967168649264619323392}$	64	$\frac{9835460990954460589934774119982926 07}{32667107224410092492483962313449748299776}$

(monomer-polymer mixtures, other types of lattices, etc.) in order to push computational limits further.

The rest of the paper is organized as follows. In Sec. II we collect some known results and approaches that connect power series expansion of $f_2(p)$ to the combinatorial data. Section III contains Theorem 1, a main combinatorial ingredient of our computation. Section IV contains the description of our algorithm together with all computer algebra machinery used to speed it up. Finally, in Sec. V we describe our implementation, provide numerical results, and compare them to those of previous work.

II. REDUCTION TO GRAND-CANONICAL PARTITION FUNCTION

For every fixed m and n, consider the *grand-canonical partition function*

$$\Theta_{m,n}(z) = \sum_{s=0}^{\lfloor mn/2 \rfloor} a_{2s/mn}(m,n) z^s. \tag{3}$$

We consider a thermodynamic limit of $\Theta_{m,n}(z)$ (its existence is proved in [10, VIII]),

$$\Theta(z) = \lim_{m,n\to\infty} [\Theta_{m,n}(z)]^{\frac{1}{mn}}, \text{ and } \ln\Theta(z) = \lim_{m,n\to\infty} \frac{\ln\Theta_{m,n}(z)}{mn}. \tag{4}$$

Since $\Theta'_{m,n}(0) = 2mn - m - n$, $\ln\Theta(z) = 2z + O(z^2)$. Theorems [10, 8.8A, 8.8B] rewritten in our notation [i.e., replacing μ with $\ln z$, $g(\mu)$ with $\ln\Theta(z)$, ρ with $\frac{p}{2}$, and $h(\rho)$ with $f_2(p)$] state that

$$f_2(p) = \inf_{z\in\mathbb{R}_+} \left\{ -\frac{p}{2}\ln z + \ln\Theta(z) \right\}. \tag{5}$$

We compute z, where the expression $-\frac{p}{2}\ln z + \ln\Theta(z)$ is minimal,

$$\frac{d}{dz}\left[-\frac{p}{2}\ln z + \ln\Theta(z) \right] = -\frac{p}{2z} + [\ln\Theta(z)]' = 0;$$

hence,

$$p(z) = 2z[\ln\Theta(z)]'. \tag{6}$$

Since $\Theta(z) = 1 + 2z + O(z^2)$, then $2z[\ln\Theta(z)]' = 4z + O(z^2)$. Hence, there exists a unique compositional inverse $z(p)$ of $p(z)$ and $z(p)$ is a formal power series (see [18, Theorem 1.8]). Moreover, $z(p) = \frac{p}{4} + O(p^2)$.

Due to [10, Eq. 8.24], $\ln\Theta(z)$ is a convex function in $\ln z$ for all $z \in \mathbb{R}_+$. Then Eq. (5) can be seen as a fact that $-f_2(p)$ as a function of $\frac{p}{2}$ is a Legendre transform of $\ln\Theta(z)$ as a function of $\ln z$ (for introduction to Legendre transform, see [19] and [20, Sec. 14]). Involutivity of Legendre transform (see [20, Sec. 14.C]) implies that

$$\ln\Theta(z) = \sup_{p\in\mathbb{R}} \left\{ \frac{p}{2}\ln z + f_2(p) \right\}, \tag{7}$$

and the supremum on the right-hand side is reached at $p = p(z)$. On the other hand, we can find this p also by differentiation:

$$\frac{d}{dp}\left[\frac{p}{2}\ln z + f_2(p)\right] = \frac{\ln z}{2} + f_2'(p) = 0.$$

Hence, $f_2'[p(z)] = -\frac{\ln z}{2}$. Substituting $z(p)$ into z, we obtain $f_2'(p) = -\frac{1}{2}\ln z(p)$. Integrating with respect to p, and using the initial condition $f_2(0) = 0$, we conclude that

$$f_2(p) = \frac{1}{2}\int_0^p z(p)dp, \text{ where } p(z) = 2z[\ln \Theta(z)]'. \quad (8)$$

The same formula is obtained in [17, IV.A] using another argument.

Finally, we deduce expansion (2) from (8). Using $z(p) = \frac{p}{4} + O(p^2)$, we conclude that

$$\begin{aligned} f_2(p) &= \frac{1}{2}\int_0^p \ln\left[\frac{p}{4} + O(p^2)\right]dp \\ &= \frac{1}{2}\int_0^p [\ln p - 2\ln 2 + O(p)]dp \\ &= \frac{1}{2}\ln p - \frac{1}{2}p - p\ln 2 + \sum_{k=2}^{\infty} a_k p^k. \end{aligned}$$

The latter expression is exactly of the same form as the right-hand side in (2).

III. COMPUTATION OF $\Theta(z)$ USING $\Theta_{m,n}(z)$

The goal of the present section is to prove the following theorem, which provides a way to compute the thermodynamical limit $\ln\Theta(z)$.

Theorem 1. For every integer $N \geq 4$,

$$\ln\Theta(z) - (S_N - 3S_{N-1} + 3S_{N-2} - S_{N-3}) = O(z^{N-1}), \quad (9)$$

where $S_M = \sum_{m+n=M} \ln\Theta_{m,n}(z)$.

In what follows, we will use some properties of the Mayer expansion following [21, Sec. 2.2].

Let $R_{\infty,\infty}$ be the first quadrant of the plane. We denote the $m \times n$ rectangle whose lower-left corner is the origin by $R_{m,n}$. By definition, $R_{m,n} \subset R_{\infty,\infty}$ for all m and n. We denote the set of all dimers in $R_{m,n}$ by $D_{m,n}$. By definition, the cardinality of $D_{m,n}$ is $2mn - m - n$, and $D_{m,n} \subset D_{\infty,\infty}$ for every $m,n \in \mathbb{Z}_{>0}$. For $d_1,d_2 \in D_{\infty,\infty}$, we introduce $W(d_1,d_2)$ by

$$W(d_1,d_2) = \begin{cases} 1, \text{if } d_1 \text{ and } d_2 \text{ do not overlap,} \\ 0, \text{if } d_1 \text{ and } d_2 \text{ overlap.} \end{cases}$$

Using this notation, the grand-canonical partition function introduced in (3) can be written as (see formula (1.1a) in [21])

$$\Theta_{m,n}(z) = \sum_{s=0}^{\infty}\frac{z^s}{s!}\sum_{(d_1,\ldots,d_s)\in D_{m,n}^s}\left[\prod_{1\leq i<j\leq s} W(d_i,d_j)\right], \quad (10)$$

where $D_{m,n}^s$ stands for the set of all ordered s-tuples of elements of $D_{m,n}$. Unlike (3), formula (10) includes an infinite sum. However, since among all $\lfloor mn/2\rfloor + 1$ dimers there exists at least one pair of overlapping dimers, all terms with $s > \lfloor mn/2\rfloor$ vanish. We introduce $F(d_1,d_2) = W(d_1,d_2) - 1$ for every $d_1,d_2 \in D_{\infty,\infty}$. Then (10) can be rewritten as (see formula (2.7) in [21])

$$\begin{aligned} \Theta_{m,n}(z) = \sum_{s=0}^{\infty}\frac{z^s}{s!}\sum_{\mathbf{d}=(d_1,\ldots,d_s)\in D_{m,n}^s}\ \sum_{G\in\mathcal{G}_s} F(\mathbf{d},G), \text{ where } F(\mathbf{d},G) \\ = \prod_{(ij)\in E(G)} F(d_i,d_j), \quad (11) \end{aligned}$$

and $\mathcal{G}_s$ denotes the set of all graphs on $\{1,\ldots,s\}$, and $E(G)$ is the set of edges of a graph G. Changing the order of summation, we obtain $\Theta_{m,n}(z) = \sum_{s=0}^{\infty}\frac{z^s}{s!}\sum_{G\in\mathcal{G}_s}\mathcal{W}_{m,n}(G)$, where

$$\mathcal{W}_{m,n}(G) = \sum_{\mathbf{d}\in D_{m,n}} F(\mathbf{d},G). \quad (12)$$

In [21, p. 1161] it is shown that [see formula (2.11a)]

$$\ln\Theta_{m,n}(z) = \sum_{s=0}^{\infty}\frac{z^s}{s!}\sum_{G\in\mathcal{C}_s}\mathcal{W}_{m,n}(G), \quad (13)$$

where $\mathcal{C}_s$ is the set of all connected graphs on $\{1,\ldots,s\}$. The above calculations work in quite general context and do not exploit the structure of $D_{m,n}$. Now we will perform a more careful analysis of (12) and (13) in our setting.

For a tuple $\mathbf{d} = (d_1,\ldots,d_s) \in (D_{m,n})^s$, we construct a graph with vertices labeled $1,\ldots,s$ such that there is an edge between i and j if and only if d_i and d_j overlap. We call a tuple $\mathbf{d}$ *connected* if the corresponding graph is connected. The set of connected tuples in $D_{m,n}$ of length s is denoted by $(D_{m,n})_c^s$. For $\mathbf{d} \in (D_{m,n})_c^s$, we define *the height* [*the width*] of $\mathbf{d}$ to be the number of rows [columns] having nontrivial intersection with at least one of the dimers in $\mathbf{d}$. We denote it by $h(\mathbf{d})$ [$w(\mathbf{d})$]. Two tuples $\mathbf{d}_1 = (d_1^1,\ldots,d_s^1)$ and $\mathbf{d}_2 = (d_1^2,\ldots,d_s^2)$ are said to be *translation equivalent* if there exists a translation π of the plane by some vector such that $\pi(d_i^1) = d_i^2$ for every $1 \leq i \leq s$. This is an equivalence relation, and we write it as $\mathbf{d}_1 \sim \mathbf{d}_2$.

The following facts follow straightforwardly from the definitions.

Lemma 1.

(i) For every tuple $(d_1,\ldots,d_s) \in (D_{m,n})^s \setminus (D_{m,n})_c^s$, the corresponding summand in (12) vanishes.

(ii) If $\mathbf{d}_1 = (d_1^1,\ldots,d_s^1)$ and $\mathbf{d}_2 = (d_1^2,\ldots,d_s^2)$ are translation equivalent, then $F(\mathbf{d}_1,G) = F(\mathbf{d}_2,G)$ for every graph $G \in \mathcal{G}_s$.

(iii) For every connected tuple $\mathbf{d} \in (D_{\infty,\infty})_c^s$, the number of tuples $\mathbf{d}' \in (D_{m,n})_c^s$ such that $\mathbf{d} \sim \mathbf{d}'$ is exactly

$$[m - h(\mathbf{d}) + 1]_+[n - w(\mathbf{d}) + 1]_+,$$

where $(x)_+ := \max(x,0)$.

We denote by $\mathcal{T}_s$ a set of tuples in $(D_{\infty,\infty})_c^s$ that contains exactly one representative of every equivalence class of translation-equivalent connected tuples. Due to Lemma 1, we can rewrite (12) as

$$\mathcal{W}_{m,n}(G) = \sum_{\mathbf{d}\in D_{m,n}} F(\mathbf{d},G) \overset{\text{(i)}}{=} \sum_{\mathbf{d}\in (D_{m,n})_c^s} F(\mathbf{d},G)$$
$$\overset{\text{(ii)}}{=} \sum_{\mathbf{d}\in\mathcal{T}_s}\left[\sum_{\mathbf{d}'\in D_{m,n}^s, \mathbf{d}'\sim\mathbf{d}} F(\mathbf{d}',G)\right]$$
$$\overset{\text{(iii)}}{=} \sum_{\mathbf{d}\in\mathcal{T}_s} [m-h(T)+1]_+[n-w(T)+1]_+F(\mathbf{d},G). \tag{14}$$

For $\mathbf{d}\in (D_{m,n})_c^s$, we define $\mathcal{W}(\mathbf{d}) = \sum_{G\in\mathcal{C}_s} F(\mathbf{d},G)$. Using this notation and (14), we can rewrite (13) as

$$\begin{aligned}\ln\Theta_{m,n}(z) &= \sum_{s=0}^{\infty}\frac{z^s}{s!}\sum_{G\in\mathcal{C}_s}\mathcal{W}_{m,n}(G)\\ &= \sum_{s=0}^{\infty}\frac{z^s}{s!}\sum_{G\in\mathcal{C}_s}\Big\{\sum_{\mathbf{d}\in\mathcal{T}_s}[m-h(\mathbf{d})+1]_+\\ &\quad\times[n-w(\mathbf{d})+1]_+F(\mathbf{d},G)\Big\}\\ &= \sum_{s=0}^{\infty}\frac{z^s}{s!}\sum_{\mathbf{d}\in\mathcal{T}_s}[m-h(\mathbf{d})+1]_+\\ &\quad\times[n-w(\mathbf{d})+1]_+\mathcal{W}(\mathbf{d}).\end{aligned}$$

Hence,

$$\ln\Theta_{m,n}(z) = \sum_{s=0}^{\infty}\frac{z^s}{s!}\sum_{\mathbf{d}\in\mathcal{T}_s}[m-h(\mathbf{d})+1]_+[n-w(\mathbf{d})+1]_+ \cdot\mathcal{W}(\mathbf{d}). \tag{15}$$

Now we want to obtain a similar expression for $\ln\Theta(z)$ defined in (4):

$$\ln\Theta(z) = \lim_{m,n\to\infty}\frac{\ln\Theta_{m,n}(z)}{mn} = \sum_{s=0}^{\infty}\frac{z^s}{s!}\sum_{\mathbf{d}\in\mathcal{T}_s}\lim_{m,n\to\infty} \times\left\{\frac{[m-h(\mathbf{d})+1]_+[n-w(\mathbf{d})+1]_+}{mn}\right\}\mathcal{W}(\mathbf{d}).$$

Since $\lim_{m,n\to\infty}\frac{[m-h(\mathbf{d})+1]_+[n-w(\mathbf{d})+1]_+}{mn} = 1$, we obtain

$$\ln\Theta(z) = \sum_{s=0}^{\infty}\frac{z^s}{s!}\sum_{\mathbf{d}\in\mathcal{T}_s}\mathcal{W}(\mathbf{d}). \tag{16}$$

We are now ready to deduce Theorem 1 from (15) and (16).

Lemma 2. For every $N\in\mathbb{Z}_{>0}$,

$$S_N = \sum_{s=0}^{\infty}\frac{z^s}{s!}\sum_{\mathbf{d}\in\mathcal{T}_s}\binom{[N-w(\mathbf{d})-h(\mathbf{d})+3]_+}{3}\mathcal{W}(\mathbf{d}).$$

Proof. By Lemma 1, the coefficient of $\frac{z^s}{s!}\mathcal{W}(\mathbf{d})$ is equal to

$$\sum_{m+n=N}[m-h(\mathbf{d})+1]_+[n-w(\mathbf{d})+1]_+.$$

If $N < p := w(\mathbf{d})+h(\mathbf{d})$, the above expression is equal to $0 = \binom{[N-w(\mathbf{d})-h(\mathbf{d})+3]_+}{3}$. Otherwise, it is equal to

$$\sum_{k=1}^{N-p+1}k(N-p+2-k) = (N-p+2)\left(\sum_{k=1}^{N-p+1}k\right) - \left(\sum_{k=1}^{N-p+1}k^2\right).$$

It can be verified by direct computation using the formula for the sum of squares that the latter expression is equal to $\binom{N-w(\mathbf{d})-h(\mathbf{d})+3}{3}$. This proves the lemma. ■

Fix some $s\le N-2$ and $\mathbf{d}\in\mathcal{T}_s$. We will prove that all summands of the form $\frac{z^s}{s!}\mathcal{W}(\mathbf{d})$ on the left-hand side of (9) cancel. Since $\mathbf{d}$ is connected, it contains at least $w(\mathbf{d})-1$ horizontal dimers and at least $h(\mathbf{d})-1$ vertical dimers. Hence, $w(\mathbf{d})+h(\mathbf{d})-2\le s\le N-2$; so $p := w(\mathbf{d})+h(\mathbf{d})\le N$. This inequality together with Lemma 2 and (16) implies that the coefficient of $\frac{z^s}{s!}\mathcal{W}(\mathbf{d})$ on the left-hand side of (9) is equal to

$$1-\left[\binom{N-p+3}{3}-3\binom{N-p+2}{3} +3\binom{N-p+1}{3}-\binom{N-p}{3}\right].$$

Expanding the brackets, we verify that this expression is zero for every $N-p\ge 0$. This concludes the proof of Theorem 1 .

IV. DESCRIPTION OF THE ALGORITHM

A. General algorithm

Combining (8) and Theorem 1, we obtain Algorithm 1, the first version of an algorithm for computing the first n terms of $f_2(p)$. Note that

(i) line 1 is correct due to Theorem 1;

(ii) line 1 is correct due to (8);

(iii) procedure ComputeTheta is described in Sec. IV B;

(iv) procedure InversePowerSeries$[a(z)]$ computes a power series $z(p)$ given a power series $p(z)$ (see [18, Theorem 1.8]).

Algorithm 1: Nonoptimized version of the algorithm

Input: Nonnegative integer n.
Output: $f_2(p)$ modulo $O(p^n)$.

```
1  for i from 1 to ⌊(n+1)/2⌋ do
2      [Θ_{i,1}(z), ..., Θ_{i,n0+1−i}(z)] := ComputeTheta(i, n0 + 1 − i);
3      for j from 1 to n0 + 1 − i do
4          Θ_{j,i}(z) := Θ_{i,j}(z);
5  for k from 0 to 3 do
6      S_{n+1−k} := Σ_{i+j=n+1−k} ln Θ_{i,j}(z);
7  ln Θ(z) := S_{n+1} − 3S_n + 3S_{n−1} − S_{n−2};
8  z(p) := InversePowerSeries(2z (ln Θ(z))′);
9  f2(p) := −(1/2) ∫ ln z(p) dp;
10 return f2(p);
```

Several improvements can be made:

(i) Computation of ComputeTheta(i,j) deals with a very long vector of possibly very large numbers (see Sec. IV B).

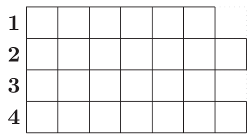

FIG. 1. $F_5^{(6)}$.

In order to fit into the memory, we perform computation modulo several primes and use Chinese remaindering and rational reconstruction to obtain the final result (see Sec. IV D).

(ii) The output of Algorithm 1 with input n [let us call it $\tilde{g}_n(p)$] coincides with $f_2(p)$ only modulo $O(z^n)$. Nevertheless, the first few nonzero coefficients of $f_2(p) - \tilde{g}_n(p)$ turn out to satisfy linear recurrence relations with respect to n, so they can be computed easily. This allows us to "correct" these terms and obtain a more precise result. See Sec. IV C for further details.

(iii) Since we need only the first n terms of $\Theta(z)$, it is sufficient to compute only the first n terms for every computed $\Theta_{i,j}(z)$. Therefore, all intermediate polynomials can also be truncated.

With these improvements, we obtain the final version of our algorithm. For more details, see the source code (see Sec. V A).

B. Computation of $\Theta_{m,n}(z)$

We will compute $\Theta_{m,n}(z)$ using an optimization of *the transfer-matrix method* (see [22, Sec. 4.7]). Fix a positive integer m. Let n be a nonnegative integer, and $0 \leq N < 2^m$. Viewing N as a vector of m bits, we denote the ith bit of N by $N[i]$. We denote by $F_N^{(m,n)}$ the polygon obtained from the $m \times n$ rectangle by adding one additional cell (we will call it *an external* cell) to the end of every row such that $N[i] = 1$, where i is the index of the row. For example, $F_5^{(4,6)}$ is shown in Fig. 1. In particular, $F_0^{(m,n+1)}$ is the same as $F_{2^m-1}^{(m,n)}$.

We introduce polynomial $P_N^{(m,n)}(z)$ to be a generating function for the number of tilings of $F_N^{(m,n)}$ such that every external cell is covered by a horizontal dimer, i.e., $P_N^{(m,n)}(z) = \sum_{j=0}^{m(n+1)} a_{N,j}^{(m,n)} z^j$, where $a_{N,j}^{(m,n)}$ is the number of monomer-dimer tilings of $F_N^{(m,n)}$ with exactly j dimers such that every external cell is covered by a horizontal dimer. We will call such tilings *rigid*. In particular, $\Theta_{m,n}(z) = P_0^{(m,n)}(z)$. We denote by $P^{(m,n)}$ the vector $[P_0^{(m,n)}(z), \ldots, P_{2^m-1}^{(m,n)}(z)]$.

Remark 1. It can be shown (using techniques from [23, Sec. V.6.]) that there exists a matrix M with entries in $\mathbb{Z}[z]$ such that $P^{(m,n+1)} = MP^{(m,n)}$. Hence, $\Theta_{m,n}(z)$ can be computed as the first coordinate of $M^n P^{(m,0)}$. However, in our computations m can be any natural number up to 30, so M can have $2^{30} \times 2^{30} = 2^{60} \approx 10^{18}$ entries. Luckily, the matrix M is highly structured (see [24]), so there exists a faster algorithm for computing $P^{(m,n+1)}$ from $P^{(m,n)}$.

We present an algorithm (Algorithm 2) that computes $P^{(m,n+1)}$ from $P^{(m,n)}$ in place [i.e., with $O(1)$ additional space] using $O(m2^m)$ arithmetic operations. We denote the number of ones in the binary representation of N by BinDig(N).

Algorithm 2: Computing $P^{(m,n+1)}$ from $P^{(m,n)}$.

Input: Vector $P = (P_0^{(m,n)}(z), \ldots, P_{2^m-1}^{(m,n)}(z))$.
Output: Vector $P = (P_0^{(m,n+1)}(z), \ldots, P_{2^m-1}^{(m,n+1)}(z))$.

1 **for** N *from* 0 *to* $2^{m-1} - 1$ **do**
2   Swap values $P[N]$ and $P[2^m - 1 - N]$;
3 **for** j *from* 1 *to* m **do**
4   **for** N *from* 0 *to* $2^m - 1$ **do**
5   **if** $N[j] = 0$ **then**
6          $P[N] \mathrel{+}= P[N + 2^{m-j}]$;
7   **if** $N[j] = 0$ *and* $j > 1$ *and* $N[j-1] = 0$ **then**
8          $P[N] \mathrel{+}= z \cdot P[N + 2^{m-j} + 2^{m-j+1}]$;
9 **for** N *from* 0 *to* $2^m - 1$ **do**
10    $d := \text{BinDig}(N)$;
11    $P[N] := z^d \cdot P[N]$;
12 **return** P;

Proposition 1. Algorithm 2 is correct.

Proof. We will prove by induction on j that after the jth iteration of the loop in line 2 (for $j = 0$ it means the moment just before the first iteration) $\widetilde{P}_N := z^{\text{BinDig}(N)} P[N]$ is the generating polynomial for the number of monomer-dimer tilings of $F_N^{(m,n)}$ satisfying the following A_j property.

A_j property. The tiling is rigid, and the rightmost cell in rows with the number greater than j is covered by a horizontal dimer.

First we prove the base case, where $j = 0$. Due to the loop in line 2, $P[N] = P_{2^m-1-N}^{(m,n)}(z)$. Since the binary representation of $2^m - 1 - N$ can be obtained from the binary representation of N by inverting all m bits, adding a horizontal dimer to the end of every row without an external cell provides us a bijection between the set of rigid tilings of $F_{2^m-1-N}^{(m,n)}$ and the set of tilings of $F_N^{(m,n+1)}$ with A_0 property (see Fig. 2). This map adds BinDig(N) new dimers, so the corresponding generating polynomials differ by the factor $z^{\text{BinDig}(N)}$.

Assume now that $j > 0$. For N such that $N[j] = 1$, properties A_{j-1} and A_j are the same, so the corresponding component of vector P should not be changed. Assume that $N[j] = 0$. We denote the last cell of the jth row in $F_N^{(m,n+1)}$ by c. Consider an arbitrary monomer-dimer tiling of $F_N^{(m,n+1)}$ with property A_j. There are three options for c.

(1) Cell c is covered by a horizontal dimer. Then this tiling has also property A_{j-1} and is already counted in $z^{\text{BinDig}(N)} P[N]$.

(2) Cell c is covered by a monomer. Replacing this monomer with a horizontal dimer, we establish a bijection

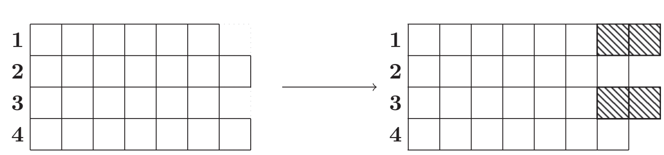

FIG. 2. Tilings of $F_5^{(6)}$ to tilings of $F_{10}^{(7)}$ with property A_0.

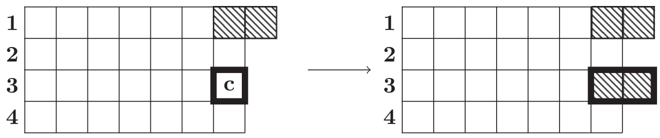

FIG. 3. Tilings of $F_8^{(7)}$ with property A_3 to tilings of $F_{10}^{(7)}$ with property A_2.

between such tilings of $F_N^{(m,n)}$ and tilings of $F_{N+2^{m-j}}^{(m,n+1)}$ with property A_{j-1} (see Fig. 3). Due to the induction hypothesis, the generating polynomial for the latter is $\widetilde{P}_{N+2^{m-j}}$. Hence, in order to take into account tilings where c is covered by a monomer, we should add $\frac{1}{z}\widetilde{P}_{N+2^{m-j}}$ to $\widetilde{P}_N$. This is equivalent to $P[N] \mathrel{+}= P[N+2^{m-j}]$ in line 2.

(3) Cell c is covered by a vertical dimer. This dimer cannot cover also the cell below c due to A_j property. Hence, it covers c and the cell above, say d, so $N[j-1]=0$. Replacing this dimer with two horizontal dimers, we establish a bijection between such tilings of $F_N^{(m,n)}$ and tilings of $F_{N+2^{m-j}+2^{m-j+1}}^{(m,n+1)}$ with property A_{j-1} (see Fig. 4). These cases are counted in line 2.

Since the A_n property is just rigidness, after multiplication by an appropriate degree of z in line 2 we obtain the vector $P^{(m,n+1)}$. ■

Remark 2. Algorithm 2 can be parallelized. Consider an iteration of the loop in line 2 with $j>0$. Then, during the iteration, coordinates of $P[N]$ with different $N[0]$ do not interact, so the whole vector can be divided into two halves (depending on $N[0]$), and these halves can be processed by separate threads. Taking into account $N[1]$, we can divide the work between four threads, and so on. In our computation, we used 32 threads (so we divided the work based on $N[0],\dots,N[4]$).

Finally, using Algorithm 2, we can write a pseudocode for procedure ComputeTheta(m,n); see Algorithm 3.

Algorithm 3: Compute Theta

```
Input: Natural numbers m and n.
Output: Vector of polynomials [Θ_{m,1}(z), ..., Θ_{m,n}(z)].
result := [];
P := zero vector of polynomials in z of length 2^m;
P[0] := 1;
for i from 1 to n do
    Apply Algorithm 2 to P;
    Append P[0] to result;
return result;
```

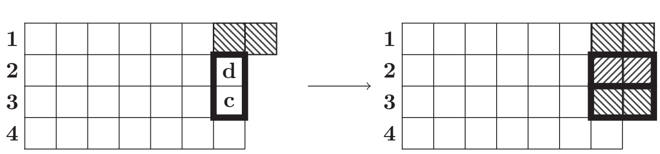

FIG. 4. Tilings of $F_8^{(7)}$ with property A_3 to tilings of $F_{14}^{(7)}$ with property A_2.

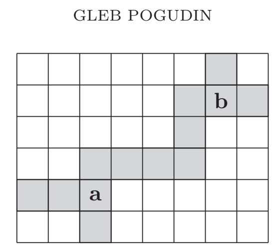

FIG. 5. "Large" and "thin" polyomino for $N=15$.

C. Correction terms

We can compute more terms of $\Theta(z)$ and, consequently, of $f_2(p)$ if we examine carefully the right-hand side of (9). Below we write the first nonzero term of the right-hand side of (9) for $N=4,5,\dots$:

$$\mathbf{11}z^3, \mathbf{-38}z^4, \mathbf{115}z^5, \mathbf{-309}z^6, \mathbf{759}z^7, \mathbf{-1748}z^8, \mathbf{3847}z^9, \mathbf{-8203}z^{10}, \mathbf{17115}z^{11}, \dots.$$

Denote the sequence of coefficients by $\{a_n\}_{n=1}^{\infty}$. Using the GUESS package ([25], for introduction to guessing, see [26, Sec. 4]), we find that this sequence (we computed first 50 terms) satisfies the following recurrence relation:

$$a_{n+5} = -6a_{n+4} - 14a_{n+3} - 16a_{n+2} - 9a_{n+1} - 2a_n. \quad (17)$$

Using (17), we can compute a_n easily, so we get one more correct term of $\Theta(z)$. Instead of giving a rigorous proof of (17), which is long and involved, we would like to explain informally why it is natural to expect such a relation.

Formula (16) shows that the coefficient of z^s in $\Theta(z)$ is a sum of weights of all connected polyominos constructed from s overlapping dimers. On the other hand, the argument after Lemma 2 shows that the coefficient of z^s in

$$S_N - 3S_{N-1} + 3S_{N-2} - S_{N-3}$$

is a sum of weights over all connected polyominos constructed from s overlapping dimers with the sum of height and width at most $N-2$. Hence, the coefficient of z^{N-1} in their difference is a sum of weights of all connected polyominos constructed from N overlapping dimers with the sum of the height and the width exactly $N-1$ (the sum cannot be larger for a connected polyomino). These requirements on a polyomino are quite restrictive, by a combinatorial argument one can see that all such polyominos are "of a similar shape" as those in Fig. 5. More precisely, there exist two cells (a and b in the figure), maybe coinciding, such that each of them is connected to two sides of an $m\times n$ ($m+n=N-1$) rectangle by straight lines, and a and b are connected by a path such that at each step the path becomes closer to b (all such paths have the same length). Counting such polyominos is a standard combinatorial problem (similar counting problems for polyominos are discussed in [22, Sec. 4.7.5]) that is very likely to result in a formula satisfying a linear recurrence.

Moreover, the same argument shows that there also should be a combinatorial description and a similar recurrence for the second nonzero term in the left-hand side of (9), the third, the fourth, and so on. Our data were enough to discover and verify five formulas of this type [from the first until the fifth nonzero term in (9)]. This is the recurrence for the second nonzero

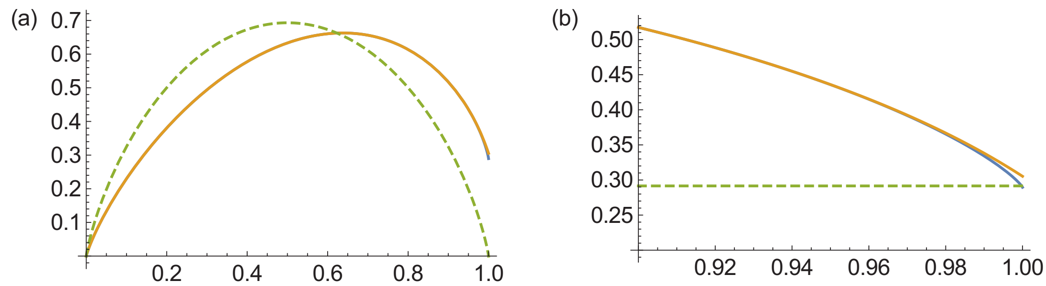

FIG. 6. Plots of $L_{64}(p)$ and $U_{64}(p)$: (a) on $[0,1]$; (b) on $[0.9,1]$.

coefficient,

$$b_{n+7} = -9b_{n+6} - 34b_{n+5} - 70b_{n+4} - 85b_{n+3} - 61b_{n+2} - 24b_{n+1} - 4b_n.$$

We omit the others because they are too large. However, in our program we do not use recurrences themselves, but the closed-form expression for their solutions. This allows us to compute five more terms of $\Theta(z)$ and, consequently, of $f_2(p)$.

D. Modular computation

The largest n we used as an input of the algorithm in our computation was 65. Taking into account correction terms, this means that ComputeTheta is invoked with parameters 30 and 31. Hence, the vector P in Algorithm 2 will have $2^{30} \approx 10^9$ entries. Every entry is a polynomial (in our computations it is a truncated polynomial with only 70 terms); hence, in total we have 7.5×10^{10} integers at every moment. Since these integers represent the number of tilings of a rectangle, they grow fast, so storing them all exactly would require at least several terabytes of memory. However, the final result is a list of coefficients of a power series for $f_2(p)$, that is just 65 rational numbers. A standard way to deal with such a situation (see [26, Sec. 4.2]) is to use computations modulo prime p for intermediate steps. If $p \leq 2^{31} - 1$, then all numbers will fit into 32 bits, and the whole vector P will occupy just 270 GB. Repeating this computation for different primes, we can reconstruct the coefficients of $f_2(p)$ using the Chinese remaindering (see [27, Sec. 5.4]) and the rational reconstruction procedure (see [27, Sec. 5.10]).

The question is how many primes we should take. We start with $2^{31} - 1$ and add new prime numbers until the result of the reconstruction stabilizes. It turned out that 15 prime numbers (from $2^{31} - 1 = 2147483647$ down to 2147483269) are enough; however, we computed several more in order to make sure that the result is correct. The correctness of the result is further justified by the comparison in Sec. V.

V. NUMERICAL RESULTS AND IMPLEMENTATION

A. Implementation

We implemented most of our algorithm in SAGE except the function ComputeTheta, which was implemented in C [28]. Computation modulo one prime with $n = 5$ took about two days using 32 cores and 270 GB of memory. Since we need 15 primes, the whole computation took about one month.

B. Numerical results

Table I contains a_k's [defied in (1)] obtained by our computation. Expanding $(1-p)\ln(1-p)$ into Taylor series at $p = 0$, we obtain the following formula expressing b_k defined in (2) via a_k:

$$b_k = a_k - \frac{1}{k(k-1)}. \tag{18}$$

We introduce following truncated versions of (1) and (2):

$$U_n(p) = \frac{1}{2}[(2\ln 2 + 1)p - p\ln p] + \sum_{j=2}^{n} b_j p^j,$$

$$L_n(p) = \frac{1}{2}[(2\ln 2 - 1)p - p\ln p] - (1-p)\ln(1-p) + \sum_{j=2}^{n} a_j p^j.$$

All computed 63 values a_k are positive; all computed 63 values b_k are negative. Assuming that this pattern persists, we can write

$$L_n(p) \leq f_2(p) \leq U_n(p).$$

This provides us with lower and upper bound for $f_2(p)$. We plot both $L_{64}(p)$ and $U_{64}(p)$ together for $p \in [0,1]$ on Fig. 6(a). The dashed curve on this plot is $-p\ln p - (1-p)\ln(1-p)$, which is the negative value of the free energy for monomer-monomer problem with two different types of monomers. We also plot both $L_{64}(p)$ and $U_{64}(p)$ for $p \in [0.9,1]$ in Fig. 6(b); the dashed line is $y = f_2(1) = \frac{G}{\pi}$.

Plots of $L_{64}(p)$ and $U_{64}(p)$ in Fig. 6(a) are indistinguishable; the difference between them in Fig. 6(b) is visible only very close to $p = 1$. On Fig. 6(b) we also see that the lower bound is much more accurate at $p = 1$. The difference $U_{64}(p) - L_{64}(p)$ does not exceed 2.3×10^{-16} for $p \in [0,0.5]$ and 2.1×10^{-6} for $p \in [0,0.9]$. Note that for $U_{24}(p) - L_{24}(p)$ (these two bounds could be computed using results of [16]) these numbers are 9.3×10^{-11} and 7.5×10^{-4}, respectively, so our bound reduces the error by several orders of magnitude.

C. Comparison with [29]

We already compared our result to the previously known best bound used power series expansion from [16]. However, another method of computing lower and upper bounds for $f_2(p)$ based on the empirically observed inequality [29, Eq. 16]

TABLE II. Comparison with [29]. Digits in square brackets mean the corresponding digit in lower and upper bounds.

p	[29]	Our estimate
10/20	0.633195588930[4 − 5]	**0.6331955889305251415416[5 − 6]**
11/20	0.650499726669[5 − 8]	**0.6504997266695759205[7 − 8]**
12/20	0.66044120984[2 − 5]	**0.66044120984322136[2 − 4]**
13/20	0.6625636470[2 − 4]	**0.66256364703101[3 − 4]**
14/20	0.65620036[0 − 1]	**0.656200361027[4 − 5]**
15/20	0.64039026[3 − 5]	**0.6403902642[8 − 9]**
16/20	0.6137181[3 − 4]	**0.613718137[2 − 7]**
17/20	0.573983[2 − 3]	**0.573983[2 − 3]**
18/20	0.51739[1 − 2]	**0.51739[1 − 3]**
19/20	0.435[8 − 9]	**0.435[8 − 9]**

for strips was proposed in [29]. In this paper bounds for $p = \frac{1}{20}, \ldots, \frac{20}{20}$ were computed (see [29, Table II]). We compare our results with this computation in Table II. The table shows that for p close to 1 Kong's results may be more accurate. On the other hand, our bound is much more precise for $p \leq \frac{17}{20}$.

ACKNOWLEDGMENTS

This work was supported by Austrian Science Fund FWF Grant No. Y464-N18. The author is grateful to Manuel Kauers and Doron Zeilberger for introducing him to the subject and for helpful discussions, as well as to the referees for the comments on the manuscript.

[1] P. W. Kasteleyn, The statistics of dimers on a lattice, Physica **27**, 1209 (1961).
[2] M. E. Fisher, Statistical mechanics of dimers on a plane lattice, Phys. Rev. **124**, 1664 (1961).
[3] H. N. Temperley and M. E. Fisher, Dimer problem in statistical mechanics—An exact result, Philos. Mag. (1798-1977) **6**, 1061 (1961).
[4] H. Cohn, R. Kenyon, and J. Propp, A variational principle for domino tilings, J. Am. Math. Soc. **14**, 297 (2000).
[5] M. Ciucu, An improved upper bound for the 3-dimensional dimer problem, Duke Math. J. **94**, 1 (1998).
[6] D. Gamarnik and D. Katz, Sequential cavity method for computing free energy and surface pressure, J. Stat. Phys. **137**, 205 (2009).
[7] M. Abért, P. Csikvári, and T. Hubai, Matching measure, Benjamini-Schramm convergence and the monomer-dimer free energy, J. Stat. Phys. **161**, 16 (2015).
[8] R. H. Fowler and G. S. Rushbrooke, An attempt to extend the statistical theory of perfect solutions, Trans. Faraday Soc. **33**, 1272 (1937).
[9] D. Everett and M. Penney, The thermodynamics of hydrocarbon solutions. ii. The systems benzene+diphenyl, benzene+diphenylmethane and benzene+dibenzyl, Proc. R. Soc. London A **212**, 164 (1952).
[10] O. J. Heilmann and E. Lieb, Theory of monomer-dimer systems, Commun. Math. Phys. **25**, 190 (1972).
[11] M. Jerrum, Two-dimensional monomer-dimer systems are computationally intractable, J. Stat. Phys. **48**, 121 (1987).
[12] S. Friedland and U. N. Peled, Theory of computation of multidimensional entropy with an application to the monomer-dimer problem, Adv. Appl. Math. **34**, 486 (2005).
[13] S. Friedland, E. Krop, P. H. Lundow, and K. Markström, On the validations of the asymptotic matching conjectures, J. Stat. Phys. **133**, 513 (2008).
[14] P. Federbush, Computation of terms in the asymptotic expansion of dimer λ_d for high dimension, Phys. Lett. A **374**, 131 (2009).
[15] P. Federbush and S. Friedland, An asymptotic expansion and recursive inequalities for the monomer-dimer problem, J. Stat. Phys. **143**, 306 (2011).
[16] P. Butera, P. Federbush, and M. Pernici, Higher-order expansions for the entropy of a dimer or a monomer-dimer system on d-dimensional lattices, Phys. Rev. E **87**, 062113 (2013).
[17] Y. Kong, Asymptotics of the monomer-dimer model on two-dimensional semi-infinite lattices, Phys. Rev. E **75**, 051123 (2007).
[18] S. K. Lando, *Lectures on Generating Functions* (American Mathematical Society, Providence, RI, 2003).
[19] R. Zia, E. F. Redish, and S. R. McKay, Making sense of the Legendre transform, Am. J. Phys. **77**, 614 (2009).
[20] V. I. Arnol'd, *Mathematical Methods of Classical Mechanics*, 2nd ed. (Springer-Verlag, Berlin, 1989).
[21] A. D. Scott and A. D. Sokal, The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma, J. Stat. Phys. **118**, 1151 (2005).
[22] R. P. Stanley, *Enumerative Combinatorics* (Cambridge University Press, Cambridge, U.K., 2012), Vol. 1.
[23] P. Flajolet and R. Sedgewick, *Analytic Combinatorics* (Cambridge University Press, Cambridge, U.K., 2009).
[24] E. H. Lieb, Solution of the dimer problem by the transfer matrix method, J. Math. Phys. **8**, 2339 (1967).
[25] M. Kauers, *Guessing Handbook*, Technical Report (RISC-Linz, Linz, Austria, 2009).
[26] M. Kauers, The holonomic toolkit, in *Computer Algebra in Quantum Field Theory: Integration, Summation and Special Functions* edited by C. Schneider and J. Blümlein (Springer-Verlag, Vienna, Austria, 2013), pp. 119–144.
[27] J. von zur Garthen and J. Gerhard, *Modern Computer Algebra* (Cambridge University Press, Cambridge, U.K., 2013).
[28] See the source code in github.com/pogudingleb/monomer_dimer_tilings.
[29] Y. Kong, Monomer-dimer model in two-dimensional rectangular lattices with fixed dimer density, Phys. Rev. E **74**, 061102 (2006).