Séminaire Lotharingien de Combinatoire **59** (2008), Article B59e

SOLUTION TO A COMBINATORIAL PUZZLE ARISING FROM MAYER'S THEORY OF CLUSTER INTEGRALS

OLIVIER BERNARDI

Abstract. Mayer's theory of cluster integrals allows one to write the partition function of a gas model as a generating function of weighted graphs. Recently, Labelle, Leroux and Ducharme have studied the graph weights arising from the one-dimensional hard-core gas model and noticed that the sum of the weights over all connected graphs with n vertices is $(-n)^{n-1}$. This is, up to sign, the number of rooted Cayley trees on n vertices and the authors asked for a combinatorial explanation. The main goal of this article is to provide such an explanation.

1. Introduction

In [9], Mayer used an algebraic identity in order to express the partition function of a gas model as a generating function of weighted graphs. By Mayer's transformation, any choice of an interaction potential between particles in the gas leads to a specific graph weight. For instance, in the case of the one-dimensional hard-core gas, Labelle, Leroux and Ducharme [6] have shown that the Mayer's weight of a connected graph G having vertex set $V(G) = \{0, \ldots, n\}$ and edge set $E(G)$ is $w(G) = (-1)^{|E(G)|}\mathrm{Vol}(\Pi_G)$ where $\mathrm{Vol}(\Pi_G)$ is the volume of the n-dimensional polytope

$$\Pi_G = \{(x_1, \ldots, x_n) \in \mathbb{R}^n / \ x_0 = 0 \text{ and } |x_i - x_j| \leq 1 \text{ for every edge } (i,j) \in E(G)\}.$$

The pressure in the model is related to Mayer's weights by

$$P \ = \ kT \sum_{G \text{ connected graph}} w(G) \frac{z^{|V(G)|}}{|V(G)|!}, \tag{1}$$

where k is Boltzmann's constant, T is the temperature and z is the *activity*.

It is known (see [3]) that the pressure of the hard-core gas is $P = kTL(z)$, where $L(z)$ is the *Lambert function* defined by the functional equation $L(z) = z\exp(-L(z))$. Comparing this expression of the pressure with (1) and extracting the coefficient of z^{n+1} gives

$$\sum_{G \in \mathcal{C}_n} w(G) \ = \ (-1)^n (n+1)^n, \tag{2}$$

where the sum is over all connected graphs with $n+1$ vertices. Labelle et al. observed that the right-hand-side of (2) is, up to sign, the number of rooted Cayley trees with $n+1$ vertices and asked for a combinatorial explanation [6, Question 1]. The main purpose of this paper is to give such an explanation.

The outline of the paper is as follows. In Section 2, we briefly review Mayer's theory of cluster integrals following the line of [8]. We illustrate this theory on a very simple model of *discrete gas* and prove the equivalence of this model with the Potts model on the complete graph. Comparing two expressions of the pressure in the discrete gas leads to a surprising combinatorial identity. In Section 3, we give a combinatorial proof of this identity. In Section 4, we recall Mayer's setting for the hard-core continuum gas and then give a combinatorial proof of Equation (2), thereby answering the question of Labelle et al.

We close this section with some notations. We denote by $\mathbb{Z}$ the set of integers and by $\mathbb{R}$ the set of real numbers. We denote $[n] = \{1, \ldots, n\}$ and by $\mathfrak{S}_n$ the set of permutations of $[n]$. In this paper, all *graphs* are *simple*, *undirected* and *labelled*. Let G be a graph. We denote by $v(G)$, $e(G)$ and $c(G)$ respectively the number of vertices, edges and connected components of G. A graph H is a *spanning subgraph* of G if the vertex sets of H and G are the same while the edge set of H is included in the edge set of G; we denote $H \subseteq G$ in this case. We denote by $e = (i, j)$ the edge with endpoints i an j and write $e \in G$ if the edge e belongs to G. For any edge e, we denote by $G \oplus e$ the graph obtained from G by either adding the edge e if $e \notin G$ or by deleting this edge if $e \in G$.

2. Review of Mayer's theory of cluster integrals

Consider a gas made of n (indistinguishable) particles in a vessel $\Omega \subset \mathbb{R}^d$. We suppose that the gas is free from outside influence and that interaction between two particles i and j at positions x_i and x_j is given by the potential $\phi(x_i, x_j)$. In the classical Boltzmann setting, the probability measure of a configuration is proportional to $\exp(-H/kT)$, where k is Boltzmann's constant, T is the temperature and H is the *Hamiltonian* of the system given by

$$H = \sum_{1 \le i \le n} \frac{m_i v_i^2}{2} + \sum_{1 \le i < j \le n} \phi(x_i, x_j),$$

where x_i, v_i, m_i and $\frac{m_i v_i^2}{2}$ are respectively the position, velocity, mass and kinetic energy of the i^{th} particle.

The *partition function* of the gas model is

$$Z(\Omega, T, n) = \frac{1}{h^{dn} n!} \iint_{x_1, \ldots, x_n \in \Omega,\ v_1, \ldots, v_n \in \mathbb{R}^d} \exp(-H/kT) dx_1 \ldots dx_n dv_1 \ldots dv_n,$$

where h is Planck's constant. After integrating over all possible velocity, the partition function becomes

$$Z(\Omega, T, n) = \frac{1}{\lambda^n n!} \iint_{\Omega^n} \prod_{i<j} \exp\left(-\frac{\phi(x_i, x_j)}{kT}\right) dx_1 \ldots dx_n,$$

where λ depends on the temperature T.

Mayer noticed that the partition function can be decomposed into a sum over graphs. Indeed, by setting $f(x_i, x_j) = \exp\left(-\frac{\phi(x_i, x_j)}{kT}\right) - 1$, one gets

$$\prod_{i<j} \exp\left(-\frac{\phi(x_i, x_j)}{kT}\right) = \prod_{i<j} 1 + f(x_i, x_j) = \sum_{G \subseteq K_n} \prod_{(i,j) \in G} f(x_i, x_j),$$

where the sum is over all graphs on n vertices (equivalently, spanning subgraph of the complete graph K_n) and the inner product is over all edges of G. In terms of the partition function, this gives *Mayer's relation*:

$$\lambda^n n! Z(\Omega, T, n) \;=\; \iint_{\Omega^n} \prod_{i<j} \exp\left(-\frac{\phi(x_i, x_j)}{kT}\right) dx_1 \dots dx_n \;=\; \sum_{G \subseteq K_n} W(G),$$

where $W(G) = \iint_{\Omega^n} \prod_{(i,j)\in G} f(x_i, x_j) dx_1 \dots dx_n$ is the *first Mayer's weight* of the graph G.

Example: the *discrete gas*. Suppose Ω is made of q distinct boxes $B_1, \dots, B_q$ of volume 1 and that the interaction potential $\phi(x_i, x_j)$ is equal to α if the particles i and j are in the same box and 0 otherwise. By definition, the Mayer's weight of a graph G is

$$W(G) = \iint_{\Omega^n} \prod_{(i,j)\in G} f(x_i, x_j) dx_1 \dots dx_n = \sum_{c:[n]\mapsto[q]} \iint_{x_1\in B_{c(1)},\dots,x_n\in B_{c(n)}} \prod_{i<j} f(x_i, x_j) dx_1 \dots dx_n.$$

We denote $u = \exp(-\alpha/kT)$ and observe that $f(x_i, x_j) = u - 1$ if i and j are in the same box and 0 otherwise. Therefore, the product $\prod_{(i,j)\in G} f(x_i, x_j)$ equals $(u-1)^{e(G)}$ if the value of c is constant over each connected component of the graph G and 0 otherwise. Summing over all mappings $c : [n] \mapsto [q]$ gives

$$W(G) \;=\; q^{c(G)}(u-1)^{e(G)},$$

since there are $q^{c(G)}$ mappings $c : [n] \mapsto [q]$ which are constant over each connected components of G.

In our discrete gas example, a direct calculation of the partition function gives

$$\lambda^n n! Z(\Omega, T, n) = \sum_{c:[n]\mapsto[q]} \iint_{x_1\in B_{c(1)}\dots x_n\in B_{c(n)}} \prod_{i<j} \exp\left(-\frac{\phi(x_i, x_j)}{kT}\right) dx_1 \dots dx_n = \sum_{c:[n]\mapsto[q]} u^{\delta(c)},$$

where $\delta(c)$ is the number of edges $(i,j) \in K_n$ such that $c(i) = c(j)$. Hence, Mayer's relation reads

$$\sum_{c:[n]\mapsto[q]} u^{\delta(c)} = \sum_{G\subseteq K_n} q^{c(G)}(u-1)^{e(G)}. \tag{3}$$

Equation (3) is a special case of the equivalence established by Fortuin and Kasteleyn [4] between the partition function of the Potts model (see e.g. [1]) and the *Tutte polynomial* (see e.g. [2]). Indeed, the right-hand-side corresponds to the partition function of the Potts model on the complete graph K_n while the left-hand-side corresponds to the subgraph expansion of the Tutte polynomial of K_n up to scaling and change of variables. The relation of Fortuin and Kasteleyn is the generalisation of (3) obtained by replacing the complete graph K_n by any graph H. This more general case relies on the observation that $\prod_{(i,j)\in H} \phi_{i,j} = \sum_{G\subseteq H} \prod_{(i,j)\in G} f_{i,j}$ as soon as $\phi_{i,j} = 1 + f_{i,j}$ for all $(i,j) \in H$.

We now return to the general theory of Mayer and consider a system with an arbitrary number of particles. The *grand canonical partition function* is defined by

$$Z_{\mathrm{gr}}(z) \;\equiv\; Z_{\mathrm{gr}}(\Omega,T,z) \;=\; \sum_{n\leq 0} z^n\lambda^n Z(\Omega,T,n),$$

where z is the *activity* of the system. In terms of Mayer's weights, the grand canonical partition function is the exponential generating functions of graphs weighted by their first Mayer's weight:

$$Z_{\mathrm{gr}}(z) \;=\; \sum_{n\leq 0}\frac{z^n}{n!}\sum_{G\subseteq K_n} W(G) \;=\; \sum_G W(G)\frac{z^{v(G)}}{v(G)!}.$$

The macroscopic parameters of the systems, such as the density ρ, or pressure P, can be obtained from $Z_{\mathrm{gr}}(z)$ by the relations

$$P = \frac{kT}{|\Omega|}\log(Z_{\mathrm{gr}}(z)) \quad\text{and}\quad \rho = \frac{z}{|\Omega|}\frac{\partial}{\partial z}\log(Z_{\mathrm{gr}}(z)).$$

Observe that the first Mayer's weight is *multiplicative* over connected components, that is, if a graph G is the disjoint union of two graphs G_1 and G_2 then $W(G) = W(G_1)W(G_2)$. This is the key property implying $\log(Z_{\mathrm{gr}}(z)) = \sum_{G \text{ connected}} W(G)z^{v(G)}$ (see [6] for a complete proof), or equivalently,

$$P = \frac{kT}{|\Omega|}\sum_{G \text{ connected}} W(G)\frac{z^{v(G)}}{v(G)!}\ . \tag{4}$$

Example: the discrete gas. For the discrete gas model introduced before, Equation (4) gives

$$\frac{P}{kT} = \sum_{G \text{ connected}} (u-1)^{e(G)}\frac{z^{v(G)}}{v(G)!}. \tag{5}$$

In the special case of an infinite repulsive interaction between particles in the same box, that is, $\alpha = \infty$ and $u = 0$, the pressure P can also be computed directly. Indeed, in this case, one gets

$$\lambda^n n! Z(\Omega,T,n) \;=\; \sum_{c:[n]\mapsto[q]} u^{\delta(c)} \;=\; \#\{c:[n]\mapsto[q] \text{ injective}\} \;=\; q(q-1)\cdots(q-n+1),$$

and

$$Z_{\mathrm{gr}}(z) \;=\; \sum_{n\geq 0} z^n\lambda^n Z(\Omega,T,n) \;=\; \sum_{n\geq 0}\binom{q}{n}z^n \;=\; (1+z)^q.$$

This expression for the grand canonical partition function comes to no surprise since each of the q boxes contains either nothing (activity 1) or one particle (activity z). Now,

$$\frac{P}{kT} \;=\; \frac{1}{q}\log(Z_{\mathrm{gr}}(z)) \;=\; \log(1+z) \;=\; \sum_{n>0}\frac{(-1)^{n-1}}{n}z^n,$$

and extracting the coefficient of z^n in both side of (5) gives

$$(-1)^{n-1} \sum_{G \subseteq K_n \text{ connected}} (-1)^{e(G)} \; = \; (n-1)!. \tag{6}$$

Identity (6) is quite surprising at first sight but can be understood by recognising in the left-hand-side the evaluation of the Tutte polynomial of K_n counting the root-connected acyclic orientations (acyclic orientation in which the vertex 1 is the only source) [5]. Indeed, in the case of the complete graph K_n, root-connected acyclic orientations are linear orderings of $[n]$ in which 1 is the least element, or equivalently, permutations of $\{2, \ldots, n\}$. In the next section, we give a combinatorial proof of Equation (6) which avoids introducing the whole theory of the Tutte polynomial (though it is based on it) and prepares for the more evolved proof of Equation (2).

3. Pressure in the hard-core discrete gas and increasing trees

In this section, we give a combinatorial proof of (6) by exhibiting an involution Φ on connected graphs which cancels the contribution of almost all graphs in the sum

$$\sum_{G \subseteq K_n \text{ connected}} (-1)^{e(G)}.$$

We consider the *lexicographic order* on the edges of K_n defined by $(i, j) < (k, l)$ if either $\min(i, j) < \min(k, l)$ or $\min(i, j) = \min(k, l)$ and $\max(i, j) < \max(k, l)$. For a graph G and an edge $e = (i, j)$ (not necessarily in G), we denote by $G^{>e}$ the spanning subgraph of G made of the edges which are greater than e. We say that $e = (i, j)$ is *G-active* if there is a path in $G^{>e}$ connecting i and j and we denote by e^*_G the least G-active edge (if there are some). We then define a mapping Ψ on the set of connected graphs by setting: $\Psi(G) = G$ if there is no G-active edge and $\Psi(G) = G \oplus e^*_G$ otherwise.

Lemma 1. *The mapping Ψ is an involution on connected graphs.*

Proof. • First observe that *the image of a connected graph is connected*. Indeed, if the edge e^*_G exists and belongs to G, then it is in a cycle of G and deleting it does not disconnect G.
• We now prove that *any edge is G-active if and only if it is $\Psi(G)$-active*. Suppose that the edge $e = (i, j)$ is G-active and let P be a path of $G^{>e}$ connecting i and j. Since $e^*_G \leq e$ the path P does not contain e^*_G, hence $P \subseteq \Psi(G)^{>e}$ and e is $\Psi(G)$-active. Suppose conversely that $e = (i, j)$ is $\Psi(G)$-active and let P be a path of $\Psi(G)^{>e}$ connecting i and j. If P does not contain e^*_G, then $P \subseteq G^{>e}$ and e is G-active. Otherwise, $e^*_G > e$ and there is a path Q of $G^{>e^*_G} \subseteq G^{>e}$ connecting the endpoints of e^*_G. Thus, there is a path contained in $(P - e^*_G) \cup Q \subseteq G^{>e}$ connecting i and j and again e is G-active.
• By the preceding point, there is a G-active edge if and only if there is a $\Psi(G)$-active edge and in this case $e^*_{\Psi(G)} = e^*_G$. Thus, $\Psi(\Psi(G)) = G$. □

The mapping Ψ is an involution and $(-1)^{e(G)} + (-1)^{e(\Psi(G))} = 0$ whenever $G \neq \Psi(G)$, hence

$$\sum_{G \subseteq K_n \text{ connected}} (-1)^{e(G)} = \sum_{G \subseteq K_n \text{ connected}, \Psi(G) = G} (-1)^{e(G)}. \tag{7}$$

We now characterise the fixed points of the involution Ψ. A tree on $\{1,\dots,n\}$ is said *increasing* if the labels of the vertices are increasing along any simple path starting from the vertex 1.

Lemma 2. *A connected graph G has no G-active edge if and only if it is an increasing tree.*

Proof. • We suppose first that G is an increasing tree and want to prove that no edge is G-active. Since G has no cycle, no edge in G is G-active. Consider now an edge $e=(i,j)\notin G$ and the nearest common ancestor k of i and j (the root vertex of G being the vertex 1). There is an edge $e'=(k,l)$ containing k on the path of G connecting i and j. Since G is an increasing tree, $k\leq \min(i,j)$ and $l\leq \max(i,j)$. Thus, $e'=(k,l)<e=(i,j)$ and e is not G-active.
• Suppose now that there is no G-active edge. First observe that G is a tree since if G had a cycle then the minimal edge in this cycle would be active. We now want to prove that the tree G is increasing. Suppose the contrary and consider a sequence of labels $1=i_1<i_2<\ldots<i_r>i_{r+1}$ on a path of G starting from the vertex $i_1=1$. Then, the edge (i_{r-1},i_{r+1}) is G-active and we reach a contradiction. ☐

By Lemma 2, the fixed points of the involution Ψ are the increasing trees. The increasing trees on $\{1,\dots,n\}$ are known to be in bijection with the permutations of $\{2,\dots,n\}$ [10]. Hence, there are $(n-1)!$ increasing trees on $[n]$ and continuing Equation (7) gives

$$(-1)^{n-1}\sum_{G\subseteq K_n \text{ connected}}(-1)^{e(G)}=\sum_{\substack{G\subseteq K_n \text{ connected}\\ \Psi(G)=G}}(-1)^{e(G)+n-1}=\#\{\text{increasing trees on } [n]\}=(n-1)!.$$

This completes the proof of Equation (6).

4. Pressure in the hard-core continuum gas and Cayley trees

In the 1-dimensional hard-core continuum gas, the vessel is an interval $\Omega=[-q/2,q/2]$ and the potential of interaction between two particles i and j is $\phi(x_i,x_j)=\infty$ if $|x_i-x_j|\leq 1$ and 0 otherwise. By definition, $f(x_i,x_j)\equiv\exp(-\phi(x_i,x_j)/kT)-1$ is equal to -1 if $|x_i-x_j|\leq 1$ and 0 otherwise. Thus, the first Mayer weight of a graph G on n vertices, is

$$W(q,T,G)=\iint_{[-\frac{q}{2},\frac{q}{2}]^n}\prod_{(i,j)\in G}f(x_i,x_j)dx_1\dots dx_n=(-1)^{e(G)}\iint_{[-\frac{q}{2},\frac{q}{2}]^n}\prod_{(i,j)\in G}\mathbb{1}_{|x_i-x_j|\leq 1}.$$

In the thermodynamic limit where the volume q of the Vessel tends to infinity, it becomes interesting to consider the *second Mayer's weight* of connected graphs defined by $w(G)=\lim_{q\to\infty}\frac{W(q,T,G)}{q}$ and related to the pressure by $P=kT\sum_{G \text{ connected}}w(G)z^G$. In [6], it is shown that for any connected graph G on $\{0,\dots,n\}$, the second Mayer weight $w(G)$ equals $(-1)^{e(G)}\mathrm{Vol}(\Pi_G)$, where $\mathrm{Vol}(\Pi_G)$ is the volume of the n-dimensional polytope

$$\Pi_G=\{(x_1,\dots,x_n)\in\mathbb{R}^n/\ x_0=0 \text{ and } |x_i-x_j|\leq 1 \text{ for all edges } (i,j)\in G\}.$$

For instance, the polytope Π_{K_3} is represented in Figure 1. The rest of this paper is devoted to the proof of Equation (2) given in the introduction.

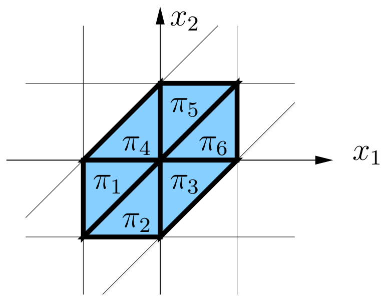

FIGURE 1. The polytope Π_{K_3} (dashed) and its decomposition into subpolytopes.

Subpolytopes. As observed by Bodo Lass [7], it is possible to decompose the polytope Π_G into subpolytopes of volume $1/n!$. Each subpolytope is defined by fixing the integral parts and the relative order of the fractional parts of the coordinates $x_1, \dots, x_n$. Let us first recall some definitions: for any real number x we write $x = h(x) + \epsilon(x)$ where $h(x) \in \mathbb{Z}$ is the *integral part* and $0 \le \epsilon(x) < 1$ is the *fractional part.* Given a vector of integers $\mathbf{h} = (h_1, \dots, h_n) \in \mathbb{Z}^n$ and a permutation $\sigma \in \mathfrak{S}_n$, we denote by $\pi(\mathbf{h}, \sigma)$ the polytope whose interior is made of the points $(x_1, \dots, x_n)$ such that $h(x_i) = h_i$ for all $i = 1 \dots n$ and $0 < \epsilon(x_{\sigma^{-1}(1)}) < \cdots < \epsilon(x_{\sigma^{-1}(n)}) < 1$. In particular, the polytope $\pi(\mathbf{h}, \sigma)$ contains the point $(h_1 + \frac{\sigma(1)}{n+1}, \dots, h_n + \frac{\sigma(n)}{n+1},)$ in its interior.

Observe that the condition $|x_i - x_j| < 1$ is equivalent to $h(x_i) - h(x_j) \in \{0, \operatorname{sign}(\epsilon(x_j) - \epsilon(x_i))\}$ where the value of $\operatorname{sign}(x)$ is -1 if $x < 0$, +1 if $x > 0$ and 0 if $x = 0$. Therefore, a point $(x_1, \dots, x_n)$ in the interior of the polytope $\pi(\mathbf{h}, \sigma)$ is in Π_G if and only if $(i, j) \in G$ implies $h_i - h_j \in \{0, \operatorname{sign}(\sigma(j) - \sigma(i))\}$ with the convention that $h_0 = 0$ and $\sigma(0) = 0$. This condition only depends on the pair $(\mathbf{h}, \sigma)$, therefore either the polytope $\pi(\mathbf{h}, \sigma)$ is included in the polytope Π_G or the interiors of the two polytopes are disjoints. Moreover, a simple calculation shows that the volume of the polytope $\pi(\mathbf{h}, \sigma)$ is $1/n!$ for all $\mathbf{h} \in \mathbb{Z}^n, \sigma \in \mathfrak{S}_n$. This proves the following lemma.

Lemma 3. *For any connected graph G on $\{0, \dots, n\}$, the value $n!\,Vol(\Pi_G)$ counts the pairs $\boldsymbol{h} \in \mathbb{Z}^n, \sigma \in \mathfrak{S}_n$ such that $\pi(\boldsymbol{h}, \sigma)$ is a subpolytope of Π_G.*

For example, the polytope Π_{K_3} represented in Figure 1 contains 6 subpolytopes $\pi_1 = \pi((-1,-1), 12)$, $\pi_2 = \pi((-1,-1), 21)$, $\pi_3 = \pi((0,-1), 12)$, $\pi_4 = \pi((-1,0), 21)$, $\pi_5 = \pi((0,0), 12)$, $\pi_6 = \pi((0,0), 21)$ each having volume 1/2.

Rearrangement. Summing the second Mayer's weights over connected graphs and using Lemma 3 gives

$$\sum_{G \in \mathcal{C}_n} w(G) = \sum_{G \in \mathcal{C}_n} (-1)^{e(G)} \mathrm{Vol}(\Pi_G) = \frac{1}{n!} \sum_{\substack{\mathbf{h} \in \mathbb{Z}^n, \sigma \in \mathfrak{S}_n, G \in \mathcal{C}_n \\ \pi(\mathbf{h}, \sigma) \subseteq \Pi_G}} (-1)^{e(G)}. \tag{8}$$

Let σ be a permutation of $[n]$. For any vector $\mathbf{h} = (h_1, \dots, h_n)$ in $\mathbb{Z}^n$, we denote $\sigma(\mathbf{h}) = (h_{\sigma(1)}, \dots, h_{\sigma(n)})$. For any graph G labelled on $\{0, \dots, n\}$, we denote by $\sigma(G)$ the graph where the label i is replaced by $\sigma(i)$ for all $i = 1, \dots, n$.

Lemma 4. *Let $\boldsymbol{h}$ be a vector in $\mathbb{Z}^n$, let σ be a permutation of $[n]$ and let G be a graph. Then, $\pi(\boldsymbol{h},\sigma) \subseteq \Pi_G$ if and only if $\pi(\sigma^{-1}(\boldsymbol{h}), Id) \subseteq \Pi_{\sigma(G)}$, where Id is the identity permutation.*

We omit the proof of Lemma 4 which is straightforward. From this lemma, one gets for any permutation σ of $[n]$,

$$\sum_{\substack{\mathbf{h}\in\mathbb{Z}^n,\ G\in\mathcal{C}_n\\ \pi(\mathbf{h},\sigma)\subseteq\Pi_G}} (-1)^{e(G)} = \sum_{\substack{\mathbf{h}\in\mathbb{Z}^n,\ G\in\mathcal{C}_n\\ \pi(\sigma^{-1}(\mathbf{h}),\mathrm{Id})\subseteq\Pi_{\sigma(G)}}} (-1)^{e(G)} = \sum_{\substack{\mathbf{h}\in\mathbb{Z}^n,\ G\in\mathcal{C}_n\\ \pi(\mathbf{h},\mathrm{Id})\subseteq\Pi_G}} (-1)^{e(\sigma^{-1}(G))} = \sum_{\substack{\mathbf{h}\in\mathbb{Z}^n,\ G\in\mathcal{C}_n\\ \pi(\mathbf{h},\mathrm{Id})\subseteq\Pi_G}} (-1)^{e(G)}.$$

where the second equality is obtained by changing the order of summations on the graphs G and on the vectors $\mathbf{h}$. Therefore, continuing Equation (8) gives

$$\sum_{G\in\mathcal{C}_n} w(G) \;=\; \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\sum_{\substack{\mathbf{h}\in\mathbb{Z}^n,G\in\mathcal{C}_n\\ \pi(\mathbf{h},\sigma)\subseteq\Pi_G}} (-1)^{e(G)} \;=\; \sum_{\substack{\mathbf{h}\in\mathbb{Z}^n,G\in\mathcal{C}_n\\ \pi(\mathbf{h},\mathrm{Id})\subseteq\Pi_G}} (-1)^{e(G)}. \tag{9}$$

Killing involution. Let $\mathbf{h}$ be a vector in $\mathbb{Z}^n$. We will now evaluate the sum

$$\sum_{\substack{G\in\mathcal{C}_n\\ \pi(\mathbf{h},\mathrm{Id})\subseteq\Pi_G}} (-1)^{e(G)}$$

thanks to a *killing involution* similar to the one defined in Section 3. To the vector $\mathbf{h} = (h_1,\ldots,h_n)$ we associate the *centroid* $\overline{\mathbf{h}} = (\overline{h}_0, \overline{h}_1, \ldots, \overline{h}_n)$, where $\overline{h}_i = h_i + \frac{i}{n+1}$ for $i = 0,\ldots,n$ with the convention that $h_0 = 0$. We also denotes by $G_{\mathbf{h}}$ the graph on $\{0,\ldots,n\}$ whose edges are the pairs (i,j) such that $|\overline{h}_i - \overline{h}_j| < 1$. Observe that for any graph G, $\pi(\mathbf{h},\mathrm{Id}) \subseteq \Pi_G$ if and only if $(\overline{h}_1,\ldots,\overline{h}_n) \in \Pi_G$ if and only if $G \subseteq G_{\mathbf{h}}$.

We order the edges $e = (i,j)$ of the graph $G_{\mathbf{h}}$ by the lexicographic order on the corresponding pairs $(\overline{\mathbf{h}}_i, \overline{\mathbf{h}}_j)$, that is, $(i,j) < (k,l)$ if either $\min(\overline{\mathbf{h}}_i,\overline{\mathbf{h}}_j) < \min(\overline{\mathbf{h}}_k,\overline{\mathbf{h}}_l)$ or $\min(\overline{\mathbf{h}}_i,\overline{\mathbf{h}}_j) = \min(\overline{\mathbf{h}}_k,\overline{\mathbf{h}}_l)$ and $\max(\overline{\mathbf{h}}_i,\overline{\mathbf{h}}_j) < \max(\overline{\mathbf{h}}_k,\overline{\mathbf{h}}_l)$. For a graph $G \subseteq G_{\mathbf{h}}$ and an edge $e = (i,j)$ in $G_{\mathbf{h}}$, we denote by $G^{>e}$ the set of edges in G that are greater than e and we say that e is $(G,\boldsymbol{h})$-*active* if there is a path in $G^{>e}$ connecting i and j. We also denote by $e^*_{G,\mathbf{h}}$ the least $(G,\mathbf{h})$-active edge if there is any. We then define a mapping $\Psi_{\mathbf{h}}$ on the set of connected spanning subgraphs of $G_{\mathbf{h}}$ by: $\Psi_{\mathbf{h}}(G) = G$ if there is no G-active edges and $\Psi_{\mathbf{h}}(G) = G \oplus e^*_{G,\mathbf{h}}$ otherwise.

Lemma 5. *For any vector $\boldsymbol{h}$ in $\mathbb{Z}^n$, the mapping $\Psi_{\boldsymbol{h}}$ is an involution on the connected subgraphs of $G_{\boldsymbol{h}}$.*

The proof of Lemma 5 is identical to the proof of Lemma 1. As a consequence, one gets

$$\sum_{G\subseteq G_{\mathbf{h}} \text{ connected}} (-1)^{e(G)} = \sum_{G\subseteq G_{\mathbf{h}} \text{ connected},\ \Psi_{\mathbf{h}}(G)=G} (-1)^{e(G)}.$$

We now characterise the fixed points of $\Psi_{\mathbf{h}}$. Let i_0 be the index of the least coordinate of the centroid $\overline{\mathbf{h}}$ (that is, $\overline{h}_{i_0} = \min_{i\in\{0,\ldots,n\}}(\overline{h}_i)$). A tree on $\{0,\ldots,n\}$ is said $\boldsymbol{h}$-*increasing* if the labels $i_0, i_1, \ldots, i_k$ of the vertices along any simple path starting at vertex i_0 are such that $\overline{h}_{i_0} < \overline{h}_{i_1} < \cdots < \overline{h}_{i_k}$.

Lemma 6. *Let $\boldsymbol{h}$ be a vector in $\mathbb{Z}^n$. A connected graph $G \subseteq G_{\boldsymbol{h}}$ is an $\boldsymbol{h}$-increasing tree if and only if there is no $(G, \boldsymbol{h})$-active edge.*

Proof. • We suppose first that G is an $\mathbf{h}$-increasing tree. Since G has no cycle, no edge in G is active. Consider now an edge $e = (i,j) \notin G$ and the nearest common ancestor k of i and j (the root vertex of G being the vertex i_0). Let also $e' = (k,l)$ be an edge containing k on the path in G between i and j. Since G is an $\mathbf{h}$-increasing tree, $\overline{\mathbf{h}}_k \leq \min(\overline{\mathbf{h}}_i, \overline{\mathbf{h}}_j)$ and $\overline{\mathbf{h}}_l \leq \max(\overline{\mathbf{h}}_i, \overline{\mathbf{h}}_j)$. Thus, $e' = (k,l) < e = (i,j)$ and e is not $(G, \mathbf{h})$-active.
• Suppose now that there is no $(G, \mathbf{h})$-active edge. First observe that G is a tree since if G had a cycle then the minimal edge in this cycle would be active. We now want to prove that the tree G is $\mathbf{h}$-increasing. Suppose the contrary and consider a sequence of labels $i_0, i_1, \ldots, i_r, i_{r+1}$ such that $\overline{h}_{i_0} < \cdots < \overline{h}_{i_{r-1}} < \overline{h}_{i_r}, \overline{h}_{i_{r+1}}$ on a path of G starting from the vertex i_0. Then, the edge (i_{r-1}, i_{r+1}) belongs to $G_{\mathbf{h}}$ (since $|\overline{h}_{i_{r-1}} - \overline{h}_{i_{r+1}}| < \max(|\overline{h}_{i_{r-1}} - \overline{h}_{i_r}|, |\overline{h}_{i_r} - \overline{h}_{i_{r+1}}|) < 1$) and is $(G, \mathbf{h})$-active. We reach a contradiction. □

From Lemma 6, one gets for any vector $\mathbf{h} \in \mathbb{Z}^n$,

$$\sum_{G \subseteq G_{\mathbf{h}} \text{ connected}} (-1)^{e(G)} \;=\; (-1)^n \, \#\{\mathbf{h}\text{-increasing trees}\}.$$

Therefore, continuing Equation (9) gives

$$\sum_{G \in \mathcal{C}_n} w(G) = \sum_{\mathbf{h} \in \mathbb{Z}^n} \; \sum_{G \subseteq G_{\mathbf{h}} \text{ connected}} (-1)^{e(G)} = (-1)^n \sum_{\mathbf{h} \in \mathbb{Z}^n} \#\{\mathbf{h}\text{-increasing trees}\}. \tag{10}$$

Cayley trees. We now relate $\mathbf{h}$-increasing trees and Cayley trees.

Lemma 7. *Any rooted Cayley tree on $\{0, \ldots, n\}$ with root i_0 is $\boldsymbol{h}$-increasing for exactly one vector $\boldsymbol{h}$ in $\mathbb{Z}^n$ such that $\overline{h}_{i_0} = \min_{i \in \{0,\ldots,n\}} (\overline{h}_i)$.*

Proof. Let T be a Cayley tree rooted on i_0. The tree T is $\mathbf{h}$-increasing with

$$\overline{h}_{i_0} = \min_{i \in \{0,\ldots,n\}} (\overline{h}_i)$$

if and only if any vertex $j \neq i_0$ satisfies $\overline{h}_i < \overline{h}_j < \overline{h}_i + 1$, where i is the father of j. The condition $\overline{h}_i < \overline{h}_j < \overline{h}_i + 1$ holds if and only if either $i < j$ and $h_j = h_i$ or $j < i$ and $h_j = h_i + 1$. Therefore, tree T is $\mathbf{h}$-increasing with $\overline{h}_{i_0} = \min(\overline{h}_i)$ if and only if for every index $i = 0, \ldots, n$, the difference $h_i - h_{i_0}$ is the number of *descents* in the sequence of labels $i_0, i_1, \ldots, i_s = i$ along the path of T from i_0 to i (a descent is an index $r < s$ such that $i_{r+1} < i_r$). Knowing that $h_0 = 0$ completes the proof. □

It is well known that the number of rooted Cayley trees on $\{0, \ldots, n\}$ is $(n+1)^n$. Thus, Lemma 7 gives

$$\sum_{G \in \mathcal{C}_n} w(G) = (-1)^n \sum_{\mathbf{h} \in \mathbb{Z}^n} \#\mathbf{h}\text{-increasing trees} = (-1)^n (n+1)^n. \tag{11}$$

This completes the proof of Equation (2) and answers a question of Labelle et al. [6].

Acknowledgements: I would like to thank Pierre Leroux who warmly encouraged me to work on the combinatorial interpretations of identities arising from Mayer's theory of cluster integrals.

References

[1] R.J. Baxter, *Exactly Solved Models in Statistical Mechanics.* Academic Press (1982).
[2] B. Bollobás, *Modern Graph Theory.* Springer–Verlag (1998).
[3] D.C. Brydge and J.Z. Imbrie, *Dimensional reduction formulas for branched polymer correlation functions.* J. Statist. Phys. **110** 503-518 (2003).
[4] C.M. Fortuin and P.W. Kasteleyn, *On the random cluster model: I. Introduction and relation to other models.* Physica **57** 536-564 (1972).
[5] C. Greene and T. Zaslavsky, *On the interpretation of Whitney numbers through arrangements of hyperplanes, zonotopes, non-Radon partitions, and orientations of graphs.* Trans. Amer. Math. Soc. **280** 97-126 (1983).
[6] G. Labelle, P. Leroux and M.G. Ducharme, *Graph weights arising from Mayer's theory of cluster integrals.* Sém. Lothar. Combin. **54** [B54m] (2007).
[7] B. Lass, Private communication.
[8] P. Leroux, *Enumerative problems inspired by Mayer's theory of cluster integrals.* Electron. J. Combin. **11** [R32] (2006).
[9] J.E, Mayer and M.G. Mayer, *Statistical Mechanics.* Wiley & Sons (1940).
[10] R.P. Stanley, *Enumerative Combinatorics, volume 1.* Wadsworth & Brooks/Cole (1986).