Enumerative problems inspired by Mayer's theory of cluster integrals

Pierre Leroux *

Département de Mathématiques et LaCIM
Université du Québec à Montréal, Canada
leroux@lacim.uqam.ca

Submitted: Jul 31, 2003; Accepted: Apr 20, 2004; Published: May 14, 2004
MR Subject Classifications: 05A15, 05C05, 05C30, 82Axx

Abstract

The basic functional equations for connected and 2-connnected graphs can be traced back to the statistical physicists Mayer and Husimi. They play an essential role in establishing rigorously the virial expansion for imperfect gases. We first review these functional equations, putting the emphasis on the structural relationships between the various classes of graphs. We then investigate the problem of enumerating some classes of connected graphs all of whose 2-connected components (blocks) are contained in a given class B. Included are the species of Husimi graphs ($B =$ "complete graphs"), cacti ($B =$ "unoriented cycles"), and oriented cacti ($B =$ "oriented cycles"). For each of these, we address the question of their labelled and unlabelled enumeration, according (or not) to their block-size distributions. Finally we discuss the molecular expansion of these species. It consists of a descriptive classification of the unlabelled structures in terms of elementary species, from which all their symmetries can be deduced.

1 Introduction

1.1 Functional equations for connected graphs and blocks

Informally, a *combinatorial species* is a class of labelled discrete structures which is closed under isomorphisms induced by relabelling along bijections. See Joyal [13] and Bergeron, Labelle, Leroux [2] for an introduction to the theory of species. Note that the present article is mostly self-contained. To each species F are associated a number of series

*With the partial support of FQRNT (Québec) and CRSNG (Canada)

expansions among which are the following. The (exponential) *generating function*, $F(x)$, for labelled enumeration, is defined by

$$F(x) = \sum_{n\geq 0} |F[n]| \frac{x^n}{n!}, \tag{1}$$

where $|F[n]|$ denotes the number of F-structures on the set $[n] = \{1, 2, \ldots, n\}$. The (ordinary) *tilde generating function* $\widetilde{F}(x)$, for unlabelled enumeration, is defined by

$$\widetilde{F}(x) = \sum_{n\geq 0} \widetilde{F}_n x^n, \tag{2}$$

where $\widetilde{F}_n$ denotes the number of isomorphism classes of F-structures of order n. The *cycle index series*, $Z_F(x_1, x_2, x_3, \cdots)$, is defined by

$$Z_F(x_1, x_2, x_3, \cdots) = \sum_{n\geq 0} \frac{1}{n!} \sum_{\sigma\in S_n} \operatorname{fix} F[\sigma]\, x_1^{\sigma_1} x_2^{\sigma_2} x_3^{\sigma_3} \cdots, \tag{3}$$

where S_n denotes the group of permutations of $[n]$, $\operatorname{fix} F[\sigma]$ is the number of F-structures on $[n]$ left fixed by σ, and σ_j is the number of cycles of length j in σ. Finally, the *molecular expansion* of F is a description and a classification of the F-structures according to their stabilizers.

Combinatorial operations are defined on species: sum, product, (partitional) composition, derivation, which correspond to the usual operations on the exponential generating functions. And there are rules for computing the other associated series, involving plethysm. See [2] for more details. An *isomorphism* $F \cong G$ between species, denoted by $F = G$, is a family of bijections between structures,

$$\alpha_U : F[U] \to G[U],$$

where U ranges over all underlying sets, which commute with relabellings. It gives rise to equalities $F(x) = G(x)$, $\widetilde{F}(x) = \widetilde{G}(x)$, $Z_F = Z_G$, ... between their series expansions.

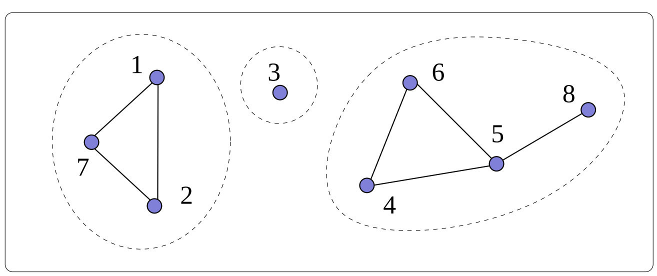

Figure 1: A simple graph g and its connected components

For example, the fact that any simple graph on a set (of vertices) U is the disjoint union of connected simple graphs (see Figure 1) is expressed by the equation

$$\mathcal{G} = E(\mathcal{C}), \tag{4}$$

where $\mathcal{G}$ denotes the species of (simple) graphs, $\mathcal{C}$, that of connected graphs, and E, the species of *Sets* (in French: *Ensembles*). There correspond the well-known relations

$$\mathcal{G}(x) = \exp(\mathcal{C}(x)), \tag{5}$$

for their exponential generating functions, and, for their tilde generating functions,

$$\begin{aligned} \widetilde{\mathcal{G}}(x) &= Z_E(\widetilde{\mathcal{C}}(x), \widetilde{\mathcal{C}}(x^2), \ldots) \\ &= \exp\left(\sum_{k\geq 1} \frac{1}{k}\widetilde{\mathcal{C}}(x^k)\right). \end{aligned} \tag{6}$$

Definitions. A *cutpoint* (or *articulation point*) of a connected graph g is a vertex of g whose removal yields a disconnected graph. A connected graph is called *2-connected* if it has no cutpoint. A *block* in a simple graph is a maximal 2-connected subgraph. The *block-graph* of a graph g is a new graph whose vertices are the blocks of g and whose edges correspond to blocks having a common cutpoint. The *block-cutpoint tree* of a connected graph g is a graph whose vertices are the blocks and the cutpoints of g and whose edges correspond to incidence relations in g. See Figure 2.

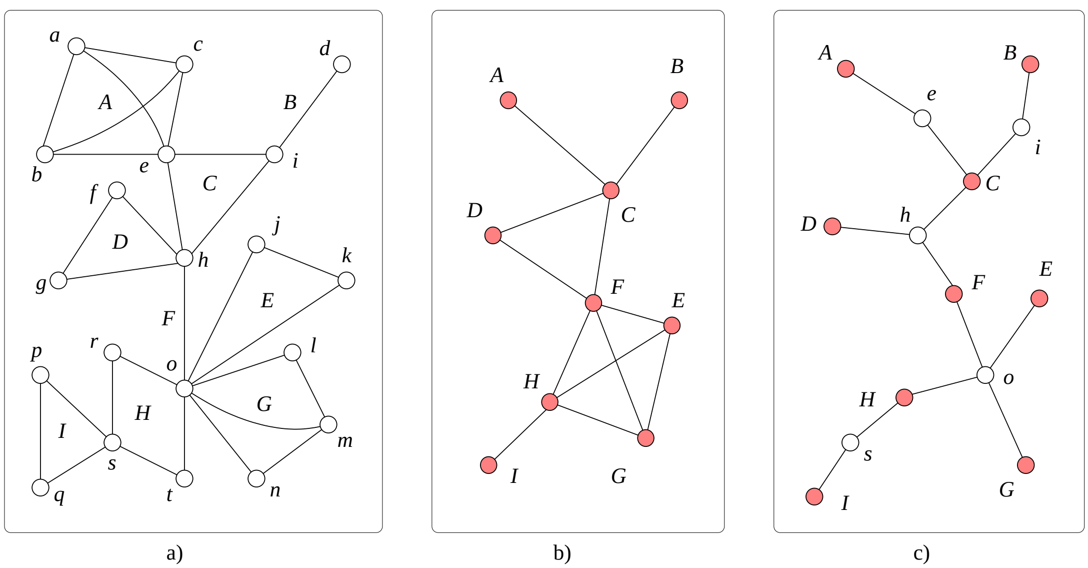

Figure 2: a) A connected graph g, b) the block-graph of g, c) the block-cutpoint tree of g

Now let $\mathcal{B}$ be a given species of 2-connected graphs. We denote by $C_{\mathcal{B}}$ the species of connected graphs all of whose blocks are in $\mathcal{B}$, called $C_{\mathcal{B}}$*-graphs*.

Examples 1.1. Here are some examples for various choices of $\mathcal{B}$:

1. If $\mathcal{B} = \mathcal{B}_a$, the class of *all* 2-connected graphs, then $C_{\mathcal{B}} = \mathcal{C}$, the species of (all) connected graphs.

2. If $\mathcal{B} = K_2$, the class of "edges", then $C_{\mathcal{B}} = \mathfrak{a}$, the species of (unrooted, free) trees ($\mathfrak{a}$ for French *arbres*).

3. If $\mathcal{B} = \{P_m, m \geq 2\}$, where P_m denotes the class of size-m polygons (by convention, $P_2 = K_2$), then $C_{\mathcal{B}} = \mathrm{Ca}$, the species of cacti. A *cactus* can also be defined as a connected graph in which no edge lies in more than one cycle. Figure 3, a), represents a typical cactus.

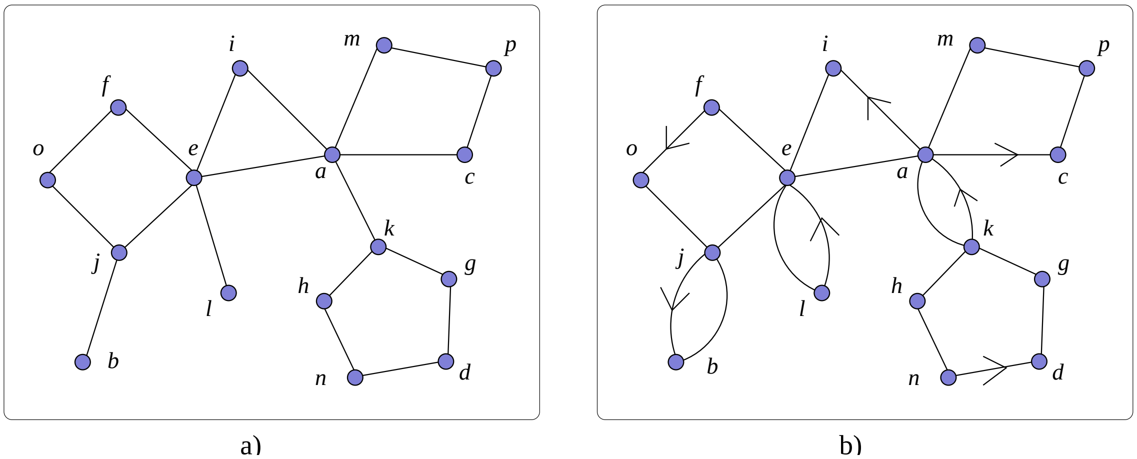

Figure 3: a) a typical cactus, b) a typical oriented cactus

4. If $\mathcal{B} = K_3 = P_3$, the class of "triangles", then $C_{\mathcal{B}} = \delta$, the class of triangular cacti.

5. If $\mathcal{B} = \{K_n, n \geq 2\}$, the family of complete graphs, then $C_{\mathcal{B}} = \mathrm{Hu}$, the species of *Husimi graphs*; that is, of connected graphs whose blocks are complete graphs. They were first (informally) introduced by Husimi in [12]. A Husimi graph is shown in Figure 2, b). See also Figure 7. It can be easily shown that any Husimi graph is the block-graph of some connected graph.

6. If $\mathcal{B} = \{C_n, n \geq 2\}$, the family of oriented cycles, then $C_{\mathcal{B}} = \mathrm{Oc}$, the species of *oriented cacti*. Figure 3, b) shows a typical oriented cactus. These structures were introduced by C. Springer [29] in 1996. Although directed graphs are involved here, the functional equations (7) and (12) given below are still valid.

Remark. Cacti were first called Husimi trees. See for example [9], [11], [27] and [30]. However this term received much criticism since they are not necessarily trees. Also, a careful reading of Husimi's article [12] shows that the graphs he has in mind and that he enumerates (see formula (42) below) are the Husimi graphs defined in item 5 above. The term cactus is now widely used, see Harary and Palmer [10]. Cacti appear regularly in the mathematical litterature, for example, in the classification of base matroids [21], and in combinatorial optimization [4].

The following functional equation (see (7)) is fairly well known. It can be found in various forms and with varying degrees of generality in [2], [10], [13], [18], [19], [20], [25], [27], [28]. In fact, it was anticipated by the physicists (see [12] and [30]) in the context of Mayers' theory of cluster integrals as we will see below. The form given here, in the structural language of species, is the most general one since all the series expansions follow. It is also the easiest form to prove.

Recall that for any species $F = F(X)$, the derivative F' of F is the species defined as follows: an F'-structure on a set U is an F-structure on the set $U \cup \{*\}$, where $*$ is an external (unlabelled) element. In other words, one sets

$$F'[U] = F[U + \{*\}].$$

Moreover, the operation $F \mapsto F^\bullet$, of pointing (or rooting) F-structures at an element of the underlying set, can be defined by

$$F^\bullet = X \cdot F'.$$

Theorem 1.1 *Let $\mathcal{B}$ be a class of 2-connected graphs and $C_{\mathcal{B}}$ be the species of connected graphs all of whose blocks are in $\mathcal{B}$. We then have the functional equation*

$$C'_{\mathcal{B}} = E(\mathcal{B}'(C^\bullet_{\mathcal{B}})). \tag{7}$$

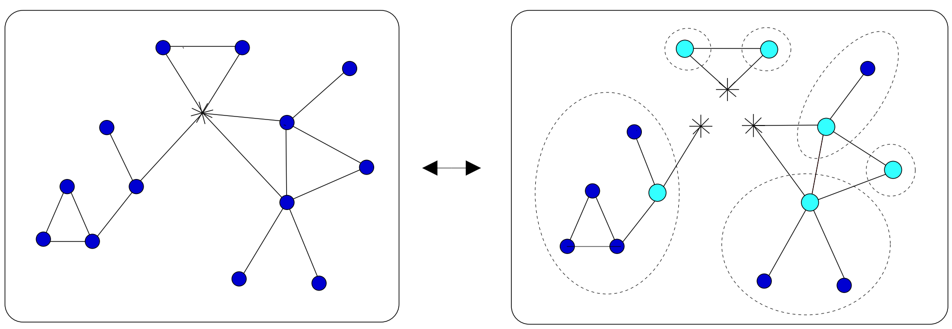

Figure 4: $C'_{\mathcal{B}} = E(\mathcal{B}'(C^\bullet_{\mathcal{B}}))$

Proof. See Figure 4. $\square$

Multiplying (7) by X, one finds

$$C^\bullet_{\mathcal{B}} = X \cdot E(\mathcal{B}'(C^\bullet_{\mathcal{B}})), \tag{8}$$

and, for the exponential generating function,

$$C^\bullet_{\mathcal{B}}(x) = x \cdot \exp(\mathcal{B}'(C^\bullet_{\mathcal{B}}(x))). \tag{9}$$

1.2 Weighted versions

Weighted versions of these equations are needed in the applications. See for example Uhlenbeck and Ford [30]. A *weighted species* is a species F together with weight functions $w = w_U : F[U] \to \mathbb{K}$ defined on F-structures, which commute with the relabellings. Here $\mathbb{K}$ is a commutative ring in which the weights are taken; usually $\mathbb{K}$ a ring of polynomials or of formal power series over a field of characteristic zero. We write $F = F_w$ to emphasize the fact that F is a weighted species with weight function w. The associated generating functions are then adapted by replacing set cardinalities $|A|$ by total weights

$$|A|_w = \sum_{a \in A} w(a).$$

The basic operations on species are also adapted to the weighted context, using the concept of Cartesian product of weighted sets: Let (A, u) and (B, v) be weighted sets. A weight function w is defined on the Cartesian product $A \times B$ by

$$w(a, b) = u(a) \cdot v(b).$$

We then have $\quad |A \times B|_w = |A|_u \cdot |B|_v$.

Definition. A weight function w on the species $\mathcal{G}$ of graphs is said to be *multiplicative on the connected components* if for any graph $g \in \mathcal{G}[U]$, whose connected components are $c_1, c_2, \ldots, c_k$, we have

$$w(g) = w(c_1)w(c_2) \cdots w(c_k).$$

Examples 1.2. The following weight functions w on the species of graphs are multiplicative on the connected components.

1. $w_1(g) := y^{e(g)}$, where $e(g)$ is the number of edges of g.
2. $w_2(g) =$ graph complexity of $g :=$ number of maximal spanning forests of g.
3. $w_3(g) := x_0^{n_0} x_1^{n_1} x_2^{n_2} \cdots$, where n_i is the number of vertices of degree i.

Theorem 1.2 *Let w be a weight function on graphs which is multiplicative on the connected components. Then we have*

$$\mathcal{G}_w = E\left(\mathcal{C}_w\right). \tag{10}$$

For the exponential generating functions, we have

$$G_w(x) = \exp(\mathcal{C}_w(x)),$$

where $G_w(x) = \sum_{n \geq 0} |\mathcal{G}[n]|_w \frac{x^n}{n!} = \sum_{n \geq 0} (\sum_{g \in G[n]} w(g)) \frac{x^n}{n!}$, and similarly for $\mathcal{C}_w(x)$.

Definition. A weight function on connected graphs is said to be *block-multiplicative* if for any connected graph c, whose blocks are $b_1, b_2, \ldots, b_k$, we have

$$w(c) = w(b_1)w(b_2) \cdots w(b_k).$$

Examples 1.3. The weight functions $w_1(g) = y^{e(g)}$ and $w_2(g) =$ graph complexity of g of Examples 1.2 are block-multiplicative, but the function $w_3(g) = x_0^{n_0} x_1^{n_1} x_2^{n_2} \cdots$ is not. Another example of a block-multiplicative weight function is obtained by introducing formal variables y_i ($i \geq 2$) marking the block sizes. In other terms, if the connected graph c has n_i blocks of size i, for $i = 2, 3, \ldots$, one sets

$$w(c) = y_2^{n_2} y_3^{n_3} \cdots . \tag{11}$$

The following result is then simply the weighted version of Theorem 1.1.

Theorem 1.3. Let w be a block-multiplicative weight function on connected graphs whose blocks are in a given species B. Then we have

$$(C_{\mathcal{B}}^{\bullet})_w = X \cdot E(\mathcal{B}'_w((C_{\mathcal{B}}^{\bullet})_w)). \tag{12}$$

1.3 Outline

In the next section, we see how equations (10) and (12) are involved in the thermodynamical study of imperfect (or non ideal) gases, following Mayers' theory of cluster integrals [22], as presented in Uhlenbeck and Ford [30]. In particular, the virial expansion, which is a kind of asymptotic refinement of the perfect gases law, is established rigourously, at least in its formal power series form; see equation (34) below. It is amazing to realize that the coefficients of the virial expansion involve directly the total valuation $|\mathcal{B}_a[n]|_w$, for $n \geq 2$, of 2-connected graphs. An important role in this theory is also played by the enumerative formula (42) for labelled Husimi trees according to their block-size distribution, which extends Cayley's formula n^{n-2} for the number of labelled trees of size n.

Motivated by this, we consider, in Section 3, the enumeration of some classes of connected graphs of the form $C_{\mathcal{B}}$, according or not to their block-size distribution. Included are the species of Husimi graphs, cacti, and oriented cacti. In the labelled case, the methods involve the Lagrange inversion formula and Prüfer-type bijections. It is also natural to examine the unlabelled enumeration of these structures. This is a more difficult problem, for two reasons. First, equation (12) deals with rooted structures and it is necessary to introduce a tool for counting the unrooted ones. Traditionally, this is done by extending Otter's Dissimilarity Charactistic formula for trees [26]. See for example [9]. Inspired by formulas of Norman ([6], (18)) and Robinson ([28], Theorem 7), we have given over the years a more structural formula which we call the Dissymmetry Theorem for Graphs, whose proof is remarkably simple and which can easily be adapted to various classes of tree-like structures; see [2], [3], [7], [14], [15]–[17], [19], [20]. Second, as for trees, it should not be expected to obtain simple closed expressions but rather recurrence formulas for the number of unlabelled $C_{\mathcal{B}}$-structures. Three examples are given in this section.

Finally, in Section 4, we present the molecular expansion of some of these species. It consists of a descriptive classification of the unlabelled structures in a given class in terms of elementary species from which all their symmetries can be deduced. This expansion can be first computed recursively for the rooted species and the Dissymmetry Theorem is then invoked for the unrooted ones. The computations can be carried out using the Maple package "Devmol" available at the URL www.lacim.uqam.ca; see also [1].

Acknowledgements. This paper is partly taken from my student Mélanie Nadeau's "Mémoire de maîtrise" [24]. I would like to thank her and Pierre Auger for their consider-

able help, and also Abdelmalek Abdesselam, André Joyal, Gilbert Labelle, Bob Robinson, and Alan Sokal, for useful discussions.

2 Some statistical mechanics

2.1 Partition functions for the non-ideal gas

Consider a non-ideal gas, formed of N particles interacting in a vessel $V \subseteq \mathrm{IR}^3$ (whose volume is also denoted by V) and whose positions are $\overrightarrow{x_1}, \overrightarrow{x_2}, \ldots, \overrightarrow{x_N}$. The Hamiltonian of the system is of the form

$$H = \sum_{i=1}^{N} \left(\frac{\overrightarrow{p_i}^2}{2m} + U(\overrightarrow{x_i}) \right) + \sum_{1 \le i < j \le N} \varphi(|\overrightarrow{x_i} - \overrightarrow{x_j}|), \tag{13}$$

where $\overrightarrow{p_i}$ is the linear momentum vector and $\frac{\overrightarrow{p_i}^2}{2m}$ is the kinetic energy of the i^{th} particle, $U(\overrightarrow{x_i})$ is the potential at position $\overrightarrow{x_i}$ due to outside forces (e.g., walls), $|\overrightarrow{x_i} - \overrightarrow{x_j}| = r_{ij}$ is the distance between the particles $\overrightarrow{x_i}$ and $\overrightarrow{x_j}$, and it is assumed that the particles interact only pairwise through the central potential $\varphi(r)$. This potential function φ has a typical form shown in Figure 5 a).

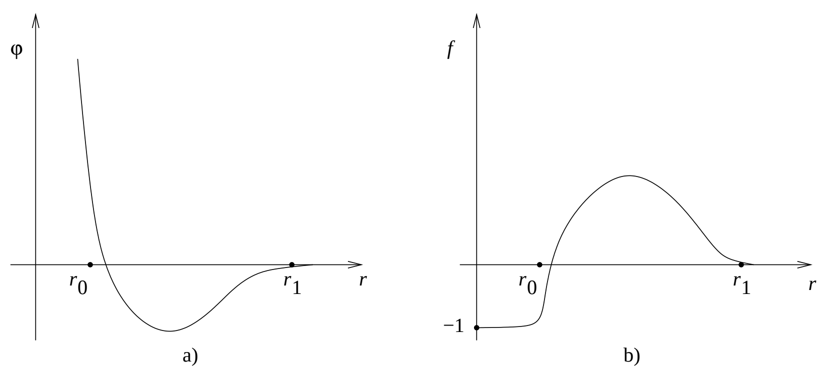

Figure 5: a) the function $\varphi(r)$, b) the function $f(r)$

The *canonical partition function* $Z(V, N, T)$ is defined by

$$Z(V, N, T) = \frac{1}{N! h^{3N}} \int \exp\left(-\beta H\right) d\Gamma, \tag{14}$$

where h is Planck's constant, $\beta = \frac{1}{kT}$, T is the absolute temperature and k is Boltzmann's constant, and Γ represents the state space $\overrightarrow{x_1}, \ldots, \overrightarrow{x_N}, \overrightarrow{p_1}, \ldots, \overrightarrow{p_N}$ of dimension $6N$. A first simplification comes from the assumption that the potential energy $U(\overrightarrow{x_i})$ is negligible or null. Secondly, the integral over the momenta $\overrightarrow{p_i}$ in (14) is a product of Gaussian integrals which are easily evaluated so that the canonical partition function can now be written as

$$Z(V, N, T) = \frac{1}{N! \lambda^{3N}} \int_V \cdots \int_V \exp\left(-\beta \sum_{i<j} \varphi(|\overrightarrow{x_i} - \overrightarrow{x_j}|) \right) d\overrightarrow{x_1} \cdots d\overrightarrow{x_N}, \tag{15}$$

where $\lambda = h(2\pi mkT)^{-\frac{1}{2}}$.

Mathematically, the *grand-canonical distribution* is simply the generating function for the canonical partition functions, defined by

$$Z_{\mathrm{gr}}(V,T,z) = \sum_{N=0}^{\infty} Z(V,N,T)(\lambda^3 z)^N, \tag{16}$$

where the variable z is called the *fugacity* or the *activity*. All the macroscopic parameters of the system are then defined in terms of this grand canonical ensemble. For example, the *pressure* P, the average number of particles $\overline{N}$, and the *density* ρ, are defined by

$$\frac{P}{kT} = \frac{1}{V}\log Z_{\mathrm{gr}}(V,T,z), \quad \overline{N} = z\frac{\partial}{\partial z}\log Z_{\mathrm{gr}}(V,T,z), \quad \text{and} \quad \rho := \frac{\overline{N}}{V}. \tag{17}$$

2.2 The virial expansion

In order to better explain the thermodynamic behaviour of non ideal gases, Kamerlingh Onnes proposed, in 1901, a series expansion of the form

$$\frac{P}{kT} = \frac{\overline{N}}{V} + \gamma_2(T)\left(\frac{\overline{N}}{V}\right)^2 + \gamma_3(T)\left(\frac{\overline{N}}{V}\right)^3 + \cdots, \tag{18}$$

called the *virial expansion*. Here $\gamma_2(T)$ is the second virial coefficient, $\gamma_3(T)$ the third, etc. This expansion was first derived theoretically from the partition function Z_{gr} by Mayer [22] around 1930. It is the starting point of Mayer's theory of "cluster integrals". Mayer's idea consists of setting

$$1 + f_{ij} = \exp\left(-\beta\varphi(|\overrightarrow{x_i} - \overrightarrow{x_j}|)\right), \tag{19}$$

where $f_{ij} = f(r_{ij})$. The general form of the function $f(r) = \exp(-\beta\varphi(r)) - 1$ is shown in Figure 5, b). In particular, $f(r)$ vanishes when r is greater than the range r_1 of the interaction potential. Alternatively, f should satisfy some integrability condition. By substituting in the canonical partition function (15), one obtains

$$Z(V,N,T) = \frac{1}{N!\lambda^{3N}}\int_V \cdots \int_V \prod_{1\le i<j\le N}(1+f_{ij})\, d\overrightarrow{x_1}\cdots d\overrightarrow{x_N}. \tag{20}$$

The terms obtained by expanding the product $\prod_{1\le i<j\le N}(1+f_{ij})$ can be represented by simple graphs where the vertices are the particles and the edges are the chosen factors f_{ij}. The partition function (20) can then be rewritten in the form

$$\begin{aligned} Z(V,N,T) &= \frac{1}{N!\lambda^{3N}}\sum_{g\in\mathcal{G}[N]}\int_V\cdots\int_V \prod_{\{i,j\}\in g} f_{ij}\; d\overrightarrow{x_1}\cdots d\overrightarrow{x_N} \\ &= \frac{1}{N!\lambda^{3N}}\sum_{g\in\mathcal{G}[N]} W(g), \end{aligned} \tag{21}$$

where the weight $W(g)$ of a graph g is given by the integral

$$W(g) = \int_V \cdots \int_V \prod_{\{i,j\} \in g} f_{ij} d\overrightarrow{x_1} \cdots d\overrightarrow{x_N}. \tag{22}$$

For the grand canonical function, we then have

$$\begin{aligned} Z_{\mathrm{gr}}(V,T,z) &= \sum_{N=0}^{\infty} Z(V,N,T)(\lambda^3 z)^N \\ &= \sum_{N=0}^{\infty} \frac{1}{N!\lambda^{3N}} \sum_{g \in \mathcal{G}[N]} W(g)(\lambda^3 z)^N \\ &= \sum_{N=0}^{\infty} \frac{1}{N!} \sum_{g \in \mathcal{G}[N]} W(g) z^N \\ &= \mathcal{G}_W(z). \end{aligned} \tag{23}$$

Proposition 2.1 *The weight function W is multiplicative on the connected components.*

For example, for the graph g of Figure 1, we have

$$\begin{aligned} W(g) &= \int_{V^8} f_{12}f_{17}f_{27}f_{45}f_{46}f_{56}f_{58} d\overrightarrow{x_1} \cdots d\overrightarrow{x_8} \\ &= \int f_{12}f_{17}f_{27} d\overrightarrow{x_1} d\overrightarrow{x_2} d\overrightarrow{x_7} \int d\overrightarrow{x_3} \int f_{45}f_{46}f_{56}f_{58} d\overrightarrow{x_4} d\overrightarrow{x_5} d\overrightarrow{x_6} d\overrightarrow{x_8} \\ &= W(c_1)W(c_2)W(c_3), \end{aligned}$$

where c_1, c_2 and c_3 represent the three connected components of g. Following Theorem 1.2, we deduce that

$$\mathcal{G}_W(z) = \exp\left(\mathcal{C}_W(z)\right), \tag{24}$$

where $\mathcal{C}_W$ denotes the weighted species of connected graphs, with

$$\mathcal{C}_W(z) = \sum_{n \geq 1} |\mathcal{C}[n]|_W \frac{z^n}{n!},$$

and

$$|\mathcal{C}[n]|_W = \sum_{c \in \mathcal{C}[n]} \int_V \cdots \int_V \prod_{\{i,j\} \in c} f_{ij} \; d\overrightarrow{x_1} \cdots d\overrightarrow{x_n}. \tag{25}$$

Historically, the quantities $b_n(V) = \frac{1}{Vn!}|\mathcal{C}[n]|_W$ are precisely the *cluster integrals* of Mayer. Equation (24) then provides a combinatorial interpretation for the quantity $\frac{P}{kT}$. Indeed, one has, by (17),

$$\begin{aligned} \frac{P}{kT} &= \frac{1}{V} \log Z_{\mathrm{gr}}(V,T,z) \\ &= \frac{1}{V} \log \mathcal{G}_W(z) \\ &= \frac{1}{V} \mathcal{C}_W(z). \end{aligned} \tag{26}$$

Proposition 2.2 *For large V, the weight function $w(c) = \frac{1}{V}W(c)$, defined on the species of connected graphs, is block-multiplicative.*

Proof. First observe that for any connected graph c on the set of vertices $[k] = \{1, 2, \dots k\}$, the value of the partial integral

$$I = I(\overrightarrow{x_k}) = \lim_{V\to\infty} \int_V \cdots \int_V \prod_{\{i,j\}\in c} f_{ij} d\overrightarrow{x_1} \cdots d\overrightarrow{x_{k-1}} \tag{27}$$

is in fact independent of $\overrightarrow{x_k}$. Indeed, since the f_{ij}'s only depend on the relative positions $r_{ij} = |\overrightarrow{x_i} - \overrightarrow{x_j}|$, and considering the short range r_1 of the interaction potential and the connectednes of c, we see that the support of the integrand in (27) lies in a ball of radius at most $(k-1)r_1$ centered at $\overrightarrow{x_k}$ and that a simple translation $\overrightarrow{x_i} \mapsto \overrightarrow{x_i} + \overrightarrow{u}$ will give the same value of the integral. It follows that

$$\begin{aligned} W(c) &= \int_V \cdots \int_V \prod_{\{i,j\}\in c} f_{ij} d\overrightarrow{x_1} \cdots d\overrightarrow{x_{k-1}} \int_V d\overrightarrow{x_k} \\ &\approx \int_V \cdots \int_V \prod_{\{i,j\}\in c} f_{ij} d\overrightarrow{x_1} \cdots d\overrightarrow{x_{k-1}} \cdot V, \end{aligned}$$

for large V, and the value of the partial integral (27) is in fact $w(c)$:

$$\lim_{V\to\infty} \int_V \cdots \int_V \prod_{\{i,j\}\in c} f_{ij} d\overrightarrow{x_1} \cdots d\overrightarrow{x_{k-1}} = \lim_{V\to\infty} \frac{1}{V} W(c) = w(c). \tag{28}$$

Now if a connected graph is decomposed into blocks $b_1, b_2, \dots, b_k$, we have

$$w(c) = w(b_1)w(b_2)\cdots w(b_k).$$

For example, for the graph c shown in Figure 6, we have

$$\begin{aligned} w(c) &= \int_{V^7} f_{12}f_{13}f_{23}f_{34}f_{56}f_{37}f_{36}f_{67}f_{68}f_{78} d\overrightarrow{x_1} d\overrightarrow{x_2} \cdots d\overrightarrow{x_7} \\ &= \int f_{12}f_{13}f_{23} d\overrightarrow{x_1} d\overrightarrow{x_2}\, f_{34} d\overrightarrow{x_4}\, f_{56} d\overrightarrow{x_5}\, f_{37}f_{36}f_{67}f_{68}f_{78} d\overrightarrow{x_3} d\overrightarrow{x_6} d\overrightarrow{x_7} \\ &= w(b_1)w(b_2)w(b_3)w(b_4). \end{aligned}$$

□

From Theorem 1.3 it follows that

$$\mathcal{C}_w^\bullet = X \cdot E(\mathcal{B}'_w(\mathcal{C}_w^\bullet)), \tag{29}$$

where $\mathcal{B} = \mathcal{B}_a$ is the species of all 2-connected graphs, and for the exponential generating functions,

$$\mathcal{C}_w^\bullet(z) = z \exp(\mathcal{B}'_w(\mathcal{C}_w^\bullet(z))). \tag{30}$$

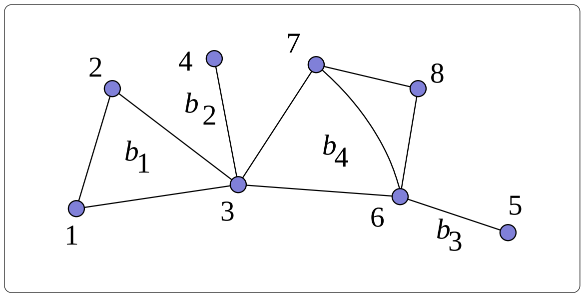

Figure 6: A connected graph with blocks b_1, b_2, b_3, b_4

2.3 Computation of the virial expansion

The virial expansion (18) can now be established, following Uhlenbeck and Ford [30]. From (17), we have, for the density $\rho(z) = \frac{\overline{N}}{V}$,

$$\begin{aligned} \rho(z) &= z\frac{\partial}{\partial z}\frac{1}{V}\log Z_{\mathrm{gr}}(V,T,z) \\ &= z\frac{\partial}{\partial z}\mathcal{C}_w(z) \\ &= \mathcal{C}_w^{\bullet}(z). \end{aligned} \tag{31}$$

Hence $\rho(z)$ satisfies the functional equation (30), that is

$$\rho(z) = z\exp\mathcal{B}_w'(\rho(z)). \tag{32}$$

The idea is then to use this relation in order to express z in terms of ρ in $\frac{P}{kT}(z)$, as follows. We have, by (26) and (31),

$$\begin{aligned} \frac{P}{kT} &= \frac{1}{V}\log Z_{\mathrm{gr}}(V,T,z) \\ &= \mathcal{C}_w(z) \\ &= \int_0^z \mathcal{C}_w'(t)\ dt \\ &= \int_0^z \frac{\rho(t)}{t}\ dt. \end{aligned} \tag{33}$$

Let us make the change of variable

$$t = t(r) = r\exp\left(-\mathcal{B}_w'(r)\right),$$

which is the inverse function of $r = \rho(t)$, by (32). Note that $\rho(0) = 0$ and $\rho(z) = \rho$, and also that

$$dt = [\exp\left(-\mathcal{B}_w'(r)\right) - r\exp\left(-\mathcal{B}_w'(r)\right)\cdot\mathcal{B}_w''(r)]dr.$$

Pursuing the computation of the integral (33), we have

$$\begin{aligned}
\frac{P}{kT} &= \int_0^z \frac{\rho(t)}{t}\ dt \\
&= \int_0^\rho \left(1 - r\mathcal{B}_w''(r)\right)\ dr \\
&= \rho - \int_0^\rho r\mathcal{B}_w''(r)\ dr \\
&= \rho - \int_0^\rho \sum_{n\geq 1} n\beta_{n+1}\frac{r^n}{n!}\ dr \\
&= \rho - \sum_{n\geq 2}(n-1)\beta_n\frac{\rho^n}{n!}, \qquad (34)
\end{aligned}$$

where we have set $\mathcal{B}_w(r) = \sum_{n\geq 2}\beta_n\frac{r^n}{n!}$. This is precisely the virial expansion, with $\rho = \overline{\frac{N}{V}}$. Hence the n^{th} virial coefficient, for $n \geq 2$, is given by

$$\begin{aligned}
\gamma_n(T) &= -\frac{(n-1)}{n!}\beta_n \\
&= -\frac{(n-1)}{n!}|\mathcal{B}[n]|_w. \qquad (35)
\end{aligned}$$

Mayer's original proof of the virial expansion is more technical, since he is not aware of a direct combinatorial proof of equation (30). The following observation is used: By grouping the connected graphs on $[n]$ whose block decomposition determines the same Husimi graph on $[n]$, and then collecting all Husimi graphs having the same block-size distribution, one obtains, using the 2-multiplicativity of w,

$$\begin{aligned}
|C[n]|_w &= \sum_{c\in\mathcal{C}[n]} w(c) \\
&= \sum_{\{B_1,B_2,\ldots\}\in\mathrm{Hu}[n]}\ \sum_{\{b_i\in\mathcal{B}[B_i]\}}\prod_i w(b_i) \\
&= \sum_{\{B_1,B_2,\ldots\}\in\mathrm{Hu}[n]}\ \prod_i \sum_{b\in\mathcal{B}[B_i]} w(b) \\
&= \sum_{\{B_1,B_2,\ldots\}\in\mathrm{Hu}[n]}\ \prod_i \beta_{|B_i|} \\
&= \sum_{\substack{n_2,n_3,\ldots \\ \sum n_i(i-1)=n-1}} \mathrm{Hu}(n_2,n_3,\ldots)\beta_2^{n_2}\beta_3^{n_3}\cdots,
\end{aligned}$$

where $\mathrm{Hu}[n]$ denotes the set of Husimi graphs on $[n]$ and $\mathrm{Hu}(n_2,n_3,\ldots)$ is the number of Husimi graphs on $[n]$ having n_i blocks of size i, for $i \geq 2$. Mayer then proves the following enumerative formula

$$\mathrm{Hu}(n_2,n_3,\ldots) = \frac{(n-1)!\, n^{\sum n_j - 1}}{\prod_{i\geq 2}(i-1)!^{n_i}\, n_i!}, \qquad (36)$$

and goes on proving (30) and (34) analytically.

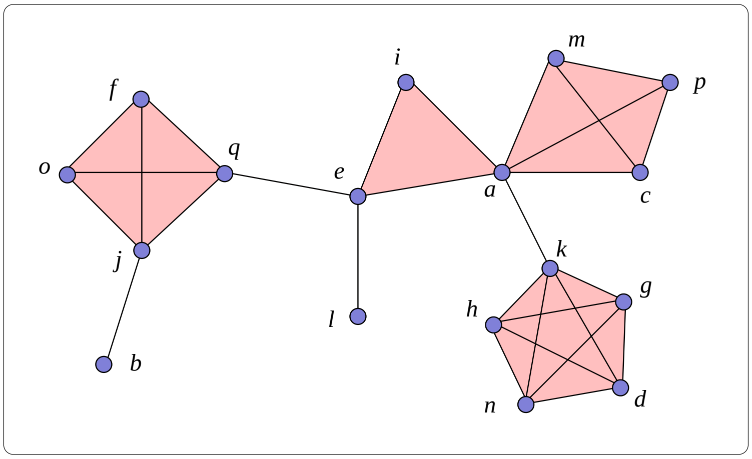

Figure 7: A Husimi graph with block-size distribution $(2^4 3^1 4^2 5^1 6^0 \cdots)$

Formula (36) is quite remarkable. It is an extension of Cayley's formula n^{n-2} for the number of trees on $[n]$ (take $n_2 = n-1, n_3 = 0, \ldots$). It has many different proofs, using, for example, Lagrange inversion or a Prüfer correspondence, and gives the motivation for the enumerative problems related to Husimi graphs, cacti, and oriented cacti, studied in the next section.

2.4 Gaussian model

It is interesting, mathematically, to consider a Gaussian model, where

$$f_{ij} = -\exp\left(-\alpha||\overrightarrow{x_i} - \overrightarrow{x_j}||^2\right), \tag{37}$$

which corresponds to a soft repulsive potential, at constant temperature. In this case, all cluster integrals can be explicitly computed (see [30]): The weight $w(c)$ of a connected graph c, defined by (28), has value

$$w(c) = (-1)^{e(c)} \left(\frac{\pi}{\alpha}\right)^{\frac{3}{2}(n-1)} \gamma(c)^{-\frac{3}{2}}, \tag{38}$$

where $e(c)$ is the number of edges of c and $\gamma(c)$ is the graph complexity of c, that is, the number of spanning subtrees of c. This formula incorporates three very descriptive weightings on connected graphs, which are multiplicative on 2-connected components, namely,

$$w_1(c) = y^{e(c)}, \quad w_2(c) = \gamma(c),$$

which were already seen in Examples 1.2, and $w_3(c) = u^{n-1}$, where n is the number of vertices.

3 Enumerative results

In this section, we investigate the enumeration of Husimi graphs, cacti and oriented cacti, according, or not, to their block size distribution. Recall that these classes can be viewed as species of connected graphs of the form $C_{\mathcal{B}}$ and that the functional equation (8) for rooted $C_{\mathcal{B}}$-structures can be invoked, as well as its weighted version (12).

3.1 Labelled enumeration

Proposition 3.1 *The number* $\mathrm{H}_n = |\,\mathrm{Hu}[n]|$ *of (labelled) Husimi graphs on* $[n]$*, for* $n \geq 1$*, is given by*

$$\mathrm{H}_n = \sum_{k\geq 0} S(n-1,k)\, n^{k-1}, \tag{39}$$

where k *represents the number of blocks and* $S(m,k)$ *denotes the Stirling number of the second kind.*

Proof. The species Hu of Husimi graphs is of the form $C_{\mathcal{B}}$ with $\mathcal{B} = K_{\geq 2}$, the class of complete graphs of size ≥ 2. From the species point of view, it is equivalent to take $\mathcal{B} = E_{\geq 2}$, the species of sets of size ≥ 2. We then have $\mathcal{B}' = E_{\geq 1}$, with exponential generating function $E_{\geq 1}(x) = e^x - 1$. Hence the species $\mathrm{Hu}^\bullet$ of rooted Husimi graphs satisfies the functional equation

$$\mathrm{Hu}^\bullet = X\, E(E_{\geq 1}(\mathrm{Hu}^\bullet)). \tag{40}$$

which, for the generating function $\mathrm{Hu}^\bullet(x)$, translates into

$$\mathrm{Hu}^\bullet(x) = xR(\mathrm{Hu}^\bullet(x)), \tag{41}$$

with $R(x) = \exp(e^x - 1)$. The Lagrange inversion formula then gives

$$\begin{aligned}
[x^n]\,\mathrm{Hu}^\bullet(x) &= \frac{1}{n}[t^{n-1}]\left(e^{e^t-1}\right)^n \\
&= \frac{1}{n}[t^{n-1}]e^{n(e^t-1)} \\
&= \frac{1}{n}[t^{n-1}]\sum_{k\geq 0} n^k \frac{(e^t-1)^k}{k!} \\
&= \frac{1}{n}[t^{n-1}]\sum_{k\geq 0} n^k \sum_m S(m,k)\frac{t^m}{m!} \\
&= \frac{1}{n}\sum_{k\geq 0} n^k \frac{S(n-1,k)}{(n-1)!} \\
&= \sum_{k\geq 0} n^k \frac{S(n-1,k)}{n!}.
\end{aligned}$$

Since we are dealing with exponential generating functions, we should multiply by $n!$ to get the coefficient of $\frac{x^n}{n!}$. Moreover we should divide by n to obtain the number of unrooted

Husimi graphs. In conclusion, we have $\mathrm{H}_n = \frac{1}{n} n! [x^n] \mathrm{Hu}^\bullet(x) = \sum_{k\geq 0} S(n-1,k)\, n^{k-1}$. The fact that k represents the number of blocks will appear more clearly in the bijective proof given below. $\square$

We now come to Mayer's enumerative formula (36).

Proposition 3.2 (Mayer [23], Husimi [12]) *Let $(n_2, n_3, \ldots)$ be a sequence of non-negative integers and $n = \sum_{i\geq 2} n_i(i-1)+1$. Then the number $\mathrm{Hu}(n_2, n_3, \ldots)$ of Husimi graphs on $[n]$ having n_i blocks of size i is given by*

$$\mathrm{Hu}(n_2, n_3, \ldots) = \frac{(n-1)!}{(1!)^{n_2} n_2! (2!)^{n_3} n_3! \cdots}\, n^{k-1}, \tag{42}$$

where $k = \sum_{i\geq 2} n_i$ is the total number of blocks.

Proof. Here we are dealing with the weighted species Hu_w of Husimi graphs weighted by the function $w(h) = y_2^{n_2} y_3^{n_3} \cdots$, which describes the block-size distribution of the Husimi graph h. Technically, we should take $B_w = \sum_{m\geq 2} (E_m)_{y_m}$, where the index y_m indicates that the sets of size m have weight y_m, for which

$$B_v(x) = \sum_{m\geq 2} y_m \frac{x^m}{m!} \quad \text{and} \quad \mathcal{B}'_v(x) = \sum_{m\geq 1} y_{m+1} \frac{x^m}{m!}.$$

The functional equation (12) then gives

$$\mathrm{Hu}^\bullet_w(x) = x\,(\exp(\sum_{m\geq 1} y_{m+1} \frac{x^m}{m!}) \circ \mathrm{Hu}^\bullet_w(x)) \tag{43}$$

and Lagrange inversion formula can be used. We find

$$\begin{aligned}
[x^n]\,\mathrm{Hu}^\bullet_w(x) &= \frac{1}{n}[t^{n-1}] \left(\exp(\sum_{m\geq 1} y_{m+1} \frac{t^m}{m!}) \right)^n \\
&= \frac{1}{n}[t^{n-1}] \exp(n \sum_{m\geq 1} y_{m+1} \frac{t^m}{m!}) \\
&= \frac{1}{n}[t^{n-1}] \sum_{k\geq 0} \frac{n^k}{k!} \left(\sum_{m\geq 1} y_{m+1} \frac{t^m}{m!} \right)^k \\
&= \frac{1}{n}[t^{n-1}] \sum_{k\geq 0} \frac{n^k}{k!} \left(y_2 \frac{t}{1!} + y_3 \frac{t^2}{2!} + y_4 \frac{t^3}{3!} \cdots \right)^k \\
&= \frac{1}{n}[t^{n-1}] \sum_{k\geq 0} \frac{n^k}{k!} \sum_{k_1+k_2+\cdots=k} \frac{k!}{k_1! k_2! \cdots} \left(y_2 \frac{t}{1!} \right)^{k_1} \left(y_3 \frac{t^2}{2!} \right)^{k_2} \cdots \\
&= \frac{1}{n}[t^{n-1}] \sum_{k\geq 0} \sum_{k_1+k_2+\cdots=k} \frac{n^k}{k_1! k_2! \cdots} \frac{y_2^{k_1} y_3^{k_2} \cdots}{(1!)^{k_1} (2!)^{k_2} \cdots} t^{k_1+2k_2+\cdots} \\
&= \sum_{k\geq 0} \sum_{\substack{k_1+k_2+\cdots=k \\ k_1+2k_2+\cdots=n-1}} \frac{n^{k-1}}{k_1!(1!)^{k_1} k_2! (2!)^{k_2} \cdots} y_2^{k_1} y_3^{k_2} \cdots.
\end{aligned}$$

Extracting the coefficient of the monomial $y_2^{n_2}y_3^{n_3}\cdots$ and multiplying by $(n-1)! = \frac{1}{n}n!$ then gives the result. $\square$

There are many other proofs of (42). Husimi [12] establishes a recurrence formula which, in fact, characterizes the numbers $\mathrm{Hu}(n_2, n_3, \ldots)$. He then goes on to prove the functional equation (43). Mayer gives a direct proof which becomes more convincing when coupled with a Prüfer-type correspondence. Such a correspondence is given by Springer in [29] for the number $\mathrm{Oc}(n_2, n_3, \ldots)$ of labelled oriented cacti having block size distribution $2^{n_2}3^{n_3}\cdots$.

It is easy to adapt Springer's bijection to Husimi graphs. To each Husimi graph h on $[n]$ having k blocks one assigns a pair (λ, π), where λ is a sequence $(j_1, \ldots, j_{k-1})$ of elements of $[n]$ of length $k-1$, and π is a partition of the set $[n]\backslash\{j_{k-1}\}$ into k parts. Moreover, if h has block size distribution $2^{n_2}3^{n_3}\cdots$, then π has part-size distribution $1^{n_2}2^{n_3}\cdots$. This is done as follows. A *leaf-block* b of a Husimi graph h is a block of h containing exactly one articulation point, denoted by j_b. Let $b = b(h)$ be the leaf-block of h for which the set $b(h)\backslash\{j_{b(h)}\}$ contains the smallest element among all sets of the form $b\backslash\{j_b\}$.

The correspondence proceeds recursively with the following steps:

1. Add $j_{b(h)}$ to the sequence λ.

2. Add the set $b(h)\backslash\{j_{b(h)}\}$ to the partition π.

3. Remove the block $b(h)$ (but not the articulation point $j_{b(h)}$) from h.

4. Resume with the remaining Husimi graph.

The procedure stops after the $(k-1)^{\text{th}}$ iteration when, in supplement, the last remaining block minus the $(k-1)^{\text{th}}$ articulation point j_{k-1} is added to the partition π. An example is shown in Figure 8. This procedure can easily be reversed and the resulting bijection proves both (39) and (42).

We now turn to the species of oriented cacti, defined in Example 1.1.6. These structures were introduced by Springer in [29] for the purpose of enumerating "ordered short factorizations" of a circular permutation ρ of length n into circular permutations, a_i of length i for each i. Such a factorization is called *short* if $\sum_{i\geq 2}(i-1)a_i = n-1$.

Proposition 3.3 *The number* $\mathrm{Oc}_n = |\,\mathrm{Oc}[n]|$ *of oriented cacti on* $[n]$*, for* $n \geq 2$*, is given by*

$$\mathrm{Oc}_n = \sum_{k\geq 1} \frac{(n-1)!}{k!}\binom{n-2}{k-1} n^{k-1}, \tag{44}$$

where $k = \sum_{i\geq 2} n_i$ *is the number of cycles.*

Proof. The proof is similar to that of (39). Here $\mathcal{B} = \sum_{k\geq 2} C_k$, the species of oriented cycles of size $k \geq 2$ and $\mathcal{B}' = \sum_{k\geq 1} X^k$, the species of "lists" (totally ordered sets) of

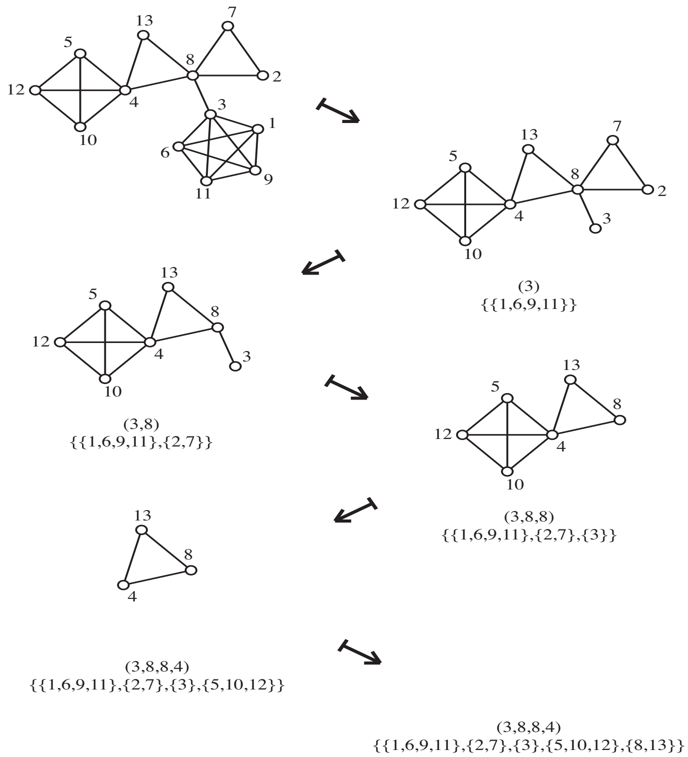

Figure 8: Prüfer correspondence for a Husimi graph

size $k \geq 1$. One can use Lagrange inversion formula, with $R(x) = \exp(\frac{x}{1-x})$. Alternately, observe that the factor $\frac{(n-1)!}{k!}\binom{n-2}{k-1}$ in (44) represents the number of partitions of a set of size $n-1$ into k totally ordered parts so that the Prüfer-type bijection of Springer [29] can be used here. □

Proposition 3.4 [29] *Let* $(n_2, n_3, \ldots)$ *be a sequence of non-negative integers and* $n = \sum_{i\geq 2} n_i(i-1) + 1$. *Then the number* $\mathrm{Oc}(n_2, n_3, \ldots)$ *of oriented cacti on* $[n]$ *having* n_i *cycles of size* i *for each* i, *is given by*

$$\mathrm{Oc}(n_2, n_3, \cdots) = \frac{(n-1)!}{n_2!n_3!\cdots}\, n^{k-1}, \tag{45}$$

where $k = \sum_{i\geq 2} n_i$ *is the number of cycles.*

Proof. Again, it is possible to use Lagrange inversion or the Prüfer-type correspondence of Springer. However the result now follows simply from equation (42) since it is easy to see that

$$\mathrm{Oc}(n_2, n_3, \ldots) = \prod_{i\geq 2}(i-1)!^{n_i}\, \mathrm{Hu}(n_2, n_3, \ldots).$$

Indeed, a set of size i can be structured into an oriented cycle in $(i-1)!$ ways. □

Finally, let us consider the species Ca of cacti which is of the form $C_{\mathcal{B}}$, where $B = \sum_{k\geq 2} P_k$ is the species of polygons. By convention, a polygon of size 2 is simply an edge ($K_2 = E_2$). See Example 1.1.3 and Figure 3. These structures frequently appear in mathematics, for example, more recently, in the context of the Traveling Salesman Problem (see for instance [4]) and in the characterization of graphic matroids [21].

Proposition 3.5 (Ford and Uhlenbeck [5]) *Let $(n_2, n_3, \ldots)$ be a sequence of non-negative integers and $n = \sum_{i\geq 2} n_i(i-1)+1$. Then the number $\mathrm{Ca}(n_2, n_3, \ldots)$ of cacti on $[n]$ having n_i polygons of size i for each i, is given by*

$$\mathrm{Ca}(n_2, n_3, \cdots) = \frac{1}{2^{\sum_{j\geq 3} n_j}} \frac{(n-1)!}{\prod_{j\geq 2} n_j!} n^{k-1}, \tag{46}$$

where $k = \sum_{i\geq 2} n_i$ is the number of polygons.

Proof. Indeed, since any (labelled) polygon of size ≥ 3 has 2 orientations, we see that

$$\mathrm{Oc}(n_2, n_3 \ldots) = 2^{\sum_{j\geq 3} n_j} \mathrm{Ca}(n_2, n_3, \cdots).$$

The result then follows from (45). $\square$

As observed in [5], this corrects the formula given in the introduction of [11] which is rather Husimi's formula (42) for Husimi graphs. By summing over the polygon-size distribution, we finally obtain:

Proposition 3.6 *The number $\mathrm{Ca}_n = |\,\mathrm{Ca}[n]|$ of (labelled) cacti on $[n]$, for $n \geq 2$, is given by*

$$\mathrm{Ca}_n = \sum_{k\geq 0} \sum_{\substack{n_2+n_3+\cdots=k \\ n_2+2n_3+\cdots=n-1}} \frac{(n-1)!\ n^{k-1}}{2^{n_3+n_4+\cdots}\ n_2!n_3!\cdots} \,. \tag{47}$$

3.2 Unlabelled enumeration

For unlabelled enumeration, the cases of rooted and of unrooted $C_{\mathcal{B}}$-graphs must be treated separately. There is a basic species relationship which permits the expression of the unrooted species in terms of the rooted ones. It plays the role of the classical Dissimilarity characteristic theorem for graphs (see [10]). We introduce the following notations:

1. $C_{\mathcal{B}}^{\diamondsuit}$ is the species of $C_{\mathcal{B}}$-graphs with a distinguished block,
2. $C_{\mathcal{B}}^{\bullet\diamondsuit}$ is the species of $C_{\mathcal{B}}$-graphs with a distinguished vertex-rooted block.

Theorem 3.7 (Dissymmetry Theorem for Graphs [19], [20], [2]). *The species $C_{\mathcal{B}}$ of connected graphs whose blocks are in $\mathcal{B}$ and its associated rooted species are related by the following isomophism:*

$$C_{\mathcal{B}}^{\bullet} + C_{\mathcal{B}}^{\diamondsuit} = C_{\mathcal{B}} + C_{\mathcal{B}}^{\bullet\diamondsuit}. \tag{48}$$

This identity can also be written as

$$C_{\mathcal{B}}^{\bullet} + \mathcal{B}(C_{\mathcal{B}}^{\bullet}) = C_{\mathcal{B}} + C_{\mathcal{B}}^{\bullet} \cdot \mathcal{B}'(C_{\mathcal{B}}^{\bullet}). \tag{49}$$

Proof. The proof of (48) is remarkably simple. It uses the concept of *center* of a $C_{\mathcal{B}}$-graph, which is defined as the center of the associated block-cutpoint tree (see Figure 2, c)). The center will necessarily be either a vertex (in fact an articulation point), or a block of the $C_{\mathcal{B}}$-graph. Now a structure s belonging to the left-hand-side of (48) is a $C_{\mathcal{B}}$-graph which is rooted at either a vertex or a block (a cell). It can happen that the rooting is performed right at the center. This is canonically equivalent to doing nothing and is represented by the first term in the right-hand-side of (48). On the other hand, if the rooting is done at an off-center cell, a vertex or a block, then there is a unique incident cell of the other kind (a block for a vertex and vice-versa) which is located towards the center, thus defining a unique $C_{\mathcal{B}}^{\bullet\diamondsuit}$-structure. It is easily checked that this correspondence is bijective and independent of any labelling, giving the desired species isomorphism.

For (49), it suffices to verify the species identities (isomophisms) $C_{\mathcal{B}}^{\diamondsuit} = \mathcal{B}(C_{\mathcal{B}}^{\bullet})$ and $C_{\mathcal{B}}^{\bullet\diamondsuit} = C_{\mathcal{B}}^{\bullet} \cdot \mathcal{B}'(C_{\mathcal{B}}^{\bullet})$. Details are left as an exercise. $\square$

The method of enumeration of unlabelled $C_{\mathcal{B}}$-graphs then consists in first enumerating the rooted unlabelled $C_{\mathcal{B}}$-graphs using the functional equation (8) or (12), and then enumerating the unrooted ones, using (49). Notice that these steps will lead not to explicit but rather to recursive formulas for the desired numbers or total weights. Below we illustrate the method in detail in the case of triangular cacti (Example 1.1.4). We will also give some results for Husimi graphs and oriented cacti.

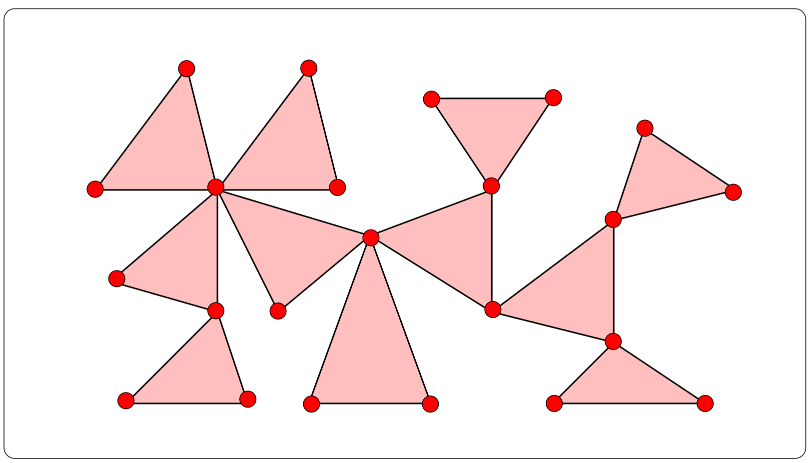

Figure 9: Unlabelled triangular cactus

Recall that a triangular cactus is a connected graph all of whose blocks are triangles. See Figure 9 for an example. These structures (and also the quadrangular cacti) were enumerated by Harary, Norman, and Uhlenbeck [9], [11]. Here we go further by giving recurrence formulas for their numbers. Let $\delta = C_{\mathcal{B}}$ and $\Delta = \delta^{\bullet}$ denote the species of triangular cacti and of rooted triangular cacti, respectively. Also set

$$\widetilde{\delta}(x) = \sum_{n\geq 1} d_n x^n \quad \text{and} \quad \widetilde{\Delta}(x) = \sum_{n\geq 1} D_n x^n, \tag{50}$$

where d_n and D_n denote the numbers of unlabelled triangular cacti and rooted triangular cacti, respectively. Here we have $\mathcal{B} = K_3 = E_3$, and $\mathcal{B}' = E_2$. The functional equations (8) and (49) give

$$\Delta = X\, E(E_2(\Delta)) \tag{51}$$

and

$$\Delta + E_3(\Delta) = \delta + \Delta\, E_2(\Delta), \tag{52}$$

respectively. Note that

$$\begin{aligned} Z_{E\circ E_2}(x_1, x_2, x_3, \ldots) &= Z_E(Z_{E_2}(x_1, x_2, \ldots), Z_{E_2}(x_2, x_4, \ldots), \ldots) \\ &= Z_E(\frac{1}{2}(x_1^2 + x_2), \frac{1}{2}(x_2^2 + x_4), \ldots) \\ &= \exp\left(\sum_{k\geq 1} \frac{1}{k} \frac{(x_k^2 + x_{2k})}{2}\right). \end{aligned}$$

From (51), we deduce (see [11])

$$\begin{aligned} \widetilde{\Delta}(x) &= (XE(E_2(\Delta)))^{\sim}(x) \\ &= xZ_{E\circ E_2}(\widetilde{\Delta}(x), \widetilde{\Delta}(x^2), \widetilde{\Delta}(x^3), \ldots) \\ &= x\exp\left(\sum_{k\geq 1} \frac{1}{2k}\left(\widetilde{\Delta}^2(x^k) + \widetilde{\Delta}(x^{2k})\right)\right). \end{aligned} \tag{53}$$

A recurrence formula for D_n can be obtained by taking the derivative of (53) times x. Set

$$b(x) = x\frac{d}{dx}\sum_{k\geq 1} \frac{1}{2k}\left(\widetilde{\Delta}^2(x^k) + \widetilde{\Delta}(x^{2k})\right).$$

We then have $x\widetilde{\Delta}'(x) = \widetilde{\Delta}(x) + \widetilde{\Delta}(x)b(x)$ and

$$x\widetilde{\Delta}'(x) - \widetilde{\Delta}(x) = \widetilde{\Delta}(x)b(x). \tag{54}$$

But

$$\begin{aligned} b(x) &= x\frac{d}{dx}\sum_{k\geq 1} \frac{1}{2k}\left(\widetilde{\Delta}^2(x^k) + \widetilde{\Delta}(x^{2k})\right) \\ &= x \cdot \sum_{k\geq 1} \frac{1}{2k}\left(2\widetilde{\Delta}(x^k)\widetilde{\Delta}'(x^k)kx^{k-1} + \widetilde{\Delta}'(x^{2k})2kx^{2k-1}\right) \\ &= \sum_{k\geq 1}\left(\widetilde{\Delta}(x^k)\widetilde{\Delta}'(x^k)x^k + \widetilde{\Delta}'(x^{2k})x^{2k}\right) \\ &= \sum_{k\geq 1}\left(\sum_{h\geq 1}\sum_{j\geq 1} jD_hD_jx^{k(h+j)} + \sum_{j\geq 1} jD_jx^{2kj}\right) \\ &= \sum_{m\geq 1}\left(\sum_{\substack{d\mid m\\ d\geq 2}}\sum_{j=1}^{d} jD_{d-j}D_j + \sum_{\substack{d\mid m\\ d\ \text{even}}} \frac{d}{2}D_{\frac{d}{2}}\right)x^m. \end{aligned}$$

By extracting the coefficient of x^{n+1} in (54), we obtain

Proposition 3.8 *The numbers D_n, of unlabelled rooted triagular cacti on n vertices satisfy $D_1 = 1$ and the recurrence formula, for $n \geq 1$,*

$$D_{n+1} = \frac{1}{n}\sum_{m=1}^{n} D_{n-m+1}\left(\sum_{\substack{d|m\\ d\geq 2}}\sum_{j=1}^{d} jD_{d-j}D_j + \sum_{\substack{d|m\\ d\ \ even}}\frac{d}{2}D_{\frac{d}{2}}\right). \tag{55}$$

In order to enumerate unlabelled *unrooted* triangular cacti, we use (49), that is,

$$\delta = \Delta + E_3(\Delta) - \Delta E_2(\Delta).$$

The passage to the tilde generating functions yields (see [9])

$$\begin{aligned}
\widetilde{\delta}(x) &= \widetilde{\Delta}(x) + (E_3(\Delta(x)))^{\sim} - \widetilde{\Delta}(x)(E_2(\Delta(x)))^{\sim} \\
&= \widetilde{\Delta}(x) + Z_{E_3}(\widetilde{\Delta}(x), \widetilde{\Delta}(x^2), \ldots) - \widetilde{\Delta}(x)Z_{E_2}(\widetilde{\Delta}(x), \widetilde{\Delta}(x^2), \ldots) \\
&= \widetilde{\Delta}(x) + \frac{1}{6}(\widetilde{\Delta}^3(x) + 3\widetilde{\Delta}(x)\widetilde{\Delta}(x^2) + 2\widetilde{\Delta}(x^3)) - \frac{1}{2}\widetilde{\Delta}(x)(\widetilde{\Delta}^2(x) + \widetilde{\Delta}(x^2)) \\
&= \widetilde{\Delta}(x) + \frac{1}{3}\left(\widetilde{\Delta}(x^3) - \widetilde{\Delta}^3(x)\right). \qquad (56)
\end{aligned}$$

By extracting the coefficient of x^n, we finally obtain

Proposition 3.9 *The numbers d_n, of unlabelled (unrooted) triangular cacti on n vertices satisfy, for $n \geq 1$,*

$$d_n = D_n + \frac{1}{3}\left(\chi(3|n)D_{\frac{n}{3}} - \sum_{i+j+k=n} D_iD_jD_k\right). \tag{57}$$

We now turn our attention to the species Hu of Husimi graphs. Let us denote by h_n and H_n, the numbers of unlabelled Husimi graphs and rooted Husimi graphs, respectively, and set

$$h(x) = \sum_{n\geq 1} h_n x^n = \widetilde{\mathrm{Hu}}(x) \quad \text{and} \quad H(x) = \sum_{n\geq 1} H_n x^n = \widetilde{\mathrm{Hu}^{\bullet}}(x). \tag{58}$$

Here, as seen in Section 3.1, $\mathcal{B} = E_{\geq 2}$ and $\mathcal{B}' = E_{\geq 1} = E_+$, and the basic functional equation (40) translates, for the tilde generating function, into (see [25], p. 51)

$$H(x) = x\exp\sum_{k\geq 1}\frac{1}{k}\left(\exp\left(\sum_{m\geq 1}\frac{1}{m}H(x^{mk})\right) - 1\right). \tag{59}$$

We introduce the auxiliary series

$$b(x) = \sum_{n\geq 1} b_n x^n = x\frac{d}{dx}\sum_{k\geq 1}\frac{1}{k}\left(\exp\left(\sum_{m\geq 1}\frac{1}{m}H(x^{mk})\right) - 1\right), \tag{60}$$

$$\varphi(x) = \sum_{n\geq 1}\varphi_n x^n = x\exp\left(\sum_{j\geq 1}\frac{1}{j}\,\mathrm{H}(x^j)\right). \tag{61}$$

Then, after some computations similar to those for triangular cacti, we find:

Proposition 3.10 *The numbers H_n, of unlabelled rooted Husimi graphs on n vertices can be computed by the following recursive scheme:* $\varphi_1 = 1$, $H_1 = 1$, *and, for* $n \geq 1$,

$$\begin{aligned} b_n &= \sum_{d|n} \sum_{h=1}^{d} \sum_{l|d-h+1} lH_l \ \varphi_h, &(62)\\ H_{n+1} &= \frac{1}{n} \sum_{k=1}^{n} H_{n-k+1} b_k, &(63)\\ \varphi_{n+1} &= \frac{1}{n} \sum_{m=1}^{n} \varphi_{n-m+1} \sum_{d|m} dH_d. &(64) \end{aligned}$$

For unlabelled *unrooted* Husimi graphs, we use the Dissymmetry Theorem, in the form (49), which gives

$$\begin{aligned} \mathrm{Hu}^{\bullet} + E_{\geq 2}(\mathrm{Hu}^{\bullet}) &= \mathrm{Hu} + \mathrm{Hu}^{\bullet} \cdot E_{+}(\mathrm{Hu}^{\bullet}),\\ E_{+}(\mathrm{Hu}^{\bullet}) &= \mathrm{Hu} + \mathrm{Hu}^{\bullet} \cdot E_{+}(\mathrm{Hu}^{\bullet}), \end{aligned}$$

and finally

$$\mathrm{Hu} = (E_{+}(\mathrm{Hu}^{\bullet})) \cdot (1 - \mathrm{Hu}^{\bullet}). \tag{65}$$

For the tilde generating series, we deduce that

$$\begin{aligned} h(x) &= \left(\exp\left(\sum_{k\geq 1} \frac{1}{k} H(x^k) \right) - 1 \right) \cdot (1 - H(x))\\ &= \left(\frac{\varphi(x)}{x} - 1 \right) \cdot (1 - H(x)), \end{aligned}$$

and we obtain the following result.

Proposition 3.11 *The numbers h_n, of unlabelled Husimi graphs on n vertices satisfy*

$$h_n = \varphi_{n+1} - \sum_{k=1}^{n-1} \varphi_{k+1} H_{n-k}, \tag{66}$$

where the numbers H_n and φ_n are given by Proposition 3.10.

The same method can be applied to weighted $C_{\mathcal{B}}$-graphs where the weight is defined by (11), that is, with the variables $(y_2, y_3, \ldots)$ marking the block sizes. We have done the computations for the weighted species Oc_w, of oriented cacti, where $\mathcal{B}_w = \sum_{m\geq 2} (C_m)_{y_m}$ and $\mathcal{B}'_w = \sum_{m\geq 1} (L_m)_{y_{m+1}}$, where $(C_m)_{y_m}$ denotes the species of oriented cycles of length m and weight y_m, and $L_m = X^m$ is the species of lists of size m, endowed here with the weight y_{m+1}. Introduce the following tilde generating series, whose coefficients are polynomials in the variables $\mathbf{y} = (y_2, y_3, \ldots)$:

$$\begin{aligned} o(x;\mathbf{y}) := \widetilde{\mathrm{Oc}_w}(x) &= \sum_{n\geq 1} o_n(\mathbf{y})x^n, \\ O(x;\mathbf{y}) := \widetilde{\mathrm{Oc}_w^\bullet}(x) &= \sum_{n\geq 1} O_n(\mathbf{y})x^n. \end{aligned}$$

The two basic species functional equations are

$$\mathrm{Oc}_w^\bullet = X \cdot E\left(\sum_{m\geq 1}(L_m)_{y_{m+1}}(\mathrm{Oc}_w^\bullet)\right), \tag{67}$$

$$C_w(\mathrm{Oc}^\bullet) = \mathrm{Oc} + L_{\geq 2,w}(\mathrm{Oc}^\bullet), \tag{68}$$

where $C_w = \sum_{m\geq 1}(C_m)_{y_m}$, with $C_1 = X,\ y_1 = 1$, and $L_{\geq 2,w} = \sum_{m\geq 2}(X^m)_{y_m}$. Note that

$$Z_{C_w} = \sum_{m\geq 1}\frac{y_m}{m}\sum_{d|m}\phi(d)x_d^{\frac{m}{d}} = \sum_{d\geq 1}\sum_{h\geq 1}\frac{y_{dh}}{dh}\phi(d)x_d^h,$$

where ϕ is Euler's totient function, and that $Z_{L_{\geq 2,w}} = \sum_{m\geq 2} y_m x_1^m$. From (67), we deduce, using the plethystic composition rule for the tilde generating function of weighted species (see (4.3.1) of [2]),

$$O(x;\mathbf{y}) = x\exp\left(\sum_{k\geq 1}\frac{1}{k}\sum_{j\geq 1}y_{j+1}^k O^j(x^k;\mathbf{y}^k)\right), \tag{69}$$

where $\mathbf{y}^k = (y_2^k, y_3^k, \ldots)$. We then set

$$O^j(x;\mathbf{y}) = (O(x;\mathbf{y}))^j = \sum_{n\geq 1} O_n^{(j)}(\mathbf{y})x^n, \tag{70}$$

$$b(x;\mathbf{y}) = \sum_{m\geq 1} b_m(\mathbf{y})x^m = x\frac{\partial}{\partial x}\sum_{k\geq 1}\frac{1}{k}\sum_{j\geq 1}y_{j+1}^k O^j(x^k;\mathbf{y}^k). \tag{71}$$

From (68), we deduce

$$o(x;\mathbf{y}) = \sum_{d\geq 1}\phi(d)\sum_{h\geq 1}\frac{y_{dh}}{dh}O(x^d;\mathbf{y}^d) - \sum_{m\geq 2} y_m O(x;\mathbf{y})^m \tag{72}$$

and we obtain, after some computations, the following result.

Proposition 3.12 *The generating polynomials $O_n(\boldsymbol{y})$ and $o_n(\boldsymbol{y})$ for unlabelled rooted and unrooted (resp.) oriented cacti on n vertices are given by the following recursive scheme:*

$$O_{n+1}(\boldsymbol{y}) = \frac{1}{n}\sum_{m=1}^{n} O_{n-m+1}(\boldsymbol{y})b_m(\boldsymbol{y}), \tag{73}$$

where

$$b_m(\boldsymbol{y}) = \sum_{d|m} \sum_{i=1}^{d} \sum_{j=1}^{d-i+1} ij \; y_{j+1}^{\frac{m}{d}} O_i(\boldsymbol{y}^{\frac{m}{d}}) O_{d-i}^{(j-1)}(\boldsymbol{y}^{\frac{m}{d}}), \tag{74}$$

and

$$o_n(\boldsymbol{y}) = \sum_{d|n} \phi(d) \sum_{h=1}^{\frac{n}{d}} \frac{y_{dh}}{dh} O_{\frac{n}{d}}^{(h)}(\boldsymbol{y}^d) - \sum_{m=2}^{n} y_m O_n^{(m)}(\boldsymbol{y}). \tag{75}$$

4 Molecular expansions

The molecular expansion of a species F is a description and a classification of the unlabelled F-structures according to their stabilizers, within the language of species. It is interesting and useful to have at hand the first few terms of the molecular expansion of $C_{\mathcal{B}}$-graphs. This is possible with the Maple package Devmol, developed at LaCIM; see [1].

As an example, consider again the species Hu_w of Husimi graphs, weighted according to their block-size distribution. Here $\mathcal{B}_w = E_{\geq 2,w} = \sum_{k\geq 2}(E_k)_{y_k}$. This is the species BB in the following extract from a Maple session, which uses the package Devmol. The Devmol procedure "`CBgraphes`(BB,n)" produces the molecular expansion of the species of $\mathcal{C}_{BB}$-graphs, up to the given truncation order n. The expansion is then collected by degree (vertex number). The weight variables $y_2, y_3, \ldots, y_n$ are first declared to Devmol. Afterwards, the monomial weights in the y_i's appear multiplicatively in the expressions. Here is the extract:

```
> n := 6;
```

$$n := 6$$

```
> ajoutvv(seq(y[k], k = 1..n));
```

$$\{t,\, y_1,\, y_2,\, y_3,\, y_4,\, y_5,\, y_6\}$$

```
> BB := sum(y[k]*E[k](X), k = 2..n);
```

$$BB := y_2\, E_2(X) + y_3\, E_3(X) + y_4\, E_4(X) + y_5\, E_5(X) + y_6\, E_6(X)$$

```
> HuW := CBgraphes(BB,n):
> affichertable(tablephom(HuW));
```

$$1$$
$$X$$
$$2$$
$$y_2\, E_2(X)$$

$$3$$
$$X\,{y_2}^2\,E_2(X)+y_3\,E_3(X)$$
$$4$$
$$y_2\,X^2\,y_3\,E_2(X)+X\,{y_2}^3\,E_3(X)+{y_2}^3\,E_2(X^2)+y_4\,E_4(X)$$
$$5$$
$$(y_3\,{y_2}^2+{y_2}^4)\,X^3\,E_2(X)+X\,{y_2}^2\,E_2(X)^2\,y_3+(y_3\,{y_2}^2+{y_2}^4)\,X\,E_2(X^2)$$
$$+X\,{y_2}^4\,E_4(X)+X\,{y_3}^2\,E_2(E_2(X))+y_5\,E_5(X)+y_2\,X^2\,y_4\,E_3(X)$$
$$6$$
$${y_2}^5\,X^2\,E_2(X^2)+({y_2}^2\,y_4+y_3\,y_4+y_3\,{y_2}^3)\,X\,E_2(X)\,E_3(X)$$
$$+y_4\,E_2(X)\,{y_2}^2\,E_2(X^2)+({y_3}^2\,y_2+{y_2}^5+3\,y_3\,{y_2}^3)\,X^4\,E_2(X)$$
$$+{y_2}^5\,E_2(X^3)+({y_3}^2\,y_2+{y_2}^5)\,E_2(X\,E_2(X))+y_2\,{y_3}^2\,X^2\,E_2(E_2(X))$$
$$+y_3\,{y_2}^3\,E_3(X^2)+y_2\,y_5\,X^2\,E_4(X)+y_6\,E_6(X)+y_3\,{y_2}^3\,X^2\,E_2(X)^2$$
$$+\,y_3\,{y_2}^3\,X^6+({y_2}^5+{y_2}^2\,y_4)\,X^3\,E_3(X)+{y_2}^5\,X\,E_5(X)$$

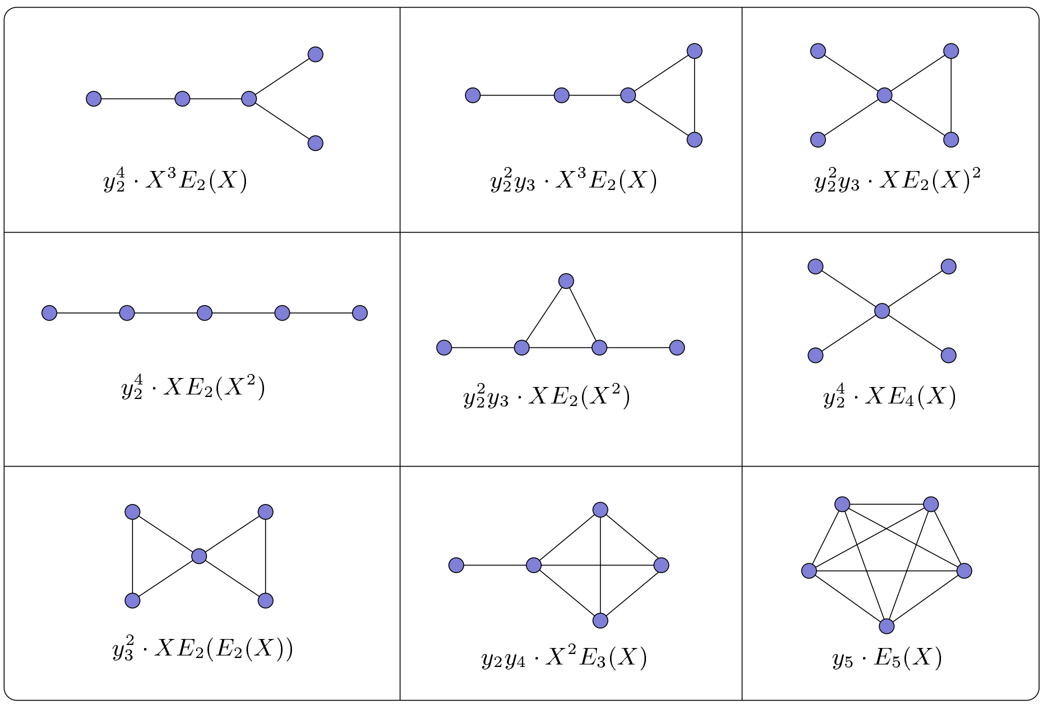

Figure 10: Husimi graphs of size 5

As an illustration, the nine terms of size 5 are shown in Figure 10. Observe that multiciplicities will frequently occur in these expansions. For size 6, for example, the term $3\,y_3\,y_2^3\,X^4E_2(X)$ corresponds to the 3 similar types of Husimi graphs shown in Figure 11.

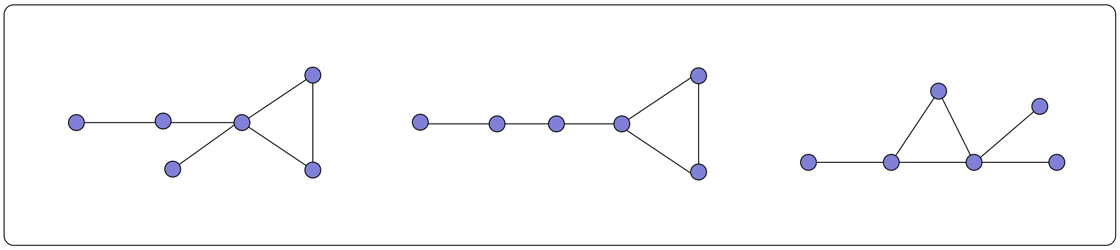

Figure 11: The three unlabelled Husimi graphs of molecular type $y_2^3 y_3 \cdot X^4 E_2(X)$

References

[1] P. Auger, P. Leroux and G. Labelle, *Computing the molecular expansion of species with the Maple Package Devmol*, Séminaire Lotharingien de Combinatoire, B49z (2003), 34 pp. `http://euler.univ-lyon1.fr/home/slc`

[2] F. Bergeron, G. Labelle and P. Leroux, *Combinatorial Species and Tree-Like Structures*, Coll. Encyclopedia of mathematics and its applications, Vol. 67, Cambridge University Press, 1998, 457 p.

[3] M. Bóna, M. Bousquet, G. Labelle and P. Leroux, *Enumeration of m-ary cacti*, Advances in Applied Mathematics, 24, 2000, 22–56.

[4] Lisa Fleischer, *Building chain and cactus representations of all minimum cuts from Hao-Orlin in the same asymptotic run time*, J. Algorithms, 33, 1999, 51–72.

[5] G. W. Ford and G. E. Uhlenbeck, *Combinatorial Problems in the Theory of Graphs, I*, Proc. Nat. Acad. Sciences, 42 (1956), 122–128.

[6] G. W. Ford, R. Z. Norman and G. E. Uhlenbeck, *Combinatorial Problems in the Theory of Graphs, II*, Proc. Nat. Acad. Sciences, 42 (1956), 203–208.

[7] T. Fowler, I. Gessel, G. Labelle and P. Leroux, *The specification of 2-tress*, Advances in Applied Mathematics, 28, 2002, 145–168.

[8] F. Harary, *Graph Theory*, Addison-Wesley, 1969, 274 p.

[9] F. Harary and R. Z. Norman. *The dissimilarity characteristic of Husimi trees*, Ann. of Math., 58, 1953, 134–141.

[10] F. Harary and E. M. Palmer. *Graphical Enumeration*, Academic Press, New York and London, 1973, 271 p.

[11] F. Harary and G. E. Uhlenbeck. *On the number of Husimi trees*, Proc. Nat. Aca. Sci., 39, 1953, 315–322.

[12] Kodi Husimi. *Note on Mayers' theory of cluster integrals*, The Journal of Chemical Physics, 18, 1950, 682–684.

[13] A. Joyal, *Une théorie combinatoire des séries formelles*, Advances in Mathematics, 42, 1981, 1–82.

[14] G. Labelle, *Counting asymmetric enriched trees*, Journal of Symbolic Computation, 14, 1992, 211–242.

[15] G. Labelle, C. Lamathe, P. Leroux, *Énumération des 2-arbres k-gonaux*, Proceedings of the conference Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities, Versailles, September 16–19, 2002, in *Mathematics and Computer Science II*, B. Chauvin, P. Flajolet, D. Gardy, and A. Mokkadem, ed., Birkhäuser, Basel, 2002, 95–109.

[16] G. Labelle, C. Lamathe and P. Leroux, *A classification of plane and planar 2-trees*, Theoretical Computer Science, 307 (2003), 337–363.

[17] G. Labelle, C. Lamathe, P. Leroux, *Enumeration k-gonal 2-trees*, Journal of Combinatorial Theory, Series A, Article in Press, 2004.

[18] J. Labelle, *Applications diverses de la théorie combinatoire des espèces de structures*, Annales des Sciences Mathématiques du Québec, 7, 1983, 59–94.

[19] P. Leroux, *Methoden der Anzahlbestimmung für einige Klassen von Graphen*, Bayreuther Mathematische Schriften, 26, 1988, 1–36.

[20] P. Leroux and B. Miloudi *Généralisations de la formule d'Otter*, Annales des Sciences Mathématiques du Québec, 16, 1992, 53–80.

[21] F. Maffioli and N. Zagaglia Salvi, *A characterization of the base-matroids of a graphic matroid*, Preprint, 2003.

[22] J. E. Mayer and M. G. Mayer, *Statistical Mechanics*, Wiley, New York, 1940.

[23] J. E. Mayer, *Equilibrium Statistical Mechanics*, The international encyclopedia of physical chemistry and chemical physics, Pergamon Press, Oxford, 1968, 242 p.

[24] M. Nadeau, *Graphes de Husimi, Cactus, et mécanique statistique*. Mémoire de maîtrise, UQAM, 2002.

[25] R. Z. Norman, *On the number of linear graphs with given blocks*, Dissertation, University of Michigan, 1954.

[26] R. Otter, *The number of trees*, Ann. of Math, 49, 1948, 583–599.

[27] R. J. Riddell . *Contributions to the theory of condensation*, Dissertation, University of Michigan, Ann Arbor, 1951.

[28] R. W. Robinson *Enumeration of non-separable graphs*, Journal of Combinatorial Theory, 9, 1970, 327–356.

[29] Colin Springer, *Factorizations, Trees, and Cacti*, Proceedings of the Eighth International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC), University of Minnesota, 1996, 427–438.

[30] G. E. Uhlenbeck et G. W. Ford, *Lectures in Statistical Mechanics*, Amer. Math. Soc., Providence, Rhode Island, 1963, 181 p.