Enumeration of externally labelled trees

1991 J. Phys. A: Math. Gen. 24 929

(http://iopscience.iop.org/0305-4470/24/5/008)

J. Phys. A: Math. Gen. 24 (1991) 929–936. Printed in the UK

Enumeration of externally labelled trees

M J A M Brummelhuis and H Kuijf
Instituut-Lorentz, Rijksuniversiteit te Leiden, PO Box 9506, 2300 RA Leiden, The Netherlands

Received 24 September 1990

Abstract. We enumerate a special class of trees, called externally labelled trees. This is done by formulating a recursion relation in two variables, the number of internal lines, i, and the number of external lines, n, for these trees, and subsequently solving it. The solution is discussed in detail. In particular we consider the case where n becomes large, while summing on i. The externally labelled trees considered correspond to the Feynman diagrams one encounters in perturbation theories for fundamental forces, involving one type of particle that self-interacts.

1. Introduction and problem

We begin by defining the notion of an externally labelled tree. An externally labelled tree is a connected graph without cycles, of which the external branches are labelled. Here, an external branch is understood to be a line arriving in a point in which no other lines arrive. The lines in the graph which are not external will be called 'internal branches'. Internal and external branches are connected at given points, called vertices. The minimum number of branches that can come together in a vertex is three, the maximum number is m (≥ 3), an adjustable integer valued parameter. Both the internal branches and the vertices are unlabelled. External branches connected to the same vertex can be interchanged without producing a different externally labelled tree. We point out that for a given externally labelled tree, letting N_k denote the number of vertices where k branches ($k = 3, \ldots, m$) come together, n the number of external branches, and i the number of internal ones, one has

$$\sum_{k=3}^{m} N_k = i + 1 \tag{1.1}$$

and

$$\sum_{k=3}^{m} kN_k = n + 2i\,. \tag{1.2}$$

The problem we shall address in this paper is to calculate how many different externally labelled trees there exist for given, but general, m and n. The number of internal branches i at given, fixed, m and n is allowed to vary. To our knowledge, this question of enumeration has not been studied before in the literature, for general m. The cases $m = 3$ and $m = 4$ have already received some attention from other authors [1,2]. For some interesting general texts on graph theory see, e.g., [3, 4].

0305-4470/91/050929+08$03.50

The problem of enumerating externally labelled trees is of some physical relevance from the point of view of high energy physics. This is because externally labelled trees as defined above are just the Feynman diagrams one encounters in a perturbation theory for fundamental forces, involving only one type of particle that self-interacts. There, in order to calculate a scattering process of n of these particles, one needs to have all the diagrams with n external branches. It is thus important to know how many of them exist. The parameter m introduced above, controls the degree of self-interaction, and basically determines which theory one is studying. For example, $m = 3$ corresponds to Φ^3 theory, $m = 4$ to Yang–Mills theory, whereas the limit $m \to \infty$ corresponds to so-called linearized gravitation. In fact, these three cases are the most important ones from the point of view of high energy physics, since theories corresponding to other values of m are as yet unknown. Therefore we shall pay special attention to them in the following. In practical calculations performed up till now, n has always been less than 10. The record for the largest number of diagrams calculated presently lies at 34 300, being the case $m = 4$ and $n = 8$ [5].

The outline of the paper is as follows. We first formulate recursion relations, expressing the number of externally labelled trees with n external and i internal branches in the number of externally labelled trees with lesser branches. Then these recursion relations are solved, leading to a formal solution of the problem stated above. The solution is subsequently studied in more detail. In particular, we address the question as to how the number of externally labelled trees grows with n, at fixed m, as n becomes large, thus indicating the impossible task one faces when trying to calculate all Feynman diagrams for larger n.

2. Enumeration

2.1. Recursion relations and solution

Let $D_n^{(m)}$ denote the number of externally labelled trees, with n external branches, there exist when allowing $3, 4, \ldots, m$-point vertices to be present. With $D_{n,i}^{(m)}(N_3, \ldots, N_m)$ denoting the number of these trees with i internal lines and *exactly* N_3 3-point vertices, N_4 4-point vertices, up to and including N_m m-point vertices, we can write

$$D_n^{(m)} = \sum_{i=0}^{\infty} \sum_{N_3=0}^{\infty} \cdots \sum_{N_m=0}^{\infty} D_{n,i}^{(m)}(N_3, \ldots, N_m) \,. \tag{2.1}$$

For given n, the right-hand side of (2.1) converges, since $D_{n,i}^{(m)}$ only differs from zero if conditions (1.1) and (1.2) are fulfilled. It is easy to write down a recursion relation for $D_{n,i}^{(m)}(N_3, \ldots, N_m)$, valid for $n \geq 4$ and $i \geq 1$:

$$D_{n,i}^{(m)}(N_3, \ldots, N_m) = (n + i - 2) D_{n-1,i-1}^{(m)}(N_3 - 1, N_4, \ldots, N_m) + \sum_{k=3}^{m-1} (N_k + 1) D_{n-1,i}^{(m)}(N_3, \ldots, N_k + 1, N_{k+1} - 1, \ldots, N_m) \,. \tag{2.2}$$

This is so because an externally labelled tree with n external and i internal branches can be formed in two different ways. Namely, by taking an externally labelled tree with $(n-1)$ external and $(i-1)$ internal branches, then choosing a point on any one of the already existing internal or external branches and attaching the new external branch to it. This creates a new 3-vertex (the first term on the right-hand side of (2.2)). Or by taking an externally labelled tree with $(n-1)$ external and i internal branches, attaching the new external branch to an already existing vertex (the second term on the right-hand side of (2.2)). In the procedure described here, every externally labelled tree one can think of can be constructed. Also it is easy to convince oneself that there is no double counting. This is because, due to the fact that the external branches are labelled, the procedure itself has a tree structure. One can see this remarking that for a given externally labelled tree one can find its predecessor by just removing the external branch with the highest number, so that its predecessor is unique. Therefore the solution of (2.2) is indeed what we are looking for. However, in order for $D_{n,i}^{(m)}$ to be uniquely determined for all n and i, the recursion relation (2.2) must be supplemented with boundary conditions. We have

$$\begin{aligned} &D_{3,0}^{(m)}(1,0,\ldots,0)=1 \\ &D_{3,i}^{(m)}(N_3,\ldots,N_m)=0 \qquad \text{otherwise} \end{aligned} \tag{2.3}$$

expressing that there is only one externally labelled tree having three external branches and no internal ones, and, for $j=3,\ldots,m$

$$D_{j,0}^{(m)}(0,\ldots,0,N_j=1,0,\ldots,0)=1 \tag{2.4}$$

expressing that a j-vertex is unique. The solution for $D_{n,i}^{(m)}(N_3,\ldots,N_m)$ then reads

$$D_{n,i}^{(m)}(N_3,\ldots,N_m)=\frac{(\sum_{k=3}^{m}(k-1)N_k)!}{\prod_{k=3}^{m}[((k-1)!)^{N_k}N_k!]}\,\delta_{n-2,\sum_{j=3}^{m}(j-2)N_j}\,\delta_{i+1,\sum_{j=3}^{m}N_j} \tag{2.5}$$

as can be verified by a direct substitution in (2.2). It is easy to check that the boundary conditions (2.3) and (2.4) are indeed fulfilled. Substituting (2.5) in (2.1), the sum on i can be performed, yielding

$$D_n^{(m)}=\sum_{N_3=0}^{\infty}\cdots\sum_{N_m=0}^{\infty}\frac{(n-2+\sum_{k=3}^{m}N_k)!}{\prod_{k=3}^{m}[((k-1)!)^{N_k}N_k!]}\,\delta_{n-2,\sum_{j=3}^{m}(j-2)N_j}\,. \tag{2.6}$$

This, then, is the formal solution to the problem. However, since it still involves $(m-2)$ sums, it is not very transparent what it implies. In the following we shall therefore study some aspects of it more closely for some special cases. In particular, we shall study the growth factor $\Xi^{(m)}$, defined as

$$\Xi^{(m)}=\lim_{n\to\infty}\frac{D_n^{(m)}}{n\,D_{n-1}^{(m)}} \tag{2.7}$$

and show that it is finite for all m. The implications of the finiteness of $\Xi^{(m)}$ for $D_n^{(m)}$ itself will be discussed later.

2.2. The case m = 3

For $m = 3$, the sum in (2.6) in fact only consists of one non-zero term, and we find

$$\begin{aligned} D_n^{(3)} &= (2n-5)!! && \text{all } n \geq 3 \\ &\simeq \frac{\sqrt{2}}{\mathrm{e}^2}\left(\frac{2n}{\mathrm{e}}\right)^{n-2} && n \to \infty \end{aligned} \tag{2.8}$$

in accordance with [1]. From this one obtains

$$\Xi^{(3)} = 2\,. \tag{2.9}$$

2.3. The case m = 4

Performing the sum on N_3 in (2.6) we arrive at

$$D_n^{(4)} = \sum_{N_4=0}^{\mathrm{int}[(n/2)-1]} \frac{(2n-N_4-4)!}{2^{n-2-2N_4}(n-2N_4-2)!\,6^{N_4}N_4!} \qquad n \geq 3 \tag{2.10}$$

where 'int' stands for 'integer part of'. Let us write $N_4 = \alpha n$, thus introducing the new variable α. Then we have

$$D_n^{(4)} = \sum_\alpha \frac{[(2-\alpha)n-4]!}{2^{(1-2\alpha)n-2}[(1-2\alpha)n-2]!\,6^{\alpha n}(\alpha n)!} \tag{2.11}$$

where α now takes the values $0, (1/n), \ldots, (1/n)\mathrm{int}[(n/2)-1]$. In order to evaluate this sum asymptotically, we use Stirling's formula in the form

$$(an+b)! \simeq \sqrt{2\pi a n}\,(an)^{an+b}\,\mathrm{e}^{-an} \tag{2.12}$$

for the factorials in the summand, and note that $(1/n)\sum_\alpha \to \int_0^{1/2} \mathrm{d}\alpha$ as $n \to \infty$. We then obtain

$$D_n^{(4)} \simeq n^{n-3/2} \int_0^{1/2} \mathrm{d}\alpha\, f(\alpha)\,[g(\alpha)]^n \tag{2.13}$$

with

$$f(\alpha) = \frac{4(1-2\alpha)^{3/2}}{(2\pi\alpha)^{1/2}\,(2-\alpha)^{7/2}} \tag{2.14}$$

and

$$g(\alpha) = \frac{(2-\alpha)^{2-\alpha}}{2^{1-2\alpha}(1-2\alpha)^{1-2\alpha}6^\alpha\alpha^\alpha\,\mathrm{e}}\,. \tag{2.15}$$

By considering $\ln g(\alpha)$ one easily shows that $g(\alpha)$ has a maximum on $[0, \frac{1}{2}]$ for $\alpha = \alpha_0 = \frac{1}{11}(7-3\sqrt{3})$, the value at the maximum being $g(\alpha_0) = (3/11\mathrm{e})(4+3\sqrt{3})$. Taylor expanding the integrand in (2.13) around α_0 and then introducing a new variable x according to $\alpha = \alpha_0 + x/\sqrt{n}$ we find

$$D_n^{(4)} \simeq f(\alpha_0)\sqrt{\frac{2\pi g(\alpha_0)}{-g''(\alpha_0)}}\,[g(\alpha_0)]^n n^{n-2}\,. \tag{2.16}$$

Now

$$-\frac{g''(\alpha_0)}{g(\alpha_0)} = -\left(\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2}\ln g(\alpha)\right)_{\alpha_0} = \frac{1}{\alpha_0} - \frac{1}{2-\alpha_0} + \frac{4}{1-2\alpha_0} = 4 + \frac{13}{\sqrt{3}} \tag{2.17}$$

so that one ultimately gets

$$D_n^{(4)} \simeq \frac{1}{\mathrm{e}}\sqrt{\frac{g(\alpha_0)}{\mathrm{e}\sqrt{3}}}\,[g(\alpha_0)\,n]^{n-2} = (0.16285\cdots)[0.92265\cdots n]^{n-2}\,. \tag{2.18}$$

The growth factor as defined in (2.7) is given by

$$\Xi^{(4)} = \mathrm{e}\,g(\alpha_0) = \tfrac{3}{11}(4+3\sqrt{3}) \doteq 2.5080416\cdots\,. \tag{2.19}$$

2.4. The general case

Finding an asymptotic expression for $D_n^{(m)}$ as $n \to \infty$ in this case, would amount to starting with (2.6), then performing, e.g., the sum on N_3 to eliminate the Kronecker delta, and subsequently applying the method of steepest descent on an $(m-3)$-fold integral. This is a procedure that is very similar to the one sketched above for the case $m=4$. However, as it turns out, it is already reasonably difficult to find where the maximum used in the method of steepest descent is located. Therefore, we shall at first content ourselves with calculating $\Xi^{(m)}$, since here a knowledge of the maximum itself suffices, as we have already seen in the case $m=4$. To do so we return to (2.5), which we rewrite as

$$D_{n,i}^{(m)}(N_3,\ldots,N_m) = \frac{(n+i-1)!}{\Pi_{k=3}^m[((k-1)!)^{N_k}N_k!]}\,\delta_{n-2,\sum_{j=3}^m(j-2)N_j}\,\delta_{i+1,\sum_{j=3}^m N_j}\,. \tag{2.20}$$

Obviously, what we are looking for are the values of i and $N_3,\ldots,N_m$ which maximize $(n+i-1)!/\Pi_{k=3}^m[((k-1)!)^{N_k}N_k!]$, in a space restricted by conditions (1.1) and (1.2). With this in mind we write

$$i = \gamma n \qquad N_k = \beta_k n \tag{2.21}$$

so that, using (2.12), we arrive at

$$\frac{(n+i-1)!}{\Pi_{k=3}^m[((k-1)!)^{N_k}N_k!]} \simeq \frac{f(\gamma,\{\beta_k\})}{n^{(m+1)/2}}\,[g(\gamma,\{\beta_k\})\,n]^n\,. \tag{2.22}$$

Here

$$g(\gamma,\{\beta_k\}) = \frac{(1+\gamma)^{1+\gamma}}{\mathrm{e}\,\Pi_{k=3}^m[(k-1)!\,\beta_k]^{\beta_k}} \tag{2.23}$$

while the expression for $f(\gamma,\{\beta_k\})$ will not be given, since it is not important for our purposes. Thus, we must maximize $g(\gamma,\{\beta_k\})$, or what is easier but amounts to the same, $\ln g(\gamma,\{\beta_k\})$, subject to the conditions

$$\sum_{j=3}^m \beta_j = \gamma \tag{2.24a}$$

$$\sum_{j=3}^m j\beta_j = 2\gamma + 1 \tag{2.24b}$$

following from (1.1) and (1.2) when using the scaling (2.21) and taking $n \to \infty$ at fixed γ and β_k. Therefore we consider

$$\Lambda_{\lambda,\mu}(\gamma,\{\beta_k\}) = \ln g(\gamma,\{\beta_k\}) - \lambda\left(\sum_{j=3}^{m}\beta_j - \gamma\right) - \mu\left(\sum_{j=3}^{m} j\beta_j - 2\gamma - 1\right) \tag{2.25}$$

where λ and μ are Lagrange multipliers, and we shall want to extremize this expression. Solving

$$\frac{\partial}{\partial\gamma}\Lambda_{\lambda,\mu} = \frac{\partial}{\partial\beta_k}\Lambda_{\lambda,\mu} = 0 \qquad \text{for all } k \tag{2.26}$$

one finds

$$\beta_k^0 = \frac{e^{\lambda+\mu k-1}}{(k-1)!} \tag{2.27a}$$

$$\gamma^0 = e^{\lambda+2\mu-1} - 1 \tag{2.27b}$$

which express γ^0 and β_k^0, the values at the extremum, in terms of λ and μ. The values of λ and μ are themselves obtained by substituting (2.27*a*) and (2.27*b*) back into (2.24*a*) and (2.24*b*), solving the resulting equations for λ and μ. In particular, we find that

$$\sum_{k=1}^{m-2}\frac{e^{\mu k}}{k!} = 1 \tag{2.28a}$$

$$\sum_{k=1}^{m-2}\frac{e^{\mu k}}{(k+1)!} = \frac{\gamma^0}{1+\gamma^0} \tag{2.28b}$$

implicitly giving the values of μ and γ^0. The growth factor $\Xi^{(m)}$ is found to be given by

$$\Xi^{(m)} = e\, g(\gamma^0,\{\beta_k^0\}) = (1+\gamma^0)\,e^{-\mu} = \left(2e^{\mu} - 1 - \frac{e^{\mu(m-1)}}{(m-1)!}\right)^{-1} \tag{2.29}$$

where the different equalities can be established using the definition (2.7) of $\Xi^{(m)}$ and formulas (2.22)–(2.28*b*). It is directly determined by that solution of (2.28*a*) for which $e^{\mu} \in [0,1]$. For example, we find

$$\begin{aligned}
\Xi^{(3)} &= 2\\
\Xi^{(4)} &= \tfrac{3}{11}(4+3\sqrt{3}) = 2.508\cdots\\
\Xi^{(5)} &= \left(2x - 1 - \frac{x^4}{24}\right)^{-1} \qquad \text{with } x = (5+\sqrt{26})^{1/3} - (5+\sqrt{26})^{-1/3} - 1 \\
&= 2.578\cdots\\
\Xi^{(\infty)} &= \frac{1}{2\ln 2 - 1} = 2.588\cdots .
\end{aligned} \tag{2.30}$$

Also, using (2.28a) and (2.29), it is not hard to show that $\Xi^{(\infty)}$ is approached as

$$\Xi^{(\infty)} - \Xi^{(m)} \simeq \frac{\sum_{k=m}^{\infty}[(\ln 2)^k/k!]}{(2\ln 2 - 1)^2} \qquad m \to \infty\,. \tag{2.31}$$

The existence of $\Xi^{(m)}$ implies that asymptotically, i.e., for $n \to \infty$

$$D_n^{(m)} \simeq H_m(n)\left(\Xi^{(m)}\right)^n n! \tag{2.32}$$

where $H_m(n)$ has the property that $\lim_{n\to\infty} H_m(n+1)/H_m(n) = 1$. Therefore we see that the number of externally labelled trees for fixed m and $n \to \infty$ grows factorially fast. However, we can even do better than this: we can find the n-dependence of $H_m(n)$ quite easily. The only thing one has to do is to follow the procedure outlined at the beginning of this section for calculating $D_n^{(m)}$ asymptotically, keeping track of powers of n. We stress that it is not necessary to do the whole actual calculation itself. It is not hard to convince oneself that one finds

$$D_n^{(m)} \sim n^{n-2} \tag{2.33}$$

as $n \to \infty$, for all m. Therefore we obtain

$$D_n^{(m)} \simeq C_m\left(\Xi^{(m)}\frac{n}{e}\right)^{n-2} \qquad n \to \infty \qquad \text{all } m \tag{2.34}$$

where C_m is an m-dependent constant. It is amusing to remark here that the number of ways in which one can join n labelled points to form a tree (also allowing only two lines to arrive in a given labelled point) is given exactly by n^{n-2} [6]. The result (2.34) is consistent with (2.32). One may check that the results (2.8) and (2.18) for the cases $m = 3$ and $m = 4$, respectively, are indeed of the form (2.34). Also, there, the constants C_3 and C_4 can be read off. The corrections to (2.34) are of relative order $1/n$. Because of this one can actually calculate $\Xi_n^{(m)} \equiv D_n^{(m)}/nD_{n-1}^{(m)}$ up to first order in $1/n$ using (2.34), since the corrections to (2.34) only contribute to second and higher orders in the large-n expansion of $\Xi_n^{(m)}$. We find

$$\Xi_n^{(m)} = \Xi^{(m)}\left[1 - \frac{5}{2n} + \cdots\right]. \tag{2.35}$$

Thus, the coefficient of the $1/n$ term in the expansion of $\Xi_n^{(m)}/\Xi^{(m)}$ in powers of $1/n$ does not depend on m. This, however, is no longer true for the higher-order terms. For the case $m = 3$ there are in fact no higher-order terms, as can be seen from the exact result in (2.8) valid for all $n \geq 3$. But a numerical evaluation of $\Xi_n^{(m)}/\Xi^{(m)}$ for the cases $m = 4$ and $m = \infty$ yields that then there are indeed higher-order corrections. They are equal to $k_m/n^2 + l_m/n^3 + \cdots$, say, and we find numerically

$$\begin{aligned} k_4 &= 0.015\,986\,086\,1860\cdots & l_4 &= 0.052\,271\,31\cdots \\ k_\infty &= 0.016\,095\,598\,3798\cdots & l_\infty &= 0.056\,334\,59\cdots\,. \end{aligned} \tag{2.36}$$

Because these coefficients are so small, one might expect that the approximation to $\Xi_n^{(m)}$ one obtains when neglecting the terms represented by the dots in (2.35), is already quite good for fairly small n. We have for $m = 4$ and $m = \infty$ explicitly verified that this is indeed the case.

Finally we remark that the case $m = \infty$ can in principle be tackled more directly. Here, namely, one can immediately write down a recursion relation for $E_n(i)$, the number of externally labelled trees with n external and i internal branches, since all types of vertices are allowed. It reads

$$E_n(i) = (n+i-2)E_{n-1}(i-1) + (i+1)E_{n-1}(i) \qquad n \geq 4 \qquad i \geq 1 \tag{2.37}$$

with boundary conditions

$$E_3(i) = \delta_{i,0} \qquad E_n(0) = 1 \qquad n \geq 3\,. \tag{2.38}$$

The first term on the right-hand side of (2.37) originates from creating a new vertex in an externally labelled tree with $n-1$ external and $i-1$ internal branches, and attaching the nth external branch to it. The second term comes from adding the new external branch to an already existing vertex in an externally labelled tree with $n-1$ external and i internal branches. Unfortunately, we have not been able to solve (2.37) with boundary conditions (2.38) directly. However, from the point of view of a numerical evaluation, it is much easier to obtain values for $D_n^{(\infty)} = \sum_{i=0}^{n-3} E_n(i)$ using (2.37) than using (2.2). Also, starting from (2.37), one may show that

$$\Xi_n^{(\infty)} = 1 + 2\frac{\sum_i iE_{n-1}(i)}{n\sum_i E_{n-1}(i)} \equiv 1 + 2\frac{\langle i \rangle_{n-1}}{n}\,. \tag{2.39}$$

This implies that as $n \to \infty$ the growth factor $\Xi^{(\infty)}$ is directly determined by the position of the maximum of $E_n(i)$ as a function of i. Since $0 < \langle i \rangle_{n-1}/n < 1$ one immediately sees that $\Xi^{(\infty)} \in (1, 3)$, consistent with (2.30).

Acknowledgment

It is a pleasure to thank Professor P W Kasteleyn for useful conversations.

References

[1] Kleiss R and Kuijf H 1989 *Nucl. Phys.* B **312** 616
[2] Kübleck J 1989 *Master's Thesis* Würzburg University
[3] Harray F (ed) 1967 *Graph Theory and Theoretical Physics* (London: Academic)
[4] Biggs N L, Lloyd E K and Wilson R J 1976 *Graph Theory 1736–1936* (Oxford: Clarendon)
[5] Berends F, Giele W and Kuijf H 1989 *Phys. Lett.* **232B** 266
[6] Cayley A 1889 *Quart. J. Pure Appl. Math.* **23** 376