Revisiting Catalan Numbers: An AB_2 Polymerization Perspective

Surbhi Khewle

Department of Chemical Engineering, Indian Institute of Technology Gandhinagar, Gandhinagar, Gujarat-382355, India
surbhi.khewle@iitgn.ac.in

<https://orcid.org/0000-0001-6264-0693>

Pratyush Dayal¹

Department of Chemical Engineering, Indian Institute of Technology Gandhinagar, Gandhinagar, Gujarat-382355, India
pdayal@iitgn.ac.in

<https://orcid.org/0000-0001-5173-2250>

Neeldhara Misra

Department of Computer Science and Engineering, Indian Institute of Technology Gandhinagar, Gandhinagar, Gujarat-382355, India
neeldhara.m@iitgn.ac.in

<https://orcid.org/0000-0003-1727-5388>

Ravi Anand Singh

Department of Chemical Engineering, Indian Institute of Technology Gandhinagar, Gandhinagar, Gujarat-382355, India
ravi.singh@iitgn.ac.in

Abstract

Number sequences, like Fibonacci, Fermat, Markov, Euler, Bernoulli, etc., have been popular in exemplifying a variety of scientific phenomena. Here, we explore the Catalan numbers in the context of AB_m step polymerisation and develop a framework to derive its alternative closed form expression. Our approach harnesses the concepts of combinatorics and graph theory, in conjunction with kinetics of AB_2 polymerisation to obtain the chain length distribution that directly gives the closed-form expression of Catalan number expressed as a bivariate distribution function. Furthermore, we validate our expression by comparing first 5000 Catalan numbers obtained from its traditional closed-form. As an offshoot, we discuss “pathwidth”, a construct used in graph theory, as a better metrics for describing topology of polymer chains. The framework developed in this work can be extended to AB_m step polymerisation and thus, facilitates topological characterisation of hyperbranched polymers (HPs) that ultimately, dictates their structure-property relationships.

Keywords and phrases Catalan number, AB_2 polymerisation, step polymerisation, polymer chain topology

Related Version Our detailed approach to obtain the topological information of HPs can be found at <https://pubs.acs.org/doi/10.1021/acs.macromol.1c01608>.

Funding The Authors acknowledge IIT Gandhinagar for funding.

¹ To whom correspondence should be addressed.

1 Introduction

Hyperbranched polymers (HPs) are an emerging class of synthetic soft materials that have been in demand due to their tailorable topological structures that provide them versatile physicochemical properties. The design of branching architecture of HPs affords a convenient mechanism to tune their properties [1–4] for intended applications in nanotechnology [5, 6], biomaterials [7–9], coatings [9–11], self-assembly [6, 12, 13], tissue engineering [14, 15], to name a few. Depending on their topologies, these polymers fall primarily into two broad categories, namely compact and segmented hyperbranched polymers (CHPs and SHPs, respectively). [16, 17] One of the techniques to synthesize homo HPs is single monomer methodology (SMM) wherein the polymerization is carried out using only one type of monomer molecules. SMM via one-pot step polymerization of multifunctional AB_m monomers [17], however, is the widely employed method of synthesis because of its significant advantages over other methods. For instance, it is possible to design HPs through click chemistry with precise control on chain topology using this mechanism. Needless to say that the attachments of the individual monomeric moieties onto the polymeric backbone, which govern the chain topology, bears one-to-one correspondence to the physicochemical properties of HPs.

Over a period of time various metrics, like degree of branching (DB) [18, 19], Weiner number [20], contraction factors [21, 22] etc., have been developed to describe the polymer chain topologies; the information obtained from these metrics is then used to extract the structural information of HPs. One of the approaches, uses the graph theory constructs to define fundamental structural units and then uses the time evolution rate equations to obtain the structural information of the polymer chains; degree of branching is an example of this approach [18, 19]. Here, however, there are two common methodologies namely, usage of bivariate distribution function (BDF) [23, 24] and random graph theory (RGT) [25, 26]; the former uses kinetic rate equations of functional groups and structural units to obtain chain length distributions, whereas the latter obtains size distribution using temporal degree distribution functions of directed graphs. An alternate approach harnesses fundamental constructs already defined in the graph theory but uses combinatorics to obtain the branching characteristics of polymer chains directly. For instance, binary Catalan numbers (2C_n) have been used for describing polymer chain topologies synthesised via AB_2 step polymerization. [27, 28] On similar front, we have recently demonstrated that “pathwidth” [29], a metrics used to measure the width of any graph, can be used to derive chain topological information for HPs synthesised by AB_m step polymerisation. [30] In particular, we have used the information from pathwidth along with Catalan numbers to calculate exact numbers of linear and branched structures. Furthermore, we have also combined Half Catalan numbers with the pathwidth to obtain total numbers of linear and branched isomers for a special case of AB_2 step polymerisation. [30]

The focus of this work is to use the constructs of graph theory and combinatorics, in conjunction with the kinetics of AB_2 step polymerisation to obtain an alternate closed form expression for binary Catalan numbers. In addition, we validate this alternate expression by comparing the values from the traditional closed form expression of Catalan numbers. We, however, note that the closed form expressions for obtaining the chain topology have been derived using BDF approach [23, 24] but no connection to Catalan numbers have been proposed so far.

1.1 Catalan numbers

Binary Catalan numbers, or just Catalan numbers, are one of the ubiquitous recursive sequences of natural numbers in enumerative combinatorics that are represented by the following closed form expression [31, 32]

$${}^2C_n = \frac{1}{n+1} \binom{2n}{n} = \frac{(2n)!}{(n+1)!n!}, \text{ where } n \geq 0 \quad (1)$$

On the other hand, the m -ary Catalan numbers, $(^m C_n)$ are the generalizations of binary Catalan numbers given by [32,33]

$$^m C_n = \frac{1}{mn - n + 1} \binom{mn}{n} = \frac{(mn)!}{(mn - n + 1)!n!}, \text{ where } n \geq 0 \quad (2)$$

It can be easily verified, by comparing the above equations, that the generalisation gives the expression for Catalan numbers for $m = 2$. Alternatively, the Catalan numbers are also expressed by the recursive relation

$$^2 C_{n+1} = \sum_{k=0}^n C_k C_{n-k}, \text{ where } n \geq 0 \text{ and } C_0 = 1, C_1 = 1 \quad (3)$$

Figures 1 and 2 shows two different interpretations of Catalan numbers that have been used in the fields other than polymerisation. In particular, n^{th} Catalan number in Fig. 1 depicts the number of unique Dyck paths, defined as the steps of unit lengths forming a “staircase” from $(0,0)$ to (n,n) , that are strictly restricted to one side of the diagonal. [34] Dyck paths have been used for generating pseudo-random numbers, identifying sequential algorithms etc., and hence, are used in stenography, cryptography, to name a few. [34–36] On the other hand, in Fig. 2 the n^{th} Catalan number gives the number of all the possible structures of a rooted full binary trees (FBTs) with $2n + 1$ vertices (represented by circles). [31] It is this notion of Catalan numbers that we harness for obtaining structural information of polymer chains.

■ **Figure 1** n^{th} Catalan number represents all possible Dyck paths in a $n \times n$ square cell.

1.2 Polymer chains as m -ary trees

We begin our discussion by portraying one-to-one correspondence between FMTs and acyclic polymer chains in Fig. 3. Specifically, we represent a linear structure of a pentamer ($n = 5$) as FMTs for two cases, $m = 2$ and $m = 3$, known as full binary and full ternary trees (FBTs and FTTs), respectively. The topmost vertex of the trees in Fig. 3 is known as the root node (orange circle) and the tree grows in the downward direction through this root node by addition of child nodes. The FMTs have either m or 0 children at each parent node and thus, the root node has no parent but has m children. The vertices with no children are called leaves (green circle) whereas, all the other vertices (black circles) are named as internal nodes, or simply nodes.

■ **Figure 2** n^{th} Catalan number represents all possible structures of FBTs with $n + 1$ leaves.

Let us now examine FMTs in the context of AB_m polymerisation. Here, monomer consists of two functional groups A and B with functionalities of $m + 1$ and 2, respectively. In essence, a single AB_m monomer consists of a root node and m leaves. With the advent of polymerisation, the A functional group of one monomer reacts with their B equivalents in another monomer, to form an AB unit, thereby adding internal nodes to the FMT. With each monomer addition, therefore, the number of leaves in the FMT increase by 1. In short, difference between the chain length of the polymer and the number of internal nodes in the corresponding FMT always remains unity. Alternatively, the relationship between the polymer chain length n and the number of leaves k for a FMT is $k = mn - n + 1$. We note that these relationships remain valid irrespective of whether the polymer chain is linear or branched.

■ **Figure 3** FMT representation of hexamer formed by AB_m monomers; (left) FBT for $m = 2$ and (right) FTT for $m = 3$. The orange, black and green circles represent root, internal and leaf nodes, respectively.

In order to use DB for characterising the chain topology, the polymer chain is divided into $m + 1$ structural units as shown in Fig. 4, depending upon the number of leaves attached to it. For instance, in Fig. 4, the units with m , $m - 1$ and 0 leaves are known as terminal (T), linear (L), and dendritic (D) units; these units are depicted by blue, yellow and orange triangles, respectively. It is, however, clear from Fig. 4 that polymer chains with AB_2 monomers only have D, T and L units whereas AB_3 monomers form an additional semi-dendritic unit. Extending this line of argument, therefore, AB_4 monomers form two additional semi-dendritic units and so on and so forth. Thus, in

general, the number of structural units increases as m increases, however, it is not necessary that each configuration of the polymer chain consists of all the structural units. On the other hand, for a fixed m , the total number of structural units present in a polymer chain is equal to its chain length, n . For the specific examples of $n = 6$, shown in Fig. 4, the AB_2 polymer chain consists of two 3 L, 2 T and 1D units whereas, the AB_3 polymer chains consists of 3 L, 2 T and 1 semi-dendritic unit.

Figure 4 Identification of structural units in branched pentamer chain formed by AB_m monomers; FBT for $m = 2$ and (right) FTT for $m = 3$. The dashed triangles in yellow, orange, blue and black colors represent linear (L), perfect dendritic (D), terminal (T), and semi-dendritic structural units, respectively.

The increase of the number of structural units with m has implications on the calculation of DB as well as the chain length distribution. Recall that the calculation of DB and chain length distributions involve the solution of kinetic rate equations of the structural units and thus, as the number of structural units increases, the calculations become cumbersome and computationally expensive. Here, we restrict ourselves to AB_2 step polymerisation, i.e., FBTs with only D, T, and L structural units, and show that the calculation of chain length distribution function from the distribution function of linear units directly results in an alternate (non-traditional) expression of Catalan numbers. In other words, we show that the closed form expression of Catalan numbers can directly be derived from the BDF.

In order to establish the alternate expression for Catalan number, we first describe the kinetics involved in AB_2 step polymerisation and thereby obtain the rate equation for polymer molecule and the structural units. We then solve these rate equations using direct sequential solution and obtain an expression for chain length distribution containing the alternate expression for Catalan number. At the end, we present the summary and conclusion of our work.

2 Methodology

2.1 Kinetics of AB_2 step polymerization

Although, the kinetics of AB_2 step polymerisation has been discussed extensively in the literature, [18, 24, 37, 38] here, we present the basic framework required for the obtaining the alternate expression for Catalan numbers. In general, any entity in the step polymerisation is considered as “polymer” and thus, all the monomers, dimers and n -mers etc., are considered as polymers. In AB_2 step polymerisation, each monomer contains one A, and two B, functional groups and the polymerisation takes place when A and B functional groups combine from different molecules, as shown in Fig. 5, to form a bigger molecule. From another perspective, AB_2 polymerisation involves the reaction between either L or T structural units as both these units, unlike D structural unit, have B functional group available for reaction.

■ **Figure 5** Illustrating the reaction mechanism between A and B functional groups in AB_2 type step polymerization. The first and second reactions show the formation of a polymer molecule via reaction between a L-T units and L-L units, respectively.

To develop the kinetic equations for AB_2 step polymerisation, we follow the standard methodology and define polymeric unit $P_{n,l}$, of chain length n containing l numbers of L structural units. Thus, the total polymer concentration, $P = \sum_{n=1}^{\infty} P_n$ where, P_n is the concentration of the polymers with chain length n , i.e., $P_n = \sum_{l=0}^{n-1} P_{n,l}$ such that $n - l$ is odd. [23] In Fig. 5, we show the possible scenarios of the reaction of L unit of a $P_{i,j}$ molecule with another molecule; the top and bottom rows depict the reaction of L unit with T and L units of another molecule, respectively. Both these reactions result in the formation of a $P_{n,l}$ molecule and can be written as, [23]

where, k_L and k_T represent the rate constants for the formation of L and T structural units, respectively. Note that in the first reaction, the number of L units increases by 1 at the expense of the T unit, while in the second reaction the number of L units decreases by 1 at the expense of D unit. In other words, the corresponding parent molecules have a total number of $l - 1$ and $l + 1$ L units for first and second reaction, respectively. Thus, the rate of formation of the $P_{n,l}$ molecule is given by [23]

$$\left(\frac{dP_{n,l}}{dt} \right)_{\text{form}} = \frac{k_T}{2} \sum_{i,j} (n - l + 3) P_{i,j} P_{n-i,l-1-j} + \frac{k_L}{2} \sum_{i,j} (l + 1) P_{i,j} P_{n-i,l+1-j}$$

Here, the coefficients represent the total number of ways the reaction between structural units can occur. [23] Following the similar line of arguments, a $P_{n,l}$ molecule disappears when it reacts with any $P_{i,j}$ molecule and thus, the rate of disappearance is written as

$$\begin{aligned} \left(\frac{dP_{n,l}}{dt} \right)_{\text{diss}} &= k_T \left((n + 1 - l) P_{n,l} \sum_{i,j} P_{i,j} + P_{n,l} \sum_{i,j} (i + 1 - j) P_{i,j} \right) \\ &\quad + k_L \left(l P_{n,l} \sum_{i,j} P_{i,j} + P_{n,l} \sum_{i,j} j P_{i,j} \right) \end{aligned}$$

Combining the above equations, the net change in concentration of $P_{n,l}$ molecules is written as, [23]

$$\begin{aligned} \frac{dP_{n,l}}{dt} = & \frac{k_L}{2} \sum_{i,j} (l+1) P_{i,j} P_{n-i,l+1-j} + \frac{k_T}{2} \sum_{i,j} (n-l+3) P_{i,j} P_{n-i,l-1-j} \\ & - k_T \left((n+1-l) P_{n,l} \sum_{i,j} P_{i,j} + P_{n,l} \sum_{i,j} (i+1-j) P_{i,j} \right) \\ & - k_L \left(l P_{n,l} \sum_{i,j} P_{i,j} + P_{n,l} \sum_{i,j} j P_{i,j} \right) \end{aligned} \quad (5)$$

Equation 5 represents a system of differential equations whose solution ultimately gives the chain length distribution $P_{n,l}$ of polymer chains containing l numbers of L units. Assuming that at the beginning of polymerisation the amount of the AB_2 monomer is M_0 , the initial conditions for the above equation is written as,

$$P_{1,0}(t=0) = M_0; P_{n \geq 1, l \geq 0}(t=0) = 0 \quad (6)$$

Recall that in the step polymerisation all molecules are polymers and therefore, the above equation depicts that the initial amount of polymer molecules ($P_{1,0}$) is same as the monomer molecules, M_0 . As the polymerisation proceeds and monomers combine, the total amount of A functional groups keeps on decreasing. As each polymer molecule exactly has one unreacted A functional group, the total amount of polymer, P , is equal to the amount of A functional groups, or $P = A$. Likewise, the amounts of all the L structural units, denoted by L , is the sum all the linear units (l) in each polymer molecule, $P_{n,l}$, and the total number of monomer units, M_0 is the sum of the total number of monomeric units attached to each polymer chain. Mathematically, therefore, one can write,

$$\sum_{n,l} P_{n,l} = A; \sum_{n,l} n P_{n,l} = M_0; \sum_{n,l} l P_{n,l} = L; \sum_{n,l} (n+1-l) P_{n,l} = 2T \quad (7)$$

The last expression in Eq. 7 is obtained from the relationship, $t = d + 1$, and that the sum of the various structural units is the same as total amount of monomer, i.e., $L + D + T = M_0$. [18]

In view of the above treatment, it is clear that in order to find the expression for $P_{n,l}$ from the system of differential equations, given in Eq. 5, the specific amounts of A , L and T needs to be determined. Thus, we write the rate equations of these units as,

$$\frac{dA}{dt} = -(k_L L + 2k_T T) A \quad (8)$$

$$\frac{dL}{dt} = 2k_T T A - k_L L A \quad (9)$$

$$\frac{dT}{dt} = -2k_T T A \quad (10)$$

The rationale for the rate equations is that the A functional groups always disappear upon polymerisation through reaction with, either L or T units. Similarly, the T units disappear upon reaction with A functional groups (see top row of Fig. 5), to form L unit. In addition, the L units are formed when the A functional group attaches to another L unit (see bottom row of Fig. 5). The solutions of the Eqs. 8-10 are obtained using the following initial conditions, [18, 37]

$$A(t=0) = M_0; T(t=0) = M_0; L(t=0) = 0 \quad (11)$$

XX:8 Revisiting Catalan Numbers

The analytical solution of Eq. 5 is obtained by introducing the transformation $z = k_T \int_0^t A dt$ (alternatively, $dz/dt = k_T A$) that results in [37]

$$\frac{dP_{n,l}}{dz} = \frac{1}{2A} \left[r(l+1) \sum_{i,j} P_{i,j} P_{n-i,l+1-j} + (n-l+3) \sum_{i,j} P_{i,j} P_{n-i,l-1-j} \right] - \left[(r-1)l + n + 1 - \frac{d \ln A}{dz} \right] P_{n,l} \quad (12)$$

where, $r = k_L/k_T$. Subsequently, the expression for $P_{n,l}$ is obtained as [37]

$$\frac{P_{n,l}}{A} = \left[\frac{(1 + (-1)^{n-l+1})}{2} \right] \left[\frac{2^l (n-1)!}{l! \left(\frac{n-l-1}{2}\right)! \left(\frac{n-l+1}{2}\right)!} \right] \left[\frac{e^{(r-2)z} - 1}{r-2} \right]^l \left[e^{-((r-1)l+n+1)z} \right] \left[\frac{2(e^{-rz} - 1) - r(e^{-2z} - 1)}{r-2} \right]^{\frac{n-l-1}{2}} \quad (13)$$

As we shall show below, the above expression serves as the genesis for the alternate expression for Catalan numbers.

3 Results

To derive the alternate non-traditional expression of Catalan numbers, we assume that the rate constants of formation of linear and terminal structural units are independent of the chain topology and therefore, $k_L = k_T = k$. To begin with, we obtain the expressions for the amounts of A, T and L by solving the rate equations given in Eqs. 8- 10 along with the initial conditions of Eq. 11, to get,

$$A = \frac{M_0}{2e^\tau - 1} \quad (14)$$

$$T = \frac{M_0 e^{2\tau}}{(2e^\tau - 1)^2} \quad (15)$$

$$L = \frac{M_0 2e^{2\tau}(e^\tau - 1)}{(2e^\tau - 1)^2} \quad (16)$$

where, the dimensionless time, $\tau = M_0 kt$. We then substitute the relations, from Eq. 7 and Eqs. 14-16, in Eq.5 to obtain a simplified kinetic equation of $P_{n,l}$ as

$$\frac{dP_{n,l}}{dt} = \frac{k}{2} \left[\sum_{i,j} (l+1) P_{i,j} P_{n-i,l+1-j} + \sum_{i,j} (n-l+3) P_{i,j} P_{n-i,l-1-j} \right] - kP_{n,l} ((n+2)P + M_0) \quad (17)$$

where, the total amount of polymer, $P = A$. The equation above represents a system of differential equations, where each equation is obtained by choosing the value of n and all the corresponding valid values of l . Recall that in the above framework $n-l$ is always odd. Thus, for $n = 3$, $l = 0$ & 2 are valid whereas for $n = 4$, $l = 1$ & 3 are valid. In essence, the solution of the n^{th} differential equation depends upon the solutions of previous $n-1$ equations. The solution of Eq. 17 is thus obtained using a Direct Sequential Solution (DSS) methodology and gives us the analytical expression for the amount of polymer molecules of chain length n with l linear units. To elucidate DSS let us choose the rate equation for the monomer, i.e., $n = 1$ in Eq 17, to get,

$$\frac{dP_{1,0}}{dt} = k [3P_{1,0}P + P_{1,0}M_0] \quad (18)$$

The solution of the above equation is obtained by direct integration, using the initial conditions given in Eq. 6, as

$$P_{1,0} = \frac{M_0 e^{2\tau}}{(2e^\tau - 1)^3} \quad (19)$$

The exponential terms in the above expression are typically replaced by the extent of reaction, measured in terms of the conversion, x , of A functional groups. For the irreversible reaction in Eq. 4, the conversion x is written as, [39]

$$\frac{x}{2} = \frac{e^\tau - 1}{2e^\tau - 1} \quad (20)$$

Substituting the above relationship in Eq. 19, we obtain the expression for the amount of monomer with $l = 0$ L units as,

$$P_{1,0} = M_0(1 - x) \left(1 - \frac{x}{2}\right)^2 \quad (21)$$

On similar lines, the differential equations and its solution for $n = 2$ is obtained as,

$$\begin{aligned} \frac{dP_{2,1}}{dt} &= 2kP_{1,0}^2 - kP_{2,1}(4P + M_0) \\ P_{2,1} &= 2M_0(1 - x) \left(\frac{x}{2}\right) \left(1 - \frac{x}{2}\right)^3 \end{aligned}$$

Using the DSS approach, one can thus obtain a general expression for $P_{n,l}$ in terms of conversion x as follows,

$$P_{n,l} = wM_0(1 - x) \left(\frac{x}{2}\right)^{n-1} \left(1 - \frac{x}{2}\right)^{n+1} \quad (22)$$

Here, the w is the coefficient that varies with the value of n and l as

$$w = \left[\frac{(1 + (-1)^{n-l+1})}{2} \frac{2^l(n-1)!}{l! \left(\frac{n-l-1}{2}\right)! \left(\frac{n-l+1}{2}\right)!} \right] \quad (23)$$

Note that the factor $\frac{(1+(-1)^{n-l+1})}{2}$ in the above expression guarantees that $(n-l)$ is odd for all the polymeric units. The values of w from $n = 2$ to $n = 6$ are summarised in Table 1. The validity of the $P_{n,l}$ expression given Eq. 22 can be checked by putting $r = 1$ in the BDF expression of Eq. 13. [23,24]

n	Value of w at different l					
	l = 0	l = 1	l = 2	l = 3	l = 4	l = 5
1	1	-	-	-	-	-
2	-	2	-	-	-	-
3	1	-	4	-	-	-
4	-	6	-	8	-	-
5	2	-	24	-	16	-
6	-	20	-	80	-	32

■ **Table 1** Coefficient, w , in Eq. 23, at valid values of l for different chain lengths, n .

Ultimately, we find the chain length distribution, P_n , for AB_2 polymerization, by adding the amount of $P_{n,l}$ for all the possible values of l , i.e., $P_n = \sum_{l=0}^{n-1} P_{n,l}$ as follows,

$$\sum_{l=0}^{n-1} \frac{P_{n,l}}{M_0(1-x)} = \left[\sum_{l=0}^{n-1} \frac{(1 + (-1)^{n-l+1})}{2} \frac{2^l(n-1)!}{l! \left(\frac{n-l-1}{2}\right)! \left(\frac{n-l+1}{2}\right)!} \right] \left(\frac{x}{2}\right)^{n-1} \left(1 - \frac{x}{2}\right)^{n+1} \quad (24)$$

Alternatively, we can write

$$\frac{P_n}{M_0(1-x)} = C_n \left(\frac{x}{2}\right)^{n-1} \left(1 - \frac{x}{2}\right)^{n+1} \quad (25)$$

Upon comparison of the above equation with the traditionally obtained expression [39] for chain length distribution, we can conclude that the constant, C_n , is the n^{th} Catalan number, which is given by

$$C_n = \sum_{l=0}^{n-1} \frac{(1 + (-1)^{n-l+1})}{2} \frac{2^l(n-1)!}{l! \left(\frac{n-l-1}{2}\right)! \left(\frac{n-l+1}{2}\right)!} \quad (26)$$

To verify that the above expression indeed represents the Catalan number, we compare its values calculated from the Eq. 1 upto $n = 5000$. In Table 2, we give the calculations for selected values of $n = 1 \dots 1000$.

4 Conclusion

In this work, we develop a basic framework to obtain an alternate expression for Catalan numbers, which have been known to depict the total number of configurations of FBTs. To establish this framework, we explore the role of Catalan numbers in various fields of engineering and describe their usage in the context of step polymerisation. Subsequently, we formulate one-to-one correspondence between the topology of polymer chains with FMTs and develop relationships between the polymer chain length and number of leaves of FMTs for a general AB_m step polymerisation. Furthermore, we identify key structural units in FMTs to obtain topological information of polymer chains through the calculation of DB. It turns out that the number of structural units needed for the calculation of DB, for AB_m step polymerisation, is $m + 1$. Here, we briefly discuss the advantages of “pathwidth”, a construct used in graph theory, over DB.

Restricting ourselves to AB_2 step polymerisation, we harness the BDF approach to calculate the chain length distribution. In particular, we develop the rate equations for various structural units, and functional groups, assuming that their topology plays an insignificant role on their reactivity. Using DSS methodology, we then obtain the closed form expression for the distribution of linear structural units, $P_{n,l}$, that affords the calculation of DB. Ultimately, we determine the chain length distribution, P_n , that directly leads to an alternate closed form expression for binary Catalan numbers. Needless to say, this alternate expression is the BDF representation of the n^{th} Catalan number and, to the best of our knowledge, cannot be derived from its traditional closed-form expression. Finally, we verify the validity of our expression by calculating the Catalan numbers upto 5000. Although our work primarily focusses on obtaining an alternate expression for binary Catalan numbers, the approach used herein can provide more insights into obtaining the structural information of polymer chains synthesised by AB_m polymerisation.

References

- 1 H. Magnusson, E. Malmström, A. Hult, and M. Johansson, “The effect of degree of branching on the rheological and thermal properties of hyperbranched aliphatic polyethers,” *Polymer*, vol. 43, no. 2, pp. 301–306, 2002.

n	Valid l values	C _n by Eqs. 1 & 26
1	0	1
2	1	2
3	0, 2	5
4	1, 3	14
5	0, 2, 4	42
⋮	⋮	⋮
10	1, 3, ..., 7, 9	16 796
11	0, 2, ..., 8, 10	58 786
⋮	⋮	⋮
50	1, 3, ..., 47, 49	1 978 261 657 756 160 653 623 774 456
100	1, 3, ..., 97, 99	896 519 947 090 131 496 687 170 070 074 100 632 420 837 521 538 745 909 320
⋮	⋮	⋮
500	1, 3, ..., 497, 499	539 497 486 917 039 060 909 410 566 119 711 128 734 834 348 196 703 167 679 426 896 420 410 037 336 371 644 508 208 550 747 509 720 888 947 317 534 973 145 917 768 881 736 628 103 627 844 100 238 921 194 561 723 883 202 123 256 952 806 711 505 149 177 419 849 031 086 149 939 116 975 191 706 558 395 784 192 643 914 160 118 616 272 189 452 807 591 091 542 120 727 401 415 762 287 153 293 056 320
1000	1, 3, ..., 997, 999	2 046 105 521 468 021 692 642 519 982 997 827 217 179 245 642 339 057 975 844 538 099 572 176 010 191 891 863 964 968 026 156 453 752 449 015 750 569 428 595 097 318 163 634 370 154 637 380 666 882 886 375 203 359 653 243 390 929 717 431 080 443 509 007 504 772 912 973 142 253 209 352 126 946 839 844 796 747 697 638 537 600 100 637 918 819 326 569 730 982 083 021 538 057 087 711 176 285 777 909 275 869 648 636 874 856 805 956 580 057 673 173 655 666 887 003 493 944 650 164 153 396 910 927 037 406 301 799 052 584 663 611 016 897 272 893 305 532 116 292 143 271 037 140 718 751 625 839 812 072 682 464 343 153 792 956 281 748 582 435 751 481 498 598 087 586 998 603 921 577 523 657 477 775 758 899 987 954 012 641 033 870 640 665 444 651 660 246 024 318 184 109 046 864 244 732 001 962 029 120

■ **Table 2** Calculation of Catalan numbers from Eq. 26.

- 2 X. Zhu, Y. Zhou, and D. Yan, "Influence of branching architecture on polymer properties," *Journal of Polymer Science Part B: Polymer Physics*, vol. 49, pp. 1277–1286, sep 2011.
- 3 G. Arya and A. Z. Panagiotopoulos, "Impact of Branching on the Phase Behavior of Polymers," *Macromolecules*, vol. 38, pp. 10596–10604, dec 2005.
- 4 H. Cheng, S. Xie, Y. Zhou, W. Huang, D. Yan, J. Yang, and B. Ji, "Effect of Degree of Branching on the Thermoresponsive Phase Transition Behaviors of Hyperbranched Multiarm Copolymers: Comparison of Systems with LCST Transition Based on Coil-to-Globule Transition or HydrophilicHydrophobic Balance," *The Journal of Physical Chemistry B*, vol. 114, pp. 6291–6299, may 2010.
- 5 S.-E. Stiriba, H. Kautz, and H. Frey, "Hyperbranched Molecular Nanocapsules: Comparison of the Hyperbranched Architecture with the Perfect Linear Analogue," *Journal of the American Chemical*

- Society*, vol. 124, pp. 9698–9699, aug 2002.
- 6 Y. Zhou and D. Yan, “Supramolecular self-assembly of amphiphilic hyperbranched polymers at all scales and dimensions: progress, characteristics and perspectives,” *Chemical Communications*, no. 10, p. 1172, 2009.
 - 7 Y. Zhou, W. Huang, J. Liu, X. Zhu, and D. Yan, “Self-Assembly of Hyperbranched Polymers and Its Biomedical Applications,” *Advanced Materials*, vol. 22, pp. 4567–4590, nov 2010.
 - 8 M. Hu, M. Chen, G. Li, Y. Pang, D. Wang, J. Wu, F. Qiu, X. Zhu, and J. Sun, “Biodegradable Hyperbranched Polyglycerol with Ester Linkages for Drug Delivery,” *Biomacromolecules*, vol. 13, pp. 3552–3561, nov 2012.
 - 9 H. Zhang, T. Zhao, B. Newland, W. Liu, W. Wang, and W. Wang, “Catechol functionalized hyperbranched polymers as biomedical materials,” *Progress in Polymer Science*, vol. 78, pp. 47–55, mar 2018.
 - 10 K. Mańczyk and P. Szweczyk, “Highly branched high solids alkyd resins,” *Progress in Organic Coatings*, vol. 44, pp. 99–109, apr 2002.
 - 11 J. Lange, E. Stenroos, M. Johansson, and E. Malmström, “Barrier coatings for flexible packaging based on hyperbranched resins,” *Polymer*, vol. 42, pp. 7403–7410, aug 2001.
 - 12 W. Jiang, Y. Zhou, and D. Yan, “Hyperbranched polymer vesicles: from self-assembly, characterization, mechanisms, and properties to applications,” *Chemical Society Reviews*, vol. 44, no. 12, pp. 3874–3889, 2015.
 - 13 W. Chen, X. Li, C. Liu, J. He, M. Qi, Y. Sun, B. Shi, H. Sepehrpour, H. Li, W. Tian, and P. J. Stang, “ β -Cyclodextrin modified Pt(II) metallacycle-based supramolecular hyperbranched polymer assemblies for DOX delivery to liver cancer cells,” *Proceedings of the National Academy of Sciences*, vol. 117, pp. 30942–30948, dec 2020.
 - 14 S. T. Meikle, V. Perugini, A. L. Guildford, and M. Santin, “Synthesis, Characterisation and in vitro Anti-Angiogenic Potential of Dendron VEGF Blockers,” *Macromolecular Bioscience*, vol. 11, pp. 1761–1765, dec 2011.
 - 15 V. Perugini, A. L. Guildford, J. Silva-Correia, J. M. Oliveira, S. T. Meikle, R. L. Reis, and M. Santin, “Anti-angiogenic potential of VEGF blocker dendron loaded on to gellan gum hydrogels for tissue engineering applications,” *Journal of Tissue Engineering and Regenerative Medicine*, vol. 12, pp. e669–e678, feb 2018.
 - 16 M. Zhang, S. M. June, T. E. Long, and J. Kong, “Principles of Step-Growth Polymerization (Polycondensation and Polyaddition),” in *Reference Module in Materials Science and Materials Engineering*, Elsevier, 2016.
 - 17 Y. Zheng, S. Li, Z. Weng, and C. Gao, “Hyperbranched polymers: advances from synthesis to applications,” *Chemical Society Reviews*, vol. 44, no. 12, pp. 4091–4130, 2015.
 - 18 H. Frey, “Degree of branching in hyperbranched polymers. 2. Enhancement of the db: Scope and limitations,” *Acta Polymerica*, vol. 48, pp. 298–309, 1997.
 - 19 C. J. Hawker, R. Lee, and J. M. Fréchet, “One-step synthesis of hyperbranched dendritic polyesters,” *Journal of the American Chemical Society*, vol. 113, no. 12, pp. 4583–4588, 1991.
 - 20 H. Wiener, “Structural Determination of Paraffin Boiling Points,” *Journal of the American Chemical Society*, vol. 69, pp. 17–20, jan 1947.
 - 21 B. H. Zimm and W. H. Stockmayer, “The Dimensions of Chain Molecules Containing Branches and Rings,” *The Journal of Chemical Physics*, vol. 17, pp. 1301–1314, dec 1949.
 - 22 P. Kratochvíl and M. Netopilík, “On the contraction factors of long-chain branched macromolecules,” *European Polymer Journal*, vol. 51, pp. 177–181, feb 2014.
 - 23 Z. Zhou and D. Yan, “Distribution function of hyperbranched polymers formed by AB₂ type polycondensation with substitution effect,” *Polymer*, vol. 47, pp. 1473–1479, feb 2006.
 - 24 H. Tobita and K. Yamazaki, “Bivariate Distribution and Related Analytical Solutions for Batch Step-Growth Polymerization of AB₂-Type Monomer,” *Macromolecular Theory and Simulations*, vol. 29, p. 1900049, jan 2020.

- 25 V. Schamboeck, P. D. Iedema, and I. Kryven, "Dynamic Networks that Drive the Process of Irreversible Step-Growth Polymerization," *Scientific Reports*, vol. 9, p. 2276, dec 2019.
- 26 I. Kryven, J. Duivenvoorden, J. Hermans, and P. D. Iedema, "Random Graph Approach to Multi-functional Molecular Networks," *Macromolecular Theory and Simulations*, vol. 25, pp. 449–465, sep 2016.
- 27 S. Costeux, P. Wood-Adams, and D. Beigzadeh, "Molecular Structure of Metallocene-Catalyzed Polyethylene: Rheologically Relevant Representation of Branching Architecture in Single Catalyst and Blended Systems," *Macromolecules*, vol. 35, pp. 2514–2528, mar 2002.
- 28 S. K. Gupta and A. Kumar, "Simulation of Step Growth Polymerizations.," *Chemical Engineering Communications*, vol. 20, pp. 1–52, feb 1983.
- 29 N. Robertson and P. Seymour, "Graph minors. V. Excluding a planar graph," *Journal of Combinatorial Theory, Series B*, vol. 41, pp. 92–114, aug 1986.
- 30 P. Dayal, N. Misra, S. Khewle, and R. A. Singh, "On the Role of Half-Catalan Numbers and Pathwidth in Hyperbranched Polymers Synthesized by AB_m Step Polymerization," *Macromolecules*, vol. 54, pp. 10456–10467, nov 2021.
- 31 R. P. Stanley, *Enumerative Combinatorics*. Boston, MA: Springer US, 1986.
- 32 R. P. Stanley and S. Fomin, *Enumerative Combinatorics*, vol. 2. Cambridge: Cambridge University Press, jan 1999.
- 33 S. Heubach, N. Y. Li, and T. Mansour, "Staircase tilings and k-Catalan structures," *Discrete Mathematics*, vol. 308, no. 24, pp. 5954–5964, 2008.
- 34 T. Mansour, E. Y. P. Deng, and R. R. X. Du, "Dyck paths and restricted permutations," *Discrete Applied Mathematics*, vol. 154, no. 11, pp. 1593–1605, 2006.
- 35 M. Saračević, A. Selimi, and F. Selimović, "Generation of cryptographic keys with algorithm of polygon triangulation and Catalan numbers," *Computer Science*, vol. 19, no. 3, pp. 243–256, 2018.
- 36 M. Saračević, S. Adamović, V. Miškovic, N. Maček, and M. Šarac, "A novel approach to steganography based on the properties of Catalan numbers and Dyck words," *Future Generation Computer Systems*, vol. 100, pp. 186–197, 2019.
- 37 Z. Zhou and D. Yan, "Kinetic analysis of the polycondensation of AB_g type monomer with a multifunctional core," *Polymer*, vol. 41, pp. 4549–4558, jun 2000.
- 38 H. Galina, J. B. Lechowicz, and K. Kaczmarek, "Kinetic Models of the Polymerization of an AB₂ Monomer," *Macromolecular Theory and Simulations*, vol. 10, pp. 174–178, mar 2001.
- 39 P. J. Flory, *Principles of polymer chemistry*. Ithaca : Cornell University Press, 1953.