On rotation distance between binary coupling trees and applications for $3nj$-coefficients

V. Fack, S. Lievens and J. Van der Jeugt[1]
Department of Applied Mathematics and Computer Science,
University of Ghent, Krijgslaan 281-S9, B-9000 Gent, Belgium

Abstract

Generalized recoupling coefficients or $3nj$-coefficients for a Lie algebra (with su(2), the Lie algebra for the quantum theory of angular momentum, as generic example) can always be expressed as multiple sums over products of Racah coefficients (i.e. $6j$-coefficients). In general there exist many such expressions; we say that such an expression is optimal if the number of Racah coefficients in such a product (and, correlated, the number of summation indices) is minimal. The problem of finding an optimal expression for a given $3nj$-coefficient is equivalent to finding a shortest path in a graph G_n. The vertices of this graph G_n consist of binary coupling trees, representing the coupling schemes in the bra/kets of the $3nj$-coefficients. This is the graph of rooted (unordered) binary trees with labelled leaves, and has order $(2n-1)!!$. As the order increases so rapidly, finding a shortest path is computationally achievable only for $n < 11$. We present some mathematical tools to compute or estimate the length of such shortest paths between binary coupling trees. The diameter of G_n is determined explicitly upto $n < 11$, and it is shown to grow like $n\log(n)$. Thus for n large enough, the number of Racah coefficients in the expansion of a $3nj$-coefficient is of order $n\log(n)$. We also show that this problem in Racah-Wigner theory is equivalent to a problem in mathematical biology, where one is concerned with the quantitative comparison of classifications or dendrograms. From this context, some algorithms for approximating the shortest path can be deduced.

PACS : 02.20 (Group Theory); 02.70 (Computational techniques); 03.65F (Algebraic methods in Quantum Theory).

Corresponding author : V. Fack, Department of Applied Mathematics and Computer Science, University of Ghent, Krijgslaan 281-S9, B-9000 Gent, Belgium.
Tel. ++ 32 9 2644808; Fax ++ 32 9 2644995; E-mail Veerle.Fack@rug.ac.be.

[1]Research Associate of the Fund for Scientific Research – Flanders (Belgium)

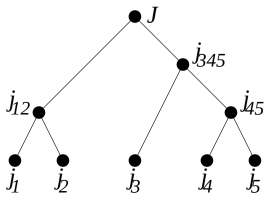
1 Introduction

The subject of the coupling of $n+1$ angular momenta, and the related $3nj$-coefficients, is a difficult one, and the literature is extensive. Classical monographs [1, 2] deal primarily with techniques for implementing graphical methods (known as Yutsis graphs) for carrying out summations over projection quantum numbers in products of Wigner coefficients. In [3, Topic 12], recoupling theory is considered from the point of view of *binary coupling schemes*. A binary coupling scheme is the rooted binary tree representing the order of coupling of a state vector in the tensor product of $n+1$ angular momentum multiplets, labelled respectively by the angular momenta $j_1, j_2, \ldots, j_{n+1}$. The leaves of the binary tree are labelled by these angular momenta $j_1, j_2, \ldots, j_{n+1}$, and the remaining vertices of the tree can be labelled by the intermediate angular momenta. For example, in the tensor product $V_1 \otimes \cdots \otimes V_5$, where each V_i carries a representation of the angular momentum algebra labelled by j_i, the following vector can be considered :

$$\begin{aligned} &|((j_1,j_2)j_{12},(j_3,(j_4,j_5)j_{45})j_{345})J,M\rangle = \\ &\quad \sum_{\text{all } m_i} C^{j_1,j_2,j_{12}}_{m_1,m_2,m_{12}} C^{j_4,j_5,j_{45}}_{m_4,m_5,m_{45}} C^{j_3,j_{45},j_{345}}_{m_3,m_{45},m_{345}} C^{j_{12},j_{345},J}_{m_{12},m_{345},M} \\ &\quad \times |j_1m_1\rangle \otimes |j_2m_2\rangle \otimes |j_3m_3\rangle \otimes |j_4m_4\rangle \otimes |j_5m_5\rangle. \end{aligned} \tag{1}$$

Herein, $C^{j,j',j''}_{m,m',m''}$ is a vector-coupling (Wigner or Clebsch-Gordan) coefficient [3, 4]. The binary coupling scheme representing the above vector is given in Figure 1. The projection

Figure 1: Binary coupling scheme representing (1)



quantum number M is not represented in this binary coupling scheme for reasons that will soon become apparent.

There are obviously several ways in which $n+1$ angular momenta can be coupled, and the quantities that typically appear in atomic and nuclear structure computations are the related general recoupling coefficients or $3nj$-coefficients. A general recoupling coefficient (or a generalized $3nj$-coefficient) is defined to be the transformation coefficient between any such two coupling schemes, e.g.

$$\langle(((j_1,j_4)j_{14},(j_2,j_3)j_{23})j_{1423},j_5)J|((j_1,j_2)j_{12},(j_3,(j_4,j_5)j_{45})j_{345})J\rangle. \tag{2}$$

The M-dependence is dropped since such coefficients are independent of M by the Wigner-Eckart theorem [4]. It is a fundamental theorem of recoupling theory [3, p. 455] that each such transformation coefficient (i.e. every generalized $3nj$-coefficient) can be expressed in terms of sums over products of Racah coefficients ($6j$-coefficients). A famous program of Burke [5], `NJSYM`, is already dealing with this problem. Burke's approach is equivalent to finding a certain path between the two binary coupling schemes (representing the bra- and ket-part of the general recoupling coefficient) by successive elementary transformations on the trees. As we shall explain later, the shorter this path, the better the resulting formula. The path found by `NJSYM` is generally rather long, thus yielding expressions which are far from optimal. In order to improve `NJSYM`, Bar-Shalom and Klapisch [6] developed a new program `NJGRAF` by implementing graphical methods due to Yutsis [1]. Recently, both methods were re-examined and a better implementation was given [7, 8, 9].

In the present paper we consider the method of binary coupling tree transformations as used in `NJSYM` [5] and `NJFORMULA` [7], and relate it to problems in graph theory, mathematical biology, and computer science. For fixed n, we consider the set of all binary coupling trees with $n+1$ leaves (i.e. $n+1$ basic angular momenta). This set is shown to have the natural structure of a graph, G_n, with each vertex of G_n representing a binary coupling tree. Two vertices in G_n are connected by an edge if there exists an elementary transformation (to be defined later) between the corresponding binary coupling trees. In order to find an optimal expression for a general recoupling coefficient (generalized $3nj$-coefficient), it is then sufficient to consider the two vertices in G_n corresponding to the bra- and ket-vector, and to find a shortest path between them in G_n. Although this is a simple reduction of the original problem, the new graph theoretical problem turns out to be as hard as the original problem. One advantage of the equivalent graph theoretical problem is that it has appeared in a number of different contexts, such as computer science and mathematical biology, and thus some properties of the graphs G_n can be found in the literature. In pure graph theory, the problem was first considered by Robinson [10]. In computer science, the equivalent problem is known as finding the rotation distance between unordered rooted binary trees with labelled leaves [11]. In mathematical biology, the problem is known as computing the nearest neighbour interchange metric between dendrograms [12, 13, 14, 15].

Since our main problem is now reduced to finding shortest paths in the graph G_n, we shall study some properties of G_n that are related to distance [16]. In particular we shall be concerned with calculating or estimating the diameter $d(G_n)$ of G_n, since this gives an upper bound for the number of Racah coefficients appearing in the expressions of our generalized $3nj$-coefficients. For $n < 11$, $d(G_n)$ is computed explicitly by means of a computer program. Since the number of vertices of G_n grows rapidly (the order of G_n is $(2n-1)!!$), $d(G_n)$ can no longer be computed for $n \geq 11$. Then, we use a number of techniques from computer science to give upper and lower bounds for $d(G_n)$. We show that $d(G_n)$ grows like $n \log(n)$. Some properties of G_n that are known in the literature are then summarized and converted to our context of $3nj$-coefficients. Finally, we propose some ideas on how to compute approximations for the shortest path problem in G_n.

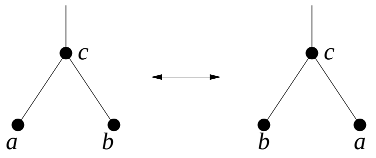
2 Transformations on binary coupling trees

The two parts of a general recoupling coefficient or a $3nj$-coefficient (i.e. the bra- and ket-vector) consist of binary coupling schemes. As shown in the example in Figure 1, the vertices of a binary coupling scheme are labelled by the angular momentum values. The leaves of the binary coupling scheme are labelled by basic angular momentum labels j_i; the other vertices (the coupled vertices) are labelled by the intermediate angular momentum values; and the root or top vertex is labelled by the final angular momentum value J.

As observed by Burke [5] and used in [7], to find an expression of a general recoupling coefficient as a (multiple) sum over products of Racah coefficients, it is sufficient to find a sequence of elementary operations which transform the binary coupling scheme of the bra-vector into the binary coupling scheme of the ket-vector. There are two elementary operations, both corresponding to simple recoupling coefficients and thus to simple contributions in the summation formula for the $3nj$-coefficient.

We refer to [7] for a detailed description of the two elementary operations, and just recall their main properties here. The first elementary operation is called an *exchange* (terminology of [7]) or a *twist* (computer science terminology). It corresponds to the transformation of a state vector of the form $|(a,b)c\rangle$ to $|(b,a)c\rangle$. Its effect on a binary coupling scheme is shown in Figure 2, where a and b can be leaves or coupled vertices. Its

Figure 2: Twist operation



value is determined by the recoupling coefficient

$$\langle (a,b)c \mid (b,a)c\rangle = (-1)^{a+b-c}, \tag{3}$$

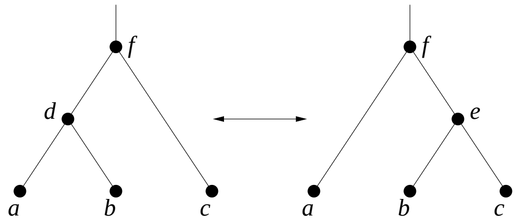
following from the Clebsch-Gordan coefficient property

$$C^{a,b,c}_{\alpha,\beta,\gamma} = (-1)^{a+b-c} C^{b,a,c}_{\beta,\alpha,\gamma}. \tag{4}$$

The second elementary operation is called a *flop* (terminology of [7]) or a *rotation* (computer science terminology in the context of binary search trees). This is a transformation of a state vector of the form $|((a,b)d,c)f\rangle$ to $|(a,(b,c)e)f\rangle$ or vice versa. Its effect on a binary coupling scheme is shown in Figure 3; here again, a, b or c can be leaves or coupled vertices. Its value is determined by the recoupling coefficient

$$\begin{aligned} &\langle ((a,b)d,c)f \mid (a,(b,c)e)f\rangle = \langle (a,(b,c)e)f \mid ((a,b)d,c)f\rangle \\ &= U^{a,b,d}_{c,f,e} = (-1)^{a+b+c+f}\sqrt{(2d+1)(2e+1)} \left\{ \begin{array}{ccc} a & b & d \\ c & f & e \end{array} \right\}. \end{aligned} \tag{5}$$

Figure 3: Rotation on binary coupling scheme



Herein, U is a Racah coefficient, and the last symbol is a $6j$-coefficient [4]. The Racah coefficient can be defined as

$$U^{a,b,d}_{c,f,e} = \sum_{m_1,\ldots,m_5} C^{a,b,d}_{m_1,m_2,m_3} C^{d,c,f}_{m_3,m_4,m} C^{b,c,e}_{m_2,m_4,m_5} C^{a,e,f}_{m_1,m_5,m}. \qquad (6)$$

In order to obtain an expression in terms of Racah coefficients for a general recoupling coefficient (or $3nj$-coefficient), we start from the binary coupling scheme of the bra-vector and try to transform it into the binary coupling scheme of the ket-vector, by applying a sequence of elementary operations on subtrees of the coupling scheme [7]. Each operation contributes a part in the final expression. For a twist, the only contribution is a sign factor of the form (3). For a rotation, the contribution is a factor $U^{a,b,d}_{c,f,e}$; moreover (when going from left to right in Figure 3, for example) if e does not yet appear as a vertex in the binary coupling scheme of the ket-vector, then this operation also gives rise to a new summation variable $\sum_e$. In this case the rotation is said to create a new vertex e. The final expression for the general recoupling coefficient is then a (multiple) summation formula over the products of all contributions corresponding to the sequence of elementary operations. For an example, see [7, Section 2]. This leads to the following :

Theorem 1 *Consider a general recoupling coefficient or $3nj$-coefficient $\langle I \mid F\rangle$, with I and F two couplings of $n+1$ basic angular momenta. Let i and f be the binary coupling schemes corresponding to I and F respectively. If S is a sequence of elementary operations consisting of s_t twists and s_r rotations transforming i to f, then there exists an expression for the $3nj$-coefficient as a multiple sum with each term consisting of a product of s_r Racah coefficients (and a phase factor). Moreover, the number of summation variables is equal to the number of rotations in S that create a new vertex.*

We shall refer to such an expression as *an expansion of the $3nj$-coefficient in terms of Racah coefficients.*

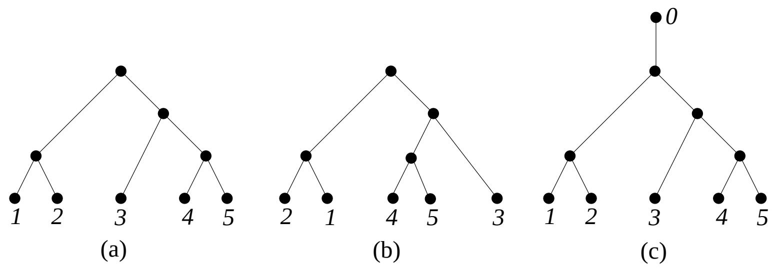
In order to determine an optimal expression for a $3nj$-coefficient, one should find a sequence of elementary operations consisting of the minimum number of rotations. Indeed, a twist is inexpensive since it contributes only a sign (and never an extra summation variable). A rotation however is expensive since it contributes a Racah coefficient (computationally expensive since this involves the evaluation of a single sum expression), and since it can give rise to an extra summation variable. In the terminology of angular momentum, a twist is irrelevant since the corresponding state vectors are related by a phase factor, whereas a rotation is crucial since the corresponding state vectors are related through different intermediate angular momenta. Thus, we shall say that an expansion of a $3nj$-coefficient is optimal if the number of Racah coefficients appearing in such an expansion is minimal.

With this in mind we can redefine our problem and the basic structure that it is dealing with. A *binary coupling tree* on $n+1$ leaves is a rooted binary tree such that

- the $n+1$ leaves are labelled $1, 2, \ldots, n+1$;
- the internal vertices are not labelled;
- for each internal vertex, one can exchange the left and right children of that vertex.

These are sometimes referred to as *unordered rooted binary trees with labelled leaves.* An example is given in Figure 4. Note that in this figure, (a) and (b) represent the same

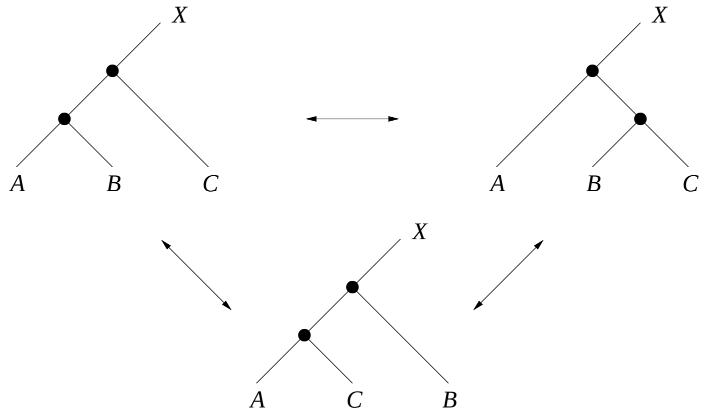
Figure 4: Binary coupling trees



binary coupling tree, since one can freely exchange the left and right children. Sometimes it will be convenient to attach an extra vertex with label 0 to the root of the binary coupling tree, such as in (c). The only elementary operation that now remains is rotation on binary coupling trees (since by definition a twist does not change the binary coupling tree). This is illustrated in Figure 5, where A, B and C represent subtrees and X is a part of the binary coupling tree containing the root (or, equivalently, the label 0).

The relation with binary coupling schemes is obvious. The leaf labels $1, 2, \ldots, n+1$ refer to the angular momentum values $j_1, j_2, \ldots, j_{n+1}$. An internal vertex is no longer explicitly labelled, but it is implicitly labelled by the collection of leaves underneath it. For a given $3nj$-coefficient with binary coupling schemes i and f and for every sequence S consisting of s_t twists and s_r rotations transforming i into f, there exists a sequence of s_r rotations transforming the corresponding binary coupling trees into each other and vice versa. Clearly, from the sequence of rotations between the binary coupling trees, the sequence of twists and rotations between the binary coupling schemes can be reconstructed and hence no information is lost for determining the summation formula for the $3nj$-coefficient. Our basic problem is now reduced to finding a shortest sequence of rotations transforming one binary coupling tree into another.

Figure 5: Binary coupling trees related by a rotation



A binary coupling tree can be given either explicitly as a graph, see Figure 4, or as a bracketing of the leaf labels. For example, the binary coupling tree of Figure 4 could be represented as

$$((1,2),(3,(4,5))) \qquad \text{or} \qquad ((2,1),((4,5),3)). \tag{7}$$

Henceforth we shall use this notation for a binary coupling tree. Sometimes we shall even use it to refer to the underlying binary coupling scheme itself, if the intermediate angular momentum values play no explicit role.

In the following section we shall show that finding a sequence of rotations transforming one binary coupling tree into another one is equivalent to finding a path in a graph G_n. First, there are two important observations.

Remark 2 We wish to draw attention to the fact that this method of binary coupling trees is more generally applicable than the case of the angular momentum algebra considered here. The angular momentum algebra is the Lie algebra $su(2)$, and the multiplets correspond to finite-dimensional irreducible representations of $su(2)$. Also for an arbitrary finite-dimensional semi-simple Lie algebra g, one can consider the $(n+1)$-fold tensor product $V_1 \otimes \cdots \otimes V_{n+1}$. Just as in (1), one can define vectors in this tensor product using the Clebsch-Gordan coefficients of g; then j_i stands for the representation labels of V_i and m_i for the internal labels of the vector. Since the tensor product is in general no longer multiplicity-free, the coupled vectors are labelled by representation labels and an additional label (see, e.g. [17, Section 19.6], or [18] for the example of $su(N)$). But the formal problem of writing a general recoupling coefficient of g [17, Section 19.11] in terms of Racah coefficients of g remains exactly the same as for $su(2)$, and thus the method of binary coupling trees holds here as well. Thus, all the following results in this paper hold for the expansion of a $3nj$-coefficient of an arbitrary semi-simple Lie algebra g in terms of Racah coefficients of g. One can even extend the applicability to non-compact Lie groups, or to infinite-dimensional representations. For example, the method also works for tensor products of positive discrete series representations of $su(1,1)$, since such a tensor product is completely decomposable into a direct sum of positive discrete series representations (even without multiplicity labels). Even though we continue to use the terminology of angular momentum coupling in the following sections, we wish to emphasize that we have this extended coupling problem in mind.

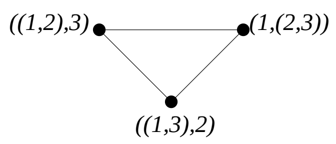
Remark 3 For the case of $su(2)$, the powerful method of Yutsis graphs was developed [1]. This graphical method is extremely useful to find (optimal) expansions for $3nj$-coefficients of $su(2)$ [6, 9], or to classify them. However, intrinsically this method uses various symmetry properties of Clebsch-Gordan or Racah coefficients that are valid for the case of $su(2)$ only. Thus it is no longer valid for the extended case described in the previous Remark and considered in the rest of this paper. For the extended case, only two properties are needed : (4) (or an equivalent one, see [18, (5.17)]) and (6), which is always valid (see eq. (19.49) of [17]).

3 The graph G_n

Let $n > 1$ be fixed, and consider the set of all binary coupling trees with $n + 1$ leaves. Since our only basic operation is rotation, we shall consider the *rotation graph* of binary coupling trees. This graph G_n has as vertex set the set of all binary coupling trees with $n + 1$ leaves, and there is an edge between two vertices if and only if the corresponding binary coupling trees are related through a single rotation. It follows that an optimal expression for a $3nj$-coefficient corresponds to finding a shortest path in G_n between the two binary coupling trees related to the bra- and ket-vector of the $3nj$-coefficient.

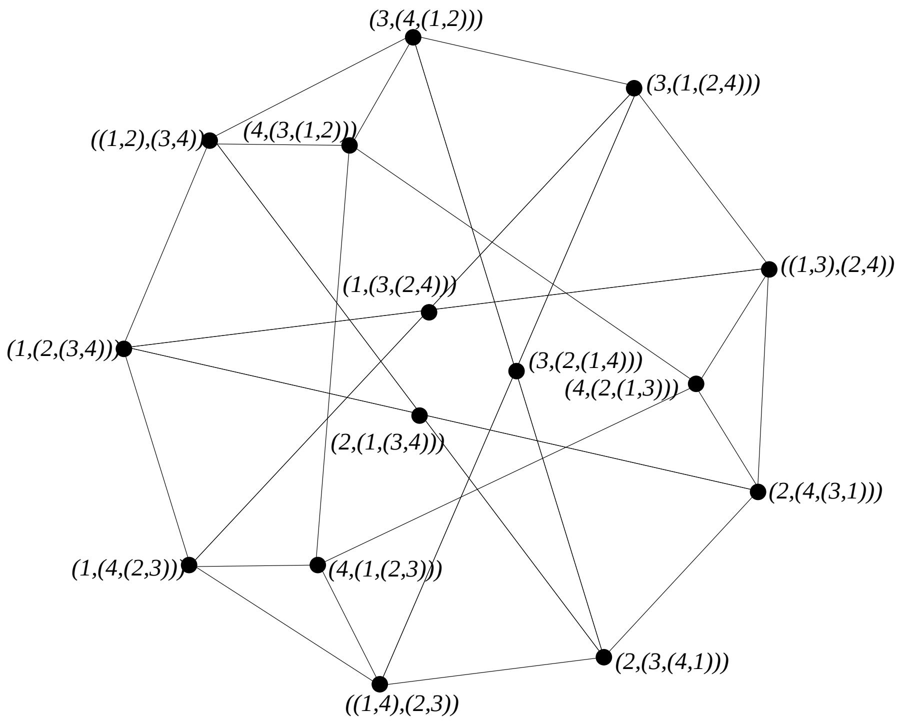
We shall now consider some examples, and deduce some general properties of G_n. For $n = 2$ this graph is simply a triangle. In Figure 6 we give G_2, and use the convention (7) to label the corresponding binary coupling trees. The next graph, G_3, has order 15. This

Figure 6: The graph G_2



graph is shown in Figure 7, using the bracket representation (7) for the binary coupling trees. The equivalence between optimal expressions for $3nj$-coefficients and shortest paths in G_n can be illustrated in Figure 7 for the classical $9j$-coefficient. For this coefficient, the corresponding binary coupling trees are $((1,2),(3,4))$ and $((1,3),(2,4))$. The shortest path in G_3 is of length 3, and thus this implies that this $9j$-coefficient can be written as a single sum expression over the product of three $6j$-coefficients (which is, at least for $su(2)$, a well-known fact). Note that in our terminology, we refer to other coefficients such as the ones corresponding to $\langle(1,(2,(3,4)))|(3,(2,(1,4))\rangle$, $\langle(1,(2,(3,4)))|(2,(3,(1,4)))\rangle$ or $\langle(1,(2,(3,4)))|((1,2),(3,4))\rangle$ also as $9j$-coefficients, even though the first two reduce to the product of two $6j$-coefficients and the last one to a single $6j$-coefficient (as can be seen in Figure 7 from the corresponding distances in G_3).

Figure 7: The graph G_3



Let us now consider some general properties of the graph G_n. An arbitrary element of G_n, i.e. a binary coupling tree on $n+1$ leaves, has $n-1$ internal edges (i.e. edges containing no leaf). Two rotations can be performed with respect to each internal edge, thus every binary coupling tree is connected by an edge to $2(n-1)$ other binary coupling trees. In other words, G_n is a regular graph of degree $2(n-1)$. For example, G_3 is regular of degree 4.

To determine the number of binary coupling trees on $n+1$ leaves (or the order $|G_n|$ of G_n), consider first a binary coupling tree T on n leaves (with labels $1,\ldots,n$), and extend the root of T with an extra edge ending in the leaf 0 (as in Figure 4). This tree has $2n-1$ edges in total. Therefore, there are $2n-1$ different ways of adding an extra edge ending with leaf label $n+1$ to this tree, namely by attaching it to each consisting edge, see e.g. Figure 8. Thus we have $|G_n| = (2n-1)|G_{n-1}|$, and find (see also [10])

$$|G_n| = (2n-1)!! = (2n-1)(2n-3)\cdots 3\cdot 1. \tag{8}$$

This implies that the order of G_n grows exponentially. Table 1 gives the degree and the order of G_n for $n < 11$.

It is also easy to show that G_n is a connected graph, i.e. for any two binary coupling trees T_1 and T_2, there exists at least one path between T_1 and T_2 [10]. In a sense, this statement is equivalent to the fundamental theorem of recoupling theory [3, p. 455] that

Figure 8: Five ways of attaching an extra leaf label 4 to a given binary coupling tree on labels 1,2,3

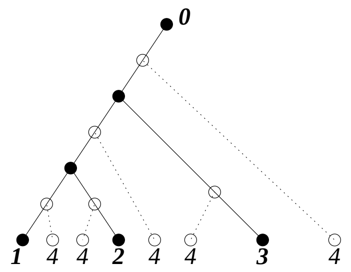


Table 1: Degree and order of G_n

n	2	3	4	5	6	7	8	9	10
$\deg(G_n)$	2	4	6	8	10	12	14	16	18
$\lvert G_n \rvert$	3	15	105	945	10395	135135	2027025	34459425	654729075

each generalized $3nj$-coefficient can be expressed in terms of sums over products of Racah coefficients ($6j$-coefficients).

Clearly, for $n \geq 11$ the order of G_n becomes too large to represent G_n in the RAM of a computer, which is necessary for the computation of shortest paths or of the diameter of G_n. In the following section we shall describe some of the results of our computations for $n < 11$. After that, we shall concentrate on the diameter of G_n, and give some approximations.

4 Distance in G_n

So far, we have reduced our problem to the following : given two binary coupling trees T_1 and T_2 from G_n, find a shortest path between T_1 and T_2, and in particular determine the length of this path (since this determines the number of Racah coefficients in the expansion). The length of a shortest path between T_1 and T_2 is known as the distance $d(T_1, T_2)$ between T_1 and T_2, since this induces a distance function or metric [12, 15].

Let us consider again the example $n = 3$, with G_3 given in Figure 7. Starting from the binary coupling tree $((1,2),(3,4))$, one finds that

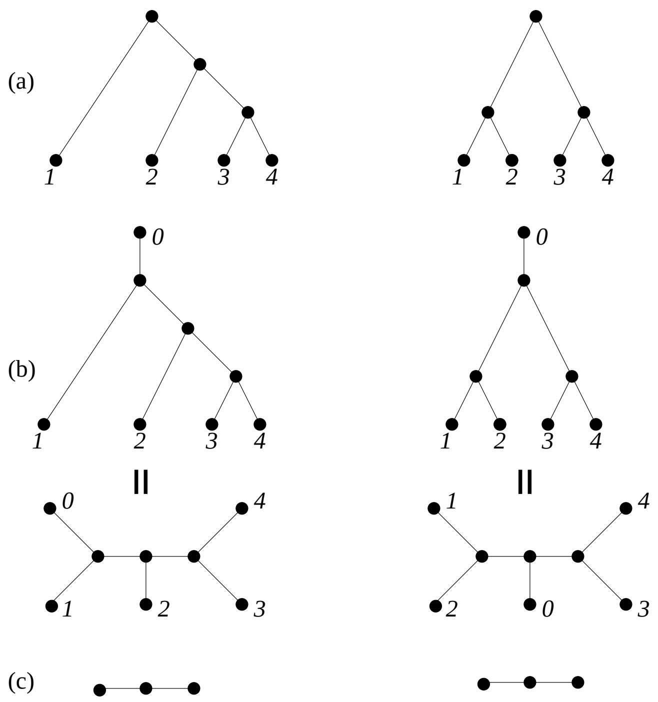
- there are 4 elements of G_3 at distance 1, namely
  $(4,(3,(1,2)))$, $(3,(4,(1,2)))$, $(2,(1,(3,4)))$ and $(1,(2,(3,4)))$;
- there are 8 vertices of G_3 at distance 2, namely
  $(3,(1,(2,4)))$, $(3,(2,(1,4)))$, $(4,(1,(2,3)))$, $(1,(3,(2,4)))$, $(1,(4,(2,3)))$,
  $(4,(2,(1,3)))$, $(2,(3,(4,1)))$ and $(2,(4,(3,1)))$;

- there are 2 vertices at distance 3, namely $((1,3),(2,4))$ and $((1,4),(2,3))$.

The sequence giving the number of elements at distance k ($k = 0, 1, \ldots$) from a given vertex T is called the *distance degree sequence* (DDS) for that vertex T. Thus, in G_3, the distance degree sequence of $((1,2),(3,4))$ is $(1,4,8,2)$. The two elements at maximum distance, in casu $((1,3),(2,4))$ and $((1,4),(2,3))$, give rise to $3nj$-coefficients with the maximum number of Racah coefficients in an optimal expression; in this case $\langle((1,2),(3,4))|((1,3),(2,4))\rangle$ and $\langle((1,2),(3,4))|((1,4),(2,3))\rangle$ give rise to genuine $9j$-coefficients that have as optimal expansion a single sum over products of three $6j$-coefficients.

It is not surprising that the distance degree sequence of $((1,3),(2,4))$ or $((1,4),(2,3))$ is also $(1,4,8,2)$. After all, a permutation of the leaf labels of the binary coupling trees in G_n does not change the structure of G_n. On the other hand, it is at first sight surprising that also the other vertices of G_3 of the form $(a,(b,(c,d)))$ have the same distance degree sequence as $((1,2),(3,4))$. Indeed, the binary coupling trees for $(1,(2,(3,4)))$ or $((1,2),(3,4))$ look different, see Figure 9(a). Thus G_3 has two different *types* of binary

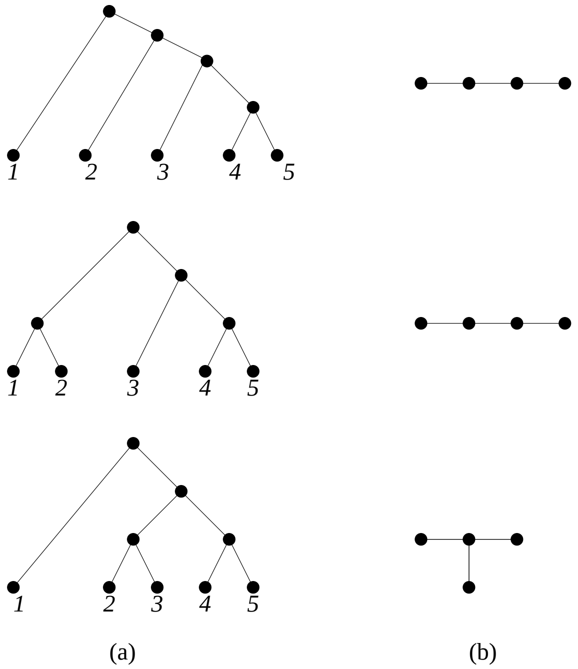
Figure 9: Two different types of binary coupling trees with the same skeleton



coupling trees, the first of the form $(a,(b,(c,d)))$ and the second of the form $((a,b),(c,d))$. This distinction changes however when one attaches an extra leaf label 0 to the root (Figure 9(b)). The full binary trees with labelled leaves (labels from $0, 1, \ldots, 4$) are now the same, upto a permutation of the labels. Since distance is governed by rotations over internal edges, it follows that the corresponding binary coupling trees will indeed have the same distance degree sequence. Note that there is another way of saying this, by introducing the skeleton [13]. Generally, the *skeleton* of a binary coupling tree T of G_n with labelled leaves (labels from $0, 1, \ldots, n+1$) is obtained by deleting all leaves and corresponding edges from T. The result is thus an unlabelled tree (of maximum degree 3) with $n-1$ edges, see Figure 9(c) for $n = 3$.

Let us consider the next case $n = 4$ (corresponding to $12j$-coefficients). G_4 has 105 vertices. It is easy to verify that there are now three different types of binary coupling trees, given in Figure 10(a). When one considers their corresponding skeletons, there turn

Figure 10: Binary coupling trees for $n = 4$ and their skeletons



out to be 2 different ones, given in Figure 10(b). Thus, to have a complete picture of the distance in G_4, one should determine the distance degree sequences only for two vertices in G_4, e.g. for $T_1 = (1,(2,(3,(4,5))))$ and for $T_2 = (1,((2,3),(4,5)))$. We have found that $\mathrm{DDS}(T_1) = (1,6,20,40,34,4)$ and $\mathrm{DDS}(T_2) = (1,6,24,30,44)$. Thus for T_1 there are 4 vertices at distance 5, whereas for T_2 there are no vertices at distance 5, but 44 vertices at distance 4. The vertices at maximum distance of $T_1 = (1,(2,(3,(4,5))))$ are given by

$$(5,(2,(3,(1,4)))), \qquad (5,(3,(2,(1,4)))), \qquad (4,(2,(3,(1,5)))), \qquad (4,(3,(2,(1,5)))).$$

As a consequence, for the $12j$-coefficient corresponding to

$$\langle (1,(2,(3,(4,5))))|(5,(2,(3,(1,4))))\rangle,$$

the optimal expansion is a double sum over products of 5 Racah coefficients. This may seem to contradict the fact that the classical $12j$-coefficients for $su(2)$ have an expression in terms of a single sum over products of only 4 Racah coefficients, which can be found using the technique of Yutsis. In this context, however, recall the observation in Remark 3. The fact that the optimal expansions for $12j$-coefficients of $su(2)$ can even further be reduced is related to symmetry properties that hold only for $su(2)$ coupling coefficients, and not for the general Lie algebra case considered here.

From the previous examples $n = 3$ and $n = 4$ it is clear that in order to determine distance properties for given G_n it is sufficient (a) to determine the number of skeletons; (b) to determine the distance degree sequence for each skeleton. Of course, this does not yet give the shortest path between any two given elements of G_n. But it does at least yield many other distance concepts (eccentricity, radius, center, periphery, ... [16]), and it also determines one of the most important distance characteristics of G_n, its diameter $d(G_n)$. The *diameter* of G_n is defined as follows :

$$d(G_n) = \max\{d(T_1, T_2)|T_1, T_2 \in G_n\}. \tag{9}$$

Thus it is the maximum value over all possible shortest path lengths of G_n; in other words : it is the length of the longest distance degree sequence.

To determine the number of skeletons t_n is an easy task, since for given n the skeletons are the unlabelled trees of maximum degree 3 (the so-called trivalent trees) with $n-1$ edges (or, equivalently, with n vertices). This number is known, see e.g. sequence number A000672 of [19], or [20]. The first few values are given in Table 2. In Figure 11 we list the

Table 2: Number of trivalent trees with n vertices

n	2	3	4	5	6	7	8	9	10	11	12
t_n	1	1	2	2	4	6	11	18	37	66	135

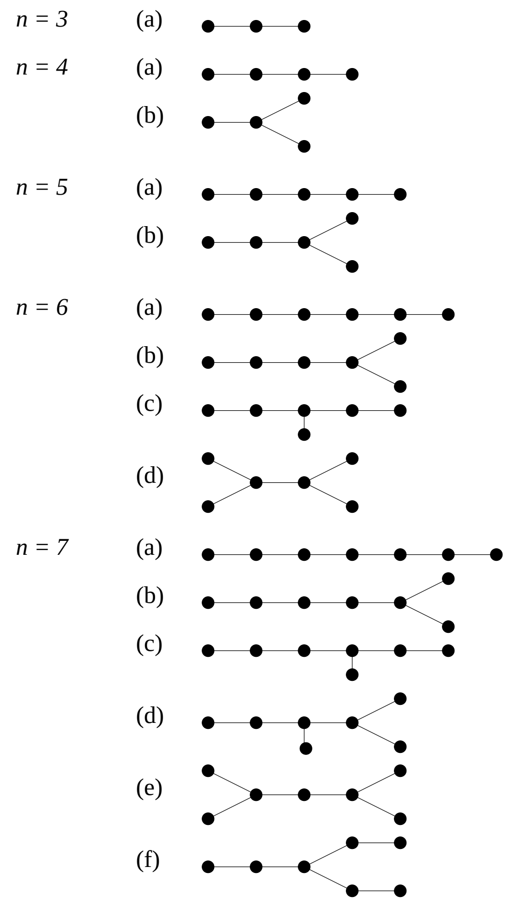
trivalent trees with n vertices (skeletons) upto $n = 7$ (see also [21]).

The purpose is then to calculate the distance degree sequence for a binary coupling tree corresponding to a skeleton. Our method for doing this is described in the following section.

5 The diameter $d(G_n)$

Let T be a given binary coupling tree of G_n. We wish to compute the distance degree sequence DDS(T). Let D_i be the set of elements of G_n at distance i from T. There is one element at distance 0, namely T itself; thus we have $D_0 = \{T\}$. Observe that it is easy to determine the neighbours of T in G_n : these are the elements of G_n at distance 1 from T, i.e. they are the binary coupling trees obtained from T by performing one rotation. Such rotations are easy to perform; as we have already observed in Section 3, every binary coupling tree has $2n-2$ neighbours. Thus D_1 has $2n-2$ elements. Next we compute the set of neighbours of the elements of D_1, and delete from this set the ones that were

Figure 11: Skeletons (trivalent trees) with n vertices for $n = 3, \ldots, 7$



already in D_0 or D_1, yielding D_2. Continuing this way, one can determine all D_i, and their orders give DDS(T).

Such a computation requires : (a) a simple data structure for a binary coupling tree, that allows an easy determination of its neighbours (i.e. perform rotations); (b) a proper way of keeping track of the elements of G_n that have already been encountered in D_0, D_1, ..., D_i while D_{i+1} is being determined. We have written a C program to calculate the distance degree sequence in G_n for any given binary coupling tree. Our program is inspired by some techniques of [22] (or, equivalently, [23]), and we shall not go into the details of this program here. Note that by the second requirement it is necessary that the elements of G_n can be stored in the RAM of the computer. Knowing the order of G_n, see (8) or Table 1, it is clear that on any present-day computer the computation cannot be performed beyond $n = 10$. We have calculated all distance degree sequences upto $n = 10$. Table 3 gives the results upto $n = 7$. For the complete results upto $n = 10$ see URL `http://allserv.rug.ac.be/~jvdjeugt/BCT`.

Table 3: Distance degree sequences in G_n. The skeleton types (a), (b), etc. refer to Figure 11.

	0	1	2	3	4	5	6	7	8	9	10	11	12
$n=3$													
(a)	1	4	8	2									
$n=4$													
(a)	1	6	20	40	34	4							
(b)	1	6	24	30	44								
$n=5$													
(a)	1	8	36	110	244	328	198	20					
(b)	1	8	40	120	228	312	220	16					
$n=6$													
(a)	1	10	56	220	670	1616	2810	3064	1708	236	4		
(b)	1	10	60	238	730	1604	2652	3060	1736	304			
(c)	1	10	60	250	732	1608	2598	2972	1880	276	8		
(d)	1	10	64	268	752	1648	2516	2672	2192	272			
$n=7$													
(a)	1	12	80	378	1408	4344	11210	23028	34630	35050	20518	4320	156
(b)	1	12	84	404	1520	4688	11546	22420	33584	34748	20832	5104	192
(c)	1	12	84	416	1586	4796	11548	22188	32688	34588	21936	5100	192
(d)	1	12	88	454	1724	5096	11864	21808	30520	33200	24624	5712	32
(e)	1	12	88	430	1688	4912	11844	22352	31616	34224	22368	5248	352
(f)	1	12	84	428	1652	4920	11550	21752	32088	34372	22804	5320	152

From the calculation of the distance degree sequences, one easily deduces the diameter of G_n. Table 4 gives the diameter up to $n = 10$, which is as far as one can compute on a present-day computer, and which goes beyond previously calculated diameters [13, 22].

Table 4: Diameter of G_n

n	2	3	4	5	6	7	8	9	10
$d(G_n)$	1	3	5	7	10	12	15	18	21

The diameter is important as it gives an upper bound for the number of Racah coefficients appearing in an optimal expansion of a $3nj$-coefficient, see Section 2. Since it is so difficult to calculate the diameter explicitly beyond $n = 10$, one may wonder whether a proper approximation of $d(G_n)$ can be determined. This is indeed the case. In this paper we shall give a new lower and upper bound for $d(G_n)$. The details of the proofs are omitted here, and will be given in a separate comprehensive study of diameter properties of G_n [24].

Lemma 4 *The number of elements within distance i from any given binary coupling tree in G_n is less than or equal to $\binom{n+2i}{i}4^i$.*

This can be shown using so-called short encodings [25]; for a detailed proof see [24].

Theorem 5 *The diameter of G_n satisfies*

$$d(G_n) > \frac{1}{4}\log(n!) > \frac{1}{4}n\log(n/e). \tag{10}$$

Herein (and in what follows), $\log = \log_2$ is the logarithm in basis 2, and e is the basis of the natural logarithm.
To prove the theorem, let $\delta = d(G_n)$, then by the above lemma and (8) we have that

$$\binom{n+2\delta}{\delta}4^\delta \geq (2n-1)!! = \frac{(2n)!}{2^n n!}. \tag{11}$$

The lhs of (11) can be enlarged by

$$\frac{2^{n+2\delta}}{2n} > \binom{n+2\delta}{\delta},$$

which holds for all integers $n > 0$ and $\delta > 1$, see [24]. The rhs of (11) can be bounded using $\binom{2n}{n} \geq 2^{2n}/(2n)$. Thus (11) yields :

$$2^{4\delta} > n!,$$

from which the theorem follows.

An upper bound follows from

Theorem 6 *The diameter of G_n satisfies ($n > 1$)*

$$d(G_n) < n\lceil\log(n)\rceil - 2^{\lceil\log(n)\rceil} + 2(n - \lceil\log(n+1)\rceil) + 1 < n\lceil\log(n)\rceil + n - 2\lceil\log(n)\rceil + 1. \tag{12}$$

Herein, $\lceil x \rceil$ is the smallest integer larger than or equal to $x > 0$.

The proof of this theorem follows the lines indicated in [22]. Let a *spine* be a binary coupling tree of the form

$$(1, (2, (3, (\ldots (n, n+1)) \cdots); \tag{13}$$

see Figure 12 for its general shape. First, one determines an upper bound for the number of rotations needed to transform such a spine with arbitrary ordered labels into the spine with ordered labels (13). This upper bound is given by $n\lceil\log(n)\rceil - 2^{\lceil\log(n)\rceil} + 1$, see [24]. Next, it is not difficult to see that there are at most $n - \lceil\log(n+1)\rceil$ rotations needed to transform an arbitrary binary coupling tree into a spine. Thus to transform two binary coupling trees T_1 and T_2 into each other, first transform both T_1 and T_2 into a spine, and then transform one spine into the second one. This leads to (12).

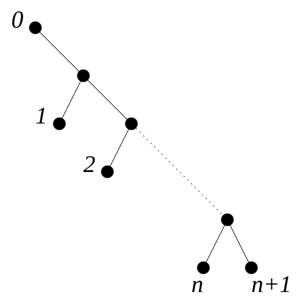
Together, the above two theorems imply that the diameter of G_n is of order $n\log(n)$, i.e.

$$d(G_n) = \Theta(n\log(n)).$$

Note that a weaker upper limit has been given earlier in [11], and weaker upper and lower limits were determined in [22].

The for us important consequence is :

Figure 12: Binary coupling tree which is a spine



Corollary 7 *Consider an optimal expansion of a $3nj$-coefficient in terms of Racah coefficients. Then the number s_r of Racah coefficients appearing in a term of the expansion is of order $n\log(n)$; more explicitly, it is bounded by*

$$\frac{1}{4}n\log(n/e) < s_r < n\lceil\log(n)\rceil + n - 2\lceil\log(n)\rceil + 1.$$

6 Equivalent and related problems

In this paper, we have reduced the problem of finding an optimal expansion of a $3nj$-coefficient to the graph theoretical problem of finding a shortest path between binary coupling trees in G_n. This last problem has been encountered before in different contexts. One of the first papers where this problem is stated, with applications in mind, is [12]. In that context, our binary coupling trees are called *dendrograms*, and the purpose is the computation of a similarity measure or distance (coinciding with our distance $d(T_1, T_2)$) between dendrograms.

Dendrograms can be defined as rooted trees where each of the terminal vertices (leaves) represent an object and where the root vertex represents the entire object-set [26]. According to [26], binary dendrograms can have labelled or unlabelled leaves, and can be ranked or non-ranked (ordered or unordered, in our terminology). An enumeration of four types of binary dendrograms was given in [26]. The ones of interest to us are the non-ranked (unordered) dendrograms with labelled leaves, since they coincide with our binary coupling trees. These are also the dendrograms appearing in the paper of Waterman and Smith [12], and are of importance in mathematical biology. The first area of application is taxonomy. Here, various hierarchical cluster methods are used to construct taxonomic dendrograms. A cluster algorithm can result in dendrograms with differing initial ordering, thus a method of measuring the degree of similarity between dendrograms is of importance [12]. The similarity is computed by means of the distance between dendrograms, coinciding with distance in G_n. A second area of biological research involving dendrograms is in morphogenesis and/or cell differentiation studies, where the development of systems is represented by a tree (decision tree; equivalent to our binary coupling tree). Here again, a tree similarity measure is of importance [12], and is given by distance in G_n.

Some ways of computing or estimating the similarity of dendrograms have been considered in the literature. Waterman and Smith [12], who introduced similarity measure, also use the term *nearest neighbour interchange metric* to refer to the distance d in G_n. Realizing that d is difficult to compute in general, they introduced another measure c which afterwards turned out to violate the triangle inequality [13, 14, 15]. Brown and Day showed that both d and c are difficult to compute [27]. They designed an approximation to d, and analysed the algorithm for the computation of this approximation. Another approximation was considered in [28], requiring only $O(n)$ time to compute. That in general d is difficult to calculate was indicated by Křivánek [29], who showed that computing d is an NP-complete problem. It should be mentioned that Li *et al* [22] found a mistake in Křivánek's proof, so the question of NP-completeness remains open. Still, at present there is no simple algorithm to compute the distance $d(T_1, T_2)$ between given binary coupling trees in G_n. Therefore, we plan to reconsider the methods of [27] or [28] in order to develop programs that produce close approximations for the distance $d(T_1, T_2)$ (and thus programs that produce expansions of $3nj$-coefficients which are nearly optimal).

The problem of computing the distance d in G_n was also considered by computer scientists [11, 23]. Often, however, computer scientists are more interested in a closely related problem : calculating the rotation distance d in the graph H_n consisting of binary search trees (rooted ordered binary trees, with unlabelled leaves) with n internal vertices. In [11], it was already shown that the diameter $d(H_n) \leq 2n-2$ (so a linear upper bound, instead of a $n\log(n)$ upper bound for $d(G_n)$). A detailed study [30] revealed that $d(H_n) \leq 2n-6$, later confirmed by more elementary methods [31, 32]. Although the diameter is easier to estimate in this case, computing the actual distance d in H_n once again turns out to be difficult [33].

One of the properties of the distance function d for G_n, which has received attention in the mathematical biology literature, is that of non-decomposability. This peculiar property is also of interest in our context of $3nj$-coefficients. Let $T, T' \in G_n$ be two binary coupling trees with leaves labelled $1, 2, \ldots, n+1$. Suppose that, in the notation of (7), $T = (t_1, t_2)$ and $T' = (t_1', t_2')$, where $t_1, t_1' \in G_k$ are binary coupling trees with leaves labelled by $1, 2, \ldots, k+1$, and $t_2, t_2' \in G_{n-k-1}$ are binary coupling trees with leaves labelled by $k+2, \ldots, n+1$. The distance d is said to satisfy the decomposition property if for all such T and T' :

$$d(T, T') = d(t_1, t_1') + d(t_2, t_2'),$$

where (by abuse of notation) the first d in the rhs is the distance function in G_k, and the second d in G_{n-k-1}. Otherwise, d is non-decomposable. It was indicated in [14] and shown in [22] that the distance function d for G_n does not satisfy the decomposition property. This implies that for general $3nj$-coefficients, a so-called cut on two lines [1, Figure 14.2][6] in the corresponding Yutsis graph does not necessarily yield the most optimal expansion in Racah coefficients (although it will do so for $n \leq 7$).

7 Conclusion

We have considered the problem of finding an optimal expansion of a general $3nj$-coefficient (for finite-dimensional representations of a semi-simple Lie algebra g) in terms of Racah coefficients of g. This problem was reduced to the shortest path problem in the graph G_n, of which the vertices are given by binary coupling trees on $n+1$ leaves and the edges correspond to rotations. Finding shortest paths in the rotation graph of binary coupling trees turns out to be a difficult problem. Upto $n = 10$, the distance degree sequences of G_n have been calculated explicitly, yielding many distance properties of G_n and in particular implying the diameter $d(G_n)$. This diameter is an upper bound for the number of Racah coefficients appearing in an optimal expression for a $3nj$-coefficient. We have shown that $d(G_n)$ grows like $n \log(n)$ by giving upper and lower bounds for it. Finally, we have shown that our shortest path problem has already appeared in other contexts, such as mathematical biology and computer science, where it has important applications. Methods to find approximations of the shortest path, developed in these areas, can be useful in our context of $3nj$-coefficients and will be studied in the future.

References

[1] A.P. Yutsis, I.B. Levinson and V.V. Vanagas, Mathematical Apparatus of the Theory of Angular Momentum (Israel Program for Scientific Translation, Jerusalem, 1962).

[2] E. El Baz and B. Castel, Graphical Methods of Spin Algebra in Atomic, Nuclear, and Particle Physics (Marcel Dekker, New York, 1972).

[3] L.C. Biedenharn and J.D. Louck, The Racah-Wigner Algebra in Quantum Theory, Encyclopedia of Mathematics and its Applications, Vol. 9, ed. G.-C. Rota (Addison-Wesley, Reading, MA, 1981).

[4] A.R. Edmonds, Angular Momentum in Quantum Physics (Princeton Univ. Press, Princeton, 1957).

[5] P.G. Burke, A program to calculate a general recoupling coefficient, Comput. Phys. Commun. 1 (1970) 241-250.

[6] A. Bar-Shalom and M. Klapisch, NJGRAF – An efficient program for calculation of general recoupling coefficients by graphical analysis, compatible with NJSYM. Comput. Phys. Commun. 50 (1988) 375-393.

[7] V. Fack, S.N. Pitre and J. Van der Jeugt, New efficient programs to calculate general recoupling coefficients. Part I: Generation of a summation formula, Comput. Phys. Commun. 83 (1994) 275-292.

[8] V. Fack, S.N. Pitre and J. Van der Jeugt, New efficient programs to calculate general recoupling coefficients. Part II: Evaluation of a summation formula, Comput. Phys. Commun. 86 (1995) 105-122.

[9] V. Fack, S.N. Pitre and J. Van der Jeugt, Calculation of general recoupling coefficients using graphical methods, Comput. Phys. Commun. 101 (1997) 155-170.

[10] D.F. Robinson, Comparison of labeled trees with valency three, J. Combinatorial Theory Ser. B 11 (1971) 105-119.

[11] K. Culik II and D. Wood, A note on some tree similarity measures, Inform. Process. Lett. 15 (1982) 39-42.

[12] M.S. Waterman and T.F. Smith, On the similarity of dendrograms, J. Theor. Biol. 73 (1978) 789-800.

[13] J.P. Jarvis, J.K. Luedeman and D.R. Shier, Comments on computing the similarity of binary trees, J. Theor. Biol. 100 (1983) 427-433.

[14] W.H.E. Day, Properties of the nearest neighbor interchange metric for trees of small size, J. Theor. Biol. 101 (1983) 275-288.

[15] R.P. Boland, E.K. Brown and W.H.E. Day, Approximating minimum-length-sequence metrics: a cautionary note, Math. Soc. Sci. 4 (1983) 261-270.

[16] F. Buckley and F. Harary, Distance in Graphs (Addison-Wesley, Reading MA, 1990).

[17] B.G. Wybourne, Classical Groups for Physicists (John Wiley & Sons, New York, 1974).

[18] J.-Q. Chen, P.-N. Wang, Z.-M. Lü and X.-B. Wu, Tables of the Clebsch-Gordan, Racah and subduction coefficients of SU(n) groups (World Scientific, Singapore, 1987).

[19] N.J.A. Sloane, On-Line Encyclopedia of Integer Sequences (URL : `http://www.research.att.com/~njas/sequences/`).

[20] A. Cayley, On the analytical forms called trees, with application to the theory of chemical combinations, Reports British Assoc. Advance. Sci. 45 (1875), 257-305.

[21] F. Harary, Graph Theory (Addison-Wesley, Reading MA, 1969).

[22] M. Li, J. Tromp and L. Zhang, On the nearest neighbour interchange distance between evolutionary trees, J. Theor. Biol. 182 (1996) 463-467.

[23] M. Li, J. Tromp and L. Zhang, Some notes on the nearest neighbour interchange distance, Lect. Notes Computer Science 1090 (1996), 343-351.

[24] V. Fack, S. Lievens and J. Van der Jeugt, On the diameter of the rotation graph of binary coupling trees, University of Ghent preprint (1999).

[25] D.D. Sleator, R.E. Tarjan and W.P. Thurston, Short encodings of evolving structures, SIAM J. Disc. Math. 5 (1992) 428-450.

[26] F. Murtagh, Counting dendrograms : a survey, Discrete Appl. Math. 7 (1984) 191-199.

[27] E.K. Brown and W.H.E. Day, A computationally efficient approximating to the nearest neighbor interchange metric, J. Class. 1 (1984) 93-124.

[28] W.H.E. Day, Optimal algorithms for comparing trees with labeled leaves, J. Class. 2 (1985) 7-28.

[29] M. Křivánek, Computing the nearest neighbor interchange metric for unlabeled binary trees is NP-complete, J. Class. 3 (1986) 55-60.

[30] D.D. Sleator, R.E. Tarjan and W.P. Thurston, Rotation distance, triangulations and hyperbolic geometry, J. Amer. Math. Soc. 1 (1988) 647-681.

[31] E. Mäkinen, On the the rotation distance of binary trees, Inform. Process. Lett. 26 (1987) 271-272.

[32] F. Luccio and L. Pagli, On the upper bound on the rotation distance of binary trees, Inform. Process. Lett. 31 (1989) 57-60.

[33] R.O. Rogers and R.D. Dutton, On distance in the rotation graph of binary trees, Congr. Numer. 120 (1996) 103-113.