Subschemes of the Johnson Scheme

M. E. Muzychuk

Let \( \mathcal{F} = (X, \{ R_i \}_{i=0}^{d'}) \) and \( \mathcal{F}' = (X', \{ R'_i \}_{i=0}^{d'}) \) be two association schemes defined on the same set \( X \). We say that \( \mathcal{F}' \) is a subscheme of \( \mathcal{F} \) if each relation \( R'_i \) is a union of some \( R_i \). The subscheme lattice of the Johnson scheme \( J(n, m) \) is studied. We prove that it is trivial for \( m = 3n + 4 \geq 13 \).

1. Association Schemes and Their Subschemes

This section contains the basic definitions and properties of association schemes [2].

Let \( \mathcal{F} = (X, \{ R_i \}_{i=0}^{d'}) \) be a symmetric association scheme (or simply scheme) with \( d \) classes defined on finite set \( X \) or cardinality \( v \). We refer to [2] for the general theory of association schemes. If \( d = 1 \) then the scheme \( \mathcal{F} \) will be called trivial.

Let \( A_i \) be the adjacency matrix with respect to the relation \( R_i \). The Bose-Mesner algebra \( \mathcal{A} = \{ A_0, A_1, \ldots, A_d \} \) is a semisimple commutative algebra (over \( \mathbb{C} \)) of dimension \( d + 1 \). It has a unique set of primitive idempotents \( E_0 = (1/v)I, E_1, \ldots, E_d \) (here, \( J \) is the matrix the entries of which are all 1). We shall say that \( A_0, \ldots, A_d \) and \( E_0, \ldots, E_d \) are first and second bases of \( \mathcal{A} = \{ A_0, \ldots, A_d \} \). Let \( P = (P_i(j)) \) and \( Q = vP^{-1} = (Q_i(j)), i, j = 0, d \) be the first and second eigenmatrices of the scheme \( \mathcal{F} \).

The following orthogonality conditions hold [2]:

\[
P_i(j) - \frac{Q_i(j)}{v} = \frac{\mu_i}{\mu_j}, \quad i, j = 0, 1, \ldots, d,
\]

where \( \mu_i = P_i(0) \) and \( \mu_j = Q_j(0) \).

\( \mathcal{F}' \) is a subscheme of \( \mathcal{F} \) if its Bose-Mesner algebra is contained in the Bose-Mesner algebra of \( \mathcal{F} \). Let us denote by \( \{ A'_0, A'_0, \ldots, A'_d \} \) and \( \{ E'_0, E'_0, \ldots, E'_d \} \), the first and second bases of the Bose-Mesner algebra of the subscheme \( \mathcal{F}' \). Then each matrix \( A'_i \) (dually \( E'_i \)) is a \((0, 1)\)-linear combination of \( A_i \) (resp. \( E_i \)). So we have two partitions \( \tau = \{ T_0 = \{ 0 \}, T_1, \ldots, T_{d'} \} \) and \( \pi = \{ \Pi_0 = \{ 0 \}, \Pi_1, \ldots, \Pi_{d'} \} \) of the index set \( \{ 0, 1, \ldots, d \} \) such that

\[
A'_i = \sum_{j \in T_i} A_j, \quad i = 0, 1, \ldots, d', \quad E'_i = \sum_{j \in \Pi_i} E_j, \quad i = 0, 1, \ldots, d'.
\]

It is clear that any subscheme of the scheme \( \mathcal{F} \) is uniquely determined by either of these two partitions.

The following lemma was independently proved in [1] and [8].

**Lemma 1.** Let \( \mathcal{F}' \) be a subscheme of a scheme \( \mathcal{F} \) and let \( \tau, \pi \) be the above-defined partitions of the set \( \{ 0, 1, \ldots, d \} \). Then, for each \( 0 \leq k \leq d' \) and any pair \( i, j \in T_k \) (dually \( i, j \in \Pi_k \)) the following holds:

\[
\sum_{h \in \Pi_l} Q_h(i) = \sum_{h \in \Pi_l} Q_h(j), \quad l = 0, 1, \ldots, d'
\]

(dually \( \sum_{h \in T_l} P_h(i) = \sum_{h \in T_l} P_h(j), l = 0, 1, \ldots, d' \)).

We omit the proof because it is very simple.
This lemma has a useful corollary.

**Corollary 1.** Let \( \mathcal{A} = (X, \{R_i\}_{i=0}^d) \) be an association scheme the first eigenmatrix \( P \) of which (dually the second eigenmatrix \( Q \)) satisfies the following condition.† There exist two numbers \( 0 < k, m \leq d \) such that, for each pair \( i, j \) with \( j \neq 0, m, i \neq 0, k \) it holds that
\[
P_j(k) > P_j(i)
\]
(dually \( Q_j(k) > Q_j(i) \)). Then the Bose–Mesner algebra of any non-trivial subscheme \( \mathcal{A}' \) contains the element \( E_k \) (dually \( R_k \)).

**Proof.** We shall consider the direct case only, because the dual one has the analogous proof.

Let \( \mathcal{A}' = (X, \{R'_i\}_{i=0}^d) \) be a non-trivial subscheme, let \( \mathcal{A}' = (A'_i)_{i=0}^d \) be its Bose–Mesner algebra, and let \( \tau \) and \( \pi \) be two partitions of the set \( \{0, 1, \ldots, d\} \) associated with \( \mathcal{A}' \). Since \( \mathcal{A}' \) is non-trivial, then there exists \( T \in \tau \) such that \( 0, m \notin T \). Let \( \Pi \) be a set from partition \( \pi \) containing the index \( k \). By definition, \( 0 \notin \Pi \). To prove our statement it is sufficient to show that \( \Pi = \{k\} \).

Suppose that \( \Pi \) contains another element \( s \neq 0, k \). Then, by Lemma 1, we have
\[
\sum_{j \in T} P_j(k) = \sum_{j \in T} P_j(s).
\]
On the other hand, the assumption (3) gives
\[
\sum_{j \in T} P_j(k) > \sum_{j \in T} P_j(s).
\]
This is a contradiction. Hence, \( \Pi = \{k\} \).

## 2. The Johnson Scheme and its Subschemes

Let \( M \) be a finite set of cardinality \( m \). For any integer \( n, 1 \leq n \leq m/2 \), we define the set \( (\begin{smallmatrix} n \\ m \end{smallmatrix}) = \{N \subseteq M \mid |N| = n\} \). The Johnson distance \( \rho(N_1, N_2) \) between two subsets \( N_1, N_2 \in (\begin{smallmatrix} n \\ m \end{smallmatrix}) \) is defined by formula \( \rho(N_1, N_2) = n - |N_1 \cap N_2| \) [2]. It is well known that the family of relations \( R_i = \{(N_1, N_2) \mid \rho(N_1, N_2) = i\} \) forms an association scheme, which is called the Johnson scheme [2]. We shall denote it by \( J(n, m) \).

The enumeration problem of subschemes of the Johnson scheme was first considered in [4]. It was proved in that paper that there is a function \( d(n) \) such that Johnson scheme \( J(n, m) \) with \( m \geq d(n) \) does not contain non-trivial subschemes. The authors used this result to prove the asymptotic maximality of permutation group \( (S(M), (\begin{smallmatrix} n \\ m \end{smallmatrix})) \) either in symmetric or alternating groups of degree \( C_m^n \). The complete list of \( (S(M), (\begin{smallmatrix} n \\ m \end{smallmatrix})) \) supergroups was obtained in [9]. One can use this result for the construction of non-trivial subschemes of the Johnson scheme. However, this list of subschemes will not be complete. The non-trivial subschemes of \( J(n, m) \) are known only for the following set \( R \) of pairs \( (n, m) \):
\[
R = \{(n, 2n) \mid n \in \mathbb{N}\} \cup \{(n, 2n + 1) \mid n \in \mathbb{N}\} \cup \{(3, 10), (4, 11), (4, 12), (6, 13)\}.
\]

All sporadic examples of non-trivial subschemes were discovered by M. H. Klin, using manual and computer calculations [6]. He also proved in [5] that the existence of a non-trivial subscheme in \( J(n, m) \) implies \( (n, m) \in R \) for pairs \( (n, m) \) satisfying either \( m \geq \sqrt[3]{n} n^3 - 4n^3 + 2n \) or \( n \leq 6 \). V. A. Ustimenko, using a computer, proved in [10] the non-existence of new non-trivial subschemes with \( d' = 2 \) for \( n \leq 20, m \leq 60 \). The problem of the existence of new non-trivial examples of subschemes in Johnson schemes mentioned in [7] is one of the most intriguing questions about subschemes of association schemes. The main goal of this paper is to prove the non-existence of non-trivial subschemes for \( m \geq 3n + 4 \geq 13 \) (Theorem 1).

† The identity \( \sum_{i=0}^{d} P_i(j) = \sum_{i=0}^{d} P_i(k) \) (\( i, k \neq 0 \)) implies the existence of at least one index \( m \) with \( P_m(k) \leq P_m(i) \). We demand the uniqueness of such \( m \).
Let us recall some numerical parameters of $J(n, m)$ [3]. We shall use the following notation:

- $P_i^{n,m}(j)$ is the $(i, j)$-entry of the first eigenmatrix of $J(n, m)$;
- $Q_i^{n,m}(j)$ is the $(i, j)$-entry of the second eigenmatrix of $J(n, m)$;
- $v_k(n, m)$ is the valency of the relation $R_k$;
- $\mu_k(n, m)$ is the rank of the idempotent $E_k$.

The values of these parameters are given by the following formulas [3]:

$$P_i^{n,m}(j) = \sum_{p=0}^{j} (-1)^p C_p^{i} C_{n-p}^{i} C_{m-n-p}^{i-p},$$

$$v_k(n, m) = C_k^k C_{m-n}, \quad \mu_k(n, m) = C_k^k - C_k^{k-1},$$

where

$$C_a^b = \begin{cases} \frac{a!}{b!(a-b)!} & \text{if } 0 \leq b \leq a \\ 0 & \text{otherwise} \end{cases}$$

**Theorem 1.** The Johnson scheme $J(n, m)$ contains no subscheme for $13 \leq 3n + 4 \leq m$.

The key lemma for the proof of this theorem is as follows:

**Lemma 2.** For any pair $(j, i)$, $0 < j < n$, $1 < i \leq n$, and $m \geq 3n + 4 \geq 13$ it holds that

$$Q_j^{n,m}(1) > |Q_j^{n,m}(i)|.$$  \hspace{1cm} (5)

We shall prove inequality (5) in the next section, but now we will prove Theorem 1.

**Proof of Theorem 1.** Let $\mathcal{E}' = (X, \{R'_i\}_{i=0}^{d'})$ be any subscheme of $J(n, m)$. It follows from Corollary 1 and Lemma 2 that $R_i$ belongs to the set $\{R'_i\}_{i=0}^{d'}$. The scheme $J(n, m)$ is $P$-polynomial, so the inclusion $R_i \in \{R'_i\}_{i=0}^{d'}$ implies $\mathcal{E}' = J(n, m)$. \hfill $\Box$

3. **The Proof of Lemma 2**

We shall prove inequality (5) by induction on $n$. This can be done due to the recursive formula [3]:

$$p_i^{n,m}(j) = \begin{cases} p_i^{n-1,m-2}(j-1) - p_i^{n-1,m-2}(j-1), & i = 1, \ldots, n - 1, \\ -p_i^{n-1,m-2}(j-1), & i = n, \end{cases}$$

which holds for each $1 \leq j \leq n$.

First we shall consider several special values of $j, i$ in (5).

**Proposition 1.** If $m \geq 3n + 4 \geq 13$, then

(i) $Q_j^{n,m}(1) > |Q_j^{n,m}(i)|, i = 2, \ldots, n$;

(ii) $Q_j^{n,m}(1) > |Q_j^{n,m}(n)|, j = 1, \ldots, n - 1$;

(iii) $Q_i^{n,m}(1) > |Q_i^{n,m}(i)|, i = 2, \ldots, n$. 

PROOF. Identity (1) enables us to replace inequality (5) by the equivalent one:

\[ \frac{P^i_{m,n}(j)}{v_i(n, m)} > \frac{|P^i_{m,n}(j)|}{v_i(n, m)}. \]

(i) By (4) we obtain

\[ \frac{P^i_{m,n}(n - 1)}{v_i(n, m)} = \frac{(-1)^iC_{n-1}^i + (-1)^{i-1}C_{n-1}^{i-1}(m - 2n + 1)}{C_n C_m^{i-1}} = (-1)^i C_{m-n}^{i-1} \frac{i(m - 2n + 2) - n}{n C_m^{i-1}}. \]

In particular, we have

\[ \frac{P^i_{m,n}(n - 1)}{v_i(n, m)} = \frac{m - 3n + 2}{n(m - n)}, \quad \frac{|P^i_{m,n}(n - 1)|}{v_i(n, m)} = \frac{i(m - 2n + 2) - n}{n C_m^{i-1}}. \]

Now we write the sequence of inequalities to estimate \(|P^i_{m,n}(n - 1)|/v_i(n, m)|.

\[ \frac{|P^i_{m,n}(n - 1)|}{v_i(n, m)} < \frac{(i + 1)(m - 2n + 2)}{n C_m^{i-1}} = \frac{(m - 2n + 2)(m - n + 1)}{n C_m^{i-1} + 1}. \]

Since \(3 \leq i + 1 \leq n + 1 < (m - n + 1)/2\), then \(C_m^{i+1} \geq C_m^{i-1} \) and, therefore,

\[ \frac{|P^i_{m,n}(n - 1)|}{v_i(n, m)} < \frac{(m - 2n + 2)(m - n + 1)}{n C_m^{i+1}} = \frac{m - 2n + 2}{n(m - n)(m - n - 1)}. \]

The number \(n\) is greater or equal to 3, so

\[ 6 \frac{m - 2n + 2}{n(m - n)(m - n - 1)} \leq \frac{m - 3n + 2}{n(m - n)} = \frac{P^i_{m,n}(n - 1)}{v_i(n, m)}. \]

This completes the proof of case (i).

(ii) The direct calculation gives us the equalities:

\[ P^i_{m,n}(j) = (-1)^jC_{m-n-j}, \quad v_i(n, m) = C_m^{i-1}, \]

\[ \frac{|P^i_{m,n}(j)|}{v_i(m, n)} = \frac{C_{m-n-j}}{C_{m-n+j}} = \frac{C_m^{2n-2j}}{C_{m-n}^{2j}} = \prod_{k=n+1}^{m-n}(1 - j/k). \]

By the assumption \(j \leq n - 1\), therefore, each factor in the product \(\prod_{k=n+1}^{m-n}(1 - j/k)\) is strictly less than 1. So we can write the inequality:

\[ \frac{|P^i_{m,n}(j)|}{v_i(m, n)} < \left(1 - \frac{j}{m - n}\right)\left(1 - \frac{j}{n + 3}\right)\left(1 - \frac{j}{n + 1}\right). \]

Now we will prove that the right-hand part is less than or equal to \(P^m_{1,n}(j)/v_1(n, m):

\[ \frac{P^m_{1,n}(j)}{v_1(n, m)} = \frac{(n - j)(m - n - j) - j}{n(m - n)} = \left(1 - \frac{j}{m - n}\right)\left(1 - \frac{j}{n}\right)\left(1 - \frac{j}{(n - j)(m - n - j)}\right). \]

Since \(j \leq n - 1\), then \((n - j)(m - n - j) \geq m - 2n + 1 > n + 3\). Therefore

\[ 1 - \frac{j}{n - j}(m - n - j) > 1 - \frac{j}{n + 3}. \]

To complete the proof it is sufficient to show that

\[ 1 - \frac{j}{n} \geq \left(1 - \frac{j}{n + 1}\right)\left(1 - \frac{j}{n + 2}\right). \]
Let us consider the difference

$$1 - \frac{j}{n} - \left(1 - \frac{j}{n+1}\right)\left(1 - \frac{j}{n+2}\right),$$

which is equal to

$$j \frac{(2n + 3 - j)n - (n + 1)(n + 2)}{n(n + 1)(n + 2)}.$$

Since \( j \leq n - 1 \), the last term is greater or equal to

$$j \frac{(2n + 3 - (n - 1))n - (n + 1)(n + 2)}{n(n + 1)(n + 2)} = \frac{(n - 2)}{n(n + 1)(n + 2)} \geq 0.$$

This inequality completes the proof of case (ii).

(iii) One can easily calculate the numbers \( P_i^{m,n}(1)/v_i(n, m) \) by (4):

$$P_i^{m,n}(1) = \frac{-mi + n(m - n)}{n(m - n)}.$$

Since \( \frac{-mi + n(m - n)}{n(m - n)} \) decreases when \( i \) increases, then it holds for each \( 2 \leq i \leq n \):

$$\frac{|P_i^{m,n}(1)|}{v_i(n, m)} \leq \max\left(\frac{-2m + n(m - n)}{n(m - n)}, \frac{n^2}{n(m - n)}\right).$$

But the right-hand part of this inequality is less than \( \frac{-m + n(m - n)}{n(m - n)} = P_i^{m,n}(1)/v_i(n, m) \) when \( m \geq 3n + 4 \geq 13 \).

Now we are able to prove Lemma 2.

It was mentioned above that we shall prove the inequality \( P_i^{m,n}(j)/v_i(n, m) > |P_i^{m,n}(j)|/v_i(n, m), 2 \leq i \leq n, 1 \leq j \leq n - 1 \) by induction on \( n \).

First we verify the inductive hypothesis for \( n = 3 \). But there is nothing to prove, because all the possible values of \( j \) in this case were considered in Proposition 1.

Therefore, our hypothesis is valid for \( n = 3 \).

Let us now consider the general case. Due to Proposition 1 we may assume that \( 1 < j < n - 1, 2 \leq i < n \). So we can write \( P_i^{m,n}(j) = P_i^{n-1,m-2}(j - 1) - P_i^{n-1,m-2}(j - 1) \) according to (6). Since \( m - 2 > 3(n - 1) + 4 \), then we can use the inductive hypothesis for estimating both of \( P_i^{n-1,m-2}(j - 1) \) and \( P_i^{n-1,m-2}(j - 1) \):

$$\left| P_i^{n-1,m-2}(j - 1) \right| < \frac{P_i^{n-1,m-2}(j - 1)}{v_i(n - 1, m - 2)} \leq \frac{P_i^{n-1,m-2}(j - 1)}{v_i(n - 1, m - 2)}.$$

From these inequalities we obtain:

$$\frac{\left| P_i^{n,m}(j) \right|}{v_i(n, m)} < \frac{v_i(n - 1, m - 2) + v_{i-1}(n - 1, m - 2)}{v_i(n, m)} \frac{P_i^{n-1,m-2}(j - 1)}{v_i(n - 1, m - 2)} \frac{P_i^{n-1,m-2}(j - 1)}{v_i(n - 1, m - 2)}.$$

Therefore, our statement will be proved if we can show that

$$\frac{v_i(n - 1, m - 2) + v_{i-1}(n - 1, m - 2)}{v_i(n, m)} \frac{P_i^{n-1,m-2}(j - 1)}{v_i(n - 1, m - 2)} \frac{P_i^{n-1,m-2}(j - 1)}{v_i(n - 1, m - 2)} \leq 1.$$
To do this, let us use (4), which gives the expressions:

\[ v_i(n-1, m-2) = C_{n-1}^{i-1}C_{m-n-1}^{i-1}, \quad v_{i-1}(n-1, m-2) = C_{n-1}^{i-1}C_{m-n-1}^{i-1}, \]

\[ v_i(n, m) = C_n^iC_{m-n}^i, \quad v_i(n-1, m-2) = (n-1)(m-n-1). \]

\[ v_1(n, m) = n(m-n), \]

\[ P_{i-1}^{n-1,m-2}(j-1) = (n-j)(m-n-j) - (j-1), \]

\[ P_1^{n,m}(j) = (n-j)(m-n-j) - j. \]

After substitution of these expressions into the left-hand part of the above inequality, we obtain the following one:

\[ \frac{(n-i)(m-n-i) + i^2}{(n-1)(m-n-1)} \cdot \frac{n(m-n) + j^2 - j + 1 - jm}{n(m-n) + j^2 - j - jm} \leq 1. \]

Since \( 1 < j < n-1, \leq i \leq n-1 \) and \( m \geq 3n + 4 \geq 13 \), then the left-hand part reaches its maximum if \( j = n-2, i = 2 \). Its value is equal to

\[ \left(1 + \frac{1}{2m-5n+6}\right)\left(1 - \frac{m-7}{(n-1)(m-n-1)}\right). \]

To complete the proof it is sufficient to show that:

\[ \frac{1}{2m-5n+6} \leq \frac{m-7}{(n-1)(m-n-1)}. \]

Since \( m \geq 3n + 4 \), then

\[ \frac{1}{2m-5n+6} < \frac{1}{n+1}. \]

But it is not difficult to prove that

\[ \frac{1}{n+1} < \frac{m-7}{(n-1)(m-n-1)} \quad \text{if} \quad m \geq 3n + 4. \]

\[ \square \]

ACKNOWLEDGEMENTS

The author would like to thank M. H. Klin, A. A. Ivanov and V. A. Ustimenko for their support of this work.

REFERENCES

3. P. Delsarte, Properties and applications of the recurrence \( F(i+1, k+1, n+1) = q^{k+1}F(i, k+1, n) - q^kF(i, k, n) \), SIAM J. Appl. Math., 31(2) (1976), 262–270.
Subschemes of the Johnson scheme


Received 17 April 1991 and accepted in revised form 10 December 1991

M. E. MUZYCHUK

Department of Mathematics and Computer Science,
Bar-Ilan University,
Ramat-Gan, 52900, Israel