BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 43, Number 2, Pages 249–253
S 0273-0979(05)01077-3
Article electronically published on October 7, 2005

*Association schemes. Designed experiments, algebra and combinatorics*, by Rosemary A. Bailey, Cambridge University Press, Cambridge, 2004, xvii+387 pp., US$70.00, ISBN 0-521-82446-X

An *association scheme* (or a *scheme* as we shall say briefly) is a mathematical structure that has been created by statisticians [3], which, during the last three decades, has been intensively investigated by combinatorialists [2], and which, finally, turned out to be an algebraic object generalizing in a particularly natural way the concept of a group [4], [8].

The definition of a scheme could easily catch the interest of an undergraduate student: given a set $X$, one calls a partition $S$ of the cartesian product $X \times X$ a scheme on $X$ if $S$ satisfies the following three conditions:

- The identity $1_X := \{(x, x) \mid x \in X\}$ belongs to $S$.
- For each element $s$ in $S$, $s^* := \{(y, z) \mid (z, y) \in s\}$ belongs to $S$.
- Let $s$ be an element in $S$, and let $(y, z)$ be an element in $s$. Then, for any two elements $p$ and $q$ in $S$, the number of elements $x$ with $(y, x) \in p$ and $(x, z) \in q$ does not depend on $y$ or $z$, only on $s$.

The cardinal numbers arising from the last condition are usually denoted by $a_{pqs}$ and called the *structure constants* of $S$.

The elements of $S$ are defined to be subsets of $X \times X$. This means that they are binary relations on $X$. Thus, assuming (for the moment) that $X$ is a finite set, the elements of $S$ can be viewed as 0-1-matrices. Let us (in this case) denote by $M_s$ the 0-1-matrix which is associated (in the usual way) to the element $s$ in $S$. It is easy to see that the last condition is equivalent to the fact that, for any two elements $p$ and $q$ in $S$,

$$M_p M_q = \sum_{s \in S} a_{pqs} M_s.$$

In particular, the matrices $M_s$ with $s \in S$ form a linear basis of the (semisimple) $F$-algebra $FS$ generated by the matrices $M_s$ ($F$ a field of characteristic 0).

It is clear that the structure constants $a_{ss^*1}$ have a distinguished meaning. Indeed, it is easy to see that, for each element $s$ in $S$, $n_s := a_{ss^*1}$ is the valency of $s$ if $s$ is viewed as a (directed) graph on $X$.

Let us look at an example. Let $V$ stand for the ten subsets of cardinality 2 of a given set of cardinality 5. The *Petersen graph* is defined to be the graph on $V$ which connects two elements in $V$ by an edge if and only if their set theoretic intersection is empty.

The Petersen graph gives rise to an association scheme. By $p$ we denote the set of all pairs of elements in $V$ having distance 1, and by $q$ we denote the set of all pairs of elements in $V$ having distance 2. Then $P := \{1_V, p, q\}$ is a scheme on $V$ (with $n_p = 2$ and $n_q = 6$). Clearly, we also have $p^* = p$ and $q^* = q$, and a scheme $S$ which satisfies $s^* = s$ for each of its elements $s$ is called *symmetric*. It is easy to see that the (above-mentioned) associative algebra $FS$ is commutative if $S$ is symmetric. Thus, the representation theory of $FS$ reduces to computations with eigenvalues.

2000 *Mathematics Subject Classification.* Primary 05E30, 62K10.

The scheme $P$ which we obtained from the Petersen graph has another interesting feature. One easily computes that $M_p^2 = M_q + 3M_{1_V}$, and schemes satisfying a polynomial equation similar to this one always come from so-called *distance-regular graphs*. (The Petersen graph is distance-regular.) The interest of combinatorialists in symmetric schemes originates from their interest in distance-regular graphs, since distance-regular graphs are related to certain aspects of codes and designs.

Let us now look at a major variety of examples. Let $G$ be a group. For each element $g$ in $G$, we define $\tilde{g}$ to be the set of all pairs $(e, f)$ where $e$ and $f$ are elements in $G$ satisfying $eg = f$. It is easy to see that the set of all sets $\tilde{g}$ with $g \in G$ is a scheme on $G$. Moreover, each element of this scheme has valency 1, and nonempty subsets of a scheme in which each element has valency 1 are called *thin*. This shows that each group gives rise to a thin scheme. It does not require much to see that, conversely, each thin scheme gives rise to a group and that these two constructions are inverse to each other. Thus, the class of groups can be viewed (in a natural way) as a distinguished class of schemes, namely as the class of the thin schemes.

It is tempting to consider this observation as a justification for far-reaching and ambitious conjectures. One would like to know to which extent basic group theoretic definitions and results can be generalized to scheme theory in such a way that the thin version of the scheme theoretic generalizations correspond to the group theoretic originals one starts with.

In fact, scheme theory allows us to do quite a few steps in this direction. There is a *Lagrange Theorem* for schemes, there is a *Homomorphism Theorem* for schemes, there are two *Isomorphism Theorems*, there is a *Jordan-Hölder Theorem* for schemes; and even *Sylow's Theorem* [6] on finite groups can be generalized in a natural way to scheme theory. Moreover, all of these theorems satisfy the above-mentioned requirement that their thin version is exactly the group theoretic result one starts with.

The generalization of Lagrange's observation (that the order of a subgroup of a finite group $G$ divides the order of $G$) to scheme theory is easy to understand. One just has to find the appropriate generalization of the notion of a subgroup of a group. In order to establish this generalization one defines, for any two elements $p$ and $q$ of a scheme $S$, $pq$ to be the set of all elements $s$ in $S$ such that $a_{pqs} \neq 0$. A nonempty subset $R$ of a scheme $S$ is called *closed* if, for any two elements $p$ and $q$ in $R$, $p^*q \subseteq R$.[1] If $S$ is thin, closed subsets of $S$ correspond to subgroups of the group which corresponds to $S$.

For each nonempty subset $R$ of $S$, we now write $n_R$ in order to denote the sum of the cardinal numbers $n_r$ with $r \in R$ and call $n_R$ the *valency* of $R$. If $S$ is thin, we have $n_T = |T|$ for each closed subset $T$ of $S$. Thus, in this case, Lagrange's Theorem on finite groups says that, for each closed subset $T$ of a scheme $S$ of finite valency, $n_T$ divides $n_S$. This divisibility condition holds, in fact, for any scheme $S$ of finite valency and any closed subset $T$ of $S$ (and not only for thin schemes of finite valency).

The above-mentioned generalization of Sylow's theorems to schemes is particularly amazing. In order to look at this generalization we fix a prime number $p$ and a scheme $S$. An element $s$ in $S$ is called *$p$-valenced* if $n_s$ is a power of $p$. A subset of $S$

[1] Recall that $p^*$ is defined to be the set of all pairs $(y, z)$ of elements in $X$ such that $(z, y) \in p$.

is called *$p$-valenced* if each of its elements is $p$-valenced, and a nonempty $p$-valenced subset $R$ of $S$ is called a *$p$-subset* if $n_R$ is a power of $p$.

Assume $n_S$ to be finite, and let $\mathrm{Syl}_p(S)$ stand for the set of all closed $p$-subsets $T$ of $S$ such that $p$ does not divide $n_S/n_T$. Then, if $S$ is $p$-valenced, $|\mathrm{Syl}_p(S)| \equiv 1$ $(p)$; cf. [5]. Since thin schemes are obviously $p$-valenced for each prime number $p$, this result generalizes Sylow's famous theorems on finite groups.

Let us now briefly show how far Tits' theory of buildings can be considered as part of scheme theory. In order to do so we fix a scheme $S$. For each subset $R$ of $S$, we define $\langle R\rangle$ to be the intersection of all closed subsets of $S$ which contain $R$. An element $s$ of $S$ will be called an *involution* if $|\langle\{s\}\rangle| = 2$.[2]

Let us fix a set $L$ of involutions of $S$, and let us assume that $\langle L\rangle = S$. Simulating group theoretic reasoning, one easily proves that $S$ is the union of the sets $L^n$, $n$ a nonnegative integer. Thus, for each element $s$ in $S$, there exists a smallest integer $n$ such that $s \in L^n$. We denote this element by $\ell(s)$.

For each element $q$ in $S$, we define $S_L(q)$ to be the set of all elements $p$ in $S$ such that there exists an element $r$ in $pq$ with $\ell(r) = \ell(p) + \ell(q)$. One calls $S$ *constrained with respect to* $L$ if, for any two elements $q$ in $S$ and $p$ in $S_L(q)$, $|pq| = 1$.

Let us assume $S$ to be constrained with respect to $L$. The scheme $S$ is called a *Coxeter scheme with respect to* $L$ if, for any three elements $h$, $k$ in $L$ and $s$ in $S_L(k)$, $h \in S_L(s)$ implies that $hs \subseteq sk \cup S_L(k)$.[3]

We saw earlier that groups may be viewed as thin schemes. Similarly, one can prove that buildings (in the sense of Tits) may be viewed as Coxeter schemes. (The identification is based on the concept of a *coset geometry.*) Referring to this identification, one may wish to know how Tits' famous result on buildings of spherical type [7] sounds scheme theoretically.

Assume that $S$ is a Coxeter scheme with respect to $L$. Assume that $S$ has finitely many elements and that $L$ has at least three elements, none of them thin. Tits' theorem says then that $S$ arises from a (simple algebraic or classical) group.

There are several other sufficient conditions for schemes to arise from groups, and the question which schemes (precisely) arise from groups is one of the more challenging questions in scheme theory.

The fact that, indeed, large classes of schemes arise from groups seems to be a hint that there is more behind the relationship between groups and schemes than we presently know. Who knows, maybe one day in the future, geometric arguments in finite group theory can be replaced with a fully developed scheme theory and classification theorems in group theory will emerge naturally as parts of a well-understood scheme theory.

Presently, we are still far away from this. The literature about scheme theory is quite limited and still modest in comparison with other branches of algebra. Under these circumstances, it is gratifying to learn that Rosemary Bailey wrote a book about schemes (cf. [1]), and we curiously ask ourselves how her book *Association Schemes. Designed Experiments, Algebra and Combinatorics* fits into the picture.

The author gives the answer to this question right in the beginning of the preface: the book is supposed to bridge a gap, she says, between practicing statisticians and pure mathematicians. According to our introductory remarks this tells us that the

[2]Clearly, involutions of thin schemes correspond to group theoretic involutions.

[3]Again, the definition of Coxeter schemes is made in such a way that thin Coxeter schemes correspond to Coxeter groups.

author comes back to the historical roots of scheme theory, namely to commutative association schemes in statistics. And here, the author shows how much theory and terminology have developed since the relationship between schemes and statistics was first considered.

The content of Bailey's book is an interesting collection of interplays between statistics and scheme theory. It starts with three chapters of basic scheme theory. The definition (carefully presented under all possible aspects) comes first, together with a selection of standard examples in commutative scheme theory (the Petersen graph included). The second chapter is a standard introduction to basic (real-valued) representation theoretic aspects of finite commutative schemes. The third chapter deals with *direct products* of finite commutative schemes, *wreath products*, and combinations of the two. At the end of the third chapter, we see how schemes are used in order to design experiments.

The next four chapters refer only loosely to the first three chapters. They bring the terminology of statistics into the game and deal with *incomplete-block designs*, with *partial balance*, with *orthogonal block structures*, and with the relationship between the notion of a *design* and the notion of an orthogonal block structure. Key words are *efficiency factors*, estimating the *treatment parameters*, *variance* of certain *estimators*, *canonical efficiency factors* of partially balanced incomplete-block designs, *efficiency of estimation*, and *randomization* of the choice of designs.

In the last four of the eleven theoretical chapters of the book, the author comes back to a more abstract approach to schemes. Indeed, Chapters 8, 10, and 11 can be considered as a contribution to the theory of symmetric finite schemes. The author talks about the correspondence between certain partitions of a finite group $G$ and schemes defined on $G \times G$ via convolution, about subschemes and quotient schemes of finite commutative schemes, and about the question of how schemes on the same set are related.

Bailey's book concludes with a chapter on possible generalizations of some of the topics which have been discussed in the book (on the statistical side as well as on the scheme theoretic side) and with a chapter containing a collection of names, little stories, references, and concepts which are related to schemes and/or the design of experiments which have been discussed in the previous chapters.

The (main) title of Bailey's book is *Association Schemes*. However, everybody who picks up Bailey's book will realize immediately that this title does not exactly reflect the content of the book. In fact, taking into account the efforts which have been made by so many graph theorists, combinatorialists, and algebraists in order to explore the notion of a scheme and to understand its significance, one cannot call Bailey's book a book about association schemes. It would be more appropriate to look at it as an introduction to commutative association schemes as they occur in designed experiments. A title like *Association Schemes in Designed Experiments* or *Commutative Association Schemes in Designed Experiments* would describe the content of Bailey's book more precisely. Association schemes as they occur in designed experiments form the heart of the book, as the author says correctly in the preface. As such Rosemary Bailey's book is a valuable contribution to one of the important aspects of commutative scheme theory: the application of commutative scheme theory to statistics. Scheme theorists will be grateful for this contribution, because it fills one of the many gaps in the global view of schemes and their central role in mathematics.

## References


[1] R. A. Bailey: *Association Schemes. Designed Experiments, Algebra and Combinatorics*, Cambridge Stud. Adv. Math. **84**, University Press, Cambridge, 2004. MR2047311 (2005d:05001)

[2] E. Bannai and T. Ito: *Algebraic Combinatorics. I*, Benjamin-Cummings Lecture Note Ser. **58**, The Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA, 1984. MR0882540 (87m:05001)

[3] R. C. Bose and T. Shimamoto: Classification and analysis of partially balanced incomplete block designs with two associate classes, *J. Amer. Statist. Assoc.* **47**, 151-184 (1952). MR0048772 (14:67b)

[4] D. G. Higman: Coherent configurations. I. Ordinary representation theory, *Geom. Dedicata* **4**, 1-32 (1975). MR0398868 (53:2719)

[5] M. Hirasaka, M. Muzychuk, and P.-H. Zieschang: A generalization of Sylow's theorems on finite groups to association schemes, *Math. Z.* **241**, 665-672 (2002). MR1942234 (2003k:05141)

[6] L. Sylow: Théorèmes sur les groupes de substitutions, *Math. Ann.* **5**, 584-94 (1872).

[7] J. Tits: *Buildings of Spherical Type and Finite BN-Pairs*, Springer Lecture Notes in Math. **386**, Berlin-Heidelberg-New York (1974). MR0470099 (57:9866)

[8] P.-H. Zieschang: *An Algebraic Approach to Association Schemes*, Springer Lecture Notes in Math. **1628**, Berlin-Heidelberg-New York (1996). MR1439253 (98h:05185)



Paul-Hermann Zieschang

University of Texas at Brownsville

*E-mail address*: zieschang@utb.edu