Design Theory

Second Edition

Thomas Beth
Universität Karlsruhe
Dieter Jungnickel
Universität Augsburg
Hanfried Lenz
Freie Universität Berlin

Volume I

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011-4211, USA www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain

First published 1999

Printed in the United Kingdom at the University Press, Cambridge

Typeset in Times Roman 10/13pt. in LaTeX 2ε [TB]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data

Beth, Thomas, 1949–
Design theory / Thomas Beth, Dieter Jungnickel, Hanfried Lenz. – 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0 521 44432 2 (hardbound)
1. Combinatorial designs and configurations. I. Jungnickel, D. (Dieter), 1952– . II. Lenz, Hanfried. III. Title.
QA166.25.B47 1999
511′.6 – dc21 98-29508 CIP

ISBN 0 521 44432 2 hardback

Contents

Contents of Volume II

I

Examples and Basic Definitions

It's elementary, Watson
(Conan Doyle)

§1. Incidence Structures and Incidence Matrices

The most basic notion in (finite) geometry is that of an incidence structure. It contains nothing more than the idea that two objects from distinct classes of things (say points and lines) may be "incident" with each other. The only requirement will be that the classes do not overlap. We now make this more precise.

1.1 Definitions. An *incidence structure* is a triple $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ where V and $\boldsymbol{B}$ are any two disjoint sets and I is a binary relation between V and $\boldsymbol{B}$, i.e. $I \subseteq V \times \boldsymbol{B}$. The elements of V will be called *points*, those of $\boldsymbol{B}$ *blocks* and those of I *flags*. Instead of $(p, B) \in I$, we will simply write $p\,I\,B$ and use such geometric language as "the point p lies on the block B", "B passes through p", "p and B are incident", etc.

For reasons of convenience, we will usually not state whether a given object is a point or a block; this will be clear from the context and we will always use lower case letters (e.g. $p, q, r, \ldots$) to denote points and upper case letters (e.g. $B, C, \ldots$) to denote blocks. Now let us look at some examples! Of course, familiar (euclidean) geometry provides examples, e.g. taking points and lines (as blocks) in the euclidean plane or points and planes (as blocks) in 3-space. One reason for choosing the term "block" instead of "line" is that we will very often consider planes or hyperplanes as blocks. These classical examples are of course infinite; in this book we will almost exclusively deal with *finite* incidence structures (i.e. both V and $\boldsymbol{B}$ are finite). According to our definition, any binary relation between two disjoint finite sets will give an example. Naturally, this is much too general to be of any interest in itself and, guided by a series

of examples, we will single out classes of incidence structures having more interesting properties. Before doing so, let us introduce some notation. If p is any point, (p) will denote the set of blocks incident with p, i.e.

(1.1.a) $\qquad (p) := \{B \in \boldsymbol{B} : p\, I\, B\}$

and more generally for any subset Q of the point set

(1.1.b) $\qquad (Q) := \{B \in \boldsymbol{B} : p\, I\, B \text{ for each } p \in Q\}.$

Instead of $(\{p_1, \ldots, p_m\})$, we simply write $(p_1, \ldots, p_m)$ when it is clear that this symbol is not intended to denote the ordered m-tuple of these points. Similarly, we write

(1.1.c) $\qquad (\boldsymbol{C}) := \{p \in V : p\, I\, B \text{ for each } B \in \boldsymbol{C}\}$

for any subset $\boldsymbol{C}$ of the block set $\boldsymbol{B}$. For a point p, the number $|(p)|$ is called the *degree* of p, and similarly for blocks. Distinct blocks G and H may well be incident with the same point set; thus $(V, \boldsymbol{B}, I)$ with $V = \{1, 2, 3\}$, $\boldsymbol{B} = \{B_1, B_2, B_3\}$ and $(B_1) = (B_2) = \{1, 2\}$, $(B_3) = \{1, 3\}$ is a perfectly respectable incidence structure. But in this case one may not just list the sets of points incident with a block but one has to give their multiplicities too. If there are distinct blocks with the same point set one speaks of "repeated blocks". Often we will consider incidence structures where distinct blocks have distinct point sets.

1.2 Definition. An incidence structure is called *simple* if $(B) \neq (C)$ whenever B and C are distinct blocks. Here the *trace* (B) of a block B is the set $\{x \in V : x\, I\, B\}$ of its points, cf. (1.1.c).

1.3 Example. Take as point set $V = \{0, \ldots, 6\}$, as block set $\boldsymbol{B} = \{\{0, 1, 3\}, \{1, 2, 4\}, \{2, 3, 5\}, \{3, 4, 6\}, \{4, 5, 0\}, \{5, 6, 1\}, \{6, 0, 2\}\}$ and as incidence relation the membership relation $\in$.

For any simple incidence structure, we can (and usually will) identify each block B with the corresponding point set (B) and the incidence relation I with the membership relation $\in$. However, the way we have written our example is not intuitively appealing. So one can sometimes draw a picture, representing points by points in the real plane and blocks by point sets, usually drawn somewhat like

Venn diagrams or alternatively represented by curves in the plane. For instance, we may draw Example 1.3 as follows.

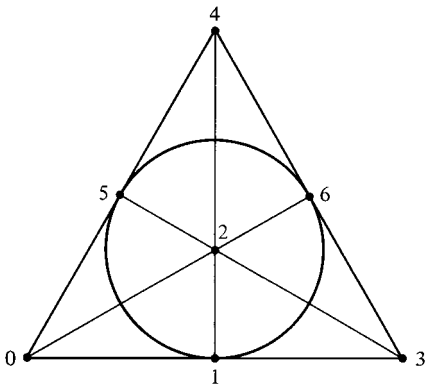

Figure 1.1

Later in this book, we will usually simplify notation and omit the symbol I for the incidence relation. Thus we just write $\boldsymbol{D} = (V, \boldsymbol{B})$ instead of $\boldsymbol{D} = (V, \boldsymbol{B}, I)$, even if $\boldsymbol{D}$ is not simple. Then $\boldsymbol{B}$ should be considered as a "multiset" or "list"; see III.9.2 for a formal definition. Finally, we introduce another way of representing an incidence structure, which will later allow us to use methods from linear algebra.

1.4 Definition. Let $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ be a finite incidence structure and label the points as $p_1, \ldots, p_v$ and the blocks as $B_1, \ldots, B_b$. Then the matrix $M = (m_{ij})$ $(i = 1, \ldots, v;\ j = 1, \ldots, b)$ defined by

$$m_{ij} := \begin{cases} 1 & \text{if } p_i I B_j \\ 0 & \text{otherwise} \end{cases}$$

is called an *incidence matrix* for $\boldsymbol{D}$. The column of M belonging to a block B is called the *incidence vector* of B.

Instead of m_{ij} one may write $M(p, B)(p \in V, B \in \boldsymbol{B})$. Then M is a mapping of $V \times \boldsymbol{B}$ into $\{0, 1\}$.

Of course, M depends on the labelling used, but up to column and row permutations it is unique. Conversely, every matrix with entries from $\{0, 1\}$ determines an incidence structure.

1.5 Example. The incidence structure of 1.3 yields the following incidence matrix.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

An interesting question one may ask about finite incidence structures in general is to find necessary and sufficient conditions for the existence of an incidence structure with given point degrees and block degrees (or in matrix language, for the existence of a (0, 1)-matrix of type $v \times b$ with given row and column sums). The answer to this question is not trivial and its proof requires the methods of transversal theory. Proofs may be found in the books of Mirsky (1971) and Jungnickel (1999). Here we will only give one necessary (in general not sufficient) condition, as its proof is an example of a very important method in finite geometry.

1.6 Proposition. *Assume the existence of an incidence structure with point degrees $r_1, \ldots, r_v$ and block degrees $k_1, \ldots, k_b$. Then necessarily*

$$\sum_{i=1}^{v} r_i = \sum_{j=1}^{b} k_j. \tag{1.6.a}$$

Proof. We count all flags (i.e. pairs (p_i, B_j) with $p_i I B_j$) in two ways. Using the fact that the i-th point is on exactly r_i blocks, the left-hand side of (1.6.a) counts the total number of flags. Similarly, since the j-th block B_j contains exactly k_j points, the right-hand side of (1.6.a) also counts the total number of flags. Hence these two quantities agree. ■

1.7 Corollary. *Let $\mathbf{D}$ be an incidence structure with v points, b blocks, all point degrees equal to r and all block degrees equal to k. Then*

$$vr = bk. \tag{1.7.a}$$ ■

The principle of counting in two ways will be used throughout this book and is – in spite of its simplicity – an extremely powerful tool (especially when combined with statistical ideas, like computing means and variances).

We conclude this section with a first non-trivial result, namely by settling the existence problem for simple incidence structures with constant point degrees r and constant block degrees k; in the terminology we shall introduce in Definition 3.1 below, such a structure will be called a *simple* 1-*design*. The simple proof we present is due to Billington (1982). We begin with the following lemma concerning the case of not necessarily uniform point degrees.

1.8 Lemma. *Suppose the existence of a simple incidence structure $\boldsymbol{D}$ on v points with b blocks, constant block degrees k and point degrees $r_1, \ldots, r_v$. If one has $r_i > r_j$ for a pair of indices i, j (with $i > j$, say), then there also exists a simple incidence structure $\boldsymbol{D}'$ on v points with b blocks, constant block degrees k and point degrees $r_1, \ldots, r_i - 1, r_{i+1}, \ldots, r_{j-1}, r_j + 1, \ldots, r_v$.*

Proof. The hypothesis $r_i > r_j$ implies that $\boldsymbol{D}$ has more blocks containing the point p_i but not the point p_j than blocks containing p_j but not p_i. This guarantees the existence of some block B of $\boldsymbol{D}$ which contains p_i but not p_j and such that $B^* := (B \setminus \{p_i\}) \cup \{p_j\}$ is not a block of $\boldsymbol{D}$. We now replace the block B by B^* and obtain the desired incidence structure $\boldsymbol{D}'$. ■

1.9 Theorem. *A simple incidence structure on v points with b blocks, constant block degrees k and constant point degrees r exists if and only if*

$$\text{(1.9.a)} \qquad vr = bk \quad \textit{and} \quad b \leq \binom{v}{k}.$$

Proof. In view of Corollary 1.7 and the simplicity requirement, condition (1.9.a) is obviously necessary for the existence of the desired incidence structure. Conversely, assume the validity of condition (1.9.a) for some integers k, r, v, b. We first define a simple incidence structure $\boldsymbol{D}_0$ by choosing any set of b pairwise distinct k-subsets $\{B_1, \ldots, B_b\}$ from some v-set $V = \{p_1, \ldots, p_v\}$. Denote the point degrees of $\boldsymbol{D}_0$ by $r_1, \ldots, r_v$; obviously,

$$\text{(1.9.b)} \qquad vr = bk = r_1 + \ldots + r_v.$$

If all point degrees happen to agree with r, we are finished. Otherwise, we may apply Lemma 1.8 to reduce one degree larger than r by 1 while simultaneously enlarging some other degree smaller than r by 1. Applying this process recursively yields the desired simple incidence structure $\boldsymbol{D}$. ■

1.10 Definition. The *complementary structure* of an incidence structure $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ is the incidence structure $\overline{\boldsymbol{D}} = (V, \boldsymbol{B}, J)$ with $J = (V \times \boldsymbol{B}) \setminus I$, that

is,

(1.10.a) $\quad x\,J\,B \iff x \not\mathrel{I} B \quad$ for all $x \in V$ and all $B \in \boldsymbol{B}$.

If M is an incidence matrix for $\boldsymbol{D}$, then an incidence matrix for $\overline{\boldsymbol{D}}$ is obtained by interchanging 0's and 1's in M.

1.11 Exercise. Let $\boldsymbol{D}$ be an incidence structure for which each point is on exactly r blocks and any two points are on exactly λ common blocks. Show that any two points are on exactly $b - 2r + \lambda$ common blocks of $\overline{\boldsymbol{D}}$.

1.12 Remark. For simplicity of notation, we will very often write incidence structures in the form $\boldsymbol{D} = (V, \boldsymbol{B})$, even if $\boldsymbol{D}$ is not simple. Then $\boldsymbol{B}$ is a multiset of blocks. Thus the incidence relation will not be mentioned explicitly in general.

§2. Block Designs and Examples From Affine and Projective Geometry

We now study some geometric examples which will lead us to the fundamental concept of a block design. We first look at (finite) projective and affine planes and prove some of their basic properties, although we will not make any detailed geometric investigations. The interested reader may consult the standard text books by Pickert (1975) and by Hughes and Piper (1982). Let us define a projective plane:

2.1 Definition. An incidence structure $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ is called a *projective plane* if and only if it satisfies the following axioms:

(2.1.a) Any two distinct points are joined by exactly one line.[1]

(2.1.b) Any two distinct lines intersect in a unique point.

(2.1.c) There exists a *quadrangle*, i.e. four points no three of which are on a common line.

We have already met an example in the previous section: Figure 1.1 represents a projective plane with seven points and seven lines. (In fact this is the smallest

[1] In geometry texts one usually speaks of "lines" instead of "blocks". In design theory, this will only be done for incidence structures for which any two points are joined by at most one block.

possible projective plane as the reader may check by considering the axioms.) This example shows a remarkable degree of uniformity: each point is on three lines, each line contains three points, and the number of points and lines agrees. Before giving more examples, we show that this is not a coincidence:

2.2 Proposition. *Let $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ be a finite projective plane. Then there exists a natural number n, called the* order *of $\boldsymbol{D}$, satisfying:*

(2.2.a) $|(p)| = |(G)| = n + 1$ *for all $p \in V$ and $G \in \boldsymbol{B}$;*

(2.2.b) $|V| = |\boldsymbol{B}| = n^2 + n + 1$.

Proof. Consider any point p and any line G with $p \not I G$. By (2.1.a) and (2.1.b), the mapping $\pi : (G) \to (p)$ with $q^\pi := pq$ for all $q I G$ is a bijection (here pq denotes the unique line joining p and q). Hence $|(p)| = |(G)|$ whenever $p \not I G$. Thus (2.2.a) follows if we can show that for any two distinct lines G, H there is a point p with $p \not I G, H$. But by (2.2.c) there exists a quadrangle q, r, s, t. If each of these points is on G or H, we may assume that $q, r I G$ and $s, t I H$. Then we can choose p to be the intersection of the lines qs and rt. To verify (2.2.b) we again use counting in two ways. We choose a fixed point p and count all flags (q, G) with pIG and $p \neq q$. By (2.1.a), we obtain $|V| - 1$ such flags; and by (2.2.a) we have $(n + 1)n$ such flags. Thus $|V| = n^2 + n + 1$. The assertion on $|\boldsymbol{B}|$ follows similarly (or using (2.2.a), the value for $|V|$ and (1.7.a)). ■

Now let us construct some more examples:

2.3 Proposition. *For each prime power q, there exists a projective plane of order q.*

Proof. Let F be the Galois field on q elements and W the vector space of dimension 3 over F. Choose as points all 1-dimensional subspaces and as lines all 2-dimensional subspaces of W. Using the dimension formula of linear algebra, one checks that the axioms (2.1.a) and (2.1.b) are satisfied. For (2.1.c), one may choose the points e_1F, e_2F, e_3F and $(e_1 + e_2 + e_3)F$ where e_1, e_2, e_3 is any basis of W. This argument works in fact for any 3-dimensional vector space. The fact that F is the field on q elements is only needed to show that the resulting projective plane has order q: the number of 1-dimensional subspaces of W is then $(q^3 - 1)/(q - 1) = q^2 + q + 1$. ■

The projective plane of order q thus constructed will be denoted by the symbol $PG(2, q)$. These are the most important projective planes. However there are others. It is possible to tell just from the geometry of the plane whether it arises from this construction, and below in 3.6 we shall mention such a characterisation. Now let us look at affine planes.

2.4 Definition. An incidence structure $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ is called an *affine plane* if and only if it satisfies the following axioms:

(2.4.a) Any two distinct points are joined by exactly one line.

(2.4.b) Given any point p and any line G with $p \not I G$, there is precisely one line H with $p\, I\, H$ and not intersecting G.

(2.4.c) There is a *triangle*, i.e. three points not on a common line.

We say that two lines G and H are *parallel* if $G = H$ or $|(G, H)| = 0$, and we write $G \parallel H$. Thus (2.4.b) is Euclid's parallel axiom.

2.5 Proposition. *Let $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ be an affine plane. Then parallelism is an equivalence relation on $\boldsymbol{B}$. If $\boldsymbol{D}$ is finite, there exists a natural number n (called the* order *of $\boldsymbol{D}$) satisfying:*

(2.5.a) $|(p)| = n + 1$ *for all points p;*

(2.5.b) $|(G)| = n$ *for all lines G;*

(2.5.c) $|V| = n^2, \quad |\boldsymbol{B}| = n^2 + n.$

Proof. We only need to check the transitivity of $\parallel$. Thus assume $G \parallel H$ and $H \parallel K$. W.l.o.g. (without loss of generality) we may suppose that G, H and K are mutually distinct. If $G \nparallel K$ there would be a point $p I G, K$; but then G and K would be two parallels through p to H, contradicting (2.4.b). The remaining assertion follows as in the proof of Proposition 2.2 (or from Propositions 2.2 and 2.7 below). We leave the details to the reader. ■

We can provide examples of affine planes by showing that they are essentially the same as projective planes. To state the connection precisely, we need a definition.

2.6 Definition. Let $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ be an incidence structure and $Q \subseteq V$ and $\boldsymbol{C} \subseteq \boldsymbol{B}$. Then the incidence structure *induced* by $\boldsymbol{D}$ on Q and $\boldsymbol{C}$ is $\boldsymbol{D}' =$

$(Q, \boldsymbol{C}, I \mid Q \times \boldsymbol{C})$, and $\boldsymbol{D}'$ is called an *induced substructure* of $\boldsymbol{D}$. Instead of $I \mid Q \times \boldsymbol{C}$ we will usually again write I.

2.7 Proposition. *Let $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ be a projective plane and G a line of $\boldsymbol{D}$. Then the substructure $\boldsymbol{D}_G = (V \setminus (G), \boldsymbol{B} \setminus \{G\}, I)$ is an affine plane. Conversely, every affine plane may be obtained in this way from a projective plane. In the finite case, the orders of $\boldsymbol{D}$ and $\boldsymbol{D}_G$ are the same.*

Proof. Let $\boldsymbol{D}$ be a projective plane and remove a line U together with all its points. Then $\boldsymbol{D}_U$ satisfies (2.4.a), because $\boldsymbol{D}$ satisfies (2.1.a). Let p be a point of $\boldsymbol{D}_U$ and G a line of $\boldsymbol{D}_U$ with $p \nmid G$. If q is the point of intersection of G and U in $\boldsymbol{D}$, then the line pq is the desired parallel to G through p. Certainly, it is unique. Now, (2.4.c) follows from (2.1.c) and the existence of a point which is on neither of two given lines (cf. the proof of Proposition 2.2).

Conversely, given an affine plane, we can obtain a projective plane. Extend the point set by choosing an additional point corresponding to each parallel class of lines. Then choose one additional line ("the line at infinity"). The new incidence extends the old incidence as follows: The line at infinity contains each new point and a new point is contained in each of the lines in the corresponding parallel class. (This can be thought of as decreeing that each parallel class meets at some "point at infinity".) Clearly the given affine plane arises from the new structure by deleting the line at infinity. We leave the verification that the enlarged structure is indeed a projective plane to the reader. ■

We remark that this process of embedding an affine plane is in fact unique up to isomorphism, whereas non-isomorphic affine planes may in general be obtained from a given projective plane, depending on which line is deleted.

As an immediate consequence of Propositions 2.3 and 2.7 we have:

2.8 Corollary. *For any prime power q, there exists an affine plane of order q.* ■

These affine planes may, of course, be constructed directly, in a way similar to the proof of Proposition 2.3. One then uses a 2-dimensional vector space W over the field F and takes as points the vectors and as lines the cosets of 1-dimensional subspaces. The reader should verify this assertion and draw pictures of the affine planes of orders 2 and 3. We now generalise the common combinatorial features of finite affine and projective planes and arrive (by replacing "exactly one" in (2.1.a) and (2.4.a) by "exactly λ" where λ is an arbitrary natural number) at the following definition of a block design.

2.9 Definition. A finite incidence structure $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ is called a *block design* with parameters v, k, λ ($v, k, \lambda \in \mathbb{N}$) if it satisfies the following conditions:

(2.9.a) $|V| = v$;

(2.9.b) $|(p, q)| = \lambda$ for all $\{p, q\} \in \binom{V}{2}$, i.e. any two distinct points are joined by exactly λ blocks.

(2.9.c) $|(B)| = k$ for any block B.

For reasons which will be explained later we will often call $\boldsymbol{D}$ briefly an $S_\lambda(2, k; v)$ or in case of $\lambda = 1$ simply an $S(2, k; v)$. The letter S abbreviates "Steiner system", cf. Definition 3.1.

Thus in this terminology, projective planes of order n are examples of $S(2, n+1; n^2+n+1)$ and affine planes of $S(2, n; n^2)$. In fact, these are the only examples of block designs with these parameters. The reader should prove this as an exercise, see Exercise 2.21 below.

We have found examples of projective planes for all prime power values of n. It is not known whether there exist projective planes of any other order. We will prove a famous non-existence result due to Bruck and Ryser (1949) in Section II.4.

The reader might wonder why we have postulated only the constancy of the block degrees in Definition 2.9 and not also of the point degrees, which was also satisfied in our examples. The reason is that this property is a simple consequence of the next result.

2.10 Theorem. *Let $\boldsymbol{D}$ be an $S_\lambda(2, k; v)$. Then we have:*

(2.10.a) $|(p)| = \lambda(v-1)/(k-1) =: r$ *for all points p;*

(2.10.b) $|\boldsymbol{B}| = \lambda v(v-1)/k(k-1) =: b$.

Proof. Let p be a fixed point of $\boldsymbol{D}$. As in the proof of Proposition 2.2, we count all flags (q, G) with pIG and $q \neq p$ in two ways and obtain the equation $\lambda(v-1) = |(p)|(k-1)$ by (2.9.a), (2.9.b), and (2.9.c). This yields (2.10.a); then (2.10.b) follows from (1.7.a) using (2.10.a). ■

As r and b must be natural numbers, we have the following

2.11 Corollary. *Let $v, k, \lambda \in \mathbb{N}$. Necessary conditions for the existence of an $S_\lambda(2, k; v)$ are*

(2.11.a) $\lambda(v-1) \equiv 0 \pmod{k-1}$;

(2.11.b) $\lambda v(v-1) \equiv 0 \pmod{k(k-1)}$. ■

Much of this book is devoted to the problem of finding sufficient existence conditions for block designs, the basic problem in design theory. We will in fact prove in Chapter IX that for $k = 3$, 4, and 5 the conditions (2.11.a) and (2.11.b) are sufficient (Hanani), with one exception for $k = 5$. Also, by a deep result of Wilson (1975), they are asymptotically sufficient, i.e. for all sufficiently large v, given a fixed pair (k, λ). This will be proved in Chapter XI.

Given v and k, there is a trivial block design with these parameters, where one takes as blocks all k-subsets of a v-set; hence in this case $\lambda = \binom{v-2}{k-2}$. More generally, we shall call any "multiple" of such a design (that is, in the terminology to be introduced in §3, any t-design with $k = t$) a *trivial design*. For obvious reasons, these designs are sometimes called "complete designs"; for instance, this terminology is used by Cameron (1976a). However, this contradicts the original use of the term in the statistical literature on designs, where a *complete* design is a design with $k = v$; we shall use this term in the same sense. Note that a complete design is a special type of trivial design.

Affine and projective spaces provide further families of block designs, generalising the constructions of projective and affine planes given above.

2.12 Definition. Let F be a field and W an n-dimensional vector space over F. Then the set of all cosets of subspaces of W, ordered by inclusion, is called the *n-dimensional affine space* over F. If F is the field on q elements, it will be denoted by $AG(n, q)$. The cosets of $\{0\}$ are called *points*, those of 1-dimensional subspaces *lines*, those of 2-dimensional subspaces *planes*, those $(n-1)$-dimensional subspaces *hyperplanes*, and in general the cosets of i-dimensional subspaces are called *i-dimensional flats* or just *i-flats*. The points of $AG(n, q)$ together with the d-dimensional flats of $AG(n, q)$ as blocks and incidence by natural containment form an incidence structure denoted by $AG_d(n, q)$. By abuse of notation, we often write $AG(n, q)$ instead of $AG_1(n, q)$. A justification for this will be given in Section 8.

2.13 Proposition. *$AG_d(n, q)$ is a block design with parameters $v = q^n$, $k = q^d$, $r = \begin{bmatrix} n \\ d \end{bmatrix}_q$, $\lambda = \begin{bmatrix} n-1 \\ d-1 \end{bmatrix}_q$ and $b = q^{n-d}\begin{bmatrix} n \\ d \end{bmatrix}_q$. Here $\begin{bmatrix} n \\ i \end{bmatrix}_q$ denotes the number*

of i-dimensional subspaces of an n-dimensional vector space over $GF(q)$, the so-called Gaussian coefficients.

Proof. The values for v and k are trivial. Note that the number of d-flats containing a point x equals the number of d-dimensional linear subspaces containing $0 = x - x$, i.e. $\begin{bmatrix} n \\ d \end{bmatrix}_q$. Similarly, the number of d-flats containing two points x and y equals the number of d-dimensional linear subspaces containing 0 and $x - y$. But this is the same as the number of $(d-1)$-dimensional subspaces of W/U, where U is the subspace generated by $x - y$. Hence $\lambda = \begin{bmatrix} n-1 \\ d-1 \end{bmatrix}_q$. ■

For the convenience of the reader, we will compute an explicit formula for the Gaussian coefficients:

2.14 Lemma. *Let q be a prime power and n and d positive integers with $d \leq n$. Then*

$$\begin{bmatrix} n \\ d \end{bmatrix}_q = \frac{(q^n - 1)(q^{n-1} - 1) \cdots (q^{n-d+1} - 1)}{(q^d - 1)(q^{d-1} - 1) \cdots (q - 1)}. \tag{2.14.a}$$

Proof. Let W be an n-dimensional vector space over $GF(q)$. The number of ordered d-tuples of linearly independent elements of W is

$$(q^n - 1)(q^n - q) \cdots (q^n - q^{d-1}).$$

The first vector may be chosen as any of the $q^n - 1$ non-zero vectors. Then a 1-dimensional subspace of W is excluded, and the second vector may be chosen in $q^n - q$ ways. Continuing in this way, one arrives at the number given above. Now each d-tuple determines a d-dimensional subspace of W, and each such subspace is determined by $(q^d - 1)(q^d - q) \cdots (q^d - q^{d-1})$ d-tuples, since this is the number of d-tuples of linearly independent ele- ments in a d-dimensional vector space. Thus $\begin{bmatrix} n \\ d \end{bmatrix}_q$ is the quotient given in the assertion. ■

2.15 Definition. Let F be a field and W an $(n+1)$-dimensional vector space over F. Then the set of all subspaces of W, ordered by inclusion, is called the *n-dimensional projective space* over F. If F is the Galois field $GF(q)$, it will be denoted by $PG(n, q)$. The 1-dimensional (2-dimensional, 3-dimensional, n-dimensional) subspaces of W are called *points (lines, planes, hyperplanes)*; in general, the $(i+1)$-dimensional subspaces are called *i-flats*. The points of $PG(n, q)$ together with the d-dimensional flats of $PG(n, q)$ as blocks and incidence by natural containment form an incidence structure denoted by $PG_d(n, q)$.

By abuse of notation, we often write $PG(n, q)$ instead of $PG_1(n, q)$. A justification for this will be given in Section 8.

2.16 Proposition. *$PG_d(n, q)$ is a block design with parameters $v = \begin{bmatrix} n+1 \\ 1 \end{bmatrix}_q = (q^{n+1} - 1)/(q - 1)$, $k = \begin{bmatrix} d+1 \\ 1 \end{bmatrix}_q = (q^{d+1} - 1)/(q - 1)$, $r = \begin{bmatrix} n \\ d \end{bmatrix}_q$, $\lambda = \begin{bmatrix} n-1 \\ d-1 \end{bmatrix}_q$ and $b = \begin{bmatrix} n+1 \\ d+1 \end{bmatrix}_q$.*

The proof is similar to that of Proposition 2.13 and will be left to the reader.

For more about the Gaussian coefficients (and for the Gaussian polynomials), we refer the reader to Constantine (1987) . We have only given the most basic combinatorial properties of finite affine and projective geometries by showing that they yield designs. More material on the combinatorics of finite geometries will be provided later in §VIII.5 (subplanes, ovals, unitals, maximal arcs) and in §X.9 (partial spreads). For a detailed treatment of the combinatorial aspects of finite geometries, we refer the reader to Batten (1986), Beutelspacher (1982a, 1983) and Hirschfeld (1985, 1998), Hirschfeld and Thas (1991) .

We will conclude this section with some more definitions and notation. As will be quite clear from what follows, sometimes the postulates for block designs are too restrictive. It is in fact useful to consider conditions (2.9.b) and (2.9.c) separately. This yields the following notions:

2.17 Definition. Let λ be a natural number and $K \subseteq \mathbb{N}$. A finite incidence structure $\boldsymbol{D} = (P, \boldsymbol{B}, I)$ is called a *pairwise balanced design* (*PBD*) with block sizes from K iff it satisfies

(2.17.a) $|(p, q)| = \lambda$ for any two points p, q,

(2.17.b) $|(B)| \in K$ for any block B.

If $\boldsymbol{D}$ has v points, it is called an $S_\lambda(2, K; v)$. For $\lambda = 1$, we also use the simpler notation $S(2, K; v)$ and the term *linear space* (instead of *PBD*). Note that not all numbers in K actually need occur as block sizes. In particular, K may be infinite. If K is a singleton $\{k\}$, we arrive at the notion of a block design again.

2.18 Definition. Let k be a natural number and $\boldsymbol{D} = (P, \boldsymbol{B}, I)$ a finite incidence structure. $\boldsymbol{D}$ is called a *k-hypergraph* iff it satisfies

(2.18.a) $|(B)| = k$ for any block B.

If $k = 2$, we speak of a *graph*. Then the points are called *vertices* and the blocks are called *edges*. (This is often called a "multigraph" in the literature; then "graph" means "simple graph" in our terminology.)

2.19 Notation. Let λ be a natural number and $K \subseteq \mathbb{N}$. Then the set of all $v \in \mathbb{N}$ for which an $S_\lambda(2, K; v)$ exists is denoted by $S_\lambda(2, K)$ or alternatively $B(K, \lambda)$ (which is Hanani's notation). In case $\lambda = 1$, we use the simpler notation $B(K)$. If K is a singleton $\{k\}$, we just write $S_\lambda(2, k)$ or $B(k)$.

Thus we may consider B as an operator on $2^{\mathbb{N}}$, mapping $K \subseteq \mathbb{N}$ onto $B(K)$. This idea will play a fundamental role in the existence theory for block designs and *PBD*'s in later chapters. In fact, B will turn out to be what is known as a closure operator. For now we will only translate the results of this section into our new notation and collect the following facts.

(2.19.a) $q^n \in B(q)$ for all prime powers q;

(2.19.b) $q^n + q^{n-1} + \ldots + q + 1 \in B(q+1)$ for all prime powers q;

(2.19.c) $q^n \in B\left(q^d, \begin{bmatrix} n-1 \\ d-1 \end{bmatrix}_q\right)$ for all prime powers q and all $n > d$;

(2.19.d) $q^n + q^{n-1} + \ldots + q + 1 \in B\left(q^d + \ldots + q + 1, \begin{bmatrix} n-1 \\ d-1 \end{bmatrix}_q\right)$ for all prime powers q and all $n > d$.

In particular, we have

(2.19.e) $7, 9, 15, 27, \ldots, 2^n - 1, 3^n \in B(3)$ for all n;

(2.19.f) $13, 16, 40, 64 \in B(4)$;

(2.19.g) $21, 25, 85, 125 \in B(5)$.

2.20 Exercise. Show that $r \geq k$, and hence $b \geq v$, in any $S(2, k; v)$. This will be generalised in Sections II.2 and II.6.

2.21 Exercises. (a) Prove that any $S(2, n+1; n^2 + n + 1)$ with $n \geq 2$ is a projective plane of order n. Hint: Compute r and assume the existence of two non-intersecting blocks to obtain a contradiction.

(b) Prove that any $S(2, n; n^2)$ with $n \geq 2$ is an affine plane of order n.

§3. *t*-Designs, Steiner Systems and Configurations

Let us reconsider one of our examples a bit more closely. In Proposition 2.13, we observed that $AG_2(n, 2)$ (i.e. the incidence structure formed by points and planes in the n-dimensional affine space over $GF(2)$) is an $S_\lambda(2, 4; 2^n)$ with $\lambda = 2^{n-1} - 1$. In particular, any two points are on $2^{n-1} - 1$ planes. But there is an even more remarkable uniformity: any three points are on a unique common plane. Here we use the property that no three points are on a common line in a vector space over $GF(2)$. This observation leads us to the following generalisation of our previous concepts:

3.1 Definition. Let t and λ be positive integers and $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ a finite incidence structure. Then $\boldsymbol{D}$ is called *t-balanced* with parameter λ iff

(3.1.a) $\quad |(Q)| = \lambda \quad$ for any t-subset $Q \subseteq V$.

If $|V| = v$, $|(B)| \in K$ for each $B \in \boldsymbol{B}$, then $\boldsymbol{D}$ is called an $S_\lambda(t, K; v)$. If $\boldsymbol{D}$ is also k-hypergraph (Definition 2.18), then it is called a *t-design* with parameters k and λ. A t-design on v points is called an $S_\lambda(t, k; v)$. In case $\lambda = 1$, it is called a *Steiner system* $S(t, k; v)$. This is the reason for the notation $S_\lambda(t, k; v)$.

Thus affine and projective planes are Steiner systems. The *PBD*'s are 2-balanced incidence structures and block designs are 2-designs. The designs $AG_2(n, 2)$ are simultaneously 1-, 2- and 3-designs. This is accounted for by the following result.

3.2 Theorem. *Let $\boldsymbol{D}$ be a t-design and let $s < t$ be a positive integer. Then $\boldsymbol{D}$ is also an s-design. More specifically, if $\boldsymbol{D}$ has parameters v, k and λ_t (where λ_t is the number of blocks through a t-set), then the parameter λ_s (the number of blocks through an s-set) is given by*

$$\text{(3.2.a)} \qquad \lambda_s = \lambda_t \binom{v-s}{t-s} \Big/ \binom{k-s}{t-s}.$$

Proof. Let S be an s-set of points. We count all pairs (X, B), where X is a $(t - s)$-set of points with $S \cap X = \emptyset$ and B is a block with $S \cup X \subseteq (B)$ in two ways. By (3.1.a) we obtain $\lambda_t \binom{v-s}{t-s}$ such pairs, as there are $\binom{v-s}{t-s}$ ways of choosing X. On the other hand, for given B with $S \subseteq (B)$ there are $\binom{k-s}{t-s}$ ways of choosing X; hence we also obtain $|(S)|\binom{k-s}{t-s}$ such pairs. This yields the desired formula with $\lambda_s = |(S)|$. ■

An interesting improvement of this result will be given in Theorem II.6.1. As in Section 2, we introduce some notation.

3.3 Notation. Let t and λ be positive integers and $K \subset \mathbb{N}$. Then $S_\lambda(t, K)$, or simply $S(t, K)$ when $\lambda = 1$, denotes the set of all $v \in \mathbb{N}$ for which a t-balanced incidence structure with v points, parameter λ, and block sizes from K exists. For $K = \{k\}$, we simply write $S_\lambda(t, k)$ or $S(t, k)$. For $t = 2$, we will retain the earlier notation 2.19. Hanani's notation is not used for $t \neq 2$. There are also some special names for particular types of Steiner systems:

3.4 Definition. An $S(2, 3; v)$ is called a *Steiner triple system STS*(v). An $S(3, 4; v)$ is called a *Steiner quadruple system SQS*(v).

Our example above has shown that $AG_2(n, 2)$ is in fact a Steiner quadruple system. Thus we have

(3.4.a) $\quad 2^n \in S(3, 4) \quad$ for all $n > 1$.

We will meet more examples of t-designs with $t \geq 3$ later. There are in fact other "classical" examples (such as the Möbius planes), but their construction needs a knowledge of automorphisms and will be postponed to Chapter III. We now look at 1-designs.

3.5 Definition. A 1-design $S_r(1, k; v)$ is called a *tactical configuration* (or simply *configuration*) with parameters v, r, k, and $b := vr/k$.

In terms of the incidence matrix, a configuration is characterised by constant row sums r and constant column sums k. A special case of Theorem 3.2 is the fact that every 2-design is a configuration (already proved in Theorem 2.10). We mention some more examples: any graph of constant degree r is an $S_r(1, 2; v)$ (such graphs are usually called *r-regular*). Next, we give two examples which are of fundamental importance for classical projective geometry.

3.6 Example. The configuration with parameters $v = b = 10$, $r = k = 3$ given by the incidence matrix below and drawn in Figure 3.1 is called the *Desargues configuration*. Its geometric importance is a follows. A projective plane is isomorphic to one constructed from a field as in the proof of Proposition 2.3 if and only if for any two triangles $\{p_5, p_6, p_7\}$ and $\{p_8, p_9, p_{10}\}$, which are "in perspective" from a point p_1 (as drawn), the intersection points p_2, p_3, p_4 of corresponding sides are on a common line G_1, see e.g. Hughes and Piper (1982).

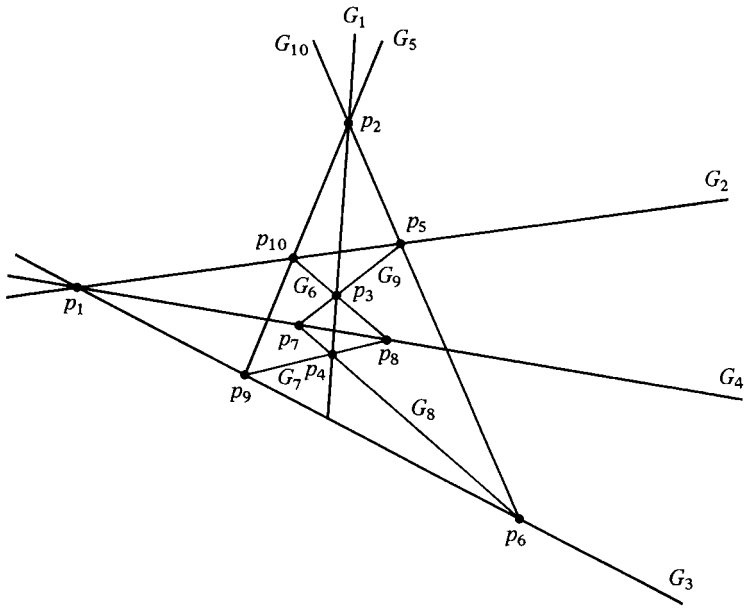

Figure 3.1

The corresponding incidence matrix is

(3.6.a)
$$\left(\begin{array}{ccccccccccc}
0&1&1&1&0&0&0&0&0&0\\
1&0&0&0&1&0&0&0&0&1\\
1&0&0&0&0&1&0&0&1&0\\
1&0&0&0&0&0&1&1&0&0\\
&&&&&&&&&\\
0&1&0&0&0&0&0&0&1&1\\
0&0&1&0&0&0&0&1&0&1\\
0&0&0&1&0&0&0&1&1&0\\
&&&&&&&&&\\
0&0&0&1&0&1&1&0&0&0\\
0&0&1&0&1&0&1&0&0&0\\
0&1&0&0&1&1&0&0&0&0
\end{array}\right).$$

In view of the importance of this characterisation, planes arising from fields are called *Desarguesian* and all others are called *non-Desarguesian*. In defining projective planes in Proposition 2.3, we did not require that the field is commutative. Interestingly, whether or not the field is actually commutative can be discerned from another configuration.

3.7 Example. The *Pappos configuration*, with parameters $v = b = 9$, $r = k = 3$, is given by the incidence matrix

$$\text{(3.7.a)} \qquad \left(\begin{array}{ccc|ccc|ccc} 1&1&1&0&0&0&0&0&0\\ 0&0&0&1&1&1&0&0&0\\ 0&0&0&0&0&0&1&1&1\\ \hline 1&0&0&1&0&0&1&0&0\\ 0&1&0&0&1&0&0&1&0\\ 0&0&1&0&0&1&0&0&1\\ \hline 1&0&0&0&0&1&0&1&0\\ 0&1&0&1&0&0&0&0&1\\ 0&0&1&0&1&0&1&0&0 \end{array}\right)$$

and looks as follows (check this!):

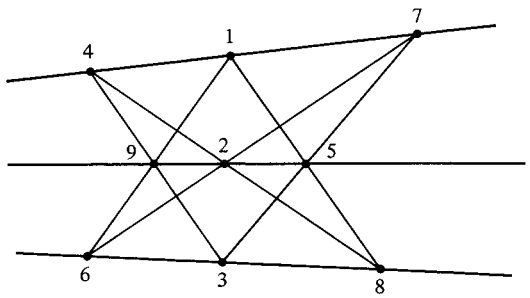

Figure 3.2

Its geometric meaning is the following. A projective plane is isomorphic to one constructed from a commutative field as in the proof of Proposition 2.3 if and only if for any six distinct points $p_1, \ldots, p_6$ (with p_1, p_3, p_5 and p_2, p_4, p_6 resp. collinear) the intersection points s_i of $p_i p_{i+1}$ and $p_{i+3} p_{i+4}$ ($i = 1, 2, 3$; indices modulo 6) are collinear, see e.g. Hughes and Piper (1982).

Planes satisfying this condition are called *Pappian*. By the remark about the Desargues configuration, the validity of the Pappos configuration must imply that of the Desargues configuration. Also, it is well known that all finite fields are commutative; hence every finite Desarguesian plane is actually Pappian. Some finite non-Desarguesian projective planes will be constructed in Corollary X.9.20.

We conclude this section with some interesting digressions. We have seen that it is possible to represent the Pappos and Desargues configurations in the real plane using straight lines. However it is impossible to do this for linear spaces except for trivial spaces. Specifically:

3.8 Proposition (Sylvester's Problem). *An $S(2, K; n)$ with at least three points on every line cannot be represented in the real space (using straight lines only) unless all points lie on one line.*

Proof. Assume the contrary and choose a line L and a point $p \notin L$ such that the euclidean distance from p to L is minimal, say d. As L has at least three points, there is a triangle $\{p, a, b\}$ with $a, b \in L$ such that the angle at a is at least $\pi/2$. But then the distance of a from pb is smaller than d, a contradiction. ■

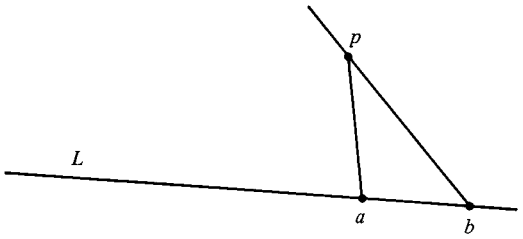

Figure 3.3

Proposition 3.8 explains why the "circle" in Figure 1.1 is "necessary" when representing Example 1.3. In connection with Proposition 3.8, we refer the reader to Coxeter (1961).

3.9 Remark. Gropp (1992) gives a report on the birth of design theory in British India. Some comments on the early history of what is now (unfortunately) called Steiner systems can be found in the paper by de Vries (1984). An $S(2, 3; 9)$ was already known to Plücker (1835), realised on the nine points of inflection of a cubic curve without singular points in the complex plane. It should be noted that Plücker's purported construction of an $S(3, 4; 28)$ on the double tangents of a curve of degree 4 (which is mentioned by de Vries) is incorrect, as pointed out by Felix Klein. Special cases of the existence problem for Steiner systems were posed and studied by Woolhouse (1844) and Kirkman (1847, 1850a), well before the note by Steiner (1853). In particular, Kirkman (1847) already settled the existence problem for Steiner triple systems. An interesting paper on Kirkman's mathematical work is due to Biggs (1981).

We finally give some general comments on the existence problem for t-designs $S_\lambda(t, k; v)$. Theorem 3.2 gives some necessary conditions, since the parameters λ_s are obviously integers. However, these conditions are in general not sufficient, as we shall see in Chapter II. One of the main research areas in design theory lies in finding necessary and sufficient conditions for the existence of an $S_\lambda(t, k; v)$, given the parameters t, k and λ. Of course, t-designs with $k = t$ or $k = v$ always exist; we have called such examples *trivial*. In spite of much effort, no non-trivial Steiner system $S(t, k; v)$ with $t \geq 6$ has yet been found; the analogous statement also applies for t-wise balanced designs $S(t, K; v)$ with $\min K > t$. The situation changes drastically if blocks of size t are allowed; then there is an enormous number of such designs, as the following result of Colbourn, Hoffman et al. (1991) shows. We here anticipate the formal definition of an isomorphism which will be given in the next section. Also, we use the Landau symbol $o(n)$ to denote a quantity x depending on a natural number n such that the ratio x/n tends to 0 as n tends to infinity.

3.10 Theorem. *The number $N_t(v)$ of non-isomorphic t-wise balanced designs on a v-set (allowing blocks of size t) with $\lambda = 1$ is given by the asymptotic formula*

$$N_t(v) = v^{[\binom{v}{t}/(t+1)](1+o(1))}. \quad \blacksquare \tag{3.10.a}$$

§4. Isomorphisms, Duality and Correlations

After having introduced some of the main objects of this book in the previous few sections, we now consider the notion of an isomorphism. The formal definition is the following one:

4.1 Definition. Let $\boldsymbol{D} = (V, \boldsymbol{B}, I)$ and $\boldsymbol{D}' = (V', \boldsymbol{B}', I')$ be incidence structures and let $\pi : V \cup \boldsymbol{B} \to V' \cup \boldsymbol{B}'$ be a bijection. π is called an *isomorphism* iff it satisfies:

$$V^\pi = V' \quad \text{and} \quad \boldsymbol{B}^\pi = \boldsymbol{B}'; \tag{4.1.a}$$

$$p\,I\,B \iff p^\pi\,I'\,B^\pi \quad \text{for all } p \in V \text{ and all } B \in \boldsymbol{B}. \tag{4.1.b}$$

In this case, $\boldsymbol{D}$ and $\boldsymbol{D}'$ are said to be *isomorphic*. If $\boldsymbol{D} = \boldsymbol{D}'$, then π is called an *automorphism* (or, if $\boldsymbol{D}$ is 2-balanced with $\lambda = 1$, a *collineation*, since it preserves lines).

4.2 Observation. In terms of their incidence matrices M and M', the structures $\boldsymbol{D}$ and $\boldsymbol{D}'$ are isomorphic if and only if there exist row and column permutations

transforming M into M', i.e. if and only if there are permutation matrices P, Q satisfying

(4.2.a) $\qquad PMQ = M'.$

Clearly the set of all automorphisms of a given incidence structure $\boldsymbol{D}$ forms a group. ■

We introduce the following

4.3 Convention. Let $\boldsymbol{D}$ be an incidence structure. Then the group of all automorphisms of $\boldsymbol{D}$ is called the *full automorphism group* of $\boldsymbol{D}$ and denoted by Aut $\boldsymbol{D}$. Any subgroup of Aut $\boldsymbol{D}$ will be called *an automorphism group* of $\boldsymbol{D}$. The terms "full collineation group" and "collineation group" are defined similarly.

As an example, let us consider the automorphisms of the projective plane of order 2 (see Examples 1.3, 1.5 and Figure 1.1).

4.4 Proposition. *Let $\boldsymbol{D}$ be the projective plane of order 2 (Example* 1.3*). Then* $|\text{Aut } \boldsymbol{D}| = 168$ *and* Aut $\boldsymbol{D}$ *is regular*[2] *on ordered triangles.*

Proof. We will sketch the proof and leave the details to the reader. Let (a_1, a_2, a_3) and (b_1, b_2, b_3) be ordered triangles and assume $a_i^\alpha = b_i$ $(i = 1, 2, 3)$ for $\alpha \in \text{Aut } \boldsymbol{D}$. Let a_4 be the third point on a_1a_2 and b_4 the third point on b_1b_2. If α is to be an automorphism, we have to have $a_4^\alpha = b_4$. Arguing similarly for a_1a_3, a_2a_3 and a_3a_4, one sees how α has to be defined. One then has to check that this definition indeed yields an automorphism of $\boldsymbol{D}$. To determine the order of Aut $\boldsymbol{D}$, we thus only have to count the number of ordered triangles (a, b, c) of $\boldsymbol{D}$. This is easily seen to be 168, as there are seven possibilities for choosing a, then six possibilities for choosing b, and four possibilities for choosing c (the third point on ab is not allowed!). ■

By results from elementary group theory (Sylow theorems) we conclude the existence of elements of order 3 and 7 and of a subgroup of order 8 of Aut $\boldsymbol{D}$. It may be useful to see them explicitly. In the equilateral triangle representation of Figure 1.1 an automorphism of order 3 is given by the rotation about the

[2] This means that for any two ordered triangles (a_1, a_2, a_3), (b_1, b_2, b_3) there is precisely one element $\alpha \in \text{Aut } \boldsymbol{D}$ with $a_i^\alpha = b_i$ for $i = 1, 2, 3$. More details on permutation groups will be given in Chapter III.