ELEMENTS OF THE TOPOLOGY OF PLANE SETS OF POINTS

by

M. H. A. NEWMAN, M.A., F.R.S.
Fellow of St John's College

ELEMENTS OF THE TOPOLOGY OF PLANE SETS OF POINTS

PREFACE

The properties of plane sets of points which form the principal subject-matter of this book are those that depend essentially on the fact that a plane can be cut up into polygonal pieces, and therefore belong to what is usually called the combinatorial theory of sets of points. The 2-dimensional part of the theory seems particularly well suited to form a first introduction to Topology, since it shews clearly the advantages of an algebraic treatment of complexes (modulo 2 addition) without involving any but the simplest theorems of linear algebra.

I hope the book may also be useful to analysts, by making accessible the simple proof of Jordan's Theorem, and other separation theorems, that are based on the "Alexander Lemma".

Since the book is meant to be an account of methods, rather than a comprehensive collection of results, special topics, such as the theory of prime ends, have been omitted. It was also not possible to give any account of the remarkable methods by which S. Eilenberg has recently proved many of the results of Chapters v–viii, in some cases extending them to sets which are not open or closed (see Note 24).

The combinatorial theory occupies Part II. In Part I the necessary general theorems on closed and open sets, compactness, connected sets, etc. are established for sets in any metric space, since the proofs are no more difficult than for plane sets. Readers who are acquainted with the simplest properties of closed and open sets may find it convenient to turn straight to Part II, using Part I as the need arises. (A list of symbols used precedes the Index.)

I am indebted for many suggestions throughout the book to Dr J. H. C. Whitehead, who read both the manuscript and the proofs, and to Dr N. E. Steenrod, of Princeton University, and Professor J. H. Roberts, of Duke University, who also read the work in manuscript. I wish also to thank the University Printer and his staff for the great trouble they have taken to ensure the best possible arrangement of each page.

M. H. A. N.

November

CONTENTS

PART II. TOPOLOGY OF PLANE SETS OF POINTS

PART I

GENERAL PROPERTIES OF SETS OF POINTS

Chapter I

SETS*

§1. THE CALCULUS OF SETS

1. The object of the calculus described in § 1 is a practical one—to shew how complicated properties of sets may be deduced by formal rules from a small number of properties which are sufficiently simple to be self-evident. The propositions accepted as self-evident (lettered A to F) are not intended as a system of axioms—a much smaller number would suffice for that purpose—but as a convenient body of "standard forms" for use in symbolical work.

2. A set (or class or aggregate) is to be thought of not as a heap of things specified by enumerating its members one after another, but as something determined by a *property*, which can be used to test the claim of any object to be a member of the set. Thus the set of even integers is determined by the property of being twice some other integer, the algebraic numbers by the property of satisfying a polynomial equation with integral coefficients. Two properties determine the same set if they are "formally equivalent", i.e. if no object has one property without having the other.

The symbol $x \epsilon A$ means "x is a member (or element) of A". The set which has only the single member a is denoted by (a). Thus $x \epsilon (a)$ means simply $x = a$.

3. The symbol $A \subseteq B$ (or $B \supseteq A$) means that, for every x,

$$x \epsilon A \quad \text{implies} \quad x \epsilon B.$$

It is usually read "A is contained in B", or "A is a subset of B", but, as the form of the symbol suggests, identity is not excluded.

* The subjects dealt with in this preliminary chapter are (1) the calculus or algebra of sets, and (2) the distinction between enumerable sets and others. Readers who are familiar with these matters should omit this chapter, but should note the definitions here adopted for the symbols $A \subseteq B$ and $A \subset B$, pp. 3, 4; $B - A$, p. 7; and $\mathscr{C}A$, p. 7.

The symbol $A \subset B$ (rarely used in this book) means "$A \subseteq B$ but $A \neq B$", and is read "A is a proper subset of B".[1]*

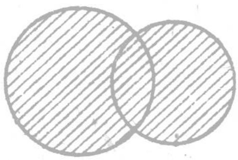
A 1. *$A \subseteq A$.*

2. *If $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.*

3. *$A = B$ if, and only if, $A \subseteq B$ and $B \subseteq A$.*

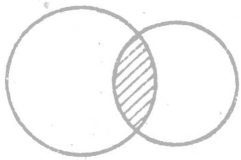
These properties of sets must be accepted as self-evident. (A 3 may, if preferred, be regarded as a definition of equality between sets.)

4. The *sum*, $A+B$, of the sets A and B is the set of all members of either set. Thus "$x \in (A+B)$" means "$x \in A$ or $x \in B$ (or both)"



$A+B$

Fig. 1



AB

Fig. 2

For example, if A is the set of real numbers between 0 and 2, and B the set between 1 and 3, $A+B$ is the set between 0 and 3.

The *common part*,† or intersection, AB, of the sets A and B is the set of things belonging to both A and B, i.e. "$x \in AB$" means "$x \in A$ and $x \in B$". A dot is sometimes inserted, thus: $A \cdot B$. The usual convention about bracketing is adopted, i.e. $AB + C$ means $(AB)+C$.

The operations $+$ and $\cdot$ have all the formal properties of ordinary addition and multiplication.

B 1. $A+(B+C) = (A+B)+C, \quad (AB)C = A(BC)$,

2. $A+B = B+A, \quad AB = BA$,

3. $A(B+C) = AB+AC$;

* The numbers (1), (2), etc. refer to the notes at the back of the book.

† There are so many different symbolic "products" in the theory of sets and topology that the word is best avoided in connection with the common part.

and they have the further properties

B 4. $A+A=A$, $AA=A$,

5. $A \subseteq A+B$, $AB \subseteq A$,

6·1. If $A \subseteq C$ and $B \subseteq C$ then $A+B \subseteq C$,

6·2. If $A \supseteq C$ and $B \supseteq C$ then $AB \supseteq C$.

All these formulae may be accepted as self-evident as they stand, or referred back to propositions of logic. For example, if the "definitions" of $A \subseteq B$ and of $x \epsilon A+B$ are inserted in **B** 6·1, it becomes "if $x \epsilon A$ implies $x \epsilon C$, and $x \epsilon B$ implies $x \epsilon C$, then ($x \epsilon A$ or $x \epsilon B$) implies $x \epsilon C$".

If A is a finite set, with members $a, b, \ldots, k$, then

$$A=(a)+(b)+\ldots+(k),$$

which will be abbreviated to $(a, b, \ldots, k)$.

Examples. 1. $(A+B)\cdot(A+C)=A+BC$.

For

$$(A+B)(A+C)=A+AB+AC+BC, \quad \text{by } \mathbf{B}\,1\text{–}4.$$

Since $AB \subseteq A$, $A+AB=A$ by **B** 6·1, and hence

$$(A+AB)+AC=A+AC=A,$$

giving the result. (This is the "dual distributive law", obtained from **B** 3 by interchanging + and ·.)

2. A necessary and sufficient condition that $A \subseteq B$ is that $A+B=B$.

If $A+B=B$, $A \subseteq B$ by **B** 5. If $A \subseteq B$, $A+B \subseteq B$ by **A** 1 and **B** 6·1, and $B \subseteq A+B$ by **B** 5

Exercises. 1. Prove **B** 4 formally from **A** and the rest of **B** 1–6.

2. A necessary and sufficient condition that $A \subseteq B$ is that $AB=A$.

5. The *null-set*, denoted by 0, has no members and is contained in every set:

C. $0 \subseteq A$: *the null-set is a subset of every set.*

When sets are regarded as collections or heaps of things a set with no members is a rather shadowy or even paradoxical entity, but its mysterious quality disappears if statements about sets are interpreted as statements about properties. Let p be called a *null-property* if it is not possessed by any object. Examples are: being greater than 3 and less than 2, or being a zero of e^z. Such properties are frequently considered in mathematics, particularly in proofs by *reductio ad absurdum*. Any two null-properties are "formally equivalent", in the sense of par. 2, and therefore all these properties determine the same set, which is called the null-set.

To arrive from this definition at proposition C it is necessary to consider more closely the interpretation of the symbol $A \subseteq B$. The meaning assigned to it was: for every x, $x \epsilon A$ implies $x \epsilon B$. This means that $x \epsilon B$ unless "$x \epsilon A$" is false, i.e. "$x \epsilon B$" *is true, or* "$x \epsilon A$" *is false*. This final form may be taken as the basic meaning of $A \subseteq B$, and from it it is clear that $0 \subseteq A$. For since, for all x, "$x \epsilon 0$" *is* false, the proposition

$$\text{"}x \epsilon A\text{" is true or "}x \epsilon 0\text{" is false}$$

is true, whatever the set A may be.

From C and B 6·1 and 6·2 it follows that

$$A + 0 = A, \quad A0 = 0.$$

Thus the formal properties of the null-set justify the symbol 0.

Two sets are said to *meet* (or intersect) if they have at least one common member. It follows from the definition of the null-set that the necessary and sufficient condition that A and B meet is that $AB \neq 0$.

Note. In work that is not purely symbolical the symbol $A \subseteq B$ is often replaced by the words "all a's are b's"—for example, "the set of all parabolas is contained in the set of all conics" by "all parabolas are conics". It must, however, be borne in mind that if A is the null-set "all a's are b's" is to be regarded as true whatever B may be. Thus all zeros of e^z are real and positive, because e^z has no zeros. All zeros of e^z are also real and negative, and there is no contradiction between the statements, because it is not asserted that any actual number is both positive and negative, but only that the set of zeros of e^z (i.e. the null-set) is a subset of both the other sets of numbers.

6. *Subtraction.* The difference, $B-A$, between any two sets, B and A, is the set of elements of B not contained in A; i.e.

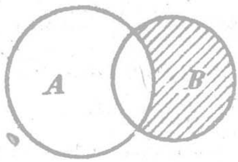
"$x \in (B-A)$" means "$x \in B$, but not $(x \in A)$".

Evidently $A-A=0$, and $A-0=A$.

The agreeable similarity so far observable between the algebra of sets and ordinary algebra breaks down with the introduction of subtraction, which is not even associative; for

$$A+(A-A)=A+0=A,$$

but $$(A+A)-A=A-A=0.$$



$B-A$

Fig. 3

This is due to the fact that $B-A$ is not the solution of the equation $A+X=B$, which may have an infinity of solutions (e.g. if $B=A$), or none (if $B=0$). It is possible nevertheless to maintain a workable calculus by operating with *complements* with respect to a fixed set S.

If $A \subseteq S$ the set $S-A$ is called the *complement*, or residual set, of A with respect to S. If S is supposed fixed, $S-A$ may be denoted by $\mathscr{C}A$.

Besides the obvious properties

D 1. $\mathscr{C}S=0, \quad \mathscr{C}0=S,$

2. $A+\mathscr{C}A=S, \quad A\cdot\mathscr{C}A=0,$

3. $\mathscr{C}(\mathscr{C}A)=A,$

4. *If* $A \subseteq B$ *then* $\mathscr{C}B \subseteq \mathscr{C}A$,

the complement has the important property of interchanging $+$ and $\cdot$:

D 5. $\mathscr{C}(A+B)=\mathscr{C}A\cdot\mathscr{C}B, \quad \mathscr{C}(AB)=\mathscr{C}A+\mathscr{C}B.$

(See Figs. 1 and 2, where S may be taken to be the whole plane.) This proposition corresponds to the theorem in logic that "not (p or q)" is equivalent to "not p and not q", and "not (p and q)" to "not p or not q".

* Theorem 6·1. *If* $A+X=S$, *and* $AX=0$, *then* $X=\mathscr{C}A$.

By the first of the given equations, $X \subseteq S$. Multiplying the first equation of D 2 by X and using $AX=0$, we obtain $X\cdot\mathscr{C}A=X$.

* I.e. Theorem 1 of par. 6 (of Ch. I), referred to in this chapter as 6·1, in others as I. 6·1.

Multiplying the first of the given equations by $\mathscr{C}A$ and using $A\cdot\mathscr{C}A=0$, we get $X\cdot\mathscr{C}A=\mathscr{C}A$. Hence $X=\mathscr{C}A$.

7. We now return to the difference, $B-A$, between any two sets. If S is any set containing A and B, the property "x belongs to B but not to A", defining $B-A$, is evidently equivalent to "x belongs to B and S, but not to A", i.e. to

$$(x\,\epsilon\,B) \quad \text{and} \quad (x\,\epsilon\,S \text{ but not } x\,\epsilon\,A).$$

The first bracket is the determining property of B, the second that of $\mathscr{C}A$. Hence

D 6. *If complements are taken with respect to any set containing both A and B, $B-A=B\cdot\mathscr{C}A$.*

By means of **D 6** all differences occurring in any formula can be expressed in terms of complements with respect to a fixed set, S, containing all the sets involved, and the formulae **D 1–5** applied. This is the method recommended for proving formal identities.

Examples. (All complements are formed with respect to an arbitrary set, S, containing A, B and C, except in Example 4.)

1. $A(B-C)=B(A-C)=AB-C$, for all three sets are $AB\cdot\mathscr{C}C$.

2. $B-A=B-AB$.

$$\begin{aligned} B-AB &= B\cdot\mathscr{C}(AB)\\ &= B\cdot\mathscr{C}A+B\cdot\mathscr{C}B\\ &= B\cdot\mathscr{C}A=B-A. \end{aligned}$$

3. A necessary and sufficient condition that $AB=0$ is that $A\subseteq\mathscr{C}B$.

If it is given that $A\subseteq\mathscr{C}B$, multiply both sides by B. If it is given that $AB=0$, multiply both sides of $A\subseteq B+\mathscr{C}B$ by A.

4. If $C\subseteq B\subseteq A$ then $(A-B)+(B-C)=A-C$. Take complements with respect to A. The left-hand side is $\mathscr{C}B+B\mathscr{C}C$. Since $\mathscr{C}B\subseteq\mathscr{C}C$, this is

$$\mathscr{C}B\cdot\mathscr{C}C+B\mathscr{C}C=\mathscr{C}C=A-C.$$

5. *The propositions* **D 1** *and* **D 3–5** *can be deduced formally from* **A–C** *and* **D 2**.

The theorem 6·1 is first proved, as in the text. It uses only **A–C** and **D 2**.

Proof of **D 1**. Since $S+0=S$ and $S0=0$, it follows from 6·1 that $\mathscr{C}S=0$ and $\mathscr{C}0=S$.

Proof of **D** 3. By **D** 2 the equations

$$\mathscr{C}A+X=S,\quad \mathscr{C}A\cdot X=0$$

are satisfied by $X=A$. Therefore by 6·1, $A=\mathscr{C}(\mathscr{C}A)$.

Proof of **D** 4. Since $A\subseteq B$, $S=A+\mathscr{C}A\subseteq B+\mathscr{C}A$. Multiply both sides by $\mathscr{C}B$:

$$\mathscr{C}B\subseteq \mathscr{C}B\cdot\mathscr{C}A\subseteq\mathscr{C}A.$$

Proof of **D** 5. Let $A'=\mathscr{C}A$ and $B'=\mathscr{C}B$. Then

$$(A+B)\cdot(A'B')=AA'\cdot B'+BB'\cdot A'=0,$$

$$(A+B)+(A'B')\supseteq(AB'+A'B+AB)+A'B',$$

since all the sets in the bracket on the right are contained in A or B,

$$=(A+A')\cdot(B+B')=S.$$

Since A, B and $A'B'$ are all contained in S, it follows that

$$(A+B)+A'B'=S.$$

Hence by 6·1 $A+B$ and $A'B'$ are complementary sets. This gives the first part of **D** 5 immediately, and the second part on interchanging dashed and undashed letters.

Exercises. 1. $A(B-C)=AB-AC$.

2. $(A-B)+(A-C)=A-BC$.

3. $(A-B)(A-C)=A-(B+C)$.

4. $(A-C)+(B-C)=(A+B)-C$.

5. $(A-B)+(B-A)=(A+B)-AB$.

6. $A-(A-B)=AB$.

7. $\mathscr{C}(A_1+A_2+\ldots+A_k)=\mathscr{C}A_1\cdot\mathscr{C}A_2\cdot\ldots\mathscr{C}A_k$.

8. $A+(B-A)=A+B$, $A(B-A)=0$. Prove that the whole of **D** 1–6 can be deduced formally from these two relations together with **A**–**C** and the definition "$\mathscr{C}A=S-A$ if $A\subseteq S$". [First prove that the equations $A+X=A+B$, $AX=0$ have at most one solution. Cf. 6·1 and Example 5 above.]

8. *Duality.* The calculus that has so far been developed (sometimes called the Algebra of Sets, or Boolean Algebra) has a duality property which has probably already been observed by the reader. If in any theorem of the Algebra all differences are expressed in terms of complements with respect to a fixed set S, and then the symbols

$$\left.\begin{array}{l}+\text{ and }\cdot\\ 0\text{ and }S\\ \subseteq\text{ and }\supseteq\end{array}\right\}\text{ are everywhere interchanged,}$$

the result is also a true theorem of the Algebra.

Since no appeal is made to the duality property in this book a general proof (which would require a more exact delimitation of the "Algebra of Sets") is not given. (Cf. note 2.)

9. It is frequently necessary to consider sets whose members are themselves sets of things. If **M** is such a set of sets, the members of **M** (i.e. the sets of "things") are usually denoted by a suffix notation, A_x. The suffix x may range through any set B, e.g. the integers from 1 to k, all the positive integers, all the real numbers, etc. When this notation is used for the members of **M** the set **M** itself is denoted by $\{A_x\}$.

The *sum-set*

$$\sum_{x \epsilon B} A_x$$

is the set of all members of the sets A_x; i.e.

$$\text{“}z \epsilon \sum_{x \epsilon B} A_x\text{”}$$

means "for some x of B, $z \epsilon A_x$". The *product-set*

$$\prod_{x \epsilon B} A_x$$

is the set of elements that belong to all the A_x; i.e.

$$\text{“}z \epsilon \prod_{x \epsilon B} A_x\text{”}$$

means "for every x of B, $z \epsilon A_x$". The notations for sum-set and product-set may be abbreviated to $\sum_x A_x$ and $\prod_x A_x$, or even ΣA and ΠA, when the meaning is clear. When the suffixes are positive integers the sum is denoted by

$$\sum_1^k A_n \quad \text{or} \quad \sum_1^\infty A_n,$$

and the product similarly; but it is to be emphasised that the infinite sum and product are not derived from the finite ones by any limiting process, but have an independent definition of their own.

Example. If A_n is the set of roots of the equation $z^n = 1$, $\sum_1^\infty A_n$ is the set of numbers $e^{2\pi i a}$, where a takes all rational values; and $\prod_1^\infty A_n$ is the single number 1.

The sum-set and product-set evidently coincide with the earlier sum and common part when the number of sets is finite. ΣA_x and ΠA_x may therefore in all cases be called simply the sum and common part of the sets A_x.

The formal properties of Σ and Π are:

E 1. *If* $a \epsilon B$, $\prod_{x \epsilon B} A_x \subseteq A_a \subseteq \sum_{x \epsilon B} A_x$.

2·1. *If, for every a of B,* $A_a \subseteq C$, *then* $\Sigma A_x \subseteq C$.

2·2. *If, for every a of B,* $A_a \supseteq C$, *then* $\Pi A_x \supseteq C$.

These propositions may be "translated" in the usual way; e.g. E 2·1 states that if $a \epsilon B$ implies $A_a \subseteq C$, then "$z \epsilon A_a$ and $a \epsilon B$" implies $z \epsilon C$.

F 1. *If* $A_x \subseteq B_x$ *for each x,* $\Sigma A_x \subseteq \Sigma B_x$ *and* $\Pi A_x \subseteq \Pi B_x$.

2. $\Sigma(A_x + B_x) = \Sigma A_x + \Sigma B_x$.

3·1. $\Pi(A + B_x) = A + \Pi B_x$.

3·2. $\Sigma(AB_x) = A \Sigma B_x$.

4. *If S contains all the sets* A_x, *then* ΣA_x *and* $\Pi(S - A_x)$ *are complementary sets in S; i.e. if* $\mathscr{C}$ *denotes the complement in S,* $\mathscr{C}(\Sigma A_x) = \Pi(\mathscr{C} A_x)$.

As a final example of the use of the "calculus" it will now be shewn that the propositions F are formally derivable from A–E. From this and other examples that have been given in this section it follows that all the "standard forms" A–F can be derived formally from A, B (without B 4), C, D 2 and E.(2)

Proof of F 1. For every a, $A_a \subseteq B_a \subseteq \Sigma B_x$ and $\Pi A_x \subseteq A_a \subseteq B_a$: apply E 2.

Proof of F 2. Since $A_a \subseteq \Sigma A_x$ and $B_a \subseteq \Sigma B_x$, $A_a + B_a \subseteq \Sigma A_x + \Sigma B_x$. Therefore by E 2·1

$$\Sigma(A_x + B_x) \subseteq \Sigma A_x + \Sigma B_x.$$

The other half follows from F 1.

Proof of F 3·1. Let $X = \Pi(A + B_x)$. Then, by F 1, $A \subseteq X$ and $\Pi B_x \subseteq X$, and therefore

$$A + \Pi B_x \subseteq X.$$

If B_a is one of the sets B_x, $X \subseteq A + B_a$, and therefore

$$X - A \subseteq (A + B_a) - A \subseteq B_a.$$

Hence $X - A \subseteq \Pi B_x$, and therefore $X \subseteq A + \Pi B_x$.