LATTICES AND THEIR APPLICATIONS*

GARRETT BIRKHOFF

It is my privilege to introduce to this Society a vigorous and promising younger brother of group theory, by name, lattice theory. Among other things, I shall try to bring out the family resemblance.

It is generally recognized that some familiarity with the notions of group, subgroup, normal subgroup, inner automorphism, commutator, and their technical properties, in a word, with group theory, is an essential preliminary to the understanding of algebraic equations, of differential equations, of the relation between the different branches of geometry, of automorphic functions, of crystallography, and of many other parts of mathematics and mathematical physics.

I shall try to convince you that, in the same way, some familiarity with the notions of lattice, sublattice, the modular identity, dual automorphism, chain, and their technical properties, in a word, with lattice theory, is an essential preliminary to the full understanding of logic, set theory, probability, functional analysis, projective geometry, the decomposition theorems of abstract algebra, and many other branches of mathematics.

It is often said that mathematics is a language. If so, group theory provides the proper vocabulary for discussing *symmetry*. In the same way, lattice theory provides the proper vocabulary for discussing *order*, and especially systems which are in any sense *hierarchies*. One might also say that just as group theory deals with permutations, so lattice theory deals with *combinations*.

One difference between the two is that whereas our knowledge of group theory has increased by not more than fifty per cent in the last thirty years, our knowledge of lattice theory has increased by perhaps two hundred per cent in the last ten years.

Lattice theory is based on a single undefined relation, the *inclusion* relation $x \leqq y$. In this it resembles group theory, which is based on one undefined operation, group multiplication. The relation of inclusion is assumed to satisfy three primary postulates:

---

* This paper and the five papers which follow it constitute a partial record of the Symposium on Lattice Theory arranged by the Program Committee and held at the Charlottesville meeting of the Society, April 15, 1938. The first three papers present the principal addresses, and the other three contain the remarks of the leaders of the discussion.

P1. *For all* x, $x \leqq x$ (reflexiveness).

P2. *If* $x \leqq y$ *and* $y \leqq x$, *then* $x \doteq y$ (anti-symmetry).

P3. *If* $x \leqq y$ *and* $y \leqq z$, *then* $x \leqq z$ (transitivity).

Following Hausdorff, we shall term a relation which satisfies P1–P3, a "partial ordering."

Mathematics abounds with examples of partial orderings. For instance, set-inclusion is a partial ordering, whether we consider all subsets of a class, or merely subsets "distinguished" by some special property (as for example the subalgebras of an abstract algebra). Again, the relation "x divides y" partially orders the integers. The real numbers are partially ordered by the relation $x \leqq y$ as usually interpreted. Since time is isomorphic with the real number system, the relation of time-priority is also a partial ordering. Curiously, priority even defines a partial ordering when understood in the sense of the special theory of relativity. Real functions are partially ordered if we let $f \leqq g$ mean that $f(x) \leqq g(x)$ for all x. And finally, partitions are partially ordered if we let $\Pi \leqq \Pi'$ mean that Π is a refinement (that is, subpartition) of Π'.

Hausdorff introduced the definition of a "partially ordered system" in the first edition of his *Mengenlehre*, but omitted it in the later editions. In view of the examples just mentioned, this diffidence seems unjustified, at least if we admit the philological principle of Zipf that it is reasonable to have a word for any frequently used concept.*

It is clear from the symmetry of conditions P1–P3 that the relation $y \geqq x$ meaning $x \leqq y$ defines from any given partial ordering another "dual" partial ordering. This "duality" pervades all lattice theory.

In partially ordered systems, special roles are played by elements 0 and I which satisfy $0 \leqq x$ and $x \leqq I$ for all x. They are clearly unique, and we shall consistently denote them by 0 and I. Moreover they are even important in philosophy; thus to say that we are all descended from Adam is simply to say that our genealogical tree has a 0.

It is also easy to define the notion of a *simply* ordered system or "chain" in terms of the inclusion relation. By a chain, we mean a partially ordered system in which the following postulate is satisfied:

P4. *Given* x *and* y, *either* $x \leqq y$ *or* $y \leqq x$.

A few partially ordered systems, such as the real numbers, are simply ordered, but the majority are not.

---

* Thus although there is no word signifying one's "stepmother's second cousin's son-in-law," we would coin a word to describe this relationship, if we had frequent occasion to talk about it.

On the other hand, almost all partially ordered systems contain chains as subsystems; and, in fact, the so-called "chain conditions" of abstract algebra are simply the conditions that such chains be well-ordered.

In any partially ordered system, one can define the "join" of a set of elements x_α as an element which (1) contains every x_α, and (2) is contained in all other elements which contain every x_α. One can define the "meet" of the x_α dually.*

These definitions specialize in many interesting ways. In set theory, they specialize to the usual definitions of the sum and product of sets. With subgroups, they define the subgroup generated by and the intersection of the x_α. With divisibility, they define the least common multiple and highest common factor of the x_α. Applied to real numbers, they define the l.u.b. and gr.l.b., and to real functions, the supremum and infimum of the x_α. Finally, what is usually called the "product" of two partitions is their "meet" in the sense just defined.

In general partially ordered systems not all sets of x have joins and meets. This leads us to define a "complete lattice" as a partially ordered system P, every subset of which has a join and a meet. It may be that every countable subset of P has a join and a meet, although P is not a complete lattice; in this case P is called a σ-lattice, by analogy with the usual notions of σ-rings and σ-fields of sets. If every two elements x, y of P have a join $x \cup y$ and a meet $x \cap y$, then every finite subset of P has the same property, and P is called a *lattice.*

The fact that $x \cup y$ and $x \cap y$ are (single-valued) binary operations, suggests regarding lattices as *abstract algebras* and indicates a relationship not only to groups, but also to rings, hypercomplex algebras, and so on.

Inclusion, and therefore all lattice definitions,† can be defined in terms of either operation; for example, $x \leqq y$ if and only if $x = x \cup y$. Moreover the two lattice operations have a number of important properties, such as the following:

L1. $x \cup x = x$ *and* $x \cap x = x$.

L2. $x \cup y = y \cup x$ *and* $x \cap y = y \cap x$.

L3. $x \cup (y \cup z) = (x \cup y) \cup z$ *and* $x \cap (y \cap z) = (x \cap y) \cap z$.

L4. $x = x \cup (x \cap y) = x \cap (x \cup y)$.

* There is some confusion as to the origin of these definitions; they are due to C. S. Peirce, American Journal of Mathematics, vol. 3 (1880), p. 33.

† Thus 0 can be defined through the identity $0 \cap x = 0$, and I through the identity $x \cap I = x$.

Conversely, it is easy to show that any system in which two operations are defined which satisfy L1–L4 is a lattice.

The analogy of lattices with groups and rings suggests obvious definitions of such notions as sublattice, homomorphism, automorphism, and so on. Also, the existence of special classes of rings (for example, commutative rings and fields) suggests looking for special classes of lattices. Among these, three may be cited: modular lattices, distributive lattices, and complemented lattices. These are defined by the assertion that they satisfy the following postulates, respectively:

L5. *If* $x \leqq z$, *then* $x \cup (y \cap z) = (x \cup y) \cap z$.

L6. $x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$ *and* $x \cup (y \cap z) = (x \cup y) \cap (x \cup z)$.

L7. *To every* x *corresponds an* x', *such that* $x \cap x' = 0$ *and* $x \cup x' = I$.

We shall use this classification later; we only note now that each condition is self-dual, and that any distributive lattice is modular.

In this connection, it is a curious thing that C. S. Peirce should have believed that every lattice was distributive. In fact he says "this is easy to prove, but the proof is too tedious to give here." Actually, it is hard to imagine a more serious blunder; I leave it to you to draw your own moral.

Two auxiliary notions play an important role in lattice theory. The first is that of a "modular" functional as a real function defined on a lattice and satisfying the condition:

M1. $m[x]+m[y]=m[x \cap y]+m[x \cup y]$.

Measure and dimension functions are modular.

The second is that of an intrinsic *lattice topology*. This is easy to define in chains by letting intervals be neighborhoods of their interior points; we shall only hint at the general definition by pointing out that if we define lim sup $\{x_k\}$ as the meet of the joins S_n of the sets S_n of x_k, $(k \geqq n)$, and lim inf $\{x_k\}$ dually, then (1) lim sup $\geqq$ lim inf for any sequence, and (2) when the two are equal, we may say that $\{x_k\}$ converges to the common limit.

I think it is now evident that lattice theory provides one with a useful language for discussing order and related concepts. Lattice theory also has much more specific applications.

Thus consider the isomorphism between sets and qualities, first propounded by Boole. With each quality* Q Boole associated the hypothetical set S of all objects having the quality Q; conversely,

* Qualities are called in logic "propositional functions"; a nontechnical synonym is the word "property."

with each set S of objects he associated the hypothetical "quality" Q of membership in S.

Boole's correspondence has many properties. In the first place, it is effectively one-one. Again, it identifies the logical proposition "Q implies Q_1" with the set-theoretical proposition $S \leqq S_1$. Similarly, it identifies the quality "Q or Q_1" with the set-theoretical sum $S \cup S_1$, the quality "Q and Q_1" with the product $S \cap S_1$, and the quality "not Q" with the set complement S' of the corresponding set.

But it is easy to show that the subsets of any class I satisfy P1–P3 and L1–L7 if the above notation is used; that is, in technical language, if they form a Boolean algebra. We infer that the algebra of qualities (or "propositional functions") is also a Boolean algebra.

This has interesting immediate consequences. Thus it reduces the "law of contradiction" and the "law of the excluded middle" (tertium non datur) to simple theorems on Boolean algebra, namely, $x \leqq (x')'$ and $(x')' \leqq x$.

Even more interesting is the light which it sheds on proposed modifictions of logic. One can very easily take exception to Boole's primitive ideas. For example, the existence of "categorical" propositions, distinguishing one object from all others, is questionable. Also, Brouwer and the intuitionist school have attacked the law of the excluded middle in a fairly convincing way. More recently, Tarski has shown that the unrestricted distributive law* implies the existence of "categorical" propositions; so this also is suspect. In the same vein, it has been pointed out by von Neumann and myself that quantum mechanics suggests a propositional calculus in which all laws except the distributive law L6 hold; even L5 and the law of the excluded middle are valid.

So much for logic, and this is a good place to emphasize the fact that not all logic is Boolean algebra; logic cannot be taught as a branch of lattice theory. But neither has Felix Klein's Erlanger Programm reduced geometry to the status of a branch of group theory. Lattice theory is like group theory in providing some, but not all, of the leading ideas of the parts of mathematics to which it applies.

The applications of lattice theory to set theory are of a more technical nature; I shall confine myself to a single illustration. Consider

* Obtained from the simple distributive law L6, as follows. By induction, L6 can be shown to imply

$$\prod_{i=1}^{n}\left\{\sum_{j=1}^{r} x_{i,j}\right\} = \sum_{\phi}\left\{\prod_{i=1}^{n} x_{i,\phi(i)}\right\},$$

summed over all ϕ. If we remove the restriction that the number of indices be finite, we obtain the transfinite distributive law.

the Boolean algebra B of all Borel sets of (say) a unit square. It is well known that if one ignores sets of measure zero, one gets a homomorphic image B/N of B, and that if one ignores sets of first (Baire) category, then one gets another B/F. It is a theorem on Boolean algebra due to Stone, and so a part of lattice theory, that this is because the sets of measure zero (and likewise those of first category) form an "ideal," that is, contain with any set all its subsets and with any two subsets their sum. Thus lattice theory tells us, at least in principle, how to find *all* the homomorphic images of a given Boolean algebra.

Lattice theory also suggests the rather remarkable theorems that B/N and B/F are "complete lattices." It suggests the question "are they isomorphic?" The answer is, that they are not. It then suggests that they form propositional calculi with the remarkable property of being "atom-free," that is, of containing no categorical propositions.

I think you are all aware of the recent changes in the foundations of the theory of probability and the tendency to identify it with the theory of measure. The necessity for this is indeed suggested directly by geometrical probabilities: if I is any region of unit area, then the measure of any subset S of I can be identified with the probability that a point thrown into I at random will come to rest in S.

From the axiomatic point of view, the common features can be easily described. Both measure functions and probability functions are additive functionals on Boolean algebras; both are "positive" in the sense that $x \geqq y$ implies $m[x] \geqq m[y]$; in both, $m[0]=0$. Probability is distinguished only by the special assumption $m[I]=1$.

Lattice theory makes it easy to discuss the "completeness" of these postulates, and gives one an easy vantage point from which to compare Jordan with Lebesgue measure, or Tornier's with Kolmogoroff's postulates for probability. It also leads to the neat concept of "stochastic distance," defined as $m[x \cup y]-m[x \cap y]$, and correlates it with the "dimensional distance" $d[x \cup y]-d[x \cap y]$ used by von Neumann in an entirely different connection. Besides being metric in the sense of Fréchet, stochastic distance has many other properties and uses.

These remarks do not go very deep; I principally wish to show that the new axiomatic foundations of general probability are lattice-theoretic, without discussing their importance.

I should next like to show how similar ideas lead to a vastly improved mathematical theory of dependent probabilities. In the theory of dependent probabilities (alias Markoff chains, alias stochastic processes), one expresses one's knowledge about a system Σ at any instant t by a *probability function* $p[x; t]$; $p[x; t]$ expresses the probability that Σ has the property x. This point of view is familiar in quantum mechanics where $p[S; t] = \int_S \psi(t)\psi^*(t)dV$, and it is used in classical statistical mechanics.

One also assumes that one's knowledge of Σ at time t can be projected into the future, but only imperfectly; in the simplest (finite) case, the dependence is expressed by a matrix of transition probabilities.† In the general case, the usual formulation is highly technical and involves integral equations with Lebesgue integrals. Lattice theory suggests a very much simpler formulation, in terms of linear operators on "partially ordered function spaces" which I shall discuss later.

Time prevents my giving further details‡ of this theory; I shall only mention that (1) it makes possible a ten-line proof of Markoff's fundamental theorem on convergence to the case of independent probabilities, and (2) allows one to generalize the mean ergodic theorem of von Neumann, through the method of G. D. Birkhoff, from deterministic to non-deterministic mechanics.

The application of lattice theory to functional analysis begins with the observation that there is a natural partial ordering for the elements of every significant space of *real* functions. Moreover this order is preserved under linear translations $x \rightarrow x+a$, and under multiplication $x \rightarrow \lambda x$ by positive scalars; it is inverted by the transformation $x \rightarrow -x$.

Indeed, under this partial ordering most function spaces become *lattices*; we shall call such function spaces "linear lattices." It is easy to prove that any linear lattice satisfies the distributive law L6, and that its elements x admit a Jordan decomposition into their positive and negative parts.

Again, the intrinsic topology (mentioned earlier) of most linear lattices is decidedly interesting. In the case of Banach spaces, not only are all additive functionals "modular," but an additive functional is *bounded* (in the sense of Banach) if and only if it is the difference of "positive" functionals, or equivalently, numerically bounded on all sets which are "bounded" in the lattice-theoretic sense of having upper and lower bounds. This leads to a purely lattice-theoretic notion of conjugate space, which generalizes Banach's essentially metrical notion.

Most function spaces thus satisfy L5 and L6 without satisfying L7.

† These matrices play a familiar role in Bayes' theorem.

‡ They are sketched in my paper *Dependent probabilities and spaces* (L), Proceedings of the National Academy of Sciences, vol. 24 (1938), pp. 154–159.

Projective geometries, on the other hand, satisfy L5 and L7 without satisfying L6; they are thus "complemented modular lattices." One can say that L5 and L7 characterize the algebra of classes; in this connection, it is interesting that projective geometry has been known traditionally as the "geometry of intersection and union" (Geometrie des Schneidens und Verbindens).

How completely L5 and L7 characterize projective geometry is shown by the fact that any complemented modular lattice whose chains are of bounded length is in a precise sense the "direct sum" of projective geometries, and conversely. This enables one to build up projective geometry in terms of the algebra of combination (L1–L5 and L7), of a finite chain condition, and of a condition of algebraic irreducibility. These conditions are all self-dual, and so make the celebrated "duality principle" of projective geometry apparent from the start. Incidentally, the irreducibility condition is the only condition which is not purely lattice-theoretic;† it is equivalent to the usual postulate that every line contains at least three points.

Von Neumann has shown how omission of the chain condition leads one to envisage *point-free*‡ projective geometries, in which a (modular) "dimension function" ranging continuously from 0 to 1 is defined. These "continuous-dimensional projective geometries" are beautiful analogues of the Boolean algebra of Borel sets modulo null sets; the role of measure is exactly replaced by that of dimension.

In these brief remarks, I have ignored the important applications of lattice theory to the foundations of abstract algebra, and especially of modular lattices to decomposition theory. I have also ignored the role played by lattices in the general theory of bicompact spaces.

HARVARD UNIVERSITY

† Besides, it is equivalent to the lattice-theoretic condition that every x not 0 or I has at least two complements.

‡ The word "point" (or "atom") is understood in the sense of Euclid, as an element $a>0$ which is indivisible, that is, is such that $a>x>0$ has no solution. In logic, the same condition characterizes "categorical propositions."