Z. Wahrscheinlichkeitstheorie 2, 340 368 (1964)

On the Foundations of Combinatorial Theory
I. Theory of Möbius Functions

By

GIAN-CARLO ROTA

Contents

1. Introduction

One of the most useful principles of enumeration in discrete probability and combinatorial theory is the celebrated *principle of inclusion-exclusion* (cf. FELLER*, FRÉCHET, RIORDAN, RYSER). When skillfully applied, this principle has yielded the solution to many a combinatorial problem. Its mathematical foundations were thoroughly investigated not long ago in a monograph by FRÉCHET, and it might at first appear that, after such exhaustive work, little else could be said on the subject.

One frequently notices, however, a wide gap between the bare statement of the principle and the skill required in recognizing that it applies to a particular combinatorial problem. It has often taken the combined efforts of many a combinatorial analyst over long periods to recognize an inclusion-exclusion pattern. For example, for the ménage problem it took fifty-five years, since CAYLEY's attempts, before JACQUES TOUCHARD in 1934 could recognize a pattern, and thence readily obtain the solution as an explicit binomial formula. The situation becomes bewildering in problems requiring an enumeration of any of the numerous collections of combinatorial objects which are nowadays coming to the fore. The counting of trees, graphs, partially ordered sets, complexes, finite sets on which groups act, not to mention more difficult problems relating to permutations with restricted position, such as Latin squares and the coloring of maps, seem to lie beyond present-day methods of enumeration. The lack of a systematic

theory is hardly matched by the consummate skill of a few individuals with a natural gift for enumeration.

This work begins the study of a very general principle of enumeration, of which the inclusion-exclusion principle is the simplest, but also the typical case. It often happens that a set of objects to be counted possesses a natural ordering, in general only a partial order. It may be unnatural to fit the enumeration of such a set into a linear order such as the integers: instead, it turns out in a great many cases that a more effective technique is to work with the natural order of the set. One is led in this way to set up a "difference calculus" relative to an arbitrary partially ordered set.

Looked at in this way, a surprising variety of problems of enumeration reveal themselves to be instances of the general problem of inverting an "indefinite sum" ranging over a partially ordered set. The inversion can be carried out by defining an analog of the "difference operator" relative to a partial ordering. Such an operator is the Möbius function, and the analog of the "fundamental theorem of the calculus" thus obtained is the Möbius inversion formula on a partially ordered set. This formula is here expressed in a language close to that of number theory, where it appears as the well-known inverse relation between the Riemann zeta function and the Dirichlet generating function of the classical Möbius function. In fact, the algebra of formal Dirichlet series turns out to be the simplest non-trivial instance of such a "difference calculus", relative to the order relation of divisibility.

Once the importance of the Möbius function in enumeration problems is realized, interest will naturally center upon relating the properties of this function to the structure of the ordering. This is the subject of the first paper of this series; we hope to have at least begun the systematic study of the remarkable properties of this most natural invariant of an order relation.

We begin in Section 3 with a brief study of the incidence algebra of a locally finite partially ordered set and of the invariants associated with it: the zeta function, Möbius function, incidence function, and Euler characteristic. The language of number theory is kept, rather than that of the calculus of finite differences, and the results here are quite simple.

The next section contains the main theorems: Theorem 1 relates the Möbius functions of two sets related by a Galois connection. By suitably varying one of the sets while keeping the other fixed one can derive much information. Theorem 2 of this section is suggested by a technique that apparently goes back to RAMANUJAN. These two basic results are applied in the next section to a variety of special cases; although a number of applications and special cases have been left out, we hope thereby to have given an idea of the techniques involved.

The results of Section 6 stem from an "Ideenkreis" that can be traced back to Whitney's early work on linear graphs. Theorem 3 relates the Möbius function to certain very simple invariants of "cross-cuts" of a finite lattice, and the analogy with the Euler characteristic of combinatorial topology is inevitable. Pursuing this analogy, we were led to set up a series of homology theories, whose Euler characteristic does indeed coincide with the Euler characteristic which we had introduced by purely combinatorial devices.

Some of the work in lattice theory that was carried out in the thirties is useful in this investigation; it turns out, however, that modular lattices are not combinatorially as interesting as a type of structure first studied by WHITNEY, which we have called geometric lattices following BIRKHOFF and the French school. The remarkable property of such lattices is that their Möbius function alternates in sign (Section 7).

To prevent the length of this paper from growing beyond bounds, we have omitted applications of the theory. Some elementary but typical applications will be found in the author's expository paper in the American Mathematical Monthly. Towards the end, however, the temptation to give some typical examples became irresistible, and Sections 9 and 10 were added. These by no means exhaust the range of applications, it is our conviction that the Möbius inversion formula on a partially ordered set is a fundamental principle of enumeration, and we hope to implement this conviction in the successive papers of this series. One of them will deal with structures in which the Möbius function is multiplicative, —that is, has the analog of the number-theoretic property $\mu(mn) = \mu(m)\,\mu(n)$ if m and n are coprime — and another will give a systematic development of the Ideenkreis centering around POLYA's Hauptsatz, which can be significantly extended by a suitable Möbius inversion.

A few words about the history of the subject. The statement of the Möbius inversion formula does not appear here for the first time: the first coherent version—with some redundant assumptions—is due to WEISNER, and was independently rediscovered shortly afterwards by PHILIP HALL. Ward gave the statement in full generality. Strangely enough, however, these authors did not pursue the combinatorial implications of their work; nor was an attempt made to systematically investigate the properties of Möbius functions. Aside from HALL's applications to p-groups, and from some applications to statistical mechanics by M. S. GREEN and NETTLETON, little has been done; we give a hopefully complete bibliography at the end.

It is a pleasure to acknowledge the encouragement of G. BIRKHOFF and A. GLEASON, who spotted an error in the definition of a cross-cut, as well as of SEYMOUR SHERMAN and KAI-LAI CHUNG. My colleagues D. KAN, G. WHITEHEAD, and especially F. PETERSON gave me essential help in setting up the homological interpretation of the cross-cut theorem.

2. Preliminaries

Little knowledge is required to read this work. The two notions we shall not define are those of a *partially ordered set* (whose order relation is denoted by $\leqq$) and a *lattice*, which is a partially ordered set where max and min of two elements (we call them join and meet, as usual, and write them $\vee$ and $\wedge$) are defined. We shall use instead the symbols $\cup$ and $\cap$ to denote union and intersection *of sets* only. A *segment* $[x, y]$, for x and y in a partially ordered set P, is the set of all elements z between x and y, that is, such that $x \leqq z \leqq y$. We shall occasionally use open or half-open segments such as $[x, y)$, where one of the endpoints is to be omitted. A segment is endowed with the induced order structure; thus, a segment of a lattice is again a lattice. A partially ordered set is *locally finite* if every segment is finite. We shall only deal with locally finite partially ordered sets.

The *product* $P \times Q$ of partially ordered sets P and Q is the set of all ordered pairs (p, q), where $p \in P$ and $q \in Q$, endowed with the order $(p, q) \geqq (r, s)$ whenever $p \geqq r$ and $q \geqq s$. The product of any number of partially ordered sets is defined similarly. The *cardinal power* Hom (P, Q) is the set of all monotonic functions from P to Q, endowed with the partial order structure $f \geqq g$ whenever $f(p) \geqq g(p)$ for every p in P.

In a partially ordered set, an element p *covers* an element q when the segment $[q, p]$ contains two elements. An *atom* in P is an element that covers a minimal element, and a *dual atom* is an element that is covered by a maximal element.

If P is a partially ordered set, we shall denote by P^* the partially ordered set obtained from P by inverting the order relation.

A *closure relation* in a partially ordered set P is a function $p \to \bar{p}$ of P into itself with the properties (1) $\bar{p} \geqq p$; (2) $\bar{\bar{p}} = \bar{p}$; (3) $p \geqq q$ implies $\bar{p} \geqq \bar{q}$. An element is *closed* if $p = \bar{p}$. If P is a finite Boolean algebra of sets, then a closure relation on P defines a lattice structure on the closed elements by the rules $p \wedge q = p \cap q$ and $p \vee q = \overline{p \cup q}$, and it is easy to see that every finite lattice is isomorphic to one that is obtained in this way. A *Galois connection* (cf. ORE, p. 182ff.) between two partially ordered sets P and Q is a pair of functions $\zeta: P \to Q$ and $\pi: Q \to P$ with the properties: (1) both ζ and π are order-inverting; (2) for p in P, $\pi(\zeta(p)) \geqq p$, and for q in Q, $\zeta(\pi(q)) \geqq q$. Under these circumstances the mappings $p \to \pi(\zeta(p))$ and $q \to \zeta(\pi(q))$ are closure relations, and the two partially ordered sets formed by the closed sets are isomorphic.

In Section 7, the notion of a closure relation with the *Mac Lane-Steinitz exchange property* will be used. Such a closure relation is defined on the Boolean algebra P of subsets of a finite set E and satisfies the following property: if p and q are points of E, and S a subset of E, and if $p \notin \bar{S}$ but $p \in \overline{S \cup q}$, then $q \in \overline{S \cup p}$. Such a closure relation can be made the basis of WHITNEY's theory of independence, as well as of the theory of geometric lattices. The closed sets of a closure relation satisfying the MAC LANE-STEINITZ exchange property where every point is a closed set form a geometric (= matroid) lattice in the sense of BIRKHOFF (Lattice Theory, Chapter IX).

A partially ordered set P is said to have a 0 or a I if it has a unique minimal or maximal element. We shall always assume $0 \neq I$. A partially ordered set P having a 0 and a I satisfies the *chain condition* (also called the JORDAN-DEDEKIND chain condition) when all totally ordered subsets of P having a maximal number of elements have the same number of elements. Under these circumstances one introduces the *rank* $r(p)$ of an element p of P as the length of a maximal chain in the segment $[o, p]$, minus one. The rank of 0 is 0, and the rank of an atom is 1. The height of P is the rank of any maximal element, plus one.

Let P be a finite partially ordered set satisfying the chain condition and of height $n + 1$. The *characteristic polynomial* of P is the polynomial $\sum_{x \in P} \mu(0, x)\,\lambda^{n-r(x)}$, where r is the rank function (see the def. of μ below).

If A is a finite set, we shall write $n(A)$ for the number of elements of A.

3. The incidence algebra

Let P be a locally finite partially ordered set. The *incidence algebra* of P is defined as follows. Consider the set of all real-valued functions of two variables $f(x, y)$, defined for x and y ranging over P, and with the property that $f(x, y) = 0$ if $x \not\leq y$. The sum of two such functions f and g, as well as multiplication by scalars, are defined as usual. The product $h = fg$ is defined as follows:

$$h(x, y) = \sum_{x \leq z \leq y} f(x, z)\, g(z, y).$$

In view of the assumption that P is locally finite, the sum on the right is well-defined. It is immediately verified that this is an associative algebra over the real field (any other associative ring could do). The incidence algebra has an identity element which we write $\delta(x, y)$, the Kronecker delta.

The *zeta function* $\zeta(x, y)$ of the partially ordered set P is the element of the incidence algebra of P such that $\zeta(x, y) = 1$ if $x \leq y$ and $\zeta(x, y) = 0$ otherwise. The function $n(x, y) = \zeta(x, y) - \delta(x, y)$ is called the *incidence function*.

The idea of the incidence algebra is not new. The incidence algebra is a special case of a semigroup algebra relative to a semigroup which is easily associated with the partially ordered set. The idea of taking "interval functions" goes back to DEDEKIND and E. T. BELL; see also WARD.

Proposition 1. *The zeta function of a locally finite partially ordered set is invertible in the incidence algebra.*

Proof. We define the inverse $\mu(x, y)$ of the zeta function by induction over the number of elements in the segment $[x, y]$. First, set $\mu(x, x) = 1$ for all x in P. Suppose now that $\mu(x, z)$ has been defined for all z in the open segment $[x, y)$. Then set

$$\mu(x, y) = -\sum_{x \leq z < y} \mu(x, z).$$

Clearly μ is an inverse of ζ.

The function μ, inverse to ζ, is called the *Möbius function* of the partially ordered set P.

The following result, simple though it is, is fundamental:

Proposition 2. (Möbius inversion formula). *Let $f(x)$ be a real-valued function, defined for x ranging in a locally finite partially ordered set P. Let an element p exist with the property that $f(x) = 0$ unless $x \geq p$.*

Suppose that

(*) $$g(x) = \sum_{y \leq x} f(y).$$

Then

(**) $$f(x) = \sum_{y \leq x} g(y)\, \mu(y, x).$$

Proof. The function g is well-defined. Indeed, the sum on the right can be written as $\sum_{p \leq y \leq x} f(y)$, which is finite for a locally finite ordered set.

Substituting the right side of (*) into the right side of (**) and simplifying,

we get

$$\sum_{y \leq x} g(y)\, \mu(y, x) = \sum_{y \leq x} \sum_{z \leq y} f(z)\, \mu(y, x) = \sum_{y \leq x} \sum_{z} f(z)\, \zeta(z, y)\, \mu(y, x).$$

Interchanging the order of summation, this becomes

$$\sum_{z} f(z) \sum_{y \leq x} \zeta(z, y)\, \mu(y, x) = \sum_{z} f(z)\, \delta(z, x) = f(x), \quad \text{q. e. d.}$$

Corollary 1. *Let $r(x)$ be a function defined for x in P. Suppose there is an element q such that $r(x)$ vanishes unless $x \leq q$. Suppose that*

$$s(x) = \sum_{y \geq x} r(y).$$

Then

$$r(x) = \sum_{y \geq x} \mu(x, y)\, s(y).$$

The proof is analogous to the above and is omitted.

Proposition 3. (Duality). *Let P^* be the partially ordered set obtained by inverting the order of a locally finite partially ordered set P, and let μ^* and μ be the Möbius functions of P^* and P. Then $\mu^*(x, y) = \mu(y, x)$.*

Proof. We have, in virtue of Proposition 2 and Corollary 1,

$$\sum_{x \geq^* y \geq^* z} \mu^*(x, y) = \delta(x, z).$$

Letting $q(x, y) = \mu^*(y, x)$, it follows that q is an inverse of ζ in the incidence algebra of P. Since the inverse is unique, $q = \mu$, q. e. d.

Proposition 4. *The Möbius function of any segment $[x, y]$ of P equals the restriction to $[x, y]$ of the Möbius function of P.*

The proof is omitted.

Proposition 5. *Let $P \times Q$ be the direct product of locally finite partially ordered sets P and Q. The Möbius function of $P \times Q$ is given by*

$$\mu((x, y), (u, v)) = \mu(x, u)\, \mu(y, v), \; x, u \in P;\; y, v \in Q.$$

The proof is immediate and is omitted.

The same letter μ has been used for the Möbius functions of three partially ordered sets, and we shall take this liberty whenever it will not cause confusion.

Corollary (Principle of Inclusion-Exclusion). *Let P be the Boolean algebra of all subsets of a finite set of n elements. Then, for x and y in P,*

$$\mu(x, y) = (-1)^{n(y) - n(x)}, \qquad y \geq x,$$

where $n(x)$ denotes the number of elements of the set x.

Indeed, a Boolean algebra is isomorphic to the product of n chains of two elements, and every segment $[x, y]$ in a Boolean algebra is isomorphic to a Boolean algebra.

Aside of the simple result of Proposition 5, little can be said in general about how the Möbius function varies by taking subsets and homomorphic images of a partially ordered set. We shall see that more sophisticated notions will be required to relate the Möbius functions of two partially ordered sets.

Let P be a finite partially ordered set with 0 and I. The *Euler characteristic* E of P is defined as

$$E = 1 + \mu(0,1).$$

The simplest result relating to the computation of the Euler characteristic was proved by PHILIP HALL by combinatorial methods. We reprove it below with a very simple proof which shows one of the uses of the incidence algebra:

Proposition 6. *Let P be a finite partially ordered set with 0 and I. For every k, let C_k be the number of chains with k elements stretched between 0 and I. Then*

$$E = 1 - C_2 + C_3 - C_4 + \cdots.$$

Proof. $\mu = \zeta^{-1} = (\delta + n)^{-1} = \delta - n + n^2 \ldots$. It is easily verified that $n^{k-1}(x, y)$ equals the number of chains of k elements stretched between x and y. Letting $x = 0$ and $y = I$, the result follows at once.

It will be seen in section 6 that the Euler characteristic of a partially ordered set can be related to the classical Euler characteristic in suitable homology theories built on the partially ordered set.

Proposition 6 is a typical application of the incidence algebra. Several other results relating the number of chains and subsets with specified properties can often be expressed in terms of identities for functions in the incidence algebra. In this way, one obtains generalizations to an arbitrary partially ordered set of some classical identities for binomial coefficients. We shall not pursue this line here further, since it lies out of the track of the present work.

Example 1. The classical Möbius function $\mu(n)$ is defined as $(-1)^k$ if n is the product of k distinct primes, and 0 otherwise. The classical inversion formula first derived by Möbius in 1832 is:

$$g(m) = \sum_{n|m} f(n); \quad f(m) = \sum_{n|m} g(n)\,\mu\left(\frac{m}{n}\right).$$

It is easy to see (and will follow trivally from later results) that $\mu\left(\frac{m}{n}\right)$ is the Möbius function of the set of positive integers, with divisibility as the partial order. In this case the incidence algebra has a distinguished subalgebra, formed by all functions $f(n, m)$ of the form $f(n, m) = G\left(\frac{m}{n}\right)$. The product $H = FG$ of two functions in this subalgebra can be written in the simpler form

$$(*) \qquad H(m) = \sum_{kn=m} F(k)\,G(n).$$

If we associate with the element F of this subalgebra the *formal Dirichlet series* $\hat{F}(s) = \sum_{n=1}^{\infty} F(n)/n^s$, then the product (*) corresponds to the product of two formal Dirichlet series considered as functions of s, $\hat{H}(s) = \hat{F}(s)\,\hat{G}(s)$. Under this representation, the zeta function of the partially ordered set is the classical *Riemann zeta function* $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$, and the statement that the Möbius function is

the inverse of the zeta function reduces to the classical identity $1/\zeta(s) = \sum_{n=1}^{\infty} \mu(n)/n^s$. It is hoped this example justifies much of the terminology introduced above.

Example 2. If P is the set of ordinary integers, then $\mu(m, n) = -1$ if $m = n - 1$, $\mu(m, m) = 1$, and $\mu(m, n) = 0$ otherwise. The Möbius inversion formula reduces to a well known formula of the calculus of finite differences, which is the discrete analog of the fundamental theorem of calculus.

The Möbius function of a partially ordered set can be viewed as the analog of the classical difference operator $\Delta f(n) = f(n + 1) - f(n)$, and the incidence algebra serves as a calculus of finite differences on an arbitrary partially ordered set.

4. Main results

It turns out that the Möbius functions of two partially ordered sets can be compared, when the sets are related by a Galois connection. By keeping one of the sets fixed, and varying the other from among sets with a simpler structure, such as Boolean algebras, subspaces of a finite vector space, partitions, etc., one can derive much information about a Möbius function. This is the program we shall develop. The basic result is the following:

Theorem 1. *Let P and Q be finite partially ordered sets, where P has a 0 and Q has a 0 and a 1. Let μ_P and μ be their Möbius functions. Let*

$$\pi: Q \to P; \quad \varrho: P \to Q$$

be a Galois connection such that

(1) $\qquad \pi(x) = 0$ if and only if $x = 1$.

(2) $\qquad \varrho(0) = 1$.

Then

$$\mu(0, 1) = \sum_{a>0} \mu_P(0, a)\,\zeta(\varrho(a), 0) = \sum_{[a:\varrho(a)=0]} \mu_P(0, a).$$

One gets a significant summand on the right for every $a > 0$ in P which is mapped into 0 by ϱ. One therefore expects the right side to contain "few" terms. In general, μ_P is a known function and μ is the function to be determined.

Proof. We shall first establish the identity

$$(*) \qquad \sum_{a \geqq b} \delta(\pi(x), a) = \zeta(x, \varrho(b))$$

for every b in P. Here ζ on the right stands for the zeta function of Q. Equation (*) is equivalent to the following statement: $\pi(x) \geqq b$ if and only if $x \leqq \varrho(b)$. But this latter statement is immediate from the properties of a Galois connection. Indeed, if $\pi(x) \geqq b$, then $\varrho(\pi(x)) \leqq \varrho(b)$, but $x \leqq \varrho(\pi(x))$, hence $x \leqq \varrho(b)$, and similarly for the converse implication.

To identity (*) we apply the Möbius inversion formula relative to P, thereby obtaining the identity

$$(**) \qquad \delta(\pi(x), 0) = \sum_{a \geqq 0} \mu_P(0, a)\,\zeta(x, \varrho(a)).$$

Now, $\delta(\pi(x), 0)$ takes the value 1 if and only if $\pi(x) = 0$, that is, in view of assumption (1), if and only if $x = 1$. For all other values of x, we have $\delta(\pi(x), 0) = 0$. Therefore,

$$\delta(\pi(x), 0) = 1 - n(x, 1).$$

We can now rewrite equation (**) in the form

$$1 - n(x, 1) = \zeta(x, \varrho(0)) + \sum_{a>0} \mu_P(0, a)\zeta(x, \varrho(a))$$

However, in view of assumption (2), $\zeta(x, \varrho(0)) = \zeta(x, 1)$, and this is identically one for all x in Q. Therefore, simplifying,

$$-n(x, 1) = \sum_{a>0} \mu_P(0, a)\zeta(x, \varrho(a)).$$

Now, since $\zeta = \delta + n$, we have $\mu = \delta - \mu n$, hence, recalling that $0 \neq 1$,

$$\mu(0, 1) = -\sum_{0 \leq x \leq 1} \mu(0, x) n(x, 1) = \sum_{0 \leq x \leq 1} \sum_{a>0} \mu_P(0, a)\mu(0, x)\zeta(x, \varrho(a)).$$

Interchanging the order of summation, we get

$$\mu(0, 1) = \sum_{a>0} \mu_P(0, a) \sum_{0 \leq x \leq 1} \mu(0, x)\zeta(x, \varrho(a)).$$

The last sum on the right equals $\delta(0, \varrho(a))$, and this equals $\zeta(\varrho(a), 0)$. The proof is therefore complete.

For simplicity of application, we restate Theorem 1 inverting the order of P.

Corollary. *Let $p : Q \to P$; $q : P \to Q$ be order preserving functions between P and Q such that*

(1) If $p(x) = 1$ then $x = 1$, and conversely.

(2) $q(1) = 1$.

(3) $p(q(x)) \leq x$ and $q(p(x)) \geq x$.

Then

$$\mu(0, 1) = \sum_{a<1} \mu_P(a, 1)\zeta(q(a), 0) = \sum_{[a: q(a)=0]} \mu_P(a, 1)$$

where μ is the Möbius function of Q.

The second result is suggested by a technique which apparently goes back to RAMANUJAN (cf. HARDY, RAMANUJAN, page 139).

Theorem 2. *Let Q be a finite partially ordered set with 0, and let P be a partially ordered set with 0. Let $p : Q \to P$ be a monotonic function of Q onto P. Assume that the inverse image of every interval $[0, a]$ in P is an interval $[0, x]$ in Q, and that the inverse image of 0 contains at least two points.*

Then

$$\sum_{[x: p(x)=a]} \mu(0, x) = 0$$

for every a in P.

The *proof* is by induction over the set P. Since $[0, 0]$ is an interval and its inverse image is an interval $[0, q]$ with $q > 0$, we have

$$\sum_{[x: p(x)=0]} \mu(0, x) = \sum_{0 \leq x \leq q} \mu(0, x) = 0.$$

Suppose now the statement is true for all b such that $b < a$ in P. Then

$$\sum_{b<a} \sum_{[x: p(x)=b]} \mu(0, x) = 0.$$

It follows that

$$\sum_{[x: p(x)=a]} \mu(0, x) = \sum_{b \leq a} \sum_{[x: p(x)=b]} \mu(0, x).$$

The last sum equals the sum over some interval $[0, r]$ which is the inverse image of the segment $[0, a]$, that is

$$\sum_{b \leq a} \sum_{[x: p(x)=b]} \mu(0, x) = \sum_{0 \leq x \leq r} \mu(0, x) = \delta(0, r).$$

But $r > 0$ because a is strictly greater than 0. Hence $\delta(r, 0) = 0$, and this concludes the proof.

5. Applications

The simplest (and typical) application of Theorem 1 is the following:

Proposition 1. *Let R be a subset of a finite lattice L with the following properties: $1 \notin R$, and for every x of L, except $x = 1$, there is an element y of R such that $y \geq x$.*

For $k \geq 2$, let q_k be the number of subsets of R containing k elements whose meet is 0. Then $\mu(0, 1) = q_2 - q_3 + q_4 + \cdots$.

Proof. Let $B(R)$ be the Boolean algebra of subsets of R. We take $P = B(R)$ and $Q = L$ in Theorem 1, and establish a Galois connection as follows. For x in L, let $\pi(x)$ be the set of elements of R which dominate x. In particular, $\pi(1)$ is the empty set. For A in $B(R)$, set $\varrho(A) = \bigwedge A$, namely, the meet of all elements of A, an empty meet giving as usual the element 1. This is evidently a Galois connection. Conditions (1) and (2) of the Theorem are obviously satisfied.

The function μ_P is given by the Corollary of Proposition 5 of Section 3, and hence the conclusion is immediate.

Two noteworthy special cases are obtained by taking R to be the set of dual atoms of Q, or the set of all elements < 1 (cf. also WEISNER).

Closure relations. A useful application of Theorem 1 is the following:

Proposition 2. *Let $x \to \bar{x}$ be a closure relation on a partially ordered set Q having 1, with the property that $\bar{x} = 1$ only if $x = 1$. Let P be the partially ordered subset of all closed elements of Q. Then:* (a) *If $\bar{x} > x$, then $\mu(x, 1) = 0$;* (b) *If $\bar{x} = x$, then $\mu(x, 1) = \mu_P(x, 1)$, where μ_P is the Möbius function of P.*

Proof. Considering $[x, 1]$, it may be assumed that P has a 0 and $x = 0$. We apply Corollary 1 of Theorem 1, setting $p(x) = \bar{x}$ and letting q be the injection map of P into Q. It is then clear that the assumptions of the Corollary are satisfied, and the set of all a in P such that $q(a) = 0$ is either the empty set or the single element 0, q. e. d.

Corollary (Ph. Hall). *If 0 is not the meet of dual atoms of a finite lattice L, or if 1 is not the join of atoms, then $\mu(0, 1) = 0$.*

Proof. Set $\bar{x} = \bigwedge A(x)$, where $A(x)$ is the set of dual atoms of Q dominating x, and apply the preceding result. The second assertion is obtained by inverting the order.

Example 1. *Distributive lattices.* Let L be a locally finite distributive lattice. Using Proposition 2, we can easily compute its Möbius function. Taking an interval

$[x, y]$ and applying Proposition 4 of Section 3, we can assume that L is finite. For $a \in L$, define $\bar{a}$ to be the join of all atoms which a dominates. Then $a \to \bar{a}$ is a closure relation in the inverted lattice L^*. Furthermore, the subset of closed elements is easily seen to be isomorphic to a finite Boolean algebra (cf. BIRKHOFF Lattice Theory, Ch. IX) Applying Proposition 5 of Section 3, we find: $\mu(x, y) = 0$ if y is not the join of elements covering x, and $\mu(x, y) = (-1)^n$ if y is the join of n distinct elements covering x.

In the special case of the integers ordered by divisibility, we find the formula for the classical Möbius function (cf. Example 1 of Section 3.).

The Möbius function of cardinal products. Let P and Q be finite partially ordered sets. We shall determine the Möbius function of the partially ordered set Hom(P, Q) of monotonic functions from P to Q, in terms of the Möbius function of Q. It turns out that very little information is needed about P.

A few preliminaries are required for the statement.

Let R be a subset of a partially ordered set Q with 0, and let $\bar{R}$ be the ideal generated by R, that is, the set of all elements x in Q which are below ($<$) some element of R. We denote by Q/R the partially ordered set obtained by removing off all the elements of $\bar{R}$, and leaving the rest of the order relation unchanged. There is a natural order-preserving transformation of Q onto Q/R which is one-to-one for elements of Q not in $\bar{R}$. We shall call Q/R the *quotient* of Q by the ideal generated by R.

Lemma. *Let $f: P \to Q$ be monotonic with range $R \subset Q$. Then the segment $[f, 1]$ in* Hom (P, Q) *is isomorphic with* Hom $(P, Q/R)$.

Proof. For g in $[f, 1]$, set $g'(x) = g(x)$ to obtain a mapping $g \to g'$ of $[f, 1]$ to Hom $(P, Q/R)$. Since $g \geqq f$, the range of g lies above R, so the map is an isomorphism.

Proposition 3. *The Möbius function μ of the cardinal product* Hom (P, Q) *of the finite partially ordered set P with the partially ordered set Q with 0 and 1 is determined as follows:*

(a) *If $f(p) \neq 0$ for some element p of P which is not maximal, then $\mu(0, f) = 0$.*

(b) *In all other cases,*

$$\mu(0, f) = \prod_m \mu(0, f(m)), \qquad f \in P,$$

where the product ranges over all maximal elements of P, and where μ on the right stands for the Möbius function of Q.

(c) *For $f \leqq g$, $\mu(f, g) = \mu(0, g')$, where g' is the image of g under the canomial map of $[f, 1]$ onto Hom $(P, Q/R)$, provided Q/R has a 0.*

Proof. Define a closure relation in $[0, f]^*$, namely the segment $[0, f]$ with the inverted order relation, as follows. Set $\bar{g}(m) = g(m)$ if m is a maximal element of P, and $g(a) = 0$ if a is not a maximal element of P. If $\bar{g} = 0$, then $g(m) = 0$ for all maximal elements m, hence $g(a) = 0$ for all $a <$ some maximal element, since g is monotonic. Hence $g = 0$, and the assumption of Proposition 2 is satisfied. The set of closed elements is isomorphic to Hom (M, P), where M is a set of as many elements as there are maximal elements in P. Conclusion (*a*) now follows from Proposition 2, and conclusion (b) from Proposition 5 of Section 3. Conclusion (c) follows at once from the Lemma.

We pass now to some applications of Theorem 2.

Proposition 4. *Let $a \to \bar{a}$ be a closure relation on a finite lattice Q, with the property that $\overline{a \vee b} = \bar{a} \vee \bar{b}$ and $\bar{0} > 0$. Then for all $a \in Q$,*

$$\sum_{[x: \bar{x} = a]} \mu(0, x) = 0.$$

Proof. Let P be a partially ordered set isomorphic to the set of closed elements of L. We define $p(x)$, for x in Q, to be the element of P corresponding to the closed element $\bar{x}$. Since $\bar{0} > 0$, any x between 0 and $\bar{0}$ is mapped into $\bar{0}$. Hence the inverse image of 0 in P under the homomorphism p is the nontrival interval $[0, \bar{0}]$.

Now consider an interval $[0, a]$ in P. Then $p^{-1}([0, a]) = [0, \bar{x}]$, where $\bar{x}$ is the closed element of L corresponding to a. Indeed, if $0 \leqq y \leqq \bar{x}$ then $\bar{y} \leqq \bar{\bar{x}} = \bar{x}$, hence $p(y) \leqq a$. Conversely, if $p(y) \leqq a$, then $\bar{y} \leqq \bar{x}$ but $y \leqq \bar{y}$, hence $y \leqq \bar{x}$. Therefore the condition of Theorem 2 is satisfied, and the conclusion follows at once.

Corollary (Weisner).

(a) *Let $a > 0$ in a finite lattice L. Then, for any b in L,*

$$\sum_{x \vee a = b} \mu(0, x) = 0$$

(b) *Let $a < 1$ in L. Then, for any b in L,*

$$\sum_{x \wedge a = b} \mu(x, 1) = 0.$$

Proof. Take $\bar{x} = x \vee a$. Part (b) is obtained by inverting the order.

Example 2. Let V be a finite-dimensional vector space of dimension n over a finite field with q elements. We denote by $L(V)$ the lattice of subspaces of V. We shall use Proposition 4 to compute the Möbius function of $L(V)$.

In the lattice $L(V)$, every segment $[x, y]$, for $x \leqq y$, is isomorphic to the lattice $L(W)$, where W is the quotient space of the subspace y by the subspace x. If we denote by $\mu_n = \mu_n(q)$ the value of $\mu(0, 1)$ for $L(V)$, it follows that $\mu(x, y) = \mu_j$, when j is the dimension of the quotient space W. Therefore once μ_n is known for for every n, the entire Möbius function is known.

To determine μ_n, consider a subspace a of dimension $n - 1$. In view of the preceding Corollary, we have for all $a < 1$ (where 1 stands for the entire space V):

$$\sum_{x \wedge a = 0} \mu(x, 1) = 0$$

where 0 stands of course for the 0-subspace. Let a be a dual atom of $L(V)$, that is, a subspace of dimension $n - 1$. Which subspaces x have the property that $x \wedge a = 0$? x must be a line in V, and such a line must be disjoint except for 0 from a. A subspace of dimension $n - 1$ contains q^{n-1} distinct points, so there will be $q^n - q^{n-1}$ points outside of a. However, every line contains exactly $q - 1$ points. Therefore, for each subspace a of dimension $n - 1$ there are

$$\frac{q^n - q^{n-1}}{q - 1} = q^{n-1}$$

distinct lines x such that $x \wedge a = 0$. Since each interval $[x, 1]$ is isomorphic to a space of dimension $n-1$, we obtain

$$\mu_n = \mu(0,1) = -\sum_{\substack{x\wedge a=0\\ x\neq 0}} \mu(x,1) = -q^{n-1}\mu_{n-1}.$$

This is a difference equation for μ_n which is easily solved by iteration. We obtain the result, first established by PHILIP HALL (see also WEISNER and S. DELSARTE):

$$\mu_n(q) = (-1)^n q^{n(n-1)/2} = (-1)^n q^{\binom{n}{2}}.$$

6. The Euler characteristic

Sharper results relating $\mu(0,1)$ to combinatorial invariants of a finite lattice can be obtained by application of Theorem 1, when the "comparison set" P remains a Boolean algebra.

A *cross-cut* C of a finite lattice L is a subset of L with the following properties:

(a) C does not contain 0 or 1.

(b) no two elements of C are comparable (that is, if x and y belong to C, then neither $x<y$ nor $x>y$ holds).

(c) Any maximal chain stretched between 0 and 1 meets the set C.

A *spanning subset* S of L is a subset such that $\vee S = 1$ and $\wedge S = 0$.

The main result is the following *Cross-cut Theorem*:

Theorem 3. *Let μ be the Möbius function and E the Euler characteristic of a non-trivial finite lattice L, and let C be a cross-cut of L. For every integer $k \geqq 2$, let q_k denote the number of spanning subsets of C containing k distinct elements. Then*

$$E-1 = \mu(0,1) = q_2 - q_3 + q_4 - q_5 + \cdots$$

The *proof* is by induction over the distance of a cross-cut C from the element 1.

Define the distance $d(x)$ of an element x from the element 1 as the maximum length of a chain stretched between x and 1. For example, the distance of a dual atom is two. If C is a cross-cut of L, define the distance $d(C)$ as $\max d(x)$ as x ranges over C. Thus, the distance of the cross-cut consisting of all dual atoms is two, and conversely, this is the only cross-cut having distance two.

It follows from Proposition 1 of Section 5 that the result holds when $d(C)=2$ (take $R=C$ in the assertion of the Proposition). Thus, we shall assume the truth of the statement for all cross-cuts whose distance is less than n, and prove it for a cross-cut with $d(C)=n$.

If C is a subset of L, we shall write $x>C$ or $x \leqq C$ to mean that there is an element y or C such that $x>y$, or that there is an element y of C such that $x \leqq y$. For a general C, these possibilities may not be mutually exclusive; they are mutually exclusive when C is a cross-cut. We shall repeatedly make use of this remark below.

Define a modified lattice L' as follows. Let L' contain all the elements x such that $x \leqq C$ in the same order. On top of C, add an element 1 covering all the elements of C, but no others; this defines L'.

In L', consider the cross-cut C and apply Proposition 1 of section 5 again. If μ' is the Möbius function of L', then

$$\mu'(0,1) = p_2 - p_3 + p_4 \ldots,$$

where p_k is the number of all subsets $A \subset C \subset L'$ of k elements, such that $\wedge A = 0$.

Comparing the lattices L and L', we have

$$0 = \sum_{x \leqq C} \mu(0,x) + \sum_{x>C} \mu(0,x) = \sum_{x \leqq C} \mu'(0,x) + \mu'(0,1).$$

However, for $x \leqq C$, we have $\mu'(0,x) = \mu(0,x)$ by construction of L'. Hence

$$\sum_{x \leqq C} \mu(0,x) = -p_2 + p_3 - p_4 + \cdots$$

Since the sets $(x/x \leqq C)$ and $(x/x > C)$ are disjoint, we can write

$$\mu(0,1) = -\sum_{x<1} \mu(0,x) = -\Big[\sum_{x \leqq C} \mu(0,x) + \sum_{1>x>C} \mu(0,x)\Big].$$

We now simplify the first summation on the right:

(*) $$\mu(0,1) = p_2 - p_3 + p_4 \cdots - \sum_{1>x>C} \mu(0,x).$$

Now let $q_k(x)$ be the number of subsets of C having k elements, whose meet is 0 and whose join is x. In particular, $q_k(1) = q_k$. Then clearly

$$p_k = \sum_{x>C} q_k(x), \quad k \geqq 2,$$

the summation in (*) can be simplified to

(**) $$\mu(0,1) = (q_2 - q_3 + q_4 - \cdots) - \sum_{1>x>C} [-q_2(x) + q_3(x) - q_4(x) + + \cdots + \mu(0,x)].$$

For x above C and unequal to 1, consider the segment $[0,x]$. We prove that $C(x) = C \cap [0,x]$ is a cross-cut of the lattice $[0,x]$ such that $d(C(x)) < d(C)$. Once this is done, it follows by the induction hypothesis that every term in brackets on the right of (**) vanishes, and the proof will be complete.

Conditions (a) and (b) in the definition of a cross-cut are trivially satisfied by $C(x)$, and condition (c) is verified as follows. Suppose Q is a maximal chain in $[0,x]$ which does not meet $C(x)$. Choose a maximal chain R in the segment $[x,1]$; then the chain $Q \cup R$ is maximal in L, and does not intersect C.

It remains to verify that $d(C(x)) < d(C)$, and this is quite simple. There is a chain Q stretched between C and x whose length is $d(C(x))$. Then $d(C)$ exceeds the length of the chain $Q \cup R$, and since $x<1$, R has length at least 2, hence the length of $Q \cup R$ exceeds that of Q by at least one. The proof is therefore complete.

Theorem 3 gives a relation between the value $\mu(0,1)$ and the width of narrow cross-cuts or *bottlenecks* of a lattice. The proof of the following statement is immediate.

Corollary 1. (a) *If L has a cross-cut with one element, then $\mu(0,1) = 0$.*

(b) *If L has a cross-cut with two elements, then the only two possible values of $\mu(0,1)$ are 0 and 1.*

(c) *If L has a cross-cut having three elements, then the only possible values of $\mu(0,1)$ are 2, 1, 0 and -1.*

In this connection, an interesting combinatorial problem is to determine all possible values of $\mu(0,1)$, given that L has a cross-cut with n elements.

Reduction of the main formula. In several applications of the cross-cut theorem, the computation of the number q_k of spanning sets may be long, and systematic procedures have to be devised. One such procedure is the following:

Proposition 1. *Let C be a cross-cut of a finite lattice L. For every integer $k \geqq 0$, and for every subset $A \subset C$, let $q(A)$ be the number of spanning sets containing A, and let $S_k = \sum_A q(A)$, where A ranges over all subsets of C having k elements. Set S_0 to be the number of elements of C. Then*

$$\mu(0,1) = S_0 - 2S_1 + 2^2 S_2 - 2^3 S_3 + \cdots.$$

Proof. For every subset $B \subset C$, set $p(B) = 1$ if B is a spanning set, and $p(B) = 0$ otherwise. Then

$$q(A) = \sum_{C \supseteq B \supseteq A} p(B).$$

Applying the Möbius inversion formula on the Boolean algebra of subsets of C, we get

$$p(A) = \sum_{B \supseteq A} q(B)\,\mu(A,B),$$

where μ is the Möbius function of the Boolean algebra. Summing over all subsets $A \subset C$ having exactly k elements,

$$q_k = \sum_{n(A)=k} p(A) = \sum_{n(A)=k} \sum_{B \supseteq A} q(B)\,\mu(A,B).$$

Interchanging the order of summation on the right, recalling Proposition 5 of Section 3 and the fact that a set of $k+l$ elements possesses $\binom{k+l}{l}$ subsets of k elements, we obtain

$$q_k = S_k - \binom{k+1}{1} S_{k+1} + \binom{k+2}{2} S_{k+2} \cdots + (-1)^{n-k} \binom{n}{k} S_n.$$

A convenient way of recasting this expression in a form suitable for computation is the following. Let V be the vector space of all polynomials in the variable x over the real field. The polynomials $1, x, x^2, \ldots$, are linearly independent in V. Hence there exists a linear functional L in V such that

$$L(x^k) = S_k, \quad k = 0, 1, 2, \ldots.$$

Formula (*) can now be rewritten in the concise form

$$q_k = L\left(x^k - (k+1)x^{k+1} + \binom{k+2}{2} x^{k+2} - \cdots\right) = L\left(\frac{x^k}{(1+x)^{k+1}}\right).$$

Upon applying the cross-cut theorem, we find the expression (where q_0 and q_1 are also given by (*), but turn out to be 0)

$$\begin{aligned}\mu(0,1) &= L\left(\frac{1}{1+x} - \frac{x}{(1+x)^2} + \frac{x^2}{(1+x)^3} - \cdots\right)\\ &= L\left(\frac{1}{1+2x}\right) = L(1 - 2x + 4x^2 - 8x^3 + \cdots)\\ &= S_0 - 2S_1 + 4S_2 - \cdots, \quad \text{q.e.d.}\end{aligned}$$

The cross-cut theorem can be applied to study which alterations of the order relation of a lattice preserve the Euler characteristic. Every alteration which preserves meets and joins of the spanning subsets of some cross-cut will preserve the Euler characteristic. There is a great variety of such changes, and we shall not develop a systematic theory here. The following is a simple case.

Following BIRKHOFF and JÓNSSON and TARSKI we define the *ordinal sum* of lattices as follows. Given a lattice L and a function assigning to every element x of L a lattice $L(x)$, (all the $L(x)$ are distinct) the *ordinal sum* $P = \sum_L L(x)$ of the lattices $L(x)$ *over* the lattice L is the partially ordered set P consisting of the set $\bigcup_{x \in L} L(x)$, where $u \leqq v$ if $u \in L(x)$ and $v \in L(x)$ and $u \leqq v$ in $L(x)$, or if $u \in L(x)$ and $v \in L(y)$ and $x < y$. It is clear that P is a lattice if all the $L(x)$ are finite lattices.

Proposition 2. *If the finite lattice P is the ordinal sum of the lattices $L(x)$ over the non-trivial lattice L, and μ_P, μ_x and μ_L are the corresponding Möbius functions, then: If $L(0)$ is the one element lattice, then $\mu_P(0,1) = \mu_L(0,1)$.*

Proof. The atoms of P are in one-to-one correspondence with the atoms of L and the spanning subsets are the same. Hence the result follows by applying the cross-cut theorem to the atoms.

In virtue of a theorem of JÓNSSON and TARSKI, every lattice P has a unique maximal decomposition into an ordinal sum over a "skeleton" L. This can be used in connection with the preceding Corollary to further simplify the computation of $\mu(0,n)$ as n ranges through P.

Homological interpretation. The alternating sums in the Cross-Cut Theorem suggest that the Euler characteristic of a lattice be interpreted as the Euler characteristic in a suitable homology theory. This is indeed the case. We now define* a *homology theory* $H(C)$ relative to an arbitrary cross-cut C of a finite lattice L. For the homological notions, we refer to Eilenberg-Steenrod.

Order the elements of C, say $a_1, a_2, \ldots, a_n$. For $k \geqq 0$, let a *k-simplex* σ be any subset of C of $k+1$ elements which does *not* span. Let C_k be the free abelian group generated by the k-simplices. We let $C_{-1} = 0$; for a given simplex σ, let σ_i be the set obtained by omitting the $(i+1)$-st element of σ, when the elements of σ are ordered according to the given ordering of C. The boundary of a k-simplex is defined as usual as $\partial_k \sigma = \sum_{i=0}^{k} (-1)^i \sigma_i$, and is extended by linearity to all of C_k, giving a linear mapping of C_k into C_{k-1}. The k-th homology group H_k is defined as the abelian group obtained by taking the quotient of the kernel of ∂_k by the image of ∂_{k+1}. The rank b_k of the abelian group H_k, that is, the number of independent generators of infinite cyclic subgroups of H_k, is the k-th *Betti number*.

Let α_k be the rank of C_k, that is, the number of k-simplices. The *Euler characteristic* of the homology $H(C)$ is defined in homology theory as

$$E(C) = \sum_{k=0}^{\infty} (-1)^k \alpha_k.$$

* This definition was obtained jointly with D. KAN, F. PETERSON and G. WHITEHEAD, whom I now wish to thank.

25*

It follows from well-known results in homology theory that

$$E(C) = \sum_{k=0}^{\infty} (-1)^k b_k .$$

Let q_k be the number of spanning subsets with k elements as in Theorem 3. Then $q_{k+1} + \alpha_k$ is the total number of subsets of C having $k+1$ elements; if C has N elements, then $\alpha_k = \binom{N}{k+1} - q_{k+1}$. It follows from the Cross-Cut Theorem that

$$E(C) = \sum_{k=0}^{\infty} (-1)^k \binom{N}{k+1} - \sum_{k=0}^{\infty} (-1)^k q_{k+1}$$
$$= \sum_{k=0}^{\infty} (-1)^k \binom{N}{k+1} + \mu(0,1).$$

We have however

$$\sum_{k=0}^{\infty} (-1)^k \binom{N}{k+1} = -\sum_{i=1}^{\infty} (-1)^i \binom{N}{i} = 1 - \sum_{i=0}^{\infty} (-1)^i \binom{N}{i} = 1 - (1-1)^N = 1,$$

and hence

$$E(C) = 1 + \mu(0,1) = E;$$

in other words:

Proposition 3. *In a finite lattice, the Euler characteristic of the homology of any cross-cut C equals the Euler characteristic of the lattice.*

This result can sometimes be used to compute the Möbius functions of "large" lattices. In general, the numbers q_k are rather redundant, since any spanning subset of k elements gives rise to several spanning subsets with more than k elements. A method for eliminating redundant spanning sets is then called for. One such method consists precisely in the determination of the Betti numbers b_k.

We conjecture that the Betti numbers of $H(C)$ are themselves independent of the cross-cut C, and are also "invariants" of the lattice L, like the Euler characteristic $E(C)$. In the special case of lattices of height 4 satisfying the chain condition, this conjecture has been proved (in a different language) by DOWKER.

Example 1. *The Betti numbers of a Boolean algebra.* We take the cross-cut C of all atoms. If the height of the Boolean algebra is $n+1$, then every k-cycle, for $k < n-2$, bounds, so that $b_0 = 1$ and $b_k = 0$ for $0 < k < n-2$. On the other hand, there is only one cycle in dimension $n-2$. Hence $b_{n-2} = 1$ and we find $E = 1 + (-1)^{n-2}$, which agrees with Proposition 5 of Section 3.

A notion of Euler characteristic for *distributive* lattices has been recently introduced by HADWIGER and KLEE. For finite distributive lattices, KLEE's Euler characteristic is related to the one introduced in this work. We refer to KLEE's paper for details.

7. Geometric lattices

An ordered structure of very frequent occurrence in combinatorial theory is the one that has been variously called matroid (WHITNEY), matroid lattice (BIRKHOFF), closure relation with the exchange property (MACLANE), geometric lattice

(BIRKHOFF), abstract linear dependence relation (BLEICHER and PRESTON). Roughly speaking, these structures arise in the study of combinatorial objects that are obtained by piecing together smaller objects with a particularly simple structure. The typical such case is a linear graph, which is obtained by piecing together edges. Several counting problems associated with such structures can often be attacked by Möbius inversion, and one finds that the Möbius functions involved have particularly simple properties.

We briefly summarize the needed facts out of the theory of such structures, referring to any of the works of the above authors for the proofs.

A finite lattice L is a *geometric lattice* when every element of L is the join of atoms, and whenever if a and b in L cover $a \wedge b$, then $a \vee b$ covers both a and b. Equivalently, a geometric lattice is characterized by the existence of a rank function satisfying $r(a \wedge b) + r(a \vee b) \leqq r(a) + r(b)$. Notice that this implies the chain condition. In particular if a is an atom, then $r(a \vee c) = r(c)$ or $r(c) + 1$. If M is a semimodular lattice, then the partially ordered subset of all elements which are joins of atoms is a geometric sublattice.

Geometric lattices are most often obtained from a closure relation on a finite set which satisfies the MACLANE-STEINITZ exchange property. The lattice L of closed sets in such a closure relation is a geometric lattice whenever every one-element set is closed. Conversely, every geometric lattice can be obtained in this way by defining one such closure relation on the set of its atoms.

The fundamental property of the Möbius function of geometric lattices is the following:

Theorem 4. *Let μ be the Möbius function of a finite geometric lattice L. Then:*

(a) *$\mu(x,y) \neq 0$ for any pair x, y in L, provided $x \leqq y$.*

(b) *If y covers z, then $\mu(x,y)$ and $\mu(x,z)$ have opposite signs.*

Proof. Any segment $[x,y]$ of a geometric lattice is also a geometric lattice. It will therefore suffice to assume that $x = 0$, $y = 1$ and that z is a dual atom of L.

We proceed by induction. The theorem is certainly true for lattices of height 2, where $\mu(0,1) = -1$. Assume it is true for all lattices of height $n-1$, and let L be a lattice of height n. By the Corollary to Proposition 4 of Section 5, with $b = 1$, and a an atom of L, we have

$$\mu(0,1) = -\sum_{\substack{x \vee a = 1 \\ x \neq 1}} \mu(0,x).$$

Now from the subadditive inequality

$$r(x \wedge a) + r(x \vee a) \leqq r(x) + r(a)$$

we infer that if $x \vee a = 1$, then $n \leqq \dim x + \dim a$, hence $\dim x \geqq n-1$. The element x must therefore be a dual atom. It follows from the induction assumption and from the fact that L satisfies the chain condition, that all the $\mu(0,x)$ in the sum on the right have the same sign, and none of them is zero. Therefore, $\mu(0,1)$ is not zero, and its sign is the opposite of that $\mu(0,x)$ for any dual atom x. This concludes the proof.

Corollary. *The coefficients of the characteristic polynomial of a geometric lattice alternate in sign.*

We next derive a combinatorial interpretation of the Euler characteristic of a geometric lattice, which generalizes a technique first used by WHITNEY in the study of linear graphs.

A subset $\{a, b, \ldots, c\}$ of a geometric lattice L is *independent* when

$$r(a \vee b \vee \cdots \vee c) = r(a) + r(b) + \cdots + r(c).$$

Let C_k be the cross-cut of L of all elements of rank $k > 0$. A maximal independent subset $\{a, b, \ldots, c\} \subset C_k$ is a *basis* of C_k. All bases of C_k have the same number of elements, namely, $n - k$ if the lattice has height n. A subset $A \subset C_k$ is a *circuit* (WHITNEY) when it is not independent but every proper subset is independent. A set is independent if and only if it contains no circuits.

Order the elements of L of rank k in a linear order, say $a_1, a_2, \ldots, a_l$. This ordering induces a lexicographic ordering of the circuits of C_k.

If the subset $\{a_{i_1}, a_{i_2}, \ldots, a_{i_j}\}$ $(i_1 < i_2 < \cdots < i_j)$ is a circuit, the subset $a_{i_1}, a_{i_2}, \ldots, a_{i_{j-1}}$ will be called a *broken circuit.*

Proposition 1. *Let L be a geometric lattice of height $n + 1$, and let C_k be the cross-cut of all elements of rank k. Then $\mu(0, 1) = (-1)^n m_k$, where m_k is the number of subsets of C_k whose meet is 0, containing $n - k + 1$ elements each, and not containing all the arcs of any broken circuit.*

Again, the assertion implies that $m_1 = m_2 = m_3 = \cdots$.

Proof. Let the lexicographically ordered broken circuits be $P_1, P_2, \ldots, P_\sigma$, and let S_i be the family of all spanning subsets of C_k containing P_i but not $P_1, P_2, \ldots,$ or P_{i-1}. In particular, $S_{\sigma+1}$ is the family of all those spanning subsets not containing all the arcs of any broken circuit. Let q_j^i be the number of spanning subsets of j elements and not belonging to S_i. We shall prove that for each $i \geqq 1$

$$\mu(0, 1) = q_2^i - q_3^i + q_4^i \cdots. \qquad (*)$$

First, set $i = 1$. The set S_1 contains all spanning subsets containing the broken circuit P_1. Let $\bar{P}_1$ be the cicuit obtained by completing the broken circuit P_1. — A spanning set contained in S_1 contains either $\bar{P}_1$ or else P_1 but not $\bar{P}_1$; call these two families of spanning subsets A and B, and let q_j^A and q_j^B be defined accordingly. Then $q_j = q_j^1 + q_j^A + q_j^B$, and

$$\mu(0, 1) = q_2 - q_3 + q_4 \cdots = q_2^1 - q_3^1 + \cdots +$$
$$+ q_2^A + (q_2^B - q_3^A) - (q_3^B - q_4^A) + \cdots.$$

Now, $q_2^A = 0$, because no circuit can contain two elements; there is a one-to-one correspondence between the elements of A and those of B, obtained by completing the broken circuit P_1. Thus, all terms in parentheses cancel and the identity (*) holds for $i = 1$.

To prove (*) for $i > 1$, remark that the element c_i of C_k, which is dropped from a circuit to obtain the broken circuit P_i, does not occur in any of the previous circuits, because of the lexicographic ordering of the circuits. Hence the induction can be continued up to $i = \sigma + 1$.

Any set belonging to $S_{\sigma+1}$ does not contain any circuit. Hence, it is an independent set. Since it is a spanning set, it must contain $n - k + 1$ elements. Thus, all the integers $q_{\sigma+1}$ vanish except $q_{n-k+1}^{\sigma+1}$ and the statement follows from (*), q.e.d.

Corollary 1. *Let $q(\lambda) = \lambda^n + m_1\lambda^{n-1} + m_2\lambda^{n-2} + \cdots + m_n$ be the characteristic polynomial of a geometric lattice of height $n + 1$. Then $(-1)^k m_k$ is a positive integer for $1 \leqq k \leqq n$, equal to the number of independent subsets of k atoms not containing any broken circuit.*

The *proof* is immediate: take $k = 1$ in the preceding Proposition.

The homology of a geometric lattice is simpler than that of a general lattice:

Proposition 2. *In the homology relative to the cross-cut C_k of all elements of rank $k = 1$, the Betti numbers $b_1, b_2, \ldots, b_{k-2}$ vanish.*

The *proof* is not difficult.

Example 1. *Partitions of a set.*

Let S be a finite set of n elements. A *partition* π of S is a family of disjoint subsets $B_1, B_2, \ldots, B_k$, called *blocks*, whose union is S. There is a (well-known) natural ordering of partitions, which is defined as follows: $\pi \leqq \sigma$ whenever every block of π is contained in a block of partition σ. In particular, 0 is the partition having n blocks, and I is the partition having one block. In this ordering, the partially ordered set of partitions is a geometric lattice (cf. BIRKHOFF).

The Möbius function for the lattice of partitions was first determined by SCHÜTZENBERGER and independently by ROBERTO FRUCHT and the author. We give a new proof which uses a recursion. If π is a partition, the *class* of π is the (finite) sequence $(k_1, k_2, \ldots)$, where k_i is the number of blocks with i elements.

Lemma. *Let L_n be the lattice of partitions of a set with n elements. If $\pi \in L_n$ is of rank k, then the segment $[\pi, 1]$ is isomorphic to L_{n-k}. If π is of class $(k_1, k_2, \ldots)$, then the segment $[0, \pi]$ is isomorphic to the direct product of k_1 lattices isomorphic to L_1, k_2 lattices isomorphic to L_2,* etc.

The *proof* is immediate.

It follows from the Lemma that if $[x, y]$ is a segment of L_n, then it is isomorphic to a product of k_i lattices isomorphic to L_i, $i = 1, 2, \ldots$. We call the sequence $(k_1, k_2, \ldots)$ the *class* of the *segment* $[x, y]$.

Proposition 3. *Let $\mu_n = \mu(0, 1)$ for the lattice of partitions of a set with n elements. Then $\mu_n = (-1)^{n-1}(n-1)!$.*

Proof. By the Corollary to Proposition 4 of Section 5, $\sum_{x \wedge a = 0} \mu(x, 1) = 0$. Let a be the dual atom consisting of a block C_1 containing $n - 1$ points, and a second block C_2 containing one point. Which non-zero partitions x have the property that $x \wedge a = 0$? Let the blocks of such a partition x be $B_1, \ldots, B_k$. None of the blocks B_i can contain two distinct points of the block C_1, otherwise the two points would still belong to the same block in the intersection. Furthermore, only one of the B_i can contain the block C_2. Hence, all the B_i contain one point, except one, which contains C_2 and an extra point. We conclude that x must be an atom, and there are $n - 1$ such atoms. Hence, $\mu_n = \mu(0, 1) = -\sum_x \mu(x, 1)$, where x ranges over a set of $n - 1$ atoms. By the Lemma, the segment $[x, 1]$ is isomorphic to the lattice of partitions of a set with $n-1$ elements, hence $\mu_n = -(n-1)\mu_{n-1}$. Since $\mu_2 = -1$, the conclusion follows.

Corollary. *If the segment* $[x, y]$ *is of class* $(k_1, k_2, \ldots, k_n)$, *then*

$$\mu(x,y) = \mu_1^{k_1}\mu_2^{k_2}\ldots\mu_n^{k_n} = (-1)^{k_1+k_2+\cdots+k_n-n}(2!)^{k_3}(3!)^{k_4}\ldots((n-1)!)^{k_n}.$$

The Möbius inversion formula on the partitions of a set has several combinatorial applications; see the author's expository paper on the subject.

8. Representations

There is, as is well known, a close analogy between combinatorial results relating to Boolean algebras and those relating to the lattice of subspaces of a vector space. This analogy is displayed for example in the theory of q-difference equations developed by F. H. JACKSON, and can be noticed in many number-theoretic investigations. In view of it, we are led to surmise that a result analogous to Proposition 1 of Section 5 exists, in which the Boolean algebra of subsets of R is replaced by a lattice of subspaces of a vector space over a finite field. Such a result does indeed exist; in order to establish it a preliminary definition is needed.

Let L be a finite lattice, and let V be a finite-dimensional vector space over a finite field with q elements. A *representation* of L over V is a monotonic map p of L into the lattice M of subspaces of V, having the following properties:

(1) $p(0) = 0$.

(2) $p(a \vee b) = p(a) \vee p(b)$.

(3) Each atom of L is mapped to a line of the vector space V, and the set of lines thus obtained spans the entire space V.

A representation is *faithful* when the mapping p is one-to-one. We shall see in Section 9 that a great many ordered structures arising in combinatorial problems admit faithful representations. Given a representation $p: L \to M$, one defines the *conjugate map* $q: M \to L$ as follows.

Let K be the set of atoms of M (namely, lines of V), and let A be the image under p of the set of atoms of L. For $s \in M$, let $K(s)$ be the set of atoms of M dominated by s, and let $B(s)$ be a minimal subset of A which spans (in the vector space sense) every element of $K(s)$. Let $A(s)$ be the subset of A which is spanned by $B(s)$. A simple vector-space argument, which is here omitted, shows that the set $A(s)$ is well defined, that is, that it does not depend upon the choice of $B(s)$, but only upon the choice of s.

Let $C(s)$ be the set of atoms of L which are mapped by p onto $A(s)$. Set $q(s) = \vee C(s)$ in the lattice L; this defines the map q. It is obviously a monotonic function.

Lemma. *Let* $p: L \to M$ *be a faithful representation and let* $q: M \to L$ *be the conjugate map. Assume that every element of* L *is a join of atoms. Then* $p(q(s)) \geq s$ *and* $q(p(x)) \leq x$.

Proof. By definition, $q(s) = \vee C(s)$, where $C(s)$ is the inverse image of $A(s)$ under p. By property (2) of a representation,

$$p(q(s)) = p(\vee C(s)) = \vee p(C(s)) = \vee A(s).$$

But this join of the set of lines $A(s)$ in the lattice M is the same as their span in the vector space V. Hence $\vee A(s) \geq s$, and we conclude that $p(q(s)) \geq s$.

To prove that $q(p(x)) \leq x$, it suffices to show that $A(p(x)) = B$, where B is the set of atoms in A dominated by $p(x)$. Clearly $B \subset A(p(x))$, and it will suffice to establish the converse implication. By (2), and by the fact that x is a join of atoms, we have $p(x) = \vee B$. Therefore every line l dominated by $p(x)$ is spanned by a subset of B. If in addition $l \in A$, then $l \leq \vee C$ for some subset $C \subset B$, hence $l \in B$. This shows $B \supset A(p(x))$, q.e.d.

Theorem 5. *Let* L *be a finite lattice, where every element is a join of atoms, let* $p: L \to M$ *be a faithful representation of* L *into the lattice* M *of subspaces of a vector space* V *over a finite field with* q *elements, and let* $q: M \to L$ *be the conjugate map. For every* $k \geq 2$, *let* m_k *be the number of* k*-dimensional subspaces* s *of* V *such that* $q(s) = I$. *Then*

$$(*) \qquad \mu(0,1) = q^{\binom{2}{2}} m_2 - q^{\binom{3}{2}} m_3 + q^{\binom{4}{2}} m_4 - \cdots,$$

where μ *is the Möbius function of* L.

Proof. Let $Q = L^*$, let $c: L \to Q$ and $c^*: Q \to L$ be the canonical isomorphisms between L and Q. Define $\pi: Q \to M$ as $\pi = pc^*$, and $\varrho: M \to Q$ as $\varrho = cq$. We verify that π and ϱ give a Galois connection between Q and M satisfying the hypothesis of Theorem 1. If $\pi(x) = 0$, then there is a $y \in L$ such that $y = c^*(x)$ and $p(y) = 0$. It follows from the definition of a representation that $y = 0$. Hence $x = c(y) = 1$. Furthermore, $\varrho(0) = c(q(0)) = 1$. It follows from the preceding Lemma that π and ϱ are a Galois connection. Applying Theorem 1 and the result of Example 2 of Section 5, formula (*) follows at once.

Remark. It is easy to see that every lattice having a faithful representation is a geometric lattice. The converse is however not true, as an example of T. LAZARSON shows.

A reduction similar to that of Proposition 1 of Section 7 can be carried out with Theorem 5 and representations, and another combinatorial property of the Euler characteristic is obtained.

9. The coloring of graphs

By way of illustration of the preceding theory, we give some applications to the classic problem of coloring of graphs, and to the problem of constructing flows in networks with specified properties. Our results extend previous work of G. D. BIRKHOFF, D. C. LEWIS, W. T. TUTTE and H. WHITNEY.

A *linear graph* $G = (V, E)$ is a structure consisting of a finite set V, whose elements are called vertices, together with a family E of two-element subsets of V, called edges. Two vertices a and b are adjacent when the set (a, b) is an edge; the vertices a and b are called the endpoints of (a, b). Alternately, one calls the vertices *regions* and calls the graph a *map*, and we use the two terms interchangeably, considering them as two words for the same object. If S is a set of edges, the *vertex set* $V(S)$ consists of all vertices which are incident to some edge in S.

A set of edges S is *connected* when in any partition $S = A \cup B$ into disjoint non-empty sets A and B, the vertex sets $V(A)$ and $V(B)$ are not disjoint. Every set of edges is the union of disjoint connected *blocks*.

The *bond closure* on a graph $G = (V, E)$ is a closure relation defined on the set E of edges as follows. If $S \subset E$, let $\bar{S}$ be the set of all edges both of whose endpoints belong to one and the same block of S. Every set consisting of a single edge is closed, and these are the only minimal non-empty closed sets.

Lemma 1. *The bond closure $S \to \bar{S}$ has the exchange property.*

Proof. Suppose e and f are edges, $S \subset E$, and $e \in \overline{S \cup f}$ but $e \notin \bar{S}$. Then every endpoint of e which is not in $V(S)$ is an endpoint of f; on the other hand, S and f have at least one point in common, otherwise $e \in \bar{S}$. Thus both e and f either connect the same two blocks of S, or else they have one endpoint in S and one common endpoint; hence $f \in \overline{S \cup e}$, q.e.d.

The lattice $L = L(G)$ of bond-closed subsets of E is called the *bond lattice* of the graph G. Suppose that E has n blocks and $p(\lambda)$ is the characteristic polynomial of L, then the polynomial $\lambda^n p(\lambda)$ is the *chromatic polynomial* of the graph G, first studied by G. D. BIRKHOFF. From Theorem 4 we infer at once the theorem of WHITNEY that the coefficients of the chromatic polynomial alternate in sign.

The chromatic polynomial has the following combinatorial interpretation. Let C be a set of n elements, called colors. A function $f: V \to C$ is a *proper coloring* of the graph, when no two adjacent vertices are assigned the same color. To every coloring f — not necessarily proper — there corresponds a subset of E, the *bond* of f, defined as the set of all edges whose endpoints are assigned the same color by f. The bond of f is a closed set of edges. For every closed set S, let $p(\lambda, S)$ be the number of colorings whose bond is S. Then we shall prove that $p(\lambda, S) = \lambda^n q(\lambda, S)$, where $q(\lambda, S)$ is the characteristic polynomial of the segment $[S, I]$ in the lattice L. Since every coloring has a bond $\sum_{T \geq S} p(\lambda, T)$ equals the total number of colorings having some bond $T \geq S$. But this number is evidently $\lambda^{k \cdot r(s)}$, where k is the number of vertices of the graph and $r(S)$ is the rank of S in L. Applying the Möbius inversion formula on the bond-lattice, we get

$$(*) \qquad p(\lambda) = p(\lambda, 0) = \sum_{T \in L} \lambda^{k - r(T)} \mu(0, T).$$

But the number of colorings whose bond is the null set 0 is exactly the number of proper colorings.

WHITNEY's evaluation (cf. A logical expansion in Mathematics) of the chromatic polynomials of a graph in terms of the number of subgraphs of s edges and p connected components is an immediate consequence of the cross-cut theorem applied to the atoms of the bond-lattice of G. This result of WHITNEY's can now be sharpened in two directions: first, a cross-cut other than that of the atoms can be taken; secondly, the computation of the coefficients of the chromatic polynomial can be simplified by Proposition 1 of Section 8. The cross-cut of all elements of rank 2 is particularly suited for computation, and can be programmed. The interested reader may wish to explicitly translate the cross-cut theorem and the results of Section 8 into the geometric language of graphs.

Example 1. For a *complete graph* on n vertices, where every two-element subset is an edge, the bond-lattice is isomorphic to the lattice of partitions of a set with n elements. The chromatic polynomial is evidently $(\lambda)_n = \lambda(\lambda - 1) \ldots (\lambda - n + 1)$, and the coefficients $s(n, k)$ are the *Stirling numbers of the first kind.*

Thus, $\sum_{r(\pi)=k} \mu(0, \pi) = s(n, k)$. This gives a combinatorial interpretation to the Stirling numbers of the first kind.

For a map $\mathfrak{m}$ embedded in the plane, where regions and boundaries have their natural meaning and no region bounds with itself, one obtains an interesting geometric result by applying the cross-cut theorem to the dual atoms of the bond lattice $L(\mathfrak{m})$.

Let $\mathfrak{m}$ be a connected map in the plane; without loss of generality we can assume: (a) that all the regions of $\mathfrak{m}$, except one which is unbounded, lie inside a convex polygon, the outer boundary of $\mathfrak{m}$; (b) that all boundaries are segments of straight lines. The *dual graph* of $\mathfrak{m}$ is the linear graph made up of the boundaries of $\mathfrak{m}$. A *circuit* in a linear graph is defined as a simple closed curve contained in the graph. We give an expression of the polynomial $P(\lambda, \mathfrak{m})$ in terms of the circuits of the dual graph. The outer boundary is always a circuit.

A set of circuits of a map $\mathfrak{m}$ in the plane *spans*, when their union — in the set-theoretic sense — is the entire boundary of $\mathfrak{m}$.

Proposition 1. *For every integer $k \geq 1$, let C_k be the number of spanning sets of k distinct circuits of a map $\mathfrak{m}$ in the plane. Then*

$$\mu_{\mathfrak{m}}(0, 1) = -C_1 + C_2 - C_3 + C_4 - \cdots$$

Proof. If the map has two regions, then $C_1 = 1$ and all other $C_c = 0$, so the result is trivial. Assume now that $\mathfrak{m}$ has at least 3 regions. Then $C_1 = 0$. All we have to prove is that the integers C_k are the integers q_k of Theorem 3, relative to the cross-cut of $L(\mathfrak{m})$ consisting of all the dual atoms.

By the Jordan curve theorem, every circuit divides the plane into two regions; this gives a one-to-one correspondence of the circuits with the dual atoms of $L(\mathfrak{m})$. Conversely, because we can assume that the map is of the special type described above, every dual atom in $L(\mathfrak{m})$ is a map with two connected regions, and so must have as a boundary a simple closed curve, q.e.d.

It has been shown by RICHARD RADO (p. 312) that the bond-lattice $L(G)$ of any linear graph G has a faithful representation. Accordingly, Theorem 5 can also be applied to obtain expression for $\mu(0,1)$. These expressions usually give sharper bounds than similar expressions based upon the cross-cut of atoms.

Farther-reaching techniques for the computation of the Möbius function of $L(G)$ are obtained by applying Theorem 1 to situations where P and Q are both bond-lattices of graphs. This we shall now do. A *monomorphism* of a graph G into a graph H is a one-to-one function f of the vertices of G *onto* the vertices of H, which induces a map $\bar{f}$ of the edges of G *into* the edges of H. Every monomorphism $f: G \to H$ induces a monotonic map $p: L(G) \to L(H)$, where $p(S)$ is defined as the closure of the image $\bar{f}(S)$ in H. It also induces a monotonic map $q: L(H) \to L(G)$, where $q(T)$ is defined as the set of edges of G whose image is in T.

Lemma 2. *$q(p(S)) = S$ for S in $L(G)$ and $p(q(T)) \leq T$ for T in $L(H)$.*

Proof. Intuitively, $p(S)$ is obtained by "adding edges" to S, and $q(p(S))$ simply removes the added edges. Thus, the first statement is graphically clear. The second one can be seen as follows. $q(T)$ is obtained from T by removing a number of edges. Taking $p(q(T))$, some of the edges may be replaced, but in general not all. Thus, $p(q(T)) \leqq T$.

Taking $M = L(H)^*$ and $c : L(H) \to M$ to be the canonical order-inverting map, we see that $\pi = cp$ and $\zeta = qc$ give a Galois connection between $L(G)$ and M. Now, $\pi(x) = 0$ is equivalent to $p(x) = 1$ for $x \in L(G)$. This can happen only if x has only one component, that is — since x is closed — only if $x = 1$ in $L(G)$. Thus $\pi(x) = 0$ if and only if $x = 1$. Secondly, $\varrho(0) = q(1) = 1$, evidently. We have verified all the hypotheses of Theorem 1, and we therefore obtain:

Proposition 2. *Let $f : G \to H$ be a monomorphism of a linear graph G into a linear graph H, and let μ_G and μ_H be the Möbius functions of the bond-lattices. Then*

$$\mu_G(0, 1) = \sum_{[a \in L(H);\, q(a)=0]} \mu_H(a, 1),$$

where q is the map of $L(H)$ into $L(G)$ naturally associated with f, as above.

Proposition 1 can be used to derive a great many of the reductions of G. D. BIRKHOFF and D. C. LEWIS, and provides a systematic way of investigating the changes of Möbius functions — and hence of the chromatic polynomial — when edges of a graph are removed. It has a simple geometric interpretation.

An interesting application is obtained by taking H to be the complete lattice on n elements. We then obtain a formula for μ which completes the statements of Theorems 3 and 5. Let G be a linear graph on n vertices. Let C be the family of two-element subsets of G which are not edges of G. Let F be the family of all subsets of C which are closed sets in the bond-lattice of the complete graph on n vertices built on the vertices of G. Then,

Corollary.

$$\mu_G(0, 1) = \sum_{a \in F} \mu(a, 1),$$

where μ is the Möbius function of the lattice of partitions (cf. Example 5) of a set of n elements.

Stronger results can be obtained by considering "epimorphisms" rather than "monomorphisms" of graphs, relating μ_G to the Möbius function obtained from G by "coalescing" points. In this way, one makes contact with G. A. DIRAC's theory of critical graphs. We leave the development of this topic to a later work.

10. Flows in networks

A network $N = (V, E)$ is a finite set V of vertices, together with a set of ordered pairs of vertices, called edges.

We shall adopt for networks the same language as for linear graphs.

A *circuit* is a sequence of edges S such that every vertex in $V(S)$ belongs to exactly two edges of S. Every edge has a positive and a negative endpoint. Given a function Φ from E to the integers from 0 to $\lambda - 1$, let for each vertex v, $\overline{\Phi}(v)$ be defined as

$$\overline{\Phi}(v) = \sum_e \eta(e, v)\, \Phi(e),$$

where the sum ranges over all edges incident to v, and the function $\eta(e, v)$ takes

the value $+1$ or -1 according as the positive or negative end of the edge e abuts at the vertex v, and the value zero otherwise. The function Φ is a *flow* (mod. λ) when $\overline{\Phi}(v) = 0$ (mod. λ) for every vertex v. The value $\Phi(e)$ for an edge e is called the *capacity* of the *flow* through e. The mod. λ restriction is inessential, but will be kept throughout.

A *proper flow* is one in which no edge is assigned zero capacity. TUTTE was the first to point out the importance of the problem of counting proper flows (cf. A contribution to the theory of chromatic polynomials) in combinatorial theory.

We shall reduce the solution of the problem to a Möbius inversion on a lattice associated with the network. This will give an expression for the number of proper flows as a polynomial in λ, whose coefficients are the values of a Möbius function.

Every flow through N is a proper flow of a suitable subnetwork of N, obtained by removing those edges which are assigned capacity 0. However, the converse of this assertion is not true: given a subnetwork S of N, it may not be possible to find a flow which is proper on the complement of N. This happens because every flow which assigns capacity zero to each edge of S may assign capacity zero to some further edges. We are therefore led to define a closure relation on the set of all subgraphs as follows: $\bar{S}$ shall be the set of all edges which necessarily are assigned capacity zero, in any flow of N which assigns capacity zero to every edge of S. In other words, if $e \notin \bar{S}$, then there is a flow in N which assigns capacity $\neq 0$ to the edge e, but which assigns capacity zero to all the edges of $\bar{S}$. It is immediately verified that $S \to \bar{S}$ is a closure relation. We call it the *circuit closure* of S. The circuit closure *has the exchange property*: if $e \in \overline{\bar{S} \cup p}$ but $e \notin \bar{S}$, then $p \in \overline{\bar{S} \cup e}$. Before verifying it, we first derive a geometric characterization of the circuit closure. A set S is circuit closed ($S = \bar{S}$) if and only if through every edge e not in S there passes a circuit which is disjoint from S. For if S is closed and $e \notin S$, then there is a flow through e and disjoint from S. But this can happen only if there is a circuit through e.

If there is a circuit through the edge p disjoint from $\overline{\bar{S} \cup e}$, and a circuit through e disjoint from $\bar{S}$ and containing p, then there is — as has been observed by WHITNEY — also a circuit through e not containing $\bar{S} \cup p$. This implies that e is not in the closure of $\bar{S} \cup p$, and verifies the exchange property.

The lattice $C(N)$ of closed subsets of edges of the network N is the *circuit lattice of N*. An atom in this lattice is not necessarily a single edge.

Proposition 1. *The number of proper flows,* (mod. λ) *on a network N with v vertices, e edges and p connected components is a polynomial $p(\lambda)$ of degree $e - v + p$. This polynomial is the characteristic polynomial of the circuit lattice of N. The coefficients alternate in sign.*

Proof. The last statement is an immediate consequence of Theorem 4 of Section 8.

The total number of flows on N (not necessarily proper) is determined as follows. Assume for simplicity that N is connected. Remove a set D of $v - 1$ edges from N, one adjacent to each but one of the vertices.

Every flow on N can be obtained by first assigning to each of the edges not in D an arbitrary capacity, between 0 and $\lambda - 1$, and then filling in capacities for the edges in D to match the requirement of zero capacity through each vertex. There are λ^{e-v+1} ways of doing this, and this is therefore the total number of flows mod. λ. If the network is in p connected components, the same argument gives λ^{e-v+p}. Now, every flow on G is a proper flow on a unique closed subset $\bar{S}$, obtained by removing all edges having capacity zero.

Hence

$$\lambda^{e-v+p} = \sum_{\bar{S} \in \underline{C}(G)} p(\bar{S}, \lambda),$$

where $p(\bar{S}, \lambda)$ is the characteristic polynomial of the closed subgraph $\bar{S}$. Setting $n(s) = e(s) - v(s) + p(s)$, the number of edges, vertices and components of s, and applying the inversion formula, we get

$$p(G, \lambda) = \sum_{S \in \underline{C}(G)} \lambda^{n(s)} \mu(S, G), \quad \text{q.e.d.}$$

In the course of the proof we have also shown that $n(s)$ is the *rank* of S in the circuit lattice of G. The rank of the null subgraph is one.

The four-color problem is equivalent to the statement that every planar network without an isthmus has a proper flow mod 5. (An isthmus is an edge that disconnects a component of the network when removed.)

Most of the results of the preceding section extend to circuit lattices of a network, and give techniques for computation of the flow polynomials of networks. We shall not write down their translation into the geometric language of networks.

References

AUSLANDER, L., and H. M. TRENT: Incidence matrices and linear graphs. J. Math. Mech. 8, 827 — 835 (1959).

BELL, E. T.: Algebraic Arithmetic. New York: Amer. Math. Soc. (1927).

— Exponential polynomials. Ann. of Math., II. Ser. **35**, 258—277 (1934).

BERGE, C.: Théorie des graphes et ses applications. Paris: Dounod 1958.

BIRKHOFF, GARRETT: Lattice Theory, third preliminary edition. Harvard University, 1963.

— Lattice Theory, revised edition. American Mathematical Society, 1948.

BIRKHOFF, G. D.: A determinant formula for the number of ways of coloring a map. Ann. of Math., II. Ser. **14**, 42—46 (1913).

—, and D. C. LEWIS: Chromatic polynomials. Trans. Amer. math. Soc. **60**, 355—451 (1946).

BLEICHER, M. N., and G. B. PRESTON: Abstract linear dependence relations. Publ. Math., Debrecen **8**, 55—63 (1961).

BOUGAYEV, N. V.: Theory of numerical derivatives. Moscow, 1870—1873, pp. 1—222.

BRUIJN, N. G. DE: Generalization of Polya's fundamental theorem in enumerative combinatorial analysis. Indagationes math. **21**, 59—69 (1959).

CHUNG, K.-L., and L. T. C. HSU: A combinatorial formula with its application to the theory of probability of arbitrary events. Ann. math. Statistics **16**, 91—95 (1945).

DEDEKIND, R.: Gesammelte Mathematische Werke, volls. I—II—III. Hamburg: Deutsche Math. Verein. (1930).

DELSARTE, S.: Fonctions de Möbius sur les groupes abéliens finis. Ann. of Math., II. Ser. **49**, 600—609 (1948).

DILWORTH, R. P.: Proof of a conjecture on finite modular lattices. Ann. of Math., II. Ser. **60**, 359—364 (1954).

DIRAC, G. A.: On the four-color conjecture. Proc. London math. Society, III. Ser. **13**, 193 to 218 (1963).

DOWKER, C. H.: Homology groups of relations. Ann. of Math., II. Ser. **56**, 84—95 (1952).

DUBREIL-JACOTIN, M.-L., L. LESIEUR et R. CROISOT: Leçons sur la théorie des treilles des structures algebriques ordonnées et des treilles géometriques. Paris: Gauthier-Villars 1953.

EILENBERG, S., and N. STEENROD: Foundations of algebraic topology. Princeton: University Press 1952.

FARY, I.: On straight-line representation of planar graphs. Acta Sci. math. Szeged **11**, 229—233 (1948).

FELLER, W.: An introduction to probability theory and its applications, second edition. New York: Wiley 1960.

FRANKLIN, P.: The four-color problem. Amer. J. Math. **44**, 225—236 (1922).

FRÉCHET, M.: Les probabilités associées à un système d'évenements compatibles et dépendants. Actualitées scientifiques et industrielles, nos. 859 et 942. Paris: Hermann 1940 et 1943.

FRONTERA MARQUÉS, B.: Una función numérica en los retículos finitos que se anula para los retículos reducibles. Actas de la 2a, Reunión de matemáticos españoles. Zaragoza 103—111 1962.

FRUCHT, R., and G.-C. ROTA: La función de Möbius para el retículo di particiones de un conjunto finito. To appear in Scientia (Chile).

GOLDBERG, K., M. S. GREEN and R. E. NETTLETON: Dense subgraphs and connectivity. Canadian J. Math. **11** (1959).

GOLOMB, S. W.: A mathematical theory of discrete classification. Fourth Symposium in Information Theory, London, 1961.

GREEN, M. S., and R. E. NETTLETON: Möbius function on the lattice of dense subgraphs. J. Res. nat. Bur. Standards **64B**, 41—47 (1962).

— — Expression in terms of modular distribution functions for the entropy density in an infinite system. J. Chemical Physisc **29**, 1365—1370 (1958).

HADWIGER, H.: Eulers Charakteristik und kombinatorische Geometrie. J. reine angew. Math. **194**, 101—110 (1955).

HALL, PHILIP: A contribution to the theory of groups of prime power order. Proc. London math. Soc., II. Ser. **36**, 39—95 (1932).

— The Eulerian functions of a group. Quart. J. Math. Oxford Ser. 134—151, 1936.

HARARY, F.: Unsolved problems in the enumeration of graphs. Publ. math. Inst. Hungar. Acad. Sci. **5**, 63—95 (1960).

HARDY, G. H.: Ramanujan. Cambridge: University Press 1940.

—, and E. M. WRIGHT: An introduction to the theory of numbers. Oxford: University Press 1954.

HARTMANIS, J.: Lattice theory of generalized partitions. Canadian J. Math. **11**, 97—106 (1959).

HILLE, E.: The inversion problems of Möbius. Duke math. J. **3**, 549—568 (1937).

HSU, L. T. C.: Abstract theory of inversion of iterated summation. Duke math. J. **14**, 465 to 473 (1947).

— On Romanov's device of orthogonalization. Sci. Rep. Nat. Tsing Hua Univ. **5**, 1—12 (1948).

— Note on an abstract inversion principle. Proc. Edinburgh math. Soc. (2) **9**, 71—73 (1954).

JACKSON, F. H.: Series connected with the enumeration of partitions. Proc. London math. Soc., II. Ser. **1**, 63—88 (1904).

— The q-form of Taylor's theorem. Messenger of Mathematics **38**, 57—61 (1909).

JÓNSSON, B.: Lattice-theoretic approach to projective and affine geometry. Symposium on the Axiomatic Method. Amsterdam, North-Holland Publishing Company, 1959, 188—205.

—, and A. TARSKI: Direct decomposition of finite algebraic systems. Notre Dame Mathematical lectures, no. 5. Indiana: Notre Dame 1947.

KAC, M., and J. C. WARD: A combinatorial solution of the two-dimensional Ising model. Phys. Review **88**, 1332—1337 (1952).

KAPLANSKI, I., and J. RIORDAN: The problème des ménages. Scripta math. **12**, 113—124 (1946).

KLEE, V.: The Euler characteristic in combinatorial geometry. Amer. math. Monthly **70**, 119—127 (1963).

LAZARSON, T.: The representation problem for independence functions. J. London math. Soc. **33**, 21–25 (1958).
MACLANE, S.: A lattice formulation of transcendence degrees and p-bases. Duke math. J. **4**, 455–468 (1938).
MACMILLAN, B.: Absolutely monotone functions. Ann. of Math., II. Ser. **60**, 467–501 (1954).
MÖBIUS, A. F.: Über eine besondere Art von Umkehrung der Reihen. J. reine angew. Math. **9**, 105–123 (1832).
ORE, O.: Theory of graphs. Providence: American Mathematical Society 1962.
POLYA, G.: Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. Acta math. **68**, 145–253 (1937).
RADO, R.: Note on independence functions. Proc. London math. Soc., III. Ser. **7**, 300–320 (1957).
READ, R. C.: The enumeration of locally restricted graphs, I. J. London math. Soc. **34**, 417 to 436 (1959).
REDFIELD, J. H.: The theory of group-reduced distributions. Amer. J. Math. **49**, 433–455 (1927).
REVUZ, ANDRÉ: Fonctions croissantes et mesures sur les espaces topologiques ordonnés. Ann. Inst. Fourier **6** 187–268 (1955).
RIORDAN, J.: An introduction to combinatorial analysis. New York: Wiley 1958.
ROMANOV, N. P.: On a special orthonormal system and its connection with the theory of primes. Math. Sbornik, N. S. **16**, 353–364 (1945).
ROTA, G.-C.: Combinatorial theory and Möbius functions. To appear in Amer. math. Monthly.
— The number of partitions of a set. To appear in Amer. math. Monthly.
RYSER, H. J.: Combinatorial Mathematics. Buffalo: Mathematical Association of America 1963.
SCHÜTZENBERGER, M. P.: Contribution aux applications statistiques de la théorie de l'information. Publ. Inst. Stat. Univ. Paris, **3**, 5–117 (1954).
TARSKI, A.: Ordinal algebras. Amsterdam: North-Holland Publishing Company 1956.
TOUCHARD, J.: Sur un problème de permutations. C. r. Acad. Sci., Paris, **198**, 631–633 (1934).
TUTTE, W. T.: A contribution to the theory of chromatic polynomials. Canadian J. Math. **6**, 80–91 (1953).
A class of Abelian group. Canadian J. Math. **8**, 13–28 (1956).
A homotopy theorem for matroids. I. and II. Trans. Amer. math. Soc. **88**, 144–140 (1958).
Matroids and graphs. Trans. Amer. math. Soc. **90**, 527–552 (1959).
WARD, M.: The algebra of lattice functions. Duke math. J. **5**, 357–371 (1939).
WEISNER, L.: Abstract theory of inversion of finite series. Trans. Amer. math. Soc. **38**, 474–484 (1935).
— Some properties of prime-power groups. Trans. Amer. math. Soc. **38**, 485–492 (1935).
WHITNEY, H.: A logical expansion in mathematics. Bull. Amer. math. Soc. **38**, 572–579 (1932).
— Characteristic functions and the algebra of logic. Ann. of Math., II. Ser. **34**, 405–414 (1933).
— The abstract properties of linear dependence. Amer. J. Math. **57**, 507–533 (1935).
WIELANDT, H.: Beziehungen zwischen den Fixpunktzahlen von Automorphismengruppen einer endlichen Gruppe. Math. Z. **73**, 146–158 (1960).
WINTNER, A.: Eratosthenian Averages. Baltimore (privately printed) 1943.

Department of Mathematics
Massachusetts Institute of Technology
Cambridge 39, Massachusetts

(Received September 2, 1963)

Reprinted by TRUEXpress Oxford England

PATHS, TREES, AND FLOWERS

JACK EDMONDS

1. Introduction. A *graph G* for purposes here is a finite set of elements called *vertices* and a finite set of elements called *edges* such that each edge *meets* exactly two vertices, called the *end-points* of the edge. An edge is said to *join* its end-points.

A *matching* in G is a subset of its edges such that no two meet the same vertex. We describe an efficient algorithm for finding in a given graph a matching of maximum cardinality. This problem was posed and partly solved by C. Berge; see Sections 3.7 and 3.8.

Maximum matching is an aspect of a topic, treated in books on graph theory, which has developed during the last 75 years through the work of about a dozen authors. In particular, W. T. Tutte **(8)** characterized graphs which do not contain a *perfect* matching, or *1-factor* as he calls it—that is a set of edges with exactly one member meeting each vertex. His theorem prompted attempts at finding an efficient construction for perfect matchings.

This and our two subsequent papers will be closely related to other work on the topic. Most of the known theorems follow nicely from our treatment, though for the most part they are not treated explicitly. Our treatment is independent and so no background reading is necessary.

Section 2 is a philosophical digression on the meaning of "efficient algorithm." Section 3 discusses ideas of Berge, Norman, and Rabin with a new proof of Berge's theorem. Section 4 presents the bulk of the matching algorithm. Section 7 discusses some refinements of it.

There is an extensive combinatorial-linear theory related on the one hand to matchings in bipartite graphs and on the other hand to linear programming. It is surveyed, from different viewpoints, by Ford and Fulkerson in **(5)** and by A. J. Hoffman in **(6)**. They mention the problem of extending this relationship to non-bipartite graphs. Section 5 does this, or at least begins to do it. There, the König theorem is generalized to a matching-duality theorem for arbitrary graphs. This theorem immediately suggests a polyhedron which in a subsequent paper **(4)** is shown to be the convex hull of the vectors associated with the matchings in a graph.

Maximum matching in non-bipartite graphs is at present unusual among combinatorial extremum problems in that it is very tractable and yet not of the "unimodular" type described in **(5** and **6)**.

Received November 22, 1963. Supported by the O.N.R. Logistics Project at Princeton University and the A.R.O.D. Combinatorial Mathematics Project at N.B.S.