*Multiplication Operators on Incidence Algebras**

DAVID A. SMITH

Communicated by G.-C. ROTA

1. Introduction. If $\mathcal{A}$ is an algebra over a field K , (left or right) multiplication by an element of $\mathcal{A}$ is a linear operator on $\mathcal{A}$. This assertion is completely trivial, of course, but there are situations in which it is convenient, even interesting, to consider multiplication by a fixed element as an operator.

For example, if $\mathcal{A}$ is the incidence algebra of a locally finite partially ordered set P , then multiplications by ζ (the characteristic function of the ordering) and μ (the Möbius function) are, respectively, summation and differencing operators, and the Möbius Inversion Formula is the “fundamental theorem of (difference) calculus” on P .

In this paper we consider multiplication operators on an incidence algebra induced in a natural way by any subset of the partially ordered set, by a closure operator on the set, by a Galois connection (residuated mapping) between two partially ordered sets, and by an increasing function from one set to another.

The principal results concerning these operators, while basically very simple, provide simplifications and generalizations of some results of Rota and Crapo on Möbius inversion. The fundamental idea of studying Möbius inversion across a Galois connection is due to Rota [8], and the present work in part continues the refinement of this idea by Crapo [4]. There are also some generalizations of results on arithmetic functions, and in this sense this paper is a sequel to the author’s series of papers on that subject [12, 13, 14]. Some of the results have combinatorial interest, as indicated briefly in §4.

Most of the operators considered in this paper are multiplications by idempotent elements, and hence are Baxter operators. In an Appendix, we give several necessary and sufficient conditions for an averaging, Baxter operator on an associative algebra with 1 to be a multiplication by an element of the algebra.

2. Multiplication operators induced by a subset. Let $(P, \leq)$ be a locally finite partially ordered set and $\mathcal{A} = \mathcal{A}(P)$ its incidence algebra over a field K .

* Research supported in part by National Science Foundation grant GP-8902.

Let Q be a non-empty subset of P , with the induced order, and consider $\alpha_Q = \alpha(Q)$ as a subalgebra of α , i.e. imbedded in α as in [13, Theorem 5].

Proposition 1. α_Q has a faithful representation as an algebra of linear transformations on α .

Indeed, this is true for any subalgebra of an algebra with identity. The representation is by either left or right multiplications, depending on which way one chooses to define composition of mappings. A left (respectively, right) multiplication is also a left (right) averaging operator (see Appendix), so we may describe the representation as by averaging operators on A .

Let $\alpha_1(P, Q) = \{f \in \alpha \mid f(x, y) = 0 \text{ unless } y \in Q\}$. Where P and Q are fixed, we will write simply α_1 . α_1 is clearly a subalgebra, and $\alpha_Q \subseteq \alpha_1$. For $f \in \alpha, g \in \alpha_Q$, we have $fg \in \alpha_1$. On the other hand, the Kronecker delta function δ_Q on Q (the identity element of α_Q) is a right identity for $\alpha_1 : f\delta_Q = f$ for $f \in \alpha_1$. Hence

$$(1) \quad \alpha \alpha_Q = \alpha_1.$$

Similarly, if we let $\alpha_2 = \alpha_2(P, Q) = \{f \in \alpha \mid f(x, y) = 0 \text{ unless } x \in Q\}$, then α_2 is a subalgebra containing α_Q and

$$(2) \quad \alpha_Q \alpha = \alpha_2.$$

Furthermore, $\alpha_Q = \alpha_1 \cap \alpha_2$.

If $Q \neq P$, no element of α_1 or α_2 can be invertible in α . However, partial inversion formulas using elements of these subalgebras can be obtained in the following way: Let $\mu_1 = \mu\delta_Q$, $\mu_2 = \delta_Q\mu$, $\bar{\mu}_1 = \mu\zeta_Q$, and $\bar{\mu}_2 = \zeta_Q\mu$, where μ is the Möbius function on P and ζ_Q is the zeta function on Q . Then $\mu_i, \bar{\mu}_i \in \alpha_i, i = 1, 2$.

Proposition 2. Let $f \in \alpha$, and let

$$F(x, y) = f\zeta(x, y) = \sum_{x \leq z \leq y} f(x, z),$$

$$G(x, y) = \zeta f(x, y) = \sum_{x \leq z \leq y} f(z, y),$$

for all $x, y \in P$. Then

$$(3) \quad \sum_{x \leq z \leq y} F(x, z)\mu_1(z, y) = f(x, y) \text{ if } y \in Q, 0 \text{ otherwise};$$

$$(4) \quad \sum_{x \leq z \leq y} \mu_2(x, z)G(z, y) = f(x, y) \text{ if } x \in Q, 0 \text{ otherwise};$$

$$(5) \quad \sum_{x \leq z \leq y} F(x, z)\bar{\mu}_1(z, y) = \sum_{\substack{x \leq z \leq y \\ z \in Q}} f(x, z) \text{ if } y \in Q, 0 \text{ otherwise};$$

$$(6) \quad \sum_{x \leq z \leq y} \bar{\mu}_2(x, z)G(z, y) = \sum_{\substack{x \leq z \leq y \\ z \in Q}} f(z, y) \text{ if } x \in Q, 0 \text{ otherwise}.$$

Proof. $F = f\zeta$ implies $F\mu_1 = f\zeta\mu\delta_Q = f\delta_Q$, which is formula (3). $F\bar{\mu}_1 = f\zeta\mu\zeta_Q = f\zeta_Q$, which is (5). The other two statements are similar.

The significance of the functions μ_i is that only values of μ on intervals with an endpoint in Q are needed for their evaluation, with (3) and (4) yielding corresponding information about f . The functions $\bar{\mu}_i$ are suggested by a generalization of the ordinary Möbius function due to H. Gupta [6] (see §4 below).

3. A multiplication operator induced by a closure operator. Suppose $x \rightarrow \bar{x}$ is a closure operator on P , and let Q be the set of closed elements. We define an operator B on $\mathcal{A}$ by

$$(7) \quad Bf(x, y) = \sum_{\{z \mid \bar{z} = y\}} f(x, z),$$

for all $f \in \mathcal{A}$, $x, y \in P$. The right-hand member of (7) may be rewritten as

$$\sum_{x \leq z \leq y} f(x, z) \delta(\bar{z}, y) = f\bar{\delta}(x, y),$$

where $\bar{\delta}(x, y) = \delta(\bar{x}, y)$, so B is a right multiplication by an element of $\mathcal{A}$.

Theorem 1. $Bf = f\zeta\mu_Q$, for all $f \in \mathcal{A}$, where μ_Q is the Möbius function on Q .

Proof. It suffices to show that $\bar{\delta} = \zeta\mu_Q$. We have, for $x, y \in P$,

$$\begin{aligned} \delta(\bar{x}, y) &= \delta_Q(\bar{x}, y) = \zeta_Q\mu_Q(x, y) \\ &= \sum_{\substack{\bar{x} \leq z \leq y \\ z \text{ closed}}} \mu_Q(z, y) \\ &= \sum_{x \leq z \leq y} \mu_Q(z, y) \\ &= \zeta\mu_Q(x, y), \end{aligned}$$

since non-zero terms in the sums can occur only for z closed, in which case $z \geq x$ if and only if $z \geq \bar{x}$.

Corollary 1. (Rota) $B\mu = \mu_Q$, i.e.

$$\sum_{\{z \mid \bar{z} = y\}} \mu(x, z) = \begin{cases} \mu_Q(x, y), & \text{if } x, y \in Q, \\ 0, & \text{otherwise.} \end{cases}$$

Corollary 2. For all $f \in \mathcal{A}$, $B(f\mu) = f\mu_Q$, i.e.

$$\sum_{x \leq t \leq z \leq \bar{z} = y} f(x, t) \mu(t, z) = \sum_{x \leq v \leq y} f(x, v) \mu_Q(v, y).$$

Another useful consequence of Theorem 1 is obtained by taking $f = \zeta$. $B\zeta(x, y)$ is clearly the number of z in $[x, y]$ such that $\bar{z} = y$. (Statements about the “number” of elements in some set being some element of the base field K are to be interpreted modulo the characteristic of K .) Let $\tau = \zeta^2$; $\tau(x, y)$ is the number of elements in $[x, y]$.

Corollary 3. $B\zeta = \tau\mu_Q$, i.e. the number of elements in $[x, y]$ with closure y is given by

$$(8) \quad \sum_{\substack{x \leq z \leq y \\ z \in Q}} \tau(x, z) \mu_Q(z, y).$$

Theorem 1 allows us to identify several other images under B which are of interest. For example, for two of the functions introduced in Section 2, we have:

Corollary 4. $B\mu_2 = \mu_Q$, $B\bar{\mu}_2 = \delta_Q$.

Suppose $H: P \rightarrow \mathbf{N}^*$ is an arbitrary increasing mapping of P into the natural numbers, ordered by divisibility. As in [12, §6], we may define power functions, for each integer $k \geq 0$:

$$\nu_{k,H}(x, y) = H(y)^k / H(x)^k;$$

“sum-of-divisors” functions:

$$(9) \quad \sigma_{k,H} = \nu_{k,H} \zeta;$$

and “Jordan totient” functions;

$$(10) \quad \phi_{k,H} = \nu_{k,H} \mu.$$

(In the last, we have reversed the order of the factors from that given in [12]; the choice is clearly arbitrary.)

By replacing ζ and μ by ζ_Q and μ_Q we obtain the following analogous functions (which are not the same as those given by (9) and (10) by restricting H to Q):

$$(11) \quad \sigma_{k,Q} = \nu_{k,Q} \zeta_Q;$$

$$(12) \quad \phi_{k,Q} = \nu_{k,Q} \mu_Q.$$

(We have suppressed H in the notation here, and will continue to do so below.)

Corollary 5. $B\phi_k = \phi_{k,Q}$.

Corollary 6. $B\sigma_k = \nu_k \tau \mu_Q$; $B(\nu_k \tau_Q \mu) = \sigma_{k,Q}$.

In the statement of Corollary 6, $\tau_Q = \zeta_Q^2$. If ν_k and τ_Q commute, the last statement takes the form $B(\tau_Q \phi_k) = \sigma_{k,Q}$.

Some examples of these formulas will be given below. In particular, we shall note that the generalized Euler function given by Klee [7] is a special case of (12) for $k = 1$.

Several of the formulas listed so far have concerned ζ , ζ_Q , and their squares. Higher powers are also of interest: $\zeta^k(x, y)$ is the number of chains of length $\leq k$ from x to y , and a similar statement holds for ζ_Q^k .

Corollary 7. $B\zeta^k = \zeta^{k+1} \mu_Q$; $B(\zeta_Q^k \mu) = \zeta_Q^{k-1}$, for all k .

Theorem 2. For arbitrary $x, y \in P, f \in \mathfrak{A}$,

$$(13) \quad (Bf)\zeta(x, y) = f\zeta(x, y) - \sum_{\substack{x \leq w \\ z \leq w}} f(x, z)$$

$$(14) \quad = (Bf)\zeta_Q(x, y) - \sum_{\substack{w = \bar{w} \leq y \\ w \in Q}} Bf(x, w).$$

Proof. We have

$$\begin{aligned} (Bf)\zeta(x, y) &= \sum_{\substack{x \leq t \leq u \leq y \\ t = u}} f(x, t) = \sum_{\substack{x \leq t \leq y \\ t \leq y}} f(x, t) \\ &= \sum_{x \leq t \leq y} f(x, t) - \sum_{\substack{x \leq z \leq y \\ z \leq y}} f(x, z), \end{aligned}$$

which proves (13). Also, since $Bf \in \mathcal{Q}_1$,

$$\begin{aligned} (Bf)\zeta(x, y) &= \sum_{x \leq u = u \leq y} Bf(x, u) \\ &= \sum_{x \leq u = u \leq \bar{y}} Bf(x, u) - \sum_{\substack{x \leq w = \bar{y} \leq y \\ w \leq y}} Bf(x, w), \end{aligned}$$

which proves (14).

Corollary 1. For all $x \in P, y \in Q, f \in \mathcal{Q}$,

$$(15) \quad f\zeta(x, y) = (Bf)\zeta(x, y) = (Bf)\zeta_Q(x, y).$$

By setting $f = \mu$ in Theorem 2, we obtain a result complementary to the assertion $\zeta\mu_Q = \bar{\delta}$ of Theorem 1.

Corollary 2. For all $x, y \in P$,

$$(16) \quad \mu_Q\zeta(x, y) = \delta(x, y) - \sum_{\substack{z \leq y \\ z \leq y}} \mu(x, z)$$

$$(17) \quad = \delta(x, \bar{y}) - \sum_{\substack{w \leq \bar{y} \\ w \leq y}} \mu_Q(x, w).$$

In particular, if $x = \bar{x} < y < \bar{y}$, then

$$(18) \quad \sum_{\substack{z \leq y \\ z \leq y}} \mu(x, z) = \sum_{\substack{w \leq \bar{y} \\ w \leq y}} \mu_Q(x, w);$$

and if $x < \bar{x}, x < y < \bar{y}$, then

$$(19) \quad \sum_{\substack{z \leq y \\ z \leq y}} \mu(x, z) = 0.$$

Substitution of $f = \zeta$ in Theorem 2 produces a formula relating the number of $t \in [x, y]$ such that $\bar{t} \in [x, y]$ to $\tau(x, \bar{y})$. Another interesting formula results from setting $f = \bar{\mu}_2 = \zeta_Q\mu$. (Recall that $B\bar{\mu}_2 = \delta_Q$).

Corollary 3. For all $x, y \in P$,

$$(20) \quad \zeta_Q(x, y) - \sum_{\substack{z \leq y \\ z \leq y}} \bar{\mu}_2(x, z) = \delta(x, \bar{x}).$$

In particular, if x is closed and y is not, we have

$$(21) \quad \sum_{\substack{z \leq y \\ z \leq y}} \bar{\mu}_2(x, z) = -1.$$

Still other identities may be obtained by combining the results of Corollaries 5, 6, and 7 of Theorem 1 with Theorem 2.

Finally, we note some special properties of the operator B . It is easy to verify that $\bar{\delta}$ is an idempotent element of $\mathcal{A}$. Thus B is a projection onto its image, which is just $\mathcal{A}_1$. Any right multiplication is a right averaging operator, and multiplication by an idempotent is easily seen to be a Baxter operator (see Appendix). We already know that $\mathcal{A}_1$ is a subalgebra, and it follows from Theorem 1 that the complementary subspace $\ker B$ is also a subalgebra (cf. Atkinson [1, Theorem 7]). These subalgebras are not, in general, ideals, so $\mathcal{A}$ is not the algebra direct sum of $\mathcal{A}_1$ and $\ker B$.

It would be natural to generalize the operator B by considering, for arbitrary $f, g \in \mathcal{A}$, sums of the form

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

This is just multiplication of f by the function $\bar{g}$ defined by $\bar{g}(x, y) = \delta(\bar{x}, y)g(x, y)$. One may verify that $\bar{g}$ is idempotent if and only if $g(x, x) = 1$ for all x . After the case of $g = \zeta$ considered above, the most interesting case would be that of $g = \mu$. Essentially nothing is known about the corresponding Baxter operator, but [3, Theorem 1] is a theorem about a closely related operator which we shall define and study in §6 below.

4. Examples and applications. We consider at this point two specific examples of closure operators and the corresponding Baxter operators.

First consider the positive integers $\mathbf{N}^*$, ordered by divisibility. Let r be a fixed element of $\mathbf{N}^*$, and for each $x \in \mathbf{N}^*$, let $\bar{x}$ = the smallest r -th power multiple of x . Clearly, $x \rightarrow \bar{x}$ is a closure operator, and the set Q of closed elements is the set of r -th powers. Note also that the mapping $x \rightarrow x^r$ is an isomorphism of $(\mathbf{N}^*, |)$ onto $(Q, |)$. Thus $\mu_Q(x^r, y^r) = \mu(x, y)$.

For convenience in stating results, we shall restrict our attention to the reduced incidence algebra of functions f such that $f(x, y) = f(z, w)$ whenever $y/x = w/z$. If we write $f(y/x)$ for $f(x, y)$, this subalgebra becomes the ordinary algebra of arithmetic functions with the Dirichlet product.

Consider the function $\bar{\mu}_2$ of Section 2: $\bar{\mu}_2\zeta(1, m) = \zeta_Q(1, m) = 1$ or 0 according as m is or is not an r -th power. Thus $\bar{\mu}_2(1, m) = v_r(m)$, where v_r is Gupta's generalized Möbius function [6]. Formula (6) of Proposition 2 becomes, in this context, Theorem 3 of [6]. From Corollary 4 to Theorem 1, we have, for arbitrary $m \in \mathbf{N}^*$,

$$(22) \quad \sum_{\bar{k}=m^r} v_r(k) = 1.$$

From Corollary 3 to Theorem 2, if m is not an r -th power,

$$(23) \quad \sum_{k|m, \bar{k} \nmid m} v_r(k) = -1.$$

The number of elements $k \in \mathbf{N}^*$ such that $\bar{k} = m^r$ can be counted directly:

it is $r^{\lambda(m)}$, where $\lambda(m)$ = the number of distinct prime divisors of m . By Corollary 3 to Theorem 1,

$$(24) \quad r^{\lambda(m)} = \sum_{d|m} \tau(d^r) \mu\left(\frac{m}{d}\right),$$

where $\tau(k)$ = the number of divisors of k . By (ordinary) Möbius inversion, (24) is equivalent to

$$(25) \quad \tau(m^r) = \sum_{d|m} r^{\lambda(d)}.$$

Neither of these formulas is obvious. They are easy to prove directly by making use of the factorability of the functions involved, but it is of interest to note that factorability considerations are not involved in the proof just given.

For the present example we may take the function H of §3 to be the identity function. Then ϕ_k is the Jordan totient function (in particular, $\phi = \phi_1$ is the Euler function), and $\sigma_k(m)$ is the sum of the k -th powers of divisors of m . The function $\phi_{1,q}$ defined by (12) for $k = 1$ is Klee's generalization of the Euler function [7] (in his notation, Φ_r). In general,

$$(26) \quad \phi_{k,q}(m^r) = \sum_{d|m} d^{rk} \mu\left(\frac{m}{d}\right).$$

Corollary 5 to Theorem 1 relates the functions defined by (26) to the Jordan functions:

$$(27) \quad \sum_{n=m^r} \phi_k(n) = \phi_{k,q}(m^r).$$

For the functions defined by (11), we have $\sigma_{k,q}(m^r)$ = the sum of the kr -th powers of divisors of m . The remark following Corollary 6 to Theorem 1 applies, and we have

$$(28) \quad \sum_{d^r|m} \tau(d) \phi_k(m/d^r) = \sigma_{k,q}(m).$$

(The right-hand side is 0 unless m is an r -th power.)

$\zeta^k(m)$ is the number of (ordered) factorizations of m into k factors (the function τ_k of Beumer [2]) and $\zeta_q^k(m)$ is the number of such factorizations with each factor an r -th power (the function $\tau_{k,r}$ of Sivaramakrishnan [11]). Corollary 7 to Theorem 1 gives some identities involving these functions.

Finally, formula (18) provides the following identity for the Möbius function:

$$(29) \quad \sum_{d|m, d \nmid m} \mu(d) = \sum_{d^r|m, d^r \nmid m} \mu(d).$$

For the second example, let V be a finite dimensional vector space over a finite field of q elements. Let $(S, \subseteq)$ be the Boolean algebra of subsets of V . For each $x \in S$, let $\bar{x}$ be the subspace of V spanned by x . It is easy to verify (and well-known) that $x \rightarrow \bar{x}$ is a closure operator on S , with image Q the lattice of subspaces of V .

Both μ and μ_Q are well-known [8]:

$$(30) \quad \mu(x, y) = (-1)^{|z|-|x|},$$

$$(31) \quad \mu_Q(x, y) = (-1)^k q^{\binom{k}{2}}, \quad k = \dim y - \dim x.$$

Corollary 3 to Theorem 1, evaluated for $x = 0$, counts the number of subsets which span y :

$$(32) \quad \sum_{z \subseteq y} 2^{|z|} \mu_Q(z, y),$$

where the sum is over subspaces z . Let $n = \dim y$. For a subspace z of dimension k , $|z| = q^k$. The number of subspaces of dimension k is given by the q -binomial (Gaussian) coefficient:

$$\begin{bmatrix} n \\ k \end{bmatrix} = \frac{(q^n - 1)(q^{n-1} - 1) \cdots (q^{n-k+1} - 1)}{(q^k - 1)(q^{k-1} - 1) \cdots (q - 1)}.$$

Collecting like terms in (32), and substituting from (31), we get the following result.

Proposition 3. *The number of spanning subsets of an n -dimensional vector space over a field of q elements is*

$$(33) \quad \sum_{k=0}^n (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} 2^{q^k} q^{\binom{n-k}{2}}.$$

Corollary 1 of Theorem 1, evaluated for $x = 0$ and y of dimension n , and using (30), yields the following formula:

$$(34) \quad \sum_{k=0}^{\infty} (-1)^k r_k = (-1)^n q^{\binom{n}{2}},$$

where r_k denotes the number of spanning subsets of k elements. Both of these formulas can be obtained directly by Möbius inversion on the appropriate lattices. Here the inversion has been concealed in the proof of Theorem 1: $\xi \mu_Q = \bar{\delta}$.

5. Operators induced by residuated mappings. Let $\sigma : P_1 \rightarrow P_2$ be a residuated mapping [5] between two partially ordered sets; i.e. σ is increasing, and there exists an increasing mapping $\sigma^+ : P_2 \rightarrow P_1$ such that $\sigma^+ \sigma$ is a closure operator on P_1 and $\sigma \sigma^+$ is a coclosure operator on P_2 (a closure operator on the dual P_2^* of P_2). Let Q_1 (respectively, Q_2) be the set of closed (coclosed) elements in $P_1(P_2)$. Then $Q_1 = \sigma^+(P_2)$, $Q_2 = \sigma(P_1)$, and the restrictions of σ , σ^+ are inverse isomorphisms between Q_1 and Q_2 . It is convenient to identify Q_1 and Q_2 as a common subset Q of P_1 and P_2 , so that $\mathfrak{A}(Q)$ is imbedded as a subalgebra in both $\mathfrak{A}(P_1)$ and $\mathfrak{A}(P_2)$. The closure operator $\sigma^+ \sigma$ induces a Baxter operator B_1 on $\mathfrak{A}(P_1)$ as in §3, whose image is $\mathfrak{A}_1(P_1)$. Dually, the coclosure operator $\sigma \sigma^+$ induces a Baxter operator B_2 on $\mathfrak{A}(P_2)$:

$$(35) \quad B_2 f(x, y) = \sum_{\{z \mid \sigma\sigma^+(z) = x\}} f(z, y),$$

and the image of B_2 is $\mathfrak{A}_2(P_2)$. Note that $\mathfrak{A}(Q) = \mathfrak{A}_1(P_1) \cap \mathfrak{A}_2(P_2)$.

It is convenient to extend the definition of multiplication of functions slightly: if $f \in \mathfrak{A}(P_1)$ and $g \in \mathfrak{A}(P_2)$, with either $f \in \mathfrak{A}_1(P_1)$ or $g \in \mathfrak{A}_2(P_2)$, then we define a product $fg \in \mathfrak{A}(Q)$ by

$$(36) \quad fg(a, b) = \sum_x f(a, x)g(x, b),$$

where x runs over P_1 or P_2 or Q , since in any case, one factor of each term in the sum is 0 unless $x \in Q$.

The mappings σ, σ^+ induce linear operators between $\mathfrak{A}(P_1)$ and $\mathfrak{A}(P_2)$ in the following way:

$$(37) \quad f^\sigma(x, y) = f(\sigma^+(x), \sigma^+(y)) \quad \text{for } f \in \mathfrak{A}(P_1), x, y \in P_2;$$

$$(38) \quad g^{\sigma^+}(z, w) = g(\sigma(z), \sigma(w)) \quad \text{for } g \in \mathfrak{A}(P_2), z, w \in P_1.$$

These mappings are not, in general, algebra homomorphisms. However, the restriction of each to $\mathfrak{A}(Q)$ is just the identity.

Lemma 1. *If $f \in \mathfrak{A}_1(P_1)$ and $g \in \mathfrak{A}(P_2)$, then fg^{σ^+} and fg agree on intervals in Q . Dually, if $f \in \mathfrak{A}(P_1)$ and $g \in \mathfrak{A}_2(P_2)$, then $f^\sigma g$ and fg agree on Q .*

Proof. Let $a, b \in Q, f \in \mathfrak{A}_1(P_1), g \in \mathfrak{A}(P_2)$. Then

$$\begin{aligned} fg^{\sigma^+}(a, b) &= \sum_{x \in P_1} f(a, x)g(\sigma(x), b) \\ &= \sum_{x \in Q} f(a, x)g(x, b) \\ &= fg(a, b), \end{aligned}$$

in the sense of (36).

Lemma 2. *Let $f \in \mathfrak{A}(P_1), g \in \mathfrak{A}(P_2)$. If $x \in P_1, b \in Q$, then $fg^{\sigma^+}(x, b) = (B_1 f)g^{\sigma^+}(x, b)$. Dually, if $a \in Q, y \in P_2$, then $f^\sigma g(a, y) = f^\sigma(B_2 g)(a, y)$.*

Proof. From the definition of B_1 , we have

$$(B_1 f)g^{\sigma^+}(x, b) = \sum_{x \leq y \leq z \leq b} f(x, y)\delta(\sigma(y), z)g(\sigma(z), b).$$

In each term, the middle factor is 0 unless $z \in Q$, so the sum reduces to

$$\sum_{x \leq y \leq b} f(x, y)g(\sigma(y), b) = fg^{\sigma^+}(x, b).$$

Theorem 3. *Let $f, h_1 \in \mathfrak{A}(P_1), g, h_2 \in \mathfrak{A}(P_2)$, and suppose $B_1 h_1 = B_2 h_2 \in \mathfrak{A}(Q)$. Then $fh_1 g^{\sigma^+}$ and $f^\sigma h_2 g$ agree on intervals in Q .*

Proof. Let $a, b \in Q$.

$$\begin{aligned}
fh_1g^{\sigma^+}(a, b) &= \sum_{a \leq x \leq b} f(a, x)h_1g^{\sigma^+}(x, b) \\
&= \sum_{x \in Q} f(a, x)(B_1h_1)g^{\sigma^+}(x, b), \text{ by Lemma 2,} \\
&= \sum_{x \in Q} f(a, x)(B_1h_1)g^{\sigma^+}(x, b), \text{ since } B_1h_1 \in \mathcal{Q}, \\
&= \sum_{x \in Q} f(a, x)(B_1h_1)g(x, b), \text{ by Lemma 1,} \\
&= f(B_1h_1)g(a, b) \\
&= f(B_2h_2)g(a, b).
\end{aligned}$$

$f^{\sigma}h_2g(a, b)$ is equal to the same thing for the same (dual) reasons.

Corollary 1. (Crapo [4, Theorem 3]) For $f \in \mathcal{Q}(P_1)$, $g \in \mathcal{Q}(P_2)$, $a, b \in Q$, we have

$$(39) \quad \sum_{x \in P_1} f\mu_{P_1}(a, x)g(\sigma(x), b) = \sum_{y \in P_2} f(a, \sigma^+(y))\mu_{P_2}g(y, b).$$

Proof. By Rota's Theorem (§3), we have $B_i\mu_{P_i} = \mu_Q$, $i = 1, 2$.

Corollary 2. (Crapo [4, Cor. 4 to Thm. 3]) For $a \in P_1$, $b \in P_2$,

$$(40) \quad \sum_{a \leq x \leq \sigma^+(b)} \mu_{P_1}(a, x)\delta(\sigma(x), b) = \sum_{\sigma(a) \leq y \leq b} \delta(a, \sigma^+(y))\mu_{P_2}(y, b).$$

Proof. Note that the left-hand member of (40) is clearly 0 unless b is closed, and it is 0 unless a is closed, by Rota's Theorem. Similarly, the right-hand member is 0 unless $a, b \in Q$, and (40) is then the special case of (39) with $f = \delta_{P_1}$, $g = \delta_{P_2}$.

We remark that Corollary 2 also follows directly from Rota's Theorem, since for $a, b \in Q$ it asserts:

$$(41) \quad B_1\mu_{P_1}(a, b) = \mu_Q(a, b) = B_2\mu_{P_2}(a, b).$$

It has been observed by Crapo [4] that the result " $B\mu = \mu_Q$ " is the basic result for the theory of Möbius inversion across a Galois connection, which he refined substantially. In [8], Rota stated two apparently different "main results;" both are consequences of (40) or (41).

Corollary 3. (Rota [8, Theorem 1 and Corollary]) If $[a, \sigma^+(b)]$ is an interval in P_1 in which the only element whose closure is $\sigma^+(b)$ is $\sigma^+(b)$ itself, then

$$\mu_{P_1}(a, \sigma^+(b)) = \sum_{\{y \mid \sigma^+(y) = \sigma^+(b)\}} \mu_{P_2}(y, b).$$

Proof. The hypothesis asserts that the left-hand member of (40) has only one term.

Corollary 4. (Rota [8, Theorem 2]) Let P_1 and P_2 each have a smallest element 0, and suppose σ is an increasing mapping of P_1 onto P_2 such that $\sigma^{-1}([0, b])$ is an interval $[0, x]$ in P_1 for each $b \in P_2$. Then σ is a residuated mapping with closed set $Q \cong P_2$ and

$$(42) \quad \sum_{\{x \mid \sigma(x) = b\}} \mu_{P_1}(0, x) = \mu_Q(0, b).$$

Proof. For each $b \in P_2$, if $\sigma^{-1}([0, b]) = [0, x]$, let $x = \sigma^+(b)$. Then $\sigma^+: P_2 \rightarrow P_1$ is a one-to-one increasing mapping with $\sigma^+\sigma(a) \geq a$ for all $a \in P_1$ and $\sigma\sigma^+$ the identity on P_2 . In this case, the right-hand side of (40) has only one term, and we have (42).

The form in which Rota stated this result included the extra hypothesis that $\sigma^{-1}(\{0\})$ has more than one element; in this case $0 \notin Q$ and the right-hand side of (42) is 0.

6. Operators induced by increasing mappings. Let P be a locally finite partially ordered set, Q an arbitrary partially ordered set, and $\sigma: P \rightarrow Q$ an increasing mapping. For each $y \in P$, let $P_y = \{x \in P \mid x \leq y\}$, and on each subset P_y define an operator $x \rightarrow \bar{x}$ by

$$\bar{x} = \begin{cases} x, & \text{if } \sigma(x) < \sigma(y), \\ y, & \text{if } \sigma(x) = \sigma(y). \end{cases}$$

One verifies easily that this is a closure operator on P_y . We may then define an operator B on $\mathcal{A} (= \mathcal{A}(P))$ by formula (7), interpreted to mean the value $Bf(x, y)$ is determined by the closure operator on P_y . That is,

$$(43) \quad Bf(x, y) = \sum_{\substack{x \leq z \leq y \\ \sigma(z) = \sigma(y)}} f(x, z) = \sum_{x \leq y \leq z} f(x, z) \delta(\sigma(z), \sigma(y)).$$

Thus B is a right multiplication by an element of $\mathcal{A}$, namely $\delta_\sigma(x, y) = \delta(\sigma(x), \sigma(y))$. This operator is analogous to that of §3 determined by a closure operator on P , but δ_σ is not, in general, idempotent (hence B is not a Baxter operator).

Proposition 4. *If K has characteristic 0, then δ_σ is idempotent only in the trivial case $\delta_\sigma = \delta$.*

Proof. $\delta_\sigma = \delta$ if and only if $x < y$ implies $\sigma(x) < \sigma(y)$. If $x < y$ and $\sigma(x) = \sigma(y)$, then

$$\delta_\sigma^2(x, y) = \sum_{x \leq z \leq y} \delta(\sigma(x), \sigma(z)) \delta(\sigma(z), \sigma(y)) = \tau(x, y) \geq 2,$$

while $\delta_\sigma(x, y) = 1$, so δ_σ is not idempotent.

For each $y \in P$, let μ_y be the Möbius function on the subset of closed elements in P_y . Define a function $\mu_\sigma \in \mathcal{A}$ as follows:

$$\mu_\sigma(x, y) = \begin{cases} \mu_y(x, y), & \text{if } \sigma(x) < \sigma(y), \\ \delta(x, y), & \text{if } \sigma(x) = \sigma(y). \end{cases}$$

(This is an extension of the function introduced by Crapo [3, p. 127].) The following theorem is analogous to Theorem 1, and the proof is similar.

Theorem 4. $\delta_\sigma = \zeta \mu_\sigma$, hence $Bf = f \zeta \mu_\sigma$ for all $f \in \mathcal{A}$.

Proof. If $\sigma(x) = \sigma(y)$, then $\zeta \mu_\sigma(x, y) = \zeta(x, y) = 1 = \delta_\sigma(x, y)$. If $x < y$ and $\sigma(x) < \sigma(y)$, then x and y are closed elements in P_y , and

$$\begin{aligned}\delta_\sigma(x, y) &= 0 = \sum_{\substack{x \leq z \leq y \\ z \text{ closed}}} \mu_\sigma(z, y) \\ &= \sum_{x \leq z \leq y} \mu_\sigma(z, y) = \zeta \mu_\sigma(x, y),\end{aligned}$$

since $\mu_\sigma(z, y) = \delta(z, y) = 0$ for z not closed, i.e. $z < y$, but $\sigma(z) = \sigma(y)$.

Corollary 1. (Crapo [3, formula (4)]) $B\mu = \mu_\sigma$.

Corollary 2. For all $f \in \mathfrak{A}$, $B(f\mu) = f\mu_\sigma$.

Corollary 3. $B\zeta = \tau\mu_\sigma$.

Note that $B\zeta$ counts the number of elements $z \in [x, y]$ such that $\sigma(z) = \sigma(y)$. Unlike the function $\tilde{\delta}$ of §3, δ_σ is invertible, since $\delta_\sigma(x, x) = 1$ for all $x \in P$.

Theorem 5. $\delta_\sigma^{-1}(x, y) = \mu(x, y)\delta(\sigma(x), \sigma(y))$ for all $x, y \in P$.

Proof. If $x \leq y$ and $\sigma(x) = \sigma(y)$, then

$$\sum_{x \leq z \leq y} \delta_\sigma(x, z)\mu(z, y)\delta_\sigma(z, y) = \sum_{x \leq z \leq y} \mu(z, y) = \delta(x, y).$$

If $x < y$ and $\sigma(x) < \sigma(y)$, then

$$\sum_{x \leq z \leq y} \delta_\sigma(x, z)\mu(z, y)\delta_\sigma(z, y) = 0 = \delta(x, y),$$

since, for each term in the sum, either $\delta_\sigma(x, z)$ or $\delta_\sigma(z, y)$ is 0. Thus the indicated function is a right inverse for δ_σ , hence is the inverse.

Let M be the operator on $\mathfrak{A}$ given by right multiplication by δ_σ^{-1} :

$$\begin{aligned}Mf(x, y) &= \sum_{x \leq z \leq y} f(x, z)\mu(z, y)\delta(\sigma(z), \sigma(y)) \\ &= \sum_{\substack{x \leq z \leq y \\ \sigma(z) = \sigma(y)}} f(x, z)\mu(z, y).\end{aligned}$$

This is the analog in this context of the Baxter operator corresponding to μ defined at the end of §3.

Corollary 1. $MB = BM = 1$, i.e.

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

Corollary 2. (Crapo [3, Theorems 1 and 2]).

$$\sum_{\sigma(z) = \sigma(w) = \sigma(y)} \mu(x, z)\zeta(z, w)\mu(w, y) = \sum_{\sigma(w) = \sigma(y)} \mu_\sigma(x, w)\mu(w, y) = \mu(x, y).$$

Corollary 3. μ_σ is invertible, and

$$\mu_\sigma^{-1}(x, y) = \delta_\sigma^{-1}\zeta(x, y) = \sum_{\substack{x \leq z \leq y \\ \sigma(z) = \sigma(x)}} \mu(x, z).$$

Appendix. Averaging and Baxter operators. Let $\mathcal{A}$ be an associative algebra with 1 over a field K and let T be a linear operator on $\mathcal{A}$. T is called a *right averaging* operator if, for all $x, y \in \mathcal{A}$,

$$(1) \quad (Tx)(Ty) = T[(Tx)y].$$

Similarly, *left averaging* is defined by the identity

$$(2) \quad (Tx)(Ty) = T[x(Ty)].$$

T is called a *Baxter operator* on $\mathcal{A}$ if, for all $x, y \in \mathcal{A}$,

$$(3) \quad (Tx)(Ty) = T[x(Ty) + (Tx)y - \theta xy],$$

where θ is some fixed element of K . If T satisfies (3) with $\theta \neq 0$, then $\theta^{-1}T$ satisfies (3) with $\theta = 1$. Hence, without loss of generality, one may consider only the cases $\theta = 0$ and $\theta = 1$. The $\theta = 0$ case is essentially that of antiderivations (see [1, p. 14]), and of no interest here, so we shall reserve the term Baxter operator for linear operators T on $\mathcal{A}$ satisfying the identity

$$(4) \quad (Tx)(Ty) + T(xy) = T[x(Ty) + (Tx)y].$$

There is a substantial literature dealing with averaging and Baxter operators, which arise naturally in probability theory. Most of the known results concern operators on commutative algebras. In particular, Rota [9, 10] has recently demonstrated the significance for combinatorial theory of Baxter operators on commutative algebras. [9] also includes a good bibliography on the subject.

An operator T on $\mathcal{A}$ which is both Baxter and right averaging also clearly satisfies the identity

$$(5) \quad T(xy) = T[x(Ty)],$$

and any two of the identities (1), (4), and (5) imply the other. The operators considered in the main body of this paper were all multiplications by elements of $\mathcal{A}$, hence certainly averaging operators. A right multiplication satisfies (5), hence is a Baxter operator, if and only if it is idempotent. The question we consider here is: When is a Baxter, right averaging operator T on an algebra $\mathcal{A}$ necessarily a multiplication by an element of $\mathcal{A}$?

Lemma 1. *If T is a Baxter operator, so is $I - T$.*

The proof is a straightforward verification that $I - T$ satisfies (4) when T does.

Lemma 2. *If T is Baxter and right averaging, then T is idempotent, hence T and $I - T$ are projections onto $\text{Im } T$ and $\text{Ker } T$.*

Proof. Set $x = 1$ in (5).

Lemma 3. *If T is an idempotent right averaging operator, then $T1$ is a right identity in $\text{Im } T$, and is idempotent.*

Proof. Note that (1) implies that $\text{Im } T$ is a subalgebra. Setting $y = 1$ in (1), we have $(Tx)(T1) = T^2x = Tx$. Setting $x = 1$, we have $(T1)^2 = T1$.

Remark. Atkinson has proved [1, Theorem 7] that an operator T on an algebra $\mathcal{A}$ is Baxter if and only if $\text{Im } T$ and $\text{Im}(I - T)$ are subalgebras and $\mathcal{A}$ is isomorphic to a subdirect sum of these two algebras. If T is idempotent, $\text{Im}(I - T) = \text{Ker } T$. Thus it follows from Atkinson's Theorem that if T is an idempotent Baxter operator, then $\text{Im } T$ and $\text{Ker } T$ are subalgebras, and $\mathcal{A}$ is their (vector space) direct sum.

Theorem. Let T be a Baxter, right averaging operator on $\mathcal{A}$. The following are equivalent:

- (a) T is the right multiplication by $T1$.
- (b) T is a right multiplication by some element of $\mathcal{A}$.
- (c) $T(xy) = x(Ty)$, for all $x, y \in \mathcal{A}$.
- (d) $T[x(Ty)] = x(Ty)$, for all $x, y \in \mathcal{A}$.
- (e) $x(T1) \in \text{Im } T$, for all $x \in \mathcal{A}$.
- (f) $x(T1) = 0$, for all $x \in \text{Ker } T$.
- (g) $I - T$ is right averaging.

Proof. Clearly (a) $\Rightarrow$ (b) $\Rightarrow$ (c). Suppose (c) is true. Then $T[x(Ty)] = T[T(xy)] = T(xy) = x(Ty)$, by Lemma 2. Hence (c) $\Rightarrow$ (d). Clearly (d) $\Rightarrow$ (e). Suppose (e) is true, and $x \in \text{Ker } T$. Then $T[x(T1)] = T(x) = 0$, by (5), so $x(T1) \in \text{Ker } T \cap \text{Im } T = \{0\}$, by Lemma 2. Hence (e) $\Rightarrow$ (f). Suppose (f) is true, and let x be an arbitrary element of $\mathcal{A}$. Then $x = u + v$, where $u = Tx \in \text{Im } T$ and $v = (I - T)x \in \text{Ker } T$. Hence $Tx = u = u(T1) + v(T1) = x(T1)$, by Lemma 3. Thus (f) $\Rightarrow$ (a).

Finally, by direct substitution of $I - T$ for T in (1), one finds that $I - T$ is right averaging if and only if

$$(6) \quad (Tx)(Ty) + T(xy) = T[(Tx)y] + x(Ty),$$

for all $x, y \in \mathcal{A}$. By (4), this is equivalent to (d), and the proof is complete.

If $\mathcal{A}$ is a commutative algebra, then the notions of left and right averaging coincide. Furthermore, it follows from the defining identities that a linear operator T on a commutative algebra is Baxter and averaging if and only if it is an idempotent algebra endomorphism. A simple example will demonstrate that such an operator need not be a multiplication by an element of the algebra: Let $\mathcal{A}$ be a 2-dimensional vector space with basis $\{e_1, e_2\}$, and define a product on $\mathcal{A}$ by $e_1^2 = e_1$, $e_1e_2 = e_2e_1 = e_2^2 = e_2$. Then $\mathcal{A}$ is a commutative, associative algebra with identity e_1 . The projection T of $\mathcal{A}$ onto the span of e_1 along e_2 is an idempotent endomorphism, and clearly not a multiplication by an element of $\mathcal{A}$.

REFERENCES

- [1] F. V. ATKINSON, Some aspects of Baxter's functional equation, *J. Math. Anal. Appl.*, **7** (1963) 1-30.
- [2] M. G. BEUMER, The arithmetic function $\tau_k(n)$, *Amer. Math. Monthly*, **69** (1962) 777-781.
- [3] H. H. CRAPO, The Möbius function of a lattice, *J. Comb. Theory* **1** (1966) 126-131.

- [4] H. H. CRAPO, Möbius inversion in lattices, *Archiv. der Math.*, **19** (1968) 595–607.
- [5] J.-C. DERDÉRIAN, Residuated mappings, *Pacific J. Math.*, **20** (1967) 35–43.
- [6] H. GUPTA, A generalization of the Möbius function, *Scripta Math.*, **19** (1953) 121–126.
- [7] V. L. KLEE, A generalization of Euler's φ — function, *Amer. Math. Monthly*, **55** (1948) 358–359.
- [8] G.-C. ROTA, On the foundations of combinatorial theory, I. Theory of Möbius functions, *Z. Wahrsch.*, **2** (1964) 340–368.
- [9] G.-C. ROTA, Baxter algebras and combinatorial identities, I, *Bull. Amer. Math. Soc.*, **75** (1969) 326–329.
- [10] G.-C. ROTA, Baxter algebras and combinatorial identities, II, *Bull. Amer. Math. Soc.*, **75** (1969) 330–334.
- [11] R. SIVARAMAKRISHNAN, The arithmetic function $\tau_{k,r}(n)$, *Amer. Math. Monthly*, **75** (1968) 988–989.
- [12] D. A. SMITH, Incidence functions as generalized arithmetic functions, I., *Duke Math. J.*, **34** (1967) 617–634.
- [13] D. A. SMITH, Incidence functions as generalized arithmetic functions, II., *Duke Math. J.*, **36** (1969) 15–30.
- [14] D. A. SMITH, Incidence functions as generalized arithmetic functions, III., *Duke Math. J.*, **36** (1969) 353–368.

Duke University

Date communicated: NOVEMBER 17, 1969