18.312: Algebraic Combinatorics	Lionel Levine
Lecture 10	
Lecture date: March 8, 2011	*Notes by: Alex Zhai*

1 Review of incidence algebras

Recall that for a poset P and field K, we have defined the incidence algebra $I(P)$ to be the set of functions mapping intervals of P to K. To simplify notation, for $\alpha \in I(P)$ and $x, y \in P$ with $x \leq y$, we will write $\alpha(x, y)$ for $\alpha([x, y])$. For $\alpha, \beta \in I(P)$, we defined the product $\alpha\beta$ by

$$(\alpha\beta)(x, y) = \sum_{x \leq z \leq y} \alpha(x, z)\beta(z, y).$$

This can be interpreted as a matrix multiplication. Associate to each $\alpha \in I(P)$ a $|P| \times |P|$ matrix $M(\alpha)$ with rows and columns indexed by elements of P, in which the xy entry is $\alpha(x, y)$ if $x \leq y$ and 0 otherwise. Then, it is a routine exercise to verify that $M(\alpha)M(\beta) = M(\alpha\beta)$.

From the matrix perspective, it is clear that the multiplication we have defined is associative with identity element $1(x, y) = \delta_{xy}$. Thus, the incidence algebra $I(P)$ has the structure of a K-algebra, justifying its appellation. Furthermore, we may naturally say that $\beta = \alpha^{-1}$ if $\alpha\beta = 1$ (or, equivalently, $\beta\alpha = 1$). In more explicit terms, this condition is

$$\sum_{x \leq z \leq y} \alpha(x, z)\beta(z, y) = \delta_{xy}$$

for all $x, y \in P$ with $x \leq y$. Rearranging the equation yields

$$\beta(x, y) = \frac{1}{\alpha(x, x)} \left(\delta_{xy} - \sum_{x < z \leq y} \alpha(x, z)\beta(z, y) \right),$$

which provides us a way of solving for the inverse of α so long as $\alpha(x, x) \neq 0$ for all $x \in P$. In particular, since $\beta(x, y)$ can be expressed in terms of $\beta(z, y)$ for $z > x$, we may start by solving for $\beta(y, y)$ and inductively solve for values $\beta(z, y)$ with $z < y$ until we reach $\beta(x, y)$. Note that $\beta(x, y)$ only depends on the values of $\alpha(x', y')$ for $x', y' \in [x, y]$.

Exercise 1 *We have shown that $\alpha(x,x) \neq 0$ is a sufficient condition for α to have an inverse. Observe that it is also necesary.*

2 Mobius inversion for posets

We now shift our attention to two particular elements of $I(P)$. Consider the element $\zeta \in I(P)$ given by $\zeta(x,y) = 1$ if $x \leq y$ and $\zeta(x,y) = 0$ otherwise. Since $\zeta(x,x) \neq 0$ for all $x \in P$, ζ has an inverse, which will be denoted by μ. As will be seen later, the suggestive naming of these elements is meant to draw a connection to the zeta and Mobius functions we have studied before in the context of Dirichlet series.

The first hint towards this connection is the Mobius inversion formula for posets.

Theorem 2 (Mobius inversion formula (posets)) *Let P be a poset and K a field, and let f and g be functions from P to K. If f satisfies*

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

for all $x \in P$, then g is given by

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

Conversely, if the second equation holds, then so does the first.

Proof: It is possible to verify this directly by substituting the equation for f into the equation for g that is to be proven. However, it is perhaps more insightful to take a different approach.

Just as we interpreted elements of $I(P)$ as $|P| \times |P|$ matrices, for a function $f : P \to K$, we may associate to f a row vector $M(f)$ with $|P|$ columns indexed by the elements of P, where the column x entry is $f(x)$. It is not hard to verify, then, that the condition

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

is equivalent to the equation $M(f) = M(g)M(\zeta)$. We can then multiply both sides on the right by $M(\mu)$, and because μ is the inverse of ζ, we obtain $M(f)M(\mu) = M(g)$. Writing this matrix equation out entry-by-entry, we find that

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

as desired. A similar argument proves the converse. □

The compactness of the equations in the preceding proof when we converted to matrices should not be taken as mere algebraic coincidence. It is sometimes fruitful to think of elements of $I(P)$ as acting on the $|P|$-dimensional vector space K^P of all functions from P to K. We have just seen the right action of $I(P)$ on K^P, giving K^P the structure of a right $I(P)$-module.

In fact, it is also possible to give K^P the structure of a left $I(P)$-module by expressing the elements of K^P as column vectors instead of row vectors. This leads us to another version of the Mobius inversion formula.

Theorem 3 (Mobius inversion formula (dual version)) *Let P be a poset and K a field, and let f and g be functions from P to K. If f satisfies*

$$f(x) = \sum_{y \geq x} g(y)$$

for all $x \in P$, then g is given by

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

Conversely, if the second equation holds, then so does the first.

Proof: The proof follows along the same lines as the proof of Theorem 2. Alternatively, setting $Q = P^*$, this follows from Mobius inversion on Q upon noting that $\mu_Q(x, y) = \mu_P(y, x)$. The proof of this last statement is left as an exercise for the reader. □

Example 4 *Let us apply Mobius inversion to the case $P = \mathbf{n}$. In the previous section, we outlined a procedure by which the inverse of an element $\alpha \in I(P)$ may be computed. Applying this to $\alpha = \zeta$, we find that $\mu(x, x) = 1$, $\mu(x, x+1) = -1$, and $\mu(x, y) = 0$ for $y > x + 1$.*

Plugging this into the Mobius inversion formula yields

$$f(k) = g(1) + \ldots + g(k) \iff g(k) = \begin{cases} f(k) - f(k-1) & \text{if } k \geq 2 \\ f(1) & \text{if } k = 1 \end{cases},$$

which expresses a familiar relationship between a series and its partial sums.

Exercise 5 *In the above example, compute μ instead by computing the matrix inverse of $M(\zeta)$. (It may be helpful to think of $M(\zeta)$ as a change of basis for the vector space of polynomials of degree less than n.)*

It was fairly easy to compute μ in the last example, but for more complicated posets, it may not be so simple. However, in the case of products of posets, the following lemma can make the computation more tractable.

Lemma 6 *Let P and Q be posets, and consider any elements $x, x' \in P$ and $y, y' \in Q$, where $x \leq x'$ and $y \leq y'$. Then,*

$$\mu_{P\times Q}\left((x,y),(x',y')\right) = \mu_P(x,x')\mu_Q(y,y'),$$

where the subscript on μ indicates which poset is being used to define the Mobius function.

Proof: Let $\mu'\left((x,y),(x',y')\right) = \mu_P(x,x')\mu_Q(y,y')$. We may compute

$$\begin{aligned}
&\sum_{(x,y)\leq(u,v)\leq(x',y')} \mu'\left((x,y),(u,v)\right) \\
&= \sum_{(x,y)\leq(u,v)\leq(x',y')} \mu_P(x,u)\mu_Q(y,v) \\
&= \left(\sum_{x\leq u\leq x'} \mu_P(x,u)\right)\left(\sum_{y\leq v\leq y'} \mu_Q(y,v)\right) \\
&= \delta_{xx'}\delta_{yy'} = \delta_{(x,y)(x',y')}.
\end{aligned}$$

This equation says exactly that μ' is the inverse of $\zeta_{P\times Q}$, so $\mu_{P\times Q} = \mu'$. □

It may be somewhat surprising that the above calculation worked out so conveniently. Whenever calculations work out nicely in mathematics, there is some hope that a more abstract underlying theory is at work. Such is the case here, and we shall take a digression to develop briefly the abstract viewpoint. (The uninterested reader may skip ahead.)

3 A digression on tensor products

Before we define the tensor product of two vector spaces, it will be useful to review the concept of a dual vector space. If V is a K-vector space, then we define V^* to be the space of all linear maps from V to K, and we call V^* the *dual* of V. In the case that V is finite-dimensional, it is not hard to check that V^* has the same dimension as V.

Since all vector spaces of the same finite dimension are isomorphic, defining V^* in this way may not seem like a particularly useful notion. However, an isomorphism between two arbitrary vector spaces of the same finite dimension is not necessarily *natural*: it requires picking a basis for each vector space, and a different choice of bases will result in a different isomorphism. If the definition of tensor product depended on a choice of basis, then it would be nothing more than a computational tool. Instead, we shall see that the tensor product can be defined in an intrinsic way; we may then specialize to particular bases to reap the insight gained from the abstract viewpoint.

Suppose now that W is another vector space, and T is a linear map from V to W. Then, we can associate to T a linear map $T^* : W^* \to V^*$ defined by $T^*(f)(v) = f(T(v))$ for $f \in W^*$ and $v \in V$. The reader should think of T and T^* as two sides of the same coin. If S is a map from another vector space U to V, then it is not hard to check that $(ST)^* = T^*S^*$.

Given vector spaces V and W, define the *tensor product* $V \otimes W$ to be the space of all bilinear forms on $V^* \times W^*$. If $S : V \to V$ and $T : W \to W$ are linear maps, we can define a linear map $S \otimes T : V \otimes W \to V \otimes W$ as follows: for $B \in V \otimes W$, we define

$$((S \otimes T)B))(f, g) = B(S^*f, T^*g)$$

for all $f \in V^*$ and $g \in W^*$. The use of dual spaces here seems a bit onerous, but there is a good reason for it. Suppose that we have two more maps $S' : V \to V$ and $T' : W \to W$. It is not hard to check from the definitions that

$$(S' \otimes T') \circ (S \otimes T) = (S'S \otimes T'T).$$

If we had instead defined $V \otimes W$ as the space of bilinear forms on $V \times W$ and proceeded in an analogous way, this composition law would be reversed.

Now that we have established some basic properties of the tensor product, it is useful to see what it looks like in coordinates. Let $v_1, \ldots, v_n$ be a basis of V, and let $w_1, \ldots, w_m$ be a basis of W. Let $v_1^*, \ldots, v_n^*$ be the corresponding basis for V^*, where $v_i^*(v_j) = 1$ if $j = i$ and $v_i^*(v_j) = 0$ otherwise. Define $w_1^*, \ldots, w_m^*$ in an analogous manner, and let $B_{ij} \in V \otimes W$ be the bilinear form given by $B_{ij}(v_{i'}^*, w_{j'}^*) = 1$ if $(i, j) = (i', j')$ and $B_{ij}(v_{i'}^*, w_{j'}^*) = 0$ otherwise.

The B_{ij} form a basis of $V \otimes W$, and we will analyze $S \otimes T$ in the B_{ij} coordinates. Note that

$$\begin{aligned} ((S \otimes T)B_{ij})(v_{i'}^*, w_{j'}^*) &= B_{ij}(S^*(v_{i'}^*), T^*(w_{j'}^*)) \\ &= (S^*(v_{i'}^*)(v_i)) \left(T^*(w_{j'}^*)(w_j)\right) \\ &= v_{i'}^*(Sv_i) w_{j'}^*(Tw_j) = S_{i'i}T_{j'j}. \end{aligned}$$

The first expression is the $B_{i'j'}$ component of $(S \otimes T)B_{ij}$, so we conclude that the matrix entries of $S \otimes T$ are products of matrix entries of S and T; in particular, $(S \otimes T)_{(i',j')(i,j)} = S_{i'i}T_{j'j}$.

We can apply the theory developed thus far to the case of the ζ and μ functions on product posets. For posets P and Q, it is trivial to check that

$$\zeta_{P \times Q}\left((x, y), (x', y')\right) = \zeta_P(x, x')\zeta_Q(y, y').$$

Thus, $M(\zeta_{P\times Q}) = M(\zeta_P) \otimes M(\zeta_Q)$. From the composition properties we have shown for tensor products of maps, we can invert this equation to obtain

$$\begin{aligned} M(\zeta_{P\times Q})^{-1} &= M(\zeta_P)^{-1} \otimes M(\zeta_Q)^{-1} \\ M(\mu_{P\times Q}) &= M(\mu_P) \otimes M(\mu_Q). \end{aligned}$$

In coordinates, this is

$$\mu_{P\times Q}\left((x, y), (x', y')\right) = \mu_P(x, x')\mu_Q(y, y'),$$

which gives us another proof of Lemma 6. This was the purported purpose of our digression. Of course, it may seem rather anticlimactic for all the theory developed here to culminate in a result that can be proven by calculation within a few lines. However, the theory of tensor products[1] is not limited to the context of posets and incidence algebras. Tensor products appear in diverse areas of mathematics and once learned, are often a useful tool for understanding results that may otherwise seem unmotivated.

Exercise 7 *Let V and W be vector spaces. Suppose that we specify bases on V and W; as we saw above, these give rise naturally to corresponding bases for V^* and W^*. With these bases, we can write $T : V \to W$ and $T^* : W^* \to V^*$ as matrices. Describe the relationship between these two matrices.*

[1]Note: We have only developed the theory of tensor products of vector spaces. More generally, one can define tensor products of modules.

Exercise 8 *Let $T : V \to W$ be a linear map, and let $\pi : W \to W/T(V)$ be the projection.*

1. *Show that $\pi \circ T$ is the zero map, and conclude that $(\pi \circ T)^*$ is also the zero map.*
2. *Show that for any surjective linear map S, S^* is injective. In particular, π^* is injective.*
3. *Based on the previous two parts, compute a lower bound on the dimension of the image of T^* in terms of the dimensions of V, W, and $W/T(V)$.*
4. *Using the previous exercise and a symmetry argument, deduce that the row and column ranks of a matrix are equal.*

Exercise 9 *Suppose that $S : V \to V'$ and $T : W \to W'$ are linear maps between vector spaces. Come up with a definition of $S \otimes T$ and state a composition law for it.*

Exercise 10 *We derived Lemma 6 using the tensor product $V \otimes W$, which we defined as the space of bilinear forms on $V^* \times W^*$. Come up with an alternate proof using instead the space of bilinear forms on $V \times W$. (In fact, this proof is perhaps simpler than the one we have given, but it does not directly introduce the tensor product.)*

4 Returning to posets

Let us now return to our main topic, the Mobius function on posets. Lemma 6 gives us a convenient way to compute the Mobius functions for some more complicated posets, as in the next example.

Example 11 *Let $P = B_n = \mathbf{2} \times \cdots \times \mathbf{2}$. The poset P can be naturally thought of as subsets of $[n]$ under inclusion; a subset $S \subset [n]$ represents the element $(b_1, \ldots, b_n) \in P$, where $b_i = 2$ if $i \in S$ and $b_i = 1$ if $i \notin S$.*

If $T \subset [n]$ represents $(c_1, \ldots, c_n) \in P$, and $S \subset T$, then we have by Lemma 6

$$\begin{aligned}\mu_P(S,T) &= \mu_P\left((b_1, \ldots, b_n), (c_1, \ldots, c_n)\right) \\ &= \prod_{i=1}^{n} \mu_{\mathbf{2}}(b_i, c_i) \\ &= (-1)^{|T|-|S|},\end{aligned}$$

where we have used the formula for $\mu_{\mathbf{2}}$ from Example 4. Applying Mobius inversion, we find that

$$f(T) = \sum_{S \subset T} g(S) \iff g(T) = \sum_{S \subset T} (-1)^{|T|-|S|} f(S).$$

This formula reminds us of the principle of inclusion-exclusion, and in fact, it is a generalization. To derive the principle of inclusion-exclusion as a special case, however, it is more convenient to work with Theorem 3, which states that

$$f(T) = \sum_{S \supset T} g(S) \iff g(T) = \sum_{S \supset T} (-1)^{|S|-|T|} f(S).$$

To enter the setting of inclusion-exclusion, suppose we have sets $E_1, \ldots, E_n$, and we wish to compute the cardinality of $X = E_1 \cup \cdots \cup E_n$. Let $g(S)$ be the number of elements in X which are contained in E_s if and only if $s \in S$. If $f(T) = \sum_{S \supset T} g(S)$, then $f(T)$ counts the number of elements in $\bigcup_{t \in T} E_t$. Rearranging the inversion formula gives

$$f(\emptyset) = g(\emptyset) + \sum_{S \neq \emptyset} (-1)^{|S|+1} f(S).$$

Noting that $f(\emptyset) = |X|$ and $g(\emptyset) = 0$, this gives us the same formula for $|X|$ as the principle of inclusion-exclusion.

Exercise 12 *Derive the principle of inclusion-exclusion directly from the first Mobius inversion formula in the above example. Can you derive Theorem 3 directly from Theorem 2?*

Example 13 *Let $P = D_n$ be the poset of divisors of an integer n, partially ordered by divisibility. If n has prime factorization $p_1^{n_1} \cdots p_k^{n_k}$, then $P \cong (\mathbf{n_1+1}) \times \cdots \times (\mathbf{n_k+1})$.*

Using Lemma 6 and Example 4, we find that $\mu_P(1, n) = \mu(n)$, where the second μ is the familiar number theoretic Mobius function defined on natural numbers. For $d_1, d_2 \in P$ with $d_1 | d_2$, we find that $[d_1, d_2] \cong D_{d_2/d_1}$. Thus, $\mu_P(d_1, d_2) = \mu_{D_{d_2/d_1}}(1, d_2/d_1) = \mu(d_2/d_1)$.

This shows the connection between the Mobius function for posets and the number theoretic Mobius function; in some sense, the number theoretic ζ and Mobius functions are the infinite versions of the functions of the same name we have defined on finite posets. (The analogy is slightly off, since the Riemann ζ function refers to a Dirichlet series, while the Mobius function refers to the coefficients on the Dirichlet series inverse of ζ.)

These examples conclude our basic development of the Mobius function for posets. We end with a lemma that will be used in the next lecture and also revisits the chain counting results from the previous lecture.

Lemma 14 *Let P be a finite rank n poset, and let $\widehat{P} = \mathbf{1} \oplus P \oplus \mathbf{1}$. Denote the minimal and maximal elements of $\widehat{P}$ by $\widehat{0}$ and $\widehat{1}$, respectively. If c_i is the number of length i chains from $\widehat{0}$ to $\widehat{1}$, then*

$$\mu_{\widehat{P}}(\widehat{0}, \widehat{1}) = \sum_{i=0}^{n+1} (-1)^i c_i,$$

where we define $c_0 = 1$.

Proof: Recall from the previous lecture that $(\zeta_{\widehat{P}} - 1)^k(x, y)$ counts the number of length k chains from x to y. In particular, $(\zeta_{\widehat{P}} - 1)^k(\widehat{0}, \widehat{1}) = c_k$.

This provides the connection between $\mu_{\widehat{P}}$ and the c_i. We can expand $\mu_{\widehat{P}}$ as a power series as follows.

$$\begin{aligned} \mu_{\widehat{P}} &= \big(1 + (\zeta_{\widehat{P}} - 1)\big)^{-1} \\ &= \sum_{i=0}^{\infty} (-1)^i (\zeta_{\widehat{P}} - 1)^i. \end{aligned}$$

The sum is actually finite, since $\zeta_{\widehat{P}} - 1$ is nilpotent. Evaluating both sides at $[\widehat{0}, \widehat{1}]$ and applying our initial observation yields the lemma. $\Box$