Introduction to Möbius Functions

Bruce E. Sagan
Department of Mathematics
Michigan State University
East Lansing, MI 48824-1027
sagan@math.msu.edu

1. Motivation and Introduction

2. Möbius functions of posets

3. The Möbius Inversion Theorem

4. The incidence algebra

5. The Möbius algebra

6. The order complex

Motivation and Introduction

I. Number Theory. The number-theoretic Möbius function is $\mu : \mathbb{Z}_{>0} \to \mathbb{Z}$ defined as

$$\mu(n) = \begin{cases} 0 & \text{if } n \text{ is not square free,} \\ (-1)^k & \text{if } n = \text{product of } k \text{ distinct primes,} \end{cases}$$

The importance of μ lies in the number-theoretic Möbius Inversion Theorem.

Theorem 1 *Let $f, g : \mathbb{Z}_{>0} \to \mathbb{Z}$ satisfy*

$$f(n) = \sum_{d|n} g(d)$$

for all $n \in \mathbb{Z}_{>0}$. Then

$$g(n) = \sum_{d|n} \mu(n/d) f(d). \qquad \blacksquare$$

II. Combinatorics. A very useful tool is the Principle of Inclusion-Exclusion or PIE.

Theorem 2 *Let S be a finite set and $S_1, \ldots, S_n \subseteq S$.*

$$\begin{aligned} |S - \bigcup_{i=1}^{n} S_i| &= |S| - \sum_{1 \le i \le n} |S_i| + \sum_{1 \le i < j \le n} |S_i \cap S_j| \\ & \qquad - \cdots + (-1)^n |\bigcap_{i=1}^{n} S_i|. \qquad \blacksquare \end{aligned}$$

III. Theory of Finite Differences. If one takes a function $f : \mathbb{Z}_{\geq 0} \to \mathbb{C}$ then there is an analog of the derivative, namely the difference operator

$$\Delta f(n) = f(n) - f(n-1)$$

(where $f(-1) = 0$ by definition), and an analog of the integral, namely the summation operator

$$Sf(n) = \sum_{i=0}^{n} f(n).$$

The Fundamental Theorem of the Difference Calculus then states

Theorem 3 (FTDC) *If* $f : \mathbb{Z}_{\geq 0} \to \mathbb{C}$ *then*

$$\Delta Sf(n) = f(n). \qquad \blacksquare$$

One of the advantages of the combinatorial Möbius function is that its inversion theorem unifies and generalizes the previous three results. In addition, it makes the number-theoretic definition transparant, encodes topological information about poets, and has even been used to bound the running time of certain algorithms.

Möbius functions of posets

Let P be a finite *poset* (<u>p</u>artially <u>o</u>rdered <u>set</u>) which is a set P with an order relation $\leq$ satisfying the three axioms for all $x, y, z \in P$:

1. (reflexsive) $x \leq x$,

2. (antisymmetry) $x \leq y$ and $y \leq x$ implies $x = y$,

3. (transitivity) $x \leq y$ and $y \leq z$ implies $x \leq z$,

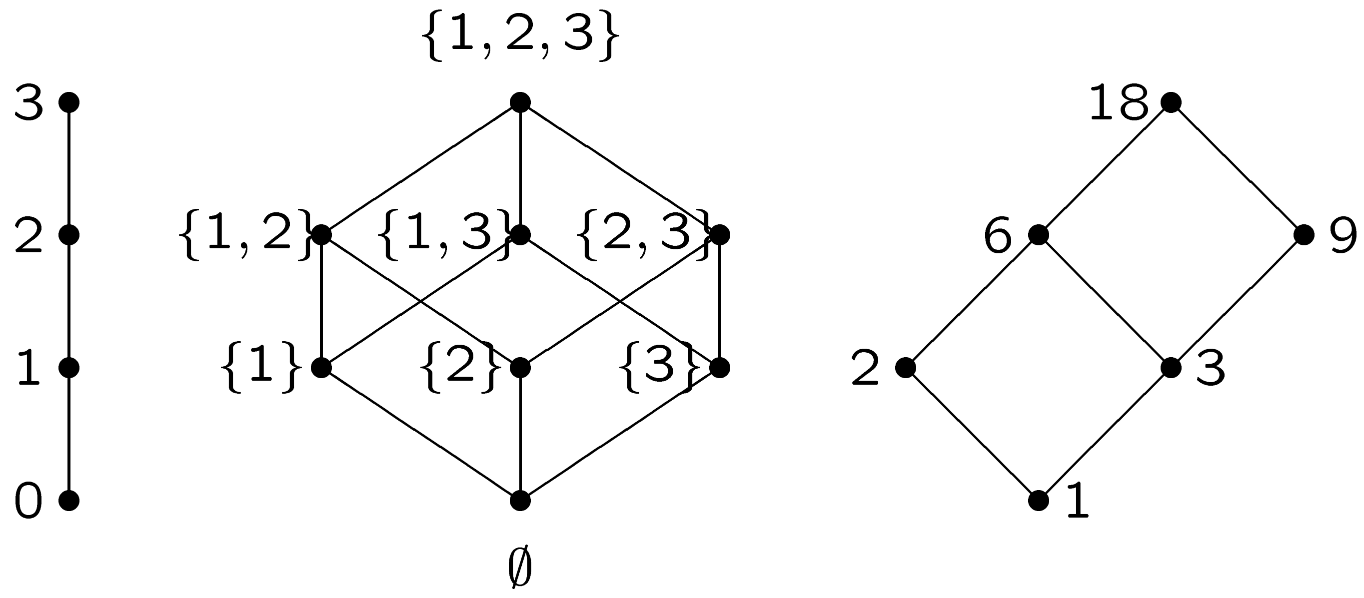
A poset is represented by its *Hasse diagram* which is a graph whose vertices are the elements of P and with an edge from x up to y if x is *covered* by y (i.e., $x < y$ and there is no z with $x < z < y$).

Examples:

1. The *chain*, C_n, consists of the integers $\{0, 1, \ldots, n\}$ ordered in the usual manner.

2. The *Boolean algebra*, B_n, has as elements all subsets of $[n] := \{1, 2, \ldots, n\}$ and $\subseteq$ as order relation.

3. The *divisor poset*, D_n, consists of all $d|n$ ordered by $c \leq d$ if $c|d$.



C_3 B_3 D_{18}

If P has a unique minimal element then it will be denoted $\hat{0} = \hat{0}_P$, and if it has a unique maximal element then we will use the notation $\hat{1} = \hat{1}_P$.

If P has a $\hat{0}$ then its *Möbius function*, $\mu : P \to \mathbf{Z}$, is defined recursively by

$$\mu(x) = \begin{cases} 1 & \text{if } x = \hat{0}, \\ -\sum_{y<x} \mu(y) & \text{if } x > \hat{0}. \end{cases}$$

Equivalently

$$\sum_{y \leq x} \mu(y) = \delta_{x,\hat{0}}$$

where $\delta_{x,\hat{0}}$ is the Kronecker delta.

Computations: It is immediate directly from the definition of μ that in C_n we have

$$\mu(x) = \begin{cases} 1 & \text{if } x = 0 \\ -1 & \text{if } x = 1 \\ 0 & \text{if } x \geq 2 \end{cases}$$

1. If $x \in B_n$ then $\mu(x) = (-1)^{|x|}$.

2. If $x \in D_n$ then $\mu(x)$ is the number theory function.

If P, Q are posets then order $P \times Q$ componentwise.

Proposition 4 *If $P \times Q$ has Möbius function μ then*

$$\mu(x, y) = \mu_P(x)\mu_Q(y)$$

Proof. It suffices to show that $\mu'(x, y) := \mu_P(x)\mu_Q(y)$ satisfies the defining equation for μ.

$$\begin{aligned} \sum_{(a,b)\leq(x,y)} \mu'(a, b) &= \sum_{(a,b)\leq(x,y)} \mu_P(a)\mu_Q(b) \\ &= \sum_{a\leq x} \mu_P(a) \sum_{b\leq y} \mu_Q(b) \\ &= \delta_{x,\hat{0}_P}\delta_{y,\hat{0}_Q} = \delta_{(x,y),(\hat{0}_P,\hat{0}_Q)}. \end{aligned} \qquad \blacksquare$$

P and Q are *isomorphic*, $P \cong Q$, if there is a bijection $f : P \to Q$ with both f and f^{-1} order-preserving.

Proposition 5 *If $f : P \to Q$ is an isomorphism then $\mu_P(x) = \mu_Q(f(x))$.* $\blacksquare$

1. B_n is isomorphic as a poset to the n-fold product $(C_1)^n$ where a set corresponds to its bit-string. Now for each bit $\mu(0) = 1, \mu(1) = -1$, so $\mu(x) = (-1)^{|x|}$.

2. If n has prime factorization $n = \prod_i p_i^{n_i}$ then we have an isomorphism $D_n \cong \times_i C_{n_i}$. So $p_i^{n_i}$ contributes -1 or 0 to the product for μ depending on whether $n_i = 1$ or $n_i \geq 2$.

The Möbius Inversion Theorem (MIT)

It is convenient to extend the definition of μ. If $x \le y$ in P then we have the *interval*

$$[x,y] = \{z \in P \;:\; x \le z \le y\}.$$

Let $\operatorname{Int} P$ denote the set of intervals of P.

If P is any finite poset, then its *Möbius function*, $\mu : \operatorname{Int} P \to \mathbf{Z}$, is defined recursively by

$$\mu(x,y) = \begin{cases} 1 & \text{if } x = y, \\ -\sum_{x \le z < y} \mu(x,z) & \text{if } x < y. \end{cases}$$

Equivalently

$$\sum_{x \le z \le y} \mu(x,z) = \delta_{x,y}.$$

So if P has a $\hat{0}$ then $\mu(x) = \mu(\hat{0}, x)$. Furthermore, we can consider $[x,y]$ as a poset in it's own right with $x = \hat{0}_{[x,y]}$ and in this poset the new definition reduces to the old one.

Note that our previous computations also give the two-variable version of μ since in all cases an interval $[x,y]$ in P_n ($P = C$, B, or D) satisfies

$$[x,y] \cong P_k$$

for some $k \le n$.

Theorem 6 *Let P be a poset and $f, g : P \to \mathbb{C}$.*

MIT1 If for all $x \in P$ we have $f(x) = \sum_{y \le x} g(y)$ then

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

MIT2 If for all $x \in P$ we have $f(x) = \sum_{y \ge x} g(y)$ then

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

Proof. We prove MIT2 as MIT1 is similar.

$$\begin{aligned}
\sum_{y \ge x} \mu(x,y) f(y) &= \sum_{y \ge x} \mu(x,y) \sum_{z \ge y} g(z) \\
&= \sum_{z \ge x} g(z) \sum_{x \le y \le z} \mu(x,y) \\
&= \sum_{z \in P} g(z) \delta_{x,z} \\
&= g(x). \qquad \blacksquare
\end{aligned}$$

Note (a) $\mathbb{C}$ above can be replaced by any vector space.

(b) The converses of MIT1 & 2 are also true.

It is now easy to obtain the three theorems from the introduction as corollaries by using Möbius inversion over D_n, B_n, and C_n, respectively.

I. If $f(m) = \sum_{d|m} g(d)$ for all m then in D_n we have $f(m) = \sum_{d \le m} g(d)$. So by MIT and $[d,m] \cong D_{m/d}$

$$g(m) = \sum_{d \le m} \mu(d,m) f(d) = \sum_{d|m} \mu(m/d) f(d).$$

II. For the PIE, given $S_1, \ldots, S_n \subseteq S$ and $x \in B_n$ let $S_x = \cap_{i \in x} S_i$. Define $f, g : B_n \to \mathbb{Z}_{\ge 0}$ by

$$\begin{aligned} f(x) &= |S_x|, \\ g(x) &= |S_x - \bigcup_{i \notin x} S_i|. \end{aligned}$$

For example in B_2

$$\begin{array}{ll} f(\emptyset) = |S| & g(\emptyset) = |S - (S_1 \cup S_2)| \\ f(1) = |S_1| & g(1) = |S_1 - S_2| \\ f(2) = |S_2| & g(2) = |S_2 - S_2| \\ f(12) = |S_1 \cap S_2| & g(12) = |S_1 \cap S_2|. \end{array}$$

By definition $f(x) = \sum_{y \ge x} g(y)$, so by MIT2

$$\left| S - \bigcup_{i=1}^{n} S_i \right| = g(\emptyset) = \sum_{y \ge \emptyset} \mu(y) f(y) = \sum_{y \in B_n} (-1)^{|y|} \left| \bigcap_{i \in y} S_i \right|.$$

III. For the FTDC, use $g(m) := Sf(m) = \sum_{i \le m} f(i)$ in C_n. So by MIT1 and $[i,j] \cong C_{j-i}$

$$f(m) = \sum_{i \le m} \mu(i,m) g(i) = g(m) - g(m-1)$$

$$= \Delta g(m) = \Delta S f(m).$$

The incidence algebra

Consider the *incidence algebra of* P, $I(P)$, which consists of all functions $f : \mathrm{Int}(P) \to \mathbb{C}$. The multiplication in this algebra is *convolution* defined by

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

It will be convenient to extend a $f \in I(P)$ to any pair $(x, y) \in P \times P$ by letting $f(x, y) = 0$ if $x \not\leq y$. So

$$f * g(x, y) = \sum_{z \in P} f(x, z) g(z, y).$$

Let $x_1, x_2, \ldots, x_n$ be a fixed linear extension of P (so $P = \{x_1, \ldots, x_n\}$ and $x_i < x_j$ in P implies $i < j$). Define a set of complex matrices $\mathrm{Mat}(P)$ with rows and columns indexed by $x_1, \ldots, x_n$:

$$M \in \mathrm{Mat}(P) \iff m_{x,y} = 0 \text{ if } x \not\leq y.$$

Proposition 7 *$I(P) \cong M(P)$ as algebras via the isomorphism*

$$f \in I(P) \leftrightarrow M_f = (f(x, y)). \quad \blacksquare$$

For example, B_2 has linear extension $\emptyset, 1, 2, 12$ and

$$M_\mu = \begin{pmatrix} 1 & -1 & -1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Corollary 8 *1. $I(P)$ has an identity element denoted* $1 : \mathrm{Int}(P) \to \mathbb{C}$ *given by* $1(x,y) = \delta_{x,y}$.

2. $f \in I(P)$ is invertible iff $f(x,x) \neq 0$ for all $x \in P$.

Proof. 1. M_1 is the identity matrix.

2. M_f is invertible iff $\det M_f \neq 0$. Since M_f is upper triangular, $\det M_f = \prod_{x \in P} f(x,x)$. ■

Define $\zeta \in I(P)$ by

$$\zeta(x,y) = 1 \quad \forall [x,y] \in \mathrm{Int}\, P.$$

Corollary 9 *We have $\zeta^{-1} = \mu$.*

Proof. It suffices to show $\mu * \zeta = 1$ but

$$(\mu * \zeta)(x,y) = \sum_{x \le z \le y} \mu(x,z)\zeta(z,y)$$

$$= \sum_{x \le z \le y} \mu(x,z) = 1(x,y). \quad ■$$

The Möbius algebra

A *lattice* is a poset L where every $x, y \in L$ have a greatest lower bound or meet, $x \wedge y$, and a least upper bound or joint, $x \vee y$. For example,

Poset	$x \wedge y$	$x \vee y$
C_n	$\min\{x, y\}$	$\max\{x, y\}$
B_n	$x \cap y$	$x \cup y$
D_n	$\gcd(x, y)$	$\operatorname{lcm}(x, y)$

Consider the complex vector space of formal sums

$$M(L) = \{\sum_{x \in L} c_x x \ : \ c_x \in \mathbb{C}\}.$$

So the elements of L are a basis for $M(L)$. For $x \in L$, define

$$\epsilon_x := \sum_{z \leq x} \mu(z, x) x$$

in $M(L)$. For example in $A(B_2)$ where we have $\mu(z, x) = (-1)^{|x-z|}$

$$\begin{aligned}
&\epsilon_\emptyset = \mu(\emptyset, \emptyset)\emptyset = \emptyset, \\
&\epsilon_1 = \mu(1, 1)1 - \mu(\emptyset, 1)\emptyset = 1 - \emptyset, \\
&\epsilon_2 = \mu(2, 2)2 - \mu(\emptyset, 2)\emptyset = 2 - \emptyset, \\
&\epsilon_{12} = \sum_{x \in B_2} \mu(x, 12)x = 12 - 1 - 2 + \emptyset.
\end{aligned}$$

Lemma 10 *a. If $x \in L$ then $x = \sum_{z \le x} \epsilon_z$.*

b. The set $\mathcal{B} = \{\epsilon_x \ : \ x \in L\}$ is a basis for $M(L)$.

Proof. 1. This follows immediately by applying the converse of MIT1 to the definition of the ϵ_x.

2. Since $|\mathcal{B}| = |L|$ we need only prove that the ϵ_x span, which follows from part 1. $\blacksquare$

The *Möbius algebra* is $M(L)$ with multiplication

$$\epsilon_x \cdot \epsilon_y = \delta_{x,y} \epsilon_x.$$

Proposition 11 *If $x, y \in L$ then $x \cdot y = x \wedge y$.*

Proof. Using Lemma 10.a

$$\begin{aligned} x \cdot y &= \left(\sum_{z \le x} \epsilon_z \right) \cdot \left(\sum_{w \le y} \epsilon_w \right) \\ &= \sum_{\substack{z \le x \\ w \le y}} \epsilon_z \cdot \epsilon_w \\ &= \sum_{z \le x, y} \epsilon_z \\ &= \sum_{z \le x \wedge y} \epsilon_z \\ &= x \wedge y. \quad \blacksquare \end{aligned}$$

Theorem 12 (Wiesner's Theorem) *Suppose L is a finite lattice and $c \in L \setminus \hat{1}$. Then*

$$\sum_{x \wedge c = \hat{0}} \mu(x, \hat{1}) = 0.$$

Proof. Expand $c \cdot \epsilon_{\hat{1}}$ in two ways

$$c \cdot \epsilon_{\hat{1}} = \sum_{d \leq c} \epsilon_d \cdot \epsilon_{\hat{1}} = 0$$

since $d \leq c < \hat{1}$ so $d \neq \hat{1}$ and $\epsilon_d \cdot \epsilon_{\hat{1}} = 0$. Now by the Proposition

$$0 = c \cdot \epsilon_{\hat{1}} = c \cdot \sum_{x \leq \hat{1}} \mu(x, \hat{1}) x = \sum_{x \in L} \mu(x, \hat{1}) x \wedge c.$$

Taking the coefficient of $\hat{0}$ on both sides finishes the proof. ■

Theorem 13 (Dual Weisner) *Suppose L is a finite lattice and $d \in L \setminus \hat{0}$. Then*

$$\sum_{x \vee c = \hat{1}} \mu(\hat{0}, x) = 0.$$

The *atoms* of poset P are

$$A = A(P) = \{a \in P \ : \ a \text{ covers } \hat{0}\}.$$

The *coatoms* of poset P are

$$A^* = A^*(P) = \{b \in P \ : \ \hat{1} \text{ covers } b\}.$$

For example

P	$A(P)$	$A^*(P)$
C_n	$\{1\}$	$\{n-1\}$
B_n	$\{x \ : \ \|x\| = 1\}$	$\{x \ : \ \|x\| = n-1\}$
D_n	$\{p \text{ prime} : \ p\|n\}$	$\{n/p : \ p \text{ prime}\}$

Theorem 14 (P. Hall) *Let L be a finite lattice.*

a. If $\bigvee A(L) \neq \hat{1}$ then $\mu(\hat{0}, \hat{1}) = 0$.

b. If $\bigwedge A^*(L) \neq \hat{0}$ then $\mu(\hat{0}, \hat{1}) = 0$.

Proof. a. Let $c = \bigvee A(L)$ in Weisner which is OK since $c \neq \hat{1}$.

$$\therefore \sum_{x \wedge c = \hat{0}} \mu(x, \hat{1}) = 0.$$

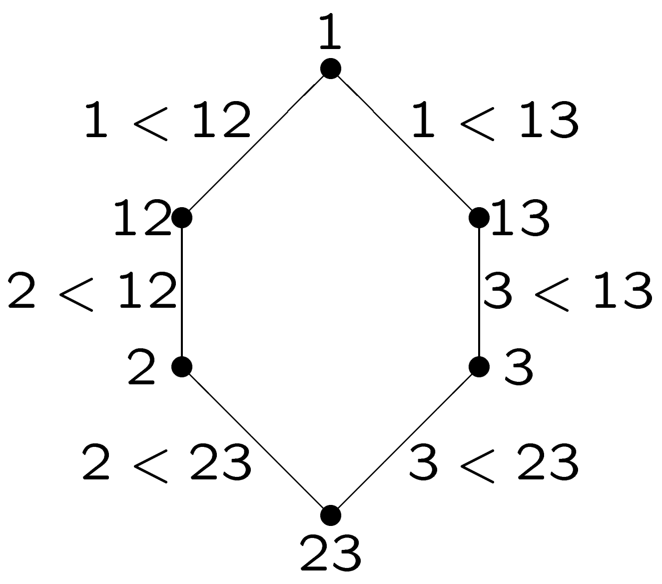
But $x \wedge c = \hat{0}$ implies $x = \hat{0}$. (If $x \neq \hat{0}$ then $x \geq a$ for some $a \in A(L)$ so $x \wedge c \geq a$.) But then the sum above has only one term corresponding to $x = \hat{0}$ and so reduces to $\mu(\hat{0}, \hat{1}) = 0$. ■

The order complex

A *chain C of length l* in P is $C : x_0 < x_1 < \ldots < x_l$. Equivalently $C \cong C_l$. Let P have a $\hat{0}$ and a $\hat{1} \neq \hat{0}$. The *order complex* is

$$\Delta P = \{C \ : \ x_0 > \hat{0},\ x_l < \hat{1}\}$$

Then ΔP is a simplicial complex since any subset of a chain is a chain. (We do not include $\hat{0}$ or $\hat{1}$ in our chains since if we did ΔP would be contractible.) For example



The simplices of ΔP of dimension l are

$$\Delta^l P = \{C \in \Delta P \ : \ C \text{ has length } l\}.$$

Also $\Delta^{-1} P = \{\emptyset\}$ and $\Delta^l P = \emptyset$ for $l \leq -2$.

Lemma 15 *For* $l \geq 0$: $(\zeta - 1)^l(\hat{0}, \hat{1}) = |\Delta^{l-2}P|$.

Proof. From the definitions

$$(\zeta - 1)(x, y) = \begin{cases} 1 & \text{if } x < y \\ 0 & \text{if } x = y \end{cases}$$

$$\begin{aligned}
\therefore\ & (\zeta - 1)^l(\hat{0}, \hat{1}) \\
&= \sum_{\hat{0} \leq x_0 \leq \ldots \leq x_{l-2} \leq \hat{1}} (\zeta - 1)(\hat{0}, x_0) \cdots (\zeta - 1)(x_{l-2}, \hat{1}) \\
&= \sum_{\hat{0} < x_0 < \ldots < x_{l-2} < \hat{1}} 1 \\
&= |\Delta^{l-2}P|. \quad \blacksquare
\end{aligned}$$

Proposition 16 (Reduced Euler Characteristic)

$$\mu(\hat{0}, \hat{1}) = \sum_{l \geq -1} (-1)^l |\Delta^l P|.$$

Proof. Using the lemma and $(\zeta - 1)(\hat{0}, \hat{1}) = 0$:

$$\begin{aligned}
\mu(\hat{0}, \hat{1}) &= \zeta^{-1}(\hat{0}, \hat{1}) \\
&= [1 + (\zeta - 1)]^{-1}(\hat{0}, \hat{1}) \\
&= \sum_{l \geq 0} (-1)^l (\zeta - 1)^l (\hat{0}, \hat{1}) \\
&= \sum_{l \geq 0} (-1)^l |\Delta^{l-2}P| \\
&= \sum_{l \geq -1} (-1)^l |\Delta^l P|. \quad \blacksquare
\end{aligned}$$