

Partially Ordered Sets and their Möbius Functions I: The Möbius Inversion Theorem

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

June 10, 2014

Lecture 1: The Möbius Inversion Theorem.

Introduction to partially ordered sets and Möbius functions.

Lecture 2: Graph Coloring.

The chromatic polynomial of a graph and the characteristic polynomial of its bond lattice.

Lecture 3: Topology of Posets.

The order complex and shellability.

Lecture 4: Factoring the Characteristic Polynomial

Quotients of posets and applications.

Example A: Combinatorics.

Given a set, S , let

$$\#S = |S| = \text{cardinality of } S.$$

The Principle of Inclusion-Exclusion or PIE is a very useful tool in enumerative combinatorics.

Theorem (PIE)

Let U be a finite set and $U_1, \dots, U_n \subseteq U$. We have

$$\begin{aligned} \left| U - \bigcup_{i=1}^n U_i \right| &= |U| - \sum_{1 \leq i \leq n} |U_i| + \sum_{1 \leq i < j \leq n} |U_i \cap U_j| \\ &\quad - \dots + (-1)^n \left| \bigcap_{i=1}^n U_i \right|. \end{aligned}$$

□

Example B: Theory of Finite Differences.

\mathbb{N} = the nonnegative integers.

\mathbb{P} = the positive integers.

\mathbb{R} = the real numbers.

If one takes a function $f : \mathbb{N} \rightarrow \mathbb{R}$ then there is an analogue of the derivative, namely the difference operator

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

(where $f(-1) = 0$ by definition). There is also an analogue of the integral, namely the summation operator

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

The Fundamental Theorem of the Difference Calculus states:

Theorem (FTDC)

If $f : \mathbb{N} \rightarrow \mathbb{R}$ then

$$\Delta Sf(n) = f(n). \quad \square$$

Example C: Number Theory

If $d, n \in \mathbb{Z}$ then write $d|n$ if d divides evenly into n . The number-theoretic Möbius function is $\mu : \mathbb{P} \rightarrow \mathbb{Z}$ 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 or MIT.

Theorem (Number Theory MIT)

Let $f, g : \mathbb{P} \rightarrow \mathbb{R}$ satisfy

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

for all $n \in \mathbb{P}$. Then

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

Möbius inversion over partially ordered sets (posets) is important for the following reasons.

1. It unifies and generalizes the three previous examples.
2. It makes the number-theoretic definition transparent.
3. It encodes topological information about partially ordered sets.
4. It can be used to solve combinatorial problems.

A *partially ordered set* or *poset* is a set P together with a binary relation \leq such that for all $x, y, z \in P$:

1. (reflexivity) $x \leq x$,
2. (antisymmetry) $x \leq y$ and $y \leq x$ implies $x = y$,
3. (transitivity) $x \leq y$ and $y \leq z$ implies $x \leq z$.

Given any poset notation, if we wish to be specific about the poset P involved, we attach P as a subscript. For example, using \leq_P for \leq . We also adopt the usual conventions for inequalities. For example, $x < y$ means $x \leq y$ and $x \neq y$. We write $x \parallel y$ if x, y are *incomparable*, that is $x \not\leq y$ and $y \not\leq x$. All posets will be finite unless otherwise stated.

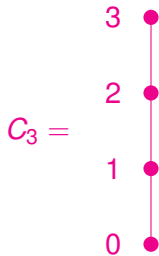
If $x, y \in P$ then x is covered by y or y covers x , written $x \triangleleft y$, if $x < y$ and there is no z with $x < z < y$. The *Hasse diagram* of P is the (directed) graph with vertices P and an edge from x up to y if $x \triangleleft y$.

Example: The Chain.

The *chain of length n* is

$$C_n = \{0, 1, \dots, n\}$$

with the usual \leq on the integers.



Example: The Boolean Algebra.

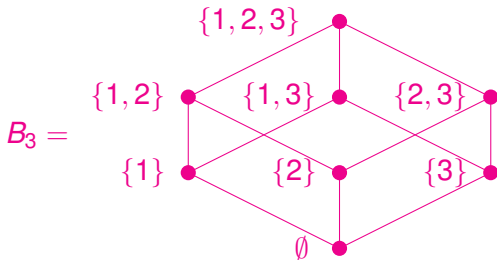
Let

$$[n] = \{1, 2, \dots, n\}.$$

The *Boolean algebra* is

$$B_n = \{S : S \subseteq [n]\}$$

partially ordered by $S \leq T$ if and only if $S \subseteq T$.



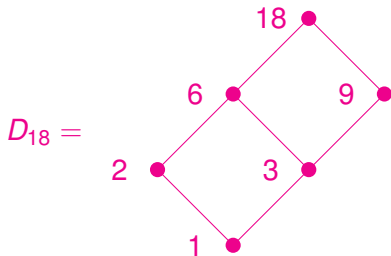
Note that B_3 looks like a cube.

Example: The Divisor Lattice.

Given $n \in \mathbb{P}$ the corresponding *divisor lattice* is

$$D_n = \{d \in \mathbb{P} : d|n\}$$

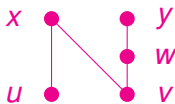
partially ordered by $c \leq_{D_n} d$ if and only if $c|d$.



Note that D_{18} looks like a rectangle.

In a poset P , a *minimal* element is $x \in P$ such that there is no $y \in P$ with $y < x$. A *maximal* element is $x \in P$ such that there is no $y \in P$ with $y > x$.

Example. The poset on the left has minimal elements u and v , and maximal elements x and y .



A poset *has a zero* if it has a unique minimal element, $\hat{0}$. A poset *has a one* if it has a unique maximal element, $\hat{1}$. A poset is *bounded* if it has both a $\hat{0}$ and a $\hat{1}$.

Example. Our three fundamental examples are bounded:

$$\hat{0}_{C_n} = 0, \quad \hat{1}_{C_n} = n, \quad \hat{0}_{B_n} = \emptyset, \quad \hat{1}_{B_n} = [n], \quad \hat{0}_{D_n} = 1, \quad \hat{1}_{D_n} = n.$$

If $x \leq y$ in P then the corresponding *closed interval* is

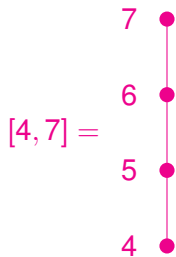
$$[x, y] = \{z : x \leq z \leq y\}.$$

Open and half-open intervals are defined analogously. Note that $[x, y]$ is a poset in its own right and it has a zero and a one:

$$\hat{0}_{[x,y]} = x, \quad \hat{1}_{[x,y]} = y.$$

Example: The Chain.

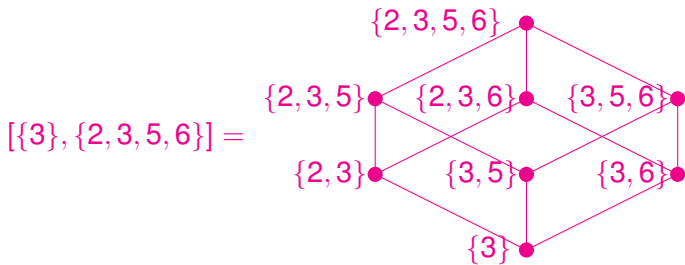
In C_9 we have the interval



This interval looks like C_3 .

Example: The Boolean Algebra.

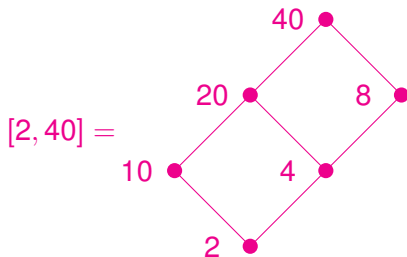
In B_7 we have the interval



Note that this interval looks like B_3 .

Example: The Divisor Lattice.

In D_{80} we have the interval



Note that this interval looks like D_{18} .

For posets P and Q , an *order preserving (op) map* is $f : P \rightarrow Q$ with

$$x \leq_P y \implies f(x) \leq_Q f(y).$$

An *isomorphism* is a bijection $f : P \rightarrow Q$ such that both f and f^{-1} are op. In this case P and Q are *isomorphic*, written $P \cong Q$.

Proposition

If $i \leq j$ in C_n then $[i, j] \cong C_{j-i}$.

If $S \subseteq T$ in B_n then $[S, T] \cong B_{|T-S|}$.

If $c|d$ in D_n then $[c, d] \cong D_{d/c}$.

Proof for C_n . Define $f : [i, j] \rightarrow C_{j-i}$ by $f(k) = k - i$. Then f is op since

$$k \leq l \implies k - i \leq l - i \implies f(k) \leq f(l).$$

Also f is bijective with inverse $f^{-1}(k) = k + i$. Similarly, one can prove that f^{-1} is op. □

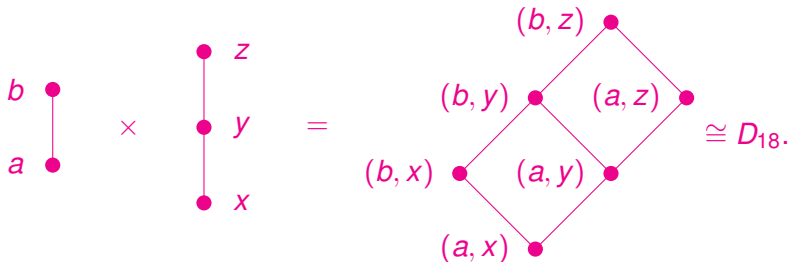
If P and Q are posets, then their *product* is

$$P \times Q = \{(a, x) : a \in P, x \in Q\}$$

partially ordered by

$$(a, x) \leq_{P \times Q} (b, y) \iff a \leq_P b \text{ and } x \leq_Q y.$$

Ex.



If P is a poset then let $P^n = \overbrace{P \times \dots \times P}^n$.

Proposition

For the Boolean algebra

$$B_n \cong (C_1)^n.$$

If the prime factorization of n is $n = p_1^{m_1} \cdots p_k^{m_k}$, then

$$D_n \cong C_{m_1} \times \cdots \times C_{m_k}.$$

Proof for B_n . Since $C_1 = \{0, 1\}$, define $f : B_n \rightarrow (C_1)^n$ by

$$f(S) = (b_1, b_2, \dots, b_n) \quad \text{where} \quad b_i = \begin{cases} 1 & \text{if } i \in S, \\ 0 & \text{if } i \notin S. \end{cases}$$

for $1 \leq i \leq n$. To show f is op, suppose that we have $f(S) = (b_1, \dots, b_n)$ and $f(T) = (c_1, \dots, c_n)$. Now $S \leq T$ in B_n means $S \subseteq T$. Equivalently, $i \in S$ implies $i \in T$ for every $1 \leq i \leq n$. So for each $1 \leq i \leq n$ we have $b_i \leq c_i$ in C_1 . But then $(b_1, \dots, b_n) \leq (c_1, \dots, c_n)$ in $(C_1)^n$, that is, $f(S) \leq f(T)$. Constructing f^{-1} and proving it op is similar. □

The *incidence algebra* of a finite poset P is the set

$$I(P) = \{\alpha : P \times P \rightarrow \mathbb{R} \mid \alpha(x, y) = 0 \text{ if } x \not\leq y\},$$

together with the operations:

1. (addition) $(\alpha + \beta)(x, y) = \alpha(x, y) + \beta(x, y)$,
2. (scalar multiplication) $(k\alpha)(x, y) = k \cdot \alpha(x, y)$ for $k \in \mathbb{R}$,
3. (convolution) $(\alpha * \beta)(x, y) = \sum_{z \in P} \alpha(x, z)\beta(z, y)$.

Ex. $I(P)$ has Kronecker's delta: $\delta(x, y) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{if } x \neq y. \end{cases}$

Proposition

*For all $\alpha \in I(P)$: $\alpha * \delta = \delta * \alpha = \alpha$.*

Proof of $\alpha * \delta = \alpha$. For any $x, y \in P$:

$$(\alpha * \delta)(x, y) = \sum_z \alpha(x, z)\delta(z, y) = \alpha(x, y)\delta(y, y) = \alpha(x, y). \quad \square$$

Note. We have

$$(\alpha * \beta)(x, y) = \sum_{z \in [x, y]} \alpha(x, z)\beta(z, y)$$

since $\alpha(x, z) \neq 0$ implies $x \leq z$ and $\beta(z, y) \neq 0$ implies $z \leq y$.

An *algebra* over a field F is a set A together with operations of sum (+), product (\bullet), and scalar multiplication (\cdot) such that

1. $(A, +, \bullet)$ is a ring,
2. $(A, +, \cdot)$ is a vector space over F ,
3. $k \cdot (a \bullet b) = (k \cdot a) \bullet b = a \bullet (k \cdot b)$ for all $k \in F, a, b \in A$.

Ex. The $n \times n$ matrix algebra over \mathbb{R} is

$$\text{Mat}_n(\mathbb{R}) = \text{all } n \times n \text{ matrices with entries in } \mathbb{R}.$$

Ex. The Boolean algebra is an algebra over \mathbb{F}_2 where, for all $S, T \in B_n$:

1. $S + T = (S \cup T) - (S \cap T)$,
2. $S \bullet T = S \cap T$,
3. $0 \cdot S = \emptyset$ and $1 \cdot S = S$.

Ex. The incidence algebra $I(P)$ is an algebra with convolution as the product.

Note. Often \cdot and \bullet are suppressed since context makes it clear which multiplication is meant.

Let $L : x_1, \dots, x_n$ be a list of the elements of P . An $L \times L$ matrix has rows and columns indexed by L . The *matrix algebra* of P is

$$M(P) = \{M \in \text{Mat}_n(\mathbb{R}) \mid M \text{ is } L \times L \text{ and } M_{x,y} = 0 \text{ if } x \not\leq y.\}$$

Note that $M(P)$ is a subalgebra of $\text{Mat}_n(\mathbb{R})$.

Ex. For B_2 , let $L : \emptyset, \{1\}, \{2\}, \{1,2\}$. Then a typical element of $M(B_2)$ is

$$M = \begin{matrix} & \emptyset & \{1\} & \{2\} & \{1,2\} \\ \emptyset & \heartsuit & \heartsuit & \heartsuit & \heartsuit \\ \{1\} & 0 & \heartsuit & 0 & \heartsuit \\ \{2\} & 0 & 0 & \heartsuit & \heartsuit \\ \{1,2\} & 0 & 0 & 0 & \heartsuit \end{matrix}$$

where the \heartsuit 's can be replaced by any real numbers.

The list $L : x_1, \dots, x_n$ is a *linear extension* of P if $x_i \leq x_j$ in P implies $i \leq j$, that is, x_i comes before x_j in L . Henceforth we will take L to be a linear extension. This makes each $M \in M(P)$ upper triangular:

$$i > j \implies x_i \not\leq x_j \implies M_{x_i, x_j} = 0.$$

An *isomorphism* of algebras A and B is a bijection $f : A \rightarrow B$ such that for all $a, b \in A$ and $k \in F$,

$$f(a + b) = f(a) + f(b), \quad f(a \bullet b) = f(a) \bullet f(b), \quad f(k \cdot a) = k \cdot f(a).$$

Given any $\alpha \in I(P)$ we let M^α be the matrix with entries

$$M_{x,y}^\alpha = \alpha(x, y).$$

Ex. We have $M^\delta = I$ where I is the identity matrix.

Theorem

The map $\alpha \mapsto M^\alpha$ is an algebra isomorphism $I(P) \rightarrow M(P)$.

Proof that product is preserved. We wish to show

$M^{\alpha * \beta} = M^\alpha M^\beta$. But given $x, y \in P$:

$$M_{x,y}^{\alpha * \beta} = (\alpha * \beta)(x, y) = \sum_z \alpha(x, z) \beta(z, y) = (M^\alpha M^\beta)_{x,y}. \quad \square$$

Proposition

If $\alpha \in I(P)$ then α^{-1} exists if and only if $\alpha(x, x) \neq 0$ for all $x \in P$.

Proof. By the previous theorem

$$\exists \alpha^{-1} \iff \exists (M^\alpha)^{-1} \iff \det M^\alpha \neq 0 \iff \prod_{x \in P} \alpha(x, x) \neq 0. \quad \square$$

The *zeta function* of P is $\zeta \in I(P)$ defined by

$$\zeta(x, y) = \begin{cases} 1 & \text{if } x \leq y, \\ 0 & \text{if } x \not\leq y. \end{cases}$$

The *Möbius function* of P is $\mu = \zeta^{-1}$. Note that μ is well defined by the previous proposition. From the definition of μ :

$$\delta(x, y) = (\mu * \zeta)(x, y) = \sum_{z \in [x, y]} \mu(x, z) \zeta(z, y) = \sum_{z \in [x, y]} \mu(x, z).$$

Equivalently,

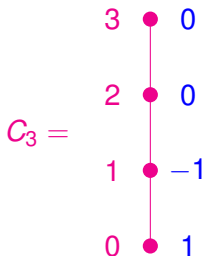
$$\begin{aligned} &\text{if } x = y \text{ then } \mu(x, x) = 1, \\ &\text{if } x < y \text{ then } \sum_{z \in [x, y]} \mu(x, z) = 0. \end{aligned}$$

Note. If P has a zero then we write $\mu(y) = \mu(\hat{0}, y)$, and so

$$\mu(\hat{0}) = 1, \text{ and if } y > \hat{0} \text{ then } \sum_{z \leq y} \mu(z) = 0.$$

$$\mu(\hat{0}) = 1, \text{ and if } y > \hat{0} \text{ then } \sum_{z \leq y} \mu(z) = 0.$$

Example The Chain.



$$\mu(0) = \mu(\hat{0}) = 1,$$

$$\mu(1) + \mu(0) = 0 \implies \mu(1) = -1,$$

$$\mu(2) + \mu(1) + \mu(0) = 0 \implies \mu(2) = 0,$$

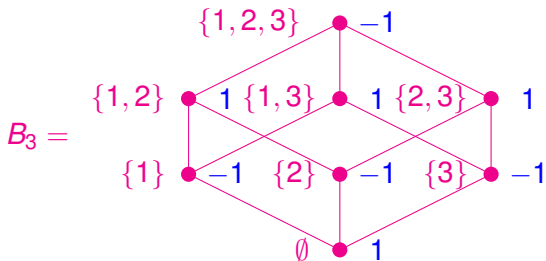
$$\mu(3) + \mu(2) + \mu(1) + \mu(0) = 0 \implies \mu(3) = 0.$$

Proposition

In C_n we have $\mu(i) =$

$$\begin{cases} 1 & \text{if } i = 0 \\ -1 & \text{if } i = 1, \\ 0 & \text{else.} \end{cases}$$


Example: The Boolean Algebra.



$$\mu(\emptyset) = \mu(\hat{0}) = 1,$$

$$\mu(\{1\}) + \mu(\emptyset) = 0 \implies \mu(\{1\}) = -1,$$

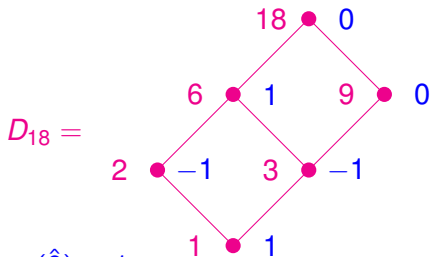
$$\mu(\{1, 2\}) + \mu(\{1\}) + \mu(\{2\}) + \mu(\emptyset) = 0 \implies \mu(\{1, 2\}) = 1,$$

$$\mu(\{1, 2, 3\}) + \cdots + \mu(\emptyset) = 0 \implies \mu(\{1, 2, 3\}) = -1.$$

Conjecture

In B_n we have $\mu(S) = (-1)^{|S|}$.

Example: The Divisor Lattice.



$$\mu(1) = \mu(\hat{0}) = 1,$$

$$\mu(2) = \mu(3) = -1,$$

$$\mu(6) + \mu(2) + \mu(3) + \mu(1) = 0 \implies \mu(6) = 1,$$

$$\mu(9) + \mu(3) + \mu(1) = 0 \implies \mu(9) = 0,$$

$$\mu(18) + \dots + \mu(1) = 0 \implies \mu(18) = 0.$$

Conjecture

If $d \in D_n$ has prime factorization $d = p_1^{m_1} \cdots p_k^{m_k}$ then

$$\mu(d) = \begin{cases} (-1)^k & \text{if } m_1 = \dots = m_k = 1, \\ 0 & \text{if } m_i \geq 2 \text{ for some } i. \end{cases}$$

Theorem

1. If $f : P \rightarrow Q$ is an isomorphism and $x, y \in P$ then

$$\mu_P(x, y) = \mu_Q(f(x), f(y)).$$

2. If $a, b \in P$ and $x, y \in Q$ then

$$\mu_{P \times Q}((a, x), (b, y)) = \mu_P(a, b)\mu_Q(x, y). \quad (1)$$

Proof for $P \times Q$. For any poset R , the equation

$\sum_{t \in [r, s]} \mu(r, t) = \delta(r, s)$ uniquely defines μ . So it suffices to show that the right-hand side of (??) satisfies the defining equation.

$$\begin{aligned} \sum_{(c, z) \in [(a, x), (b, y)]} \mu_P(a, c)\mu_Q(x, z) &= \sum_{c \in [a, b]} \mu_P(a, c) \sum_{z \in [x, y]} \mu_Q(x, z) \\ &= \delta_P(a, b)\delta_Q(x, y) \\ &= \delta_{P \times Q}((a, x), (b, y)). \quad \square \end{aligned}$$

Theorem

1. If $S \in B_n$ then $\mu(S) = (-1)^{|S|}$
2. If $d = p_1^{m_1} \cdots p_k^{m_k} \in D_n$ then

$$\mu(d) = \begin{cases} (-1)^k & \text{if } m_1 = \dots = m_k = 1, \\ 0 & \text{if } m_i \geq 2 \text{ for some } i. \end{cases}$$

Proof for B_n . We have an isomorphism $f : B_n \rightarrow (C_1)^n$. Also

$$\mu_{C_1}(0) = 1 \quad \text{and} \quad \mu_{C_1}(1) = -1.$$

Now if $f(S) = (b_1, \dots, b_n)$ then by the previous theorem

$$\begin{aligned} \mu_{B_n}(S) &= \mu_{(C_1)^n}(b_1, \dots, b_n) \\ &= \prod_i \mu_{C_1}(b_i) \\ &= (-1)^{\#\text{ of } b_i = 1} \\ &= (-1)^{|S|}. \quad \square \end{aligned}$$

Theorem (Möbius Inversion Thm - MIT, Weisner (1935))

Consider a finite poset P and two functions $f : P \rightarrow \mathbb{R}$ and $g : P \rightarrow \mathbb{R}$. Then the following are equivalent statements.

1. $f(y) = \sum_{x \leq y} g(x)$ for all $y \in P$.
2. $g(y) = \sum_{x \leq y} \mu(x, y)f(x)$ for all $y \in P$.

Proof. Let $L : x_1, \dots, x_n$ be the linear extension used for $I(P)$. Consider vectors $v^f = [f(x_1) \dots f(x_n)]$, $v^g = [g(x_1), \dots, g(x_n)]$.

$$f(y) = \sum_{x \leq y} g(x) \quad \forall y \in P \iff f(y) = \sum_{x \in P} g(x)\zeta(x, y) \quad \forall y \in P$$

$$\iff v^f = v^g M^\zeta \iff v^g = v^f (M^\zeta)^{-1} = v^f M^\mu$$

$$\iff g(y) = \sum_{x \in P} f(x)\mu(x, y) \quad \forall y \in P$$

$$\iff g(y) = \sum_{x \leq y} f(x)\mu(x, y) \quad \forall y \in P. \quad \square$$

Theorem (MIT)

$$f(y) = \sum_{x \leq y} g(x) \forall y \in P \iff g(y) = \sum_{x \leq y} \mu(x, y) f(x) \forall y \in P. \quad \square$$

Ex. Theory of Finite Differences.

For $g : \mathbb{N} \rightarrow \mathbb{R}$: $\Delta g(n) = g(n) - g(n-1)$, $Sg(n) = \sum_{i=0}^n g(i)$.

Theorem (FTDC)

If $g : \mathbb{N} \rightarrow \mathbb{R}$ then: $\Delta Sg(n) = g(n)$.

Proof. Consider the chain C_n and the restriction $g : C_n \rightarrow \mathbb{R}$.
For each $k \in C_n$, define

$$f(k) = \sum_{i \leq k} g(i) = Sg(k).$$

Then by the MIT applied to C_n

$$\begin{aligned} g(n) &= \sum_{i \leq n} \mu(i, n) f(i) = \mu(n, n) f(n) + \mu(n-1, n) f(n-1) \\ &= f(n) - f(n-1) = \Delta f(n) = \Delta Sg(n). \quad \square \end{aligned}$$

Theorem (Dual MIT)

$$f(x) = \sum_{y \geq x} g(y) \forall x \in P \iff g(x) = \sum_{y \geq x} \mu(x, y) f(y) \forall x \in P. \quad \square$$

Ex. Principle of Inclusion-Exclusion.

Theorem (PIE)

Let U be a finite set and $U_1, \dots, U_n \subseteq U$.

$$\left| U - \bigcup_{i=1}^n U_i \right| = |U| - \sum_{1 \leq i \leq n} |U_i| + \dots + (-1)^n \left| \bigcap_{i=1}^n U_i \right|.$$

Proof. For the Boolean algebra B_n , define $f, g : B_n \rightarrow \mathbb{R}$ by

$f(S) = \#$ of elements in all $U_i, i \in S$, and possibly other U_j ,

$g(S) = \#$ of elements in all $U_i, i \in S$, and no other U_j .

Now $f(S) = \left| \bigcap_{i \in S} U_i \right|$ and $f(S) = \sum_{T \supseteq S} g(T)$. Thus

$$\left| U - \bigcup_{i=1}^n U_i \right| = g(\emptyset) = \sum_{T \supseteq \emptyset} \mu(\emptyset, T) f(T) = \sum_{T \in B_n} (-1)^{|T|} \left| \bigcap_{i \in T} U_i \right|. \quad \square$$

Theorem (MIT)

$$f(y) = \sum_{x \leq y} g(x) \forall y \in P \iff g(y) = \sum_{x \leq y} \mu(x, y) f(x) \forall y \in P. \quad \square$$

Ex. Number Theory

Theorem (Number Theory MIT)

Let $f, g : \mathbb{P} \rightarrow \mathbb{R}$ satisfy $f(n) = \sum_{d|n} g(d)$ for all $n \in \mathbb{P}$. Then

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

Proof. The restrictions $f, g : D_n \rightarrow \mathbb{R}$ satisfy, for all $m \in D_n$:

$$f(m) = \sum_{d|m} g(d) = \sum_{d \leq_{D_n} m} g(d).$$

Apply the poset MIT to D_n and use $[d, n] \cong [1, n/d]$:

$$g(n) = \sum_{d \leq_{D_n} n} \mu(d, n) f(d) = \sum_{d|n} \mu(d, n) f(d) = \sum_{d|n} \mu(n/d) f(d). \quad \square$$