6 Möbius Inversion and Lattices

Motivation

Problem 1: We are interested in counting the number of surjections from \([n]\) to \([k]\). Although this is difficult to count directly, it is rather easy to count the number of (arbitrary) functions from \([n]\) to a given subset \(H\) of \([k]\), as this is given by \(|H|^n\). The solution to this problem is a standard representative of an inclusion-exclusion argument.

Euler’s \(\phi\): For \(n \in \mathbb{N} \setminus \{0\}\) we define

\[
\phi(n) = \{1 \leq k \leq n : GCD(k, n) = 1\}
\]

Note that \(\phi(n)\) is the order of the multiplicative group \(\mathbb{Z}_n^*\).

Proposition 6.1 \(\sum_{d|n} \phi(d) = n\)

Proof: The number of integers \(m \in [n]\) with \((m, n) = d\) (i.e. \(m = m'd\) and \(n = n'd\) with \((m', n') = 1\) and \(m' \leq n'\)) is precisely \(\phi(n') = \phi(n/d)\). Thus

\[
n = \sum_{d|n} \phi(n/d) = \sum_{d|n} \phi(d) \quad \Box
\]

Problem 2: Use the above proposition to find \(\phi(n)\).

Problem 3: We are interested in counting the number of linear functions from \(\mathbb{F}_q^n\) onto \(\mathbb{F}_q^k\). Although this is difficult to do directly, it is easy to count the number of (arbitrary) linear functions from \(\mathbb{F}_q^n\) to \(\mathbb{F}_q^k\) as we may freely assign the image of a basis, and this forces the remaining elements, this is given by \((q^k)^n = q^{nk}\).

General Problem: Consider a poset \((P, \leq)\) and a function \(f : P \rightarrow \mathbb{C}\) and let \(g : P \rightarrow \mathbb{C}\) be given by \(g(x) = \sum_{y \leq x} f(x)\). We would like to use \(g\) to find \(f\).
Möbius Inversion

Incidence Algebra: If \((P, \leq)\) is a poset, the associated incidence algebra is

\[ \mathbb{A}(P) = \{ \alpha \in \mathbb{C}^{P \times P} : \alpha(x, y) = 0 \text{ unless } x \leq y \} \]

Note that if \(\alpha, \beta \in \mathbb{A}(P)\) then by matrix multiplication we have

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

It is immediate that \(\mathbb{A}(P)\) is closed under scalar multiplication, addition, and multiplication. We let \(I \in \mathbb{A}(P)\) denote the function with \(I(x, y) = 0\) if \(x \neq y\) and \(I(x, y) = 1\) if \(x = y\) and note that \(I\) is a multiplicative identity.

Zeta Function: The zeta function is given by

\[ \zeta(x, y) = \begin{cases} 1 & \text{if } x \leq y \\ 0 & \text{otherwise} \end{cases} \]

Note that \(\zeta \in \mathbb{A}(P)\).

Möbius Function: We define a function \(\mu \in \mathbb{A}(P)\) by the rule that \(\mu(x, y) = 0\) whenever \(x \not\leq y\) and \(\mu(x, x) = 1\) and then define the other values inductively by the rule:

\[ \mu(x, y) = -\sum_{x \leq z < y} \mu(x, z) \]

It is immediate from this definition that

\[ \sum_{x \leq z \leq y} \mu(x, z) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases} \tag{1} \]

This last equation is equivalent to the \(\mu \zeta = I\). So, in fact, \(\mu\) is the inverse of \(\zeta\). Note further that by this construction, \(\mu\) is always integer valued. The equation \(\zeta \mu = I\) yields the following similar identity for the Möbius function.

\[ \sum_{x \leq z \leq y} \mu(z, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases} \tag{2} \]
Theorem 6.2 (Möbius Inversion) Let \((P, \leq)\) be a poset, let \(\mu\) be its Möbius function, let \(f : P \to \mathbb{C}\) and let \(g : P \to \mathbb{C}\) be given by \(g(x) = \sum_{y \leq x} f(y)\). Then

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

Proof:

\[
\sum_{y \leq x} \mu(y, x) g(y) = \sum_{y \leq x} \mu(y, x) \left( \sum_{z \leq y} f(z) \right)
\]

\[
= \sum_{z \leq y \leq x} \mu(y, x) f(z)
\]

\[
= \sum_{z \leq x} f(z) \left( \sum_{z \leq y \leq x} \mu(y, x) \right)
\]

\[
= f(x) \quad \square
\]

Note: The above formula is precisely what we need to solve our general problem, so long as we can find the Möbius function.

Proposition 6.3 The Möbius function for the poset of subsets of \([n]\) is given by:

\[
\mu(A, B) = \begin{cases} 
(-1)^{|B| - |A|} & \text{if } A \subseteq B \\
0 & \text{otherwise}
\end{cases}
\]

Proof: It suffices to show that \(\mu(\emptyset, B) = (-1)^{|B|}\) as the other cases are similar. This case we prove by induction on \(|B|\). As a base, note that it holds trivially when \(|B| = 0\). For the inductive step, let \(|B| = b\) and assume that the result holds for all sets with size less than \(|B|\). Then using \(0 = (1 - 1)^{|B|} = \sum_{C \subseteq B} (-1)^{|C|}\) and the formula for the Möbius function we have

\[
\mu(\emptyset, B) = -\sum_{C \subseteq B} \mu(\emptyset, C)
\]

\[
= -\sum_{C \subseteq B} (-1)^{|C|}
\]

\[
= (-1)^{|B|} \quad \square
\]

Answer to Problem 1: For every \(H \subseteq [k]\) let \(f(H)\) denote the number of surjections from \([n]\) to \(H\). Then define for every \(H \subseteq [k]\) the function \(g(H) = \sum_{J \subseteq H} f(J)\). Now, \(g(H)\) is
precisely the total number of functions from \([n]\) to \(H\) so \(g(H) = |H|^n\). By Möbius inversion we get

\[
f([k]) = \sum_{H \subseteq [k]} (-1)^{k-|H|} |H|^n = \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i} n^i
\]

More generally, using the Möbius function on the poset of subsets of \([n]\) is the general technique of inclusion-exclusion.

**Lattice:** A lattice is a poset \((L, \leq)\) with the following additional properties:

(i) for every \(x, y \in L\) there is a unique element, denoted \(x \vee y\), with the property that \(z \geq x\) and \(z \geq y\) implies \(z \geq x \vee y\).

(ii) for every \(x, y \in L\) there is a unique element, denoted \(x \wedge y\), with the property that \(z \leq x\) and \(z \leq y\) implies \(z \leq x \wedge y\).

It follows immediately from (i) and (ii) that for any subset \(X \subseteq L\) there is a unique element which is minimal (maximal) subject to being greater than (less than) or equal to all elements in \(X\). It follows from this that every finite lattice has a unique minimal element which we denote by \(0_L\) and a unique maximal element which we denote by \(1_L\).

**Theorem 6.4 (Weisner)** Let \(\mu\) be the Möbius function of a finite lattice \((L, \leq)\) and let \(a \in L\) with \(a > 0_L\). Then

\[
\sum_{x : x \vee a = 1_L} \mu(0_L, x) = 0
\]

**Proof:** Now, using the Möbius formula (2) in the first equality, and the Möbius formula (1) together with the observation that \(y \neq 0_L\) in the third we have:

\[
\sum_{x : x \vee a = 1_L} \mu(0_L, x) = \sum_x \mu(0_L, x) \sum_{y \geq x \vee a} \mu(y, 1_L)
\]

\[
= \sum_{y \geq a} \mu(y, 1_L) \sum_{x \leq y} \mu(0_L, x)
\]

\[
= 0 \quad \square
\]
Divisibility Poset

**Proposition 6.5** In the divisibility poset on the positive integers we have

\[
\mu(a, b) = \begin{cases} 
(-1)^k & \text{if } \frac{b}{a} = p_1p_2\ldots p_k \text{ where } p_1, \ldots, p_k \text{ are distinct primes} \\
0 & \text{otherwise}
\end{cases}
\]

(note that \(\mu(a, a) = 1\) since 1 is a product of zero distinct primes).

**Proof:** To compute \(\mu(a, b)\) it suffices to consider the case when \(a = 1\) (the 0 of the lattice). For this we proceed by induction on \(b\). The base case, when \(b = 1\) holds trivially. For the inductive step, choose a prime \(p\) which divides \(b\). Now, by Weisner’s Theorem we have

\[
\mu(0, b) = -\sum_{d \neq b : \gcd(d, b) = b} \mu(0, d) = -\sum_{d \neq b : \gcd(d, b) = b} \mu(0, d)
\]

Now, if \(p^2 | b\) the sum on the right is empty and we have \(\mu(0, b) = 0\) as desired. Otherwise, there is only one possible choice for \(d\), namely \(d = b/p\) and for this choice we have: \(\mu(0, b) = -\mu(0, b/p)\) and now the result follows by induction. \(\square\)

**Number Theoretic Möbius:** Noting that the above function depends only on \(b/a\), we now define a one parameter Möbius function on the positive integers as follows:

\[
\mu(n) = \begin{cases} 
(-1)^k & \text{if } n = p_1p_2\ldots p_k \text{ where } p_1, \ldots, p_k \text{ are distinct primes} \\
0 & \text{otherwise}
\end{cases}
\]

**Proposition 6.6 (Solution to Problem 2)** For every positive integer \(n\) we have

\[
\frac{\phi(n)}{n} = \sum_{d|n} \frac{\mu(d)}{d}
\]

**Proof:** We have shown \(\sum_{d|n} \phi(d) = n\) so by Möbius inversion (applied with \(f = \phi\) and \(g\) the identity) we get

\[
\frac{\phi(n)}{n} = \frac{1}{n} \sum_{d|n} \mu(d, n)d = \sum_{d|n} \mu(n/d) \frac{d}{n} = \sum_{d|n} \frac{\mu(d)}{d}
\]

**Theorem 6.7** For a prime power \(q\), the number of monic irreducible polynomials of degree \(n\) over \(\mathbb{F}_q\) which we denote by \(N_n\) is given by

\[
N_n = \frac{1}{n} \sum_{d|n} \mu(n/d)q^d
\]
Proof: We previously proved that $q^n = \sum_{d|n} dN_d$. Applying Möbius inversion (for the functions $f(n) = nN_n$ and $g(n) = q^n$) gives us

$$nN_n = \sum_{d|n} \mu(n/d)q^d$$

from which the result follows. \[\Box\]

Note: There is a connection between the Möbius & zeta functions for a poset and the number theoretic Möbius and Riemann zeta function. This is given by the following theorem (the Riemann $\zeta$ function is given by $\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s}$ and defined on all $s \in \mathbb{C}$ with $\Re(s) > 1$).

**Theorem 6.8**

$$\frac{1}{\zeta(s)} = \sum_{n=1}^\infty \frac{\mu(n)}{n^s}$$

Proof: Letting $p_i$ denote the $i^{th}$ prime, the Euler product expression for $\zeta$ is

$$\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s}$$

$$= 1 + \frac{1}{2^s} + \frac{1}{3^s} \ldots$$

$$= \left(1 + \frac{1}{p_1^s} + \frac{1}{p_2^{2s}} + \ldots\right) \left(1 + \frac{1}{p_2^s} + \frac{1}{p_2^{2s}} + \ldots\right) \ldots$$

$$= \prod_{i=1}^\infty \frac{1}{1 - \frac{1}{p_i^s}}$$

Using this we find

$$\frac{1}{\zeta(s)} = \prod_{i=1}^\infty \left(1 - \frac{1}{p_i^s}\right)$$

$$= \sum_{n=1}^\infty \frac{\mu(n)}{n^s} \quad \Box$$
Subspace Poset

Gaussian Numbers: If $n, k$ are positive integers and $q$ is a power of a prime, we define

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

**Proposition 6.9** The number of $k$-dimensional subspaces of $\mathbb{F}_q^n$ is $\begin{bmatrix} n \\ k \end{bmatrix}_q$. More generally, the number of $k$-dimensional subspaces of $\mathbb{F}_q^n$ containing a given $\ell$-dimensional subspace is given by $\begin{bmatrix} n-\ell \\ k-\ell \end{bmatrix}_q$.

*Proof:* Fix subspaces $U \subseteq V \subseteq \mathbb{F}_q^n$ and suppose that $\text{dim}(U) = a$ and $\text{dim}(V) = b$. The number of maximal chains in the subspace poset with smallest element $U$ and largest element $V$ depends only on $a, b$. To construct such a chain, we may choose the next term above $U$ by taking the space spanned by $U$ and any of the $q^b - q^a$ vectors in $V \setminus U$. However, every such subspace will be formed by $q^{b-a+1} - q^a$ different vectors, so the total number of ways to choose this next term is $\frac{q^{b-a+1} - q^a}{q-1}$. Continuing in this manner we find that the total number of chains between $U$ and $V$ is

$$\frac{(q^{b-a} - 1)(q^{b-a-1} - 1) \ldots (q - 1)}{(q-1)^{b-a}}$$

This depends only on the difference of $a$ and $b$ so defining $M(c) = \frac{(q^{c-1})(q^{c-2} - 1) \ldots (q - 1)}{(q-1)^c}$ we have that $M(c)$ is the number of maximal chains between any pair of subspaces $U \subseteq V$ with $\text{dim}(V) - \text{dim}(U) = c$. Now, the total number of maximal chains is $M(n)M(n-k)$ and the number containing any fixed $k$-dimensional subspace is $M(k)M(n-k)$ which gives us

$$\# \{ V \leq \mathbb{F}_q^n : \text{dim}(V) = k \} = \frac{M(n)}{M(k)M(n-k)} = \frac{(q^n - 1)(q^{n-1} - 1) \ldots (q - 1)}{(q^k - 1)(q^{k-1} - 1) \ldots (q - 1)(q^{n-k} - 1)(q^{n-k-2} - 1) \ldots (q - 1)}$$

$$= \frac{(q^n - 1)(q^{n-1} - 1) \ldots (q^{n-k+1} - 1)}{(q^k - 1)(q^{k-1} - 1) \ldots (q - 1)}$$

$$= \begin{bmatrix} n \\ k \end{bmatrix}_q$$

A similar computation yields the more general result that the number of $k$-dimensional subspaces containing a fixed $\ell$-dimensional subspace is $\begin{bmatrix} n-\ell \\ k-\ell \end{bmatrix}_q$. □
Proposition 6.10 The M"obius function for the subspace poset of $\mathbb{F}_q^n$ is

$$
\mu(U, V) = \begin{cases} 
(-1)^k q^{\binom{k}{2}} & \text{if } U \subseteq V \text{ and } \dim(V) - \dim(U) = k \\
0 & \text{otherwise}
\end{cases}
$$

Proof: Again we may assume without loss that $U = \{0\}$ and that $V = \mathbb{F}_q^n$. Now, we proceed by induction on $n$. As a base, note the result is trivial if $n = 1$. For the inductive step, choose a 1-dimensional subspace $P \subseteq V$ and apply Weisner’s theorem to get

$$
\mu(0, V) = - \sum_{W < V: W \lor P = V} \mu(0, W)
$$

$$
= - \sum_{W < V: P \not\subseteq W, \dim(W) = n-1} (-1)^{n-1} q^{\binom{n-1}{2}}
$$

$$
= - \left( \binom{n}{n-1}_q - \binom{n-1}{n-2}_q \right) (-1)^{n-1} q^{\binom{n-1}{2}}
$$

$$
= - \left( \binom{n}{1}_q - \binom{n-1}{1}_q \right) (-1)^{n-1} q^{\binom{n-1}{2}}
$$

$$
= - \left( \frac{q^n - 1}{q - 1} - \frac{q^{n-1} - 1}{q - 1} \right) (-1)^{n-1} q^{\binom{n-1}{2}}
$$

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

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

Answer to Problem 3: Let $V = \mathbb{F}_q^k$. For every subspace $U \subseteq V$ let $f(U)$ denote the number of surjective linear mappings from $\mathbb{F}_q^n$ to $U$ and let $g(U)$ denote the number of arbitrary linear mappings from $\mathbb{F}_q^n$ to $U$. Then $g(U) = q^{\dim(U)}$ so by M"obius inversion we have

$$
f(V) = \sum_{U \subseteq V} \mu(U, V) g(U)
$$

$$
= \sum_{i=0}^{k} \binom{k}{i}_q (-1)^{k-i} q^{\binom{k-i}{2} + ni}
$$