Math 116 Professor: Padraic Bartlett

Lecture 4: Möbius Inversion

Weeks 5-7 *UCSB 2015*

In last week's talks, we looked at the concept of **inclusion-exclusion**, and showed how the sieve method that we developed for generating functions can be thought of as a generalization of this idea. In these lectures, we look at another way to generalize the idea of inclusion-exclusion: the concept of Möbius inversion!

1 Posets

1.1 Preliminary Concepts

Definition. Take any set S. A **relation** R on this set S is a map that takes in ordered pairs of elements of S, and outputs either true or false for each ordered pair.

You know many examples of relations:

- Equality ($=$), on any set you want, is a relation; it says that $x = x$ is true for any x, and that $x = y$ is false whenever x and y are not the same objects from our set.
- "Mod n" ($\equiv \mod n$) is a relation on the integers: we say that $x \equiv y \mod n$ is true whenever $x - y$ is a multiple of n, and say that it is false otherwise.
- "Less than" ($<$) is a relation on many sets, for example the real numbers; we say that $x < y$ is true whenever x is a smaller number than y (i.e. when $y - x$ is positive,) and say that it is false otherwise.
- "Beats" is a relation on the three symbols (rock, paper, scissors) in the game Rock-Paper-Scissors. It says that the three statements "Rock beats scissors," "Scissors beats paper," and "Paper beats rock" are all true, and that all of the other pairings of these symbols are false.

Definition. A **partially ordered set** $P = (X, <)$, often called a **poset** for short, is a set X with a relation $<$ on P that satisfies the following two properties:

- **Antisymmetry**: For all $x, y \in P$, if $x < y$, we do not have $y < x$.
- **Transitivity**: For all $x, y \in P$, if $x < y$ and $y < z$, then we have $x < z$.

Notice that we do **not** inherently know that any two elements are related: if we had the third property that for any $x \neq y$ in P¡ we have $x < y$ or $y < x$, we would have something called a **total** ordering.

For instance, of the four example sets and relations listed above, the only pair that forms a poset is $(\mathbb{R}, <)$.

For another example of a poset, consider the set P of all breakfast foods, with the relation $>$ defined by "is tastier than." For instance, we definitely have

$$\text{(delicious perfect pancakes)} > \text{(horribly burnt pancakes)},$$

so these two objects are comparable. However, some other objects are **not comparable**: i.e.

$$\text{(delicious perfect pancakes)} \text{ , } \text{(delicious perfect french toast)}$$

are two different objects such that neither are really obviously "tastier" than the other; this is OK, because in a poset we do not know that any two elements are comparable. Most things in life that we put orderings on are usually posets.

If a partially ordered set has the property that any two elements in are comparable, we call this a **totally ordered set**. This is infrequently abbreviated to the phrase **toset**, because "toset" is a silly word.

Given a poset $P = (X, <)$, it can be very useful to visualize P by drawing it as a diagram! We do this as follows:

- Let $M_0 = \{x \in P \mid \nexists y \in P \text{ with } x < y\}$; that is, M_0 is the collection of all "maximal" elements. At the top of our paper, draw points in a row, one for each element of M_0.
- Now, take the collection M_1 of all of the elements "directly beneath" M_0; that is, form the set

  $$M_1 = \{x \in P \mid \exists y \in M_0 \text{ with } x < y, \text{ but } \nexists y \in (P \setminus M_0) \text{ with } x < y\}$$

  of all elements with only M_0-objects greater than them under our relation. Draw these elements in a row beneath the M_0 vertices, and draw a line from any element in M_0 to any element in M_1 whenever they are comparable.
- Now, take the collection

  $$M_2 = \{x \in P \mid \exists y \in M_0 \text{ with } x < y, \text{ but } \nexists y \in (P \setminus (M_0 \cup M_1)) \text{ with } x < y\}$$

  of all points directly beneath the M_1 points. Draw these points beneath the M_1 points, and connect points in M_1 to points in M_2 if they are comparable.
- Repeat this until you've drawn all of P!

We call this diagram the **lattice diagram** for our poset.

This is perhaps overly abstract / best understood with an example. Consider the **divisor poset** defined as follows: for any natural number n, let X be the collection of all of the natural numbers that divide n. For any two divisors a, b of n, write $a < b$ if and only if $a|b$ and $a \neq b$, that is if a divides b and $a \neq b$. This creates a poset:

- We have antireflexivity by definition, as we said that $a < b$ cannot hold if $a = b$.
- We have antisymmetry because it is impossible for two numbers a, b to satisfy $a|b, b|a$ and $a \neq b$ simultaneously.
- We have transitivity because for any three numbers a, b, c, if $a|b$ and $b|c$ then a is clearly a factor of c; i.e. $a|c$.

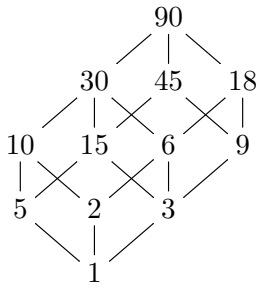
To give an example of this poset, let's consider $n = 90$. The collection of all divisors of 90 are the following:

$$90, 45, 30, 18, 15, 10, 9, 6, 5, 3, 2, 1.$$

If we group elements as suggested above, we get

- $M_0 = 90$,
- $M_1 = 45, 30, 18$,
- $M_2 = 15, 10, 9, 6$,
- $M_3 = 5, 3, 2$,
- $M_4 = 1$.

If we draw the diagram as described, we get the following lattice:



Cool! Notice that the lattice above actually contains all of the information about our poset: if we want to know if any two elements x, y are comparable, we just need to find them in the diagram above, and see if there is a strictly ascending path from the lower of our elements to the higher; the existence of any such path will guarantee (by transitivity) that those elements are comparable! So, for example, $2 < 18$, as the ascending path $2 \to 6 \to 18$ demonstrates. However, 45 and 2 are incomparable, as we cannot find any path from 2 to 45 that is going up. This makes sense; because $2 \not| 45$ and $45 \not| 2$, we know that these two elements are incomparable!

There are some beautiful open questions we can talk about right now with these diagrams. In particular, consider the following two-player game:

- Player 1 picks out an element on the lattice, and deletes it, along with every element "beneath" that element (i.e. if player 1 pick out x, they delete x and every $y < x$ from our lattice.)
- Player 2 then picks out a remaining element in the lattice, and deletes it along with all elements beneath it.
- Players repeat this process until there are no elements left in the lattice. The player who chose the last element loses.

Determining who wins this game on many families of posets, and I believe even on the divisor posets that we've discussed above, is an open problem!

1.2 The Incidence Algebra; Convolution

This, however, is not what I want to focus on here; it's just an interesting diversion. What I'd like us to focus on is actually the following set of definitions:

Definition. Take any poset $P = (X, <)$ on a finite set X. We define the **incidence algebra** of this poset over the real numbers[1] $\mathbb{R}$ as the following:

$$\mathbb{A}(P) = \{f : P^2 \to \mathbb{R} \mid f(x, y) = 0 \text{ whenever } x \not\leq y\}.$$

(If you were concerned about the definition of $\leq$ above; we say that $x \leq y$ if and only if we either have $x < y$ or $x = y$.)

In other words, $\mathbb{A}(P)$ is the collection of all functions that take in two elements of P and output real numbers, such that their output on "incomparable" pairs (i.e. whenever their input is (x, y) for $x \not< y, x \neq y$) is 0.

There are a few commonly-occurring and useful functions that we will want to consider in most incidence algebras:

1. The "zero" map $0(x, y) = 0$, which is in every incidence algebra.
2. The **Kronecker delta function** of P, defined by

$$\delta(x, y) = \begin{cases} 1, & x = y \\ 0, & x \neq y \end{cases}.$$

   This function simply detects when two elements are the same.

3. The **zeta function** of P, defined by

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

   This function simply detects when $x \leq y$.

As always in mathematics, whenever we are given a collection of objects, we will want to know how to combine them! One natural way to combine elements from $\mathbb{A}(P)$ is via addition: that is, given any two elements $f, g \in \mathbb{A}(P)$, we can define

$$(f + g)(x, y) = f(x, y) + g(x, y),$$

where addition on the right is performed in $\mathbb{R}$. For any x, y with $x \not\leq y$, we clearly have $f(x, y) + g(x, y) = 0 + 0 = 0$; so this sum of functions is also in $\mathbb{A}(P)$. We also have a notion of scalar multiplication: given any $c \in \mathbb{R}$ and $f \in \mathbb{A}(P)$, we can define

$$(c \cdot f)(x, y) = c \cdot (f(x, y)),$$

where the multiplication on the right is performed in $\mathbb{R}$. Again, the result is in $\mathbb{A}(P)$; so we have made a set that is closed under addition and scalar multiplication! In fact, it is not hard to see that what we've made here is a **vector space** over $\mathbb{R}$, which is kinda cool.

[1] You can do this over many other fields, if you are so inclined.

This, however, is not the thing I want to talk about; instead of addition or scalar multiplication, I want to define a third operation, called **convolution**, that will let us take a pair of elements of $\mathbb{A}(P)$ and get something that is a sort of "hybrid" between the two. Consider the following definition:

Definition. Take any finite poset P and any $f, g \in \mathbb{A}(P)$. Define $f * g$, the **convolution** of f and g, as follows:

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

The motivation for this definition (especially if you haven't worked with convolutions in other contexts before) is likely unclear at the moment. That's OK; for now, just accept it as a strange thing that we'll make useful in a page or two. For now, however, let's calculate a few examples:

Example. The zero map convolved with any element of $\mathbb{A}(P)$ is just the zero map, as

$$(f * 0)(x, y) = \sum_{z : x \leq z \leq y} f(x, z) \cdot 0(z, y) = 0,$$

for any $x, y \in P$.

Similarly, the delta function convolved with any element $f \in \mathbb{A}(P)$ is just f, as

$$\begin{aligned}
(f * \delta)(x, y) &= \sum_{z : x \leq z \leq y} f(x, z) \cdot \delta(z, y) \\
&= f(x, y) \cdot \delta(y, y) + \sum_{z : x \leq z < y} f(x, z) \cdot \delta(z, y) \\
&= f(x, y) + \sum_{z : x \leq z < y} f(x, z) \cdot 0 \\
&= f(x, y).
\end{aligned}$$

Finally, the zeta function convolved with any element $f \in \mathbb{A}(P)$ is also something with a nice closed form:

$$\begin{aligned}
(f * \zeta)(x, y) &= \sum_{z : x \leq z \leq y} f(x, z) \cdot \zeta(z, y) \\
&= \sum_{z : x \leq z < y} f(x, z) \cdot 1 \\
&= \sum_{z : x \leq z < y} f(x, z).
\end{aligned}$$

In other words, $(f * \zeta)(x, y)$ is just the sum of all of the $f(x, z)$ values for z between x and y, which is a thing we will often want to study.

Whenever we define an operation in mathematics, there are a pair of natural questions we want to ask:

1. Is there an identity? In other words: is there a function $id \in \mathbb{A}(P)$ such that $f * id = f$ for any f?

2. Is this a nice operation? In other words: is it associative? Commutative? Does it distribute over our addition operation?

3. Is it invertible? In other words: for any $f \in \mathbb{A}(P)$, is there a function $f^{-1} \in \mathbb{A}(P)$ such that $f^{-1} * f = id$?

The answer to the first is yes; as we showed above, the delta function is an identity, as $\delta * f = f$ for any f. The second is a question you explore on the HW: in general, you can show that this operation is associative without too much work (and I'll leave the questions of commutativity and distributivity for you to determine!)

The answer to the third, sadly, is a no — as demonstrated above, the zero function has no inverse, as $0 * f = 0$ for any f, and therefore in particular there is no f such that $0 * f = \delta$. In general, if $f \in \mathbb{A}(P)$ is a function such that there is **any** $x \in P$ with $f(x, x) = 0$, it is impossible for f to have an inverse, as for any $g \in \mathbb{A}(P)$ we have

$$f * g(x, x) = \sum_{z : x \leq z \leq x} f(x, z) \cdot g(z, x) = f(x, x) \cdot g(x, x) = 0 \neq \delta(x, x).$$

However, I claim that these are the only kinds of maps that are noninvertible; that is, if $f(x, x) \neq 0$ for all x, I claim that f is invertible. We prove this here:

Theorem. Suppose that P is a finite poset and $f \in \mathbb{A}(P)$ has $f(x, x) \neq 0$ for any $x \in P$. Then there is some $f^{-1} \in \mathbb{A}(P)$ such that $f^{-1} * f = \delta$, the Kronecker delta function.

Proof. Take any such map f. We build f^{-1} inductively, by first defining it on all of the pairs (x, x) with $x \in X$, and then extending it to pairs[2] (x, y) in our poset with $x < y$.

To do the first step: for any $x \in P$, define $f^{-1}(x, x) = \frac{1}{f(x,x)}$. Then, by definition, we have

$$(f * f^{-1})(x, x) = \sum_{z : x \leq z \leq x} f(x, z) \cdot f^{-1}(z, x) = f(x, x) \cdot \frac{1}{f(x, x)} = 1 = \delta(x, x).$$

For any $x < y \in P$, we say that x, y are **distance** k from each other, for some $k \in \mathbb{N}$, if and only if the following holds:

- There is a set of $k - 1$ elements $z_1, \ldots z_{k-1}$ such that $x < z_1 < z_2 < \ldots z_{k-1} < y$. (If $k = 1$, then we just ask that $x < y$; if $k = 0$ we just ask that $x = y$.)

- There are no sets of k elements $z_1, \ldots z_k$ such that $x < z_1 < \ldots < z_k < y$.

[2] We know that pairs (x, y) with $x \leq y$ are the only values we need to define f^{-1} on; this is by definition, as because f^{-1} is an element of $\mathbb{A}(P)$ we know that $f^{-1}(x, y) = 0$ for all $x \not\leq y$.

If you revisit our lattice-picture from before, we can define the distance between x, y to be k if and only if $x \in M_i, y \in M_{i+k}$ for some i, where the M_i's were the "levels" of our lattice.

We now define our f^{-1} inductively on the distances between x and y, with our base case (where the distance is 0) done for us already.

Inductively, if we have defined f^{-1} on all of the pairs (x, z) with distance at most k, extend this to any (x, y) at distance $k + 1$ by the following definition:

$$f^{-1}(x, y) = -\frac{1}{f(y, y)} \cdot \left(\sum_{x \leq z < y} f^{-1}(x, z) f(z, y) \right).$$

Notice that the sum on the inside is over all z with $x \leq z < y$, where the right-inequality is specifically a strict inequality; as a result, if the distance from x to y is $k + 1$, the set of all (x, z) pairs in this sum have distance at most k, as the z-part is always less than y. So this sum is actually well-defined, by induction! So we've built f^{-1}, and defined it for all values $x \leq y \in P$.

Our last step, then, is verifying that we actually have made an inverse: that is, that $(f^{-1} * f)(x, y) = 0$ for $x < y$. This is straightforward from our definitions:

$$\begin{aligned}
(f^{-1} * f)(x, y) &= \sum_{z : x \leq z \leq y} f^{-1}(x, z) \cdot f(z, x) \\
&= \left(\sum_{z : x \leq z < y} f^{-1}(x, z) \cdot f(z, y) \right) + f^{-1}(x, y) f(y, y) \\
&= \left(\sum_{z : x \leq z < y} f^{-1}(x, z) \cdot f(z, y) \right) - \frac{1}{f(y, y)} \cdot \left(\sum_{x \leq z < y} f^{-1}(x, z) f(z, y) \right) f(y, y) \\
&= \left(\sum_{z : x \leq z < y} f^{-1}(x, z) \cdot f(z, y) \right) - \left(\sum_{x \leq z < y} f^{-1}(x, z) f(z, y) \right) \\
&= 0 = \delta(x, y).
\end{aligned}$$

□

Success! It bears noting that this left inverse is a right inverse, as well:

Theorem. Suppose that G is a set with an operation $* : G \times G \to G$ that is associative, has an identity, and has left inverses. Then its left inverses are also right inverses: that is, for any $f \in G$, if f^{-1} is an element such that $f^{-1} * f = id$, then $f * f^{-1} = id$ as well.

Proof. Take any $f \in G$. Let f^{-1} be the left inverse of f; by definition, we have $f^{-1} * f = id$. We want to show that

$$f * f^{-1} = id.$$

To do this: we know that any element of G has a left inverse. So, in particular, the element $f * f^{-1}$ has a left inverse; call it $(f * f^{-1})^{-1}$. Again, by definition, we have $(f * f^{-1})^{-1} * (f *$

$f^{-1}) = id$. Then, notice that several applications of our associative properties and identity property gives us that

$$\begin{aligned} f * f^{-1} &= id * (f * f^{-1}) \\ &= ((f * f^{-1})^{-1} * (f * f^{-1})) * (f * f^{-1}) \\ &= (f * f^{-1})^{-1} * (f * f^{-1} * f * f^{-1}) \\ &= (f * f^{-1})^{-1} * (f * (f^{-1} * f) * f^{-1}) \\ &= (f * f^{-1})^{-1} * (f * (id) * f^{-1}) \\ &= (f * f^{-1})^{-1} * (f * f^{-1}) \\ &= id. \end{aligned}$$

In other words, we've proven our claim! □

The reason we care about this is the following:

Definition. Take the zeta function ζ, as defined before. By definition, $\zeta(x, x) = 1$ for all x, so this function has an inverse! Call this function the **Möbius function** $\mu = \zeta^{-1}$. By definition, we have

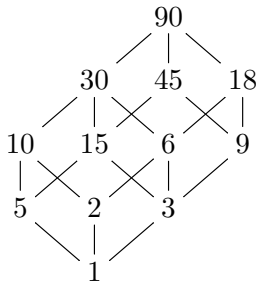
$$\mu(x, x) = \frac{1}{\zeta(1, 1)} = 1,$$

for any $x \in P$; as well, for any $x < y$, we have

$$\mu(x, y) = -\sum_{x \leq z < y} \mu(x, z).$$

Finally, because μ is in our incidence algebra, we have $\mu(x, y) = 0$ whenever $x \not\leq y$.

To give an example, let's return to the factorization poset for 90 from before, and calculate $\mu(1, d)$ for all of the divisors of 90:



We know that by definition,

- $\mu(1, 1) = 1$.

Therefore, we can use this to calculate $\mu(1, 2), \mu(1, 3), \mu(1, 5)$:

- $\mu(1, 2) = -\sum_{1 \leq z < 2} \mu(1, z) = -\mu(1, 1) = -1$.

- Similar logic holds for all of the $\mu(1,p)$, where p is any prime factor of μ; if we're summing over all of the factors betwen 1 and p not including p, then the only factor we count is 1; i.e. our sum is just $-\mu(1,1) = -1$.

Now, let's calculate $\mu(1,pq)$ for any two primes p,q; that is, let's find $\mu(1,d)$ for $d = 10, 15, 6, 9$.

- For $d = 10, 15, 6$ — that is, for $d = pq$ for two distinct primes p, q — we have $\mu(1,d) = -(\mu(1,1) + \mu(1,p) + \mu(1,q)) = -1 + 1 + 1 = 1$.
- For $d = 9$ — that is, for $d = p^2$ for some prime p — we have $\mu(1,d) = -(\mu(1,1) + \mu(1,p)) = 0$.

We can now calculate $\mu(1,pqr)$ for any three primes pqr; that is, we can find $\mu(1,d)$ for $d = 30, 45, 18$;

- For $d = p^2q$ for two distinct primes, we have $\mu(1,d) = -(\mu(1,1) + \mu(1,p) + \mu(1,q) + \mu(1,pq) + \mu(1,p^2)) = 0$.
- For $d = pqr$ for three distinct primes, we have $\mu(1,d) = -(\mu(1,1) + \mu(1,p) + \mu(1,q) + +\mu(1,r) + \mu(1,pq) + \mu(1,pr) + \mu(1,qr)) = -1$.

Finally, we can use this to find $\mu(1,60)$:

$$\begin{aligned}\mu(1,60) = -\sum_{1\le z<60} \mu(1,z) = &-(\mu(1,1) + \mu(1,2) + \mu(1,3) + \mu(1,5) \\ &+\mu(1,10) + \mu(1,15) + \mu(1,6) + \mu(1,9) \\ &+\mu(1,30) + \mu(1,45) + \mu(1,18)) \\ =&-(1-1-1-1+1+1+1+0-1+0+0) \\ =&0.\end{aligned}$$

By itself, this may not look too useful. The following theorem, however, may make this theorem look more useful:

Theorem. Let P be any poset, and let e be any function $P \to \mathbb{R}$. Suppose that P has a unique minimal element: that is, there is some $m \in P$ such that for all $x \in P, m < x$.

Define the function $n : P \to \mathbb{R}$ as follows: for any $a \in P$, set

$$n(a) = \sum_{x \le a} e(x).$$

Then we can "invert" the formula above: that is, for any $a \in P$, we have

$$e(a) = \sum_{x \le a} n(x)\mu(x,a),$$

where μ is the Möbius function.

Proof. Take any function $e : P \to \mathbb{R}$, and define $n : P \to \mathbb{R}$ as above. Let m be the unique minimal element of P. Define the functions $f, g \in \mathbb{A}(P)$ as follows:

- $f(m, a) = e(a)$ for all a, and $f(x, y) = 0$ for all other undefined values.
- $g(m, a) = n(a)$ for all a, and $f(x, y) = 0$ for all other undefined values.

Now, by definition, we have for any $a \in P$,

$$\begin{aligned} g(m, a) = n(a) &= \sum_{x \leq a} e(x) \\ &= \sum_{x \leq a} f(m, x) \\ &= \sum_{m \leq x \leq a} f(m, x)\zeta(x, a) \\ &= (f * \zeta)(m, a), \end{aligned}$$

where ζ is the zeta function from before. As well, we know that for any $x, y \in P$ with $x \neq m$, that $g(x, y) = 0$ by definition, and similarly that

$$(f * \zeta)(x, y) = \sum_{z : x \leq z \leq y} f(x, z) \cdot \zeta(z, y) = \sum_{z : x \leq z \leq y} 0 \cdot \zeta(z, y) = 0,$$

so they are equal in fact for every $x, y \in P$. In other words, we've shown that $g = f * \zeta$.

But we know that the inverse of ζ is the Möbius function μ: that is, we know that $g * \mu = f$! In other words, we know that

$$\begin{aligned} e(a) = f(m, a) &= (g * \mu)(m, a) \\ &= \sum_{m \leq x \leq a} g(m, x)\mu(x, a) \\ &= \sum_{x \leq a} n(x)\mu(x, a), \end{aligned}$$

as claimed. □

This turns out to be a pretty big deal, as we explore in our next section:

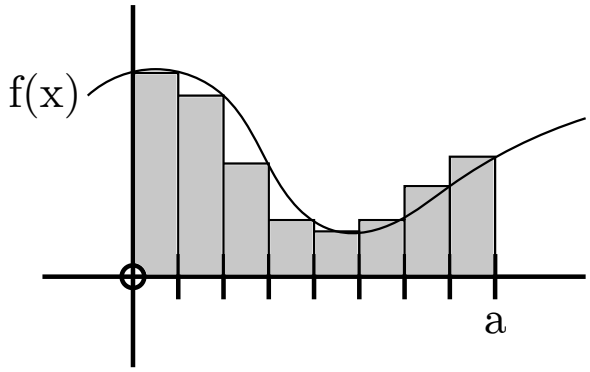
2 Möbius Inversion

2.1 Möbius Inversion and Calculus

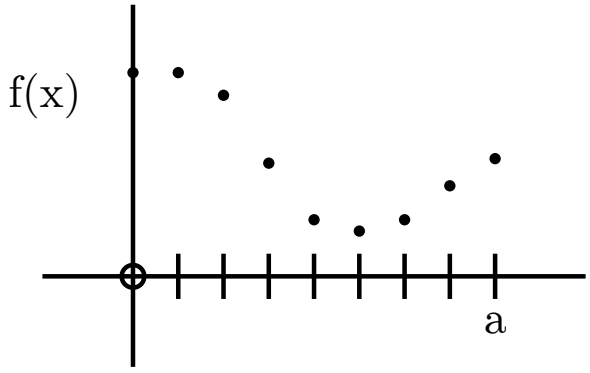
One (surprising) application of Mobius inversion here is to **calculus**! Specifically: think back to the first definition you saw for the integral of a function. If you're like most people, you saw the integral defined first as the Riemann integral: that is, if we have a function $f : [0, a] \to \mathbb{R}$ and points $0 = x_1 < x_1 < x_2 < \ldots < x_{n-1} < x_n = a$, we can approximate the integral of f on $[0, a]$ as the "area" under all of the rectangles $[x_i, x_{i+1}] \times [0, f(x_i)]$. That is, we can write

$$\int_0^a f(x)dx \approx \sum_{i=1}^{n}(x_i - x_{i+1}) * f(x_i).$$

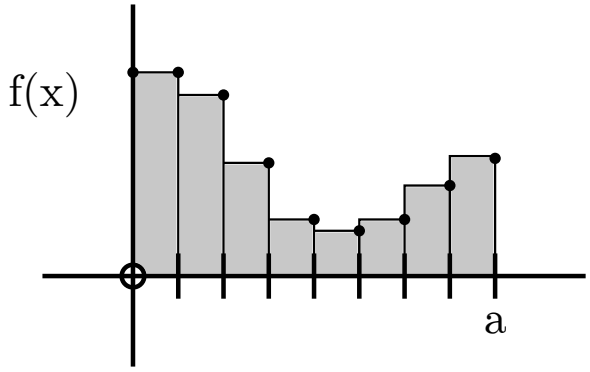
This is something that is perhaps best understood via a picture: we're saying that the area under $f(x)$ is approximated by the gray rectangles below!



A natural question to ask, given that we're in a discrete mathematics class, is how we can define an integral over other sorts of intervals. For example, suppose we had a function $f : \mathbb{N} \to \mathbb{R}$. Then we don't have a curve, but rather a set of points:



In this sense, we can still think of the "integral" of this function as the area of these rectangles: that is, we can define the integral of f from 0 to a as the area under the rectangles in the picture here:



This has a nice definition: because the base of each of these rectangles is 1 (as our function is defined on the natural numbers,) we get the formal definition

$$\int_0^a f(x)dx = \sum_{k=0}^{a-1} f(k),$$

for any function $f : \mathbb{N} \to \mathbb{R}$.

This looks interesting from a Möbius function perspective! In particular, if we think of $\mathbb{N}$ as a poset under its normal ordering $<$, we have two functions $f(x) : \mathbb{N} \to \mathbb{R}$ and $F(x) = \sum_{k=0}^{a-1} f(k)$, one of which is just a sum of the other! Therefore, we can conclude that

$$f(x) = \sum_{k=0}^{a-1} F(x)\mu(x, a-1);$$

that is, we can relate a function's value to its integral! To understand the way this works, let's actually calculate what μ is. First, we know by definition that $\mu(a, a) = 1$ for any $a \in \mathbb{N}$. Then, again by definition, we have that

$$\mu(a, a+1) = -\sum_{a \le x < a+1} \mu(a, x) = -\mu(a, a) = -1,$$

and

$$\mu(a, a+2) = -\sum_{a \le x < a+2} \mu(a, x) = -(\mu(a, a) + \mu(a, a+1)) = -(1-1) = 0,$$

and in general by induction that

$$\begin{aligned}\mu(a, a+k) = -\sum_{a \le x < a+k} \mu(a, x) &= -\left(\mu(a, a) + \mu(a, a+1) + \mu(a, a+2) + \ldots + \mu(a, a+k)\right) \\ &= -\left(1 - 1 + 0 + \ldots + 0\right) = 0.\end{aligned}$$

So on the poset given by $\mathbb{N}$, we have that $\mu(a,a) = 1, \mu(a,a+1) = -1$, and $\mu(x,y) = 0$ otherwise; therefore we have

$$\begin{aligned} f(x) =& \sum_{k=0}^{a-1} F(x)\mu(x, a-1) \\ =& F(0)\mu(0,a-1) + F(1)\mu(1,a-1) + \ldots + F(a-2)\mu(a-2,a-1) + F(a-1)\mu(a-1,a-1) \\ =& 0 + \ldots + 0 + -F(a-2) + F(a-1) \\ =& F(a-1) - F(a-2). \end{aligned}$$

In other words, we've "undone" integration by looking at $F(a-1) - F(a-2)$; if we write this in the form

$$\frac{F(a-1) - F(a-2)}{(a-1) - (a-2)},$$

we can perhaps recognize this as a discrete version of the **derivative**! Möbius inversion: the strangest way of studying the fundamental theorem of calculus.

We explore more of the connections between Möbius inversion and calculus operations on the homework. Here, however, we turn to a second application that we've already studied with different tools: the technique of **inclusion-exclusion**!

2.2 Inclusion-Exclusion via Möbius Inversion

Theorem. (Inclusion-exclusion.) Suppose that X is any finite set and $A_1, \ldots A_n$ are a collection of subsets of X. Then the number of elements in X but not in any of the subsets A_i can be expressed in terms of the sizes of the A_i sets and their intersections. In specific, we have the following equality:

$$\left| X \setminus \left(\bigcup_{k=1}^{n} A_k \right) \right| = |X| + \sum_{k=1}^{n} \left((-1)^k \cdot \sum_{1 \leq i_1 < \ldots < i_k \leq n} |A_{i_1} \cap \ldots \cap A_{i_k}| \right).$$

We have proven this theorem twice; once directly by using an "overcounting-undercounting" argument, and once by using the method of sieves! In this section, we prove it a third time via **Möbius inversion**.

To start, we should come up with an appropriate poset to study! We do this here:

Definition. Take any set A. Look at the power set $\mathcal{P}(A)$; that is, the collection of all of the subsets of A. Define a partial ordering on $\mathcal{P}(A)$ by using the inclusion relation: that is, for any $X, Y \in \mathcal{P}(A)$, use the the relation $X \subsetneq Y$ to create our poset. (Check for yourself that this relation satisfies the antisymmetric/antireflexive/transitive properties that we ask for in a poset!)

We call this poset the **Boolean lattice**[3] on A.

[3] A **lattice** is a special kind of poset, in which any two elements have a unique smallest common upper bound and a unique largest common lower bound. For our purposes here, lattice is synonymous with poset; later, if we have time, we might work more formally with lattices!

A specific kind of Boolean lattice is the one where we let $A = \{1, 2, 3, \dots n\}$: we call this lattice B_n for short.

Given this lattice, a natural question to ask is the following: how is the Möbius function defined? We answer this in the following theorem:

Theorem. Take any set A. The Möbius function on the Boolean lattice associated to A is defined as follows: for any two subsets $X, Y \subseteq A$, we have

$$\mu(X, Y) = \begin{cases} 0, & X \not\subseteq Y, \\ (-1)^{|Y \setminus X|}, & X \subseteq Y. \end{cases}$$

Proof. We prove this by using the definition of the Möbius function and the technique of induction. Recall that on an arbitrary poset, we have

$$\mu(X, Y) = \begin{cases} 0, & X \not\subseteq Y, \\ 1, & X = Y, \\ -\sum_{X \subseteq Z \subsetneq Y} \mu(X, Z), & X \subsetneq Y. \end{cases}$$

Therefore, by definition we know that for $X \not\subseteq Y$ our claim holds, and also for $X = Y$ our claim holds (as $(-1)^{X \setminus Y} = (-1)^0 = 1$ for $X = Y$.)

We proceed by induction on the size of $X \setminus Y$. We've already done our base case of $|X \setminus Y| = 0$; so let's proceed to our inductive step. We assume that for any $X \subseteq Y$ with $|X \setminus Y| \leq k$, our claim holds; that is, that $\mu(X, Y) = (-1)^{|Y \setminus X|}$.

We want to prove that our claim holds for $k + 1$; that is, for any $X \subset Y$ with $|X \setminus Y| = k + 1$. We do this by applying the definition of the Möbius function, and then using our inductive assumption:

$$\begin{aligned} \mu(X, Y) &= -\sum_{X \subseteq Z \subsetneq Y} \mu(X, Z) \\ &= -\left(\sum_{X \subseteq Z \subsetneq Y} (-1)^{|Z \setminus X|} \right) \\ &= -\left(\left(\sum_{X \subseteq Z \subseteq Y} (-1)^{|Z \setminus X|} \right) - (-1)^{|Y \setminus X|} \right), \end{aligned}$$

where our last step was just adding and subtracting the term $(-1)^{|Y \setminus X|}$ to make the sum cleaner.

We now can use a result we proved on the midterm:

Problem. Take any two finite sets A, B such that $A \subseteq B$. Prove that

$$\sum_{A \subseteq K \subseteq B} (-1)^{|K \setminus A|} = \begin{cases} 1, & A = B, \\ 0, & A \neq B. \end{cases}$$

If you don't remember how to solve this, here's a solution!

Solution. In the case where $A = B$, the above equation is immediate, as the only subset K between A and B is A, and $A \setminus A = \emptyset$ has cardinality 0.

When $A \subsetneq B$, let $n = |B \setminus A|$. Any subset $A \subseteq K \subseteq B$ has $|K \setminus A| = k$, for some $0 \leq k \leq n$. Moreover, for fixed k, there are $\binom{n}{k}$-many possible subsets K such that $A \subseteq K \subseteq B$ with $|K \setminus A| = k$; this is because there are n elements in $B \setminus A$, and any way to choose k of them gives us one of our possible subsets.

Therefore, if we group subsets in the sum above by the size of $|K \setminus A$, we get that

$$\sum_{A \subseteq K \subseteq B} (-1)^{|K \setminus A|} = \sum_{k=0}^{n} (-1)^k \binom{n}{k}.$$

But, if you remember the binomial formula, this is just

$$(-1+1)^n,$$

which we know is 0 for any $n \geq 1$. So we've proven our claim!

If we apply this to our expression for μ, we get the following:

$$\begin{aligned}
\mu(X,Y) =& -\left(\left(\sum_{X \subseteq Z \subseteq Y} (-1)^{|Z \setminus X|}\right) - (-1)^{|Y \setminus X|}\right) \\
=& -\left(0 - (-1)^{|Y \setminus X|}\right) \\
=& (-1)^{|Y \setminus X|}.
\end{aligned}$$

So we've proven our claim! □

We can use this result to actually **prove** the inclusion-exclusiopn theorem! We do this here:

Theorem. (Inclusion-exclusion.) Suppose that X is any finite set and $A_1, \ldots A_n$ are a collection of subsets of X. Then we have the following equality:

$$\left| X \setminus \left(\bigcup_{k=1}^{n} A_k \right) \right| = |X| + \sum_{k=1}^{n} \left((-1)^k \cdot \sum_{1 \leq i_1 < \ldots < i_k \leq n} |A_{i_1} \cap \ldots \cap A_{i_k}| \right).$$

Proof. Take our set X, and its n subsets $A_1, \ldots A_n$. Define a collection of properties $P = \{p_1, p_2, \ldots p_n\}$ on this set as follows: for any x, we say that x has property p_i if and only if x is an element in A_i.

Consider the Boolean lattice B_n on $\{1, 2, \ldots n\}$; as noted before, elements $I \in B_n$ are subsets of $\{1, 2, \ldots n\}$. Define a function $e : B_n \to \mathbb{R}$ as follows: for any $I \subset \{1, \ldots n\}$, set

$$e(I) = |\{x \in X \mid x \text{ satisfies } p_i \text{ if and only if } i \in I\}|.$$

In other words, $e(I)$ counts all of the elements that have exactly the properties $\{p_i \mid i \in I\}$ and no others.

As well, set

$$n(I) = |\{x \in X \mid x \text{ satisfies } p_i \text{ for every } i \in I\}|.$$

In other words, $n(I)$ counts all of the elements that have **at least** the properties $\{p_i \mid i \in I\}$; it also counts elements that have properties beyond this set.

On one hand, we can easily see that

$$n(I) = \sum_{J : I \subseteq J \subseteq \{1 \ldots n\}} e(J);$$

the number of things that satisfy **at least** I is just the sum of all of the collections of things that satisfy J, for every $J \supseteq I$.

On the other, we can observe that we know $n(I)$; as shown in our sieve method notes before, for any $I \neq \emptyset$ we have

$$n(I) = \left| \bigcap_{i \in I} A_i \right|,$$

because this is precisely the collection of all elements that satisfy all of the properties $\{p_i \mid i \in I\}$ and perhaps others. (When $I = \emptyset$ we have $n(I) = |X|$, as every element has at least no properties.)

If we use a result on the fourth homework that states a "converse" of the Möbius inversion theorem[4], we can get the following: for any $I \subseteq \{1, \ldots n\}$, we have

$$e(I) = \sum_{J : I \subseteq J} n(J)\mu(I, J).$$

But we know that $\mu(I, J) = (-1)^{|J \setminus I|}$ from our earlier work; so this sum is just

$$e(I) = \sum_{J : I \subseteq J} n(I)(-1)^{|J \setminus I|}.$$

In particular, when $I = \emptyset$, we can plug in our knowledge of what $n(I)$ actually is to get the following formula:

[4]Specifically, you were asked to prove the following theorem:

Theorem. Let P be any poset with a unique maximal element: that is, there is some $M \in P$ such that for all $x \in P, M > x$. Let e be any function $P \to \mathbb{R}$.

Define the function $n : P \to \mathbb{R}$ as follows: for any $a \in P$, set

$$n(a) = \sum_{x \geq a} e(x).$$

Then we can "invert" the formula above: that is, for any $a \in P$, we have

$$e(a) = \sum_{x \geq a} n(x)\mu(a, x).$$

The proof of this theorem is similar to the result we showed earlier in these notes; try to prove it!

$$e(\emptyset) = |X| + \sum_{J:\emptyset\subsetneq J} (-1)^{|J|} \cdot \left|\bigcap_{i\in I} A_i\right|$$

If we group the J-subsets by their sizes (as the expressions above only care about the size of J,) we get the following expression:

$$e(\emptyset) = |X| + \sum_{k=1}^{n} \left((-1)^k \cdot \sum_{I\subset\{1,\ldots n\},|I|=k} \left|\bigcap_{i\in I} A_i\right| \right).$$

But $e(\emptyset)$ is just the collection of things with **no** properties: that is, the collection $|X \setminus (\bigcup_{k=1}^{n} A_k)|$, which is what we want! So we've proven our claim. □

For our third and last set of examples, we turn to one of the first posets we studied in these notes: the **divisor poset**!

2.3 Möbius Inversion and the Divisor Poset

We actually studied some values of the Möbius function on the divisor poset earlier: in fact, we proved that on the $n = 90$ poset, that

- $\mu(1, p) = -1$ for any prime p,
- $\mu(1, pq) = 1$ for any two distinct primes pq,
- $\mu(1, pqr) = -1$ for any three distinct primes pqr, and
- $\mu(1, p^2) = \mu(1, p^2q) = 0$ for any primes p, q.

We generalize these results to the divisor poset for any n here, in the following claim:

Theorem. Take any natural number n, let P be the divisor poset associated to n, and let μ be its Möbius function. Then, for any $x, y \in P$, we have

$$\mu(x, y) = \begin{cases} 0, & x \nmid y, \\ (-1)^k, & y/x \text{ is the product of } k \text{ distinct primes}, \\ 0, & \text{otherwise.} \end{cases}$$

In particular $\mu(x, y) = \mu(1, y/x)$ for any two $x, y \in P$ with $x|y$, and also $\mu(1, d) = 0$ whenever d has a repeated prime factor.

Proof. We start by proving that $\mu(x, y) = \mu(1, y/x)$ for any $x|y \in P$, and prove this by induction on the number of prime factors of y/x. For $y/x = 1$, i.e. we have no prime factors, this is immediate: we have $y = x$ and therefore that $\mu(x, y) = 1 = \mu(1, 1)$.

Inductively, assume our claim holds for all $x|y$ where y/x has at most k prime factors, and take any $x|y$ where y/x has $k + 1$ prime factors.

Notice that by definition, the Möbius function on any poset has the form

$$\mu(x,y) = \begin{cases} 0, & x \not\leq y, \\ 1, & x = y, \\ -\sum_{z \in P : x \leq z < y} \mu(x,z), & x < y. \end{cases}$$

If $x|y \in P$ then by definition we have that $x < y$ in P, as our relation was "is a divisor of." Therefore, for any $x|y \in P$ we have

$$\mu(x,y) = -\sum_{z \in P : x \leq z < y} \mu(x,z).$$

As well, because 1 is a divisor of everything, we have

$$\mu(1, y/x) = -\sum_{z \in P : 1 \leq z < y} \mu(1, z/x).$$

By induction, we know that $\mu(x,z) = \mu(1, z/x)$, and thus we've proven our claim!

We use proofs by induction on the number of prime factors of y/x to prove our other claims as well: specifically, that

$$\mu(x,y) = \begin{cases} 0, & x \not| y, \\ (-1)^k, & y/x \text{ is the product of } k \text{ distinct primes}, \\ 0, & \text{otherwise.} \end{cases}$$

(We use induction here, and in general do this with most Möbius function proofs, because the Möbius function is defined recursively in terms of its values at earlier stages! Whenever you see a recursively-defined object, induction is almost always the most natural proof technique to apply.)

As before, our base case (when $y/x = 1$ is immediate): in this case we have $x = y$ and therefore that $\mu(x,y) = 1 = (-1)^0$. Also by definition we know that $\mu(x,y) = 0$ whenever $x \not| y$; so it suffices to study the case when $x|y, x \neq y$. By our work above, we can simply study the values of $\mu(1,d)$ for any $d \in P$, as $\mu(x,y) = \mu(1, y/x)$ for any $x|y$.

In this case, assume inductively that $\mu(1,d) = (-1)^l$ for all d with $\leq k$ prime factors, and take any d with $k+1$ prime factors.

There are two cases:

1. None of the prime factors of d are repeated. In this case, for any $z|d, z \neq d$ none of the prime factors of x are repeated as well: so, when we calculate

   $$\mu(1,d) = -\sum_{z \in P : 1 \leq z < d} \mu(1,z),$$

   we can use induction to notice that $\mu(1,z) = (-1)^l$, where l is the number of prime factors of z.

   Suppose that we group z's above by the number of prime factors that each z has: that is, we write

   $$\mu(1,d) = -\sum_{l=0}^{k} \sum_{\substack{z \in P : 1 \leq z < d, \\ z \text{ has l prime factors}}} (-1)^l.$$

For any l, how many elements z of our poset are divisors of d and have l prime factors? Well: there are $k+1$ prime factors of d in total, and we are asking for all of the ways to pick out a subset of l of them to make a number, so there are $\binom{k+1}{l}$-many ways to do this in total! This gives us

$$\mu(1,d) = -\sum_{l=0}^{k} \binom{k+1}{l}(-1)^l.$$

As used many many times in proofs in the past, we know from the binomial theorem that

$$0 = (-1+1)^{k+1} = -\sum_{l=0}^{k+1} \binom{k+1}{l}(-1)^l.$$

Subtracting $\binom{k+1}{k+1}(-1)^{k+1}$ from both sides yields

$$\Rightarrow -(-1)^{k+1} = -\binom{k+1}{k+1}(-1)^{k+1} = \sum_{l=0}^{k} \binom{k+1}{l}(-1)^l.$$

But this is just $-\mu(1,d)$ by our earlier work! Therefore, we've proven our claim: that $\mu(1,d) = (-1)^{k+1}$.

2. d has a repeated prime factor; call it p. In this case, for any $z|d, z \neq d$, if z has p^2 as a factor we know that by induction $\mu(1,z) = 0$; so we can ignore all of these values of z! In particular, this lets us write

$$\mu(1,d) = -\left(\sum_{\substack{z \in P : 1 \leq z < d, \\ p \text{ is not a factor of } z}} \mu(1,z) \right) - \left(\sum_{\substack{z \in P : 1 \leq z < d, \\ p \text{ is a factor of } z, \\ p^2 \text{ is not a factor of } z}} \mu(1,z) \right),$$

by getting rid of all of the other terms, as they're zero anyways.

Also by induction, we can get rid of all of the terms $\mu(1,z)$ where z is a multiple of q^2 for any other prime q:

$$\mu(1,d) = -\left(\sum_{\substack{z \in P : 1 \leq z < d, \\ p \text{ is not a factor of } z, \\ z \text{ has no repeated prime factors}}} \mu(1,z) \right) - \left(\sum_{\substack{z \in P : 1 \leq z < d, \\ p \text{ is a factor of } z, \\ p^2 \text{ is not a factor of } z, \\ z \text{ has no repeated prime factors}}} \mu(1,z) \right).$$

We can "pair up" terms in the first sum with terms in the second sum as follows: to any $z = q_1 \cdot \ldots \cdot q_l$ in the first sum, where the q_i are all distinct primes, associate

$z' = q_1 \cdot \ldots \cdot q_l \cdot p$ in the second sum. This pairing covers all of the terms in each sum, as we can write any z used in the first sum in the form $q_1 \cdot \ldots \cdot q_l$ for some primes q_i, and similarly write any z used in the second sum in the form $q_1 \cdot \ldots \cdot q_l \cdot p$; as well, it's a bijective pairing.

But what does this mean? Well: for any $z = q_1 \cdot \ldots \cdot q_l$, we know by induction that $\mu(1, z) = (-1)^z$. As well, for any $z' = q_1 \cdot \ldots \cdot q_l \cdot p$, we know by induction that $\mu(1, z') = (-1)^{l+1} = -(-1)^l$. So, when we add these sums together, **everything cancels out**, because every term in the first sum has a corresponding opposite in the second! In other words, we get $\mu(1, d) = 0$, as claimed.

□

On the homework this week, you study some applications of this Möbius function, and in particular prove the following two claims:

Theorem. For any $n \in \mathbb{N}$,

$$n = \sum_{d \in \mathbb{N} : d | n} \varphi(d).$$

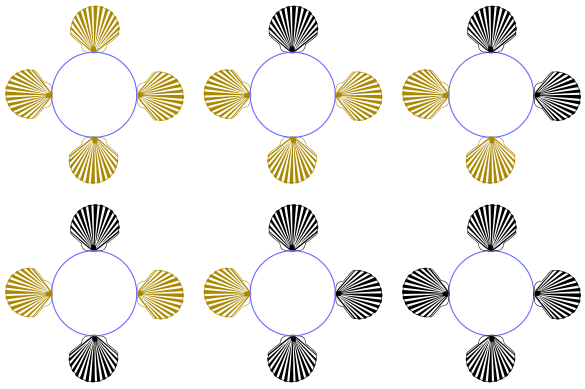
Theorem. For any $n \in \mathbb{N}$,

$$\frac{\varphi(n)}{n} = \sum_{d \in \mathbb{N} : d | n} \frac{\mu(1, d)}{d}.$$

We look at a third problem on the divisor poset here, that can also be studied with Möbius inversion:

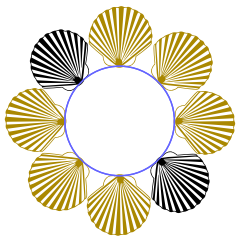
Question. Over the weekend, you collected a stack of seashells from the seashore. Some of them are tan and some are black; you have tons of each color.

You want to arrange them in a necklace! You consider two necklaces to be the same if one can be rotated so that it is the other (you don't allow flips, though, because then your seashells would be backwards.) For example, here are all of the distinct necklaces you can make with four shells:



How many different necklaces on n shells can you make, for any n?

Answer. We start off by making some useful definitions. Given any necklace N, we define the **period** of that necklace as the smallest natural number d such that shifting N counterclockwise d places gives us a necklace that is equal to N. So, for example, the following necklace on 8 shells has period 4:



The reason we do this is the following. Our problem, as stated, seems somewhat tricky; so let's first consider an easier problem! Suppose that we don't consider two necklaces as being the same if they can be rotated into each other. Then the total number of necklaces we can make on n shells is just 2^n; we have two choices for each shell's color, and n shells in total.

Now, for any $d \in \mathbb{N}$, let $f(d, n)$ denote the collection of all period-d necklaces on n shells that are inequivalent under rotation. Notice that $f(d, n) = 0$ for any $d \not| n$, as we cannot have a period that is not a divisor of the total length of our necklace; so the only values we care about are those d such that $d|n$. For $n = 4$, in particular, we have that $f(1, 4) = 2, f(2, 4) = 1, f(4, 4) = 3$ by looking at the six necklaces we drew earlier.

If we consider the function $g(d, n) = d \cdot f(d, n)$, then, we can see that $g(d, n)$ is simply the number of necklaces of period d of length n; we've taken all of the inequivalent necklaces and for each one, generated all of the different equivalent necklaces under rotation! Therefore, if we sum up the values of $g(d, n)$ for all of the divisors of d, we should get the collection of all possible necklaces: that is, we get

$$2^n = \sum_{d|n} g(d, n) = \sum_{d|n} df(d, n). \tag{1}$$

This almost looks like something we can apply Möbius inversion to; but the left-hand side isn't a function on our poset yet, while the right-hand side takes in two arguments! We can fix this with the following observation:

Observation. For any n, d, $f(d, n) = f(d, d)$. That is; the collection of all inequivalent necklaces on n beads with period d is the same size as the collection of all inequivalent necklaces on n/d beads with period n/d.

Proof. Write the collection of all rotation-inequivalent necklaces on n shells with period d as the set of ordered n-strings

$$\mathcal{N}_{n,d} = \{(s_1, \ldots s_n) \mid s_i \in \{W, B\}\}.$$

Notice that because these necklaces have period d, we must have $s_i = s_j$ for every $i \equiv j \mod d$, and therefore that any necklace in $\mathcal{N}_{n,d}$ is of the form

$$(s_1, s_2, \ldots s_d, s_1, s_2, \ldots s_d, \ldots s_1, s_2, \ldots s_d).$$

Consider the map f that takes any necklace $N \in \mathcal{N}_d$ and sends it to $(s_1, \dots s_d)$. I claim that this necklace is in $\mathcal{N}_{d,d}$. To see this, notice that the length is clearly right, so it is in $\mathcal{N}_{d,k}$ for some divisor of d; but if $k \neq d$ then this smaller necklace would have period strictly less than n/d. In particular, this would force our original necklace to have period $k < d$, as rotating it k steps causes each $s_1, \dots s_d$ string to return to itself, and therefore causes the entire necklace to return to itself.

I also claim that f is a bijection, which will prove our claim. It clearly is a surjection; given any $N = (s_1, \dots s_d) \in \mathcal{N}_{d,d}$ we have

$$f(\overbrace{s_1, \dots s_d, s_1, \dots s_d, \dots, s_1, \dots s_d}^{n/d \text{ repetitions}}) = (s_1, \dots s_d).$$

As well, I claim it is an injection. Take any two

$$(s_1, s_2, \dots s_d, s_1, s_2, \dots s_d, \dots s_1, s_2, \dots s_d),$$
$$(t_1, t_2, \dots t_d, t_1, t_2, \dots t_d, \dots t_1, t_2, \dots t_d),$$

in $\mathcal{N}_{d,n}$ such that $(s_1, \dots s_d)$ is rotation-equivalent to $(t_1, \dots t_d)$.

Then there is some rotation of $(s_1, \dots s_d)$ that makes it equal to $(t_1, \dots t_d)$; but applying that rotation to

$$(s_1, s_2, \dots s_d, s_1, s_2, \dots s_d, \dots s_1, s_2, \dots s_d)$$

will yield

$$(t_1, t_2, \dots t_d, t_1, t_2, \dots t_d, \dots t_1, t_2, \dots t_d).$$

In other words, we've shown that no two distinct inputs can yield the same output, which gives us injectivity and finishes our proof. □

If we apply this to our earlier observations, we get that

$$2^n = \sum_{d|n} df(d, d). \tag{2}$$

Notice that f is no longer a function of two variables now! So, if we shorten it to $f(d)$, and let m denote any element in our divisor poset for n, we actually have

$$2^m = \sum_{d|m} df(d). \tag{3}$$

In other words, we now have two functions on our divisor poset, $f(d)$ and 2^m, such that one is a sum of the other! Therefore, we can apply Möbius inversion to get

$$mf(m) = \sum_{d|m} 2^d \mu(d, m).$$

As noted above, we have $\mu(d, m) = \mu(1, m/d)$, and therefore can rewrite this as

$$mf(m) = \sum_{d|m} 2^d \mu(1, m/d).$$

Finally, we can use the fact that summing over all of the divisors d of m is just the same as summing over all the divisors m/d of m to get that

$$mf(m) = \sum_{d|m} 2^{m/d} \mu(1, d).$$

This is... not what we want. It's close, though! We wanted all of the inequivalent ways to make a necklace on n beads: that is, we want

$$\sum_{m|n} f(m).$$

We have information about this sum now: we have (by switching the order of summation)

$$\begin{aligned}
\sum_{m|n} f(m) &= \sum_{m|n} \frac{1}{m} \sum_{d|m} 2^{m/d} \mu(1, d) \\
&= \sum_{m,d:d|m,m|n} \frac{2^{m/d} \mu(1, d)}{m} \\
&= \sum_{d:d|n} \sum_{m:d|m,m|n} \frac{2^{m/d} \mu(1, d)}{m}
\end{aligned}$$

We can reparametrize the inner sum as actually being taken over all factors l of n/d, instead of over all factors m between d and n, by just dividing through by d. If we do this, we replace each m in our summands with a ld, which gives us

$$\sum_{m|n} f(m) = \sum_{d:d|n} \sum_{l:l|(n/d)} \frac{2^l \mu(1, d)}{ld}$$

We can change the order of summation again to get

$$\begin{aligned}
\sum_{m|n} f(m) &= \sum_{l:l|n} \sum_{d:d|(n/l)} \frac{2^l \mu(1, d)}{ld} \\
&= \sum_{l:l|n} \frac{2^l}{l} \sum_{d:d|(n/l)} \frac{\mu(1, d)}{d}
\end{aligned}$$

On the homework, you will prove (if you do this problem!) that

$$\frac{\varphi(n)}{n} = \sum_{d\in\mathbb{N}:d|n} \frac{\mu(1,d)}{d},$$

and therefore that

$$\frac{l\varphi(n/l)}{n} = \sum_{d\in\mathbb{N}:d|(n/l)} \frac{\mu(1,d)}{d}.$$

Applying this result here to the inner sum gives us

$$\sum_{m|n} f(m) = \sum_{l:l|n} \frac{2^l\varphi(n/l)}{n}.$$

An answer! An odd answer, but an answer nonetheless. It even looks fairly nice for certain values of n; for example, if n is a prime, we have that the only divisors of n are $1, n$ and therefore that the sum above is

$$\frac{1}{n}(2^n\varphi(1) + 2\varphi(n)) = \frac{1}{n}(2^n + 2(n-1)).$$

So, for instance, the number of such necklaces for $n = 3$ is $\frac{1}{3}(8+4) = 4$, which makes sense; we have four distinct necklaces by counting the number of black shells (there are either 0, 1, 2 or 3) such shells, and any two necklaces with the same number of black shells are the same under rotation (check it!)