Winter 2020

The Mobius Function and Mobius Inversion

Carl Lienert

The Möbius Function and Möbius Inversion

Carl Lienert*

January 17, 2021

August Ferdinand Möbius (1790–1868) is perhaps most well known for the one-sided *Möbius strip* and, in geometry and complex analysis, for the *Möbius transformation*. In number theory, Möbius' name can be seen in the important technique of *Möbius inversion*, which utilizes the important *Möbius* function. In this PSP we'll study the problem that led Möbius to consider and analyze the Möbius function. Then, we'll see how other mathematicians, Dedekind, Laguerre, Mertens, and Bell, used the Möbius function to solve a *different* inversion problem.[1] Finally, we'll use Möbius inversion to solve a problem concerning Euler's totient function.

1 Möbius: the Möbius function

All excerpts of Möbius' work in this project are from [Möbius, 1832], *Uber eine besondere Art von Umkehrung der Reihen* (*On a special type of series inversion*). The following excerpt, from the beginning of Möbius' paper, sets up the basic form of Möbius' inversion problem:

The famous problem of series inversion is that, when a function of a variable is given as a consecutive series of powers of the variable, one inversely requires the variable itself, or even any other function of it, expressed as an ongoing series of powers of the original function. One knows that it requires no small analytical ingenuity to discover the rule according to which the coefficients of the second series depend on the coefficients of the first. The following task is much easier to solve.

Suppose a function fx of a variable x is given as a series according to the powers of x:

$$fx = a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \dots \qquad (1)$$

One should represent x as an ongoing series, not according to the powers of the function fx, but rather according to the function f of the powers of x:

$$x = b_1fx + b_2f(x^2) + b_3f(x^3) + b_4f(x^4) + \dots \qquad (2)$$

*Department of Mathematics, Fort Lewis College, Durango, CO, 81301; lienert_c@fortlewis.edu.

[1] All translations of excerpts from the German works of Möbius, Dedekind, and Mertens in this project were done by David Pengelley, New Mexico State University (retired), 2021. The author of this PSP is responsible for the translation of the French excerpt from the work of Laguerre.

Task 1 Where have you seen a function written in the form in (1)?

Task 2 The expression in (2) is the inversion of the expression in (1). Why would this be called an *inversion*?

Möbius continued, and stated the goal of the problem:

The main demand of our problem is: Express the coefficients $b_1, b_2, b_3, \ldots$ of the series (2) as functions of the coefficients $a_1, a_2, a_3, \ldots$ of the series (1); and this occurs through the following very easy calculation.

Task 3 In your own words, what is the objective?

Ok, now it's time to get our hands dirty. Given that

$$f(x) = a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \ldots,$$

we'll find expressions for $b_1, b_2, b_3, b_4, \ldots$ in terms of the given coefficients $a_1, a_2, a_3, a_4, \ldots$.

The symbolic equations have been removed from the following two excerpts. The tasks that follow ask you to fill in the missing sets of equations.

From (1) flows:

Equation Set A

If one substitutes these values of $f\left(x^2\right), f\left(x^3\right), \ldots$ and of fx itself from (1) into the equation (2), one gets:

Equation B

Task 4 Give expressions for $f(x^2)$, $f(x^3)$, $f(x^4)$, $f(x^5)$, and $f(x^6)$. These are Equation Set A.

Task 5 Do what Möbius instructed: "substitute these values of $f\left(x^2\right), f\left(x^3\right), \ldots$ and of fx itself from (1) into the equation (2)" *and* then rearrange the expression you obtain on the right so that it's in the form

$$\_\_\_\_x + \_\_\_\_x^2 + \_\_\_\_x^3 + \_\_\_\_x^4 + \_\_\_\_x^5 + \_\_\_\_x^6 + \ldots . \tag{3}$$

This is Equation B.

Task 6 Give the coefficient of x^{23} in terms of a's and b's.

Task 7 Give the coefficient of x^{24} in terms of a's and b's.

Möbius continued:

The law of progression of the coefficients in this series is clear. Namely, to determine the coefficient of x^m, partition the number m in all possible ways into two positive whole factors. Each of these products then gives a term of the coefficient sought, in that one takes the two factors of the product as indices of an a and b to multiply together.

Because the equation above must hold for every value of x, we have:

Equation Set C

through which every b can be calculated with the aid of the previous b's.

In order therefrom to find the individual b's independently from one another, one sets $a_1 = 1$ for the sake of greater simplicity, and obtains:

Equation Set D

Task 8 Möbius explained how to obtain the coefficient of x^m in this expansion in the first paragraph of this excerpt. Compare his explanation to your answers *and* to your work for Tasks 6 and 7.

Remember, the expression you found, (3), is the right hand side of (2):

$$x = b_1 f(x) + b_2 f(x^2) + b_3 f(x^3) + b_4 f(x^4) + \ldots .$$

Task 9 Next, Möbius stated "Because the equation above must hold for every value of $x\ldots$" So, match (3) with the left hand side of (2) in order to obtain conditions on all the coefficients you found in (3). This will be a list of *equations* with a's and b's on the one side of the equality, and a number on the other. This is Equation Set C.

At this point, Möbius decided to let $a_1 = 1$ for convenience. We'll do the same. There's no harm done; if the function you are interested in doesn't have $a_1 = 1$ use the function $\frac{1}{a_1} f(x)$ instead and adjust accordingly in the end.

From Equation Set C, you can now find values for the b's in terms of the a's.

Task 10 What is b_1?

Task 11 Use the value of b_1 to find the value of b_2 in terms of a's. Continue: find b_3, b_4, b_5, b_6, b_7, and b_8 in terms of a's (no b's). These are Equation Set D.

Next, Möbius made an observation about how to form the b's without the need to extend the process above indefinitely:

These few developments are sufficient to take away how also the values of the succeeding b's are put together from $a_2, a_3, \ldots$. Namely one decomposes the index m of b_m in all possible ways into factors, in which one takes m itself as the largest factor, but omitting 1, and also considers any two decompositions, that differ only in the order of their factors, as different; or as one can express briefly in the language of combinatorial theory: One builds all variations with repetition to the product m. Each of these variations then gives a term in the value of b_m, taking the elements of the variation as the indices of a's, and this term receives the positive or negative sign, according to whether the number of elements is even or odd.

So for example all variations of the product 12 are:

$$12,\ 2\cdot 6,\ 3\cdot 4,\ 4\cdot 3,\ 6\cdot 2,\ 2\cdot 2\cdot 3,\ 2\cdot 3\cdot 2,\ 3\cdot 2\cdot 2,$$

and thus

$$b_{12} = -a_{12} + 2a_2a_6 + 2a_3a_4 - 3a_2a_2a_3.$$

The general correctness of this rule flows from the recurrence formula (Equation Set C) so easily that it would be superfluous for us to tarry for a proof.

Task 12 Use Möbius' observation to give an expression for b_8 in terms of a's. Compare both your answer and your process to those of Task 11.

Task 13 Use Möbius' observation to give an expression for b_{31} in terms of a's.

Task 14 Give an expression for b_{45} in terms of a's.

Möbius presented several generalizations of the basic problem stated in the first excerpt. We'll look at one of these to get the idea.

The same relations between the coefficients a's and b's would incidentally also be obtained if, like (1) and (2), one had in the same way compared the general equalities

$$fx = a_1 Fx + a_2 F(x^2) + a_3 F(x^3) + \ldots$$
$$Fx = b_1 fx + b_2 f(x^2) + b_3 f(x^3) + \ldots$$

with one another. Supposing therefore that the relations (Equation Set C) hold between the a's and the b's, then for these two equalities the second is a consequence of the first, and the first a consequence of the second, where in the first case Fx, and in the latter fx, may be a function of x.

Task 15 What function, $F(x)$, would make this generalized problem the same as the basic problem presented in the first excerpt?

Task 16 Möbius wrote "Supposing therefore that the relations (Equation Set C) hold between the a's and the b's. . . ." Verify that this is, in fact, true. That is, repeat the analysis of Tasks 4, 5, and 9 for this generalized problem.

Perhaps the fact that Equation Set C is the same for the basic problem as it is for the generalized problem made Möbius think something interesting was happening. In fact, Möbius presented two further generalizations in which the same pattern continued to occur.

What Möbius did next is a valuable lesson: work out a simple example. Maybe the example will provide insight, maybe the example will be important in its own right. That is, he didn't try to analyze the most general case he had presented, which would have been very difficult. Instead he returned to the basic example, and in fact made it even easier for himself and for his readers:

In order now to give a very simple example of this new kind of series inversion, we want to set

$$a_1 = a_2 = a_3 = \cdots = 1,$$

so that from (1)

$$fx = x + x^2 + x^3 + \ldots \text{ and therefore } fx = \frac{x}{1-x}$$

But with these values for a's, according to (Equation Set D):

$$b_1 = 1, b_2 = -1, b_3 = -1, b_4 = 0, b_5 = -1, b_6 = 1, b_7 = -1, b_8 = 0, \text{ etc,}$$

and thus from (2):

$$x = \frac{x}{1-x} - \frac{x^2}{1-x^2} - \frac{x^3}{1-x^3} - \frac{x^5}{1-x^5} + \frac{x^6}{1-x^6} - \frac{x^7}{1-x^7} + \frac{x^{10}}{1-x^{10}} - \frac{x^{11}}{1-x^{11}} - \frac{x^{13}}{1-x^{13}} + \dots \tag{4}$$

Task 17 Möbius claimed

$$f(x) = x + x^2 + x^3 + \cdots = \frac{x}{1-x}.$$

What kind of series is $x + x^2 + x^3 + \dots$? Show that, in fact,

$$x + x^2 + x^3 + \cdots = \frac{x}{1-x}.$$

Task 18 Show how the result from Task 17 yields (4).

Task 19 What are the next 3 non-zero terms in (4)?

Task 20 Can you predict b values without going through the entire process? Try to predict values for b_{37}, b_{64}, b_{65}, b_{105}, and b_{128}. Explain how you arrived at your predictions.

Here is the observation Möbius gave for b values:

In the series (4), whose general term is $\frac{x^m}{1-x^m}$ and whose sum is $= x$, the law therefore reigns, that for $m = 1$ and for every m that is a product of an even number of distinct prime numbers, the coefficient of the term is $= 1$, that every term, whose m is itself a prime number, or a product of an odd number of distinct primes, has the coefficient -1, and finally that all terms are dropped, whose exponents have quadratic or higher powers of prime numbers as factors.

Task 21 Compare your predictions in Task 20 to Möbius' observation. Do they agree? If not, explain how your prediction method differs from that of Möbius.

These b values are the values of what is known today as the Möbius function.

Task 22 The Möbius function, b_i, is defined for positive integers, i. Write the function that Möbius described using modern notation.

Möbius dedicated the middle portion of his paper to a careful and thorough examination of the this function. Having found the b values, Möbius had solved the inversion problem, but not what is known today as Möbius Inversion. He then used this inversion technique to produce interesting series results. Some examples are:

$$
\begin{aligned}
e &= 2^{1/2} \cdot 3^{1/3} \cdot 5^{1/5} \cdot 6^{-1/6} \cdot 7^{1/7} \cdot 10^{-1/10} \cdot 11^{1/11} \cdots \\
\frac{4}{\pi} &= 1 + \frac{1}{3} - \frac{1}{5} + \frac{1}{7} + \frac{1}{11} - \frac{1}{13} - \frac{1}{15} - \frac{1}{17} + \dots \\
\frac{4}{\pi} &= \frac{3}{4} \cdot \frac{5}{4} \cdot \frac{7}{8} \cdot \frac{11}{12} \cdot \frac{13}{12} \cdot \frac{17}{16} \cdot \frac{19}{20} \cdot \frac{23}{24} \cdots .
\end{aligned}
$$

Notice, these examples give a method to approximate the transcendental numbers e and π.

Task 23 In the product expansion given above for e what are the next 3 *missing* factors? Why?

Task 24 (Optional) Use technology to experiment with how quickly these series converge to e or to π. How many factors or terms do you need to achieve 5 decimal places of accuracy?

The development of these examples can be found at the end of Möbius' paper and with a little work you can follow it, even if you don't read German. However, this type of inversion is not what is known today at *Möbius inversion* and so we won't follow that detour here.

2 Dedekind: Möbius inversion

Today, Möbius inversion concerns a different kind of sum: a divisor sum of an arithmetic function.

An *arithmetic function* is one whose domain is positive integers only. You might notice that the Möbius function is an arithmetic function. In fact, the values a_i in the first excerpt from Möbius define an arithmetic function: the domain is $i = 1, 2, 3, \dots$. If $f(k)$ is an arithmetic function, then a *divisor sum* is:

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

For example, if $n = 6$,

$$F(6) = f(1) + f(2) + f(3) + f(6).$$

Julius Wilhelm Richard Dedekind (1831–1916) studied mathematics under Carl Friedrich Gauss (1777–1855) and later worked closely with Peter Gustav Lejeune Dirichlet (1805–1859) who took Gauss' chair upon his death. Dedekind was the first to state and prove Möbius inversion in his paper *Abrifs einer Theorie der höhern Congruenzen in Bezug auf einen reelen Primzahl-Modulus* (Outline of a theory of higher congruences in connection with a real prime-modulus) [Dedekind, 1857]:

The shared source of the theorem in section 18 and the analogous theorem just now used is the following. Let m be any whole number; further $a, b, c, \ldots, k$ all the distinct prime numbers that divide into m; one forms two separate complexes D, D' of divisors of the number m according to the following principle. In the complex D one initially includes all divisors of the number m; in the complex D' all divisors of $\frac{m}{a}$, all divisors of $\frac{m}{b}$ and so on; then again in the complex D all divisors of $\frac{m}{ab}$, of $\frac{m}{ac}$, of $\frac{m}{bc}$ and so on; then again in the complex D' all divisors of $\frac{m}{abc}$ and so on, until finally one has included also all divisors of $\frac{m}{abc\cdots k}$ either in the complex D or in the complex D', depending on whether the number of prime numbers $a, b, c, \ldots, k$ is even or odd. Then it is easy to show that each divisor of the number m occurs just as often in one complex as in the other, with the exception of the divisor m itself, which occurs solely and only once in the complex D. It requires only one look to derive from this the inversion of the equalities

$$\sum f(\delta) = F(m) \text{ or } \prod f(\delta) = F(m)$$

in which the sum or product sign $\sum$ or $\prod$ refers to all divisors δ of an arbitrary number m; these solutions are contained in the formulas

$$f(m) = F(m) - \sum F\left(\frac{m}{a}\right) + \sum F\left(\frac{m}{ab}\right) - \text{etc.} \ldots \tag{5}$$

We are only interested in the equations involving summation, and not those with products.

Task 25 What does this have to do with Möbius?! Rewrite the expression on right hand side of (5) using the Möbius function.

Task 26 What did Dedekind mean by "complex?"

Task 27 Determine D and D' for $m = 60$.

Task 28 Explain Dedekind's strategy. What does Dedekind's discussion about the complexes D and D' have to do with (5)?

Task 29 Write the expression on the right side of (5) for $m = 6$, without summation notation. Simplify and confirm that, in fact, you obtain $f(6)$.

Task 30 Write the expression on the right side of (5) for $m = 2^3 \cdot 3^2$ without summation notation. Simplify and confirm that, in fact, you obtain $f(2^3 \cdot 3^2)$. Try to pay careful attention to how many times a given term $f(k)$, for any divisor k of m, is added or subtracted in the expression (5).

Dedekind's setup of the statement with "$a, b, c, \ldots k$" would, today, be replaced by writing m in its prime factor form:

$$m = p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} \cdots p_l^{\alpha_l}$$

where p_i is prime.

Task 31 Translate the theorem that Dedekind stated using modern "if ... then ..." presentation and modern notation. In particular use $\sum_{d|m}$ notation and write m in its prime factor form.

Dedekind claimed that "it only requires one look" to see that

$$F(m) - \sum F\left(\frac{m}{a}\right) + \sum F\left(\frac{m}{ab}\right) - \text{etc.} \tag{6}$$

simplifies to $f(m)$.

Tasks 29 and 30 provide reason to believe that Dedekind's claim is, in fact, true, but certainly do not constitute a proof. Let's try a slightly more complicated example with the hope that it will give us an idea how to prove (5) is true for *any* value of m.

Task 32 Let $m = 2^4 \cdot 3^{12} \cdot 5^3 \cdot 7^{11}$. Let $k = 2^2 \cdot 3^{12} \cdot 5^2 \cdot 7^{10}$. We'll count how many times $f(k)$ appears in each of the sums in (6).

(a) Remember $F(m) = \sum_{d|m} f(d)$. How many times does $f(k)$ appear in this sum?

(b) How many times does $f(k)$ appear in

$$\sum_a F\left(\frac{m}{a}\right) = \sum_a \sum_{d|\frac{m}{a}} f(d)?$$

The sum on the left of the equality is over distinct primes, a, in the factorization of m.

(The way to read double sums like this is to fix a prime a in the outer sum, then work through the inner sum for that value of a. Then change to another value of a, work through the inner sum, etc., until you've exhausted all the primes that divide m.)

(c) How many times does $f(k)$ appear in

$$\sum_{a,b} F\left(\frac{m}{ab}\right) = \sum_{a,b} \sum_{d|\frac{m}{ab}} f(d)?$$

The sum on the left of the equality is over pairs of distinct primes in the factorization of m.

(d) How many times does $f(k)$ appear in

$$\sum_{a,b,c} F\left(\frac{m}{abc}\right) = \sum_{a,b,c} \sum_{d|\frac{m}{abc}} f(d)?$$

(e) How many times does $f(k)$ appear in

$$\sum_{a,b,c,e} F\left(\frac{m}{abce}\right) = \sum_{a,b,c,e} \sum_{d|\frac{m}{abce}} f(d)?$$

(f) This is where (6) ends for this example. Why?

(g) Finally, add and subtract your answers to parts 1, 2, 3, 4, and 5 appropriately. Do you obtain the value you expected?

(h) There is only one divisor of m for which something similar *won't* happen. What divisor is this?

Task 33 Repeat Task 32 with $k = 2^4 \cdot 3^{12}$.

We're ready to make the idea above general. Let

$$m = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdots p_l^{\alpha_l}.$$

We may as well assume that $\alpha_i > 0$ for each i. Any divisor of m is of the form

$$k = p_1^{\beta_1} \cdot p_2^{\beta_2} \cdots p_l^{\beta_l}$$

where $0 \leq \beta_i \leq \alpha_i$ for each i.

Task 34 We've allowed the β's to be zero, but not the α's. Why?

Notice, the key to the counting in Task 32 was to look at the exponents in the prime factorizations of m and k.

One of two things happens for any given i: either $\beta_i < \alpha_i$ or $\beta_i = \alpha_i$.

Task 35 If $\beta_i < \alpha_i$ for M different values of i, how many terms in (6) will count $f(k)$? That is, in how many of the sums in (6) does $f(k)$ appear? It doesn't matter whether the difference $\alpha_i - \beta_i$ is 1 or 21. Why?

Task 36 In the case $\beta_i = \alpha_i$ for all values of i, which of the terms in (6) will count $f(k)$? Why?

Now, let M be the size of the set $\{\alpha_i - \beta_i \neq 0\}$. That is, M is the number exponents in the prime factorization of k that are smaller than those of m. Keep in mind, we are focusing on one particular divisor, k, at a time.

In order to count effectively we'll use a function from combinatorics. The function $\binom{M}{t}$ [2] is the number of ways to choose t items from a collection of M items when the order in which they are chosen doesn't matter.

Task 37 We're ready to repeat the idea in Task 32 in general. For any particular divisor k, of m:

(a) Remember $F(m) = \sum_{d|m} f(d)$. How many times does $f(k)$ appear in this sum?

(b) How many times does $f(k)$ appear in

$$\sum_{p_i} F\left(\frac{m}{p_i}\right) = \sum_{p_i} \sum_{d|\frac{m}{p_i}} f(d)?$$

[2] read "M choose t"

(c) How many times does $f(k)$ appear in

$$\sum_{p_i,p_j} F\left(\frac{m}{p_i p_j}\right) = \sum_{p_i,p_j} \sum_{d|\frac{m}{p_i p_j}} f(d)?$$

$(p_i \neq p_j)$

(d) How many times does $f(k)$ appear in

$$\sum_{p_i,p_j,p_r} F\left(\frac{m}{p_i p_j p_r}\right) = \sum_{p_i,p_j,p_r} \sum_{d|\frac{m}{p_i p_j p_r}} f(d)?$$

(e) How many times does $f(k)$ appear in

$$\sum_{p_i,p_j,p_r,p_s} F\left(\frac{m}{p_i p_j p_r p_s}\right) = \sum_{p_i,p_j,p_r,p_s} \sum_{d|\frac{m}{p_i p_j p_r p_s}} f(d)?$$

(f) What is the form of the last, non-empty sum? That is, what is the last term in (6) that counts an occurrence of $f(k)$? How many times does $f(k)$ appear in this last sum?

(g) Add and subtract your answers appropriately to the above parts.

To finish we'll need the Binomial Theorem:

$$(x-y)^M = 1 - \binom{M}{1} x^{M-1} y + \binom{M}{2} x^{M-2} y^2 - \binom{M}{3} x^{M-3} y^3 + \cdots + (-1)^M \binom{M}{M} y^M.$$

Task 38 Compare your answer in part ((g)) above to the Binomial Theorem. Pick values for x and for y that make the two expressions the same, and hence compute the sum.

Task 39 If you haven't already taken into account the case when $k = m$, explain what happens in this case.

This ends a proof of Dedekind's presentation of Möbius inversion.

3 Laguerre and Mertens: evolution of Möbius inversion

Next, Edmond Laguerre (1834–1886) and Franz Carl Joseph Mertens (1840–1927) contributed to the story of Möbius inversion. Their papers appeared at similar times; we'll look at Laguerre's work first.

Laguerre was the first to present the theorem of Möbius inversion in the format used today. This excerpt is from his paper *Sur quelques théorèmes d'arithmétique* (*On several theorems of arithmetic*) [Laguerre, 1872/73].

Let $\lambda(n)$ designate a number equal to 0 if n is divisible by a square, and in the other case, equal to ± 1 according to whether the number of factors of n is even or odd. Suppose two functions $f(m)$ and $\varphi(m)$ are connected by the following relation

$$f(m) = \sum \varphi(d)$$

where in the second part the summation extends over all the divisors of the integer m. Reciprocally, one has

$$\varphi(m) = \sum \lambda\left(\frac{m}{d}\right) f(d), \tag{7}$$

the summation also extending over all the divisors of m.

Task 40 Which function in Laguerre's statement is the Möbius function?

Task 41 The first sentence in this excerpt should be more precise. How?

Task 42 Explain how the right hand sides of equations (5) and (7) are the same.

Laguerre used the letter φ because he was, in particular, interested in deriving a formula for Euler's totient function $\varphi(n)$. However, the statement holds for any arithmetic function. [3]

Mertens introduced the modern notation for the Möbius function, namely the choice of the letter μ, and provided a more succinct definition than that of Möbius. You might notice he also took advantage of prime factor notation. The following excerpt is from *Ueber einige asymptotische Gesetze der Zahlentheorie* (On several asymptotic laws in number theory) [Mertens, 1874].

Denote by μn a number depending on n in such a way that $\mu n = 0$ if n admits a quadratic divisor (other than 1), but otherwise possesses the value $+1$ or -1, according to whether n is composed of an even (the case of 1 belongs here) or odd number of different prime factors. If m is decomposed into its prime factors $= a^\alpha b^\beta \ldots$, then φm, as is generally known, is given via the formula

$$\begin{aligned}\varphi m &= m\left(1 - \frac{1}{a}\right)\left(1 - \frac{1}{b}\right)\cdots \\ &= \sum \mu T \, \frac{m}{T},\end{aligned}$$

where the summation extends over all divisors of m.

[3] We'll return to $\varphi(n)$ a little later.

Task 43 There is an interesting result at the end of the excerpt. Don't worry about the first equality; we'll come back to that later. The second equality is known as an Euler product. Prove this equality by expanding the product

$$m\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\cdots.$$

In particular, make sure to explain why the Möbius function appears.

Mertens defined a function, which today bears his name, in the following excerpt from the same paper. Mertens defined the function for complex numbers, but we will only be interested in the case where n is a real integer.

While retaining all the notations of the two previous sections, if one forms the sum $\sum \mu d$ extending over all primary divisors d of a given whole complex number n, then this sum is always $= 0$, except when $n = i^\rho$. Namely, when n contains at least one primary complex prime number q as factor, and setting $n = q^k n'$, where n' is no longer divisible by q, one obtains all primary divisors of n, when each primary divisor δ of q^k is multipled by each primary divisor δ' of n'. And since moreover $\mu\delta\delta' = \mu\delta\mu\delta'$, one has

$$\sum \mu d = \sum \mu\delta \sum \mu\delta'.$$

But

$$\sum \mu\delta = 1 + \mu q = 1 - 1 = 0$$

and consequently

$$\sum \mu d = 0.$$

If however $n = i^\rho$, then n contains only the one primary divisor 1 and one has

$$\sum \mu d = \mu 1 = 1.$$

For our purposes, the Mertens function is

$$M(n) = \sum_{d|n} \mu(d).$$

Task 44 Experiment first: compute $M(1)$, $M(2)$, $M(3)$, etc.

Task 45 Mertens claimed that $M(n)$ takes on only two values. Write $M(n)$ in piecewise notation.

Task 46 We'll prove your answer to the previous task. The proof should remind you of Task 37.

(a) Write n in its prime factor form: $n = p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} \cdots p_k^{\alpha_k}$. (nothing to do here)

(b) How many divisors of n have the form $n = q$ where q is prime? What is $\mu(q)$ for each of these divisors?

(c) How many divisors of n are of the form $n = q \cdot w$ where q and w are both prime? What is $\mu(q \cdot w)$ for each of these divisors?

(d) Write $\sum_{d|n} \mu(d)$ as a signed sum of your answers to parts 2 and 3 (and other similar questions) according to the form of the divisors d. Hint: Look back at Dedekind. How did he break up a similar sum?

(e) Use the Binomial Theorem to simplify and evaluate the expression from step 4.

We'll use the Mertens function to present a slight modification to Dedekind's proof of Möbius inversion. For convenience, here is a statement of Möbius inversion:

If $g(d)$ is an arithmetic function, and $G(n) = \sum_{d|n} g(d)$, then $g(n) = \sum_{d|n} \mu(d) G\left(\frac{n}{d}\right)$.

We need to evaluate the sum

$$\sum_{d|n} \mu(d) G\left(\frac{n}{d}\right),$$

and show it equals $g(n)$. The steps are shown below; the tasks that follow ask you to explain each step.

$$\begin{aligned}
\sum_{d|n} \mu(d) G\left(\frac{n}{d}\right) &= \sum_{d|n} \mu(d) \sum_{k|\frac{n}{d}} g(k) && \text{the second sum is } \textit{inside} \text{ the first sum} && (8)\\
&= \sum_{d|n} \sum_{k|\frac{n}{d}} \mu(d) g(k) && && (9)\\
&= \sum_{k|n} \sum_{d|\frac{n}{k}} \mu(d) g(k) && && (10)\\
&= \sum_{k|n} g(k) \sum_{d|\frac{n}{k}} \mu(d) && \textit{notice}, \text{the second (inside) sum is Mertens' function!} && (11)\\
&= g(n) && && (12)
\end{aligned}$$

Task 47 Explain the equality in (8).

Task 48 Explain how to get from (8) to (9). It may help to write out a small example, say for $n = 6$.

Task 49 The step from (9) to (10) is a common and very useful technique in number theory. Explain what happened. Again, it may help to continue your example from above for a small value of n.

Task 50 Explain how to get from (10) to (11).

Task 51 Finally, why is (11) simply equal to $g(n)$. Of course, make use of the hint about Mertens' function.

This proof is not fundamentally different than Dedekind's. The counting in Task 37 is more directly connected to the divisor sum. In the proof using Mertens' function the counting is "hidden" in the evaluation of $M(n)$. One presentation is not necessarily better than the other. Dedekind's proof more explicitly illustrates the issues in play. The proof presented here, using Mertens' function, is easier to follow (perhaps), but in the end one might not *see* what's going on.

4 Bell: Möbius inversion from an abstract algebra perspective

Eric Temple Bell (188–1960) introduced the idea of treating arithmetic functions from a ring structure perspective. We won't worry too much about what a *ring* is; we'll focus instead on the aspects that help us provide another proof of Möbius inversion.

We've seen divisor sums above. A special kind of divisor sum appears frequently in number theory. Today, this type of sum is called a *Dirichlet convolution*.[4] If $f(n)$ and $g(n)$ are arithmetic functions, then the *Dirichlet convolution* is:

$$(f \star g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right).$$

Notice, Möbius inversion is a Dirichlet convolution.

Arithmetic functions together with the binary operation $\star$ form a structure (ring) that is similar to the structure of the integers together with the binary operation of multiplication. (Don't worry about addition, we won't need that.) In particular, $\star$ is commutative and associative.

Task 52 Show that Dirichlet convolution is commutative.

Task 53 (optional) Show that Dirichlet convolution is associative.

In the following excerpt, Bell showed an illustration of Dirichlet convolution, introduced two arithmetic functions we haven't seen yet, and illustrated a typical use of these two functions. The excerpt is from Bell's book *An Arithmetic Theory of Certain Numerical Functions* [Bell, 1915].

3.00. The concept of ideal multiplication may be briefly illustrated by the first (historically) example in which it is implicit. Denoting by $\varphi(n)$ the totient of n (viz., the number of integers not greater than, and prime to, n), and by $u_0(n)$ a numerical function of n which has the

[4] We will not read any excerpts from the work of Dirichlet, but it should be noted that he played no small role in this story. Dirichlet and Dedekind collaborated closely. Dedekind published a book of Dirichlet's notes after Dirichlet's death titled *Vorlesungen über Zahlentheorie* (*Lectures on Number Theory*). This book, according the the translator into English, is "one of the most important mathematics books of the 19th century." The papers we have considered by both Mertens and Laguerre were motivated by work of Dirichlet. Dirichlet convolutions arise naturally in Dirichlet's remarkable proof that there are infinitely many primes in any arithmetic progression of the form $a + bm$ with a and b relatively prime.

value of unity for all values of the argument n, and by $u_1(n)$, a numerical function which has the value n for all values of the argument n, there is the well-known theorem $\sum_{(n)} \varphi(d) = n$; or what is equivalent,

$$\sum_{(n)} \varphi(d) u_0\left(\frac{n}{d}\right) = u_1(n);$$

the notation $\sum_{(n)}$ etc. meaning (as customary) that the summation is extended over all divisors d, including 1 and n, of n.

For our purposes the word "unity" simply means the number 1. Bell defined a *numerical function* to be an arithmetic function, $g(n)$, with $g(1) = 1$. The distinction between *numerical* and *arithmetic* will not be important for us. Bell called Dirichlet convolution "ideal multiplication."

Task 54 Evaluate the divisor sums:

$$\sum_{d|12} u_0(d)$$

$$\sum_{d|12} u_1(d)$$

(We won't need this later, it's just to get a little practice with these two functions.)

Task 55 Notice, Bell's example showed how to write $\sum_{(n)} \varphi(d) = n$ as a Dirichlet convolution. (We'll prove $\sum_{(n)} \varphi(d) = n$ later; that's not our focus here.) Our statement of Möbius inversion has the set-up: $G(n) = \sum_{d|n} g(d)$. Follow Bell's example and write $G(n)$ as a convolution of two arithmetic functions. Write this convolution both as a divisor sum and using the $\star$ notation.

Bell defined another class of functions in the following excerpt. For our purposes, it is just one function:

3.10. If $\psi(n)$ is a numerical function which vanishes for all values of n greater than unity, then $\psi(n)$ is a *unit function*, and ψ is a *unit functional form*; briefly, ψ is a *unit*.

Task 56 Write $\psi(n)$ in piecewise notation.

Task 57 If $g(n)$ is any arithmetic function, what is $(g \star \psi)(n)$?

Task 58 Compute $\mu \star u_0$.

Task 59 In Task 55, you wrote G as a convolution. Convolve both sides of this equation with μ and simplify the right side. Remember, Dirichlet convolution is both commutative and associative.

Task 60 Write the equation you found in Task 59 as a divisor sum. You should recognize this divisor sum. What is it?

This proof is even slicker than the one using Mertens' function and it illustrates the power of modern abstract algebra. As before, there seems to be a trade off between being slick and *really* seeing what's going on.

5 One example

Euler's totient function, $\varphi(n)$, has appeared several times in this project. Reread the first two excerpts from Section 3 to see where $\varphi(n)$ appeared in the works of Laguerre and Mertens. The value of $\varphi(n)$ is the number of integers between 1 and n that are relatively prime to n. For example, $\varphi(6) = 2$ since 1 and 5 are relatively prime to 6 (and 2, 3, 4, and 6 are not relatively prime to 6). We'll use Möbius inversion to find a formula for $\varphi(n)$.

Task 61 Compute $\varphi(7)$, $\varphi(12)$, and $\varphi(36)$.

If $f(n)$ is an arithmetic function, it is sometimes easier to find a formula for $\sum_{d|n} f(d)$ than it is to find a formula for $f(n)$. With this in mind, we'll first find a formula for $\sum_{d|n} \varphi(d)$.

Task 62 Let $n = 12$. Compute $\sum_{d|12} \varphi(d)$.

(Of course, the answer can be found earlier in the project, but do the computation as if you didn't know the answer.)

Task 63 Let $n = 15$. Compute $\sum_{d|15} \varphi(d)$.

Clearly, the conjecture is that $\sum_{d|n} \varphi(d) = n$. Let's prove it:

Task 64 It will help to work a small example first to illustrate the idea. Let $n = 12$.

(a) Write the fractions: $\dfrac{1}{12}, \dfrac{2}{12}, \ldots, \dfrac{12}{12}$.

(b) What is the length of this list of fractions?

(c) Reduce each of the fractions to lowest terms. That is, so the numerator and denominator are relatively prime.

(d) How many fractions have a denominator of 6? What is $\varphi(6)$?

(e) How many fractions have a denominator of 12? What is $\varphi(12)$?

(f) Argue that $\varphi(1) + \varphi(2) + \varphi(3) + \varphi(4) + \varphi(6) + \varphi(12) = 12$ *without* computing the values of $\varphi(d)$ for each divisor d of 12.

Let's reproduce the argument in general:[5]

Task 65

(a) Write the fractions: $\dfrac{1}{n}, \dfrac{2}{n}, \ldots, \dfrac{n}{n}$.

(b) What is the length of the list of fractions?

(c) Reduce the fractions to lowest common terms. Of course, you can't actually do this, so you'll have to imagine the list of fractions in lowest common terms.

(d) For any divisor d of n, what can you say about the numerator of a fraction in the lowest terms list with d in the denominator?

(e) If d is a divisor of n, and b is an integer between 1 and d *and* relatively prime to d, explain why the fraction $\frac{b}{d}$ must be in the list.

(f) How many fractions in the list have a divisor of d?

(g) Finally, explain why $\sum_{d|n} \varphi(d) = n$.

We're ready to find a formula for $\varphi(n)$:

Task 66

(a) Apply Möbius inversion to $\sum_{d|n} \varphi(d) = n$. Your answer probably looks like a Dirichlet convolution.

(b) If you haven't already, expand your answer from the previous step so that it looks something like (6). What is the function F in this case?

(c) Factor! Hint: see Task 43.

Task 67 Use the formula for $\varphi(n)$ to compute $\varphi(7)$, $\varphi(12)$, and $\varphi(36)$. Compare your answers *and* the work to those from Task 61.

If you read Mertens' first excerpt carefully, you'll notice he made the argument in the opposite direction. That is, he started with a formula for $\varphi(n)$ and *then* showed that $\sum_{d|n} \varphi(d) = n$. If you're interested, find a direct derivation of the formula for $\varphi(n)$ ([Rosen, 2005], for example) and compare the approaches.

[5] This particular presentation of the argument is from [Ireland and Rosen, 1990].

References

T. M. Apostol. *Introduction to Analytic Number Theory.* Springer, 1976.

E. T. Bell. *An Arithmetical Theory of Certain Numerical Functions.* University of Washington Publications, 1915.

R. Dedekind. Abriss einer Theorie der höhern Congruenzen in Bezug auf einen reellen Primzahl-Modulus. *Journal für die Reine und Angewandte Mathematik. (Crelle's Journal)*, 54:1 – 26, 1857.

L. E. Dickson. *History of the Theory of Numbers, Divisibility and Primality*, volume 1. Carnegie Institution of Washington, 1919.

K. Ireland and M. Rosen. *A Classical Introduction to Modern Number Theory.* Springer, 2nd edition, 1990.

E. Laguerre. Sur quelques théorèmes d'arithmétique. *Bulletin de la Société Mathématique de France*, 1:77 – 81, 1872/73.

F. Mertens. Ueber einige asymptotische Gesetze der Zahlentheorie. *Journal für die Reine und Angewandte Mathematik. (Crelle's Journal)*, 77:289 – 338, 1874.

A. F. Möbius. Über eine besondere Art von Umkehrung der Reihen. *Journal für die Reine und Angewandte Mathematik (Crelles Journal)*, 9:105 – 123, 1832.

K. H. Rosen. *Elementary Number Theory and its Applications.* Pearson, 5th edition, 2005.

Notes to Instructors

PSP Content: Topics and Goals

This PSP introduces the Möbius function and Möbius inversion. The Möbius function is developed from a motivating question. Three proofs of Möbius inversion are given: a direct counting argument from Dedekind, an indirect counting argument using the Mertens' functions, and an abstract algebra proof following the ideas of Bell. The goal of showing three proofs is to both illustrate the development of the topic and its treatment, and also to compare the relative merits of different types of proofs.

Along the way several typical number theory topics are either reinforced or introduced: divisors, divisor sums, Dirichlet convolution, arithmetic functions, the Fundamental Theorem of Arithmetic, changing the order of summation in order to evaluate a divisor sum, etc.

The primary target course is an Introduction to Number Theory course. The PSP could possibly be used in a Discrete Math course, an Introduction to Proofs course, or a Combinatorics course.

Student Prerequisites

With the exception of Tasks 37 and 46 this PSP should be accessible to any student in an Introduction to Number Theory. Basic familiarity and practice with divisibility and prime factorization are all that is needed.

There are a couple more challenging Tasks, but steps and hints are provided for these.

Tasks 37 and 46 require the combinatorial function $C(n, k)$ (but not evaluation of the function) and the Binomial Theorem. These combinatorial tools are presented in the PSP to the extent they are needed.

For the most part, the proofs and demonstrations are computational in nature.

PSP Design, and Task Commentary

This PSP consists of five sections.

- **Möbius**

  The Möbius function is developed from a motivating question. It's perhaps not clear why the motivating question is an interesting one. However, textbooks typically present the function out of the blue, so the fact that there is a motivating question is interesting. This section is largely one of careful algebra and bookkeeping, and pattern recognition.

  The sum of a geometric series appears; students may need a reminder about what a geometric series is and how to sum it.

  - Task 1: The intent here is simply that students recognize this as a Taylor series.
  - Task 9: I solved this by equating coefficients. A more rigorous treatment might include a discussion about bases, but that might distract from the main story line.
  - Task 16: Again, I suggest treating this naively and equating coefficients.

- **Dedekind**

  This section introduces divisor sums. If appropriate for your class and students, you might pick some other arithmetic functions not involved in this PSP and ask the students to compute several divisor sums, just for practice, so they understand how divisor sums work. Task 25 asks students to write Dedekind's expression using μ. However, I think it will be easier to use

Dedekind's expression (without μ) for the subsequent tasks. Task 37 may be challenging. A counting argument is presented here and the Binomial Theorem is required.

The double sums that appear in this section may cause some confusion at first.

The PSP discusses the minimum amount of combinatorics required. The function $\binom{M}{t}$ can be treated literally and simply as the number of ways to choose t items from M items. No computation of this value is needed. The Binomial Theorem is presented as a result to be used without proof. The intent is to follow the story line without distraction, but of course, each instructor should modify the project as appropriate for their class. Use of relevant material from the course's textbook might serve this purpose.

- **Laguerre and Mertens**

  It's possible that students will have already written Möbius inversion in the format given by Laguerre. The part from Laguerre is simply to make explicit the connection to Möbius inversion as it appears in modern textbooks.

  Mertens did not introduce the function, $M(n) = \sum_{d|n} \mu(n)$, in order to prove Möbius inversion. I'm not sure where this application of Mertens' function first appeared, perhaps in a textbook. Students should notice that Tasks 37 and 46 are similar. Task 49 will be the most difficult of the section.

- **Bell**

  Dirichlet convolution is introduced. It might help students to compute some convolutions of other, unrelated, arithmetic functions. No experience with abstract algebra is needed; the comments about associativity and commutativity are only meant to help students with Task 59.

- **One example**

  This section is included since $\varphi(n)$ appears several times in the primary source excerpts. It doesn't seem fair to leave that unresolved.

Suggestions for Classroom Implementation

My experience with PSPs is that students are unpredictable. My primary advice is to find a balance between allowing students to dictate the pace and getting stuck in one spot.

My general advice is to assign primary sources as advance reading and the following couple tasks which tend to either serve as reading comprehension tasks or to be computational. If class ends in the middle of a sequence of tasks, pick one or two to assign as homework.

Start each day with a short discussion to clear up any confusion and preview any techniques that will be needed that day (eg. prime factorization, Dirichlet convolution, etc.). Then, as much as possible, have students complete the Tasks in groups.

Sample Implementation Schedule (based on a 50-minute class period)

- Day 0: Assign first two Möbius excerpts as reading and Tasks 1–3.
- Day 1:
  - In class, in groups: Tasks 4–14.
  - For homework: read Möbius excerpt that follows Task 14. Complete Tasks 15 and 16. Also, read excerpt that follows Task 16.
- Day 2:
  - In class, in groups: remainder of Section 1: Tasks 17–23.
  - For homework: read first part of section 2, and complete Tasks 25 and 28.
- Day 3:
  - Whole class: Discuss answer to Tasks 25 and 28.
  - Discuss prime factorization, divisor sums, and double sums as appropriate.
  - In groups: start Tasks 29–39.
  - Homework: assign one Task or part of a Task depending on progress made in-class.
- Day 4:
  - Whole class: Discuss combinations and the Binomial theorem as appropriate.
  - In groups: finish Tasks 29–39.
  - Whole class: address any common questions, misunderstandings, etc.
  - Homework: read excerpts from both Laguerre and Mertens. Complete Tasks 40 and 41.
- Day 5:
  - In groups: Tasks 42–46.
  - Homework: have students write out (9) for, say $n = 12$, distribute, regroup, factor out $g(k)$, and write as a new double sum.
- Day 6:
  - Discuss changing order of double sum.
  - In groups: Tasks 47–51.
  - Homework: Read section 4 up to Task 54 and complete Task 54. Warn students there will be some unfamiliar language, but not to worry about it too much.
- Day 7:
  - Discussion as necessary to clarify Bell's excerpt.
  - In groups: remainder of section 4: Tasks 55–60.
  - Homework: Section 5 up to Task 63.
- Day 8:
  - In groups: Complete section 5: Tasks 64–66.

The actual number of class periods spent on each section naturally depends on the instructor's goals and on how the PSP is actually implemented with students. Estimates on the high end of the range assume most PSP work is completed by students working in small groups during class time. Some possible time saving options are suggested below.

Possible Modifications of the PSP

One could definitely do only sections 1 and 2. Sections 3 and 4 present alternate proofs for Möbius inversion. If you prefer one of the proofs from sections 3 or 4 you could do section 1, and one of 2, 3, or 4 with some slight modifications. Section 3 requires the combinatorics discussion which is in section 2. Section 4 is independent with the exception of the referenced hint in Task 66.

Section 5 could be omitted, or perhaps used as a pathway back to a textbook.

I suggest having students complete as many Tasks as possible working in groups. However, if you either want to speed things up a little, or believe this does not provide enough closure/authority for students you could present the main proofs of sections 2 and 3 as lecture. The relevant Tasks are:

- Section 2: Tasks 37–39.
- Section 3: Tasks 46–51.

LaTeX code of this entire PSP is available from the author by request to facilitate preparation of advanced preparation / reading guides or 'in-class worksheets' based on tasks included in the project. The PSP itself can also be modified by instructors as desired to better suit their goals for the course.

Connections to other Primary Source Projects

The following additional projects based on primary sources (written by the indicated authors) are also freely available for use in teaching standard topics in an introductory number theory course. "Mini-PSPs," designed to be completed in 1–2 class periods, are designated with an asterisk (*). The implementation time required for the other "Full PSPs" ranges from 3–9 classroom periods; the Notes to Instructors section of each project provides a detailed sample schedule and suggestions for possible modifications. Classroom-ready versions of these projects can be downloaded from `https://digitalcommons.ursinus.edu/triumphs_number/`.

- *Gaussian Integers and Dedekind's Creation of an Ideal: A Number Theory Project* (Janet Heine Barnett)
- *Generating Pythagorean Triples: A Gnomonic Exploration** (Janet Heine Barnett)
- *Greatest Common Divisor: Algorithm and Proof* (Mary K. Flagg)
- *Primes, Divisibility, and Factoring* (Dominic Klyve)
- *The Origin of the Prime Number Theorem** (Dominic Klyve)
- *The Pell Equation in India* (Toke Knudsen and Keith Jones)

Recommendations for Further Reading

Dickson's book [Dickson, 1919, p. 441-451] has a chapter on the Möbius function and Möbius inversion which both summarizes the development in this PSP as well as indicates other developments of interest.

Acknowledgments

The development of this student project has been partially supported by the TRansforming Instruction in Undergraduate Mathematics via Primary Historical Sources (TRIUMPHS) Program with funding from the National Science Foundation's Improving Undergraduate STEM Education Program under Grant Nos. 1523494, 1523561, 1523747, 1523753, 1523898, 1524065, and 1524098. Any opinions, findings, and conclusions or recommendations expressed in this project are those of the author and do not necessarily represent the views of the National Science Foundation.

For more information about TRIUMPHS, visit `https://blogs.ursinus.edu/triumphs/`.