The Möbius Function and Möbius Inversion

Carl Lienert*

March 11, 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 *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 $f(x)$ of a variable x is given as a series according to the powers of x:

$$f(x) = a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \dots \quad (1)$$

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

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

*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 $f(x^2), f(x^3), \ldots$ and of $f(x)$ itself from (1) into the equation (2)" *and* then rearrange the expression you obtain on the right so that it's in the form

$$\_\_\_\_x + \_\_\_\_x^2 + \_\_\_\_x^3 + \_\_\_\_x^4 + \_\_\_\_x^5 + \_\_\_\_x^6 + \ldots. \qquad (3)$$

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..." 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

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

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

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

so that from (1)

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

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 $x + x^2 + x^3 + \ldots$? 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:

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

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?

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) [**?**]:

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

Task 25 What did Dedekind mean by "complex?"

Task 26 Determine D and D' for $m = 60$.

Task 27 Explain Dedekind's strategy. What does Dedekind's discussion about the complexes D and D' have to do with (5)?

Task 28 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 29 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 30 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 28 and 29 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 31 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 32 Repeat Task 31 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 33 We've allowed the β's to be zero, but not the α's. Why?

Notice, the key to the counting in Task 31 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 34 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 35 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 36 We're ready to repeat the idea in Task 31 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

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

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

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

[2] read "M choose t"

$(p_i \neq p_j)$

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

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

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

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

(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 37 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 38 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*) [**?**].

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 39 Which function in Laguerre's statement is the Möbius function?

Task 40 The first sentence in this excerpt should be more precise. How?

Task 41 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) [**?**].

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 42 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.