6.S096: Number Theory January 18, 2024

Lecture 3— More Elementary Number Theory

A. Anas Chentouf, M. Wacyl Meddour, M.A.O. *Scribe:*

1 P-adic valuation

We will now discuss another fundamental way to look at integers through the lens of primes. As we have discussed in the Fundamental Theorem of Arithmetic, every integer n can be uniquely written as a product of primes of this form :

$$n = q_1^{\alpha_1} q_2^{\alpha_2} \dots q_k^{\alpha_k}$$

where q_i is a prime number and α_i is a positive integer, $1 \leq i \leq k$. Since $x^0 = 1$ for every positive integer x, we can "supplement" this expression by adding all primes in order, with exponents ≥ 0. Indeed, every positive integer n can be written as :

$$n = 2^{k_1} \cdot 3^{k_2} \cdot 5^{k_3} \dots$$

where each $k_i \geq 0$. The unique expression from the FTA comes from the k_i's strictly greater than 0, and guarantees to us the uniqueness of the expression above.

As such, we can identify each integer n with a sequence of non-negative integers $(k_i)_{i \in \mathbb{N}}$ such that $n = \prod_{i=1}^{+\infty} p_i^{k_i}$ where p_i is the i'th prime. Since n is finite, the sequence of k_i's is eventually constant and equal to 0.

For example, one can write $12 = 2^2 \cdot 3^1 \cdot 5^0 \cdot 7^0 \cdot 11^0 \dots$, and the corresponding sequence of exponents in this case is $(2, 1, 0, 0, 0, \dots)$. Again, fundamental theorem of arithmetic guarantees that only this sequence would yield 12 as a result. In that sense, FTA effectively provides the existence of a "DNA" of integers consisting of that sequence of exponents described above. Knowing the DNA of biological beings can help us study them, and the same goes for integers. For example, if d is the number of positive divisors of n, then : $d = (k_1 + 1)(k_2 + 1)(k_3 + 1) \dots$ where $(k_1, k_2, k_3, \dots)$ is the "DNA" sequence of n (further explanation below). The expression on the right makes sense because k_i is eventually always 0 and if $k_i = 0$ then $k_i + 1 = 1$, and multiplying by 1 does nothing, so that the expression on the right is effectively a finite number of multiplications.

Continuing with the example of 12, notice that $12 = 2^2 \cdot 3^1 \cdot 5^0 \cdot 7^0 \ldots$, the number of positive divisors of 12 is exactly $(2+1)(1+1)(0+1)(0+1)\ldots = 3 \cdot 2 = 6$, and this is indeed correct since the positive divisors of 12 are $1, 2, 3, 4, 6, 12$. The reason this process works can be explained by a shift in perspective. Multiplying a number by another can only increase each element of its DNA. Indeed, given an integer x, the integer $12x$ will have atleast a 2 in the exponent of the prime 2 and atleast a 1 in the exponent of the prime 3 in its DNA expansion. As such, a number divides another if and only if its DNA expansion is pointwise smaller.

Here is another example : $180 = 2^2 \cdot 3^2 \cdot 5^1 \cdot 7^0 \cdot 11^0 \ldots$ and $12 = 2^2 \cdot 3^1 \cdot 5^0 \ldots$. The DNA of 180 is $(2, 2, 1, 0, 0, \ldots)$ and the DNA of 12 is $(2, 1, 0, 0, 0, 0, \ldots)$. As we can see, the DNA of 12 is pointwise smaller than the DNA of 180, and as such $12|180$ which is exactly the case since $180 = 12 \cdot 15$. Notice also that $15 = 2^0 \cdot 3^1 \cdot 5^1 \cdot 7^0 \ldots$, and that the DNA of 15 is exactly what remains from the DNA of 180 after substracting the DNA of 12.

From all this, we can deduce that the number of positive divisors of a positive integer is the number of sequences of DNA smaller than its own DNA. Now if the DNA of n is $(k_1, k_2, k_3 \ldots)$, then we have $k_1 + 1$ choices for the first slot (all integers from 0 to k_1 included), $k_2 + 1$ choices for the second slot (all integers from 0 to k_2 included), and so on. By the multiplicative principle, this gives us the formula above.

The whole point of this discussion is to showcase how important the DNA of positive integers is, and how knowing it effectively tells you a lot of information on it (like the number of divisors or its image by φ). Some cryptographic concepts are even based on the idea of not knowing the prime factorization of a number, i.e its DNA.

As such, we would like to extract some information or elements of the DNA sequence of integers. To do so, we would like to call a function on an integer to extract a certain exponent from its prime factorization. This is the whole idea behing *p-adic functions*.

Definition 1.1: p-adic Valuation

Let p be a prime number and let n be a positive integer. We denote by $\nu_p(n)$ the greatest non-negative integer k for which $p^k | n$. The number $\nu_p(n)$ is called the *p-adic valuation of n*.

The definition above reflects the discussion we've been having so far. $\nu_p(n)$ is simply the exponent of the prime p in the "DNA expansion" of n. For example, $12 = 2^2 \cdot 3^1 \cdot 5^0 \cdot 7^0 \cdot 11^0 \ldots$, and as such $\nu_2(12) = 2, \nu_3(12) = 1, \nu_5(12) = 0, \nu_7(12) = 0$ and so on. ν_p itself can be viewed as a function from positive integers to non-negative integers. For example, here are the first 16 values of ν_2 :

n	$\nu_2(n)$
1	0
2	1
3	0
4	2
5	0
6	1
7	0
8	3
9	0
10	1
11	0
12	2
13	0
14	1
15	0
16	4

Try to convince yourself why each entry of the table above is correct. Again, remember that it simply represents the exponent of 2 in the prime factorization of n. Some easy but important remarks are that $\nu_p(n) = 0$ if and only if $p \nmid n$, and that naturally $\nu_p(n) \geq 1$ if and only if $p|n$. These follow from the very definition of ν_p as the greatest exponent of p in n's prime factorization. Restating the result of number of divisors above, we find ourselves with the following proposition:

Proposition 1.2: Number of Divisors

Let n be a positive integer and let $\tau(n)$ denote its number of positive divisors. Then :

$$\tau(n) = \prod_{\substack{p|n \\ p \text{ prime}}} (\nu_p(n) + 1)$$

Try to convince yourself that this is just a restatement of our original discussion. We have limited our range in the product over the primes p dividing n, but this is, in fact, not necessary. If we range over all primes p, then $\nu_p(n) + 1 = 1$ for $p \nmid n$ so

it won't affect the product, as discussed above.

Another reason why ν_p is an interesting function is because of its interesting set of properties, some of which I have compiled in the following proposition:

Proposition 1.3: Properties of p-adic Valuation

Let m, n be two positive integers and let p be a prime number. Then :
(i) $n = \prod_{\substack{q|n \\ q \text{ prime}}} q^{\nu_q(n)}$.
(ii) $\nu_p(mn) = \nu_p(m) + \nu_p(n)$
(iii) $\nu_p(m+n) \geq \min(\nu_p(n), \nu_p(m))$ with equality if $\nu_p(m) \neq \nu_p(n)$.

Before proving the proposition above, let us explore some of the properties that it describes in order to get a proper understanding of what they are saying. Property (i) is simply a restatement of what ν_p is : the exponent of n in its "DNA" expansion. The same remark said to Proposition 1.2 applies here : we could range over all primes and it wouldn't matter since $q^0 = 1$.

The second property (ii) is by far the most interesting of the bunch. It encompasses the whole "divisor" discussion that we had above, and even beyond that. Remember when we saw that 180's DNA is $180 = 2^2 \cdot 3^2 \cdot 5^1 \cdot 7^0 \ldots$ i.e $(2, 2, 1, 0, 0, \ldots)$ and that 12's DNA is $(2, 1, 0, 0, 0, \ldots)$? Well, $180 = 12 \cdot 15$ and 15's DNA is $(0, 1, 1, 0, 0, \ldots)$. As noticed before, 15's DNA is exactly the "difference" between 180's DNA and 12's. This is a direct restatement of property (ii). Indeed, $\nu_3(180) = \nu_3(15) + \nu_3(12)$, i.e that the second element of 12 and 15's DNA sum up to the second element of 180's DNA.

The third property is fundamental on a theoretical sense as it can help find some bounds on the p-adic valuation of sums. What the property is saying is that if p^a divides n and p^b divides m, $p^{\min(a,b)}$ divides $n+m$, which makes sense since it divides both p^a and p^b so that it divides both n and m. For example, $4|12$ and $4|16$ so that $4|12 + 16 = 28$. Since $\nu_2(16) = 4$ and $\nu_2(12) = 2$, then we can directly say that $\nu_2(28)$ is the minimum of the two, i.e 2.

Proof of Proposition 1.3:

(i) Direct consequence of FTA and definition of $\nu_q(n)$.

(ii) Let $a = \nu_p(n), b = \nu_p(m)$ and $c = \nu_p(mn)$. We want to show $c = a + b$.

First of all, notice by definition that $p^a|n$ aka $n = p^a n'$ and $p^b|m$ aka $m = p^b m'$ so that $p^{a+b}|mn = p^{a+b}n'm'$ and hence $c \geq a + b$. Now suppose for the sake of contradiction that $c > a + b$, that means that $p^{a+b+1}|p^{a+b}m'n'$ and hence $p|m'n'$. But since $p \nmid m'$ and $p \nmid n'$ and p is prime then this is impossible. So $c = a+b$, as desired.

(iii) Let $a = \nu_p(n)$ and $b = \nu_p(m)$. Let $c = \min(a, b)$. Since $c \leq a$ and $c \leq b$, then $p^c|p^a$ and $p^c|p^b$. But $p^a|n$ and $p^b|m$. Thus $p^c|n$ and $p^c|m$. Hence $p^c|m + n$ so $\nu_p(m + n) \geq c$. Now if $a \neq b$, then write $n = p^a n'$ and $m = p^b m'$. Thus $\dfrac{m+n}{p^c} = p^{a-c}n' + p^{b-c}m'$. Since $a \neq b$, either $a - c > 0$ or $b - c > 0$ (but not both since either $a = c$ or $b = c$, as $c = \min(a, b)$). As such, p divides one of the two terms but not the other (since $p \nmid m'n'$). Hence $p \nmid \dfrac{m+n}{p^c}$ and so $\nu_p(m + n) = c$, as desired. □

The properties of the p-adic valuation as a function are fundamental when it comes to the study of integers. It turns out that ν_p can be generalized as a concept to the set of rational numbers and can lead to an extension of rationals into a new number system called the p-adic numbers, but this is out of the scope of this course.

2 Arithmetic Functions

In this chapter, we will discuss a bunch of arithmetic functions and talk about some of their properties. Let us recall the definition of an arithmetic function:

Definition 2.1: Arithmetic Function

An *arithmetic function* f is a function $f : \mathbb{N} \to \mathbb{C}$, i.e a function taking positive integer inputs and that returns complex numbers.

Even if you are not familiar with complex numbers, this definition should not scare you. Technically speaking, 2 *is* a complex number, so when we say that f returns complex numbers, we simply mean that the range of f is a subset of $\mathbb{C}$, which will usually consist of $\mathbb{N}$, the set of positive integers, but not always.

We have already seen multiple examples of arithmetic functions, such as φ, τ and ν_p for p prime. We will now discuss on general properties of arithmetic functions and introduce others. But before that, here is another definition:

Definition 2.2: Multiplicative Functions

Let f be an arithmetic function. We say that f is *multiplicative* if for every positive integers m, n for which $\gcd(m, n) = 1$ we have that

$$f(mn) = f(m)f(n)$$

.

This definition may seem quite random at first, but it turns out that it matches the behavior of a lot of arithmetic functions. What we are saying here is that we can know $f(mn)$ given $f(m)$ and $f(n)$ for as long as $\gcd(m, n) = 1$. For example, if I tell you that f is multiplicative, and I give you $f(2)$ and $f(3)$, then $f(6) = f(2)f(3)$ since $\gcd(3, 2) = 1$.

We already know a multiplicative arithmetic function from the second lecture, which is Euler's totient function φ. Indeed, if $\gcd(m, n) = 1$ then $\varphi(mn) = \varphi(m)\varphi(n)$.

We now explore some other examples of multiplicative arithmetic functions. As explained above, τ is the function that maps each integer n to its number of positive divisors. One of the ways we can write τ is the following:

$$\tau(n) = \sum_{d|n} 1$$

The big sigma means sum, and the indice below it means that we're summing over all positive divisors d of n. As such, for each positive divisor d we "touch" or "range over", we add 1. This is quite an algorithmic formula and doesn't really tell us much about the divisors function at first sight. However, it can help us show a very important property of τ.

Proposition 2.3: Multiplicative Sigma Functions

Let α be any real number. Define $\sigma_\alpha : \mathbb{N} \to \mathbb{C}$ an arithmetic function that sends a positive integer n to the following :

$$\sigma_\alpha(n) = \sum_{d|n} d^\alpha$$

Then, σ_α is multiplicative.

To prove the proposition above, we will make use of the following short but

useful lemma:

Lemma 2.4

Let m and n be two positive integers s.t $\gcd(m,n) = 1$, and let $d|mn$. Then d can be written uniquely as a product $d = ab$ where $a|n$ and $b|m$.

Proof :

Let $a = \gcd(n,d)$ and $b = \gcd(m,d)$. Since $\gcd(m,n) = 1$, then $d = \gcd(mn,d) = \gcd(m,d)\gcd(n,d) = ab$. Since $a|n$ and $b|m$, this proves existence. Now suppose $ab = a'b'$ where $a, a'|n$ and $b, b'|m$. We have that $a|a'b'$ but $\gcd(a,b') = 1$ since $\gcd(m,n) = 1$. Hence $a|a'$. Similarly, $a'|a$ so $a = a'$. Thus $b = b'$ and this proves uniqueness. □

This lemma is quite useful as it establishes a one-to-one correspondence between pairs (a,b) st $a|n$ and $b|m$ and the divisors of mn whenever $\gcd(m,n) = 1$. We shall make use of this correspondence in the following proof.

Proof of Proposition 2.3 :

Let m, n be two coprime positive integers. Then by **Lemma 2.4** :

$$\sigma_\alpha(mn) = \sum_{d|mn} d^\alpha = \sum_{\substack{a|n \\ b|m}} (ab)^\alpha = (\sum_{a|n} a^\alpha)(\sum_{b|m} b^\alpha) = \sigma_\alpha(n)\sigma_\alpha(m)$$

The third equality can be understood as follows : To get all possible products $(ab)^\alpha$, we sum all possible a^α's, sum all possible b^α's, and take the product of the two sum. This gives us all possible pairs $(ab)^\alpha$ after distributing so the equality is true. □

By letting $\alpha = 0$ in **Proposition 2.3**, we see that $\sigma_0 = \tau$, and as such, τ is a multiplicative arithmetic function. But this proposition gives us a whole family of multiplicative arithmetic functions. For example, letting $\alpha = 1$ gives us the "sum of divisors" function. What the proposition is effectively stating in that case is that the sum of divisors of mn whenever $\gcd(m,n) = 1$ is the product of the sum of divisors of m and of n. $\alpha = 2$ yields the sum of square of divisors function, and so on.

Having seen a whole bunch of multiplicative functions, let us now take a look at one important property that they all share and that can help us prove more than one useful identity.

Proposition 2.5

Let f and g be two arithmetic multiplicative function such that $f(p^k) = g(p^k)$ for all positive integer k and for all prime number p. Then $f = g$ (i.e f and g are the same function).

Proof : Let n be a positive integer. Write $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$ where p_i's are primes and $\alpha_i \geq 1$. Then :

$$f(n) = f(p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}) = f(p_1^{\alpha_1}) f(p_2^{\alpha_2}) \dots f(p_k^{\alpha_k}) = g(p_1^{\alpha_1}) g(p_2^{\alpha_2}) \dots g(p_k^{\alpha_k}) = g(n)$$

so that $f = g$, as desired. □

This means that multiplicative functions are fully defined by their images at prime powers. If you know that f is multiplicative, and you know $f(p^k)$ for every prime power p^k, then you can effectively recover the value of f at any integer n. This is simply a generalization of what we have seen with the φ function to arbitrary multiplicative functions.

Let us note that some arithmetic functions f are called *completely multiplicative* if they satisfy $f(mn) = f(m)f(n)$ for every pair of positive integers (m, n) regardless of what their greatest common divisor is. An example of such a function would be $f(x) = x^2$ since $f(mn) = (mn)^2 = m^2 n^2 = f(m)f(n)$ for all positive integers m, n. In that case, f is fully determined by its image at primes instead of prime powers.

Despite not being multiplicative, the p-adic valuation ν_p for a prime p is a function that is fully determined by its images at primes, due to the fact that $\nu_p(xy) = \nu_p(x) + \nu_p(y)$ for every positive integers x and y. In fact, ν_p is the only function that satisfies this "logarithmic" condition and $\nu_p(p) = 1$ and $\nu_q(p) = 0$ for all primes $q \neq p$.

3 Dirichlet Convolution

We now consider an intriguing way of combining arithmetic functions. But before, let us look at an interesting property of φ.

Theorem 3.1

Let n be a positive integer, and let φ denote Euler's totient function. Then:

$$\sum_{d|n} \varphi(d) = n$$

Proof:

Consider the set $S = \{1, 2, \ldots, n-1, n\}$ of integers from 1 to n. We now classify these integers depending on their gcd with n, by defining a set $S(d)$ associated to each divisor d of n as follows:

$$S(d) = \{1 \leq k \leq n | \gcd(k, n) = d\}$$

meaning $S(d)$ is the set of integers k from 1 to n who'se GCD with n is d. Define a_d as the number of element in $S(d)$, where $d|n$. Then :

$$n = \sum_{d|n} a_d$$

because the sets $S(d)$'s partition the set S. But $a_d = \varphi(\frac{n}{d})$ since $\gcd(n, k) = d \iff \gcd(\frac{n}{d}, \frac{k}{d}) = 1$, so there is a one-to-one correspondence between the set $\{1 \leq k \leq n | \gcd(k, n) = d\}$ and $\{1 \leq i \leq \frac{n}{d} | \gcd(i, \frac{n}{d}) = 1\}$.

This means that $n = \sum_{d|n} \varphi(\frac{n}{d})$. Since $\frac{n}{d}$ ranges over all divisors of d, we can just write it as $n = \sum_{d|n} \varphi(d)$, as desired. □

In order to get a deeper understanding of such expressions, we would like to get a better handle on expressions that sum over divisors. To do so, we define the following operation on arithmetic function.

Definition 3.2: Dirichlet Convolution

Let f and g be two arithmetic functions. We define their *Dirichlet convolution* h as the following arithmetic function:

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

and we write $h = f * g$.

Under this definition, **Theorem 3.1** states that if f is defined as $f(n) = n$ for

all integers n, and g is defined as $g(n) = 1$ for all integers n, then $f = \varphi * g$. This function g is well-known as the *unit* function and is denoted by u.

$$u(n) = 1 \text{ for all positive integers } n.$$

Here are the fundamental properties of Dirichlet convolution as an operation, namely *commutativity* and *associativity*.

Proposition 3.3: Commutativity / Associativity

Let f, g and h be three arithmetic functions. Then:
(i) $f * g = g * f$ (Commutativity)
(ii) $f * (g * h) = (f * g) * h$ (Associativity)

Commutativity essentially says that the order of operation does not matter

Proof :

(i) This is a consequence of the definition of Dirichlet convolution, since ranging over divisors "from left to right" is the same as ranging from divisors "from right to left" so that for all positive integers n :

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

(ii) Let n be a positive integer. Then:

$$(f * (g * h))(n) = \sum_{d|n} f(d) \sum_{q|\frac{n}{d}} g(q)h(\frac{n}{dq}) = \sum_{abc=n} f(a)g(b)h(c)$$

so $(f * (g * h))(n)$ is simply the sum of $f(a)g(b)h(c)$ for all triples (a, b, c) for which $abc = n$. Since this expression does not depend on the order, its clear that $((f * g) * h)(n)$ yields the same result. □

These essential properties of Dirichlet convolution ensures that it works the same way as our usual operations like addition or multiplication.

We now try to expand the similarities between Dirichlet convolution and usual multiplication. Is there a function I for which $f * I = f$? Well, it turns out there is one! If we pick $I(n) = 0$ for every positive integer $n \geq 2$, and $I(1) = 1$, then for

every arithmetic function f and positive integer n:

$$(f * I)(n) = \sum_{d|n} f(d)I(\frac{n}{d}) = f(n)$$

since $I(k) = 0$ for $k > 1$. Hence $f * I = f$, and I serves as our "identity function".

Having an identity, one may consider the following concept: Given a function f, can we find an inverse function g for which $f * g = I$? If so, this would make us able to cancel out functions on certain relations as well as "bring them to the other side", as per usual mathematical operations.

Well, the answer to this question is that an inverse exists for *almost* every f.

Proposition 3.4: Existence of Inverses

Let f be an arithmetic function. There exists a function g for which $f * g = I$ if and only if $f(1) \neq 0$.

Proof :
If $f * g = I$ for some function g, then $I(1) = f(1)g(1) = 1$ and hence $f(1) \neq 0$.
Now suppose that $f(1) \neq 0$, then we can see that if a function g satisfies $f * g = I$, then $g(1) = \frac{1}{f(1)}$. In general for $n \geq 2$, knowing $g(1), g(2), \ldots, g(n-1)$, we can see that $\sum_{d|n} f(d)g(\frac{n}{d}) = 0$ so that $g(n) = -\frac{1}{f(1)} \cdot \sum_{\substack{d|n \\ d>1}} f(d)g(\frac{n}{d})$. So we can recursively construct g, and hence it exists. □

3.1 Möbius Inversion Formula

To conclude, we introduce the following arithmetic function called *the Möbius function.*

Definition 3.5: Möbius function μ

We define $\mu : \mathbb{N} \to \{-1; 0; 1\}$ as follows :
(i) $\mu(1) = 1$
(ii) $\mu(n) = 0$ if $p^2|n$ for some prime p.
(iii) $\mu(p_1p_2p_3 \ldots p_k) = (-1)^k$ for $p_1, p_2, \ldots p_k$ distinct primes.

This definition seems very odd at first sight. Try to compute the first 10 values of μ and compare them with the table below.

To understand this function, we need to make sure to understand what it really counts. First, let us recall that a an integer n is called *squarefree* if n is a product of distinct primes. First, μ ignores non-squarefree integers. Then, it returns the parity of the number of primes that divide d if n is a product of distinct primes. It returns 1 if that number is even, and -1 otherwise.

n	$\mu(n)$
1	1
2	-1
3	-1
4	0
5	-1
6	1
7	-1
8	0
9	0
10	1

The reason to define this convoluted function is that it has the very interesting following property:

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases}$$

Indeed, it is clear that this is true if $n = 1$. However, if $n > 1$, then every non-square free divisor disappears in the sum, so we may as well assume that n is squarefree (reduce every exponent to 1). There is a prime $p|n$, and thus :

$$\sum_{d|n} \mu(d) = \sum_{d|\frac{n}{p}} \mu(d) + \sum_{d|\frac{n}{p}} \mu(pd) = \sum_{d|\frac{n}{p}} \mu(d) - \mu(d) = 0$$

since $p \nmid d$ when $d|\dfrac{n}{p}$

In the language of Dirichlet Convolution, this means that $\mu * u = I$, i.e that μ is the inverse of the unit function u. And thus, we get the following relation : If f denotes the function for which $f(n) = n$ for every integer n, then $f = \varphi * u$, and hence $f * \mu = \varphi * u * \mu = \varphi * I = \varphi$, and hence for every positive integer n:

$$\varphi(n) = \sum_{d|n} d \cdot \mu(\frac{n}{d})$$

In general, this gives us the following formula, known as *Möbius' Inversion Formula.*

Proposition 3.6: Möbius' Inversion Formula

Let f and g be two arithmetic functions such that:

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

for every positive integer n. Then:

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

Proof:
$f = g * u$ and as such $f * \mu = g * u * \mu = g * I = g$, as desired. □

There also exists a product version for this formula, which can be described as follows:

$$f(n) = \prod_{d|n} g(d) \iff g(n) = \prod_{d|n} f(d)^{\mu(\frac{n}{d})}$$

which can be seen by simply taking the log of both sides, giving us the summation version.

The point of discussing Dirichlet convolution and Möbius Inversion Formula is to study expressions that range over divisors of an integer. The powerful ideas behind this operation can help us reduce problems and their complexity by "avoiding" the direct approach of simply ranging over divisors, when we can get parallel information by looking at inverted expressions.