Analytic Number Theory, undergrad, Courant Institute, Spring 2017
http://www.math.nyu.edu/faculty/goodman/teaching/NumberTheory2017/index.html

Section 2, Euler products

version 1.2 (latest revision February 8, 2017)

1 Introduction.

This section serves two purposes. One is to cover the Euler product formula for the zeta function and prove the fact that

$$\sum_p p^{-1} = \infty \, . \tag{1}$$

The other is to develop skill and tricks that justify the calculations involved.

The zeta function is the sum

$$\zeta(s) = \sum_1^\infty n^{-s} \, . \tag{2}$$

We will show that the sum converges as long as $s > 1$. The Euler product formula is

$$\zeta(s) = \prod_p \left(1 - p^{-s}\right)^{-1} \, . \tag{3}$$

This formula expresses the fact that every positive integers has a unique representation as a product of primes. We will define the infinite product, prove that this one converges for $s > 1$, and prove that the infinite sum is equal to the infinite product if $s > 1$.

The derivative of the zeta function is

$$\zeta'(s) = -\sum_1^\infty \log(n)\, n^{-s} \, . \tag{4}$$

This formula is derived by differentiating each term in (2), as you would do for a finite sum. We will prove that this calculation is valid for the zeta sum, for $s > 1$. We also can differentiate the product (3) using the Leibnitz rule as

though it were a finite product. In the following calculation, r is another prime:

$$\begin{aligned}
\zeta'(s) &= \sum_p \left\{ \left[\frac{d}{ds} \left(1 - p^{-s}\right)^{-1} \right] \sum_{r \neq p} \left(1 - r^{-s}\right)^{-1} \right\} \\
&= -\sum_p \left\{ \left[\log(p) p^{-s} \left(1 - p^{-s}\right)^{-2} \right] \sum_{r \neq p} \left(1 - r^{-s}\right)^{-1} \right\} \\
&= -\left\{ \sum_p \log(p)\, p^{-s} \left(1 - p^{-s}\right)^{-1} \right\} \zeta(s) \\
\zeta'(s) &= -\left\{ \sum_p \left(\log(p) \sum_1^\infty p^{-ks} \right) \right\} \zeta(s) \ . \qquad (5)
\end{aligned}$$

If we divide by $\zeta(s)$, the sum on the right is a sum over prime powers (numbers n that have a single p in their prime factorization). This is expressed using the *von Mangoldt* function

$$\Lambda(n) = \begin{cases} \log(p) & \text{if } n = p^k \text{ for an integer } k > 0 \text{ and prime } p \\ 0 & \text{otherwise.} \end{cases} \qquad (6)$$

The final formula is

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_n \Lambda(n)\, n^{-s} \ . \qquad (7)$$

To justify these calculations, we prove a theorem about differentiating an infinite series term by term and a theorem about changing the order of summation.

The product and sum formulas for $\zeta(s)$ work together to prove the divergence of the sum of (1). Using the sum formula (2), we show that $\zeta(s) \to \infty$ as $s \downarrow 1$. The notation $s \downarrow 1$ means that we take the limit as s approaches 1 through values $s > 1$. However, we analyze the product formula (3) to see that $\zeta(s)$ is bounded as $s \downarrow 1$ unless the sum p^{-1} diverges.

Mathematicians often do things in more generality and with more abstraction than necessary for a particular problem. They do this to find the simplest version of a problem. The abstract version may be simpler because it has less irrelevant structure and detail. The abstract version also may serve to "kill two birds with one stone". One general fact justifies many particular calculations. Mathematicians search for general formulations when they find more than one problem of a similar form. For example, a mathematician may look at the zeta function sum (2) and the derivative formula (7) and start talking about general *Dirichlet series* of the form

$$f(s) = \sum_1^\infty a_n n^{-s} \ . \qquad (8)$$

Any theorem about general Dirichlet series automatically applies to both examples, and to many more that are coming. When differentiating a series like (2),

a mathematician may ask about the more general problem of differentiating a function that is a general sum of functions like

$$f(s) = \sum_1^\infty b_n(s) \; . \tag{9}$$

The question would be: find "nice" condition on the functions $b_n(s)$ that imply that

$$f'(s) = \sum_1^\infty b_n'(s) \; . \tag{10}$$

Nice means that the condition is be easy to check and applies in many specific examples we are interested in.

2 Sizes and conventions

For many of you this section will be as interesting as the safety instructions at the beginning of a plane flight. But it has to be said, so ...

These definitions are for the limit $x \to \infty$. The term *asymptotic behavior* of a $f(x)$ refers to approximate descriptions of $f(x)$ that become increasingly accurate as $x \to \infty$. I will write $f(x) \approx g(x)$ to mean that $g(x)$ is an increasingly accurate approximation of $f(x)$ as $x \to \infty$.[1] "Big Oh" and "little Oh" notation are devices for being a little more formal with statements about relative size. If $g(x) \geq 0$ for $x > 0$ we write

$$f(x) = O(g(x))$$

if there is a C and an x_0 so that if $x > x_0$, then $f(x) \leq Cg(x)$. Some people might want $f(x) \geq 0$ in this definition, so we don't say $-x^2 = O(x)$. If you want to talk about absolute value, use absolute value signs as in $|f(x)| = O(g(x))$. You are supposed to say "$f(x)$ is of the order of $g(x)$", but this can be misleading. The actual definition allows the possibility that $f(x)$ is much smaller than $g(x)$. For example, $x = O(x^2)$. This is the "big Oh".

"Little Oh" is for saying one quantity is (asymptotically) smaller than another. We say $f(x) = o(g(x))$ if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0 \; .$$

This presumes $f(x) \geq 0$, or doesn't quite capture our intent otherwise. An equivalent definition of little Oh is that for any $\epsilon > 0$ there is an x_0 so that

$$f(x) \leq \epsilon g(x) \quad , \quad \text{if } x > x_0 \; .$$

[1] Specialists in analytic number theory use notations for this that are not shared by people in other branches of mathematics or science. For example, they write $f(x) \ll x$ not to mean that $f(x)$ is much smaller than x, but to mean $f(x) = O(x)$. You may encounter notation like this in the references. This class avoids these in favor of notation more common to other parts of math and science.

Bigger powers of x are always bigger in this sense, and constants and logs don't change that. For example

$$(\log(x))^2 \, x^{\frac{1}{2}} = o(x) \; .$$

Applied mathematicians often write $f(x) \sim g(x)$ to mean that $f(x)$ and $g(x)$ are *asymptotically equivalent*, which means that the difference between them is smaller than they are:

$$|f(x) - g(x)| = o(g(x)) \; .$$

Unfortunately, others use the $\sim$ symbol to mean other things, so I will use the less formal $f(x) \approx g(x)$ for this. As an example, consider the *logarithmic integral*

$$\mathrm{li}(x) = \int_2^x \frac{1}{\log(y)} \, dy \; .$$

This satisfies (it's an exercise, literally. Look for it in the homework.)

$$\mathrm{li}(x) \approx \frac{x}{\log(x)} \;\; , \text{ which means } \;\; \left| \mathrm{li}(x) - \frac{x}{\log(x)} \right| = o\left(\frac{x}{\log(x)} \right) \; .$$

Computer scientists write $f(x) = \Theta(g(x))$ to mean $f(x) = O(g(x))$ and $g(x) = O(f(x))$. It would be great if analytic number theorists would use this notation, but they generally don't.

Big Oh notation is used in formulas to indicate the size of a discrepancy. For example, we will show that if $s > 1$, then

$$\sum_N^\infty n^{-s} = \int_N^\infty x^{-s} \, dx + O(N^{-s}) \; .$$

This means that there is an N_0 and a C so that if $N > N_0$, then

$$\left| \sum_N^\infty n^{-s} - \int_N^\infty x^{-s} \, dx \right| \leq C N^{-s} \; .$$

Note that we are using $N \to \infty$ instead of x. The expression $O(1)$ is interpreted as $g(x) = 1$ being the constant function. For example, we will show that

$$\zeta(s) = \frac{1}{s-1} + O(1) \; , \;\; \text{as } s \downarrow 1 \; .$$

This means that there is an $s_0 > 1$ and a C so that if $1 < s < s_0$, then

$$\left| \zeta(s) - \frac{1}{s-1} \right| \leq C \cdot 1 = C \; .$$

3 The zeta function and the prime number theorem, informally

Mathematicians spend as much time guessing as they do proving. One form of guessing is "optimistic" formal manipulation of the kind that led to (7). You will get a sense of which calculations are easy formulate in a mathematically rigorous way. It may be argued that the non-rigorous version should be called *informal*. If so, it is possible to become even less formal and more speculative. Here is such an informal discussion of the role of the zeta function and the logarithmic derivative in the prime number theorem.

The prime number theorem is about the large x behavior of prime numbers. The counting function for prime numbers, the number of primes less than x, is

$$\pi(x) = \sum_{p<x} 1 \; . \tag{11}$$

Two related functions are

$$\phi(x) = \sum_{p<x} \log(p) \; , \tag{12}$$

and

$$\psi(x) = \sum_{n<x} \Lambda(n) = \sum_{p^k<x} \log(p) \; . \tag{13}$$

Informally, let $\rho(x)$ be the "probability" that a given number about x is prime. Or, we could say that $\rho(x)$ is the density of primes near x. The theorem, informally, is that for large x,

$$\rho(x) \approx \frac{1}{\log(x)} \; . \tag{14}$$

If this is true, then the counting function should be approximately[2]

$$\pi(x) \approx \int_2^x \rho(y)\, dy \approx \int_2^x \frac{1}{\log(y)}\, dy \; . \tag{15}$$

The last integral is the "logarithmic integral", or

$$\mathrm{li}(x) = \int_2^x \frac{1}{\log(y)}\, dy \; . \tag{16}$$

It is not hard to show that

$$\mathrm{li}(x) = \frac{x}{\log(x)} + o\left(\frac{x}{\log(x)}\right) \; . \tag{17}$$

[2]The integral diverges as $x \to \infty$. The statement is that $\pi(x)$ and $\mathrm{li}(x)$ diverge in the same way. Therefore, the lower limit $y = 2$ is irrelevant. We use 2 because $\log(y)$ blows up at $y = 1$, which otherwise might have been a more natural lower limit.

The actual, rigorous, prime number theorem is

$$\pi(x) = \frac{x}{\log(x)} + o\left(\frac{x}{\log(x)}\right) \ . \tag{18}$$

The prime number theorem also gives the large x behavior of $\phi(x)$ and $\psi(x)$. For $\phi(x)$, we have the informal formula

$$\phi(x) \approx \int_2^x \frac{\log(y)}{\rho(y)}\, dy \approx x \ .$$

The theorem is

$$\phi(x) = x + o(x) \ . \tag{19}$$

It is hard to prove (18) or (19), but is is easier to show that either of these statements implies the other. The third equivalent statement of the prime number theorem is

$$\psi(x) = x + o(x) \ . \tag{20}$$

If you believe either the ψ estimate (20) or the ϕ estimate (19) then it is routine[3] to prove that ϕ and ψ are close to each other

$$\psi(x) - \phi(x) \leq C \log(x) x^{\frac{1}{2}} \ . \tag{21}$$

Note that $\psi(x) \geq \phi(x)$ because the ψ sum (13) includes all the prime terms in the ϕ sum (12), plus the extra terms from prime powers. "Close" is a relative term, but the right side of the inequality (21) is smaller than ϕ or ψ. This is because $x^{\frac{1}{2}}$ is smaller than x, and putting in the log doesn't change this.

We come back to the log derivative function function, particularly as $s \downarrow 1$. The sum diverges when $s = 1$, so the behavior as $s \downarrow 1$ may tell us how it diverges. The closeness of $\phi(x)$ and $\psi(x)$ is related to the fact that prime powers are more rare than primes. Let us use this, and the idea of replacing sums with integrals, in log derivative formula (7). In the integral, the $\log(y)$ in the numerator is from $\Lambda(n)$ and the $\log(y)$ in the denominator is from $\rho(x)$. With all these "approximations", we get

$$\frac{\zeta'(s)}{\zeta(s)} \approx \int_2^\infty \frac{\log(y)}{\log(y)}\, y^{-s}\, dy \approx \frac{1}{s-1} \ . \tag{22}$$

It is "routine" to show that

$$\frac{\zeta'(s)}{\zeta(s)} - \frac{1}{s-1}$$

is bounded as $s \downarrow 1$. This is "evidence" in favor of the prime number theorem. For example, if $\rho(x)$ were $\frac{2}{\log(x)}$, then the right side of (22) would be $\frac{2}{s-1}$. In fact (22) is one of several ways of seeing that if the primes have a simple density, then that density must be $\frac{1}{\log(x)}$.

[3] *Routine* means that it's the kind of thing a practiced professional can do in an hour or so.

This motivates Riemann's approach to the prime number theorem. The density of primes for large x (should such a density be meaningful) is related to the behavior of $\zeta(s)$ (actually $\zeta'(s)/\zeta(s)$) for s near 1. The hard part is to go the other way, from understanding ζ near $s = 1$ to the large x density of primes. Riemann's approach to that was to look at $\zeta(s)$ when $s = \sigma + it$ is a complex number. It's traditional to use σ and t for the real and imaginary parts of s when you talk about $\zeta(s)$. It isn't just s near 1 that matters, but σ near 1 for all t. Much of this class is devoted to this.

4 Prime factorization and Euclid's algorithm.

[*Instructor's warning: I am not very good at basic algebra and am reluctant to copy it from a book. Therefore, this section might be longer than the slicker version in another book. Feel free to look up gcd and unique prime factorization in another source rather than reading this. However, I want to convey the sense, which has been important in my mathematical career, that it is possible to make proofs of lemmas once you have a general idea how things work. You understand things better that way. It's how I got the sense that I possibly could do mathematics. This is what it looks like.*

We say p is a prime number (a prime) if p is an integer $p \geq 2$ and if $p = ab$ with a and b being positive integers, then either $a = 1$ or $b = 1$. Any integer $n \geq 2$ may be *factored* into primes, written as a product of primes. This is "easy" to show. It is harder to show that prime factorization is unique. For example, if $n = p_1 p_2$ and $n = p_3 p_4$, (all p_k being primes), then either $p_1 = p_3$ or $p_1 = p_4$. Primes in the prime factorization may be repeated, as in $28 = 2 \cdot 2 \cdot 7 = 2^2 \cdot 7$. There may be more primes, as in $90 = 2 \cdot 3^2 \cdot 5$. Let $p_1 < p_2 < \cdots$ be the list of all primes. Unique prime factorization is the statement that for any n, there is a sequence of non-negative integers r_k so that

$$n = p_1^{r_1} \cdot p_2^{r_2} \cdot \ \cdots \ .$$

If $r_k = 0$, then p_k does not occur in the prime factorization of n. There are only finitely many non-zero r_k. The r_k for a given n are unique.

Unique prime factorization will seem natural once you try many examples. The same can be said for the *Goldbach conjecture*, which says that every even

integer $n > 2$ may be expressed as the sum of two primes. Try it:

$$\begin{aligned}
4 &= 2 + 2 \\
6 &= 3 + 3 \\
8 &= 5 + 3 \\
10 &= 7 + 3 = 5 + 5 \\
12 &= 7 + 5 \\
14 &= 11 + 3 = 7 + 7 \\
&\vdots \\
50 &= 31 + 19 = 37 + 13 = 43 + 7 = 47 + 3 \\
52 &= 29 + 23 = 41 + 11 = 47 + 5 \\
&\vdots
\end{aligned}$$

This section contains a proof of unique prime factorization of integers. Attempts to prove the Goldbach conjecture have not succeeded yet, but have led to beautiful contributions to mathematics.

Here is the strategy of the proof, which goes back to Euclid in Alexandria (Egypt, not Virginia) about twenty two centuries ago. Suppose, for example, that $n = p_1p_2$ and $n = p_3p_4$. We want to show that $p_1 = p_3$ or $p_1 = p_4$. We prove something more general. Suppose p is a prime and p divides $n = ab$ (asking that $a \geq 2$ and $b \geq 2$ but not necessarily prime). Then either p divides a or p divides b. Said differently, if p is prime and does not divide a, then p divides b. We will show (this is the heart of the matter) that if p does not divide a, and if p is prime, then there are integers i and j so that $1 = ia + jp$. This may be written as We multiply $n = ab$ by i, and use $ia = 1 - jp$ to get

$$\begin{aligned}
in &= (1 - jp)b \\
b &= in - jbp\ .
\end{aligned}$$

If this formula is true, and if p divides n, then p divides both terms on the right side of the equation and therefore p divides b.

Here is the argument done more formally with the definitions and details. If c and a are integers greater than one, and if $a = kc$ for some (integer) $k > 1$, then we write $c \mid a$ and say c divides a or that c is a *divisor* of a. We write $c \nmid a$ if c is not a divisor of a. We say d is a *common divisor* of a and b if $d \mid a$ and $d \mid b$. The *greatest common divisor* of a and b is written $c = \gcd(a, b)$, or just $c = (a, b)$. It is basic fact that there is a gcd and it is unique. If a and b are positive integers, there is a unique positive integer c so that $c \mid a$ and $c \mid b$. Moreover, if d is any common divisor then $d \mid c$. For example, $(30, 24) = 6$, and, since $3 \mid 30$ and $3 \mid 24$, we know $3 \mid 6$. We say that a and b are *relatively prime* if $(a, b) = 1$.

The *euclidean algorithm*[4] is a way to find the gcd by looking for integers i

[4] You're really successful as a mathematician when your name stops being capitalized.

and j (not necessarily greater than one or even positive) so that

$$(a, b) = ia + jb \; . \tag{23}$$

(In particular, if a and b are relatively prime, then you can write $ia + jb = 1$.) This is done by finding the i and j that gives $ia + jb$ its smallest positive value. This smallest value is the gcd.

We first show that $c \,|\, (ia + jb)$ if $ia + jb > 0$. Since $c \,|\, a$, there is a k with $a = kc$. Similarly, $b = lc$ for some l. Therefore

$$ia + jb = ikc + jlc = (ik + jl)c \; .$$

This shows that $ia + jb = mc$, with $m = ik + jl$.

Let L be the set of all numbers of the form $ia + jb$ for integer i and j. Let c be the smallest positive element[5] of L. Then $c = 1$ or c is some larger integer. Either way, we claim that every other positive element of L is of the form kc. This means that the positive elements of L have the form $c, 2c, 3c, \ldots$, so L may be called a *lattice*. All the numbers kc are in L, because $kc = (ki)a + (kj)b$. Therefore, we have to show that there are no other positive numbers in L. Suppose $d = la + mb$ is an element of L that is not of the form kc for any integer k. Then there is a k so that $kc < d < (k + 1)c$. The crux of the proof is that $d - kc \in L$ and $0 < d - kc < c$, which contradicts c being the smallest positive element of L. The second part, $0 < d - kc < c$, is the same as $kc < d < (k + 1)c$. The first part, $d - kc \in L$, follows from the calculation

$$d - kc = la + mb - k(ia + jb) = (l - ki)a + (m - kj)b \; .$$

To put all this together, we have shown that a and b are positive elements of L. We have shown that c is a divisor of every positive element of L. We also shown that if d divides both a and b, then d divides c. This implies that c is a common divisor of a and b and that any other common divisor of a and b is a divisor of c. That makes c the greatest common divisor of a and b. It also shows that the representation equation (23) may be satisfied. Since L, and the smallest positive element of L, are uniquely defined, this proves that $c = (a, b)$ is unique.

Euclid gave this argument more in terms of an algorithm. You start with $0 < a < b$ (interchange a and b if necessary). You want to find $c = (a, b)$. For any integer k, c is also the gcd of a and $b - ka$. If there is a k with $b - ka = 0$, then b is a multiple of a and $(a, b) = a$. Otherwise, (as in the main step the other way we said it), there is a k so that $0 < b - ka < a$. Choose that k and solve the gcd problem with the new pair $c = (b - ka, a)$. This continues either until $b - ka = 1$ (with the current a and b, not the original ones) or until $b = ka$. You can check that at every step of the current a and b are *integer linear combinations* of the original a and b. We say that a' is an integer linear combination of a and b if

[5]The principle of induction may be viewed as saying that any set of positive integers has a smallest element. You prove something by induction by looking at the set of numbers for which it is false and taking the smallest element of that.

there are integers so that $a' = ia + jb$. In the terminology above, $a' \in L$. The algorithm seeks successively smaller positive elements of L until it's impossible to get any smaller.

For example, if $b = 33$ and $a = 7$, then subtracting a repeatedly gives the sequence $26, 19, 12, 5, -2$. We see that there is no $b - ka = 0$, so b is not a multiple of a (33 is not a multiple of 7). We stop at $b - 4a = 5$ because 5 is in the range $\{0, 1, \ldots, 6\}$. But if we start with $b = 35$ and $a = 7$, the sequence is $28, 21, 14, 7, 0$. We stop at 0 because it is in the range. We learn that $35 = 5 \cdot 7$, so $7 \,|\, 35$.

A prime number is an integer $p \geq 2$ that has no *non-trivial* divisors. The trivial divisors are 1 and p. Let $n \geq 2$ be another integer. Since (n, p) is a divisor of p, the only possibilities are $(n, p) = p$, (p is a divisor of n), or $(n, p) = 1$ (p and n are relatively prime).

We return to unique prime factorization, first in the example $n = p_1 p_2 = p_3 p_4$. We want to show that $p_1 = p_3$ or $p_1 = p_4$. Said another way, if $p \,|\, n$, and if $n = p_3 p_4$, then $p = p_3$ or $p = p_4$. Suppose $p \neq p_3$. Then $(p, p_3) = 1$ and $ip + jp_3 = 1$ for some integers i and j. We multiply the equation $n = p_3 p_4$ by j and substitute $jp_3 = 1 - ip$ (as before) to get

$$\begin{aligned} jn &= jp_3 p_4 \\ jn &= (1 - ip)p_4 \\ p_4 &= jn + ipp_4 \,. \end{aligned}$$

But p is a divisor of both terms on the right side, the first by assumption and the second by formula. Therefore, p is a divisor of p_4, which means $p = p_4$. The same argument works more generally. Suppose $n = ab$ and $p \,|\, n$, then either $p \,|\, a$ or $p \,|\, b$. In fact, if p does not divide a, then $(p, a) = 1$, so $ip + ja = 1$, and $jn + ipa = b$, so p is a divisor of b.

More generally, if $p \,|\, abc$ then $p \,|\, a$ or $p \,|\, b$, or $p \,|\, c$. To see this, note that we already showed that if $p \nmid a$, then $p \,|\, bc$. Therefore, if $p \nmid a$, then $p \,|\, b$, or $p \,|\, c$. This reasoning extends to any number of factors.

Finally, the general theorem. Suppose that n has a prime factorization of the form

$$n = p_1^{r_1} \cdot p_2^{r_2} \cdot \cdots \,. \tag{24}$$

We want to show that the powers r_k are uniquely determined by n. We do this by showing that the numbers r_k are determined by the extent to which p_k divides n. That makes the r_k uniquely determined by n and makes the prime factorization (24) unique. A step in this direction is that if $r_k = 0$, then $p_k \nmid n$. This is where we use the divisibility stuff we just did. There are only finitely many terms in (24) that are not equal to 1. If $p_k \,|\, n$, then p_k divides one of these terms (we just showed). The only term p_k possibly could divide is $p_k^{r_k}$, which implies that $r_k > 0$. Now you apply the same argument to $n/(p_k^{r_k})$, which shows that p_k does not divide $n/(p_k^{r_k}$. Therefore, $n = p_k^{r_k} \cdot m$, where p_k does not divide m, and m has fewer prime factors (primes p_j with $r_j > 0$) than n. Eventually, we see that the whole prime factorization expression (24) is determined by divisibility.

This paragraph and the next are for people who have taken or may take more abstract algebra. In the discussion of gcd and prime factorization, we used addition, subtraction, multiplication, and a cancellation property (if $ab = ac$, then $b = c$) but not division. Addition and multiplication are commutative, associative, and distributive. A *ring* is a system of objects with these properties. Another ring is the *Gaussian integers*, which is the set of complex numbers with integer real and imaginary parts. The set of polynomials in one variable $p(x) = a_n x^n + \cdots + a_0$ is another ring.

We found the gcd (a, b) by considering the set $L = \{ia + jb\}$, i and j elements of the ring. This set has the properties that (i) if k is any element (not necessarily in L) and $d \in L$, then $kd \in L$, (ii) if $d \in L$ and $e \in L$, then $d + e \in L$. A subset like this is called an *ideal*. If c is an element of the ring, there is the *principal* ideal $(c) = \{kc\,, k \text{ in the ring}\}$. For the ring of integers, we showed that any ideal is a principal ideal. That makes the integers a *principal ideal domain*, or *PID*. We did this by showing that there is a measure of size and finding c as the smallest non-zero element of L. A ring with a measure of size that works like this (look up details if interested) is called a *euclidean domain*. We proved that a euclidean domain is a principal ideal domain. The Gaussian integers (see Homework) and ring of polynomials are other examples of euclidean domains. Any Gaussian integer may be factored, uniquely (see homework for the fine print on uniqueness) as a product of powers of prime Gaussian integers. Any polynomial may be factored, again uniquely, as a product of irreducible polynomials. Many rings that some up in more advanced number theory are not principal ideal domains. These rings do not have unique prime factorization. The theory of rings and ideals was invented for these. The big theorem is that (for the a big class of rings used in number theory) any ideal may be factored, uniquely, as a product of prime ideals. Find a book on algebraic number theory for details.

5 Infinite sums

Some of the material here is review of mathematical analysis. How much depends on your analysis background.

Suppose a_n is a sequence of numbers. The infinite sum is[6]

$$A = \sum_1^\infty a_n \ . \tag{25}$$

The *partial sums* are

$$A_N = \sum_1^N a_n = a_1 + \cdots + a_N \ . \tag{26}$$

[6]The numbers form a *sequence* the sum is a *series*. The Taylor *series* is the infinite sum.

The infinite sum is defined to be the limit of the partial sums:

$$A = \lim_{N\to\infty} A_N \ . \tag{27}$$

If the limit exists, we say the sum *converges*. The sum is not defined if the limit does not exist. If $A_N \to \infty$ as $N \to \infty$, we say the sum *diverges*, or possible "diverges to infinity". For example, consider the sum

$$\sum_0^\infty (-1)^n \ ,$$

whose partial sums form the series

$$1,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1,\ \ldots \ .$$

I would say the partial sums "fail to converge" but I probably wouldn't say they diverge. Others might.

The sum converges *absolutely* if the sum

$$\sum_1^\infty |a_n|$$

converges. More generally, suppose $b_n \geq 0$ is a non-negative sequence. Then the partial sums B_N are a monotone non-decreasing sequence. Such a sequence either has a finite limit or diverges to infinity. We write

$$\sum_1^\infty b_n = \infty$$

if the partial sums diverge to infinity. Otherwise, we say

$$\sum_1^\infty b_n < \infty$$

A sum converges absolutely if

$$\sum_1^\infty |a_n| < \infty \ . \tag{28}$$

If a sum converges absolutely, then the sum converges. You can prove this (if you don't remember the proof) by showing that the partial sums form a Cauchy sequence.

A sum converges *conditionally* if it converges but does not converge absolutely. The "condition" is that we not change the order of the summands. It is a theorem of mathematical analysis that the terms in a conditionally convergent sum may be re-arranged to give any answer. An absolutely convergent sum converges *unconditionally*, which means you can add the terms in any order and get

the same answer. If a sum converges conditionally, we say it has *cancellation*, the positive terms and the negative terms "cancel" each other to a large extent. The sum of the positive terms alone would be infinite.

The *geometric series* makes a good example. If $|z| < 1$, then

$$S(z) = \sum_0^\infty z^n = \frac{1}{1-z} \; . \tag{29}$$

We can prove this using the partial sums

$$S_N = \sum_0^N z^n = \frac{1 - z^{N+1}}{1-z} \; .$$

As a reminder, here is the algebraic trick that gives the formula for the partial sums:

$$\begin{aligned} S_N &= 1 + z + \quad \cdots \quad + z^N \\ zS_N &= \quad z + z^2 + \cdots + z^N + z^{N+1} \\ (1-z)S_N &= 1 + 0 + \quad \cdots \quad + 0 \quad - z^{N+1} \\ S_N &= \frac{1 - z^{N+1}}{1-z} \; . \end{aligned}$$

The geometric series converges absolutely for $|z| < 1$, because

$$\sum_0^\infty |z^n| = \sum_0^\infty |z|^n = \frac{1}{1-|z|} < \infty \; .$$

Therefore, for example, you can add the even and odd terms separately, as in

$$\begin{aligned} \sum_0^\infty z^n &= \sum_0^\infty z^{2n} + \sum_0^\infty z^{2n+1} \\ &= \sum_0^\infty z^{2n} + z \sum_0^\infty z^{2n} \\ &= \frac{1}{1-z^2} + z \, \frac{1}{1-z^2} \\ &= \frac{1+z}{1-z^2} \\ &= \frac{1}{1-z} \; . \end{aligned}$$

This is a more complicated route to the same answer.

If a sum converges absolutely, you usually show it by "comparing" it to (bounding it by, more properly) another sum you know converges absolutely. For example if $|z| < 1$, then $\left|z^2\right| < |z|$ and

$$\sum_0^\infty \left|z^{n^2}\right| \le \sum_0^\infty |z|^n < \infty \; .$$

This means that the function

$$f(z) = \sum_0^\infty z^{n^2}$$

is well defined, at least for $|z| < 1$.

For Dirichlet series, the comparison may be to an integral. The zeta function sum (2) is "like" the corresponding integral

$$\int_1^\infty x^{-s}\,dx = \frac{1}{s-1}\ .$$

A term n^{-s} is the area of a rectangle of width 1 and height n^{-s}. This is "like" the area under the curve x^{-s}:

$$\int_n^{n+1} x^{-s}\,dx\ . \tag{30}$$

We will look at this in more detail soon. But this isn't quite what we need to prove the sum converges. For that we need an "upper bound", something larger than n^{-s}, but $x^{-s} < n^{-s}$ if $x > n$, so the integral is smaller than n^{-s}. There are several ways to do this. One way is to find conditions under which

$$2x^{-s} > n^{-s}\ .$$

Some algebra shows that this is true if

$$x < 2^{\frac{1}{s}} n\ .$$

Since $2^{\frac{1}{s}} > 1$, this is true for x in the range $n \le x \le n+1$ if n is large enough. For n that large, we have

$$n^{-s} \le 2\int_n^{n+1} x^{-s}\,dx\ .$$

Another way[7] is to note that of $x \le n$ then $x^{-s} \ge n^{-s}$. Therefore

$$n^{-s} \le \int_{n-1}^{n} x^{-s}\,dx\ .$$

[7] In situations like this there are usually many ways to do things. If you compare your homework with others in the class, you will discover that they often find clever tricks that make your approach look unnecessarily complicated. Sometimes, you may be the one with the more efficient approach. Efficiency in this kind of thing is not very important in the long run.

This implies that

$$\begin{aligned}\sum_1^\infty n^{-s} &= 1 + \sum_2^\infty n^{-s} \\ &\le 1 + \sum_{n=2}^\infty \int_{n-1}^n x^{-s}\, dx \\ &= 1 + \int_1^\infty x^{-s}\, dx \\ &= 1 + \frac{1}{s-1} < \infty\ .\end{aligned}$$

This proves that the zeta sum (2) converges absolutely. If we use the lower bound (30) in the same way, we prove that

$$\frac{1}{s-1} \le \zeta(s) \le 1 + \frac{1}{s-1}\quad , \quad \text{for all } s > 1\ .$$

This implies, for example, that

$$\zeta(s) = \frac{1}{s-1} + O(1)\ , \quad \text{as } s \downarrow 1\ .$$

As $s \downarrow 1$ the sum $\zeta(s)$ and the integral blow up in the same way.

A convergence proof is called "efficient" if it doesn't take very many words to say it. An estimate[8] is called *sharp* if the two sides are close to each other or if there is no stronger inequality of the same type. Sometimes we use estimates that are far from being sharp for the sake of efficiency. As an example, take the convergence of the sum on the right side of (7). We can use the inequality $\Lambda(n) \le \log(n)$. For most n, $\Lambda(n) = 0$. If n is a prime power $n = p_k$, then $\Lambda(n)$ is k times smaller than $\log(n)$. Nevertheless, we have (the second inequality is in the homework)

$$\sum_1^\infty \Lambda(n) n^{-s} \le \sum_1^\infty \log(n) n^{-s} = \frac{1}{(s-1)^2} + O(1)\ , \quad \text{as } s \downarrow 1\ .$$

This is efficient (quick and easy), but the final inequality isn't sharp. We will see that

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_1^\infty \Lambda(n) n^{-s} = \frac{1}{s-1} + O(1)\ , \quad \text{as } s \downarrow 1\ .$$

This is sharper because $(s-1)^{-1}$ is much smaller than $(s-1)^{-2}$ when s is close to 1. You might think you need at least some number theory to prove the sharper bound, because it depends on the fact that $\Lambda(n) = 0$ for most n. The sum of $\log(n) n^{-s}$ really is about $(s-1)^{-2}$. Mathematicians spend a lot of time guessing stuff like this, and they typically guess wrong, at least at first.

[8] Inequalities are often called *estimates* in this context.

6 Limits of sums, differentiating a sum, dominated convergence

We come to the promised formula for differentiating the Dirichlet series (4). We do this much more generally by giving it as an example of term-by-term differentiation (10). We do (10) as an example of the more general problem of interchanging a sum and a limit. Define the sum

$$S(h) = \sum_1^\infty c_n(h) \ , \tag{31}$$

and suppose that

$$\lim_{h \to 0} c_n(h) = d_n \ , \quad \text{for every } n \ .$$

We want conditions that insure that

$$\sum_1^\infty d_n = \lim_{h \to 0} S(h) \ . \tag{32}$$

Part of what we want to know is that the limit on the right exists.

Here's how the general pair (31) and (32) is related to the original issue (10). Suppose the sum (9) converges absolutely for every s. Then we can write the difference quotient as a sum of difference quotients

$$\frac{f(s+h) - f(s)}{h} = \sum_1^\infty \frac{b_n(s+h) - b_n(s)}{h} \ .$$

This has the form (31), if we take

$$S(h) = \frac{f(s+h) - f(s)}{h} \ ,$$

and

$$c_n(h) = \frac{b_n(s+h) - b_n(s)}{h} \ . \tag{33}$$

If b_n is differentiable, then $c_n(h) \to b_n'(s)$ as $h \to 0$. Therefore, the left side of (32) is the right side of (10). The right side of (32) is the left side of (10).

Now, back to the general limit problem (32). It is a tradition in math exposition to begin a discussion like this with a counterexample. The counterexample shows that the theorem is serious, and that you really do need some extra hypothesis to insure that the desired conclusion is true. Not to disappoint, consider the example (with $h > 0$)

$$c_n(h) = he^{-nh} \ .$$

We start the sum from $n = 0$ instead of $n = 1$ to make the algebra simpler:

$$S(h) = h \sum_0^\infty c_n(h) = h \sum_0^\infty \left(e^{-h}\right)^n = \frac{h}{1 - e^{-h}} \ .$$

You can calculate that $S(h) \to 1$ as $h \to 0$. You can also see that

$$d_n = \lim_{h \to 0} c_n(h) = 0 \ , \quad \text{for all } n \ .$$

In this example, the left side of (10) is zero and the right side is 1. We need more hypotheses. It would be nice if the hypothesis were *convenient*, simple to state and easy to check in the examples we care about, such as differentiating the Dirichlet series.

One hypothesis that often is convenient is *dominated convergence*. Dominated convergence means there is a *dominating sequence*, D_n that does not depend on h, but so that

$$|c_n(h)| \leq D_n \ , \quad \text{for all } n \text{ and } h, \tag{34}$$

and

$$\sum_{1}^{\infty} D_n < \infty \ . \tag{35}$$

The inequality (34) states that the sequence D_n *dominates* the sequence $c_n(h)$. for all h. The *dominated convergence theorem* states[9] that if there is a dominating sequence (35) with a finite sum (35), then (32) is true.

There are two ways to find a dominating sequence. One is be clever inequalities. The other is by calculating the *maximal function*

$$M_n = \sup_h |c_n(h)| \ .$$

Clearly M_n is a dominating sequence. If D_n is any other dominating sequence, then $D_n \geq M_n$ for all n. That makes the maximal sequence the sharpest possible dominating sequence. Using the maximal sequence serves two purposes. First, it gives you a concrete way to look for a dominating sequence, which may stop you from spending hours lost in a forest of inequalities. Second, you may calculate M_n and discover that the sum (35) is infinite. If that happens, you know there is no dominating sequence.

In the example, the maximal function is

$$M_n = \max_h h e^{-nh} \ .$$

You can find the max by calculus. Differentiate with respect to h and set the derivative to zero, and you find that the maximizing h is $h_* = n^{-1}$. Plug this in, and you find

$$M_n = h_* e^{-nh_*} = n^{-1} e^{-1} \ .$$

[9] If you look up the dominated convergence theorem on the web or in a book, you are likely to find something involving Lebesgue integration, or something involving abstract measure theory. The version of the dominated convergence theorem given here is easier than those general theorems but uses the same basic idea. It is an instance of the general measure theory theorem, if the measure space is $\mathbb{N}$ and the measure is counting measure. It isn't important or possibly even helpful for you to understand this footnote.

The sum is

$$\sum_1^\infty M_n = e^{-1} \sum_1^\infty n^{-1} = \infty \ .$$

This shows that there is no convergent dominating sequence, so the dominated convergence theorem does not apply.

The proof of the dominated convergence theorem uses basic tricks for limits and sums. You want to show that for any $\epsilon > 0$ there is a $\delta > 0$ so that

$$\left| \sum_1^\infty d_n - \sum_1^\infty c_n(h) \right| \leq \epsilon \ , \quad \text{if } |h| < \delta \ .$$

Suppose there is a dominating sequence with a finite sum. First, choose N so that

$$\sum_N^\infty D_n \leq \frac{\epsilon}{3} \ .$$

Because of the individual limits, we may choose $\delta_n > 0$ so that

$$|c_n(h) - d_n| \leq \frac{\epsilon}{3N} \ , \quad \text{if } |h| < \delta_n \ .$$

Now take δ to be the minimum of the N numbers δ_n. This is larger than zero because the min of finitely many positive numbers is a positive number. If $|h| < \delta$, then all the inequalities below are satisfied:

$$\begin{aligned} \left| \sum_1^\infty d_n - \sum_1^\infty c_n(h) \right| &\leq \sum_N^\infty |d_n| + \sum_N^\infty |c_n(h)| + \sum_1^N |d_n - c_n(h)| \\ &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + N \frac{\epsilon}{3N} \\ &= \epsilon \ . \end{aligned}$$

When I was a grad student, this kind of proof was called an "epsilon over three argument". The sums from N to infinity are *tail sums*. You bound the tail sums using the dominating sequence. You get convergence in the central part (the terms below N) using the individual limits. The numbers $\frac{\epsilon}{3}$ and $\frac{\epsilon}{3N}$ in arguments like this are found after lots of trial and error.

We apply this to the problem of differentiating an infinite sum term by term, as (10). We just need there to be a single dominating sequence for the difference quotients (33). One form of the intermediate value theorem is that there is a ξ_n between s and $s + h$ so that

$$\frac{b_n(s+h) - b_n(s)}{h} = b_n'(\xi_n)$$

If $h_0 > 0$ and $|h| \leq h_0$, then $|\xi_n - s| \leq h_0$. Define the derivative maximal function to be

$$M_n(s, h_0) = \sup_{|\xi - s| \leq h_0} |b_n'(\xi)| \ .$$

The dominated convergence theorem then leads to the differentiation theorem. If, for some $h_0 > 0$,

$$\sum_1^\infty M_n(s, h_0) < \infty \ ,$$

then the term by term differentiation formula (10) is true.

We apply this to the zeta derivative formula (4). This is supposed to be true for any $s > 1$. The derivative maximal function in this problem is (Drop minus signs because you take the absolute value. The max is at the left endpoint.):

$$M_n = \max_{|\xi - s| \le h_0} \log(n) n^{-\xi} = \log(n) n^{-(s-h_0)} \ .$$

If you choose h_0 so that $s - h_0 > 1$, then the M_n have a finite sum (homework exercise).

7 Infinite products

An infinite product is the limit of finite partial products, just as an infinite sum is the limit of finite partial sums. For a sequence of numbers a_n, the infinite product is

$$P = \prod_1^\infty b_n \ . \tag{36}$$

The finite partial products are

$$P_N = \prod_1^N b_n = b_1 a_2 \cdots b_N \ . \tag{37}$$

The infinite product converges if the limit of the partial products exists. In that case, the infinite product is

$$P = \lim_{N \to \infty} P_N \ . \tag{38}$$

If any of the factors b_n is equal to zero, then the infinite product is equal to zero. This is because $P_N = 0$ for $N \ge n$.

If the partial products converge, the limit may be zero or some non-zero number. If the limit exists and is not zero, it is necessary that

$$\lim_{n \to \infty} b_n = 1 \ . \tag{39}$$

Here is the "routine" proof. If the b_n do not converge to 1, there are infinitely many $n_k \to \infty$, and an $\epsilon > 0$ so that $|b_{n_k} - 1| \ge \epsilon$. If $P_N \to P \ne 0$, there is an N_0 so that if $N > N_0$, then (tried $\epsilon/2$ first, that didn't work)

$$|P_N - P| < \frac{\epsilon P}{4} \quad \text{if } N \ge N_0 \ . \tag{40}$$

Now choose $n_k > N_0$ and suppose $b_{n_k} > 1 + \epsilon$. By assumption,

$$P_{n_k - 1} > \left(1 - \frac{\epsilon}{4}\right) P \ .$$

But, if ϵ is small enough so that $\frac{3}{4}\epsilon - \frac{1}{4}\epsilon^2 > \frac{1}{2}\epsilon$,

$$P_{n_k} = b_{n_k} P_{n_k - 1} > (1 + \epsilon) \left(1 - \frac{\epsilon}{4}\right) P > \left(1 + \frac{\epsilon}{2}\right) P \ .$$

This contradicts (40). If you get lost in this simple argument (because I explained it badly), if b_n is not close to 1, then $P_n = b_n P_{n-1}$ is not close to P_{n-1}. Therefore, it's impossible for both numbers to be very close to P. This presumes that $P \neq 0$. You can multiply 0 by a_n not close to 1 and still get 0. For the rest of this section we make three assumptions: first, that $P_N \to P \neq 0$ as $N \to \infty$, second, that $b_n > 0$ for all n, third, that $b_n \to 0$ as $n \to \infty$. The third assumption almost implies the second, except for finitely many n. Those don't effect the arguments here if $b_n \neq 0$ there.

A simple and efficient way to understand infinite products, under the above conditions, is to transform them to infinite sums using the log. Since $b_n > 0$ for all n,

$$a_n = \log(b_n)$$

is well defined. The partial sums of the logs are related to the partial products

$$A_N = \sum_1^N a_n = \sum_1^N \log(b_n) = \log\left(\prod_1^N b_n\right) = \log(P_N) \ .$$

If

$$A = \lim_{N \to \infty} A_N = \sum_1^\infty a_n$$

exists, because $a \mapsto b = \exp(a)$ is continuous, then

$$P = e^A = \lim_{N \to \infty} e^{A_N} = \prod_1^\infty b_n$$

also exists. If the a_n sum converges absolutely, then the b_n product also converges absolutely. That means that the b_n is the same no matter how the terms are re-arranged.

There is a simple test for absolute convergence of an infinite product. The test for absolute convergence of the sum is (28). Absolute convergence is a property of the tails of the sequence, where a_n is close to 0 and b_n is close to 1. In that "regime",

$$b_n = e^{a_n} \approx 1 + a_n$$

and

$$a_n \approx b_n - 1 \ .$$

It is "easy" to use this approximate relationship to prove that (28) holds if and only if

$$\sum_1^\infty |b_n - 1| < \infty \; . \tag{41}$$

Early in this section we assumed that $b_n > 0$ for all n. That, "clearly", is irrelevant if the absolute convergence criterion (41) is satisfied. It is irrelevant, because convergence or absolute convergence is a property of the tail of the sum. If the absolute convergence criterion (41) is satisfied, than all but a finite number of the b_n are positive. If $a_n \neq 0$ for all n, and if the product converges absolutely, then the product is not zero. But consider Euler's product formula for $\sin(x)$,

$$\sin(x) = x \prod_1^\infty \left(1 - \frac{x^2}{\pi^2 n^2}\right) \; . \tag{42}$$

This formula is the basis for some homework. The product converges absolutely (exercise) and is equal to zero only if one of the factors is equal to zero. Those are the zeros of $\sin(x)$.

Differentiation forumlas for infinite sums may also be found using the log transform, if you also use the chain rule. The resulting formula is given in Theorem 6 below. Now look back at the derivation of the zeta derivative formula (5). You will see that we found this without using the log trick. You would, hopefully, get the same answer using the log trick. In fact, you would get it more quickly. If you try to justify (5), without knowing the log trick, it might be frustrating but you could do it. Sometime during this process you might discover the log trick see how easy it makes things.

8 The Euler product.

We can now give the proof that the Euler product (3) converges absolutely for $s > 1$. If $b_n = (1 - p_n^{-s})^{-1}$, then clearly $b_n \to 1$ as $n \to \infty$. A Taylor series calculation shows that $b_n \approx 1 + p_n^{-s}$, so

$$\sum_1^\infty |b_n(s) - 1| \sim \sum_1^\infty p_n^{-s} < \infty \; .$$

I write $\sim$ above, rather than $\approx$ because the right side may not be a quantitative approximation to the left side, but they are "like" each other as far as convergence or divergence are concerned. We can add some "slop" and get the rigorous upper bound

$$\left|\left(1 - p_n^{-s}\right)^{-1}\right| < 2p_n^{-s} \; ,$$

which holds for all but finitely many p_n for any range of s of the form $|s - s_0| < \delta$, as long as $s - \delta > 1$. This is done using the mean value theorem, for example.

At last, we come to the Euler formula (3), which is the analytic number theory of this set of notes. We already saw the geometric series formula, which we write in the form

$$\left(1 - p_n^{-s}\right)^{-1} = \sum_{r=0}^{\infty} p_n^{-rs} \; .$$

We substitute this into the Euler product on the right side of (3) and multiply it out

$$\begin{aligned} \prod_{n=1}^{\infty} \left(1 - p_n^{-s}\right)^{-1} &= \prod_{n=1}^{\infty} \left(\sum_{r=0}^{\infty} p_n^{-rs} \right) \\ &= \sum \left(\prod_k p_k^{r_k} \right)^{-s} \; . \end{aligned}$$

The sum in the second line is over all sequences $r_2, r_3, \ldots$, so that $r_k = 1$ for all but finitely many k. Because of unique prime factorization of integers, each positive integer may be expressed in one and only one way as a product like this. Therefore, the right side is equal to zeta function sum (2).

The formal proof has to show somehow that

$$\lim_{N \to \infty} P_N = \lim_{N \to \infty} A_N \; ,$$

where

$$P_N = \prod_1^N \left(1 - p_n^{-s}\right)^{-1} \; ,$$

and

$$A_N = \sum_1^N n^{-s} \; .$$

We know both limits exist, because the sum and the product both converge absolutely for $s > 1$. But this still isn't comparing finite sum to finite sum, because $\left(1 - p_n^{-s}\right)^{-1}$ is an infinite sum. To compare finite sums to finite sums, we replace the infinite geometric series with a finite approximation

$$\left(1 - p_n^{-s}\right)^{-1} \approx \sum_{r=0}^{M} p_n^{-rs} \; .$$

The right side converges to the left side as $M \to \infty$. Therefore, for any $s > 1$ and any N and any $\epsilon > 0$, there is an M so that (it's an "epsilon over three" argument, but with four parts instead of three)

$$\left| \prod_{n \le N} \left(\sum_{r=0}^{M} p_n^{-rs} \right) - \prod_{n \le N} \left(1 - p_n^{-s}\right)^{-1} \right| \le \frac{\epsilon}{4} \; .$$

Define

$$P_{N,M} = \prod_{n \le N} \left(\sum_{r=0}^{M} p_n^{-rs} \right) . \tag{43}$$

We will choose N, M, and L large enough so that

$$|P_{N,M} - A_L| \le \frac{\epsilon}{4} \qquad \text{(i)}$$
$$|P_{N,M} - P_N| \le \frac{\epsilon}{4} \qquad \text{(ii)}$$
$$|P - P_N| \le \frac{\epsilon}{4} \qquad \text{(iii)}$$
$$|A - A_L| \le \frac{\epsilon}{4} \qquad \text{(iv)}$$

This will prove that $|P - A| \le \epsilon$ for any $\epsilon > 0$, which proves that $P = A$. The last two are the convergence of the zeta sum and product. The second we just did. The first is simple, but takes some notation to explain.

The set of positive numbers is denoted by $\mathbb{N}$, which stands for "natural" numbers. For a given N and M, let $H \subset \mathbb{N}$ be the set of numbers with

$$n = \prod p_k^{r_k} \tag{44}$$

so that $r_k = 0$ for $k > N$, and $r_k \le M$ for all k. That is, $n \in H$ if its prime factors are no larger than P_N and the multiplicities are no larger than M. The informal argument above for the Euler product formula is a rigorous proof of

$$P_{N,M} = \sum_{n \in H} n^{-s} .$$

Clearly, for any L_1, there are sufficiently large N and M so $n \in H$ for every $n \le L_1$. Also, there is an L_2 so that $n \le L_2$ for every $n \in H$. Therefore $A_{L_1} \le P_{N,M} \le L_{L_2}$. Since A_L has a limit as $L \to \infty$ (and therefore is a Cauchy sequence), it is possible to make (i) above hold. This proves the Euler product formula (3). The differentiation formula is a consequence of the product formula and the differentiation theorem Theorem 6 below.

In the end, we come back to the beginning and the divergence of the prime sequence (1). You can prove this by contradiction. If the prime series would converge, then the Euler product

$$\prod_{1}^{\infty} \left(1 - p_n^{-1}\right)$$

would converge to a finite number. But we know that

$$\lim_{N \to \infty} \prod_{1}^{N} \left(1 - p_n^{-1}\right) = \infty .$$

Another way to say this, now that we have the dominated convergence theorem, is that if the sum (1) were finite, it would be a dominating sequence for the zeta sum with $s \geq 1$, because

$$p^{-s} \leq p^{-1} \quad \text{if} \quad s \geq 1 \, .$$

This would make the zeta Euler product converge uniformly in the interval $[1, s_0]$ with $s_0 > 1$.

9 Summary of theorems

Theorem 1.(*Greatest common divisor*). For each pair of positive integers, a and b, there is a positive integer $c = (a, b)$ so that if $d \,|\, a$ and $d \,|\, b$, then $d \,|\, c$. There are integers i and j, not necessarily positive, so that $c = ia + jb$.

Theorem 2. (*Unique prime factorization of integers*). For each integer $n \geq 2$ there is a unique set of primes $p_1, \ldots, p_k$ and exponents $r_1, \ldots, r_k$ so that $n = p_1^{r_1} \cdot p_2^{r_2} \cdots$.

Theorem 3.(*Zeta convergence*). The sum (2) converges for $s > 1$ and satisfies the estimate

$$\zeta(s) = \frac{1}{s-1} + O(1) \quad \text{as } s \downarrow 1 \, .$$

Theorem 4.(*Dominated convergence theorem*). If $a_n(s)$ is a sequence of continuous functions, and if there is a $\delta > 0$ and a sequence D_n with $|a_n(s)| \leq D_n$ for all n and $|s - s_0| < \delta$, and

$$\sum_1^\infty D_n < \infty \, ,$$

then

$$\lim_{s \to s_0} \sum_1^\infty a_n(s) = \sum_1^\infty a_n(s_0) \, .$$

Theorem 5.(*Differentiating a sum*). Let $a_n(s)$ be a sequence of continuous functions that satisfies the conditions of Theorem 4. Suppose the functions a_n are differentiable and

$$|a_n'(s)| \leq E_n$$

for all $|s - s_0| < \delta$, and

$$\sum_1^\infty E_n < \infty \, .$$

Then, if $|s - s_0| < \delta$,

$$\frac{d}{ds} \left(\sum_1^\infty a_n(s) \right) = \sum_1^\infty a_n'(s_0) \, .$$

Theorem 6.(*Infinite products*). Let $b_n(s)$ be a sequence of differentiable functions. If there is a sequence D_n so that if $|s - s_0| < \delta$ then

$$|b_n(s) - 1| < D_n$$

and

$$\sum_1^\infty D_n < \infty \ ,$$

then

$$\prod_1^\infty b_n(s) = \lim_{N\to\infty} \left(\prod_1^N b_n(s) \right)$$

exists for every $|s - s_0| < \delta$. Moreover

$$\lim_{s\to s_0} \prod_1^\infty b_n(s) = \prod_1^\infty b_n(s_0) \ .$$

Suppose that, for all n and $|s - s_0| < \delta$,

$$\left| \frac{b_n'(s)}{b_n(s)} \right| \le E_n \ ,$$

and

$$\sum_1^\infty E_n < \infty \ .$$

Then, if $|s - s_0| < \delta$, the product is differentiable and

$$\frac{d}{ds} \left(\prod_1^\infty b_n(s) \right) = \left(\sum_1^\infty \frac{b_n'(s)}{b_n(s)} \right) \left(\prod_1^\infty b_n(s) \right) \ .$$

10 Notes on references

Both Jameson and Apostol have good introductory material to complement Section 1. The "big Oh" stuff in Section 2 you also can find in Jameson or Apostol. The zeta function inequalities are also there. I don't want to dwell on this as much as they do in order to get to the number theory material sooner. Elkies takes a different route to this material and assumes more mathematical sophistication, so he is helpful in motivation and less helpful in mathematical background. The number theoretic functions $\pi(x)$, $\phi(x)$ and $\psi(x)$ are defined in all these sources, and also in Hardy and Wright. Jameson and Apostol present the summation-by-parts trick (called *Abel summation*) that relates these. I chose postpone Abel summation to make Section 3 less technical. We will do it later. I recommend Hardy and Wright as the best source for prime factorization and the euclidean algorithm, and anything to do with basic number theory.

The analytical material in Section 5 and Section 6 may be found in Apostol or Jameson, but I don't know a simple reference for the version of the dominated convergence theorem given here. That theorem is usually presented in the context of measure theory. I recommend looking at Elkies for the Euler product material. It's expertly motivated and given in fewer words than these notes or the other references.

11 Exercises.

1. Show that
$$\int_x^{2x} \frac{1}{\log(y)}\, dy = \frac{x}{\log(x)} + o\left(\frac{x}{\log(x)}\right) .$$

2. A "hundred bit number" is a number whose expression base 2 has 100 bits, with a 1 in the 2^{99} bit. Such a number is in the range $2^{99} \leq n < 2^{100}$. Use $\mathrm{li}(x)$ and its large x approximation to give a quantitative estimate of the probability that a random number in this range is prime. Hint: The hard part of this problem may be converging from log base 2 to log base e.

3. Show that the sum
$$-\zeta'(s) = \sum_1^\infty \log(n) n^{-s}$$
converges absolutely if $s > 1$. Do this by comparing to the integral
$$\int_2^\infty \log(x) x^{-s}\, dx .$$
Use integral comparisons to show that
$$\zeta'(s) = \frac{-1}{(s-1)^2} + O\left(\frac{1}{s-1}\right) \quad , \quad \text{as} \quad s \downarrow 1 .$$
Hint: integrate by parts to remove the log from the integral. Make the following calculation complete and rigorous:
$$\frac{\zeta'(s)}{\zeta(s)} = \frac{(s-1)^{-2} + O\left((s-1)^{-1}\right)}{(s-1)^{-1} + O(1)} = (s-1)^{-1} + O(1) .$$

4. You can guess the sizes sums over primes informally by imagining that the density of primes is $\rho(x) \approx 1/\log(x)$. That means that
$$\sum_{p \leq x} f(p) \approx \int^x f(y)\rho(y)\, dy = \int^x \frac{f(y)}{\log(y)}\, dy .$$
Use this to "guess" the large x behavior of
$$\beta(x) = \sum_{p \leq x} \frac{1}{p} .$$

(*Notes*: (1) The wording of this exercise is bad. There should never be an algorithm for guessing. (2) The notation $\beta(x)$ for this sum is not standard. The notation $\pi(x)$, $\theta(x)$ and $\psi(x)$ for other prime sums is standard.)

5. The *Gaussian integers* is the set of complex numbers of the form $a + bi$ where a and b are integers. The set of ordinary integers is called $\mathbb{Z}$ (for the German word "Zahlen", which means "numbers"). The set of Gaussian integers is written $\mathbb{Z}[i]$, which is the integers with i "adjoined". This exercise goes through the proof that any Gaussian integer has a unique factorization as a product of "Gaussian primes".

   (a) Show that if $x = a + bi$ and $y = c + di$ are Gaussian integers, then $-x$ and xy are Gaussian integers. A set with these properties (and commutative and associative and distributive) is called a *ring*. The Gaussian integers are a ring. You don't have to check distributivity, associativity and commutativity because Gaussian integers are complex numbers and we're using the complex number operations. *Domain* is used to mean "ring" in this context, sometimes.

   (b) Let $x = a + bi$ and $y = c + di$ be Gaussian integers. Show that there is a Gaussian integer $z = e + fi$ so that $y = zx + w$, with w being in the square with corners 0, x, ix, and $x + ix$, but not on the outer edge between x and $x + ix$ or the edge between ix and $x + ix$. *Hint:* you can move y toward the square in the $x = a + ib$ direction by subtracting x. You can move y toward the box in the perpendicular ix direction by subtracting ix.

   (c) Show that there are $n = a^2 + b^2$ possible values of w in part (b), and that all of them are possible.

   (d) $\mathbb{Z}[i]$ has a set of four *units*: $U = \{1, i, -1, -i\}$. Show that if $x \in \mathbb{Z}[i]$ and there is a $y \in \mathbb{Z}[i]$ with $xy = 1$ (i.e., if $x^{-1} \in \mathbb{Z}[i]$), then $x \in U$. That is, if x is invertible in $\mathbb{Z}[i]$ then x is a unit. Conversely, if x is a unit then x is invertible.

   (e) A $p \in \mathbb{Z}[i]$ is a *Gaussian prime*,[10] if it has the property that if $p = xy$ (x and y being Gaussian integers), then either $x \in U$ or $y \in U$. Show that if $x = p_1 p_2$ and $x = q_1 q_2$ with p_1, p_2, q_1, and q_2 being Gaussian primes, then either $p_1 = uq_1$ or $p_1 = uq_2$. *Hint:* Use the Euclidean algorithm. A domain (ring) with unique prime factorization is a UFD. A ring where the Euclidean algorithm works is a Euclidean domain. A Euclidean domain is a UFD (don't prove this). The Gaussian integers are a Euclidean domain and therefore a UFD (you just proved this).

   (f) Show that 3 is a Gaussian prime but 2 is not.

[10] The units in the ordinary integers $\mathbb{Z}$ are $U = \{1, -1\}$. This definition would make both 7 and -7 primes in $\mathbb{Z}$. We could say $p \in \mathbb{Z}[i]$ is prime if it has no non-trivial factors and if $p = a + bi$ with $a > 0$ and $b \geq 0$. That would be closer to the usual definition for $\mathbb{Z}$, but it is less traditional.

(g) Show that any Gaussian integer has a unique factorization, modulo units, into a product of Gaussian primes.

6. There is a zeta function and Euler product formula for the Gaussian integers. In this exercise, we write $\sum'$ to be the sum modulo units. This means that we choose one of the four Gaussian integers $y = ux$ to include in the sum. The zeta function is[11]

$$\zeta_K(s) = \sum{}' |x|^{-s} \ . \tag{45}$$

This may also be written

$$\zeta_K(s) = \sum_{a=1}^{\infty} \sum_{b=0}^{\infty} \left|a^2 + b^2\right|^{-s/2} \ .$$

The Euler product is

$$\zeta_K(s) = \prod_p{}' \left(\frac{1}{1 - |p|^{-s}} \right) \ . \tag{46}$$

The product is over all Gaussian primes $p = a + bi$ with $a > 0$ and $b \geq 0$.

(a) Show that the sum (45) converges if $s > 2$.

(b) Show that the product (46) converges if $s > 2$.

(c) Show that the infinite product is equal to the infinite sum if $s > 2$. Show that $\zeta_K(s) \to \infty$ as $s \downarrow 2$.

(d) Show that

$$\sum{}' \frac{1}{|p|^2} = \infty$$

where the sum is over all Gaussian primes with $a > 0$ and $b \geq 0$.

7. There is a convergence theorem for series called the *monotone convergence theorem* related to the dominated convergence theorem. The theorem is about a family of sequences $c_{n,x}$. The monoticity hypothesis that gives the theorem its name is that $c_{n,y} \geq c_{n,x} \geq 0$ for all n and $y > x$. We also suppose that

$$\lim_{x \to \infty} c_{n,x} = d_n < \infty$$

for each n. The theorem states that under these conditions

$$\lim_{x \to \infty} \sum_{n=1}^{\infty} c_{n,x} = \sum_{1}^{\infty} d_n \ .$$

There are two possibilities, if $\sum d_n = \infty$ then $\sum c_{n,x} \to \infty$ as $x \to \infty$. The other possibility is that $\sum d_n < \infty$, in which case the sum $\sum c_{n,x}$ converges to that.

[11] The subscript K tells us that this is not the Riemann zeta function, which is *the* zeta function. The K means something.

(a) Prove this theorem in both parts. You may look at the proof of the dominated convergence theorem in the notes to see how to prove this kind of thing.

(b) Consider the partial Euler products

$$\zeta_x(s) = \prod_{p \le x} \left(1 - p^{-s}\right)^{-1}$$

Show that

$$\zeta_x(s) = \sum_1^\infty a_n(x) n^{-s} \ ,$$

where $a_n(x) = 1$ if all the prime factors of n have $p \le x$, and $a_n(x) = 0$ if n has a prime factor with $p > x$.

(c) Use the monotone convergence theorem to give a different proof that

$$\lim_{x \to \infty} \zeta_x(s) = \zeta(s) = \sum_1^\infty n^{-s} \ .$$

8. If a function $f(x)$ can be represented as a Taylor series, it is tempting to think of f as being a fancy version of a polynomial. If $f(x)$ is a polynomial of degree N, and the roots of f are $x_1, \ldots, x_N$, then[12]

$$f(x) = C \prod_1^N \left(1 - \frac{x}{x_n}\right) \ . \tag{47}$$

If $f(x)$ is an odd polynomial ($f(-x) = -f(x)$) of order $2N+1$ and positive roots $x_1, \ldots, x_N$, then

$$f(x) = Cx \prod_1^N \left(1 - \frac{x^2}{x_n^2}\right) \ . \tag{48}$$

Since $\sin(x)$ is a fancy polynomial in the sense of having a Taylor series representation, maybe it also has a product representation:

$$\sin(x) = Cx \prod_n^\infty \left(1 - \frac{x^2}{\pi^2 n^2}\right) \ . \tag{49}$$

To talk about the product without proving the formula is correct, define

$$f(x) = x \prod_n^\infty \left(1 - \frac{x^2}{\pi^2 n^2}\right) \ . \tag{50}$$

This exercise does not include a proof that $f(x) = \sin(x)$, though it is.

[12] Mathematicians use C to represent a "generic" constant whose value need not be given for the purpose at hand. Moreover, the value of C may be different in different places. For example, if $f(x) \le Cx$, then $f(x)^2 \le Cx^2$. If the first C is 10, then the second C is 100.

(a) Prove that a representation of the form (48) follows from the representation (47) if f is an odd polynomial. *Hint 1:* The main difficulty in doing this is the conflict of notation. N and x_N mean different things in (47) and (48). You can explain the reasoning more clearly if you remove the conflict, for example, by defining $M = 2N + 1$ to be the degree of an odd f, and $y_k = 0$, or $y_k = \pm x_n$ (with the relationship between n and k depending on the sign). *Hint 2:* This is not exactly true, as you will discover when you try to prove it. You need the hypothesis that $f'(0) \neq 0$.

(b) Prove that the infinite product (50) converges for any x.

(c) Show that $f(x) = x + O(x^3)$ as $x \to 0$. Use this to identify C in (49), assuming (49) is true.

(d) Show that $f(x) = x + f_3 x^3 + O(x^5)$ as $x \to 0$, with the explicit formula

$$f_3 = \frac{1}{\pi^2} \sum_1^\infty \frac{1}{n^2} \ . \tag{51}$$

Hint: You multiply out the infinite product (50) to get a formal derivation of (51). The rigorous proof with the $O(x^5)$ error estimate is less interesting "technique".

(e) Assuming (49) is true, show that

$$\sum_1^\infty \frac{1}{n^2} = \frac{\pi^2}{6} \ . \tag{52}$$

The young Leonard Euler gained recognition as a mathematical talent by discovering this formula. We will give more proofs of it later.

9. Show that the following derivative exists for $s > 1$ and find an infinite sum representation for it as a Dirichlet series involving the square of the von Mangoldt $\Lambda(n)$.

$$\frac{d^2}{ds^2} \log(\zeta(s))$$

Prove that your differentiation formula is correct using the dominated convergence theorem or some other argument. (Complex analysis gives a simple way to know that all the derivatives of functions like this exist, but you can prove it without complex analysis.)

10. Suppose the prime factorization is written

$$n = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k} \ ,$$

where only powers $r_j \geq 1$ are written. For example, $75 = 3 \cdot 5^2$, so $p_1 = 3$, $p_2 = 5$, $r_1 = 1$ and $r_2 = 2$. The *Möbius* function $\mu(n)$ is defined in terms of the prime factorization by

$$\mu(n) = \begin{cases} (-1)^k & \text{if } r_j = 1 \text{ for all } j \\ 0 & \text{otherwise.} \end{cases}$$

For example, $\mu(1) = 1$ (because $k = 0$), $\mu(2) = \mu(3) = 1$, $\mu(4) = 0$, $\mu(6) = 1$, etc. A number that satisfies the first condition (no prime powers $r_j > 1$ is called *square free*. An equivalent definition of square free is that d^2 does not divide n for any integer d.

(a) Show that if $s > 1$, then

$$\frac{1}{\zeta(s)} = \sum_1^\infty \mu(n)\, n^{-s} \ .$$

(*Hint*: Use the Euler product representation of $\zeta(s)$ and prove that its OK to multiply it out. Part of the problem (not the hard part) is to show that the sum converges absolutely for $s > 1$.)

(b) Suppose functions $f(s)$ and $g(s)$ are given as *Dirichlet series*, which means that

$$f(s) = \sum_1^\infty a_n n^{-s} \quad , \quad g(s) = \sum_1^\infty b_n n^{-s} \ .$$

Show formally (i.e. without worrying about convergence) that $h(s) = f(s)g(s)$ also is a Dirichlet series with

$$h(s) = \sum_1^\infty c_n n^{-s} \quad , \quad c_n = \sum_{jk=n} a_j b_k \ .$$

The second formula (which is the point of this part) is often written

$$c_n = \sum_{d|n} a_d b_{\frac{n}{d}} \ .$$

(c) Use formal manipulations like this on the product

$$\zeta'(s) = \zeta(s)\, \frac{\zeta'(s)}{\zeta(s)} = \left(\sum_1^\infty n^{-s} \right) \left(-\sum_1^\infty \Lambda(n) n^{-s} \right) \ .$$

The result should be the formula

$$\log(n) = \sum_{d|n} \Lambda(d) \ . \tag{53}$$

(d) Once you have found a formula in some possibly illegitimate way, you can try to give it a direct legitimate proof. Find a direct proof of (53). *Hint*: Calculate $\log(p_1^{r_1} \cdots p_k^{r_k})$.

(e) Show that

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1. \end{cases} \tag{54}$$

Hint: Use parts (a) and (b) and the formula

$$\zeta(s) \frac{1}{\zeta(s)} = 1 \ .$$

(f) (*Discussion, nothing to hand in for this part*) The formula (54) is a special case of the *Möbius inversion formula*. It is a formula about finite sums that has nothing to do with convergence of Dirichlet series. A rigorous proof of (54) using Diriclhet series should make you uncomfortable, since convergence should have nothing to do with finite algebraic sums. It would seem more fitting to give a direct proof that doesn't use Dirichlet series, which would be called a *combinatorial* proof. In this case, it isn't too hard to do that. There are many such formulas that are discovered manipulating infinite series but are then given combinatorial proofs.

11. The *Gamma function* is defined by

$$\Gamma(s) = \int_0^\infty x^{s-1} e^{-x}\, dx \ . \tag{55}$$

This exercise and the next give two ways the Gamma function is used in analytic number theory.

(a) Show that the integral (55) converges absolutely if $s > 0$. Show that $\Gamma(s)$ is a positive differentiable increasing function of s in this range.

(b) Show that if $s > 0$, then[13]

$$\int_0^\infty x^{s-1} e^{-nx}\, dx = n^{-s}\Gamma(s) \ . \tag{56}$$

(c) Show that if $s > 1$, then

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^\infty \frac{x^{s-1} e^{-x}}{1 - e^{-x}}\, dx \ . \tag{57}$$

Hint: The formula (56) gives an integral formula for n^{-s}. Interchange the order of summation and integration in the sum formula for $\zeta(s)$, then sum the resulting geometric series. Do the analysis to show that the integral converges for $s > 1$ and to justify doing the sum inside the integral.

(d) The denominator in the integral "looks like" $\frac{1}{x}$ and the numerator looks like x^{s-1} (because e^{-x} looks like 1) for x near zero. Let $R(x)$ be the error in this small x approximation:

$$R(x) = \frac{e^{-x}}{1 - e^{-x}} - \frac{1}{x} \ .$$

Show that there is a C so that

$$|R(x)| \leq C \quad , \quad \text{for } \ 0 < x < 1 \ . \tag{58}$$

[13]The idea in this exercise is due to Riemann. He used the slightly different function: $\Pi(s) = \int x^s e^{-x} dx = \Gamma(s+1)$. The Gamma function is more natural in that there are s "factors" of x in the integral (55), the last one being the dx term. Riemann's version of the formula had $n^{s+1}\Pi(s)$.

(e) Use these facts to show that

$$\left|\zeta(s) - \frac{1}{s-1}\right| \leq C \text{ for all } s > 1 .$$

Hint: You can start by proving the result for the integral without the 1Γ factor. For this, you may formulate and prove a lemma that says something like: if $g(s) - C/(s-1) = O(1)$ as $s \downarrow 1$ and if $h(s)$ is continuous at $s = 1$ with $h(1) \neq 0$, then $h(s)g(s) - C/(s-1) = O(1)$ (different C). It may be convenient to split the integral into two parts, $\int_0^1$ and $\int_1^\infty$. You may need one argument for small s ($1 < s \leq 1+\epsilon$), and another (much simpler) argument for $s > 1 + \epsilon$.

This proof may be easier or harder than the one given in the notes. It has more steps, but the steps may or may not seem more natural.