Two classic theorems from number theory: The Prime Number Theorem and Dirichlet's Theorem

Senior Exercise in Mathematics

Lee Kennard

15 November, 2006

Contents

0 Notes and Notation

Proving the Prime Number Theorem (PNT) might have been sufficient for a senior exercise; it is a beautiful fact which is nontrivial to prove. I had thought about Dirichlet's theorem on primes in arithmetic progressions as well, but I thought that both would be too much, so my real plan was to do one or the other. Once I finished the proof of the PNT, however, I had some momentum, and I was having fun. Sure, I had learned about Abel's Lemma, the Dirichlet test for convergence, and the Leibniz integral rule, but all of these seemed to me disconnected tools, used together here only because they happened to be necessary for the proof. I longed for unity or, as least, something which I could add to make this exercise a whole. Since the Dirichlet theorem interested me, here is what I came up with. I asked the questions "Do there exist infinitely many primes" (in the set of natural numbers $\mathbb{N}$ or in arithmetic progressions in $\mathbb{N}$) and, if so, how are these primes (in $\mathbb{N}$ or in these arithmetic progressions) distributed – that is, how dense are they in $\mathbb{N}$? Euclid and, later, Euler answered the first half of the first question; Dirichlet answered the second half; the Prime Number Theorem (PNT) answered the first half of the second question; and the generalized PNT answered the second half.

The answers to these four questions form the outline of this exercise. Along the way, we will do some analytic number theory: we will define the Riemann ζ-function and the Dirichlet L-functions; we will do some complex analysis: we will use the notion of compactness several times (thank you Professor Schumacher), continuously deform contours of integration (thank you Professor Holdener), and "analytically continue" functions; we will do some abstract algebra: we will use Lagrange's theorem multiple times[1] (thank you Professor Aydin), prove that $\mathbb{Z}_{p^\alpha}$ is cyclic for odd primes p, and define Dirichlet characters on $\mathbb{Z}_q$ for all positive integers q; and, at one point, we will even do some probability theory (thank you Professor Hartlaub).[2]

The following notation will be used throughout the paper:

$\mathbb{N}$	the set of positive integers $1, 2, 3, \ldots$
$\mathbb{Z}$	the set of integers $\ldots, -1, 0, 1, \ldots$
$\mathbb{R}$	the set of real numbers
$\mathbb{C}$	the set of complex numbers
$\mathbb{Z}_q$	the ring $\mathbb{Z}/q\mathbb{Z}$
$\mathbb{Z}_q^*$	the multiplicative group of invertible elements of $\mathbb{Z}_q$
$\Re(s)$	denotes the real part of $s \in \mathbb{C}$
$\Im(s)$	denotes the imaginary part of $s \in \mathbb{C}$
$\phi(n)$	the Euler totient function
$\zeta(s)$	the Riemann zeta function (defined below)

[1] In the words of John B. Fraleigh, author of *A first course in abstract algebra*: "Never underestimate the power of a theorem which counts something!"

[2] I also want to thank Professor Milnikel and Professor Brown; a large portion of what I understand is understood because I understand linear algebra.

1 Introduction

Prime numbers are the building blocks of the natural numbers, just as atoms are the building blocks of molecules. Though we cannot understand everything about the natural numbers via meditations on the primes, we can surely learn a lot!

A prime natural number is an irreducible number. That is, the only divisors of a prime are 1 and itself.[3] Also true is the fact that if a prime number p dividing a product ab of two integers a and b, then either p divides a or p divides b. This fact can be used to show that the factorization of natural numbers into primes is unique. Since every natural number can be factored into primes (using the fact that the quotient obtained by divided a number by a prime is strictly less than the number), we conclude that, for all natural numbers $n > 1$, there exists a unique set of primes $p_1 < p_2 < \cdots < p_r$ and positive integers $a_1, \ldots, a_r$ such that

$$n = \prod_{k=1}^{r} p_k^{a_k}.$$

Here the nature of primes as building blocks is clear.

In this exercise, we shall take interest in the distribution of primes in the natural numbers. We ask: How many primes are there? About how many primes are there less than a given x? How many primes are there of the form $a+bn$ where a and b are integers? And finally: About how many primes of the form $a + bn$ are there less than a given x? These four questions form the skeleton of this exercise.

[3] We do not count 1 as a prime for a few reasons, the most important of which is to preserve unique factorization.

2 Primes in the odd integers

Since we do not count 1 as a prime, and since we do not count 0 as a natural number, we might ask if 2 is prime. This is obvious, since 1 and 2 divide 2, and there are no other natural numbers less than 2 which could divide it. Therefore 2 is the smallest prime. In fact, it is the only even prime. For if n is an even natural number besides two, then $n = 2m$ for some natural number $m > 1$, and this implies that (at least) three natural numbers – namely, 1, 2, and m – all divide n. Two is the smallest prime; two is the only even prime; and many ways, two is the *oddest of all primes*. Because two is so special, we choose to ignore it throughout most of this paper. So we start by studying the distribution of primes in the odd integers.

2.1 The infinitude of primes

From the definition, we see also that 3, 5, 7, 11, and 13 are all prime and that (by skipping to triple-digit numbers) 101, 103, and 107 are all prime. If we work a little harder, we would find that 1009, 1013, 1019, and 1021 are prime. In fact, no matter what power of 10 we go up to, we will always be able to find another prime. The reason:

There exist infinitely many primes.

One incredible proof of this was given by Euclid. I might say that no senior exercise on the distribution of primes would be complete if it lacked this proof. It makes the non-mathematician smile; it makes us squeal.

The proof goes as follows: Suppose there are only finitely many primes. Denote them by $p_1, p_2, \ldots, p_n$. Define the natural number N by

$$N = \prod_{k=1}^{n} p_k + 1.$$

As we mentioned (but did not prove) above, we can write

$$N = \prod_{k=1}^{n} p_k^{a_k},$$

where the a_k are nonnegative integers.[4] But then if $a_k > 0$ for some k, then p_k divides both N and $N - 1$, which means that $p_k = 1$, a contradiction. If $a_k = 0$ for all k, then $N = 1$, contradicting the fact that $N \geq p_1 + 1 > 1$. Therefore, there must exist infinitely many primes.

Like all great mathematical theorems, this one has multiple proofs. A second proof is in order for two reasons. First, it will prove a stronger fact from which the infinitude of primes follows immediately. Second, it will introduce us to the Riemann zeta function, a fundamental function in analytic number theory. But before we get to this, we prove an important result which we will use throughout the paper:

Theorem 1. *Let $\chi : \mathbb{N} \to \mathbb{C}$ be a completely multiplicative[5] bounded function, and (formally) define $L(s, \chi)$ for all $s \in \mathbb{C}$ by*

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

[4] Before we specified that each $a_k > 0$, but we allow some a_k to be 0 here so we can more easily and clearly write N as a product of the primes.

[5] That is, $\chi(mn) = \chi(m)\chi(n)$ for all $m, n \in \mathbb{N}$ (without restriction).

Then $L(s,\chi)$ converges absolutely for all $\Re(s) > 1$ and has the following product representation which converges absolutely for all $\Re(s) > 1$:

$$L(s,\chi) = \prod_p \frac{1}{1-\chi(p)p^{-s}}$$

where the product is over all primes p.

Proof. For a prime p, consider the series

$$\sum_{k=0}^{\infty} \left(\frac{\chi(p)}{p^s}\right)^k = \sum_{k=0}^{\infty} \frac{\chi(p^k)}{p^{ks}}$$

for a variable $s \in \mathbb{C}$. Since χ is bounded, there exists a real number M such that $|\chi(n)| \leq M$ for all n. Therefore,

$$\sum_{k=0}^{\infty} \left|\frac{\chi\left(p^k\right)}{p^{ks}}\right| \leq M \sum_{k=0}^{\infty} \frac{1}{|p^{ks}|} = M \sum_{k=0}^{\infty} \left(\frac{1}{p^{\Re(s)}}\right)^k = \frac{M}{1-p^{-\Re(s)}} < \infty$$

since $p^{-\Re(s)} < 1$, so this series absolutely converges for all prime p. If we enumerate the primes $p_1 = 2, p_2 = 3, p_3 = 5, \ldots$, then for all $N \in \mathbb{N}$,

$$\prod_{j=1}^{N} \frac{1}{1-\chi(p_j)p_j^{-s}} = \prod_{j=1}^{N} \left(\sum_{k_j=0}^{\infty} \left(\frac{\chi(p_j)}{p_j^s}\right)^{k_j}\right) = \prod_{j=1}^{N} \left(\sum_{k_j=0}^{\infty} \frac{\chi\left(p_j^{k_j}\right)}{p_j^{k_j s}}\right).$$

Since we can multiply convergent series in the natural way (see Appendix A.4.2), this product is equal to

$$\sum_{k_1=1}^{\infty} \cdots \sum_{k_N=1}^{\infty} \left(\prod_{j=1}^{N} \frac{\chi\left(p_j^{k_j}\right)}{p_j^{k_j s}}\right) = \sum_{k_1=1}^{\infty} \cdots \sum_{k_N=1}^{\infty} \frac{\chi\left(p_1^{k_1} p_2^{k_2} \cdots p_N^{k_N}\right)}{\left(p_1^{k_1} p_2^{k_2} \cdots p_N^{k_N}\right)^s} = \sum_{n \in S_N} \frac{\chi(n)}{n^s},$$

where S_N is the set of positive integers whose greatest prime divisor is less than or equal to N. (That is, S_N is the set of $n \in \mathbb{N}$ such that $p \leq N$ for all prime p which divide n.) The justification for this last step is what makes this proof beautiful. We first note that the multiplication of absolutely convergent series yields another absolutely convergent series. So rearranging the terms is justified.

To show that the sums are equal, we show that each term in each sum shows up in the other, and that no term shows up twice in either sum. Starting on the left side of the equality, to each $(k_1, \ldots, k_N)$-term corresponds a unique $n \in S_N$, namely, $n = p_1^{k_1} \cdots p_N^{k_N}$, such that the $(k_1, \ldots, k_N)$-term on the left equals the n-term on the right. On the other hand, given the n-th term on the right, since $n \in S_N$, n has a factorization into primes less than or equal to p_N. Since this factorization is *unique*, there exists a unique N-tuple $(k_1, \ldots, k_N)$ such that the n-term on the right equals the $(k_1, \ldots, k_N)$-term on the left.

Since the terms in each series show up exactly once in the other series, the series are equal. Finally, since every $n \in \mathbb{N}$ can be factored into primes,

$$\lim_{N \to \infty} \sum_{n \in S_N} \frac{\chi(n)}{n^s} = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = L(s,\chi),$$

and since

$$\sum_{n=0}^{\infty}\left|\frac{\chi(n)}{n^s}\right| \leq M\sum_{n=0}^{\infty}\frac{1}{n^{\Re(s)}} < \infty,$$

the series $L(s,\chi)$ converges (absolutely) for all $\Re(s) > 1$, giving us

$$L(s,\chi) = \lim_{N\to\infty}\sum_{n\in S_N}\frac{\chi(n)}{n^s} = \lim_{N\to\infty}\prod_{j=1}^{N}\frac{1}{1-\chi(p_j)p_j^{-s}} = \prod_p \frac{1}{1-\chi(p)p^{-s}},$$

as claimed.

Finally, to show that the product converges absolutely, it will be sufficient (see Appendix A.4.1) to show that the series

$$\sum_p\left(\frac{1}{1-\chi(p)p^{-s}} - 1\right)$$

converges absolutely. But this is straightforward. We have that $|\chi(n)| \leq M$ for all $n \in \mathbb{N}$. For a technical reason, redefine M if necessary such that $|\chi(n)| \leq M$ and $p^{\Re(s)} \neq M$ for all primes p. (This is always possible because the set of all $p^{\Re(s)}$ is a discrete subset of $\mathbb{R}$.) Then

$$\sum_p\left|\frac{1}{1-\chi(p)p^{-s}} - 1\right| = \sum_p\left|\frac{\chi(p)}{p^s-\chi(p)}\right| \leq M\sum_{n=1}^{\infty}\frac{1}{n^{\Re(s)}-M},$$

and this finite because $\Re(s) > 1$, so the product converges absolutely for $\Re(s) > 1$. □

With this fact, we make a definition and derive an amazing corollary.

Definition 2. *Define the Riemann zeta function $\zeta(s)$ for all $s \in \mathbb{C}$ by*

$$\zeta(s) = \sum_{n=1}^{\infty}\frac{1}{n^s}.$$

(Note: we will often omit the bounds on the sum; when we do so, it should be assumed the sum over n is over $n \in \mathbb{N}$.)

Clearly, if $\chi(n) = 1$ for all $n \in \mathbb{N}$, then $L(s,\chi) = \zeta(s)$ (in the notation of the previous theorem). Since the constant function $\chi(n) = 1$ is completely multiplicative and bounded, we obtain the following corollary:

Corollary 3 (Euler's product formula). *The Riemann zeta function converges absolutely for all $\Re(s) > 1$, and it has the product representation*

$$\sum_n \frac{1}{n^s} = \zeta(s) = \prod_p \frac{1}{1-p^{-s}},$$

which also converges absolutely for $\Re(s) > 1$.

Meditations on the equality of this infinite sum and infinite product are good for the soul, I believe, if and only if one understands the equality. Euler sure did, and his soul was well. The equality in Corollary 3 has been called the "all-important formula" and the "golden key" by Havil[6] and Derbyshire,[7] respectively.

[6] Havil, J. *The All-Important Formula.* Princeton, NJ: Princeton University Press, 1979.

[7] Derbyshire, J. *Prime Obsession: Bernhard Riemann and the Greatest Unsolved Problem in Mathematics.* New York, NY: Penguin, 2004.

Furthermore, this equality is an analytic one, yet it is equivalent to the Fundamental Theorem of Arithmetic. For we used the Fundamental Theorem in our proof of the equality, and, conversely, if the equality holds, then every $n \in \mathbb{N}$, the n-term in the series must correspond to one and only one term in the infinite product, once the multiplication is performed.

Using Corollary 3, we can easily provide our second proof of the infinitude of the primes. Considering the series representation of $\zeta(s)$, we see that $\zeta(1)$ is the harmonic series which diverges to infinity. Suppose, there existed finitely many primes. Then by Corollary 3,

$$\lim_{s\to 1^+} \zeta(s) = \lim_{s\to 1^+} \prod_p \frac{1}{1-p^{-s}} = \prod_p \frac{1}{1-p^{-1}} < \infty,$$

which contradicts $\zeta(s) \to \infty$ as $s \to 1^+$.

2.2 The Prime Number Theorem

With the infinitude of the primes proven, one might ask how "dense" the primes are in $\mathbb{N}$. For example, the sequences $2, 4, 6, 8, \ldots$ and $4, 16, 64, 256, \ldots$ each contain an infinite number of terms, but the first sequence, in a sense, contains many more terms. To make this notion rigorous, we could define

$$\psi_1(x) = \#\{2m | m \in \mathbb{N}, 2m \leq x\}$$

and

$$\psi_2(x) = \#\{4^m | m \in \mathbb{N}, 4^m \leq x\}.$$

Then $\psi_1(x)/x$ is approximately $\frac{1}{2}$ while $\psi_2(x)/x$ is approximately $\frac{1}{\log 4}\frac{\log x}{x}$, which goes to 0 as $x \to \infty$ and is therefore less than $\frac{1}{2}$.

In a similar way, we give one (famous) answer to the density of primes question by defining the function

$$\pi(x) = \#\{p \leq x | p \text{ is prime}\}.$$

Then, for example, $\pi(6) = \#\{2, 3, 5\} = 3$ and $\pi(7) = \#\{2, 3, 5, 7\} = 4$. We will use this definition of $\pi(x)$ throughout the paper.

The Prime Number Theorem (or PNT) gives us an asymptotic formula for counting the number of primes less than x. The PNT states that $\pi(x) \sim \frac{x}{\log x}$ where the logarithm is taken base e (here and throughout the paper) and the notation $f(x) \sim g(x)$ means that

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

That is, to count the number of primes less than, say, 1000, we could calculate $1000/\log 1000 \approx 144.76$. The actual value is $\pi(1000) = 168$. The ratio of the actual value to the approximate is about 1.161. If we do the same for a million, we have $\pi(10^6) = 78,498$ and $10^6/\log(10^6) \approx 72,382$. The ratio between these values is 1.084. As we will prove, this ratio converges to 1.

As an interesting aside, one might notice the presence of the *transcendental* number e in the formula given by the PNT. Given how uncomfortable the Greeks were when they discovered the irrational numbers, imagine how distraught they would be upon discovering that e shows up in a theorem describing the density in $\mathbb{N}$ of the multiplicative building blocks of $\mathbb{N}$. This is an elementary question, yet the answer requires, not just irrational numbers, but transcendental numbers! Perhaps this is part of the reason why algebraic prodigy Niels Abel described the PNT as "perhaps the most remarkable theorem in mathematics."

2.2.1 History of the Prime Number Theorem

A timeline of the history of the PNT is as follows:

1793 Gauss begins thinking the function $\pi(x)$.

1798 Gauss and Legendre conjecture something close to the PNT; Legendre specifically conjectures that $\pi(x) = \frac{x}{A \log x + B}$ where $A = 1$ and $B = -1.08366$.

1848 Chebyshev proved that if $\pi(x) \sim x/\log x^N$ for some positive integer N, then $N = 1$.

1852 Chebyshev uses elementary properties of the factorial and logarithmic functions to show that there existed x_0 such that for all $x > x_0$,

$$B < \frac{\pi(x)}{x/\log x} < \frac{6}{5}B,$$

where $B \approx 0.921$ and $\frac{6}{5}B \approx 1.11$.

1892 Sylvester improved on Chebyshev's bound, increasing the value of B to about 0.956, but he closes his paper with a pessimistic comment about the generalizability of Chebyshev and his (elementary) methods.

1896 Hadamard and de la Vallée Poussin applied Hadamard's theory of integral functions to the Riemann zeta function, as well as a simple trigonometric identity to prove the PNT.

1921 Hardy asks whether he should believe that there exists an elementary proof (i.e., one which does not involve complex analysis) to what seems like an elementary fact about the natural numbers.

1932 Landau and Wiener provide simplified proofs of the PNT.

1948 Paul Erdös and Atle Selberg find a truly elementary proof, to the surprise of Hardy and many others. (The development of the proof and of the controversy over who got the credit is an interesting story, which is summarized in [G].)

1980 Donald Newman, following in the footsteps of Wiener and Ikehara's simplifications, simplifies contour integral methods used in older proofs, to both avoid estimates at ∞ and the use of Fourier transforms. D. Zagier's summarized proof spans only three pages.

Many great theorems have multiple proofs. As this brief history hopefully shows, the PNT is not lacking in proofs. The proof we give is essentially Newman's, but we follow Zagier's paper, which pulls results from Euler, Riemann, Chebyshev, Hadamard, de la Vallée Poussin, Mertens, and Newman.

2.2.2 An analytic theorem

Before we get to Newman's proof of the PNT, we first state and prove what we, following Newman, will call the Analytic Theorem. It is, for his purposes, a lemma, so we state and prove it here. The result is the most difficult to prove of the sequence of facts Newman proves on his way to proving the PNT. It might seem unmotivated, but rest assured: it is just what we need! But first, for *our* purposes, we prove the following lemma (the proof of which must have been so obvious to Zagier that he felt no need to mention it in his paper):

Lemma 4. *Fix $R > 0$. Suppose $g : \mathbb{C} \to \mathbb{C}$ is holomorphic on $\Re(s) \geq 0$. Then there exists $\eta > 0$ such that g is holomorphic at s for all $s \in \mathbb{C}$ such that $|s| \leq R$ and $\Re(s) \geq -\eta$.*

Proof. Since g is holomorphic for $\Re(s) \geq 0$, it is (in particular) holomorphic on

$$A_0 = \{s \in \mathbb{C} : \Re(s) = 0, |s| \leq R\}.$$

By definition, for each $s \in A_0$, there exists $r_s > 0$ such that g is holomorphic at each point in the disk $B_{r_s}(s) = \{w \in \mathbb{C} : |w - s| < r_s\}$. Because this disk is open, it makes sense to say that g is holomorphic on the disk or, more loosely, at each point the disk. Since A_0 can be written

$$A_0 = \{iy : -R \leq y \leq R\},$$

A_0 is clearly closed and bounded, which means that it is compact. Note that

$$\bigcup_{s \in A_0} B_{r_s}(s) \supseteq A_0,$$

so by the definition of compactness, there exist $s_1, \ldots, s_n \in A_0$ such that

$$\bigcup_{k=1}^{n} B_{r_k}(s_k) \supseteq A_0,$$

where we have set $r_k = r_{s_k}$ for each k to simplify the notation. If needed, relabel the n points $s_1, \ldots, s_n$ such that $y_1 < y_2 < \cdots < y_n$, where $y_k = \Im(s_k)$ for each k. Also, if $B_{r_k}(s_k) \subseteq B_{r_l}(s_l)$ for some $k \neq l$, then throw out s_k from the list and relabel.

Next, define $w_0 = -Ri$, $w_n = Ri$, and

$$w_j = \frac{(s_{j+1} - ir_{j+1}) + (s_j + ir_j)}{2}$$

for all $j = 1, \ldots, n-1$. As is clear from Figure 1, $w_0 \in B_{r_1}(s_1)$, $w_n \in B_{r_n}(s_n)$, and $w_j \in B_{r_j}(s_j) \cap B_{r_{j+1}}(s_{j+1})$ for each $j = 1, \ldots, n-1$. Therefore, since each $B_{r_k}(s_k)$ is open, which also implies that the intersection of any two is also open, we can choose positive numbers $\eta_0, \ldots, \eta_n$ such that $B_{\eta_0}(w_0) \subseteq B_{r_1}(s_1)$, $B_{\eta_n}(w_n) \subseteq B_{r_n}(s_n)$, and $B_{\eta_j}(w_j) \subseteq B_{r_j}(s_j) \cap B_{r_{j+1}}(s_{j+1})$ for all $j = 1, \ldots, n-1$. Set $\eta = \frac{1}{2}\min\{\eta_0, \ldots, \eta_n\}$. Then for each $j = 1, \ldots, n$, both $w_j - \eta \in B_{r_j}(s_j)$ and $w_{j-1} - \eta \in B_{r_j}(s_j)$. (The $\frac{1}{2}$ in the definition of η was so that we could claim that the $w_j - \eta$ are in the balls.) Since balls are convex, this implies that, for each $j = 1, \ldots, n$, every point on the segment connecting w_j and w_{j-1} is in $B_{r_j}(s_j)$ and is therefore in $\bigcup_{k=1}^{n} B_{r_k}(s_k)$. Since the union of these segments is the segment connecting $w_0 = -Ri$ and $w_n = Ri$, we have that

$$A_{-\eta} \subseteq \bigcup_{k=1}^{n} B_{r_k}(s_k),$$

where

$$A_{-\eta} = \{-\eta + iy : -R \leq y \leq R\}.$$

Furthermore, given any $s = \sigma + it \in \mathbb{C}$ such that $|s| \leq R$ and $-\eta < \sigma < 0$, we have $it \in A_0$ and $-\eta + iy \in A_{-\eta}$, which implies that $iy, -\eta + iy \in \bigcup_{k=1}^{n} B_{r_k}(s_k)$. In fact, since $-\eta + iy$ is further from each s_k than iy, there exists k such that $iy, -\eta + iy \in B_{r_k}(s_k)$. Thus by the convexity of $B_{r_k}(s_k)$,

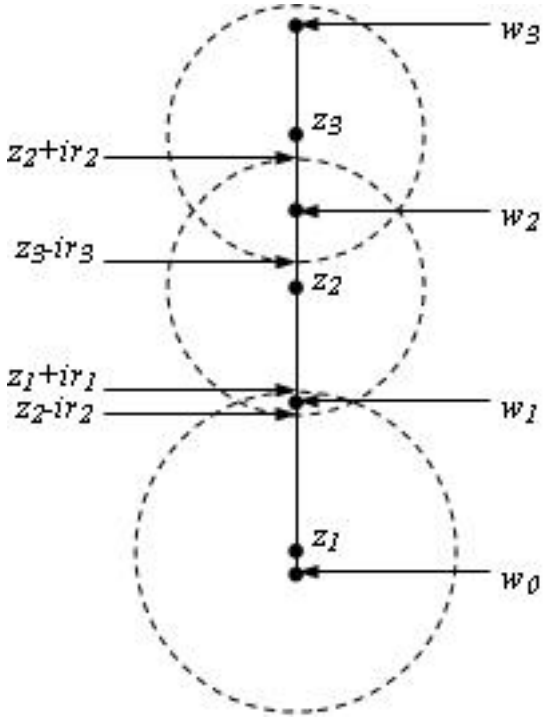

Figure 1: This figure demonstrates how we choose the points w_j; here, the balls centered at s_1, s_2, and s_3 cover A_0.

$-\eta' + iy \in B_{r_k}(s_k)$ for all $\eta' \in (0, \eta)$. Since $s = \sigma + iy$ was an arbitrary complex number such that $|s| \leq R$ and $-\eta < \sigma < 0$, we have that

$$A_{-\eta'} \subseteq \bigcup_{k=1}^{n} B_{r_k}(s_k)$$

for all $-\eta < -\eta' \leq 0$ (where $A_{-\eta'}$ is defined in the same way as $A_{-\eta}$). So, in fact, the union of all such $A_{-\eta'}$ is contained in $\bigcup_{k=1}^{n} B_{r_k}(s_k)$, which (remember) contained only points at which $g(s)$ is holomorphic. Therefore, $g(s)$ is holomorphic on

$$\bigcup_{-\eta \leq -\eta' \leq 0} A_{-\eta'} = \{\sigma + iy : -\eta \leq \sigma \leq 0, -R \leq y \leq R\}.$$

In particular, $g(s)$ is holomorphic for all $s \in \mathbb{C}$ such that $|s| \leq R$ and $\Re(s) \geq -\eta$. □

Now we move on to the theorem.

Theorem 5 (Analytic Theorem). *Let $f(x)$ be a bounded and locally integrable function*[8] *and suppose that the function $g(s) = \int_0^\infty f(x)e^{-sx}dx$, defined for $\Re(s) > 0$, extends holomorphically to $\Re(s) \geq 0$. Then $\int_0^\infty f(x)dx$ exists and is equal to $g(0)$.*

Proof. For all $T > 0$, define

$$g_T(s) = \int_0^T f(x)e^{-sx}dx.$$

[8] A *locally integrable function* is one whose integral is finite over any compact subset of the function's domain.

We wish to show that $\lim_{T\to\infty} g_T(0) = g(0)$, as this would prove the theorem. Our strategy will be to estimate the difference between $g_T(0)$ and $g(0)$ in terms of a contour integral in $\mathbb{C}$, and to prove that it goes to 0. To this end, let $\epsilon > 0$.

Our first step is choosing real numbers M, $R = R(M, \epsilon)$, $\eta = \eta(R)$, $M' = M'(\eta, R)$, $\delta = \delta(\eta, M')$, and $T_0 = T_0(\delta, \epsilon, M', R)$. We will then let $T > T_0$, and show that this implies $|g_T(0) - g(0)| < \epsilon$ by means of an estimate of a contour integral.

To begin, we know that $f(x)$ is bounded. Choose $M > 0$ such that $|f(x)| \leq M$ for all $x \geq 0$. Next, choose $R > 3M/\epsilon$. Given R, use Lemma 4 and choose $\eta > 0$ such that $g(s)$ is holomorphic at all $s \in \mathbb{C}$ such that $|s| \leq R$ and $-\eta \leq \sigma$.

Since $g(s)$ is holomorphic on the closed and bounded – that is, compact – subset containing all $s \in \mathbb{C}$ such that $|s| \leq R$ and $\sigma > -\eta$, $g(s)$ is bounded on this set. Choose $M' > 0$ such that $|g(s)| \leq M'$ for all s in this region.

Next, choose $\delta > 0$ such that $\delta < \eta$ and

$$\sin^{-1}\left(\frac{\delta}{R}\right) < \frac{\pi\epsilon}{9M'}.$$

(Note that such a δ exists, because $\sin^{-1}(\delta/R) \to 0$ continuously as $\delta \to 0$.) Finally, choose $T_0 > 0$ such that

$$e^{-\delta T_0} < \frac{\pi\delta\epsilon}{18RM'}.$$

(Again, $\delta > 0$, so $e^{-\delta T_0} \to 0$ continuously as $T_0 \to \infty$, so this is possible.) Let $T > T_0$. We claim that $|g_T(0) - g(0)| < \epsilon$.

To compute this estimate, we rewrite this difference as the value of a contour integral. Define the contour C in $\mathbb{C}$ as the boundary of the region

$$\{s \in \mathbb{C} : |s| \leq R, \Re(s) \geq -\delta\},$$

positively oriented. Around this contour, we will integrate

$$h(s) = (g_T(s) - g(s))e^{sT}\left(1 + \frac{s^2}{R^2}\right)\frac{1}{s}.$$

Note the following

- $g(s)$ is holomorphic on C (by our choice of δ, which is less that η, so $g(s)$ is bounded on C.
- By the Leibniz Integral Rule (see Appendix A.2, since 0 and T do not depend on s,

  $$\frac{d}{ds}g_T(s) = \int_0^T \frac{\partial}{\partial s}\left(f(x)e^{-sx}\right)dx = \int_0^T f(x)(-x)e^{-sx}dx.$$

  Since $f(x)$ is locally integrable, the integrand is locally integrable. Therefore, this integral, which is the derivative of $g_T(s)$, exists for all $s \in \mathbb{C}$. So $g_T(s)$ is entire and, as a result, bounded on any bounded subset of $\mathbb{C}$.
- $e^{sT}\left(1 + s^2/R^2\right)$ is entire and, as a result, bounded on any bounded subset of $\mathbb{C}$.
- s^{-1} has a simple pole at $s = 0$.
- $h(s)$ is holomorphic except for a simple pole at 0, since it is the product of s^{-1} and functions which are holomorphic inside C.

From these considerations, since 0 is in the region of which C is a boundary, Cauchy's Residue Theorem implies that

$$\int_C h(s)ds = 2\pi i \text{Res}(h,0) = 2\pi i \lim_{s\to 0} sh(s) = 2\pi i(g_T(0) - g(0)),$$

where the last equality follows because $g(s)$ and $g_T(s)$ are holomorphic and therefore continuous at 0. From this, we reevaluate our goal: We must show that

$$\frac{1}{2\pi}\left|\int_C h(s)\right| = |g_T(0) - g(0)| < \epsilon.$$

We split the estimate into three pieces. Define $C_\pm = C \cap \{s \in \mathbb{C} : \pm\Re(s) > 0\}$. Then

$$\begin{aligned}\left|\int_C h(s)ds\right| &\leq \left|\int_{C_+} h(s)ds\right| \\ &\quad + \left|\int_{C_-} g_T(s)e^{sT}\left(1+\frac{s^2}{R^2}\right)\frac{ds}{s}\right| \\ &\quad + \left|\int_{C_-} g(s)e^{sT}\left(1+\frac{s^2}{R^2}\right)\frac{ds}{s}\right| \end{aligned} \tag{1}$$

For the first integral, we note first that, since $\Re(s) > 0$ for $s \in C_+$,

$$\begin{aligned}|g(s) - g_T(s)| &= \left|\int_T^\infty f(x)e^{-sx}dx\right| \leq \int_T^\infty |f(x)|e^{-\Re(s)x}dx \\ &\leq M\int_T^\infty e^{-\Re(s)x}dx = \frac{M}{\Re(s)}e^{-\Re(s)T}\end{aligned}$$

for all $s \in C_+$. Also, since $|s| = R$ for $s = \sigma + it \in C_+$,

$$\begin{aligned}\left|e^{sT}\left(1+\frac{s^2}{R^2}\right)\frac{1}{s}\right| &= e^{\Re(s)T}|\sigma^2 + t^2 + (\sigma+it)^2|R^{-3} \\ &= e^{\Re(s)T}|2\sigma||\sigma+it|R^{-3} \\ &= 2e^{\Re(s)T}|\Re(s)|R^{-2}.\end{aligned}$$

Therefore, if we denote the arc length of C_+ by $\ell(C_+)$,

$$\begin{aligned}\left|\int_{C_+} h(s)ds\right| &\leq \ell(C_+)\max_{s\in C_+}|h(s)| \\ &\leq \left(\frac{2\pi R}{2}\right)\left(\frac{M}{\Re(s)}e^{-\Re(s)T}\right)\left(e^{\Re(s)T}\frac{2\Re(s)}{R^2}\right) \\ &= \frac{2\pi M}{R} \\ &< 2\pi\frac{\epsilon}{3}\end{aligned}$$

(since $R > 3M/\epsilon$). So far, so good. Now we estimate the second piece. Since $g_T(s)$ is entire, we can continuously deform[9] C_- into $C'_- = \{s \in \mathbb{C} : |s| = R,\ \Re(s) < 0\}$ and replace the integral over

[9]See Appendix A.1.1.

C_- with the integral over C'_-. Now for all $s \in C'_-$,

$$\begin{aligned} |g_T(s)| &\leq \int_0^T |f(x)| e^{-\Re(s)x} dx \leq M \int_0^T e^{-\Re(s)x} dx \\ &= \frac{M}{-\Re(s)} \left(e^{-\Re(s)T} - 1 \right) \leq \frac{M}{|\Re(s)|} e^{-\Re(s)T} \end{aligned}$$

and, as in the first integral,

$$\left| e^{sT} \left(1 + \frac{s^2}{R^2} \right) \frac{1}{s} \right| = 2e^{\Re(s)T} \frac{|\Re(s)|}{R^2}.$$

So

$$\begin{aligned} \left| \int_{C'_-} g_T(s) e^{sT} \left(1 + \frac{s^2}{R^2} \right) \frac{ds}{s} \right| &\leq \ell(C'_-) \max_{s \in C'_-} \left| g_T(s) e^{sT} \left(1 + \frac{s^2}{R^2} \right) \frac{1}{s} \right| \\ &\leq \frac{2\pi R}{2} \left(\frac{M}{|\Re(s)|} e^{-\Re(s)T} \right) \left(2e^{\Re(s)T} \frac{|\Re(s)|}{R^2} \right) \\ &= \frac{2\pi M}{R} \\ &< 2\pi \frac{\epsilon}{3}. \end{aligned}$$

Finally, we estimate the third integral in three parts. Let

$$\begin{aligned} C_1 &= \{ s \in C_- : \Re(s) = -\delta \}, \\ C_2 &= \{ s \in C_- : \Re(s) > -\delta, \Im(s) > 0 \}, \\ C_3 &= \{ s \in C_- : \Re(s) > -\delta, \Im(s) < 0 \}. \end{aligned}$$

These three contours clearly partition C_-, so

$$\left| \int_{C_-} g(s) e^{sT} \left(1 + \frac{s^2}{R^2} \right) \frac{ds}{s} \right| \leq \sum_{j=1}^{3} \left| \int_{C_j} g(s) e^{sT} \left(1 + \frac{s^2}{R^2} \right) \frac{ds}{s} \right|.$$

For the integral over C_1, we have $\ell(C_1) \leq 2R$, $|g(s)| \leq M'$,

$$\left| e^{sT} \right| = e^{-\delta T} \leq e^{-\delta T_0} < \frac{\pi \delta \epsilon}{18RM'}$$

(by our choice of T_0), and

$$\left| \left(1 + \frac{s^2}{R^2} \right) \frac{1}{s} \right| \leq (1+1) \frac{1}{\delta} = \frac{2}{\delta}.$$

So

$$\begin{aligned} \left| \int_{C_1} g(s) e^{sT} \left(1 + \frac{s^2}{R^2} \right) \frac{ds}{s} \right| &\leq \ell(C_1) \max_{s \in C_1} \left| g(s) e^{sT} \left(1 + \frac{s^2}{R^2} \right) \frac{1}{s} \right| \\ &\leq (2R) M' e^{-\delta T} \frac{2}{\delta} \\ &< \frac{4RM'}{\delta} \left(\frac{\pi \delta \epsilon}{18M'} \right) \\ &= 2\pi \frac{\epsilon}{9}. \end{aligned}$$

For the integral over C_2, we have

$$\ell(C_2) = R\sin^{-1}\left(\frac{\delta}{R}\right) < \frac{\pi\epsilon R}{9M'}$$

(by our choice of δ), $|g(s)| \leq M'$, $\left|e^{sT}\right| = e^{\Re(s)T} \leq 1$, and

$$\left|\left(1+\frac{s^2}{R^2}\right)\frac{1}{s}\right| \leq (1+1)\frac{1}{R} = \frac{2}{R}.$$

So

$$\begin{aligned}
\left|\int_{C_2} g(s)e^{sT}\left(1+\frac{s^2}{R^2}\right)\frac{ds}{s}\right| &\leq \ell(C_2)\max_{s\in C_2}\left|g(s)e^{sT}\left(1+\frac{s^2}{R^2}\right)\frac{1}{s}\right| \\
&\leq \left(R\sin^{-1}\left(\frac{\delta}{R}\right)\right)M'\left(\frac{2}{R}\right) \\
&< \left(\frac{\pi\epsilon R}{9M'}\right)\frac{2M'}{R} \\
&= 2\pi\frac{\epsilon}{9}.
\end{aligned}$$

A similar – no, an identical – calculation yields the same bound for the integral over C_3. Therefore, the third integral in Equation 1 is bounded by

$$2\pi\frac{\epsilon}{9} + 2\pi\frac{\epsilon}{9} + 2\pi\frac{\epsilon}{9} = 2\pi\frac{\epsilon}{3},$$

which means the the integral over C is bounded by

$$2\pi\frac{\epsilon}{3} + 2\pi\frac{\epsilon}{3} + 2\pi\frac{\epsilon}{3} = 2\pi\epsilon.$$

So

$$2\pi|g_T(0) - g(0)| = \left|\int_C h(s)ds\right| < 2\pi\epsilon,$$

which, after dividing both sides by 2π, gives us our result. □

2.2.3 Newman's proof of the Prime Number Theorem

The proof follows Zagier's sequence of steps, up to some reordering of the lemmas (including the analytic theorem). Zagier was kind enough to leave plenty of details for the reader to work out. For example, his proofs of Lemmas 10, 12, and 13 are two sentences long, and his proof of Lemma 11 is a mere three. I have *greatly* expanded on his proofs with the intention of making it understandable to the undergraduate who has had some analysis. My hope is that one could read this and understand it in one (possibly long) sitting, as opposed to the multiple (infinitely long) sittings which I have required. We begin with a definition:

Definition 6. *For all $x \in \mathbb{R}$ and $s \in \mathbb{C}$, define*

$$\Phi(s) = \sum_p \frac{\log p}{p^s} \quad \textit{and} \quad \vartheta(x) = \sum_{p\leq x} \log p,$$

where the sums over p run over the prime natural numbers.

With these definitions in hand, we prove the Prime Number Theorem (following [Z]). The proof follows easily from a sequence of steps, which we prove here as lemmas. The first two lemmas build up to the third:

Lemma 7. *For $\Re(s) > 1$, $\zeta(s)$ and $\Phi(s)$ are absolutely convergent.*

Proof. Corollary 3 proves that $\zeta(s)$ is absolutely convergent, so we must show that $\Phi(s)$ is too. Suppose $\sigma = \Re(s) > 1$. Then consider that

$$\sum_p \left| \frac{\log p}{p^s} \right| \leq \sum_{n=2}^{\infty} \frac{\log n}{n^\sigma} \leq \frac{\log 2}{2^\sigma} + \int_2^\infty \frac{\log n}{n^\sigma} dn.$$

A computation shows that the value of the integral is $(\sigma - 1)^{-2}$ when $\sigma > 1$, from which we conclude that $\Phi(s)$ absolutely converges for all $\sigma > 1$. □

Lemma 8. *For all $\delta > 1$, $\zeta(s)$ and $\Phi(s)$ are uniformly convergent on $\sigma = \Re(s) \geq \delta$. (When this is the case, we say these series are locally uniformly convergent).*

Proof. We use the Weierstrass M-test. Fix $\delta > 1$. For all $n \in \mathbb{N}$, set $M_n = n^{-\delta}$ and $M'_n = n^{-\delta} \log n$. Then because $\delta > 1$, both $\sum_{n=1}^\infty M_n$ and $\sum_{n=1}^\infty M'_n$ converge. Also, $|n^{-s}| = n^{-\sigma} \leq M_n$ and $|n^{-s} \log n| = n^{-\sigma} \log n \leq M'_n$ for all $\sigma > 1$. So by the Weierstrass M-test, $\zeta(s)$ and $\Phi(s)$ converge uniformly for $\sigma \geq \delta$. □

From these two lemmas we can prove the following lemma:

Lemma 9. *For $\Re(s) > 1$, $\zeta(s)$ and $\Phi(s)$ are holomorphic.*

Proof. Let $\sigma = \Re(s) > 1$. By Lemma 8, $\zeta(s)$ and $\Phi(s)$ converge uniformly at s. Then if we can show that the "termwise derivatives" of these series also converge uniformly at s, then we could conclude that $\zeta'(s)$ and $\Phi'(s)$ exist and are equal to their respective termwise derivatives (see Appendix A.1.2).

To do this, we compute the series of termwise derivatives, obtaining

$$\sum_{n=1}^\infty \frac{d}{ds} \frac{1}{n^s} = \sum_{n=1}^\infty -\frac{\log n}{n^s} \quad \text{and} \quad \sum_p \frac{d}{dx} \frac{\log p}{p^s} = \sum_p -\frac{\log^2 p}{p^s}$$

These are clearly uniformly convergent by the Weierstrass M-test; the argument is similar to that which proved Lemma 8. □

In the next two lemmas, we modify $\zeta(s)$ and $\Phi(s)$ slightly so as to enlarge the region over which it is holomorphic. More specifically, we subtract $(s-1)^{-1}$ from both functions and prove that the "modified" functions are holomorphic over some larger region.

Lemma 10. *For $\sigma = \Re(s) > 0$, $\zeta(s) - 1/(s-1)$ is holomorphic. (That is, $\zeta(s) - \frac{1}{s-1}$ extends holomorphically to $\Re(s) > 0$.)*

Proof. For $\sigma > 1$, we have

$$\frac{1}{s-1} = \int_1^\infty \frac{dx}{x^s} = \sum_{n=1}^\infty \int_n^{n+1} \frac{dx}{x^s}.$$

Furthermore, this series absolutely converges for $\sigma > 1$ since

$$\sum_{n=1}^{\infty} \left| \int_n^{n+1} \frac{dx}{x^s} \right| \leq \sum_{n=1}^{\infty} \int_n^{n+1} \frac{dx}{x^\sigma} = \sum_{n=1}^{\infty} \frac{1}{\sigma - 1} \left(\frac{1}{n+1} - \frac{1}{n} \right) = \frac{1}{\sigma - 1} < \infty.$$

Therefore, since the series for $\zeta(s)$ also converges absolutely for $\sigma > 1$ by Corollary 3, we can subtract the series termwise, obtaining

$$\zeta(s) - \frac{1}{s-1} = \sum_{n=1}^{\infty} \int_n^{n+1} \frac{dx}{n^s} - \sum_{n=1}^{\infty} \int_n^{n+1} \frac{dx}{x^s} = \sum_{n=1}^{\infty} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s} \right) dx, \tag{2}$$

which is an absolutely convergent, holomorphic function for $\sigma > 1$ (because $\zeta(s)$ and $(s-1)^{-1}$ are for $\sigma > 1$).

Now we show that this series converges absolutely for all $\sigma > 0$. We evaluate the following in two different ways. On one hand,

$$\left| s \int_n^{n+1} \int_n^x \frac{du}{u^{s+1}} dx \right| = \left| \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s} \right) dx \right|.$$

On the other hand,

$$\left| s \int_n^{n+1} \int_n^x \frac{du}{u^{s+1}} dx \right| \leq ((n+1) - n)(x - n) \max_{\substack{n \leq x \leq n+1 \\ n \leq u \leq x}} \left| \frac{s}{u^{s+1}} \right| \leq (1)(1) \frac{s}{n^{\sigma+1}},$$

so

$$\sum_{n=1}^{\infty} \left| \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s} \right) dx \right| \leq \sum_{n=1}^{\infty} \frac{s}{n^{\sigma+1}},$$

which converges because $\sigma + 1 > 1$. Hence the series in Equation 2 converges absolutely for $\sigma > 0$. Now consider the termwise derivative of the series in Equation 2:

$$\frac{d}{ds} \left(\zeta(s) - \frac{1}{s-1} \right) = \sum_{n=1}^{\infty} \frac{d}{ds} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s} \right) dx = \sum_{n=1}^{\infty} \int_n^{n+1} \left(-\frac{\log n}{n^s} + \frac{\log x}{x^s} \right) dx.$$

(Note the use of the Leibniz integral rule again!)

If this series converges uniformly for $\sigma > 0$, we will conclude that $\zeta(s) - (s-1)^{-1}$ is holomorphic for $\sigma > 0$. To prove this, we use the same trick as before:

$$\begin{aligned} \left| \int_n^{n+1} \left(-\frac{\log n}{n^s} + \frac{\log x}{x^s} \right) dx \right| &= \left| \int_n^{n+1} \int_n^x \left(\frac{1 - s \log u}{u^{s+1}} \right) du dx \right| \\ &\leq ((n+1) - n)(x - n) \max_{\substack{n \leq x \leq n+1 \\ n \leq u \leq x}} \left| \frac{1 - s \log u}{u^{s+1}} \right| \\ &\leq \frac{1 + \sigma \log(n+1)}{n^{\sigma+1}}, \end{aligned}$$

which implies

$$\sum_{n=1}^{\infty} \left| \int_n^{n+1} \left(-\frac{\log n}{n^s} + \frac{\log x}{x^s} \right) dx \right| \leq \sum_{n=1}^{\infty} \frac{1 + \sigma \log(n+1)}{n^{\sigma+1}}.$$

Since $\sigma + 1 > 1$, the series for $\frac{d}{ds}(\zeta(s) - (s-1)^{-1})$ converges absolutely and therefore exists. $\square$

Lemma 11. *For $\sigma = \Re(s) \geq 1$, $\zeta(s) \neq 0$ and $\Phi(s) - 1/(s-1)$ is holomorphic.*

Proof. First, for $\sigma > 1$, the product representing $\zeta(s)$ converges absolutely by Corollary 3, hence it converges to a nonzero finite number (see Appendix A.4.1). Later in the proof, we show that $\zeta(s) \neq 0$ for $\sigma = 1$ on our way to showing the second half of the result.

By Lemma 9, $\Phi(s)$ is holomorphic for $\Re(s) > 1$. So since the same is true for $(s-1)^{-1}$, the result clearly follows for $\Re(s) > 1$. It is left to show that $\Phi(s) - (s-1)^{-1}$ is holomorphic for $\Re(s) = 1$.

As a first step, note that $\Phi(s)$ and the series $\sum_p p^{-s}(p^s - 1)^{-1} \log p$ converge absolutely for $\sigma > 1$. Hence

$$\Phi(s) + \sum_p \frac{\log p}{p^s(p^s-1)} = \sum_p \log p \left(\frac{1}{p^s} + \frac{1}{p^s(p^s-1)} \right) = \sum_p \frac{\log p}{p^s - 1}. \tag{3}$$

Now consider that

$$\begin{aligned} -\frac{\zeta'(s)}{\zeta(s)} &= -\frac{1}{\zeta(s)} \frac{d}{dz} \prod_p \frac{1}{1-p^{-s}} \\ &= -\frac{1}{\zeta(s)} \sum_p \frac{d}{dz} \left(\frac{1}{1-p^{-s}} \right) \prod_{q \neq p} \frac{1}{1-q^{-s}} \\ &= -\sum_p \left(\frac{-p^{-s} \log p}{(1-p^{-s})^2} \right) (1-p^{-s}) \\ &= \sum_p \frac{\log p}{p^s - 1}. \end{aligned} \tag{4}$$

Equating Equations 3 and 4, solving for $\Phi(s)$, and subtracting $(s-1)^{-1}$, we obtain

$$\Phi(s) - \frac{1}{s-1} = -\left(\frac{\zeta'(s)}{\zeta(s)} + \frac{1}{s-1} \right) + \sum_p \frac{\log p}{p^s(p^s-1)}. \tag{5}$$

Considering that

$$\left| \frac{\log p}{p^s(p^s-1)} \right| \leq \frac{\log p}{(p^\sigma - 1)^2} \leq \frac{\log p}{4p^{2\sigma}},$$

we could prove that the sum over p on the right-hand side of Equation 5 is uniformly convergent for $\sigma > \frac{1}{2}$ by using the Weierstrass M-test with $M_p = \frac{\log p}{4p^{2\sigma}}$. We would then prove the same about the sequence of derivatives of the partial sums and conclude that it is holomorphic for $\sigma > \frac{1}{2}$.

Also, as we now show, the other term on the right-hand side of Equation 5 is holomorphic on $\sigma > 0$, except where it is undefined. This would prove that the left-hand side of Equation 5, $\Phi(s) - (s-1)^{-1}$, is holomorphic (except where it is undefined) for $\sigma > 0$.

By Lemma 10, the function

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

is holomorphic for $\sigma(s) > 0$. Hence the numerator of this function is holomorphic for $\sigma > 0$, which implies that $(s-1)\zeta(s)$ is holomorphic for $\sigma > 0$. Since complex-differentiability implies that derivatives of all orders exist (see Appendix A.1.2), we conclude that $\frac{d}{ds}((s-1)\zeta(s))$ is also holomorphic. Hence

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

is holomorphic for $\sigma > 0$ except where it is undefined, which could only be where $(s-1)\zeta(s) = 0$.

We first check that $(s-1)\zeta(s) \neq 0$ at $s = 1$. Since $\zeta(s) - (s-1)^{-1}$ is holomorphic on $\sigma > 0$ by Lemma 10, it is holomorphic at $s = 1$. In particular, the (finite) value at 1 exists and the limit of $\zeta(s) - (s-1)^{-1}$ is that value. Then we have

$$\lim_{s\to 1}\left(\zeta(s) - \frac{1}{s-1}\right) = \lim_{s\to 1}\left(\frac{(s-1)\zeta(s)-1}{s-1}\right). \tag{6}$$

Because $s-1$ converges to 0 as $s \to 1$, it must be the case that $(s-1)\zeta(s) - 1$ converges to 0 as well. So $(s-1)\zeta(s)$ is not 0 at $s = 1$.

Now we check that $(s-1)\zeta(s) \neq 0$ for $\sigma \neq 1$ and $t = \Im(s) \neq 0$. That is, we want to show that $\zeta(1+it) \neq 0$ for all real $t \neq 0$. (This will also complete the proof of the first part of this theorem.) Suppose on the contrary that $\zeta(s)$ has a zero of order μ at $s = 1+it$. Because $\zeta(s) - \frac{1}{s-1}$ is holomorphic at $s = 1+it$ and since $t \neq 0$, $\zeta(s)$ is holomorphic at $s = 1+it$. Therefore, by Lemma 10, $\mu \geq 0$. Write $\zeta(s) = (s-1-it)^{\mu}g(s)$. Showing that $\mu = 0$ would show that $\zeta(s)$ has no zeros on the line $\sigma = 1$, so we proceed to this now.

We need to get our hands on μ. To do this, note that $\zeta'(s) = \mu(s-1-it)^{\mu-1}g(s) + (s-1-it)^{\mu}g'(s)$. This implies, for all $\epsilon > 0$,

$$\frac{\zeta'(1 \pm it + \epsilon)}{\zeta(1 \pm it + \epsilon)} = \frac{\mu}{(1 \pm it + \epsilon) - 1 - it} + \frac{g'(1 \pm it + \epsilon)}{g(1 \pm it + \epsilon)}.$$

Multiplying both sides by ϵ and taking $\epsilon \to 0$ yields

$$\lim_{\epsilon\to 0^+} \epsilon\frac{\zeta'(1 \pm it + \epsilon)}{\zeta(1 \pm it + \epsilon)} = \mu + \lim_{\epsilon\to 0^+} \epsilon\frac{g'(1 \pm it + \epsilon)}{g(1 \pm it + \epsilon)}. \tag{7}$$

But $\zeta(s)$, being once differentiable for $\sigma > 1$, is infinitely differentiable for $\sigma > 1$, so $g'(s)$ is differentiable and therefore bounded for $\sigma > 1$. Also $g(1+it) \neq 0$ because otherwise $1+it$ would have order strictly greater than μ. Finally, since

$$(\overline{s} - 1 - it)^{\mu}g(\overline{s}) = \zeta(\overline{s}) = \sum_{n=1}^{\infty}\frac{1}{n^{\overline{s}}} = \overline{\zeta(s)} = (\overline{s} - 1 - it)^{\mu}\overline{g(s)},$$

it follows that $g(1+it) = \overline{g(1-it)} \neq \overline{0} = 0$. Thus the limit involving $g(s)$ and $g'(s)$ in Equation 7 is 0, and the right-hand side is equal to μ. Great! Now to interpret the left-hand side, we use Equation 5. Set $s = 1 \pm it$ in this equation, multiply both sides by a positive ϵ and take $\epsilon \to 0^+$. Then the $\epsilon(s-1)^{-1}$ terms vanish, as does the term involving the sum over p because it is bounded. This leaves us with

$$\lim_{\epsilon\to 0^+} \epsilon\Phi(1 \pm it + \epsilon) = -\lim_{\epsilon\to 0^+} \epsilon\frac{\zeta'(1 \pm it + \epsilon)}{\zeta(1 \pm it + \epsilon)} = -\mu. \tag{8}$$

Now it will be convenient to consider the point $1 + 2it$ as a zero of order ν. As we argued for μ, we have that $\nu \geq 0$. (The actual value of ν has no bearing on our proof; it is merely a convenience to introduce it.) Now by analogous reasoning, we could prove an analogue of Equation 8 for ν:

$$\lim_{\epsilon\to 0^+} \epsilon\Phi(1 \pm 2it + \epsilon) = -\nu. \tag{9}$$

Finally, if we substitute $1 + \epsilon$ into Equation 5, multiply by $\epsilon > 0$, and take $\epsilon \to 0^+$, we obtain

$$\lim_{\epsilon\to 0^+} \epsilon\Phi(1 + \epsilon) = 1. \tag{10}$$

To complete the proof – which, recall, required us to show that $\mu = 0$ – we come up with an equality in terms of μ and ν, from which we will conclude that $\mu = 0$. To do this, note that

$$\begin{aligned}
0 &\leq \sum_p \frac{\log p}{p^{1+\epsilon}}(2\Re(p^{it/2}))^4 \\
&= \sum_p \frac{\log p}{p^{1+\epsilon}}(p^{it/2} + p^{-it/2})^4 \\
&= \sum_p \frac{\log p}{p^{1+\epsilon+2it}} + 4\sum_p \frac{\log p}{p^{1+\epsilon+it}} + 6\sum_p \frac{\log p}{p^{1+\epsilon}} \\
&\quad +4\sum_p \frac{\log p}{p^{1+\epsilon-it}} + \sum_p \frac{\log p}{p^{1+\epsilon-2it}} \\
&= \Phi(1+2it+\epsilon) + 4\Phi(1+it+\epsilon) + 6\Phi(1+\epsilon) \\
&\quad +4\Phi(1-it+\epsilon) + \Phi(1-2it+\epsilon). \qquad (11)
\end{aligned}$$

Now multiply both sides of Equation 11 by $\epsilon > 0$, take $\epsilon \to 0^+$, and use Equations 8, 9, and 10 to arrive at

$$0 \leq -\nu - 4\mu + 6 - 4\mu - \nu = 6 - 8\mu - 2\nu.$$

Since $\nu \geq 0$, this inequality implies $8\mu \leq 6$, from which we (finally!) conclude that $\mu = 0$. □

The next step uses some new notation, which we introduce now before continuing. Given functions defined on $\mathbb{R}$, we say that $f(x) = \mathcal{O}(g(x))$ if there exist $C \in \mathbb{R}$ and $x_0 \in \mathbb{R}$ such that $|f(x)| \leq C|g(x)|$ for all $x > x_0$. For example, the function $f(x) = \log x + \log^2 x + x$ is $\mathcal{O}(x)$ because, in the limit as $x \to \infty$, the x term dominates, from which we could prove that $f(x) = \mathcal{O}(x)$ using the definition.

Now we go back to the proof of the Prime Number Theorem.

Lemma 12. $\vartheta(x) = \mathcal{O}(x)$.

Proof. First let $n \in \mathbb{N}$. Then consider that

$$e^{\vartheta(2n)-\vartheta(n)} = \exp\left(\sum_{n<p\leq 2n} \log p\right) = \prod_{n<p\leq 2n} p$$

where k ranges over the integers (in the specified range) and p over the primes. Now let

$$a = \prod_{n<k\leq 2n} k, \quad b = \prod_{1\leq k\leq n} k, \quad \text{and} \quad c = \prod_{n<p\leq 2n} p.$$

Then $\exp(\vartheta(2n) - \vartheta(n)) = c$, and $\frac{a}{b} = \binom{2n}{n}$, where $\binom{2n}{n}$ is the binomial coefficient. Now note that $b \mid a$ since $\binom{2n}{n}$ is always an integer – a cool fact in itself! Also note that $c \mid a$ since c is a product over a subset of the set of terms over which a is a product. Furthermore, note that if p is a prime dividing c, then $p > n$, which implies $p \nmid k$ for all terms k in the product a. Hence b and c are relatively prime, from which we conclude that bc divides a. In particular, $bc \leq a$, so

$$e^{\vartheta(2n)-\vartheta(n)} = c \leq \frac{a}{b} = \binom{2n}{n} \leq \sum_{k=0}^{2n} \binom{2n}{k} = (1+1)^{2n} = 2^{2n}.$$

Taking logarithms, we have $\vartheta(2n) - \vartheta(n) \leq 2n \log 2$.

Next, we use this last equality to prove that $\vartheta(x) - \vartheta(\frac{x}{2}) \leq 2x \log 2$ for all $x \in (0, \infty)$. If $x \in (0, 2]$, the inequality holds trivially because $\vartheta(x) = \vartheta(\frac{x}{2}) = 0$ unless $x = 2$, in which case the inequality reduces to $\log 2 \leq 4 \log 2$. So let $x \in (2, \infty)$, and choose $m \in \mathbb{N}$ such that $2m < x \leq 2m + 1$. Then

$$\vartheta(x) = \sum_{p \leq x} \log p \leq \sum_{p \leq 2m+1} \log p \leq \vartheta(2m) + \log(2m + 1)$$

(where the last inequality is strict if and only if $2m + 1$ is prime). Also,

$$\vartheta\left(\frac{x}{2}\right) = \sum_{p \leq \frac{x}{2}} \log p \geq \sum_{p \leq m} \log p = \vartheta(m).$$

Hence

$$\begin{aligned} \vartheta(x) - \vartheta\left(\frac{x}{2}\right) &\leq \vartheta(2m) - \vartheta(m) + \log(2m + 1) \\ &\leq 2m \log 2 + \log(2m + 1) \\ &\leq x \log 2 + \log(x + 1) \\ &\leq 2x \log 2, \end{aligned}$$

where the last inequality follows since $\log(x + 1) \leq x \log 2$ for all $x > x_0 = 2$.

Finally, we use this estimate to prove the lemma! Choose $C = 2 \log 2$ and $x_0 = 2$, and let $x > x_0$. Next, choose $r \in \mathbb{N}$ such that $2^r \leq x < 2^{r+1}$. Then $\vartheta(x/2^r) = 0$, and hence

$$\begin{aligned} \vartheta(x) &= \sum_{k=1}^{r} \left(\vartheta\left(\frac{x}{2^{k-1}}\right) - \vartheta\left(\frac{x}{2^k}\right) \right) \leq \sum_{k=1}^{r} \log 2 \left(\frac{x}{2^{k-1}}\right) \\ &= x \log 2 \left(\frac{1 - 2^{-r}}{1 - 2^{-1}}\right) \leq 2x \log 2 = Cx, \end{aligned}$$

which proves the lemma. □

Lemma 13. *The improper integral* $\displaystyle\int_1^\infty \frac{\vartheta(x) - x}{x^2} dx$ *converges.*

Proof. Let I be the value of the integral, be it finite or infinite. We are going to apply the analytic theorem to prove that this integral convergence, so, in order to do this, we must check that the hypotheses of Theorem 5. We change variables $x = e^t$ in the integral to get

$$I = \int_1^\infty \left(\frac{\vartheta(x)}{x} - 1\right) \frac{dx}{x} = \int_0^\infty \left(\vartheta\left(e^t\right) e^{-t} - 1\right) dt = \int_0^\infty f(t) dt,$$

where we have defined $f(t) = \vartheta(e^t)e^{-t} - 1$. Note that, by Lemma 12, for some constant C,

$$|f(t)| \leq |\vartheta(e^t)|e^{-t} + 1 \leq (Ce^t)e^{-t} + 1 = C + 1$$

for all t past some t_0. Since $f(t)$ has only jump discontinuities at the prime t, it is clearly bounded on $(0, t_0]$ and hence bounded on $(0, \infty)$. Furthermore, for any compact $K \subset (0, \infty)$, $f(t)$ has only finitely many discontinuities, so the integral of $f(t)$ over K is finite. Hence f is locally integrable.

Now define $g(z) = \int_0^\infty f(t)e^{-zt} dt$ for all $\Re(z) \geq 0$. (As we will show, $g(z)$ is holomorphic on $\Re(z) \geq 0$, so the convergence of the integral is not a problem.) As we will also show,

$$\int_0^\infty \vartheta(e^t)e^{-(z+1)t} dt = \frac{\Phi(z + 1)}{z + 1}, \tag{12}$$

which implies

$$g(z) = \int_0^\infty \vartheta(e^t)e^{-(z+1)t}dt - \int_0^\infty e^{-zt}dt = \frac{\Phi(z+1)}{z+1} - \frac{1}{z}$$

or, equivalently,

$$g(z) = \frac{1}{z+1}\left(\Phi(z+1) - \frac{1}{z} - 1\right).$$

If this holds for $\Re(z) > 0$, then $\Re(z+1) > 1$, then Lemma 11 would then imply that $\Phi(z+1) - z^{-1}$ is holomorphic for $\Re(z) \geq 0$. The assumptions of the analytic theorem would therefore be satisfied, from which we could conclude that I, the integral in question which equals $\int_0^\infty f(t)dt$, exists and is equal to $g(0)$ (a bounded quantity since $g(z)$ is holomorphic at $z = 0$), thus proving the theorem.

All we need to do is prove is Equation 12. To do this, we let J be the value of the integral in Equation 12, set $s = z + 1$, and make the variable substitution $x = e^t$ to get

$$sJ = \int_1^\infty \frac{s\vartheta(x)}{x^{s+1}}dx = \sum_{k=0}^\infty \int_{p_k}^{p_{k+1}} \frac{s\vartheta(x)}{x^{s+1}}dx,$$

where $p_1, p_2, \ldots$ is an enumeration of the primes and $p_0 = 1$ for convenience. The reason for splitting up the integral in this way is because, for $x \in (p_k, p_{k+1})$,

$$\vartheta(x) = \sum_{p \leq x} \log p = \sum_{p \leq p_k} \log p = \vartheta(p_k).$$

This allows us to pull the $\vartheta(x)$ out of the integral in order to evaluate it:

$$sJ = \sum_{k=0}^\infty \vartheta(p_k) \int_{p_k}^{p_{k+1}} \frac{s}{x^{s+1}}dx = \sum_{k=0}^\infty \vartheta(p_k)\left(\frac{1}{p_k^s} - \frac{1}{p_{k+1}^s}\right).$$

This *almost* looks like a telescoping series – how we would love to see one of those right now! In fact, if we note that $\vartheta(p_k) = \vartheta(p_{k+1}) - \log p_{k+1}$ and recall that the series $\sum_p \vartheta(p)p^{-s}$ and $\Phi(s)$ converge absolutely, we arrive at exactly this:

$$sJ = \sum_{k=0}^\infty \left(\frac{\vartheta(p_k)}{p_k^s} - \frac{\vartheta(p_{k+1})}{p_{k+1}^s} + \frac{\log p_{k+1}}{p_{k+1}^s}\right) = \sum_{k=0}^\infty \left(\frac{\vartheta(p_k)}{p_k^s} - \frac{\vartheta(p_{k+1})}{p_{k+1}^s}\right) + \Phi(s).$$

Since the value of the telescoping series is $\vartheta(1)/1 = 0$ for all $\Re(s) > 1$, we have $J = \Phi(z+1)/(z+1)$ for $\Re(z) > 0$, from which Equation 12 follows immediately. □

Lemma 14. *$\vartheta(x) \sim x$.*

Proof. We first show the following: For all $\epsilon > 0$, there exists $x_0 > 0$ such that $x > x_0$ implies $\vartheta(x)/x - 1 < \epsilon$. A parallel proof (which we omit) will show the corresponding inequality $1 - \vartheta(x)/x < \epsilon$, which will prove (straight from the definition of convergence) that $\vartheta(x)/x \to 1$ as $x \to \infty$.

We proceed by contradiction. Suppose there exists an $\epsilon > 0$ such that for all $x_0 > 0$, there exists $x > x_0$ for which $\vartheta(x)/x - 1 \geq \epsilon$ or, equivalently, $\vartheta(x) \geq \lambda x$ where $\lambda = 1 + \epsilon > 1$. Fix such an ϵ and let $x_0 = 1$. Choose $x > x_0$ such that $\vartheta(x) \geq \lambda x$. Then, since $\vartheta(x)$ is an increasing function,

$$\int_x^{\lambda x} \frac{\vartheta(t) - t}{t^2}dt \geq \int_x^{\lambda x} \frac{\vartheta(x) - t}{t^2}dt \geq \int_x^{\lambda x} \frac{\lambda x - t}{t^2}dt.$$

This integral's value is $C(\lambda) = \lambda - \log\lambda - 1$, which is a strictly positive quantity since $\lambda > 1$.

Now we repeat this process. Set $x_1 = x$, $x_0 = \lambda x$, and find an $x > x_0$ such that $\vartheta(x) \geq \lambda x$. Set $x_2 = x$ and continue inductively to define a sequence $x_1, x_2, \ldots$ where each $x_{n+1} > \lambda x_n$ and

$$\int_{x_{n+1}}^{\lambda x_{n+1}} \frac{\vartheta(t) - t}{t^2} dt \geq C(\lambda)$$

for each n. Then

$$\int_1^{\infty} \frac{\vartheta(t) - t}{t^2} dt \geq \sum_{n=1}^{\infty} \int_{x_n}^{\lambda x_n} \frac{\vartheta(t) - t}{t^2} dt \geq \sum_{n=1}^{\infty} C(\lambda),$$

which clearly diverges since $C(\lambda) > 0$. This contradicts Lemma 13, which concludes the proof. □

With all of the hard work done, we can prove the Prime Number Theorem:

Theorem 15 (The Prime Number Theorem). *Let $\pi(x)$ denote the number of primes less than or equal to x. Then $\pi(x) \sim x/\log x$.*

Proof. Let $\epsilon > 0$. Without loss of generality, assume that $\epsilon < 1$. To help with the estimates, we

- Choose $\eta \in (0, \epsilon)$ such that $\eta < \frac{\epsilon}{4+\epsilon}$. (This inequality implies $\frac{1+\eta}{1-\eta} < 1 + \frac{\epsilon}{2}$.)
- Choose $x_1 > 0$ such that $x > x_1$ implies $(1-\eta)x < \vartheta(x) < (1+\eta)x$. (This is possible by Lemma 14.)
- Choose $x_2 > 0$ such that $x^{-\eta} \log x < 1 + \frac{\epsilon}{3}$ for all $x > x_2$. (This is possible because $x^{-\eta} \log x \to 0$ as $x \to \infty$, as can be seen by applying l'Hôpital's rule.)

Set $x_0 = \max\{x_1, x_2\}$, and let $x > x_0$. Then from

$$(1-\eta)x < \vartheta(x) = \sum_{p \leq x} \log p \leq \sum_{p \leq x} \log x = \pi(x) \log x,$$

we deduce $\pi(x) \log x/x - 1 > -\eta > -\epsilon$.

Now we obtain the other estimate. Since $\log x$ is an increasing function, $p > x^{1-\eta}$ implies $\log p > (1-\eta) \log x$. Hence

$$\begin{aligned} (1+\eta)x &> \vartheta(x) \geq \sum_{x^{1-\eta} < p \leq x} \log p \geq \sum_{x^{1-\eta} < p \leq x} \log x \\ &= (1-\eta) \log x \left(\pi(x) - \pi(x^{1-\eta})\right) \geq (1-\eta) \log x \left(\pi(x) - x^{1-\eta}\right). \end{aligned}$$

Dividing both sides by $(1-\eta)x$ yields

$$\frac{1+\eta}{1-\eta} > \frac{\pi(x) \log x}{x} - \frac{\log x}{x^{\eta}},$$

from which we conclude

$$\frac{\pi(x) \log x}{x} - 1 < \frac{1+\eta}{1-\eta} - 1 + \frac{\log x}{x^{\eta}} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Thus $\pi(x) \log x/x$ is within ϵ of 1 for all $x > x_0$, as required! □

Sit. Meditate. Enjoy the beauty which lies before us.[10]

[10] What comes next is sure to shock those who do not know and pleasantly remind those who do. Either way, the task before us is a great one, so if a break is desired, now is the time.

3 Primes in arithmetic progressions

We move on now to a topic which motivated the birth of analytic number theory: primes in arithmetic progressions. So far, we have proven that there exist infinitely many primes, and we have given a formula for the density of these primes in the set of natural numbers. Ignoring the prime 2, we could have equivalently considered these questions about the odd numbers – that is, the *arithmetic progression* $\{1+2n\}_{n=0}^{\infty}$ – and come to the same conclusions.

In this section, we generalize to arbitrary arithmetic progressions $\{a+qn\}_{n=0}^{\infty}$ where $a, q \in \mathbb{N}$, and we ask the same two questions: Are there infinitely many primes, and what is their density in $\mathbb{N}$?

3.1 Dirichlet's Theorem

Dirichlet's theorem, or, more properly, Dirichlet's theorem on primes in arithmetic progressions, states that there exist infinitely many primes in any arithmetic progression

$$a, a+q, a+2q, a+3q, \ldots$$

so long as $\gcd(a,q) = 1$. (If a and q have a common factor, then $a+qn$ for all $n \in \mathbb{N}$ is composite.)

We will prove this result in a sequence of steps. Our proof will follow Davenport's, which primarily follows Dirichlet's original. We give Euler's proof for $q = 2$. We already have seen the result for $q = 2$, both from Euclid's argument and by considering the Euler product formula for $\zeta(s)$, but we give a third proof which, for one thing, proves a stronger fact, and, for another thing, ultimately inspired Dirichlet who generalized the argument to compose a proof of his theorem. (Note: This is the way Dirichlet did it, and it is how Davenport presents it.)

3.1.1 Euler's proof for $q = 2$

In this section, we provide a third proof of the fact that there exist infinitely many primes. Recall that $\zeta(s) = \sum_n n^{-s} = \prod_p (1-p^{-s})^{-1}$ for all $s \in (1,\infty)$. Taking logarithms[11] and expanding the logarithms into power series (which will be allowed since $|p^{-s}| < 1$, implies that p^{-s} is within the radius of convergence of 1), we obtain

$$\begin{aligned}
\log\zeta(s) &= \log\prod_p (1-p^{-s})^{-1} = -\sum_p \log(1-p^{-s}) \\
&= -\sum_p \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m}(-p^{-s})^m = \sum_p \sum_{m=1}^{\infty} m^{-1}p^{-sm} \\
&= \sum_p p^{-s} + \delta(s) \qquad (13)
\end{aligned}$$

where

$$\begin{aligned}
\delta(s) &= \sum_p \sum_{m=2}^{\infty} m^{-1}p^{-sm} < \sum_p \sum_{m=2}^{\infty} p^{-sm} = \sum_p p^{-2s}\left(\frac{1}{1-p^{-s}}\right) \\
&= \sum_p \frac{1}{p^s(p^s-1)} < \sum_p \frac{1}{p^2} < \zeta(2).
\end{aligned}$$

[11] Technically, we are about to turn the logarithm of an infinite product into an infinite sum, which may cause problems that do not arise with finite sums and products. We excuse ourselves in Appendix A.4.1 by proving that this works as it does in the finite case.

(Note that we have sneakily rearranged the order of summation, but this is okay since the series converges absolutely, as can be seen since $|m^{-1}p^{-sm}| < p^{-sm}$ and $-sm < -1$.) Since $\log \zeta(s) \to \infty$ as $s \to 1^+$, and since $\delta(s)$ is bounded in the same limit, we conclude from Equation 13 that

$$\sum_p \frac{1}{p} \text{ diverges.}$$

This fact clearly implies the infinitude of the primes, and it is interesting in its own right – the primes are not too sparsely scattered in $\mathbb{N}$.

In order to prove Dirichlet's theorem, we will show the corresponding (generalized) fact:

$$\sum_{p \equiv_q a} \frac{1}{p} \text{ diverges,}$$

where the notation $p \equiv_q a$ means that p is congruent to a modulo q, and the sum over $p \equiv_q a$ is over all primes p congruent to a modulo q. This is quite a task, however, so we build our bridge to the theorem in steps. First, we let $q > 2$ be prime, define Dirichlet characters and Dirichlet L-functions for q, and prove the theorem for q. We then let q be composite, define general Dirichlet characters and Dirichlet L-functions, and prove the theorem in general. Each step for the second case is a generalization of the corresponding step in the first case. Let's begin!

3.1.2 Dirichlet characters and L-functions (prime modulus)

In this section, we construct Dirichlet characters of a prime modulus and prove a key property about them. Dirichlet characters for the modulus q are arithmetic functions $\chi : \mathbb{N} \to \mathbb{C}$, which are q-periodic (that is, periodic with period q) and completely multiplicative (that is, $\chi(mn) = \chi(m)\chi(n)$ for all $m, n \in \mathbb{N}$). We also require that there exists a linear combination whose value is 1 if $n \equiv_q a$ and 0 otherwise. (Note: Dirichlet characters satisfy the hypotheses of Theorem 1.)

To construct the Dirichlet characters for the prime modulus q, consider the multiplicative group $\mathbb{Z}_q^* = \mathbb{Z}_q \backslash \{0\}$. As is well known, this group is cyclic (see Appendix A.3.1), so there exists an element $g \in \mathbb{Z}_q$ such that g generates $\mathbb{Z}_q^*$. For all $n \in \mathbb{Z}_q^*$, define $\nu(n)$ to be the index of n relative to g; that is, define $\nu(n)$ such that $g^{\nu(n)} \equiv_q n$. Then, for each of the $q-1$ primitive $(q-1)$-th roots of unity ω, define the Dirichlet character on $\mathbb{Z}_q^*$ corresponding to ω by $\omega^{\nu(n)}$. Finally, periodically extend the definition of $\omega^{\nu(n)}$ to define the character for all $n \in \mathbb{N}$ which are not divisible by q, and for those n for which $q \mid n$, define $\omega^{\nu(n)} = 0$.

The character $\omega^{\nu(n)}$ is clearly well-defined, for $n = m$ implies $n \equiv_q m$, which implies $g^{\nu(n)} \equiv n \equiv m \equiv g^{\nu(m)}$ modulo q, which implies $\nu(n) = \nu(m) + k(q-1)$ for some integer k. Therefore, since $\omega^{q-1} = 1$, $\omega^{\nu(n)} = \omega^{\nu(m)}$.

Also, the function is periodic by construction. To show complete multiplicativity, note that $k = mn$, for $m \neq 0 \neq n$, implies $k \equiv mn$ modulo q, which implies

$$g^{\nu(k)} \equiv k \equiv mn \equiv g^{\nu(m)} g^{\nu(n)} = g^{\nu(m)+\nu(n)}.$$

Since $g^{q-1} \equiv 1$ modulo q, $\nu(k) = \nu(m) + \nu(n) + l(q-1)$ for some $l \in \mathbb{Z}$, which implies

$$\omega^{\nu(k)} = \omega^{\nu(m)} \omega^{\nu(n)} \left(\omega^{q-1}\right)^l = \omega^{\nu(m)} \omega^{\nu(n)}.$$

Finally, if m or n is 0, then so is k, which implies that

$$\omega^{\nu(k)} = 0 = \omega^{\nu(m)} \omega^{\nu(n)}.$$

Hence $\omega^{\nu(n)}$ is completely multiplicative.

Finally, we must show that some linear combination of the $\omega^{\nu(n)}$ gives us 1 or 0, according to whether $n \equiv_q a$ or $n \not\equiv_q a$, respectively. But first, how many $\omega^{\nu(n)}$ are there? When we defined the functions, we specified that $\omega^{q-1} = 1$, so each ω is one of $q-1$ primitive $(q-1)$-th roots of unity. On the other hand, if $\omega_1^{\nu(n)} = \omega_2^{\nu(n)}$, then, in particular, they are equal for $n = g$, for which $\nu(n) = 1$, which implies $\omega_1 = \omega_2$. So there are $q-1$ functions, each of them distinct.

Before we construct such a linear combination, we give an example of a set of characters.

Example 16. *Let $q = 5$. We choose 2 as our generator. (This choice is arbitrary, but this will not present a problem to us later on. We typically choose some generator at the beginning and it is assumed that we keep the same generator.) Then*

$$\nu(1) = 0,\ \nu(2) = 1,\ \nu(3) = 3,\ \text{and}\ \nu(4) = 2.$$

The $(q-1)$-th roots of unity are 1, i, -1, and $-i$, and the corresponding characters are 1, $i^{\nu(n)}$, $(-1)^{\nu(n)}$, and $(-i)^{\nu(n)}$. A table of these values is given here:

	1	2	3	4
1	1	1	1	1
$i^{\nu(n)}$	1	i	$-i$	-1
$(-1)^{\nu(n)}$	1	-1	-1	1
$(-i)^{\nu(n)}$	1	$-i$	i	-1

Moving on, we prove the key property about Dirichlet characters.

Theorem 17. *For a prime $q > 2$,*

$$\frac{1}{q-1}\sum_{\omega} \omega^{-\nu(a)}\omega^{\nu(n)} = \begin{cases} 1, & n \equiv_q a; \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

where the sum is over all $(q-1)$-th roots of unity.

Proof. First, note that if $q-1 \mid k$, for some $k \in \mathbb{Z}$, then $\omega^k = 1$ for all $(q-1)$-th roots of unity ω, and hence $\sum_\omega \omega^k = q-1$. On the other hand, if $q-1 \nmid k$, then $\omega^k \neq 1$, and powers of ω^k generate all of the $\left(\frac{q-1}{g}\right)$-th roots of unity, where $g = \gcd(k, q-1)$. From this, we have that

$$\{\omega^{kn}\}_{n=0}^{l-1} = \{\omega^{kn}\}_{n=l}^{2l-1} = \cdots = \{\omega^{kn}\}_{n=(g-1)l}^{gl-1}.$$

So

$$\sum_{\omega} \omega^k = \sum_{j=1}^{g}\sum_{n=0}^{l-1} \omega^{kn} = \sum_{j=1}^{g}\left(\frac{\omega^{kl}-1}{\omega^k - 1}\right) = 0$$

since $\omega^{kl} = \left(\omega^k\right)^l = 1$. Therefore, if $q-1$ does not divide k, then $\sum_\omega \omega^k = 0$. We have proven

$$\sum_{\omega} \omega^k = \begin{cases} q-1, & q-1 \text{ divides } k; \\ 0, & q-1 \text{ does not divide } k. \end{cases}$$

If we divide both sides by $q-1$, we set $k = -\nu(a) + \nu(n)$, and note that $q-1 \mid k$ if and only if $k \equiv_{q-1} 0$, which holds if and only if $\nu(a) \equiv_q \nu(n)$, which in turn holds if and only if $n \equiv_q a$, we see that our result follows from the above equality. □

Hence the $q-1$ arithmetic functions $\omega^{\nu(n)}$ form a set of Dirichlet characters for the prime modulus q. With these characters in hand, Dirichlet defined what are now called Dirichlet L-functions. Following Dirichlet, we now proceed with the definition of these functions.

Definition 18. *For each Dirichlet character $\omega^{\nu(n)}$, define the Dirichlet L-function by*

$$L_\omega(s) = \sum_{n=1}^{\infty} \omega^{\nu(n)} n^{-s}$$

for all $s \in \mathbb{C}$.

By Theorem 1, $L_\omega(s)$ absolutely converges for $\Re(s) > 1$ and has the following product representation

$$L_\omega(s) = \prod_p \frac{1}{1 - \omega^{\nu(p)} p^{-s}},$$

which also converges absolutely for $\Re(s) > 1$. Now recall that $\omega^{\nu(n)} = 0$ for all n which are divisible by q, so, in fact,

$$L_\omega(s) = \sum_{n \not\equiv_q 0} \omega^{\nu(n)} n^{-s}. \tag{15}$$

Also, since $\omega^{\nu(q)} = 0$, the $p = q$ term in the product representation is 1, so we have

$$L_\omega(s) = \prod_{p \neq q} \frac{1}{1 - \omega^{\nu(p)} p^{-s}}. \tag{16}$$

Now note that since

$$|\omega^{\nu(p)} p^{-s}| = p^{-\Re(s)} < p^{-1} < 1/2,$$

no term in the product is 0. So since $L_\omega(s)$ converges absolutely by Theorem 1, we have that $L_\omega(s) \neq 0$ for all $s \in \mathbb{C}$ with $\Re(s) > 1$. (This is a general fact about infinite products. See Appendix A.4.1 for an explanation of why this is true.)

Similar to our calculation of $\log \zeta(s)$ in Section 3.1.1, since $L_\omega(s) \neq 0$, we have

$$\log L_\omega(s) = \sum_{p \neq q} \sum_{m=1}^{\infty} m^{-1} \left(\omega^{\nu(p)} p^{-s} \right)^m = \sum_{p \neq q} \sum_{m=1}^{\infty} m^{-1} \omega^{\nu(p^m)} p^{-ms}$$

by the multiplicativity of $\omega^{\nu(n)}$. Now we multiply both sides by $\omega^{-\nu(a)}$, sum over all $(q-1)$-th roots of unity ω, rearrange infinite sums at will (because they are all absolutely convergent!), and obtain

$$\begin{aligned} \sum_\omega \omega^{-\nu(a)} \log L_\omega(s) &= \sum_\omega \omega^{-\nu(a)} \sum_{p \neq q} \sum_{m=1}^{\infty} m^{-1} \omega^{\nu(p^m)} p^{-ms} \\ &= \sum_{p \neq q} \sum_{m=1}^{\infty} m^{-1} p^{-ms} \sum_\omega \omega^{-\nu(a)} \omega^{\nu(p^m)}. \end{aligned}$$

Because of Equation 14, this implies that

$$\sum_\omega \omega^{-\nu(a)} \log L_\omega(s) = (q-1) \sum_{p \neq q} \sum_{\substack{m=1 \\ p^m \equiv_q a}}^{\infty} m^{-1} p^{-ms}. \tag{17}$$

On the right-hand side of Equation 17, we have

$$(q-1)\sum_{p\equiv_q a} p^{-s} + (q-1)\sum_{p\neq q}\sum_{\substack{m=2\\ p^m\equiv_q a}}^{\infty} m^{-1}p^{-ms}.$$

Since the double sum over p and m has all positive terms, it is bounded below by 0 and above by $\delta(s) < \zeta(2)$, as we saw in Section 3.1.1. Therefore, we have proven the following lemma to Dirichlet's theorem:

Lemma 19. *If*

$$\sum_{\omega} \omega^{-\nu(a)} \log L_\omega(s)$$

diverges to infinity as $s \to 1^+$, then the series $\sum_p p^{-s}$ diverges to infinity.

The conclusion of this theorem, if true, would prove Dirichlet's theorem. Since we want to prove the conclusion, let us consider this lemma's hypothesis. One of the $(q-1)$ roots of unity is 1 itself, so the sum on the left-hand side of Equation 17 is

$$\sum_{\omega} \omega^{-\nu(a)} \log L_\omega(s) = \log L_1(s) + \sum_{\omega\neq 1} \omega^{-\nu(a)} \log L_\omega(s). \tag{18}$$

If we consider the $\omega = 1$ term, we see from Equation 16 and Corollary 3 that

$$L_1(s) = \prod_{p\neq q} \frac{1}{1-p^{-s}} = (1-q^{-s})\zeta(s), \tag{19}$$

so since $\zeta(s) \to \infty$ as $s \to 1^+$ yet $1 - p^{-s}$ remains bounded, $L_\omega(s) \to \infty$ as $s \to 1^+$. This is a good start! If we can show that each of the remaining terms is bounded as $s \to 1^+$, we will apply Lemma 19 to conclude Dirichlet's theorem.

Before embarking on this venture, however, we can simplify our task a bit. First, note that the $\omega^{\nu(n)}$ have magnitude 1, so they remain bounded as $s \to 1^+$. This means that what we have to show is that $\log L_\omega(s)$ is bounded as $s \to 1^+$. Since the logarithm function is continuous (for $\Re(s) > 0$ – we have chosen the principle branch of the logarithm without mentioning it), we have that

$$\lim_{s\to 1^+} \log L_\omega(s) = \log \lim_{s\to 1^+} L_\omega(s).$$

Hence our task reduces to proving that $L_\omega(s)$ remains finite but nonzero as $s \to 1^+$.

But we can simplify our task once more. We can show that $L_\omega(s)$ is continuous and therefore finite at $s = 1$, so $\lim_{s\to 1^+} L_\omega(s) = L_\omega(1)$. Once we do so, we will have the following precursor to Dirichlet's theorem:

Lemma 20. *If $L_\omega(s) \neq 0$ for all $\omega \neq 1$, then Dirichlet's theorem holds.*

Proof. We first prove that $L_\omega(s)$ is continuous using the Dirichlet test for convergence and the series representation of $L_\omega(s)$ given in Equation 15. To do this, first note that n^{-s} decreases as n increases and converges to 0. Second, note that for all $N \in \mathbb{N}$, $\sum_{n=1}^{N} \omega^{\nu(n)}$ is bounded above by $q-2$. The reason for this is that as n runs through $q-1$ consecutive integers, $\nu(n)$ runs through each possible index $0, 1, \ldots, q-2$, so

$$\sum_{n=1}^{q-1} \omega^{\nu(n)} = \sum_{\nu=0}^{q-2} \omega^{\nu} = 0$$

since each $(q-1)$-th root of unity shows up exactly once in the sum. Similarly, $\sum_{n=1}^{N} \omega^{\nu(n)}$ is periodically 0, and is, in fact, periodic with period q. Since periodic sequences are bounded, the second claim above follows.

Therefore, the Dirichlet test for convergence implies that $L_\omega(s)$ converges for all $\Re(s) > 0$. Furthermore,[12] this convergence is uniform for all $\Re(s) \geq \delta > 0$. Since the partial sums of $L_\omega(s)$ are continuous, their uniform limit is continuous, so $L_\omega(s)$ is continuous on $\Re(s) \geq \delta$. Supposing we chose $\delta < 1$, which we assume anyway because our choice of $\delta > 0$ was arbitrary, we have that $L_\omega(s)$ is continuous at $s = 1$.

Given now that $L_\omega(s)$ is continuous at 1 for all $\omega \neq 1$, we have that $L_\omega(s) \to L_\omega(1)$ as $s \to 1^+$ for all $\omega \neq 1$. Hence, if our assumption holds, then $L_\omega(1) \neq 0$, which implies that $\lim_{s\to 1^+} L_\omega(s) \neq 0$, which in turn implies that $\log L_\omega(s)$ remains bounded as $s \to 1^+$, for all $\omega \neq 1$. Hence, as we mentioned, since each $\omega^{\nu(a)}$ is bounded, the $\omega \neq 1$ terms on the left-hand side of Equation 17 remains bounded as $s \to 1^+$. Hence the left-side diverges to infinity since the $\omega = 1$ term does, which means that Lemma 19 applies and Dirichlet's theorem holds. □

It is hard to say where Dirichlet's proof *begins*. He created the huge body of work we have been discussing in this section. To say that the proof has already started would be fair. However, I feel that what we have shown so far are merely consequences of the definitions of Dirichlet characters and the L-functions. These are of interest to people independent of the fact that they are used to prove Dirichlet's Theorem. I think it is best to think about "Dirichlet's proof" as the proof of the fact that $L_\omega(1) \neq 0$ for all $\omega \neq 1$. Besides, the proof is already long; if we included everything so far in "the proof" it would appear more unwieldy than is necessary. Enough of this discussion – let us proceed with the proof!

3.1.3 Dirichlet's proof (prime modulus)

In this section, we show that $L_\omega(1) \neq 0$ for all $\omega \neq 1$. We consider two cases: ω is complex and $\omega = -1$. (Note: if ω is not complex, then it must be -1 since we are assuming $\omega \neq 1$.)

Suppose first that ω is complex. Since Equation 17 holds for any value of a, it, in particular, holds for $a = 1$. Substituting $a = 1$ into Equation 17 gives us

$$\sum_{\omega} \log L_\omega(s) = \sum_{p \neq q} \sum_{m=1}^{\infty} m^{-1} p^{-ms} \geq 0,$$

where the second inequality follows from $m^{-1}p^{-ms} > 0$. (The reason why the inequality can not be made strict – which would make this case a triviality – is that we do not know for sure that the sum is not empty!)

Exponentiating both sides yields

$$1 \leq \prod_{\omega} e^{\log L_\omega(s)} = \prod_{\omega} L_\omega(s). \tag{20}$$

Suppose that for some complex ω, $L_\omega(1) = 0$. Then

$$L_{\bar{\omega}}(1) = \sum_{n=1}^{\infty} \overline{\omega}^{-\nu(n)} \frac{1}{n} = \sum_{n=1}^{\infty} \overline{\omega^{-\nu(n)}} \frac{1}{n} = \overline{\sum_{n=1}^{\infty} \omega^{-\nu(n)} \frac{1}{n}} = \overline{L_\omega(1)} = 0.$$

[12] We claim this without proof, not because it is difficult, but because it seems repetitive. We have already seen how this sort of proof would go for the functions $\zeta(s)$ and $\Phi(s)$ in the proof of Lemma 8.

Now by Equation 19, the fact that $(s-1)\zeta(s) \to 1$ as $s \to 1^+$ implies that

$$(s-1)L_1(s) = (1-q^{-s})(s-1)\zeta(s) \to 1 - \frac{1}{q}$$

as $s \to 1^+$. Thus $L_1(s)$ has a simple pole at $s=1$. Consider the limit of the product of the ω-term, the $\overline{\omega}$-term, and the 1-term in the product:

$$\lim_{s\to 1^+} L_\omega(s)L_{\overline{\omega}}(s)L_1(s) \quad = \quad \lim_{s\to 1^+} \frac{L_\omega(s)L_{\overline{\omega}}(s)}{s-1}\left(L_1(s)(s-1)\right).$$

The limit of $L_1(s)(s-1)$ is bounded as we mentioned. The limit of the other part, using l'Hôpital's rule, is equal to

$$\lim_{s\to 1^+} \left(L'_\omega(s)L_{\overline{\omega}}(s) + L_\omega(s)L'_{\overline{\omega}}(s)\right).$$

The derivative of the L-functions is computed term-wise and shown to be uniformly convergent by Dirichlet's test (as we did above for $L_\omega(s)$, so we omit the prove here). Therefore, $L_\omega(s)$ and $L'_\omega(s)$, as uniform limits of continuous functions, are continuous at $s=1$. This continuity at $s=1$ implies that $L_\omega(1)$ and $L'_\omega(1)$ are finite, so since $L_\omega(1) = L_{\overline{\omega}}(1) = 0$, the limit is equal to 0. Since the other terms in the product in Equation 20 are bounded (because each $L_\omega(s)$ is bounded!), the product is 0, contradicting the inequality in Equation 20. So $L_\omega(1) \neq 0$ for each complex ω. For large q, we have just covered the vast majority of the cases for Dirichlet's proof. Unfortunately, the amount of work required to prove these cases is not proportional to the amount of work required to prove the single case in which $\omega = -1$. We proceed to this now.

Suppose $\omega = -1$. Since $\overline{-1} = -1$, the above proof fails, so we must prove $L_{-1}(s) \neq 0$ directly. What is to our advantage is that we know exactly what ω is; the Dirichlet character for -1 reduces to $\omega^{\nu(n)} = (-1)^{\nu(n)}$. Note that $q \mid n$ if and only if $\omega^{\nu(n)} = 0$. Also, from this equality, we have that $\omega^{\nu(n)} = 1$ if and only if $\nu(n) = 2k$ for some $k \in \mathbb{Z}$, which is true if and only if $n \equiv g^{\nu(n)} \equiv g^{2k} \equiv \left(g^k\right)^2$ modulo q, which in turn is true if and only if n is a square modulo q. Negating these statements, we also have that $\omega^{\nu(n)} = -1$ if and only if n is not a square modulo q. Hence

$$\omega^{\nu(n)} = \left(\frac{n}{q}\right) = \begin{cases} 0, & \text{if } q \mid n; \\ 1, & \text{if } q \nmid n \text{ and } n \text{ is a square modulo } q; \\ -1, & \text{if } q \nmid n \text{ and } n \text{ is not a square modulo } q. \end{cases}$$

where $\left(\frac{n}{q}\right)$ is the Legendre symbol.[13]

In what follows, we define the Gaussian sum $G(n)$, partially evaluate it to get $\left(\frac{n}{q}\right)$ in terms of $G(1)$, quote the value of $G(1)$, and substitute the value of $G(1)$ into the expression for $L_{-1}(1)$. At that point, we will show that $L_{-1}(1) \neq 0$. Before continuing, we specify $e^{2\pi i n/q}$ (for $n = 0, 1, \ldots, q-1$) as an arbitrary q-th root of unity, because we will need to address specific roots. We will denote this particular root by $e_q(n)$.

First, we give the definition of the Gaussian sums.

Definition 21. *Define the Gaussian sum $G(n)$ for all $n \in \mathbb{N}$ by*

$$G(n) = \sum_{m=1}^{q-1} \left(\frac{m}{q}\right) e_q(mn).$$

[13]The definition and properties of the Legendre symbol which we will use are discussed in Appendix A.3.3.

Next, we show how to rewrite the Legendre symbols in $L_{-1}(s)$ in terms of the Gaussian sums. Our first step is to prove the following lemma:

Lemma 22. *If we let G denote the numeric value of $G(1)$, then*

$$L_{-1}(1) = \frac{1}{G}\sum_{m=1}^{q-1}\left(\frac{m}{q}\right)\sum_{n=1}^{\infty}\frac{1}{n}e_q(mn) \tag{21}$$

Proof. For $n \not\equiv_q 0$, then n is a unit in $\mathbb{Z}_q^*$, which implies

$$\{m\}_{m=1}^{q-1} = \{mn\}_{m=1}^{q-1},$$

where these sets are considered as subsets of $\mathbb{Z}_q$. In this case, then, since the Legendre symbol is multiplicative,

$$\left(\frac{n}{q}\right)G(n) = \sum_{m=1}^{q-1}\left(\frac{nm}{q}\right)e_q(mn) = \sum_{m'=1}^{q-1}\left(\frac{m'}{q}\right)e_q(m'\cdot 1) = G(1).$$

Since $(\frac{n}{q})$ is its own inverse, this implies that

$$G(n) = \left(\frac{n}{q}\right)G, \tag{22}$$

where we have set $G = G(1)$. Since, for $n \equiv_q 0$,

$$G(0) = \sum_{m=1}^{q-1}\left(\frac{m}{q}\right) = \sum_{m=1}^{q-1}\omega^{\nu(m)} = 0 = \left(\frac{n}{q}\right) = \left(\frac{n}{q}\right)G(1),$$

Equation 22 holds for all $n \in \mathbb{N}$.

As it turns out, G is $\sqrt{q}$ if $q \equiv_4 1$ and $i\sqrt{q}$ if $q \equiv_4 3$. Dirichlet finished what Gauss started by showing this. We omit the (clever) calculation, which can be found in Davenport, Chapter 2. For now, we use the fact that $G \neq 0$. Then

$$\left(\frac{n}{q}\right) = \frac{1}{G}\sum_{m=1}^{q-1}\left(\frac{m}{q}\right)e_q(mn),$$

which implies

$$L_{-1}(1) = \sum_{n=1}^{\infty}\left(\frac{1}{G}\sum_{m=1}^{q-1}\left(\frac{m}{q}\right)e_q(mn)\right)\frac{1}{n}.$$

Swapping the order of summation proves Equation 21. □

The reader who knows his/her Taylor series well might notice a logarithm hidden in Equation 21. Our next step is to note this and use it to rewrite our expression for $L_{-1}(1)$.

Lemma 23. *Following the notation of the previous lemma, we have*

$$L_{-1}(1) = -\frac{1}{G}\sum_{m=1}^{q-1}\left(\frac{m}{q}\right)\left(\log\left(2\sin\frac{\pi m}{q}\right) + i\left(\frac{\pi m}{q} - \frac{\pi}{2}\right)\right). \tag{23}$$

Proof. Consider Equation 21. To evaluate the inner sum, consider the Taylor series for the function

$$-\log(1-z) = \sum_{n=1}^{\infty} \frac{1}{n} z^n,$$

which converges for all $|z| \leq 1$ except $z = 1$. (See Appendix A.4.4.)

For $|z| \leq 1$, we have $\Re(z) \leq 1$. Since $z \neq 1$, we have $\Re(z) < 1$, $\Re(1-z) > 0$, and hence $-\pi/2 < \arg(1-z) < \pi/2$. If we use the principle branch of the logarithm, we will avoid all disgusting discontinuity issues completely. Write $z = e^{i\theta}$. Thinking of 1, z, and $1-z$ as vectors in the complex plane, we can draw the triangle whose sides are 1, z, and $1-z$. Since $|z| = |1| = 1$, the triangle is isosceles, and the two equal angles α sum to $\pi - \theta$. Extending the line coincident with $1-z$, we see that $\arg(1-z) = -\alpha$. Therefore, $-2\arg(1-z) = 2\alpha = \pi - \theta$, which gives us $\arg(1-z) = \frac{1}{2}(\theta - \pi)$. Also, we can calculate

$$|1-z| = |(1-\cos\theta) - i\sin\theta| = \sqrt{2-2\cos\theta} = 2\sqrt{\frac{1-\cos\theta}{2}} = 2\sin\frac{\theta}{2}.$$

With the magnitude and the argument of $1-z$, we can calculate

$$\sum_{n=1}^{\infty} \frac{1}{n} e^{in\theta} = -\log(1-e^{i\theta}) = -\log|1-z| - i\arg(1-z),$$

which gives us

$$\sum_{n=1}^{\infty} \frac{1}{n} e^{in\theta} = -\log(2\sin\frac{\theta}{2}) - \frac{1}{2}(\theta - \pi)i. \tag{24}$$

Since $e_q(mn) = e_q(m)^n$, and since $e_q(m)$ has magnitude 1 yet is not equal to 1 (since $m \not\equiv_q 0$), we can set $\theta = 2\pi m/q$ in Equation 24 and substitute the result into Equation 21 to obtain Equation 23. □

Whew! If I had arrived at this equation myself, I would have been sure that I had made a mistake. For $L_{-1}(1)$ is a real number! In no way is it obvious that the expression on the right-hand side of Equation 23 is real. Despite all doubt, however, we have proven this equation, so it is correct.

Returning to our goal – namely, to prove that $L_{-1}(1) \neq 0$ – we use Equation 23 to complete Dirichlet's proof in the case that $q \equiv_4 3$.

Lemma 24. *For $q \equiv_4 3$, $L_{-1}(1) \neq 0$.*

Proof. In this case, we have $G = i\sqrt{q}$, so ignoring the (zero) imaginary part of Equation 23, we obtain

$$\begin{aligned} L_{-1}(1) &= -\frac{1}{\sqrt{q}} \sum_{m=1}^{q-1} \left(\frac{\pi m}{q} - \frac{\pi}{2} \right) \\ &= -\frac{\pi}{q^{3/2}} \sum_{m=1}^{q-1} m \left(\frac{m}{q} \right) + \frac{\pi}{2\sqrt{q}} \sum_{m=1}^{q-1} \left(\frac{m}{q} \right) \\ &= -\frac{\pi}{q^{3/2}} \sum_{m=1}^{q-1} m \left(\frac{m}{q} \right). \end{aligned}$$

The factor $-\pi q^{-3/2}$ is certainly not 0, and neither is the sum, for if we set $q = 4k+3$ for some $k \in \mathbb{Z}$, and

$$\sum_{m=1}^{q-1} m\left(\frac{m}{q}\right) \equiv_2 \sum_{m=1}^{q-1} m \equiv_2 \frac{q(q-1)}{2} \equiv_2 (4k+1)(2k+1) \equiv_2 1.$$

To consider an integer modulo 2 is to blur one's vision of the integer as much as possible without losing sight of it. But in this case, that this sum is congruent to 1 modulo 2 means that it is non-zero, which proves $L_{-1}(1) \neq 0$, as we wished! □

Now what about the case where $q \equiv_4 1$? We prove that $L_{-1}(1) \neq 0$, as required, but the proof is in a few steps. First, we make a definition.

Definition 25. *Let R and N range over all $m = 1, \ldots, q-1$ such that $(\frac{m}{q}) = 1$ and $(\frac{m}{q}) = -1$, respectively. (Note: R stands for residues, and N stands for nonresidues.) Define the complex number Q by*

$$Q = \frac{\prod_N \sin\frac{\pi N}{q}}{\prod_R \sin\frac{\pi R}{q}}. \tag{25}$$

Second, we translate our requirement that $L_{-1}(1) \neq 0$ into the requirement that $Q \neq 1$.

Lemma 26. *If $Q \neq 1$, then $L_{-1}(1) \neq 0$.*

Proof. Since $q \equiv_4 1$, $G = \sqrt{q}$, so since $L_{-1}(1)$ is a real number, Equation 23 becomes

$$L_{-1}(1) = -\frac{1}{\sqrt{q}} \sum_{m=1}^{q-1} \left(\frac{m}{q}\right) \log\left(2\sin\frac{\pi m}{q}\right).$$

Now let R and N range over all $m \in \{1, 2, \ldots, q-1\}$ such that $(\frac{m}{q}) = 1$ and $(\frac{m}{q}) = -1$. Then

$$\begin{aligned}
\exp(\sqrt{q}L_{-1}(1)) &= \exp\left(-\sum_R \log\left(2\sin\frac{\pi R}{q}\right) + \sum_N \log\left(2\sin\frac{\pi N}{q}\right)\right) \\
&= \exp\log\left(\frac{\prod_N 2\sin\frac{\pi N}{q}}{\prod_R 2\sin\frac{\pi R}{q}}\right) \\
&= \frac{\prod_N \sin\frac{\pi N}{q}}{\prod_R \sin\frac{\pi R}{q}},
\end{aligned}$$

where, to obtain the last equality, we used the fact that the number of quadratic resides is equal to the number of nonresidues (see Appendix A.3.3). But this is equal to Q, so if $Q \neq 1$, then $\sqrt{q}L_{-1}(1) \neq \log 1 = 0$, which proves $L_{-1}(1) \neq 0$. □

Third, we show that $Q \neq 1$, which by Lemma 26 proves that $L_{-1}(1) \neq 0$ for the case where $q \equiv_4 1$.

Lemma 27. *In the notation of the previous lemma, $Q \neq 1$.*

Proof. Note that R is a quadratic residue modulo q if and only if $q - R$ is as well, because $q \equiv_4 1$ implies that[14]

$$\left(\frac{q-R}{q}\right) = \left(\frac{-R}{q}\right) = \left(\frac{-1}{q}\right)\left(\frac{R}{q}\right) = (-1)^{(q-1)/2}\left(\frac{R}{q}\right) = \left(\frac{R}{q}\right). \tag{26}$$

Hence $\sum_R R = \sum_R (q - R)$, which implies that

$$2\sum_R R = \sum_R R + \sum_R (q - R) = q\sum_R 1 = q\frac{q-1}{2}.$$

Since the corresponding statement of each of the above statements is true when R is replaced by N, a similar conclusion holds as well. Therefore, $\sum_R R = \sum_N N = \frac{q(q-1)}{4}$. This implies

$$\begin{aligned} e_q\left(\frac{q(q-1)}{8}\right)(-2i)^{(q-1)/2}\prod_R \sin\frac{\pi R}{q} &= e_q\left(\frac{1}{2}\sum_R R\right)\prod_R\left(-2i\sin\frac{\pi R}{q}\right) \\ &= \prod_R e_q(R/2)\left(-2i\sin\frac{\pi R}{q}\right) \\ &= \prod_R e_q(R/2)(e_q(-R/2) - e_q(R/2)) \\ &= \prod_R (1 - e_q(R)), \end{aligned} \tag{27}$$

and, similarly,

$$e_q\left(\frac{q(q-1)}{8}\right)(-2i)^{(q-1)/2}\prod_N \sin\frac{\pi R}{q} = \prod_N (1 - e_q(N)). \tag{28}$$

Dividing Equation 28 by Equation 27, we arrive at

$$Q = \frac{\prod_N (1 - e_q(N))}{\prod_R (1 - e_q(R))}. \tag{29}$$

We now appeal to the result Gauss proved: Given an indeterminant x, there exist polynomials $Y(x), Z(x) \in \mathbb{Z}[x]$ such that

$$\prod_R (x - e_q(R)) = \frac{Y(x) - \sqrt{q}Z(x)}{2}; \tag{30}$$

$$\prod_N (x - e_q(N)) = \frac{Y(x) + \sqrt{q}Z(x)}{2}. \tag{31}$$

Evaluating these expressions at $x = 1$ and substituting the results into Equation 29 yields

$$Q = \frac{Y + q^{1/2}Z}{Y - q1/2Z},$$

where $Y = Y(1)$ and $Z = Z(1)$. Since Y and Z are real numbers (in fact, integers!), having $Z \neq 0$ would imply that $Q \neq 1$, as we need.

[14] See Appendix A.3.3 to read the rules for computing Legendre symbols

To prove $Z \neq 1$, multiply Equations 30 and 31 to obtain

$$\frac{Y^2(x) - qZ^2(x)}{4} = \left(\prod_R (x - e_q(R))\right)\left(\prod_N (x - e_q(N))\right),$$

which is simply a product of $x - e_q(m)$ over all of the q-th roots of unity $e_q(m)$ which are not equal to 1. Thus it is further true that

$$\frac{Y^2(x) - qZ^2(x)}{2} = \prod_{m=1}^{q-1} (x - e_q(m)) = \frac{x^q - 1}{x - 1} = \sum_{m=0}^{q-1} x^m.$$

This middle step follows because each $e_q(m) \neq 1$ and is a q-th root of unity, and the last step by recognizing this as the sum of a partial geometric sum.

Evaluating this expression involving $Y(x)$ and $Z(x)$ at $x = 1$, we arrive at $Y^2 - qZ^2 = 4q$. Now q is prime, which means $4q$ is not a perfect square, which in turn means $Z \neq 0$, as required! □

These last two lemmas prove that $L_{-1}(1) \neq 0$ in the case where $q \equiv_4 1$, and Lemma 24 proves that $L_{-1}(1) \neq 0$ in the case where $q \equiv_4 3$. Combining this with our proof of the fact that $L_\omega(s) \neq 0$ for all complex ω, we conclude that $L_\omega(1) \neq 0$ for all $\omega \neq 1$. Hence, the hypotheses of Lemma 20 are satisfied, from which we deduce Dirichlet's theorem in the case where q is a prime.

3.1.4 Dirichlet characters and L-functions (general modulus)

We now move on to the proof of Dirichlet's theorem for a general q. Our first step, as in the proof of the special case will be to define the Dirichlet characters for the modulus q. We will define the characters for powers of primes in two steps, then we will define the characters for a general modulus.

Recall that a Dirichlet character, which we will now denote $\chi(n)$ is a complex-valued function defined on the positive integers such that

1. $\chi(n + q) = \chi(n)$ for all $n \in \mathbb{N}$,
2. $\chi(mn) = \chi(m)\chi(n)$ for all $m, n \in \mathbb{N}$, and
3. there exists a linear combination of them whose value is 1 if $n \equiv_q a$ and 0 otherwise.

(Note: requirement 3 is usually omitted, but it will be the key property of our character construction that will allow us to prove Dirichlet's theorem.)

To begin the construction, consider the multiplicative groups of units in $\mathbb{Z}_q$, which we denote by

$$\mathbb{Z}_q^* = \{n \in \mathbb{Z}_q : (n, q) = 1\}.$$

We now consider two cases.

Suppose $q = p^\alpha$ where p is an odd prime, and $\alpha \in \mathbb{N}$. The construction in this case is parallel to that where q is prime. As we show in Appendix A.3.1, $\mathbb{Z}_q^*$ is cyclic. Let g be a generator of $\mathbb{Z}_q^*$, and define $\nu(n)$ to be the index of n relative to g – that is, define $\nu(n)$ such that $g^{\nu(n)} = n$. Let ω be any $\phi(q)$-th root of unity, where $\phi(q)$ is the order of the group $\mathbb{Z}_q^*$ and is called the Euler totient function. Then the Dirichlet character corresponding to ω is $\chi(n) = \omega^{\nu(n)}$ for all n such that $(n, q) = 1$ and $\chi(n) = 0$ otherwise. (In the case where q is prime, we have $\chi(n) = \omega^{\nu(n)}$ when

$q \mid n$ and $\chi(n) = 0$ otherwise; this is exactly how we defined $\chi(n)$ in that case.) To complete the definition, we extend $\chi(n)$ periodically so that it is defined for all positive integers.

That $\chi(n)$ is well-defined, q-periodic, and completely multiplicative follows by arguments identical to those given for the case where q was prime; one simply uses $\phi(q)$ in place of $q-1$. That leaves us to show that $\chi(n)$ satisfy property 3, but we will do this for every modulus at the end. We now move on to the case where q is a power of 2.

Suppose $q = 2^\alpha$ where $\alpha \in \mathbb{N}$. For $\alpha = 1, 2$, the construction is similar. For $\alpha = 1$, $q = 2$, which means there is $\phi(2) = 1$ character, namely,

$$\chi(n) = \begin{cases} 1, & \text{if } (n,2) = 1; \\ 0, & \text{otherwise.} \end{cases}$$

For $\alpha = 2$, $q = 4$, and there are $\phi(q) = 2$ characters, the first of which is 1 when $(n,4) = 1$ and 0 otherwise, and the other of which is 1 when $n \equiv_4 1$, -1 when $n \equiv_4 3$, and 0 otherwise. (You can easily see this! In the notation above, the generator g is 3, which means $\nu(1) = 2$ and $\nu(3) = 1$. The square roots of unity are 1 and -1, from which you can obtain the two characters.)

For $\alpha \geq 3$, the situation is not quite as nice, because $\mathbb{Z}_{2^\alpha}^*$ is not cyclic! What is true, however, is that the subgroup generated by 5 and -1 is the entire group! (See Appendix A.3.2 for a proof of this fact.) That is,

$$\mathbb{Z}_{2^\alpha}^* = \{(-1)^a 5^b : a = 0, 1;\ b = 0, 1, \ldots, 2^{\alpha-2} - 1\}.$$

Because there are $2 \cdot 2^{\alpha-2} = 2^{\alpha-1} = \phi(2^\alpha)$ elements in the set generated by powers of -1 and 5, we have that, for all $n \in \mathbb{Z}_{2^\alpha}^*$, there exists a unique pair $(\nu_0(n), \nu_0'(n))$ such that $n = (-1)^{\nu_0} 5^{\nu_0'}$. Using $(\nu_0(n), \nu_0'(n))$ as our substitute for an index $\nu(n)$ relative to a generating element of $\mathbb{Z}_{2^\alpha}^*$, we define the Dirichlet characters by

$$\chi(n) = \begin{cases} \omega^{\nu_0(n)} \left(\omega'\right)^{\nu_0'(n)}, & \text{if } (n, 2^\alpha) = 1; \\ 0, & \text{otherwise.} \end{cases}$$

where ω is a square root of unity and ω' is a $(2^{\alpha-2})$-th root of unity.

Here is an example set of characters for the modulus 8.

Example 28. *Let $q = 2^3$. Then $\mathbb{Z}_q^* = \{1, 3, 5, 7\}$. Since every element in this group squares to 1, it is clearly not cyclic. However, we note that $5^0 \equiv 1$, $5^1 \equiv 5$, $(-1)5^0 \equiv 7$, and $(-1)5^1 \equiv 3$, so $(\nu_0(n), \nu_0'(n))$ for $n = 1, 3, 5, 7$ is $(0,0), (1,1), (0,1), (-1,0)$, respectively.*

Now the four characters will correspond to the four combinations of a square root ω of unity and a 2^{3-2}-th root (or square root) ω' of unity. That is,

$$\begin{aligned} \chi_0(n) &= (1)^{\nu_0(n)}(1)^{\nu_0'(n)} = 1 \\ \chi_1(n) &= (1)^{\nu_0(n)}(-1)^{\nu_0'(n)} = (-1)^{\nu_0'(n)} \\ \chi_2(n) &= (-1)^{\nu_0(n)}(1)^{\nu_0'(n)} = (-1)^{\nu_0(n)} \\ \chi_3(n) &= (-1)^{\nu_0(n)}(-1)^{\nu_0'(n)} = (-1)^{\nu_0(n)+\nu_0'(n)}. \end{aligned}$$

The values of these characters are given here:

	1	3	5	7
χ_0	1	1	1	1
χ_1	1	-1	-1	1
χ_2	1	-1	1	-1
χ_3	1	1	-1	-1

Interestingly, you might notice that $\chi_0\chi_k = \chi_k$ for all k, that $\chi_3 = \chi_1\chi_2$, that $\chi_k^2 = \chi_0$ for all k, and so on. Hence the set of characters forms an abelian group in which every element has order 2. Since there only one such group up to isomorphism, the set of characters is isomorphic to the group $\mathbb{Z}_{2^3}^*$.

Another interesting note: This group of characters has order 4 just as the group of characters defined in Example 28, but they are not isomorphic groups. The group in Example 28, for example, is cyclic whereas the group in this example is not.

Proving that each $\chi(n)$ is well-defined, completely multiplicative, and q-periodic is an easy extension of the argument for the case where q is prime, so we omit it.

We have defined characters now for all prime powers. Given that we want our general characters to be multiplicative, this will be sufficient to define them in general. We do this now.

Suppose q is composite. In fact, since the following will hold for prime q as well, we could in fact take the definition that follows as the definition of Dirichlet characters for any modulus q. Let $q = 2^\alpha \prod_{k=1}^n p_k^{\alpha_k}$. Let $\chi(n; p^\beta)$ denote one of the $\phi(p^\beta)$ Dirichlet characters modulo p^β, as we have defined above. For each combination of these characters defined for a prime power modulus, define a Dirichlet character for all $n \in \mathbb{N}$ by

$$\chi(n) = \chi(n; 2^\alpha) \prod_{k=1}^n \chi(n; p_k^{\alpha_k}).$$

Note that $\chi(n) = 0$ whenever $(n, q) > 1$ because, in this case, $(n, p^\beta) > 1$ for some prime factor p of n, which implies that $\chi(n; p^\beta) = 0$.

Does this collection of $\chi(n)$ give us what we want? Let us see. First, χ is well-defined because each $\chi(n; p^\beta)$ is well-defined. Second, for all p^β which divide q, the fact that $\chi(n; p^\beta)$ is p^β-periodic implies that it is also q-periodic; hence, since each term in the product is q-periodic, $\chi(n)$ is q-periodic. Third, each $\chi(n; p^\beta)$ in the product is completely multiplicative, so the product $\chi(n)$ is multiplicative:

$$\begin{aligned}
\chi(mn) &= \chi(mn; 2^\alpha) \prod_{k=1}^n \chi(mn; p_k^{\alpha_k}) \\
&= \chi(m; 2^\alpha)\chi(n; 2^\alpha) \prod_{k=1}^n \left(\chi(m; p_k^{\alpha_k})\chi(n; p_k^{\alpha_k})\right) \\
&= \chi(m)\chi(n).
\end{aligned}$$

Thus the collection of the $\chi(n)$ characters are q-periodic completely multiplicative complex-valued functions defined on $\mathbb{N}$. All that is left to show is that some linear combination of these $\chi(n)$ outputs 1 when $n \equiv_q a$ and outputs 0 otherwise. We prove this in a sequence of steps.

Lemma 29. *There are $\phi(q)$ characters modulo q.*

Proof. In Section 3.1.2, we proved that there are $q - 1 = \phi(q)$ characters of a prime modulus q. In an identical fashion, we could prove that there are $\phi(q)$ characters of prime power modulus q. Now in the general case, we let $q = 2^\alpha p_1^{\alpha_1} \cdots p_n^{\alpha_n}$, and let

$$\chi(n) = \chi(n; 2^\alpha)\chi(n; p_1^{\alpha_1}) \cdots \chi(n; p_n^{\alpha_n}),$$

where $\chi(n;2^\alpha)$ is one of $\phi(2^\alpha)$ characters of modulus 2^α, $\chi(n;p_1^{\alpha_1})$ is one of $\phi(p_1^{\alpha_1})$ characters of modulus $p_1^{\alpha_1}$, and so on. This gives us

$$\phi(2^\alpha)\phi(p_1^{\alpha_1})\cdots\phi(p_n^{\alpha_n}) = \phi(2^\alpha p_1^{\alpha_1}\cdots p_n^{\alpha_n}) = \phi(q)$$

choices. □

We now construct a linear combination which will almost be what we need.

Lemma 30. *Consider the linear combination $\ell(n) = \sum_\chi \overline{\chi(a)}\chi(n)$, which is a sum over every Dirichlet character of modulus q. Then $\ell(n) = \phi(q)$ if $n \equiv_q a$ and $\ell(n) = 0$ otherwise.*

Proof. Suppose first that $n \equiv_q a$. Then by the q-periodicity of $\chi(n)$, $\chi(n) = \chi(a)$, which implies

$$\ell(n) = \sum_\chi \overline{\chi(a)}\chi(a) = \sum_\chi |\chi(a)|^2.$$

Since $\chi(a)$ is some power of a root of unity, $|\chi(a)| = 1$. Hence, since there are $\phi(q)$ Dirichlet characters, $\ell(n) = \phi(q)$.

Now suppose that $n \not\equiv_q a$. Let $q = p_0^{\alpha_0}p_1^{\alpha_1}\cdots p_n^{\alpha_n}$, where $p_0 = 2$. Then for some $k = 0,\ldots,n$, $p_k^{\alpha_k} \nmid (n-a)$. Now the sum over all characters χ of modulus q can be rewritten as $n+1$ sums over all of the characters χ_k, say, of modulus $p_k^{\alpha_k}$:

$$\ell(n) = \sum_{\chi_0}\cdots\sum_{\chi_n}\overline{\chi(a)}\chi(n).$$

Now we evaluate the sum over χ_k, because we know that $p_k^{\alpha_k} \nmid (n-a)$. Considering just this part, we have

$$\sum_{\chi_k}\overline{\chi(a)}\chi(n) = \prod_{j\neq k}\overline{\chi_k(a)}\prod_{j\neq k}\chi_k(n)\sum_{\chi_k}\overline{\chi_k(a)}\chi_k(n).$$

Now for $k > 0$, let ω be one of the $\phi(p_k^{\alpha_k})$-th roots of unity. Since $(a,q) = 1$, $a \in \mathbb{Z}_{p_k^{\alpha_k}}$, and hence

$$\overline{\chi_k(a)}\chi_k(n) = \overline{\omega^{\nu(a)}}\omega^{\nu(n)} = \omega^{-\nu(a)+\nu(n)} = \omega^{\nu(a^{-1})+\nu(n)} = \omega^{\nu(a^{-1}n)}.$$

Therefore, since the sum over the characters χ_k of modulus $p_k^{\alpha_k}$ is sum over the $\phi(q)$-th roots of unity, we have

$$\sum_{\chi_k}\overline{\chi_k(a)}\chi_k(n) = \sum_{\chi_k}\omega^{\nu(a^{-1}n)} = \sum_\omega \omega^{\nu(a^{-1}n)}.$$

Now $a \not\equiv_q n$, so $a^{-1}n \not\equiv_q 1$, which implies $\nu(a^{-1}n) \neq 0$, which in turn implies $\omega^{\nu(a^{-1}n)}$ is a $\phi(q)$-th root of unity besides 1. Hence the sum on the right-hand side is 0. Summing 0 over the other characters χ_j for $j \neq k$ gives us that $\ell(n) = 0$. □

Thus, by dividing $\ell(n)$ by $\phi(q)$, we obtain the required linear combination:

$$\frac{1}{\phi(q)}\sum_\chi \overline{\chi(a)}\chi(n) = \begin{cases} 1, & \text{if } n \equiv_q a; \\ 0, & \text{otherwise.} \end{cases}$$

At this point, we are ready to move on with our proof of Dirichlet's theorem. However, we have a closely related fact about sums of characters which we will use later. So we prove it here.

Theorem 31. *For any q consecutive natural numbers indexed by n,*

$$\sum_n \chi(n) = 0.$$

Proof. First, by the q-periodicity of $\chi(n)$, it clearly suffices to show that $\sum_{n=0}^{q-1} \chi(n) = 0$. Let $q = 2^\alpha p_1^{\alpha_1} \cdots p_r^{\alpha_r}$. Consider the definition of the Dirichlet character:

$$\chi(n) = \chi(n; 2^\alpha)\chi(n; p_1^{\alpha_1}) \cdots \chi(n; p_r^{\alpha_r}).$$

Each $\chi(n; p^\beta)$ corresponds to a unique $(r+2)$-tuple $(\omega_0, \omega_0', \omega_1, \ldots, \omega_r)$ where $\omega_0^2 = 1$, $(\omega_0')^{\phi(2^{\alpha-1})} = 1$, and $\omega_k^{\phi(p_k^{\alpha_k})} = 1$ for all $k = 1, \ldots, r$. For each of these roots of unity, there will exist a unique power of a given primitive root of unity that will equal the original root. Hence this $(r+2)$-tuple corresponds to a unique $(m+2)$-tuple $(m_0, m_0', m_1, \ldots, m_r) \in A$ where[15]

$$A = \mathbb{Z}_2 \times \mathbb{Z}_{\phi(2^{\alpha-1})} \times \mathbb{Z}_{\phi(p_1^{\alpha_1})} \times \cdots \times \mathbb{Z}_{\phi(p_r^{\alpha_r})},$$

$\left(e^{2\pi i/2}\right)^{m_0} = \omega_0$, $\left(e^{2\pi i/\phi(2^{\alpha-1})}\right)^{m_0'} = \omega_0'$, and $\left(e^{2\pi i/\phi(p_k^{\alpha_k})}\right)^{m_k} = \omega_k$ for all $k = 1, \ldots, r$.

Hence, using this complex-exponential representation,

$$\chi(n; 2^\alpha) = \omega_0^{\nu_0(n)}(\omega_0')^{\nu_0'(n)} = \exp\left(2\pi i\left(\frac{m_0\nu_0(n)}{2} + \frac{m_0'\nu_0'(n)}{\phi(2^{\alpha-1})}\right)\right)$$

and

$$\chi(n; p_1^{\alpha_1}) = \omega_1^{\nu_1(n)} = \exp\left(2\pi i \frac{m_k\nu_k(n)}{\phi(p_k^{\alpha_k})}\right)$$

for all $k = 1, \ldots, r$. So

$$\chi(n) = \exp\left(2\pi i\left(\frac{m_0\nu_0(n)}{2} + \frac{m_0'\nu_0'(n)}{\phi(2^{\alpha-1})} + \sum_{k=1}^{r} \frac{m_k\nu_k(n)}{\phi(p_k^{\alpha_k})}\right)\right).$$

Now what happens when we sum over $n \in \mathbb{Z}_q$ (where again, $\mathbb{Z}_q$ here denotes nothing more than the set $\{0, 1, \ldots, q-1\}$)?

Now recall that $\nu_k(n)$, for each $k = 1, \ldots, r$, is the unique element modulo $\phi(p_k^{\alpha_k})$ such that $g_k^{\nu_k(n)} \equiv n$ modulo $p_k^{\alpha_k}$, where g_k is a generator of $\mathbb{Z}^*_{p_k^{\alpha_k}}$ which we fixed (way back) when we defined $\chi(n; p_k^{\alpha_k})$. Also, $\nu_0(n)$ and $\nu_0'(n)$ are the unique elements modulo 2 and $\phi(2^{\alpha-1})$ such that $(-1)^{\nu_0(n)}5^{\nu_0'(n)} \equiv n$ modulo 2^α. So by the definition of the index functions $\nu_0(n)$, $\nu_0'(n)$, and so on, $(\nu_0(n), \nu_0'(n))$ is a bijection from $\mathbb{Z}^*_{2^\alpha}$ onto $\mathbb{Z}_2 \times \mathbb{Z}_{\phi(2^{\alpha-1})}$, and $\nu_k(n)$ is a bijection from $\mathbb{Z}_{p_k^{\alpha_k}}$ onto $\mathbb{Z}_{\phi(p_k^{\alpha_k})}$ for all $k = 1, \ldots, r$. Hence the function $\vec{\nu}(n)$, defined by

$$\vec{\nu}(n) = (\nu_0(n), \nu_0'(n), \nu_1(n), \ldots, \nu_r(n))$$

is a bijection from

$$\mathbb{Z}_q^* \cong \mathbb{Z}^*_{2^\alpha} \oplus \mathbb{Z}^*_{p_1^{\alpha_1}} \oplus \cdots \oplus \mathbb{Z}^*_{p_r^{\alpha_r}}$$

onto

$$A = \mathbb{Z}_2 \times \mathbb{Z}_{\phi(2^{\alpha-1})} \times \mathbb{Z}_{\phi(p_1^{\alpha_1})} \times \cdots \times \mathbb{Z}_{\phi(p_r^{\alpha_r})}.$$

[15]Here we are abusing the notation slightly. The sets of the form $\mathbb{Z}_n$ in the Cartesian product for A are meant here to specify the set $\{0, 1, \ldots, n-1\}$ and nothing more.

(We have used $\oplus$ to denote the direct product of groups, as opposed to $\times$, which we used to denote the Cartesian products of sets.)

Therefore, the sum over n involving $\vec{\nu}(n)$ can be rewritten as a sum over the values $\vec{\nu} = (\nu_0, \nu_0', \nu_1, \ldots, \nu_r)$ in the range of $\vec{\nu}(n)$. That is,

$$\sum_{n=0}^{q-1} \chi(n) = \sum_{\vec{\nu} \in A} \chi_{\vec{\nu}}$$

where

$$\chi_{\vec{\nu}} = \exp\left(2\pi i \left(\frac{m_0\nu_0}{2} + \frac{m_0'\nu_0'}{\phi(2^{\alpha-1})} + \sum_{k=1}^{r} \frac{m_k\nu_k}{\phi(p_k^{\alpha_k})}\right)\right). \tag{32}$$

Next, we reindex our sum once more. For clarity, denote $\chi_{\vec{\nu}}$ by $\chi_{\vec{\nu},\vec{m}}$ where $\vec{m} = (m_0, m_0', m_1, \ldots, m_r)$. What we have is a fixed $\vec{m}$, and we are summing over $\vec{\nu}$. But, as is clear from Equation 32, $\chi_{\vec{\nu},\vec{m}} = \chi_{\vec{m},\vec{\nu}}$. So if we swap the the values of $\vec{\nu}$ and $\vec{m}$, we obtain

$$\sum_{\vec{\nu} \in A} \chi_{\vec{\nu}} = \sum_{\vec{\nu} \in A} \chi_{\vec{\nu},\vec{m}} = \sum_{\vec{\nu} \in A} \chi_{\vec{m},\vec{\nu}} = \sum_{\vec{m} \in A} \chi_{\vec{\nu},\vec{m}}.$$

We have (seemingly magically) turned this sum into one over the $(r+2)$-tuples $\vec{m} \in A$. But the set of $\vec{m}$ is in one-to-one correspondence with the set of characters, as mentioned above. Hence this last sum is one over all possible characters χ. The terms in the sum are obtained by choosing (the unique) $n' \in \{0, 1, \ldots, q-1\}$ such that $\nu(n') = \vec{\nu}$. So

$$\sum_{\vec{m} \in A} \chi_{\vec{\nu},\vec{m}} = \sum_{\chi} \chi(n') = \sum_{\chi} \overline{\chi(1)}\chi(n').$$

But if we apply Lemma 30 where $a = 1$, then, since $n \not\equiv_q 1$, the sum is 0, from which the theorem follows. □

The next step in the proof is just like the step we made in the proof of the special case. Refer to Section 3.1.4 to see how our definitions and various steps relate to those made in the proof of the special case.

We begin by defining the Dirichlet L-functions for our general modulus.

Definition 32. *For $q \in \mathbb{N}$ and a character $\chi(n)$ of modulus q, define $L(s, \chi)$ for all $s \in \mathbb{C}$ by*

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

By Theorem 1, we have

$$L(s, \chi) = \prod_{p} \frac{1}{1 - \chi(p)p^{-s}}$$

for all $\Re(s) > 1$. Now $L(s, \chi) \neq 0$, as we could see by considering the product formula for $L(s, \chi)$ (which we did in Section 3.1.4). Therefore, by a calculation similar to the one in Section 3.1.4, this time using Lemma 30, we obtain

$$\begin{aligned} \log L(s, \chi) &= -\sum_{p} \log(1 - \chi(p)p^{-s}) \\ &= \sum_{p} \sum_{m=1}^{\infty} m^{-1}\chi(p^m)p^{-ms}. \end{aligned}$$

Next, we multiply by $\overline{\chi(a)}$ and sum over all characters χ, arriving at

$$\begin{aligned}\frac{1}{\phi(q)}\sum_{\chi}\overline{\chi(a)}\log L(s,\chi) &= \sum_{p}\sum_{m=1}^{\infty}\left(\sum_{\chi}\overline{\chi(a)}\chi(p^m)\right)p^{-ms} \\ &= \sum_{p}\sum_{m} m^{-1}p^{-ms}, \end{aligned} \tag{33}$$

where the sum over m runs over $m \in \mathbb{N}$ such that $p^m \equiv_q a$.

On the right-hand side, we can pull out the $m = 1$ term to obtain

$$\sum_{p\equiv_q a}\frac{1}{p^s} + \delta, \tag{34}$$

where $|\delta| < \zeta(2)$ for all $s > 1$, in the same way as before. On the left-hand side, consider the $\chi = \chi_0$ term, where χ_0 is the character corresponding $\omega = 1$. We have $\chi_0(n) = 1$ for all $(n, q) = 1$ and $\chi_0(n) = 0$ otherwise, which implies that

$$\prod_{p}\frac{1}{1-\chi_0(p)p^{-s}} = \prod_{p\nmid q}\frac{1}{1-p^{-s}} = \zeta(s)\prod_{p|q}(1-p^{-s}).$$

Since the finite product remains a finite quantity as $s \to 1^+$, this diverges to ∞ because $\zeta(s) \to \infty$ as $s \to 1^+$. Thus the $\chi = \chi_0$ term on the left-hand side diverges to ∞ as $s \to 1^+$.

If the entire left-side converges to infinity as well, then the sum $\sum_{p\equiv_q a} p^{-1}$ must necessarily diverge, which would imply the existence of infinitely many primes congruent to a modulo q. Dirichlet's theorem will be proved! All we need to show is that the other terms (the $\chi \neq \chi_0$ terms) are bounded, which is to say that $\lim_{s\to 1^+} L(s,\chi) \in (0,\infty)$. Using continuity arguments similar to those in the proof of the special case, we see that the limit is bounded and, in particular, equal to $L(1,\chi)$. Hence our proof requires only that

$$L(s,\chi) \neq 0 \text{ for all } \chi \neq \chi_0.$$

Proving this is the task of the next section.

3.1.5 Dirichlet's proof (general modulus)

As in the proof of the case where q is prime, we split our task into two subtasks.

Suppose χ is a complex character. The first is to suppose χ is a complex character; that is, suppose that $\chi(n)$ is complex-valued for some $n \in \mathbb{N}$. Then, as in the proof of the special case, we consider Equation 33, evaluated at $a = 1$. (This equation holds for all a which are relatively prime to q; we are trying to prove a property about $L(s,\chi)$, so the fact that we are setting $a = 1$ does not somehow restrict ourselves to a special case.) This gives us an analogue to Equation 17. The right-hand side is nonnegative, so we have

$$\sum_{\chi}\log L(s,\chi) \geq 0,$$

from which we conclude that that

$$\prod_{\chi} L(s,\chi) \geq 1.$$

The rest of the proof continues similarly.

Suppose χ is a real character The case where χ is a real character (besides the principle character χ_0) is challenging (as it was in the special case where q was prime). We wish to prove that $L(1,\chi) \neq 0$. To do this, we extend the definition of $L(s,\chi)$ to all $s \in \mathbb{C}$. Throughout the rest of this section, let s denote an arbitrary complex number with real part σ and imaginary part t, so that $s = \sigma + it$. By Theorem 1, $L(s,\chi)$ converges absolutely for all $\sigma > 1$. We now note a couple of properties about the function $L(s,\chi)$.

First, for all $\delta > 0$, the series $L(s,\chi)$ converges uniformly for $\sigma \geq \delta$. The proof of this is just like the proof of the fact that $\zeta(s)$ converges uniformly (see proof of Lemma 8). Since the partial sums of the series for $L(s,\chi)$ are holomorphic functions, and since we could (again, easily) show that the sequence of derivatives of the partial sums converge uniformly, we have the following result:

Lemma 33. *$L(s,\chi)$ is holomorphic for $\sigma > 1$.*

Using this corollary, we prove a crucial property of $L(s,\chi)$. But first, we make a definition.

Definition 34. *Let V and $U \subseteq V$ be subsets of $\mathbb{C}$. Let $f : U \to \mathbb{C}$ be a holomorphic function and $F : V \to \mathbb{C}$ be a holomorphic function such that $F(s) = f(s)$ for all $s \in U$. Then F is an* holomorphic continuation *of f to V.*

With this terminology, we state the following important property:

Theorem 35. *$L(s,\chi)$ can be holomorphically continued to $\sigma > 0$.*

Before we give the proof, we note that this theorem holds for $\chi = \chi_0$ as well, except that $L(s,\chi_0)$ has a simple pole at $s = 1$. The proof, which is identical to the one below except for a step at the end, might provide an exercise for the reader.

Proof. Let $X(n) = \sum_{m=1}^{n} \chi(m)$. More generally, define $X(x) = \sum_{m \leq x} \chi(m)$, where the sum over m runs over all positive integers less than or equal to x. Then, by Abel's Lemma (see Appendix A.4.3),

$$\begin{aligned} L(s,\chi) &= \lim_{N\to\infty} \sum_{n=1}^{N} \chi(n)\frac{1}{n^s} \\ &= \lim_{N\to\infty} \left(\sum_{n=1}^{N-1} X(n)\left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right) + \chi(N)\frac{1}{N^s} \right) \\ &= \sum_{n=1}^{\infty} X(n)\left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right). \end{aligned}$$

Next, we consider the fact that $X(x) = X(n)$ for all $x \in (n, n+1)$. This allows us to write

$$L(s,\chi) = \sum_{n=1}^{\infty} \int_n^{n+1} \frac{sX(x)}{x^{s+1}} dx = s\int_1^{\infty} \frac{X(x)}{x^{s+1}} dx.$$

Now by Theorem 31,

$$0 = \sum_{n=0}^{q-1} \chi(n) = \sum_{n=q}^{2q-1} \chi(n) = \cdots.$$

Hence $X(x)$ is a bounded function. If $M > 0$ such that $|X(x)| < M$ for all $x > 1$, then $\sigma > 0$ implies that

$$\int_1^{\infty} \left|\frac{X(x)}{x^{s+1}}\right| dx \leq M\int_1^{\infty} \frac{dx}{x^{\sigma+1}} = \frac{M}{\sigma},$$

which means this integral converges and takes on a finite value for all $\sigma > 0$. Furthermore, using the Leibniz integral rule,

$$\frac{d}{ds}L(s,\chi) = \int_1^\infty \frac{X(x)}{x^{s+1}}dx - s(s+1)\int_1^\infty \frac{X(x)}{x^{s+2}}dx.$$

This is clearly convergent for $\sigma > 0$, so $L(s,\chi)$ is holomorphic for $\sigma > 0$, as claimed. □

We are almost there! Here is one more quick lemma:

Lemma 36. *$L(s,\chi_0)$ has simple pole at $s = 1$.*

Proof. We showed in the proof of Lemma 11 that $(s-1)\zeta(s) \to 1$ as $s \to 1$. (Specifically, we knew that the limit in Equation 6 was finite, from which we concluded that $(s-1)\zeta(s) - 1$ converged to 0 as $s \to 1$.) Hence $\zeta(s)$ has a simple pole at 1. But we can write

$$L(s,\chi_0) = \prod_p \frac{1}{1-\chi_0(p)p^{-s}} = \prod_{p\nmid q} \frac{1}{1-p^{-s}} = \zeta(s)\prod_{p|q}(1-p^{-s}).$$

Since only finitely many primes divide q, the product on the right will be finite if and only if $\zeta(s)$ is. Hence $L(s,\chi_0) \to \infty$ as $s \to 1^+$ because $\zeta(s) \to \infty$ in the same limit, and $(s-1)L(s,\chi_0)$ remains finite as $s \to 1^+$ because $(s-1)\zeta(s)$ does. Hence $L(s,\chi_0)$ has a simple pole at $s = 1$. □

The proof of the theorem – which we have reduced to the claim that $L(1,\chi) \neq 0$ for all real $\chi \neq \chi_0$ – is in two steps. We define a function $\psi(s)$ on $\mathbb{C}$ and prove two properties of it in the following two lemmas. Dirichlet's theorem will follow immediately, as we show below.

Lemma 37 (The first $\psi(s)$ lemma). *Define the function $\psi(s)$ for all $s \in \mathbb{C}$ by*

$$\psi(s) = \frac{L(s,\chi)L(s,\chi_0)}{L(2s,\chi_0)}.$$

If $L(1,\chi) = 0$, then $\psi(s) \to 0$ as $s \to \frac{1}{2}^+$.

Proof. By assumption, $L(s,\chi)$ has a zero at $s = 1$. So by Lemma 36, $L(s,\chi)L(s,\chi_0)$ is holomorphic at $s = 1$. By Theorem 35, $L(s,\chi)L(s,\chi_0)$ is, in fact, holomorphic for $\sigma > 0$. Define $\psi(s)$ for $s \in \mathbb{C}$ by

$$\psi(s) = \frac{L(s,\chi)L(s,\chi_0)}{L(2s,\chi_0)}.$$

Note that $L(2s,\chi_0)$ is holomorphic for all $\sigma > \frac{1}{2}$. So $\psi(s)$ is holomorphic for all $\sigma > \frac{1}{2}$. Also, because $L(s,\chi)L(s,\chi_0)$ is holomorphic at $\frac{1}{2}$, it is bounded at that point; hence, since $L(s,\chi_0) \to \infty$ as $s \to \frac{1}{2}$ from the right, we have that $\psi(s) \to 0$ as $s \to \frac{1}{2}$ from the right. □

Lemma 38 (The second $\psi(s)$ lemma). *Given $\psi(s)$ as defined in the previous lemma, $\psi(s) \geq 1$ for all $\frac{1}{2} < s < 2$.*

Proof. As we have seen previously, we can write

$$L(s,\chi_0) = \prod_{p\nmid q} \frac{1}{1-p^{-s}}.$$

So

$$L(2s, \chi_0) = \prod_{p \nmid q} \frac{1}{1 - p^{-2s}}.$$

Now consider the product for $L(s, \chi) = \prod_p (1 - \chi(p)p^{-s})^{-1}$. If $p \mid q$, then $\chi(p) = 0$ (just as it is in the product representation for $L(s, \chi)$), so

$$L(s, \chi) = \prod_{p \nmid q} \frac{1}{1 - \chi(p)p^{-s}}.$$

Hence $\psi(s)$ has the following product representation:

$$\psi(s) = \prod_{p \nmid q} \frac{1 - p^{-2s}}{(1 - \chi(p)p^{-s})(1 - p^{-s})}.$$

If $\chi(p) = -1$ for some p in this product, then the p-term is clearly 1. Hence we have

$$\psi(s) = \prod_{\chi(p)=1} \frac{1 - p^{-2s}}{(1 - p^{-s})^2} = \prod_{\chi(p)=1} \frac{1 + p^{-s}}{1 - p^{-s}},$$

where the product is understood to be over all prime p such that $\chi(p) = 1$. Now we recognize $(1 - p^{-s})^{-1}$ as the sum of a geometric series with ratio p^{-s}. Hence

$$\psi(s) = \prod_{\chi(p)=1} \left(1 + \frac{1}{p^s}\right) \left(\sum_{m=0}^{\infty} \left(\frac{1}{p}\right)^{ms}\right) = \prod_{\chi(p)=1} \left(1 + 2\sum_{m=1}^{\infty} \frac{1}{p^{ms}}\right). \tag{35}$$

If we wrote out this product as a series, every term would be of the form $a_n n^{-s}$ for some $a_n \in \mathbb{C}$ and $n \in \mathbb{N}$. (If there are no terms in this product, which could conceivably happen if $\chi(p) \neq 1$ for all prime $p \nmid q$, then the empty product would equal 1, in which case our claim holds trivially.) Let

$$\psi(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}. \tag{36}$$

By Equation 35, we see that $a_n \geq 0$ for all $n \in \mathbb{N}$ and that $a_1 = 1$.

Next, we consider another series representation of $\psi(s)$, the Taylor series representation. Since $L(s, \chi)L(s, \chi_0)$ is holomorphic for $\sigma > 0$ and $L(2s, \chi_0)$ is holomorphic for $\sigma > \frac{1}{2}$, their quotient (that is, $\psi(s)$) is holomorphic for $\sigma > \frac{1}{2}$. Hence if we center the Taylor series at 2, writing

$$\psi(s) = \sum_{m=0}^{\infty} \frac{\psi^{(m)}(2)}{m!}(s - 2)^m, \tag{37}$$

then the Taylor series converges for all $|s - 2| < \frac{3}{2}$. We now calculate $\psi^{(m)}(2)$ using the Dirichlet series, and substitute the result into the Taylor series, which will then provide us with the desired estimate.

Using Equation 36, we have

$$\begin{aligned}
\psi'(s) &= -s\sum_{n=1}^{\infty}\frac{a_n}{n^{s+1}};\\
\psi''(s) &= s(s+1)\sum_{n=1}^{\infty}\frac{a_n}{n^{s+2}}\\
&\vdots\\
\psi^{(m)}(s) &= (-1)^m s(s+1)\cdots(s+m-1)\sum_{m=1}^{\infty}\frac{a_n}{n^{s+m}},
\end{aligned}$$

as a simple induction argument would show. Hence

$$\psi^{(m)}(2) = (-1)^m(m+1)!\sum_{n=1}^{\infty}\frac{a_n}{n^{2+m}}.$$

If we let $b_m = (m+1)\sum_{n=1}^{\infty} a_n n^{-(2+m)}$, then $\psi^{(m)}(2) = (-1)^m m! b_m$. Substituting this into Equation 37, we have

$$\psi(s) = \sum_{m=0}^{\infty}(-1)^m b_m(s-2)^m = \sum_{m=0}^{\infty} b_m(2-s)^m.$$

Now the fact that $a_n \geq 0$ for all $n \in \mathbb{N}$ implies that $b_m \geq 0$ for all $m = 0, 1, \ldots$. Also, $b_0 = a_1 = 1$. Hence for all $s \in (\frac{1}{2}, 2)$, $(2-s)^m \geq 0$, from which we conclude

$$\psi(s) \geq \sum_{m=0}^{\infty} b_m(0)^m = b_0 = 1,$$

as desired. □

We have made it to the top of the mountain, and we are ready to make the last step. Let us do so!

Theorem 39 (Dirichlet's Theorem on Primes in Arithmetic Progressions). *Let $(a, q) = 1$. Then there exist infinitely many primes of the form $a + qn$ where $n \in \mathbb{N}$.*

Proof. By what we have shown, all we need to do is show that $L(1, \chi) \neq 0$ where $\chi \neq \chi_0$ is a real character. We proceed by contradiction. Assume $L(1, \chi) = 0$. Define the function $\psi(s)$ as in Lemmas 37 and 38. Then by Lemma 37, $\psi(s) \to 0$ as $s \to \frac{1}{2}^+$. But by Lemma 38, the limit of $\psi(s)$ as $s \to \frac{1}{2}^+$ – if it even exists – is greater than or equal to 1. Hence we have a contradiction, and the theorem is proved. □

To state it once more, every residue class modulo q, in which each element is relatively prime to q, has infinitely many primes.

3.1.6 A probabilistic interpretation of Dirichlet's theorem

Pick any pair of relatively prime integers a and q, say, $10^{100}+1$ and 10^{100}. We have just proven that there are infinitely many primes of the form

$$1+10^{100}, 1+2\cdot 10^{100}, 1+3\cdot 10^{100}, 1+4\cdot 10^{100}, \ldots.$$

This is amazing! The probability of randomly choose a prime between $n\cdot 10^{100}$ and $(n+1)\cdot 10^{100}$ is approximately

$$\begin{aligned}
\frac{\text{number of primes in range}}{\text{number of integers in range}} &\approx \frac{\pi((n+1)\cdot 10^{100}) - \pi(n\cdot 10^{100})}{10^{100}} \\
&= \frac{\frac{(n+1)\cdot 10^{100}}{\log((n+1)\cdot 10^{100})} + \frac{n\cdot 10^{100}}{\log(n\cdot 10^{100})}}{10^{100}} \\
&= \frac{(n+1)\log n - n\log(n+1)}{100\log 10\log n\log(n+1)} \\
&= \frac{(n\log n - n\log n + 1) + \log n}{100\log 10\log n\log(n+1)}.
\end{aligned}$$

This clearly goes to 0 as $n\to\infty$. So, in particular, the probability that $1+n\cdot 10^{100}$ is prime goes to 0. One interpretation of Dirichlet's theorem is that this probability does not go to 0 *too quickly*. For if it did, then after testing whether $1+n\cdot 10^{100}$ is prime for all $n\in\mathbb{N}$, the expected number of primes found might be finite. But Dirichlet's theorem implies that the expected number – which is equal to the actual number since what we have is really a deterministic space of outcomes – is infinite!

The preceding discussion is the same if we replace 10^{100} with any $q\in$ ıı (besides 1) and 1 with any a which is relatively prime to q. The purpose of choosing particular (large) values is to express how amazing this fact is.

We have proven that there exist infinitely many primes, and we have given an asymptotic expression for how dense they are in the $\mathbb{N}$. Now that we have proven that there are infinitely many primes in (certain) arithmetic progressions, we would like to give a related asymptotic expression for their density in $\mathbb{N}$.

3.2 The Generalized Prime Number Theorem

We have provided answers (with proof!) to three of the four questions posed in the introduction:

Q: How many primes are there?
A: Infinitely many! (Proofs due to Euclid and Euler.)
Q: How dense are they?
A: The PNT says $\pi(x)\sim x/\log x$.
Q: Given an arithmetic progression (with $(a,q)=1$), how many primes are there?
A: Infinitely many! (Proof due to Dirichlet.)

Our last question is: Given such an arithmetic progression, how dense in $\mathbb{N}$ are the primes in that progression? What function might tell us the approximate number of primes less than some fixed number in a given arithmetic progression? To begin, the actual value of the function must be bounded above by $\pi(x)$, for there cannot exist fewer primes less than x than primes of a particular kind less than x. Let $\pi_{a,q}(x)$ denote the number of primes less than or equal to x which are congruent to a modulo q. Now given an prime p, p corresponds to a unique residue class

modulo q. Since the p could only show up in a residue class whose members are relatively prime to q, we must have that

$$\pi(x) = \sum_{a \in \mathbb{Z}_q^*} \pi_{a,q}(x),$$

where the notation in the sum indicates a sum over all integers $0, 1, \ldots, q-1$ which are relatively prime to q.

Beyond this, we are left to conjecture: what is $\pi_{a,q}(x)$? One easy guess would be that $\pi_{1,q}(x) = \pi(x)$ and that $\pi_{a,q}(x) = 0$ for all $a \neq 1$. But this does not seem to be the case for small q. For example, when $q = 4$, there are primes congruent to 1 (such as 5, 13, and 17) as well as primes congruent to 3 (such as 7, 11, and 19). In fact, for small q, it appears that all residue classes contribute primes. Then next natural conjecture might be that each residue class contributes *equally*. Since $\pi_{a,q}(x)$, then, would not depend on a, we could write this function (approximately) as $\pi_q(x)$. Furthermore, since there $\phi(q)$ relatively prime residue classes modulo q, we would have that

$$\pi(x) = \sum_{a \in \mathbb{Z}_q^*} \pi_{a,q}(x) \sim \pi_q(x) \sum_{a \in \mathbb{Z}_q^*} 1 = \pi_q(x)\phi(q).$$

Hence our conjecture is that

$$\pi_q(x) \sim \frac{1}{\phi(q)}\pi(x) \sim \frac{1}{\phi(q)} \frac{x}{\log x}.$$

Because this *is* a statement of the Prime Number Theorem when $q = 2$, we call this (for our purposes) the generalized Prime Number Theorem, or gPNT.

The proof of the gPNT given in [So] is parallel – step for step – to Zagier's presentation of Newman's proof. I omit it here, and instead discuss why we have a good reason to be optimistic about our conjecture. We prove two lemmas here, then we prove a theorem, which follows trivially from the gPNT and does not quite prove the gPNT, but it has independent interest. This discussion follows that in [C].

Lemma 40. *As $s \to 1^+$,*

$$\left(\sum_p \frac{1}{p^s}\right) \cdot \left(\log \frac{1}{s-1}\right)^{-1} \to 1.$$

Proof. From Equation 13 that

$$\log \zeta(s) = \sum_p \frac{1}{p^s} + \delta(s),$$

where $|\delta(s)| < \zeta(2)$. We also know from the proof of Lemma 11 that $(s-1)\zeta(s) \to 1$ as $s \to 1^+$. Hence

$$\sum_p \frac{1}{p^s} = \log((s-1)\zeta(s)) + \log \frac{1}{s-1} - \delta(s) = \log \frac{1}{s-1} + \delta_1(s),$$

where $\delta_1(s)$ remains bounded as $s \to 1^+$. Dividing by $\log(s-1)^{-1}$ and taking $s \to 1^+$ proves the lemma. □

The next lemma is similar.

Lemma 41. *As $s \to 1^+$,*

$$\left(\sum_{p \equiv_q a} \frac{1}{p^s}\right) \cdot \left(\log \frac{1}{s-1}\right)^{-1} \to 1.$$

Proof. From Equations 33 and 34, we have that

$$\frac{1}{\phi(q)} \log L(s, \chi_0) + \frac{1}{\phi(q)} \sum_{\chi \neq \chi_0} \overline{\chi(a)} L(s, \chi) = \sum_{p \equiv_q a} \frac{1}{p^s} + \delta(s) \tag{38}$$

where again $\delta(s)$ remains bounded as $s \to \infty$. Now we proved that the sum over $\chi \neq \chi_0$ remains bounded as $s \to 1^+$ – that was the hard part about proving Dirichlet's theorem! Furthermore, we proved in Lemma 36 that $(s-1)L(s, \chi_0)$ is bounded as $s \to 1^+$. From our proof it is also clear that $(s-1)L(s, \chi_0)$ does not converge to 0 as $s \to 1^+$. Hence $\log((s-1)L(s, \chi_0))$ remains bounded as $s \to 1^+$. Rewriting Equation 38 and combining the terms which remain bounded as $s \to 1^+$, we obtain

$$\begin{aligned} \phi(q) \sum_{p \equiv_q a} \frac{1}{p^s} &= \log((s-1)L(s, \chi_0)) + \log \frac{1}{s-1} + \sum_{\chi \neq \chi_0} \overline{\chi(a)} L(s, \chi) - \delta(s) \\ &= \log \frac{1}{s-1} + \delta_2(s), \end{aligned}$$

where $\delta_2(s)$ remains bounded as $s \to \infty$. Dividing by $\log \frac{(}{s} - 1)^{-1}$ and taking $s \to 1^+$, we obtain our lemma. □

When we combine Lemmas 40 and 41, we obtain the following result:

Theorem 42. *For relatively prime integers a and q, the primes in the arithmetic progression $\{a + qn\}_{n=1}^{\infty}$ are evenly distributed in the following sense:*

$$\frac{\sum_{p \equiv_q a} p^{-s}}{\sum_p p^{-s}} \to \frac{1}{\phi(q)} \quad as \quad s \to 1^+. \tag{39}$$

Hence the primes really are, in some sense, evenly distributed among the residue classes.

A Appendix

Given the length of the main proofs in this exercise, I felt that it was appropriate to put off proving some of the facts I used until later. Thus we have this appendix. Some of the sections are purely technical, and others are definitional. There are some gems to be found here, however, such as the theorems proven in Sections 49 and A.4.4.

A.1 Complex Integration and Differentiation

In this section we give some definitions which are required to state two important theorems from complex integration theory, Cauchy's Residue Theorem and the Deformation Invariance Theorem, both of which we use earlier. The definitions and proofs can be found in [Sa]. Also, we discuss the methods by which one proves that a function is differentiable.

A.1.1 Complex Integration

A *domain* D is an open subset of $\mathbb{C}$ that is also connected. A *contour* C (for our purposes) is a continuous parameterizable curve in a domain D that is differentiable except possibly at finitely many points; that is, C is a curve for which there exists a continuous bijection $s(t)$ from $[0,1]$ onto D such that $s'(t)$ exists for all $t \in [0,1] \setminus A$ where A is a finite subset of $[0,1]$. Finally, a contour integral $\int_C f(s)ds$ is defined just as a line integral is in multivariable calculus; we let $s(t)$ be a parameterization of C and write

$$\int_C f(s)ds = \sum_{i=0}^{n} \int_{t_i}^{t_{i+1}} f(s(t))s'(t)dt$$

and integrate, where $t_0 = 0$ and $t_n = 1$, and where $t_1, \ldots, t_{n-1}$ are the points (if there are any) at which $s(t)$ is not differentiable.

Now line integrals over $\mathbb{R}^2$ and contour integrals over $\mathbb{C}$ are not entirely the same; in fact, our analogies end here: the key difference is related to the notion of homotopy. Two contours C_1 and C_2 are homotopic in a domain D if there exists a continuous function $s(s,t)$ such that

$$\begin{aligned} C_1 &= \{s(0,t) : t \in [0,1]\}, \\ C_2 &= \{s(1,t) : t \in [0,1]\}, \end{aligned}$$

and $s(s,t) \in D$ for all $(s,t) \in [0,1] \times [0,1]$. With this definition we have the following:

Theorem 43 (Deformation Invariance Theorem). *If $f(s)$ is holomorphic on a domain D, and if C_1 and C_2 are homotopic contours in D, then $\int_{C_1} f(s)ds = \in_{C_2} f(s)ds$.*

I omit the proof, as it is proven in most introductory courses on complex analysis, including the one taught at Kenyon.

Theorem 43 is so far from true for line integrals in $\mathbb{R}^2$, that it seems impossible that it is true for contour integrals in $\mathbb{C}$. The difference only becomes clearer as one learns more about integration over $\mathbb{C}$. Another amazing property of contour integrals in $\mathbb{C}$ is Cauchy's Residue Theorem. To state it, we need a few more definitions. A contour is *simple* if the parameterization $s(t)$ of it is one-to-one on $[0,1)$. A *closed* contour is one for which the ends of the contour meet: $s(0) = s(1)$. A simple closed contour, therefore, is a *loop* in the complex plane that does not cross itself. Finally, a *positively oriented* contour is less easily defined. For our purposes, such a contour is defined as

a simple closed contour such that $s(t)$ moves in a generally counterclockwise direction around the contour as t increases.

Finally, if $f(s)$ has a pole of order j at s_j, then the residue of $f(s)$ at s_j is

$$\text{Res}(f, s_j) = \lim_{s \to s_j} \frac{1}{(m-1)!} \frac{d^{m-1}}{ds^{m-1}} \left((s - s_j)^m f(s) \right).$$

This large expression is quite simple in the case that $f(s)$ has a simple pole – that is, a pole of order 1 – at s_j, because in this case, we have $\text{Res}(f, s_j) = \lim_{s \to s_j} (s - j) f(s)$. With these definitions, we state this amazing theorem.

Theorem 44 (Cauchy's Residue Theorem). *Let D be a domain in which $f(s)$ is holomorphic except at the points $s_1, \ldots, s_n$, and let C be a simple closed positively oriented contour in D containing some of these points, say, $w_1, \ldots, w_m$. Then*

$$\int_C f(s) ds = 2\pi i \sum_{k=1}^{m} \text{Res}(f, w_k).$$

If there are no such points w_k (or s_k), then the value of the integral is 0.

Once again, we omit the proof, as it is on the syllabus of most introductory complex analysis courses.

A.1.2 Complex Differentiation

The property of function which is complex-differentiable at a point s goes by many names: complex-differentiable at s, analytic at s, holomorphic at s, and regular at s. We use the third option. There are many ways one can prove that a given function f is holomorphic, some of which are quite obvious: show $f = g + h$ or $f = gh$ for holomorphic functions g and h, etc. Here is a method which is useful for the purpose of this exercise:

Theorem 45. *If f_n is a sequence of holomorphic functions such that $f_n \to f$ uniformly and $f_n' \to g$ uniformly, then $\frac{df}{ds} = g$. That is,*

$$\frac{d}{ds} \lim_{n \to \infty} f_n = \lim_{n \to \infty} \frac{df_n}{ds}.$$

One surprising fact about complex differentiation is the following:

Theorem 46. *If a function is holomorphic at s, then derivatives of all orders exist at s.*

This is not true about real differentiation! The existence of the first derivative *does not imply* the existence of the second derivative! Some facts in complex analysis seem too good to be true – but they are! This one in particular was used multiple times in this exercise.

A.2 The Leibniz Integral Rule

We prove the following theorem, which has theoretical uses (see the proof of Theorem 5) and practical uses (as an integration tool!). It is stated as follows:

Theorem 47 (The Leibniz Integral Rule). *Let $a(s)$ and $b(s)$ differentiable real-valued functions on $\mathbb{C}$, and let $f(s,t)$ be a complex-valued function defined on $\mathbb{C} \times \mathbb{R}$. Assume that $a(s)$, $b(s)$, and*

$f(s,x)$ are differentiable with respect to s, that $f(s,x)$ is continuous with respect to x, and that $\frac{\partial}{\partial s}f(s,x)$ is integrable with respect to x. Then

$$\frac{d}{ds}\int_a^b f(s,x)dx = \int_a^b \frac{\partial}{\partial s}f(s,x)dx - f(s,a)\frac{da}{ds} + f(s,b)\frac{db}{ds}.$$

Proof. Let $I = \int_a^b f(s,x)dx$. Regard I as a function of a, b, and s. Then note three things:

1. Since $f(s,x)$ is continuous in x, the Fundamental Theorem of Calculus implies that

$$\frac{\partial I}{\partial a} = -\frac{\partial}{\partial a}\int_b^a f(s,x)dx = -f(s,a).$$

2. Similarly, we have

$$\frac{\partial I}{\partial b} = f(s,b).$$

3. Fix some s. In the case that $a(s) = b(s)$, it follows trivially that

$$\frac{\partial}{\partial s}I(s,a,b) = \frac{\partial}{\partial s}\int_a^b f(s,x)dx = 0 = \int_a^b \frac{\partial}{\partial s}f(s,x)dx. \tag{40}$$

Now suppose that $a(s) \neq b(s)$. Let $\epsilon > 0$. Since $f(s,x)$ is differentiable with respect to s, there exists $\delta > 0$ such that

$$\left|\frac{f(s',x) - f(s,x)}{|s'-s|} - \frac{\partial}{\partial s}f(s,x)\right| < \frac{\epsilon}{|b-a|}$$

whenever $|s'-s| < \delta$. Fix such a δ, and let $s' \in \mathbb{C}$ such that $|s'-s| < \delta$. Then

$$\left|\int_a^b \frac{f(s',x) - f(s,x)}{|s'-s|}dx - \int_a^b \frac{\partial}{\partial s}f(s,x)dx\right|$$

is equal to

$$\left|\int_a^b \left(\frac{f(s',x) - f(s,x)}{|s'-s|} - \frac{\partial}{\partial s}f(s,x)\right)dx\right|,$$

which is less than or equal to

$$\int_a^b \left|\frac{f(s',x) - f(s,x)}{|s'-s|} - \frac{\partial}{\partial s}f(s,x)\right|dx \leq \int_a^b \frac{\epsilon}{|b-a|}dx = \epsilon.$$

Thus we conclude that

$$\lim_{s'\to s}\int_a^b \frac{f(s',x) - f(s,x)}{|s'-s|}dx = \int_a^b \frac{\partial}{\partial s}f(s,x)dx.$$

Now the left-hand side of this expression is equal to

$$\lim_{s'\to s}\frac{\left(\int_a^b f(s',x)dx - \int_a^b f(s,x)dx\right)}{|s'-s|} = \lim_{s'\to s}\frac{I(s',a,b) - I(s,a,b)}{|s'-s|} = \frac{\partial}{\partial s}I(s,a,b).$$

So Equation 40 holds for $a(s) \neq b(s)$ and, therefore, always.

So the partial derivatives of $I(s,a,b)$ with respect to s, a, and b exist. Therefore, since $a(s)$ and $b(s)$ are differentiable, the we can apply the chain rule to differentiate $I(s,a,b)$ with respect to s:

$$\begin{aligned}\frac{dI}{ds} &= \frac{\partial I}{\partial s} + \frac{\partial I}{\partial a}\frac{da}{ds} + \frac{\partial I}{\partial b}\frac{db}{ds} \\ &= \int_a^b \frac{\partial}{\partial s} f(s,x)dx - f(s,a)\frac{da}{ds} + f(s,b)\frac{db}{ds}.\end{aligned}$$

This proves the theorem. □

In the case that $a(s)$ and $b(s)$ are constants, then $a'(s) = b'(s) = 0$, and we have the (more) common special case:

Corollary 48. *If $f(s,x)$ is a complex-valued function defined on $\mathbb{C} \times \mathbb{R}$ such that $f(s,x)$ is differentiable with respect to s, $f(s,x)$ is continuous with respect to x, and $\frac{\partial}{\partial s} f(s,x)$ is integrable with respect to x, then*

$$\frac{d}{ds}\int_a^b f(s,x)dx = \int_a^b \frac{\partial}{\partial s} f(s,x)dx.$$

A.3 Useful algebraic concepts and facts

We prove two group-theoretic facts in the first two sections of this appendix, and in the third section we define the Legendre symbol, and give the method one uses to compute them: quadratic reciprocity. The first two sections follow [Ap], and the third section follows [L]. I must make a cautionary note here: my proofs which follow [Ap] are quite close to the originals, which are very beautifully conveyed. Though I do not feel as though I have done justice to the exposition in [Ap], I have nevertheless shamelessly included the proofs here. My reasoning was *not*, in this case, because I added anything to the original presentation; rather, this inclusion ix both for completeness and for myself. By working through these theorems, I understand two important theorems from group theory which I would otherwise not know.

A.3.1 The multiplicative group $\mathbb{Z}_{p^\alpha}^*$ is cyclic for all odd prime p

We prove that $\mathbb{Z}_{p^\alpha}^*$ is cyclic for all odd primes p and $\alpha \in \mathbb{N}$. In fact, $\mathbb{Z}_2^*$, $\mathbb{Z}_4^*$, and $\mathbb{Z}_{2p^\alpha}^*$ (for odd primes p and $\alpha \in \mathbb{N}$) are also cyclic; interestingly, these are the *only* other cyclic groups of the form $\mathbb{Z}_q^*$. All of this is proven in [Ap], but we only prove here what we need for this exercise. We start by proving that $\mathbb{Z}_p^*$ is cyclic for all primes p.

Lemma 49. *Let p be an odd prime. Then $\mathbb{Z}_p^*$ is cyclic.*

Proof. We assume two facts from elementary group theory: that the order $|a^k|$ of a^k is equal to $|a|/\gcd(|a|,k)$ and that $\sum_{d|n}\phi(d) = n$, where the sum is over the positive divisors of d and $\phi(d)$ is (as usual) the Euler totient function.

For each $d|p-1$, define $A(d) = \{x \in \mathbb{Z}_p^* : |x| = d\}$, where $|x|$ denotes the order of x in $\mathbb{Z}_p^*$. Since every element $x \in \mathbb{Z}_p^*$ has a unique order which divides $|\mathbb{Z}_p^*| = \phi(p) = p-1$ (by Lagrange's theorem), the collection of $A(d)$ partitions $\mathbb{Z}_p^*$. Hence $\sum_{d|p-1}|A(d)| = p-1$. From our assumed fact from group theory, we have

$$\sum_{d|p-1}(\phi(d) - |A(d)|) = 0. \tag{41}$$

We claim that $\phi(d) = |A(d)|$ for all $d|p-1$. To obtain this from Equation 41, we prove that either $|A(d)| = 0$ or $|A(d)| = \phi(d)$. This would imply that $\phi(d) - |A(d)| \geq 0$ for all $d|p-1$, in which case we have equality in Equation 41 if and only if $\phi(d) = |A(d)|$ for each $d|p-1$.

So fix some $d|p-1$ and suppose $|A(d)| \neq 0$. Then choose some $a \in A(d)$. Then $|a| = d$, so $a^d \equiv_p 1$ and no smaller power of a is congruent to 1 modulo p. Hence $a, a^2, \ldots, a^d$ are d distinct elements, all of which satisfy $x^d - 1 \equiv_p 0$. Since the polynomial $x^d - 1$ has as most d solutions in $\mathbb{Z}_p^*$, *every* solution is of the form a^k for some $k = 1, \ldots, p-1$. Therefore, since every $x \in A(d)$ satisfies $x^d - 1 \equiv_p 0$, $A(d)$ is a subset of the set generated by a. We can express this as

$$A(d)\{a^k : k = 1, \ldots, d;\ |a^k| = d\}.$$

But $|a^k| = |a|/\gcd(|a|, k) = d/\gcd(d, k)$, as one proves in elementary group theory, so $|a^k| = d$ if and only if $\gcd(d, k) = 1$. So $A(d)$ is a set, indexed by all $k = 1, \ldots, p-1$ such that $\gcd(k, d) = 1$. Hence $|A(d)| = \phi(d)$, as claimed.

What does this have to do with proving that $\mathbb{Z}_p^*$ is cyclic? Well, $|A(d)| = \phi(d)$ for all $d = 1, \ldots, p-1$, so, in particular, $|A(p-1)| = \phi(p-1)$. Since 1 is relatively prime to every natural number, $\phi(p-1) \geq 1$, and since $A(p-1)$, the set of elements in $\mathbb{Z}_p^*$ with order $p-1$, has at least one element, there exists an element in $\mathbb{Z}_p^*$ with order $p-1$. This element, having order $p-1$ must generate a subgroup of $\mathbb{Z}_p^*$ with $p - 1 = |\mathbb{Z}_p^*$ elements; hence, this element generates $\mathbb{Z}_p^*$, and the proof is complete. □

Next we prove another lemma, which is of little interest besides the fact that we need it:

Lemma 50. *Let p be an odd prime. If g is a generator of $\mathbb{Z}_p$ such that $g^{p-1} \not\equiv_{p^2} 1$, then for all $\alpha \geq 2$, $g^{\phi(p^{\alpha-1})} \not\equiv_{p^\alpha} 1$.*

Proof. Let g be as given. We proceed by induction on α. For $\alpha = 2$,

$$g^{\phi(p^{\alpha-1})} \equiv_{p^2} g^{\phi(p)} \equiv_{p^2} g^{p-1} \not\equiv_{p^2} 1,$$

so the base case holds. Now suppose $g^{\phi(p^{\alpha-1})} \not\equiv_{p^\alpha} 1$ for α. Considering g as an element of $\mathbb{Z}_{p^{\alpha-1}}$, we have by Lagrange's theorem that $|g|$ divides $\phi(p^{\alpha-1})$. So $g^{\phi(p^{\alpha-1})} \equiv_{p^{\alpha-1}} 1$. Translating this congruence into a fact that holds in $\mathbb{Z}$, we have

$$g^{\phi(p^{\alpha-1})} = 1 + kp^{\alpha-1} \tag{42}$$

for some $k \in \mathbb{Z}$. But by our induction hypothesis, $p \nmid k$ because $g^{\phi(p^{\alpha-1})} \not\equiv_{p^\alpha} 1$. That $p \nmid k$ is crucial, as we will see in a moment.

Using Equation 42, $g^{\phi(p^\alpha)}$. Since

$$\phi(p^\alpha) = p^{\alpha-1}(p-1) = p(p^{\alpha-2}(p-1)) = p\phi(p^{\alpha-1}),$$

we have

$$g^{\phi(p^\alpha)} = \left(g^{\phi(p^{\alpha-1})}\right)^p = (1 + kp^{\alpha-1})^p.$$

Expanding this binomial yields

$$\begin{aligned} g^{\phi(p^\alpha)} &= 1 + pkp^{\alpha-1} + \frac{1}{2}p(p-1)(kp^{\alpha-1})^2 \\ &\quad + \frac{1}{6}p(p-1)(p-2)(kp^{\alpha-3})^3 + \ldots + (kp^{\alpha-1})^p \\ &= 1 + kp^\alpha + k^2\left(\frac{p-1}{2}\right)p^{2\alpha-1} + p^{3\alpha-3}t \end{aligned}$$

for some integer t. Now $\alpha \geq 2$, so $2\alpha - 1$ and $3\alpha - 3$ are both at least $\alpha + 1$. Hence $p^{\alpha+1}$ divides both $p^{2\alpha-1}$ and $p^{3\alpha-3}$, which implies $g^{\phi(p^\alpha)} \equiv_{p^{\alpha+1}} 1 + kp^\alpha$. Recalling the key property about k, namely, that $p \nmid k$, we have that $g^{\phi(p^\alpha)} \not\equiv_{p^{\alpha+1}} 1$, and the induction is complete. □

We now prove the main theorem of this section:

Theorem 51. *Let p be an odd prime, and let $\alpha \geq 1$. Then the following hold:*

(a) If g is a generator of $\mathbb{Z}_p^*$, then g generates $\mathbb{Z}_{p^\alpha}^*$ if and only if $g^{p-1} \not\equiv_{p^2} 1$.

(b) One of the generators of $\mathbb{Z}_p^*$ satisfies $g^{p-1} \not\equiv_{p^2} 1$.

(c) There exists a generator of $\mathbb{Z}_{p^\alpha}^*$; that is, $\mathbb{Z}_{p^\alpha}^*$ is cyclic.

Proof. We prove (a) first. By Lemma 49, $\mathbb{Z}_p^*$ is cyclic. Let g be a generator of $\mathbb{Z}_p^*$. If g in fact generates $\mathbb{Z}_{p^\alpha}^*$ for all $\alpha \geq 1$, then g generates $\mathbb{Z}_{p^2}^*$, which implies $g^{p-1} \not\equiv_{p^2} 1$ because $p - 1 < p^2$.

Now suppose that $g^{p-1} \not\equiv_{p^2} 1$. For $\alpha = 1$, then, g generates $\mathbb{Z}_{p^\alpha}^* = \mathbb{Z}_p^*$. So consider $\alpha \geq 2$. We need to show that g generates $\mathbb{Z}_{p^\alpha}^*$ or, equivalently, that $|g| = \phi(p^\alpha)$, where $|g|$ is the order of g in $\mathbb{Z}_{p^\alpha}^*$.

We have that $g^{|g|} \equiv_{p^\alpha} 1$, so $g^{|g|} \equiv_p 1$. Since g is a generator of $\mathbb{Z}_p^*$, $g^{|g|} \equiv_p 1$ implies that $|g| \equiv 0$ modulo $\phi(p) = p - 1$, the order of $\mathbb{Z}_p^*$. So $\phi(p)$ divides $|g|$. Write $|g| = m\phi(p)$. By Lagrange's theorem, $|g|$ divides $\phi(p^\alpha)$, the order of $\mathbb{Z}_{p^\alpha}^*$, so $m\phi(p)|\phi(p^\alpha)$. Since $\phi(p) = p - 1$ and $\phi(p^\alpha) = p^{\alpha-1}(p-1)$, this implies $m(p-1)|p^{\alpha-1}(p-1)$. If we rewrite this as $p^{\alpha-1}(p-1) = m'm(p-1)$ for some $m' \in \mathbb{Z}$, we see that $m|p^{\alpha-1}$. So $m = p^\beta$ for some $\beta = 0, 1, \ldots, \alpha - 1$. We will be done if we can show that $\beta = \alpha - 1$, for this would imply that

$$|g| = m\phi(p) = p^\beta(p-1) = p^{\alpha-1}(p-1) = \phi(p^\alpha).$$

We proceed here by contradiction. Specifically, assume that $\beta \leq \alpha - 2$. Then $|g| = p^\beta(p-1)$ divides $p^{\alpha-1}(p-1) = \phi(p^\alpha)$, which implies that $g^{\phi(p^{\alpha-1})} \equiv_{p^\alpha} 1$. Since $\alpha \geq 2$, since g generates $\mathbb{Z}_p^*$, and since $g^{p-1} \not\equiv_{p^2} 1$ by assumption, Lemma 50 implies that we have a contradiction. Hence $\beta = \alpha - 1$, which is sufficient, as we proved above.

To prove part (b), we let g be a generator of $\mathbb{Z}_p^*$. (Again, a generator exists by Lemma A.3.1.) If $g^{p-1} \not\equiv_{p^2} 1$, then g generates $\mathbb{Z}_{p^\alpha}^*$ by (a). If, on the other hand, $g^{p-1} \equiv_{p^2} 1$, then $(g+p)^{p-1} \not\equiv_{p^2} 1$, as we will show. Since $g + p \equiv_p g$, $g + p$ is a generator of $\mathbb{Z}_p^*$, so (a) would imply that $g + p$ generates $\mathbb{Z}_{p^\alpha}^*$.

To show that $(g+p)^{p-1} \not\equiv_{p^2} 1$, consider that

$$(g+p)^{p-1} = g^{p-1} + p(p-1)g^{p-2} + \frac{1}{2}p^2(p-1)(p-2)g^{p-3} + \ldots + p^{p-1}.$$

Now if we reduce this modulo p^2, then the right-hand side becomes $1 + p(p-1)g^{p-2}$. So we have $(g+p)^{p-1} - 1 \equiv_{p^2} p(p-1)g^{p-2}$. Since g generates $\mathbb{Z}_p^*$, we have that $p \nmid g^{p-2}$, from which we conclude that $p^2 \nmid p(p-1)g^{p-2}$. Hence $(g+p)^{p-1} - 1 \not\equiv_{p^2} 0$ or, equivalently, $(g+p)^{p-1} \not\equiv_{p^2} 1$, as desired.

Part (c) clearly follows from (a) and (b), which concludes the theorem. □

A.3.2 The multiplicative group $\mathbb{Z}_{2^\alpha}^*$ for $n \geq 3$ is generated by 5 and -1.

We prove the useful fact that, although $\mathbb{Z}_{2^\alpha}$ is not cyclic, it is generated by the elements 5 and -1. To do this, we first show that the order of 5, denoted by $|5|$, is $2^{\alpha-2}$.

Since $|5|$ divides the order $\phi(2^\alpha) = 2^{\alpha-1}$ of $\mathbb{Z}_{2^\alpha}$, $|5| = 2^\beta$ for some $\beta = 1, 2, \ldots, \alpha - 1$. Now write $5 = 1 + 2^2$ and consider that

$$5^{2^\beta} = (1 + 2^2)^{2^\beta} = 1 + 2^\beta 2^2 + \frac{1}{2}2^\beta(2^\beta - 1)(2^2)^2 + \ldots + 2^{2^{\beta+1}}.$$

So

$$5^{2^\beta} - 1 = 2^{\beta+2} + 2^{\beta+3}r = 2^{\beta+2}(2r + 1)$$

for some integer r. Now since $|5| = 2^\beta$ in the group $\mathbb{Z}_{2^\alpha}$, we have $5^{2^\beta} \equiv 1$ modulo 2^α. So $2^\alpha \mid (5^{2^\beta} - 1)$, from which we conclude that $2^\alpha \mid 2^{\beta+2}$ since $2r + 1$ is odd. Hence $\alpha \leq \beta + 2$, or $\beta \geq \alpha - 2$. To conclude that $\beta = \alpha - 2$ (as opposed to $\beta = \alpha - 1$), we note that $\beta = \alpha - 1$ would imply that $|5| = 2^{\alpha-1}$, which is the order of the group, which in turn would imply that 5 generates the group. But no power of 5 is congruent to 3 modulo 2^α since all powers of 5 are congruent to 1 modulo 4 whereas $3 \equiv_4 1$.

Thus $|5| = 2^{\alpha-2}$, and hence $\{5, 5^2, \ldots, 5^{2^{\alpha-2}}\}$ is a collection of $2^{\alpha-2}$ distinct elements. Now consider the collection $\{-5, -5^2, \ldots, -5^{2^{\alpha-2}}\}$. Since negation is a one-to-one function, this is another collection of $2^{\alpha-2}$ distinct elements. Moreover, the first collection is made up of elements congruent to 1 modulo 4 whereas the second collection is made up of elements congruent to -1, or 3, modulo 4. Hence the union of these two collections, which is the set generated by 5 and -1, is a set of $2 \cdot 2^{\alpha-2} = 2^{\alpha-1}$ distinct elements. Since $\mathbb{Z}^*_{2^\alpha}$ has exactly this many elements, we conclude 5 and -1 generate $\mathbb{Z}_{2^\alpha}$.

A.3.3 The Legendre Symbol

Given a prime p, consider the group of units $\mathbb{Z}^*_p$. For example, let $p = 7$. We might ask which of the elements in $\mathbb{Z}_7$ are squares in $\mathbb{Z}_7$, or quadratic residues modulo 7. Computing, we see that $1^1 \equiv 1$, $2^2 \equiv 4$, $3^2 \equiv 2$, $4^2 \equiv 2$, $5^2 \equiv 4$, and $6^2 \equiv 1$. So 1, 2, and 4 are called *quadratic residues* modulo 7, and 3, 5, and 6 and called *quadratic nonresidues* modulo 7. Note that there are equally as many quadratic residues as nonresidues. As it turns out, this is always true when $\mathbb{Z}_q$ is cyclic but $q \neq 2$ or 4:

Theorem 52. *If $\mathbb{Z}^*_q$ is cyclic but not $\mathbb{Z}^*_2$ or $\mathbb{Z}^*_4$, then exactly half of the elements are squares. Equivalently, if $q = p^\alpha$ or $2p^\alpha$ where p is and odd prime, then there are $\phi(q)/2$ quadratic residues (and therefore $\phi(q)/2$ nonresidues) modulo q.*

Proof. Since $\mathbb{Z}^*_q$ is cyclic, there exists a generating element g such that

$$\mathbb{Z}^*_q = \{g^m\}_{m=1}^{\phi(q)} = \{g^{2m}\}_{m=1}^{\phi(q)/2} \cup \{g^{2m+1}\}_{m=1}^{\phi(q)/2}.$$

From this it is clear that at least half of the elements are squares. On the other hand, $x^2 \equiv (-x)^2$ for all $x \in \mathbb{Z}^*_q$, so if $x \not\equiv -x$ for all $x \in \mathbb{Z}^*_q$, then at most half of the elements would be squares. Hence exactly half would be squares in $\mathbb{Z}^*_q$. But if $x \equiv_q -x$ for some $x \in \mathbb{Z}^*_q$, then $2x \equiv_q 0$, which implies that q, which is p^α or $2p^\alpha$, divides $2x$. Hence we would have that p, an odd prime, would divide x. But this would imply that $x \notin \mathbb{Z}^*_q$, a contradiction, so the theorem is proved. □

Legendre came up with some handy notation for dealing with quadratic residues:

Definition 53. *For all odd primes p, define the Legendre symbol $(\frac{n}{p})$ for all $n \in \mathbb{N}$ by*

$$\left(\frac{n}{p}\right) = \begin{cases} 1, & \text{if } n \text{ is a quadratic residue modulo } p; \\ -1, & \text{if } n \text{ is not a quadratic residue modulo } p; \\ 0, & \text{if } p \mid n. \end{cases}$$

Note from this definition that $(\frac{n^2}{p}) = 1 = (\frac{n}{p})^2$ for all n which are not divisible by p, and that, if $p \mid n$, then $(\frac{n^2}{p}) = 0 = (\frac{n}{p})^2$. So $(\frac{n^2}{p}) = (\frac{n}{p})^2$ for all $n \in \mathbb{Z}$. As it turns out, a stronger fact holds: $(\frac{mn}{p}) = (\frac{m}{p})(\frac{n}{p})$ for all $m, n \in \mathbb{N}$. We enumerate this and a few other important properties of the Legendre symbol here:

Theorem 54. *Let p and q be odd primes, and let m and n be integers. Then the following are true of the Legendre symbol:*

1. $\left(\frac{mn}{p}\right) = \left(\frac{m}{p}\right)\left(\frac{n}{p}\right)$.
2. $\left(\frac{m+p}{p}\right) = \left(\frac{m}{p}\right)$.
3. $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$.
4. $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}$.

Proof. We gave a proof of 1 above for the special case of $m = n$; we refer the reader to [L] or [Ap] for a proof of the general case. The second fact follows because the Legendre symbol is computed using information about $m + p$ and m as elements in $\mathbb{Z}_p^*$. But $m + p$ and m are the same elements in $\mathbb{Z}_p^*$, so the Legendre symbol involving each is the same.

The third and fourth points, or sometimes just the fourth point, are called the quadratic reciprocity law. I read my first proof of this fact in [Da], but others are found in [L] and [Ap]. □

Using this theorem, it is clear how we could compute the Legendre symbol of any odd integer n modulo p. We would first factor the integer and use the multiplicativity to rewrite $(\frac{n}{p})$ as a product of symbols $(\frac{q}{p})$ over all (odd) primes $q|p$. If n is negative, we would use part 3 to calculate the $(\frac{-1}{p})$ term. For the remaining terms, we could use part 4 to "flip" each $(\frac{q}{p})$ for which $q < p$ into $\pm(\frac{p}{q})$ so as to make the computation easier. Finally, one starts squaring elements in $\mathbb{Z}_p^*$ and other various $\mathbb{Z}_q^*$ until all of the symbols are computed. To see an example, see our calculation in Equation 26.

A.4 Carefully considered convergence issues

A general rule I have learned and now live by – at least when I do mathematics – is, *be careful, no, be very careful when you see infinities in your mathematics.* To be sure, we have been very careful this entire time. Only at a few points in this exercise have we taken a slippery step or two. To give some traction to these steps, I have added this appendix.

A.4.1 Infinite products

Although we see infinite sums (which we call series) in our calculus courses, we seldom see infinite products. In contrast to a first or second calculus course, infinite products abound in complex analysis. Convergence issues are translated into convergence issues about series, which we understand very well. This appendix contains a brief overview of infinite products and their convergence properties. The presentation follows [Ah].

An infinite product $P = \prod_{n=1}^{\infty} p_n$ is defined as the limit of the partial products $P_n = \prod_{k=1}^{n} p_k$. This is similar to the definition of an infinite sum. However, there is an added technical requirement. An infinite product is convergent if at most finitely many p_n are zero, and if the product of the nonzero terms converges to something finite and nonzero. There are good reasons for these

technicality. For example, a series has just one zero, all of the partial products (past a certain point) are zero – no matter what the other terms in the product are!

Note that, for a convergent product, $p_n = P_n/P_{n-1} \to 1$ as $n \to \infty$. With some thought, we would expect this; it is analogous to the fact that the terms in a convergent series converge to 0. Hence it makes sense to let $p_n = 1 + a_n$ for all $n \in \mathbb{N}$.

Consider a convergent product in which every $p_n = 1 + a_n \neq 0$. We can consider

$$\log \prod_{n=1}^{\infty}(1 + a_n) = \sum_{n=1}^{\infty} \log(1 + a_n). \tag{43}$$

(To prove this equality, one considers the partial product from 1 to N, uses the corresponding property of logarithms of finite products, takes $N \to \infty$, and uses the fact that logarithms are continuous functions.) As it turns out, the infinite product converges if and only if the series converges. In this way, the problem of testing for convergence in an infinite product is reduced to testing for convergence in a series.

An infinite product *absolutely converges* if the corresponding, given in Equation 43, absolutely converges. Since the series involves logarithms, testing for absolute convergence might be difficult. Fortunately, we have the following theorem:

Theorem 55. *An infinite product $\prod_{n=1}^{\infty}(1 + a_n)$ converges absolutely if and only if the series $\sum_{n=1}^{\infty} a_n$ converges absolutely.*

With these definitions of convergence being so closely tied to the definitions of convergence for a series, statements about termwise differentiability of series, for example, are easily translated translated into statements about termwise differentiability of products.

A.4.2 Combining absolutely convergent series

If the sums $\sum_k a_k$ and $\sum_k b_k$ are finite sums over the same index k, then

$$\sum_k (a_k + b_k) = \sum_k a_k + \sum_k b_k.$$

We would like to say the same about series:

Theorem 56 (Adding absolutely convergent series). *Let $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ be absolutely convergent series. Then*

$$\sum_{k=1}^{\infty}(a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k,$$

and this converges absolutely.

Proof. For all $n \in \mathbb{N}$ let $r_n = \sum_{k=1}^{n} a_k$, $s_n = \sum_{k=1}^{\infty} b_k$, and $t_n = \sum_{k=1}^{n}(a_k + b_k) = r_n + s_n$. By hypothesis, r_n and t_n converge as $n \to \infty$, so

$$\lim_{n\to\infty}(r_n + s_n) = \lim_{n\to\infty} r_n + \lim_{n\to\infty} s_n = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k.$$

On the other hand, this limit is equal to

$$\lim_{n\to\infty} t_n = \sum_{k=1}^{\infty}(a_k + b_k).$$

So the first part of the theorem is proved. To show that the convergence is absolute, consider that

$$\sum_{k=1}^{\infty} |a_k + b_k| \leq \sum_{k=1}^{\infty} (|a_k| + |b_k|) \leq \sum_{k=1}^{\infty} |a_k| + \sum_{k=1}^{\infty} |b_k|,$$

which is finite by the theorem's hypothesis. □

Next, we consider whether we can multiply series in the natural way. We prove the following theorem:

Theorem 57 (Multiplying absolutely convergent series). *The product equality*

$$\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} a_j b_k = \left(\sum_{j=1}^{\infty} a_j \right) \left(\sum_{k=1}^{\infty} b_k \right) \tag{44}$$

holds and the sum on the left-hand side converges absolutely.

Proof. Using the same notation for r_n and s_n, and letting

$$t_n = \sum_{j=1}^{n} \sum_{k=1}^{\infty} a_j b_k = r_n s_n,$$

the fact that r_n and s_n converge as $n \to \infty$ implies

$$\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} a_j b_k = \lim_{n \to \infty} t_n = \left(\lim_{n \to \infty} r_n \right) \left(\lim_{n \to \infty} s_n \right).$$

This proves Equation 44. To prove absolute convergence, consider that, for all $n \in \mathbb{N}$,

$$\sum_{j=1}^{n} \sum_{k=1}^{n} |a_j b_k| = \left(\sum_{j=1}^{n} |a_j| \right) \left(\sum_{k=1}^{n} |b_k| \right) \leq \left(\sum_{j=1}^{\infty} |a_j| \right) \left(\sum_{k=1}^{\infty} |b_k| \right).$$

The absolute convergence of the sums r_n and s_n imply that the right-hand side is finite. Hence the left-hand side is bounded for all $n \in \mathbb{N}$, which proves that the limit as $n \to \infty$ is finite. □

A.4.3 The Dirichlet test for convergence

We first prove a finite analogue to integration by parts. Our functions "u" and "v" will be the sequences defined on the nonnegative integers.

Theorem 58 (Abel summation). *If (a_k) and (b_k) are sequences of complex numbers, and if (A_k) is the sequence of partial sums of the terms in (a_k), that is, $A_n = \sum_{k=0}^{n} a_k$, then*

$$\sum_{k=1}^{n} a_k b_k = \sum_{k=0}^{n-1} A_k (b_k - b_{k+1}) + A_n b_n$$

for all $n \in \mathbb{N}$.

Proof. We proceed by induction on n. For $n = 0$, the sum on the right-hand side is empty, so the base case is trivial. Now suppose the equality holds for n. Then, using our induction hypothesis right away, adding and subtracting $A_n(b_n - b_{n+1})$, and replacing $A_n + a_{n+1}$ by A_{n+1}, we obtain

$$\begin{aligned}
\sum_{k=0}^{n+1} a_k b_k &= \sum_{k=0}^{n} A_k(b_k - b_{k+1}) + a_{n+1}b_{n+1} \\
&= \sum_{k=0}^{n-1} A_k(b_k - b_{k+1}) + A_n b_n + a_{n+1}b_{n+1} \\
&= \sum_{k=0}^{n} A_k(b_k - b_{k+1}) - A_n(b_n - b_{n+1}) + A_n b_n + a_{n+1}b_{n+1} \\
&= \sum_{k=0}^{n} A_k(b_k - b_{k+1}) + (A_n + a_{n+1})b_{n+1} \\
&= \sum_{k=0}^{(n+1)-1} A_k(b_k - b_{k+1}) + A_{n+1}b_{n+1}.
\end{aligned}$$

We recognize this as our claim for $n + 1$, so the result follows by induction. □

By setting $a_0 = a_1 = \cdots = a_m = 0$, we have that $A_0 = A_1 = \cdots = A_m = 0$, from which we derive the following corollary:

Corollary 59. *Let (a_n), (b_n), and (b_n) be as given in Theorem 58. Then*

$$\sum_{k=m+1}^{n} a_k b_k = \sum_{k=m+1}^{n-1} A_k(b_k - b_{k+1}) + A_n b_n$$

for all $n \geq m$.

The reader should now recall the alternating series test for the convergence of series. It says that if $c_n = (-1)^n b_n$ for all $n \in \mathbb{N}$, where the b_n are decreasing and converging to 0 as $n \to \infty$, then the series $\sum_{n=1}^{\infty} c_n$ converges. The theorem below is a generalization of this test. In short, it states that the $(-1)^n$ in the definition of c_n can be replaced by any sequence a_n as long as the partial sums $\sum_{k=1}^{n} a_k$ remain bounded as $n \to \infty$.

Theorem 60 (Dirichlet's test for convergence). *Let (a_n) and (b_n) be sequences in $\mathbb{R}$ such that*

- *there exists $M \in \mathbb{R}$ such that $\left|\sum_{k=1}^{n} a_k\right| \leq M$ for all $n \in \mathbb{N}$, and*
- *b_n decreases and converges to 0*

Then the series $\sum_{n=1}^{\infty} a_n b_n$ converges.

Proof. Since $\mathbb{R}$ is complete, proving that the series converges is equivalent to showing that it is Cauchy. Since we do not know (or care) what the point of convergence is, it will be easiest to prove that the sequence is Cauchy and use the completeness of $\mathbb{R}$ to conclude the theorem.

To this end, let $\epsilon > 0$. Since the partial sums are bounded, let $M > 0$ such that $|A_n| \leq M$ for all $n \in \mathbb{N}$. Now $b_n \to 0$, so there exists $N_1 \in \mathbb{N}$ such that $|b_n| < \epsilon/(3M)$ for all $n > N_1$. Also, the sequence (b_n) is necessarily Cauchy (because it converges!), so there exists $N_2 \in \mathbb{N}$ such that $n \geq m > N_2$ implies that $b_m - b_n = |b_m - b_n| < \epsilon/(3M)$. Choose $N \in \mathbb{N}$ to be larger than N_1 and N_2. Then for all $n \geq m > N$, using Abel's summation lemma, we have

$$\begin{aligned}\left|\sum_{k=0}^{n} a_k b_k - \sum_{k=0}^{m} a_k b_k\right| &= \left|\sum_{k=0}^{n-1} A_k(b_k - b_{k+1}) + A_n b_n - \sum_{k=0}^{m-1} A_k(b_k - b_{k+1}) - A_m b_m\right| \\ &\leq \left|\sum_{k=m}^{n-1} A_k(b_k - b_{k+1})\right| + |A_n b_n| + |A_m b_m| \\ &\leq M\sum_{k=m}^{n-1} |b_k - b_{k+1}| + M|b_n| + M|b_m|.\end{aligned}$$

Now (b_n) is decreasing, so $\sum_{k=m}^{n-1} |b_k - b_{k+1}| = \sum_{k=m}^{n-1}(b_k - b_{k+1}) = b_m - b_n < \epsilon/(3M)$ by our choice of N. Also by our choice of N, $|b_n|$ and $|b_m|$ are less than $\epsilon/(3M)$. So

$$\left|\sum_{k=0}^{n} a_k b_k - \sum_{k=0}^{m} a_k b_k\right| \leq M\left(\frac{\epsilon}{3M}\right) + M\left(\frac{\epsilon}{3M}\right) + M\left(\frac{\epsilon}{3M}\right) = \epsilon.$$

Hence the partial sums form a Cauchy and, therefore, convergent sequence. This is what it means for the series to converge, so our proof is complete. □

A.4.4 The Taylor series for $-\log(1-z)$ converges for all $z \neq 1$ such that $|z| \leq 1$

That this series converges for $|z| < 1$ is obvious by comparison the geometric series z^n. The convergence for $|z| = 1$ follows from Dirichlet's test for convergence since

1. n^{-1} decreases and converges to 0 as $n \to \infty$, and
2. $\sum_{n=1}^{\infty} z^n = \sum_{n=1}^{\infty} e^{i\phi n}$, for some $0 < \phi < 2\pi$, and the series on the right is bounded.

The proof of the second point shall be my final treat for the reader. We do this by proving the following theorem:

Theorem 61. *Let $\theta > 1$. Then*

$$\sum_{n=1}^{\infty} e^{2\pi i n/\theta}$$

is bounded.

Proof. First note that if θ is an integer, then the sum of any consecutive θ terms – which are θ-th roots of unity – is 0. Intuitively, then, for general θ, the sum of $\lfloor\theta\rfloor$ consecutive terms should be about 0. So to prove this fact, we first consider partial sums up to integer multiples of $\lfloor\theta\rfloor$ and

break the sums up into blocks of $\lfloor\theta\rfloor$ terms. That is, for all $m \in \mathbb{N}$,

$$\begin{aligned}\sum_{n=1}^{m\lfloor\theta\rfloor} e^{2\pi in/\theta} &= \sum_{j=0}^{m-1}\sum_{n=j\lfloor\theta\rfloor+1}^{(j+1)\lfloor\theta\rfloor} e^{2\pi in/\theta} \\ &= \sum_{j=0}^{m-1}\sum_{n'=1}^{\lfloor\theta\rfloor} e^{2\pi in'/\theta}e^{2\pi ij\lfloor\theta\rfloor/\theta} \\ &= \left(\sum_{j=0}^{m-1}\left(e^{2\pi i\lfloor\theta\rfloor/\theta}\right)^j\right)\left(\sum_{n'=1}^{\lfloor\theta\rfloor}\left(e^{2\pi i/\theta}\right)^{n'}\right),\end{aligned}$$

where we have reindexed the inner sum by the variable $n' = n - j\lfloor\theta\rfloor$. Note that, for $\theta \in \mathbb{Z}$, the second term in the product is 0, and hence the whole sum is 0. For $\theta \notin \mathbb{Z}$, $\lfloor\theta\rfloor/\theta$ is not an integer, and hence, after evaluating the geometric sums,

$$\left|\sum_{n=1}^{m\lfloor\theta\rfloor} e^{2\pi in/\theta}\right| = \left|\left(\frac{1-e^{2\pi i\lfloor\theta\rfloor m/\theta}}{1-e^{2\pi i\lfloor\theta\rfloor/\theta}}\right)\left(e^{2\pi i/\theta}\frac{1-e^{2\pi i\lfloor\theta\rfloor/\theta}}{1-e^{2\pi i/\theta}}\right)\right| \leq (1+1)|f(\theta)|,$$

where $f(\theta) = \left(e^{2\pi i/\theta}-1\right)^{-1}$ is a function of θ and is independent of m. Therefore, for all $N \in \mathbb{N}$, we can use the division algorithm to write $N = m\lfloor\theta\rfloor + r$ where $0 \leq r < \lfloor\theta\rfloor$ and conclude that

$$\left|\sum_{n=1}^{N} e^{2\pi in/\theta}\right| \leq \left|\sum_{n=1}^{m\lfloor\theta\rfloor} e^{2\pi in/\theta}\right| + \left|\sum_{n=m\lfloor\theta\rfloor+1}^{m\lfloor\theta\rfloor+r} e^{2\pi in/\theta}\right| \leq 2|f(\theta)| + r < 2|f(\theta)| + \lfloor\theta\rfloor.$$

Given any θ, therefore, the series $\sum_{n=1}^{\infty} e^{2\pi in/\theta}$ is bounded, as claimed. □

References

[Ah] Ahlfors, L.V. *Complex Analysis: An Introduction to the Theory of Analytic Functions of One Complex Variable*. York, PA: McGraw-Hill Book Company, Inc., 1953.

[Ap] Apostol, Tom M. *Introduction to Analytic Number Theory*. New York, NY: Springer-Verlag, 1976.

[C] Cohn, C. *Advanced Number Theory*. New York, NY: Dover Publications, 1962.

[D] Davenport, H. *Multiplicative Number Theory, Third Edition*. New York, NY: Springer-Verlag, 2000.

[Da] Davidoff, G. *Elementary Number Theory, Group Theory, and Ramanujan Graphs*. Cambridge, UK: Cambridge University Press, 2003.

[G] Goldfeld, D., The elementary proof of the Prime Number Theorem: An historical perspective, (preprint).

[L] LeVeque, W.J., *Topics in Number Theory, Volume 1*. Reading, MA: Addison-Wesley Publishing Company, Inc., 1956.

[N] Newman, D.J., Simple analytic proof of the Prime Number Theorem, *Amer. Math. Monthly*, Vol. **87** (1980), 693-696.

[Sa] Saff, E.B., and Snider, A.D., *Fundamentals of Complex Analysis with Applications to Engineering and Science, Third Edition*. Upper Saddle River, NJ: Pearson Education, Inc., 2003.

[So] Soprounov, I., A short proof of the Prime Number Theorem for arithmetic progressions, (preprint).

[Z] Zagier, D., Newman's short proof of the Prime Number Theorem, *Amer. Math. Monthly*, Vol. **104** (1997), No. 8, 705-708.