2. INFINITELY MANY PRIMES, WITH ANALYSIS

The ides of this chapter is to introduce the reader to some basic counting concepts while: at first, proving that there are infinitely many primes, then giving upper and lower bounds for the number of primes, or certain functions of the number of primes, culminating in a discussion of the inter-relation between these counting functions.

2.1. First Counting Proofs. It is instructive to see several proofs involving simple counting arguments, as these will give a flavor of things to come.

Suppose that there are only finitely many primes, say $p_1 < p_2 < \ldots < p_k$. Let $m = p_1 p_2 \ldots p_k$. If d is squarefree then d divides m, and so the number of integers up to m that are divisible by d is m/d. Therefore, by the inclusion-exclusion principle, the number of positive integers up to m that are not divisible by any prime is

$$m - \sum_{i=1}^{k} \frac{m}{p_i} + \sum_{1\le i<j\le k} \frac{m}{p_i p_j} - \cdots = m \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right) = \prod_{i=1}^{k} (p_i - 1).$$

Now 3 is a prime so this quantity is $\geq 3 - 1 = 2$. However 1 is the only positive integer that is not divisible by any prime, and so we obtain a contradiction.

The astute reader will observe that this counting argument holds with or without the assumption: We have simply evaluated Euler's function $\phi(m)$, the number of positive integers $\leq m$ that are coprime with m.

Another version of this proof is to count more accurately the number of integers up to x: By the fundamental theorem of arithmetic we wish to count the number, N, of k-tuples of non-negative integers $(e_1, e_2, \ldots, e_k)$ such that $p_1^{e_1} p_2^{e_2} \ldots p_k^{e_k} \leq x$. In other words we wish to count the number or lattice points $(e_1, e_2, \ldots, e_k) \in \mathbb{Z}^k$ inside the k-dimensional tetrahedron $T(x)$ defined by $e_1, \ldots, e_k \geq 0$ with

$$e_1 \log p_1 + e_2 \log p_2 + \cdots + e_k \log p_k \leq \log x. \tag{2.1.1}$$

If one adjoins to each lattice point a unit cube in the positive direction, one obtains a shape S whose volume is exactly the number of lattice points, and which is roughly the same shape as the original tetrahedron itself. The tetrahedron $T(x)$ has volume

$$\frac{1}{k!} \prod_{i=1}^{k} \frac{\log x}{\log p_i}, \tag{2.1.2}$$

so we might expect this to be a good approximation to the number of such lattice points; actually it is a lower bound on the number of lattice points since $T(x) \subset S$. On the other

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX hand S is contained in the tetrahedron given by adding 1 in each direction to the diagonal boundary of the tetrahedron, that is $T(xp_1 \dots p_k)$. Therefore if $p_1 < p_2 < \dots < p_k$ are the only primes then $x = N = |S| \leq |T(xp_1 \dots p_k)|$ for all x; and so if $m = p_1 p_2 \dots p_k$ then

$$x \leq \frac{1}{k!} \prod_{i=1}^{k} \frac{\log(mx)}{\log p_i}, \tag{2.1.3}$$

which is certainly false if x is large enough (see exercise 1.4).

One can easily obtain good enough upper bounds on the number of solutions to (2.1.1) with less work: In any solution to (2.1.1) we must have $e_j \log p_j$ is no more than the left side of (2.1.1) and thus than $\log x$, so that $0 \leq e_j \leq (\log x)/(\log p_j)$, and therefore $N \leq \prod_{i=1}^{k} (\log xp_j)/(\log p_j)$.

Exercises

2.1a. Prove that (2.1.3) is false for x sufficiently large. Find a constant C such that (2.1.3) is false for $x = (C \log m)^k$.

2.1b. It is useful to be able to count the number of integers up to x divisible by a given integer $d \geq 1$. These are the set of integers of the form dn where $n \geq 1$ and $dn \leq x$, in other words these are in 1-to-1 correspondence with the set of integers n in the range $1 \leq n \leq x/d$. There are $[x/d]$ such integers (where $[t]$, the *integer part of t*, denotes the largest integer $\leq t$), and $[x/d] \leq x/d$. What about the number of positive integers up to x which are $\equiv a \pmod d$? Can you come up with a precise expression for this in terms of the least positive residue of $a \pmod d$? Can you provide an approximation, involving a smooth function involving only the variables x and d, that is out from the correct count by at most 1?

2.2. Euler's proof and the Riemann zeta-function.

In the seventeenth century Euler gave a different proof that there are infinitely many primes, one which would prove highly influential in what was to come later. Suppose again that the list of primes is $p_1 < p_2 < \cdots < p_k$. Euler observed that the fundamental theorem of arithmetic implies that there is a 1-to-1 correspondence between the sets $\{n \geq 1 : \ n \text{ is a positive integer}\}$ and $\{p_1^{a_1} p_2^{a_2} \dots p_k^{a_k} : \ a_1, a_2, \dots, a_k \geq 0\}$. Thus a sum involving the elements of the first set should equal the analogous sum involving the elements of the second set:

$$\begin{aligned}
\sum_{\substack{n \geq 1 \\ n \text{ a positive integer}}} \frac{1}{n^s} &= \sum_{a_1, a_2, \dots, a_k \geq 0} \frac{1}{(p_1^{a_1} p_2^{a_2} \dots p_k^{a_k})^s} \\
&= \left(\sum_{a_1 \geq 0} \frac{1}{(p_1^{a_1})^s} \right) \left(\sum_{a_2 \geq 0} \frac{1}{(p_2^{a_2})^s} \right) \cdots \left(\sum_{a_k \geq 0} \frac{1}{(p_k^{a_k})^s} \right) \\
&= \prod_{j=1}^{k} \left(1 - \frac{1}{{p_j}^s} \right)^{-1}.
\end{aligned}$$

The last equality holds because each sum in the second-to-last line is over a geometric progression. Euler then noted that if we take $s = 1$ then the right side equals some rational number (since each $p_j > 1$) whereas the left side equals ∞, a contradiction (and thus there cannot be finitely many primes). We prove that $\sum_{n\geq 1} 1/n$ diverges in exercise 2.2a below

What is wonderful about Euler's formula is that something like it holds without assumption, involving the infinity of primes; that is

$$\sum_{\substack{n\geq 1\\ n \text{ a positive integer}}} \frac{1}{n^s} = \prod_{p \text{ prime}} \left(1-\frac{1}{p^s}\right)^{-1}. \tag{2.2.1}$$

One does need to be a little careful about convergence issues. It is safe to write down such a formula when both sides are "absolutely convergent", which takes place when $s > 1$. In fact they are absolutely convergent even if s is a complex number so long as $\mathrm{Re}(s) > 1$.

We have just seen that (2.2.1) makes sense when s is to the right of the horizontal line in the complex plane going through the point 1. Like Euler, we want to be able to interpret what happens to (2.2.1) when $s = 1$. To not fall afoul of convergence issues we need to take the limit of both sides as $s \to 1^+$, since (2.2.1) holds for values of s larger than (2.2.1). To do this it is convenient to note that the left side of (2.2.1) is well approximated by $\int_1^\infty \frac{dt}{t^s} = \frac{1}{s-1}$, and thus diverges as $s \to 1^+$. We deduce that

$$\prod_{p \text{ prime}} \left(1-\frac{1}{p}\right) = 0 \tag{2.2.2}$$

which, upon taking logarithms, implies that

$$\sum_{p \text{ prime}} \frac{1}{p} = \infty. \tag{2.2.3}$$

So how numerous are the primes? One way to get an idea is to determine the behaviour of the sum analogous to (2.2.3) for other sequences of integers. For instance $\sum_{n\geq 1} \frac{1}{n^2}$ converges, so the primes are, in this sense, more numerous than the squares. We can do better than this from our observation, just above, that $\sum_{n\geq 1} \frac{1}{n^s} \approx \frac{1}{s-1}$ is convergent for any $s > 1$ (see exercise 2.2b below). In fact, since $\sum_{n\geq 1} \frac{1}{n(\log n)^2}$ converges, we see that the primes are in the same sense more numerous than the numbers $\{n(\log n)^2 : \ n \geq 1\}$, and hence there are infinitely many integers x for which there are more than $x/(\log x)^2$ primes $\leq x$.

Exercises

2.2a. The box with corners at $(n, 0), (n + 1, 0), (n, 1/n), (n + 1, 1/n)$ has area $1/n$ and contains the area under the curve $y = 1/x$ between $x = n$ and $x = n + 1$. Therefore $\sum_{n\leq N} 1/n \geq \int_1^{N+1} \frac{1}{t} dt = \log(N + 1)$. Deduce that the sum of the reciprocals of the positive integers diverges.

a) Now draw the box of height $1/n$ and width 1 to the left of the line $x = n$, and obtain the upper bound $\sum_{n\leq N} 1/n \leq \log(N) + 1$.

b) Our goal is to prove that $\lim_{N\to\infty}(1/1 + 1/2 + 1/3 + \cdots + 1/N - \log N)$ exists — it is usually denoted by γ and called the *Euler-Mascheroni constant*. Now let $x_n = 1/1 + 1/2 + 1/3 + \cdots + 1/n - \log n$ for each integer $n \geq 1$. By the same argument as in part (a) show that if $n > m$ then $0 \leq x_m - x_n \leq \log(1 + 1/m) - \log(1 + 1/n) < 1/m$. Thus x_m is a Cauchy sequence and converges to a limit as desired. It can be shown that $\gamma = .5772156649\ldots$.

c) Prove that $0 \le 1/1 + 1/2 + 1/3 + \cdots + 1/N - \log N - \gamma \le 1/N$.

d) Let $\{t\} = t - [t]$ denote the *fractional part* of t. Prove that

$$\gamma = 1 - \int_1^\infty \frac{\{t\}}{t^2}\, dt.$$

2.2b. Use the method of 2.2a to show that if $\sigma > 1$ then

$$\frac{1}{\sigma - 1} \le \sum_{n \ge 1} \frac{1}{n^\sigma} \le \frac{\sigma}{\sigma - 1}.$$

2.2c. Given that $\sum_p 1/p$ diverges, deduce that there are arbitrarily large values of x for which $\#\{p \le x : p \text{ prime}\} \ge \sqrt{x}$. Improve the $\sqrt{x}$ here as much as you can using these methods.

2.2d. Prove that there are infinitely many primes p with a 1 in their decimal expansion.

2.3. Upper bound on the number of primes up to x. Fix $\epsilon > 0$. By (2.2.2) we know that there exists y such that $\prod_{p \le y}(1 - 1/p) < \epsilon/3$. Let m be the product of the primes $\le y$, and select $x > 3y/\epsilon$. If $k = [x/m]$, so that $km \le x < (k+1)m < 2km$, then the number of primes up to x is no more than the number of primes up to $(k+1)m$, which is no more than the number of primes up to y plus the number of integers up to $(k+1)m$ which have all of their prime factors $> y$. Since there are no more than y primes up to y, and since the set of integers up to $(k+1)m$ is $\{jm + i : \ 1 \le i \le m, \ 0 \le j \le k\}$, we deduce that the number of primes up to x is

$$\begin{aligned} &\le y + \sum_{j=0}^{k} \sum_{\substack{1 \le i \le m \\ (jm+i,m)=1}} 1 = y + (k+1)\phi(m) \\ &< \epsilon x/3 + 2km \prod_{p \le y} (1 - 1/p) < \epsilon x/3 + 2x\epsilon/3 = \epsilon x. \end{aligned}$$

In other words

$$\lim_{x \to \infty} \frac{1}{x} \ \#\{p \le x : \ p \text{ prime}\} \to 0.$$

2.4. An explicit lower bound on the sum of reciprocals of the primes. Every integer n up to x can be written as a squarefree integer m times a square, r^2. Moreover any squarefree integer $\le x$ can be written as a product of distinct primes $\le x$, so that

$$\sum_{n \le x} \frac{1}{n} \le \sum_{r \ge 1} \frac{1}{r^2} \ \sum_{m \le x} \frac{\mu^2(m)}{m} \le 2 \prod_{\substack{p \text{ prime} \\ p \le x}} \left(1 + \frac{1}{p}\right),$$

where the last inequality follows from exercise 2.2b above and $\mu(m)$ is the Mobius function. Since $1 + 1/p \le e^{1/p}$ we may take the logarithm of both sides to obtain, using exercise 2.2a,

$$\sum_{\substack{p \text{ prime} \\ p \le x}} \frac{1}{p} \ge \log\log(x+1) - \log 2. \tag{2.4.1}$$

This gives a good quantitative lower bound on the number of primes, with a certain weight.

2.5. Another explicit lower bound. An easier argument follows from writing $\log n = \sum_{p^a|n} \log p$ where the sum is over all of the prime powers that divide n. We obtain the identity

$$\sum_{n\le N} \log n = \sum_{n\le N} \sum_{p^a|n} \log p = \sum_{\substack{p \text{ prime, } a\ge 1\\ p^a\le N}} \log p \sum_{\substack{n\le N\\ p^a|n}} 1. \tag{2.5.1}$$

Use a modification of the method of exercise 2.2a to prove that $\sum_{n\le N} \log n \ge \int_1^N \log t dt = N(\log N - 1) + 1$. Then (2.5.1) yields, using exercise 2.1b,

$$N(\log N - 1) + 1 \le \sum_{\substack{p \text{ prime}\\ p\le N}} \sum_{a\ge 1} \frac{N}{p^a} \log p \le N \sum_{\substack{p \text{ prime}\\ p\le N}} \frac{\log p}{p-1}.$$

Therefore

$$\sum_{\substack{p \text{ prime}\\ p\le N}} \frac{\log p}{p-1} \ge \log N - 1. \tag{2.5.2}$$

2.6. Binomial coefficients: First bounds. Every prime in $(n+1, 2n+1]$ divides the numerator of the binomial coefficient $\binom{2n+1}{n}$. Therefore the product of these primes is $\le \binom{2n+1}{n} \le \frac{1}{2} \times 2^{2n+1}$. We deduce by induction that the product of the primes up to N is $\le 4^{N-1}$ for all $N \ge 1$. Taking logarithms (in base e) we obtain

$$\sum_{\substack{p \text{ prime}\\ p\le x}} \log p \le (x-1)\log 4. \tag{2.6.1}$$

If $n > 1$ then $\binom{n}{[n/2]}$ is the largest of the binomial coefficients $\binom{n}{a}$; and is at least as large as the sum of the two smallest. Therefore $2^n \le \sum_{a=0}^n \binom{n}{a} \le n\binom{n}{[n/2]}$. Combining this with exercise 2.6b.c below we obtain that

$$\frac{2^n}{n} \le \binom{n}{[n/2]} \le \prod_{\substack{p \text{ prime}\\ p\le n}} p^e \le n^{\#\{p \text{ prime: } p\le n\}},$$

so that

$$\#\{p \text{ prime}: \ p \le n\} \ge (\log 2)\ \frac{n}{\log n} - 1. \tag{2.6.2}$$

Exercises

2.6a. Modify the above argument to prove that

$$\#\{p \le n :\ p \text{ prime}\} \le \log 4\ \frac{n}{\log n} + 2(\log 4)^2\ \frac{n}{(\log n)^2} \tag{2.6.3}$$

for all $n \ge 2$. (Hint: First prove this for all $n \le 100$, and then by induction noting that $\log(2n+1) - \log(n+1) \le \log 2$, and that $1/\log(n+1) \le 13/(11\log(2n+1))$ for all $n \ge 50$.)

2.6b. a) Use exercise 2.1b to prove that the power of prime p which divides $n!$ is $\sum_{e\ge 1}[n/p^e]$.

b) Deduce that the power of p that divides the binomial coefficient $\binom{a+b}{a}$ is given by the number of carries when adding a and b in base p.

c) Deduce that if p^e divides the binomial coefficient $\binom{n}{m}$ then $p^e \le n$.

2.6c. Improve (2.6.1) by considering the prime divisors of $\binom{6n}{3n}\binom{3n}{2n}$ (note: check this!).

2.7. Bertrand's postulate. This states that there is a prime number between n and $2n$ for every integer n. Equations (2.6.2) and (2.6.3), taken together, just fail to imply Bertrand's postulate. Paul Erdős's ingenious approach involves a more detailed analysis of the prime factors of binomial coefficients, similar to those that arose in section 2.6, by considering the primes in different intervals: Let p^{e_p} be the exact power of prime p dividing $\binom{2n}{n}$. From exercise 2.6b.b we know that $e_p = 1$ if $n < p \le 2n$, and $e_p = 0$ if $2n/3 < p \le n$ (verify this as an exercise). By exercise 2.6b.c we know that $e_p \le 1$ if $\sqrt{2n} < p \le 2n/3$, and that $p^{e_p} \le 2n$ if $p \le \sqrt{2n}$. Therefore we have

$$\begin{aligned}\frac{2^{2n}}{2n} \le \binom{2n}{n} &= \prod_{p\le 2n} p^{e_p} \le \prod_{n<p\le 2n} p \prod_{p\le 2n/3} p \prod_{p\le\sqrt{2n}} 2n \\ &\le \left(\prod_{n<p\le 2n} p\right) \times 4^{2n/3-1} \times (2n)^{(\sqrt{2n}+1)/2}\end{aligned}$$

using estimates proved in section 2.6, since the number of primes up to $\sqrt{2n}$ is no more than $(\sqrt{2n}+1)/2$ (as neither 1 nor any even integer > 2 is prime). Taking logarithms we deduce that

$$\sum_{\substack{p \text{ prime} \\ n<p\le 2n}} \log p > \frac{\log 4}{3}\ n - \frac{\sqrt{2n}+3}{2}\ \log(2n).$$

This implies that

$$\sum_{\substack{p \text{ prime} \\ n<p\le 2n}} \log p \ge \frac{1}{3}\ n \tag{2.7.1}$$

for all $n \ge 2349$. It is a simple matter to write a computer program to check that (2.7.1) holds for all n in the range $1 \le n \le 2348$. Therefore (2.7.1) holds for all $n \ge 1$ which implies a strong form of Bertrand's postulate.

Exercises

2.7a. Write a computer program to verify (2.7.1) for all positive integers $n \leq 10000$.

2.7b. Verify Bertrand's postulate for all n up to 20000 using *only* the primes $2, 3, 5, 7, 13, 23, 43, 83, 163, 317, 631, 1259, 2503, 5$

2.7c. Prove that there are infinitely many primes p with a 1 as the leftmost digit in their decimal expansion.

2.7d. Use Bertrand's postulate to show, by induction, that every integer $n > 6$ can be written as the sum of distinct primes.

2.7e. Our goal is to prove the theorem of Sylvester and Schur, that any k consecutive integers, beginning at $n+1$ with $n \geq k$, is divisible by a prime $> k$. We suppose that this is false.
a) Show that $\binom{n+k}{n}$ only has prime factors $\leq k$ and so is $\leq (n+k)^{\pi(k)}$.

b) Since $\binom{n+k}{n} > (n/k)^k$, use (2.6.3) deduce that $n \leq 8k$ if $k \geq 62503$. Improve on this as best you can, perhaps using an explicit form of Stirling's formula.

c) Sylvester actually proved that if $(m, d) = 1$ where $d \geq 1$ and $m > n$ then $(m+d)(m+2d)\dots(m+nd)$ has a prime factor larger than n. Modify your work above to go some way to proving this.

2.8. Big Oh and other notation. Obtaining an upper bound on the sum in (2.5.2) of the form $\log N + C$ for some constant C, by the methods of section 2.5, is feasible but complicated. Moreover one might guess that $x_N := \sum_{p \leq N} (\log p)/p - \log N$ tends to a limit as $N \to \infty$, so the value of C that we can obtain is of limited interest. If instead we focus on proving that x_N is bounded, then our work is easier, and it pays to have notation that reflects this new objective:

If $A(x)$ and $B(x)$ are functions of x, and there exists some constant $c > 0$ such that $|A(x)| \leq cB(x)$ for all $x \geq 1$ then we write $A(x) = O(B(x))$ (we say "$A(x)$ is big oh of $B(x)$"), or $A(x) \ll B(x)$ ("$A(x)$ is less than, less than $B(x)$"), or even $B(x) \gg A(x)$ ("$B(x)$ is greater than, greater than $A(x)$"). If $A(x) \ll B(x)$ and $B(x) \ll A(x)$, that is $A(x), B(x) > 0$ and there exist constants $c_1, c_2 > 0$ such that $c_1 B(x) \leq A(x) \leq c_2 B(x)$, then we write $A(x) \asymp B(x)$. For example, $\#\{p \text{ prime} : \ p \leq n\} \asymp n/\log n$ by (2.6.2) and (2.6.3).

Now the left side of (2.5.1) equals

$$\int_1^N \log t \; dt + O(\log N) = N(\log N - 1) + O(\log N).$$

The right side equals

$$\sum_{\substack{p \text{ prime},\ a \geq 1 \\ p^a \leq N}} \log p \left[\frac{N}{p^a}\right] = \sum_{\substack{p \text{ prime},\ a \geq 1 \\ p^a \leq N}} \log p \left(\frac{N}{p^a} + O(1)\right) = N \sum_{\substack{p \text{ prime},\ a \geq 1 \\ p^a \leq N}} \frac{\log p}{p^a} + O(N)$$

by (2.6.1), and

$$\sum_{\substack{p \text{ prime},\ a \geq 2 \\ p^a \leq N}} \frac{\log p}{p^a} \leq \sum_{p \text{ prime}} \frac{\log p}{p(p-1)} = O(1),$$

by exercise 2.2b. Therefore, combining the above displayed equations with (2.5.1), we obtain

$$\sum_{\substack{p \text{ prime} \\ p \leq N}} \frac{\log p}{p} = \log N + O(1). \tag{2.8.1}$$

We define $E(N) := \sum_{p\leq N} \frac{\log p}{p} - \log N$ so that $E(N) = O(1)$ by (2.8.1). From the identity

$$\sum_{\substack{p \text{ prime}\\ p\leq N}} \frac{1}{p} = \frac{1}{\log N} \sum_{\substack{p \text{ prime}\\ p\leq N}} \frac{\log p}{p} + \int_2^N \frac{1}{t(\log t)^2} \left(\sum_{\substack{p \text{ prime}\\ p\leq t}} \frac{\log p}{p} \right) dt \tag{2.8.2}$$

we obtain, using (2.8.1),

$$\begin{aligned}\sum_{\substack{p \text{ prime}\\ p\leq N}} \frac{1}{p} &= \frac{1}{\log N}(\log N + E(N)) + \int_2^N \frac{1}{t(\log t)^2}(\log t + E(t))\ dt \\ &= 1 + \frac{E(N)}{\log N} + \log\left(\frac{\log N}{\log 2}\right) + \int_2^\infty \frac{E(t)}{t(\log t)^2}dt - \int_N^\infty \frac{E(t)}{t(\log t)^2}dt.\end{aligned}$$

Now since $|E(t)| \ll 1$ we have $\int_2^\infty \frac{E(t)}{t(\log t)^2}dt \ll \int_2^\infty \frac{1}{t(\log t)^2}dt = -\left[\frac{1}{\log t}\right]_2^\infty = \frac{1}{\log 2}$, and that $\int_N^\infty \frac{E(t)}{t(\log t)^2}dt \ll \frac{1}{\log N}$. So let c be the constant $1 - \log\log 2 + \int_2^\infty \frac{E(t)}{t(\log t)^2}dt$ and then the above reads:

$$\sum_{\substack{p \text{ prime}\\ p\leq N}} \frac{1}{p} = \log\log N + c + O\left(\frac{1}{\log N}\right), \tag{2.8.3}$$

a big improvement on (2.4.1), although determining c from this proof would be difficult.

How difficult is it to obtain the seemingly magical identity (2.8.2)? Can we find such identities in other circumstances? This leads us to the subject of partial summation: Suppose that $a(n)$ is some function on the integers (for example $a(p) = (\log p)/p$ if p is prime, and $a(n) = 0$ otherwise), and let $A(x) := \sum_{n\leq x} a(n)$. If $f(t)$ is a differentiable function on $\mathbb{R}^+$ then, formally, we have

$$\begin{aligned}\sum_{n\leq x} a(n)f(n) &= \int_{1^-}^x f(t)dA(t) = [f(t)A(t)]_{1^-}^x - \int_{1^-}^x f'(t)A(t)dt \\ &= f(x)\sum_{n\leq x} a(n) - \int_1^x f'(t)\left(\sum_{n\leq t} a(n)\right) dt \end{aligned} \tag{2.8.4}$$

after integrating by parts, since $A(1^-) = 0$. Calculating a sum in this manner is called *partial summation*. We see that (2.8.2) is an example of this, with $f(t) = 1/\log t$, a decreasing function. If $a(n) \geq 0$ for all n and $f(x) > 0$ for all $x > x_0$, then this formula is most successful with decreasing $f(t)$ since then $f'(t) < 0$ and all but finitely many of the terms in (2.8.4) take the same sign. Let's look at an example where $f(t)$ is increasing: Again let $a(p) = (\log p)/p$ if p is prime, and $a(n) = 0$ otherwise, but this time take $f(t) = t/\log t$, so as to count the number of primes up to N. Then, by (2.8.4) we have

$$\sum_{\substack{p \text{ prime} \\ p \le N}} 1 = \frac{N}{\log N} \sum_{\substack{p \text{ prime} \\ p \le N}} \frac{\log p}{p} - \int_1^N \left(\frac{1}{\log t} - \frac{1}{(\log t)^2} \right) \left(\sum_{\substack{p \text{ prime} \\ p \le t}} \frac{\log p}{p} \right) dt,$$

so by (2.8.1) the right side becomes

$$\begin{aligned} &= \frac{N}{\log N}(\log N + O(1)) - \int_1^N \left(\frac{1}{\log t} - \frac{1}{(\log t)^2} \right) (\log t + O(1))\ dt \\ &= N + O\left(\frac{N}{\log N} \right) - \left\{ N + O\left(\int_1^N \frac{dt}{\log t} \right) \right\} \ll \frac{N}{\log N} \end{aligned} \tag{2.8.5}$$

(we could have just as well have written $O(N/\log N)$ at the final equality). We see here that the main terms cancel and that the secondary terms are not known accurately enough to get an asymptotic estimate. Typically, when $f(t)$ is increasing there will be this cancellation between the main terms of the two halves of the formula, so we usually use partial summation when f is decreasing.

Exercises

2.8a. Prove that $\sum_p (\log p)/p \ - \ \sum_p (\log p)/(p-1) \ll 1$.

2.8b. Prove (2.8.4) by considering the coefficient of $a(n)$ on both sides, for each n.

2.8c. Let $a_N = \log N! - N(\log N - 1)$. We have seen that $a_N \ll \log N$. Prove that $\log n = \int_{n-1/2}^{n+1/2} \log t dt + O(1/n^2)$, and deduce that $a_N - a_M = \int_M^N \log t dt + \frac{1}{2}\log(N/M) + O(1/M)$. Finally deduce that there exists a constant C such that $N! = (N/e)^N (CN)^{1/2}\{1 + O(1/N)\}$. (In fact $C = 2\pi$, which is *Stirling's formula.*)

2.9. How many prime factors does a typical integer have?. We begin this section by estimating the average number of prime factors of an integer $\le x$. To be clear we must distinguish whether we are counting the prime factors with their multiplicities or not; that is, one might count $12 = 2^2 \times 3$ as having two or three primes factors depending on whether one counts the 2^2 as one or two primes. So define

$$\omega(n) = \sum_{\substack{p \text{ prime} \\ p | n}} 1 \quad \text{and} \quad \Omega(n) = \sum_{\substack{p \text{ prime}, a \ge 1 \\ p^a | n}} 1.$$

First note that

$$\begin{aligned} \sum_{n \le x} \omega(n) &= \sum_{n \le x} \sum_{\substack{p \text{ prime} \\ p | n}} 1 = \sum_{p \text{ prime}} \sum_{\substack{n \le x \\ p | n}} 1 = \sum_{p \text{ prime}} \left[\frac{x}{p} \right] \\ &= \sum_{\substack{p \text{ prime} \\ p \le x}} \left(\frac{x}{p} + O(1) \right) = x \left(\log\log x + c + O\left(\frac{1}{\log x} \right) \right). \end{aligned} \tag{2.9.1}$$

by (2.8.3) and (2.8.5). Therefore the average number of *distinct* prime factors of an integer up to x is about $\log\log x + c + o(1)$. The difference when we allow multiplicities is

$$\sum_{n\le x}\{\Omega(n)-\omega(n)\} = \sum_{n\le x}\ \sum_{\substack{p \text{ prime}, a\ge 2\\ p^a|n}} 1 = \sum_{\substack{p \text{ prime}\\ a\ge 2}}\ \sum_{\substack{n\le x\\ p^a|n}} 1 = \sum_{\substack{p \text{ prime}\\ a\ge 2}} \left[\frac{x}{p^a}\right]$$

$$= \sum_{\substack{p \text{ prime}, a\ge 2\\ p^a\le x}} \left(\frac{x}{p^a}+O(1)\right). \tag{2.9.2}$$

Now there are terms in this sum only for primes $p \le \sqrt{x}$, and at most $O(\log x)$ terms for each such prime so that the error terms contribute $O(\pi(\sqrt{x})\log x) = O(\sqrt{x})$ by (2.8.5). For each such prime p we can extend the sum to the infinite arithmetic progression with sum $1/p(p-1)$ by adding in the terms with $p^a > x$. These extra terms sum to no more than $2/x$ thus contributing a total of $\ll \pi(\sqrt{x}) \ll \sqrt{x}$. Now let us add in the sum of $1/p(p-1)$ over all primes $p > \sqrt{x}$ which contributes $\le x\sum_{n>\sqrt{x}} 1/n(n-1) \ll \sqrt{x}$. Therefore we have proved that the quantity in (2.9.2) equals $c'x + O(\sqrt{x})$ where $c' = \sum_{p \text{ prime}} 1/p(p-1)$. This implies that the average number of (not necessarily distinct) prime factors of an integer up to x is about $\log\log x + c + c' + o(1)$, not much different from the distinct case.

We are going to go one step further and ask how much $\omega(n)$ varies from its mean, that is we are going to compute the statistical quantity, the *variance*. We begin with a standard identity for the variance (which we will use repeatedly, see exercise 2.9b):

$$\frac{1}{x}\sum_{n\le x}\left(\omega(n) - \frac{1}{x}\sum_{m\le x}\omega(m)\right)^2 = \frac{1}{x}\sum_{n\le x}\omega(n)^2 - \left(\frac{1}{x}\sum_{m\le x}\omega(m)\right)^2. \tag{2.9.3}$$

Now the sum in the first term here is

$$\sum_{n\le x}\omega(n)^2 = \sum_{n\le x}\ \sum_{\substack{p \text{ prime}\\ p|n}}\ \sum_{\substack{q \text{ prime}\\ q|n}} 1 = \sum_{p \text{ prime}}\left(\sum_{\substack{q \text{ prime}\\ q=p}}\left[\frac{x}{p}\right] + \sum_{\substack{q \text{ prime}\\ q\ne p}}\left[\frac{x}{pq}\right]\right)$$

$$\le \sum_{\substack{p \text{ prime}\\ p\le x}}\frac{x}{p} + \sum_{p \text{ prime}}\ \sum_{\substack{q \text{ prime},\ q\ne p\\ pq\le x}}\frac{x}{pq} \le x\sum_{\substack{p \text{ prime}\\ p\le x}}\frac{1}{p} + x\left(\sum_{\substack{p \text{ prime}\\ p\le x}}\frac{1}{p}\right)^2.$$

Therefore (2.9.1) implies that the variance is

$$\sum_{\substack{p \text{ prime}\\ p\le x}}\frac{1}{p} + \left(\sum_{\substack{p \text{ prime}\\ p\le x}}\frac{1}{p}\right)^2 - \left(\sum_{\substack{p \text{ prime}\\ p\le x}}\frac{1}{p} + O\left(\frac{\pi(x)}{x}\right)\right)^2 \ll \log\log x. \tag{2.9.4}$$

We are now going to "clean up" the statement for the variance to get a very elegant theorem of Hardy and Ramanujan, in the form given given by Turán: First note that, by applying the inequality $(a+b)^2 \leq 2(a^2+b^2)$ to each term in the sum below and then by (2.9.1), we obtain

$$\frac{1}{x}\sum_{n\leq x}(\omega(n)-\log\log x)^2 \ll (2.9.3)+\frac{1}{x}\sum_{n\leq x}(c+o(1))^2 \ll \log\log x. \tag{2.9.5}$$

If S is the set of $n \leq x$ for which $|\omega(n)-\log\log x| \geq (\log\log x)^{2/3}$, then the left side of (2.9.5) is

$$\geq \frac{1}{x}\sum_{n\in S}(\log\log x)^{4/3}+\frac{1}{x}\sum_{\substack{n\leq x\\ n\notin S}}0 = (|S|/x)(\log\log x)^{4/3}.$$

Therefore, by (2.9.5) we have $|S| \ll x/(\log\log x)^{1/3}$ which is, of course $o(x)$. Therefore *almost all* integers $n \leq x$ have $\log\log x + O((\log\log x)^{2/3})$ distinct prime factors. (By *almost all* integers up to x, we mean all but at most $o(x)$ integers up to x.) One more thing, $\log\log n = \log\log x + o(1)$ for almost all integers $n \leq x$ (see exercise 2.9d), and so almost all integers $n \leq x$ have $\log\log n + O((\log\log n)^{2/3})$ distinct prime factors. So we have proved

Theorem 2.9. *Almost all integers n have $\{1+o(1)\}\log\log n$ distinct prime factors.*

We need to explain the notation. The term "$o(1)$" stands for a function $A(n)$ for which $A(n) \to 0$ as $n \to \infty$. We could equally have written $\log\log n + o(\log\log n)$ where "$o(\log\log n)$" stands for a function $B(n)$ for which $B(n)/\log\log n \to 0$ as $n \to \infty$. By "almost all integers n" we mean "almost all integers $n \leq x$, for all sufficiently large x"; and by that we mean that there are no more than $o(x)$ integers $n \leq x$ for which this is false.

Exercises

2.9a. Use (2.9.1) to prove that the average number of prime factors of an integer in $[x,2x]$ is $\log\log x + c + O\left(\frac{1}{\log x}\right)$. Show that if n is in this interval then this quantity equals $\log\log n + c + O\left(\frac{1}{\log n}\right)$. Continue on to prove that $\log\log n = \log\log x + o(1)$ for almost all integers $n \leq x$ (Hint: One way to do so would be to start by considering integers n in the two ranges $x \geq n \geq x/\log x$, and then $n \leq x/\log x$).

2.9b. If $a_1,\ldots,a_N$ have average m show that $\frac{1}{N}\sum_{n\leq N}(a_n-m)^2 = \frac{1}{N}\sum_{n\leq N}a_n^2 - m^2$.

2.9c. Justify (2.9.3).

2.9d. Show that if $f(x)$ is a function such that, for every $\epsilon > 0$, we have $f(x) = 1+O(\epsilon)$ then $f(x) = 1+o(1)$.

2.9e. (Erdős's *multiplication table theorem*) Show that there are $o(x^2)$ distinct integers $n \leq x^2$ that can be written as the product of two integers $\leq x$. (Hint: Compare the typical number of prime factors of ab where $a,b \leq x$ with the typical number of prime factors of n.)

2.10. How many primes are there up to x? Let $\pi(x)$ denote the number of primes up to x. After (2.6.2) and (2.6.3) we know that for any $\epsilon > 0$ if x is sufficiently large then

$$(\log 2-\epsilon)\frac{x}{\log x} \leq \pi(x) \leq (\log 4+\epsilon)\frac{x}{\log x};$$

so one might guess that there exists some numbers L, between $\log 2$ and $\log 4$, such that $\pi(x)/(\frac{x}{\log x}) \to L$ as $x \to \infty$. Such a suggestion was first published by Legendre in 1808, with $L = 1$, in fact with the more precise assertion that there exists a constant b such that $\pi(x)$ is well approximated by $x/(\log x - B)$ for large enough x. Actually, back in 1792 or 1793, at the age of 15 or 16, Gauss had already made a much better guess, based on studying tables of primes.[1] His observation may be best quoted as

About 1 in $\log x$ of the integers near x are prime.

This suggest that a good approximation to the number of primes up to x is $\sum_{n=2}^{x} 1/\log n$. Using the method of exercise 2.2a this is, up to an error of $O(1)$, equal to

$$\int_2^x \frac{dt}{\log t}. \tag{2.10.1}$$

We denote this quantity by $\mathrm{Li}(x)$ and call it *the logarithmic integral.* Here is a comparison of Gauss's prediction with the actual count of primes up to various values of x:

x	$\pi(x) = \#\{\text{primes} \leq x\}$	Overcount: $\mathrm{Li}(x) - \pi(x)$
10^3	168	10
10^4	1229	17
10^5	9592	38
10^6	78498	130
10^7	664579	339
10^8	5761455	754
10^9	50847534	1701
10^{10}	455052511	3104
10^{11}	4118054813	11588
10^{12}	37607912018	38263
10^{13}	346065536839	108971
10^{14}	3204941750802	314890
10^{15}	29844570422669	1052619
10^{16}	279238341033925	3214632
10^{17}	2623557157654233	7956589
10^{18}	24739954287740860	21949555
10^{19}	234057667276344607	99877775
10^{20}	2220819602560918840	222744644
10^{21}	21127269486018731928	597394254
10^{22}	201467286689315906290	1932355208
10^{23}	1925320391606818006727	7236148412

TABLE 1. Primes up to various x, and the overcount in Gauss's prediction.

We see that Gauss's prediction is amazingly accurate. It does seem to always be an overcount, and since the width of the last column is about half that of the central one it appears that the difference is no bigger than $\sqrt{x}$, perhaps multiplied by a constant. The data certainly suggests that $\pi(x)/\mathrm{Li}(x) \to 1$ as $x \to \infty$.

[1] On Christmas Eve 1847 he wrote to Encke, "In 1792 or 1793 ... I turned my attention to the decreasing frequency of primes ... counting the primes in intervals of length 1000. I soon recognized that behind all of the fluctuations, this frequency is on average inversely proportional to the logarithm...

If we integrate (2.10.1) by parts then we find that for any fixed integer $N > 0$ that

$$\mathrm{Li}(x) = \frac{x}{\log x} + \frac{x}{(\log x)^2} + \frac{2!\; x}{(\log x)^3} + \ldots + \frac{(N-1)!\; x}{(\log x)^{N-1}} + O_N\left(\frac{x}{(\log x)^N}\right) \tag{2.10.2}$$

and so $\mathrm{Li}(x)/(\frac{x}{\log x}) \to 1$ as $x \to \infty$. Combining this with Gauss's prediction gives that $\pi(x)/(\frac{x}{\log x}) \to 1$ as $x \to \infty$. The notation of limits is rather cumbersome notation – it is easier to write

$$\pi(x) \sim \frac{x}{\log x} \tag{2.10.3}$$

as $x \to \infty$, "$\pi(x)$ is asymptotic to $x/\log x$". In general, $A(x) \sim B(x)$ is equivalent to $\lim_{x\to\infty} A(x)/B(x) = 1$. If one wishes to be more precise than (2.10.3) about the difference between the two sides one might write

$$\pi(x) = \frac{x}{\log x} + O\left(\frac{x}{(\log x)^2}\right), \tag{2.10.4}$$

so the difference between $\pi(x)$ and $x/\log x$ is bounded by a constant multiple of $x/(\log x)^2$. Confronted with a formula like (2.10.4) one might only be concerned as to whether the secondary term of the right side, $O(x/(\log x)^2)$ is actually much smaller than the main term, $x/\log x$. Specifically whether $(x/(\log x)^2)/(x/\log x) \to 0$ as $x \to \infty$: of course it does and we use the notation $x/(\log x)^2 = o(x/\log x)$. In general, $A(x) = o(B(x))$ is equivalent to $\lim_{x\to\infty} A(x)/B(x) = 0$ and we say "$A(x)$ is little oh of $B(x)$". Thus the right side of (2.10.4) can we rewritten as

$$\frac{x}{\log x} + o\left(\frac{x}{\log x}\right), \quad \text{or even} \quad \{1 + o(1)\}\frac{x}{\log x}$$

which is equivalent to (2.10.3).

The asymptotic (2.10.3) is called *The Prime Number Theorem* and its proof had to wait until the end of the nineteenth century, requiring various remarkable developments. The proof was a high point of nineteenth century mathematics and there is still no straightforward proof. There are reasons for this: Surprisingly the prime number theorem is equivalent to a statement about zeros of an analytic continuation, and although a proof can be given that hides this fact, it is still lurking somewhere just beneath the surface, perhaps inevitably so. A proof of the prime number theorem will be the central focus of the next few chapters of this book.

We will now that if L exists then it must equal 1: Suppose that $\pi(x) = \{L+o(1)\}x/\log x$. Using partial summation (taking $f(t) = (\log t)/t$ in (2.8.4) where $a(n) = 1$ if n is prime and 0 otherwise) we obtain

$$\begin{aligned}\sum_{p\le x} \frac{\log p}{p} &= \frac{\log x}{x}\pi(x) + \int_1^x \frac{\log t - 1}{t^2}\pi(t)dt \\ &= O(1) + \{L + o(1)\}\int_1^x \frac{\log t - 1}{t^2}\frac{t}{\log t}dt = L\log x + o(\log x).\end{aligned}$$

Comparing this to (2.8.1), we deduce that $L = 1$.

Exercises

2.10a. Prove that the difference between Gauss's prediction, $\mathrm{Li}(x)$, and Legendre's prediction, $x/(\log x - B)$, is $\geq x/(2(\log x)^3)$ for x sufficiently large, no matter what the choice of B.

2.10b. Determine, with proof, the asymptotic series in (2.10.2).

2.10c. Assuming the prime number theorem, show that for all $\epsilon > 0$ there are primes between x and $x + \epsilon x$ is x is sufficiently large. Deduce that $\mathbb{R}_{\geq 0}$ is the set of limit points of the set $\{p/q : \; p, q \text{ primes}\}$.

2.10d. Let $A(x) = x + \sqrt{x}$. Show that $\{1 + o(1)\}A(x) - \{1 + o(1)\}x$ is not equal to $\{1 + o(1)\}\sqrt{x}$, and indicate what it is equal to.

2.10e. Assume that $\pi(x) = x/\log x + \{1 + o(1)\}x/(\log x)^2$. Prove that if x is sufficiently large then there are more primes up to x than between x and $2x$.

2.10f. Show that if $\pi(x) = x/\log x + x/(\log x)^2 + O(x/(\log x)^{2+\epsilon})$ for some $\epsilon > 0$ then there exists a contant c such that $\sum_{p \leq x} (\log p)/p = \log x + c + o(1)$.

2.11. Smoothing out the prime counting function.

Table 1 indicates that Gauss's guesstimate $\mathrm{Li}(x)$ is indeed a startling good approximation to $\pi(x)$, the number of primes up to x. Although $\mathrm{Li}(x)$, as defined in (2.10.1) can be written in a rather compact form, its value is the complicated asymptotic series (2.10.2) which is far from easy to work with. Early researchers realized that if one counts primes with the *weight* $\log p$ then Gauss's guesstimate looks a lot nicer: We define

$$\theta(x) := \sum_{\substack{p \text{ prime} \\ p \leq x}} \log p$$

and expect that this should be well-approximated by $\sum_{n \leq x} \frac{1}{\log n} \cdot \log n = x$. Most easily we have

$$\theta(x) \leq \sum_{\substack{p \text{ prime} \\ p \leq x}} \log x = \log x \pi(x).$$

More precisely

$$\theta(x) - x = \int_1^x \log t d\{\pi(t) - \mathrm{Li}(t)\} = (\pi(x) - \mathrm{Li}(x)) \log x - \int_1^x \frac{\pi(t) - \mathrm{Li}(t)}{t} dt,$$

from which we deduce that

$$|\theta(x) - x| \leq 2 \log x \cdot \max_{1 \leq t \leq x} |\pi(t) - \mathrm{Li}(t)|. \tag{2.11.1}$$

Therefore if $\pi(x)$ is well approximated by $\mathrm{Li}(x)$ then $\theta(x)$ is well approximated by x. In the other direction

$$\pi(x) - \mathrm{Li}(x) = \int_{2^-}^x \frac{1}{\log t} d\{\theta(t) - t\} = \frac{\theta(x) - x}{\log x} + \frac{2}{\log 2} + \int_2^x \frac{\theta(t) - t}{t(\log t)^2} dt. \tag{2.11.2}$$

Riemann recognized, for technical reasons that we will come to later, that including the prime powers up to x leads to a "better" counting function. So define

$$\psi(x) := \sum_{\substack{p \text{ prime}, \ a\geq 1 \\ p^a \leq x}} \log p.$$

It is not difficult to relate $\psi(x)$ to $\theta(x)$, and so to $\pi(x)$; see exercise 2.11b.

How about the size of the nth prime: Let $p_1 = 2 < p_2 = 3 < \dots$ be the sequence of primes. By inverting the relation $\pi(x) \sim \mathrm{Li}(x)$ we can deduce that $p_n \sim n \log n$, and obtain better estimates for p_n the smaller that $|\pi(x) - \mathrm{Li}(x)|$ is.

Exercises

2.11a. Use (2.11.1) and (2.11.2) to show that

$$\max_{\sqrt{x}\leq t\leq x} |\theta(t) - t| + \sqrt{x} \asymp \log x \max_{\sqrt{x}\leq t\leq x} |\pi(t) - \mathrm{Li}(t)| + \sqrt{x}$$

2.11b. Prove that $\psi(x) = \theta(x) + \theta(x^{1/2}) + \theta(x^{1/3}) + \dots$. Deduce that $\psi(x) = \theta(x) + O(\sqrt{x})$.

2.11c. Prove that $\pi(x) \sim \mathrm{Li}(x)$ if and only if $\psi(x) \sim x$.

2.11d. Suppose that $\pi(x)$ is asymptotic to the first three terms of the asymptotic expansion of $\mathrm{Li}(x)$ as in (2.10.2). What does this imply about p_n?

2.11e. Prove that $\mathrm{lcm}[1, 2, \dots, n] = e^{\psi(n)}$.

2.12. The count of primes. Studying the table in section 2.10 we quickly notice two important things about the difference $\mathrm{Li}(x) - \pi(x)$ in the range in which $\pi(x)$ has been successfully calculated. Firstly that $\mathrm{Li}(x) - \pi(x)$ is positive throughout the range and we might guess that this is always so; this question has some interesting twists and will be discussed in detail in chapter 22. The second observation is that the second column in the table is about half the width of the first column, which suggest that $\mathrm{Li}(x) - \pi(x)$ is about the square root of $\pi(x)$. If we do not worry about the power of $\log x$ (which is small compared to the power of x by which this error term appears to be smaller than the main term) then we would conjecture that there exists a constant $A > -1$ such that

$$\pi(x) = \mathrm{Li}(x) + O(\sqrt{x}(\log x)^A). \tag{2.12.1}$$

This is equivalent to the two estimates $\theta(x) = x + O(\sqrt{x}(\log x)^{A+1})$, and $\psi(x) = x + O(\sqrt{x}(\log x)^{A+1})$, using partial summation. One of the great conjectures of mathematics is the *Riemann Hypothesis* which we shall discuss in detail throughout this book – if you are reading this book there is a good chance you have heard of it and its notoriety. It is a statement concerning the zeros of the analytic continuation of the function we defined in (2.2.1) in half of the complex plane, $\mathrm{Re}(s) > 1$, at first sight a rather difficult and esoteric question to appreciate. Riemann developed this question because he was thinking about the distribution of primes and, in particular, the difference $\mathrm{Li}(x) - \pi(x)$. In fact the Riemann Hypothesis is *equivalent* to (2.12.1) and to each of the two statements following it. It is intriguing how a question of esoteric complex analysis could be equivalent to something as down-to-earth as the count of primes (as we will discuss in section 23).

2.13. Mertens' theorem. This states that

$$\prod_{\substack{p \text{ prime} \\ p\leq x}} \left(1-\frac{1}{p}\right) = \frac{e^{-\gamma}}{\log x}\left(1+O\left(\frac{1}{\log x}\right)\right). \tag{2.13.1}$$

Here γ is the same constant that we encountered in exercise 2.2a. To prove this we proceed as follows:

After (2.8.3), let $E(N) = \sum_{p\leq N} 1/p - \log\log N - c$ so that $E(N) \ll 1/\log N$. We substitute this in below so as to obtain for any $\delta > 0$ that

$$\sum_{p \text{ prime}} \frac{1}{p^{1+\delta}} = \delta \int_2^\infty \frac{1}{t^{1+\delta}} \sum_{\substack{p \text{ prime} \\ p\leq t}} \frac{1}{p}\, dt = \delta \int_2^\infty \frac{\log\log t + c}{t^{1+\delta}} dt + \delta \int_2^\infty \frac{E(t)}{t^{1+\delta}} dt.$$

Now making the change of variable $u = \delta \log t$ we obtain

$$\delta \int_1^\infty \frac{\log\log t + c}{t^{1+\delta}} dt = \int_0^\infty \frac{\log(u/\delta)+c}{e^u} du = c - \log\delta + \int_0^\infty \frac{\log u}{e^u} du.$$

We shall prove in exercise 7.9a that $\int_0^\infty e^{-u}\log u \ du = -\gamma$. It is not difficult to show that the contribution of the integral between 1 and 2 is $\ll \delta$, and using the fact that $E(t) \ll 1/\log t$ we obtain

$$\delta \int_2^\infty \frac{E(t)}{t^{1+\delta}} dt \ll \delta \int_2^{e^{1/\delta}} \frac{1}{t\log t} dt + \delta \int_{e^{1/\delta}}^\infty \frac{1}{t^{1+\delta}\log t} dt \ll \delta \log(1/\delta).$$

Now, by exercise 2.2b, we know that $\zeta(1+\delta) = 1/\delta + O(1)$ so that $\log\zeta(1+\delta) = \log(1/\delta) + O(\delta)$. Combining all these estimates we obtain

$$\sum_{p \text{ prime}} \left\{ \frac{1}{p^{1+\delta}} + \log\left(1-\frac{1}{p^{1+\delta}}\right)\right\} = c - \gamma + O(\delta\log(1/\delta)).$$

Letting $\delta \to 0$, we note that the left side converges so that

$$\sum_{p \text{ prime}} \left\{ \frac{1}{p} + \log\left(1-\frac{1}{p}\right)\right\} = c - \gamma. \tag{2.13.2}$$

Now the terms with $p > x$ contribute $\ll \sum_{p>x} 1/p^2 \ll 1/x$ to the left side, and so we can truncate that sum at x, and use (2.8.3) to obtain

$$\sum_{\substack{p \text{ prime} \\ p\leq x}} \log\left(1-\frac{1}{p}\right) = -\log\log x - \gamma + O\left(\frac{1}{\log x}\right).$$

Exponentiating gives Mertens' theorem.

Exercises

2.13a. Use Mertens' theorem and (2.7.1) to prove that $\phi(n) \geq \{e^{-\gamma} + o(1)\}n/\log\log n$.

2.14. A few more analytic estimates.

What is the average number of divisors of integers up to x? The easiest way to do this is to write the appropriate sums out, using exercise 2.2a:

$$\sum_{n\leq x}\sum_{d|n} 1 = \sum_{d\leq x}\sum_{\substack{n\leq x\\ d|n}} 1 = \sum_{d\leq x}\left[\frac{x}{d}\right] = \sum_{d\leq x}\left(\frac{x}{d} + O(1)\right) = x\sum_{d\leq x}\frac{1}{d} + O(x) = x(\log x + O(1)).$$

Dirichlet, however, noted a nice trick to improve the error term here: The poor error term was caused by summing over the integers d all the way up to x. What Dirichlet noted was that divisors come in pairs $ab = n$ with $a \leq b$; so instead of counting 1 for each of a and b, rather count 2 for a (unless it is $= b = \sqrt{n}$ in which case we count 1). Therefore, using exercise 2.2a,

$$\begin{aligned}\sum_{n\leq x}\sum_{d|n} 1 &= \sum_{n\leq x}\sum_{\substack{d|n\\ d<\sqrt{n}}} 2 + \sum_{\substack{a\geq 1\\ a^2=d\leq x}} 1 = \sum_{d<\sqrt{x}}\sum_{\substack{d^2<n\leq x\\ d|n}} 2 + [\sqrt{x}] = 2\sum_{d<\sqrt{x}}\left(\left[\frac{x}{d}\right] - d\right) + O(\sqrt{x})\\ &= 2\sum_{d<\sqrt{x}}\left(\frac{x}{d} - d + O(1)\right) + O(\sqrt{x}) = 2x\sum_{d<\sqrt{x}}\frac{1}{d} - x + O(\sqrt{x})\\ &= x(\log x + 2\gamma - 1) + O(\sqrt{x}).\end{aligned}$$

A remarkable improvement in the error term! Getting an even better error term is an important open challenge.

What proportion of pairs of integers are pairwise coprime? In combinatorics the inclusion-exclusion principle often boils down to the identity $(1-1)^n = 1$ if $n = 1$ and 0 otherwise. In analytic number theory it often boils down to the identity

$$\sum_{d|m}\mu(d) = \begin{cases} 1 & \text{if } m = 1\\ 0 & \text{otherwise.}\end{cases}$$

So, to determine the number of pairs of integers up to x that are pairwise coprime we have

$$\sum_{\substack{a,b\leq x\\ (a,b)=1}} 1 = \sum_{a,b\leq x}\sum_{d|(a,b)}\mu(d) = \sum_{d\leq x}\mu(d)\sum_{\substack{a,b\leq x\\ d|(a,b)}} = \sum_{d\leq x}\mu(d)\left[\frac{x}{d}\right]^2.$$

Finally we replace $[t]$ by $t + O(1)$ and it is the readers challenge to complete the proof that this comes to $cx^2 + O(x\log x)$ where $c = \prod_{p \text{ prime}}(1 - 1/p^2)$, so that the proportion of pairs of integers that are pairwise coprime is c (which happens to equal $6/\pi^2$).

What is the average size of the gcd of two integers $\leq x$? We saw that it paid off above to swap the order of summation. In this problem we can do this if we write $g = \sum_{d|g}\phi(d)$

where $g = (a, b)$. Prove that the average is $c(x) + O(1)$ where $c(x) := \sum_{d \leq x} \phi(d)/d^2$. Prove that $c(x) \to \infty$ as $x \to \infty$, but that $c(x) \leq \log x + 1$. In chapter 7 we will see that $c(x) = c \log x + O(1)$, and therefore our average is $c \log x + O(1)$.

What is the average size of the gcd of three integers $\leq x$? Or four integers? Or any k integers $\leq x$?, where $k \geq 3$ is fixed.