Analytic Number Theory Solutions

Sean Li
Cornell University
sxl6@cornell.edu

Jan. 2013

Introduction

This document is a work-in-progress solution manual for Tom Apostol's *Introduction to Analytic Number Theory*. The solutions were worked out primarily for my learning of the subject, as Cornell University currently does not offer an analytic number theory course at either the undergraduate or graduate level. However, this document is public and available for use by anyone. If you are a student using this document for a course, I recommend that you first try work out the problems by yourself or in a group. My math documents are stored on a math blog at www.epicmath.org.

4 Some Elementary Theorems on the Distribution of Prime Numbers

4.1. Let $S = \{1, 5, 9, 13, 17, \dots\}$ denote the set of all positive integeers of the form $4n + 1$. An element p of S is called an *S-prime* if $p > 1$ and if the only divisors of p among the elements of S, are 1 and p. (For example, 49 is an S-prime.) An element $n > 1$ in S which is not an S-prime is called an *S-composite*.

(a) Prove that every S-composite is a product of S-primes.

Proof. Let $a \in S$ be an S-composite. Then we can write $a = bc$, where $b, c \in S$, $b, c \neq 1$. Hence we have $b, c < a$, and we may repeat the decomposition to b and c, ending when we get to primes. This process terminates since the number of elements in S less than a is finite, and at each stage of the decomposition we either have a prime or two smaller elements of S.

(b) Find the smallest S-composite that can be expressed in more than one way as a product of S-primes.

Solution: $693 = 21 \cdot 33 = 9 \cdot 77 = 3^2 \cdot 7 \cdot 11$. The idea is to use the lowest prime numbers which are not in S but whose products are primes in S. The lowest such primes are 3, 7, 11 as $3^2 = 9, 3 \cdot 7 = 21, 3 \cdot 11 = 33, 7 \cdot 11 = 77$ are all in primes in S. This gives the solution 693.

4.2. Consider the following finite set of integers:

$$T = \{1, 7, 11, 13, 17, 19, 23, 29\}.$$

(a) For each prime p in the interval $30 < p < 100$ determine a pair of integers m, n, where $m \geq 0$ and $n \in T$, such that $p = 30m + n$.

Solution.

$$\begin{array}{lll} 31 = 1 \cdot 30 + 1 & 53 = 1 \cdot 30 + 23 & 73 = 2 \cdot 30 + 13 \\ 37 = 1 \cdot 30 + 7 & 59 = 1 \cdot 30 + 29 & 79 = 2 \cdot 30 + 19 \\ 41 = 1 \cdot 30 + 11 & 61 = 2 \cdot 30 + 1 & 83 = 2 \cdot 30 + 23 \\ 43 = 1 \cdot 30 + 13 & 67 = 2 \cdot 30 + 7 & 89 = 2 \cdot 30 + 29 \\ 47 = 1 \cdot 30 + 17 & 71 = 2 \cdot 30 + 11 & 97 = 3 \cdot 30 + 7 \end{array}$$

(b) Prove the following statement or exhibit a counter example: Every prime $p > 5$ can be expressed in the form $30m + n$, where $m \geq 0$ and $n \in T$.

Proof. Let $n = 30m + r$, where $0 \leq r < 30$. We exclude all cases that force n composite. Clearly $2|r$ implies $2|n$, $3|r$ implies $3|n$, and $5|r$ implies $5|n$. This excludes the numbers

$$r = 0, 2, 3, 4, 5, 6, 8, 9, 10, 12, 14, 15, 16, 18, 20, 21, 22, 24, 25, 26, 27, 28.$$

Hence the remaining values that a prime number must take are

$$r = 1, 7, 11, 13, 17, 23, 29,$$

exactly as desired.

4.3. Let $f(x) = x^2 + x + 41$. Find the smallest integer $x \geq 0$ for which $f(x)$ is composite.

Solution: 40. Note that $f(40) = 40^2 + 40 + 41 = 41^2$ is composite. If $0 \leq x \leq 39$, then $f(x)$ is prime. One can check this exhaustively.

4.4. Let $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ be a polynomial with integer coefficients, where $a_n > 0$ and $n \geq 1$. Prove that $f(x)$ is composite for infinitely many n.

Proof. Suppose $a_0 \neq 0$. Then $a_0 | f(a_0), a_0 | f(2a_0), a_0 | f(3a_0), \ldots$. If $a_0 = 0$, then let i be the lowest power for which there is a nonzero coefficient. This is guaranteed to exist as $a_n > 0$. Then $a_i | f(a_i), a_i | f(2a_i), \ldots$.

4.5. Prove that for every $n \geq 1$ there exist n consecutive composite numbers.

Proof. Given $n \geq 1$, consider the sequence of consecutive numbers $(n+1)! + 2, (n+1)!+3, \ldots, (n+1)!+(n+1)$. Note that $2|(n+1)!+2, 3|(n+1)!+3, \ldots, (n+1)|(n+1)!+(n+1)$, thus this sequence contains n consecutive composite numbers.

4.6. Prove that there do not exist polynomials P and Q such that

$$\pi(x) = \frac{P(x)}{Q(x)} \text{ for } x = 1, 2, 3, \ldots$$

Proof. This is impossible because the asymptotics are different. From the prime number theorem we have that

$$\pi(x) \sim \frac{x}{\log x}.$$

Letting m be the degree of $P(x)$ and a_m be the coefficient of the x^m term, and letting n be the degree of $Q(x)$ and b_n be the coefficient of its x^n term, we have

$$\frac{P(x)}{Q(x)} \sim \frac{a_m x^m}{b_n x^n} = \frac{a^m}{b^n} x^{m-n}.$$

4.7. Let $a_1 < a_2 < \cdots < a_n \leq x$ be a set of positive integers such that no a_i divides the product of the others. Prove that $n \leq \pi(x)$.

We first prove a lemma.

Lemma: Let A be an $n \times m$ matrix with nonnegative integer entries that satisfies condition H, defined as follows: for each row i, there is a column j such that the element in the (i, j) position is greater than the sum of the elements in column j in all the other rows, i.e. for each j, there exists an i such that $a_{i,j} > \sum_{k \neq i} a_{k,j}$. Then $n \leq m$.

Proof of Lemma. We proceed by induction. When $n = 1$ the lemma is trivial. Suppose then that the lemma holds for $n - 1$, we show it holds for n, $n \geq 2$. In general, we note the following: Given an $n \times m$ matrix that satisfies condition H, given an arbitrary column, removing that column and the row that contains the largest element in that column results in a $(n-1) \times (m-1)$ matrix that satisfies condition H. The proof is by contradiction. Suppose some row in the resulting matrix contained entries that were $\leq$ the sum of the entries in that column in all the other rows. Then that row in the original matrix also contains every entry $\leq$ the sum of the entries in that column in all the other rows, including in the column that was removed.

We then take an arbitrary $n \times m$ matrix, and suppose that it satisfies condition H but $n > m$. Removing a column and the row that contains the largest entry in that column results in a $(n-1) \times (m-1)$ matrix that also satisfies condition H, which contradicts the induction hypothesis as $n-1 > m-1$. Hence, $n \leq m$.

Proof of Exercise. Let there be n positive integers in our set $a_1, \dots, a_n$. There are precisely $\pi(x)$ primes which are $\leq x$. Let M be the $n \times \pi(x)$ matrix where each row consists of the exponents of prime powers in the prime factorization number for an a_i. That is, let $a_i = \prod_j p_j^{b_{i_j}}$, then the (i, j)th entry in the matrix is b_{i_j}. The condition that no a_i divides the product of the others is equivalent to condition H on the matrix. Then by the lemma, we have $n \leq \pi(x)$.

4.8. Calculate the highest power of 10 that divides 1000!.

Solution: 249. Denote by k the highest power of 10 that divides 1000!, and let m be the highest power of 5 that divides 1000! and n be the highest power of 2 that divides 1000!. Then $k = \min\{m, n\}$. Since $5 > 2$, we have $m \leq n$, so it suffices to determine m. The algorithm is to sum through all the numbers from 1 to 1000, adding the highest power of 5 for each multiple of 5. This is the same as summing from 1 to 1000 several times, first adding 1 for each multiple of 5, then adding 1 for each multiple of 5^2, then 1 for each multiple of 5^3, etc. This gives

$$\begin{aligned} m &= \sum_{\substack{i=1 \\ 5|i}}^{1000} 1 + \sum_{\substack{i=1 \\ 25|i}}^{1000} 1 + \sum_{\substack{i=1 \\ 125|i}}^{1000} 1 + \sum_{\substack{i=1 \\ 625|i}}^{1000} 1 \\ &= \left[\frac{1000}{5}\right] + \left[\frac{1000}{25}\right] + \left[\frac{1000}{125}\right] + \left[\frac{1000}{625}\right] \\ &= 200 + 40 + 8 + 1 \\ &= 249. \end{aligned}$$

Note: The same method shows that

$$\begin{aligned} n = &\left[\frac{1000}{2}\right] + \left[\frac{1000}{4}\right] + \left[\frac{1000}{8}\right] + \left[\frac{1000}{16}\right] + \left[\frac{1000}{32}\right] \\ &+ \left[\frac{1000}{64}\right] + \left[\frac{1000}{128}\right] + \left[\frac{1000}{256}\right] + \left[\frac{1000}{512}\right] \\ = &\ 500 + 250 + 125 + 62 + 31 + 15 + 7 + 3 + 1 \\ = &\ 994. \end{aligned}$$

4.9. Given an arithmetic progression of integers

$$h, h+k, h+2k, \ldots, h+nk, \ldots,$$

where $0 < k < 2000$. If $h+nk$ is prime for $n = t, t+1, \ldots t+r$ prove that $r \leq 9$. In other words, at most 10 consecutive terms of this progression can be primes.

Proof. Suppose we had 11 or more consecutive terms of such a progression all prime. We go through the primes ≤ 11. Since there are at least 11 consecutive, we have $2 \nmid n + kt$ and $2 \nmid n + k(t+1)$, this implies $2|k$. Similarly, we have $3 \nmid n+kt, n+k(t+1), n+k(t+2)$, thus $3|k$. Continuing, we have $5|k, 7|k$, and $11|k$. This implies $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 = 2310|k$, which contradicts $0 < k < 2000$.

4.10. Let s_n denote the nth partial sum of the series

$$\sum_{r=1}^{\infty} \frac{1}{r(r+1)}.$$

Prove that for every integer $k > 1$ there exist integers m and n such that $s_m - s_n = 1/k$.

Proof. Note that we may rewrite

$$\frac{1}{r(r+1)} = \frac{1}{r} - \frac{1}{r+1}.$$

Hence the sum is actually a telescoping sum given by

$$s_n = \sum_{r=1}^{n} \left(\frac{1}{r} - \frac{1}{r+1} \right) = 1 - \frac{1}{n+1}.$$

We have $1/(2k) + 1/(2k) = 1/k$, so let $m = -2k - 1$ and $n = 2k - 1$.

4.11. Let s_n denote the sum of the first n primes. Prove that for each n there exists an integer whoses square lies between s_n and s_{n+1}.

Proof. The gap between successive squares n^2 and $(n+1)^2$ is $2n-1$. It suffices to show that p_n is greater than $2n-1$ after a certain point. This is obvious, as $2n-1$ enumerates the odd numbers, while p_n enumerates the primes which are all odd except 2. Indeed, after $p_4 = 2(4) - 1 = 7$, $p_n > 2n+1$. We still need to check that the claim holds for $2, 3, 5, 7$. In these cases, we have $s_n = 2, 5, 10, 17$, and the squares $4, 9, 16$ fit between them. Thus there is always a square between s_n and s_{n+1}.

4.12.–4.16. Prove each of the statements in Exercise 12 through 16. In this group of exercises you may use the prime number theorem.

4.12. If $a > 0$ and $b > 0$, then $\pi(ax)/\pi(bx) \sim a/b$ as $x \to \infty$.

Proof. By the prime number theorem we have

$$\begin{aligned}\frac{\pi(ax)}{\pi(bx)} &\sim \frac{ax\log bx}{bx\log ax} \\ &= \frac{ax(\log b+\log x)}{bx(\log a+\log x)} \\ &= \frac{a}{b}\end{aligned}$$

as $(\log b + \log x)/(\log a + \log x) \to 1$ as $x \to \infty$.

4.13. If $0 < a < b$, there exists an x_0 such that $\pi(ax) < \pi(bx)$ if $x \geq x_0$.

Proof. By part (a) we have

$$\frac{\pi(bx)}{\pi(ax)} \sim \frac{b}{a} > 1.$$

That is, the asymptotic gives that for $\epsilon > 0$, there exists M such that for all $x \geq M$,

$$\frac{b}{a}(1-\epsilon) < \frac{\pi(bx)}{\pi(ax)} < \frac{b}{a}(1+\epsilon).$$

Choosing an ϵ such that $(1-\epsilon)b/a > 1$ gives the desired result.

4.14. If $0 < a < b$, there exists an x_0 such that for all $x \geq x_0$ there is at least one prime between ax and bx.

Proof. It follows directly from Exercise 4.13.

4.15. Every interval $[a, b]$ with $0 < a < b$ contains a rational number of the form p/q, where p and q are primes.

Proof. Let $d = b - a$. Choose prime n sufficiently large enough such that $1/n < d/2$. Then there exist integers c and d such that $a \leq c/n < d/n \leq b$. By Exercise 4.14, there exists sufficiently large m_0 such that for integers $x \geq m_0$ there is at least one prime between cx and dx. Let y be such a prime when $x = m_0$, then we have $a \leq y/(nm_0) \leq b$ where y and n are prime. Now, let q be a prime larger than nm_0, and let l be such that $q = lm_0n$. Note that $l > 1$. Then we have

$$a \leq \frac{c}{n} = \frac{clm_0}{q} < \frac{p}{q} < \frac{dlm_0}{q} = \frac{d}{n} \leq b,$$

where the existence of p is guaranteed, as by Exercise 4.14, there must exist a prime between clm_0 and dlm_0.

4.16.

(a) Given a positive integer n there exists a positive integer k and a prime p such that $10^k n < p < 10^k(n+1)$.

Proof. This follows from Exercise 4.14. We have $0 < n < n+1$, hence there is large enough x_0 such that for all $x \geq x_0$, there is at least one prime between nx and $(n+1)x$. Let k be such that $10^k \geq x$, and the proof is done.

(b) Given m integers $a_1, \ldots, a_m$ such that $0 \leq a_i \leq 9$ for $i = 1, 2, \ldots, m$, there exists a prime p whose decimal expansion has $a_1, \ldots, a_m$ for its first m digits.

Proof. Let $n = \sum_{i=1}^{m} a_i 10^{m-i}$. The result then follows immediately from part (a).

4.17. Given an integer $n > 1$ with two factorizations $n = \prod_{i=1}^{r} p_i$ and $n = \prod_{i=1}^{t} q_i$ where the p_i are primes (not necessarily distinct) and the q_i are arbitrary integers > 1. Let α be a nonnegative integer.

(a) If $\alpha \geq 1$ prove that

$$\sum_{i=1}^{r} p_i^{\alpha} \leq \sum_{i=1}^{t} q_i^{\alpha}.$$

Proof. Each q_i is the product of p_i's, say $q_i = \prod_j p_{i_j}$. Then each $q_i^{\alpha} \geq \sum_j p_{i_j}^{\alpha}$, so the equation holds.

(b) Obtain a corresponding inequality relating these sums if $0 \leq \alpha < 1$.

Solution. The inequality in part (a) is still valid for low n if α is close to 1, yet the inequality is clearly reversed if $\alpha = 0$.

4.18. Prove that the following relations are equivalent:

(a) $\pi(x) = \dfrac{x}{\log x} + O\left(\dfrac{x}{\log^2 x}\right)$.

(b) $\vartheta(x) = x + O\left(\dfrac{x}{\log x}\right)$.

Proof. This follows from the relation $\pi(x) \sim \vartheta(x)/\log x$, which is a consequence of Theorem 4.4.

4.19. If $x \geq 2$, let

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t} (\text{the } \textit{logarithmic integral} \text{ of } x).$$

(a) Prove that

$$\mathrm{Li}(x) = \frac{x}{\log x} + \int_2^x \frac{dt}{\log^2 t} - \frac{2}{\log 2},$$

and that, more generally,

$$\mathrm{Li}(x) = \frac{x}{\log x}\left(1 + \sum_{k=1}^{n-1} \frac{k!}{\log^k x}\right) + n! \int_2^x \frac{dt}{\log^{n+1} t} + C_n,$$

where C_n is independent of x.

Proof. Using integration by parts, we obtain

$$\int \frac{dt}{\log t} = \frac{t}{\log t} + \int \frac{dt}{\log^2 t},$$

from which the first statement follows when making the integral definite from 2 to x. The general equation holds from repeatedly invoking integration by parts, and from the observation that

$$\frac{d}{dt}\frac{1}{\log^k t} = \frac{\log^{k-1} t}{t}\frac{1}{\log^{2k} t} = \frac{1}{t \log^{k+1} t}.$$

(b) If $x \geq 2$ prove that

$$\int_2^x \frac{dt}{\log^n t} = O\left(\frac{x}{\log^n x}\right).$$

Proof. This follows from integration by parts as in part (a).

4.20. Let f be an arithmetical fucntion such that

$$\sum_{p \leq x} f(p) \log p = (ax + b) \log x + cx + O(1) \text{ for } x \geq 2.$$

Prove that there is a constant A (depending on f) such that, if $x \geq 2$,

$$\sum_{p \leq x} f(p) = ax + (a + c)\left(\frac{x}{\log x} + \int_2^x \frac{dt}{\log^2 t}\right) + b \log(\log x) + A + O\left(\frac{1}{\log x}\right).$$

Proof. We use Abel's identity (Theorem 4.2). Letting $f(p)$ be the arithmetical function and $1/\log p$ as the function with a continuous derivative, we have

$$\sum_{p \leq x} f(p) = \frac{A(x)}{\log x} + \int_2^x \frac{A(t)}{t \log^2 t} dt + A_0,$$

where $A(x) = (ax + b) \log x + cx + O(1)$ and A_0 is a constant. The first term gives

$$\frac{A(x)}{\log x} = ax + b + \frac{cx}{\log x} + O\left(\frac{1}{\log x}\right),$$

while the integral term results in

$$\int_2^x \frac{A(t)}{t\log^2 t}dt = \int_2^x \frac{(at+b)\log t + ct + O(1)}{t\log^2 t}dt$$
$$= a\int_2^x \frac{dt}{\log t} + b\int_2^x \frac{dt}{t\log t} + c\int_2^x \frac{dt}{\log^2 t} + O\left(\frac{1}{\log x}\right).$$

Now, the integral of $dt/(t\log t)$ is $\log(\log t)$, so the second term is $b\log(\log x) - b\log(\log 2)$, and by Exercise 4.19, the first term is the logarithmic integral and is equal to

$$a\int_2^x \frac{dt}{\log t} = a\frac{x}{\log x} + a\int_2^x \frac{dt}{\log^2 t} - a\frac{2}{\log 2}.$$

Combining all the constants together into A gives

$$\sum_{p\le x} f(p) = ax + (a+c)\left(\frac{x}{\log x} + \int_2^x \frac{dt}{\log^2 t}\right) + b\log(\log x) + A + O\left(\frac{1}{\log x}\right).$$

4.21. Given two real-valued functions $S(x)$ and $T(x)$ such that

$$T(x) = \sum_{n\le x} S\left(\frac{x}{n}\right) \text{ for all } x \ge 1.$$

If $S(x) = O(x)$ and if c is a positive coonstant, prove that the relation

$$S(x) \sim cx \text{ as } x \to \infty$$

implies

$$T(x) \sim cx\log x \text{ as } x \to \infty.$$

Proof. Since $S(x) \sim cx$, we have $S(x/n) \sim cx/n$. Hence we have

$$T(x) \sim \sum_{n\le x} \frac{cx}{n} \sim cx\log x$$

as $\sum_{n\le x} 1/n \sim \log x$.

Exercises 22-30 are not available at this time.