Apostol's Analytic Number Theory

Stefanus[1]

[1] Samsung Semiconductor Inc

San Jose, CA 95134 USA

(Dated: January 31, 2017)

Abstract

Just for fun :)

Chapter 1

Problem 1.11. Prove that $n^4 + 4$ is composite if $n > 1$.

This problem can be easily solved by completing the square

$$n^4 + 4 = (n^2)^2 + 2^2$$
$$= (n^2 + 2)^2 - (2n)^2$$
$$= (n^2 + 2 - 2n)(n^2 + 2 + 2n)$$

But, suppose you're not aware of completing the square trick, here's what you can do. Since 4 is even if n is also even then $n^4 + 4$ is guaranteed to be composite. What if n is odd? Let's see

$$3^4 + 4 = 81 + 4 = 85$$
$$5^4 + 4 = 625 + 4 = 629$$
$$7^4 + 4 = 2401 + 4 = 2405$$
$$9^4 + 4 = 6561 + 4 = 6565$$
$$11^4 + 4 = 14641 + 4 = 14645$$
$$13^4 + 4 = 28561 + 4 = 28565$$
$$15^4 + 4 = 50625 + 4 = 50629$$

so we see that if $n \not\equiv 0 \pmod 5$ then $n^4 \equiv 1 \pmod{10}$ and therefore $n^4 + 4 \equiv 0 \pmod 5$ and if $n \equiv 0 \pmod 5$ then $n^4 + 4 \equiv -1 \pmod 5$, does this always apply? well, let's do everything in mod 5 then

$$n \equiv 1 \pmod 5 \rightarrow 1^4 + 4 \equiv 0 \pmod 5$$
$$n \equiv 2 \pmod 5 \rightarrow 2^4 + 4 \equiv 0 \pmod 5$$
$$n \equiv 3 \pmod 5 \rightarrow 3^4 + 4 \equiv 0 \pmod 5$$
$$n \equiv 4 \pmod 5 \rightarrow 4^4 + 4 \equiv 0 \pmod 5$$
$$n \equiv 5 \pmod 5 \rightarrow 5^4 + 4 \equiv 4 \pmod 5$$

So it is true, for $n \not\equiv 0 \pmod 5$, $n^4 + 4 \equiv 0 \pmod 5$ and thus $5 | n^4 + 4$ and $n^4 + 4$ is composite, except for $n = 1$ of course where $1^4 + 4 = 5$ which is prime.

Our task is to now see if we can prove that $n^4 + 4 \equiv 4 \pmod 5$ is composite. We know that if n is even then we are done, our task now is to see what happens if n is odd and a multiple of 5, *i.e.* to check whether $(5(2k + 1))^4 + 4$ is composite.

A strategy would be to show explicitly that we can factorize $(5(2k+1))^4 + 4 = (5a + 2)(5b+2)$. The way I did it was to try out different k's and see if there's a pattern to a and b

$$k = 0 \quad \rightarrow \quad a = 3 \quad , \ b = 7 \qquad \frac{a}{2k+1} = \underline{0}3 \ , \ b - a = 4 + 8 \cdot \underline{0}$$
$$k = 1 \quad \rightarrow \quad a = 39 \quad , \ b = 51 \qquad \frac{a}{2k+1} = \underline{1}3 \ , \ b - a = 4 + 8 \cdot \underline{1}$$
$$k = 2 \quad \rightarrow \quad a = 115 \quad , \ b = 135 \qquad \frac{a}{2k+1} = \underline{2}3 \ , \ b - a = 4 + 8 \cdot \underline{2}$$
$$k = 3 \quad \rightarrow \quad a = 231 \quad , \ b = 259 \qquad \frac{a}{2k+1} = \underline{3}3 \ , \ b - a = 4 + 8 \cdot \underline{3}$$
$$k = 4 \quad \rightarrow \quad a = 387 \quad , \ b = 423 \qquad \frac{a}{2k+1} = \underline{4}3 \ , \ b - a = 4 + 8 \cdot \underline{4}$$
$$k = 5 \quad \rightarrow \quad a = 583 \quad , \ b = 627 \qquad \frac{a}{2k+1} = \underline{5}3 \ , \ b - a = 4 + 8 \cdot \underline{5}$$
$$k = 6 \quad \rightarrow \quad a = 819 \quad , \ b = 871 \qquad \frac{a}{2k+1} = \underline{6}3 \ , \ b - a = 4 + 8 \cdot \underline{6}$$
$$k = 7 \quad \rightarrow \quad a = 1095 \ , \ b = 1155 \qquad \frac{a}{2k+1} = \underline{7}3 \ , \ b - a = 4 + 8 \cdot \underline{7}$$

so it seems like $a = (2k+1)(10k+3) = 20k^2 + 16k + 3$ and $b = a + 4 + 8k = 20k^2 + 24k + 7$ and indeed if we expand

$$(5a+2)(5b+2) = (5(20k^2 + 16k + 3) + 2)(5(20k^2 + 24k + 7) + 2)$$
$$= 10000k^4 + 20000k^3 + 15000k^2 + 5000k + 625 + 4$$
$$= (5(2k+1))^4 + 4$$

so indeed $(5(2k+1))^4 + 4$ can be factorized into $(5(20k^2 + 16k + 3) + 2)(5(20k^2 + 24k + 7) + 2)$ therefore they must be composite.

So in short, if $n \not\equiv 0 \pmod 5$ then $5 | n^4 + 4$ and if $5 | n$ then $(5(2k+1))^4 + 4 = (5(20k^2 + 16k + 3) + 2)(5(20k^2 + 24k + 7) + 2)$ and therefore $n^4 + 4$ is composite for $n > 1$.

Problem 1.15. Prove that every $n \geq 12$ is the sum of two composite numbers.

Here I assume the composite numbers are positive otherwise it is just a restatement of Bezout's lemma. Although this question looks tough initially it's actually very simple, for any even numbers ≥ 12 we have

$$2(6 + n) = 4 + 2(4 + n)$$

with $n \geq 0$ and for any odd number ≥ 13 we have

$$2(6 + n) + 1 = 9 + 2(2 + n)$$

again with $n \geq 0$ and the reason we start with 12 is because between 1 and 12 only $8 = 4 + 4$ and $10 = 4 + 6$ are sums of composite numbers but for 12 and above, every number is a sum of composite numbers.

Problem 1.16. Prove that if $2^n - 1$ is prime then n is prime

The trick we need to know here is that

$$x^n - 1 = (x - 1)(x^{n-1} + x^{n-2} + \cdots + 1)$$

Thus if n is not prime we must have $n = ab$

$$2^n - 1 = 2^{ab} - 1 = (2^a)^b - 1$$
$$\rightarrow (2^a)^b - 1 = (2^a - 1)((2^a)^{b-1} + (2^a)^{b-2} + \cdots + 1)$$

and thus $2^n - 1$ is not prime, we can immediately generalize this to if $a^m - b^n$ is prime with $a, b > 1$ then $\gcd(m, n) = 1$ because otherwise

$$a^m - b^n = a^{m'd} - b^{n'd} = (a^{m'})^d - (b^{n'})^d$$
$$= \left(a^{m'} - b^{n'}\right)\left((a^{m'})^{d-1} + (a^{m'})^{d-2}(b^{n'}) + \cdots + (a^{m'})(b^{n'})^{d-2} + (b^{n'})^{d-1}\right)$$

Problem 1.17. Prove that if $2^n + 1$ is prime, then n is a power of 2.

The key point we need to know here is

$$x^m + 1 = (x^h + 1)(x^{m-h} - x^{m-h-h} + x^{m-h-h-h} - \ldots + 1)$$

since the signs alternate from plus to minus on the RHS, the above factorization can only occur if $h|m$ and that m/h is **odd**. Another piece of info we need is that $(x^h - 1) \nmid (x^m + 1)$ for any h due to the same reason and so we need not care about nor consider $x^h - 1$.

Now say n in the $2^n + 1$ is even $\rightarrow n = 2w \rightarrow 2^{2w} + 1$, now if w is odd then $2^{2w} + 1$ is divisible by $2^2 + 1$, if w is even we can repeat the process $w = 2t \rightarrow 2^{4t} + 1$ and again, if t is odd then $2^{4t} + 1$ is divisible by $2^4 + 1$ but if t is even then we again repeat the process. So the only way it is not divisible by $2^z + 1$ is that the remaining factors are all even and it only happens when n is a power of 2.

If n is odd and composite then all factors of n are odd and therefore $2^n + 1$ is divisible by $2^u + 1$ where $u|n$, the only case left is when n is prime but in this case $2^n + 1$ is

immediately divisible by $2^1 + 1$ which is kind of a cool fact actually that any $2^n + 1$ where n is prime is divisible by 3 :)

Problem 1.18. If $m \neq n$ compute $\gcd(a^{2m} + 1, a^{2n} + 1)$ in terms of a. [*Hint:* Let $A_n = a^{2n} + 1$ and show that $A_n | (A_m - 2)$ if $m > n$.]

Without loss of generality let's assume that $m > n$. If $d = \gcd(a^{2n} + 1, a^{2m} + 1)$ then $d|(a^{2m} + 1) - (a^{2n} + 1) \rightarrow d|a^{2n}(a^{2(m-n)} - 1)$, if $d|a^{2n}$ then since $d|(a^{2n} + 1)$ this means that $d|(a^{2n} + 1 - a^{2n}) \rightarrow d|1$ which is not that interesting.

The interesting part is when $d|(a^{2(m-n)} - 1)$ and since $m > n \rightarrow m = nq + r$, in this case

$$d|((a^{2(nq+r-n)} - 1) + (a^{2n} + 1))$$
$$d|(a^{2(n(q-1)+r)} + a^{2n})$$
$$d|a^{2n}(a^{2(n(q-2)+r)} + 1)$$

just like before, if $d|a^{2n}$ then $d|1$, so we ignore this case, the other case is $d|(a^{2(n(q-2)+r)} + 1)$. Now, if $q - 2 > 2$ we can repeat the process but this time by taking the difference

$$d|((a^{2(n(q-2)+r)} + 1) - (a^{2n} + 1))$$
$$d|(a^{2(n(q-2)+r)} - a^{2n})$$
$$d|a^{2n}(a^{2(n(q-3)+r)} - 1)$$

again, we can repeat the process but this time by taking the sum just like before, once we reach $n(q - q) + r = r$ (or if $q = 1$ to begin with we will reach this step immediately) since $0 \leq r < n$ we will now express $n = rq' + r'$ (depending on whether q was even or odd we might have plus or minus in front of the constant 1 for simplicity we will choose plus)

$$d|((a^{2n} + 1) - (a^{2r} + 1))$$
$$d|((a^{2(rq'+r')} + 1) - (a^{2r} + 1))$$
$$d|(a^{2(rq'+r')} - a^{2r})$$
$$d|a^{2r'}(a^{2(r(q'-1)+r')} - 1)$$

and the process continues with the same exact pattern as Euclid's algorithm for finding the gcd between two numbers and that's what we will find, the $\gcd(m, n) = g$, *i.e.* we

will eventually get $d|(a^{2g} \pm 1)$ we can of course continue one step further to get $d|0$ (a tautology) or $d|2$ but this depends on a few things that we'll discuss later.

The next piece of information we need is the following

$$a^{2s} - 1 = (a^{2r} + 1)(a^{2s-2r} - a^{2s-2r-2r} + a^{2s-2r-2r-2r} - \ldots - 1)$$

so $(a^{2r} + 1)|(a^{2s} - 1)$ if $s = rk$ where k is *even* because we want the constant 1 in the second bracket of the RHS to have a minus sign and the pattern of the second bracket is $+ - + - + - + - \ldots$, next,

$$a^{2s} + 1 = (a^{2r} + 1)(a^{2s-2r} - a^{2s-2r-2r} + a^{2s-2r-2r-2r} - \ldots + 1)$$

by the same token $(a^{2r} + 1)|(a^{2s} + 1)$ only if $s = rk$ where k is *odd* and last one $(a^{2r} - 1) \nmid (a^{2s} + 1)$ this is because the second bracket of the RHS must all have positive sign but the -1 requires a minus sign. And of course we know about

$$a^{2s} - 1 = (a^{2r} - 1)(a^{2s-2r} + a^{2s-2r-2r} + a^{2s-2r-2r-2r} + \ldots + 1)$$

as long as $r|s$.

Going back to our case $d|(a^{2g} \pm 1)$, first thing to note is that this means $d \leq a^{2g} \pm 1$. Next, let's tackle it case by case, first case, $d|(a^{2g}+1)$, we know that d divides both $a^{2m}+1$ and $a^{2n}+1$ and we also know that since $m = gk_m$ and $n = gk_n$ (note that $g = \gcd(m, n)$) if k_m and k_n are both odd then since $d \leq a^{2g} + 1$ and $(a^{2g} + 1)$ divides both this means that $a^{2m} + 1$ and $a^{2n} + 1$, $d = a^{2g} + 1$.

But if one of or both of k_m or k_n are even then we have something different :) Let's just assume that k_n is even (this will not change the conclusion and only simplify things), now let $h = \gcd(a^{2g} + 1, a^{2n} + 1)$ where $n = gk_n$, this also means

$$h|(a^{2n} + 1) - (a^{2g} + 1)$$
$$\rightarrow h|(a^{2g(k_n-1)} - 1)$$
$$h|(a^{2g(k_n-2)} + 1)$$
$$h|(a^{2g(k_n-3)} - 1)$$
$$\vdots$$
$$h|(a^{2g(k_n-k_n)} + 1)$$
$$\rightarrow h|2$$

6

the last constant 1 has a plus sign in front of it because k_n is even, this means that since $d|a^{2g} + 1$ and $d|a^{2n} + 1$, $d \leq h$ but since $h \leq 2 \to d \leq 2$, therefore if a is odd $a^w + 1$ is even and $d = 2$ (because d is the *greatest* common divisor) but if a is even then $a^w + 1$ is odd and $d = 1$.

If $d|a^{2g} - 1$ then we know that from the above discussion that $(a^{2g} - 1) \nmid (a^{2m} + 1)$, $(a^{2g} - 1) \nmid (a^{2n} + 1)$, however since $g|m$, $g|n$ the following is true $(a^{2g} - 1)|(a^{2m} - 1)$, $(a^{2g} - 1)|(a^{2n} - 1)$. But this means that any common divisor x between $a^{2g} - 1$ and $a^{2m,n} + 1$ must also divide $a^{2m,n} - 1$ that is

$$x|(a^{2m,n} + 1) - (a^{2m,n} - 1)$$

$$x|2$$

which means that $d \leq x \to d \leq 2$, so again depending on whether a is even or odd we get $d = 1$ or $d = 2$ respectively.

Problem 1.19. The *Fibonacci sequence* $1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$ is defined by the recursion formula $a_{n+1} = a_n + a_{n-1}$, with $a_1 = a_2 = 1$. Prove that $(a_n, a_{n+1}) = 1$ for each n.

The suitable action is to induce :) We know that $(a_1, a_2) = (1, 1) = 1$, that's the base case, now say it is true that up to $n = m$, $(a_{m-1}, a_m) = 1$, now assume the opposite for say $d = (a_m, a_{m+1}) > 1$ then

$$a_{m+1} = a_n + a_{m-1}$$
$$(a_{m+1} - a_n) = a_{m-1}$$
$$d(a'_{m+1} - a'_n) = a_{m-1}$$
$$d \mid a_{m-1}$$

thus $(a_{m-1}, a_m) = d > 1$ which is a contradiction, therefore $d = 1$

Problem 1.24. Prove the following multiplicative property of the gcd:

$$(ah, bk) = (a, b)(h, k) \left(\frac{a}{(a, b)}, \frac{k}{(h, k)} \right) \left(\frac{b}{(a, b)}, \frac{h}{(h, k)} \right)$$

In particular this shows that $(ah, bk) = (a, k)(b, h)$ whenever $(a, b) = (h, k) = 1$.

Let's solve this the way a physicist might have done it. First, the whole situation is symmetrical under $a \leftrightarrow h$ and $b \leftrightarrow k$ and $ah \leftrightarrow bk$ therefore a first guess would be

$$(ah, bk) = (a, b)(h, k)(a, k)(b, h)$$

but this immediately fails when $a = p, h = p, b = p, k = 1$

$$(p \cdot p, p \cdot 1) = (p, p)(p, 1)(p, 1)(p, p)$$
$$= p^2$$

so this means that we have to remove some overcountings, *i.e.* we have to divide the whole thing with something like this

$$(ah, bk) = (a, b)(h, k)\frac{(a, k)(b, h)}{(a, b)(h, k)}$$

we divide by both (a, b) and (h, k) because the $ab \leftrightarrow hk$ symmetry, but again this is overdoing it, for example for $a = pp, h = 1, b = p, k = p$

$$(pp \cdot 1, p \cdot p) = (pp, p)(1, p)\frac{(pp, p)(1, p)}{(pp, p)(1, p)}$$
$$= (p)(1)\frac{(p)(1)}{(p)(1)}$$
$$= p$$

the next choice that makes sense would then be to absorb those denominators into the numerators and to still respect the symmetry we will have

$$(ah, bk) = (a, b)(h, k)\left(\frac{a}{(a, b)}, \frac{k}{(h, k)}\right)\left(\frac{b}{(a, b)}, \frac{h}{(h, k)}\right)$$

and that turns out to be correct :) Well for consistency we should do the nested test

$$(a(hh'), b) = ((ah)h', b)$$
$$(a, b)(hh', 1)\left(\frac{a}{(a, b)}, \frac{1}{(hh', 1)}\right)\left(\frac{b}{(a, b)}, \frac{hh'}{(hh', 1)}\right) = (ah, b)(h', 1)\left(\frac{ah}{(ah, b)}, \frac{1}{(h', 1)}\right)\left(\frac{b}{(ah, b)}, \frac{h'}{(h', 1)}\right)$$
$$(a, b)\left(\frac{b}{(a, b)}, hh'\right) = (ah, b)\left(\frac{b}{(ah, b)}, h'\right)$$

we now reapply the same formula to both LHS and RHS to see if this is consistent, first the LHS

$$(a, b)\left(\frac{b}{(a, b)}, hh'\right) = (a, b)\left\{\left(\frac{b}{(a, b)}, h\right)\left(\frac{b}{(a, b)\left(\frac{b}{(a,b)}, h\right)}, h'\right)\right\}$$

8

and now the RHS

$$(ah, b)\left(\frac{b}{(ah,b)}, h'\right) = \left\{(a,b)\left(\frac{b}{(a,b)}, h\right)\right\}\left(\frac{b}{(a,b)\left(h, \frac{b}{(a,b)}\right)}, h'\right)$$

and they are apparently the same :)

There are two reasons I chose to do the above problem, one, it's fun to apply what I learned during my physics PhD on a math problem and two we will need this result to solve Problem 1.28.

Problem 1.27. (a) If $(a, b) = 1$ then for every $n > ab$ there exist positive x and y such that $n = ax + by$.

(b) If $(a, b) = 1$ there are no positive x and y such that $ab = ax + by$.

(a) This question is actually quite tricky, at first I tried using Bezout's lemma (or the extended gcd method) but didn't get anywhere fast, the problem is, say we start with Bezout, since $(a, b) = 1$ we have

$$as + bt = 1$$

where one of s, t must be negative, next we add ab to both sides

$$ab + as + bt = ab + 1$$

we now need to show that the negative one among s, t is smaller than a or b and therefore we can add ab to it and make it positive, but it's not that straightforward to show that this is the case, the break for me came when I realized that we can use congruence modulo mumbo jumbo :)

Without loss of generality assume $b > a$, next, we need to lay some ground work. First some notation, we call a set $S = \{0, 1, 2, \ldots, b-1\}$ as a set of the remainders of b. Now we can shift S by multiples of b to get say $T = nb + S = \{nb+0, nb+1, \ldots, nb+b-1\}$ where n is any integer and we call T to be equivalent to S because if we take the remainders of the elements of T with respect to b we'll get back S. We can also shift the elements by *different* multiples of b with the same result. Another property of the set of the remainders of b is that every element is unique. By unique we mean that the difference between any two elements is *not* a multiple of b. To simplify things we'll use the symbol $c \equiv d$ if c and

9

d differ by a multiple of b and therefore equivalent, and $c \not\equiv d$ if c and d do **not** differ by a multiple of b and therefore c and d are unique (not equivalent).

With the above set, we can now move forward to the next important result. If we multiply a set of remainders of b, say S with a number say a that is co-prime to b, *i.e.* $(a, b) = 1$, then aS is also a set of remainders of b and aS is equivalent to any set of remainders of b.

Proof. First we will show that the elements of aS are all unique as defined above, to prove this let's assume the opposite, say aw and au differ by a multiple of b, where $w, u \in S$ and $w \not\equiv u$ where the symbol $\not\equiv$ means w and u do **not** differ by a multiple of b, then

$$aw = bq + au$$

$$a(w - u) = bq$$

$$\rightarrow b \mid a(w - u)$$

but since $(a, b) = 1$ this means that $b|(w - u) \rightarrow w \equiv u$ but this is a contradiction since $w \not\equiv u$, therefore every member of aS is unique but since the total number of elements of $\#aS = \#S$, this means aS is also a set of remainders of b.

The next thing we need to note is that by the same argument above since $(a, b) = 1$, $aw \equiv 0$ only if $w \equiv 0$ and if $w \not\equiv 0$ then $aw \not\equiv 0$.

Using the above result, we can say that for every $0 < h < b$ (we skip $h = 0$ for now) we can express

$$ah = bf + r$$

where again $(a, b) = 1$ and $0 < r < b$. Now since $a, h > 0$ therefore r cannot be 0 (recall that $ah \equiv 0 \rightarrow r \equiv 0$ only if $h \equiv 0$) instead $0 < r < b$, this means that $f > 0$. Since both sides are positive $\rightarrow ah > bf$, but since $b > h$, for $ah > bf$, f must be smaller than a, this result $f < a$ is the key. We can now move things around

$$ah - bf = r$$

$$\rightarrow ah + b(a - f) = ab + r$$

10

since $f < a$, $(a - f) > 0$ and this is our positive x and y, almost. We need to show it for **all** $n > ab$. But we can now build the case for **all** n.

In the above we have shown that there are positive x and y such that $ax + by = n$ for $ab < n < ab + b$. First, the above can be easily modified for any $n = vab + r$, $v > 1$

$$(v - 1)ab + ah + b(a - f) = (v - 1)ab + ab + r$$
$$a((v - 1)b + h) + b(a - f) = vab + r$$

so the only cases left are the ones where $n = vab$, $v > 1$ but this can be easily dealt with

$$a(b) + b\{(v - 1)a\} = vab$$

and that's all there is to it :) how about $ax + by = ab$? that will be dealt with in part (b)

(b) Say assume the opposite that we have positive x and positive y

$$ab = ax + by$$
$$a(b - x) = by$$
$$\rightarrow a \mid y$$

and by the same token $b \mid x$ therefore

$$ab = a(bx') + b(ay')$$
$$\cancel{ab} = \cancel{ab}(x' + y')$$
$$1 = x' + y'$$

therefore one of x and y has to be zero and it is a contradiction since x and y have to be positive

Problem 1.28. If $a > 1$ then $(a^m - 1, a^n - 1) = a^{(m,n)} - 1$.

I have actually solved this problem when I was doing Stein's elementary number theory book, it's Problem 2.25 of that book but here I'll do it differently :)

Proposition 1.28.1, $\gcd(x^n + x^{n-1} + \ldots + x + 1, x - 1) = \gcd(x - 1, n + 1)$

Proof. Say there's a common divisor d of $\sum_{i=1}^{n+1} x^{i-1} = x^n + x^{n-1} + \ldots + x + 1$ and

11

$x - 1$, since it divides $x - 1$ it also divides

$$d | x - 1$$

$$d | (x - 1)(x + 1) = x^2 - 1$$

$$d | (x - 1)(x^2 + x + 1) = x^3 - 1$$

$$\vdots$$

$$d | (x - 1)(x^{n-1} + x^{n-2} + \ldots + 1) = x^n - 1$$

if we sum all of them, d will still be a divisor of the sum

$$\left(x^n + x^{n-1} + \ldots + 1 \right) - \left\{ (x^n - 1) + (x^{n-1} - 1) + \ldots + (x - 1) \right\} = n + 1$$

therefore $d | (n + 1)$ since this is true for any common divisor d, it is also true for the gcd so if $g = \gcd\left(\sum_{i=1}^{n+1} x^{i-1}, x - 1 \right)$ then $g | (n + 1)$ but this still doesn't guarantee that $g = \gcd(x - 1, n + 1)$ but we can turn the above argument on its head, say $h = \gcd(x - 1, n + 1)$, since h divides $x - 1$, just like before it must also divide $x^2 - 1, x^3 - 1, \ldots, x^n + 1$, if we sum all of these and also $n + 1$ we get

$$\left\{ (x^n - 1) + (x^{n-1} - 1) + \ldots + (x - 1) \right\} + (n + 1) = x^n + x^{n-1} + \ldots + 1$$

and so h must also divide $x^n + x^{n-1} + \ldots + 1$ and that is that :) well almost, just one more little detail, what this says is that $h | \sum_{i=1}^{n+1} x^{i-1}$ and $g | (n + 1)$ but it doesn't say $h = g$? say they are different, $h \neq g$, then one of them must be smaller, say $h > g$, now from the above discussion we know that h is also a common divisor of $\left(\sum_{i=1}^{n+1} x^{i-1}, x - 1 \right)$ but if $h > g$ then g is not the **greatest** common divisor and we can choose h to be the $\gcd(\sum_{i=1}^{n+1} x^{i-1}, x - 1)$, the other way is also true, if $g > h$ then since g is a common divisor of $(x - 1, n + 1)$ then we can choose g as the gcd instead of h and therefore $\gcd(\sum_{i=1}^{n+1} x^{i-1}, x - 1) = \gcd(x - 1, n + 1)$.

One important thing to note here is that **the gcd depends on the value of** x, this is somewhat interesting.

There's another way to see why the above is true, note that $\sum_{i=1}^{n+1} x^{i-1}$ is just a geometric series

$$\sum_{i=1}^{n+1} x^{i-1} = x^n + x^{n-1} + \cdots + x + 1 = \frac{x^{n+1} - 1}{x - 1}$$

now if we set $x = (n+1) + 1$, from the numerator we get

$$\{(n+1)+1\}^{n+1} - 1 = \binom{n+1}{0}(n+1)^{n+1} + \binom{n+1}{1}(n+1)^n + \ldots + \binom{n+1}{n}(n+1) + 1 - 1$$

if we divide the above with $x - 1 = (n+1) + 1 - 1 = (n+1)$ we get

$$\frac{x^{n+1} - 1}{x - 1} = \binom{n+1}{0}(n+1)^n + \binom{n+1}{1}(n+1)^{n-1} + \ldots + \binom{n+1}{n}$$

$$= \binom{n+1}{0}(n+1)^n + \binom{n+1}{1}(n+1)^{n-1} + \ldots + (n+1)$$

$$= (n+1)\left\{ \binom{n+1}{0}(n+1)^{n-1} + \binom{n+1}{1}(n+1)^{n-2} + \ldots + 1 \right\}$$

$$\rightarrow = (x-1)\left\{ \binom{n+1}{0}(x-1)^{n-1} + \binom{n+1}{1}(x-1)^{n-2} + \ldots + 1 \right\}$$

so the key here is that although the last binomial coefficient $\binom{n+1}{n}$ is independent of what value we choose for x, if we judiciously choose $x - 1$ we can pull out a factor of $x - 1$ out of the terms like above, on the other hand, if $x - 1$ has no common factor with $\binom{n+1}{n} = (n+1)$ then we cannot pull a common factor out and that's the key.

Proposition 1.28.2, $\gcd(x^a, \sum_{i=0}^{N} x^{b_i}) = 1$ where $a \geq 0, b_0 = 0, \; b_{i \neq 0} \neq 0$

Proof. Say we have a *non-composite* common divisor, d, of x^a and $\sum_{i=0}^{N} x^{b_i}$ (non-composite so we cover primes and 1), since d is non-composite and $d | x^a$ we have $d | x$ and

$$\sum_{i=0}^{N} x^{b_i} = 1 + d \sum_{i=1}^{N} d^{b_i - 1} x^{b_i}$$

therefore since $d | \sum_{i=0}^{N} x^{b_i} \rightarrow d | 1$, therefore since *every* non-composite divisor is 1 the $\gcd(x^a, \sum_{i=0}^{N} x^{b_i}) = 1$ (note that any other divisor is a (possibly duplicate) product of the non-composite divisors)

Proposition 1.28.3, $\gcd(x^n + x^{n-1} + \ldots + 1, x^m + x^{m-1} + \ldots + 1) = \sum_{i=0}^{\gcd(n+1,m+1)-1} x^i$

Proof. This one is actually quite interesting, the sequence $x^n + x^{n-1} + \ldots + x + 1$ can be thought of just a number $\rightarrow n + 1$, let's see a concrete example, say we have

$$x^{11} + x^{10} + x^9 + x^8 + x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$$

and we want to know what we get if we divide it by $x^2 + x + 1$, in this we can easily factorize the above

$$(x^2 + x + 1)(x^9 + x^6 + x^3 + 1)$$

13

and there you go, we can immediately see the pattern, if $(n+1)|(m+1)$ then

$$x^m + x^{m-1} + \ldots + x + 1 = (x^n + x^{n-1} + \ldots + x + 1)(x^{(n+1)\cdot(s-1)} + x^{(n+1)\cdot(s-2)} + \ldots + x^{(n+1)} + 1)$$

$$\rightarrow \sum_{i=0}^{m} x^i = \left(\sum_{j=0}^{n} x^j\right)\left(\sum_{k=0}^{s-1} x^{(n+1)\cdot k}\right)$$

where $s = (m+1)/(n+1)$ but this also means something else, say we have $\sum_{i=0}^{n} x^i$, we can always produce a higher powered sequence by multiplying it with another sequence according to the rules above, *i.e.*

$$\left(\sum_{j=0}^{n} x^j\right)\left(\sum_{k=0}^{w} x^{(n+1)\cdot k}\right) = \sum_{i=0}^{((n+1)\cdot(w+1))-1} x^i$$

since what matters here is whether $(n+1)|(m+1)$, let's come up with a better notation to eliminate this ± 1 offset, say $f_n = \sum_{i=0}^{n-1} x^i$ and $g_w = \sum_{i=0}^{w-1} x^{(n+1)\cdot i}$ such that

$$f_n \cdot g_w = f_{n\cdot w}$$

in other words what we care most here is the number of terms in the sequence and **not** the highest power of the sequence. With this notation we can treat these sequences just like natural numbers, the most important thing we want to apply to them is Bezout's lemma that says for any number a and b there are c and d such that

$$ac + bd = \gcd(a, b)$$

we are now ready to utilize Bezout, say we have two sequences f_m and f_n where $\gcd(m, n) = 1$ and say we have a *non-composite* common divisor, $d \rightarrow d|f_m, \; d|f_n$, of those two (non-composite because I want to include primes and 1, recall that 1 is not prime nor composite), this means that

$$d|f_n \cdot g_y = f_{n\cdot y}$$
$$d|f_m \cdot g_z = f_{m\cdot z}$$

such that $ny - mz = 1$ which is guaranteed by Bezout (here we have assumed that $ny > mz$ but the conclusion still applies even if $mz > ny$), now let's take the difference

$$f_{n\cdot y} - f_{m\cdot z} = (x^{ny-1} + x^{ny-2} + \ldots + x + 1) - (x^{mz-1} + x^{mz-2} + \ldots + x + 1)$$
$$= (x^{mz} + x^{mz-1} + \ldots + x + 1) - (x^{mz-1} + x^{mz-2} + \ldots + x + 1)$$
$$= x^{mz}$$

which in turn means

$$d | f_{n \cdot y} - f_{m \cdot z}$$
$$\rightarrow d | x^{mz}$$

now since d is non-composite $\rightarrow d|x$ but since $d|f_n$ this also means that $d|1$, since *every* non-composite common divisor of f_n and f_m, $\gcd(f_n, f_m) = 1$ when $\gcd(m, n) = 1$.

Now if $g = \gcd(m, n) > 1$ what we have from Bezout is $ny - mz = g$

$$
\begin{aligned}
f_{n \cdot y} - f_{m \cdot z} &= (x^{ny-1} + x^{ny-2} + \ldots + x + 1) - (x^{mz-1} + x^{mz-2} + \ldots + x + 1) \\
&= (x^{mz+g-1} + x^{mz+g-2} + \ldots + x + 1) - (x^{mz-1} + x^{mz-2} + \ldots + x + 1) \\
&= x^{mz+g-1} + x^{mz+g-2} + \ldots + x^{mz} \\
&= x^{mz} \left(x^{g-1} + x^{g-2} + \ldots + 1 \right) \\
f_{n \cdot y} - f_{m \cdot z} &= x^{mz} \cdot f_g
\end{aligned}
$$

Also note that

$$f_n = f_g \cdot g_{n/g}$$
$$f_m = f_g \cdot g_{m/g}$$

this all means that since $g|f_n$, $g|f_m$

$$g | (f_{n \cdot y} - f_{m \cdot z})$$
$$g | (x^{mz} \cdot f_g)$$

Now $g = \gcd(f_n, f_m)$ and by definition gcd is divisible by all common divisors so in this case it is divisible by $f_g \rightarrow f_g | g$ and since $g|(x^{mz} \cdot f_g)$ and by Proposition 1.28.2 $\gcd(x^{mz}, f_g) = 1$ this means that $g = hf_g$ where $h|x^{mz}$ and we now need to figure out what h is.

Now note that $g|f_n \rightarrow hf_g|f_g \cdot g_{n/g} \rightarrow h|g_{n/g}$ but again by the proof Proposition 1.28.2 noting that $g_n = \sum_{i=0}^{N} x^{b_i}$, $\gcd(h, g_{n/g}) = 1 \rightarrow h = 1$ and thus $g = f_g$.

Therefore, the final conclusion is $\gcd(x^n + x^{n-1} + \ldots + 1, x^m + x^{m-1} + \ldots + 1) = \sum_{i=0}^{\gcd(n+1, m+1)-1} x^i$

Now back to **Problem 1.28**, first step is to utilize the result of Problem 1.24, let's denote $d = \gcd(m, n)$ and utilizing a similar notation to that of Proposition 1.28.3

$$(a^m - 1, a^n - 1) = ((a^d - 1)f_{m/d}, (a^d - 1)f_{n/d})$$

$$= (a^d - 1, a^d - 1)(f_{m/d}, f_{n/d}) \times$$
$$\left(\frac{a^d - 1}{(a^d - 1, a^d - 1)}, \frac{f_{n/d}}{(f_{m/d}, f_{n/d})}\right)\left(\frac{a^d - 1}{(a^d - 1, a^d - 1)}, \frac{f_{m/d}}{(f_{m/d}, f_{n/d})}\right)$$

$$= (a^d - 1)(f_{m/d}, f_{n/d})\left(1, \frac{f_{n/d}}{(f_{m/d}, f_{n/d})}\right)\left(1, \frac{f_{m/d}}{(f_{m/d}, f_{n/d})}\right)$$

$$= (a^d - 1)(f_{m/d}, f_{n/d})$$

where now $f_{m/d} = \left((a^d)^{m/d-1} + (a^d)^{m/d-2} + \ldots + a^d + 1\right)$, by Proposition 1.28.3 $\gcd(f_{m/d}, f_{n/d}) = \gcd(m/d, n/d) = 1$ and we are done :) well of course if we have chosen a, h, b, k differently we will not get the answer as easily, *e.g.*

$$(a^m - 1, a^n - 1) = ((a^d - 1)f_{m/d}, f_{n/d}(a^d - 1))$$

$$= (a^d - 1, f_{n/d})(f_{m/d}, a^d - 1) \times$$
$$\left(\frac{a^d - 1}{(a^d - 1, f_{n/d})}, \frac{a^d - 1}{(f_{m/d}, a^d - 1)}\right)\left(\frac{f_{n/d}}{(a^d - 1, f_{n/d})}, \frac{f_{m/d}}{(f_{m/d}, a^d - 1)}\right)$$

$$= (a^d - 1, n/d)(m/d, a^d - 1)\left(\frac{a^d - 1}{(a^d - 1, n/d)}, \frac{a^d - 1}{(m/d, a^d - 1)}\right)$$

$$= (a^d - 1)$$

from line 2 to 3 we have used the fact that since $(f_{m/d}, f_{n/d}) = 1$ therefore $(f_{m/d}/t, f_{n/d}/s) = 1$, going into the last line might need a bit more explanation, let's denote

$$(a^d - 1, n/d) = c_n$$
$$(a^d - 1, m/d) = c_m$$

now since $\gcd(m/d, n/d) = 1$ we have

$$a^d - 1 = c_a c_m c_n$$

therefore

$$(a^d - 1, n/d)(m/d, a^d - 1)\left(\frac{a^d - 1}{(a^d - 1, n/d)}, \frac{a^d - 1}{(m/d, a^d - 1)}\right) = c_n c_m(c_a c_m, c_a c_n)$$

$$= c_n c_m c_a$$

$$= a^d - 1$$

so depending on how stupidly you make your choices, you might have to suffer a bit more

:)

Chapter 2

Problem 2.1 Find all integers n such that

$$\text{(a)} \ \varphi(n) = n/2 \qquad \text{(b)} \ \varphi(n) = \varphi(2n) \qquad \text{(c)} \ \varphi(n) = 12$$

For (a), using the definition of $\varphi(n)$, $\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$

$$\frac{\not{n}}{2} = \not{n} \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

$$\frac{1}{2} = \frac{\prod_{p|n} (p-1)}{\prod_{p|n} p}$$

$$\prod_{p|n} p = 2 \prod_{p|n} (p-1)$$

if n is odd the LHS is odd while the RHS is even, so it can't be. If n is even the LHS only has one factor of 2 while the RHS has many so it will only work if $n = 2$.

For (b)

$$\not{n} \prod_{p|n} \left(1 - \frac{1}{p}\right) = 2\not{n} \prod_{p|2n} \left(1 - \frac{1}{p}\right)$$

If n is even then

$$\prod_{p|2n} \left(1 - \frac{1}{p}\right) = \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

and so

$$\prod_{p|n} \left(1 - \frac{1}{p}\right) = 2 \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

$$\rightarrow 1 = 2$$

which is impossible, so n has to be odd, in that case

$$\prod_{p|2n} \left(1 - \frac{1}{p}\right) = \left(1 - \frac{1}{2}\right) \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

$$= \frac{1}{2} \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

17

and therefore

$$\prod_{p|n}\left(1-\frac{1}{p}\right) = 2\frac{1}{2}\prod_{p|n}\left(1-\frac{1}{p}\right)$$
$$\rightarrow 1 = 1$$

and therefore $\varphi(n) = \varphi(2n)$ for all odd n.

For (c)

$$\varphi(n) = 12 = 2\cdot 2\cdot 3$$
$$= \prod_{p|n} p^{\alpha_p} - p^{\alpha_p-1}$$

$$\varphi\left(\prod_{p|n} p^{\alpha_p}\right) = \prod_{p|n} p^{\alpha_p-1}(p-1)$$

the only possible solution is $n = 13$

Problem 2.2. For each of the following statements either give a proof or exhibit a counter example.

(a) If $(m,n) = 1$ then $(\varphi(m), \varphi(n)) = 1$

(b) If n is composite, then $(n, \varphi(n)) > 1$

(c) If the same primes divide m and n, then $n\varphi(m) = m\varphi(n)$

For (a) a counter example will be $(3,4) = 1$, while $\varphi(3) = 2$, $\varphi(4) = 2$

For (b) a counter example would be $n = 15$ which means that $\varphi(15) = 8$ and $(15, 8) = 1$

For (c) I think what it means by "the same primes divide m and n" is that $m = \prod p^{\alpha_p}$ and $n = \prod p^{\beta_p}$, so they both have the same primes but they might have different exponents for each prime, in this case $\prod_{p|n} = \prod_{p|m}$

$$n\varphi(m) = n\left(m\prod_{p|m}\left(1-\frac{1}{p}\right)\right)$$
$$= m\left(n\prod_{p|n}\left(1-\frac{1}{p}\right)\right)$$
$$n\varphi(m) = m\varphi(n)$$

Problem 2.3. Prove that

$$\frac{n}{\varphi(n)} = \sum_{d|n}\frac{\mu^2(d)}{\varphi(d)}$$

18

Since $\mu(n)$ and $\varphi(n)$ are both multiplicative so is μ^2/φ, in that case $g(n) = \sum_{d|n} \frac{\mu^2(d)}{\varphi(d)}$ is also multiplicative. To determine $g(n)$ we need only compute $g(p^\alpha)$ for prime powers

$$
\begin{aligned}
g(p^\alpha) &= \sum_{d|p^\alpha} \frac{\mu^2(d)}{\varphi(d)} \\
&= \frac{\mu^2(1)}{\varphi(1)} + \frac{\mu^2(p)}{\varphi(p)} + \ldots + \frac{\mu^2(p^\alpha)}{\varphi(p^\alpha)} \\
&= 1 + \frac{1}{p-1} \\
&= \frac{p}{p-1} \\
&= p^\alpha \cdot \frac{p}{p^\alpha(p-1)} \\
\rightarrow \sum_{d|p^\alpha} \frac{\mu^2(d)}{\varphi(d)} &= \frac{p^\alpha}{\varphi(p^\alpha)}
\end{aligned}
$$

We can also prove it the other way around by assuming the LHS, to do this it is easiest to use the Mobius inversion formula

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

and we want to find out what this $g(d)$ is, which is

$$
g(n) = \sum_{d|n} \frac{d}{\varphi(d)} \mu\left(\frac{n}{d}\right)
$$

The RHS is multiplicative so like above we just need to evaluate $g(p^\alpha)$ for prime powers

$$
\begin{aligned}
g(p^\alpha) &= \sum_{d|p^\alpha} \frac{d}{\varphi(d)} \mu\left(\frac{p^\alpha}{d}\right) \\
&= \frac{p^{\alpha-1}}{\varphi(p^{\alpha-1})} \mu\left(\frac{p^\alpha}{p^{\alpha-1}}\right) + \frac{p^\alpha}{\varphi(p^\alpha)} \mu\left(\frac{p^\alpha}{p^\alpha}\right) \\
&= -\frac{p^{\alpha-1}}{\varphi(p^{\alpha-1})} + \frac{p^\alpha}{\varphi(p^\alpha)} \\
&= -\frac{p^\alpha}{\varphi(p^\alpha)} + \frac{p^\alpha}{\varphi(p^\alpha)} \\
&= 0
\end{aligned}
$$

19

if $\alpha > 1$ and if $\alpha = 1$ we get

$$g(p) = \sum_{d|p} \frac{d}{\varphi(d)} \mu\left(\frac{p}{d}\right)$$

$$= \frac{1}{\varphi(1)} \mu\left(\frac{p}{1}\right) + \frac{p}{\varphi(p)} \mu\left(\frac{p}{p}\right)$$

$$= -1 + \frac{p}{\varphi(p)}$$

$$= -1 + \frac{p}{p-1}$$

$$= \frac{1}{p-1}$$

$$g(p) = \frac{1}{\varphi(p)}$$

This means that $g(p^\alpha) = 1/\varphi(p^\alpha)$ is $\alpha = 1$ and $g(p^\alpha) = 0$ if $\alpha > 1$, in other words $g(p^\alpha) = \mu^2(p^\alpha)/\varphi(p^\alpha)$

Problem 2.4. Prove that $\varphi(n) > n/6$ for all n with at most 8 distinct prime factors.

First, let's demystify this number 8, the reason 8 is involved is because if you multiply out $(p-1)/p$ for the first eight primes we get

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdot \frac{10}{11} \cdot \frac{12}{13} \cdot \frac{16}{17} \cdot \frac{18}{19} = \frac{55296}{323323} \sim 0.171 > \frac{1}{6}$$

but if we multiply the first nine

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdot \frac{10}{11} \cdot \frac{12}{13} \cdot \frac{16}{17} \cdot \frac{18}{19} \cdot \frac{22}{23} = \frac{110592}{676039} \sim 0.164 < \frac{1}{6}$$

So that's how we got the eight and of course if we chose any other eight primes we will get something bigger than $55296/323323 > 1/6$ because $n/(n+1)$ converges to 1 as $n \to \infty$, i.e. $n/(n+1)$ gets bigger as n gets bigger.

Another reason we have to limit it to eight is because $n/(n+1) < 1$ and if we keep multiplying them we'll get a smaller and smaller number and after some point we will reach $< 1/6$.

The rest is straightforward,

$$\frac{\varphi(n)}{n} = \prod_{p|n} \frac{p-1}{p}$$

so the argument above holds

Problem 2.5. Define $\nu(1) = 0$, and for $n > 1$ let $\nu(n)$ be the number of distinct prime factors of n. Let $f = \mu * \nu$ and prove that $f(n)$ is either 0 or 1.

As the inverse of μ is $\mu^{-1} = u$, this means that

$$u * f = (u * \mu) * \nu$$

$$= I * \nu$$

$$u * f = \nu$$

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

ν is obviously not multiplicative since $\nu(1) \neq 1$, $\nu(pq) \neq \nu(p)\nu(q)$ but it is actually additive since $\nu(p^\alpha q^\beta) = \nu(p^\alpha) + \nu(q^\beta) = \nu(p) + \nu(q)$ where $p \neq q$ are distinct primes, so let's decompose n into its primal constituents, $n = \prod_i p_i^{\alpha_i}$

$$\nu \left(\prod_i p_i^{\alpha_i} \right) = \sum_{d|n} f(d)$$

$$\sum_i \nu(p_i^{\alpha_i}) = \sum_{d|n} f(d)$$

$$\sum_i \nu(p_i) = \sum_{d|n} f(d)$$

from here we can immediately see that $f(n)$ is given by

$$f(n) = \begin{cases} 1 \text{ if } n \text{ is prime} \\ 0 \text{ otherwise} \end{cases}$$

Problem 2.6. Prove that

$$\sum_{d^2|n} \mu(d) = \mu^2(n)$$

and, more generally,

$$\sum_{d^k|n} \mu(d) = \begin{cases} 0 & \text{if } m^k|n \text{ for some } m > 1 \\ 1 & \text{otherwise} \end{cases}$$

The last sum is extended over all positive divisors d of n whose kth power also divide n.

21

The key point here is again "multiplicative", since $\mu(d)$ is multiplicative so is $\sum_{d^2|n}\mu(d)$ so we need to only consider $g(p^\alpha) = \sum_{d^2|p^\alpha}\mu(d)$ but note that even though the sum is over $d^2 \to \sum_{d^2|n}$, μ is only taking d, $\mu(d)$ and not $\mu(d^2)$

$$\sum_{d^2|p^\alpha}\mu(d) = \mu(1) + \mu(p)$$
$$= 1 - 1$$
$$= 0$$

The above holds if $\alpha > 1$ otherwise for $0 \leq \alpha \leq 1 \to \sum_{d^2|p^\alpha}\mu(d) = \mu(1) = +1$, in short

$$g(p^\alpha) = \sum_{d^2|p^\alpha}\mu(d) = \begin{cases} 0 & \text{if } \alpha > 1 \\ 1 & \text{if } 0 \leq \alpha \leq 1 \end{cases}$$
$$= \mu^2(p^\alpha)$$

The second part follows closely, again since it is multiplicative and again note that even though the sum is over $d^k \to \sum_{d^k|n}$, μ is only taking d, $\mu(d)$ and not $\mu(d^k)$

$$\sum_{d^k|p^\alpha}\mu(d) = \mu(1) + \mu(p)$$
$$= 1 - 1$$
$$= 0$$

if $\alpha > k$ otherwise for $0 \leq \alpha \leq k \to \sum_{d^k|p^\alpha}\mu(d) = \mu(1) = +1$, the only difference now is that we can't say it is equal to $\mu^2(p^\alpha)$ because say $\alpha = k - 1 > 0 \to \mu(p^{k-1}) = 0$ but $\sum_{d^k|p^{k-1}}\mu(d) = \mu(1) = +1$

Problem 2.7. Let $\mu(p, d)$ denote the value of the Mobius function at the gcd of p and d. Prove that for every prime p we have

$$\sum_{d|n}\mu(d)\mu(p, d) = \begin{cases} 1 & \text{if } n = 1 \\ 2 & \text{if } n = p^a, a \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

The thing is the gcd (p, mn) is multiplicative as long as $(m, n) = 1$ because p is prime and once we expand m and n in their primal constituents it is evident, *i.e.* $(p, mn) = (p, m)(p, n)$, therefore $\mu(p, mn) = \mu(p, m)\mu(p, n)$

The first case is obvious $\sum_{d|1} \mu(d)\mu(p,d) = \mu(1)\mu(1) = 1$.

The second case

$$\sum_{d|p^a} \mu(d)\mu(p,d) = \mu(1)\mu(p,1) + \mu(p)\mu(p,p)$$

$$= \mu(1)\mu(1) + \mu(p)\mu(p)$$

$$= (1)(1) + (-1)(-1)$$

$$= 2$$

To show the last case it's easiest to utilize the fact that $g(n) = \sum_{d|n} \mu(d)\mu(p,d)$ is multiplicative and now we just need to show $g(q^b)$, $q \neq p$ as $g(p^a)$ is already covered above

$$g(q^b) = \sum_{d|q^b} \mu(d)\mu(p,d) = \mu(1)\mu(p,1) + \mu(q)\mu(p,q)$$

$$= \mu(1)\mu(1) + \mu(q)\mu(1)$$

$$= (1)(1) + (-1)(1)$$

$$= 0$$

Problem 2.8. Prove that

$$\sum_{d|n} \mu(d) \log^m d = 0$$

if $m \geq 1$ and n has more than m distinct prime factors. [*Hint:* Induction.]

To use induction we need to prove the base case, the thing is that log is not multiplicative, so that's a bit hard. The base case should be $m = 1$ and then we go up from there to bigger m ?!? ¯\\(°_o)/¯

But one thing I notice is that we only need to consider numbers with one power of distinct primes, *i.e.* $n = p_1 p_2 \ldots p_k$ because $\mu(d)$ is zero if the powers of the primes are not zero that is

$$\sum_{d|n} \mu(d) \log^m d = \cancel{\mu(1) \log^m(1)} + \mu(p_1) \log^m(p_1) + \ldots + \mu(p_k) \log^m(p_k) +$$

$$\mu(p_1 p_2) \log^m(p_1 p_2) + \ldots + \mu(p_{k-1} p_k) \log^m(p_{k-1} p_k) + \ldots +$$

$$\mu(p_1 p_2 \ldots p_k) \log^m(p_1 p_2 \ldots p_k)$$

23

and from the definition of $\mu(d)$ we know that if it has odd number of primes it's negative and it there are an even number of distinct primes μ is positive, therefore

$$\sum_{d|n} \mu(d) \log^m d = -(\log^m(p_1) + \ldots + \log^m(p_k))$$
$$+(\log^m(p_1 p_2) + \ldots + \log^m(p_{k-1} p_k)) +$$
$$-(\log^m(p_1 p_2 p_3) + \ldots + \log^m(p_{k-2} p_{k-1} p_k)) +$$
$$(-1)^k \log^m(p_1 p_2 \ldots p_k)$$

Since log is additive we can expand them but before we do that let's denote $\log(p_k) = l_k$

$$\sum_{d|n} \mu(d) \log^m d = -\sum_{i_1=(k|1)} l_{i_1}^m + \sum_{i_1,i_2=(k|2)} (l_{i_1} + l_{i_2})^m - \sum_{i_1,i_2,i_3=(k|3)} (l_{i_1} + l_{i_2} + l_{i_3})^m +$$
$$\ldots + (-1)^k \sum_{i_1,i_2,\ldots,i_k=(k|k)} (l_{i_1} + l_{i_2} + \ldots + l_{i_k})$$

where the notation $(k|j)$ means that all combinations of k choose j, as a concrete example, say $m = 4$, $k = 5$ which is the minimum k required

$$\sum_{d|n} \mu(d) \log^m d = -(l_1^4 + l_2^4 + l_3^4 + l_4^4 + l_5^4) +$$
$$+((l_1 + l_2)^4 + (l_1 + l_3)^4 + (l_1 + l_4)^4 + (l_1 + l_5)^4 + (l_2 + l_3)^4 + (l_2 + l_4)^4 +$$
$$(l_2 + l_5)^4 + (l_3 + l_4)^4 + (l_3 + l_5)^4 + (l_4 + l_5)^4) +$$
$$-((l_1 + l_2 + l_3)^4 + (l_1 + l_2 + l_4)^4 + (l_1 + l_2 + l_5)^4 + (l_1 + l_3 + l_4)^4 +$$
$$(l_1 + l_3 + l_5)^4 + (l_1 + l_4 + l_5)^4 + (l_2 + l_3 + l_4)^4 + (l_2 + l_3 + l_5)^4 +$$
$$(l_2 + l_4 + l_5)^4 + (l_3 + l_4 + l_5)^4)$$
$$+((l_1 + l_2 + l_3 + l_4)^4 + (l_1 + l_2 + l_3 + l_5)^4 + (l_1 + l_2 + l_4 + l_5)^4 +$$
$$(l_1 + l_3 + l_4 + l_5)^4 + (l_2 + l_3 + l_4 + l_5)^4) +$$
$$-((l_1 + l_2 + l_3 + l_4 + l_5)^4)$$

Now we gather coefficients of same powers, say we collect all l_1^4,

$$(5|1) \rightarrow (-1)l_1^4$$
$$(5|2) \rightarrow (+4)l_1^4$$
$$(5|3) \rightarrow (-6)l_1^4$$
$$(5|4) \rightarrow (+4)l_1^4$$
$$(5|5) \rightarrow (-1)l_1^4$$

so they're basically the Pascal triangle coefficients, why is this? Well, for example, for $(5|1)$, first we fix **one** l and then choose a partner for it from the remaining **four**, however in this case we only need one l and we already fixed it, so we will just need **zero** partner, *i.e.* $\binom{4}{0} = 1$.

For $(5|2)$ we first pick an l and then choose a partner (again because $(5|\mathbf{2})$ means we need **2** l's in total) for it from 4 available choices, which is $\binom{4}{1}$, *i.e.* this l will appear $\binom{4}{1} = 4$ times, for $(5|3)$ it's the same thing we first pick an l and then choose *two* partners for it, *i.e.* this l will then appear $\binom{4}{2} = 6$ times, and for $(5|3)$, it's pick an l and choose $\binom{4}{3} = 4$ partners and so on and therefore the coefficients of l_1 is just those of Pascal triangle's but with the signs alternating between plus and minus. And this is true for other l's not just l_1.

We now need to tackle the cross terms say $l_1^3 l_2$, first thing to note that this cross product is always preceded by a constant (which again is from Pascal triangle), for $(l_1 + l_2)^4$ it is $4 l_1^3 l_2$, note that this coefficient is the same no matter how many terms are being exponentiated, *i.e.* even for $(l_1 + l_2 + l_3 + \ldots + l_w)^4$, the coefficient for $l_1^3 l_2$ is still 4 because it is still $\binom{4}{3}$ no matter what, this is because

$$(l_1 + l_2 + \ldots)^4 = \underbrace{(l_1 + l_2 + \ldots)}_{\text{bin } \#1} \underbrace{(l_1 + l_2 + \ldots)}_{\text{bin } \#2} \underbrace{(l_1 + l_2 + \ldots)}_{\text{bin } \#3} \underbrace{(l_1 + l_2 + \ldots)}_{\text{bin } \#4}$$

To get $l_1^3 l_2$ we need to gather **three** l_1's and we have **four** bins to choose for as shown above that's why we have 4 choose 3, $\binom{4}{3} = 4$ possibilities. And as the number of bins are the same no matter how many l's we have the number of possibilities is still the same.

We also have other cross terms like $l_1^2 l_3 l_4$, in this case, we need to gather **two** l_1's from **four** bins so it's $\binom{4}{2} = 6$, next we need to choose **one** l_3 from the remaining **two** bins

which is $\binom{2}{1} = 2$ and once we've chosen the bin for l_2, the other bin will definitely contain l_3, so in total there are

$$\binom{4}{2} \times \binom{2}{1} = 6 \times 2 = 12$$

and since the number of bins is constant no matter what this coefficient remains the same no matter how many l's we have.

So now for $4l_1^3 l_2$ we have

$$(5|1) \rightarrow (0)$$
$$(5|2) \rightarrow (+1)4l_1^3 l_2$$
$$(5|3) \rightarrow (-3)4l_1^3 l_2$$
$$(5|4) \rightarrow (+3)4l_1^3 l_2$$
$$(5|5) \rightarrow (-1)4l_1^3 l_2$$

again Pascal triangle, why is this? This time we fix **two** l's (instead of just one for l^4 above), and then calculate how many partners this couple might have, for $(5|2)$, we only need **two** in total so because we already fixed two of them we just need **zero** partner from the three remaining ones, *i.e.* $\binom{3}{0} = 1$. For $(5|3)$, again we fix **two** l's and choose one more partner (because in total we need 3) from the remaining three, *i.e.* $\binom{3}{1} = 3$ and so on. This is also true for any two-term cross terms.

And this pattern continues for higher cross terms like $l_1^2 l_2 l_3$, *e.g.* for $(5|4)$ we fix **three** l's and then choose one partner from the remaining **two**, which means $\binom{2}{1} = 2$.

This pattern continues for any k, say we now have $k = 6$ while m stays the same, $m = 4$, in this case we have $(6|1), (6|2), (6|3), (6|4), (6|5), (6|6)$, and to get the coefficients for different l powers we use the same method as described above.

Say you want to know the coefficient l_1^4 for each $(6|1), (6|2), (6|3), (6|4), (6|5), (6|6)$, then fix an l and choose a partner for it depending on which combination $(6|j)$ you're on; for just a single l the combination is $\binom{6-1}{j-1}$ and for three l's like $l_1 l_2 l_3^2$ we fix three and then choose a partner resulting in $\binom{6-3}{j-3}$ combo.

And here we immediately see why the number of distinct primes k must be larger than m, the exponent of log, it's because if $k = m$ then on the last combo $(k = m|j = m)$

we will have $\binom{m-m}{m-m} = 1$ but these l's, $l_1^{a_1} l_2^{a_2} \ldots l_m^{a_m}$ can only be found once and there'll be nothing to cancel it, the same is true if $k < m$, the longest l combo $l_1^{a_1} l_2^{a_2} \ldots l_k^{a_k}$ is only generated once and there's nothing to cancel it to zero.

As a concrete example take $m = 3$ and $k = 2$, we will then have

$$-(l_1^3 + l_2^3) + (l_1 + l_2)^3 = 3l_1^2 l_2 + 3l_1 l_2^2$$

and for $m = 3$, $k = 3$

$$-(l_1^3 + l_2^3 + l_3^3) + ((l_1 + l_2)^3 + (l_1 + l_3)^3 + (l_2 + l_3)^3) - (l_1 + l_2 + l_3)^3 = -6l_1 l_2 l_3$$

so you see the longest l combo is not canceled whenever $k \leq m$. But if $k > m$ then the longest l combo is still m and for every combo of the form $(k|m \leq j \leq k)$ we have a coefficient of $\binom{k-m}{j-m}$ which is just the Pascal triangle for $(1 - 1)^{k-m} = 0$.

In summary, there are three numbers involved, the exponent m, the number of distinct primes k, and lastly the dummy index j as indicated below

$$\sum_{j=1}^{k} \sum_{(k|j)} \left(l_{i_1} + \ldots + l_{i_j} \right)^m$$

and the pattern of these different combinations of l's can be seen in Table I and we immediately see why they all sum to zero.

In Exercises 10, 11, and 12, $d(n)$ denotes the number of positive divisors of n.

Problem 2.10. Prove that $\prod_{t|n} t = n^{d(n)/2}$.

Again, let's decompose n into its primal constituents $n = \prod_i^N p_i^{\alpha_i}$ then $d(n)$ is given by

$$d(n) = d \left(\prod_i^N p_i^{\alpha_i} \right) = \prod_i^N (\alpha_i + 1)$$

To see why this is we just need to recall that the number of combinations an N-digit (base-10) number has is

$$\# \text{ of combo} = \underbrace{10 \times 10 \times 10 \times \ldots \times 10}_{N \text{ of them}}$$

because each digit can take 10 possible different values. For our case, each prime factor plays the role of a digit, however, each has different possible values, which is $(\alpha_i + 1)$

TABLE I: Coefficients of various combo of l's for different j's with a given k and m, note that the sum of the exponents of l's is always m, $\sum_i a_i = m$

	l^m (one l)	$l_{b_1}^{a_1} l_{b_2}^{a_2}$ (2 l's)	$l_{b_1}^{a_1} l_{b_2}^{a_2} l_{b_3}^{a_3}$ (3 l's)	$\cdots$	$l_{b_1}^{a_1} l_{b_2}^{a_2} \ldots l_{b_m}^{a_m}$ (m l's)
$(k\|1)$	$\binom{k-1}{1-1}$				
$(k\|2)$	$\binom{k-1}{2-1}$	$\binom{k-2}{2-2}$			
$(k\|3)$	$\binom{k-1}{3-1}$	$\binom{k-2}{3-2}$	$\binom{k-3}{3-3}$		
$\vdots$	$\vdots$	$\vdots$	$\vdots$		
$(k\|m)$	$\binom{k-1}{m-1}$	$\binom{k-2}{m-2}$	$\binom{k-3}{m-3}$		$\binom{k-m}{m-m}$
$(k\|m+1)$	$\binom{k-1}{(m+1)-1}$	$\binom{k-2}{(m+1)-2}$	$\binom{k-3}{(m+1)-3}$		$\binom{k-m}{(m+1)-m}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
$(k\|k)$	$\binom{k-1}{k-1}$	$\binom{k-2}{k-2}$	$\binom{k-3}{k-3}$		$\binom{k-m}{m-m}$

because we can have $p_i^0, p_i^1, p_i^2, \ldots, p_N^{\alpha_N}$ so the total number of combinations for $\prod_i^N p_i^{\alpha_i}$ is

$$\text{\# of combo} = \underbrace{(\alpha_1 + 1)(\alpha_2 + 1)(\alpha_3 + 1)\ldots(\alpha_N + 1)}_{N \text{ prime factors}}$$

Next, we can decompose $\prod_{t|n} t$ in terms of its primal constituents as well, say we focus on p_1 of $\prod_i^N p_i^{\alpha_i}$, the divisors of $p_1^{\alpha_1}$ are $p_1^0, p_1^1, \ldots, p_1^{\alpha_1}$, so if we multiply all of them we have $p_1^{1+2+3+\ldots+\alpha_1} = p_1^{\frac{\alpha_1(\alpha_1+1)}{2}} = (p_1^{\alpha_1})^{\frac{\alpha_1+1}{2}}$.

But here $p_1^{\alpha_1}$ is not alone, each divisor of $p_1^{\alpha_1}$, i.e. p_1^j, $0 \le j \le \alpha_1$, occurs $(\alpha_2 + 1)(\alpha_3 + 1)\ldots(\alpha_N + 1)$ times, so the final exponent for p_1 in $\prod_{t|n} t$ is

$$(p_1^{\alpha_1})^{\frac{(\alpha_1+1)}{2}(\alpha_2+1)(\alpha_3+1)\ldots(\alpha_N+1)} = (p_1^{\alpha_1})^{d(n)/2}$$

the same case goes for any other p_i, thus $\prod_{t|n} t = n^{d(n)/2}$. As a concrete example, take $n = p_1^2 p_2^3$, the divisors of n are

$$
\begin{array}{cccc}
p_1^0 \; p_2^0 & p_1^0 \; p_2^1 & p_1^0 \; p_2^2 & p_1^0 \; p_2^3 \\
p_1^1 \; p_2^0 & p_1^1 \; p_2^1 & p_1^1 \; p_2^2 & p_1^1 \; p_2^3 \\
p_1^2 \; p_2^0 & p_1^2 \; p_2^1 & p_1^2 \; p_2^2 & p_1^2 \; p_2^3
\end{array}
$$

so you can see that $(p_1^0 \; p_1^1 \; p_1^2)$ occurs $4 = (\alpha_2 + 1)$ times $\rightarrow (p_1^0 \; p_1^1 \; p_1^2)^{\alpha_2+1}$.

Problem 2.11. Prove that $d(n)$ is odd if, and only if, n is square.

As shown above for $n = \prod_i^N p_i^{\alpha_i}$, $d(n) = \prod_i^N (\alpha_i + 1)$, so to get $d(n)$ to be odd we need *all* of α_i to be even so that $(\alpha_i + 1)$ is odd, therefore n must be even

Problem 2.12. Prove that $\sum_{t|n} d(t)^3 = \left(\sum_{t|n} d(t)\right)^2$.

The above relationship is evidently not true in general, we therefore need to utilize the properties of $d(t)$ to derive it. One thing to note is that $g(n) = \sum_{t|n} d(t)^3$ is multiplicative as $d(t)$ is. Therefore we just need to consider $g(p^\alpha) = \sum_{t|p^\alpha} d(t)^3$.

My strategy would be to utilize induction. Assume that $\sum_{t|p^\alpha} d(t)^3 = \left(\sum_{t|p^\alpha} d(t)\right)^2$ is true up to some p^α, we now want to know what happens with $p^{\alpha+1}$

$$\sum_{t|p^{\alpha+1}} d(t)^3 = d(p^{\alpha+1})^3 + \sum_{t|p^\alpha} d(t)^3$$

and $d(p^{\alpha+1}) = \alpha + 2$ thus

$$d(p^{\alpha+1})^3 + \sum_{t|p^\alpha} d(t)^3 = (\alpha + 2)^3 + \left(\sum_{t|p^\alpha} d(t)\right)^2$$

$$= (\alpha + 2)^2(\alpha + 2) + \left(\sum_{t|p^\alpha} d(t)\right)^2$$

$$= (\alpha + 2)^2 + (\alpha + 2)^2(\alpha + 1) + \left(\sum_{t|p^\alpha} d(t)\right)^2$$

$$= d(p^{\alpha+1})^2 + (\alpha + 2) \cdot 2\frac{(\alpha + 2)(\alpha + 1)}{2} + \left(\sum_{t|p^\alpha} d(t)\right)^2$$

$$= d(p^{\alpha+1})^2 + 2d(p^{\alpha+1})\left(\sum_{t|p^\alpha} d(t)\right) + \left(\sum_{t|p^\alpha} d(t)\right)^2$$

$$= \left(d(p^{\alpha+1}) + \sum_{t|p^\alpha} d(t)\right)^2$$

$$\sum_{t|p^{\alpha+1}} d(t)^3 = \left(\sum_{t|p^{\alpha+1}} d(t)\right)^2$$

Going to line 5 we have used the fact that $\sum_{t|p^\alpha} d(t) = \sum_{i=1}^{\alpha+1} i = \frac{(\alpha+1)(\alpha+2)}{2}$ since $d(p^j) = j + 1$. We can of course dispel induction for a bruter force approach by ex-

panding $\sum_{t|p^{\alpha}+1} d(t)^3 = \sum_{i=1}^{\alpha+1} i^3$ but this requires us to know the formula for a sum of consecutive cubes ¯\\(°_o)/¯

Chapter 3

Section 3.3, Page 54. "Euler's summation formula, Theorem 3.1", if you look carefully enough at the definition the lower limit of the sum, $\sum_{y<n\leq x}$, is just a less than sign and **not** and less than **equal** sign, this is crucial as we go to the proof of Theorem 3.3

PROOF of Theorem 3.1, Page 55, there's some "fudging" going on here and not just once :) in the first line

$$\int_{n-1}^{n} [t]f'(t)dt \to \int_{n-1}^{n} (n-1)f'(t)dt$$

so we assume that $[t] = n-1$ for the whole interval $[n-1,n]$, well this is true for if the interval excludes n, *i.e.* $[n-1,n)$. So here's where the "fudging" comes in, in the definition of Riemann integral we can always remove a point since the area under a curve of just one point is zero $\to dt = 0$ so in this case we remove the end point n.

The second "fudging" happens in Eq (6) when we go from $-\int_{m}^{k} \to -\int_{y}^{x}$, the area under the curve is definitely different for those two integrals unless $m = [y] = y$ and $k = [x] = x$, let's see this with a concrete example say, $f(t) = t \to f'(t) = 1$ with $y = 1.5$ and $x = 3.1$, the integral $\int_{y}^{x} [t]f'(t)dt$ is given by

$$\int_{1.5}^{3.1} [t]f'(t)dt = \int_{1.5}^{2} dt + \int_{2}^{3} 2dt + \int_{3}^{3.1} 3dt$$
$$= (2-1.5) + 2(3-2) + 3(3.1-3)$$
$$= 0.5 + 2 + 0.3$$
$$= 2.8$$

while the integral $\int_{[y]}^{[x]} [t]f'(t)dt$ is given by

$$\int_{[1.5]}^{[3.1]} [t]f'(t)dt = \int_{1}^{2} dt + \int_{2}^{3} 2dt$$
$$= (2-1) + 2(3-2)$$
$$= 3$$

30

so obviously $-\int_m^k \neq -\int_y^x$ but this is ok because

$$-\int_y^x [t]f'(t)dt = -\int_m^k [t]f'(t)dt + \int_m^y [t]f'(t)dt - \int_k^x [t]f'(t)dt$$

$$\rightarrow -\int_m^k [t]f'(t)dt + kf(k) - mf(m) = -\int_y^x [t]f'(t)dt + \left(kf(k) + \int_k^x [t]f'(t)dt\right)$$

$$+ \left(-\int_m^y [t]f'(t)dt - mf(m)\right)$$

$$= -\int_y^x [t]f'(t)dt + \left(\cancel{kf(k)} + [x]f(x) - \cancel{[k]f(k)}\right)$$

$$+ \left(-[y]f(y) + \cancel{[m]f(m)} - \cancel{mf(m)}\right)$$

$$= -\int_y^x [t]f'(t)dt + kf(x) - mf(y)$$

recall that $k = [x]$ and $m = [y]$. Note that in the last line the parameters for f have changed from $f(k), f(m) \rightarrow f(x), f(y)$ and so it is a legit move.

Page 55, "Integration by parts gives us ... and when this is combined with (6) we obtain (5)", what we want to do here is to move everything to the LHS

$$\int_y^x f(t)dt = xf(x) - yf(y) - \int_y^x tf'(t)dt$$

$$\rightarrow \int_y^x f(t)dt - xf(x) + yf(y) + \int_y^x tf'(t)dt = 0$$

and we add this zero to (6) to get (5) sans the "fudgings" discussed above :)

PROOF of Theorem 3.2, Page 55 The peculiar thing here is the constant 1 which we get from $-f(y)([y] - y)$. The problem here is that the lower limit $y = 1$ and so $[y] - y = 1 - 1 = 0$, however, as explained above, the lower limit must not be an integer as $y < n$ and n starts with $n = 1$, the less than sign is crucial here.

This means that $0 < y < 1$ and in this case y has to be $y = 1^-$ such that we can take the limit $y \rightarrow 1$, so even though the lower limit of the integrals is $y = 1$ we cannot just simply choose $y = 1$, this manifests more strongly in the proof of Theorem 3.2 part (b) where in this case $f(y) = x^{-s}$

$$-f(y)([y] - y) = -\frac{1}{y^s}(0 - y)$$

$$= \frac{1}{y^{s-1}}$$

31

it won't equal 1 unless we take the limit $y \to 1$, we can however, choose any value of y in the range $0 < y < 1$ say $y = 1/w$ where $0 < w < \infty$

$$\sum_{n \leq x} \frac{1}{n^s} = \int_{1/w}^{x} \frac{dt}{t^s} - s \int_{1/w}^{x} \frac{t - [t]}{t^{s+1}} + \frac{[x] - x}{x^s} - \frac{0 - (1/w)}{(1/w)^s}$$

$$= \frac{x^{1-s}}{1-s} - \frac{(1/w)^{1-s}}{1-s} - s \int_{1/w}^{1} \frac{t - [t]}{t^{s+1}} - s \int_{1}^{x} \frac{t - [t]}{t^{s+1}} - \frac{x - [x]}{x^s} + (1/w)^{1-s}$$

$$= \frac{x^{1-s}}{1-s} - \frac{(1/w)^{1-s}}{1-s} - s \int_{1/w}^{1} \frac{t - [t]}{t^{s+1}} - s \int_{1}^{x} \frac{t - [t]}{t^{s+1}} - \frac{x - [x]}{x^s} + \frac{(1-s)(1/w)^{1-s}}{1-s}$$

we now focus on the first integral on the RHS

$$-s \int_{1/w}^{1} \frac{t - [t]}{t^{s+1}} = -s \int_{1/w}^{1} \frac{t - [1/w]}{t^{s+1}}$$

$$= -s \int_{1/w}^{1} \frac{t}{t^{s+1}}$$

$$= -s \int_{1/w}^{1} \frac{1}{t^s}$$

$$= -\frac{s}{1-s} + \frac{s(1/w)^{1-s}}{1-s}$$

Combining everything we get

$$\sum_{n \leq x} \frac{1}{n^s} = \frac{x^{1-s}}{1-s} - \cancel{\frac{(1/w)^{1-s}}{1-s}} - \frac{s}{1-s} + \cancel{\frac{s(1/w)^{1-s}}{1-s}} - s \int_{1}^{x} \frac{t - [t]}{t^{s+1}} - \frac{x - [x]}{x^s} + \frac{(1 - \cancel{s})(1/w)^{1-s}}{1-s}$$

$$= \frac{x^{1-s}}{1-s} + \frac{-1 + (1-s)}{1-s} - s \int_{1}^{x} \frac{t - [t]}{t^{s+1}} - \frac{x - [x]}{x^s}$$

$$= \frac{x^{1-s}}{1-s} - \frac{1}{1-s} + 1 - s \int_{1}^{x} \frac{t - [t]}{t^{s+1}} - \frac{x - [x]}{x^s}$$

and we get the same result.

The lesson here is that we need to be very careful in choosing and processing these limits otherwise we'll get the wrong result.

Page 56, first equation, the key here is that $t - [t] < 1$ and so the inequality holds

Page 56, the value of C (and $C(s)$), in the roughly middle section of the page we have

$$\lim_{x \to \infty} \left(\sum_{n \leq x} \frac{1}{n} - \log x \right) = 1 - \int_{1}^{\infty} \frac{t - [t]}{t^2} dt,$$

and the RHS is C and therefore C equals the Euler's constant since the LHS is Euler's constant but it seems like this is only true when $x \to \infty$, which is to say that C is

32

independent of x, but recall that from earlier in the page

$$\sum_{n\leq x}\frac{1}{n} = \log x + C + O\left(\frac{1}{x}\right)$$

$$\to C = \sum_{n\leq x}\frac{1}{n} - \log x - O\left(\frac{1}{x}\right)$$

It therefore seems like C depends on x after all, how can this be? The answer is that

$$\sum_{n\leq x}\frac{1}{n} - \log x - O\left(\frac{1}{x}\right) = \lim_{x\to\infty}\left(\sum_{n\leq x}\frac{1}{n} - \log x\right)$$

Thus the LHS does **not** depend on x after all even though it looks like it, this argument also applies to $C(s)$ lower down in the page

Page 58, Figure 3.1, note that this figure is a bit deceptive, to calculate $\sum_{qd\leq 10}$ we need to draw one hyperbola for every $qd = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, in the figure there are only 3 hyperbolas drawn. Also, the way the sum is calculated is by counting **all** lattice points underneath the largest hyperbola, true that we have to count only points that lie on the hyperbolas but if we draw **all** hyperbolas with $qd \leq x$ we will cover all lattice points under the largest hyperbolic curve.

Page 60, "It can be shown that $\zeta(2) = \pi^2/6$", an elementary proof of this fact can be obtained from the Fourier series (*not* Fourier transform) of x^2

$$x^2 = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)$$

$b_n = 0$ for all n because x^2 is an even function while

$$a_0 = \int_{-\pi}^{+\pi} x^2 \frac{dx}{\pi}$$
$$= \frac{1}{3\pi}\left(\pi^3 - (-\pi)^3\right)$$
$$= \frac{2\pi^2}{3}$$

and

$$a_n = \int_{-\pi}^{+\pi} x^2 \cos(nx)\frac{dx}{\pi}$$
$$= \frac{2x}{n\pi}\sin(nx)\Big|_{-\pi}^{+\pi} - 2\int_{-\pi}^{+\pi} x\sin(nx)\frac{dx}{n\pi}$$

33

the boundary terms are zero since $\sin(\pm n\pi) = 0$, we now do another integration by parts on the remaining integral

$$-2\int_{-\pi}^{+\pi} x\sin(nx)\frac{dx}{n\pi} = \frac{2x}{n^2\pi}\cos(nx)\bigg|_{-\pi}^{+\pi} - 2\int_{-\pi}^{+\pi}\cos(nx)\frac{dx}{n^2\pi}$$

$$= \frac{2x}{n^2\pi}\cos(nx)\bigg|_{-\pi}^{+\pi} + \frac{-2}{n^3\pi}\sin(nx)\bigg|_{-\pi}^{+\pi}$$

$$a_n = (-1)^n\frac{4}{n^2}$$

where $\cos(\pm n\pi) = \cos(n\pi) = (-1)^n$ and of course $\sin(\pm n\pi) = 0$, thus

$$x^2 = \frac{\pi^2}{3} + \sum_{n=1}^{\infty}(-1)^n\frac{4}{n^2}\cos(nx)$$

setting $x = \pi$ we get

$$\pi^2 = \frac{\pi^2}{3} + \sum_{n=1}^{\infty}(-1)^n\frac{4}{n^2}\cos(n\pi)$$

$$\frac{2\pi^2}{3} = \sum_{n=1}^{\infty}(-1)^{2n}\frac{4}{n^2}$$

$$\rightarrow \frac{\pi^2}{6} = \sum_{n=1}^{\infty}\frac{1}{n^2}$$

Page 61, PROOF of Theorem 3.6, this is rather odd the sum is usually split into the following (see Theorem 3.4 and 3.5)

$$\sum_{n\leq x}\sigma_\alpha(n) \rightarrow \sum_{n\leq x}\sum_{q|n}q^\alpha = \sum_{d\leq x}\left(\sum_{q\leq x/d}q^\alpha\right)$$

but for Theorem 3.6 it's split another way (with a dramatically different final result)

$$\sum_{n\leq x}\sigma_{-\beta}(n) \rightarrow \sum_{n\leq x}\sum_{d|n}\frac{1}{d^\beta} = \sum_{d\leq x}\left(\frac{1}{d^\beta}\sum_{q\leq x/d}1\right)$$

although the two are equivalent counting wise but once you use Euler's summation formula you get a completely different result. To see that the two are equivalent let's just calculate

34

something simple, say $\sum_{n \leq 5} \sigma_\alpha(n)$ where $\alpha = -\beta$, the original definition gives us

$$\sum_{n \leq 5} \sigma_{-\beta}(n) \to \sum_{n \leq x} \sum_{d|n} \frac{1}{d^\beta}$$

$$= \sum_{d|1} \frac{1}{d^\beta} + \sum_{d|2} \frac{1}{d^\beta} + \sum_{d|3} \frac{1}{d^\beta} + \sum_{d|4} \frac{1}{d^\beta} + \sum_{d|5} \frac{1}{d^\beta}$$

$$= \left(\frac{1}{1^\beta} \right) + \left(\frac{1}{1^\beta} + \frac{1}{2^\beta} \right) + \left(\frac{1}{1^\beta} + \frac{1}{3^\beta} \right) + \left(\frac{1}{1^\beta} + \frac{1}{2^\beta} + \frac{1}{4^\beta} \right) + \left(\frac{1}{1^\beta} + \frac{1}{5^\beta} \right)$$

$$= \left(\frac{1}{1^\beta} \times 5 \right) + \left(\frac{1}{2^\beta} \times 2 \right) + \left(\frac{1}{3^\beta} \times 1 \right) + \left(\frac{1}{4^\beta} \times 1 \right) + \left(\frac{1}{5^\beta} \times 1 \right)$$

while Theorem 3.6 decomposition gives us

$$\sum_{d \leq 5} \left(\frac{1}{d^\beta} \sum_{q \leq 5/d} 1 \right) = \left(\frac{1}{1^\beta} \sum_{q \leq 5/1} 1 \right) + \left(\frac{1}{2^\beta} \sum_{q \leq 5/2} 1 \right) + \left(\frac{1}{3^\beta} \sum_{q \leq 5/3} 1 \right) + \left(\frac{1}{4^\beta} \sum_{q \leq 5/4} 1 \right) + \left(\frac{1}{5^\beta} \sum_{q \leq 5/5} 1 \right)$$

$$= \left(\frac{1}{1^\beta} \times 5 \right) + \left(\frac{1}{2^\beta} \times 2 \right) + \left(\frac{1}{3^\beta} \times 1 \right) + \left(\frac{1}{4^\beta} \times 1 \right) + \left(\frac{1}{5^\beta} \times 1 \right)$$

and lastly Theorem 3.4 and 3.5 decomposition generates

$$\sum_{d \leq 5} \left(\sum_{q \leq 5/d} q^\alpha \right) = \left(\sum_{q \leq 5/1} q^\alpha \right) + \left(\sum_{q \leq 5/2} q^\alpha \right) + \left(\sum_{q \leq 5/3} q^\alpha \right) + \left(\sum_{q \leq 5/4} q^\alpha \right) + \left(\sum_{q \leq 5/5} q^\alpha \right)$$

$$= (5 \times 1^\alpha) + (2 \times 2^\alpha) + (1 \times 3^\alpha) + (1 \times 4^\alpha) + (1 \times 5^\alpha)$$

so you see that all three are equivalent but now let's apply Theorem 3.4 and 3.5's decomposition tot he proof of Theorem 3.6

$$\sum_{n \leq x} \sigma_{-\beta}(n) = \sum_{n \leq x} \sum_{d|n} \frac{1}{d^\beta} = \sum_{d \leq x} \sum_{q \leq x/d} \frac{1}{q^\beta}$$

if $\beta \neq 1$ we have

$$\sum_{d \leq x} \sum_{q \leq x/d} \frac{1}{q^\beta} = \sum_{d \leq x} \frac{(x/d)^{1-\beta}}{1 - \beta} + \zeta(\beta) + O\left((x/d)^{-\beta} \right)$$

$$= x\zeta(\beta) + \frac{x^{1-\beta}}{1 - \beta} \sum_{d \leq x} d^{\beta-1} + O\left(\frac{1}{x^\beta} \sum_{d \leq x} d^\beta \right)$$

$$= x\zeta(\beta) + \frac{x^{1-\beta}}{1 - \beta} \left(\frac{x^\beta}{\beta} + O(x^{\beta-1}) \right) + O\left(\frac{1}{x^\beta} \left(\frac{x^{\beta+1}}{\beta + 1} + O(x^\beta) \right) \right)$$

$$= x\zeta(\beta) + O(x) + O(1) + O(x) + O(1)$$

$$= x\zeta(\beta) + O(x)$$

For $\beta = 1$ first we need to calculate $\sum_{n \leq x} \log n$

$$\sum_{n \leq x} \log n = \int_1^x \log t \; dt + \int_1^x \frac{(t - [t])}{t} dt + \log x([x] - x) - \lim_{y \to 1} \log y([y] - y)$$

$$= \cancel{x \log x} - x + 1 + \int_1^x \frac{(t - [t])}{t} dt - \cancel{x \log x} + [x] \log x$$

$$= [x] \log x - x + 1 + O(\log x)$$

$$= O(x \log x)$$

the above is actually given in the book on Page 68 and on that page the $x \log x$ wasn't canceled using $-\cancel{x \log x} + [x] \log x$ instead, the latter terms were grouped as $\log x([x] - x) \to O(\log x)$ since $|[x] - x| \leq 1$, so this summation integration thingy is kinda somewhat random

While if $\beta = 1$ we have

$$\sum_{d \leq x} \sum_{q \leq x/d} \frac{1}{q^\beta} = \sum_{d \leq x} \log \left(\frac{x}{d} \right) + C + O \left(\frac{d}{x} \right)$$

$$= xC + x \log x - \sum_{d \leq x} \log d + O \left(\frac{1}{x} \sum_{d \leq x} d \right)$$

$$= xC + x \log x - O(x \log x) + O \left(\frac{1}{x} \left(\frac{x^2}{2} + O(x) \right) \right)$$

$$= xC + x \log x - O(x \log x) + O(x) + O(1)$$

$$= xC + O(x \log x)$$

so the result is very different from the book's, now if we apply Theorem 3.6's decomposition to Theorem 3.5 we will get

$$\sum_{n \leq x} \sigma_\alpha(n) = \sum_{n \leq x} \sum_{d | n} d^\alpha = \sum_{d \leq x} d^\alpha \sum_{q \leq x/d} 1$$

$$= \sum_{d \leq x} d^\alpha \left(\frac{x}{d} + O(1) \right) = x \sum_{d \leq x} d^{\alpha - 1} + O \left(\sum_{d \leq x} d^\alpha \right)$$

$$= \left(\frac{x^{\alpha+1}}{\alpha} + O(x^\alpha) \right) + O \left(\frac{x^{\alpha+2}}{\alpha + 1} + O \left(x^{\alpha+2} \right) \right)$$

$$= \frac{x^{\alpha+1}}{\alpha} + O(x^\alpha) + O \left(x^{\alpha+2} \right)$$

so here we see why the different decomposition, the thing is that we want to maximize the leading terms and minimize the big Oh terms, in our examples above for $\sigma_{-\beta}$ our big

36

Oh is bigger than the ones in the book and for σ_α not only our big Oh is bigger but our leading term is also smaller, so that's the reason why we have different decompositions. But of course the trick is knowing which decomposition to use beforehand which is not that obvious for someone new.

Page 62, PROOF of Theorem 3.8, the thing to note here is that the visibility between two points is translationally invariant (as long as we translate using integer increments), you can translate the two points to any two other points and the visibility will be the same. And if we think of them as vectors then $(a - m, b - n)$ is the the vector pointing to (a, b) from (m, n).

It is also symmetric under reflection w.r.t x or y or x followed by y or vice versa but not under rotation, maybe under boost? ¯\\(°_o)/¯although I'm not sure if these symmetries will bring about any new information about the Euler totient function or any other arithmetical functions ...

Page 65, Theorem 3.10, let's recap some definitions, the asterisk product is defnied as

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

and the circle product is defined as

$$(\alpha \circ F)(x) = \sum_{n \leq x} \alpha(n)F\left(\frac{x}{n}\right)$$

where α is an aritmetical function while F is a real function with $F(x) = 0$ for $0 < x < 1$. Therefore

$$
\begin{aligned}
H(x) &= \sum_{n \leq x} h(n) \\
&= \sum_{n \leq x} (f * g)(n)U\left(\frac{x}{n}\right) \\
&= (f * g) \circ U
\end{aligned}
$$

Page 67, Theorem 3.14, this is actually a pretty neat theorem because it relates the factorial of x to prime numbers less than x, but I believe it can be straightforwardly proven using induction. And as to why this works, visually we can see it below with an

example for $\alpha(p)$ with $x = 10$ and $p = 2$, each column shows the number of 2's in each number (we only consider even numbers here obviously)

$$2 \qquad \longrightarrow x/p^3 = 10/2^3 = 1 \text{ factor of } 2$$
$$2 \quad 2 \qquad \longrightarrow x/p^2 = 10/2^2 = 2 \text{ factor of } 2$$
$$2\ 2\ 2\ 2\ 2 \quad \longrightarrow x/p^1 = 10/2^1 = 5 \text{ factor of } 2$$
$$\uparrow\ \uparrow\ \uparrow\ \uparrow\ \uparrow$$
$$2\ 4\ 6\ 8\ 10 \longrightarrow \text{a total of } 5 + 2 + 1 = 8 \text{ factors of } 2$$

and I believe there's a small typo in the *Note* under Eq (20), it should've read "The sum for $\alpha(p)$ is finite since $[x/p^m] = 0$ for $p^m > x$." and **not** "for $p > x$".

Page 68, last line

$$x \sum_{n=2}^{\infty} \frac{\log n}{n(n-1)} = O(x).$$

this is because $\sum_{n=2}^{\infty} \frac{\log n}{n(n-1)}$ is a convergent series, how do we see this, first we use integral test to show that $\sum_{n=2}^{\infty} \frac{\log n}{n^2}$ is convergent and then use limit comparison test to show that $\sum_{n=2}^{\infty} \frac{\log x}{n(n-1)}$ is also convergent. First

$$\int_2^{\infty} \frac{\log x}{x^2} = -\frac{\log x}{x} - \frac{1}{x}\Big|_2^{\infty}$$
$$= 0 - \left(-\frac{\log 2}{2} - \frac{1}{2}\right)$$
$$= \frac{1 + \log 2}{2}$$

because $\lim_{x\to\infty} \frac{\log x}{x} = 0$ so $\int_2^{\infty} \frac{\log x}{x^2}$ is convergent. Now, the integral test says that if $\int_k^{\infty} f(x)dx$ is convergent then $\sum_{n=k}^{\infty} a(n)$ is also convergent, and if the integral is divergent the sum is too, and of course here $f(n) = a(n)$, therefore $\sum_{n=2}^{\infty} \frac{\log n}{n^2}$ is also convergent.

The limit comparison test says that if we have two series $\sum a(n)$ and $\sum b(n)$ with $a(n) \geq 0$, $b(n) > 0$, if the following limit exists

$$c = \lim_{n\to\infty} \frac{a(n)}{b(n)}$$

and c is **positive** and finite then either both series diverge or both converge. So we now need to take the limit

$$\lim_{n\to\infty} \frac{\frac{\log n}{n^2}}{\frac{\log n}{n(n-1)}} = \lim_{n\to\infty} \frac{n^2 - n}{n^2}$$
$$= 1$$

38

so the limit is finite and positive and since $\sum_{n=2}^{\infty} \frac{\log n}{n^2}$ is convergent therefore $\sum_{n=2}^{\infty} \frac{\log n}{n(n-1)}$ is must also be convergent.

I actually tried doing integral test straight on $\sum_{n=2}^{\infty} \frac{\log n}{n(n-1)}$ but the integral turned out to be complicated and so I used the two step method above :)

Problem 3.20. If n is a positive integer prove that $[\sqrt{n} + \sqrt{n+1}] = [\sqrt{4n+2}]$

This problem turns out to be trickier than I initially thought, the first order of business here is to see how $4n + 2$ behaves, the $4n$ scream modulo 4 to me, so let's see.

One thing we know is that only $4n + 1$ and $4n$ can be perfect squares, for one thing the square of any odd number is $(2x + 1)^2 = 4(x^2 + x) + 1$ and the square of any even number is $4x^2$ while we know that 2 (mod 4) and 3 (mod 4) are not quadratic residues modulo 4.

To see why $4n+2$ can't be a perfect square, without going into congruence and modulo stuff, let's consider the opposite

$$2(2n + 1) = m^2$$

which means that $2|m^2$ and that in turn means $m = 2w$ and

$$2(2n + 1) = 2{\kern-6pt/}2w^2$$
$$1 = 2(w^2 - n)$$
$$\rightarrow 2 \mid 1$$

which is a contradiction. For $4n + 3$, it is an odd number and thus as shown above, if it is to be a perfect square

$$4n + 3 = 4(x^2 + x) + 1$$
$$4n + 2 = 4(x^2 + x)$$
$$2n + 1 = 2(x^2 + x)$$
$$1 = 2((x^2 + x) - n)$$
$$\rightarrow 2 \mid 1$$

which again is a contradiction and thus $4n+2$ and $4n+3$ cannot be perfect squares ever.

Next, we will list the natural numbers (including zero) and their corresponding square roots

$\sqrt{m}$	m	$m = 4n + j$
0	**0**	$\mathbf{4 \cdot 0 + 0}$
1	**1**	$\mathbf{4 \cdot 0 + 1}$
1.414	2	$4 \cdot 0 + 2$
1.732	3	$4 \cdot 0 + 3$
2	**4**	$\mathbf{4 \cdot 1 + 0}$
2.236	5	$4 \cdot 1 + 1$
2.449	6	$4 \cdot 1 + 2$
2.646	7	$4 \cdot 1 + 3$
2.828	8	$4 \cdot 2 + 0$
3	**9**	$\mathbf{4 \cdot 2 + 1}$
3.162	10	$4 \cdot 2 + 2$
3.317	11	$4 \cdot 2 + 3$
3.464	12	$4 \cdot 3 + 0$
3.606	13	$4 \cdot 3 + 1$
3.742	14	$4 \cdot 3 + 2$
3.873	15	$4 \cdot 3 + 3$
4	**16**	$\mathbf{4 \cdot 4 + 0}$

so we see that $\sqrt{m} = \sqrt{4n + j}$ hits integer points only at $m = 4n + 0$ or $m = 4n + 1$ and in between any two consecutive integer $\sqrt{m}$'s (*e.g.* between $\sqrt{m} = 3$ and $\sqrt{m} = 4$) the values of $\sqrt{m}$ in between them all share the same floor because the two consecutive integer $\sqrt{m}$'s only differ by one, this only means one thing

$$[\sqrt{4n + 1}] = [\sqrt{4n + 2}] = [\sqrt{4n + 3}]$$

Next we tackle $\sqrt{n} + \sqrt{n + 1}$, if we square it

$$\left(\sqrt{n} + \sqrt{n + 1}\right)^2 = 2n + 1 + 2\sqrt{n^2 + n}$$

Now the thing to note is that $n^2 < (n^2 + n) < (n + 1/2)^2$, at least for $n > 0$, and so

$$
\begin{aligned}
2n + 1 + 2\sqrt{n^2} \;&<\; 2n + 1 + 2\sqrt{n^2 + n} \;<\; 2n + 1 + 2\sqrt{(n + 1/2)^2} \\
4n + 1 \;&<\; \left(\sqrt{n} + \sqrt{n + 1}\right)^2 \;<\; 4n + 2 \\
\sqrt{4n + 1} \;&<\; \sqrt{n} + \sqrt{n + 1} \;<\; \sqrt{4n + 2}
\end{aligned}
$$

and if we take the floor of all three items above

$$
[\sqrt{4n + 1}] \;\leq\; [\sqrt{n} + \sqrt{n + 1}] \;\leq\; [\sqrt{4n + 2}]
$$

note that the inequality signs have all changed into less than **equal** signs. At this point we need to be extra careful because even though $\sqrt{4n + 1}$ and $\sqrt{4n + 2}$ differ by less than one it can be that the three numbers $\sqrt{4n + 1}, \sqrt{n} + \sqrt{n + 1}, \sqrt{4n + 2}$ straddle through an integer, for example say $\sqrt{4n + 1} = 15.91\ldots$ and $\sqrt{n} + \sqrt{n + 1} = 15.92\ldots$ while $\sqrt{4n + 2} = 16.01\ldots$, in this case $[\sqrt{4n + 1}] = [\sqrt{n} + \sqrt{n + 1}] \neq [\sqrt{4n + 2}]$ which works against us, so our argument has to be solid to establish the result we want.

Lucky for us, from the above discussion we know that $[\sqrt{4n + 1}] = [\sqrt{4n + 2}] = [\sqrt{4n + 3}]$ therefore since $[\sqrt{4n + 1}] \leq [\sqrt{n} + \sqrt{n + 1}] \leq [\sqrt{4n + 2}]$ the only possibility is $[\sqrt{4n + 1}] = [\sqrt{n} + \sqrt{n + 1}] = [\sqrt{4n + 2}]$ and also

$$
[\sqrt{n} + \sqrt{n + 1}] = [\sqrt{4n + 1}] = [\sqrt{4n + 2}] = [\sqrt{4n + 3}]
$$

which is what we are after and more, for one thing we get equalities with $[\sqrt{4n + 1}], [\sqrt{4n + 3}]$ and it's obvious that this equality holds for $n = 0$, therefore the question could've have been, for any integer $n \geq 0$, prove that the following is true $[\sqrt{n} + \sqrt{n + 1}] = [\sqrt{4n + 1}] = [\sqrt{4n + 2}] = [\sqrt{4n + 3}]$

Various Failed Attempts

First thing I tried was the famous physicist formula for a square root $\sqrt{1 + x} \approx 1 + \frac{1}{2}x$ and so

$$
\begin{aligned}
\left[\sqrt{n} + \sqrt{n + 1}\right] &= \left[\sqrt{n} + \sqrt{n}\sqrt{1 + \frac{1}{n}}\right] = \left[\sqrt{n} + \sqrt{n}\left(1 + \frac{1}{2n}\right)\right] \\
&= \left[2\sqrt{n} + \frac{1}{2\sqrt{n}}\right]
\end{aligned}
$$

and also

$$\left[\sqrt{4n+2}\right] = \left[2\sqrt{n}\sqrt{1 + \frac{1}{2n}}\right] = \left[2\sqrt{n}\left(1 + \frac{1}{4n}\right)\right]$$
$$= \left[2\sqrt{n} + \frac{1}{2\sqrt{n}}\right]$$

so to a physicist they're the same :)

But joking aside I did try the Taylor expansion trick above, first I tried to estimate how big (or small) $\frac{1}{2\sqrt{n}}$ is, but this turns out not to be fruitful because I don't know how close $2\sqrt{n}$ is to an integer, because depending on how close it is, the correction term might or might not push it up to the next integer, and I don't see a clear way to tell how close $2\sqrt{n}$ is to an integer.

Next, I did the Taylor expansion above to relate $\sqrt{n} + \sqrt{n+1}$ to $\sqrt{4n+2}$, but first, a couple of things to note, for small x

$$1 + \frac{1}{2}x < \sqrt{1+x}$$
$$1 - \frac{1}{2}x > \sqrt{1-x}$$

Expanding $\sqrt{n} + \sqrt{n+1}$

$$\sqrt{n} + \sqrt{n+1} = \sqrt{\left(\sqrt{n} + \sqrt{n+1}\right)^2}$$
$$= \sqrt{2n + 1 + 2\sqrt{n(n+1)}}$$
$$= \sqrt{2n + 1 + 2\sqrt{\left(n + \frac{1}{2}\right)^2 - \frac{1}{4}}}$$
$$= \sqrt{2n + 1 + \sqrt{(2n+1)^2 - 1}}$$
$$= \sqrt{2n + 1 + \left((2n+1) - \frac{1}{2(2n+1)}\right)}$$
$$= \sqrt{4n + 2 - \frac{1}{2(2n+1)}}$$
$$\sqrt{n} + \sqrt{n+1} < \sqrt{4n+2} - \frac{1}{2(4n+2)^{3/2}}$$

42

a somewhat interesting result, next, let's denote $\Delta = \frac{1}{2(4n+2)^{3/2}}$, and I wanted to show that $\sqrt{n} + \sqrt{n+1} + \sqrt{\Delta}$ is bigger than $\sqrt{4n+2}$

$$\sqrt{n} + \sqrt{n+1} + \sqrt{\Delta} = \sqrt{4n+2-\Delta} + \sqrt{\Delta}$$
$$\rightarrow \left(\sqrt{4n+2-\Delta} + \sqrt{\Delta}\right)^2 = 4n+2-\Delta+\Delta + 2\sqrt{\Delta}(4n+2-\Delta)$$
$$\rightarrow \left(\sqrt{4n+2-\Delta} + \sqrt{\Delta}\right)^2 > \left(\sqrt{4n+2}\right)^2$$
$$\rightarrow \sqrt{4n+2-\Delta} + \sqrt{\Delta} > \sqrt{4n+2}$$
$$\rightarrow \sqrt{n} + \sqrt{n+1} + \sqrt{\Delta} > \sqrt{4n+2}$$

and since $0 < \Delta < 1$ it means that $0 < \sqrt{\Delta} < 1$ and so

Thus what we have is

$$\sqrt{n} + \sqrt{n+1} \quad < \quad \sqrt{4n+2} \quad < \quad \sqrt{n} + \sqrt{n+1} + \sqrt{\Delta}$$
$$[\sqrt{n} + \sqrt{n+1}] \quad \leq \quad [\sqrt{4n+2}] \quad \leq \quad [\sqrt{n} + \sqrt{n+1} + \sqrt{\Delta}]$$

and I ran into the same straddle thing again, $\sqrt{4n+2}$ might not be an integer and it can be that say, $\sqrt{n} + \sqrt{n+1} = 15.9\ldots$ while $\sqrt{4n+2} = 16.001\ldots$ and $\sqrt{n} + \sqrt{n+1} + \sqrt{\Delta} = 16.002\ldots$, in this case $[\sqrt{n} + \sqrt{n+1}] \neq [\sqrt{4n+2}] = [\sqrt{n} + \sqrt{n+1} + \sqrt{\Delta}]$.

I wrongfully tried to remedy the situation by claiming that

$$[\sqrt{n} + \sqrt{n+1}] \quad < \quad \sqrt{4n+2} \quad < \quad [\sqrt{n} + \sqrt{n+1} + \sqrt{\Delta}]$$
$$[\sqrt{n} + \sqrt{n+1}] \quad < \quad \sqrt{4n+2} \quad < \quad [\sqrt{n} + \sqrt{n+1}] + 1$$

we can substitute $[\sqrt{n} + \sqrt{n+1} + \sqrt{\Delta}]$ with $[\sqrt{n} + \sqrt{n+1}] + 1$ because $[\sqrt{n} + \sqrt{n+1}] + 1 \geq [\sqrt{n} + \sqrt{n+1} + \sqrt{\Delta}]$ and here since we are moving the upper bound we want the bigger one. The problem with the above argument is that for our numerical example $\sqrt{4n+2} = 16.001\ldots$ and $\sqrt{n} + \sqrt{n+1} + \sqrt{\Delta} = 16.002\ldots$, $\sqrt{4n+2} = [\sqrt{n} + \sqrt{n+1}] + 1$ and the inequality doesn't hold.

So the key here is to make sure that there's no such straddling and that means that $\sqrt{4n+1}$ and $\sqrt{4n+2}$ lie in between the same two integers $[a, b)$ and that's how I got the observation about $4n + j$ above.

For next attempts I was thinking that maybe I should process this whole thing in a different way maybe multiply and add them RHS and LHS or maybe I should get some

identity that involves $[\sqrt{n} + \sqrt{n+1}] - [\sqrt{4n+2}]$ and requires this difference to be zero, for example $f(x) = g(x) + [\sqrt{n} + \sqrt{n+1}] - [\sqrt{4n+2}] = 0$ and we already know that $g(x) = 0$ so the difference between those two square brackets must be zero, but I couldn't find such identity.

The next thing I tried was to maybe multiply or divide, I tried division together with Taylor expansion

$$\frac{\sqrt{4n+2}}{\sqrt{n+1}+\sqrt{n}} = \frac{\sqrt{4n+2}}{\sqrt{n+1}+\sqrt{n}} \times \frac{\sqrt{n+1}-\sqrt{n}}{\sqrt{n+1}-\sqrt{n}}$$

$$= \frac{\sqrt{4n+2}\left(\sqrt{n}\sqrt{1+\frac{1}{n}}-\sqrt{n}\right)}{n+1-n}$$

$$= \frac{\sqrt{4n+2}\left(\sqrt{n}\left(1+\frac{1}{2n}\right)-\sqrt{n}\right)}{1}$$

$$= \sqrt{4n+2}\left(\frac{1}{2\sqrt{n}}\right)$$

$$= \sqrt{1+\frac{1}{2n}}$$

$$\frac{\sqrt{4n+2}}{\sqrt{n+1}+\sqrt{n}} = 1+\frac{1}{4n}$$

and I wanted to see if the numerical values compute, *i.e.* $1+\frac{1}{4n}$ multiplied with $\sqrt{n}+\sqrt{n+1}$ will never produce a number that crosses/straddle an integer point but it was actually a silly idea, there's no easy way to verify this fact and along this silly line of thinking I was also trying to show that the difference between $\lceil\sqrt{n}+\sqrt{n+1}\rceil - \sqrt{n}+\sqrt{n+1}$ is always bigger than $d = [\sqrt{4n+2}] - [\sqrt{n}+\sqrt{n+1}]$ so that adding d to $[\sqrt{n}+\sqrt{n+1}]$ would never cross an integer point.

Problem 3.21. Determine all positive integers n such that $[\sqrt{n}]$ divides n

Say we have a perfect square $m = a^2$, take a number $n = m + k$, if $k = aw$ then $n = a^2 + aw$ is divisible by $a = [\sqrt{n}]$, the caveat here is that to make sure $a = [\sqrt{n}]$, we must restrict w such that $n < (a+1)^2$ which means that

$$
\begin{array}{ccccc}
a^2 & < & n & < & (a+1)^2 \\
0 & < & aw & < & 2a+1 \\
0 & < & w & < & 2+\dfrac{1}{a}
\end{array}
$$

now let's solve for a generously provided by our friend the quadratic formula for $a^2 + aw - n = 0$

$$a = \frac{-w \pm \sqrt{w^2 + 4n}}{2}$$

for a to have an integer solution $w^2 + 4n$ must be a perfect square and as described in Problem 3.20, $4n + j$ can be a perfect square if and only if $j = 0, 1$ so therefore as long as $w = 0, 1, 2$ (which are the valid values for w) the above will have a solution. Next thing we need to check is that the solution is an integer, if w is even then $\sqrt{w^2 + 4n}$ is even and the whole numerator is even, if w is odd $\sqrt{w^2 + 4n}$ is also odd and therefore the whole numerator will be even as well so it checks out.

Also $w = 0$ means that $4n$ is a perfect square which means that n is a perfect square. Therefore as long as either n is a perfect square or $4n + 1$ or $4n + 4$ is a perfect square, n is divisible by $[\sqrt{n}]$.

Problem 3.22. If n is a positive integer, prove that

$$\left[\frac{8n + 13}{25}\right] - \left[\frac{n - 12 - \left[\frac{n-17}{25}\right]}{3}\right]$$

is independent of n

n	$\left[\frac{8n+13}{25}\right]$	$\left[\frac{n-12-\left[\frac{n-17}{25}\right]}{3}\right]$	$\left[\frac{n-17}{25}\right]$	$8n+13 \pmod{25}$	$n-17 \pmod{25}$
-1	0	-4	-1	5	7
0	0	-4	-1	13	8
1	0	-4	-1	$\mathbf{-4}$	9
2	1	-3	-1	4	10
3	1	-3	-1	12	11
4	1	-3	-1	$\mathbf{-5}$	12
5	2	-2	-1	3	13
6	2	-2	-1	11	14
7	2	-2	-1	$\mathbf{-6}$	15
8	3	-1	-1	2	16
9	3	-1	-1	10	17
10	3	-1	-1	$\mathbf{-7}$	18
11	4	0	-1	1	19
12	4	0	-1	9	20
13	4	0	-1	$\mathbf{-8}$	21
14	5	1	-1	$\mathbf{0}$	22
15	5	1	-1	8	23
16	5	1	-1	16	24
$\mathbf{17}$	$\mathbf{5}$	$\mathbf{1}$	$\mathbf{0}$	$\mathbf{-1}$	$\mathbf{0}$
18	6	2	0	7	1
19	6	2	0	15	2
20	6	2	0	$\mathbf{-2}$	3

OK, the strategy here is simple, although initially I wanted to see if I can utilize something from chapter 3 to solve this problem, maybe like putting a sum in front of those two terms and see what I get but let's just stay simple.

The strategy here is to show that when we increase n the change in the first square bracket is the same as the change in the second square bracket, the above table says it all

and it looks like it is true even when n is negative. As for the reason why this works the above table also says it all :)

For the first square bracket, the denominator is 25 and the numerator is $8n$, $8 \cdot 3 = 24$ so roughly every 3 increments the quotient increases by one while for the second square bracket, the -12 does nothing as it's a multiple of 3, the numerator (barring the other square bracket) is just n and thus the quotient increases every 3 increments as well.

The only exception to the rules is when $n = 14$ as $8 \cdot 14 + 13$ is a multiple of 25 and thus the quotient only increases when n increases by 4 instead of three, but this shift in pattern is also followed by the second square bracket thanks to $- \left[\frac{n-17}{25} \right]$, this bracket plays into effect when $n = 17$, it increases its value by one as to delay (due to the minus sign in front of it) the second square bracket from increasing its value by one.

We just need to show this pattern for the first 25 entries because the whole thing repeats every 25 counts, $i.e.$ $n \to n + 25$, as

$$8m + 13 = 25q + r$$

$$\to 8(m + 25) + 13 = 25q + r + (8 \cdot 25)$$

$$8(m + 25) + 13 = 25(q + 8) + r$$

$$\to \left[\frac{8(n + 25) + 13}{25} \right] = \left[\frac{8n + 13}{25} \right] + 8$$

thus the first square bracket increases by 8 every time $n \to n + 25$ while the nested square bracket changes by

$$h - 17 = 25q' + r'$$

$$\to (h + 25) - 17 = 25q' + r' + 25$$

$$(h + 25) - 17 = 25(q' + 1) + r'$$

$$\to \left[\frac{(n + 25) - 17}{25} \right] = \left[\frac{n - 17}{25} \right] + 1$$

so in total the second square bracket also changes by 8 as shown below

$$l - 12 - \left[\frac{l-17}{25}\right] = 3q'' + r''$$

$$\to (l+25) - 12 - \left[\frac{(l+25)-17}{25}\right] = 3q'' + r'' + 25 - 1$$

$$(l+25) - 12 - \left[\frac{(l+25)-17}{25}\right] = 3(q''+8) + r''$$

$$\to \left[\frac{(n+25) - 12 - \left[\frac{(n+25)-17}{25}\right]}{3}\right] = \left[\frac{n - 12 - \left[\frac{n-17}{25}\right]}{3}\right] + 8$$

Problem 3.24. Prove that

$$\sum_{n \leq x} \left[\sqrt{\frac{x}{n}}\right] = \sum_{n \leq \sqrt{x}} \left[\frac{x}{n^2}\right]$$

Here I'll just give the outline of my proof since the pattern is pretty obvious, there's however a much smarter solution due to Greg Hurst that I'll reproduce here.

My strategy is to use induction since the base case $x = 1$ is obvious, the next thing to show is that what happens when x is not an integer, it turns out that for this particular problem $\sum_{n \leq x} = \sum_{n \leq [x]}$ and we'll see why for a minute.

The pattern is more manifest from the RHS, let's see it with a couple of examples, say $x = 13$

$$\sum_{n \leq \sqrt{13}} \left[\frac{x}{n^2}\right] = \sum_{n \leq 3.6055} \left[\frac{x}{n^2}\right] = \left[\frac{13}{1^2}\right] + \left[\frac{13}{2^2}\right] + \left[\frac{13}{3^2}\right]$$

$$= \left[\frac{13}{1}\right] + \left[\frac{13}{4}\right] + \left[\frac{13}{9}\right]$$

from the above example we can learn a few things, one, it is now straightforward to see that $\sum_{n \leq x} = \sum_{n \leq [x]}$ because say we add $0 < \Delta < 1$ to 13 above since Δ is **less than one**, $[\sqrt{13 + \Delta}] = [\sqrt{13}]$ since if x is not a perfect square the only chance is to add one to x, if the Δ is less than one then $x + \Delta$ has no hope of becoming a perfect square, and also

$$\left[\frac{x}{n^2}\right] = \left[\frac{x + \Delta}{n^2}\right]$$

if $0 < \Delta < 1$ for the same reason that if we express $x = n^2 q + r$, since $0 \leq r < n^2$ adding $0 < \Delta < 1$ still maintains the range of $0 < r + \Delta < n^2$ and thus

$$\left[\frac{x + \Delta}{n^2} \right] = \left[\frac{n^2 q}{n^2} + \frac{r + \Delta}{n^2} \right]$$
$$= \left[q + \frac{r + \Delta}{n^2} \right]$$
$$= \left[q + \frac{r}{n^2} \right]$$

since $0 < \frac{r+\Delta}{n^2} < 1$ just as $0 \leq \frac{r}{n^2} < 1$. And here we also see that if x is already a multiple of a perfect square adding $0 < \Delta < 1$ doesn't do anything.

For the LHS it's a bit harder to see that $\sum_{n \leq x} = \sum_{n \leq [x]}$ but it's practically the same, if $\frac{x}{n}$ is not a perfect square adding $0 < \Delta < 1$ to x doesn't do anything because the only way $\left[\sqrt{\frac{x+\Delta}{n}} \right] \neq \left[\sqrt{\frac{x}{n}} \right]$ is if $\frac{x+\Delta}{n}$ is a perfect square while $\frac{x}{n}$ is not.

But if $\frac{x}{n}$ is not a perfect square, to make it a perfect square we have to increment $\frac{x}{n}$ by at least one, in our case however, $0 < \frac{\Delta}{n} < 1$ so it doesn't do anything. And of course if $\frac{x}{n}$ is already a perfect square, adding $0 < \Delta < 1$ to x doesn't change anything.

So we have seen that the case for non-integer x is the same as integer x, therefore we can just consider integer x, so let's begin our induction. The base case is $x = 1$ and it's self-evident :)

Notice that increasing $x \to x + 1$ will increase both LHS and RHS by at *least* one. For the RHS this is due to $\left[\frac{x}{1^2} \right]$, increasing x bby one will definitely increase this term, for the LHS this is because increasing x by one means that we have one more value for the dummy index n and the additional term $\sqrt{\frac{x+1}{n=x+1}}$ is a new term for $x + 1$ that wasn't there for x.

Now what happens if we increase $x \to x + 1$? For the RHS it's simple, if x were not a multiple of some n^2 and $x + 1$ is then $\left[\frac{x+1}{n^2} \right] = \left[\frac{x}{n^2} \right] + 1$ otherwise there's no change for that **particular** n^2.

How about the LHS? it's pretty much the same although not as apparent as the RHS. For $\sqrt{\frac{x+1}{n}} \neq \sqrt{\frac{x}{n}}$ for a **particular** n, we must have $\frac{x+1}{n}$ to be a perfect square while $\frac{x}{n}$ is not.

To see it more clearly let us relabel things, say for the RHS we are focusing on a particular $n^2 = m^2$, so now for the LHS we want to see when $\frac{x}{n}$ will be equal m^2, not that

x is the same for both LHS and RHS, so if x is not a multiple of m^2, on the RHS $\left[\frac{x}{m^2}\right]$ and $x + 1$ is, then on the RHS we get an increment of 1.

On the LHS since $m^2 \nmid x$, there's no n such that $\frac{x}{n}$ is m^2, but since $m^2 | x + 1$, there's now an $n = y$ on the LHS such that $\frac{x+1}{y} = m^2$ and thus $\left[\sqrt{\frac{x+1}{y}}\right] = \left[\sqrt{\frac{x}{y}}\right] + 1$ and so the LHS also increases by one. Note that $\frac{x+1}{y}$ must be exactly m^2 due to the square root

But now there's a catch, if $y > 1$ then what happens to all $w \neq y$, $\left[\sqrt{\frac{x+1}{w}}\right]$? will they also change? Well, the answer is no because $x + 1 = ym^2$ so it can't also be wm^2 with $w \neq y$.

Next question will be why $\frac{x+1}{y}$ has to be exactly m^2 why not some km^2? Now if k is not a perfect square what we have is that neither $\frac{x+1}{y}$ nor $\frac{x}{y}$ is a perfect square and by the reasoning explained above, increasing x doesn't do anything, $i.e.$ $\left[\sqrt{\frac{x+1}{y}}\right] = \left[\sqrt{\frac{x}{y}}\right]$.

Now how about $k = s^2 \rightarrow \frac{x+1}{y} = s^2m^2$? in this case there would be $n = s^2$ such that $\frac{x+1}{s^2} = m^2$ and this is possible since obviously $s^2 < x + 1$. But this also means that $\frac{x}{s^2}$ is not a perfect square and increasing $x \rightarrow x + 1$ has the same effect as discussed above except that we have s^2 instead of m^2. There's also $n = m^2$ such that $\frac{x+1}{m^2} = s^2$ and in this case we have another increment in the LHS and of course there's also $\frac{x+1}{y} = s^2m^2$ and so we have three separate increments of $\left[\sqrt{\frac{x}{n}}\right]$ in the LHS which are also mirrored on the RHS, $\left[\frac{x+1}{s^2}\right]$, $\left[\frac{x+1}{m^2}\right]$, $\left[\frac{x+1}{s^2m^2}\right]$, so the increments are always in sync between LHS and RHS.

And that's all there is to it, to reduce obscurity an example is fitting. Take for example going from $x = 7 \rightarrow 8 \rightarrow 9$ for the LHS with $x = 7$ we have

$$\left[\sqrt{\frac{7}{2}}\right] = \left[\sqrt{\frac{7}{3}}\right] = \ldots = \left[\sqrt{\frac{7}{7}}\right] = 1$$

while

$$\left[\sqrt{\frac{7}{1}}\right] = 2$$

while the RHS is

$$\left[\frac{7}{1}\right] + \left[\frac{7}{4}\right]$$

now going from $7 \rightarrow 8$ we have on the LHS

$$\left[\sqrt{\frac{8}{3}}\right] = \left[\sqrt{\frac{8}{4}}\right] = \ldots = \left[\sqrt{\frac{8}{8}}\right] = 1$$

while
$$\left[\sqrt{\frac{8}{1}}\right] = \left[\sqrt{\frac{8}{2}}\right] = 2$$

because now while $\frac{7}{2} \neq 4$, $\frac{8}{2} = 4$ and so here $y = 2$, $x = 7$, and $m^2 = 4$ and as advertised
$$\left[\sqrt{\frac{x+1}{y}}\right] = \left[\sqrt{\frac{x}{y}}\right] + 1$$

only for $y = 2$ and for the RHS we have
$$\left[\frac{8}{4}\right] = \left[\frac{7}{4}\right] + 1$$

Now going from $8 \to 9$, with $x = 9$ we have
$$\left[\sqrt{\frac{9}{3}}\right] = \left[\sqrt{\frac{9}{4}}\right] = \ldots = \left[\sqrt{\frac{9}{9}}\right] = 1$$

but now
$$\left[\sqrt{\frac{9}{1}}\right] = 3 \qquad \left[\sqrt{\frac{9}{2}}\right] = 2$$

and in this case $x = 8$, $y = 1$, and $m^2 = 9$
$$\left[\sqrt{\frac{x+1}{y}}\right] = \left[\sqrt{\frac{x}{y}}\right] + 1$$

only for $y = 1$ and for the RHS we have and additional term
$$\left[\frac{9}{9}\right] = 1$$

apart from $\left[\frac{9}{4}\right]$ and $\left[\frac{9}{1}\right]$.

The clever proof by Greg Hurst utilizes Theorem 3.11 and by introducing a new square indicator arithmetical function, $s(n)$ where $s(n) = 1$ if n is a perfect square and 0 otherwise, then

$$\left[\sqrt{x}\right] = \# \text{ of perfect squares } \leq x = S(x) = \sum_{n \leq x} s(n)$$

the above is an astute and important observation, in that $\left[\sqrt{x}\right]$ itself already counts the number of perfect squares less than x.

From Theorem 3.11 that states, if $F(x) = \sum_{n \leq x} f(n)$ we have

$$\sum_{n \leq x} \sum_{d \mid n} f(d) = \sum_{n \leq x} f(n) \left[\frac{x}{n}\right] = \sum_{n \leq x} F\left(\frac{x}{n}\right)$$

so here we set $F(x) = S(x) = [\sqrt{x}]$

$$\sum_{n \leq x} \left[\frac{x}{n}\right] = \sum_{n \leq x} S\left(\frac{x}{n}\right) = \sum_{n \leq x} s(n) \left[\frac{x}{n}\right]$$

but since $s(n)$ is either 1 or 0 and it's only 1 if $n = m^2$ we have

$$\sum_{n \leq x} s(n) \left[\frac{x}{n}\right] = \sum_{m^2 \leq x} \left[\frac{x}{m^2}\right] = \sum_{m \leq \sqrt{x}} \left[\frac{x}{m^2}\right]$$

So at first glance it seems to be a lot simpler to utilize Theorem 3.11 to solve this problem, however, recall that to prove Theorem 3.11 we need to have the concept of arithmetical functions, Dirichlet products, and generalized Dirichlet products, to say the least while induction is well basic :) but on the other hand, we do need proper tools to solve problems and in this case induction might not be the proper tool.

Chapter 4

Page 80, Theorem 4.5, Eq (12)m the proof of this theorem assumes that Eq (11) is true, *i.e.* $\lim_{x \to \infty} \frac{\pi(x) \log x}{x} = 1$ is true, but if this is true we can immediately substitute it, $\lim_{x \to \infty} \pi(x) = \frac{x}{\log x}$, to Eq (12)

$$
\begin{aligned}
\lim_{x \to \infty} \frac{\pi(x) \log \pi(x)}{x} &= \lim_{x \to \infty} \frac{\pi(x) \log \left(\frac{x}{\log x}\right)}{x} \\
&= \lim_{x \to \infty} \frac{\pi(x) \log x}{x} - \frac{\pi(x) \log \log x}{x} \\
&= \lim_{x \to \infty} \frac{\pi(x) \log x}{x}
\end{aligned}
$$

since $\lim_{x\to\infty} \frac{\pi(x)\log\log\ x}{x} = 0$ which can be easily seen by applying L'Hospital's rule

$$\lim_{x\to\infty} \frac{\pi(x)\log\log\ x}{x} = \lim_{x\to\infty} \left(\frac{\cancel{x}}{\log\ x}\right)\left(\frac{\log\log\ x}{\cancel{x}}\right)$$

$$= \lim_{x\to\infty} \frac{\log\log\ x}{\log\ x} \quad\longrightarrow\quad \text{apply L'Hospital next}$$

$$= \lim_{x\to\infty} \frac{\left(\frac{1}{\log\ x}\right)\left(\frac{1}{x}\right)}{\frac{1}{x}}$$

$$= \lim_{x\to\infty} \frac{1}{\log\ x}$$

$$= 0$$

Page 81, somewhere mid page, this is quite a neat trick

$$\frac{p_{n+1}}{n\log\ n} = \frac{p_{n+1}}{(n+1)\log\ (n+1)}\frac{(n+1)\log\ (n+1)}{n\log\ n}$$

but we have assumed that $\lim_{n\to\infty} \frac{p_n}{n\log\ n} = 1$ thus $\lim_{n+1\to\infty} \frac{p_{n+1}}{(n+1)\log\ (n+1)} = 1$ as well, the other term we use L'Hospital again

$$\lim_{n\to\infty} \frac{(n+1)\log\ (n+1)}{n\log\ n} = \lim_{n\to\infty} \frac{\log\ (n+1) + (n+1)/(n+1)}{\log\ n + n/n}$$

$$= \lim_{n\to\infty} \frac{\log\ (n+1) + 1}{\log\ n + 1}$$

$$= \lim_{n\to\infty} \frac{\frac{1}{n+1}}{\frac{1}{n}}$$

$$= 1$$

where we have used L'Hospital twice.

Page 82, PROOF of Theorem 4.6, "and the other one is easily verified by induction" and this refers to $2^n < \binom{2n}{n}$.

For the RHS, going from $n \to n+1$ we get

$$\binom{2(n+1)}{(n+1)} = \frac{2(n+1)!}{(n+1)!(n+1)!}$$

$$= \frac{(2n+2)(2n+1)(2n)!}{(n+1)(n+1)n!n!}$$

$$= 2\left(\frac{2n+1}{n+1}\right)\binom{2n}{n}$$

$$> 2\binom{2n}{n}$$

53

while for the LHS $2^{n+1} = 2 \cdot 2^n$ and thus the RHS is always bigger.

Page 82, PROOF of Theorem 4.6, last equation bottom of page, m only needs to run from 1 to

$$p^m = n$$

$$\log \ p^m = \log \ n$$

$$m = \frac{\log \ n}{\log \ p}$$

page 83, Eq (18), in the last inequality he has used the fact that

$$\frac{\log 2}{2} = 0.34657 > \frac{1}{4}$$

but it is also bigger than $\frac{1}{3}$, if it's me I'd choose $\frac{\log 2}{2} > \frac{1}{3}$ for a tighter bound.

Next, for $\pi(2n + 1)$, for the last inequality

$$\frac{2n}{2n+1} \geq \frac{2}{3}$$

is because the minimum is achieved when $n = 1 \to \frac{2n}{2n+1} = \frac{2}{3}$, but if had we used $\frac{1}{3}$ for Eq(18) we would get as the lower bound

$$\frac{2}{9}\frac{n}{\log n} < \pi(n)$$

which is slightly better than $\frac{1}{6}$:) and if we are willing to put a condition, by $n = 13 \to \pi(2n + 1) = \pi(27)$

$$\frac{\log 2}{2} \cdot \frac{2n}{2n+1} \to \frac{\log 2}{2}\frac{26}{27} = 0.33373\ldots > \frac{1}{3}$$

so we have

$$\pi(n) > \frac{1}{3}\frac{n}{\log n}, \ \text{for } n \geq 27$$

however, we don't actually need the qualifier $n \geq 27$ after all, if we list all values of $\pi(n)$ and $\frac{1}{3}\frac{n}{\log n}$ for $2 \leq n \leq 27$, it is immediately evident that $\pi(n) > \frac{1}{3}\frac{n}{\log n}$ and so it is true for all $n \geq 2$.

Using the technique of the PROOF of Theorem 4.6 the best we can do is when

$$\lim_{n\to\infty} \frac{\log 2}{2}\frac{2n}{2n+1} = \frac{\log 2}{2} = 0.34657359\ldots$$

54

and so $\frac{1}{3}\frac{n}{\log n}$ seems to be the best we can get

Page 83, mid page, "For those primes p in the interval $n < p \leq 2n$ we have $[2n/p] - 2[n/p] = 1$", just note that $p > n$ and so $[n/p] = 0$ while for $[2n/p]$ if $p = 2n$ it's 1, otherwise since $p > n$, $2n/p$ it will be greater than 1 but less than 2.

Page 83, bottom page, "the sum of the left telescopes", a telescoping series is a series whose partial sums eventually only have a fixed number of terms after cancellation (taken from Wikipedia). In this case

$$\{\vartheta(2^{k+1}) - \vartheta(2^k)\} + \{\vartheta(2^k) - \vartheta(2^{k-1})\} + \ldots + \{\vartheta(2^1) - \vartheta(2^0)\}$$
$$= \vartheta(2^{k+1}) - \vartheta(2^0)$$
$$= \vartheta(2^{k+1}) - \vartheta(1)$$

but

$$\vartheta(x) = \sum_{p \leq x} \log\ p$$
$$\to \vartheta(1) = 0$$

on the RHS

$$2^{k+1}\log\ 2 + 2^k \log\ 2 + \ldots + 2^{0+1}\log\ 2 = \left(2^{k+1} + 2^k + \ldots + 2\right)\log\ 2$$
$$= 2\left(2^k + 2^{k-1} + \ldots + 1\right)\log\ 2$$
$$= 2\left(\frac{2^{k+1} - 1}{2 - 1}\right)\log\ 2$$
$$= (2^{k+2} - 2)\log\ 2$$
$$< 2^{k+2}\log\ 2$$

Page 84, "Taking $\alpha = 2/3$", near the end of the PROOF of Theorem 4.7, the way I think he's got this number is to take the derivative with respect to α and setting it to zero

$$0 = \frac{d}{d\alpha}\left(\frac{4\log 2}{\alpha} + \frac{1}{e(1 - \alpha)}\right) = -\frac{4\log 2}{\alpha^2} + \frac{1}{e(1 - \alpha)^2}$$
$$\to e(1 - \alpha)^2 4\log 2 = \alpha^2$$
$$\to \frac{4e\log 2 \pm 2\sqrt{e\log 2}}{4e\log 2 - 1} = \alpha$$

55

which gives actual numerical values of $\alpha_+ = 1.572967072$ or $\alpha_- = 0.732998762$, substituting these valus back

$$\frac{4\log 2}{\alpha_+} + \frac{1}{e(1-\alpha_+)} = \frac{\sqrt{e}\log 2(4e\log 2 - 1)^2}{\sqrt{e}\log 2(4e^2\log 2 + e) + 4e^2\log 2} = 1.120588573$$

$$\frac{4\log 2}{\alpha_-} + \frac{1}{e(1-\alpha_-)} = \frac{\sqrt{e}\log 2(4e\log 2 - 1)^2}{\sqrt{e}\log 2(4e^2\log 2 + e) - 4e^2\log 2} = 5.160347754$$

so the choice of $2/3$ for α is so that we can get the constant factor of 6, to get a tighter bound we can choose α_+, however, $0 < \alpha < 1$ and so we have to choose α_- and in this way we'll get $\pi(n) < 5.160347754\frac{n}{\log n}$ and so we have tighter lower and upper bounds

$$\frac{1}{3}\frac{n}{\log n} < \pi(n) < 5.160347754\frac{n}{\log n}$$

Page 84, last step of PROOF of Theorem 4.6, just some comments

$$6\log 2 + \frac{3}{e} = 5.2625$$

this is definitely less than 6 and will give a tighter bound but 6 looks neater :) However, this will give a better bound for Theorem 4.7 as well so the upper bound will become

$$p_n < \left(12\log 2 + \frac{6}{e}\right)\left(n\log n + n\log\frac{\left(12\log 2 + \frac{6}{e}\right)}{e}\right)$$

But as I've shown above the lower bound can be improved further to $\frac{1}{3}\frac{n}{\log n} < \pi(n)$ and thus the upper bound of Theorem 4.7 can be reduced to

$$p_n < 6\left(n\log n + n\log\frac{6}{e}\right)$$

Page 82-84, some comments on the PROOF of Theorem 4.6, this is a pretty sick proof, so far it's the longest in the book, it started with something unrelated to $\pi(x)$, *i.e.* $\binom{2n}{n}$, and it moves on from there.

One key point that is used repeatedly is the fact that $\pi(n) \le n$, the other is we need a summation involving only primes $\sum_{p \le x}$, this we can get from $\vartheta(x)$ or sums involving von Mangoldt functions

But the final step utilizing $0 < \alpha < 1 \to (\pi(n) - \pi(n^\alpha))\log n^\alpha$ is really sick, not to mention using the fact that the function $f(x) = x^{-c}\log x$ has a peak at $x = e^{1/c}$. There's a lot to learn here people ....

56

Page 85, top of page, another ridiculous trick, $\log x \leq (2/e)\sqrt{x}$, let's see how this might come about, maybe we need something involving $\sqrt{p_n}$, and we want to see if we have a function $f(x) = C\sqrt{x} - \log x$ we want to find out what constant C will guarantee that $f(x) > 0$, $x \geq 1$, first let's take the derivative

$$\frac{df(x)}{dx} = C/2\frac{1}{\sqrt{x}} - \frac{1}{x}$$

it's zero when

$$C/2\frac{1}{\sqrt{x}} - \frac{1}{x} = 0$$
$$C/2x = \sqrt{x}$$
$$C^2/4x^2 = x$$
$$\rightarrow x = 4/C^2$$

and also when $x \rightarrow \infty$, now let's see if this is a maximum or minimum

$$C/2\frac{1}{\sqrt{x}} - \frac{1}{x} \rightarrow -C/4\frac{1}{(\sqrt{x})^3} + \frac{1}{x^2}$$
$$= -C/4\frac{1}{\left(\sqrt{4/C^2}\right)^3} + \frac{1}{(4/C^2)^2}$$
$$= \frac{-C^4}{32} + \frac{C^4}{16}$$
$$= \frac{C^4}{32}$$

since this is positive it is a minimum point, good so now we need to make sure that this point is positive to ensure that $C\sqrt{x} - \log x \geq 0$

$$C\sqrt{4/C^2} - \log(4/C^2) = 2 - 2\log(2/C)$$
$$= 2 - 2\log 2 + 2\log C$$

and what we want is

$$2 - 2\log 2 + 2\log C \geq 0$$
$$\log C \geq \log 2 - 1$$
$$C \geq \frac{2}{e}$$

but since we want the tightest bound we pick the smallest $C = 2/e$ so at least some of the mystery of this $2/e$ number is solved :)

Page 85, upper bound of Theorem 4.7 can be improved (significantly?) by noting that the lower bound of Theorem 4.6 gives the upper bound for Theorem 4.7 and the upper bound for Theorem 4.6 gives the lower bound for Theorem 4.7.

Above the lower bound of Theorem 4.6 has been improved from $\frac{1}{6}\frac{n}{\log n} \to \frac{1}{3}\frac{n}{\log n}$ but I believe we can do better, from the empirical values of the prime counting function $\pi(n)$ we see that $\pi(n) > \frac{n}{\log n}$ for $n \geq 11$ and so

$$\pi(p_n) > \frac{n}{\log n}$$
$$n > \frac{p_n}{\log p_n}$$
$$\to p_n < n \log p_n$$

at least for $n \geq 11$ but direct inspection shows that this is already true for $n \geq 4$, to include all n we need $p_n < 2.89n \log p_n$ which is roughtly $3 \log p_n$ which was the implication of $\frac{1}{3}\frac{n}{\log n} < \pi(n)$ as shown above, so I'll take $n \geq 4$:)

Continuing with $p_n < n \log p_n$ with the steps outlined in the book we get

$$p_n < 2n \log n + 2n \log \frac{2}{e}$$

But four numbers are not too much to check, so listing all the numbers we see that with a slight modification we get

$$p_n \leq 2n \log n + 2n$$

with equality only for $n = 1, p_n = 2$, the only dodgy thing about this assertion is that we based it on the empirical fact of $\frac{n}{\log n} < \pi(n)$ for $n \geq 11$ and it's not a solid proof. Combining this with our earlier result $5.160347754\frac{n}{\log n} < \pi(n)$ we get

$$0.193785389 \, n \log n < p_n < 2n \log n + 2n$$

where 0.193785389 is just $1/5.160347754$.

Page 85, before Section 4.6, "$\sum_{n=1}^{\infty} 1/p_n$ diverges, by comparison with $\sum_{n=2}^{\infty} 1/(n \log n)$"

I think here we need to use the limit comparison test, since $p_n < 12(n \log n + n \log 12/e)$, and we can do this because $a_n, b_n > 0$ where $\sum a_n \equiv \sum_{n=2}^{\infty} 1/(n \log n)$ and

58

$\sum b_n \equiv \sum_{n=1}^{\infty} 1/\{12(n \log\ n + n \log\ 12/e)\}$. Note that since $p_n < 12(n \log\ n + n \log\ 12/e)$, $1/p_n > 1/12(n \log\ n + n \log\ 12/e)$, so if the series $\sum_{n=1}^{\infty} 1/\{12(n \log\ n + n \log\ 12/e)\}$ diverges so does the series for $1/p_n$.

What we want to show is first that $\sum_{n=2}^{\infty} 1/(n \log\ n)$ diverges, we can do this using the integral test

$$\int_{2}^{\infty} \frac{1}{x \log x} dx = \log \log x \big|_{2}^{\infty}$$
$$= \log \log \infty - \log \log 2$$
$$= \infty$$

so $\sum_{n=2}^{\infty} 1/(n \log\ n)$ diverges, next we calculate the ratio $c = \lim_{n \to \infty} \frac{a_n}{b_n}$

$$\lim_{n \to \infty} \frac{\frac{1}{n \log\ n}}{\frac{1}{12(n \log\ n + n \log\ 12/e)}} = \lim_{n \to \infty} \frac{12n \log\ n}{n \log\ n} + \frac{12n \log\ 12/e}{n \log\ n}$$
$$= \lim_{n \to \infty} 12 + \frac{12 \log\ 12/e}{\log\ n}$$
$$= 12$$

so since this is a positive finite number and one series diverges the other must too.

Some comments on Theorem 4.6 and Theorem 4.7, taking the limit of the ratio of the upper bound we get

$$\lim_{n \to \infty} \frac{12n \log n + 12n \log \frac{12}{e}}{p_n} = \lim_{n \to \infty} 12 \frac{p_n}{p_n \log p_n} \log \left(\frac{p_n}{\log p_n} \right) + 12 \frac{p_n}{p_n \log p_n} \log \frac{12}{e}$$
$$= \lim_{n \to \infty} 12 \frac{\log p_n}{\log p_n} - 12 \frac{\log \log p_n}{\log p_n} + \frac{12}{\log p_n} \log \frac{12}{e}$$
$$= 12$$

since $\lim_{x \to \infty} \frac{\log \log x}{\log x} = 0$ just by doing L'Hospital twice and note that here $n = \pi(p_n)$ and so we can use the asymptotic form of $\pi(p_n)$. This gives me an idea, this means that

$$\lim_{n \to \infty} \frac{n \log n}{p_n} = \lim_{n \to \infty} \frac{p_n}{p_n \log p_n} \log \left(\frac{p_n}{\log p_n} \right)$$
$$= \lim_{n \to \infty} \frac{\log p_n}{\log p_n} - \frac{\log \log p_n}{\log p_n}$$
$$= 1$$

and so $n \log n$ actually gives quite a good estimate of p_n when n is big enough but when n is small it's no good of course, *e.g.* $n \log n = 0$ when $n = 1$, even when $n = 10000$,

$n \log n = 92103.4$ while $p_{10000} = 104729$ but we are close enough and it gives a starting point to search for the next prime number. The curious thing is that up to $n = $ one billion, $n \log n < p_n$ still, does this mean that our lower bound is no good? ¯\\(°_o)/¯

Another interesting thing is that if we take the limit of the ratio of the difference of the upper and lower bound to p_n itself we have

$$\lim_{n\to\infty} \frac{12n \log n + 12n \log \frac{12}{e} - \frac{1}{6}n \log n}{p_n} = \lim_{n\to\infty} 11\tfrac{5}{6}\frac{\cancel{p_n}}{\cancel{p_n} \log p_n} \log\left(\frac{p_n}{\log p_n}\right) + 12\frac{\cancel{p_n}}{\cancel{p_n} \log p_n} \log \frac{12}{e}$$
$$= \lim_{n\to\infty} 11\tfrac{5}{6}\frac{\cancel{\log p_n}}{\cancel{\log p_n}} - 11\tfrac{5}{6}\frac{\log \log p_n}{\log p_n} + \frac{12}{\log p_n} \log \frac{12}{e}$$
$$= 11\tfrac{5}{6}$$

so the **difference** between the upper and lower bounds is bigger than the prime number itself, almost 12 times bigger.

Another interesting find is that when we set a constant multiplier $Cn \log n$, if $C = 1$ it seems that $n \log n < p_n$, but if $C = 1.1$, $Cn \log n > p_n$ when $n > 43$ so once $C > 1$, $Cn \log n > p_n$ after some n, or in other words

Proposition, if $C > 1$ then there's a positive $n_0 > 1$ such that $Cn \log n > p_n$ whenever $n \geq n_0$ and $Cn \log n < p_n$ when $n < n_0$ and if $C \leq 1$ then $Cn \log n < p_n$ for all n

Incomplete Proof, the reasoning goes as follows, we know that

$$\lim_{n\to\infty} \frac{Cn \log n}{p_n} = \lim_{n\to\infty} \frac{C\cancel{p_n}}{\cancel{p_n} \log p_n} \log\left(\frac{p_n}{\log p_n}\right)$$
$$= \lim_{n\to\infty} \frac{C\cancel{\log p_n}}{\cancel{\log p_n}} - \frac{C \log \log p_n}{\log p_n}$$
$$= C$$

we exclude $n_0 = 1$ because $\log 1 = 0$:) Now if $C2 \log 2 < p_2$ and from above we know that as $n \to \infty$, $Cn \log n = Cp_n > p_n$, and if $C > 1$ then it must mean that there's some n_0 where the inequality changes sign, and this point is our n_0. And of course if $C2 \log 2 > p_2$ then $n_0 = 2$.

The only thing I can't prove right now is that this ratio $(Cn \log n)/p_n$ is a monotonic smooth function and it doesn't vacillate too much from less than one to greater than one. But I believe the following is true, for $n \geq 1$

$$n \log n < p_n \leq 2n \log n + 2$$

where equality is only achieved when $n = 1$.

One last thing I want to say is that Theorem 4.6 and 4.7 do not use anything special to deduce the bounds for $\pi(n)$ and p_n, not even complex analysis, they are just using careful observations about arithmetical functions which make them really cool theorems :)

Page 86, third last equation, this is a trick that is used over and over again

$$\sum_{x/2 < n \leq x} \left[\frac{x}{n}\right] = \sum_{x/2 < n \leq x} 1$$

the thing to note here is the limits of the sum, since $x/2 < n$, this meant that $x/n < 2$ and since $n \leq x$, this also means that x/n is at least one, and so since this is less than two and at least one, $[x/n] = 1$.

Page 87, top of page, this telescoping series technique is used again here, it was first used in the PROOF of Theorem 4.6, bottom of Page 83, so looks like it is a useful tool to be kept in our arsenal :)

Page 90, mid page, "Therefore if we take $f(t) = 1/\log t$ in Theorem 4.2", Theorem 4.2 is Abel's lemma, Eq (4)

$$\sum_{y < n \leq x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)dt$$

with $A(x) = \sum_{n \leq x} a(n)$. The confusing thing here is that the same symbols in Theorem 4.2 are used for different things in Theorem 4.12, here $a(n)$ in Theorem 4.2 is $1/p = (\log p/p) \times (1/\log p)$

$$\sum_{2^- < n \leq x} \frac{\log p}{p} \frac{1}{\log p} = \frac{A(x)}{\log x} - \frac{A(2^-)}{\log 2^-} + \int_2^x A(t)f'(t)dt$$

where 2^- is something less than 2 such that 2 is included in the sum, but then $A(2^-) = 0$.

Page 92, the derivation of Eq (34) bottom half of page, here Theorem 3.10 is utilized, Theorem 3.10 is

$$H(x) = \sum_{n \leq x} f(n)G\left(\frac{x}{n}\right) = \sum_{n \leq x} g(n)F\left(\frac{x}{n}\right)$$

61

as mentioned in the text, here $f = \mu$, $g = \Lambda$ and so

$$-\sum_{n \leq x} \mu(n) \log n = \sum_{n \leq x} \sum_{d|n} \mu(d) \Lambda \left(\frac{n}{d}\right)$$

$$= \sum_{n \leq x} \mu * \Lambda = \sum_{n \leq x} f * g, \quad h = f * g$$

$$= \sum_{n \leq x} h(n)$$

$$-\sum_{n \leq x} \mu(n) \log n = H(x)$$

but from Theorem 3.10 we have

$$H(x) = -\sum_{n \leq x} \mu(n) \log n = \sum_{n \leq x} f(n) G \left(\frac{x}{n}\right)$$

$$= \sum_{n \leq x} \mu(n) \sum_{m \leq x/n} \Lambda(m)$$

$$= \sum_{n \leq x} \mu(n) \psi \left(\frac{x}{n}\right)$$

where we have used the definition of Chebyshev's function, Page 75, $\psi(x) = \sum_{n \leq x} \Lambda(n)$.

Page 92, bottom page last inequality, this is just the definition of limit for $x \to \infty$, the definition of limit for $x \to a$ with a finite is, $\lim_{x \to a} f(x) = L$ if for every number $\varepsilon > 0$ there's a number $\delta > 0$ such that $|f(x) - L| < \varepsilon$ whenever $0 < |x - a| < \delta$.

While the definition of limit $x \to -\infty$ is $\lim_{x \to \infty} f(x) = L$ if for every number $\varepsilon > 0$ there is $N < 0$ such that $|f(x) - L| < \varepsilon$ whenever $x < N$.

Page 95, middish page, "express each summand", the Mobius inversion formula is

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

$$\sigma_0 = \sum_{d|n} 1 \longleftrightarrow 1 = \sum_{d|n} \sigma_0(d) \mu \left(\frac{n}{d}\right) = \sum_{d|n} \mu(d) \sigma_0 \left(\frac{n}{d}\right)$$

since $f * g = g * f$. Next, $\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d}$, here we need to start with Theorem 2.10, $\log n = \sum_{d|n} \Lambda(d)$ and then use the Mobius inversion formula

$$\log n = \sum_{d|n} \Lambda(d) \longleftrightarrow \Lambda(n) = \sum_{d|n} (\log d) \mu \left(\frac{n}{d}\right) = \sum_{d|n} \mu(d) \log \left(\frac{n}{d}\right)$$

62

while $\left[\frac{1}{n}\right] = \sum_{d|n} \mu(d)$ is just Theorem 2.1 or just the restatement of $\mu * u = I$ where $I(n) = \left[\frac{1}{n}\right]$.

Page 95, "This implies (38)", mid bot page, what it's saying is that

$$[x] - \psi(x) - 2C = \sum_{\substack{q,d \\ qd \leq x}} \mu(d)f(q)$$

$$\to \psi(x) = [x] - \sum_{\substack{q,d \\ qd \leq x}} \mu(d)f(q) - 2C$$

$$= x + O(1) - \sum_{\substack{q,d \\ qd \leq x}} \mu(d)f(q) - 2C$$

$$\psi(x) = x - \sum_{\substack{q,d \\ qd \leq x}} \mu(d)f(q) + O(1)$$

where we have used the fact that $[x] = x + O(1)$ and $O(1) - 2C = O(1)$

Page 96, expansion of $F(x)$ top mid page, the expansion of $\sum_{n \leq x} \log n$ is given on top of Page 68 using the Euler's summation formula

$$\sum_{n \leq x} \log n = \int_1^x \log t \, dt + \int_1^x \frac{t - [t]}{t} dt - (x - [x]) \log x$$

$$= x \log x - x + 1 + \int_1^x \frac{t - [t]}{t} dt + O(\log x)$$

and since $t - [t] \leq 1$ we have

$$\int_1^x \frac{t - [t]}{t} dt = O\left(\int_1^x \frac{1}{t} dt\right) = O(\log x)$$

and so

$$1 + \int_1^x \frac{t - [t]}{t} dt + O(\log x) = 1 + O(\log x) + O(\log x) = O(\log x)$$

$$\to x \log x - x + 1 + \int_1^x \frac{t - [t]}{t} dt + O(\log x) = x \log x - x + O(\log x)$$

while for $-2C \sum_{n \leq x} 1$ we use Theorem 3.2 (d), $\sum_{n \leq x} n^\alpha = \frac{x^{\alpha+1}}{\alpha+1} + O(x^\alpha)$, $\alpha \geq 0$ so in this case $\alpha = 0$ and that's why we get the $O(1)$.

Page 96, "$|F(x)| \leq B\sqrt{x}$ for all $x \geq 1$", why does x starts with 1? it is because of the definition of the Big Oh notation, see page 53, "If $g(x) > 0$ for all $x \geq a \dots$", here $g(x) = \sqrt{x}$, and $\sqrt{x} > 0$ starting from $x \geq 1$ (we are considering integer x here)

Page 96, Eq (41), to get the last equality please recall that $ab = x$ as mentioned on bottom of Page 95

Page 96, bottom half of page, "Since $M(x) = O(x)$ as $x \to \infty$", it's interesting that he didn't even use the definition of the limit for $\lim_{x \to \infty} M(x)/x = 0$ although that is what we are trying to prove, that $\lim_{x \to \infty} M(x)/x = 0$ implies the prime number theorem but here we are just using the definition of Big Oh.

The definition of Big Oh is that for all $x \geq c$ such that $g(x) > 0$ ..., here $g(x) = x$ since $f(x) = O(x)$ where $f(x) = M(x)$, so that must mean there's a positive constant W such that

$$|M(x)| \geq Wx \quad \text{for all } x \geq c$$

in this case the constant $W = \varepsilon/K$, the purpose of this whole Big Oh thing is because we want to re-use ε for all three terms on RHS of Eq (40), so in this case we set W and go about looking for c.

Had he used the definition of limit, he wouldn't have achieved much, this is because the definition of limit is quite strict, "for **every** number $\varepsilon > 0$", while with the Big Oh we can fix ε to a certain value.

Page 97, after Eq (44), the chain of inequalities, recall that

$$M(x) = \sum_{n \leq x} \mu(n)$$

and since $\mu(n)$ is just $\pm 1, 0$, $M(x) < x$ and so we can replace $M(b)$ with $< b$, this is how we get the second inequality. To get third inequality we must recall from Page 95 (before Eq (42)) that

$$\frac{A}{\sqrt{a}} < \varepsilon$$

and so $A < \varepsilon\sqrt{a}$, but the assumption here is that since $x = ab$ we choose a such that $a < b$ or $x > a^2$ and so we can replace $\varepsilon\sqrt{a}$ with $< \varepsilon\sqrt{b}$.

Page 92-97, comments on Theorem 4.14 and Theorem 4.15, note that what we are trying to prove here is that the prime number theorem implies $\lim_{x \to \infty} M(x)/x = 0$ and vice versa, but note that in the PROOF of Theorem 4.14 what we are doing was

$$\text{prime number theorem} \implies \lim_{x \to \infty} \frac{\psi(x)}{x} = 1 \implies \lim_{x \to \infty} \frac{H(x)}{x \log x} = 0 \implies \lim_{x \to \infty} \frac{M(x)}{x} = 0$$

64

so it's quite a treacherous path :) while for the PROOF of Theorem 4.15 it's even crazier as we start with

$$\psi(x) = x - \sum_{\substack{q,d \\ qd \leq x}} \mu(d)f(q) + O(1)$$

to Theorem 3.17

$$\sum_{\substack{q,d \\ qd \leq x}} \mu(d)f(q) = \sum_{n \leq b} \mu(n)F\left(\frac{x}{n}\right) + \sum_{n \leq a} f(n)M\left(\frac{x}{n}\right) - F(a)M(b)$$

where $M(x) = \sum_{n \leq x} \mu(n)$ and then we used various other Theorems and finally utilized $\lim_{x \to \infty} M(x)/x = 0$ in the form of $M(x) = o(x)$ as $x \to \infty$ (and there's a typo in the book, it uses big Oh instead of small oh) to finally prove the prime number theorem in the form of $\psi(x) \sim x$:)

Another important thing needed for Theorem 4.15 is that since we deal with the three terms of the RHS of Eq (40) separately, the condition for x, *i.e.* for the first term $x \geq 1$, for the second $x > c$, and for the third and last one $x > a^3, x > ac$, these conditions on x have to be consistent, so the argument used to prove Theorem 4.15 is actually a quite delicate one.

Page 97, Eq (45), this is the small *oh* notation, in this case

$$\lim_{x \to \infty} \frac{A(x)}{1} = 0$$

so it just means $\lim_{x \to \infty} A(x) = 0$:)

Page 100, bottom half of page, "Actually, we shall only use the weaker estimate", the thing is that

$$\log^2 x > \sqrt{x} \quad \text{for } x < 5504$$
$$\log^2 x < \sqrt{x} \quad \text{for } x \geq 5504$$

so in this case $\log^2 x = O(\sqrt{x})$.

Page 101 Theorem 4.9 is

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \log x + O(1)$$

65

but first we need to know

$$\sum_{n \le x} \Lambda(n) = \sum_{n \le x} \frac{\Lambda(n)}{n} n = A(x)x - \int_2^x A(t)dt$$

$$= (\log x + O(1))x - \int_2^x \{\log t + O(1)\}dt$$

$$= x \log x + O(x) - \left\{ t \log t - t|_2^x + O\left(\int_2^x dt\right)\right\}$$

$$= \bcancel{x \log x} + O(x) - \{\bcancel{x \log x} - x - (2\log 2 - 2) + O(x)\}$$

$$= O(x)$$

where $A(x) = \sum_{n \le x} \Lambda(n)/n$ and $A(2^-) = 0$ because $\Lambda(2^-) = 0$ and therefore $A(y)y = A(2^-)(2^-) = 0$, we have used Abel's lemma of course :)

And so

$$\sum_{n \le x} \left(\frac{x}{n} - C - 1\right) \Lambda(n) = x \sum_{n \le x} \frac{\Lambda(n)}{n} - (C+1) \sum_{n \le x} \Lambda(n)$$

$$= x(\log x + O(1)) - (C+1)O(x)$$

$$= x \log x + O(x)$$

combining everything

$$(x - C - 1)\log x + \sum_{n \le x} \left(\frac{x}{n} - C - 1\right) \Lambda(n) + O(x)$$

$$= x \log x - (C+1)\log x + x \log x + O(x) + O(x)$$

$$= 2x \log x + O(x) - (C+1)\log x$$

$$= 2x \log x + O(x)$$

the thing here is that $\log x < x$ so we can lump $\log x$ into $O(x)$ me hope :)