Apostol's Analytic Number Theory

Stefanus[1]

[1] Samsung Semiconductor Inc

San Jose, CA 95134 USA

(Dated: January 2, 2017)

Abstract

Just for fun :)

Chapter 2

Problem 2.1 Find all integers n such that

(a) $\varphi(n) = n/2$ (b) $\varphi(n) = \varphi(2n)$ (c) $\varphi(n) = 12$

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

$$\begin{aligned}\frac{\cancel{n}}{2} &= \cancel{n}\prod_{p|n}\left(1-\frac{1}{p}\right)\\ \frac{1}{2} &= \frac{\prod_{p|n}(p-1)}{\prod_{p|n}p}\\ \prod_{p|n}p &= 2\prod_{p|n}(p-1)\end{aligned}$$

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)

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

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

$$\begin{aligned}\prod_{p|n}\left(1-\frac{1}{p}\right) &= 2\prod_{p|n}\left(1-\frac{1}{p}\right)\\ &\to 1 = 2\end{aligned}$$

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

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

and therefore

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

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

For (c)

$$\begin{aligned}\varphi(n) = 12 &= 2\cdot 2\cdot 3\\ &= \prod_{p|n} p^{\alpha_p} - p^{\alpha_p - 1}\\ \varphi\left(\prod_{p|n} p^{\alpha_p}\right) &= \prod_{p|n} p^{\alpha_p-1}(p-1)\end{aligned}$$

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 $\prod_{p|n} = \prod_{p|m}$

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

Problem 2.3. Prove that

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

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

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

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

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

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

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

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

$$\begin{aligned} u * f &= (u * \mu) * \nu \\ &= I * \nu \\ u * f &= \nu \\ \to \nu(n) &= \sum_{d|n} f(d) \end{aligned}$$

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, $n = \prod_i p_i^{\alpha_i}$

$$\begin{aligned} \nu\left(\prod_i p_i^{\alpha_i}\right) &= \sum_{d|n} f(d) \\ \sum_i \nu\left(p_i^{\alpha_i}\right) &= \sum_{d|n} f(d) \\ \sum_i \nu\left(p_i\right) &= \sum_{d|n} f(d) \end{aligned}$$

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.

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 $\mu(d^2)$

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

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

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

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 $\mu(d^k)$

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

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

$$\begin{aligned}\sum_{d|p^a} \mu(d)\mu(p,d) &= \mu(1)\mu(p,1) + \mu(p)\mu(p,p) \\ &= \mu(1)\mu(1) + \mu(p)\mu(p) \\ &= (1)(1) + (-1)(-1) \\ &= 2\end{aligned}$$

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

$$\begin{aligned}g(q^b) = \sum_{d|q^b} \mu(d)\mu(p,d) &= \mu(1)\mu(p,1) + \mu(q)\mu(p,q) \\ &= \mu(1)\mu(1) + \mu(q)\mu(1) \\ &= (1)(1) + (-1)(1) \\ &= 0\end{aligned}$$

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_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

$$\begin{aligned}\sum_{d|n} \mu(d) \log^m d = \cancel{\mu(1) \log^m (1)} &+ \mu(p_1) \log^m (p_1) + \ldots + \mu(p_k) \log^m (p_k) + \\ &\mu(p_1 p_2) \log^m (p_1 p_2) + \ldots + \mu(p_{k-1} p_k) \log^m (p_{k-1} p_k) + \ldots + \\ &\mu(p_1 p_2 \ldots p_k) \log^m (p_1 p_2 \ldots p_k)\end{aligned}$$

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

$$\begin{aligned}\sum_{d|n} \mu(d)\log^m d = &-(\log^m(p_1) + \ldots + \log^m(p_k))\\ &+(\log^m(p_1p_2) + \ldots + \log^m(p_{k-1}p_k)) +\\ &-(\log^m(p_1p_2p_3) + \ldots + \log^m(p_{k-2}p_{k-1}p_k)) +\\ &(-1)^k \log^m(p_1p_2\ldots p_k)\end{aligned}$$

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

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

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

$$
\begin{aligned}
(5|1) &\rightarrow (-1)l_1^4 \\
(5|2) &\rightarrow (+4)l_1^4 \\
(5|3) &\rightarrow (-6)l_1^4 \\
(5|4) &\rightarrow (+4)l_1^4 \\
(5|5) &\rightarrow (-1)l_1^4
\end{aligned}
$$

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 $4l_1^3 l_2$, 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

$$\begin{aligned}
&(5|1) \to (0) \\
&(5|2) \to (+1)4l_1^3 l_2 \\
&(5|3) \to (-3)4l_1^3 l_2 \\
&(5|4) \to (+3)4l_1^3 l_2 \\
&(5|5) \to (-1)4l_1^3 l_2
\end{aligned}$$

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, $l_1^{a_1} l_2^{a_2} \dots l_m^{a_m}$ can only be found once and there'll be nothing to cancel it, the same is true if $k < m$, the longest l combo $l_1^{a_1} l_2^{a_2} \dots l_k^{a_k}$ 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

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

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

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

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, $\sum_i a_i = m$

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

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

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

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 $p_1^0, p_1^1, \dots, p_1^{\alpha_1}$, so if we multiply all of them we have $p_1^{1+2+3+\dots+\alpha_1} = p_1^{\frac{\alpha_1(\alpha_1+1)}{2}} = (p_1^{\alpha_1})^{\frac{\alpha_1+1}{2}}$.

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 $(\alpha_2+1)(\alpha_3+1)\dots(\alpha_N+1)$ times, so the final exponent for p_1 in $\prod_{t|n} t$ is

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

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 $\to (p_1^0\ p_1^1\ p_1^2)^{\alpha_2+1}$.

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

$$\begin{aligned}
d(p^{\alpha+1})^3 + \sum_{t|p^\alpha} d(t)^3 &= (\alpha+2)^3 + \left(\sum_{t|p^\alpha} d(t)\right)^2 \\
&= (\alpha+2)^2(\alpha+2) + \left(\sum_{t|p^\alpha} d(t)\right)^2 \\
&= (\alpha+2)^2 + (\alpha+2)^2(\alpha+1) + \left(\sum_{t|p^\alpha} d(t)\right)^2 \\
&= d(p^{\alpha+1})^2 + (\alpha+2)\cdot 2\frac{(\alpha+2)(\alpha+1)}{2} + \left(\sum_{t|p^\alpha} d(t)\right)^2 \\
&= d(p^{\alpha+1})^2 + 2d(p^{\alpha+1})\left(\sum_{t|p^\alpha} d(t)\right) + \left(\sum_{t|p^\alpha} d(t)\right)^2 \\
&= \left(d(p^{\alpha+1}) + \sum_{t|p^\alpha} d(t)\right)^2 \\
\sum_{t|p^{\alpha+1}} d(t)^3 &= \left(\sum_{t|p^{\alpha+1}} d(t)\right)^2
\end{aligned}$$

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 expanding $\sum_{t|p^\alpha+1} d(t)^3 = \sum_{i=1}^{\alpha+1} i^3$ but this requires us to know the formula for a sum of consecutive cubes ¯\(º_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, $\sum_{y<n\le x}$, 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

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

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

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

so .... ¯\(º_o)/¯

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

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

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

PROOF of Theorem 3.2, Page 55 The peculiar thing here is the contant 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 $y \to 1$, 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 $f(y) = x^{-s}$

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

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 $0 < w < \infty$

$$\sum_{n \le x} \frac{1}{n^s} = \int_{1/w}^x \frac{dt}{t^s} - s\int_{1/w}^x \frac{t - [t]}{t^{s+1}} + \frac{[x] - x}{x^s} - \frac{0 - (1/w)}{(1/w)^s}$$
$$= \frac{x^{1-s}}{1-s} - \frac{(1/w)^{1-s}}{1-s} - s\int_{1/w}^1 \frac{t - [t]}{t^{s+1}} - s\int_1^x \frac{t - [t]}{t^{s+1}} - \frac{x - [x]}{x^s} + (1/w)^{1-s}$$
$$= \frac{x^{1-s}}{1-s} - \frac{(1/w)^{1-s}}{1-s} - s\int_{1/w}^1 \frac{t - [t]}{t^{s+1}} - s\int_1^x \frac{t - [t]}{t^{s+1}} - \frac{x - [x]}{x^s} + \frac{(1-s)(1/w)^{1-s}}{1-s}$$

we now focus on the first integral on the RHS

$$\begin{aligned}
-s\int_{1/w}^{1}\frac{t-[t]}{t^{s+1}} &= -s\int_{1/w}^{1}\frac{t-[1/w]}{t^{s+1}}\\
&= -s\int_{1/w}^{1}\frac{t}{t^{s+1}}\\
&= -s\int_{1/w}^{1}\frac{1}{t^{s}}\\
&= -\frac{s}{1-s}+\frac{s(1/w)^{1-s}}{1-s}
\end{aligned}$$

Combining everything we get

$$\begin{aligned}
\sum_{n\le x}\frac{1}{n^s} &= \frac{x^{1-s}}{1-s}-\cancel{\frac{(1/w)^{1-s}}{1-s}}-\frac{s}{1-s}+\cancel{\frac{s(1/w)^{1-s}}{1-s}}-s\int_{1}^{x}\frac{t-[t]}{t^{s+1}}-\frac{x-[x]}{x^s}+\frac{(\cancel{1}-\cancel{s})(1/w)^{1-s}}{1-s}\\
&= \frac{x^{1-s}}{1-s}+\frac{-1+(1-s)}{1-s}-s\int_{1}^{x}\frac{t-[t]}{t^{s+1}}-\frac{x-[x]}{x^s}\\
&= \frac{x^{1-s}}{1-s}-\frac{1}{1-s}+1-s\int_{1}^{x}\frac{t-[t]}{t^{s+1}}-\frac{x-[x]}{x^s}
\end{aligned}$$

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 60, "It can be shown that $\xi(2)=\pi^2/6$", an elementary proof of this fact is thorugh 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

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

and

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

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

$$\begin{aligned}
-2\int_{-\pi}^{+\pi} x\sin(nx)\frac{dx}{n\pi} &= \frac{2x}{n^2\pi}\cos(nx)\Big|_{-\pi}^{+\pi} - 2\int_{-\pi}^{+\pi}\cos(nx)\frac{dx}{n^2\pi} \\
&= \frac{2x}{n^2\pi}\cos(nx)\Big|_{-\pi}^{+\pi} + \frac{-2}{n^3\pi}\sin(nx)\Big|_{-\pi}^{+\pi} \\
a_n &= (-1)^n\frac{4}{n^2}
\end{aligned}$$

where $\cos(\pm n\pi) = \cos(n\pi) = (-1)^n$, thus

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

setting $x = \pi$ we get

$$\begin{aligned}
\pi^2 &= \frac{\pi^2}{3} + \sum_{n=1}^{\infty}(-1)^n\frac{4}{n^2}\cos(n\pi) \\
\frac{2\pi^2}{3} &= \sum_{n=1}^{\infty}(-1)^{2n}\frac{4}{n^2} \\
\rightarrow \frac{\pi^2}{6} &= \sum_{n=1}^{\infty}\frac{1}{n^2}
\end{aligned}$$