# MODEL ANSWERS TO THE THIRD HOMEWORK

6.4.3. Note that

$$\frac{\mathrm{d}}{\mathrm{d}x}\eta(x) = 1,$$

so that integrating by parts, we have

$$\begin{aligned}
\int_k^{k+1/2} f(x)\,\mathrm{d}x &= \Big[\eta(x)f(x)\Big]_k^{k+1/2} - \int_k^{k+1/2} \eta(x)f'(x)\,\mathrm{d}x \\
&= \eta(k+1/2)f(k+1/2) - \eta(k)f(k) - \int_k^{k+1/2} \eta(x)f'(x)\,\mathrm{d}x \\
&= f(k)/2 - \int_k^{k+1/2} \eta(x)f'(x)\,\mathrm{d}x
\end{aligned}$$

and

$$\begin{aligned}
\int_{k+1/2}^{k+1} f(x)\,\mathrm{d}x &= \Big[\eta(x)f(x)\Big]_{k+1/2}^{k+1} - \int_{k+1/2}^{k+1} \eta(x)f'(x)\,\mathrm{d}x \\
&= \eta(k+1)f(k+1) - \eta(k+1/2)f(k) - \int_{k+1/2}^{k+1} \eta(x)f'(x)\,\mathrm{d}x \\
&= f(k+1)/2 - \int_{k+1/2}^{k+1} \eta(x)f'(x)\,\mathrm{d}x
\end{aligned}$$

Therefore

$$\begin{aligned}
\int_m^n f(x)\,\mathrm{d}x &= \sum_{k=m}^{n-1} \int_k^{k+1/2} f(x)\,\mathrm{d}x + \int_{k+1/2}^{k+1} f(x)\,\mathrm{d}x \\
&= \sum_{k=m}^{n-1} f(k)/2 - \int_k^{k+1/2} \eta(x)f'(x)\,\mathrm{d}x + f(k+1)/2 - \int_{k+1/2}^{k+1} \eta(x)f'(x)\,\mathrm{d}x \\
&= \frac{1}{2}\sum_{k=m}^{n-1} f(k+1) + f(k) - \int_k^{k+1} \eta(x)f'(x)\,\mathrm{d}x \\
&= \sum_{k=m}^{n} f(x) - f(m)/2 - f(n)/2 - \int_m^n \eta(x)f'(x)\,\mathrm{d}x.
\end{aligned}$$

Thus

$$\sum_{k=m}^{n} f(x) = \int_m^n f(x)\,\mathrm{d}x + f(m)/2 + f(n)/2 + \int_m^n \eta(x) f'(x)\,\mathrm{d}x.$$

(a) Note that

$$\begin{aligned}\left|\int_n^\infty \frac{\eta(x)}{x^2}\,\mathrm{d}x\right| &\le \int_n^\infty \frac{\mathrm{d}x}{x^2}\\ &= \left[-\frac{1}{x}\right]_n^\infty\\ &= \frac{1}{n}.\end{aligned}$$

Therefore, if we put

$$f(x) = \frac{1}{x},$$

and $m = 1$ then we get

$$\begin{aligned}\sum_{k=1}^{n} \frac{1}{k} &= \int_1^n \frac{\mathrm{d}x}{x} + \frac{1}{2} + \frac{1}{2n} + \int_1^n \frac{-\eta(x)}{x^2}\,\mathrm{d}x\\ &= \log n + \gamma + O\left(\frac{1}{n}\right).\end{aligned}$$

(b) Note that

$$\begin{aligned}\left|\int_n^\infty \frac{\eta(x)}{x^{\alpha-1}}\,\mathrm{d}x\right| &\le \int_n^\infty \frac{\mathrm{d}x}{x^{\alpha-1}}\\ &= \left[-\frac{x^{-\alpha}}{\alpha}\right]_n^\infty\\ &= O(n^{-\alpha}).\end{aligned}$$

Therefore, if we put

$$f(x) = x^{-\alpha}$$

and $m = 1$ then we get

$$\begin{aligned}\sum_{k=1}^{n} k^{-\alpha} &= \int_1^n \frac{\mathrm{d}x}{x^\alpha} + 1^{-\alpha}/2 + n^{-\alpha}/2 + \int_1^n \frac{-\eta(x)}{x^{\alpha-1}}\,\mathrm{d}x\\ &= \frac{n^{1-\alpha}}{1-\alpha} + c_\alpha + O(n^{-\alpha}).\end{aligned}$$

(c) Note that

$$\begin{aligned}\left|\int_1^n \frac{\eta(x)}{x}\,\mathrm{d}x\right| &\le \int_1^n \frac{\mathrm{d}x}{x}\\ &= \Big[\log x\Big]_1^n\\ &= \log n.\end{aligned}$$

Therefore, if we put

$$f(x) = \log x.$$

and $m = 1$ then we get

$$\begin{aligned}\log n! &= \sum_{k=1}^n \log k\\ &= \int_1^n \log x\,\mathrm{d}x + 1/2\log 1 + 1/2\log n + \int_1^n \frac{\eta(x)}{x}\,\mathrm{d}x\\ &= n\log n - n + 1 + 1/2\log n + O(\log n)\\ &= (n+1/2)\log n - n + O(\log n).\end{aligned}$$

6.4.4. We have

$$\begin{aligned}\mathrm{li}(x) &= \int_2^x \frac{\mathrm{d}t}{\log t}\\ &= \left[\frac{t}{\log t}\right]_2^x + \int_2^x \frac{\mathrm{d}t}{\log^2 t}\\ &\le \frac{x}{\log x} - \frac{2}{\log 2} + \int_2^{\sqrt{x}} \frac{\mathrm{d}t}{\log^2 t} - \int_{\sqrt{x}}^x \frac{\mathrm{d}t}{\log^2 t}\\ &\le \frac{x}{\log x} + \sqrt{x}\frac{1}{\log 2} + x\frac{1}{\log^2\sqrt{x}}\\ &= \frac{x}{\log x} + \sqrt{x}\frac{1}{\log 2} + 4x\frac{1}{\log^2 x}\\ &= \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right).\end{aligned}$$

6.4.5. Imagine approximating the area under the graph of $y = f(x)$ between $x = 1$ and $x = n$. If we draw $n-1$ rectangles of heights $f(1)$, $f(2)$, ..., $f(n-1)$ over the $n-1$ intervals $[1,2]$, $[2,3]$, ..., $[n-1,n]$, then their area is bigger than the area under the graph, as $f(x)$ is decreasing. But if we draw rectangles draw $n-1$ rectangles of heights $f(2)$, $f(3)$, ..., $f(n)$ over the same intervals then their area is smaller than the area under the graph.

It follows that

$$\sum_{k=2}^{n} f(x) \leq \int_1^n f(t)\,\mathrm{d}t \leq \sum_{k=1}^{n-1} f(x).$$

Since the difference between the third and first term is $f(1) - f(n)$, the difference between

$$\sum_{k=1}^{n} f(x) \qquad \text{and} \qquad \int_1^n f(t)\,\mathrm{d}t$$

is bounded. As the second term tends to infinity, it follows that

$$\sum_{k=1}^{n} f(x) \sim \int_1^n f(t)\,\mathrm{d}t.$$

Now suppose that $f(x) = o(g(x))$. Then

$$\begin{aligned} \left| \int_1^n f(t)\,\mathrm{d}t \right| &\leq \int_1^n \epsilon g(t)\,\mathrm{d}t \\ &= \epsilon \int_1^n g(t)\,\mathrm{d}t, \end{aligned}$$

for all $\epsilon > 0$, so that

$$\int_1^n f(t)\,\mathrm{d}t = o\left( \int_1^n g(t)\,\mathrm{d}t \right).$$

We have

$$\begin{aligned} \lim_{n\to\infty} \frac{\sum_{k=1}^n f(x)}{\sum_{k=1}^n g(x)} &= \lim_{n\to\infty} \frac{\sum_{k=1}^n f(x)}{\int_1^n f(t)\,\mathrm{d}t} \cdot \frac{\int_1^n f(t)\,\mathrm{d}t}{\int_1^n g(t)\,\mathrm{d}t} \cdot \frac{\int_1^n g(t)\,\mathrm{d}t}{\sum_{k=1}^n g(x)} \\ &= \lim_{n\to\infty} \frac{\sum_{k=1}^n f(x)}{\int_1^n f(t)\,\mathrm{d}t} \cdot \lim_{n\to\infty} \frac{\int_1^n f(t)\,\mathrm{d}t}{\int_1^n g(t)\,\mathrm{d}t} \cdot \lim_{n\to\infty} \frac{\int_1^n g(t)\,\mathrm{d}t}{\sum_{k=1}^n g(x)} \\ &= 1 \cdot 0 \cdot 1 \\ &= 0. \end{aligned}$$

Thus

$$\sum_{k=1}^{n} f(x) = o\left( \sum_{k=1}^{n} g(x) \right).$$

6.4.6. Note that if $g(x) = (\log\log x)^2$ then

$$a \log\log x = o(g(x)) \qquad \text{and} \qquad g(x) = o(\delta \log x).$$

If we put $f(x) = e^{g(x)}$ and we exponentiate these inequalities we get

$$\log^a x = o(f(x)) \qquad \text{and} \qquad f(x) = o(x^\delta).$$

On the other hand, $f(x)$ is increasing as $g(x)$ is increasing, so that $f'(x) > 0$.

6.4.9. The idea is to splice together copies of $\log_n x$ over intervals. To get continuity at the endpoints of the intervals we need to rescale (or shift). But if we do that we need to make the intervals longer and longer to make sure the limit of $f(x)$ as $x$ increases is infinity.

Define sequences of positive real numbers $a_1, a_2, \dots$ and $x_1, x_2, \dots$ recursively by the rule:

$$\begin{aligned} x_1 &= 1 & a_{n-1}\log_{n-1} x_n &= n \\ a_1 &= 1 & a_n \log_n x_n &= n. \end{aligned}$$

In the second equation we suppose that $a_{n-1}$ is known and we solve for $x_n$. Note that $\log_n x$ is a composition of invertible functions, so that it is invertible. In the fourth equation we suppose that $x_n$ is known and we solve for $a_n$ (this is easy).

Let $f(x)$ be the function, defined for $x \geq 1$, by the rule

$$f(x) = a_n \log_n x \qquad \text{when} \qquad x \in [x_{n-1}, x_n).$$

The sequences $a_1, a_2, \dots$ and $x_1, x_2, \dots$ are chosen so that $f(x_n) = n$ and $f(x)$ is continuous.

Indeed as $x_n \in [x_n, x_{n+1})$ we have

$$\begin{aligned} f(x_n) &= a_{n+1} \log_{n+1} x_n \\ &= n. \end{aligned}$$

On the other hand, if we check to see that happens as $x$ approaches $x_n$ from below, we see that

$$\begin{aligned} \lim_{x \to x_n^-} f(x) &= a_{n-1} \log_{n-1} x \\ &= a_{n-1} \log_{n-1} x_n \\ &= n, \end{aligned}$$

as $x \in [x_{n-1}, x_n)$ if $x$ is approaching $x_n$ from below.

Note that $f(x_n) = n$ and $f(x)$ is equal to $\log_n x$ until $f(x) = n + 1$. Thus $x_1, x_2, \dots$ is unbounded.

Note that $f(x)$ is monotonic increasing, since $\log_n x$ is increasing, for all $n$. As $f(x_n) = n$ and $\lim_{n\to\infty} x_n = \infty$ it follows that $\lim_{x\to\infty} f(x) = \infty$. As $a_n \log_n x_n = a_{n-1} \log_{n-1} x_n$ and $\log_n$ grows slower than $\log_{n-1}$, it follows that $f(x) \leq a_n \log_n x$ for all $x \geq x_n$. Thus $f(x) = o(\log_n(x))$.

6.5.1. We follow the argument of the sieve of Eratosthenes given in §5. Note that if $m$ is not square-free then $m$ must be divisible by the square of a prime. Let $\psi(x)$ be the number of square-free integers up $x$.

Let $B(x,r)$ be the number of integers up to $x$ not divisible by the square of one of the first $r$ primes. We want to estimate $B(x,r)$. Let $M_i$ be the set of integers from 1 to $n$ which are multiples of $p_i^2$. Let $M_{ij}$ be the set of integers from 1 to $n$ which are multiples of both $p_i^2$ and $p_j^2$. As $p_i$ and $p_j$ are coprime,

$$M_{ij} = M_i \cap M_j.$$

Note that

$$|M_i| = \llcorner \frac{x}{p_i^2} \lrcorner \qquad \text{and} \qquad |M_{ij}| = \llcorner \frac{x}{p_i^2 p_j^2} \lrcorner,$$

and so on. It follows by inclusion-exclusion that

$$B(x,r) = \llcorner x \lrcorner - \sum_{i=1}^{r} \llcorner \frac{x}{p_i^2} \lrcorner + \sum_{i \neq j \leq r} \llcorner \frac{x}{p_i^2 p_j^2} \lrcorner + \cdots + (-1)^r \llcorner \frac{x}{p_1^2 p_2^2 \ldots p_r^2} \lrcorner.$$

Suppose that we approximate the RHS by simply ignoring all of the round downs,

$$x - \sum_{i=1}^{r} \frac{x}{p_i^2} + \sum_{i \neq j \leq r} \frac{x}{p_i^2 p_j^2} + \cdots + (-1)^r \frac{x}{p_1^2 p_2^2 \ldots p_r^2} = x \prod_{i=1}^{r} \left(1 - \frac{1}{p_i^2}\right).$$

The worse case scenario for the error is

$$1 + \binom{r}{1} + \binom{r}{2} + \cdots + \binom{r}{r} = 2^r.$$

Thus

$$\begin{aligned}
\psi(x) &\leq r + x \prod_{i=1}^{r} \left(1 - \frac{1}{p_i^2}\right) + 2^r \\
&\leq x \prod_{i=1}^{r} \left(1 - \frac{1}{p_i^2}\right) + 2^{r+1} \\
&\leq x \prod_{i=1}^{r} \left(1 - \frac{1}{p_i^2}\right) + O(\frac{x}{\log \log x}),
\end{aligned}$$

where we take $r = \log x$.
Now

$$\prod_{p>y} \left(1 - \frac{1}{p^2}\right)^{-1} = \prod_{p>y} \left(1 + \frac{1}{p^2} + \frac{1}{p^4} + \ldots\right),$$

If we expand the product on the RHS we get

$$1 + \sum_{p|k \implies p>y} \frac{1}{k^2} \geq \frac{y+1}{y},$$

where we used the fact that $\sum k^{-2}$, as a Riemman sum for $1/x^2$, gives a lower bound for the area under the graph from $y$ to $\infty$. Thus

$$\begin{aligned}
\prod_{i=1}^{r}\left(1-\frac{1}{p_i^2}\right) &= \prod_{p}\left(1-\frac{1}{p^2}\right)\cdot\prod_{p>p_r}\left(1-\frac{1}{p^2}\right)^{-1} \\
&> \frac{p_r+1}{p_r}\prod_{p}\left(1-\frac{1}{p^2}\right) \\
&= \prod_{p}\left(1-\frac{1}{p^2}\right)+\frac{1}{p_r}\prod_{p}\left(1-\frac{1}{p^2}\right) \\
&= \prod_{p}\left(1-\frac{1}{p^2}\right)+\frac{1}{\log x}\prod_{p}\left(1-\frac{1}{p^2}\right) \\
&= \prod_{p}\left(1-\frac{1}{p^2}\right)+o(x).
\end{aligned}$$