6/5/2016

# Closing Comments

[Assorted remarks, revisions, augmentations for Lecs 1–30]

## Part I (Lecs. 1–16)

1. (Lec 1, p. 1) "The primary reference for these lectures will be Ingham, Distr. of Prime Numbers, from 1932." This sentence somehow failed to make its way into the record of Lec 1. ☹

2. (Lec 1, p. 18) The fact that $\psi(x) = \theta(x) + O(x^{\frac{1}{2}})$ is important and probably should have been highlighted via some kind of box.

3. (Lec 3, pp. 7+8) Apostol's book, Mathematical Analysis, 1st edition, 1957, can be cited as a good reference for Riemann–Stieltjes integrals.

On page 8 (bottom half), it is important to note that the stated integration-by-parts formula holds equally well when $f$ is merely continuous + piecewise $C^1$ on $[a, b]$. Compare Lec 11, ① bot–② middle.

4. (Lec 5, p. 9) In regard to formula #1 on this page, I should have paused to note that letting $z \to 1$ immediately gives

$$\sum_1^N n^{-1} = 1 + \ln N - \int_1^N \frac{r(t)}{t^2}\,dt$$

$$\Rightarrow \text{[Euler constant]} \quad \gamma = 1 - \int_1^\infty \frac{r(t)}{t^2}\,dt$$

$$\Rightarrow \sum_1^N n^{-1} - \gamma = \ln N + \int_N^\infty \frac{r(t)}{t^2}\,dt \,,$$

hence

$$\boxed{\begin{aligned} \sum_{n=1}^N \frac{1}{n} &= \ln N + \gamma + \int_N^\infty \frac{r(t)}{t^2}\,dt \,; \\ \sum_{n=1}^N \frac{1}{n} &= \ln N + \gamma + O\left(\frac{1}{N}\right) \end{aligned}} \; .$$

In formula #2 on p. 9, application of the above equation for $\gamma$ quickly leads to

$$\zeta(z) - \gamma = \frac{1}{z-1} - \left\{ z \int_1^\infty \frac{r(t)}{t^{z+1}}\,dt - \int_1^\infty \frac{r(t)}{t^2}\,dt \right\} .$$

The brace is simply $H(z) - H(1)$ with

$$H(z) \equiv z \int_1^\infty \frac{r(t)}{t^{z+1}}\,dt \,.$$

The function $H(z)$ is analytic on $\{x > \delta > 0\}$ as already noted. Accordingly, on page 10 (top), after introducing $F$, one can simultaneously assert that

$$\boxed{\zeta(z) = \frac{1}{z-1} + \gamma + O(z-1) \quad \text{near } z = 1} \ .$$

Compare Lec 18 (40) – (41).

5. (Lec 6, p. 10 top) This subtraction trick is the "$z = z$" counterpart of what was just obtained for $\sum_1^N n^{-1}$ in item #4. Its <u>importance</u> can thus be said to have been recognized very early on.

6. (Lec 8, p. 14) Taking $f = \frac{1}{1+t}$ and $N \hookrightarrow N-1$ in E-M version I leads to:

$$\left\{ \begin{array}{l} \displaystyle\sum_{n=1}^{N} \frac{1}{n} = \frac{1}{2} + \frac{1}{2N} + \ln N - \int_1^N \frac{\beta(t)}{t^2}\, dt \\ \beta(t) \equiv t - [\![ t ]\!] - \frac{1}{2} \end{array} \right\} .$$

This agrees with item #4 above.

7. (Lec 9, p. 21) Regarding EULER and the special values $\zeta(-2k) = 0$, $\zeta(2k)$, $\zeta(0)$, $\zeta(1-2k)$ for $k \geqq 1$, it is very illuminating to actually have a look in Euler's collected works. (Cf., e.g.,

Leonhardi Euleri, Opera Omnia, Series I,

vol 14 pp. 73 – 86 (1734);
114 (§19) (1736);
424 – 434, 440 – 443 (1740);
477 – 479 (1750);

vol 15 pp. 72 – 78 (1749).

One keeps in mind here the (modified) function $\phi(s) \equiv \sum_1^\infty (-1)^{n-1} n^{-s} = (1-2^{1-s})\zeta(s)$. The paper by R. Ayoub, "Euler and the Riemann Zeta Function", Amer. Math. Monthly 81 (1974) 1067 – 1086 is also very worthwhile, as is A. Weil's, "Prehistory of the Zeta-Function", in Number Theory, Trace Formulas, and Discrete Groups (ed. by K. Aubert, et al.), Acad. Press, 1989, pp. 1 – 9.

8. (Lec 11, p. 24) Concerning the functional equation $\xi(s) = \xi(1-s)$ and the alternate version

$$\zeta(1-s) = 2(2\pi)^{-s} \cos\left(\frac{\pi s}{2}\right) \Gamma(s) \zeta(s)$$

noted in Lec 16, p. 7 (top), a look in Euler is again very revealing. (f., e.g.,

vol. 14 p. 443 (1740);

vol. 15 pp. 79 ~ 90 (1749);
131 ~ 138 (1772).

Also of interest:

vol. 16 (part 2, preface) pp. XXVIII, XXXII,
LXXXII ~ LXXXV,

and the aforementioned works by R. Ayoub and A. Weil.
The extent to which Euler more-or-less "stumbled on" the
functional equation already around 1740 - 1749 is
striking indeed!! Remember Riemann is ≈ 1859.

9. (Lec 11, p. 19) It is worthwhile to show how a bare bones
form of Poisson summation follows nearly immediately from
Euler - Maclaurin version I (lec 8, p. 14). *

Given any $\varphi \in C^1(\mathbb{R})$ such that $\varphi \in L_1(\mathbb{R})$, $\varphi' \in L_1(\mathbb{R})$.
We then have:

(a) $\varphi \to 0$ as $x \to \pm\infty$;

(b) $\sum_{-\infty}^{\infty} |\varphi(x+n)|$ conv uniformly on $\mathbb{R}$-compacta;

(c) $\sum_{-\infty}^{\infty} \varphi(x+n) = \lim_{N\to\infty} \sum_{-N}^{N} \hat{\varphi}(k)\, e^{2\pi i k x}$, each $x \in \mathbb{R}$.

* The TRICK will be very similar to pp. 3 (bottom) - 4 (line 8) of Lec 11.

Since $\varphi \in L_1(\mathbb{R})$, clearly $\liminf_{x \to \infty} |\varphi(x)| = 0$. Since $\varphi' \in L_1(\mathbb{R})$, $\varphi(y)$ is uniformly Cauchy as $y \to \infty$. Hence $\lim_{x \to \infty} \varphi(x) = 0$. The case $x \to -\infty$ is similar. $\Rightarrow$ (a) OK

Notice that $\int_0^1 \left( \sum_{-\infty}^{\infty} |\varphi(x+n)| \right) dx < \infty$. Hence $\sum_{-\infty}^{\infty} \varphi(x+n)$ conv absolutely almost everywhere on $[0,1]$. E.g., at point $x_0$. Consider any $x_1 \in [0,1]$ with, say, $x_1 > x_0$. (The case $x_1 < x_0$ will be similar.) For $N \geqq M$ large, observe that:

$$\sum_{M}^{N} |\varphi(x_1+n)| = \sum_{M}^{N} \left| \varphi(x_0+n) + \int_{x_0}^{x_1} \varphi'(v+n)\, dv \right|$$

$$\leqq \sum_{M}^{N} |\varphi(x_0+n)| + \sum_{M}^{N} \int_{x_0+n}^{x_1+n} |\varphi'(w)|\, dw$$

$$\left\{ \frac{\text{but the intervals } [n, n+1] \text{ are non-overlapping}}{\text{and } x_j + n \in [n, n+1]} \right\}$$

$$\leqq \sum_{M}^{N} |\varphi(x_0+n)| + \int_{M}^{\infty} |\varphi'(w)|\, dw .$$

Negative $N$ and $M$ are treated similarly; the "new" $w$-integral will be

$$\int_{-\infty}^{N+1} |\varphi'(w)|\, dw .$$

This proves (b) on $K = [0,1]$. By virtue of (a), we then get (b) on a general $K$.

To prove (c), a standard translation shows that $x = 0$ wlog. Let $M$ be big. We have: $\boxed{\text{by E-M}}$

$$\frac{1}{2}\varphi(-M) + \sum_{|n|<M} \varphi(n) + \frac{1}{2}\varphi(M) = \int_{-M}^{M} \varphi(x)\,dx$$
$$+ \int_{-M}^{M} \varphi'(x)\beta(x)\,dx$$

$$\left\{ \beta(x) = x - [\![x]\!] - \frac{1}{2} \right\}.$$

Let $M \to \infty$ remembering that $|\beta(x)| \le \frac{1}{2}$. Get:

$$\sum_{-\infty}^{\infty} \varphi(n) = \hat{\varphi}(0) + \int_{-\infty}^{\infty} \varphi'(x)\beta(x)\,dx.$$

Let $s_N(x)$ be the usual partial sum for $\beta(x)$; recall the bounded convergence properties of $s_N$. Hence,

$$\lim_{N\to\infty} \int_{-\infty}^{\infty} |\varphi'(x)|\,|\beta(x) - s_N(x)|\,dx = 0$$

$$\int_{-\infty}^{\infty} \varphi'(x)\beta(x)\,dx = \lim_{N\to\infty} \sum_{k=1}^{N} \int_{-\infty}^{\infty} \varphi'(x)\frac{\sin(2\pi k x)}{-\pi k}\,dx.$$

When $k \ge 1$, we immediately check via (a) and integ-by-parts that

$$\int_{-\infty}^{\infty} \frac{\sin(2\pi k x)}{-\pi k}\, d\varphi(x) = \frac{1}{\pi k}\int_{-\infty}^{\infty} \varphi(x)(2\pi k)\cos(2\pi k x)\, dx$$

$$= \int_{-\infty}^{\infty} \varphi(x)\left[e^{2\pi i k x} + e^{-2\pi i k x}\right] dx$$

$$= \hat{\varphi}(k) + \hat{\varphi}(-k) .$$

Hence:

$$\sum_{-\infty}^{\infty} \varphi(n) = \hat{\varphi}(0) + \lim_{N\to\infty} \Big\{ \sum_{1 \le |\ell| \le N} \hat{\varphi}(\ell) \Big\}$$

$$= \lim_{N\to\infty} \sum_{-N}^{N} \hat{\varphi}(\ell) . \qquad \blacksquare$$

10. (Lec 16, p.7) Once the 2nd box was obtained on p.7, had I not been in a rush, it would have been useful to stop for a moment and obtain

$$\boxed{\frac{\zeta'(0)}{\zeta(0)} = \ln(2\pi)}$$

by letting $s \to 0$. See Lec 18 (41) (42) and item #4 above.

Notice incidentally that

$$\operatorname*{Res}_{s=0} \left[ \frac{x^{s+1}}{s(s+1)} \left( -\frac{\zeta'(s)}{\zeta(s)} \right) \right] = \left( -\frac{\zeta'(0)}{\zeta(0)} \right) x .$$

for p.1 in Lec 16

## Part II (Lecs. 17–30)

**11.** (Lec 17+18, pp. 28–35) A slight improvement can be made on pp. 28(top) – 30(bottom) with very little effort. Recall that $x \geq 1+\delta_0$, $1 < c \leq 2$, $T \geq 3$ as on pp. 10+11(top). Consider those $\rho$ for which

$$T-1 \leq \gamma < T.$$

On p. 28 (bot), it is preferable to replace $\mathcal{C}$ by a new path $\tilde{\mathcal{C}}_\rho$ which changes with $\rho$.

$$\boxed{\rho = \beta + i\gamma}$$


$$\tilde{\mathcal{C}}_\rho = \mathcal{C}_1 \cup \mathcal{C}_2$$

semicircular part ($\mathcal{C}_1$)

linear part ($\mathcal{C}_2$)

Keep $0 < \varepsilon \leq \frac{c-1}{4} \leq \frac{1}{4}$. Notice that

$$\beta + \varepsilon \leq 1 + \varepsilon < 1 + \frac{c-1}{2} = \frac{c+1}{2} < c.$$

On $\mathcal{C}_1$, clearly $|s-\rho| \geqq |s-(\beta+iT)| = \varepsilon$ by elementary geometry. On $\mathcal{C}_2$, clearly $|s-\rho| \geqq |\sigma-\beta|$.

Observe that:

$$\left| \frac{1}{2\pi i} \int_{\mathcal{C}_1} \frac{x^s}{s(s-\rho)} ds \right| \leqq \frac{1}{2\pi} \int_{\mathcal{C}_1} \frac{x^\sigma}{|s||s-\rho|} |ds|$$

$$\leqq \frac{1}{2\pi} \int_{\mathcal{C}_1} \frac{x^\sigma}{(T-1)\,\varepsilon} |ds|$$

$$\leqq \frac{1}{2} \frac{x^{\beta+\varepsilon}}{(T-1)} = O\left(\frac{1}{T}\right) x^c .$$

At the same time,

$$\left| \frac{1}{2\pi i} \int_{\mathcal{C}_2} \frac{x^s}{s(s-\rho)} ds \right| \leqq \frac{1}{2\pi} \int_{\mathcal{C}_2} \frac{x^\sigma}{(T-1)|\sigma-\beta|} d\sigma$$

$$\leqq \frac{1}{2\pi\varepsilon(T-1)} \int_{\mathcal{C}_2} x^\sigma d\sigma$$

$$\leqq \frac{1}{2\pi\varepsilon(T-1)} \int_{-\infty}^{c} x^\sigma d\sigma \qquad (x>1)$$

$$= \frac{1}{2\pi\varepsilon(T-1)} \frac{x^c}{\ln x}$$

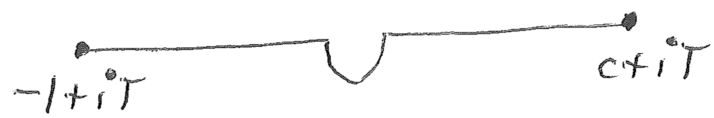
$$= O\left(\frac{1}{T}\right) \frac{x^c}{\varepsilon \ln x} .$$

Accordingly,

$$\frac{1}{2\pi i}\int_{\tilde{\mathcal{C}}_\rho} \frac{x^s}{s(s-\rho)}\,ds = O\left(\frac{1}{T}\right) x^c \left[1 + \frac{1}{\varepsilon \ln x}\right].$$

The case $T < \gamma \leq T+1$ is addressed similarly via

$\tilde{\mathcal{C}}_\rho$




.

As such, p. 30 line 4 now becomes

$$O\left(\frac{\ln T}{T}\right) x^c \left[1 + \frac{1}{\varepsilon \ln x}\right]$$

(**provided** $\varepsilon$ is kept independent of $\rho$);

line 7 becomes

$$O\left(\frac{\ln T}{T}\right) x^c \left[1 + 1 + \frac{1}{\varepsilon \ln x}\right];$$

and, lastly, 30 (bottom) becomes

$$O\left(\frac{\ln T}{T}\right) x^c \left[1 + \frac{1}{\varepsilon \ln x}\right].$$

On pp. 31–32, we can now make the <u>replacement</u>

$$O\left(\frac{\ln T}{T}\right)\frac{x^c}{c-1} \hookrightarrow O\left(\frac{\ln T}{T}\right)x^c\left[1+\frac{1}{\varepsilon \ln x}\right].$$

Note that the term $O\left(\frac{x^c \ln x}{T(c-1)}\right)$ is still present: minimization of this term leads to

$$c \approx 1 + \frac{b}{\ln x} \qquad (b = \text{tiny constant}).$$

As far as $\varepsilon$ goes, it is (clearly) natural to put

$$\varepsilon \approx \frac{c-1}{4} = \frac{b}{4 \ln x}.$$

The <u>error</u> term on p. 33 (bottom) thus becomes

$$O\left(\frac{x \ln T}{T}\right) + O\left(\frac{x \ln^2 x}{T}\right) + O(\ln x)\min\left\{1, \frac{x}{T\langle x\rangle}\right\}.$$

The implied constants will depend solely on $\delta_0$.

On p. 35, <u>in</u> the statement of the explicit formula, we finally reach:

$$\boxed{\begin{aligned}\psi^*(x) &= x - \sum_{|\gamma|\le T} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\ln(1-x^{-2}) \\ &+ O\left(\frac{x\ln^2 x}{T}\right) + O\left(\frac{x \ln T}{T}\right) \\ &+ O(\ln x)\min\left\{1, \frac{x}{T\langle x\rangle}\right\}.\end{aligned}}$$

This is the improved version that was tacitly referred to in the footnote on p. 37.

12. (Lec 19+20, p. 25) Concerning Prop 3 ($\sum \frac{\mu(n)}{n} = 0$) and the Euler product developments for $\zeta(s)$ and $1/\zeta(s)$, see:

(Lec 6, 3)

L. Euler, Opera Omnia, Ser. I

$\frac{\text{products}}{\text{sum} = 0}$ vol. 8 pp. 286 (§269), 288 (§274), 300 (1748) / vol. 8 291, 307 for $\mu(n)$, $\lambda(n)$, resp.;

$\frac{\text{products}}{\text{sum} = 0}$ vol. 14 pp. 230–231 and, à la logs, 243–244 (1737) / vol. 14 241–242 for $\lambda(n)$ with methodology; compare, however, 227–229 (1737).

{vol. 8 = Introductio in analysin infinitorum}

13. (Lec 19 + 20, p. 28) Concerning item $\boxed{D}$ : during the lecture, I misstated the relation to $\sum \frac{\mu(n)\ln n}{n} = -1$ (i.e. p. 27 box). I wanted the implication for this aspect to "go" only one way.

This point is now correct in the revised pdf for Lec 19 + 20.

Incidentally: note that if $\sum \frac{\mu(n)\ln n}{n}$ converges, its value must be $-1$ thanks to p. 27 (bot) and Lec 21, p. 11 (Fact 2b). Similarly for $\sum \frac{\mu(n)}{n}$ and $0$.

14. (Lec 24, p. 18) The current fraction [after much technical effort] is $\frac{13}{84} = .15476^{+}$. Note that $\frac{1}{7} = .142857^{+}$.

In 2005, the fraction was $\frac{32}{205} = .15609^{+}$.

15. (Lec 24, pp. 16–18) As the [upper] bound for $\mu(\sigma)$ gradually improves, it is only natural to wonder what can be obtained via Perron's formula (Lec 19, p. 4) in a variety of problems utilizing just a crude absolute value technique over a rectangle – akin to what we did in Lec 19, p. 18 ff with $M(x)$.

My <u>original</u> thought was to give another homework problem or two touching on this matter ; alas, time (<u>and</u> endurance?) constraints intervened.

In the for what it's worth "department", I'll now scratch the surface on this topic by <u>sketching</u> what happens for

$$\sum_{n \leq x} d(n) \quad .$$

Here, of course, we have $f(s) = \zeta(s)^2 = \sum_1^\infty \frac{d(n)}{n^s}$. Nothing is lost by taking $x$ à la Lec 19, pp. 14–15. Nothing is lost by taking $x$ to have form $N + \frac{1}{2}$. For Lec 19, p. 4, we want:

$$\left\{ \begin{array}{l} a_n = d(n), \ \alpha = 2, \ \Phi(v) = \mathcal{M} v^\varepsilon, \ c = 1 + \frac{1}{\ln x} \\ x = big \end{array} \right\} .$$

Get:

$$\sum_{n \leq x} d(n) = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \zeta(w)^2 \frac{x^w}{w} dw + O\left[\frac{x \ln^2 x}{T}\right]$$

$$+ O\left[\frac{x^{1+\varepsilon} \ln x}{T}\right] + O\left[\frac{x^{1+\varepsilon}}{T}\right] .$$

We'll write this with a minor <u>abuse</u> of language in the equivalent form

$$(\star) \qquad \sum_{n \leq x} d(n) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} \zeta(w)^2 \frac{x^w}{w}\,dw + O\left[\frac{x^{1+\varepsilon}}{T}\right].$$

The residue at $w=1$ was computed earlier as

$$x \ln x + (2\gamma - 1)x$$

in Lec 21, p. 9. It seems reasonable to now select any $\lambda \in (0,2]$ and push $\mathrm{Re}(w) = c$ over to $\mathrm{Re}(w) = 1-\lambda$. If $\lambda = 1$, we make a minor indentation at $w = 0$.

Prior to continuing, we recall that the fcn $\mu(\sigma)$ satisfies $\mu(\sigma) = \mu(1-\sigma) + \frac{1}{2} - \sigma$, which can be rewritten as

$$(\star\star) \qquad \mu(\sigma) + \frac{1}{2}\sigma = \mu(1-\sigma) + \frac{1}{2}(1-\sigma).$$

Put:

$$k(\sigma) = \left\{\begin{array}{ll} \frac{1}{2} - \sigma, & \sigma \leq 0 \\ \frac{1-\sigma}{2}, & 0 \leq \sigma \leq 1 \\ 0, & \sigma \geq 1 \end{array}\right\}, \quad \nu(\sigma) = \left\{\begin{array}{ll} \frac{1}{2} - \sigma, & \sigma \leq \frac{1}{2} \\ 0, & \sigma \geq \frac{1}{2} \end{array}\right\}.$$

The fcns $k$ and $\nu$ both satisfy $(\star\star)$; it is obvious that $k(\sigma) \geq \nu(\sigma)$. One <u>hopes</u> that $\mu(\sigma) = \nu(\sigma)$.

Let $A > 0$ satisfy

$$|\zeta(1-\lambda+it)| = O(1)\,(1+|t|)^{A}\,.$$

To avoid trivia in $(\star)$ when $c \hookrightarrow 1-\lambda$, we insist that

$$x^{1-\lambda}\,T^{2A} \lesssim x\,, \quad \underline{\text{i.e.}} \quad T^{2A} \lesssim x^{\lambda}\,.$$

In short order, by employing convexity à la pp. 8 + 12 in Lec 24, estimate $(\star)$ transforms into

horiz. contr.

$$\sum_{n \lesssim x} d(n) = O(T^{2A}x^{1-\lambda}) + O\!\left(T^{\varepsilon}\frac{x}{T}\right) + O(x^{\frac{1}{10}}) + x\ln x + (2\gamma-1)\,x + O\!\left(\frac{x^{1+\varepsilon}}{T}\right),$$

at least when $x > x_0(A,\varepsilon,\lambda)$. [ $x_0 = \exp\left(\frac{2A}{\varepsilon\lambda}\right)$ is fine.] The term $O(x^{\frac{1}{10}})$ is present whenever $\lambda \in [\frac{9}{10}, 2]$ in order to accommodate either the residue at $w=0$ or an indentation made to avoid any issues at $w=0$.

To optimize, we imagine $\varepsilon$ as $0$ and set

$$x^{1-\lambda} T^{2A} = \frac{x}{T} \, .$$

Thus

$$\boxed{T = x^{\frac{\lambda}{2A+1}}} \, ,$$

which then leads to a collective error term of

$$O(x^{\frac{1}{10}}) + O\left[ x^{\frac{2A+(1-\lambda)}{2A+1} + 2\varepsilon} \right] . \;//$$

We remark here that $T \leq x^{\lambda}$ and that $\frac{\lambda}{2A+1} < \frac{\lambda}{2A}$, this last relation showing that $T$ is admissible.

Because of the $2\varepsilon$, one is now free to substitute any number $\geq \mu(1-\lambda)$ for $A$ and still have a true result (i.e., valid remainder term).

We'll go with $A = k(1-\lambda)$. Accordingly,

$$0 < \lambda < 1 \Rightarrow A = \frac{\lambda}{2} \Rightarrow O\left[ x^{\frac{1}{1+\lambda} + 2\varepsilon} \right] \, ;$$

$$1 \leq \lambda \leq 2 \Rightarrow A = \lambda - \frac{1}{2} \Rightarrow O\left[ x^{\frac{1}{2} + 2\varepsilon} \right] .$$

The optimal estimate for $\sum_{n \leq x} d(n)$ with this [crude] Perron-type technique is therefore:

$$(\star\star\star) \qquad \sum_{n \leq x} d(n) = x \ln x + (2\gamma - 1)x + O(x^{\frac{1}{2}+2\varepsilon})$$

attained basically for any $\lambda \in [1, 2]$.

Under the Lindelöf Hypothesis, one can take $A = v(1-\lambda)$. This produces

$$0 < \lambda < \tfrac{1}{2} \Rightarrow A = 0 \Rightarrow O[x^{1-\lambda+2\varepsilon}]$$

$$\tfrac{1}{2} \leq \lambda \leq 2 \Rightarrow A \approx \lambda - \tfrac{1}{2} \Rightarrow O[x^{\frac{1}{2}+2\varepsilon}] \quad .$$

In other words: we <u>still</u> get $(\star\star\star)$, only beginning already at $\lambda \approx \frac{1}{2}$.

---

One can SUMMARIZE by saying that a crude application of Perron's method essentially leads to <u>just</u>

$$\sum_{n \leq x} d(n) = x \ln x + (2\gamma - 1)x + O(\sqrt{x}),$$

a result which was proved earlier, (indeed) almost effortlessly, in Lec 21, pages 2 + 3.

To obtain a remainder term of size $O(x^{1/3})$ ⟦which is the <u>classical</u> <u>nontrivial</u> estimate⟧, it is basically necessary to exploit the functional equation along $\mathrm{Re}(w) = 1-\lambda$ with $\lambda > 1$ and then use techniques resembling those in Lec 22, p. 15 (Lemma IV) and Lec 23, pp. 11–13. See Titchmarsh, Theory of $\zeta(s)$, § 12.2 with $k = 2$.

Compare: Landau, Vorlesungen, Sätze 508 + 509 (where an important, much more intrinsic method is used).

---

Out of curiosity, it is natural to wonder what taking $f(s) = \zeta(s)^1$ produces with this crude Perron-type technique. One would *hope* that $[\![x]\!]$ could be estimated reasonably accurately!

One has

$$\left\{ \begin{array}{c} a_n = 1,\ q = 1,\ \Phi(v) = 1,\ c = 1 + \dfrac{1}{\ln x} \\ x = \text{big} \end{array} \right\}.$$

It is convenient to keep $\lambda \in (0, M]$ with the tacit restriction that $|\lambda - 1| > 10^{-3}$ (say). Here $M$ = some large integer.

The earlier procedure is now easily mimicked. One insists that $T^A \lesssim x^\lambda$ (at the very least). The collective error term is easily seen to be

$$O(1) + O(T^A x^{1-\lambda}) + O(T^\varepsilon \tfrac{x}{T}) + O(\tfrac{x^{1+\varepsilon}}{T}).$$

One optimizes with $T = x^{\frac{\lambda}{A+1}}$. We have $T \lesssim x^M$. This leads to

$$O(1) + O\left[x^{\frac{A+(1-\lambda)}{A+1} + M\varepsilon}\right].$$

We have:

$$0 < \lambda < 1 \Rightarrow k(1-\lambda) = \frac{\lambda}{2} \Rightarrow O\left[x^{\frac{2-\lambda}{2+\lambda} + M\varepsilon}\right]$$
$$1 < \lambda \lesssim M \Rightarrow k(1-\lambda) = \lambda - \frac{1}{2} \Rightarrow O\left[x^{\frac{1}{2\lambda+1} + M\varepsilon}\right];$$

$$0 < \lambda < \frac{1}{2} \Rightarrow \nu(1-\lambda) = 0 \Rightarrow O\left[x^{1-\lambda+M\varepsilon}\right]$$
$$\frac{1}{2} \lesssim \lambda \lesssim M \Rightarrow \nu(1-\lambda) = \lambda - \frac{1}{2} \Rightarrow O\left[x^{\frac{1}{2\lambda+1} + M\varepsilon}\right].$$

In each instance, as $\lambda$ grows, the exponent <u>decreases</u>. Since $\varepsilon$ is arbitrary, we get

$$[\![x]\!] = x + O\left(x^{\frac{1}{2M+1} + \frac{\varepsilon}{2}}\right).$$

Letting $M$ grow, it emerges that

$$[\![x]\!] = x + O(x^\varepsilon).$$

$\boxed{\text{Not bad!}}$

16. (Lec 25, p. 20 last line) The proof can be seen in Selberg's *Collected Papers*, vol. 2, p. 225.

17. (Lec 26, p. 11) A very nice complex variable proof of this THEOREM is given in Titchmarsh, *Theory of $\zeta(s)$*, §4.14.

18. (Lec 28, p. 9) In Newman's General Thm, it goes virtually without saying that $g(ir)$ can be addressed simply by setting $s = \sigma + ir$.

19. (Lec 29, p. 29) In regard to the use of other kernel functions $k(w)$, it is worth mentioning that an interesting *variant* of the 1936 Ingham method is pushed through in Montgomery and Vaughan, *Multiplicative Number Theory*, vol. I, pages 477–479, 482 (bottom) – 483 (top).

Cf. also: Atle Selberg Archive, Hong Kong Lecture Series, 1998, Lec 4, pp. 2–10.
http://publications.ias.edu/selberg/section/2479

20. (Lec 30, p. 3) Prior to beginning the proof of THM, it would have been wise to step back and draw attention to an immediate Corollary; viz.,

First

{ Let $q \geq 1$. **If** $\sum_1^\infty \frac{\mu(n)(\ln n)^q}{n}$ converges, then $\sum_1^\infty \frac{\mu(n)}{n}$ also converges and its value must be $0$; as such, one will again get (i) + (ii) in an elementary fashion. }

Pf

That $\sum_{n=1}^\infty \frac{\mu(n)}{n}$ converges is self-evident by Dirichlet's test; see Lec 7, p. 1 (bottom). By item #13 above (i.e., Fact 2(b) in Lec 21), the summation's value is immediately ascertained to be $0$. ▨

**21.** (Lec 30, p. 3, proof of THM) The (following Landau's Handbuch) proof clearly possesses a certain beauty. After finishing it, however, one is left wondering how the "$q$-condition" of item #20 fits into the (overall) scheme.

This issue was ultimately clarified by Kienast in Math. Annalen 95 (1925) 427-445. See also Landau, Handbuch der Primzahlen, CHELSEA EDITION, Appendix (by P. Bateman), p. 941 § 159.

A bit of "diagram-chasing" is necessary in order to unravel Kienast's paper. Let $\sim$ mean "elementarily equivalent"; let

$$[M+k] \quad \text{mean} \quad M(x) = o(1)\frac{x}{(\ln x)^k}$$

$$[\psi+k] \quad \text{mean} \quad \psi(x) - x = o(1)\frac{x}{(\ln x)^k}$$

$$g_k \quad \text{mean} \quad \sum_1^\infty \frac{\mu(n)(\ln n)^k}{n} \quad \text{converges.}$$

(Here $k \geqq 0$.) Note that $A \sim B$ and $B \sim C \Rightarrow A \sim C$.

<u>Lemma 1</u> (baby calculus — very useful)

Let $\varphi \in C^1[a,b]$ be positive and monotonic. Let $F$ be real and <u>piecewise</u> $C^1$ on $[a,b]$. Assume that $|F(x)| \leqq \Phi(x)$, where $\Phi$ is positive + monotonic + $C^1$ on $[a,b]$. We then have

$$\left|\int_a^b \varphi(x)\,dF(x)\right| \leqq 2[\varphi(a)\Phi(a) + \varphi(b)\Phi(b)] + \left|\int_a^b \varphi(x)\,d\Phi(x)\right| .$$

<u>Pf</u>

The fcn $F$ has only a finite number of <u>actual</u> discontinuities. Because of the <u>Lipschitz condition</u> on the "chunks" of the rest of the graph, $F(x)$ has bounded total variation on $[a,b]$. Any such fcn is expressible as the difference of two monotonic increasing $\varphi_j(x)$. (The R–S integral of $\varphi dF$ is thus fine.)

We now exploit integ-by-parts twice.

$$\left| \int_a^b \varphi(x)\, dF(x) \right| = \left| \varphi(b)F(b) - \varphi(a)F(a) - \int_a^b F(x)\varphi'(x)\,dx \right|$$

$$\leq \varphi(b)\Phi(b) + \varphi(a)\Phi(a)$$

$$+ \left| \int_a^b \Phi(x)\varphi'(x)\,dx \right|$$

{this is correct when the sign of $\varphi'$ is fixed}

$$= \varphi(b)\Phi(b) + \varphi(a)\Phi(a)$$

$$+ \left| \Phi(b)\varphi(b) - \Phi(a)\varphi(a) - \int_a^b \varphi(x)\, d\Phi(x) \right|$$

$$\leq 2\varphi(b)\Phi(b) + 2\varphi(a)\Phi(a)$$

$$+ \left| \int_a^b \varphi(x)\, d\Phi(x) \right| . \quad ▨$$

## Lemma 2

Let $k \geqq 0$. Put $F(x) = \sum_{n \leqq x} \mu(n)(\ln n)^k$. We then have

$$M(x) = \frac{o(x)}{(\ln x)^k} \iff F(x) = o(x)$$

elementarily [i.e. via a sequence of elementary techniques].

Pf

$k = 0$ is trivial, so take $k \geqq 1$ wlog.

Assume first that $F(x) = o(x)$. Keep $T$ large.
Notice that

$$M(T) = O(1) + \int_2^T (\ln x)^{-k} dF(x)$$

$$= O(1) + O_\varepsilon(1) + \int_{T_\varepsilon}^T (\ln x)^{-k} dF(x)$$

$$\left\{ \frac{\text{apply Lemma 1 with } \Phi(x) = \varepsilon x}{\text{let } |\Theta| \leqq 1 \text{ as usual}} \right\}$$

$$= O_\varepsilon(1) + 2\Theta \left[ (\ln T)^{-k} \varepsilon T + O_\varepsilon(1) + \int_{T_\varepsilon}^T (\ln x)^{-k} d(\varepsilon x) \right]$$

$$= O_\varepsilon(1) + 2\Theta \frac{\varepsilon T}{(\ln T)^k} + 2\Theta\varepsilon \int_2^T \frac{dx}{(\ln x)^k}$$

familiar

$$= O_\varepsilon(1) + 2\Theta \frac{\varepsilon T}{(\ln T)^k} + 2\Theta\varepsilon \frac{O(T)}{(\ln T)^k}$$ (27)

$$= O_\varepsilon(1) + \varepsilon\, O(1) \frac{T}{(\ln T)^k}.$$

Hence $\mathcal{M}(T) = o(1) \frac{T}{(\ln T)^k}$. $\boxed{\text{OK}}$

Next, suppose that $\mathcal{M}(x) = \frac{o(x)}{(\ln x)^k}$. Keep $T$ large.
Notice that

$$F(T) = O(1) + \int_2^T (\ln x)^k\, d\mathcal{M}(x)$$

$$= O(1) + O_\varepsilon(1) + \int_{T_\varepsilon}^T (\ln x)^k\, d\mathcal{M}(x)$$

$$\left\{ \frac{\text{apply lemma 1 with } \Phi(x) = \varepsilon \frac{x}{(\ln x)^k}}{\text{let } |\Theta| \leqq 1 \text{ as usual}} \right\}$$

$$= O_\varepsilon(1) + 2\Theta\Big[ (\ln T)^k \varepsilon \frac{T}{(\ln T)^k} + O_\varepsilon(1)$$

$$+ \int_{T_\varepsilon}^T (\ln x)^k\, d\left(\frac{\varepsilon x}{(\ln x)^k}\right) \Big]$$

$$= O_\varepsilon(1) + 2\Theta\varepsilon T$$

$$+ 2\Theta\varepsilon \int_{\exp(k)}^T (\ln x)^k \left[ \frac{(\ln x)^k - k(\ln x)^{k-1}}{(\ln x)^{2k}} \right] dx$$

$$= O_\varepsilon(1) + 2\Theta\varepsilon T + 2\Theta\varepsilon \int_{\exp(k)}^{T} \left[1 - \frac{k}{\ln x}\right] dx$$

$$= O_\varepsilon(1) + 4\Theta\varepsilon T .$$

Hence $F(T) = o(T)$. $\blacksquare$

Going back (now) to $[M+k]$, $[\Psi+k]$, $g_k$, one observes that — in his paper — Kienast either recalls or proves:

Satz 1. $[M+0] \sim [\Psi+0] \sim g_0$ ;

Satz 3. $[\Psi+k] \sim g_k$, all $k \geq 0$ ;

Lemma 2 + Satz 10. Anytime $[\Psi+k]$ is true, we have $[M+(k+1)] \sim g_{k+1}$ .

above

The following assertion is now the key!

<u>CLAIM</u>: we have

$$[M+q] \sim g_q \sim [\Psi+q] \quad \text{for each } q \geq 0 .$$

Proof

Suppose not. Let $k$ be the smallest case where the proposed 3-way relation is FALSE. By Satz 1, $k \geq 1$.

Suppose $[M+k]$ holds. Clearly $[M+(k-1)]$ holds. But 3-way relation $[M+(k-1)] \sim g_{k-1} \sim [\Psi+(k-1)]$ is TRUE. Hence we get $[\Psi+(k-1)]$ elementarily. By Satz 10, the truth of $[M+k] \Rightarrow$ that of $g_k$ elementarily. So we can write $[M+k] \overset{e}{\Rightarrow} g_k$, the "e" meaning <u>elementary</u>.

Suppose next that $g_k$ holds. By Satz 3, we get $[\Psi+k]$ in an elementary fashion. So, $g_k \overset{e}{\Rightarrow} [\Psi+k]$.

Finally, suppose $[\Psi+k]$ holds. By Satz 3, $g_k$ holds elementarily. Trivially, of course, $[\Psi+(k-1)]$ holds. By Satz 10, we then have $g_k \Longrightarrow [M+k]$ elementarily. So, $[\Psi+k] \overset{e}{\Longrightarrow} [M+k]$.

All told, we have seen $[M+k] \overset{e}{\Rightarrow} g_k \overset{e}{\Longrightarrow} [\Psi+k] \overset{e}{\Rightarrow} [M+k]$. This <u>contradicts</u> the definition of $k$.

The CLAIM is thus proved. ▨

22. (Lec 30, p. 3, "already knew") Both here and in regard to Lec 2, p. 20 line 2, it is thought-provoking to have a look at Euler [1737], Opera Omnia, Ser. I, Vol. 14, pp. 242 (thm 19, second sentence) and 243–244. Cf. also item #12 above.

Euler's assertion that $\sum p^{-1-\varepsilon} = \ln\left(\frac{C}{\varepsilon} + O(1)\right)$ is but 1 or 2 steps away from heuristically concluding that $T(e^u) \sim \ln u$, wherein

$$T(x) \equiv \sum_{p \leq x} \frac{1}{p},$$

— and, then, via an analogous reasoning, engendering/sparking the suspicion that [in all likelihood] $\pi(x)$ must be roughly of size $\int_2^x \frac{dt}{\ln t} \sim \frac{x}{\ln x}$. Alas, none of this is said there explicitly.

In modern parlance, the $T(e^u)$ deduction (based strictly on real $\varepsilon > 0$) is a standard example of a Karamata-type Tauberian theorem; in this regard, cf. also Titchmarsh, Theory of $\zeta(s)$, (7.12.1)–(7.12.5).*
See Ingham, p. 10, for an Euler-style proof that $T(e^u) > \ln u - \frac{1}{2}$; note too p. 11 (bot, "infinitely fewer than the integers").

* and Hardy, *Divergent Series*, theorem 108

For some additional perspective on both aspects of this matter, $T$ and $\pi$, see Edwards, Riemann's Zeta Fcn, pp. 1–2.