مقاله جدید

اثباتی مقدماتی از قضیهٔ اعداد اول•

نورمن لوینسن
ترجمهٔ محمدصادق منتخب

۱. سرآغاز. یکی از برجسته‌ترین قضیه‌های ریاضیْ قضیهٔ اعداد اول است که نخست، پیش از سال ۱۸۰۰ میلادی، گاوس و لژاندر مستقل از هم آن را حدس زدند و در ۱۸۹۶، آدامار و پواسن مستقل از هم آن‌را اثبات کردند. چبیشف و ریمان در زمرهٔ ریاضیدانان برجستهٔ متعددی در سدهٔ نوزدهم بودند که موفق به اثبات این قضیه نشدند، ولی به نتیجه‌های مهمی در این مورد دست یافتند. ریمان نشان داد که قضیهٔ اعداد اول به رفتار تابع زتا در صفحهٔ مختلط بستگی دارد، و بسیاری از ویژگیهای این تابع را، که از همان زمان به‌نام وی معروف است، به‌دست آورد. یافته‌های ریمان در اثباتهای آدامار و پواسن به‌کار رفت و تعمیم یافت.

در سال ۱۹۴۹، اردوش و سلبرگ مشترکاً، و با استفاده از فرمولی که سلبرگ پیشتر به روشی مقدماتی اثبات کرده بود، چند اثبات مقدماتی از قضیهٔ اعداد اول ارائه کردند، [۳]. ولی نه این اثباتها و نه اثبات دیگر سلبرگ [۶]، علی‌رغم مقدماتی بودن، ساده نیستند.

به دلیلِ گسترش فوق‌العادهٔ ریاضیات، بسیاری از ریاضیدانها به مطالعهٔ نظریهٔ اعداد نمی‌پردازند. بنابراین، به نظر می‌رسد ارائهٔ شرحی خودکفا از یک اثبات مقدماتی قضیهٔ اعداد اول، همراه با ذکر زمینه و انگیزهٔ گامهای اثبات، سودمند باشد.

عددهای اول $\{۲, ۳, ۵, ۷, ۱۱, ۱۳, \cdots\}$ را از دوران باستان می‌شناختند و در آثار اقلیدس اثباتی هست از اینکه تعداد آنها نامتناهی است. تعداد عددهای اول نابیشتر از x، $\pi(x)$ نامیده می‌شود و می‌توان آن‌را به‌صورت

$$\pi(x) = \sum_{p\leq x} ۱ \qquad (۱.۱)$$

نشان داد که در آن نماد p بر دنبالهٔ عددهای اول، به‌ترتیب صعودی، تغییر می‌کند. ساده‌ترین صورت حدس لژاندر چنین است

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x} = ۱ \qquad (۲.۱)$$

معلوم شده است حدس گاوس، که می‌گوید مقدار $\pi(x)$ به‌ازای xهای بزرگ تقریباً برابر $\int_۲^x \frac{dt}{\log t}$ است، عمیقتر است. گاوس با بررسی جدولی از عددهای اول به این نتیجه رسید که ظاهراً عددهای اول یک چگالی مجانبی دارند که مقدارش در x برابر $\frac{۱}{\log x}$ است. به این دلیل، برای بعضی مقاصد، بهتر است برای تعیین رفتار مجانبی عددهای اول به هر p وزن $\log p$ را بدهیم. این کار به‌کمک تابع

$$\theta(x) = \sum_{p\leq x} \log p \qquad (۳.۱)$$

انجام می‌گیرد. درواقع، به‌طوری که به‌زودی خواهیم دید، حتی مناسبتر است که نه از θ بلکه از تابع ψای که به θ بستگی بسیار نزدیکی دارد، استفاده کنیم.

نورمن لوینسن (Norman Levinson) در سال ۱۹۱۲ متولد شد و در سال ۱۹۷۵ درگذشت. وی بیشتر عمر علمی خود را در دانشگاه ام. آی. تی گذراند و از آغاز تحصیل تحت تأثیر نوربرت وینر بود، هرچند که دورهٔ دکتری ریاضی خود را در دانشگاه پرینستن زیر نظر جان فون نویمان گذراند. نفوذ علمی این دو ریاضیدان نامدار در تحقیقات علمی لوینسن مشهود است. او آثار مهمی در زمینه‌های احتمال، پردازش سیگنال، معادلات دیفرانسیل، آنالیز مختلط، و نظریهٔ اعداد دارد. لوینسن در اواخر عمرش با استفاده از یک رهیافت احتمالاتی، آثار بدیعی در ارتباط با فرضیهٔ ریمان در مورد صفرهای تابع زتا پدید آورد. مقالهٔ حاضر، در تشریح اثبات مقدماتی اردوش وسلبرگ از قضیهٔ اعداد اول، که ترجمه‌اش به مناسبت درگذشت اردوش در اینجا به چاپ می‌رسد، از آثار توصیفی لوینسن است که در سال ۱۹۷۱ به دریافت جایزهٔ شوونه (Chauvenet) از جامعهٔ ریاضی آمریکا نائل شده است.

شرحی که می‌آید، با تجزیهٔ عدد صحیح به حاصلضرب توانهای عددهای اول آغاز می‌شود و با توصیف اثباتهایی از کشفهای مربوطه در سدهٔ نوزدهم در بخشهای ۲ و ۳ ادامه می‌یابد. این رهیافت در بخش ۴ با اثبات فرمول سلبرگ، و سرانجام در بخش ۵، با اثبات سلبرگ [۶] به شکلی که به‌وسیلهٔ رایت [۴]، [۹]، ساده شده و نویسندهٔ این مقاله آن‌را ساده‌تر هم کرده [۵]، دنبال می‌شود.

اثبات مقدماتی قضیهٔ اعداد اول را برویش [۲]، بمبیری [۱]، و ویرزینگ [۸] و دیگران بسط داده و اثباتهایی مقدماتی برای صورتهای قویتری از قضیه همراه با باقیمانده به‌دست آورده‌اند.

۲. اتحاد چبیشف و معکوس‌سازیِ آن. در آغاز خاطرنشان می‌کنیم که هر عدد صحیح مثبت را می‌توان به حاصلضربی از توانهای عددهای اول متمایز تجزیه کرد. بنابراین، به‌ازای هر عدد صحیح مثبت n

$$n = p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m} \tag{۱.۲}$$

که p_jها، $1 \le j \le m$، عددهای اول متمایزند و هر k_j عدد صحیح مثبتی است. چون جمع ساده‌تر از ضرب است، صورت مفیدتری از (۱.۲) عبارت است از

$$\log n = k_1 \log p_1 + k_2 \log p_2 + \cdots + k_m \log p_m \tag{۲.۲}$$

فایدهٔ این فرمول در صورت استفاده از نماد فون من‌گولت[1]، $\Lambda(n)$، خیلی بیشتر می‌شود. این نماد در ۱۸۹۵ معرفی شد و به‌صورت زیر تعریف می‌شود

$$\Lambda(n) = \begin{cases} \log p & \text{اگر } n \text{ توانی صحیح و مثبت از عدد اول } p \text{ باشد} \\ 0 & \text{در غیر این‌صورت} \end{cases} \tag{۳.۲}$$

بنابراین، فقط وقتی‌که n توانی از عددی اول باشد، $\Lambda(n) \neq 0$.

نماد $\sum_{j|n}$ برای نشان دادن مجموعی روی j، وقتی j همهٔ مقسوم‌علیه‌های مثبت عدد صحیح مثبت n را اختیار می‌کند، به‌کار می‌رود. با این نماد، نشان خواهم داد که (۲.۲) را می‌توان به‌صورت

$$\log n = \sum_{j|n} \Lambda(j) \tag{۴.۲}$$

نوشت. برای اثباتِ (۴.۲) توجه می‌کنیم که بنا به (۱.۲) و تعریف $\Lambda(j)$، $\log p_1$، $\log p_2$، ...، $\log p_m$ تنها جمله‌های ناصفری هستند که در طرف راست (۴.۲) ظاهر می‌شوند. به‌علاوه $\log p_1$ به‌ازای $j = p_1$، به‌ازای $j = p_1^2$، ...، و به‌ازای $j = p_1^{k_1}$ ظاهر می‌شود. بنابراین، $\log p_1$ دقیقاً k_1 بار ظاهر می‌شود؛ همین‌طور $\log p_2$ دقیقاً k_2 بار ظاهر می‌شود، و ...، که نشان می‌دهد (۴.۲) نتیجه‌ای از (۲.۲) است. فرمول (۴.۲) صورتی فوق‌العاده قوی از (۱.۲) است و ویژگیهایی از عددهای اول را که در اینجا مورد نیازند، دربر می‌گیرد. تبدیل (۲.۲) به‌صورت (۴.۲) واضح نیست و از نظر تاریخی نسبتاً جدیدتر است.

فرمول (۴.۲) را می‌توان به‌صورتِ معادلِ

$$\log n = \sum_{ij=n} \Lambda(j) \tag{۵.۲}$$

نوشت که در آن i و j عددهایی صحیح مثبت‌اند که هر یک از آنها همهٔ مقدارهای ممکنِ صادق در $ij = n$ را اختیار می‌کند، بنابراین j همهٔ مقسوم‌علیه‌های مثبت n را اختیار می‌نماید (i نیز چنین می‌کند).

میان تعداد عددهای اول نابیشتر از x، یعنی $\pi(x)$، و مجموع

$$\psi(x) = \sum_{j \le x} \Lambda(j) \tag{۶.۲}$$

رابطهٔ نزدیکی برقرار است. بنا به تعریف $\Lambda(j)$

$$\begin{aligned} \psi(x) &= \sum_{p \le x} \log p + \sum_{p^2 \le x} \log p + \sum_{p^3 \le x} \log p + \cdots \\ &= \theta(x) + \theta(x^{\frac{1}{2}}) + \theta(x^{\frac{1}{3}}) + \cdots \end{aligned}$$

چبیشف تابع $\psi(x)$ را، به این صورتِ اخیر، می‌شناخت و به‌روشی ساده ثابت کرد که قضیهٔ اعداد اول (۲.۱) معادل با

$$\lim_{x \to \infty} \frac{\psi(x)}{x} = 1 \tag{۷.۲}$$

است. این اثبات در لم ۴.۳ می‌آید، و (۷.۲) در بخش ۵ ثابت خواهد شد. (با قدری تسامح، $\psi(x)$ به‌ازای xهای بزرگ مانند $\pi(x) \log x$ عمل می‌کند، زیرا، با توجه به (۳.۲)، $\psi(x)$ هر $p \le x$ را با وزن $\log p$ می‌شمارد و به‌ازای «بیشترِ» pهای نابیشتر از x، مقدار $\log p$ به $\log x$ نزدیک است. به علاوه، $\psi(x)$، $\log p$ را به‌ازای $p \le x^{\frac{1}{2}}$، به‌ازای $p \le x^{1/3}$، و ... نیز می‌شمارد، گرچه طبق اثباتی که بعداً ارائه خواهد شد، اینها بسیار پراکنده‌اند.)

برای استفاده از (۵.۲) به‌منظور استخراج اطلاعی دربارهٔ $\psi(x)$ مجموع هر یک از طرفین (۵.۲) را به‌ازای nهای نابزرگتر از x تشکیل می‌دهیم، یعنی

$$\sum\nolimits_{n \le x} \log n = \sum\nolimits_{n \le x} \sum\nolimits_{ij=n} \Lambda(j)$$

به‌طوری که اگر تعریف کنیم

$$T(x) = \sum_{n \le x} \log n \tag{۸.۲}$$

آنگاه

$$T(x) = \sum_{ij \le x} \Lambda(j) \tag{۹.۲}$$

چون لگاریتم تابع همواری است، $T(x)$ را می‌توان به‌ازای xهای بزرگ به‌سادگی محاسبه کرد، و این کار در (۴.۳) انجام خواهد یافت.

مجموع دوگانهٔ (۹.۲) به‌ازای آن دسته از نقطه‌های شبکه‌ایِ ربع مثبت صفحهٔ (i, j) محاسبه می‌گردد که رویِ یا زیرِ هذلولی $ij = x$ قرار دارند.

1. von Mangoldt

اگر مجموع دوگانهٔ (۹.۲) مجموعی مکرر به‌حساب آورده شود، و مجموعه‌یابی نخست روی j انجام گیرد، داریم

$$T(x)=\sum_{i\le x}\sum_{j\le \frac{x}{i}}A(j)=\sum_{i\le x}\psi\left(\frac{x}{i}\right)$$

این اتحاد را چبیشف (در ۱۸۵۰) کشف کرد و به‌صورت

$$T(x)=\sum_{n\le x}\psi\left(\frac{x}{n}\right) \qquad (۱۰.۲)$$

نوشته می‌شود.

درواقع، اتحاد چبیشف (۱۰.۲) رابطه‌ای تبدیلی است؛ و از آن چنین برمی‌آید که می‌توان به‌ازای هر تابع $F(x)$، $x>1$، تابع وابستهٔ $G(x)$ای، $x>1$، با ضابطهٔ

$$\begin{aligned}G(x)=\sum_{n\le x}F\left(\frac{x}{n}\right)&=F(x)+F\left(\frac{x}{2}\right)\\&+F\left(\frac{x}{3}\right)+\cdots+F\left(\frac{x}{[x]}\right)\end{aligned} \qquad (۱۱.۲)$$

تعریف کرد که در آن $[x]$، طبق معمول، به‌معنی بزرگترین عدد صحیح نابیشتر از x است. $G(x)$ را می‌توان تبدیلی از $F(x)$ به حساب آورد. به‌طوری که دیده می‌شود، G تابع خطی همگنی از F است. رابطه‌های این تبدیلات از جملهٔ کارامدترین ابزارهای ریاضیدان است و این یکی هم مستثنی نیست.

چون $T(x)$ تابع نسبتاً ساده‌ای است، کوشش در معکوس‌سازی رابطهٔ (۱۰.۲) برای بیان $\psi(x)$ برحسب T، با استفاده از نمادهای کلیتر، معکوس‌سازی (۱۱.۲) برای محاسبهٔ F برحسب G، کار جالب و مفیدی خواهد بود. یک گام کوچک اولیه برای محاسبهٔ F برحسب G، حذف $F(\frac{x}{2})$ از طرف راست (۱۱.۲) است. این‌کار را می‌توان به‌سادگی با قراردادن $\frac{x}{2}$ به‌جای x در (۱۱.۲) انجام داد. داریم

$$G\left(\frac{x}{2}\right)=F\left(\frac{x}{2}\right)+F\left(\frac{x}{4}\right)+F\left(\frac{x}{6}\right)+\cdots$$

با تفریق این رابطه از (۱۱.۲)، $F(\frac{x}{2})$ حذف می‌شود. این فرایند را بلافاصله می‌توان با قراردادن $\frac{x}{2}$، و سپس $\frac{x}{3}$، و آنگاه ... به‌جای x در (۱۱.۲)، تعمیم داد. در این صورت

$$\begin{array}{llllllll}G(x) = & F(x)+F(\frac{x}{2})+ & F(\frac{x}{3}) & +F(\frac{x}{4}) & +F(\frac{x}{5}) & +F(\frac{x}{6}) & +\cdots\\ G(\frac{x}{2}) = & F(\frac{x}{2}) & & +F(\frac{x}{4}) & & +F(\frac{x}{6}) & +\cdots\\ G(\frac{x}{3}) = & & F(\frac{x}{3}) & & & +F(\frac{x}{6}) & +\cdots\\ G(\frac{x}{4}) = & & & F(\frac{x}{4}) & & & +\cdots\\ G(\frac{x}{5}) = & & & & F(\frac{x}{5}) & & +\cdots\\ G(\frac{x}{6}) = & & & & & F(\frac{x}{6}) & +\cdots\end{array} \qquad (۱۲.۲)$$

اگر معادله‌ها را به‌ترتیب بنویسیم، می‌توانیم نخست $F(\frac{x}{2})$، سپس $F(\frac{x}{3})$، سپس $F(\frac{x}{4})$، ... را از سمت راست حذف کنیم. به‌عنوان مثال تا $G(\frac{x}{6})$ به‌دست می‌آوریم

$$F(x)=G(x)-G\left(\frac{x}{2}\right)-G\left(\frac{x}{3}\right)-G\left(\frac{x}{5}\right)+G\left(\frac{x}{6}\right)+\cdots$$

از این رابطه چنین برمی‌آید، و درواقع به‌دلایل صورتِ قطری طرف راست (۱۲.۲) ثابت می‌شود، که (۱۱.۲) را می‌توان به‌کمک فرمولی از نوع

$$F(x)=\sum_{k\le x}\mu(k)G\left(\frac{x}{k}\right) \qquad (۱۳.۲)$$

معکوس کرد به شرط اینکه $\mu(k)$ مشخص شود (موبیوس، ۱۸۳۲). برای تعیین $\mu(k)$ توجه می‌کنیم که بنا به (۱۱.۲)، $G(\frac{x}{k})=\sum_{j\le \frac{x}{k}}F(\frac{x}{jk})$، که اگر در (۱۳.۲) جایگزین شود

$$F(x)=\sum_{k\le x}\mu(k)\sum_{j\le\frac{x}{k}}F\left(\frac{x}{jk}\right)=\sum_{jk\le x}\mu(k)F\left(\frac{x}{jk}\right)$$

اگر این مجموع دوگانه نخست بر نقاط شبکه‌ای روی هذلولیهای $jk=n$ و سپس به‌ازای n، $1\le n\le x$ محاسبه شود، داریم

$$F(x)=\sum_{n\le x}F(x/n)\sum_{jk=n}\mu(k) \qquad (۱۴.۲)$$

معادلهٔ (۱۴.۲) به شرط $\mu(1)=1$، و در صورتِ قراردادن $k|n$ به‌جای $jk=n$ به شرط

$$\sum_{k|n}\mu(k)=0 \text{ ، } n\ge 2 \qquad (۱۵.۲)$$

به اتحاد تبدیل می‌شود. اگر متوالیاً قرار دهیم $n=2,3,4,\ldots$، $\mu(k)$ به‌طور یکتا معین می‌شود. برای تعیین صریح $\mu(k)$ حالت $n=p$ را امتحان می‌کنیم. در این صورت $k=1$ و $k=p$، که نتیجه می‌شود $\mu(1)+\mu(p)=0$، و بنابراین $\mu(p)=-1$. در حالت $n=p_1p_2$ نتیجه می‌شود

$$\mu(1)+\mu(p_1)+\mu(p_2)+\mu(p_1p_2)=0$$

و بنابراین $\mu(p_1p_2)=1$. به‌همین نحو به‌آسانی معلوم می‌شود که

$$\mu(p_1p_2p_3)=-1 \text{، } \mu(p^2)=0 \text{، } \mu(p^3)=0 \text{، } \ldots \text{، } \mu(p_1^2p_2)=0$$

از اینجا چنین برمی‌آید که اگر p_1، ...، p_m عددهای اول متمایزی باشند و $n=p_1p_2\cdots p_m$، آنگاه

$$\mu(n)=(-1)^m \qquad (۱۶.۲)$$

و اگر $p^2|n$، و p عددی اول باشد

$$\mu(n)=0 \qquad (۱۷.۲)$$

تابع $\mu(n)$ به تابع موبیوس موسوم است.

اکنون ثابت می‌کنیم که اگر $\mu(n)$ به‌صورت (۱۶.۲) و (۱۷.۲) تعریف شود، آنگاه (۱۵.۲) نتیجه می‌شود. خاطرنشان می‌کنیم تابعی که در (۱۵.۲) صدق کند یکتاست. به‌ازای $n=p_1^{k_1}\cdots p_m^{k_m}$، بنا به (۱۷.۲)

$$\sum_{j|n}\mu(j)=\sum_{j|p_1\cdots p_m}\mu(j)$$

بنابراین، کافی است طرف راست را برای اثبات (۱۵.۲) به‌کار بریم. اگر $m = 1$، (۱۵.۲) برقرار است، زیرا $\mu(1) = 1$ و $\mu(p_1) = -1$. اگر $m \geq 2$،

$$\sum_{j|p_1\cdots p_m}\mu(j) = \sum_{k|p_1\cdots p_{m-1}}(\mu(k)+\mu(kp_m)) \qquad (۱۸.۲)$$

ولی، اگر k در (۱۸.۲) حاصلضرب r عدد اول باشد، بنا به (۱۶.۲)

$$\mu(kp_m) = (-1)^{r+1} = -\mu(k)$$

پس هر جملهٔ طرف راست (۱۸.۲) صفر است و بنابراین (۱۵.۲) ثابت می‌شود. به‌علاوه، بنا به (۱۶.۲) و (۱۷.۲)

$$|\mu(n)| \leq 1 \qquad (۱۹.۲)$$

(که تنها استفاده‌ای است که از مطالبِ میان (۱۵.۲) و (۱۹.۲) می‌کنیم).

از اعمال فرمول معکوس‌سازی موبیوس (۱۳.۲) بر اتحاد چبیشف (۱۰.۲) فرمول معکوس‌سازی

$$\psi(x) = \sum_{k\leq x}\mu(k)T(\frac{x}{k}) \qquad (۲۰.۲)$$

به‌دست می‌آید. (۲۰.۲) را می‌توان با استفاده از تعریفهای ψ و T به‌صورت

$$\begin{aligned}\sum_{n\leq x}\Lambda(n) &= \sum_{k\leq x}\mu(k)\sum_{j\leq \frac{x}{k}}\log j = \sum_{jk\leq x}\mu(k)\log j\\ &= \sum_{n\leq x}\sum_{jk=n}\mu(k)\log j = \sum_{n\leq x}\sum_{k|n}\mu(k)\log\frac{n}{k}\end{aligned}$$

نوشت. از اینجا به‌ازای $x = 1, 2, 3, \ldots$ نتیجه می‌گیریم

$$\Lambda(n) = \sum_{k|n}\mu(k)\log\frac{n}{k}, \quad n\geq 1 \qquad (۲۱.۲)$$

که فرمول معکوس‌سازی (۴.۲) است. ملاحظه می‌شود که (۲۱.۲) را می‌توان مستقیماً با استفاده از (۴.۲) و (۱۵.۲) نیز به‌دست آورد.

۳. چند حکم مقدماتی. گرچه همهٔ حکمها مربوط به عددهای اول که در این بخش می‌آیند در سدهٔ نوزدهم کشف شدند، برخی از آنها تا هفتاد سال بعد از آنکه لژاندر و گاوس قضیهٔ اعداد اول را حدس زدند به‌دست، نیامده بودند.

در این بحث استفاده از لم معروف زیر، که در آن، طبق معمول، $[x]$ به‌معنی بزرگترین عدد صحیح نابیشتر از x است، مناسب است.

لم ۱.۳ فرض می‌کنیم $f(t)$ دارای مشتق پیوستهٔ $f'(t)$، $t \geq 1$، باشد. همچنین فرض می‌کنیم c_n، $n \geq 1$، ثابتهایی باشند و $C(u) = \sum_{n\leq u} c_n$. در این صورت

$$\sum_{n\leq x} c_n f(n) = f(x)C(x) - \int_1^x f'(t)C(t)dt \qquad (۱.۳)$$

و

$$\begin{aligned}\sum_{n\leq x} f(n) = \int_1^x f(t)dt + \int_1^x (t-[t])f'(t)dt\\ + f(1) - (x-[x])f(x) \qquad (۲.۳)\end{aligned}$$

اثبات. چون $C(u)$ تابعی پله‌ای است، $C(n) - C(n-1) = c_n$ و $C(u) = C([u])$. بنابراین، اگر $[x] = N$

$$\begin{aligned}\sum_{n\leq x} c_n f(n) &= \sum_{n\leq x}(C(n)-C(n-1))f(n)\\ &= \sum_{n\leq x-1} C(n)(f(n)-f(n+1)) + C(x)f(N)\\ &= -\sum_{n\leq x-1} C(n)\int_n^{n+1} f'(t)dt + C(x)f(N)\\ &= -\int_1^N C(t)f'(t)dt + C(x)f(N)\end{aligned}$$

چون $C(t)$ بر $N \leq t < x$ ثابت است،

$$\int_N^x C(t)f'(t)dt = C(x)(f(x)-f(N))$$

با جمع کردن این رابطه با رابطهٔ قبلی و انتقال انتگرالِ طرف چپ به راست، (۱.۳) نتیجه می‌شود.

در حالت $c_n = 1$، (۱.۳) به این صورت در می‌آید

$$\begin{aligned}\sum_{n\leq x} f(n) &= [x]f(x) - \int_1^x f'(t)[t]dt\\ &= [x]f(x) - \int_1^x f'(t)tdt + \int_1^x f'(t)(t-[t])dt\end{aligned}$$

با انتگرالگیری از نخستین انتگرال سمت راست به روش جزءبه‌جزء، (۲.۳) ثابت می‌شود.

در اینجا استفاده از نماد زیر مناسب است. فرض می‌کنیم $f(x)$ به‌ازای xهای متناهی کراندار باشد و به‌ازای xهای بزرگ، عدد ثابت K و تابع $g(x)$ موجود باشد به‌طوری که

$$|f(x)| \leq Kg(x)$$

در این صورت می‌نویسیم

$$f(x) = O(g(x)) \qquad (۳.۳)$$

و هر جا مناسب باشد، طرف راست (۳.۳) را جایگزین $f(x)$ می‌کنیم.

با به کار بردن (۲.۳) در مورد $f(t) = \log t$ و استفاده از $0 \leq t - [t] < 1$ نتیجه می‌گیریم

$$T(x) = x\log x - x + O(\log x) \qquad (۴.۳)$$

که صورت ضعیفی از فرمول استرلینگ است.

لم ۲.۳ (چبیشف، ۱۸۵۰). به‌ازای xهای بزرگ

$$\psi(x) < \frac{۳}{۲}x \tag{۵.۳}$$

اثبات. با استفاده از اتحاد چبیشف (۱۰.۲)

$$T(x) - ۲T(\frac{x}{۲}) = \psi(x) - \psi(\frac{x}{۲}) + \psi(\frac{x}{۳}) - \psi(\frac{x}{۴}) + \cdots \geq \psi(x) - \psi(\frac{x}{۲})$$

زیرا ψ یکنوای غیرنزولی است و بنابراین $\psi(\frac{x}{۲n-۱}) - \psi(\frac{x}{۲n}) \geq ۰$. بنا به (۳.۴)، به‌ازای Kی ثابتی، اگر $x \geq ۲$، آنگاه

$$\psi(x) - \psi(\frac{x}{۲}) \leq x\log ۲ + K\log x$$

در صورت قراردادن $\frac{x}{۲^j}$ به‌جای x، به شرط $۲ \leq \frac{x}{۲^j}$ که معادل است با $j < \frac{\log x}{\log ۲}$، به‌دست می‌آوریم

$$\psi(\frac{x}{۲^j}) - \psi(\frac{x}{۲^{j+۱}}) \leq \frac{x}{۲^j}\log ۲ + K\log x \tag{۶.۳}$$

چون $\psi(t) = ۰$، $t < ۲$، با مجموعیابی از هر یک از طرفین (۶.۳) روی $۰ \leq j < \frac{\log x}{\log ۲}$ داریم

$$\begin{aligned}\psi(x) &\leq x\log ۲(۱ + \frac{۱}{۲} + \frac{۱}{۲^۲} + \cdots) + \frac{\log x}{\log ۲}K\log x \\ &= ۲x\log ۲ + K\frac{\log^۲ x}{\log ۲}\end{aligned}$$

چون $\log ۲ < ۰٫۷$، حکم (۵.۳) نتیجه می‌شود.

لم ۳.۳ (این لم را مرتنس در ۱۸۷۴ با کمی تفاوت ثابت کرده است)

$$\sum_{n\leq x}\frac{\Lambda(n)}{n} = \log x + O(۱) \tag{۷.۳}$$

اثبات. در مجموع دوگانهٔ (۹.۲)، عمل مجموعیابی را نخست روی i و سپس روی j انجام می‌دهیم (برعکس ترتیبی که در استخراج (۱۰.۲) به‌کار رفت) و به‌دست می‌آوریم

$$\begin{aligned}T(x) &= \sum_{j\leq x}\Lambda(j)\sum_{i\leq x/j}۱ = \sum_{j\leq x}\Lambda(j)\left[\frac{x}{j}\right] \\ &= x\sum_{j\leq x}\frac{\Lambda(j)}{j} - \sum_{j\leq x}\Lambda(j)\left(\frac{x}{j} - \left[\frac{x}{j}\right]\right)\end{aligned} \tag{۸.۳}$$

به‌علاوه، بنا به (۵.۳)

$$۰ \leq \sum_{j\leq x}\Lambda(j)\left(\frac{x}{j} - \left[\frac{x}{j}\right]\right) \leq \sum_{j\leq x}\Lambda(j) = \psi(x) = O(x) \tag{۹.۳}$$

ملحوظ کردن (۴.۳) و (۹.۳) در (۸.۳)، (۷.۳) را نتیجه می‌دهد.

لم ۴.۳

$$\psi(x) = \pi(x)\log x + O(\frac{x\log\log x}{\log x}) \tag{۱۰.۳}$$

بنابراین، (۷.۲) با قضیهٔ اعداد اول معادل است.

اثبات. بنا به (۳.۲) و تعریف (۶.۲)

$$\psi(x) = \sum_{p\leq x}\log p + \sum_{p\leq x^{۱/۲}}\log p + \sum_{p\leq x^{۱/۳}}\log p + \cdots \tag{۱۱.۳}$$

که فقط به شرط $۲ \leq x^{\frac{۱}{j}}$ یا $j \leq \frac{\log x}{\log ۲}$، مجموعهای $p \leq x^{\frac{۱}{j}}$ فوق ناصفرند. پس

$$\psi(x) \leq \sum_{p\leq x}\log p + \frac{\log x}{\log ۲}\sum_{p\leq x^{\frac{۱}{۲}}}\log p$$

بنابراین، طبق تعریف (۱.۱) از $\pi(x)$

$$\psi(x) \leq \log x.\pi(x) + \frac{\log x}{\log ۲}\pi(x^{۱/۲})\log x^{۱/۲}$$

چون $\pi(y) \leq y$، نتیجه می‌شود

$$\psi(x) \leq \log x\,\pi(x) + \frac{x^{۱/۲}\log^۲ x}{۲\log ۲} \tag{۱۲.۳}$$

بنا به (۱۱.۳)

$$\begin{aligned}\psi(x) &\geq \sum_{x/\log^۲ x < p\leq x}\log p \geq \log\left(\frac{x}{\log^۲ x}\right)\sum_{x/\log^۲ x < p\leq x}۱ \\ &= \log\left(\frac{x}{\log^۲ x}\right)\left(\pi(x) - \pi\left(\frac{x}{\log^۲ x}\right)\right)\end{aligned}$$

چون $\pi(y) \leq y$، به‌دست می‌آوریم

$$\frac{\psi(x)}{\log x - ۲\log\log x} \geq \pi(x) - \frac{x}{\log^۲ x}$$

یا

$$\begin{aligned}\pi(x)\log x &\leq \psi(x)\frac{\log x}{\log x - ۲\log\log x} + \frac{x}{\log x} \\ &= \psi(x) + \psi(x)\frac{۲\log\log x}{\log x - ۲\log\log x} + \frac{x}{\log x}\end{aligned}$$

با استفاده از (۵.۳) و $۲\log\log x < \frac{\log x}{۲}$ به‌ازای xهای بزرگ

$$\pi(x)\log x \leq \psi(x) + \frac{۴x\log\log x}{\log x} + \frac{x}{\log x} \tag{۱۳.۳}$$

که همراه با (۱۲.۳) لم را نتیجه می‌دهد.

لم ۵.۳ اگر γ ثابت اویلر باشد،

$$\sum_{n\le x}\frac{1}{n}=\log x+\gamma+O(1/x) \qquad (۱۴.۳)$$

اثبات. با اِعمال (۲.۳) بر $f(t)=\frac{1}{t}$ داریم

$$\sum_{n\le x}\frac{1}{n}=\log x-\frac{x-[x]}{x}+1-\int_1^x\frac{t-[t]}{t^2}dt$$

اگر

$$\gamma=1-\int_1^{+\infty}\frac{t-[t]}{t^2}dt \qquad (۱۵.۳)$$

آنگاه $\sum_{n\le x}\frac{1}{n}=\log x+\gamma+H$، که چون $0\le t-[t]<1$،

$$H=\int_x^{+\infty}\frac{t-[t]}{t^2}dt-\frac{x-[x]}{x}=O(\frac{1}{x})$$

و حکم (۱۴.۳) نتیجه می‌شود.

توجه. بنا به (۱۵.۳)، $0<\gamma<1$.

۴. نابرابری مقدماتی سلبرگ. اکنون فرمول معکوس‌سازیِ موبیوس (۲۰.۲) را، که ψ را برحسب T بیان می‌کند، به منظور بررسی رفتار $\psi(x)$ به‌ازای xهای بزرگ به‌کار می‌بریم. اگر بتوانیم $F(x)$ ساده‌ای، که آن را $\bar{F}(x)$ می‌نامیم، با تبدیل $\bar{G}(x)$ای نزدیک به $T(x)$ پیدا کنیم، محاسبه ساده‌تر خواهد شد. در این حالت، فرمول معکوس‌سازی موبیوس برای $\bar{F}$، (۱۳.۲)، را از فرمول مربوط به ψ کم می‌کنیم

$$\psi(x)-\bar{F}(x)=\sum_{k\le x}\mu(k)\left(T\left(\frac{x}{k}\right)-\bar{G}\left(\frac{x}{k}\right)\right) \qquad (۱.۴)$$

اگر بتوانیم نشان دهیم که طرف راست کوچک است، آنگاه $\psi(x)$ به $\bar{F}(x)$ نزدیک خواهد بود.

اگر بتوانیم ثابت کنیم که $\frac{\psi(x)}{x}\to 1$، $\psi(x)$ به ازای xهای بزرگ به x نزدیک خواهد بود. بنابراین $\bar{F}(x)=F_0(x)=x$ را امتحان می‌کنیم. پس $G_0(x)=\sum_{n\le x}F_0(\frac{x}{n})=x\sum_{n\le x}n^{-1}$، که بنا به (۱۴.۳)، نتیجه می‌شود $G_0(x)=x\log x+\gamma x+O(1)$. این به اندازهٔ کافی به $T(x)$ به‌صورت داده شده در (۴.۳) نزدیک نیست. برای اصلاح آن قرار می‌دهیم $\bar{F}(x)=F_1(x)=x-C$، که C عدد ثابتی است. (انتخابهای بسیاری جز C وجود دارند که در اینجا مناسب‌اند.) پس

$$\begin{aligned}G_1(x)&=x\sum_{n\le x}\frac{1}{n}-C\sum_{n\le x}1\\&=x\log x+\gamma x+O(1)-C[x]\\&=x\log x-(C-\gamma)x+O(1)\end{aligned}$$

پس اگر $C=1+\gamma$، بنا به (۴.۳)

$$T(x)-G_1(x)=O(\log x) \qquad (۲.۴)$$

که نسبتاً کوچک است. با استفاده از (۱.۴) و به ازای $\bar{F}=x-C$ داریم

$$\psi(x)-x+C=\sum_{k\le x}\mu(k)\left(T\left(\frac{x}{k}\right)-G_1\left(\frac{x}{k}\right)\right) \qquad (۳.۴)$$

حتی اگر طرف راست (۲.۴) به‌صورت قویتر $O(1)$ می‌بود (که نیست)، از اینکه $|\mu(k)|\le 1$ (بنا به (۱۹.۲))، فقط نتیجه می‌شود که طرف راست (۳.۴)، $O(x)$ است. بنابراین، فرمول معکوس‌سازی موبیوس (۳.۴) قضیهٔ اعداد اول را نتیجه نمی‌دهد بلکه حداکثر، نتیجهٔ به مراتب ضعیفترِ

$$\psi(x)=O(x) \qquad (۴.۴)$$

را (که پیشتر در لم ۲.۳ از راه ساده‌تری ثابت شد) به‌دست می‌دهد. در واقع، به‌طوری که استدلال خام زیر نشان می‌دهد، (۴.۴) از (۳.۴) نتیجه می‌شود.

چون رشد لگاریتم به‌مراتب از هر توان جبری مثبت آهسته‌تر است، $\log x=O(x^{\frac{1}{2}})$. بنابراین، به‌عنوان مثال، از (۲.۴) نتیجهٔ به‌مراتب ضعیفتر

$$T(x)-G_1(x)=O(x^{\frac{1}{2}}) \qquad (۵.۴)$$

به‌دست می‌آید. بنابر این رابطه و رابطهٔ $|\mu(k)|\le 1$، ثابت Kای وجود دارد به‌طوری که طرف راست (۳.۴) کوچکتر یا مساوی است با

$$\begin{aligned}Kx^{1/2}\sum_{k\le x}k^{-1/2}&<Kx^{1/2}\left(1+\sum_{2\le k\le x}\int_{k-1}^{k}u^{-1/2}du\right)\\&\le Kx^{1/2}\left(1+\int_1^x u^{-1/2}du\right)=O(x)\end{aligned} \qquad (۶.۴)$$

که در واقع (۴.۴) را نتیجه می‌دهد.

بنابراین، معکوس‌سازی فرمول چبیشف با استفاده از فرمول موبیوس فقط نتیجهٔ خام (۴.۴) را به‌دست می‌دهد، و دلیل تأخیر طولانی در کشف اثباتی مقدماتی از قضیهٔ عددهای اول در همین نکته نهفته است.

توجه کنید که در براورد طرف راست (۳.۴)، نتیجهٔ خام (۵.۴) به‌خوبیِ نتیجهٔ (۲.۴) که بسیار ظریفتر است، عمل می‌کند. این نکته فکر زیر را القا می‌کند.

در فرمول معکوس‌سازیِ موبیوس

$$F(x)=\sum_{k\le x}\mu(k)G(\frac{x}{k}) \qquad (۷.۴)$$

(که در اینجا $F=\psi-x+C$ و $G=T-G_1$)، تا اندازه‌ای می‌توانیم جمله‌های مجموع طرف راست را افزایش دهیم، زیرا افزایش جمله‌ها تغییری در براورد خام $O(x)$، (۶.۴)، در این طرف نمی‌دهد. از سوی دیگر، افزایش مناسب جمله‌های طرف راست ممکن است باعث شود که به‌جای $F(x)$ (و بنابراین $\psi(x)$) در طرف چپ، حاصلضرب $F(x)$ درتابعی با رشد مثبت قرار گیرد، که این، براورد $O(x)$ در طرف راست را مفید می‌سازد.

کمی آزمایش نشان می‌دهد که ساده‌ترین حالت برای محاسبهٔ صریح وقتی است که به‌جای طرف راست (۷.۴)

$$J(x)=\sum_{k\le x}\mu(k)\log\frac{x}{k}G\left(\frac{x}{k}\right) \qquad (۸.۴)$$

قرار گیرد. این باید برحسب F محاسبه شود. بنا به تعریف G،

$$\begin{aligned}J(x)&=\sum_{k\le x}\mu(k)\log\frac{x}{k}\sum_{j\le x/k}F\left(\frac{x}{jk}\right)\\&=\sum_{jk\le x}\mu(k)\log\frac{x}{k}F\left(\frac{x}{jk}\right)\\&=\sum_{n\le x}F\left(\frac{x}{n}\right)\sum_{jk=n}\mu(k)\log\frac{x}{k}\\&=\sum_{n\le x}F\left(\frac{x}{n}\right)\sum_{k|n}\mu(k)\log\frac{x}{k}\end{aligned}$$

با استفاده از $\log\frac{x}{k}=\log\frac{x}{n}+\log\frac{n}{k}$ داریم

$$\begin{aligned}J(x)=&\sum_{n\le x}F\left(\frac{x}{n}\right)\log\frac{x}{n}\sum_{k|n}\mu(k)\\&+\sum_{n\le x}F\left(\frac{x}{n}\right)\sum_{k|n}\mu(k)\log\frac{n}{k}\end{aligned}$$

بنا به (۱۵.۲) و (۲۱.۲)، نتیجه می‌شود

$$J(x)=F(x)\log x+\sum_{n\le x}F\left(\frac{x}{n}\right)\Lambda(n)$$

پس، بنا به (۸.۴)

$$F(x)\log x+\sum_{n\le x}F\left(\frac{x}{n}\right)\Lambda(n)=\sum_{k\le x}\mu(k)\log\frac{x}{k}G\left(\frac{x}{k}\right) \qquad (۹.۴)$$

و این نابرابریِ تاتوزاوا-ایزکی[1] است [۷] که به‌سادگی نابرابری اتل سلبرگ را نتیجه می‌دهد. در واقع بنا به (۲.۴)

$$\log x(T(x)-G_1(x))=O(\log^2 x)=O(x^{\frac{1}{2}})$$

و بنابراین، به‌طوری که پیشتر در (۶.۴) دیده‌ایم

$$\sum_{k\le x}\mu(k)\log\frac{x}{k}\left(T\left(\frac{x}{k}\right)-G_1\left(\frac{x}{k}\right)\right)=O(x)$$

پس (۹.۴) به‌ازای $F(x)=\psi(x)-x+C$ به‌صورت

$$(\psi(x)-x)\log x+\sum_{n\le x}\left(\psi\left(\frac{x}{n}\right)-\frac{x}{n}\right)\Lambda(n)=O(x) \qquad (۱۰.۴)$$

درمی‌آید. در اینجا از (۴.۴) برای ادغام $C\psi(x)$ و $C\log x$ در $O(x)$ استفاده شده است. رابطهٔ (۱۰.۴) صورتی از نابرابری معروف اتله سلبرگ است [۶].

1. Tatuzawa-Iseki

با استفاده از لم ۳.۳، (۱۰.۴) را می‌توان به‌صورت

$$\psi(x)\log x+\sum_{n\le x}\Lambda(n)\psi(x/n)=2x\log x+O(x) \qquad (۱۱.۴)$$

نوشت. از (۱.۳) و (۵.۳) به‌ازای $c_n=\Lambda(n)$ نتیجه می‌شود

$$\begin{aligned}\sum_{n\le x}\Lambda(n)\log n&=\psi(x)\log x-\int_1^x\frac{\psi(t)}{t}dt\\&=\psi(x)\log x+O(x) \qquad (۱۲.۴)\end{aligned}$$

همچنین

$$\sum_{j\le x}\Lambda(j)\psi\left(\frac{x}{j}\right)=\sum_{j\le x}\Lambda(j)\sum_{k\le x/j}\Lambda(k)=\sum_{jk\le x}\Lambda(j)\Lambda(k) \qquad (۱۳.۴)$$

بنابراین اگر

$$\Lambda_2(n)=\Lambda(n)\log n+\sum_{jk=n}\Lambda(j)\Lambda(k) \qquad (۱۴.۴)$$

آنگاه از کاربرد (۱۲.۴) و (۱۳.۴) در (۱۱.۴) نتیجه می‌شود $\sum_{n\le x}\Lambda_2(n)=2x\log x+O(x)$، که معادل با (۱۱.۴) است. بنا به (۴.۳)، $\sum_{n\le x}\log n=x\log x+O(x)$. با ترکیب دو نابرابری فوق به‌دست می‌آید

$$Q(n)=\sum_{k\le n}(\Lambda_2(k)-2\log k)=O(n),\quad n\ge 2,\quad Q(1)=0 \qquad (۱۵.۴)$$

۵. اثبات قضیهٔ اعداد اول. اگر $R(x)=\psi(x)-x$، $x\ge 2$، و $R(x)=0$، $x<2$، آنگاه (۱۰.۴) به‌صورت

$$R(x)\log x+\sum\Lambda(n)R\left(\frac{x}{n}\right)=O(x) \qquad (۱.۵)$$

درمی‌آید که در آن، مجموعْ متناهی است چون به ازای $n>\frac{x}{2}$ داریم $R(\frac{x}{n})=0$. هدف، یعنی (۷.۲)، به‌صورت

$$\lim_{x\to\infty}\frac{R(x)}{x}=0 \qquad (۲.۵)$$

در می‌آید. استخراج (۲.۵) از (۱.۵) پیچیده است، زیرا وزنهای $\Lambda(n)$ در مجموع وزندار (۱.۵) به مکان عددهای اول بستگی دارند و این همان چیزی است که در صدد یافتن آن هستیم. بدلیل این پیچیدگی است که تاکنون اثبات ساده‌ای برای (۲.۵) به کمک (۱.۵) به‌دست نیامده است.

در اثباتی که عرضه می‌شود، از «عملهای هموارکنندهٔ» متعددی روی (۱.۵) استفاده می‌شود تا نابرابری به‌دست آید که کارکردن با آن آسانتر باشد. بسیاری از این عملها موجب از دست رفتن اطلاعات می‌شوند، و هدف از آنها هموارسازی برای سهولت استفاده است بی‌آنکه (۱.۵) کاملاً از دست رود.

نخست به‌جای $R(x)$، هموارکنندهٔ

$$S(y)=\int_2^y \frac{R(x)}{x}dx,\qquad y\ge 2 \qquad (۳.۵)$$

و $S(y)=0$ به‌ازای $y<2$، را قرار می‌دهیم. خوشبختانه اگر بتوانیم ثابت کنیم

$$\lim_{y\to\infty}\frac{S(y)}{y}=0 \qquad (۴.۵)$$

به‌سادگی می‌توانیم (۲.۵) را نتیجه بگیریم (این کار بعداً انجام خواهد شد).

لم ۱.۵ ثابت cای وجود دارد به‌طوری که

$$|S(y)|\le cy,\qquad y\ge 2 \qquad (۵.۵)$$

و

$$|S(y_2)-S(y_1)|\le c|y_1-y_2| \qquad (۶.۵)$$

به‌علاوه، بنا به (۱.۵)،

$$S(y)\log y+\sum\Lambda(j)S(\frac{y}{j})=O(y) \qquad (۷.۵)$$

اثبات. بنا به (۵.۳)، به‌ازای xهای بزرگ، $-x\le\psi(x)-x\le\frac{x}{2}$. پس

$$\limsup_{x\to\infty}\frac{|R(x)|}{x}\le 1 \qquad (۸.۵)$$

و، چون $|R(x)|$ به‌ازای xهای متناهی کراندار است، ثابت cای باید موجود باشد به‌طوری که

$$|R(x)|\le cx,\qquad x\ge 2 \qquad (۹.۵)$$

بنا به (۳.۵)، $S'(y)=\frac{R(y)}{y}$، بجز در $y=p^j$، که در آن $R(y)$ ناپیوسته است. پس بنا به (۹.۵)

$$|S'(y)|\le c,\qquad y\ne p^j \qquad (۱۰.۵)$$

بنابراین، نخست توجه کنید که (۶.۵) در حالتی که بازهٔ $y_1<y<y_2$ فاقد p^jای باشد، برقرار است. ولی، چون $S(y)$ پیوسته است، با توجه به اینکه قدر مطلق مجموع کوچکتر یا برابر با مجموع قدر مطلقهاست، می‌توانیم (۶.۵) را به‌ازای هر y_1 و y_2ای تعمیم دهیم. شرط (۶.۵) به‌شرط لیپشیتس معروف است. نتیجهٔ (۵.۵) از (۶.۵) به‌ازای $y_1=2$ به‌دست می‌آید.

چون $||a|-|b||\le|a-b|$، بنا به (۶.۵) داریم

$$||S(y_2)|-|S(y_1)||\le c|y_2-y_1| \qquad (۱۱.۵)$$

برای اثبات (۷.۵)، (۱.۵) را به x تقسیم می‌کنیم و انتگرالگیری می‌کنیم

$$\int_2^y \frac{R(x)}{x}\log x\,dx+\sum\Lambda(n)\int_2^y R(\frac{x}{n})\frac{dx}{x}=O(y) \qquad (۱۲.۵)$$

با انتگرالگیری از جملهٔ اول به‌روش جزءجزء و بنا به (۵.۵)

$$\int_2^y\frac{R(x)}{x}\log x\,dx=\log y\,S(y)-\int_2^y\frac{S(x)}{x}dx=\log y S(y)+O(y)$$

همچنین اگر $\xi=\frac{x}{n}$

$$\int_2^y R(\frac{x}{n})\frac{dx}{x}=\int_2^{\frac{y}{n}}\frac{R(\xi)}{\xi}d\xi=S(\frac{y}{n})$$

با جایگزینی اینها در (۱۲.۵)، (۷.۵) نتیجه می‌شود.

برای قابل استفاده‌ترکردن مجموع وزندار (۷.۵)، چگالی مجموعهٔ نقطه‌هایی را که به‌ازای آنها $S(\frac{y}{j})$ در مجموع ظاهر می‌شود، با تکرار (۷.۵) افزایش می‌دهیم.

لم ۲.۵ اگر مانند (۱۴.۴)، $\Lambda_2(n)=\Lambda(n)+\sum_{ij=n}\Lambda(i)\Lambda(j)$، و K_1 عدد ثابتی باشد،

$$\log^2 y|S(y)|\le\sum\Lambda_2(m)|S(y/m)|+K_1 y\log y \qquad (۱۳.۵)$$

اثبات. در (۷.۵)، $\frac{y}{k}$ را جایگزین y می‌کنیم، در $\Lambda(k)$ ضرب می‌کنیم و روی $k\le y$ مجموعیابی می‌کنیم. در این صورت

$$\sum\Lambda(k)S\left(\frac{y}{k}\right)\log\frac{y}{k}+\sum\sum\Lambda(k)\Lambda(j)S\left(\frac{y}{jk}\right)=O(y)\sum_{k\le y}\frac{\Lambda(k)}{k}$$

با قرار دادن $jk=m$ در مجموع دوم و مجموعیابی روی m، و با ملحوظ کردن $\log\frac{y}{k}=\log y-\log k$ در مجموع اول و قرار دادن m به‌جای این kای آخری، و با استفاده از (۷.۳) داریم

$$\log y\sum\Lambda(k)S\left(\frac{y}{k}\right)-\sum_{m\le y}S\left(\frac{y}{m}\right)\left\{\Lambda(m)\log m-\sum_{jk=m}\Lambda(j)\Lambda(k)\right\}=O(y\log y)$$

پس بنا به (۷.۵)

$$S(y)\log^2 y=-\sum S\left(\frac{y}{m}\right)\left\{\Lambda(m)\log m-\sum_{jk=m}\Lambda(j)\Lambda(k)\right\}+O(y\log y)$$

که با قدر مطلق گرفتن از طرفین آن، (۱۳.۵) به‌دست می‌آید.

از نابرابری (۱۵.۴) برمی‌آید که $\Lambda_2(m)$ به‌طور متوسط مانند $2\log m$ عمل می‌کند. مجموع وزنداری با وزنهای $2\log m$ کاملاً برای عملیات مناسب

است و بنابراین، می‌توان با قرار دادن $2\log m$ به‌جای $\Lambda_2(m)$، (۱۳.۵) را اصلاح کرد.

لم ۳.۵ ثابت K_2ای وجود دارد به‌طوری که

$$\log^2 y|S(y)| \le 2\sum |S(y/m)|\log m + K_2 y\log y \qquad (۱۴.۵)$$

اثبات.

$$\sum |S(y/m)|\Lambda_2(m) = 2\sum_{m\le y}|S(y/m)|\log m + J(y) \qquad (۱۵.۵)$$

که چون بنا به (۱۵.۴)، $\Lambda_2(m) - 2\log m = Q(m) - Q(m-1)$، و چون به‌ازای $y < 2$، $S(y) = 0$، داریم

$$\begin{aligned} J(y) &= \sum_{m\le y}(Q(m)-Q(m-1))|S(y/m)| \\ &= \sum_{m\le y} Q(m)|S(y/m)| - \\ &\quad \sum_{m\le y} Q(m)|S(y/(m+1))| \\ &= \sum_{2\le m\le y} Q(m)(|S(y/m)| - |S(y/(m+1))|) \end{aligned}$$

بنا به (۱۵.۴) و (۱۱.۵)، ثابت K_3ای هست به‌طوری که

$$\begin{aligned} J(y) &\le K_3 \sum_{2\le m\le y} m\left(\frac{y}{m} - \frac{y}{m+1}\right) \\ &= K_3 y \sum_{2\le m\le y}\frac{1}{m+1} < K_3 y\int_1^y \frac{dv}{v} = K_3 y\log y \end{aligned}$$

از این و (۱۵.۵)، ثابت می‌شود که (۱۴.۵) نتیجه‌ای از (۱۳.۵) است.

تعویض مجموع در (۱۴.۵) با انتگرال، موجب سادگی بیشتر می‌شود.

لم ۴.۵ ثابت K_4ای وجود دارد به‌طوری که

$$\log^2 y|S(y)| \le 2\int_2^y |S(y/u)|\log u\, du + K_4 y\log y \qquad (۱۶.۵)$$

اثبات. چون $\log u$ صعودی است

$$\log m|S(y/m)| \le \int_m^{m+1}\log u|S(y/m)|du$$

ولی $|S(y/m)| \le |S(y/u)| + |S(y/m) - S(y/u)|$، پس

$$\log m|S(y/m)| \le \int_m^{m+1}\log u|S(y/u)|du + J_m$$

$$J_m = \int_m^{m+1}\log u|S(y/m) - S(y/u)|du \qquad (۱۷.۵)$$

با استفاده از (۶.۵)

$$J_m \le c\left(\frac{y}{m} - \frac{y}{m+1}\right)\int_m^{m+1}\log u\, du \le \frac{cy\log(m+1)}{m(m+1)}$$

چون $\log(m+1) \le m$، با منظور کردن نتیجهٔ بالا در (۱۷.۵) به‌دست می‌آوریم

$$\log m\left|S\left(\frac{y}{m}\right)\right| \le \int_m^{m+1}\log u\left|S\left(\frac{y}{u}\right)\right|dy + \frac{cy}{m+1}$$

با به‌کار بردن این در (۱۴.۵)، (۱۶.۵) به‌ازای $K_4 = K_2 + c$ به‌دست می‌آید.

با تعویض متغیری نمایی، نابرابری (۱۶.۵) به‌صورتی ساده‌تر در می‌آید. $v = \log\frac{y}{u}$ را جایگزین u می‌کنیم. همچنین قرار می‌دهیم $x = \log y$. در این صورت (۱۶.۵) به‌صورت

$$x^2|S(e^x)| \le 2\int_0^{x-\log 2}|S(e^v)|(x-v)e^{(x-v)}dv + K_4 x e^x \qquad (۱۸.۵)$$

در می‌آید. اگر

$$W(x) = e^{-x}S(e^x) \qquad (۱۹.۵)$$

آنگاه (۱۸.۵) به‌صورت زیر در می‌آید

$$|W(x)| \le \frac{2}{x^2}\int_0^x (x-v)|W(v)|dv + \frac{K_4}{x} \qquad (۲۰.۵)$$

این نابرابری حاوی اطلاع باارزشی است، زیرا در واقع می‌گوید که $|W(x)|$ از یک میانگین وزندار $|W|$ بیشتر نیست. در نتیجه، لم زیر را داریم: (توجه کنید که γ در این لم، ثابت اویلر نیست.)

لم ۵.۵ فرض می‌کنیم

$$\begin{aligned} \gamma &= \limsup_{x\to\infty}\frac{1}{x}\int_0^x |W(\xi)|d\xi \\ \alpha &= \limsup_{x\to\infty}|W(x)| \qquad (۲۱.۵) \end{aligned}$$

در این صورت $\alpha \le 1$ و

$$\alpha \le \gamma \qquad (۲۲.۵)$$

توجه. بر اساس (۴.۵) و (۱۹.۵)، هدف این است که نشان دهیم $\alpha = 0$.

اثبات. اینکه $\alpha \le 1$، از (۱۹.۵) و از این واقعیت که (۸.۵) و (۵.۳) نابرابری

$$\limsup_{y\to\infty}\frac{|S(y)|}{y} \le 1 \qquad (۲۳.۵)$$

را به‌دست می‌دهند، نتیجه می‌شود. حکم اصلی یعنی $\alpha \le \gamma$ با استفاده از (۲۰.۵) ثابت می‌شود و این تنها استفاده‌ای است که از لمهای ۲.۵، ۳.۵ و ۴.۵ می‌شود. توجه کنید که (۲۰.۵) را می‌توان به‌صورت

$$|W(x)| \le \frac{2}{x^2}\int_0^x u\,du\left(\frac{1}{u}\int_0^u |W(v)|dv\right) + \frac{K_4}{x} \qquad (۲۴.۵)$$

نوشت و این امر را با تغییر ترتیب انتگرال‌گیری می‌توان تحقیق کرد. ولی

$$\frac{2}{x^2}\int_0^x u\,du = 1$$

و بنا به (۵.۵)، انتگرال طرف راست (۲۴.۵) صرفاً میانگین وزنداری از $\frac{1}{u}\int_0^u |W(v)|dv = \left(\frac{1}{u}\right)\int_0^u e^{-v}|S(e^v)|dv \le c$ است. پس به‌ازای هر x_1 ثابت و $x > x_1$

$$\begin{aligned} I(x) &= \frac{2}{x^2}\int_0^x u\,du\left(\frac{1}{u}\int_0^u |W(v)|dv\right) \\ &\le \frac{2c}{x^2}\int_0^{x_1} u\,du + \frac{2}{x^2}\int_{x_1}^x u\,du\left(\frac{1}{u}\int_0^u |W(v)|dv\right) \end{aligned} \qquad (۲۵.۵)$$

فرض می‌کنیم $\epsilon > 0$. بنا به تعریف γ و به‌ازای x_1 به اندازهٔ کافی بزرگ

$$\frac{1}{u}\int_0^u |W(v)|dv < \gamma + \epsilon, \qquad u \ge x_1$$

پس، بنا به (۲۵.۵)

$$I(x) \le \frac{cx_1^2}{x^2} + (\gamma+\epsilon)\left(1 - \frac{x_1^2}{x^2}\right)$$

بنابراین به‌ازای xهای بزرگ، از (۲۴.۵) نتیجه می‌شود

$$|W(x)| \le \gamma + \epsilon + \frac{cx_1^2}{x^2} + \frac{K_4}{x}$$

اگر $x \to \infty$، $\alpha \le \gamma + \epsilon$، و چون این به‌ازای هر $\epsilon > 0$ برقرار است، (۲۲.۵) نتیجه می‌شود.

برای اثبات $\alpha = 0$ دو ویژگی دیگر W مورد نیاز است.

لم ۶.۵ اگر $k = 2c$، آنگاه

$$|W(x_2) - W(x_1)| \le k|x_2 - x_1| \qquad (۲۶.۵)$$

و بنابراین

$$||W(x_2)| - |W(x_1)|| \le k|x_2 - x_1| \qquad (۲۷.۵)$$

اثبات. چون $W(x) = e^{-x}S(e^x)$، به‌ازای $x \ne j\log p$،

$$|W'(x)| \le e^{-x}|S(e^x)| + |S'(e^x)|$$

پس، اگر $x \ne j\log p$، بنا به (۵.۵) و (۱۰.۵)، $|W'(x)| \le 2c = k$. از این رابطه به (۲۶.۵) می‌رسیم همان‌طور که از (۱۰.۵) به (۶.۵) رسیدیم.

لم ۷.۵ اگر به ازای $v_1 < v < v_2$، $W(v) \ne 0$، آنگاه عدد Mای وجود دارد به‌طوری که

$$\int_{v_1}^{v_2} |w(v)|dv \le M \qquad (۲۸.۵)$$

اثبات. با توجه به (۱.۳) قرار می‌دهیم $c_n = \Lambda(n)$ و $f(n) = \frac{1}{n}$. در این صورت، بنا به (۷.۳)

$$\int_2^x \frac{\psi(t)}{t^2}dt = \log x + O(1)$$

یا، چون $R(t) = \psi(t) - t$

$$\int_2^x \frac{R(t)}{t^2}dt = O(1) \qquad (۲۹.۵)$$

ولی

$$\begin{aligned} \int_2^x \frac{S(y)}{y^2}dy &= \int_2^x \frac{dy}{y^2}\int_2^y \frac{R(t)}{t}dt = \int_2^x \frac{R(t)}{t}\left(\int_t^x \frac{dy}{y^2}\right)dt \\ &= \int_2^x \frac{R(t)}{t^2}dt - \frac{1}{x}\int_2^x \frac{R(t)}{t}dt \end{aligned}$$

با استفاده از (۲۹.۵) و (۹.۵)، $\int_2^x \frac{S(y)}{y^2}dy = O(1)$، یا با قراردادن $y = e^u$، $x = e^v$ داریم

$$\int_{\log 2}^v W(u)du = O(1)$$

اگر این را بر حسب $v = v_1$ و $v = v_2$ بنویسیم و از هم کم کنیم، نتیجه می‌شود که انتگرال حاصل کراندار است و بنابراین ثابت Mای وجود دارد که

$$\left|\int_{v_1}^{v_2} W(u)du\right| \le M$$

ولی اگر به ازای $v_1 < u < v_2$، $W(u) \ne 0$، این را می‌توان به‌صورت (۲۸.۵) نوشت. چون M را در صورت اقتضا می‌توان بزرگتر کرد، می‌توانیم فرض کنیم $Mk > 1$.

لم ۸.۵ در مورد تابع $W(x)$ای که در سه شرط (۲۲.۵)، (۲۷.۵)، و (۲۸.۵) صدق می‌کند، باید $\alpha = 0$.

اثبات. انتخاب می‌کنیم $\beta > \alpha$. بنا به تعریف α، x_βای وجود دارد به‌طوری که

$$|W(x)| \le \beta, \quad x \ge x_\beta \qquad (۳۰.۵)$$

اگر به ازای هر x بزرگ، $W(x) \ne 0$، از (۲۸.۵) نتیجه می‌شود $\gamma = 0$ و بنابراین $\alpha = 0$. پس فرض می‌کنیم $W(x)$ دارای صفرهای به‌دلخواه بزرگ باشد. تصور می‌کنیم a و b دو صفر متوالی $W(x)$ به ازای $x > x_\beta$ هستند.

حالت ۱. $b-a \ge \frac{2M}{\beta}$. بنا به (۲۸.۵)، چون $W(x) \neq 0$، $a<x<b$،

$$\int_a^b |W(x)| \le M \le \frac{b-a}{2}\beta$$

(پس میانگین $|W|$ بر (a,b) کمتر از $\frac{1}{2}\beta$ است.)

حالت ۲. $b-a \le \frac{2\beta}{k}$. در این صورت بنا به (۲۷.۵)، اگر نمودار $|W(x)|$ با حداکثر سرعت ممکن از $x=a$ به راست و از $x=b$ به چپ برود، نمی‌تواند بالای مثلثی به ارتفاع $k\frac{(b-a)}{2} \le \beta$ قرار بگیرد و بنابراین

$$\int_a^b |W(x)|dx \le \frac{1}{2}(b-a)\beta$$

حالت ۳. $\frac{2\beta}{k} < b-a < \frac{2M}{\beta}$. با استدلالی مشابه حالت ۲ برای فاصله‌ای چون $\frac{\beta}{k}$ از هر نقطهٔ انتهایی و با استفاده از (۳۰.۵) در موارد دیگر

$$\begin{aligned}\int_a^b |W(x)| &\le \frac{\beta^2}{k} + \left(b-a-\frac{2\beta}{k}\right)\beta \\ &= (b-a)\beta\left(1-\frac{\beta}{k(b-a)}\right) \\ &\le (b-a)\beta\left(1-\frac{\beta^2}{2Mk}\right) \\ &< (b-a)\beta\left(1-\frac{\alpha^2}{2Mk}\right) \qquad (۳۱.۵)\end{aligned}$$

چون $Mk \ge 1$، و چون $\alpha \le 1$، (۳۱.۵) در حالتهای ۱ و ۲ نیز برقرار است. اگر x_1 نخستین صفر $W(x)$ در طرف راست x_β و $\tilde{x}$ بزرگترین صفر $W(x)$ در طرف چپ y باشد، آنگاه از (۲۸.۵) و (۳۱.۵) نتیجه می‌شود که

$$\int_0^y |W(x)|dx \le \int_0^{x_1} |W(x)|dx + (\tilde{x}-x_1)\beta\left(1-\frac{\alpha^2}{2Mk}\right) + M$$

با تقسیم بر y و توجه به اینکه $\tilde{x} \le y$

$$\frac{1}{y}\int_0^y |W(x)|dx \le \frac{1}{y}\int_0^{x_1} |W(x)|dx + \beta\left(1-\frac{\alpha^2}{2Mk}\right) + \frac{M}{y}$$

اگر $y \to \infty$، $\gamma \le \beta(1-\alpha^2/2Mk)$، و چون $\gamma \ge \alpha$

$$\alpha \le \beta\left(1-\frac{\alpha^2}{2Mk}\right)$$

چون این به‌ازای هر $\beta > \alpha$ برقرار است، باید به‌ازای $\beta = \alpha$ نیز برقرار باشد. پس $\alpha^3 \le 0$، و چون $\alpha \ge 0$، نتیجه می‌شود $\alpha = 0$. از آنجا که $W(x) = e^{-x}S(e^x)$، نتیجه می‌شود

$$\lim_{y\to\infty} \frac{|S(y)|}{y} \to 0$$

پس، اگر $\epsilon > 0$ مفروض باشد، به‌ازای y به‌اندازهٔ کافی بزرگ

$$|S(y)| \le \frac{1}{3}\epsilon^2 y$$

بنابراین $S(y(1+\epsilon)) - S(y) \le \frac{1}{3}\epsilon^2(y(1+\epsilon)+y) < \epsilon^2 y$ ، یا

$$\int_y^{y(1+\epsilon)} \frac{R(u)}{u}du \le \epsilon^2 y$$

چون $R(u) = \psi(u) - u$ و ψ غیر نزولی است،

$$\frac{\psi(y)}{y(1+\epsilon)}\int_y^{y(1+\epsilon)} du - \int_y^{y(1+\epsilon)} du \le \epsilon^2 y$$

پس $\frac{\psi(y)}{y} \le (1+\epsilon)^2$. به همین نحو، به‌ازای yهای به‌اندازهٔ کافی بزرگ، از $S(y) - S(y(1-\epsilon)) \ge -\epsilon^2 y$ نتیجه می‌شود $\frac{\psi(y)}{y} \ge (1-\epsilon)^2$. پس چون ϵ دلخواه است، (۷.۲) ثابت می‌شود.

مراجع

1. E. Bombieri, Sulle formule di A. Selberg generalizzate per classi di funzioni aritmetiche e le applicazioni al problema del resto nel "Primzahlsatz", Riv. Mat. Univ. Parma, (2) 3 (1962) 393-440.

2. R. Breusch, An elementary proof of the prime number theorem with remainder term, Pacific J. Math., 10(1960) 487-497.

3. P. Erdös, On a new method in elementary number theory which leads to an elementary proof of the prime number theorem, Proc. Nat. Acad. Sci. U.S.A., 35 (1949) 374-384.

4. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Oxford University Press, New York, 1960.

5. N. Levinson, On the elementary proof of the prime number theorem, Proc. Edinburgh Math. Soc., (2) 15 (1966) 141-146.

6. A. Selberg, An elementary proof of the prime number theorem, Ann. of Math., (2) 50 (1949) 305-313.

7. T. Tatuzawa and K. Iseki, On Selberg's elementary proof of the prime number theorem, Proc. Japan Acad., 27 (1951) 340-342.

8. E. Wirsing, Elementare Beweise des Primzahlsatzes mit Restglied, I, J. Reine Angew. Math., 211 (1962) 205-214.

9. E. M. Wright, The elementary proof of the prime number theorem, Proc. Roy. Soc. Edinburgh. Sect. A., 63 (1951) 257-267.

******

- Norman Levinson, "A motivated account of an elementary proof of the prime number theorem", *Amer. Math. Monthly*, (3) **۷۶** (1969) 225-245.