Der Primzahlsatz

Zusammenfassung

Im Jahr 1896 wurde von Hadamard und de la Vallée Poussin der Primzahlsatz bewiesen: Die Anzahl der Primzahlen kleiner gleich x verhält sich asymptotisch wie $x/\log x$. Für ihren Beweis war allerdings die Verwendung von komplex-analytischen Hilfsmittel wesentlich, und es war eine Sensation als 1948 ein elementarer Beweis von Erdős und Selberg angekündigt wurde (siehe [2] für einen historischen Abriß). Hier stellen wir eine Variante dieses elementaren Beweises von Hardy und Wright vor [1].

Wir möchten den Primzahlsatz beweisen, d.h. die Aussage

$$\lim_{x\to\infty} \pi(x) \Big/ \frac{x}{\log x} = 1,$$

wobei $\pi(x)$ die Anzahl der Primzahlen $p \leq x$ ist. Im folgenden bezeichnen m und n immer natürliche Zahlen, p ist immer eine Primzahl während x und t reelle Zahlen sind.

1 Bemerkung (asymptotische Notation). Seien f und g Funktionen von $\mathbb{N}$, $\mathbb{Z}$ oder $\mathbb{R}$ nach $\mathbb{R}$. Falls $f(x)$ für alle x positiv ist, vereinbaren wir die folgenden Kurzschreibweisen.

- $g = O(f)$ falls es eine positive Konstante A gibt, so daß $|g(x)| \leq Af(x)$ für alle genügend großen x. Die Gleichung $f(x) = g(x) + O(h(x))$ ist zu lesen als: „Es gibt eine Konstante A, so daß $|f(x) - g(x)| \leq Ah(x)$ für alle genügend großen x“.
- $g \asymp f$ falls es positive Konstanten A und B gibt, so daß $Af(x) \leq g(x) \leq Bf(x)$ für alle genügend großen x.
- $g \sim f$ falls $\lim\limits_{x\to\infty} \frac{g(x)}{f(x)} = 1$.

In dieser Notation ist unser Ziel also $\pi(x) \sim \frac{x}{\log x}$.

2 Bemerkung. Die beste bisher bewiesene Abschätzung ist

$$\pi(x) = \mathrm{Li}(x) + O\left(x \exp\left(-\frac{A(\log x)^{3/5}}{(\log\log x)^{1/5}}\right)\right)$$

mit $\mathrm{Li}(x) = \int\limits_2^x \frac{\mathrm{d}t}{\log t}$. Äquivalent zur Riemannschen Vermutung ist die schönere Abschätzung

$$\pi(x) = \mathrm{Li}(x) + O(\sqrt{x}\log x),$$

oder genauer

$$|\pi(x) - \mathrm{Li}(x)| < \frac{\sqrt{x}\log x}{8\pi}.$$

Der elementare Beweis

Wie weit Tschebyschow gekommen ist

Um das das asymptotische Verhalten von $\pi(x)$ zu analysieren betrachtete Tschebyschow die folgenden zahlentheoretischen Funktionen.

3 Definition. Die Funktion $\Lambda : \mathbb{N} \to \mathbb{R}$ ist definiert durch

$$\Lambda(n) = \begin{cases} \log p & \text{falls } n = p^m, \\ 0 & \text{sonst.} \end{cases}$$

Die Tschebyschow-Funktionen sind

$$\vartheta(x) = \sum_{p \le x} \log p = \log \prod_{p \le x} p, \qquad \psi(x) = \sum_{n \le x} \Lambda(n) = \sum_{p^m \le x} \log p. \tag{1}$$

Als erstes zeigen wir, daß ϑ und ψ die gleiche Größenordnung haben.

4 Satz. $\psi(x) = \vartheta(x) + O(\sqrt{x}(\log x)^2)$

Beweis. Es ist

$$\psi(x) = \vartheta(x) + \vartheta\left(x^{1/2}\right) + \vartheta\left(x^{1/3}\right) + \cdots = \sum_{m \ge 1} \vartheta\left(x^{1/m}\right)$$

wobei höchstens $\lceil \log x / \log 2 \rceil$ Summanden von 0 verschieden sind. Jetzt schätzen wir die Summanden für $m \ge 2$ ganz grob ab:

$$\vartheta(x^{1/m}) \le x^{1/m} \log x \le x^{1/2} \log x \qquad \text{für } m \ge 2,$$

und damit folgt schon die Behauptung. ∎

Und zwar wachsen beide linear.

5 Satz. $\vartheta(x) \asymp x$ und $\psi(x) \asymp x$.

Beweis. Wir zeigen die Abschätzungen $\vartheta(n) < 2n \log 2$ und $\psi(x) \ge \frac{1}{4} x \log 2$. Nach Satz 4 reicht das aus.

- Für $M = \binom{2m+1}{m}$ haben wir $M < 2^m$ und $\left(\prod_{m+1 < p \le 2m+1} p \right) \mid M$ und daraus schließen wir

  $$\vartheta(2m+1) - \vartheta(m+1) = \sum_{m+1 < p \le 2m+1} \log p \le \log M < 2m \log 2.$$

  Jetzt machen wir Induktion. Der Anfang $n = 1$ und $n = 2$ ist klar, für gerades $n > 2$ ist

  $$\vartheta(n) = \vartheta(n-1) < 2(n-1) \log 2 < 2n \log 2$$

  und für ungerades $n = 2m + 1 > 2$

  $$\vartheta(2m+1) < \theta(m+1) + 2m \log 2 < 2(m+1) \log 2 + 2m \log 2 = 2n \log 2.$$

- Wir schreiben

$$N := \binom{2n}{n} = \prod_{p \le 2n} p^{k_p} \qquad \text{mit } k_p = \sum_{m \ge 1} \left(\left\lfloor \frac{2n}{p^m} \right\rfloor - 2 \left\lfloor \frac{n}{p^m} \right\rfloor \right).$$

In der Summe für k_p ist jeder Summand 0 oder 1 und für $m > \log(2n)/\log p$ passiert nichts mehr, d.h. $k_p \le \lfloor \log(2n)/\log p \rfloor$ und damit

$$\log N = \sum_p k_p \log p \le \sum_p \left\lfloor \frac{\log(2n)}{\log p} \right\rfloor \log p = \psi(2n).$$

Andrerseits ist aber $N \ge 2n$ und damit folgt $\psi(2n) \ge n \log 2$ und die behauptete schwächere Abschätzung für reelle x ist leicht. ∎

6 Satz (Tschebyschow). $\pi(x) \asymp \frac{x}{\log x}$.

Beweis. Wegen $\vartheta(x) = \sum_{p \le x} \log p \le \pi(x) \log x$ folgt mit Satz 5, daß $\pi(x) \ge Ax/\log x$ für eine Konstante A. Um die Abschätzung in die andere Richtung hinzukriegen müssen wir die kleinen Primzahlen ignorieren. Für $0 < \delta < 1$ ist

$$\vartheta(x) \ge \sum_{x^{1-\delta} < p \le x} \log p \ge (1-\delta) \log x (\pi(x) - x^{1-\delta}),$$

und damit $\pi(x) < Bx/\log x$. ∎

Indem wir uns mit dem δ im letzten Beweis ein bißchen mehr Mühe geben, kriegen wir die asymptotische Beziehung zwischen $\pi(x)$ und den Tschebyschow-Funktionen.

7 Satz. $\pi(x) \sim \frac{\vartheta(x)}{\log x} \sim \frac{\psi(x)}{\log x}$.

Beweis. Wegen Satz 5 reicht es, das erste $\sim$ zu beweisen. Wie im Beweis von Satz 6 ist

$$1 \le \frac{\pi(x) \log x}{\vartheta(x)} \le \frac{x^{1-\delta} \log x}{\vartheta(x)} + \frac{1}{1-\delta}.$$

Für ein beliebiges $\varepsilon > 0$ können wir nun zuerst ein $\delta > 0$ mit $1/(1-\delta) < 1 + \varepsilon/2$ wählen und dann ein x_0, so daß für alle $x > x_0$

$$\frac{x^{1-\delta} \log x}{\vartheta(x)} < \varepsilon/2.$$

Und damit folgt $\pi(x) \log x/\vartheta(x) \to 1$. ∎

8 Satz. Sei $c_1, c_2, \ldots$ eine Folge und sei die Funktion C definiert durch $C(x) = \sum_{n \le x} c_n$. Dann gilt für jede Funktion f

$$\sum_{n \le x} c_n f(n) = \sum_{n \le x-1} C(n) \left[f(n) - f(n+1) \right] + C(x) f(\lfloor x \rfloor). \tag{2}$$

Falls außerdem $c_n = 0$ für $n < n'$ und $f(x)$ für $x \ge n'$ stetig differenzierbar ist, so gilt

$$\sum_{n \le x} c_n f(n) = C(x) f(x) - \int_{n'}^{x} C(t) f'(t) \, \mathrm{d}t. \tag{3}$$

Beweis. Es ist

$$\sum_{n\le x} c_n f(n) = C(1)f(1) + (C(2) - C(1))f(2) + \cdots + (C(\lfloor x\rfloor) - C(\lfloor x\rfloor - 1))\, f(\lfloor x\rfloor)$$

und damit steht (2) da. Der zweite Teil folgt mit

$$C(n)\,[f(n) - f(n+1)] = -\int_n^{n+1} C(x)f'(x)\,\mathrm{d}x.$$

■

9 Lemma. Für $h > 0$ ist $\sum\limits_{n\le x} \log^h(x/n) = O(x)$.

Beweis. $\sum\limits_{n=2}^{\lfloor x\rfloor} \log^h(x/n) \le \int\limits_1^x \log^h(x/t)\,\mathrm{d}t = x\int\limits_1^x \frac{\log^h(u)}{u^2}\,\mathrm{d}u = O(x).$ ■

10 Satz.

$$\sum_{n\le x} \frac{\Lambda(n)}{n} = \log x + O(1). \tag{4}$$

Beweis. Mit $h = 1$ folgt aus Lemma 9, daß $\sum\limits_{n\le x} \log n = x\log x + O(x)$. Andrerseits ist

$$\sum_{n\le x} \log n = \log(\lfloor x\rfloor!) = \sum_{p^m\le x} \lfloor x/p^m\rfloor \log p = \sum_{n\le x} \frac{x}{n}\Lambda(n).$$

Mit $\sum\limits_{n\le x} \Lambda(n) = \psi(x) = O(x)$ folgt die Behauptung. ■

Und schließlich kriegen wir, daß 1 der einzige Kandidat für den Grenzwert $\lim\limits_{x\to\infty} \pi(x)\log x/x$ ist.

11 Satz.

$$\liminf_{x\to\infty} \frac{\pi(x)\log x}{x} \le 1 \le \limsup_{x\to\infty} \frac{\pi(x)\log x}{x}.$$

Beweis. Wegen Satz 7 reicht es, $\liminf \psi(x)/x \le 1 \le \limsup \psi(x)/x$ zu zeigen. Gleichung (3) liefert mit $f(t) = 1/t$ und $c_n = \Lambda(n)$

$$\sum_{n\le x} \frac{\Lambda(n)}{n} = \frac{\psi(x)}{x} + \int\limits_2^x \frac{\psi(t)}{t^2}\,\mathrm{d}t.$$

Mit den Sätzen 10 und 5 folgt

$$\int\limits_2^x \frac{\psi(t)}{t^2}\,\mathrm{d}t = \log x + O(1).$$

Angenommen $\liminf \psi(x)/x = 1 + \delta$ für ein positives δ. Dann gibt's ein x_0, so daß $\psi(x)/x \ge 1 + \delta/2$ für alle $x \ge x_0$, und es folgt der Widerspruch

$$\int\limits_2^x \frac{\psi(t)}{t^2}\,\mathrm{d}t > \int\limits_{x_0}^x \frac{1+\frac{\delta}{2}}{t^2}\,\mathrm{d}t = \left(1 + \frac{\delta}{2}\right)\log x + O(1).$$

Der lim sup-Teil geht analog. ■

Selberg und Erdős

Die Grundlage zur Vervollständigung des Beweises ist der folgende Satz.

12 Satz (Selberg-Formeln).

$$\psi(x)\log x + \sum_{n\le x}\Lambda(n)\psi(x/n) = 2x\log x + O(x), \tag{5}$$

$$\sum_{n\le x}\Lambda(n)\log n + \sum_{mn\le x}\Lambda(m)\Lambda(n) = 2x\log x + O(x). \tag{6}$$

Die beiden Gleichungen sind äquivalent, denn

$$\sum_{n\le x}\Lambda(n)\psi(x/n) = \sum_{n\le x}\Lambda(n)\sum_{m\le x/b}\Lambda(n)\Lambda(n) = \sum_{mn\le x}\Lambda(m)\Lambda(n),$$

und mit $c_n = \Lambda(n)$, $f(t) = \log t$ in (3) ist

$$\sum_{n\le x}\Lambda(n)\log n = \psi(x)\log x - \int_2^x \frac{\psi(t)}{t}\,\mathrm{d}t = \psi(x)\log x + O(x).$$

Bevor wir Satz 12 beweisen, müssen wir ein paar Werkzeuge zusammensammeln.

13 Definition. Die Möbiusfunktion $\mu : \mathbb{N} \to \{0, \pm 1\}$ ist definiert durch $\mu(1) = 1$ und

$$\mu(n) = \begin{cases} (-1)^k & \text{falls } n = p_1 \dots p_k \ (p_i \text{ prim, } p_i \neq p_j \text{ für } i \neq j), \\ 0 & \text{falls } p^2 \mid n \text{ für eine Primzahl } p. \end{cases}$$

14 Lemma. Es ist $\sum_{d\mid n}\mu(d) = \begin{cases} 1 & \text{für } n = 1, \\ 0 & \text{für } n > 1. \end{cases}$

15 Satz (Möbius-Inversion). Für Funktionen $f, g : \mathbb{N} \to \mathbb{R}$ gilt

$$f(n) = \sum_{d\mid n} g(d) \iff g(n) = \sum_{d\mid n}\mu(d)g(n/d).$$

Beweis. Die Dirichlet-Multiplikation für zahlentheoretische Funktionen ist definiert durch

$$(f * g)(n) = \sum_{d\mid n} f(d)g(n/d).$$

Man rechnet leicht nach, daß diese Operation assoziativ und kommutativ ist, und daß die Funktion I mit $I(1) = 1$ und $I(n) = 0$ für $n > 1$ ein neutrales Element ist, d.h. $f * I = I * f = f$ für alle f. Für die Funktion u mit $u(n) = 1$ für alle n sagt Lemma 14, daß $\mu * u = I$. Nun gilt

$$f = u * g \iff \mu * f = \mu * u * g = I * g = g,$$

und das ist genau die Behauptung. ∎

Aus (3) mit $c_n = 1$ und $f(x) = 1/x$ folgt eine asymptotische Formel für die harmonische Reihe.

16 Lemma. $\sum_{n\le x}\frac{1}{n}=\log x+\gamma+O(1/x)$ mit der Eulerschen Konstante $\gamma=1-\int\limits_1^\infty\frac{t-\lfloor t\rfloor}{t^2}\,\mathrm{d}t$.

Beweis (von Satz 12). Für $n=p_1^{\alpha_1}p_2^{\alpha_2}\dots p_k^{\alpha_k}$ gilt

$$\log n=\sum_{i=1}^k\alpha_i\log p_i=\sum_{d\mid n}\Lambda(d). \tag{7}$$

Möbius-Inversion liefert

$$\Lambda(n)=\sum_{d\mid n}\mu(d)\log(n/d)=-\sum_{d\mid n}\mu(d)\log d. \tag{8}$$

Das benutzen wir um für die Funktion

$$S(x)=\sum_{n\le x}\sum_{d\mid n}\mu(d)\log^2(x/d)$$

nachzurechnen, daß

$$S(x)=\sum_{n\le x}\Lambda(n)\log n+\sum_{mn\le x}\Lambda(m)\Lambda(n)+O(x)$$

gilt. $S(x)$ ist also im wesentlichen die linke Seite von (6) und es reicht, $S(x)=2x\log x+O(x)$ zu zeigen. Das geht indem man $S(x)-\gamma^2$ hinschreibt, und unter mehrfacher Verwendung von Lemma 16 eine Weile umformt. ∎

Wir betrachten nun die Funktionen

$$R(x)=\psi(x)-x,\qquad V(\xi)=e^{-\xi}R(e^\xi)=e^{-\xi}\psi(e^\xi)-1.$$

17 Bemerkung. Primzahlsatz $\iff$ $\psi(x)\sim x$ $\iff$ $R(x)=o(x)$ $\iff$ $\lim\limits_{\xi\to\infty}V(\xi)=0$.

Um die letzte dieser äquivalenten Aussagen anzusteuern, schätzen wir $|V(\xi)|$ mit Hilfe von Satz 12 ab.

18 Lemma.

$$\xi^2|V(\xi)|\le 2\int\limits_0^\xi\int\limits_0^\zeta|V(\eta)|\,\mathrm{d}\eta\,\mathrm{d}\zeta+O(\xi).$$

Beweis. Wenn wir $\psi(x)=R(x)+x$ in (5) einsetzen erhalten wir mit Lemma 10

$$R(x)\log x+\sum_{n\le x}\Lambda(n)R(x/n)=O(x), \tag{9}$$

und dementsprechend auch

$$R(x/n)\log(x/n)+\sum_{m\le x/n}\Lambda(m)R\left(\frac{x}{mn}\right)=O(x/n), \tag{10}$$

Wir multiplizieren (9) mit $\log x$ und (10) mit $\Lambda(n)$ (für jedes n), subtrahieren das voneinander und erhalten nach ein bißchen Umstellen

$$R(x)\log^2 x = -\sum_{n\le x}\Lambda(n)R(x/n)\log n + \sum_{mn\le x}\Lambda(m)\Lambda(n)R\left(\frac{x}{mn}\right)$$

und damit die Abschätzung

$$|R(x)|\log^2 x \le \sum_{n\le x} a_n|R(x/n)| + O(x\log x)$$

mit $a_n = \Lambda(n)\log n + \sum_{hk=n}\Lambda(h)\Lambda(k)$. Wegen (6) ist $\sum_{n\le x} a_n = 2x\log x + O(x)$ und durch ein paar Umformungen kann man die Summe auf der rechten Seite der Abschätzung durch ein Integral ersetzen:

$$\sum_{n\le x} a_n|R(x/n)| = 2\int_1^x |R(x/t)|\log t\,\mathrm{d}t + O(x\log x).$$

Mit den Substitutionen $x = e^\xi$ und $t = xe^{-\eta}$ wird das Integral auf der rechten Seite zu $x\int_0^\xi\int_0^\zeta |V(\eta)|\,\mathrm{d}\eta\,\mathrm{d}\zeta$ und damit sind wir fertig. ■

Um zu beweisen, daß $V(\xi)$ gegen 0 geht, brauchen wir noch zwei Abschätzungen.

19 Lemma. Es gibt eine Konstante A_1, so daß für alle $\left|\int_{\xi_1}^{\xi_2} V(\eta)\,\mathrm{d}\eta\right| \le A_1$.

20 Lemma. Sei η_0 eine Nullstelle von V, und sei $\alpha > 0$. Dann ist

$$\int_{\eta_0}^{\eta_0+\alpha} |V(\xi)|\,\mathrm{d}\xi \le \alpha^2/2 + O(1/\eta_0).$$

21 Satz. $\lim\limits_{\xi\to\infty} |V(\xi)| = 0$.

Beweis. Wir setzen $\alpha = \limsup\limits_{\xi\to\infty}|V(\xi)|$ und $\beta = \limsup\limits_{\xi\to\infty}\frac{1}{\xi}\int_0^\xi |V(\eta)|\,\mathrm{d}\eta$. Aus Lemma 18 folgt mit $\int_0^\zeta |V(\eta)|\,\mathrm{d}\eta = \beta\zeta + o(\zeta)$, daß $\alpha \le \beta$. Wir nehmen nun $\alpha > 0$ an und leiten daraus einen Widerspruch ab. Die Funktion $V(\eta)$ ist monoton fallend außer an den Unstetigkeitsstellen. Auf einem gegebenen Intervall hat V daher eine Nullstelle oder höchstens einen Vorzeichenwechsel. Wir betrachten nun das Intervall $I = [\zeta, \zeta+\delta]$ mit einem $\delta > \alpha$, für das wir gleich einen cleveren konkreten Wert wählen.

Fall 1. $V(\eta_0) = 0$ für ein $\eta_0 \in I$. Wir zerlegen das Integral $\int_\zeta^{\zeta+\delta} = \int_\zeta^{\eta_0} + \int_{\eta_0}^{\eta_0+\alpha} + \int_{\eta_0+\alpha}^{\zeta+\delta}$. Auf dem ersten und dem dritten Teilintervall schätzen wir $|V(\xi)|$ durch $\alpha + o(1)$ an, in der Mitte benutzen wir Lemma 20, und schließlich erhalten wir

$$\int_\zeta^{\zeta+\delta} |V(\eta)|\,\mathrm{d}\eta \le \alpha'\delta + o(1) \qquad (\text{für } \zeta\to\infty) \tag{11}$$

mit $\alpha' = \alpha\left(1 - \frac{\alpha}{2\delta}\right) < \alpha$.

Fall 2. $V(\eta)$ wechselt auf I höchstens einmal das Vorzeichen. Dann können wir mit Lemma 19 das Integral von ζ bis $\zeta+\delta-\alpha$ durch $2A_1$ abschätzen. Auf $[\zeta+\delta-\alpha, \zeta+\delta]$ ist ja $|V(\xi)|$ durch $\alpha + o(1)$ beschränkt, und somit

$$\int\limits_{\zeta}^{\zeta+\delta} |V(\eta)| \,\mathrm{d}\eta \leq 2A_1 + \alpha^2 + o(1).$$

Wir setzen jetzt $\alpha'\delta$ mit $2A_1 + \alpha^2$ gleich, stellen nach δ um und erhalten $\delta = (3\alpha^2 + 4A_1)/(2\alpha) > \alpha$. Für dieses δ gilt dann also in jedem Fall die Ungleichung (11) und jetzt zerhacken wir einfach das Integral von 0 bis ξ in lauter Stücke der Länge δ:

$$\int\limits_0^{\xi} |V(\eta)| \,\mathrm{d}\eta = \sum_{m=0}^{\lfloor \xi/\delta \rfloor - 1} \int\limits_{m\delta}^{(m+1)\delta} |V(\eta)| \,\mathrm{d}\eta + \int\limits_{\lfloor \xi/\delta \rfloor \delta}^{\xi} |V(\eta)| \,\mathrm{d}\eta \leq \left\lfloor \frac{\xi}{\delta} \right\rfloor \alpha'\delta + o\left(\left\lfloor \frac{\xi}{\delta} \right\rfloor \right) + O(1)$$

und daraus folgt der benötigte Widerspruch $\beta = \limsup_{\xi\to\infty} \frac{1}{\xi} \int\limits_0^{\xi} |V(\eta)| \,\mathrm{d}\eta \leq \alpha' < \alpha$. ∎

Literatur

[1] G.H. Hardy & E.M. Wright, *An Introduction to the Theory of Numbers*, 5th Ed., Oxford University Press, 1979.

[2] D. Goldfeld, The elementary proof of the prime number theorem: an historical perspective. `http://www.math.columbia.edu/~goldfeld/ErdosSelbergDispute.pdf`