Seminar der WE AlZAGK

WS 2010/11

Glatte Zahlen I

von Michelle Sinning

Mail: michellesinning@googlemail.com

Inhaltsverzeichnis

1 Glatte Zahlen

Glatte Zahlen sind Zahlen, deren Primfaktorzerlegung nur kleine Zahlen enthält. Diese werden für viele verschiedene Algorithmen benötigt, deren Geschwindigkeit von der Anzahl an glatten Zahlen in einem bestimmten Intervall bestimmt wird. Für diesen Anteil an glatten Zahlen in einem Intervall gibt es viele Schätzungen, von denen ich ein paar im Folgenden ausführlicher behandeln werde.
Insgesamt lehne ich mich zu großen Teilen an die Werke [1] von A. Granville und [2] von C. Pomerance an.

Definition 1 Für $y \neq 0$ in $\mathbb{N}$ heißt eine von Null verschiedene natürliche Zahl *y-glatt*, wenn sie keine Primteiler größer y hat.
Man schreibt:

$$S(x,y) = \text{Menge der y-glatten Zahlen} \leq x \tag{1}$$
$$\Psi(x,y) = \#S(x,y) \tag{2}$$

Beispiele

1. Die Primfaktorzerlegung der Zahl 720 ist: $720 = 2^4 \cdot 3^2 \cdot 5$. 720 ist also weder 1-, 2-, 3- noch 4-glatt, aber 5-glatt.
2. Die Anzahl der 5-glatten Zahlen ≤ 10 ist 9, denn nur 7 hat einen Primfaktor > 5. Also

$$\Psi(10,5) = 9$$

1.1 Die Dickman Funktion $\rho(u)$

Man interessiert sich nun für die relative Häufigkeit der y-glatten Zahlen unter den Zahlen $\leq x$, das heißt für den Quotienten

$$\frac{\Psi(x,y)}{x}. \tag{3}$$

Man setzt $x = y^u$, das heißt $u = \frac{\log x}{\log y}$.
Es ist klar, dass bei $0 \leq u \leq 1$ die Zahlen $\leq y^u = x$ keine Primteiler größer y haben, das heißt in diesem u-Intervall gilt:

$$\frac{\Psi(x,y)}{x} = 1 \text{ für } x = y^u \quad \text{ mit } 0 \leq u \leq 1 \tag{4}$$

Nach K. Dickman hat man bei $u \geq 1$ eine asymptotische Aussage:

$$\Psi(x,y) \sim x\rho(u), \text{ wenn } x \to \infty, \text{ wobei } x = y^u. \tag{5}$$

Hierbei ist ρ eine differenzierbare Funktion, die einer Differentialgleichung mit Zeitverzögerung genügt:

$$u\rho'(u) + \rho(u-1) = 0, \tag{6}$$

äquivalent:

$$\rho(u) = \frac{1}{u} \int_{u-1}^{u} \rho(t)dt \quad \text{für } u > 1 \tag{7}$$

(mit der Anfangsbedingung $\rho(u) = 1$ für $0 \leq u \leq 1$).

Die Differentialgleichung (6) gibt für $1 \leq u \leq 2$:

$$\rho(u-1) \equiv 1, \text{ also } \rho'(u) = -\frac{1}{u} \tag{8}$$

und damit $\rho(u) = 1 - \log u$, wobei der konstante Term so gewählt werden muss, dass $\rho(1) = 1$. Das kann man iterieren: für $2 \leq u \leq 3$ hat man $1 \leq u-1 \leq 2$ und also $\rho(u-1) = 1 - \log(u-1)$, sodass (6) liefert:

$$\rho'(u) = -\frac{1}{u}(1 - \log(u-1)), \tag{9}$$

und es kommt:

$$\rho(u) = 1 - \log u + \int_{2}^{u} \frac{\log(v-1)}{v} dv. \tag{10}$$

Mit diesen und ähnlichen Methoden hat Dan Bernstein eine Wertetabelle für die Funktion $\rho(u)$ erstellt (vgl. [1]).
Man hat zum Beispiel $\rho(2) = 3,0685281945 \cdot 10^{-1}$, was nach dem Satz von Dickman bedeutet: asymptotisch sind etwa $\frac{3}{10}$ aller Zahlen $\leq y^2$ ohne Primteiler $> y$, bei $n = 50$ hat man dann entsprechend: asymptotisch sind etwa $\frac{7}{10^{97}}$ aller Zahlen $\leq y^{50}$ ohne Primteiler $> y$.

Einschub Der Primzahlsatz gibt für $\pi(x) =$ Anzahl der Primzahlen $\leq x$:

$$\pi(x) \sim \frac{x}{\log x} \quad \text{für } x \to \infty \tag{11}$$

Vor dem Beweis des Primzahlsatzes hatte Tschebycheff asymptotische Schranken für $\pi(x)$ erhalten, und damit hat Mertens zwei nützliche zahlentheoretische Resultate gewonnen:

1. $\sum_{p \leq x} \frac{\log(p)}{p} = \log x + \mathcal{O}(1)$
2. $\sum_{p \leq x} \frac{1}{p} = \log\log(x) + c + o(1)$
   mit der Meissel-Mertens-Konstanten $c \approx 0,26149721\ldots$

(vgl. [3])

Der Satz von Dickmann besagt für $1 \leq u \leq 2$ mit $\rho(u) = 1 - \log u$:

$$\Psi(x, x^{\frac{1}{u}}) \sim x(1 - \log u) \quad \text{für } x \to \infty. \tag{12}$$

Dies wird im Folgenden bewiesen, und der Beweis wird indirekt geführt.
Annahme: Sei n eine positive Zahl kleiner gleich x und sei $n \notin S(x, x^{\frac{1}{u}})$.
$\Rightarrow$ n hat einen Primfaktor $p > x^{\frac{1}{u}}$
$\Rightarrow$ n kann nicht zwei solche Faktoren p und q haben, denn $x \geq n \geq pq > x^{\frac{2}{u}} \geq x$
An dieser Stelle fließt die Vorgabe $1 \leq u \leq 2$ ein.
$\Rightarrow$ n kann also als $n = pm$ dargestellt werden, mit p Primzahl aus $[x^{\frac{1}{u}}, x]$ und $m \leq \frac{x}{p}$
Das heißt die Anzahl der Zahlen bis x, deren Primfaktoren kleiner als $x^{\frac{1}{u}}$ sind, ist

$$\Psi(x, x^{\frac{1}{u}}) = \lfloor x \rfloor - \sum_{x^{\frac{1}{u}} < p \leq x} \left\lfloor \frac{x}{p} \right\rfloor,$$

denn genau $\left\lfloor \frac{x}{p} \right\rfloor$ Vielfache von p sind im Intervall $[1, x]$ enthalten. Diese sind aber wie p auch nicht $x^{\frac{1}{u}}$-glatt und werden abgezogen. Wenn jetzt die Gaußklammern entfernt werden, entsteht ein Fehler, der durch $\pi(x) \overset{(12)}{\sim} \frac{x}{log(x)}$ beschränkt wird.

$$\begin{aligned}
\Psi(x, x^{\frac{1}{u}}) &= x - x \sum_{x^{\frac{1}{u}} < p \leq x} \frac{1}{p} + \mathcal{O}\left(\frac{x}{log(x)}\right), \text{ also mit dem 2. Satz von Mertens:} \\
&= x\left(1 - (\log\log(x) - \log\log(x^{\frac{1}{u}}) + o(1))\right) + \mathcal{O}\left(\frac{x}{log(x)}\right) \\
&= x\left(1 - (\log\left(\frac{\log(x)}{\frac{1}{u}\log(x)}\right) + o(1))\right) + \mathcal{O}\left(\frac{x}{log(x)}\right) \\
&= x\left(1 - (\log\left(\frac{\log(x)}{\frac{1}{u}\log(x)}\right) + o(1)) + \mathcal{O}\left(\frac{1}{log(x)}\right)\right) \\
&= x\left(1 - \log(u) + o(1)\right)
\end{aligned}$$

und damit

$$\frac{\Psi(x, x^{\frac{1}{u}})}{x} \approx 1 - \log(u), \text{ für } x \to \infty \text{ bei } 1 \leq u \leq 2.$$

□

In der Tabelle von Dan Bernstein ist gut zu sehen, dass $\rho(u)$ sehr schnell gegen Null konvergiert, in etwa wie $\frac{1}{u^u}$. Es existieren verschiedene Schätzungen für $\rho(u)$. Eine grobe Schätzung ist

$$\rho(u) = \frac{1}{u^{u+o(u)}} \qquad \text{für} \quad u \to \infty, \tag{13}$$

dies kann aber genauer dargestellt werden durch

$$\rho(u) = \left(\frac{e + o(1)}{u\log(u)}\right)^u \qquad \text{für} \quad u \to \infty. \tag{14}$$

Ich zeige (14) $\Rightarrow$ (13)

$$\begin{aligned}
\rho(u) &= \left(\frac{e+o(1)}{u\log(u)}\right)^u \\
&= \frac{1}{u^u}\left(\frac{e+o(1)}{\log(u)}\right)^u \qquad \text{mit } \frac{\log u}{u} = o(1): \\
&\approx \frac{1}{u^u}\left(\frac{1}{u^{uo(1)}}\right) \\
&= \frac{1}{u^u}\frac{1}{u^{o(u)}} \\
&= \frac{1}{u^{u+o(u)}}
\end{aligned}$$

2 Wertebereich von $\Psi(x,y)$

De Bruijn zeigte 1951, bzw. 1966, dass

$$\Psi(x,y) = x\rho(u)\left\{1+\mathcal{O}\left(\frac{\log(u+1)}{\log(y)}\right)\right\} \qquad \text{mit } x = y^u \tag{15}$$

für $1 \le u \le (\log(y)^{3/5-\varepsilon})$, also

$$\begin{aligned}
\frac{\log(x)}{\log(y)} &\le \log(y)^{3/5-\varepsilon} \\
\log(x) &\le \log(y)^{8/5-\varepsilon} \\
\log(x)^{\frac{1}{8/5-\varepsilon}} &\le \log(y) \\
\log(x)^{\frac{1}{8/5(1-\frac{\varepsilon}{8/5})}} &\le \log(y) \\
\log(x)^{\frac{5}{8}\left(\frac{1}{1-\frac{\varepsilon}{8/5}}\right)} &\le \log(y) \\
\log(x)^{\frac{5}{8}\left(1+\frac{\varepsilon}{8/5}+(\frac{\varepsilon}{8/5})^2+\ldots\right)} &\le \log(y) \\
e^{(\log(x))^{5/8+\varepsilon}} &< y.
\end{aligned}$$

Dabei gilt (15) gleichmäßig in u für $1 \le u \le y^{1/2-\epsilon}$, das heißt $y \ge (\log x)^{2+\epsilon}$, genau dann, wenn die Riemannsche Vermutung richtig ist (siehe dazu [11]).

3 Schranken für $\Psi(x,y)$

In diesem Abschnitt geht es um schärfere Schranken für $\Psi(x,y)$. Diese werden benötigt, um besser einschätzen zu können, wieviele glatte Zahlen es in einem bestimmten Intervall gibt. Das bedeutet auch, dass die Rechenzeiten der Algorithmen besser abgeschätzt werden können.

Konyagin und Pomerance zeigten 1997 in „On primes recognizable in deterministic polynomial time“, dass $\Psi(x,y)$ für alle $x \geq y \geq 2$ und $x \geq 4$ nach unten folgendermaßen beschränkt ist

$$\Psi(x,y) \geq \frac{x}{(\log(x))^u} \tag{16}$$

Auch Hildebrand und de Bruijn haben sich mit den Schranken für $\Psi(x,y)$ beschäftigt, und die folgenden beiden Ergebnisse für obere und untere Schranken sind zu großen Teilen ihnen zuzuschreiben.

Für $x = y^u$ mit $u \geq 0$ hat man die folgenden Abschätzungen für die Anzahl $\Psi(x,y)$ der y-glatten Zahlen $\leq x$.

Satz 1 Es gilt für die Anzahl der y-glatten Zahlen bis x bei $x = y^u$:

$$\binom{\left\lfloor \frac{\log(x)}{\log(2)} \right\rfloor + \pi(y)}{\pi(y)} \geq \Psi(x,y) \geq \binom{\lfloor u \rfloor + \pi(y)}{\pi(y)} \geq \frac{\pi(y)^{\lfloor u \rfloor}}{\lfloor u \rfloor !} \tag{17}$$

und

Satz 2 Die Anzahl der y-glatten Zahlen bis x wird begrenzt durch

$$\frac{1}{k!} \prod_{p \leq y} \frac{\log(X)}{\log(p)} \geq \Psi(x,y) \geq \frac{1}{k!} \prod_{p \leq y} \frac{\log(x)}{\log(p)}, \qquad \text{mit } \log(X) = \log(x) + \sum_{p \leq y} \log(p) \tag{18}$$

Beweis zu Satz 1 Sei $p_1 < p_2 < p_3 < ... < p_k$ die Folge der Primzahlen $\leq y$, also $\pi(y) = k$. Für $n \in S(x,y)$, also $n \leq x$, kann man dann schreiben

$$n = p_1^{a_1} \cdot \ldots \cdot p_k^{a_k} \text{ mit } a_k \geq 0 \tag{19}$$

Dann folgt

$$a_1 \log(p_1) + a_2 \log(p_2) + ... + a_k \log(p_k) \leq \log(x). \tag{20}$$

Alle p_i sind größer oder gleich zwei, also ergibt sich mit $\log(p_i) \geq \log(2)$ weiterhin

$$a_1 + a_2 + ... + a_k \leq \left\lfloor \frac{\log(x)}{\log(2)} \right\rfloor = A.$$

Die Anzahl der y-glatten Zahlen bis x ist nach oben beschränkt durch die Anzahl der Möglichkeiten, k aus $A + k$ auszuwählen, also ist

$$\Psi(x,y) \leq \binom{A+k}{k}$$

eine obere Schranke.
Analog dazu erhält man

$$a_1 + a_2 + ... + a_k \geq \left\lfloor \frac{\log(x)}{\log(y)} \right\rfloor = \lfloor u \rfloor ,$$

da alle $p_i \leq y$ sein müssen und somit

$$\log(p_i) \leq \log(y), \text{ wenn } x = y^u, \text{ mit } \log(x) = u\log(y).$$

Es ergibt sich also hier die untere Schranke für $\Psi(x, y)$

$$\Psi(x, y) \geq \binom{\lfloor u \rfloor + k}{k}.$$

□

Die obere Schranke wird aber wertlos, wenn wir ein zu großes k wählen, denn dadurch errechnet sich $\binom{A+k}{k} > x$.
Die untere Schranke kann noch anders aufgeschrieben bzw. abgeschätzt werden.

$$\begin{aligned}
\binom{\lfloor u \rfloor + k}{k} &= \frac{(\lfloor u \rfloor + k)!}{(k)!(\lfloor u \rfloor)!} \\
&= \frac{(1) \cdot ... \cdot (k) \cdot (k+1) \cdot ... \cdot (k + \lfloor u \rfloor)}{(1) \cdot ... \cdot (k) \cdot \lfloor u \rfloor!} \\
&= \frac{(k+1) \cdot ... \cdot (k + \lfloor u \rfloor)}{\lfloor u \rfloor!} \\
&\geq \frac{k^{\lfloor u \rfloor}}{\lfloor u \rfloor!}
\end{aligned}$$

Dies kann weiterhin über die Stirling Formel approximiert werden, dazu zunächst die Formel

Stirling Formel $\mathrm{n}! \simeq \sqrt{2\pi n}\ \left(\frac{n}{\mathrm{e}}\right)^n, \qquad n \to \infty$

Wenden wir nun also diese an:

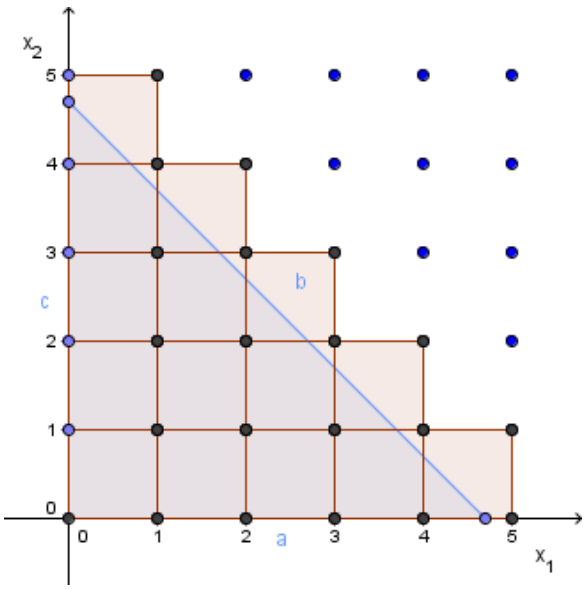
$$\begin{aligned}
\frac{k^{\lfloor u \rfloor}}{\lfloor u \rfloor!} &\approx \frac{k^u}{\sqrt{2\pi u}\ \left(\frac{u}{\mathrm{e}}\right)^u} \\
&\approx \frac{\left(\frac{y}{\log(y)}\right)^u}{\sqrt{2\pi u}\ \left(\frac{u}{\mathrm{e}}\right)^u} \\
&\approx \left(\frac{ey}{u\log(y)}\right)^u \\
&= \left(\frac{ey}{\log(x)}\right)^u \\
&= \frac{x}{\left(\left(\frac{1}{e}\right)\log(x)\right)^u}
\end{aligned}$$

Wir erhalten also

$$\Psi(x, y) \gtrapprox \frac{x}{\left(\left(\frac{1}{e}\right)\log(x)\right)^u},$$

was stark an (16) erinnert.

Beweis zu Satz 2 Sei ein Dreieck beschrieben durch $\{x_1, x_2 \geq 0, \alpha_1 x_1 + \alpha_2 x_2 \leq \tau\}$, über das ein Einheitsgitter gelegt wurde. Bildet man von jedem Gitterpunkt, der im Dreieck liegt, nach rechts oben ein Quadrat und schattiert dieses, liegt das Dreieck ganz in der schattierten Fläche. Die Anzahl der Gitterpunkte gibt offensichtlich den Flächeninhalt der schattierten Fläche an. Dazu folgende Abbildung:



Es gibt also mehr Gitterpunkte als der Flächeninhalt des Dreiecks groß ist. Dieser lässt sich berechnen durch $\frac{\tau}{\alpha_1} \cdot \frac{\tau}{\alpha_2} \cdot \frac{1}{2!}$, also gibt es mehr als $\frac{\tau^2}{(2!\alpha_1\alpha_2)}$ Gitterpunkte im Dreieck. Es gibt auch ein Dreieck, das die schattierte Fläche beinhaltet, beschrieben durch $\{x_1, x_2 \geq 0, \alpha_1(x_1 - 1) + \alpha_2(x_2 - 1) \leq \tau\}$, damit ist die Anzahl der Gitterpunkte im Dreieck höchstens $\frac{(\tau+\alpha_1+\alpha_2)^2}{2!\alpha_1\alpha_2}$.
Erweitern wir dies auf k-dimensionale Tetraeder, beschreibt

$$\{x_1, ..., x_k \geq 0, \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_k x_k \leq \tau\}$$

den inneren k-dimensionalen Tetraeder und

$$\{x_1, ..., x_k \geq 0, \alpha_1(x_1 - 1) + \alpha_2(x_2 - 1) + ... + \alpha_k(x_k - 1) \leq \tau\}$$

den äußeren.
Vergleicht man diese Darstellung mit (20), so entspricht der Flächeninhalt der schattierten Fläche also $\Psi(x, y)$ und durch die Dreiecke lassen sich Schranken finden. Das τ entspricht dabei dem $\log x$, die α entsprechen den $\log p$.
Das Volumen eines solchen Simplex ist $\frac{1}{k!} \prod_{p \leq y} \frac{\log(x)}{\log(p)}$ an der unteren Schranke, während bei der oberen das Simplex

$$\sum (a_i - 1) \log p_i \quad \leq \log x \quad \text{, das heißt} \tag{21}$$

$$\sum a_i \log p_0 \quad \leq \log x + \sum \log p_i \tag{22}$$

zu betrachten ist. Somit ergibt sich

$$\frac{1}{k!}\prod_{p\leq y}\frac{\log(X)}{\log(p)} \geq \Psi(x,y) \geq \frac{1}{k!}\prod_{p\leq y}\frac{\log(x)}{\log(p)}, \qquad \text{mit } \log(X) = \log(x) + \sum_{p\leq y}\log(p).$$

□

Literaturverzeichnis

[1] Andrew Granville, *Smooth numbers: computational number theory and beyond*, S. 267 bis 323 in Surveys in algorithmic number theory, editiert von J. P. Buhler und P. Stevenhagen, Math. Sci. Res. Inst. Publ. 44, Cambridge University Press, New York, 2008

[2] Carl Pomerance, *Smooth numbers and the quadratic sieve*, S. 69 bis 81 in Surveys in algorithmic number theory, editiert von J. P. Buhler und P. Stevenhagen, Math. Sci. Res. Inst. Publ. 44, Cambridge University Press, New York, 2008

[3] Friedrich Pillichshammer, *Verteilung der Primzahlen*, Universität Linz, 2008

[4] Karl Dickman, *On the frequency of numbers containing prime factors of a certain relative magnitude* in Arkiv for Matematik, Astronomi och Fysik, Band 22, 1930-1932

[5] Nicolaas Govert de Bruijn, *On the numbers of positive integers $\leq x$ and free of prime factors $\geq y$*, S. 50 bis 60 in Indagationes Mathematicae 13, 1951

[6] Nicolaas Govert de Bruijn, *On the numbers of positive integers $\leq x$ and free of prime factors $\geq y$ II*, S. 239 bis 247 in Indagationes Mathematicae 28, 1966

[7] A. Hildebrand, G. Tenenbaum, *Integers without large primefactors*, S. 411 bis 484 in J. Th. Nombres Bordeaux 5, 1993

[8] E. Canfield, D. Erdös, C. Pomerance, *On a problem of Oppenheim concerning "factorisatio numerorum"*, S. 1 bis 28 in J. Number Theory 17, 1983

[9] V. Ramaswami, *On the number of positive integers less than x and free of prime divisors greater than x^c*, S. 1122 bis 1127 in Bull AMS 55, 1949

[10] S. Konyagin, C. Pomerance, *On primes recognizable in deterministic polynomial time*, S. 176 bis 198 in The Mathematics of Paul Erdös I, R. L. Graham and J. Nešetřil, eds., Springer, 1997

[11] A. Hildebrand, *Integers free of large prime factors and the Riemann hypothesis*, S. 258 bis 271 in Mathematika 31, 1984