3.5 Der probabilistische Primzahltest von Rabin

Grundlage ist die Übertragung einer Idee von Solovay und Strassen auf den Millerschen Test, die Rabin vorgeschlagen hat. Wählt man a zufällig in $[2 \dots n-1]$, so fällt n „in der Regel“ durch den strengen Pseudoprimzahltest zur Basis a, wenn es zusammengesetzt ist. Was heißt aber „in der Regel“? Wie groß ist die Wahrscheinlichkeit? Diese Frage wurde schon im Korollar zu Satz 2 beantwortet, wobei die schärfere Schranke $\frac{1}{4}$ ohne Beweis angegeben wurde.

Zu bemerken ist, dass die Schranke $\frac{1}{4}$ scharf ist. Das sieht man an den Zahlen der Form

$$n = (1+2t)(1+4t)$$

mit ungeradem t (sofern die Faktoren $p = 1+2t$ und $q = 1+4t$ prim sind – Beispiel: $t = 24969$, $p = 49939$, $q = 99877$). Dann ist $n-1 = 2r$ mit $r = 3t+4t^2$,

$$B_u = \{a \mid a^r \equiv 1 \pmod{n}\} \cup \{a \mid a^r \equiv -1 \pmod{n}\}.$$

Da $\mathrm{ggT}(r, p-1) = \mathrm{ggT}(3t+4t^2, 2t) = t = \mathrm{ggT}(r, q-1)$, hat jede dieser beiden Kongruenzen genau t^2 Lösungen. Also ist $\#B_u = 2t^2$,

$$\frac{\#B_u}{n-1} = \frac{2t^2}{2\cdot(3t+4t^2)} = \frac{t}{3+4t} = \frac{1}{4+\frac{3}{t}}.$$

Die Grenze $\frac{1}{4}$ wird allerdings für die meisten zusammengesetzten Zahlen längst nicht erreicht.

Allgemein sei eine Familie $(B_{(n)})_{n\geq 1}$ von Mengen $B_{(n)} \subseteq [1 \dots n-1]$ und ein $\varepsilon \in]0,1[$ gegeben mit

1. $B_{(n)} = [1 \dots n-1]$, wenn n prim,
2. $\#B_{(n)} \leq \varepsilon \cdot (n-1)$ für alle genügend großen ungeraden zusammengesetzten Zahlen n.

Ferner soll die Entscheidung, ob $a \in B_{(n)}$, für alle $a \in [1 \dots n-1]$ effizient möglich sein, also mit einem Aufwand, der höchstens polynomial mit $\ln(n)$ wächst. Dann gibt es einen zugehörigen (abstrakten) Pseudoprimzahltest:

1. Wähle $a \in [1 \dots n-1]$ zufällig.
2. Prüfe, ob $a \in B_{(n)}$.
3. Ausgabe:
   (a) Falls **nein**: n ist sicher zusammengesetzt.
   (b) Falls **ja**: n ist pseudoprim für a.

Der zugehörige **probabilistische Primzahltest** besteht aus der Aneinanderreihung von k solchen Pseudoprimzahltests mit unabhängig voneinander zufällig gewählten Basen a (die sich also auch wiederholen dürfen). Ist stets $a \in B_{(n)}$, so ist n fast sicher prim – man kann diesem Ergebnis die „Irrtumswahrscheinlichkeit“ δ zuweisen, die aber nicht $= \varepsilon^k$ ist, sondern sich so berechnet:

In der Menge der ungeraden Zahlen $< 2^r$, also mit höchstens r Bits sei X die Teilmenge der *zusammengesetzten* Zahlen und Y_k die Menge der Zahlen, die k unabhängige (abstrakte) Pseudoprimzahltests bestehen. Die Wahrscheinlichkeit, dass das einer zusammengesetzten Zahl gelingt, ist dann gegeben durch die bedingte Wahrscheinlichkeit $P(Y_k|X) \leq \varepsilon^k$. Für die praktische Anwendung interessanter ist allerdings die „umgekehrte“ Wahrscheinlichkeit $\delta = P(X|Y_k)$ dafür, dass eine Zahl n, die bestanden hat, trotzdem zusammengesetzt ist. Diese kann man mit Hilfe der Bayesschen Formel abschätzen:

$$P(X|Y_k) = \frac{P(X) \cdot P(Y_k|X)}{P(Y_k)} \leq \frac{P(Y_k|X)}{P(Y_k)} \leq \frac{1}{q} \cdot \varepsilon^k \leq r \cdot \ln(2) \cdot \varepsilon^k,$$

wobei die Dichte der Primzahlen nach dem Primzahlsatz eingeht:

$$P(Y_k) \geq P(\text{prim}) =: q \geq \frac{1}{r \cdot \ln(2)}$$

(die letzte Ungleichung ist sogar großzügig, da wir nur ungeade Zahlen betrachten). Die gesuchte „Irrtumswahrscheinlichkeit“ $\delta = P(X|Y_k)$ könnte also durchaus größer als ε^k sein. Man kann (und) sollte sie dadurch verringern, dass man die Grundmenge, in der man nach Primzahlen sucht, einschränkt und damit $P(Y_k)$ vergrößert, z. B. indem man vor Anwendung des Pseudoprimzahltests durch alle Primzahlen etwa $< 100r$ probeweise dividiert.

Beim **Primzahltest von** Rabin ist $B_{(n)}$ die Menge der Basen a, für die n den strengen Pseudoprimzahltest zur Basis a besteht, und $\varepsilon = \frac{1}{4}$. Übersteht n die 25-fache Anwendung, so ist es mit einer sehr kleinen Irrtumswahrscheinlichkeit prim. Es ist eher wahrscheinlich, dass die Berechnung wegen eines Hard- oder Software-Fehlers falsch ist, als dass der Algorithmus die Primzahl-Eigenschaft falsch schätzt. Knuth bezweifelt auch, dass ein veröffentlichter Beweis der erweiterten Riemannschen Vermutung jemals eine so hohe Glaubwürdigkeit haben kann. Dennoch ist es natürlich mathematisch unbefriedigend, nicht mit absoluter Sicherheit sagen zu können, dass wirklich eine Primzahl vorliegt.

Als weiterführende Literatur zur Frage, wie „gut“ ein probabilistischer Primzahltest ist, sei

- S. H. Kim/ C. Pomerance: The probability that a random probable prime is composite. Math Comp. 53 (1989), 721–741,

empfohlen.