KAPITEL 6

Die Pollardsche ρ-Methode zur Faktorisierung

1. Einführung

Um die Primfaktorzerlegung einer natürlichen Zahl n zu erhalten, geht man normalerweise folgendermaßen vor:
Man teilt zunächst alle kleinen Teiler von n heraus und macht dann einen Primzahltest. Nun gibt es zwei Möglichkeiten:

- Ist n wahrscheinlich prim, so kann man beweisen (bzw. zu beweisen versuchen), dass n prim ist, oder sich mit dem Ergebnis 'wahrscheinlich prim' zufrieden geben.
- Ist n zusammengesetzt, sucht man einen nichttrivialen Teiler t von n und beginnt mit t und $\frac{n}{t}$ von vorne.

Unser Hauptproblem ist nun: Wie findet man einen nichttrivialen Teiler einer zusammengesetzten Zahl?

Wir haben bereits die naive Faktorisierungsmethode kennengelernt. Um damit den kleinsten Primteiler p einer Zahl n zu bestimmen, braucht man allerdings $O(p)$ Schritte. Wir wollen jetzt ein besseres Verfahren kennenlernen. (Wir haben auch gesehen, dass das Fermatsche Faktorisierungsverfahren nur in sehr speziellen Fällen zum Ziel führt.)

2. Die Pollardsche ρ-Methode

Wir haben folgende Situation: n ist eine zusammengesetzte natürliche Zahl ohne kleine Teiler. Wir wollen einen nichttrivialen Teiler von n finden.

Eine Idee: Sei n zusammengesetzt ohne kleine Teiler und p ein Primteiler von n. Ist x_i eine Folge von Zahlen modulo n, d.h. $x_i \in \{0, 1, \ldots, n-1\}$ und gilt für zwei Indizes $k < l$ die Beziehung $x_k \equiv x_l \bmod p$, so gilt $p | \mathrm{ggT}(x_k - x_l, n)$, man kann also hoffen, so einen nichttrivialen Teiler von n gefunden zu haben.

Beispiel: Wir betrachten $n = 9797 = 97 \cdot 101$. Mit Maple erzeugen wir eine Zufallsfolge x_i modulo n:

i	1	2	3	4	5	6	7	8	9	10
x_i	3302	8560	6905	8869	4675	9247	1456	148	6081	4485
$x_i \bmod 97$	4	24	18	42	19	32	1	51	67	23
$x_i \bmod 101$	70	76	37	82	29	56	42	47	21	41

i	11	12	13	14	15	16	17	18	19	20
x_i	9478	6095	722	2723	8544	3776	920	2943	746	5010
$x_i \bmod 97$	69	81	43	7	8	90	47	33	67	63
$x_i \bmod 101$	85	35	15	97	60	39	11	14	39	61

Man sieht

$$x_9 \equiv x_{19} \equiv 67 \bmod 97, \quad x_{16} \equiv x_{19} \equiv 39 \bmod 101$$

und findet mittels

$$\mathrm{ggT}(x_9 - x_{19}, n) = 97, \quad \mathrm{ggT}(x_{16} - x_{19}, n) = 101$$

die Faktorisierung von n.

Wir wollen jetzt die obige Idee genauer untersuchen.

Datei: ki_rho.tex. Version vom 16.6.2016

LEMMA. *Sei M eine Menge mit m Elementen und $x_1, x_2, \ldots, x_k$ eine zufällig gewählte Folge von Elementen aus M. Dann ist die Wahrscheinlichkeit, dass alle Folgenglieder verschieden sind, gleich*

$$p_{m,k} = (1 - \frac{1}{m})(1 - \frac{2}{m}) \ldots (1 - \frac{k-1}{m}).$$

(Ziehen mit Zurücklegen.)

Beweis: Insgesamt gibt es m^k Möglichkeiten für $(x_1, \ldots, x_k)$. Wieviele günstige Möglichkeiten gibt es? x_1 kann beliebig gewählt werden, also m Möglichkeiten, für x_2 bleiben noch $m-1$ Möglichkeiten, für x_3 dann noch $m-2$ Möglichkeiten, etc. Also gibt es $m(m-1)(m-2)\ldots(m-(k-1))$ günstige Möglichkeiten, woraus sich durch Division sofort die Behauptung ergibt. ■

Beispiel: Wir wählen M mit $|M| = 365$ und eine zufällige Folge $x_1, \ldots, x_k$. Die Wahrscheinlichkeit, dass alle Folgenglieder verschieden sind, kann man dann für verschiedene Werte von k der folgenden Tabelle entnehmen:

k	10	20	30	40	50	60
Wahrscheinlichkeit in %	88.3	58.9	29.4	10.9	3.0	0.6

Eine Interpretation: Hat man eine Gruppe von 30 Leuten, so ist die Wahrscheinlichkeit, dass zwei davon am gleichen Tag Geburtstag haben, größer als 70 Prozent, bei einer Gruppe von 60 Leuten erhöht sich die Wahrscheinlichkeit schon auf über 99 Prozent, etc.

Überlegung: Wir wollen die Wahrscheinlichkeit $p_{m,k}$ abschätzen für den Fall, daß m groß gegen k ist. Zunächst ist

$$\ln p_{m,k} = \sum_{i=1}^{k-1} \ln(1 - \frac{i}{m}).$$

Ist k groß gegen m und $1 \leq i \leq k-1$, so ist wegen $\ln(1-x) \approx -x$ (für kleine x) dann $\ln(1 - \frac{i}{m}) \approx -\frac{i}{m}$ und damit

$$\ln p_{m,k} \approx -\sum_{i=1}^{k-1} \frac{i}{m} = -\frac{k(k-1)}{2m},$$

d.h.

$$p_{m,k} \approx e^{-\frac{k(k-1)}{2m}}.$$

Ist z.B. m groß und $k \approx 3.1\sqrt{m}$, so wird

$$p_{m,k} \approx e^{-\frac{3.1\sqrt{m}(3.1\sqrt{m}-1}{2m}} \approx e^{-3.1^2/2} \approx 0.0082.$$

Indem man kleine Werte von m explizit betrachtet, kann man dann folgendes Lemma beweisen:

LEMMA. *Ist M eine Menge mit m Elementen und $x_1, \ldots, x_k$ eine zufällig gewählte Folge in M mit*

$$k \geq 3.1\sqrt{m},$$

so ist die Wahrscheinlichkeit, dass zwei Folgenglieder gleich sind, größer als 99 Prozent.

Beweis:

(1) Die Funktion $x \mapsto \ln(1-x)$ wird im Intervall $(-1, 1)$ durch ihre Taylorreihe beschrieben, d.h. es gilt

$$\ln(1-x) = -\sum_{n=1}^{\infty} \frac{1}{n} x^n \quad \text{für} \quad -1 < x < 1.$$

Daraus folgt sofort

$$\ln(1-x) \leq -x \quad \text{für} \quad 0 \leq x < 1.$$

(2) Mit den Bezeichnungen des vorangegangenen Lemmas ist

$$p_{m,k} = \prod_{i=1}^{k-1}(1 - \frac{i}{m}).$$

Für $k \geq m+1$ ist $p_{m,k} = 0$, sodass wir uns hier auf $1 \leq k \leq m$ beschränken können. Wir erhalten mit (1)

$$\ln p_{m,k} = \sum_{i=1}^{k-1} \ln(1 - \frac{i}{m}) \leq \sum_{i=1}^{k-1}(-\frac{i}{m}) = -\frac{k(k-1)}{2m},$$

also

$$p_{m,k} \leq e^{-\frac{k(k-1)}{2m}}.$$

(3) Wir wollen untersuchen, wann $p_{m,k} < \frac{1}{100}$ gilt. Wir formen äquivalent um:

$$\begin{aligned} e^{-\frac{k(k-1)}{2m}} < \frac{1}{100} \quad &\Longleftrightarrow \quad -\frac{k(k-1)}{2m} < -2\ln 10 \quad \Longleftrightarrow \quad k(k-1) > 4m\ln 10 \quad \Longleftrightarrow \\ &\Longleftrightarrow \quad (k-\frac{1}{2})^2 > 4m\ln 10 + \frac{1}{4} \quad \Longleftrightarrow \quad k > \sqrt{4m\ln 10 + \frac{1}{4}} + \frac{1}{2} \quad \Longleftrightarrow \\ &\Longleftrightarrow \quad \frac{k}{\sqrt{m}} > \sqrt{4\ln 10 + \frac{1}{4m}} + \frac{1}{2\sqrt{m}}. \end{aligned}$$

Die Funktion

$$f(m) = \sqrt{4\ln 10 + \frac{1}{4m}} + \frac{1}{2\sqrt{m}}$$

ist streng monoton fallend mit

$$f(60) = 3.1000903725\ldots \quad \text{und} \quad f(61) = 3.0995478390\ldots$$

Wir erhalten daher für $m \geq 61$ folgende Implikationen:

$$\begin{aligned} k \geq 3.1\sqrt{m} \quad &\Longrightarrow \quad \frac{k}{\sqrt{m}} \geq 3.1 > f(61) \geq f(m) = \sqrt{4\ln 10 + \frac{1}{4m}} + \frac{1}{2\sqrt{m}} \\ &\Longrightarrow \quad p_{m,k} \leq e^{-\frac{k(k-1)}{2m}} < \frac{1}{100}. \end{aligned}$$

Dies beweist die Behauptung für $m \geq 61$.

(4) Durch explizites Ausrechnen von $p_{m,\lceil 3.1\sqrt{m}\rceil}$ sieht man, dass die behauptete Aussage auch für $m \geq 60$ gilt. ■

Bemerkung: Statt 99 Prozent kann man natürlich auch eine Abschätzung für andere Werte herleiten. Für uns ist wichtig: Hat man eine zufällig gewählte Folge x_i, so ist es sehr wahrscheinlich, dass unter den ersten $c\sqrt{m}$ Folgenglieder zwei gleiche sind, z.B. mit $c = 3.1$.

Anwendung: Sei n eine zusammengesetzte Zahl und p ein Primteiler von n. Sei $x_1, x_2, x_3, \ldots, x_k$ eine zufällig gewählte Folge von Zahlen modulo n. Dann erhält man durch Reduktion modulo p eine Folge in $\mathbb{Z}/(p)$. Ist nun $k \geq 3.1\sqrt{p}$, so ist nach unseren Überlegungen die Wahrscheinlichkeit, dass es verschiedene Indizes i, j gibt mit $x_i \equiv x_j \bmod p$, größer als 99 Prozent. In diesem Fall ist wieder $ggT(x_i - x_j, n)$ ein von 1 verschiedener Teiler von n.
Will man dies in die Praxis umsetzen, stellen sich folgende beiden Probleme:

(1) Wie erhält man eine Zufallsfolge $x_1, x_2, x_3, \ldots$?
(2) Wie findet man Indizes $i < j$ mit $1 < \mathrm{ggT}(x_i - x_j, n) < n$?

Wir wenden uns zunächst der Konstruktion von Zufallsfolgen in einer endlichen Menge M zu. (Dabei darf man zunächst den Begriff *Zufallsfolge* nicht zu eng sehen, d.h. wir werden keine wahrscheinlichkeitstheoretischen Aussagen beweisen.) Ein Ansatz ist folgender:

- Wähle $x_0 \in M$,
- wähle eine Abbildung $f : M \to M$,
- definiere rekursiv $x_i = f(x_{i-1})$ für $i \geq 1$.

Beispiel: Wir betrachten $M = \mathbb{Z}/(61)$, $x_0 = 2$ und $f(x) = x^2 + 2 \bmod 61$. Dann ergibt sich die Folge

$$2,\quad 6,\quad 38,\quad 43,\quad 21,\quad 16,\quad 14,$$
$$15,\quad 44,\quad 47,\quad 15,\quad 44,\quad 47,\quad 15,\quad 44,\quad 47,\quad 15,\quad 44,\quad 47,\quad 15,\quad 44,\quad \ldots$$

Die Folge ist also irgendwann periodisch. Das gleiche Phänomen passiert allgemein:

LEMMA. *Für eine endliche Menge M, eine Abbildung $f : M \to M$, ein Element $x_0 \in M$ werde eine Folge $(x_i)_{i\geq 0}$ rekursiv definiert durch*

$$x_i = f(x_{i-1}) \quad \text{für} \quad i \geq 1.$$

(1) *Es gibt ganze Zahlen $s \geq 0$ und $t \geq 1$ mit der Eigenschaft*

$$x_s = x_{s+t}.$$

(2) *Sind $s \geq 0$ und $t \geq 1$ Zahlen mit $x_s = x_{s+t}$, so gilt*

$$x_j = x_{j+mt} \quad \text{für alle} \quad j \geq s, \quad m \geq 1.$$

(3) *Wählt man ein minimales $t \geq 1$ und, dazu passend, ein minimales $s \geq 0$ mit obigen Eigenschaften, so gilt für alle ganzen Zahlen $j \geq 0$ und $u \geq 1$:*

$$x_j = x_{j+u} \quad \Longleftrightarrow \quad t \mid u \text{ und } j \geq s.$$

s und t sind durch diese Eigenschaft eindeutig bestimmt. $x_0, x_1, \ldots, x_{s-1}$ heißt die Vorperiode der Folge und $x_s, x_{s+1}, \ldots, x_{s+t-1}$ die Periode der Folge. t wird Periodenlänge oder auch einfach Periode genannt.

Beweis:

(1) Da M als endlich vorausgesetzt ist, muss es Indizes $j < k$ geben mit $x_j = x_k$. Setzen wir $s = j$ und $t = k - j$, so folgt die Behauptung.

(2) Aus $x_s = x_{s+t}$ folgt durch Anwendung von f nacheinander

$$\begin{aligned} x_{s+1} &= f(x_s) = f(x_{s+t}) = x_{s+t+1} = x_{s+1+t}, \\ x_{s+2} &= f(x_{s+2}) = f(x_{s+1+t}) = x_{s+1+t+1} = x_{s+2+t}, \\ x_{s+3} &= f(x_{s+2}) = f(x_{s+2+t}) = x_{s+2+t+1} = x_{s+3+t}, \\ &\vdots \qquad\quad \vdots, \end{aligned}$$

woraus man durch Induktion sofort $x_j = x_{j+t}$ für $j \geq s$ erhält. Damit folgt dann für $j \geq s$

$$x_j = x_{j+t} = x_{j+t+t} = x_{j+2t} = x_{j+2t+t} = x_{j+3t} = \ldots,$$

woraus man durch Induktion schließlich $x_j = x_{j+mt}$ für $j \geq s$ und $m \geq 1$ erhält.

(3) Sei jetzt t minimal mit obiger Eigenschaft, dazu $s \geq 0$ minimal mit $x_s = x_{s+t}$.

(a) Wir definieren

$$P = \{u \geq 1 : \text{ es gibt ein } i \geq 0 \text{ mit } x_i = x_{i+u}\}.$$

Ist $u \in P$ und $i \geq 0$ mit $x_i = x_{i+u}$, so folgt wie oben

$$x_j = x_{j+u} \quad \text{für alle} \quad j \geq i,$$

also können wir P alternativ durch

$$P = \{u \geq 1 : \text{ es gibt ein } i_u \geq 0 \text{ mit } x_i = x_{i+u} \text{ für alle } i \geq i_u\}$$

beschreiben.

(b) Nach Definition von t gilt $t = \min P$. Mit (2) folgt sofort, dass auch $mt \in P$ für alle $m \in \mathbb{N}$ gilt. Sei nun umgekehrt $u \in P$. Wir dividieren u durch t und erhalten eine Darstellung $u = qt + r$ mit $q \geq 1$ und $0 \leq r < t$. Wäre $r > 0$, so würde für hinreichend große i folgen

$$x_i = x_{i+u} = x_{i+qt+r} = x_{i+r},$$

also $r \in P$, ein Widerspruch zur Minimalität von t. Daher muss $r = 0$ sein, $u = qt$, d.h. $t \mid u$, und damit $P = \mathbb{N}t$.

(c) Wir wollen jetzt (3) beweisen. Die Richtung $\Longleftarrow$ in (3) folgt sofort aus (2).

(d) Wir zeigen $\Longrightarrow$ in (3). Sei also $x_j = x_{j+u}$ mit $j \geq 0$ und $u \geq 1$. Dann ist $u \in P$, also $t \mid u$, $u = qt$ für ein $q \in \mathbb{N}$. Angenommen, es wäre $j < s$. Aus (2) folgt dann $x_{s-1} = x_{s-1+u} = x_{s-1+qt}$. Wegen $s - 1 + t \geq s$ ist aber

$$x_{s-1+t} = x_{s-1+2t} = \cdots = x_{s-1+qt} = x_{s-1+u} = x_{s-1},$$

und damit $x_{s-1} = x_{s-1+t}$, ein Widerspruch zur Minimalität von s. Also folgt $j \geq s$, wie behauptet. Damit ist auch Aussage (3) bewiesen.■

Bemerkung: Stellt man eine Folge wie im Lemma graphisch dar, so wird man an den griechischen Buchstaben ρ erinnert, wobei die Vorperiode dem Schwanz von ρ, die Periode dem Kreis von ρ entspricht. Daher kommt der Name des Pollardschen ρ-Verfahrens.

Bemerkung: Es ist nicht klar, wann eine durch Vorgabe von $x_0 \in \mathbb{Z}/n\mathbb{Z}$ und $f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ konstruierte Folge $x_i = f(x_{i-1})$ (im wahrscheinlichkeitstheoretischen Sinn) zufällig ist. In der Praxis hat sich der Ansatz $f(x) = x^2 + a \bmod n$ bewährt - mit gewissen Einschränkungen an a, z.B. $a \neq 0$, $a \neq -2$

Unser nächstes Problem ist, wie man zu einer Folge x_i in einer endlichen Menge M, die durch eine Abbildung $f : M \to M$ mittels $x_{i+1} = f(x_i)$ definiert ist, Indizes $j < k$ finden kann mit $x_j = x_k$.
Natürlich könnte man so vorgehen: Man vergleicht x_1 mit x_0, dann x_2 mit x_0 und x_1, dann x_3 mit x_0, x_1, x_2, etc. Ist man bei x_k angelangt, so braucht man insgesamt $\binom{k+1}{2} = \frac{k(k+1)}{2}$ Vergleiche, außerdem muss man die Elemente $x_0, x_1, \ldots, x_k$ speichern. Dies ist im allgemeinen zu aufwendig.
Einen eleganten Ausweg liefert folgendes Lemma:

LEMMA. *Sei M eine endliche Menge, $f : M \to M$ eine Abbildung, $x_0 \in M$ und die Folge $(x_i)_{i \geq 0}$ rekursiv definiert durch $x_i = f(x_{i-1})$ für $i \geq 1$. Die Folge habe eine Vorperiode der Länge s und Periodenlänge t.*

(1) *Für $\ell \geq 1$ gilt:*

$$x_\ell = x_{2\ell} \quad \Longleftrightarrow \quad t \mid \ell \text{ und } \ell \geq s.$$

Bezeichnet ℓ_0 das kleinste derartige ℓ, so gilt

$$\ell_0 = \begin{cases} t & \text{für } s = 0, \\ \lceil \frac{s}{t} \rceil \cdot t & \text{für } s > 0. \end{cases}$$

In jedem Fall ist $\ell_0 \leq s + t$.

(2) *Definiert man eine Folge y_i durch $y_0 = x_0$ und $y_i = f(f(y_{i-1}))$, so gilt $y_i = x_{2i}$. (Insbesondere ist $x_\ell = x_{2\ell}$ gleichwertig mit $x_\ell = y_\ell$.)*

Beweis:

(1) Mit dem letzten Lemma gilt für $\ell \geq 1$:

$$x_\ell = x_{2\ell} = x_{\ell+\ell} \quad \Longleftrightarrow \quad t \mid \ell \text{ und } \ell \geq s.$$

Wir können daher $\ell = mt$ mit $m \geq 1$ ansetzen. Die zweite Bedingung wird dann

$$\ell \geq s \quad \Longleftrightarrow \quad mt \geq s \quad \Longleftrightarrow \quad m \geq \frac{s}{t} \quad \Longleftrightarrow \quad m \geq \lceil \frac{s}{t} \rceil,$$

was auch die Aussage über ℓ_0 beweist. Die behauptete Abschätzung ist trivial für $s = 0$, für $s \geq 1$ folgt sie aus

$$\ell_0 = \lceil \frac{s}{t} \rceil \cdot t < (\frac{s}{t} + 1) \cdot t = s + t.$$

Dies war zu zeigen.

(2) Wir zeigen $y_i = x_{2i}$ durch Induktion, wobei der Fall $i = 0$ aufgrund der Definition klar ist. Nun ist

$$y_{i+1} = f(f(y_i)) = f(f(x_{2i})) = f(x_{2i+1}) = x_{2i+2} = x_{2(i+1)},$$

was die Behauptung liefert. ■

Beispiel: Wir betrachten wieder die durch $x_0 = 2$, $x_i = x_{i-1}^2 + 2 \bmod 61$ definierte Folge.

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
x_i	2	6	38	43	21	16	14	15	44	47	15	44	47	15	44	47	15	44	47	15	44
y_i	2	38	21	14	44	15	47	44	15	47	44	15	47	44	15	47	44	15	47	44	15

Die Periode ist 15, 44, 47, also $t = 3$ und $s = 7$. Das kleinste $\ell \geq 1$ mit $x_\ell = x_{2\ell}$ ist $\ell = 9$.

Bemerkungen: Sei M eine endliche Menge, $x_0 \in M$ und eine Folge x_i rekursiv definiert durch $x_{i+1} = f(x_i)$.

(1) Um Indizes $j < k$ mit $x_j = x_k$ zu finden, kann man nach dem Lemma so vorgehen: Man berechnet parallel zu x_i eine Folge y_i, die durch $y_0 = x_0$ und $y_{i+1} = f(f(y_i))$ definiert wird, und vergleicht x_i mit y_i. Gilt für $\ell \geq 1$ nun $x_\ell = y_\ell$, so ist $x_\ell = x_{2\ell}$ und wir haben unser Problem gelöst. Ein großer Vorteil ist, dass man keine Speicherprobleme hat.
(2) Aus unseren wahrscheinlichkeitstheoretischen Überlegungen folgt weiter: Ist $x_0, x_1, \ldots, x_{[3.1\sqrt{m}]}$ eine Zufallsfolge, dann ist die Wahrscheinlichkeit, dass ein ℓ mit $1 \leq \ell \leq 3.1\sqrt{m}$ und $x_\ell = y_\ell = x_{2\ell}$ existiert, größer als 99 Prozent.

Wir fassen unsere Überlegungen zusammen: Ist $x_0, x_1, x_2, \ldots$ eine wie oben definierte Zufallsfolge modulo n und p ein Primteiler von n, so ist die Wahrscheinlichkeit, dass es einen Index ℓ mit $1 \leq \ell \leq 3.1\sqrt{p}$ gibt, sodass $x_\ell \equiv x_{2\ell} = y_\ell \bmod p$ gilt, größer als 99 Prozent. Dann haben wir $p | \mathrm{ggT}(x_\ell - x_{2\ell}, n)$. Wir erhalten damit eine Version der Pollardschen ρ-Methode zur Faktorisierung:

Faktorisierung mit der Pollardschen ρ-Methode: Sei n eine zusammengesetzte natürliche Zahl ohne kleine Teiler.

(1) Wähle $x, a \in \mathbb{Z}$, setze $y := x$. (Beispielsweise $x := 2$, $a = 2$.)
(2) Berechne $x := x^2 + a \bmod n$, $y := y^2 + a \bmod n$, $y := y^2 + a \bmod n$.
(3) Berechne $d = \mathrm{ggT}(x - y, n)$.
(4) Ist $d = 1$, gehe zurück zu (2).
(5) Ist $1 < d < n$ gib d als nichttrivialen Teiler von n aus, andernfalls gehe zu 2.
(6) Ist $d = n$, kann man es noch einmal versuchen, man kann allerdings auch a und den Startwert für x in (1) verändern.

Beispiel: $n = 9797 = 97 \cdot 101$. Wir betrachten die durch $x_0 = 2$, $x_i^2 = x_{i-1}^2 + 2$ definierte Folge. Modulo 9797 hat sie Vorperiode 12, Periode 24 und $\ell_0 = 24$, modulo 97 Vorperiode 2, Periode 8 und $\ell_0 = 8$, modulo 101 Vorperiode 12, Periode 6 und $\ell_0 = 12$.

i	x_i	y_i	ggT$(x_i - y_i, 9797)$	x_i mod 97	y_i mod 97	x_i mod 101	y_i mod 101
1	6	38	1	6	38	6	38
2	38	4157	1	38	83	38	16
3	1446	2734	1	88	18	32	7
4	4157	7047	1	83	63	16	78
5	8540	4112	1	4	38	56	72
6	2734	1732	1	18	83	7	15
7	9444	7293	1	35	18	51	21
8	7047	3846	97	63	63	78	8
9	9015	3045	1	91	38	26	15
10	4112	6485	1	38	83	72	21
11	8721	5159	1	88	18	35	8
12	1732	1227	101	83	63	15	15
13	1944	7798	1	4	38	25	21
14	7293	4351	1	18	83	21	8
15	9735	2540	1	35	18	39	15
16	3846	5980	97	63	63	8	21
17	8045	5664	1	91	38	66	8
18	3045	1732	101	38	83	15	15
19	4065	7293	1	88	18	25	21
20	6485	3846	1	83	63	21	8
21	6503	3045	1	4	38	39	15
22	5159	6485	1	18	83	8	21
23	6631	5159	1	35	18	66	8
24	1227	1227	9797	63	63	15	15
25	6590	7798	1	91	38	25	21
26	7798	4351	1	38	83	21	8
27	8624	2540	1	88	18	39	15
28	4351	5980	1	83	63	8	21
29	3399	5664	1	4	38	66	8
30	2540	1732	101	18	83	15	15
31	5176	7293	1	35	18	25	21
32	5980	3846	97	63	63	21	8
33	1352	3045	1	91	38	39	15
34	5664	6485	1	38	83	8	21
35	5520	5159	1	88	18	66	8
36	1732	1227	101	83	63	15	15
37	1944	7798	1	4	38	25	21
38	7293	4351	1	18	83	21	8
39	9735	2540	1	35	18	39	15
40	3846	5980	97	63	63	8	21
41	8045	5664	1	91	38	66	8
42	3045	1732	101	38	83	15	15
43	4065	7293	1	88	18	25	21
44	6485	3846	1	83	63	21	8
45	6503	3045	1	4	38	39	15
46	5159	6485	1	18	83	8	21
47	6631	5159	1	35	18	66	8
48	1227	1227	9797	63	63	15	15
49	6590	7798	1	91	38	25	21
50	7798	4351	1	38	83	21	8
51	8624	2540	1	88	18	39	15
52	4351	5980	1	83	63	8	21
53	3399	5664	1	4	38	66	8
54	2540	1732	101	18	83	15	15
55	5176	7293	1	35	18	25	21
56	5980	3846	97	63	63	21	8
57	1352	3045	1	91	38	39	15
58	5664	6485	1	38	83	8	21
59	5520	5159	1	88	18	66	8
60	1732	1227	101	83	63	15	15

Beispiel: Wir wollen mit $x_0 = 2$, $x_i = x_{i-1}^2 + 2 \bmod n$ faktorisieren. Unter der ℓ-Folge einer Zahl n verstehen wir die Indizes ℓ, für die $x_\ell \equiv x_{2\ell} \bmod n$ gilt.

(1) 7 hat die ℓ-Folge $(3, 4, 5, \dots)$, 19 und 23 haben die gleiche ℓ-Folge $(3, 6, 9, \dots)$.
(2) Die ersten ggT's $\neq 1$ beim Faktorisieren von $133 = 7 \cdot 19$ sind 133 und dann 7.
(3) Die ggT's $\neq 1$ beim Faktorisieren von $437 = 19 \cdot 23$ sind immer 437.

Beispiel: Ein Gefühl für die Anzahl der Schritte und die Rechenzeit kann man bei folgenden Zahlen der Bauart pq bekommen:

$n = pq$	Rechenzeit	Schritte
$(10^7 + 19)(10^7 + 79)$	0 sec	734
$(10^8 + 7)(10^8 + 37)$	0 sec	14073
$(10^9 + 7)(10^9 + 9)$	1 sec	15406
$(10^{10} + 19)(10^{10} + 33)$	2 sec	90027
$(10^{11} + 3)(10^{11} + 19)$	3 sec	235892
$(10^{12} + 39)(10^{12} + 61)$	2 sec	89074
$(10^{13} + 37)(10^{13} + 51)$	10 sec	584003
$(10^{14} + 31)(10^{14} + 67)$	106 sec	5602275
$(10^{15} + 37)(10^{15} + 91)$	909 sec	40772022

Man sieht, dass die Schrittzahl wie $\sqrt{p}$ wächst - gemäß unseren wahrscheinlichkeitstheoretischen Überlegungen.

Bemerkungen:

(1) Das Pollardsche ρ-Verfahren ist nicht deterministisch, d.h. man kann nicht genau vorhersagen, nach wievielen Schritten ein Teiler gefunden wird.
(2) Ist p der kleinste Primteiler von n, so wächst statistisch gesehen die Anzahl der Schritte um p zu finden, wie $\sqrt{p}$.

Beispiel: Um ein Gefühl dafür zu erhalten, wie sich Zahlen im allgemeinen faktorisieren lassen, haben wir das Programm auf die Zahlen $10^{30} + i$ mit $i = 1, \dots, 100$ angewendet (in der Version rho_t.c). Das Ergebnis findet sich in der Tabelle am Schluss des Kapitels. Wir wollen die Tabelle etwas anschauen und definieren zu diesem Zweck:

Sei n eine natürliche Zahl mit der Primfaktorzerlegung $n = p_1^{e_1} \dots p_r^{e_r}$. Dann wird die Anzahl r der verschiedenen Primteiler von n mit $\omega(n)$ bezeichnet.

Bemerkung: Man kann zeigen, dass der durchschnittliche Wert von $\omega(n)$ um n ungefähr $\ln \ln n$ beträgt.

Beispiel: Wir die Zahlen n mit $10^{30} + 1 \leq n \leq 10^{30} + 100$ erhalten wir folgende Verteilung:

ω	1	2	3	4	5	6	7	8
$\#\{10^{30} + 1 \leq n \leq 10^{30} + 100 : \omega(n) = \omega\}$	2	8	20	22	25	12	8	3

Der Mittelwert der Anzahl von $\omega(n)$ beträgt 4.43. Dies passt recht gut zu $\ln \ln 10^{30} \approx 4.24$.

Wir wollen jetzt noch betrachten, wie groß der größte Primteiler einer Zahl durchschnittlich ist.

Heuristische Überlegung: (Dies ist nicht mathematisch exakt.) Sei $n = p_1 p_2 \dots p_r$ $(p_1 \leq \dots \leq p_r)$ eine Zahl mit einer durchschnittlichen Faktorzerlegung, insbesondere $r \approx \ln \ln n$. Dann ist auch $p_1 \dots p_{r-1} = \frac{n}{p_r}$ durchschnittlich und daher $r - 1 = \omega(p_1 \dots p_{r-1}) \approx \ln \ln \frac{n}{p_r}$. Dies ergibt $\ln \ln n - 1 \approx \ln \ln \frac{n}{p_r}$, also $1 \approx \ln \frac{\ln n}{\ln n - \ln p_r}$ und damit $e \approx \frac{\ln n}{\ln n - \ln p_r}$. Aufgelöst ergibt dies

$$\ln p_r \approx (1 - \frac{1}{e}) \ln n \approx 0.63 \ln n$$

oder

$$p_r \approx n^{0.63}.$$

Interpretation: Durchschnittlich hat der größte Primfaktor einer k-stelligen Zahl $0.63k$ Stellen.

Beispiel: Sei p_n der größte Primteiler von n. Für den Mittelwert von $\frac{\ln p_n}{\ln n}$, wo n alle Zahlen obiger Tabelle durchläuft, erhält man dann 0.63, wie es unsere heuristische Überlegung nahelegte.

Überlegung: Für das RSA-Verfahren nimmt man natürliche Zahlen n mit einer Faktorzerlegung $n = pq$, p, q, wo p und q gleiche Größenordnung haben. Das ρ-Verfahren braucht dann statistisch gesehen $c\sqrt{p}$ Schritte, ist also bei großem n unbrauchbar. Da das ρ-Verfahren aber probabilistisch arbeitet, gibt es keine Garantie für eine Mindestlaufzeit. Dies zeigt auch folgendes Beispiel:

Beispiel: Folgende 200-stellige Zahl wird in 7 Schritten (mit $a = 2$, $x_0 = 2$) faktorisiert.

$$\begin{aligned} N &= 27428598456891939993701066266109748264775935125764701398975513736354404148578388 \\ &\quad 01209223046919555550302211051542066344632386893604614257222215748531604399336761 \\ &\quad 13071478884489836950636250489165158387 09 = \\ &= 30334674888300381250713630117576138414293059640646447575108969049807876663476729 \\ &\quad 35773504604009153441 \cdot \\ &\quad 90419951945721131313783691015019807415105527918283251771244025353074677979580700 \\ &\quad 53868031863724136149 \end{aligned}$$

Frage: Gibt es eine Möglichkeit, RSA-Zahlen $n = pq$ zu konstruieren, die sich mit der ρ-Methode sicher nicht (schnell) faktorisieren lassen?

Primfaktorzerlegung der Zahlen $10^{30} + r$ für $1 \leq r \leq 100$:

r	Primfaktorzerlegung von $10^{30} + r$	Zeit
1	$61 \cdot 101 \cdot 3541 \cdot 9901 \cdot 27961 \cdot 4188901 \cdot 39526741$	0 sec
2	$2 \cdot 3 \cdot 43 \cdot 84623843 \cdot 45802327746425579083$	0 sec
3	$1859827 \cdot 537684419034673655130289$	0 sec
4	$2 \cdot 2 \cdot 157 \cdot 18737561 \cdot 1267985441 \cdot 67021328093$	0 sec
5	$3 \cdot 5 \cdot 4723 \cdot 1399606163 \cdot 10085210079364883$	0 sec
6	$2 \cdot 7 \cdot 3919 \cdot 758405810021 \cdot 24032284101871$	8 sec
7	$251897 \cdot 387727 \cdot 10238844796821566353$	0 sec
8	$2 \cdot 2 \cdot 2 \cdot 3 \cdot 3 \cdot 379 \cdot 485899 \cdot 1666666667 \cdot 45251558227$	0 sec
9	$17 \cdot 1021 \cdot 57613642910641239845595437$	0 sec
10	$2 \cdot 5 \cdot 11 \cdot 59 \cdot 154083204930662557781201849$	0 sec
11	$3 \cdot 6275371 \cdot 53117709428388111767947$	0 sec
12	$2 \cdot 2 \cdot 13 \cdot 19 \cdot 97 \cdot 2251 \cdot 4635491895871499269567$	0 sec
13	$7 \cdot 7 \cdot 10641095639 \cdot 1917862968030234283$	3 sec
14	$2 \cdot 3 \cdot 487083044419 \cdot 342173000223132751$	20 sec
15	$5 \cdot 541 \cdot 2659 \cdot 139031879314767479609237$	0 sec
16	$2 \cdot 2 \cdot 2 \cdot 2 \cdot 29 \cdot 3821 \cdot 914813630641 \cdot 616555753529$	5 sec
17	$3 \cdot 3 \cdot 53 \cdot 2339 \cdot 896295878024886551349239$	0 sec
18	$2 \cdot 500000000000000000000000009$	0 sec
19	$2531 \cdot 3323 \cdot 26737 \cdot 28409 \cdot 156534074818411$	0 sec
20	$2 \cdot 2 \cdot 3 \cdot 5 \cdot 7 \cdot 17041 \cdot 445847 \cdot 313378923550840603$	0 sec
21	$11 \cdot 23 \cdot 677 \cdot 58383591875339354627775 41$	0 sec
22	$2 \cdot 897132419 \cdot 557331325243302794969$	1 sec
23	$3 \cdot 288317 \cdot 1156134856194165912288673$	0 sec
24	$2 \cdot 2 \cdot 2 \cdot 193 \cdot 727997 \cdot 4381493 \cdot 203049064151651$	0 sec
25	$5 \cdot 5 \cdot 13 \cdot 15384616923077 \cdot 199999980000001$	110 sec
26	$2 \cdot 3 \cdot 3 \cdot 3 \cdot 17 \cdot 2574211 \cdot 7603259 \cdot 55656179666743$	0 sec
27	$7 \cdot 67 \cdot 73 \cdot 109 \cdot 5737 \cdot 15991 \cdot 2044633 \cdot 1428571429$	0 sec
28	$2 \cdot 2 \cdot 163 \cdot 153374233128834355828 2208589$	0 sec
29	$3 \cdot 20746013 \cdot 16067344281203975594411$	0 sec
30	$2 \cdot 5 \cdot 31 \cdot 10928153 \cdot 2951831340220898 46821$	0 sec
31	$19 \cdot 49995853 \cdot 1052718891452225468633$	1 sec
32	$2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 3 \cdot 11 \cdot 601 \cdot 971 \cdot 166667 \cdot 9736273423807121$	0 sec
33	$79 \cdot 980149 \cdot 397212031 \cdot 32513102512333$	0 sec
34	$2 \cdot 7 \cdot 321403 \cdot 222239902641143450967877$	0 sec
35	$3 \cdot 3 \cdot 5 \cdot 32552664510871 \cdot 682654478707913$	141 sec
36	$2 \cdot 2 \cdot 37 \cdot 1093 \cdot 6181845157142503894562449$	0 sec
37	$198904178873 \cdot 5027546458128958669$	13 sec
38	$2 \cdot 3 \cdot 13 \cdot 13 \cdot 1237 \cdot 74009461 \cdot 10772217263415281$	0 sec
39	$405047 \cdot 665002827539 \cdot 3712539543883$	9 sec
40	$2 \cdot 2 \cdot 2 \cdot 5 \cdot 41 \cdot 16183 \cdot 53731 \cdot 701248902272927557$	0 sec
41	$3 \cdot 7 \cdot 19267 \cdot 13646771 \cdot 142413307 \cdot 1271704279$	0 sec
42	$2 \cdot 47776669 \cdot 10465359148416144290009$	0 sec
43	$11 \cdot 17 \cdot 7245349079 \cdot 738072593132499991$	3 sec
44	$2 \cdot 2 \cdot 3 \cdot 3 \cdot 23 \cdot 10273261 \cdot 117560477495805257593$	0 sec
45	$5 \cdot 29 \cdot 29 \cdot 43 \cdot 47 \cdot 47 \cdot 89 \cdot 2094523 \cdot 13430577524641$	0 sec
46	$2 \cdot 83 \cdot 65250739 \cdot 14145130711 \cdot 6526788089$	1 sec
47	$3 \cdot 100379 \cdot 3320747699552031135330431$	0 sec
48	$2 \cdot 2 \cdot 2 \cdot 2 \cdot 7 \cdot 87113497 \cdot 102493548486194149357$	0 sec
49	$5261 \cdot 19007793195210036114807 0709$	0 sec
50	$2 \cdot 3 \cdot 5 \cdot 5 \cdot 19 \cdot 347 \cdot 577 \cdot 1752467013532462655147$	0 sec

51	$13 \cdot 71 \cdot 523 \cdot 439261127 \cdot 4716000458119597$	1 sec
52	$2 \cdot 2 \cdot 213219257 \cdot 1172501975278902693109$	0 sec
53	$3 \cdot 3 \cdot 3 \cdot 3 \cdot 1451 \cdot 2593 \cdot 3281293301279429086991$	0 sec
54	$2 \cdot 11 \cdot 173 \cdot 439 \cdot 1019 \cdot 23482390033 \cdot 25012103353$	0 sec
55	$5 \cdot 7 \cdot 33073 \cdot 930287 \cdot 974879 \cdot 952556395837$	0 sec
56	$2 \cdot 2 \cdot 2 \cdot 3 \cdot 79043 \cdot 3998741 \cdot 454197539 \cdot 290240017$	0 sec
57	$10000000000000000000000000000057$	0 sec
58	$2 \cdot 5927 \cdot 408479 \cdot 206521534283520766213$	0 sec
59	$3 \cdot 12841021 \cdot 25958475835631242510493$	0 sec
60	$2 \cdot 2 \cdot 5 \cdot 17 \cdot 353 \cdot 8331944675887352107982003$	0 sec
61	$31 \cdot 229 \cdot 1694831 \cdot 279949 \cdot 296891322502381$	0 sec
62	$2 \cdot 3 \cdot 3 \cdot 7 \cdot 7 \cdot 61 \cdot 139 \cdot 436283893 \cdot 306490910893253$	0 sec
63	$1429 \cdot 9377 \cdot 527545979 \cdot 141463219586809$	1 sec
64	$2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 13 \cdot 8317 \cdot 2500000001 \cdot 57805606658281$	2 sec
65	$3 \cdot 5 \cdot 11 \cdot 606060606060606060606060606061$	0 sec
66	$2 \cdot 5000000000000000000000000033$	0 sec
67	$23 \cdot 149 \cdot 283 \cdot 1433 \cdot 30139 \cdot 23873958310352401$	0 sec
68	$2 \cdot 2 \cdot 3 \cdot 131 \cdot 6361323155216284987277353 69$	0 sec
69	$7 \cdot 19 \cdot 59 \cdot 127437237160698356059640627$	0 sec
70	$2 \cdot 5 \cdot 53 \cdot 827 \cdot 228149026944400082133 6497$	0 sec
71	$3 \cdot 3 \cdot 122098703 \cdot 191409233 \cdot 4754266877681$	0 sec
72	$2 \cdot 2 \cdot 2 \cdot 257 \cdot 480260611 \cdot 1012744563716049667$	1 sec
73	$37 \cdot 72534977159 \cdot 372606817918788931$	1 sec
74	$2 \cdot 3 \cdot 29 \cdot 171576747469 \cdot 33495951645895279$	14 sec
75	$5 \cdot 5 \cdot 1831772893 \cdot 21836768167526322271$	2 sec
76	$2 \cdot 2 \cdot 7 \cdot 11 \cdot 443 \cdot 7329014101023130368502829$	0 sec
77	$3 \cdot 13 \cdot 17 \cdot 323679953621 \cdot 4659836387979599$	10 sec
78	$2 \cdot 733 \cdot 28631 \cdot 227467 \cdot 104739648684528079$	0 sec
79	$6091 \cdot 7047989699 \cdot 23294110958062831$	1 sec
80	$2 \cdot 2 \cdot 2 \cdot 2 \cdot 3 \cdot 3 \cdot 3 \cdot 5 \cdot 103 \cdot 1361 \cdot 2647 \cdot 53047 \cdot 23519935674229$	0 sec
81	$41 \cdot 929 \cdot 26254299141484418073459529$	0 sec
82	$2 \cdot 113 \cdot 10607 \cdot 417156477897798331540951$	0 sec
83	$3 \cdot 7 \cdot 476190476190476190476190476 23$	0 sec
84	$2 \cdot 2 \cdot 107 \cdot 10289 \cdot 227082184675949180823727$	0 sec
85	$5 \cdot 2000000000000000000000000000017$	0 sec
86	$2 \cdot 3 \cdot 3083 \cdot 111149 \cdot 109458421 \cdot 4443451929883$	0 sec
87	$11 \cdot 11 \cdot 68494927 \cdot 120658027855367381761$	1 sec
88	$2 \cdot 2 \cdot 2 \cdot 19 \cdot 43 \cdot 607303 \cdot 2519315333693262204 61$	0 sec
89	$3 \cdot 3 \cdot 1597 \cdot 2141 \cdot 1299230293 \cdot 25012077694661$	0 sec
90	$2 \cdot 5 \cdot 7 \cdot 13 \cdot 23 \cdot 6526241 \cdot 12047047 \cdot 607696954219$	0 sec
91	$368148733 \cdot 27162934715301600671 27$	1 sec
92	$2 \cdot 2 \cdot 3 \cdot 31 \cdot 47 \cdot 487 \cdot 117443839530374969993099$	0 sec
93	$605209663 \cdot 1652319949822083392611$	1 sec
94	$2 \cdot 17 \cdot 67 \cdot 1272318379 \cdot 345024932453925887$	0 sec
95	$3 \cdot 5 \cdot 127 \cdot 13417 \cdot 39124572050540339683447$	0 sec
96	$2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 949111 \cdot 2181581 \cdot 37748411 \cdot 399818603$	0 sec
97	$7 \cdot 3319 \cdot 4799 \cdot 8968998629330722469791$	0 sec
98	$2 \cdot 3 \cdot 3 \cdot 11 \cdot 179 \cdot 1931 \cdot 905429153 \cdot 16137826098283$	0 sec
99	$10000000000000000000000000000099$	0 sec
100	$2 \cdot 2 \cdot 5 \cdot 5 \cdot 73 \cdot 137 \cdot 7841 \cdot 1275220010201505037 61$	0 sec