ÜBER EINIGE FUNKTIONALDIFFERENTIALGLEICHUNGEN AUS DER ANALYSE VON ALGORITHMEN

P.Kirschenhofer, H.Prodinger und R.F.Tichy

Abstract. In the analysis of algorithms the study of moment generating functions often leads to functional differential equations. In this paper a number of different cases are treated in a unique framework which allows to establish explicit formulae as well as the asymptotic behaviour of the (factorial) moments.

1. EINLEITUNG

Bei der Analyse von Algorithmen trifft man oft auf folgende Situation: Die Menge A der zu analysierenden Objekte zerfällt vermöge einer "Größenfunktion" in die Klassen $A = \bigcup_{n\geq 0} A_n$ mit den Anzahlen a_n. Wichtige Beispiele sind etwa: A_n, die Familie der Binärbäume mit n Knoten oder A_n, die Familie der Permutationen von n Elementen. Ein gewisser (ganzzahliger) Parameter der gegebenen Objekte aus A_n soll hinsichtlich seiner statistischen Eigenschaften zumindest asymptotisch (für $n\to\infty$) untersucht werden. Dazu betrachtet man die Wahrscheinlichkeiten

$$(1.1) \qquad p_{nk} = \frac{a_{nk}}{a_n},$$

wobei a_{nk} die Anzahl der Objekte in A_n angibt, für die der betrachtete Parameter den Wert k annimmt. Zum Beispiel: a_{nk} ist die Anzahl der Binärbäume mit n Knoten und Pfadlänge k oder a_{nk} ist die Anzahl der Permutationen von n Elementen mit k Inversionen.

In allen hier betrachteten Fällen erweist sich die folgende Vorgangsweise als günstig:

Wir betrachten die wahrscheinlichkeitserzeugende Funktion

$$(1.2) \qquad G_n(z) = \sum_{k\geq 0} p_{nk}\, z^k$$

und versuchen, eine Rekursion für die Folge $G_n(z)$ aufzustellen. Dabei wird die inhaltliche Bedeutung der zumeist rekursiven Algorithmen benützt.

Im nächsten Schritt geht man über zur doppelt-erzeugenden Funktion

(1.3) $$H(z,u) = \sum_{n\geq 0} G_n(z)u^n.$$

Oftmals läßt sich nun die obige Rekursion durch eine Funktionaldifferentialgleichung für $H(z,u)$ wiedergeben. Einige Beispiele dieser Art werden im folgenden näher untersucht.

Sei nun

(1.4) $$\beta_s(n) = \frac{d^s}{dz^s} G_n(z)\Big|_{z=1}$$

das s-te faktorielle Moment des betrachteten zufälligen Parameters. Dann benützen wir zur Bestimmung des asymptotischen Verhaltens der $\beta_s(n)$ die Beziehung

(1.5) $$H(z,u) = \sum_{s\geq 0} f_s(u)(z-1)^s/s!,$$

wobei

(1.6) $$f_s(u) = \sum_{n\geq 0} \beta_s(n)u^n$$

die erzeugende Funktion der s-ten faktoriellen Momente bezeichnet. Es sei bemerkt, daß man die gewöhnlichen Momente auf übliche Weise aus den faktoriellen Momenten gewinnen kann.

In den folgenden Kapiteln betrachten wir jeweils einen speziellen Parameter, behalten jedoch stets die eben eingeführte Notation bei.

$[x^n]f(x)$ bezeichnet den Koeffizienten von x^n in der Potenzreihe $f(x)$.

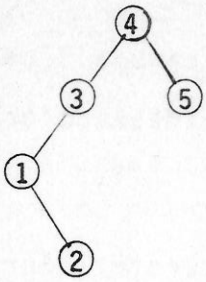
2. PFADLÄNGE IN SUCHBÄUMEN

Ein Suchbaum mit n Knoten ist ein binärer Baum mit n internen Knoten, die mit den Zahlen $1,\ldots,n$ markiert sind: vorgegeben ist dabei eine Permutation π von $1,\ldots,n$; der Suchbaum wird dann rekursiv wie folgt aufgebaut:

die Wurzel ist mit $\pi(1)$ markiert;

sind die mit $\pi(1),\ldots,\pi(j-1)$ bezeichneten Knoten schon vorhanden, so wird der mit $\pi(j)$ bezeichnete Knoten auf folgende Weise gebildet: ist $\pi(j)$ größer (bzw. kleiner) als $\pi(1)$, so wendet man sich nach rechts (bzw. links) und verfährt rekursiv so lange weiter, bis eine freie Stelle im Baum gefunden wird.

Zum Beispiel: $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 1 & 2 \end{pmatrix}$ ergibt den Suchbaum



Ein wichtiger Parameter ist die sogenannte interne Pfadlänge, d.h. die Summe der Abstände aller Knoten von der Wurzel. Im obigen Beispiel ist die interne Pfadlänge 7.

Sei a_{nk} die Anzahl der Permutationen von $1,\ldots,n$, sodaß der zugehörige Suchbaum die interne Pfadlänge k hat. Dann gilt folgende Rekursion:

$$(2.1)\qquad \begin{aligned} a_{nk} &= \sum_{1\le j\le n} \binom{n-1}{j-1} \sum_{l+m+n-1=k} a_{j-1,l}\cdot a_{n-j,m} \qquad (n\ge 1) \\ a_{0k} &= \delta_{0,k}. \end{aligned}$$

Zur Erklärung betrachte man alle Permutationen mit $\pi(1)=j$. Dann müssen in den zugehörigen Suchbäumen die Elemente $1,\ldots,j-1$ im linken Teilbaum und die Elemente $j+1,\ldots,n$ im rechten Teilbaum aufscheinen. Es gibt $\binom{n-1}{j-1}$ Möglichkeiten, aus $2,\ldots,n$ diejenigen $j-1$ Elemente auszuwählen, für die das Bild unter π in $1,2,\ldots,j-1$ liegt. Die Pfadlänge eines Suchbaumes mit n Knoten ergibt sich als Summe der Pfadlängen seiner beiden Teilbäume vermehrt um die Anzahl $n-1$ ihrer Knoten.

Aus (2.1) ergibt sich unmittelbar

$$(2.2)\qquad \begin{aligned} G_n(z) &= \frac{z^{n-1}}{n} \sum_{1\le j\le n} G_{j-1}(z)\, G_{n-j}(z) \qquad (n\ge 1) \\ G_0(z) &= 1\ . \end{aligned}$$

Multipliziert man (2.2) mit nu^{n-1} und summiert man über alle $n\ge 1$, so erhält man (vgl.[13],p.145)

$$\begin{aligned} \sum_{n\ge 1} n\, G_n(z) u^{n-1} &= \sum_{n\ge 1} (uz)^{n-1} \sum_{1\le j\le n} G_{j-1}(z) G_{n-j}(z) \\ &= \sum_{n\ge 0} (uz)^n \sum_{0\le j\le n} G_j(z) G_{n-j}(z), \end{aligned}$$

also

$$(2.3)\qquad \frac{\partial}{\partial u} H(z,u) = H^2(z,zu) \text{ mit } \quad H(1,u) = \frac{1}{1-u}\,.$$

Aus dieser Funktionaldifferentialgleichung erhalten wir wegen

$$f_j(zu) = \sum_{i\geq 0} f_j^{(i)}(u)(z-1)^i\, u^i/i!$$

wobei $f^{(l)}$ die l-te Ableitung von f bedeutet, für die Funktion $f_s(u)$ die lineare Differentialgleichung

$$f_s'(u) - \frac{2}{1-u} f_s(u) = h_s(u) \qquad (s\geq 1) \text{ mit}$$

$$(2.4)\qquad h_s(u) = \sum_{\substack{k_1+k_2+i_1+i_2=s,\\ k_1,k_2\neq s}} \frac{s!}{i_1!i_2!k_1!k_2!} f_{k_1}^{(i_1)}(u) f_{k_2}^{(i_2)}(u) u^{i_1+i_2},$$

$$f_0(u) = \frac{1}{1-u}.$$

Die Lösung von (2.4) läßt sich rekursiv aus der Formel

$$(2.5)\qquad f_s(u) = \frac{1}{(1-u)^2} \int_0^u h_s(t)(1-t)^2\, dt\,.$$

gewinnen. Für kleine Werte von s läßt sich daraus $f_s(u)$ unschwer explizit ermitteln. Zum Beispiel:

$$(2.6)\qquad f_1(u) = \frac{2L(u)}{(1-u)^2} - \frac{2}{(1-u)^2} + \frac{1}{1-u} \text{ mit } \quad L(u) = \log \frac{1}{1-u}\,,$$

sodaß wegen der Formel (vgl.[4],p.14)

$$(2.7)\qquad L(u)(1-u)^{-m-1} = \sum_{n\geq 0} (H_{n+m}-H_m) \binom{n+m}{m} u^n$$

$$\left(H_m = \sum_{1\leq j\leq m} \frac{1}{j}\,, m\geq 0\right)$$

der Erwartungswert der internen Pfadlänge

$$(2.8)\qquad \beta_1(n) = 2(n+1)H_n - 4n$$

beträgt (vgl.z.B. [13]).

Wir zeigen nun durch Induktion, daß $f_s(u)$ allgemein die folgende Bauform besitzt:

$$(2.9) \qquad f_s(u) = \frac{2^s s! L^s(u)}{(1-u)^{s+1}} + \frac{c_s 2^s s! L^{s-1}(u)}{(1-u)^{s+1}} + R_{s-2,s+1}(u),$$

wobei $R_{p,q}(u)$ eine unspezifizierte Linearkombination von Termen der Form $L^i(u)\cdot(1-u)^{-j}$ ist, mit natürlichen Zahlen i,j, für die gilt: $j<q$ und i beliebig, oder $j=q$ und $i\leq p$. Für die Folge c_s ergibt sich aus dem Induktionsbeweis eine Rekursionsformel, die später behandelt wird.

Der Fall s=0 ist trivial mit $c_0=0$. Wir nehmen nun an, daß (2.9) richtig ist für alle $f_j(u)$ mit $0\leq j\leq s-1$. Um (2.9) für $f_s(u)$ zu beweisen, beachten wir:

Für $g(u) = dq!L^p(u)(1-u)^{-q-1}+R_{p-1,q+1}(u)$ (d eine Konstante), so erfüllen die Ableitungen

$$g^{(i)}(u) = d(q+i)!L^p(u)\cdot(1-u)^{-q-i-1}+R_{p-1,q+i+1}(u).$$

Damit erhalten wir für $h_s(u)$ aus (2.4)

$$h_s(u)=s! \sum_{\substack{k_1+k_2=s \\ k_1,k_2\neq s}} \frac{f_{k_1}(u)}{k_1!}\frac{f_{k_2}(u)}{k_2!} +2s! \sum_{k_1+k_2=s-1} \frac{uf'_{k_1}(u)}{k_1!}\frac{f_{k_2}(u)}{k_2!} +R_{s-2,s+2}(u)=$$

$$=2^s s! \sum_{\substack{k_1+k_2=s \\ k_1,k_2\neq s}} \left(\frac{L^{k_1}(u)}{(1-u)^{k_1+1}} +c_{k_1}\frac{L^{k_1-1}(u)}{(1-u)^{k_1+1}}\right)\left(\frac{L^{k_2}(u)}{(1-u)^{k_2+1}} +c_{k_2}\frac{L^{k_2-1}(u)}{(1-u)^{k_2+1}}\right)$$

$$+2^s s! \sum_{k_1+k_2=s-1} u\cdot(k_1+1)\frac{L^{k_1}(u)}{(1-u)^{k_1+2}}\cdot\frac{L^{k_2}(u)}{(1-u)^{k_2+1}} + R_{s-2,s+2}(u) =$$

$$= \frac{2^s s!}{(1-u)^{s+2}}\left((s-1)L^s(u)+\left(\frac{(s+1)s}{2}+2\sum_{k=1}^{s-1}c_k\right)L^{s-1}(u)\right) +R_{s-2,s+2}(u).$$

Daraus erhalten wir mit (2.5)

$$f_s(u)= \frac{2^s s!}{(1-u)^2}\int_0^u \frac{1}{(1-t)^s}\left((s-1)L^s(t)+\left(\frac{(s+1)s}{2} +2\sum_{k=1}^{s-1}c_k\right)L^{s-1}(t)\right)dt+$$

$$+R_{s-2,s+1}(u)= \frac{2^s s! L^s(u)}{(1-u)^{s+1}} +\left(\frac{s}{2} + \frac{2}{s-1}\sum_{k=1}^{s-1}c_k\right)\frac{2^s s! L^{s-1}(u)}{(1-u)^{s+1}} +R_{s-2,s+1}(u),$$

woraus sich (2.9) und die Rekursion

$$(2.10)\qquad (s-1)c_s = \binom{s}{2} + 2\sum_{k=1}^{s-1} c_k\ ,\ c_0=0$$

ergibt. Durch Subtraktion von

$$sc_{s+1} = \binom{s+1}{2} + 2\sum_{k=1}^{s} c_k$$

folgt $\quad sc_{s+1} = (s-1)c_s + 2c_s+s,$

und damit

$$\frac{c_{s+1}}{s+1} = \frac{c_s}{s} + \frac{1}{s+1}\ ,\ c_1 = -1.$$

Durch Aufsummierung ergibt sich als Lösung von (2.10)

$$(2.11)\qquad c_s = s(H_s-2).$$

Zur asymptotischen Auswertung von $\beta_s(n) = [u^n]f_s(u)$ könnte man entweder Taubersche Sätze verwenden (vgl.z.B.[1],[3]) oder die explizite Kenntnis der Taylor-Koeffizienten von Funktionen der folgenden Art (vgl.[14]):

$$L^p(u)\cdot(1-u)^{-q-1} = \sum_{n\geq 0} P_p(H_q^{(1)},\ldots,H_{n+q}^{(p)}-H_q^{(p)})\cdot\binom{n+q}{q}u^n,$$

(2.12) wobei $P_p(s_1,\ldots,s_p)$ definiert ist durch $P_0=1$ und

$$P_p(s_1,\ldots,s_p) = (-1)^p Y_p(-s_1,-s_2,-2s_3,\ldots,-(p-1)!s_p)$$

wobei Y_p das p-te Bellpolynom ist und $H_m^{(j)} = \sum_{1\leq k\leq m} \frac{1}{k^j}$.

Aus (2.9) und (2.12) ergibt sich

$$\beta_s(n)=2^s\cdot s!P_s(H_{n+s}^{(1)}-H_s^{(1)},\ldots,H_{n+s}^{(s)}-H_s^{(s)})\binom{n+s}{s} +$$
$$+ s(H_s-2)2^s s!P_{s-1}(H_{n+s}^{(1)}-H_s^{(1)},\ldots,H_{n+s}^{(s-1)}-H_s^{(s-1)})\binom{n+s}{s} +$$
$$+ O(n^s\cdot\log^{s-2}n),$$

denn

$$P_p(H_{n+s}^{(1)}-H_s^{(1)},\ldots,H_{n+s}^{(p)}-H_s^{(p)}) = O(n^s\cdot H_n^p) = O(n^s\cdot\log^p n).$$

Beachtet man

$$P_p(s_1,\ldots,s_p) = s_1^p - \binom{p}{2}s_1^{p-2}s_2 + \ldots\ ,$$

so ergibt sich

$$\beta_s(n)=\binom{n+s}{s}[2^s s!(H_{n+s}-H_s)^s+2^s s!s(H_s-2)(H_{n+s}-H_s)^{s-1}]+O(n^s\cdot\log^{s-2}n)=$$
$$=2^s n^s[H_{n+s}^s-sH_{n+s}^{s-1}+s(H_s-2)H_{n+s}^{s-1}] + O(n^s\cdot\log^{s-2}n) =$$

$$=2^s n^s[(\log(n+s)+\gamma)^s-2s(\log(n+s)+\gamma)^{s-1}]+O(n^s\cdot\log^{s-2}n)=$$
$$=2^s n^s[\log^s n+s(\gamma-2)\log^{s-1}n]+O(n^s\cdot\log^{s-2}n).$$

(γ die Eulersche Konstante)

Das Ergebnis dieses Kapitels wollen wir im folgenden Satz zusammenfassen:

SATZ. Für das s-te faktorielle Moment $\beta_s(n)$ der internen Pfadlänge von allen binären Suchbäumen, die aus den Permutationen von 1,...,n entstehen, gilt die asymptotische Formel

$$\beta_s(n)=2^s n^s\log^s n + 2^s s(\gamma-2)n^s\log^{s-1}n + O(n^s\cdot\log^{s-2}n),$$

wobei γ = 0,57721... die Eulersche Konstante ist.

3. IN SITU PERMUTATION

Sei $\pi =. \begin{pmatrix} 1 & \dots & n \\ \pi(1) & \dots & \pi(n)\end{pmatrix}$ eine Permutation von 1,...,n; man betrachtet das folgende Programmstück

(3.1)
```
for j:=1 to n do
begin k:=π(j);
      while k>j do
            k:=π(k)
end;
```

Diese Anweisungen können verwendet werden um zu überprüfen, ob j ein *Zyklenführer* ist (d.h. die kleinste Zahl im Zyklus). Dazu muß man nach Verlassen der while-Schleife "k=j?" abfragen.

Die Bestimmung des Zyklenführers ist nützlich, wenn man ein Feld x[1],...,x[n] gemäß der Permutation π umordnen und im wesentlichen ohne weiteren Speicherbedarf auskommen will (in situ Permutation). Für jeden Zyklus $(i_1,\dots,i_k)$ müssen die Elemente $x[i_1],\dots,x[i_k]$ durch $x[\pi(i_1)],\dots,x[\pi(i_k)]$ ersetzt werden. Falls dies genau dann gemacht wird, wenn i_1 ein Zyklenführer ist, wird dies für jeden Zyklus genau einmal durchgeführt. Der vollständige Algorithmus wurde von Mac Leod [10] entwickelt und von Knuth [9] analysiert.

Einer der drei interessanten Parameter dieses Algorithmus wird mit $A(\pi)$ bezeichnet und gibt an, wie oft die Anweisung "k:=π(k)" ausgeführt wird. Mit der in dieser Arbeit verwendeten Technik können nun sämtliche faktoriellen Momente bestimmt werden, wobei alle Permutationen von $1,\dots,n$ als gleichwahrscheinlich angesehen werden. Erwartungswert und Varianz wurden auf andere Weise in [9] berechnet. Wir fassen uns im folgenden kurz und verweisen für eine ausführliche Darstellung auf [6].

Unter Benützung der "kanonischen Zyklenschreibweise $\pi=q(1)\dots q(n)$", wobei die Zyklen nach absteigender Größe der Zyklenführer und jeweils mit den Zyklenführern beginnend unter Weglassen der Klammern aufgeführt werden (vgl.[7],p.176) gilt

(3.2) $A(\pi)=\mathrm{card}\{(i,j):1\le i<j\le n, q(i)<q(k) \text{ für alle } k \text{ mit } i<k\le j\}$.

Man kann sich überlegen, daß $A(\pi)$ auch als "einseitige Pfadlänge" in der Terminologie der binären Suchbäume (vgl.Kapitel 2) gedeutet werden kann.

Wie in der Einleitung vereinbart, sei a_{nk} die Anzahl der Permutationen π von $1,\dots,n$ mit $A(\pi)=k$, und $a_n=n!$. Für die wahrscheinlichkeitserzeugenden Funktionen ergibt sich die Rekursion $(n\ge1)$

$$G_n(z) = \frac{1}{n}\sum_{j=0}^{n-1} z^j G_j(z)G_{n-1-j}(z),\quad G_0(z)=1, \tag{3.3}$$

sowie die Funktionaldifferentialgleichung

$$\begin{aligned} &\frac{\partial}{\partial u} H(z,u) = H(z,u)\cdot H(z,zu) \\ &\text{mit } H(1,u) = \frac{1}{1-u}\,. \end{aligned} \tag{3.4}$$

Wie im 2.Kapitel erhält man für die momenten-erzeugenden Funktionen $(s\ge1)$

$$f_s'(u)-2(1-u)^{-1}f_s(u)=h_s(u), \text{ mit}$$

$$h_s(u)=\sum_{i=1}^{s-1}\binom{s}{i}f_i(u)\sum_{r=0}^{s-i}\binom{s-i}{r}u^r f_{s-r-i}^{(r)}(u)+(1-u)^{-1}\sum_{r=1}^{s}\binom{s}{r}u^r f_{s-r}^{(r)}(u), \tag{3.5}$$

$$f_0(u)=\frac{1}{1-u},\quad h_0(u)= -\frac{1}{(1-u)^2} \text{ und } f_s(0)=0.$$

Wie vorher gilt daher

$$f_s(u)=\frac{1}{(1-u)^2}\int_0^u h_s(t)(1-t)^2\,dt. \tag{3.6}$$

Zum Beispiel erhält man für $s=1$

$$f_1(u) = \frac{L(u)}{(1-u)^2} - \frac{1}{(1-u)^2} + \frac{1}{1-u}\,,$$

und daher für den Erwartungswert

$$(3.7) \qquad \beta_1(n) = (n+1)H_n-2n.$$

Für die momenten-erzeugenden Funktionen kann in Analogie zu (2.9) die Entwicklung

$$(3.8) \quad f_s(u) = \frac{s!L^s(u)}{(1-u)^{s+1}} + \frac{s!s(H_s-2)L^{s-1}(u)}{(1-u)^{s+1}} + R_{s-2,s+1}(u)$$

gezeigt werden, woraus sich der folgende Satz ergibt:

SATZ. Für das s-te faktorielle Moment $\beta_s(n)$ des Parameters $A(\pi)$ aller Permutationen π der Zahlen $1,\dots,n$ gilt die asymptotische Formel

$$\beta_s(n) = n^s\log^s n + s(\gamma-2)n^s\log^{s-1}n + O(n^s\log^{s-2}n).$$

4. DIE ANZAHL DER VERGLEICHE IN QUICKSORT

Quicksort ist ein rekursiver Sortieralgorithmus, der eine Permutation π von $1,\dots,n$ in einem Schritt zu einer Permutation $\rho m\sigma$ umformt, wobei ρ eine Permutation von $1,\dots,(m-1)$ und σ eine Permutation von $(m+1),\dots,n$ bezeichnet. Die "Teilpermutationen" ρ,σ werden nach dem Algorithmus Quicksort in rekursiver Weise weiter behandelt (vgl.[8]). Für die Anzahl $C(\pi)$ der Vergleiche in Quicksort erhält man die Formel

$$(4.1) \qquad C(\pi) = C(\rho)+C(\sigma)+(n+1).$$

Wie in der Einleitung sei a_{nk} die Anzahl der Permutationen π mit $C(\pi)=k$ und $a_n=n!$. Man erhält dann für die wahrscheinlichkeitserzeugenden Funktionen $(n\geq 2)$

$$(4.2) \quad G_n(z) = \frac{z^{n+1}}{n}\sum_{j=0}^{n-1} G_j(z)\cdot G_{n-1-j}(z), \quad G_0(z)=1, G_1(z)=1.$$

Wir gewinnen die Funktionaldifferentialgleichung

$$(4.3) \qquad \begin{aligned} \frac{\partial}{\partial u} H(z,u) &= z^2\, H^2(z,zu) \\ H(1,u) &= \frac{1}{1-u} \end{aligned}$$

und mit der Substitution $\tilde{H}(z,u) = z^2H(z,u)$ die Gleichung

$$(4.4) \qquad \begin{aligned} \frac{\partial}{\partial u} \tilde{H}(z,u) &= \tilde{H}^2(z,zu) \text{ mit} \\ \tilde{H}(1,u) &= \frac{1}{1-u}\ . \end{aligned}$$

Also erfüllt $\tilde{H}(z,u)$ die in Kapitel 2 ausführlich behandelte Funktionaldifferentialgleichung. Wir schreiben

$$\tilde{H}(z,u) = \sum_{s\geq 0} \tilde{f}_s(u) \frac{(z-1)^s}{s!}$$

mit den in Kapitel 2 behandelten momenten-erzeugenden Funktionen $\tilde{f}_s(u)$. Aus der obigen Substitution ergibt sich

$$H(z,u) = \sum_{s\geq 0} f_s(u) \frac{(z-1)^s}{s!} = \frac{\tilde{H}(z,u)}{z^2} = \frac{\tilde{H}(z,u)}{(1+(z-1))^2} =$$

$$= \sum_{i\geq 0} (i+1)! \frac{(z-1)^i}{i!} \sum_{j\geq 0} \tilde{f}_j(u) \frac{(z-1)^j}{j!} ,$$

und daher

$$(4.5) \qquad f_s(u) = \sum_{j=0}^{s} \binom{s}{j} (s-j+1)! \tilde{f}_j(u).$$

Die für die asymptotische Entwicklung relevanten Hauptbeiträge kommen nur vom Summanden $\tilde{f}_s(u)$ in (4.5). Also gilt:

<u>SATZ. Für das s-te faktorielle Moment $\beta_s(n)$ der Anzahl der Vergleiche in Quicksort, wobei alle Permutationen von $1,\ldots,n$ als gleichwahrscheinlich angesehen werden, gilt die asymptotische Formel</u>

$$\beta_s(n) = 2^s n^s \log^s n + 2^s s(\gamma-2) n^s \log^{s-1} n + O(n^s \log^{s-2} n).$$

5. INVERSIONEN

Sei eine Permutation $\pi = \begin{pmatrix} 1 & \ldots & n \\ \pi(1) & \ldots & \pi(n) \end{pmatrix}$ gegeben. Man spricht von einer Inversion, falls $i<j$ und $\pi(i)>\pi(j)$ gilt. Die Anzahl sämtlicher Inversionen von π bezeichnet man mit $I(\pi)$. Diese Größe ist bei einer Reihe von Algorithmen von fundamentaler Bedeutung. Sei also a_{nk} die Anzahl der Permutationen π von $1,\ldots,n$ mit $I(\pi)=k$, sowie $a_n=n!$ Wir wollen im folgenden kurz dartun, daß die Methode der früheren Kapitel auch geeignet ist, die faktoriellen Momente zu bestimmen. Die Resultate sind nicht neu, vgl.[2],[11],[12]; dort wurde jedoch jeweils eine andere Vorgangsweise gewählt.

Es gilt (was auch leicht zu zeigen ist), vgl.[7]

$$(5.1) \quad G_n(z) = \prod_{k=1}^{n} \frac{1-z^k}{k(1-z)} = \frac{1-z^n}{n(1-z)} G_{n-1}(z), n\geq 1, G_0(z)=1,$$

also

$$\sum_{n\geq 1} n(1-z) G_n(z) u^{n-1} = \sum_{n\geq 1} (1-z^n) G_{n-1}(z)$$

oder

$$(5.2)\qquad (1-z)\frac{\partial}{\partial u}H(z,u) = H(z,u) - zH(z,zu)$$

$$H(1,u) = \frac{1}{1-u}\,.$$

Mit einer analogen Vorgangsweise wie in den vorigen Kapiteln erhalten wir für die momenten-erzeugenden Funktionen $(s\geq 1)$

$$(5.3)\qquad f_s'(u) - \frac{1}{1-u}f_s(u) = h_s(u)$$

mit

$$h_s(u) = \frac{1}{1-u}\sum_{j=1}^{s}\binom{s}{j}f_{s-j}^{(j)}u^j + \frac{1}{(s+1)(1-u)}\sum_{j=2}^{s+1}\binom{s+1}{j}f_{s+1-j}^{(j)}u^j,\quad f_0(u) = \frac{1}{1-u}$$

und der Lösung

$$(5.4)\qquad f_s(u) = \frac{1}{1-u}\int_0^u h_s(t)(1-t)dt.$$

Man beweist nun durch Induktion, daß

$$(5.5)\qquad f_s(u) \doteq \frac{(2s)!}{4^s(1-u)^{2s+1}} + \frac{c_s(2s-1)!}{(1-u)^{2s}} + R_{0,2s-1},$$

wobei die Folge c_s die Rekursion

$$(5.6)\qquad c_s = c_{s-1}\cdot\frac{s}{2\cdot(2s-1)} + \frac{s(1-4s)}{3(2s-1)4^{s-1}}\quad\text{mit } c_0=0$$

erfüllt.

Setzt man $d_s=c_s\cdot\binom{2s}{s}$, erhält man aus (5.6)

$$d_s = d_{s-1} + \frac{2}{3\cdot 4^{s-1}}\binom{2s-2}{s-1}(1-4s),$$

sodaß

$$(5.7)\qquad d_s = -\frac{2}{3}\sum_{0\leq t<s}\frac{1}{4^t}\binom{2t}{t}(3+4t).$$

Man benützt nun die Identitäten

$$(5.8)\qquad \sum_{0\leq t<s}\frac{1}{4^t}\binom{2t}{t} = \sum_{0\leq t<s}(-1)^t\binom{-1/2}{t} =$$

$$= (-1)^{s-1}\binom{-3/2}{s-1} = 2s\binom{2s}{s}\frac{1}{4^s}\,,$$

bzw. (ähnlich)

$$(5.9)\qquad \sum_{0\leq t<s}\frac{t}{4^t}\binom{2t}{t} = \frac{2s(s-1)}{3}\binom{2s}{s}\frac{1}{4^s}\,,$$

da allgemein

$$(5.10)\qquad \sum_{0\leq t<s}(-1)^t\binom{a}{t} = \binom{a-1}{s-1}(-1)^{s-1}$$

gilt. Damit erhält man

(5.11) $$c_s = -\frac{s}{9\cdot 4^{s-1}}(4s+5).$$

Die asymptotische Auswertung der Koeffizienten $\beta_s=[u^n]f_s(u)$ ergibt dann nach kurzer Rechnung das folgende Ergebnis.

SATZ. Für das s-te faktorielle Moment $\beta_s(n)$ der Anzahl der Inversionen in einer Permutation von $1,\dots,n$ (wobei alle Permutationen als gleichwahrscheinlich angesehen werden) gilt

$$\beta_s(n) = \frac{1}{4^s}n^{2s} + \frac{s(2s-11)}{9\cdot 4^s}n^{2s-1} + O(n^{2s-2}).$$

6. LINKS-RECHTS-MAXIMA VON PERMUTATIONEN

Sei $\pi = \begin{pmatrix} 1 & \dots & n \\ \pi(1) & \dots & \pi(n)\end{pmatrix}$ eine Permutation von $1,\dots,n$. Unter einem Links-Rechts-Maximum von π versteht man eine Zahl $j=\pi(i)$, sodaß $\pi(k)<j$ für alle $1\le k<i$ gilt. Es bezeichne a_{nk} die Anzahl aller Permutationen von $1,\dots,n$ mit genau k Links-Rechts-Maxima und $a_n=n!$. Dann gilt(vgl.[7;p.176]) für die wahrscheinlichkeitserzeugenden Funktionen:

(6.1) $$G_n(z) = \frac{z+n-1}{n}G_{n-1}(z), n\ge 1, G_0(z)=1.$$

Geht man analog wie in den früheren Kapiteln vor, erhält man die Differentialgleichung

(6.2) $$\frac{\partial}{\partial u}H(z,u) = \frac{z}{1-u}H(z,u),\quad H(1,u) = \frac{1}{1-u}$$

mit der Lösung

(6.3) $$H(z,u) = (1-u)^{-z},$$

was man natürlich auch direkt aus (6.1) gewinnen kann. Nun ist

$$f_s(u) = \frac{\partial^s}{\partial z^s}H(z,u)\bigg|_{z=1} = \frac{1}{1-u}L^s(u).$$

Mit (2.12) ergibt sich daher

(6.4) $$\beta_s(n) = (-1)^s Y_s(-H_n^{(1)}, -H_n^{(2)}, -2H_n^{(3)},\dots,-(s-1)!H_n^{(s)})$$

und somit als Ergebnis:

SATZ. Für das s-te faktorielle Moment $\beta_s(n)$ der Anzahl der Links-Rechts-Maxima in einer Permutation von $1,2,\dots,n$, wobei alle Permutationen von $1,\dots,n$ als gleichwahrscheinlich angesehen werden, gilt

$$\beta_s(n) = \log^s n + \gamma s\log^{s-1}n + O(\log^{s-2}n).$$

Es sei bemerkt, daß der Hauptterm der gewöhnlichen s-ten Momente in [5] studiert wurde.

LITERATUR:

[1] DE BRUIJN N.G., Asymptotic methods in analysis, Dover, New York 1981.

[2] DOBERKAT E.E., Asymptotic estimates for the higher moments of the expected behaviour of straight insertion sort, Informat. Process.Letters 14 (1982), 179-182.

[3] FELLER W., An Introduction to Probability Theory, Vol.II,Wiley, New York 1971.

[4] GREENE D.H. und D.E.KNUTH, Mathematics for the analysis of algorithms, Birkhäuser, Boston 1982.

[5] KEMP R., Fundamentals of the average case analysis of particular algorithms, Wiley-Teubner, Stuttgart 1984.

[6] KIRSCHENHOFER P.,H.PRODINGER und R.F.TICHY, A contribution to the analysis of in situ permutation, preprint, TU Wien (1985).

[7] KNUTH D.E., The art of computer programming, Vol.I, Addison Wesley 1969.

[8] KNUTH D.E., The art of computer programming, Vol.III, Addison Wesley 1973.

[9] KNUTH D.E., Mathematical analysis of algorithms, in: Information Processing 71, 19-27, North Holland 1972.

[10] MAC LEOD L.D.G., An algorithm for in-situ permutation, Austral. Comput.J.2 (1970), 16-19.

[11] PANNY W., A note on the higher moments of the expected behaviour of straight insertion sort, Inform.Process.Letters 22(1986), 175-178.

[12] PRODINGER H., Eine Bemerkung zum asymptotischen Verhalten der höheren Momente der Anzahl der Inversionen, Anz.Österr.Akad. Wiss.Math.Nat.Kl.121(1984), 141-143.

[13] SEDGEWICK R., Mathematical Analysis of Combinatorial Algorithms, in: Probability Theory and Computer Science(G.Louchard, G.Latouche eds.), 123-206, Academic Press, London 1983.

[14] ZAVE D.A., A series expansion involving the harmonic numbers, Inform.Process.Letters 5(1976), 75-77.

Peter Kirschenhofer
Helmut Prodinger
Inst.f.Algebra u.Diskrete Mathematik
der Technischen Universität Wien
Wiedner Hauptstr.8-10
1040 WIEN

Robert F.Tichy
Inst.f.Analysis,Techn.Mathem.
u.Versicherungsmathematik der
Technischen Universität Wien
Wiedner Hauptstr.8-10
1040 WIEN