# Cayleys Formel
# Drei Beweise durch geschicktes Zählen

Marc Wagner
mcwagner@stud.informatik.uni-erlangen.de
Ferienakademie, September 1999

## 1 Vorbetrachtungen

### Labeled Trees (nummerierte Bäume)

Ein Labeled Tree ist ein zusammenhängender, ungerichteter Graph ohne Zyklen, dessen Knoten eindeutig bezeichnet sind. Es gibt weder eine ausgezeichnete Wurzel noch sind die Äste eines Knotens in irgendeiner Weise geordnet. Einen solchen Baum kann man eindeutig beschreiben, indem man zu jedem Knoten die Menge seiner Nachbarn angibt.

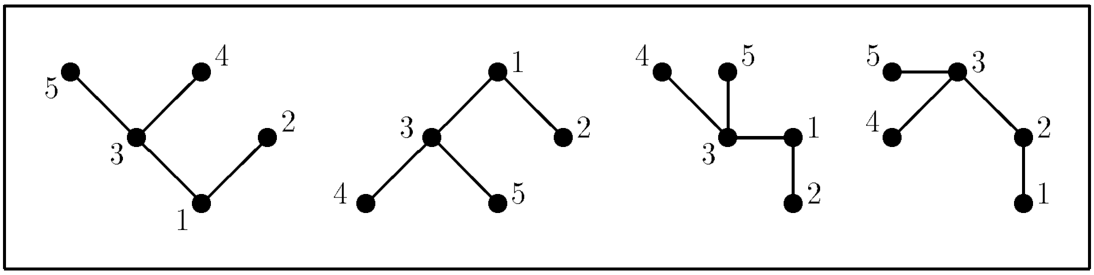


Abbildung 1: Labeled Trees (die ersten drei sind gleich)

### Cayleys Formel

Sei $V = \{1, 2, \ldots, n\}$ die Menge der Knoten. Mit $t_n$ wird die Anzahl der verschiedenen Labeled Trees mit Knotenmenge $V$ bezeichnet. Für kleine $n$ kann man $t_n$ noch relativ einfach durch Konstruktion aller möglichen Bäume und anschliessendes Zählen ermitteln.

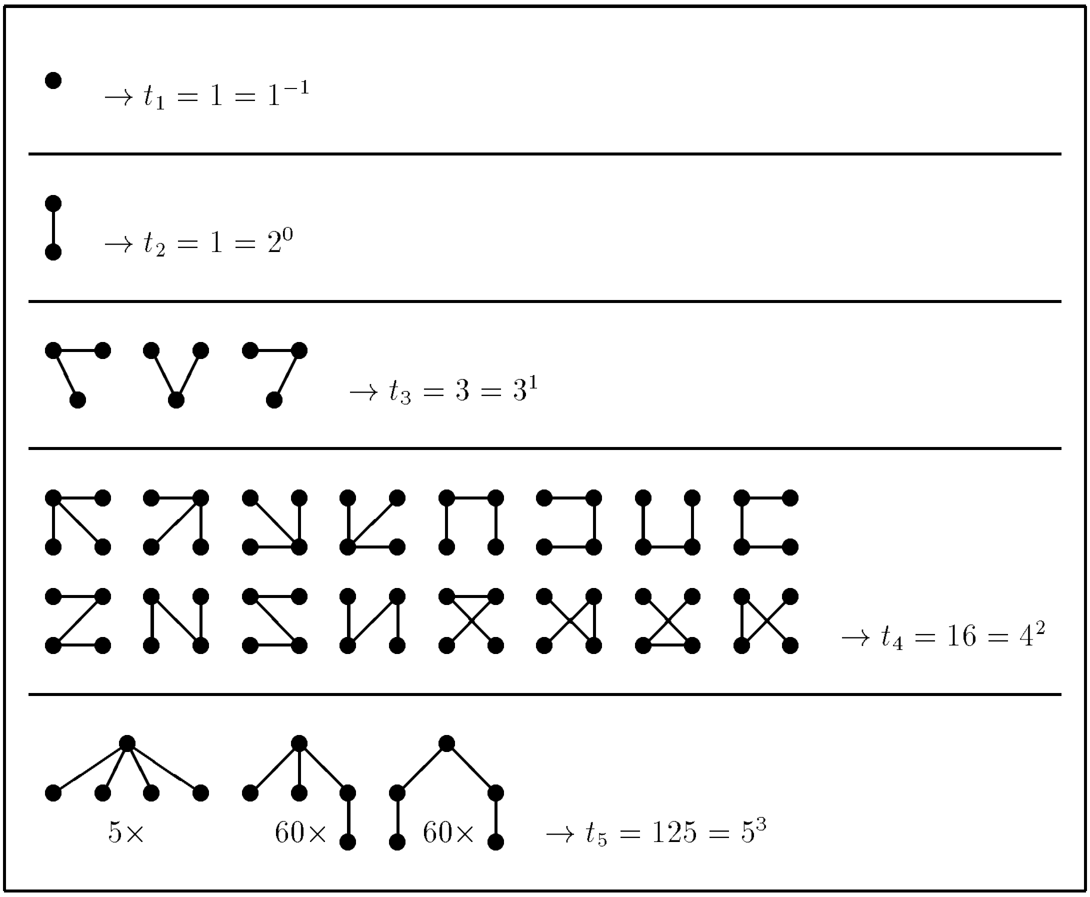


Abbildung 2: Alle Labeled Trees für $n = 1, 2, 3, 4, 5$

Für beliebige $n \geq 1$ liefert Arthur Cayleys berühmte Formel,

$$t_n \quad = \quad n^{n-2} \quad ,$$

konkrete Werte für $t_n$. Es gibt mittlerweile eine ganze Reihe von Beweisen dieser Gleichung, die mit unterschiedlichsten Methoden geführt wurden. Drei davon, basierend auf relativ einfachen Zählargumenten, werden im folgenden präsentiert.

# 2 Beweis durch die Formel von Moon

## Multinomialkoeffizienten

Multinomialkoeffizienten $\binom{n}{r_1, \ldots, r_k}$ werden definiert durch

(1) $$(x_1 + \ldots + x_k)^n \quad = \quad \sum \binom{n}{r_1, \ldots, r_k} x_1^{r_1} \ldots x_k^{r_k} \quad ,$$

wobei über alle $k$-Tupel $(r_1, \ldots, r_k)$ mit ganzzahligen, nichtnegativen $r_i$ summiert wird, deren Summe $n$ beträgt.

Es ist nicht schwierig eine Rekursionsformel für Multinomialkoeffizienten herzuleiten. Die Beziehung

$$(x_1 + \ldots + x_k)^n \quad = \quad (x_1 + \ldots + x_k)^{n-1}(x_1 + \ldots + x_k)$$

gilt offensichtlich. Gleichung (1) eingesetzt ergibt

$$\begin{aligned} &\sum \binom{n}{r_1, \ldots, r_k} x_1^{r_1} \ldots x_k^{r_k} \quad = \\ &= \left[ \sum \binom{n-1}{s_1, \ldots, s_k} x_1^{s_1} \ldots x_k^{s_k} \right] (x_1 + \ldots + x_k) \quad , \end{aligned}$$

wobei links über alle $k$-Tupel $(r_1, \ldots, r_k)$ summiert wird, deren Summe $n$ beträgt, und rechts über alle $k$-Tupel $(s_1, \ldots, s_k)$ summiert wird, deren Summe $n-1$ beträgt. Betrachtet man nur noch ein Glied aus der linken Summe erhält man

$$\begin{aligned} &\binom{n}{r_1, \ldots, r_k} x_1^{r_1} \ldots x_k^{r_k} \quad = \quad \left[ \binom{n-1}{r_1 - 1, \ldots, r_k} x_1^{r_1 - 1} \ldots x_k^{r_k} \right] x_1 + \ldots + \\ &+ \left[ \binom{n-1}{r_1, \ldots, r_k - 1} x_1^{r_1} \ldots x_k^{r_k - 1} \right] x_k \quad , \end{aligned}$$

was zu

$$(2) \qquad \binom{n}{r_1, \ldots, r_k} \quad = \quad \sum_{i=1}^{k} \binom{n-1}{r_1, \ldots, r_i - 1, \ldots, r_k}$$

führt, einer Rekursionsformel die im folgenden Beweis verwendet wird.

Ferner gilt

$$(3) \qquad \binom{n}{r_1, \ldots, r_k} \quad = \quad \binom{n}{r_1, \ldots, r_k, 0} \quad ,$$

wie man unschwer aus Gleichung (1) folgern kann.

## Formel von Moon

Als Grad eines Knotens bezeichnet man die Anzahl der von ihm ausgehenden Kanten. $t(n; d_1, \ldots, d_n)$ sei die Anzahl der verschiedenen Labeled Trees mit der Knotenmenge $V = \{v_1, v_2, \ldots, v_n\}$ für die $d_i = Grad(v_i)$ gilt $(1 \leq i \leq n)$, das heisst jedes $d_i$ ist einem speziellen Knoten zugeordnet. Es ist klar, dass für jeden Baum mit $n$ Knoten $\sum_{i=1}^{n} d_i = 2n - 2$ gilt, da jeder solche Baum genau $n-1$ Kanten hat.

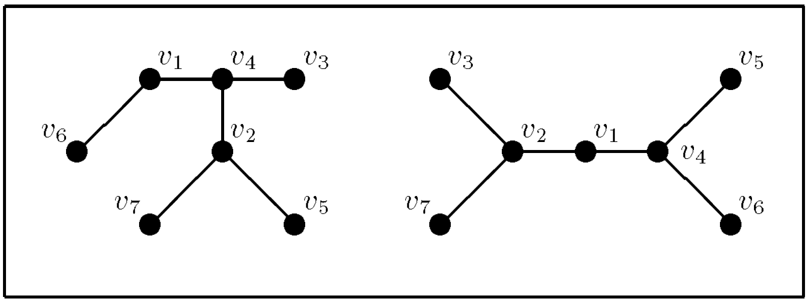


Abbildung 3: Zwei der von $t(7; 2, 3, 1, 3, 1, 1, 1)$ gezählten Bäume

Ohne Beschränkung der Allgemeinheit kann man annehmen $d_1 \geq d_2 \geq \ldots \geq d_n$, da sich der Wert von $t(n; \ldots)$ nicht ändert, wenn man die $d_i$ untereinander vertauscht. Mit Sicherheit ist bei einem Baum, der mehr als einen Knoten hat, $d_n = 1$, denn jeder solche Baum hat mindestens zwei Knoten von denen nur eine Kante ausgeht. $v_n$ ist also mit genau einem anderen Knoten verbunden. Man kann daher die Rekursionsformel

$$(4) \qquad t(n; d_1, \ldots, d_n) \;=\; \sum_{i=1}^{n-1} t(n-1; d_1, \ldots, d_i - 1, \ldots, d_{n-1})$$

aufstellen. Auf der rechten Seite dieser Gleichung werden alle Bäume berücksichtigt, die durch entfernen des Knotens $v_n$ und dessen einziger Kante aus den von $t(n; d_1, \ldots, d_n)$ gezählten Bäumen entstehen können. Die Summe läuft von 1 bis $n-1$, weil $v_n$ mit jedem der anderen Knoten $v_1$ bis $v_{n-1}$ verbunden sein kann, wobei der Grad des entsprechenden Knotens wegen der nicht zu betrachtenden Kante um 1 verringert werden muss.

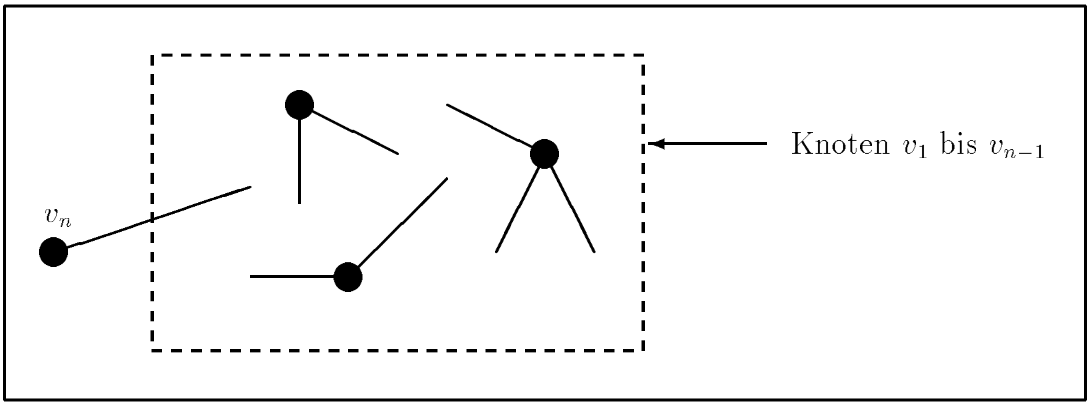


Abbildung 4: Zur Veranschaulichung von Gleichung (4)

Die Formel von Moon,

$$t(n; d_1, \ldots, d_n) \;=\; \binom{n-2}{d_1 - 1, \ldots, d_n - 1} \quad ,$$

für $n \geq 3$, $d_i \geq 1$ lässt sich nun leicht durch vollständige Induktion beweisen.

### Induktionsanfang

Induktionsanfang bei $n = 3$. Ein Baum mit drei Knoten hat genau zwei Kanten. Daraus folgt $\sum_{i=1}^{3} d_i = 4$. Man verifiziert nun problemlos

$$\begin{aligned}
t(3; 2, 1, 1) &= 1 = \binom{1}{1, 0, 0} \quad, \\
t(3; 1, 2, 1) &= 1 = \binom{1}{0, 1, 0} \quad, \\
t(3; 1, 1, 2) &= 1 = \binom{1}{0, 0, 1} \quad.
\end{aligned}$$

### Induktionsannahme

Die Formel von Moon gilt für Bäume mit $n - 1$ Knoten.

### Induktionsschluss

$$\begin{aligned}
& t(n; d_1, \ldots, d_{n-1}, d_n) = \\
&= t(n; d_1, \ldots, d_{n-1}, 1) = && \mid d_n = 1 \text{ da } d_1 \geq d_2 \geq \ldots \geq d_n \\
&= \sum_{i=1}^{n-1} t(n-1; d_1, \ldots, d_i - 1, \ldots, d_{n-1}) = && \mid \text{Gleichung (4)} \\
&= \sum_{i=1}^{n-1} \binom{n-3}{d_1 - 1, \ldots, d_i - 2, \ldots, d_{n-1} - 1} = && \mid \text{Induktionsannahme} \\
&= \binom{n-2}{d_1 - 1, \ldots, d_{n-1} - 1} = && \mid \text{Gleichung (2)} \\
&= \binom{n-2}{d_1 - 1, \ldots, d_{n-1} - 1, 0} = && \mid \text{Gleichung (3)} \\
&= \binom{n-2}{d_1 - 1, \ldots, d_{n-1} - 1, d_n - 1} && \mid d_n = 1 \text{ da } d_1 \geq d_2 \geq \ldots \geq d_n
\end{aligned}$$

## Beweis von Cayleys Formel

In Gleichung (1) wird nun $n$ durch $n-2$, $k$ durch $n$, $r_i$ durch $d_i - 1$ und $x_i$ durch 1 ersetzt.

$$(x_1 + \ldots + x_k)^n = \sum \binom{n}{r_1, \ldots, r_k} x_1^{r_1} \ldots x_k^{r_k} \qquad \mid \text{Gleichung (1)}$$

$$n^{n-2} = \sum \binom{n-2}{d_1 - 1, \ldots, d_n - 1} \qquad | \text{ Ersetzen (s. o.)}$$

$$n^{n-2} = \sum t(n; d_1, \ldots, d_n) \quad \text{für } n \geq 3 \qquad | \text{ Formel von Moon} \tag{5}$$

Während in der obersten Gleichung noch über alle $k$-Tupel $(r_1, \ldots, r_k)$ summiert wird, deren Summe $n$ beträgt, wird, bedingt durch die Ersetzungen, in der untersten Gleichung über alle $n$-Tupel $(d_1, \ldots, d_n)$ summiert, deren Summe $2n-2$ beträgt.

Wie bereits erwähnt gilt für einen Baum mit $n$ Knoten $\sum_{i=1}^{n} d_i = 2n - 2$. Daraus folgt

$$t_n = \sum t(n; d_1, \ldots, d_n) \tag{6}$$

wenn über alle $n$-Tupel $(d_1, \ldots, d_n)$ summiert wird, deren Summe $2n-2$ beträgt.

Die Gleichungen (5) und (6) führen sofort auf

$$t_n = n^{n-2}$$

für $n \geq 3$. Da die Korrektheit dieser Beziehung für $n = 1$ und $n = 2$ leicht von Hand verfizierbar ist, folgt

$$t_n = n^{n-2} \quad ,$$

Cayleys Formel.

# 3 Beweis durch Rekursion von Riordan und Rényi

## Aufstellen einer rekursiven Formel

Sei $A$ eine beliebige $k$-elementige Teilmenge der Knotenmenge $V = \{1, 2, \ldots, n\}$. Mit $t_{n,k}$ wird die Anzahl der verschiedenen Planted Forests (gewurzelte Wälder) mit Knotenmenge $V$ bezeichnet, bestehend aus $k$ Bäumen, bei denen die $k$ Wurzeln gerade die Menge $A$ bilden.

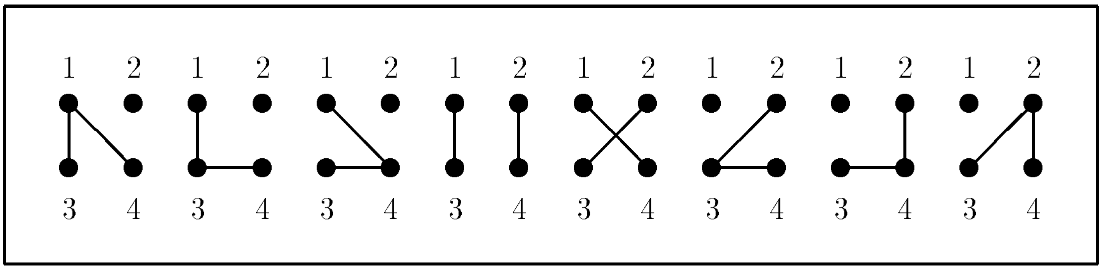


Abbildung 5: $t_{4,2} = 8$ für $A = \{1, 2\}$

Es ist offensichtlich, dass es nicht auf einzelnen Elemente der Menge $A$ ankommt, sondern nur auf deren Mächtigkeit.

Im folgenden wird gezeigt, dass die rekursive Beziehung

(7)
$$t_{n,k} = \underbrace{\sum_{i=0}^{n-k}}_{\alpha)} \underbrace{\binom{n-k}{i}}_{\beta)} \underbrace{t_{n-1,k-1+i}}_{\gamma)}$$

für $n \geq k \geq 1$ gilt.

$\alpha)$ Sei ohne Beschränkung der Allgemeinheit $A = \{1, \ldots, k\}$. Der Knoten 1 kann mit $i$ Knoten, $0 \leq i \leq n-k$, durch jeweils eine Kante verbunden sein.

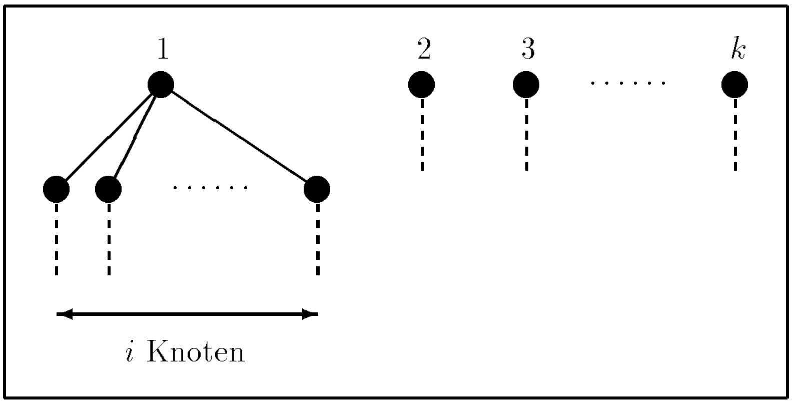


Abbildung 6: Zur Verdeutlichung von $\alpha)$

$\beta)$ Es gibt $\binom{n-k}{i}$ Möglichkeiten die $i$ Nachbarn von Knoten 1 aus der Knotenmenge $\{k+1, \ldots, n\}$ auszuwählen.

$\gamma)$ Entfernt man Knoten 1 und alle seine Kanten erhält man einen Planted Forest mit $n-1$ Knoten, bestehend aus $k-1+i$ Bäumen, dessen Wurzeln gerade die Knoten 2 bis $k$ und die $i$ Nachbarn von Knoten 1 sind. Durch diese Knoten wird automatisch eine neue Menge $A$ festgelegt.

Damit die Rekursion endet wird $t_{0,0} = 1$ (es gibt genau einen Wald mit 0 Knoten und 0 Bäumen) und $t_{n,0} = 0$ für $n > 0$ (es gibt keinen Wald mit einer nichtleeren Knotenmenge und 0 Bäumen) gesetzt.

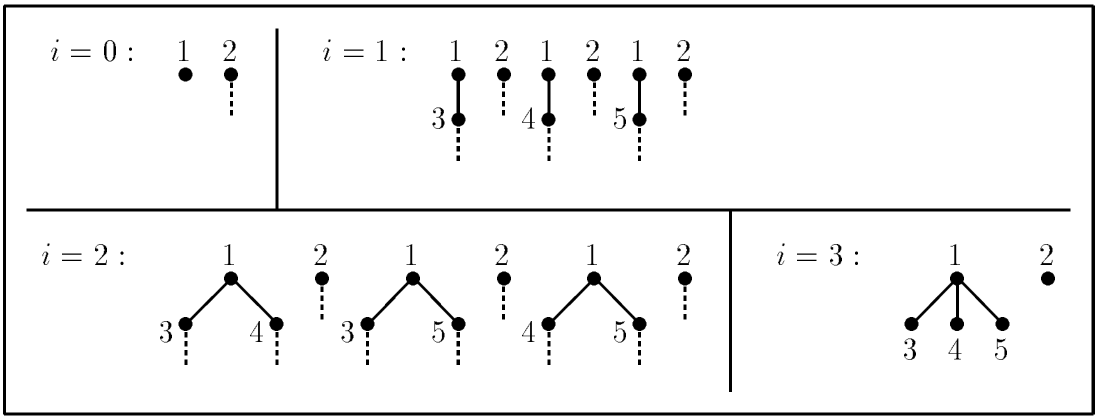


Abbildung 7: Der erste Rekursionsschritt für $n = 5$ und $A = \{1, 2\}$

## Beweis von Cayleys Formel

Die Gleichung

$$t_{n,k} = kn^{n-k-1} \tag{8}$$

für $n \geq k \geq 1$ wird durch vollständige Induktion bewiesen.

### Induktionsanfang

Induktionsanfang bei $n = 1$. $t_{1,0} = 0$ und $t_{1,1} = 1$ ist offensichtlich korrekt.

### Induktionsannahme

Gleichung (8) gilt für $n - 1$.

### Induktionsschluss

$$t_{n,k} = \sum_{i=0}^{n-k} \binom{n-k}{i} t_{n-1,k-1+i} = \qquad | \text{ Gleichung (7)}$$
$$= \sum_{i=0}^{n-k} \binom{n-k}{i} (k-1+i)(n-1)^{n-1-k-i} = \qquad | \text{ Induktionsannahme}$$

$$
\begin{aligned}
&= \sum_{i=n-k}^{0} \binom{n-k}{n-k-i} (n-1-i)(n-1)^{i-1} \;= && \mid i \to n-k-i \\
&= \sum_{i=0}^{n-k} \binom{n-k}{i} (n-1)^{i} - \sum_{i=0}^{n-k} \binom{n-k}{i} i(n-1)^{i-1} \;= \\
&= \sum_{i=0}^{n-k} \binom{n-k}{i} (n-1)^{i} 1^{n-k-i} - \sum_{i=1}^{n-k} \frac{(n-k)!}{i!(n-k-i)!} i(n-1)^{i-1} \;= \\
&= (n-1+1)^{n-k} - \\
&- (n-k) \sum_{i=1}^{n-k} \frac{(n-k-1)!}{(i-1)!(n-k-i)!} (n-1)^{i-1} \;= && \mid \text{binomischer Satz} \\
&= n^{n-k} - (n-k) \sum_{i=1}^{n-k} \binom{n-k-1}{i-1} (n-1)^{i-1} \;= \\
&= n^{n-k} - (n-k) \sum_{i=0}^{n-k-1} \binom{n-k-1}{i} (n-1)^{i} 1^{n-k-1-i} \;= && \mid i \to i+1 \\
&= n^{n-k} - (n-k)(n-1+1)^{n-k-1} \;= && \mid \text{binomischer Satz} \\
&= n^{n-k} - (n-k)n^{n-k-1} \;=\; kn^{n-k-1}
\end{aligned}
$$

Setzt man in Gleichung (8) $k = 1$ erhält man

$$t_{n,1} \;=\; n^{n-2} \quad .$$

Man sieht problemlos, dass $t_{n,1} = t_n$. Diese Erkenntnis führt direkt zu

$$t_n \;=\; n^{n-2} \quad ,$$

Cayleys Formel.

# 4 Beweis durch doppeltes Zählen

Dieser von Jim Pitman geführte Beweis von Cayleys Formel ist vermutlich der eleganteste von allen bisher bekannten.

## Planted Forests

Ein Planted Forest ist ein Labeled Forest, bei dem jeder Baum einen ausgezeichneten Wurzelknoten enthält. Sei $\mathcal{F}_{n,k}$ die Menge aller Planted Forests mit $n$ Knoten und $k$ Bäumen. Dann ist $\mathcal{F}_{n,1}$ die Menge aller Planted Trees. Sicher ist $|\mathcal{F}_{n,1}| = nt_n$, denn zu jedem Labeled Tree mit $n$ Knoten gibt es $n$ Möglichkeiten einen Wurzelknoten auszuwählen.

Ein Planted Forest $F_{n,k} \in \mathcal{F}_{n,k}$ kann auch als gerichteter Graph betrachtet werden, bei dem alle Kanten von den Wurzeln weg gerichtet sind. Ein Wald $F_{n,i}$ enthält einen Wald $F_{n,j}$ falls alle gerichteten Kanten von $F_{n,j}$ auch in $F_{n,i}$ vorhanden sind. In einem solchen Fall ist $i \leq j$, mit anderen Worten $F_{n,i}$ hat höchstens so viele Bäume wie $F_{n,j}$.

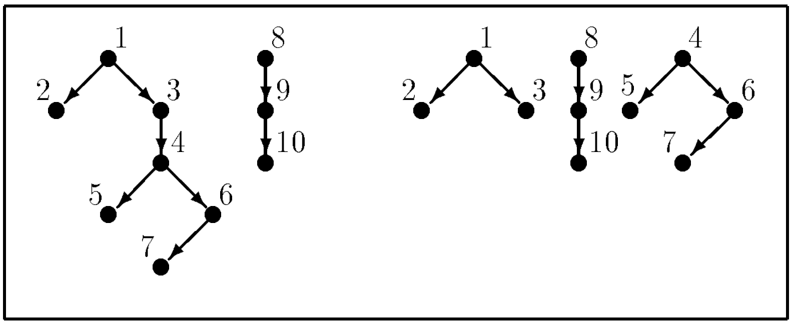


Abbildung 8: Der linke Planted Forest enthält den rechten

## Verfeinerungssequenzen

Eine Folge $F_1, F_2, \ldots, F_k$ von Planted Forests wird als Verfeinerungssequenz bezeichnet, falls für alle $i$ gilt: $F_i \in \mathcal{F}_{n,i}$ und $F_i$ enthält $F_{i+1}$. Bei vorgegebenem $F_k \in \mathcal{F}_{n,k}$ sei

- $N(F_k)$ die Anzahl der verschiedenen Planted Trees, die $F_k$ enthalten und
- $N^*(F_k)$ die Anzahl der verschiedenen Verfeinerungssequenzen, die mit $F_k$ enden.

$N^*(F_k)$ kann auf zwei Arten ermittelt werden, je nachdem ob man mit den Überlegungen bei $F_1$ oder bei $F_k$ beginnt.

### Start bei $F_1$

Bei einem Planted Tree, der $F_k$ enthält, können die $k-1$ Kanten, die nicht zu $F_k$ gehören, in beliebiger Reihenfolge, also auf $(k-1)!$ Möglichkeiten, entfernt werden, um eine Verfeinerungssequenz zu bekommen, die mit $F_k$ endet. Da es $N(F_k)$ verschiedene Planted Trees gibt, die $F_k$ enthalten, folgt

$$(9) \qquad N^*(F_k) \quad = \quad N(F_k)(k-1)! \quad .$$

**Start bei $F_k$**

Um von $F_k$ nach $F_{k-1}$ zu kommen muss eine gerichtete Kante von einem beliebigen Knoten, dafür gibt es $n$ Möglichkeiten, zu einer Wurzel in einem anderen Baum, es bleiben $k-1$ Möglichkeiten, eingefügt werden. Folglich kann dies auf $n(k-1)$ Arten geschehen.

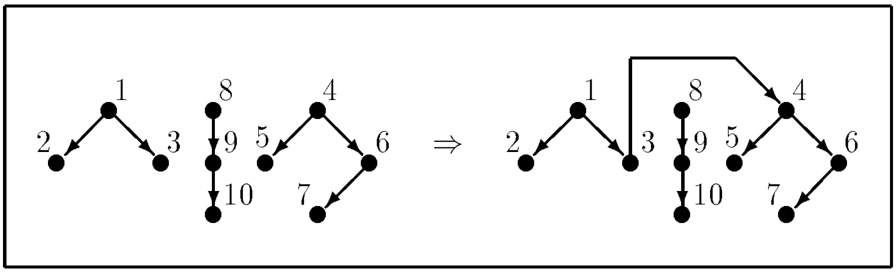


Abbildung 9: Einfügen einer Kante führt von $F_3$ zu $F_2$

Analog ermittelt man $n(k-2)$ Alternativen eine Kante einzufügen, die von $F_{k-1}$ nach $F_{k-2}$ führt. Wendet man diese Überlegung $k-1$ mal an, erhält man schliesslich

$$N^*(F_k) \quad = \quad n^{k-1}(k-1)! \quad . \tag{10}$$

## Beweis von Cayleys Formel

Die Gleichungen (9) und (10) führen sofort auf die Beziehung

$$N(F_k) \quad = \quad n^{k-1} \tag{11}$$

für $F_k \in \mathfrak{F}_{n,k}$. $F_n$ kann nur aus $n$ isolierten Knoten bestehen, ist also eindeutig festgelegt. Daher entspricht $N(F_n)$ der Anzahl aller Planted Trees die genau $n$ Knoten enthalten, also $N(F_n) = |\mathfrak{F}_{n,1}|$. Wie bereits festgestellt ist $|\mathfrak{F}_{n,1}| = nt_n$. Daraus kann man

$$N(F_n) \quad = \quad nt_n$$

folgern und erhält dann unter Verwendung von Gleichung (11)

$$t_n \quad = \quad n^{n-2} \quad ,$$

Cayleys Formel.

## 5 Literaturempfehlung

Weitere Informationen zu diesem und ähnlichen Themen enthalten die Bücher

- Proofs from the Book, M. Aigner, G. M. Ziegler, Springer 1998 und
- A Course in Combinatorics, J. H. van Lint, R. M. Wilson, Cambridge UP 1992.