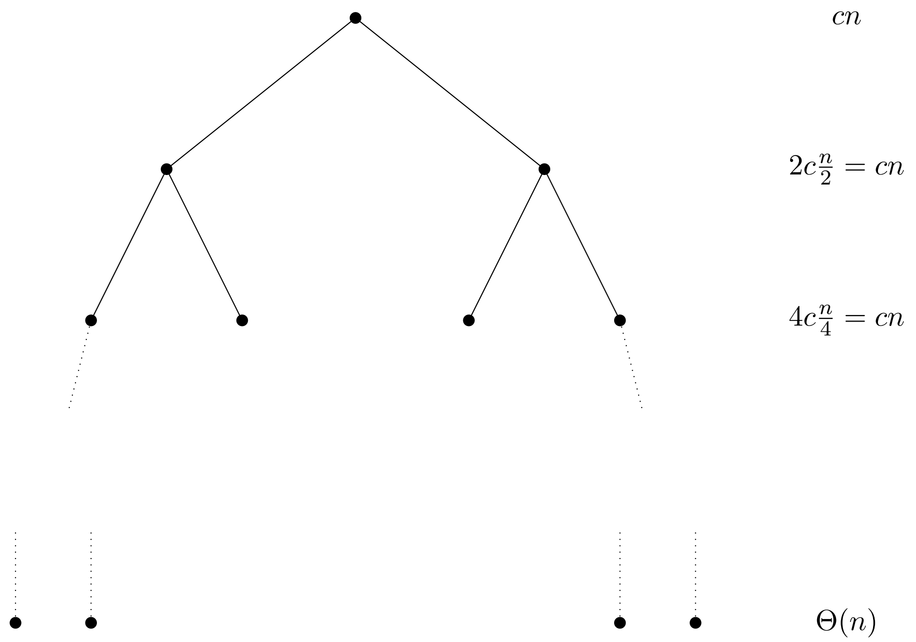

Abbildung 4.1: Rekursionsbaum von $T(n) = 2T(\frac{n}{2}) + cn$

4.4 Die Rekursionsbaummethode

Die Rekursionsbaummethode besteht im Aufzeichnen des durch eine Rekursionsgleichung induzierten Rekursionsbaums sowie in der anschließenden Summierung der im gesamten Baum anfallenden Kosten. Sie wird oft verwendet, um zu einer Vermutung zu gelangen, die danach mit der Substitutionsmethode bewiesen wird. Deshalb wird mit Rekursionsbäumen oft recht ungenau gearbeitet, z.B. lässt man die Rundungsoperatoren $\lfloor \cdot \rfloor$ und $\lceil \cdot \rceil$ weg oder man ersetzt Ausdrücke wie $\Theta(f(n))$ durch $cf(n)$.

Beispiel 4.6. Die Rekursion von Sortieren durch Verschmelzen war

$$T(n) = \begin{cases} \Theta(1) & \text{für } n = 1 \\ T(\lceil \frac{n}{2} \rceil) + T(\lfloor \frac{n}{2} \rfloor) + \Theta(n) & \text{für } n \geq 2 \end{cases}.$$

Wir vereinfachen diese zu $T(n) = 2T(\frac{n}{2}) + cn$ und zeichnen den Rekursionsbaum auf, siehe Abbildung 4.1.

In der ersten Ebene, d.h. am Wurzelknoten, treten Kosten von cn auf. In der zweiten Ebene treten an jedem Knoten Kosten von $c\frac{n}{2}$ auf. Da es in der zweiten Ebene zwei Knoten gibt, treten insgesamt in der zweiten Ebene Kosten von cn auf. In der dritten Ebene gibt es vier Konten mit Kosten jeweils $c\frac{n}{4}$, also hat auch diese Ebene Kosten von cn, usw. Jede Ebene hat also Kosten cn. Wie viele Ebenen gibt es? Bis n auf eine Konstante (z.B. konkret auf 1) reduziert wurde benötigen wir $\log n$ Ebenen. Der Rekursionsbaum ist also ein vollständiger Binärbaum der Tiefe $\log n$ wobei wir auch hier, im Sinne einer vereinfachenden Annahme, die Tatsache vernachlässigen dass $\log n$ nicht notwendigerweise eine natürliche Zahl ist und nicht klar ist was mit einem Baum reeller Tiefe gemeint sein soll. Ein solcher Baum hat $\Theta(n)$ Blätter. So kommen wir also zur Vermutung $T(n) = cn \log n + \Theta(n) = \Theta(n \log n)$.

Beispiel 4.7. Betrachten wir nun die vereinfachte Rekursionsgleichung

$$T(n) = 3T(\frac{n}{4}) + cn^2$$

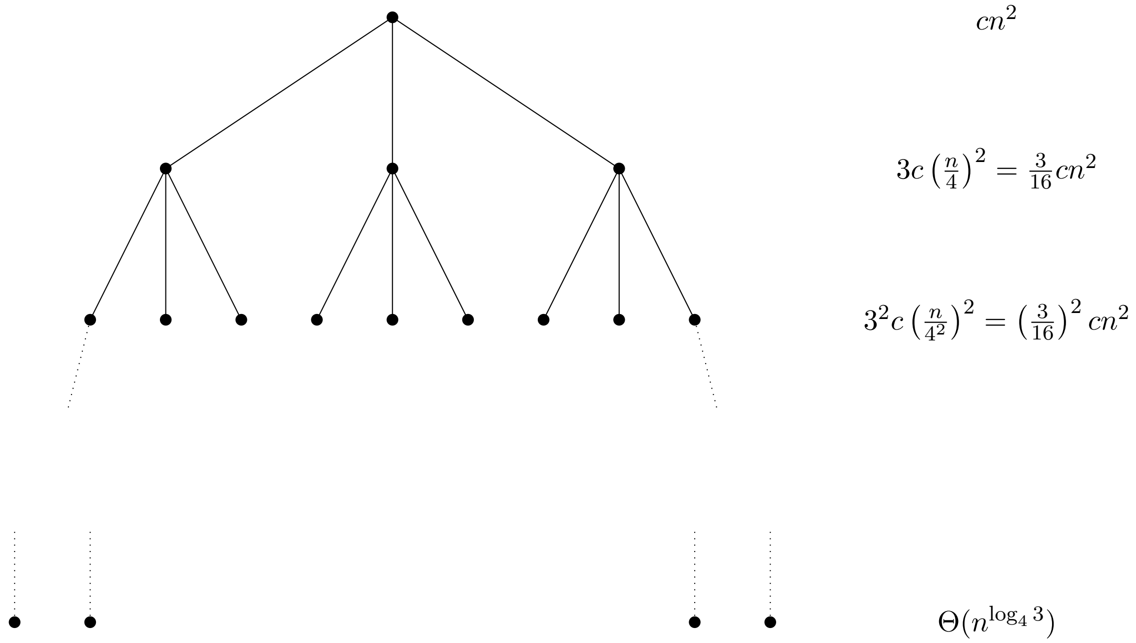

Abbildung 4.2: Rekursionsbaum von $T(n) = 3T(\frac{n}{4}) + cn^2$

Der Rekursionsbaum für diese Gleichung ist in Abbildung 4.2 zu finden. Die erste Ebene hat Kosten cn^2, die zweite $\frac{3}{16}cn^2$, die dritte $\left(\frac{3}{16}\right)^2 cn^2$ usw. Dieser Rekursionsbaum ist ein vollständiger Baum der Arität 3 mit Tiefe $\log_4 n$, er hat also $3^{\log_4 n} = n^{\log_4 3}$ Blätter. Nun beobachten wir dass

$$\sum_{i=0}^{\log_4 n-1} \left(\frac{3}{16}\right)^i cn^2 + n^{\log_4 3} < cn^2 \sum_{i=0}^{\infty} \left(\frac{3}{16}\right)^i + n^{\log_4 3} = cn^2 \frac{1}{1-\frac{3}{16}} + n^{\log_4 3} = O(n^2)$$

und erhalten somit die Vermutung $T(n) = O(n^2)$.

4.5 Die Mastermethode

Summenbildungen über Rekursionsbäume, wie sie in den beiden Beispielen im vorherigen Abschnitt auftreten, folgen immer dem selben Muster. Wir werden nun einen allgemeinen Satz über solche Summen beweisen, der dann erlaubt die Lösungen der allermeisten Teile-und-herrsche Rekursionsgleichungen als Korollare zu erhalten. Dazu betrachten wir $a_0, a_1 \in \mathbb{N}$ mit $a = a_0 + a_1 \geq 1$, $b > 1$ und eine Rekursionsgleichung der Form

$$T(n) = a_0 T(\left\lfloor \frac{n}{b} \right\rfloor) + a_1 T(\left\lceil \frac{n}{b} \right\rceil) + f(n) \tag{4.3}$$

Da wir uns nur für asymptotische Lösungen interessieren, entfällt die Angabe von Anfangswerten. Der Anfang des Rekursionsbaums der Gleichung (4.3) ist in Abbildung 4.3 dargestellt. Um das iterierte auf- und abrunden darzustellen verwenden wir Zeichenketten im Alphabet $\{0, 1\}$ wobei 0 für Abrunden und 1 für Aufrunden steht.

Definition 4.6. $\{0, 1\}^*$ ist die Menge aller endlich langen aus 0 und 1 bestehenden Zeichenketten (Wörtern) inklusive dem Leerwort ε. Die Länge $|w|$ eines Wortes $w \in \{0, 1\}^*$ ist die

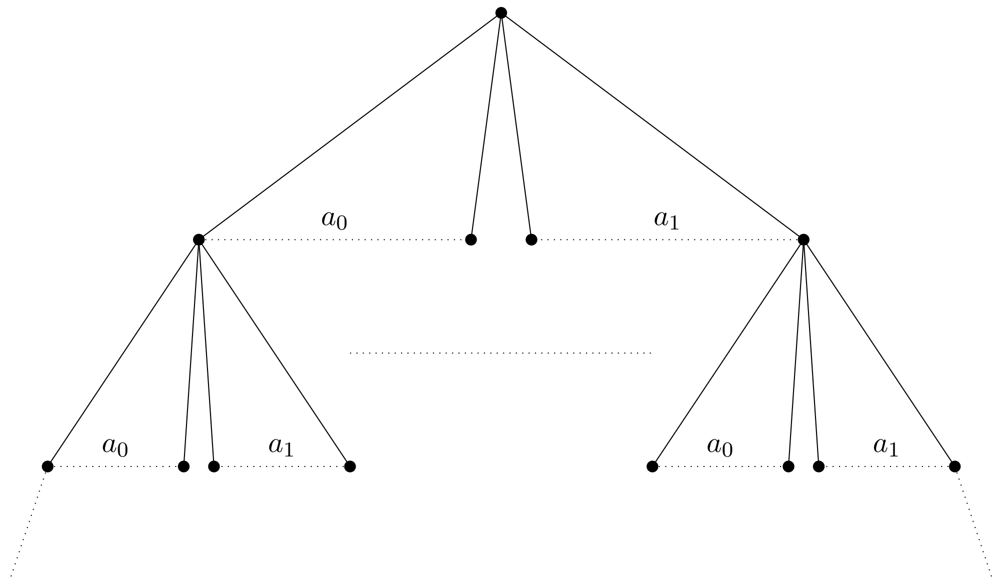

Abbildung 4.3: Anfang des Rekursionsbaums von Gleichung (4.3)

Anzahl der Zeichen aus denen w besteht. Die Länge des Leerwortes ist $|\varepsilon| = 0$. Für $w \in \{0,1\}^*$ bezeichnet $n_0(w)$ die Anzahl von Nullern in w und $n_1(w)$ die Anzahl von Einsern in w. Für einen Buchstaben x und ein $i \in \mathbb{N}$ bezeichnet x^i das Wort das aus i Vorkommen des Buchstaben x besteht.

Definition 4.7. Sei $b > 1$. Wir definieren für $n \in \mathbb{N}$ und $w \in \{0,1\}^*$ die Zahl $n_w \in \mathbb{N}$ als

$$n_w = \begin{cases} n & \text{falls } w = \varepsilon \\ \left\lfloor \frac{n_v}{b} \right\rfloor & \text{falls } w = 0v \\ \left\lceil \frac{n_v}{b} \right\rceil & \text{falls } w = 1v \end{cases}$$

Jeder Knoten im Baum entspricht also T angewandt auf ein bestimmtes n_w. Ein w kommt im Baum an $a_0^{n_0(w)} a_1^{n_1(w)}$ vielen Stellen vor. Die lokalen Kosten an einem Knoten der $T(n_w)$ entspricht sind $f(n_w)$. Wir wollen diese lokalen Kosten über alle Knoten im Baum summieren.

Wie tief ist dieser Baum? Falls n eine Potenz von b ist, und somit niemals gerundet wird, ist klar: nach $\log_b n$ Schritten in die Tiefe ist $T(1)$ erreicht. Aber auch der Fall wo n keine Potenz von b ist, ist nicht wesentlich komplizierter: n_w verhält sich bis auf eine additive Konstante so wie $\frac{n}{b^{|w|}}$, siehe Abbildung 4.4.

Lemma 4.2. *Sei $b > 1$, dann existiert ein $c > 0$ so dass für alle $n \in \mathbb{N}$ und $w \in \{0,1\}^*$: $\frac{n}{b^{|w|}} - c < n_w < \frac{n}{b^{|w|}} + c$.*

Beweis. Sei $|w| = j$, dann ist $n_{0^j} \leq n_w \leq n_{1^j}$. Wir behaupten weiters dass

$$n_{0^j} \geq \frac{n}{b^j} - \sum_{i=0}^{j-1} \frac{1}{b^i} \quad \text{und} \quad n_{1^j} \leq \frac{n}{b^j} + \sum_{i=0}^{j-1} \frac{1}{b^i}.$$

Wir zeigen die Ungleichung für n_{1^j} mit Induktion nach j. Falls $j = 0$, dann ist $n_\varepsilon = n$. Für den Induktionsschritt gilt

$$n_{1^{j+1}} = \left\lceil \frac{n_{1^j}}{b} \right\rceil \leq \frac{n_{1^j}}{b} + 1 \leq^{\text{IH}} \frac{1}{b} \left(\frac{n}{b^j} + \sum_{i=0}^{j-1} \frac{1}{b^i} \right) + 1 = \frac{n}{b^{j+1}} + \sum_{i=0}^{j} \frac{1}{b^i}.$$

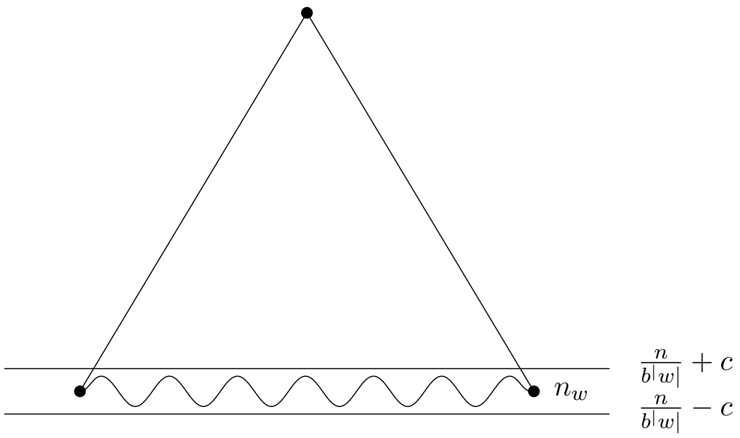

Abbildung 4.4: Illustration von Lemma 4.2 für $w \in \{0,1\}^j$

Analog dazu wird die Ungleichung für n_{0^j} gezeigt. Das Resultat folgt nun aus

$$\sum_{i=0}^{j-1} \frac{1}{b^i} < \sum_{i=0}^{\infty} \left(\frac{1}{b}\right)^i = \frac{1}{1-\frac{1}{b}} = \frac{b}{b-1} = c.$$

□

Für die Tiefe $d(n)$ des Baums können wir nun eine beliebige Funktion wählen, die $\frac{n}{b^{d(n)}} = \Theta(1)$ erfüllt. Dann ist nämlich nach Lemma 4.2 auch für alle $w \in \{0,1\}^{d(n)}$ erreicht dass $n_w = \Theta(1)$ und damit auch für alle $w \in \{0,1\}^{d(n)}$ dass $T(n_w) = \Theta(1)$. Wir setzen $d(n) = \lfloor \log_b n \rfloor$ und beobachten dass $\frac{n}{b^{d(n)}} = \Theta(1)$. Dementsprechend definieren wir die Indizes der inneren Knoten und die Indizes der Blattknoten als:

Definition 4.8. $X_I(n) = \{0,1\}^{\leq d(n)-1}$ und $X_B(n) = \{0,1\}^{d(n)}$.

Die Kosten des gesamten Baums setzen sich zusammen aus den Kosten an den inneren Knoten $\sum_{w \in X_I(n)} a_0^{n_0(w)} a_1^{n_1(w)} f(n_w)$ und den Kosten an den Blättern. Jedes Blatt verursacht konstante Kosten. Für die Bestimmung der Anzahl der Blätter spielt der Unterschied zwischen a_0 und a_1 keine Rolle und somit können wir zu diesem Zweck den Baum als vollständigen Baum der Arität a mit Tiefe $d(n)$ betrachten. Dieser hat $a^{d(n)} = a^{\lfloor \log_b n \rfloor} = \Theta(a^{\log_b n}) = \Theta(n^{\log_b a})$ Blätter. Also erhalten wir

$$T(n) = \Theta(n^{\log_b a}) + \sum_{w \in X_I(n)} a_0^{n_0(w)} a_1^{n_1(w)} f(n_w).$$

für die Summe über den gesamten Baum.

Eine weitere nützliche Beobachtung ist: Falls wir eine Funktion φ schichtweise summieren wollen die nur von der Länge von w, nicht aber von w selbst abhängt, dann spielt der Unterschied zwischen a_0 und a_1 ebenfalls keine Rolle und wir können ebenfalls den Baum als einen vollständigen Baum mit Arität a betrachten und entsprechend aufsummieren, d.h.

$$\sum_{w \in \{0,1\}^j} a_0^{n_0(w)} a_1^{n_1(w)} \varphi(|w|) = a^j \varphi(j).$$

Damit ist also

$$\sum_{w \in X_I(n)} a_0^{n_0(w)} a_1^{n_1(w)} \varphi(|w|) = \sum_{j=0}^{d(n)-1} a^j \varphi(j).$$

Satz 4.4 (Master-Theorem für Teile-und-herrsche-Rekursionsgleichungen)**.** *Seien $a_0, a_1 \in \mathbb{N}$ mit $a = a_0 + a_1 \geq 1$, sei $b > 1$ und sei $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$. Dann hat die Rekursionsgleichung*

$$T(n) = a_0 T(\left\lfloor \frac{n}{b} \right\rfloor) + a_1 T(\left\lceil \frac{n}{b} \right\rceil) + f(n)$$

die folgende asymptotische Lösung.

1. *Falls $f(n) = O(n^{\log_b a - \varepsilon})$ für ein $\varepsilon > 0$, dann $T(n) = \Theta(n^{\log_b a})$.*
2. *Falls $f(n) = \Theta(n^{\log_b a})$, dann $T(n) = \Theta(n^{\log_b a} \log n)$.*
3. *Falls es ein $c < 1$ gibt, so dass für alle bis auf endlich viele $n \in \mathbb{N}$ die Ungleichung $a_0 f(\lfloor \frac{n}{b} \rfloor) + a_1 f(\lceil \frac{n}{b} \rceil) \leq c f(n)$ gilt, dann ist $f(n) = \Omega(n^{\log_b a + \varepsilon})$ für ein $\varepsilon > 0$ und $T(n) = \Theta(f(n))$.*