Datenkompression mit deterministischen Experten

und

Systeme teilweise unabhängiger Zufallsvariablen

Dissertation zur Erlangung des Doktorgrades der

Fakultät für Mathematik der Universität Bielefeld

vorgelegt von

Bernhard Balkenhol

1995

1. Begutachter: Priv. Doz. Dr. I. Althöfer

2. Begutachter: Prof. Dr. R. Ahlswede

Datum der mündlichen Prüfung: 31.3.1995

Vorwort :

In dieser Arbeit wird der Einsatz von Experten in der Datenkompression betrachtet.
Hat man zum Beispiel eine Datenbank, in welcher die Spielverläufe von Schachpartien gespeichert werden sollen, so können zur Speicherung der einzelnen Züge einer Partie Experten in Form von Schachcomputern oder Schachprogrammen benutzt werden. Dabei liegen dem Ein – und Decodierer die gleichen deterministischen Experten vor.

Hat man n Experten $E_1, \ldots, E_n$, wobei der i–te Experte den zu speichernden Zug mit einer Wahrscheinlichkeit von p_i korrekt vorhersagt, so stellt sich je nach betrachtetem Modell die Frage, wie effizient der Einsatz der Experten ist:

- Welche Reihenfolge der Experten ist die günstigste in Modellen, in denen die Experten anzuordnen sind ?
- Welche Experten sollte man benutzen, wenn sie nicht unabhängig voneinander sind ?
- Wie viele der Experten sollte man benutzen ?
- Wie stark hängt das zu erzielende Resultat von dem Grad der Unabhängigkeit der Experten ab ?

Durch obige Fragen läßt sich die Komplexität dieses Problems erahnen. Daher werden von mir in dieser Arbeit nur zwei Teilprobleme analysiert, die unabhängig voneinander gelesen werden können.

1. In dem ersten Teil dieser Arbeit werden einige Modelle zur Datenkompression vorgestellt und analysiert.

    - Welche Anzahl von unabhängigen und identisch verteilten Experten liefert eine minimale durchschnittliche Codewortlänge ?

2. Im zweiten Teil dieser Arbeit werden Systeme teilweise unabhängiger Zufallsvariablen betrachtet.

    Kennt man von einem komplizierten System mit vielen Koordinaten nur etwas über das Zusammenspiel von jeweils kleinen Koordinatengruppen, wie kann sich dann das Gesamtsystem verhalten (z.B. wenn sich Unabhängigkeiten nicht fortsetzen) ?

Seien n Zufallsvariablen $X_1, \ldots, X_n$ mit Wertebereichen $\mathcal{X} = \{0, 1\}$ und Randverteilungen $(\frac{1}{2}, \frac{1}{2})$ gegeben. Sei eine Funktion $f : \mathcal{X}^n \to \mathbb{R}$ gegeben.

1. Durch welche Wahrscheinlichkeitsverteilung P auf $\mathcal{X}^n$ wird der Erwartungswert $E_f(P) = \sum_{v \in \mathcal{X}^n} p_v \cdot f(v)$ maximiert ?

Sei $E_f^*(k,n)$ dieses Maximum bei n Zufallsvariablen, die k– unabhängig sind.

2. Welche quantitativen Unterschiede im Resultat werden durch den Grad der Unabhängigkeit der Zufallsvariablen erzielt ?

3. Wie schnell wächst für jedes feste k die Folge $(E_f^*(k,n))_{n=k}^{\infty}$?

Besonders danken möchte ich Priv. Doz. Dr. Ingo Althöfer und Prof. Dr. Rudolf Ahlswede für die Anregung zu dieser Arbeit und für die ausgesprochen motivierende Betreuung.

Inhaltsverzeichnis

1 Datenkompression mit Experten

Datenkompression unter Verwendung von unabhängigen deterministischen Experten

1.1 Einleitung

In den letzten Jahren wurden in der Technologie der Informationsverarbeitung und der Kommunikation erhebliche Fortschritte erzielt (z.B. Glasfaserkabel, neue Speichermedien wie CD' s, kleinere und schnellere Microchips, TV–Satelliten, usw.). Die mathematische Disziplin zur Behandlung von Information und Kommunikation nennt sich **Informationstheorie**.
Bei der ständig wachsenden Datenmenge ist neben der elektronischen Weiterentwicklung auch eine auf Software beruhende Optimierung des Datenaustausches von Interesse. Eine Reduzierung der Datenmenge ohne Informationsverlust ist daher eines der wesentlichen, durch Software zu lösenden Probleme. Mit der theoretischen Behandlung dieses Problems beschäftigt sich ein Teilbereich der Informationstheorie: die **Datenkompression**.

Im weiteren Verlauf dieser Arbeit soll ein Modell der Datenkompression vorgestellt und analysiert werden. In diesem Modell sollen Daten komprimiert werden, die durch Experten beschrieben werden können.

Man stelle sich zum Beispiel eine Datenbank vor, in welcher die Spielverläufe von Schachpartien gespeichert werden sollen. Zum Speichern eines Zuges in einer gegebenen Schachstellung benötige man durchschnittlich b Bits.

Nun soll gezeigt werden, daß bei der Speicherung oder Übertragung des Zuges Experten in Form von Schachcomputern oder Programmen benutzt werden können. Hat man einen deterministischen Schachcomputer, so heißt deterministisch, daß dieser Schachcomputer in der gleichen Situation auch immer den gleichen Zug vorschlägt. Möchte man einen Zug der Datenbank hinzufügen, so speichert man, falls der Schachcomputer den gleichen Zug vorschlägt, wie man ihn der Datenbank hinzufügen möchte, eine 1, anderenfalls eine 0 gefolgt von den b Bits, die benötigt werden, wenn kein Experte benutzt wird. Beim Lesen der Datenbank erhält man den gewünschten Zug, falls eine 1 gelesen wurde, durch den Vorschlag des Schachcomputers. Falls jedoch eine 0 als erste Ziffer übermittelt wurde, so erhält man den Zug aus den nachfolgenden b Bits (siehe [Alt 91]). Ist nun b sehr groß, und stimmt der zu speichernde Zug sehr häufig mit dem vom Schachcomputer vorgeschlagenen Zug überein, so ist im Durchschnitt die zu speichernde Länge sehr viel kürzer als b Bits. Man hat also eine Reduzierung des benötigten Speicherplatzes.

Als ein anderes Beispiel betrachte man das Speichern oder Übertragen von natürlichsprachlichen Texten mittels eines Sprachprogramms, in dem die Wörter des Dudens gespeichert sind. Die Funktionsweise des Programms soll sein, bei einem gegebenen Wortanfang den Buchstaben zu nennen, der am wahrscheinlichsten Fortsetzung des Wortanfangs ist. Möchte man nun ein Wort speichern oder übertragen, so sendet man eine 1, falls der zu übertragende Buchstabe mit dem Vorschlag des Sprachprogramms übereinstimmt, eine 0 gefolgt von b Bits sonst. Dabei entsprechen die b Bits einer anderen gegebenen Codierung des Buchstabens (z.B. durch den ASCII-Code).

Ein anderes Modell wird durch die Codierung der Differenz zwischen dem Vorschlag des Experten und dem zu übertragenden Zeichen geliefert. Man habe zum Beispiel Binärfolgen konstanter Länge und einen Experten, der zu jedem Zeitpunkt die zu übertragende Binärfolge bis auf eine Abweichung von maximal einem Bit vorhersagen kann. In einem solchen Modell ist es günstig, nur die Abweichung der Vorhersage des Experten von der zu übermittelnden Binärfolge zu codieren. Ein derartiges Modell wurde in [Eli 55] betrachtet und analysiert.

Natürlich können auch solche Modelle betrachtet werden, in denen mehrere Experten befragt werden. Zum Beispiel kann die Übertragung eines Zuges in einer festen Schachstellung dadurch erfolgen, daß man den Zug selbst in irgendeiner Form codiert überträgt, oder aber, daß man die Nummer des Experten überträgt, der diesen Zug in dieser vorgegebenen Situation vorschlägt. Dieses Verfahren setzt voraus, daß Sender und Empfänger die gleichen deterministischen Experten in der gleichen Reihenfolge benutzen.

Eine anderes Modell erhält man dadurch, daß man denjenigen Zug als Vorschlag betrachtet, der von der Mehrheit der Experten bevorzugt wird. Ist diese Mehrheit nicht eindeutig, so entscheidet man nach einer geeigneten deterministischen Tie–Break– Regel. Damit ist auch dieses Modell deterministisch und kann zur Komprimierung benutzt werden.

In diesem Teil der Arbeit wird ein Codierungsverfahren in folgendem stochastischen Modell analysiert:

- Man hat n deterministische Experten in einer fest vorgegebenen Reihenfolge.
- Der Vorschlag eines jeden Experten wird nicht durch die Vorschläge der anderen beeinflußt (Unabhängigkeit).
- Jeder Experte schlägt mit der gleichen Wahrscheinlichkeit p $(0 < p < 1)$ den richtigen, d.h. den zu übertragenden Zug vor.
- Der i–te Experte wird nur dann gefragt, wenn alle vorhergehenden Experten einen anderen als den zu übertragenden Zug vorgeschlagen haben.
- Wenn kein Experte eingesetzt wird, benötigt man durchschnittlich b Bits, um den Zug anders zu übertragen.

Da der erste Experte den richtigen Zug mit einer Wahrscheinlichkeit p vorschlägt, wird auch mit einer Wahrscheinlichkeit p der erste Experte bei der Übertragung des Zuges angegeben. Der zweite Experte wird nur dann gefragt, wenn der erste Experte **nicht** den richtigen Zug vorgeschlagen hat. Da auch dieser zweite Experte mit einer Wahrscheinlichkeit von p den richtigen Zug vorschlägt und die Experten unabhängig sind, ist die Wahrscheinlichkeit, daß der zweite Experte bei der Übertragung des Zuges angegeben wird, $(1-p) \cdot p$, usw. Die Wahrscheinlichkeit, daß der i–te Experte bei der Übertragung des Zuges angegeben wird, ist $(1-p)^{i-1} \cdot p$. Es bleibt noch der Fall, daß keiner der Experten den zu übertragenden Zug vorgeschlagen hat. Die Wahrscheinlichkeit hierfür ist $(1-p)^n$. Man erhält also aus der Anzahl der eingesetzten Experten n und der Wahrscheinlichkeit p eine Wahrscheinlichkeitsverteilung auf der Menge $\{0, \ldots, n\}$

$$Prob(p,n) := \big((1-p)^n, (1-p)^{n-1} \cdot p, ..., (1-p) \cdot p, p\big) .$$

Notation :
Die Wahrscheinlichkeit, daß der i–te Experte bei der Übertragung eines Zuges angegeben wird, sei p_i für alle $1 \le i \le n$, und die Wahrscheinlichkeit, daß kein Experte den zu übertragenden Zug angibt, sei p_0. Es gilt also $p_0 = (1-p)^n$ und $p_i = (1-p)^{i-1} \cdot p$ für alle $1 \le i \le n$. Jedem p_i wird nun ein Codewort C_i zugeordnet, indem man die Wahrscheinlichkeitsverteilung $Prob(p,n)$ mit dem **Huffman**–Verfahren codiert (siehe [Ash 65]).

Die Funktionsweise des Huffman–Verfahrens wird in Kapitel 3 beschrieben[1]. Dort werden auch die aus der Informationstheorie benutzten Theoreme vorgestellt und erläutert.

Nun gilt für die Übertragung:

- Muß der i–te Experte bei der Übertragung angegeben werden, so wird C_i übermittelt.
- Hat keiner der Experten den zu übertragenden Zug vorgeschlagen, so wird zunächst C_0 übertragen, um dem Empfänger diese Situation mitzuteilen, dann werden die b Bits übertragen.

Die Kosten $kost(p,b,n)$ dieses Codierungsverfahrens werden durch die durchschnittliche Codewortlänge repräsentiert.
Sei l_i die Länge von C_i für alle $0 \le i \le n$. Dann ist $L(p,n) = (l_0, l_n, l_{n-1}, \ldots, l_1)$ die Liste der Codewortlängen, die man erhält, wenn man die Wahrscheinlichkeitsverteilung $Prob(p,n)$ mit dem Huffman–Verfahren codiert.

Es gilt: $\qquad kost(p,b,n) = p_0 \cdot (l_0 + b) + \sum_{i=1}^{n} p_i \cdot l_i.$

[1] Dies geschieht, da nur mit detaillierter Kenntnis der Funktionsweise die Beweise innerhalb der folgenden Kapitel verstanden werden können.

Zu vorgegebenem p und b ist nun $n \in \mathbb{N} \cup \{0, \infty\}$ so zu wählen, daß $kost(p, b, n)$ minimal wird.

1.2 Ergebnisse

Notation :
Die Wahrscheinlichkeiten p_i und die Längen l_i hängen sowohl von p als auch von n ab. Im Zusammenhang mit fest vorgegebenen p und n, werde ich auf die Schreibweise $p_i(p,n)$ anstatt p_i und $l_i(p,n)$ anstatt l_i verzichten.

Nach der Analyse dieses Problems vermute ich, daß $kost(p,b,n)$ minimal wird, wenn man $n=0$ oder $n=\infty$ wählt, d.h., daß man je nach Vorgabe von p und b entweder keinen oder unendlich viele Experten einsetzen sollte.

Genauer:
Sei $t(p) = \left\lceil \frac{-log_2(2-p)}{log_2(1-p)} \right\rceil$ und $s(p) = \left\lfloor \frac{log_2(p)}{log_2(1-p)} \right\rfloor$.

Sei $c_0(p) = 1 + \lfloor log_2 t(p) \rfloor + \frac{(1-p)^{2^{\lfloor log_2 t(p)\rfloor+1}-t(p)}}{1-(1-p)^{t(p)}}$.

Dann vermute ich, daß $kost(p,b,n)$ für $n=0$ minimal ist, wenn $b \leq c_0(p)$, und sonst für $n=\infty$.

Von dieser Vermutung sind die in dem folgenden Theorem enthaltenen Ergebnisse bewiesen.

Sei $\quad H(q_1,\ldots,q_n) = -\sum_{i=1}^{n} q_i \cdot log_2 q_i \quad$ die Entropie der Wahrscheinlichkeitsverteilung $Q=(q_1,\ldots,q_n)$.

Theorem 1:

1. Für $b < \frac{1}{p} \cdot H(p, 1-p)$ ist $kost(p,b,n)$ streng monoton steigend in n.
2. Für $b > \frac{1}{p}$ ist $kost(p,b,n)$ streng monoton fallend in n.
3. Für alle $p \geq \frac{4}{3} - \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}} - \frac{1}{9 \cdot \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}}} \approx 0.245122$ gilt:
   (a) Für $b < c_0(p)$ ist $kost(p,b,n)$ minimal genau dann, wenn $n=0$.
   (b) Für $b > c_0(p)$ ist $kost(p,b,n)$ minimal genau dann, wenn $n=\infty$.
   (c) Für $b = c_0(p)$ ist $kost(p,b,n)$ minimal genau dann, wenn $n = i \cdot t(p)$ für ein $i \in \mathbb{N} \cup \{0,\infty\}$.
4. Ist $kost(p,b,n)$ minimal, so ist $n \in \{0,\infty\}$ oder es gilt
   $t(p) < n < s(p) + 3 \cdot t(p)$.

Um dieses Theorem zu beweisen, werden zunächst einige Hilfsmittel aus der Informationstheorie benötigt (siehe Kapitel 3). Speziell mit Hilfe des Gruppierungsaxioms und des Noiseless Coding Theorem erhält man:

1. $H(Prob(p,n)) = H(p, 1-p) \cdot \frac{1-(1-p)^n}{p}$,

2. $H(Prob(p,n)) + (1-p)^n \cdot b \leq kost(p,b,n) < H(Prob(p,n)) + (1-p)^n \cdot b + 1$.

Nun kann in Kapitel 4 eine Ungleichung für die Kostenfunktion präsentiert werden. Es gilt für alle $n \in \mathbb{N}$ und alle $p \in (0,1)$:

$$kost(p,b,n) \leq kost(p,b,n-1) + (1-p)^{n-1} - p \cdot b \cdot (1-p)^{n-1} \tag{1.1}$$

mit Gleichheit genau dann, wenn $p \geq \frac{3-\sqrt{5}}{2}$.

Unter Verwendung dieses Ergebnisses lassen sich die ersten beiden Aussagen von Theorem 1 beweisen. Es verbleibt, den Fall $\frac{1}{p} \cdot H(p, 1-p) \leq b \leq \frac{1}{p}$ zu analysieren.

Für die weitere Analyse werden die Codewortlängen, also die Liste $L(p,n)$, betrachtet. Für verschiedene Wahrscheinlichkeiten p und q können die Listen $L(p,n)$ und $L(q,n)$ verschieden sein. Als Beispiel betrachte man $L(\frac{1}{3}, 3) = (2,2,2,2)$ und $L(\frac{1}{4}, 3) = (1,3,3,2)$.

Zu jeder Wahrscheinlichkeit p existiert ein Intervall $I(p) \subset (0,1)$, mit der Eigenschaft, daß für alle $q \in I(p)$ gilt: $L(p,n) = L(q,n)$ für alle $n \in \mathbb{N}$.
Um eine weitere Analyse der Kostenfunktion vornehmen zu können, wäre es von Interesse, die Grenzen dieser Intervalle bestimmen zu können. Diese Aufgabe erweist sich jedoch als äußerst schwierig. Aus diesem Grund wird zunächst auf eine gröbere Partitionierung des Intervalls $(0,1)$ zurückgegriffen, die durch den Begriff „Periodenlänge“ definiert werden soll.
In Kapitel 5 wird der Begriff der Periodenlänge im Zusammenhang mit der Folge $L(p) := \{L(p,n)\}_{n \in \mathbb{N}}$ definiert. Die Bedeutung dieses Begriffs möchte ich an dieser Stelle anhand zweier Beispiele andeuten.

Wie man an dem folgenden Beispiel für $p = \frac{1}{2}$ sehen kann, tritt für jedes $n \geq 2$ jede Codewortlänge in $L(\frac{1}{2}, n)$, bis auf die größte, genau einmal auf. Daher möchte ich $L(p)$ für $p = \frac{1}{2}$ die Periodenlänge 1 zuordnen.
In dem Beispiel für $p = \frac{1}{3}$ tritt für jedes $n \geq 4$ jede Codewortlänge in $L(\frac{1}{3}, n)$, bis auf jeweils eine Ausnahme, genau zweimal auf. Daher möchte ich $L(p)$ für $p = \frac{1}{3}$ die Periodenlänge 2 zuordnen.

Die Analyse in Kapitel 5 ergibt für allgemeine p, daß $L(p)$ die Periodenlänge

$$t(p) = \left\lceil \frac{-ln(2-p)}{ln(1-p)} \right\rceil$$

hat.

$$
\begin{array}{llll}
L(\frac{1}{2},0) & = (1) & L(\frac{1}{3},0) & = (1) \\
L(\frac{1}{2},1) & = (1,1) & L(\frac{1}{3},1) & = (1,1) \\
L(\frac{1}{2},2) & = (2,2,1) & L(\frac{1}{3},2) & = (1,2,2) \\
L(\frac{1}{2},3) & = (3,3,2,1) & L(\frac{1}{3},3) & = (2,2,2,2) \\
L(\frac{1}{2},4) & = (4,4,3,2,1) & L(\frac{1}{3},4) & = (2,3,3,2,2) \\
L(\frac{1}{2},5) & = (5,5,4,3,2,1) & L(\frac{1}{3},5) & = (3,3,3,3,2,2) \\
L(\frac{1}{2},6) & = (6,6,5,4,3,2,1) & L(\frac{1}{3},6) & = (3,4,4,3,3,2,2) \\
L(\frac{1}{2},7) & = (7,7,6,5,4,3,2,1) & L(\frac{1}{3},7) & = (4,4,4,4,3,3,2,2) \\
L(\frac{1}{2},8) & = (8,8,7,6,5,4,3,2,1) & L(\frac{1}{3},8) & = (4,5,5,4,4,3,3,2,2)
\end{array}
$$

In Kapitel 6 wird unter Verwendung der bisherigen Ergebnisse die obere Schranke für n aus Aussage 4 von Theorem 1 bewiesen. Es wird für $n \geq s(p) + 3 \cdot t(p)$ eine Rekursionsformel für die Listen $L(p, n)$ und damit auch für die Kostenfunktion entwickelt.

Es gilt für alle $n \geq s(p) + 3 \cdot t(p)$:

$$kost(p, b, n + t(p)) = (1 - p)^{t(p)} \cdot kost(p, b, n) + kost(p, b, t(p)). \tag{1.2}$$

Mit Hilfe von 1.1 wird Aussage 3 von Theorem 1 für alle $p \geq \frac{3-\sqrt{5}}{2}$ bewiesen.

In Kapitel 7 wird die untere Schranke für n von Aussage 4 aus Theorem 1 bewiesen. Sei $k = \lfloor log_2 n \rfloor$. Dann gilt für alle $n \leq t(p)$:

$$kost(p, b, n) = b \cdot p_0 + (k + 1) \cdot (1 - p_0) + (1 - p)^{2^{k+1}-n}.$$

In Kapitel 7 wird Aussage 3 aus Theorem 1 unter Verwendung von Aussage 4 für alle $p \geq \frac{4}{3} - \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{100}}} - \frac{1}{9 \cdot \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{100}}}}$ bewiesen. Es wird gezeigt, daß die Rekursionsvorschrift 1.2 für alle $n \geq 2$ gilt.

Schließlich werden Vermutungen und noch offene Fragestellungen behandelt.

1.3 Hilfsmittel aus der Informationstheorie

In diesem Abschnitt sollen einige Theoreme aus der Informationstheorie vorgestellt und erläutert werden. Dabei werde ich die folgenden Theoreme und Definitionen speziell für die Anwendung auf das von mir betrachtete Problem formulieren, obwohl sie in der Informationstheorie allgemeiner vorhanden sind und benutzt werden (siehe [CT 91], [LH 87] und [Sha 51]).

Man habe ein endliches Alphabet gegeben durch $\mathcal{X} = \{1, \ldots, a\}$ und eine Wahrscheinlichkeitsverteilung $P = (P(1), \ldots, P(a))$ auf $\mathcal{X}$. Die Quelle $(\mathcal{X}, P)$ kann beschrieben werden durch eine Zufallsvariabel $X : \mathcal{X} \to \mathcal{X}$, wobei
$Prob(X = x) = P(x)$ für alle $x \in \mathcal{X}$. Ein Wort $(x_1, \ldots, x_n) \in \mathcal{X}^n$ ist die Realisierung der Zufallsvariablen $X^n = (X_1, \ldots X_n)$, wobei die X_i's identisch verteilt und unabhängig voneinander sind.
Daher ist die Wahrscheinlichkeit $P^n(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i)$ das Produkt der Wahrscheinlichkeiten der einzelnen Zeichen.

Sei $\mathcal{Y} = \{0, 1\}$ das binäre Alphabet und $\mathcal{Y}^* = \cup_{n=0}^{\infty} \mathcal{Y}^n$ die Menge aller endlichen Sequenzen (Wörter) über $\mathcal{Y}$. (Konvention: das Wort der Länge 0 bezeichne ich als leeres Wort.)

Definition : Ein **Code** (variabler Länge) ist eine Funktion $c : \mathcal{X} \to \mathcal{Y}^*$.
$\{c(1), \ldots, c(a)\}$ ist die Menge der **Codeworte**, wobei für $x = 1, \ldots, a$
$c(x) = \big(c_1(x), \ldots, c_{L(x)}(x)\big)$ ist. $L(x)$ bezeichnet man als die **Codewortlänge** von $c(x)$.

Zu einem gegebenen Code c stellt $\overline{L}(c) = \sum_{i=1}^{a} P(i) \cdot L(i)$ die durchschnittliche Codewortlänge dar.

Definition : Ein Code c heißt **eindeutig entzifferbar**, wenn jedes Wort in $\mathcal{Y}^*$ durch höchstens eine Sequenz von Codeworten repräsentiert wird.
D.h., es ist notwendig, daß die Funktion $C : \mathcal{X}^* \to \mathcal{Y}^*$ gegeben durch
$C(x_1, \ldots, x_t) = c(x_1) \ldots c(x_t)$ (für jedes beliebige $t \in \mathbb{N}$ mit $(x_1, \ldots, x_t) \in \mathcal{X}^*$) injektiv ist.

Definition : Ein Code c heißt **Präfixcode**, wenn für je zwei Codeworte $c(x)$ und $c(y)$ mit $x \neq y$ und $L(x) \leq L(y)$ gilt:
$\big(c_1(x), \ldots, c_{L(x)}(x)\big) \neq \big(c_1(y), \ldots c_{L(x)}(y)\big)$. D.h., $c(x)$ und $c(y)$ müssen sich in mindestens einer der ersten $L(x)$ Komponenten unterscheiden.

Definition : Ein Präfixcode c heißt **optimal**, wenn kein eindeutig entzifferbarer Code d mit $\overline{L}(d) < \overline{L}(c)$ existiert.

Theorem 2: (Kraftungleichung für eindeutig entzifferbare Codes)
Ein eindeutig entzifferbaren Code mit Codewortlängen $L(1), \ldots, L(a)$ existiert genau dann, wenn $\sum_{x \in \mathcal{X}} 2^{-L(x)} \leq 1$.

Theorem 3: (Kraftungleichung für Präfixcodes)
Ein Präfixcode mit Codewortlängen $L(1), \ldots, L(a)$ existiert genau dann, wenn $\sum\limits_{x \in \mathcal{X}} 2^{-L(x)} \leq 1$.

Theorem 4: Gegeben sei eine Quelle $(\mathcal{X}, P)$, wobei $\mathcal{X} = \{1, \ldots, a\}$ und die Wahrscheinlichkeiten in absteigender Reihenfolge geordnet sind, d.h.
$P(1) \geq \ldots \geq P(a)$. Eine neue Wahrscheinlichkeitsverteilung P' sei definiert durch $P' = (P(1), \ldots, P(a-2), P(a-1) + P(a))$. Sei $c' = (c'(1), \ldots, c'(a-1))$ ein optimaler Präfixcode für P'. Aus c' definieren wir einen Code für die Wahrscheinlichkeitsverteilung P durch

$c(x) = c'(x)$ für alle $1 \leq x \leq a-2$, $c(a-1) = c'(a-1)0$ und $c(a) = c'(a-1)1$.

So stellt c einen optimalen Präfixcode für P dar.

Dieses Theorem suggeriert den folgenden rekursiven Algorithmus, der **Huffman**–Verfahren genannt wird (siehe [Ash 65]).

Zunächst wird die Quelle schrittweise um ein Element reduziert, indem man die beiden kleinsten Wahrscheinlichkeiten $P(a-1)$ und $P(a)$ der Wahrscheinlichkeitsverteilung P addiert und das Resultat in die verbleibenden Wahrscheinlichkeiten einsortiert. Somit erhält man eine neue Wahrscheinlichkeitsverteilung P' aus P, mit der man rekursiv fortsetzt, bis schließlich nur noch zwei Wahrscheinlichkeiten übrig bleiben. Dem ersten Element dieser beiden Wahrscheinlichkeiten ordnet man eine 0 zu, dem anderen eine 1.

Nun wird die Quelle schrittweise wieder zerlegt, bis man letztendlich die ursprüngliche Quelle restauriert hat. Dabei erhält man in jedem Schritt die Codeworte $c(a-1)$ und $c(a)$, indem man dem zu $P(a-1) + P(a)$ gehörenden Codewort der Wahrscheinlichkeitsverteilung P' in dem ersten Fall eine 0 und im zweiten Fall eine 1 anhängt.

Definition : Als **Entropie** einer Wahrscheinlichkeitsverteilung P bezeichnet man die Funktion

$$H(P) = -\sum_{i=1}^{a} P(i) \cdot log_2 P(i)$$

(zu den Eigenschaften der Entropie siehe z.B. [AW 79]).

Theorem 5: (Noiseless Coding Theorem)
Für eine Quelle $(\mathcal{X}, P)$ mit $\mathcal{X} = \{1, \ldots, a\}$ existiert ein Präfixcode c, dessen durchschnittlicher Codewortlänge erfüllt:

$$H(P) \leq \overline{L}(c) < H(P) + 1.$$

Bemerkung 1: Der Huffmancode ist ein optimaler Code und erfüllt somit das Noiseless Coding Theorem. Die Länge eines jeden Codewortes entspricht der um

1 vergrößerten Häufigkeit des Zusammenfassens der zu diesem Codewort gehörenden Wahrscheinlichkeit.

Theorem 6:
Die Entropiefunktion hat unter anderen folgende Eigenschaften:

- $H(P(1),\ldots,P(n))$ ist symmetrisch.
- $H(P(1),\ldots,P(n),0) = H(P(1),\ldots,P(n))$.
- (Gruppierungsaxiom) Ist $P(n) = Q(1) + Q(2) > 0$, dann gilt

$$H(P(1),\ldots,P(n-1),Q(1),Q(2)) = H(P(1),\ldots,P(n)) + P(n)\cdot H(\frac{Q(1)}{P(n)},\frac{Q(2)}{P(n)}).$$

Bemerkung 2:
Die durchschnittliche Codewortlänge eines Codes hängt nicht von den Codeworten selbst, sondern nur von den Codewortlängen ab. Aufgrund der beiden Kraftungleichungen existiert zu jedem eindeutig entzifferbaren Code ein Präfixcode mit den gleichen Codewortlängen. Daher ist es bei der Suche nach optimalen Codes ausreichend, sich auf Präfixcodes zu beschränken.

Mit diesen Hilfsmitteln lassen sich nun für die in dieser Arbeit betrachtete Wahrscheinlichkeitsverteilung $Prob(p,n)$ einige nützliche Aussagen herleiten.

Lemma 1:
Es gilt

1. $H(Prob(p,n)) = H(p,1-p)\cdot\frac{1-(1-p)^n}{p}$,
2. $\lim\limits_{n\to\infty} H(Prob(p,n)) = \frac{1}{p}\cdot H(p,1-p)$,
3. $H(Prob(p,n)) + (1-p)^n\cdot b \leq kost(p,b,n) < H(Prob(p,n)) + (1-p)^n\cdot b + 1$.

Beweis :
Es gilt, daß $p_0 + p_n = (1-p)^n + (1-p)^{n-1}\cdot p = (1-p)^{n-1}$.
Daher gilt $(p_0 + p_n, p_{n-1},\ldots,p-1) = Prob(p,n-1)$.
Nach dem Gruppierungsaxiom gilt für die Unterteilung in $n-1$ Gruppen:

$$H(Prob(p,n)) = H(Prob(p,n-1)) + (1-p)^{n-1}\cdot H(p,1-p).$$

Durch mehrfache Anwendung dieses Axioms ergibt sich mit
$H(Prob(p,1)) = H(p,1-p)$:

$$H(Prob(p,n)) = H(p,1-p)\cdot\sum_{i=0}^{n-1}(1-p)^i \text{ , also}$$

$$H(Prob(p,n)) = H(p,1-p)\cdot\tfrac{1-(1-p)^n}{p}.$$

Damit ist 1. bewiesen. Da $\lim\limits_{n\to\infty}(1-p)^n = 0$, ist auch 2. mit 1. bewiesen.

Nun gilt nach dem Noiseless Coding Theorem:

$$H(Prob(p,n)) \leq kost(p,0,n) < H(Prob(p,n)) + 1, \text{ also}$$

$$H(Prob(p,n)) + (1-p)^n \cdot b \leq kost(p,b,n) < H(Prob(p,n)) + (1-p)^n \cdot b + 1. \ \square$$

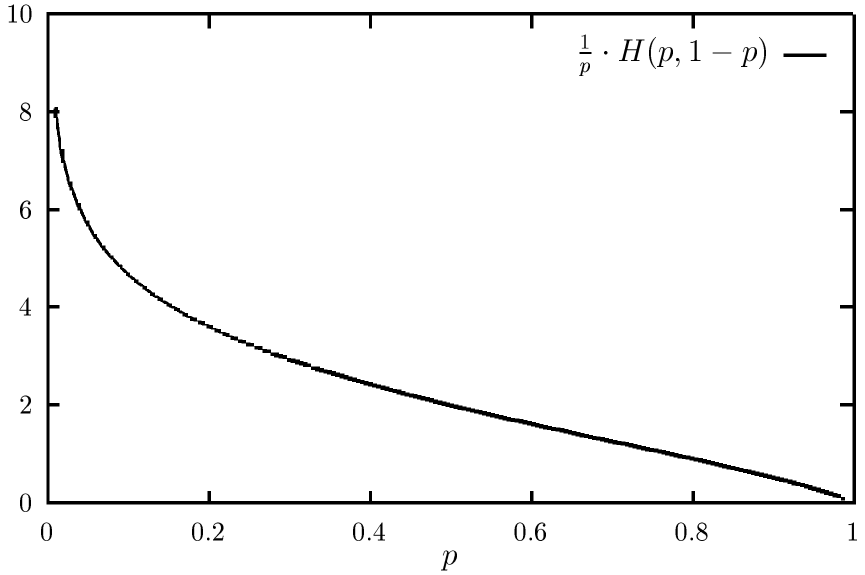
Lemma 2:
$\frac{1}{p} \cdot H(p, 1-p)$ ist für $0 < p < 1$ monoton fallend in p.

Beweis :
$\frac{1}{p} \cdot H(p,1-p) = \frac{1}{p} \cdot \left(-p \cdot \frac{ln(p)}{ln(2)} - (1-p) \cdot \frac{ln(1-p)}{ln(2)}\right) = -\frac{ln(p)}{ln(2)} - \frac{ln(1-p)}{p \cdot ln(2)} + \frac{ln(1-p)}{ln(2)}$.

$\frac{1}{p} \cdot H(p,1-p)$ ist monoton fallend in p, wenn $\frac{\partial}{\partial p}\left(\frac{1}{p} \cdot H(p,1-p)\right) < 0$.

$$\frac{\partial}{\partial p}\left(\frac{1}{p} \cdot H(p,1-p)\right) = -\frac{2}{p \cdot ln(2)} + \frac{ln(1-p)}{p^2 \cdot ln(2)} < 0$$

da beide Summanden negativ sind für alle $0 < p < 1$. $\square$



Definition :
Bei der Huffman–Codierung sei das Zusammenfassen der beiden kleinsten Wahrscheinlichkeiten und das Einsortieren der daraus resultierenden neuen Wahrscheinlichkeit als **Schritt** bezeichnet. Die Anzahl Wahrscheinlichkeiten dieser Restliste, die kleiner sind als die Summe der beiden kleinsten Wahrscheinlichkeiten, sei als **Schrittweite** bezeichnet.

1.4 Schranken für b

In diesem Kapitel sollen unter Verwendung der Hilfsmittel aus der Informationstheorie (Kapitel 3) die Schranken für b, d.h. die Aussagen 1 und 2 von Theorem 1 bewiesen werden.

Bemerkung 3:
Es gilt $p_i < p_{i-1}$ für alle $2 \leq i \leq n$, denn $p_i = (1-p) \cdot p_{i-1}$ für alle $2 \leq i \leq n$ und $(1-p) < 1$.
Daher ist $Prob(p,n)$ bis auf p_0 bereits sortiert für alle p und n.

Lemma 3:
Es gilt für alle p, b und alle $n \in \mathbb{N}$:

$$kost(p,b,n) \leq kost(p,b,n-1) + (1-p)^{n-1} - p \cdot b \cdot (1-p)^{n-1}$$

mit Gleichheit genau dann, wenn $p \geq \frac{3-\sqrt{5}}{2}$ ist.

Beweis :
Sind p_0 und p_n die beiden kleinsten Wahrscheinlichkeiten, so werden diese bei der Huffman–Codierung im ersten Schritt zusammengefaßt.

$$p_0 + p_n = (1-p)^n + (1-p)^{n-1} \cdot p = (1-p)^{n-1} \cdot (1-p+p) = (1-p)^{n-1}$$

Durch Zusammenfassen von p_0 und p_n erhält man also aus der Wahrscheinlichkeitsverteilung $Prob(p,n)$ die Wahrscheinlichkeitsverteilung $Prob(p,n-1)$.

Sind also p_0 und p_n die beiden kleinsten Wahrscheinlichkeiten in der Wahrscheinlichkeitsverteilung $Prob(p,n)$, so gilt für $kost(p,0,n)$:

$$\begin{aligned} kost(p,0,n) &= kost(p,0,n-1) + (1-p)^n \cdot (l+1) \\ &\quad +(1-p)^{n-1} \cdot p \cdot (l+1) - (1-p)^{n-1} \cdot l \\ &= kost(p,0,n-1) + (1-p)^{n-1} \end{aligned}$$

wobei l die Länge des Codewortes p_0 von $Prob(p,n-1)$ darstellt.
Sind p_0 und p_n nicht die beiden kleinsten Wahrscheinlichkeiten, so wird die durchschnittliche Codewortlänge größer, wenn man p_0 und p_n zusammenfaßt und dann mit der Huffmancodierung fortfährt, als wenn die beiden kleinsten Wahrscheinlichkeiten zusammengefaßt werden. Es gilt also in diesem Fall:

$$kost(p,0,n) < kost(p,0,n-1) + (1-p)^{n-1}$$

und daher gilt allgemein:

$$kost(p,0,n) \leq kost(p,0,n-1) + (1-p)^{n-1}.$$

Nun gilt

$$\begin{aligned}
kost(p,b,n) &= kost(p,0,n) + (1-p)^n \cdot b \\
&\leq kost(p,0,n-1) + (1-p)^{n-1} + (1-p)^n \cdot b \\
&= kost(p,b,n-1) - (1-p)^{n-1} \cdot b + (1-p)^{n-1} + (1-p)^n \cdot b \\
&= kost(p,b,n-1) + (1-p)^{n-1} - p \cdot b \cdot (1-p)^{n-1}.
\end{aligned}$$

Also gilt Gleichheit in Lemma 3, wenn p_0 und p_n die beiden kleinsten Wahrscheinlichkeiten der Wahrscheinlichkeitsverteilung $Prob(p,n)$ sind. Da, wie bereits erwähnt, $Prob(p,n)$ bis auf p_0 bereits sortiert ist, ist diese Bedingung erfüllt, wenn $p_0 \leq p_{n-1}$.

Damit muß gelten: $(1-p)^n \leq (1-p)^{n-2} \cdot p$,

also $(1-p)^2 - p \leq 0$ und somit

$$p \geq \tfrac{3-\sqrt{5}}{2}.$$

Damit ist Lemma 3 bewiesen. □

Proposition 1:
Falls $b \geq \frac{1}{p}$, ist $kost(p,b,n)$ monoton fallend in n.

Beweis :

Nach Lemma 3 gilt:

$$kost(p,b,n) \leq kost(p,b,n-1) + (1-p)^{n-1} - p \cdot b \cdot (1-p)^{n-1}.$$

Damit ist die Behauptung sicherlich bewiesen, wenn

$$(1-p)^{n-1} - p \cdot b \cdot (1-p)^{n-1} \leq 0, \quad \text{also } 1 \leq p \cdot b \quad \text{und somit } b \geq \tfrac{1}{p}.$$

□

Lemma 4:
Für $p \geq \frac{3-\sqrt{5}}{2}$ gilt:

(i) $kost(p,b,n)$ ist monoton steigend in $n \Leftrightarrow b < \frac{1}{p}$.

(ii) $kost(p,b,n)$ ist konstant in $n \Leftrightarrow b = \frac{1}{p}$.

(iii) $kost(p,b,n)$ ist monoton fallend in $n \Leftrightarrow b > \frac{1}{p}$.

Beweis :
Da nach der Bemerkung die Wahrscheinlichkeitsverteilung $Prob(p,n)$ bis auf p_0 bereits sortiert ist, sind p_0 und p_n die beiden kleinsten Wahrscheinlichkeiten von $Prob(p,n)$, wenn $p_0 < p_{n-1}$ ist.
Dies gilt, wie bereits im Beweis von Lemma 3 gezeigt wurde, wenn $p \geq \frac{3-\sqrt{5}}{2}$.

Für ein p mit dieser Eigenschaft gilt in Lemma 3 Gleichheit, d.h.

$$kost(p,b,n) = kost(p,b,n-1) + (1-p)^{n-1} - p \cdot b \cdot (1-p)^{n-1}.$$

Damit ist Lemma 4 bewiesen. □

Proposition 2:
Für alle $p \in (0,1)$ und $b \leq \frac{1}{p} \cdot H(p, 1-p)$ gilt:
$kost(p,b,n)$ ist minimal für $n = 0$.

Beweis :
Zu beweisen ist also $kost(p,b,n) \geq kost(p,b,0)$ für alle $n \geq 1$.
Es gilt $kost(p,b,0) = b$.

$$\begin{aligned} kost(p,b,n) &= kost(p,0,n) + (1-p)^n \cdot b \\ &\geq H(Prob(p,n)) + (1-p)^n \cdot b \\ &= H(p,1-p) \cdot \tfrac{1-(1-p)^n}{p} + (1-p)^n \cdot b. \end{aligned}$$

Folglich gilt $kost(p,b,n) \geq b$, wenn

$$H(p,1-p) \cdot \frac{1-(1-p)^n}{p} + (1-p)^n \cdot b \geq b,$$

also

$$\frac{1}{p} \cdot H(p,1-p) \geq b.$$

□

$kost(p,b,n)$ ist also noch zu analysieren für $\frac{1}{p} \cdot H(p,1-p) < b < \frac{1}{p}$.

1.5 Periodenlänge der Codierung

In diesem Abschnitt wird der Begriff der Periodenlänge definiert und behandelt.

Definition :
Sei L eine Liste mit den Elementen $L = (e_1, \ldots, e_k)$, und seien $a, b \in \mathbb{N}$ mit $0 \leq a, b \leq k$ gegeben.

Dann definiere ich

$$Anfang(L, a) := (e_1, \ldots, e_a) \quad ,$$
$$Ende(L, b) := (e_{k+1-b}, \ldots, e_k) \quad ,$$
$$Mitte(L, a, b) := (e_{a+1}, \ldots, e_{k-b}).$$

D.h. $Anfang(L, a)$ stellt die Liste der ersten a Elemente von L und $Ende(L, b)$ die Liste der letzten b Elemente von L dar. Dabei wird durch $Anfang(L, 0)$ und durch $Ende(L, 0)$ die leere Liste () dargestellt.

$Anfang(L, a)$ werde ich als **Anfang** von L der Länge a und $Ende(L, b)$ als **Ende** von L der Länge b bezeichnen.

Definition :
Sei $A = \{A_n\}_{n \in \mathbb{N}}$ eine Liste von Listen, wobei jede Liste A_n die Länge $n+1$ habe. Dann hat A die Periodenlänge t, wenn gilt:
Es existieren Größen $a(i) \in \mathbb{N}$ für alle $0 \leq i < t$, $c \in \mathbb{N}$ und $n_0 \in \mathbb{N} \cup \{0\}$, so daß für alle $n \geq n_0$ jeder in $Mitte(A_n, a(n \bmod t), c)$ auftretende Wert genau t mal auftritt.

Nun soll gezeigt werden, daß man $L(p) := \{L(p, n)\}_{n \in \mathbb{N}}$ eine Periodenlänge $t(p)$ mit den Konstanten $s_1(p, i), s_2(p)$ und $n_0(p)$ zuordnen kann. Da $t(p)$ nur von p abhängt, werde ich auch die Schreibweise „$L(p, n)$ hat Periodenlänge $t(p)$" benutzen.

Bemerkung 4:
Mit Hilfe der Konstanten $s_1(p, i)$ und $s_2(p)$ wird die Liste $L(p, n)$ für $n \geq n_0(p)$ in die drei Teile

$Anfang(L(p, n), s_1(p, n \bmod i))$, $Mitte(L(p, n), s_1(p, n \bmod i), s_2(p))$
und $Ende(L(p, n), s_2(p))$ zerlegt.

Im Beispiel für $p = \frac{1}{2}$ (Seite 10) tritt nur $i = 0$ auf, da $t(p) = 1$. Man setze $n_0(\frac{1}{2}) = 2$, $s_1(\frac{1}{2}, 0) = 2$ und $s_2(\frac{1}{2}) = 0$. Dann ist der Anfang gegeben durch (n, n) und das Ende durch ( ). Jede weitere Länge (von 1 bis $n-1$) tritt genau einmal auf. Für $p = \frac{1}{2}$ hat $L(p, n)$ also die Periodenlänge 1.

In dem Beispiel für $p = \frac{1}{3}$ (Seite 10) muß man zwischen $i = 0$ und $i = 1$ unterscheiden, da $t(p) = 2$. Man setze $n_0(\frac{1}{3}) = 4$, $s_1(\frac{1}{3}, 0) = 1$, $s_1(\frac{1}{3}, 1) = 2$ und $s_2(\frac{1}{3}) = 0$. Dann ist der Anfang gegeben durch $(\frac{n}{2})$, falls n gerade ist, durch

$(\frac{n+1}{2}, \frac{n+1}{2})$, falls n ungerade ist, und das Ende durch ( ). Jede weitere Länge tritt genau zweimal auf. Für $p = \frac{1}{3}$ hat $L(p,n)$ also die Periodenlänge 2.

Um nun den Listen $L(p,n)$ eine Periodenlänge zuordnen zu können, werden die Nullstellen der im folgenden Lemma auftretenden Funktion benötigt. Daher wird nun zunächst gezeigt, daß diese Funktion im Intervall $(0,1)$ genau eine Nullstelle hat.

Lemma 5:
Sei $f_t(p) = (1-p)^{t+1} + (1-p)^t - 1$.

Dann gilt:

1. $f_t(p)$ hat im Intervall $[0,1]$ genau eine Nullstelle in p und
2. $f_t(p)$ ist streng monoton fallend in t.

Beweis :

1. Es gilt $f_t(0) = 1$ und $f_t(1) = -1$, d.h. nach dem Mittelwertsatz existiert mindestens eine Nullstelle.

   $f_t(p)$ ist bezüglich p im Intervall $[0,1]$ monoton, denn
   $$\frac{\partial}{\partial p} f_t(p) = (-1) \cdot (t+1) \cdot (1-p)^t + (-1) \cdot t \cdot (1-p)^{t-1} < 0,$$
   da $0 < p < 1$. Damit ist die Nullstelle eindeutig.

2. Es gilt:
   $$\frac{\partial}{\partial t} f_t(p) = ln(1-p) \cdot (1-p)^{t+1} + ln(1-p) \cdot (1-p)^t < 0.$$
   Damit ist $f_t(p)$ streng monoton fallend in t. □

Definition :
Für eine Funktion $f(p)$ mit der Eigenschaft, daß $f(p)$ im Intervall $[0,1]$ genau eine Nullstelle hat, bezeichne $NS(f(p))$ diese Nullstelle.

Proposition 3:
$L(p,n)$ hat eine Periodenlänge von t genau dann, wenn
$$NS(f_t(p)) \leq p < NS(f_{t-1}(p)),$$
d.h. zu festem p hat $L(p,n)$ eine Periodenlänge
$$t(p) = \left\lceil \frac{-ln(2-p)}{ln(1-p)} \right\rceil.$$

Beweis :
Zunächst soll zu vorgegebenem t ein Intervall I_t gefunden werden, so daß für alle $p \in I_t$ gilt: $L(p,n)$ hat eine Periodenlänge von t. Dabei ist zu diesem Zeitpunkt noch nicht von Interesse, ob I_t alle p mit dieser Eigenschaft enthält. Ein solches Intervall I_t wird konstruiert, indem einige Forderungen an die Huffman–Codierung der Wahrscheinlichkeitsverteilung gestellt werden. Die wesentliche Idee dieses Beweises ist, zunächst eine Situation $(*)$ zu konstruieren, aus der sich mittels der Forderungen an die Huffman–Codierung eine Periodenlänge von t ergibt. Nach der Betrachtung dieser konstruierten Situation $(*)$ wird gezeigt, daß man beim Zusammenfassen der Wahrscheinlichkeitsverteilung $Prob(p,n)$ zwangsläufig in die Situation $(*)$ gelangt.

Um eine Periodenlänge von t nachzuweisen, muß gezeigt werden, daß alle zwischen einem Anfang und einem Ende auftretenden Längen genau t–mal auftreten. Die zu einer Wahrscheinlichkeit p_i gehörende Codewortlänge l_i kann aufgrund des Huffman–Verfahrens aus der Anzahl des Zusammenfassens der Wahrscheinlichkeit p_i ermittelt werden (siehe Bemerkung 4 aus Kapitel 3).

Man betrachte nun die folgende Situation:
Beim Zusammenfassen der Wahrscheinlichkeitsverteilung $Prob(p,n)$ habe man bereits s Schritte durchgeführt. Nach diesen s Schritten habe man die Wahrscheinlichkeit p_0 mit mindestens einer anderen Wahrscheinlichkeit zusammengefaßt. Sei k_0 die Summe der beiden kleinsten Wahrscheinlichkeit nach s Schritten. Nach dem Einsortieren von k_0 sei k_1 die Summe der beiden kleinsten Elemente der neu entstandenen Wahrscheinlichkeitsverteilung. Man erhält eine neue Wahrscheinlichkeitsverteilung durch Weglassen dieser beiden kleinsten Elemente und Einsortieren von k_1. Sei k_i die Summe der beiden kleinsten Wahrscheinlichkeiten der im Schritt $s+i-1$ entstandenen Wahrscheinlichkeitsverteilung nach Einsortieren von k_{i-1}. Man habe nun nach $s+t$ Schritten eine Wahrscheinlichkeitsverteilung P, die genau t Summen $k_0, \ldots, k_{t-1}$ enthalten soll. Diese t Summen sollen mit den t kleinsten noch nicht zusammengefaßten Wahrscheinlichkeiten von P die $2 \cdot t$ kleinsten noch vorhandenen Wahrscheinlichkeiten von P bilden, derart, daß diese $2 \cdot t$ Wahrscheinlichkeiten sortiert t Paare aus jeweils einem p_j und einem k_i bilden. Sei p_j die kleinste noch nicht zusammengefaßte Wahrscheinlichkeit von P. Dann werden die $2 \cdot t$ kleinsten Wahrscheinlichkeiten von P gebildet durch

$$k_0, p_j, k_1, p_{j-1}, \ldots, k_{t-1}, p_{j+1-t} \qquad (*)$$

bzw. die k_i dürfen die Positionen mit den nächstgrößeren ursprünglichen Wahrscheinlichkeiten p_{j-i} tauschen, da in diesen Fällen die in jedem Schritt zusammenzufassenden Wahrscheinlichkeiten die gleichen sind. So ist also auch zum Beispiel die Reihenfolge

$$p_j, k_0, k_1, p_{j-1}, \ldots, k_{t-1}, p_{j+1-t}$$

eine zulässige Situation $(*)$.

Nun soll in jedem weiteren Schritt gelten: Die Summe der beiden kleinsten Wahrscheinlichkeiten hat eine Schrittweite von $2 \cdot t - 2$ oder $2 \cdot t - 1$. Damit wird in jedem Schritt das neu entstandene k_i derart einsortiert, daß man wieder eine Situation wie $(*)$ enthält. Wird nun $Prob(p, n)$ soweit zusammengefaßt, bis nur noch t Wahrscheinlichkeiten $w_1, \ldots w_t$ in der resultierenden Wahrscheinlichkeitsverteilung enthalten sind, so lassen sich die Wahrscheinlichkeiten von P in t Gruppen unterteilen: Alle Wahrscheinlichkeiten von P sollen der gleichen Gruppe angehören, die nach dem Zusammenfassen in dem gleichen w_i enthalten sind. Bekommt nun jedes w_i durch das Huffman–Verfahren eine Codewortlänge zugewiesen und trennt man die w_i wieder auf, so wird in jedem Schritt einer Wahrscheinlichkeit und einer noch zusammengefaßten Summe von Wahrscheinlichkeiten die um 1 vergrößerte Codewortlänge von w_i zugewiesen. Für die Wahrscheinlichkeiten einer Gruppe gilt also, daß jede Codewortlänge, ab der um 1 vergrößerten Codewortlänge von w_i bis zu der um 1 verringerten Codewortlänge des größten noch vorhandenen k_i, genau einmal auftritt.
Sei l_1^* die größte der Codewortlängen der w_i und sei l_2^* die kleinste der Codewortlängen der in der Wahrscheinlichkeitsverteilung P noch vorhandenen k_i. Dann gilt für die Codewortlängen von P: Jede Codewortlänge l mit $l_1^* < l < l_2^*$ tritt in P genau t–mal auf.
Mit dieser betrachteten Vorschrift des Zusammenfassens läßt sich also aus der Situation $(*)$ eine Periodenlänge von t konstruieren.

Zur Vereinfachung betrachte ich nun eine Periodenlänge von $t+1$. Die Schrittweite soll in jedem Schritt $2 \cdot t$ oder $2 \cdot t + 1$ betragen. Man befinde sich nach s Schritten in der Situation $(*)$.

Bei einer Schrittweite von $2 \cdot t$ muß gelten:

$$k_i = k_{i-t-1} + (1-p)^{n-i-1-s} \cdot p$$

damit auch k_i einen Schritt von $2 \cdot t$ durchführt, muß gelten:

$$(1-p)^{n-i-t-s-1} \cdot p \leq k_i < (1-p)^{n-i-t-s-2} \cdot p.$$

Für eine Schrittweite von $2 \cdot t + 1$ gilt die gleiche Rekursionsvorschrift für k_i, jedoch ist die Bedingung:

$$(1-p)^{n-i-t-s-2} \cdot p \leq k_i < (1-p)^{n-i-t-s-3} \cdot p.$$

Analyse der Bedingungen:

a) $(1-p)^{n-i-t-s-1} \cdot p \leq k_i$

b) $k_i < (1-p)^{n-i-t-s-3} \cdot p$

zu a) $k_i \geq (1-p)^{n-i-t-s-1} \cdot p$

$\Leftrightarrow k_i = k_{i-t-1} + (1-p)^{n-i-s-1} \cdot p \geq (1-p)^{n-i-t-s-1} \cdot p$

$\Rightarrow (1-p)^{n-s-i} \cdot p + (1-p)^{n-i-s-1} \cdot p \geq (1-p)^{n-i-t-s-1} \cdot p$

$\Rightarrow (1-p)^{n-i-t-s-1} \cdot p \cdot ((1-p)^{t+1} + (1-p)^{t} - 1) \geq 0$

$\Leftrightarrow p \leq NS(f_t(p))$.

zu b) $k_i < (1-p)^{n-i-t-s-3} \cdot p$

$\Leftrightarrow k_i = k_{i-t-1} + (1-p)^{n-i-s-1} \cdot p < (1-p)^{n-i-t-s-3} \cdot p$

$\Rightarrow (1-p)^{n-i-s-2} \cdot p + (1-p)^{n-i-s-1} \cdot p < (1-p)^{n-i-t-s-3} \cdot p$

$\Rightarrow (1-p)^{n-i-t-s-3} \cdot p \cdot ((1-p)^{t+2} + (1-p)^{t+1} - 1) < 0$

$\Leftrightarrow p > NS(f_{t+1}(p))$.

Bleibt zu zeigen, daß man beim Zusammenfassen der Wahrscheinlichkeitsverteilung $Prob(p, n)$ in die Situation $(*)$ gelangt.

Würde man $Prob(p, n)$ ohne die Wahrscheinlichkeit p_0 mit dem Huffman–Verfahren codieren, so ergäbe sich eine Periodenlänge von t (siehe [GV 75]). Daher ist lediglich zu zeigen, daß durch das Vorhandensein von p_0 diese Eigenschaft nicht gestört wird.

In dem Schritt, in dem p_0 zusammengefaßt wird, können nur zwei Fälle eintreten:

1. Entweder p_0 wird mit einem bereits in einem vorhergehenden Schritt entstandenen k_i zusammengefaßt, oder
2. p_0 wird mit einer Wahrscheinlichkeit p_j zusammengefaßt.

Zu 1. :
Die vorhandene Wahrscheinlichkeitsverteilung wird wie folgt gebildet:
Die beiden kleinsten Wahrscheinlichkeiten sind p_0 und k_i, da sie in dem nun folgenden Schritt zusammengefaßt werden.Da jedes k_i mit einer ursprünglichen Wahrscheinlichkeit p_j zusammengefaßt würde, wird durch p_0 die ursprüngliche Wahrscheinlichkeit p_j ersetzt. D.h. die vier kleinsten Wahrscheinlichkeiten sind

$$p_0, k_i, p_j, k_{i+1} \quad \text{oder} \quad k_i, p_0, p_j, k_{i+1}.$$

Bei einer Periodenlänge von t wird nun k_{i+t} gebildet durch $p_0 + k_i$. Da p_0 eine größere Wahrscheinlichkeit ersetzt, kann die Schrittweite von k_{i+t} nicht größer werden als die Schrittweite von k_{i+t-1}. Es gilt aber $k_{i+t} \geq k_{i+t-1}$ aufgrund der Konstruktion. Daher bildet k_{i+t} mit der oberhalb von k_{i+t-1} befindlichen Wahrscheinlichkeit p_{j-t+1} ein neues Paar und man befindet sich in der Situation $(*)$.

Zu 2. :
In diesem Fall sind die vier kleinsten Wahrscheinlichkeiten

$$p_0, p_j, k_{i+1}, p_{j-1} \quad \text{oder} \quad p_j, p_0, k_{i+1}, p_{j-1}.$$

Bildet man nun $k_{i+t+1} := p_0 + p_j$, so können beim Einsortieren wieder zwei Fälle eintreten:

Entweder man erhält einen Schritt nach dem Einsortieren von k_{i+t+1}

$$\text{a)}\ \ldots, k_{i+t}, p_{j-t}, k_{i+t+1}, k_{i+t+2}, p_{j-t-1}, \ldots$$

oder

$$\text{b)}\ \ldots, k_{i+t}, k_{i+t+1}, p_{j-t}, k_{i+t+2}, p_{j-t-1}, \ldots$$

In Fall a) befindet man sich nach dem Zusammenfassen von k_{i+t} mit p_{j-t} in einer Situation mit einem Störelement ähnlich Fall 1, in Fall b) befindet man sich in dem Schritt in einer Situation wie Fall 1, in dem k_{i+t} und k_{i+t+1} zusammengefaßt werden müssen.

Anhand der bisherigen Ergebnisse lassen sich nun einige Feststellungen treffen:

- Hat man eine Wahrscheinlichkeitsverteilung $Prob(p, n)$, bei der man sich beim Zusammenfassen nach s Schritten in der Situation $(*)$ befindet, so befindet man sich auch bei allen $Prob(p, n + i \cdot t)$ für alle $i \in \mathbb{N}$ nach s Schritten in der Situation $(*)$.
- Die Länge des Anfangs von $L(p, n)$ hängt von der Schrittweite von p_0 ab.
- Aus Fall 2 läßt sich entnehmen, daß, wenn zwei Wahrscheinlichkeiten p und q zu der gleichen Periodenlänge bei $L(p)$ und $L(q)$ führen, dies nicht bedeutet, daß $L(p, n)$ und $L(q, n)$ gleich sein müssen.
- Betrachtet man den Schritt von n nach $n + t$, so gilt für den Anfang von $L(p, n+t)$, daß er aus dem um t Werten verlängerten Mittelteil von $L(p, n)$ entstanden ist. Daher enthält $L(p, n+t)$ aufgrund der rekursiven Konstruktion die gleiche Anzahl an Elementen wie $L(p, n)$. Die einzelnen Längen des Anfanges von $L(p, n)$ werden jeweils um 1 vergrößert und bilden dann den Anfang von $L(p, n + t)$. Daher existiert eine Konstante $s_1(p, \mathrm{n \bmod t})$, die die Länge des Anfanges von $L(p, n)$ darstellt.

Aufgrund der Konstruktion haben die Summen der beiden jeweils kleinsten Wahrscheinlichkeiten eines jeden Schrittes nach Eintreten der Situation $(*)$ konstante Schrittweite. Nachdem die Wahrscheinlichkeiten von $Prob(p, n)$ soweit zusammengefaßt sind, daß weniger vorhanden sind als die Schrittweite der beiden kleinsten Wahrscheinlichkeiten darstellen würde, so ist die Summe der beiden kleinste Wahrscheinlichkeiten jeweils die größte der noch vorhandenen. Daher verhält sich der Rest der Wahrscheinlichkeitsverteilung beim Zusammenfassen ebenso wie die Gleichverteilung. Deshalb sind die Codewortlängen des Endes von $L(p, n)$ vorherbestimmbar und das Ende von $L(p, n)$ hat konstante Länge für alle n.
Die Auswertung dieser Eigenschaften erfolgt zu einem späteren Zeitpunkt. Für den Beweis von Proposition 3 reicht jedoch die Existenz der Konstanten.

Zu jedem t haben wir nun ein Intervall I_t konstruiert, so daß für alle $p \in I_t$ gilt: $L(p, n)$ hat Periodenlänge t. Da die Intervalle I_t das gesamte Intervall $(0, 1)$

überdecken, ist damit bewiesen, daß $L(p,n)$ für jedes $p \in (0,1)$ eine Periodenlänge hat und daß die Zuordnung zwischen Wahrscheinlichkeit p und Periodenlänge t bijektiv ist.

Bleibt zu zeigen, daß $t = \left\lceil \frac{-ln(2-p)}{ln(1-p)} \right\rceil$.

Es gilt: $f_t(p)$ ist streng monoton fallend in p, da

$$\begin{aligned}\frac{\partial}{\partial p} f_t(p) &= \frac{\partial}{\partial p}\left((1-p)^{t+1} + (1-p)^t - 1\right) \\ &= -(t+1)\cdot(1-p)^t - t\cdot(1-p)^{t-1} < 0.\end{aligned}$$

Seien $q_1 = NS(f_t(p))$ und $q_2 = NS(f_{t-1}(p))$. Dann gilt $q_1 \leq p < q_2$.
Da $f_t(p)$ streng monoton fallend ist in p gilt:

$f_t(q_1) \geq f_t(p) > f_t(q_2)$ und $f_{t-1}(q_1) \geq f_{t-1}(p) > f_{t-1}(q_2)$.

Damit gilt also: $0 = f_t(q_1) \geq f_t(p)$ und $f_{t-1}(p) > f_{t-1}(q_2) = 0$

$$\Leftrightarrow f_t(p) \leq 0 < f_{t-1}(p)$$

$$\Leftrightarrow (1-p)^{t+1} + (1-p)^t - 1 \leq 0 < (1-p)^t + (1-p)^{t-1} - 1$$

$$\Rightarrow t \geq \frac{-ln(2-p)}{ln(1-p)} > t-1 \quad \Rightarrow t = \left\lceil \frac{-ln(2-p)}{ln(1-p)} \right\rceil.$$

Damit ist Proposition 3 bewiesen. □

Wenn auch in diesen Intervallen $I_t = [NS(f_t(p)), NS(f_{t-1}(p)))$ eine gleiche Periodenlänge gilt, so heißt dies jedoch nicht, daß die gleiche Verteilung der Codewortlängen vorliegt.

Man betrachte zum Beispiel die Wahrscheinlichkeiten $p = \frac{1}{3}$ und $q = \frac{1}{4}$.

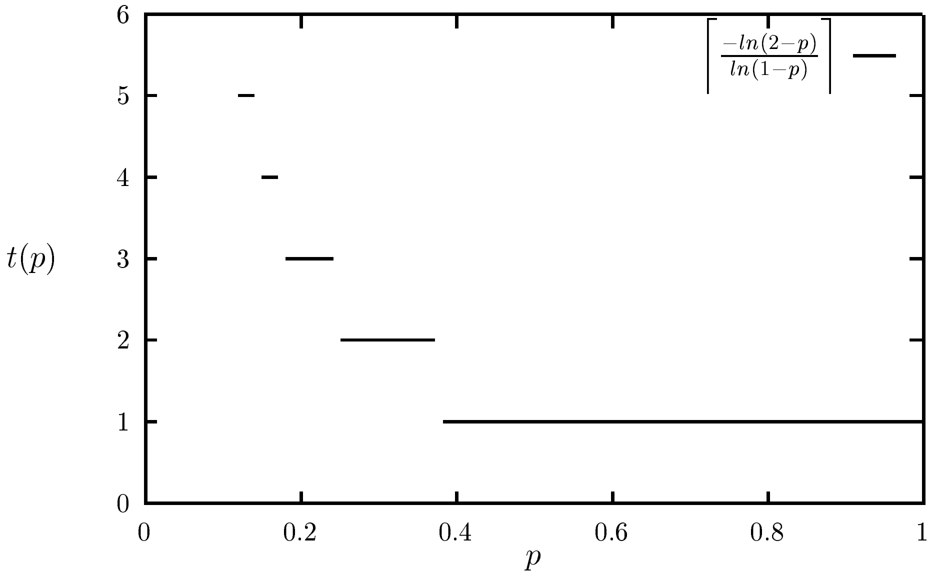
Nach Proposition 3 gilt, daß $L(p,n)$ und $L(q,n)$ eine Periodenlänge von 2 haben.

Betrachtet man aber die Codewortlängen für $n = 3$, so ergibt sich:
$L(\frac{1}{3}, 3) = (2,2,2,2)$ und $L(\frac{1}{4}, 3) = (1,3,3,2)$.

Einige Intervalle, d.h. einige Nullstellen obigen Polynoms, seien in folgender Tabelle aufgeführt.

Periodenlänge	untere Intervallgrenze	obere Intervallgrenze
1	0.381966	1.000000
2	0.245122	0.381966
3	0.180828	0.245122
4	0.143325	0.180828
5	0.118728	0.143325
6	0.101346	0.118728
7	0.088408	0.101346
8	0.078401	0.088408
9	0.070430	0.078401
10	0.063931	0.070430
11	0.058530	0.063931
12	0.053972	0.058530
13	0.050071	0.053972
14	0.046697	0.050071
15	0.043749	0.046697
16	0.041152	0.043749
17	0.038845	0.041152
18	0.036784	0.038845
19	0.034930	0.036784
20	0.033253	0.034930

Zur Veranschaulichung wird das Verhältnis von p zu t in einer Grafik dargestellt:



Um zu einem späteren Zeitpunkt die Konstante $n_0(p)$ abschätzen zu können, ist es notwendig, die Schrittweite der Wahrscheinlichkeit p_0 zu kennen.

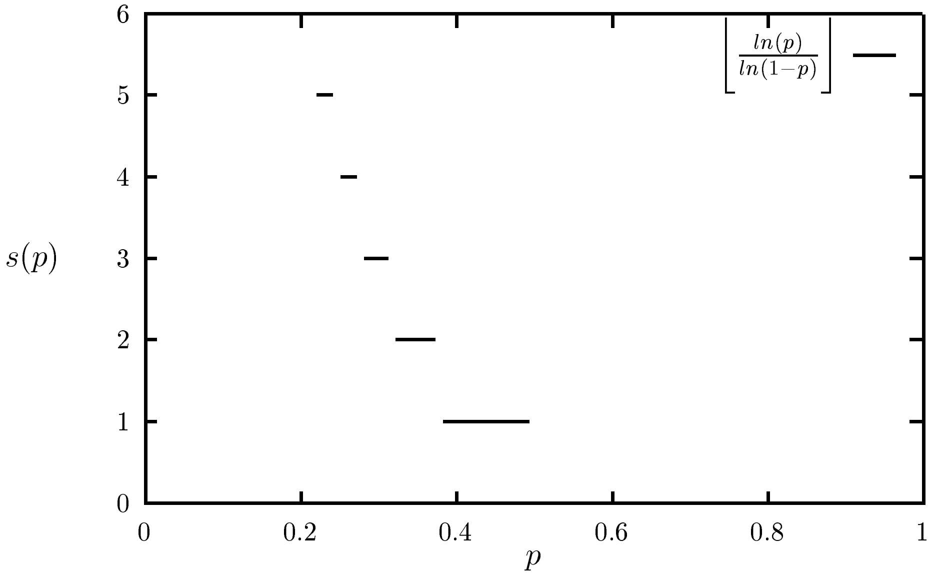
Lemma 6:
p_0 hat eine Schrittweite von s genau dann, wenn

$$s = \left\lfloor \frac{ln(p)}{ln(1-p)} \right\rfloor .$$

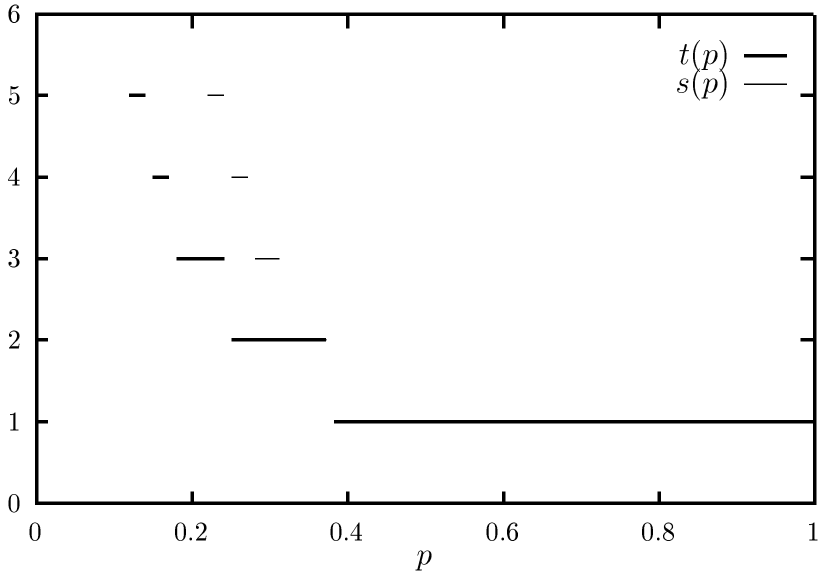
Beweis :
p_0 hat eine Schrittweite von s genau dann, wenn $p_{n-s+1} \leq p_0 < p_{n-s}$.
D.h., $(1-p)^{n-s} \cdot p \leq (1-p)^n < (1-p)^{n-s-1} \cdot p$.
Daher gilt $s = \left\lfloor \frac{ln(p)}{ln(1-p)} \right\rfloor$. □

Sei auch hier das Verhältnis von p zur Schrittweite von p_0 dargestellt:

Schrittweite	untere Intervallgrenze	obere Intervallgrenze
0	0.500000	1.000000
1	0.381966	0.500000
2	0.317672	0.381966
3	0.275508	0.317672
4	0.245122	0.275508
5	0.221910	0.245122
6	0.203455	0.221910
7	0.188347	0.203455
8	0.175700	0.188347
9	0.164921	0.175700
10	0.155603	0.164921
11	0.147449	0.155603
12	0.140243	0.147449
13	0.133819	0.140243
14	0.128049	0.133819
15	0.122833	0.128049
16	0.118090	0.122833
17	0.113755	0.118090
18	0.109774	0.113755
19	0.106104	0.109774
20	0.102708	0.106104



Man kann den Grafiken und den Tabellen entnehmen, daß bei fallendem p die Schrittweite von p_0 schneller wächst, als die Periodenlänge der Codierung.
Sei $t(p)$ die Periodenlänge, die sich bei einer Wahrscheinlichkeit von p für die Codierung der Wahrscheinlichkeitsverteilung $Prob(p, n)$ ergibt. Sei $s(p)$ die Schrittweite von p_0 in der Wahrscheinlichkeitsverteilung $Prob(p, n)$ bei einer vorgegebenen Wahrscheinlichkeit p. Dann wird das Wachstumsverhalten beider Funktionen an folgender Grafik deutlich:



Die innerhalb gleicher Periodenlänge unterschiedlich auftretenden Schrittweiten von p_0 sind dafür verantwortlich, daß auch bei gleicher Periodenlänge der Codierungen zu unterschiedlichen Wahrscheinlichkeiten p und q die Codewortlängen nicht gleich sein müssen.

1.6 Eine obere Schranke für n

Zu einer gegebenen Wahrscheinlichkeit p habe die Liste $L(p,n)$ eine Periodenlänge von t. Es wird gezeigt, daß $kost(p,b,n)$ nur dann minimal sein kann, wenn $0 \leq n < n_0(p) + t$ oder $n = \infty$ gilt. Der Wert $n_0(p)$ wird am Ende dieses Kapitels nach oben abgeschätzt. Diese Behauptung wird bewiesen, indem man $kost(p,b,n)$ mit $kost(p,b,n+t)$ in Beziehung setzt.

Lemma 7:
Sei $k = \lfloor log_2 t \rfloor$. Für alle $n \geq n_0(p)$ gilt:
Das Ende von $L(p,n)$ wird durch $(2^{k+1} - t)$– mal die Länge $k+1$ gebildet.

Beweis :
Für $n \geq n_0(p)$ ist l_0 größer als die größte im Mittelteil von $L(p,n)$ auftretende Codewortlänge. Daher ist l_0 auch größer als jede im Ende von $L(p,n)$ auftretende Länge. Daher hat die Codierung von p_0 keinen Einfluß mehr auf die Codierung des Endes.
Wie in Kapitel 5 (erster Teil des Beweises von Proposition 3) gezeigt wurde, haben die dort benutzten k_i in jedem Schritt eine Schrittweite von $2 \cdot t - 1$ oder $2 \cdot t$. Von der Situation $(*)$ (Seite 21) ausgehend, geschieht dies so lange, bis die zu dem Mittelteil von $L(p,n)$ gehörenden Wahrscheinlichkeiten durchlaufen wurden. Anschließend sieht die Situation wie folgt aus:

$$k_i, p_t, k_{i+1}, p_{t-1}, \ldots, k_{i+t-1}, p_1 \text{ oder } p_t, k_i, p_{t-1}, k_{i+1}, \ldots, p_1, k_{i+t-1}.$$

Da in den folgenden Schritten weniger als $2 \cdot t - 1$ Wahrscheinlichkeiten vorhanden sind, kann auch die Schrittweite nicht mehr $2 \cdot t - 1$ betragen. Die in jedem Schritt neu entstandene Wahrscheinlichkeit ist nun jeweils die größte der noch vorhandenen Wahrscheinlichkeiten. D.h. bei dem weiteren Zusammenfassen ergibt sich:

$$k_{i+1}, p_{t-1}, \ldots, k_{i+t-1}, p_1, k_{i+t} \text{ und nach } t \text{ Schritten: } k_{i+t}, k_{i+t+1}, \ldots, k_{i+2\cdot t-1}.$$

Dieses Schema setzt sich aufgrund der Konstruktion der k_i fort, d.h. in jedem Schritt ergibt die Summe der beiden kleinsten Wahrscheinlichkeiten eine Wahrscheinlichkeit, die die größte der noch vorhandenen Wahrscheinlichkeiten darstellt. Eine Wahrscheinlichkeitsverteilung, die sich beim Zusammenfassen ebenso verhält, ist die Gleichverteilung. Somit lassen sich aber die Codewortlängen vorhersagen.

Sei $k = \lfloor log_2 t \rfloor$. Ist $t = 2^k$, so nehme man als Zahlenfolge t mal den Wert k.

Ist $2^k < t < 2^{k+1}$, so nehme man $2^{k+1} - t$ mal den Wert k und setze $2 \cdot t - 2^{k+1}$ mal den Wert $k+1$ davor.

Beim weiteren Trennen der Wahrscheinlichkeiten entsteht in jedem Schritt eine originale und eine noch zusammengefaßte Wahrscheinlichkeit. Dabei entstehen in jedem Schritt zwei Codeworte, deren Längen um 1 größer ist, als die Länge des eliminierten Codewortes (siehe Huffman–Verfahren Seite 13). Damit ist Lemma 7 bewiesen. $\square$

Proposition 4:
Sei $k = \lfloor log_2 t \rfloor$. Für alle $n \geq n_0(p)$ gilt:

$$kost(p, b, n+t) = (1-p)^t \cdot kost(p, b, n) + 1 + k + (1-p)^{2^{k+1}-t} - (k+1) \cdot (1-p)^t.$$

Beweis :
Am Ende des Beweises von Proposition 3 wurde gezeigt, daß man für alle $n \geq n_0(p)$ den Anfang von $L(p, n+t)$ erhält, indem man jeden im Anfang von $L(p, n)$ enthaltenen Wert um 1 vergrößert. In Lemma 7 wurde gezeigt, daß für alle $n \geq n_0(p)$ das Ende von $L(p, n)$ gleich ist. Damit wird also der Mittelteil von $L(p, n)$ um genau t Einträge vergrößert.
Mit diesen Ergebnissen läßt sich nun die Kostenfunktion für $n+t$ aus der Kostenfunktion für n für alle $n \geq n_0(p)$ berechnen. Man erhält also den Wert für $kost(p, b, n+t)$ aus $kost(p, b, n)$, indem man $kost(p, b, n)$ mit $(1-p)^t$ multipliziert, anschließend, da jede Codewortlänge um eins vergrößert wird, muß jede Wahrscheinlichkeit einmal hinzuaddiert werden. Dies geschieht durch die Addition von 1, da die Summe der Wahrscheinlichkeiten 1 ergibt. Da die Wahrscheinlichkeiten von p_1 bis p_{t-1} zuviel hinzuaddiert wurden, addieren wir anstatt der Länge $k+2$ nur $(k+1)$– mal die entsprechenden Wahrscheinlichkeiten und anstatt $k+1$ nur k–mal.

Es gilt also:

$$\begin{aligned} kost(p, b, n+t) &= (1-p)^t \cdot kost(p, b, n) + 1 + \sum_{i=1}^{2^{k+1}-t} p_i \cdot k + \sum_{i=2^{k+1}-t+1}^{t} p_i \cdot (k+1) \\ &= (1-p)^t \cdot kost(p, b, n) + 1 + k + (1-p)^{2^{k+1}-t} - (k+1) \cdot (1-p)^t. \end{aligned}$$

Damit ist Proposition 4 bewiesen. $\square$

Proposition 5:

Sei $c_0 = 1 + \lfloor log_2 t \rfloor + \frac{(1-p)^{2^{\lfloor log_2 t \rfloor + 1} - t}}{1-(1-p)^t}$.

Es gilt für alle $n \geq n_0(p)$:

- Ist $kost(p, b, n) = c_0$, so ist auch $kost(p, b, n + i \cdot t) = c_0$ für alle $i \in \mathbb{N}$.
- Ist $kost(p, b, n) > c_0$, so ist die Folge $kost(p, b, n + i \cdot t)$ als Folge in i streng monoton fallend.
- Ist $kost(p, b, n) < c_0$, so ist die Folge $kost(p, b, n + i \cdot t)$ als Folge in i streng monoton steigend.

- $\lim_{n\to\infty} kost(p,b,n) = c_0$.

Beweis :
Sei $k = \lfloor log_2 t \rfloor$. Aus Proposition 4 folgt, daß $kost(p,b,n+t) = kost(p,b,n)$ genau dann, wenn

$$\begin{aligned} kost(p,b,n) &= \tfrac{1+k+(1-p)^{2^{k+1}-t}-(k+1)\cdot(1-p)^t}{1-(1-p)^t} \\ &= 1+k+\tfrac{(1-p)^{2^{k+1}-t}}{1-(1-p)^t}. \end{aligned}$$

Da c_0 nicht von n abhängt, gilt für alle $n \geq n_0(p)$:

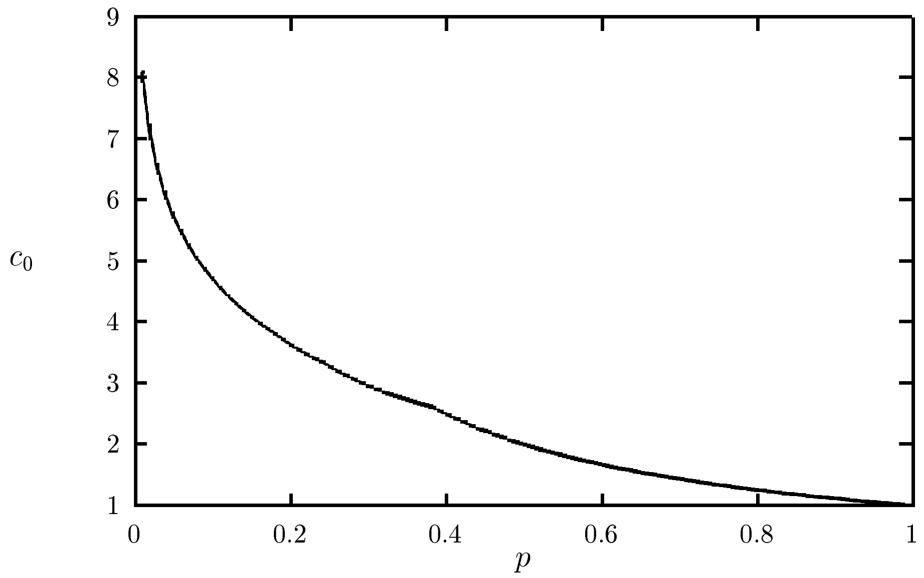
- Ist $kost(p,b,n) = c_0$, so auch $kost(p,b,n+i\cdot t)$ für alle $i \in \mathbb{N}$.
- Ist $kost(p,b,n) > c_0$, so ist die Folge $kost(p,b,n+i\cdot t)$ streng monoton fallend in i.
- Ist $kost(p,b,n) < c_0$, so ist die Folge $kost(p,b,n+i\cdot t)$ streng monoton steigend in i.
- $\lim_{n\to\infty} kost(p,b,n) = c_0$.

Damit ist Proposition 5 bewiesen. □

Für die Periodenlänge t, die sich bei einer gegebenen Wahrscheinlichkeit p für $L(p,n)$ ergibt, gilt nach Proposition 3:

$$t = \left\lceil \frac{-ln(2-p)}{ln(1-p)} \right\rceil.$$

Damit läßt sich c_0 als eine nur von p abhängige Funktion darstellen.



Die in diesem Graf sichtbare „Delle“ beruht nicht auf Rundungsfehlern. Sie entsteht aufgrund der Gauss–Klammern in der Funktion c_0.

Um den bei minimalem $kost(p, b, n)$ möglichen Bereich für n etwas genauer bestimmen zu können, soll nun $n_0(p)$ nach oben begrenzt werden. Damit man ein n hat, das größer ist als $n_0(p)$, muß die, der Wahrscheinlichkeit p_0 zuzuordnende Codewortlänge, größer sein als die für das Ende berechnete Codewortlänge.
Seien q_1 und q_2 zwei beliebige Wahrscheinlichkeiten einer gegebenen Wahrscheinlichkeitsverteilung. Seien l_1^* die zu q_1 und l_2^* die zu q_2 gehörenden Codewortlängen, die man erhält, wenn man die Wahrscheinlichkeitsverteilung mit dem Huffman–Verfahren codiert.
Dann gilt:

$$\text{Ist } q_1 \geq q_2, \text{ so ist } l_1^* \leq l_2^*.$$

In Kapitel 5 wurde die Schrittweite von p_0 bestimmt als

$$s := \left\lfloor \frac{ln(p)}{ln(1-p)} \right\rfloor.$$

Lemma 8:
Für alle $p \in (0, 1)$ gilt: $n_0(p) < s + 3 \cdot t$.

Beweis :
Das Ende von $L(p, n)$ hat eine Länge, die kleiner oder gleich t ist (nach Lemma 8). Die Situation $(*)$ (erster Teil des Beweises von Proposition 3) tritt nach spätestens $s+2\cdot t$ Schritten ein, d.h., der Anfang von $L(p, n)$ hat eine Länge, die kleiner ist als $s+2\cdot t$ (siehe Teil 2 des Beweises von Proposition 3). In der Situation $(*)$ gehören die noch nicht zusammengefaßten Wahrscheinlichkeiten entweder zum Mittelteil oder zum Ende. Damit wird durch die Summe der Länge des Anfangs von $L(p, n)$

und der Länge des Endes von $L(p, n)$ der Wert $n_0(p)$ nach oben beschränkt (beide Größen sind für alle $n \geq n_0(p)$ konstant). Daher kann $n_0(p)$ durch $s + 3 \cdot t$ nach oben begrenzt werden. □

Aus Lemma 8 und Proposition 5 zusammen ergibt sich:

Proposition 6:
Ist $kost(p, b, n)$ für n minimal, so gilt $0 \leq n < s + 3 \cdot t$ oder $n = \infty$.

1.7 Eine untere Schranke für n

Die Codierung von p_0 hat auch dann keinen Einfluß auf die Codierung aller anderen Wahrscheinlichkeiten, wenn p_0 größer oder gleich der Summe aller anderen Wahrscheinlichkeiten ist. Das passiert, wenn nur wenige Experten eingesetzt werden. Speziell für $n < t$ kann die Liste der Codeworte $L(p, n)$ bestimmt werden. In dem verbleibenden Rest dieses Kapitels werden die in Proposition 6 möglichen Werte für n auf die Werte $n = 0$ oder $t \leq n < s + 3 \cdot t$ oder $n = \infty$ beschränkt.

Lemma 9:

Sei $k = \lfloor log_2 n \rfloor$. Dann ergibt sich als Codewortlängen für $1 \leq n < t$:

- $l_0 = 1$,
- $l_i = k + 1$ für alle $1 \leq i \leq 2^{k+1} - n$ und
- falls $2^{k+1} - n < n$, so ist $l_i = k + 2$ für alle $2^{k+1} - n < i \leq n$.

Beweis :
Die Aussage ist bewiesen, wenn p_0 größer ist als die Summe aller anderen vorhandenen Wahrscheinlichkeiten für alle $1 \leq n < t$.
In Kapitel 5 (Proposition 3) wurde gezeigt, daß $L(p, n)$ eine Periodenlänge von t hat, wenn

$$(1-p)^{t+1} + (1-p)^t - 1 < 0 \leq (1-p)^t + (1-p)^{t-1} - 1, \text{ d.h.}$$
$$(1-p)^{t+1} + (1-p)^t < 1 \leq (1-p)^t + (1-p)^{t-1}.$$

Nun gilt, daß $(1-p)^t$ streng monoton fallend ist in t, da $0 < (1-p) < 1$.

Da $1 \leq (1-p)^t + (1-p)^{t-1}$, gilt: $(1-p)^{t-1} \geq \frac{1}{2}$.

Damit gilt: $(1-p)^i \geq \frac{1}{2}$ für alle $0 \leq i < t$

$$\Leftrightarrow 1 - (1-p)^i \leq (1-p)^i \text{ für alle } 0 \leq i < t$$
$$\Leftrightarrow \sum_{j=0}^{i-1} (1-p)^j \cdot p \leq (1-p)^i \text{ für alle } 1 \leq i < t.$$

D.h. bei einer Periodenlänge von t gilt für $0 \leq n < t$, daß $(1-p)^n$ größer ist als die Summe aller anderen Wahrscheinlichkeiten von $Prob(p, n)$. Damit wird die Codewortlänge von $p_0 = (1-p)^n$ eine Länge von 1 haben. Für die anderen Wahrscheinlichkeiten gilt, daß die Summe der beiden kleinsten Wahrscheinlichkeiten eine Sprungweite hat, die entweder $2 \cdot t$ oder $2 \cdot t - 1$ entsprechen muß. Nun kann ein solcher Sprung für $0 \leq n < t$ nicht erfolgen. Hier ist die Summe der beiden kleinsten Wahrscheinlichkeiten in jedem Schritt die zweitgrößte Zahl, da sie nicht größer werden kann als $(1-p)^n$. Da dies nun in jedem Schritt gilt, ergeben sich als Codewortlängen für $0 \leq n < t$:

- Für $n = 0$ ist kein Codewort vorhanden.
- Für $n = 1$ gilt als Codierung $l_0 = l_1 = 1$.
- Für $1 < n < t$ gilt $l_0 = 1$. Die anderen Codewortlängen werden analog der Bestimmung des Endes für $n \geq n_0(p)$ durchgeführt: Sei für $n-1$ die Liste $L(p, n-1)$ gegeben. Man streiche nun l_1 und schreibe dafür zweimal den Wert $1 + l_1$ zwischen l_0 und l_n in die Liste $L(p, n-1)$. Die neu entstandene Liste entspricht nun $L(p, n)$.

Für $n = 2$ ergibt sich also $L(p, 2) = (1, 2, 2)$ als Codewortlängen für alle p, die eine Periodenlänge größer als 1 haben.

Sei $k = \lfloor log_2 n \rfloor$. Dann ergibt sich als Codewortlängen für $1 \leq n < t$:

- $l_0 = 1$.
- $l_i = k + 1$ für alle $1 \leq i \leq 2^{k+1} - n$.
- falls $2^{k+1} - n < n$, so ist $l_i = k + 2$ für alle $2^{k+1} - n < i \leq n$. $\square$

Nun soll die Kostenfunktion für $1 \leq n < t$ mit Hilfe von Lemma 9 in eine geschlossene Form überführt werden:

$$
\begin{aligned}
kost(p, b, n) &= (1 + b) \cdot p_0 + \sum_{i=1}^{n} p_i \cdot l_i \\
&= (1 + b) \cdot p_0 + \sum_{i=1}^{2^{k+1}-n} p_i \cdot (k+1) + \sum_{2^{k+1}-n+1}^{n} p_i \cdot (k+2) \\
&= (1 + b) \cdot p_0 + (k+1) \cdot \sum_{i=1}^{n} p_i + \sum_{i=2^{k+1}-n+1}^{n} p_i \\
&= (1 + b) \cdot p_0 + (k+1) \cdot (1 - p_0) \\
&+ (1-p)^{2^{k+1}-n} \cdot (1 - (1-p)^{2 \cdot n - 2^{k+1}}) \\
&= b \cdot p_0 + (k+1) \cdot (1 - p_0) + (1-p)^{2^{k+1}-n}.
\end{aligned}
$$

Mit dieser Darstellung der Kostenfunktion für $1 \leq n < t$ gilt:

$$kost(p, b, n) \geq b$$

$$\Leftrightarrow b \cdot p_0 + (k+1) \cdot (1 - p_0) + (1-p)^{2^{k+1}-n} \geq b$$

$$\Leftrightarrow b \leq \frac{(k+1)\cdot(1-p_0)+(1-p)^{2^{k+1}-n}}{1-p_0}$$

$$\Leftrightarrow b \leq 1+k+\frac{(1-p)^{2^{k+1}-n}}{1-p_0}.$$

Sei

$$f(p,n) = 1 + \lfloor log_2 n\rfloor + \frac{(1-p)^{2^{\lfloor log_2 n\rfloor+1}-n}}{1-(1-p)^n}.$$

Proposition 7:

Sei $c_0(p) = 1 + \lfloor log_2 t\rfloor + \frac{(1-p)^{2^{\lfloor log_2 t\rfloor+1}-t}}{1-(1-p)^t}$.

Dann ist $kost(p,b,n) \geq b$ für alle $b \leq c_0$ und $1 \leq n < t$.

In diesem Fall kann es also nicht optimal sein, zwischen 1 und $t-1$ Experten einzusetzen.

Beweis :
Obige Behauptung ist für $0 \leq n < t$ bewiesen, wenn gilt

$$f(p,i) \geq f(p,t) \text{ für alle } 1 \leq i < t.$$

Dabei ist $f(p,t) = c_0(p)$.

Um die obige Proposition beweisen zu können, wird die Funktion $f(p,i)$ für alle $i \in \mathbb{R}^+$ betrachtet. Es wird gezeigt, daß $f(p,i)$ für $i \in \mathbb{R}^+ \setminus \{1,2,4,8,16,\ldots\}$ differenzierbar in i ist. Die Ableitung besitzt nur eine Nullstelle. Sei i_0 diese Nullstelle. Dann ist $f(p,i)$ monoton fallend in i für alle $i < i_0$ und $f(p,i)$ monoton steigend in i für alle $i > i_0$. Weiter wird gezeigt, daß $f(p,i)$ für $i \in \mathbb{R}^+$ stetig ist in i. Es gilt: $i_0 > t-1$. Mit $f(p,t) > f(p,t-1)$ und dem Monotonieverhalten von $f(p,i)$ folgt, daß $f(p,i)$ für $i \in \mathbb{N}$ minimal ist genau dann, wenn $i = t$.

Es gilt für alle i aus $\mathbb{R}^+ \setminus \{1,2,4,8,16,\ldots\}$:

$$\frac{\partial}{\partial i} f(p,i) = (1-p)^{2^{\lfloor log i\rfloor+1}} \cdot \frac{\partial}{\partial i} \cdot \frac{1}{(1-p)^i-(1-p)^{2\cdot i}}$$

$$\Leftrightarrow \frac{\partial}{\partial i} f(p,i) = (1-p)^{2^{\lfloor log i\rfloor+1}} \cdot \frac{-ln(1-p)}{((1-p)^i-(1-p)^{2\cdot i})^2} \cdot ((1-p)^i - 2\cdot(1-p)^{2\cdot i})$$

$$\Rightarrow \frac{\partial}{\partial i} f(p,i) = 0$$

$$\Leftrightarrow (1-p)^i - (1-p)^{2\cdot i} = 0$$

$$\Leftrightarrow i = \frac{-ln2}{ln(1-p)}.$$

Sei $i_0 := \frac{-ln2}{ln(1-p)} \notin \{1, 2, 4, 8, \ldots\}$. Dann gilt für alle $i \in \mathbb{R}^+$:

- Für alle $i < i_0$ ist $f(p,i)$ für alle p streng monoton fallend in i.
- Für alle $i > i_0$ ist $f(p,i)$ für alle p streng monoton steigend in i.

Nun soll gezeigt werden, daß $f(p,i)$ eine auf $\mathbb{R}^+$ stetige Funktion in i ist.

Sei $i_1 = 2^k + \epsilon$ und $i_2 = 2^k - \epsilon$.

$$f(p,i_1) = 1 + k + \frac{(1-p)^{2^k-\epsilon}}{1-(1-p)^{i_1}} = k + \frac{(1-p)^{2^k-\epsilon} + 1 - (1-p)^{i_1}}{1-(1-p)^{i_1}}$$

$$\epsilon \to 0 \Rightarrow f(p,i_1) \to k + \frac{1}{1-(1-p)^{2^k}}$$

$$f(p,i_2) = k + \frac{(1-p)^{\epsilon}}{1-(1-p)^{i_2}}$$

$$\epsilon \to 0 \Rightarrow f(p,i_2) \to k + \frac{1}{1-(1-p)^{2^k}}$$

$\Rightarrow f(p,i)$ ist stetig in i für alle $i \in \mathbb{R}^+$.

Daher gilt für alle $i \in \mathbb{R}^+$:

- Für alle $i < i_0$ ist $f(p,i)$ für alle p streng monoton fallend in i.
- Für alle $i > i_0$ ist $f(p,i)$ für alle p streng monoton steigend in i.

$\Rightarrow f(p,i)$ ist minimal genau dann, wenn $i = \frac{-ln2}{ln(1-p)} = i_0$.

Um i ganzzahlig zu wählen, ist $f(p,i)$ minimal für ein ganzzahliges i, wenn i in der Nähe von i_0 liegt.

Ist nun $t \leq i_0$, so ist also $f(p,i)$ für $i = t$ minimal.

Nun soll gezeigt werden, daß $f(p,t) < f(p,t-1)$. Mit dieser Bedingung gilt, daß $i_0 > t-1$. Daher ist dann auch in diesem Fall $f(p,i)$ für $i = t$ minimal.

1. $\lfloor log_2(t-1)\rfloor = \lfloor log_2 t\rfloor = k$

$$f(p,t) - f(p,t-1) = \frac{(1-p)^{2^{k+1}-t}}{1-(1-p)^t} - \frac{(1-p)^{2^{k+1}-t+1}}{1-(1-p)^{t-1}}$$

$$< (1-p)^{2^{k+1}-t+1} \cdot \frac{-p\cdot(1-p)^{t-1}}{(1-(1-p)^t)\cdot(1-(1-p)^{t-1})} < 0.$$

2. $\lfloor log_2(t-1)\rfloor + 1 = \lfloor log_2 t\rfloor = k+1$

$$f(p,t) - f(p,t-1) = 1 + \frac{(1-p)^{2^{k+2}-t}}{1-(1-p)^t} - \frac{(1-p)^{2^{k+1}-t+1}}{1-(1-p)^{t-1}}$$

$$= 1 + (1-p)^{2^{k+1}-t} \cdot \left(\frac{(1-p)^{2^{k+1}}}{1-(1-p)^t} - \frac{1-p}{1-(1-p)^{t-1}}\right)$$

$$= 1 - \left(\frac{(1-p)\cdot(1-(1-p)^t) - (1-p)^{2^{k+1}}\cdot(1-(1-p)^{t-1})}{(1-(1-p)^{t-1})\cdot(1-(1-p)^t}\right)$$

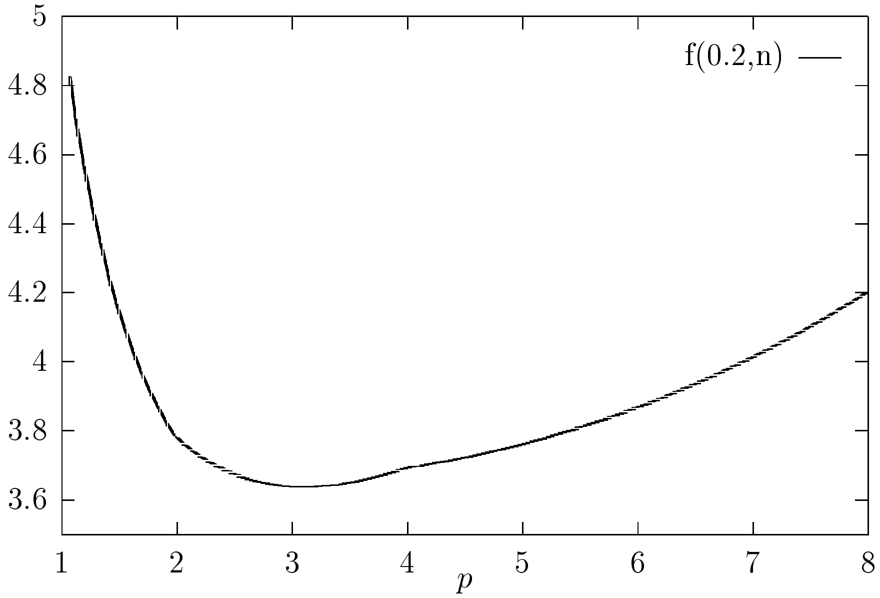
$$= \frac{p+(1-p)^{t+1}-(1-p)^{t-1}}{(1-(1-p)^t)\cdot(1-(1-p)^{t-1})} \leq 0$$

$$\Leftrightarrow 1 - (1-p)^{t-1}\cdot(2-p) \leq 0$$

$$\Leftrightarrow t \leq 1 - \frac{log_2(2-p)}{log_2(1-p)}$$

Damit ist Proposition 7 bewiesen. □

Zur Veranschaulichung sei die Funktion $f(p,n)$ für $p = 0.2$ grafisch illustriert:



1.8 Analyse der Kostenfunktion

In diesem Kapitel wird nun Aussage 3 von Theorem 1 bewiesen. Dazu werden die Intervalle genauer betrachtet, in denen gleich codiert wird.
Ein solches Intervall I wird definiert durch die Eigenschaft: $L(p,n) = L(q,n)$ für alle $n \in \mathbb{N}$ für alle p, $q \in I$.
Durch Proposition 3 wurde gezeigt, daß einige Intervallgrenzen gegeben werden durch die Nullstellen der Funktion $f_t(p) = (1-p)^{t+1} + (1-p)^t - 1$, also durch die Menge $\{p \in (0,1) : \text{es existiert } t \in \mathbb{N}, \text{ so daß } f_t(p) = 0\}$.
Nun sollen weitere Intervallgrenzen angegeben werden. Da zur Berechnung der Kostenfunktion nicht die Codeworte selbst, sondern nur die Codewortlängen benötigt werden, sollen nun die Listen von Codewortlängen optimaler Codes klassifiziert werden. Nach Bemerkung 2 (Seite 14) existiert zu jeder Liste von Codewortlängen eines optimalen Codes ein Präfixcode mit diesen Längen. Zunächst betrachte man alle Listen von Codewortlängen, die durch das Huffman–verfahren konstruiert werden können.
Hat man zum Beispiel die Wahrscheinlichkeitsverteilung $q_1 = q_2 = q_3 = 0.2$ und $q_4 = 0.4$, so wird im ersten Schritt q_1 und q_2 zusammengefaßt. Sei $q^* = q_1 + q_2$. Nun kann das Einsortieren von q^* auf zweierlei Art geschehen. Entweder man erhält (q_3, q^*, q_4) oder (q_3, q_4, q^*). Bezüglich der Werte sind diese beiden Wahrscheinlichkeitsverteilungen identisch. Werden jedoch die beiden Wahrscheinlichkeiten q_1 und q_2 wieder getrennt, so erhält man im ersten Fall die Codewortlängen $(3,3,2,1)$ und im zweiten Fall $(2,2,2,2)$.

Definition :
Eine Liste L^* von Codewortlängen eines optimalen Codes zu einer Wahrscheinlichkeitsverteilung Q heiße **Huffman–darstellbar** oder kurz **H–darstellbar**, wenn man die Positionen des Einsortierens so wählen kann, daß man L^* als Codewortlängen erhält.

Lemma 10:
Jede Liste von Codewortlängen eines optimalen Codes zu einer Wahrscheinlichkeitsverteilung Q ist H–darstellbar.

Beweis :
Es ist bekannt, daß zu jedem optimalen Präfixcode zwei Codeworte maximaler Länge existieren, die sich nur in dem letzten Bit unterscheiden. Daher kann zu jeder Liste von Codewortlängen ein Präfixcode derart konstruiert werden, daß die zu den beiden kleinsten Wahrscheinlichkeiten gehörenden Codeworte sich nur in dem letzten Bit unterscheiden. Ersetzt man nun die beiden kleinsten Wahrscheinlichkeiten durch ihre Summe und ordnet man dieser neuen Wahrscheinlichkeit als Codewort das um ein Bit verkürzte Codewort eines der beiden ursprünglichen Wahrscheinlichkeiten zu, so erhält man nach Theorem 4 (Seite 13) einen optimalen Code für die neu entstandene Wahrscheinlichkeitsverteilung. Das heißt aber, jeder optimale Präfixcode ist rekursiv konstruierbar.

Man nehme an, es existiert eine Liste von Codewortelängen eines optimalen Codes, die nicht H–darstellbar ist. Sei m die minimale Anzahl von Codeworten, bei der ein nicht H–darstellbarer optimaler Code existiert. Dann existiert nach Bemerkung 2 (Seite 14) eine optimaler Präfixcode, der nicht H–darstellbar ist. Nach obigen Ausführungen kann dieser Präfixcode derart konstruiert werden, daß sich die zu den beiden kleinsten Wahrscheinlichkeiten gehörenden Codeworte nur in dem letzten Bit unterscheiden. Faßt man nun die beiden kleinsten Wahrscheinlichkeiten zusammen und ordnet der Summe entsprechend obigen Ausführungen ein Codewort zu, so erhält man einen optimalen Präfixcode mit $m-1$ Codeworten. Da m minimal gewählt wurde, ist die Liste der Codewortlängen dieses Präfixcodes H–darstellbar. Damit ist aber auch die ursprüngliche Liste von Codewortlängen H–darstellbar. Durch diesen Widerspruch erhält man, daß ein solches minimales m nicht existiert. □

Bemerkung 5: Aus diesem Lemma erhält man die folgenden Eigenschaften:

- Wenn durch eine Liste von Codewortlängen die zu einer Wahrscheinlichkeitsverteilung gehörende durchschnittliche Codwortlänge minimiert wird, so ist diese Liste H–darstellbar. Gilt zusätzlich, daß in jedem Schritt die Summe der beiden kleinsten Wahrscheinlichkeiten keine bereits enthaltene Wahrscheinlichkeit darstellt, so ist diese Liste eindeutig.

- Die Intervalle, in denen gleich codiert wird, werden durch Wahrscheinlichkeiten p begrenzt, bei denen $L(p,n)$ für ein n nicht eindeutig ist. Nach Proposition 4 (Seite 30) und Lemma 8 (Seite 33) reicht es, nur solche n zu betrachten, die kleiner sind als $3 * t(p)$.

- Für jede beliebige aber feste Vorgabe von b und n ist $kost(p,b,n)$ stetig in p.

Lemma 11:
Ist für eine Liste $L(p,n)$ die Codewortlänge $l_0 = 1$, so ist $n < 2 \cdot t(p)$.

Beweis :
Damit l_0 eine Länge von 1 hat, darf p_0 im letzten Schritt nicht zusammengefaßt werden. Man hat im letzten Schritt drei Wahrscheinlichkeiten p_0, k_0 und k_1 und es muß gelten: $k_0 \leq k_1 \leq p_0$. Mit $k_0 + k_1 = 1 - p_0$ folgt: $k_1 \geq \frac{1-p_0}{2}$. Daher muß gelten $p_0 > \frac{1-p_0}{2}$. Diese Bedingung gilt genau dann, wenn $n < \frac{-log_2 3}{log_2(1-p)}$.
Nun wird gezeigt, daß $\frac{-log_2 3}{log_2(1-p)} < 2 \cdot \frac{-log_2(2-p)}{log_2(1-p)} \leq 2 \cdot t(p)$.

$\frac{-log_2 3}{log_2(1-p)} < 2 \cdot \frac{-log_2(2-p)}{log_2(1-p)}$ genau dann, wenn

$$p < 2 - \sqrt{3} \approx 0.267949 > \frac{4}{3} - \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}} - \frac{1}{9 \cdot \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}}} \approx 0.245122.$$ □

Nun kann Aussage 3 von Theorem 1 bewiesen werden.

Nach Lemma 4 (Seite 17) gilt die zu beweisende Aussage bereits für alle $p \geq \frac{3-\sqrt{5}}{2}$. In diesem Intervall ist p_0 eine der beiden kleinsten Wahrscheinlichkeiten.

Proposition 8:
Für alle $p \geq \frac{3-\sqrt{5}}{2}$ gilt:

- $l_0 = n$ und $l_i = i$ für alle $1 \leq i \leq n$.
- $kost(p, b, n) = (1-p)^n \cdot (b - \frac{1}{p}) + \frac{1}{p}$.

Für $p < \frac{3-\sqrt{5}}{2}$ werden p_n und p_{n-1} im ersten Schritt zusammengefaßt.

Für alle $\frac{4}{3} - \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}} - \frac{1}{3 \cdot \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}}} \leq p \leq \frac{3-\sqrt{5}}{2}$ gilt nach Proposition 3, daß $t(p) = 2$.

Die Position des Einsortierens von p_0 ist nicht eindeutig, wenn der Wert von p_0 mit einer anderen Wahrscheinlichkeit p_i übereinstimmt. Dies führt zu verschiedenen Codewortlängen, wenn durch das Ändern der Schrittweite ein Schritt existiert, bei dem im ersten Fall p_{i+1} mit p_0 zusammengefaßt wird und im zweiten Fall p_{i+1} mit p_i. Falls $p_n + p_{n-1} \geq p_0$ geschieht dies genau dann, wenn die Srungweite von p_0 von einem ungeraden zu einem geraden Wert wechselt.
Aus Lemma 6 (Seite 27) erhält man, daß dieser Wechsel bei den Nullstellen der Funktion $g_i(p) := (1-p)^i - p$ stattfindet.
$p_n + p_{n-1} \geq p_0$ genau dann, wenn $p \geq 1 - \sqrt{\frac{1}{2}}$.
Damit erhält man als Intervallgrenzen:
$\{p \geq 1 - \sqrt{\frac{1}{2}}$: es existiert ein $i \in \mathbb{N}$, so daß $g_{2\cdot i}(p) = 0\}$.

Ist $p_n + p_{n-1} < p_0$, so kann eine Intervallgrenze durch den Wechsel der Schrittweite von p_0 von einem geraden zu einem ungeraden Wert entstehen.

Ist $p < \frac{3-\sqrt{5}}{2}$ so werden im ersten Schritt p_n und p_{n-1} zusammengefaßt. Weitere Intervallgrenze knnen nun entstehen, wenn die Summe der beiden kleinsten Wahrscheinlichkeiten eines Schrittes gleich p_0 ist. Aus dem ersten Schritt erhlt man: $p_n + p_{n-1} = p_0$. Diese Gleichung wird nur durch die Wahrscheinlichkeit $p = 1 - \sqrt{\frac{1}{2}}$ erfüllt.

Damit erhält man nun als mgliche Intervallgrenzen:

$$\{p \in (0,1)\text{: es existiert } t \in \mathbb{N} \text{ mit } f_f(p) = 0\} \quad \cup$$
$$\{p \geq 1 - \sqrt{\tfrac{1}{2}} \text{ : es existiert } i \in \mathbb{N} \text{ mit } g_{2\cdot i}(p) = 0\} \cup \{1 - \sqrt{\tfrac{1}{2}}\}$$

Das Intervall $\frac{4}{3} - \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}} - \frac{1}{9 \cdot \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}}} \leq p \leq \frac{3-\sqrt{5}}{2}$ wird durch die Nullstelle von $g_4(p)$ geteilt.
Sei p^* die Nullstelle von $g_4(p)$.
Im Intervall $p^* \leq p < \frac{3-\sqrt{5}}{2}$ ist die Schrittweite von p_0 entweder 2 oder 3.
Es gilt $p_{n-3} \leq k_1 := p_n + p_{n-1} < p_{n-4}$. Daher wird p_0 im zweiten Schritt mit p_{n-2} zusammengefaßt. Weiter gilt $p_{n-4} \leq k_2 := p_0 + p_{n-2} < p_{n-5}$. Nun liegt die Situation $(*)$ (Beweis von Proposition 3 Seite 21) vor. Die k_i werden nach Proposition 3 in jedem zweiten Schritt wieder zusammengefaßt. Damit gilt die Gleichung (1.2) von Seite 11 für alle $n \in \mathbb{N}$. Man erhält also

Proposition 9:
Für alle $p^* \leq p < \frac{3-\sqrt{5}}{2}$ gilt:

- Für gerade n ist
  $l_0 = \frac{n}{2}$, $l_{2 \cdot i-1} = l_{2 \cdot i} = i + 1$ für alle $1 \leq i \leq \frac{n}{2}$.

$$kost(p, b, n) = (1-p)^n \cdot \left(b - 2 - \frac{(1-p)^2}{(2-p) \cdot p} \right) + 2 + \frac{(1-p)^2}{(2-p) \cdot p}.$$

- Für ungerade n ist
- $l_0 = l_n = \frac{n+1}{2}$, $l_{2 \cdot i-1} = l_{2 \cdot i} = i + 1$ für alle $1 \leq i \leq \frac{n-1}{2}$.

$$kost(p, b, n) = (1-p)^n \cdot \left(b + \frac{1}{p \cdot (1-p) \cdot (2-p)} \right) + 2 + \frac{(1-p)^2}{(2-p) \cdot p}.$$

Damit gilt Proposition 5 (Seite 31) für alle $n \in \mathbb{N}$ und damit ist Aussage 3 von Theorem 1 für das Intervall $p^* \leq p < \frac{3-\sqrt{5}}{2}$ bewiesen.

Es bleibt das Intervall $\frac{4}{3} - \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}} - \frac{1}{9 \cdot \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}}} \leq p < p^*$ zu analysieren.

In diesem Intervall gilt $p_{n-3} \leq p_0 < p_{n-4}$ und $p_{n-2} \leq k_1 := p_n + p_{n-1} < p_{n-3}$. Daher wird k_1 im zweiten Schritt wieder zusammengefaßt.
Es folgt, daß $p_{n-4} \leq k_2 := k_1 + p_{n-2} < p_{n-5}$, daher wird im dritten Schritt p_0 mit p_{n-3} zusammengefaßt und einsortiert: $p_{n-5} \leq k_3 := p_0 + p_{n-3} < p_{n-6}$. Nun liegt die Situation $(*)$ (Beweis von Proposition 3 Seite 21) vor. Damit werden die k_i nach Proposition 3 in jedem zweiten Schritt wieder zusammengefaßt, und es gilt die Gleichung (1.2) von Seite 11 für alle $n > 1$.
Man erhält also

Proposition 10:
Für alle $\frac{4}{3} - \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}} - \frac{1}{9 \cdot \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}}} \leq p < p^*$ gilt:

- Für gerade n ist
  $l_0 = \frac{n}{2}$, $l_{2 \cdot i-1} = l_{2 \cdot i} = i + 1$ für alle $1 \leq i \leq \frac{n}{2}$.
  $$kost(p, b, n) = (1-p)^n \cdot \left(b - 2 - \frac{(1-p)^2}{(2-p) \cdot p} \right) + 2 + \frac{(1-p)^2}{(2-p) \cdot p}.$$
- $L(p, 3) = (1, 3, 3, 2)$.
- Für ungerade $n > 3$ ist
  $l_0 = \frac{n-1}{2}$, $l_n = l_{n-1} = \frac{n+3}{2}$, $l_{n-2} = \frac{n+1}{2}$,
  $l_{2 \cdot i-1} = l_{2 \cdot i} = i + 1$ für alle $1 \leq i \leq \frac{n-3}{2}$.
  $kost(p, b, n) = (1-p)^n \cdot \left(b + \frac{3-p}{(1-p)^3} - 2 - \frac{2-(1-p)^2}{(1-p) \cdot p \cdot (2-p)} \right) + 2 + \frac{(1-p)^2}{(2-p) \cdot p}$.

Damit gilt Proposition 5 (Seite 31) für alle $n > 1$ und damit ist Aussage 3 von Theorem 1 für das Intervall $\frac{4}{3} - \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}} - \frac{1}{9 \cdot \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}}} \leq p < p^*$ bewiesen.

Damit ist Aussage 3 von Theorem 1 bewiesen. □

1.9 Zusammenfassung und offene Fragen

Dieser Abschnitt ist im wesentlichen schon verständlich, wenn man die Einleitung gelesen hat.

1. Durch Aussage 3 von Theorem 1 wurde bewiesen, daß für alle
   $p \geq \frac{4}{3} - \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}} - \frac{1}{9 \cdot \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}}} \approx 0.245122$ der Einsatz von keinem oder unendlich vielen Experten optimal ist.

   Falls $kost(p, b, 0) = kost(p, b, \infty)$, so ist $kost(p, b, 0) = kost(p, b, i \cdot t(p))$ für alle $i \in \mathbb{N}$. Dabei stellt $t(p)$ die in Proposition 3 bestimmte Periodenlänge von $L(p, n)$ dar.

   Die Hauptvermutung ist, daß dieses Ergebnis für alle $0 < p < 1$ gilt.

   Genauer:

   Sei $c_0(p) = 1 + \lfloor log_2 t(p) \rfloor + \frac{(1-p)^{2^{\lfloor log_2 t(p) \rfloor + 1} - t(p)}}{1-(1-p)^{t(p)}}$.

Ich vermute, daß $kost(p, b, n)$ für $n = 0$ minimal ist, wenn $b \leq c_0$ und sonst für $n = \infty$.

Von dieser Vermutung sind die in Theorem 1 aufgeführten Aussagen bewiesen.

In den Fällen, in denen der Einsatz von unendlich vielen Experten optimal ist, ist die Folge $kost(p, b, n)$ in n nicht notwendig monoton fallend.

2. Von den Intervallen in denen gleich codiert wird, d.h. Intervalle I, in denen gilt $L(p, n) = L(q, n)$ für alle $n \in \mathbb{N}$ für alle $q, p \in I$, werden in Kapitel 8 Seite 41 einige Intervallgrenzen angegeben. Diese erhält man aus den Nullstellen der Funktionen $f_t(p) = (1-p)^{t+1} + (1-p)^t - 1$ und

   $g_i(p) := (1-p)^i - p$ und dem Wert $p = 1 - \sqrt{\frac{1}{2}}$.

   Welche Nullstellen bilden eine Intervallgrenze ?

3. Angeregt durch die Aussage von Proposition 3, habe ich folgende allgemeine abstrakte Vermutung:

   Seien $r \in \mathbb{R}$ mit $r > 1$, sei $a_i = r^i$ für alle $i \in \mathbb{N}$ und seien $b_1, \ldots, b_t$ beliebige positive reelle Zahlen ($t < \infty$).

   Zunächst sortiert man die (unendlich vielen) a_i und die (endlich vielen) b_j aufsteigend zu einer Liste $(c_k)_{k=1}^{\infty}$ mit $c_1 \leq c_2 \leq \ldots$.

   Auf der Folge (c_k) führe man folgende Operation durch:

   Addiere die beiden kleinsten Elemente und sortiere das Ergebnis der Größe entsprechend in die Restliste ein. Das Zusammenfassen soll also nach der gleichen Vorschrift wie der des Huffman–Verfahrens erfolgen. Man iteriere diese Operation.

   Vermutung: Es gibt ein $s \in \mathbb{N}$ und ein $t^* \in \mathbb{N}$, sodaß nach s Startschritten die Summe der beiden kleinsten Werte immer genau t^* Stellen vorrückt.

   Der Fall $r \geq 2$ ist einfach. Möglicherweise läßt sich diese Vermutung für $1 < r < 2$ mit einem zu dem Beweis von Proposition 3 (Kapitel 5) ähnlichen Verfahren beweisen.

4. Sei $s(p)$ die in Lemma 6 (Seite 27) bestimmte Schrittweite von p_0. Nach Proposition 4 (Seite 30) und Lemma 8 (Seite 33) gilt:

$$kost(p, b, n + t(p)) = (1-p)^{t(p)} \cdot kost(p, b, n) + (1 - (1-p)^{t(p)}) \cdot c_0(p)$$

für alle $n \geq s(p) + 3 \cdot t(p)$. Durch Lemma 9 (Seite 34) erhält man:

$$kost(p, 0, t(p)) = (1 - (1 - p)^{t(p)}) \cdot c_0(p).$$

Daher gilt:

$$kost(p, b, n + t(p)) = (1 - p)^{t(p)} \cdot kost(p, b, n) + kost((p, 0, t(p)). \tag{1.3}$$

Mit Hilfe dieses Resultats erhält man eine weitere Konstruktion der Listen $L(p, n + t(p))$ für alle $n \geq s(p) + 3 \cdot t(p)$:

- $l_i(p, n + t(p)) = l_i(p, t(p)$ für alle $1 \leq i \leq t(p)$,
- $l_i(p, n + t(p)) = 1 + l_{i-t(p)}(p, n)$ für alle $t(p) + 1 \leq i \leq n + t(p)$,
- $l_0(p, n + t(p)) = 1 + l_0(p, n)$.

In Kapitel 8 wurde durch die Propositionen 8,9 und 10 gezeigt, daß für alle $p \geq \frac{4}{3} - \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}} - \frac{1}{9 \cdot \sqrt[3]{\frac{25}{54} + \sqrt{\frac{23}{108}}}}$ diese Konstruktion für alle $n > 1$ das optimale Ergebnis liefert.

Nach Proposition 4 (Seite 30) und Lemma 8 (Seite 33) gilt die Gleichung 1.3 für alle $n \geq s(p) + 3 \cdot t(p)$.

Eine weitere Vermutung ist, daß die Gleichheit in 1.3 für jedes ganzzahlige positive Vielfache von $t(p)$ gilt, d.h.

$$kost(p, b, i \cdot t(p)) = (1 - p)^{t(p)} \cdot kost(p, b, (i - 1) \cdot t(p)) + kost(p, 0, t(p))$$

für alle $i \in \mathbb{N}$. Mit Hilfe dieser Rekursion erhält man:

$$kost(p, b, i \cdot t(p)) = (1 - p)^{i \cdot t(p)} \cdot (b - c_0(p)) + c_0(p).$$

Für welche weiteren n gilt Gleichheit in 1.3 ?

1.10 Literatur

[Alt 91] I. Althöfer, Data compression using an intelligent generator: the storage of chess games as an example, Artificial Intelligence 52 (1991), 109–113.

[Ash 65] R.B. Ash, Information Theory, Interscience Tracts in Pure and Applied Mathematics 19, Interscience, New York, 1965.

[AW 79] R. Ahlswede and I. Wegener, Suchprobleme, Teubner, Stuttgart, 1979.

[CT 91] T.M. Cover and J.A. Thomas, Elements of Information Theory, Wiley, New York, 1991.

[Eli 55] P. Elias, Predictive Coding, IRE Transactions– Information Theory (1955), 16–33.

[GV 75] R.G. Gallager and D.C. van Voorhis, Optimal Source Codes for Geometrically Distributed Integer Alphabets, IEEE Transactions on Information Theory (1975), 228–230.

[LH 87] D.A. Lelewer and D.S. Hirschberg, Data compression, ACM Computing Surveys 19 (1987), 261–296.

[Sha 51] C.E. Shannon, Prediction and entropy of printed English, Bell Systems Technical Journal 30 (1951), 50–64.

2 k–Unabhängigkeit von Zufallsvariablen

2.1 Einleitung

In diesem Teil der Arbeit werden Systeme teilweise unabhängiger Zufallsvariablen betrachtet.
Kennt man von einem komplizierten System mit vielen Koordinaten nur etwas über das Zusammenspiel von jeweils kleinen Koordinatengruppen, wie kann sich dann das Gesamtsystem verhalten (z.B. wenn sich Unabhängigkeiten nicht fortsetzen) ?

Um dieses Problem zu verdeutlichen, sei ein Beispiel aus der Physik angeführt.

Gegeben seien drei unabhängige Elementarmagneten Y_1, Y_2, Y_3 mit Spins aus $\{-1, +1\}$ und $Prob\{Y_i = -1\} = Prob\{Y_i = +1\} = \frac{1}{2}$. Weiter seien drei Zufallsvariablen definiert durch

$$X_{12} = Y_1 \cdot Y_2 \qquad X_{23} = Y_2 \cdot Y_3 \qquad X_{31} = Y_3 \cdot Y_1.$$

Die X_i ($i \in \{12, 23, 31\}$) sind Konjunktionen von jeweils zwei Spins.
Es gilt:

- $Prob\{X_i = -1\} = Prob\{X_i = +1\} = \frac{1}{2}$ für alle $i \in \{12, 23, 31\}$
- $Prob\{X_i = x_i, X_j = x_j\} = Prob\{X_i = x_i\} \cdot Prob\{X_j = x_j\}$ für alle $\{i, j\} \subset \{12, 23, 31\}$.
- $Prob\{X_{12} = x_{12}, X_{23} = x_{23}, X_{31} = x_{31}\} \neq \prod_i Prob\{X_i = x_i\}$.

Wie man an diesem Beispiel feststellt, verhalten sich jeweils zwei der Zufallsvariablen X_i unabhängig, obwohl nicht alle drei Zufallsvariablen unabhängig voneinander sind. Daher bezeichne ich die drei Zufallsvariablen X_{12}, X_{23}, X_{31} als 2–unabhängig oder paarweise unabhängig.

Eine Anwendung k–unabhängiger Zufallsvariablen wird in [AS 92] vorgestellt. In diesem Buch werden unter anderem effiziente Zufalls–Algorithmen vorgestellt, die in deterministische Algorithmen umgewandelt werden. Diese Umwandlung (Derandomisierung) wird durch ein Verfahren vorgenommen, das
Method of Conditional Probabilities genannt wird. Dabei werden k–unabhängige Zufallsvariablen für beliebige, aber feste $k \in \mathbb{N}$, benötigt.
Dieses Verfahren der Method of Conditional Probabilities kann ebenso zur Parallelisierung von Algorithmen und zur Färbung von Hypergraphen benutzt werden (siehe auch [NN 90] und [AGHP 90]).

In dieser Arbeit werden n Zufallsvariablen $X_1, \ldots, X_n$ mit Wertebereichen $\mathcal{X} = \{0,1\}$ und Randwahrscheinlichkeiten $Prob\{X_i = 0\} = Prob\{X_i = 1\} = \frac{1}{2}$ für alle $1 \leq i \leq n$ betrachtet.

Sei $X^n = (X_1, \ldots, X_n)$ und $p_v = Prob\{X^n = v\}$. Durch die Wahrscheinlichkeiten p_v ist eine Wahrscheinlichkeitsverteilung P auf $\{0,1\}^n$ gegeben.
Sei eine Funktion $f : \{0,1\}^n \to \mathbb{R}$ gegeben.
Der Erwartungswert der Funktion f ist

$$E_f(P) = \sum_{v \in \{0,1\}^n} f(v) \cdot p_v.$$

Ziel dieser Arbeit ist es, den Erwartungswert $E_f(P)$ für verschiedene Funktionen f unter der Nebenbedingung der k–Unabhängigkeit der n Zufallsvariablen zu maximieren. Diesen maximalen Erwartungswert werde ich mit $E_f^*(k,n)$ bezeichnen.

2.2 Unabhängigkeit von Zufallsvariablen

Definition :
n diskrete Zufallsvariablen $X_1, X_2, \ldots, X_n$ mit Wertebereichen $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n$ heißen **unabhängig**, falls

$$Prob\{X_i = x_i \text{ für alle } 1 \leq i \leq n\} = \prod_{i=1}^{n} Prob\{X_i = x_i\}$$

für alle $(x_1, \ldots, x_n) \in \mathcal{X}_1 \times \ldots \times \mathcal{X}_n$.

Definition :
n diskrete Zufallsvariablen heißen **k–unabhängig** für ein k mit $1 \leq k \leq n$, falls für jede k–elementige Teilmenge $\{i_1, \ldots, i_k\}$ von $\{1, \ldots, n\}$ und alle $(x_{i_1}, \ldots, x_{i_k}) \in \mathcal{X}_{i_1} \times \ldots \times \mathcal{X}_{i_k}$ gilt

$$Prob\{X_{i_1} = x_{i_1}, \ldots, X_{i_k} = x_{i_k}\} = \prod_{j=1}^{k} Prob\{X_{i_j} = x_{i_j}\}.$$

Notation :
Als 0–unabhängig bezeichne ich den Fall, in dem der Erwartungswert über alle Wahrscheinlichkeitsverteilungen maximiert wird. Da in dieser Arbeit ansonsten nur Wahrscheinlichkeitsverteilungen mit den Randverteilungen der Zufallsvariablen $(\frac{1}{2}, \frac{1}{2})$ betrachtet werden, ist dieser Fall der 0–Unabhängigkeit von dem Fall der 1–Unabhängigkeit verschieden.

Definition :
Sei $v = (v_1, \ldots, v_n)$ ein Vektor aus $\{0,1\}^n$. Dann heißt
$d_H(v) := \sum_{i=1}^{n} v_i$ das **Hamming–Gewicht** des Vektors v.

Definition :
Eine Wahrscheinlichkeitsverteilung P heiße **gewichts–symmetrisch**, wenn für alle $v, w \in \{0,1\}^n$ mit $d_H(v) = d_H(w)$ gilt: $p_v = p_w$.
Entsprechend heißt eine Funktion f gewichts–symmetrisch, wenn $f(v) = f(w)$ für alle $v, w \in \{0,1\}^n$ mit $d_H(v) = d_H(w)$.

Notation :
Bei einer gewichts–symmetrischen Wahrscheinlichkeitsverteilung P hängt die Wahrscheinlichkeit p_v nicht von dem Vektor v selbst, sondern nur von dessen Hamming–Gewicht ab. Daher werde ich bei gewichts–symmetrischen Wahrscheinlichkeitsverteilungen nur das Hamming–Gewicht als Index benutzen und jeden Index nur einmal aufführen, d.h. eine solche Wahrscheinlichkeitsverteilung wird dargestellt durch $Q = (q_0, \ldots, q_n)$ mit $\sum\limits_{i=0}^{n} \binom{n}{i} \cdot q_i = 1$.

Definition :
Eine gewichts–symmetrische Wahrscheinlichkeitsverteilung
$Q = (q_0, \ldots, q_n)$ heiße **spiegel–symmetrisch**, wenn zusätzlich gilt: $q_i = q_{n-i}$ für alle $0 \leq i \leq n$.
Entsprechend heißt eine gewichts–symmetrische Funktion f spiegel–symmetrisch, wenn für alle Vektoren $v, w \in \{0,1\}^n$ mit $d_H(v) = i$ und $d_H(w) = n - i$ gilt: $f(v) = f(w)$. Eine spiegel–symmetrische Funktionen f kann als Abbildung von $\{0, \ldots, n\}$ nach $\mathbb{R}$ mit der Eigenschaft $f(i) = f(n-i)$ aufgefaßt werden.

2.3 Ergebnisse

Sei die Betragsfunktion definiert durch $f(v) = \mid d_H(v) - \frac{n}{2} \mid$ und die Mehrheitsfunktion durch

$$f(v) = \begin{cases} 1 & \text{, falls } d_H(v) > \frac{n}{2} \\ 0 & \text{, falls } d_H(v) < \frac{n}{2} \\ \frac{1}{2} & \text{, falls } d_H(v) = \frac{n}{2} \end{cases}$$

Dann gilt:

Theorem 1:

1. Zu jeder gewichts–symmetrischen Funktion f existiert eine gewichts–symmetrische Wahrscheinlichkeitsverteilung, die E_f maximiert.

2. Zu jeder spiegel–symmetrischen Funktion f existiert eine spiegel–symmetrische Wahrscheinlichkeitsverteilung, die E_f maximiert.

3. Für alle spiegel–symmetrischen Funktion f gilt $E_f^*(2 \cdot l, n) = E_f^*(2 \cdot l + 1, n)$ für alle $n \in \mathbb{N}$ und für alle $0 \leq l \leq \frac{n-1}{2}$.

4. Bei der Betragsfunktion stimmt das Wachstumsverhalten von $E_f^*(2, n)$ bis auf einen konstanten Faktor mit dem Wachstumsverhalten von $E_f^*(n, n)$ überein.

5. Bei der Mehrheitsfunktion gilt für alle $0 \leq k \leq 3$: $\lim_{n \to \infty} E_f^*(k, n) = 1$.

2.4 Betrachtung des Modells

In diesem Kapitel werden die Aussagen 1,2 und 3 von Theorem 1 bewiesen.

Lemma 1:
Ist f eine gewichts–symmetrische Funktion, dann existiert eine gewichts–symmetrische Verteilung, die E_f maximiert.

Beweis :
Sei P eine Wahrscheinlichkeitsverteilung, die E_f maximiert. Man erhält aus P eine gewichts–symmetrische Wahrscheinlichkeitsverteilung Q.
Zu jedem Vektor $v \in \{0,1\}^n$ existieren $\binom{n}{d_H(v)}$ viele Vektoren in $\{0,1\}^n$ mit demselben Hamming–Gewicht $d_H(v)$.
Nun berechne man Q wie folgt aus P:

Man setzt

$$q_{d_H(v)} = \frac{1}{\binom{n}{d_H(v)}} \cdot \sum_{x:\ d_H(x) = d_H(v)} p_x \text{ für alle } v \in \{0,1\}^n.$$

Damit ist Q gewichts–symmetrisch. Da f nach Voraussetzung gewichts–symmetrisch ist, gilt $E_f(P) = E_f(Q)$.

Es bleibt zu zeigen, daß Q k–unabhängig ist.
Da die X_i Randverteilung $(\frac{1}{2}, \frac{1}{2})$ haben, gilt für jede Permutation $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$, daß die Wahrscheinlichkeitsverteilung P' gegeben durch $p'_v = p_{\pi(v)}$ k–unabhängig und daß jede Konvexkombination k–unabhängiger Wahrscheinlichkeitsverteilungen k–unabhängig ist. Man erhält nun eine gewichts–symmetrische Wahrscheinlichkeitsverteilung P^* aus P durch $p_v^* = \frac{1}{n!} \cdot \sum_\pi p_{\pi(v)}$. Da nun $q_{d_H(v)} = p_v^*$ für alle $v \in \{0,1\}^n$ ist Lemma 1 bewiesen. $\square$

Bemerkung 1:
Wie aus dem Beweis von Lemma 1 ersichtlich, läßt sich die k–Unabhängigkeit bei in den Koordinaten symmetrischen Wahrscheinlichkeitsverteilungen durch folgendes Gleichungssystem darstellen:

$$\sum_{i=0}^{n-k} \binom{n-k}{i} \cdot q_{i+j} = \frac{1}{2^k} \text{ für alle } 0 \leq j \leq k. \tag{2.1}$$

Bei der Mehrheitsfunktion und der Betragsfunktion wird daher der Nachweis der k–Unabhängigkeit der jeweiligen Wahrscheinlichkeitsverteilungen mittels dieses Gleichungssystems erbracht.

Lemma 2:
Sei $z_s = 2^{-k}$ die s' te Zeile des linearen Gleichungssystems 2.1. Dann gilt: $\sum\limits_{s=0}^{k} \binom{k}{s} z_s = 1$. Dieses Ergebnis beinhaltet, daß nur Wahrscheinlichkeitsverteilungen das Gleichungssystem 2.1 lösen können.

Lemma 3:
Ist f spiegel–symmetrisch, so existiert eine spiegel–symmetrische Wahrscheinlichkeitsverteilung, die $E_f(Q)$ maximiert.

Lemma 4:
Sei $V(k,n) = \{q = (q_0, \ldots, q_n) : q \text{ erfüllt die Gleichung 2.1}\}$.
Dann gilt: $V(k+1,n) \subset V(k,n)$ für alle $0 \leq k < n$.

Beweis :
Man rechnet direkt nach, daß aus 2.1 für $k+1$ auch 2.1 für k folgt.

Bemerkung 2:
Die 0–Unabhängigkeit einer Wahrscheinlichkeitsverteilung Q bedeutet nach Gleichungssystem 2.1 einfach nur, daß Q eine beliebige Wahrscheinlichkeitsverteilung ist. Jede Wahrscheinlichkeitsverteilung, die Randverteilung
$Prob\{X_1 = 0\} = Prob\{X_i = 1\} = \frac{1}{2}$ hat, ist nach Gleichungssystem 2.1 schon 1–unabhängig.

Bemerkung 3:
Das Hinzufügen einer Gleichung zu einem linearen Gleichungssystem schränkt den Lösungsraum des Gleichungssystems um maximal eine Dimension ein oder macht ihn gleich ganz leer. Da die Lösung des Systems bei n–Unabhängigkeit aus einem einzelnen Punkt, nämlich der Gleichverteilung, besteht (also Dimension 0 hat) und der Lösungsraum bei 0–Unabhängigkeit n–dimensional ist, gilt:

- Die Gleichungen des linearen Gleichungssystems für die k–Unabhängigkeit sind linear unabhängig und

- bei jedem Schritt von k nach $k+1$ wird die Dimension des Lösungsraums um genau 1 reduziert.

Bemerkung 4:
Aus der Linearen Optimierung ist bekannt, daß zu jeder linearen Funktion E_f eine k–unabhängige Wahrscheinlichkeitsverteilung Q existiert, die E_f maximiert, aber nur $k+1$ von 0 verschiedene Einträge hat.

Proposition 1:
Für alle spiegel–symmetrischen Funktionen f gilt $E_f^*(2\cdot l, n) = E_f^*(2\cdot l+1, n)$ für alle $n \in \mathbb{N}$ und für alle $0 \leq l \leq \frac{n-1}{2}$.

Beweis : Sei Q eine Wahrscheinlichkeitsverteilung, die $E_f(2\cdot l, n)$ maximiert. Gezeigt wird nun, daß man durch Symmetrisierung der $2\cdot l$– unabhängigen gewichts–symmetrischen Wahrscheinlichkeitsverteilung eine $2\cdot l+1$ – unabhängige spiegel–symmetrische erhält. Da Q nach Voraussetzung $2\cdot l$– unabhängig ist, gilt:

$$\sum_{i=0}^{n-2\cdot l} \binom{n-2\cdot l}{i} q_{i+j} = 2^{-2\cdot l} \qquad \text{für alle } \ 0 \leq j \leq 2\cdot l\,.$$

Sei P gegeben durch $p_i = \frac{1}{2}q_i + \frac{1}{2}q_{n-i}$ für alle $0 \leq i \leq n$. Dann gilt:

$$\begin{aligned}
E_f(P) \quad &= \sum_{i=0}^{n} p_i \cdot \tbinom{n}{i} \cdot f(i) \\
&= \sum_{i=0}^{n} (\tfrac{1}{2}q_i + \tfrac{1}{2}q_{n-i}) \cdot \tbinom{n}{i} \cdot f(i) \\
&= \tfrac{1}{2} \cdot E_f^*(2\cdot l, n) + \tfrac{1}{2} \cdot \sum_{i=0}^{n} q_{n-i} \cdot \tbinom{n}{n-i} \cdot f(n-i) \\
&= \tfrac{1}{2} \cdot E_f^*(2\cdot l, n) + \tfrac{1}{2} \cdot \sum_{i=0}^{n} q_i \cdot \tbinom{n}{i} \cdot f(i) \\
&= \tfrac{1}{2} \cdot E_f^*(2\cdot l, n) + \tfrac{1}{2} \cdot E_f^*(2\cdot l, n) \\
&= E_f^*(2\cdot l, n).
\end{aligned}$$

Nach Lemma 4 gilt $E_f^*(2\cdot l+1, n) \leq E_f^*(2\cdot l, n)$.
Es bleibt also zu zeigen, daß P die Bedingung der $2\cdot l+1$– Unabhängigkeit erfüllt, d.h., daß

$$\sum_{i=0}^{n-2\cdot l-1} \binom{n-2\cdot l-1}{i} \cdot p_{i+j} = 2^{-2\cdot l-1} \qquad \text{für alle } 0 \leq j \leq 2\cdot l+1.$$

Da das Gleichungssystem linear ist, ist jede Konvexkombination zweier Lösungen wieder eine Lösung des Gleichungssystems. Damit erfüllt P die Bedingung der $2\cdot l$– Unabhängigkeit. Es gilt also:

$$
\begin{aligned}
2^{-2\cdot l} &= \sum_{i=0}^{n-2\cdot l} \tbinom{n-2\cdot l}{i} \cdot p_{i+j} \\
&= p_j + p_{n-2\cdot l+j} + \sum_{i=1}^{n-2\cdot l-1} \tbinom{n-2\cdot l}{i} \cdot p_{i+j} \\
&= p_j + p_{n-2\cdot l+j} + \sum_{i=1}^{n-2\cdot l-1} \left(\tbinom{n-2\cdot l-1}{i-1} + \tbinom{n-2\cdot l-1}{i}\right) p_{i+j} \\
&= p_j + \sum_{i=1}^{n-2\cdot l-1} \tbinom{n-2\cdot l-1}{i} p_{i+j} + p_{n-2\cdot l+j} + \sum_{i=1}^{n-2\cdot l-1} \tbinom{n-2\cdot l-1}{i-1} p_{i+j} \\
&= \sum_{i=0}^{n-2\cdot l-1} \tbinom{n-2\cdot l-1}{i} p_{i+j} + \sum_{i=0}^{n-2\cdot l-1} \tbinom{n-2\cdot l-1}{i} p_{i+(j+1)}.
\end{aligned}
$$

Betrachtet man nun den Fall, daß $j = l$ ist, so gilt:

$$p_{n-2\cdot l-1+j+1} = p_{n-l} = p_l = p_j.$$

Damit entspricht für den Fall $j = l$ die zweite Summe der ersten in umgekehrter Reihenfolge der Summanden. Daher gilt, daß die beiden Summen gleich sein müssen. Daher ist nun

$$2^{-2\cdot l} = 2 \cdot \sum_{i=0}^{n-2\cdot l-1} \binom{n-2\cdot l-1}{i} p_{i+l}$$

also

$$\sum_{i=0}^{n-2\cdot l-1} \binom{n-2\cdot l-1}{i} p_{i+l} = 2^{-2\cdot l-1}.$$

Da nun, wie in Lemma 4 bewiesen, die Summe zweier aufeinanderfolgender Zeilen des Gleichungssystems der $(k+1)$–Unabhängigkeit eine Zeile des Gleichungssystems der k–Unabhängigkeit ergibt, gilt:

$$\sum_{i=0}^{n-2\cdot l-1} \binom{n-2\cdot l-1}{i} p_{i+j} = 2^{-2\cdot l-1} \quad \text{für alle } 0 \leq j \leq k+1.$$

Damit stellt also P eine Lösung des Gleichungssystem der $(k+1)$–Unabhängigkeit dar, die E_f maximiert. $\square$

2.5 Die Betragsfunktion

Für die Betragsfunktion ist der Erwartungswert

$$E_f(Q) = \sum_{i=0}^{n} \binom{n}{i} \cdot q_i \cdot \mid i - \frac{n}{2} \mid$$

durch geeignete Wahl der Wahrscheinlichkeitsverteilung zu maximieren.

Bemerkung 5:
Die Funktion $f(i) = \mid i - \frac{n}{2} \mid$ ist spiegel–symmetrisch.
Daher gilt nach Proposition 1: $E_f^*(2 \cdot l, n) = E_f^*(2 \cdot l + 1, n)$ für alle $n \in \mathbb{N}$ und für alle $0 \leq l \leq \frac{n-1}{2}$.

Bemerkung 6:
Es gilt $E_f^*(0, n) = \frac{n}{2}$.

Beweis :
Setze $q_0 = q_n = \frac{1}{2}$ und $q_i = 0$ für alle $1 \leq i \leq n-1$. □

Lemma 5:
Für alle geraden $n \in \mathbb{N}$ gilt:

$$E_f^*(n, n) = 2^{-n} \cdot \tfrac{n}{2} \cdot \tbinom{n}{\frac{n}{2}} = \Omega(\sqrt{n}).$$

und für alle ungeraden $n \in \mathbb{N}$ gilt:

$$E_f^*(n, n) = 2^{-n} \cdot n \cdot \tbinom{n-1}{\frac{n-1}{2}} = \Omega(\sqrt{n}).$$

Beweis :
Für $k = n$ erfüllt nur $p_i = 2^{-n}$ für alle $0 \leq i \leq n$ die Bedingung der n–Unabhängigkeit.

Daher gilt: $\qquad E_f(n, n) = 2^{-n} \cdot \sum_{i=0}^{n} \binom{n}{i} \cdot \mid i - \frac{n}{2} \mid.$

Mit Hilfe der Stirlingformel erhält man, daß $E_f^*(n, n) \approx \sqrt{\frac{n}{2 \cdot \pi}}$ ist. □

Proposition 2:
Bei der Betragsfunktion gilt für die 2–Unabhängigkeit $\quad E_f^*(2, n) \leq \frac{\sqrt{n}}{2}$.

Beweis :

$$E_f(Q) = E \mid \left(\sum_{i=1}^{n} X_i \right) - \frac{n}{2} \mid .$$

Mit Hilfe der Cauchy– Schwarzschen– Ungleichung erhält man

$$E_f(Q) \leq \sqrt{Var(\sum_{i=1}^{n} X_i)}.$$

Für paarweise unabhängige Zufallsvariablen gilt:

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i).$$

Daher gilt:

$$E_f^*(2,n) \leq \frac{\sqrt{n}}{2}.$$

Das Wachstumsverhalten von $E_f^*(2,n)$ stimmt also bis auf einen konstanten Faktor mit dem Wachstumsverhalten von $E_f^*(n,n)$ überein. $\square$

Beobachtung und Vermutung:
Für $1 \leq n \leq 50$ wurden die exakten Werte mit Hilfe des Simplexalgorithmus durch Mathematica bestimmt.
Diese Werte führen zu der folgenden exakten Berechnung für $E_f^*(2,n)$.
Ist $n = a^2$ für ein $a \in \mathbb{N}$, so ist $E_f^*(2,n) = \frac{a}{2}$.
Ist $n = a^2 + 1$ für ein $a \in \mathbb{N}$, so ist $E_f^*(2,n) = \frac{a}{2}$.
In allen anderen Fällen sei $d = n - \lfloor\sqrt{n}\rfloor^2$ und $b = \lfloor\sqrt{n}\rfloor + 1$.
Falls nun $d - b - 1 \leq 0$, dann setze $c =\mid d - b - 1 \mid \cdot(b-1)$, sonst setze $c =\mid d - b - 1 \mid \cdot b$.
Nun gilt:

$$E_f^*(2,n) = \frac{1 + 2\cdot(b-1) + c}{4 + 2\cdot \mid d - b - 1 \mid}.$$

Diese Formel erhlt man, wenn man die Werte von $E_f^*(2,n)$ wie folgt in ein Diagramm einträgt.

n	a											
3	1						$\frac{3}{4}$					
6...8	2					$\frac{7}{6}$	$\frac{5}{4}$	$\frac{8}{6}$				
11...15	3				$\frac{13}{8}$	$\frac{10}{6}$	$\frac{7}{4}$	$\frac{11}{6}$	$\frac{15}{8}$			
18...24	4			$\frac{21}{10}$	$\frac{17}{8}$	$\frac{13}{6}$	$\frac{9}{4}$	$\frac{14}{6}$	$\frac{19}{8}$	$\frac{24}{10}$		
27...35	5		$\frac{31}{12}$	$\frac{26}{10}$	$\frac{21}{8}$	$\frac{16}{6}$	$\frac{11}{4}$	$\frac{17}{6}$	$\frac{23}{8}$	$\frac{29}{10}$	$\frac{35}{12}$	
38...48	6	$\frac{43}{14}$	$\frac{37}{12}$	$\frac{31}{10}$	$\frac{25}{8}$	$\frac{19}{6}$	$\frac{13}{4}$	$\frac{20}{6}$	$\frac{27}{8}$	$\frac{34}{10}$	$\frac{41}{12}$	$\frac{48}{14}$

Folgerung:
$E_f^*(2,n)$ ist monoton wachsend in n.

2.6 Die Mehrheitsfunktion

In dem vorhergehenden Kapitel ist gezeigt worden, daß für die dort betrachtete Funktion das Wachstumsverhalten von $E_f^*(2,n)$ größenordnungsmäßig mit dem Wachstumsverhalten von $E_f^*(n,n)$ übereinstimmt.

Da $E_f^*(k,n)$ in k monoton fallend ist, existiert ein minimales k, so daß

$$E_f^*(0,n) - E_f^*(k,n) \geq E_f^*(k,n) - E_f^*(n,n).$$

Sei zu gegebener Funktion f und gegebenem $n \in \mathbb{N}$ dieses k bezeichnet durch $k_f(n)$. Im vorhergehenden Kapitel ist somit $k_f(n) = 2$ für alle $n \in \mathbb{N}$.

In diesem Kapitel soll nun eine Funktion f betrachtet werden, für die $k_f(n)$ vermutlich nicht konstant in n ist.

Für die Mehrheitsfunktion ist der Erwartungswert für alle ungeraden n und alle $Q = (q_0, \ldots, q_n)$

$$E_f(Q) = \sum_{i=\frac{n+1}{2}}^{n} q_i \cdot \binom{n}{i}$$

und für alle geraden n

$$E_f(Q) = \tfrac{1}{2} \cdot q_{\frac{n}{2}} \cdot \binom{n}{\frac{n}{2}} + \sum_{i=\frac{n}{2}+1}^{n} q_i \cdot \binom{n}{i}.$$

Lemma 6:
$E_f^*(0,n) = 1$.

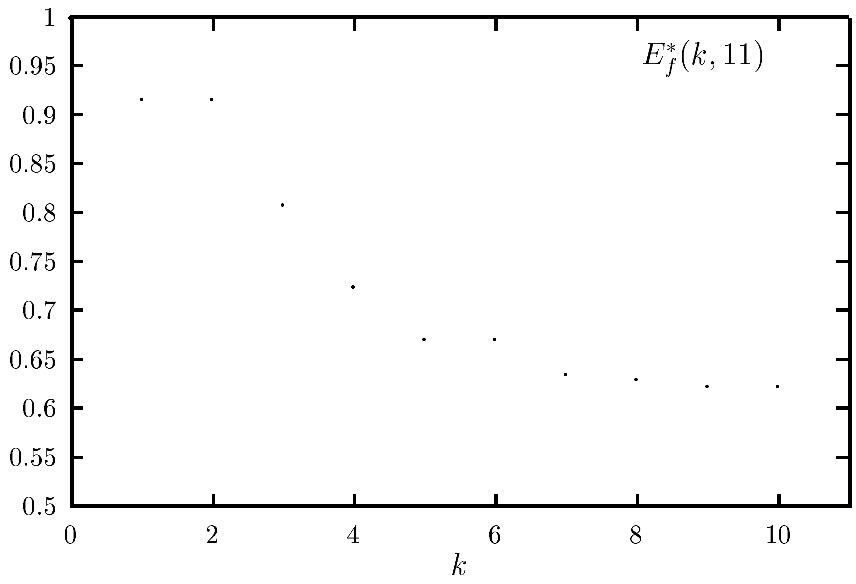
Beweis :
Sei $q_n = 1$ und $q_i = 0$ für alle $0 \leq i < n$.
Dann gilt $E_f(Q) = 1$.

Da $\sum_{i=0}^{n} q_i \cdot \binom{n}{i} = 1$, gilt $E_f^*(0,n) \leq 1$ für alle Wahrscheinlichkeitsverteilungen Q.□

Lemma 7:
$E_f^*(n,n) = \frac{1}{2}$.

Beweis :
Für $k = n$ erfüllt nur die Gleichverteilung auf $\mathcal{X}^n$ das Gleichungssystem 2.1. □

Der Verlauf der Funktion $E_f^*(k,n)$ sei am Beispiel $n = 11$ grafisch dargestellt.



Für $n = 11$ ist $k_f(11) = 4$.

In den nun folgenden drei Teilabschnitten werden die Fälle $k = 1$, $k = 2$ und $k = 3$ behandelt. In jedem dieser Teilabschnitte werden Wahrscheinlichkeitsverteilungen konstruiert, mit deren Hilfe gezeigt wird, daß $\lim\limits_{n\to\infty} E_f^*(k, n) = 1$ für alle $0 \leq k \leq 3$.

2.6.1 Der Fall $k = 1$

Proposition 3:
Für alle ungeraden $n \in \mathbb{N}$ gilt: $E_f^*(1, n) \geq \frac{n}{n+1}$.

Beweis : Sei Q gegeben durch

$q_0 = \frac{1}{n+1}$, $q_{\frac{n+1}{2}} = \frac{1}{2\cdot\binom{n-1}{\frac{n-1}{2}}}$ und $q_i = 0$ für alle $i \notin \{0, \frac{n+1}{2}\}$.

Es ist zu zeigen, daß Q 1–unabhängig ist und daß $E_f(1, n)$ für Q den Wert $\frac{n}{n+1}$ annimmt.

Zur 1–Unabhängigkeit ist zu zeigen:

$$\sum_{i=0}^{n-1} \binom{n-1}{i} q_{i+j} = \tfrac{1}{2} \text{ für alle } 0 \leq j \leq 1.$$

Für die einzelnen Zeilen ergibt sich:

Zu $j = 0$: $\quad q_0 + \binom{n-1}{\frac{n+1}{2}} \cdot q_{\frac{n+1}{2}} = \frac{1}{n+1} + \binom{n-1}{\frac{n+1}{2}} \cdot \frac{1}{2\cdot\binom{n-1}{\frac{n-1}{2}}} = \frac{1}{2}.$

Zu $j = 1$: $\quad \binom{n-1}{\frac{n-1}{2}} \cdot q_{\frac{n+1}{2}} = \binom{n-1}{\frac{n-1}{2}} \cdot \frac{1}{2\cdot\binom{n-1}{\frac{n-1}{2}}} = \frac{1}{2}.$

Also folgt: $\quad E_f(Q) = \binom{n}{\frac{n+1}{2}} \cdot q_{\frac{n+1}{2}} = \binom{n}{\frac{n+1}{2}} \cdot \frac{1}{2\cdot\binom{n-1}{\frac{n-1}{2}}} = \frac{n}{n+1}.$ $\square$

Proposition 4:
Für alle geraden $n \in \mathbb{N}$ gilt: $E_f^*(1,n) \geq \frac{n}{n+2}$.

Beweis : Sei Q gegeben durch

$q_0 = \frac{2}{n+2}$, $q_{\frac{n+2}{2}} = \frac{1}{2\cdot\binom{n-1}{\frac{n}{2}}}$ und $q_i = 0$ für alle $i \notin \{0, \frac{n}{2}+1\}$.

Es ist zu zeigen, daß Q 1–unabhängig ist und $E_f(Q)$ für Q den Wert $\frac{n}{n+1}$ annimmt.

Zur 1–Unabhängigkeit ist zu zeigen:

$$\sum_{i=0}^{n-1} \binom{n-1}{i} q_{i+j} = \tfrac{1}{2} \text{ für alle } 0 \leq j \leq 1.$$

Für die einzelnen Zeilen ergibt sich:

Zu $j = 0$: $\quad q_0 + \binom{n-1}{\frac{n+2}{2}} \cdot q_{\frac{n+2}{2}} = \frac{1}{n+2} + \binom{n-1}{\frac{n+2}{2}} \cdot \frac{1}{2\cdot\binom{n-1}{\frac{n}{2}}} = \frac{1}{2}$.

Zu $j = 1$: $\quad \binom{n-1}{\frac{n}{2}} \cdot q_{\frac{n+2}{2}} = \binom{n-1}{\frac{n}{2}} \cdot \frac{1}{2\cdot\binom{n-1}{\frac{n}{2}}} = \frac{1}{2}$.

Daraus folgt: $\quad E_f(Q) = \binom{n}{\frac{n+2}{2}} \cdot q_{\frac{n+2}{2}} = \binom{n}{\frac{n+2}{2}} \cdot \frac{1}{2\cdot\binom{n-1}{\frac{n}{2}}} = \frac{n}{n+2}$. $\square$

Proposition 5:
Es gilt: $E_f^*(1,n) = \frac{n}{n+1}$ für alle ungeraden $n \in \mathbb{N}$ und $E_f^*(1,n) = \frac{n}{n+2}$ für alle geraden $n \in \mathbb{N}$.

Beweis :
Wie bereits erwähnt, existiert eine Wahrscheinlichkeitsverteilung Q mit maximal 2 positiven Einträgen, die $E_f(Q)$ maximiert.
Annahme, Q sei nur an einer Koordinate q_a positiv. Da für $k = 1$ gilt:

$$\sum_{i=0}^{n-1} \binom{n-1}{i} \cdot q_{i+j} = \tfrac{1}{2} \quad \text{für alle } 0 \leq j \leq 1,$$

ist $0 < a < n$. Für ein solches a muß gelten:

$$\binom{n-1}{a} \cdot q_a = \tfrac{1}{2} \quad \text{und} \quad \binom{n-1}{a-1} \cdot q_a = \tfrac{1}{2}.$$

Dieses Gleichungssystem kann nur für gerade n durch $a = \frac{n}{2}$ gelöst werden. In diesem Fall gilt:

$$q_{\frac{n}{2}} = \frac{1}{2\cdot\binom{n-1}{\frac{n}{2}}} \quad \text{und} \quad E_f(Q) = \tfrac{1}{2} \cdot \binom{n}{\frac{n}{2}} \cdot q_{\frac{n}{2}} = \tfrac{1}{2}.$$

Für alle $n > 2$ ist $\frac{1}{2} < \frac{n}{n+2}$. Daher kann $E_f(Q)$ für $n > 2$ nicht durch ein Q mit nur einer positiven Koordinate maximiert werden.
Damit existiert für alle $n \in \mathbb{N}$ eine Wahrscheinlichkeitsverteilung Q, die $E_f(Q)$ maximiert und genau zwei positive Wahrscheinlichkeiten hat.
Seien diese beiden Wahrscheinlichkeiten gegeben durch q_a und q_b mit $a < b$. Wir gehen die verschiedenen Möglichkeiten nacheinander durch.

1. Sei $a = 0$ und $b = n$. Dann ist $q_0 = q_n = \frac{1}{2}$. Daher ist $E_f(Q) = \frac{1}{2}$.

2. Sei $a = 0$ und $b < n$. Dann ist $q_b = \frac{1}{2\cdot\binom{n-1}{b-1}}$ und $q_0 = \frac{2\cdot b-n}{2\cdot b}$.
   Aus $q_0 > 0$ folgt: $b > \frac{n}{2}$. Damit ist $E_f(Q) = \binom{n}{b} \cdot q_b = \frac{n}{2\cdot b}$.
   $E_f(Q)$ wird maximal, wenn b minimal wird. Für gerade n gilt also: $b = \frac{n}{2}+1$ und für ungerade n: $b = \frac{n+1}{2}$. Diese Wahrscheinlichkeitsverteilungen wurden in Proposition 3 und 4 analysiert.

3. Sei $a \geq 0$ und $b = n$. Dann ist $q_a = \frac{1}{2\cdot\binom{n-1}{a}}$ und $q_n = \frac{n-2\cdot a}{2\cdot(n-a)}$.
   Aus $q_n > 0$ folgt: $a < \frac{n}{2}$. Damit ist $E_f(Q) = q_n$.
   Es gilt: $\frac{\partial}{\partial a}E_f(Q) = -\frac{n}{2\cdot(n-a)^2} < 0$. Daher wird $E_f(Q)$ maximal für $a = 0$ mit $E_f(Q) = \frac{1}{2}$.

4. Sei $a > 0$ und $b < n$. Dann gilt $q_a = \frac{n-2\cdot b}{2\cdot(a-b)\cdot\binom{n}{a}}$ und $q_b = \frac{n-2\cdot a}{2\cdot(a-b)\cdot\binom{n}{b}}$.
   Aus $q_a > 0$ folgt: $b > \frac{n}{2}$ und aus $q_b > 0$ folgt: $a < \frac{n}{2}$.
   Daher ist $E_f(Q) = \binom{n}{b} \cdot q_b = \frac{n-2\cdot a}{2\cdot(a-b)}$. Es gilt: $\frac{\partial}{\partial b}E_f(Q) = \frac{2\cdot a-n}{2\cdot(b-a)^2} < 0$, da $a < \frac{n}{2}$. D.h., für gerade n wird $E_f(Q)$ maximal durch $b = \frac{n}{2} + 1$ und für ungerade n durch $b = \frac{n+1}{2}$. Daher gilt für gerade n: $E_f(Q) = \frac{n-2\cdot a}{n+1-2\cdot a}$. Daraus folgt: $\frac{\partial}{\partial a}E_f(Q) = -\frac{2}{(n+1-2\cdot a)^2} < 0$. Daher ist $E_f(Q) < \frac{n}{n+1}$. Entsprechend ergibt sich für ungerade n: $E_f(Q) < \frac{n}{n+2}$.

Damit ist Proposition 5 bewiesen. □

2.6.2 Der Fall $k = 2$

Proposition 6:
Für alle ungeraden $n \in \mathbb{N}$ gilt: $E_f^*(2, n) \geq \frac{n}{n+1}$.

Beweis : Sei Q gegeben durch die maximierende Wahrscheinlichkeitsverteilung bei 1–Unabhängigkeit, also
$q_0 = \frac{1}{n+1}$, $q_{\frac{n+1}{2}} = \frac{n}{(n+1)\cdot\binom{n}{\frac{n-1}{2}}}$ und $q_i = 0$ für alle $i \notin \{0, \frac{n+1}{2}\}$.
Nun zeige ich, daß dieses Q schon 2–unabhängig ist, also, daß:

$$\sum_{i=0}^{n-2} \binom{n-2}{i} q_{i+j} = 2^{-2} \text{ für alle } 0 \leq j \leq 2.$$

Für die einzelnen Zeilen ergibt sich:

Zu $j = 0$: $\quad q_0 + \binom{n-2}{\frac{n+1}{2}} q_{\frac{n+1}{2}} = \frac{1}{n+1} + \binom{n-2}{\frac{n+1}{2}} \cdot \frac{n}{(n+1)\cdot\binom{n}{\frac{n-1}{2}}} = \frac{1}{4}$.

Zu $j = 1$: $\quad \binom{n-2}{\frac{n-1}{2}} q_{\frac{n+1}{2}} = \binom{n-2}{\frac{n-1}{2}} \cdot \frac{n}{(n+1)\cdot\binom{n}{\frac{n-1}{2}}} = \frac{1}{4}$.

Zu $j = 2$: $\quad \binom{n-2}{\frac{n-3}{2}} q_{\frac{n+1}{2}} = \binom{n-2}{\frac{n-3}{2}} \cdot \frac{n}{(n+1)\cdot\binom{n}{\frac{n-1}{2}}} = \frac{1}{4}$.

Damit erfüllt Q die 2–Unabhängigkeit. □

Vermutung :
Für alle ungeraden n ist $E_f^*(2,n) = \frac{n}{n+1}$.

Diese Aussage wurde für alle ungeraden $n \leq 19$ mit Hilfe von Mathematica bestätigt.

Bemerkung 7:
Für $n = 2$ gilt $E_f^*(2,2) = \frac{1}{2}$, da nur die Gleichverteilung 2–unabhängig ist.

Lemma 8:
Für $n = 4$ gilt $E_f^*(2,4) = \frac{7}{12}$.

Beweis :
Damit eine Wahrscheinlichkeitsverteilung 2–unabhängig ist, muß gelten:

$$q_0 + 2 \cdot q_1 + q_2 = \tfrac{1}{4}, \quad q_1 + 2 \cdot q_2 + q_3 = \tfrac{1}{4} \quad \text{und} \quad q_2 + 2 \cdot q_3 + q_4 = \tfrac{1}{4}.$$

Sei $q_4 = \lambda$ und $q_3 = \mu$. Dann ist

$$q_2 = \tfrac{1}{4} - \lambda - 2 \cdot \mu, \quad q_1 = -\tfrac{1}{4} + 3 \cdot \mu + 2 \cdot \lambda \quad \text{und} \quad q_0 = \tfrac{1}{2} - 3 \cdot \lambda - 4 \cdot \mu.$$

Nun gilt also:

$$\begin{aligned} E_f(Q) &= q_0 + 4 \cdot q_1 + \tfrac{6}{2} \cdot q_2 \\ &= \tfrac{1}{2} - 3 \cdot \lambda - 4 \cdot \mu + 4 \cdot (-\tfrac{1}{4} + 3 \cdot \mu + 2 \cdot \lambda) + 3 \cdot (\tfrac{1}{4} - \lambda - 2 \cdot \mu) \\ &= \tfrac{1}{4} + 2 \cdot (\mu + \lambda). \end{aligned}$$

Damit wird $E_f(Q)$ maximal, wenn $\mu + \lambda$ maximal wird.

Da $\mu \geq 0$ und $\lambda \geq 0$, folgt aus $q_0 \geq 0$, daß $\lambda \leq \frac{1}{6} - \frac{4}{3} \cdot \mu$.

Damit gilt: $\mu + \lambda \leq \frac{1}{6} - \frac{4}{3} \cdot \mu + \mu = \frac{1}{6} - \frac{\mu}{3}$.

Da $\mu \geq 0$, gilt also $\mu + \lambda \leq \frac{1}{6}$. Man erhält nun aus $\mu = 0$ und $\lambda = \frac{1}{6}$ die Wahrscheinlichkeitsverteilung $Q = (0, \frac{1}{12}, \frac{1}{12}, 0, \frac{1}{6})$ und es gilt:

$$E_f^*(2,4) = \frac{1}{4} + 2 \cdot \frac{1}{6} = \frac{7}{12}.$$

□

Proposition 7:
Für alle geraden n mit $n \equiv 0 \bmod 4$ gilt: $E_f^*(2,n) \geq \frac{n}{n+4}$.

Beweis :

Für $n = 4$ wurde die Aussage bereits bewiesen. Sei nun $n \equiv 0 \bmod 4$ und größer als 4. Setze

$$q_{\frac{n+2}{2}} = \frac{n+2}{2 \cdot (n+4) \cdot \binom{n-1}{\frac{n}{2}}} \quad \text{und} \quad q_{\frac{1}{4} \cdot n} = \frac{n-4}{4 \cdot (n-1) \cdot (n+4) \cdot \binom{n-2}{\frac{3}{4} \cdot n}}$$

und $q_i = 0$ für alle $i \notin \{\frac{n+2}{2}, \frac{1}{4} \cdot n\}$.

Für $n > 4$ ist $q_{\frac{1}{4} \cdot n} > 0$ und $q_{\frac{n+2}{2}} > 0$.

Zur 2–Unabhängigkeit ist, muß gezeigt werden, daß

$$\sum_{i=0}^{n-2} \binom{n-2}{i} \cdot q_{i+j} = \tfrac{1}{4} \text{ für alle } 0 \leq j \leq 2.$$

Für die einzelnen Zeilen ergibt sich:

Zu $j = 0$:

$$\begin{aligned}\binom{n-2}{\frac{1}{4} \cdot n} \cdot q_{\frac{n}{4}} + \binom{n-2}{\frac{n}{2}+1} \cdot q_{\frac{n}{2}+1} &= \binom{n-2}{\frac{1}{4} \cdot n} \cdot \frac{n-4}{4 \cdot (n-1) \cdot (n+4) \cdot \binom{n-2}{\frac{3}{4} \cdot n}} + \binom{n-2}{\frac{n}{2}+1} \cdot \frac{n+2}{2 \cdot (n+4) \cdot \binom{n-1}{\frac{n}{2}}} \\ &= \frac{3 \cdot (3 \cdot n - 4)}{4 \cdot (n-1) \cdot (n+4)} + \frac{(n-2) \cdot (n-4)}{4 \cdot (n-1) \cdot (n+4)} = \frac{1}{4}.\end{aligned}$$

Zu $j = 1$:

$$\begin{aligned}\binom{n-2}{\frac{1}{4} \cdot n - 1} \cdot q_{\frac{n}{4}} + \binom{n-2}{\frac{n}{2}} \cdot q_{\frac{n}{2}+1} &= \frac{\binom{n-2}{\frac{1}{4} \cdot n - 1} \cdot (n-4)}{4 \cdot (n-1) \cdot (n+4) \cdot \binom{n-2}{\frac{3}{4} \cdot n}} + \frac{\binom{n-2}{\frac{n}{2}} \cdot (n+2)}{2 \cdot (n+4) \cdot \binom{n-1}{\frac{n}{2}}} \\ &= \frac{3 \cdot n}{4 \cdot (n-1) \cdot (n+4)} + \frac{(n-2) \cdot (n+2)}{4 \cdot (n-1) \cdot (n+4)} = \frac{1}{4}.\end{aligned}$$

Zu $j = 2$:

$$\binom{n-2}{\frac{1}{4} \cdot n - 2} \cdot q_{\frac{n}{4}} + \binom{n-2}{\frac{n}{2}-1} \cdot q_{\frac{n}{2}+1} = \frac{\binom{n-2}{\frac{1}{4} \cdot n - 2} \cdot (n-4)}{4 \cdot (n-1) \cdot (n+4) \cdot \binom{n-2}{\frac{3}{4} \cdot n}} + \frac{\binom{n-2}{\frac{n}{2}-1} \cdot (n+2)}{2 \cdot (n+4) \cdot \binom{n-1}{\frac{n}{2}}} = \frac{1}{4}.$$

Damit ist also $(q_0, \ldots, q_n)$ 2–unabhängig. Es gilt:

$E_f(Q) = \binom{n}{\frac{n+2}{2}} \cdot q_{\frac{n+2}{2}} = \frac{n! \cdot (\frac{n}{2}-1)! \cdot (\frac{n}{2})!}{(\frac{n}{2}+1)! \cdot (\frac{n}{2}-1)! \cdot (n-1)!} \cdot \frac{1}{2} \cdot \frac{n+2}{(n+4)} = \frac{n}{n+4}$. □

Vermutung :
Für alle $n > 4$ mit $n \equiv 0 \bmod 4$ ist $E_f^*(2, n) = \frac{n}{n+4}$.

Diese Ergebnis wurde für $n \in \{8, 12, 16\}$ mit Hilfe von Mathematica bestätigt.

Lemma 9:
Für $n = 6$ gilt $E_f^*(2, 6) \geq \frac{5}{8}$.

Beweis :

Setzt $Q = (0, \frac{1}{24}, 0, \frac{1}{80}, \frac{1}{30}, 0, 0)$. □

Bemerkung 8:
Nach Berechnung mit Hilfe von Mathematica ist $E_f^*(2, 6) = \frac{5}{8}$.

Proposition 8:
Für alle geraden $n > 6$ mit $n \equiv 2 \bmod 4$ gilt: $E_f^*(2,n) \geq \frac{n\cdot(n-8)}{(n-2)\cdot(n-2)}$.

Beweis : Sei

$$q_{\frac{n+2}{2}} = \frac{n-8}{2\cdot(n-2)\cdot\binom{n-1}{\frac{n}{2}+1}} \quad \text{und} \quad q_{\frac{n+6}{4}} = \frac{3\cdot(3n-10)}{4\cdot(n-1)\cdot(n-2)\cdot\binom{n-2}{\frac{n+6}{4}}} \quad \text{und} \quad q_0 = \frac{4}{(n-2)\cdot(n-2)}$$

und $q_i = 0$ für alle $i \notin \{0, \frac{n+2}{2}, \frac{n+6}{4}\}$.

Für $n > 6$ mit $n \equiv 2 \bmod 4$ ist es klar, daß $q_i \geq 0$ für alle $0 \leq i \leq n$.

Zur 2–Unabhängigkeit muß gezeigt werden, daß

$$\sum_{i=0}^{n-2} \binom{n-2}{i} \cdot q_{i+j} = \tfrac{1}{4} \text{ für alle } 0 \leq j \leq 2.$$

Für die einzelnen Zeilen ergibt sich:

Zu $j = 0$:

$$\binom{n-2}{\frac{n+6}{4}} \cdot q_{\frac{n+6}{4}} + \binom{n-2}{\frac{n+2}{2}} \cdot q_{\frac{n+2}{2}} = \frac{3\cdot(3n-10)}{4\cdot(n-1)\cdot(n-2)} + \frac{(n-2)!\cdot(\frac{n}{2}+1)!\cdot(\frac{n}{2}-2)!\cdot(n-8)}{(\frac{n}{2}+1)!\cdot(\frac{n}{2}-3)!\cdot(n-1)!\cdot 2\cdot(n-2)}$$

$$= \frac{3\cdot(3n-10)}{4\cdot(n-1)\cdot(n-2)} + \frac{(n-4)\cdot(n-8)}{4\cdot(n-1)\cdot(n-2)} = \frac{1}{4}.$$

Zu $j = 1$:

$$\binom{n-2}{\frac{n+2}{4}} \cdot q_{\frac{n+6}{4}} + \binom{n-2}{\frac{n}{2}} \cdot q_{\frac{n+2}{2}} = \frac{(n-2)!\cdot(\frac{n+6}{4})!\cdot(\frac{3\cdot n-14}{4})!\cdot 3\cdot(3\cdot n-10)}{(\frac{n+2}{4})!\cdot(\frac{3\cdot n-10}{4})!\cdot(n-2)!\cdot 4\cdot(n-1)\cdot(n-2)}$$

$$+\frac{(n-2)!\cdot(\frac{n}{2}+1)!\cdot(\frac{n}{2}-2)!\cdot(n-3)}{(\frac{n}{2})!\cdot(\frac{n}{2}-2)!\cdot(n-1)!\cdot 2\cdot(n-2)}$$

$$= \frac{4\cdot(n+6)\cdot 3\cdot(3\cdot n-10)}{4\cdot(3\cdot n-10)\cdot 4\cdot(n-1)\cdot(n-2)} + \frac{(n+2)\cdot(n-8)}{4\cdot(n-1)\cdot(n-2)} = \frac{1}{4}.$$

Zu $j = 2$:

$$\binom{n-2}{\frac{n-2}{4}} \cdot q_{\frac{n+6}{4}} + \binom{n-2}{\frac{n}{2}-1} \cdot q_{\frac{n+2}{2}} + q_0 = \frac{(n-2)!\cdot(\frac{n+6}{4})!\cdot(\frac{3\cdot n-14}{4})!\cdot 3\cdot(3\cdot n-10)}{(\frac{n-2}{4})!\cdot(\frac{3\cdot n-6}{4})!\cdot(n-2)!\cdot 4\cdot(n-1)\cdot(n-2)}$$

$$+\frac{(n-2)!\cdot(\frac{n}{2}+1)!\cdot(\frac{n}{2}-2)!\cdot(n-8)}{(\frac{n}{2}-1)!\cdot)\frac{n}{2}-1)!\cdot(n-1)!\cdot 2\cdot(n-2)} + \frac{4}{(n-2)\cdot(n-2)}$$

$$= \frac{(n+6)\cdot(n+2)\cdot(3\cdot n-10)\cdot 3}{(3\cdot n-6)\cdot(3\cdot n-10)\cdot 4\cdot(n-1)\cdot(n-2)}$$

$$+\frac{(n+2)\cdot 2\cdot n\cdot(n-8)}{4\cdot(n-2)\cdot(n-1)\cdot 2\cdot(n-2)} + \frac{4}{(n-2)\cdot(n-2)}$$

$$= \frac{(n+6)\cdot(n+2)+(n+2)\cdot n\cdot(n-8)+16\cdot(n-1)}{4\cdot(n-1)\cdot(n-2)\cdot(n-2)} = \frac{1}{4}.$$

Es ergibt sich:

$$E_f(Q) = \binom{n}{\frac{n+2}{2}} \cdot q_{\frac{n+2}{2}} = \frac{n!\cdot(\frac{n}{2}+1)!\cdot(\frac{n}{2}-2)!\cdot(n-8)}{(\frac{n}{2}+1)!\cdot(\frac{n}{2}-1)!\cdot(n-1)!\cdot 2\cdot(n-2)} = \frac{n^2-8\cdot n}{n^2-4\cdot n+4}. \qquad \square$$

Damit ist für den Fall $k = 2$ bewiesen, daß

$$\lim_{n\to\infty} E_f^*(2,n) = 1.$$

Vermutung :
Für alle geraden $n > 6$ mit $n \equiv 2 \bmod 4$ gilt:

$$E_f^*(2,n) = \frac{n\cdot(n-8)}{(n-2)\cdot(n-2)} \approx \frac{n}{n-4}.$$

2.6.3 Der Fall $k = 3$

Für die in diesem Teilabschnitt erzielten Resultate wurde die Maximalität von $E_f^*(3,n)$ für alle $n \leq 19$ mit Hilfe von Mathematica bestätigt. Daher werde ich in den einzelnen Vermutungen die bestätigten Werte nicht nochmals erwähnen.

Proposition 9:
Für alle $n \in \mathbb{N}$ mit $n > 1$ und $n \equiv 1 \bmod 4$ gilt: $E_f^*(3,n) \geq \frac{n}{n+3} + \frac{1}{3\cdot n+1}$.

Beweis : Sei

$$q_{\frac{n-1}{4}} = \frac{3\cdot(3\cdot n-7)}{8\cdot(n-2)\cdot(n+3)\cdot\binom{n-3}{\frac{n-1}{2}}},\; q_{\frac{n+1}{2}} = \frac{n-3}{4\cdot(n+3)\cdot\binom{n-2}{\frac{n+1}{2}}},\; q_n = \frac{1}{3\cdot n+1} \text{ und } q_i = 0 \text{ sonst.}$$

Für $n > 1$ mit $n \equiv 1 \bmod 4$ ist es klar, daß $q_i \geq 0$ für alle $0 \leq i \leq n$.

Damit nun $(q_0,\ldots,q_n)$ 3–unabhängig ist, muß gezeigt werden, daß

$$\sum_{i=0}^{n-3} \binom{n-3}{i}\cdot q_{i+j} = \frac{1}{8} \text{ für alle } 0 \leq j \leq 3.$$

Für die einzelnen Zeilen ergibt sich:

Zu $j = 0$:

$$\begin{aligned}\binom{n-3}{\frac{n-1}{4}}\cdot q_{\frac{n-1}{4}} + \binom{n-3}{\frac{n+1}{2}}\cdot q_{\frac{n+1}{2}} &= \frac{(n-3)!\cdot(\frac{n-1}{4})!\cdot(\frac{3\cdot n-11}{4})!\cdot 3\cdot(3\cdot n-7)}{(\frac{n-1}{4})!\cdot(\frac{3\cdot n-11}{4})!\cdot(n-3)!\cdot 8\cdot(n-2)\cdot(n+3)} \\ &+ \frac{(n-3)!\cdot(\frac{n+1}{2})!\cdot(\frac{n-5}{2})!\cdot(n-3)}{(\frac{n+1}{2})!\cdot(\frac{n-7}{2})!\cdot(n-2)!\cdot 4\cdot(n+3)} \\ &= \frac{3\cdot(3\cdot n-7)}{8\cdot(n-2)\cdot(n+3)} + \frac{(n-5)\cdot(n-3)}{2\cdot(n-2)\cdot(n+3)\cdot 4} = \frac{1}{8}.\end{aligned}$$

Zu $j = 1$:

$$\begin{aligned}\binom{n-3}{\frac{n-5}{4}}\cdot q_{\frac{n-1}{4}} + \binom{n-3}{\frac{n-1}{2}}\cdot q_{\frac{n+1}{2}} &= \frac{(n-3)!\cdot(\frac{n-1}{4})!\cdot(\frac{3\cdot n-11}{4})!\cdot 3\cdot(3\cdot n-7)}{(\frac{n-5}{4})!\cdot(\frac{3\cdot n-7}{4})!\cdot(n-3)!\cdot 8\cdot(n-2)\cdot(n+3)} \\ &+ \frac{(n-3)!\cdot(\frac{n+1}{2})!\cdot(\frac{n-5}{2})!\cdot(n-3)}{(\frac{n-1}{2})!\cdot(\frac{n-5}{2})!\cdot(n-2)!\cdot 4\cdot(n+3)} \\ &= \frac{4\cdot(n-1)\cdot 3\cdot(3\cdot n-7)}{8\cdot(n-2)\cdot(n+3)\cdot 4\cdot(3\cdot n-7)} + \frac{(n+1)\cdot(n-3)}{2\cdot(n-2)\cdot(n+3)\cdot 4} = \frac{1}{8}.\end{aligned}$$

Zu $j = 2$:

$$\binom{n-3}{\frac{n-9}{4}} \cdot q_{\frac{n-1}{4}} + \binom{n-3}{\frac{n-3}{2}} \cdot q_{\frac{n+1}{2}} = \frac{(n-3)! \cdot (\frac{n-1}{4})! \cdot (\frac{3 \cdot n-11}{4})! \cdot 3 \cdot (3 \cdot n-7)}{(\frac{n-9}{4})! \cdot (\frac{3 \cdot n-3}{4})! \cdot (n-3)! \cdot 8 \cdot (n-2) \cdot (n+3)}$$

$$+ \frac{(n-3)! \cdot (\frac{n+1}{2})! \cdot (\frac{n-5}{2})! \cdot (n-3)}{(\frac{n-3}{2})! \cdot (\frac{n-3}{2})! \cdot (n-2)! \cdot 4 \cdot (n+3)}$$

$$= \frac{3 \cdot 4 \cdot 4 \cdot (n-1) \cdot (n-5)}{4 \cdot 4 \cdot 3 \cdot (n-1) \cdot 8 \cdot (n-2) \cdot (n+3)} + \frac{(n+1) \cdot (n-1) \cdot 2 \cdot (n-3)}{4 \cdot (n-3) \cdot 4 \cdot (n+3) \cdot (n-2)} = \frac{1}{8}.$$

Zu $j = 3$:

$$\binom{n-3}{\frac{n-13}{4}} \cdot q_{\frac{n-1}{4}} + \binom{n-3}{\frac{n-5}{2}} \cdot q_{\frac{n+1}{2}} + q_n = \frac{(n-3)! \cdot (\frac{n-1}{4})! \cdot (\frac{3 \cdot n-11}{4})! \cdot 3 \cdot (3 \cdot n-7)}{(\frac{n-13}{4})! \cdot (\frac{3 \cdot n+1}{4})! \cdot (n-3)! \cdot 8 \cdot (n-2) \cdot (n+3)}$$

$$+ \frac{(n-3)! \cdot (\frac{n+1}{2})! \cdot (\frac{n-5}{2})! \cdot (n-3)}{(\frac{n-5}{2})! \cdot (\frac{n-1}{2})! \cdot (n-2)! \cdot 4 \cdot (n+3)} + \frac{1}{3 \cdot n+1}$$

$$= \frac{(n-1) \cdot (n-5) \cdot (n-9)}{(3 \cdot n+1) \cdot (n-1) \cdot 8 \cdot (n-2) \cdot (n+3)} + \frac{(3 \cdot n+1) \cdot (n+1) \cdot (n-3)}{2 \cdot (n-2) \cdot 4 \cdot (n+3)} + \frac{1}{3 \cdot n+1}$$

$$= \frac{(n-5) \cdot (n-9) + (3 \cdot n+1) \cdot (n+1) \cdot (n-3) + 8 \cdot (n-2) \cdot (n+3)}{8 \cdot (n-2) \cdot (n+3) \cdot (3 \cdot n+1)} = \frac{1}{8}.$$

Es ergibt sich:

$$E_f(Q) = q_n + \binom{n}{\frac{n+1}{2}} \cdot q_{\frac{n+1}{2}} = \frac{1}{3 \cdot n+1} + \binom{n}{\frac{n+1}{2}} \cdot \frac{n-3}{\binom{n-2}{\frac{n+1}{2}} \cdot 4 \cdot (n+3)} = \frac{3 \cdot n^2+2 \cdot n+3}{3 \cdot n^2+10 \cdot n+3} \approx \frac{n}{n+3}. \qquad \square$$

Vermutung :
Für alle $n \in \mathbb{N}$ mit $n > 1$ und $n \equiv 1 \bmod 4$ gilt: $E_f^*(3, n) = \frac{n}{n+3} + \frac{1}{3 \cdot n+1}$.

Proposition 10:
Für alle $n \in \mathbb{N}$ mit $n \equiv 3 \bmod 4$ gilt:

$$E_f^*(3, n) \geq \frac{9 \cdot n^3 + 27 \cdot n^2 - 29 \cdot n - 15}{9 \cdot n^3 + 51 \cdot n^2 + 27 \cdot n - 15}.$$

Beweis : Sei

$$q_{\frac{n-3}{4}} = \frac{4 \cdot n}{(n+1) \cdot (n+5) \cdot \binom{n}{\frac{n+1}{2}}}, \; q_{\frac{n+1}{4}} = \frac{4 \cdot n}{(3 \cdot n-1) \cdot (n+1) \cdot \binom{n}{\frac{n+1}{2}}},$$

$$q_{\frac{n+1}{2}} = \frac{n \cdot (n^2+2 \cdot n-7)}{(n-1) \cdot (n+1) \cdot (n+5) \cdot \binom{n}{\frac{n+1}{2}}}, \; q_n = \frac{3 \cdot n^2-2 \cdot n+3}{3 \cdot (n+1) \cdot (n-1) \cdot (3 \cdot n-1)} \text{ und } q_i = 0 \text{ sonst.}$$

Für $n \equiv 3 \bmod 4$ ist es klar, daß $q_i \geq 0$ für alle $0 \leq i \leq n$.

Damit nun $(q_0, \ldots, q_n)$ 3–unabhängig ist, muß gezeigt werden, daß

$$\sum_{i=0}^{n-3} \binom{n-3}{i} \cdot q_{i+j} = \frac{1}{8} \text{ für alle } 0 \leq j \leq 3.$$

Für die einzelnen Zeilen ergibt sich:

Zu $j = 0$:

$$\binom{n-3}{\frac{n-3}{4}} \cdot q_{\frac{n-3}{4}} + \binom{n-3}{\frac{n+1}{4}} \cdot q_{\frac{n+4}{2}} + \binom{n-3}{\frac{n+1}{2}} \cdot q_{\frac{n+1}{2}}$$

$$= \frac{(n-3)!\cdot(\frac{n+1}{4})!\cdot(\frac{3\cdot n-1}{4})!\cdot 4\cdot n}{(\frac{n-3}{4})!\cdot(\frac{3\cdot n-9}{4})!\cdot n!\cdot(n+1)\cdot(n+5)} + \frac{(n-3)!\cdot(\frac{n+1}{4})!\cdot(\frac{3\cdot n-1}{4})!\cdot 4\cdot n}{(\frac{n+1}{4})!\cdot(\frac{3\cdot n-13}{4})!\cdot n!\cdot(3\cdot n-1)\cdot(n+1)}$$

$$+ \frac{(n-3)\cdot(\frac{n+1}{2})!\cdot(\frac{n-1}{2})!\cdot n\cdot(n^2+2\cdot n-7)}{(\frac{n+1}{2})!\cdot(\frac{n-7}{2})!\cdot n!\cdot(n-1)\cdot(n+1)\cdot(n+5)}$$

$$= \frac{(n+1)\cdot(3\cdot n-1)\cdot(3\cdot n-5)}{16\cdot(n-1)\cdot(n-2)\cdot(n+1)\cdot(n+5)} + \frac{(3\cdot n-1)\cdot(3\cdot n-5)\cdot(3\cdot n-9)}{16\cdot(3\cdot n-1)\cdot(n-1)\cdot(n-2)\cdot(n+1)}$$

$$+ \frac{(n-1)\cdot(n-3)\cdot(n-5)\cdot(n^2+2\cdot n-7)}{8\cdot(n-2)\cdot(n-1)\cdot(n-1)\cdot(n+1)\cdot(n+5)}$$

$$= \frac{(3\cdot n-1)\cdot(3\cdot n-5)}{16\cdot(n-1)\cdot(n-2)\cdot(n+5)} + \frac{3\cdot(3\cdot n-5)\cdot(n-3)}{16\cdot(n-1)\cdot(n+1)\cdot(n-2)}$$

$$+ \frac{(n-3)\cdot(n-5)\cdot(n^2+2\cdot n-7)}{8\cdot(n-2)\cdot(n-1)\cdot(n+1)\cdot(n+5)}$$

$$= \frac{2\cdot(3\cdot n-5)\cdot(3\cdot n^2+4\cdot n-23)}{16\cdot(n-1)\cdot(n+1)\cdot(n-2)\cdot(n+5)} + \frac{(n^2-8\cdot n+15)\cdot(n^2+2\cdot n-7)}{8\cdot(n-2)\cdot(n-1)\cdot(n+1)\cdot(n+5)}$$

$$= \frac{n^4+3\cdot n^3-11\cdot n^2-3\cdot n+10}{8\cdot(n^4+3\cdot n^3-11\cdot n^2-3\cdot n+10)} = \frac{1}{8}.$$

Zu $j = 1$:

$$\binom{n-3}{\frac{n-7}{4}} \cdot q_{\frac{n-3}{4}} + \binom{n-3}{\frac{n-3}{4}} \cdot q_{\frac{n+4}{2}} + \binom{n-3}{\frac{n-1}{2}} \cdot q_{\frac{n+1}{2}}$$

$$= \frac{(n-3)!\cdot(\frac{n+1}{4})!\cdot(\frac{3\cdot n-1}{4})!\cdot 4\cdot n}{(\frac{n-7}{4})!\cdot(\frac{3\cdot n-5}{4})!\cdot n!\cdot(n+1)\cdot(n+5)} + \frac{(n-3)!\cdot(\frac{n+1}{4})!\cdot(\frac{3\cdot n-1}{4})!\cdot 4\cdot n}{(\frac{n-3}{4})!\cdot(\frac{3\cdot n-9}{4})!\cdot n!\cdot(3\cdot n-1)\cdot(n+1)}$$

$$+ \frac{(n-3)\cdot(\frac{n+1}{2})!\cdot(\frac{n-1}{2})!\cdot n\cdot(n^2+2\cdot n-7)}{(\frac{n-1}{2})!\cdot(\frac{n-5}{2})!\cdot n!\cdot(n-1)\cdot(n+1)\cdot(n+5)}$$

$$= \frac{(n+1)\cdot(3\cdot n-1)\cdot(n-3)}{16\cdot(n-1)\cdot(n-2)\cdot(n+1)\cdot(n+5)} + \frac{(3\cdot n-1)\cdot(3\cdot n-5)\cdot(n+1)}{16\cdot(3\cdot n-1)\cdot(n-1)\cdot(n-2)\cdot(n+1)}$$

$$+ \frac{(n-1)\cdot(n-3)\cdot(n+1)\cdot(n^2+2\cdot n-7)}{8\cdot(n-2)\cdot(n-1)\cdot(n-1)\cdot(n+1)\cdot(n+5)}$$

$$= \frac{(3\cdot n-1)\cdot(n-3)}{16\cdot(n-1)\cdot(n-2)\cdot(n+5)} + \frac{3\cdot n-5}{16\cdot(n-1)\cdot(n-2)} + \frac{(n-3)\cdot(n^2+2\cdot n-7)}{8\cdot(n-2)\cdot(n-1)\cdot(n+5)}$$

$$= \frac{2\cdot(3\cdot n^2-11)}{16\cdot(n-1)\cdot(n-2)\cdot(n+5)} + \frac{n^3-n^2-13\cdot n+21}{8\cdot(n-2)\cdot(n-1)\cdot(n+5)}$$

$$= \frac{n^3+2\cdot n^2-13\cdot n+10}{8\cdot(n^3+2\cdot n^2-13\cdot n+10)} = \frac{1}{8}.$$

Zu $j = 2$:

$$\binom{n-3}{\frac{n-11}{4}} \cdot q_{\frac{n-3}{4}} + \binom{n-3}{\frac{n-7}{4}} \cdot q_{\frac{n+4}{2}} + \binom{n-3}{\frac{n-3}{2}} \cdot q_{\frac{n+1}{2}}$$

$$= \frac{(n-3)! \cdot (\frac{n+1}{4})! \cdot (\frac{3\cdot n-1}{4})! \cdot 4 \cdot n}{(\frac{n-11}{4})! \cdot (\frac{3\cdot n-1}{4})! \cdot n! \cdot (n+1) \cdot (n+5)}$$

$$+ \frac{(n-3)! \cdot (\frac{n+1}{4})! \cdot (\frac{3\cdot n-1}{4})! \cdot 4 \cdot n}{(\frac{n-7}{4})! \cdot (\frac{3\cdot n-5}{4})! \cdot n! \cdot (3\cdot n-1) \cdot (n+1)} + \frac{(n-3) \cdot (\frac{n+1}{2})! \cdot (\frac{n-1}{2})! \cdot n \cdot (n^2+2\cdot n-7)}{(\frac{n-3}{2})! \cdot (\frac{n-3}{2})! \cdot n! \cdot (n-1) \cdot (n+1) \cdot (n+5)}$$

$$= \frac{(n+1) \cdot (n-3) \cdot (n-7)}{16 \cdot (n-1) \cdot (n-2) \cdot (n+1) \cdot (n+5)} + \frac{(3\cdot n-1) \cdot (n-3) \cdot (n+1)}{16 \cdot (3\cdot n-1) \cdot (n-1) \cdot (n-2) \cdot (n+1)}$$

$$+ \frac{(n-1) \cdot (n-1) \cdot (n+1) \cdot (n^2+2\cdot n-7)}{8 \cdot (n-2) \cdot (n-1) \cdot (n-1) \cdot (n+1) \cdot (n+5)}$$

$$= \frac{(n-3) \cdot (n-7)}{16 \cdot (n-1) \cdot (n-2) \cdot (n+5)} + \frac{n-3}{16 \cdot (n-1) \cdot (n-2)} + \frac{n^2+2\cdot n-7}{8 \cdot (n-2) \cdot (n+5)}$$

$$= \frac{2 \cdot (n-3) \cdot (n-1)}{16 \cdot (n-1) \cdot (n-2) \cdot (n+5)} + \frac{n^2+2\cdot n-7}{8 \cdot (n-2) \cdot (n+5)} = \frac{1}{8}.$$

Zu $j = 3$:

$$\binom{n-3}{\frac{n-15}{4}} \cdot q_{\frac{n-3}{4}} + \binom{n-3}{\frac{n-11}{4}} \cdot q_{\frac{n+4}{2}} + \binom{n-3}{\frac{n-5}{2}} \cdot q_{\frac{n+1}{2}} + q_n$$

$$= \frac{(n-3)! \cdot (\frac{n+1}{4})! \cdot (\frac{3\cdot n-1}{4})! \cdot 4 \cdot n}{(\frac{n-15}{4})! \cdot (\frac{3\cdot n+3}{4})! \cdot n! \cdot (n+1) \cdot (n+5)} + \frac{(n-3)! \cdot (\frac{n+1}{4})! \cdot (\frac{3\cdot n-1}{4})! \cdot 4 \cdot n}{(\frac{n-11}{4})! \cdot (\frac{3\cdot n-1}{4})! \cdot n! \cdot (3\cdot n-1) \cdot (n+1)}$$

$$+ \frac{(n-3) \cdot (\frac{n+1}{2})! \cdot (\frac{n-1}{2})! \cdot n \cdot (n^2+2\cdot n-7)}{(\frac{n-5}{2})! \cdot (\frac{n-1}{2})! \cdot n! \cdot (n-1) \cdot (n+1) \cdot (n+5)} + \frac{3\cdot n^2-2\cdot n+3}{3 \cdot (n+1) \cdot (n-1) \cdot (3\cdot n-1)}$$

$$= \frac{(n-11) \cdot (n-3) \cdot (n-7)}{16 \cdot 3 \cdot (n-1) \cdot (n-2) \cdot (n+1) \cdot (n+5)} + \frac{(n-7) \cdot (n-3) \cdot (n+1)}{16 \cdot (3\cdot n-1) \cdot (n-1) \cdot (n-2) \cdot (n+1)}$$

$$+ \frac{(n-1) \cdot (n-3) \cdot (n^2+2\cdot n-7)}{8 \cdot (n-2) \cdot (n-1) \cdot (n-1) \cdot (n+5)} + \frac{3\cdot n^2-2\cdot n+3}{3 \cdot (n+1) \cdot (n-1) \cdot (3\cdot n-1)}$$

$$= \frac{(n-11) \cdot (n-3) \cdot (n-7)}{16 \cdot 3 \cdot (n-1) \cdot (n-2) \cdot (n+1) \cdot (n+5)} + \frac{(n-7) \cdot (n-3)}{16 \cdot (3\cdot n-1) \cdot (n-1) \cdot (n-2)}$$

$$+ \frac{(n-3) \cdot (n^2+2\cdot n-7)}{8 \cdot (n-2) \cdot (n-1) \cdot (n+5)} + \frac{3\cdot n^2-2\cdot n+3}{3 \cdot (n+1) \cdot (n-1) \cdot (3\cdot n-1)}$$

$$= \frac{(n-3) \cdot (n-7) \cdot (3 \cdot (n+1) \cdot (n+5) + (n-11) \cdot (3\cdot n-1))}{16 \cdot 3 \cdot (n+1) \cdot (n-1) \cdot (n-2) \cdot (n+5) \cdot (3\cdot n-1)}$$

$$+ \frac{6 \cdot (n-3) \cdot (n+1) \cdot (3\cdot n-1) \cdot (n^2+2\cdot n-7) + (3\cdot n^2-2\cdot n+3) \cdot 16 \cdot (n+5) \cdot (n-2)}{16 \cdot 3 \cdot (n+1) \cdot (n-1) \cdot (n-2) \cdot (n+5) \cdot (3\cdot n-1)}$$

$$= \frac{6 \cdot (3\cdot n^5+8\cdot n^4-36\cdot n^3+2\cdot n^2+33\cdot n-10}{3 \cdot (n+1) \cdot (n-1) \cdot (3\cdot n-1) \cdot 16 \cdot (n+5) \cdot (n-2)} = \frac{1}{8}.$$

Es ergibt sich:

$$\begin{aligned} E_f(Q) &\geq q_n + \binom{n}{\frac{n+1}{2}} \cdot q_{\frac{n+1}{2}} \\ &= \frac{3 \cdot n^2 - 2 \cdot n + 3}{3 \cdot [n+1) \cdot (n-1) \cdot (3 \cdot n - 1)} + \binom{n}{\frac{n+1}{2}} \cdot \frac{n \cdot (n^2 + 2 \cdot n - 7)}{(n-1) \cdot (n+1) \cdot (n+5) \cdot \binom{n}{\frac{n+1}{2}}} \\ &= \frac{(3 \cdot n^2 - 2 \cdot n + 3) \cdot (n+5) + 3 \cdot n \cdot (n^2 + 2 \cdot n - 7) \cdot (3 \cdot n - 1)}{3 \cdot (n+1) \cdot (n-1) \cdot (3 \cdot n - 1) \cdot (n+5)} = \frac{9 \cdot n^3 + 27 \cdot n^2 - 29 \cdot n - 15}{9 \cdot n^3 + 51 \cdot n^2 + 27 \cdot n - 15}. \end{aligned}$$

□

Vermutung :
Für alle $n \in \mathbb{N}$ mit $n \equiv 3 \bmod 4$ gilt:

$$E_f^*(3, n) = \frac{9 \cdot n^3 + 27 \cdot n^2 - 29 \cdot n - 15}{9 \cdot n^3 + 51 \cdot n^2 + 27 \cdot n - 15} \approx \frac{n}{n+6}.$$

Proposition 11:
Für alle $n \in \mathbb{N}$ mit $n \equiv 2 \bmod 6$ gilt:

$$E_f^*(3, n) \geq \frac{4 \cdot n^4 + 31 \cdot n^3 + 24 \cdot n^2 - 2 \cdot n - 80}{4 \cdot n^4 + 58 \cdot n^3 + 186 \cdot n^2 + 52 \cdot n - 80}.$$

Beweis :

Sei

$$q_{\frac{n-2}{3}} = \frac{9 \cdot n}{(n+1) \cdot (n+10) \cdot \binom{n}{\frac{n+1}{3}}}, \; q_{\frac{n+1}{3}} = \frac{9 \cdot n}{2 \cdot (2 \cdot n - 1) \cdot (n+4) \cdot \binom{n}{\frac{n+1}{3}}},$$

$$q_{\frac{n+2}{2}} = \frac{(n+2) \cdot (n+7)}{(n+4) \cdot (n+10) \cdot \binom{n}{\frac{n}{2}}}, \; q_n = \frac{n-2}{2 \cdot (n+1) \cdot (2 \cdot n - 1)} \text{ und } q_i = 0 \text{ sonst.}$$

Für $n \in \mathbb{N}$ mit $n \equiv 2 \bmod 6$ ist es klar, daß $q_i \geq 0$ für alle $0 \leq i \leq n$.

Damit nun $(q_0, \ldots, q_n)$ 3–unabhängig ist, muß gezeigt werden, daß

$$\sum_{i=0}^{n-3} \binom{n-3}{i} \cdot q_{i+j} = \tfrac{1}{8} \text{ für alle } 0 \leq j \leq 3.$$

Für die einzelnen Zeilen ergibt sich:

Zu $j = 0$:

$$\binom{n-3}{\frac{n-2}{3}} \cdot q_{\frac{n-2}{3}} + \binom{n-3}{\frac{n+1}{3}} \cdot q_{\frac{n+1}{3}} + \binom{n-3}{\frac{n+2}{2}} \cdot q_{\frac{n+2}{2}} = \frac{(n-3)! \cdot (\frac{n+1}{3})! \cdot (\frac{2 \cdot n-1}{3})! \cdot 9 \cdot n}{(\frac{n-2}{3})! \cdot (\frac{2 \cdot n-7}{3})! \cdot n! \cdot (n+1) \cdot (n+10)}$$

$$+ \frac{(n-3)! \cdot (\frac{n+1}{3})! \cdot (\frac{2 \cdot n-1}{3})! \cdot 9 \cdot n}{(\frac{n+1}{3})! \cdot (\frac{2 \cdot n-10}{3})! \cdot n! \cdot 2 \cdot (2 \cdot n-1) \cdot (n+4)} + \frac{(n-3)! \cdot (\frac{n}{2})! \cdot (\frac{n}{2})! \cdot (n+2) \cdot (n+7)}{(\frac{n+2}{2})! \cdot (\frac{n-8}{2})! \cdot n! \cdot (n+4) \cdot (n+10)}$$

$$= \frac{(2 \cdot n-1) \cdot (2 \cdot n-4) \cdot n}{3 \cdot n \cdot (n-1) \cdot (n-2) \cdot (n+10)} + \frac{(2 \cdot n-1) \cdot (n-2) \cdot 2 \cdot (2 \cdot n-7)}{3 \cdot (n-1) \cdot (n-2) \cdot 2 \cdot (2 \cdot n-1) \cdot (n+4)} + \frac{2 \cdot (n-4) \cdot (n-6) \cdot (n+2) \cdot (n+7)}{(n+2) \cdot 16 \cdot (n-1) \cdot (n+4) \cdot (n+10)}$$

$$= \frac{(2 \cdot n-1) \cdot 2}{3 \cdot (n-1) \cdot (n+10)} + \frac{2 \cdot n-7}{3 \cdot (n-1) \cdot (n+4)} + \frac{(n-4) \cdot (n-6) \cdot (n+7)}{8 \cdot (n-1) \cdot (n+4) \cdot (n+10)} = \frac{1}{8}.$$

Zu $j = 1$:

$$\binom{n-3}{\frac{n-5}{3}} \cdot q_{\frac{n-2}{3}} + \binom{n-3}{\frac{n-2}{3}} \cdot q_{\frac{n+1}{3}} + \binom{n-3}{\frac{n}{2}} \cdot q_{\frac{n+2}{2}} = \frac{(n-3)! \cdot (\frac{n+1}{3})! \cdot (\frac{2 \cdot n-1}{3})! \cdot 9 \cdot n}{(\frac{n-5}{3})! \cdot (\frac{2 \cdot n-4}{3})! \cdot n! \cdot (n+1) \cdot (n+10)}$$

$$+ \frac{(n-3)! \cdot (\frac{n+1}{3})! \cdot (\frac{2 \cdot n-1}{3})! \cdot 9 \cdot n}{(\frac{n-2}{3})! \cdot (\frac{2 \cdot n-7}{3})! \cdot n! \cdot 2 \cdot (2 \cdot n-1) \cdot (n+4)} + \frac{(n-3)! \cdot (\frac{n}{2})! \cdot (\frac{n}{2})! \cdot (n+2) \cdot (n+7)}{(\frac{n}{2})! \cdot (\frac{n}{2}-3)! \cdot n! \cdot (n+4) \cdot (n+10)}$$

$$= \frac{(2 \cdot n-1) \cdot (n-2)}{3 \cdot (n-1) \cdot (n-2) \cdot (n+10)} + \frac{2 \cdot (n+1) \cdot (n-2)}{3 \cdot (n-1) \cdot (n-2) \cdot 2 \cdot (n+4)} + \frac{n \cdot (n-2) \cdot (n-4) \cdot (n+2) \cdot (n+7)}{8 \cdot n \cdot (n-1) \cdot (n-2) \cdot (n+4) \cdot (n+10)} = \frac{1}{8}.$$

Zu $j = 2$:

$$\binom{n-3}{\frac{n-8}{3}} \cdot q_{\frac{n-2}{3}} + \binom{n-3}{\frac{n-5}{3}} \cdot q_{\frac{n+1}{3}} + \binom{n-3}{\frac{n-2}{2}} \cdot q_{\frac{n+2}{2}} = \frac{(n-3)! \cdot (\frac{n+1}{3})! \cdot (\frac{2 \cdot n-1}{3})! \cdot 9 \cdot n}{(\frac{n-8}{3})! \cdot (\frac{2 \cdot n-1}{3})! \cdot n! \cdot (n+1) \cdot (n+10)}$$

$$+ \frac{(n-3)! \cdot (\frac{n+1}{3})! \cdot (\frac{2 \cdot n-1}{3})! \cdot 9 \cdot n}{(\frac{n-5}{3})! \cdot (\frac{2 \cdot n-4}{3})! \cdot n! \cdot 2 \cdot (2 \cdot n-1) \cdot (n+4)} + \frac{(n-3)! \cdot (\frac{n}{2})! \cdot (\frac{n}{2})! \cdot (n+2) \cdot (n+7)}{(\frac{n}{2}-1)! \cdot (\frac{n}{2}-2)! \cdot n! \cdot (n+4) \cdot (n+10)}$$

$$= \frac{(n-5) \cdot (n-2)}{3 \cdot (n-1) \cdot (n-2) \cdot (n+10)} + \frac{(n+1) \cdot (n-2) \cdot (2 \cdot n-1)}{3 \cdot (n-1) \cdot (n-2) \cdot 2 \cdot (n+4) \cdot (2 \cdot n-1)} + \frac{n \cdot (n-2) \cdot n \cdot (n+2) \cdot (n+7)}{8 \cdot n \cdot (n-1) \cdot (n-2) \cdot (n+4) \cdot (n+10)}$$

$$= \frac{n-5}{3 \cdot (n-1) \cdot (n+10)} + \frac{n+1}{6 \cdot (n-1) \cdot (n+4)} + \frac{n \cdot (n+2) \cdot (n+7)}{8 \cdot (n-1) \cdot (n+4) \cdot (n+10)} = \frac{1}{8}.$$

Zu $j = 3$:

$$\binom{n-3}{\frac{n-11}{3}} \cdot q_{\frac{n-2}{3}} + \binom{n-3}{\frac{n-8}{3}} \cdot q_{\frac{n+1}{3}} + \binom{n-3}{\frac{n-4}{2}} \cdot q_{\frac{n+2}{2}} + q_n = \frac{(n-3)! \cdot (\frac{n+1}{3})! \cdot (\frac{2 \cdot n-1}{3})! \cdot 9 \cdot n}{(\frac{n-11}{3})! \cdot (\frac{2 \cdot n+2}{3})! \cdot n! \cdot (n+1) \cdot (n+10)}$$

$$+ \frac{(n-3)! \cdot (\frac{n+1}{3})! \cdot (\frac{2 \cdot n-1}{3})! \cdot 9 \cdot n}{(\frac{n-8}{3})! \cdot (\frac{2 \cdot n-1}{3})! \cdot n! \cdot 2 \cdot (2 \cdot n-1) \cdot (n+4)} + \frac{(n-3)! \cdot (\frac{n}{2})! \cdot (\frac{n}{2})! \cdot (n+2) \cdot (n+7)}{(\frac{n}{2}-1)! \cdot (\frac{n}{2}-2)! \cdot n! \cdot (n+4) \cdot (n+10)} + \frac{n-2}{2 \cdot (n+1) \cdot (2 \cdot n-1)}$$

$$= \frac{(n-5) \cdot (n-8)}{6 \cdot (n-1) \cdot (n+1) \cdot (n+10)} + \frac{(n+1) \cdot (n-5)}{6 \cdot (n-1) \cdot (2 \cdot n-1) \cdot (n+4)}$$

$$+ \frac{n \cdot (n-2) \cdot n \cdot (n+2) \cdot (n+7)}{8 \cdot n \cdot (n-1) \cdot (n-2) \cdot (n+4) \cdot (n+10)} + \frac{n-2}{2 \cdot (n+1) \cdot (2 \cdot n-1)}$$

$$= \frac{(n-5) \cdot (n-8) \cdot 4 \cdot (n+4) \cdot (2 \cdot n-1) + (n+1) \cdot (n-5) \cdot 4 \cdot (n+10) \cdot (n+1)}{24 \cdot (n+1) \cdot (n-1) \cdot (n+10) \cdot (2 \cdot n-1) \cdot (n+4)}$$

$$+ \frac{n \cdot (n+2) \cdot (n+7) \cdot 3 \cdot (2 \cdot n-1) \cdot (n+1) + (n-2) \cdot 12 \cdot (n+4) \cdot (n-1) \cdot (n+10)}{24 \cdot (n+1) \cdot (n-1) \cdot (n+10) \cdot (2 \cdot n-1) \cdot (n+4)}$$

$$= \frac{6 \cdot n^5 + 84 \cdot n^4 + 192 \cdot n^3 - 201 \cdot n^2 - 198 \cdot n + 120}{8 \cdot (6 \cdot n^5 + 84 \cdot n^4 + 192 \cdot n^3 - 201 \cdot n^2 - 198 \cdot n + 120)} = \frac{1}{8}.$$

Es ergibt sich:

$$E_f(Q) = q_n + \binom{n}{\frac{n+2}{2}} \cdot q_{\frac{n+2}{2}} = \frac{n-2}{2\cdot(n+1)\cdot(2\cdot n-1)} + \frac{n!\cdot(\frac{n}{2})!\cdot(\frac{n}{2})!\cdot(n+2)\cdot(n+7)}{(\frac{n}{2}+1)!\cdot(\frac{n}{2}-1)!\cdot n!\cdot(n+4)\cdot(n+10)}$$

$$= \frac{n-2}{2\cdot(n+1)\cdot(2\cdot n-1)} + \frac{n\cdot(n+7)}{(n+4)\cdot(n+10)} = \frac{4\cdot n^4+31\cdot n^3+24\cdot n^2-2\cdot n-80}{4\cdot n^4+58\cdot n^3+186\cdot n^2+52\cdot n-80}.$$

□

Vermutung :
Für alle $n \in \mathbb{N}$ mit $n \equiv 2 \bmod 6$ gilt:

$$E_f^*(3,n) = \frac{4\cdot n^4 + 31\cdot n^3 + 24\cdot n^2 - 2\cdot n - 80}{4\cdot n^4 + 58\cdot n^3 + 186\cdot n^2 + 52\cdot n - 80} \approx \frac{n}{n+15}.$$

Proposition 12:
Für alle $n \in \mathbb{N}$ mit $n \equiv 4 \bmod 6$ gilt:

$$E_f^*(3,n) \geq \frac{4\cdot n^3 - n^2 + 2\cdot n - 32}{4\cdot n^3 + 26\cdot n^2 - 52\cdot n - 32}.$$

Beweis : Sei

$$q_{\frac{n-1}{3}} = \frac{27\cdot n}{2\cdot(n+2)\cdot(n+8)\cdot\binom{n}{\frac{n+2}{3}}},\ q_{\frac{n+2}{2}} = \frac{(n+2)\cdot(n-1)}{(n-2)\cdot(n+8)\cdot\binom{n}{\frac{n}{2}}},$$

$$q_n = \frac{n-4}{2\cdot(2\cdot n+1)\cdot(n-2)} \text{ und } q_i = 0 \text{ sonst.}$$

Für $n \in \mathbb{N}$ mit $n \equiv 4 \bmod 6$ ist es klar, daß $q_i \geq 0$ für alle $0 \leq i \leq n$.

Damit nun $(q_0, \ldots, q_n)$ 3–unabhängig ist, muß gezeigt werden, daß

$$\sum_{i=0}^{n-3} \binom{n-3}{i} \cdot q_{i+j} = \frac{1}{8} \text{ für alle } 0 \leq j \leq 3.$$

Für die einzelnen Zeilen ergibt sich:

Zu $j = 0$:

$$\binom{n-3}{\frac{n-1}{3}} \cdot q_{\frac{n-1}{3}} + \binom{n-3}{\frac{n+2}{2}} \cdot q_{\frac{n+2}{2}} = \frac{(n-3)!\cdot(\frac{n+2}{3})!\cdot(\frac{2\cdot n-2}{3})!\cdot 27\cdot n}{(\frac{n-1}{3})!\cdot(\frac{2\cdot n-8}{3})!\cdot n!\cdot 2\cdot(n+2)\cdot(n+8)} + \frac{(n-3)!\cdot(\frac{n}{2})!\cdot(\frac{n}{2})!\cdot(n+2)\cdot(n-1)}{(\frac{n}{2}+1)!\cdot(\frac{n}{2}-4)!\cdot n!\cdot(n-2)\cdot(n+8)}$$

$$= \frac{(n+2)\cdot(2\cdot n-2)\cdot(2\cdot n-5)\cdot 27}{27\cdot(n-1)\cdot(n-2)\cdot 2\cdot(n+2)\cdot(n+8)} + \frac{2\cdot n\cdot(n-2)\cdot(n+2)\cdot(n-4)\cdot(n-6)}{(n+2)\cdot 16\cdot n\cdot(n-2)\cdot(n-2)\cdot(n+8)}$$

$$= \frac{2\cdot n-5}{(n-2)\cdot(n+8)} + \frac{(n-4)\cdot(n-6)}{8\cdot(n-2)\cdot(n+8)} = \frac{1}{8}.$$

Zu $j = 1$:

$$\binom{n-3}{\frac{n-4}{3}} \cdot q_{\frac{n-1}{3}} + \binom{n-3}{\frac{n}{2}} \cdot q_{\frac{n+2}{2}} = \frac{(n-3)!\cdot(\frac{n+2}{3})!\cdot(\frac{2\cdot n-2}{3})!\cdot 27\cdot n}{(\frac{n-4}{3})!\cdot(\frac{2\cdot n-5}{3})!\cdot n!\cdot 2\cdot(n+2)\cdot(n+8)} + \frac{(n-3)!\cdot(\frac{n}{2})!\cdot(\frac{n}{2})!\cdot(n+2)\cdot(n-1)}{(\frac{n}{2})!\cdot(\frac{n}{2}-3)!\cdot n!\cdot(n-2)\cdot(n+8)}$$

$$= \frac{(n+2)\cdot(2\cdot n-2)\cdot 6\cdot(n-1)}{3\cdot(n-1)\cdot(n-2)\cdot 2\cdot(n+2)\cdot(n+8)} + \frac{(n-2)\cdot(n+2)\cdot(n-4)}{8\cdot(n-2)\cdot(n-2)\cdot(n+8)}$$

$$= \frac{n-1}{(n-2)\cdot(n+8)} + \frac{(n-4)\cdot(n+2)}{8\cdot(n-2)\cdot(n+8)} = \frac{1}{8}.$$

Zu $j = 2$:

$$\binom{n-3}{\frac{n-7}{3}} \cdot q_{\frac{n-1}{3}} + \binom{n-3}{\frac{n}{2}-1} \cdot q_{\frac{n+2}{2}} = \frac{(n-3)!\cdot(\frac{n+2}{3})!\cdot(\frac{2\cdot n-2}{3})!\cdot 27\cdot n}{(\frac{n-7}{3})!\cdot(\frac{2\cdot n-2}{3})!\cdot n!\cdot 2\cdot(n+2)\cdot(n+8)} + \frac{(n-3)!\cdot(\frac{n}{2})!\cdot(\frac{n}{2})!\cdot(n+2)\cdot(n-1)}{(\frac{n}{2}-1)!\cdot(\frac{n}{2}-2)!\cdot n!\cdot(n-2)\cdot(n+8)}$$

$$= \frac{(n+2)\cdot(n-1)\cdot(n-4)}{2\cdot(n-1)\cdot(n-2)\cdot(n+2)\cdot(n+8)} + \frac{(n-2)\cdot(n+2)\cdot n}{8\cdot(n-2)\cdot(n-2)\cdot(n+8)} = \frac{1}{8}.$$

Zu $j = 3$:

$$\binom{n-3}{\frac{n-10}{3}} \cdot q_{\frac{n-1}{3}} + \binom{n-3}{\frac{n}{2}-2} \cdot q_{\frac{n+2}{2}} + q_n = \frac{(n-3)!\cdot(\frac{n+2}{3})!\cdot(\frac{2\cdot n-2}{3})!\cdot 27\cdot n}{(\frac{n-10}{3})!\cdot(\frac{2\cdot n+1}{3})!\cdot n!\cdot 2\cdot(n+2)\cdot(n+8)}$$

$$+\frac{(n-3)!\cdot(\frac{n}{2})!\cdot(\frac{n}{2})!\cdot(n+2)\cdot(n-1)}{(\frac{n}{2}-1)!\cdot(\frac{n}{2}-2)!\cdot n!\cdot(n-2)\cdot(n+8)} + \frac{n-4}{2\cdot(2\cdot n+1)\cdot(n-2)}$$

$$= \frac{(n+2)\cdot(n-1)\cdot(n-4)\cdot(n-7)}{2\cdot(2\cdot n+1)\cdot(n-1)\cdot(n-2)\cdot(n+2)\cdot(n+8)} + \frac{(n-2)\cdot(n+2)\cdot n}{8\cdot(n-2)\cdot(n-2)\cdot(n+8)} + \frac{n-4}{2\cdot(2\cdot n+1)\cdot(n-2)}$$

$$= \frac{(n-4)\cdot(n-7)}{(2\cdot n+1)\cdot(n-2)\cdot(n+8)\cdot 2} + \frac{n\cdot(n+2)}{8\cdot(n-2)\cdot(n+8)} + \frac{n-4}{2\cdot(2\cdot n+1)\cdot(n-2)}$$

$$= \frac{4\cdot(n^2-11\cdot n+28)+(n^2+2\cdot n)\cdot(2\cdot n+1)+4\cdot(n^2+4\cdot n-32)}{8\cdot(2\cdot n+1)\cdot(n^2+6\cdot n-16)} = \frac{1}{8}.$$

Es ergibt sich:

$$E_f(Q) = q_n + \binom{n}{\frac{n+2}{2}} \cdot q_{\frac{n+2}{2}}$$

$$= \frac{n-4}{2\cdot(2\cdot n+1)\cdot(n-2)} + \frac{n!\cdot(\frac{n}{2})!\cdot(\frac{n}{2})!\cdot(n+2)\cdot(n-1)}{(\frac{n}{2}+1)!\cdot(\frac{n}{2}-1)!\cdot n!\cdot(n-2)!\cdot(n+8)}$$

$$= \frac{n-4}{2\cdot(2\cdot n+1)\cdot(n-2)} + \frac{2\cdot n\cdot(n+2)\cdot(n-1)}{(n+2)\cdot 2\cdot(n-2)\cdot(n+8)} = \frac{4\cdot n^3-n^2+2\cdot n-32}{4\cdot n^3+26\cdot n^2-52\cdot n-32}.$$

□

Vermutung :
Für alle $n \in \mathbb{N}$ mit $n \equiv 4 \bmod 6$ gilt:

$$E_f^*(3,n) = \frac{4\cdot n^3 - n^2 + 2\cdot n - 32}{4\cdot n^3 + 26\cdot n^2 - 52\cdot n - 32} \approx \frac{n}{n+7}.$$

Proposition 13:
Für alle $n \in \mathbb{N}$ mit $n \equiv 0 \bmod 6$ gilt:

$$E_f^*(3,n) \geq \frac{4\cdot n^3 + 51\cdot n^2 + 99\cdot n + 108}{4\cdot n^3 + 78\cdot n^2 + 396\cdot n + 432}.$$

Beweis :

Sei

$$q_{\frac{n-3}{3}} = \frac{9}{2\cdot(n+12)\cdot\binom{n}{\frac{n}{3}}},\ q_{\frac{n}{3}} = \frac{9}{2\cdot(n+6)\cdot\binom{n}{\frac{n}{3}}},$$

$$q_{\frac{n+2}{2}} = \frac{(n+2)\cdot(n^2+9\cdot n-18)}{(n-2)\cdot(n+6)\cdot(n+12)\cdot\binom{n}{\frac{n}{2}}},\ q_n = \frac{n-3}{2\cdot(2\cdot n+3)\cdot(n-2)}$$

und $q_i = 0$ sonst.

Für $n \in \mathbb{N}$ mit $n \equiv 0 \bmod 6$ ist es klar, daß $q_i \geq 0$ für alle $0 \leq i \leq n$.

Damit nun $(q_0, \ldots, q_n)$ 3–unabhängig ist, muß gezeigt werden, daß

$$\sum_{i=0}^{n-3}\binom{n-3}{i}\cdot q_{i+j} = \tfrac{1}{8} \text{ für alle } 0 \leq j \leq 3.$$

Für die einzelnen Zeilen ergibt sich:

Zu $j = 0$:

$$\binom{n-3}{\frac{n-3}{3}}\cdot q_{\frac{n-3}{3}} + \binom{n-3}{\frac{n}{3}}\cdot q_{\frac{n}{3}} + \binom{n-3}{\frac{n+2}{2}}\cdot q_{\frac{n+2}{2}} = \tfrac{(n-3)!\cdot(\frac{n}{3})!\cdot(\frac{2\cdot n}{3})!\cdot 9}{(\frac{n-3}{3})!\cdot(\frac{2\cdot n-6}{3})!\cdot n!\cdot 2\cdot(n+12)}$$

$$+\tfrac{(n-3)!\cdot(\frac{n}{3})!\cdot(\frac{2\cdot n}{3})!\cdot 9}{(\frac{n}{3})!\cdot(\frac{2\cdot n}{3}-3)!\cdot n!\cdot 2\cdot(n+6)} + \tfrac{(n-3)!\cdot(\frac{n}{2})!\cdot(\frac{n}{2})!\cdot(n+2)\cdot(n^2+9\cdot n-18)}{(\frac{n+2}{2})!\cdot(\frac{n}{2}-4)!\cdot n!\cdot(n-2)\cdot(n+6)\cdot(n+12)}$$

$$= \tfrac{(2\cdot n-3)\cdot n}{3\cdot(n-1)\cdot(n-2)\cdot(n+12)} + \tfrac{(2\cdot n-3)\cdot 2\cdot(n-3)}{3\cdot(n-1)\cdot(n-2)\cdot(n+6)} + \tfrac{(n-4)\cdot(n-6)\cdot(n^2+9\cdot n-18)}{8\cdot(n-1)\cdot(n+6)\cdot(n+12)}$$

$$= \tfrac{8\cdot(2\cdot n-3)\cdot(n+6)\cdot n+16\cdot(2\cdot n-3)\cdot(n^2+9\cdot n-36)+3\cdot(n^2-10\cdot n+24)\cdot(n^2+9\cdot n-18)}{24\cdot(n-1)\cdot(n-2)\cdot(n+6)\cdot(n+12)}$$

$$= \tfrac{n^4+15\cdot n^3+20\cdot n^2-180\cdot n+144}{8\cdot(n^4+15\cdot n^3+20\cdot n^2-180\cdot n+144)}$$

$$= \tfrac{1}{8}.$$

Zu $j = 1$:

$$\binom{n-3}{\frac{n-6}{3}}\cdot q_{\frac{n-3}{3}} + \binom{n-3}{\frac{n-3}{3}}\cdot q_{\frac{n}{3}} + \binom{n-3}{\frac{n}{2}}\cdot q_{\frac{n+2}{2}} = \tfrac{(n-3)!\cdot(\frac{n}{3})!\cdot(\frac{2\cdot n}{3})!\cdot 9}{(\frac{n-6}{3})!\cdot(\frac{2\cdot n-3}{3})!\cdot n!\cdot 2\cdot(n+12)}$$

$$+\tfrac{(n-3)!\cdot(\frac{n}{3})!\cdot(\frac{2\cdot n}{3})!\cdot 9}{(\frac{n-3}{3})!\cdot(\frac{2\cdot n}{3}-2)!\cdot n!\cdot 2\cdot(n+6)} + \tfrac{(n-3)!\cdot(\frac{n}{2})!\cdot(\frac{n}{2})!\cdot(n+2)\cdot(n^2+9\cdot n-18)}{(\frac{n}{2})!\cdot(\frac{n}{2}-3)!\cdot n!\cdot(n-2)\cdot(n+6)\cdot(n+12)}$$

$$= \tfrac{(n-3)\cdot n}{3\cdot(n-1)\cdot(n-2)\cdot(n+12)} + \tfrac{(2\cdot n-3)\cdot n}{3\cdot(n-1)\cdot(n-2)\cdot(n+6)} + \tfrac{(n-4)\cdot(n+2)\cdot(n^2+9\cdot n-18)}{8\cdot(n-1)\cdot(n+6)\cdot(n+12)\cdot(n-2)}$$

$$= \tfrac{3\cdot(n^4+15\cdot n^3+20\cdot n^2-180\cdot n+144)}{24\cdot(n^4+15\cdot n^3+20\cdot n^2-180\cdot n+144)} = \tfrac{1}{8}.$$

Zu $j = 2$:

$$\binom{n-3}{\frac{n-9}{3}} \cdot q_{\frac{n-3}{3}} + \binom{n-3}{\frac{n-6}{3}} \cdot q_{\frac{n}{3}} + \binom{n-3}{\frac{n-2}{2}} \cdot q_{\frac{n+2}{2}} = \frac{(n-3)! \cdot (\frac{n}{3})! \cdot (\frac{2 \cdot n}{3})! \cdot 9}{(\frac{n-9}{3})! \cdot (\frac{2 \cdot n}{3})! \cdot n! \cdot 2 \cdot (n+12)}$$

$$+ \frac{(n-3)! \cdot (\frac{n}{3})! \cdot (\frac{2 \cdot n}{3})! \cdot 9}{(\frac{n-6}{3})! \cdot (\frac{2 \cdot n}{3}-1)! \cdot n! \cdot 2 \cdot (n+6)} + \frac{(n-3)! \cdot (\frac{n}{2})! \cdot (\frac{n}{2})! \cdot (n+2) \cdot (n^2+9 \cdot n-18)}{(\frac{n}{2}-1)! \cdot (\frac{n}{2}-2)! \cdot n! \cdot (n-2) \cdot (n+6) \cdot (n+12)}$$

$$= \frac{(n-3) \cdot (n-6)}{6 \cdot (n-1) \cdot (n-2) \cdot (n+12)} + \frac{(n-3) \cdot n}{3 \cdot (n-1) \cdot (n-2) \cdot (n+6)} + \frac{n \cdot (n+2) \cdot (n^2+9 \cdot n-18)}{8 \cdot (n-1) \cdot (n+6) \cdot (n+12) \cdot (n-2)}$$

$$= \frac{3 \cdot (n^4+15 \cdot n^3+20 \cdot n^2-180 \cdot n+144)}{24 \cdot (n^4+15 \cdot n^3+20 \cdot n^2-180 \cdot n+144)} = \frac{1}{8}.$$

Zu $j = 3$:

$$\binom{n-3}{\frac{n-12}{3}} \cdot q_{\frac{n-3}{3}} + \binom{n-3}{\frac{n-9}{3}} \cdot q_{\frac{n}{3}} + \binom{n-3}{\frac{n-4}{2}} \cdot q_{\frac{n+2}{2}} + q_n = \frac{(n-3)! \cdot (\frac{n}{3})! \cdot (\frac{2 \cdot n}{3})! \cdot 9}{(\frac{n-12}{3})! \cdot (\frac{2 \cdot n+3}{3})! \cdot n! \cdot 2 \cdot (n+12)}$$

$$+ \frac{(n-3)! \cdot (\frac{n}{3})! \cdot (\frac{2 \cdot n}{3})! \cdot 9}{(\frac{n-9}{3})! \cdot (\frac{2 \cdot n}{3})! \cdot n! \cdot 2 \cdot (n+6)} + \frac{(n-3)! \cdot (\frac{n}{2})! \cdot (\frac{n}{2})! \cdot (n+2) \cdot (n^2+9 \cdot n-18)}{(\frac{n}{2}-2)! \cdot (\frac{n}{2}-1)! \cdot n! \cdot (n-2) \cdot (n+6) \cdot (n+12)} + \frac{n-3}{2 \cdot (n-2) \cdot (2 \cdot n+3)}$$

$$= \frac{(n-3) \cdot (n-6) \cdot (n-9)}{6 \cdot (n-1) \cdot (n-2) \cdot (n+12) \cdot (23 \cdot n+3)} + \frac{(n-3) \cdot (n-6)}{6 \cdot (n-1) \cdot (n-2) \cdot (n+6)}$$

$$+ \frac{n \cdot (n+2) \cdot (n^2+9 \cdot n-18)}{8 \cdot (n-1) \cdot (n+6) \cdot (n+12) \cdot (n-2)} + \frac{n-3}{2 \cdot (n-2) \cdot (2 \cdot n+3)}$$

$$= \frac{3 \cdot (2 \cdot n^5+33 \cdot n^4+85 \cdot n^3-300 \cdot n^2-252 \cdot n+432)}{24 \cdot (2 \cdot n^5+33 \cdot n^4+85 \cdot n^3-300 \cdot n^2-252 \cdot n+432)} = \frac{1}{8}.$$

Es ergibt sich:

$$E_f(Q) = q_n + \binom{n}{\frac{n+2}{2}} \cdot q_{\frac{n+2}{2}}$$

$$= \frac{n-3}{2 \cdot (n-2) \cdot (2 \cdot n+3)} + \frac{n! \cdot (\frac{n}{2})! \cdot (\frac{n}{2})! \cdot (n+2) \cdot (n^2+9 \cdot n-18)}{(\frac{n}{2}+1)! \cdot (\frac{n}{2}-1)! \cdot n! \cdot (n-2) \cdot (n+6) \cdot (n+12)}$$

$$= \frac{n-3}{2 \cdot (n-2) \cdot (2 \cdot n+3)} + \frac{n \cdot (n^2+9 \cdot n-18)}{(n-2) \cdot (n+6) \cdot (n+12)} = \frac{4 \cdot n^3+51 \cdot n^2+99 \cdot n+108}{4 \cdot n^3+78 \cdot n^2+396 \cdot n+432}.$$

□

Vermutung :
Für alle $n \in \mathbb{N}$ mit $n \equiv 0 \mod 6$ gilt:

$$E_f^*(3, n) = \frac{4 \cdot n^3 + 51 \cdot n^2 + 99 \cdot n + 108}{4 \cdot n^3 + 78 \cdot n^2 + 396 \cdot n + 432} \approx \frac{n}{n+20}.$$

Durch die Propositionen 9–13 ist für den Fall $k = 3$ bewiesen, daß

$$\lim_{n \to \infty} E_f^*(3, n) = 1.$$

2.7 Zusammenfassung und offene Fragen

Da $E_f^*(k,n)$ in k monoton fallend ist, existiert ein minimales k, so daß

$$E_f^*(0,n) - E_f^*(k,n) \geq E_f^*(k,n) - E_f^*(n,n).$$

Sei zu gegebener Funktion f und gegebenem $n \in \mathbb{N}$ dieses k bezeichnet durch $k_f(n)$. $k_f(n)$ ist also der kleinste Wert für k, so daß $E_f^*(k,n)$ näher an $E_f^*(n,n)$ liegt als an $E_f^*(0,n)$.
Im Kapitel 5 ist gezeigt worden, daß für die Betragsfunktion $k_f(n) = 2$ für alle $n \in \mathbb{N}$.

In Kapitel 6 wurde die Mehrheitsfunktion betrachtet, für die $k_f(n)$ vermutlich nicht konstant ist.

Allgemein habe ich folgende Vermutungen und offene Fragen:

1. Gilt für die Mehrheitsfunktion und für jedes feste $k \in \mathbb{N}$:
$$\lim_{n\to\infty} E_f^*(k,n) = 1 \quad ?$$
2. Wie ist der Verlauf von $k_f(n)$ für die Mehrheitsfunktion ?
3. Existiert zu jedem Paar $k \leq n$ eine Funktion f, so daß $k_f(n) = k$?
4. Erhält man die gleichen Werte $k_f(n)$, wenn man den Erwartungswert nicht maximiert sondern minimiert ?
5. Lassen sich Ergebnisse und Phänomene vom binären Alphabet (siehe z.B. Lemma 1 und Proposition 1) auf größere Alphabete mit den Randverteilungen $(\frac{1}{m}, \ldots, \frac{1}{m})$ übertragen ?

2.8 Literatur

[AGHP 90] N. Alon, O. Goldreich, J. Hastad and R. Peralta, Simple constructions of almost k–wise independent random variables, Proceedings of the 31st FOCS, St. Louis, MO, IEEE, New York (1990), 544–553.

[AS 92] N. Alon and J.H. Spencer, The Probabilistic Method, Wiley–Interscience Series in Discrete Mathematics and Optimization, John Wiley and Sons, Inc. (1992)

[NN 90] J. Noar and M. Noar, Small–bias probability spaces: efficient constructions and applications, Proceedings of the 22nd annual ACM STOC, ACM Press, New York (1990), 213–223.