

$$\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d}$$

(deutsche Version)

Darij Grinberg

Erweitertes Skript zu einem Vortrag beim QED-Seminar in
Augsburg, April 2010
Version 10.5.2010

Inhalt

1.	$\frac{1}{n} \sum_{d n} \phi(d) q^{n/d} \in \mathbb{Z}$ für $n \in \mathbb{N}_+$ und $q \in \mathbb{Z}$	2
2.	Zahlentheorie I: Die Dirichletfaltung und der Äquivalenzsatz für Geisterwittfolgen	5
2.1.	Der Äquivalenzsatz	5
2.2.	Aus dem Äquivalenzsatz folgt Satz 1.1	7
2.3.	Beweis von Satz 2.1: Einleitung	9
2.4.	Die Dirichletfaltung	9
2.5.	Beweis von Satz 2.1: $\mathcal{D} \iff \mathcal{E} \iff \mathcal{F} \iff \mathcal{G}$	16
2.6.	Beweis von Satz 2.1: $\mathcal{A} \iff \mathcal{D}$	19
2.7.	Beweis von Satz 2.1: $\mathcal{G} \iff \mathcal{H}$	22
2.8.	Beweis von Satz 2.1: $\mathcal{A} \iff \mathcal{B}$	23
2.9.	Beweis von Satz 2.1: $\mathcal{B} \iff \mathcal{C}$	26
3.	Potenzreihen und ein alternativer Beweis von $\mathcal{B} \iff \mathcal{D}$	27
3.1.	Was sind Potenzreihen?	28
3.2.	Die binomische Formel und die Vandermonde-Faltungsformel: Eine erste Anwendung von Potenzreihen	33
3.3.	Einsetzung von Potenzreihen ineinander	38
3.4.	Rationale Potenzreihen und Ableitung	41
3.5.	Der natürliche Logarithmus einer rationalen Potenzreihe	42
3.6.	Ausblicke zu Potenzreihen	49
3.7.	Zwei "Normalformen" für Potenzreihen	50
3.8.	Beweis von Satz 2.1: ein alternativer Beweis von $\mathcal{B} \iff \mathcal{D}$	53

4. Binomialkoeffizienten statt Potenzen	56
4.1. $\sum_{d n} \phi(d) \binom{qn/d}{rn/d} \in n\mathbb{Z}$ und Analoga	56
4.2. Lemmata zum Beweis	59
4.3. Beweis von Satz 4.1 und Satz 4.3	69
4.4. Ein Zusatz zum Äquivalenzsatz 2.1	70
5. Kombinatorik I: Das Pólya-Burnsidesche Abzählverfahren	70
5.1. Perlenketten mit n Perlen in q Farben	70
5.2. Anzahlen von Äquivalenzklassen: ein fast triviales Lemma	72
5.3. Abbildungsgruppen und induzierte Äquivalenzrelationen: das Nicht-Burnside-Lemma	73
5.4. Beweis von Satz 5.1	79
5.5. Aperiodische Perlenketten und $\frac{1}{n} \sum_{d n} \mu(d) q^{n/d}$	85
6. Der negative Fall und die Kombinatorik	93
7. Kombinatorik II: Nester und Fixpunktzahlen	98
8. Zahlentheorie III: Irreduzible Polynome über \mathbb{F}_q	100
9. Freie Liealgebren und Dimensionen	100
10. Wittpolynome	101
11. Zahlentheorie II: Geisterburnsidefolgen	101
12. Spuren von Matrixpotenzen	101
13. Die Perlenketten-Identität	101
14. Kombinatorik III: Teilmengen von $\{1, 2, \dots, n\}$ mit durch n teilbarer Summe	102

$$\boxed{\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d}} \quad \text{(deutsche Version)}$$

[not completed, not proofread]

[derzeit bis einschließlich Abschnitt 5.5 fertig]

[Fehler im ersten Beweis von Lemma 4.4 korrigiert]

1. $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} \in \mathbb{Z}$ **für** $n \in \mathbb{N}_+$ **und** $q \in \mathbb{Z}$

Der Titel dieses Vortrages ist auch sein roter Faden: es geht um die Zahlen

$$\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} \quad \text{für } n \in \mathbb{N}_+ \text{ und } q \in \mathbb{Z}$$

(was das Zeichen ϕ bedeutet, wird einige Zeilen weiter unten erklärt). Vordergründig geht es darum, zu beweisen, daß diese Zahlen immer ganz sind. In Wirklichkeit sind die verschiedenen Beweise jeweils für sich interessanter, als ebendiese Aussage, die sie beweisen. Insbesondere sind mehrere Verallgemeinerungen in verschiedene Richtungen möglich.

Wir führen zuallererst die benötigten Begriffe ein:

Definition (natürliche Zahlen): Wir verstehen unter \mathbb{N} die Menge $\{0, 1, 2, \dots\}$, und unter \mathbb{N}_+ die Menge $\{1, 2, 3, \dots\}$. (Einige andere Autoren schreiben stattdessen \mathbb{N} für $\{1, 2, 3, \dots\}$.) Wir bezeichnen die Elemente der Menge $\mathbb{N} = \{0, 1, 2, \dots\}$ als *natürliche Zahlen*.

Definition ($\mathbb{N}_{|n}$): Für jedes $n \in \mathbb{N}_+$ bezeichne $\mathbb{N}_{|n}$ die Menge aller positiven Teiler von n , also $\mathbb{N}_{|n} = \{d \in \mathbb{N}_+ \mid (d \mid n)\}$. Wir verwenden das Zeichen $\sum_{d|n}$ als Synonym für $\sum_{d \in \mathbb{N}_{|n}}$. Das heißt zum Beispiel: $\sum_{d|n} d$ ist die Summe aller positiven Teiler der Zahl n .

Definition (Eulersche Phi-Funktion ϕ): Die *Eulersche Phi-Funktion* $\phi : \mathbb{N}_+ \rightarrow \mathbb{Z}$ wird definiert durch

$$\phi(n) = (\text{Anzahl aller Zahlen } m \in \{1, 2, \dots, n\}, \text{ die zu } n \text{ teilerfremd sind})$$

für jedes $n \in \mathbb{N}_+$.

Diese Funktion ϕ hat eine Reihe bekannter Eigenschaften: Es gilt $\phi(1) = 1$, sowie

$$\phi(uv) = \phi(u)\phi(v) \text{ für je zwei teilerfremde Zahlen } u \in \mathbb{N}_+ \text{ und } v \in \mathbb{N}_+; \quad (1)$$

$$\phi(p) = p - 1 \text{ für jede Primzahl } p; \quad (2)$$

$$\phi(p^k) = (p - 1)p^{k-1} = p^k - p^{k-1} \text{ für jede Primzahl } p \text{ und jedes } k \in \mathbb{N}_+; \quad (3)$$

$$\phi(n) = n \cdot \prod_{\substack{p \text{ Primteiler} \\ \text{von } n}} \left(1 - \frac{1}{p}\right) \text{ für jedes } n \in \mathbb{N}_+. \quad (4)$$

Nun behauptet unsere Hauptaussage:

Satz 1.1: Für alle $q \in \mathbb{Z}$ und $n \in \mathbb{N}_+$ ist $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} \in \mathbb{Z}$.

Es ist hilfreich, sich diesen Satz anhand von einfachen Beispielen zu verdeutlichen: Für $n = 1$ hat die Summe $\sum_{d|n} \phi(d) q^{n/d}$ nur einen Summanden (denn $n = 1$ hat nur einen Teiler), und somit ist $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} = \frac{1}{1} \phi(1) q^{1/1} = q$ für $n = 1$. Wir erhalten also $q \in \mathbb{Z}$, was nicht erstaunlich ist, haben wir es doch als Bedingung gefordert. Interessanter wird der Fall $n = p$ für eine Primzahl p ; in diesem Fall ist nämlich

$$\begin{aligned} \frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} &= \frac{1}{p} \left(\phi(1) q^{p/1} + \phi(p) q^{p/p} \right) && \left(\begin{array}{l} \text{denn } n = p \text{ ist eine Primzahl} \\ \text{und hat daher nur die Teiler } 1 \text{ und } p \end{array} \right) \\ &= \frac{1}{p} \left(1q^p + (p-1)q^1 \right) = \frac{1}{p} (q^p + (p-1)q). \end{aligned}$$

Satz 1.1 ergibt also $\frac{1}{p}(q^p + (p-1)q) \in \mathbb{Z}$. Mit anderen Worten: $p \mid q^p + (p-1)q$. Wir können dies auch in der Form $q^p \equiv -(p-1)q \pmod{p}$ umschreiben. Da $-(p-1)q \equiv -(-1)q = q \pmod{p}$ ist, vereinfacht sich dies zu $q^p \equiv q \pmod{p}$. Dies ist eine Formulierung des kleinen Satzes von Fermat. Satz 1.1 verallgemeinert also den kleinen Satz von Fermat.

Die Allgemeinheit kommt jedoch um einen Preis: Je "weniger prim" n ist (also je mehr Primfaktoren in je höheren Potenzen in n vorkommen), desto komplizierter wird der Ausdruck $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d}$. Beispielsweise ist

$$\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} = \frac{1}{p^2} (q^{p^2} + (p-1)q^p + p(p-1)q) \quad \text{für } n = p^2 \text{ mit } p \text{ prim;}$$

$$\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} = \frac{1}{pp'} (q^{pp'} + (p-1)q^{p'} + (p'-1)q^p + (p-1)(p'-1)q)$$

für $n = pp'$ mit zwei verschiedenen Primzahlen p und p' .

Die rechten Seiten dieser beiden Gleichungen sind also nach Satz 1.1 ganzzahlig.

Wir wollen noch eine äquivalente Form von Satz 1.1 formulieren, die ohne den Begriff der Phi-Funktion auskommt:

Satz 1.2: Für alle $q \in \mathbb{Z}$ und $n \in \mathbb{N}_+$ ist $\frac{1}{n} \sum_{i=1}^n q^{\text{ggT}(i,n)} \in \mathbb{Z}$.

Daß dieser Satz 1.2 zu Satz 1.1 äquivalent ist, folgt aus der leicht zu beweisenden Identität $\sum_{i=1}^n q^{\text{ggT}(i,n)} = \sum_{d|n} \phi(d) q^{n/d}$ (der Beweis ist der Sonderfall des Beweises der Äquivalenz $\mathcal{G} \iff \mathcal{H}$ weiter unten, wenn man $(b_1, b_2, b_3, \dots) = (q^1, q^2, q^3, \dots)$ einsetzt).

Wir werden uns nun auf verschiedene Weisen Satz 1.1 nähern. Wer nur an schnellen Beweisen von Satz 1.1 interessiert ist, kann z. B. in [6] mehrere finden (wo eigentlich Satz 1.2 bewiesen wird, der aber, wie wir wissen, zu Satz 1.1 äquivalent ist). Unser Ziel ist es nicht, Satz 1.1 zu beweisen, sondern dabei einige Begriffe aus der Kombinatorik, Algebra und Zahlentheorie kennenzulernen. Unser grober Plan ist folgender:

- In Abschnitt 2 werden wir den ersten Beweis von Satz 1.1 geben (in zwei Varianten), indem wir den Term $q^{n/d}$ in Satz 1.1 durch ein allgemeineres $b_{n/d}$ (wobei (b_1, b_2, b_3, \dots) eine Folge ganzer Zahlen ist) ersetzen, und schauen, wann dann Satz 1.1 noch gilt. Dafür konstruieren wir viele äquivalente Kriterien (die Aussagen $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}$ und \mathcal{H} von Satz 2.1).
- In Abschnitt 3 werden wir den Begriff der formalen Potenzreihen kennenlernen¹. Damit wird ein Teil der verallgemeinerten Version von Satz 1.1 aus Abschnitt 2 (also Satz 2.1) erneut bewiesen.

¹Den ich mit einiger Ausführlichkeit behandle, weil ich feststellen musste, daß es zwar vielfältige Literatur gibt, in der Potenzreihen verwendet und anschaulich erklärt werden, aber nur wenige, in der sie formal sauber eingeführt und ihre wesentlichen Eigenschaften bewiesen werden.

- In Abschnitt 4 zeigen wir zwei merkwürdige Analoga von Satz 1.1: Für beliebige ganze q und r und natürliche n gilt

$$\sum_{d|n} \phi(d) \binom{qn/d}{rn/d} \in n\mathbb{Z} \quad \text{und} \quad \sum_{d|n} \phi(d) \binom{qn/d}{rn/d} \in \frac{q}{r}n\mathbb{Z}.$$

Wir zeigen in Wirklichkeit sogar mehr. (Was Binomialkoeffizienten wie $\binom{qn/d}{rn/d}$ bedeuten, wenn q oder r negativ ist, wird in Abschnitt 4 erklärt.) Als Nebeneffekt erhalten wir zwei zusätzliche äquivalente Kriterien, die man zu Satz 2.1 hinzufügen könnte (die Aussagen \mathcal{I} und \mathcal{J} von Satz 4.7).

- In Abschnitt 5 beweisen wir $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} \in \mathbb{Z}$ erneut, indem wir die Zahl $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} \in \mathbb{Z}$ kombinatorisch deuten (d. h. als eine Anzahl gewisser Objekte darstellen). Dieser Beweis hat einen kleinen Haken - er funktioniert nur im Fall von $q \geq 0$ (dieser Defekt wird in Abschnitt 6 auf zwei Weisen ausgemerzt). Der Ansatz, mit dem sich $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} \in \mathbb{Z}$ als Anzahl darstellen läßt, führt in seiner Verallgemeinerung auf das sogenannte *Pólya-Burnsidesche Abzählverfahren*.
- In Abschnitt 6 ergänzen wir den kombinatorischen Ansatz von Abschnitt 5 zu einem vollständigen Beweis von Satz 2.1. Dies kann auf zwei Arten geschehen; eine davon führt über eine Charakterisierung aller Polynome, die auf ganzen Zahlen ganze Werte liefern (Satz 6.2).
- In Abschnitt 7 wird Satz 2.1 auf sogenannte Nester verallgemeinert und erhält im Falle eines endlichen Nestes einen kombinatorischen Aspekt, der den kombinatorischen Zugang von Abschnitt 5 verallgemeinert. Dadurch wird die zahlentheoretische Verallgemeinerung der Zahlen $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d}$ (Abschnitt 2) mit ihrer kombinatorischen Interpretation (Abschnitt 5) vereinigt. [...]

2. Zahlentheorie I: Die Dirichletfaltung und der Äquivalenzsatz für Geisterwittfolgen

2.1. Der Äquivalenzsatz

Der erste Beweis von Satz 1.1 ist ein zahlentheoretischer. Er fragt sich allgemeiner, welche Folgen $(b_1, b_2, b_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ von ganzen Zahlen die Eigenschaft haben, daß $\frac{1}{n} \sum_{d|n} \phi(d) b_{n/d} \in \mathbb{Z}$ für jedes $n \in \mathbb{N}_+$ ist. Und tatsächlich erweist sich, daß diese Folgen durch eine Fülle weiterer Eigenschaften beschreibbar sind:

Satz 2.1 (der Äquivalenzsatz für Geisterwittfolgen): Sei $(b_1, b_2, b_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ eine Folge ganzer Zahlen. Dann sind folgende Aussagen \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} und \mathcal{H} äquivalent:

Aussage A: Für jede Zahl $n \in \mathbb{N}_+$ und jeden Primteiler p von n gilt

$$b_{n/p} \equiv b_n \pmod{p^{v_p(n)}}.$$

Hierbei ist die Zahl $v_p(n)$ weiter unten definiert.

Aussage B: Es gibt eine Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ ganzer Zahlen, die

$$\left(b_n = \sum_{d|n} dx_d^{n/d} \text{ für jedes } n \in \mathbb{N}_+ \right) \quad (5)$$

erfüllt.

Aussage C: Es gibt *genau eine* Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ ganzer Zahlen, die (5) erfüllt.

Aussage D: Es gibt eine Folge $(y_1, y_2, y_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ ganzer Zahlen, die

$$\left(b_n = \sum_{d|n} dy_d \text{ für jedes } n \in \mathbb{N}_+ \right) \quad (6)$$

erfüllt.

Aussage E: Es gibt *genau eine* Folge $(y_1, y_2, y_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ ganzer Zahlen, die (6) erfüllt.

Aussage F: Für jedes $n \in \mathbb{N}_+$ gilt

$$\sum_{d|n} \mu(d) b_{n/d} \in n\mathbb{Z}, \quad (7)$$

wobei $\mu : \mathbb{N}_+ \rightarrow \mathbb{Z}$ die sogenannte Möbiusfunktion ist (siehe ihre Definition weiter unten).

Aussage G: Für jedes $n \in \mathbb{N}_+$ gilt

$$\sum_{d|n} \phi(d) b_{n/d} \in n\mathbb{Z}. \quad (8)$$

Aussage H: Für jedes $n \in \mathbb{N}_+$ gilt

$$\sum_{i=1}^n b_{\text{ggT}(i,n)} \in n\mathbb{Z}.$$

Wir müssen zwei der in diesem Satz verwendeten Begriffe erst einmal definieren:

Definition (Möbiusfunktion μ): Die *Möbiusfunktion* $\mu : \mathbb{N}_+ \rightarrow \mathbb{Z}$ ist definiert durch

$$\mu(n) = \begin{cases} 1, & \text{wenn } n \text{ ein Produkt von gerade vielen } \textit{paarweise verschiedenen} \text{ Primzahlen ist;} \\ -1, & \text{wenn } n \text{ ein Produkt von ungerade vielen } \textit{paarweise verschiedenen} \text{ Primzahlen ist;} \\ 0, & \text{wenn in der Primfaktorzerlegung von } n \text{ (mindestens) eine Primzahl mehrfach vorkommt} \end{cases}.$$

² Diese Definition mag auf den ersten Blick sehr willkürlich aussehen; allerdings werden wir später sehen, wie natürlich die Funktion μ eigentlich ist.

Definition (p -Wertigkeit): Für jede von 0 verschiedene ganze Zahl n bezeichnen wir mit $v_p(n)$ die größte Zahl $k \in \mathbb{N}$, für die $p^k \mid n$ gilt. Mit anderen Worten: $v_p(n)$ ist die Potenz, in der der Primfaktor p in der Primfaktorzerlegung von n vorkommt. Die Zahl $v_p(n)$ wird oft als p -Wertigkeit (oder p -adische Wertigkeit) der Zahl n bezeichnet.

Man setzt oft auch $v_p(0) = \infty$, um dem Ausdruck $v_p(n)$ für jede ganze Zahl n (und nicht nur für jede von 0 verschiedene ganze Zahl n) Sinn zu verleihen.³

Bemerkung: Der Name von Satz 2.1 rührt daher, daß eine Folge $(b_1, b_2, b_3, \dots) \in \mathbb{Z}^{\mathbb{N}^+}$, die die äquivalenten Aussagen $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}$ und \mathcal{H} erfüllt, als *Geisterwittfolge* (oder *Nebenwittfolge*) bezeichnet wird (zumindest in [1]).⁴

2.2. Aus dem Äquivalenzsatz folgt Satz 1.1

Um Satz 1.1 aus Satz 2.1 zu erhalten, reicht es aus, *eine* der Aussagen $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}$ und \mathcal{H} für die Folge $(b_1, b_2, b_3, \dots) = (q^1, q^2, q^3, \dots)$ nachzuweisen (denn dann ergeben sich wegen Satz 2.1 sofort alle anderen Aussagen, darunter insbesondere Aussage \mathcal{G} , welche sofort Satz 1.1 ergibt). Am einfachsten läßt sich dies mit Aussage \mathcal{B} erledigen; denn für die Folge $(b_1, b_2, b_3, \dots) = (q^1, q^2, q^3, \dots)$ gibt es offensichtlich eine Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}^+}$, die (5) erfüllt (nämlich die Folge $(x_1, x_2, x_3, \dots) = (q, 0, 0, 0, \dots)$ mit $x_1 = q$ und $x_i = 0$ für alle $i > 1$ ⁵). Man kann alternativ auch Aussage \mathcal{A} nachprüfen:

Lemma 2.2: Seien $q \in \mathbb{Z}$ und $n \in \mathbb{N}_+$. Sei p ein Primfaktor von n . Dann gilt $q^{n/p} \equiv q^n \pmod{p^{v_p(n)}}$.

Anstatt dieses Lemma direkt zu zeigen, beweisen wir etwas allgemeineres (und nützlicheres):

²Insbesondere bedeutet dies $\mu(1) = 1$, denn 1 ist das Produkt von null Primzahlen, und null ist gerade.

³Man sollte dabei natürlich bedenken, daß ∞ ein Ausdruck ist, der gewissen Regeln gehorcht (wie etwa $n + \infty = \infty$ für alle $n \in \mathbb{N}$), mit dem man aber nicht alles machen kann, was man mit ganzen Zahlen kann (so kann man natürlich nicht ∞ von der Gleichung $n + \infty = \infty$ subtrahieren, um $n = 0$ zu erhalten). Im Zweifel sollte man sich immer bewusst sein, ob die Umformungen, die man mit ∞ vollzieht, auch ohne die Verwendung des Zeichens ∞ sinnvoll umgeschrieben werden könnten.

⁴Denn Aussage \mathcal{B} läßt sich umschreiben als "Die Folge (b_1, b_2, b_3, \dots) ist die Folge der Geisterkomponenten (a. k. a. Nebenkomponten) eines ganzzahligen (großen a. k. a. universellen) Wittvektors". (Der Begriff eines Wittvektors wird in [1] erklärt.)

⁵Denn für diese Folge $(x_1, x_2, x_3, \dots) = (q, 0, 0, 0, \dots)$ ist

$$\sum_{d|n} dx_d^{n/d} = \sum_{d \in \mathbb{N}_n} dx_d^{n/d} = \underbrace{1x_1^{n/1}}_{=x_1^n=q^n} + \sum_{\substack{d \in \mathbb{N}_n; \\ d \neq 1}} d \underbrace{x_d^{n/d}}_{=0 \text{ (da } x_d=0 \text{ für } d \neq 1)} = q^n + \underbrace{\sum_{\substack{d \in \mathbb{N}_n; \\ d \neq 1}} d \cdot 0}_{=0} = q^n$$

für jedes $n \in \mathbb{N}_+$.

Lemma 2.3: Seien $p \in \mathbb{Z}$, $u \in \mathbb{Z}$ und $v \in \mathbb{Z}$. Seien $k \in \mathbb{N}_+$ und $\ell \in \mathbb{N}$.
Wenn $u \equiv v \pmod{p^k}$ ist, dann ist $u^{p^\ell} \equiv v^{p^\ell} \pmod{p^{k+\ell}}$.

Dieses Lemma, oft als *kleines Henselsches Lemma* bezeichnet, ist eines der wichtigsten Hilfsmittel in der Zahlentheorie, um Resultate über Restklassen modulo Primpotenzen zu erreichen. Unter anderem beweist man mit seiner Hilfe, daß es Primitivwurzeln modulo p^k für primes $p > 2$ und $k \in \mathbb{N}_+$ gibt⁶.

(Kleine *Anmerkung* am Rande: Für Lemma 2.3 braucht p nicht prim zu sein, auch wenn es meist auf primes p angewendet wird.)

Beweis von Lemma 2.3: Wir beweisen Lemma 2.3 durch Induktion nach ℓ . Für $\ell = 0$ ist nichts zu beweisen. Also zum Induktionsschritt: Angenommen, für irgendein $\ell \in \mathbb{N}$ haben wir bereits $u^{p^\ell} \equiv v^{p^\ell} \pmod{p^{k+\ell}}$ gezeigt, und jetzt ist es an der Zeit, $u^{p^{\ell+1}} \equiv v^{p^{\ell+1}} \pmod{p^{k+\ell+1}}$ zu beweisen.

Aus $u \equiv v \pmod{p^k}$ folgt insbesondere $u \equiv v \pmod{p}$ (denn $k \in \mathbb{N}_+$ und damit $p \mid p^k$). Nun ist

$$u^{p^{\ell+1}} - v^{p^{\ell+1}} = (u^{p^\ell})^p - (v^{p^\ell})^p = (u^{p^\ell} - v^{p^\ell}) \cdot \sum_{i=0}^{p-1} (u^{p^\ell})^i (v^{p^\ell})^{p-1-i}$$

(nach der dritten binomischen Formel). Der erste Faktor dieses Produktes ist durch $p^{k+\ell}$ teilbar (denn $u^{p^\ell} \equiv v^{p^\ell} \pmod{p^{k+\ell}}$ ergibt $p^{k+\ell} \mid u^{p^\ell} - v^{p^\ell}$). Der zweite Faktor ist durch p teilbar (denn aus $u \equiv v \pmod{p}$ folgt

$$\sum_{i=0}^{p-1} (u^{p^\ell})^i (v^{p^\ell})^{p-1-i} \equiv \sum_{i=0}^{p-1} (v^{p^\ell})^i (v^{p^\ell})^{p-1-i} = \sum_{i=0}^{p-1} (v^{p^\ell})^{p-1} = p (v^{p^\ell})^{p-1} \equiv 0 \pmod{p}$$

). Also ist das Produkt durch $p^{k+\ell} \cdot p = p^{k+\ell+1}$ teilbar. Das heißt, $u^{p^{\ell+1}} - v^{p^{\ell+1}}$ ist durch $p^{k+\ell+1}$ teilbar. Mit anderen Worten: $u^{p^{\ell+1}} \equiv v^{p^{\ell+1}} \pmod{p^{k+\ell+1}}$. Damit ist der Induktionsschritt abgeschlossen, und Lemma 2.3 bewiesen.

Aus Lemma 2.3 folgt nun schnell Lemma 2.2:

Beweis von Lemma 2.2: Da p ein Primfaktor von n ist, ist $v_p(n) \geq 1$. Sei nun $\ell = v_p(n) - 1$; dann ist also $\ell \geq 0$ und damit $\ell \in \mathbb{N}$. Seien ferner $k = 1$, $u = q$ und $v = q^p$. Dann ist $u \equiv v \pmod{p^k}$ (denn dies bedeutet nichts anderes als $q \equiv q^p \pmod{p^1}$, was aus dem kleinen Satz von Fermat folgt). Nach Lemma 2.3 ist also $u^{p^\ell} \equiv v^{p^\ell} \pmod{p^{k+\ell}}$. Mit anderen Worten: $q^{p^\ell} \equiv (q^p)^{p^\ell} \pmod{p^{k+\ell}}$. Wegen $k + \ell = 1 + (v_p(n) - 1) = v_p(n)$ wird dies zu $q^{p^\ell} \equiv (q^p)^{p^\ell} \pmod{p^{v_p(n)}}$.

Wir wollen aber zeigen, daß $q^{n/p} \equiv q^n \pmod{p^{v_p(n)}}$ ist. Dies ist nicht mehr weit: wir werden zeigen, daß $q^{n/p}$ eine Potenz von q^{p^ℓ} , und q^n dieselbe Potenz von $(q^p)^{p^\ell}$ ist.

In der Tat gibt es ein $m \in \mathbb{N}_+$ mit $n = mp^{v_p(n)}$ (denn nach der Definition von $v_p(n)$ ist $p^{v_p(n)} \mid n$). Somit ist $n/p = mp^{v_p(n)-1} = mp^\ell$ (da $v_p(n) - 1 = \ell$). Also ist $q^{n/p} = q^{mp^\ell} = (q^{p^\ell})^m$ und $q^n = (q^p)^{n/p} = (q^p)^{mp^\ell} = ((q^p)^{p^\ell})^m$. Mithilfe von $q^{p^\ell} \equiv (q^p)^{p^\ell} \pmod{p^{v_p(n)}}$ ergibt sich nun $q^{n/p} = (q^{p^\ell})^m \equiv ((q^p)^{p^\ell})^m = q^n \pmod{p^{v_p(n)}}$, und Lemma 2.2 ist bewiesen.

Und schließlich zwei Beweise von Satz 1.1 unter Berufung auf Satz 2.1:

Erster Beweis von Satz 1.1: Wenden wir Satz 1.1 auf die Folge $(b_1, b_2, b_3, \dots) = (q^1, q^2, q^3, \dots)$ an, dann erhalten wir, daß die Aussagen \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} und \mathcal{H}

⁶sobald man erst einmal gezeigt hat, daß es Primitivwurzeln modulo p gibt

für diese Folge äquivalent sind. Da die Aussage \mathcal{B} für diese Folge erfüllt ist (dies folgt aus Lemma 2.2), ist also auch die Aussage \mathcal{G} für diese Folge erfüllt. Das heißt, $\sum_{d|n} \phi(d) b_{n/d} \in n\mathbb{Z}$ für jedes $n \in \mathbb{N}_+$. Mit anderen Worten: $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} \in \mathbb{Z}$ für jedes $n \in \mathbb{N}_+$. Doch dies heißt nichts anderes, als daß Satz 1.1 gilt.

Zweiter Beweis von Satz 1.1: Wenden wir Satz 1.1 auf die Folge $(b_1, b_2, b_3, \dots) = (q^1, q^2, q^3, \dots)$ an, dann erhalten wir, daß die Aussagen $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}$ und \mathcal{H} für diese Folge äquivalent sind. Da die Aussage \mathcal{A} für diese Folge erfüllt ist (wie wir bereits festgestellt haben), ist also auch die Aussage \mathcal{G} für diese Folge erfüllt. Wie im Ersten Beweis von Satz 1.1 folgern wir hieraus, daß Satz 1.1 gilt.

2.3. Beweis von Satz 2.1: Einleitung

Jetzt haben wir bereits zwei Beweise von Satz 1.1, aber zu beiden fehlt noch eine Herleitung von Satz 2.1. Es ist auf verschiedene Arten möglich, Satz 2.1 zu beweisen; auch verallgemeinern läßt er sich mehrfach. Wir werden ihn später (in Abschnitt 4.4) auch noch um zwei äquivalente Aussagen erweitern, und danach (Abschnitt 7) auch noch weiter ausdehnen. Ein Beweis von Satz 2.1, sogar mehrfach verallgemeinert, ist in [3] zu finden.⁷ Wir wollen hier einen (in Teilen) anderen Beweis geben.

Zum Beweis von Satz 1.1 nötig ist allerdings nur die Äquivalenz $\mathcal{A} \iff \mathcal{G}$ bzw. die Äquivalenz $\mathcal{B} \iff \mathcal{G}$ (je nachdem, ob wir den Ersten oder den Zweiten Beweis nehmen wollen); die Aussagen \mathcal{D} und \mathcal{F} sind ihrerseits als Zwischenschritte zum Beweis dieser Äquivalenz förderlich. Die Aussagen \mathcal{C}, \mathcal{E} und \mathcal{H} könnten wir getrost weglassen, wenn es uns nur um den Beweis von Satz 1.1 ginge - aber der Vollständigkeit halber haben wir sie hier aufgeführt und werden sie auch in den Beweis einbeziehen.

Übrigens ist Aussage \mathcal{D} (bzw., je nach Lesart, \mathcal{E}) für die Folge $(b_1, b_2, b_3, \dots) = (q^1, q^2, q^3, \dots)$ auch recht bekannt - beispielsweise kam sie (in etwas anderer Formulierung) in [6] als Aufgabe 48 in Kapitel 5 vor (allerdings nur im Sonderfall $q = 2$, wobei die Lösung auch für jedes $q \geq 0$ durchgeht, jedoch nicht für $q < 0$; wir werden uns diesen Lösungsweg noch in Abschnitt 5 genauer ansehen).

2.4. Die Dirichletfaltung

Zum Beweis von Satz 2.1 wollen wir erstmal eine Notation einführen. Mehrere der Aussagen von Satz 2.1 enthalten Summen, in denen über Teiler d von n summiert wird, und in der Summe kommen d und $\frac{n}{d}$ vor. Einige solche Summen kennt man in der Zahlentheorie unter dem Begriff *Dirichlet-Faltungen*. Diese Dirichlet-Faltungen sind für die elementare und die analytische Zahlentheorie recht wichtig und kommen in den meisten Büchern darüber (z. B. [5], Kapitel 2, Abschnitt 3) vor. Bevor wir definieren, was eine Dirichlet-Faltung ist, vereinbaren wir eine Notation:

Definition (Zahlenfunktion, ε , id und $\underline{1}$): Unter einer *Zahlenfunktion* verstehen wir im Folgenden eine Funktion von \mathbb{N}_+ nach \mathbb{Z} . Insbesondere

⁷Und zwar erhält man Satz 2.1, indem man man Theorem 15 in [3] auf den Fall $N = \mathbb{N}_+$ anwendet. Man beachte, daß die Aussagen, die wir hier mit $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}$ und \mathcal{H} bezeichnet haben, in [3] auf die Namen $\mathcal{C}_\emptyset, \mathcal{D}_\emptyset, \mathcal{D}'_\emptyset, \mathcal{E}_\emptyset, \mathcal{E}'_\emptyset, \mathcal{F}_\emptyset, \mathcal{G}_\emptyset$ bzw. \mathcal{H}_\emptyset hören.

sind ϕ und μ zwei Zahlenfunktionen. Wir definieren eine weitere Zahlenfunktion $\varepsilon : \mathbb{N}_+ \rightarrow \mathbb{Z}$ durch

$$\left(\varepsilon(n) = \begin{cases} 1, & \text{wenn } n = 1; \\ 0, & \text{wenn } n > 1 \end{cases} \quad \text{für jedes } n \in \mathbb{N}_+ \right).$$

Eine weitere Zahlenfunktion ist $\text{id} : \mathbb{N}_+ \rightarrow \mathbb{N}_+$, definiert durch ($\text{id}(n) = n$ für alle $n \in \mathbb{N}_+$). Eine andere Zahlenfunktion ist $\underline{1} : \mathbb{N}_+ \rightarrow \{1\}$, definiert durch ($\underline{1}(n) = 1$ für alle $n \in \mathbb{N}_+$).

Definition (Gleichsetzung von Folgen mit Zahlenfunktionen): Für jede unendliche Folge (a_1, a_2, a_3, \dots) ganzer Zahlen können wir eine Zahlenfunktion $a : \mathbb{N}_+ \rightarrow \mathbb{Z}$ definieren durch

$$(a(n) = a_n \text{ für jedes } n \in \mathbb{N}_+).$$

Umgekehrt entsteht jede Zahlenfunktion $a : \mathbb{N}_+ \rightarrow \mathbb{Z}$ aus genau einer unendlichen Folge (a_1, a_2, a_3, \dots) auf diese Weise. Somit können wir unendliche Folgen ganzer Zahlen, deren Folgenglieder mit $1, 2, 3, \dots$ indiziert sind, gleichsetzen mit Zahlenfunktionen. Die Zahlenfunktion ε ist auf diese Weise gleichgesetzt mit der Folge $(1, 0, 0, 0, \dots)$ (mit einer Eins am Anfang und sonst nur Nullen). Die Zahlenfunktion id entspricht der Folge $(1, 2, 3, 4, \dots)$. Die Zahlenfunktion $\underline{1}$ entspricht der Folge $(1, 1, 1, 1, \dots)$. Auch die Folge (b_1, b_2, b_3, \dots) aus Satz 2.1 entspricht einer Zahlenfunktion; wir werden im Beweis von Satz 2.1 diese Zahlenfunktion mit b bezeichnen (und ebenso die anderen vorkommenden Folgen durch die entsprechenden Zahlenfunktionen ersetzen).

Definition (Dirichlet-Faltung): Nun ist die *Dirichlet-Faltung* zweier Zahlenfunktionen f und g eine neue Zahlenfunktion, genannt $f * g$ und definiert durch

$$\left((f * g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right) \quad \text{für alle } n \in \mathbb{N}_+ \right). \quad (9)$$

Diese Definition läßt sich auch folgendermaßen "symmetrischer" hinschreiben:

$$\left((f * g)(n) = \sum_{\substack{(d,e) \in \mathbb{N}_+^2; \\ de=n}} f(d) g(e) \quad \text{für alle } n \in \mathbb{N}_+ \right). \quad (10)$$

Aus dieser "symmetrischen" Definition (10) folgt sofort $f * g = g * f$ für zwei beliebige Zahlenfunktionen f und g ⁸. Das heißt, die Dirichlet-Faltung $*$ ist kommutativ.

⁸denn für jedes $n \in \mathbb{N}_+$ ist

$$\begin{aligned} (f * g)(n) &= \sum_{\substack{(d,e) \in \mathbb{N}_+^2; \\ de=n}} f(d) g(e) = \sum_{\substack{(d,e) \in \mathbb{N}_+^2; \\ de=n}} g(e) f(d) \\ &= \sum_{\substack{(e,d) \in \mathbb{N}_+^2; \\ ed=n}} g(d) f(e) \quad (\text{hier haben wir } (d, e) \text{ in } (e, d) \text{ umbenannt}) \\ &= \sum_{\substack{(d,e) \in \mathbb{N}_+^2; \\ de=n}} g(d) f(e) = (g * f)(n) \end{aligned}$$

Ferner sieht man leicht ein, daß $*$ assoziativ ist, d. h. daß $(f * g) * h = f * (g * h)$ für je drei Zahlenfunktionen f, g und h gilt⁹. Folglich können wir von der Zahlenfunktion $f * g * h$ sprechen, ohne Klammern setzen zu müssen (gemeint ist damit natürlich die Zahlenfunktion $(f * g) * h$, die gleichzeitig auch die Zahlenfunktion $f * (g * h)$ ist). Schließlich gilt für jede Zahlenfunktion f die Gleichung $\varepsilon * f = f * \varepsilon = f$, wobei ε die Zahlenfolge $(1, 0, 0, 0, \dots)$ (mit $\varepsilon_1 = 1$ und $\varepsilon_i = 0$ für alle $i > 1$) ist¹⁰. Dies alles bedeutet, daß die Zahlenfunktionen mit der Verknüpfung $*$ und dem neutralen Element ε ein abelsches Monoid bilden; das heißt, man kann die Dirichlet-Faltung $*$ als eine Art Multiplikation von Zahlenfunktionen deuten, und die Zahlenfolge ε als eine Art Eins.

Nun stellt sich heraus, daß die Zahlenfunktionen $\phi, \mu, \varepsilon, \underline{1}$ und id unter der Verknüpfung $*$ sich auf diverse Weisen interessant verhalten:

Satz 2.4: (a) Es gilt $\phi * \underline{1} = \text{id}$. Mit anderen Worten: Für jedes $n \in \mathbb{N}_+$

⁹Beweis: Für jedes $n \in \mathbb{N}_+$ ist

$$\begin{aligned}
((f * g) * h)(n) &= \sum_{\substack{(d,e) \in \mathbb{N}_+^2; \\ de=n}} (f * g)(d) h(e) && \text{(nach (10), angewendet auf } f * g \text{ und } h \text{ anstelle von } f \text{ bzw. } g) \\
&= \sum_{\substack{(d,e) \in \mathbb{N}_+^2; \\ de=n}} \sum_{\substack{(r,s) \in \mathbb{N}_+^2; \\ rs=d}} f(r) g(s) h(e) \\
&\quad \left(\begin{array}{l} \text{denn für jedes } d \text{ ist } (f * g)(d) = \sum_{\substack{(r,s) \in \mathbb{N}_+^2; \\ rs=d}} f(r) g(s) \text{ nach (10) (angewandt auf } d \text{ statt } n), \\ \text{wobei die Indizes } d \text{ und } e \text{ aus der Gleichung (10) hier in } r \text{ bzw. } s \text{ umbenannt wurden} \end{array} \right) \\
&= \sum_{\substack{d \in \mathbb{N}_+ \\ de=n}} \sum_{\substack{e \in \mathbb{N}_+ \\ de=n}} \sum_{\substack{r \in \mathbb{N}_+ \\ rs=d}} \sum_{\substack{s \in \mathbb{N}_+ \\ rs=d}} f(r) g(s) h(e) = \sum_{r \in \mathbb{N}_+} \sum_{s \in \mathbb{N}_+} \sum_{\substack{d \in \mathbb{N}_+; \\ rs=d}} \sum_{\substack{e \in \mathbb{N}_+; \\ de=n}} f(r) g(s) h(e) = \sum_{r \in \mathbb{N}_+} \sum_{\substack{s \in \mathbb{N}_+ \\ rse=n}} \sum_{\substack{e \in \mathbb{N}_+; \\ rse=n}} f(r) g(s) h(e) \\
&\quad \left(\begin{array}{l} \text{denn } \sum_{\substack{d \in \mathbb{N}_+; \\ rs=d}} \sum_{\substack{e \in \mathbb{N}_+; \\ de=n}} f(r) g(s) h(e) = \sum_{\substack{e \in \mathbb{N}_+; \\ rse=n}} f(r) g(s) h(e), \text{ weil eine Summe der Form } \sum_{\substack{d \in \mathbb{N}_+; \\ rs=d}} \dots \\ \text{immer nur aus einem Summanden besteht (nämlich aus dem Summanden für } d = rs) \end{array} \right) \\
&= \sum_{\substack{(r,s,e) \in \mathbb{N}_+^3; \\ rse=n}} f(r) g(s) h(e)
\end{aligned}$$

und analog

$$(f * (g * h))(n) = \sum_{\substack{(d,r,s) \in \mathbb{N}_+^3; \\ drs=n}} f(d) g(r) h(s).$$

Da sich die Summen $\sum_{\substack{(r,s,e) \in \mathbb{N}_+^3; \\ rse=n}} f(r) g(s) h(e)$ und $\sum_{\substack{(d,r,s) \in \mathbb{N}_+^3; \\ drs=n}} f(d) g(r) h(s)$ nur in der Benennung ihrer Indizes unterscheiden, haben sie den gleichen Wert, und somit ist $((f * g) * h)(n) = (f * (g * h))(n)$ für alle n , also $(f * g) * h = f * (g * h)$, was zu beweisen war.

¹⁰denn für alle $n \in \mathbb{N}_+$ ist

$$\begin{aligned}
(f * \varepsilon)(n) &= \sum_{\substack{(d,e) \in \mathbb{N}_+^2; \\ de=n}} f(d) \varepsilon(e) && \text{(nach (10))} \\
&= f(n) \underbrace{\varepsilon(1)}_{=1} + \sum_{\substack{(d,e) \in \mathbb{N}_+^2; \\ de=n; e>1}} f(d) \underbrace{\varepsilon(e)}_{=0 \text{ (da } e>1)} = f(n) \cdot 1 + \sum_{\substack{(d,e) \in \mathbb{N}_+^2; \\ de=n; e>1}} f(d) \cdot 0 = f(n) + 0 = f(n)
\end{aligned}$$

und analog $(\varepsilon * f)(n) = f(n)$

ist

$$\sum_{d|n} \phi(d) = n.$$

(b) Es gilt $\mu * \underline{1} = \varepsilon$. Mit anderen Worten: Für jedes $n \in \mathbb{N}_+$ ist

$$\sum_{d|n} \mu(d) = \varepsilon(n) = \begin{cases} 1, & \text{wenn } n = 1; \\ 0, & \text{wenn } n > 1 \end{cases}.$$

(c) Es gilt $\mu * \text{id} = \phi$. Mit anderen Worten: Für jedes $n \in \mathbb{N}_+$ ist

$$\sum_{d|n} \mu(d) \frac{n}{d} = \phi(n).$$

Dies sind nur 3 von den $\frac{5(5+1)}{2} = 15$ möglichen Dirichlet-Faltungen, die man aus den fünf Zahlenfunktionen $\phi, \mu, \varepsilon, \underline{1}$ und id bilden kann. Natürlich sind die 5 Faltungen, die man mit ε bilden kann, trivial (denn $\varepsilon * f = f * \varepsilon = f$ für jede Zahlenfunktion f). Über einige der anderen $15 - 3 - 5 = 7$ Faltungen läßt sich auch etwas aussagen. So ist $\phi * \text{id}$ die Funktion, die jedes $n \in \mathbb{N}_+$ in $\sum_{k=1}^n \text{ggT}(k, n)$ überführt (Übungsaufgabe!); diese Funktion war Gegenstand einer für die IMO 2004 vorgeschlagenen Aufgabe (siehe [4]). Die Funktion $\underline{1} * \underline{1}$ ordnet jedem $n \in \mathbb{N}_+$ die Anzahl der Teiler von n zu. Die Funktion $\underline{1} * \text{id}$ ordnet jedem $n \in \mathbb{N}_+$ die Summe der Teiler von n zu.

Wir wenden uns aber nun dem *Beweis von Satz 2.4* zu:

(a) Sei $n \in \mathbb{N}_+$. Für jeden Teiler d von n sei V_d die Menge aller Zahlen $i \in \{1, 2, \dots, \frac{n}{d}\}$, die zu $\frac{n}{d}$ teilerfremd sind. Die Mächtigkeit $|V_d|$ dieser Menge V_d ist dann die Anzahl aller solchen Zahlen, also $\phi\left(\frac{n}{d}\right)$. Andererseits gibt es eine Bijektion

$$V_d \rightarrow \{j \in \{1, 2, \dots, n\} \mid \text{ggT}(j, n) = d\}, \quad \text{gegeben durch } i \mapsto id.$$

¹¹. Somit ist $|\{j \in \{1, 2, \dots, n\} \mid \text{ggT}(j, n) = d\}| = |V_d|$. Wegen $|V_d| = \phi\left(\frac{n}{d}\right)$ wird dies zu

$$|\{j \in \{1, 2, \dots, n\} \mid \text{ggT}(j, n) = d\}| = \phi\left(\frac{n}{d}\right). \quad (11)$$

Andererseits ist die Menge $\{1, 2, \dots, n\}$ die Vereinigung der paarweise disjunkten Mengen $\{j \in \{1, 2, \dots, n\} \mid \text{ggT}(j, n) = d\}$ für $d \in \mathbb{N}_{|n}$ ¹². Also ist

$$|\{1, 2, \dots, n\}| = \sum_{d \in \mathbb{N}_{|n}} |\{j \in \{1, 2, \dots, n\} \mid \text{ggT}(j, n) = d\}| = \sum_{d \in \mathbb{N}_{|n}} \phi\left(\frac{n}{d}\right) \quad (\text{nach (11)}).$$

¹¹Denn einerseits liegt für jedes $i \in V_d$ die Zahl id in der Menge $\{1, 2, \dots, n\}$ und erfüllt $\text{ggT}(id, n) = d$ (denn $\text{ggT}(id, n) = \text{ggT}\left(id, \frac{n}{d}\right) = \text{ggT}\left(i, \frac{n}{d}\right) d = d$, weil $\text{ggT}\left(i, \frac{n}{d}\right) = 1$, denn $i \in V_d$), andererseits aber hat jedes Element j der Menge $\{j \in \{1, 2, \dots, n\} \mid \text{ggT}(j, n) = d\}$ die Form id für ein $i \in V_d$ (nämlich für $i = \frac{j}{d}$).

¹²Denn jedes Element $x \in \{1, 2, \dots, n\}$ liegt in der Menge $\{j \in \{1, 2, \dots, n\} \mid \text{ggT}(j, n) = d\}$ für genau ein $d \in \mathbb{N}_{|n}$, nämlich für $d = \text{ggT}(x, n)$.

Wegen $|\{1, 2, \dots, n\}| = n$ wird dies zu $n = \sum_{d \in \mathbb{N}_{|n}} \phi\left(\frac{n}{d}\right)$. Also ist

$$\begin{aligned} (\phi * \underline{1})(n) &= (\underline{1} * \phi)(n) && \text{(da } \phi * \underline{1} = \underline{1} * \phi, \text{ wegen der Kommutativitat von } *) \\ &= \sum_{d|n} \underbrace{\underline{1}(d)}_{=1} \phi\left(\frac{n}{d}\right) = \sum_{d|n} \phi\left(\frac{n}{d}\right) = \sum_{d \in \mathbb{N}_{|n}} \phi\left(\frac{n}{d}\right) = n = \text{id}(n). \end{aligned}$$

Da dies fur alle $n \in \mathbb{N}_+$ gilt, haben wir also gezeigt, da $\phi * \underline{1} = \text{id}$ gilt. Mit anderen Worten: Fur jedes $n \in \mathbb{N}_+$ ist $\sum_{d|n} \phi(d) = n$ (denn $(\phi * \underline{1})(n) = \sum_{d|n} \phi(d) \underbrace{\underline{1}\left(\frac{n}{d}\right)}_{=1} = \sum_{d|n} \phi(d)$

und $\text{id}(n) = n$). Damit ist Satz 2.4 **(a)** gezeigt.

(b) Sei $\text{PF } n$ die Menge aller Primteiler von n . Offensichtlich ist $\text{PF } n$ eine endliche Menge. Wir bezeichnen mit $\mathcal{P}(\text{PF } n)$ ihre Potenzmenge. Wir definieren eine Abbildung

$$L : \mathcal{P}(\text{PF } n) \rightarrow \mathbb{N}_{|n} \quad \text{durch} \quad L(S) = \prod_{p \in S} p \text{ fur jede Teilmenge } S \text{ von } \text{PF } n.$$

Diese Abbildung ist wohldefiniert¹³ und injektiv¹⁴.

Andererseits ist

$$\sum_{d|n} \mu(d) = \sum_{d \in \mathbb{N}_{|n}} \mu(d) = \sum_{\substack{d \in \mathbb{N}_{|n}; \\ \mu(d) \neq 0}} \mu(d) + \sum_{\substack{d \in \mathbb{N}_{|n}; \\ \mu(d) = 0}} \underbrace{\mu(d)}_{=0} = \sum_{\substack{d \in \mathbb{N}_{|n}; \\ \mu(d) \neq 0}} \mu(d) + \sum_{\substack{d \in \mathbb{N}_{|n}; \\ \mu(d) = 0}} 0 = \sum_{\substack{d \in \mathbb{N}_{|n}; \\ \mu(d) \neq 0}} \mu(d). \quad (12)$$

Nun erfullt ein Teiler d von n aber genau dann $\mu(d) \neq 0$, wenn d ein Produkt von paarweise verschiedenen Primzahlen ist (nach der Definition von μ), d. h. wenn d ein Produkt von paarweise verschiedenen Elementen der Menge $\text{PF } n$ ist (denn Primfaktoren, die nicht in $\text{PF } n$ liegen, kann d nicht haben, weil d ein Teiler von n ist), d. h. wenn $d = \prod_{p \in S} p$ fur eine Teilmenge S von $\text{PF } n$ ist, d. h. wenn $d = L(S)$ fur ein $S \in \mathcal{P}(\text{PF } n)$ ist, d. h. wenn $d \in L(\mathcal{P}(\text{PF } n))$ ist.

Wir haben also

$$\sum_{\substack{d \in \mathbb{N}_{|n}; \\ \mu(d) \neq 0}} \mu(d) = \sum_{d \in L(\mathcal{P}(\text{PF } n))} \mu(d) = \sum_{S \in \mathcal{P}(\text{PF } n)} \mu(L(S)) \quad \text{(weil } L \text{ injektiv ist).}$$

Zusammen mit (12) fuhrt dies auf

$$\sum_{d|n} \mu(d) = \sum_{S \in \mathcal{P}(\text{PF } n)} \mu(L(S)).$$

¹³Denn fur jede Menge $S \in \mathcal{P}(\text{PF } n)$ ist $\prod_{p \in S} p \in \mathbb{N}_{|n}$, denn jede der paarweise verschiedenen

Primzahlen $p \in S$ ist ein Primteiler von n , und somit ist auch ihr Produkt $\prod_{p \in S} p$ ein Teiler von n .

¹⁴Denn waren S und S' zwei verschiedene Teilmengen von $\text{PF } n$ mit $L(S) = L(S')$, dann ware $\prod_{p \in S} p = L(S) = L(S') = \prod_{p \in S'} p$ eine naturliche Zahl mit zwei verschiedenen Primfaktorzerlegungen, was nicht moglich ist.

Doch für jede Menge $S \in \mathcal{P}(\text{PF } n)$ ist $\mu(L(S)) = (-1)^{|S|}$ ¹⁵. Somit ist

$$\begin{aligned}
\sum_{d|n} \mu(d) &= \sum_{S \in \mathcal{P}(\text{PF } n)} \underbrace{\mu(L(S))}_{=(-1)^{|S|}} = \sum_{S \in \mathcal{P}(\text{PF } n)} (-1)^{|S|} = \sum_{k=0}^{|\text{PF } n|} \sum_{\substack{S \in \mathcal{P}(\text{PF } n); \\ |S|=k}} \underbrace{(-1)^{|S|}}_{=(-1)^k} \\
&= \sum_{k=0}^{|\text{PF } n|} \underbrace{\sum_{\substack{S \in \mathcal{P}(\text{PF } n); \\ |S|=k}} (-1)^k}_{=|\{S \in \mathcal{P}(\text{PF } n) \mid |S|=k\}| \cdot (-1)^k} = \sum_{k=0}^{|\text{PF } n|} \underbrace{|\{S \in \mathcal{P}(\text{PF } n) \mid |S|=k\}|}_{=\binom{|\text{PF } n|}{k}} \cdot (-1)^k \\
&= \sum_{k=0}^{|\text{PF } n|} \binom{|\text{PF } n|}{k} \cdot (-1)^k = (1 + (-1))^{|\text{PF } n|} \quad (\text{nach der binomischen Formel}) \\
&= 0^{|\text{PF } n|} = \begin{cases} 1, & \text{wenn } |\text{PF } n| = 0; \\ 0, & \text{wenn } |\text{PF } n| > 0 \end{cases} \\
&= \begin{cases} 1, & \text{wenn } n = 1; \\ 0, & \text{wenn } n > 1 \end{cases} \quad \left(\begin{array}{l} \text{denn } |\text{PF } n| = 0 \text{ gilt genau dann, wenn } n = 1, \\ \text{und sonst gilt } |\text{PF } n| > 0 \end{array} \right) \\
&= \varepsilon(n).
\end{aligned}$$

Wegen $(\mu * \underline{1})(n) = \sum_{d|n} \mu(d) \underbrace{\underline{1}\left(\frac{n}{d}\right)}_{=1} = \sum_{d|n} \mu(d)$ ist dies äquivalent zu $(\mu * \underline{1})(n) = \varepsilon(n)$.

Da dies für jedes $n \in \mathbb{N}_+$ gilt, ist also $\mu * \underline{1} = \varepsilon$, und der Beweis von Satz 2.4 **(b)** ist komplett.

(c) Wir könnten Satz 2.4 **(c)** "zu Fuß" beweisen, indem wir $\sum_{d|n} \mu(d) \frac{n}{d} = \phi(n)$ durch kombinatorische Überlegungen nachweisen (diesen Beweis habe ich in [3], im Beweis von Theorem 6 durchgezogen; der Leser kann ihn sich auch als Übungsaufgabe überlegen). Hier wollen wir aber einen anderen Beweis von Satz 2.4 **(c)** zeigen; und zwar wollen wir Satz 2.4 **(c)** in wenigen Zeilen aus dem bereits bewiesenen herleiten: Nach Satz 2.4 **(a)** ist $\text{id} = \phi * \underline{1} = \underline{1} * \phi$ (wegen der Kommutativität von $*$). Also ist

$$\begin{aligned}
\mu * \text{id} &= \mu * (\underline{1} * \phi) = (\mu * \underline{1}) * \phi && (\text{nach der Assoziativität von } *) \\
&= \varepsilon * \phi && (\text{da } \mu * \underline{1} = \varepsilon \text{ nach Satz 2.4 (b)}) \\
&= \phi,
\end{aligned}$$

und Satz 2.4 **(c)** ist bewiesen. (Auswertung von $\mu * \text{id} = \phi$ an der Stelle n führt auf $\sum_{d|n} \mu(d) \frac{n}{d} = \phi(n)$.) Dieser Beweis zeigt, wie man Aussagen über die Dirichlet-Faltung recht schnell beweisen kann, sobald man einige grundlegende Eigenschaften der Dirichlet-Faltung (Assoziativität und Kommutativität, sowie Satz 2.4 **(a)** und **(b)**) erst einmal nachgewiesen hat.

¹⁵Denn $L(S) = \prod_{p \in S} p$ ist ein Produkt von paarweise verschiedenen Primzahlen, und nach der Definition von μ ist somit $\mu(L(S)) = 1$, wenn es gerade viele Primzahlen sind (also wenn $|S|$ gerade ist), und $\mu(L(S)) = -1$, wenn es ungerade viele Primzahlen sind (also wenn $|S|$ ungerade ist); in beiden Fällen ist also $\mu(L(S)) = (-1)^{|S|}$.

Definition (punktweises Produkt): Nun definieren wir eine weitere Verknüpfung von Zahlenfunktionen, und zwar eine viel einfachere: das *punktweise Produkt*. Und zwar definieren wir für je zwei Zahlenfunktionen f und g eine Zahlenfunktion $f \cdot g$ durch

$$((f \cdot g)(n) = f(n)g(n) \text{ für alle } n \in \mathbb{N}_+).$$

Diese Zahlenfunktion $f \cdot g$ nennen wir das *punktweise Produkt* der Zahlenfunktionen f und g .

Es ist klar, daß das punktweise Produkt kommutativ ist (d. h. daß $f \cdot g = g \cdot f$ für beliebige Zahlenfunktionen f und g gilt), assoziativ ist (also $(f \cdot g) \cdot h = f \cdot (g \cdot h)$ für je drei Zahlenfunktionen f, g und h) und daß $f \cdot \underline{1} = \underline{1} \cdot f = f$ für jede Zahlenfunktion f ist. Die Eigenschaft des punktweisen Produktes, die uns am meisten interessiert, ist aber folgende:

Satz 2.5: Für je zwei Zahlenfunktionen f und g ist

$$(\text{id} \cdot f) * (\text{id} \cdot g) = \text{id} \cdot (f * g).$$

Beweis von Satz 2.5: Für jedes $n \in \mathbb{N}_+$ ist

$$\begin{aligned} ((\text{id} \cdot f) * (\text{id} \cdot g))(n) &= \sum_{d|n} \underbrace{(\text{id} \cdot f)(d)}_{=\text{id}(d)f(d)=df(d)} \cdot \underbrace{(\text{id} \cdot g)\left(\frac{n}{d}\right)}_{=\text{id}\left(\frac{n}{d}\right)g\left(\frac{n}{d}\right)=\frac{n}{d}g\left(\frac{n}{d}\right)} = \sum_{d|n} df(d) \frac{n}{d} g\left(\frac{n}{d}\right) \\ &= \sum_{d|n} n f(d) g\left(\frac{n}{d}\right) = \underbrace{\sum_{d|n} f(d) g\left(\frac{n}{d}\right)}_{=(f * g)(n)} = \text{id}(n) (f * g)(n) = (\text{id} \cdot (f * g))(n). \end{aligned}$$

Also ist $(\text{id} \cdot f) * (\text{id} \cdot g) = \text{id} \cdot (f * g)$, und Satz 2.5 ist bewiesen.

Schließlich brauchen wir noch eine Eigenschaft der Dirichlet-Faltung: Nämlich kann man sich fragen, wann eine Zahlenfunktion f ein Inverses bezüglich $*$ hat, also eine Zahlenfunktion f' mit $f * f' = f' * f = \varepsilon$ existiert. Klar ist, daß μ und $\underline{1}$ je ein Inverses haben (denn nach Satz 2.4 (b) sind μ und $\underline{1}$ gegenseitig invers). Aber auch id und ϕ haben Inverse, und allgemeiner gilt:

Satz 2.6: Sei $f : \mathbb{N}_+ \rightarrow \mathbb{Z}$ eine Zahlenfunktion. Genau dann existiert eine Zahlenfunktion $f' : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $f * f' = f' * f = \varepsilon$, wenn $(f(1) = 1$ oder $f(1) = -1)$ ist.

Beweis von Satz 2.6: Eine Richtung von Satz 2.6 ist klar: wenn eine Zahlenfunktion $f' : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $f * f' = f' * f = \varepsilon$ existiert, dann ist $(f(1) = 1$ oder $f(1) = -1)$ ¹⁶.

¹⁶Denn aus $f * f' = \varepsilon$ folgt $(f * f')(1) = \varepsilon(1) = 1$, was wegen $(f * f')(1) = \sum_{d|1} f(d) f'\left(\frac{1}{d}\right) = f(1) f'(1)$ zu $f(1) f'(1) = 1$ wird, und somit muß $f(1)$ ein Teiler von 1 sein, also $f(1) = 1$ oder $f(1) = -1$.

Wir wollen nun die umgekehrte Richtung beweisen: Angenommen, $(f(1) = 1 \text{ oder } f(1) = -1)$. Dann ist zu zeigen, daß eine Zahlenfunktion $f' : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $f * f' = f' * f = \varepsilon$ existiert.

Wir wollen diese Zahlenfunktion $f' : \mathbb{N}_+ \rightarrow \mathbb{Z}$ konstruieren, indem wir ihre Werte $f'(n)$ für alle $n \in \mathbb{N}_+$ definieren. Und zwar definieren wir sie rekursiv: Sei $n \in \mathbb{N}_+$. Wir nehmen an, wir haben alle Werte $f'(m)$ für $m \in \mathbb{N}_+$ mit $m < n$ bereits definiert (für $n = 1$ heißt dies, daß wir noch gar nichts definiert haben). Dann definieren wir den Wert $f'(n)$ durch

$$f'(n) = \frac{1}{f(1)} \left(\varepsilon(n) - \sum_{\substack{d \in \mathbb{N}_{|n}; \\ d < n}} f'(d) f\left(\frac{n}{d}\right) \right) \quad (13)$$

(wobei $\frac{1}{f(1)}$ eine ganze Zahl ist, da $(f(1) = 1 \text{ oder } f(1) = -1)$ gilt). Auf diese Weise haben wir eine Zahlenfunktion $f' : \mathbb{N}_+ \rightarrow \mathbb{Z}$ definiert. Jetzt ist nur noch zu zeigen, daß für diese Zahlenfunktion f' auch tatsächlich $f * f' = f' * f = \varepsilon$ erfüllt ist. In der Tat gilt für jedes $n \in \mathbb{N}_+$ die Gleichung

$$\begin{aligned} (f' * f)(n) &= \sum_{d|n} f'(d) f\left(\frac{n}{d}\right) = \sum_{d \in \mathbb{N}_{|n}} f'(d) f\left(\frac{n}{d}\right) = \sum_{\substack{d \in \mathbb{N}_{|n}; \\ d < n}} f'(d) f\left(\frac{n}{d}\right) + \underbrace{f'(n) f\left(\frac{n}{n}\right)}_{=f(1)} \\ &= \sum_{\substack{d \in \mathbb{N}_{|n}; \\ d < n}} f'(d) f\left(\frac{n}{d}\right) + f'(n) f(1) = \sum_{\substack{d \in \mathbb{N}_{|n}; \\ d < n}} f'(d) f\left(\frac{n}{d}\right) + \left(\varepsilon(n) - \sum_{\substack{d \in \mathbb{N}_{|n}; \\ d < n}} f'(d) f\left(\frac{n}{d}\right) \right) \\ &\quad \left(\text{denn nach (13) ist } f'(n) f(1) = \varepsilon(n) - \sum_{\substack{d \in \mathbb{N}_{|n}; \\ d < n}} f'(d) f\left(\frac{n}{d}\right) \right) \\ &= \varepsilon(n), \end{aligned}$$

und somit ist $f' * f = \varepsilon$, und wegen der Kommutativität von $*$ folgt hieraus $f * f' = f' * f = \varepsilon$. Damit ist auch die zweite Richtung von Satz 2.6 bewiesen.

2.5. Beweis von Satz 2.1: $\mathcal{D} \iff \mathcal{E} \iff \mathcal{F} \iff \mathcal{G}$

Nun können wir zum *Beweis von Satz 2.1* schreiten. Wir identifizieren die Folge (b_1, b_2, b_3, \dots) mit der durch $(b(n) = b_n \text{ für jedes } n \in \mathbb{N}_+)$ definierten Zahlenfunktion $b : \mathbb{N}_+ \rightarrow \mathbb{Z}$. Nun führen wir vier neue Aussagen ein:

Aussage \mathcal{D}_1 : Es gibt eine Zahlenfunktion $y : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $b = (\text{id} \cdot y) * \underline{1}$.

Aussage \mathcal{E}_1 : Es gibt *genau eine* Zahlenfunktion $y : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $b = (\text{id} \cdot y) * \underline{1}$.

Aussage \mathcal{F}_1 : Es gibt eine Zahlenfunktion $z : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $\mu * b = \text{id} \cdot z$.

Aussage \mathcal{G}_1 : Es gibt eine Zahlenfunktion $w : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $\phi * b = \text{id} \cdot w$.

Wir werden nun zeigen, daß diese Aussagen \mathcal{D}_1 , \mathcal{E}_1 , \mathcal{F}_1 und \mathcal{G}_1 jeweils äquivalent zu den Aussagen \mathcal{D} , \mathcal{E} , \mathcal{F} bzw. \mathcal{G} sind.

Wir wollen erstmal beweisen, daß $\mathcal{D} \iff \mathcal{D}_1$ ist. In der Tat haben wir folgende Äquivalenz von Aussagen¹⁷:

$$\begin{aligned} \mathcal{D} &\iff \left(\text{es gibt eine Folge } (y_1, y_2, y_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}, \text{ die } b_n = \sum_{d|n} dy_d \text{ für jedes } n \in \mathbb{N}_+ \text{ erfüllt} \right) \\ &\iff \left(\text{es gibt eine Zahlenfunktion } y : \mathbb{N}_+ \rightarrow \mathbb{Z}, \text{ die } b(n) = \sum_{d|n} dy(d) \text{ für jedes } n \in \mathbb{N}_+ \text{ erfüllt} \right) \\ &\iff (\text{es gibt eine Zahlenfunktion } y : \mathbb{N}_+ \rightarrow \mathbb{Z}, \text{ die } b = (\text{id} \cdot y) * \underline{1} \text{ erfüllt}) \iff \mathcal{D}_1. \end{aligned}$$

Hierbei gilt der erste Äquivalenzpfeil wegen der Definition von Aussage \mathcal{D} ; der zweite Äquivalenzpfeil ergibt sich aus der Identifikation von Folgen mit Zahlenfunktionen; der dritte Äquivalenzpfeil kommt davon, daß $\left(b(n) = \sum_{d|n} dy(d) \text{ für jedes } n \in \mathbb{N}_+ \right)$ äquivalent zu $b = (\text{id} \cdot y) * \underline{1}$ ist¹⁸; der vierte Äquivalenzpfeil kommt wiederum von der Definition von Aussage \mathcal{D}_1 . Somit ist $\mathcal{D} \iff \mathcal{D}_1$ bewiesen.

Genauso, wie wir gerade $\mathcal{D} \iff \mathcal{D}_1$ gezeigt haben, läßt sich $\mathcal{E} \iff \mathcal{E}_1$ zeigen (man muß nur an den richtigen Stellen "eine Folge" durch "genau eine Folge" bzw. "eine Zahlenfunktion" durch "genau eine Zahlenfunktion" ersetzen).

Kommen wir nun zum Beweis von $\mathcal{F} \iff \mathcal{F}_1$: Wir haben folgende Äquivalenz von Aussagen¹⁹:

$$\begin{aligned} \mathcal{F} &\iff \left(\text{für jedes } n \in \mathbb{N}_+ \text{ ist } \sum_{d|n} \mu(d) b_{n/d} \in n\mathbb{Z} \right) \\ &\iff \left(\text{für jedes } n \in \mathbb{N}_+ \text{ gibt es ein } z_n \in \mathbb{Z} \text{ mit } \sum_{d|n} \mu(d) b_{n/d} = nz_n \right) \\ &\iff \left(\text{es gibt eine Folge } (z_1, z_2, z_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}, \text{ die } \sum_{d|n} \mu(d) b_{n/d} = nz_n \text{ für jedes } n \in \mathbb{N}_+ \text{ erfüllt} \right) \\ &\iff \left(\text{es gibt eine Zahlenfunktion } z : \mathbb{N}_+ \rightarrow \mathbb{Z}, \text{ die } \sum_{d|n} \mu(d) b\left(\frac{n}{d}\right) = nz(n) \text{ für jedes } n \in \mathbb{N}_+ \text{ erfüllt} \right) \\ &\iff (\text{es gibt eine Zahlenfunktion } z : \mathbb{N}_+ \rightarrow \mathbb{Z}, \text{ die } \mu * b = \text{id} \cdot z \text{ erfüllt}) \iff \mathcal{F}_1. \end{aligned}$$

¹⁷Die Äquivalenzen werden weiter unten genauer begründet.

¹⁸Denn für jedes $n \in \mathbb{N}_+$ ist

$$((\text{id} \cdot y) * \underline{1})(n) = \sum_{d|n} \underbrace{(\text{id} \cdot y)(d)}_{=\text{id}(d)y(d)=dy(d)} \underbrace{\underline{1}\left(\frac{n}{d}\right)}_{=1} = \sum_{d|n} dy(d),$$

und somit ist $\left(b(n) = \sum_{d|n} dy(d) \text{ für jedes } n \in \mathbb{N}_+ \right)$ äquivalent zu

$(b(n) = ((\text{id} \cdot y) * \underline{1})(n) \text{ für jedes } n \in \mathbb{N}_+)$, also zu $b = (\text{id} \cdot y) * \underline{1}$.

¹⁹Die Äquivalenzen werden weiter unten genauer begründet.

Hierbei folgt der erste Äquivalenzpfeil aus der Definition von Aussage \mathcal{F} ; der zweite und der dritte sind offensichtlich; der vierte Äquivalenzpfeil ist der Identifizierung zwischen Folgen und Zahlenfunktionen geschuldet; der fünfte Äquivalenzpfeil rührt von der Äquivalenz der Aussagen $\left(\sum_{d|n} \mu(d) b\left(\frac{n}{d}\right) = nz(n) \text{ für jedes } n \in \mathbb{N}_+\right)$ und $\mu * b = \text{id} \cdot z$ ²⁰; der sechste Äquivalenzpfeil kommt von der Definition von Aussage \mathcal{F}_1 . Somit ist $\mathcal{F} \iff \mathcal{F}_1$ bewiesen. Völlig analog zeigt sich, daß $\mathcal{G} \iff \mathcal{G}_1$ ist (im Beweis muss lediglich μ durch ϕ und z durch w ersetzt werden).

Als nächstes sind die Äquivalenzen zwischen den Aussagen \mathcal{D}_1 , \mathcal{E}_1 , \mathcal{F}_1 und \mathcal{G}_1 dran. Diese sind jetzt, wo wir die Eigenschaften der Dirichlet-Faltung kennen, sehr leicht:

Beweis von $\mathcal{D}_1 \implies \mathcal{F}_1$: Angenommen, Aussage \mathcal{D}_1 gilt. Dann gibt es eine Zahlenfunktion $y : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $b = (\text{id} \cdot y) * \underline{1}$. Wir haben also $b = (\text{id} \cdot y) * \underline{1} = \underline{1} * (\text{id} \cdot y)$ (wegen der Kommutativität der Dirichlet-Faltung $*$) und somit

$$\begin{aligned} \mu * b &= \mu * (\underline{1} * (\text{id} \cdot y)) = \underbrace{(\mu * \underline{1})}_{=\varepsilon \text{ (nach Satz 2.4 (b))}} * (\text{id} \cdot y) && \text{(wegen der Assoziativität von *)} \\ &= \varepsilon * (\text{id} \cdot y) = \text{id} \cdot y. \end{aligned}$$

Also gibt es eine Zahlenfunktion $z : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $\mu * b = \text{id} \cdot z$, nämlich $z = y$. Also gilt Aussage \mathcal{F}_1 . Damit ist $\mathcal{D}_1 \implies \mathcal{F}_1$ gezeigt.

Beweis von $\mathcal{F}_1 \implies \mathcal{G}_1$: Angenommen, Aussage \mathcal{F}_1 gilt. Dann gibt es eine Zahlenfunktion $z : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $\mu * b = \text{id} \cdot z$. Nach Satz 2.4 (c) ist nun $\phi = \mu * \text{id} = \text{id} * \mu$ (da $*$ kommutativ ist), also

$$\begin{aligned} \phi * b &= (\text{id} * \mu) * b = \underbrace{\text{id}}_{=\text{id} \cdot \underline{1}} * \underbrace{(\mu * b)}_{=\text{id} \cdot z} && \text{(weil * assoziativ ist)} \\ &= (\text{id} \cdot \underline{1}) * (\text{id} \cdot z) = \text{id} \cdot (\underline{1} * z) && \text{(nach Satz 2.5)}. \end{aligned}$$

Also gibt es eine Zahlenfunktion $w : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $\phi * b = \text{id} \cdot w$, nämlich $w = \underline{1} * z$. Also gilt Aussage \mathcal{G}_1 . Damit ist $\mathcal{F}_1 \implies \mathcal{G}_1$ nachgewiesen.

Beweis von $\mathcal{G}_1 \implies \mathcal{D}_1$: Angenommen, Aussage \mathcal{G}_1 gelte. Es gibt also eine Zahlenfunktion $w : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $\phi * b = \text{id} \cdot w$. Sei nun $y = \mu * w$. Dann ist $y = \mu * w = w * \mu$ (da $*$ kommutativ ist), also

$$\begin{aligned} y * \underline{1} &= (w * \mu) * \underline{1} = w * \underbrace{(\mu * \underline{1})}_{=\varepsilon \text{ (nach Satz 2.4 (b))}} && \text{(da * assoziativ ist)} \\ &= w * \varepsilon = w, \end{aligned}$$

²⁰Denn für jedes $n \in \mathbb{N}_+$ ist $\sum_{d|n} \mu(d) b\left(\frac{n}{d}\right) = (\mu * b)(n)$ und $\underbrace{\sum_{d|n} z(n)}_{=\text{id}(n)} = \text{id}(n) z(n) =$

$(\text{id} \cdot z)(n)$, und somit ist die Aussage $\left(\sum_{d|n} \mu(d) b\left(\frac{n}{d}\right) = nz(n) \text{ für jedes } n \in \mathbb{N}_+\right)$ äquivalent zu $((\mu * b)(n) = (\text{id} \cdot z)(n) \text{ für jedes } n \in \mathbb{N}_+)$, also zu $\mu * b = \text{id} \cdot z$.

und somit wird $\phi * b = \text{id} \cdot w$ zu

$$\begin{aligned} \phi * b &= \text{id} \cdot (y * \underline{1}) = (\text{id} \cdot y) * \left(\underbrace{\text{id} \cdot \underline{1}}_{=\text{id}=\phi * \underline{1} \text{ (nach Satz 2.4 (a))}} \right) && \text{(nach Satz 2.5)} \\ &= (\text{id} \cdot y) * (\phi * \underline{1}) = (\phi * \underline{1}) * (\text{id} \cdot y) && \text{(da } * \text{ kommutativ ist)} \\ &= \phi * (\underline{1} * (\text{id} \cdot y)) && \text{(da } * \text{ assoziativ ist)}. \end{aligned}$$

Nach Satz 2.6 (angewandt auf $f = \phi$) gibt es nun eine Zahlenfunktion $\phi' : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $\phi * \phi' = \phi' * \phi = \varepsilon$ (denn $\phi(1) = 1$). Wegen der Assoziativität von $*$ ist nun $\phi' * (\phi * g) = \underbrace{(\phi' * \phi)}_{=\varepsilon} * g = \varepsilon * g = g$ für jede Zahlenfunktion $g : \mathbb{N}_+ \rightarrow \mathbb{Z}$. Insbesondere

folgt daraus: Wenn $\phi * g = \phi * h$ für zwei Zahlenfunktionen g und h ist, dann ist $g = h$ (denn $\phi' * (\phi * g) = g$ und analog $\phi' * (\phi * h) = h$). Angewandt auf $g = b$ und $h = \underline{1} * (\text{id} \cdot y)$ ergibt dies $b = \underline{1} * (\text{id} \cdot y)$ (denn $\phi * b = \phi * (\underline{1} * (\text{id} \cdot y))$). Also ist $b = \underline{1} * (\text{id} \cdot y) = (\text{id} \cdot y) * \underline{1}$ (wegen der Kommutativität von $*$). Aussage \mathcal{D}_1 gilt also. Damit ist $\mathcal{G}_1 \implies \mathcal{D}_1$ bewiesen.

Aus $\mathcal{D}_1 \implies \mathcal{F}_1$, $\mathcal{F}_1 \implies \mathcal{G}_1$ und $\mathcal{G}_1 \implies \mathcal{D}_1$ folgt schon einmal die Äquivalenz $\mathcal{D}_1 \iff \mathcal{F}_1 \iff \mathcal{G}_1$. Jetzt müssen wir die Äquivalenz zwischen \mathcal{D}_1 und \mathcal{E}_1 nachprüfen:

Beweis von $\mathcal{D}_1 \implies \mathcal{E}_1$: Angenommen, Aussage \mathcal{D}_1 gilt. Wir wollen Aussage \mathcal{E}_1 nachweisen; d. h. wir wollen zeigen, daß es *genau eine* Zahlenfunktion $y : \mathbb{N}_+ \rightarrow \mathbb{Z}$ mit $b = (\text{id} \cdot y) * \underline{1}$ gibt. Dazu nehmen wir an, $y : \mathbb{N}_+ \rightarrow \mathbb{Z}$ und $\tilde{y} : \mathbb{N}_+ \rightarrow \mathbb{Z}$ seien *zwei* Zahlenfunktionen mit $b = (\text{id} \cdot y) * \underline{1}$ und $b = (\text{id} \cdot \tilde{y}) * \underline{1}$. Unser Ziel besteht darin, $y = \tilde{y}$ zu beweisen.

In der Tat ist $\mu * b = \text{id} \cdot y$ (dies beweist man genau so wie im Beweis von $\mathcal{D}_1 \implies \mathcal{F}_1$) und $\mu * b = \text{id} \cdot \tilde{y}$ (analogerweise), und damit $\text{id} \cdot y = \text{id} \cdot \tilde{y}$. Dies führt auf $y = \tilde{y}$ (denn für jedes $n \in \mathbb{N}_+$ ist $(\text{id} \cdot y)(n) = \text{id}(n)y(n) = ny(n)$ und analog $(\text{id} \cdot \tilde{y})(n) = n\tilde{y}(n)$, und somit folgt aus $\text{id} \cdot y = \text{id} \cdot \tilde{y}$ sofort $ny(n) = n\tilde{y}(n)$ und damit $y(n) = \tilde{y}(n)$ für jedes $n \in \mathbb{N}_+$). Folglich ist Aussage \mathcal{E}_1 wahr. Damit ist $\mathcal{D}_1 \implies \mathcal{E}_1$ gezeigt.

Zusammen mit der offensichtlichen Implikation $\mathcal{E}_1 \implies \mathcal{D}_1$ ergibt dies die Äquivalenz $\mathcal{D}_1 \iff \mathcal{E}_1$. Diese Äquivalenz läßt sich mit $\mathcal{D}_1 \iff \mathcal{F}_1 \iff \mathcal{G}_1$ zu $\mathcal{D}_1 \iff \mathcal{E}_1 \iff \mathcal{F}_1 \iff \mathcal{G}_1$ verknüpfen. Wegen $\mathcal{D} \iff \mathcal{D}_1$, $\mathcal{E} \iff \mathcal{E}_1$, $\mathcal{F} \iff \mathcal{F}_1$ und $\mathcal{G} \iff \mathcal{G}_1$ führt dies auf $\mathcal{D} \iff \mathcal{E} \iff \mathcal{F} \iff \mathcal{G}$. Jetzt fehlt es uns noch, die anderen Aussagen \mathcal{A} , \mathcal{B} , \mathcal{C} und \mathcal{H} an diese Äquivalenz anzuschließen.

Für Aussage \mathcal{A} kann man dies entweder, indem man $\mathcal{A} \iff \mathcal{D}$ zeigt, oder indem man $\mathcal{A} \iff \mathcal{F}$ zeigt. Wir wählen den ersteren Weg:

2.6. Beweis von Satz 2.1: $\mathcal{A} \iff \mathcal{D}$

Beweis von $\mathcal{A} \implies \mathcal{D}$: Angenommen, Aussage \mathcal{A} sei wahr. Das heißt, für jede Zahl $n \in \mathbb{N}_+$ und jeden Primteiler p von n gilt

$$b_{n/p} \equiv b_n \pmod{p^{v_p(n)}}. \quad (14)$$

Wir wollen Aussage \mathcal{D} beweisen; wir wollen also zeigen, daß es eine Folge $(y_1, y_2, y_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ gibt, die

$$b_n = \sum_{d|n} dy_d \quad (15)$$

für jedes $n \in \mathbb{N}_+$ erfüllt. Dazu definieren wir eine solche Folge $(y_1, y_2, y_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ rekursiv: Sei $m \in \mathbb{N}_+$. Angenommen, alle Folgenglieder $y_n \in \mathbb{Z}$ für $n \in \{1, 2, \dots, m-1\}$ sind bereits definiert, und zwar so, daß (15) für jedes $n \in \{1, 2, \dots, m-1\}$ erfüllt ist. Wir wollen jetzt ein Folgenglied $y_m \in \mathbb{Z}$ definieren, welches dazu führt, daß (15) auch für $n = m$ gilt. Wenn wir damit Erfolg haben, können wir auf diese Weise eine Folge $(y_1, y_2, y_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ rekursiv konstruieren, die (15) für jedes $n \in \mathbb{N}_+$ erfüllt; damit wird Aussage \mathcal{D} bewiesen sein.

Versuchen wir also, y_m zu definieren. Und zwar wollen wir y_m als

$$y_m = \frac{1}{m} \left(b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d \right) \quad (16)$$

definieren; dazu müssen wir beweisen, daß $m \mid b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d$ ist.

In der Tat sei $\text{PF } m$ die Menge aller Primteiler von m . Dann ist $m = \prod_{p \in \text{PF } m} p^{v_p(m)}$

²¹. Wir werden nun zeigen, daß $p^{v_p(m)} \mid b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d$ für jedes $p \in \text{PF } m$ gilt;

daraus wird natürlich $\prod_{p \in \text{PF } m} p^{v_p(m)} \mid b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d$ folgen (denn die Zahlen $p^{v_p(m)}$ für

verschiedene $p \in \text{PF } m$ sind alle teilerfremd, weil alle diese $p \in \text{PF } m$ prim sind), und damit $m \mid b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d$, womit unser Ziel erreicht sein wird.

Wir haben angenommen, daß (15) für jedes $n \in \{1, 2, \dots, m-1\}$ erfüllt ist. Insbesondere ist also (15) für $n = m/p$ erfüllt (denn wegen $p \in \text{PF } m$ ist $m/p \in \mathbb{Z}$, also $m/p \in \{1, 2, \dots, m-1\}$); mit anderen Worten: $b_{m/p} = \sum_{d|(m/p)} dy_d$. Nun zerlegen wir

die Summe $\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d$ in zwei Teilsommen:

$$\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d = \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d|(m/p)}} dy_d + \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d \nmid (m/p)}} dy_d.$$

Die erste dieser Teilsommen ist

$$\begin{aligned} \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d|(m/p)}} dy_d &= \sum_{d \in \mathbb{N}_{|(m/p)}} dy_d \quad \left(\begin{array}{l} \text{denn die Teiler von } m, \text{ die kleiner als } m \text{ sind und } m/p \text{ teilen,} \\ \text{sind schlicht und einfach die Teiler von } m/p \end{array} \right) \\ &= \sum_{d|(m/p)} dy_d = b_{m/p}. \end{aligned}$$

Die zweite dieser Teilsommen ist

$$\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d \nmid (m/p)}} dy_d \equiv 0 \pmod{p^{v_p(m)}},$$

²¹Denn für jedes $p \in \text{PF } m$ ist $v_p(m)$ die höchste Zahl $k \in \mathbb{N}$, die $p^k \mid m$ erfüllt, also der Exponent, mit dem der Primfaktor p in der Primfaktorzerlegung von m vorkommt.

denn für jeden Teiler d von m , der $d \nmid (m/p)$ erfüllt, gilt $p^{v_p(m)} \mid d$ ²² und damit $dy_d \equiv 0 \pmod{p^{v_p(m)}}$. Die Gesamtsumme ist somit

$$\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d = \underbrace{\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d \mid (m/p)}} dy_d}_{= b_{m/p}} + \underbrace{\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d \nmid (m/p)}} dy_d}_{\equiv 0 \pmod{p^{v_p(m)}}} \equiv b_{m/p} + 0 = b_{m/p} \equiv b_m \pmod{p^{v_p(m)}}$$

(nach Aussage \mathcal{A} , angewandt auf $n = m$, denn Aussage \mathcal{A} haben wir ja als wahr angenommen). Das heißt, $p^{v_p(m)} \mid b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d$. Wie wir wissen, folgt hieraus $m \mid$

$b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d$, und somit ist durch (16) tatsächlich eine ganze Zahl $y_m \in \mathbb{Z}$ definiert.

Diese Zahl y_m erfüllt tatsächlich (15) für $n = m$, weil

$$\begin{aligned} \sum_{d \mid m} dy_d &= \sum_{\substack{d \in \mathbb{N}_{|m} \\ d < m}} dy_d = my_m + \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d = m \cdot \frac{1}{m} \left(b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d \right) + \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dy_d & \quad (\text{nach (16)}) \\ &= b_m \end{aligned}$$

gilt.

Wir können also rekursiv eine Folge $(y_1, y_2, y_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ konstruieren, und diese Folge erfüllt (15) für jedes $n \in \mathbb{N}_+$. Damit ist Aussage \mathcal{D} bewiesen. Wir haben also gezeigt, daß $\mathcal{A} \implies \mathcal{D}$ ist.

Beweis von $\mathcal{D} \implies \mathcal{A}$: Nun wollen wir die Aussage \mathcal{D} als wahr annehmen. Es gibt also eine Folge $(y_1, y_2, y_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ ganzer Zahlen, die (6) erfüllt. Wir wollen nun Aussage \mathcal{A} herleiten.

In der Tat sei $n \in \mathbb{N}_+$ ganz, und p ein Primteiler von n . Nach (6) gilt dann $b_n = \sum_{d \mid n} dy_d$, und wenden wir (6) auf n/p statt n an, so erhalten wir $b_{n/p} = \sum_{d \mid (n/p)} dy_d$.

Nun zerlegen wir die Summe $\sum_{d \mid n} dy_d = \sum_{d \in \mathbb{N}_{|n}} dy_d$ in zwei Teilsommen:

$$\sum_{d \in \mathbb{N}_{|n}} dy_d = \sum_{\substack{d \in \mathbb{N}_{|n}; \\ d \mid (n/p)}} dy_d + \sum_{\substack{d \in \mathbb{N}_{|n}; \\ d \nmid (n/p)}} dy_d.$$

Die erste dieser Teilsommen ist

$$\begin{aligned} \sum_{\substack{d \in \mathbb{N}_{|n}; \\ d \mid (n/p)}} dy_d &= \sum_{d \in \mathbb{N}_{|(n/p)}} dy_d & \left(\begin{array}{l} \text{denn die Teiler von } n, \text{ die } n/p \text{ teilen,} \\ \text{sind schlicht und einfach die Teiler von } n/p \end{array} \right) \\ &= \sum_{d \mid (n/p)} dy_d = b_{n/p}. \end{aligned}$$

²²Denn aus $d \nmid (m/p)$ folgt $\frac{m/p}{d} \notin \mathbb{Z}$, also $\frac{m/d}{p} \notin \mathbb{Z}$ (denn $\frac{m/d}{p} = \frac{m}{dp} = \frac{m/p}{d}$), und damit $p \nmid (m/d)$, also $v_p(m/d) = 0$ und damit $v_p(m) = v_p(d)$ (denn $v_p(m) = v_p(d \cdot m/d) = v_p(d) + v_p(m/d)$, weil für zwei beliebige natürliche Zahlen α und β stets $v_p(\alpha\beta) = v_p(\alpha) + v_p(\beta)$ gilt), also $p^{v_p(m)} = p^{v_p(d)} \mid d$.

Die zweite dieser Teilsummen ist

$$\sum_{\substack{d \in \mathbb{N}_{|n}; \\ d \nmid (n/p)}} dy_d \equiv 0 \pmod{p^{v_p(n)}},$$

denn für jeden Teiler d von n , der $d \nmid (n/p)$ erfüllt, gilt $p^{v_p(n)} \mid d$ ²³ und damit $dy_d \equiv 0 \pmod{p^{v_p(n)}}$. Die Gesamtsumme ist somit

$$\sum_{d \in \mathbb{N}_{|n}} dy_d = \underbrace{\sum_{\substack{d \in \mathbb{N}_{|n}; \\ d \mid (n/p)}} dy_d}_{=b_{n/p}} + \underbrace{\sum_{\substack{d \in \mathbb{N}_{|n}; \\ d \nmid (n/p)}} dy_d}_{\equiv 0 \pmod{p^{v_p(n)}}} \equiv b_{n/p} + 0 = b_{n/p} \pmod{p^{v_p(n)}}.$$

Hieraus folgt

$$b_n = \sum_{d|n} dy_d = \sum_{d \in \mathbb{N}_{|n}} dy_d \equiv b_{n/p} \pmod{p^{v_p(n)}},$$

und damit ist Aussage \mathcal{A} nachgewiesen. Wir haben damit gezeigt, daß $\mathcal{D} \implies \mathcal{A}$ ist.

Da wir nun $\mathcal{A} \implies \mathcal{D}$ und $\mathcal{D} \implies \mathcal{A}$ nachgewiesen haben, wissen wir also, daß $\mathcal{A} \iff \mathcal{D}$ ist. Zusammen mit $\mathcal{D} \iff \mathcal{E} \iff \mathcal{F} \iff \mathcal{G}$ ergibt dies die Äquivalenzkette $\mathcal{A} \iff \mathcal{D} \iff \mathcal{E} \iff \mathcal{F} \iff \mathcal{G}$.

Hiermit haben wir zwar noch nicht den ganzen Satz 2.1 gezeigt, aber bereits genug, um daraus Satz 1.1 herzuleiten. *Unser zweiter Beweis von Satz 1.1 ist also fertig.* Aber zum ersten Beweis von Satz 1.1 benötigen wir noch Aussage \mathcal{B} , und Aussagen \mathcal{C} und \mathcal{G} sind ebenfalls von Interesse.

2.7. Beweis von Satz 2.1: $\mathcal{G} \iff \mathcal{H}$

Der *Beweis von $\mathcal{G} \iff \mathcal{H}$* ist ganz einfach, weil Aussage \mathcal{H} nur eine Umschreibung von Aussage \mathcal{G} ist, denn für jedes $n \in \mathbb{N}_+$ ist

$$\begin{aligned} \sum_{i=1}^n b_{\text{ggT}(i,n)} &= \sum_{j=1}^n b_{\text{ggT}(j,n)} = \sum_{j \in \{1,2,\dots,n\}} b_{\text{ggT}(j,n)} = \sum_{d|n} \sum_{\substack{j \in \{1,2,\dots,n\}; \\ \text{ggT}(j,n)=d}} \underbrace{b_{\text{ggT}(j,n)}}_{=b_d \text{ (denn } \text{ggT}(j,n)=d)} = \sum_{d|n} \sum_{\substack{j \in \{1,2,\dots,n\}; \\ \text{ggT}(j,n)=d}} b_d \\ &= \sum_{d|n} \underbrace{|\{j \in \{1,2,\dots,n\} \mid \text{ggT}(j,n)=d\}|}_{=\phi\left(\frac{n}{d}\right) \text{ (nach (11))}} \cdot b_d = \sum_{d|n} \phi\left(\frac{n}{d}\right) b_d = \sum_{d \in \mathbb{N}_{|n}} \phi\left(\frac{n}{d}\right) b_d \\ &= \sum_{d \in \mathbb{N}_{|n}} \phi\left(\frac{n}{n/d}\right) b_{n/d} \quad \left(\begin{array}{l} \text{hier haben wir } d \text{ für } \frac{n}{d} \text{ in der Summe substituiert, denn} \\ \text{die Abbildung } \mathbb{N}_{|n} \rightarrow \mathbb{N}_{|n}, d \mapsto \frac{n}{d} \text{ ist bijektiv} \end{array} \right) \\ &= \sum_{d \in \mathbb{N}_{|n}} \phi(d) b_{n/d} = \sum_{d|n} \phi(d) b_{n/d}. \end{aligned}$$

Die gerade bewiesene Äquivalenz $\mathcal{G} \iff \mathcal{H}$ erweitert die Äquivalenzkette $\mathcal{A} \iff \mathcal{D} \iff \mathcal{E} \iff \mathcal{F} \iff \mathcal{G}$ um die Aussage \mathcal{H} . Wir haben damit gezeigt: $\mathcal{A} \iff \mathcal{D} \iff \mathcal{E} \iff \mathcal{F} \iff \mathcal{G} \iff \mathcal{H}$. Es fehlen aber noch Aussagen \mathcal{B} und \mathcal{C} .

²³Dies läßt sich genauso beweisen, wie wir im Beweis von $\mathcal{A} \implies \mathcal{D}$ gezeigt haben, daß für jeden Teiler d von m , der $d \nmid (m/p)$ erfüllt, $p^{v_p(m)} \mid d$ gilt. Wir müssen nur im Beweis immer m durch n ersetzen.

2.8. Beweis von Satz 2.1: $\mathcal{A} \iff \mathcal{B}$

An dieser Stelle gibt es zwei Wege, weiterzukommen. Wir werden sie beide begehen, jedoch verschieben wir den zweiten auf später.

Der erste Weg besteht darin, $\mathcal{A} \iff \mathcal{B}$ zu zeigen. Der Beweis ist sehr ähnlich zu dem oben für $\mathcal{A} \iff \mathcal{D}$ gegebenen; wir werden, soweit möglich, den letzteren Wort für Wort übernehmen, um die Ähnlichkeiten zu verdeutlichen.

Beweis von $\mathcal{A} \implies \mathcal{B}$: Angenommen, Aussage \mathcal{A} sei wahr. Das heißt, für jede Zahl $n \in \mathbb{N}_+$ und jeden Primteiler p von n gilt

$$b_{n/p} \equiv b_n \pmod{p^{v_p(n)}}. \quad (17)$$

Wir wollen Aussage \mathcal{B} beweisen; wir wollen also zeigen, daß es eine Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ gibt, die

$$b_n = \sum_{d|n} dx_d^{n/d} \quad (18)$$

für jedes $n \in \mathbb{N}_+$ erfüllt. Dazu definieren wir eine solche Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ rekursiv: Sei $m \in \mathbb{N}_+$. Angenommen, alle Folgenglieder $x_n \in \mathbb{Z}$ für $n \in \{1, 2, \dots, m-1\}$ sind bereits definiert, und zwar so, daß (18) für jedes $n \in \{1, 2, \dots, m-1\}$ erfüllt ist. Wir wollen jetzt ein Folgenglied $x_m \in \mathbb{Z}$ definieren, welches dazu führt, daß (18) auch für $n = m$ gilt. Wenn wir damit Erfolg haben, können wir auf diese Weise eine Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ rekursiv konstruieren, die (18) für jedes $n \in \mathbb{N}_+$ erfüllt; damit wird Aussage \mathcal{B} bewiesen sein.

Versuchen wir also, x_m zu definieren. Und zwar wollen wir x_m als

$$x_m = \frac{1}{m} \left(b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d} \right) \quad (19)$$

definieren; dazu müssen wir beweisen, daß $m \mid b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d}$ ist.

In der Tat sei $\text{PF } m$ die Menge aller Primteiler von m . Dann ist $m = \prod_{p \in \text{PF } m} p^{v_p(m)}$

²⁴. Wir werden nun zeigen, daß $p^{v_p(m)} \mid b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d}$ für jedes $p \in \text{PF } m$ gilt;

daraus wird natürlich $\prod_{p \in \text{PF } m} p^{v_p(m)} \mid b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d}$ folgen (denn die Zahlen $p^{v_p(m)}$

für verschiedene $p \in \text{PF } m$ sind alle teilerfremd, weil alle diese $p \in \text{PF } m$ prim sind), und damit $m \mid b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d}$, womit unser Ziel erreicht sein wird.

Wir haben angenommen, daß (18) für jedes $n \in \{1, 2, \dots, m-1\}$ erfüllt ist. Insbesondere ist also (18) für $n = m/p$ erfüllt (denn wegen $p \in \text{PF } m$ ist $m/p \in \mathbb{Z}$, also $m/p \in \{1, 2, \dots, m-1\}$); mit anderen Worten: $b_{m/p} = \sum_{d|(m/p)} dx_d^{(m/p)/d}$. Nun

²⁴Dies sieht man wie im Beweis von $\mathcal{A} \implies \mathcal{B}$.

zerlegen wir die Summe $\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d}$ in zwei Teilsommen:

$$\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d} = \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d|(m/p)}} dx_d^{m/d} + \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d \nmid (m/p)}} dx_d^{m/d}. \quad (20)$$

Die erste dieser Teilsommen ist

$$\begin{aligned} \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d|(m/p)}} dx_d^{m/d} &= \sum_{d \in \mathbb{N}_{|(m/p)}} dx_d^{m/d} && \left(\begin{array}{l} \text{denn die Teiler von } m, \text{ die kleiner als } m \text{ sind und } m/p \text{ teilen,} \\ \text{sind schlicht und einfach die Teiler von } m/p \end{array} \right) \\ &= \sum_{d|(m/p)} dx_d^{m/d}. \end{aligned} \quad (21)$$

In Analogie zum Beweis von $\mathcal{A} \implies \mathcal{D}$ würden wir nun erwarten, daß sich dies zu $b_{m/p}$ vereinfacht; dies wird aber diesmal schwieriger als im Beweis von $\mathcal{A} \implies \mathcal{D}$, und funktioniert nicht mehr in \mathbb{Z} , sondern nur modulo $p^{v_p(m)}$ (was für uns natürlich ausreicht). Und zwar haben wir $x_d^{(m/d)/p} \equiv x_d^{m/d} \pmod{p^{v_p(m/d)}}$ für jeden Teiler d von m/p (nach Lemma 2.2, angewandt auf $q = x_d$ und $n = m/d$), weil p ein Primfaktor von m/d ist (denn $\frac{m/d}{p} = \frac{m}{dp} = \frac{m/p}{d} \in \mathbb{Z}$, da $d \mid (m/p)$). Wegen $(m/d)/p = (m/p)/d$ läßt sich das zu $x_d^{(m/p)/d} \equiv x_d^{m/d} \pmod{p^{v_p(m/d)}}$ umschreiben. Das heißt, $p^{v_p(m/d)} \mid x_d^{(m/p)/d} - x_d^{m/d}$. Folglich ist $dp^{v_p(m/d)} \mid d(x_d^{(m/p)/d} - x_d^{m/d}) = dx_d^{(m/p)/d} - dx_d^{m/d}$. Doch $p^{v_p(m)} \mid dp^{v_p(m/d)}$.²⁵ Somit ist $p^{v_p(m)} \mid dx_d^{(m/p)/d} - dx_d^{m/d}$, also $dx_d^{m/d} \equiv dx_d^{(m/p)/d} \pmod{p^{v_p(m)}}$. Damit wird (21) zu

$$\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d|(m/p)}} dx_d^{m/d} = \sum_{\substack{d|(m/p) \\ \equiv dx_d^{(m/p)/d} \pmod{p^{v_p(m)}}}} \underbrace{dx_d^{m/d}}_{\equiv dx_d^{(m/p)/d} \pmod{p^{v_p(m)}}} \equiv \sum_{d|(m/p)} dx_d^{(m/p)/d} = b_{m/p} \pmod{p^{v_p(m)}}.$$

Damit haben wir die erste der beiden Teilsommen auf der rechten Seite von (20) vereinfacht (modulo $p^{v_p(m)}$ zumindest). Die zweite dieser Teilsommen ist

$$\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d \nmid (m/p)}} dx_d^{m/d} \equiv 0 \pmod{p^{v_p(m)}},$$

²⁵Denn

$$\begin{aligned} v_p\left(dp^{v_p(m/d)}\right) &= v_p(d) + v_p\left(\underbrace{p^{v_p(m/d)}}_{=v_p(m/d)}\right) && \text{(da } v_p(\alpha\beta) = v_p(\alpha) + v_p(\beta) \text{ für alle natürlichen Zahlen } \alpha \text{ und } \beta) \\ &= v_p(d) + v_p(m/d) \\ &= v_p(d \cdot m/d) && \text{(da } v_p(\alpha) + v_p(\beta) = v_p(\alpha\beta) \text{ für alle natürlichen Zahlen } \alpha \text{ und } \beta) \\ &= v_p(m). \end{aligned}$$

denn für jeden Teiler d von m , der $d \nmid (m/p)$ erfüllt, gilt $p^{v_p(m)} \mid d$ ²⁶ und damit $dx_d^{m/d} \equiv 0 \pmod{p^{v_p(m)}}$. Die Gesamtsumme ist somit

$$\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d} = \underbrace{\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d|(m/p)}} dx_d^{m/d}}_{\equiv b_{m/p} \pmod{p^{v_p(m)}}} + \underbrace{\sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m; \\ d \nmid (m/p)}} dx_d^{m/d}}_{\equiv 0 \pmod{p^{v_p(m)}}} \equiv b_{m/p} + 0 = b_{m/p} \equiv b_m \pmod{p^{v_p(m)}}$$

(nach Aussage \mathcal{A} , angewandt auf $n = m$, denn Aussage \mathcal{A} haben wir ja als wahr angenommen). Das heißt, $p^{v_p(m)} \mid b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d}$. Wie wir wissen, folgt hieraus

$m \mid b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d}$, und somit ist durch (19) tatsächlich eine ganze Zahl $x_m \in \mathbb{Z}$

definiert. Diese Zahl x_m erfüllt tatsächlich (18) für $n = m$, weil

$$\begin{aligned} \sum_{d|m} dx_d^{m/d} &= \sum_{d \in \mathbb{N}_{|m}} dx_d^{m/d} = m \underbrace{x_m^{m/m}}_{=x_m^1=x_m} + \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d} = mx_m + \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d} \\ &= m \cdot \frac{1}{m} \left(b_m - \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d} \right) + \sum_{\substack{d \in \mathbb{N}_{|m}; \\ d < m}} dx_d^{m/d} \quad (\text{nach (19)}) \\ &= b_m \end{aligned}$$

gilt.

Wir können also rekursiv eine Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ konstruieren, und diese Folge erfüllt (18) für jedes $n \in \mathbb{N}_+$. Damit ist Aussage \mathcal{B} bewiesen. Wir haben also gezeigt, daß $\mathcal{A} \implies \mathcal{B}$ ist.

Beweis von $\mathcal{B} \implies \mathcal{A}$: Nun wollen wir die Aussage \mathcal{B} als wahr annehmen. Es gibt also eine Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ ganzer Zahlen, die (5) erfüllt. Wir wollen nun Aussage \mathcal{A} herleiten.

In der Tat sei $n \in \mathbb{N}_+$ ganz, und p ein Primteiler von n . Nach (5) gilt dann $b_n = \sum_{d|n} dx_d^{n/d}$, und wenden wir (5) auf n/p statt n an, so erhalten wir $b_{n/p} = \sum_{d|(n/p)} dx_d^{(n/p)/d}$.

Nun zerlegen wir die Summe $\sum_{d|n} dx_d^{n/d} = \sum_{d \in \mathbb{N}_{|n}} dx_d^{n/d}$ in zwei Teilsommen:

$$\sum_{d \in \mathbb{N}_{|n}} dx_d^{n/d} = \sum_{\substack{d \in \mathbb{N}_{|n}; \\ d|(n/p)}} dx_d^{n/d} + \sum_{\substack{d \in \mathbb{N}_{|n}; \\ d \nmid (n/p)}} dx_d^{n/d}.$$

²⁶Dies zeigt sich wieder wie im Beweis von $\mathcal{A} \implies \mathcal{D}$.

Die erste dieser Teilsummen ist

$$\begin{aligned}
\sum_{\substack{d \in \mathbb{N}_n; \\ d|(n/p)}} dx_d^{n/d} &= \sum_{d \in \mathbb{N}_{(n/p)}} dx_d^{n/d} \quad \left(\begin{array}{l} \text{denn die Teiler von } n, \text{ die } n/p \text{ teilen,} \\ \text{sind schlicht und einfach die Teiler von } n/p \end{array} \right) \\
&= \sum_{d|(n/p)} dx_d^{n/d} \equiv \sum_{d|(n/p)} dx_d^{(n/p)/d} \\
&\quad \left(\begin{array}{l} \text{denn für jeden Teiler } d \text{ von } n/p \text{ gilt } dx_d^{n/d} \equiv dx_d^{(n/p)/d} \pmod{p^{v_p(n)}} \\ \text{(dies beweist man genauso, wie wir } dx_d^{m/d} \equiv dx_d^{(m/p)/d} \pmod{p^{v_p(m)}} \text{ im} \\ \text{Beweis von } \mathcal{A} \implies \mathcal{B} \text{ gezeigt haben)} \end{array} \right) \\
&= b_{n/p} \pmod{p^{v_p(n)}}.
\end{aligned}$$

Die zweite dieser Teilsummen ist

$$\sum_{\substack{d \in \mathbb{N}_n; \\ d \nmid (n/p)}} dx_d^{n/d} \equiv 0 \pmod{p^{v_p(n)}},$$

denn für jeden Teiler d von n , der $d \nmid (n/p)$ erfüllt, gilt $p^{v_p(n)} \mid d$ ²⁷ und damit $dx_d^{n/d} \equiv 0 \pmod{p^{v_p(n)}}$. Die Gesamtsumme ist somit

$$\begin{aligned}
\sum_{d \in \mathbb{N}_n} dx_d^{n/d} &= \underbrace{\sum_{\substack{d \in \mathbb{N}_n; \\ d|(n/p)}} dx_d^{n/d}}_{\equiv b_{n/p} \pmod{p^{v_p(n)}}} + \underbrace{\sum_{\substack{d \in \mathbb{N}_n; \\ d \nmid (n/p)}} dx_d^{n/d}}_{\equiv 0 \pmod{p^{v_p(n)}}} \equiv b_{n/p} + 0 = b_{n/p} \pmod{p^{v_p(n)}}.
\end{aligned}$$

Hieraus folgt

$$b_n = \sum_{d|n} dx_d^{n/d} = \sum_{d \in \mathbb{N}_n} dx_d^{n/d} \equiv b_{n/p} \pmod{p^{v_p(n)}},$$

und damit ist Aussage \mathcal{A} nachgewiesen. Wir haben damit gezeigt, daß $\mathcal{B} \implies \mathcal{A}$ ist.

Da wir nun $\mathcal{A} \implies \mathcal{B}$ und $\mathcal{B} \implies \mathcal{A}$ nachgewiesen haben, wissen wir also, daß $\mathcal{A} \iff \mathcal{B}$ ist. Damit haben wir die Aussage \mathcal{B} an die Äquivalenzkette $\mathcal{A} \iff \mathcal{D} \iff \mathcal{E} \iff \mathcal{F} \iff \mathcal{G} \iff \mathcal{H}$ angeschlossen. Es gibt auch eine alternative Möglichkeit, diesen Anschluss zu vollziehen - nämlich durch direkten Beweis von $\mathcal{B} \iff \mathcal{D}$ (ohne Umweg durch \mathcal{A}); diese Möglichkeit werden wir später (in Abschnitt 3) vorstellen.

2.9. Beweis von Satz 2.1: $\mathcal{B} \iff \mathcal{C}$

Jetzt fehlt nur noch Aussage \mathcal{C} . Wir werden beweisen, daß $\mathcal{B} \iff \mathcal{C}$ ist. In der Tat ist $\mathcal{C} \implies \mathcal{B}$ klar, und zum *Beweis von $\mathcal{B} \implies \mathcal{C}$* müssen wir nur zeigen, daß je zwei Folgen $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}^+}$, die beide (5) erfüllen, notwendigerweise gleich sein müssen. Dazu bezeichnen wir diese zwei Folgen mit (x_1, x_2, x_3, \dots) und $(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \dots)$. Da sie beide (5) erfüllen, gilt

$$\left(b_n = \sum_{d|n} dx_d^{n/d} \text{ für jedes } n \in \mathbb{N}_+ \right) \text{ und } \left(b_n = \sum_{d|n} d\tilde{x}_d^{n/d} \text{ für jedes } n \in \mathbb{N}_+ \right).$$

²⁷Dies läßt sich genauso beweisen, wie wir im Beweis von $\mathcal{A} \implies \mathcal{D}$ gezeigt haben, daß für jeden Teiler d von m , der $d \nmid (m/p)$ erfüllt, $p^{v_p(m)} \mid d$ gilt. Wir müssen nur im Beweis immer m durch n ersetzen.

Wir beweisen nun durch starke Induktion nach n , daß $x_n = \tilde{x}_n$ für alle $n \in \mathbb{N}_+$ gilt. Dazu sei $n \in \mathbb{N}_+$ beliebig; wir nehmen an, daß wir $x_m = \tilde{x}_m$ für alle $m \in \{1, 2, \dots, n-1\}$ bereits bewiesen haben. Wir müssen dann $x_n = \tilde{x}_n$ zeigen. Wir haben

$$b_n = \sum_{d|n} dx_d^{n/d} = \sum_{\substack{d \in \mathbb{N}_{|n} \\ d < n}} dx_d^{n/d} = \sum_{\substack{d \in \mathbb{N}_{|n} \\ d < n}} dx_d^{n/d} + n \underbrace{x_n^{n/n}}_{=x_n^1=x_n} = \sum_{\substack{d \in \mathbb{N}_{|n} \\ d < n}} dx_d^{n/d} + nx_n$$

und analog

$$b_n = \sum_{\substack{d \in \mathbb{N}_{|n} \\ d < n}} d\tilde{x}_d^{n/d} + n\tilde{x}_n,$$

also

$$\sum_{\substack{d \in \mathbb{N}_{|n} \\ d < n}} dx_d^{n/d} + nx_n = \sum_{\substack{d \in \mathbb{N}_{|n} \\ d < n}} d\tilde{x}_d^{n/d} + n\tilde{x}_n. \quad (22)$$

Doch für jedes $d \in \mathbb{N}_{|n}$ mit $d < n$ gilt $x_d = \tilde{x}_d$ (wegen unserer Induktionsannahme, daß $x_m = \tilde{x}_m$ für alle $m \in \{1, 2, \dots, n-1\}$ bereits bewiesen ist). Also ist $\sum_{\substack{d \in \mathbb{N}_{|n} \\ d < n}} dx_d^{n/d} =$

$\sum_{\substack{d \in \mathbb{N}_{|n} \\ d < n}} d\tilde{x}_d^{n/d}$, und aus (22) folgt somit $nx_n = n\tilde{x}_n$, also $x_n = \tilde{x}_n$. Damit ist der

Induktionsbeweis von $x_n = \tilde{x}_n$ für alle $n \in \mathbb{N}_+$ komplett. Folglich sind die beiden Folgen (x_1, x_2, x_3, \dots) und $(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \dots)$ gleich, und damit ist $\mathcal{B} \implies \mathcal{C}$ bewiesen. Da trivialerweise $\mathcal{C} \implies \mathcal{B}$ gilt, ist also $\mathcal{B} \iff \mathcal{C}$. Zusammen mit $\mathcal{A} \iff \mathcal{B}$ und $\mathcal{A} \iff \mathcal{D} \iff \mathcal{E} \iff \mathcal{F} \iff \mathcal{G} \iff \mathcal{H}$ führt dies auf $\mathcal{A} \iff \mathcal{B} \iff \mathcal{C} \iff \mathcal{D} \iff \mathcal{E} \iff \mathcal{F} \iff \mathcal{G} \iff \mathcal{H}$, und Satz 2.1 ist bewiesen.

(Übrigens könnten wir genauso, wie wir gerade $\mathcal{B} \iff \mathcal{C}$ bewiesen haben, auch $\mathcal{D} \iff \mathcal{E}$ zeigen; aber wir haben $\mathcal{D} \iff \mathcal{E}$ bereits anders nachgewiesen.)

Anmerkungen: Unser obiger Beweis der Äquivalenz zwischen den Aussagen \mathcal{A} , \mathcal{B} und \mathcal{C} ist mehr oder weniger der gleiche wie in [3]²⁸. Auch die Äquivalenz $\mathcal{A} \iff \mathcal{D} \iff \mathcal{E} \iff \mathcal{F} \iff \mathcal{G} \iff \mathcal{H}$ haben wir hier etwa genauso wie in [3] bewiesen, wenn auch in [3] der Begriff der Dirichlet-Faltung vermieden wurde²⁹.

3. Potenzreihen und ein alternativer Beweis von

$$\mathcal{B} \iff \mathcal{D}$$

Was wir oben in Abschnitt 2 gemacht haben, hat zum Beweis von Satz 2.1 (und damit auch Satz 1.1) gereicht, aber es läßt Wünsche offen: In den Aussagen \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{G} und \mathcal{H} von Satz 2.1 kommt der Begriff einer Primzahl nie vor (und auch in Aussage \mathcal{F} nur indirekt über die Definition der Möbiusfunktion μ), doch zum Beweis der Äquivalenz dieser Aussagen sind wir den Umweg über Aussage \mathcal{A} gegangen, in der Primzahlen maßgeblich sind. Man kann sich fragen, ob es nicht einen anderen Weg gibt, die Äquivalenz $\mathcal{B} \iff \mathcal{C} \iff \mathcal{D} \iff \mathcal{E} \iff \mathcal{H}$ zu beweisen, ohne die Aussage \mathcal{A}

²⁸In der Tat wurde in [3] ein allgemeinerer Sachverhalt gezeigt (Theorem 4 in [3]), aus dem die Äquivalenz zwischen den Aussagen \mathcal{A} , \mathcal{B} und \mathcal{C} in unserem Fall durch Anwendung auf $A = \mathbb{Z}$, $N = \mathbb{N}_+$ und $\varphi_p = \text{id}$ für alle p folgt.

²⁹Auch in diesem Fall wurde in [3] etwas Allgemeineres gezeigt (Theorem 5 in [3]).

(und damit die Primfaktorzerlegung von n) als Zwischenschritt heraufzubeschwören. In der Tat gibt es einen solchen Beweis, und er ist in meinen Augen interessant genug, daß er bessere Bekanntheit verdient, als eine (nur indirekte) Erwähnung in den Tiefen (Abschnitt 17) eines algebraischen Übersichtsartikels [1]. Dieser Beweis verwendet den Begriff von *Potenzreihen* (genauer gesagt, *formalen Potenzreihen*); wir wollen eine Einführung in diesen Begriff geben. Wir werden also in den folgenden Abschnitten definieren, was eine Potenzreihe sind, was die Summe und das Produkt von zwei Potenzreihen sind, was die Ableitung und der Logarithmus einer Potenzreihe sind, usw. Diese Begriffe stammen alle historisch aus der Analysis, wo sie ursprünglich den Zweck hatten, Funktionen (wie die Exponentialfunktion \exp , oder die trigonometrischen Funktionen \sin und \cos , und allgemein jede reell-analytische Funktion) annähernd zu berechnen. Aber in der Form, wie wir sie hier einführen, sind Potenzreihen ein von der Analysis völlig unabhängiges Konzept - wir werden sie nur als Kurzschreibweise für Zahlenfolgen betrachten, und nicht als unendliche Summen, in die man wirklich irgendeine Zahl "einsetzen" könnte.³⁰ Infolgedessen müssen wir uns auch so gut wie um Konvergenzfragen kümmern, außer wenn wir unendliche Summen von Potenzreihen betrachten - und auch dann ist es eine formale, von der Analysis losgelöste Art von Konvergenz, die (im Gegensatz zu der Analysis) meistens sehr leicht nachzuprüfen ist.

Unsere Einführung in Potenzreihen wird weitgehend in sich abgeschlossen sein. Um allerdings zu einem tiefergreifenden Verständnis der Materie zu kommen, sollte man Beispiele für die Anwendung von Potenzreihen kennengelernt haben. Davon werden wir hier recht wenige geben; in der Hinsicht sei der Leser auf [8] verwiesen (ein ganzes Buch über Potenzreihen und ihre Anwendung in der Kombinatorik) und auch auf [16] (ein Standardwerk zur abzählenden Kombinatorik, welches momentan neu aufgelegt wird - und welches Potenzreihen in Abschnitt 1.1 einführt und sie ab dann immer wieder verwendet).³¹ Beginnen wir erst einmal mit der Definition des Begriffes einer *Potenzreihe*.

3.1. Was sind Potenzreihen?

Bildlich gesprochen handelt es sich bei Potenzreihen um eine Art "Polynome", die aber unendlich viele Koeffizienten haben dürfen. Ein Polynom in einer Variable X über den ganzen Zahlen hat immer die Form $\sum_{i=0}^n a_i X^i$ für ein bestimmtes $n \in \mathbb{N}$ und ganze Koeffizienten a_0, a_1, \dots, a_n . Eine Potenzreihe in einer Variable X über den ganzen Zahlen hat hingegen die Form $\sum_{i=0}^{\infty} a_i X^i$ für ganze Koeffizienten a_0, a_1, a_2, \dots . Es ist wichtig, daran zu denken, daß man in Potenzreihen - im Allgemeinen - keine Zahlen einsetzen kann (d. h. man kann nicht einfach 2 für X in $\sum_{i=0}^{\infty} a_i X^i$ einsetzen und hoffen, daß der entstehende Ausdruck $\sum_{i=0}^{\infty} a_i 2^i$ Sinn macht), im Gegensatz zu Polynomen. Die Analysis bietet eine Reihe³² von Kriterien, wann Potenzreihen beim Einsetzen von

³⁰Wir werden allerdings lernen, daß man in eine Potenzreihe eine andere Potenzreihe einsetzen kann - solange letztere bestimmte Eigenschaften hat.

³¹Auch eine Google-Suche nach dem Begriff "*generating function*" (zu deutsch: "*erzeugende Funktion*" - der Fachbegriff für die Darstellung einer Zahlenfolge in Form einer Potenzreihe) sollte viel ans Licht bringen.

³²Verzeihung...

Zahlen doch noch konvergente Reihen ergeben; doch wir interessieren uns dafür nicht, sondern wir wollen mit Potenzreihen *formal rechnen*, ohne in sie irgendwelche Zahlen einzusetzen.

Wir müssen allerdings zuerst klären, was eine Potenzreihe überhaupt ist. Um eine formale Definition einer Potenzreihe zu bekommen, müssen wir eine Potenzreihe durch die Folge ihrer Koeffizienten kodieren (wie das auch bei Polynomen gemacht wird!). Hier eine Definition:

Definition (ganzzahlige (formale) Potenzreihe): Unter einer *ganzzahligen formalen Potenzreihe in der Variable X* (oder einfach *ganzzahligen Potenzreihe in der Variable X*) verstehen wir, abstrakt gesprochen, eine Folge $(a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ ganzer Zahlen. Es ist zu beachten, daß wir solche Folgen $(a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ (deren Folgenglieder mit $0, 1, 2, \dots$ durchnummeriert sind) **nicht** mit Zahlenfunktionen identifizieren, im Gegensatz zu Folgen der Form $(a_1, a_2, a_3, \dots) \in \mathbb{Z}^{\mathbb{N}^+}$ (deren Folgenglieder mit $1, 2, 3, \dots$ durchnummeriert sind). Dies mag nach einer begrifflichen Lappalie aussehen, aber es ist wichtig, denn wir werden für Folgen $(a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ ein Produkt einführen, das nichts mit dem punktweisen Produkt von Zahlenfunktionen zu tun hat, und dieses Produkt mit \cdot bezeichnen; um das Produkt nicht mit dem (ebenfalls \cdot genannten) punktweisen Produkt von Zahlenfunktionen zu verwechseln, ist es wichtig, daß wir Folgen $(a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ nicht mit Zahlenfunktionen gleichsetzen.

Statt von einer "ganzzahligen Potenzreihe in der Variable X " reden wir im Folgenden kurz von einer "ganzzahligen Potenzreihe", da wir hier nur mit einer Variablen arbeiten. Solange wir nur ganzzahlige (und nicht etwa rationale, oder noch andere) Potenzreihen kennen, werden wir auch kurz einfach "Potenzreihe" statt "ganzzahlige Potenzreihe" schreiben.

Definition (Koeffizient einer Potenzreihe): Für jede ganzzahlige Potenzreihe $(a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ bezeichnen wir die Folgenglieder a_0, a_1, a_2, \dots als *Koeffizienten* der Potenzreihe (a_0, a_1, a_2, \dots) . Genauer gesagt: Für jedes $i \in \mathbb{N}$ bezeichnen wir die Zahl a_i als den *i -ten Koeffizienten der Potenzreihe* (a_0, a_1, a_2, \dots) . Also ist "Koeffizient" nur ein anderes Wort für "Folgenglied". Insbesondere: Wenn zwei Potenzreihen gleich sind, dann sind alle ihre entsprechenden Koeffizienten gleich.

Definition (Summe zweier Potenzreihen): Wir definieren die Summe $a + b$ zweier Potenzreihen (d. h. Folgen) $a = (a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ und $b = (b_0, b_1, b_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ als die Folge

$$(a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots) \in \mathbb{Z}^{\mathbb{N}}.$$

(Das n -te Glied dieser Folge ist $a_n + b_n$ für jedes $n \in \mathbb{N}$.)

Definition (Produkt zweier Potenzreihen): Wir definieren das Produkt $a \cdot b$ zweier Potenzreihen (d. h. Folgen) $a = (a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ und $b = (b_0, b_1, b_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ als die Folge

$$(a_0b_0, a_0b_1 + a_1b_0, a_0b_2 + a_1b_1 + a_2b_0, \dots) \in \mathbb{Z}^{\mathbb{N}}.$$

(Das n -te Glied dieser Folge ist $\sum_{k=0}^n a_k b_{n-k}$ für jedes $n \in \mathbb{N}$.) Statt $a \cdot b$ schreiben wir auch ab . Die Addition $+$ und die Multiplikation \cdot von Potenzreihen (also Folgen in $\mathbb{Z}^{\mathbb{N}}$) sind kommutativ, assoziativ und distributiv.

Wie gesagt, hat das Produkt von zwei Folgen in $\mathbb{Z}^{\mathbb{N}}$ nichts zu tun mit dem punktwweisen Produkt von Zahlenfunktionen. Es ist leicht zu sehen, daß für die Addition und die Multiplikation von Potenzreihen alle gängigen Gesetze (die Assoziativgesetze, die Distributivgesetze und die Kommutativgesetze) gelten.

Definition (Potenzreihen $\mathbf{0}$ und $\mathbf{1}$): Die durch $\mathbf{0} = (0, 0, 0, \dots)$ definierte Folge $\mathbf{0} \in \mathbb{Z}^{\mathbb{N}}$ erfüllt $\mathbf{0} + a = a + \mathbf{0} = a$ für jede Folge $a \in \mathbb{Z}^{\mathbb{N}}$, und die durch $\mathbf{1} = (1, 0, 0, 0, \dots)$ (alle Glieder bis auf das erste sind Null) definierte Folge $\mathbf{1} \in \mathbb{Z}^{\mathbb{N}}$ erfüllt $\mathbf{1}a = a\mathbf{1} = a$ für jede Folge $a \in \mathbb{Z}^{\mathbb{N}}$.

Wir werden öfters die Folge $\mathbf{0} = (0, 0, 0, \dots)$ auch einfach mit 0 bezeichnen (ohne Fettdruck), und die Folge $\mathbf{1} = (1, 0, 0, 0, \dots)$ auch einfach mit 1 (ebenfalls ohne Fettdruck). Allgemeiner bezeichnen wir für jedes $a \in \mathbb{Z}$ die Potenzreihe $(a, 0, 0, 0, \dots)$ (alle Glieder ab dem zweiten sind Null) mit a .

Definition (Multiplikation einer Potenzreihe mit einer ganzen Zahl):

Ferner definieren wir für jede ganze Zahl λ und jede Potenzreihe (d. h. Folge) $a = (a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ eine Potenzreihe $\lambda \cdot a \in \mathbb{Z}^{\mathbb{N}}$ als die Folge

$$(\lambda a_0, \lambda a_1, \lambda a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}.$$

Diese Folge wird auch als λa bezeichnet.

Definition (Potenzreihe X): Wir bezeichnen ferner mit $X \in \mathbb{Z}^{\mathbb{N}}$ die Folge $(0, 1, 0, 0, 0, \dots)$ (alle Glieder bis auf das zweite sind Null).

Für jedes $i \in \mathbb{N}$ ist dann X^i die Folge $(0, 0, \dots, 0, 1, 0, 0, \dots)$, deren i -tes Glied 1 und alle anderen Glieder 0 sind. (Dabei ist X^i definiert als i -te Potenz von X bezüglich der Multiplikation \cdot von Folgen; das heißt, $X^i = \underbrace{X \cdot X \cdot \dots \cdot X}_{i \text{ mal}}$.)

Unser Ziel ist nun, die Folge $(a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ in der Form $\sum_{i=0}^{\infty} a_i X^i$ schreiben und damit rechnen zu dürfen. Dazu müssen wir klarmachen, was eine unendliche Summe wie $\sum_{i=0}^{\infty} a_i X^i$ überhaupt bedeutet. Wir werden sie als Grenzwert $\lim_{n \rightarrow \infty} \sum_{i=0}^n a_i X^i$ deuten.

Doch dazu müssen wir erstmal den Begriff der Konvergenz einer Folge von Folgen (!) einführen. Er hat eine gewisse Ähnlichkeit mit dem Begriff der Konvergenz einer Folge von rationalen Zahlen: Während die Konvergenz einer Folge von rationalen Zahlen zu einem Grenzwert anschaulich gesehen bedeutet, daß diese Zahlen "immer näher" an den Grenzwert werden und beliebig nah an ihn herankommen, bedeutet die Konvergenz einer Folge von Folgen zu einer Grenzfolge, daß diese Folgen "immer gleicher" zur Grenzfolge werden (also immer größere Anfangsabschnitte gleich werden) und "beliebig gleich" zu ihr werden. Formal gesehen ist die Definition folgende:

Definition (Konvergenz einer Folge von Folgen/Potenzreihe zu einer Folge/Potenzreihe): Sei M eine Menge. Man sagt, eine Folge

$(\alpha_0, \alpha_1, \alpha_2, \dots) \in (M^{\mathbb{N}})^{\mathbb{N}}$ von Folgen von Elementen von M konvergiere zu einer Folge $\beta \in M^{\mathbb{N}}$, wenn folgendes gilt:

Für jedes $g \in \mathbb{N}$ gibt es ein $N \in \mathbb{N}$, sodaß für jedes $n \in \mathbb{N}$ mit $n \geq N$ gilt:

(Die Folgen β und α_n unterscheiden sich nicht vor ihrem g -ten Glied), d. h.

(für jedes $k \in \{0, 1, \dots, g-1\}$ ist das k -te Glied der Folge β gleich dem der Folge α_n).

In diesem Fall schreiben wir $\beta = \lim_{n \rightarrow \infty} \alpha_n$ und bezeichnen β als *Grenzwert* oder *Grenzfolge* der Folge $(\alpha_0, \alpha_1, \alpha_2, \dots)$. Es ist leicht zu sehen, daß eine konvergente Folge immer nur zu einem Grenzwert (und nicht etwa zu zwei verschiedenen) konvergieren kann.

Wir wollen nun ein wichtiges Beispiel für eine konvergente Folge von Folgen geben: Ist $(a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ eine Folge, dann konvergiert die Folge

$$\begin{aligned} & ((a_0, 0, 0, 0, 0, 0, \dots), \\ & (a_0, a_1, 0, 0, 0, 0, \dots), \\ & (a_0, a_1, a_2, 0, 0, 0, \dots), \\ & (a_0, a_1, a_2, a_3, 0, 0, \dots), \\ & \dots) \in (\mathbb{Z}^{\mathbb{N}})^{\mathbb{N}} \end{aligned}$$

gegen $(a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$. Wir haben also $(a_0, a_1, a_2, \dots) = \lim_{n \rightarrow \infty} (a_0, a_1, \dots, a_n, 0, 0, 0, 0, \dots)$.
Wegen

$$\begin{aligned} (a_0, a_1, \dots, a_n, 0, 0, 0, 0, \dots) &= \sum_{i=0}^n \underbrace{(0, 0, \dots, 0, a_i, 0, 0, \dots)}_{\substack{\text{die Folge, deren } i\text{-tes Glied } a_i \text{ ist} \\ \text{und alle anderen Glieder 0}}} \\ &= \sum_{i=0}^n a_i \underbrace{(0, 0, \dots, 0, 1, 0, 0, \dots)}_{\substack{\text{die Folge, deren } i\text{-tes Glied 1 ist} \\ \text{und alle anderen Glieder 0}}} = \sum_{i=0}^n a_i X^i \end{aligned}$$

können wir dies als $(a_0, a_1, a_2, \dots) = \lim_{n \rightarrow \infty} \sum_{i=0}^n a_i X^i$ schreiben. Wenn wir jetzt noch die

Abkürzung $\sum_{i=0}^{\infty} p_i$ für $\lim_{n \rightarrow \infty} \sum_{i=0}^n p_i$ (wobei $p_i \in \mathbb{Z}^{\mathbb{N}}$ für jedes $i \in \mathbb{N}$ ist) einführen, haben

wir es geschafft: Wir erhalten $(a_0, a_1, a_2, \dots) = \sum_{i=0}^{\infty} a_i X^i$. Da wir den Summationsindex

beliebig umbenennen dürfen, bedeutet dies auch $(a_0, a_1, a_2, \dots) = \sum_{j=0}^{\infty} a_j X^j$ und

$(a_0, a_1, a_2, \dots) = \sum_{n=0}^{\infty} a_n X^n$ und so weiter. Statt $\sum_{i=0}^{\infty} a_i X^i$ schreibt man auch $\sum_{i \in \mathbb{N}} a_i X^i$

oder $a_0 + a_1 X + a_2 X^2 + \dots$ (aber die Notation $a_0 + a_1 X + a_2 X^2 + \dots$ sollte man vermeiden, wenn die Folgenglieder a_0, a_1, a_2, \dots kompliziert aussehen; stattdessen ist in diesem Fall die Notation $\sum_{i=0}^{\infty} a_i X^i$ oder $\sum_{i \in \mathbb{N}} a_i X^i$ vorzuziehen).

Wenn wir von Potenzreihen reden, wollen wir nach Möglichkeit immer die Notation $\sum_{i=0}^{\infty} a_i X^i$ benutzen, und nicht die Notation (a_0, a_1, a_2, \dots) . Der Grund ist, daß

die Rechenregeln für Potenzreihen in der Notation $\sum_{i=0}^{\infty} a_i X^i$ wesentlich klarer sind als in der Notation (a_0, a_1, a_2, \dots) . So haben wir das Produkt $a \cdot b$ zweier Potenzreihen $a = (a_0, a_1, a_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ und $b = (b_0, b_1, b_2, \dots) \in \mathbb{Z}^{\mathbb{N}}$ definiert als die Potenzreihe

$$(a_0 b_0, a_0 b_1 + a_1 b_0, a_0 b_2 + a_1 b_1 + a_2 b_0, \dots) \in \mathbb{Z}^{\mathbb{N}},$$

deren n -tes Glied $\sum_{k=0}^n a_k b_{n-k}$ ist für jedes $n \in \mathbb{N}$. Diese Definition wirkt ein wenig unanschaulich, aber in der $\sum_{i=0}^{\infty} a_i X^i$ -Notation bedeutet sie: Das Produkt $a \cdot b$ zweier Potenzreihen $a = \sum_{i=0}^{\infty} a_i X^i$ und $b = \sum_{i=0}^{\infty} b_i X^i$ ist definiert als die Potenzreihe $\sum_{n=0}^{\infty} \sum_{k=0}^n a_k b_{n-k} X^n$. Dies ist aber genau das Ergebnis, was wir erhalten, wenn wir (nach dem Distributivgesetz) die Klammern auflösen:

$$\begin{aligned} a \cdot b &= \left(\sum_{i=0}^{\infty} a_i X^i \right) \cdot \left(\sum_{i=0}^{\infty} b_i X^i \right) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \underbrace{a_i X^i \cdot b_j X^j}_{=a_i b_j X^{i+j}} = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_i b_j X^{i+j} \\ &= \sum_{i=0}^{\infty} \sum_{n=i}^{\infty} a_i b_{n-i} X^n \quad (\text{hier haben wir } i+j \text{ durch } n \text{ substituiert}) \\ &= \sum_{n=0}^{\infty} \sum_{i=0}^n a_i b_{n-i} X^n = \sum_{n=0}^{\infty} \sum_{k=0}^n a_k b_{n-k} X^n \quad (\text{hier haben wir den Index } i \text{ in } k \text{ umbenannt}). \end{aligned}$$

In der $\sum_{i=0}^{\infty} a_i X^i$ -Notation bekommt also das (recht seltsam definierte) Produkt zweier Potenzreihen eine ganz natürliche Erklärung: Man erhält das Produkt $a \cdot b$ zweier Potenzreihen a und b , indem man diese Potenzreihen a und b in der Form $\sum_{i=0}^{\infty} a_i X^i$ bzw. $\sum_{i=0}^{\infty} b_i X^i$ darstellt, und im Produkt $\left(\sum_{i=0}^{\infty} a_i X^i \right) \cdot \left(\sum_{i=0}^{\infty} b_i X^i \right)$ (nach dem Distributivgesetz) die Klammern auflöst.

Da die Potenzreihe $\sum_{i=0}^{\infty} a_i X^i$ gleich der Folge (a_0, a_1, a_2, \dots) ist, sind die Koeffizienten der Potenzreihe $\sum_{i=0}^{\infty} a_i X^i$ einfach die Zahlen a_0, a_1, a_2, \dots . Somit gilt: Sind zwei Potenzreihen $\sum_{i=0}^{\infty} a_i X^i$ und $\sum_{i=0}^{\infty} b_i X^i$ einander gleich, dann gilt $a_i = b_i$ für jedes $i \in \mathbb{N}$ (denn wenn zwei Potenzreihen gleich sind, dann sind alle ihre entsprechenden Koeffizienten gleich).

Eine *Warnung* für alle, die den Begriff der Potenzreihen aus der Funktionentheorie kennen: Unser Begriff von "Potenzreihen" ist in vielerlei Hinsicht ähnlich zum Begriff "Potenzreihen (um den Punkt 0)" in der Funktionentheorie, und oftmals auch einfacher zu benutzen als letzterer (so muß man sich um Konvergenz der formalen Potenzreihen selber keine Gedanken machen, und auch die Konvergenz von Folgen von formalen Potenzreihen ist meist viel einfacher zu zeigen als Konvergenz von Folgen von holomorphen Funktionen in der Funktionentheorie). Allerdings hat die Einfachheit einen Preis: So kann man eine formale Potenzreihe nicht "um einen anderen Punkt" entwickeln (so wie man in der Funktionentheorie z. B. die Exponentialfunktion $z \mapsto \exp z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$ auch um den Punkt 1 herum entwickeln kann und die Funktion

$z \mapsto \exp z = \sum_{k=0}^{\infty} \frac{(z+1)^k}{ek!}$ erhält), und auch eine Potenzreihe nicht immer in eine andere Potenzreihe einsetzen (aber letzteres geht zumindest, wenn der 0-te Koeffizient der ersteren Potenzreihe gleich 0 ist - siehe Satz 3.4a).

3.2. Die binomische Formel und die Vandermonde-Faltungsformel: Eine erste Anwendung von Potenzreihen

Wozu nun das Ganze? Man stellt fest, daß man mit Potenzreihen auf elegante Weise rechnen kann, und die entstandenen Ergebnisse wieder potenzreihenfrei umformulieren und gut benutzen kann. Hier ein Beispiel, das uns später noch nützlich sein wird:

Satz 3.1 (die binomische Formel): Für jedes $n \in \mathbb{Z}$ ist die Potenzreihe

$$(1 + X)^n \text{ gleich } \sum_{k=0}^{\infty} \binom{n}{k} X^k.$$

Hierzu sind ein paar erklärende Worte angesagt:

1) Viele kennen den Binomialkoeffizienten $\binom{n}{k}$ nur für $n \in \mathbb{N}$ und $k \in \{0, 1, \dots, n\}$

(dies ist auch der gebräuchlichste Fall für die Verwendung von $\binom{n}{k}$); in diesem Fall ist

$\binom{n}{k}$ nämlich die Anzahl aller k -elementigen Teilmengen einer gegebenen n -elementigen

Menge). Doch man kann $\binom{n}{k}$ genauso für beliebige ganze, rationale, reelle oder sogar komplexe Zahlen n und $k \in \mathbb{Z}$ definieren: nämlich folgendermaßen:

Definition (Binomialkoeffizient $\binom{n}{k}$ mit $n \in \mathbb{Z}$ und $k \in \mathbb{N}$): Wir

definieren den sogenannten *Binomialkoeffizienten* $\binom{n}{k}$ durch

$$\binom{n}{k} = \begin{cases} \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!}, & \text{wenn } k \geq 0; \\ 0, & \text{wenn } k < 0 \end{cases}$$

für alle $n \in \mathbb{Z}$ (oder $n \in \mathbb{Q}$, oder $n \in \mathbb{R}$, oder $n \in \mathbb{C}$) und $k \in \mathbb{Z}$. Diese Definition sorgt dafür, daß die Rekursionsgleichung $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$ für alle $n \in \mathbb{Z}$ (oder $n \in \mathbb{Q}$, etc.) und $k \in \mathbb{Z}$ gilt (der Beweis ist sehr leicht).

33

³³ *Warnung:* Diese Definition von $\binom{n}{k}$ für $n \in \mathbb{Z}$ und $k \in \mathbb{Z}$ ist nicht die einzige, die in der Literatur vorkommt, und manche benutzen Definitionen, die zu anderen Werten führen! (Und zwar sind die Werte von $\binom{n}{k}$ für negatives k kontrovers. Die Werte von $\binom{n}{k}$ für $k \geq 0$ sind unumstritten - jeder, der den Binomialkoeffizienten $\binom{n}{k}$ für $n \in \mathbb{Z}$ und $k \geq 0$ überhaupt definiert, definiert

Es sei angemerkt, daß $\binom{n}{k} \in \mathbb{Z}$ für alle $n \in \mathbb{Z}$ und $k \in \mathbb{N}$ gibt (mit obiger Definition von $\binom{n}{k}$). Dies werden wir weiter unten (Satz 3.3 (b)) beweisen.

2) Was bedeutet der Ausdruck $(1+X)^n$ für $n \in \mathbb{Z}$? Für $n \geq 0$ ist dies klar. Für $n < 0$ ist dieser Ausdruck nur dann sinnvoll, wenn die Potenzreihe $1+X$ ein multiplikatives Inverses hat. Aber es stellt sich fest: Ja, die Potenzreihe $1+X$ hat ein multiplikatives Inverses, nämlich die Potenzreihe $1-X+X^2-X^3+\dots = \sum_{k=0}^{\infty} (-1)^k X^k$.

Denn

$$\begin{aligned} (1+X) \cdot (1-X+X^2-X^3+\dots) &= (1-X+X^2-X^3+\dots) + X \cdot (1-X+X^2-X^3+\dots) \\ &= (1-X+X^2-X^3+\dots) + (X-X^2+X^3-X^4+\dots) = 1, \end{aligned}$$

weil sich alle Terme bis auf 1 gegenseitig wegkürzen³⁴.

Natürlich ist Satz 3.1 für $n \geq 0$ einfach nur die altbekannte binomische Formel, denn $\sum_{k=0}^{\infty} \binom{n}{k} X^k = \sum_{k=0}^n \binom{n}{k} X^k$ (weil $\binom{n}{k} = 0$ für alle $k > n$). Interessant ist für uns eher der Fall $n < 0$, in dem $\sum_{k=0}^{\infty} \binom{n}{k} X^k$ eine echte unendliche Potenzreihe ist.

Nun zum *Beweis von Satz 3.1*: Wir zeigen erstmal, daß für jedes $n \in \mathbb{Z}$ gilt:

$$\sum_{k=0}^{\infty} \binom{n+1}{k} X^k = (1+X) \cdot \left(\sum_{k=0}^{\infty} \binom{n}{k} X^k \right). \quad (23)$$

ihn als $\frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!}$. Aber für $k < 0$ wird $\binom{n}{k}$ nicht von allen Autoren als 0 festgelegt; manche ziehen andere Werte vor (die dann allerdings zur Folge haben, daß $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$ nicht mehr immer gilt!). Unsere Konvention ($\binom{n}{k} = 0$ für $k < n$) hat meines Erachtens die meisten Vorteile, und sie wird auch von Knuth in [13] und von Wilf in [8] benutzt.)
³⁴Formaler ausgedrückt sieht dieser Beweis folgendermaßen aus:

$$\begin{aligned} (1+X) \cdot \sum_{k=0}^{\infty} (-1)^k X^k &= \sum_{k=0}^{\infty} (-1)^k X^k + X \cdot \sum_{k=0}^{\infty} (-1)^k X^k = \sum_{k=0}^{\infty} (-1)^k X^k + \sum_{k=0}^{\infty} (-1)^k X^{k+1} \\ &= \sum_{k=0}^{\infty} (-1)^k X^k + \sum_{k=1}^{\infty} (-1)^{k-1} X^k \\ &\quad \text{(hier haben wir in der zweiten Summe } k \text{ durch } k-1 \text{ substituiert)} \\ &= \sum_{k=0}^{\infty} (-1)^k X^k + \left(\sum_{k=0}^{\infty} (-1)^{k-1} X^k - (-1)^{0-1} X^0 \right) = \left(\sum_{k=0}^{\infty} (-1)^k X^k + \sum_{k=0}^{\infty} (-1)^{k-1} X^k \right) - \underbrace{(-1)^{0-1}}_{=-1} \underbrace{X^0}_{=1} \\ &= \sum_{k=0}^{\infty} \underbrace{\left((-1)^k X^k + (-1)^{k-1} X^k \right)}_{=(-1)^k + (-1)^{k-1}} X^k - (-1) = 0 + 1 = 1. \end{aligned}$$

In der Tat ist

$$\begin{aligned}
(1 + X) \cdot \left(\sum_{k=0}^{\infty} \binom{n}{k} X^k \right) &= \sum_{k=0}^{\infty} \binom{n}{k} X^k + X \cdot \sum_{k=0}^{\infty} \binom{n}{k} X^k \\
&= \sum_{k=0}^{\infty} \binom{n}{k} X^k + \sum_{k=0}^{\infty} \binom{n}{k} X^{k+1} = \sum_{k=0}^{\infty} \binom{n}{k} X^k + \sum_{k=1}^{\infty} \binom{n}{k-1} X^k \\
&\quad \text{(hier haben wir } k-1 \text{ für } k \text{ in der zweiten Summe substituiert)} \\
&= \sum_{k=0}^{\infty} \binom{n}{k} X^k + \sum_{k=0}^{\infty} \binom{n}{k-1} X^k \\
&\quad \left(\begin{array}{l} \text{hier haben wir die zweite Summe um den Summanden für } k=0 \\ \text{erweitert, was nichts ändert, da dieser Summand 0 ist (denn } \binom{n}{0-1} = 0 \text{)} \end{array} \right) \\
&= \sum_{k=0}^{\infty} \underbrace{\left(\binom{n}{k} + \binom{n}{k-1} \right)}_{= \binom{n+1}{k}} X^k = \sum_{k=0}^{\infty} \binom{n+1}{k} X^k.
\end{aligned}$$

Aus (23) und $(1 + X)^0 = 1 = \sum_{k=0}^{\infty} \binom{0}{k} X^k$ (denn $\binom{0}{k} = \begin{cases} 1, & \text{wenn } k=0; \\ 0, & \text{wenn } k \neq 0 \end{cases}$) folgt

nun $(1 + X)^n = \sum_{k=0}^{\infty} \binom{n}{k} X^k$ für alle $n \geq 0$ per Induktion nach n , und $(1 + X)^n = \sum_{k=0}^{\infty} \binom{n}{k} X^k$ für alle $n < 0$ per "Rückwärtsinduktion" nach n (also Schritt von $n+1$ auf n , statt wie in der gewöhnlichen Induktion von n auf $n+1$). Somit ist $(1 + X)^n = \sum_{k=0}^{\infty} \binom{n}{k} X^k$ für alle $n \in \mathbb{Z}$ bewiesen, und Satz 3.1 ist gezeigt.

Eine *Bemerkung* am Rande: Oft *definiert* man auch gebrochene oder gar komplexe Potenzen von $1 + X$, indem man $(1 + X)^n = \sum_{k=0}^{\infty} \binom{n}{k} X^k$ für gebrochene bzw. komplexe n setzt. Solche Potenzen haben, wenn man für X eine reelle Zahl einsetzt, nicht unbedingt ihren Sinn, aber als Potenzreihen kann man sie problemlos definieren.

Wir wollen nun als Beispiel dafür, wie nützlich Potenzreihen sein können, die sogenannte *Vandermonde-Faltungsformel* beweisen (eine bekannte Formel aus der Kombinatorik). Diese Formel ist nicht notwendig für unsere weiteren Pläne; deshalb kann der Leser auch gleich zum Anfang von Abschnitt 3.3 springen.

Hier ist also die Formel:

Satz 3.2 (die Vandermonde-Faltungsformel): Für beliebige $a \in \mathbb{Z}$, $b \in \mathbb{Z}$ und $n \in \mathbb{Z}$ ist

$$\binom{a+b}{n} = \sum_{k=0}^n \binom{a}{k} \binom{b}{n-k}.$$

(Die Summe $\sum_{k=0}^n$ ist für $n < 0$ einfach als die leere Summe zu verstehen, und die leere

Summe hat per definitionem den Wert 0, was sich natürlich mit $\binom{a+b}{n} = 0$ für $n < 0$

deckt.)

Beweis von Satz 3.2: Nach Satz 3.1 ist $(1 + X)^a = \sum_{k=0}^{\infty} \binom{a}{k} X^k = \sum_{i=0}^{\infty} \binom{a}{i} X^i$ und analog $(1 + X)^b = \sum_{i=0}^{\infty} \binom{b}{i} X^i$. Also ist

$$(1 + X)^a (1 + X)^b = \sum_{i=0}^{\infty} \binom{a}{i} X^i \cdot \sum_{i=0}^{\infty} \binom{b}{i} X^i = \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{a}{k} \binom{b}{n-k} X^n$$

(nach der Rechenregel für das Produkt zweier Potenzreihen). Andererseits ist

$$\begin{aligned} (1 + X)^a (1 + X)^b &= (1 + X)^{a+b} = \sum_{k=0}^{\infty} \binom{a+b}{k} X^k && \text{(nach Satz 3.1)} \\ &= \sum_{n=0}^{\infty} \binom{a+b}{n} X^n. \end{aligned}$$

Somit ist

$$\sum_{n=0}^{\infty} \sum_{k=0}^n \binom{a}{k} \binom{b}{n-k} X^n = \sum_{n=0}^{\infty} \binom{a+b}{n} X^n.$$

Doch wenn zwei Potenzreihen gleich sind, sind alle ihre entsprechenden Koeffizienten gleich; somit folgt hieraus

$$\sum_{k=0}^n \binom{a}{k} \binom{b}{n-k} = \binom{a+b}{n} \quad \text{für alle } n \in \mathbb{N}.$$

Doch auch für negative ganze n gilt $\sum_{k=0}^n \binom{a}{k} \binom{b}{n-k} = \binom{a+b}{n}$ (denn für negative ganze n ist $\binom{a+b}{n} = 0$ und $\sum_{k=0}^n \binom{a}{k} \binom{b}{n-k} = (\text{leere Summe}) = 0$). Somit ist

$$\sum_{k=0}^n \binom{a}{k} \binom{b}{n-k} = \binom{a+b}{n} \quad \text{für alle } n \in \mathbb{Z}.$$

Damit ist Satz 3.2 bewiesen.

Nun könnte sich ein kritischer Leser fragen, was wir eigentlich gewonnen haben, indem wir Binomialkoeffizienten $\binom{n}{k}$ mit negativem (oder gar rationalem oder reellem) n eingeführt und mithilfe von Potenzreihen Satz 3.2 gezeigt haben. Denn für die "gewöhnlichen" Binomialkoeffizienten (also $\binom{n}{k}$ mit $n \in \mathbb{N}$ und $k \in \{0, 1, \dots, n\}$) benötigt man für den Beweis von Satz 3.2 keine Potenzreihen (man kann in diesem Fall nämlich $(1 + X)^a$ und $(1 + X)^b$ genauso gut als Polynome auffassen, weil a und b natürliche Zahlen sind). Die Benutzung von Potenzreihen war also nur in dem Fall notwendig, wenn die Binomialkoeffizienten nicht "gewöhnlich" sind, also $a < 0$ oder $b < 0$ ist. Aber warum interessieren uns solche Binomialkoeffizienten?

Wie sich herausstellt, gibt es dafür einen guten Grund: Solche Binomialkoeffizienten sagen einiges aus über die "gewöhnlichen" Binomialkoeffizienten. Denn es gilt:

Satz 3.3 (Obere Negation): (a) Für $n \in \mathbb{Z}$ (oder $n \in \mathbb{Q}$, oder $n \in \mathbb{R}$, oder $n \in \mathbb{C}$ - je nachdem, was man will) und $k \in \mathbb{N}$ gilt stets

$$\binom{-n}{k} = (-1)^k \binom{n+k-1}{k}. \quad (24)$$

(b) Für jedes $n \in \mathbb{Z}$ und jedes $k \in \mathbb{N}$ ist $\binom{n}{k} \in \mathbb{Z}$.

Beweis von Satz 3.3: (a) Nach der Definition von Binomialkoeffizienten ist

$$\binom{-n}{k} = \begin{cases} \frac{(-n) \cdot ((-n) - 1) \cdot \dots \cdot ((-n) - k + 1)}{k!}, & \text{wenn } k \geq 0; \\ 0, & \text{wenn } k < 0 \end{cases} \quad \text{und}$$

$$\binom{n+k-1}{k} = \begin{cases} \frac{(n+k-1) \cdot ((n+k-1) - 1) \cdot \dots \cdot ((n+k-1) - k + 1)}{k!}, & \text{wenn } k \geq 0; \\ 0, & \text{wenn } k < 0 \end{cases}.$$

Um (24) zu beweisen, müssen wir somit nur noch zeigen, daß

$$(-n) \cdot ((-n) - 1) \cdot \dots \cdot ((-n) - k + 1) = (-1)^k \cdot ((n+k-1) \cdot ((n+k-1) - 1) \cdot \dots \cdot ((n+k-1) - k + 1))$$

gilt, wenn $k \geq 0$ ist. Doch dies ist klar, denn

$$\begin{aligned} (-n) \cdot ((-n) - 1) \cdot \dots \cdot ((-n) - k + 1) &= (-1)^k \cdot (n \cdot (n+1) \cdot \dots \cdot (n+k-1)) \\ &= (-1)^k \cdot ((n+k-1) \cdot (n+k-2) \cdot \dots \cdot n) \\ &= (-1)^k \cdot ((n+k-1) \cdot ((n+k-1) - 1) \cdot \dots \cdot ((n+k-1) - k + 1)). \end{aligned}$$

Damit ist (24) bewiesen, also Satz 3.3 (a) gezeigt.

(b) Im Falle von $n \geq 0$ ist

$$\binom{n}{k} = (\text{Anzahl aller } k\text{-elementigen Teilmenge einer vorgegebenen } n\text{-elementigen Menge}) \in \mathbb{Z},$$

und im Falle von $n < 0$ ist

$$\begin{aligned} \binom{n}{k} &= (-1)^k \binom{-n+k-1}{k} \quad (\text{nach Satz 3.3 (a), angewandt auf } -n \text{ statt } n) \\ &\in \mathbb{Z} \end{aligned}$$

(denn setzen wir $\nu = -n + k - 1$, dann ist $\nu \geq 0$ und damit

$$\begin{aligned} \binom{-n+k-1}{k} &= \binom{\nu}{k} \\ &= (\text{Anzahl aller } k\text{-elementigen Teilmenge einer vorgegebenen } \nu\text{-elementigen Menge}) \in \mathbb{Z} \end{aligned}$$

). In beiden möglichen Fällen ist also $\binom{n}{k} \in \mathbb{Z}$. Daher muss $\binom{n}{k} \in \mathbb{Z}$ immer gelten.

Damit ist Satz 3.3 (b) gezeigt, und der Beweis von Satz 3.3 ist vollständig.

Mithilfe von Satz 3.3 (a) können wir Satz 3.2 ein wenig mutieren:

Folgerung 3.4 (die zweimal negierte Vandermonde-Faltungsformel):

Für beliebige $a \in \mathbb{Z}$, $b \in \mathbb{Z}$ und $n \in \mathbb{Z}$ ist

$$\binom{(a+b)+n-1}{n} = \sum_{k=0}^n \binom{a+k-1}{k} \binom{b+(n-k)-1}{n-k}.$$

Beweis von Folgerung 3.4: Wenden wir Satz 3.2 auf $-a$ und $-b$ statt a bzw. b an, so erhalten wir

$$\binom{(-a)+(-b)}{n} = \sum_{k=0}^n \binom{-a}{k} \binom{-b}{n-k}. \quad (25)$$

Doch nach Satz 3.3 (a) gilt

$$\begin{aligned} \binom{(-a)+(-b)}{n} &= \binom{-(a+b)}{n} = (-1)^n \binom{(a+b)+n-1}{n}; \\ \binom{-a}{k} &= (-1)^k \binom{a+k-1}{k}; \\ \binom{-b}{n-k} &= (-1)^{n-k} \binom{b+(n-k)-1}{n-k}. \end{aligned}$$

Setzen wir dies in (25) ein, und kürzen wir $(-1)^n$ heraus, erhalten wir die Behauptung von Folgerung 3.4.

Nun betrachten wir Folgerung 3.4 genauer: Wenn a und b natürliche Zahlen sind, dann sind die Binomialkoeffizienten in Folgerung 3.4 ganz gewöhnliche Binomialkoeffizienten (außer gelegentlich auftretenden $\binom{-1}{0} = 1$ im Falle von $a = 0$ oder $b = 0$), und wir haben damit eine Identität zwischen gewöhnlichen Binomialkoeffizienten bewiesen, aber im Beweis "ungewöhnliche" Binomialkoeffizienten wie $\binom{-a}{k}$ und Potenzreihen verwendet. Dies ist ein starkes Argument dafür, daß der Begriff einer Potenzreihe kein Irrweg ist.

Man kann übrigens Satz 3.2 auch etwas anders variieren, indem man nur a durch $-a$ ersetzt (aber b unverändert läßt).

3.3. Einsetzung von Potenzreihen ineinander

Wir führen nun eine simple Notation ein:

Definition (Potenzreihe lower a): Ist a eine Potenzreihe, dann definieren wir eine neue Potenzreihe lower a folgendermaßen: Wir schreiben a in der Form $a = \sum_{i=0}^{\infty} a_i X^i$ mit ganzen Zahlen a_0, a_1, a_2, \dots . Dann bezeichnen wir mit lower a die Potenzreihe $\sum_{i=1}^{\infty} a_i X^{i-1}$.

Wir haben dann

$$a = \sum_{i=0}^{\infty} a_i X^i = a_0 \underbrace{X^0}_{=1} + \sum_{i=1}^{\infty} a_i X^i = a_0 + \sum_{i=1}^{\infty} a_i X^i, \quad \text{also}$$

$$a - a_0 = \sum_{i=1}^{\infty} a_i X^i = X \underbrace{\sum_{i=1}^{\infty} a_i X^{i-1}}_{=\text{lower } a} = X \text{ lower } a.$$

Da a_0 der 0-te Koeffizient der Potenzreihe a ist, läßt sich dies umschreiben als

$$a - (\text{der 0-te Koeffizient der Potenzreihe } a) = X \text{ lower } a. \quad (26)$$

Insbesondere folgt hieraus $a = X \text{ lower } a$, wenn der 0-te Koeffizient der Potenzreihe a gleich 0 ist.

Das folgende einfache Resultat über Potenzreihen wird oft (teilweise stillschweigend) gebraucht:

Satz 3.4a (Satz vom Einsetzen von Potenzreihen): Sei P eine Potenzreihe, deren 0-ter Koeffizient gleich 0 ist.

(a) Sei $b = \sum_{i=0}^{\infty} b_i X^i$ eine weitere Potenzreihe. Dann konvergiert die Folge $\left(\sum_{i=0}^n b_i P^i \right)_{n \in \mathbb{N}}$ gegen eine Potenzreihe.

Diese Potenzreihe bezeichnet man mit $b \circ P$ oder mit $\sum_{i=0}^{\infty} b_i P^i$; man nennt sie auch die *Einsetzung von P in b* , oder die *Auswertung von P an b* .

(Anschaulich gesprochen ist diese Potenzreihe $b \circ P$ das, was man erhält, wenn man die Potenzreihe P an Stelle von X in die Potenzreihe b einsetzt. Satz 3.4a (a) besagt also, daß man eine Potenzreihe, deren 0-ter Koeffizient gleich 0 ist - so eine Potenzreihe ist in unserem Fall P - in eine Potenzreihe einsetzen darf. Aber man sollte nicht vergessen, daß man sonst im Allgemeinen nichts - nicht einmal Zahlen! - in eine Potenzreihe einsetzen darf.)

(b) Sei P eine Potenzreihe, deren 0-ter Koeffizient gleich 0 ist. Seien b und c zwei weitere Potenzreihen. Dann ist $-(b \circ P) = (-b) \circ P$, $(b + c) \circ P = b \circ P + c \circ P$ und $(bc) \circ P = (b \circ P)(c \circ P)$.

Beweis von Satz 3.4a: Da P eine Potenzreihe ist, deren 0-ter Koeffizient gleich 0 ist, gilt $P = X \text{ lower } P$ (wegen (26)).

(a) Sei Q die Potenzreihe, deren k -ter Koeffizient gleich dem k -ten Koeffizienten der Potenzreihe $\sum_{i=0}^k b_i P^i$ für jedes $k \in \mathbb{N}$ ist. Wir wollen dann zeigen, daß die Folge

$\left(\sum_{i=0}^n b_i P^i \right)_{n \in \mathbb{N}}$ gegen die Potenzreihe Q konvergiert. Dazu müssen wir beweisen, daß folgendes gilt:

Für jedes $g \in \mathbb{N}$ gibt es ein $N \in \mathbb{N}$, sodaß für jedes $n \in \mathbb{N}$ mit $n \geq N$ gilt:

$\left(\text{Die Folgen } Q \text{ und } \sum_{i=0}^n b_i P^i \text{ unterscheiden sich nicht vor ihrem } g\text{-ten Glied} \right)$, d. h.
 $\left(\text{für jedes } k \in \{0, 1, \dots, g-1\} \text{ ist das } k\text{-te Glied der Folge } Q \text{ gleich dem der Folge } \sum_{i=0}^n b_i P^i \right).$

Da wir "Potenzreihe" statt "Folge" und "Koeffizient" statt "Glied" sagen, bedeutet diese Bedingung also:

Für jedes $g \in \mathbb{N}$ gibt es ein $N \in \mathbb{N}$, sodaß für jedes $n \in \mathbb{N}$ mit $n \geq N$ gilt:

$$\left(\text{Die Potenzreihen } Q \text{ und } \sum_{i=0}^n b_i P^i \text{ unterscheiden sich nicht vor ihrem } g\text{-ten Koeffizienten} \right), \text{ d. h.}$$

$$\left(\begin{array}{l} \text{für jedes } k \in \{0, 1, \dots, g-1\} \text{ ist der } k\text{-te Koeffizient der Potenzreihe } Q \\ \text{gleich dem der Potenzreihe } \sum_{i=0}^n b_i P^i \end{array} \right).$$

Und in der Tat ist diese Bedingung erfüllt, nämlich (zum Beispiel) für $N = g$, denn für $N = g$ gilt: Für jedes $n \in \mathbb{N}$ mit $n \geq N$ (also $n \geq g$, weil ja $N = g$ ist) und jedes $k \in \{0, 1, \dots, g-1\}$ ist

$$\begin{aligned} & \text{(der } k\text{-te Koeffizient der Potenzreihe } Q) \\ &= \left(\text{der } k\text{-te Koeffizient der Potenzreihe } \sum_{i=0}^k b_i P^i \right) \quad \text{(nach der Definition von } Q) \\ &= \left(\text{der } k\text{-te Koeffizient der Potenzreihe } \sum_{i=0}^n b_i P^i \right) - \left(\text{der } k\text{-te Koeffizient der Potenzreihe } \sum_{i=k+1}^n b_i P^i \right) \\ & \quad \left(\text{denn wegen } k < g = N \leq n \text{ ist } \sum_{i=0}^k b_i P^i = \sum_{i=0}^n b_i P^i - \sum_{i=k+1}^n b_i P^i \right) \\ &= \left(\text{der } k\text{-te Koeffizient der Potenzreihe } \sum_{i=0}^n b_i P^i \right) \end{aligned}$$

(weil

$$\begin{aligned} & \left(\text{der } k\text{-te Koeffizient der Potenzreihe } \sum_{i=k+1}^n b_i P^i \right) \\ &= \sum_{i=k+1}^n b_i \underbrace{\left(\text{der } k\text{-te Koeffizient der Potenzreihe } P^i \right)}_{=0, \text{ denn } P^i = (X \text{ lower } P)^i = X^i (\text{lower } P)^i \text{ und } i > k} = \sum_{i=k+1}^n b_i 0 = 0 \end{aligned}$$

gilt). Damit ist Satz 3.4a **(a)** bewiesen.

(b) Die Gleichungen $-(b \circ P) = (-b) \circ P$ und $(b + c) \circ P = b \circ P + c \circ P$ sind sehr leicht zu zeigen. Um $(bc) \circ P = (b \circ P)(c \circ P)$ zu beweisen, müssen wir genauer mit Grenzwerten arbeiten; wir überlassen den Beweis dem Leser³⁵. Wir werden die Aussage von Satz 3.4a **(b)** nicht zum Beweis von $\mathcal{B} \iff \mathcal{D}$ gebrauchen; daher belasse ich es hierbei.

³⁵Hinweis: wenn man b in der Form $b = \sum_{i=0}^{\infty} b_i X^i$ und c in der Form $c = \sum_{i=0}^{\infty} c_i X^i$ schreibt, dann ist

$$\left(\left(\sum_{i=0}^n b_i X^i \right) \cdot \left(\sum_{i=0}^n c_i X^i \right) \right) \circ P = \left(\left(\sum_{i=0}^n b_i X^i \right) \circ P \right) \left(\left(\sum_{i=0}^n c_i X^i \right) \circ P \right)$$

für jedes $n \in \mathbb{N}$. Jetzt muss man sich überlegen, daß für $n \rightarrow \infty$ die linke Seite dieser Gleichung gegen $(bc) \circ P$ und die rechte gegen $(b \circ P)(c \circ P)$ strebt. Für die linke ist dies trivial; für die rechte muss man ein wenig nachdenken.

3.4. Rationale Potenzreihen und Ableitung

Nach diesen Fingerübungen im Arbeiten mit Potenzreihen wollen wir nun ein paar weitere Definitionen geben, die uns den Weg zum alternativen Beweis von $\mathcal{B} \iff \mathcal{D}$ ebnen. Zuerst wollen wir unsere Theorie von ganzzahligen Potenzreihen (d. h. von Folgen, deren Glieder ganze Zahlen sind) auf sogenannte rationale Potenzreihen ausdehnen:

Definition (rationale Potenzreihe): Wir haben bislang über ganzzahlige Potenzreihen geredet; dies sind Folgen, deren Glieder ganze Zahlen sind. Wir können genauso *rationale Potenzreihen* definieren als Folgen, deren Glieder rationale Zahlen sind. Wir definieren für rationale Potenzreihen genau die gleichen Notationen wie für ganzzahlige Potenzreihen; insbesondere schreiben wir eine rationale Potenzreihe $(a_0, a_1, a_2, \dots) \in \mathbb{Q}^{\mathbb{N}}$ auch als unendliche Summe $\sum_{i=0}^{\infty} a_i X^i$. Die Menge $\mathbb{Z}^{\mathbb{N}}$ der ganzzahligen Potenzreihen ist eine Teilmenge der Menge $\mathbb{Q}^{\mathbb{N}}$ der rationalen Potenzreihen.

Wenn wir im Folgenden von einer "Potenzreihe" sprechen, meinen wir eine rationale Potenzreihe (außer, wir sagen extra dazu, daß wir eine ganzzahlige Potenzreihe wollen!).

Nun definieren wir die sogenannte *Ableitung* einer Potenzreihe. Obwohl dieser Begriff sehr ähnlich zum Begriff der Ableitung einer Funktion in der Analysis ist (er ist an diesen Begriff angelehnt - siehe Anmerkung 1) nach der Definition), benötigen wir für seine Definition überhaupt keine Analysis (keine Grenzwerte, keine Differenzierbarkeitsbedingungen, etc.)³⁶:

Definition (Ableitung einer Potenzreihe): Sei $a = \sum_{i=0}^{\infty} a_i X^i$ eine ganzzahlige oder rationale Potenzreihe. Dann definieren wir eine neue ganzzahlige bzw. rationale (je nachdem, ob a ganzzahlig oder rational war) Potenzreihe a' durch $a' = \sum_{i=0}^{\infty} (i+1) a_{i+1} X^i$. Statt a' nennt man diese Potenzreihe oft auch $\frac{d}{dX} a$.

Diese Potenzreihe $a' = \frac{d}{dX} a = \sum_{i=0}^{\infty} (i+1) a_{i+1} X^i$ nennt man oft die *Ableitung* (oder *formale Ableitung*) der Potenzreihe a nach X .

Anmerkungen: 1) Der Begriff "Ableitung" kommt hier natürlich daher, daß man in der Funktionentheorie die Ableitung einer Funktion, die durch eine Potenzreihe $\sum_{i=0}^{\infty} a_i x^i$ gegeben ist, nach der Formel $\sum_{i=0}^{\infty} (i+1) a_{i+1} x^i$ berechnen kann; doch wir brauchen uns nicht um die Funktionentheorie zu kümmern, um mit Ableitungen formaler Potenzreihen zu arbeiten. Historisch war es so, daß der Begriff der Ableitung einer Funktion

³⁶weshalb dieser Begriff auch *formale Ableitung* genannt wird, um hervorzuheben, daß er rein formal durch Manipulation der Potenzreihe definiert wird und nicht (wie der Ableitungsbegriff in der Analysis) durch Grenzwerte

(von \mathbb{R} nach \mathbb{R} , oder von \mathbb{C} nach \mathbb{C}) zuerst erfunden wurde, und dann der Begriff der Ableitung einer formalen Potenzreihe "nach seinem Abbild" erschaffen wurde, indem einfach die Formel $\sum_{i=0}^{\infty} (i+1) a_{i+1} x^i$ als die Definition genommen wurde. Ableiten von Potenzreihen ist aber viel einfacher als Ableiten von Funktionen: Wir brauchen uns nicht um Differenzierbarkeit zu kümmern.

2) Die Ableitung von Potenzreihen erfüllt die gleichen Rechenregeln wie die Ableitung von Funktionen: Für jede Potenzreihe P und jedes $\lambda \in \mathbb{Q}$ gilt $\frac{d}{dX}(\lambda P) = \lambda \frac{d}{dX}P$. Für jedes $\lambda \in \mathbb{Q}$ gilt $\frac{d}{dX}\lambda = 0$ (wobei wir hier λ als "konstante Potenzreihe" betrachten, d. h. als Folge $(\lambda, 0, 0, 0, \dots)$, in der alle Folgenglieder bis auf das nullte gleich 0 sind). Für je zwei Potenzreihen P und Q gilt

$$\frac{d}{dX}(P+Q) = \frac{d}{dX}P + \frac{d}{dX}Q \quad \text{und} \quad (27)$$

$$\frac{d}{dX}(PQ) = \left(\frac{d}{dX}P\right) \cdot Q + P \cdot \left(\frac{d}{dX}Q\right). \quad (28)$$

Durch mehrfache Anwendung von (28) erhalten wir

$$\frac{d}{dX}(P^i) = iP^{i-1} \cdot \left(\frac{d}{dX}P\right) \quad (29)$$

für jede Potenzreihe P und jedes $i \in \mathbb{N}_+$.

3.5. Der natürliche Logarithmus einer rationalen Potenzreihe

Nun eine weitere Definition:

Definition (natürlicher Logarithmus einer rationalen Potenzreihe):

Für jede rationale Potenzreihe a , deren 0-ter Koeffizient gleich 1 ist, definieren wir eine neue rationale Potenzreihe $\ln a$ durch

$$\ln a = - \sum_{i=1}^{\infty} \frac{(1-a)^i}{i}.$$

Diese neue Potenzreihe $\ln a$ heißt der *natürliche Logarithmus* der Potenzreihe a .

Hierzu sind einige *Bemerkungen* nötig:

1) Die unendliche Summe $\sum_{i=1}^{\infty} \frac{(1-a)^i}{i}$ ist definiert als $\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{(1-a)^i}{i}$. Dabei nutzen wir aus, daß die Folge $\left(\sum_{i=1}^n \frac{(1-a)^i}{i}\right)_{n \in \mathbb{N}}$ konvergiert, was aus Satz 3.4a (a) (angewandt auf die Potenzreihen $P = 1 - a$ und $b = \sum_{i=1}^{\infty} \frac{X^i}{i}$) folgt³⁷, weil der 0-te Koeffizient der Potenzreihe $1 - a$ gleich 0 ist (denn der 0-te Koeffizient der Potenzreihe a ist 1).

³⁷Strenggenommen müssten wir in Satz 3.4a (a) überall $i = 0$ durch $i = 1$ und "Potenzreihe" durch "rationale Potenzreihe" ersetzen, um Satz 3.4a (a) auf unsere Situation anzuwenden, doch auch mit diesen Änderungen läßt sich der Beweis von Satz 3.4a (a) genauso durchführen wie oben.

2) Für jede rationale Potenzreihe a , deren 0-ter Koeffizient gleich 1 ist, haben wir nun eine neue rationale Potenzreihe $\ln a$ definiert. Es sollte jedoch darauf hingewiesen werden, daß für eine ganzzahlige Potenzreihe a die Potenzreihe $\ln a$ nicht notwendigerweise ganzzahlig sein muss.

3) Oftmals wird auch $\log a$ statt $\ln a$ geschrieben.

4) Man sieht sofort ein, daß $\ln 1 = 0$ ist.

5) Die Bezeichnung \ln ist an die Analysis angelehnt, wo der ganzzahlige Logarithmus \ln die Potenzreihendarstellung $\ln x = - \sum_{i=1}^{\infty} \frac{(1-x)^i}{i}$ in der Nähe von $x = 1$ hat (diese

Darstellung ist besser bekannt in der Form $\ln(1-x) = - \sum_{i=1}^{\infty} \frac{x^i}{i}$ in der Nähe von $x = 0$). Wie immer, brauchen wir uns aber im Falle formaler Potenzreihen keine Gedanken darüber zu machen, wo eine solche Potenzreihe konvergiert.

In der Analysis ist der Logarithmus dafür bekannt, Produkte in Summen umzuwandeln: $\ln(ab) = \ln a + \ln b$, soweit $\ln a$ und $\ln b$ definiert sind. Auch unser "formaler" Logarithmus \ln von Potenzreihen hat diese Eigenschaft:

Satz 3.5: (a) Sind a und b zwei Potenzreihen, deren 0-te Koeffizienten gleich 1 sind, dann ist $\ln(ab) = \ln a + \ln b$.

(b) Für jedes $d \in \mathbb{N}_+$ sei $a(d)$ eine Potenzreihe, deren 0-ter Koeffizient 1 ist. Angenommen, die Folge $\left(\prod_{d=1}^n a(d) \right)_{n \in \mathbb{N}}$ konvergiert. Dann ist $\ln \prod_{d=1}^{\infty} a(d) = \sum_{d=1}^{\infty} \ln a(d)$, wobei wir mit $\prod_{d=1}^{\infty} a(d)$ den Grenzwert der Folge $\left(\prod_{d=1}^n a(d) \right)_{n \in \mathbb{N}}$ bezeichnen.

Diesen Satz direkt zu beweisen ist gar nicht so einfach (es geht aber). Wir werden ihn durch einen Trick beweisen, indem wir zuerst einen Hilfssatz (den wir sowieso später noch einmal brauchen werden) zeigen:

Satz 3.6 (der Satz von der logarithmischen Ableitung): Sei a eine rationale Potenzreihe, deren 0-ter Koeffizient gleich 1 ist. Dann gibt es eine rationale Potenzreihe b , die $b \cdot a = 1$ erfüllt. Wir bezeichnen diese Potenzreihe b mit a^{-1} (weil sie ja das multiplikative Inverse der Potenzreihe a ist). Der 0-te Koeffizient dieser Potenzreihe a^{-1} ist 1.

Ferner gilt $\frac{d}{dX} (\ln a) = a^{-1} \cdot \frac{d}{dX} a$. Wenn die Potenzreihe a ganzzahlig ist, dann sind auch die Potenzreihen a^{-1} und $\frac{d}{dX} (\ln a)$ ganzzahlig.

Dies ist besonders interessant, weil nicht für jede ganzzahlige Potenzreihe a , deren 0-ter Koeffizient gleich 1 ist, auch die Potenzreihe $\ln a$ ganzzahlig ist - aber (nach Satz 3.6) ihre Ableitung doch.

Beweis von Satz 3.6: Zuerst beweisen wir, daß es eine rationale Potenzreihe b gibt, die $b \cdot a = 1$ erfüllt. Und zwar konstruieren wir eine solche Potenzreihe direkt: Die Folge $\left(\sum_{i=0}^n (1-a)^i \right)_{n \in \mathbb{N}}$ konvergiert (dies folgt auf Satz 3.4a (a), angewandt auf die

Potenzreihen $P = 1 - a$ und $b = \sum_{i=0}^n X^i$ ³⁸, weil der 0-te Koeffizient der Potenzreihe $1 - a$ gleich 0 ist³⁹). Daher ist die unendliche Summe $\sum_{i=0}^{\infty} (1 - a)^i = \lim_{n \rightarrow \infty} \sum_{i=0}^n (1 - a)^i$ wohldefiniert. Nun definieren wir die Potenzreihe b durch $b = \sum_{i=0}^{\infty} (1 - a)^i$. Dann ist b eine rationale Potenzreihe, und erfüllt

$$\begin{aligned} b \cdot (1 - a) &= \sum_{i=0}^{\infty} (1 - a)^i \cdot (1 - a) = \sum_{i=0}^{\infty} (1 - a)^{i+1} = \sum_{i=1}^{\infty} (1 - a)^i \\ &\quad \text{(hier haben wir } i + 1 \text{ in der Summe durch } i \text{ substituiert)} \\ &= \underbrace{\sum_{i=0}^{\infty} (1 - a)^i}_{=b} - \underbrace{(1 - a)^0}_{=1} = b - 1, \end{aligned}$$

also $1 = b - b \cdot (1 - a) = b \cdot (1 - (1 - a)) = b \cdot a$. Wir haben also die Existenz einer rationalen Potenzreihe b , die $b \cdot a = 1$ erfüllt, bewiesen. Und wir wissen jetzt auch, wie man diese Potenzreihe b findet: $b = \sum_{i=0}^{\infty} (1 - a)^i$. Da wir diese Potenzreihe b mit a^{-1} bezeichnen, haben wir also $a^{-1} = \sum_{i=0}^{\infty} (1 - a)^i$.

Der 0-te Koeffizient dieser Potenzreihe a^{-1} ist 1 (denn unter Verwendung der weiter oben definierten Potenzreihe a gilt

$$\begin{aligned} a^{-1} &= \sum_{i=0}^{\infty} (1 - a)^i = \sum_{i=0}^{\infty} (-X \cdot \text{lower } a)^i \quad \left(\begin{array}{l} \text{denn aus (26) folgt } a - 1 = X \text{ lower } a \\ \text{(da der 0-te Koeffizient der Potenzreihe } a \text{ gleich} \\ \text{1 ist), also } 1 - a = -X \cdot \text{lower } a \end{array} \right) \\ &= \sum_{i=0}^{\infty} (-\text{lower } a)^i X^i \end{aligned}$$

und damit

$$\begin{aligned} \left(\text{der 0-te Koeffizient der Potenzreihe } a^{-1} \right) &= \left(\text{der 0-te Koeffizient der Potenzreihe } \sum_{i=0}^{\infty} (-\text{lower } a)^i X^i \right) \\ &= \sum_{i=0}^{\infty} \left(\underbrace{\text{der 0-te Koeffizient der Potenzreihe } (-\text{lower } a)^i X^i}_{\substack{\text{dieser Koeffizient ist 1, wenn } i=0 \text{ ist,} \\ \text{und sonst ist er immer 0}}} \right) = 1 \end{aligned}$$

).

³⁸Strenggenommen müssten wir in Satz 3.4a (a) überall "Potenzreihe" durch "rationale Potenzreihe" ersetzen, um Satz 3.4a (a) auf unsere Situation anzuwenden, doch auch mit diesen Änderungen läßt sich der Beweis von Satz 3.4a (a) genauso durchführen wie oben.

³⁹denn der 0-te Koeffizient der Potenzreihe a ist 1

Nun wollen wir zeigen, daß $\frac{d}{dX}(\ln a) = a^{-1} \cdot \frac{d}{dX}a$ ist. Dazu rechnen wir:

$$\begin{aligned}
& \frac{d}{dX}(\ln a) \\
&= \frac{d}{dX} \left(- \sum_{i=1}^{\infty} \frac{(1-a)^i}{i} \right) \quad (\text{nach der Definition von } \ln a) \\
&= - \sum_{i=1}^{\infty} \frac{d}{dX} \frac{(1-a)^i}{i} \quad (\text{nach (27), auf unendliche Summen übertragen}) \\
&= - \sum_{i=1}^{\infty} \frac{1}{i} \frac{d}{dX} \left((1-a)^i \right) \quad \left(\text{denn } \frac{d}{dX}(\lambda P) = \lambda \frac{d}{dX}P \text{ für jedes } \lambda \in \mathbb{Q} \text{ und jede Potenzreihe } P \right) \\
&= - \sum_{i=1}^{\infty} \underbrace{\frac{1}{i}}_{=1} (1-a)^{i-1} \cdot \left(\underbrace{\frac{d}{dX}(1-a)}_{= \frac{d}{dX}1 - \frac{d}{dX}a = -\frac{d}{dX}a} \right) \\
&\quad \left(\text{denn (29) (angewandt auf } P = 1-a) \text{ ergibt } \frac{d}{dX}((1-a)^i) = i(1-a)^{i-1} \cdot \left(\frac{d}{dX}(1-a) \right) \right) \\
&= - \sum_{i=1}^{\infty} (1-a)^{i-1} \cdot \left(-\frac{d}{dX}a \right) = \sum_{i=1}^{\infty} (1-a)^{i-1} \cdot \frac{d}{dX}a = \underbrace{\sum_{i=0}^{\infty} (1-a)^i}_{=a^{-1}} \cdot \frac{d}{dX}a \\
&\quad (\text{hier haben wir } i \text{ für } i-1 \text{ in der Summe substituiert}) \\
&= a^{-1} \cdot \frac{d}{dX}a.
\end{aligned}$$

Schließlich wollen wir zeigen, daß für jede ganzzahlige Potenzreihe a auch die Potenzreihen a^{-1} und $\frac{d}{dX}(\ln a)$ ganzzahlig sind. Doch dies ist klar: Denn $a^{-1} = \sum_{i=0}^{\infty} (1-a)^i = \lim_{n \rightarrow \infty} \sum_{i=0}^n (1-a)^i$ ist ganzzahlig (als Grenzwert einer Folge ganzzahliger Potenzreihen), und $\frac{d}{dX}(\ln a) = a^{-1} \cdot \frac{d}{dX}a$ ist ebenfalls ganzzahlig (als Produkt ganzzahliger Potenzreihen). Somit ist der Beweis von Satz 3.6 vollständig.

Aus Satz 3.6 können wir nun Satz 3.5 schnell mithilfe von folgendem Lemma folgern:

Satz 3.7 (Ableitungskriterium): Seien $u = \sum_{i=0}^{\infty} u_i X^i$ und $v = \sum_{i=0}^{\infty} v_i X^i$ zwei rationale Potenzreihen. Wenn $u_0 = v_0$ und $\frac{d}{dX}u = \frac{d}{dX}v$ gilt, dann ist $u = v$.

Beweis von Satz 3.7: Wegen $u = \sum_{i=0}^{\infty} u_i X^i$ ist $\frac{d}{dX}u = \sum_{i=0}^{\infty} (i+1) u_{i+1} X^i$, und analog $\frac{d}{dX}v = \sum_{i=0}^{\infty} (i+1) v_{i+1} X^i$. Somit wird $\frac{d}{dX}u = \frac{d}{dX}v$ zu $\sum_{i=0}^{\infty} (i+1) u_{i+1} X^i = \sum_{i=0}^{\infty} (i+1) v_{i+1} X^i$. Wenn zwei Potenzreihen gleich sind, sind jeweils entsprechende Koeffizienten gleich; also folgt hieraus $(i+1) u_{i+1} = (i+1) v_{i+1}$ für jedes $i \in \mathbb{N}$. Wegen $i+1 > 0$ ergibt dies

$u_{i+1} = v_{i+1}$ für jedes $i \in \mathbb{N}$. Das heißt, $u_j = v_j$ für jedes $j \in \mathbb{N}_+ = \mathbb{N} \setminus \{0\}$. Aber auch für $j = 0$ muß $u_j = v_j$ gelten (da $u_0 = v_0$). Also ist $u_j = v_j$ für jedes $j \in \mathbb{N}$. Folglich ist $u = \sum_{i=0}^{\infty} \underbrace{u_i}_{=v_i} X^i = \sum_{i=0}^{\infty} v_i X^i = v$, und Satz 3.7 ist bewiesen.

Schließlich noch eine Trivialität:

Satz 3.8: Für jede rationale Potenzreihe a , deren 0-ter Koeffizient gleich 1 ist, ist der 0-te Koeffizient der Potenzreihe $\ln a$ gleich 0.

Beweis von Satz 3.8: Wie wir weiter oben gesehen haben, ist $1 - a = -X \cdot \text{lower } a$. Laut der Definition von $\ln a$ ist nun

$$\begin{aligned} \ln a &= \sum_{i=1}^{\infty} \frac{(1-a)^i}{i} = \sum_{i=1}^{\infty} \frac{(-X \cdot \text{lower } a)^i}{i} && \text{(da } 1-a = -X \cdot \text{lower } a) \\ &= \sum_{i=1}^{\infty} \frac{(-\text{lower } a)^i X^i}{i} = X \sum_{i=1}^{\infty} \frac{(-\text{lower } a)^i X^{i-1}}{i}. \end{aligned}$$

Es ist somit klar, daß der 0-te Koeffizient der Potenzreihe $\ln a$ gleich 0 ist (denn für jedes $i \in \{1, 2, 3, \dots\}$ ist $\frac{(-\text{lower } a)^i X^i}{i}$ eine Potenzreihe, deren 0-ter Koeffizient gleich 0 ist, und somit hat die Summe $\sum_{i=1}^{\infty} \frac{(-\text{lower } a)^i X^i}{i}$ all dieser Potenzreihen ebenfalls diese Eigenschaft). Damit ist Satz 3.8 bewiesen.

Nun können wir den *Beweis von Satz 3.5 (a)* durchführen: Bezeichnen wir mit u die Potenzreihe $\ln(ab)$, und mit v die Potenzreihe $\ln a + \ln b$. Wir haben dann

$$\begin{aligned} \frac{d}{dX} u &= \frac{d}{dX} (\ln(ab)) = (ab)^{-1} \cdot \frac{d}{dX} (ab) && \text{(nach Satz 3.5)} \\ &= (ab)^{-1} \cdot \left(\left(\frac{d}{dX} a \right) \cdot b + a \cdot \left(\frac{d}{dX} b \right) \right) \\ &\quad \left(\text{denn } \frac{d}{dX} (ab) = \left(\frac{d}{dX} a \right) \cdot b + a \cdot \left(\frac{d}{dX} b \right) \text{ nach (28)} \right) \\ &= (ab)^{-1} \cdot \left(\frac{d}{dX} a \right) \cdot b + (ab)^{-1} \cdot a \cdot \left(\frac{d}{dX} b \right) = \underbrace{(ab)^{-1} b}_{=a^{-1}} \cdot \frac{d}{dX} a + \underbrace{(ab)^{-1} a}_{=b^{-1}} \cdot \frac{d}{dX} b \\ &= \underbrace{a^{-1} \cdot \frac{d}{dX} a}_{=\frac{d}{dX}(\ln a)} + \underbrace{b^{-1} \cdot \frac{d}{dX} b}_{=\frac{d}{dX}(\ln b)} = \frac{d}{dX} (\ln a) + \frac{d}{dX} (\ln b) \\ &\quad \text{(nach Satz 3.5) \quad (nach Satz 3.5)} \\ &= \frac{d}{dX} \left(\underbrace{\ln a + \ln b}_{=v} \right) && \text{(nach (27))} \\ &= \frac{d}{dX} v. \end{aligned}$$

Andererseits gilt $u_0 = v_0$, wenn wir die Potenzreihe u in der Form $\sum_{i=0}^{\infty} u_i X^i$ und die Potenzreihe v in der Form $\sum_{i=0}^{\infty} v_i X^i$ schreiben⁴⁰. Laut Satz 3.7 folgt hieraus $u = v$, also $\ln(ab) = \ln a + \ln b$, womit Satz 3.5 (a) endlich bewiesen ist.

Satz 3.5 (b) folgt aus Satz 3.5 (a) durch Induktion und Grenzübergang (weil \ln stetig ist⁴¹). Damit ist Satz 3.5 komplett bewiesen.

Eine Aufgabe für Leser, die nichts zu tun haben: Man kann Satz 3.5 (a) auch anders, direkter (also ohne Umweg über die Ableitungen) beweisen, indem man $x = 1 - a$ und $y = 1 - b$ setzt, und $\ln(ab)$ mithilfe von $x + y(1 - x) = 1 - ab$ vereinfacht. Auf die Weise landet man bei einer (recht komplizierten) kombinatorischen Identität, die äquivalent zu $\ln(ab) = \ln a + \ln b$ ist. Jetzt kann man entweder diese Identität beweisen, und damit einen direkten Beweis von Satz 3.5 erhalten. Oder, wenn man mit dem bereits oben gegebenen Beweis von Satz 3.5 zufrieden ist, erhält man diese kombinatorische Identität "gratis" als Folgerung. Noch eine kombinatorische Identität also, die man durch Rechnen mit Potenzreihen hergeleitet hat.

Noch eine banale Folgerung aus dem Vorherigen:

Satz 3.9: Sei a eine rationale Potenzreihe, deren 0-ter Koeffizient gleich 1 ist.

(a) Dann ist $\ln(a^{-1}) = -\ln a$. (Hierbei ist $\ln a^{-1}$ wohldefiniert, weil aus Satz 3.6 folgt, daß der 0-te Koeffizient der Potenzreihe a^{-1} gleich 1 ist).

(b) Für jedes $k \in \mathbb{Z}$ ist $\ln(a^k) = k \ln a$.

Beweis von Satz 3.9: Nach Satz 3.5 (a) (angewandt auf $b = a^{-1}$) ist $\ln(aa^{-1}) = \ln a + \ln(a^{-1})$. Wegen $\ln(aa^{-1}) = \ln 1 = 0$ ist also $0 = \ln a + \ln(a^{-1})$, und daraus folgt Satz 3.9 (a). Satz 3.9 (b) folgt für $k \geq 0$ durch Induktion nach k aus Satz 3.5 (a), und der Fall $k < 0$ läßt sich auf den Fall $k \geq 0$ zurückführen, indem man a^{-1} für a einsetzt (hier benötigt man Satz 3.9 (a)). Damit ist der Beweis von Satz 3.9 komplett.

Da wir schon den "formalen Logarithmus" \ln eingeführt haben, ist es naheliegend, auch seiner Umkehrung - die, genauso wie in der Analysis, \exp heißt und dieselbe Potenzreihe hat wie in der Analysis - einige Worte zu widmen. Der Leser beachte jedoch, daß wir im Folgenden \exp nie mehr benötigen werden; wer sich nur für Satz 2.1 interessiert, kann von hier direkt zu Satz 3.11 springen.

Für jede rationale Potenzreihe P , deren 0-ter Koeffizient gleich 0 ist, definieren wir eine rationale Potenzreihe $\exp P$ durch $\exp P = \sum_{i=0}^{\infty} \frac{P^i}{i!}$. Diese unendliche Summe $\sum_{i=0}^{\infty} \frac{P^i}{i!}$

ist wieder als Grenzwert $\lim_{n \rightarrow \infty} \sum_{i=0}^n \frac{P^i}{i!}$ zu verstehen, wobei die Konvergenz der Folge

$\left(\sum_{i=0}^n \frac{P^i}{i!} \right)_{n \in \mathbb{N}}$ aus Satz 3.4a (a) (angewandt auf $b = \sum_{i=0}^{\infty} \frac{X^i}{i!}$) folgt. Man sieht leicht ein,

⁴⁰Denn u_0 ist der 0-te Koeffizient der Potenzreihe $u = \ln(ab)$, und v_0 ist der 0-te Koeffizient der Potenzreihe $v = \ln a + \ln b$, und nach Satz 3.8 müssen diese beiden Koeffizienten gleich 0 sein.

⁴¹Damit meinen wir, daß für jede konvergente Folge $(a_n)_{n \in \mathbb{N}}$ von Potenzreihen, deren 0-te Koeffizienten alle gleich 1 sind, die Gleichung $\ln\left(\lim_{n \rightarrow \infty} a_n\right) = \lim_{n \rightarrow \infty} (\ln a_n)$ gilt. Dies ist sehr leicht zu beweisen.

daß $\frac{d}{dX}(\exp P) = (\exp P) \cdot \frac{d}{dX}P$ ist⁴²; insbesondere folgt hieraus $\frac{d}{dX}(\exp X) = \exp X$.

Die interessanteste Eigenschaft von \exp und \ln ist nun:

Satz 3.10: (a) Für jede rationale Potenzreihe P , deren 0-ter Koeffizient gleich 0 ist, gilt $\ln(\exp P) = P$.

(b) Für jede rationale Potenzreihe a , deren 0-ter Koeffizient gleich 1 ist, gilt $\exp(\ln a) = a$.

Beweis von Satz 3.10 (nur grob skizziert):

(a) Wir schreiben die Potenzreihe $\ln(\exp P)$ in der Form $\ln(\exp P) = \sum_{i=0}^{\infty} u_i X^i$, und die Potenzreihe P in der Form $P = \sum_{i=0}^{\infty} v_i X^i$. Dann ist $u_0 = 0 = v_0$ (hierbei ist $v_0 = 0$, denn v_0 ist der 0-te Koeffizient der Potenzreihe P , und laut Annahme ist dieser

⁴²*Beweis:* Für jede konvergente Folge (p_0, p_1, p_2, \dots) von Potenzreihen ist $\frac{d}{dX} \lim_{n \rightarrow \infty} p_n = \lim_{n \rightarrow \infty} \frac{d}{dX} p_n$ (dies folgt aus der Definition von $\frac{d}{dX}$ und der Definition von $\lim_{n \rightarrow \infty}$). Angewandt auf die konvergente Folge (p_0, p_1, p_2, \dots) , die durch $p_n = \sum_{i=0}^n \frac{P^i}{i!}$ für alle $n \in \mathbb{N}$ gegeben ist, ergibt dies $\frac{d}{dX} \lim_{n \rightarrow \infty} \sum_{i=0}^n \frac{P^i}{i!} = \lim_{n \rightarrow \infty} \frac{d}{dX} \sum_{i=0}^n \frac{P^i}{i!}$. Da aber

$$\begin{aligned} \frac{d}{dX} \sum_{i=0}^n \frac{P^i}{i!} &= \sum_{i=0}^n \frac{d}{dX} \frac{P^i}{i!} && \text{(nach der Formel (27), mehrfach angewandt)} \\ &= \underbrace{\frac{d}{dX} \frac{P^0}{0!}}_{= \frac{d}{dX} \frac{1}{1} = \frac{d}{dX} 1 = 0} + \sum_{i=1}^n \frac{d}{dX} \frac{P^i}{i!} \\ &= \sum_{i=1}^n \frac{1}{i!} \frac{d}{dX} P^i = \sum_{i=1}^n \frac{1}{i!} i P^{i-1} \cdot \left(\frac{d}{dX} P \right) && \text{(nach (29))} \\ &= \sum_{i=1}^n \frac{1}{i! / i} P^{i-1} \cdot \left(\frac{d}{dX} P \right) = \sum_{i=1}^n \frac{1}{(i-1)!} P^{i-1} \cdot \left(\frac{d}{dX} P \right) && \text{(denn } i! / i = (i-1)! \text{)} \\ &= \sum_{i=0}^{n-1} \frac{1}{i!} P^i \cdot \left(\frac{d}{dX} P \right) && \text{(hier haben wir } i-1 \text{ in der Summe durch } i \text{ substituiert)} \\ &= \sum_{i=0}^{n-1} \frac{P^i}{i!} \cdot \left(\frac{d}{dX} P \right) \end{aligned}$$

für jedes $n \in \mathbb{N}$ gilt, wird dies zu

$$\frac{d}{dX} \lim_{n \rightarrow \infty} \sum_{i=0}^n \frac{P^i}{i!} = \lim_{n \rightarrow \infty} \left(\sum_{i=0}^{n-1} \frac{P^i}{i!} \cdot \left(\frac{d}{dX} P \right) \right) = \left(\lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \frac{P^i}{i!} \right) \cdot \left(\frac{d}{dX} P \right)$$

(denn für jede konvergente Folge (p_0, p_1, p_2, \dots) von Potenzreihen und jede Potenzreihe q ist $\lim_{n \rightarrow \infty} (p_n q) = \left(\lim_{n \rightarrow \infty} p_n \right) q$, wie man leicht einsieht). Wegen $\lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \frac{P^i}{i!} = \lim_{n \rightarrow \infty} \sum_{i=0}^n \frac{P^i}{i!} = \sum_{i=0}^{\infty} \frac{P^i}{i!} = \exp P$ wird dies zu $\frac{d}{dX} \exp P = (\exp P) \cdot \left(\frac{d}{dX} P \right)$, was zu beweisen war.

Koeffizient gleich 0) und

$$\begin{aligned} \frac{d}{dX} (\ln (\exp P)) &= (\exp P)^{-1} \cdot \underbrace{\frac{d}{dX} (\exp P)}_{=(\exp P) \cdot \frac{d}{dX} P} && \text{(nach Satz 3.6, angewandt auf } a = \exp P) \\ &= \frac{d}{dX} P. \end{aligned}$$

Nach Satz 3.7 ist also $\ln (\exp P) = P$, und der Beweis von Satz 3.10 (a) ist komplett.

(b) Aus Satz 3.10 (a) (angewandt auf $P = \ln a$) folgt, daß $\ln (\exp (\ln a)) = \ln a$ gilt. Wenn wir eine Potenzreihe b durch $b = a^{-1} \cdot \exp (\ln a)$ definieren, dann ist

$$\begin{aligned} \ln b &= \ln (a^{-1} \cdot \exp (\ln a)) = \underbrace{\ln (a^{-1})}_{=-\ln a} + \underbrace{\ln (\exp (\ln a))}_{=\ln a} && \text{(nach Satz 3.5 (a))} \\ &= 0. \end{aligned}$$

Nach Satz 3.6 ist aber $\frac{d}{dX} (\ln b) = b^{-1} \cdot \frac{d}{dX} b$, und wegen $\frac{d}{dX} (\ln b) = 0$ wird dies zu $0 = b^{-1} \cdot \frac{d}{dX} b$, also $0 = \frac{d}{dX} b$. Hieraus folgt schnell, daß $b = 1$ ist (denn der 0-te Koeffizient von b ist 1). Also ist $a^{-1} \cdot \exp (\ln a) = b = 1$, daher $\exp (\ln a) = a$, und Satz 3.10 (b) ist bewiesen.

3.6. Ausblicke zu Potenzreihen

Soweit unsere kleine Einführung in Potenzreihen in einer Variablen. Wer weitere Anwendungen dieser Methode in der Kombinatorik kennenlernen will, sei auf [8] verwiesen. Wer sich hingegen für weitere Verfeinerungen des Begriffes von formalen Potenzreihen interessiert, braucht nur in ein höheres Algebrabuch zu schauen (wie Abschnitt IV.§9 von [9] oder Abschnitt III.7 von [10] oder Abschnitt III.11 von [11]); hier sind einige Hinweise, was man in der Richtung tun kann:

- Der Ring aller ganzzahligen Potenzreihen wird meistens $\mathbb{Z}[[X]]$ genannt. Der Ring aller rationalen Potenzreihen wird meistens $\mathbb{Q}[[X]]$ genannt. Entsprechend kann man für einen Ring A einen Ring $A[[X]]$ aller Potenzreihen mit Koeffizienten aus A definieren. Diese Ringe haben eine Reihe interessanter Eigenschaften.
- Man kann auch Potenzreihen in n Variablen betrachten. Eine *ganzzahlige Potenzreihe in n Variablen* X_1, X_2, \dots, X_n wird definiert als eine n -"Multifolge" von ganzen Zahlen, also eine Familie $(a_{(\nu_1, \nu_2, \dots, \nu_n)})_{(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n}$, deren Elemente mit n -Tupeln $(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n$ indiziert sind. Man definiert Addition solcher Potenzreihen durch

$$(a_{(\nu_1, \nu_2, \dots, \nu_n)})_{(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n} + (b_{(\nu_1, \nu_2, \dots, \nu_n)})_{(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n} = (a_{(\nu_1, \nu_2, \dots, \nu_n)} + b_{(\nu_1, \nu_2, \dots, \nu_n)})_{(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n}$$

und Multiplikation durch

$$\begin{aligned} & \left(a_{(\nu_1, \nu_2, \dots, \nu_n)} \right)_{(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n} \cdot \left(b_{(\nu_1, \nu_2, \dots, \nu_n)} \right)_{(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n} \\ &= \left(\sum_{\substack{(\mu_1, \mu_2, \dots, \mu_n) \in \mathbb{N}^n; \\ (\mu_i \leq \nu_i \text{ für jedes } i)}} a_{(\mu_1, \mu_2, \dots, \mu_n)} b_{(\nu_1 - \mu_1, \nu_2 - \mu_2, \dots, \nu_n - \mu_n)} \right)_{(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n}. \end{aligned}$$

Diese Definitionen ergeben sich ganz natürlich, wenn man die Potenzreihen $\left(a_{(\nu_1, \nu_2, \dots, \nu_n)} \right)_{(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n}$ und $\left(b_{(\nu_1, \nu_2, \dots, \nu_n)} \right)_{(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n}$ in den Formen $\sum_{(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n} a_{(\nu_1, \nu_2, \dots, \nu_n)} X_1^{\nu_1} X_2^{\nu_2} \dots X_n^{\nu_n}$ bzw. $\sum_{(\nu_1, \nu_2, \dots, \nu_n) \in \mathbb{N}^n} b_{(\nu_1, \nu_2, \dots, \nu_n)} X_1^{\nu_1} X_2^{\nu_2} \dots X_n^{\nu_n}$ schreibt (so wie man Potenzreihen in einer Variablen in der Form $\sum_{i=0}^{\infty} a_i X^i$ geschrieben hat) und mithilfe des Distributivgesetzes ausmultipliziert.

- Für Potenzreihen in einer Variablen gilt: Man kann in eine Potenzreihe zwar (im Allgemeinen) keine Zahl einsetzen, aber man kann in sie eine andere Potenzreihe P einsetzen, solange der 0-te Koeffizient dieser Potenzreihe P gleich 0 ist. Dies wissen wir aus Satz 3.4a. Auch für Potenzreihen in n Variablen gilt eine analoge Aussage: Man kann in eine Potenzreihe in n Variablen zwar (im Allgemeinen) kein n -Tupel von Zahlen einsetzen, aber man kann in sie ein n -Tupel (P_1, P_2, \dots, P_n) von anderen Potenzreihen (in m Variablen, für ein beliebiges m) einsetzen, solange der $(0, 0, \dots, 0)$ -te Koeffizient von jeder der Potenzreihen P_1, P_2, \dots, P_n gleich 0 ist.

3.7. Zwei "Normalformen" für Potenzreihen

Als nächstes nun ein Satz, auf dem unser Beweis von $\mathcal{B} \iff \mathcal{D}$ beruht.

Satz 3.11: Sei ρ eine ganzzahlige Potenzreihe, deren 0-ter Koeffizient gleich 1 ist.

(a) Es gibt genau eine Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ von ganzen Zahlen, für die

$$\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} = \rho \quad (30)$$

gilt.

(b) Es gibt genau eine Folge $(y_1, y_2, y_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ von ganzen Zahlen, für die

$$\prod_{d=1}^{\infty} (1 - X^d)^{-y_d} = \rho \quad (31)$$

gilt.

Hierbei muss man die unendlichen Produkte $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ und $\prod_{d=1}^{\infty} (1 - X^d)^{-y_d}$ natürlich als Grenzwerte von endlichen Produkten verstehen (also $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} = \lim_{n \rightarrow \infty} \prod_{d=1}^n (1 - x_d X^d)^{-1}$ und $\prod_{d=1}^{\infty} (1 - X^d)^{-y_d} = \lim_{n \rightarrow \infty} \prod_{d=1}^n (1 - X^d)^{-y_d}$).

Wir wollen den *Beweis von Satz 3.11* nur andeuten. Der Leser, der einen komplett ausformulierten Beweis vorzieht, kann einen solchen für Satz 3.11 **(a)** (sogar in etwas allgemeinerer Form) in [12] (Theorem 7 **(a)**) finden⁴³.

Wir wollen zuerst Satz 3.11 **(a)** beweisen. Wir schreiben die Potenzreihe ρ in der Form $\rho = \sum_{n=0}^{\infty} \rho_n X^n$. Für jedes $i \in \mathbb{N}$ ist also ρ_i der i -te Koeffizient der Potenzreihe ρ ; insbesondere ist ρ_0 der 0-te Koeffizient dieser Potenzreihe, und folglich gleich 1.

Um Satz 3.11 **(a)** zu beweisen, müssen wir nun eine Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}^+}$ von ganzen Zahlen finden, für die (30) gilt, und beweisen, daß diese Folge die einzige solche Folge ist.

Es ist klar, daß die 0-ten Koeffizienten der Potenzreihen $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ und $\sum_{n=0}^{\infty} \rho_n X^n$ immer gleich sind, egal wie man die Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}^+}$ wählt (denn $1 = \rho_0$).

Wir wollen jetzt sehen, was es bedeutet, daß die 1-ten Koeffizienten dieser Potenzreihen gleich sind: Wegen

$$\begin{aligned} & \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} \\ &= \prod_{d=1}^{\infty} (1 + x_d X^d + x_{2d} X^{2d} + x_{3d} X^{3d} + \dots) \quad (\text{nach der geometrischen Reihenformel}) \\ &= (1 + x_1 X^1 + x_1^2 X^2 + x_1^3 X^3 + \dots) (1 + x_2 X^2 + x_2^2 X^4 + x_2^3 X^6 + \dots) (1 + x_3 X^3 + x_3^2 X^6 + x_3^3 X^9 + \dots) \dots \\ &= 1 + x_1 X^1 + (\text{Glieder mit } X^i \text{ für } i > 1) \end{aligned}$$

ist der 1-te Koeffizient der Potenzreihe $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ gleich x_1 , und der 1-te Koeffizient der Potenzreihe $\sum_{n=0}^{\infty} \rho_n X^n$ ist natürlich ρ_1 . Damit die 1-ten Koeffizienten der Potenzreihen $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ und $\sum_{n=0}^{\infty} \rho_n X^n$ gleich sind, muss also $x_1 = \rho_1$ sein.

Nun wollen wir wissen, wann die 2-ten Koeffizienten dieser Potenzreihen gleich sind:

$$\begin{aligned} & \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} = \prod_{d=1}^{\infty} (1 + x_d X^d + x_{2d} X^{2d} + x_{3d} X^{3d} + \dots) \quad (\text{nach der geometrischen Reihenformel}) \\ &= (1 + x_1 X^1 + x_1^2 X^2 + x_1^3 X^3 + \dots) (1 + x_2 X^2 + x_2^2 X^4 + x_2^3 X^6 + \dots) (1 + x_3 X^3 + x_3^2 X^6 + x_3^3 X^9 + \dots) \dots \\ &= 1 + x_1 X^1 + x_1^2 X^2 + x_2 X^2 + (\text{Glieder mit } X^i \text{ für } i > 2) \\ &= 1 + x_1 X^1 + (x_1^2 + x_2) X^2 + (\text{Glieder mit } X^i \text{ für } i > 2) \end{aligned}$$

ist der 2-te Koeffizient der Potenzreihe $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ gleich $x_1^2 + x_2$, und der 2-te Koeffizient der Potenzreihe $\sum_{n=0}^{\infty} \rho_n X^n$ ist natürlich ρ_2 . Damit die 2-ten Koeffizienten der Potenzreihen $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ und $\sum_{n=0}^{\infty} \rho_n X^n$ gleich sind, muss also $x_1^2 + x_2 = \rho_2$ sein.

⁴³Dieses Theorem 7 steht zwar recht weit am Ende von [12], jedoch braucht man den Text davor nicht zu lesen, um den Beweis zu verstehen.

Als nächstes schauen wir, wann die 3-ten Koeffizienten dieser Potenzreihen gleich sind:

$$\begin{aligned} \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} &= \prod_{d=1}^{\infty} (1 + x_d X^d + x_{2d} X^{2d} + x_{3d} X^{3d} + \dots) && \text{(nach der geometrischen Reihenformel)} \\ &= (1 + x_1 X^1 + x_1^2 X^2 + x_1^3 X^3 + \dots) (1 + x_2 X^2 + x_2^2 X^4 + x_2^3 X^6 + \dots) (1 + x_3 X^3 + x_3^2 X^6 + x_3^3 X^9 + \dots) \dots \\ &= 1 + x_1 X^1 + x_1^2 X^2 + x_2 X^2 + x_1^3 X^3 + x_1 x_2 X^3 + x_3 X^3 + (\text{Glieder mit } X^i \text{ für } i > 3) \\ &= 1 + x_1 X^1 + (x_1^2 + x_2) X^2 + (x_1^3 + x_1 x_2 + x_3) X^3 + (\text{Glieder mit } X^i \text{ für } i > 3) \end{aligned}$$

ist der 3-te Koeffizient der Potenzreihe $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ gleich $x_1^3 + x_1 x_2 + x_3$, und der 3-te Koeffizient der Potenzreihe $\sum_{n=0}^{\infty} \rho_n X^n$ ist natürlich ρ_3 . Damit die 3-ten Koeffizienten der Potenzreihen $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ und $\sum_{n=0}^{\infty} \rho_n X^n$ gleich sind, muss also $x_1^3 + x_1 x_2 + x_3 = \rho_3$ sein.

Dieses Spiel kann man weitertreiben: Sei $i \in \mathbb{N}_+$. Um zu sehen, wann die i -ten Koeffizienten der Potenzreihen $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ und $\sum_{n=0}^{\infty} \rho_n X^n$ gleich sind, multiplizieren wir das Produkt $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ soweit aus, daß wir den i -ten Koeffizient dieses Produktes erkennen. Dieser i -te Koeffizient ist dann ein festes Polynom (mit ganzzahligen Koeffizienten) von x_1, x_2, \dots, x_i . Nennen wir dieses Polynom P_i . Dann sind also die i -ten Koeffizienten der Potenzreihen $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ und $\sum_{n=0}^{\infty} \rho_n X^n$ genau dann gleich, wenn $P_i(x_1, x_2, \dots, x_i) = \rho_i$ ist. Beispielsweise ist

$$\begin{aligned} P_1(x_1) &= x_1; \\ P_2(x_1, x_2) &= x_1^2 + x_2; \\ P_3(x_1, x_2, x_3) &= x_1^3 + x_1 x_2 + x_3. \end{aligned}$$

Man erkennt hieran eine Gesetzmäßigkeit: Für jedes $i \in \mathbb{N}_+$ hat das Polynom P_i die Eigenschaft, daß darin x_i nur einmal vorkommt, und zwar mit Koeffizient 1 davor und in der ersten Potenz. Mit anderen Worten: $P_i(x_1, x_2, \dots, x_i) = x_i + (\text{ein Polynom in } x_1, x_2, \dots, x_{i-1})$. (Dies erklärt sich auch sehr schnell, wenn man $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ ausmultipliziert.)

Wir wollen nun eine Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ finden, für welche die Potenzreihen $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ und $\sum_{n=0}^{\infty} \rho_n X^n$ gleich sind. Das heißt, daß für jedes $i \in \mathbb{N}_+$ die i -ten Koeffizienten dieser beiden Potenzreihen gleich sind (denn die 0-ten Koeffizienten dieser Reihen sind sowieso immer gleich, wie wir schon gesehen haben). Wie wir wissen, ist dies äquivalent zu

$$P_i(x_1, x_2, \dots, x_i) = \rho_i \quad \text{für jedes } i \in \mathbb{N}_+.$$

Dies ist ein Gleichungssystem in den Unbekannten x_1, x_2, x_3, \dots mit unendlich vielen Gleichungen. Da nun

$$P_i(x_1, x_2, \dots, x_i) = x_i + (\text{ein Polynom in } x_1, x_2, \dots, x_{i-1}) \quad \text{für jedes } i \in \mathbb{N}_+$$

gilt, können wir dieses Gleichungssystem rekursiv nach x_1, x_2, x_3, \dots auflösen: Aus der i -ten Gleichung $P_i(x_1, x_2, \dots, x_i) = \rho_i$ erhalten wir x_i , wenn wir die bereits gefundenen Werte von x_1, x_2, \dots, x_{i-1} einsetzen. Auf diese Weise erhalten wir eine eindeutige Lösungsfolge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$, für welche die Potenzreihen $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ und $\sum_{n=0}^{\infty} \rho_n X^n$ gleich sind. Damit ist Satz 3.11 (a) bewiesen.

Der Beweis von Satz 3.11 (b) ergibt sich analog, wenn man statt der Formel

$$(1 - x_d X^d)^{-1} = 1 + x_d X^d + x_{2d} X^{2d} + x_{3d} X^{3d} + \dots$$

immer die Formel

$$(1 - X^d)^{-y_d} = \sum_{i=0}^{\infty} \binom{-y_d}{i} (-X^d)^i = 1 + y_d X^d + \binom{-y_d}{2} X^{2d} - \binom{-y_d}{3} X^{3d} \pm \dots$$

anwendet (diese Formel folgt aus Satz 3.1). Es ist zu beachten, daß wir hierbei statt den Polynomen P_i neue Polynome bekommen, die nicht mehr ganzzahlige, sondern nunmehr rationale Koeffizienten haben, aber immer noch beim Einsetzen ganzer Zahlen ganze Werte ergeben (denn diese Polynome sind aus Binomialkoeffizienten zusammengesetzt); genaueres über solche Polynome werden wir in Abschnitt 6 erfahren.

3.8. Beweis von Satz 2.1: ein alternativer Beweis von $\mathcal{B} \iff \mathcal{D}$

Nun endlich zum eigentlichen alternativen Beweis von $\mathcal{B} \iff \mathcal{D}$:

Beweis von $\mathcal{B} \implies \mathcal{D}$: Angenommen, Aussage \mathcal{B} von Satz 2.1 sei wahr. Dann gibt es eine Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ ganzer Zahlen, die (5) erfüllt. Sei nun ρ die Potenzreihe $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$. Nach Satz 3.5 (b) ist dann

$$\begin{aligned} \ln \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} &= \sum_{d=1}^{\infty} \underbrace{\ln (1 - x_d X^d)^{-1}}_{\substack{= -\ln(1 - x_d X^d) \\ \text{(nach Satz 3.9 (a))}}} = - \sum_{d=1}^{\infty} \ln (1 - x_d X^d) = - \sum_{d=1}^{\infty} \left(- \sum_{i=1}^{\infty} \frac{x_d^i X^{di}}{i} \right) \\ &\left(\begin{array}{l} \text{denn nach der Definition von } \ln \text{ ist} \\ \ln (1 - x_d X^d) = - \sum_{i=1}^{\infty} \frac{(1 - (1 - x_d X^d))^i}{i} = - \sum_{i=1}^{\infty} \frac{(x_d X^d)^i}{i} = - \sum_{i=1}^{\infty} \frac{x_d^i X^{di}}{i} \end{array} \right) \\ &= \sum_{d=1}^{\infty} \sum_{i=1}^{\infty} \frac{x_d^i X^{di}}{i} = \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+ \\ d|n}} \frac{x_d^{n/d} X^n}{n/d} \\ &\quad \text{(hier haben wir in der zweiten Summe } di \text{ durch } n \text{ substituiert)} \\ &= \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+ \\ d|n}} \frac{d}{n} x_d^{n/d} X^n \end{aligned}$$

und somit

$$\frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} \right) = \frac{d}{dX} \left(\sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+ \\ d|n}} \frac{d}{n} x_d^{n/d} X^n \right) = \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+ \\ d|n}} \frac{d}{n} x_d^{n/d} \underbrace{\frac{d}{dX} X^n}_{=nX^{n-1}} = \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+ \\ d|n}} dx_d^{n/d} X^{n-1},$$

also

$$X \cdot \frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} \right) = X \cdot \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+; \\ d|n}} dx_d^{n/d} X^{n-1} = \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+; \\ d|n}} dx_d^{n/d} X^n = \sum_{n=1}^{\infty} \sum_{d|n} dx_d^{n/d} X^n.$$

Nach (5) ist aber $\sum_{d|n} dx_d^{n/d} = b_n$ für jedes $n \in \mathbb{N}_+$; somit wird dies zu

$$X \cdot \frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} \right) = \sum_{n=1}^{\infty} \underbrace{\sum_{d|n} dx_d^{n/d} X^n}_{=b_n} = \sum_{n=1}^{\infty} b_n X^n. \quad (32)$$

Nun ist $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ eine ganzzahlige Potenzreihe, deren 0-ter Koeffizient gleich 1 ist. Nach Satz 3.11 (b) (angewandt auf $\rho = \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$) gibt es daher genau eine Folge $(y_1, y_2, y_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ von ganzen Zahlen, für die $\prod_{d=1}^{\infty} (1 - X^d)^{-y_d} = \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ gilt. Nach Satz 3.5 (b) ist

$$\begin{aligned} \ln \prod_{d=1}^{\infty} (1 - X^d)^{-y_d} &= \sum_{d=1}^{\infty} \underbrace{\ln (1 - X^d)^{-y_d}}_{\substack{=-y_d \ln(1 - X^d) \\ \text{(nach Satz 3.9 (b))}}} = - \sum_{d=1}^{\infty} y_d \ln (1 - X^d) = - \sum_{d=1}^{\infty} y_d \left(- \sum_{i=1}^{\infty} \frac{X^{di}}{i} \right) \\ &\left(\begin{array}{l} \text{denn nach der Definition von } \ln \text{ ist} \\ \ln (1 - X^d) = - \sum_{i=1}^{\infty} \frac{(1 - (1 - X^d))^i}{i} = - \sum_{i=1}^{\infty} \frac{(X^d)^i}{i} = - \sum_{i=1}^{\infty} \frac{X^{di}}{i} \end{array} \right) \\ &= \sum_{d=1}^{\infty} \sum_{i=1}^{\infty} y_d \frac{X^{di}}{i} = \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+; \\ d|n}} y_d \frac{X^n}{n/d} \\ &\quad \text{(hier haben wir in der zweiten Summe } di \text{ durch } n \text{ substituiert)} \\ &= \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+; \\ d|n}} \frac{d}{n} y_d X^n \end{aligned}$$

und somit

$$\frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - X^d)^{-y_d} \right) = \frac{d}{dX} \left(\sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+; \\ d|n}} \frac{d}{n} y_d X^n \right) = \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+; \\ d|n}} \frac{d}{n} y_d \underbrace{\frac{d}{dX} X^n}_{=nX^{n-1}} = \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+; \\ d|n}} dy_d X^{n-1},$$

also

$$X \cdot \frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - X^d)^{-y_d} \right) = X \cdot \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+; \\ d|n}} dy_d X^{n-1} = \sum_{d=1}^{\infty} \sum_{\substack{n \in \mathbb{N}_+; \\ d|n}} dy_d X^n = \sum_{n=1}^{\infty} \sum_{d|n} dy_d X^n. \quad (33)$$

Wegen $\prod_{d=1}^{\infty} (1 - X^d)^{-y_d} = \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1}$ ist aber

$$X \cdot \frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - X^d)^{-y_d} \right) = X \cdot \frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} \right),$$

was im Lichte von (33) und (32) zu

$$\sum_{n=1}^{\infty} \sum_{d|n} dy_d X^n = \sum_{n=1}^{\infty} b_n X^n$$

wird. Da zwei gleiche Potenzreihen auch immer gleiche Koeffizienten haben, folgt hieraus $\sum_{d|n} dy_d = b_n$ für jedes $n \in \mathbb{N}_+$. Folglich gilt Aussage \mathcal{D} . Wir haben also die Implikation $\mathcal{B} \implies \mathcal{D}$ gezeigt.

Der *Beweis von $\mathcal{D} \implies \mathcal{B}$* verläuft genau umgekehrt: Wir nehmen an, Aussage \mathcal{D} von Satz 2.1 sei wahr. Dann gibt es eine Folge $(y_1, y_2, y_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ ganzer Zahlen, die (6) erfüllt. Sei nun ρ die Potenzreihe $\prod_{d=1}^{\infty} (1 - X^d)^{-y_d}$. Genauso wie im Beweis von $\mathcal{B} \implies \mathcal{D}$ vorhin zeigen wir, daß

$$X \cdot \frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - X^d)^{-y_d} \right) = \sum_{n=1}^{\infty} \sum_{d|n} dy_d X^n$$

gilt. Doch da $\sum_{d|n} dy_d = b_n$ für jedes $n \in \mathbb{N}_+$ gilt (nach (6)), wird dies zu

$$X \cdot \frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - X^d)^{-y_d} \right) = \sum_{n=1}^{\infty} \underbrace{\sum_{d|n} dy_d}_{=b_n} X^n = \sum_{n=1}^{\infty} b_n X^n. \quad (34)$$

Andererseits ist $\prod_{d=1}^{\infty} (1 - X^d)^{-y_d}$ eine ganzzahlige Potenzreihe, deren 0-ter Koeffizient gleich 1 ist. Nach Satz 3.11 (a) (angewandt auf $\rho = \prod_{d=1}^{\infty} (1 - X^d)^{-y_d}$) gibt es daher genau eine Folge $(x_1, x_2, x_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ von ganzen Zahlen, für die $\prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} = \prod_{d=1}^{\infty} (1 - X^d)^{-y_d}$ gilt. Hieraus folgt

$$X \cdot \frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} \right) = X \cdot \frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - X^d)^{-y_d} \right).$$

Aufgrund von (34) und der Identität

$$X \cdot \frac{d}{dX} \left(\ln \prod_{d=1}^{\infty} (1 - x_d X^d)^{-1} \right) = \sum_{n=1}^{\infty} \sum_{d|n} dx_d^{n/d} X^n$$

(welche man genauso wie im Beweis von $\mathcal{B} \implies \mathcal{D}$ vorhin zeigen kann) wird diese Gleichung zu

$$\sum_{n=1}^{\infty} b_n X^n = \sum_{n=1}^{\infty} \sum_{d|n} dx_d^{n/d} X^n.$$

Da zwei gleiche Potenzreihen auch immer gleiche Koeffizienten haben, folgt hieraus $b_n = \sum_{d|n} dx_d^{n/d}$ für jedes $n \in \mathbb{N}_+$. Folglich gilt Aussage \mathcal{B} . Damit ist die Implikation $\mathcal{D} \implies \mathcal{B}$ bewiesen.

Jetzt, nachdem beide Implikationen $\mathcal{B} \implies \mathcal{D}$ und $\mathcal{D} \implies \mathcal{B}$ bewiesen sind, ist unser alternativer Beweis von $\mathcal{B} \iff \mathcal{D}$ komplett.

Da Satz 3.11 nicht nur Existenzaussagen, sondern auch Eindeutigkeitsaussagen liefert, können wir diesen Beweis übrigens zu einem Beweis von $\mathcal{B} \iff \mathcal{C} \iff \mathcal{D} \iff \mathcal{E}$ erweitern. Die Details hiervon seien dem Leser überlassen.

4. Binomialkoeffizienten statt Potenzen

4.1. $\sum_{d|n} \phi(d) \binom{qn/d}{rn/d} \in n\mathbb{Z}$ und Analoga

Satz 2.1 gibt eine Reihe von äquivalenten Eigenschaften einer Folge $(b_1, b_2, b_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$. Aber wie sehen Folgen aus, die diese Eigenschaften erfüllen? Wir kennen schon eine Reihe von Beispielen für solche Folgen: Für jedes $q \in \mathbb{Z}$ ist $(b_1, b_2, b_3, \dots) = (q^1, q^2, q^3, \dots)$ eine solche Folge. Aber es gibt auch ein weiteres Beispiel:

Satz 4.1: Seien $q \in \mathbb{Z}$ und $r \in \mathbb{Q}$. Dann erfüllt die Folge (b_1, b_2, b_3, \dots) , welche durch $b_n = \binom{qn}{rn}$ für alle $n \in \mathbb{N}_+$ definiert ist, die Aussagen \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} und \mathcal{H} von Satz 2.1.

An dieser Stelle ist erstmal Mißtrauen angebracht: Was bedeutet $\binom{qn}{rn}$ für ein $r \in \mathbb{Q}$? Wir kennen Binomialkoeffizienten $\binom{n}{k}$ mit natürlichem k , und seit Abschnitt 3 auch welche mit ganzem k (auch wenn wir sie für negatives k einfach als 0 definiert haben). Aber mit rationalem k ? Nun, die Antwort ist eine sehr billige:

Definition (Binomialkoeffizient $\binom{n}{k}$ mit $n \in \mathbb{Z}$ und $k \in \mathbb{Q}$): Wir definieren den sogenannten *Binomialkoeffizienten* $\binom{n}{k}$ für alle $n \in \mathbb{Z}$ (oder $n \in \mathbb{Q}$, oder $n \in \mathbb{R}$, oder $n \in \mathbb{C}$) und $k \in \mathbb{Q}$ durch

$$\binom{n}{k} = \begin{cases} \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!}, & \text{wenn } k \in \mathbb{N}; \\ 0, & \text{wenn } k \notin \mathbb{N} \end{cases} .$$

Wieder sorgt diese Definition dafür, daß die Rekursionsgleichung $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$ für alle $n \in \mathbb{Z}$ (oder $n \in \mathbb{Q}$, etc.) und $k \in \mathbb{Q}$ gilt.

Diese Definition sieht ein wenig inhaltsleer aus: Was haben wir davon gewonnen, daß wir einige "neue" Werte von $\binom{n}{k}$ definiert haben, indem wir sie einfach zu 0 gesetzt haben? Das mag so sein, aber diese Definition führt dazu, daß Satz 4.1 für alle $r \in \mathbb{Q}$ (und nicht nur für alle $r \in \mathbb{Z}$) gilt, was tatsächlich eine gewisse Bereicherung darstellt. (Für gebrochenes r hat die Folge (b_1, b_2, b_3, \dots) zwar viele Nullen, aber sie ist immer noch nicht die Folge $(0, 0, 0, \dots)$, außer wenn $r > q \geq 0$ oder $r < 0$ ist. Insofern ist es keine Trivialität, daß sie die Aussagen $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}$ und \mathcal{H} von Satz 4.1 erfüllt!)

Bevor wir zum Beweis von Satz 4.1 kommen, machen wir uns klar, was Satz 4.1 impliziert. So folgt aus Aussage \mathcal{G} von Satz 2.1, angewandt auf diese Folge, daß

$$\sum_{d|n} \phi(d) \binom{qn/d}{rn/d} \in n\mathbb{Z}$$

gilt; entsprechend läßt sich aus Aussage \mathcal{F} auf

$$\sum_{d|n} \mu(d) \binom{qn/d}{rn/d} \in n\mathbb{Z}$$

schließen, und Aussage \mathcal{H} von Satz 2.1, angewandt auf dieselbe Folge, ergibt

$$\sum_{i=1}^n \binom{q \operatorname{ggT}(i, n)}{r \operatorname{ggT}(i, n)} \in n\mathbb{Z}.$$

Nun stellt man nach ein wenig Experimentieren fest, daß diese Aussagen Analoga zulassen:

Satz 4.2: Seien $q \in \mathbb{Z}$ und $r \in \mathbb{Z}$ (hier, im Unterschied zum Satz 4.1, fordern wir $r \in \mathbb{Z}$ und nicht nur $r \in \mathbb{Q}$). Für jedes $n \in \mathbb{N}_+$ gilt dann

$$\sum_{d|n} \phi(d) \binom{qn/d}{rn/d} \in \frac{q}{r} n\mathbb{Z}; \quad \sum_{d|n} \mu(d) \binom{qn/d}{rn/d} \in \frac{q}{r} n\mathbb{Z}; \quad \sum_{i=1}^n \binom{q \operatorname{ggT}(i, n)}{r \operatorname{ggT}(i, n)} \in \frac{q}{r} n\mathbb{Z}.$$

Dieser Satz 4.2 folgt nicht mehr direkt aus Satz 4.1, jedoch aus folgendem Analogon:

Satz 4.3: Seien $q \in \mathbb{Z}$ und $r \in \mathbb{Z}$ (hier fordern wir wieder $r \in \mathbb{Z}$, wie in Satz 4.2). Dann erfüllt die Folge (b_1, b_2, b_3, \dots) , welche durch $b_n = \binom{qn-1}{rn-1}$ für alle $n \in \mathbb{N}_+$ definiert ist, die Aussagen $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}$ und \mathcal{H} von Satz 2.1.

Warum folgt Satz 4.2 aus Satz 4.3?

Beweis von Satz 4.2: Wir stellen erst einmal fest, daß für beliebige ganze a und b mit $a \neq 0$ die Formel

$$\binom{a-1}{b-1} = \binom{a}{b} \frac{b}{a} \tag{35}$$

gilt.⁴⁴

Nun definiere man eine Folge (b_1, b_2, b_3, \dots) durch $\left(b_n = \binom{qn-1}{rn-1}\right)$ für alle $n \in \mathbb{N}_+$. Laut Satz 4.3 erfüllt diese Folge dann die Aussage \mathcal{G} von Satz 2.1. Das heißt,

$$\sum_{d|n} \phi(d) \binom{qn/d-1}{rn/d-1} \in n\mathbb{Z}.$$

Dies läßt sich aber als

$$\sum_{d|n} \phi(d) \binom{qn/d}{rn/d} \frac{r}{q} \in n\mathbb{Z}$$

umschreiben, denn $\binom{qn/d-1}{rn/d-1} = \binom{qn/d}{rn/d} \frac{r}{q}$ (nach (35), angewandt auf $a = qn/d$ und $b = rn/d$). Mit anderen Worten:

$$\sum_{d|n} \phi(d) \binom{qn/d}{rn/d} \in \frac{q}{r}n\mathbb{Z}.$$

Analog folgert man aus Aussage \mathcal{F} von Satz 2.1 die Relation

$$\sum_{d|n} \mu(d) \binom{qn/d}{rn/d} \in \frac{q}{r}n\mathbb{Z},$$

und aus Aussage \mathcal{H} erhält man

$$\sum_{i=1}^n \binom{q \text{ ggT}(i, n)}{r \text{ ggT}(i, n)} \in \frac{q}{r}n\mathbb{Z}.$$

Damit ist Satz 4.2 unter Verwendung von Satz 4.3 bewiesen.

Übrigens verallgemeinert Satz 4.2 eine recht bekannte Olympiadeaufgabe (siehe [14]).

Man könnte wegen den Sätzen 4.1 und 4.3 auf den Gedanken kommen, man könnte auch eine Folge (b_1, b_2, b_3, \dots) durch $\left(b_n = \binom{qn-\alpha}{rn-\alpha}\right)$ für alle $n \in \mathbb{N}_+$ definieren, wobei α eine feste ganze Zahl ist, und würde auch für sie die Aussagen \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} und \mathcal{H} von Satz 2.1 feststellen. Tatsache ist aber, daß dies für $\alpha \notin \{0, 1\}$ nicht unbedingt wahr ist (so ist $\alpha = -1$ ein Gegenbeispiel).

⁴⁴Diese Formel folgt im Falle von $b \in \mathbb{N}_+$ aus

$$\begin{aligned} \binom{a}{b} &= \frac{a(a-1)\dots(a-b+1)}{b!} = \frac{a \cdot (a-1)(a-2)\dots(a-b+1)}{b \cdot (b-1)!} \\ &= \frac{a}{b} \cdot \underbrace{\frac{(a-1)(a-2)\dots(a-b+1)}{(b-1)!}}_{= \binom{a-1}{b-1}}, \end{aligned}$$

und im Falle von $b \notin \mathbb{N}_+$ ist sie sowieso trivial.

4.2. Lemmata zum Beweis

Wir wollen nun die Sätze 4.1 und 4.3 beweisen. Beidesmal wird dies geschehen, indem wir Aussage \mathcal{A} für unsere Folge (b_1, b_2, b_3, \dots) nachweisen. Dazu benötigen wir die folgenden zwei Lemmata⁴⁵:

Lemma 4.4: Seien $q \in \mathbb{Z}$, $r \in \mathbb{Q}$ und $n \in \mathbb{N}_+$. Sei p ein Primfaktor von n .

Dann gilt $\binom{qn/p}{rn/p} \equiv \binom{qn}{rn} \pmod{p^{v_p(n)}}$.

Lemma 4.5: Seien $q \in \mathbb{Z}$, $r \in \mathbb{Z}$ und $n \in \mathbb{N}_+$. Sei p ein Primfaktor von n .

Dann gilt $\binom{qn/p-1}{rn/p-1} \equiv \binom{qn-1}{rn-1} \pmod{p^{v_p(n)}}$.

Wir wollen zuerst Lemma 4.4 zeigen; dafür gibt es zwei Wege:

Erster Beweis von Lemma 4.4: Zuerst führen wir eine einfache (aber sehr nützliche) Sprechweise ein. Und zwar geht es um den Begriff einer *p-ganzen Zahl*. Die Definition ist die folgende:

Definition (p-ganze Zahl): Sei p eine Primzahl. Eine rationale Zahl $u \in \mathbb{Q}$ heißt *p-ganz*, wenn man u in der Form $u = \frac{\alpha}{\beta}$ schreiben kann, wobei α und β ganze Zahlen sind mit $p \nmid \beta$. In Worten: Eine ganze Zahl heißt *p-ganz*, wenn man sie als Quotient zweier ganzer Zahlen schreiben kann, und der Nenner nicht durch p teilbar ist.

Es ist leicht zu sehen, daß die Summe zweier *p-ganzer Zahlen* wieder *p-ganz* ist⁴⁶, und daß das Produkt zweier *p-ganzer Zahlen* wieder *p-ganz* ist⁴⁷ (für beide diese Aussagen brauchen wir die Annahme, daß p prim ist!). Ferner ist klar, daß jede ganze Zahl auch *p-ganz* ist.

Beispiele: Für $p = 2$ sind die Zahlen 1, 14 und $\frac{2}{3}$ alle *p-ganz*, aber die Zahl $\frac{3}{2}$ nicht.

Nun eine weitere Definition:

⁴⁵Lemma 4.4 ist identisch mit Lemma 19 in [3]. Lemma 4.5 ist ein Sonderfall von Lemma 21 in [3].

⁴⁶*Beweis:* Seien u und v zwei *p-ganze Zahlen*. Wir müssen dann beweisen, daß $u + v$ wieder eine *p-ganze Zahl* ist. In der Tat kann man u in der Form $u = \frac{\alpha}{\beta}$ schreiben, wobei α und β ganze Zahlen sind mit $p \nmid \beta$ (denn u ist *p-ganz*). Analog ergibt sich, daß man v in der Form $v = \frac{\gamma}{\delta}$ schreiben kann, wobei γ und δ ganze Zahlen sind mit $p \nmid \delta$ (denn v ist *p-ganz*). Somit ist $u + v = \frac{\alpha}{\beta} + \frac{\gamma}{\delta} = \frac{\alpha\delta + \beta\gamma}{\beta\delta}$. Setzen wir $\varepsilon = \alpha\delta + \beta\gamma$ und $\phi = \beta\delta$, dann ist also $u + v = \frac{\varepsilon}{\phi}$. Da $p \nmid \phi$ ist (denn da p prim ist, folgt aus $p \nmid \beta$ und $p \nmid \delta$ sofort $p \nmid \beta\delta = \phi$), folgt hieraus, daß $u + v$ eine *p-ganze Zahl* ist, was zu beweisen war.

⁴⁷*Beweis:* Seien u und v zwei *p-ganze Zahlen*. Wir müssen dann beweisen, daß $u \cdot v$ wieder eine *p-ganze Zahl* ist. In der Tat kann man u in der Form $u = \frac{\alpha}{\beta}$ schreiben, wobei α und β ganze Zahlen sind mit $p \nmid \beta$ (denn u ist *p-ganz*). Analog ergibt sich, daß man v in der Form $v = \frac{\gamma}{\delta}$ schreiben kann, wobei γ und δ ganze Zahlen sind mit $p \nmid \delta$ (denn v ist *p-ganz*). Somit ist $u \cdot v = \frac{\alpha}{\beta} \cdot \frac{\gamma}{\delta} = \frac{\alpha\gamma}{\beta\delta}$. Setzen wir $\varepsilon = \alpha\gamma$ und $\phi = \beta\delta$, dann ist also $u \cdot v = \frac{\varepsilon}{\phi}$. Da $p \nmid \phi$ ist (denn da p prim ist, folgt aus $p \nmid \beta$ und $p \nmid \delta$ sofort $p \nmid \beta\delta = \phi$), folgt hieraus, daß $u \cdot v$ eine *p-ganze Zahl* ist, was zu beweisen war.

Definition ($u \stackrel{p}{\equiv} v \pmod{p^k}$): Für zwei p -ganze Zahlen u und v und eine natürliche Zahl $k \in \mathbb{N}$ schreiben wir $u \stackrel{p}{\equiv} v \pmod{p^k}$ genau dann, wenn die Zahl $\frac{u-v}{p^k}$ auch p -ganz ist. Man sieht nun leicht, daß folgendes gilt: Wenn die Zahlen u und v ganz sind (und nicht nur p -ganz sind), dann ist genau dann $u \stackrel{p}{\equiv} v \pmod{p^k}$, wenn $u \equiv v \pmod{p^k}$ ist⁴⁸. Somit können wir (für zwei p -ganze Zahlen u und v eine natürliche Zahl $k \in \mathbb{N}$) statt unserer Schreibweise $u \stackrel{p}{\equiv} v \pmod{p^k}$ auch einfach $u \equiv v \pmod{p^k}$ schreiben, ohne (im Falle ganzer Zahlen u und v) mit der altbekannten Notation $u \equiv v \pmod{p^k}$ in Konflikt zu geraten (denn im Falle ganzer Zahlen u und v sind die neue Notation $u \stackrel{p}{\equiv} v \pmod{p^k}$ und die altbekannte Notation $u \equiv v \pmod{p^k}$ gleichbedeutend). Für zwei p -ganze Zahlen u und v und eine natürliche Zahl $k \in \mathbb{N}$ hat man also folgende Äquivalenz von Aussagen:

$$\left(u \equiv v \pmod{p^k}\right) \iff \left(u \stackrel{p}{\equiv} v \pmod{p^k}\right) \iff \left(\frac{u-v}{p^k} \text{ ist } p\text{-ganz}\right).$$

Beispiele: Für $p = 2$ gilt $\frac{1}{3} \equiv \frac{5}{3} \pmod{p^2}$ (denn $\frac{\frac{1}{3} - \frac{5}{3}}{p^2} = \frac{-\frac{4}{3}}{2^2} = \frac{-1}{3}$ ist p -ganz) und $\frac{1}{9} \equiv \frac{11}{3} \pmod{p^4}$ (denn $\frac{\frac{1}{9} - \frac{11}{3}}{p^4} = \frac{-\frac{32}{9}}{2^4} = \frac{-2}{9}$ ist p -ganz), aber $\frac{1}{3} \not\equiv 1 \pmod{p^2}$.

Wir können nun genauso mit p -ganzen Zahlen "modulo p^k rechnen", wie wir es mit ganzen Zahlen gewohnt sind: Wenn u, v, u' und v' vier p -ganze Zahlen und $k \in \mathbb{N}$ eine natürliche Zahl sind, und wenn $u \equiv v \pmod{p^k}$ und $u' \equiv v' \pmod{p^k}$ erfüllt ist, dann gilt auch $u + u' \equiv v + v' \pmod{p^k}$ und $uu' \equiv vv' \pmod{p^k}$. Dies beweist man genauso wie für ganze Zahlen.

Was ist aber der Nutzen davon, mit p -ganzen Zahlen modulo p^k zu rechnen? Ganz einfach: Um eine Kongruenz zwischen zwei ganzen Zahlen modulo p^k nachzuprüfen, ist es oft hilfreich, unterwegs mit p -ganzen Zahlen zu rechnen, obwohl am Ende alles

⁴⁸*Beweis:* Wenn die Zahl $\frac{u-v}{p^k}$ eine p -ganze Zahl ist, dann ist sie auch ganz (denn da die Zahl $\frac{u-v}{p^k}$ eine p -ganze Zahl ist, läßt sie sich in der Form $\frac{u-v}{p^k} = \frac{\alpha}{\beta}$ schreiben, wobei α und β ganze Zahlen sind mit $p \nmid \beta$; somit ist $\alpha p^k = \beta(u-v)$, also $p^k \mid \beta(u-v)$ und folglich $p^k \mid u-v$ (denn da p prim und $p \nmid \beta$ ist, ist β teilerfremd zu p^k), also $\frac{u-v}{p^k} \in \mathbb{Z}$). Umgekehrt: Wenn die Zahl $\frac{u-v}{p^k}$ ganz ist, dann ist sie auch p -ganz (weil jede ganze Zahl automatisch auch p -ganz ist). Somit haben wir folgende Äquivalenz von Aussagen:

$$\left(\frac{u-v}{p^k} \text{ ist } p\text{-ganz}\right) \iff \left(\frac{u-v}{p^k} \text{ ist ganz}\right).$$

Damit ergibt sich folgende Kette von Äquivalenzen:

$$\left(u \stackrel{p}{\equiv} v \pmod{p^k}\right) \iff \left(\frac{u-v}{p^k} \text{ ist } p\text{-ganz}\right) \iff \left(\frac{u-v}{p^k} \text{ ist ganz}\right) \iff (p^k \mid u-v) \iff (u \equiv v \pmod{p^k}),$$

was zu beweisen war.

wieder ganz ist. Ein Beispiel dafür ist die Aufgabe 1 der Internationalen Mathematik-Olympiade 2005⁴⁹; ein anderes Beispiel dafür ist der Beweis von Lemma 4.4, den wir jetzt geben werden.

Zum Beweis von Lemma 4.4 unterscheiden wir drei Fälle:

Fall 1: Es gilt $rn/p \in \mathbb{N}$.

Fall 2: Es gilt $rn/p \notin \mathbb{N}$ und $rn \in \mathbb{N}$.

Fall 3: Es gilt $rn/p \notin \mathbb{N}$ und $rn \notin \mathbb{N}$.

In Fall 3 ist Lemma 4.4 offensichtlich, denn wegen $rn/p \notin \mathbb{N}$ ist $\binom{qn/p}{rn/p} = 0$ und wegen $rn \notin \mathbb{N}$ ist $\binom{qn}{rn} = 0$.

Betrachten wir nun den Fall 1. In diesem Fall sei $\gamma = \min \{v_p(rn), v_p(qn)\}$. Dann ist $\gamma \leq v_p(rn)$ und damit $p^\gamma \mid rn$, und genauso ergibt sich $\gamma \leq v_p(qn)$ und damit $p^\gamma \mid qn$. Andererseits ist im Fall 1 trivialerweise $rn \in \mathbb{N}$ (da $rn/p \in \mathbb{N}$ und $p \in \mathbb{N}$), und somit können wir den Binomialkoeffizienten $\binom{qn}{rn}$ als

$$\binom{qn}{rn} = \frac{(qn)(qn-1)\dots(qn-rn+1)}{(rn)!}$$

darstellen. Wegen

$$(qn)(qn-1)\dots(qn-rn+1) = \prod_{k \in \{0,1,\dots,rn-1\}} (qn-k) = \prod_{\substack{k \in \{0,1,\dots,rn-1\}; \\ p \mid k}} (qn-k) \cdot \prod_{\substack{k \in \{0,1,\dots,rn-1\}; \\ p \nmid k}} (qn-k)$$

und

$$(rn)! = \prod_{k \in \{0,1,\dots,rn-1\}} (rn-k) = \prod_{\substack{k \in \{0,1,\dots,rn-1\}; \\ p \mid k}} (rn-k) \cdot \prod_{\substack{k \in \{0,1,\dots,rn-1\}; \\ p \nmid k}} (rn-k)$$

⁴⁹*Aufgabe:* Man bestimme alle ganzen Zahlen, die teilerfremd zu jedem Element der durch $(a_n = 2^n + 3^n + 6^n - 1$ für alle $n \in \mathbb{N}_+$) definierten Folge (a_1, a_2, a_3, \dots) sind.

Lösung mithilfe p-ganzer Zahlen: Die einzige solche ganze Zahl ist 1, denn für jede Primzahl p gibt es ein $n \in \mathbb{N}_+$ mit $p \mid a_n$. In der Tat kann man $n = 2$ für $p \in \{2, 3\}$ setzen, und $n = p - 2$ für alle $p > 3$. Ersteres läßt sich einfach nachrechnen; um letzteres zu beweisen, müssen wir zeigen, daß $p \mid a_{p-2}$ für alle Primzahlen $p > 3$ gilt. Dazu bemerken wir, daß für jede zu p teilerfremde ganze Zahl u die Zahl $\frac{1}{u}$ eine p -ganze Zahl ist, und $u^{p-2} \equiv \frac{1}{u} \pmod p$ erfüllt (denn nach dem kleinen Satz von Fermat ist $u^{p-1} \equiv 1 \pmod p$, und Multiplikation dieser Kongruenz mit der Kongruenz $\frac{1}{u} \equiv \frac{1}{u} \pmod p$ ergibt $u^{p-2} \equiv \frac{1}{u} \pmod p$), und somit ist

$$a_{p-2} = 2^{p-2} + 3^{p-2} + 6^{p-2} - 1 \equiv \frac{1}{2} + \frac{1}{3} + \frac{1}{6} - 1 = 0 \pmod p,$$

also $p \mid a_{p-2}$, was zu beweisen war.

Dieser Beweis läßt sich zwar leicht ohne Benutzung von p -ganzen Zahlen umschreiben (man multipliziert einfach mit 6 und kürze am Ende 6 wieder heraus), aber man findet ihn am leichtesten, wenn man sich bewusst ist, daß man mit p -ganzen Zahlen modulo p^k (in unserem Fall einfach modulo p) genauso rechnen kann, wie mit ganzen Zahlen.

wird dies zu

$$\begin{aligned}
\binom{qn}{rn} &= \frac{\prod_{k \in \{0,1,\dots, rn-1\}; p|k} (qn - k) \cdot \prod_{k \in \{0,1,\dots, rn-1\}; p \nmid k} (qn - k)}{\prod_{k \in \{0,1,\dots, rn-1\}; p|k} (rn - k) \cdot \prod_{k \in \{0,1,\dots, rn-1\}; p \nmid k} (rn - k)} \\
&= \frac{\prod_{k \in \{0,1,\dots, rn-1\}; p|k} (qn - k)}{\prod_{k \in \{0,1,\dots, rn-1\}; p|k} (rn - k)} \cdot \frac{\prod_{k \in \{0,1,\dots, rn-1\}; p \nmid k} (qn - k)}{\prod_{k \in \{0,1,\dots, rn-1\}; p \nmid k} (rn - k)}. \tag{36}
\end{aligned}$$

Der erste Bruch auf der rechten Seite dieser Gleichung vereinfacht sich zu

$$\begin{aligned}
\frac{\prod_{k \in \{0,1,\dots, rn-1\}; p|k} (qn - k)}{\prod_{k \in \{0,1,\dots, rn-1\}; p|k} (rn - k)} &= \frac{\prod_{\ell \in \{0,1,\dots, rn/p-1\}} (qn - \ell p)}{\prod_{\ell \in \{0,1,\dots, rn/p-1\}} (rn - \ell p)} \\
&\quad \text{(hier haben wir } k \text{ durch } \ell p \text{ substituiert, da } p \mid k \text{ vorausgesetzt wurde)} \\
&= \frac{\prod_{\ell \in \{0,1,\dots, rn/p-1\}} ((qn/p - \ell) \cdot p)}{\prod_{\ell \in \{0,1,\dots, rn/p-1\}} ((rn/p - \ell) \cdot p)} = \frac{\prod_{\ell \in \{0,1,\dots, rn/p-1\}} (qn/p - \ell)}{\prod_{\ell \in \{0,1,\dots, rn/p-1\}} (rn/p - \ell)} \\
&\quad \text{(hier haben wir aus Zähler und Nenner jeweils } rn/p \text{ mal den Faktor } p \text{ herausgekürzt)} \\
&= \frac{(qn/p)(qn/p - 1) \dots (qn/p - rn/p + 1)}{(rn/p)!} = \binom{qn/p}{rn/p} \quad (\text{da } rn/p \in \mathbb{N}),
\end{aligned}$$

und der zweite Bruch auf der rechten Seite erfüllt

$$\begin{aligned}
\frac{\prod_{k \in \{0,1,\dots, rn-1\}; p \nmid k} (qn - k)}{\prod_{k \in \{0,1,\dots, rn-1\}; p \nmid k} (rn - k)} &= \prod_{k \in \{0,1,\dots, rn-1\}; p \nmid k} \frac{qn - k}{rn - k} \equiv \prod_{k \in \{0,1,\dots, rn-1\}; p \nmid k} 1 \\
&\quad \left(\begin{array}{l} \text{denn für jedes } k \in \{0, 1, \dots, rn - 1\} \text{ mit } p \nmid k \text{ ist } \frac{qn - k}{rn - k} \text{ eine } p\text{-ganze Zahl,} \\ \text{und erfüllt } \frac{qn - k}{rn - k} \equiv 1 \pmod{p^\gamma}, \text{ weil } qn - k \equiv rn - k \pmod{p^\gamma} \text{ ist} \\ \text{(denn } qn \text{ und } rn \text{ sind beide durch } p^\gamma \text{ teilbar)} \end{array} \right) \\
&= 1 \pmod{p^\gamma};
\end{aligned}$$

somit wird die Gleichung (36) zu

$$\begin{aligned}
\binom{qn}{rn} &= \frac{\prod_{\substack{k \in \{0,1,\dots, rn-1\}; \\ p|k}} (qn-k)}{\prod_{\substack{k \in \{0,1,\dots, rn-1\}; \\ p|k}} (rn-k)} \cdot \frac{\prod_{\substack{k \in \{0,1,\dots, rn-1\}; \\ p \nmid k}} (qn-k)}{\prod_{\substack{k \in \{0,1,\dots, rn-1\}; \\ p \nmid k}} (rn-k)} \\
&= \underbrace{\binom{qn/p}{rn/p}}_{\equiv 1 \pmod{p^\gamma}} \\
&\equiv \binom{qn/p}{rn/p} \cdot 1 = \binom{qn/p}{rn/p} \pmod{p^\gamma}.
\end{aligned}$$

Doch $\binom{qn/p}{rn/p} p^\gamma$ ist durch $p^{v_p(n)}$ teilbar⁵⁰, und somit folgt hieraus

$$\binom{qn}{rn} \equiv \binom{qn/p}{rn/p} \pmod{p^{v_p(n)}}.$$

Damit ist Lemma 4.4 auch im Fall 1 bewiesen.

⁵⁰*Beweis:* Wir haben $\gamma = \min\{v_p(rn), v_p(qn)\}$. Daher sind zwei Fälle möglich: der Fall $\gamma = v_p(rn)$ und der Fall $\gamma = v_p(qn)$. Im Fall $\gamma = v_p(qn)$ ist offensichtlich $\binom{qn/p}{rn/p} p^\gamma$ durch $p^{v_p(n)}$ teilbar (denn $\gamma = v_p(qn) = \underbrace{v_p(q)}_{\geq 0} + v_p(n) \geq v_p(n)$). Wir betrachten also fortan nur den Fall $\gamma = v_p(rn)$. In diesem Fall wenden wir (35) auf $a = qn/p$ und $b = rn/p$ an, und erhalten $\binom{qn/p-1}{rn/p-1} = \frac{rn/p}{qn/p} \binom{qn/p}{rn/p}$, also

$$\begin{aligned}
\binom{qn/p}{rn/p} &= \frac{qn/p}{rn/p} \binom{qn/p-1}{rn/p-1} = \frac{qn}{rn} \binom{qn/p-1}{rn/p-1}, & \text{damit} \\
rn \binom{qn/p}{rn/p} &= qn \binom{qn/p-1}{rn/p-1}, & \text{also} \\
v_p\left(rn \binom{qn/p}{rn/p}\right) &= v_p\left(qn \binom{qn/p-1}{rn/p-1}\right) = \underbrace{v_p(q)}_{\geq 0} + v_p(n) + \underbrace{v_p\left(\binom{qn/p-1}{rn/p-1}\right)}_{\geq 0} \geq v_p(n),
\end{aligned}$$

was aber wegen

$$v_p\left(rn \binom{qn/p}{rn/p}\right) = \underbrace{v_p(rn)}_{=\gamma=v_p(p^\gamma)} + v_p\left(\binom{qn/p}{rn/p}\right) = v_p(p^\gamma) + v_p\left(\binom{qn/p}{rn/p}\right) = v_p\left(p^\gamma \cdot \binom{qn/p}{rn/p}\right) = v_p\left(\binom{qn/p}{rn/p} p^\gamma\right)$$

zu

$$v_p\left(\binom{qn/p}{rn/p} p^\gamma\right) \geq v_p(n)$$

wird. Dies bedeutet, daß $\binom{qn/p}{rn/p} p^\gamma$ durch $p^{v_p(n)}$ teilbar ist. Somit ist auch im Fall $\gamma = v_p(rn)$

bewiesen, daß $\binom{qn/p}{rn/p} p^\gamma$ durch $p^{v_p(n)}$ teilbar ist, und unser Beweis ist fertig.

Es bleibt nur noch der Fall 2. In diesem Fall ist einerseits $\binom{qn/p}{rn/p} = 0$ (da $rn/p \notin \mathbb{N}$), andererseits

$$\binom{qn}{rn} = \frac{qn}{rn} \binom{qn-1}{rn-1}$$

(denn aus (35), angewandt auf $a = qn$ und $b = rn$, folgt $\binom{qn-1}{rn-1} = \binom{qn}{rn} \frac{rn}{qn}$). Doch da wir im Fall 2 sind, gilt $rn/p \notin \mathbb{N}$ und damit $p \nmid rn$; folglich ist $\frac{q}{rn}$ eine p -ganze Zahl, und somit ist $\frac{qn}{rn} \equiv 0 \pmod{p^{v_p(n)}}$ (dies folgt aus der Kongruenz $n \equiv 0 \pmod{p^{v_p(n)}}$, multipliziert mit der Kongruenz $\frac{q}{rn} \equiv \frac{q}{rn} \pmod{p^{v_p(n)}}$). Wir haben damit

$$\binom{qn}{rn} = \underbrace{\frac{qn}{rn}}_{\equiv 0 \pmod{p^{v_p(n)}}} \underbrace{\binom{qn-1}{rn-1}}_{\text{eine ganze Zahl}} \equiv 0 = \binom{qn/p}{rn/p} \pmod{p^{v_p(n)}},$$

und somit ist Lemma 4.4 auch im Fall 2 bewiesen.

Damit ist der erste Beweis von Lemma 4.4 vollständig. Dieser Beweis war eigentlich sehr naheliegend, zumindest wenn man weiß, daß man mit p -ganzen Zahlen genauso wie mit ganzen Zahlen modulo p^k rechnen darf. Allerdings hat dieser Beweis den Nachteil, daß wir drei Fälle unterscheiden mussten. Der folgende zweite Beweis von Lemma 4.4 hat diesen Nachteil nicht; dafür ist er weniger intuitiv.

*Zweiter Beweis von Lemma 4.4.*⁵¹ Wir nehmen o. B. d. A. an, daß $rn \in \mathbb{N}$ ist (denn im Falle von $rn \notin \mathbb{N}$ ist $\binom{qn/p}{rn/p} = 0$ und $\binom{qn}{rn} = 0$, und somit ist Lemma 4.4 in diesem Falle trivial). Daraus folgt $rn \geq 0$ und somit $r \geq 0$ (da $n > 0$).

Wir setzen $m = qn$. Nach der binomischen Formel (Satz 3.1) (angewandt auf m/p statt n) ist dann

$$(1 + X)^{m/p} = \sum_{k=0}^{\infty} \binom{m/p}{k} X^k \quad (37)$$

(als Gleichheit zwischen zwei ganzzahligen Potenzreihen). Wenn wir in dieser Gleichung überall X durch X^p ersetzen⁵², erhalten wir

$$(1 + X^p)^{m/p} = \sum_{k=0}^{\infty} \binom{m/p}{k} (X^p)^k.$$

⁵¹Dieser zweite Beweis von Lemma 4.4 ist weitgehend identisch mit dem "Proof of Lemma 19" in [3].

⁵²Wir dürfen dies tun, weil es (abstrakt gesehen) auf ein Einsetzen von X^p in die Potenzreihen $(1 + X)^{m/p}$ und $\sum_{k=0}^{\infty} \binom{m/p}{k} X^k$ hinausläuft, und wenn zwei Potenzreihen gleich sind, dann ergeben sie auch nach Einsetzen von X^p zwei gleiche Potenzreihen.

Wir haben also

$$\begin{aligned}
(1 + X^p)^{m/p} &= \sum_{k=0}^{\infty} \underbrace{\binom{m/p}{k}}_{=X^{pk}} \underbrace{(X^p)^k}_{=X^{pk}} = \sum_{k \in \mathbb{N}} \binom{m/p}{pk/p} X^{pk} = \sum_{\lambda \in p\mathbb{N}} \binom{m/p}{\lambda/p} X^\lambda \\
&= \sum_{k \in \mathbb{N}} \binom{m/p}{pk/p} \\
&\quad \text{(hier haben wir } pk \text{ durch } \lambda \text{ substituiert)} \\
&= \sum_{\lambda \in \mathbb{N}} \binom{m/p}{\lambda/p} X^\lambda - \sum_{\lambda \in \mathbb{N} \setminus p\mathbb{N}} \underbrace{\binom{m/p}{\lambda/p}}_{=0, \text{ denn aus } \lambda \notin p\mathbb{N} \text{ folgt } \lambda/p \notin \mathbb{N}} X^\lambda = \sum_{\lambda \in \mathbb{N}} \binom{m/p}{\lambda/p} X^\lambda - \underbrace{\sum_{\lambda \in \mathbb{N} \setminus p\mathbb{N}} 0 X^\lambda}_{=0} = \sum_{\lambda \in \mathbb{N}} \binom{m/p}{\lambda/p} X^\lambda \\
&= \sum_{k \in \mathbb{N}} \binom{m/p}{k/p} X^k = \sum_{k=0}^{\infty} \binom{m/p}{k/p} X^k. \tag{38}
\end{aligned}$$

Andererseits ist $1 + X^p \equiv (1 + X)^p \pmod p$ ⁵³, weil

$$\begin{aligned}
(1 + X)^p &= \sum_{k=0}^{\infty} \binom{p}{k} X^k \quad \text{(nach Satz 3.1, angewandt auf } p \text{ statt } n) \\
&= \underbrace{\binom{p}{0}}_{=1} \underbrace{X^0}_{=1} + \sum_{k=1}^{p-1} \underbrace{\binom{p}{k}}_{\substack{=0 \pmod p, \\ \text{weil bekanntlich} \\ p \mid \binom{p}{k} \text{ ist}}} X^k + \underbrace{\binom{p}{p}}_{=1} X^p + \sum_{k=p+1}^{\infty} \underbrace{\binom{p}{k}}_{=0 \text{ (da } k > p)} X^k \\
&\equiv 1 + \underbrace{\sum_{k=1}^{p-1} 0 X^k}_{=0} + X^p + \underbrace{\sum_{k=p+1}^{\infty} 0 X^k}_{=0} = 1 + X^p \pmod p
\end{aligned}$$

ist.

Nun wollen wir hieraus folgern, daß $(1 + X^p)^{m/p} \equiv ((1 + X)^p)^{m/p} \pmod{p^{v_p(m)}}$ ist. Hierzu benötigen wir eine Verallgemeinerung von Lemma 2.3 auf Potenzreihen⁵⁴:

Lemma 4.6: Seien $p \in \mathbb{Z}$, und seien u und v zwei ganzzahlige Potenzreihen. Seien $k \in \mathbb{N}_+$ und $\ell \in \mathbb{N}$. Wenn $u \equiv v \pmod{p^k}$ ist⁵⁵, dann ist $u^{p^\ell} \equiv$

⁵³Hierbei bedeutet die Kongruenz $P \equiv Q \pmod p$ für zwei ganzzahlige Potenzreihen P und Q , daß *jeder Koeffizient* der Potenzreihe $P - Q$ durch p teilbar ist. Mit anderen Worten: Die Kongruenz $P \equiv Q \pmod p$ bedeutet, daß für jedes $i \in \mathbb{N}$ die Kongruenz

$$(\text{der } i\text{-te Koeffizient der Potenzreihe } P) \equiv (\text{der } i\text{-te Koeffizient der Potenzreihe } Q) \pmod p$$

gilt.

⁵⁴Übrigens ist Lemma 4.6, genauso wie Lemma 2.3, ein Sonderfall von Lemma 3 in [3].

⁵⁵Hierbei bedeutet die Kongruenz $P \equiv Q \pmod{p^k}$ für zwei ganzzahlige Potenzreihen P und Q , daß *jeder Koeffizient* der Potenzreihe $P - Q$ durch p^k teilbar ist. Mit anderen Worten: Die Kongruenz $P \equiv Q \pmod{p^k}$ bedeutet, daß für jedes $i \in \mathbb{N}$ die Kongruenz

$$(\text{der } i\text{-te Koeffizient der Potenzreihe } P) \equiv (\text{der } i\text{-te Koeffizient der Potenzreihe } Q) \pmod{p^k}$$

gilt.

$$v^{p^\ell} \bmod p^{k+\ell}.$$

Beweis von Lemma 4.6: Der Beweis von Lemma 4.6 verlauft vollig analog zum Beweis von Lemma 2.3 (wobei wir diesmal die Tatsache benotigen, da das Produkt einer Potenzreihe, deren Koeffizienten alle durch $p^{k+\ell}$ teilbar sind, mit einer Potenzreihe, deren Koeffizienten alle durch p teilbar sind, selber eine Potenzreihe ist, deren Koeffizienten alle durch $p^{k+\ell} \cdot p$ teilbar sind; aber dies folgt trivial aus der Definition des Produktes zweier Potenzreihen).

Nun zuruck zum Beweis von Lemma 4.4: Wir setzen $\ell = v_p(m/p)$. Hierbei ist m/p eine ganze Zahl, da $m/p = \underbrace{q}_{\in \mathbb{Z}} \underbrace{n/p}_{\in \mathbb{Z}} \in \mathbb{Z}$.

Ferner setzen wir $u = 1+X^p$ und $v = (1+X)^p$. Damit wird $1+X^p \equiv (1+X)^p \bmod p$ zu $u \equiv v \bmod p^1$. Wenden wir Lemma 4.6 auf $k = 1$ an, dann erhalten wir also $u^{p^\ell} \equiv v^{p^\ell} \bmod p^{1+\ell}$. Wegen $v_p(m) = v_p(p \cdot (m/p)) = \underbrace{v_p(p)}_{=1} + \underbrace{v_p(m/p)}_{=\ell} = 1 + \ell$ wird

dies zu $u^{p^\ell} \equiv v^{p^\ell} \bmod p^{v_p(m)}$.

Nun ist $m/p = p^\ell \cdot \kappa$ fur eine ganze Zahl κ (denn $\ell = v_p(m/p)$ ergibt $p^\ell \mid (m/p)$). Also ist

$$\begin{aligned} u^{m/p} &= u^{p^\ell \cdot \kappa} = (u^{p^\ell})^\kappa \equiv (v^{p^\ell})^\kappa && \left(\text{da } u^{p^\ell} \equiv v^{p^\ell} \bmod p^{v_p(m)} \right) \\ &= v^{p^\ell \cdot \kappa} = v^{m/p} \bmod p^{v_p(m)} && \left(\text{da } p^\ell \cdot \kappa = m/p \right). \end{aligned}$$

Wegen $u = 1+X^p$ und $v = (1+X)^p$ wird dies zu $(1+X^p)^{m/p} \equiv ((1+X)^p)^{m/p} \bmod p^{v_p(m)}$. Wir haben somit

$$(1+X^p)^{m/p} \equiv ((1+X)^p)^{m/p} = (1+X)^m = \sum_{k=0}^{\infty} \binom{m}{k} X^k \bmod p^{v_p(m)}$$

(nach Satz 3.1, angewandt auf m statt n). Aus (38) folgt somit

$$\sum_{k=0}^{\infty} \binom{m}{k} X^k \equiv \sum_{k=0}^{\infty} \binom{m/p}{k/p} X^k \bmod p^{v_p(m)}.$$

Da bei zwei Potenzreihen, die kongruent zueinander modulo $p^{v_p(m)}$ sind, auch die entsprechenden Koeffizienten jeweils kongruent zueinander modulo $p^{v_p(m)}$ sein mussen, erhalten wir hieraus

$$\binom{m}{k} \equiv \binom{m/p}{k/p} \bmod p^{v_p(m)} \quad \text{fur alle } k \in \mathbb{N}.$$

Setzen wir hier $k = rn$ ein und erinnern wir uns daran, da $m = qn$ ist, dann wird dies zu

$$\binom{qn}{rn} \equiv \binom{qn/p}{rn/p} \bmod p^{v_p(qn)}. \quad (39)$$

Nun ist aber $v_p(qn) = \underbrace{v_p(q)}_{\geq 0} + v_p(n) \geq v_p(n)$ und somit $p^{v_p(n)} \mid p^{v_p(qn)}$. Aus dieser

Kongruenz folgt also

$$\binom{qn}{rn} \equiv \binom{qn/p}{rn/p} \bmod p^{v_p(n)}.$$

Somit ist Lemma 4.4 erneut bewiesen.

Es ist nicht schwer, Lemma 4.5 auf ähnliche Weise zu zeigen:

Erster Beweis von Lemma 4.5: Der folgende Beweis von Lemma 4.5 ist eine Variation von unserem ersten Beweis von Lemma 4.4.

Wir nehmen o. B. d. A. an, daß $r > 0$ ist. Diese Annahme ist in der Tat legitimiert, denn Lemma 4.5 ist trivial im Falle von $r \leq 0$ (denn im Falle von $r \leq 0$ ist $\binom{qn/p-1}{rn/p-1} = 0$ (da $\underbrace{rn/p}_{\leq 0} - 1 \leq -1 < 0$) und $\binom{qn-1}{rn-1} = 0$ (da $\underbrace{rn}_{\leq 0} - 1 \leq -1 < 0$)).

Aus $r \in \mathbb{Z}$ und $r > 0$ folgt $r \in \mathbb{N}$, also $\underbrace{r}_{\in \mathbb{N}} \underbrace{n}_{\in \mathbb{N}} \in \mathbb{N}$ und $\underbrace{r}_{\in \mathbb{N}} \underbrace{n/p}_{\in \mathbb{N}} \in \mathbb{N}$. Aus $rn \in \mathbb{N}$ und $rn > 0$ (da $r > 0$ und $n > 0$) folgt $rn - 1 \in \mathbb{N}$, und somit können wir den Binomialkoeffizienten $\binom{qn-1}{rn-1}$ als

$$\binom{qn-1}{rn-1} = \frac{(qn-1)(qn-2)\dots(qn-rn+1)}{(rn-1)!}$$

darstellen. Wegen

$$(qn-1)(qn-2)\dots(qn-rn+1) = \prod_{k \in \{1,2,\dots,rn-1\}} (qn-k) = \prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p|k}} (qn-k) \cdot \prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p \nmid k}} (qn-k)$$

und

$$(rn-1)! = \prod_{k \in \{1,2,\dots,rn-1\}} (rn-k) = \prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p|k}} (rn-k) \cdot \prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p \nmid k}} (rn-k)$$

wird dies zu

$$\begin{aligned} \binom{qn-1}{rn-1} &= \frac{\prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p|k}} (qn-k) \cdot \prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p \nmid k}} (qn-k)}{\prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p|k}} (rn-k) \cdot \prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p \nmid k}} (rn-k)} \\ &= \frac{\prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p|k}} (qn-k)}{\prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p|k}} (rn-k)} \cdot \frac{\prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p \nmid k}} (qn-k)}{\prod_{\substack{k \in \{1,2,\dots,rn-1\}; \\ p \nmid k}} (rn-k)}. \end{aligned} \quad (40)$$

Der erste Bruch auf der rechten Seite dieser Gleichung vereinfacht sich zu

$$\frac{\prod_{k \in \{1, 2, \dots, rn-1\};} (qn - k)}{p \nmid k} = \frac{\prod_{\ell \in \{1, 2, \dots, rn/p-1\}} (qn - \ell p)}{\prod_{\ell \in \{1, 2, \dots, rn/p-1\}} (rn - \ell p)}$$

(hier haben wir k durch ℓp substituiert, da $p \mid k$ vorausgesetzt wurde)

$$= \frac{\prod_{\ell \in \{1, 2, \dots, rn/p-1\}} ((qn/p - \ell) \cdot p)}{\prod_{\ell \in \{1, 2, \dots, rn/p-1\}} ((rn/p - \ell) \cdot p)} = \frac{\prod_{\ell \in \{1, 2, \dots, rn/p-1\}} (qn/p - \ell)}{\prod_{\ell \in \{1, 2, \dots, rn/p-1\}} (rn/p - \ell)}$$

(hier haben wir aus Zähler und Nenner jeweils $rn/p - 1$ mal den Faktor p herausgekürzt)

$$= \frac{(qn/p - 1)(qn/p - 2) \dots (qn/p - rn/p + 1)}{(rn/p - 1)!} = \binom{qn/p - 1}{rn/p - 1} \quad (\text{da } rn/p - 1 \in \mathbb{N}),$$

und der zweite Bruch auf der rechten Seite erfüllt

$$\frac{\prod_{k \in \{1, 2, \dots, rn-1\};} (qn - k)}{p \nmid k} = \prod_{k \in \{1, 2, \dots, rn-1\};} \frac{qn - k}{rn - k} \equiv \prod_{k \in \{1, 2, \dots, rn-1\};} 1$$

\left(\begin{array}{l} \text{denn für jedes } k \in \{1, 2, \dots, rn - 1\} \text{ mit } p \nmid k \text{ ist } \frac{qn - k}{rn - k} \text{ eine } p\text{-ganze Zahl,} \\ \text{und erfüllt } \frac{qn - k}{rn - k} \equiv 1 \pmod{p^{v_p(n)}}, \text{ weil } qn - k \equiv rn - k \pmod{p^{v_p(n)}} \text{ ist} \\ \text{(denn } n \text{ ist durch } p^{v_p(n)} \text{ teilbar, und wir haben } q \in \mathbb{Z} \text{ und } r \in \mathbb{Z}) \end{array} \right)

$$= 1 \pmod{p^{v_p(n)}};$$

somit wird die Gleichung (40) zu

$$\begin{aligned} \binom{qn - 1}{rn - 1} &= \underbrace{\frac{\prod_{k \in \{1, 2, \dots, rn-1\};} (qn - k)}{p \nmid k}}_{= \binom{qn/p - 1}{rn/p - 1}} \cdot \underbrace{\frac{\prod_{k \in \{1, 2, \dots, rn-1\};} (qn - k)}{\prod_{k \in \{1, 2, \dots, rn-1\};} (rn - k)}}_{\equiv 1 \pmod{p^{v_p(n)}}} \\ &\equiv \binom{qn/p - 1}{rn/p - 1} \cdot 1 = \binom{qn/p - 1}{rn/p - 1} \pmod{p^{v_p(n)}}. \end{aligned}$$

Damit ist der Beweis von Lemma 4.5 abgeschlossen.

*Zweiter Beweis von Lemma 4.5:*⁵⁶ Wie im zweiten Beweis von Lemma 4.4 gezeigt wurde, gilt (39). Damit ist $p^{v_p(qn)} \mid \binom{qn}{rn} - \binom{qn/p}{rn/p}$, also

$$v_p \left(\binom{qn}{rn} - \binom{qn/p}{rn/p} \right) \geq v_p(qn) = v_p(q) + v_p(n). \quad (41)$$

⁵⁶Der folgende Zweite Beweis von Lemma 4.5 richtet sich im Wesentlichen nach dem Beweis von Lemma 21 in [3].

Nun gilt

$$\begin{aligned}
 & \left(\begin{array}{c} \binom{qn-1}{rn-1} - \binom{qn/p-1}{rn/p-1} \\ = \binom{qn}{rn} \frac{rn}{qn} \text{ (nach (35), angewandt auf } a=qn \text{ und } b=rn) \\ = \binom{qn/p}{rn/p} \frac{rn/p}{qn/p} \text{ (nach (35), angewandt auf } a=qn/p \text{ und } b=rn/p) \end{array} \right) q = \left(\binom{qn}{rn} \frac{rn}{qn} - \binom{qn/p}{rn/p} \frac{rn/p}{qn/p} \right) q \\
 & = \binom{qn}{rn} \underbrace{\frac{rn}{qn}}_{=r} q - \binom{qn/p}{rn/p} \underbrace{\frac{rn/p}{qn/p}}_{=r} q = \binom{qn}{rn} r - \binom{qn/p}{rn/p} r = \left(\binom{qn}{rn} - \binom{qn/p}{rn/p} \right) r.
 \end{aligned}$$

Folglich ist

$$\begin{aligned}
 v_p \left(\left(\binom{qn-1}{rn-1} - \binom{qn/p-1}{rn/p-1} \right) q \right) &= v_p \left(\left(\binom{qn}{rn} - \binom{qn/p}{rn/p} \right) r \right) \\
 &= v_p \left(\binom{qn}{rn} - \binom{qn/p}{rn/p} \right) + v_p(r) \\
 &\quad \geq v_p(q) + v_p(n) \text{ (nach (41))} \\
 &\geq v_p(q) + v_p(n) + \underbrace{v_p(r)}_{\geq 0} \geq v_p(q) + v_p(n).
 \end{aligned}$$

Wegen

$$v_p \left(\left(\binom{qn-1}{rn-1} - \binom{qn/p-1}{rn/p-1} \right) q \right) = v_p \left(\binom{qn-1}{rn-1} - \binom{qn/p-1}{rn/p-1} \right) + v_p(q)$$

wird dies zu

$$v_p \left(\binom{qn-1}{rn-1} - \binom{qn/p-1}{rn/p-1} \right) + v_p(q) \geq v_p(q) + v_p(n),$$

was sich zu

$$v_p \left(\binom{qn-1}{rn-1} - \binom{qn/p-1}{rn/p-1} \right) \geq v_p(n)$$

vereinfacht. Doch dies bedeutet einfach, daß $p^{v_p(n)} \mid \binom{qn-1}{rn-1} - \binom{qn/p-1}{rn/p-1}$ gilt; das heißt, $\binom{qn/p-1}{rn/p-1} \equiv \binom{qn-1}{rn-1} \pmod{p^{v_p(n)}}$. Damit ist Lemma 4.5 erneut bewiesen.

4.3. Beweis von Satz 4.1 und Satz 4.3

Nun können wir zum Beweis der Sätze 4.1 und 4.3 schreiten:

Beweis von Satz 4.1: Die Folge (b_1, b_2, b_3, \dots) , welche durch $b_n = \binom{qn}{rn}$ für alle $n \in \mathbb{N}_+$ definiert ist, erfüllt die Aussage \mathcal{A} von Satz 2.1 (denn die Aussage \mathcal{A} für diese Folge ist

nichts anderes als die Aussage von Lemma 4.4, und wir wissen, daß Lemma 4.4 wahr ist). Da die Aussagen \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} und \mathcal{H} äquivalent sind (laut Satz 2.1), erfüllt diese Folge somit alle Aussagen \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} und \mathcal{H} von Satz 2.1. Dadurch ist Satz 4.1 bewiesen.

Der *Beweis von Satz 4.3* ist völlig analog zu dem gerade gegebenen Satz von 4.1, mit dem einzigen Unterschied, daß wir jetzt Lemma 4.5 statt Lemma 4.4 anwenden.

4.4. Ein Zusatz zum Äquivalenzsatz 2.1

An dieser Stelle wollen wir eigentlich mit der Untersuchung der Folgen von Satz 4.1 und 4.3 aufhören, aber nicht ohne eine Erweiterung von Satz 2.1 zu konstatieren:

Satz 4.7: Sei $(b_1, b_2, b_3, \dots) \in \mathbb{Z}^{\mathbb{N}^+}$ eine Folge ganzer Zahlen. Dann sind die Aussagen \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} , \mathcal{H} , \mathcal{I} und \mathcal{J} äquivalent, wobei \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} und \mathcal{H} die in Satz 2.1 festgelegten Aussagen sind, und \mathcal{I} und \mathcal{J} die folgenden zwei Aussagen sind:

Aussage \mathcal{I} : Es gibt eine Folge $(q_1, q_2, q_3, \dots) \in \mathbb{Z}^{\mathbb{N}^+}$ ganzer Zahlen, die

$$\left(b_n = \sum_{d|n} d \binom{q_d n / d}{n / d} \text{ für jedes } n \in \mathbb{N}_+ \right) \quad (42)$$

erfüllt.

Aussage \mathcal{J} : Es gibt *genau eine* Folge $(q_1, q_2, q_3, \dots) \in \mathbb{Z}^{\mathbb{N}^+}$ ganzer Zahlen, die (42) erfüllt.

Wir wollen den *Beweis von Satz 4.7* nicht ausführen, aber er sollte dem Leser nicht schwerfallen. Er muss nur $\mathcal{A} \iff \mathcal{I}$ beweisen (analog zu unserem Beweis von $\mathcal{A} \iff \mathcal{B}$ weiter oben) und $\mathcal{I} \iff \mathcal{J}$ beweisen (analog zu unserem Beweis von $\mathcal{B} \iff \mathcal{C}$ weiter oben). Lemma 4.4 kommt zur Hilfe.

Ein Analogon von Satz 4.7 mit $\binom{q_d n / d - 1}{n / d - 1}$ statt $\binom{q_d n / d}{n / d}$ funktioniert übrigens nicht.

5. Kombinatorik I: Das Pólya-Burnsidesche Abzählverfahren

5.1. Perlenketten mit n Perlen in q Farben

Nach all der Algebra und Zahlentheorie, die wir nun betrieben haben, wird es Zeit für Kombinatorik. Aus dieser stammt nämlich ein ganz anderer Ansatz zum Beweis von Satz 1.1, nämlich folgender:

Betrachten wir erstmal den Fall $q \geq 0$. Wir haben eine kreisförmige Kette aus n Perlen, und wollen jede Perle mit einer von q Farben färben⁵⁷. Wieviele solche Färbungen gibt es? Die Antwort hängt natürlich davon ab, wie man die Frage versteht. Wenn man eine Färbung einfach als Abbildung aus der Menge der Perlen in die

⁵⁷Wobei nicht notwendigerweise jede der q Farben auch in jeder Färbung vorkommen muss!

Menge der Farben ansieht, gibt es q^n Färbungen (denn für jede der n Perlen sind q Farben möglich). Doch diese Interpretation des Begriffes einer Färbung ist weder interessant, noch realistisch. Uns interessiert folgende Interpretation: Wieviele verschiedene solche Färbungen gibt es, wenn man zwei Färbungen, die eine zyklische Rotation der Perlenkette (in der Ebene) ineinander überführt, als gleich ansieht?⁵⁸ So sollen z. B. die Färbungen "Perle 1 schwarz und Perlen 2, 3, ..., n weiß" und "Perle 2 schwarz und Perlen 3, 4, ..., n , 1 weiß" (wenn wir die Perlen mit 1, 2, ..., n durchnummerieren) als gleich angesehen werden.

Wir werden nun zeigen, daß es insgesamt $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d}$ solche Färbungen gibt.

Natürlich wird daraus folgen, daß $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} \in \mathbb{N}$ ist, und somit wird Satz 1.1 für $q \geq 0$ bewiesen sein. Zum Fall $q < 0$ werden wir später (in Abschnitt 6) noch zurückkommen.

Wir wollen zuerst den Begriff einer Perlenkette formalisieren⁵⁹. Wir nummerieren die Perlen mit den Elementen $\bar{0}, \bar{1}, \dots, \overline{n-1}$ von $\mathbb{Z}/n\mathbb{Z}$ ⁶⁰ durch (wobei wir sie gegen den Uhrzeigersinn durchgehen), und wir nummerieren die Farben mit den ganzen Zahlen $1, 2, \dots, q$ durch. Eine *strikte Färbung* unserer Perlenkette mit unseren q Farben werde definiert als eine Abbildung von $\mathbb{Z}/n\mathbb{Z}$ nach $\{1, 2, \dots, q\}$ (die jeder Perle ihre Farbe zuordnet). Wie gesagt, gibt es q^n strikte Färbungen, aber dies ist nicht besonders interessant, denn eigentlich wollen wir zwei Färbungen, welche eine zyklische Rotation der Perlenkette ineinander überführt, miteinander gleichsetzen.

Wir können aber in der Mathematik nicht einfach zwei verschiedene Abbildungen von $\mathbb{Z}/n\mathbb{Z}$ nach $\{1, 2, \dots, q\}$ gleichsetzen. Stattdessen müssen wir, um formal korrekt zu arbeiten, nicht *Abbildungen*, sondern *Äquivalenzklassen von Abbildungen* betrachten, und zwar modulo der Äquivalenzrelation, die dadurch definiert ist, daß wir zwei Abbildungen $f : \mathbb{Z}/n\mathbb{Z} \rightarrow \{1, 2, \dots, q\}$ und $g : \mathbb{Z}/n\mathbb{Z} \rightarrow \{1, 2, \dots, q\}$ genau dann äquivalent nennen, wenn es ein $a \in \mathbb{Z}/n\mathbb{Z}$ gibt, welches

$$(f(x+a) = g(x) \quad \text{für alle } x \in \mathbb{Z}/n\mathbb{Z}) \quad (43)$$

erfüllt. (Denn diese Bedingung (43) sagt genau aus, daß die strikte Färbung g aus der strikten Färbung f durch Rotation um den Winkel $\frac{a}{n} \cdot 360^\circ$ hervorgeht.) Wir definieren also eine *Färbung* als eine Äquivalenzklasse von strikten Färbungen modulo der gerade definierten Äquivalenzrelation.⁶¹ Die formale Formulierung unserer Behauptung ist

⁵⁸Aber zwei Färbungen, die eine Geradenspiegelung (bzw. eine Drehung im Raum) ineinander überführt, wollen wir hier *nicht* als gleich ansehen. Wir werden später kurz darauf zurückkommen, was passieren würde, würden wir es tun.

⁵⁹*Bemerkung:* Der englische Begriff für "Färbung einer Perlenkette" in diesem Kontext heißt "necklace". Genauer gesagt heißt das, was wir hier als eine Färbung einer Perlenkette mit n Perlen und (potentiell) q Farben bezeichnen, im Englischen eine " q -ary necklace of length n ". Mit diesem Suchbegriff wird der Leser eine Masse an Literatur zu diesem Thema finden.

⁶⁰Mit $\mathbb{Z}/n\mathbb{Z}$ bezeichnen wir die Menge der Restklassen der ganzen Zahlen modulo n . (Viele Autoren schreiben gerne \mathbb{Z}/n oder $\mathbb{Z}/(n)$ oder \mathbb{Z}_n statt $\mathbb{Z}/n\mathbb{Z}$, aber es ist zu beachten, daß die Bezeichnung \mathbb{Z}_n zuweilen auch etwas anderes bedeutet, weshalb wir sie hier vermeiden.) Die Restklassen der Zahlen $0, 1, \dots, n-1$ modulo n heißen $\bar{0}, \bar{1}, \dots, \overline{n-1}$. Allgemein bezeichnen wir mit \bar{a} die Restklasse einer ganzen Zahl a modulo n .

⁶¹Auch wenn das Wort "Färbung" Teil des Begriffes "strikte Färbung" ist, sind Färbungen natürlich keine strikten Färbungen, sondern Äquivalenzklassen von strikten Färbungen.

also:

Satz 5.1 (MacMahons Perlenkettensatz): Seien $q \in \mathbb{N}$ und $n \in \mathbb{N}_+$. Wir definieren eine Relation $\overset{\text{neck}}{\sim}$ auf der Menge $\{1, 2, \dots, q\}^{\mathbb{Z}/n\mathbb{Z}}$ (dies ist die Menge aller Abbildungen von $\mathbb{Z}/n\mathbb{Z}$ nach $\{1, 2, \dots, q\}$) folgendermaßen: Für zwei Abbildungen $f : \mathbb{Z}/n\mathbb{Z} \rightarrow \{1, 2, \dots, q\}$ und $g : \mathbb{Z}/n\mathbb{Z} \rightarrow \{1, 2, \dots, q\}$ sei genau dann $f \overset{\text{neck}}{\sim} g$, wenn es ein $a \in \mathbb{Z}/n\mathbb{Z}$ gibt, welches

$$(f(x+a) = g(x) \quad \text{für alle } x \in \mathbb{Z}/n\mathbb{Z}) \quad (44)$$

erfüllt.

(a) Diese Relation $\overset{\text{neck}}{\sim}$ ist eine Äquivalenzrelation auf der Menge $\{1, 2, \dots, q\}^{\mathbb{Z}/n\mathbb{Z}}$.

(b) Die Menge $\{1, 2, \dots, q\}^{\mathbb{Z}/n\mathbb{Z}}$ zerfällt in genau $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d}$ Äquivalenzklassen bezüglich der Relation $\overset{\text{neck}}{\sim}$.

Dieser Satz 5.1 kommt z. B. in Abschnitt 4.9 von [13] vor. Wir werden ihn nun beweisen, und dabei einige auch anderweitig hilfreiche Lemmata aufsammeln.

5.2. Anzahlen von Äquivalenzklassen: ein fast triviales Lemma

Satz 5.1 (a) ist einfach einzusehen; die Hauptschwierigkeit am Beweis von Satz 5.1 besteht darin, Teil (b) zu beweisen. Dazu könnte man sich erst einmal folgende ganz allgemeine Frage stellen: Wieviele Äquivalenzklassen gibt es bezüglich einer beliebigen Äquivalenzrelation auf einer beliebigen Menge? So allgemein, wie diese Frage gestellt ist, kann man keine wirklich inhaltsreiche Antwort auf sie geben, aber zumindest kann man diese Anzahl auch durch eine Summe ausdrücken:

Satz 5.2: Sei F eine endliche Menge, und \sim eine Äquivalenzrelation auf dieser Menge F . Dann ist

$$|F/\sim| = \sum_{f \in F} \frac{1}{|\text{Äquivalenzklasse von } f|}.$$

Bemerkungen: Hierbei bezeichnet F/\sim die Menge aller Äquivalenzklassen bezüglich der Relation \sim auf der Menge F . Ferner bezeichnen wir für jede endliche Menge X die Mächtigkeit der Menge X mit $|X|$. (Somit bezeichnet $|F/\sim|$ die Anzahl aller Äquivalenzklassen bezüglich der Relation \sim auf der Menge F .)

Beweis von Satz 5.2: Die Menge F ist die Vereinigung aller Äquivalenzklassen bezüglich der Relation \sim auf der Menge F , und all diese Äquivalenzklassen sind paar-

weise disjunkt. Hieraus folgt

$$\begin{aligned}
 \sum_{f \in F} \frac{1}{|\text{Äquivalenzklasse von } f|} &= \sum_{\substack{K \text{ Äquivalenzklasse} \\ \text{bezüglich der Relation } \sim \\ \text{auf der Menge } F}} \sum_{f \in K} \frac{1}{|\text{Äquivalenzklasse von } f|} \\
 &= \sum_{K \in F/\sim} \underbrace{1}_{\substack{\text{(nach der} \\ \text{Definition von } F/\sim)}} \frac{1}{|K|} \quad \left(\text{denn die Äquivalenzklasse} \right. \\
 &\quad \left. \text{von } f \text{ ist } K, \text{ denn } f \in K \right) \\
 &= \sum_{K \in F/\sim} \underbrace{\sum_{f \in K} \frac{1}{|K|}}_{\substack{= |K| \cdot \frac{1}{|K|} = 1}} = \sum_{K \in F/\sim} 1 = |F/\sim| \cdot 1 = |F/\sim|.
 \end{aligned}$$

Damit ist Satz 5.2 bewiesen.

5.3. Abbildungsgruppen und induzierte Äquivalenzrelationen: das Nicht-Burnside-Lemma

Satz 5.2 ist allgemein und (wie der Beweis gezeigt hat) trivial - aber trotzdem nicht zu unterschätzen, wenn es darum geht Anzahlen von irgendwelchen Objekten bis auf eine bestimmte Äquivalenzrelation (also formal gesprochen, Anzahlen von Äquivalenzklassen) zu berechnen. Wir haben es aber im Falle von Satz 5.1 nicht mit irgendeiner willkürlichen Äquivalenzrelation zu tun, sondern mit einer, die von einer Abbildungsgruppe induziert ist. Was dies bedeutet, müssen wir erst einmal erklären:

Definition (Abbildungsgruppe): Sei F eine beliebige Menge. Wir bezeichnen mit S_F die Menge aller bijektiven Abbildungen von F nach F . (Diese Menge S_F wird öfters auch mit $\text{End}_{\text{Set}} F$ oder $\text{Bij } F$ bezeichnet.) Eine Teilmenge H dieser Menge S_F heißt eine *Abbildungsgruppe* von F , wenn sie folgende drei Eigenschaften erfüllt:

- 1) Die Identitätsabbildung $\text{id} : F \rightarrow F$ liegt in H .
- 2) Für je zwei Abbildungen $\alpha \in H$ und $\beta \in H$ liegt auch die Abbildung $\alpha \circ \beta$ in H .
- 3) Für jede Abbildung $\alpha \in H$ liegt auch ihre Umkehrabbildung α^{-1} in H .

Beispiele für Abbildungsgruppen sind leicht zu finden: So ist die gesamte Menge S_F eine Abbildungsgruppe von F . Ferner ist die Menge $\{\text{id}\}$, die nur aus der Identitätsabbildung besteht, auch eine Abbildungsgruppe von F .

Ein anderes Beispiel für Abbildungsgruppen haben wir in der Situation von Satz 5.1 vor uns. Dort können wir nämlich $F = \{1, 2, \dots, q\}^{\mathbb{Z}/n\mathbb{Z}}$ setzen, und für jedes $a \in \mathbb{Z}/n\mathbb{Z}$ eine Abbildung $D_a : F \rightarrow F$ definieren, die jeder Funktion $f : \mathbb{Z}/n\mathbb{Z} \rightarrow \{1, 2, \dots, q\}$ eine neue Funktion $D_a(f) : \mathbb{Z}/n\mathbb{Z} \rightarrow \{1, 2, \dots, q\}$ zuordnet, welche durch

$$((D_a(f))(x) = f(x + a) \quad \text{für alle } x \in \mathbb{Z}/n\mathbb{Z}) \quad (45)$$

definiert ist. Anschaulich gesprochen entspricht die Abbildung D_a für jedes $a \in \mathbb{Z}$ der Rotation der Perlenkette um den Winkel $\frac{a}{n} \cdot 360^\circ$, bzw. der Wirkung, die diese

Drehung auf Färbungen⁶² der Perlenkette ausübt (also: jede Farbe wird um a Perlen im Uhrzeigersinn verschoben).

Sei nun H die Teilmenge $\{D_{\bar{0}}, D_{\bar{1}}, \dots, D_{\overline{n-1}}\} = \{D_a \mid a \in \mathbb{Z}/n\mathbb{Z}\}$ von S_F . Dann ist H eine Abbildungsgruppe von F , denn sie erfüllt die drei Eigenschaften **1)**, **2)** und **3)** aus der Definition einer Abbildungsgruppe:

1) Die Identitätsabbildung $\text{id} : F \rightarrow F$ liegt in H (denn $D_{\bar{0}} = \text{id}$ ⁶³).

2) Für je zwei Abbildungen $\alpha \in H$ und $\beta \in H$ liegt auch die Abbildung $\alpha \circ \beta$ in H . (Denn wegen $\alpha \in H$ ist $\alpha = D_a$ für ein $a \in \mathbb{Z}/n\mathbb{Z}$, und wegen $\beta \in H$ ist $\beta = D_b$ für ein $b \in \mathbb{Z}/n\mathbb{Z}$, und daraus folgt $\alpha \circ \beta = D_a \circ D_b = D_{a+b} \in H$. Hierbei haben wir die Gleichheit

$$D_a \circ D_b = D_{a+b} \quad \text{für alle } a \in \mathbb{Z}/n\mathbb{Z} \text{ und } b \in \mathbb{Z}/n\mathbb{Z} \quad (46)$$

benutzt, welche aber leicht zu beweisen ist⁶⁴.)

3) Für jede Abbildung $\alpha \in H$ liegt auch ihre Umkehrabbildung α^{-1} in H . (Denn wegen $\alpha \in H$ ist $\alpha = D_a$ für ein $a \in \mathbb{Z}/n\mathbb{Z}$, und daraus folgt $\alpha^{-1} = D_a^{-1} = D_{-a} \in H$. Hierbei haben wir die Gleichheit

$$D_a^{-1} = D_{-a} \quad \text{für alle } a \in \mathbb{Z}/n\mathbb{Z} \quad (47)$$

benutzt, welche aber leicht zu beweisen ist⁶⁵.)

Alle drei Eigenschaften **1)**, **2)** und **3)**, die eine Teilmenge H von S_F benötigt, um eine Abbildungsgruppe von F zu sein, sind somit nachgeprüft. Also ist H eine Abbildungsgruppe von F .

Nun wollen wieder in den allgemeinen Fall zurückgehen, wo eine beliebige Menge F (nicht notwendigerweise $\{1, 2, \dots, q\}^{\mathbb{Z}/n\mathbb{Z}}$) und eine beliebige Abbildungsgruppe H von F (nicht notwendigerweise $\{D_{\bar{0}}, D_{\bar{1}}, \dots, D_{\overline{n-1}}\}$) gegeben sind:

Definition (induzierte Äquivalenzrelation $\overset{H}{\sim}$): Sei F eine beliebige Menge, und H eine Abbildungsgruppe von F . Wir definieren eine Relation $\overset{H}{\sim}$ auf der Menge F folgendermaßen: Für zwei Elemente $f \in F$ und $g \in F$ sei genau dann $f \overset{H}{\sim} g$, wenn es ein $h \in H$ mit $h(f) = g$ gibt.

⁶²Strenggenommen meinen wir hier nicht wirklich Färbungen, sondern Abbildungen $\mathbb{Z}/n\mathbb{Z} \rightarrow \{1, 2, \dots, q\}$ (denn eine Färbung ist ja keine solche Abbildung, sondern eine Äquivalenzklasse von solchen Abbildungen).

⁶³Denn für jede Funktion $f : \mathbb{Z}/n\mathbb{Z} \rightarrow \{1, 2, \dots, q\}$ ist $D_{\bar{0}}(f) = f$, weil

$$(D_{\bar{0}}(f))(x) = f(x + \bar{0}) = f(x)$$

für alle $x \in \mathbb{Z}/n\mathbb{Z}$ gilt.

⁶⁴In der Tat ist $(D_a \circ D_b)(f) = (D_{a+b})(f)$ für jede Funktion $f : \mathbb{Z}/n\mathbb{Z} \rightarrow \{1, 2, \dots, q\}$, denn für alle $x \in \mathbb{Z}/n\mathbb{Z}$ ist

$$\begin{aligned} ((D_a \circ D_b)(f))(x) &= (D_a(D_b(f)))(x) = (D_b(f))(x + a) = f((x + a) + b) \\ &= f(x + (a + b)) = (D_{a+b}(f))(x). \end{aligned}$$

⁶⁵In der Tat haben wir weiter oben gezeigt, daß $D_a \circ D_b = D_{a+b}$ ist. Angewandt auf $b = -a$ ergibt dies $D_a \circ D_{-a} = D_{a+(-a)}$. Doch wegen $D_{a+(-a)} = D_{\bar{0}} = \text{id}$ wird dies zu $D_a \circ D_{-a} = \text{id}$, also zu $D_a^{-1} = D_{-a}$.

Diese Relation $\overset{H}{\sim}$ heißt die *von der Abbildungsgruppe H induzierte Äquivalenzrelation* auf der Menge F .

Wir haben noch nicht gezeigt, daß diese Relation $\overset{H}{\sim}$ eine Äquivalenzrelation ist, aber dies holen wir schnell nach - und beweisen dabei gleich eine Formel für die Anzahl der Äquivalenzklassen:

Satz 5.3 (Nicht-Burnside-Lemma für Abbildungsgruppen): Sei F eine Menge, und H eine Abbildungsgruppe von F .

(a) Die vorhin definierte Relation $\overset{H}{\sim}$ ist eine Äquivalenzrelation auf der Menge F .

(b) Ist F eine endliche Menge, dann gilt

$$|F / \overset{H}{\sim}| = \frac{1}{|H|} \sum_{h \in H} |\text{Fix } h|.$$

Bemerkung: Hierbei bezeichnet $\text{Fix } h$ die Menge $\{f \in F \mid h(f) = f\}$ für jedes $h \in H$.

Der zugegebenerweise merkwürdige Name von Satz 5.3 geht auf die Tatsache zurück, daß dieser Satz oft in der Literatur als "Burnside-Lemma" bezeichnet wurde, obwohl er nicht von William Burnside stammt (dieser hat ihn lediglich benutzt, und dabei selber auf einen früheren Entdecker - Ferdinand Georg Frobenius - verwiesen). Bekannt war Satz 5.3 indes bereits Augustin Louis Cauchy 1845.

Bevor wir zum Beweis von diesem Satz kommen, wollen wir denjenigen Lesern, die mit dem Begriff einer (abstrakten) *Gruppe* (und nicht nur einer Abbildungsgruppe) vertraut sind, eine leichte Umformulierung von diesem Satz 5.3 vorstellen, die auch heutzutage deutlich bekannter ist als Satz 5.3 selber:

Satz 5.4 (Nicht-Burnside-Lemma): Sei F eine Menge, und H eine endliche Gruppe, die auf der Menge F operiert. Wir definieren eine Relation $\overset{H}{\sim}$ auf der Menge F folgendermaßen: Für zwei Elemente $f \in F$ und $g \in F$ sei genau dann $f \overset{H}{\sim} g$, wenn es ein $h \in H$ mit $hf = g$ gibt.

(a) Diese Relation $\overset{H}{\sim}$ ist eine Äquivalenzrelation auf der Menge F .

(b) Ist F eine endliche Menge, dann gilt

$$|F / \overset{H}{\sim}| = \frac{1}{|H|} \sum_{h \in H} |\text{Fix } h|.$$

Bemerkung: Hierbei bezeichnet $\text{Fix } h$ die Menge $\{f \in F \mid hf = f\}$ für jedes $h \in H$.

Wenn in der Literatur vom "Burnside-Lemma", "orbit counting lemma" oder "the Lemma that isn't Burnside's" die Rede ist, ist meistens Satz 5.4 gemeint.

Offensichtlich ist Satz 5.3 ein Sonderfall von Satz 5.4 (denn eine Abbildungsgruppe von der Menge F operiert immer kanonisch auf der Menge F), doch man kann auch

sehr leicht Satz 5.4 aus Satz 5.3 folgern. Wir wollen aber auf diesen Beweis verzichten, da man stattdessen auch einfach Satz 5.4 völlig analog zu Satz 5.3 beweisen könnte⁶⁶, und da wir Satz 5.4 sowieso nie verwenden werden.

Beweis von Satz 5.3: **(a)** Die Relation $\overset{H}{\sim}$ ist reflexiv (denn für jedes $f \in F$ ist $f \overset{H}{\sim} f$, weil es ein $h \in H$ mit $h(f) = f$ gibt (nämlich $h = \text{id}$ ⁶⁷)), symmetrisch (denn für beliebige $f \in F$ und $f' \in F$, die $f \overset{H}{\sim} f'$ erfüllen, gilt auch $f' \overset{H}{\sim} f$ ⁶⁸) und transitiv (denn für beliebige $f \in F$, $f' \in F$ und $f'' \in F$, die $f \overset{H}{\sim} f'$ und $f' \overset{H}{\sim} f''$ erfüllen, gilt auch $f \overset{H}{\sim} f''$ ⁶⁹). Daher ist die Relation $\overset{H}{\sim}$ eine Äquivalenzrelation, und Satz 5.3 **(a)** ist bewiesen.

(b) Wir haben

$$|\{h \in H \mid h(f) = g\}| = 0 \quad \text{für beliebige } f \in H \text{ und } g \in H \text{ mit } f \not\overset{H}{\sim} g \quad (48)$$

(denn wegen $f \not\overset{H}{\sim} g$ gibt es kein $h \in H$, das $h(f) = g$ erfüllt).

Ferner haben wir

$$|\{h \in H \mid h(f) = g\}| = |\{h \in H \mid h(f) = f\}| \quad \text{für beliebige } f \in H \text{ und } g \in H \text{ mit } f \overset{H}{\sim} g. \quad (49)$$

Beweis von (49): Seien $f \in H$ und $g \in H$ mit $f \overset{H}{\sim} g$ gegeben. Wegen $f \overset{H}{\sim} g$ gibt es ein $h_0 \in H$, das $h_0(f) = g$ erfüllt. Somit ist die Abbildung

$$\begin{aligned} \{h \in H \mid h(f) = f\} &\rightarrow \{h \in H \mid h(f) = g\}, & \text{gegeben durch} \\ \eta &\mapsto h_0 \circ \eta \end{aligned}$$

wohldefiniert⁷⁰, und ferner ist die Abbildung

$$\begin{aligned} \{h \in H \mid h(f) = g\} &\rightarrow \{h \in H \mid h(f) = f\}, & \text{gegeben durch} \\ \varepsilon &\mapsto h_0^{-1} \circ \varepsilon \end{aligned}$$

⁶⁶Man müsste nur immer die Verkettung \circ durch die Multiplikation in der Gruppe H ersetzen, und Terme wie $h(f)$ durch Terme wie hf ersetzen.

⁶⁷Denn da H eine Abbildungsgruppe ist, gilt $\text{id} \in H$.

⁶⁸Denn wegen $f \overset{H}{\sim} f'$ gibt es ein $h \in H$ mit $h(f) = f'$, und somit gibt es auch ein $\tilde{h} \in H$ mit $\tilde{h}(f') = f$ (nämlich $\tilde{h} = h^{-1}$, denn da H eine Abbildungsgruppe ist, folgt aus $h \in H$ auch $h^{-1} \in H$), und daher ist $f' \overset{H}{\sim} f$.

⁶⁹*Beweis:* Wegen $f \overset{H}{\sim} f'$ gibt es ein $h_1 \in H$ mit $h_1(f) = f'$, und wegen $f' \overset{H}{\sim} f''$ gibt es ein $h_2 \in H$

mit $h_2(f') = f''$. Also ist $(h_2 \circ h_1)(f) = h_2 \left(\underbrace{h_1(f)}_{=f'} \right) = h_2(f') = f''$. Ferner ist $h_2 \circ h_1 \in H$

(da H eine Abbildungsgruppe ist, und $h_2 \in H$ und $h_1 \in H$ gilt). Somit gibt es ein $\tilde{h} \in H$ mit $\tilde{h}(f) = f''$ (nämlich $\tilde{h} = h_2 \circ h_1$). Daher ist $f \overset{H}{\sim} f''$.

⁷⁰Denn für jedes $\eta \in \{h \in H \mid h(f) = f\}$ ist $\eta(f) = f$, also $(h_0 \circ \eta)(f) = h_0 \left(\underbrace{\eta(f)}_{=f} \right) = h_0(f) = g$,

und daher $h_0 \circ \eta \in \{h \in H \mid h(f) = g\}$ (denn aus $h_0 \in H$ und $\eta \in H$ folgt $h_0 \circ \eta \in H$, weil H eine Abbildungsgruppe ist).

wohldefiniert⁷¹. Diese beiden Abbildungen sind zueinander invers⁷²; also sind sie beide bijektiv, und folglich gibt es Bijektionen zwischen den Mengen $\{h \in H \mid h(f) = f\}$ und $\{h \in H \mid h(f) = g\}$. Hieraus folgt $|\{h \in H \mid h(f) = g\}| = |\{h \in H \mid h(f) = f\}|$. Damit ist (49) bewiesen.

Für jedes $f \in F$ haben wir nun

$$\begin{aligned}
|H| &= \sum_{g \in F} |\{h \in H \mid h(f) = g\}| = \sum_{\substack{g \in F; \\ f \not\sim^H g}} \underbrace{|\{h \in H \mid h(f) = g\}|}_{=0 \text{ (nach (48))}} + \sum_{\substack{g \in F; \\ f \sim^H g}} \underbrace{|\{h \in H \mid h(f) = g\}|}_{=|\{h \in H \mid h(f)=f\}| \text{ (nach (49))}} \\
&= \sum_{\substack{g \in F; \\ f \not\sim^H g}} 0 + \sum_{\substack{g \in F; \\ f \sim^H g}} |\{h \in H \mid h(f) = f\}| = \left| \underbrace{\{g \in F \mid f \sim^H g\}}_{\text{dies ist genau die Äquivalenzklasse von } f \text{ (bezüglich der Relation } \sim^H)} \right| \cdot |\{h \in H \mid h(f) = f\}| \\
&= \left| \text{Äquivalenzklasse von } f \right| \cdot |\{h \in H \mid h(f) = f\}|,
\end{aligned}$$

also

$$\frac{|H|}{|\text{Äquivalenzklasse von } f|} = |\{h \in H \mid h(f) = f\}|.$$

Doch nach Satz 5.2 (angewandt auf die Äquivalenzrelation \sim^H anstelle von \sim) ist

$$|F / \sim^H| = \sum_{f \in F} \frac{1}{|\text{Äquivalenzklasse von } f|},$$

also

$$\begin{aligned}
|H| \cdot |F / \sim^H| &= |H| \cdot \sum_{f \in F} \frac{1}{|\text{Äquivalenzklasse von } f|} = \sum_{f \in F} \frac{|H|}{|\text{Äquivalenzklasse von } f|} \\
&= \sum_{f \in F} |\{h \in H \mid h(f) = f\}| = |\{(h, f) \in H \times F \mid h(f) = f\}| \\
&= \sum_{h \in H} \left| \underbrace{\{f \in F \mid h(f) = f\}}_{=\text{Fix } h} \right| = \sum_{h \in H} |\text{Fix } h|,
\end{aligned}$$

also $|F / \sim^H| = \frac{1}{|H|} \sum_{h \in H} |\text{Fix } h|$, und damit ist Satz 5.3 bewiesen.

⁷¹Denn für jedes $\varepsilon \in \{h \in H \mid h(f) = g\}$ ist $\varepsilon(f) = g$, also $(h_0^{-1} \circ \varepsilon)(f) = h_0^{-1} \left(\underbrace{\varepsilon(f)}_{=g} \right) = h_0^{-1}(g) =$

f (denn $h_0(f) = g$), und daher $h_0^{-1} \circ \varepsilon \in \{h \in H \mid h(f) = f\}$ (denn aus $h_0 \in H$ folgt $h_0^{-1} \in H$ (weil H eine Abbildungsgruppe ist), und zusammen mit $\varepsilon \in H$ ergibt dies $h_0^{-1} \circ \varepsilon \in H$ (wieder weil H eine Abbildungsgruppe ist)).

⁷²Denn für jedes $\eta \in \{h \in H \mid h(f) = f\}$ ist $h_0^{-1} \circ (h_0 \circ \eta) = \eta$, und für jedes $\varepsilon \in \{h \in H \mid h(f) = g\}$ ist $h_0 \circ (h_0^{-1} \circ \varepsilon) = \varepsilon$.

Bevor wir mithilfe von Satz 5.3 den Satz 5.1 beweisen, wollen wir ein einfacheres Beispiel für die Anwendung von Satz 5.3 kennenlernen, nämlich die erste Aufgabe der Internationalen Mathematik-Olympiade 1987:

Satz 5.5: Sei $n \in \mathbb{N}_+$. Dann gilt

$$n! = \sum_{k=0}^n k \cdot (\text{Anzahl aller Permutationen der Menge } \{1, 2, \dots, n\}, \text{ die genau } k \text{ Fixpunkte haben}).$$

Beweis von Satz 5.5: Sei N die Menge $\{1, 2, \dots, n\}$. Anwendung von Satz 5.3 (b) auf den Fall $F = N$ und $H = S_N$ ergibt

$$\left| N / \overset{S_N}{\sim} \right| = \frac{1}{|S_N|} \sum_{h \in S_N} |\text{Fix } h|. \quad (50)$$

Doch $\left| N / \overset{S_N}{\sim} \right| = 1$ (denn es gibt nur eine Äquivalenzklasse bezüglich der Relation $\overset{S_N}{\sim}$ auf der Menge N , weil alle Elemente der Menge N zueinander äquivalent bezüglich dieser Relation sind⁷³). Ferner ist S_N die Menge aller bijektiven Abbildungen von N nach N , also die Menge aller Permutationen von N ; hieraus folgt $|S_N| = |N|! = n!$ (da $|N| = |\{1, 2, \dots, n\}| = n$). Für jedes $h \in S_N$ ist schließlich

$$\text{Fix } h = \{f \in \{1, 2, \dots, n\} \mid h(f) = f\} = \{f \in \{1, 2, \dots, n\} \mid f \text{ ist ein Fixpunkt von } h\}$$

die Menge aller Fixpunkte von h . All dies zusammen ergibt

$$\begin{aligned} 1 &= \left| N / \overset{S_N}{\sim} \right| = \frac{1}{|S_N|} \sum_{h \in S_N} \underbrace{|\text{Fix } h|}_{\substack{=|\text{Menge aller Fixpunkte von } h| \\ =(\text{Anzahl aller Fixpunkte von } h)}} && \text{(nach (50))} \\ &= \frac{1}{n!} \quad (\text{denn } |S_N| = n!) \\ &= \frac{1}{n!} \sum_{h \in S_N} (\text{Anzahl aller Fixpunkte von } h), \end{aligned}$$

also

$$\begin{aligned} n! &= \sum_{h \in S_N} (\text{Anzahl aller Fixpunkte von } h) = \sum_{k=0}^n \sum_{\substack{h \in S_N; \\ \text{die Anzahl aller} \\ \text{Fixpunkte von } h \text{ ist } k}} \underbrace{(\text{Anzahl aller Fixpunkte von } h)}_{=k} \\ &= \sum_{k=0}^n \sum_{\substack{h \in S_N; \\ \text{die Anzahl aller} \\ \text{Fixpunkte von } h \text{ ist } k}} k \\ &= \underbrace{|\{h \in S_N \mid \text{die Anzahl aller Fixpunkte von } h \text{ ist } k\}|}_{\substack{=|\text{Anzahl aller Abbildungen } h \in S_N, \text{ die genau } k \text{ Fixpunkte haben}| \\ =(\text{Anzahl aller Permutationen der Menge } \{1, 2, \dots, n\}, \text{ die genau } k \text{ Fixpunkte haben}) \\ (\text{denn } S_N \text{ ist die Menge aller Permutationen der Menge } N = \{1, 2, \dots, n\})}} \cdot k \\ &= \sum_{k=0}^n k \cdot (\text{Anzahl aller Permutationen der Menge } \{1, 2, \dots, n\}, \text{ die genau } k \text{ Fixpunkte haben}). \end{aligned}$$

⁷³Denn für je zwei Elemente $a \in N$ und $b \in N$ gilt $a \overset{S_N}{\sim} b$ (weil es eine Abbildung $h \in S_N$ gibt, die $h(a) = b$ erfüllt (denn S_N ist die Menge aller bijektiven Abbildungen von N nach N , und es gibt stets eine bijektive Abbildung von N nach N , die a auf b abbildet)).

Damit ist Satz 5.5 bewiesen.

5.4. Beweis von Satz 5.1

Nun wollen wir Satz 5.1 beweisen. Dazu schauen wir uns an, was Satz 5.3 für unsere Perlenketten besagt. Kommen wir also zu unserem Beispiel zurück, wo $F = \{1, 2, \dots, q\}^{\mathbb{Z}/n\mathbb{Z}}$ und $H = \{D_{\bar{0}}, D_{\bar{1}}, \dots, D_{\overline{n-1}}\} = \{D_a \mid a \in \mathbb{Z}/n\mathbb{Z}\}$ war, wobei die Abbildungen $D_a : F \rightarrow F$ für alle $a \in \mathbb{Z}/n\mathbb{Z}$ gemäß (45) definiert waren. Wir wissen bereits, daß H eine Abbildungsgruppe von F ist. Wir wollen nun zeigen, daß die von H induzierte Äquivalenzrelation $\overset{H}{\sim}$ genau unsere Äquivalenzrelation $\overset{\text{neck}}{\sim}$ ist. Dazu stellen wir fest, daß für je zwei Funktionen $f \in F$ und $g \in F$ folgende Kette von Äquivalenzen gilt:

$$\begin{aligned}
(f \overset{H}{\sim} g) &\iff (\text{es gibt ein } h \in H \text{ mit } h(f) = g) \\
&\iff (\text{es gibt ein } a \in \mathbb{Z}/n\mathbb{Z} \text{ mit } D_a(f) = g) \\
&\iff (\text{es gibt ein } a \in \mathbb{Z}/n\mathbb{Z} \text{ mit } ((D_a(f))(x) = g(x) \text{ für alle } x \in \mathbb{Z}/n\mathbb{Z})) \\
&\iff (\text{es gibt ein } a \in \mathbb{Z}/n\mathbb{Z} \text{ mit } (f(x+a) = g(x) \text{ für alle } x \in \mathbb{Z}/n\mathbb{Z})) \\
&\iff (f \overset{\text{neck}}{\sim} g). \tag{51}
\end{aligned}$$

Hierbei folgt der erste Äquivalenzpfeil aus der Definition der Relation $\overset{H}{\sim}$; der zweite Äquivalenzpfeil folgt aus $H = \{D_a \mid a \in \mathbb{Z}/n\mathbb{Z}\}$; der dritte ist trivial (denn zwei Funktionen sind genau dann gleich, wenn alle ihre entsprechenden Werte gleich sind); der vierte folgt aus (45); schließlich folgt der fünfte aus der Definition der Relation $\overset{\text{neck}}{\sim}$.

Wir haben damit gezeigt, daß die Äquivalenzrelation $\overset{H}{\sim}$ identisch mit der Äquivalenzrelation $\overset{\text{neck}}{\sim}$ ist. Somit ist

$$\begin{aligned}
|F / \overset{\text{neck}}{\sim}| &= |F / \overset{H}{\sim}| = \frac{1}{|H|} \sum_{h \in H} |\text{Fix } h| \quad (\text{nach Satz 5.3}) \\
&= \frac{1}{n} \sum_{a=0}^{n-1} |\text{Fix } D_{\bar{a}}| \tag{52}
\end{aligned}$$

(denn da $H = \{D_{\bar{0}}, D_{\bar{1}}, \dots, D_{\overline{n-1}}\}$ gilt, und da die Abbildungen $D_{\bar{0}}, D_{\bar{1}}, \dots, D_{\overline{n-1}}$ paarweise verschieden sind, haben wir $|H| = n$ und $\sum_{h \in H} |\text{Fix } h| = \sum_{a=0}^{n-1} |\text{Fix } D_{\bar{a}}|$). Wir wollen nun $|\text{Fix } D_{\bar{a}}|$ für jedes $a \in \{0, 1, \dots, n-1\}$ ausrechnen. Genauer gesagt: Wir wollen zeigen, daß $|\text{Fix } D_{\bar{a}}| = q^{\text{ggT}(a,n)}$ für jedes $a \in \{0, 1, \dots, n-1\}$ ist.

Bevor wir dies beweisen, geben wir zuerst einen unformalen (aber hoffentlich halbwegs anschaulichen) Beweisplan (wer nur den formalen Beweis will, kann diesen Absatz überspringen): Die Menge $\text{Fix } D_{\bar{a}}$ war ja definiert als die Menge aller strikten Färbungen⁷⁴ $f \in F$, welche von der Abbildung $D_{\bar{a}}$ in sich selbst überführt werden, d. h. die Menge aller strikten Färbungen $f \in F$, welche von der Rotation der Perlenkette um den Winkel $\frac{a}{n} \cdot 360^\circ$ in sich selbst überführt werden (denn die Abbildung $D_{\bar{a}}$ entspricht der Rotation der Perlenkette um den Winkel $\frac{a}{n} \cdot 360^\circ$, bzw. der

⁷⁴Zur Erinnerung: Unter einer "strikten Färbung" verstehen wir eine Abbildung von $\mathbb{Z}/n\mathbb{Z}$ nach $\{1, 2, \dots, q\}$. Strikte Färbungen sind nicht zu verwechseln mit Färbungen (ohne das Attribut "strikt"), welche als Äquivalenzklassen von solchen Abbildungen definiert sind.

Wirkung, die diese Drehung auf strikte Färbungen der Perlenkette ausübt). Nun ist eine strikte Färbung eine Abbildung von $\mathbb{Z}/n\mathbb{Z}$ nach $\{1, 2, \dots, q\}$, und solche Abbildungen entsprechen eineindeutig n -periodischen Abbildungen von \mathbb{Z} nach $\{1, 2, \dots, q\}$. Wenn die strikte Färbung f von der Rotation der Perlenkette um den Winkel $\frac{a}{n} \cdot 360^\circ$ in sich selbst überführt wird, muß die entsprechende n -periodische Abbildung von \mathbb{Z} nach $\{1, 2, \dots, q\}$ auch noch a -periodisch sein; sie ist also auch $\text{ggT}(n, a)$ -periodisch, und die Anzahl aller $\text{ggT}(n, a)$ -periodischen Abbildungen von \mathbb{Z} nach $\{1, 2, \dots, q\}$ ist $q^{\text{ggT}(n, a)}$.

Hier der formale *Beweis von* $|\text{Fix } D_{\bar{a}}| = q^{\text{ggT}(a, n)}$: Zuerst eine Definition:

Definition (k -periodische Abbildung): Sei $k \in \mathbb{Z}$. Sei Q eine Menge. Sei $f : \mathbb{Z} \rightarrow Q$ eine Abbildung. Die Abbildung f heie k -periodisch, wenn $(f(x + k) = f(x))$ fur alle $x \in \mathbb{Z}$ gilt.

Wir bezeichnen mit $Q_{\text{per } k}^{\mathbb{Z}}$ die Menge aller k -periodischen Abbildungen von \mathbb{Z} nach Q .

Folgende Eigenschaften k -periodischer Abbildungen sind leicht einzusehen:

- Ist $k \in \mathbb{Z}$ eine ganze Zahl, Q eine Menge, und $f : \mathbb{Z} \rightarrow Q$ eine k -periodische Abbildung, und ist $n \in \mathbb{N}$, dann gilt $f(x + nk) = f(x)$ fur alle $x \in \mathbb{Z}$.⁷⁵
- Ist $k \in \mathbb{Z}$ eine ganze Zahl, Q eine Menge, und $f : \mathbb{Z} \rightarrow Q$ eine k -periodische Abbildung, und sind x und y zwei Elemente von \mathbb{Z} mit $x \equiv y \pmod{k}$, dann ist

$$f(x) = f(y). \quad (53)$$

76

- Ist $k \in \mathbb{Z}$ eine ganze Zahl, Q eine Menge, und $f : \mathbb{Z} \rightarrow Q$ eine k -periodische Abbildung, dann gilt

$$f(x) = f(x \bmod k) \quad \text{fur alle } x \in \mathbb{Z}. \quad (54)$$

⁷⁷ Hierbei bezeichnen wir mit $x \bmod k$ den Rest, den die Zahl x bei Division durch k lat.

- Jede Abbildung $f : \mathbb{Z} \rightarrow Q$ ist 0-periodisch.

⁷⁵Dies beweist man durch vollstandige Induktion nach n , unter Ausnutzung der Definition von "k-periodisch".

⁷⁶*Beweis:* Ohne Beschrankung der Allgemeinheit durfen wir annehmen, da $x \leq y$ ist (denn sonst konnen wir einfach x mit y vertauschen). Dann ist $y - x \geq 0$. Ferner nehmen wir ebenfalls ohne Beschrankung der Allgemeinheit an, da $k > 0$ ist (denn fur $k = 0$ ist die Aussage trivial, und fur $k < 0$ konnen wir einfach k durch $-k$ ersetzen). Damit wird $\frac{y-x}{k} \geq 0$ (da $y - x \geq 0$).

Aus $x \equiv y \pmod{k}$ folgt aber $k \mid y - x$ und damit $\frac{y-x}{k} \in \mathbb{Z}$. Zusammen mit $\frac{y-x}{k} \geq 0$ wird dies zu $\frac{y-x}{k} \in \mathbb{N}$. Nun wissen wir bereits, da $f(x + nk) = f(x)$ fur jedes $n \in \mathbb{N}$ ist; angewandt auf $n = \frac{y-x}{k}$ ergibt dies $f\left(x + \frac{y-x}{k}k\right) = f(x)$, also $f(y) = f(x)$ (da $x + \frac{y-x}{k}k = y$), und damit $f(x) = f(y)$, was zu beweisen war.

⁷⁷Dies folgt aus (53), angewandt auf $y = x \bmod k$ (denn $x \equiv (x \bmod k) \pmod{k}$).

- Ist $k \in \mathbb{Z}$ eine ganze Zahl, Q eine Menge, und $f : \mathbb{Z} \rightarrow Q$ eine k -periodische Abbildung, dann gilt:

$$\begin{aligned} & \text{Ist } k' \in \mathbb{Z} \text{ eine ganze Zahl mit } k \mid k', \\ & \text{dann ist die Abbildung } f : \mathbb{Z} \rightarrow Q \text{ auch } k'\text{-periodisch.} \end{aligned} \quad (55)$$

78

- Sind k und ℓ zwei von 0 verschiedene ganze Zahlen, Q eine Menge, und $f : \mathbb{Z} \rightarrow Q$ eine Abbildung, dann gilt:

$$\begin{aligned} & \text{Genau dann ist die Abbildung } f \text{ gleichzeitig } k\text{-periodisch und } \ell\text{-periodisch,} \\ & \text{wenn sie } \text{ggT}(k, \ell)\text{-periodisch ist.} \end{aligned} \quad (56)$$

79

Nun wollen wir sehen, daß k -periodische Abbildungen $\mathbb{Z} \rightarrow Q$ (für ein festes $k \in \mathbb{N}_+$) in eindeutiger Zuordnung zu Abbildungen $\mathbb{Z}/k\mathbb{Z} \rightarrow Q$ stehen. Die Idee dahinter ist die folgende: Wenn $f : \mathbb{Z} \rightarrow Q$ eine k -periodische Abbildung ist, dann hängt der Wert von $f(x)$ (für $x \in \mathbb{Z}$) nur von der Restklasse von x modulo k ab (wegen (53)), und somit kann man f als Funktion der Restklasse von x modulo k (statt als Funktion der ganzen Zahl x) auffassen. Umgekehrt: Hat man eine Abbildung $F : \mathbb{Z}/k\mathbb{Z} \rightarrow Q$,

⁷⁸*Beweis:* Für jedes $x \in \mathbb{Z}$ gilt $f(x+k') = f(x)$ (denn (53) (angewandt auf $y = x+k'$) ergibt $f(x) = f(x+k')$, da $x \equiv x+k' \pmod{k}$ ist (was wiederum aus $k \mid k'$ folgt)).

⁷⁹*Beweis von (56):* Um (56) zu beweisen, müssen wir zwei Aussagen nachweisen:

Aussage 1: Ist die Abbildung f gleichzeitig k -periodisch und ℓ -periodisch, dann ist sie $\text{ggT}(k, \ell)$ -periodisch.

Aussage 2: Ist f eine $\text{ggT}(k, \ell)$ -periodische Abbildung, dann ist f gleichzeitig k -periodisch und ℓ -periodisch.

Beweis: Angenommen, f sei gleichzeitig k -periodisch und ℓ -periodisch. Wir müssen dann zeigen, daß f auch $\text{ggT}(k, \ell)$ -periodisch ist.

In der Tat sei $x \in \mathbb{Z}$. Der Satz von Bezout besagt, daß für zwei beliebige teilerfremde ganze Zahlen α und β gilt: Es gibt ganze Zahlen u und v mit $u\alpha + v\beta = 1$. Angewandt auf $\alpha = \frac{k}{\text{ggT}(k, \ell)}$

und $\beta = \frac{\ell}{\text{ggT}(k, \ell)}$ (diese beiden Zahlen α und β sind ganz, denn $\text{ggT}(k, \ell)$ teilt sowohl k als auch

ℓ) ergibt dies, daß es ganze Zahlen u und v mit $u \cdot \frac{k}{\text{ggT}(k, \ell)} + v \cdot \frac{\ell}{\text{ggT}(k, \ell)} = 1$ gibt. Also ist

$\frac{uk + v\ell}{\text{ggT}(k, \ell)} = u \cdot \frac{k}{\text{ggT}(k, \ell)} + v \cdot \frac{\ell}{\text{ggT}(k, \ell)} = 1$, und damit $uk + v\ell = \text{ggT}(k, \ell)$. Nun ist

$$f(x) = f(x + uk) \quad (\text{nach (53), angewandt auf } y = x + uk, \text{ denn } x \equiv x + uk \pmod{k})$$

$$\begin{aligned} & = f\left(x + \underbrace{uk + v\ell}_{=\text{ggT}(k, \ell)}\right) \quad \left(\begin{array}{l} \text{nach (53), angewandt auf } \ell, x + uk \text{ und } x + uk + v\ell \text{ statt } k, x \text{ und } y, \\ \text{denn } x + uk \equiv x + uk + v\ell \pmod{\ell} \end{array} \right) \\ & = f(x + \text{ggT}(k, \ell)). \end{aligned}$$

Da dies für alle $x \in \mathbb{Z}$ gilt, folgt hieraus, daß f eine $\text{ggT}(k, \ell)$ -periodische Abbildung ist, und Aussage 1 ist bewiesen.

Beweis von Aussage 2: Ist f eine $\text{ggT}(k, \ell)$ -periodische Abbildung, dann ist f auch k -periodisch (denn $\text{ggT}(k, \ell) \mid k$) und ℓ -periodisch (denn $\text{ggT}(k, \ell) \mid \ell$). Damit ist Aussage 2 bewiesen.

Nachdem nun beide Aussagen 1 und 2 gezeigt sind, ist (56) nachgewiesen.

dann kann man aus ihr eine k -periodische Abbildung $f : \mathbb{Z} \rightarrow Q$ erhalten, wenn man $f(x) = F(\bar{x}_k)$ für alle $x \in \mathbb{Z}$ setzt (wobei \bar{x}_k die Restklasse von x modulo k bedeutet). Wir wollen diese Zuordnung formal definieren:

Definition ($R_{Q,k}$ und $R'_{Q,k}$): Sei $k \in \mathbb{N}_+$ eine positive ganze Zahl, und Q eine endliche Menge. Für jedes $x \in \mathbb{Z}$ bezeichnen wir mit \bar{x}_k die Restklasse von x modulo k . (Insbesondere wird das, was wir weiter oben mit \bar{x} bezeichnet haben, in unserer neuen Notation \bar{x}_n genannt).

(a) Wir definieren nun eine Abbildung

$$R_{Q,k} : Q_{\text{per } k}^{\mathbb{Z}} \rightarrow Q^{\mathbb{Z}/k\mathbb{Z}}$$

dadurch, daß wir für jede k -periodische Abbildung $f : \mathbb{Z} \rightarrow Q$ eine Abbildung $R_{Q,k}(f) : \mathbb{Z}/k\mathbb{Z} \rightarrow Q$ durch

$$((R_{Q,k}(f))(\bar{x}_k) = f(x) \quad \text{für alle } x \in \mathbb{Z})$$

definieren. (Diese Abbildung $R_{Q,k}(f)$ ist hierdurch tatsächlich wohldefiniert, denn jedes Element von $\mathbb{Z}/k\mathbb{Z}$ läßt sich in der Form \bar{x}_k für ein $x \in \mathbb{Z}$ schreiben, und wenn ein Element von $\mathbb{Z}/k\mathbb{Z}$ sich auf zweierlei Art als \bar{x}_k für $x \in \mathbb{Z}$ schreiben läßt, dann sind die Werte von $f(x)$ für diese beiden möglichen x gleich⁸⁰.)

(b) Wir definieren ferner eine Abbildung

$$R'_{Q,k} : Q^{\mathbb{Z}/k\mathbb{Z}} \rightarrow Q_{\text{per } k}^{\mathbb{Z}}$$

dadurch, daß wir für jede Abbildung $F : \mathbb{Z}/k\mathbb{Z} \rightarrow Q$ eine k -periodische Abbildung $R'_{Q,k}(F) : \mathbb{Z} \rightarrow Q$ durch

$$((R'_{Q,k}(F))(x) = F(\bar{x}_k) \quad \text{für alle } x \in \mathbb{Z})$$

definieren. (Diese Abbildung $R'_{Q,k}(F) : \mathbb{Z} \rightarrow Q$ ist in der Tat k -periodisch,

denn für jedes $x \in \mathbb{Z}$ gilt $(R'_{Q,k}(F))(x+k) = F\left(\underbrace{\overline{x+k}_k}_{=\bar{x}_k \text{ (denn } x+k \equiv x \pmod{k})}\right) =$

$$F(\bar{x}_k) = (R'_{Q,k}(F))(x) .)$$

Für jedes $k \in \mathbb{N}_+$ und jede endliche Menge Q gilt $R_{Q,k} \circ R'_{Q,k} = \text{id}$ (denn für jede Abbildung $F : \mathbb{Z}/k\mathbb{Z} \rightarrow Q$ ist $(R_{Q,k} \circ R'_{Q,k})(F) = \text{id}(F)$ ⁸¹) und $R'_{Q,k} \circ R_{Q,k} = \text{id}$

⁸⁰Denn sind x und y zwei Elemente von \mathbb{Z} mit $\bar{x}_k = \bar{y}_k$, dann ist $f(x) = f(y)$ (denn aus $\bar{x}_k = \bar{y}_k$ folgt $x \equiv y \pmod{k}$, und somit ist $f(x) = f(y)$ gemäß (53)).

⁸¹denn für jedes $x \in \mathbb{Z}$ ist

$$\begin{aligned} & \left(\underbrace{(R_{Q,k} \circ R'_{Q,k})(F)}_{=R_{Q,k}(R'_{Q,k}(F))} \right) (\bar{x}_k) \\ &= (R_{Q,k}(R'_{Q,k}(F)))(\bar{x}_k) = (R'_{Q,k}(F))(x) \quad (\text{nach der Definition von } R_{Q,k}) \\ &= F(\bar{x}_k) \quad (\text{nach der Definition von } R'_{Q,k}(F)) \\ &= (\text{id}(F))(\bar{x}_k), \end{aligned}$$

(denn für jede k -periodische Abbildung $f : \mathbb{Z} \rightarrow Q$ ist $(R'_{Q,k} \circ R_{Q,k})(f) = \text{id}(f)$ ⁸²). Somit sind die Abbildungen $R_{Q,k}$ und $R'_{Q,k}$ zueinander invers, und folglich ist $R_{Q,k}$ eine Bijektion. Hieraus folgt: Ist $k \in \mathbb{N}_+$ eine positive ganze Zahl, und Q eine endliche Menge, dann ist

$$|Q_{\text{per } k}^{\mathbb{Z}}| = |Q|^k. \quad (57)$$

83

Nun zum Beweis von $|\text{Fix } D_{\bar{a}}| = q^{\text{ggT}(a,n)}$: Setzen wir $Q = \{1, 2, \dots, q\}$. Dann ist $|Q| = q$. Für eine Abbildung $f \in F$ gilt folgende Kette von Äquivalenzen:⁸⁴

$$\begin{aligned} (f \in \text{Fix } D_{\bar{a}}) &\iff (D_{\bar{a}}(f) = f) \iff ((D_{\bar{a}}(f))(y) = f(y) \text{ für alle } y \in \mathbb{Z}/n\mathbb{Z}) \\ &\iff ((D_{\bar{a}}(f))(\bar{x}) = f(\bar{x}) \text{ für alle } x \in \mathbb{Z}) \\ &\iff (f(\bar{x} + \bar{a}) = f(\bar{x}) \text{ für alle } x \in \mathbb{Z}) \\ &\iff ((R'_{Q,n}(f))(x + a) = (R'_{Q,n}(f))(x) \text{ für alle } x \in \mathbb{Z}) \\ &\iff (\text{die Abbildung } R'_{Q,n}(f) : \mathbb{Z} \rightarrow Q \text{ ist } a\text{-periodisch}) \\ &\iff (\text{die Abbildung } R'_{Q,n}(f) : \mathbb{Z} \rightarrow Q \text{ ist gleichzeitig } a\text{-periodisch und } n\text{-periodisch}) \\ &\iff (\text{die Abbildung } R'_{Q,n}(f) : \mathbb{Z} \rightarrow Q \text{ ist } \text{ggT}(a, n)\text{-periodisch}) \\ &\iff (R'_{Q,n}(f) \in Q_{\text{per } \text{ggT}(a,n)}^{\mathbb{Z}}). \end{aligned} \quad (58)$$

Hierbei gilt der erste Äquivalenzpfeil wegen $\text{Fix } D_{\bar{a}} = \{f \in F \mid D_{\bar{a}}(f) = f\}$; der zweite Äquivalenzpfeil ist trivial; der dritte Äquivalenzpfeil folgt daraus, daß jedes Element $y \in \mathbb{Z}/n\mathbb{Z}$ sich in der Form \bar{x} für ein $x \in \mathbb{Z}$ schreiben läßt; der vierte Äquivalenzpfeil folgt aus $(D_{\bar{a}}(f))(\bar{x}) = f(\bar{x} + \bar{a})$ (dies gilt nach (45)); der fünfte Äquivalenzpfeil folgt aus $(R'_{Q,n}(f))(x + a) = f(\bar{x} + \bar{a}_n) = f(\bar{x} + \bar{a}) = f(\bar{x} + \bar{a})$ und $(R'_{Q,n}(f))(x) = f(\bar{x}_n) = f(\bar{x})$ (was beides aus der Definition von $R'_{Q,k}$ folgt); der sechste Äquivalenzpfeil folgt aus der Definition einer a -periodischen Abbildung; der siebte Äquivalenzpfeil

und somit gilt $(R_{Q,k} \circ R'_{Q,k})(F) = \text{id}(F)$ (denn jedes Element von $\mathbb{Z}/k\mathbb{Z}$ hat die Form \bar{x}_k für ein $x \in \mathbb{Z}$)

⁸²denn für jedes $x \in \mathbb{Z}$ ist

$$\begin{aligned} &\left(\underbrace{(R'_{Q,k} \circ R_{Q,k})(f)}_{= R'_{Q,k}(R_{Q,k}(f))} \right) (x) \\ &= (R'_{Q,k}(R_{Q,k}(f)))(x) = (R_{Q,k}(f))(\bar{x}_k) \quad (\text{nach der Definition von } R'_{Q,k}) \\ &= f(x) \quad (\text{nach der Definition von } R_{Q,k}) \\ &= (\text{id}(f))(x) \end{aligned}$$

⁸³Denn da die Abbildung $R_{Q,k} : Q_{\text{per } k}^{\mathbb{Z}} \rightarrow Q^{\mathbb{Z}/k\mathbb{Z}}$ eine Bijektion ist, gilt

$$|Q_{\text{per } k}^{\mathbb{Z}}| = |Q^{\mathbb{Z}/k\mathbb{Z}}| = |Q|^{|\mathbb{Z}/k\mathbb{Z}|} = |Q|^k$$

(da $|\mathbb{Z}/k\mathbb{Z}| = k$).

⁸⁴Begründungen der jeweiligen Äquivalenzen finden sich weiter unten.

folgt daraus, daß die Abbildung $R'_{Q,n}(f)$ immer n -periodisch ist (nach der Definition der Abbildung $R'_{Q,n}$); der achte Äquivalenzpfeil folgt daraus, daß eine Abbildung von \mathbb{Z} nach Q genau dann gleichzeitig a -periodisch und n -periodisch ist, wenn sie $\text{ggT}(a, n)$ -periodisch ist (dies folgt aus (56)); der neunte Äquivalenzpfeil folgt daraus, daß $Q_{\text{per ggT}(a,n)}^{\mathbb{Z}}$ die Menge aller $\text{ggT}(a, n)$ -periodischen Abbildungen von \mathbb{Z} nach Q ist.

Nun ist $\text{Fix } D_{\bar{a}}$ eine Teilmenge von $F = \{1, 2, \dots, q\}^{\mathbb{Z}/n\mathbb{Z}} = Q^{\mathbb{Z}/n\mathbb{Z}}$. Indes ist $Q_{\text{per ggT}(a,n)}^{\mathbb{Z}}$ eine Teilmenge von $Q_{\text{per } n}^{\mathbb{Z}}$ (denn jedes Element von $Q_{\text{per ggT}(a,n)}^{\mathbb{Z}}$ ist eine $\text{ggT}(a, n)$ -periodische Abbildung, also auch eine n -periodische Abbildung (denn $\text{ggT}(a, n) \mid n$) und damit ein Element von $Q_{\text{per } n}^{\mathbb{Z}}$). Die Äquivalenzkette (58) besagt also:

$$\text{Fix } D_{\bar{a}} = \left(R'_{Q,n}\right)^{-1} \left(Q_{\text{per ggT}(a,n)}^{\mathbb{Z}}\right). \quad (59)$$

Da $R'_{Q,n}$ eine Bijektion ist, folgt hieraus $|\text{Fix } D_{\bar{a}}| = \left|Q_{\text{per ggT}(a,n)}^{\mathbb{Z}}\right|$. Doch

$$\begin{aligned} \left|Q_{\text{per ggT}(a,n)}^{\mathbb{Z}}\right| &= |Q|^{\text{ggT}(a,n)} && \text{(nach (57), angewandt auf } k = \text{ggT}(a, n)) \\ &= q^{\text{ggT}(a,n)} \end{aligned}$$

(da $|Q| = q$). Zusammen mit $|\text{Fix } D_{\bar{a}}| = \left|Q_{\text{per ggT}(a,n)}^{\mathbb{Z}}\right|$ ergibt dies $|\text{Fix } D_{\bar{a}}| = q^{\text{ggT}(a,n)}$. Damit haben wir $|\text{Fix } D_{\bar{a}}| = q^{\text{ggT}(a,n)}$ bewiesen.

Damit wird (52) zu

$$\begin{aligned} \left|F / \overset{\text{neck}}{\sim}\right| &= \frac{1}{n} \sum_{a=0}^{n-1} \underbrace{|\text{Fix } D_{\bar{a}}|}_{=q^{\text{ggT}(a,n)}} = \frac{1}{n} \sum_{a=0}^{n-1} q^{\text{ggT}(a,n)} = \frac{1}{n} \sum_{i=0}^{n-1} q^{\text{ggT}(i,n)} \\ &\text{(hier haben wir den Summationsindex } a \text{ in } i \text{ umbenannt)} \\ &= \frac{1}{n} \left(q^{\text{ggT}(0,n)} + \sum_{i=1}^{n-1} q^{\text{ggT}(i,n)} \right) = \frac{1}{n} \left(q^{\text{ggT}(n,n)} + \sum_{i=1}^{n-1} q^{\text{ggT}(i,n)} \right) \\ &\text{(denn } \text{ggT}(0, n) = \text{ggT}(n, n), \text{ weil } \text{ggT}(0, n) = n = \text{ggT}(n, n)) \\ &= \frac{1}{n} \sum_{i=1}^n q^{\text{ggT}(i,n)}. \end{aligned}$$

Da aber $\sum_{i=1}^n q^{\text{ggT}(i,n)} = \sum_{d \mid n} \phi(d) q^{n/d}$ ist (wie wir bereits seit Abschnitt 1 wissen), wird dies zu

$$\left|F / \overset{\text{neck}}{\sim}\right| = \frac{1}{n} \sum_{d \mid n} \phi(d) q^{n/d}.$$

Nun ist $\left|F / \overset{\text{neck}}{\sim}\right|$ die Anzahl der Äquivalenzklassen bezüglich der Relation $\overset{\text{neck}}{\sim}$, in die die Menge $F = \{1, 2, \dots, q\}^{\mathbb{Z}/n\mathbb{Z}}$ zerfällt. Wir haben also gezeigt: Die Menge $\{1, 2, \dots, q\}^{\mathbb{Z}/n\mathbb{Z}}$ zerfällt in genau $\frac{1}{n} \sum_{d \mid n} \phi(d) q^{n/d}$ Äquivalenzklassen bezüglich der Relation $\overset{\text{neck}}{\sim}$. Damit ist Satz 5.1 (b) bewiesen.

Aus Satz 5.1 (b) folgt, daß $\frac{1}{n} \sum_{d \mid n} \phi(d) q^{n/d} \in \mathbb{Z}$ für jedes $q \in \mathbb{N}$ und $n \in \mathbb{N}_+$ gilt (denn laut Satz 5.1 (b) ist $\frac{1}{n} \sum_{d \mid n} \phi(d) q^{n/d}$ eine Anzahl, und Anzahlen sind immer

ganze Zahlen). Damit haben wir einen neuen Beweis von Satz 1.1 im Falle von $q \in \mathbb{N}$ erhalten. Doch dieser Fall ist nicht der einzig mögliche - es gibt auch noch den Fall $q \in \mathbb{Z} \setminus \mathbb{N}$ (also $q < 0$), und in diesem Fall können wir Satz 1.1 nicht mehr aus Satz 5.1 (b) folgern (denn Satz 5.1 setzt $q \in \mathbb{N}_+$ voraus, und ergibt für $q < 0$ keinen Sinn - was soll denn eine Färbung einer Perlenkette mit n Perlen in q Farben sein, wenn $q < 0$ ist?). Kann man trotzdem unseren Beweis so ergänzen, daß er für alle $q \in \mathbb{Z}$ (also auch für die negativen) funktioniert? Die Antwort ist "ja", und zwei Möglichkeiten, wie man diese Ergänzung durchführen kann, werden in Abschnitt 6 vorgestellt. Doch jetzt in Abschnitt 5 wollen wir uns weiter mit der Kombinatorik der Perlenketten und ihrer Verallgemeinerungen beschäftigen.

5.5. Aperiodische Perlenketten und $\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}$

Wir haben aus Satz 5.1 (b) gefolgert, daß $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} \in \mathbb{Z}$ für jedes $q \in \mathbb{N}$ und $n \in \mathbb{N}_+$ gilt. Nun wissen wir aus Abschnitt 2, daß auch $\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d} \in \mathbb{Z}$ gilt - und es stellt sich die Frage, ob auch die Zahl $\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}$ eine ähnliche kombinatorische Interpretation hat, d. h. Anzahl von irgendwelchen Färbungen ist. Und tatsächlich ist $\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}$ eine Anzahl - nämlich die der sogenannten *aperiodischen* Färbungen einer Perlenkette aus n Perlen in q Farben. Wir wollen erst einmal den Begriff einer aperiodischen Färbung einer Perlenkette aus n Perlen in q Farben (d. h. einer aperiodischen Abbildung von $\mathbb{Z}/n\mathbb{Z}$ nach $\{1, 2, \dots, q\}$, bzw. richtiger: einer Äquivalenzklasse solcher Abbildungen) formal definieren:

Definition (aperiodische Abbildung): Sei $n \in \mathbb{N}_+$, und sei Q eine Menge. Eine Abbildung $f : \mathbb{Z}/n\mathbb{Z} \rightarrow Q$ heie *aperiodisch*⁸⁵, wenn es keine von $\bar{0}$ verschiedene Restklasse $a \in \mathbb{Z}/n\mathbb{Z}$ gibt, die

$$(f(x+a) = f(x) \quad \text{für alle } x \in \mathbb{Z}/n\mathbb{Z})$$

erfüllt.

Wir bezeichnen mit $Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$ die Menge aller aperiodischen Abbildungen von $\mathbb{Z}/n\mathbb{Z}$ nach Q . Offensichtlich ist $Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$ eine Teilmenge von $Q^{\mathbb{Z}/n\mathbb{Z}}$.

Satz 5.6 (Aperiodischer Perlenkettensatz): Seien $q \in \mathbb{N}$ und $n \in \mathbb{N}_+$.

(a) Es gilt

$$|\{1, 2, \dots, q\}_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}| = \sum_{d|n} \mu(d) q^{n/d}.$$

(b) Wir betrachten die in Satz 5.1 definierte Äquivalenzrelation $\overset{\text{neck}}{\sim}$ auf der Menge $\{1, 2, \dots, q\}^{\mathbb{Z}/n\mathbb{Z}}$. Sei $\overset{\text{neck}}{\sim}_{\text{aper}}$ die Einschränkung dieser Relation $\overset{\text{neck}}{\sim}$ auf

⁸⁵Man verwechsle den Begriff "aperiodisch" nicht mit dem Begriff "a-periodisch" für ein $a \in \mathbb{Z}$!

die Teilmenge $\{1, 2, \dots, q\}_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$. Dann zerfällt die Menge $\{1, 2, \dots, q\}_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$ in genau $\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}$ Äquivalenzklassen bezüglich der Relation $\overset{\text{neck}}{\sim}_{\text{aper}}$.

Bevor wir Satz 5.6 beweisen, wollen wir eine Umschreibung von Satz 5.6 **(a)** geben, die von Abbildungen $\mathbb{Z} \rightarrow \{1, 2, \dots, q\}$ handelt, anstatt von Abbildungen $\mathbb{Z}/n\mathbb{Z} \rightarrow \{1, 2, \dots, q\}$. Diese Umschreibung (Satz 5.7 weiter unten) wird einfacher zu beweisen sein als Satz 5.6 **(a)**, und wir werden danach Satz 5.6 **(a)** aus ihr herleiten.

Wir haben bereits oben in Abschnitt 5.4 den Begriff einer k -periodischen Abbildung eingeführt. Wir führen jetzt zwei damit verbundene Begriffe ein:

Definition (periodische Abbildung): Sei Q eine Menge. Sei $f : \mathbb{Z} \rightarrow Q$ eine Abbildung. Die Abbildung f heiße *periodisch*, wenn es ein $k \in \mathbb{N}_+$ gibt, für das f eine k -periodische Abbildung ist.

Definition (Minimalperiode einer periodischen Abbildung): Sei Q eine Menge. Sei $f : \mathbb{Z} \rightarrow Q$ eine periodische Abbildung. Unter der *Minimalperiode* der Abbildung f verstehen wir dann die kleinste Zahl $k \in \mathbb{N}_+$, für die gilt, daß f eine k -periodische Abbildung ist.⁸⁶

Für jedes $k \in \mathbb{N}_+$ bezeichnen wir mit $Q_{\text{minper } k}^{\mathbb{Z}}$ die Menge aller periodischen Abbildungen von \mathbb{Z} nach Q , deren Minimalperiode gleich k ist.

Aus dieser Definition folgt natürlich, daß die Mengen $Q_{\text{minper } k}^{\mathbb{Z}}$ für verschiedene Werte von k paarweise disjunkt sind.

Nun unsere Umschreibung von Satz 5.6⁸⁷ (zumindest von Satz 5.6 **(a)**, doch Satz 5.6 **(b)** folgt recht schnell aus **(a)**):

Satz 5.7: Für alle $q \in \mathbb{N}$ und $n \in \mathbb{N}_+$ ist $|\{1, 2, \dots, q\}_{\text{minper } n}^{\mathbb{Z}}| = \sum_{d|n} \mu(d) q^{n/d}$.

Beweis von Satz 5.7: Erstmal sehen wir, daß für jedes $n \in \mathbb{N}_+$, jede Menge Q und jede Abbildung $f : \mathbb{Z} \rightarrow Q$ gilt:

Genau dann ist f eine n -periodische Abbildung, wenn folgendes gilt:
(die Abbildung f ist periodisch, und die Minimalperiode von f ist ein Teiler von n).
(60)

⁸⁶Solche Zahlen $k \in \mathbb{N}_+$ existieren wirklich, denn f ist eine periodische Abbildung.

⁸⁷Genauer es dazu im Beweis von Satz 5.6.

⁸⁸ Für jedes $n \in \mathbb{N}_+$ und jede Menge Q gilt also

$$\begin{aligned} & \{f \in Q^{\mathbb{Z}} \mid f \text{ ist eine } n\text{-periodische Abbildung}\} \\ &= \{f \in Q^{\mathbb{Z}} \mid \text{die Abbildung } f \text{ ist periodisch, und die Minimalperiode von } f \text{ ist ein Teiler von } n\} \\ &= \bigcup_{d \in \mathbb{N}_{|n}} \{f \in Q^{\mathbb{Z}} \mid \text{die Abbildung } f \text{ ist periodisch, und die Minimalperiode von } f \text{ ist } d\}. \end{aligned} \quad (61)$$

Doch für jedes n ist $\{f \in Q^{\mathbb{Z}} \mid f \text{ ist eine } n\text{-periodische Abbildung}\} = Q_{\text{per } n}^{\mathbb{Z}}$ (denn so wurde die Menge $Q_{\text{per } k}^{\mathbb{Z}}$ definiert), und für jedes d ist

$$\{f \in Q^{\mathbb{Z}} \mid \text{die Abbildung } f \text{ ist periodisch, und die Minimalperiode von } f \text{ ist } d\} = Q_{\text{minper } d}^{\mathbb{Z}} \quad (62)$$

(denn so wurde die Menge $Q_{\text{minper } k}^{\mathbb{Z}}$ definiert). Also wird (61) zu

$$Q_{\text{per } n}^{\mathbb{Z}} = \bigcup_{d \in \mathbb{N}_{|n}} Q_{\text{minper } d}^{\mathbb{Z}}.$$

Ist Q eine endliche Menge, so folgt hieraus

$$|Q_{\text{per } n}^{\mathbb{Z}}| = \sum_{d \in \mathbb{N}_{|n}} |Q_{\text{minper } d}^{\mathbb{Z}}|$$

(denn die Mengen $Q_{\text{minper } k}^{\mathbb{Z}}$ für verschiedene Werte von k sind paarweise disjunkt).

Wegen $|Q_{\text{per } n}^{\mathbb{Z}}| = |Q|^n$ (dies folgt aus (57), angewandt auf $k = n$) wird dies zu

$$|Q|^n = \sum_{d \in \mathbb{N}_{|n}} |Q_{\text{minper } d}^{\mathbb{Z}}|.$$

Sei nun q eine feste nichtnegative ganze Zahl. Wir setzen $Q = \{1, 2, \dots, q\}$, und definieren eine Zahlenfunktion $\tau : \mathbb{N}_+ \rightarrow \mathbb{Z}$ durch $(\tau(n) = |Q_{\text{minper } n}^{\mathbb{Z}}| \text{ für alle } n \in \mathbb{N}_+)$.

⁸⁸*Beweis von (60):* Um (60) nachzuweisen, müssen wir zwei Aussagen zeigen:

Aussage 3: Ist f eine n -periodische Abbildung, dann gilt: Die Abbildung f ist periodisch, und die Minimalperiode von f ist ein Teiler von n .

Aussage 4: Ist f eine periodische Abbildung, und ist die Minimalperiode von f ein Teiler von n , dann ist f eine n -periodische Abbildung.

Beweis von Aussage 3: Nehmen wir an, f sei eine n -periodische Abbildung. Dann ist f eine periodische Abbildung. Sei m die Minimalperiode von f . Das heißt, m ist die kleinste natürliche Zahl $k \in \mathbb{N}_+$, für die gilt, daß f eine k -periodische Abbildung ist. Insbesondere ist also f eine m -periodische Abbildung. Somit ist f gleichzeitig n -periodisch und m -periodisch. Gemäß (56) (angewandt auf n und m statt k und ℓ) ist f also $\text{ggT}(n, m)$ -periodisch. Daher ist $\text{ggT}(n, m) \geq m$ (denn die kleinste natürliche Zahl $k \in \mathbb{N}_+$, für die gilt, daß f eine k -periodische Abbildung ist, ist m). Zusammen mit $\text{ggT}(n, m) \mid m$ führt dies auf $\text{ggT}(n, m) = m$, und somit ist $m \mid n$. Das heißt, die Minimalperiode von f ist ein Teiler von n (denn die Minimalperiode von f ist m , und m ist ein Teiler von n). Damit ist Aussage 3 bewiesen.

Beweis von Aussage 4: Angenommen, die Abbildung f ist periodisch, und die Minimalperiode von f ist ein Teiler von n . Bezeichnen wir diese Minimalperiode mit m , dann ist also $m \mid n$. Andererseits ist f eine m -periodische Abbildung (denn m ist die Minimalperiode von f). Wegen $m \mid n$ folgt hieraus, daß f auch eine n -periodische Abbildung ist (nach (55), angewandt auf m und n statt k und k'). Damit ist Aussage 4 gezeigt.

Da nun beide Aussagen 3 und 4 nachgewiesen sind, ist der Beweis von (60) vollständig.

Wir erinnern uns an die Definition der Dirichlet-Faltung in Abschnitt 2.4. Für jedes $n \in \mathbb{N}_+$ ist dann

$$\begin{aligned} (\tau * \underline{1})(n) &= \sum_{d|n} \tau(d) \underbrace{\underline{1}\left(\frac{n}{d}\right)}_{=1} = \sum_{d|n} \tau(d) = \sum_{d \in \mathbb{N}_{|n}} \tau(d) = \sum_{d \in \mathbb{N}_{|n}} |Q_{\minper d}^{\mathbb{Z}}| \\ &\quad \left(\text{denn } \tau(d) = |Q_{\minper d}^{\mathbb{Z}}| \text{ nach der Definition der Zahlenfunktion } \tau \right) \\ &= |Q|^n = q^n \end{aligned}$$

(denn $|Q| = |\{1, 2, \dots, q\}| = q$). Wenn wir eine weitere Zahlenfunktion $\kappa : \mathbb{N}_+ \rightarrow \mathbb{Z}$ durch $(\kappa(n) = q^n$ für jedes $n \in \mathbb{N}_+$) definieren, dann lässt sich das umschreiben als $(\tau * \underline{1})(n) = \kappa(n)$. Da dies für alle $n \in \mathbb{N}_+$ gilt, ist also $\tau * \underline{1} = \kappa$. Daher ist

$$\begin{aligned} \kappa * \mu &= (\tau * \underline{1}) * \mu = \tau * (\underline{1} * \mu) && \text{(denn die Dirichlet-Faltung } * \text{ ist assoziativ)} \\ &= \tau * \varepsilon \end{aligned}$$

(denn da die Dirichlet-Faltung $*$ kommutativ ist, gilt $\underline{1} * \mu = \mu * \underline{1} = \varepsilon$ (nach Satz 2.4 **(b)**)). Wegen $\tau * \varepsilon = \tau$ wird dies zu $\kappa * \mu = \tau$. Somit ist $\mu * \kappa = \tau$ (denn $\kappa * \mu = \mu * \kappa$, da die Dirichlet-Faltung $*$ kommutativ ist). Für jedes $n \in \mathbb{N}_+$ ist also $(\mu * \kappa)(n) = \tau(n)$.

Wegen $\tau(n) = |Q_{\minper n}^{\mathbb{Z}}|$ und $(\mu * \kappa)(n) = \sum_{d|n} \mu(d) \kappa\left(\frac{n}{d}\right) = \sum_{d|n} \mu(d) q^{n/d}$ (denn $\kappa\left(\frac{n}{d}\right) = \kappa(n/d) = q^{n/d}$ nach der Definition von κ) wird dies zu $|Q_{\minper n}^{\mathbb{Z}}| = \sum_{d|n} \mu(d) q^{n/d}$. Wegen $Q = \{1, 2, \dots, q\}$ bedeutet dies $|\{1, 2, \dots, q\}_{\minper n}^{\mathbb{Z}}| = \sum_{d|n} \mu(d) q^{n/d}$, und damit ist Satz 5.7 bewiesen.

Nun zum *Beweis von Satz 5.6*: **(a)** Wir wollen Satz 5.6 **(a)** auf Satz 5.7 zurückführen. Sei Q die Menge $\{1, 2, \dots, q\}$. Wir setzen $F = \{1, 2, \dots, q\}^{\mathbb{Z}/n\mathbb{Z}} = Q^{\mathbb{Z}/n\mathbb{Z}}$, und wir führen für jedes $a \in \mathbb{Z}/n\mathbb{Z}$ die Abbildung $D_a : F \rightarrow F$ wie in Abschnitt 5.3 ein. Ferner wollen wir die in Abschnitt 5.4 definierten Abbildungen $R_{Q,k}$ und $R'_{Q,k}$ für jedes $k \in \mathbb{N}_+$ verwenden.

Für jede Abbildung $f : \mathbb{Z}/n\mathbb{Z} \rightarrow Q$ und jede Restklasse $a \in \mathbb{Z}/n\mathbb{Z}$ gilt dann folgende Äquivalenz von Aussagen⁸⁹:

$$\begin{aligned} &(f(x+a) = f(x) \text{ für alle } x \in \mathbb{Z}/n\mathbb{Z}) \\ &\iff ((D_a(f))(x) = f(x) \text{ für alle } x \in \mathbb{Z}/n\mathbb{Z}) \\ &\iff (D_a(f) = f) \iff (f \in \text{Fix } D_a). \end{aligned} \tag{63}$$

Hierbei ist der erste Äquivalenzpfeil trivial (denn $f(x+a) = (D_a(f))(x)$ nach der Definition von D_a), und der zweite ebenfalls (denn zwei Abbildungen sind genau dann einander gleich, wenn alle ihre Werte jeweils übereinstimmen); der dritte folgt aus der Definition von $\text{Fix } D_a$.

⁸⁹Weiter unten werden diese Äquivalenzpfeile auch genauer begründet.

Für jede Abbildung $f : \mathbb{Z}/n\mathbb{Z} \rightarrow Q$ gilt nun folgende Äquivalenz von Aussagen⁹⁰:

$$\begin{aligned}
& (f \in Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}) \\
& \iff (\text{die Abbildung } f \text{ ist aperiodisch}) \\
& \iff \left(\begin{array}{l} \text{es gibt keine von } \bar{0} \text{ verschiedene Restklasse } a \in \mathbb{Z}/n\mathbb{Z}, \text{ die} \\ (f(x+a) = f(x) \text{ für alle } x \in \mathbb{Z}/n\mathbb{Z}) \text{ erfüllt} \end{array} \right) \\
& \iff (\text{es gibt keine von } \bar{0} \text{ verschiedene Restklasse } a \in \mathbb{Z}/n\mathbb{Z}, \text{ die } f \in \text{Fix } D_a \text{ erfüllt}) \\
& \iff (\text{es gibt kein } a \in \{\bar{1}, \bar{2}, \dots, \overline{n-1}\}, \text{ das } f \in \text{Fix } D_a \text{ erfüllt}) \\
& \iff (\text{es gibt kein } i \in \{1, 2, \dots, n-1\}, \text{ das } f \in \text{Fix } D_{\bar{i}} \text{ erfüllt}) \\
& \iff \left(f \notin \bigcup_{i=1}^{n-1} \text{Fix } D_{\bar{i}} \right) \iff \left(f \in Q^{\mathbb{Z}/n\mathbb{Z}} \setminus \bigcup_{i=1}^{n-1} \text{Fix } D_{\bar{i}} \right). \tag{64}
\end{aligned}$$

Hierbei ist der erste Äquivalenzpfeil eine Konsequenz der Definition von $Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$; der zweite Äquivalenzpfeil ergibt sich aus der Definition des Begriffes "aperiodisch"; der dritte folgt aus (63); der vierte ist offensichtlich (denn die von $\bar{0}$ verschiedenen Restklassen $a \in \mathbb{Z}/n\mathbb{Z}$ sind genau die Restklassen $\bar{1}, \bar{2}, \dots, \overline{n-1}$); der fünfte ist völlig trivial (da haben wir einfach $a = \bar{i}$ substituiert); der sechste bedarf keiner Erklärung; der siebte folgt schließlich aus $f \in Q^{\mathbb{Z}/n\mathbb{Z}}$.

Aus der Äquivalenz (64) folgt nun $Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}} = Q^{\mathbb{Z}/n\mathbb{Z}} \setminus \bigcup_{i=1}^{n-1} \text{Fix } D_{\bar{i}}$. Mit der in Abschnitt 5.4 definierten Abbildung $R'_{Q,n}$ ist also

$$\begin{aligned}
R'_{Q,n} \left(Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}} \right) &= R'_{Q,n} \left(Q^{\mathbb{Z}/n\mathbb{Z}} \setminus \bigcup_{i=1}^{n-1} \text{Fix } D_{\bar{i}} \right) = R'_{Q,n} \left(Q^{\mathbb{Z}/n\mathbb{Z}} \right) \setminus \bigcup_{i=1}^{n-1} R'_{Q,n} \left(\text{Fix } D_{\bar{i}} \right) \\
& \quad \left(\text{denn } R'_{Q,n} \text{ ist eine Bijektion, da } R_{Q,n} \circ R'_{Q,n} = \text{id} \text{ und } R'_{Q,n} \circ R_{Q,n} = \text{id} \right) \\
&= Q_{\text{per } n}^{\mathbb{Z}} \setminus \bigcup_{i=1}^{n-1} Q_{\text{per ggT}(i,n)}^{\mathbb{Z}}
\end{aligned}$$

(denn $R'_{Q,n} \left(Q^{\mathbb{Z}/n\mathbb{Z}} \right) = Q_{\text{per } n}^{\mathbb{Z}}$ ⁹¹ und $R'_{Q,n} \left(\text{Fix } D_{\bar{i}} \right) = Q_{\text{per ggT}(i,n)}^{\mathbb{Z}}$ für jedes $i \in \{1, 2, \dots, n-1\}$ ⁹²). Wegen $\bigcup_{i=1}^{n-1} Q_{\text{per ggT}(i,n)}^{\mathbb{Z}} = \bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}}$ (dies folgt aus

⁹⁰Weiter unten werden diese Äquivalenzpfeile auch genauer begründet.

⁹¹denn die Abbildung $R'_{Q,n} : Q^{\mathbb{Z}/n\mathbb{Z}} \rightarrow Q_{\text{per } n}^{\mathbb{Z}}$ ist eine Bijektion, da $R_{Q,n} \circ R'_{Q,n} = \text{id}$ und $R'_{Q,n} \circ R_{Q,n} = \text{id}$

⁹²denn nach (59), angewandt auf i statt a , gilt $\text{Fix } D_{\bar{i}} = (R'_{Q,n})^{-1} \left(Q_{\text{per ggT}(i,n)}^{\mathbb{Z}} \right)$, und damit $R'_{Q,n} \left(\text{Fix } D_{\bar{i}} \right) = Q_{\text{per ggT}(i,n)}^{\mathbb{Z}}$ (weil $R'_{Q,n}$ eine Bijektion ist)

$$\bigcup_{i=1}^{n-1} Q_{\text{per ggT}(i,n)}^{\mathbb{Z}} \subseteq \bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}} \quad {}^{93} \text{ und } \bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}} \subseteq \bigcup_{i=1}^{n-1} Q_{\text{per ggT}(i,n)}^{\mathbb{Z}} \quad {}^{94}$$

wird dies zu

$$R'_{Q,n} \left(Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}} \right) = Q_{\text{per } n}^{\mathbb{Z}} \setminus \bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}}. \quad (65)$$

Doch wir wissen, daß

$$Q_{\text{per } n}^{\mathbb{Z}} = \bigcup_{d \in \mathbb{N}_{|n}} Q_{\text{minper } d}^{\mathbb{Z}} = \left(\bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}} \right) \cup Q_{\text{minper } n}^{\mathbb{Z}}$$

ist, und daß die Mengen $\bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}}$ und $Q_{\text{minper } n}^{\mathbb{Z}}$ disjunkt sind (denn

$$\begin{aligned} \left(\bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}} \right) \cap Q_{\text{minper } n}^{\mathbb{Z}} &= \bigcup_{d \in \mathbb{N}_{|n}; d < n} \underbrace{\left(Q_{\text{minper } d}^{\mathbb{Z}} \cap Q_{\text{minper } n}^{\mathbb{Z}} \right)}_{=\emptyset \text{ (denn die Mengen } Q_{\text{minper } d}^{\mathbb{Z}} \text{ und } Q_{\text{minper } n}^{\mathbb{Z}} \text{ sind disjunkt, weil die Mengen } Q_{\text{minper } k}^{\mathbb{Z}} \text{ für verschiedene Werte von } k \text{ paarweise disjunkt sind)}} = \bigcup_{d \in \mathbb{N}_{|n}; d < n} \emptyset = \emptyset \end{aligned}$$

⁹³*Beweis:* Sei f ein Element von $\bigcup_{i=1}^{n-1} Q_{\text{per ggT}(i,n)}^{\mathbb{Z}}$. Dann gibt es ein $i \in \{1, 2, \dots, n-1\}$ mit $f \in Q_{\text{per ggT}(i,n)}^{\mathbb{Z}}$. Die Abbildung f ist also eine $\text{ggT}(i, n)$ -periodische Abbildung. Doch nach (60) (angewandt auf $\text{ggT}(i, n)$ statt n) ist f genau dann eine $\text{ggT}(i, n)$ -periodische Abbildung, wenn gilt:

(die Abbildung f ist periodisch, und die Minimalperiode von f ist ein Teiler von $\text{ggT}(i, n)$).

Da f eine $\text{ggT}(i, n)$ -periodische Abbildung ist, folgern wir hieraus, daß die Abbildung f periodisch ist, und die Minimalperiode von f ein Teiler von $\text{ggT}(i, n)$ ist. Bezeichnen wir mit δ die Minimalperiode von f , dann ist also $\delta \mid \text{ggT}(i, n)$. Daraus folgt $\delta \in \mathbb{N}_{|n}$ (denn $\delta \mid \text{ggT}(i, n) \mid n$) und $\delta \mid \text{ggT}(i, n) \leq i \leq n-1 < n$. Damit ist $Q_{\text{minper } \delta}^{\mathbb{Z}} \subseteq \bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}}$. Wegen

$f \in Q_{\text{minper } \delta}^{\mathbb{Z}}$ (denn die Abbildung f ist periodisch, und die Minimalperiode von f ist δ) ist also $f \in \bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}}$.

Wir haben also für jedes Element f von $\bigcup_{i=1}^{n-1} Q_{\text{per ggT}(i,n)}^{\mathbb{Z}}$ gezeigt, daß $f \in \bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}}$

ist. Daraus folgt $\bigcup_{i=1}^{n-1} Q_{\text{per ggT}(i,n)}^{\mathbb{Z}} \subseteq \bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}}$, was zu beweisen war.

⁹⁴*Beweis:* Sei f ein Element von $\bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}}$. Dann gibt es ein $d \in \mathbb{N}_{|n}$ mit $d < n$, welches

$f \in Q_{\text{minper } d}^{\mathbb{Z}}$ erfüllt. Somit ist f eine periodische Abbildung, und die Minimalperiode von f ist d . Also ist f eine d -periodische Abbildung (denn d ist die Minimalperiode von f , also die kleinste Zahl $k \in \mathbb{N}_+$, für die gilt, daß f eine k -periodische Abbildung ist). Das heißt, $f \in Q_{\text{per } d}^{\mathbb{Z}}$. Wegen $d = \text{ggT}(d, n)$ (denn $d \mid n$) wird dies zu $f \in Q_{\text{per ggT}(d,n)}^{\mathbb{Z}}$. Doch wegen $d \in \{1, 2, \dots, n-1\}$

(denn $d \in \mathbb{N}_{|n}$ und $d < n$) ist $Q_{\text{per ggT}(d,n)}^{\mathbb{Z}} \subseteq \bigcup_{i=1}^{n-1} Q_{\text{per ggT}(i,n)}^{\mathbb{Z}}$. Aus $f \in Q_{\text{per ggT}(d,n)}^{\mathbb{Z}}$ folgt also

$$f \in \bigcup_{i=1}^{n-1} Q_{\text{per ggT}(i,n)}^{\mathbb{Z}}.$$

Wir haben also gezeigt: Für jedes Element f von $\bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}}$ gilt $f \in \bigcup_{i=1}^{n-1} Q_{\text{per ggT}(i,n)}^{\mathbb{Z}}$.

Hieraus folgt $\bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}} \subseteq \bigcup_{i=1}^{n-1} Q_{\text{per ggT}(i,n)}^{\mathbb{Z}}$, was zu beweisen war.

). Hieraus folgt

$$Q_{\text{per } n}^{\mathbb{Z}} \setminus \bigcup_{d \in \mathbb{N}_{|n}; d < n} Q_{\text{minper } d}^{\mathbb{Z}} = Q_{\text{minper } n}^{\mathbb{Z}}.$$

Somit wird (65) zu

$$R'_{Q,n} \left(Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}} \right) = Q_{\text{minper } n}^{\mathbb{Z}}.$$

Da $R'_{Q,n}$ eine Bijektion ist (denn $R_{Q,n} \circ R'_{Q,n} = \text{id}$ und $R'_{Q,n} \circ R_{Q,n} = \text{id}$), folgt hieraus $|Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}| = |Q_{\text{minper } n}^{\mathbb{Z}}|$. Doch wegen $\{1, 2, \dots, q\} = Q$ ist

$$|\{1, 2, \dots, q\}_{\text{minper } n}^{\mathbb{Z}}| = |Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}| = |Q_{\text{minper } n}^{\mathbb{Z}}| = |\{1, 2, \dots, q\}_{\text{minper } n}^{\mathbb{Z}}| = \sum_{d|n} \mu(d) q^{n/d}$$

(nach Satz 5.7), und damit ist Satz 5.6 **(a)** gezeigt.

(b) Sei ρ die Anzahl der Äquivalenzklassen bezüglich der Relation $\overset{\text{neck}}{\sim}_{\text{aper}}$, in die die Menge $\{1, 2, \dots, q\}_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$ zerfällt. Wir bezeichnen diese Äquivalenzklassen mit $\ddot{A}_1, \ddot{A}_2, \dots, \ddot{A}_\rho$ (wobei jede Äquivalenzklasse genau einmal in der Liste $(\ddot{A}_1, \ddot{A}_2, \dots, \ddot{A}_\rho)$ vorkommen soll).

Wir wollen zuerst einmal beweisen: Für jedes $i \in \{1, 2, \dots, \rho\}$ ist $|\ddot{A}_i| = n$.

In der Tat setzen wir $Q = \{1, 2, \dots, q\}$. Wie wir im Beweis von Satz 5.6 **(a)** gezeigt haben, gilt $Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}} = Q^{\mathbb{Z}/n\mathbb{Z}} \setminus \bigcup_{i=1}^{n-1} \text{Fix } D_{\bar{i}}$. Wegen

$$\begin{aligned} \bigcup_{i=1}^{n-1} \text{Fix } D_{\bar{i}} &= \bigcup_{i \in \{1, 2, \dots, n-1\}} \text{Fix } D_{\bar{i}} = \bigcup_{b \in \{\bar{1}, \bar{2}, \dots, \overline{n-1}\}} \text{Fix } D_b \quad (\text{hier haben wir } b \text{ für } \bar{i} \text{ substituiert}) \\ &= \bigcup_{b \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}} \text{Fix } D_b \quad (\text{denn } \{\bar{1}, \bar{2}, \dots, \overline{n-1}\} = (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}) \end{aligned}$$

wird dies zu $Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}} = Q^{\mathbb{Z}/n\mathbb{Z}} \setminus \bigcup_{b \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}} \text{Fix } D_b$.

Nun seien f und g zwei Elemente von $Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$. Dann gilt folgende Kette von Äquivalenzen:

$$\left(f \overset{\text{neck}}{\sim}_{\text{aper}} g \right) \iff \left(f \overset{\text{neck}}{\sim} g \right) \iff (\text{es gibt ein } a \in \mathbb{Z}/n\mathbb{Z} \text{ mit } D_a(f) = g). \quad (66)$$

Hierbei folgt der erste Äquivalenzpfeil aus der Definition von $\overset{\text{neck}}{\sim}_{\text{aper}}$ (nämlich haben wir die Relation $\overset{\text{neck}}{\sim}_{\text{aper}}$ als die Einschränkung der Relation $\overset{\text{neck}}{\sim}$ auf die Teilmenge $\{1, 2, \dots, q\}_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}} = Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$ definiert), und der zweite Äquivalenzpfeil folgt aus (51).

Für jedes $f \in Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$ gilt nun:

$$\begin{aligned} & \left(\text{Äquivalenzklasse von } f \text{ bezüglich der Relation } \overset{\text{neck}}{\sim}_{\text{aper}} \right) \\ &= \left\{ g \in Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}} \mid f \overset{\text{neck}}{\sim}_{\text{aper}} g \right\} \\ &= \left\{ g \in Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}} \mid \text{es gibt ein } a \in \mathbb{Z}/n\mathbb{Z} \text{ mit } D_a(f) = g \right\} \quad (\text{nach (66)}) \\ &= \{ D_a(f) \mid a \in \mathbb{Z}/n\mathbb{Z} \} \quad (67) \end{aligned}$$

(denn für jedes $a \in \mathbb{Z}/n\mathbb{Z}$ ist $D_a(f) \in Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$ ⁹⁵). Für verschiedene Elemente $a \in \mathbb{Z}/n\mathbb{Z}$ sind auch die Abbildungen $D_a(f)$ verschieden⁹⁶; hieraus folgt

$$|\{D_a(f) \mid a \in \mathbb{Z}/n\mathbb{Z}\}| = |\mathbb{Z}/n\mathbb{Z}| = n.$$

Wegen (67) wird dies zu

$$\left| \text{Äquivalenzklasse von } f \text{ bezüglich der Relation } \underset{\text{aper}}{\overset{\text{neck}}{\sim}} \right| = n. \quad (68)$$

Hieraus folgt $|\ddot{A}_i| = n$ für jedes $i \in \{1, 2, \dots, \rho\}$ (denn ist f ein Element der Äquivalenzklasse \ddot{A}_i , dann ist \ddot{A}_i die Äquivalenzklasse von f bezüglich der Relation $\underset{\text{aper}}{\overset{\text{neck}}{\sim}}$, und aus (68) folgt nun $|\ddot{A}_i| = n$). Da die Äquivalenzklassen $\ddot{A}_1, \ddot{A}_2, \dots, \ddot{A}_\rho$ paarweise disjunkt sind (weil verschiedene Äquivalenzklassen stets disjunkt sind), ist nun $|\ddot{A}_1 \cup \ddot{A}_2 \cup \dots \cup \ddot{A}_\rho| = \underbrace{|\ddot{A}_1| + |\ddot{A}_2| + \dots + |\ddot{A}_\rho|}_{\rho \text{ mal } n} = \rho n$. Doch wegen $\ddot{A}_1 \cup \ddot{A}_2 \cup \dots \cup \ddot{A}_\rho = \{1, 2, \dots, q\}_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$ (denn $\ddot{A}_1, \ddot{A}_2, \dots, \ddot{A}_\rho$ sind alle Äquivalenzklassen bezüglich der Relation $\underset{\text{aper}}{\overset{\text{neck}}{\sim}}$, in die die Menge $\{1, 2, \dots, q\}_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$ zerfällt) ist $|\ddot{A}_1 \cup \ddot{A}_2 \cup \dots \cup \ddot{A}_\rho| = |\{1, 2, \dots, q\}_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}| = \sum_{d|n} \mu(d) q^{n/d}$ (nach Satz 5.6 (a)), und damit wird dies zu $\sum_{d|n} \mu(d) q^{n/d} = \rho n$, also zu

⁹⁵*Beweis:* Wegen $f \in Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}} = Q^{\mathbb{Z}/n\mathbb{Z}} \setminus \bigcup_{b \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}} \text{Fix } D_b$ ist $f \notin \bigcup_{b \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}} \text{Fix } D_b$. Für jedes

$b \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}$ ist also $f \notin \text{Fix } D_b$, also (mit anderen Worten) $D_b(f) \neq f$, und damit auch $D_a(D_b(f)) \neq D_a(f)$ (denn D_a ist eine bijektive Abbildung, also $D_b(D_a(f)) \neq D_a(f)$ (denn

$$D_b(D_a(f)) = \underbrace{(D_b \circ D_a)}_{=D_{b+a} \text{ (nach (46))}}(f) = \underbrace{D_{b+a}}_{=D_{a+b}}(f) = \underbrace{D_{a+b}}_{=D_a \circ D_b \text{ (nach (46))}}(f) = (D_a \circ D_b)(f) = D_a(D_b(f))$$

), und damit $D_a(f) \notin \text{Fix } D_b$. Da dies für alle $b \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}$ gilt, ist also $D_a(f) \notin \bigcup_{b \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}} \text{Fix } D_b$. Das heißt, $D_a(f) \in Q^{\mathbb{Z}/n\mathbb{Z}} \setminus \bigcup_{b \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}} \text{Fix } D_b = Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$, was zu beweisen war.

⁹⁶*Beweis:* Seien a und A zwei verschiedene Elemente von $\mathbb{Z}/n\mathbb{Z}$. Wir müssen dann beweisen, daß auch die Abbildungen $D_a(f)$ und $D_A(f)$ verschieden sind. Dazu bemerken wir, daß

$$\begin{aligned} D_A \circ D_{a-A} &= D_{(a-A)+A} && \text{(nach (46))} \\ &= D_a \end{aligned}$$

gilt, und wir

$$D_A(D_{a-A}(f)) = \left(\underbrace{D_A \circ D_{a-A}}_{=D_a} \right)(f) = D_a(f)$$

haben. Doch aus $f \in Q_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}} = Q^{\mathbb{Z}/n\mathbb{Z}} \setminus \bigcup_{b \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}} \text{Fix } D_b$ folgt $f \notin \bigcup_{b \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}} \text{Fix } D_b$, und

daher $f \notin \text{Fix } D_b$ für alle $b \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}$. Angewandt auf $b = a - A$ ergibt dies $f \notin \text{Fix } D_{a-A}$ (denn $a - A \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{\bar{0}\}$, denn $a \neq A$ liefert $a - A \neq 0$), also $D_{a-A}(f) \neq f$. Da D_A bijektiv ist, folgt hieraus $D_A(D_{a-A}(f)) \neq D_A(f)$. Wegen $D_A(D_{a-A}(f)) = D_a(f)$ bedeutet dies $D_a(f) \neq D_A(f)$. Folglich sind die Abbildungen $D_a(f)$ und $D_A(f)$ verschieden, was zu beweisen war.

$\rho = \frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}$. Da ρ definiert war als die Anzahl der Äquivalenzklassen bezüglich der Relation $\overset{\text{neck}}{\underset{\text{aper}}{\sim}}$, in die die Menge $\{1, 2, \dots, q\}_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$ zerfällt, heißt dies also: Die Menge $\{1, 2, \dots, q\}_{\text{aper}}^{\mathbb{Z}/n\mathbb{Z}}$ zerfällt in genau $\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}$ Äquivalenzklassen bezüglich der Relation $\overset{\text{neck}}{\underset{\text{aper}}{\sim}}$. Damit ist Satz 5.6 (b) bewiesen.

[...]

ab hier BAUSTELLE

- * mu und aperiodische perlenketten
- * allgemeine fragestellung
- * daraus 2.4
- * beweis
- * gewichtete fragestellung
- * daraus 4.1
- * beweis
- * hinweis: das ist der engel-beweis
- * bracelets

6. Der negative Fall und die Kombinatorik

Wie wir am Ende von Abschnitt 5.4 gesehen haben, folgt aus Satz 5.1 "beinahe" Satz 1.1, denn Anzahlen sind stets ganze Zahlen. Das einzige, was diesen Beweis behindert, ist die Tatsache, daß $q \in \mathbb{N}$ gelten muss, damit wir Satz 5.1 anwenden können, und somit der Fall $q < 0$ nicht durch diesen Beweis abgedeckt wird. Doch diese Beweislücke ist nicht schwer zu flicken: Wir müssen den Fall $q < 0$ auf den Fall $q \geq 0$ zurückführen. Mit anderen Worten müssen wir folgendes Lemma zeigen:

Lemma 6.1: Wenn Satz 1.1 für alle $q \in \mathbb{N}$ gilt, dann gilt Satz 1.1 auch allgemein (das heißt, für alle $q \in \mathbb{Z}$).

Wir können dieses Lemma auf zwei verschiedene Weisen beweisen: Die erste ist sehr kurz und einfach, aber nicht so einfach auf andere Situationen (wie Satz 4.2) verallgemeinerbar; hingegen ist die zweite nicht so kurz, aber trivial zu verallgemeinern (sie zeigt allgemein, daß ein Polynom, das an allen nichtnegativen ganzen Zahlen ganze Werte annimmt, auch an den negativen ganzen Zahlen ebenfalls ganze Werte annehmen muss).

Erster Beweis von Lemma 6.1: Wir nehmen an, daß Satz 1.1 für alle $q \in \mathbb{N}$ gilt. Nun wollen wir zeigen, daß Satz 1.1 auch allgemein, also für alle $q \in \mathbb{Z}$ gilt. Seien $q \in \mathbb{Z}$ und $n \in \mathbb{N}_+$ beliebig gewählt. Sei r der Rest, den q bei Division durch n lässt. Dann ist $r \in \mathbb{N}$ und $q \equiv r \pmod{n}$. Wenden wir Satz 1.1 auf r statt q an (dies dürfen wir tun, denn $r \in \mathbb{N}$, und wir haben ja angenommen, daß Satz 1.1 für alle $q \in \mathbb{N}$ gilt), so erhalten wir $\frac{1}{n} \sum_{d|n} \phi(d) r^{n/d} \in \mathbb{Z}$. Das heißt, $\sum_{d|n} \phi(d) r^{n/d} \in n\mathbb{Z}$, also $\sum_{d|n} \phi(d) r^{n/d} \equiv 0 \pmod{n}$. Wegen $q \equiv r \pmod{n}$ folgt hieraus $\sum_{d|n} \phi(d) q^{n/d} \equiv 0 \pmod{n}$, also $\sum_{d|n} \phi(d) q^{n/d} \in n\mathbb{Z}$,

und damit $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d} \in \mathbb{Z}$. Damit ist Satz 1.1 bewiesen (und zwar im allgemeinen Fall, also im Fall einer beliebigen ganzen Zahl $q \in \mathbb{Z}$). Also ist Lemma 6.1 gezeigt.

Zweiter Beweis von Lemma 6.1: Unser zweiter Beweis von Lemma 6.1 benötigt folgendes Hilfsresultat:

Satz 6.2: Sei n eine ganze Zahl. Sei P ein Polynom mit rationalen Koeffizienten in der Variablen X . Angenommen, $\deg P < n$. Sei $\alpha \in \mathbb{Z}$ eine fest gewählte ganze Zahl. Dann sind folgende drei Aussagen \mathcal{P}_1 , \mathcal{P}_2 und \mathcal{P}_3 äquivalent:

Aussage \mathcal{P}_1 : Es gilt $P(q) \in \mathbb{Z}$ für alle $q \in \mathbb{Z}$.

Aussage \mathcal{P}_2 : Es gilt $P(q) \in \mathbb{Z}$ für alle $q \in \{\alpha, \alpha + 1, \dots, \alpha + n - 1\}$.

Aussage \mathcal{P}_3 : Es gibt ganze Zahlen a_0, a_1, \dots, a_{n-1} , so daß

$$P(X) = \sum_{k=0}^{n-1} a_k \binom{X}{k} \quad (69)$$

gilt. Dabei bezeichnet $\binom{X}{k}$ das Polynom $\frac{X(X-1)\dots(X-k+1)}{k!}$, und die Gleichung (69) ist als eine Gleichung zwischen Polynomen zu verstehen.

Dieser Satz ist im Übrigen nicht nur für uns hier von Bedeutung; so taucht die Äquivalenz der Aussagen \mathcal{P}_1 und \mathcal{P}_3 als Proposition 7.3 (a) in [15] auf, wo sie zur Konstruktion des Hilbertpolynoms verwendet wird.

Beweis von Satz 6.2: Es ist klar, daß $\mathcal{P}_1 \implies \mathcal{P}_2$ und $\mathcal{P}_3 \implies \mathcal{P}_1$ ist⁹⁷. Um die Äquivalenz der Aussagen \mathcal{P}_1 , \mathcal{P}_2 und \mathcal{P}_3 nachzuweisen, müssen wir also nur noch zeigen, daß $\mathcal{P}_2 \implies \mathcal{P}_3$ ist.

Beweis von $\mathcal{P}_2 \implies \mathcal{P}_3$: Wir beweisen nun die Implikation $\mathcal{P}_2 \implies \mathcal{P}_3$. Dazu nehmen wir an, daß Aussage \mathcal{P}_2 erfüllt ist, und versuchen Aussage \mathcal{P}_3 zu beweisen. Und zwar definieren wir ganze Zahlen a_0, a_1, \dots, a_{n-1} durch

$$\left(a_k = \sum_{\ell=0}^k (-1)^{k-\ell} \binom{k}{\ell} P(\alpha + \ell) \quad \text{für alle } k \in \{0, 1, \dots, n-1\} \right).$$

Wir wollen nun zeigen, daß diese ganzen Zahlen die Gleichung (69) erfüllen.

In der Tat erfüllen diese ganzen Zahlen für jedes $x \in \{0, 1, \dots, n-1\}$ folgendes:

$$\begin{aligned} \sum_{k=0}^{n-1} a_k \binom{x}{k} &= \sum_{k=0}^{n-1} \underbrace{\sum_{\ell=0}^k (-1)^{k-\ell} \binom{k}{\ell} P(\alpha + \ell)}_{= \sum_{\ell=0}^{n-1} \sum_{k=\ell}^{n-1}} \binom{x}{k} && \text{(nach der Definition von } a_k) \\ &= \sum_{\ell=0}^{n-1} \sum_{k=\ell}^{n-1} (-1)^{k-\ell} \binom{k}{\ell} P(\alpha + \ell) \binom{x}{k} = \sum_{\ell=0}^{n-1} P(\alpha + \ell) \sum_{k=\ell}^{n-1} (-1)^{k-\ell} \binom{k}{\ell} \binom{x}{k}. \end{aligned} \quad (70)$$

⁹⁷Zum Beweis von $\mathcal{P}_3 \implies \mathcal{P}_1$ benötigt man Satz 3.3 (b).

Nun ist aber

$$\binom{k}{\ell} \binom{x}{k} = \binom{x}{\ell} \binom{x-\ell}{k-\ell} \quad \text{für beliebige ganze } x, k \text{ und } \ell \text{ mit } k \geq \ell \geq 0 \quad (71)$$

⁹⁸. Somit gilt für jedes natürliche ℓ folgendes:

$$\begin{aligned} \sum_{k=\ell}^{n-1} (-1)^{k-\ell} \binom{k}{\ell} \binom{x}{k} &= \sum_{k=\ell}^{n-1} (-1)^{k-\ell} \binom{x}{\ell} \binom{x-\ell}{k-\ell} && \text{(nach (71))} \\ &= \sum_{u=0}^{n-1-\ell} (-1)^u \binom{x}{\ell} \binom{x-\ell}{u} && \text{(hier haben wir } k-\ell \text{ in der Summe durch } u \text{ substituiert)} \\ &= \binom{x}{\ell} \sum_{u=0}^{n-1-\ell} (-1)^u \binom{x-\ell}{u}. && (72) \end{aligned}$$

Wir wollen nun zeigen, daß für jedes natürliche ℓ gilt:

$$\binom{x}{\ell} \sum_{u=0}^{n-1-\ell} (-1)^u \binom{x-\ell}{u} = \begin{cases} 1, & \text{wenn } x = \ell; \\ 0, & \text{wenn } x \neq \ell \end{cases}. \quad (73)$$

Um dies zu beweisen, unterscheiden wir zwei Fälle: den Fall $x < \ell$, und den Fall $x \geq \ell$. Im Fall $x < \ell$ ist (73) offensichtlich (denn im Fall $x < \ell$ ist $\binom{x}{\ell} = 0$ (weil $x < \ell$ und $x \geq 0$) und $\begin{cases} 1, & \text{wenn } x = \ell; \\ 0, & \text{wenn } x \neq \ell \end{cases} = 0$ (denn aus $x < \ell$ folgt $x \neq \ell$), und somit sind beide Seiten der Gleichung (73) gleich Null, weshalb diese Gleichung offensichtlich gelten

⁹⁸denn aus $\binom{k}{\ell} = \frac{k(k-1)\dots(k-\ell+1)}{\ell!}$ und $\binom{x}{k} = \frac{x(x-1)\dots(x-k+1)}{k!}$ folgt

$$\begin{aligned} \binom{k}{\ell} \binom{x}{k} &= \frac{k(k-1)\dots(k-\ell+1)}{\ell!} \cdot \frac{x(x-1)\dots(x-k+1)}{k!} \\ &= \frac{x(x-1)\dots(x-k+1)}{\ell!} \cdot \frac{k!}{\underbrace{k(k-1)\dots(k-\ell+1)}} \\ &= \frac{x(x-1)\dots(x-k+1)}{\ell!} \cdot \frac{k(k-1)\dots 1}{k(k-1)\dots(k-\ell+1)} = \frac{x(x-1)\dots(x-k+1)}{\ell!} \cdot \frac{1}{(k-\ell)!} \\ &= \frac{x(x-1)\dots(x-k+1)}{\ell!} \cdot \frac{\underbrace{(x(x-1)\dots(x-\ell+1)) \cdot \underbrace{((x-\ell)(x-\ell-1)\dots(x-k+1))}}_{(k-\ell)!}}{\ell! (k-\ell)!} \\ &= \underbrace{\frac{x(x-1)\dots(x-\ell+1)}{\ell!}}_{=\binom{x}{\ell}} \cdot \frac{\underbrace{(x-\ell)(x-\ell-1)\dots((x-\ell)-(k-\ell)+1)}}_{(k-\ell)!}}{\ell! (k-\ell)!} = \binom{x}{\ell} \binom{x-\ell}{k-\ell} \end{aligned}$$

muss). Im Fall $x \geq \ell$ ist $x - \ell \geq 0$, und somit folgt (73) in diesem Fall aus

$$\begin{aligned}
\binom{x}{\ell} \sum_{u=0}^{n-1-\ell} (-1)^u \binom{x-\ell}{u} &= \binom{x}{\ell} \left(\sum_{u=0}^{x-\ell} (-1)^u \binom{x-\ell}{u} + \underbrace{\sum_{u=x-\ell+1}^{n-1-\ell} (-1)^u \binom{x-\ell}{u}}_{=0 \text{ (denn } x-\ell \geq 0 \text{ und } u > x-\ell)} \right) \\
&\left(\begin{array}{l} \text{denn wegen } x \in \{0, 1, \dots, n-1\} \text{ ist } x \leq n-1, \text{ also } x-\ell \leq n-1-\ell \\ \text{und damit } \sum_{u=0}^{n-1-\ell} (-1)^u \binom{x-\ell}{u} = \sum_{u=0}^{x-\ell} (-1)^u \binom{x-\ell}{u} + \sum_{u=x-\ell+1}^{n-1-\ell} (-1)^u \binom{x-\ell}{u} \end{array} \right) \\
&= \binom{x}{\ell} \left(\underbrace{\sum_{u=0}^{x-\ell} (-1)^u \binom{x-\ell}{u}}_{=(1+(-1))^{x-\ell} \text{ (nach der binomischen Formel)}} + \underbrace{\sum_{u=x-\ell+1}^{n-1-\ell} (-1)^u 0}_{=0} \right) = \binom{x}{\ell} ((1+(-1))^{x-\ell} + 0) \\
&= \binom{x}{\ell} (1+(-1))^{x-\ell} = \binom{x}{\ell} 0^{x-\ell} = \binom{x}{\ell} \begin{cases} 1, & \text{wenn } x = \ell; \\ 0, & \text{wenn } x \neq \ell \end{cases} \\
&\left(\text{denn wegen } x - \ell \geq 0 \text{ ist } 0^{x-\ell} = \begin{cases} 1, & \text{wenn } x - \ell = 0; \\ 0, & \text{wenn } x - \ell \neq 0 \end{cases} = \begin{cases} 1, & \text{wenn } x = \ell; \\ 0, & \text{wenn } x \neq \ell \end{cases} \right) \\
&= \begin{cases} \binom{x}{\ell} \cdot 1, & \text{wenn } x = \ell; \\ \binom{x}{\ell} \cdot 0, & \text{wenn } x \neq \ell \end{cases} = \begin{cases} 1, & \text{wenn } x = \ell; \\ 0, & \text{wenn } x \neq \ell \end{cases} \\
&\left(\text{denn im Falle von } x = \ell \text{ ist } \binom{x}{\ell} \cdot 1 = \binom{x}{\ell} = \binom{\ell}{\ell} = 1, \text{ und im Falle von } x \neq \ell \text{ ist } \binom{x}{\ell} \cdot 0 = 0 \right).
\end{aligned}$$

Somit ist die Gleichung (73) in beiden Fällen $x < \ell$ und $x \geq \ell$ bewiesen. Damit ist der Beweis von (73) vollständig.

Aus (72) und (73) folgt

$$\sum_{k=\ell}^{n-1} (-1)^{k-\ell} \binom{k}{\ell} \binom{x}{k} = \begin{cases} 1, & \text{wenn } x = \ell; \\ 0, & \text{wenn } x \neq \ell \end{cases}.$$

für jedes natürliche ℓ . Somit wird (70) zu

$$\begin{aligned}
\sum_{k=0}^{n-1} a_k \binom{x}{k} &= \sum_{\ell=0}^{n-1} P(\alpha + \ell) \sum_{k=\ell}^{n-1} (-1)^{k-\ell} \binom{k}{\ell} \binom{x}{k} = \sum_{\ell=0}^{n-1} P(\alpha + \ell) \begin{cases} 1, & \text{wenn } x = \ell; \\ 0, & \text{wenn } x \neq \ell \end{cases} \\
&= \sum_{\ell \in \{0, 1, \dots, n-1\}} P(\alpha + \ell) \begin{cases} 1, & \text{wenn } x = \ell; \\ 0, & \text{wenn } x \neq \ell \end{cases} \\
&= \sum_{\ell \in \{0, 1, \dots, n-1\}; x=\ell} P(\alpha + \ell) \underbrace{\begin{cases} 1, & \text{wenn } x = \ell; \\ 0, & \text{wenn } x \neq \ell \end{cases}}_{=1 \text{ (da } x=\ell)} + \sum_{\ell \in \{0, 1, \dots, n-1\}; x \neq \ell} P(\alpha + \ell) \underbrace{\begin{cases} 1, & \text{wenn } x = \ell; \\ 0, & \text{wenn } x \neq \ell \end{cases}}_{=0 \text{ (da } x \neq \ell)} \\
&= \underbrace{\sum_{\ell \in \{0, 1, \dots, n-1\}; x=\ell} P(\alpha + \ell) 1}_{=P(\alpha+x)1 \text{ (denn } x \in \{0, 1, \dots, n-1\})} + \underbrace{\sum_{\ell \in \{0, 1, \dots, n-1\}; x \neq \ell} P(\alpha + \ell) 0}_{=0} = P(\alpha + x) 1 + 0 = P(\alpha + x).
\end{aligned}$$

Mit anderen Worten:

$$\sum_{k=0}^{n-1} a_k \binom{x}{k} - P(\alpha + x) = 0.$$

Das heißt, das Polynom $\sum_{k=0}^{n-1} a_k \binom{X}{k} - P(\alpha + X) \in \mathbb{Q}[X]$ hat x als Nullstelle. Da dies für alle $x \in \{0, 1, \dots, n-1\}$ gilt, hat dieses Polynom $\sum_{k=0}^{n-1} a_k \binom{X}{k} - P(\alpha + X)$ also die Zahlen $0, 1, \dots, n-1$ als Nullstellen. Doch der Grad des Polynoms $\sum_{k=0}^{n-1} a_k \binom{X}{k} - P(\alpha + X)$ ist $< n$ ⁹⁹. Ein Polynom von Grad $< n$, welches mindestens n paarweise verschiedene Nullstellen hat, muss bekanntermaßen identisch Null sein. Da unser Polynom $\sum_{k=0}^{n-1} a_k \binom{X}{k} - P(\alpha + X)$ einen Grad $< n$ hat und mindestens n paarweise verschiedene Nullstellen hat (nämlich die n Nullstellen $0, 1, \dots, n-1$), muss dieses Polynom also identisch Null sein.

Wir haben also gezeigt: Das Polynom $\sum_{k=0}^{n-1} a_k \binom{X}{k} - P(\alpha + X)$ ist identisch Null.

Mit anderen Worten: $P(\alpha + X) = \sum_{k=0}^{n-1} a_k \binom{X}{k}$. Damit ist gezeigt, daß unsere ganzen Zahlen a_0, a_1, \dots, a_{n-1} die Gleichung (69) erfüllen. Somit ist Aussage \mathcal{P}_3 gültig. Wir haben damit die Implikation $\mathcal{P}_2 \implies \mathcal{P}_3$ nachgewiesen. Wie schon gesagt, folgt daraus sofort die Äquivalenz aller drei Aussagen $\mathcal{P}_1, \mathcal{P}_2$ und \mathcal{P}_3 . Somit ist Satz 6.2 bewiesen.

(Ein anderer Beweis von Satz 6.2, genauer gesagt von der Implikation $\mathcal{P}_2 \implies \mathcal{P}_3$, läßt sich durch Induktion nach n führen.)

[...]

* problem: burnside-beweis geht nur für $q \geq 0$ - WARUM kombinatorischer beweis schwer wäre (\neq unmöglich)

* binoeffizienten auch

* die billige methode: modulo n geht nicht mehr für binkühe, aber 6.2 geht

⁹⁹denn

$$\deg \left(\sum_{k=0}^{n-1} a_k \binom{X}{k} \right) \leq \max \left\{ \underbrace{\deg \binom{X}{k}}_{=k} \mid k \in \{0, 1, \dots, n-1\} \right\} = \max \{k \mid k \in \{0, 1, \dots, n-1\}\} = n-1 < n$$

und somit

$$\deg \left(\sum_{k=0}^{n-1} a_k \binom{X}{k} - P(\alpha + X) \right) \leq \max \left\{ \underbrace{\deg \left(\sum_{k=0}^{n-1} a_k \binom{X}{k} \right)}_{< n}, \underbrace{\deg (P(\alpha + X))}_{=\deg P < n} \right\} < n$$

7. Kombinatorik II: Nester und Fixpunktzahlen

Satz 2.1 handelt von unendlichen Folgen. Zwar waren bislang alle unsere Anwendungen (die Folge $(b_1, b_2, b_3, \dots) = (q^1, q^2, q^3, \dots)$ in Abschnitt 2, und zwei weitere Folgen in Abschnitt 4) unendliche Folgen, doch die Erfahrung in der Mathematik zeigt, daß oft endliche Folgen mehr und bessere Eigenschaften besitzen als unendliche Folgen. Man könnte nun versuchen, Satz 2.1 auf endliche Folgen zu übertragen (also (b_1, b_2, \dots, b_ν) statt (b_1, b_2, b_3, \dots) , und (x_1, x_2, \dots, x_ν) statt (x_1, x_2, x_3, \dots) , und so weiter, wobei ν eine feste natürliche Zahl ist). Tatsächlich ist aber ein wenig mehr möglich: Statt mit Folgen, die mit positiven Zahlen Zahlen indiziert sind, kann man in Satz 2.1 auch mit Familien arbeiten, die mit den Elementen eines *Nests* indiziert sind. Hier die notwendige Definition:

Definition (Nest): Unter einem *Nest* verstehen wir eine nichtleere Teilmenge N der Menge \mathbb{N}_+ , die die Eigenschaft hat, daß für jedes Element $n \in N$ auch jeder positive Teiler von n in N liegt.

Es ist klar, daß jedes Nest das Element 1 enthält¹⁰⁰. Beispiele für Nester sind die Mengen $\{1\}$, \mathbb{N}_+ , $\{1, 2, \dots, \nu\}$ (wobei ν eine natürliche Zahl ist), $\{1, p, p^2, p^3, \dots\}$ (wobei p eine Primzahl ist), \mathbb{N}_n (dies ist die Menge aller positiven Teiler von n , wobei n eine natürliche Zahl ist) und $\{1, 2, 3, 4, 6, 7, 9\}$.

Und hier ist die Verallgemeinerung von Satz 2.1 und Satz 4.7 auf Familien, die mit Nestelementen indiziert sind:

Satz 7.1 (der Äquivalenzsatz für vernestete Geisterwitfolgen): Sei N ein Nest, und $(b_n)_{n \in N} \in \mathbb{Z}^N$ eine Familie ganzer Zahlen.¹⁰¹ Dann sind folgende Aussagen \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} , \mathcal{H} , \mathcal{I} und \mathcal{J} äquivalent:

Aussage \mathcal{A} : Für jede Zahl $n \in N$ und jeden Primteiler p von n gilt

$$b_{n/p} \equiv b_n \pmod{p^{v_p(n)}}.$$

Aussage \mathcal{B} : Es gibt eine Familie $(x_n)_{n \in N} \in \mathbb{Z}^N$ ganzer Zahlen, die

$$\left(b_n = \sum_{d|n} dx_d^{n/d} \text{ für jedes } n \in N \right) \quad (74)$$

erfüllt.

Aussage \mathcal{C} : Es gibt *genau eine* Familie $(x_n)_{n \in N} \in \mathbb{Z}^N$ ganzer Zahlen, die (74) erfüllt.

¹⁰⁰denn ist N ein Nest, dann ist N nichtleer, und hat somit ein Element, und da 1 notwendigerweise ein Teiler dieses Elementes ist, muß also 1 in N liegen

¹⁰¹*Beispiele:* Im Falle von $N = \mathbb{N}_+$ ist so eine Familie $(b_n)_{n \in N} \in \mathbb{Z}^{\mathbb{N}_+}$ nichts anderes als eine Folge $(b_1, b_2, b_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ ganzer Zahlen. Im Falle von $N = \{1, 2, \dots, \nu\}$ (wobei ν eine natürliche Zahl ist) ist so eine Familie $(b_n)_{n \in N} \in \mathbb{Z}^N$ ein ν -Tupel $(b_1, b_2, \dots, b_\nu) \in \mathbb{Z}^\nu$ ganzer Zahlen. Im Falle von $N = \{1, 2, 3, 4, 6, 7, 9\}$ ist so eine Familie $(b_n)_{n \in N} \in \mathbb{Z}^N$ ein 7-Tupel $(b_1, b_2, b_3, b_4, b_6, b_7, b_9)$ ganzer Zahlen.

Aussage \mathcal{D} : Es gibt eine Familie $(y_n)_{n \in N} \in \mathbb{Z}^N$ ganzer Zahlen, die

$$\left(b_n = \sum_{d|n} dy_d \text{ für jedes } n \in N \right) \quad (75)$$

erfüllt.

Aussage \mathcal{E} : Es gibt *genau eine* Familie $(y_n)_{n \in N} \in \mathbb{Z}^N$ ganzer Zahlen, die (75) erfüllt.

Aussage \mathcal{F} : Für jedes $n \in N$ gilt

$$\sum_{d|n} \mu(d) b_{n/d} \in n\mathbb{Z},$$

wobei $\mu : \mathbb{N}_+ \rightarrow \mathbb{Z}$ die Möbiusfunktion ist.

Aussage \mathcal{G} : Für jedes $n \in N$ gilt

$$\sum_{d|n} \phi(d) b_{n/d} \in n\mathbb{Z}.$$

Aussage \mathcal{H} : Für jedes $n \in N$ gilt

$$\sum_{i=1}^n b_{\text{ggT}(i,n)} \in n\mathbb{Z}.$$

Aussage \mathcal{I} : Es gibt eine Familie $(q_n)_{n \in N} \in \mathbb{Z}^N$ ganzer Zahlen, die

$$\left(b_n = \sum_{d|n} d \binom{q_d n/d}{n/d} \text{ für jedes } n \in N \right) \quad (76)$$

erfüllt.

Aussage \mathcal{J} : Es gibt *genau eine* Familie $(q_n)_{n \in N} \in \mathbb{Z}^N$ ganzer Zahlen, die (76) erfüllt.

Der Beweis dieses Satzes verläuft völlig analog zum Beweis von Satz 2.1 und Satz 4.7, mit dem einzigen Unterschied, daß wir jetzt nicht mehr von Folgen, sondern von Familien reden müssen. Mehr soll hier zu dem Beweis nicht gesagt werden.

Es stellt sich nun heraus, daß wir im Falle eines *endlichen* Nests N zu den äquivalenten Aussagen \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} , \mathcal{H} , \mathcal{I} und \mathcal{J} noch eine zusätzliche, kombinatorische Aussage hinzufügen können:

Satz 7.2: Sei N ein *endliches* Nest, und $(b_n)_{n \in N} \in \mathbb{Z}^N$ eine Familie ganzer Zahlen. Dann sind die Aussagen \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} , \mathcal{H} , \mathcal{I} , \mathcal{J} und \mathcal{K} äquivalent, wobei \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \mathcal{F} , \mathcal{G} , \mathcal{H} , \mathcal{I} und \mathcal{J} die in Satz 7.1 festgelegten Aussagen sind, und \mathcal{K} die folgende Aussage ist:

Aussage \mathcal{K} : Es gibt eine endliche Menge M , und zwei Abbildungen $f : M \rightarrow M$ und $g : M \rightarrow M$, so daß

$$(|\text{Fix}(f^n)| - |\text{Fix}(g^n)|) = b_n \quad \text{für jedes } n \in N$$

gilt. Hierbei bezeichnet $\text{Fix } h$ die Menge $\{m \in M \mid h(m) = m\}$, wobei $h : M \rightarrow M$ eine Abbildung ist.

Und wenn $(b_n)_{n \in \mathbb{N}} \in \mathbb{Z}^{\mathbb{N}}$ eine Familie *natürlicher* Zahlen ist, läßt sich dies noch verschärfen:

Satz 7.3: Sei N ein *endliches* Nest, und $(b_n)_{n \in \mathbb{N}} \in \mathbb{Z}^{\mathbb{N}}$ eine Familie *natürlicher* Zahlen. Dann sind die Aussagen $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}, \mathcal{H}, \mathcal{I}, \mathcal{J}, \mathcal{K}$ und $\mathcal{K}_{\mathbb{N}}$ äquivalent, wobei $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}, \mathcal{H}, \mathcal{I}$ und \mathcal{J} die in Satz 7.1 festgelegten Aussagen sind, \mathcal{K} die in Satz 7.2 festgelegte Aussage ist, und $\mathcal{K}_{\mathbb{N}}$ die folgende Aussage ist:

Aussage $\mathcal{K}_{\mathbb{N}}$: Es gibt eine endliche Menge M , und eine Abbildungen $f : M \rightarrow M$, so daß

$$(|\text{Fix}(f^n)| = b_n \quad \text{für jedes } n \in \mathbb{N})$$

gilt. Hierbei bezeichnet $\text{Fix } h$ die Menge $\{m \in M \mid h(m) = m\}$, wobei $h : M \rightarrow M$ eine Abbildung ist.

[...]

* beweise

* dold erwähnen?

* perlenketten als sonderfall

* phi-summe = anzahl der ÄK mod aktion; mu-summe = anzahl der afixpunkte

...

8. Zahlentheorie III: Irreduzible Polynome über \mathbb{F}_q

* leider nur für q primpotenz

Damit haben wir eine kombinatorische (oder, in unserem Fall eher algebraische oder zahlentheoretische - denn endliche Körper mit q Elementen sind mit reiner Kombinatorik schwer zu beschreiben, außer wenn q Primzahl ist) Interpretation für die Zahlen $\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}$ gefunden, wenn $n \in \mathbb{N}_+$ und q Primpotenz ist. Folglich sind diese Zahlen ganz, wenn $n \in \mathbb{N}_+$ und q Primpotenz ist. Leider ist es nicht möglich, hieraus zu folgern, daß sie auch für *alle ganzen* q (und nicht nur für Primpotenzen) ganz sind; es gibt nämlich durchaus Polynome, die auf allen Primpotenzen ganzzahlige Werte annehmen, aber nicht auf allen ganzen Zahlen. Außerdem: Auch wenn dies möglich wäre, hätten wir damit nur die Ganzzahligkeit von $\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}$ bewiesen und nicht

die Ganzzahligkeit von $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d}$; um letztere daraus zu folgern, müsste man wieder das Argument bemühen, mit dem die Äquivalenz $\mathcal{F} \iff \mathcal{G}$ von Satz 2.1 gezeigt wurde.

...

9. Freie Liealgebren und Dimensionen

...

10. Wittpolynome

...

11. Zahlentheorie II: Geisterburnsidefolgen

In Abschnitt 2 haben wir Satz 1.1 verallgemeinert, indem wir $q^{n/d}$ im Term $\frac{1}{n} \sum_{d|n} \phi(d) q^{n/d}$ durch $b_{n/d}$ ersetzt haben, wobei $(b_1, b_2, b_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ eine Folge ganzer Zahlen ist. Es ist aber auch eine andere Verallgemeinerung möglich, indem man q durch q_d ersetzt für eine Folge $(q_1, q_2, q_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ ganzer Zahlen. Wieder kann man sich dann fragen, wann $\frac{1}{n} \sum_{d|n} \phi(d) q_d^{n/d} \in \mathbb{Z}$ für jedes $n \in \mathbb{N}_+$ gilt. Folgender Satz (den ich nirgendwo in der Literatur gesehen habe) gibt die Antwort:

Satz 11.1 (der Äquivalenzsatz für Geisterburnsidefolgen): Sei $(q_1, q_2, q_3, \dots) \in \mathbb{Z}^{\mathbb{N}_+}$ eine Folge ganzer Zahlen. Dann sind folgende Aussagen $\mathcal{A}_{\text{burn}}$ und $\mathcal{G}_{\text{burn}}$ äquivalent:

Aussage $\mathcal{A}_{\text{burn}}$: Für jede Zahl $n \in \mathbb{N}_+$ und jeden Primteiler p von n gilt

$$q_{n/p} \equiv q_n \pmod{p}.$$

Aussage $\mathcal{G}_{\text{burn}}$: Für jedes $n \in \mathbb{N}_+$ gilt

$$\sum_{d|n} \phi(d) q_d^{n/d} \in n\mathbb{Z}.$$

[...]

[endlicher fall vll auch]

[möbius etc]

* wohl eine beweisskizze

* verallg. auf polya

12. Spuren von Matrixpotenzen

...

13. Die Perlenketten-Identität

[...]

14. Kombinatorik III: Teilmengen von $\{1, 2, \dots, n\}$ mit durch n teilbarer Summe

* auch durch kn teilbare summe betrachten

* anzahlen von elementen fixen, oder: wie kommt man an die binomialkoeffizienten

* stirlingnumberversion

Literaturhinweise

[1] Michiel Hazewinkel, *Witt vectors. Part 1*, revised version: 20 April 2008.

[2] Siegfried Bosch, *Algebra*, Sechste Auflage, Springer-Verlag 2006.

[3] Darij Grinberg, *Witt#5: Around the integrality criterion 9.93*.

<http://www.cip.ifi.lmu.de/~grinberg/algebra/witt5.pdf>

[4] Tales et al., *MathLinks topic #30906 ("Multiplicative function")*.

<http://www.mathlinks.ro/Forum/viewtopic.php?t=30906>

[5] Reinhold Remmert, Peter Ulrich, *Elementare Zahlentheorie*, 3. Auflage 2008.

[6] Arthur Engel, *Problem-Solving Strategies*, Springer 1998.

[7] Xenon et al., *Matheplanet topic #130535*.

<http://matheplanet.com/matheplanet/nuke/html/viewtopic.php?topic=130535>

[8] Herbert S. Wilf, *generatingfunctionology*, 2004.

<http://www.math.upenn.edu/~wilf/DownldGF.html>

[9] Serge Lang, *Algebra*, Third Edition, Springer 2002.

[10] Pierre Antoine Grillet, *Abstract Algebra*, Second Edition, Springer 2007.

[11] Nicolas Bourbaki, *Algebra I, Chapters 1-3*, 2nd printing, Springer 1989.

[12] Darij Grinberg, *Witt#4: Some computations with symmetric functions*.

<http://www.cip.ifi.lmu.de/~grinberg/algebra/witt4.pdf>

[13] Ronald L. Graham, Donald E. Knuth, Oren Patashnik, *Concrete Mathematics*, 2nd Edition, 1994.

[14] *Sum of binomial coefficients [with gcd] (MathLinks topic #91364)*,

<http://www.mathlinks.ro/Forum/viewtopic.php?t=91364>

[15] Robin Hartshorne, *Algebraic Geometry*, Springer 1977.

[16] Richard Stanley, *Enumerative Combinatorics, Volume 1*, second edition, preliminary version 2010.

<http://math.mit.edu/~rstan/ec/ec1/>