
Vorwort

The world is continuous, but the mind is discrete.

David Mumford

Unser Ziel ist es, einige kritische Lücken zwischen diversen Gebieten der Mathematik zu schließen, indem wir das Zusammenspiel zwischen dem stetigen und dem diskreten Volumen von Polytopen untersuchen. Beispiele für Polytope in drei Dimensionen sind unter anderem Kristalle, Quader, Tetraeder und beliebige konvexe Objekte, deren Oberflächen alle flach sind. Es ist unterhaltsam zu sehen, wie viele Probleme aus der Kombinatorik, Zahlentheorie und vielen weiteren mathematischen Gebieten in die Sprache von Polytopen, die in einem euklidischen Raum existieren, übersetzt werden können. Umgekehrt liefert uns die flexible Struktur von Polytopen zahlentheoretische und kombinatorische Informationen, die auf natürliche Weise aus ihrer Geometrie hervorquellen.

Das *diskrete Volumen* eines Körpers \mathcal{P} kann intuitiv als die Anzahl der Rasterpunkte, die in \mathcal{P} liegen, beschrieben werden, wenn ein festes Raster im euklidischen Raum gegeben ist. Das *stetige Volumen* von \mathcal{P} hat die übliche intuitive Bedeutung des Volumens, das wir alltäglichen Gegenständen in der wirklichen Welt zuordnen.



Abb. 0.1. Stetiges und diskretes Volumen.

Den Unterschied zwischen den beiden Volumenbegriffen kann man sich physikalisch wie folgt denken. Auf der einen Seite liefert uns das Raster auf Quantenebene, das von der molekularen Struktur der Wirklichkeit vorgegeben wird, einen diskreten Begriff des Raums und damit ein diskretes Volumen. Auf der anderen Seite liefert uns der Newton'sche Begriff des stetigen Raums das stetige Volumen. Wir betrachten die Dinge stetig auf der Newton'schen Ebene, aber in der Praxis berechnen wir oft Dinge diskret auf der Quantenebene. Mathematisch gesehen hilft uns das Raster, das wir in den Raum legen – entsprechend dem durch die Atome, aus denen ein Gegenstand besteht, gebildeten Raster – auf überraschende Weise dabei, das übliche stetige Volumen zu berechnen, wie wir noch sehen werden.

Um das stetig-diskrete-Zusammenspiel der drei Felder Kombinatorik, Zahlentheorie und Geometrie in Aktion zu sehen, konzentrieren wir uns zunächst auf das leicht zu stellende *Münzenproblem* von Frobenius. Die Schönheit dieses konkreten Problems besteht darin, dass es leicht zu verstehen ist, ein nützliches Berechnungstool liefert und trotzdem die meisten Zutaten der tiefergehenden Theorien, die hier entwickelt werden, enthält.

Im ersten Kapitel geben wir detaillierte Formeln, die sich auf natürliche Weise aus dem Frobenius'schen Münzwechselproblem ergeben, an, um die Verbindungen zwischen den drei oben genannten Gebieten aufzuzeigen. Das Münzenproblem gibt uns ein Gerüst, um die Verbindungen zwischen diesen Gebieten zu identifizieren. In den nachfolgenden Kapiteln entfernen wir dieses Gerüst und konzentrieren uns auf die Verbindungen selbst:

- (1) Aufzählung ganzzahliger Punkte in Polyedern – Kombinatorik,
- (2) Dedekind-Summen und endliche Fourier-Reihen – Zahlentheorie und
- (3) Polygone und Polytope – Geometrie.

Wir legen besonderen Wert auf Berechnungstechniken und auf die Berechnung von Volumina durch Zählen ganzzahliger Punkte unter Benutzung diverser alter und neuer Ideen. Daher sollen die Formeln, die wir erhalten, nicht nur schön sein (was sie wahrlich sind!), sondern sie sollen es uns auch erlauben, Volumina effizient zu berechnen, indem wir einige schöne Funktionen verwenden. In den wirklich seltenen Fällen mathematischer Darstellungen, in denen wir eine Formulierung haben, die sowohl „leicht zu schreiben“ als auch „schnell berechenbar“ ist, haben wir ein mathematisches Juwel gefunden. Wir haben uns bemüht, dieses Buch mit solchen mathematischen Juwelen zu füllen.

Vieles vom Material in diesem Buch wird vom Leser in den mehr als 200 Aufgaben entwickelt. Die meisten Kapitel enthalten Aufwärmübungen, die nicht auf dem Material in dem Kapitel aufbauen und gestellt werden können, bevor das Kapitel gelesen wird. Einige Aufgaben sind von zentraler Bedeutung, in dem Sinn, dass das aktuelle oder spätere Material von ihnen abhängt. Diese Aufgaben sind mit ♣ markiert, und wir geben detaillierte Lösungshinweise dazu am Ende des Buches. Die meisten Kapitel enthalten auch eine Liste offener Forschungsprobleme.

Es stellt sich heraus, dass sogar ein Fünftklässler eine interessante Arbeit über Gitterpunktaufzählung schreiben kann [144], während das Thema sich zur tiefergehenden Untersuchung anbietet, die die aktuellen Anstrengungen führender Forscher anzieht. Es handelt sich also um ein Gebiet der Mathematik, das sowohl unsere unschuldigen Kindheitsfragen als auch unsere verfeinerte Einsicht und tiefere Neugierde anzieht. Das Niveau der Untersuchung ist sehr angemessen für eine vertiefende Grundstudiumsvorlesung in Mathematik. Da die drei oben skizzierten Themen sich zur weiteren Untersuchung anbieten, wurde unser Buch auch mit Erfolg für einen einführenden Hauptstudiumskurs verwendet.

Um dem Leser dabei zu helfen, die Tragweite der Verbindungen zwischen dem stetigen und dem diskreten Volumen voll zu erfassen, beginnen wir unsere Abhandlung in zwei Dimensionen, wo wir leicht Skizzen machen und schnell experimentieren können. Wir führen behutsam die Funktionen, die wir in höheren Dimensionen brauchen (Dedekind-Summen) ein, indem wir das Münzenproblem geometrisch als das diskrete Volumen eines verallgemeinerten Dreiecks, auch Simplex genannt, betrachten.

Die Techniken sind am Anfang recht einfach, im Wesentlichen nichts weiter als Partialbruchzerlegung rationaler Funktionen. Daher ist das Buch für einen Studenten, der die üblichen Vorlesungen über Analysis und lineare Algebra gehört hat, leicht verständlich. Hilfreich wären ein grundlegendes Verständnis der Partialbruchzerlegung, unendlicher Reihen, offener und abgeschlossener Mengen im \mathbb{R}^d , komplexer Zahlen (insbesondere Einheitswurzeln) und modularer Arithmetik.

Ein wichtiges Berechnungstool, das wir uns das ganze Buch hindurch zu Nutze machen werden, ist die *Erzeugendenfunktion* $f(x) = \sum_{m=0}^{\infty} a(m)x^m$, wobei die $a(m)$ eine beliebige Folge von Zahlen bilden, die wir untersuchen möchten. Wenn die unendliche Folge von Zahlen $a(m), m = 0, 1, 2, \dots$, in einer einzigen Funktion $f(x)$ zusammengefasst wird, stellt sich oft heraus, dass wir aus bis dahin unvorhergesehenen Gründen die ganze Reihe $f(x)$ in überraschend kompakter Form aufschreiben können. Es ist diese Umformulierung der Erzeugendenfunktionen, die es uns erlaubt, die Kombinatorik der zugrundeliegenden Folge $a(m)$ zu verstehen. Für uns könnte die Folge zum Beispiel die Anzahl der möglichen Zerlegungen einer ganzen Zahl mit gegebenen Münzwerten oder die Anzahl der Punkte in einem immer größer werdenden Körper sein, und so weiter. Hier finden wir noch ein weiteres Beispiel für das Zusammenspiel zwischen dem Diskreten und dem Stetigen: Wir bekommen eine *diskrete* Menge von Zahlen $a(m)$ und führen dann die Untersuchung auf der Erzeugendenfunktion $f(x)$ in der *stetigen* Variable x durch.

Was ist das diskrete Volumen?

Die oben gegene physikalisch intuitive Beschreibung des diskreten Volumens steht auf einem soliden mathematischen Fundament, sobald wir den Begriff des Gitters einführen. Das Raster wird mathematisch durch die Sammlung

aller ganzzahligen Punkte im euklidischen Raum beschrieben, nämlich $\mathbb{Z}^d = \{(x_1, \dots, x_d) : \text{alle } x_k \in \mathbb{Z}\}$. Diese diskrete Sammlung gleichmäßig verteilter Punkte wird *Gitter* genannt. Zu einem gegebenen geometrischen Körper \mathcal{P} ist sein diskretes Volumen einfach als die Anzahl der Gitterpunkte in \mathcal{P} definiert, also als Anzahl der Elemente der Menge $\mathbb{Z}^d \cap \mathcal{P}$.

Intuitiv erhalten wir, wenn wir das Gitter um einen Faktor k verkleinern und die Anzahl der so geschrumpften Gitterpunkte in \mathcal{P} zählen, eine bessere Annäherung an das Volumen von \mathcal{P} , relativ zum Volumen einer einzelnen Zelle des geschrumpften Gitters. Es stellt sich heraus, dass, nachdem das Gitter um einen ganzzahligen Faktor k geschrumpft wurde, die Anzahl $\#(\mathcal{P} \cap \frac{1}{k}\mathbb{Z}^d)$ geschrumpfter Gitterpunkte in einem *ganzzahligen Polytop* \mathcal{P} wie von Geisterhand ein Polynom in k ist. Diese Zählfunktion $\#(\mathcal{P} \cap \frac{1}{k}\mathbb{Z}^d)$ ist als *Ehrhart-Polynom* von \mathcal{P} bekannt. Wenn wir das Gitter durch Grenzwertbildung immer weiter schrumpfen, dann kommen wir natürlich beim durch das Riemann-Integral aus der Analysis definierten stetigen Volumen heraus:

$$\text{vol } \mathcal{P} = \lim_{k \rightarrow \infty} \# \left(\mathcal{P} \cap \frac{1}{k}\mathbb{Z}^d \right) \frac{1}{k^d}.$$

Wenn wir aber bei festgelegten Streckungen des Gitters stehenbleiben, erhalten wir überraschende Flexibilität für die Berechnung des Volumens von \mathcal{P} und für die Anzahl der Gitterpunkte, die in \mathcal{P} enthalten sind.

Wenn also der Körper \mathcal{P} ein ganzzahliges Polytop ist, verhalten sich die Fehlerterme, die die Diskrepanz zwischen dem diskreten und dem stetigen Volumen messen, recht erfreulich; sie sind durch Ehrhart-Polynome gegeben, und diese Aufzählungspolynome bilden den Inhalt von Kapitel 3

Die Fourier-Dedekind-Summen sind die Bausteine: Zahlentheorie

Jedes Polytop hat ein diskretes Volumen, das durch gewisse endliche Summen, die als *Dedekind-Summen* bekannt sind, ausgedrückt werden kann. Bevor wir deren Definition geben, motivieren wir diese Summen zunächst mit einigen Beispielen, die ihr Verhalten als Bausteine der Gitterpunkt-Aufzählung illustrieren. Konkret betrachten wir als Beispiel ein 1-dimensionales Polytop, gegeben durch das Intervall $\mathcal{P} = [0, a]$, wobei a eine beliebige positive reelle Zahl ist. Es ist klar, dass wir die Gauß-Klammer $[x]$ benötigen, um die Gitterpunkte in \mathcal{P} zu zählen, und tatsächlich ist die Antwort $[a] + 1$.

Als nächstes betrachten wir einen 1-dimensionalen Geradenabschnitt in der 2-dimensionalen Ebene. Wir wählen unseren Geradenabschnitt \mathcal{P} so, dass er im Koordinatenursprung beginnt und im Gitterpunkt (c, d) endet. Wie nach kurzem Nachdenken klar wird, enthält die Anzahl der Gitterpunkte auf diesem endlichen Geradenabschnitt einen alten Bekannten, nämlich den größten gemeinsamen Teiler von c und d . Die genaue Anzahl der Gitterpunkte auf dem Geradenabschnitt ist $\text{ggT}(c, d) + 1$.

Um diese beiden Beispiele zu vereinheitlichen, betrachten wir ein Dreieck \mathcal{P} in der Ebene, dessen Ecken rationale Koordinaten haben. Es stellt sich heraus,

dass eine bestimmte endliche Summe völlig natürlich ist, da sie gleichzeitig die Gauß-Klammer und den größten gemeinsamen Teiler verallgemeinert, obwohl letzteres weniger offensichtlich ist. Ein Beispiel für eine Dedekind-Summe in zwei Dimensionen, die auf natürliche Weise in der Formel für das diskrete Volumen eines rationalen Dreiecks \mathcal{P} auftaucht, ist das Folgende:

$$s(a, b) = \sum_{m=1}^{b-1} \left(\frac{m}{b} - \frac{1}{2} \right) \left(\frac{ma}{b} - \left\lfloor \frac{ma}{b} \right\rfloor - \frac{1}{2} \right).$$

Die Definition benutzt die Gauß-Klammer. Warum ähneln diese Summen auch dem größten gemeinsamen Teiler? Glücklicherweise genügen die Dedekind-Summen einem bemerkenswerten Reziprozitätsgesetz, ganz ähnlich dem euklidischen Algorithmus, der den ggT berechnet. Dieses Reziprozitätsgesetz erlaubt es, Dedekind-Summen in etwa $\log(b)$ Schritten zu berechnen, anstatt der b Schritte, die die obige Definition nahelegt. Das Reziprozitätsgesetz für $s(a, b)$ bildet das Herzstück einiger erstaunlicher Zahlentheorie, die wir elementar behandeln, die aber auch aus dem tiefergehenden Gebiet der Modulformen und anderer moderner Hilfsmittel kommt.

Wir befinden uns in der glücklichen Position, eine wichtige Spitze eines enormen Ideenbergs zu sehen, der in die Wasser der Geometrie getaucht ist. Während wir immer tiefer in diese Gewässer eintauchen, zeigt sich uns immer mehr versteckte Schönheit, und die Dedekind-Summen sind ein unverzichtbares Hilfsmittel, das es uns erlaubt, je weiter zu sehen, desto tiefer wir dringen.

Die relevanten Körper sind Polytope: Geometrie

Die Beispiele, die wir benutzt haben, nämlich Geradenabschnitte und Polygone in der Ebene, sind Spezialfälle von Polytopen in beliebigen Dimensionen. Ein Weg, Polytope zu definieren, ist es, die *konvexe Hülle* einer endlichen Menge von Punkten im euklidischen Raum \mathbb{R}^d zu betrachten. Das heißt, angenommen, jemand gibt uns eine Menge von Punkten $\mathbf{v}_1, \dots, \mathbf{v}_n$ im \mathbb{R}^d . Das durch die gegebenen Punkte \mathbf{v}_j bestimmte Polytop ist definiert als die Menge aller Linearkombinationen $c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n$, wobei die Koeffizienten c_j nichtnegative reelle Zahlen sind, die der Bedingung $c_1 + c_2 + \dots + c_n = 1$ genügen. Diese Konstruktion wird *\mathcal{V} -Beschreibung* des Polytops genannt.

Es gibt eine weitere äquivalente Definition, die *\mathcal{H} -Beschreibung* des Polytops. Wenn uns nämlich jemand die linearen Ungleichungen gibt, die eine Sammlung von Halbräumen im \mathbb{R}^d definieren, dann können wir das dazugehörige Polytop als Durchschnitt aller durch die gegebenen Ungleichungen definierten Halbräume definieren.

Es gibt einige „offensichtliche“ Tatsachen über Polytope, die den meisten Studenten intuitiv klar sind, die aber in Wirklichkeit vertrackt sind und deren Beweis aus elementaren Axiomen nicht trivial ist. Zwei dieser Tatsachen, nämlich dass jedes Polytop sowohl eine \mathcal{V} - als auch eine \mathcal{H} -Beschreibung hat, und dass jedes Polytop trianguliert werden kann, bilden eine unabdingbare

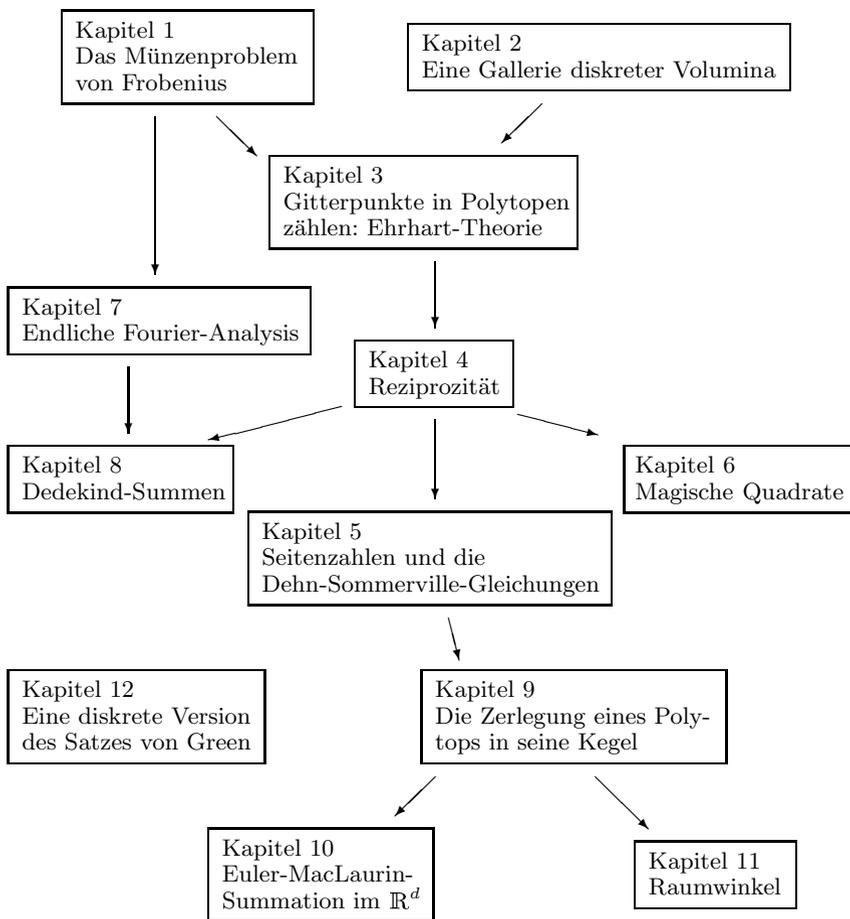


Abb. 0.2. Die partiell geordnete Menge der Kapitelabhängigkeiten.

Basis für das Material, das wir in diesem Buch erarbeiten werden. Wir beweisen beide Tatsachen sorgfältig in den Anhängen. Die beiden Hauptaussagen der Anhänge sind intuitiv klar, so dass Neulinge die Beweise überspringen können, ohne dass ihre Fähigkeit zur Berechnung stetiger und diskreter Volumina Schaden nimmt. Alle Sätze im Text (auch die in den Anhängen) werden aus grundlegenden Axiomen hergeleitet, mit Ausnahme des letzten Kapitels, in dem wir einige Grundbegriffe der Funktionentheorie voraussetzen.

Der Text ist in zwei Teile gegliedert, wie wir jetzt erläutern.

Teil I

Wir haben uns sehr bemüht, den Inhalt der ersten sechs Kapitel nahtlos zu einem Text zusammenfließen zu lassen.

- Die Kapitel 1 und 2 führen einige Grundkonzepte von Erzeugendenfunktionen ein, im visuell ansprechenden Kontext diskreter Geometrie, mit einer Fülle detaillierter motivierender Beispiele.
- Die Kapitel 3, 4 und 5 entwickeln die gesamte Ehrhart-Theorie diskreter Volumina rationaler Polytope.
- Kapitel 6 ist ein „Nachtischkapitel“ in dem Sinn, dass es uns Gelegenheit gibt, die entwickelte Theorie auf die Aufzählung *magischer Quadrate* anzuwenden, einem antiken Thema, das sich aktiver gegenwärtiger Forschung erfreut.

Teil II

Wir fangen jetzt noch einmal von Vorne an.

- Nachdem wir Erfahrung mit einer Vielzahl von Beispielen und Ergebnissen über ganzzahlige Polytope gesammelt haben, sind wir bereit, mehr über die *Dedekind-Summen* aus Kapitel 8 zu lernen, die die atomaren Einheiten der diskreten Volumenpolynome bilden. Auf der anderen Seite müssen wir uns, um Dedekind-Summen vollständig zu verstehen, mit *endlicher Fourier-Analysis* vertraut machen, die wir daher aus elementaren Grundlagen in Kapitel 7 entwickeln, wobei wir lediglich Partialbruchzerlegungen verwenden.
- Kapitel 9 beantwortet eine einfache aber vertrackte Frage: Wie lässt sich die endliche geometrische Reihe in einer Dimension auf höherdimensionale Polytope erweitern? Der *Satz von Brion* gibt die elegante und endgültige Antwort auf diese Frage.
- Kapitel 10 erweitert das Zusammenspiel zwischen dem stetigen und dem diskreten Volumen eines Polytops (das wir bereits im ersten Teil im Detail untersucht haben) durch Einführung von *Euler-Maclaurin-Summiationsformeln* in allen Dimensionen. Diese Formeln vergleichen die stetige Fourier-Transformation eines Polytops mit dessen diskreter Fourier-Transformation, dabei ist das Material völlig in sich abgeschlossen.
- Kapitel 11 entwickelt eine spannende Erweiterung der Ehrhart-Theorie, die *Raumwinkel* eines Polytops definiert und untersucht; diese sind natürliche Erweiterungen 2-dimensionaler Winkel auf höhere Dimensionen.
- Schließlich enden wir mit einem weiteren „Nachtischkapitel“, das funktionentheoretische Methoden verwendet, um eine Integralformel für die Diskrepanz zwischen diskreten und stetigen Flächen, die von einer geschlossenen Kurve in der Ebene umschlossen sind, zu finden.

Da Polytope sowohl von theoretischem Nutzen (in triangulierten Mannigfaltigkeiten zum Beispiel) als auch in der Praxis unverzichtbar (in Computergrafik zum Beispiel) sind, werden wir sie benutzen, um Ergebnisse aus der Zahlentheorie und der Kombinatorik zu verbinden. Viele Forschungsarbeiten werden über diese Zusammenhänge geschrieben, noch während wir dies schreiben, und es ist unmöglich, diese alle hier abzudecken; wir hoffen allerdings, dass diese bescheidenen Anfänge dem Leser, der nicht mit diesen Gebieten vertraut ist, ein gewisses Gefühl für ihre Schönheit, grenzenlose Verbundenheit und Nützlichkeit geben. Wir laden den allgemeinen mathematischen Leser herzlich ein in das, was wir für eine atemberaubende Welt des Zählens und der Verbindungen zwischen Gebieten der Kombinatorik, Zahlentheorie und Geometrie halten.

Es gibt eine Reihe hervorragender Bücher, die sich nichttrivial mit unserem überschneiden und Material enthalten, das die hier behandelten Themen ergänzt. Wir empfehlen herzlich die Monografien von Barvinok [12] (über Konvexität allgemein), Ehrhart [80] (die historische Einführung in Ehrhart-Theorie), Ewald [81] (über Verbindungen zur algebraischen Geometrie), Hibi [95] (über das Zusammenspiel algebraischer Geometrie mit Polytopen), Miller-Sturmfels [131] (über algorithmische kommutative Algebra) und Stanley [171] (über allgemeine Aufzählungsprobleme in der Kombinatorik).

Danksagungen

Wir hatten das große Glück, während des Schreibens an diesem Buch Hilfe von vielen freundlichen Menschen zu erhalten. Zu allererst bedanken wir uns bei den Studenten der Kurse, in denen wir dieses Material ausprobieren konnten, an der Binghamton University (SUNY), der San Francisco State University und der Temple University. Unser Dank gilt unseren Studenten bei der MSRI/Banff 2005 Graduate Summer School. Insbesondere danken wir Kristin Camenga und Kevin Woods, die die Tutorien zu diesem Sommerkurs abgehalten, zahlreiche Tippfehler entdeckt und uns viele interessante Anregungen für dieses Buch gegeben haben. Wir sind dankbar für die großzügige Unterstützung des Sommerkurses durch das Mathematical Sciences Research Institute, das Pacific Institute of Mathematics und die Banff International Research Station.

Viele Kollegen haben dieses Unterfangen unterstützt, und wir sind besonders all denen dankbar, die uns über (Tipp-)Fehler informiert und uns gute Vorschläge gemacht haben: Daniel Antonetti, Alexander Barvinok, Nathanael Berglund, Andrew Beyer, Tristram Bogart, Garry Bowlin, Benjamin Braun, Robin Chapman, Yitwah Cheung, Jessica Cuomo, Dimitros Dais Aaron Dall, Jesus De Loera, David Desario, Mike Develin, Ricardo Diaz, Michael Dobbins, Jeff Doker, Han Minh Duong, Richard Ehrenborg, David Einstein, Joseph Gubeladze, Christian Haase, Mary Halloran, Friedrich Hirzebruch, Brian Hopkins, Serkan Hoşten, Benjamin Howard, Piotr Maciak, Evgeny Materov, Asia Matthews, Peter McMullen, Martín Mereb, Ezra Miller, Mel Nathanson,

Julian Pfeifle, Peter Pleasants, Jorge Ramírez Alfonsín, Bruce Reznick, Adrian Riskin, Steven Sam, Junro Sato, Kim Seashore, Melissa Simmons, Richard Stanley, Bernd Sturmfels, Thorsten Theobald, Read Vanderbilt, Andrew Van Herick, Sven Verdoolaege, Michèle Vergne, Julie Von Bergen, Neil Weickel, Carl Woll, Zhiqiang Xu, Jon Yaggie, Ruriko Yoshida, Thomas Zaslavsky, Günter Ziegler und zwei anonymen Referees. Wir werden Errata, Aktualisierungen usw. auf der Internetseite

`math.sfsu.edu/beck/ccd.html`

sammeln.

Wir sind der Redaktion des Springer-Verlags dankbar, allen voran Mark Spencer, der uns den Prozess der Veröffentlichung stets auf freundliche und unterstützende Art vereinfacht hat. Wir danken David Kramer für das makellose Redigieren, Frank Ganz dafür, dass er uns an seinem L^AT_EX-Wissen teilhaben ließ, und Felix Pertnoy für den nahtlosen Produktionsprozess.

Matthias Beck möchte seine tiefste Dankbarkeit gegenüber Tendai Chitewere ausdrücken, für ihre Geduld, Unterstützung und bedingungslose Liebe. Er dankt seiner Familie dafür, dass sie immer für ihn da ist. Sinai Robins möchte Michal Robins, Shani Robins und Gabriel Robins für ihre unermüdliche Unterstützung und ihr Verständnis während der Fertigstellung dieses Projekts danken. Wir beide danken allen Cafés, die wir in den letzten fünf Jahren besucht haben, dafür, dass sie uns ihren Kaffee in Theoreme umwandeln ließen.

San Francisco
Philadelphia
June 2007

Matthias Beck
Sinai Robins

Das Münzenproblem von Frobenius

The full beauty of the subject of generating functions emerges only from tuning in on both channels: the discrete and the continuous.

Herb Wilf [186]

Angenommen, wir untersuchen eine unendliche Folge von Zahlen $(a_k)_{k=0}^{\infty}$, die geometrisch oder rekursiv definiert ist. Gibt es eine „gute Formel“ für a_k in Abhängigkeit von k ? Erfüllt die Folge Gleichungen in mehreren a_k s? Indem wir die Folge in ihre **Erzeugendenfunktion**

$$F(z) = \sum_{k \geq 0} a_k z^k$$

einbetten, erhalten wir auf erstaunlich einfache und elegante Art Antworten auf die obigen Fragen. In gewisser Weise hebt $F(z)$ dabei unsere Folge a_k aus ihrem diskreten Kontext in die stetige Welt der Funktionen.

1.1 Warum Erzeugendenfunktionen?

Um diese Konzepte zu veranschaulichen, wärmen wir uns mit der klassischen Folge der **Fibonacci-Zahlen** f_k auf, die nach Leonardo Pisano Fibonacci (1170–1250?)¹ benannt und durch die Rekursion

$$f_0 = 0, f_1 = 1 \quad \text{und} \quad f_{k+2} = f_{k+1} + f_k \quad \text{für} \quad k \geq 0$$

definiert sind. Dadurch erhalten wir eine Folge $(f_k)_{k=0}^{\infty} = (0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \dots)$ (siehe auch [164, Sequence A000045]). Jetzt wollen wir schauen, was Erzeugendenfunktionen für uns leisten können. Wir setzen

¹ Für mehr Informationen über Fibonacci siehe <http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Fibonacci.html>.

$$F(z) = \sum_{k \geq 0} f_k z^k$$

und übersetzen beide Seiten der Rekursionsgleichung in Aussagen über ihre Erzeugendenfunktionen:

$$\sum_{k \geq 0} f_{k+2} z^k = \sum_{k \geq 0} (f_{k+1} + f_k) z^k = \sum_{k \geq 0} f_{k+1} z^k + \sum_{k \geq 0} f_k z^k. \quad (1.1)$$

Die linke Seite von (1.1) lautet

$$\sum_{k \geq 0} f_{k+2} z^k = \frac{1}{z^2} \sum_{k \geq 0} f_{k+2} z^{k+2} = \frac{1}{z^2} \sum_{k \geq 2} f_k z^k = \frac{1}{z^2} (F(z) - z),$$

während die rechte Seite von (1.1) gleich

$$\sum_{k \geq 0} f_{k+1} z^k + \sum_{k \geq 0} f_k z^k = \frac{1}{z} F(z) + F(z)$$

ist. Also kann (1.1) als

$$\frac{1}{z^2} (F(z) - z) = \frac{1}{z} F(z) + F(z)$$

umformuliert werden, bzw.

$$F(z) = \frac{z}{1 - z - z^2}.$$

Es ist eine nette Übung, (z.B. mit dem Computer) nachzurechnen, dass wir, wenn wir F als Potenzreihe entwickeln, tatsächlich die Fibonacci-Zahlen als Koeffizienten bekommen:

$$\frac{z}{1 - z - z^2} = z + z^2 + 2z^3 + 3z^4 + 5z^5 + 8z^6 + 13z^7 + 21z^8 + 34z^9 + \dots$$

Wir wenden nun unser Lieblingswerkzeug im Umgang mit rationalen Funktionen an: Die Partialbruchzerlegung. In unserem Fall zerfällt der Nenner in die Faktoren $1 - z - z^2 = \left(1 - \frac{1+\sqrt{5}}{2}z\right) \left(1 - \frac{1-\sqrt{5}}{2}z\right)$, und die Partialbruchzerlegung lautet (s. Aufgabe 1.1)

$$F(z) = \frac{z}{1 - z - z^2} = \frac{1/\sqrt{5}}{1 - \frac{1+\sqrt{5}}{2}z} - \frac{1/\sqrt{5}}{1 - \frac{1-\sqrt{5}}{2}z}. \quad (1.2)$$

Die beiden Terme legen es nahe, die **geometrische Reihe**

$$\sum_{k \geq 0} x^k = \frac{1}{1 - x} \quad (1.3)$$

(siehe Aufgabe 1.2) mit $x = \frac{1+\sqrt{5}}{2}z$ bzw. $x = \frac{1-\sqrt{5}}{2}z$ zu verwenden:

$$\begin{aligned} F(z) &= \frac{z}{1-z-z^2} = \frac{1}{\sqrt{5}} \sum_{k \geq 0} \left(\frac{1+\sqrt{5}}{2} z \right)^k - \frac{1}{\sqrt{5}} \sum_{k \geq 0} \left(\frac{1-\sqrt{5}}{2} z \right)^k \\ &= \sum_{k \geq 0} \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^k - \left(\frac{1-\sqrt{5}}{2} \right)^k \right) z^k. \end{aligned}$$

Indem wir die Koeffizienten der z^k in der Definition von $F(z) = \sum_{k \geq 0} f_k z^k$ und dem neuen Ausdruck oben vergleichen, erhalten wir eine geschlossene Form für die Fibonacci-Zahlen:

$$f_k = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^k - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2} \right)^k.$$

Diese Methode, von einer rationalen Erzeugendenfunktion zu ihrer Partialbruchzerlegung überzugehen, ist eines unserer wichtigsten Werkzeuge. Da wir Partialbruchzerlegungen das ganze Buch hindurch immer wieder verwenden werden, halten wir das Ergebnis, auf dem diese Methode beruht, fest.

Satz 1.1 (Partialbruchzerlegung). *Zu jeder rationalen Funktion*

$$F(z) := \frac{p(z)}{\prod_{k=1}^m (z - a_k)^{e_k}},$$

wobei p ein Polynom vom Grad kleiner als $e_1 + e_2 + \dots + e_m$ ist und die a_k s paarweise verschieden sind, gibt es eine Zerlegung

$$F(z) = \sum_{k=1}^m \left(\frac{c_{k,1}}{z - a_k} + \frac{c_{k,2}}{(z - a_k)^2} + \dots + \frac{c_{k,e_k}}{(z - a_k)^{e_k}} \right),$$

wobei die $c_{k,j} \in \mathbb{C}$ eindeutig bestimmt sind.

Ein möglicher Beweis dieses Satzes basiert auf der Tatsache, dass die Polynome einen euklidischen Ring bilden. Für Leser, die mit diesem Konzept vertraut sind, skizzieren wir den Beweis in Aufgabe 1.35.

1.2 Zwei Münzen

Angenommen, wir führen ein neues Münzsystem ein: Anstatt 1, 2 und 5 Cent sowie Zehnerpotenzvielfache davon zu verwenden, einigen wir uns auf Münzen zu 4, 7, 9 und 34 Cent. Der aufmerksame Leser wird eine Schwachstelle in diesem System bemerken: Bestimmte Beträge lassen sich damit nicht herausgeben, z.B. 2 oder 5 Cent. Gerade diese Unzulänglichkeit macht unser neues

System aber auch interessant, denn sie wirft folgende Frage auf: „Welche Beträge können wir mit diesem Münzsystem darstellen?“. In Aufgabe 1.20 werden wir zeigen, dass es nur endlich viele ganzzahlige Beträge gibt, die *nicht* herausgegeben werden können. Eine naheliegende Frage, die als erstes von Georg Frobenius (1849–1917)² und Joseph Sylvester (1814–1897)³ bearbeitet wurde, lautet: „Welches ist der *größte* Betrag, der in unserem neuen Münzsystem nicht dargestellt werden kann?“. Als Mathematiker möchten wir unsere Fragestellungen so allgemein wie möglich halten und suchen für Münzen mit Werten a_1, a_2, \dots, a_d , die positive ganze Zahlen ohne einen gemeinsamen Teiler sind, nach einer Formel für den größten Betrag, der mit diesem Münzsystem nicht herausgegeben werden kann. Dieses Problem ist als *Münzenproblem von Frobenius* bekannt.

Genauer nehmen wir an, dass eine Menge

$$A = \{a_1, a_2, \dots, a_d\}$$

mit $\text{ggT}(a_1, a_2, \dots, a_d) = 1$ gegeben sei. Wir nennen eine ganze Zahl n **darstellbar**, falls es nichtnegative ganze Zahlen m_1, m_2, \dots, m_d gibt, so dass

$$n = m_1 a_1 + \dots + m_d a_d$$

gilt. Für unser Münzsystem bedeutet das, dass wir den Betrag n mit Münzen mit den Werten a_1, a_2, \dots, a_d herausgeben können. Das Frobenius-Problem (oft auch *lineares diophantisches Problem von Frobenius* genannt) besteht nun darin, die größte nicht darstellbare ganze Zahl zu finden. Wir nennen diese Zahl die **Frobenius-Zahl** und bezeichnen sie mit $g(a_1, \dots, a_d)$. Der folgende Satz gibt uns eine elegante Formel für den Fall $d = 2$.

Satz 1.2. *Für teilerfremde natürliche Zahlen a_1 und a_2 gilt*

$$g(a_1, a_2) = a_1 a_2 - a_1 - a_2.$$

Inspiziert durch diese einfach aussehende Formel für g wurde mit beträchtlichem Aufwand nach Formeln für $g(a_1, a_2, \dots, a_d)$ geforscht, allerdings nur mit begrenztem Erfolg; siehe die Anmerkungen am Ende dieses Kapitels. Für $d = 2$ fand Sylvester folgendes Resultat:

Satz 1.3 (Satz von Sylvester). *Seien a_1 und a_2 teilerfremde natürliche Zahlen. Dann ist genau die Hälfte der ganzen Zahlen zwischen 1 und $(a_1 - 1)(a_2 - 1)$ darstellbar.*

Unser Ziel in diesem Kapitel ist der Beweis dieser beiden Sätze (und etwas mehr), indem wir die Maschinerie der Partialbruchzerlegung anwenden. Wir

² Für mehr Informationen über Frobenius siehe <http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Frobenius.html>.

³ Für mehr Informationen über Sylvester siehe <http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Sylvester.html>.

nähern uns dem Frobenius-Problem, indem wir die **eingeschränkte Partitionsfunktion**

$$p_A(n) := \# \{ (m_1, \dots, m_d) \in \mathbb{Z}^d : \text{alle } m_j \geq 0 \text{ und } m_1 a_1 + \dots + m_d a_d = n \},$$

also die Anzahl der additiven Partitionen von n mit Teilen aus A , untersuchen.⁴ Im Hinblick auf diese Partitionsfunktion ist $g(a_1, \dots, a_d)$ die größte ganze Zahl n , für die $p_A(n) = 0$ gilt.

Es gibt eine schöne geometrische Interpretation der eingeschränkten Partitionsfunktion. Dazu definieren wir zunächst die Menge

$$\mathcal{P} = \{ (x_1, \dots, x_d) \in \mathbb{R}^d : \text{alle } x_j \geq 0 \text{ und } x_1 a_1 + \dots + x_d a_d = 1 \}. \quad (1.4)$$

Die n -te Streckung einer beliebigen Menge $S \subseteq \mathbb{R}^d$ ist

$$\{ (n x_1, n x_2, \dots, n x_d) : (x_1, \dots, x_d) \in S \}.$$

Die Funktion $p_A(n)$ zählt genau diejenigen Gitterpunkte, die in der n -ten Streckung von \mathcal{P} liegen. Der Streckungsprozess ist in diesem Zusammenhang gleichbedeutend damit, $x_1 a_1 + \dots + x_d a_d = 1$ in der Definition von \mathcal{P} durch $x_1 a_1 + \dots + x_d a_d = n$ zu ersetzen. Die Menge \mathcal{P} stellt sich als ein *Polytop* heraus. Man kann es für Dimensionen $d \leq 3$ leicht graphisch darstellen; Abbildung 1.1 zeigt den dreidimensionalen Fall.

1.3 Partialbrüche und eine überraschende Formel

Wir konzentrieren uns zunächst auf den Fall $d = 2$ und betrachten

$$p_{\{a,b\}}(n) = \# \{ (k, l) \in \mathbb{Z}^2 : k, l \geq 0, ak + bl = n \},$$

wobei a und b wie bisher teilerfremd sind. Zunächst experimentieren wir ein wenig mit Erzeugendenfunktionen. Wir betrachten das folgende Produkt zweier geometrischer Reihen:

$$\left(\frac{1}{1 - z^a} \right) \left(\frac{1}{1 - z^b} \right) = (1 + z^a + z^{2a} + \dots) (1 + z^b + z^{2b} + \dots)$$

(siehe Aufgabe 1.2). Wenn wir alle Terme ausmultiplizieren, erhalten wir eine Potenzreihe, in der alle Exponenten Linarkombinationen von a und b sind. Der Koeffizient von z^n in dieser Potenzreihe zählt nämlich gerade, auf wieviele Arten n als nichtnegative Linearkombination von a und b geschrieben werden

⁴ Eine (**additive**) **Partition** einer natürlichen Zahl n ist eine Multimenge (d.h. eine Menge, in der Wiederholungen erlaubt sind) $\{n_1, n_2, \dots, n_k\}$ natürlicher Zahlen, so dass $n = n_1 + n_2 + \dots + n_k$. Die Zahlen n_1, n_2, \dots, n_k heißen die **Teile** der Partition.

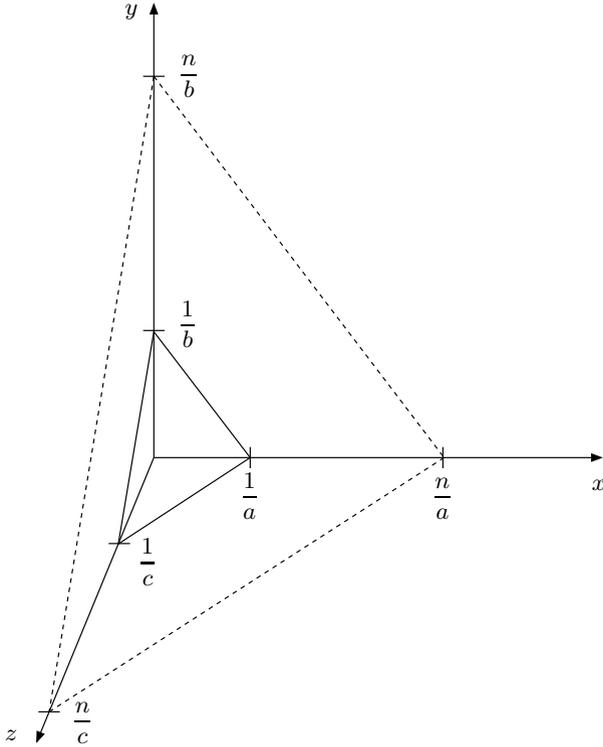


Abb. 1.1. $d = 3$.

kann. Mit anderen Worten heißt das, dass diese Koeffizienten exakt die Werte unserer Zählfunktion $p_{\{a,b\}}$ sind:

$$\left(\frac{1}{1-z^a}\right) \left(\frac{1}{1-z^b}\right) = \sum_{k \geq 0} \sum_{l \geq 0} z^{ak} z^{bl} = \sum_{n \geq 0} p_{\{a,b\}}(n) z^n.$$

Also ist diese Funktion die Erzeugendenfunktion der Folge $(p_{\{a,b\}}(n))_{n=0}^{\infty}$ von ganzen Zahlen. Die Idee besteht nun darin, die kompakte Funktion auf der linken Seite zu untersuchen.

Wir werden nun eine interessante Formel für $p_{\{a,b\}}(n)$ aufdecken, indem wir uns die Erzeugendenfunktion auf der linken Seite genauer ansehen. Um unsere Berechnungen zu vereinfachen, untersuchen wir den *konstanten Term* einer verwandten Reihe; es ist nämlich $p_{\{a,b\}}(n)$ der konstante Term von

$$f(z) := \frac{1}{(1-z^a)(1-z^b)z^n} = \sum_{k \geq 0} p_{\{a,b\}}(k) z^{k-n}.$$

Die letzte Reihe in dieser Gleichungskette ist nicht mehr ganz eine Potenzreihe, da sie auch Terme mit negativen Exponenten enthält. Solche Reihen nennt man *Laurent-Reihen*, nach Pierre Alphonse Laurent (1813–1854). Für eine (um 0 zentrierte) Potenzreihe könnten wir die dazugehörige Funktion einfach bei $z = 0$ auswerten, um den konstanten Term zu erhalten; sobald wir aber auch negative Exponenten haben ist dies nicht mehr möglich. Wenn wir jedoch zunächst alle Terme mit negativen Exponenten abziehen, bekommen wir eine Potenzreihe, deren (unveränderter) konstanter Term nun durch Auswerten der restlichen Funktion bei $z = 0$ errechnet werden kann.

Um diesen konstanten Term berechnen zu können, werden wir f in Partialbrüche zerlegen. Als Aufwärmübung im Umgang mit Partialbruchzerlegungen betrachten wir zunächst ein eindimensionales Beispiel. Wir bezeichnen die erste a -te Einheitswurzel mit

$$\xi_a := e^{2\pi i/a} = \cos \frac{2\pi}{a} + i \sin \frac{2\pi}{a}.$$

Die a -ten Einheitswurzeln sind dann $1, \xi_a, \xi_a^2, \xi_a^3, \dots, \xi_a^{a-1}$.

Beispiel 1.4. Wir wollen die Partialbruchzerlegung von $\frac{1}{1-z^a}$ bestimmen. Die Pole dieser Funktion liegen bei den a -ten Einheitswurzeln ξ_a^k für $k = 0, 1, \dots, a - 1$. Also erweitern wir zu

$$\frac{1}{1-z^a} = \sum_{k=0}^{a-1} \frac{C_k}{z - \xi_a^k}.$$

Wie können wir nun die Koeffizienten C_k bestimmen? Es gilt:

$$C_k = \lim_{z \rightarrow \xi_a^k} (z - \xi_a^k) \left(\frac{1}{1-z^a} \right) = \lim_{z \rightarrow \xi_a^k} \frac{1}{-a z^{a-1}} = -\frac{\xi_a^k}{a},$$

wobei wir die Regel von de l'Hospital in der vorletzten Gleichung verwendet haben. Also gelangen wir zu der Zerlegung

$$\frac{1}{1-z^a} = -\frac{1}{a} \sum_{k=0}^{a-1} \frac{\xi_a^k}{z - \xi_a^k}. \quad \square$$

Wir kommen zurück zu den eingeschränkten Partitionsfunktionen. Die Pole von f liegen bei $z = 0$ mit Vielfachheit n , bei $z = 1$ mit Vielfachheit 2, und bei allen anderen a -ten und b -ten Einheitswurzeln mit Vielfachheit 1, da a und b teilerfremd sind. Also sieht unsere Partialbruchzerlegung wie folgt aus:

$$f(z) = \frac{A_1}{z} + \frac{A_2}{z^2} + \dots + \frac{A_n}{z^n} + \frac{B_1}{z-1} + \frac{B_2}{(z-1)^2} + \sum_{k=1}^{a-1} \frac{C_k}{z - \xi_a^k} + \sum_{j=1}^{b-1} \frac{D_j}{z - \xi_b^j}. \quad (1.5)$$

Dem Leser sei die Berechnung der Koeffizienten zur Übung empfohlen (Aufgabe 1.21):

$$C_k = -\frac{1}{a(1-\xi_a^{kb})\xi_a^{k(n-1)}}, \quad (1.6)$$

$$D_j = -\frac{1}{b(1-\xi_b^{ja})\xi_b^{j(n-1)}}.$$

Um B_2 zu berechnen, multiplizieren wir beide Seiten von (1.5) mit $(z-1)^2$ und bestimmen den Grenzwert für $z \rightarrow 1$. Wir erhalten

$$B_2 = \lim_{z \rightarrow 1} \frac{(z-1)^2}{(1-z^a)(1-z^b)z^n} = \frac{1}{ab},$$

indem wir beispielsweise die Regel von de l'Hospital zweimal anwenden. Um die interessantere Konstante B_1 zu bestimmen, rechnen wir

$$B_1 = \lim_{z \rightarrow 1} (z-1) \left(\frac{1}{(1-z^a)(1-z^b)z^n} - \frac{\frac{1}{ab}}{(z-1)^2} \right) = \frac{1}{ab} - \frac{1}{2a} - \frac{1}{2b} - \frac{n}{ab},$$

wieder mit der Regel von de l'Hospital.

Wir brauchen die Koeffizienten A_1, \dots, A_n gar nicht auszurechnen, da sie nur zu den Termen mit negativen Koeffizienten beitragen, und die können wir einfach vernachlässigen, da sie sich nicht auf den konstanten Term von f auswirken. Sobald wir die anderen Koeffizienten haben, ergibt sich der konstante Term der Laurent-Reihe von f – wie oben erläutert – durch Auswerten der folgenden Funktion bei 0:

$$p_{\{a,b\}}(n) = \left(\frac{B_1}{z-1} + \frac{B_2}{(z-1)^2} + \sum_{k=1}^{a-1} \frac{C_k}{z-\xi_a^k} + \sum_{j=1}^{b-1} \frac{D_j}{z-\xi_b^j} \right) \Bigg|_{z=0}$$

$$= -B_1 + B_2 - \sum_{k=1}^{a-1} \frac{C_k}{\xi_a^k} - \sum_{j=1}^{b-1} \frac{D_j}{\xi_b^j}.$$

Mit Hilfe von (1.6) vereinfachen wir dies zu

$$p_{\{a,b\}}(n) = \frac{1}{2a} + \frac{1}{2b} + \frac{n}{ab} + \frac{1}{a} \sum_{k=1}^{a-1} \frac{1}{(1-\xi_a^{kb})\xi_a^{kn}} + \frac{1}{b} \sum_{j=1}^{b-1} \frac{1}{(1-\xi_b^{ja})\xi_b^{jn}}. \quad (1.7)$$

Ermuntert von diesem anfänglichen Erfolg machen wir uns nun daran, die einzelnen Summen in (1.7) zu untersuchen, in der Hoffnung, sie als bekannte Objekte zu erkennen.

Für den nächsten Schritt müssen wir zunächst die **Gauß-Klammer** $\lfloor x \rfloor$ definieren, die die größte ganze Zahl kleiner als oder gleich x bezeichnet. Eng mit ihr verwandt ist die **Nachkommaanteilsfunktion** $\{x\} = x - \lfloor x \rfloor$. Lesern, die mit den Funktionen $\lfloor x \rfloor$ und $\{x\}$ nicht vertraut sind, seien die Aufgaben 1.3 bis 1.5 zur Bearbeitung empfohlen.

Als nächstes betrachten wir einen Spezialfall, nämlich $b = 1$. Er ist deswegen interessant, weil $p_{\{a,1\}}(n)$ gerade die Gitterpunkte in einem Intervall zählt:

$$\begin{aligned} p_{\{a,1\}}(n) &= \#\{(k,l) \in \mathbb{Z}^2 : k,l \geq 0, ak + l = n\} \\ &= \#\{k \in \mathbb{Z} : k \geq 0, ak \leq n\} \\ &= \#\left\{k \in \mathbb{Z} : 0 \leq k \leq \frac{n}{a}\right\} \\ &= \left\lfloor \frac{n}{a} \right\rfloor + 1. \end{aligned}$$

(siehe Aufgabe 1.3). Auf der anderen Seite haben wir in (1.7) nur einen anderen Ausdruck für diese Funktion berechnet, so dass

$$\frac{1}{2a} + \frac{1}{2} + \frac{n}{a} + \frac{1}{a} \sum_{k=1}^{a-1} \frac{1}{(1 - \xi_a^k) \xi_a^{kn}} = p_{\{a,1\}}(n) = \left\lfloor \frac{n}{a} \right\rfloor + 1.$$

Mit Hilfe der Nachkommaanteilsfunktion $\{x\} = x - \lfloor x \rfloor$ haben wir eine Formel für die folgende Summe über a -te Einheitswurzeln hergeleitet:

$$\frac{1}{a} \sum_{k=1}^{a-1} \frac{1}{(1 - \xi_a^k) \xi_a^{kn}} = -\left\{ \frac{n}{a} \right\} + \frac{1}{2} - \frac{1}{2a}. \quad (1.8)$$

Damit sind wir fast fertig: Wir laden den Leser in Aufgabe 1.22 ein, zu zeigen, dass

$$\frac{1}{a} \sum_{k=1}^{a-1} \frac{1}{(1 - \xi_a^{bk}) \xi_a^{kn}} = \frac{1}{a} \sum_{k=1}^{a-1} \frac{1}{(1 - \xi_a^k) \xi_a^{b^{-1}kn}}, \quad (1.9)$$

wobei b^{-1} eine ganze Zahl ist, für die $b^{-1}b \equiv 1 \pmod{a}$ gilt, und folgern, dass

$$\frac{1}{a} \sum_{k=1}^{a-1} \frac{1}{(1 - \xi_a^{bk}) \xi_a^{kn}} = -\left\{ \frac{b^{-1}n}{a} \right\} + \frac{1}{2} - \frac{1}{2a}. \quad (1.10)$$

Jetzt müssen wir nur noch diesen Ausdruck zurück in (1.7) einsetzen und wir erhalten die folgende schöne Formel, die auf Tiberiu Popoviciu (1906–1975) zurückgeht.

Satz 1.5 (Satz von Popoviciu). *Für teilerfremde a und b gilt*

$$p_{\{a,b\}}(n) = \frac{n}{ab} - \left\{ \frac{b^{-1}n}{a} \right\} - \left\{ \frac{a^{-1}n}{b} \right\} + 1,$$

wobei $b^{-1}b \equiv 1 \pmod{a}$ und $a^{-1}a \equiv 1 \pmod{b}$. □

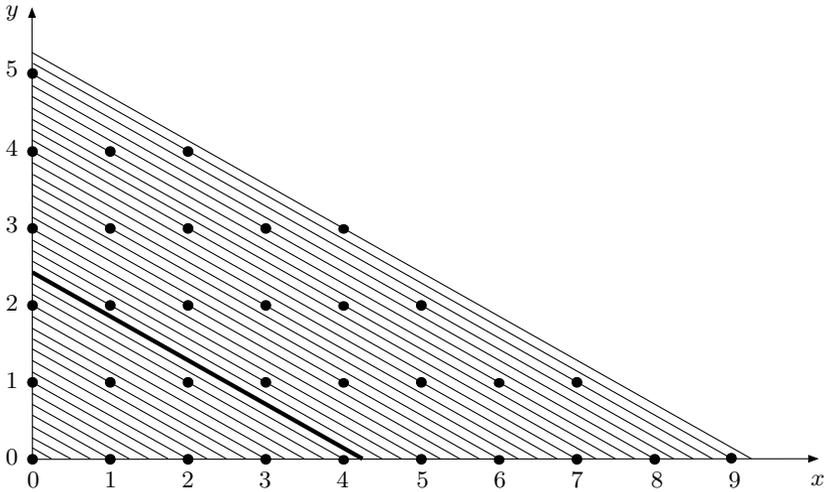


Abb. 1.2. $4x + 7y = n, n = 1, 2, \dots$

1.4 Der Satz von Sylvester

Bevor wir Satz 1.5 anwenden, um die klassischen Sätze 1.2 und 1.3 zu beweisen, kehren wir für einen Moment zur Geometrie hinter der eingeschränkten Partitionsfunktion $p_{\{a,b\}}(n)$ zurück. Im zweidimensionalen Fall (über den Satz 1.5 eine Aussage macht) zählen wir Gitterpunkte $(x, y) \in \mathbb{Z}^2$ auf Geradenabschnitten, die durch die Bedingungen

$$ax + by = n, \quad x, y \geq 0$$

bestimmt sind. Wenn n größer wird, werden diese Geradenabschnitte gestreckt. Es ist nicht zu abwegig (obwohl Aufgabe 1.13 uns lehrt, mit solchen Aussagen vorsichtig zu sein) zu erwarten, dass die Wahrscheinlichkeit, mit der ein Gitterpunkt auf dem Geradenabschnitt liegt, mit wachsendem n größer wird. Tatsächlich könnte man sogar annehmen, dass die Anzahl der Punkte auf dem Geradenabschnitt linear mit n ansteigt, da der Geradenabschnitt ein eindimensionales Objekt ist. Satz 1.5 quantifiziert diese Aussage sehr präzise: $p_{\{a,b\}}(n)$ hat den „Leiterterm“ n/ab , und die weiteren Terme sind als Funktionen von n beschränkt. Abbildung 1.2 zeigt die Geometrie hinter der Zählfunktion $p_{\{4,7\}}(n)$ für die ersten paar Werte von n . Man beachte, dass die dickgedruckte Strecke für $n = 17 = 4 \cdot 7 - 4 - 7$ die letzte ist, die überhaupt keinen Gitterpunkt enthält.

Lemma 1.6. Falls a und b teilerfremde positive ganze Zahlen sind und $n \in [1, ab - 1]$ kein Vielfaches von a oder b ist, gilt

$$p_{\{a,b\}}(n) + p_{\{a,b\}}(ab - n) = 1.$$

Mit anderen Worten, für n zwischen a und $ab - 1$, das weder durch a noch durch b teilbar ist, ist genau eine der beiden Zahlen n und $ab - n$ als Kombination von a und b darstellbar.

Beweis. Diese Gleichung folgt direkt aus Satz 1.5:

$$\begin{aligned} p_{\{a,b\}}(ab - n) &= \frac{ab - n}{ab} - \left\{ \frac{b^{-1}(ab - n)}{a} \right\} - \left\{ \frac{a^{-1}(ab - n)}{b} \right\} + 1 \\ &= 2 - \frac{n}{ab} - \left\{ \frac{-b^{-1}n}{a} \right\} - \left\{ \frac{-a^{-1}n}{b} \right\} \\ &\stackrel{(*)}{=} -\frac{n}{ab} + \left\{ \frac{b^{-1}n}{a} \right\} + \left\{ \frac{a^{-1}n}{b} \right\} \\ &= 1 - p_{\{a,b\}}(n). \end{aligned}$$

Dabei folgt $(*)$ aus der Tatsache, dass $\{-x\} = 1 - \{x\}$ falls $x \notin \mathbb{Z}$ (siehe Aufgabe 1.5). \square

Beweis von Satz 1.2. Wir müssen zeigen, dass $p_{\{a,b\}}(ab - a - b) = 0$ und dass $p_{\{a,b\}}(n) > 0$, falls $n > ab - a - b$. Die erste Behauptung folgt aus Aufgabe 1.24, die besagt, dass $p_{\{a,b\}}(a + b) = 1$, und Lemma 1.6. Um die zweite Behauptung zu zeigen, benutzen wir, dass für jede ganze Zahl m die Ungleichung $\left\{ \frac{m}{a} \right\} \leq 1 - \frac{1}{a}$ gilt. Daher gilt für jede positive ganze Zahl n , dass

$$p_{\{a,b\}}(ab - a - b + n) \geq \frac{ab - a - b + n}{ab} - \left(1 - \frac{1}{a}\right) - \left(1 - \frac{1}{b}\right) + 1 = \frac{n}{ab} > 0. \quad \square$$

Beweis von Satz 1.3. Wir erinnern uns zunächst an Lemma 1.6, das besagt, dass für n zwischen 1 und $ab - 1$, das weder a noch b als Teiler hat, genau eine der beide Zahlen n und $ab - n$ darstellbar ist. Es gibt

$$ab - a - b + 1 = (a - 1)(b - 1)$$

Zahlen zwischen 1 und $ab - 1$, die weder durch a noch durch b teilbar sind. Schließlich beachten wir, dass $p_{\{a,b\}}(n) > 0$ falls n ein Vielfaches von a oder b ist, was aus der Definition von $p_{\{a,b\}}(n)$ folgt. Daher ist die Anzahl der nicht darstellbaren ganzen Zahlen genau $\frac{1}{2}(a - 1)(b - 1)$. \square

Damit haben wir sogar mehr gezeigt. Im wesentlichen wegen Lemma 1.6 gilt, dass jede positive ganze Zahl kleiner als ab höchstens eine Darstellung hat. Daher sind die darstellbaren Zahlen, die kleiner als ab sind, *eindeutig* darstellbar (siehe auch Aufgabe 1.25).

1.5 Drei und mehr Münzen

Was passiert mit der Komplexität des Frobenius-Problems, wenn wir mehr als zwei Münzen haben? Wir kommen auf die eingeschränkte Partitionsfunktion

$$p_A(n) = \# \{ (m_1, \dots, m_d) \in \mathbb{Z}^d : \text{alle } m_j \geq 0, m_1 a_1 + \dots + m_d a_d = n \}$$

zurück, wobei $A = \{a_1, \dots, a_d\}$. Mit genau der gleichen Argumentation wie in Abschnitt 1.3 können wir ganz leicht die Erzeugendenfunktion von $p_A(n)$ aufschreiben:

$$\sum_{n \geq 0} p_A(n) z^n = \left(\frac{1}{1 - z^{a_1}} \right) \left(\frac{1}{1 - z^{a_2}} \right) \cdots \left(\frac{1}{1 - z^{a_d}} \right).$$

Wir wenden die gleichen Methoden an, die wir in Abschnitt 1.3 benutzt haben, um unsere Funktion $p_A(n)$ als konstanten Term einer nützlichen Erzeugendenfunktion zu erhalten, nämlich

$$p_A(n) = \text{const} \left(\frac{1}{(1 - z^{a_1})(1 - z^{a_2}) \cdots (1 - z^{a_d}) z^n} \right).$$

Jetzt zerlegen wir die Funktion auf der rechten Seite in Partialbrüche. Der Einfachheit halber nehmen wir im Folgenden an, dass a_1, \dots, a_d paarweise teilerfremd sind; d.h. keine zwei der ganzen Zahlen a_1, \dots, a_d haben einen gemeinsamen Teiler. Dann sieht unsere Partialbruchzerlegung wie folgt aus:

$$\begin{aligned} f(z) &= \frac{1}{(1 - z^{a_1}) \cdots (1 - z^{a_d}) z^n} \\ &= \frac{A_1}{z} + \frac{A_2}{z^2} + \cdots + \frac{A_n}{z^n} + \frac{B_1}{z - 1} + \frac{B_2}{(z - 1)^2} + \cdots + \frac{B_d}{(z - 1)^d} \quad (1.11) \\ &\quad + \sum_{k=1}^{a_1-1} \frac{C_{1k}}{z - \xi_{a_1}^k} + \sum_{k=1}^{a_2-1} \frac{C_{2k}}{z - \xi_{a_2}^k} + \cdots + \sum_{k=1}^{a_d-1} \frac{C_{dk}}{z - \xi_{a_d}^k}. \end{aligned}$$

Inzwischen sind wir geübt im Umgang mit Partialbruchkoeffizienten, so dass der Leser leicht nachprüfen kann, dass (Aufgabe 1.29)

$$C_{1k} = - \frac{1}{a_1 \left(1 - \xi_{a_1}^{ka_2} \right) \left(1 - \xi_{a_1}^{ka_3} \right) \cdots \left(1 - \xi_{a_1}^{ka_d} \right) \xi_{a_1}^{k(n-1)}}. \quad (1.12)$$

Wie vorher müssen wir auch hier die Koeffizienten A_1, \dots, A_n nicht berechnen, da sie nicht zum konstanten Term von f beitragen. Für die Berechnung von B_1, \dots, B_d können wir ein Computeralgebraprogramm wie **Maple** oder **Mathematica** benutzen. Wieder gilt, dass wir, sobald wir diese Koeffizienten berechnet haben, den konstanten Term von f berechnen können, indem wir alle negativen Exponenten weglassen und die übrig bleibende Funktion bei 0 auswerten:

$$\begin{aligned} p_A(n) &= \left(\frac{B_1}{z - 1} + \cdots + \frac{B_d}{(z - 1)^d} + \sum_{k=1}^{a_1-1} \frac{C_{1k}}{z - \xi_{a_1}^k} + \cdots + \sum_{k=1}^{a_d-1} \frac{C_{dk}}{z - \xi_{a_d}^k} \right) \Bigg|_{z=0} \\ &= -B_1 + B_2 - \cdots + (-1)^d B_d - \sum_{k=1}^{a_1-1} \frac{C_{1k}}{\xi_{a_1}^k} - \sum_{k=1}^{a_2-1} \frac{C_{2k}}{\xi_{a_2}^k} - \cdots - \sum_{k=1}^{a_d-1} \frac{C_{dk}}{\xi_{a_d}^k}. \end{aligned}$$

Wenn wir zum Beispiel die Ausdrücke, die wir für C_{lk} gefunden haben, in die letzte Summe über die nichttrivialen a_l -ten Einheitswurzeln einsetzen, ergibt sich

$$\frac{1}{a_1} \sum_{k=1}^{a_1-1} \frac{1}{(1 - \xi_{a_1}^{ka_2}) (1 - \xi_{a_1}^{ka_3}) \dots (1 - \xi_{a_1}^{ka_d}) \xi_{a_1}^{kn}}.$$

Dies motiviert die Definition der **Fourier-Dedekind-Summe**

$$s_n(a_1, a_2, \dots, a_m; b) := \frac{1}{b} \sum_{k=1}^{b-1} \frac{\xi_b^{kn}}{(1 - \xi_b^{ka_1}) (1 - \xi_b^{ka_2}) \dots (1 - \xi_b^{ka_m})}. \quad (1.13)$$

Wir werden diese Summen im Detail in Kapitel 8 untersuchen. Mit dieser Definition sind wir bei folgendem Resultat angelangt.

Satz 1.7. *Die eingeschränkte Partitionsfunktion für $A = \{a_1, a_2, \dots, a_d\}$, wobei die a_k s paarweise teilerfremd sind, kann durch*

$$p_A(n) = -B_1 + B_2 - \dots + (-1)^d B_d + s_{-n}(a_2, a_3, \dots, a_d; a_1) + s_{-n}(a_1, a_3, a_4, \dots, a_d; a_2) + \dots + s_{-n}(a_1, a_2, \dots, a_{d-1}; a_d)$$

berechnet werden. Dabei sind B_1, B_2, \dots, B_d die Partialbruchkoeffizienten in der Zerlegung von (1.11). □

Beispiel 1.8. Wir geben die eingeschränkte Partitionsfunktion für $d = 3$ und 4 an. Diese geschlossenen Formeln haben sich in der verfeinerten Analyse der Periodizität, die der Partitionsfunktion $p_A(n)$ inhärent ist, als nützlich erwiesen. Zum Beispiel kann man den Graph von $p_{\{a,b,c\}}(n)$ als eine „gewellte Parabel“ visualisieren, wie aus der Formel offensichtlich wird.

$$p_{\{a,b,c\}}(n) = \frac{n^2}{2abc} + \frac{n}{2} \left(\frac{1}{ab} + \frac{1}{ac} + \frac{1}{bc} \right) + \frac{1}{12} \left(\frac{3}{a} + \frac{3}{b} + \frac{3}{c} + \frac{a}{bc} + \frac{b}{ac} + \frac{c}{ab} \right) + \frac{1}{a} \sum_{k=1}^{a-1} \frac{1}{(1 - \xi_a^{kb}) (1 - \xi_a^{kc}) \xi_a^{kn}} + \frac{1}{b} \sum_{k=1}^{b-1} \frac{1}{(1 - \xi_b^{kc}) (1 - \xi_b^{ka}) \xi_b^{kn}} + \frac{1}{c} \sum_{k=1}^{c-1} \frac{1}{(1 - \xi_c^{ka}) (1 - \xi_c^{kb}) \xi_c^{kn}},$$

$$p_{\{a,b,c,d\}}(n) = \frac{n^3}{6abcd} + \frac{n^2}{4} \left(\frac{1}{abc} + \frac{1}{abd} + \frac{1}{acd} + \frac{1}{bcd} \right) + \frac{n}{12} \left(\frac{3}{ab} + \frac{3}{ac} + \frac{3}{ad} + \frac{3}{bc} + \frac{3}{bd} + \frac{3}{cd} + \frac{a}{bcd} + \frac{b}{acd} + \frac{c}{abd} + \frac{d}{abc} \right)$$

$$\begin{aligned}
& + \frac{1}{24} \left(\frac{a}{bc} + \frac{a}{bd} + \frac{a}{cd} + \frac{b}{ad} + \frac{b}{ac} + \frac{b}{cd} + \frac{c}{ab} + \frac{c}{ad} + \frac{c}{bd} \right. \\
& \quad \left. + \frac{d}{ab} + \frac{d}{ac} + \frac{d}{bc} \right) - \frac{1}{8} \left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} \right) \\
& + \frac{1}{a} \sum_{k=1}^{a-1} \frac{1}{(1 - \xi_a^{kb}) (1 - \xi_a^{kc}) (1 - \xi_a^{kd}) \xi_a^{kn}} \\
& + \frac{1}{b} \sum_{k=1}^{b-1} \frac{1}{(1 - \xi_b^{kc}) (1 - \xi_b^{kd}) (1 - \xi_b^{ka}) \xi_b^{kn}} \\
& + \frac{1}{c} \sum_{k=1}^{c-1} \frac{1}{(1 - \xi_c^{kd}) (1 - \xi_c^{ka}) (1 - \xi_c^{kb}) \xi_c^{kn}} \\
& + \frac{1}{d} \sum_{k=1}^{d-1} \frac{1}{(1 - \xi_d^{ka}) (1 - \xi_d^{kb}) (1 - \xi_d^{kc}) \xi_d^{kn}}. \quad \square
\end{aligned}$$

Anmerkungen

1. Die Theorie der Erzeugendenfunktionen hat eine lange und mächtige Tradition. Ihre Nützlichkeit streifen wir hier nur. Den Lesern, die ein wenig tiefer im weitläufigen Garten der Erzeugendenfunktionen graben möchten, seien Herb Wilfs *Generatingfunctionology* [186] und László Lovász's *Combinatorial Problems and Exercises* [121] sehr empfohlen. Der Leser mag sich vielleicht wundern, dass wir Konvergenzaspekte der Erzeugendenfunktionen, mit denen wir hantieren, nicht betonen. Die sind jedoch ausnahmslos geometrische Reihen, die triviale Konvergenzeigenschaften haben. Um eine klare mathematische Darstellung nicht zu trüben, lassen wir solche Konvergenzdetails weg.

2. Das Frobenius-Problem ist nach Georg Frobenius benannt, der es anscheinend gerne in Vorlesungen stellte [40]. Satz 1.2 ist ein berühmtes Resultat in der mathematischen Folklore und vielleicht eines der am häufigsten falsch zitierten mathematischen Ergebnisse überhaupt. Viele Autoren zitieren James J. Sylvesters Problem in [176], aber sein Aufsatz enthält Satz 1.3, nicht Satz 1.2. Tatsächlich war Sylvesters Problem bereits als ein Satz in [175] aufgetaucht. Es ist nicht bekannt, wer Satz 1.2 als erstes entdeckt oder bewiesen hat. Es ist durchaus denkbar, dass Sylvester diesen Satz kannte, als er Satz 1.3 aufstellte.

3. Das lineare diophantische Problem von Frobenius sollte nicht mit dem *Briefmarkenproblem* verwechselt werden. Letzteres fragt nach der Bestimmung einer ähnlichen Größe, beschränkt allerdings unabhängig die Größe der ganzzahligen Lösungen der linearen Gleichung.

4. Satz 1.5 hat eine interessante Geschichte. Das früheste uns bekannte Auftauchen dieses Ergebnisses ist in einem Aufsatz von Tiberiu Popoviciu [147]. Popovicius Formel ist seitdem mindestens zweimal wiederbelebt worden [160, 182].

5. Fourier-Dedekind-Summen sind das erste Mal implizit in Sylvesters Arbeit aufgetreten (siehe z.B. [174]), und das erste Mal explizit im Zusammenhang mit eingeschränkten Partitionsfunktionen in [103]. Sie wurden in [24] erneut entdeckt, in Verbindung mit dem Frobenius-Problem. Die Aufsätze [82, 156] enthalten interessante Verbindungen mit Bernoulli- und Euler-Polynomen. Wir werden die Untersuchung von Fourier-Dedekind-Summen in Kapitel 8 wieder aufnehmen.

6. Wie wir oben bereits erwähnt haben, ist das Frobenius-Problem für $d \geq 3$ wesentlich schwieriger als im Fall $d = 2$, den wir behandelt haben. Jenseits von $d = 3$ ist das Problem zweifelsohne noch völlig offen, obwohl bereits erhebliche Anstrengungen zu seiner Untersuchung unternommen wurden. Die Literatur zu diesem Problem ist weitläufig, und es gibt immer noch viel Raum für Verbesserungen. Der interessierte Leser sei auf die umfassende Monographie [152] verwiesen, die Referenzen auf fast alle Aufsätze, die das Frobenius-Problem behandeln, sichtet und etwa 40 offene Probleme und Vermutungen im Zusammenhang mit dem Frobenius-Problem angibt. Als Kostprobe erwähnen wir zwei Meilensteine, die über $d = 2$ hinaus gehen:

Der erste behandelt die Erzeugendenfunktion $r(z) := \sum_{k \in R} z^k$, wobei R die Menge aller ganzen Zahlen, die durch eine gegebene Menge teilerfremder positiver ganzer Zahlen a_1, a_2, \dots, a_d darstellbar sind, angibt. Es ist nicht schwer einzusehen (Aufgabe 1.34), dass $r(z) = p(z)/(1 - z^{a_1})(1 - z^{a_2}) \dots (1 - z^{a_d})$ für ein Polynom p . Diese rationale Erzeugendenfunktion enthält sämtliche Informationen über das Frobenius-Problem, z.B. ist die Frobenius-Zahl gerade der Totalgrad der Funktion $\frac{1}{1-z} - r(z)$. Daher wird das Frobenius-Problem darauf reduziert, das Polynom p , das im Zähler von r steht, zu finden. Marcel Morales [133, 134] und Graham Denham [72] haben die bemerkenswerte Tatsache entdeckt, dass für $d = 3$ das Polynom p entweder 4 oder 6 Terme hat. Darüberhinaus gaben sie halbexplizite Formeln für p . Der Satz von Morales-Denham impliziert, dass die Frobenius-Zahl im Fall $d = 3$ schnell berechnet werden kann; ein Ergebnis, das ursprünglich, in unterschiedlichen Gestalten, auf Jürgen Herzog [94], Harold Greenberg [88] und J. Leslie Davison [64] zurückgeht. Genauso wie es scheinbar eine klare Grenze zwischen den Fällen $d = 2$ und $d = 3$ gibt, scheint es auch zwischen den Fällen $d = 3$ und $d = 4$ eine Grenze zu geben: Henrik Bresinsky [42] hat bewiesen, dass für $d \geq 4$ keine absolute Schranke für die Anzahl der Terme im Zähler p existiert, im starken Kontrast zum Satz von Morales-Denham.

Auf der anderen Seite haben Alexander Barvinok und Kevin Woods [14] bewiesen, dass für feste d die rationale Erzeugendenfunktion $r(z)$ als „kurze“ Summe rationaler Funktionen geschrieben werden kann; insbesondere kann r

effizient berechnet werden, wenn d fest ist. Ein Korollar dazu ist, dass die Frobenius-Zahl für feste d effizient berechnet werden kann; dieser Satz geht auf Ravi Kannan zurück [104]. Andererseits hat Jorge Ramírez-Alfonsín [151] gezeigt, dass es aussichtslos ist, die Frobenius-Zahl effizient berechnen zu wollen, solange d variabel gelassen wird.

Während die obigen Ergebnisse die theoretische Komplexität der Berechnung der Frobenius-Zahl klären, sind praktische Algorithmen ein völlig anderes Thema. Sowohl Kannans als auch Barvinok-Woods Ideen scheinen komplex genug zu sein, dass bisher niemand versucht hat, sie zu implementieren. Der derzeit schnellste Algorithmus wird in [31] vorgestellt.

Aufgaben

1.1. ♣ Überprüfen Sie die folgende Partialbruchzerlegung (1.2):

$$\frac{z}{1-z-z^2} = \frac{1/\sqrt{5}}{1-\frac{1+\sqrt{5}}{2}z} - \frac{1/\sqrt{5}}{1-\frac{1-\sqrt{5}}{2}z}.$$

1.2. ♣ Es sei z eine komplexe Zahl und n eine positive ganze Zahl. Zeigen Sie, dass

$$(1-z)(1+z+z^2+\cdots+z^n) = 1-z^{n+1},$$

und benutzen Sie dies, um zu zeigen, dass für $|z| < 1$ gilt:

$$\sum_{k \geq 0} z^k = \frac{1}{1-z}$$

1.3. ♣ Finden Sie eine Formel für die Anzahl der Gitterpunkte in $[a, b]$ für beliebige reelle Zahlen a und b .

1.4. Zeigen Sie das Folgende. Soweit nicht anders angegeben, sei $n \in \mathbb{Z}$ und $x, y \in \mathbb{R}$.

(a) $\lfloor x+n \rfloor = \lfloor x \rfloor + n$.

(b) $\lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x+y \rfloor \leq \lfloor x \rfloor + \lfloor y \rfloor + 1$.

(c) $\lfloor x \rfloor + \lfloor -x \rfloor = \begin{cases} 0 & \text{falls } x \in \mathbb{Z}, \\ -1 & \text{sonst.} \end{cases}$

(d) Für $n \in \mathbb{Z}_{>0}$, $\lfloor \frac{\lfloor x \rfloor}{n} \rfloor = \lfloor \frac{x}{n} \rfloor$.

(e) $-\lfloor -x \rfloor$ ist die kleinste ganze Zahl größer als oder gleich x , geschrieben $\lceil x \rceil$.

(f) $\lfloor x+1/2 \rfloor$ ist die nächste ganze Zahl zu x (und falls es zwei solche Zahlen gibt, die größere der beiden).

(g) $\lfloor x \rfloor + \lfloor x+1/2 \rfloor = \lfloor 2x \rfloor$.

(h) Falls m und n positive ganze Zahlen sind, ist $\lfloor \frac{m}{n} \rfloor$ die Anzahl der ganzen Zahlen im Bereich $1, \dots, m$, die durch n teilbar sind.

- (i) ♣ Falls $m \in \mathbb{Z}_{>0}, n \in \mathbb{Z}$, dann $\lfloor \frac{n-1}{m} \rfloor = - \lfloor \frac{-n}{m} \rfloor - 1$.
 (j) ♣ Falls $m \in \mathbb{Z}_{>0}, n \in \mathbb{Z}$, dann ist $\lfloor \frac{n-1}{m} \rfloor + 1$ die kleinste ganze Zahl größer als oder gleich n/m .

1.5. Schreiben Sie so viele der obigen Gleichungen in Erzeugendenfunktionsgleichungen um, wie Ihnen sinnvoll erscheinen.

1.6. Es seien m und n teilerfremde positive ganze Zahle. Zeigen Sie, dass

$$\sum_{k=0}^{m-1} \left\lfloor \frac{kn}{m} \right\rfloor = \sum_{j=0}^{n-1} \left\lfloor \frac{jm}{n} \right\rfloor = \frac{1}{2}(m-1)(n-1).$$

1.7. Zeigen Sie die folgenden Gleichungen. Wir werden sie mindestens zweimal gebrauchen: Wenn wir Partialbrüche untersuchen, und wenn wir endliche Fourierreihen behandeln. Für $\phi, \psi \in \mathbb{R}, n \in \mathbb{Z}_{>0}, m \in \mathbb{Z}$ gilt:

- (a) $e^{i0} = 1$,
 (b) $e^{i\phi} e^{i\psi} = e^{i(\phi+\psi)}$,
 (c) $1/e^{i\phi} = e^{-i\phi}$,
 (d) $e^{i(\phi+2\pi)} = e^{i\phi}$,
 (e) $e^{2\pi i} = 1$,
 (f) $|e^{i\phi}| = 1$,
 (g) $\frac{d}{d\phi} e^{i\phi} = i e^{i\phi}$,
 (h) $\sum_{k=0}^{n-1} e^{2\pi i k m/n} = \begin{cases} n & \text{falls } n|m, \\ 0 & \text{sonst,} \end{cases}$
 (i) $\sum_{k=1}^{n-1} k e^{2\pi i k/n} = \frac{n}{e^{2\pi i/n} - 1}$.

1.8. Es seien $m, n \in \mathbb{Z}$ und $n > 0$. Finden Sie eine geschlossene Formel für $\sum_{k=0}^{n-1} \left\{ \frac{k}{n} \right\} e^{2\pi i k m/n}$ (als Funktion von m und n).

1.9. ♣ Es seien m und n teilerfremde ganze Zahlen, und n positiv. Zeigen Sie, dass

$$\left\{ e^{2\pi i k m/n} : 0 \leq k < n \right\} = \left\{ e^{2\pi i j/n} : 0 \leq j < n \right\}$$

und

$$\left\{ e^{2\pi i k m/n} : 0 < k < n \right\} = \left\{ e^{2\pi i j/n} : 0 < j < n \right\}.$$

Folgern Sie daraus, dass für eine beliebige komplexwertige Funktion f gilt

$$\sum_{k=0}^{n-1} f\left(e^{2\pi i k m/n}\right) = \sum_{j=0}^{n-1} f\left(e^{2\pi i j/n}\right)$$

und

$$\sum_{k=1}^{n-1} f\left(e^{2\pi i k m/n}\right) = \sum_{j=1}^{n-1} f\left(e^{2\pi i j/n}\right).$$

1.10. Es sei n eine positive ganze Zahl. Falls Sie wissen, was eine *Gruppe* ist, zeigen Sie, dass die Menge $\{e^{2\pi ik/n} : 0 \leq k < n\}$ eine zyklische Gruppe der Ordnung n bildet (unter Multiplikation in \mathbb{C}).

1.11. Wir halten ein $n \in \mathbb{Z}_{>0}$ fest. Für eine ganze Zahl m sei $(m \bmod n)$ die kleinste nichtnegative ganze Zahl in $G_1 := \mathbb{Z}_n$, zu der m kongruent ist. Wir bezeichnen mit \star die Addition modulo n , und mit \circ folgende Verknüpfung:

$$\left\{ \frac{m_1}{n} \right\} \circ \left\{ \frac{m_2}{n} \right\} = \left\{ \frac{m_1 + m_2}{n} \right\},$$

die auf der Menge $G_2 := \left\{ \left\{ \frac{m}{n} \right\} : m \in \mathbb{Z} \right\}$ definiert ist. Wir definieren die folgenden Funktionen:

$$\begin{aligned} \phi((m \bmod n)) &= e^{2\pi im/n}, \\ \psi\left(e^{2\pi im/n}\right) &= \left\{ \frac{m}{n} \right\}, \\ \chi\left(\left\{ \frac{m}{n} \right\}\right) &= (m \bmod n). \end{aligned}$$

Zeigen Sie das Folgende:

$$\begin{aligned} \phi((m_1 \bmod n) \star (m_2 \bmod n)) &= \phi((m_1 \bmod n)) \phi((m_2 \bmod n)), \\ \psi\left(e^{2\pi im_1/n} e^{2\pi im_2/n}\right) &= \psi\left(e^{2\pi im_1/n}\right) \circ \psi\left(e^{2\pi im_2/n}\right), \\ \chi\left(\left\{ \frac{m_1}{n} \right\} \circ \left\{ \frac{m_2}{n} \right\}\right) &= \chi\left(\left\{ \frac{m_1}{n} \right\}\right) \star \chi\left(\left\{ \frac{m_2}{n} \right\}\right). \end{aligned}$$

Zeigen Sie, dass die drei oben definierten Abbildungen, nämlich ϕ , ψ und χ , injektiv sind. Wieder für Leser, die mit dem Begriff einer *Gruppe* vertraut sind, sei G_3 die Gruppe der n -ten Einheitswurzeln. Was wir gezeigt haben ist, dass die drei Gruppen G_1 , G_2 und G_3 isomorph zueinander sind. Es ist sehr hilfreich, zwischen diesen drei isomorphen Gruppen herumzuwechseln.

1.12. ♣ Zu gegebenen ganzen Zahlen a , b , c und d bilden Sie den Geradenabschnitt in \mathbb{R}^2 , der den Punkt (a, b) mit (c, d) verbindet. Zeigen Sie, dass die Anzahl der Gitterpunkte auf diesem Geradenabschnitt $\text{ggT}(a - c, b - d) + 1$ ist.

1.13. Geben Sie ein Beispiel für eine Gerade mit

- (a) keinem Gitterpunkt;
- (b) einem Gitterpunkt;
- (c) unendlich vielen Gitterpunkten.

Geben Sie in jedem der Fälle, sofern es angemessen ist, notwendige Bedingungen an die (Ir)Rationalität der Steigung an.

1.14. Angenommen, eine Gerade $y = mx + b$ geht durch die Gitterpunkte (p_1, q_1) und (p_2, q_2) . Zeigen Sie, dass sie auch durch die Gitterpunkte

$$(p_1 + k(p_2 - p_1), q_1 + k(q_2 - q_1)), \quad k \in \mathbb{Z}$$

geht.

1.15. Zeigen Sie für gegebene irrationale Zahlen p und q mit $\frac{1}{p} + \frac{1}{q} = 1$, dass $\mathbb{Z}_{>0}$ die disjunkte Vereinigung der beiden Ganzzahlfolgen $\{\lfloor pn \rfloor : n \in \mathbb{Z}_{>0}\}$ und $\{\lfloor qn \rfloor : n \in \mathbb{Z}_{>0}\}$ ist. Dieser Satz von 1894 geht auf Lord Rayleigh zurück und wurde im Jahr 1926 von Sam Beatty erneut entdeckt. Folgen der Form $\{\lfloor pn \rfloor : n \in \mathbb{Z}_{>0}\}$ werden oft *Beatty-Folgen* genannt.

1.16. Seien $a, b, c, d \in \mathbb{Z}$. Wir nennen $\{(a, b), (c, d)\}$ eine *Gitterbasis* von \mathbb{Z}^2 , falls jeder Gitterpunkt $(m, n) \in \mathbb{Z}^2$ als

$$(m, n) = p(a, b) + q(c, d)$$

für bestimmte $p, q \in \mathbb{Z}$ geschrieben werden kann. Zeigen Sie, dass, wenn $\{(a, b), (c, d)\}$ und $\{(e, f), (g, h)\}$ Gitterbasen von \mathbb{Z}^2 sind, es eine Matrix M mit ganzzahligen Einträgen und Determinante ± 1 gibt, so dass

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = M \begin{pmatrix} e & f \\ g & h \end{pmatrix}.$$

Folgern Sie daraus, dass die Determinante von $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ gleich ± 1 ist.

1.17. ♣ Zeigen Sie, dass ein Dreieck mit Eckpunkten auf dem Gitter der ganzen Zahlen genau dann keine weiteren Gitterpunkte im Inneren bzw. auf dem Rand hat, wenn seine Fläche $\frac{1}{2}$ misst. (*Hinweis:* Beginnen Sie damit, das Dreieck zu einem Parallelogramm zu „verdoppeln“.)

1.18. Wir definieren einen *Nordost-Gitterpfad* als einen Pfad durch Gitterpunkte, der nur die Schritte $(1, 0)$ und $(0, 1)$ benutzt. Sei L_n die Gerade, die durch $x + 2y = n$ definiert ist. Zeigen Sie, dass die Anzahl der Nordost-Gitterpfade vom Ursprung zu einem Gitterpunkte auf L_n gerade die $(n + 1)$ -te Fibonacci-Zahl f_{n+1} ist.

1.19. Berechnen Sie die Koeffizienten der Taylorreihe von $1/(1 - z)^2$ um $z = 0$...

- (a) ... durch ein Abzählargument,
- (b) ... durch Ableiten der geometrischen Reihe.

Verallgemeinern Sie.

1.20. ♣ Zeigen Sie, dass für $a_1, a_2, \dots, a_d \in \mathbb{Z}_{>0}$, die keinen gemeinsamen Teiler haben, die Frobenius-Zahl $g(a_1, \dots, a_d)$ wohldefiniert ist.

1.21. ♣ Berechnen Sie die Partialbruchkoeffizienten (1.6).

1.22. ♣ Zeigen Sie (1.9): Für teilerfremde positive ganze Zahlen a und b gilt

$$\frac{1}{a} \sum_{k=1}^{a-1} \frac{1}{(1 - \xi_a^{bk}) \xi_a^{kn}} = \frac{1}{a} \sum_{k=1}^{a-1} \frac{1}{(1 - \xi_a^k) \xi_a^{b^{-1}kn}},$$

wobei $b^{-1}b \equiv 1 \pmod{a}$, und folgern Sie daraus (1.10), nämlich

$$\frac{1}{a} \sum_{k=1}^{a-1} \frac{1}{(1 - \xi_a^{bk}) \xi_a^{kn}} = - \left\{ \frac{b^{-1}n}{a} \right\} + \frac{1}{2} - \frac{1}{2a}.$$

(*Hinweis:* Benutzen Sie Aufgabe 1.9.)

1.23. Beweisen Sie, dass für teilerfremde positive ganze Zahlen a und b die Aussage

$$p_{\{a,b\}}(n + ab) = p_{\{a,b\}}(n) + 1$$

gilt.

1.24. ♣ Zeigen Sie, dass, wenn a und b teilerfremde positive ganze Zahlen sind,

$$p_{\{a,b\}}(a + b) = 1$$

gilt.

1.25. Um das Frobenius-Problem zu erweitern, bezeichnen wir eine ganze Zahl n als k -darstellbar, falls $p_A(n) = k$; d.h. falls n auf genau k Arten unter Benutzung der Zahlen aus A dargestellt werden kann. Wir definieren $g_k = g_k(a_1, \dots, a_d)$ als die größte k -darstellbare ganze Zahl. Zeigen Sie:

- Sei $d = 2$. Zu jedem $k \in \mathbb{Z}_{\geq 0}$ gibt es ein N , so dass alle ganzen Zahlen größer als N mindestens k Darstellungen haben (und daher $g_k(a, b)$ wohldefiniert ist).
- $g_k(a, b) = (k + 1)ab - a - b$.
- Zu gegebenem $k \geq 2$ ist die kleinste k -darstellbare ganze Zahl $ab(k - 1)$.
- Das kleinste Intervall, das alle eindeutig darstellbaren ganzen Zahlen enthält, ist $[\min(a, b), g_1(a, b)]$.
- Zu gegebenem $k \geq 2$ ist das kleinste Intervall, das alle k -darstellbaren ganzen Zahlen enthält, das Intervall $[g_{k-2}(a, b) + a + b, g_k(a, b)]$.
- Es gibt genau $ab - 1$ ganze Zahlen, die eindeutig darstellbar sind. Zu gegebenem $k \geq 2$ gibt es genau ab k -darstellbare ganze Zahlen.
- Erweitern Sie all dies auf $d \geq 3$ (see auch die offenen Probleme).

1.26. Finden Sie eine Formel für $p_{\{a\}}(n)$.

1.27. Beweisen Sie die folgende Rekursionsformel:

$$p_{\{a_1, \dots, a_d\}}(n) = \sum_{m \geq 0} p_{\{a_1, \dots, a_{d-1}\}}(n - ma_d)$$

(Dabei benutzen wir die Vereinbarung, dass $p_A(n) = 0$ falls $n < 0$). Benutzen Sie dies im Fall $d = 2$, um einen alternativen Beweis von Satz 1.2 zu geben.

1.28. Zeigen Sie die folgende Erweiterung von Satz 1.5: Sei $\text{ggT}(a, b) = d$. Dann gilt

$$p_{\{a, b\}}(n) = \begin{cases} \frac{nd}{ab} - \left\{ \frac{\beta n}{a} \right\} - \left\{ \frac{\alpha n}{b} \right\} + 1 & \text{falls } d|n, \\ 0 & \text{sonst,} \end{cases}$$

wobei $\beta \frac{b}{d} \equiv 1 \pmod{\frac{a}{d}}$, und $\alpha \frac{a}{d} \equiv 1 \pmod{\frac{b}{d}}$.

1.29. ♣ Berechnen Sie den Partialbruchkoeffizienten (1.12).

1.30. Finden Sie eine Formel für $p_{\{a, b, c\}}(n)$ für den Fall $\text{ggT}(a, b, c) \neq 1$.

1.31. ♣ Für $A = \{a_1, a_2, \dots, a_d\} \subset \mathbb{Z}_{>0}$ sei

$$p_A^\circ(n) := \# \{ (m_1, \dots, m_d) \in \mathbb{Z}^d : \text{alle } m_j > 0, m_1 a_1 + \dots + m_d a_d = n \};$$

d.h. $p_A^\circ(n)$ zählt die Anzahl der Partitionen von n mit Elementen von A als Teilen, wobei jedes Element mindestens einmal verwendet wird. Finden Sie Formeln für p_A° für $A = \{a\}$, $A = \{a, b\}$, $A = \{a, b, c\}$ und $A = \{a, b, c, d\}$, wobei a, b, c und d paarweise teilerfremde positive ganze Zahlen sind. Beachten Sie, dass in allen Beispielen die Zählfunktionen p_A und p_A° die algebraische Gleichung

$$p_A^\circ(-n) = (-1)^{d-1} p_A(n)$$

erfüllen.

1.32. Zeigen Sie, dass $p_A^\circ(n) = p_A(n - a_1 - a_2 - \dots - a_d)$ (hier sei, wie üblich, $A = \{a_1, a_2, \dots, a_d\}$). Folgern Sie, dass in den Beispielen von Aufgabe 1.31 die algebraische Gleichung

$$p_A(-t) = (-1)^{d-1} p_A(t - a_1 - a_2 - \dots - a_d)$$

erfüllt ist.

1.33. Für teilerfremde positive ganze Zahlen a und b sei

$$R := \{am + bn : m, n \in \mathbb{Z}_{\geq 0}\}$$

die Menge aller durch a und b darstellbaren ganzen Zahlen. Zeigen Sie, dass

$$\sum_{k \in R} z^k = \frac{1 - z^{ab}}{(1 - z^a)(1 - z^b)}.$$

Benutzen Sie diese Erzeugendenfunktion, um alternative Beweise der Sätze 1.2 und 1.3 zu geben.

1.34. Für teilerfremde positive ganze Zahlen a_1, a_2, \dots, a_d sei

$$R := \{m_1 a_1 + m_2 a_2 + \dots + m_d a_d : m_1, m_2, \dots, m_d \in \mathbb{Z}_{\geq 0}\}$$

die Menge aller durch a_1, a_2, \dots, a_d darstellbaren ganzen Zahlen. Zeigen Sie, dass

$$r(z) := \sum_{k \in R} z^k = \frac{p(z)}{(1 - z^{a_1})(1 - z^{a_2}) \dots (1 - z^{a_d})}$$

für ein Polynom p .

1.35. Beweisen Sie Satz 1.1: Zu jeder rationalen Funktion $\frac{p(z)}{\prod_{k=1}^m (z - a_k)^{e_k}}$, wobei p ein Polynom von kleinerem Grad als $e_1 + e_2 + \dots + e_m$ ist und die a_k s verschieden sind, gibt es eine Zerlegung

$$\sum_{k=1}^m \left(\frac{c_{k,1}}{z - a_k} + \frac{c_{k,2}}{(z - a_k)^2} + \dots + \frac{c_{k,e_k}}{(z - a_k)^{e_k}} \right),$$

wobei die $c_{k,j} \in \mathbb{C}$ eindeutig bestimmt sind.

Wir skizzieren einen möglichen Beweis: Zunächst erinnern wir uns daran, dass die Menge aller Polynome (über \mathbb{R} oder \mathbb{C}) einen *euklidischen Ring* bildet, d.h. zu beliebigen Polynomen $a(z)$ und $b(z) \neq 0$ gibt es Polynome $q(z)$ und $r(z)$ mit $\deg(r) < \deg(b)$, so dass

$$a(z) = b(z)q(z) + r(z).$$

Wenn wir diese Prozedur wiederholt anwenden (also den *euklidischen Algorithmus* durchführen), erhalten wir den größten gemeinsamen Teiler von $a(z)$ und $b(z)$ als Linearkombination der beiden. Das bedeutet, dass es Polynome $c(z)$ und $d(z)$ gibt, für die $a(z)c(z) + b(z)d(z) = \text{ggT}(a(z), b(z))$.

Schritt 1: Wenden Sie den euklidischen Algorithmus an, um zu zeigen, dass es Polynome u_1 und u_2 gibt, für die

$$u_1(z)(z - a_1)^{e_1} + u_2(z)(z - a_2)^{e_2} = 1.$$

Schritt 2: Folgern Sie daraus, dass es Polynome v_1 und v_2 mit $\deg(v_k) < e_k$ gibt, so dass

$$\frac{p(z)}{(z - a_1)^{e_1} (z - a_2)^{e_2}} = \frac{v_1(z)}{(z - a_1)^{e_1}} + \frac{v_2(z)}{(z - a_2)^{e_2}}.$$

(*Hinweis:* schriftliche Division.)

Schritt 3: Wiederholen Sie diese Prozedur, um zu einer Partialbruchzerlegung für

$$\frac{p(z)}{(z - a_1)^{e_1} (z - a_2)^{e_2} (z - a_3)^{e_3}}$$

zu gelangen.

Offene Probleme

1.36. Entwerfen Sie einen neuen Ansatz oder einen neuen Algorithmus für das Frobenius-Problem im Fall $d = 4$.

1.37. Es gibt eine sehr gute untere [64] und mehrere obere Schranken [152, Chapter 3] für die Frobenius-Zahl. Finden Sie eine bessere obere Schranke.

1.38. Lösen Sie Vladimir I. Arnolds Probleme 1999-8 bis 1999-11 [7]. Um einen Geschmack davon zu geben, erwähnen wir zwei davon explizit:

- (a) Untersuchen Sie das statistische Verhalten von $g(a_1, a_2, \dots, a_d)$ für typische große a_1, a_2, \dots, a_d . Es wird vermutet, dass $g(a_1, a_2, \dots, a_d)$ asymptotisch wie eine Konstante mal $\sqrt[d]{a_1 a_2 \cdots a_d}$ wächst.
- (b) Bestimmen Sie für typische große a_1, a_2, \dots, a_d , welcher Bruchteil der ganzen Zahlen im Intervall $[0, g(a_1, a_2, \dots, a_d)]$ darstellbar ist. Es wird vermutet, dass dieser Bruchteil asymptotisch gleich $\frac{1}{d}$ ist. (Satz 1.3 impliziert, dass diese Vermutung im Fall $d = 2$ wahr ist.)

1.39. Untersuchen Sie vektorielle Verallgemeinerungen des Frobenius-Problems [154, 163].

1.40. Es gibt einige Spezialfälle für $A = \{a_1, a_2, \dots, a_d\}$, in denen das Frobenius-Problem gelöst ist, z.B. arithmetische Folgen [152, Kapitel 3]. Betrachten Sie diese Spezialfälle im Lichte der Erzeugendenfunktion $r(x)$, die in den Anmerkungen und in Aufgabe 1.34 definiert wurde.

1.41. Untersuchen Sie die verallgemeinerte Frobenius-Zahl g_k (definiert in Aufgabe 1.25), z.B. im Zusammenhang mit dem in den Anmerkungen erwähnten Satz von Morales-Dunham. Leiten Sie Formeln für Spezialfälle, z.B. arithmetische Folgen, her.

1.42. Für welche $0 \leq n \leq b - 1$ ist $s_n(a_1, a_2, \dots, a_d; b) = 0$?